Devising superconvergent HDG methods for partial differential equations

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Ke Shi

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
Doctor of Philosophy

Thesis Adviser: Professor Bernardo Cockburn

August, 2012

Acknowledgments

First and foremost, I want to thank my advisor, Professor Bernardo Cockburn. I appreciate all his time, inspiration and his generous support during the past four years in Minnesota. The enthusiasm he has for both research and life has been a lighthouse to lead my way to success.

My sincere thanks also go to Dr. Fatih Celiker and Dr. Weifeng Qiu for helpful discussions. Our joint work [1, 2, 3, 4] is included in this thesis.

Last but not the least, I would like to thank my wife Luqiong Li, our son Jason, our parents and friends for their graceful support throughout my life.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 The HDG methods for diffusion equations

To better describe hybridizable discontinuous Galerkin (HDG) methods, we consider the following model problem:

$$\boldsymbol{q} + \nabla u = 0 \qquad \text{in } \Omega, \tag{1.1a}$$

$$\nabla \cdot \boldsymbol{q} = f \qquad \text{in } \Omega \tag{1.1b}$$

$$u = g \qquad \text{on } \partial\Omega. \tag{1.1c}$$

Here $\Omega \subset \mathbb{R}^n$ $(n = 2, 3)$ is a bounded polyhedral domain, $f \in L^2(\Omega)$ and $g \in H^{\frac{1}{2}}(\partial\Omega)$.

Let $\mathcal{T}_h := \{K\}$ denote a triangulation of Ω, where K is a polyhedral element. We denote the set of faces F of an element $K \in \mathcal{T}_h$ by $\mathcal{F}(K)$, and the set of faces F of all elements $K \in \mathcal{T}_h$ by $\mathcal{E}_h$. In general, DG methods are seeking the approximate solution $(\boldsymbol{q}_h, u_h)$ in some finite element space $\boldsymbol{V}_h \times W_h \subset \boldsymbol{L}^2(\Omega) \times L^2(\Omega)$, and determine it as the only solution of the following weak formulation:

$$-(u_h\,,\,\nabla \cdot \boldsymbol{v})_{\mathcal{T}_h} + (\boldsymbol{q}_h\,,\,\boldsymbol{v})_{\mathcal{T}_h} + \langle \widehat{u}_h\,,\,\boldsymbol{v} \cdot \boldsymbol{n}\rangle_{\partial\mathcal{T}_h} = 0, \tag{1.2a}$$

$$-(\boldsymbol{q}_h\,,\,\nabla w)_{\mathcal{T}_h} + \langle \widehat{\boldsymbol{q}}_h \cdot \boldsymbol{n}\,,\,w\rangle_{\partial\mathcal{T}_h} = (f\,,\,w)_{\mathcal{T}_h}, \tag{1.2b}$$

for all $(w, \boldsymbol{v}) \in W_h \times \boldsymbol{V}_h$. Here we write $(\eta\,,\,\zeta)_{\mathcal{T}_h} := \sum_{K\in\mathcal{T}_h}(\eta, \zeta)_K$, where $(\eta, \zeta)_D$ denotes the integral of $\eta\zeta$ over the domain $D \subset \mathbb{R}^n$. We also write $\langle \eta\,,\,\zeta\rangle_{\partial\mathcal{T}_h} := \sum_{K\in\mathcal{T}_h}\langle \eta\,,\,\zeta\rangle_{\partial K}$, where $\langle \eta\,,\,\zeta\rangle_D$ denotes the integral of $\eta\zeta$ over the domain $D \subset \mathbb{R}^{n-1}$

and $\partial\mathcal{T}_h := \{\partial K : K \subset \mathcal{T}_h\}$. The unknowns $\widehat{u}_h, \widehat{\boldsymbol{q}}_h$ are called the numerical traces. To complete the specification of a DG method we must express the numerical traces $\widehat{u}_h$, $\widehat{\boldsymbol{q}}_h$ in terms of u_h and $\boldsymbol{q}_h$. In order to provide good approximate solutions, we need to carefully define the numerical traces as well as the finite element spaces. We refer readers to [5] for more details.

The DG methods are ideally suited for numerically solving hyperbolic problems. However this is not the case for diffusion problems,even though they are ideally suited for *hp*-adaptivity. Indeed, when compared with the classical continuous Galerkin methods on the same mesh, they have many more global degrees of freedom and they are not easy to implement. When compared with the mixed methods, they do not provide optimally convergent approximations to the flux and do not display superconvergence properties of the scalar variable. As a response to these disadvantages, the HDG methods were introduced in [6]. Therein, it was shown that HDG methods can be implemented as efficiently as the mixed methods. Later in [7] it was proven that the HDG methods do share with mixed methods their superior convergence properties while retaining the advantages typical of the DG methods.

The main feature of the HDG methods is that it treats one or more numerical traces as *new* unknowns. At the same time, we complete the system by some additional *conservativity conditions*. In the case of the model problem (1.1), we consider $\widehat{u}_h$ as a new unknown and we seek the numerical solution $(u_h, \boldsymbol{q}_h, \widehat{u}_h)$, in the finite element space $W_h \times \boldsymbol{V}_h \times M_h$, where

$$\begin{aligned}
\boldsymbol{V}_h &:= \{\boldsymbol{v} \in \boldsymbol{L}^2(\mathcal{T}_h) : \boldsymbol{v}|_K \in \boldsymbol{V}(K),\ K \in \mathcal{T}_h\},\\
W_h &:= \{w \in L^2(\mathcal{T}_h) : w|_K \in W(K), K \in \mathcal{T}_h\},\\
M_h &:= \{\mu \in L^2(\mathcal{E}_h) : \mu|_F \in W(F),\ \ F \in \mathcal{E}_h\},
\end{aligned}$$

and determine it as the only solution of the following weak formulation:

$$\begin{aligned}
-(u_h\,,\,\nabla\cdot\boldsymbol{v})_{\mathcal{T}_h} + (\boldsymbol{q}_h\,,\,\boldsymbol{v})_{\mathcal{T}_h} + \langle \widehat{u}_h\,,\,\boldsymbol{v}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} &= 0, && \text{(1.3a)}\\
-(\boldsymbol{q}_h\,,\,\nabla w)_{\mathcal{T}_h} \qquad\qquad + \langle \widehat{\boldsymbol{q}}_h\cdot\boldsymbol{n}\,,\,w\rangle_{\partial\mathcal{T}_h} &= (f\,,\,w)_{\mathcal{T}_h}, && \text{(1.3b)}\\
\langle \widehat{\boldsymbol{q}}_h\cdot\boldsymbol{n}, \mu\rangle_{\partial\mathcal{T}_h\setminus\partial\Omega} &= 0, && \text{(1.3c)}\\
\langle \widehat{u}_h, \mu\rangle_{\partial\Omega} &= \langle g, \mu\rangle_{\partial\Omega}, && \text{(1.3d)}
\end{aligned}$$

for all $(w, \boldsymbol{v}, \mu) \in W_h \times \boldsymbol{V}_h \times M_h$. The definition of the method is completed with the definition of the normal component of the numerical trace:

$$\widehat{\boldsymbol{q}}_h \cdot \boldsymbol{n} = \boldsymbol{q}_h \cdot \boldsymbol{n} + \alpha(u_h - \widehat{u}_h) \quad \text{on} \quad \partial\mathcal{T}_h. \tag{1.4}$$

By taking particular choices of the local spaces $\boldsymbol{V}(K)$, $W(K)$ and $M(F)$, and the *linear local stabilization* operator α, the different HDG methods are obtained. Here (1.3c) is called the *conservativity condition* which forces the normal component of the numerical trace $\widehat{\boldsymbol{q}}_h$ continuous on $\mathcal{E}_h$. It seems like we are facing a much bigger system to solve. In fact, in [6] it was shown that, in the above framework, the only global unknown is the numerical trace $\widehat{u}_h$. We can solve all other unknowns element by element. A simple calculation shows that the size of the global system is significant reduced if we employ high order polynomial spaces. In addition, in [7], they analyzed a specific family of such HDG methods. Namely, they consider the triangulation with simplexes and the local spaces are defined as

$$W(K) = P^k(K), \quad \boldsymbol{V}(K) = \boldsymbol{P}^k(K), \quad M(F) = P^k(F),$$

Here $P^k(D)$ denotes the space of polynomials with degree no more than k on D and $\boldsymbol{P}^k(D) = [P^k(D)]^n$. They define the local stablization operator α by

$$\alpha = \tau = \begin{cases} 0, & \text{on } \partial K \backslash e_K^\tau, \\ \tau_K, & \text{on } e_K^\tau, \end{cases}$$

where e_K^τ is an arbitrary face of K and τ_K is a strictly positive real number. They proved that if we choose $\tau_K = \mathcal{O}(\frac{1}{h})$, the numerical solution $\boldsymbol{q}_h, u_h$ converges to $\boldsymbol{q}, u$ with order $k+1$. Moreover, by applying a simple postprocessing technique, we can obtain a better approximation $u_h^\star$ with order $k+2$ ($k \geq 1$). This shows that for the model problem (1.1), the accuracy of HDG methods is as good as those well known mixed methods [8].

Inspired by these results, in this Thesis we are trying to explore HDG methods in a wider circumstance. We briefly discuss our main contributions in the following sections.

1.2 The HDG methods for Timoshenko Beams

Our long term goal for this project is to devise optimal HDG methods for plate and shell models. As a starting point, we began by testing our methods for Timoshenko beams.

The first discontinuous Galerkin (DG) methods for Timoshenko beams were introduced in [9]. They were numerically explored in [10] and then theoretically studied in [10]. Therein, these DG methods were proven to provide approximations for the displacement, rotation, bending moment and shear force simultaneously converging with the optimal order of $k+1$ when polynomials of degree $k \geq 0$ were used to define the method. For the corresponding numerical traces at the nodes, the order of convergence of $2k+1$ was also established for $k \geq 0$. Finally, it was proved that these properties hold uniformly with respect to the thickness of the beam.

In Chapter 2, we analyze the HDG methods proposed in [1] by using the approach proposed in [11] to study HDG methods for diffusion problems when all the variables are approximated by piecewise polynomials of degree $k \geq 0$. We show that results similar to those proven in [10] for the DG methods [9] hold for a wide class of HDG methods. The approach has three main steps. The first is to find a suitably defined projection such that the equations of the *projection of the approximation errors* becomes extremely simple. This is achieved by tailoring the definition of the projection to the very structure of the numerical traces of the HDG methods. The second step is to study the approximation properties of the projection in a single typical element. It is in this step that the information of the particular definition of the numerical traces is captured and allows us to determine for what choices we obtain the optimal order of convergence of $k+1$ for each of the four variables for $k \geq 0$. The third step consists in bounding the projection of the errors in terms of the *approximation properties of the projection* only. This step is rendered particularly concise because of the simple form of the equations of the projection of the errors.

The analysis we present here has two striking features. The first is that no positivity of the stabilization function is required; let us elaborate. In the analysis of numerical methods for symmetric elliptic problems, it is standard to take advantage of the fact that they have an underlying *energy*. The delicate devising of the numerical traces for the DG methods is then geared towards making sure that the *discrete energy* associated with the jumps of the approximation is actually positive. This is the approach taken in [10] to analyze the first DG methods for Timoshenko beams; it could have been easily taken to carry out the analysis of the methods under consideration. However, the alternative approach we present here has an unexpected and surprising feature. It is the fact that, unlike the analysis of any other DG method, we do *not* rely on the stabilization function having any positivity property.

The second striking feature is that the projection of the errors of each of *all* the components of the approximation *superconverge* with order $k+2$ when $k \geq 1$ and with order 1 for $k=0$.

These above two properties are due to the fact that we are only using duality arguments and that the problem is one dimensional. The price to pay, however, is that the maximum meshsize of the partition, h, needs to be *small enough*.

The Chapter 2 is organized as follows. In Section 2.1, we display the HDG methods, define the projection employed in the error analysis and state and briefly discuss our main results. Detailed proofs of these results are presented in Sections 2.2 (approximation properties of the projection), 2.3 (estimates of the projection of the errors) and 2.4 (estimates of the errors of the numerical traces). Numerical results verifying the theoretical orders of convergence are presented in Section 2.5. We end in Section 2.6 with some concluding remarks.

1.3 Superconvergent HDG methods for second order elliptic equations

In Chapter 3, we propose a projection-based a priori error analysis of finite element methods for second-order elliptic problems. The analysis is *unifying* because it applies to a large class of methods including the hybridized version of most well-known mixed methods as well as several hybridizable discontinuous Galerkin (HDG) methods. The novelty of the approach is that it reduces the whole error analysis to the element-by-element construction of an auxiliary projection satisfying certain orthogonality and approximation properties, and to the verification of very simple inclusion properties of the local spaces defining the methods.

Two ideas led to this approach. The first is that many mixed methods, including the method of Raviart-Thomas (RT), [12, 13, 14], Brezzi-Douglas-Marini, [15], and Brezzi-Douglas-Fortin-Marini, [16], were successfully analyzed by using suitably defined auxiliary projections; see also [8]. The second is that both mixed and HDG methods can be seen as particular cases of a single, general numerical method uncovered in [6]. This suggested the possibility of using a similar projection-based approach to analyze HDG methods. Recently, this was actually achieved, first for a particular case of HDG methods whose local solvers are defined by the local discontinuous Galerkin method (LDG-H) (defined on simplexes) in [7], and then for the whole family of those methods in [11]. In this paper, we continue this effort and show that a single error analysis of many of the methods fitting in the general framework proposed in [6] can be realized.

The Chapter 3 is organized as follows. In Section 1 we introduce the general HDG methods for diffusion problems. In Section 2, we describe the conditions on the auxiliary projection Π_h and the local spaces associated with our finite element methods and present our a priori error estimates. In Section 3, we describe a template for the devising of superconvergent HDG methods. In Section 4, we use it to we give various particular examples of hybridized mixed and HDG methods with superconvergent properties. In Section 5, we provide a detailed proof of the a priori error estimates. We then end with some extensions and concluding remarks in Section 6.

1.4 Superconvergent HDG methods on isoparametric elements for second order elliptic equations

The analysis proposed in Chapter 3 and [3] applies to a large class of methods including the hybridized version of most well known mixed methods and the HDG methods. The elements K these methods use are related to the reference element $\underline{K}$ by simple affine mappings. Thus, the analysis can be applied when the elements are simplexes, rectangles and cubes, or prisms. In this paper, we extend this approach to the case in which the elements K are related to the reference element $\underline{K}$ by *nonlinear* mappings. In particular, we find new superconvergent mixed and HDG methods defined on general quadrilateral, hexahedral or isoparametric elements.

We show that for all these methods, we can construct an operator mapping the exact solution $(\boldsymbol{q}, u)$ into an approximation $(\boldsymbol{\Pi}_V \boldsymbol{q}, \Pi_W u)$ in the finite element space such that

$$\|\boldsymbol{q} - \boldsymbol{q}_h\|_{\mathcal{T}_h} \leq 2\, \|\boldsymbol{q} - \boldsymbol{\Pi}_V \boldsymbol{q}\|_{\mathcal{T}_h},$$
$$\|\Pi_W u - u_h\|_{\mathcal{T}_h} \leq C\, h \|\boldsymbol{q} - \boldsymbol{\Pi}_V \boldsymbol{q}\|_{\mathcal{T}_h},$$

where $\|\cdot\|_{\mathcal{T}_h}$ denotes the $L^2(\mathcal{T}_h)$-norm, provided certain conditions on the mesh, and on the reference element are satisfied. The conditions on the mesh are, roughly speaking, that the mappings from the reference element $\underline{K}$ to the adjacent elements K and K', G_K and $G_{K'}$, satisfy simple *continuity* conditions when restricted to $K \cap K'$. The conditions on the reference element are almost exactly those obtained in [3] for the case in which the mappings G_K are affine. The only additional condition is a condition on the space of *traces* which allow the corresponding global space to be a space of single-valued function on the borders of the elements. This condition appears because of the nonlinear nature of the mappings G_K.

Note that if the error $\Pi_W u - u_h$ converges to zero *faster* than the error $u - u_h$, this *superconvergence* property can be advantageously exploited; see [14, 15, 17, 18, 19]. Indeed, in [3] we showed that , for all the methods under consideration, it is possible to compute, in an element-by-element fashion, a new approximation $u_h^* \in W_h^*$ such that

$$\|u - u_h^*\|_{\mathcal{T}_h} \leq \|\Pi_W u - u_h\|_{\mathcal{T}_h} + C\, h\, (\|\boldsymbol{q} - \boldsymbol{q}_h\|_{\mathcal{T}_h} + \inf_{\omega \in W_h^*} \|\nabla(u - \omega)\|_{\mathcal{T}_h}),$$

which means that, if we choose correctly the postprocessing space W_h^*, it is possible to define u_h^* converging to u as fast as $\Pi_W u - u_h$ converges to zero. Here we show that this can be also done.

Let us briefly mention that the main feature of our approach is that, thanks to the introduction of two new, suitably defined projections, the error analysis for the nonlinear mappings G_K is almost *identical* to that of the case considered in Chapter 3 and [3] in which the mappings G_K are affine and the two above-mentioned projections coincide. Let us also mention that the approach we propose here can be applied to other partial differential equations. In particular, it can be extended to convection-diffusion equations [20, 21], the heat equation [22], to Stokes flow [23], to linear elasticity [24] and to the wave equation [25].

The Chapter 4 is organized as follows. In Section 1, we introduce the methods under consideration. In particular, we describe the type of meshes we are going to consider and the form of the finite element spaces in which we seek the approximations. In Section 2, we describe the conditions on the reference element and present our a priori error estimates. In Section 3, we provide detailed proofs of our error estimates. Therein, we only present the proofs of the properties of the projection. We omit the rest of the proofs since they are almost the same as in Chapter 3. We refer readers to [4] for more details.

1.5 Superconvergent HDG methods for Stokes equations

In Chapter 5, we propose a projection-based analysis of superconvergent HDG methods for the velocity gradient-velocity-pressure formulation of the Stokes equations, namely,

$$\mathrm{L} - \nabla \boldsymbol{u} = 0 \qquad \text{on } \Omega, \tag{1.5a}$$

$$-\nabla \cdot (\nu \mathrm{L}) + \nabla p = \boldsymbol{f} \qquad \text{on } \Omega, \tag{1.5b}$$

$$\nabla \cdot \boldsymbol{u} = 0 \qquad \text{on } \Omega, \tag{1.5c}$$

$$\boldsymbol{u} = \boldsymbol{g} \qquad \text{on } \partial\Omega, \tag{1.5d}$$

$$\int_\Omega p = 0, \tag{1.5e}$$

where $\int_{\partial\Omega} \boldsymbol{g} \cdot \boldsymbol{n} = 0$. Here $\Omega \subset \mathbb{R}^n$ $(n = 2, 3)$ is a bounded polygonal domain if $n = 2$ or a Lipschitz polyhedral domain if $n = 3$. We assume that ν is a constant.

We extend the methodology used in Chapter 3 and [3] to the Stokes equations of incompressible flow. Our main result is that if we can construct an auxiliary projection $\Pi_h(\mathrm{L}, \boldsymbol{u}, p) = (\Pi_G \mathrm{L}, \boldsymbol{\Pi}_V \boldsymbol{u}, \mathit{\Pi}_P p)$ satisfying certain orthogonality and approximation conditions, and if the local spaces $\mathrm{G}(K), \boldsymbol{V}(K), P(K)$ and $\boldsymbol{M}(F)$, for $F \in \mathcal{F}(K)$, satisfy some inclusion properties, then the method is well defined and numerical solution is optimally convergent to the exact solution. Moreover, thanks to the projection, we are able to construct a new approximate solution $\boldsymbol{u}*_h$ which is superconvergent to the exact velocity $\boldsymbol{u}$.

The a priori error analysis of the HDG method proposed in [26] is now a particular case of our general approach. However, we provide an *alternative* definition of the auxiliary projection to carry it out. The key property of this new projection is that the definition of $\mathit{\Pi}_P p$ is completely *decoupled* from the definition of $(\Pi_G \mathrm{L}, \boldsymbol{\Pi}_V \boldsymbol{u})$. This not only considerably simplifies the study of its approximation properties but allows us to see how to systematically construct the HDG methods under consideration from the HDG methods for diffusion considered in [3]. To the best of the authors' knowledge, the methods obtained by using this construction seem to be *new*, except, of course, for the HDG methods on simplexes in [27, 26]. It is worth noting that a particular case of the family of methods employing k-th order Raviart-Thomas elements for the approximate velocity gradient is already known. Indeed, the two-dimensional case with $k = 0$ is nothing but the mixed method proposed in [28] for triangles and in [29] for rectangles; general, convex quadrilaterals were also considered in [29].

The Chapter 5 is organized as follows. In Section 1, we will present the Stokes equation and the general HDG methods under consideration. In Section 2, we describe the conditions on the auxiliary projection Π_h and on the local spaces associated with our finite element methods; we then state and discuss the corresponding a priori error estimates. In Section 3, we show how to systematically construct HDG methods with superconvergent velocities from superconvergence HDG methods for diffusion. In Section 4, we provide detailed proofs for the main results. In Section 5, we give a simple proof for the approximation property of the projection mentioned above. We then end with some concluding remarks in Section 6 which include the sketch of the extension of our approach of the HDG methods for linear elasticity recently proposed in [30].

1.6 Superconvergent HDG methods for linear elasticity with weakly symmetric stresses

In Chapter 6, we introduce a systematic way of devising new superconvergent mixed and HDG methods for the system of linear elasticity

$$\mathcal{A}\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\epsilon}}(\boldsymbol{u}) = 0 \quad \text{in } \Omega \subset \mathbb{R}^n, \tag{1.6a}$$

$$\nabla \cdot \underline{\boldsymbol{\sigma}} = \boldsymbol{f} \quad \text{in } \Omega, \tag{1.6b}$$

$$\boldsymbol{u} = \boldsymbol{0} \quad \text{on } \partial\Omega, \tag{1.6c}$$

Here $\underline{\boldsymbol{\epsilon}}(\boldsymbol{u}) := \frac{1}{2}(\nabla \boldsymbol{u} + \nabla^t \boldsymbol{u})$ is the strain, $\mathcal{A}$ is a bounded symmetric positive definite tensor, $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega)$ and Ω is a polyhedral domain.

The main feature of our approach to devising of superconvergent mixed or HDG methods is that it is based on the construction of superconvergent mixed or HDG methods for diffusion obtained in [3]. Thus, for any superconvergent mixed or HDG method for diffusion, we can immediately obtain a superconvergent mixed or HDG method for linear elasticity based on a weak symmetry formulation. This result is an extension of similar work carried out for HDG methods for the Stokes system of incompressible fluid flow and for the equations of isotropic elasticity [23]. It incorporates a new technique for devising mixed methods for linear elasticity based on weak symmetry formulations introduced in [31, 32, 33].

The weak formulation of these methods is based on the following form of the system of liner elasticity:

$$\begin{aligned} \mathcal{A}\underline{\boldsymbol{\sigma}} - \nabla \boldsymbol{u} + \underline{\boldsymbol{\rho}} &= 0 \quad \text{in } \Omega, \\ \nabla \cdot \underline{\boldsymbol{\sigma}} &= \boldsymbol{f} \quad \text{in } \Omega, \\ \boldsymbol{u} &= 0 \quad \text{on } \partial\Omega, \end{aligned}$$

where $\underline{\boldsymbol{\rho}} = \frac{1}{2}(\nabla \boldsymbol{u} - \nabla^t \boldsymbol{u})$ is called the *rotation*. Many mixed methods had been proposed which use the above system. In 1984 [34] the mixed method called PEERS for two-space dimensions was introduced. The space of the pseudo-stress $\underline{\boldsymbol{\sigma}}$ used the so-called Raviart-Thomas space of lowest index with additional bubble functions on each row. Optimal convergence orders were obtained for all unknowns. In 1988 [17] several new methods for the linear elasticity problem in both two- and three-space dimensions were proposed. It was shown how to use the Brezzi-Douglas-Marini and Raviart-Thomas spaces, originally devised for solving diffusion problems, to obtain the pseudo-stress space. By adding suitably defined bubble functions to the pseudo-stress stress space, optimal convergence orders were obtained.

In 2007, the de Rham complex was used to construct mixed methods [35]. The space of polynomials of degree no more than $k+1$ was used for the pseudo-stress, and polynomials of degree no more than k for the displacement $\boldsymbol{u}$ and the rotation $\underline{\boldsymbol{\rho}}$. The order of convergence for all unknowns was proven to be $k+1$. In 2009, an alternative analysis for these methods using an $\inf - \sup$ argument was proposed in [36]. More recently, two new mixed methods were introduced [31, 32] which use new bubble functions thanks to which both methods were proven to have optimal orders of convergence for all unknowns. Soon after, a unified analysis for all the mixed methods mentioned above was proposed [33].

We incorporate the use of the above-mentioned bubble functions into our approach for devising superconvergent methods. However, we do not restrict ourselves to simplexes (the elements we use can be also be squares, cubes or prisms), nor we require the commutativity property

$$\nabla \cdot \underline{\mathbf{\Pi}_{\boldsymbol{V}}^{\boldsymbol{D}}} = \underline{\boldsymbol{P}} \nabla \cdot,$$

where $\underline{\mathbf{\Pi}_{\boldsymbol{V}}^{\boldsymbol{D}}}$ is a suitably defined projection into the space of pseudo-stresses and $\underline{\boldsymbol{P}}$ is the L^2-projection onto the space of displacements, typically required in the analysis of mixed methods. In our approach, we relax this property so that the analysis can be used for methods whose approximate pseudo-stress does not necessarily lie in $\underline{\boldsymbol{H}}(div, \Omega)$.

The Chapter 6 is organized as follows, In Section 1, we describe the finite element spaces and present our a priori error estimates. In Section 2, we provide a detailed proof of the a priori error estimates.

Chapter 2

Superconvergent HDG methods for Timoshenko Beams

2.1 Main Result

The dimensionless form of the Timoshenko beam model [37], is given by the differential equations, see [10],

$$w' = \theta - d^2 \frac{T}{GA}, \qquad \theta' = \frac{M}{EI}, \qquad M' = T, \qquad T' = q, \tag{2.1a}$$

in $\Omega := (0,1)$, and the boundary conditions

$$w = w_D, \qquad \theta = \theta_N \quad \text{on } \partial\Omega = \{0,1\}. \tag{2.1b}$$

Here, the unknowns are the transverse displacement w, the rotation of the transverse cross-section of the beam θ, the bending moment M, and the shear force T. The material and geometrical properties of the beam are characterized by the shear modulus G, the cross-section area A, the Young modulus E, and the moment of inertia I. The transverse load is denoted by q. It was shown in [10] that we can assume without loss of generality that EI and GA are very smooth functions. The parameter $0 < d < 1$ represents the thickness of the beam.

2.1.1 The HDG methods

Let us describe the HDG methods under consideration, We begin by introducing the notation. To each partition of the domain Ω,

$$\mathcal{T}_h := \{(x_{j-1}, x_j) : \ 0 = x_0 < x_1 < \cdots < x_{N-1} < x_N = 1\},$$

we associate the set of nodes, $\mathcal{E}_h := \{x_0, x_1, \dots, x_N\}$, and the set of interior nodes, $\mathcal{E}_h^0 := \mathcal{E}_h \backslash \partial\Omega$; we also set $\partial\mathcal{T}_h := \{\partial K : \ K \in \mathcal{T}_h\}$. For each element $K \in \mathcal{T}_h$, let h_K denote the length of K, and set $h := \max_{K \in \mathcal{T}_h}\{h_K\}$. Finally, the space of piecewise polynomials of degree k on Ω is defined as

$$V_h^k := \{v : \mathcal{T}_h \mapsto \mathbb{R} : \quad v|_K \in P^k(K) \text{ for all } K \in \mathcal{T}_h\}.$$

We also set

$$L_0^2(\mathcal{E}_h) := \{\mathsf{w} \in L^2(\mathcal{E}_h) : \mathsf{w} = 0 \text{ on } \partial\Omega\}.$$

The HDG method seeks an approximation $(T_h, M_h, \theta_h, w_h, \widehat{M}_h, \widehat{w}_h)$ to the exact solution $(T, M, \theta, w, M|_{\mathcal{E}_h}, w|_{\mathcal{E}_h})$, in the finite dimensional space $[V_h^k]^4 \times L^2(\mathcal{E}_h) \times L^2(\mathcal{E}_h)$. It is determined by requiring that

$$-(w_h, v_1')_{\mathcal{T}_h} + \langle \widehat{w}_h \,,\, v_1 \boldsymbol{n}\rangle_{\partial\mathcal{T}_h} = (\theta_h, v_1)_{\mathcal{T}_h} - d^2\,(T_h/GA, v_1)_{\mathcal{T}_h}, \tag{2.2a}$$

$$-(\theta_h, v_2')_{\mathcal{T}_h} + \langle \widehat{\theta}_h \,,\, v_2 \boldsymbol{n}\rangle_{\partial\mathcal{T}_h} = (M_h/EI, v_2)_{\mathcal{T}_h}, \tag{2.2b}$$

$$-(M_h, v_3')_{\mathcal{T}_h} + \langle \widehat{M}_h \,,\, v_3 \boldsymbol{n}\rangle_{\partial\mathcal{T}_h} = (T_h, v_3)_{\mathcal{T}_h}, \tag{2.2c}$$

$$-(T_h, v_4')_{\mathcal{T}_h} + \langle \widehat{T}_h \,,\, v_4 \boldsymbol{n}\rangle_{\partial\mathcal{T}_h} = (q, v_4)_{\mathcal{T}_h}, \tag{2.2d}$$

$$\langle \widehat{\theta}_h \,,\, \mathsf{m}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} = \langle \theta_N \,,\, \mathsf{m}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}, \tag{2.2e}$$

$$\langle \widehat{T}_h \,,\, \mathsf{w}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} = 0, \tag{2.2f}$$

hold for all $(v_1, v_2, v_3, v_4, \mathsf{m}, \mathsf{w}) \in [V_h^k]^4 \times L^2(\mathcal{E}_h) \times L_0^2(\mathcal{E}_h)$. Here, the outward unit normal vectors are $n(x^{\mp}) := \pm 1$ for $x \in \mathcal{E}_h$. The "volume" inner product is defined as

$$(u, v)_{\mathcal{T}_h} := \sum_{j=1}^{N} (u, v)_{I_j} \quad \text{where} \quad (u, v)_{I_j} := \int_{I_j} u(x)v(x)\,dx,$$

and the boundary inner product is defined as

$$\langle u \,,\, v\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} := \sum_{j=1}^{N} \langle u \,,\, v\boldsymbol{n}\rangle_{\partial I_j} \quad \text{where} \quad \langle u \,,\, v\boldsymbol{n}\rangle_{\partial I_j} := u(x_j^-)v(x_j^-) + u(x_{j-1}^+)v(x_{j-1}^+).$$

Here, we are using the following notation. For any $\varphi \in L^2(\Omega_h)$ we set $\varphi(x_l^{\mp}) := \lim_{\epsilon\downarrow 0}\varphi(x_l \mp \epsilon)$ for $x_l \in \mathcal{E}_h$. For any $\varphi \in L^2(\partial\Omega_h)$, the value of φ at x_l is denoted by $\varphi(x_l^-)$, respectively $\varphi(x_l^+)$, when x_l is a boundary of I_l, respectively, of I_{l+1}.

Note that the boundary condition (2.1b) on θ is imposed by equation (2.2e). The boundary condition (2.1b) on w is imposed as follows:

$$\widehat{w}_h = w_D \quad \text{on } \partial\Omega. \tag{2.3a}$$

To complete the definition of the HDG method, we need to express the numerical traces $\widehat{T}_h$ and $\widehat{\theta}_h$ in terms of the unknowns:

$$\begin{bmatrix}\widehat{\theta}_h\\ \widehat{T}_h\end{bmatrix} = \begin{bmatrix}\theta_h\\ T_h\end{bmatrix} - \mathrm{S}\begin{bmatrix}M_h - \widehat{M}_h\\ w_h - \widehat{w}_h\end{bmatrix}\boldsymbol{n}, \tag{2.3b}$$

where the so-called stabilization function S is a matrix-valued function defined on $\partial\mathcal{T}_h$. It has to be suitably defined to guarantee the existence and uniqueness of the approximate solution; for details see [1].

Let us note that the hallmark of these methods lies in the fact that the only globally coupled degrees of freedom are the values of $\widehat{M}_h$ and $\widehat{w}_h$ on $\mathcal{E}_h$. The remaining degrees of freedom can then be recovered in an element-by-element fashion; see [1].

2.1.2 The projection

Next, we introduce the main *tool* of our error analysis, namely, a new projection

$$\boldsymbol{\Pi}_h = (\Pi_t, \Pi_M, \Pi_\theta, \Pi_w) : [L^2(\Omega)]^4 \to [V_h^k]^4,$$

associated with the HDG methods. It is a generalization of the projection introduced in [11] for the error analysis of HDG methods for second order elliptic problems.

It is defined as follows.Given a function $(u_1, u_2, u_3, u_4) \in [L^2(\Omega)]^4$ and an arbitrary subinterval $K \in \mathcal{T}_h$, the restriction of $(\Pi_T u_1, \Pi_M u_2, \Pi_\theta u_3, \Pi_w u_4)$ to K is defined as the element of $[P^k(K)]^4$ that satisfies

$$(\Pi_T u_1, v_1)_K = (u_1, v_1)_K, \tag{2.4a}$$

$$(\Pi_M u_2, v_2)_K = (u_2, v_2)_K, \tag{2.4b}$$

$$(\Pi_\theta u_3, v_3)_K = (u_3, v_3)_K, \tag{2.4c}$$

$$(\Pi_W u_4, v_4)_K = (u_4, v_4)_K, \tag{2.4d}$$

for all $(v_1, v_2, v_3, v_4) \in [P^{k-1}(K)]^4$, and

$$\begin{bmatrix} u_3 \\ u_1 \end{bmatrix} = \begin{bmatrix} \Pi_\theta u_3 \\ \Pi_T u_1 \end{bmatrix} - \mathrm{S} \begin{bmatrix} \Pi_M u_2 - u_2 \\ \Pi_w u_4 - u_4 \end{bmatrix} \boldsymbol{n} \qquad \text{on } \partial K. \tag{2.4e}$$

Note that when $k = 0$, the projection is defined solely by (2.4e). Note also that the last set of equations reflects the form of the equations (2.3) defining the numerical traces $\widehat{\theta}_h$ and $\widehat{T}_h$. As we are going to see in the next subsection, this is what allows us to obtain a very simple set of equations for the projection of the errors.

Finally, let us point out that the projection is well defined under mild conditions on the stabilization function S. To see this, note that the total number of unknowns involved in the linear system that is needed to be solved for computing $(\Pi_T u_1, \Pi_M u_2, \Pi_\theta u_3, \Pi_w u_4)$ is $4(k+1)$ since each component of the projection has $k+1$ degrees of freedom. On the other hand, the total number of linearly independent equations provided by the definition of the projection is also $4(k+1)$. The existence and uniqueness of the projection then follows from the approximation properties of the projection; see below.

2.1.3 The equations for the projection of the errors

As we said in the Introduction, the projection should be devised in such a way that the equations of the projection of the errors be as simple as possible. Let us show that this is indeed the case.

So, since the exact solution (T, M, θ, w) of the governing equations (2.1) satisfies the formulation of the HDG approximation, (2.2), we immediately see that the equations for the errors

$$e_u := u - u_h \quad \text{and} \quad \widehat{e}_u := u - \widehat{u}_h \quad \text{for } u = T, M, \theta, w,$$

are

$$
\begin{aligned}
-(e_{\mathrm{w}}, v_1')_{\mathcal{T}_h} + \langle \widehat{e}_{\mathrm{w}}\,,\, v_1\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} &= (e_\theta, v_1)_{\mathcal{T}_h} - d^2(e_T/GA, v_1)_{\mathcal{T}_h},\\
-(e_\theta, v_2')_{\mathcal{T}_h} + \langle \widehat{e}_\theta\,,\, v_2\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} &= (e_M/EI, v_2)_{\mathcal{T}_h},\\
-(e_M, v_3')_{\mathcal{T}_h} + \langle \widehat{e}_M\,,\, v_3\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} &= (e_T, v_3)_{\mathcal{T}_h},\\
-(e_T, v_4')_{\mathcal{T}_h} + \langle \widehat{e}_T\,,\, v_4\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} &= 0,\\
\langle \widehat{e}_\theta\,,\, \mathsf{m}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} &= 0,\\
\langle \widehat{e}_T\,,\, \mathsf{w}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} &= 0,
\end{aligned}
$$

hold for all $(v_1, v_2, v_3, v_4, \mathsf{m}, \mathsf{w}) \in [V_h^k]^4 \times L^2(\mathcal{E}_h) \times L_0^2(\mathcal{E}_h)$. Hence, setting

$$\delta_u := u - \Pi_u u \quad \text{for} \quad u = T, M, \theta, \mathrm{w},$$

we obtain

$$
\begin{aligned}
-(\Pi_{\mathrm{w}} e_{\mathrm{w}}, v_1')_{\mathcal{T}_h} + \langle \widehat{e}_{\mathrm{w}}\,,\, v_1\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} &= (\Pi_\theta e_\theta, v_1)_{\mathcal{T}_h} - d^2(\Pi_T e_T/GA, v_1)_{\mathcal{T}_h} \\
&\quad + (\delta_\theta, v_1)_{\mathcal{T}_h} - d^2(\delta_T/GA, v_1)_{\mathcal{T}_h}, && \text{(2.5a)}\\
-(\Pi_\theta e_\theta, v_2')_{\mathcal{T}_h} + \langle \widehat{e}_\theta\,,\, v_2\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} &= (\Pi_M e_M/EI, v_2)_{\mathcal{T}_h} + (\delta_M/EI, v_2)_{\mathcal{T}_h}, && \text{(2.5b)}\\
-(\Pi_M e_M, v_3')_{\mathcal{T}_h} + \langle \widehat{e}_M\,,\, v_3\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} &= (\Pi_T e_T, v_3)_{\mathcal{T}_h} + (\delta_T, v_3)_{\mathcal{T}_h}, && \text{(2.5c)}\\
-(\Pi_T e_T, v_4')_{\mathcal{T}_h} + \langle \widehat{e}_T\,,\, v_4\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} &= 0, && \text{(2.5d)}\\
\langle \widehat{e}_\theta\,,\, \mathsf{m}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} &= 0, && \text{(2.5e)}\\
(\widehat{e}_T\,,\, \mathsf{w}\boldsymbol{n})_{\mathcal{T}_h} &= 0, && \text{(2.5f)}
\end{aligned}
$$

for all $(v_1, v_2, v_3, v_4, \mathsf{m}, \mathsf{w}) \in [V_h^k]^4 \times L^2(\mathcal{E}_h) \times L_0^2(\mathcal{E}_h)$. Note that we have used the orthogonality property of the projection (2.4) in each of the first terms of the first four equations.

To complete the error equations, we have to add the boundary conditions

$$\widehat{e}_{\mathrm{w}} = 0 \quad \text{on } \partial\Omega, \tag{2.6a}$$

as well as the equations relating the errors inside the elements to the errors of the numerical traces, namely,

$$\begin{bmatrix} \widehat{e}_\theta \\ \widehat{e}_T \end{bmatrix} = \begin{bmatrix} \Pi_\theta e_\theta \\ \Pi_\theta e_T \end{bmatrix} - \mathrm{S} \begin{bmatrix} \Pi_M e_M - \widehat{e}_M \\ \Pi_{\mathrm{w}} e_{\mathrm{w}} - \widehat{e}_{\mathrm{w}} \end{bmatrix} \boldsymbol{n} \qquad \text{on } \partial\mathcal{T}_h. \tag{2.6b}$$

These equations hold as a direct consequence of the parallelism between the definition of the numerical traces of the HDG method, (2.3b), and the definition of the projection, (2.4e).

The *simplicity* of the above equations we have been referring to resides in the fact that they differ from the HDG approximation *only* in the definition of their right-hand sides which, moreover, involve only *volume* integrals of the approximation errors δ_T, δ_M and δ_θ.

2.1.4 Approximation properties of the projection $\mathbf{\Pi}_h$

In this subsection we state a theorem displaying the approximation properties of the projection $\mathbf{\Pi}_h = (\Pi_T, \Pi_M, \Pi_\theta, \Pi_{\mathrm{w}})$. First, we need to introduce some notation. Let $K = (x_L, x_R)$ be an element of $\mathcal{T}_h$. For any function z on K, we define $z^- := z(x_L)$, $z^+ := z(x_R)$, We denote the usual norm and seminorm on a Sobolev space $H^s(D)$ by $\|\cdot\|_{s,D}$ and $|\cdot|_{s,D}$, respectively. We drop the first subindex if $s = 0$, and the second one if $D = \Omega$ or $D = \mathcal{T}_h$. We also define the seminorm of a vector-valued function (u_1, u_2, u_3, u_4) as

$$|(u_1, u_2, u_3, u_4)|_{s,D} := |u_1|_{s,D} + |u_2|_{s,D} + |u_3|_{s,D} + |u_4|_{s,D}.$$

Its norm is defined similarly.

Theorem 2.1.1. *On each $K \in \mathcal{T}_h$ with stabilization function S, we have for any s in $[1, k+1]$ that*

$$\|(\delta_T, \delta_M, \delta_\theta, \delta_{\mathrm{w}})\| \le C\, C_{\mathrm{S}} h^s |(T, M, \theta, \mathrm{w})|_s$$

Here, C is a constant independent of the discretization parameters and $(T, M, \theta, \mathrm{w})$, and C_{S} is given by

$$\begin{aligned} C_{\mathrm{S}} := {} & \|(\mathrm{S}^+ + \mathrm{S}^-)^{-1}\|_\infty + \|(\mathrm{S}^+ + \mathrm{S}^-)^{-1}\mathrm{S}^+\|_\infty + \|(\mathrm{S}^+\mathrm{S}^-)^{-1}\mathrm{S}^-\|_\infty \\ & + \|\mathrm{S}^+(\mathrm{S}^+ + \mathrm{S}^-)^{-1}\|_\infty + \|\mathrm{S}^-(\mathrm{S}^+ + \mathrm{S}^-)^{-1}\|_\infty + \|\mathrm{S}^+(\mathrm{S}^+ + \mathrm{S}^-)^{-1}\mathrm{S}^-\|_\infty. \end{aligned}$$

A detailed proof of this result is given in Section 2.2. Let us note that we stated the above result for $(T, M, \theta, \mathrm{w})$ merely for notational convenience. In fact, the result remains valid if we replace $(T, M, \theta, \mathrm{w})$ with any $(u_1, u_2, u_3, u_4) \in [H^{s+1}(K)]^4$.

2.1.5 Superconvergence of the projection of the errors

Here, we present an estimate of the following norm of the projection of the errors,

$$\|\boldsymbol{\Pi e}\| := (\|\Pi_T e_T\|^2 + \|\Pi_M e_M\|^2 + \|\Pi_\theta e_\theta\|^2 + \|\Pi_{\mathrm{w}} e_{\mathrm{w}}\|^2)^{\frac{1}{2}},$$

in terms of the following norm of the approximation error of the projection

$$\|\boldsymbol{\delta}\| := (\|\delta_T\|^2 + \|\delta_M\|^2 + \|\delta_\theta\|^2 + \|\delta_{\mathrm{w}}\|^2)^{\frac{1}{2}}.$$

It is stated in terms of the solution of the so-called dual problem we define next.

For any given $(\eta_T, \eta_M, \eta_\theta, \eta_{\mathrm{w}}) \in [L^2(\Omega)]^4$, the associated dual-problem is:

$$\psi_{\mathrm{w}}' = \psi_\theta - d^2\psi_T/GA + \eta_T \quad \text{in } \Omega, \tag{2.7a}$$

$$\psi_\theta' = \psi_M/EI - \eta_M \quad \text{in } \Omega, \tag{2.7b}$$

$$\psi_M' = \psi_T + \eta_\theta \quad \text{in } \Omega, \tag{2.7c}$$

$$\psi_T' = -\eta_{\mathrm{w}} \quad \text{in } \Omega, \tag{2.7d}$$

$$\psi_{\mathrm{w}} = 0 \quad \text{on } \partial\Omega, \tag{2.7e}$$

$$\psi_\theta = 0 \quad \text{on } \partial\Omega. \tag{2.7f}$$

The key inequality we use about the solution of this problem is the following elliptic regularity result:

Theorem 2.1.2.

$$\|\psi_T\|_1 + \|\psi_M\|_1 + \|\psi_\theta\|_1 + \|\psi_{\mathrm{w}}\|_1 \le C_{reg} \left\|(\eta_T, \eta_M, \eta_\theta, \eta_{\mathrm{w}})\right\|, \tag{2.8}$$

where the constant C_{reg} is independent of the data $(\eta_T, \eta_M, \eta_\theta, \eta_{\mathrm{w}})$ and of the thickness d.

The proof is based on the following simple lemma.

Lemma 2.1.3. *Let $f \in L^2(\Omega)$ and set $f^n(x) := \int_0^x f^{n-1}(t)\,dt$, for $n \ge 1$ with $f^0 := f$. Then $\|f^n\| \le \|f\|$.*

We are now ready to prove Theorem 2.1.2.

Proof. A straightforward computation shows that the solution of (2.7) is given by

$$\begin{aligned}
\psi_T &= -\eta^1_{\mathcal{W}} + c_T, \quad \psi_M = -\eta^2_{\mathcal{W}} + \eta^1_\theta + c_T x + c_M, \\
\psi_\theta &= \int_0^x \frac{1}{EI}(-\eta^2_{\mathcal{W}} + \eta^1_\theta + c_T t + c_M)\, dt - \eta^1_M + c_\theta, \\
\psi_{\mathcal{W}} &= \int_0^x \int_0^t \frac{1}{EI}(-\eta^2_{\mathcal{W}} + \eta^1_\theta + c_T s + c_M)\, dsdt \\
&\quad + \int_0^x \frac{d^2}{GA}\eta^1_{\mathcal{W}}\, dt - \eta^2_M + c_\theta x + c_{\mathcal{W}}.
\end{aligned} \tag{2.9}$$

where c_t, c_M, c_θ, and $c_{\mathcal{W}}$ are constants of integration. Using the boundary conditions $\psi_\theta(0) = \psi_{\mathcal{W}}(0) = 0$, we immediately get $c_\theta = c_{\mathcal{W}} = 0$. The remaining boundary conditions, $\psi_\theta(1) = \psi_{\mathcal{W}}(1) = 0$, yield the following linear system for c_T and c_M:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} \begin{pmatrix} c_T \\ c_M \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix},$$

where

$$a_{11} = \int_0^1 \frac{1}{EI} t\, dt,\; a_{12} = \int_0^1 \frac{1}{EI}\, dt,\; a_{21} = \int_0^1\int_0^t \frac{1}{EI} s\, dsdt,\; a_{22} = \int_0^1\int_0^t \frac{1}{EI}\, dsdt,$$

and $b_1 = \int_0^1 \frac{1}{EI}(\eta^2_{\mathcal{W}} - \eta^1_\theta)\, dt + \int_0^1 \eta_M\, dt$ and

$$b_2 = \int_0^1 \int_0^t \frac{1}{EI}(\eta^2_{\mathcal{W}} - \eta^1_\theta)\, dsdt - \int_0^1 \frac{d^2}{GA}\eta^1_{\mathcal{W}}\, dt + \int_0^1 \eta^1_M\, dt.$$

By Lemma 2.1.3 we can easily show that $|b_i| \leq C \|\eta\|$ for $i = 1, 2$, where the constant C depends solely on EI and GA. Note that the dependence on d can be suppressed since $d \leq 1$. Furthermore, we can also suppress the dependence of C on EI and GA since these, and hence their reciprocals, are functions of order one on Ω. Thus, we have that $|c_T| \leq C \|\eta\|$, and that $|c_M| \leq C \|\eta\|$. Inserting these estimates into (2.9) and applying Lemma 2.1.3 once more, we finally obtain (2.8).

□

We are now ready to state our main result.

Theorem 2.1.4. *For $k \geq 1$, we have that, if h sufficiently small,*

$$\|\mathbf{\Pi} \boldsymbol{e}\| \leq C\, C_{reg} h \, \|\boldsymbol{\delta}\|\,.$$

For $k = 0$, we have

$$\|\boldsymbol{\Pi e}\| \leq C\, C_{reg}\, \|\boldsymbol{\delta}\|\,.$$

Here C is a constant independent of the data of the problem and of the discretization parameters.

2.1.6 A priori error estimates

Next we present an estimate for the following norm of the errors:

$$\|\boldsymbol{e}\| := (\|e_T\|^2 + \|e_M\|^2 + \|e_\theta\|^2 + \|e_{\mathrm{w}}\|^2)^{\frac{1}{2}},$$

which is a direct consequence of the last result.

Theorem 2.1.5. *Suppose that the exact solution $(T, M, \theta, \mathrm{w})$ belongs to $[H^{k+1}(\Omega)]^4$. Then, for $k \geq 1$ and h sufficiently small, we have*

$$\|\boldsymbol{e}\| \leq (1 + C\, C_{reg} h)\, \|\boldsymbol{\delta}\|\,.$$

For $k = 0$, we have

$$\|\boldsymbol{e}\| \leq (1 + C\, C_{reg})\, \|\boldsymbol{\delta}\|\,.$$

Here C is a constant independent of the data of the problem and of the discretization parameters.

Note that the error estimate appearing in the above theorem shows that, if the matrix-valued function S is chosen is such a way that C_{S} is uniformly bounded, the HDG method is optimally convergent, that is, $\|u - u_h\| = \mathcal{O}(h^{k+1})$ for smooth solutions for each $u = T, M, \theta, \mathrm{w}$, and *locking-free.* The method is locking-free because the constant C_{S} does not depend on the parameter d and because the seminorms appearing on the right-hand side of the estimate can be bounded uniformly with respect to d; see [10].

2.1.7 Superconvergence at the nodes

Our next result states that the numerical traces of the HDG solution superconverge. To state this result we need to introduce the Green's functions associated with the problem under consideration. For any superindex $\star = T, M, \theta, w$ and any $y \in (0,1)$ we define $(\Phi^\star_{T,y}, \Phi^\star_{M,y}, \Phi^\star_{\theta,y}, \Phi^\star_{w,y})$ as the solution of

$$\frac{d\Phi^\star_{w,y}}{dx} = \Phi^\star_{\theta,y} - d^2\frac{\Phi^\star_{T,y}}{GA}, \quad \frac{d\Phi^\star_{\theta,y}}{dx} = \frac{\Phi^\star_{M,y}}{EI}, \quad \frac{d\Phi^\star_{M,y}}{dx} = \Phi^\star_{T,y}, \quad \frac{d\Phi^\star_{T,y}}{dx} = 0, \tag{2.10a}$$

in $(0,y)\cup(y,1)$ that satisfies the boundary conditions

$$\Phi^\star_{w,y}(0) = \Phi^\star_{w,y}(1) = \Phi^\star_{\theta,y}(0) = \Phi^\star_{\theta,y}(1) = 0, \tag{2.10b}$$

and the jump conditions

$$[\![\Phi^\star_{w,y}]\!](y) = -\delta_{\star T},\ \ [\![\Phi^\star_{\theta,y}]\!](y) = \delta_{\star M},\ \ [\![\Phi^\star_{M,y}]\!](y) = -\delta_{\star\theta},\ \ [\![\Phi^\star_{T,y}]\!](y) = \delta_{\star w}. \tag{2.10c}$$

Here, $\delta_{ab} = 1$ if $a = b$, and $\delta_{ab} = 0$ otherwise. The *jump* operator, $[\![\cdot]\!]$, is defined by

$$[\![\varphi]\!](x) := \varphi(x^-) - \varphi(x^+) \qquad \text{for } x \in \mathscr{E}_h$$

where $\varphi(x^\mp) := \lim_{\epsilon\downarrow 0}\varphi(x \mp \epsilon)$. We define, for $z \in \{0,1\}$,

$$(\Phi^\star_{T,z}, \Phi^\star_{M,z}, \Phi^\star_{\theta,z}, \Phi^\star_{w,z}) := \lim_{y\to z}(\Phi^\star_{T,y}, \Phi^\star_{M,y}, \Phi^\star_{\theta,y}, \Phi^\star_{w,y}).$$

We are now ready to present our superconvergence result of the numerical traces.

Theorem 2.1.6. *Under the same assumptions as in Theorem 2.1.5, we have*

$$|(u - \widehat{u}_h)(x_i)| \le C_{k-1}h^k|(T,M,\theta,w)|_{k+1}\,\|\boldsymbol{\delta}^u_i\| + C\,\|\boldsymbol{e}\|\,\|\boldsymbol{\delta}^u_i\|$$

for $u = T, M, \theta, w$, and $i = 0, 1, \ldots, N$. Here

$$\begin{aligned}\boldsymbol{\delta}^u_i := (&\left\|\Phi^u_{T,x_i} - \Pi_T\Phi^u_{T,x_i}\right\|^2 + \left\|\Phi^u_{M,x_i} - \Pi_M\Phi^u_{M,x_i}\right\|^2 \\ &+ \left\|\Phi^u_{\theta,x_i} - \Pi_\theta\Phi^u_{\theta,x_i}\right\|^2 + \left\|\Phi^u_{w,x_i} - \Pi_w\Phi^u_{w,x_i}\right\|^2)^{\frac{1}{2}},\end{aligned}$$

and C_{k-1} is a constant that depends solely on the polynomial degree k.

Note that, for any given $k \ge 0$, if q, EI, GA are very smooth functions in $\mathcal{T}_h$, the exact solution (T,M,θ,w) belongs to $\left[H^{k+1}(\Omega)\right]^4$; see [10]. This regularity result is also valid for the Green's Functions since in this case we take $q = 0$. Hence, we can assume that $(\Phi^u_{T,x_i}, \Phi^u_{M,x_i}, \Phi^u_{\theta,x_i}, \Phi^u_{w,x_i})$ belongs to $\left[H^{k+1}(\Omega)\right]^4$. As a consequence,

$\|\boldsymbol{\delta}_i^u\| = \boldsymbol{O}(h^{k+1})$ and the above result states that, if the constant $C_{\rm S}$ is uniformly bounded, *all* the numerical traces superconverge with order $2k+1$ at each node. A similar result was proved for the DG methods for Timoshenko beams studied in [10] and [38].

Let us point out that if the data EI and GA are constants on each $K \in \mathcal{T}_h$, then, for $k \geq 3$, Theorem 2.1.6 implies that $\widehat{e}_u(x_i) = 0$, for $u = T, M, \theta$, or w and for any node x_i. Indeed, in this case the Green's functions are piecewise polynomials of degree at most 3 and hence $|(\Phi^u_{T,x_i}, \Phi^u_{M,x_i}, \Phi^u_{\theta,x_i}, \Phi^u_{\mathrm{w},x_i})|_{k+1} = 0$.

An immediate application of the superconvergence result of Theorem 2.1.6 is an element-by-element postprocessing of the approximate solution provided by the HDG method. *All* the four components of the postprocessed solution converge to the exact solution with order $2k+1$, not only at the nodes, but also *uniformly* at the interior of $\mathcal{T}_h$. For details, see [39].

2.2 Approximation properties of the projection: Proof of Theorem 2.1.1

In this section we provide a detailed proof of Theorem 2.1.1. We only give the proof for $k \geq 1$. The proof for $k = 0$ is easier.

Fix an interval $K \in \mathcal{T}_h$ and set $d_u := u_k - \boldsymbol{\Pi}_u u$, $\;g_u := u - u_k$, for $u = T, M, \theta, \mathrm{w}$; Here u_k denotes the L^2-projection into $\mathcal{P}_k(K)$. Since $\delta_u = g_u + d_u$, we only need to estimate d_u. To do that, we proceed as follows. From the definition of the projection (2.4a)-(2.4e) and the definition of the L^2-projection into $\mathcal{P}_k(K)$, we have

$$(d_T, v_1)_K = (d_M, v_2)_K = (d_\theta, v_3)_K = (d_{\mathrm{w}}, v_4)_K = 0, \tag{2.11a}$$

for all $(v_1, v_2, v_3, v_4) \in [P^{k-1}(K)]^4$, and

$$\begin{bmatrix} d_\theta \\ d_T \end{bmatrix} n - \mathrm{S} \begin{bmatrix} d_M \\ d_{\mathrm{w}} \end{bmatrix} = \begin{bmatrix} g_\theta \\ g_T \end{bmatrix} n - \mathrm{S} \begin{bmatrix} g_M \\ g_{\mathrm{w}} \end{bmatrix} \qquad \text{on } \partial K. \tag{2.11b}$$

By equations (2.11a), we see that we can write $d_u = C_u \, L_k$ for $u = T, M, \theta, \mathrm{w}$, where L_k denotes the scaled Legendre polynomial of degree k. Hence, we can write (2.11b) in the following form

$$\begin{bmatrix} \mathrm{I} & -\mathrm{S}^{+} \\ \mathrm{I} & \mathrm{S}^{-} \end{bmatrix} \begin{bmatrix} C_\theta \\ C_T \\ C_M \\ C_{\mathrm{w}} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} g_\theta \\ g_T \end{bmatrix}^{+} - \mathrm{S}^{+} \begin{bmatrix} g_M \\ g_{\mathrm{w}} \end{bmatrix}^{+} \\ (-1)^k \begin{bmatrix} g_\theta \\ g_T \end{bmatrix}^{-} + (-1)^k \mathrm{S}^{-} \begin{bmatrix} g_M \\ g_{\mathrm{w}} \end{bmatrix}^{-} \end{bmatrix},$$

from above equation we can see that, the system has unique solution if and only if the matrix $(\mathrm{S}^{+} + \mathrm{S}^{-})$ is not singular. Hence we obtain, after some algebraic manipulation,

$$\begin{aligned} \begin{bmatrix} C_M \\ C_{\mathrm{w}} \end{bmatrix} &= (-1)^k (\mathrm{S}^{-} + \mathrm{S}^{+})^{-1} \begin{bmatrix} g_\theta \\ g_T \end{bmatrix}^{-} + (-1)^k (\mathrm{S}^{-} + \mathrm{S}^{+})^{-1} \mathrm{S}^{-} \begin{bmatrix} g_M \\ g_{\mathrm{w}} \end{bmatrix}^{-} \\ &\quad - (\mathrm{S}^{-} + \mathrm{S}^{+})^{-1} \begin{bmatrix} g_\theta \\ g_T \end{bmatrix}^{+} + (\mathrm{S}^{-} + \mathrm{S}^{+})^{-1} \mathrm{S}^{+} \begin{bmatrix} g_M \\ g_{\mathrm{w}} \end{bmatrix}^{+} \\ \begin{bmatrix} C_\theta \\ C_T \end{bmatrix} &= (-1)^k \mathrm{S}^{+} (\mathrm{S}^{-} + \mathrm{S}^{+})^{-1} \begin{bmatrix} g_\theta \\ g_T \end{bmatrix}^{-} + (-1)^k \mathrm{S}^{+} (\mathrm{S}^{-} + \mathrm{S}^{+})^{-1} \mathrm{S}^{-} \begin{bmatrix} g_M \\ g_{\mathrm{w}} \end{bmatrix}^{-} \\ &\quad + \mathrm{S}^{-} (\mathrm{S}^{-} + \mathrm{S}^{+})^{-1} \begin{bmatrix} g_\theta \\ g_T \end{bmatrix}^{+} - \mathrm{S}^{+} (\mathrm{S}^{-} + \mathrm{S}^{+})^{-1} \mathrm{S}^{-} \begin{bmatrix} g_M \\ g_{\mathrm{w}} \end{bmatrix}^{+}, \end{aligned}$$

and conclude that

$$\begin{aligned} \|(d_T, d_M, d_\theta, d_{\mathrm{w}})\|_K &= \|L_k\|_K (|C_T| + |C_M| + |C_\theta| + |C_{\mathrm{w}}|) \\ &\le C_{\mathrm{S}} \, \|L_k\|_K \|(g_T, g_M, \ g_\theta, g_{\mathrm{w}})\|_{\partial K} \\ &\le C_{\mathrm{S}} \, h^{1/2} \, \|(g_T, g_M, \ g_\theta, g_{\mathrm{w}})\|_{\partial K} \\ &\le C \, C_{\mathrm{S}} \, h^s \, |(T, M, \theta, \mathrm{w})|_{s,K}, \end{aligned}$$

for all $1 \le s \le k+1$, by the trace inequality and the approximation properties of the L^2-projection.

By the triangle inequality, we have

$$\|(\delta_T, \delta_M, \delta_\theta, \delta_{\mathrm{w}})\|_K \le \|d_T, d_M, d_\theta, d_{\mathrm{w}}\|_K + \|g_T, g_M, g_\theta, g_{\mathrm{w}}\|_K,$$

and the estimate of Theorem 2.1.1 readily follows. This completes the proof.

2.3 Estimates of the projection of the errors: Proof of Theorem 2.1.4

This subsection is devoted to the proof of Theorem 2.1.4. We proceed in two steps. In the first, we use a key identity obtained by duality to prove Theorem 2.1.4. In the second, we prove the identity.

2.3.1 Step 1: The duality identity and the proof of Theorem 2.1.4

Our proof will is based on the following auxiliary result.

Lemma 2.3.1. *For any* $(\eta_T, \eta_M, \eta_\theta, \eta_w) \in [L^2(\Omega)]^4$, *set*

$$\mathcal{E} := (\Pi_T e_T, \eta_T)_{\mathcal{T}_h} + (\Pi_M e_M, \eta_M)_{\mathcal{T}_h} + (\Pi_\theta e_\theta, \eta_\theta)_{\mathcal{T}_h} + (\Pi_w e_w, \eta_w)_{\mathcal{T}_h}.$$

Then

$$\begin{aligned}
\mathcal{E} = \quad & (\delta_T, \Pi_\theta \psi_\theta)_{\mathcal{T}_h} && - (\Pi_T e_T, \delta_{\psi_\theta})_{\mathcal{T}_h} \\
& -(\delta_T, d^2 \Pi_T \psi_T / GA)_{\mathcal{T}_h} && + (\Pi_T e_T, d^2 \delta_{\psi_T} / GA)_{\mathcal{T}_h} \\
& -(\delta_M, \Pi_M \psi_M / EI)_{\mathcal{T}_h} && + (\Pi_M e_M, \delta_{\psi_M} / EI)_{\mathcal{T}_h} \\
& +(\delta_\theta, \Pi_T \psi_T)_{\mathcal{T}_h} && - (\Pi_\theta e_\theta, \delta_{\psi_T})_{\mathcal{T}_h}.
\end{aligned}$$

Here, on each K, *we take* S^t *as the stabilization function for defining the projection* $(\Pi_T \psi_T, \Pi_M \psi_M, \Pi_\theta \psi_\theta, \Pi_w \psi_w)$.

We delay the proof of this identity to the end of this subsection. We are now ready to prove Theorem 2.1.4.

Proof. (Theorem 2.1.4) We first consider the case $k \geq 1$. Setting

$$(\eta_T, \eta_M, \eta_\theta, \eta_w) = (\Pi_T e_T, \Pi_M e_M, \Pi_\theta e_\theta, \Pi_w e_w)$$

in the identity of Lemma 2.3.1 gives

$$\begin{aligned}
\|\boldsymbol{\Pi e}\|^2 = \quad & (\delta_T, \Pi_\theta \psi_\theta)_{\mathcal{T}_h} && - (\Pi_T e_T, \delta_{\psi_\theta})_{\mathcal{T}_h} \\
& -(\delta_T, d^2 \Pi_T \psi_T / GA)_{\mathcal{T}_h} && + (\Pi_T e_T, d^2 \delta_{\psi_T} / GA)_{\mathcal{T}_h} \\
& -(\delta_M, \Pi_M \psi_M / EI)_{\mathcal{T}_h} && + (\Pi_M e_M, \delta_{\psi_M} / EI)_{\mathcal{T}_h} \\
& +(\delta_\theta, \Pi_T \psi_T)_{\mathcal{T}_h} && - (\Pi_\theta e_\theta, \delta_{\psi_T})_{\mathcal{T}_h}.
\end{aligned}$$

Using the fact that $\Pi_u \psi_u = \psi_u - \delta_{\psi_u}$ for $u = T, M, \theta$, and w, we get

$$\|\mathbf{\Pi e}\|^2 = T_1 + T_2 + T_3 + T_4$$

where

$$\begin{aligned}
T_1 &= \quad (\delta_T, \psi_\theta)_{\mathcal{T}_h} && - (\delta_T, \delta_{\psi_\theta})_{\mathcal{T}_h} && - (\Pi_T e_T, \delta_{\psi_\theta})_{\mathcal{T}_h} \\
T_2 &= -(\delta_T, d^2\psi_T/GA)_{\mathcal{T}_h} && + (\delta_T, d^2\delta_{\psi_T}/GA)_{\mathcal{T}_h} && + (\Pi_T e_T, d^2\delta_{\psi_T}/GA)_{\mathcal{T}_h} \\
T_3 &= -(\delta_M, \psi_M/EI)_{\mathcal{T}_h} && + (\delta_M, \delta_{\psi_M}/EI)_{\mathcal{T}_h} && + (\Pi_M e_M, \delta_{\psi_M}/EI)_{\mathcal{T}_h} \\
T_4 &= \quad (\delta_\theta, \psi_T)_{\mathcal{T}_h} && - (\delta_\theta, \delta_{\psi_T})_{\mathcal{T}_h} && - (\Pi_\theta e_\theta, \delta_{\psi_T})_{\mathcal{T}_h}.
\end{aligned} \tag{2.12}$$

By the orthogonality property of the projection, (2.4), we can rewrite these equation as

$$\begin{aligned}
T_1 &= \quad (\delta_T, \psi_\theta - (\psi_\theta)_{k-1})_{\mathcal{T}_h} && - (\delta_T, \delta_{\psi_\theta})_{\mathcal{T}_h} - (\Pi_T e_T, \delta_{\psi_\theta})_{\mathcal{T}_h} \\
T_2 &= -(\delta_T, d^2(\psi_T/GA - (\psi_T/GA)_{k-1}))_{\mathcal{T}_h} && + (\delta_T, d^2\delta_{\psi_T}/GA)_{\mathcal{T}_h} \\
& && + (\Pi_T e_T, d^2\delta_{\psi_T}/GA)_{\mathcal{T}_h} \\
T_3 &= -(\delta_M, \psi_M/EI - (\psi_M/EI)_{k-1})_{\mathcal{T}_h} && + (\delta_M, \delta_{\psi_M}/EI)_{\mathcal{T}_h} \\
& && + (\Pi_M e_M, \delta_{\psi_M}/EI)_{\mathcal{T}_h} \\
T_4 &= \quad (\delta_\theta, \psi_T - (\psi_T)_{k-1})_{\mathcal{T}_h} && - (\delta_\theta, \delta_{\psi_T})_{\mathcal{T}_h} - (\Pi_\theta e_\theta, \delta_{\psi_T})_{\mathcal{T}_h}.
\end{aligned} \tag{2.13}$$

An estimate on $\|\mathbf{\Pi e}\|$ now follows by estimating T_i for $i = 1, 2, 3, 4$. We only show the details of how to estimate T_2, since the remaining terms can be estimated in a similar fashion. Applying the Cauchy-Schwarz inequality to each term in T_2, we get

$$|T_2| \leq \|\delta_T\| \left\|d^2(\psi_T/GA - (\psi_T/GA)_{k-1})\right\| + (\|\delta_T\| + \|\Pi_T e_T\|) \left\|d^2\delta_{\psi_T}/GA\right\|.$$

By the approximation properties of the L^2-projection, we get that

$$|T_2| \leq C\, hd^2 \, \|\delta_T\| \, \|\psi_T/GA\|_1 + \|d^2/GA\|_\infty (\|\delta_T\| + \|\Pi_T e_T\|) \, \|\delta_{\psi_T}\|,$$

by the fact that $0 < d < 1, GA$ is very smooth, and by Theorem 2.1.1 we get that

$$|T_2| \leq C\, h \, \|\delta_T\| \, \|\psi_T\|_1 \|1/GA\|_1 + C\, h(\|\delta_T\| + \|\Pi_T e_T\|)\, |(\psi_T, \psi_M, \psi_\theta, \psi_w)|_1.$$

By the elliptic regularity estimate (2.8), we have $|(\psi_T, \psi_M, \psi_\theta, \psi_{\mathrm{w}})|_1 \leq C_{reg}\, \|\boldsymbol{\Pi e}\|$, we get that

$$|T_2| \leq C\, C_{reg} h\, \|\delta_T\|\, \|\boldsymbol{\Pi e}\| + C C_{reg} h\, \|\boldsymbol{\Pi e}\|^2 .$$

The remaining terms T_1, T_3, and T_4 can be estimated similarly, and hence we obtain

$$\|\boldsymbol{\Pi e}\|^2 \leq |T_1| + |T_2| + |T_3| + |T_4| \leq C\, C_{reg} h\, \|\boldsymbol{\delta}\|\, \|\boldsymbol{\Pi e}\| + C\, C_{reg} h\, \|\boldsymbol{\Pi e}\|^2 .$$

If we assume that h is small enough so that $C C_{reg} h < 1$ then

$$\|\boldsymbol{\Pi e}\|^2 \leq C\, C_{reg} h\, \|\boldsymbol{\delta}\|\, \|\boldsymbol{\Pi e}\| ,$$

and the first estimate of Theorem (2.1.4) follows.

Next we consider the case $k = 0$. In this case (2.12) is still valid, but we do not have (2.13) since the L^2-projection into polynomials of degree $k-1$ is no longer defined. Nevertheless, we can still estimate T_i for $i = 1, 2, 3, 4$ in their form given by (2.12). We provide the details for only T_1. Applying Cauchy-Schwarz inequality to each term in T_1 we get

$$|T_1| \leq \|\delta_T\|\, \|\psi_\theta\| + (\|\delta_T\| + \|\Pi_T e_T\|)\, \|\delta_{\psi_\theta}\| .$$

By Theorem 2.1.1 we have that

$$|T_1| \leq \|\delta_T\|\, \|\psi_\theta\| + (\|\delta_T\| + \|\Pi_T e_T\|)\, C\, h\, |(\psi_T, \psi_M, \psi_\theta, \psi_{\mathrm{w}})|_1,$$

and, by the elliptic regularity inequality (2.8) we have

$$|T_1| \leq C\, C_{reg}\, \|\boldsymbol{\delta}\|\, \|\boldsymbol{\Pi e}\| + C C_{reg} h\, \|\boldsymbol{\Pi e}\|^2 .$$

Since the remaining terms can be estimated in a similar fashion, we obtain

$$\|\boldsymbol{\Pi e}\|^2 \leq C C_{reg}\, \|\boldsymbol{\delta}\|\, \|\boldsymbol{\Pi e}\| + C C_{reg} h\, \|\boldsymbol{\Pi e}\|^2 .$$

The second estimate of Theorem (2.1.4) now follows if we assume that $C C_{reg} h < 1$. This completes the proof. $\square$

2.3.2 Step 2: Proof of the duality identity of Lemma 2.3.1

To prove Lemma 2.3.1, we begin by obtaining a couple of auxiliary identities. The first is the following.

Lemma 2.3.2. *Let* $(v_1, v_2, v_3, v_4) \in [H^1(\Omega)]^4$ *and we take* S^t *as the stabilization function of the projection* $(\Pi_T v_1, \Pi_M v_2, \Pi_\theta v_3, \Pi_{\mathrm{w}} v_4)$ *. Then*

$$\begin{aligned} &-\langle \widehat{e}_\theta - e_\theta \,,\, \delta_{v_2}\boldsymbol{n}\rangle_{\partial\mathfrak{T}_h} \;+\langle \widehat{e}_M - e_M \,,\, \delta_{v_3}\boldsymbol{n}\rangle_{\partial\mathfrak{T}_h} \\ &-\langle \widehat{e}_T - e_T \,,\, \delta_{v_4}\boldsymbol{n}\rangle_{\partial\mathfrak{T}_h} + \langle \widehat{e}_{\mathrm{w}} - e_{\mathrm{w}} \,,\, \delta_{v_1}\boldsymbol{n}\rangle_{\partial\mathfrak{T}_h} = 0. \end{aligned}$$

Proof. Let Θ be the left-hand side of the identity we want to prove, that is,

$$\Theta := -\langle \begin{bmatrix} \widehat{e}_\theta - e_\theta \\ \widehat{e}_T - e_T \end{bmatrix} , \begin{bmatrix} \delta_{v_2} \\ \delta_{v_4} \end{bmatrix} \boldsymbol{n}\rangle_{\partial\mathfrak{T}_h} + \langle \begin{bmatrix} \widehat{e}_M - e_M \\ \widehat{e}_{\mathrm{w}} - e_{\mathrm{w}} \end{bmatrix} , \begin{bmatrix} \delta_{v_3} \\ \delta_{v_1} \end{bmatrix} \boldsymbol{n}\rangle_{\partial\mathfrak{T}_h}.$$

Noting that $\widehat{e}_u - e_u = u_h - \widehat{u}_h$ for $u = T, M, \theta, \mathrm{w}$, and that, by the definition of the numerical traces (2.3b), we have

$$\begin{bmatrix} \widehat{e}_\theta - e_\theta \\ \widehat{e}_T - e_T \end{bmatrix} = \begin{bmatrix} \theta_h - \widehat{\theta}_h \\ T_h - \widehat{T}_h \end{bmatrix} = \mathrm{S} \begin{bmatrix} M_h - \widehat{M}_h \\ \mathrm{w}_h - \widehat{\mathrm{w}}_h \end{bmatrix} \boldsymbol{n},$$

we get

$$\Theta = -\langle \mathrm{S} \begin{bmatrix} M_h - \widehat{M}_h \\ \mathrm{w}_h - \widehat{\mathrm{w}}_h \end{bmatrix} , \begin{bmatrix} \delta_{v_2} \\ \delta_{v_4} \end{bmatrix} \rangle_{\partial\mathfrak{T}_h} + \langle \begin{bmatrix} M_h - \widehat{M}_h \\ \mathrm{w}_h - \widehat{\mathrm{w}}_h \end{bmatrix} , \mathrm{S}^t \begin{bmatrix} \delta_{v_2} \\ \delta_{v_4} \end{bmatrix} \rangle_{\partial\mathfrak{T}_h} = 0$$

because $\begin{bmatrix} \delta_{v_3} \\ \delta_{v_1} \end{bmatrix} = \mathrm{S}^t \begin{bmatrix} \delta_{v_2} \\ \delta_{v_4} \end{bmatrix} \boldsymbol{n}$, by (2.4e). This completes the proof. □

Lemma 2.3.3. *Let* $(u_1, u_2, u_3, u_4), (v_1, v_2, v_3, v_4) \in [L^2(\Omega)]^4$ *with the stabilization functions* S, S^t *respectively. Then*

$$-\langle \delta_{u_3} \,,\, \delta_{v_2}\boldsymbol{n}\rangle_{\partial\mathfrak{T}_h} + \langle \delta_{u_2} \,,\, \delta_{v_3}\boldsymbol{n}\rangle_{\partial\mathfrak{T}_h} - \langle \delta_{u_1} \,,\, \delta_{v_4}\boldsymbol{n}\rangle_{\partial\mathfrak{T}_h} + \langle \delta_{u_4} \,,\, \delta_{v_1}\boldsymbol{n}\rangle_{\partial\mathfrak{T}_h} = 0.$$

Proof. Let Θ be the left-hand side of the identity we want to prove, that is,

$$\Theta := -\langle \begin{bmatrix} \delta_{u_3} \\ \delta_{u_1} \end{bmatrix} , \begin{bmatrix} \delta_{v_2} \\ \delta_{v_4} \end{bmatrix} \boldsymbol{n}\rangle_{\partial\mathfrak{T}_h} + \langle \begin{bmatrix} \delta_{u_2} \\ \delta_{u_4} \end{bmatrix} , \begin{bmatrix} \delta_{v_3} \\ \delta_{v_1} \end{bmatrix} \boldsymbol{n}\rangle_{\partial\mathfrak{T}_h}.$$

Since, by (2.4e), we have that $\begin{bmatrix}\delta_{u_3}\\ \delta_{u_1}\end{bmatrix} = \mathrm{S}\begin{bmatrix}\delta_{u_2}\\ \delta_{u_4}\end{bmatrix}\boldsymbol{n}$, and $\begin{bmatrix}\delta_{v_3}\\ \delta_{v_1}\end{bmatrix} = \mathrm{S}^t\begin{bmatrix}\delta_{v_2}\\ \delta_{v_4}\end{bmatrix}\boldsymbol{n}$, we readily obtain that

$$\Theta = -\langle \mathrm{S}\begin{bmatrix}\delta_{u_2}\\ \delta_{u_4}\end{bmatrix} n\,,\ \begin{bmatrix}\delta_{v_2}\\ \delta_{v_4}\end{bmatrix}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} + \langle \begin{bmatrix}\delta_{u_2}\\ \delta_{u_4}\end{bmatrix}\,,\ \mathrm{S}^t\begin{bmatrix}\delta_{v_2}\\ \delta_{v_4}\end{bmatrix}\rangle_{\partial\mathcal{T}_h} = 0.$$

This completes the proof. $\square$

We are now ready to prove Lemma 2.3.1.

Proof. (Lemma 2.3.1) By the definition of $\mathcal{E}$ and the equations defining the dual solution (2.7), we have

$$\begin{aligned}\mathcal{E} = \quad &(\Pi_T e_T, \psi'_{\mathrm{w}})_{\mathcal{T}_h} - (\Pi_T e_T, \psi_\theta)_{\mathcal{T}_h} + (\Pi_T e_T, d^2\psi_T/GA)_{\mathcal{T}_h}\\ &- (\Pi_M e_M, \psi'_\theta)_{\mathcal{T}_h} + (\Pi_M e_M, \psi_M/EI)_{\mathcal{T}_h}\\ &+ (\Pi_\theta e_\theta, \psi'_M)_{\mathcal{T}_h} - (\Pi_\theta e_\theta, \psi_T)_{\mathcal{T}_h} - (\Pi_{\mathrm{w}} e_{\mathrm{w}}, \psi'_T)_{\mathcal{T}_h}.\end{aligned}$$

Since, for any pair (e_u, ψ_v), we have that

$$\begin{aligned}(\Pi_u e_u, \psi'_v)_{\mathcal{T}_h} &= (\Pi_u e_u, (\Pi_v\psi_v)')_{\mathcal{T}_h} + (\Pi_u u, \delta'_{\psi_v})\\ &= (\Pi_u e_u, (\Pi_v\psi_v)')_{\mathcal{T}_h} - ((\Pi_u e_u)', \delta_{\psi_v})_{\mathcal{T}_h} + \langle \Pi_u u\,,\ \delta_{\psi_v}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}\\ &= (\Pi_u e_u, (\Pi_v\psi_v)')_{\mathcal{T}_h} + \langle \Pi_u u\,,\ \delta_{\psi_v}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h},\end{aligned}$$

by the orthogonality properties (2.4a) to (2.4d) of the projection. Hence

$$\begin{aligned}\mathcal{E} = \quad &(\Pi_T e_T, (\Pi_{\mathrm{w}}\psi_{\mathrm{w}})')_{\mathcal{T}_h} - (\Pi_T e_T, \psi_\theta)_{\mathcal{T}_h} + (\Pi_T e_T, d^2\psi_T/GA)_{\mathcal{T}_h}\\ &- (\Pi_M e_M, (\Pi_\theta\psi_\theta)')_{\mathcal{T}_h} + (\Pi_M e_M, \psi_M/EI)_{\mathcal{T}_h}\\ &+ (\Pi_\theta e_\theta, (\Pi_M\psi_M)')_{\mathcal{T}_h} - (\Pi_\theta e_\theta, \psi_T)_{\mathcal{T}_h} - (\Pi_{\mathrm{w}} e_{\mathrm{w}}, (\Pi_T\psi_T)')_{\mathcal{T}_h}\\ &+ \langle \Pi_T e_T\,,\ \delta_{\psi_{\mathrm{w}}}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle \Pi_M e_M\,,\ \delta_{\psi_\theta}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}\\ &+ \langle \Pi_\theta e_\theta\,,\ \delta_{\psi_M}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle \Pi_{\mathrm{w}} e_{\mathrm{w}}\,,\ \delta_{\psi_T}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}.\end{aligned}$$

Taking $(v_1, v_2, v_3, v_4) = (-\Pi_T\psi_T, \Pi_M\psi_M, -\Pi_\theta\psi_\theta, \Pi_{\mathrm{w}}\psi_{\mathrm{w}})$ in the error equations (2.5) and carrying out some very simple algebraic manipulations, we obtain

$$
\begin{aligned}
\mathcal{E} = H &+ (\delta_T, \Pi_\theta \psi_\theta)_{\mathcal{T}_h} && - (\Pi_T e_T, \delta_{\psi_\theta})_{\mathcal{T}_h} \\
&+ (\Pi_T e_T, d^2 \delta_{\psi_T}/GA)_{\mathcal{T}_h} && - (\delta_T, d^2 \Pi_T \psi_T/GA)_{\mathcal{T}_h} \\
&+ (\Pi_M e_M, \delta_{\psi_M}/EI)_{\mathcal{T}_h} && - (\delta_M, \Pi_M \psi_M/EI)_{\mathcal{T}_h} \\
&+ (\delta_\theta, \Pi_T \psi_T)_{\mathcal{T}_h} && - (\Pi_\theta e_\theta, \delta_{\psi_T})_{\mathcal{T}_h}
\end{aligned}
$$

where

$$
\begin{aligned}
H := \quad & \langle \widehat{e}_T \,,\, \Pi_{\mathcal{W}} \psi_{\mathcal{W}} \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} && + \langle \Pi_T e_T \,,\, \delta_{\psi_{\mathcal{W}}} \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} \\
- & \langle \widehat{e}_M \,,\, \Pi_\theta \psi_\theta \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} && - \langle \Pi_M e_M \,,\, \delta_{\psi_\theta} \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} \\
+ & \langle \widehat{e}_\theta \,,\, \Pi_M \psi_M \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} && + \langle \Pi_T e_\theta \,,\, \delta_{\psi_M} \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} \\
- & \langle \widehat{e}_{\mathcal{W}} \,,\, \Pi_T \psi_T \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} && - \langle \Pi_{\mathcal{W}} e_{\mathcal{W}} \,,\, \delta_{\psi_T} \boldsymbol{n} \rangle_{\partial \mathcal{T}_h}.
\end{aligned}
$$

It remains to show that $H = 0$.

Since ψ_M and $\psi_{\mathcal{W}}$ are single-valued on $\mathscr{E}_h$, and $\psi_{\mathcal{W}} = 0$ on $\partial\Omega$, we can take $\mathsf{m} := \psi_M$ and $\mathsf{w} := \psi_{\mathcal{W}}$ in the error equations (2.5e) and (2.5f), respectively to get

$$
\langle \widehat{e}_\theta \,,\, \psi_M \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} = \langle \widehat{e}_T \,,\, \psi_{\mathcal{W}} \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} = 0.
$$

Moreover, since $\widehat{e}_M$ and $\widehat{e}_{\mathcal{W}}$ are single valued on $\mathscr{E}_h$, and $\widehat{e}_{\mathcal{W}} = 0$, $\psi_\theta = 0$ on $\partial\mathcal{T}_h$, we have

$$
\langle \widehat{e}_M \,,\, \psi_\theta \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} = \langle \widehat{e}_{\mathcal{W}} \,,\, \psi_T \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} = 0.
$$

This implies that

$$
\begin{aligned}
H = \quad & \langle \widehat{e}_T \,,\, (\Pi_{\mathcal{W}} \psi_{\mathcal{W}} - \psi_{\mathcal{W}}) \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} && + \langle \Pi_T e_T \,,\, \delta_{\psi_{\mathcal{W}}} \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} \\
& - \langle \widehat{e}_M \,,\, (\Pi_\theta \psi_\theta - \psi_\theta) \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} && - \langle \Pi_M e_M \,,\, \delta_{\psi_\theta} \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} \\
& + \langle \widehat{e}_\theta \,,\, (\Pi_M \psi_M - \psi_M) \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} && + \langle \Pi_\theta e_\theta \,,\, \delta_{\psi_M} \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} \\
& - \langle \widehat{e}_{\mathcal{W}} \,,\, (\Pi_T \psi_T - \psi_T) \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} && - \langle \Pi_{\mathcal{W}} e_{\mathcal{W}} \,,\, \delta_{\psi_T} \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} \\
= & - \langle \widehat{e}_T - \Pi_T e_T \,,\, \delta_{\psi_{\mathcal{W}}} \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} && + \langle \widehat{e}_M - \Pi_M e_M \,,\, \delta_{\psi_\theta} \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} \\
& - \langle \widehat{e}_\theta - \Pi_\theta e_\theta \,,\, \delta_{\psi_M} \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} && + \langle \widehat{e}_{\mathcal{W}} - \Pi_{\mathcal{W}} e_{\mathcal{W}} \,,\, \delta_{\psi_T} \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} \\
= \quad & H_1 + H_2,
\end{aligned}
$$

where

$$
\begin{aligned}
H_1 = &- \langle \widehat{e}_T - e_T \, , \, \delta_{\psi_{\mathtt{w}}} \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} + \langle \widehat{e}_M - e_M \, , \, \delta_{\psi_\theta} \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} \\
&- \langle \widehat{e}_\theta - e_\theta \, , \, \delta_{\psi_M} \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} \; + \langle \widehat{e}_{\mathtt{w}} - e_{\mathtt{w}} \, , \, \delta_{\psi_T} \boldsymbol{n} \rangle_{\partial \mathcal{T}_h}, \\
H_2 = &- \langle \delta_T \, , \, \delta_{\psi_{\mathtt{w}}} \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} \qquad + \langle \delta_M \, , \, \delta_{\psi_\theta} \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} \\
&- \langle \delta_\theta \, , \, \delta_{\psi_M} \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} \qquad + \langle \delta_{\mathtt{w}} \, , \, \delta_{\psi_T} \boldsymbol{n} \rangle_{\partial \mathcal{T}_h}.
\end{aligned}
$$

But $H_1 = 0$ by Lemma 2.3.2 with $(v_1, v_2, v_3, v_4) = (\psi_T, \psi_M, \psi_\theta, \psi_{\mathtt{w}})$, and $H_2 = 0$ by Lemma 2.3.3 with $(u_1, u_2, u_3, u_4) = (T, M, \theta, \mathtt{w})$ and $(v_1, v_2, v_3, v_4) = (\psi_T, \psi_M, \psi_\theta, \psi_{\mathtt{w}})$. This completes the proof. $\square$

2.4 Error of the numerical traces: Proof of Theorem 2.1.6

To prove this theorem we proceed in two steps. In the first, we obtain representation formulas for the errors in the numerical traces. In the second, we use approximation results to estimate them. We prove the result only for $k \geq 1$; the proof for the case $k = 0$ is not difficult.

2.4.1 Step 1: Representation of the errors

The following lemma provides a representation formula for the errors in the numerical traces.

Lemma 2.4.1. *Let $x_i \in \mathscr{E}_h$ be an arbitrary node and let Φ^u_{T,x_i}, Φ^u_{M,x_i}, Φ^u_{θ,x_i}, $\Phi^u_{\mathtt{w},x_i}$, for $u = T, M, \theta$, or $\mathtt{w}$, be the functions defined by* (2.10a), (2.10b), *and* (2.10c). *Then*

$$\widehat{e}_u(x_i) = \Gamma^u_1(x_i) + \Gamma^u_2(x_i)$$

where

$$
\begin{aligned}
\Gamma^u_1(x_i) = &\quad (\mathtt{w}' - (\mathtt{w}')_{k-1}, \Phi^u_{T,x_i} - \Pi_T \Phi^u_{T,x_i})_{\mathcal{T}_h} \\
&- (\theta' - (\theta')_{k-1}, \Phi^u_{M,x_i} - \Pi_M \Phi^u_{M,x_i})_{\mathcal{T}_h} \\
&+ (M' - (M')_{k-1}, \Phi^u_{\theta,x_i} - \Pi_\theta \Phi^u_{\theta,x_i})_{\mathcal{T}_h} \\
&- (T' - (T')_{k-1}, \Phi^u_{\mathtt{w},x_i} - \Pi_{\mathtt{w}} \Phi^u_{\mathtt{w},x_i})_{\mathcal{T}_h}, \\
\Gamma^u_2(x_i) = &\quad (e_M/EI, \Phi^u_{M,x_i} - \Pi_M \Phi^u_{M,x_i})_{\mathcal{T}_h} - (e_\theta, \Phi^u_{T,x_i} - \Pi_T \Phi^u_{T,x_i})_{\mathcal{T}_h} \\
&+ (d^2 e_T/GA, \Phi^u_{T,x_i} - \Pi_T \Phi^u_{T,x_i})_{\mathcal{T}_h} - (e_T, \Phi^u_{\theta,x_i} - \Pi_\theta \Phi^u_{\theta,x_i})_{\mathcal{T}_h}.
\end{aligned}
$$

To prove this lemma we need an auxiliary result which establishes a relation between the errors in the numerical traces and the Green's functions.

Lemma 2.4.2. *Set*

$$\Theta_i^u := \langle \widehat{e}_{\mathcal{W}} \, , \, \Phi^u_{T,x_i}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle \widehat{e}_{\theta} \, , \, \Phi^u_{M,x_i}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} + \langle \widehat{e}_{M} \, , \, \Phi^u_{\theta,x_i}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle \widehat{e}_{T} \, , \, \Phi^u_{\mathcal{W},x_i}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}.$$

Then, we have $\Theta_i^u = \Theta_{i,1}^u + \Theta_{i,2}^u + \Theta_{i,3}^u$ *where*

$$\begin{aligned}
\Theta_{i,1}^u = \quad & \langle \widehat{e}_{\mathcal{W}} - e_{\mathcal{W}} \, , \, (\Phi^u_{T,x_i} - v_1)\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle \widehat{e}_{\theta} - e_{\theta} \, , \, (\Phi^u_{M,x_i} - v_2)\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} \\
& + \langle \widehat{e}_{M} - e_{M} \, , \, (\Phi^u_{\theta,x_i} - v_3)\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle \widehat{e}_{T} - e_{T} \, , \, (\Phi^u_{\mathcal{W},x_i} - v_4)\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}, \\
\Theta_{i,2}^u = \quad & (e'_{\mathcal{W}}, \Phi^u_{T,x_i} - v_1)_{\mathcal{T}_h} - (e'_{\theta}, \Phi^u_{M,x_i} - v_2)_{\mathcal{T}_h} \\
& + (e'_{M}, \Phi^u_{\theta,x_i} - v_3)_{\mathcal{T}_h} - (e'_{T}, \Phi^u_{\mathcal{W},x_i} - v_4)_{\mathcal{T}_h}, \\
\Theta_{i,3}^u = \quad & (e_M/EI, \Phi^u_{M,x_i} - v_2)_{\mathcal{T}_h} - (e_{\theta}, \Phi^u_{T,x_i} - v_1)_{\mathcal{T}_h} \\
& + (d^2 e_T/GA, \Phi^u_{T,x_i} - v_1)_{\mathcal{T}_h} - (e_T, \Phi^u_{\theta,x_i} - v_3)_{\mathcal{T}_h},
\end{aligned}$$

for $(v_1, v_2, v_3, v_4) \in [V_h^k]^4$.

Proof. Adding and subtracting the term

$$\langle \widehat{e}_{\mathcal{W}} \, , \, v_1\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle \widehat{e}_{\theta} \, , \, v_2\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} + \langle \widehat{e}_{M} \, , \, v_3\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle \widehat{e}_{T} \, , \, v_4\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}$$

to the original expression for Θ_i^u, we see that

$$\begin{aligned}
\Theta_i^u = \quad & \langle \widehat{e}_{\mathcal{W}} \, , \, (\Phi^u_{T,x_i} - v_1)\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle \widehat{e}_{\theta} \, , \, (\Phi^u_{M,x_i} - v_2)\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} \\
& + \langle \widehat{e}_{M} \, , \, (\Phi^u_{\theta,x_i} - v_3)\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle \widehat{e}_{T} \, , \, (\Phi^u_{\mathcal{W},x_i} - v_4)\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} \\
& + \langle \widehat{e}_{\mathcal{W}} \, , \, v_1\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle \widehat{e}_{\theta} \, , \, v_2\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} \\
& + \langle \widehat{e}_{M} \, , \, v_3\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle \widehat{e}_{T} \, , \, v_4\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}.
\end{aligned}$$

Rewriting the last four terms above by using the error equations (2.5a)-(2.5d), we obtain

$$\begin{aligned}
\Theta_i^u = \quad & \langle \widehat{e}_{\mathcal{W}} \, , \, (\Phi^u_{T,x_i} - v_1)\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle \widehat{e}_{\theta} \, , \, (\Phi^u_{M,x_i} - v_2)\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} \\
& + \langle \widehat{e}_{M} \, , \, (\Phi^u_{\theta,x_i} - v_3)\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle \widehat{e}_{T} \, , \, (\Phi^u_{\mathcal{W},x_i} - v_4)\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} \\
& + (e_{\mathcal{W}}, v'_1)_{\mathcal{T}_h} - (e_{\theta}, v'_2)_{\mathcal{T}_h} + (e_M, v'_3)_{\mathcal{T}_h} - (e_T, v'_4)_{\mathcal{T}_h} \\
& + (e_{\theta}, v_1)_{\mathcal{T}_h} - (d^2 e_T/GA, v_1)_{\mathcal{T}_h} - (e_M/EI, v_2)_{\mathcal{T}_h} + (e_T, v_3)_{\mathcal{T}_h}.
\end{aligned}$$

Note that, by the definition of the Green's functions, we have

$$
\begin{aligned}
(e_{\mathcal{W}}, (\Phi^u_{T,x_i})')_{\mathcal{T}_h} &= 0,\\
(e_\theta, (\Phi^u_{M,x_i})')_{\mathcal{T}_h} &= (e_\theta, \Phi^u_{T,x_i})_{\mathcal{T}_h},\\
(e_M, (\Phi^u_{\theta,x_i})')_{\mathcal{T}_h} &= (e_M, \Phi^u_{M,x_i}/EI)_{\mathcal{T}_h},\\
(e_T, (\Phi^u_{\mathcal{W},x_i})')_{\mathcal{T}_h} &= (e_T, \Phi^u_{\theta,x_i} - d^2\Phi^u_{T,x_i}/GA)_{\mathcal{T}_h}.
\end{aligned}
$$

Inserting these equations into the last expression for Θ^u_i, and rearranging terms, we obtain

$$
\begin{aligned}
\Theta^u_i = \Theta^u_{i,3} &+ \langle \widehat{e}_{\mathcal{W}}\,,\, (\Phi^u_{T,x_i} - v_1)\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle \widehat{e}_\theta\,,\, (\Phi^u_{M,x_i} - v_2)\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}\\
&+ \langle \widehat{e}_M\,,\, (\Phi^u_{\theta,x_i} - v_3)\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle \widehat{e}_T\,,\, (\Phi^u_{\mathcal{W},x_i} - v_4)\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}\\
&- (e_{\mathcal{W}}, (\Phi^u_{T,x_i} - v_1)')_{\mathcal{T}_h} \quad + (e_\theta, (\Phi^u_{M,x_i} - v_2)')_{\mathcal{T}_h}\\
&- (e_M, (\Phi^u_{\theta,x_i} - v_3)')_{\mathcal{T}_h} \quad + (e_T, (\Phi^u_{\mathcal{W},x_i} - v_4)')_{\mathcal{T}_h}.
\end{aligned}
$$

It remains to show that

$$
\begin{aligned}
\Theta^u_{i,1} + \Theta^u_{i,2} = \quad & \langle \widehat{e}_{\mathcal{W}}\,,\, (\Phi^u_{T,x_i} - v_1)\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle \widehat{e}_\theta\,,\, (\Phi^u_{M,x_i} - v_2)\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}\\
&+ \langle \widehat{e}_M\,,\, (\Phi^u_{\theta,x_i} - v_3)\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle \widehat{e}_T\,,\, (\Phi^u_{\mathcal{W},x_i} - v_4)\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}\\
&- (e_{\mathcal{W}}, (\Phi^u_{T,x_i} - v_1)')_{\mathcal{T}_h} \quad + (e_\theta, (\Phi^u_{M,x_i} - v_2)')_{\mathcal{T}_h}\\
&- (e_M, (\Phi^u_{\theta,x_i} - v_3)')_{\mathcal{T}_h} \quad + (e_T, (\Phi^u_{\mathcal{W},x_i} - v_4)')_{\mathcal{T}_h}.
\end{aligned}
$$

This follows by integrating by parts on each of the last four terms. This completes the proof. □

We are now ready to prove our representation result.

Proof. (Lemma 2.4.1) We begin by noting that, by the definition of the Green's functions, (2.10b) and (2.10c), we have

$$\Theta^u_i = \widehat{e}_u(x_i).$$

On the other hand, setting

$$(v_1, v_2, v_3, v_4) = (\Pi_T\Phi^u_{T,x_i}, \Pi_M\Phi^u_{M,x_i}, \Pi_\theta\Phi^u_{\theta,x_i}, \Pi_{\mathcal{W}}\Phi^u_{\mathcal{W},x_i})$$

in Lemma (2.4.2), we obtain

$$\widehat{e}_u(x_i) = \Theta^u_{i,1} + \Theta^u_{i,2} + \Theta^u_{i,3} \tag{2.14}$$

with

$$\begin{aligned}
\Theta^u_{i,1} = \quad & \langle \widehat{e}_{\mathcal{W}} - e_{\mathcal{W}} \,,\, (\Phi^u_{T,x_i} - \Pi_T \Phi^u_{T,x_i})\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle \widehat{e}_\theta - e_\theta \,,\, (\Phi^u_{M,x_i} - \Pi_M \Phi^u_{M,x_i})\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} \\
& + \langle \widehat{e}_M - e_M \,,\, (\Phi^u_{\theta,x_i} - \Pi_\theta \Phi^u_{\theta,x_i})\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle \widehat{e}_T - e_T \,,\, (\Phi^u_{\mathcal{W},x_i} - \Pi_{\mathcal{W}} \Phi^u_{\mathcal{W},x_i})\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}, \\
\Theta^u_{i,2} = \quad & (e'_{\mathcal{W}}, \Phi^u_{T,x_i} - \Pi_T \Phi^u_{T,x_i})_{\mathcal{T}_h} \qquad - (e'_\theta, \Phi^u_{M,x_i} - \Pi_M \Phi^u_{M,x_i})_{\mathcal{T}_h} \\
& + (e'_M, \Phi^u_{\theta,x_i} - \Pi_\theta \Phi^u_{\theta,x_i})_{\mathcal{T}_h} \qquad - (e'_T, \Phi^u_{\mathcal{W},x_i} - \Pi_{\mathcal{W}} \Phi^u_{\mathcal{W},x_i})_{\mathcal{T}_h}, \\
\Theta^u_{i,3} = \quad & (e_M/EI, \Phi^u_{M,x_i} - \Pi_M \Phi^u_{M,x_i})_{\mathcal{T}_h} \qquad - (e_\theta, \Phi^u_{T,x_i} - \Pi_T \Phi^u_{T,x_i})_{\mathcal{T}_h} \\
& + (d^2 e_T/GA, \Phi^u_{T,x_i} - \Pi_T \Phi^u_{T,x_i})_{\mathcal{T}_h} \qquad - (e_T, \Phi^u_{\theta,x_i} - \Pi_\theta \Phi^u_{\theta,x_i})_{\mathcal{T}_h},
\end{aligned}$$

Clearly,

$$\Theta^u_{i,3} = \Gamma^u_2(x_i). \tag{2.15}$$

By Lemma 2.3.2 with $(v_1, v_2, v_3, v_4) = (\Phi_{T,x_i}, \Phi_{M,x_i}, \Phi_{\theta,x_i}, \Phi_{\mathcal{W},x_i})$ we have that

$$\Theta^u_{i,1} = 0. \tag{2.16}$$

By the orthogonality property, (2.4), of the projection we have

$$\begin{aligned}
\Theta^u_{i,2} = \quad & (e'_{\mathcal{W}} - (e'_{\mathcal{W}})_{k-1}, \Phi^u_{T,x_i} - \Pi_T \Phi^u_{T,x_i})_{\mathcal{T}_h} \\
& - (e'_\theta - (e'_\theta)_{k-1}, \Phi^u_{M,x_i} - \Pi_M \Phi^u_{M,x_i})_{\mathcal{T}_h} \\
& + (e'_M - (e'_M)_{k-1}, \Phi^u_{\theta,x_i} - \Pi_\theta \Phi^u_{\theta,x_i})_{\mathcal{T}_h} \\
& - (e'_T - (e'_T)_{k-1}, \Phi^u_{\mathcal{W},x_i} - \Pi_{\mathcal{W}} \Phi^u_{\mathcal{W},x_i})_{\mathcal{T}_h}.
\end{aligned}$$

Since

$$\begin{aligned}
e'_u - (e'_u)_{k-1} &= (u' - u'_h) - (u' - u'_h)_{k-1} \\
&= u' - (u')_{k-1} + (u'_h)_{k-1} - u'_h \\
&= u' - (u')_{k-1},
\end{aligned}$$

we see that

$$\Theta^u_{i,2} = \Gamma^u_1(x_i). \tag{2.17}$$

The result now follows from (2.14), (2.15), (2.16), and (2.17). □

2.4.2 Step 2: Proof of Theorem 2.1.6

We are now ready to prove Theorem 2.1.6.

By Lemma 2.4.1 we have that

$$|\widehat{e}_u(x_i)| \leq |\Gamma_1^u(x_i)| + |\Gamma_2^u(x_i)| \tag{2.18}$$

where $\Gamma_1^u(x_i) = T_1 + T_2 + T_3 + T_4$ with

$$\begin{aligned}
T_1 &= (w' - (w')_{k-1}, \Phi^u_{T,x_i} - \Pi_T \Phi^u_{T,x_i})_{\mathcal{T}_h},\\
T_2 &= -(\theta' - (\theta')_{k-1}, \Phi^u_{M,x_i} - \Pi_M \Phi^u_{M,x_i})_{\mathcal{T}_h},\\
T_3 &= (M' - (M')_{k-1}, \Phi^u_{\theta,x_i} - \Pi_\theta \Phi^u_{\theta,x_i})_{\mathcal{T}_h},\\
T_4 &= -(T' - (T')_{k-1}, \Phi^u_{w,x_i} - \Pi_w \Phi^u_{w,x_i})_{\mathcal{T}_h},
\end{aligned}$$

and $\Gamma_2^u(x_i) = S_1 + S_2 + S_3 + S_4$ with

$$\begin{aligned}
S_1 &= (e_M/EI, \Phi^u_{M,x_i} - \Pi_M \Phi^u_{M,x_i})_{\mathcal{T}_h},\\
S_2 &= -(e_\theta, \Phi^u_{T,x_i} - \Pi_T \Phi^u_{T,x_i})_{\mathcal{T}_h},\\
S_3 &= (d^2 e_T/GA, \Phi^u_{T,x_i} - \Pi_T \Phi^u_{T,x_i})_{\mathcal{T}_h},\\
S_4 &= -(e_T, \Phi^u_{\theta,x_i} - \Pi_\theta \Phi^u_{\theta,x_i})_{\mathcal{T}_h}.
\end{aligned}$$

Let us estimate the term T_1. By the approximation properties of the L^2-projection and Theorem 2.1.1, we get

$$\begin{aligned}
T_1 &\leq \left\|w' - (w')_{k-1}\right\| \;\cdot\; \left\|\Phi^u_{T,x_i} - \Pi_T \Phi^u_{T,x_i}\right\|\\
&\leq C\, h^s |w'|_s \;\left\|\boldsymbol{\delta_i^u}\right\|\\
&\leq C\, h^s |(T, M, \theta, w)|_{s+1} \left\|\boldsymbol{\delta_i^u}\right\|,
\end{aligned}$$

for any $s \in [1, k]$. We estimate the terms T_2, T_2, and T_3 in a similar fashion and obtain

$$|\Gamma_1^u(x_i)| \leq C\, h^s |(T, M, \theta, w)|_{s+1}| \, \|\boldsymbol{\delta}_i^u\|\,.$$

Next, we obtain an estimate on $\Gamma_2^u(x_i)$. We only show how to estimate S_1 since the remaining terms in $\Gamma_2^u(x_i)$ can be estimated similarly. By Theorem 2.1.5, and Lemma 2.1.1, we get

$$S_1 \leq \|e_M/EI\| \;\cdot\; \left\|\Phi^u_{M,x_i} - \Pi_M \Phi^u_{T,x_i}\right\| \leq C\, \|\boldsymbol{e}\| \; \|\boldsymbol{\delta}_i^u\|\,.$$

This implies that

$$|\Gamma_2^u(x_i)| \le C \,\|\boldsymbol{e}\|\, \|\boldsymbol{\delta}_i^u\|\,.$$

Inserting the estimates of $|\Gamma_1^u(x_i)|$ and $|\Gamma_2^u(x_i)|$ with $s = k$ into (2.18) completes the proof of Theorem 2.1.6.

2.5 Numerical results

In this section, we display numerical results to verify our theoretical findings. We solve the equations (2.1) with $q(x) = e^x$, $(EI)(x) = e^x$, $(GA)(x) = e^{-x}$, together with the boundary conditions (2.1b) $w(0) = w(1) = \theta(0) = \theta(1) = 0$. We take $\mathrm{S} := \begin{bmatrix} \alpha_\theta & \tau \\ \tau & -\alpha_T \end{bmatrix}$ to be constant on $\partial\Omega_h$.

We display our numerical results in Tables 2.1 and 2.2. In Table 2.1, we present a history of convergence study for the projection of the errors, namely,

$$\boldsymbol{\Pi e} := (\Pi_T e_T, \Pi_M e_M, \Pi_\theta e_\theta, \Pi_{\mathtt{w}} e_{\mathtt{w}}).$$

Therein, "mesh $= i$" means we employed a uniform mesh with 2^i elements to obtain the results of that particular row of the table. For polynomials degrees $k = 0, 1, 2, 3$ we display the error $\|\boldsymbol{\Pi e}\|$. We also display numerical orders of convergence which are computed as follows. Let $e_u(i)$ denote the error where a mesh with 2^i elements has been employed to obtain the HDG solution. As usual, the order of convergence, r_i, at level i is defined as $r_i := \log\big(e_u(i-1)/e_u(i)\big)/\log 2$. Observe that the results displayed in Table 2.1 validate the superconvergence of order $k+2$ for $k \ge 1$, and optimal convergence for $k = 0$, predicted by Theorem 2.1.4.

In Table 2.2 we carry out a similar study for the errors in the numerical traces. We display the error

$$\|\,\widehat{e}\,\|_\infty := \max_{u\in\{T,M,\theta,\mathtt{w}\}} \left(\max_{x\in\mathscr{E}_h} |(u - \widehat{u}_h)(x)| \right),$$

and its order of convergence. We see that the $2k+1$ order superconvergence of the numerical traces predicted by Theorem 2.1.6 is verified.

In these two tables we took the thickness parameter $d = 10^{-2}$ but let us note that in the numerical experiments we observed similar results and exactly the same convergence orders when we took $d = 10^{-8}$.

Table 2.1: History of convergence of the projection of the error.

	$k=0$		$k=1$		$k=2$		$k=3$	
mesh	$\lVert\boldsymbol{\Pi e}\rVert$	order	$\lVert\boldsymbol{\Pi e}\rVert$	order	$\lVert\boldsymbol{\Pi e}\rVert$	order	$\lVert\boldsymbol{\Pi e}\rVert$	order
$\alpha_\theta \equiv 0,\ \alpha_T \equiv 0,\ \tau \equiv 1$								
3	1.44E-01	0.94	4.56E-04	2.95	3.87E-07	4.54	1.78E-09	4.97
4	7.36E-02	0.97	5.81E-05	2.97	2.06E-08	4.24	5.62E-11	4.99
5	3.72E-02	0.98	7.34E-06	2.99	1.22E-09	4.07	1.76E-12	4.99
6	1.87E-02	0.99	9.22E-07	2.99	7.54E-11	4.02	5.50E-14	5.00
$\alpha_\theta \equiv 1,\ \alpha_T \equiv 0,\ \tau \equiv 1$								
3	9.75E-02	0.98	4.62E-04	2.95	5.49E-07	4.31	3.11E-09	4.98
4	4.92E-02	0.99	5.88E-05	2.97	3.20E-08	4.10	9.80E-11	4.99
5	2.47E-02	0.99	7.43E-06	2.99	1.96E-09	4.03	3.07E-12	5.00
6	1.24E-02	1.00	9.33E-07	2.99	1.22E-10	4.01	9.59E-14	5.00
$\alpha_\theta \equiv 0,\ \alpha_T \equiv 1,\ \tau \equiv 1$								
3	1.43E-01	0.94	3.79E-04	2.94	5.77E-07	4.60	1.75E-09	4.96
4	7.29E-02	0.97	4.87E-05	2.96	2.92E-08	4.30	5.52E-11	4.99
5	3.68E-02	0.98	6.17E-06	2.98	1.70E-09	4.10	1.73E-12	5.00
6	1.85E-02	0.99	7.77E-07	2.99	1.04E-10	4.03	5.41E-14	5.00
$\alpha_\theta \equiv 1,\ \alpha_T \equiv 1,\ \tau \equiv 0$								
3	5.56E-02	0.57	1.07E-04	3.16	3.46E-07	4.31	2.16E-09	5.01
4	3.60E-02	0.62	1.25E-05	3.10	1.99E-08	4.12	6.73E-11	5.00
5	2.16E-02	0.74	1.51E-06	3.05	1.21E-09	4.04	2.10E-12	5.00
6	1.21E-02	0.84	1.84E-07	3.03	7.54E-11	4.01	6.56E-14	5.00
$\alpha_\theta \equiv 1,\ \alpha_T \equiv 1,\ \tau \equiv 1$								
3	5.42E-02	0.89	1.74E-04	2.87	5.20E-07	4.33	2.76E-09	4.97
4	2.95E-02	0.88	2.29E-05	2.93	2.99E-08	4.12	8.68E-11	4.99
5	1.59E-02	0.89	2.94E-06	2.96	1.83E-09	4.03	2.72E-12	5.00
6	8.41E-03	0.92	3.72E-07	2.98	1.14E-10	4.01	8.50E-14	5.00

Table 2.2: History of convergence of the numerical traces.

	$k=0$		$k=1$		$k=2$		$k=3$	
mesh	$\Vert\widehat{e}\Vert_\infty$	order	$\Vert\widehat{e}\Vert_\infty$	order	$\Vert\widehat{e}\Vert_\infty$	order	$\Vert\widehat{e}\Vert_\infty$	order
				$\alpha_\theta\equiv 0,\ \alpha_T\equiv 0,\ \tau\equiv 1$				
3	1.43E-01	0.96	4.33E-04	2.95	2.36E-07	4.99	3.21E-11	6.98
4	7.23E-02	0.98	5.52E-05	2.97	7.42E-09	4.99	2.49E-13	7.05
5	3.64E-02	0.99	6.97E-06	2.99	2.33E-10	5.00	1.95E-15	7.00
6	1.83E-02	1.00	8.76E-07	2.99	7.29E-12	5.00	1.52E-17	7.00
				$\alpha_\theta\equiv 1,\ \alpha_T\equiv 0,\ \tau\equiv 1$				
3	1.01E-01	0.97	4.33E-04	2.95	2.36E-07	4.99	3.21E-11	6.98
4	5.10E-02	0.99	5.52E-05	2.97	7.42E-09	4.99	2.43E-13	7.04
5	2.56E-02	0.99	6.97E-06	2.99	2.33E-10	5.00	1.91E-15	6.99
6	1.28E-02	1.00	8.76E-07	2.99	7.29E-12	5.00	1.50E-17	6.99
				$\alpha_\theta\equiv 0,\ \alpha_T\equiv 1,\ \tau\equiv 1$				
3	1.42E-01	0.97	3.58E-04	2.94	3.84E-07	4.97	9.76E-12	6.92
4	7.17E-02	0.98	4.60E-05	2.96	1.21E-08	4.99	7.73E-14	6.98
5	3.61E-02	0.99	5.84E-06	2.98	3.80E-10	4.99	6.08E-16	6.99
6	1.81E-02	1.00	7.36E-07	2.99	1.19E-11	5.00	4.78E-18	6.99
				$\alpha_\theta\equiv 1,\ \alpha_T\equiv 1,\ \tau\equiv 0$				
3	5.30E-02	0.58	8.87E-05	3.22	1.48E-07	4.95	4.21E-11	6.98
4	3.43E-02	0.63	1.01E-05	3.14	4.70E-09	4.98	3.31E-13	6.99
5	2.06E-02	0.74	1.19E-06	3.08	1.48E-10	4.99	2.60E-15	6.99
6	1.15E-02	0.84	1.44E-07	3.04	4.64E-12	4.99	2.03E-17	7.00
				$\alpha_\theta\equiv 1,\ \alpha_T\equiv 1,\ \tau\equiv 1$				
3	5.32E-02	0.92	1.48E-04	2.84	2.36E-07	4.97	4.04E-12	6.88
4	2.81E-02	0.92	1.97E-05	2.91	7.44E-09	4.99	3.24E-14	6.96
5	1.49E-02	0.92	2.53E-06	2.96	2.33E-10	4.99	2.57E-16	6.98
6	7.75E-03	0.94	3.22E-07	2.98	7.31E-12	4.99	2.02E-18	6.99

To study the effect of the penalization parameters τ, α_T, and α_θ on the performance of the method, we display results for a variety of choices of these parameters. Also, to find out if the HDG method is free from shear locking, the thickness of the beam, d, is taken be 10^{-2} and then decreased down to 10^{-8}.

We display our numerical results in Tables 2.3 through 2.12. In Table 2.3, we display results corresponding to the HDG method obtained by setting $\alpha_T = \alpha_\theta = \tau = 1$ on $\partial\mathcal{T}_h$.

The first part of the table shows the numerical results for the problem (2.1) with $d = 10^{-2}$, and the second part for $d = 10^{-8}$. We see that the method converges optimally for all the variables. We also notice that the convergence is independent of the parameter d. This indicates that the method is free from shear locking.

In Table 2.4, we carry out a similar history of convergence study for the SFH method. We see that the method converges optimally for all the variables, and the convergence remains invariant as we decrease d. It is worth noting that we do not see the suboptimal convergence of the same method that was proved, and observed numerically, for the biharmonic problems in two-space dimensions [40].

In order to be able to display an extensive set of numerical results, in Tables 2.5 through 2.12 we report only a summary of what we have observed as a result of our study of the history of convergence. In these tables the column k again shows the polynomial degree of the approximation, and the columns $\|e_u\|_{L^2(\Omega)}$, for $u = T, M, \theta, w$ shows the order of convergence we have observed as a result of the numerical experiments whose details we suppress. The dash "$-$" means that the method failed to converge for that particular combination of the polynomial degree and the choice of the stabilization parameters.

From Table 2.5 we see that whenever $\alpha_T \equiv 0$, $\alpha_\theta \equiv 0$, and τ identically equal to a constant on $\partial\mathcal{T}_h$, the only choice which converges optimally for all the variables is obtained by setting $\tau \equiv O(1)$. Table 2.7 shows that in an analogous situation we have a little more flexibility if we employ an SFH method. That is, if we set $\alpha_T \equiv 0$, $\alpha_\theta \equiv 0$, and $\tau^- \equiv 0$ then we observe optimal convergence for $\tau^+ = 1/h^\mu$ if $\mu = 0, 1, 2$.

In Table 2.6 we consider the other extreme, namely, we take $\tau \equiv 0$, and $\alpha_T = \alpha_\theta \equiv \nu$ on $\partial\mathcal{T}_h$ for some constant $\nu \geq 0$. Once again, we see that the only choice to get optimal convergence for all the unknowns is obtained by setting $\nu = O(1)$.

In Table 2.8 we carry out a more extensive study of what we have done for Table 2.3. We take $\tau \equiv \alpha_\theta \equiv \alpha_T \equiv \nu$ for $\nu = h^2, h, 1, h^{-1}, h^{-2}$. We see that the only choice among these options to get optimal convergence for all the unknowns is the one we considered in Table 2.3, namely, take $\nu = 1$. All the other options result in suboptimal convergence for either M and w, or for T and θ.

In all of the Tables 2.3 through 2.8 we have considered cases where $\alpha_T \equiv \alpha_\theta$. Therefore, in Tables 2.9 to 2.12 we consider possible cases in which these two parameters take different values. We observe similar convergence behavior to what we already observed in the cases we considered in Tables 2.3 through 2.8. An interesting observation that might be worth noting that we did not observe any convergence at all for w_h if $\alpha_\theta \equiv 1$, $\alpha_T \equiv 0$, and $\tau = h^2$, see Table 2.9. Another observation we would like to emphasize is that the choice $\alpha_T \equiv 0$, $\tau \equiv 1$ converges optimally for $\alpha_\theta = h^\mu$ with $\mu = 0, 1, 2$. We have a similar behavior if we keep τ the same but switch the roles of α_T and α_θ. These conclusions are based on Tables 2.11 and 2.12.

2.6 Concluding remarks

We have shown that optimal HDG methods can be devised which are free from shear-locking. We achieved this by a careful study of the relation between the definition of the numerical traces and the corresponding convergence properties of the methods. Key to our analysis was a new projection operator which is tailored to fit the structure of the numerical traces of the HDG method. We have shown that HDG solution superconverges to the projection of the exact solution for all the unknowns. This immediately results in optimal error estimates for all the unknowns. In this sense, the error analysis is simplified only to the study of the approximation properties of the projection operator.

This provides a powerful approach for devising locking-free HDG methods for more challenging problems arising in solid mechanics, like the Reissner-Mindlin plates problem. This constitutes the subject of ongoing work.

Table 2.3: $\tau \equiv 1,\ \alpha_\theta \equiv 1,\ \alpha_T \equiv 1$.

		$\|e_T\|_{L^2(\Omega)}$		$\|e_M\|_{L^2(\Omega)}$		$\|e_\theta\|_{L^2(\Omega)}$		$\|e_W\|_{L^2(\Omega)}$	
k	mesh	error	order	error	order	error	order	error	order
				$d = 10^{-2}$					
	5	1.37E-02	0.82	7.34E-03	0.86	7.41E-03	0.94	1.61E-02	0.99
0	6	7.45E-03	0.88	3.89E-03	0.91	3.78E-03	0.97	8.05E-03	1.00
	7	3.90E-03	0.93	2.01E-03	0.95	1.91E-03	0.99	4.03E-03	1.00
	8	2.00E-03	0.96	1.03E-03	0.97	9.60E-04	0.99	2.02E-03	1.00
	3	1.32E-03	1.98	6.63E-04	1.91	1.27E-03	1.92	6.69E-04	2.01
1	4	3.33E-04	1.99	1.70E-04	1.96	3.27E-04	1.96	1.67E-04	2.00
	5	8.36E-05	1.99	4.31E-05	1.98	8.29E-05	1.98	4.16E-05	2.00
	6	2.09E-05	2.00	1.08E-05	1.99	2.09E-05	1.99	1.04E-05	2.00
	3	1.27E-05	2.98	9.93E-06	2.96	1.33E-05	2.98	4.57E-06	2.99
2	4	1.60E-06	2.99	1.26E-06	2.98	1.68E-06	2.99	5.73E-07	3.00
	5	2.00E-07	2.99	1.58E-07	2.99	2.10E-07	3.00	7.16E-08	3.00
	6	2.51E-08	3.00	1.98E-08	3.00	2.63E-08	3.00	8.94E-09	3.00
	3	6.79E-08	3.95	3.87E-08	3.96	1.49E-07	3.96	1.04E-07	3.97
3	4	4.29E-09	3.98	2.44E-09	3.99	9.46E-09	3.98	6.55E-09	3.99
	5	2.69E-10	3.99	1.53E-10	4.00	5.95E-10	3.99	4.11E-10	3.99
	6	1.69E-11	4.00	9.55E-12	4.00	3.73E-11	4.00	2.57E-11	4.00
				$d = 10^{-8}$					
	5	1.39E-02	0.81	7.40E-03	0.86	7.42E-03	0.94	1.61E-02	0.99
0	6	7.55E-03	0.88	3.93E-03	0.91	3.78E-03	0.97	8.05E-03	1.00
	7	3.96E-03	0.93	2.03E-03	0.95	1.91E-03	0.99	4.03E-03	1.00
	8	2.03E-03	0.96	1.04E-03	0.97	9.61E-04	0.99	2.02E-03	1.00
	3	1.32E-03	1.98	6.64E-04	1.91	1.28E-03	1.92	6.69E-04	2.01
1	4	3.33E-04	1.99	1.70E-04	1.96	3.28E-04	1.96	1.67E-04	2.00
	5	8.35E-05	1.99	4.31E-05	1.98	8.30E-05	1.98	4.16E-05	2.00
	6	2.09E-05	2.00	1.08E-05	1.99	2.09E-05	1.99	1.04E-05	2.00
	3	1.27E-05	2.98	9.92E-06	2.96	1.34E-05	2.98	4.57E-06	2.99
2	4	1.60E-06	2.99	1.26E-06	2.98	1.68E-06	2.99	5.73E-07	3.00
	5	2.00E-07	2.99	1.58E-07	2.99	2.11E-07	3.00	7.16E-08	3.00
	6	2.51E-08	3.00	1.98E-08	3.00	2.64E-08	3.00	8.95E-09	3.00
	3	6.78E-08	3.95	3.87E-08	3.96	1.49E-07	3.96	1.04E-07	3.97
3	4	4.29E-09	3.98	2.43E-09	3.99	9.46E-09	3.98	6.55E-09	3.99
	5	2.69E-10	3.99	1.53E-10	4.00	5.95E-10	3.99	4.11E-10	3.99
	6	1.68E-11	4.00	9.54E-12	4.00	3.73E-11	4.00	2.57E-11	4.00

Table 2.4: $\tau^+ \equiv 1,\ \tau^- \equiv 0,\ \alpha_\theta \equiv 0,\ \alpha_T \equiv 0.$

		$\|e_T\|_{L^2(\Omega)}$		$\|e_M\|_{L^2(\Omega)}$		$\|e_\theta\|_{L^2(\Omega)}$		$\|e_{\mathcal{W}}\|_{L^2(\Omega)}$	
k	mesh	error	order	error	order	error	order	error	order
				$d = 10^{-2}$					
	5	7.09E-02	0.86	4.95E-03	1.27	4.95E-04	0.96	1.17E-03	1.09
0	6	3.72E-02	0.93	1.98E-03	1.33	2.70E-04	0.88	5.75E-04	1.02
	7	1.91E-02	0.97	7.99E-04	1.31	1.43E-04	0.92	2.87E-04	1.00
	8	9.64E-03	0.98	3.37E-04	1.24	7.36E-05	0.95	1.44E-04	1.00
	3	1.94E-03	2.55	3.52E-03	1.78	2.33E-04	2.04	4.66E-04	2.10
1	4	3.84E-04	2.34	9.43E-04	1.90	5.74E-05	2.02	1.14E-04	2.03
	5	8.75E-05	2.13	2.44E-04	1.95	1.43E-05	2.00	2.83E-05	2.01
	6	2.13E-05	2.04	6.20E-05	1.98	3.58E-06	2.00	7.07E-06	2.00
	3	1.31E-05	3.00	3.42E-05	2.81	8.41E-06	2.91	1.59E-05	2.92
2	4	1.64E-06	3.00	4.57E-06	2.90	1.08E-06	2.96	2.03E-06	2.97
	5	2.04E-07	3.00	5.90E-07	2.95	1.37E-07	2.98	2.57E-07	2.98
	6	2.55E-08	3.00	7.50E-08	2.98	1.73E-08	2.99	3.23E-08	2.99
	3	9.87E-08	3.99	2.58E-07	3.81	1.24E-07	3.98	1.89E-07	4.04
3	4	6.16E-09	4.00	1.72E-08	3.90	7.81E-09	3.99	1.17E-08	4.02
	5	3.84E-10	4.00	1.11E-09	3.95	4.89E-10	4.00	7.24E-10	4.01
	6	2.40E-11	4.00	7.07E-11	3.98	3.06E-11	4.00	4.51E-11	4.00
				$d = 10^{-8}$					
	5	7.11E-02	0.86	4.95E-03	1.27	4.98E-04	0.95	1.16E-03	1.09
0	6	3.73E-02	0.93	1.97E-03	1.33	2.72E-04	0.87	5.74E-04	1.02
	7	1.91E-02	0.97	7.96E-04	1.31	1.44E-04	0.92	2.87E-04	1.00
	8	9.67E-03	0.98	3.36E-04	1.25	7.43E-05	0.95	1.43E-04	1.00
	3	1.94E-03	2.55	3.52E-03	1.78	2.33E-04	2.04	4.66E-04	2.09
1	4	3.85E-04	2.34	9.43E-04	1.90	5.73E-05	2.02	1.14E-04	2.03
	5	8.75E-05	2.14	2.44E-04	1.95	1.43E-05	2.00	2.83E-05	2.01
	6	2.13E-05	2.04	6.20E-05	1.98	3.57E-06	2.00	7.06E-06	2.00
	3	1.31E-05	3.00	3.42E-05	2.81	8.40E-06	2.91	1.59E-05	2.92
2	4	1.64E-06	3.00	4.57E-06	2.90	1.08E-06	2.96	2.03E-06	2.97
	5	2.04E-07	3.00	5.90E-07	2.95	1.37E-07	2.98	2.57E-07	2.98
	6	2.55E-08	3.00	7.50E-08	2.98	1.73E-08	2.99	3.23E-08	2.99
	3	9.87E-08	3.99	2.58E-07	3.81	1.24E-07	3.98	1.89E-07	4.04
3	4	6.16E-09	4.00	1.72E-08	3.90	7.80E-09	3.99	1.16E-08	4.02
	5	3.84E-10	4.00	1.11E-09	3.95	4.89E-10	4.00	7.23E-10	4.01
	6	2.40E-11	4.00	7.07E-11	3.98	3.06E-11	4.00	4.51E-11	4.00

Table 2.5: $\alpha_\theta \equiv 0,\ \alpha_T \equiv 0$.

k	$\|e_T\|_{L^2(\Omega)}$	$\|e_M\|_{L^2(\Omega)}$	$\|e_\theta\|_{L^2(\Omega)}$	$\|e_{\mathcal{W}}\|_{L^2(\Omega)}$
		$\tau \equiv h^2$		
0	$-$	$-$	$-$	$-$
1	k	$-$	k	$-$
2	$k+1$	$k-1$	k	$-$
3	$k+1$	$k-1$	k	$k-2$
		$\tau \equiv h$		
0	$-$	$-$	$-$	$-$
1	$k+1$	k	$k+1$	k
2	$k+1$	k	$k+1$	k
3	$k+1$	k	$k+1$	k
		$\tau \equiv 1$		
0	$k+1$	$k+1$	$k+1$	$k+1$
1	$k+1$	$k+1$	$k+1$	$k+1$
2	$k+1$	$k+1$	$k+1$	$k+1$
3	$k+1$	$k+1$	$k+1$	$k+1$
		$\tau \equiv 1/h$		
0	$-$	$-$	$-$	$-$
1	k	$k+1$	k	$k+1$
2	k	$k+1$	k	$k+1$
3	k	$k+1$	k	$k+1$
		$\tau \equiv 1/h^2$		
0	$-$	$-$	$-$	$-$
1	k	$k+1$	k	$k+1$
2	k	$k+1$	k	$k+1$
3	k	$k+1$	k	$k+1$

Table 2.6: $\tau \equiv 0$.

k	$\|e_T\|_{L^2(\Omega)}$	$\|e_M\|_{L^2(\Omega)}$	$\|e_\theta\|_{L^2(\Omega)}$	$\|e_\mathcal{W}\|_{L^2(\Omega)}$
		$\alpha_\theta \equiv \alpha_T \equiv h^2$		
0	$-$	$-$	$-$	$-$
1	$k+1$	k	$k+1$	$-$
2	$k+1$	k	$k+1$	$k-1$
3	$k+1$	k	$k+1$	$k-1$
		$\alpha_\theta \equiv \alpha_T \equiv h$		
0	$-$	$-$	$-$	$-$
1	$k+1$	k	$k+1$	k
2	$k+1$	k	$k+1$	k
3	$k+1$	k	$k+1$	k
		$\alpha_\theta \equiv \alpha_T \equiv 1$		
0	$k+1$	$k+1$	$k+1$	$k+1$
1	$k+1$	$k+1$	$k+1$	$k+1$
2	$k+1$	$k+1$	$k+1$	$k+1$
3	$k+1$	$k+1$	$k+1$	$k+1$
		$\alpha_\theta \equiv \alpha_T \equiv 1/h$		
0	$-$	$-$	$-$	$-$
1	k	$k+1$	k	$k+1$
2	k	$k+1$	k	$k+1$
3	k	$k+1$	k	$k+1$
		$\alpha_\theta \equiv \alpha_T \equiv 1/h^2$		
0	$-$	$-$	$-$	$-$
1	k	$k+1$	k	$k+1$
2	k	$k+1$	k	$k+1$
3	k	$k+1$	k	$k+1$

Table 2.7: $\tau^- \equiv 0,\ \alpha_\theta \equiv 0,\ \alpha_T \equiv 0.$

k	$\|e_T\|_{L^2(\Omega)}$	$\|e_M\|_{L^2(\Omega)}$	$\|e_\theta\|_{L^2(\Omega)}$	$\|e_{\mathcal{W}}\|_{L^2(\Omega)}$
		$\tau^+ \equiv h^2$		
0	$-$	$-$	$-$	$-$
1	k	$-$	k	$-$
2	$k+1$	$k-1$	k	$-$
3	$k+1$	$k-1$	k	$k-2$
		$\tau^+ \equiv h$		
0	$-$	$-$	$-$	$-$
1	$k+1$	k	$k+1$	k
2	$k+1$	k	$k+1$	k
3	$k+1$	k	$k+1$	k
		$\tau^+ \equiv 1$		
0	$k+1$	$k+1$	$k+1$	$k+1$
1	$k+1$	$k+1$	$k+1$	$k+1$
2	$k+1$	$k+1$	$k+1$	$k+1$
3	$k+1$	$k+1$	$k+1$	$k+1$
		$\tau^+ \equiv 1/h$		
0	$k+1$	$k+1$	$k+1$	$k+1$
1	$k+1$	$k+1$	$k+1$	$k+1$
2	$k+1$	$k+1$	$k+1$	$k+1$
3	$k+1$	$k+1$	$k+1$	$k+1$
		$\tau^+ \equiv 1/h^2$		
0	$k+1$	$k+1$	$k+1$	$k+1$
1	$k+1$	$k+1$	$k+1$	$k+1$
2	$k+1$	$k+1$	$k+1$	$k+1$
3	$k+1$	$k+1$	$k+1$	$k+1$

Table 2.8: $\tau \equiv \alpha_\theta \equiv \alpha_T \equiv \nu$.

k	$\|e_T\|_{L^2(\Omega)}$	$\|e_M\|_{L^2(\Omega)}$	$\|e_\theta\|_{L^2(\Omega)}$	$\|e_{\mathcal{W}}\|_{L^2(\Omega)}$
		$\nu = h^2$		
0	$-$	$-$	$-$	$-$
1	$k+1$	k	$k+1$	$-$
2	$k+1$	k	$k+1$	$k-1$
3	$k+1$	k	$k+1$	$k-1$
		$\nu = h$		
0	$-$	$-$	$-$	$-$
1	$k+1$	k	$k+1$	k
2	$k+1$	k	$k+1$	k
3	$k+1$	k	$k+1$	k
		$\nu = 1$		
0	$k+1$	$k+1$	$k+1$	$k+1$
1	$k+1$	$k+1$	$k+1$	$k+1$
2	$k+1$	$k+1$	$k+1$	$k+1$
3	$k+1$	$k+1$	$k+1$	$k+1$
		$\nu = 1/h$		
0	$-$	$-$	$-$	$-$
1	k	$k+1$	k	$k+1$
2	k	$k+1$	k	$k+1$
3	k	$k+1$	k	$k+1$
		$\nu = 1/h^2$		
0	$-$	$-$	$-$	$-$
1	k	$k+1$	k	$k+1$
2	k	$k+1$	k	$k+1$
3	k	$k+1$	k	$k+1$

Table 2.9: $\alpha_\theta \equiv 1,\ \alpha_T \equiv 0.$

k	$\|e_T\|_{L^2(\Omega)}$	$\|e_M\|_{L^2(\Omega)}$	$\|e_\theta\|_{L^2(\Omega)}$	$\|e_{\mathcal{W}}\|_{L^2(\Omega)}$
		$\tau = h^2$		
0	$-$	$-$	$-$	$-$
1	k	$-$	k	$-$
2	$k+1$	$k-1$	k	$-$
3	$k+1$	$k-1$	k	$-$
		$\tau = h$		
0	$-$	$-$	$-$	$-$
1	$k+1$	k	$k+1$	$-$
2	$k+1$	k	$k+1$	$k-1$
3	$k+1$	k	$k+1$	$k-1$
		$\tau = 1$		
0	$k+1$	$k+1$	$k+1$	$k+1$
1	$k+1$	$k+1$	$k+1$	$k+1$
2	$k+1$	$k+1$	$k+1$	$k+1$
3	$k+1$	$k+1$	$k+1$	$k+1$
		$\tau = 1/h$		
0	$-$	$-$	$-$	$-$
1	k	$k+1$	k	$k+1$
2	k	$k+1$	k	$k+1$
3	k	$k+1$	k	$k+1$
		$\tau = 1/h^2$		
0	$-$	$-$	$-$	$-$
1	k	$k+1$	k	$k+1$
2	k	$k+1$	k	$k+1$
3	k	$k+1$	k	$k+1$

Table 2.10: $\alpha_\theta \equiv 0,\ \alpha_T \equiv 1$.

k	$\|e_T\|_{L^2(\Omega)}$	$\|e_M\|_{L^2(\Omega)}$	$\|e_\theta\|_{L^2(\Omega)}$	$\|e_{\mathcal{W}}\|_{L^2(\Omega)}$
		$\tau = h^2$		
0	–	–	–	–
1	$k+1$	k	$k+1$	$k+1$
2	$k+1$	k	$k+1$	$k+1$
3	$k+1$	k	$k+1$	$k+1$
		$\tau = h$		
0	–	–	–	–
1	$k+1$	k	$k+1$	$k+1$
2	$k+1$	k	$k+1$	$k+1$
3	$k+1$	k	$k+1$	$k+1$
		$\tau = 1$		
0	$k+1$	$k+1$	$k+1$	$k+1$
1	$k+1$	$k+1$	$k+1$	$k+1$
2	$k+1$	$k+1$	$k+1$	$k+1$
3	$k+1$	$k+1$	$k+1$	$k+1$
		$\tau = 1/h$		
0	–	–	–	–
1	k	$k+1$	k	$k+1$
2	k	$k+1$	k	$k+1$
3	k	$k+1$	k	$k+1$
		$\tau = 1/h^2$		
0	–	–	–	–
1	k	$k+1$	k	$k+1$
2	k	$k+1$	k	$k+1$
3	k	$k+1$	k	$k+1$

Table 2.11: $\alpha_T \equiv 0,\ \tau \equiv 1$.

k	$\|e_T\|_{L^2(\Omega)}$	$\|e_M\|_{L^2(\Omega)}$	$\|e_\theta\|_{L^2(\Omega)}$	$\|e_{\mathcal{W}}\|_{L^2(\Omega)}$
		$\alpha_\theta = h^2$		
0	$k+1$	$k+1$	$k+1$	$k+1$
1	$k+1$	$k+1$	$k+1$	$k+1$
2	$k+1$	$k+1$	$k+1$	$k+1$
3	$k+1$	$k+1$	$k+1$	$k+1$
		$\alpha_\theta = h$		
0	$k+1$	$k+1$	$k+1$	$k+1$
1	$k+1$	$k+1$	$k+1$	$k+1$
2	$k+1$	$k+1$	$k+1$	$k+1$
3	$k+1$	$k+1$	$k+1$	$k+1$
		$\alpha_\theta = 1$		
0	$k+1$	$k+1$	$k+1$	$k+1$
1	$k+1$	$k+1$	$k+1$	$k+1$
2	$k+1$	$k+1$	$k+1$	$k+1$
3	$k+1$	$k+1$	$k+1$	$k+1$
		$\alpha_\theta = 1/h$		
0	$-$	$-$	$-$	$-$
1	$k+1$	$k+1$	k	k
2	$k+1$	$k+1$	k	k
3	$k+1$	$k+1$	k	k
		$\alpha_\theta = 1/h^2$		
0	$-$	$-$	$-$	$-$
1	$k+1$	$k+1$	$-$	$-$
2	$k+1$	$k+1$	$k-1$	$k-1$
3	$k+1$	$k+1$	$k-1$	$k-1$

Table 2.12: $\alpha_\theta \equiv 0,\ \tau \equiv 1$.

k	$\|e_T\|_{L^2(\Omega)}$	$\|e_M\|_{L^2(\Omega)}$	$\|e_\theta\|_{L^2(\Omega)}$	$\|e_{\mathcal{W}}\|_{L^2(\Omega)}$
		$\alpha_T = h^2$		
0	$k+1$	$k+1$	$k+1$	$k+1$
1	$k+1$	$k+1$	$k+1$	$k+1$
2	$k+1$	$k+1$	$k+1$	$k+1$
3	$k+1$	$k+1$	$k+1$	$k+1$
		$\alpha_T = h$		
0	$k+1$	$k+1$	$k+1$	$k+1$
1	$k+1$	$k+1$	$k+1$	$k+1$
2	$k+1$	$k+1$	$k+1$	$k+1$
3	$k+1$	$k+1$	$k+1$	$k+1$
		$\alpha_T = 1$		
0	$k+1$	$k+1$	$k+1$	$k+1$
1	$k+1$	$k+1$	$k+1$	$k+1$
2	$k+1$	$k+1$	$k+1$	$k+1$
3	$k+1$	$k+1$	$k+1$	$k+1$
		$\alpha_T = 1/h$		
0	$-$	$-$	$-$	$-$
1	k	k	$k+1$	$k+1$
2	k	k	$k+1$	$k+1$
3	k	k	$k+1$	$k+1$
		$\alpha_T = 1/h^2$		
0	$-$	$-$	$-$	$-$
1	$-$	k	k	k
2	$k-1$	k	$k+1$	$k+1$
3	$k-1$	k	$k+1$	$k+1$

Chapter 3

Superconvergent HDG methods for second order elliptic equations

3.1 Introduction

In this Chapter, we propose a projection-based a priori error analysis of finite element methods for second-order elliptic problems. The analysis is *unifying* because it applies to a large class of methods including the hybridized version of most well-known mixed methods as well as several HDG methods. For the sake of simplicity, we present our approach in the framework of the following diffusion problem:

$$\boldsymbol{q} + \nabla u = 0 \qquad \text{in } \Omega, \tag{3.1a}$$

$$\nabla \cdot \boldsymbol{q} = f \qquad \text{in } \Omega \tag{3.1b}$$

$$u = g \qquad \text{on } \partial\Omega. \tag{3.1c}$$

Here $\Omega \subset \mathbb{R}^n$ $(n = 2, 3)$ is a bounded polyhedral domain, $f \in L^2(\Omega)$ and $g \in H^{\frac{1}{2}}(\partial\Omega)$.

To better describe our results, let us begin by introducing the general form of the methods we are going to consider; we follow [6]. Let $\mathcal{T}_h := \{K\}$ denote a conforming triangulation of Ω, where K is a polyhedral element. We denote the set of faces F of an element $K \in \mathcal{T}_h$ by $\mathcal{F}(K)$, and the set of faces F of all elements $K \in \mathcal{T}_h$ by $\mathcal{E}_h$. The methods we are interested in seek an approximation to $(u, \boldsymbol{q}, u|_{\mathcal{E}_h})$, $(u_h, \boldsymbol{q}_h, \widehat{u}_h)$, in the finite element space $W_h \times \boldsymbol{V}_h \times M_h$, where

$$\begin{aligned}
\boldsymbol{V}_h &:= \{\boldsymbol{v} \in \boldsymbol{L}^2(\mathcal{T}_h) : \boldsymbol{v}|_K \in \boldsymbol{V}(K),\ \ K \in \mathcal{T}_h\},\\
W_h &:= \{w \in L^2(\mathcal{T}_h) : w|_K \in W(K), K \in \mathcal{T}_h\},\\
M_h &:= \{\mu \in L^2(\mathcal{E}_h) : \mu|_F \in W(F),\ \ F \in \mathcal{E}_h\},
\end{aligned}$$

and determine it as the only solution of the following weak formulation:

$$-(u_h\,,\, \nabla\cdot \boldsymbol{v})_{\mathcal{T}_h} + (\boldsymbol{q}_h\,,\, \boldsymbol{v})_{\mathcal{T}_h} + \langle \widehat{u}_h\,,\, \boldsymbol{v}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} = 0, \tag{3.2a}$$

$$-(\boldsymbol{q}_h\,,\, \nabla w)_{\mathcal{T}_h} + \langle \widehat{\boldsymbol{q}}_h\cdot\boldsymbol{n}\,,\, w\rangle_{\partial\mathcal{T}_h} = (f\,,\, w)_{\mathcal{T}_h}, \tag{3.2b}$$

$$\langle \widehat{\boldsymbol{q}}_h\cdot\boldsymbol{n}, \mu\rangle_{\partial\mathcal{T}_h\setminus\partial\Omega} = 0, \tag{3.2c}$$

$$\langle \widehat{u}_h, \mu\rangle_{\partial\Omega} = \langle g, \mu\rangle_{\partial\Omega}, \tag{3.2d}$$

for all $(w, \boldsymbol{v}, \mu) \in W_h \times \boldsymbol{V}_h \times M_h$. The definition of the method is completed with the definition of the normal component of the numerical trace:

$$\widehat{\boldsymbol{q}}_h\cdot\boldsymbol{n} = \boldsymbol{q}_h\cdot\boldsymbol{n} + \alpha(u_h - \widehat{u}_h) \quad \text{on} \quad \partial\mathcal{T}_h. \tag{3.3}$$

By taking particular choices of the local spaces $\boldsymbol{V}(K)$, $W(K)$ and $M(F)$, and the *linear local stabilization* operator α, the different mixed ($\alpha = 0$) and HDG ($\alpha \neq 0$) methods are obtained.

Our main result is to show that if we can construct, in an element-by-element fashion, an auxiliary projection $\Pi_u(\boldsymbol{q}, u) := (\boldsymbol{\Pi}_V\boldsymbol{q}, \Pi_W u)$ satisfying certain orthogonality and approximation conditions, and the local spaces $\boldsymbol{V}(K)$, $W(K)$ and $M(F)$, for all the faces F of the element K, satisfy some inclusion properties, then the method is well defined and we have the estimates

$$\begin{aligned}
\|\boldsymbol{q} - \boldsymbol{q}_h\|_{\mathcal{T}_h} &\le 2\,\|\boldsymbol{q} - \boldsymbol{\Pi}_V\boldsymbol{q}\|_{\mathcal{T}_h},\\
\|\Pi_W u - u_h\|_{\mathcal{T}_h} &\le C\,h\|\boldsymbol{q} - \boldsymbol{\Pi}_V\boldsymbol{q}\|_{\mathcal{T}_h},
\end{aligned}$$

where $\|\cdot\|_{\mathcal{T}_h}$ denotes the $L^2(\mathcal{T}_h)$-norm.

Note that if the error $\Pi_W u - u_h$ converges to zero *faster* than the error $u - u_h$, this *superconvergence* property can be advantageously exploited; see [14, 15, 17, 18, 19]. Indeed, following [17, 18, 19], we define a new approximation to u, u_h^*, in the space

$$W_h^* := \{w \in L^2(\mathcal{T}_h) : w|_K \in W^*(K), K \in \mathcal{T}_h\}, \tag{3.4a}$$

as follows. On each element $K \in \mathcal{T}_h$, the function u_h^* is the element of $W^*(K)$ such that

$$
\begin{aligned}
(\nabla u_h^*, \nabla \omega)_K &= -\,(\boldsymbol{q}_h, \nabla \omega)_K \qquad \forall\, \omega \in W^*(K):\ (\omega, 1)_K = 0, && \text{(3.4b)}\\
(u_h^*, 1)_K &= (u_h, 1)_K. && \text{(3.4c)}
\end{aligned}
$$

It is not difficult to prove that u_h^* is well defined and that we have

$$
\|u - u_h^*\|_{\mathcal{T}_h} \leq \|\Pi_W u - u_h\|_{\mathcal{T}_h} + C\, h\, (\|\boldsymbol{q} - \boldsymbol{q}_h\|_{\mathcal{T}_h} + \inf_{\omega \in W_h^*} \|\nabla(u - \omega)\|_{\mathcal{T}_h}),
$$

which means that it is possible to define u_h^* converging to u as fast as $\Pi_W u - u_h$ converges to zero.

Moreover, we do provide a *single* template for the choice of the local spaces $\boldsymbol{V}(K)$, $W(K)$ and $M(F)$, and for the stabilization operator α which *guarantees* the existence of the auxiliary projection Π_h with the above-mentioned properties. Using this template, we give many examples of superconvergent methods, old an new, fitting our general framework. The old ones are the (hybridized versions of) the main mixed methods, the LDG-H and BMMPR-H methods for simplexes; see [6]. The last method, which had never been analyzed before, is proven to superconverge even though it uses a local stabilization operator α which is different from those of the previous examples.

The new methods are several LDG-H methods for squares, cubes and prisms. The definition of these methods had remained elusive in the last few years and it is thanks to the our projection-based approach that it became clear. It is important to emphasize that, although it is very easy to devise HDG methods that are well defined, it is far from obvious to devise them so that they display the above-mentioned superconvergence property. The technique we propose here is a *new and effective* tool to achieve this goal.

Yet another new method is what seems to be the *smallest* superconvergent mixed method, on squares and cubes, whose local space $\boldsymbol{V}(K) \times W(K)$ contains the tensor-product space $\boldsymbol{Q}^k(K) \times Q^k(K)$. In [41], such a method with $\boldsymbol{V}(K) \times W(K)$ exactly equal to $\boldsymbol{Q}^k(K) \times Q^k(K)$ (in two dimensions) was experimentally shown to provide an approximate flux $\boldsymbol{q}_h$ converging with order k only; moreover, the function u_h^* was shown to converge with order $k+1$ only. Here, we prove that, by only adding a *fixed* number (3 in two-space dimensions and 7 in three-space dimensions) of extra basis functions to $\boldsymbol{Q}^k(K)$, we recover the orders of convergence of $k+1$ for the approximate flux $\boldsymbol{q}_h$ and of $k+2$ for postprocessing u_h^*, provided $k \geq 1$. This new method has convergence properties similar to those of the corresponding RT method, see [8], but uses *significantly smaller* spaces.

3.2 Main results

In this section we show how an a priori error analysis of the HDG methods can be *reduced* to the verification of a few conditions on the local spaces and on some properties of an associated, auxiliary projection Π_h defined in an element-by-element fashion.

3.2.1 A priori error estimates

The main idea of our error analysis is to estimate the projection of the errors $\Pi_h(\boldsymbol{q} - \boldsymbol{q}_h, u - u_h)$ and then deduce bounds of the $L^2(\Omega)$-norm of the errors $\boldsymbol{q} - \boldsymbol{q}_h$, $u - u_h$ and $u - u_h^*$.

Estimate of $\boldsymbol{q} - \boldsymbol{q}_h$

Our first result gives an estimate of the projection of the error $\boldsymbol{\Pi}_V \boldsymbol{q} - \boldsymbol{q}_h$ solely in terms of the approximation error of the projection $\boldsymbol{q} - \boldsymbol{\Pi}_V \boldsymbol{q}$. To state it, we need to describe our assumptions on the projection Π_h and on the local finite element spaces $\boldsymbol{V}(K)$, $W(K)$, and $M(F)$.

Assumptions A:

- *Orthogonality properties of* Π_h. On each element K, there exist a projection $\Pi_h(\boldsymbol{q}, u) = (\boldsymbol{\Pi}_V \boldsymbol{q}, \Pi_W u) \in \boldsymbol{V}(K) \times W(K)$ satisfying the following properties:

(A.1) $(\boldsymbol{\Pi}_V \boldsymbol{q}, \boldsymbol{v})_K = (\boldsymbol{q}, \boldsymbol{v})_K$ for all $\boldsymbol{v} \in \nabla W(K)$,

(A.2) $(\Pi_W u, w)_K = (u, w)_K$ for all $w \in \nabla \cdot \boldsymbol{V}(K)$,

(A.3) For all faces F of the element K,

$$\langle \boldsymbol{\Pi}_V \boldsymbol{q} \cdot \boldsymbol{n} + \alpha(\Pi_W u), \mu \rangle_F = \langle \boldsymbol{q} \cdot \boldsymbol{n} + \alpha(P_M u), \mu \rangle_F \quad \text{for all } \mu \in M(F).$$

We also need to assume suitable relations between the traces on the faces F of the local spaces $\boldsymbol{V}(K)$ and $W(K)$ with the local space $M(F)$.

• *Properties of the traces of the local spaces.* For each element K, and for any of its faces F,

(A.4) $\boldsymbol{V}(K) \cdot \boldsymbol{n}|_F \subset M(F)$,

(A.5) $W(K)|_F \subset M(F)$.

Here, $\boldsymbol{V}(K) \cdot \boldsymbol{n}|_F$ denotes the space of the *traces* of normal components of functions of $\boldsymbol{V}(K)$ on the face F of K. Similarly, $W(K)|_F$ denotes the space of *traces* of functions of $W(K)$ on the face F.

Finally, we need a simple assumption reflecting the stabilizing role of the linear operator α.

• *The semi-positivity property of α.* For each element K and any of its faces F,

(A.6) $\langle \alpha(\mu), \mu \rangle_F \geq 0$ for all $\mu \in M(F)$.

We are now ready to state our first result. In what follows, we use $\|\cdot\|_{k,D}$, $|\cdot|_{k,D}$ to denote the standard norm and seminorm on any Sobolev space $H^k(D)$, respectively. For simplicity, we use $\|\cdot\|_D$ to denote the $L^2(D)-$norm on any D.

Theorem 3.2.1. *Suppose that the Assumptions A are satisfied. Then we have*

$$\|\mathbf{\Pi}_V \boldsymbol{q} - \boldsymbol{q}_h\|_{\mathcal{T}_h} \leq \|\boldsymbol{q} - \mathbf{\Pi}_V \boldsymbol{q}\|_{\mathcal{T}_h}.$$

Note that, since this implies that

$$\|\boldsymbol{q} - \boldsymbol{q}_h\|_{\mathcal{T}_h} \leq 2\|\boldsymbol{q} - \mathbf{\Pi}_V \boldsymbol{q}\|_{\mathcal{T}_h},$$

the quality of the approximation $\boldsymbol{q}_h$ depends on the approximation properties of the first component of the projection Π_h *only.*

Estimate of $u - u_h$

Our next result shows that $\Pi_W u - u_h$ can *also* be controlled solely in terms of the approximation error of the projection $\boldsymbol{q} - \boldsymbol{\Pi}_V \boldsymbol{q}$.

It is valid under a typical elliptic regularity property we state next. We assume that, for any given $\eta \in L^2(\Omega)$, we have

$$\|\phi\|_{2,\Omega} + \|\boldsymbol{\theta}\|_{1,\Omega} \leq C\|\eta\|_{\Omega}, \tag{3.5}$$

where C only depends on the domain Ω, and $(\boldsymbol{\theta}, \phi)$ is the solution of the *dual* problem:

$$\boldsymbol{\theta} + \nabla\phi = 0 \quad \text{in } \Omega, \tag{3.6a}$$

$$\nabla \cdot \boldsymbol{\theta} = \eta \quad \text{in } \Omega, \tag{3.6b}$$

$$\phi = 0 \quad \text{on } \partial\Omega. \tag{3.6c}$$

We also need a couple of additional assumptions.

Assumptions B:

The first is an approximation property of a projection $\Pi_h^*(\boldsymbol{q}, u) = (\boldsymbol{\Pi}_V^* \boldsymbol{q}, \Pi_W^* u)$ which satisfies the assumptions *(A.1)*, *(A.2)*, and *(A.3)* where the local stabilization operator $\alpha(\cdot)$ is *replaced* by its dual $\alpha^*(\cdot)$, that is, by the linear function defined by

$$\langle \eta\,, \alpha^*(\mu)\rangle_F = \langle \alpha(\eta)\,, \mu\rangle_F \quad \text{for all} \quad \eta, \mu \in M(F) \;\; \forall\, F \in \mathcal{F}(K).$$

• *The approximation property of the projection* Π_h^*. For each element K and any $(\boldsymbol{q}, u) \in \boldsymbol{H}^1(K) \times H^2(K)$,

(B.1) $\|\boldsymbol{\Pi}_V^* \boldsymbol{q} - \boldsymbol{q}\|_K \leq C_{app}^* \, h_K(|u|_{1,K} + |\boldsymbol{q}|_{1,K})$.

The second assumption is a condition on the local space $W(K)$.

• *The local space $W(K)$ is not too small.* For each element K, we have that

(B.2) $\boldsymbol{P}^0(K) \subset \nabla W(K)$.

Here $\boldsymbol{P}^0(K) := [P^0(K)]^n$ and $P^0(K)$ is the space of constants defined on K.

We are now ready to state our second result.

Theorem 3.2.2. *Suppose that the Assumptions A and B are satisfied. Also, suppose that the elliptic regularity property* (3.5) *holds. Then we have*

$$\|\Pi_W u - u_h\|_{\mathcal{T}_h} \le C\, h \|\boldsymbol{q} - \boldsymbol{\Pi}_V \boldsymbol{q}\|_{\mathcal{T}_h},$$

for some constant C depending on C^*_{app} but independent of h and the exact solution.

From this result, we immediately get that

$$\|u - u_h\|_{\mathcal{T}_h} \le \|u - \Pi_W u\|_{\mathcal{T}_h} + C\, h\, \|\boldsymbol{q} - \boldsymbol{\Pi}_V \boldsymbol{q}\|_{\mathcal{T}_h},$$

and we see that the quality of the approximation u_h only depends on the approximation error of the projection.

Estimate of $u - u_h^*$

Note that if the second term of the above right-hand side converges faster than the first, the convergence of u_h to $\Pi_W u$ is *faster* than that of u_h to u. As mentioned before, we can take advantage of this *superconverge* result to show that the postprocessing u_h^* defined by (3.4) converges to u as fast as u_h superconverges to $\Pi_W u$. For that purpose, we need the following assumption.

Assumption C:

- *The local space $\boldsymbol{V}(K)$ is not too small.* For each element K,

(C.1) $P^0(K) \subset \nabla \cdot \boldsymbol{V}(K)$.

We can now state our third and last result.

Theorem 3.2.3. *Suppose that the Assumptions A, B, and C are satisfied. Then, we have*

$$\|u - u_h^*\|_{\mathcal{T}_h} \leq \|\Pi_W u - u_h\|_{\mathcal{T}_h} + C\,h\,(\|\boldsymbol{q} - \boldsymbol{\Pi}_V \boldsymbol{q}\|_{\mathcal{T}_h} + \inf_{\omega \in W_h^*} \|\nabla(u - \omega)\|_{\mathcal{T}_h}).$$

3.3 A template for the construction of superconvergent methods

In this section, we provide a particular way of constructing HDG methods satisfying *Assumptions A, B* and *C*. It turns out that *all* the already known methods in the literature (which we also gather in this section) can be constructed by using this template. Moreover, all the *new* superconvergent methods we introduce here were obtained by using it.

3.3.1 The choice of the local spaces and the stabilization operator

To construct our superconvergent methods, we pick an arbitrary element $K \in \mathcal{T}_h$, and proceed as follows:

Step 1: The local space $\boldsymbol{V}(K) \times W(K)$ We begin by taking a local space $\boldsymbol{V}(K) \times W(K)$ such that

$$\boldsymbol{P}^0(K) \subset \nabla W(K) \quad \subset \boldsymbol{V}(K), \tag{3.7a}$$

$$P^0(K) \subset \nabla \cdot \boldsymbol{V}(K) \subset W(K). \tag{3.7b}$$

Step 2: The local space $M(F)$ Then, for each face F of the element K, we find a space $M(F)$ such that

$$\boldsymbol{V}(K) \cdot \boldsymbol{n}|_F \subset M(F), \tag{3.8a}$$

$$W(K)|_F \subset M(F). \tag{3.8b}$$

This choice has to be made so that

$$\sum_{F \in \mathcal{F}(K)} \dim M(F) \leq (\dim \boldsymbol{V}(K) - \dim \nabla W(K)) + (\dim W(K) - \nabla \cdot \boldsymbol{V}(K)). \tag{3.9}$$

Step 3: The auxiliary local space $\widetilde{\boldsymbol{V}}(K) \times \widetilde{W}(K)$ Next, we find an *auxiliary* space $\widetilde{\boldsymbol{V}}(K) \times \widetilde{W}(K)$ satisfying

$$\nabla W(K) \subset \widetilde{\boldsymbol{V}}(K) \subset \boldsymbol{V}(K), \tag{3.10a}$$

$$\nabla \cdot \boldsymbol{V}(K) \subset \widetilde{W}(K) \subset W(K), \tag{3.10b}$$

such that, if we set

$$\boldsymbol{V}^{\perp}(K) := \{\boldsymbol{v} \in \boldsymbol{V}(K) : (\boldsymbol{v}, \widetilde{\boldsymbol{v}})_K = 0 \quad \forall\, \widetilde{\boldsymbol{v}} \in \widetilde{\boldsymbol{V}}(K)\}, \tag{3.11a}$$

$$W^{\perp}(K) := \{w \in W(K) : (w, \widetilde{w})_K = 0 \quad \forall\, \widetilde{w} \in \widetilde{W}(K)\}, \tag{3.11b}$$

we have that

$$\sum_{F \in \mathcal{F}(K)} \dim M(F) = \dim \boldsymbol{V}^{\perp}(K) + \dim W^{\perp}(K), \tag{3.12}$$

and that

$$\|\boldsymbol{v}^{\perp}\|_K \le C_V\, h_K^{1/2} \|\, \boldsymbol{v}^{\perp} \cdot \boldsymbol{n}\, \|_{\partial K_V} \qquad \text{for all } \boldsymbol{v}^{\perp} \in \boldsymbol{V}^{\perp}(K), \tag{3.13a}$$

$$\|w^{\perp}\|_K \le C_W\, h_K^{1/2} \|w^{\perp}\|_{\partial K_W} \qquad \text{for all } w^{\perp} \in W^{\perp}(K), \tag{3.13b}$$

for some subsets ∂K_V and ∂K_W of $\mathcal{F}(K)$.

Step 4: The stabilization operator α Finally, we pick the local stabilization operator α such that

$$\langle \alpha(\eta),\, \mu \rangle_F = \langle \eta, \alpha(\mu) \rangle_F \qquad \text{for all } \eta, \mu \in M(F), \tag{3.14a}$$

$$\langle \alpha(w^{\perp}), w^{\perp} \rangle_{\partial K} \ge C_\alpha\, \|w^{\perp}\|^2_{\partial K_W} \text{ for all } w^{\perp} \in W^{\perp}(K). \tag{3.14b}$$

Let us briefly discuss the most difficult points of the above steps. The first concerns the inequality (3.9) which states, roughly speaking, that the kernel of the divergence operator in $\boldsymbol{V}(K)$ has to be *big* enough. Indeed, if we assume that $W(K) \supset P^0(K)$, such inequality can be rewritten as

$$\sum_{F \in \mathcal{F}(K)} \dim M(F) \le \dim\{\boldsymbol{v} \in \boldsymbol{V}(K) : \nabla \cdot \boldsymbol{v} = 0\} + 1.$$

As a consequence, if this condition is not satisfied, we can simply *add* to the basis of the space $\boldsymbol{V}(K)$ divergence-free functions whose normal component on each the faces F of

K lies on $M(F)$. Note that the more faces an element K has, the harder is to satisfy this inequality, and so, the more of such basis functions will have to be added.

The second concerns the construction of the auxiliary space $\widetilde{\boldsymbol{V}}(K) \times \widetilde{W}(K)$. Once the inequality (3.9) is satisfied, the existence of at least one auxiliary space satisfying the inclusion properties (3.10) and the equality (3.12) is *guaranteed*, assuming that the inclusion properties in Step 1 and the inequality in Step 2 are satisfied. However, the inequalities (3.13) *still* need to be satisfied. This means that the above-mentioned additional functions have to be controlled by their normal components.

3.3.2 Verification of the *Assumptions A, B,* and *C*

We claim that the HDG method determined by the above local spaces and stabilization operator does satisfy *Assumptions A, B,* and *C*. Let us show that this is indeed the case.

It is easy to see that *Assumption* (*A.4*) is nothing but condition (3.8a), that *Assumption* (*A.5*) is nothing but condition (3.8b), that *Assumption* (*A.6*) follows from condition (3.14b), that *Assumption* (*B.2*) is nothing but the first inclusion in condition (3.7a), and that *Assumption* (*C.1*) is nothing but the first inclusion in condition (3.7b).

To verify the remaining *Assumptions*, we must introduce an auxiliary projection Π_h. We define, for any element of $\boldsymbol{H}^1(K) \times H^1(K)$, $(\boldsymbol{q}, u)$, the projection $\Pi_h(\boldsymbol{q}, u) := (\boldsymbol{\Pi}_V \boldsymbol{q}, \Pi_W u)$ as the element of $\boldsymbol{V}(K) \times W(K)$ satisfying the equations

$$(\boldsymbol{\Pi}_V \boldsymbol{q}, \widetilde{\boldsymbol{v}})_K = (\boldsymbol{q}, \widetilde{\boldsymbol{v}})_K \qquad \forall\, \widetilde{\boldsymbol{v}} \in \widetilde{\boldsymbol{V}}(K), \tag{3.15a}$$

$$(\Pi_W u, \widetilde{w})_K = (u, \widetilde{w})_K \qquad \forall\, \widetilde{w} \in \widetilde{W}(K), \tag{3.15b}$$

$$\langle \boldsymbol{\Pi}_V \boldsymbol{q} \cdot \boldsymbol{n} + \alpha(\Pi_W u), \mu \rangle_F = \langle \boldsymbol{q} \cdot \boldsymbol{n} + \alpha(P_M u), \mu \rangle_F \quad \forall\, \mu \in M(F), \tag{3.15c}$$

for all faces F of the element K.

If this projection were well defined, *Assumption* (*A.1*) would follow from the first equation defining the projection, (3.15a), and from the first inclusion in condition (3.10a); *Assumption* (*A.2*) would follow from the second equation defining the projection, (3.15b), and from the first inclusion in condition (3.10b); and *Assumption* (*A.3*) from the third equation defining the projection, (3.15c). Thus, it only remains to prove that the projection is well defined and that it satisfies *Assumption* (*B.1*).

Note that since we are assuming that the stabilization operator α is self-adjoint, see condition (3.14a), we have that $\Pi_h^* = \Pi_h$. Note also that, by condition (3.12), the system of equations defining the projection Π_h is square. Hence, it is well defined if and only if, when $(\boldsymbol{q}, u) = (\boldsymbol{0}, 0)$, we have that $\Pi_h(\boldsymbol{q}, u) = (\boldsymbol{0}, 0)$. As a consequence, both the existence of the projection Π_h as well as *Assumption*(*B.1*) follow from the approximation result we state next.

To do it, we need to introduce some notation. We denote by $(\boldsymbol{P}_V, P_W, P_{\widetilde{W}})$ the L^2-projection into the local space $\boldsymbol{V}(K) \times W(K) \times \widetilde{W}(K)$. For any face F of the element K, we set

$$\|\alpha\|_F := \sup_{\mu \in M(F)\setminus\{0\}} \|\alpha(\mu)\|_F / \|\mu\|_F,$$

and define $\|\alpha\|_D := \max_{F\in D} \|\alpha\|_F$ where D is any union of faces of K. Finally, we set, for $W^\perp(K) \neq \{0\}$, $R_{W^\perp} := \sup_{w\in W^\perp(K)\setminus\{0\}} h_K^{1/2} \|w\|_{\partial K} / \|w\|_K$.

We are now ready to state our result.

Theorem 3.3.1. *We have*

$$\begin{aligned}
\|\boldsymbol{q} - \boldsymbol{\Pi}_V \boldsymbol{q}\|_K \leq{}& \|\boldsymbol{q} - \boldsymbol{P}_V\, \boldsymbol{q}\|_K + \mathsf{C}_1\, h_K^{1/2}\, \|(\boldsymbol{q} - \boldsymbol{P}_V \boldsymbol{q}) \cdot \boldsymbol{n}\|_{\partial K_V} \\
&+ \mathsf{C}_2\, h_K\, \|\nabla \cdot \boldsymbol{q} - \boldsymbol{P}_{\widetilde{W}} \nabla \cdot \boldsymbol{q}\|_K + \mathsf{C}_3\, h_K^{1/2}\, \|u - P_W u\|_{\partial K_W}, \\
\|u - \Pi_W u\|_K \leq{}& \|u - P_W u\|_K + \mathsf{C}_4\, h_K^{1/2}\, \|u - P_W u\|_{\partial K} \\
&+ \mathsf{C}_5\, h_K\, \|\nabla \cdot \boldsymbol{q} - \boldsymbol{P}_{\widetilde{W}} \nabla \cdot \boldsymbol{q}\|_K,
\end{aligned}$$

where $\mathsf{C}_1 := C_V, \mathsf{C}_2 := 0, \mathsf{C}_3 := \mathsf{C}_V \|\alpha\|_{\partial K_V}, \mathsf{C}_4 := 0,$ *and* $\mathsf{C}_5 := 0$ *whenever* $\widetilde{W}(K) = W(K)$. *Otherwise,*

$$\begin{aligned}
&\mathsf{C}_1 := C_V, \quad \mathsf{C}_2 := C_V\, C_W\, R_{W^\perp}\, \|\alpha\|_{\partial K_V} / C_\alpha, \quad \mathsf{C}_5 := C_W^2 / C_\alpha, \\
&\mathsf{C}_4 := \mathsf{C}_5\, R_{W^\perp}\, \|\alpha\|_{\partial K}, \quad \mathsf{C}_3 := \mathsf{C}_V \|\alpha\|_{\partial K_V} (1 + R_{W^\perp}^2\, C_W\, \|\alpha\|_{\partial K} / C_\alpha).
\end{aligned}$$

This result contains the information of how the choice of local spaces and stabilization operator affects the approximation properties of the projection. It indicates how to choose them to obtain optimal orders of convergence. Let us focus our discussion on the estimate of $\|\boldsymbol{q} - \boldsymbol{\Pi}_V \boldsymbol{q}\|_K$ as it is the only relevant one for the convergence properties described in the theorems of Section 2.

Note that if $\widetilde{W}(K)$ coincides with $W(K)$, then $\mathsf{C}_2 = \mathsf{C}_3 = 0$. When $\widetilde{W}(K)$ is a strict subdomain of $W(K)$, this is still true if we have that $\partial K_V \cap \partial K_W = \emptyset$ and if we take

α in such a way that $\|\alpha\|_{\partial K_V} = 0$. In these two cases, the approximation properties of $\mathbf{\Pi}_V$ are, roughly speaking, those of the L^2-projection $\boldsymbol{P}_V$.

In the general case, it is enough to take the stabilization operator α such that $\|\alpha\|_{\partial K_V}$ and $\|\alpha\|_{\partial K}/C_\alpha$ are uniformly bounded to ensure that the constants $\mathsf{C}_1, \mathsf{C}_2$ and C_3 are independent of α. In this case, we see that $\boldsymbol{Id}-\mathbf{\Pi}_V$ converges to zero, roughly speaking, as fast as the projections $\boldsymbol{Id} - \boldsymbol{P}_V$, $h_K\,(Id - P_{\widetilde{W}})$ and $Id - P_W$ do. In particular, by the inclusion conditions (3.7), we readily get that

$$\|\boldsymbol{q} - \mathbf{\Pi}_V \boldsymbol{q}\,\|_K \leq (1 + C\,\mathsf{C}_1 + 2\,\mathsf{C}_2)\,h_K\,|\,\boldsymbol{q}\,|_{1,K} + C\,\mathsf{C}_3\,h_K\,|\,u\,|_{1,K},$$

and *Assumption* (*B.1*) is verified, as claimed. Here, the constant C depends on the shape regularity of the element K and the dimension of the local space $\boldsymbol{V}(K) \times W(K)$.

3.3.3 Proof of the approximation properties of Π_h, Theorem 3.3.1

To prove the estimates of Theorem 3.3.1, we follow [11]. The idea is to estimate the quantities $\boldsymbol{\delta_q} := \mathbf{\Pi}_V \boldsymbol{q} - \boldsymbol{P}_V \boldsymbol{q}$ and $\delta_u := \Pi_W u - P_W u$, and then use the triangle inequality to obtain the desired estimates. We proceed in three steps.

Step 1: The equations for $\boldsymbol{\delta_q}$ and δ_u

By the equations defining the projection Π_h, (3.15), we have that

$$(\boldsymbol{\delta_q}, \widetilde{\boldsymbol{v}})_K = 0 \qquad \forall\, \widetilde{\boldsymbol{v}} \in \widetilde{\boldsymbol{V}}(K), \tag{3.16a}$$

$$(\delta_u, \widetilde{w})_K = 0 \qquad \forall\, \widetilde{w} \in \widetilde{W}(K), \tag{3.16b}$$

$$\langle \boldsymbol{\delta_q} \cdot \boldsymbol{n} + \alpha(\delta_u), \mu \rangle_F = \langle \boldsymbol{I_q} \cdot \boldsymbol{n} + \alpha(I_u), \mu \rangle_F \quad \forall\, \mu \in M(F), \tag{3.16c}$$

for all faces F of the element K. Here $\boldsymbol{I_q} := \boldsymbol{q} - \boldsymbol{P}_V \boldsymbol{q}$ and $I_u := P_M u - P_W u$.

Step 2: The estimate of δ_u

Next, we obtain an estimate of δ_u. By the definition of $W^\perp(K)$, (3.11b), we see that $\delta_u \in W^\perp(K)$, by the equation (3.16b). If $W^\perp(K) = \{0\}$, then

$$\|\delta_u\|_K = 0.$$

If $W^\perp(K) \neq \{0\}$, we claim that δ_u is the element of $W^\perp(K)$ satisfying

$$\langle \alpha(\delta_u), w\rangle_{\partial K} = ((\boldsymbol{Id} - \boldsymbol{P}_{\widetilde{W}})\nabla\cdot \boldsymbol{q}, w)_K + \langle \alpha(I_u), w\rangle_{\partial K} \quad \forall\, w \in W^\perp(K).$$

Taking $w := \delta_u$ and applying the Cauchy-Schwarz inequality, we get

$$\begin{aligned}\langle \alpha(\delta_u), \delta_u\rangle_{\partial K} &\leq \|(\boldsymbol{Id} - \boldsymbol{P}_{\widetilde{W}})\nabla\cdot \boldsymbol{q}\|_K \|\delta_u\|_K + \|\alpha(I_u)\|_{\partial K}\|\delta_u\|_{\partial K}\\ &\leq C_W\big(h_K^{1/2}\,\|(\boldsymbol{Id} - \boldsymbol{P}_{\widetilde{W}})\nabla\cdot \boldsymbol{q}\|_K + R_{W^\perp}\,\|\alpha\|_{\partial K}\,\|I_u\|_{\partial K}\big)\|\delta_u\|_{\partial K_W},\end{aligned}$$

by the condition (3.13b), and, by the condition (3.14b) on the stabilization operator α,

$$\|\delta_u\|_{\partial K_W} \leq \frac{C_W}{C_\alpha}\,\Big(h_K^{1/2}\,\|(\boldsymbol{Id} - \boldsymbol{P}_{\widetilde{W}})\nabla\cdot \boldsymbol{q}\|_K + R_{W^\perp}\,\|\alpha\|_{\partial K}\,\|I_u\|_{\partial K}\Big).$$

Finally, using once again condition (3.13b), we get that

$$\|\delta_u\|_K \leq \frac{C_W^2}{C_\alpha}\,\big(h_K\,\|(\boldsymbol{Id} - \boldsymbol{P}_{\widetilde{W}})\nabla\cdot \boldsymbol{q}\|_K + h_K^{1/2}\,R_{W^\perp}\,\|\alpha\|_{\partial K}\,\|I_u\|_{\partial K}\big).$$

The estimate of $\|\Pi_W u - u\|_K$ now follows by using the triangle inequality and by noting that $\|I_u\|_F \leq \|u - P_W u\|_F$ for any face F of K since, by the inclusion property (3.8b), we have that $I_u = P_M u - P_W u = P_M(u - P_W u)$.

It remains to prove the claim. By equation (3.16c), we have that

$$\langle \boldsymbol{\delta_q}\cdot \boldsymbol{n} + \alpha(\delta_u), w\rangle_{\partial K} = \langle \boldsymbol{I_q}\cdot\boldsymbol{n} + \alpha(I_u), w\rangle_{\partial K} \quad \forall\, w\in W^\perp(K),$$

because $w|_F \in M(F)$ by the inclusion condition (3.8b).

But

$$\langle \boldsymbol{\delta_q}\cdot\boldsymbol{n}, w\rangle_{\partial K} = (\nabla\cdot\boldsymbol{\delta_q}, w)_K + (\boldsymbol{\delta_q}, \nabla w)_K = 0.$$

Indeed, we have that $(\nabla\cdot\boldsymbol{\delta_q}, w)_K = 0$ by the first inclusion in condition (3.10b) and the fact that $w \in W^\perp(K)$. We also have that $(\boldsymbol{\delta_q}, \nabla w)_K = 0$ by equation (3.16a) and the first inclusion in condition (3.10a) .

Similarly,

$$\langle \boldsymbol{I_q}\cdot\boldsymbol{n}, w\rangle_{\partial K} = (\nabla\cdot\boldsymbol{I_q}, w)_K + (\boldsymbol{I_q}, \nabla w)_K = ((\boldsymbol{Id} - \boldsymbol{P}_{\widetilde{W}})\nabla\cdot\boldsymbol{q}, w)_K.$$

Indeed, $(\nabla\cdot\boldsymbol{I_q}, w)_K = (\nabla\cdot\boldsymbol{q}, w)_K = ((\boldsymbol{Id} - \boldsymbol{P}_{\widetilde{W}})\nabla\cdot\boldsymbol{q}, w)_K$, by the first inclusion in condition (3.10b) and the fact that $w\in W^\perp(K)$. Moreover, $(\boldsymbol{I_q}, \nabla w)_K = 0$ by the first inclusion in condition (3.10a) and the definition of $\boldsymbol{I_q}$. This proves the claim.

Step 3: The estimate of $\boldsymbol{\delta_q}$

Finally, let us estimate of $\boldsymbol{\delta_q}$. By the definition of $\boldsymbol{V}^{\perp}(K)$, (3.11a), we see that $\boldsymbol{\delta_q} \in \boldsymbol{V}^{\perp}(K)$, by the equation (3.16a). By the condition (3.13a), this implies that

$$\|\boldsymbol{\delta_q}\|_K \leq C_V\, h_K^{1/2}\, \|\boldsymbol{\delta_q} \cdot \boldsymbol{n}\|_{\partial K_V},$$

and by equation (3.16c), that

$$\|\boldsymbol{\delta_q}\|_K \leq C_V\, h_K^{1/2}\, (\|\boldsymbol{I_q} \cdot \boldsymbol{n}\|_{\partial K_V} + \|\alpha(\delta_u)\|_{\partial K_V} + \|\alpha(I_u)\|_{\partial K_V})\,.$$

If $\widetilde{W}(K) = W(K)$, $\delta_u = 0$, and we get that

$$\|\boldsymbol{\delta_q}\|_K \leq C_V\, h_K^{1/2}\, (\|\boldsymbol{I_q} \cdot \boldsymbol{n}\|_{\partial K_V} + \|\alpha\|_{\partial K_V}\, \|I_u\|_{\partial K_V})\,.$$

If $\widetilde{W}(K) \neq W(K)$, then

$$\|\boldsymbol{\delta_q}\|_K \leq C_V\, h_K^{1/2}\, \big(\|\boldsymbol{I_q} \cdot \boldsymbol{n}\|_{\partial K_V} + \|\alpha\|_{\partial K_V} \big(C_W\, R_{W^{\perp}}\, \|\delta_u\|_{\partial K_W} + \|I_u\|_{\partial K_V}\big)\big)\,,$$

by condition (3.13b). The estimate of $\|\boldsymbol{\Pi}_V \boldsymbol{q} - \boldsymbol{q}\,\|_K$ follows after using the triangle inequality and inserting the estimate of $\|\delta_u\|_{\partial K_W}$. This completes the proof of Theorem 3.3.1.

3.4 Examples of superconvergent methods

In this section, we use the template described in the previous section to construct superconvergent HDG methods. We begin by discussing the three main examples of stabilization operator α. We then give many examples of superconvergent methods using simplexes, squares, cubes and prisms. They include old and new (hybridized versions of) mixed and HDG methods. Finally, we end by briefly discussing how we used the template described in the previous section to construct the new superconvergent methods.

The verification of the inclusion properties in Steps 1 to 3 is fairly simple and will be left to the reader. However, for the most important cases, we are going to carry out the verification of the dimension count (3.12) and of the estimates (3.13) for the auxiliary spaces $\boldsymbol{V}^{\perp} \times W^{\perp}(K)$. To carry out the latter task, we only need to show that

$$\boldsymbol{v} \in \boldsymbol{V}^{\perp}(K) \ \text{ and } \boldsymbol{v}\cdot\boldsymbol{n}|_{\partial K_V} = 0 \quad \text{implies } \boldsymbol{v} = 0 \text{ on } K, \tag{3.17a}$$

$$w \in W^{\perp}(K) \text{ and } w|_{\partial K_W} = 0 \qquad \text{implies } w = 0 \text{ on } K. \tag{3.17b}$$

Indeed, the above conditions guarantee that $\|\boldsymbol{v}\cdot\boldsymbol{n}\|_{\partial K_V}, \|w\|_{\partial K_W}$ define norms on $\boldsymbol{V}^{\perp}(K)$ and $W^{\perp}(K)$ respectively. The estimates (3.13) then follow by the finite dimensionality of the spaces under consideration and by standard scaling arguments.

3.4.1 The main local stabilization operators α

There are three main examples of local stabilization operators α satisfying conditions (3.14).

(1) No stabilization

The first example is the trivial choice $\alpha := 0$ which is used in all the mixed methods.

(2) Pointwise stabilization

The second example is $\alpha := \tau\, Id$. If τ is taken to be a non-negative constant on each face F of each of the elements $K \in \mathcal{T}_h$ and strictly positive on one arbitrary face $F_K \in \partial K$, it is very easy to see that this operator satisfies the condition (3.14a) of being self-adjoint as well as the coercivity condition (3.14b) with $C_\alpha := \tau|_{F_K}$ and $\partial K_W := F_K$. Finally, note that $\|\alpha\|_F = \tau|_F$.

(3) Averaging stabilization

The third and last example is the local stabilization operator used by the so-called BMMPR-H methods; see also [6] and the references therein. It is is given by $\alpha := \tau \boldsymbol{r}\cdot\boldsymbol{n}$, where $\boldsymbol{r}$ is a suitably defined *lifting* operator. In [42] and [43], see also [5], such operator was introduced in the case in which K is a simplex and $\boldsymbol{V}(K)\times W(K) := \boldsymbol{P}^k(K)\times P^k(K)$ and $M(F) := P^k(F)$. We can extend its definition as follows. Given an element K, for any $\mu \in M(F)$, we define $\boldsymbol{r}(\mu)$ on K as the element of $\boldsymbol{V}(K)$ satisfying

$$\frac{1}{|K|}(\boldsymbol{r}(\mu), \boldsymbol{v})_K = \frac{1}{|F|}\langle \mu, \boldsymbol{v}\cdot\boldsymbol{n}\rangle_F \quad \text{for all } \boldsymbol{v} \in \boldsymbol{V}(K).$$

It is not difficult to see that the operator α satisfies the condition of being self-adjoint (3.14a) since, by taking $\boldsymbol{v} := \alpha(\eta)\,\boldsymbol{n}$ in the definition of $\boldsymbol{r}$, we get

$$\langle \alpha(\eta),\, \mu\rangle_F = \tau_F\, \langle \boldsymbol{r}(\eta)\cdot\boldsymbol{n},\, \mu\rangle_F = \tau_F\, \frac{|F|}{|K|}(\boldsymbol{r}(\eta), \boldsymbol{r}(\mu))_K.$$

Unlike the previous example, the verification of the positivity condition (3.14b) depends on the choice of local spaces. It is guaranteed if, for any $w^\perp \in W^\perp(K)$, we can find an element $\boldsymbol{v}$ of $\boldsymbol{V}(K)$ such that

$$\forall w^\perp \in W^\perp(K)\ \exists \boldsymbol{v} \in \boldsymbol{V}(K) :\ \boldsymbol{v}\cdot\boldsymbol{n}|_F = w^\perp|_F, \|\boldsymbol{v}\|_K \le \mathsf{C}\,\frac{|K|^{1/2}}{|F|^{1/2}}\|w^\perp\|_F. \tag{3.18}$$

Indeed, we have that

$$\|w^\perp\|_F^2 = \langle w^\perp, \boldsymbol{v}\cdot\boldsymbol{n}\rangle_F = \frac{|F|}{|K|}(\boldsymbol{r}(w^\perp), \boldsymbol{v})_K \le\ \mathsf{C}\,\frac{|F|^{1/2}}{|K|^{1/2}}\|\boldsymbol{r}(w^\perp)\|_K\, \|w^\perp\|_F,$$

and so $\|w^\perp\|_F \le\ \mathsf{C}\,\frac{|F|^{1/2}}{|K|^{1/2}}\|\boldsymbol{r}(w^\perp)\|_K$. On the other hand,

$$\langle \alpha(w^\perp), w^\perp\rangle_F = \tau\frac{|F|}{|K|}\,\|\boldsymbol{r}(w^\perp)\|^2 \ge \frac{\tau}{\mathsf{C}^2}\|w^\perp\|_F^2.$$

We thus see that the coercivity condition (3.14b) is satisfied with $C_\alpha := \tau/\mathsf{C}^2$ and $\partial K_W := F$.

It turns out that the condition (3.18) is verified for *all* the examples we present in this Chapter.

3.4.2 Methods using simplexes

We begin by considering methods for which the element K is a simplex.

Description of the methods

To describe the methods, we use the following notation: $P^k(D)$ denotes the space of polynomials of total degree k defined on D, $\widetilde{P}^k(D)$ denotes the space of homogeneous polynomials of degree k defined on D, $\boldsymbol{P}^k(D)$ denotes the space $[P^k(D)]^n$, $\mathcal{R}^k(\partial K)$ denotes the functions whose restriction to each face F of K belong to $P^k(F)$, and $\boldsymbol{\Phi}_k(K)$ denotes the space of functions in $\boldsymbol{P}^k(K)$ which are divergence-free and whose normal component on ∂K is zero.

Table 3.1: Methods for which $M(F) = P^k(F), k \geq 1$, and K is a simplex.

method	$\boldsymbol{V}(K)$	$W(K)$	$\widetilde{\boldsymbol{V}}(K)$	$\widetilde{W}(K)$
$\mathbf{BDFM}_{k+1}$	$\{\boldsymbol{q} \in \boldsymbol{P}^{k+1}(K) :$ $\boldsymbol{q} \cdot \boldsymbol{n}\|_{\partial K} \in \mathcal{R}^k(\partial K)\}$	$P^k(K)$	$\nabla P^k(K) \oplus \boldsymbol{\Phi}_{k+1}(K)$	$P^k(K)$
$\mathbf{RT}_k$	$\boldsymbol{P}^k(K) \oplus \boldsymbol{x}\widetilde{P}^k(K)$	$P^k(K)$	$\boldsymbol{P}^{k-1}(K)$	$P^k(K)$
$\mathbf{HDG}_k$	$\boldsymbol{P}^k(K)$	$P^k(K)$	$\boldsymbol{P}^{k-1}(K)$	$P^{k-1}(K)$
$\underset{k\geq 2}{\mathbf{BDM}_k}$	$\boldsymbol{P}^k(K)$	$P^{k-1}(K)$	$\nabla P^{k-1}(K) \oplus \boldsymbol{\Phi}_k(K)$	$P^{k-1}(K)$

Table 3.2: Orders of convergence for methods for which $M(F) = P^k(F), k \geq 1$, and K is a simplex.

method	∂K_V	∂K_W	τ	$\\|\boldsymbol{q} - \boldsymbol{q}_h\\|_{\mathcal{T}_h}$	$\\|\Pi_W u - u_h\\|_{\mathcal{T}_h}$	$\\|u - u_h^\star\\|_{\mathcal{T}_h}$
$\mathbf{BDFM}_{k+1}$	∂K	-	0	$k+1$	$k+2$	$k+2$
$\mathbf{RT}_k$	∂K	-	0	$k+1$	$k+2$	$k+2$
$\mathbf{HDG}_k$	$\partial K \setminus F_K$	F_K	$\mathcal{O}(1), > 0$	$k+1$	$k+2$	$k+2$
$\underset{k\geq 2}{\mathbf{BDM}_k}$	∂K	-	0	$k+1$	$k+2$	$k+2$

In Tables 3.1 and 3.2, we give methods that satisfy all the conditions in Steps 1 to 4 in our template. The orders of convergence, for smooth solutions, follow by combining our main results with the approximation properties of the auxiliary projection given by Theorem 3.3.1.

The methods are the well-known mixed methods of Raviart-Thomas, $\mathbf{RT}_k$, of Brezzi-Douglas-Marini, $\mathbf{BDM}_k$, and of Brezzi-Douglas-Fortin-Marini, $\mathbf{BDFM}_{k+1}$. In the three-dimensional case, the $\mathbf{RT}_k$ method was introduced in [44], and the method $\mathbf{BDM}_k$ in [16]. The $\mathbf{HDG}_k$ method with $\alpha := \tau \, Id$ was proposed in [6]; since the condition (3.18) is satisfied, we can also use $\alpha := \tau \boldsymbol{r} \cdot \boldsymbol{n}$.

Verification of the conditions on $\boldsymbol{V}^\perp(K) \times W^\perp(K)$

The projections Π_h associated with the mixed methods can be also constructed as indicated by our template; see [8]. In particular, the property (3.17a) is satisfied with $\partial K_V = \partial K$ and, since $W^\perp(K) = \emptyset$ (and $\alpha \equiv 0$), the property (3.17b) is automatically satisfied for any $\partial K_W \subset \partial K$. The same remark can be made about the projection Π_h

for the $\mathbf{HDG}_k$ method proposed in [11].

For the sake of completeness, let us verify that the conditions on the local space $\boldsymbol{V}^{\perp}(K) \times W^{\perp}(K)$ of the $\mathbf{HDG}_k$ method are satisfied. We begin by verifying the dimension count (3.12):

$$\sum_{F \in \mathcal{F}(K)} \dim M(F) = (d+1)\, \dim P_k(F),$$

$$\begin{aligned}
\dim \boldsymbol{V}^{\perp}(K) &= \dim \boldsymbol{V}(K) - \dim \widetilde{\boldsymbol{V}}(K) \\
&= d\,(\dim P_k(K) - \dim P_{k-1}(K)) \\
&= d\, \dim P_k(F),
\end{aligned}$$

$$\begin{aligned}
\dim W^{\perp}(K) &= \dim W(K) - \dim \widetilde{W}(K) \\
&= \dim P_k(K) - \dim P_{k-1}(K) \\
&= \dim P_k(F),
\end{aligned}$$

and the condition (3.12) follows.

Let us now verify the estimates (3.13). Take $w \in W^{\perp}(K)$ such that $w|_{F_K} = 0$. Then, we can write that $w = \lambda_{F_K} w'$, where $w' \in P^{k-1}(K)$ and λ_{F_K} is the barycentric coordinate function associated with the vertex of the simplex K opposite to the face F_K. Since $w \in W^{\perp}(K)$,

$$0 = (w, w')_K = (\lambda_{F_K} w', w')_K,$$

and this implies that $w' = 0$ on K since λ_{F_K} is always positive on K. This shows that (3.17b) holds.

We can verify the estimate (3.17a) in a similar way. Given any $\boldsymbol{v} \in \boldsymbol{V}^{\perp}(K)$, we can apply the above argument for each $\boldsymbol{v} \cdot \boldsymbol{n}_F$, $F \in \partial K \backslash F_K$, to conclude that $\boldsymbol{v} \cdot \boldsymbol{n}_F = 0$ in K. Since $\{\boldsymbol{n}_F : F \in \partial K \backslash F_K\}$ is a basis of $\mathbb{R}^d$, we have that $\boldsymbol{v} = \mathbf{0}$.

3.4.3 Methods using squares and cubes with $M(F) = P^k(F)$

Next, we consider methods for which the element K is a square ($n = 2$) or a cube ($n = 3$) and for which the space $M(F)$ is $P^k(F)$.

Description of the methods

Here, $P^{\ell_1,\ell_2}(D)$ for $n=2$ and P^{ℓ_1,ℓ_2,ℓ_3} for $n=3$ denote the space of polynomials of degree ℓ_i on the $i-th$ variable, $i=1,\dots,n$.

Table 3.3: Methods for which $M(F)=P^k(F), k\geq 1$, and K is a square.

method	$\boldsymbol{V}(K)$	$W(K)$	$\widetilde{\boldsymbol{V}}(K)$	$\widetilde{W}(K)$
$\mathbf{BDFM}_{[k+1]}$	$P^{k+1}(K)\backslash\{y^{k+1}\}$ $\times(P^{k+1}(K)\backslash\{x^{k+1}\})$	$P^k(K)$	$\boldsymbol{P}^{k-1}(K)$	$P^k(K)$
$\mathbf{HDG}^P_{[k]}$	$\boldsymbol{P}^k(K)$ $\oplus\nabla\times(xy\widetilde{P}^k(K))$	$P^k(K)$	$\boldsymbol{P}^{k-1}(K)$	$P^{k-1}(K)$
$\mathbf{BDM}_{[k]}$ $k\geq 2$	$\boldsymbol{P}^k(K)$ $\oplus\nabla\times(xy\,x^k)$ $\oplus\nabla\times(xy\,y^k)$	$P^{k-1}(K)$	$\boldsymbol{P}^{k-2}(K)$	$P^{k-1}(K)$

Table 3.4: Methods for which $M(F)=P^k(F), k\geq 1$, and K is a cube.

method	$\boldsymbol{V}(K)$	$W(K)$	$\widetilde{\boldsymbol{V}}(K)$	$\widetilde{W}(K)$
$\mathbf{BDFM}_{[k+1]}$	$P^{k+1}(K)\backslash\widetilde{P}^{k+1}(y,z)$ $\times P^{k+1}(K)\backslash\widetilde{P}^{k+1}(x,z)$ $\times P^{k+1}(K)\backslash\widetilde{P}^{k+1}(x,y)$	$P^k(K)$	$\boldsymbol{P}^{k-1}$	$P^k(K)$
$\mathbf{HDG}^P_{[k]}$	$\boldsymbol{P}^k(K)$ $\oplus\nabla\times(yz\widetilde{P}^k(K),0,0)$ $\oplus\nabla\times(0,zx\widetilde{P}^k(K),0)$	$P^k(K)$	$\boldsymbol{P}^{k-1}(K)$	$P^{k-1}(K)$
$\mathbf{BDM}_{[k]}$ $k\geq 2$	$\boldsymbol{P}^k(K)$ $\oplus\nabla\times(0,0,xy\widetilde{P}^k(y,z))$ $\oplus\nabla\times(0,xz\widetilde{P}^k(x,y),0)$ $\oplus\nabla\times(yz\widetilde{P}^k(x,z),0,0)$	$P^{k-1}(K)$	$\boldsymbol{P}^{k-2}(K)$	$P^{k-1}(K)$

In Table 3.3, we display the methods using squares and in Table 3.4, those using cubes. In Table 3.5, we display their orders of convergence. The mixed methods $\mathbf{BDFM}_{[k+1]}$ and $\mathbf{BDM}_{[k]}$ are well known but the $\mathbf{HDG}^P_{[k]}$ method is *new*.

Table 3.5: Orders of convergence for methods for which $M(F) = P^k(F), k \geq 1$, and K is a square or a cube.

method	∂K_V	∂K_W	τ	$\Vert \boldsymbol{q} - \boldsymbol{q}_h \Vert_{\mathcal{T}_h}$	$\Vert \Pi_W u - u_h \Vert_{\mathcal{T}_h}$	$\Vert u - u_h^\star \Vert_{\mathcal{T}_h}$
$\mathbf{BDFM}_{[k+1]}$	∂K	-	0	$k+1$	$k+2$	$k+2$
$\mathbf{HDG}^P_{[k]}$	∂K	F_K	$\mathcal{O}(1), > 0$	$k+1$	$k+2$	$k+2$
$\mathbf{BDM}_{[k]}$ $k \geq 2$	∂K	-	0	$k+1$	$k+2$	$k+2$

Verification of the conditions on $\boldsymbol{V}^\perp(K) \times W^\perp(K)$

For the mixed methods, see [8]. Let us now consider the $\mathbf{HDG}^P_{[k]}$ method for the case in which K is a cube; the case for which K is a square is simpler.

We begin by verifying the dimension count (3.12):

$$\sum_{F \in \mathcal{F}(K)} \dim M(F) = 6 \dim P_k(F),$$

$$\begin{aligned} \dim \boldsymbol{V}^\perp(K) &= \dim \boldsymbol{V}(K) - \dim \widetilde{\boldsymbol{V}}(K) \\ &= (3 \dim P_k(K) + 2 \dim \widetilde{P}_k(K)) - 3 \dim P_{k-1}(K)) \\ &= 5 \dim P_k(F), \end{aligned}$$

$$\begin{aligned} \dim W^\perp(K) &= \dim W(K) - \dim \widetilde{W}(K) \\ &= \dim P_k(K) - \dim P_{k-1}(K) \\ &= \dim P_k(F), \end{aligned}$$

and the condition (3.12) follows.

The proof for (3.17b) is exactly the same as the proof in the simplex case since the spaces $W(K)$ and $\widetilde{W}(K)$ are the same. Let us verify (3.17a). Take $\boldsymbol{v} \in \boldsymbol{V}^\perp(K)$ and assume that $\boldsymbol{v} \cdot \boldsymbol{n} = 0$ on ∂K. If two parallel faces of the cube K lie on the the planes $x = a, x = b$, respectively, we can conclude that $v_1 = 0$ on $x = a, x = b$, where v_1 is the first component of $\boldsymbol{v}$. So we can write $v_1 = (x-a)(x-b)v_1'$ where $v_1' \in P^{k-1}(K)$. Since $\boldsymbol{v} \in \boldsymbol{V}^\perp(K)$ and $\boldsymbol{v}' := (v_1', 0, 0) \in \widetilde{\boldsymbol{V}}(K)$, we have that

$$0 = (\boldsymbol{v}, \boldsymbol{v}')_K = ((x-a)(x-b)v_1', v_1')_K,$$

and we can conclude that $v_1' = 0$ on K. This implies that $v_1 = 0$ on K. A similar argument can be applied for the other two components. This means that $\boldsymbol{v} = \mathbf{0}$ on K.

3.4.4 Methods using squares and cubes with $M(F) = Q^k(F)$

Next, we consider methods for which the element K is a square ($n = 2$) or a cube ($n = 3$) and for which the space $M(F)$ is $Q^k(F)$. The main motivation for exploring this type of methods is that their tensor-product structure can be exploited to achieve a very efficient implementation; see [45, 46, 47] and the references therein.

Description of the methods

To describe the methods, w use the following notation. We denote by $Q^k(D)$ the space of polynomials of degree k in each variable defined on D, and by $\boldsymbol{Q}^k(D)$ the space $[Q^k(D)]^n$. We also set

$$\begin{aligned}
\boldsymbol{V}_0(K) &:= \{\boldsymbol{v} \in \boldsymbol{V}(K) : \ \boldsymbol{v} \cdot \boldsymbol{n}|_{\partial K} = 0\},\\
\boldsymbol{S}^k(K) &:= \{\boldsymbol{v} \in \boldsymbol{V}_0(K) : \ \nabla \cdot \boldsymbol{v} = 0\},\\
\boldsymbol{H}^k(K) &:= \{((x^2-x)x^{k-1}(aL_k(y)+b), (y^2-y)y^{k-1}(cL_k(z)+d),\\
&\qquad ((z^2-z)z^{k-1}(eL_k(x)+f))) : (a,b,c,d,e,f) \in \mathbb{R}^6\},\\
\boldsymbol{H}^k_M(K) &:= \boldsymbol{H}^k(K) \oplus \{((x^2-x)x^{k-1}L_k(y)L_k(z), 0, 0)\},
\end{aligned}$$

here $L_i(x)$ denotes the scaled *Legendre* polynomial of degree i on the interval $[0,1]$.

In Table 3.6, we display the methods using squares and in Table 3.7, those using cubes. Without lost of generality, we present the spaces on the reference domain $K = [0,1]^3$. In Table 3.8, we present their orders of convergence.

The mixed method $\mathbf{RT}_{[k]}$ is well known but the methods $\mathbf{TNT}_{[k]}$ and $\mathbf{HDG}^Q_{[k]}$ are *new*. As we see in Table 3.8, these three methods achieve the same orders of convergence. However, as pointed out in the Introduction, the spaces of the new methods are *significantly smaller* than those of the $\mathbf{RT}_{[k]}$ method. For example, in the case of cubic elements, only by adding 7 or 6 new basis functions to the space $\boldsymbol{Q}^k(K)$ we obtain the superconvergent methods $\mathbf{TNT}_{[k]}$ and $\mathbf{HDG}^Q_{[k]}$, respectively. Note that the method $\mathbf{TNT}_{[k]}$ (whose name stems from the fact that its local space $\boldsymbol{V}(K)$ is a **tiny** space containing the **tensor** product space $\boldsymbol{Q}^k(K)$) is a mixed method, as its stabilization function can be taken to be identically zero.

Table 3.6: Methods for which $M(F) = Q^k(F), k \geq 1$, and K is a square.

method	$\boldsymbol{V}(K)$	$W(K)$	$\widetilde{\boldsymbol{V}}(K)$	$\widetilde{W}(K)$
$\mathbf{RT}_{[k]}$	$P^{k+1,k}(K)$ $\times P^{k,k+1}(K)$	$Q^k(K)$	$P^{k-1,k}(K)$ $\times P^{k,k-1}(K)$	$Q^k(K)$
$\mathbf{TNT}_{[k]}$	$\boldsymbol{Q}^k(K)$ $\oplus\{(x^{k+1},0),(0,y^{k+1})\}$ $\oplus\{(x^{k+1}y^k,0)\}$	$Q^k(K)$	$\nabla Q^k(K) \oplus \boldsymbol{S}^k(K)$	$Q^k(K)$
$\mathbf{HDG}^Q_{[k]}$	$\boldsymbol{Q}^k(K)$ $\oplus\{(x^{k+1},0),(0,y^{k+1})\}$	$Q^k(K)$	$\nabla Q^k(K) \oplus \boldsymbol{S}^k(K)$	$Q^k(K)\backslash\{x^k y^k\}$

Table 3.7: Methods for which $M(F) = Q^k(F), k \geq 1$, and K is a cube.

method	$\boldsymbol{V}(K)$	$W(K)$	$\widetilde{\boldsymbol{V}}(K)$	$\widetilde{W}(K)$
$\mathbf{RT}_{[k]}$	$P^{k+1,k,k}(K)$ $\times P^{k,k+1,k}(K)$ $\times P^{k,k,k+1}(K)$	$Q^k(K)$	$P^{k-1,k,k}(K)$ $\times P^{k,k-1,k}(K)$ $\times P^{k,k,k-1}(K)$	$Q^k(K)$
$\mathbf{TNT}_{[k]}$	$\boldsymbol{Q}^k(K) \oplus \boldsymbol{H}^k_M(K)$	$Q^k(K)$	$\nabla Q^k(K) \oplus \boldsymbol{S}_k(K)$	$Q^k(K)$
$\mathbf{HDG}^Q_{[k]}$	$\boldsymbol{Q}^k(K) \oplus \boldsymbol{H}^k(K)$	$Q^k(K)$	$\nabla Q^k(K) \oplus \boldsymbol{S}_k(K)$	$Q^k(K)\backslash\{x^k y^k z^k\}$

Verification of the conditions on $\boldsymbol{V}^{\perp}(K) \times W^{\perp}(K)$

Again for the mixed method, see [8]. We only consider the method $\mathbf{HDG}^Q_{[k]}$ for cubic elements since the verification of the conditions on $\boldsymbol{V}^{\perp}(K) \times W^{\perp}(K)$ for squares is much simpler. The verification of those properties for the method $\mathbf{TNT}_{[k]}$ is almost identical.

We begin by verifying the dimension count (3.12). To do that, we first need to study the space $\widetilde{\boldsymbol{V}}(K) = \nabla Q^k(K) \oplus \boldsymbol{S}^k(K)$.

Lemma 3.4.1. *We have that $\nabla Q^k(K) \oplus \boldsymbol{S}^k(K)$ is a direct sum.*

Proof. Let us show that $\nabla Q^k(K) \cap \boldsymbol{S}^k(K) = \emptyset$. Assume that there is a nonzero function $\boldsymbol{v}$ in $\nabla Q^k(K) \cap \boldsymbol{S}^k(K)$. Since $\boldsymbol{v}$ in $\nabla Q^k(K)$, there is $w \in Q^k(K)$ such that $\boldsymbol{v} = \nabla w$. As

Table 3.8: Orders of convergence for methods for which $M(F) = Q^k(F), k \geq 1$, and K is a square or a cube.

method	∂K_V	∂K_W	τ	$\|\boldsymbol{q} - \boldsymbol{q}_h\|_{\mathcal{T}_h}$	$\|\Pi_W u - u_h\|_{\mathcal{T}_h}$	$\|u - u_h^\star\|_{\mathcal{T}_h}$
$\mathbf{RT}_{[k+1]}$	∂K	-	0	$k+1$	$k+2$	$k+2$
$\mathbf{TNT}_{[k]}$	∂K	-	0	$k+1$	$k+2$	$k+2$
$\mathbf{HDG}_{[k]}^{Q}$	∂K	F_K	$\mathcal{O}(1) > 0$	$k+1$	$k+2$	$k+2$

a consequence,

$$\|\boldsymbol{v}\|_K^2 = (\nabla w, \boldsymbol{v})_K = -(w, \nabla \cdot \boldsymbol{v}) + \langle \boldsymbol{v} \cdot \boldsymbol{n}, w \rangle_{\partial K} = 0,$$

since $\boldsymbol{v} \in \boldsymbol{S}^k(K)$. This implies that $\boldsymbol{v} = \boldsymbol{0}$ and completes the proof. □

The second lemma gives the dimension of $\boldsymbol{S}^k(K)$.

Lemma 3.4.2. *We have*

$$\dim \boldsymbol{S}^k(K) = 2 \, \dim Q^k(K) - 6 \dim Q^k(F) + 8.$$

Proof. By definition of the space $\boldsymbol{S}^k(K)$, we have that

$$\dim \boldsymbol{S}^k = \dim \boldsymbol{V}_0(K) - \dim \nabla \cdot \boldsymbol{V}_0(K).$$

Now, by definition of the space $\boldsymbol{V}_0(K)$, we can write that $\boldsymbol{V}_0(K) = \boldsymbol{E}_0(K) \oplus \boldsymbol{H}^k(K)$, where

$$\begin{aligned} \boldsymbol{E}_0(K) &:= \{\boldsymbol{v} \in \boldsymbol{Q}^k(K) : \ \boldsymbol{v} \cdot \boldsymbol{n}|_{\partial K} = 0\} \\ &= (x^2 - x)P^{k-2,k,k}(K) \times (y^2 - y)P^{k,k-2,k}(K) \times (z^2 - z)P^{k,k,k-2}(K). \end{aligned}$$

Next, we consider the spaces $\nabla \cdot \boldsymbol{E}_0(K)$ and $\nabla \cdot \boldsymbol{H}^k(K)$. Note that any function $f \in (x^2 - x)P^{k-2}(0,1)$ must be of the form

$$f(x) = \sum_{\ell=1}^{k-1} a_\ell \int_0^x L_\ell(s) \, ds.$$

This means that we can write

$$\begin{aligned} \nabla \cdot \boldsymbol{E}_0(K) = \mathit{span}\{L_{i_1}(x)L_{i_2}(y)L_{i_3}(z) : \ & 1 \leq i_1 \leq k-1, 0 \leq i_2, i_3 \leq k \quad \text{or} \\ & 1 \leq i_2 \leq k-1, 0 \leq i_1, i_3 \leq k \quad \text{or} \\ & 1 \leq i_3 \leq k-1, 0 \leq i_1, i_2 \leq k\}. \end{aligned}$$

On the other hand, we have that

$$\begin{aligned}\nabla \cdot \boldsymbol{H}^k(K) = span\{&((k+1)x^k - kx^{k-1})L_k(y), (k+1)x^k - kx^{k-1},\\ &((k+1)y^k - ky^{k-1})L_k(z), (k+1)y^k - ky^{k-1},\\ &((k+1)z^k - kz^{k-1})L_k(x), (k+1)z^k - kz^{k-1}\},\end{aligned}$$

and we can see that the basis functions in $\nabla \cdot \boldsymbol{H}^k(K)$ are linearly independent with the basis functions of $\nabla \cdot \boldsymbol{E}_0(K)$.

This implies that

$$\begin{aligned}\dim \boldsymbol{S}^k(K) &= \dim \boldsymbol{E}_0(K) - \dim \nabla \cdot \boldsymbol{E}_0(K)\\ &= (3\,\dim Q^k(K) - 6\dim Q^k(F)) - (\dim Q^k(K) - 8).\end{aligned}$$

This completes the proof. $\square$

We are now ready to verify the dimension count (3.12). We have

$$\begin{aligned}\sum_{F\in\mathcal{F}(K)} \dim M(F) &= 6\,\dim Q^k(F),\\ \dim W^{\perp}(K) &= \dim W(K) - \dim \widetilde{W}(K) = 1,\end{aligned}$$

and, by Lemma 3.4.1,

$$\begin{aligned}\dim \boldsymbol{V}^{\perp}(K) &= \dim \boldsymbol{V}(K) - \dim \widetilde{\boldsymbol{V}}(K)\\ &= (3\,\dim Q^k(K) + 6) - (\dim \nabla Q^k(K) + \dim \boldsymbol{S}^k(K))\\ &= 2\,\dim Q^k(K) + 7 - \dim \boldsymbol{S}^k(K)\\ &= 6\dim Q^k(F) - 1,\end{aligned}$$

by Lemma 3.4.2. This means that the condition (3.12) is verified.

Now, let us show that the properties (3.17) are true. Let us take $w \in W^{\perp}(K)$ such that $w|_{F_K} = 0$. Then we have that $w = \lambda_{F_K} w'$, where $w' \in \widetilde{W}(K)$. By using the same argument as in the proof in the case of simplex we can conclude that $w' = 0$ which in turn implies that $w = 0$. So (3.17b) is true.

Now take $\boldsymbol{v} \in \boldsymbol{V}^{\perp}(K)$ such that $\boldsymbol{v}\cdot\boldsymbol{n}|_{\partial K} = 0$. We are going to show that $\boldsymbol{v} \in \boldsymbol{S}^k(K)$, which would imply that $\boldsymbol{v} = \boldsymbol{0}$ on K, since $\widetilde{\boldsymbol{V}}(K) \supset \boldsymbol{S}^k(K)$. To do that, we note that, for any $w \in W(K) = Q^k(K)$,

$$(\nabla \cdot \boldsymbol{v}, w)_K = \langle \boldsymbol{v} \cdot \boldsymbol{n}, w \rangle_{\partial K} - (\boldsymbol{v}, \nabla w)_K = 0,$$

because $\widetilde{\boldsymbol{V}}(K) \supset \nabla Q^k(K)$. Since $\nabla \cdot \boldsymbol{V}(K) \subset W(K) = Q^k(K)$, this implies that $\nabla \cdot \boldsymbol{v} = 0$ and $\boldsymbol{v} \in \boldsymbol{S}^k(K)$. This shows that (3.17a) is also true.

3.4.5 Methods using prisms

Finally, we present three prismatic finite elements.

Description of the methods

We consider the prism whose base is a triangle in the $(x, y)-$plane and whose lateral faces are parallel to the $z-$axis. We denote by $P^{m|n}(K)$ the space of polynomials of degree m in the two variables x and y and of degree n in the variable z. We also denote by $M_V(F), M_H(F)$ the finite dimensional spaces $M(F)$ on the vertical and horizontal faces, respectively. Finally, we set

$$\begin{aligned}
B^{k+2}(K) &:= \{w \in P^{k+2}(K) : w|_F = 0, \text{on three vertical faces}\},\\
B^{k+2|k}(K) &:= \{w \in P^{k+2|k}(K) : w|_F = 0, \text{on three vertical faces}\},\\
\boldsymbol{Y}^{k+1}(K) &:= [\nabla_{(x,y)} P^{k-1}(K) \oplus \nabla_{(x,y)} \times B^{k+2}(K)] \times P^{k-1}(K),\\
\boldsymbol{Z}^k(K) &:= [\nabla_{(x,y)} P^{k|k}(K) \oplus \nabla_{(x,y)} \times B^{k+2|k}(K)] \times P^{k|k-1}(F).
\end{aligned}$$

Here, $\nabla_{(x,y)}, \nabla_{(x,y)}\cdot, \nabla_{(x,y)}\times$ denote the corresponding differential operators in the variables x and y.

In Table 3.9, we display the methods and is Table 3.10, we present their orders of convergence. The methods $\mathbf{BDFM}_{<\mathbf{k+1}>}$ and $\mathbf{RT}_{<\mathbf{k}>}$ were introduced in [48]; the last one is a *new* $\mathbf{HDG}_{<k>}$ element.

Verification of the conditions on $\boldsymbol{V}^\perp(K) \times W^\perp(K)$

The methods $\mathbf{BDFM}_{<k+1>}, \mathbf{RT}_{<k>}$ were introduced and studied in [48]. It is not difficult to verify that they satisfy all the conditions of the template. Here we only consider the $\mathbf{HDG}_{<k>}$ method.

Table 3.9: Methods for K is a prism.

BDFM$_{<\mathbf{k+1}>}$					
$\boldsymbol{V}(K)$	$W(K)$	$\widetilde{\boldsymbol{V}}(K)$	$\widetilde{W}(K)$	$M_V(F)$	$M_H(F)$
$P^{k+1}(K)\backslash\{z^{k+1}\}$ $\times P^{k+1}(K)\backslash\{z^{k+1}\}$ $\times P^{k+1}(K)\backslash\{\widetilde{P}^{k+1}(x,y)\}$	$P^k(K)$	$\boldsymbol{Y}^{k+1}(K)$	$P^k(K)$	$P^{k+1}(F)\backslash\{z^{k+1}\}$	$P^k(F)$
RT$_{<\mathbf{k}>}$					
$\boldsymbol{V}(K)$	$W(K)$	$\widetilde{\boldsymbol{V}}(K)$	$\widetilde{W}(K)$	$M_V(F)$	$M_H(F)$
$P^{k+1\|k}(K)$ $\times P^{k+1\|k}(K)$ $\times P^{k\|k+1}(K)$	$P^{k\|k}(K)$	$\boldsymbol{Z}^k(K)$	$P^{k\|k}(K)$	$P^{k+1,k}(F)$	$P^k(F)$
HDG$_{<\mathbf{k}>}$					
$\boldsymbol{V}(K)$	$W(K)$	$\widetilde{\boldsymbol{V}}(K)$	$\widetilde{W}(K)$	$M_V(F)$	$M_H(F)$
$\boldsymbol{P}^k(K)$ $\oplus\nabla\times\{(y,-x,0)z\widetilde{P}^k(K)\}$	$P^k(K)$	$\boldsymbol{P}^{k-1}(K)$	$P^{k-1}(K)$	$P^k(F)$	$P^k(F)$

Table 3.10: Orders of convergence for methods for which K is a prism.

method	∂K_V	∂K_W	τ	$\\|\boldsymbol{q}-\boldsymbol{q}_h\\|_{\mathcal{T}_h}$	$\\|\Pi_W u-u_h\\|_{\mathcal{T}_h}$	$\\|u-u_h^\star\\|_{\mathcal{T}_h}$
BDFM$_{<k+1>}$	∂K	-	0	$k+1$	$k+2$	$k+2$
RT$_{<k>}$	∂K	-	0	$k+1$	$k+2$	$k+2$
HDG$_{<k>}$	∂K	F_K	$\mathcal{O}(1),>0$	$k+1$	$k+2$	$k+2$

We begin by verifying the dimension count (3.12):

$$\sum_{F\in\mathcal{F}(K)} \dim M(F) = 5\,\dim P_k(F),$$

$$\begin{aligned}
\dim \boldsymbol{V}^\perp(K) &= \dim \boldsymbol{V}(K) - \dim \widetilde{\boldsymbol{V}}(K)\\
&= (3\,\dim P_k(K) + \dim \widetilde{P}_k(K)) - 3\,\dim P_{k-1}(K))\\
&= 4\,\dim P_k(F),
\end{aligned}$$

$$\begin{aligned}
\dim W^\perp(K) &= \dim W(K) - \dim \widetilde{W}(K)\\
&= \dim P_k(K) - \dim P_{k-1}(K)\\
&= \dim P_k(F),
\end{aligned}$$

and the condition (3.12) follows.

The property (3.17b) holds since the spaces $W(K)$ and $W^{\perp}(K)$ are the same as those for simplexes. It remains to verify (3.17a). Assume that the two horizontal faces of the prism K lie on the planes $z = a$ and $z = b$, respectively. By the definition of $\boldsymbol{V}(K)$, for any $\boldsymbol{v} \in \boldsymbol{V}^{\perp}(K)$, we can write

$$\boldsymbol{v} = (p_1 + x\frac{\partial z\widetilde{p}}{\partial z}, p_2 + y\frac{\partial z\widetilde{p}}{\partial z}, p_3 - z(\frac{\partial x\widetilde{p}}{\partial x} + \frac{\partial y\widetilde{p}}{\partial y})),$$

where $(p_1, p_2, p_3) \in \boldsymbol{P}^k(K)$ and that $\widetilde{p} \in \widetilde{P}^k(K)$. Noting that $v_3 \in P^{k+1}(K)$ and $\boldsymbol{v} \cdot \boldsymbol{n}|_{z=a,b} = v_3|_{z=a,b} = 0$, we can apply the same argument used in the case of cubic elements to get that $v_3 = 0$ in K. This means that

$$p_3 = 0 \ , \quad \frac{\partial x\widetilde{p}}{\partial x} + \frac{\partial y\widetilde{p}}{\partial y} = 0 \ .$$

Since $\widetilde{p} = \sum_{l+m+n=k} a_{lmn} x^l y^m z^n$, we have

$$\frac{\partial x\widetilde{p}}{\partial x} + \frac{\partial y\widetilde{p}}{\partial y} = \sum_{l+m+n=k} (l+m+2) a_{lmn} x^l y^m z^n,$$

and so, all the coefficients $a_{lmn} = 0$. We then we have $\widetilde{p} = 0$. Therefore, $v_1, v_2 \in P^k(K)$. It is now very easy to conclude that $v_1 = v_2 = 0$ by using the fact that $\boldsymbol{v} \cdot \boldsymbol{n} = 0$ on the vertical faces and $\boldsymbol{v} \in \boldsymbol{V}^{\perp}(K)$. This shows that property (3.17a) also holds.

3.4.6 A remark on the construction of the superconvergent methods

Let us briefly discuss the use of the template proposed for the construction of the superconvergent methods. In the simplex case, we have

$$\begin{aligned} &W(K) = P^k(K) \ , \ \boldsymbol{V} = \boldsymbol{P}^k(K) \ , \ M(F) = P^k(F), \\ &\widetilde{W}(K) = P^{k-1}(K) \ , \ \widetilde{\boldsymbol{V}}(K) = \boldsymbol{P}^{k-1}(K) \ , \end{aligned}$$

and

$$\sum_{F \in \mathcal{F}(F)} \dim M(F) = (\dim W^{\perp}(K) + \dim \boldsymbol{V}^{\perp}(K)).$$

If we were to keep the same local spaces for, say, rectangular elements, the condition (3.12) would not be satisfied because we have one additional face. Indeed, we would have

$$\sum_{F\in\mathcal{F}(F)} \dim M(F) = (\dim W^{\perp}(K) + \dim \boldsymbol{V}^{\perp}(K)) + \dim P^k(F).$$

As discussed in Subsection 3.1, to remedy this situation, we must modify the local space $\boldsymbol{V}(K)$ by adding new basis functions (i) which are divergence-free, (ii) which are such that their normal component on the face F lies on $M(F) = P^k(F)$, and (iii) whose behavior in the element is controlled by the behavior on the normal component on its boundary. The condition (i), suggest to add $\dim P^k(F) = k+1$ new basis functions of the form

$$\nabla \times f = (-\frac{\partial f}{\partial y}, \frac{\partial f}{\partial x}),$$

where $f \in \widetilde{P}^{k+2}(K)$. The conditions (ii) and (iii) now suggest to take f in the space $xy\widetilde{P}^k(K)$. All the *new* superconvergent elements were found in a similar manner.

3.5 Proofs of the estimates of the projection of the errors

In this section we provide detailed proofs for our a priori error estimates. The main idea is to work with the following projection of the errors:

$$\begin{aligned}
\boldsymbol{e}_q &:= \boldsymbol{\Pi}_V \boldsymbol{q} - \boldsymbol{q}_h,\\
e_u &:= \Pi_W u - u_h,\\
\boldsymbol{e}_{\widehat{q}} \cdot \boldsymbol{n} &:= P_M(\boldsymbol{q}\cdot\boldsymbol{n} - \widehat{\boldsymbol{q}}_h \cdot \boldsymbol{n}),\\
e_{\widehat{u}} &:= P_M u - \widehat{u}_h.
\end{aligned}$$

Here, P_M is the L^2-projection from $L^2(\mathcal{E}_h)$ into M_h. We abuse the notation for the sake of simplicity and denote with the *same* symbol the L^2-projection from $L^2(\partial\mathcal{T}_h)$ into the space

$$\{w \in L^2(\partial\mathcal{T}_h) : (w|_{\partial K})|_F \in M(F) \text{ for all faces } F \text{ of } K \text{ and all } K \in \mathcal{T}_h\}.$$

We begin by obtaining the equations satisfied by these projections. We then use an energy argument to obtain an estimate of $\boldsymbol{e}_q$; this would prove Theorem 3.2.1. To obtain an estimate of e_u and prove Theorem 3.2.2, we employ an elliptic duality. Finally, we obtain the estimate of $u - u_h^*$ of Theorem 3.2.3 by using a simple element-by-element argument.

Step 1: The equations for the projection of the errors

We begin our error analysis with the following auxiliary result.

Lemma 3.5.1. *Suppose that the orthogonality properties of the projection* Π_h *and the properties of the traces of the local spaces of Assumption A are satisfied. Then, we have*

$$
\begin{aligned}
(\boldsymbol{e}_q\,,\,\boldsymbol{v})_{\mathcal{T}_h} - (e_u\,,\,\nabla\cdot\boldsymbol{v})_{\mathcal{T}_h} + \langle e_{\widehat{u}}\,,\,\boldsymbol{v}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} &= (\boldsymbol{\Pi}_V\boldsymbol{q}-\boldsymbol{q}\,,\,\boldsymbol{v})_{\mathcal{T}_h}, && (3.19\text{a})\\
-(\boldsymbol{e}_q\,,\,\nabla w)_{\mathcal{T}_h} + \langle \boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}\,,\,w\rangle_{\partial\mathcal{T}_h} &= 0, && (3.19\text{b})\\
\langle \boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n},\mu\rangle_{\partial\mathcal{T}_h\setminus\partial\Omega} &= 0, && (3.19\text{c})\\
\langle e_{\widehat{u}}\,,\,\mu\rangle_{\partial\Omega} &= 0, && (3.19\text{d})
\end{aligned}
$$

for all $(\boldsymbol{v},w,\mu)\in\boldsymbol{V}_h\times W_h\times M_h$*. Moreover,*

$$
\boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n} = \boldsymbol{e}_q\cdot\boldsymbol{n} + P_M(\alpha(e_u - e_{\widehat{u}})) \quad \text{on} \quad \partial\mathcal{T}_h. \tag{3.20}
$$

Proof. Let us begin by noting that the exact solution $(\boldsymbol{q},u)$ satisfies the equations

$$
\begin{aligned}
(\boldsymbol{q}\,,\,\boldsymbol{v})_{\mathcal{T}_h} - (u\,,\,\nabla\cdot\boldsymbol{v})_{\mathcal{T}_h} + \langle u\,,\,\boldsymbol{v}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} &= 0,\\
-(\boldsymbol{q}\,,\,\nabla w)_{\mathcal{T}_h} + \langle \boldsymbol{q}\cdot\boldsymbol{n}\,,\,w\rangle_{\partial\mathcal{T}_h} &= (f\,,\,w)_{\mathcal{T}_h},\\
\langle \boldsymbol{q}\cdot\boldsymbol{n},\mu\rangle_{\partial\mathcal{T}_h\setminus\partial\Omega} &= 0,\\
\langle u\,,\,\mu\rangle_{\partial\Omega} &= \langle g\,,\,\mu\rangle_{\partial\Omega},
\end{aligned}
$$

for all $(\boldsymbol{v},w,\mu)\in\boldsymbol{V}_h\times W_h\times M_h$. By the orthogonality properties $(A.1)$ and $(A.2)$ of the projection $\Pi_h=(\boldsymbol{\Pi}_V,\Pi_W)$, we obtain that

$$
\begin{aligned}
(\boldsymbol{q}\,,\,\boldsymbol{v})_{\mathcal{T}_h} - (\Pi_W u\,,\,\nabla\cdot\boldsymbol{v})_{\mathcal{T}_h} + \langle u\,,\,\boldsymbol{v}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} &= 0,\\
-(\boldsymbol{\Pi}_V\boldsymbol{q}\,,\,\nabla w)_{\mathcal{T}_h} + \langle \boldsymbol{q}\cdot\boldsymbol{n}\,,\,w\rangle_{\partial\mathcal{T}_h} &= (f\,,\,w)_{\mathcal{T}_h},\\
\langle \boldsymbol{q}\cdot\boldsymbol{n},\mu\rangle_{\partial\mathcal{T}_h\setminus\partial\Omega} &= 0,\\
\langle u\,,\,\mu\rangle_{\partial\Omega} &= \langle g\,,\,\mu\rangle_{\partial\Omega},
\end{aligned}
$$

for all $(\boldsymbol{v},w,\mu)\in\boldsymbol{V}_h\times W_h\times M_h$. Moreover, since P_M is the L^2-projection into M_h, we get, by the properties $(A.4)$ and $(A.5)$ of the traces of the local spaces, that

$$\begin{aligned}
(\boldsymbol{q}\,,\,\boldsymbol{v})_{\mathcal{T}_h} - (\Pi_W u\,,\,\nabla\cdot\boldsymbol{v})_{\mathcal{T}_h} + \langle P_M u\,,\,\boldsymbol{v}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} &= 0,\\
-(\boldsymbol{\Pi}_V\boldsymbol{q}\,,\,\nabla w)_{\mathcal{T}_h} + \langle P_M(\boldsymbol{q}\cdot\boldsymbol{n})\,,\,w\rangle_{\partial\mathcal{T}_h} &= (f\,,\,w)_{\mathcal{T}_h},\\
\langle P_M(\boldsymbol{q}\cdot\boldsymbol{n}),\mu\rangle_{\partial\mathcal{T}_h\setminus\partial\Omega} &= 0,\\
\langle P_M u\,,\,\mu\rangle_{\partial\Omega} &= \langle g\,,\,\mu\rangle_{\partial\Omega},
\end{aligned}$$

for all $(\boldsymbol{v},w,\mu)\in\boldsymbol{V}_h\times W_h\times M_h$. Subtracting the first four equations defining the weak formulation of the HDG method (3.2) from the above equations, respectively, we obtain the equations for the projection of the errors.

It remains to prove the identity for $\boldsymbol{e}_{\widehat{q}}$. We have

$$\begin{aligned}
\boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n} &= P_M(\boldsymbol{q}\cdot\boldsymbol{n}) - P_M(\widehat{\boldsymbol{q}}_h\cdot\boldsymbol{n})\\
&= P_M(\boldsymbol{\Pi}_V\boldsymbol{q}\cdot\boldsymbol{n} + \alpha(\Pi_W u - P_M u)) - P_M(\widehat{\boldsymbol{q}}_h\cdot\boldsymbol{n}),
\end{aligned}$$

by the orthogonality property $(A.3)$ of the projection Π_h. Inserting the definition of the numerical trace $\widehat{\boldsymbol{q}}_h\cdot\boldsymbol{n}$, (3.3), we get

$$\begin{aligned}
\boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n} &= P_M(\boldsymbol{e}_q\cdot\boldsymbol{n} + \alpha(e_u - e_{\widehat{u}}))\\
&= \boldsymbol{e}_q\cdot\boldsymbol{n} + P_M(\alpha(e_u - e_{\widehat{u}})),
\end{aligned}$$

by the property $(A.4)$ of the trace of the local spaces. This completes the proof. □

Step 2: The energy argument for $\boldsymbol{e}_q$

We are now ready to obtain the upper bound of the L^2-norm of $\boldsymbol{e}_q$. We proceed as follows. Taking $\boldsymbol{v} := \boldsymbol{e}_q$ in the error equation (3.19a), $w := e_u$ in the error equation (3.19b), $\mu := -e_{\widehat{u}}$ in the error equation (3.19c), and $\mu := -P_M(\boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n})$ in the error equation (3.19d), and adding the resulting equations up, we obtain

$$(\boldsymbol{e}_q\,,\,\boldsymbol{e}_q)_{\mathcal{T}_h} + \Theta_h = (\boldsymbol{\Pi}_V\boldsymbol{q} - \boldsymbol{q}\,,\,\boldsymbol{e}_q)_{\mathcal{T}_h},$$

where

$$\begin{aligned}
\Theta_h :=\ &\langle e_{\widehat{u}}\,,\,\boldsymbol{e}_q\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} + \langle\boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}\,,\,e_u\rangle_{\partial\mathcal{T}_h}\\
&- \langle\boldsymbol{e}_q\cdot\boldsymbol{n}\,,\,e_u\rangle_{\partial\mathcal{T}_h} - \langle\boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n},e_{\widehat{u}}\rangle_{\partial\mathcal{T}_h\setminus\partial\Omega} - \langle P_M(\boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}),e_{\widehat{u}}\rangle_{\partial\Omega}.
\end{aligned}$$

By the definition of the projection P_M, we get that

$$\begin{aligned}\Theta_h &= \langle e_{\widehat{u}}\,,\, \boldsymbol{e}_q\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} + \langle \boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}\,,\, e_u\rangle_{\partial\mathcal{T}_h}\\ &\quad - \langle \boldsymbol{e}_q\cdot\boldsymbol{n}\,,\, e_u\rangle_{\partial\mathcal{T}_h} - \langle \boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}, e_{\widehat{u}}\rangle_{\partial\mathcal{T}_h\setminus\partial\Omega} - \langle \boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}, e_{\widehat{u}}\rangle_{\partial\Omega}\\ &= \langle (\boldsymbol{e}_{\widehat{q}} - \boldsymbol{e}_q)\cdot\boldsymbol{n}\,,\, e_u - e_{\widehat{u}}\rangle_{\partial\mathcal{T}_h}\\ &= \langle P_M(\alpha(e_u - e_{\widehat{u}}))\,,\, e_u - e_{\widehat{u}}\rangle_{\partial\mathcal{T}_h},\end{aligned}$$

by the identity (3.20) of Lemma 3.5.1. Finally, by the definition of the projection P_M and the property (*A.5*) of the traces of the local spaces, we obtain that

$$\Theta_h = \langle \alpha(e_u - e_{\widehat{u}})\,,\, e_u - e_{\widehat{u}}\rangle_{\partial\mathcal{T}_h}.$$

Since $\Theta_h \geq 0$, by the semi-positivity property (*A.6*) of the local stabilization operator α, we have that

$$\begin{aligned}\|\boldsymbol{e}_q\|^2_{\mathcal{T}_h} &\leq (\boldsymbol{\Pi}_V\boldsymbol{q} - \boldsymbol{q}\,,\, \boldsymbol{e}_q)_{\mathcal{T}_h}\\ &\leq \|\boldsymbol{\Pi}_V\boldsymbol{q} - \boldsymbol{q}\|_{\mathcal{T}_h}\|\boldsymbol{e}_q\|_{\mathcal{T}_h},\end{aligned}$$

and the result follows. This completes the proof of Theorem 3.2.1.

Step 3: The elliptic duality argument for e_u

The estimate of e_u will follow from the following identity.

Lemma 3.5.2. *Suppose that the assumptions of Lemma 3.5.1 are satisfied. Then, we have*

$$(e_u\,,\,\eta)_{\mathcal{T}_h} = (\boldsymbol{q} - \boldsymbol{\Pi}_V\boldsymbol{q}\,,\, \boldsymbol{\Pi}^*_V\boldsymbol{\theta})_{\mathcal{T}_h} - (\boldsymbol{e}_q\,,\, \boldsymbol{\theta} - \boldsymbol{\Pi}^*_V\boldsymbol{\theta})_{\mathcal{T}_h},$$

where $(\phi, \boldsymbol{\theta})$ *is the solution of dual problem* (3.6).

Proof. We begin by using the second equation (3.6b) of the dual problem to write that

$$\begin{aligned}(e_u\,,\,\eta)_{\mathcal{T}_h} &= (e_u\,,\,\nabla\cdot\boldsymbol{\theta})_{\mathcal{T}_h}\\ &= (e_u\,,\,\nabla\cdot\boldsymbol{\theta})_{\mathcal{T}_h} - (\boldsymbol{e}_q\,,\,\boldsymbol{\theta})_{\mathcal{T}_h} - (\boldsymbol{e}_q\,,\,\nabla\phi)_{\mathcal{T}_h},\end{aligned}$$

by the first equation (3.6a) of the dual problem. This implies that

$$\begin{aligned}(e_u\,,\,\eta)_{\mathcal{T}_h} &= (e_u\,,\,\nabla\cdot\boldsymbol{\Pi}^*_V\boldsymbol{\theta})_{\mathcal{T}_h} - (\boldsymbol{e}_q\,,\,\boldsymbol{\Pi}^*_V\boldsymbol{\theta})_{\mathcal{T}_h} - (\boldsymbol{e}_q\,,\,\nabla\Pi^*_W\phi)_{\mathcal{T}_h}\\ &\quad + (e_u\,,\,\nabla\cdot(\boldsymbol{\theta} - \boldsymbol{\Pi}^*_V\boldsymbol{\theta}))_{\mathcal{T}_h} - (\boldsymbol{e}_q\,,\,\boldsymbol{\theta} - \boldsymbol{\Pi}^*_V\boldsymbol{\theta})_{\mathcal{T}_h} - (\boldsymbol{e}_q\,,\,\nabla(\phi - \Pi^*_W\phi))_{\mathcal{T}_h}.\end{aligned}$$

Taking $\boldsymbol{v} := \boldsymbol{\Pi}_V^* \boldsymbol{\theta}$ in the first error equation, (3.19a), and $w := \Pi_W^* \phi$ in the second, (3.19b), we obtain that

$$\begin{aligned}(e_u\,,\,\eta)_{\mathcal{T}_h} = &(\boldsymbol{q} - \boldsymbol{\Pi}_V \boldsymbol{q}\,,\,\boldsymbol{\Pi}_V^* \boldsymbol{\theta})_{\mathcal{T}_h} + \langle e_{\widehat{u}}\,,\,\boldsymbol{\Pi}_V^* \boldsymbol{\theta} \cdot \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} - \langle \boldsymbol{e}_{\widehat{q}} \cdot \boldsymbol{n}\,,\,\Pi_W^* \phi \rangle_{\partial \mathcal{T}_h} \\ &+ (e_u\,,\,\nabla \cdot (\boldsymbol{\theta} - \boldsymbol{\Pi}_V^* \boldsymbol{\theta}))_{\mathcal{T}_h} - (\boldsymbol{e}_q\,,\,\boldsymbol{\theta} - \boldsymbol{\Pi}_V^* \boldsymbol{\theta})_{\mathcal{T}_h} - (\boldsymbol{e}_q\,,\,\nabla(\phi - \Pi_W^* \phi))_{\mathcal{T}_h},\end{aligned}$$

and, after simple algebraic manipulations, that

$$(e_u\,,\,\eta)_{\mathcal{T}_h} = (\boldsymbol{q} - \boldsymbol{\Pi}_V \boldsymbol{q}\,,\,\boldsymbol{\Pi}_V^* \boldsymbol{\theta})_{\mathcal{T}_h} - (\boldsymbol{e}_q\,,\,\boldsymbol{\theta} - \boldsymbol{\Pi}_V^* \boldsymbol{\theta})_{\mathcal{T}_h} + \mathbb{T},$$

where

$$\begin{aligned}\mathbb{T} :=&\langle e_{\widehat{u}}\,,\,\boldsymbol{\Pi}_V^* \boldsymbol{\theta} \cdot \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} - \langle \boldsymbol{e}_{\widehat{q}} \cdot \boldsymbol{n}\,,\,\Pi_W^* \phi \rangle_{\partial \mathcal{T}_h} \\ &+ (e_u\,,\,\nabla \cdot (\boldsymbol{\theta} - \boldsymbol{\Pi}_V^* \boldsymbol{\theta}))_{\mathcal{T}_h} - (\boldsymbol{e}_q\,,\,\nabla(\phi - \Pi_W^* \phi))_{\mathcal{T}_h}.\end{aligned}$$

It remains to prove that $\mathbb{T} = 0$.

To do that, we integrate by parts and use the orthogonality properties (*A.1*) and (*A.2*) of the projection Π_h^* to get

$$\begin{aligned}\mathbb{T} =& \langle e_{\widehat{u}}\,,\,\boldsymbol{\Pi}_V^* \boldsymbol{\theta} \cdot \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} - \langle \boldsymbol{e}_{\widehat{q}} \cdot \boldsymbol{n}\,,\,\Pi_W^* \phi \rangle_{\partial \mathcal{T}_h} \\ &+ \langle e_u\,,\,(\boldsymbol{\theta} - \boldsymbol{\Pi}_V^* \boldsymbol{\theta}) \cdot \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} - \langle \boldsymbol{e}_q \cdot \boldsymbol{n}\,,\,(\phi - \Pi_W^* \phi) \rangle_{\partial \mathcal{T}_h} \\ =& \langle e_{\widehat{u}} - e_u\,,\,(\boldsymbol{\Pi}_V^* \boldsymbol{\theta} - \boldsymbol{\theta}) \cdot \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} - \langle (\boldsymbol{e}_{\widehat{q}} - \boldsymbol{e}_q) \cdot \boldsymbol{n}\,,\,\Pi_W^* \phi - \phi \rangle_{\partial \mathcal{T}_h} \\ &+ \langle e_{\widehat{u}}\,,\,\boldsymbol{\theta} \cdot \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} - \langle \boldsymbol{e}_{\widehat{q}} \cdot \boldsymbol{n}\,,\,\phi \rangle_{\partial \mathcal{T}_h} \\ =& \langle e_{\widehat{u}} - e_u\,,\,(\boldsymbol{\Pi}_V^* \boldsymbol{\theta} - \boldsymbol{\theta}) \cdot \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} - \langle (\boldsymbol{e}_{\widehat{q}} - \boldsymbol{e}_q) \cdot \boldsymbol{n}\,,\,\Pi_W^* \phi - \phi \rangle_{\partial \mathcal{T}_h}.\end{aligned}$$

Indeed, the fact that $\langle e_{\widehat{u}}\,,\,\boldsymbol{\theta} \cdot \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} = 0$ is proven as follows. Since $e_{\widehat{u}}$ is single valued on $\mathcal{E}_h$ and $\boldsymbol{\theta}$ lies in $\boldsymbol{H}(\mathrm{div})$, we have that

$$\begin{aligned}\langle e_{\widehat{u}}\,,\,\boldsymbol{\theta} \cdot \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} &= \langle e_{\widehat{u}}, \boldsymbol{\theta} \cdot \boldsymbol{n} \rangle_{\partial \Omega} \\ &= \langle e_{\widehat{u}}, P_M(\boldsymbol{\theta} \cdot \boldsymbol{n}) \rangle_{\partial \Omega},\end{aligned}$$

by the definition of the projection P_M. The above quantity is equal to zero by the fourth error equation (3.19d) with $\mu := P_M(\boldsymbol{\theta} \cdot \boldsymbol{n})$.

We can prove that $\langle \boldsymbol{e}_{\widehat{q}} \cdot \boldsymbol{n}\,,\,\phi \rangle_{\partial \mathcal{T}_h} = 0$ as follows. By the definition of the projection P_M,

$$\begin{aligned}\langle \boldsymbol{e}_{\widehat{q}} \cdot \boldsymbol{n}\,,\,\phi \rangle_{\partial \mathcal{T}_h} &= \langle \boldsymbol{e}_{\widehat{q}} \cdot \boldsymbol{n}\,,\,P_M(\phi) \rangle_{\partial \mathcal{T}_h} \\ &= \langle \boldsymbol{e}_{\widehat{q}} \cdot \boldsymbol{n}, \phi \rangle_{\partial \Omega},\end{aligned}$$

by the third error equation (3.19c) with $\mu := P_M(\phi)$. Finally, by the third equation (3.6c) of the dual problem, we have that $\phi = 0$ on $\partial\Omega$ and the result follows.

Now, inserting the expression for $\boldsymbol{e}_{\widehat{q}}$, (3.20), given by Lemma 3.5.1 in the last expression for $\mathbb{T}$, we get

$$\begin{aligned}\mathbb{T} &= \langle e_{\widehat{u}} - e_u \,,\, (\boldsymbol{\Pi}_V^* \boldsymbol{\theta} - \boldsymbol{\theta}) \cdot \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} - \langle P_M(\alpha(e_u - e_{\widehat{u}})) \,,\, \Pi_W^* \phi - \phi \rangle_{\partial \mathcal{T}_h} \\ &= \langle e_{\widehat{u}} - e_u \,,\, (\boldsymbol{\Pi}_V^* \boldsymbol{\theta} - \boldsymbol{\theta}) \cdot \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} - \langle \alpha(e_u - e_{\widehat{u}}) \,,\, \Pi_W^* \phi - P_M \phi \rangle_{\partial \mathcal{T}_h},\end{aligned}$$

by the definition of the projection P_M and the second property (*A.5*) of the traces of the local spaces. Then

$$\begin{aligned}\mathbb{T} &= \langle e_{\widehat{u}} - e_u \,,\, (\boldsymbol{\Pi}_V^* \boldsymbol{\theta} - \boldsymbol{\theta}) \cdot \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} - \langle e_u - e_{\widehat{u}} \,,\, \alpha^*(\Pi_W^* \phi - P_M \phi) \rangle_{\partial \mathcal{T}_h} \\ &= \langle e_{\widehat{u}} - e_u \,,\, (\boldsymbol{\Pi}_V^* \boldsymbol{\theta} - \boldsymbol{\theta}) \cdot \boldsymbol{n} + \alpha^*(\Pi_W^* \phi - P_M \phi) \rangle_{\partial \mathcal{T}_h} \\ &= 0,\end{aligned}$$

by the property (*A.5*) of the traces of the local spaces and the orthogonality property (*A.3*) of the projection Π_h^*. This completes the proof. □

Step 4: The estimate for e_u

We are now ready to obtain the estimate of the L^2-norm of e_u and prove Theorem 3.2.2.

We start by taking $\eta = e_u$ in the identity of Lemma 3.5.2 to obtain

$$\begin{aligned}\|e_u\|_{\mathcal{T}_h}^2 &= (\boldsymbol{q} - \boldsymbol{\Pi}_V \boldsymbol{q} \,,\, \boldsymbol{\Pi}_V^* \boldsymbol{\theta})_{\mathcal{T}_h} - (\boldsymbol{e}_q \,,\, \boldsymbol{\theta} - \boldsymbol{\Pi}_V^* \boldsymbol{\theta})_{\mathcal{T}_h} \\ &= (\boldsymbol{q} - \boldsymbol{\Pi}_V \boldsymbol{q} \,,\, \boldsymbol{\theta})_{\mathcal{T}_h} - (\boldsymbol{q} - \boldsymbol{q}_h \,,\, \boldsymbol{\theta} - \boldsymbol{\Pi}_V^* \boldsymbol{\theta})_{\mathcal{T}_h}.\end{aligned}$$

If we now use the orthogonality property (*A.1*) of the projection Π_h and the property (*B.2*) that the space $W(K)$ is not too small, we get

$$\|e_u\|_{\mathcal{T}_h}^2 = (\boldsymbol{q} - \boldsymbol{\Pi}_V \boldsymbol{q} \,,\, \boldsymbol{\theta} - \boldsymbol{P}_0 \boldsymbol{\theta})_{\mathcal{T}_h} - (\boldsymbol{q} - \boldsymbol{q}_h \,,\, \boldsymbol{\theta} - \boldsymbol{\Pi}_V^* \boldsymbol{\theta})_{\mathcal{T}_h},$$

where $\boldsymbol{P}_0$ is the L^2-projection into $\{\boldsymbol{v} \in \boldsymbol{L}^2(\Omega) : \boldsymbol{v}|_K \in \boldsymbol{P}^0(K) \,\forall\, K \in \mathcal{T}_h\}$. By the Cauchy-Schwartz inequality, we get

$$\begin{aligned}\|e_u\|_{\mathcal{T}_h}^2 &\le \|\boldsymbol{q} - \boldsymbol{\Pi}_V \boldsymbol{q}\|_{\mathcal{T}_h} \|\boldsymbol{\theta} - \boldsymbol{P}_0 \boldsymbol{\theta}\|_{\mathcal{T}_h} + \|\boldsymbol{q} - \boldsymbol{q}_h\|_{\mathcal{T}_h} \|\boldsymbol{\theta} - \boldsymbol{\Pi}_V^* \boldsymbol{\theta}\|_{\mathcal{T}_h} \\ &\le (\|\boldsymbol{\theta} - \boldsymbol{P}_0 \boldsymbol{\theta}\|_{\mathcal{T}_h} + 2\, \|\boldsymbol{\theta} - \boldsymbol{\Pi}_V^* \boldsymbol{\theta}\|_{\mathcal{T}_h})\, \|\boldsymbol{q} - \boldsymbol{\Pi}_V \boldsymbol{q}\|_{\mathcal{T}_h} \\ &\le C\, h\, (|\boldsymbol{\theta}|_{1,\mathcal{T}_h} + |\phi|_{2,\mathcal{T}_h})\, \|\boldsymbol{q} - \boldsymbol{\Pi}_V \boldsymbol{q}\|_{\mathcal{T}_h}\end{aligned}$$

by the standard approximation properties of the projection $\boldsymbol{P}_0$ and by the approximation property (*B.1*) of the dual projection Π_h^*. Finally, by the elliptic regularity property (3.5) with $\eta := e_u$, we conclude that

$$\|e_u\|_{\mathcal{T}_h}^2 \le C\, C_{app}^*\, h\, \|e_u\|_{\mathcal{T}_h}\, \|\boldsymbol{q} - \boldsymbol{\Pi}_V \boldsymbol{q}\|_{\mathcal{T}_h},$$

and the estimate follows. This completes the proof of Theorem 3.2.2.

Step 5: The estimate for $u - u^*$

By the Poincaré-Friedrichs inequality, we have that

$$\|u - u_h^*\|_K \le \|\overline{u - u_h^*}\|_K + C\, h_K\, \|\nabla(u - u_h^*)\|_K,$$

where $\overline{w}$ is the average of w over K. But $\overline{u_h^*} = \overline{u_h}$, by the second equation defining u_h^*, (3.4c), and $\overline{u} = \overline{\Pi_W u}$ by Assumptions (*A.2*) and (*C.1*). This implies that

$$\|u - u_h^*\|_K \le \|\Pi_W u - u_h\|_K + C\, h_K\, \|\nabla(u - u_h^*)\|_K.$$

Now, for any $\omega \in \mathcal{W}(K)$, we have that

$$\begin{aligned}
\|\nabla(u - u_h^*)\|_K^2 &= (\nabla(u - u_h^*), \nabla(u - \omega))_K + (\nabla(u - u_h^*), \nabla(\omega - u_h^*))_K \\
&= (\nabla(u - u_h^*), \nabla(u - \omega))_K - (\boldsymbol{q} - \boldsymbol{q}_h, \nabla(\omega - u_h^*))_K,
\end{aligned}$$

by the first equation defining the postprocessing u_h^*, (3.4b). Applying the Cauchy-Schwarz inequality, we obtain that

$$\begin{aligned}
\|\nabla(u - u_h^*)\|_K^2 &\le \|\nabla(u - u_h^*)\|_K \|\nabla(u - \omega)\|_K + \|\boldsymbol{q} - \boldsymbol{q}_h\|_K \|\nabla(\omega - u_h^*)\|_K \\
&\le \|\nabla(u - u_h^*)\|_K (\|\nabla(u - \omega)\|_K + \|\boldsymbol{q} - \boldsymbol{q}_h\|_K) + \|\boldsymbol{q} - \boldsymbol{q}_h\|_K \|\nabla(\omega - u)\|_K,
\end{aligned}$$

and, after simple applications of Young's inequality and some algebraic manipulations, we get that

$$\|\nabla(u - u_h^*)\|_K^2 \le 3(\|\boldsymbol{q} - \boldsymbol{q}_h\|_K^2 + \|\nabla(u - \omega)\|_K^2).$$

This implies that

$$\|u - u_h^*\|_K \le \|\Pi_W u - u_h\|_K + C\, h_K\, (\|\boldsymbol{q} - \boldsymbol{q}_h\|_K + \|\nabla(u - \omega)\|_K),$$

and so,

$$\|u - u_h^*\|_{\mathcal{T}_h} \le \|\Pi_W u - u_h\|_{\mathcal{T}_h} + C\, h\, (\|\boldsymbol{q} - \boldsymbol{q}_h\|_{\mathcal{T}_h} + \|\nabla(u - \omega)\|_{\mathcal{T}_h}).$$

This completes the proof of Theorem 3.2.3. $\square$

3.6 Concluding remarks

We end this paper by discussing some variations on the theoretical results we have proposed in Section 2.

3.6.1 Other postprocessings

There are several ways to define a new approximation $u_h^* \in W^*(K)$ for which Theorem 3.2.3 does hold. The following example is particularly useful when working with the $p-version$ of the method; see also [11] and the references therein. On each element $K \in \mathcal{T}_h$, the postprocessing u_h^* is defined as the element of $W^*(K)$ such that

$$\begin{aligned}
(\nabla u_h^*, \nabla \omega)_K &= -\,(\boldsymbol{q}_h, \nabla \omega)_K \quad && \forall\, \omega \in W^*(K):\ (\omega, \widetilde{w})_k = 0 \text{ for all } \widetilde{w} \in \widetilde{W}(K),\\
(u_h^*, \widetilde{w})_K &= (u_h, \widetilde{w})_K && \forall\, \widetilde{w} \in \widetilde{W}(K).
\end{aligned}$$

3.6.2 Optimal convergence when the local space $W(K)$ is small

When the local space $W(K)$ is small, that is, when it does *not* satisfy *Assumption (B.2)*, the superconvergence of the projection of the error in the scalar variable, $\Pi_W u - u_h$, is not guaranteed, and in general it does not take place. Examples of these methods are the HDG_0 and the BDM_1 methods for simplexes.

However, in this case we can still obtain the optimal order of convergence of $\Pi_W u - u_h$. Indeed, by the identity of Lemma 3.5.2 obtained by duality, we have

$$\begin{aligned}
\|e_u\|_{\mathcal{T}_h}^2 &= (\boldsymbol{q} - \boldsymbol{\Pi}_V \boldsymbol{q}\,,\, \boldsymbol{\Pi}_V^* \boldsymbol{\theta})_{\mathcal{T}_h} - (\boldsymbol{e}_q\,,\, \boldsymbol{\theta} - \boldsymbol{\Pi}_V^* \boldsymbol{\theta})_{\mathcal{T}_h}\\
&\le \|\boldsymbol{q} - \boldsymbol{\Pi}_V \boldsymbol{q}\|_{\mathcal{T}_h} \|\boldsymbol{\Pi}_V^* \boldsymbol{\theta}\|_{\mathcal{T}_h} + \|\boldsymbol{e}_q\|_{\mathcal{T}_h} \|\boldsymbol{\theta} - \boldsymbol{\Pi}_V^* \boldsymbol{\theta}\|_{\mathcal{T}_h}\\
&\le \|\boldsymbol{q} - \boldsymbol{\Pi}_V \boldsymbol{q}\|_{\mathcal{T}_h} (\|\boldsymbol{\Pi}_V^* \boldsymbol{\theta}\|_{\mathcal{T}_h} + \|\boldsymbol{\theta} - \boldsymbol{\Pi}_V^* \boldsymbol{\theta}\|_{\mathcal{T}_h}) \qquad \text{by Theorem 3.2.1,}\\
&\le \|\boldsymbol{q} - \boldsymbol{\Pi}_V \boldsymbol{q}\|_{\mathcal{T}_h} (\|\theta\|_{\mathcal{T}_h} + 2\, C_{app}\, h\, (|\phi|_{H^2(\mathcal{T}_h)} + |\boldsymbol{\theta}\|_{H^1(\mathcal{T}_h)})
\end{aligned}$$

by *Assumption (B.1)*. Finally, after a simple application of the elliptic regularity inequality (3.5) with $\eta := e_u$, we get that

$$\|e_u\|_{\mathcal{T}_h} \le C\, \|\boldsymbol{q} - \boldsymbol{\Pi}_V \boldsymbol{q}\|_{\mathcal{T}_h}.$$

Thus, even though the *Assumption (B.2)* does not hold. The convergence in the scalar variable can be optimal.

3.6.3 Superconvergence when the local space $W(K)$ is small

Next, we show that it is still possible to obtain superconvergence of the projection of the error in the scalar variable when the local space $W(K)$ is small, that is, when it does *not* satisfy *Assumption (B.2)*. We do this for the RT_0 method, which is the only method for which this is known to happen. We begin by noting that, by the identity of Lemma 3.5.2, we have

$$\begin{aligned}
\|e_u\|^2_{\mathcal{T}_h} &= (\boldsymbol{q} - \boldsymbol{\Pi}_V\boldsymbol{q}\,,\,\boldsymbol{\Pi}^*_V\boldsymbol{\theta})_{\mathcal{T}_h} - (\boldsymbol{e}_q\,,\,\boldsymbol{\theta} - \boldsymbol{\Pi}^*_V\boldsymbol{\theta})_{\mathcal{T}_h}\\
&= (\boldsymbol{q} - \boldsymbol{\Pi}_V\boldsymbol{q}\,,\,\boldsymbol{\theta})_{\mathcal{T}_h} - (\boldsymbol{q} - \boldsymbol{q}_h\,,\,\boldsymbol{\theta} - \boldsymbol{\Pi}^*_V\boldsymbol{\theta})_{\mathcal{T}_h}\\
&= -\,(\boldsymbol{q} - \boldsymbol{\Pi}_V\boldsymbol{q}\,,\,\nabla\phi)_{\mathcal{T}_h} - (\boldsymbol{q} - \boldsymbol{q}_h\,,\,\boldsymbol{\theta} - \boldsymbol{\Pi}^*_V\boldsymbol{\theta})_{\mathcal{T}_h} \qquad \text{by (3.6a),}\\
&= -\,(\boldsymbol{q} - \boldsymbol{\Pi}_V\boldsymbol{q}\,,\,\nabla(\phi - \overline{\phi}))_{\mathcal{T}_h} - (\boldsymbol{q} - \boldsymbol{q}_h\,,\,\boldsymbol{\theta} - \boldsymbol{\Pi}^*_V\boldsymbol{\theta})_{\mathcal{T}_h}\\
&= (\nabla\cdot(\boldsymbol{q} - \boldsymbol{\Pi}_V\boldsymbol{q})\,,\,\phi - \overline{\phi})_{\mathcal{T}_h} - \langle(\boldsymbol{q} - \boldsymbol{\Pi}_V\boldsymbol{q})\cdot\boldsymbol{n}, \phi - \overline{\phi}\rangle_{\partial\Omega_h}\\
&\quad - (\boldsymbol{q} - \boldsymbol{q}_h\,,\,\boldsymbol{\theta} - \boldsymbol{\Pi}^*_V\boldsymbol{\theta})_{\mathcal{T}_h}\\
&= (\nabla\cdot(\boldsymbol{q} - \boldsymbol{\Pi}_V\boldsymbol{q})\,,\,\phi - \overline{\phi})_{\mathcal{T}_h} - \langle(\boldsymbol{q} - \boldsymbol{\Pi}_V\boldsymbol{q})\cdot\boldsymbol{n}, \phi - \overline{\phi}\rangle_{\partial\Omega_h\setminus\partial\Omega}\\
&\quad - (\boldsymbol{q} - \boldsymbol{q}_h\,,\,\boldsymbol{\theta} - \boldsymbol{\Pi}^*_V\boldsymbol{\theta})_{\mathcal{T}_h}
\end{aligned}$$

by the boundary condition of the dual problem (3.6c). For the RT_0 method, we can write

$$\|e_u\|^2_{\mathcal{T}_h} = (\nabla\cdot\boldsymbol{q} - \overline{\nabla\cdot\boldsymbol{q}}\,,\,\phi - \overline{\phi})_{\mathcal{T}_h} - (\boldsymbol{q} - \boldsymbol{q}_h\,,\,\boldsymbol{\theta} - \boldsymbol{\Pi}^*_V\boldsymbol{\theta})_{\mathcal{T}_h},$$

and, proceeding in the previous subsection, we can obtain that

$$\|e_u\|_{\mathcal{T}_h} \le C\,h^2\,|\nabla\cdot\boldsymbol{q}|_{H^1(\Omega)} + C\,h\,\|\boldsymbol{q} - \boldsymbol{\Pi}_V\boldsymbol{q}\|_{\mathcal{T}_h}.$$

Superconvergence of order two is thus achieved for the projection of the error e_u.

3.6.4 Other formulas for the numerical trace of the flux

The hybridizable DG method based on the use of the so-called interior penalty (IP) method on each element, see [6], does not use the formula for the numerical trace of the flux (3.3). Instead it uses the formula

$$\widehat{\boldsymbol{q}}_h\cdot\boldsymbol{n} = -\nabla u_h\cdot\boldsymbol{n} + \alpha(u_h - \widehat{u}_h) \quad \text{on} \quad \partial\mathcal{T}_h.$$

The application of our approach to this method remains open. However, let us point out that since when $\alpha = \tau$, this method is well defined provided τ is of order $1/h$; see [6]. As a consequence, it seems very unlikely that optimal convergence will be attained for the approximate flux.

3.6.5 Conclusion

The projection-approach we have presented here provides a *simple, unified a priori error analysis* of a large class of finite elements methods including mixed and HDG methods. It provides *sufficient conditions* on the different local spaces and by the local stabilization operator α that guarantee the *superconvergence* of the postprocessing u_h^*. In other words, it gives us guidelines for the devising of new superconvergent methods for elliptic problems.

We have also proposed a template to construct such methods and have shown that all previously known mixed methods and the $\mathbf{HDG}_k$ methods for simplexes with the local stabilization operator used in [42, 43, 6] fit in it. We have also shown how to use it to uncover several superconvergent $\mathbf{HDG}_{[k]}$ methods for squares, cubes and prisms; they are the only DG methods using those elements known to be superconvergent). We have also used the template to uncover what seems to be the smallest superconvergent mixed method, on squares and cubes, containing the tensor-product space $\boldsymbol{Q}^k(K) \times Q^k(K)$.

The extension of this approach to isoparametric elements constitutes the subject of Chapter 4.

Chapter 4

Superconvergent HDG methods on isoparametric elements for second order elliptic equations

4.1 Definition of the HDG methods

In this section, we define the HDG methods we are going to work with for the diffusion equation (3.1). We begin by describing the meshes we are going to use. We then define the general form of the corresponding finite element spaces. We finally end with by providing the weak formulation defining the HDG methods.

4.1.1 Geometry of the mesh

Next, we describe the meshes we are going to use. To do that, we first introduce the concept of generalized cell. Next, we define a C^0-compatible mesh and then the so-called ℓ-regular meshes, which are the meshes we are going to work with. Finally, we propose a way of generating ℓ-regular meshes.

Reference cells and curved cells

We denote by $\underline{K}$ the reference cell in $\mathbb{R}^d$. When $d = 3$, this closed set is the standard unit tetrahedron or a unit cube. When $d = 2$, it is the unit triangle or the unit square.

We denote by $\triangle_m(\underline{K})$ the collection of all m-dimensional subcells of $\underline{K}$. They are all the faces of $\underline{K}$ when $m = 2$, are all the edges of $\underline{K}$ when $m = 1$.

Definition A closed subset K of $\mathbb{R}^d$ is a generalized d-dimensional cell if there is a C^1-diffeomorphism G_K from the reference cell $\underline{K}$ to K such that $G_K \in \mathcal{C}^\infty(\underline{K})$.

In particular, we call K a generalized hexahedron when $\underline{K}$ is a cube, and a generalized tetrahedron when $\underline{K}$ is a tetrahedron. We denote h_K by the diameter of K. We also denote by $\triangle_m(K)$ the collection of all m-dimensional subcells of K, which are exactly $G_K(\triangle_m(\underline{K}))$. Note that all points x in K are of the form $x = G_K(\underline{x})$ where $\underline{x}$ lies in $\underline{K}$.

C^0-compatible mesh

We denote by Ω_h the finite collection of generalized cells in $\mathbb{R}^d$ such that for any two different generalized cells $K, K' \in \Omega_h$, either $K \cap K' = \emptyset$ or $K \cap K' \in \triangle_m(K) \cap \triangle_m(K')$ for some $0 \le m \le d-1$. Here, the parameter h is the maximum of the diameters h_K of the cells K in Ω_h.

We denote by $\triangle_{d-1}(\Omega_h)$ the collection of $\triangle_{d-1}(K)$ for all cells K in Ω_h. Notice that for any $F \in \triangle_{d-1}(\Omega_h)$, either $F = K \cap K'$ with $K, K' \in \Omega_h$ or $F \subset \partial\Omega$ where Ω is an open subset in $\mathbb{R}^d$ such that $\overline{\Omega} = \cup_{K \in \Omega_h} K$.

Definition We say that Ω_h is a C^0-compatible mesh if, for any two subcells $\underline{F} \in \triangle_{d-1}(\underline{K})$ and $\underline{F}' \in \triangle_{d-1}(\underline{K}')$, where $K, K' \in \Omega_h$, such that $G_K(\underline{F}) = G_{K'}(\underline{F}')$, there is an affine mapping $\mathcal{R} : \underline{F} \to \underline{F}'$ satisfying

$$G_K|_{\underline{F}} = G_{K'}|_{\underline{F}'} \circ \mathcal{R}. \tag{4.1}$$

We call K a element of Ω_h. And, we call $F \in \triangle_{d-1}(\Omega_h)$ a face in Ω_h. We also define $\mathcal{E}_h := \triangle_{d-1}(\Omega_h)$.

The ℓ-regular meshes

Now, we are ready to give the description of meshes we are going to use in this paper.

Definition (ℓ-regular meshes) Let $\{\Omega_h\}_{h\in \mathrm{I}}$ be a family of C^0-compatible meshes. We call $\{\Omega_h\}_{h\in \mathrm{I}}$ a family of ℓ-regular meshes if for any $h \in \mathrm{I}$ and any $K \in \Omega_h$,

$$\sup_{\underline{x}\in\underline{K}} \|(DG_K(\underline{x}))^{-1}\| \le c_1 h_K^{-1}, \qquad \sup_{\underline{x}\in\underline{K}} \|D^i G_K(\underline{x})\| \le c_i h_K^{\ell} \quad i = 1, \dots, \ell,$$

where $c_1, \dots, c_\ell$ are positive constants independent of h and of K.

Throughout this paper, we assume that the domain Ω admits a family of ℓ-regular meshes $\{\Omega_h\}_{h\in \mathrm{I}}$ such that $\overline{\Omega} = \cup_{K\in\Omega_h} K$ for any $h \in \mathrm{I}$. As usual, we can always pick an h in I arbitrarily close to zero.

Isoparametric refinement

Next, we present a way of generating a family of ℓ-regular meshes for Ω. We begin by obtaining a C^0-compatible mesh for Ω, Ω_{h_0}, and by setting $\boldsymbol{G}^0 := \{\boldsymbol{G}^0_K,\ \forall K \in \Omega_{h_0}\}$. To obtain a finer mesh Ω_{h_1}, we first divide the reference element $\underline{K}$ uniformly into elements $\underline{K}'$. Then we refine the actual element K via the mapping $\boldsymbol{G}^0_K$, that is $K' = \boldsymbol{G}^0_K(\underline{K}')$. The remaining meshes are obtained by repeating this process. It is not difficult to verify that the family of meshes obtained in this manner is ℓ-regular if we have that

$$\sup_{\underline{x}\in\underline{K}} \|(DG^0_K(\underline{x}))^{-1}\| \le c_1 h_K^{-1}, \qquad \sup_{\underline{x}\in\underline{K}} \|D^i G^0_K(\underline{x})\| \le c_\ell h_K^{i} \quad i = 1, \dots, \ell,$$

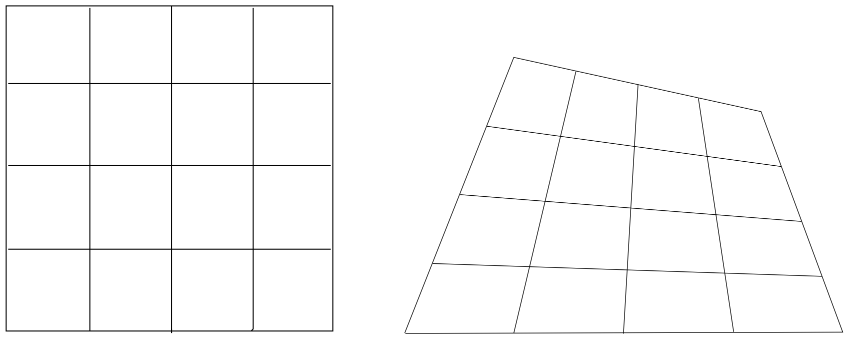
for any $K \in \Omega_{h_0}$, See Fig. 4.1 for an illustration of this process for meshes of quadrilateral elements.

4.1.2 The finite element spaces

Here, we define the finite element spaces in which we are going to seek our approximate solutions. Those spaces are defined by suitable transformations of the spaces for the reference element. We begin by defining the local spaces. We then introduce the transformations. Finally, we define the global spaces.

Notation

We denote by $\boldsymbol{V}(\underline{K})$ a finite dimensional vector-valued function space on $\underline{K}$. We denote by $W(\underline{K})$ a finite dimensional scalar function space on $\underline{K}$. For every face $\underline{F}$ of $\underline{K}$, we denote by $M(\underline{F})$ a finite dimensional scalar function space on $\underline{F}$. Finally, we denote by $\underline{\alpha}$ what we could call the *reference stabilization* function on $\partial\underline{K}$.



(a) Refinement on reference element

(b) Refinement on physical element

Figure 4.1: Isoparametric refinement on reference and physical element

Weighted Piola transform and scalar transform

Let K be a generalized n-dimensional cell. We denote by $G_K : \underline{K} \to K$ the C^1-diffeomorphism given in Definition 4.1.1. Let $\underline{\boldsymbol{q}}$ be a vector field on $\underline{K}$, then we define a vector field $\boldsymbol{q}$ on K by

$$\boldsymbol{q}(\boldsymbol{x}) = \frac{h_K^{n-1} DG_K(\underline{\boldsymbol{x}})}{\det DG_K(\underline{\boldsymbol{x}})} \underline{\boldsymbol{q}}(\underline{x}) \quad \forall \underline{\boldsymbol{x}} \in \underline{K}, \tag{4.2a}$$

where $\boldsymbol{x} = G_K(\underline{\boldsymbol{x}})$. We call this transform the weighted Piola transform. If $\underline{u}$ is a scalar function on $\underline{K}$, we define a scalar function u on K by

$$u(\boldsymbol{x}) = \underline{u}(\underline{\boldsymbol{x}}) \quad \forall \underline{\boldsymbol{x}} \in \underline{K}, \tag{4.2b}$$

where $\boldsymbol{x} = G_K(\underline{\boldsymbol{x}})$. We call this transform the scalar transform.

Let us recall that, for any subcell $F := G_K(\underline{F})$, $\underline{F} \in \triangle_{n-1}(\underline{K})$, its outward unit normal at the point $\boldsymbol{x} = G_K(\underline{\boldsymbol{x}})$ is $\boldsymbol{n}(\boldsymbol{x}) = DG_K^{-\top}(\underline{\boldsymbol{x}})\, \underline{\boldsymbol{n}} / \| DG_K^{-\top}(\underline{\boldsymbol{x}})\, \underline{\boldsymbol{n}} \|$, and we have that

$$\boldsymbol{q} \cdot \boldsymbol{n}(\boldsymbol{x}) = \frac{h_K^{n-1}}{J_F} \underline{\boldsymbol{q}} \cdot \underline{\boldsymbol{n}}(\underline{\boldsymbol{x}}) \boldsymbol{x} = G_K(\underline{\boldsymbol{x}}) \quad \forall \underline{\boldsymbol{x}} \in \underline{F}, \tag{4.3}$$

where $J_F(\underline{\boldsymbol{x}}) = \sqrt{\det[DG_K|_{\underline{F}}^{\top}(\underline{\boldsymbol{x}}) DG_K|_{\underline{F}}(\underline{\boldsymbol{x}})]}$, because it is well known that we have the identity $J_F(\underline{\boldsymbol{x}}) = \| DG_K^{-\top}(\underline{\boldsymbol{x}})\, \underline{\boldsymbol{n}} \| \det DG_K(\underline{\boldsymbol{x}})$.

Finite element spaces on C^0-compatible meshes

For any generalized cell K the C^0-compatible mesh Ω_h, we define our finite element spaces on K as follows.

Definition

$$\boldsymbol{V}(K)(\boldsymbol{x}) = \frac{h_K^{n-1} DG_K}{\det DG_K} \boldsymbol{V}(\underline{K})(\underline{\boldsymbol{x}}), \tag{4.4a}$$

$$W(K)(\boldsymbol{x}) = W(\underline{K})(\underline{\boldsymbol{x}}), \tag{4.4b}$$

$$M_K(F)(\boldsymbol{x}) = M(\underline{F})(\underline{\boldsymbol{x}}) \qquad \forall F \in \triangle_{n-1}(K), \tag{4.4c}$$

$$N_K(F)(\boldsymbol{x}) = \frac{h_K^{n-1}}{J_F} M(\underline{F})(\underline{\boldsymbol{x}}) \qquad \forall F \in \triangle_{n-1}(K), \tag{4.4d}$$

$$\alpha_K|_F(\boldsymbol{x}) = \frac{h_K^{n-1}}{J_F} \underline{\alpha}|_{\underline{F}}(\underline{\boldsymbol{x}}) \qquad \forall F \in \triangle_{n-1}(K). \tag{4.4e}$$

Here, $F = G_K(\underline{F})$ for any face $\underline{F}$ of $\underline{K}$.

Note that it is highly desirable that, when F is a subcell shared by the generalized cells K and K', we have that the spaces $M_K(F)$ and $N_K(F)$ are the *same* as the spaces $M_{K'}(F)$ and $N_{K'}(F)$, respectively. The following assumption is introduced to ensure this property.

Assumption M:

Let $\underline{F}, \underline{F}' \in \triangle_{n-1}(\underline{K})$ and $\mathcal{R} : \underline{F} \to \underline{F}'$ be any affine mapping. Then

$$M(\underline{F}') \circ \mathcal{R} = M(\underline{F}).$$

Indeed, we have the following result.

Theorem 4.1.1. *Let Ω_h be a C^0-compatible mesh and let the Assumption M be satisfied for any subcell F shared by the generalized cells K and K' of Ω_h. Then we have that $M_K(F) = M_{K'}(F)$ and that $N_K(F) = N_{K'}(F)$.*

Proof. We only prove that $N_K(F) = N_{K'}(F)$ since the other equality is similar and simpler. In fact, by symmetry, we only have to show that $N_K(F) \subset N_{K'}(F)$.

Let λ be an arbitrary element of $N_K(F)$. Then, by definition of $N_K(F)$, there is an element $\underline{\lambda}$ in $M(\underline{F})$ such that

$$\lambda(\boldsymbol{x}) = \frac{h_K^{n-1}}{J_F(\underline{\boldsymbol{x}})}\underline{\lambda}(\underline{\boldsymbol{x}}) \quad \boldsymbol{x} = G_K(\underline{\boldsymbol{x}}) \quad \forall \underline{\boldsymbol{x}} \in \underline{F}.$$

Now, let $\underline{F}'$ be the subcell of the reference element $\underline{K}$ such that $F = G_{K'}(\underline{F}')$. Since Ω_h be a C^0-compatible mesh, there is an affine mapping $\mathcal{R} : \underline{F} \to \underline{F}'$ such that $G_K|_{\underline{F}} = G_{K'}|_{\underline{F}'} \circ \mathcal{R}$. This implies that $J_F(\underline{\boldsymbol{x}}) = |detD\mathcal{R}|\, J_{F'}(\underline{\boldsymbol{x}}')$ where $\underline{\boldsymbol{x}}' = \mathcal{R}(\underline{\boldsymbol{x}})$, and so

$$\lambda(\boldsymbol{x}) = \frac{h_{K'}^{n-1}}{J_{F'}(\underline{\boldsymbol{x}}')}\underline{\lambda}'(\underline{\boldsymbol{x}}') \quad \boldsymbol{x} = G_{K'}(\underline{\boldsymbol{x}}') \quad \forall \underline{\boldsymbol{x}}' \in \underline{F}',$$

where

$$\underline{\lambda}'(\underline{\boldsymbol{x}}') := \frac{h_K^{n-1}}{h_{K'}^{n-1}} \frac{1}{|detD\mathcal{R}|}\underline{\lambda}(\mathcal{R}^{-1}(\underline{\boldsymbol{x}}')) \quad \forall \underline{\boldsymbol{x}}' \in \underline{F}'.$$

Since $\mathcal{R}$ is an affine mapping, $|detD\mathcal{R}|$ is a constant, and $\underline{\lambda} \circ \mathcal{R}^{-1}$ lies in $M(\underline{F}')$, by the *Assumption M*. We thus see that λ belongs to $N_{K'}(F)$. This completes the proof. $\square$

In what follows, whenever *Assumption M* holds, we are going to write $M(F)$ and $N(F)$ instead of $M_K(F)$ and $N_K(F)$, respectively. We can now define the finite element spaces on the whole mesh Ω_h.

Definition For any C^0-compatible mesh Ω_h, we set

$$\boldsymbol{V}_h = \{\boldsymbol{v} \in \boldsymbol{L}^2(\mathcal{T}_h) : \boldsymbol{v}|_K \in \boldsymbol{V}(K),\ K \in \mathcal{T}_h\}, \tag{4.5a}$$

$$W_h = \{w \in L^2(\mathcal{T}_h) : w|_K \in W(K),\ K \in \mathcal{T}_h\}, \tag{4.5b}$$

$$\alpha|_{\partial K} = \alpha_K \qquad \forall K \in \Omega_h. \tag{4.5c}$$

If, moreover, *Assumption M* is satisfied, we set

$$M_h = \{\mu \in L^2(\mathcal{E}_h) : \mu|_F \in M(F),\ F \in \mathcal{E}_h\}, \tag{4.5d}$$

$$N_h = \{\mu \in L^2(\mathcal{E}_h) : \mu|_F \in N(F),\ F \in \mathcal{E}_h\}. \tag{4.5e}$$

Recall that in [3], we only need to work with the space M_h for functions defined on $\mathcal{E}_h$. The introduction of the additional space N_h of functions also defined on $\mathcal{E}_h$ is needed only to be able to handle meshes using generalized cells.

4.1.3 The formulation of the HDG method on a family of ℓ-regular meshes

Let $\{\Omega_h\}_{h\in I}$ be a family of ℓ-regular meshes introduced in Definition 4.1.1. The methods we are interested in seek an approximation to $(u, \boldsymbol{q}, u|_{\mathcal{E}_h})$, $(u_h, \boldsymbol{q}_h, \widehat{u}_h)$, in the finite element space $W_h \times \boldsymbol{V}_h \times M_h$, where

$$-(u_h\,,\, \nabla\cdot \boldsymbol{v})_{\mathcal{T}_h} + (\boldsymbol{q}_h\,,\, \boldsymbol{v})_{\mathcal{T}_h} + \langle \widehat{u}_h\,,\, \boldsymbol{v}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} = 0, \tag{4.6a}$$

$$-(\boldsymbol{q}_h\,,\, \nabla w)_{\mathcal{T}_h} + \langle \widehat{\boldsymbol{q}}_h\cdot\boldsymbol{n}\,,\, w\rangle_{\partial\mathcal{T}_h} = (f\,,\, w)_{\mathcal{T}_h}, \tag{4.6b}$$

$$\langle \widehat{\boldsymbol{q}}_h\cdot\boldsymbol{n}, \mu\rangle_{\partial\mathcal{T}_h\setminus\partial\Omega} = 0, \tag{4.6c}$$

$$\langle \widehat{u}_h, \lambda\rangle_{\partial\Omega} = \langle g, \lambda\rangle_{\partial\Omega}, \tag{4.6d}$$

for all $(w, \boldsymbol{v}, \mu, \lambda) \in W_h \times \boldsymbol{V}_h \times M_h|_{\partial\mathcal{T}_h\setminus\partial\Omega} \times N_h|_{\partial\Omega}$. Here we write $(\eta\,,\,\zeta)_{\mathcal{T}_h} := \sum_{K\in\mathcal{T}_h}(\eta,\zeta)_K$, where $(\eta,\zeta)_D$ denotes the integral of $\eta\zeta$ over the domain $D\subset\mathbb{R}^n$. We also write $(\eta\,,\,\zeta)_{\mathcal{T}_h} := \sum_{K\in\mathcal{T}_h}\langle\eta\,,\,\zeta\rangle_{\partial K}$, where $\langle\eta\,,\,\zeta\rangle_D$ denotes the integral of $\eta\zeta$ over the domain D which is a $(n-1)$-dimensional surface in $\mathbb{R}^n$.

The definition of the method is completed with the definition of the normal component of the numerical trace:

$$\widehat{\boldsymbol{q}}_h\cdot\boldsymbol{n} = \boldsymbol{q}_h\cdot\boldsymbol{n} + \alpha(u_h - \widehat{u}_h) \in N_h \quad \text{on} \quad \partial\mathcal{T}_h := \{\partial K : K \subset \mathcal{T}_h\}. \tag{4.7}$$

4.1.4 Definition of u_h^*

To end this section, we introduce the postprocessed approximation u_h^*. On each element K, we first define the postprocessed u_h^* on the reference element $\underline{K}$ in the space $W^*(\underline{K})$ such that

$$(\underline{\nabla u_h^*}, \underline{\nabla\omega})_{\underline{K}} = -\,(\mathrm{M}\underline{\boldsymbol{q}}_h, \underline{\nabla\omega})_{\underline{K}} \quad \forall\, \underline{\omega}\in W^*(\underline{K}), \tag{4.8a}$$

$$(\underline{u}_h^*, 1)_{\underline{K}} = (\underline{u}_h, 1)_{\underline{K}}, \tag{4.8b}$$

where $\mathrm{M} := J_K^{-1}h_K^{n-1}D\boldsymbol{G}_K^T D\boldsymbol{G}_K$ and where $\underline{\boldsymbol{q}}_h$ and $\underline{u}_h$ are related to $\boldsymbol{q}_h$ and u_h, respectively, by the weighted Piola transform (4.2a) and the scalar transform (4.2b), respectively.

Let us point out that the motivation for the use of the operator M in the first equation defining $\underline{u}_h^*$ is that we have

$$\underline{\nabla u} = -J_K^{-1}h_K^{n-1}D\boldsymbol{G}_K^T D\boldsymbol{G}_K\underline{q} = -\mathrm{M}\underline{q}.$$

where $\underline{\boldsymbol{q}}$ and $\underline{u}$ are related to $\boldsymbol{q}$ and u, respectively, by the weighted Piola transform (4.2a) and the scalar transform (4.2b), respectively.

4.2 A priori estimates of the projection of the errors

In this section we show how an a priori error analysis of the HDG methods can be *reduced* to the verification of a few conditions on the spaces and on some properties of an associated, auxiliary projection $\underline{\Pi}$ defined on the reference element $\underline{K}$.

4.2.1 The auxiliary projection Π_h

The main idea of our error analysis is to estimate a projection of the errors $\Pi_h(\boldsymbol{q} - \boldsymbol{q}_h, u - u_h)$ and then deduce bounds of the $L^2(\Omega)$-norm of the errors $\boldsymbol{q} - \boldsymbol{q}_h$, $u - u_h$ and $u - u_h^*$. This projection is defined in term of the projection $\underline{\Pi}$ defined on the reference element $\underline{K}$ as follows.

For any $(\boldsymbol{q}, u) \in \boldsymbol{H}^1(\Omega) \times H^1(\Omega)$, we set

$$\Pi_V \boldsymbol{q}|_K(\boldsymbol{x}) = \frac{h_K^{n-1} DG_K(\underline{\boldsymbol{x}})}{\det DG_K(\underline{\boldsymbol{x}})} \underline{\Pi}_V \underline{\boldsymbol{q}}(\underline{\boldsymbol{x}}) \qquad \forall \underline{\boldsymbol{x}} \in \underline{K} \quad \forall K \in \Omega_h, \tag{4.9a}$$

$$\Pi_W u|_K(\boldsymbol{x}) = \underline{\Pi}_W \underline{u}(\underline{\boldsymbol{x}}) \qquad \forall \underline{\boldsymbol{x}} \in \underline{K} \quad \forall K \in \Omega_h, \tag{4.9b}$$

where $\boldsymbol{x} = G_K(\underline{\boldsymbol{x}})$ and, once again, $\underline{\boldsymbol{q}}$ and $\underline{u}$ are related to $\boldsymbol{q}$ and u, respectively, by the weighted Piola transform (4.2a) and the scalar transform (4.2b), respectively.

4.2.2 Estimate of $\boldsymbol{q} - \boldsymbol{q}_h$

Our first result gives an estimate of the projection of the error $\boldsymbol{\Pi}_V \boldsymbol{q} - \boldsymbol{q}_h$ solely in terms of the approximation error of the projection $\boldsymbol{q} - \boldsymbol{\Pi}_V \boldsymbol{q}$. To state it, we need to describe our assumptions on the projection $\underline{\Pi}$ and on the finite element spaces $\boldsymbol{V}(\underline{K})$, $W(\underline{K})$ and $M(\underline{F})$. Here $\underline{K}$ is the reference element and $\underline{F} \in \triangle_{n-1}(\underline{K})$.

Assumptions A:

- *Orthogonality properties of* $\underline{\Pi}$. On the reference element $\underline{K}$, there exist a projection $\underline{\Pi}(\underline{\boldsymbol{q}}, \underline{u}) = (\underline{\Pi}_V \underline{\boldsymbol{q}}, \underline{\Pi}_W \underline{u}) \in \boldsymbol{V}(\underline{K}) \times W(\underline{K})$ satisfying the following properties:

$(A.1)$ $(\underline{\Pi}_V \underline{\boldsymbol{q}}, \underline{\boldsymbol{v}})_{\underline{K}} = (\underline{\boldsymbol{q}}, \underline{\boldsymbol{v}})_{\underline{K}}$ for all $\underline{\boldsymbol{v}} \in \underline{\nabla} W(\underline{K})$,

$(A.2)$ $(\underline{\Pi}_W \underline{u}, \underline{w})_{\underline{K}} = (\underline{u}, \underline{w})_{\underline{K}}$ for all $\underline{w} \in \underline{\nabla} \cdot \boldsymbol{V}(\underline{K})$,

$(A.3)$ For all faces $\underline{F}$ of the reference element $\underline{K}$,

$$\langle \underline{\Pi}_V \underline{\boldsymbol{q}} \cdot \underline{\boldsymbol{n}} + \underline{\alpha}(\underline{\Pi}_W \underline{u}), \underline{\mu} \rangle_{\underline{F}} = \langle \underline{\boldsymbol{q}} \cdot \underline{\boldsymbol{n}} + \underline{\alpha}(\underline{P}_M \underline{u}), \underline{\mu} \rangle_{\underline{F}},$$

for all $\underline{\mu} \in M(\underline{F})$.

Here $\underline{\alpha}$ is the reference stablization function. Here the local projection $\underline{P}_M$ is standard L_2 orthogonal projection from $L^2(\underline{F})$ onto $M(\underline{F})$.

Next, we need to assume suitable relations between the traces on the faces $\underline{F}$ of the local spaces $\boldsymbol{V}(\underline{K})$ and $W(\underline{K})$ with the local space $M(\underline{F})$.

• *Properties of the traces of the spaces.* For the reference element $\underline{K}$, and for any of its faces $\underline{F}$,

$(A.4)$ $\boldsymbol{V}(\underline{K}) \cdot \underline{\boldsymbol{n}}|_{\underline{F}} \subset M(\underline{F})$,

$(A.5)$ $W(\underline{K})|_{\underline{F}} \subset M(\underline{F})$.

Here $\boldsymbol{V}(\underline{K}) \cdot \underline{\boldsymbol{n}}|_{\underline{F}}$ denotes the space of the *traces* of normal components of functions of $\boldsymbol{V}(\underline{K})$ on the face $\underline{F}$ of $\underline{K}$. Similarly, $W(\underline{K})|_{\underline{F}}$ denotes the space of *traces* of functions of $W(\underline{K})$ on the face $\underline{F}$.

Finally, we need a simple assumption reflecting the stabilizing role of the linear function $\underline{\alpha}$.

• *The semi-positivity property of* $\underline{\alpha}$. For the reference element $\underline{K}$ and any of its faces $\underline{F}$,

$(A.6)$ $\langle \underline{\alpha}(\underline{\mu}) \, , \, \underline{\mu} \rangle_{\underline{F}} \geq 0$ for all $\underline{\mu} \in M(\underline{F})$.

We are now ready to state our first result. In what follows, we use $\|\cdot\|_{k,D}$, $|\cdot|_{k,D}$ to denote the standard norm and seminorm on any Sobolev space $H^k(D)$, respectively. For simplicity, we use $\|\cdot\|_D$ to denote the $L^2(D)-$norm on any D.

Theorem 4.2.1. *Suppose that the Assumptions A are satisfied. Then we have*

$$\|\boldsymbol{\Pi}_V \boldsymbol{q} - \boldsymbol{q}_h\|_{\mathcal{T}_h} \leq \|\boldsymbol{q} - \boldsymbol{\Pi}_V \boldsymbol{q}\|_{\mathcal{T}_h}.$$

Here $\boldsymbol{\Pi}_V$ is defined by (4.9a).

Note that, since this implies that

$$\|\boldsymbol{q} - \boldsymbol{q}_h\|_{\mathcal{T}_h} \leq 2\|\boldsymbol{q} - \boldsymbol{\Pi}_V \boldsymbol{q}\|_{\mathcal{T}_h},$$

the quality of the approximation $\boldsymbol{q}_h$ only depends on the approximation properties of the first component of the projection.

4.2.3 Estimate of $u - u_h$

Our next result shows that $\Pi_W u - u_h$ can *also* be controlled solely in terms of the approximation error of the projection $\boldsymbol{q} - \boldsymbol{\Pi}_V \boldsymbol{q}$. Here Π_W is defined in (4.9b).

It is valid under a typical elliptic regularity property we state next. We assume that, for any given $\eta \in L^2(\Omega)$, we have

$$\|\phi\|_{2,\Omega} + \|\boldsymbol{\theta}\|_{1,\Omega} \leq C\|\eta\|_{\Omega}, \tag{4.10}$$

where C only depends on the domain Ω, and $(\boldsymbol{\theta}, \phi)$ is the solution of the *dual* problem:

$$\boldsymbol{\theta} + \nabla\phi = 0 \quad \text{in } \Omega, \tag{4.11a}$$

$$\nabla \cdot \boldsymbol{\theta} = \eta \quad \text{in } \Omega, \tag{4.11b}$$

$$\phi = 0 \quad \text{on } \partial\Omega. \tag{4.11c}$$

We also need two additional assumptions.

<u>*Assumptions B*</u>:

The first is an approximation property of a projection $\underline{\Pi}^*(\underline{\boldsymbol{q}}, \underline{u}) = (\underline{\Pi}_V^* \underline{\boldsymbol{q}}, \underline{\Pi}_W^* \underline{u})$ which satisfies the assumptions $(A.1)$, $(A.2)$ and $(A.3)$ where the reference stabilization function $\underline{\alpha}(\cdot)$ is *replaced* by its dual $\underline{\alpha}^*(\cdot)$, that is, by the linear function defined by

$$\langle \underline{\alpha}(\underline{p}) \,,\, \underline{q} \rangle_{\partial \underline{K}} = \langle \underline{p} \,,\, \underline{\alpha}^*(\underline{q}) \rangle_{\partial \underline{K}} \quad \text{for all} \quad \underline{p}, \underline{q} \in L^2(\partial \underline{K}).$$

For each element K and any $(\boldsymbol{q}, u) \in \boldsymbol{H}^1(K) \times H^2(K)$,

$$(B.1) \quad \|\underline{\Pi}_V^* \underline{\boldsymbol{q}} - \underline{\boldsymbol{q}}\|_{\underline{K}} \le C^*_{app} \, (|\underline{u}|_{1,\underline{K}} + |\underline{\boldsymbol{q}}|_{1,\underline{K}}).$$

• *The reference space* $\boldsymbol{W}(\underline{K})$ *is not too small.* On the reference element $\underline{K}$, we have that

$$(B.2) \quad \boldsymbol{P}^0(\underline{K}) \subset \underline{\nabla} W(\underline{K}).$$

We are now ready to state our second result.

Theorem 4.2.2. *Suppose that the Assumptions A and B are satisfied. Also, suppose that the elliptic regularity property* (4.10) *holds. Then we have*

$$\|\Pi_W u - u_h\|_{\mathcal{T}_h} \le C \, h \|\boldsymbol{q} - \boldsymbol{\Pi}_V \boldsymbol{q}\|_{\mathcal{T}_h},$$

for some constant C depending on C^*_{app} but independent of h and the exact solution.

From this result, we immediately get that

$$\|u - u_h\|_{\mathcal{T}_h} \le \|u - \Pi_W u\|_{\mathcal{T}_h} + C \, h \, \|\boldsymbol{q} - \boldsymbol{\Pi}_V \boldsymbol{q}\|_{\mathcal{T}_h},$$

and we see that the quality of the approximation u_h only depends on the approximation error of the projection.

4.2.4 Estimate of $u - u_h^*$

Note that if the second term of the above right-hand side converges faster than the first, the convergence of u_h to $\Pi_W u$ is *faster* than that of u_h to u. As mentioned before, we can take advantage of this *superconverge* result to show that the postprocessing u_h^* defined by (4.8) converges to u as fast as u_h superconverges to $\Pi_W u$. To obtain the error estimate, we need the following assumption.

Assumption C:

- *The reference space $V(\underline{K})$ is not too small.*

(*C.1*) On the reference element $\underline{K}$, we have

$$\mathcal{P}^0(\underline{K}) \subset \underline{\nabla} \cdot \boldsymbol{V}(\underline{K}).$$

We can now state our third and last result.

Theorem 4.2.3. *Suppose that the Assumptions A, B, and C are satisfied. Also, suppose that the elliptic regularity property* (4.10) *holds. Then, we have*

$$\|u - u_h^*\|_{\mathcal{T}_h} \leq \|\Pi_W u - u_h\|_{\mathcal{T}_h} + C\, h\,(\|\boldsymbol{q} - \boldsymbol{\Pi}_V \boldsymbol{q}\|_{\mathcal{T}_h} + \inf_{\omega \in W_h^*} \|\nabla(u - \omega)\|_{\mathcal{T}_h}).$$

Let us point out that the *Assumptions A, B, C* are those in [3] applied to the reference element. So on the reference element, we can use all the elements listed in [3]. They include *all* classical mixed methods as well as the new **TNT** mixed method on rectangles and cubes, and the **HDG** methods for a variety of elements. In fact, if we consider $(k+1)$-compatible meshes, our results imply that all of these methods imply that

$$\|\boldsymbol{q} - \boldsymbol{q}_h\|_{\mathcal{T}_h} \leq C\, h^{k+1},$$
$$\|u - u_h^*\|_{\mathcal{T}_h} \leq C\, h^{k+2},$$

whenever $k \geq 1$ and the exact solution is smooth enough. In other words, these methods are also superconvergent when used on suitably defined curved elements.

Let us also point out that, even though the *Assumptions A, B, C* look very technical and possibly quite restrictive, they are in fact satisfied by a wide variety of HDG and mixed methods using simplexes, prisms, square and cubic elements. This was shown in [3] where a systematic way to satisfy them is proposed.

4.3 Proofs

This Section is devoted to a detailed proof of our main error estimates. We proceed as follows. First, we rewrite *Assumptions A*, and *B* on the physical elements. (We do not need to rewrite *Assumption C*). We then proceed to obtain all the error estimates.

As mentioned in the Introduction, perhaps the main feature of our approach is that, by the introduction of the projections P_M and P_N, the error analysis is almost *identical* to that of the affine case considered in [3], case in which these two projections coincide.

To find out how *Assumptions A* and *B* look when rewritten on the physical elements, we need to introduce some auxiliary projections.

For any face $F \in \mathcal{E}_h$, we define the projections P_M, P_N from $L^2(F)$ onto $M(F), N(F)$ by

$$\langle u, \lambda \rangle_F = \langle P_M u, \lambda \rangle_F \qquad \forall \lambda \in N(F), \tag{4.12a}$$

$$\langle \boldsymbol{q} \cdot \boldsymbol{n}, \mu \rangle_F = \langle P_N(\boldsymbol{q} \cdot \boldsymbol{n}), \mu \rangle_F \quad \forall \mu \in M(F). \tag{4.12b}$$

It is not difficult to see that these projections are well defined.

The *Assumptions A* on the physical elements

The *Assumptions A* on the physical elements are contained in the following result.

Lemma 4.3.1. *For any K in the C^0-compatible triangulation Ω_h, we have*

$$(\Pi_V \boldsymbol{q}, \boldsymbol{v})_K = (\boldsymbol{q}, \boldsymbol{v})_K \qquad \forall \boldsymbol{v} \in \nabla W(K), \tag{4.13a}$$

$$(\Pi_W u, w)_K = (u, w)_K \qquad \forall w \in \nabla \cdot \boldsymbol{V}(K), \tag{4.13b}$$

$$\langle \Pi_V \boldsymbol{q} \cdot \boldsymbol{n} + \alpha(\Pi_W u), \mu \rangle_F = \langle P_V(\boldsymbol{q} \cdot \boldsymbol{n}) + \alpha(P_W u), \mu \rangle_F \quad \forall \mu \in M(F) \quad \forall F \in \triangle_{n-1}(K). \tag{4.13c}$$

Moreover, under Assumption M, we have

$$\boldsymbol{V}(K)\cdot \boldsymbol{n}|_F \subset N(F) \quad \forall F \in \triangle_{n-1}(K), \tag{4.14a}$$

$$W(K)|_F \subset M(F) \qquad \forall F \in \triangle_{n-1}(K). \tag{4.14b}$$

Finally, for any $F \in \mathcal{E}_h$, we have

$$\langle \alpha(\mu), \mu\rangle_F \geq 0 \quad \forall \mu \in M(F). \tag{4.15}$$

Proof. The equality (4.13a) follows from *Assumption (A.1)*, by simply applying the definition of the weighted Piola transform (4.2a), and the definition of the space $W(K)$.

The equality (4.13b) follows from *Assumption (A.2)*, after a simple application of the definition of the scalar transform (4.2b), and the definition of space $V(K)$.

The equality (4.13c) follows from *Assumption (A.3)*, after using the definitions of the weighted Piola and scalar transform, the definition of spaces $V(K), W(K)$, the definition of α, (4.4e), and the identity for the normal component (4.3).

The inclusion properties easily follow from *Assumption (A.4)* and *Assumption (A.5)* the definition of the spaces $M(F)$ and $N(F)$. The last property follows from *Assumption (A.6)* by the definition of α, (4.4e).

This completes the proof. □

The *Assumptions B*

We begin by rewriting the consequence of *Assumption (B.1)* we are going to use.

Lemma 4.3.2. *Let $\{\Omega_h\}_{h\in I}$ be a family of 1-regular meshes. If Assumption (B.1) holds, then for each element K and any $(\boldsymbol{q}, u) \in \boldsymbol{H}^1(K)\times H^1(K)$,*

$$\|\boldsymbol{\Pi}_V^* \boldsymbol{q} - \boldsymbol{q}\|_K \leq C^*_{app}\, h_K(|u|_{1,K} + |\boldsymbol{q}|_{1,K}).$$

Proof. By the definition of weighted Piola transform, (4.2a), we have

$$\begin{aligned}
\|\boldsymbol{\Pi}_V^* \boldsymbol{q} - \boldsymbol{q}\|_K &= h_K^{n-1}\|J_K^{-1/2} D\boldsymbol{G}_K(\underline{\Pi}_V^* \underline{\boldsymbol{q}} - \underline{\boldsymbol{q}})\|_{\underline{K}} \\
&\leq h_K^{n-1}\|J_K^{-1/2} D\boldsymbol{G}_K\|_{L^\infty(\underline{K})}\|\underline{\Pi}_V^* \underline{\boldsymbol{q}} - \underline{\boldsymbol{q}}\|_{\underline{K}} \\
&\leq C\, h_K^{\frac{n}{2}}\|\underline{\Pi}_V^* \underline{\boldsymbol{q}} - \underline{\boldsymbol{q}}\|_{\underline{K}} \\
&\leq C\, h_K^{\frac{n}{2}}(|\underline{u}|_{1,\underline{K}} + |\underline{\boldsymbol{q}}|_{1,\underline{K}}),
\end{aligned}$$

by *Assumption (B.1)*. The result now follows from the estimates

$$|\underline{\boldsymbol{q}}|_{1,\underline{K}} \le C h_K^{1-\frac{n}{2}} \|\boldsymbol{q}\|_{1,K}, \quad |\underline{u}|_{1,\underline{K}} \le C h_K^{1-\frac{n}{2}} \|u\|_{1,K},$$

which can be easily proven by expressing $\underline{\boldsymbol{q}}$ and $\underline{u}$ in terms of $\boldsymbol{q}$ and u, respectively, by using the weighted Piola transform (4.2a) and the scalar transform (4.2b), respectively. This completes the proof. □

Now, let us consider *Assumption (B.2)*. For any element K, we define the null space of the projection Π_V,

$$\boldsymbol{V}_0(K) := \{\boldsymbol{v} \in \boldsymbol{V}(K) : \ (\boldsymbol{q} - \Pi_V \boldsymbol{q}, \boldsymbol{v})_K = 0 \ \forall \boldsymbol{q} \in \boldsymbol{H}^1(K)\}.$$

Lemma 4.3.3. *Let $\{\Omega_h\}_{h\in I}$ be a family of 1-regular meshes. If Assumption (B.2) holds, then for any function $\boldsymbol{z} \in \boldsymbol{H}^1(K)$, there exists an $\boldsymbol{z}_0 \in \boldsymbol{V}_0(K)$, such that*

$$\|\boldsymbol{z} - \boldsymbol{z}_0\|_K \le C h_K \|\boldsymbol{z}\|_{1,K}.$$

This result can be proven by using a standard scaling argument. We omit the proofs for Theorem 4.2.1, Theorem 4.2.2 and Theorem 4.2.3 since they are almost a copy of Section 3.5 in Chapter 3. We refer readers to [4] for more details.

Chapter 5

Superconvergent HDG methods for Stokes equations

5.1 Introduction

In this Chapter, we propose a projection-based analysis of superconvergent hybridizable discontinuous Galerkin (**HDG**) methods for the velocity gradient-velocity-pressure formulation of the Stokes equations, namely,

$$
\begin{aligned}
\mathrm{L} - \nabla \boldsymbol{u} &= 0 && \text{on } \Omega, && (5.1a)\\
-\nabla \cdot (\nu \mathrm{L}) + \nabla p &= \boldsymbol{f} && \text{on } \Omega, && (5.1b)\\
\nabla \cdot \boldsymbol{u} &= 0 && \text{on } \Omega, && (5.1c)\\
\boldsymbol{u} &= \boldsymbol{g} && \text{on } \partial\Omega, && (5.1d)\\
\int_\Omega p &= 0, && && (5.1e)
\end{aligned}
$$

where $\int_{\partial\Omega} \boldsymbol{g} \cdot \boldsymbol{n} = 0$. Here $\Omega \subset \mathbb{R}^n$ $(n = 2, 3)$ is a bounded polygonal domain if $n = 2$ or a Lipschitz polyhedral domain if $n = 3$. We assume that ν is a constant.

In [3], we introduced a new technique to carry out the a priori error analysis of **HDG** methods for second order elliptic problem. The technique reduces the error analysis to the verification of some properties of an elementwise-defined projection and of the local spaces defining the methods. It also reduces the study of the convergence properties of the projection of the errors to that of the approximation properties of the projection. It provides sufficient conditions for the superconvergence of the projection of the error in the scalar approximation; as a consequence, a new scalar approximation can be locally computed which converges with the same order. By using this technique, the well-known mixed methods (Raviart-Thomas methods [12] and their extension by Nédélec [44], Brezzi-Douglas-Marini methods [15], Brezzi-Douglas-Durán-Fortin methods [16]) as well as old [11] and new **HDG** methods can be analyzed at once. In this paper, we extend this methodology to the Stokes equations of incompressible flow.

To better describe our results, we adapt to our setting the notation used in [3] and in [26]. Let $\mathcal{T}_h$ denote the shape-regular, conforming triangulation of Ω, set $\partial\mathcal{T}_h := \{\partial K : K \in \mathcal{T}_h\}$, and let $\mathcal{E}_h$ denote the set of all faces F of all simplexes $K \in \mathcal{T}_h$. We denote by $\mathcal{F}(K)$ the set of all faces F of K.

The methods we consider seek an approximation $(\mathrm{L}_h, \boldsymbol{u}_h, p_h, \widehat{\boldsymbol{u}}_h)$ to the exact solution $(\mathrm{L}|_\Omega, \boldsymbol{u}|_\Omega, p|_\Omega, \boldsymbol{u}|_{\mathcal{E}_h})$ in the finite dimensional space $\mathrm{G}_h \times \boldsymbol{V}_h \times P_h \times \boldsymbol{M}_h$ given by

$$\mathrm{G}_h = \{\mathrm{G} \in \mathrm{L}^2(\mathcal{T}_h) : \quad \mathrm{G}|_K \in \mathrm{G}(K) \quad \forall\, K \in \mathcal{T}_h\}, \tag{5.2a}$$

$$\boldsymbol{V}_h = \{\boldsymbol{v} \in \boldsymbol{L}^2(\mathcal{T}_h) : \quad \boldsymbol{v}|_K \in \boldsymbol{V}(K) \quad \forall\, K \in \mathcal{T}_h\}, \tag{5.2b}$$

$$P_h = \{q \in L^2(\mathcal{T}_h) : \quad q|_K \in P(K) \quad \forall\, K \in \mathcal{T}_h\}, \tag{5.2c}$$

$$\boldsymbol{M}_h = \{\boldsymbol{\mu} \in \boldsymbol{L}^2(\mathcal{E}_h) : \quad \boldsymbol{\mu}|_F \in \boldsymbol{M}(F) \quad \forall\, F \in \mathcal{E}_h\}. \tag{5.2d}$$

To describe how the approximation is defined, we need to introduce some notation related to integrals on the triangulation $\mathcal{T}_h$. Here $(\mathrm{N}\,,\,\mathrm{Z})_{\mathcal{T}_h} := \sum_{i,j=1}^n (\mathrm{N}_{i,j}\,,\,\mathrm{Z}_{i,j})_{\mathcal{T}_h}$, $(\boldsymbol{\eta}\,,\,\boldsymbol{\zeta})_{\mathcal{T}_h} := \sum_{i=1}^n (\eta_i\,,\,\zeta_i)_{\mathcal{T}_h}$, and $(\eta\,,\,\zeta)_{\mathcal{T}_h} := \sum_{K\in\mathcal{T}_h} (\eta, \zeta)_K$, where $(\eta, \zeta)_D$ denotes the integral of $\eta\zeta$ over $D \subset \mathbb{R}^n$. Similarly, we write $\langle \boldsymbol{\eta}\,,\,\boldsymbol{\zeta}\rangle_{\partial\mathcal{T}_h} := \sum_{i=1}^n \langle \eta_i\,,\,\zeta_i\rangle_{\partial\mathcal{T}_h}$ and $\langle \eta\,,\,\zeta\rangle_{\partial\mathcal{T}_h} := \sum_{K\in\mathcal{T}_h} \langle \eta\,,\,\zeta\rangle_{\partial K}$, where $\langle \eta\,,\,\zeta\rangle_D$ denotes the integral of $\eta\zeta$ over $D \subset \mathbb{R}^{n-1}$.

The approximation $(\mathrm{L}_h, \boldsymbol{u}_h, p_h, \widehat{\boldsymbol{u}}_h)$ can now be defined as the solution of the following equations:

$$(\nu \mathrm{L}_h\,,\,\mathrm{G})_{\mathcal{T}_h} + (\boldsymbol{u}_h\,,\,\nabla\cdot\mathrm{G})_{\mathcal{T}_h} - \langle \widehat{\boldsymbol{u}}_h\,,\,\mathrm{G}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} = 0, \tag{5.3a}$$

$$(\nu \mathrm{L}_h\,,\,\nabla\boldsymbol{v})_{\mathcal{T}_h} - (p_h\,,\,\nabla\cdot\boldsymbol{v})_{\mathcal{T}_h} - \langle \nu\widehat{\mathrm{L}}_h\boldsymbol{n} - \widehat{p}_h\boldsymbol{n}\,,\,\boldsymbol{v}\rangle_{\partial\mathcal{T}_h} = (\boldsymbol{f}\,,\,\boldsymbol{v})_{\mathcal{T}_h}, \tag{5.3b}$$

$$-(\boldsymbol{u}_h\,,\,\nabla q)_{\mathcal{T}_h} + \langle \widehat{\boldsymbol{u}}_h\cdot\boldsymbol{n}\,,\,q\rangle_{\partial\mathcal{T}_h} = 0, \tag{5.3c}$$

$$\langle \widehat{\boldsymbol{u}}_h\,,\,\boldsymbol{\mu}\rangle_{\partial\Omega} = \langle \boldsymbol{g}\,,\,\boldsymbol{\mu}\rangle_{\partial\Omega}, \tag{5.3d}$$

$$\langle \nu\widehat{\mathrm{L}}_h\boldsymbol{n} - \widehat{p}_h\boldsymbol{n}\,,\,\boldsymbol{\mu}\rangle_{\partial\mathcal{T}_h\setminus\partial\Omega} = 0, \tag{5.3e}$$

$$(p_h\,,\,1)_{\mathcal{T}_h} = 0, \tag{5.3f}$$

for all $(\mathrm{G}, \boldsymbol{v}, q, \boldsymbol{\mu}) \in \mathrm{G}_h \times \boldsymbol{V}_h \times P_h \times \boldsymbol{M}_h$, where

$$\nu\widehat{\mathrm{L}}_h\boldsymbol{n} - \widehat{p}_h\boldsymbol{n} = \nu\mathrm{L}_h\boldsymbol{n} - p_h\boldsymbol{n} - \boldsymbol{\alpha}(\boldsymbol{u}_h - \widehat{\boldsymbol{u}}_h) \quad \text{on } \partial\mathcal{T}_h. \tag{5.3g}$$

Note that, by taking particular choices of the local spaces $\mathrm{G}(K), \boldsymbol{V}(K), P(K)$ and $\boldsymbol{M}(F)$, and of the *linear local stabilization operator* $\boldsymbol{\alpha}$, different mixed and **HDG** methods are obtained.

Our main result is that if we can construct an auxiliary projection $\Pi_h(\mathrm{L}, \boldsymbol{u}, p) = (\Pi_G\mathrm{L}, \boldsymbol{\Pi}_V\boldsymbol{u}, \mathit{\Pi}_P p)$ satisfying certain orthogonality and approximation conditions, and if the local spaces $\mathrm{G}(K), \boldsymbol{V}(K), P(K)$ and $\boldsymbol{M}(F)$, for $F \in \mathcal{F}(K)$, satisfy some inclusion properties, then the method is well defined and we have the estimates

$$\begin{aligned}
\|\mathrm{L} - \mathrm{L}_h\|_{\mathcal{T}_h} &\le 2\|\mathrm{L} - \Pi_G\mathrm{L}\|_{\mathcal{T}_h}\\
\|p - p_h\|_{\mathcal{T}_h} &\le C_1(\|\mathit{\Pi}_P p - p\|_{\mathcal{T}_h} + \|\mathrm{L} - \Pi_G\mathrm{L}\|_{\mathcal{T}_h}),\\
\|\boldsymbol{\Pi}_V\boldsymbol{u} - \boldsymbol{u}_h\|_{\mathcal{T}_h} &\le C_2 h\|\mathrm{L} - \Pi_G\mathrm{L}\|_{\mathcal{T}_h},
\end{aligned}$$

where $\|\cdot\|_{\mathcal{T}_h}$ denotes the $L^2(\mathcal{T}_h)-$norm and the constants C_1, C_2 solely depend on the *stabilization operator* $\boldsymbol{\alpha}$.

Note that if the error $\boldsymbol{\Pi}_V\boldsymbol{u} - \boldsymbol{u}_h$ converges to zero *faster* than the error $\boldsymbol{u} - \boldsymbol{u}_h$, this *superconvergence* property can be advantageously exploited. To do that, we follow [18, 17, 19] and define a new approximation to $\boldsymbol{u}$, $\boldsymbol{u}_h^*$, in the space

$$\boldsymbol{V}_h^* := \{\boldsymbol{v} \in \boldsymbol{L}^2(\mathcal{T}_h) : \boldsymbol{v}|_K \in \boldsymbol{V}^*(K) \supset \boldsymbol{P}^0(K), K \in \mathcal{T}_h\},$$

as follows. On each element $K \in \mathcal{T}_h$, the postprocessing $\boldsymbol{u}_h^*$ is the element of $\boldsymbol{V}^*(K)$ such that

$$(\nabla \boldsymbol{u}_h^*, \nabla \boldsymbol{v})_K = (\mathrm{L}_h, \nabla \boldsymbol{v})_K \quad \forall \boldsymbol{v} \in \boldsymbol{V}^*(K), \tag{5.4a}$$

$$(\boldsymbol{u}_h^*, \boldsymbol{v})_K = (\boldsymbol{u}_h, \boldsymbol{v})_K \qquad \forall \in \boldsymbol{P}^0(K). \tag{5.4b}$$

Here $\boldsymbol{P}^0(K)$ denotes the space of vector constant functions on K. We are going to show that by properly choosing the spaces $\boldsymbol{V}_h^*(K), K \in \mathcal{T}_h$, we can make $\|\boldsymbol{u}_h^* - \boldsymbol{u}\|_{\mathcal{T}_h}$ converges as fast as $\|\boldsymbol{\Pi}_V \boldsymbol{u} - \boldsymbol{u}_h\|_{\mathcal{T}_h}$ to zero.

5.2 Main results

In this section we show how an a priori error analysis of the HDG methods can be reduced to the verification of a few conditions on the local spaces and on some properties of an associated, auxiliary projection Π_h defined in an element-by-element fashion. The main idea of our error analysis is to estimate the projection of the errors $\Pi_h(\mathrm{L}-\mathrm{L}_h, \boldsymbol{u}-\boldsymbol{u}_h, p-p_h)$ and then deduce bounds of the $L^2(\Omega)-$norm of the errors $\mathrm{L}-\mathrm{L}_h, \boldsymbol{u}-\boldsymbol{u}_h, p-p_h$ and $\boldsymbol{u}-\boldsymbol{u}^*$.

5.2.1 Estimate of $\Pi_G \mathrm{L} - \mathrm{L}_h$ and $\mathit{\Pi}_P p - p_h$

Our first result gives an estimate of the projection of the error $\Pi_G \mathrm{L} - \mathrm{L}_h$ solely in terms of the approximation error of the projection $\mathrm{L} - \Pi_G \mathrm{L}$. Our second result gives an estimate of the projection of the error $\mathit{\Pi}_P p - p_h$ in terms of the approximation errors $p - \mathit{\Pi}_P p$ and $\mathrm{L} - \Pi_G \mathrm{L}$. To state them, we need to describe our assumptions on the projection Π_h and on the local finite element spaces $\mathrm{G}(K), \boldsymbol{V}(K), P(K)$ and $\boldsymbol{M}(F)$.

Assumption A:

• *Orthogonality properties of* Π_h. On each element K, there exists a projection $\Pi_h(\mathrm{L}, \boldsymbol{u}, p) = (\Pi_G \mathrm{L}, \boldsymbol{\Pi}_V \boldsymbol{u}, \mathit{\Pi}_P p) \in \mathrm{G}(K) \times \boldsymbol{V}(K) \times P(K)$ satisfying the following properties:

(A.1) $(\Pi_G \mathrm{L}, \mathrm{G})_K = (\mathrm{L}, \mathrm{G})_K \quad \text{for all } \mathrm{G} \in \nabla \boldsymbol{V}(K),$

$(A.2)$ $(\mathbf{\Pi}_V \boldsymbol{u}, \boldsymbol{v})_K = (\boldsymbol{u}, \boldsymbol{v})_K \quad$ for all $\boldsymbol{v} \in \nabla \cdot \mathrm{G}(K) + \nabla P(K)$,

$(A.3)$ $(\mathit{\Pi}_P p, q)_K = (p, q)_K \quad$ for all $q \in \nabla \cdot \boldsymbol{V}(K)$,

$(A.4)$ For $F \in \mathcal{F}(K)$,

$\langle \nu \Pi_G \mathrm{L}\boldsymbol{n} - \mathit{\Pi}_P p\boldsymbol{n} - \boldsymbol{\alpha}(\mathbf{\Pi}_V \boldsymbol{u}) \,,\, \boldsymbol{\mu} \rangle_F = \langle \nu \mathrm{L}\boldsymbol{n} - p\boldsymbol{n} - \boldsymbol{\alpha}(\mathbf{\Pi}_M \boldsymbol{u}) \,,\, \boldsymbol{\mu} \rangle_F$ for all $\boldsymbol{\mu} \in \boldsymbol{M}(F)$.

We also need to assume suitable relations between the traces on the faces F of the local spaces $\mathrm{G}(K), \boldsymbol{V}(K)$ and $P(K)$ with the local space $\boldsymbol{M}(F)$.

- *Properties of the traces of the local spaces.* For each element K, and for any of its face F, we assume,

$(A.5)$ $\mathrm{G}(K)\boldsymbol{n}|_F \subset \boldsymbol{M}(F)$,

$(A.6)$ $\boldsymbol{V}(K)|_F \subset \boldsymbol{M}(F)$,

$(A.7)$ $P(K)\boldsymbol{n}|_F \subset \boldsymbol{M}(F)$.

Here, $\mathrm{G}(K)\boldsymbol{n}|_F$ denotes the space of the *traces* of normal components of functions of $\mathrm{G}(K)$ on the face F of K. Similarly, $\boldsymbol{V}(K)|_F$ denotes the space of the *traces* of space $\boldsymbol{V}(K)$ on the face F of K. $P(K)\boldsymbol{n}|_F$ denotes the *traces* of functions of $P(K)$ multiplied by unit normal vector $\boldsymbol{n}$ on the face F.

Finally, we need a simple assumption on the *local stabilization operator* $\boldsymbol{\alpha}$:

- *The semi-positivity property of* $\boldsymbol{\alpha}$. For each element K and any face F,

$(A.8)$ $\langle \boldsymbol{\alpha}(\boldsymbol{\mu}), \boldsymbol{\mu} \rangle_F \geq 0$ for all $\boldsymbol{\mu} \in \boldsymbol{M}(F)$.

We are now ready to state our first result. In what follows, we use $\|\cdot\|_{k,D}, |\cdot|_{k,D}$ to denote the standard norm and seminorm on any Sobolev space $H^k(D)$, respectively. When $k = 0$, we omit the index k and simply write $\|\cdot\|_D$.

Theorem 5.2.1. *Suppose that the Assumption A are satisfied. Then we have*

$$\|\Pi_G \mathrm{L} - \mathrm{L}_h\|_{\mathcal{T}_h} \le \|\mathrm{L} - \Pi_G \mathrm{L}\|_{\mathcal{T}_h}.$$

Note that, we can immediately conclude that $\|\mathrm{L} - \mathrm{L}_h\|_{\mathcal{T}_h} \le 2\|\mathrm{L} - \Pi_G \mathrm{L}\|_{\mathcal{T}_h}$, and we see that the quality of the approximation L_h depends on the approximation properties of $\Pi_G \mathrm{L}$ *only.*

Our second result is an estimate for $p - p_h$. To state it, we use the following notation for the total average of a function over Ω: $\overline{p} := \int_\Omega p/|\Omega|$, the following seminorm on $\partial\mathcal{T}_h$: $\|\boldsymbol{\mu}\|_{\boldsymbol{\alpha}/\nu} := \left\{\sum_{K\in\mathcal{T}_h} \frac{1}{\nu}\langle \boldsymbol{\alpha}(\boldsymbol{\mu})\,,\,\boldsymbol{\mu}\rangle_{\partial K}\right\}^{\frac{1}{2}}$, and the following auxiliary space

$$\begin{aligned}\mathcal{M}^{\perp}_{\boldsymbol{n},h} := \{\boldsymbol{\mu} \in \boldsymbol{L}^2(\partial\mathcal{T}_h) : \quad &\forall K \in \mathcal{T}_h : \ \ \boldsymbol{\mu}|_F \in \boldsymbol{M}(F) \ \ \forall\, F \in \mathcal{F}(K), \\ &\langle \boldsymbol{\mu}\cdot\boldsymbol{n}, q\rangle_{\partial K} = 0 \ \ \forall q \in P(K) : \ \ (q, \nabla\cdot\boldsymbol{v})_K = 0 \ \ \forall \boldsymbol{v} \in \boldsymbol{V}(K)\}.\end{aligned} \tag{5.5}$$

Theorem 5.2.2. *Suppose that the Assumption A is satisfied. Then*

$$\|\mathit{\Pi}_P p - p_h\|_{\mathcal{T}_h} \le |\overline{\mathit{\Pi}_P p - p}||\Omega|^{\frac{1}{2}} + C C_{p,\alpha/\nu}\nu \|\mathrm{L} - \Pi_G \mathrm{L}\|_{\mathcal{T}_h},$$

where

$$C_{p,\alpha/\nu} := max\left\{1, \sup_{\boldsymbol{w}\in\boldsymbol{H}^1_0(\Omega)\backslash\{\boldsymbol{0}\}} \ \sup_{\boldsymbol{\mu}\in\mathcal{M}^{\perp}_{n,h}\backslash\{0\}} \frac{\langle \boldsymbol{\alpha}(\boldsymbol{\mu})/\nu\,,\,\boldsymbol{P}\boldsymbol{w} - \boldsymbol{\Pi}_M \boldsymbol{w}\rangle_{\partial\mathcal{T}_h}}{\|\boldsymbol{\mu}\|_{\boldsymbol{\alpha}/\nu}\|\boldsymbol{w}\|_{\boldsymbol{H}^1(\Omega)}}\right\}.$$

and $\boldsymbol{P} : \boldsymbol{H}^1(\mathcal{T}_h) \to \boldsymbol{V}_h$ *is any projection such that* $(\boldsymbol{P}\boldsymbol{w} - \boldsymbol{w}, \boldsymbol{v})_{\mathcal{T}_h} = 0$ *for all* $\boldsymbol{v} \in \boldsymbol{V}_h$ *such that* $\boldsymbol{v}|_K \in \nabla P(K)$ *for all* $K \in \mathcal{T}_h$.

As a consequence, applying the triangle inequality, we immediately obtain

$$\|p - p_h\|_{\mathcal{T}_h} \le |\overline{\mathit{\Pi}_P p - p}||\Omega|^{\frac{1}{2}} + \|p - \mathit{\Pi}_P p\|_{\mathcal{T}_h} + C C_{p,\alpha/\nu}\nu\|\mathrm{L} - \Pi_G \mathrm{L}\|_{\mathcal{T}_h}.$$

5.2.2 Estimate of $\boldsymbol{\Pi}_V \boldsymbol{u} - \boldsymbol{u}_h$

Next, we provide estimates of $\boldsymbol{u} - \boldsymbol{u}_h$ and $\boldsymbol{u} - \widehat{\boldsymbol{u}}$. To state the results, we need to introduce the following dual problem. For any given $\boldsymbol{\theta}$ in $\boldsymbol{L}^2(\Omega)$, let $(\mathrm{Z}, \boldsymbol{\sigma}, \eta)$ be the solution of

$$\begin{aligned} \mathrm{Z} - \nabla \boldsymbol{\sigma} &= 0 \quad \text{on } \Omega, & (5.6a)\\ \nabla \cdot (\nu \mathrm{Z}) - \nabla \eta &= \boldsymbol{\theta} \quad \text{on } \Omega, & (5.6b)\\ -\nabla \cdot \boldsymbol{\sigma} &= 0 \quad \text{on } \Omega, & (5.6c)\\ \boldsymbol{\sigma} &= 0 \quad \text{on } \partial\Omega. & (5.6d) \end{aligned}$$

We assume that, for some real number s, we have the regularity property

$$\nu \|\mathrm{Z}\|_{\mathrm{H}^{s+1}(\omega)} + \nu \|\boldsymbol{\sigma}\|_{\boldsymbol{H}^{s+2}(\Omega)} + \|\eta\|_{H^{s+1}(\Omega)} \leq C_{reg} \|\boldsymbol{\theta}\|_{\boldsymbol{H}^{s}(\omega)}. \tag{5.7}$$

In the two-dimensional case, the above estimate with $s \leq 0$ follows from the results in [49] when the domain is convex. In the three-dimensional case, the above estimate follows from the results in [50] in the following cases: For any polyhedron with Lipschitz boundary, with $s < 1/2$; for any convex polyhedron, with $s \leq 0$; and with $s < 3/2$ if, moreover, all the edges have wedge angles at most $2\pi/3$ (a cube, for example).

Moreover, we need a couple of additional assumptions.

Assumption B:

The first assumption is an approximation property of a projection $\Pi_h^*(\mathrm{L}, \boldsymbol{u}, p) = (\Pi_G^* \mathrm{L}, \boldsymbol{\Pi}_V^* \boldsymbol{u}, \mathit{\Pi}_P^* p)$ which satisfies the assumptions $(A.1) - (A.4)$ where the local stabilization operator $\boldsymbol{\alpha}(\cdot)$ is *replaced* by its dual $\boldsymbol{\alpha}^*(\cdot)$, that is , by the linear function defined by

$$\langle \boldsymbol{\alpha}(\boldsymbol{v}) \, , \, \boldsymbol{w} \rangle_{\partial K} = \langle \boldsymbol{v} \, , \, \boldsymbol{\alpha}^*(\boldsymbol{w}) \rangle_{\partial K} \quad \text{for all } \boldsymbol{v}, \boldsymbol{w} \in \boldsymbol{L}^2(\partial K).$$

• *The approximation property of the projection* Π_h^*. For each element K and any $(\mathrm{L}, \boldsymbol{u}, p) \in \mathrm{H}^1(K) \times \boldsymbol{H}^2(K) \times H^1(K)$,

(B.1) $\|\Pi_G^* \mathrm{L} - \mathrm{L}\|_K \leq C_{app}^* h_K (|\mathrm{L}|_{1,K} + |\boldsymbol{u}|_{1,K} + |p|_{1,K}).$

The second assumption is a condition on the local space $\boldsymbol{V}(K)$.

- *The local space $\boldsymbol{V}(K)$ is not too small.* For each element K, we have that

$$(B.2)\quad \mathrm{P}^0(K) \subset \nabla \boldsymbol{V}(K).$$

Here $P^0(K)$ denotes the space of constant functions on K.

We are now ready to state our third result. To state it, we need to introduce the following norm for functions $\boldsymbol{\zeta} \in \boldsymbol{L}^2(\partial\mathcal{T}_h)$: $\|\boldsymbol{\zeta}\|_h := \left\{\sum_{K\in\mathcal{T}_h} h_K \|\boldsymbol{\zeta}\|^2_{\partial K}\right\}^{1/2}$.

Theorem 5.2.3. *Suppose that the Assumption A and B are satisfied. Also, suppose that the regularity property* (5.7) *holds. Then we have*

$$\|\boldsymbol{\Pi}_V \boldsymbol{u} - \boldsymbol{u}_h\|_{\mathcal{T}_h} + \|\boldsymbol{\Pi}_M \boldsymbol{u} - \widehat{\boldsymbol{u}}_h\|_h \le C\nu h \|\mathrm{L} - \Pi_G \mathrm{L}\|_{\mathcal{T}_h}.$$

5.2.3 Estimate of $\boldsymbol{u} - \boldsymbol{u}_h^*$

Note that if the convergence of $\boldsymbol{u}_h$ to $\boldsymbol{\Pi}_V \boldsymbol{u}$ is *faster* than that of $\boldsymbol{u}_h$ to $\boldsymbol{u}$, we can take the advantage of this *superconvergence* result to show that the postprocessing $\boldsymbol{u}_h^*$ defined by (5.4) converges to $\boldsymbol{u}$ as fast as $\boldsymbol{u}_h$ superconverges to $\boldsymbol{\Pi}_V \boldsymbol{u}$. To obtain the error estimate, we need the following assumption.

Assumption C:

- The local space $\mathrm{G}(K)$ is not too small. For each element K,

$$(C.1)\quad \boldsymbol{P}^0(K) \subset \nabla \cdot \mathrm{G}(K).$$

We can now state our last result.

Theorem 5.2.4. *Suppose that the Assumptions A, B, and C are satisfied. Also, suppose that the regularity property* (5.7) *holds. Then, we have*

$$\|\boldsymbol{u} - \boldsymbol{u}_h^*\|_{\mathcal{T}_h} \le \|\boldsymbol{\Pi}_V \boldsymbol{u} - \boldsymbol{u}_h\|_{\mathcal{T}_h} + Ch(\|\mathrm{L} - \mathrm{L}_h\|_{\mathcal{T}_h} + \inf_{\boldsymbol{v}\in \boldsymbol{V}_h^*} \|\nabla(\boldsymbol{u} - \boldsymbol{v})\|_{\mathcal{T}_h}).$$

5.3 A template for the construction of superconvergent methods

Here, we provide a template which reduces, roughly speaking, the devising of superconvergent **HDG** methods to a suitable choice of the spaces in each element of the triangulation.

5.3.1 The choice of the local spaces and the stabilization operator

To construct our superconvergent methods, we pick an arbitrary element $K \in \mathcal{T}_h$, and proceed as follows:

Step 1: The local space $\mathrm{G}(K) \times \boldsymbol{V}(K)$ We begin by taking a local space $\mathrm{G}(K) \times \boldsymbol{V}(K)$ such that

$$\mathrm{P}^0(K) \subset \nabla \boldsymbol{V}(K) \quad \subset \mathrm{G}(K), \tag{5.8a}$$

$$(\nabla\cdot \boldsymbol{V}(K))\,\mathrm{Id} \subset \mathrm{G}(K), \tag{5.8b}$$

$$\boldsymbol{P}^0(K) \subset \nabla \cdot \mathrm{G}(K) \quad \subset \boldsymbol{V}(K). \tag{5.8c}$$

Step 2: The local space $\boldsymbol{M}(F)$ Then, for each face F of the element K, we find a space $\boldsymbol{M}(F)$ such that

$$\mathrm{G}(K)\boldsymbol{n}|_F \subset \boldsymbol{M}(F), \tag{5.9a}$$

$$\boldsymbol{V}(K)|_F \subset \boldsymbol{M}(F). \tag{5.9b}$$

This choice has to be made so that

$$\sum_{F\in\mathcal{F}(K)} \dim \boldsymbol{M}(F) \leq (\dim \mathrm{G}(K) - \dim \nabla\boldsymbol{V}(K)) + (\dim \boldsymbol{V}(K) - \nabla \cdot \mathrm{G}(K)). \tag{5.10}$$

Step 3: The auxiliary local space $\widetilde{\mathrm{G}}(K) \times \widetilde{\boldsymbol{V}}(K)$ Next, we find an *auxiliary* space $\widetilde{\mathrm{G}}(K) \times \widetilde{\boldsymbol{V}}(K)$ satisfying

$$\nabla \boldsymbol{V}(K) \subset \widetilde{\mathrm{G}}(K) \subset \mathrm{G}(K), \tag{5.11a}$$

$$\nabla \cdot \mathrm{G}(K) \subset \widetilde{\boldsymbol{V}}(K) \subset \boldsymbol{V}(K), \tag{5.11b}$$

such that, if we set

$$\mathrm{G}^{\perp}(K) := \{\mathrm{G} \in \mathrm{G}(K) \,:\, (\mathrm{G}, \widetilde{\mathrm{G}})_K = 0 \quad \forall\, \widetilde{\mathrm{G}}(K)\}, \tag{5.12a}$$

$$\boldsymbol{V}^{\perp}(K) := \{\boldsymbol{v} \in \boldsymbol{V}(K) \,:\, (\boldsymbol{v}, \widetilde{\boldsymbol{v}})_K = 0 \quad \forall\, \widetilde{\boldsymbol{V}}(K)\}, \tag{5.12b}$$

we have that

$$\sum_{F \in \mathcal{F}(K)} \dim \boldsymbol{M}(F) = \dim \mathrm{G}^{\perp}(K) + \dim \boldsymbol{V}^{\perp}(K), \tag{5.13a}$$

and that

$$\|\mathrm{G}^{\perp}\|_K \le C_{\mathrm{G}}\, h_K^{1/2} \|\mathrm{G}^{\perp}\boldsymbol{n}\|_{\partial K_{\mathrm{G}}} \qquad \text{for all } \mathrm{G}^{\perp} \in \mathrm{G}^{\perp}(K), \tag{5.14a}$$

$$\|\boldsymbol{v}^{\perp}\|_K \le C_V\, h_K^{1/2} \|\, \boldsymbol{v}^{\perp} \cdot \boldsymbol{n}\,\|_{\partial K_V} \qquad \text{for all } \boldsymbol{v}^{\perp} \in \boldsymbol{V}^{\perp}(K), \tag{5.14b}$$

for some subsets ∂K_G and ∂K_V of $\mathcal{F}(K)$.

Step 4: The local space $P(K)$ Now, we take the local space $P(K)$ such that

$$\nabla \cdot \boldsymbol{V}(K) \subset P(K), \tag{5.15a}$$

$$P(K)\,\mathrm{Id} \subset \mathrm{G}(K). \tag{5.15b}$$

Step 5: The stabilization operator $\boldsymbol{\alpha}$ Finally, we pick the local stabilization operator $\boldsymbol{\alpha}$ satisfying the semi-positivity condition *(A.8)* and the following properties:

$$\langle \boldsymbol{\alpha}(\boldsymbol{\eta}),\, \boldsymbol{\mu}\rangle_F = \langle \boldsymbol{\eta}, \boldsymbol{\alpha}(\boldsymbol{\mu})\rangle_F \qquad \text{for all } \boldsymbol{\eta}, \boldsymbol{\mu} \in \boldsymbol{M}(F), \tag{5.16a}$$

$$\langle \boldsymbol{\alpha}(\boldsymbol{v}^{\perp}), \boldsymbol{v}^{\perp}\rangle_{\partial K} \ge C_\alpha \,\|\boldsymbol{v}^{\perp} \cdot \boldsymbol{n}\|^2_{\partial K_V} \text{ for all } \boldsymbol{v}^{\perp} \in \boldsymbol{V}^{\perp}(K). \tag{5.16b}$$

5.3.2 Verification of *Assumptions A, B* and *C*

We claim that the **HDG** method determined by the above local spaces and stabilization operator does satisfy *Assumptions A, B,* and *C*. Let us show that this is indeed the case.

It is easy to see that

- *Assumption* (*A.5*) is nothing but condition (5.9a),
- *Assumption* (*A.6*) is nothing but condition (5.9b),

- *Assumption* (*A.7*) follows from condition (5.15b) and *Assumption* (*A.5*),
- *Assumption* (*B.2*) is nothing but the first inclusion in condition (5.8a),
- *Assumption* (*C.1*) is nothing but the first inclusion in condition (5.8c).

Note that *Assumption* (*A.8*) was supposed to hold in Step 5.

To verify the remaining *Assumptions*, we must introduce an auxiliary projection Π_h. We define, for any element of $\mathrm{H}^1(K) \times \boldsymbol{H}^1(K) \times H^1(K)$, $(\mathrm{L}, \boldsymbol{u}, p)$, the projection $\Pi_h(\mathrm{L}, \boldsymbol{u}, p) := (\Pi_G \mathrm{L}, \boldsymbol{\Pi}_V \boldsymbol{u}, \mathit{\Pi}_P p)$ as the element of $\mathrm{G}(K) \times \boldsymbol{V}(K) \times P(K)$ defined as follows. The pressure component is defined by

$$(\mathit{\Pi}_P p, q)_K = (p, q)_K \qquad \forall q \in P(K), \tag{5.17a}$$

whereas the remaining components by

$$\begin{aligned}
(\Pi_G \mathrm{L}, \mathrm{G})_K &= (\mathrm{L}, \mathrm{G})_K && \forall \mathrm{G} \in \widetilde{\mathrm{G}}(K), && (5.17\mathrm{b})\\
(\boldsymbol{\Pi}_V \boldsymbol{u}, \boldsymbol{v})_K &= (\boldsymbol{u}, \boldsymbol{v})_K && \forall \boldsymbol{v} \in \widetilde{\boldsymbol{V}}(K), && (5.17\mathrm{c})\\
\langle \nu \Pi_G \mathrm{L}\boldsymbol{n} - \boldsymbol{\alpha}(\boldsymbol{\Pi}_V \boldsymbol{u}), \boldsymbol{\mu}\rangle_F &= \langle \nu \mathrm{L}\boldsymbol{n} - \boldsymbol{\alpha}(\boldsymbol{u}), \boldsymbol{\mu}\rangle_F && && (5.17\mathrm{d})\\
&\quad - \langle p - \mathit{\Pi}_P p, \boldsymbol{\mu} \cdot \boldsymbol{n}\rangle_F && \forall \boldsymbol{\mu} \in \boldsymbol{M}(F), &&
\end{aligned}$$

for $F \in \mathcal{F}(K)$.

If this projection were well defined, *Assumption* (*A.1*) would follow from the second equation defining the projection, (5.17b), and from the first inclusion in condition (5.11a); *Assumption* (*A.2*) would follow from the third equation defining the projection, (5.17c), from the first inclusion in condition (5.11b) and from condition (5.15b); *Assumption* (*A.3*) from the first equation defining the projection, (5.17a), and the inclusion condition (5.15a); and *Assumption* (*A.4*) from the last equation defining the projection, (5.17d). Thus, it remains to prove that the projection is well defined and that it satisfies *Assumption* (*B.1*).

Note that since we are assuming that the stabilization operator α is self-adjoint, see condition (5.16a), we have that $\Pi_h^* = \Pi_h$. Note also that, by condition (5.13), the system of equations defining the projection Π_h is square. Hence, it is well defined if and only if, when $(\mathrm{L}, \boldsymbol{u}, p) = (0, \mathbf{0}, 0)$, we have that $\Pi_h(\mathrm{L}, \boldsymbol{u}, p) = (0, \mathbf{0}, 0)$. As a consequence, both the existence of the projection Π_h as well as *Assumption(B.1)* follow from an approximation result we state next.

To do that, we need to introduce some notation. We denote by $(\boldsymbol{P}_V, P_W, \boldsymbol{P}_{\widetilde{V}})$ the L^2-projection into the local space $\boldsymbol{V}(K) \times W(K) \times \widetilde{\boldsymbol{V}}(K)$. For any face F of the element K, we set

$$\|\alpha\|_F := \sup_{\mu \in M(F)\setminus\{0\}} \|\alpha(\mu)\|_F / \|\mu\|_F,$$

and define $\|\alpha\|_D := \max_{F\in D} \|\alpha\|_F$ where D is any union of faces of K. Finally, we set, for $W^\perp(K) \neq \{0\}$, $R_{W^\perp} := \sup_{w\in W^\perp(K)\setminus\{0\}} h_K^{1/2} \|w\|_{\partial K}/\|w\|_K$.

We are now ready to state our result.

Theorem 5.3.1. *We have*

$$\begin{aligned}
\|\mathrm{L} - \Pi_G \mathrm{L}\|_K \leq \|\mathrm{I_L}\|_K + \mathrm{C}_1 h_K^{\frac{1}{2}} \|\nu \mathrm{I_L} \boldsymbol{n} - I_P \boldsymbol{n}\|_{\partial K_{\mathrm{G}}} \\
+ \mathrm{C}_2 h_K \|(\boldsymbol{Id} - \boldsymbol{P}_{\widetilde{\boldsymbol{V}}})(\nu\nabla\cdot \mathrm{L} - \nabla p)\|_K + \mathrm{C}_3 h_K^{\frac{1}{2}} \|\boldsymbol{I_u}\|_{\partial K}, \\
\|\boldsymbol{u} - \boldsymbol{\Pi}_V \boldsymbol{u}\|_K \leq \|\boldsymbol{I_u}\|_K + \mathrm{C}_4 h_K^{\frac{1}{2}} \|\boldsymbol{I_u}\|_{\partial K} + \mathrm{C}_5 h_K \|(\boldsymbol{Id} - \boldsymbol{P}_{\widetilde{\boldsymbol{V}}})(\nu\nabla\cdot \mathrm{L} - \nabla p)\|_K,
\end{aligned}$$

Here $\mathrm{I_L} := \mathrm{L} - \mathrm{P_G L}$*,* $\boldsymbol{I_u} := \boldsymbol{u} - \boldsymbol{P}_V \boldsymbol{u}$ *and* $I_p := p - P_W p$*. Moreover,* $\mathrm{C}_1 = C_G, \mathrm{C}_2 = 0, \mathrm{C}_3 = C_G C_V R_{V^\perp} \|\boldsymbol{\alpha}\|_{\partial K_{\mathrm{G}}}/C_\alpha$*,* $\mathrm{C}_4 = \mathrm{C}_5 = 0$ *whenever* $\widetilde{W}(K) = W(K)$*. Otherwise,*

$$\begin{aligned}
&\mathrm{C}_1 = C_G, \quad \mathrm{C}_2 = \frac{C_G C_V^2 R_{V^\perp}}{C_\alpha}, \quad \mathrm{C}_3 = C_G \|\boldsymbol{\alpha}\|_{\partial K_{\mathrm{G}}}(1 + \frac{C_V^2 R_{V^\perp}^2 \|\boldsymbol{\alpha}\|_{\partial K}}{C_\alpha}), \\
&\mathrm{C}_4 = \frac{C_V^2 R_{V^\perp} \|\boldsymbol{\alpha}\|_{\partial K}}{C_\alpha}, \quad \mathrm{C}_5 = \frac{C_V^2}{C_\alpha}.
\end{aligned}$$

This result contains the information of how the choice of local spaces and stabilization operator affects the approximation properties of the projection. It indicates how to choose them to obtain optimal orders of convergence. Let us focus our discussion on the estimate of $\|\mathrm{L} - \Pi_G \mathrm{L}\|_K$ as it is the only relevant one for the convergence properties described in the theorems of Section 2.

Note that if we have that $\partial K_{\mathrm{G}} \cap \partial K_{\boldsymbol{V}} = \emptyset$ and if we take $\boldsymbol{\alpha}$ in such a way that $\|\boldsymbol{\alpha}\|_{\partial K_{\mathrm{G}}} = 0$, then $\mathrm{C}_3 = 0$. In this case, the approximation properties of Π_G are solely controlled by the L^2-projections $\mathrm{P_G}, \mathit{\Pi}_P, \boldsymbol{P}_{\widetilde{\boldsymbol{V}}}$. Hence, we get optimal approximation properties if the L^2-projection $\mathit{\Pi}_P$ converges as fast as $\mathrm{P_G}$ to the identity, and if the convergence order of $\boldsymbol{P}_{\widetilde{\boldsymbol{V}}}$ to the identity is one less than that of $\mathrm{P_G}$.

In the general case, it is enough to take the stabilization operator $\boldsymbol{\alpha}$ such that $\|\boldsymbol{\alpha}\|_{\partial K_{\mathrm{G}}}$ and $\|\boldsymbol{\alpha}\|_{\partial K_{\mathrm{G}}}/C_\alpha$ are uniformly bounded to ensure that the constants $\mathrm{C}_1, \mathrm{C}_2$ and C_3 are independent of $\boldsymbol{\alpha}$.

5.3.3 Proof of the approximation properties of Π_h, Theorem 5.3.1

To prove the estimates of Theorem 5.3.1, we follow [26]. The idea is to estimate the quantities $\delta_{\mathrm{L}} := \Pi_G \mathrm{L} - \mathrm{P}_G \mathrm{L}$ and $\boldsymbol{\delta_u} := \boldsymbol{\Pi}_V \boldsymbol{u} - \boldsymbol{P}_V \boldsymbol{u}$, and then use the triangle inequality to obtain the desired estimates. We proceed in three steps.

Step 1: The equations for δ_{L} and $\boldsymbol{\delta_u}$

By the equations defining the projection Π_h, (5.17), we have that

$$(\delta_{\mathrm{L}}, \mathrm{G})_K = 0 \qquad \forall \mathrm{G} \in \widetilde{\mathrm{G}}(K), \tag{5.18a}$$

$$(\boldsymbol{\delta_u}, \boldsymbol{v})_K = 0 \qquad \forall \boldsymbol{v} \in \widetilde{\boldsymbol{V}}(K), \tag{5.18b}$$

$$\langle \nu \delta_{\mathrm{L}} \boldsymbol{n} - \boldsymbol{\alpha}(\boldsymbol{\delta_u}), \boldsymbol{\mu} \rangle_F = \langle \nu \mathrm{I}_{\mathrm{L}} \boldsymbol{n} - \boldsymbol{\alpha}(\boldsymbol{I_u}) - I_p \boldsymbol{n}, \boldsymbol{\mu} \rangle_F \quad \forall \boldsymbol{\mu} \in \boldsymbol{M}(F), \tag{5.18c}$$

for $F \in \mathcal{F}(K)$.

Step 2: The estimate of $\boldsymbol{\delta_u}$

Next, we obtain an estimate of $\boldsymbol{\delta_u}$. By the definition of $\boldsymbol{V}^\perp(K)$, (5.12b), we see that $\boldsymbol{\delta_u} \in \boldsymbol{V}^\perp(K)$, by the equation (5.18b). If $\boldsymbol{V}^\perp(K) = \{\boldsymbol{0}\}$, then

$$\|\boldsymbol{\delta_u}\|_K = \boldsymbol{0}.$$

If $\boldsymbol{V}^\perp(K) \neq \{\boldsymbol{0}\}$, we claim that $\boldsymbol{\delta_u}$ is the element of $\boldsymbol{V}^\perp(K)$, satisfying

$$\langle \boldsymbol{\alpha}(\boldsymbol{\delta_u}), \boldsymbol{v} \rangle_{\partial K} = -((\boldsymbol{Id} - \boldsymbol{P}_{\widetilde{\boldsymbol{V}}})(\nu \nabla \cdot \mathrm{L} - \nabla p), \boldsymbol{v})_K + \langle \boldsymbol{\alpha}(\boldsymbol{I_u}), \boldsymbol{v} \rangle_{\partial K}$$

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

Taking $\boldsymbol{v} := \boldsymbol{\delta_u}$ and applying the Cauchy-Schwarz inequality, we get

$$\begin{aligned}
\langle \boldsymbol{\alpha}(\boldsymbol{\delta_u}), \boldsymbol{\delta_u} \rangle_{\partial K} &\leq \|(\boldsymbol{Id} - \boldsymbol{P}_{\widetilde{\boldsymbol{V}}})(\nu \nabla \cdot \mathrm{L} - \nabla p)\|_K \|\boldsymbol{\delta_u}\|_K + \|\boldsymbol{\alpha}(\boldsymbol{I_u})\|_{\partial K} \|\boldsymbol{\delta_u}\|_{\partial K} \\
&\leq C_V\, h_K^{1/2}\, \|(\boldsymbol{Id} - \boldsymbol{P}_{\widetilde{\boldsymbol{V}}})(\nu \nabla \cdot \mathrm{L} - \nabla p)\|_K \|\boldsymbol{\delta_u} \cdot \boldsymbol{n}\|_{\partial K_V} \\
&\quad + C_V\, R_{V^\perp}\, \|\boldsymbol{\alpha}\|_{\partial K}\, \|\boldsymbol{I_u}\|_{\partial K}\, \|\boldsymbol{\delta_u} \cdot \boldsymbol{n}\|_{\partial K_V},
\end{aligned}$$

by the condition (5.14b). By the condition on the stabilization operator, (5.16b),

$$\|\boldsymbol{\delta_u}\cdot\boldsymbol{n}\|_{\partial K_V} \le \frac{C_V}{C_\alpha}\,\left(h_K^{1/2}\,\|(\boldsymbol{Id}-\boldsymbol{P}_{\widetilde{\boldsymbol{V}}})(\nu\nabla\cdot\mathrm{L}-\nabla p)\|_K + \,R_{V^\perp}\,\|\boldsymbol{\alpha}\|_{\partial K}\,\|\boldsymbol{I_u}\|_{\partial K}\right).$$

Finally, using once again condition (5.14b), we get that

$$\|\boldsymbol{\delta_u}\|_K \le \frac{C_V^2}{C_\alpha}\,\left(h_K\,\|(\boldsymbol{Id}-\boldsymbol{P}_{\widetilde{\boldsymbol{V}}})(\nu\nabla\cdot\mathrm{L}-\nabla p)\|_K + h_K^{1/2}\,R_{V^\perp}\,\|\boldsymbol{\alpha}\|_{\partial K}\,\|\boldsymbol{I_u}\|_{\partial K}\right).$$

It remains to prove the claim. By the equation (5.18c), we have that

$$\langle\nu\delta_{\mathrm{L}}\boldsymbol{n}-\boldsymbol{\alpha}(\boldsymbol{\delta_u}),\boldsymbol{v}\rangle_{\partial K} = \langle\nu\mathrm{I_L}\boldsymbol{n}-\boldsymbol{\alpha}(\boldsymbol{I_u})-I_p\boldsymbol{n},\boldsymbol{v}\rangle_{\partial K}\quad\forall\boldsymbol{v}\in\boldsymbol{V}^\perp(F),$$

because $\boldsymbol{v}|_F\in\boldsymbol{M}(F)$ by the inclusion condition (5.9b).

But

$$\langle\nu\delta_{\mathrm{L}}\boldsymbol{n},\boldsymbol{v}\rangle_{\partial K} = \nu\,(\nabla\cdot\delta_{\mathrm{L}},\boldsymbol{v})_K+\nu\,(\delta_{\mathrm{L}},\nabla\boldsymbol{v})_K = 0.$$

Indeed, we have that $(\nabla\cdot\delta_{\mathrm{L}},\boldsymbol{v})_K = 0$ by the first inclusion in condition (5.11b) and the fact that $\boldsymbol{v}\in\boldsymbol{V}^\perp(K)$. And we have that $(\delta_{\mathrm{L}},\nabla\boldsymbol{v})_K = 0$ by equation (5.18a) and the first inclusion in condition (5.11a).

Similarly,

$$\begin{aligned}\langle\nu\mathrm{I_L}\boldsymbol{n}-I_p\boldsymbol{n},\boldsymbol{v}\rangle_{\partial K} &= (\nu\nabla\cdot\mathrm{I_L}-\nabla I_p,\boldsymbol{v})_K+(\nu\mathrm{I_L},\nabla\boldsymbol{v})_K-(I_p,\nabla\cdot\boldsymbol{v})_K\\ &= ((\boldsymbol{Id}-\boldsymbol{P}_{\widetilde{\boldsymbol{V}}})(\nu\nabla\cdot\mathrm{L}-\nabla p),\boldsymbol{v})_K.\end{aligned}$$

Indeed, $(\nu\nabla\cdot\mathrm{I_L}-\nabla I_p,\boldsymbol{v})_K = (\nu\nabla\cdot\mathrm{L}-\nabla p,\boldsymbol{v})_K = ((\boldsymbol{Id}-\boldsymbol{P}_{\widetilde{\boldsymbol{V}}})(\nu\nabla\cdot\mathrm{L}-\nabla p),\boldsymbol{v})_K$ by the first inclusion in condition (5.11b), the inclusion condition (5.15a) and the fact that $\boldsymbol{v}\in\boldsymbol{V}^\perp(K)$. Moreover, $(\nu\mathrm{I_L},\nabla\boldsymbol{v})_K = 0$ by the first inclusion in condition (5.11a) and the definition of $\mathrm{I_L}$, and $(I_p,\nabla\cdot\boldsymbol{v})_K = 0$ by the inclusion condition (5.15a) and the definition of I_p. This proves the claim.

Step 3: The estimate of δ_{L}

Finally, let us estimate δ_{L}. By the definition of $\mathrm{G}^\perp(K)$, (5.12b), we see that $\delta_{\mathrm{L}}\in\mathrm{G}^\perp(K)$, by the equation (5.18a). By the condition (5.14a), this implies that

$$\|\delta_{\mathrm{L}}\|_K \le C_{\mathrm{G}}\,h_K^{1/2}\,\|\delta_{\mathrm{L}}\boldsymbol{n}\|_{\partial K_{\mathrm{G}}},$$

and by equation (5.18c), that

$$\|\delta_{\rm L}\|_K \leq C_{\rm G}\, h_K^{1/2}\, \left(\|\nu \mathrm{I}_{\rm G}\boldsymbol{n} - I_p\boldsymbol{n}\|_{\partial K_{\rm G}} + \|\boldsymbol{\alpha}(\boldsymbol{\delta_u})\|_{\partial K_{\rm G}} + \|\boldsymbol{\alpha}(\boldsymbol{I_u})\|_{\partial K_{\rm G}}\right).$$

If $\widetilde{\boldsymbol{V}}(K) = \boldsymbol{V}(K)$, $\boldsymbol{\delta_u} = \boldsymbol{0}$, and we get that

$$\|\delta_{\rm L}\|_K \leq C_{\rm G}\, h_K^{1/2}\, \left(\|\nu \mathrm{I}_{\rm G}\boldsymbol{n} - I_p\boldsymbol{n}\|_{\partial K_{\rm G}} + \|\boldsymbol{\alpha}\|_{\partial K_{\rm G}}\, \|\boldsymbol{I_u}\|_{\partial K_{\rm G}}\right).$$

If $\widetilde{\boldsymbol{V}}(K) \neq \boldsymbol{V}(K)$, then

$$\|\delta_{\rm L}\|_K \leq C_{\rm G}\, h_K^{1/2}\, \Big(\|\nu \mathrm{I}_{\rm G}\boldsymbol{n} - I_p\boldsymbol{n}\|_{\partial K_{\rm G}} + \|\boldsymbol{\alpha}\|_{\partial K_{\rm G}}\big(C_V\, R_{V^\perp}\, \|\boldsymbol{\delta_u}\|_{\partial K_{\rm G}} + \|\boldsymbol{I_u}\|_{\partial K_{\rm G}}\big)\Big),$$

by condition (5.14b). This completes the proof of Theorem 5.3.1.

5.4 Examples of superconvergent methods

In this section, we give examples of superconvergent **HDG** methods. We show how to construct them from similar methods for diffusion problems.

5.4.1 Using the superconvergent methods for diffusion

Let us begin by considering the **HDG** method for diffusion upon which we are going to construct **HDG** methods for Stokes satisfying *Assumptions A, B* and *C*.

Thus, following [3], we assume that we have local spaces

$$\boldsymbol{V}^{\rm D}(K), \quad W^{\rm D}(K) \quad \text{and} \quad M^{\rm D}(F),$$

such that

$$\boldsymbol{V}^{\rm D}(K)\cdot\boldsymbol{n}|_F \subset M^{\rm D}(F), \tag{5.19a}$$

$$W^{\rm D}(K)|_F \subset M^{\rm D}(F). \tag{5.19b}$$

$$\boldsymbol{P}^0(K) \subset \nabla W^{\rm D}(K) \subset \widetilde{\boldsymbol{V}}^{\rm D}(K) \subset \boldsymbol{V}^{\rm D}(K), \tag{5.20a}$$

$$P^0(K) \subset \nabla\cdot\boldsymbol{V}^{\rm D}(K) \subset \widetilde{W}^{\rm D}(K) \subset W^{\rm D}(K), \tag{5.20b}$$

such that, if we set

$$\boldsymbol{V}^{\perp,\mathrm{D}}(K) := \{\boldsymbol{v} \in \boldsymbol{V}^{\mathrm{D}}(K) : (\boldsymbol{v}, \widetilde{\boldsymbol{v}})_K = 0 \quad \forall\, \widetilde{\boldsymbol{v}} \in \widetilde{\boldsymbol{V}}^{\mathrm{D}}(K)\}, \tag{5.21a}$$

$$W^{\perp,\mathrm{D}}(K) := \{w \in W^{\mathrm{D}}(K) : (w, \widetilde{w})_K = 0 \quad \forall\, \widetilde{w} \in \widetilde{W}^{\mathrm{D}}(K)\}, \tag{5.21b}$$

we have that

$$\sum_{F \in \mathcal{F}(K)} \dim M^{\mathrm{D}}(F) = \dim \boldsymbol{V}^{\perp,\mathrm{D}}(K) + \dim W^{\perp,\mathrm{D}}(K), \tag{5.22}$$

and that

$$\|\boldsymbol{v}^{\perp}\|_K \le C_V\, h_K^{1/2} \|\, \boldsymbol{v}^{\perp} \cdot \boldsymbol{n}\,\|_{\partial K_V} \qquad \text{for all } \boldsymbol{v}^{\perp} \in \boldsymbol{V}^{\perp,\mathrm{D}}(K), \tag{5.23a}$$

$$\|w^{\perp}\|_K \le C_W\, h_K^{1/2} \|w^{\perp}\|_{\partial K_W} \qquad \text{for all } w^{\perp} \in W^{\perp,\mathrm{D}}(K), \tag{5.23b}$$

for some subsets ∂K_V and ∂K_W of $\mathcal{F}(K)$.

We also assume that we have a stabilization function α^{D} satisfying the following properties:

$$\langle \alpha^{\mathrm{D}}(\eta),\, \mu\rangle_F = \langle \eta, \alpha^{\mathrm{D}}(\mu)\rangle_F \quad \text{for all } \eta, \mu \in M^{\mathrm{D}}(F), \tag{5.24a}$$

$$\langle \alpha^{\mathrm{D}}(w^{\perp}), w^{\perp}\rangle_{\partial K} \ge C_\alpha\, \|w^{\perp}\|^2_{\partial K_W} \text{ for all } w^{\perp} \in W^{\perp,\mathrm{D}}(K). \tag{5.24b}$$

We can now define the **HDG** methods satisfying *Assumptions A*, *B* and *C*. To do that, we introduce some notation. In what follows, we denote by $\mathrm{G}_i(K)$ the space of all the i-th rows of functions in $\mathrm{G}(K)$, and by $\boldsymbol{V}_i(K)$ and $\boldsymbol{M}_i(F)$ the space of the i-th component of functions in $\boldsymbol{V}(K)$ and $\boldsymbol{M}(F)$, respectively, for $i = 1, \ldots, d$.

We construct the wanted **HDG** method by taking its local spaces as

$$\mathrm{G}_i(K) := \boldsymbol{V}^{\mathrm{D}}(K), \quad \boldsymbol{V}_i(K) := W^{\mathrm{D}}(K), \quad \boldsymbol{M}_i(F) := M^{\mathrm{D}}(F),$$

for $i = 1, \ldots, d$, and by taking the stabilization function as

$$\boldsymbol{\alpha}_i(\boldsymbol{a}) := \alpha^{\mathrm{D}}(a_i),$$

also for $i = 1, \ldots, d$. The choice of the space for the pressure $P(K)$ is more delicate. We have the following result.

Theorem 5.4.1. *Let $P(K)$ be such that*

$$\sum_{j=1}^{d} \partial_j W^D(K) \subset P(K) \subset \cap_{j=1}^{d}\{v_j : \boldsymbol{v} \in \boldsymbol{V}^D(K) : v_i = 0 \text{ for } i \neq j\}.$$

Then the corresponding HDG method satisfies Assumptions A, B and C.

Proof. It suffices to show that the spaces we define above satisfy all the conditions in the template in Section 3.1. In the definition of $\mathrm{G}(K), \boldsymbol{V}(K)$ we can see that the conditions (5.8a), (5.8c), (5.9), (5.11) - (5.14) and (5.16) are nothing but the vector version of the conditions (5.19) - (5.24). We only need to verify the spaces also satisfy (5.8b), (5.15).

The assumption $\sum_{j=1}^{d} \partial_j W^{\mathrm{D}}(K) \subset P(K)$ implies that $\nabla \cdot \boldsymbol{V}(K) \subset P(K)$ and the assumption $P(K) \subset \cap_{j=1}^{d}\{v_j : \boldsymbol{v} \in \boldsymbol{V}^{\mathrm{D}}(K) : v_i = 0 \text{ for } i \neq j\}$ that $P(K)Id \subset \mathrm{G}(K)$. Hence condition (5.15) holds and so

$$(\nabla \cdot \boldsymbol{V}(K))Id \subset P(K)Id \subset \mathrm{G}(K).$$

In other words, condition (5.8b) is verified. $\square$

5.4.2 Examples

We now use this construction to devise superconvergent **HDG** methods for Stokes by using the superconvergent **HDG** methods for diffusion described in detail in [3].

In the Tables 1 to 4, we display the orders of convergence for mixed methods and HDG methods using different elements K. We only show the space $P(K)$ since the other spaces are given in [3]. Note that all the methods achieve optimal orders for L_h, p_h and that superconvergence takes place for the projection of the error in $\boldsymbol{u}$. For $k = 0$ and two dimensions, the method listed in Table 1 as $\mathbf{RT}_k$ is the mixed method proposed in [28], and that the listed in Table 3 as $\mathbf{RT}_{[k]}$ is the mixed method proposed in [29]; as pointed out in the Introduction, general, convex quadrilaterals were also considered therein.

Table 5.1: Methods for which $k \geq 1$

K simplex

method for diffusion	$P(K)$	$\|\mathrm{G}-\mathrm{G}_h\|_{\mathcal{T}_h}$	$\|p-p_h\|_{\mathcal{T}_h}$	$\|\boldsymbol{u}-\boldsymbol{u}_h^\star\|_{\mathcal{T}_h}$
$\mathbf{BDFM}_{k+1}$	$P^k(K)$	$k+1$	$k+1$	$k+2$
$\mathbf{RT}_k$	$P^k(K)$	$k+1$	$k+1$	$k+2$
$\mathbf{HDG}_k$	$P^k(K)$	$k+1$	$k+1$	$k+2$
$\mathbf{BDM}_k$ $k\geq 2$	$P^k(K)$	$k+1$	$k+1$	$k+2$

5.5 Proofs of the estimates of the projection of the errors

In this section we provide detailed proofs for our a priori error estimates. The main idea is to work with the following projection of the errors:

$$\begin{aligned}
\mathrm{E}_{\mathrm{L}} &:= \Pi_G \mathrm{L} - \mathrm{L}_h, \\
\boldsymbol{e}_u &:= \boldsymbol{\Pi}_V \boldsymbol{u} - \boldsymbol{u}_h, \\
e_p &:= \mathit{\Pi}_P p - p_h, \\
\boldsymbol{e}_{\widehat{u}} &:= \boldsymbol{\Pi}_M \boldsymbol{u} - \widehat{\boldsymbol{u}}_h.
\end{aligned}$$

We begin by obtaining the equations satisfied by these projections. We then mimic the argument in [26, 51] to obtain the estimates of $\mathrm{E}_L, e_p, \boldsymbol{e}_u$ and, finally, $\boldsymbol{u}-\boldsymbol{u}_h^*$.

Step 1: The error equations. We begin by obtaining the equations satisfied by the projection of the errors.

Lemma 5.5.1.

$$(\mathrm{E}_{\mathrm{L}}\,,\,\mathrm{G})_{\mathcal{T}_h} + (\boldsymbol{e}_u\,,\,\nabla\cdot\mathrm{G})_{\mathcal{T}_h} - \langle \boldsymbol{e}_{\widehat{u}}\,,\,\mathrm{G}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} = (\Pi_G\mathrm{L}-\mathrm{L}\,,\,\mathrm{G})_{\mathcal{T}_h}, \tag{5.25a}$$

$$-(\nabla\cdot(\nu\mathrm{E}_{\mathrm{L}})\,,\,\boldsymbol{v})_{\mathcal{T}_h} + (\nabla e_p\,,\,\boldsymbol{v})_{\mathcal{T}_h} + \langle \boldsymbol{\alpha}(\boldsymbol{e}_u-\boldsymbol{e}_{\widehat{u}})\,,\,\boldsymbol{v}\rangle_{\partial\mathcal{T}_h} = 0, \tag{5.25b}$$

$$-(\boldsymbol{e}_u\,,\,\nabla q)_{\mathcal{T}_h} + \langle \boldsymbol{e}_{\widehat{u}}\,,\,q\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} = 0, \tag{5.25c}$$

$$\langle \boldsymbol{e}_{\widehat{u}}\,,\,\boldsymbol{\mu}\rangle_{\partial\mathcal{T}_h} = 0, \tag{5.25d}$$

$$\langle \nu\mathrm{E}_{\mathrm{L}}\boldsymbol{n} - e_p\boldsymbol{n} - \boldsymbol{\alpha}(\boldsymbol{e}_u-\boldsymbol{e}_{\widehat{u}})\,,\,\boldsymbol{\mu}\rangle_{\partial\mathcal{T}_h\backslash\partial\Omega} = 0, \tag{5.25e}$$

$$(e_p\,,\,1)_{\mathcal{T}_h} = (\mathit{\Pi}_P p - p\,,\,1)_{\mathcal{T}_h}, \tag{5.25f}$$

for all $(\mathrm{G},\boldsymbol{v},q,\boldsymbol{\mu}) \in \mathrm{G}_h\times\boldsymbol{V}_h\times P_h\times\boldsymbol{M}_h$.

Table 5.2: Methods for which $\boldsymbol{M}(F) = \boldsymbol{P}_k(F)$ and $k \geq 1$

K square

method for diffusion	$P(K)$	$\|\mathrm{G} - \mathrm{G}_h\|_{\mathcal{T}_h}$	$\|p - p_h\|_{\mathcal{T}_h}$	$\|\boldsymbol{u} - \boldsymbol{u}_h^\star\|_{\mathcal{T}_h}$
$\mathbf{BDFM}_{[k+1]}$	$P^k(K)$	$k+1$	$k+1$	$k+2$
$\mathbf{HDG}^P_{[k]}$	$P^k(K)$	$k+1$	$k+1$	$k+2$
$\mathbf{BDM}_{[k]}$ $k \geq 2$	$P^k(K)$	$k+1$	$k+1$	$k+2$

K cube

method for diffusion	$P(K)$	$\|\mathrm{G} - \mathrm{G}_h\|_{\mathcal{T}_h}$	$\|p - p_h\|_{\mathcal{T}_h}$	$\|\boldsymbol{u} - \boldsymbol{u}_h^\star\|_{\mathcal{T}_h}$
$\mathbf{BDFM}_{[k+1]}$	$P^k(K)$	$k+1$	$k+1$	$k+2$
$\mathbf{HDG}^P_{[k]}$	$P^k(K)$	$k+1$	$k+1$	$k+2$
$\mathbf{BDM}_{[k]}$ $k \geq 2$	$P^k(K)$	$k+1$	$k+1$	$k+2$

Proof. We begin by inserting the expression of the numerical trace given in (5.3g) into the second and the fifth equations (5.3b), (5.3e) defining the **HDG** method, then the numerical approximation $(\mathrm{L}_h, \boldsymbol{u}_h, p_h, \widehat{\boldsymbol{u}}_h)$ satisfies:

$$
\begin{aligned}
(\nu \mathrm{L}_h\,,\,\mathrm{G})_{\mathcal{T}_h} + (\boldsymbol{u}_h\,,\,\nabla\cdot\mathrm{G})_{\mathcal{T}_h} - \langle \widehat{\boldsymbol{u}}_h\,,\,\mathrm{G}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} &= 0,\\
(\nu \mathrm{L}_h\,,\,\nabla\boldsymbol{v})_{\mathcal{T}_h} - (p_h\,,\,\nabla\cdot\boldsymbol{v})_{\mathcal{T}_h} - \langle \nu\mathrm{L}_h\boldsymbol{n} - p_h\boldsymbol{n} - \boldsymbol{\alpha}(\boldsymbol{u}_h - \widehat{\boldsymbol{u}}_h)\,,\,\boldsymbol{v}\rangle_{\partial\mathcal{T}_h} &= (\boldsymbol{f}\,,\,\boldsymbol{v})_{\mathcal{T}_h},\\
-(\boldsymbol{u}_h\,,\,\nabla q)_{\mathcal{T}_h} + \langle \widehat{\boldsymbol{u}}_h\cdot\boldsymbol{n}\,,\,q\rangle_{\partial\mathcal{T}_h} &= 0,\\
\langle \widehat{\boldsymbol{u}}_h\,,\,\boldsymbol{\mu}\rangle_{\partial\Omega} &= \langle \boldsymbol{g}\,,\,\boldsymbol{\mu}\rangle_{\partial\Omega},\\
\langle \mu\mathrm{L}_h\boldsymbol{n} - p_h\boldsymbol{n} - \boldsymbol{\alpha}(\boldsymbol{u}_h - \widehat{\boldsymbol{u}}_h)\,,\,\boldsymbol{\mu}\rangle_{\partial\mathcal{T}_h\setminus\partial\Omega} &= 0,\\
(p_h\,,\,1)_{\mathcal{T}_h} &= 0,
\end{aligned}
$$

for all $(\mathrm{G}, \boldsymbol{v}, q, \boldsymbol{\mu}) \in \mathrm{G}_h \times \boldsymbol{V}_h \times P_h \times \boldsymbol{M}_h$. Next, we note that the exact solution satisfies these same equations. Hence, by the *Assumption (A.1) - (A.7)*, we can equip the exact

Table 5.3: Methods for which $\boldsymbol{M}(F) = \boldsymbol{Q}_k(F)$ and $k \geq 1$

K square

method	$P(K)$	$\|\mathrm{G} - \mathrm{G}_h\|_{\mathcal{T}_h}$	$\|p - p_h\|_{\mathcal{T}_h}$	$\|\boldsymbol{u} - \boldsymbol{u}_h^\star\|_{\mathcal{T}_h}$
$\mathbf{RT}_{[k]}$	$Q^k(K)$	$k+1$	$k+1$	$k+2$
$\mathbf{TNT}_{[k]}$	$Q^k(K)$	$k+1$	$k+1$	$k+2$
$\mathbf{HDG}_{[k]}^{Q}$	$Q^k(K)$	$k+1$	$k+1$	$k+2$

K cube

method	$P(K)$	$\|\mathrm{G} - \mathrm{G}_h\|_{\mathcal{T}_h}$	$\|p - p_h\|_{\mathcal{T}_h}$	$\|\boldsymbol{u} - \boldsymbol{u}_h^\star\|_{\mathcal{T}_h}$
$\mathbf{RT}_{[k]}$	$Q^k(K)$	$k+1$	$k+1$	$k+2$
$\mathbf{TNT}_{[k]}$	$Q^k(K)$	$k+1$	$k+1$	$k+2$
$\mathbf{HDG}_{[k]}^{Q}$	$Q^k(K)$	$k+1$	$k+1$	$k+2$

Table 5.4: Methods for which $k \geq 1$

K prism

method for diffusion	$P(K)$	$\|\mathrm{G} - \mathrm{G}_h\|_{\mathcal{T}_h}$	$\|p - p_h\|_{\mathcal{T}_h}$	$\|\boldsymbol{u} - \boldsymbol{u}_h^\star\|_{\mathcal{T}_h}$
$\mathbf{BDFM}_{<k+1>}$	$P^k(K)$	$k+1$	$k+1$	$k+2$
$\mathbf{RT}_{<k>}$	$P^{k\|k}(K)$	$k+1$	$k+1$	$k+2$
$\mathbf{HDG}_{<k>}$	$P^k(K)$	$k+1$	$k+1$	$k+2$

solution with the projection and obtain

$$
\begin{aligned}
(\nu \mathrm{L}\,,\, \mathrm{G})_{\mathcal{T}_h} + (\boldsymbol{\Pi}_V \boldsymbol{u}\,,\, \nabla \cdot \mathrm{G})_{\mathcal{T}_h} - \langle \boldsymbol{\Pi}_M \boldsymbol{u}\,,\, \mathrm{G}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} &= 0,\\
(\nu \Pi_G \mathrm{L}\,,\, \nabla \boldsymbol{v})_{\mathcal{T}_h} - (\mathit{\Pi}_P p\,,\, \nabla \cdot \boldsymbol{v})_{\mathcal{T}_h} - \langle \nu \Pi_G \mathrm{L}\boldsymbol{n} - \mathit{\Pi}_P p \boldsymbol{n} - \boldsymbol{\alpha}(\boldsymbol{\Pi}_V \boldsymbol{u} - \boldsymbol{\Pi}_M \boldsymbol{u})\,,\, \boldsymbol{v}\rangle_{\partial\mathcal{T}_h}&\\
&= (\boldsymbol{f}\,,\, \boldsymbol{v})_{\mathcal{T}_h},\\
-(\boldsymbol{\Pi}_V \boldsymbol{u}\,,\, \nabla q)_{\mathcal{T}_h} + \langle \boldsymbol{\Pi}_M \boldsymbol{u} \cdot \boldsymbol{n}\,,\, q\rangle_{\partial\mathcal{T}_h} &= 0,\\
\langle \boldsymbol{\Pi}_M \boldsymbol{u}\,,\, \boldsymbol{\mu}\rangle_{\partial\Omega} &= \langle \boldsymbol{g}\,,\, \boldsymbol{\mu}\rangle_{\partial\Omega},\\
\langle \mu \Pi_G \mathrm{L}\boldsymbol{n} - \mathit{\Pi}_P p \boldsymbol{n} - \boldsymbol{\alpha}(\boldsymbol{\Pi}_V \boldsymbol{u} - \boldsymbol{\Pi}_M \boldsymbol{u})\,,\, \boldsymbol{\mu}\rangle_{\partial\mathcal{T}_h \setminus \partial\Omega} &= 0,\\
(p\,,\, 1)_{\mathcal{T}_h} &= 0,
\end{aligned}
$$

for all $(\mathrm{G}, \boldsymbol{v}, q, \boldsymbol{\mu}) \in \mathrm{G}_h \times \boldsymbol{V}_h \times P_h \times \boldsymbol{M}_h$. If we now subtract the first set of equations from this one, we obtain the result. This completes the proof of Lemma 5.5.1. □

Step 2: Estimate of the velocity gradient. We are now ready to obtain our first estimate by using a standard energy argument.

Proposition 5.5.2. *We have*

$$\nu \|\mathrm{E_L}\|^2_{\mathcal{T}_h} + \langle \boldsymbol{\alpha}(\boldsymbol{e}_u - \boldsymbol{e}_{\widehat{u}}) \,,\, \boldsymbol{e}_u - \boldsymbol{e}_{\widehat{u}} \rangle_{\partial \mathcal{T}_h} = \nu (\Pi_G \mathrm{L} - \mathrm{L} \,,\, \mathrm{E_L})_{\mathcal{T}_h}.$$

Proof. If we take $\mathrm{G} = \nu \mathrm{E_L}, \boldsymbol{v} = \boldsymbol{e}_u$ and $q = e_p$ in the first three error equations (5.25a) - (5.25c), respectively, and add them up. We obtain, after some algebraic manipulation,

$$\nu \|\mathrm{E_L}\|^2_{\mathcal{T}_h} + \Theta_h = \nu (\Pi_G \mathrm{L} - \mathrm{L} \,,\, \mathrm{E_L})_{\mathcal{T}_h},$$

where

$$\Theta_h := -\langle \boldsymbol{e}_{\widehat{u}} \,,\, \nu \mathrm{E_L} \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} + \langle \boldsymbol{\alpha}(\boldsymbol{e_u} - \boldsymbol{e_{\widehat{u}}}) \,,\, \boldsymbol{e}_u \rangle_{\partial \mathcal{T}_h} + \langle \boldsymbol{e}_{\widehat{u}} \,,\, e_p \boldsymbol{n} \rangle_{\partial \mathcal{T}_h}.$$

We thus obtain

$$\begin{aligned}
\Theta_h &= \langle -\nu \mathrm{E_L} \boldsymbol{n} + e_p \boldsymbol{n} \,,\, \boldsymbol{e}_{\widehat{u}} \rangle_{\partial \mathcal{T}_h} + \langle \boldsymbol{\alpha}(\boldsymbol{e}_u - \boldsymbol{e}_{\widehat{u}}) \,,\, \boldsymbol{e}_u \rangle_{\partial \mathcal{T}_h} \\
&= \langle -\nu \mathrm{E_L} \boldsymbol{n} + e_p \boldsymbol{n} + \boldsymbol{\alpha}(\boldsymbol{e}_u - \boldsymbol{e}_{\widehat{u}}) \,,\, \boldsymbol{e}_{\widehat{u}} \rangle_{\partial \mathcal{T}_h} + \langle \boldsymbol{\alpha}(\boldsymbol{e}_u - \boldsymbol{e}_{\widehat{u}}) \,,\, \boldsymbol{e}_u - \boldsymbol{e}_{\widehat{u}} \rangle_{\partial \mathcal{T}_h} \\
&= \langle \boldsymbol{\alpha}(\boldsymbol{e}_u - \boldsymbol{e}_{\widehat{u}}) \,,\, \boldsymbol{e}_u - \boldsymbol{e}_{\widehat{u}} \rangle_{\partial \mathcal{T}_h},
\end{aligned}$$

by the error equations (5.25d) and (5.25e). This completes the proof. □

As a straightforward consequence of Proposition 5.5.2 and the *Assumption (A.8)*, we obtain the first error estimate for the methods:

$$\|\mathrm{E_L}\|_{\mathcal{T}_h} \leq \|\Pi_G \mathrm{L} - \mathrm{L}\|_{\mathcal{T}_h}.$$

This completes the proof of Theorem 5.2.1. Moreover, we also have the following estimate,

$$\|\boldsymbol{e}_u - \boldsymbol{e}_{\widehat{u}}\|_{\boldsymbol{\alpha}/\nu} \leq \|\mathrm{L} - \Pi_G \mathrm{L}\|_{\mathcal{T}_h}. \tag{5.26}$$

Step 3: Estimate of the pressure. Next, we show how to use the previous result to obtain the estimate of the pressure.

Proposition 5.5.3. *Let $\boldsymbol{P} : \boldsymbol{H}^1(\mathcal{T}_h) \to \boldsymbol{V}_h$ be any projection such that $(\boldsymbol{P}\boldsymbol{w} - \boldsymbol{w}, \boldsymbol{v})_{\mathcal{T}_h} = 0$ for all $\boldsymbol{v} \in \nabla P_h$. Then we have*

$$\|e_p - \overline{(p - \mathit{\Pi}_P p)}\|_{\mathcal{T}_h} \leq C C_{p,\alpha/\nu} \nu \|\mathrm{L} - \Pi_G \mathrm{L}\|_{\mathcal{T}_h},$$

where C solely depends on the shape of the domain Ω, and $C_{p,\alpha/\nu}$ is defined in Theorem 5.2.2.

Proof. It is well known [8] that for any function $q \in L^2(\Omega)$ such that $(q, 1)_\Omega = 0$ we have

$$\|q\|_\Omega \leq \kappa \sup_{\boldsymbol{w} \in \boldsymbol{H}_0^1(\Omega) \setminus \{\boldsymbol{0}\}} \frac{(q, \nabla \cdot \boldsymbol{w})_\Omega}{\|\boldsymbol{w}\|_{\boldsymbol{H}^1(\Omega)}},$$

for some constant κ independent of q. By the last error equation (5.25f), we see that we can apply the above result to $q := e_p - \overline{e_p}$. Hence we have that

$$\|e_p - \overline{e_p}\|_\Omega \leq \kappa \sup_{\boldsymbol{w} \in \boldsymbol{H}_0^1(\Omega) \setminus \{\boldsymbol{0}\}} \frac{(e_p - \overline{e_p}, \nabla \cdot \boldsymbol{w})_\Omega}{\|\boldsymbol{w}\|_{\boldsymbol{H}^1(\Omega)}}.$$

Next, we work on the numerator in the above expression. We have

$$(e_p, \nabla \cdot \boldsymbol{w})_\Omega = -(\nabla e_p \,, \boldsymbol{P}\boldsymbol{w})_{\mathcal{T}_h} + \langle e_p \,, \boldsymbol{w} \cdot \boldsymbol{n} \rangle_{\partial \mathcal{T}_h}.$$

By the second error equation (5.25b) with $\boldsymbol{v} := \boldsymbol{P}\boldsymbol{w}$, we get that

$$\begin{aligned}
(e_p, \nabla \cdot \boldsymbol{w}) &= -(\nabla \cdot (\nu \mathrm{E}_L) \,, \boldsymbol{P}\boldsymbol{w})_{\mathcal{T}_h} + \langle \boldsymbol{\alpha}(\boldsymbol{e}_u - \boldsymbol{e}_{\widehat{u}}) \,, \boldsymbol{P}\boldsymbol{w} \rangle_{\partial \mathcal{T}_h} + \langle e_p \,, \boldsymbol{w} \cdot \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} \\
&= (\nu \mathrm{E}_L \,, \nabla \boldsymbol{w})_{\mathcal{T}_h} + \langle \boldsymbol{\alpha}(\boldsymbol{e}_u - \boldsymbol{e}_{\widehat{u}}) \,, \boldsymbol{P}\boldsymbol{w} \rangle_{\partial \mathcal{T}_h} + \langle -\nu \mathrm{E}_L \boldsymbol{n} + e_p \boldsymbol{n} \,, \boldsymbol{\Pi}_M \boldsymbol{w} \rangle_{\partial \mathcal{T}_h} \\
&= (\nu \mathrm{E}_L \,, \nabla \boldsymbol{w})_{\mathcal{T}_h} + \langle \boldsymbol{\alpha}(\boldsymbol{e}_u - \boldsymbol{e}_{\widehat{u}}) \,, \boldsymbol{P}\boldsymbol{w} - \boldsymbol{\Pi}_M \boldsymbol{w} \rangle_{\partial \mathcal{T}_h},
\end{aligned}$$

by the fifth error equation (5.25e) with $\mu = \boldsymbol{\Pi}_M \boldsymbol{w}$ and by the fact that $\boldsymbol{w} \in \boldsymbol{H}_0^1(\Omega)$.

By the Cauchy-Schwarz inequality, we get that

$$\begin{aligned}
|(e_p, \nabla \cdot \boldsymbol{w})| &\leq \nu \|\mathrm{E}_L\|_{\mathcal{T}_h} \|\nabla \boldsymbol{w}\|_\Omega + \|\boldsymbol{e}_u - \boldsymbol{e}_{\widehat{u}}\|_{\boldsymbol{\alpha}/\nu} \frac{\langle \boldsymbol{\alpha}(\boldsymbol{e}_u - \boldsymbol{e}_{\widehat{u}}) \,, \boldsymbol{P}\boldsymbol{w} - \boldsymbol{\Pi}_M \boldsymbol{w} \rangle_{\partial \mathcal{T}_h}}{\|\boldsymbol{e}_u - \boldsymbol{e}_{\widehat{u}}\|_{\boldsymbol{\alpha}/\nu}} \\
&\leq \nu \, \mathbb{C} \left[\|\mathrm{E}_L\|_{\mathcal{T}_h} + \|\boldsymbol{e}_u - \boldsymbol{e}_{\widehat{u}}\|_{\boldsymbol{\alpha}/\nu} \right] \|\boldsymbol{w}\|_{\boldsymbol{H}^1(\Omega)}
\end{aligned}$$

where

$$\mathbb{C} := \max\{1, \frac{\langle \boldsymbol{\alpha}(\boldsymbol{e}_u - \boldsymbol{e}_{\widehat{u}})/\nu \,, \boldsymbol{P}\boldsymbol{w} - \boldsymbol{\Pi}_M \boldsymbol{w} \rangle_{\partial \mathcal{T}_h}}{\|\boldsymbol{e}_u - \boldsymbol{e}_{\widehat{u}}\|_{\boldsymbol{\alpha}/\nu} \, \|\boldsymbol{w}\|_{\boldsymbol{H}^1(\Omega)}}\}.$$

Now, by the estimates (5.26),

$$|(e_p, \nabla \cdot \boldsymbol{w})| \leq \nu\, \mathbb{C}\, \|\mathrm{L} - \Pi_G \mathrm{L}\|_{\mathcal{T}_h}\, \|\boldsymbol{w}\|_{\boldsymbol{H}^1(\Omega)} \leq \nu\, C_{p,\alpha/\nu}\, \|\mathrm{L} - \Pi_G \mathrm{L}\|_{\mathcal{T}_h}\, \|\boldsymbol{w}\|_{\boldsymbol{H}^1(\Omega)},$$

provided that $\boldsymbol{e}_u - \boldsymbol{e}_{\widehat{u}}$ lies in the auxiliary space $\mathcal{M}^{\perp}_{n,h}$ given by (5.5). We claim that this is indeed the case. To see that, note that for any face F of the element $K \in \Omega_h$, we have that $(\boldsymbol{e}_u - \boldsymbol{e}_{\widehat{u}})|_F \in \boldsymbol{M}(F)$ by *Assumption (A.6)*. Also, for any $q \in P(K)$ such that $(q, \nabla \cdot \boldsymbol{v})_K = 0$ whenever $\boldsymbol{v} \in \boldsymbol{V}(K)$, we have that

$$\langle (\boldsymbol{e}_u - \boldsymbol{e}_{\widehat{u}}) \cdot \boldsymbol{n}, q \rangle_{\partial K} = (\nabla \cdot \boldsymbol{e}_u, q) = 0,$$

by the error equation (5.25c). This proves the claim.

As a consequence, we obtain that

$$\|e_p - \overline{e_p}\|_{\mathcal{T}_h} \leq C C_{p,\alpha/\nu} \nu \|\mathrm{L} - \Pi_G \mathrm{L}\|_{\mathcal{T}_h}.$$

Finally, the result follows from the fact that $\overline{e_p} = \overline{\mathit{\Pi}_P p - p}$ by the error equation (5.25f). This completes the proof of Proposition 5.5.3. □

Theorem 5.2.2 follows directly from the above result.

Step 4: Estimate of the velocity. We are now ready to obtain a key identity for the projection of the error in the velocity by using a duality argument.

Lemma 5.5.4. *We have*

$$(\boldsymbol{e}_{\boldsymbol{u}}\,,\, \boldsymbol{\theta})_{\mathcal{T}_h} = \nu(\Pi_G \mathrm{L} - \mathrm{L}\,,\, \mathrm{Z})_{\mathcal{T}_h} + \nu(\mathrm{L} - \mathrm{L}_h\,,\, \mathrm{Z} - \Pi_G^* \mathrm{Z})_{\mathcal{T}_h}.$$

Proof. We have

$$(\boldsymbol{e}_u\,,\, \boldsymbol{\theta})_{\mathcal{T}_h} = \nu(\mathrm{E}_L\,,\, \mathrm{Z} - \nabla \boldsymbol{\sigma})_{\mathcal{T}_h} + (\boldsymbol{e}_u\,,\, \nabla \cdot (\nu \mathrm{Z}) - \nabla \eta)_{\mathcal{T}_h} + (e_p\,,\, \nabla \cdot \boldsymbol{\sigma})_{\mathcal{T}_h},$$

by the first three equations of the dual problem (5.6). Rearranging terms, we get

$$
\begin{aligned}
(\boldsymbol{e}_u\,,\,\boldsymbol{\theta})_{\mathcal{T}_h} &= (\mathrm{E}_L\,,\,\nu\mathrm{Z})_{\mathcal{T}_h} + (\boldsymbol{e}_u\,,\,\nabla\cdot(\nu\mathrm{Z}))_{\mathcal{T}_h}\\
&\quad-(\nu\mathrm{E}_L\,,\,\nabla\boldsymbol{\sigma})_{\mathcal{T}_h} + (e_p\,,\,\nabla\cdot\boldsymbol{\sigma})_{\mathcal{T}_h}\\
&\quad-(\boldsymbol{e}_u\,,\,\nabla\eta)_{\mathcal{T}_h}\\
&= (\mathrm{E}_L\,,\,\nu\Pi_G^*\mathrm{Z})_{\mathcal{T}_h} + (\boldsymbol{e}_u\,,\,\nabla\cdot(\nu\Pi_G^*\mathrm{Z}))_{\mathcal{T}_h}\\
&\quad-(\nu\mathrm{E}_L\,,\,\nabla\boldsymbol{\sigma})_{\mathcal{T}_h} + (e_p\,,\,\nabla\cdot\boldsymbol{\sigma})_{\mathcal{T}_h}\\
&\quad-(\boldsymbol{e}_u\,,\,\nabla\mathit{\Pi}_P^*\eta)_{\mathcal{T}_h}\\
&\quad+(\mathrm{E}_L\,,\,\nu(\mathrm{Z}-\Pi_G^*\mathrm{Z}))_{\mathcal{T}_h} + (\boldsymbol{e}_u\,,\,\nu\nabla\cdot(\mathrm{Z}-\Pi_G^*\mathrm{Z}))_{\mathcal{T}_h} - (\boldsymbol{e}_u\,,\,\nabla(\eta-\mathit{\Pi}_P^*\eta))_{\mathcal{T}_h},
\end{aligned}
$$

integrating by parts of above equation and by the *Assumption A*, we obtain

$$
\begin{aligned}
(\boldsymbol{e}_u\,,\,\boldsymbol{\theta})_{\mathcal{T}_h} &= (\mathrm{E}_L\,,\,\nu\Pi_G^*\mathrm{Z})_{\mathcal{T}_h} + (\boldsymbol{e}_u\,,\,\nabla\cdot(\nu\Pi_G^*\mathrm{Z}))_{\mathcal{T}_h}\\
&\quad+(\nabla\cdot(\nu\mathrm{E}_L)\,,\,\boldsymbol{\Pi}_V^*\boldsymbol{\sigma})_{\mathcal{T}_h} - (\nabla e_p\,,\,\boldsymbol{\Pi}_V^*\boldsymbol{\sigma})_{\mathcal{T}_h} - \langle \nu\mathrm{E}_L\boldsymbol{n} - e_p\boldsymbol{n}\,,\,\boldsymbol{\sigma}\rangle_{\partial\mathcal{T}_h}\\
&\quad-(\boldsymbol{e}_u\,,\,\nabla\mathit{\Pi}_P^*\eta)_{\mathcal{T}_h}\\
&\quad+(\mathrm{E}_L\,,\,\nu(\mathrm{Z}-\Pi_G^*\mathrm{Z}))_{\mathcal{T}_h} + \langle\boldsymbol{e}_u\,,\,\nu(\mathrm{Z}-\Pi_G^*\mathrm{Z})\boldsymbol{n} - (\eta-\mathit{\Pi}_P^*\eta)\boldsymbol{n}\rangle_{\partial\mathcal{T}_h},
\end{aligned}
$$

taking $\mathrm{G} = \nu\Pi_G^*\mathrm{Z}, \boldsymbol{v} = \boldsymbol{\Pi}_V^*\boldsymbol{\sigma}, q = \mathit{\Pi}_P^*\eta$ in the error equation (5.25a)-(5.25c) and inserting in the above identity, we get

$$
(\boldsymbol{e}_u\,,\,\boldsymbol{\theta})_{\mathcal{T}_h} = \nu(\mathrm{E}_L\,,\,\mathrm{Z}-\Pi_G^*\mathrm{Z})_{\mathcal{T}_h} + \nu(\Pi_G\mathrm{L}-\mathrm{L}\,,\,\Pi_G^*\mathrm{Z})_{\mathcal{T}_h} + \mathbb{T}_1 + \mathbb{T}_2,
$$

where

$$
\begin{aligned}
\mathbb{T}_1 &:= \langle\boldsymbol{e}_{\widehat{u}}\,,\,\Pi_G^*\mathrm{Z}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle\boldsymbol{e}_{\widehat{u}}\,,\,\mathit{\Pi}_P^*\eta\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} + \langle\boldsymbol{e}_u\,,\,\nu(\mathrm{Z}-\Pi_G^*\mathrm{Z})\boldsymbol{n} - (\eta-\mathit{\Pi}_P^*\eta)\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}\\
\mathbb{T}_2 &:= \langle\boldsymbol{\alpha}(\boldsymbol{e}_u-\boldsymbol{e}_{\widehat{u}})\,,\,\boldsymbol{\Pi}_V^*\boldsymbol{\sigma}\rangle_{\partial\mathcal{T}_h} - \langle\nu\mathrm{E}_L\boldsymbol{n} - e_p\boldsymbol{n}\,,\,\boldsymbol{\sigma}\rangle_{\partial\mathcal{T}_h}.
\end{aligned}
$$

Noting that $\langle\boldsymbol{e}_{\widehat{u}}\,,\,\mathrm{Z}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} = \langle\boldsymbol{e}_{\widehat{u}}\,,\,\mathrm{Z}\boldsymbol{n}\rangle_{\partial\Omega} = 0$ and that $\langle\boldsymbol{e}_{\widehat{u}}\,,\,\eta\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} = \langle\boldsymbol{e}_{\widehat{u}}\,,\,\eta\boldsymbol{n}\rangle_{\partial\Omega} = 0$, and inserting this two identities into $\mathbb{T}_1$, we obtain,

$$
\begin{aligned}
\mathbb{T}_1 &:= -\langle\boldsymbol{e}_{\widehat{u}}\,,\,(\mathrm{Z}-\Pi_G^*\mathrm{Z})\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} + \langle\boldsymbol{e}_{\widehat{u}}\,,\,(\eta-\mathit{\Pi}_P^*\eta)\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}\\
&\quad+\langle\boldsymbol{e}_u\,,\,\nu(\mathrm{Z}-\Pi_G^*\mathrm{Z})\boldsymbol{n} - (\eta-\mathit{\Pi}_P^*\eta)\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}\\
&= \langle\boldsymbol{e}_u-\boldsymbol{e}_{\widehat{u}}\,,\,\nu(\mathrm{Z}-\Pi_G^*\mathrm{Z})\boldsymbol{n} - (\eta-\mathit{\Pi}_P^*\eta)\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}\\
&= \langle\boldsymbol{e}_u-\boldsymbol{e}_{\widehat{u}}\,,\,\boldsymbol{\alpha}^*(\boldsymbol{\Pi}_M\boldsymbol{\sigma} - \boldsymbol{\Pi}_V^*\boldsymbol{\sigma})\rangle_{\partial\mathcal{T}_h} \qquad \text{by } \textit{Assumption (A.4)},\\
&= \langle\boldsymbol{\alpha}(\boldsymbol{e}_u-\boldsymbol{e}_{\widehat{u}})\,,\,\boldsymbol{\Pi}_M\boldsymbol{\sigma} - \boldsymbol{\Pi}_V^*\boldsymbol{\sigma}\rangle_{\partial\mathcal{T}_h}.
\end{aligned}
$$

On the other hand, we have

$$\begin{aligned}\mathbb{T}_2 &= \langle \boldsymbol{\alpha}(\boldsymbol{e}_u - \boldsymbol{e}_{\widehat{u}}) \,,\, \boldsymbol{\Pi}_V^* \boldsymbol{\sigma} \rangle_{\partial\mathcal{T}_h} - \langle \nu \mathrm{E}_L \boldsymbol{n} - e_p \boldsymbol{n} \,,\, \boldsymbol{\Pi}_M \boldsymbol{\sigma} \rangle_{\partial\mathcal{T}_h} \\ &= \langle \boldsymbol{\alpha}(\boldsymbol{e}_u - \boldsymbol{e}_{\widehat{u}}) \,,\, \boldsymbol{\Pi}_V^* \boldsymbol{\sigma} \rangle_{\partial\mathcal{T}_h} - \langle \boldsymbol{\alpha}(\boldsymbol{e}_u - \boldsymbol{e}_{\widehat{u}}) \,,\, \boldsymbol{\Pi}_M \boldsymbol{\sigma} \rangle_{\partial\mathcal{T}_h},\end{aligned}$$

by the error equation (5.25e) and by the fact that $\boldsymbol{\Pi}_M \boldsymbol{\sigma}|_{\partial\Omega} = 0$ by the boundary condition for $\boldsymbol{\sigma}$ (5.6d). Therefore, we get

$$\mathbb{T}_2 = -\langle \boldsymbol{\alpha}(\boldsymbol{e_u} - \boldsymbol{e_{\widehat{u}}}) \,,\, \boldsymbol{\Pi}_V^* \boldsymbol{\sigma} - \boldsymbol{\Pi}_M \boldsymbol{\sigma} \rangle_{\partial\mathcal{T}_h}.$$

This implies that

$$\begin{aligned}(\boldsymbol{e}_u \,,\, \boldsymbol{\theta})_{\mathcal{T}_h} &= \nu(\mathrm{E}_L \,,\, \mathrm{Z} - \Pi_G^* \mathrm{Z})_{\mathcal{T}_h} + \nu(\Pi_G \mathrm{L} - \mathrm{L} \,,\, \Pi_G^* \mathrm{Z})_{\mathcal{T}_h} \\ &= \nu(\mathrm{E}_L \,,\, \mathrm{Z} - \Pi_G^* \mathrm{Z})_{\mathcal{T}_h} - \nu(\Pi_G \mathrm{L} - \mathrm{L} \,,\, \mathrm{Z} - \Pi_G^* \mathrm{Z})_{\mathcal{T}_h} + \nu(\Pi_G \mathrm{L} - \mathrm{L} \,,\, \mathrm{Z})_{\mathcal{T}_h} \\ &= \nu(\Pi_G \mathrm{L} - \mathrm{L} \,,\, \mathrm{Z})_{\mathcal{T}_h} + \nu(\mathrm{L} - \mathrm{L}_h \,,\, \mathrm{Z} - \Pi_G^* \mathrm{Z})_{\mathcal{T}_h}.\end{aligned}$$

This completes the proof. $\square$

As a consequence, we immediately obtain an estimate on $\boldsymbol{e}_u$.

Corollary 5.5.5. *If the regularity property* (5.7) *holds and Assumption A, B are satisfied, then we have*

$$\|\boldsymbol{e}_u\|_{\mathcal{T}_h} \leq C\nu h \|\mathrm{L} - \Pi_G \mathrm{L}\|_{\mathcal{T}_h}$$

Proof. Taking $\boldsymbol{\theta} = \boldsymbol{e}_u$ in Lemma 5.5.4, we have

$$\begin{aligned}\|\boldsymbol{e}_u\|_{\mathcal{T}_h}^2 &= \nu(\Pi_G \mathrm{L} - \mathrm{L} \,,\, \mathrm{Z})_{\mathcal{T}_h} + \nu(\mathrm{L} - \mathrm{L}_h \,,\, \mathrm{Z} - \Pi_G^* \mathrm{Z})_{\mathcal{T}_h} \\ &= \nu(\Pi_G \mathrm{L} - \mathrm{L} \,,\, \mathrm{Z} - \mathrm{Z}^0)_{\mathcal{T}_h} + \nu(\mathrm{L} - \mathrm{L}_h \,,\, \mathrm{Z} - \Pi_G^* \mathrm{Z})_{\mathcal{T}_h},\end{aligned}$$

by *Assumption (A.1)* and *Assumption (B.2)*, since $\mathrm{Z}^0|_K$ is the average of Z on the element K and hence belongs to $\mathrm{P}^0(K)$.

Now, after a simple application of the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned}\|\boldsymbol{e}_u\|_{\mathcal{T}_h}^2 &\leq C\nu h \|\mathrm{L} - \Pi_G \mathrm{L}\|_{\mathcal{T}_h} |\mathrm{Z}|_{1,\mathcal{T}_h} + \nu \|\mathrm{L} - \mathrm{L}_h\|_{\mathcal{T}_h} \|\mathrm{Z} - \Pi_G^* \mathrm{Z}\|_{\mathcal{T}_h} \\ &\leq C\nu h \|\mathrm{L} - \Pi_G \mathrm{L}\|_{\mathcal{T}_h} \|\boldsymbol{e}_u\|_{\mathcal{T}_h},\end{aligned}$$

by *Assumption (B.1)* and the regularity property (5.7). This completes the proof. $\square$

Next, we can obtain the following simple estimate.

Corollary 5.5.6. *By the same assumption in Corollary 5.5.5, we have*

$$\|\boldsymbol{e}_{\widehat{u}}\|_h \le C(h\|\mathrm{L}-\Pi_G\mathrm{L}\|_{\mathcal{T}_h} + \|\boldsymbol{e}_u\|_{\mathcal{T}_h})$$

Proof. From the first error equation (5.25a), we have that

$$\langle \boldsymbol{e}_{\widehat{u}}\,, \mathrm{G}\boldsymbol{n}\rangle_{\partial K} = -(\Pi_G\mathrm{L}-\mathrm{L},\mathrm{G})_K + (\mathrm{E}_L,\mathrm{G})_K + (\boldsymbol{e}_u, \nabla\cdot\mathrm{G})_K,$$

for all $\mathrm{G}\in\mathrm{G}(K)$. Hence, by a standard scaling argument, see [15], we readily obtain that

$$h_K^{\frac{1}{2}}\|\boldsymbol{e}_{\widehat{u}}\|_{\partial K} \le C(h_K\|\mathrm{L}-\Pi_G\mathrm{L}\|_K + h_K\|\mathrm{E}_L\|_K + \|\boldsymbol{e}_u\|_K),$$

and the estimate follows by using Theorem 5.2.1. □

Finally, Theorem 5.2.3 follows by the above two corollaries.

Step 5: Estimate of the postprocessed velocity. By the Poincaré-Friedrichs inequality, we have that

$$\|\boldsymbol{u}-\boldsymbol{u}_h^*\|_K \le \|\overline{\boldsymbol{u}-\boldsymbol{u}_h^*}\|_K + C\,h_K\,\|\nabla(\boldsymbol{u}-\boldsymbol{u}_h^*)\|_K,$$

where $\overline{w}$ is the average of w over K. But $\overline{\boldsymbol{u}_h^*}=\overline{\boldsymbol{u}_h}$, by the second equation defining $\boldsymbol{u}_h^*$, (5.4b), and $\overline{\boldsymbol{u}}=\overline{\boldsymbol{\Pi}_V\boldsymbol{u}}$ by Assumptions (*A.2*) and (*C.1*). This implies that

$$\|\boldsymbol{u}-\boldsymbol{u}_h^*\|_K \le \|\boldsymbol{\Pi}_V\boldsymbol{u}-\boldsymbol{u}_h\|_K + C\,h_K\,\|\nabla(\boldsymbol{u}-\boldsymbol{u}_h^*)\|_K.$$

Now, for any $\boldsymbol{w}\in\boldsymbol{W}^*(K)$, we have that

$$\begin{aligned}\|\nabla(\boldsymbol{u}-\boldsymbol{u}_h^*)\|_K^2 &= (\nabla(\boldsymbol{u}-\boldsymbol{u}_h^*),\nabla(\boldsymbol{u}-\boldsymbol{w}))_K + (\nabla(\boldsymbol{u}-\boldsymbol{u}_h^*),\nabla(\boldsymbol{w}-\boldsymbol{u}_h^*))_K\\ &= (\nabla(\boldsymbol{u}-\boldsymbol{u}_h^*),\nabla(\boldsymbol{u}-\boldsymbol{w}))_K + (\mathrm{L}-\mathrm{L}_h,\nabla(\boldsymbol{w}-\boldsymbol{u}_h^*))_K,\end{aligned}$$

by the first equation defining the postprocessing u_h^*, (5.4a). Applying the Cauchy-Schwarz inequality, we obtain that

$$\begin{aligned}\|\nabla(\boldsymbol{u}-\boldsymbol{u}_h^*)\|_K^2 &\le \|\nabla(\boldsymbol{u}-\boldsymbol{u}_h^*)\|_K\,\|\nabla(\boldsymbol{u}-\boldsymbol{w})\|_K + \|\mathrm{L}-\mathrm{L}_h\|_K\,\|\nabla(\boldsymbol{w}-\boldsymbol{u}_h^*)\|_K,\\ &\le \|\nabla(\boldsymbol{u}-\boldsymbol{u}_h^*)\|_K\,(\|\nabla(\boldsymbol{u}-\boldsymbol{w})\|_K + \|\mathrm{L}-\mathrm{L}_h\|_K)\\ &\qquad + \|\mathrm{L}-\mathrm{L}_h\|_K\,\|\nabla(\boldsymbol{w}-\boldsymbol{u})\|_K,\end{aligned}$$

and, after simple applications of Young's inequality and some algebraic manipulations, we get that

$$\|\nabla(\boldsymbol{u}-\boldsymbol{u}_h^*)\|_K^2 \leq 3(\|\mathrm{L}-\mathrm{L}_h\|_K^2 + \|\nabla(\boldsymbol{u}-\boldsymbol{w})\|_K^2).$$

This implies that

$$\|\boldsymbol{u}-\boldsymbol{u}_h^*\|_K \leq \|\boldsymbol{\Pi}_V\boldsymbol{u}-\boldsymbol{u}_h\|_K + C\,h_K\,(\|\mathrm{L}-\mathrm{L}_h\|_K + \|\nabla(\boldsymbol{u}-\boldsymbol{w})\|_K),$$

and so,

$$\|\boldsymbol{u}-\boldsymbol{u}_h^*\|_{\mathcal{T}_h} \leq \|\boldsymbol{\Pi}_V\boldsymbol{u}-\boldsymbol{u}_h\|_{\mathcal{T}_h} + C\,h\,(\|\mathrm{L}-\mathrm{L}_h\|_{\mathcal{T}_h} + \|\nabla(\boldsymbol{u}-\boldsymbol{w})\|_{\mathcal{T}_h}).$$

This completes the proof of Theorem 5.2.4.

5.6 Concluding remarks

We have presented a technique for the a priori error analysis of **HDG** methods for Stokes equations and have shown how to use it to devise superconvergent methods from similar methods for the simple diffusion equation. Next, we discuss how to recover superconvergence when the space $\boldsymbol{V}(K)$ does not satisfy *Assumption (B.2)*. We then sketch the extension of the method to isotropic elasticity.

Superconvergence when the local space $\boldsymbol{V}(K)$ is too small

Let us consider the method listed in Table 1 as $\mathbf{RT_k}$, or the one listed in Table 3 as $\mathbf{RT_{[k]}}$, for $k=0$. In this cases, the *Assumption (B.2)* is not satisfied because $\boldsymbol{V}(K)=\boldsymbol{P}_0(K)$. However, we can still obtain superconvergence of the projection of the error in the velocity as we show next.

Then, by the identity of Lemma 5.5.4:

$$\begin{aligned}
(\boldsymbol{e_u}\,,\boldsymbol{\theta})_{\mathcal{T}_h} &= \nu(\Pi_G\mathrm{L}-\mathrm{L}\,,\mathrm{Z})_{\mathcal{T}_h} + \nu(\mathrm{L}-\mathrm{L}_h\,,\mathrm{Z}-\Pi_G^*\mathrm{Z})_{\mathcal{T}_h} \\
&= \nu(\Pi_G\mathrm{L}-\mathrm{L}\,,\nabla\boldsymbol{\sigma})_{\mathcal{T}_h} + \nu(\mathrm{L}-\mathrm{L}_h\,,\mathrm{Z}-\Pi_G^*\mathrm{Z})_{\mathcal{T}_h} && \text{by (5.6a)},\\
&= -\nu(\nabla\cdot(\Pi_G\mathrm{L}-\mathrm{L})\,,\boldsymbol{\sigma})_{\mathcal{T}_h} + \nu(\mathrm{L}-\mathrm{L}_h\,,\mathrm{Z}-\Pi_G^*\mathrm{Z})_{\mathcal{T}_h} && \text{by (5.6d)},\\
&= -\nu(\overline{\nabla\cdot\mathrm{L}}-\nabla\cdot\mathrm{L}\,,\boldsymbol{\sigma}-\overline{\boldsymbol{\sigma}})_{\mathcal{T}_h} + \nu(\mathrm{L}-\mathrm{L}_h\,,\mathrm{Z}-\Pi_G^*\mathrm{Z})_{\mathcal{T}_h},
\end{aligned}$$

and, proceeding as before, we can get that

$$\|\boldsymbol{e_u}\|_{\mathcal{T}_h} \leq C\,h^2\,\|\nabla\cdot\mathrm{L}\|_{\boldsymbol{H}^1(\Omega)} + C\,h\,\|\mathrm{L}-\Pi_G\mathrm{L}\|_{\mathcal{T}_h},$$

and we obtain the wanted superconvergence. As a consequence, the postprocessing $\boldsymbol{u}_h^*$ also satisfies the estimate of Theorem 5.2.4 and hence converges with order two whenever the solution is smooth enough. These results complement the error estimates obtained in [28, 29].

Isotropic elasticity

We end the paper by sketching the extension of our approach to the analysis of **HDG** methods for isotropic linear elasticity equations; see [30]. The governing equations of linear elasticity for isotropic materials can be written as follows

$$\mathrm{L}-\nabla\boldsymbol{u}=0,\quad \text{on }\Omega, \tag{5.27a}$$

$$-\nabla\cdot(\mu\mathrm{L})+\nabla p=0,\quad \text{on }\Omega, \tag{5.27b}$$

$$\epsilon p+\nabla\cdot\boldsymbol{u}=0,\quad \text{on }\Omega, \tag{5.27c}$$

$$\boldsymbol{u}=\boldsymbol{g},\quad \text{on }\partial\Omega, \tag{5.27d}$$

where $\epsilon=(1-2\nu)(1+\nu)/E$. Here E is the Young's modulus and $\nu\in(0,1/2]$ is the Poisson's ratio. The advantage of this formulation is that it holds for both compressible ($\nu\in(0,1/2)$) and incompressible ($\nu=1/2$) materials. Clearly, for incompressible materials, this system of equations is nothing but the Stokes system (5.1).

For the case of compressible materials($\epsilon>0$), the weak formulation of the system is similar to the Stokes equation (5.3). In fact, the only difference is that we replace the equation (5.3c) by

$$(\epsilon p_h\,,\,q)_{\mathcal{T}_h}-(\boldsymbol{u}_h\,,\,\nabla q)_{\mathcal{T}_h}+\langle\widehat{\boldsymbol{u}}_h\cdot\boldsymbol{n}\,,\,q\rangle_{\partial\mathcal{T}_h}=0.$$

Next, we argue that this difference does not affect the application of our approach to these equations.

Indeed, under the same *Assumption A*, we can obtain the following identity:

$$\epsilon\|e_p\|^2_{\mathcal{T}_h}+\nu\|\mathrm{E_L}\|^2_{\mathcal{T}_h}+\langle\boldsymbol{\alpha}(\boldsymbol{e_u}-\boldsymbol{e_{\widehat{u}}})\,,\,\boldsymbol{e_u}-\boldsymbol{e_{\widehat{u}}}\rangle_{\partial\mathcal{T}_h}=\nu(\Pi_G\mathrm{L}-\mathrm{L}\,,\,\mathrm{E_L})_{\mathcal{T}_h}.$$

Since $\epsilon > 0$ and $\boldsymbol{\alpha}$ is semi-positive definite, we still have the estimate

$$\|\mathrm{L} - \mathrm{L}_h\|_{\mathcal{T}_h} \leq 2\|\Pi_G \mathrm{L} - \mathrm{L}\|_{\mathcal{T}_h}.$$

To get the estimate of the pressure, we can use the same argument (Proposition 5.5.3) since we only need the error equation (5.25b).

Finally, if we modify the dual problem (5.6) by inserting $-\epsilon\eta$ on the left hand side of (5.6c), then, by using a similar duality argument, we obtain that

$$(\boldsymbol{e_u}\,,\,\boldsymbol{\theta})_{\mathcal{T}_h} = \nu(\Pi_G \mathrm{L} - \mathrm{L}\,,\,\mathrm{Z})_{\mathcal{T}_h} + \nu(\mathrm{L} - \mathrm{L}_h\,,\,\mathrm{Z}\Pi_G^*\mathrm{Z})_{\mathcal{T}_h} + \epsilon(e_p\,,\,\eta - \mathit{\Pi}_P^*\eta)_{\mathcal{T}_h}.$$

Finally, assuming the same regularity property, and under the same *Assumption B*, we get

$$\|\boldsymbol{\Pi}_V \boldsymbol{u} - \boldsymbol{u}_h\|_{\mathcal{T}_h} \leq Ch\|\mathrm{L} - \Pi_G \mathrm{L}\|_{\mathcal{T}_h}.$$

Clearly, the very same local postprocessing can be applied if *Assumption C* is satisfied.

In conclusion, our projection-based analysis can be naturally extended to isotropic elasticity problems. The extension of this approach to methods for linear elasticity, which are based on weak symmetry formulations, is the subject of Chapter 6.

Chapter 6

Superconvergent HDG methods for Linear Elasticity with weakly symmetric stresses

In this Chapter, we introduce a systematic way of devising new superconvergent mixed and hybridizable discontinuous Galerkin (HDG) methods for the system of linear elasticity

$$\mathcal{A}\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\epsilon}}(\boldsymbol{u}) = 0 \quad \text{in } \Omega \subset \mathbb{R}^n, \tag{6.1a}$$

$$\nabla \cdot \underline{\boldsymbol{\sigma}} = \boldsymbol{f} \quad \text{in } \Omega, \tag{6.1b}$$

$$\boldsymbol{u} = \boldsymbol{0} \quad \text{on } \partial\Omega, \tag{6.1c}$$

Here $\underline{\boldsymbol{\epsilon}}(\boldsymbol{u}) := \frac{1}{2}(\nabla \boldsymbol{u} + \nabla^t \boldsymbol{u})$ is the strain, $\mathcal{A}$ is a bounded symmetric positive definite tensor, $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega)$ and Ω is a polyhedral domain.

6.1 Main results

In this section we show how we can systematically construct superconvergent HDG (and mixed) methods form already existing superconvergent HDG (and mixed) methods for diffusion problems. We begin by introducing the general formulation for these methods.

The weak formulation of these methods is based on the following form of the system of liner elasticity:

$$\mathcal{A}\underline{\boldsymbol{\sigma}} - \nabla \boldsymbol{u} + \underline{\boldsymbol{\rho}} = 0 \quad \text{in } \Omega, \tag{6.2a}$$

$$\nabla \cdot \underline{\boldsymbol{\sigma}} = \boldsymbol{f} \quad \text{in } \Omega, \tag{6.2b}$$

$$\boldsymbol{u} = 0 \quad \text{on } \partial\Omega, \tag{6.2c}$$

where $\underline{\boldsymbol{\rho}} = \frac{1}{2}(\nabla \boldsymbol{u} - \nabla^t \boldsymbol{u})$ is called the *rotation*.

6.1.1 The general formulation

Let us begin by introducing the general form of the methods under consideration. We adapt to our setting the notation used in [3]. Let $\mathcal{T}_h$ denote a conforming triangulation of Ω made of shape-regular polyhedral elements K. Set $\partial\mathcal{T}_h := \{\partial K : K \in \mathcal{T}_h\}$, and let $\mathcal{E}_h$ denote the set of all faces F of all elements $K \in \mathcal{T}_h$. We denote by $\mathcal{F}(K)$ the set of all faces F of the element K. We also use the standard notation to denote scalar, vector and tensor spaces. Thus, if $D(K)$ denotes a space of scalar-valued functions defined on K, the corresponding space of vector-valued functions is $\boldsymbol{D}(K) := [D(K)]^d$ and the corresponding space of matrix-valued functions is $\underline{\boldsymbol{D}}(K) := [D(K)]^{d\times d}$.

The methods we consider seek an approximation $(\underline{\boldsymbol{\sigma}}_h, \boldsymbol{u}_h, \underline{\boldsymbol{\rho}}_h, \widehat{\boldsymbol{u}}_h)$ to the exact solution $(\underline{\boldsymbol{\sigma}}, \boldsymbol{u}, \underline{\boldsymbol{\rho}}, \boldsymbol{u}|_{\mathcal{E}_h})$ in the finite dimensional space $\underline{\boldsymbol{V}}_h \times \boldsymbol{W}_h \times \underline{\boldsymbol{A}}_h \times \boldsymbol{M}_h \subset \underline{\boldsymbol{L}}^2(\Omega) \times \boldsymbol{L}^2(\Omega) \times \underline{\boldsymbol{AS}}(\Omega) \times \boldsymbol{L}^2(\mathcal{E}_h)$ given by

$$\underline{\boldsymbol{V}}_h = \{\underline{\boldsymbol{v}} \in \underline{\boldsymbol{L}}^2(\mathcal{T}_h) : \quad \underline{\boldsymbol{v}}|_K \in \underline{\boldsymbol{V}}(K) \quad \forall\, K \in \mathcal{T}_h\}, \tag{6.3a}$$

$$\boldsymbol{W}_h = \{\boldsymbol{\omega} \in \boldsymbol{L}^2(\mathcal{T}_h) : \quad \boldsymbol{\omega}|_K \in \boldsymbol{W}(K) \quad \forall\, K \in \mathcal{T}_h\}, \tag{6.3b}$$

$$\underline{\boldsymbol{A}}_h = \{\underline{\boldsymbol{\eta}} \in \underline{\boldsymbol{L}}^2(\mathcal{T}_h) : \quad \underline{\boldsymbol{\eta}}|_K \in \underline{\boldsymbol{A}}(K) \quad \forall\, K \in \mathcal{T}_h\}, \tag{6.3c}$$

$$\boldsymbol{M}_h = \{\boldsymbol{\mu} \in \boldsymbol{L}^2(\mathcal{E}_h) : \quad \boldsymbol{\mu}|_F \in \boldsymbol{M}(F) \quad \forall\, F \in \mathcal{E}_h\}. \tag{6.3d}$$

Here $\underline{\boldsymbol{AS}}(\Omega) := \{\underline{\boldsymbol{\eta}} \in \underline{\boldsymbol{L}}^2(\Omega) : \ \underline{\boldsymbol{\eta}} + \underline{\boldsymbol{\eta}}^t = \underline{\boldsymbol{0}}\}$.

The approximation $(\underline{\boldsymbol{\sigma}}_h, \boldsymbol{u}_h, \underline{\boldsymbol{\rho}}_h, \widehat{\boldsymbol{u}}_h)$ can now be defined as the solution of the following equations:

$$(\mathcal{A}\underline{\boldsymbol{\sigma}}_h\,,\,\underline{\boldsymbol{v}})_{\mathcal{T}_h} + (\boldsymbol{u}_h\,,\,\nabla\cdot\underline{\boldsymbol{v}})_{\mathcal{T}_h} + (\underline{\boldsymbol{\rho}}_h\,,\,\underline{\boldsymbol{v}})_{\mathcal{T}_h} - \langle\widehat{\boldsymbol{u}}_h\,,\,\underline{\boldsymbol{v}}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} = 0, \tag{6.4a}$$

$$(\underline{\boldsymbol{\sigma}}_h\,,\,\nabla\boldsymbol{\omega})_{\mathcal{T}_h} - \langle\widehat{\underline{\boldsymbol{\sigma}}}_h\boldsymbol{n}\,,\,\boldsymbol{\omega}\rangle_{\partial\mathcal{T}_h} = -(\boldsymbol{f}\,,\,\boldsymbol{\omega})_{\mathcal{T}_h}, \tag{6.4b}$$

$$(\underline{\boldsymbol{\sigma}}_h\,,\,\underline{\boldsymbol{\eta}})_{\mathcal{T}_h} = 0, \tag{6.4c}$$

$$\langle\widehat{\underline{\boldsymbol{\sigma}}}_h\boldsymbol{n}\,,\,\boldsymbol{\mu}\rangle_{\partial\mathcal{T}_h\setminus\partial\Omega} = 0, \tag{6.4d}$$

$$\langle\widehat{\boldsymbol{u}}_h\,,\,\boldsymbol{\mu}\rangle_{\partial\Omega} = 0, \tag{6.4e}$$

for all $(\underline{\boldsymbol{v}}, \boldsymbol{\omega}, \underline{\boldsymbol{\eta}}, \boldsymbol{\mu}) \in \underline{\boldsymbol{V}}_h \times \boldsymbol{W}_h \times \underline{\boldsymbol{A}}_h \times \boldsymbol{M}_h$, where

$$\widehat{\underline{\boldsymbol{\sigma}}}_h\boldsymbol{n} = \underline{\boldsymbol{\sigma}}_h\boldsymbol{n} - \boldsymbol{\alpha}(\boldsymbol{u}_h - \widehat{\boldsymbol{u}}_h) \quad \text{on } \partial\mathcal{T}_h. \tag{6.4f}$$

Here, we write

$$(\underline{\boldsymbol{\eta}}\,,\,\underline{\boldsymbol{\zeta}})_{\mathcal{T}_h} := \sum_{i,j=1}^{n} (\underline{\boldsymbol{\eta}}_{i,j}\,,\,\underline{\boldsymbol{\zeta}}_{i,j})_{\mathcal{T}_h},$$
$$(\boldsymbol{\eta}\,,\,\boldsymbol{\zeta})_{\mathcal{T}_h} := \sum_{i=1}^{n} (\eta_i\,,\,\zeta_i)_{\mathcal{T}_h},$$
$$(\eta\,,\,\zeta)_{\mathcal{T}_h} := \sum_{K\in\mathcal{T}_h} (\eta, \zeta)_K,$$

where $(\eta, \zeta)_D$ denotes the integral of $\eta\zeta$ over $D \subset \mathbb{R}^n$. Similarly, we write $\langle\boldsymbol{\eta}\,,\,\boldsymbol{\zeta}\rangle_{\partial\mathcal{T}_h} := \sum_{i=1}^{n}\langle\eta_i\,,\,\zeta_i\rangle_{\partial\mathcal{T}_h}$ and $\langle\eta\,,\,\zeta\rangle_{\partial\mathcal{T}_h} := \sum_{K\in\mathcal{T}_h}\langle\eta\,,\,\zeta\rangle_{\partial K}$, where $\langle\eta\,,\,\zeta\rangle_D$ denotes the integral of $\eta\zeta$ over $D \subset \mathbb{R}^{n-1}$.

Note that, by taking particular choices of the local spaces $\underline{\boldsymbol{V}}(K), \boldsymbol{W}(K), \underline{\boldsymbol{A}}(K)$ and $\boldsymbol{M}(F)$, and of the *linear local stabilization operator* $\boldsymbol{\alpha}$, the different mixed and **HDG** methods are obtained. For the sake of simplicity, we assume $\boldsymbol{\alpha}$ to be self-adjoint.

6.1.2 A particular construction of the methods

Here we propose a very simple construction of the local spaces $\underline{\boldsymbol{V}}(K)$, $\boldsymbol{W}(K)$, $\underline{\boldsymbol{A}}(K)$ and $\boldsymbol{M}(F)$, and of the stabilization operator $\boldsymbol{\alpha}$.

Let us construct the local spaces. Given the spaces $\boldsymbol{V}^d(K)$, $W^d(K)$ and $M^d(F)$ defining a superconvergent HDG (or mixed) method for diffusion [3], and *any* space

$\underline{\boldsymbol{A}}(K) \subset \underline{\boldsymbol{AS}}(K)$, we take the local spaces to be of the following form:

$$\underline{\boldsymbol{V}}(K) := \underline{\boldsymbol{V}}^D(K) + \underline{\boldsymbol{B}}(K),$$
$$\boldsymbol{W}(K) := \boldsymbol{W}^D(K),$$
$$\boldsymbol{M}(F) := \boldsymbol{M}^D(F).$$

The local spaces $\underline{\boldsymbol{V}}^D(K)$, $\boldsymbol{W}^D(K)$, and $\boldsymbol{M}^D(F)$ depend solely on the spaces for diffusion and are defined as follows. The spaces $\boldsymbol{V}_i^D(K)$, $W_i^D(K)$ and $M_i^D(F)$ are nothing but $\boldsymbol{V}^d(K)$, $W^d(K)$ and $M^d(F)$ for $i = 1, ..., n$, where $\boldsymbol{V}_i^D(K)$ denotes the $i-$th row of $\underline{\boldsymbol{V}}^D(K)$ and $W_i^D(K)$ and $M_i^D(F)$ denote the $i-$th component of $\boldsymbol{W}^D(K)$ and $\boldsymbol{M}^D(F)$, respectively.

Note that, although the space of rotations $\underline{\boldsymbol{A}}(K)$ is only required to be a subset of $\underline{\boldsymbol{AS}}(K)$, it certainly *cannot* be taken to be too small, if we are to guarantee the optimal convergence of all the variables and the superconvergence of the projection of the error in the displacement. Roughly speaking, its approximation properties have to be similar to those of the space $\underline{\boldsymbol{V}}^D(K)$. A precise statement of this property is given in the next section.

The space $\underline{\boldsymbol{B}}(K)$ depends only on the space of rotations $\underline{\boldsymbol{A}}(K)$ and is given, in three-space dimensions, by

$$\underline{\boldsymbol{B}}(K) := \nabla \times ((\nabla \times \underline{\boldsymbol{A}}(K))\, \underline{\boldsymbol{b}}_K),$$

where $\underline{\boldsymbol{b}}_K$ is a suitably defined matrix-valued bubble function, and the i-th row of $\nabla \times \underline{\boldsymbol{\sigma}}$ is nothing but $\nabla\times$ applied to the i-th row of $\underline{\boldsymbol{\sigma}}$.

The actual definition of the bubble matrix functions depend only on the shape of the elements. For convex polyhedral elements, the two main examples are defined follows. For each face F of the polyhedral element K, let η_F be a linear function such that $\eta_F = 0$ on the face F and $\eta_F \geq 0$ on K. Then, we have

$$\underline{\boldsymbol{b}}_K := \left[\, \Pi_{F\in\mathcal{F}(K)}\, \eta_F \,\right] \underline{\boldsymbol{I}}, \tag{6.5a}$$

$$\underline{\boldsymbol{b}}_K := \sum_{F\in\mathcal{F}(K)} \left[\, \Pi_{F'\in\mathcal{F}(K)\setminus\{F\}}\, \eta_{F'} \,\right] \nabla\eta_F \otimes \nabla\eta_F, \tag{6.5b}$$

if K is a parallelepiped, we can reduce the above bubble function as

$$\underline{\boldsymbol{b}}_K := \sum_{F\in\mathcal{F}(K)} \left[\, \Pi_{F'\in\mathcal{F}(K)\setminus\{F,\bar{F}\}}\, \eta_{F'} \,\right] \nabla\eta_F \otimes \nabla\eta_F. \tag{6.5c}$$

Here $\underline{\boldsymbol{I}}$ is the identity matrix and $\bar{F}$ is the face parrallel to F. Note that, where the polyhedral element K has m faces, the entries of the first bubble matrix are polynomials of degree m. However, the entries of the second bubble matrix are polynomials of lower degree, namely, $(m-1)$ in the general case and $(m-2)$ if the element K is a parallelepiped. Because of this, the space $\underline{\boldsymbol{V}}(K)$ is usually smaller with the second choice of bubble matrices.

In the two-dimensional case, the space is defined by

$$\underline{\boldsymbol{B}}(K) := \nabla \times ((\nabla \times \underline{\boldsymbol{A}}(K))\, \underline{\boldsymbol{b}}_K).$$

Here the operator $\nabla\times$ is defined according to its argument as follows:

$$\nabla \times \begin{pmatrix} \tau_{11} & \tau_{12} \\ \tau_{21} & \tau_{22} \end{pmatrix} := \begin{pmatrix} -\partial_y \tau_{11} + \partial_x \tau_{12} \\ -\partial_y \tau_{21} + \partial_x \tau_{22} \end{pmatrix}, \quad \nabla \times \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} := \begin{pmatrix} -\partial_y v_1 & \partial_x v_1 \\ -\partial_y v_2 & \partial_x v_2 \end{pmatrix}.$$

Also, in the two-dimensional case, the bubble function $\underline{\boldsymbol{b}}_K$ is a *scalar*-valued function. With the natural restriction to this case of the notation used for the three-dimensional case, the main example is

$$\underline{\boldsymbol{b}}_K := \left[\, \Pi_{F \in \mathcal{F}(K)}\, \eta_F \,\right].$$

It remains to specify the stabilization operator $\boldsymbol{\alpha}$. Its construction is also very simple as we can take $\boldsymbol{\alpha} := \alpha^d\, \underline{\boldsymbol{I}}$, where α^d is the stabilization operator of the superconvergent method for diffusion [3].

6.1.3 A general construction of the methods

Next, we provide a generalization of the construction just proposed which retains the form of the local spaces. We proceed in three steps. First we utilize the local spaces for diffusion [3] to define $\underline{\boldsymbol{V}}^D(K)$, $\boldsymbol{W}^D(K)$ and $\boldsymbol{M}^D(F)$. Then we show a simple way to define the local space $\underline{\boldsymbol{A}}(K)$. Finally we introduce the space $\underline{\boldsymbol{B}}(K)$ and complete the construction of the space $\underline{\boldsymbol{V}}(K)$.

In what follows, $P^s(K)$ denotes the space of polynomials with degree no more than s on K and $Q^s(K)$ denotes the space of polynomials with degree no more than s for each variable on K.

Step 1: Construction of $\underline{\boldsymbol{V}}^D(K)$, $\boldsymbol{W}^D(K)$ and $\boldsymbol{M}^D(F)$. We require the local spaces $\underline{\boldsymbol{V}}^D(K)$, $\boldsymbol{W}^D(K)$ and $\boldsymbol{M}^D(F)$, and the stabilization operator $\boldsymbol{\alpha}$ to satisfy the following conditions:

(D.1) On each element K, there exist a projection $\Pi_h^D(\underline{\boldsymbol{\sigma}}, \boldsymbol{u}) = (\underline{\boldsymbol{\Pi}_{\boldsymbol{V}}^{\boldsymbol{D}}} \underline{\boldsymbol{\sigma}}, \boldsymbol{\Pi}_{\boldsymbol{W}}^{\boldsymbol{D}} \boldsymbol{u}) \in \underline{\boldsymbol{V}}^D(K) \times \boldsymbol{W}^D(K)$ satisfying the following properties:

$$
\begin{aligned}
(\underline{\boldsymbol{\Pi}_{\boldsymbol{V}}^{\boldsymbol{D}}} \underline{\boldsymbol{\sigma}}, \underline{\boldsymbol{v}})_K &= (\underline{\boldsymbol{\sigma}}, \underline{\boldsymbol{v}})_K \quad \text{for all } \underline{\boldsymbol{v}} \in \nabla \boldsymbol{W}^D(K), \\
(\boldsymbol{\Pi}_{\boldsymbol{W}}^{\boldsymbol{D}} \boldsymbol{u}, \boldsymbol{\omega})_K &= (\boldsymbol{u}, \boldsymbol{\omega})_K \quad \text{for all } \boldsymbol{\omega} \in \nabla \cdot \underline{\boldsymbol{V}}^D(K),
\end{aligned}
$$

For all faces F of the element K,

$$
\langle \underline{\boldsymbol{\Pi}_{\boldsymbol{V}}^{\boldsymbol{D}}} \underline{\boldsymbol{\sigma}} \boldsymbol{n} - \boldsymbol{\alpha}(\boldsymbol{\Pi}_{\boldsymbol{W}}^{\boldsymbol{D}} \boldsymbol{u}) \,,\, \boldsymbol{\mu} \rangle_F = \langle \underline{\boldsymbol{\sigma}} \boldsymbol{n} - \boldsymbol{\alpha}(\boldsymbol{P}_{\boldsymbol{M}}^{\boldsymbol{D}} \boldsymbol{u}) \,,\, \boldsymbol{\mu} \rangle_F \quad \text{for all } \boldsymbol{\mu} \in \boldsymbol{M}^D(F),
$$

where $\boldsymbol{P}_{\boldsymbol{M}}^{\boldsymbol{D}}$ denotes the L^2-projection onto $\boldsymbol{M}(F)$.

(D.2) On each face F of any element K

$$
\underline{\boldsymbol{V}}^D(K)\boldsymbol{n}|_F \subset \boldsymbol{M}^D(F), \quad \boldsymbol{W}^D(K)|_F \subset \boldsymbol{M}^D(F).
$$

(D.3) For each element K and any of its faces F,

$$
\langle \boldsymbol{\alpha}(\boldsymbol{\mu}) \,,\, \boldsymbol{\mu} \rangle_F \geq 0 \quad \text{for all } \boldsymbol{\mu} \in \boldsymbol{M}^D(F).
$$

(D.4) For each element K and any $(\underline{\boldsymbol{\sigma}}, \boldsymbol{u}) \in \underline{\boldsymbol{H}}^1(K) \times \boldsymbol{H}^1(K)$,

$$
\|\underline{\boldsymbol{\Pi}_{\boldsymbol{V}}^{\boldsymbol{D}}} \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}\|_K \leq C_{app} h_K (|\boldsymbol{u}|_{1,K} + |\underline{\boldsymbol{\sigma}}|_{1,K}).
$$

(D.5) For each element K, we have that

$$
\begin{aligned}
\underline{\boldsymbol{P}}^0(K) &\subset \nabla \boldsymbol{W}^D(K), \\
\boldsymbol{P}^0(K) &\subset \nabla \cdot \underline{\boldsymbol{V}}^D(K).
\end{aligned}
$$

Here $\boldsymbol{P}^0(K), \underline{\boldsymbol{P}}^0(K)$ space of constant vectors and the space of constant tensors, respectively.

Note that the above conditions constitute, roughly speaking, a matrix version of the corresponding *Assumptions A, B, C* on the local spaces and the stabilization operator for superconvergence (mixed) or HDG methods for diffusion [3]. If we use the particular construction described in the previous section, we can easily verify all these conditions. In particular, the i-th row of $\underline{\boldsymbol{\Pi}_{\boldsymbol{V}}^{\boldsymbol{D}}\boldsymbol{\sigma}}$ and the i-th component of $\boldsymbol{\Pi}_{\boldsymbol{W}}^{\boldsymbol{D}}\boldsymbol{u}$ are nothing but the two components of the projection for the diffusion case $(\Pi_{\boldsymbol{V}^d}, \Pi_{W^d})$ applied to the i-th row of $\underline{\boldsymbol{\sigma}}$ and the i-th component of $\boldsymbol{u}$. Also, the i-th component of the function $\boldsymbol{P}_{\boldsymbol{M}}^{\boldsymbol{D}}\boldsymbol{\mu}$ is nothing but the projection P_{M^d} of the i-th component of $\boldsymbol{\mu}$.

Step 2: Construction of $\underline{\boldsymbol{A}}(K)$. We take $\underline{\boldsymbol{A}}(K) := [A_{ij}(K)]^{d\times d} \subset \underline{\boldsymbol{AS}}(K)$, where

$$A_{ij}(K) = \begin{cases} A(K) & if \;\; i \neq j, \\ 0 & if \;\; i = j. \end{cases}$$

We require that the space $A(K)$ satisfies the following condition:

(A.1) $P^0(K) \subset A(K)$.

We can see that we only have a very little constraint on the choice of $A(K)$. However, to ensure the optimality of the methods, we need the space $A(K)$ satisfy some extra condition. We will discuss this issue at the end of this section.

Step 3: Construction of the space using bubbles. In this step we construct the local space $\underline{\boldsymbol{B}}(K)$. We follow [33] and the references therein. We begin by introducing the bubble matrices.

Definition A matrix-valued function $\underline{\boldsymbol{b}}$ defined on Ω is said to be an admissible bubble matrix if for each $K \in \mathcal{T}_h$ the matrix $\underline{\boldsymbol{b}}_K := \underline{\boldsymbol{b}}|_K$ is a matrix with polynomial entries that satisfies

(1) The tangential components of each row of $\underline{\boldsymbol{b}}_K$ vanish on all the faces of K.

(2) $C_1(\underline{\boldsymbol{v}}, \underline{\boldsymbol{v}})_K \leq (\underline{\boldsymbol{v}}\underline{\boldsymbol{b}}_K, \underline{\boldsymbol{v}})_K \quad$ for all $\underline{\boldsymbol{v}} \in \underline{\boldsymbol{L}}^2(K)$.

(3) $\|\underline{\boldsymbol{b}}_K\|_{L^\infty(K)} \leq C_2$,

where the positive constants C_1 and C_2 only depend on the shape regularity of the triangulation $\mathcal{T}_h$.

It is not difficult to verify that the two examples of bubble matrix functions given by (6.5) satisfy the above conditions. In particular, for the second example, the first property follows from the fact that $\nabla\eta_F$ is normal to the face F, and the second from the fact that $\eta_F \geq 0$ on K thanks to the convexity of K.

Note also that, if we use any admissible bubble function to define the local spaces $\underline{\boldsymbol{B}}(K)$, its not difficult to realize that any function $\underline{\boldsymbol{v}}$ lying in the space $\underline{\boldsymbol{B}}_h := \{\underline{\boldsymbol{\eta}} \in \underline{\boldsymbol{L}}^2(\Omega)| : \underline{\boldsymbol{\eta}}|_K \in \underline{\boldsymbol{B}}(K)\}$, is such that

(B.1) $\nabla \cdot \underline{\boldsymbol{v}}|_\Omega = 0,$

(B.2) $\underline{\boldsymbol{v}}\boldsymbol{n}|_{\mathcal{E}_h} = 0.$

6.1.4 A priori error estimates

To state our main result, we need to introduce some notation. We use $\|\cdot\|_D, |\cdot|_D$ to denote the usual norm and semi-norm on the Sobolev space D.

We also need to define the space $\underline{\boldsymbol{G}}_h := \{\underline{\boldsymbol{v}} \in \underline{\boldsymbol{L}}^2(\Omega) : \underline{\boldsymbol{G}}(K)\ \forall\, K \in \mathcal{T}_h\}$, where each of the local spaces $\underline{\boldsymbol{G}}(K)$ has *finite* dimension. Moreover, we require the existence of a projection into $\underline{\boldsymbol{G}}_h$ with two main properties.

Definition For any function $\underline{\boldsymbol{\eta}} \in \underline{\boldsymbol{H}}^1(\Omega)$, there is an element $\underline{\boldsymbol{\Pi_G\eta}}$ of $\underline{\boldsymbol{G}}_h$ such that

$$\langle(\underline{\boldsymbol{\eta}} - \underline{\boldsymbol{\Pi_G\eta}})\boldsymbol{n}\,,\,\boldsymbol{\mu}\rangle_F = 0 \qquad \forall\boldsymbol{\mu} \in \boldsymbol{P}^1(F)\ \forall\, F \in \mathcal{F}(K)\ \forall K \in \mathcal{T}_h, \tag{6.6a}$$

$$\nabla\cdot\underline{\boldsymbol{\Pi_G\eta}} = \boldsymbol{P}_0\nabla\cdot\underline{\boldsymbol{\eta}} \qquad \text{on } \Omega, \tag{6.6b}$$

where $\boldsymbol{P}_0$ is the L^2-projection onto $\boldsymbol{P}^0(K)$. Moreover, for each $K \in \mathcal{T}_h$,

$$\|\underline{\boldsymbol{\eta}} - \underline{\boldsymbol{\Pi_G\eta}}\|_K \leq Ch_K\|\underline{\boldsymbol{\eta}}\|_{\underline{\boldsymbol{H}}^1(K)}. \tag{6.6c}$$

In particular, when the element K is a triangle, rectangle, tetrahedron, cube or prisim, we can simply take each of the rows of the space $\underline{\boldsymbol{G}}(K)$ to be the space of vector-valued functions $\mathbf{BDDF}_1(K)$. Each of the rows of the projection $\underline{\boldsymbol{\Pi_G}}$ would then be the corresponding $\mathbf{BDDF_1}$ projection. For the precise definition of the space

$\mathbf{BDDF}_1(K)$ and its projection for different element K (in the $2D$ case, it is usually called the $\mathbf{BDM}_1(K)$), we refer the readers to [8]. For general isoparametric elements K, this choice can be extended by using the above definition on the reference element and then using the Piola transformation to obtain $\underline{\boldsymbol{G}}(K)$.

Finally, we introduce the so-called dual problem:

$$\nabla \cdot \underline{\boldsymbol{\psi}} = \boldsymbol{\theta} \quad \text{in } \Omega, \tag{6.7a}$$

$$\mathcal{A}\underline{\boldsymbol{\psi}} - \nabla\boldsymbol{\phi} + \underline{\boldsymbol{\xi}} = 0 \quad \text{in } \Omega, \tag{6.7b}$$

$$\boldsymbol{\phi} = 0 \quad \text{on } \partial\Omega. \tag{6.7c}$$

Here $\underline{\boldsymbol{\xi}} = \frac{1}{2}(\nabla\boldsymbol{\phi} - \nabla^t\boldsymbol{\phi})$. We assume the solution $(\underline{\boldsymbol{\psi}}, \boldsymbol{\phi})$ has the following elliptic regularity property:

$$\|\underline{\boldsymbol{\psi}}\|_{H^s(\Omega)} + \|\boldsymbol{\phi}\|_{H^{1+s}(\Omega)} \leq C_{reg}\|\boldsymbol{\theta}\|_\Omega, \tag{6.8}$$

for some $s \geq 0$. In the case of planar elasticity with scalar coefficients on a convex domain it holds with $s = 1$, see [52].

We are now ready to state our main result.

Theorem 6.1.1. *If $\underline{\boldsymbol{G}}(K) \subset \underline{\boldsymbol{V}}(K)$, then we have*

$$\|\underline{\boldsymbol{\Pi}_{\boldsymbol{V}}^{\boldsymbol{D}}\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_h\|_\Omega \leq C(\|\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\Pi}_{\boldsymbol{V}}^{\boldsymbol{D}}\boldsymbol{\sigma}}\|_\Omega + \|\underline{\boldsymbol{\rho}} - \underline{\boldsymbol{\Pi}_{\boldsymbol{A}}\boldsymbol{\rho}}\|_\Omega), \tag{6.9a}$$

$$\|\underline{\boldsymbol{\Pi}_{\boldsymbol{A}}\boldsymbol{\rho}} - \underline{\boldsymbol{\rho}}_h\|_\Omega \leq C(\|\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\Pi}_{\boldsymbol{V}}^{\boldsymbol{D}}\boldsymbol{\sigma}}\|_\Omega + \|\underline{\boldsymbol{\rho}} - \underline{\boldsymbol{\Pi}_{\boldsymbol{A}}\boldsymbol{\rho}}\|_\Omega), \tag{6.9b}$$

Moreover, if the elliptic regularity property (6.8) *holds for $s = 1$, then we have*

$$\|\boldsymbol{\Pi}_{\boldsymbol{W}}^{\boldsymbol{D}}\boldsymbol{u} - \boldsymbol{u}_h\|_\Omega \leq Ch(\|\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\Pi}_{\boldsymbol{V}}^{\boldsymbol{D}}\boldsymbol{\sigma}}\|_\Omega + \|\underline{\boldsymbol{\rho}} - \underline{\boldsymbol{\Pi}_{\boldsymbol{A}}\boldsymbol{\rho}}\|_\Omega). \tag{6.10}$$

Here $\underline{\boldsymbol{\Pi}_{\boldsymbol{A}}}$ is the L^2-projection onto $\underline{\boldsymbol{A}}_h$.

This result shows that the numerical errors for $\underline{\boldsymbol{\sigma}}$ and $\underline{\boldsymbol{\rho}}$ converge as fast as the projection errors. It also shows that, whenever $\boldsymbol{\Pi}_{\boldsymbol{W}}^{\boldsymbol{D}}$ converges to the identity slower than the other two components of the projection Π_h, the projection of the numerical error for $\boldsymbol{u}$ is *superconvergent* with at least an extra order. Hence, on each K, we can locally define a new approximation $\boldsymbol{u}_h^*$ in the space

$$\boldsymbol{W}_h^* := \{\boldsymbol{w} \in \boldsymbol{L}^2(\Omega) : \boldsymbol{w}|_K \in \boldsymbol{W}^*(K) \ \forall\, K \in \mathcal{T}_h\}.$$

The new postprocessed solution $\boldsymbol{u}_h^*$ superconverges to the exact solution $\boldsymbol{u}$ with at least an extra order. There are many ways to define $\boldsymbol{u}_h^*$, see [17, 18, 3]. Here we only present the one we used in [3] and omit the proof since it is almost a copy of the proof as in [3]. On each element $K \in \mathcal{T}_h$, we define the function $\boldsymbol{u}_h^*$ is the element of $\boldsymbol{W}^*(K)$ such that

$$(\nabla \boldsymbol{u}_h^*, \nabla \boldsymbol{w})_K = (\mathcal{A}\underline{\boldsymbol{\sigma}}_h + \underline{\boldsymbol{\rho}}_h, \nabla \boldsymbol{w})_K \quad \forall\, \boldsymbol{w} \in \boldsymbol{W}^*(K), \tag{6.11a}$$

$$(\boldsymbol{u}_h^*, 1)_K = (\boldsymbol{u}_h, 1)_K. \tag{6.11b}$$

We have the following superconvergence result.

Theorem 6.1.2. *Under the same conditions as in Theorem 6.1.1, we have*

$$\|\boldsymbol{u} - \boldsymbol{u}_h^*\|_\Omega \le C\, h\, (\|\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\Pi}_{\boldsymbol{V}}^{\boldsymbol{D}}\boldsymbol{\sigma}}\|_\Omega + \|\underline{\boldsymbol{\rho}} - \underline{\boldsymbol{\Pi}_{\boldsymbol{A}}\boldsymbol{\rho}}\|_\Omega + \inf_{\boldsymbol{w} \in \boldsymbol{W}_h^*} \|\nabla(\boldsymbol{u} - \boldsymbol{w})\|_{\mathcal{T}_h}).$$

The above theorems suggest that if the local spaces are constructed by following the steps in Section **2.3**, the approximation errors are bounded by the projection errors. However, this is not enough to guaranty that the methods produce optimal and superconvergent numerical approximations.

In [3], we gave many examples for which the projection error $\|\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\Pi}_{\boldsymbol{V}}^{\boldsymbol{D}}\boldsymbol{\sigma}}\|_\Omega$ is optimally convergent. Thus, we only need to choose the local space of the rotations in such a way that the terms $\|\underline{\boldsymbol{\rho}} - \underline{\boldsymbol{\Pi}_{\boldsymbol{A}}\boldsymbol{\rho}}\|_\Omega$ and $\|\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\Pi}_{\boldsymbol{V}}^{\boldsymbol{D}}\boldsymbol{\sigma}}\|_\Omega$, have similar convergence properties.

In all the examples discussed in [3], $\boldsymbol{V}^D(K)$ was either between $\boldsymbol{P}^k(K), \boldsymbol{P}^{k+1}(K)$ or between $\boldsymbol{Q}^k(K), \boldsymbol{Q}^{k+1}(K)$ and $W(K)$ was either $P^k(K)$, $P^{k-1}(K)$ or $Q^k(K)$. So, in all these cases, we can simply set $A(K) = P^k(K)$.

6.1.5 Examples

To end this section, we present some examples to illustrate the systematic construction discussed in **Section 2.3**. In what follows, all the examples are considered in $3D$ case and all the spaces are defined on the reference element K. We will use the bubble function defined by (6.5c).

Example 1: K is simplex with $\mathbf{HDG_k}$ element. We start with **Step 1.** In the case of simplex, the $\mathbf{HDG_k}$ local spaces $\boldsymbol{V}^d(K), W^d(K)$ and $M^d(F)$ for diffusion is defined as

$$\boldsymbol{V}^d(K) = \boldsymbol{P}^k(K), \quad W(K) = P^k(K), \quad M^d(F) = P^k(F).$$

The stabilization operator $\boldsymbol{\alpha} = \underline{\boldsymbol{I}}$. Hence, we can define

$$\underline{\boldsymbol{V}}^D(K) = \underline{\boldsymbol{P}}^k(K), \quad \boldsymbol{W}(K) = \boldsymbol{W}^D(K) = \boldsymbol{P}^k(K), \quad \boldsymbol{M}(F) = \boldsymbol{M}^D(F) = \boldsymbol{P}^k(F).$$

Next step is to constrcuct $\underline{\boldsymbol{A}}(K)$. We only need to specify the local space $A(K)$ due to the definition of $\underline{\boldsymbol{A}}(K)$ in **Step 2**. In order to obtain optimal and superconvergent approximation, we require that $\|\underline{\boldsymbol{\rho}} - \underline{\boldsymbol{\Pi}_{\boldsymbol{A}}\boldsymbol{\rho}}\|_\Omega$ and $\|\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\Pi}^{\boldsymbol{D}}_{\boldsymbol{V}}\boldsymbol{\sigma}}\|_\Omega$, have similar convergence properties. This suggests us to define

$$A(K) = P^k(K).$$

Finally, if we choose the bubble function $\underline{\boldsymbol{b}}_K$ as in (6.5a), we complete the construction by defining

$$\begin{aligned} \underline{\boldsymbol{V}}(K) &= \underline{\boldsymbol{P}}^k(K) + \nabla \times ((\nabla \times \underline{\boldsymbol{A}}(K))\underline{\boldsymbol{b}}_K) \\ &= \underline{\boldsymbol{P}}^k(K) \oplus \nabla \times ((\nabla \times \widetilde{\underline{\boldsymbol{A}}}(K))\underline{\boldsymbol{b}}_K), \end{aligned}$$

where $\widetilde{\underline{\boldsymbol{A}}}(K) = \underline{\boldsymbol{A}}(K) \cap \widetilde{\underline{\boldsymbol{P}}}^k(K)$.

Example 2: K is a cube with $\mathbf{RT_k}$ element. In this case, the Raviart-Thomas element for diffusion is defined as

$$\begin{aligned} &\boldsymbol{V}^d(K) = Q^{k+1,k,k}(K) \times Q^{k,k+1,k}(K) \times Q^{k,k,k+1}(K), \\ &W^d(K) = Q^k(K), \quad M^d(F) = Q^k(F). \end{aligned}$$

The stabilization operator $\boldsymbol{\alpha} = \underline{\boldsymbol{0}}$. So we can define

$$\begin{aligned} \boldsymbol{V}_i^D(K) = \boldsymbol{V}^d(K) = Q^{k+1,k,k}(K) \times Q^{k,k+1,k}(K) \times Q^{k,k,k+1}(K), \\ \boldsymbol{W}(K) = \boldsymbol{W}^D(K) = \boldsymbol{Q}^k(K), \quad \boldsymbol{M}(F) = \boldsymbol{M}^D(F) = \boldsymbol{Q}^k(F). \end{aligned}$$

Next we need to determine $A(K)$. Noting that we have $\underline{\boldsymbol{Q}}^k(K) \subset \underline{\boldsymbol{V}}^D(K)$, in order to obtain optimal numerical solutions, the smallest space we need is

$$A(K) = P^k(K).$$

Finally, if we choose the bubble function $\underline{\boldsymbol{b}}_K$ as in (6.5b), we complete the construction by defining

$$\underline{\boldsymbol{V}}(K) = \underline{\boldsymbol{V}}^D(K) + \nabla \times ((\nabla \times \underline{\boldsymbol{A}}(K))\underline{\boldsymbol{b}}_K).$$

Example 3: K **is a cube with** $\mathbf{HDG_k}, \mathbf{TNT_k}$ **elements.** When K is a cube, we have two types of HDG elements for diffusion problem. One is based on $P^k(K)$ space and the other one is based on $Q^(K)$. Here we only present the latter one. The $\mathbf{TNT_k}$ elements for diffusion was first introduced in [3]. The main feature of this type of elements is that they have the TiNiest spaces containing Tensor product spaces of polynomials of degree k. Here we follow the notation used in [53]. The spaces are defined as

$$\begin{array}{llll}
\mathbf{HDG_k} & \boldsymbol{V}^d(K) = \boldsymbol{Q}^k(K) \oplus \widetilde{\boldsymbol{\delta}_V}(K), & W^d(K) = Q^k(K), & M^d(F) = Q^k(F), \\
\mathbf{TNT_k} & \boldsymbol{V}^d(K) = \boldsymbol{Q}^k(K) \oplus \boldsymbol{\delta_V}(K), & W^d(K) = Q^k(K), & M^d(F) = Q^k(F),
\end{array}$$

Here $\widetilde{\boldsymbol{\delta}_{\boldsymbol{V}}}(K), \boldsymbol{\delta_V}(K)$ are defined as

$$\begin{aligned}
\widetilde{\boldsymbol{\delta}_{\boldsymbol{V}}}(K) := span\{ & (\tilde{B}_{k+1}(x), 0, 0),\ (0, \tilde{B}_{k+1}(y), 0),\ (0, 0, \tilde{B}_{k+1}(z)), \\
& (0,\ \tilde{B}_{k+1}(y)\tilde{P}_k(z),\ \tilde{B}_{k+1}(z)\tilde{P}_k(y)), \\
& (\tilde{B}_{k+1}(x)\tilde{P}_k(z),\ 0,\ \tilde{B}_{k+1}(z)\tilde{P}_k(x)), \\
& (\tilde{B}_{k+1}(x)\tilde{P}_k(y),\ \tilde{B}_{k+1}(y)\tilde{P}_k(x),\ 0)\}, \\
\boldsymbol{\delta_V}(K) := \widetilde{\boldsymbol{\delta}_{\boldsymbol{V}}}(K) \oplus & \\
& span(\tilde{B}_{k+1}(x)\tilde{P}_k(y)\tilde{P}_k(z),\ \tilde{P}_k(x)\tilde{B}_{k+1}(y)\tilde{P}_k(z),\ \tilde{P}_k(x)\tilde{P}_k(y)\tilde{B}_{k+1}(z))\}.
\end{aligned}$$

here $\tilde{P}_l$ denotes the *shifted* Legendre polynomial on $[0, 1]$ of degree $l \geq 0$, and $\tilde{B}_{l+1} := (\tilde{P}_{l+1} - \tilde{P}_{l-1})/(4l + 2)$, $\tilde{B}_1 = 0$. For $\mathbf{HDG_k}$ element, the stabilization operator $\boldsymbol{\alpha} = \underline{\boldsymbol{I}}$. For $\mathbf{TNT_k}$ element, the stabilization operator $\boldsymbol{\alpha} = \underline{\boldsymbol{0}}$. In both cases, we can define

$$\begin{gathered}
\boldsymbol{V}_i^D(K) = \boldsymbol{Q}^k(K) \oplus \widetilde{\boldsymbol{\delta}_{\boldsymbol{V}}}(K) \ \ or \ \ \boldsymbol{V}_i^D(K) = \boldsymbol{Q}^k(K) \oplus \boldsymbol{\delta_V}(K) \\
\boldsymbol{W}(K) = \boldsymbol{W}^D(K) = \boldsymbol{Q}^k(K), \quad \boldsymbol{M}(F) = \boldsymbol{M}^D(F) = \boldsymbol{Q}^k(F).
\end{gathered}$$

Noting that we also have $\underline{\boldsymbol{Q}}^k(K) \subset \underline{\boldsymbol{V}}^D(K)$, so we can define

$$A(K) = P^k(K).$$

Finally we complete the construction by defining

$$\underline{\boldsymbol{V}}(K) = \underline{\boldsymbol{V}}^D(K) + \nabla \times ((\nabla \times \underline{\boldsymbol{A}}(K))\underline{\boldsymbol{b}}_K).$$

Remark. In the last two examples, we can also write

$$\underline{\boldsymbol{V}}(K) = \underline{\boldsymbol{V}}^D(K) \oplus \nabla \times ((\nabla \times \widetilde{\underline{\boldsymbol{A}}}(K))\underline{\boldsymbol{b}}_K).$$

Here $\widetilde{\underline{\boldsymbol{A}}}(K)$ is a 21-dimentional space defined as follows:

$$\widetilde{\underline{\boldsymbol{A}}}(K) =: \{\underline{\boldsymbol{\eta}} = \begin{pmatrix} 0 & z_3 & -z_2 \\ -z_3 & 0 & z_1 \\ z_2 & -z_1 & 0 \end{pmatrix}, (z_1, z_2, z_3) \in A_1(K) \times A_2(K) \times A_3(K)\},$$

where

$$\begin{aligned} A_1(K) &= span\{x^k, y^k, z^k, y^{k-1}x, y^{k-1}z, z^{k-1}x, z^{k-1}y\}, \\ A_2(K) &= span\{x^k, y^k, z^k, x^{k-1}y, x^{k-1}z, z^{k-1}x, z^{k-1}y\}, \\ A_3(K) &= span\{x^k, y^k, z^k, x^{k-1}y, x^{k-1}z, y^{k-1}x, y^{k-1}z\}. \end{aligned}$$

In Table 6.1, we present the order of convergence for all the examples mentioned above.

Table 6.1: Order of convergence for different methods $k \geq 1$

method	K	$\|\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_h\|_{\mathcal{T}_h}$	$\|\underline{\boldsymbol{\rho}} - \underline{\boldsymbol{\rho}}_h\|_{\mathcal{T}_h}$	$\|\boldsymbol{u} - \boldsymbol{u}_h^\star\|_{\mathcal{T}_h}$
$\mathbf{HDG}_k$	simplex	$k+1$	$k+1$	$k+2$
$\mathbf{RT}_k$	cube	$k+1$	$k+1$	$k+2$
$\mathbf{HDG}_k$	cube	$k+1$	$k+1$	$k+2$
$\mathbf{TNT}_k$	cube	$k+1$	$k+1$	$k+2$

6.2 Error Analysis

In this section we provide detailed proofs for our a priori error estimates. We denote

$$
\begin{aligned}
\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}} &= \underline{\boldsymbol{\Pi}_{\boldsymbol{V}}^{\boldsymbol{D}}}\underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\sigma}}_h,\\
\boldsymbol{e}_{\boldsymbol{u}} &= \boldsymbol{\Pi}_{\boldsymbol{W}}^{\boldsymbol{D}}\boldsymbol{u} - \boldsymbol{u}_h,\\
\underline{\boldsymbol{e}_{\boldsymbol{\rho}}} &= \underline{\boldsymbol{\Pi}_{\boldsymbol{A}}}\underline{\boldsymbol{\rho}} - \underline{\boldsymbol{\rho}}_h,\\
\boldsymbol{e}_{\widehat{\boldsymbol{u}}} &= \boldsymbol{P}_{\boldsymbol{M}}^{\boldsymbol{D}}\boldsymbol{u} - \widehat{\boldsymbol{u}}_h,\\
\underline{\boldsymbol{e}_{\widehat{\boldsymbol{\sigma}}}}\boldsymbol{n} &= \boldsymbol{P}_{\boldsymbol{M}}^{\boldsymbol{D}}(\underline{\boldsymbol{\sigma}}\boldsymbol{n}) - \underline{\widehat{\boldsymbol{\sigma}}}_h\boldsymbol{n}.
\end{aligned}
$$

We also denote

$$
\begin{aligned}
\underline{\boldsymbol{\delta}_{\boldsymbol{\sigma}}} &= \underline{\boldsymbol{\sigma}} - \underline{\boldsymbol{\Pi}_{\boldsymbol{V}}^{\boldsymbol{D}}}\underline{\boldsymbol{\sigma}},\\
\boldsymbol{\delta}_{\boldsymbol{u}} &= \boldsymbol{u} - \boldsymbol{\Pi}_{\boldsymbol{W}}^{\boldsymbol{D}}\boldsymbol{u},\\
\underline{\boldsymbol{\delta}_{\boldsymbol{\rho}}} &= \underline{\boldsymbol{\rho}} - \underline{\boldsymbol{\Pi}_{\boldsymbol{A}}}\underline{\boldsymbol{\rho}}.
\end{aligned}
$$

Here the projection $\underline{\boldsymbol{\Pi}_{\boldsymbol{V}}^{\boldsymbol{D}}}, \boldsymbol{\Pi}_{\boldsymbol{W}}^{\boldsymbol{D}}$ is defined in (D.1). $\underline{\boldsymbol{\Pi}_{\boldsymbol{A}}}, \boldsymbol{P}_{\boldsymbol{M}}^{\boldsymbol{D}}$ are the L^2-projections onto $\underline{\boldsymbol{A}}_h, \boldsymbol{M}_h$ respectively. We proceed in several steps.

Step 1: The error equation. We first present the error equation for the analysis.

Lemma 6.2.1. *Under the hypothesis of Theorem 6.1.1, we have*

$$
\begin{aligned}
(\mathcal{A}\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\,,\,\underline{\boldsymbol{v}})_{\mathcal{T}_h} + (\boldsymbol{e}_{\boldsymbol{u}}\,,\,\nabla\cdot\underline{\boldsymbol{v}})_{\mathcal{T}_h} + (\underline{\boldsymbol{e}_{\boldsymbol{\rho}}}\,,\,\underline{\boldsymbol{v}})_{\mathcal{T}_h} - \langle \boldsymbol{e}_{\widehat{\boldsymbol{u}}}\,,\,\underline{\boldsymbol{v}}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} &= \qquad\qquad (6.12a)\\
-(\mathcal{A}\underline{\boldsymbol{\delta}_{\boldsymbol{\sigma}}}\,,\,\underline{\boldsymbol{v}})_{\mathcal{T}_h} &- (\underline{\boldsymbol{\delta}_{\boldsymbol{\rho}}}\,,\,\underline{\boldsymbol{v}})_{\mathcal{T}_h},\\
(\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\,,\,\nabla\boldsymbol{\omega})_{\mathcal{T}_h} - \langle \underline{\boldsymbol{e}_{\widehat{\boldsymbol{\sigma}}}}\boldsymbol{n}\,,\,\boldsymbol{\omega}\rangle_{\partial\mathcal{T}_h} &= 0, \qquad\qquad (6.12b)\\
(\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\,,\,\underline{\boldsymbol{\eta}})_{\mathcal{T}_h} &= -(\underline{\boldsymbol{\delta}_{\boldsymbol{\sigma}}}\,,\,\underline{\boldsymbol{\eta}})_{\mathcal{T}_h}, \qquad (6.12c)\\
\langle \underline{\boldsymbol{e}_{\widehat{\boldsymbol{\sigma}}}}\boldsymbol{n}\,,\,\boldsymbol{\mu}\rangle_{\partial\mathcal{T}_h\backslash\partial\Omega} &= 0, \qquad\qquad (6.12d)\\
\langle \boldsymbol{e}_{\widehat{\boldsymbol{u}}}\,,\,\boldsymbol{\mu}\rangle_{\partial\Omega} &= 0, \qquad\qquad (6.12e)
\end{aligned}
$$

for all $(\underline{\boldsymbol{v}}, \boldsymbol{\omega}, \underline{\boldsymbol{\eta}}, \boldsymbol{\mu}) \in \underline{\boldsymbol{V}}_h \times \boldsymbol{W}_h \times \underline{\boldsymbol{A}}_h \times \boldsymbol{M}_h$.

Proof. We notice that the exact solution $(\underline{\boldsymbol{\sigma}}, \boldsymbol{u}, \underline{\boldsymbol{\rho}}, \boldsymbol{u}|_{\mathcal{E}_h})$ also satisfies the equation (6.4).

Hence, after simple algebraic manipulations, we get that

$$\begin{aligned}
(\mathcal{A}\underline{\boldsymbol{\Pi}^{\boldsymbol{D}}_{\boldsymbol{V}}\boldsymbol{\sigma}}\,,\,\underline{\boldsymbol{v}})_{\mathcal{T}_h} + (\boldsymbol{\Pi}^{\boldsymbol{D}}_{\boldsymbol{W}}\boldsymbol{u}\,,\,\nabla\cdot\underline{\boldsymbol{v}})_{\mathcal{T}_h} + (\underline{\boldsymbol{\Pi}_{\boldsymbol{A}}\boldsymbol{\rho}}\,,\,\underline{\boldsymbol{v}})_{\mathcal{T}_h} - \langle \boldsymbol{P}^{\boldsymbol{D}}_{\boldsymbol{M}}\boldsymbol{u}\,,\,\underline{\boldsymbol{v}}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} &= \\
-(\mathcal{A}\underline{\boldsymbol{\delta}_{\boldsymbol{\sigma}}}\,,\,\underline{\boldsymbol{v}})_{\mathcal{T}_h} - (\underline{\boldsymbol{\delta}_{\boldsymbol{\rho}}}\,,\,\underline{\boldsymbol{v}})_{\mathcal{T}_h} + \langle \boldsymbol{u} - \boldsymbol{P}^{\boldsymbol{D}}_{\boldsymbol{M}}\boldsymbol{u}\,,\,\underline{\boldsymbol{v}}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - (\boldsymbol{\delta}_{\boldsymbol{u}}\,,\,\nabla\cdot\underline{\boldsymbol{v}})_{\mathcal{T}_h}&, \\
(\underline{\boldsymbol{\Pi}^{\boldsymbol{D}}_{\boldsymbol{V}}\boldsymbol{\sigma}}\,,\,\nabla\boldsymbol{\omega})_{\mathcal{T}_h} - \langle \boldsymbol{P}^{\boldsymbol{D}}_{\boldsymbol{M}}(\underline{\boldsymbol{\sigma}}\boldsymbol{n})\,,\,\boldsymbol{\omega}\rangle_{\partial\mathcal{T}_h} = -(\boldsymbol{f}\,,\,\boldsymbol{\omega})_{\mathcal{T}_h} - (\underline{\boldsymbol{\delta}_{\boldsymbol{\sigma}}}\,,\,\nabla\boldsymbol{\omega})_{\mathcal{T}_h}& \\
+\langle(\underline{\boldsymbol{\sigma}}\boldsymbol{n} - \boldsymbol{P}^{\boldsymbol{D}}_{\boldsymbol{M}}(\underline{\boldsymbol{\sigma}}\boldsymbol{n})\,,\,\boldsymbol{\omega}\rangle_{\partial\mathcal{T}_h}&, \\
(\underline{\boldsymbol{\Pi}^{\boldsymbol{D}}_{\boldsymbol{V}}\boldsymbol{\sigma}}\,,\,\underline{\boldsymbol{\eta}})_{\mathcal{T}_h} = -(\underline{\boldsymbol{\delta}_{\boldsymbol{\sigma}}}\,,\,\underline{\boldsymbol{\eta}})_{\mathcal{T}_h}&,
\end{aligned}$$

for all $(\underline{\boldsymbol{v}}, \boldsymbol{\omega}, \underline{\boldsymbol{\eta}}) \in \underline{\boldsymbol{V}}_h \times \boldsymbol{W}_h \times \underline{\boldsymbol{A}}_h$. Moreover, by the definition of the projection $\boldsymbol{P}^{\boldsymbol{D}}_{\boldsymbol{M}}$, we have that

$$\begin{aligned}
\langle \boldsymbol{P}^{\boldsymbol{D}}_{\boldsymbol{M}}(\underline{\boldsymbol{\sigma}}\boldsymbol{n})\,,\,\boldsymbol{\mu}\rangle_{\partial\mathcal{T}_h\setminus\partial\Omega} &= 0, \\
\langle \boldsymbol{P}^{\boldsymbol{D}}_{\boldsymbol{M}}\boldsymbol{u}\,,\,\boldsymbol{\mu}\rangle_{\partial\Omega} &= 0,
\end{aligned}$$

for all $\boldsymbol{\mu} \in \boldsymbol{M}_h$.

Next, we show that $(\underline{\boldsymbol{\delta}_{\boldsymbol{\sigma}}}\,,\,\nabla\boldsymbol{\omega})_{\mathcal{T}_h} = 0$ and that $(\boldsymbol{\delta}_{\boldsymbol{u}}\,,\,\nabla\cdot\underline{\boldsymbol{v}})_{\mathcal{T}_h} = 0$ for any $(\boldsymbol{w}, \underline{\boldsymbol{v}}) \in \boldsymbol{W}(K) \times \underline{\boldsymbol{V}}(K)$. The first identity is due to the projection property (D.1). By the property of the bubble function space (B.1), we have

$$\nabla\cdot(\underline{\boldsymbol{V}}(K)) = \nabla\cdot(\underline{\boldsymbol{V}}^D(K) + \underline{\boldsymbol{B}}(K)) = \nabla\cdot(\underline{\boldsymbol{V}}^D(K)).$$

The second identity now follows by the above fact and the projection property (D.1).

Now, by the inclusion property (D.2) and by the property (B.2) of the space of bubbles $\underline{\boldsymbol{B}}_h$, we have that, on each face $F \in \partial K$,

$$\boldsymbol{W}(K)|_F \subset \boldsymbol{M}(F), \quad \underline{\boldsymbol{V}}(K)\boldsymbol{n}|_F = \underline{\boldsymbol{V}}^D(K)\boldsymbol{n}|_F + \underline{\boldsymbol{B}}(K)\boldsymbol{n}|_F = \underline{\boldsymbol{V}}^D(K)\boldsymbol{n}|_F \subset \boldsymbol{M}(F).$$

This implies that $\langle(\underline{\boldsymbol{\sigma}}\boldsymbol{n} - \boldsymbol{P}^{\boldsymbol{D}}_{\boldsymbol{M}}(\underline{\boldsymbol{\sigma}}\boldsymbol{n})\,,\,\boldsymbol{\omega}\rangle_{\partial\mathcal{T}_h} = 0$ and that $\langle \boldsymbol{u} - \boldsymbol{P}^{\boldsymbol{D}}_{\boldsymbol{M}}\boldsymbol{u}\,,\,\underline{\boldsymbol{v}}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} = 0$ for all $(\underline{\boldsymbol{v}}, \boldsymbol{\omega}) \in \underline{\boldsymbol{V}}_h \times \boldsymbol{W}_{\boldsymbol{h}}$.

Using the last four identities, we obtain

$$
\begin{aligned}
(\mathcal{A}\underline{\boldsymbol{\Pi}_{\boldsymbol{V}}^{\boldsymbol{D}}\boldsymbol{\sigma}}\,,\,\underline{\boldsymbol{v}})_{\mathcal{T}_h} + (\boldsymbol{\Pi}_{\boldsymbol{W}}^{\boldsymbol{D}}\boldsymbol{u}\,,\,\nabla\cdot\underline{\boldsymbol{v}})_{\mathcal{T}_h} + (\underline{\boldsymbol{\Pi}_{\boldsymbol{A}}\boldsymbol{\rho}}\,,\,\underline{\boldsymbol{v}})_{\mathcal{T}_h} - \langle \boldsymbol{P}_{\boldsymbol{M}}^{\boldsymbol{D}}\boldsymbol{u}\,,\,\underline{\boldsymbol{v}}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} &= \\
-(\mathcal{A}\underline{\boldsymbol{\delta}_{\boldsymbol{\sigma}}}\,,\,\underline{\boldsymbol{v}})_{\mathcal{T}_h} - (\underline{\boldsymbol{\delta}_{\boldsymbol{\rho}}}\,,\,\underline{\boldsymbol{v}})_{\mathcal{T}_h}&,\\
(\underline{\boldsymbol{\Pi}_{\boldsymbol{V}}^{\boldsymbol{D}}\boldsymbol{\sigma}}\,,\,\nabla\boldsymbol{\omega})_{\mathcal{T}_h} - \langle \boldsymbol{P}_{\boldsymbol{M}}^{\boldsymbol{D}}(\underline{\boldsymbol{\sigma}}\boldsymbol{n})\,,\,\boldsymbol{\omega}\rangle_{\partial\mathcal{T}_h} &= -(\boldsymbol{f}\,,\,\boldsymbol{\omega})_{\mathcal{T}_h},\\
(\underline{\boldsymbol{\Pi}_{\boldsymbol{V}}^{\boldsymbol{D}}\boldsymbol{\sigma}}\,,\,\underline{\boldsymbol{\eta}})_{\mathcal{T}_h} &= -(\underline{\boldsymbol{\delta}_{\boldsymbol{\sigma}}}\,,\,\underline{\boldsymbol{\eta}})_{\mathcal{T}_h},\\
\langle \boldsymbol{P}_{\boldsymbol{M}}^{\boldsymbol{D}}(\underline{\boldsymbol{\sigma}}\boldsymbol{n})\,,\,\boldsymbol{\mu}\rangle_{\partial\mathcal{T}_h\setminus\partial\Omega} &= 0,\\
\langle \boldsymbol{P}_{\boldsymbol{M}}^{\boldsymbol{D}}\boldsymbol{u}\,,\,\boldsymbol{\mu}\rangle_{\partial\Omega} &= 0,
\end{aligned}
$$

for all $(\underline{\boldsymbol{v}},\boldsymbol{\omega},\underline{\boldsymbol{\eta}},\boldsymbol{\mu}) \in \underline{\boldsymbol{V}}_h \times \boldsymbol{W}_h \times \underline{\boldsymbol{A}}_h \times \boldsymbol{M}_h$. If we now subtract the equations (6.4), we obtain the result. This completes the proof. □

Step 2: Estimate of $\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}$. We are now ready to obtain our first estimate.

Proposition 6.2.2. *We have*

$$
(\mathcal{A}\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\,,\,\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}})_{\mathcal{T}_h} + \langle \boldsymbol{\alpha}(\boldsymbol{e}_{\widehat{\boldsymbol{u}}} - \boldsymbol{e}_{\boldsymbol{u}})\,,\,\boldsymbol{e}_{\widehat{\boldsymbol{u}}} - \boldsymbol{e}_{\boldsymbol{u}}\rangle_{\partial\mathcal{T}_h} = (\underline{\boldsymbol{\delta}_{\boldsymbol{\sigma}}}\,,\,\underline{\boldsymbol{e}_{\boldsymbol{\rho}}})_{\mathcal{T}_h} - (\mathcal{A}\underline{\boldsymbol{\delta}_{\boldsymbol{\sigma}}}\,,\,\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}})_{\mathcal{T}_h} - (\underline{\boldsymbol{\delta}_{\boldsymbol{\rho}}}\,,\,\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}})_{\mathcal{T}_h}
$$

Proof. Taking $\underline{\boldsymbol{v}} = \underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}$ in (6.12a), $\boldsymbol{\omega} = \boldsymbol{e}_{\boldsymbol{u}}$ in (6.12b) and adding, we have

$$
\begin{aligned}
(\mathcal{A}\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\,,\,\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}})_{\mathcal{T}_h} + \langle \underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\boldsymbol{n}\,,\,\boldsymbol{e}_{\boldsymbol{u}}\rangle_{\partial\mathcal{T}_h} - \langle \boldsymbol{e}_{\widehat{\boldsymbol{u}}}\,,\,\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle \underline{\boldsymbol{e}_{\widehat{\boldsymbol{\sigma}}}}\boldsymbol{n}\,,\,\boldsymbol{e}_{\boldsymbol{u}}\rangle_{\partial\mathcal{T}_h} + (\underline{\boldsymbol{e}_{\boldsymbol{\rho}}}\,,\,\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}})_{\mathcal{T}_h}&\\
= -(\mathcal{A}\underline{\boldsymbol{\delta}_{\boldsymbol{\sigma}}}\,,\,\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}})_{\mathcal{T}_h} - (\underline{\boldsymbol{\delta}_{\boldsymbol{\rho}}}\,,\,\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}})_{\mathcal{T}_h}&.
\end{aligned}
$$

Taking $\underline{\boldsymbol{\eta}} = \underline{\boldsymbol{e}_{\boldsymbol{\rho}}}$ in (6.12c), we have

$$
(\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\,,\,\underline{\boldsymbol{e}_{\boldsymbol{\rho}}})_{\mathcal{T}_h} = -(\underline{\boldsymbol{\delta}_{\boldsymbol{\sigma}}}\,,\,\underline{\boldsymbol{e}_{\boldsymbol{\rho}}})_{\mathcal{T}_h}.
$$

Taking $\boldsymbol{\mu} = \boldsymbol{e}_{\widehat{\boldsymbol{u}}}$ in (6.12d), $\boldsymbol{\mu} = \underline{\boldsymbol{e}_{\widehat{\boldsymbol{\sigma}}}}\boldsymbol{n}$ in (6.12e) and adding, we have

$$
\langle \underline{\boldsymbol{e}_{\widehat{\boldsymbol{\sigma}}}}\boldsymbol{n}\,,\,\boldsymbol{e}_{\widehat{\boldsymbol{u}}}\rangle_{\partial\mathcal{T}_h} = 0.
$$

Inserting the above two equations into the first equation we obtain, after some algebraic manipulations,

$$
(\mathcal{A}\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\,,\,\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}})_{\mathcal{T}_h} + \langle \boldsymbol{e}_{\boldsymbol{u}} - \boldsymbol{e}_{\widehat{\boldsymbol{u}}}\,,\,\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\boldsymbol{n} - \underline{\boldsymbol{e}_{\widehat{\boldsymbol{\sigma}}}}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} = (\underline{\boldsymbol{\delta}_{\boldsymbol{\sigma}}}\,,\,\underline{\boldsymbol{e}_{\boldsymbol{\rho}}})_{\mathcal{T}_h} - (\mathcal{A}\underline{\boldsymbol{\delta}_{\boldsymbol{\sigma}}}\,,\,\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}})_{\mathcal{T}_h} - (\underline{\boldsymbol{\delta}_{\boldsymbol{\rho}}}\,,\,\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}})_{\mathcal{T}_h}.
$$

To end the proof, we need to show that

$$
\langle \boldsymbol{e}_{\boldsymbol{u}} - \boldsymbol{e}_{\widehat{\boldsymbol{u}}}\,,\,\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\boldsymbol{n} - \underline{\boldsymbol{e}_{\widehat{\boldsymbol{\sigma}}}}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} = \langle \boldsymbol{e}_{\boldsymbol{u}} - \boldsymbol{e}_{\widehat{\boldsymbol{u}}}\,,\,\boldsymbol{\alpha}(\boldsymbol{e}_{\boldsymbol{u}} - \boldsymbol{e}_{\widehat{\boldsymbol{u}}})\rangle_{\partial\mathcal{T}_h}.
$$

On each K, by the definition of the numerical trace (6.4f) and the properties (D.1), (D.2) we have

$$\begin{aligned}\langle \boldsymbol{e_u} - \boldsymbol{e_{\widehat{u}}} \,,\, \underline{\boldsymbol{e_\sigma}}\boldsymbol{n} - \underline{\boldsymbol{e_{\widehat{\sigma}}}}\boldsymbol{n}\rangle_{\partial K} &= \langle \boldsymbol{e_u} - \boldsymbol{e_{\widehat{u}}} \,,\, \underline{\boldsymbol{\Pi^D_V}}\underline{\boldsymbol{\sigma}}\boldsymbol{n} - \underline{\boldsymbol{\sigma}}_h\boldsymbol{n} - \boldsymbol{P^D_M}(\underline{\boldsymbol{\sigma}}\boldsymbol{n}) + \underline{\widehat{\boldsymbol{\sigma}}}_h\boldsymbol{n}\rangle_{\partial K}\\ &= \langle \boldsymbol{e_u} - \boldsymbol{e_{\widehat{u}}} \,,\, \underline{\boldsymbol{\Pi^D_V}}\underline{\boldsymbol{\sigma}}\boldsymbol{n} - \boldsymbol{P^D_M}(\underline{\boldsymbol{\sigma}}\boldsymbol{n}) + \underline{\widehat{\boldsymbol{\sigma}}}_h\boldsymbol{n} - \underline{\boldsymbol{\sigma}}_h\boldsymbol{n}\rangle_{\partial K}\\ &= \langle \boldsymbol{e_u} - \boldsymbol{e_{\widehat{u}}} \,,\, \boldsymbol{\alpha}(\boldsymbol{\Pi^D_W}\boldsymbol{u} - \boldsymbol{P^D_M}\boldsymbol{u} + \widehat{\boldsymbol{u}}_h - \boldsymbol{u}_h)\rangle_{\partial K}\\ &= \langle \boldsymbol{e_u} - \boldsymbol{e_{\widehat{u}}} \,,\, \boldsymbol{\alpha}(\boldsymbol{e_u} - \boldsymbol{e_{\widehat{u}}})\rangle_{\partial K}.\end{aligned}$$

We complete the proof by taking the sum of the above equation over all $K \in \mathcal{T}_h$. □

Corollary 6.2.3. *We have*

$$\|\underline{\boldsymbol{e_\sigma}}\|^2_\Omega \leq C(\|\underline{\boldsymbol{\delta_\sigma}}\|_\Omega \|\underline{\boldsymbol{e_\rho}}\|_\Omega + \|\underline{\boldsymbol{\delta_\sigma}}\|^2_\Omega + \|\underline{\boldsymbol{\delta_\rho}}\|^2_\Omega),$$

where C is some constant which solely depends on the tensor $\mathcal{A}$.

Proof. By the semipositivity property (D.3) of the stabilization function, we know that $\langle \boldsymbol{e_u} - \boldsymbol{e_{\widehat{u}}} \,,\, \boldsymbol{\alpha}(\boldsymbol{e_u} - \boldsymbol{e_{\widehat{u}}})\rangle_{\partial\mathcal{T}_h} \geq 0$, this implies that

$$(\mathcal{A}\underline{\boldsymbol{e_\sigma}} \,,\, \underline{\boldsymbol{e_\sigma}})_{\mathcal{T}_h} \leq (\underline{\boldsymbol{\delta_\sigma}} \,,\, \underline{\boldsymbol{e_\rho}})_{\mathcal{T}_h} - (\mathcal{A}\underline{\boldsymbol{\delta_\sigma}} \,,\, \underline{\boldsymbol{e_\sigma}})_{\mathcal{T}_h} - (\underline{\boldsymbol{\delta_\rho}} \,,\, \underline{\boldsymbol{e_\sigma}})_{\mathcal{T}_h}.$$

The estimate can be obtained simply by using the Cauchy-Schwarz inequality and Young's inequality. □

Step 3: Estimate of $\boldsymbol{e_u}$. Next we use a duality argument to get an estimate for $\boldsymbol{e_u}$. First we present an important identity.

Proposition 6.2.4. *We have*

$$\begin{aligned}(\boldsymbol{e_u} \,,\, \boldsymbol{\theta})_{\mathcal{T}_h} = (\mathcal{A}\underline{\boldsymbol{e_\sigma}} \,,\, \underline{\boldsymbol{\delta_\psi}})_{\mathcal{T}_h} + (\underline{\boldsymbol{e_\rho}} \,,\, \underline{\boldsymbol{\delta_\psi}})_{\mathcal{T}_h} - (\mathcal{A}\underline{\boldsymbol{\delta_\sigma}} + \underline{\boldsymbol{\delta_\rho}} \,,\, \underline{\boldsymbol{\Pi^D_V}}\underline{\boldsymbol{\psi}})_{\mathcal{T}_h}\\ + (\underline{\boldsymbol{e_\sigma}} \,,\, \underline{\boldsymbol{\delta_\xi}})_{\mathcal{T}_h} - (\underline{\boldsymbol{\delta_\sigma}} \,,\, \underline{\boldsymbol{\Pi_A}}\underline{\boldsymbol{\xi}})_{\mathcal{T}_h},\end{aligned}$$

where $\underline{\boldsymbol{\delta_\xi}} := \underline{\boldsymbol{\xi}} - \underline{\boldsymbol{\Pi_A}}\underline{\boldsymbol{\xi}}$ and $\underline{\boldsymbol{\delta_\psi}} := \underline{\boldsymbol{\psi}} - \underline{\boldsymbol{\Pi^D_V}}\underline{\boldsymbol{\psi}}$.

Proof. By the dual equation (6.7), we can write

$$
\begin{aligned}
(\boldsymbol{e_u}\,,\,\boldsymbol{\theta})_{\mathcal{T}_h} &= (\boldsymbol{e_u}\,,\,\nabla\cdot\underline{\boldsymbol{\psi}})_{\mathcal{T}_h} + (\underline{\boldsymbol{e_\sigma}}\,,\,\mathcal{A}\underline{\boldsymbol{\psi}} - \nabla\boldsymbol{\phi} + \underline{\boldsymbol{\xi}})_{\mathcal{T}_h} \\
&= (\boldsymbol{e_u}\,,\,\nabla\cdot\underline{\boldsymbol{\psi}})_{\mathcal{T}_h} + (\mathcal{A}\underline{\boldsymbol{e_\sigma}}\,,\,\underline{\boldsymbol{\psi}})_{\mathcal{T}_h} \\
&\quad - (\underline{\boldsymbol{e_\sigma}}\,,\,\nabla\boldsymbol{\phi})_{\mathcal{T}_h} \\
&\quad + (\underline{\boldsymbol{e_\sigma}}\,,\,\underline{\boldsymbol{\xi}})_{\mathcal{T}_h} \\
&= (\boldsymbol{e_u}\,,\,\nabla\cdot\underline{\boldsymbol{\Pi}^{\boldsymbol{D}}_{\boldsymbol{V}}\boldsymbol{\psi}})_{\mathcal{T}_h} + (\mathcal{A}\underline{\boldsymbol{e_\sigma}}\,,\,\underline{\boldsymbol{\Pi}^{\boldsymbol{D}}_{\boldsymbol{V}}\boldsymbol{\psi}})_{\mathcal{T}_h} + (\boldsymbol{e_u}\,,\,\nabla\cdot\underline{\boldsymbol{\delta_\psi}})_{\mathcal{T}_h} + (\mathcal{A}\underline{\boldsymbol{e_\sigma}}\,,\,\underline{\boldsymbol{\delta_\psi}})_{\mathcal{T}_h} \\
&\quad - (\underline{\boldsymbol{e_\sigma}}\,,\,\nabla\boldsymbol{\Pi}^{\boldsymbol{D}}_{\boldsymbol{W}}\boldsymbol{\phi})_{\mathcal{T}_h} - (\underline{\boldsymbol{e_\sigma}}\,,\,\nabla\boldsymbol{\delta_\phi})_{\mathcal{T}_h} \\
&\quad + (\underline{\boldsymbol{e_\sigma}}\,,\,\underline{\boldsymbol{\xi}})_{\mathcal{T}_h}.
\end{aligned}
$$

Next, note that

$$
(\boldsymbol{e_u}\,,\,\nabla\cdot\underline{\boldsymbol{\delta_\psi}})_{\mathcal{T}_h} = \langle\boldsymbol{e_u}\,,\,\underline{\boldsymbol{\delta_\psi}}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - (\nabla\boldsymbol{e_u}\,,\,\underline{\boldsymbol{\delta_\psi}})_{\mathcal{T}_h} = \langle\boldsymbol{e_u}\,,\,\underline{\boldsymbol{\delta_\psi}}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h},
$$

by the property (D.1) of the projection and the fact that $\boldsymbol{e_u}\in\boldsymbol{W}_h$. Similarly, we have

$$
(\underline{\boldsymbol{e_\sigma}}\,,\,\nabla\boldsymbol{\delta_\phi})_{\mathcal{T}_h} = \langle\underline{\boldsymbol{e_\sigma}}\boldsymbol{n}\,,\,\boldsymbol{\delta_\phi}\rangle_{\partial\mathcal{T}_h}.
$$

Inserting these two equations into the first equation, we get

$$
\begin{aligned}
(\boldsymbol{e_u}\,,\,\boldsymbol{\theta})_{\mathcal{T}_h} &= (\boldsymbol{e_u}\,,\,\nabla\cdot\underline{\boldsymbol{\Pi}^{\boldsymbol{D}}_{\boldsymbol{V}}\boldsymbol{\psi}})_{\mathcal{T}_h} + (\mathcal{A}\underline{\boldsymbol{e_\sigma}}\,,\,\underline{\boldsymbol{\Pi}^{\boldsymbol{D}}_{\boldsymbol{V}}\boldsymbol{\psi}})_{\mathcal{T}_h} \\
&\quad - (\underline{\boldsymbol{e_\sigma}}\,,\,\nabla\boldsymbol{\Pi}^{\boldsymbol{D}}_{\boldsymbol{W}}\boldsymbol{\phi})_{\mathcal{T}_h} \\
&\quad + (\underline{\boldsymbol{e_\sigma}}\,,\,\underline{\boldsymbol{\xi}})_{\mathcal{T}_h} + (\mathcal{A}\underline{\boldsymbol{e_\sigma}}\,,\,\underline{\boldsymbol{\delta_\psi}})_{\mathcal{T}_h} + \langle\boldsymbol{e_u}\,,\,\underline{\boldsymbol{\delta_\psi}}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle\underline{\boldsymbol{e_\sigma}}\boldsymbol{n}\,,\,\boldsymbol{\delta_\phi}\rangle_{\partial\mathcal{T}_h},
\end{aligned}
$$

and taking $\underline{\boldsymbol{v}} := \underline{\boldsymbol{\Pi}^{\boldsymbol{D}}_{\boldsymbol{V}}\boldsymbol{\psi}}$ and $\boldsymbol{\omega} := \boldsymbol{\Pi}^{\boldsymbol{D}}_{\boldsymbol{W}}\boldsymbol{\phi}$ in the error equations (6.12a) and (6.12b), respectively, we obtain

$$
\begin{aligned}
(\boldsymbol{e_u}\,,\,\boldsymbol{\theta})_{\mathcal{T}_h} &= (\underline{\boldsymbol{e_\sigma}}\,,\,\underline{\boldsymbol{\xi}})_{\mathcal{T}_h} + (\mathcal{A}\underline{\boldsymbol{e_\sigma}}\,,\,\underline{\boldsymbol{\delta_\psi}})_{\mathcal{T}_h} - (\underline{\boldsymbol{e_\rho}}\,,\,\underline{\boldsymbol{\Pi}^{\boldsymbol{D}}_{\boldsymbol{V}}\boldsymbol{\psi}})_{\mathcal{T}_h} - (\mathcal{A}\underline{\boldsymbol{\delta_\sigma}} + \underline{\boldsymbol{\delta_\rho}}\,,\,\underline{\boldsymbol{\Pi}^{\boldsymbol{D}}_{\boldsymbol{V}}\boldsymbol{\psi}})_{\mathcal{T}_h} \\
&\quad + \langle\boldsymbol{e_u}\,,\,\underline{\boldsymbol{\delta_\psi}}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle\underline{\boldsymbol{e_\sigma}}\boldsymbol{n}\,,\,\boldsymbol{\delta_\phi}\rangle_{\partial\mathcal{T}_h} + \langle\boldsymbol{e_{\widehat{u}}}\,,\,\underline{\boldsymbol{\Pi}^{\boldsymbol{D}}_{\boldsymbol{V}}\boldsymbol{\psi}}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle\underline{\boldsymbol{e_{\widehat{\sigma}}}}\boldsymbol{n}\,,\,\boldsymbol{\Pi}^{\boldsymbol{D}}_{\boldsymbol{W}}\boldsymbol{\phi}\rangle_{\partial\mathcal{T}_h}.
\end{aligned}
$$

Next, note that $(\underline{\boldsymbol{e_\rho}}\,,\,\underline{\boldsymbol{\psi}})_{\mathcal{T}_h} = 0$ since $\underline{\boldsymbol{e_\rho}} \in \underline{\boldsymbol{AS}}(\Omega)$ and $\underline{\boldsymbol{\psi}}$ is a symmetric tensor. Also, note that by the regularity assumption, $(\underline{\boldsymbol{\psi}}, \boldsymbol{\phi}) \in \underline{\boldsymbol{H}}^1(\Omega)\times\boldsymbol{H}^1(\Omega)$, so $\underline{\boldsymbol{\psi}}\boldsymbol{n}, \boldsymbol{\phi}$ are single-valued on each face $F\in\mathcal{E}_h$. This implies that

$$
\begin{aligned}
&\langle\boldsymbol{e_{\widehat{u}}}\,,\,\underline{\boldsymbol{\psi}}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} = \langle\boldsymbol{e_{\widehat{u}}}\,,\,\underline{\boldsymbol{\psi}}\boldsymbol{n}\rangle_{\partial\Omega} = 0, && \text{by (6.12e)}, \\
&\langle\underline{\boldsymbol{e_{\widehat{\sigma}}}}\boldsymbol{n}\,,\,\boldsymbol{\phi}\rangle_{\partial\mathcal{T}_h} = \langle\underline{\boldsymbol{e_{\widehat{\sigma}}}}\boldsymbol{n}\,,\,\boldsymbol{P}^{\boldsymbol{D}}_{\boldsymbol{M}}\boldsymbol{\phi}\rangle_{\partial\mathcal{T}_h} = \langle\underline{\boldsymbol{e_{\widehat{\sigma}}}}\boldsymbol{n}\,,\,\boldsymbol{P}^{\boldsymbol{D}}_{\boldsymbol{M}}\boldsymbol{\phi}\rangle_{\partial\Omega} = 0 && \text{by (6.12d) and (6.7c)}.
\end{aligned}
$$

Inserting these three terms into the previous equation, we can write

$$(\boldsymbol{e_u}\,,\,\boldsymbol{\theta})_{\mathcal{T}_h} = (\underline{\boldsymbol{e_\sigma}}\,,\,\underline{\boldsymbol{\xi}})_{\mathcal{T}_h} + (\mathcal{A}\underline{\boldsymbol{e_\sigma}}\,,\,\underline{\boldsymbol{\delta_\psi}})_{\mathcal{T}_h} + (\underline{\boldsymbol{e_\rho}}\,,\,\underline{\boldsymbol{\delta_\psi}})_{\mathcal{T}_h} - (\mathcal{A}\underline{\boldsymbol{\delta_\sigma}} + \underline{\boldsymbol{\delta_\rho}}\,,\,\underline{\boldsymbol{\Pi_V^D\psi}})_{\mathcal{T}_h} + \mathbb{T}.$$

Here

$$\mathbb{T} := \langle \boldsymbol{e_u}\,,\,\underline{\boldsymbol{\delta_\psi}}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle \underline{\boldsymbol{e_\sigma}}\boldsymbol{n}\,,\,\boldsymbol{\delta_\phi}\rangle_{\partial\mathcal{T}_h} - \langle \boldsymbol{e_{\widehat{u}}}\,,\,\underline{\boldsymbol{\delta_\psi}}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} + \langle \underline{\boldsymbol{e_{\widehat{\sigma}}}}\boldsymbol{n}\,,\,\boldsymbol{\delta_\phi}\rangle_{\partial\mathcal{T}_h}.$$

We only need to show that

$$(\underline{\boldsymbol{e_\sigma}}\,,\,\underline{\boldsymbol{\xi}})_{\mathcal{T}_h} = (\underline{\boldsymbol{e_\sigma}}\,,\,\underline{\boldsymbol{\delta_\xi}})_{\mathcal{T}_h} - (\underline{\boldsymbol{\delta_\sigma}}\,,\,\underline{\boldsymbol{\Pi_A\xi}})_{\mathcal{T}_h} \quad \text{and} \quad \mathbb{T} = 0.$$

By (6.12c) we can write

$$(\underline{\boldsymbol{e_\sigma}}\,,\,\underline{\boldsymbol{\xi}})_{\mathcal{T}_h} = (\underline{\boldsymbol{e_\sigma}}\,,\,\underline{\boldsymbol{\delta_\xi}})_{\mathcal{T}_h} + (\underline{\boldsymbol{e_\sigma}}\,,\,\underline{\boldsymbol{\Pi_A\xi}})_{\mathcal{T}_h} = (\underline{\boldsymbol{e_\sigma}}\,,\,\underline{\boldsymbol{\delta_\xi}})_{\mathcal{T}_h} - (\underline{\boldsymbol{\delta_\sigma}}\,,\,\underline{\boldsymbol{\Pi_A\xi}})_{\mathcal{T}_h}.$$

Let us end the proof by showing that $\mathbb{T} = 0$. We can write

$$\begin{aligned}\mathbb{T} &= \langle \boldsymbol{e_u} - \boldsymbol{e_{\widehat{u}}}\,,\,\underline{\boldsymbol{\delta_\psi}}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle \underline{\boldsymbol{e_\sigma}}\boldsymbol{n} - \underline{\boldsymbol{e_{\widehat{\sigma}}}}\boldsymbol{n}\,,\,\boldsymbol{\delta_\phi}\rangle_{\partial\mathcal{T}_h}\\ &= \langle \boldsymbol{e_u} - \boldsymbol{e_{\widehat{u}}}\,,\,\underline{\boldsymbol{\delta_\psi}}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle \underline{\boldsymbol{e_\sigma}}\boldsymbol{n} - \underline{\boldsymbol{e_{\widehat{\sigma}}}}\boldsymbol{n}\,,\,\boldsymbol{P_M^D\phi} - \boldsymbol{\Pi_W^D\phi}\rangle_{\partial\mathcal{T}_h}\end{aligned}$$

by a similar argument as in the proof of Lemma 6.2.2,

$$\begin{aligned}\mathbb{T} &= \langle \boldsymbol{e_u} - \boldsymbol{e_{\widehat{u}}}\,,\,\underline{\boldsymbol{\delta_\psi}}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle \boldsymbol{\alpha}(\boldsymbol{e_u} - \boldsymbol{e_{\widehat{u}}})\,,\,\boldsymbol{P_M^D\phi} - \boldsymbol{\Pi_W^D\phi}\rangle_{\partial\mathcal{T}_h}\\ &= \langle \boldsymbol{e_u} - \boldsymbol{e_{\widehat{u}}}\,,\,\underline{\boldsymbol{\delta_\psi}}\boldsymbol{n} - \boldsymbol{\alpha}(\boldsymbol{P_M^D\phi} - \boldsymbol{\Pi_W^D\phi})\rangle_{\partial\mathcal{T}_h}\\ &= 0,\end{aligned}$$

by the property (D.1) of the projection. This completes the proof. $\square$

As a consequence of the result just proved, we can obtain our estimate of $\boldsymbol{e_u}$.

Corollary 6.2.5. *If the elliptic regularity property* (6.8) *holds for $s = 1$, then we have*

$$\|\boldsymbol{e_u}\|_\Omega \le Ch(\|\underline{\boldsymbol{e_\sigma}}\|_\Omega + \|\underline{\boldsymbol{e_\rho}}\|_\Omega + \|\underline{\boldsymbol{\delta_\sigma}}\|_\Omega + \|\underline{\boldsymbol{\delta_\rho}}\|_\Omega).$$

Proof. Taking $\boldsymbol{\theta} = \boldsymbol{e_u}$ in Lemma 6.2.4, we can write

$$
\begin{aligned}
\|\boldsymbol{e_u}\|_\Omega^2 &= (\mathcal{A}\underline{\boldsymbol{e_\sigma}}\,,\,\underline{\boldsymbol{\delta_\psi}})_{\mathcal{T}_h} + (\underline{\boldsymbol{e_\rho}}\,,\,\underline{\boldsymbol{\delta_\psi}})_{\mathcal{T}_h} + (\underline{\boldsymbol{e_\sigma}}\,,\,\underline{\boldsymbol{\delta_\xi}})_{\mathcal{T}_h}\\
&\quad - (\underline{\boldsymbol{\delta_\sigma}}\,,\,\underline{\boldsymbol{\Pi_A \xi}})_{\mathcal{T}_h} - (\mathcal{A}\underline{\boldsymbol{\delta_\sigma}}\,,\,\underline{\boldsymbol{\Pi_V^D \psi}})_{\mathcal{T}_h} - (\underline{\boldsymbol{\delta_\rho}}\,,\,\underline{\boldsymbol{\Pi_V^D \psi}})_{\mathcal{T}_h}\\
&= (\mathcal{A}\underline{\boldsymbol{e_\sigma}}\,,\,\underline{\boldsymbol{\delta_\psi}})_{\mathcal{T}_h} + (\underline{\boldsymbol{e_\rho}}\,,\,\underline{\boldsymbol{\delta_\psi}})_{\mathcal{T}_h} + (\underline{\boldsymbol{e_\sigma}}\,,\,\underline{\boldsymbol{\delta_\xi}})_{\mathcal{T}_h}\\
&\quad + (\underline{\boldsymbol{\delta_\sigma}}\,,\,\underline{\boldsymbol{\delta_\xi}})_{\mathcal{T}_h} + (\mathcal{A}\underline{\boldsymbol{\delta_\sigma}}\,,\,\underline{\boldsymbol{\delta_\psi}})_{\mathcal{T}_h} + (\underline{\boldsymbol{\delta_\rho}}\,,\,\underline{\boldsymbol{\delta_\psi}})_{\mathcal{T}_h}\\
&\quad - (\underline{\boldsymbol{\delta_\sigma}}\,,\,\underline{\boldsymbol{\xi}})_{\mathcal{T}_h} - (\mathcal{A}\underline{\boldsymbol{\delta_\sigma}}\,,\,\underline{\boldsymbol{\psi}})_{\mathcal{T}_h} - (\underline{\boldsymbol{\delta_\rho}}\,,\,\underline{\boldsymbol{\psi}})_{\mathcal{T}_h}.
\end{aligned}
$$

Using the dual equation (6.7b) and the fact that $\underline{\boldsymbol{\delta_\rho}}$ is antisymmetric and $\underline{\boldsymbol{\psi}}$ is symmetric, we get

$$
\begin{aligned}
\|\boldsymbol{e_u}\|_\Omega^2 &= (\mathcal{A}\underline{\boldsymbol{e_\sigma}}\,,\,\underline{\boldsymbol{\delta_\psi}})_{\mathcal{T}_h} + (\underline{\boldsymbol{e_\rho}}\,,\,\underline{\boldsymbol{\delta_\psi}})_{\mathcal{T}_h} + (\underline{\boldsymbol{e_\sigma}}\,,\,\underline{\boldsymbol{\delta_\xi}})_{\mathcal{T}_h}\\
&\quad + (\underline{\boldsymbol{\delta_\sigma}}\,,\,\underline{\boldsymbol{\delta_\xi}})_{\mathcal{T}_h} + (\mathcal{A}\underline{\boldsymbol{\delta_\sigma}}\,,\,\underline{\boldsymbol{\delta_\psi}})_{\mathcal{T}_h} + (\underline{\boldsymbol{\delta_\rho}}\,,\,\underline{\boldsymbol{\delta_\psi}})_{\mathcal{T}_h}\\
&\quad - (\underline{\boldsymbol{\delta_\sigma}}\,,\,\nabla\phi)_{\mathcal{T}_h}\\
&= (\mathcal{A}\underline{\boldsymbol{e_\sigma}}\,,\,\underline{\boldsymbol{\delta_\psi}})_{\mathcal{T}_h} + (\underline{\boldsymbol{e_\rho}}\,,\,\underline{\boldsymbol{\delta_\psi}})_{\mathcal{T}_h} + (\underline{\boldsymbol{e_\sigma}}\,,\,\underline{\boldsymbol{\delta_\xi}})_{\mathcal{T}_h}\\
&\quad + (\underline{\boldsymbol{\delta_\sigma}}\,,\,\underline{\boldsymbol{\delta_\xi}})_{\mathcal{T}_h} + (\mathcal{A}\underline{\boldsymbol{\delta_\sigma}}\,,\,\underline{\boldsymbol{\delta_\psi}})_{\mathcal{T}_h} + (\underline{\boldsymbol{\delta_\rho}}\,,\,\underline{\boldsymbol{\delta_\psi}})_{\mathcal{T}_h}\\
&\quad - (\underline{\boldsymbol{\delta_\sigma}}\,,\,\nabla\phi - \underline{\boldsymbol{P}}_0\nabla\phi)_{\mathcal{T}_h},
\end{aligned}
$$

by the properties (D.1) and (D.5) of the projection. Here $\underline{\boldsymbol{P}}_0$ is the the L^2-projection onto $\underline{\boldsymbol{P}}^0(K)$ on each $K \in \mathcal{T}_h$.

Applying the Cauchy-Schwarz inequality, we obtain

$$
\begin{aligned}
\|\boldsymbol{e_u}\|_\Omega^2 &\le C(\|\underline{\boldsymbol{e_\sigma}}\|_\Omega + \|\underline{\boldsymbol{e_\rho}}\|_\Omega + \|\underline{\boldsymbol{\delta_\sigma}}\|_\Omega + \|\underline{\boldsymbol{\delta_\rho}}\|_\Omega)\\
&\qquad \times (\|\underline{\boldsymbol{\delta_\psi}}\|_\Omega + \|\underline{\boldsymbol{\delta_\xi}}\|_\Omega + \|\nabla\phi - \underline{\boldsymbol{P}}_0\nabla\phi\|_\Omega)
\end{aligned}
$$

and using the approximation property (D.4) and the fact that, by property (A.1), we have that

$$
\|\underline{\boldsymbol{\delta_\xi}}\|_\Omega \le Ch|\underline{\boldsymbol{\xi}}|_{H^1(\Omega)},
$$

we get

$$
\begin{aligned}
\|\boldsymbol{e_u}\|_\Omega^2 &\le Ch(\|\underline{\boldsymbol{e_\sigma}}\|_\Omega + \|\underline{\boldsymbol{e_\rho}}\|_\Omega + \|\underline{\boldsymbol{\delta_\sigma}}\|_\Omega + \|\underline{\boldsymbol{\delta_\rho}}\|_\Omega)(|\underline{\boldsymbol{\psi}}|_{H^1(\Omega)} + |\underline{\boldsymbol{\xi}}|_{H^1(\Omega)})\\
&\le Ch(\|\underline{\boldsymbol{e_\sigma}}\|_\Omega + \|\underline{\boldsymbol{e_\rho}}\|_\Omega + \|\underline{\boldsymbol{\delta_\sigma}}\|_\Omega + \|\underline{\boldsymbol{\delta_\rho}}\|_\Omega)\|\boldsymbol{e_u}\|_\Omega,
\end{aligned}
$$

by the elliptic regularity property (6.8) with $s = 1$. This completes the proof. $\square$

Step 4: Estimate of $\underline{\boldsymbol{e_\rho}}$. By Corollaries 6.2.3 and 6.2.5 we can see that it only remains to obtain an estimate for the error $\underline{\boldsymbol{e_\rho}}$. Next, we show that is bounded by the projection errors $\underline{\boldsymbol{\delta_\sigma}}, \underline{\boldsymbol{\delta_\rho}}$. To do this, we write

$$\underline{\boldsymbol{e_\rho}} = \underline{\boldsymbol{e_\rho}}^0 + \underline{\boldsymbol{e_\rho}}^c,$$

where $\underline{\boldsymbol{e_\rho}}^c|_K = \frac{1}{|K|}\int_K \underline{\boldsymbol{e_\rho}} dx$ for all $K \in \mathcal{T}_h$ and $\underline{\boldsymbol{e_\rho}}^0 = \underline{\boldsymbol{e_\rho}} - \underline{\boldsymbol{e_\rho}}^c$. Next we bound these two terms separately. In what follows, we only present the proof for $3D$ case. The proof of $2D$ is a simpler version.

In order to control $\underline{\boldsymbol{e_\rho}}^0$, we first define the subspace of $\underline{\boldsymbol{A}}_h$ with average zero on each element:

$$\underline{\boldsymbol{A}}_h^0 := \{\underline{\boldsymbol{\eta}} \in \underline{\boldsymbol{A}}_h : (\underline{\boldsymbol{\eta}}, \underline{\boldsymbol{v}})_K = 0 \text{ for all } \underline{\boldsymbol{v}} \in \underline{\boldsymbol{P}}^0(K) \text{ and for all } K \in \mathcal{T}_h\}.$$

Note that the condition (A.1) ensures that the above set is a subspace of $\underline{\boldsymbol{A}}_h^0$.

We are now ready to give the estimate of $\underline{\boldsymbol{e_\rho}}^0$.

Theorem 6.2.6. *We have*

$$\|\underline{\boldsymbol{e_\rho}}^0\|_\Omega \le C(\|\underline{\boldsymbol{e_\sigma}}\|_\Omega + \|\underline{\boldsymbol{\delta_\sigma}}\|_\Omega + \|\underline{\boldsymbol{\delta_\rho}}\|_\Omega).$$

To prove this theorem, we are going to use the following result.

Lemma 6.2.7 ([33])**.** *Given $\underline{\boldsymbol{\eta}} \in \underline{\boldsymbol{A}}_h^0$, there exists $\underline{\boldsymbol{v}} \in \underline{\boldsymbol{B}}_h$ such that*

$$(\underline{\boldsymbol{\eta}}\,, \underline{\gamma})_{\mathcal{T}_h} = (\underline{\boldsymbol{v}}\,, \underline{\gamma})_{\mathcal{T}_h} \quad \textit{for all} \quad \underline{\gamma} \in \underline{\boldsymbol{A}}_h. \tag{6.13a}$$

$$\|\underline{\boldsymbol{v}}\|_\Omega \le C\|\underline{\boldsymbol{\eta}}\|_\Omega. \tag{6.13b}$$

Proof. To prove Theorem 6.2.6, we begin by rewriting the error equation (6.12a) as follows:

$$\begin{aligned}(\mathcal{A}\underline{\boldsymbol{e_\sigma}}\,, \underline{\boldsymbol{v}})_{\mathcal{T}_h} + (\boldsymbol{e_u}\,, \nabla\cdot\underline{\boldsymbol{v}})_{\mathcal{T}_h} + (\underline{\boldsymbol{e_\rho}}^0\,, \underline{\boldsymbol{v}})_{\mathcal{T}_h} + (\underline{\boldsymbol{e_\rho}}^c\,, \underline{\boldsymbol{v}})_{\mathcal{T}_h} - \langle \boldsymbol{e_{\widehat{u}}}\,, \underline{\boldsymbol{v}}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} \\ = -(\mathcal{A}\underline{\boldsymbol{\delta_\sigma}}\,, \underline{\boldsymbol{v}})_{\mathcal{T}_h} - (\underline{\boldsymbol{\delta_\rho}}\,, \underline{\boldsymbol{v}})_{\mathcal{T}_h}\end{aligned}$$

If we take $\underline{\boldsymbol{v}} \in \underline{\boldsymbol{B}}_h$ to be the function given by Lemma 6.2.7 with $\underline{\boldsymbol{\eta}} := \underline{\boldsymbol{e_\rho}}^0$, we have that

$$
\begin{aligned}
(\boldsymbol{e_u}\,,\,\nabla\cdot\underline{\boldsymbol{v}})_{\mathcal{T}_h} &= 0 && \text{by property (B.1)},\\
(\underline{\boldsymbol{e_\rho}}^c\,,\,\underline{\boldsymbol{v}})_{\mathcal{T}_h} &= (\underline{\boldsymbol{e_\rho}}^c\,,\,\underline{\boldsymbol{e_\rho}}^0)_{\mathcal{T}_h} = 0 && \text{by (6.13a)},\\
\langle \boldsymbol{e_{\widehat{u}}}\,,\,\underline{\boldsymbol{v}}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} &= 0 && \text{by property (B.2)},\\
(\underline{\boldsymbol{e_\rho}}^0\,,\,\underline{\boldsymbol{v}})_{\mathcal{T}_h} &= \|\underline{\boldsymbol{e_\rho}}^0\|_\Omega^2,
\end{aligned}
$$

by the property (6.13a) in Lemma 6.2.7. So, after a simple rearrangement of terms, we get that

$$\|\underline{\boldsymbol{e_\rho}}^0\|_\Omega^2 = -(\mathcal{A}\underline{\boldsymbol{e_\sigma}}\,,\,\underline{\boldsymbol{v}})_{\mathcal{T}_h} - (\mathcal{A}\underline{\boldsymbol{\delta_\sigma}}\,,\,\underline{\boldsymbol{v}})_{\mathcal{T}_h} - (\underline{\boldsymbol{\delta_\rho}}\,,\,\underline{\boldsymbol{v}})_{\mathcal{T}_h}.$$

The final result can be obtained by applying the Cauchy-Schwarz inequality and the estimate (6.13b) of Lemma 6.2.7. This completes the proof. □

The next result completes the estimate of $\underline{\boldsymbol{e_\rho}}$.

Theorem 6.2.8. *If $\underline{\boldsymbol{G}}(K) \subset \underline{\boldsymbol{V}}(K)$, then we have*

$$\|\underline{\boldsymbol{e_\rho}}^c\|_\Omega \le C(\|\underline{\boldsymbol{e_\sigma}}\|_\Omega + \|\underline{\boldsymbol{\delta_\sigma}}\|_\Omega + \|\underline{\boldsymbol{\delta_\rho}}\|_\Omega).$$

To prove this theorem, we need to use the following result.

Lemma 6.2.9. *Given $\underline{\boldsymbol{\eta}} \in \underline{\boldsymbol{A}}_h^c := \underline{\boldsymbol{A}}_h \cap \underline{\boldsymbol{P}}^0(\mathcal{T}_h)$, there exists $\underline{\boldsymbol{v}} \in \underline{\boldsymbol{G}}_h$ such that*

$$\nabla\cdot\underline{\boldsymbol{v}} = 0, \tag{6.14a}$$

$$(\underline{\boldsymbol{v}}\,,\,\underline{\boldsymbol{\gamma}})_{\mathcal{T}_h} = (\underline{\boldsymbol{\eta}}\,,\,\underline{\boldsymbol{\gamma}})_{\mathcal{T}_h} \quad \text{for all } \underline{\boldsymbol{\gamma}} \in \underline{\boldsymbol{A}}_h^c, \tag{6.14b}$$

$$\|\underline{\boldsymbol{v}}\|_\Omega \le C\|\underline{\boldsymbol{\eta}}\|_\Omega. \tag{6.14c}$$

Proof. To prove Theorems 6.2.8, we begin by taking $\underline{\boldsymbol{v}} \in \underline{\boldsymbol{A}}_h^c$ to be the function given by Lemma 6.2.9 with $\underline{\boldsymbol{\eta}} := \underline{\boldsymbol{e_\rho}}^c$ in the error equation (6.12a):

$$
\begin{aligned}
(\mathcal{A}\underline{\boldsymbol{e_\sigma}}\,,\,\underline{\boldsymbol{v}})_{\mathcal{T}_h} + (\boldsymbol{e_u}\,,\,\nabla\cdot\underline{\boldsymbol{v}})_{\mathcal{T}_h} + (\underline{\boldsymbol{e_\rho}}^0\,,\,\underline{\boldsymbol{v}})_{\mathcal{T}_h} + (\underline{\boldsymbol{e_\rho}}^c\,,\,\underline{\boldsymbol{v}})_{\mathcal{T}_h} - \langle \boldsymbol{e_{\widehat{u}}}\,,\,\underline{\boldsymbol{v}}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}&\\
= -(\mathcal{A}\underline{\boldsymbol{\delta_\sigma}}\,,\,\underline{\boldsymbol{v}})_{\mathcal{T}_h} - (\underline{\boldsymbol{\delta_\rho}}\,,\,\underline{\boldsymbol{v}})_{\mathcal{T}_h}&.
\end{aligned}
$$

Since we have that

$$
\begin{aligned}
(\boldsymbol{e}_{\boldsymbol{u}}\,,\, \nabla\cdot \underline{\boldsymbol{v}})_{\mathcal{T}_h} &= 0 && \text{by property (6.14a) of Lemma 6.14,}\\
\langle \boldsymbol{e}_{\widehat{\boldsymbol{u}}}\,,\, \underline{\boldsymbol{v}}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} &= \langle \boldsymbol{e}_{\widehat{\boldsymbol{u}}}, \underline{\boldsymbol{v}}\boldsymbol{n}\rangle_{\partial\Omega} && \text{since } \underline{\boldsymbol{v}} \in \underline{\mathbf{H}}(div,\Omega),\\
&= 0 && \text{by the error equation (6.12e),}\\
(\underline{\boldsymbol{e}_{\boldsymbol{\rho}}}^{c}\,,\, \underline{\boldsymbol{v}})_{\mathcal{T}_h} &= \|\underline{\boldsymbol{e}_{\boldsymbol{\rho}}}^{c}\|_{\Omega}^{2},
\end{aligned}
$$

by property (6.14b) of Lemma 6.14, we can rewrite the first error equation as follows:

$$
\|\underline{\boldsymbol{e}_{\boldsymbol{\rho}}}^{c}\|_{\Omega}^{2} = -(\mathcal{A}\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\,,\, \underline{\boldsymbol{v}})_{\mathcal{T}_h} - (\underline{\boldsymbol{e}_{\boldsymbol{\rho}}}^{0}\,,\, \underline{\boldsymbol{v}})_{\mathcal{T}_h} - (\mathcal{A}\underline{\boldsymbol{\delta}_{\boldsymbol{\sigma}}}\,,\, \underline{\boldsymbol{v}})_{\mathcal{T}_h} - (\underline{\boldsymbol{\delta}_{\boldsymbol{\rho}}}\,,\, \underline{\boldsymbol{v}})_{\mathcal{T}_h}.
$$

The estimate now is obtianed by using Cauchy-Schwarz inequality, the estimate (6.14c) of Lemma 6.14 and the estimate of Lemma 6.2.6. This completes the proof. □

Combining Theorem 6.2.6 and 6.2.8 we can obtain an estimate of $\underline{\boldsymbol{e}_{\boldsymbol{\rho}}}$:

$$
\|\underline{\boldsymbol{e}_{\boldsymbol{\rho}}}\|_{\Omega} \le C(\|\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\|_{\Omega} + \|\underline{\boldsymbol{\delta}_{\boldsymbol{\sigma}}}\|_{\Omega} + \|\underline{\boldsymbol{\delta}_{\boldsymbol{\rho}}}\|_{\Omega}). \tag{6.15}
$$

It remains to prove Lemma 6.2.9. To proceed, we need the following result.

Lemma 6.2.10. *On each K, let r denote its number of faces and s its the number of vertices. Then for every m in the space*

$$
Q_h := \{w \in L^2(\Omega) : w|_K \in P^0(K) \quad \forall\, K \in \mathcal{T}_h\},
$$

there exists a function $\boldsymbol{v}$ in the space

$$
\boldsymbol{S}_h := \{\boldsymbol{\phi} \in \boldsymbol{H}^1(\Omega) : \phi_i|_K \in E(K) + Z(K),\ i = 1, 2, \ldots, s\},
$$

such that

$$
P_Q \nabla\cdot \boldsymbol{v} = m, \tag{6.16a}
$$

$$
\|\boldsymbol{v}\|_{\boldsymbol{H}^1(\Omega)} \le C\|m\|_{\Omega}, \tag{6.16b}
$$

where P_Q is the L^2-projection onto Q_h, $E(K)$ is an r-dimensional function space and $Z(K)$ is an s-dimensional function space defined on K.

Let us first give the precise definition of the spaces $E(K)$ and $Z(K)$. For each element K and each face $F \in \mathcal{F}(K)$, we define

$$\lambda_F(\boldsymbol{x}) := \frac{dist(\boldsymbol{x}, F)}{dist(\boldsymbol{o}_K, F)}.$$

Here $dist(\boldsymbol{x}, F)$ is the function of distance between $\boldsymbol{x}$ and F, $\boldsymbol{o}_K$ denotes the barycenter of K. Next, on each *interior* face $F = \partial K^+ \cap \partial K^-$, we define the face-bubble function as

$$\Lambda_F := \prod_{F' \in \mathcal{F}(K^+) \cup \mathcal{F}(K^-), F' \neq F} \lambda_{F'},$$

and on each *boundary* face $F \in \partial\Omega$, we define

$$\Lambda_F := \prod_{F' \in \mathcal{F}(K)} \lambda_{F'}.$$

Finally, on each K, we define the local face-bubble function space as

$$E(K) := span_{F \in \mathcal{F}(K)}\{\Lambda_F\}.$$

In general, $E(K)$ is a r-dimensional function space.

Let $\widehat{K}$ be the reference element of the given triangulation $\mathcal{T}_h$ and let $\{\widehat{\phi}_i\}_{i=1}^s$ any set of function satisfying the following properties:

- $\widehat{\phi}_i(\widehat{x}_j) = \delta_{ij}$.
- $\widehat{\phi}_i|_{\widehat{F}} = 0$, if $\widehat{\boldsymbol{x}}_i \notin \widehat{F} \subset \partial\widehat{K}$.
- $\{\widehat{\phi})i\}_{i=1}^s$ is a partition of unity on $\widehat{K}$,that is, $\sum_{i=1}^s \widehat{\phi}_i(\widehat{\boldsymbol{x}}) = 1$, $\forall \widehat{\boldsymbol{x}} \in \widehat{K}$.

For each K, let $\boldsymbol{F}_K$ denote the deformation mapping from $\widehat{K}$ to K. Then we define

$$Z(K) := span\{\phi_i(\boldsymbol{x})\}_{i=1}^s, \quad \phi_i(\boldsymbol{x}) = \widehat{\phi}_i \circ \boldsymbol{F}_K^{-1}(\boldsymbol{x}).$$

We are now ready to prove Lemma 6.2.10.

Proof. It is well known that there exists $\boldsymbol{\omega} \in \boldsymbol{H}^1(\Omega)$ such that

$$\nabla \cdot \boldsymbol{\omega} = m, \tag{6.17a}$$

$$\|\boldsymbol{\omega}\|_{\boldsymbol{H}^1(\Omega)} \leq C\|m\|_\Omega. \tag{6.17b}$$

We define the function $\boldsymbol{v}$ as follows

$$\boldsymbol{v} = \boldsymbol{I}\boldsymbol{\omega} + \boldsymbol{\pi}(\boldsymbol{\omega} - \boldsymbol{I}\boldsymbol{\omega}),$$

where $\boldsymbol{I}$ is the Clément interpolant [54, 55] onto

$$\boldsymbol{Z}_h := \{\boldsymbol{\eta} \in \boldsymbol{H}^1(\Omega), \eta_i|_K \in Z(K)\ i = 1, \ldots, s\}.$$

We define the projection $\boldsymbol{\pi}$ component-wise as follows. For a scalar function u, on each $K \in \mathcal{T}_h$, $\pi u \in E(K)$ is the only function satisfying

$$\langle u - \pi u\,,\, 1\rangle_F = 0 \quad \text{for all faces } F \in \mathcal{F}(K).$$

It is easy to see the projection is well defined. Next, we show that the projection is $H^1(\Omega)-$conforming. It suffices to show that πu is continuous accross any interface $F = K^+ \cap K^-$, given $u \in H^1(\Omega)$. Let $\pi^+ u, \pi^- u$ denote the projection restricted on K^+, K^- respectively. We may write

$$\pi^+ u = \sum_{F' \in \mathcal{F}(K^+)} a^+_{F'} \Lambda^+_{F'}, \quad \pi^- u = \sum_{F' \in \mathcal{F}(K^-)} a^-_{F'} \Lambda^-_{F'}.$$

On the face F, we have

$$\langle \pi^+ u\,,\, 1\rangle_F = \langle \pi^- u\,,\, 1\rangle_F = \langle u\,,\, 1\rangle_F.$$

Note that $\pi^+ u|_F = a^+_F \Lambda^+_F$ since all other face-bubble functions vanish on the face F. Similarly, $\pi^- u|_F = a^-_F \Lambda^-_F$. This implies that

$$\int_F a^+_F \Lambda^+_F ds = \int_F a^-_F \Lambda^-_F ds.$$

By the definition of the local face-bubble functions, we know that $\Lambda^+_F = \Lambda^-_F > 0$ on F. Hence we must have $a^+_F = a^-_F$ and this implies that

$$\pi^+ u|_F = \pi^- u|_F.$$

By using the trace theorem and a scaling argument we get the following inequality:

$$\|\boldsymbol{\pi}\boldsymbol{u}\|_{\boldsymbol{H}^1(K)} \leq C(\|\boldsymbol{u}\|_{\boldsymbol{H}^1(K)} + \frac{1}{h_K}\|\boldsymbol{u}\|_K), \tag{6.18}$$

for each $K \in \mathcal{T}_h$. By the fact that $\boldsymbol{v} \in \boldsymbol{H}^1(\Omega)$ and $\boldsymbol{v}|_K \in \boldsymbol{P}^{2r-2}(K)$, we can see $\boldsymbol{v} \in \boldsymbol{S}_h$.

Now we are ready to show that $\boldsymbol{v}$ satisfies (6.16a), (6.16b). For any $w \in Q_h$ we have

$$\begin{aligned}
(\nabla \cdot \boldsymbol{v}, w)_K &= -(v, \nabla w)_K + \langle \boldsymbol{v} \cdot \boldsymbol{n}, w \rangle_{\partial K} && \text{by integration by parts,} \\
&= \langle \boldsymbol{v} \cdot \boldsymbol{n}, w \rangle_{\partial K} && \text{since } \nabla w = 0, \\
&= \langle (\boldsymbol{I\omega} + \boldsymbol{\pi}(\boldsymbol{\omega} - \boldsymbol{I\omega})) \cdot \boldsymbol{n}, w \rangle_{\partial K} && \text{by the definition of } \boldsymbol{v}, \\
&= \langle \boldsymbol{\omega} \cdot \boldsymbol{n}, w \rangle_{\partial K} && \text{by the definition of } \boldsymbol{\pi}, \\
&= (\nabla \cdot \boldsymbol{\omega}, w)_K + (\boldsymbol{\omega}, \nabla w)_K && \text{by integration by parts,} \\
&= (\nabla \cdot \boldsymbol{\omega}, w)_K && \text{since } \nabla w = 0, \\
&= (m, w)_K && \text{by (6.17a).}
\end{aligned}$$

This shows (6.16a). Next we have

$$\begin{aligned}
\|\boldsymbol{v}\|^2_{\boldsymbol{H}^1(\Omega)} &= \sum_{K \in \mathcal{T}_h} \|\boldsymbol{v}\|^2_{\boldsymbol{H}^1(K)} \\
&\leq \sum_{K \in \mathcal{T}_h} 2\|\boldsymbol{I\omega}\|^2_{\boldsymbol{H}^1(K)} + \sum_{K \in \mathcal{T}_h} 2\|\boldsymbol{\pi}(\boldsymbol{\omega} - \boldsymbol{I\omega})\|^2_{\boldsymbol{H}^1(K)} \\
&\leq 2\|\boldsymbol{I\omega}\|^2_{\boldsymbol{H}^1(\Omega)} + C \sum_{K \in \mathcal{T}_h} (\|\boldsymbol{\omega} - \boldsymbol{I\omega}\|_{\boldsymbol{H}^1(K)} + \frac{1}{h_K}\|\boldsymbol{\omega} - \boldsymbol{I\omega}\|_K)^2 \qquad \text{by (6.18),} \\
&\leq C\|\boldsymbol{\omega}\|_{\boldsymbol{H}^1(\Omega)} + C \sum_{K \in \mathcal{T}_h} (\|\boldsymbol{\omega} - \boldsymbol{I\omega}\|^2_{\boldsymbol{H}^1(K)} + \frac{1}{h_K^2}\|\boldsymbol{\omega} - \boldsymbol{I\omega}\|^2_K) \\
&\leq C\|\boldsymbol{\omega}\|_{\boldsymbol{H}^1(\Omega)},
\end{aligned}$$

by the standard property of the Clément interpolant. Finally, we get that

$$\|\boldsymbol{v}\|^2_{\boldsymbol{H}^1(\Omega)} \leq C\|m\|_\Omega,$$

by (6.17b). This completes the proof of Lemma 6.2.10. □

Now we are ready to prove Lemma 6.2.9.

Proof. of Lemma 6.2.9 For any $\underline{\boldsymbol{\eta}} \in \underline{\boldsymbol{A}}_h^c$, we can write

$$\underline{\boldsymbol{\eta}} = \begin{pmatrix} 0 & z_3 & -z_2 \\ -z_3 & 0 & z_1 \\ z_2 & -z_1 & 0 \end{pmatrix}$$

for some $z_1, z_2, z_3 \in Q_h$. Hence, we can find $\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3 \in \boldsymbol{H}^1(\Omega)$ such that

$$P_Q \nabla \cdot \boldsymbol{w}_i = -2z_i, \tag{6.19a}$$

$$\|\boldsymbol{w}_i\|_{\boldsymbol{H}^1(\Omega)} \leq C\|z_i\|_\Omega \tag{6.19b}$$

for $i = 1, 2, 3$. Let $\underline{\boldsymbol{w}}$ be the matrix whose i-th column is $\boldsymbol{w}_i$.

Since we have the following identity:

$$-2\underline{\boldsymbol{skw}}\nabla \times (\underline{\boldsymbol{w}} - \frac{\mathbf{tr}(\underline{\boldsymbol{w}})}{2}\underline{\boldsymbol{I}}) = \begin{pmatrix} 0 & \nabla \cdot \boldsymbol{w}_3 & -\nabla \cdot \boldsymbol{w}_2 \\ -\nabla \cdot \boldsymbol{w}_3 & 0 & \nabla \cdot \boldsymbol{w}_1 \\ \nabla \cdot \boldsymbol{w}_2 & -\nabla \cdot \boldsymbol{w}_1 & 0 \end{pmatrix},$$

where $\underline{\boldsymbol{skw}}\underline{\boldsymbol{v}} := (\underline{\boldsymbol{v}} - \underline{\boldsymbol{v}}^t)/2$, if we set $\underline{\boldsymbol{\Phi}} := \nabla \times (\underline{\boldsymbol{w}} - \frac{\mathbf{tr}(\underline{\boldsymbol{w}})}{2}\underline{\boldsymbol{I}})$, we can verify that we have

$$P_Q \underline{\boldsymbol{skw}}\underline{\boldsymbol{\Phi}} = \underline{\boldsymbol{\eta}}. \tag{6.20}$$

Let us now define $\underline{\boldsymbol{v}} = \underline{\boldsymbol{\Pi}_{\boldsymbol{G}}\boldsymbol{\Phi}}$, where $\underline{\boldsymbol{\Pi}_{\boldsymbol{G}}}$ is the projection into the space $\underline{\boldsymbol{G}}(K)$ defined by (6.6a). Using the projection property (6.6b) we have

$$\nabla \cdot \underline{\boldsymbol{v}} = \underline{\boldsymbol{P}}_0 \nabla \cdot \underline{\boldsymbol{\Phi}} = 0.$$

So (6.14a) holds.

Moreover, for any $\underline{\boldsymbol{\gamma}} \in \underline{\boldsymbol{A}}_h^c$, for each $K \in \mathcal{T}_h$, we have $\underline{\boldsymbol{\gamma}}|_K = \nabla \boldsymbol{p}$ for some $\boldsymbol{p} \in \boldsymbol{P}^1(K)$. Hence,

$$\begin{aligned}
(\underline{\boldsymbol{v}}\,,\,\underline{\boldsymbol{\gamma}})_{\mathcal{T}_h} &= (\underline{\boldsymbol{v}}\,,\,\nabla \boldsymbol{p})_{\mathcal{T}_h} \\
&= \langle \underline{\boldsymbol{v}}\boldsymbol{n}\,,\,\boldsymbol{p}\rangle_{\partial\mathcal{T}_h} && \text{by (6.14a),} \\
&= \langle \underline{\boldsymbol{\Pi}}_1\underline{\boldsymbol{\Phi}}\boldsymbol{n}\,,\,\boldsymbol{p}\rangle_{\partial\mathcal{T}_h} && \text{by definition of } \underline{\boldsymbol{v}}, \\
&= \langle \underline{\boldsymbol{\Phi}}\boldsymbol{n}\,,\,\boldsymbol{p}\rangle_{\partial\mathcal{T}_h} && \text{by (6.6a),} \\
&= \langle \nabla \times (\underline{\boldsymbol{w}} - \frac{\mathbf{tr}(\underline{\boldsymbol{w}})}{2}\underline{\boldsymbol{I}})\boldsymbol{n}\,,\,\boldsymbol{p}\rangle_{\partial\mathcal{T}_h} && \text{by definition of } \underline{\boldsymbol{\Phi}}, \\
&= (\nabla \times (\underline{\boldsymbol{w}} - \frac{\mathbf{tr}(\underline{\boldsymbol{w}})}{2}\underline{\boldsymbol{I}})\,,\,\nabla \boldsymbol{p})_{\mathcal{T}_h} && \text{by integration by parts,} \\
&= (\nabla \times (\underline{\boldsymbol{w}} - \frac{\mathbf{tr}(\underline{\boldsymbol{w}})}{2}\underline{\boldsymbol{I}})\,,\,\underline{\boldsymbol{\gamma}})_{\mathcal{T}_h} && \text{by definition of } \underline{\boldsymbol{\gamma}}, \\
&= (P_Q\underline{\boldsymbol{skw}}\nabla \times (\underline{\boldsymbol{w}} - \frac{\mathbf{tr}(\underline{\boldsymbol{w}})}{2}\underline{\boldsymbol{I}})\,,\,\underline{\boldsymbol{\gamma}})_{\mathcal{T}_h} && \text{by definition of } \underline{\boldsymbol{A}}_h^C, \\
&= (\underline{\boldsymbol{\eta}}\,,\,\underline{\boldsymbol{\gamma}})_{\mathcal{T}_h} && \text{by (6.20).}
\end{aligned}$$

This shows that (6.14b) holds.

Finally,

$$
\begin{aligned}
\|\underline{\boldsymbol{v}}\|_\Omega &= \sum_{K\in\mathcal{T}_h} \|\underline{\boldsymbol{\Pi}_{\boldsymbol{G}}\boldsymbol{\Phi}}\|_K \\
&\le \sum_{K\in\mathcal{T}_h} \|\underline{\boldsymbol{\Pi}_{\boldsymbol{G}}\boldsymbol{\Phi}} - \underline{\boldsymbol{\Phi}}\|_K + \|\underline{\boldsymbol{\Phi}}\|_\Omega \\
&\le C\big(\Sigma_{K\in\mathcal{T}_h} h_K^2 \|\underline{\boldsymbol{\Phi}}\|^2_{\underline{\boldsymbol{H}}^1(K)}\big)^{1/2} + \|\underline{\boldsymbol{\Phi}}\|_\Omega \quad \text{by (6.6c)} \\
&\le C\big(\Sigma_{K\in\mathcal{T}_h} h_K^2 \|\underline{\boldsymbol{w}}\|_{\boldsymbol{H}^2(K)}\big)^{1/2} + \|\underline{\boldsymbol{w}}\|_{\boldsymbol{H}^1(\Omega)} \\
&\le C\|\underline{\boldsymbol{w}}\|_{\boldsymbol{H}^1(\Omega)} \\
&\le C\|\underline{\boldsymbol{\eta}}\|_\Omega,
\end{aligned}
$$

by (6.16b). This shows that (6.14c) holds. This completes the proof. $\square$

Step 5: Final Estimates. We are now ready to complete the proof of Theorem 6.1.1. We first insert the estimate (6.15) into Corollary 6.2.3 to get:

$$
\begin{aligned}
\|\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\|_\Omega^2 &\le C(\|\underline{\boldsymbol{\delta}_{\boldsymbol{\sigma}}}\|_\Omega(\|\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\|_\Omega + \|\underline{\boldsymbol{\delta}_{\boldsymbol{\sigma}}}\|_\Omega + \|\underline{\boldsymbol{\delta}_{\boldsymbol{\rho}}}\|_\Omega) + \|\underline{\boldsymbol{\delta}_{\boldsymbol{\sigma}}}\|_\Omega^2 + \|\underline{\boldsymbol{\delta}_{\boldsymbol{\rho}}}\|_\Omega^2) \\
&\le \frac{1}{2}\|\underline{\boldsymbol{e}_{\boldsymbol{\sigma}}}\|_\Omega^2 + C(\|\underline{\boldsymbol{\delta}_{\boldsymbol{\sigma}}}\|_\Omega^2 + \|\underline{\boldsymbol{\delta}_{\boldsymbol{\rho}}}\|_\Omega^2),
\end{aligned}
$$

by Young's inequality. This proves (6.9a).

Now (6.9b) follows if we combine (6.9a) and (6.15). Finally, the estimate (6.10) can be obtained by inserting (6.9a), (6.9b) into Corollary 6.2.5. This completes the proof of Theorem 6.1.1.

References

[1] F. Celiker, B. Cockburn, and K. Shi. Hybridizable discontinuous Galerkin methods for Timoshenko beams. *J. Sci. Comput.*, 44:1–37, 2010.

[2] F. Celiker, B. Cockburn, and K. Shi. A projection-based error analysis of HDG methods for Timoshenko beams. *Math. Comp.*, 81:131–151, 2012.

[3] W. Qiu B. Cockburn and K. Shi. Conditions for superconvergence of HDG methods for second-order elliptic problems. *Math. Comp.*, in press.

[4] W. Qiu B. Cockburn and K. Shi. Superconvergent HDG methods on isoparametric elements for second-order elliptic problems. *SIAM J. Numer. Anal.*, in press.

[5] D. N. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini. Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM J. Numer. Anal.*, 39:1749–1779, 2002.

[6] B. Cockburn, J. Gopalakrishnan, and R. Lazarov. Unified hybridization of discontinuous Galerkin, mixed and continuous Galerkin methods for second order elliptic problems. *SIAM J. Numer. Anal.*, 47:1319–1365, 2009.

[7] B. Cockburn, B. Dong, and J. Guzmán. A superconvergent LDG-hybridizable Galerkin method for second-order elliptic problems. *Math. Comp.*, 77:1887–1916, 2008.

[8] F. Brezzi and M. Fortin. *Mixed and Hybrid finite element methods.* Springer Verlag, 1991.

[9] F. Celiker, B. Cockburn, S. Güzey, R. Kannapady, S.-C. Soon, H. K. Stolarski, and K. K. Tamma. Discontinuous Galerkin methods for Timoshenko beams. In *Numerical Mathematics and Advanced Applications, ENUMATH 2003*, pages 221–231. Springer, 2003.

[10] F. Celiker, B. Cockburn, and H. Stolarski. Locking-free optimal discontinuous Galerkin methods for Timoshenko beams. *SIAM J. Numer. Anal.*, 44:2297–2325, 2006.

[11] B. Cockburn, J. Gopalakrishnan, and F.-J. Sayas. A projection-based error analysis of HDG methods. *Math. Comp.*, 79:1351–1367, 2010.

[12] P. A. Raviart and J. M. Thomas. A mixed finite element method for second order elliptic problems. In I. Galligani and E. Magenes, editors, *Mathematical Aspects of Finite Element Method, Lecture Notes in Math. 606*, pages 292–315. Springer-Verlag, New York, 1977.

[13] J. Douglas, Jr. and J.E. Roberts. Global estimates for mixed methods for second order elliptic equations. *Math. Comp.*, 44:39–52, 1985.

[14] D. N. Arnold and F. Brezzi. Mixed and nonconforming finite element methods: implementation, postprocessing and error estimates. *RAIRO Modél. Math. Anal. Numér.*, 19:7–32, 1985.

[15] F. Brezzi, J. Douglas, Jr., and L. D. Marini. Two families of mixed finite elements for second order elliptic problems. *Numer. Math.*, 47:217–235, 1985.

[16] F. Brezzi, J. Douglas, Jr., M. Fortin, and L. D. Marini. Efficient rectangular mixed finite element methods in two and three space variables. *RAIRO Modél. Math. Anal. Numér.*, 21:581–604, 1987.

[17] R. Stenberg. A family of mixed finite elements for the elasticity problem. *Numer. Math.*, 53:513–538, 1988.

[18] L. Gastaldi and R.H. Nochetto. Sharp maximum norm error estimates for general mixed finite element approximations to second order elliptic equations. *RAIRO Modél. Math. Anal. Numér.*, 23:103–128, 1989.

[19] R. Stenberg. Postprocessing schemes for some mixed finite elements. *RAIRO Modél. Math. Anal. Numér.*, 25:151–167, 1991.

[20] Y. Chen and B. Cockburn. Analysis of variable-degree HDG methods for convection-diffusion equations. Part i: General nonconforming meshes. *IMA J. Num. Anal.*, to appear.

[21] Y. Chen and B. Cockburn. Analysis of variable-degree HDG methods for convection-diffusion equations. Part ii: Semimatching nonconforming meshes. *Math. Comp.*, in revision.

[22] B. Chabaud and B. Cockburn. Uniform-in-time superconvergence of HDG methods for the heat equation. *Math. Comp.*, 81:107–129, 2012.

[23] B. Cockburn and K. Shi. Conditions for superconvergence of HDG methods for Stokes flow. *Math. Comp.*, in press.

[24] B. Cockburn and K. Shi. The devising of superconvergent HDG methods for linear elasticity with weakly symmetric stresses. submitted.

[25] B. Cockburn and V. Quenneville-Bélair. Uniform-in-time superconvergence of HDG methods for the acoustic wave equation. *Math. Comp.*

[26] B. Cockburn, J. Gopalakrishnan, N.C. Nguyen, J. Peraire, and F.J. Sayas. Analysis of an HDG method for Stokes flow. *Math. Comp.* to appear.

[27] N.C. Nguyen, J. Peraire, and B. Cockburn. A hybridizable discontinuous Galerkin method for Stokes flow. *Comput. Methods Appl. Mech. Engrg.*, 199:582–597, 2010.

[28] M. Farhloul and M. Fortin. A new mixed finite element for the Stokes and elasticity problems. *SIAM J. Numer. Anal.*, 30:971–990, 1993.

[29] M. Farhloul and M. Fortin. A mixed finite element for the Stokes problem using quadrilateral elements. *Adv. Comput. Math.*, 3:101–113, 1995.

[30] N.C. Nguyen and J. Peraire. Hybridizable discontinuous Galerkin methods for partial differential equations in continuum mechanics. submitted.

[31] J. Gopalakrishnan B. Cockburn and J. Guzmán. A new elasticity element made for enforcing weak stress symmetry. *Math. Comp.*, 79:1331–1349, 2010.

[32] J. Gopalakrishnan and J. Guzmán. A second elasticity element using the matrix bubble. *IMA J. Numer. Anal.*, to appear.

[33] J. Guzmán. A unified analysis of several mixed methods for elasticity with weak stress symmetry. *J. Sci. Comput.*, 44:156–169, 2010.

[34] F. Brezzi D. N. Arnold and Jr. J. Douglas. Peers: A new mixed finite element for plane elasticity. *Japan J.Appl.Math.*, 1:347–367, 1984.

[35] R. Falk D.N. Arnold and R. Winther. Mixed finite element methods for linear elasticity with weakly imposed symmetry. *Math. Comp.*, 26:1699–1723, 2007.

[36] M. Fortin D. Boffi, F. Brezzi. Reduced symmetry elements in linear elasticity. *Commun. Pure Appl.*, 8:95–121, 2009.

[37] S. P. Timoshenko. On the correction for shear of the differential equation for transverse vibrations of prismatic bar. *Philosophical Magazine*, 41:744–746, 1921.

[38] F. Celiker. *Discontinuous Galerkin methods for Structural Mechanics.* PhD thesis, University of Minnesota, 2005.

[39] F. Celiker and B. Cockburn. Element-by-element post-processing of discontinuous Galerkin methods for Timoshenko beams. *J. Sci. Comput.*, 27:177–187, 2006.

[40] B. Dong B. Cockburn and J. Guzmán. A hybridizable and superconvergent discontinuous Galerkin method for biharmonic problems. *J. Sci. Comput.*, 40:141–187, 2009.

[41] R.M. Kirby, S.J. Sherwin, and B. Cockburn. To HDG or to CG: A comparative study. submitted.

[42] F. Bassi, S. Rebay, G. Mariotti, S. Pedinotti, and M. Savini. A high-order accurate discontinuous finite element method for inviscid and viscous turbomachinery flows. In R. Decuypere and G. Dibelius, editors, *2nd European Conference on Turbomachinery Fluid Dynamics and Thermodynamics*, pages 99–108, Antwerpen, Belgium, March 5–7 1997. Technologisch Instituut.

[43] F. Brezzi, G. Manzini, L. D. Marini, P. Pietra, and A. Russo. Discontinuous finite elements for diffusion problems. In *in Atti Convegno in onore di F. Brioschi (Milan 1997)*, pages 197–217. Istituto Lombardo, Accademia di Scienze e Lettere, 1999.

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

[45] L. F. Pavarino T. Warburton and J. S. Hesthaven. A pseudo-spectral scheme for the incompressible Navier-Stokes equations using unstructured nodal elements. *J. Comput. Phys.*, 164:1–21, 2000.

[46] S.J. Sherwin P.E.J. Vos and R.M. Kirby. From h to p efficiency: Implementing finite and spectral/hp element methods to achieve optimal performance for low- and high-order discretizations. *J. Comput. Phys.*, 229:5161–5181, 2010.

[47] R.M. Kirby C.D. Cantwell, S.J. Sherwin and P.H.J. Kelly. From h to p efficiency: Strategy selection for operator evaluation on hexahedral and tetrahedral elements. submitted.

[48] Z. Chen and J. Douglas, Jr. Prismatic mixed finite elements for second order elliptic problems. *Calcolo*, 26:135–148, 1989.

[49] R. B. Kellogg and J. E. Osborn. A regularity result for the Stokes problem in a convex polygon. *J. Functional Analysis*, 21(4):397–431, 1976.

[50] M. Dauge. Stationary Stokes and Navier-Stokes systems on two- or three-dimensional domains with corners. I. Linearized equations. *SIAM J. Math. Anal.*, 20(1):74–97, 1989.

[51] B. Cockburn and F.J. Sayas. Divergence–conforming HDG methods for Stokes flow. submitted.

[52] C. Bacuta and J. H. Bramble. Regularity estimates for solutions of the equations of linear elasticity in convex plane polygonal domains. *Z. Angew. Math. Phys.*, 54:874–878, 2003.

[53] W. Qiu B. Cockburn. Commuting diagrams for TNT elements. submitted.

[54] Ph. Clément. Approximation by finite element functions using local regularization. *RAIRO Modél. Math. Anal. Numér.*, 9:77–84, 1975.

[55] D. Braess. *Finite Elements: Theory, Fast Solvers, and Applications in Solid Mechanics*. Cambridge University Press, 2007.