2011

Analysis of HDG Methods for Stokes Flow

Bernardo Cockburn
University of Minnesota - Twin Cities

Jay Gopalakrishnan
Portland State University, gjay@pdx.edu

Ngoc Cuong Nguyen
Massachusetts Institute of Technology

Jaume Peraire
Massachusetts Institute of Technology

Francisco-Javier Sayas
Departamento de Matemática Aplicada

Citation Details

Cockburn, Bernardo et al. "Analysis of HDG methods for Stokes flow." Mathematics of Computation 80.274 (2011): 723-723.

ANALYSIS OF HDG METHODS FOR STOKES FLOW

BERNARDO COCKBURN, JAYADEEP GOPALAKRISHNAN, NGOC CUONG NGUYEN, JAUME PERAIRE, AND FRANCISCO-JAVIER SAYAS

Abstract. In this paper, we analyze a hybridizable discontinuous Galerkin method for numerically solving the Stokes equations. The method uses polynomials of degree k for all the components of the approximate solution of the gradient-velocity-pressure formulation. The novelty of the analysis is the use of a new projection tailored to the very structure of the numerical traces of the method. It renders the analysis of the projection of the errors very concise and allows us to see that the projection of the error in the velocity superconverges. As a consequence, we prove that the approximations of the velocity gradient, the velocity and the pressure converge with the optimal order of convergence of $k+1$ in L^2 for any $k \geq 0$. Moreover, taking advantage of the superconvergence properties of the velocity, we introduce a new element-by-element postprocessing to obtain a new velocity approximation which is exactly divergence-free, $\mathbf{H}(\mathrm{div})$-conforming, and converges with order $k+2$ for $k \geq 1$ and with order 1 for $k = 0$. Numerical experiments are presented which validate the theoretical results.

1. Introduction

In this paper, we carry out an *a priori* error analysis of hybridizable discontinuous Galerkin (HDG) methods proposed in [28] to numerically solve the Stokes equations of incompressible fluid flow, namely,

$$\mathrm{L} - \nabla \boldsymbol{u} = 0 \qquad \text{on } \Omega, \tag{1.1a}$$

$$-\boldsymbol{\nabla}\cdot(\nu \mathrm{L}) + \boldsymbol{\nabla} p = \boldsymbol{f} \qquad \text{on } \Omega, \tag{1.1b}$$

$$\nabla\cdot \boldsymbol{u} = 0 \qquad \text{on } \Omega, \tag{1.1c}$$

$$\boldsymbol{u} = \boldsymbol{g} \qquad \text{on } \partial\Omega, \tag{1.1d}$$

$$\int_\Omega p = 0, \tag{1.1e}$$

Received by the editor July 29, 2009 and, in revised form, January 5, 2010.

2010 *Mathematics Subject Classification.* Primary 65N30, 65M60, 35L65.

Key words and phrases. Stokes flow, mixed methods, discontinuous Galerkin methods, hybridized methods, Lagrange multipliers.

The first author was supported in part by the National Science Foundation (Grant DMS-0712955) and by the University of Minnesota Supercomputing Institute.

The second author was supported in part by the National Science Foundation under grants DMS-0713833 and SCREMS-0619080.

The third author was supported in part by the Singapore-MIT Alliance.

The fourth author was supported in part by the Singapore-MIT Alliance.

The fifth author was a Visiting Professor of the School of Mathematics, University of Minnesota, during the development of this work. He was partially supported by MEC/FEDER Project MTM2007–63204 and Gobierno de Aragón (Grupo PDIE).

where $\int_{\partial\Omega} \boldsymbol{g}\cdot\boldsymbol{n} = 0$. Here $\Omega \subset \mathbb{R}^n$ is a polygonal domain if $n = 2$ or a Lipschitz polyhedral domain if $n = 3$.

To better describe our results, let us put them in historical perspective. The HDG methods were introduced in [13] in the framework of diffusion problems as a further development of the hybridization of Raviart-Thomas (RT) [29] and Brezzi-Douglas-Marini (BDM) [3] carried out in [8]. They were also introduced as a response to the criticism that the discontinuous Galerkin (DG) methods for elliptic problems (see the unified analysis of these methods in [1] and a comparison of their performance in [6]) has too many globally coupled degrees of freedom; see [30]. Soon thereafter, it was discovered [7, 15] that the approximate flux of the HDG methods converges with order $k + 1$ for $k \geq 0$ and that the HDG methods share with the RT and BDM mixed methods superconvergence properties which allow an element-by-element computation of a *new* approximation of the scalar variable which converged with order $k+2$ for $k \geq 1$; see also the new analysis carried out in [14]. This has to be contrasted with the fact that *all* the other DG methods display the suboptimal order of convergence of k. Furthermore, most of them do not even converge for $k = 0$.

Hybridization for DG methods for Stokes was initially introduced [5] as a technique that allowed the use of globally divergence-free velocity spaces without having to actually carry out their almost-impossible construction. The technique was then further developed, with a similar intention, in the framework of mixed methods in [9, 10]. Indeed, a novel, global formulation for the method was obtained solely in terms of the tangential velocity and the pressure on the *borders of the elements*. As a further development of this approach, the first HDG methods for the Stokes equations were recently introduced in [11]. Remarkably enough, these methods were shown to be hybridizable in *four* completely different ways including a tangential-velocity/pressure formulation and a velocity/average-pressure formulation involving degrees of freedom on the borders of the elements *only*. All of the above methods used vorticity-velocity or vorticity-velocity-pressure formulations. In [28], this approach was used to devise an HDG method based on a velocity gradient-velocity-pressure formulation. This is the method we analyze in this paper.

Let us contrast this method with three of the finite element methods also based on velocity gradient-velocity-pressure formulations. First, consider the methods devised in [31] by using the RT and BDM elements developed for diffusion problems. Therein, optimal orders of convergence where obtained for each of the variables. Thus, for the method associated to BDM, the $\boldsymbol{H}(\mathrm{div})$-conforming velocity gradient and the continuous pressure were shown to converge with the optimal order of $k+1$ whereas the completely discontinuous velocity was shown to converge with the optimal order of k for $k \geq 1$. Moreover, a local postprocessing was introduced which provides a new, discontinuous approximate velocity converging with order $k+2$ for $k \geq 2$. In contrast, the HDG method presented here can be easily hybridized, and hence efficiently implemented [28], since its pressure is not continuous. It is also defined for the $k = 0$ case in which all its variables converge optimally. Also, its postprocessed velocity is $\boldsymbol{H}(\mathrm{div})$-conforming and divergence-free, and converges with order $k + 2$ for $k \geq 1$, and with order 1 for $k = 0$.

The second method we would like to compare our HDG method with is the local discontinuous Galerkin (LDG) introduced in [18] as a stepping stone towards DG methods for the Navier-Stokes equations; see [16, 17]. The method uses the same spaces and the same weak formulation as our HDG method. In fact, in spite of having been devised in completely different ways, the *only* difference between these HDG and LDG methods lies on the definition of their *numerical traces*; see the discussion in [28]; yet, this apparently minor modification produces an improvement in the way the method is implemented, renders the order of convergence of the velocity gradient and the pressure optimal, and allows for a postprocessed velocity converging with an additional order. Again, this is not quite a surprise since something similar can be said about the HDG and LDG methods for diffusion problems; see [15].

Third, the mixed finite element method developed in [21] is based on spaces where gradient and pressure satisfy a joint compatibility condition on the interfaces. The spaces are those of Raviart-Thomas finite elements for the columns of the gradient and polynomials for pressure and velocity. Because of the constraints in the interfaces the method in [21] can be implemented only after addition of Lagrange multipliers on the faces of the elements. The resulting method is hybridizable and is very similar to ours, with the difference that we obtain stability by adding the stabilization term in the interfaces instead of by adding the Raviart-Thomas degrees of freedom to the approximate gradient. We can thus use smaller local spaces. The method in [21] and the one in this paper provide the same orders of convergence for the variables and the multipliers. However, we derive a superconvergent postprocessing of the velocity by taking full advantage of those convergence properties.

Finally, let us briefly point out that the main tool in our analysis is a new projection fitting the structure of the numerical traces of the HDG method. This projection is an extension to our setting of the projection introduced for the analysis of HDG methods for diffusion problems in [14]. The use of this new tool has two important advantages. The first is that it reduces the study of the effect of the stabilization parameters to an approximation error in a single element. The second is that it renders the analysis of the projection of the errors extremely concise and sharp and allows us to easily obtain superconvergence estimates for the projection of the error in the velocity.

The paper is organized as follows. In Section 2, we describe the HDG method and state and discuss the error estimates for the approximation and for the postprocessing. Then, in Section 3, we provide detailed proofs of these results. In Section 4, we provide numerical experiments to validate the theoretical results. We end in Section 5 with some concluding remarks.

2. Main results

In this section, we state and discuss our main results. We begin by describing the HDG method, discussing three main examples and by introducing a new postprocessing of the velocity. We then introduce our main tool of analysis, namely, the above-mentioned projection, and state and discuss its main properties. The a priori error estimates on the projection of the errors of the approximations and on the postprocessed velocity are then stated and discussed.

2.1. **The HDG method.**

The weak formulation. As it is customary for finite element methods, we begin by discretizing Ω by a *shape-regular* triangulation $\mathcal{T}_h$ which, for the sake of simplicity, we take to be made of simplexes K and to be free of hanging nodes. The changes needed to deal with more general meshes are not difficult at the algorithmic level, but the presence of hanging nodes might complicate the analysis in unexpected ways. To this triangulation, we associate a finite dimensional space $\mathrm{G}_h \times \boldsymbol{V}_h \times P_h \times \boldsymbol{M}_h$ in which we seek the approximation $(\mathrm{L}_h, \boldsymbol{u}_h, p_h, \widehat{\boldsymbol{u}}_h)$ of the exact solution $(\mathrm{L}, \boldsymbol{u}, p, \boldsymbol{u}|_{\mathcal{E}_h})$ where $\mathcal{E}_h$ denotes the set of all faces F of all simplexes K of the triangulation $\mathcal{T}_h$. The space is given by

$$\mathrm{G}_h = \{\mathrm{G} \in \mathrm{L}^2(\mathcal{T}_h) : \quad \mathrm{G}|_K \in \mathrm{P}_k(K) \quad \forall\, K \in \mathcal{T}_h\}, \tag{2.1a}$$

$$\boldsymbol{V}_h = \{\boldsymbol{v} \in \boldsymbol{L}^2(\mathcal{T}_h) : \quad \boldsymbol{v}|_K \in \boldsymbol{\mathcal{P}}_k(K) \quad \forall\, K \in \mathcal{T}_h\}, \tag{2.1b}$$

$$P_h = \{q \in L^2(\mathcal{T}_h) : \quad q|_K \in \mathcal{P}_k(K) \quad \forall\, K \in \mathcal{T}_h\}, \tag{2.1c}$$

$$\boldsymbol{M}_h = \{\boldsymbol{\mu} \in L^2(\mathcal{E}_h) : \quad \boldsymbol{\mu}|_F \in \boldsymbol{\mathcal{P}}_k(F) \quad \forall\, F \in \mathcal{E}_h\}. \tag{2.1d}$$

Here $\mathcal{P}_\ell(D)$ is the space of polynomials of total degree at most ℓ defined on the domain D, $\boldsymbol{\mathcal{P}}_\ell(D) = [\mathcal{P}_\ell(D)]^n$ and $\mathrm{P}_\ell(D) = [\mathcal{P}_\ell(D)]^{n\times n}$.

Next, we describe the weak formulation determining the HDG approximation. To do this, we need to introduce some notation. 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

$$(\boldsymbol{\eta}, \boldsymbol{\zeta})_{\mathcal{T}_h} := \sum_{i=1}^{n} (\eta_i, \zeta_i)_{\mathcal{T}_h} \quad \text{and} \quad (\mathrm{N}, \mathrm{Z})_{\mathcal{T}_h} := \sum_{i,j=1}^{n} (\mathrm{N}_{ij}, \mathrm{Z}_{ij})_{\mathcal{T}_h}.$$

Finally, we write

$$\langle \eta, \zeta \rangle_{\mathcal{T}_h} := \sum_{K\in\mathcal{T}_h} \langle \eta, \zeta \rangle_{\partial K} \quad \text{and} \quad \langle \boldsymbol{\eta}, \boldsymbol{\zeta} \rangle_{\partial\mathcal{T}_h} := \sum_{i=1}^{n} \langle \eta_i, \zeta_i \rangle_{\partial\mathcal{T}_h},$$

where $\partial\mathcal{T}_h := \{\partial K : K \in \mathcal{T}_h\}$ and $\langle \eta, \zeta \rangle_D$ denotes the integral of $\eta\,\zeta$ over the domain $D \subset \mathbb{R}^{n-1}$.

We are now ready to display the equations satisfied by the HDG approximation. They are the following:

$$(\mathrm{L}_h, \mathrm{G})_{\mathcal{T}_h} + (\boldsymbol{u}_h, \boldsymbol{\nabla\cdot}\,\mathrm{G})_{\mathcal{T}_h} - \langle \widehat{\boldsymbol{u}}_h, \mathrm{G}\boldsymbol{n} \rangle_{\partial\mathcal{T}_h} = 0, \tag{2.2a}$$

$$(\nu \mathrm{L}_h, \nabla\, \boldsymbol{v})_{\mathcal{T}_h} - (p_h, \nabla\cdot \boldsymbol{v})_{\mathcal{T}_h} - \langle \nu \widehat{\mathrm{L}}_h \boldsymbol{n} - \widehat{p}_h \boldsymbol{n}, \boldsymbol{v} \rangle_{\partial\mathcal{T}_h} = (\boldsymbol{f}, \boldsymbol{v})_{\mathcal{T}_h}, \tag{2.2b}$$

$$-(\boldsymbol{u}_h, \boldsymbol{\nabla}\, q)_{\mathcal{T}_h} + \langle \widehat{\boldsymbol{u}}_h \cdot \boldsymbol{n}, q \rangle_{\partial\mathcal{T}_h} = 0, \tag{2.2c}$$

$$\langle \widehat{\boldsymbol{u}}_h, \boldsymbol{\mu} \rangle_{\partial\Omega} = \langle \boldsymbol{g}, \boldsymbol{\mu} \rangle_{\partial\Omega}, \tag{2.2d}$$

$$\langle \nu \widehat{\mathrm{L}}_h \boldsymbol{n} - \widehat{p}_h \boldsymbol{n}, \boldsymbol{\mu} \rangle_{\partial\mathcal{T}_h \setminus \partial\Omega} = 0, \tag{2.2e}$$

$$(p_h, 1)_\Omega = 0, \tag{2.2f}$$

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

$$\nu \widehat{\mathrm{L}}_h \boldsymbol{n} - \widehat{p}_h \boldsymbol{n} = \nu \mathrm{L}_h \boldsymbol{n} - p_h \boldsymbol{n} - \mathrm{S}\,(\boldsymbol{u}_h - \widehat{\boldsymbol{u}}_h) \qquad \text{on } \partial\mathcal{T}_h. \tag{2.2g}$$

General conditions on S. To complete the definition of the HDG method, it only remains to describe how to choose the *stabilization tensor* S. In [28], it was shown that if S is any definite positive tensor-valued function on $\partial\mathcal{T}_h$, the HDG method is well defined. Here, we take the stabilization tensor as follows. For each simplex K of the triangulation $\mathcal{T}_h$, we take the stabilization tensor satisfying the following four conditions:

$$\mathrm{S}|_{\partial K} \text{ is constant on each face } F \text{ of } K, \tag{2.3a}$$

$$\mathrm{S}|_{\partial K} \text{ is symmetric}, \tag{2.3b}$$

$$\mathrm{S}|_{\partial K} \text{ is positive semidefinite}. \tag{2.3c}$$

To state the last condition, we need to introduce the orthonormal basis of eigenvectors of $\mathrm{S}|_F$ for each face F of K, namely,

$$\mathrm{S}|_F\boldsymbol{\omega}_{F,i} = \lambda_{F,i}\boldsymbol{\omega}_{F,i}, \qquad \boldsymbol{\omega}_{F,i}\cdot\boldsymbol{\omega}_{F,j} = \delta_{ij}, \qquad i,j = 1,\ldots,n.$$

The last condition can be stated as follows: there exists an index set $\mathcal{I}_K \subset \{(F,i) : F \text{ is a face of } K, i = 1,\ldots,n\}$ such that

$$\mathcal{B}_K := \{\boldsymbol{\omega}_{F,i} : (F,i)\in\mathcal{I}_K\} \text{ is a basis of } \mathbb{R}^n, \tag{2.4a}$$

$$\lambda_K^{\min} := \min_{(F,i)\in\mathcal{I}_K} \lambda_{F,i} > 0, \tag{2.4b}$$

$$C_K := \max_{(F,i)\in\mathcal{I}_K} |\boldsymbol{\omega}^*_{F,i}| \le C, \tag{2.4c}$$

where $|\cdot|$ is the Euclidean norm in $\mathbb{R}^n$ and $\boldsymbol{\omega}^*_{F,i}$ is the dual basis of $\mathcal{B}_K$, that is, the basis of $\mathbb{R}^n$ (indexed in $\mathcal{I}_K$) such that $\boldsymbol{\omega}^*_{F,i}\cdot\boldsymbol{\omega}_{F',j} = \delta_{F,F'}\delta_{i,j}$ for $(F,i),(F',j)\in\mathcal{I}_K$. The constant C in (2.4c) is assumed to be independent of the mesh and of the eigenvalues $\lambda_{F,i}$. As we will see in the prototypical examples below, C_K is due to depend on the particular choice of the stabilization tensor S and on the shape regularity of the mesh. The only dependence of the error estimates on the stabilization tensor S will then appear through only two quantities, namely, $\lambda_K^{\min}$ and

$$\Lambda_K^{\max} := \max_F \max_i \lambda_{F,i}. \tag{2.5}$$

Let us briefly comment on these conditions. Conditions (2.3a) and (2.3b) can be easily relaxed, but we prefer to keep them to avoid unessential technicalities in the error analysis of the method. On the other hand, the third and fourth conditions are essential to guarantee that the HDG method is well defined.

Note that the first condition allows us to rewrite the HDG method in a more traditional way, that is, as a method given by the first five equations in (2.2) where *all* the numerical traces are expressed in terms of the approximation $(\mathrm{L}_h,\boldsymbol{u}_h,p_h)$. Indeed, it is not difficult to conclude by using the equations (2.2e) and (2.2g) (see the details in [28]), that on the face $F = \partial K^+\cap\partial K^-$, where $K^\pm\in\mathcal{T}_h$, we have

$$\widehat{\boldsymbol{u}}_h = \mathrm{AS}^+\boldsymbol{u}_h^+ + \mathrm{AS}^-\boldsymbol{u}_h^- - \mathrm{A}\,[\![(\nu\mathrm{L}_h - p_h\mathrm{I})\boldsymbol{n}]\!],$$
$$\nu\widehat{\mathrm{L}}_h - \widehat{p}_h\mathrm{I} = \mathrm{S}^-\mathrm{A}(\nu\mathrm{L}_h^+ - p_h^+\mathrm{I}) + \mathrm{S}^+\mathrm{A}(\nu\mathrm{L}_h^- - p_h^-\mathrm{I}) - \mathrm{S}^-\mathrm{AS}^+\,[\![\boldsymbol{u}_h\otimes\boldsymbol{n}]\!],$$

where $\mathrm{A} := (\mathrm{S}^-+\mathrm{S}^+)^{-1}$, $[\![\mathrm{G}\boldsymbol{n}]\!] := \mathrm{G}^+\boldsymbol{n}^+ + \mathrm{G}^-\boldsymbol{n}^-$ and $[\![\boldsymbol{v}\otimes\boldsymbol{n}]\!] := \boldsymbol{v}^+\otimes\boldsymbol{n}^+ + \boldsymbol{v}^-\otimes\boldsymbol{n}^-$. As usual, $(\mathrm{L}_h^\pm,\boldsymbol{u}_h^\pm,p_h^\pm)$ is the trace on F from the inside of the simplex $K^\pm$, $\mathrm{S}^\pm$ the value on $\partial K^\pm$, and $\boldsymbol{n}^\pm$ the unit outward normal to $K^\pm$.

For example, for the choice taken in the numerical experiments in [28], namely, for the case $\mathrm{S} := \nu\tau\,\mathrm{I}$, where τ is a constant on $\partial\mathcal{T}_h$, the numerical traces become

$$\widehat{\boldsymbol{u}}_h = \frac{1}{2}\boldsymbol{u}_h^+ + \frac{1}{2}\boldsymbol{u}_h^- - \frac{1}{2\nu\tau}\,[\![(\nu\mathrm{L}_h - p_h\mathrm{I})\boldsymbol{n}]\!],$$
$$\nu\widehat{\mathrm{L}}_h - \widehat{p}_h\mathrm{I} = \frac{1}{2}(\nu\mathrm{L}_h^+ - p_h^+\mathrm{I}) + \frac{1}{2}(\nu\mathrm{L}_h^- - p_h^-\mathrm{I}) - \frac{\nu\tau}{2}\,[\![\boldsymbol{u}_h \otimes \boldsymbol{n}]\!].$$

We thus see that the only difference between this method and the LDG method proposed in [18] is the definition of the numerical traces. In particular, the numerical traces of the velocity are independent of the velocity gradient.

Two special choices of S**.** In this paper, we are going to pay special attention to two special choices of the stabilization tensor. The first defines what we call *single face hybridizable* (SFH) method, in analogy to the method of the same name introduced and analyzed in [7] in the framework of diffusion problems. For each $K \in \mathcal{T}_h$, the stabilization tensor is given by

$$\mathrm{S}_{SF} := \begin{cases} \nu\tau\mathrm{I} & \text{on } F_K^\star, \\ 0 & \text{on } \partial K \setminus F_K^\star, \end{cases} \tag{2.6}$$

where $F_K^\star$ is an arbitrary face of K. Note that if τ is a strictly positive constant on $F_K^\star$, the three conditions (2.3) are satisfied. For (2.4) we choose as an index set the set $\mathcal{I}_K := \{(F_K^\star, i) : i = 1, \ldots, n\}$, that is, we select all eigenvectors associated to the face $F_K^\star$. Therefore, we can tag the index set $\mathcal{I}_K$ in any way to obtain as a local basis $\mathcal{B}_K$, the canonical basis of $\mathbb{R}^n$, which is its own dual basis. With this choice $\lambda_K^{\min} = \Lambda_K^{\max} = \nu\tau$ and the bound for the dual basis (2.4c) is $C_K = 1$.

The second choice is an extension of the stabilization tensor used in [28] which allows us to control in different ways the normal and tangential components of the interelement jumps of the approximate velocity. For each $K \in \mathcal{T}_h$, it is given by

$$\mathrm{S}_{nt} := \nu\tau_n\,\boldsymbol{n}\otimes\boldsymbol{n} + \nu\tau_t(\mathrm{I} - \boldsymbol{n}\otimes\boldsymbol{n}) \quad \text{on } \partial K. \tag{2.7}$$

Note that the conditions (2.3) are satisfied if on ∂K both τ_n and τ_t are are non-negative constants. Eigenvectors for $\mathrm{S}|_F$ are the normal vectors to F (the eigenvalue is $\nu\tau_n$) and every tangential vector to F (the eigenvalue being $\nu\tau_t$). Therefore $\Lambda_K^{\max} = \nu\max\{\tau_n, \tau_t\}$. If this maximum is attained with τ_n, we choose the index set $\mathcal{I}_K$ in a way that we are selecting n of the normal vectors to the faces of K. By numbering the faces of K as $\{F_i : i = 1, \ldots, n+1\}$ we can choose $\mathcal{B}_K = \{\boldsymbol{n}_{F_i} : i = 1, \ldots, n\}$. The dual basis is obtained by taking the rows of the inverse of the matrix whose columns are the elements of the basis $\mathcal{B}_K$. If the maximum is attained with τ_t, we choose one vertex and as a basis we pick the unit vectors along the edges stemming from this vertex. Its dual basis is obtained as before. In both cases C_K in (2.4c) is bounded depending only on shape regularity constants (this can be proved with basic estimates on inverses of matrices depending on the size and angles of its columns). Moreover, (2.4) is satisfied with $\lambda_K^{\min} = \Lambda_K^{\max}$.

The divergence-free condition. Next, let us point out that, for *any* of the HDG methods, we have

$$\operatorname{tr}\,\mathrm{L}_h = 0, \tag{2.8}$$

where $\operatorname{tr}\,\mathrm{A}$ denotes the trace of the matrix A. This property is a reflection of the divergence-free condition on the exact velocity $\boldsymbol{u}$ since $\operatorname{tr}\,\mathrm{L} = \nabla\cdot\boldsymbol{u} = 0$. Indeed, if in the first equation defining the method (2.2a), we take $\mathrm{G} := q\,\mathrm{I}$ where $q \in \mathcal{P}_k(K)$, we obtain that

$$(\mathrm{tr}\,\mathrm{L}_h, q)_{\mathcal{T}_h} + (\boldsymbol{u}_h, \boldsymbol{\nabla}\, q)_{\mathcal{T}_h} - \langle \widehat{\boldsymbol{u}}_h, \boldsymbol{n} q\rangle_{\partial\mathcal{T}_h} = 0.$$

We then see that $(\mathrm{tr}\,\mathrm{L}_h, q)_{\mathcal{T}_h} = 0$, by the weak divergence-free condition, equation (2.2c). This implies that $\mathrm{tr}\,\mathrm{L}_h = 0$, as claimed.

The postprocessed approximate velocity. To end this section, we show how to postprocess the approximate solution in an element-by-element way to obtain a new approximation which is exactly divergence-free, $\boldsymbol{H}(\mathrm{div})$-conforming, and converges with an additional order for $k \geq 1$. To do this, we use a modification of a *new* characterization (see the Appendix) of the Brezzi–Douglas–Marini (BDM) projection [4] (see also [3, 27]). We will first explain the more complicated three-dimensional case and then point out the modifications for the two-dimensional case.

In the *three-dimensional case* we define the postprocessed approximate velocity $\boldsymbol{u}_h^\star$ on the tetrahedron $K \in \mathcal{T}_h$ as the element of $\boldsymbol{\mathcal{P}}_{k+1}(K)$ such that

$$\langle(\boldsymbol{u}_h^\star - \widehat{\boldsymbol{u}}_h)\cdot\boldsymbol{n}, \mu\rangle_F = 0 \quad \forall\, \mu \in \mathcal{P}_k(F), \tag{2.9a}$$

$$\langle(\boldsymbol{n}\times\boldsymbol{\nabla})(\boldsymbol{u}_h^\star\cdot\boldsymbol{n}) - \boldsymbol{n}\times(\{\!\{\mathrm{L}_h^t\}\!\}\boldsymbol{n}), (\boldsymbol{n}\times\boldsymbol{\nabla})\mu\rangle_F = 0 \quad \forall\, \mu \in \mathcal{P}_{k+1}(F)^\perp, \tag{2.9b}$$

for all faces F of K, and such that

$$(\boldsymbol{u}_h^\star - \boldsymbol{u}_h, \boldsymbol{\nabla}\, w)_K = 0 \quad \forall\, w \in \boldsymbol{\mathcal{P}}_k(K), \tag{2.9c}$$

$$(\boldsymbol{\nabla}\times \boldsymbol{u}_h^\star - \mathbf{w}_h, (\boldsymbol{\nabla}\times \boldsymbol{v})\,\mathrm{B}_K)_K = 0 \quad \forall\, \boldsymbol{v} \in \boldsymbol{\mathcal{S}}_k(K). \tag{2.9d}$$

In (2.9b),

$$\mathcal{P}_{k+1}(F)^\perp := \{\mu \in \mathcal{P}_{k+1}(F) \,:\, \langle\mu, \widetilde{\mu}\rangle_F = 0, \quad \forall\widetilde{\mu} \in \mathcal{P}_k(F)\},$$

$\mathbf{n}\times\boldsymbol{\nabla}$ is the tangential gradient rotated $\pi/2$ in the positive sense (from the point of view of the normal vector) and the function $\{\!\{\mathrm{L}_h^t\}\!\}$ is the single-valued function on $\mathcal{E}_h$ equal to $((\mathrm{L}_h^t)^+ + (\mathrm{L}_h^t)^-)/2$ on the set $\mathcal{E}_h \setminus \partial\Omega$ and equal to L_h^t on $\partial\Omega$. In (2.9d),

$$\mathbf{w}_h := (\mathrm{L}_{32}^h - \mathrm{L}_{23}^h, \mathrm{L}_{13}^h - \mathrm{L}_{31}^h, \mathrm{L}_{21}^h - \mathrm{L}_{12}^h)$$

is the approximation to the vorticity and B_K is the so-called *symmetric bubble matrix* introduced in [12], namely,

$$\mathrm{B}_K := \sum_{\ell=0}^{3} \lambda_{\ell-3}\lambda_{\ell-2}\lambda_{\ell-1}\,\boldsymbol{\nabla}\,\lambda_\ell \otimes \boldsymbol{\nabla}\,\lambda_\ell,$$

where λ_i are the barycentric coordinates associated with the tetrahedron K, the subindices being counted modulo 4. Finally, to define $\boldsymbol{\mathcal{S}}_k(K)$, recall the Nedelec space of the first kind [26], defined by $\boldsymbol{N}_k = \boldsymbol{\mathcal{P}}_{k-1}(K) \oplus \boldsymbol{S}_k$, where $\boldsymbol{S}_\ell$ is the space of vector-valued homogeneous polynomials $\boldsymbol{v}$ of degree ℓ such that $\boldsymbol{v}\cdot\boldsymbol{x} = 0$. Then, define $\boldsymbol{\mathcal{S}}_k(K) := \{\boldsymbol{p} \in \boldsymbol{N}_k : (\boldsymbol{p}, \boldsymbol{\nabla}\,\phi)_K = 0 \text{ for all } \phi \in \mathcal{P}_k(K)\}$.

Note that (2.9a) and (2.9b) determine the value of $\boldsymbol{u}_h^\star\cdot\boldsymbol{n} \in \mathcal{P}_{k+1}(F)$ on F. To see this, notice that for a fixed face F, (2.9a)–(2.9b) define a square system of equations for which it is very simple to prove uniqueness of the solution with zero on the right-hand side. Note that in the proper BDM projection, the value of the normal component in each face is defined with an $L^2(F)$ projection onto $\mathcal{P}_{k+1}(F)$ instead of this more sophisticated projection. The fact that equations (2.9) define $\boldsymbol{u}_h^\star$ follows from this argument and Proposition A.1 in the Appendix.

In the *two-dimensional case*, the postprocessing is defined by the above equations if $\boldsymbol{n}\times\boldsymbol{\nabla}$ is replaced by the tangential derivative $n_2\,\partial_1+n_1\,\partial_2$, $\boldsymbol{n}\times\boldsymbol{a}$ is replaced by $n_1a_2-n_2a_1$, if $\boldsymbol{\nabla}\times\boldsymbol{u}$ is replaced by $\boldsymbol{\nabla}\times\boldsymbol{u}:=\partial_1u_2-\partial_2u_1$, and if equation (2.9d) is replaced by

$$(\boldsymbol{\nabla}\times\boldsymbol{u}_h^\star-\mathsf{w}_h,w\,b_K)_K=0\quad\forall\,w\in\mathcal{P}_{k-1}(K),$$

where $b_K:=\lambda_0\lambda_1\lambda_2$ and $\mathsf{w}_h:=\mathrm{L}_{21}^h-\mathrm{L}_{12}^h$.

2.2. **The projection.**

Definition. As we said in the Introduction, the analysis of the HDG method we are going to carry out fully exploits the special structure of its numerical traces (see the equations (2.2e) and (2.2g)) as it uses a new projection tailored to those numerical traces. Next, we introduce such a projection.

As usual, we denote by $\|\zeta\|_{H^\ell(D)}$, the sum of the squares of the L^2-norms of all the derivatives of order ℓ of the the scalar-valued function on the domain D. We set $\boldsymbol{H}^\ell(D):=[H^\ell(D)]^n$ and $\|\boldsymbol{\zeta}\|_{\boldsymbol{H}^\ell(D)}:=\sum_{i=1}^n\|\zeta_i\|_{H^\ell(D)}$; when $\ell=0$, we simply write $\|\boldsymbol{\zeta}\|_D$ instead of $\|\boldsymbol{\zeta}\|_{\boldsymbol{H}^0(D)}$. Similarly, we set $\mathrm{H}^\ell(D):=[H^\ell(D)]^{n\times n}$ and $\|\mathrm{Z}\|_{\mathrm{H}^\ell(D)}:=\sum_{i,j=1}^n\|\mathrm{Z}_{ij}\|_{H^\ell(D)}$.

Given a function $(\mathrm{L},\boldsymbol{u},p)$ in $\mathrm{H}^1(\mathcal{T}_h)\times\boldsymbol{H}^1(\mathcal{T}_h)\times H^1(\mathcal{T}_h)$, we take its projection $\mathit{\Pi}_h(\mathrm{L},\boldsymbol{u},p):=(\Pi\,\mathrm{L},\boldsymbol{\mathit{\Pi}u},\mathit{\Pi}p)$ as the element of $\mathrm{G}_h\times\boldsymbol{V}_h\times P_h$ defined as follows. On an arbitrary element K of the triangulation $\mathcal{T}_h$, the values of the projected function on the simplex K are determined by requiring that

$$(\Pi\,\mathrm{L},\mathrm{G})_K=(\mathrm{L},\mathrm{G})_K\qquad\forall\,\mathrm{G}\in\mathrm{P}_{k-1}(K),\tag{2.10a}$$
$$(\boldsymbol{\mathit{\Pi}u},\boldsymbol{v})_K=(\boldsymbol{u},\boldsymbol{v})_K\qquad\forall\,\boldsymbol{v}\in\boldsymbol{\mathcal{P}}_{k-1}(K),\tag{2.10b}$$
$$(\mathit{\Pi}p,q)_K=(p,q)_K\qquad\forall\,q\in\mathcal{P}_{k-1}(K),\tag{2.10c}$$
$$(\mathrm{tr}\,\Pi\,\mathrm{L},q)_K=(\mathrm{tr}\,\mathrm{L},q)_K\qquad\forall\,q\in\mathcal{P}_{k}(K),\tag{2.10d}$$
$$\langle\nu\Pi\,\mathrm{L}\boldsymbol{n}-\mathit{\Pi}p\,\boldsymbol{n}-\mathrm{S}\boldsymbol{\mathit{\Pi}u},\boldsymbol{\mu}\rangle_F=\langle\nu\mathrm{L}\boldsymbol{n}-p\,\boldsymbol{n}-\mathrm{S}\boldsymbol{u},\boldsymbol{\mu}\rangle_F\qquad\forall\,\boldsymbol{\mu}\in\boldsymbol{\mathcal{P}}_k(F),\tag{2.10e}$$

for all faces F of the simplex K.

Note that the projection depends on the viscosity coefficient ν and on the stabilization tensor S. Note also that the form of the equation (2.10e) is motivated by the expression of the numerical trace given in (2.2g). To see this, it is enough to put together the numerical trace given by (2.2g) and a suitable rewriting of the equation (2.10e):

$$\nu\widehat{\mathrm{L}}_h\boldsymbol{n}-\widehat{p}_h\boldsymbol{n}=\nu\mathrm{L}_h\boldsymbol{n}-p_h\boldsymbol{n}-\mathrm{S}\,(\boldsymbol{u}_h-\widehat{\boldsymbol{u}}_h),$$
$$\mathsf{P}_M(\nu\mathrm{L}\boldsymbol{n}-p\boldsymbol{n})=\nu\Pi\,\mathrm{L}\boldsymbol{n}-\mathit{\Pi}p\,\boldsymbol{n}-\mathrm{S}(\boldsymbol{\mathit{\Pi}u}-\mathsf{P}_M\boldsymbol{u}),$$

where P_M is the L^2-projection into $\boldsymbol{M}_h$. This property will allow us to obtain remarkably simple equations for the projection of the errors. As we are going to see, the estimation of these projections then becomes extremely concise and sharp.

Next, we show that the projection $\mathit{\Pi}_h$ is well defined and has reasonable approximation properties in the L^2 norm.

A general stabilization tensor S. We begin by considering any stabilization tensor S satisfying conditions (2.3) and (2.4).

Theorem 2.1. *Suppose that the conditions* (2.3) *and* (2.4) *on the stabilization tensor* S *are satisfied. Then the projection* $\mathit{\Pi}_h$ *is well defined. Moreover, on each element* $K \in \mathcal{T}_h$*, we have that*

$$\|\boldsymbol{\Pi u} - \boldsymbol{u}\|_K \leq \; C\,\frac{\Lambda_K^{\max}}{\lambda_K^{\min}}\, h_K^{\ell_{\boldsymbol{u}}+1}|\boldsymbol{u}|_{\boldsymbol{H}^{\ell_{\boldsymbol{u}}+1}(K)} + C\,\frac{h_K^{\ell_\sigma+1}}{\lambda_K^{\min}}|\,\boldsymbol{\nabla}\boldsymbol{\cdot}(\nu\mathrm{L} - p\mathrm{I})|_{\mathrm{H}^{\ell_\sigma}(K)},$$

and, if $\operatorname{tr}\mathrm{L} = 0$*, then*

$$\begin{aligned}\|\nu\Pi\,\mathrm{L} - \nu\mathrm{L}\|_K + \|\mathit{\Pi} p - p\|_K \leq \; & C\,h_K^{\ell_\sigma+1}\,|\nu\mathrm{L} - p\mathrm{I}|_{\mathrm{H}^{\ell_\sigma+1}(K)} \\ & + C\,\Lambda_K^{\max} h_K^{\ell_{\boldsymbol{u}}+1}|\boldsymbol{u}|_{\boldsymbol{H}^{\ell_{\boldsymbol{u}}+1}(K)} + C\,\Lambda_K^{\max}\|\boldsymbol{\Pi u} - \boldsymbol{u}\|_K,\end{aligned}$$

where $\ell_{\boldsymbol{u}}, \ell_\sigma \in [0, k]$.

We immediately see that if the stabilization tensor S is *independent* of the mesh size, the order of convergence of the approximation errors in each of the variables $\boldsymbol{u}$, p and L is optimal, that is, $k+1$, if they are smooth enough.

Finer results can be obtained by exploiting the structure of the stabilization tensor S. Next, we present them for the two special choices of S given in the previous subsection.

The stabilization tensor S_{SF}. In this case, we have the following result.

Theorem 2.2. *Let* S *be the stabilization tensor* S_{SF} *given by* (2.6) *and suppose that* $\tau > 0$*. Then the projection* $\mathit{\Pi}_h$ *is well defined. Moreover, on each element* $K \in \mathcal{T}_h$*, we have that*

$$\begin{aligned}\|\boldsymbol{\Pi u} - \boldsymbol{u}\|_K &\leq C\,h_K^{\ell_{\boldsymbol{u}}+1}|\boldsymbol{u}|_{\boldsymbol{H}^{\ell_{\boldsymbol{u}}+1}(K)} + C\,\frac{h_K^{\ell_\sigma+1}}{\nu\,\tau}|\,\boldsymbol{\nabla}\boldsymbol{\cdot}(\nu\mathrm{L} - p\mathrm{I})|_{\mathrm{H}^{\ell_\sigma}(K)}, \\ \|\nu\Pi\,\mathrm{L} - \nu\mathrm{L}\|_K &\leq C\,h_K^{\ell_\sigma+1}\,|\nu\mathrm{L} - p\mathrm{I}|_{\mathrm{H}^{\ell_\sigma+1}(K)}, \\ \|\mathit{\Pi} p - p\|_K &\leq C\,h_K^{\ell_\sigma+1}\,|\nu\mathrm{L} - p\mathrm{I}|_{\mathrm{H}^{\ell_\sigma+1}(K)},\end{aligned}$$

where $\ell_{\boldsymbol{u}}, \ell_\sigma \in [0, k]$. *We have assumed that* $\operatorname{tr}\mathrm{L} = 0$ *for the last two inequalities.*

Note that the error in the approximation of the pressure and the velocity gradient is *independent* of the value of the stabilization parameter τ, and is in full agreement with similar results for the SFH method for symmetric second-order elliptic problems; see [7] and [14]. The orders of convergence of the approximation errors for some key choices of the parameter τ are displayed in Table 2.1.

TABLE 2.1. Orders of convergence of the approximation errors $e_{\boldsymbol{u}} = \|\,\boldsymbol{\Pi u} - \boldsymbol{u}\,\|_K$, $e_p = \|\,\mathit{\Pi} p - p\,\|_K$ and $e_{\mathrm{L}} = \|\,\Pi\,\mathrm{L} - \mathrm{L}\,\|_K$ for $k \geq 0$ for the SFH method.

τ	$e_{\boldsymbol{u}}$	e_p, e_{L}
h	$\boldsymbol{k}$	$k+1$
1	$k+1$	$k+1$
$1/h$	$k+1$	$k+1$

The stabilization tensor S_{nt}**.** Finally, we consider the case in which S is the stabilization tensor S_{nt} given by (2.7). For the sake of simplicity, we assume that the functions τ_n and τ_t are constant on ∂K.

Theorem 2.3. *Let* S *be the stabilization tensor* S_{nt} *given by* (2.7)*. Suppose that* τ_n *and* τ_t *are nonnegative constants on* ∂K *satisfying* $\max\{\tau_n, \tau_t\} > 0$*. Then the projection* $\mathit{\Pi}_h$ *is well defined. Moreover, on each element* $K \in \mathcal{T}_h$*, we have that*

$$\|\boldsymbol{\mathit{\Pi}} \boldsymbol{u} - \boldsymbol{u}\|_K \le C\, h_K^{\ell_{\boldsymbol{u}}+1} |\boldsymbol{u}|_{\boldsymbol{H}^{\ell_{\boldsymbol{u}}+1}(K)} + C\, \frac{h_K^{\ell_\sigma+1}}{\nu\, \max\{\tau_n, \tau_t\}} |\, \boldsymbol{\nabla\cdot}(\nu \mathrm{L} - p\mathrm{I})|_{\mathrm{H}^{\ell_\sigma}(K)},$$

$$\|\nu \Pi\,\mathrm{L} - \nu \mathrm{L}\|_K \le C\, h_K^{\ell_{\mathrm{L}}+1}\, |\nu \mathrm{L}|_{\mathrm{H}^{\ell_{\mathrm{L}}+1}(K)} + C\,\nu\,\tau_t \Big(\|\boldsymbol{\mathit{\Pi}} \boldsymbol{u} - \boldsymbol{u}\|_K + h^{\ell_{\boldsymbol{u}}+1} |\boldsymbol{u}|_{\boldsymbol{H}^{\ell_{\boldsymbol{u}}+1}(K)} \Big),$$

$$\|\mathit{\Pi} p - p\|_K \le C\, h_K^{\ell_p+1}\, |p|_{H^{\ell_p+1}(K)} + \|\nu \Pi\,\mathrm{L} - \nu \mathrm{L}\|_K + C\, h_K^{\ell_{\mathrm{L}}+1}\, |\nu \mathrm{L}|_{\mathrm{H}^{\ell_{\mathrm{L}}+1}(K)},$$

where $\ell_{\boldsymbol{u}}, \ell_\sigma, \ell_{\mathrm{L}}, \ell_p \in [0, k]$*. We have assumed that* $\mathrm{tr}\ \mathrm{L} = 0$ *for the last two inequalities and that* $\nabla \cdot \boldsymbol{u} = 0$ *in the last one.*

It is interesting to see that the effect of τ_n and τ_t on the approximation properties of the projection is very different. In particular, if $\tau_n \ge \tau_t$ (as well as $\mathrm{tr}\ \mathrm{L} = \nabla\cdot \boldsymbol{u} = 0$), the convergence properties of the projection become *independent* of how large τ_n is. The orders of convergence for several choices of the parameters τ_n and τ_t are given in Table 2.2.

TABLE 2.2. Orders of convergence of the approximation errors $e_{\boldsymbol{u}} = \|\, \boldsymbol{\mathit{\Pi}} \boldsymbol{u} - \boldsymbol{u}\,\|_K$, $e_p = \|\, \mathit{\Pi} p - p\,\|_K$ and $e_{\mathrm{L}} = \|\, \Pi\,\mathrm{L} - \mathrm{L}\,\|_K$ for $k \ge 0$ when τ_n and τ_t are constant on $\partial \mathcal{T}_h$. Note that the projection is not defined when $\tau_n = \tau_t = 0$.

	$\tau_t = 0, h$		$\tau_t = 1$		$\tau_t = 1/h$	
τ_n	$e_{\boldsymbol{u}}$	e_p, e_{L}	$e_{\boldsymbol{u}}$	e_p, e_{L}	$e_{\boldsymbol{u}}$	e_p, e_{L}
$0, h$	$\boldsymbol{k}$	$k+1$	$k+1$	$k+1$	$k+1$	$\boldsymbol{k}$
1	$k+1$	$k+1$	$k+1$	$k+1$	$k+1$	$\boldsymbol{k}$
$1/h$	$k+1$	$k+1$	$k+1$	$k+1$	$k+1$	$\boldsymbol{k}$

2.3. **The a priori error estimates.** Next, we provide estimates of the *projection* of the approximation errors, namely, of $\mathrm{E}^{\mathrm{L}} := \Pi\,\mathrm{L} - \mathrm{L}_h$, $\boldsymbol{\varepsilon}^{\boldsymbol{u}} := \boldsymbol{\mathit{\Pi}} \boldsymbol{u} - \boldsymbol{u}_h$, $\varepsilon^p := \mathit{\Pi} p - p_h$ and $\boldsymbol{\varepsilon}^{\widehat{\boldsymbol{u}}} := \mathsf{P}_M \boldsymbol{u} - \widehat{\boldsymbol{u}}_h$; we also provide an estimate of $\boldsymbol{u}_h^\star - \boldsymbol{u}$.

To state the results, we need to introduce the following dual problem. For any given $\boldsymbol{\theta}$ in $\boldsymbol{L}^2(\Omega)$ let $(\Phi, \boldsymbol{\phi}, \phi)$ be the solution of

$$\Phi + \nabla \boldsymbol{\phi} = 0 \quad \text{on } \Omega, \tag{2.11a}$$

$$\boldsymbol{\nabla\cdot}(\nu\, \Phi) - \boldsymbol{\nabla}\, \phi = \boldsymbol{\theta} \quad \text{on } \Omega, \tag{2.11b}$$

$$-\nabla\cdot \boldsymbol{\phi} = 0 \quad \text{on } \Omega, \tag{2.11c}$$

$$\boldsymbol{\phi} = \boldsymbol{0} \quad \text{on } \partial\Omega. \tag{2.11d}$$

We assume that, for some real number s, we have that

$$\nu \|\Phi\|_{\mathrm{H}^{s+1}(\Omega)} + \nu \|\boldsymbol{\phi}\|_{\boldsymbol{H}^{s+2}(\Omega)} + \|\phi\|_{H^{s+1}(\Omega)} \le C_{\mathrm{reg}} \|\boldsymbol{\theta}\|_{\boldsymbol{H}^s(\Omega)}. \tag{2.12}$$

In the two-dimensional case, the above estimate with $s \le 0$ follows from the results in [24] when the domain is convex. In the three-dimensional case, the above estimate follows from the results in [20] in the following cases. For any polyhedron, with $s < -1/2$, for any convex polyhedron, with $s \le 0$, and with $s < 3/2$ if, moreover, all the edges have wedge angles at most $2\pi/3$ (a cube, for example). Henceforth, we will use the following notation for the total average of a function over Ω:

$$\overline{p} := \frac{1}{|\Omega|} \int_{\Omega} p.$$

Note that $\overline{\mathit{\Pi} p - p} = (\mathit{\Pi} p - p, 1)/|\Omega| = 0$ except in the case $k = 0$ (this term will appear in the bound for the error in p in Theorem 2.4 and in the right–hand side of the last error equation in Lemma 3.1). Also, note that $|\overline{\mathit{\Pi} p - p}| \le |\Omega|^{1/2} \|\mathit{\Pi} p - p\|$.

We are now ready to state our main results.

Theorem 2.4. *Suppose that the assumptions on the stabilization tensor* S *of Theorem* (2.1) *hold. Then*

$$\| \mathrm{E}^{\mathrm{L}} \|_{\Omega} \le \| \Pi\, \mathrm{L} - \mathrm{L} \|_{\Omega},$$
$$\|\varepsilon^{p}\|_{\Omega} \le |\overline{(\mathit{\Pi} p - p)}|\, |\Omega|^{1/2} + C\, C_p(\mathrm{S})\, \nu\, \|\Pi\, \mathrm{L} - \mathrm{L}\|_{\Omega},$$

where

$$C_p(\mathrm{S}) := \max\Big\{1, \max_{K \in \mathcal{T}_h} \{\Lambda_K^{\max}\, h_K/\nu\}^{1/2}\Big\}.$$

We can take $C_p(\mathrm{S}_{SF}) = 1$*. Moreover, if the elliptic regularity estimate* (2.12) *holds with* $s = 0$*, we have*

$$\| \boldsymbol{\varepsilon}^{\boldsymbol{u}} \|_{\Omega} + \| \boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}} \|_{h} \le C\, C_{\boldsymbol{u}}(\mathrm{S})\, h^{\min\{k,1\}}\, \| \Pi\, \mathrm{L} - \mathrm{L} \|_{\Omega},$$

where

$$C_{\boldsymbol{u}}(\mathrm{S}) = \max_{K \in \mathcal{T}_h} \left\{ \frac{\Lambda_K^{\max}}{\lambda_K^{\min}} \Big(1 + \frac{\Lambda_K^{\max} h_K}{\nu}\Big) \right\}.$$

We can take $C_{\boldsymbol{u}}(\mathrm{S}_{SF}) = 1$.

Theorem 2.5. *Under the assumptions of Theorem* 2.4*, we have that* $\boldsymbol{u}_h^{\star} \in \boldsymbol{H}(\mathrm{div}, \Omega)$ *and that* $\nabla \cdot \boldsymbol{u}_h^{\star} = 0$ *on* Ω*. Moreover,*

$$\|\boldsymbol{u}_h^{\star} - \boldsymbol{u}\|_{\Omega} \le C\, h^{\ell_{\boldsymbol{u}}+2} |\, \boldsymbol{u}\,|_{\boldsymbol{H}^{\ell_{\boldsymbol{u}}+2}(\Omega)} + C\, C_{\boldsymbol{u}}(\mathrm{S})\, h^{\min\{k,1\}}\, \|\Pi\, \mathrm{L} - \mathrm{L}\|_{\Omega}.$$

where $\ell_{\boldsymbol{u}} \in [0, k]$.

Let us briefly discuss the main consequences of these results:

- For $k = 0$, if the constants $C_p(\mathrm{S})$, $C_{\boldsymbol{u}}(\mathrm{S})$ and $C_{\boldsymbol{u}^\star}(\mathrm{S})$ are uniformly bounded, the quantities $\|\mathrm{E}^{\mathrm{L}}\|_{\Omega}$, $\|\boldsymbol{\varepsilon}^{\boldsymbol{u}}\|_{\Omega}$, $\|\boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}\|_{h}$ and $\|\boldsymbol{u}_h^{\star} - \boldsymbol{u}\|_{\Omega}$ converge with the *same* order as $\|\Pi\,\mathrm{L} - \mathrm{L}\|_{\Omega}$. The quantity $\|\varepsilon^{p}\|_{\Omega}$ converges with the *same* order of convergence as $\|\mathit{\Pi} p - p\|_{\Omega} + \|\Pi\,\mathrm{L} - \mathrm{L}\|_{\Omega}$.
- For $k \ge 1$, if the constants $C_p(\mathrm{S})$, $C_{\boldsymbol{u}}(\mathrm{S})$ and $C_{\boldsymbol{u}^\star}(\mathrm{S})$ are uniformly bounded, the quantities $\|\mathrm{E}^{\mathrm{L}}\|_{\Omega}$, $\|\varepsilon^{p}\|_{\Omega}$ converge with the *same* order of convergence as $\|\Pi\,\mathrm{L} - \mathrm{L}\|_{\Omega}$. However, the quantities $\|\boldsymbol{\varepsilon}^{\boldsymbol{u}}\|_{\Omega}$, $\|\boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}\|_{h}$ and $\|\boldsymbol{u}_h^{\star} - \boldsymbol{u}\|_{\Omega}$ converge with an *additional* order.
- The constants $C_p(\mathrm{S})$, $C_{\boldsymbol{u}}(\mathrm{S})$ and $C_{\boldsymbol{u}^\star}(\mathrm{S})$ are uniformly bounded when the stabilization tensor S is of order one, when the stabilization tensor is S_{SF}, or when the stabilization tensor is S_{nt} and $\max\{\tau_n|_{\partial K}, \tau_t|_{\partial K}\}$ is of order $2h_K^{-1}$. The corresponding orders of convergence are displayed in Tables 2.3 and 2.4.

Table 2.3. Orders of convergence of the **projection of the errors** $e_{\boldsymbol{u}} = \|\, \boldsymbol{\varepsilon}^{\boldsymbol{u}} \,\|_\Omega$, $e_{\boldsymbol{u}^\star} = \|\, \boldsymbol{u}_h^\star - \boldsymbol{u} \,\|_\Omega$, $e_{\widehat{\boldsymbol{u}}} = \|\, \boldsymbol{\varepsilon}^{\widehat{u}} \,\|_h$, $e_p = \|\, \varepsilon^p \,\|_\Omega$ and $e_{\mathrm{L}} = \|\, \mathrm{E}^{\mathrm{L}} \,\|_\Omega$ for $k \geq 0$ for a general stabilization tensor S of order one, and for the stabilization tensor S_{SF} with $\tau \in \{0, h, 1, 1/h\}$.

	$e_{\boldsymbol{u}}, e_{\boldsymbol{u}^\star}, e_{\widehat{\boldsymbol{u}}}$	e_p, e_{L}
$k = 0$	1	1
$k \geq 1$	$k+2$	$k+1$

Table 2.4. Orders of convergence of the **projection of the errors** $e_{\boldsymbol{u}} = \|\, \boldsymbol{\varepsilon}^{\boldsymbol{u}} \,\|_\Omega$, $e_{\boldsymbol{u}^\star} = \|\, \boldsymbol{u}_h^\star - \boldsymbol{u} \,\|_\Omega$, $e_{\widehat{\boldsymbol{u}}} = \|\, \boldsymbol{\varepsilon}^{\widehat{u}} \,\|_h$, $e_p = \|\, \varepsilon^p \,\|_\Omega$ and $e_{\mathrm{L}} = \|\, \mathrm{E}^{\mathrm{L}} \,\|_\Omega$ for $k \geq 0$ when τ_n and τ_t are constant on $\partial\mathcal{T}_h$, and $\tau_n \in \{0, h, 1, 1/h\}$. Note that the projection is not defined when $\tau_n = \tau_t = 0$.

	$\tau_t = 0, h$		$\tau_t = 1$		$\tau_t = 1/h$	
	$e_{\boldsymbol{u}}, e_{\boldsymbol{u}^\star}, e_{\widehat{\boldsymbol{u}}}$	e_p, e_{L}	$e_{\boldsymbol{u}}, e_{\boldsymbol{u}^\star}, e_{\widehat{\boldsymbol{u}}}$	e_p, e_{L}	$e_{\boldsymbol{u}}, e_{\boldsymbol{u}^\star}, e_{\widehat{\boldsymbol{u}}}$	e_p, e_{L}
$k = 0$	1	1	1	1	$\mathbf{0}$	$\mathbf{0}$
$k \geq 1$	$k+2$	$k+1$	$k+2$	$k+1$	$\boldsymbol{k+1}$	$\boldsymbol{k}$

Let us emphasize that Tables 2.3 and 2.4 do not show convergence orders of the HDG methods but of convergence of the numerical solutions to the HDG projection of the exact solution. The only exception is the comparison between the postprocessed approximation of $\boldsymbol{u}$ to its exact value. The orders of convergence for the total error can be tabulated putting together the information about the errors shown in these tables to the approximation errors in Tables 2.1 and 2.2. Since the case of the stabilization tensor S_{nt} offers a wide variety of situations, because we can choose parameters with different magnitudes, we show the resulting table in Section 5, for ease of comparison with the numerical results obtained therein.

3. Proofs

This section is devoted to providing detailed proofs of all the results of the previous section, except for the approximation result of Theorem 2.1 which is proven in the next section. Here, we assume that the projection is well defined and proceed as follows. We begin by obtaining the equations satisfied by the projection of the error. We then use an energy argument to obtain an estimate of the L^2-error of the velocity gradient. Then the estimate of the pressure follows by using a typical inf-sup argument. Next we obtain error estimates of the velocity by using a duality argument. The estimate of the velocity trace then follows by using a simple scaling argument. Finally, the error estimates of the postprocessed velocity are obtained.

Step 1: The error equations. We begin by obtaining the equations satisfied by the projection of the errors. They are contained in the following result.

Lemma 3.1.

$$(\mathrm{E}^{\mathrm{L}}, \mathrm{G})_{\mathcal{T}_h} + (\boldsymbol{\varepsilon}^{\boldsymbol{u}}, \boldsymbol{\nabla\cdot}\, \mathrm{G})_{\mathcal{T}_h} - \langle \boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}, \mathrm{G}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} = (\Pi\, \mathrm{L} - \mathrm{L}, \mathrm{G})_{\mathcal{T}_h}, \tag{3.1a}$$

$$-(\boldsymbol{\nabla\cdot}(\nu \mathrm{E}^{\mathrm{L}}), \boldsymbol{v})_{\mathcal{T}_h} + (\boldsymbol{\nabla}\, \varepsilon^p, \boldsymbol{v})_{\mathcal{T}_h} + \langle \mathrm{S}(\boldsymbol{\varepsilon}^{\boldsymbol{u}} - \boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}), \boldsymbol{v}\rangle_{\partial\mathcal{T}_h} = 0, \tag{3.1b}$$

$$-(\boldsymbol{\varepsilon}^{\boldsymbol{u}}, \boldsymbol{\nabla}\, q)_{\mathcal{T}_h} + \langle \boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}, q\,\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} = 0 \tag{3.1c}$$

$$\langle \boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}, \boldsymbol{\mu}\rangle_{\partial\Omega} = 0, \tag{3.1d}$$

$$\langle \nu \mathrm{E}^{\mathrm{L}}\boldsymbol{n} - \varepsilon^p\boldsymbol{n} - \mathrm{S}(\boldsymbol{\varepsilon}^{\boldsymbol{u}} - \boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}), \boldsymbol{\mu}\rangle_{\partial\mathcal{T}_h\setminus\partial\Omega} = 0, \tag{3.1e}$$

$$(\varepsilon^p, 1)_\Omega = (\mathit{\Pi} p - p, 1)_\Omega. \tag{3.1f}$$

Proof. Let us begin by noting that, if we insert the expression of the numerical trace given in (2.2g) into the second and sixth equations defining the HDG method, they read as follows:

$$\begin{aligned}
(\mathrm{L}_h, \mathrm{G})_{\mathcal{T}_h} + (\boldsymbol{u}_h, \boldsymbol{\nabla\cdot}\, \mathrm{G})_{\mathcal{T}_h} - \langle \widehat{\boldsymbol{u}}_h, \mathrm{G}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} &= 0,\\
(\nu \mathrm{L}_h, \nabla\, \boldsymbol{v})_{\mathcal{T}_h} - (p_h, \nabla\cdot\, \boldsymbol{v})_{\mathcal{T}_h} - \langle \nu \mathrm{L}_h\boldsymbol{n} - p_h\boldsymbol{n} - \mathrm{S}\,(\boldsymbol{u}_h - \widehat{\boldsymbol{u}}_h), \boldsymbol{v}\rangle_{\partial\mathcal{T}_h} &= (\boldsymbol{f}, \boldsymbol{v})_{\mathcal{T}_h},\\
-(\boldsymbol{u}_h, \boldsymbol{\nabla}\, q)_{\mathcal{T}_h} + \langle \widehat{\boldsymbol{u}}_h \cdot \boldsymbol{n}, q\rangle_{\partial\mathcal{T}_h} &= 0,\\
\langle \widehat{\boldsymbol{u}}_h, \boldsymbol{\mu}\rangle_{\partial\Omega} &= \langle \boldsymbol{g}, \boldsymbol{\mu}\rangle_{\partial\Omega},\\
\langle \nu \mathrm{L}_h\boldsymbol{n} - p_h\boldsymbol{n} - \mathrm{S}\,(\boldsymbol{u}_h - \widehat{\boldsymbol{u}}_h), \boldsymbol{\mu}\rangle_{\partial\mathcal{T}_h\setminus\partial\Omega} &= 0,\\
(p_h, 1)_\Omega &= 0,
\end{aligned}$$

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, after applying the definition of the projection (2.10), we obtain

$$\begin{aligned}
(\mathrm{L}, \mathrm{G})_{\mathcal{T}_h} + (\boldsymbol{\mathit{\Pi}} \boldsymbol{u}, \boldsymbol{\nabla\cdot}\, \mathrm{G})_{\mathcal{T}_h} - \langle \mathsf{P}_M \boldsymbol{u}, \mathrm{G}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} &= 0,\\
(\nu \Pi\, \mathrm{L}, \nabla\, \boldsymbol{v})_{\mathcal{T}_h} - (\mathit{\Pi} p, \nabla\cdot\, \boldsymbol{v})_{\mathcal{T}_h} - \langle \nu \Pi\, \mathrm{L}\boldsymbol{n} - \mathit{\Pi} p\boldsymbol{n} - \mathrm{S}\,(\boldsymbol{\mathit{\Pi}} \boldsymbol{u} - \mathsf{P}_M \boldsymbol{u}), \boldsymbol{v}\rangle_{\partial\mathcal{T}_h} &= (\boldsymbol{f}, \boldsymbol{v})_{\mathcal{T}_h},\\
-(\boldsymbol{\mathit{\Pi}} \boldsymbol{u}, \boldsymbol{\nabla}\, q)_{\mathcal{T}_h} + \langle \mathsf{P}_M \boldsymbol{u} \cdot \boldsymbol{n}, q\rangle_{\partial\mathcal{T}_h} &= 0,\\
\langle \mathsf{P}_M \boldsymbol{u}, \boldsymbol{\mu}\rangle_{\partial\Omega} &= \langle \boldsymbol{g}, \boldsymbol{\mu}\rangle_{\partial\Omega},\\
\langle \nu \Pi\, \mathrm{L}\boldsymbol{n} - \mathit{\Pi} p\boldsymbol{n} - \mathrm{S}\,(\boldsymbol{\mathit{\Pi}} \boldsymbol{u} - \mathsf{P}_M \boldsymbol{u}), \boldsymbol{\mu}\rangle_{\partial\mathcal{T}_h\setminus\partial\Omega} &= 0,\\
(p, 1)_\Omega &= 0,
\end{aligned}$$

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 3.1. □

Step 2: Estimate of the velocity gradient. We are now ready to obtain our first estimate by using a standard energy argument. Note that, from now on, we measure errors of quantities defined on $\partial\mathcal{T}_h$ with the following seminorm:

$$\|\, \boldsymbol{\mu}\, \|_{h^m \mathrm{Z}} := \left\{ \sum_{K\in\mathcal{T}_h} h_K^m \,\langle \mathrm{Z}\boldsymbol{\mu}, \boldsymbol{\mu}\rangle_{\partial K} \right\}^{1/2},$$

where $m \in \{0, 1\}$ and Z is a positive semidefinite matrix-valued function defined on $\partial\mathcal{T}_h$. When $\mathrm{Z} = \mathrm{I}$ and $m = 1$ we will simply write

$$\|\, \boldsymbol{\mu}\, \|_{h} := \left\{ \sum_{K\in\mathcal{T}_h} h_K \,\langle \boldsymbol{\mu}, \boldsymbol{\mu}\rangle_{\partial K} \right\}^{1/2},$$

Proposition 3.2. *We have*

$$\|\mathrm{E}^{\mathrm{L}}\|^2_{\mathcal{T}_h} + \|\boldsymbol{\varepsilon}^{\boldsymbol{u}} - \boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}\|^2_{\mathrm{S}/\nu} = (\Pi\,\mathrm{L} - \mathrm{L}, \mathrm{E}^{\mathrm{L}})_{\mathcal{T}_h}.$$

Proof. If we take $\mathrm{G} := \nu\mathrm{E}^{\mathrm{L}}$, $\boldsymbol{v} := \boldsymbol{\varepsilon}^{\boldsymbol{u}}$ and $q := \varepsilon^p$ in the first three error equations (3.1), respectively, and add them up, we obtain

$$\nu(\mathrm{E}^{\mathrm{L}}, \mathrm{E}^{\mathrm{L}})_{\mathcal{T}_h} + \Theta_h = \nu(\Pi\,\mathrm{L} - \mathrm{L}, \mathrm{E}^{\mathrm{L}})_{\mathcal{T}_h},$$

where

$$\begin{aligned}\Theta_h := {} & (\boldsymbol{\varepsilon}^{\boldsymbol{u}}, \boldsymbol{\nabla\cdot}(\nu\mathrm{E}^{\mathrm{L}}))_{\mathcal{T}_h} - \langle \boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}, \nu\mathrm{E}^{\mathrm{L}}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} \\ & - (\boldsymbol{\nabla\cdot}(\nu\mathrm{E}^{\mathrm{L}}), \boldsymbol{\varepsilon}^{\boldsymbol{u}})_{\mathcal{T}_h} + (\boldsymbol{\nabla}\,\varepsilon^p, \boldsymbol{\varepsilon}^{\boldsymbol{u}})_{\mathcal{T}_h} + \langle \mathrm{S}(\boldsymbol{\varepsilon}^{\boldsymbol{u}} - \boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}), \boldsymbol{\varepsilon}^{\boldsymbol{u}}\rangle_{\partial\mathcal{T}_h} \\ & - (\boldsymbol{\varepsilon}^{\boldsymbol{u}}, \boldsymbol{\nabla}\,\varepsilon^p)_{\mathcal{T}_h} + \langle \boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}, \varepsilon^p\,\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}.\end{aligned}$$

We thus obtain

$$\begin{aligned}\Theta_h &= -\,\langle \boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}, \nu\mathrm{E}^{\mathrm{L}}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} + \langle \mathrm{S}(\boldsymbol{\varepsilon}^{\boldsymbol{u}} - \boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}), \boldsymbol{\varepsilon}^{\boldsymbol{u}}\rangle_{\partial\mathcal{T}_h} + \langle \boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}, \varepsilon^p\,\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} \\ &= \langle \boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}, -\nu\mathrm{E}^{\mathrm{L}}\boldsymbol{n} + \varepsilon^p\boldsymbol{n} + \mathrm{S}(\boldsymbol{\varepsilon}^{\boldsymbol{u}} - \boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}})\rangle_{\partial\mathcal{T}_h} + \langle \mathrm{S}(\boldsymbol{\varepsilon}^{\boldsymbol{u}} - \boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}), \boldsymbol{\varepsilon}^{\boldsymbol{u}} - \boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}\rangle_{\partial\mathcal{T}_h} \\ &= \langle \mathrm{S}(\boldsymbol{\varepsilon}^{\boldsymbol{u}} - \boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}), \boldsymbol{\varepsilon}^{\boldsymbol{u}} - \boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}\rangle_{\partial\mathcal{T}_h},\end{aligned}$$

by the last two error equations (3.1) with $\boldsymbol{\mu} := \boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}$. Finally, we get that

$$\Theta_h = \nu\,\|\boldsymbol{\varepsilon}^{\boldsymbol{u}} - \boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}\|^2_{\mathrm{S}/\nu},$$

by the definition of the seminorm $\|\cdot\|_{\mathrm{S}/\nu}$. This completes the proof. $\square$

As a straightforward consequence of Proposition 3.2, we obtain the first error estimate for the method

$$\|\mathrm{E}^{\mathrm{L}}\|_{\mathcal{T}_h} + \|\boldsymbol{\varepsilon}^{\boldsymbol{u}} - \boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}\|_{\mathrm{S}/\nu} \le \|\Pi\,\mathrm{L} - \mathrm{L}\|_{\mathcal{T}_h}.$$

We can obtain a better estimate for $\|\boldsymbol{\varepsilon}^{\boldsymbol{u}} - \boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}\|_{\mathrm{S}_{SF}/\nu}$, as we show next. A similar result holds for the SFH method for second-order elliptic problems; see [7]. Recall that in the SF hybridized method, only one face $F_K^\star$ per element has been stabilized. Therefore, the seminorm $\|\boldsymbol{\mu}\|_{\mathrm{S}_{SF}/\nu}$ controls only the part of $\boldsymbol{\mu}$ defined on the faces where stabilization has been imposed.

Lemma 3.3. *We have that* $\|\boldsymbol{\varepsilon}^{\boldsymbol{u}} - \boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}\|_{\mathrm{S}_{SF}/\nu} = 0$.

Proof. By definition of the seminorm $\|\cdot\|_{\mathrm{Z}}$ and that of the definition of the stabilization tensor S_{SF} (2.6), we have that

$$\begin{aligned}\|\boldsymbol{\varepsilon}^{\boldsymbol{u}} - \boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}\|^2_{\mathrm{S}_{SF}/\nu} &= \sum_{K\in\mathcal{T}_h} \nu^{-1}\langle \mathrm{S}(\boldsymbol{\varepsilon}^{\boldsymbol{u}} - \boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}), \boldsymbol{\varepsilon}^{\boldsymbol{u}} - \boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}\rangle_{\partial K} \\ &= \sum_{K\in\mathcal{T}_h} \nu^{-1}\langle \mathrm{S}(\boldsymbol{\varepsilon}^{\boldsymbol{u}} - \boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}), \boldsymbol{\varepsilon}^{\boldsymbol{u}} - \boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}\rangle_{F_K^\star}.\end{aligned}$$

It remains to show that on $F_K^\star$, we have $\boldsymbol{\varepsilon}^{\boldsymbol{u}} - \boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}} = 0$.

To do that, we proceed as follows. Take in the error equation (3.1b) $\boldsymbol{v}$ to be a function such that, on each simplex $K \in \mathcal{T}_h$, it is orthogonal to $\boldsymbol{\mathcal{P}}_{k-1}(K)$ and satisfies $\boldsymbol{v}\cdot\boldsymbol{n}_{F_K^\star} = \boldsymbol{\varepsilon}^{\boldsymbol{u}} - \boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}$ on $F_K^\star$. Then we get that $\boldsymbol{\varepsilon}^{\boldsymbol{u}} - \boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}} = 0$ on $F_K^\star$, as wanted. This completes the proof. $\square$

Step 3: Estimate of the pressure. Next, we show how to use the previous result to obtain the estimate of the pressure.

Proposition 3.4. *Let* $\boldsymbol{P} : \boldsymbol{H}^1(\mathcal{T}_h) \mapsto \boldsymbol{V}_h$ *be any projection such that* $(\boldsymbol{P}\boldsymbol{w} - \boldsymbol{w}, \boldsymbol{v})_K = 0$ *for all* $\boldsymbol{v} \in \boldsymbol{\mathcal{P}}_{k-1}(K)$ *for all* $K \in \mathcal{T}_h$. *Then we have*

$$\|\varepsilon^p - \overline{(\varPi p - p)}\|_\Omega \leq C\,\Psi(\mathrm{S})\,\nu\,\|\Pi\,\mathrm{L} - \mathrm{L}\|_\Omega,$$

where

$$\Psi(\mathrm{S}) := \max\left\{1, \sup_{\boldsymbol{w}\in \boldsymbol{H}_0^1(\Omega)\setminus\{\boldsymbol{0}\}} \frac{\|\boldsymbol{P}\boldsymbol{w} - \mathsf{P}_M \boldsymbol{w}\|_{\mathrm{S}/\nu}}{\|\boldsymbol{w}\|_{\boldsymbol{H}^1(\Omega)}}\right\}.$$

Proof. It is well known [4] that for any function $q \in L^2(\Omega)$ such that $(q,1)_\Omega = 0$ we have

$$\|q\|_\Omega \leq \kappa \sup_{\boldsymbol{w}\in \boldsymbol{H}_0^1(\Omega)\setminus\{0\}} \frac{(q, \nabla\cdot \boldsymbol{w})_\Omega}{\|\boldsymbol{w}\|_{\boldsymbol{H}^1(\Omega)}},$$

for some constant κ independent of q. By the last error equation, we see that we can apply the above result to $q := \varepsilon^p - \overline{\varepsilon^p}$ Hence we have that

$$\|\varepsilon^p - \overline{\varepsilon^p}\|_\Omega \leq \kappa \sup_{\boldsymbol{w}\in \boldsymbol{H}_0^1(\Omega)\setminus\{\boldsymbol{0}\}} \frac{(\varepsilon^p, \nabla\cdot \boldsymbol{w})_\Omega}{\|\boldsymbol{w}\|_{\boldsymbol{H}^1(\Omega)}}.$$

Next, we work on the numerator in the above expression. We have

$$(\varepsilon^p, \nabla\cdot \boldsymbol{w})_\Omega = -(\nabla\varepsilon^p, \boldsymbol{P}\boldsymbol{w})_{\mathcal{T}_h} + \langle \varepsilon^p, \boldsymbol{w}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}.$$

By the second error equation (see Lemma 3.1), with $\boldsymbol{v} := \boldsymbol{P}\boldsymbol{w}$, we get that

$$\begin{aligned}(\varepsilon^p, \nabla\cdot \boldsymbol{w})_\Omega &= -\,(\boldsymbol{\nabla\cdot}(\nu\mathrm{E}^{\mathrm{L}}), \boldsymbol{P}\boldsymbol{w})_{\mathcal{T}_h} + \langle \mathrm{S}(\boldsymbol{\varepsilon}^{\boldsymbol{u}} - \boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}), \boldsymbol{P}\boldsymbol{w}\rangle_{\partial\mathcal{T}_h} + \langle \varepsilon^p, \boldsymbol{w}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}\\ &= (\nu\mathrm{E}^{\mathrm{L}}, \nabla\boldsymbol{w})_{\mathcal{T}_h} + \langle \mathrm{S}(\boldsymbol{\varepsilon}^{\boldsymbol{u}} - \boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}), \boldsymbol{P}\boldsymbol{w}\rangle_{\partial\mathcal{T}_h} + \langle -\nu\mathrm{E}^{\mathrm{L}}\boldsymbol{n} + \varepsilon^p\boldsymbol{n}, \mathsf{P}_M\boldsymbol{w}\rangle_{\partial\mathcal{T}_h}\\ &= (\nu\mathrm{E}^{\mathrm{L}}, \nabla\boldsymbol{w})_{\mathcal{T}_h} + \langle \mathrm{S}(\boldsymbol{\varepsilon}^{\boldsymbol{u}} - \boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}), \boldsymbol{P}\boldsymbol{w} - \mathsf{P}_M\boldsymbol{w}\rangle_{\partial\mathcal{T}_h}\end{aligned}$$

by the fifth error equation with $\boldsymbol{\mu} = \mathsf{P}_M\boldsymbol{w}$ and by the fact that $\boldsymbol{w} \in \boldsymbol{H}_0^1(\Omega)$. Applying the Cauchy-Schwarz inequality and using Proposition 3.2, we get that

$$|(\varepsilon^p, \nabla\cdot \boldsymbol{w})_\Omega| \leq \Psi(\mathrm{S})\,\nu\,\|\Pi\,\mathrm{L} - \mathrm{L}\|_\Omega\,\|\boldsymbol{w}\|_{\boldsymbol{H}^1(\Omega)}.$$

As a consequence, we obtain that

$$\|\varepsilon^p - \overline{\varepsilon^p}\|_\Omega \leq C\,\Psi(\mathrm{S})\,\nu\,\|\Pi\,\mathrm{L} - \mathrm{L}\|_\Omega,$$

and the result follows from the fact that $\overline{\varepsilon^p} = \overline{\varPi p - p}$ by the error equation (3.1f). This completes the proof of Proposition 3.4 ☐

The needed bound for $\Psi(\mathrm{S})$, which would complete the convergence analysis for the pressure, is given below in Proposition 3.9.

Step 4: Some properties of the projection. Here, in preparation for the duality argument we are going to use to obtain estimates of the velocity, we gather a few useful properties of the projection $\varPi_h$.

Lemma 3.5. *Assume that* $(\Phi, \boldsymbol{\phi}, \phi) \in \mathrm{H}^1(\mathcal{T}_h) \times \boldsymbol{H}^1(\mathcal{T}_h) \times H^1(\mathcal{T}_h)$. *Then we have*

$$(\boldsymbol{v}, \boldsymbol{\nabla\cdot}\,\Phi)_{\mathcal{T}_h} = (\boldsymbol{v}, \nabla\cdot\Pi\,\Phi)_{\mathcal{T}_h} + \langle \boldsymbol{v}, (\Phi - \Pi\,\Phi)\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}, \tag{3.2a}$$

$$(\mathrm{G}, \nabla\,\boldsymbol{\phi})_{\mathcal{T}_h} = -\,(\boldsymbol{\nabla\cdot}\,\mathrm{G}, \boldsymbol{\varPi\phi})_{\mathcal{T}_h} + \langle \mathrm{G}\boldsymbol{n}, \boldsymbol{\phi}\rangle_{\partial\mathcal{T}_h}, \tag{3.2b}$$

$$(q, \nabla\cdot\boldsymbol{\phi})_{\mathcal{T}_h} = -\,(\boldsymbol{\nabla}\,q, \boldsymbol{\varPi\phi})_{\mathcal{T}_h} + \langle q\boldsymbol{n}, \phi\rangle_{\partial\mathcal{T}_h}, \tag{3.2c}$$

$$(\boldsymbol{v}, \boldsymbol{\nabla}\,\phi)_{\mathcal{T}_h} = (\boldsymbol{v}, \boldsymbol{\nabla}\,\varPi\phi)_{\mathcal{T}_h} + \langle \boldsymbol{v}, (\phi - \varPi\phi)\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}, \tag{3.2d}$$

for all $(\mathrm{G}, \boldsymbol{v}, q, \boldsymbol{\mu}) \in \mathrm{G}_h \times \boldsymbol{V}_h \times P_h \times \boldsymbol{M}_h$.

Proof. Let us prove the first identity. We have, by integration by parts, that

$$\begin{aligned}(\boldsymbol{v}, \boldsymbol{\nabla\cdot}\,\Phi)_{\mathcal{T}_h} &= -(\nabla\,\boldsymbol{v}, \Phi)_{\mathcal{T}_h} + \langle \boldsymbol{v}, \Phi\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}\\ &= -(\nabla\,\boldsymbol{v}, \Pi\,\Phi)_{\mathcal{T}_h} + \langle \boldsymbol{v}, \Phi\boldsymbol{n}\rangle_{\partial\mathcal{T}_h},\end{aligned}$$

by the property (2.10a) of the projection $\mathit{\Pi}_h$. Finally, integrating by parts again, we get

$$(\boldsymbol{v}, \boldsymbol{\nabla\cdot}\,\Phi)_{\mathcal{T}_h} = (\boldsymbol{v}, \nabla\cdot\,\Pi\,\Phi)_{\mathcal{T}_h} + \langle \boldsymbol{v}, (\Phi - \Pi\,\Phi)\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}.$$

This proves the first identity. The remaining identities are proven in a similar fashion. This completes the proof of Lemma 3.5. $\square$

Step 5: 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 3.6. *We have*

$$(\boldsymbol{\varepsilon}^{\boldsymbol{u}}, \boldsymbol{\theta})_{\mathcal{T}_h} = \nu(\mathrm{L}_h - \mathrm{L}, \Pi\,\Phi - \Phi)_{\mathcal{T}_h} + \nu(\Pi\,\mathrm{L} - \mathrm{L}, \Phi - \mathrm{P}_{k-1}\Phi)_{\mathcal{T}_h}.$$

Proof. We have

$$(\boldsymbol{\varepsilon}^{\boldsymbol{u}}, \boldsymbol{\theta})_{\mathcal{T}_h} = \nu(\mathrm{E}^{\mathrm{L}}, \Phi + \nabla\,\boldsymbol{\phi})_{\mathcal{T}_h} + (\boldsymbol{\varepsilon}^{\boldsymbol{u}}, \boldsymbol{\nabla\cdot}(\nu\,\Phi) - \boldsymbol{\nabla}\,\phi)_{\mathcal{T}_h} - (\varepsilon^{p}, \nabla\cdot\,\boldsymbol{\phi})_{\mathcal{T}_h},$$

by the first three equations of the dual problem (2.11). Rearranging terms, we get

$$\begin{aligned}(\boldsymbol{\varepsilon}^{\boldsymbol{u}}, \boldsymbol{\theta})_{\mathcal{T}_h} = {} & \nu(\mathrm{E}^{\mathrm{L}}, \Phi)_{\mathcal{T}_h} + \nu(\boldsymbol{\varepsilon}^{\boldsymbol{u}}, \boldsymbol{\nabla\cdot}\,\Phi)\mathcal{T}_h\\ & + (\nu\mathrm{E}^{\mathrm{L}}, \nabla\,\boldsymbol{\phi})_{\mathcal{T}_h} - (\varepsilon^{p}, \boldsymbol{\nabla\cdot}\,\boldsymbol{\phi})_{\mathcal{T}_h}\\ & - (\boldsymbol{\varepsilon}^{\boldsymbol{u}}, \boldsymbol{\nabla}\,\phi)_{\mathcal{T}_h},\end{aligned}$$

and using the first four properties of the projection $\mathit{\Pi}_h$ in Lemma 3.5 on the last four terms of the right-hand side above, respectively, we obtain

$$\begin{aligned}(\boldsymbol{\varepsilon}^{\boldsymbol{u}}, \boldsymbol{\theta})_{\mathcal{T}_h} = {} & \nu(\mathrm{E}^{\mathrm{L}}, \Phi)_{\mathcal{T}_h} + \nu(\boldsymbol{\varepsilon}^{\boldsymbol{u}}, \boldsymbol{\nabla\cdot}\,\Pi\,\Phi)\mathcal{T}_h + \nu\langle\boldsymbol{\varepsilon}^{\boldsymbol{u}}, (\Phi - \Pi\,\Phi)\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}\\ & - (\boldsymbol{\nabla\cdot}(\nu\mathrm{E}^{\mathrm{L}}), \boldsymbol{\mathit{\Pi}\phi})_{\mathcal{T}_h} + \langle\nu\mathrm{E}^{\mathrm{L}}\boldsymbol{n}, \boldsymbol{\phi}\rangle_{\partial\mathcal{T}_h} + (\boldsymbol{\nabla}\,\varepsilon^{p}, \boldsymbol{\mathit{\Pi}\phi})_{\mathcal{T}_h} - \langle\varepsilon^{p}\,\boldsymbol{n}, \boldsymbol{\phi}\rangle_{\partial\mathcal{T}_h}\\ & - (\boldsymbol{\varepsilon}^{\boldsymbol{u}}, \boldsymbol{\nabla}\,\mathit{\Pi}\phi)_{\mathcal{T}_h} - \langle\boldsymbol{\varepsilon}^{\boldsymbol{u}}, (\phi - \mathit{\Pi}\phi)\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}.\end{aligned}$$

Now we use the first three error equations (3.1) with $\mathrm{G} := \nu\Pi\,\Phi$, $\boldsymbol{v} := \boldsymbol{\mathit{\Pi}\phi}$ and $q := \mathit{\Pi}\phi$, respectively, to get

$$\begin{aligned}(\boldsymbol{\varepsilon}^{\boldsymbol{u}}, \boldsymbol{\theta})_{\mathcal{T}_h} = {} & \nu(\mathrm{L}_h - \mathrm{L}, \Pi\,\Phi)_{\mathcal{T}_h} + \nu\langle\boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}, \Pi\,\Phi\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} + \nu\langle\boldsymbol{\varepsilon}^{\boldsymbol{u}}, (\Phi - \Pi\,\Phi)\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}\\ & - \langle\mathrm{S}(\boldsymbol{\varepsilon}^{\boldsymbol{u}} - \boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}), \boldsymbol{\mathit{\Pi}\phi}\rangle_{\partial\mathcal{T}_h} + \langle\nu\mathrm{E}^{\mathrm{L}}\boldsymbol{n}, \boldsymbol{\phi}\rangle_{\partial\mathcal{T}_h} - \langle\varepsilon^{p}\,\boldsymbol{n}, \boldsymbol{\phi}\rangle_{\partial\mathcal{T}_h}\\ & - \langle\boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}, \mathit{\Pi}\phi\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle\boldsymbol{\varepsilon}^{\boldsymbol{u}}, (\phi - \mathit{\Pi}\phi)\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}.\end{aligned}$$

Next, we carry out some simple algebraic manipulations to rewrite the quantity $(\boldsymbol{\varepsilon}^{\boldsymbol{u}}, \boldsymbol{\theta})_{\mathcal{T}_h}$ as $\sum_{i=1}^{5} T_i$, where

$$\begin{aligned}T_1 &:= \nu(\mathrm{L}_h - \mathrm{L}, \Pi\,\Phi - \Phi)_{\mathcal{T}_h},\\ T_2 &:= \nu(\Pi\,\mathrm{L} - \mathrm{L}, \Phi)_{\mathcal{T}_h},\\ T_3 &:= \langle\boldsymbol{\varepsilon}^{\boldsymbol{u}} - \boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}, \nu(\Phi - \Pi\,\Phi)\boldsymbol{n} - (\phi - \mathit{\Pi}\phi)\boldsymbol{n} + \mathrm{S}(\boldsymbol{\phi} - \boldsymbol{\mathit{\Pi}\phi})\rangle_{\partial\mathcal{T}_h},\\ T_4 &:= \langle\nu\mathrm{E}^{\mathrm{L}}\boldsymbol{n} - \varepsilon^{p}\boldsymbol{n} - \mathrm{S}(\boldsymbol{\varepsilon}^{\boldsymbol{u}} - \boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}), \boldsymbol{\phi}\rangle_{\partial\mathcal{T}_h},\\ T_5 &:= \langle\boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}, \Phi\boldsymbol{n} - \phi\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}.\end{aligned}$$

But, $T_2 = \nu(\Pi\,\mathrm{L} - \mathrm{L}, \Phi - \mathrm{P}_{k-1}\Phi)_{\mathcal{T}_h}$, by the property of the projection (2.10a), and $T_3 = 0$, by the property of the projection (2.10e) with $\boldsymbol{\mu} := \boldsymbol{\varepsilon}^{\boldsymbol{u}} - \boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}$. Moreover,

$$\begin{aligned} T_4 &= \langle \nu \mathrm{E}^{\mathrm{L}}\boldsymbol{n} - \varepsilon^p \boldsymbol{n} - \mathrm{S}(\boldsymbol{\varepsilon}^{\boldsymbol{u}} - \boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}), \mathsf{P}_M \boldsymbol{\phi} \rangle_{\partial \mathcal{T}_h} \\ &= \langle \nu \mathrm{E}^{\mathrm{L}}\boldsymbol{n} - \varepsilon^p \boldsymbol{n} - \mathrm{S}(\boldsymbol{\varepsilon}^{\boldsymbol{u}} - \boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}), \mathsf{P}_M \boldsymbol{\phi} \rangle_{\partial \Omega_D} && \text{by (3.1e)} \\ &= 0 && \text{by (2.11d).} \end{aligned}$$

Finally, $T_5 = \langle \boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}, \Phi \boldsymbol{n} - \phi \boldsymbol{n} \rangle_{\partial\Omega} = 0$, by the error equation (3.1d). As a consequence, we obtain

$$(\boldsymbol{\varepsilon}^{\boldsymbol{u}}, \boldsymbol{\theta})_{\mathcal{T}_h} = \nu(\mathrm{L}_h - \mathrm{L}, \Pi\,\Phi - \Phi)_{\mathcal{T}_h} + \nu(\Pi\,\mathrm{L} - \mathrm{L}, \Phi - \mathrm{P}_{k-1}\Phi)_{\mathcal{T}_h}.$$

This completes the proof. □

Step 6: Estimate of the velocity trace. Here we obtain the following simple estimate.

Lemma 3.7. *We have*

$$\|\boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}\|_h \le C\left(h\,\|\Pi\,\mathrm{L} - \mathrm{L}\,\|_\Omega + \|\,\boldsymbol{\varepsilon}^{\boldsymbol{u}}\,\|_\Omega\right).$$

Proof. From the first error equation (3.1a), we have that,

$$\langle \boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}, \mathrm{G}\boldsymbol{n} \rangle_{\partial K} = -(\Pi\,\mathrm{L} - \mathrm{L}, \mathrm{G})_K + (\mathrm{E}^{\mathrm{L}}, \mathrm{G})_K + (\boldsymbol{\varepsilon}^{\boldsymbol{u}}, \boldsymbol{\nabla}\cdot \mathrm{G})_K,$$

for all $\mathrm{G} \in \mathrm{P}_k(K)$. Hence, by a now standard scaling argument (see [3]) we readily obtain that

$$h_K^{1/2}\|\,\boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}\,\|_{\partial K} \le C\left(h_K\,\|\Pi\,\mathrm{L} - \mathrm{L}\,\|_K + h_K\,\|\,\mathrm{E}^{\mathrm{L}}\,\|_K + \|\,\boldsymbol{\varepsilon}^{\boldsymbol{u}}\,\|_K\right),$$

and the estimate follows by using Proposition 3.2. This completes the proof of Lemma 3.7. □

Step 7: Proof of Theorem 2.4. Here, we complete the proof the Theorem 2.4. To do that, we begin by gathering straightforward consequences of the lemmas obtained in the previous steps in the following result.

Corollary 3.8. *Suppose that the assumptions on the stabilization tensor* S *of Theorem* 2.1 *hold. Then*

$$\|\,\mathrm{E}^{\mathrm{L}}\,\|_\Omega \le \|\,\Pi\,\mathrm{L} - \mathrm{L}\,\|_\Omega,$$
$$\|\varepsilon^p - \overline{(\Pi p - p)}\|_\Omega \le C\,\Psi(\mathrm{S})\,\nu\,\|\Pi\,\mathrm{L} - \mathrm{L}\|_\Omega,$$

where $\Psi(\mathrm{S})$ *is defined in Proposition* 3.4. *Moreover, if the elliptic regularity estimate* (2.12) *holds for* $s = 0$, *we have*

$$\|\,\boldsymbol{\varepsilon}^{\boldsymbol{u}}\,\|_\Omega + \|\,\boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}\,\|_h \le\; C\,H(\mathrm{S})\,\|\,\Pi\,\mathrm{L} - \mathrm{L}\,\|_\Omega,$$

where

$$H(\mathrm{S}) :=\; \max\left\{ h, \nu \sup_{\boldsymbol{\theta} \in \boldsymbol{L}^2(\Omega)\setminus\{\boldsymbol{0}\}} \frac{\|\,\Pi\,\Phi - \Phi\,\|_\Omega + \|\,\Phi - \mathrm{P}_{k-1}\Phi\,\|_\Omega}{\|\,\boldsymbol{\theta}\,\|_\Omega} \right\},$$

It remains to estimate Ψ and H. This can be done by using the approximation properties of the projection Π_h and the regularity estimate (2.12). We are going to use some auxiliary projections that are introduced next.

Given a function $z \in H^1(K)$, and any given face F of K, we define $P_F z$ as the polynomial in $\mathcal{P}_k(K)$ given by

$$(P_F z, q)_K = (z, q)_K \qquad \forall\, q \in \mathcal{P}_{k-1}(K), \tag{3.3a}$$

$$\langle P_F z, \mu\rangle_F = \langle z, \mu\rangle_F \qquad \forall\, \mu \in \mathcal{P}_k(F). \tag{3.3b}$$

For a vector-valued function $\boldsymbol{z} \in \boldsymbol{H}^1(K)$, we denote by $\boldsymbol{P}_F \boldsymbol{z}$ as the vector-valued function whose i-th component is P_F applied to the i-th component of $\boldsymbol{z}$. Finally, for a matrix-valued function $\mathrm{Z} \in \mathrm{H}^1(K)$, we define $\mathrm{P}_F\mathrm{Z}$ in a similar way. To emphasize the dependence of the projection, for example, P_F on the simplex K, we write $P_{F,K}$.

Proposition 3.9. *Under the assumptions of Theorem* 2.4*, we have that*

$$\Psi(\mathrm{S}) \le \max\Big\{1, \max_{K\in\mathcal{T}_h}\{\Lambda_K^{\max} h_K/\nu\}^{1/2}\Big\},$$
$$H(\mathrm{S}) \le C\, h^{\min\{k,1\}} \max_{K\in\mathcal{T}_h}\left\{\frac{\Lambda_K^{\max}}{\lambda_K^{\min}}\Big(1+\frac{\Lambda_K^{\max} h_K}{\nu}\Big)\right\},$$

We also have $\Psi(\mathrm{S}_{SF}) \le C$ *and* $H(\mathrm{S}_{SF}) \le C\,\nu\, h^{\min\{k,1\}}$.

Proof. We begin by noting that

$$\Big((\mathsf{P}_M\boldsymbol{w})|_{\partial K}\Big)|_F = (\boldsymbol{P}_F\boldsymbol{w})|_F, \qquad \forall F\in\mathcal{E}(K), \qquad \forall K\in\mathcal{T}_h.$$

Using this identity plus elementary arguments (finite dimensionality of the discrete spaces, scaling arguments and optimality of the projections), the following chain of inequalities can be easily verified in each element $K \in \mathcal{T}_h$:

$$\begin{aligned}
\langle \mathrm{S}(\boldsymbol{P}\boldsymbol{w}-\mathsf{P}_M\boldsymbol{w}), \boldsymbol{P}\boldsymbol{w}-\mathsf{P}_M\boldsymbol{w}\rangle_{\partial K} &\le C\Lambda_K^{\max}\|\boldsymbol{P}\boldsymbol{w}-\mathsf{P}_M\boldsymbol{w}\|_{\partial K}^2\\
&= C\Lambda_K^{\max}\sum_{F\in\mathcal{E}(K)}\|\boldsymbol{P}\boldsymbol{w}-\boldsymbol{P}_F\boldsymbol{w}\|_F^2\\
&\le C\Lambda_K^{\max}\sum_{F\in\mathcal{E}(K)}h_K^{-1}\|\boldsymbol{P}\boldsymbol{w}-\boldsymbol{P}_F\boldsymbol{w}\|_K^2\\
&\le C\Lambda_K^{\max}h_K\|\boldsymbol{w}\|_{\boldsymbol{H}^1(K)}^2.
\end{aligned}$$

The bound for $\Psi(\mathrm{S})$ is an easy consequence of this.

We now proceed to bound $H(\mathrm{S})$. To abbreviate some later expressions we write $\underline{k} := \min\{k,1\}$. Using the regularity assumption (2.12) with $s=0$ we can prove that

$$\nu\|\Phi-\mathrm{P}_{k-1}\Phi\|_\Omega \le Ch^{\underline{k}}\nu\|\Phi\|_{\mathrm{H}^{\underline{k}}(\Omega)} \le Ch^{\underline{k}}\|\boldsymbol{\theta}\|_\Omega.$$

Applying the two bounds of Theorem 2.1 to the projection of the solution of the dual problem (2.11) with $\ell_{\boldsymbol{\sigma}}=0$ and $\ell_{\boldsymbol{u}}=\underline{k}$, we can prove that for all $K\in\mathcal{T}_h$,

$$\begin{aligned}
\nu\|\Phi-\Pi\Phi\|_K \le& C\Big(h_K|\nu\Phi-\phi\,\mathrm{I}|_{\mathrm{H}^1(K)} + h^{\underline{k}+1}\Lambda_K^{\max}|\boldsymbol{\phi}|_{\boldsymbol{H}^{\underline{k}+1}(K)} + \Lambda_K^{\max}\|\boldsymbol{\Pi}\boldsymbol{\phi}-\boldsymbol{\phi}\|_K\Big)\\
\le& Ch_K\Big(\nu|\Phi|_{\mathrm{H}^1(K)} + |\phi|_{H^1(K)}\Big)\\
&+ Ch_K^{\underline{k}+1}\frac{\Lambda_K^{\max}}{\nu}\Big(1+\frac{\Lambda_K^{\max}}{\lambda_K^{\min}}\Big)\nu|\boldsymbol{\phi}|_{\boldsymbol{H}^{\underline{k}+1}(K)} + Ch_K\frac{\Lambda_K^{\max}}{\lambda_K^{\min}}\|\boldsymbol{\theta}\|_K.
\end{aligned}$$

To finish the proof for the general stabilization tensor S we only have to sum over the elements of the triangulation, to overestimate $h_K \le h_K^{\underline{k}}$, to reorder terms using the fact that $\lambda_K^{\min} \le \Lambda_K^{\max}$ and to apply (2.12) with $s=0$.

The proof for the case $\mathrm{S} = \mathrm{S}_{SF}$ uses the same techniques, but specializes the bound for Ψ to sum over a single face per element and employs Theorem 2.2 instead of the more general Theorem 2.1 to bound H. $\square$

Theorem 2.4 easily follows from the two above results.

Step 8: Analysis of the postprocessed velocity. We prove the estimate of Theorem 2.5 for the three-dimensional case; we omit the proof for the two-dimensional case since it is similar and much simpler. We proceed as follows.

We have already discussed the fact that $\boldsymbol{u}_h^\star$ is well defined as a consequence of the new characterization of the BDM projection given in the appendix. The fact that $\boldsymbol{u}_h^\star$ belongs to $\boldsymbol{H}(\mathrm{div},\Omega)$ is a direct consequence of the fact that, by definition, (see equations (2.9a) and (2.9b)), its normal component depends only on the single-valued function $\widehat{\boldsymbol{u}}_h$ and on $\{\!\{\mathrm{L}_h^t \boldsymbol{n}\}\!\}$, which is also single valued.

The fact that $\boldsymbol{u}_h^\star$ is divergence-free can be seen as follows. For any $q \in P_h$, we have

$$\begin{aligned}(\nabla\cdot \boldsymbol{u}_h^\star, q)_\Omega &= (\nabla\cdot \boldsymbol{u}_h^\star, q)_{\mathcal{T}_h}\\ &= -\,(\boldsymbol{u}_h^\star, \boldsymbol{\nabla}\, q)_{\mathcal{T}_h} + \langle \boldsymbol{u}_h^\star\cdot \boldsymbol{n}, q\rangle_{\partial\mathcal{T}_h}\\ &= -\,(\boldsymbol{u}_h, \boldsymbol{\nabla}\, q)_{\mathcal{T}_h} + \langle \widehat{\boldsymbol{u}}_h\cdot \boldsymbol{n}, q\rangle_{\partial\mathcal{T}_h} && \text{by equations (2.9a) and (2.9c),}\\ &= 0, && \text{by equation (2.2c).}\end{aligned}$$

Thus, $\boldsymbol{u}_h^\star$ is divergence-free.

It remains to prove the error estimate. To do that, we proceed as follows. Let $\boldsymbol{\Pi}^{\mathrm{BDM}}\boldsymbol{u}$ be the BDM projection of $\boldsymbol{u}$ into $\boldsymbol{\mathcal{P}}_{k+1}(K)$ and set $\boldsymbol{\delta} := \boldsymbol{u}_h^\star - \boldsymbol{\Pi}^{\mathrm{BDM}}\boldsymbol{u}$. Then, by the equations defining $\boldsymbol{u}_h^\star$ (2.9) and those defining the BDM projection as given in Proposition A.1, we readily have that

$$\begin{aligned}\langle \boldsymbol{\delta}\cdot \boldsymbol{n}, \mu\rangle_F &= \langle (\widehat{\boldsymbol{u}}_h - \boldsymbol{u})\cdot \boldsymbol{n}, \mu\rangle_F && \forall\, \mu \in \mathcal{P}_k(F),\\ \langle \boldsymbol{\delta}\cdot \boldsymbol{n}, \mu\rangle_F &= \langle \boldsymbol{u}_h^\star\cdot \boldsymbol{n} - \boldsymbol{u}\cdot\boldsymbol{n}, \mu\rangle_F && \forall\, \mu \in \mathcal{P}_{k+1}(F)^\perp,\end{aligned}$$

for all faces F of K, and that

$$\begin{aligned}(\boldsymbol{\delta}, \boldsymbol{\nabla}\, w)_K &= (\boldsymbol{u}_h - \boldsymbol{u}, \boldsymbol{\nabla}\, w)_K && \forall\, w \in \boldsymbol{\mathcal{P}}_k(K),\\ (\boldsymbol{\nabla}\times \boldsymbol{\delta}, (\boldsymbol{\nabla}\times \boldsymbol{v})\mathrm{B}_K)_K &= (\mathbf{w}_h - \boldsymbol{\nabla}\times \boldsymbol{u}, (\boldsymbol{\nabla}\times \boldsymbol{v})\mathrm{B}_K)_K && \forall\, \boldsymbol{v} \in \boldsymbol{\mathcal{S}}_k(K).\end{aligned}$$

We will use two new projections: P_∂ denotes the projection that is defined face by face as the $L^2(F)$-projection onto $\mathcal{P}_k(F)$; $\mathsf{P}_\partial^\perp$ is defined similarly, but with local image in $\mathcal{P}_{k+1}(F)^\perp$. Obviously, $\mathsf{P}_\partial + \mathsf{P}_\partial^\perp$ projects face by face onto the space $\mathcal{P}_{k+1}(F)$. Inserting the definition of the projection of the errors, we get

$$\begin{aligned}\langle \boldsymbol{\delta}\cdot \boldsymbol{n}, \mu\rangle_F &= \langle -\boldsymbol{\varepsilon}^{\widehat{\boldsymbol{u}}}\cdot \boldsymbol{n}, \mu\rangle_F && \forall\, \mu \in \mathcal{P}_k(F),\\ \langle \boldsymbol{\delta}\cdot \boldsymbol{n}, \mu\rangle_F &= \langle \mathsf{P}_\partial^\perp(\boldsymbol{u}^\star\cdot \boldsymbol{n} - \boldsymbol{u}\cdot\boldsymbol{n}), \mu\rangle_F && \forall\, \mu \in \mathcal{P}_{k+1}(F)^\perp,\end{aligned}$$

for all faces F of K, and

$$\begin{aligned}(\boldsymbol{\delta}, \boldsymbol{\nabla}\, w)_K &= (\boldsymbol{\varepsilon}^{\boldsymbol{u}}, \boldsymbol{\nabla}\, w)_K && \forall\, w \in \boldsymbol{\mathcal{P}}_k(K),\\ (\boldsymbol{\nabla}\times \boldsymbol{\delta}, (\boldsymbol{\nabla}\times \boldsymbol{v})\,\mathrm{B}_K)_K &= (\mathbf{w}_h - \boldsymbol{\nabla}\times \boldsymbol{u}, (\boldsymbol{\nabla}\times \boldsymbol{v})\,\mathrm{B}_K)_K && \forall\, \boldsymbol{v} \in \boldsymbol{\mathcal{S}}_k(K).\end{aligned}$$

These equations determine $\boldsymbol{\delta}$ since they provide the degrees of freedom of the BDM projection, as given in Proposition A.1. So, a scaling argument can now be used to prove that

$$\|\boldsymbol{\delta}\|_K \leq \ C\,(h_K^{1/2}\|\boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}\|_{\partial K} + h_K^{1/2}\|\mathsf{P}_\partial^\perp(\boldsymbol{u}_h^\star\cdot\boldsymbol{n} - \boldsymbol{u}\cdot\boldsymbol{n})\|_{\partial K} \\ + \|\boldsymbol{\varepsilon}^{\boldsymbol{u}}\|_K + h_K\|\mathbf{w}_h - \boldsymbol{\nabla}\times\boldsymbol{u}\|_K).$$

Next, we estimate $\zeta := \mathsf{P}_\partial^\perp(\boldsymbol{u}_h^\star\cdot\boldsymbol{n} - \boldsymbol{u}\cdot\boldsymbol{n})$.

To do that, we begin by noting that

$$\boldsymbol{n}\times(\,\{\!\{\mathrm{L}^t\}\!\}\boldsymbol{n}) = \boldsymbol{n}\times((\boldsymbol{\nabla}\,\boldsymbol{u})^t\boldsymbol{n}) = (\boldsymbol{n}\times\boldsymbol{\nabla})(\boldsymbol{u}\cdot\boldsymbol{n})$$

and by the second equation defining the postprocessed velocity, (2.9b), we have that

$$\langle(\boldsymbol{n}\times\boldsymbol{\nabla})(\boldsymbol{u}_h^\star\cdot\boldsymbol{n} - \boldsymbol{u}\cdot\boldsymbol{n}), (\boldsymbol{n}\times\boldsymbol{\nabla})\mu\rangle_F = \langle\boldsymbol{n}\times(\,\{\!\{\mathrm{L}_h^t - \mathrm{L}^t\}\!\}\boldsymbol{n}), (\boldsymbol{n}\times\boldsymbol{\nabla})\mu\rangle_F,$$

for all $\mu\in\mathcal{P}_{k+1}(F)^\perp$ and all faces F of K. Then, we note that

$$\begin{aligned}\boldsymbol{u}_h^\star\cdot\boldsymbol{n} - \boldsymbol{u}\cdot\boldsymbol{n} &= (\mathsf{P}_\partial + \mathsf{P}_\partial^\perp)(\boldsymbol{u}_h^\star\cdot\boldsymbol{n} - \boldsymbol{u}\cdot\boldsymbol{n}) + (\mathrm{I} - \mathsf{P}_\partial - \mathsf{P}_\partial^\perp)(\boldsymbol{u}_h^\star\cdot\boldsymbol{n} - \boldsymbol{u}\cdot\boldsymbol{n})\\ &= (\mathsf{P}_\partial + \mathsf{P}_\partial^\perp)(\boldsymbol{u}_h^\star\cdot\boldsymbol{n} - \boldsymbol{u}\cdot\boldsymbol{n}) - (\mathrm{I} - \mathsf{P}_\partial - \mathsf{P}_\partial^\perp)(\boldsymbol{u}\cdot\boldsymbol{n})\\ &= \mathsf{P}_\partial(\boldsymbol{u}_h^\star\cdot\boldsymbol{n} - \boldsymbol{u}\cdot\boldsymbol{n}) + \zeta - (\mathrm{I} - \mathsf{P}_\partial - \mathsf{P}_\partial^\perp)(\boldsymbol{u}\cdot\boldsymbol{n})\\ &= \boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}} + \zeta - (\mathrm{I} - \mathsf{P}_\partial - \mathsf{P}_\partial^\perp)(\boldsymbol{u}\cdot\boldsymbol{n}),\end{aligned}$$

by the first equation defining the postprocessed velocity, (2.9a), and the definition of the projection of the error $\boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}$. So, setting $\mathsf{e}_{\boldsymbol{u}} := -(\mathrm{I} - \mathsf{P}_\partial - \mathsf{P}_\partial^\perp)(\boldsymbol{u}\cdot\boldsymbol{n})$, we see that ζ is the element of $\mathcal{P}_{k+1}(K)^\perp$ satisfying

$$\begin{aligned}\langle(\boldsymbol{n}\times\boldsymbol{\nabla})\zeta, (\boldsymbol{n}\times\boldsymbol{\nabla})\mu\rangle_F = &-\langle(\boldsymbol{n}\times\boldsymbol{\nabla})(\boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}} + \mathsf{e}_{\boldsymbol{u}}), (\boldsymbol{n}\times\boldsymbol{\nabla})\mu\rangle_F\\ &+ \langle\boldsymbol{n}\times(\,\{\!\{\mathrm{L}_h^t - \mathrm{L}^t\}\!\}\boldsymbol{n}), (\boldsymbol{n}\times\boldsymbol{\nabla})\mu\rangle_F\end{aligned}$$

for all $\mu\in\mathcal{P}_{k+1}(F)^\perp$ and all faces F of K.

So, if we pick the face $F = \partial K^+\cap\partial K^-$ and set $\{\!\{\mathrm{P}_F\mathrm{L}\}\!\} := (\mathrm{P}_{F,K^+}\mathrm{L} + \mathrm{P}_{F,K^-}\mathrm{L})/2$, where $\mathrm{P}_{F,K^\pm}$ are the matrix-valued versions of the projections defined in (3.3) (the $\pm$ sign depends on which triangle we work on), we see that we have

$$\begin{aligned}\langle(\boldsymbol{n}\times\boldsymbol{\nabla})\zeta, (\boldsymbol{n}\times\boldsymbol{\nabla})\mu\rangle_F = &-\langle(\boldsymbol{n}\times\boldsymbol{\nabla})(\boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}} + \mathsf{e}_{\boldsymbol{u}}), (\boldsymbol{n}\times\boldsymbol{\nabla})\mu\rangle_F\\ &+ \langle\boldsymbol{n}\times(\,\{\!\{\mathrm{L}_h^t - P_F\mathrm{L}^t\}\!\}\boldsymbol{n}), (\boldsymbol{n}\times\boldsymbol{\nabla})\mu\rangle_F.\end{aligned}$$

Taking $\mu := \zeta$ and applying the Cauchy-Schwarz inequality and inverse estimates on the faces, we get that

$$\begin{aligned}\|(\boldsymbol{n}\times\boldsymbol{\nabla})\,\zeta\|_F &\leq C\,h_{K^+}^{-1}(\|\boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}\|_F + \|\mathsf{e}_{\boldsymbol{u}}\|_F) + C\,\|\,\{\!\{\mathrm{L}_h - P_F\mathrm{L}\}\!\}\|_F\\ &\leq C\,h_{K^+}^{-1}(\|\boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}\|_F + \|\mathsf{e}_{\boldsymbol{u}}\|_F) + C\,h_{K^+}^{-1/2}\|\mathrm{L}_h - P_{F,K^+}\mathrm{L}\|_{K^+}\\ &\quad + C\,h_{K^-}^{-1/2}\|\mathrm{L}_h - P_{F,K^-}\mathrm{L}\|_{K^-}.\end{aligned}$$

Finally, by an inverse inequality,

$$\begin{aligned}h_{k^+}^{1/2}\,\|\zeta\|_F &\leq C\,h_{K^+}^{3/2}\,\|(\boldsymbol{n}\times\boldsymbol{\nabla})\,\zeta\|_F\\ &\leq C\,h_{K^+}^{1/2}(\|\boldsymbol{\varepsilon}^{\hat{\boldsymbol{u}}}\|_F + \|\mathsf{e}_{\boldsymbol{u}}\|_F) + C\,h_{K^+}\|\mathrm{L}_h - P_{F,K^+}\mathrm{L}\|_{K^+}\\ &\quad + C\,h_{K^-}\|\mathrm{L}_h - P_{F,K^-}\mathrm{L}\|_{K^-},\end{aligned}$$

since h_{K^+}/h_{K^-} is uniformly bounded.

This implies that

$$\|\boldsymbol{\delta}\|_K \leq C\,(h_K^{1/2}\|\boldsymbol{\varepsilon}^{\hat{u}}\|_{\partial K} + h_K^{1/2}\|\mathsf{e}_{\boldsymbol{u}}\|_{\partial K} + \|\boldsymbol{\varepsilon}^{\boldsymbol{u}}\|_K + \sum_{K'\in P(K)} h_{K'}\|\mathrm{L}_h - \mathrm{L}\|_{K'}$$
$$+ \sum_{F=\partial K\cap\partial K'} (h_K\,\|\mathrm{L} - P_{F,K}\mathrm{L}\|_{K'}) + h_{K'}\,\|\mathrm{L} - P_{F,K'}\mathrm{L}\|_{K'}),$$

where $P(K)$ is the set of simplexes K' sharing a face with K. Hence, by the approximation properties of the BDM projection and those of the projections $\mathrm{P}_{F,K}$, we get

$$\|\boldsymbol{u}_h^\star - \boldsymbol{u}\|_\Omega \leq C\,h^{\ell_{\boldsymbol{u}}+2}|\,\boldsymbol{u}\,|_{\boldsymbol{H}^{\ell_{\boldsymbol{u}}+2}(\mathcal{T}_h)} + C(\|\boldsymbol{\varepsilon}^{\hat{u}}\|_h + \|\boldsymbol{\varepsilon}^{\boldsymbol{u}}\|_\Omega + h\|\mathrm{L}_h - \mathrm{L}\|_\Omega),$$

and the estimate follows by using the estimates of Theorem 2.4. This completes the proof of Theorem 2.5.

4. Proof of the approximation properties of $\mathit{\Pi}_h$

In this section, we give a detailed proof of the approximation properties of the auxiliary projection $\mathit{\Pi}_h$ contained in Theorem 2.1.

To do that, we follow an approach similar to the one proposed in [14] for HDG methods for symmetric, second-order elliptic problems. We are thus going to proceed in several steps. We introduce the sets $\boldsymbol{\mathcal{P}}_k(K)^\perp := [\mathcal{P}_k(K)^\perp]^n$ and $\mathrm{P}_k(K)^\perp := [\mathcal{P}_k(K)^\perp]^{n\times n}$, where

$$\mathcal{P}_k(K)^\perp := \{w \in \mathcal{P}_k(K) : \ (w,\zeta)_K = 0 \quad \forall\,\zeta \in \mathcal{P}_{k-1}(K)\},$$

and state the following simple but useful auxiliary result.

Lemma 4.1 ([14])**.** *For all $p \in \mathcal{P}_k(K)^\perp$, we have that*

$$\|p\|_K \leq C\,h_K^{1/2}\|p\|_F \quad \textit{for any face } F \subseteq \partial K.$$

Step 1: Existence and uniqueness of the projection. Let us begin by proving that the projection $\mathit{\Pi}_h$ is well defined.

Let us first count the number of independent equations provided by its definition. We have that the number of equations is

$$\begin{aligned}
\dim(\mathrm{P}_{k-1}(K)) &\qquad \text{for (2.10a)},\\
\dim(\boldsymbol{\mathcal{P}}_{k-1}(K)) &\qquad \text{for (2.10b)},\\
\dim(\mathcal{P}_{k-1}(K)) &\qquad \text{for (2.10c)},\\
(n+1)\dim(\boldsymbol{\mathcal{P}}_k(F)) &\qquad \text{for (2.10e)},
\end{aligned}$$

and noting that the equations (2.10a) imply the equations (2.10d) for any $q \in \mathcal{P}_{k-1}(K)$, we get

$$\dim(\mathcal{P}_k(K)) - \dim(\mathcal{P}_{k-1}(K)) \qquad \text{for (2.10d)}.$$

On the other hand, the number of unknowns is

$$\begin{aligned}
\dim(\mathrm{P}_k(K)) &\qquad \text{for } \Pi\,\mathrm{L},\\
\dim(\boldsymbol{\mathcal{P}}_k(K)) &\qquad \text{for } \boldsymbol{\mathit{\Pi} u},\\
\dim(\mathcal{P}_k(K)) &\qquad \text{for } \mathit{\Pi} p.
\end{aligned}$$

Thus, if δ is the total number of equations minus the total number of unknowns, we have

$$\begin{aligned}\delta = \; & \dim(\mathrm{P}_{k-1}(K)) + \dim(\boldsymbol{\mathcal{P}}_{k-1}(K)) + (n+1)\dim(\boldsymbol{\mathcal{P}}_k(F)) \\ & - \dim(\mathrm{P}_k(K)) - \dim(\boldsymbol{\mathcal{P}}_k(K)) \\ = \; & (n+1)\, n \left(\dim(\mathcal{P}_{k-1}(K)) - \dim(\mathcal{P}_k(K)) + \dim(\mathcal{P}_k(F))\right) \\ = \; & 0.\end{aligned}$$

This implies that the existence of the projection follows from its uniqueness. Thus, it suffices to show that when the right-hand sides of (2.10) vanish, then $\Pi\,\mathrm{L}$, $\boldsymbol{\Pi u}$ and Πp also vanish. Since this immediately follows from the approximation estimates of Theorem 2.1, the proof of the existence and uniqueness of the projection Π_h is complete.

Step 2: Characterization of $\boldsymbol{\Pi u}$. We begin by providing a new characterization of $\boldsymbol{\Pi u}$.

Proposition 4.2. *Suppose that the assumptions on the stabilization tensor* S *of Theorem* 2.1 *hold. Then, on each simplex $K \in \mathcal{T}_h$, $\boldsymbol{\Pi u}$ is the only element of $\boldsymbol{\mathcal{P}}_k(K)$ such that*

$$(\boldsymbol{\Pi u}, \boldsymbol{v})_K = (\boldsymbol{u}, \boldsymbol{v})_K \qquad \forall\, \boldsymbol{v} \in \boldsymbol{\mathcal{P}}_{k-1}(K), \tag{4.1a}$$

$$\langle \mathrm{S}\boldsymbol{\Pi u}, \boldsymbol{v}\rangle_{\partial K} = -(\boldsymbol{\nabla\cdot}(\nu\mathrm{L} - p\mathrm{I}), \boldsymbol{v})_K + \langle \mathrm{S}\boldsymbol{u}, \boldsymbol{v}\rangle_{\partial K} \qquad \forall\, \boldsymbol{v} \in \boldsymbol{\mathcal{P}}_k(K)^{\perp}. \tag{4.1b}$$

Proof. Let us begin by noting that, from the equations (2.10b) and (2.10e) defining the projection Π_h, we have that

$$\begin{aligned}(\boldsymbol{\Pi u}, \boldsymbol{v})_K &= (\boldsymbol{u}, \boldsymbol{v})_K && \forall \boldsymbol{v} \in \boldsymbol{\mathcal{P}}_{k-1}(K), \\ \langle \mathrm{S}\boldsymbol{\Pi u}, \boldsymbol{v}\rangle_{\partial K} &= \langle -\nu(\mathrm{L} - \Pi\,\mathrm{L})\boldsymbol{n} + (p - \Pi p)\boldsymbol{n} + \mathrm{S}\boldsymbol{u}, \boldsymbol{v}\rangle_{\partial K} && \forall \boldsymbol{v} \in \boldsymbol{\mathcal{P}}_k(K)^{\perp}.\end{aligned}$$

But

$$\begin{aligned}\Theta &:= \langle -\nu(\mathrm{L} - \Pi\,\mathrm{L})\boldsymbol{n} + (p - \Pi p)\boldsymbol{n}, \boldsymbol{v}\rangle_{\partial K} \\ &= -\,\nu(\boldsymbol{\nabla\cdot}(\mathrm{L} - \Pi\,\mathrm{L}, \boldsymbol{v})_K - \nu\,(\mathrm{L} - \Pi\,\mathrm{L}, \nabla\,\boldsymbol{v})_K \\ &\quad + (\boldsymbol{\nabla}(p - \Pi p), \boldsymbol{v})_K + (p - \Pi p, \nabla\cdot\boldsymbol{v})_K \\ &= -\,\nu\,(\boldsymbol{\nabla\cdot}(\mathrm{L} - \Pi\,\mathrm{L}), \boldsymbol{v})_K + (\boldsymbol{\nabla}(p - \Pi p), \boldsymbol{v})_K \quad \text{by (2.10a) and (2.10c),} \\ &= \; -\,(\boldsymbol{\nabla\cdot}(\nu\,\mathrm{L} - p\mathrm{I}), \boldsymbol{v})_K\end{aligned}$$

since $\boldsymbol{v} \in \boldsymbol{\mathcal{P}}_k(K)^{\perp}$.

To prove that $\boldsymbol{\Pi u}$ is well defined it is enough to show that if $\boldsymbol{u} = \boldsymbol{0}$, $\mathrm{L} = 0$ and $p = 0$, the only solution, $\boldsymbol{\Pi u}$, is the trivial one. But since we have

$$\begin{aligned}(\boldsymbol{\Pi u}, \boldsymbol{v})_K &= 0 && \forall\, \boldsymbol{v} \in \boldsymbol{\mathcal{P}}_{k-1}(K), \\ \langle \mathrm{S}\boldsymbol{\Pi u}, \boldsymbol{v}\rangle_{\partial K} &= 0 && \forall\, \boldsymbol{v} \in \boldsymbol{\mathcal{P}}_k(K)^{\perp},\end{aligned}$$

we can apply Lemma 4.4 with $\boldsymbol{p} := \boldsymbol{\Pi u}$ and $b := 0$ to conclude that $\boldsymbol{\Pi u} = \boldsymbol{0}$. This completes the proof. □

Step 3: Two auxiliary results. To obtain the approximation properties of $\boldsymbol{\Pi}$, we are going to use two auxiliary results. The first result is the following.

Lemma 4.3. *Let $\mathcal{B}_K := \{\boldsymbol{\omega}_{F,i} : (F,i) \in \mathcal{I}_K\}$ be a basis of $\mathbb{R}^n$ and let C_K be the bound for its dual basis in* (2.4c). *Then, for all $\boldsymbol{p} \in \boldsymbol{\mathcal{P}}_k(K)^\perp$, we have that*

$$\|\boldsymbol{p}\|_K \le C\, h_K^{1/2} C_K \sum_{i=1}^{n} \|\boldsymbol{p}\cdot\boldsymbol{\omega}_{F,i}\|_F.$$

Proof. We can write $\boldsymbol{p} = \sum_{(F,i)\in\mathcal{I}} (\boldsymbol{p}\cdot\boldsymbol{\omega}_{F,i})\, \boldsymbol{\omega}^\star_{F,i}$ where $\{\boldsymbol{\omega}^\star_{F,i} : (F,i) \in \mathcal{I}_K\}$ is the dual basis of $\mathcal{B}_K$. This implies that

$$\|\boldsymbol{p}\|_K \le \; C_K \sum_{(F,i)\in\mathcal{I}_K} \|\boldsymbol{p}\cdot\boldsymbol{\omega}_{F,i}\|_K,$$

and the result follows by applying Lemma 4.1 with $p := \boldsymbol{p}\cdot\boldsymbol{\omega}_{F,i}$. This completes the proof. $\square$

The second auxiliary result involves the stabilization tensor S.

Lemma 4.4. *Let the stabilization tensor* S *satisfy the conditions* (2.3)–(2.4) *and let $\boldsymbol{p} \in \boldsymbol{\mathcal{P}}_k(K)^\perp$ satisfy the equation*

$$\langle \mathrm{S}\boldsymbol{p}, \boldsymbol{v}\rangle_{\partial K} = b(\boldsymbol{v}) \qquad\qquad \forall\, \boldsymbol{v} \in \boldsymbol{\mathcal{P}}_k(K)^\perp,$$

where $b(\cdot)$ is a continuous linear functional on $\boldsymbol{\mathcal{P}}_k(K)^\perp$. Then

$$\|\boldsymbol{p}\|_K \le C\, \frac{h_K}{\lambda_K^{\min}}\, \|b\|,$$

where $\|b\| := \sup_{\boldsymbol{v}\in\boldsymbol{\mathcal{P}}_k(K)^\perp\setminus\boldsymbol{0}} b(\boldsymbol{v})/\|\boldsymbol{v}\|_K$.

Proof. Taking $\boldsymbol{v} := \boldsymbol{p}$, we obtain $\langle \mathrm{S}\boldsymbol{p}, \boldsymbol{p}\rangle_{\partial K} = b(\boldsymbol{p}) \le \|\, b\,\|\, \|\boldsymbol{p}\|_K$. On the other hand, by the conditions (2.3) on the stabilization tensor S, we can write that

$$\begin{aligned}
\langle \mathrm{S}\boldsymbol{p}, \boldsymbol{p}\rangle_{\partial K} &= \sum_F \sum_{i=1}^{n} \lambda_{F,i}\, \|\boldsymbol{p}\cdot\boldsymbol{\omega}_{F,i}\|_F^2 \\
&\ge \lambda_K^{\min} \sum_{(F,i)\in\mathcal{I}} \|\boldsymbol{p}\cdot\boldsymbol{\omega}_{F,i}\|_F^2 \qquad \text{since } \lambda_{F,i} \ge 0 \text{ by (2.3c)}, \\
&\ge C\, \lambda_K^{\min}\;\, 2h_K^{-1}\, \|\boldsymbol{p}\|_K^2,
\end{aligned}$$

by Lemma 4.3. $\square$

Step 4: Estimate of $\boldsymbol{\Pi u}-\boldsymbol{u}$. We are now ready to obtain the estimate of $\boldsymbol{\Pi u}-\boldsymbol{u}$ in Theorem 2.1.

Lemma 4.5. *Suppose that the assumptions on the stabilization tensor* S *of Theorem* 2.1 *hold. Then, we have*

$$\|\boldsymbol{\Pi u} - \boldsymbol{u}\|_K \le C\, \frac{\Lambda_K^{\max}}{\lambda_K^{\min}}\; h^{\ell_{\boldsymbol{u}}+1} |\boldsymbol{u}|_{\boldsymbol{H}^{\ell_{\boldsymbol{u}}+1}(K)} + C\, \frac{h_K^{\ell_\sigma+1}}{\lambda_K^{\min}} |\,\boldsymbol{\nabla\cdot}(\nu\mathrm{L} - p\mathrm{I})|_{\boldsymbol{H}^{\ell_\sigma}(K)},$$

for $\ell_\sigma, \ell_{\boldsymbol{u}} \in [0,k]$.

Proof. To prove the result, we set $\delta^{\boldsymbol{u}} := \boldsymbol{\Pi u} - \boldsymbol{u}_k$, where $\boldsymbol{u}_k$ is the L^2-projection of $\boldsymbol{u}$ into $\boldsymbol{\mathcal{P}}_k(K)$, and note that

$$\|\boldsymbol{\Pi u} - \boldsymbol{u}\|_K \leq \|\boldsymbol{u} - \boldsymbol{u}_k\|_K + \|\delta^{\boldsymbol{u}}\|_K.$$

The first term can be readily estimated by using the approximation properties of the L^2-projection. Let us estimate the second term.

To do that, we note that, by the equation (4.1a), $\delta^{\boldsymbol{u}}$ belongs to $\boldsymbol{\mathcal{P}}_k(K)^\perp$ and, by the equation (4.1b), that it satisfies

$$\langle \mathrm{S}\,\delta^{\boldsymbol{u}}, \boldsymbol{v}\rangle_{\partial K} = b_\sigma(\boldsymbol{v}) + b_{\boldsymbol{u}}(\boldsymbol{v}) \qquad \forall \boldsymbol{v} \in \boldsymbol{\mathcal{P}}_k(K)^\perp,$$

where $b_\sigma(\boldsymbol{v}) := (-\,\boldsymbol{\nabla\cdot}(\nu\mathrm{L} - p\mathrm{I}), \boldsymbol{v})_K$, and $b_{\boldsymbol{u}}(\boldsymbol{v}) := \langle \mathrm{S}\,(\boldsymbol{u} - \boldsymbol{u}_k), \boldsymbol{v}\rangle_{\partial K}$. Thus, by Lemma 4.4 with $\boldsymbol{p} := \delta^{\boldsymbol{u}}$ and $b := b_\sigma + b_{\boldsymbol{u}}$, we get that

$$\|\delta^{\boldsymbol{u}}\|_K \leq C\,\frac{h_K}{\lambda_K^{\min}}\big(\|b_\sigma\| + \|b_{\boldsymbol{u}}\|\big).$$

It remains to estimate the norm of the linear forms b_σ and $b_{\boldsymbol{u}}$. But, since $\boldsymbol{v} \in \boldsymbol{\mathcal{P}}_k(K)^\perp$, we get

$$\|b_\sigma\| \leq C\, h_K^{\ell_\sigma}\,|\,\boldsymbol{\nabla\cdot}(\nu\mathrm{L} - p\mathrm{I})|_{\mathrm{H}^{\ell_\sigma}(K)},$$

for $\ell_\sigma \in [0,k]$. Finally, since $|b_{\boldsymbol{u}}(\boldsymbol{v})| \leq \Lambda_K^{\max}\,\|\boldsymbol{u} - \boldsymbol{u}_k\|_{\partial K}\|\boldsymbol{v}\|_{\partial K}$, we easily obtain that

$$|b_{\boldsymbol{u}}| \leq C\,\Lambda_K^{\max}\, h_K^{\ell_{\boldsymbol{u}}}|\boldsymbol{u}|_{\boldsymbol{H}^{\ell_{\boldsymbol{u}}+1}(K)},$$

for $\ell_{\boldsymbol{u}} \in [0,k]$, where $\Lambda_K^{\max}$ is given by (2.5). This completes the proof. $\square$

Step 5: Characterization of $\Pi\,\mathrm{L} - \mathit{\Pi} p\mathrm{I}$. We characterize $\Pi\,\mathrm{L} - \mathit{\Pi} p\mathrm{I}$ in terms of the projections defined in (3.3). We have the following characterization of the product $\Pi\,\mathrm{L} - \mathit{\Pi} p\mathrm{I}$ by the normal vector to any of the faces of K.

Proposition 4.6. *On the simplex K, we have*

$$(\nu\Pi\,\mathrm{L} - \mathit{\Pi} p\mathrm{I})\boldsymbol{n}_F = (\nu\mathrm{P}_F\mathrm{L} - P_F p\mathrm{I})\boldsymbol{n}_F + \mathrm{S}|_F(\boldsymbol{\Pi u} - \boldsymbol{P}_F\boldsymbol{u}),$$

for each face F of K.

Proof. Pick any face F of the simplex K and set

$$\boldsymbol{\delta} := (\nu\Pi\,\mathrm{L} - \mathit{\Pi} p\mathrm{I})\boldsymbol{n}_F - (\nu\mathrm{P}_F\mathrm{L} - P_F p\mathrm{I})\boldsymbol{n}_F - \mathrm{S}|_F(\boldsymbol{\Pi u} - \boldsymbol{P}_F\boldsymbol{u}).$$

By the condition on the stabilization tensor S (2.3a), we have that $\delta \in \mathcal{P}_k(K)$. Moreover, we claim that

$$\begin{aligned}(\boldsymbol{\delta}, \boldsymbol{w})_K &= 0 \qquad \forall\, \boldsymbol{w} \in \mathcal{P}_{k-1}(K),\\ \langle\boldsymbol{\delta}, \boldsymbol{\mu}\rangle_F &= 0 \qquad \forall\, \boldsymbol{\mu} \in \mathcal{P}_k(F).\end{aligned}$$

The first equation shows that $\boldsymbol{\delta} \in \boldsymbol{\mathcal{P}}_k(K)^\perp$, and the second implies that $\boldsymbol{\delta} = \boldsymbol{0}$ on K, by Lemma 4.1.

It remains to prove the claim. Let us prove the first equation. For any $\boldsymbol{v} \in \boldsymbol{\mathcal{P}}_{k-1}(K)$, we have that

$$(\boldsymbol{\delta}, \boldsymbol{w})_K = ((\nu\Pi\,\mathrm{L} - \mathit{\Pi} p\mathrm{I})\boldsymbol{n}_F - (\nu\mathrm{L} - p\mathrm{I})\boldsymbol{n}_F - \mathrm{S}|_F(\boldsymbol{\Pi u} - \boldsymbol{u}), \boldsymbol{w})_K,$$

by definition of $P_F, \boldsymbol{P}_F$ and P_F, (3.3), and by the fact that, by condition (2.3a), the tensor $\mathrm{S}|_F$ is constant. Hence

$$(\boldsymbol{\delta}, \boldsymbol{w})_K = (\nu(\Pi\,\mathrm{L} - \mathrm{L}), \boldsymbol{w} \otimes \boldsymbol{n}_F)_K - (\mathit{\Pi} p - p, \boldsymbol{n}_F\cdot\boldsymbol{w})_K - (\boldsymbol{\Pi u} - \boldsymbol{u}), \mathrm{S}|_F\boldsymbol{w})_K,$$

by the equations defining the projection $\mathit{\Pi}_h$, namely, (2.10a) with $\mathrm{G} := \boldsymbol{w} \otimes \boldsymbol{n}_F$, (2.10b) with $\boldsymbol{v} := \mathrm{S}|_F \boldsymbol{w}$, and (2.10c) with $q := \boldsymbol{n}_F \cdot \boldsymbol{w}$. The second equation readily follows from equation (2.10e) defining the projection $\mathit{\Pi}_h$ and the definition of $P_F, \boldsymbol{P}_F$ and P_F, (3.3). This proves the claim and completes the proof. $\square$

Note that with the results obtained so far we can also prove uniqueness for the system that defines the projection: we only need to apply Propositions 4.2, 4.6 and Lemma 4.3 to show that if L, $\boldsymbol{u}$ and p vanish, then so do $\Pi\,\mathrm{L}$, $\boldsymbol{\mathit{\Pi}} \boldsymbol{v}$ and $\mathit{\Pi} p$.

Now pick *any* face F' of K. Then the set $\{\boldsymbol{n}_F : F \neq F'\}$ is a basis of $\mathbb{R}^n$ and its dual basis is denoted by $\{\boldsymbol{n}^{\star}_{F,F'} : F \neq F'\}$. To simplify the notation, from now on, the sum over all the faces F of K is going to be denoted by $\sum_F$. Also, the sum over all the faces F of K, except the face F' will be denoted by $\sum_{F\neq F'}$.

Proposition 4.7. *In the notations above, for all F,*

$$\nu \Pi\,\mathrm{L} - \mathit{\Pi} p \mathrm{I} = \sum_{F\neq F'} (\nu \mathrm{P}_F \mathrm{L} - P_F p \mathrm{I})(\boldsymbol{n}_{F,F'} \otimes \boldsymbol{n}^{*}_{F,F'}) + \sum_{F\neq F'} \big(\mathrm{S}|_F(\boldsymbol{\mathit{\Pi}} \boldsymbol{u} - \boldsymbol{P}_F \boldsymbol{u})\big) \otimes \boldsymbol{n}^{*}_{F,F'}.$$

Proof. Using these two simple facts (A is an arbitrary square matrix)

$$\mathrm{I} = \sum_{F\neq F'} \boldsymbol{n}_{F,F'} \otimes \boldsymbol{n}^{*}_{F,F'}, \qquad \mathrm{A} = \sum_{F\neq F'} (\mathrm{A}\boldsymbol{n}_{F,F'}) \otimes \boldsymbol{n}^{*}_{F,F'}$$

and Proposition 4.6, the result follows readily. $\square$

Step 6: Estimates for the case of a general S**.** The remaining estimates of Theorem 2.1 follow from Proposition 4.7. First, note that

$$\mathrm{P}_F(\nu \mathrm{L} - p\mathrm{I}) = \nu \mathrm{P}_F \mathrm{L} - P_F p \mathrm{I}$$

and that the dual basis $\boldsymbol{n}^{*}_{F,F'}$ can be uniformly bounded (depending only on shape regularity constants of the triangulation). This and the approximation properties of P_F allow us to obtain the bound

$$\begin{aligned} \|(\nu \Pi\,\mathrm{L} - \mathit{\Pi} p \mathrm{I}) - (\nu \mathrm{L} - p\mathrm{I})\|_K \quad \leq \quad & C h^{\ell_{\boldsymbol{\sigma}}+1} |\nu \mathrm{L} - p \mathrm{I}|_{H^{\ell_{\boldsymbol{\sigma}}+1}(K)} \\ & + C \Lambda^{\max}_K (\|\boldsymbol{\mathit{\Pi}} \boldsymbol{u} - \boldsymbol{u}\|_K + h^{\ell_{\boldsymbol{u}}+1} |\boldsymbol{u}|_{H^{\ell_{\boldsymbol{u}}+1}(K)}). \end{aligned}$$

When $\mathrm{tr}\,\mathrm{L} = 0$, by (2.8) it follows that $\mathrm{tr}\,\Pi\,\mathrm{L} = 0$. Therefore

$$\mathit{\Pi} p - p = -\frac{1}{n} \mathrm{tr}\,\big((\nu \Pi\,\mathrm{L} - \mathit{\Pi} p \mathrm{I}) - (\nu \mathrm{L} - p \mathrm{I})\big)$$

and we can obtain a separate bound for $\|\mathit{\Pi} p - p\|_K$. Now, we can bound $\|\nu \Pi\,\mathrm{L} - \nu \mathrm{L}\|_K$. This completes the proof of Theorem 2.1.

Step 7: Estimates for $\mathrm{S} = \mathrm{S}_{SF}$**.** To prove this result, simply take $F' := F^{\star}_K$ in Proposition 4.7 to obtain

$$\nu \Pi\,\mathrm{L} - \mathit{\Pi} p \mathrm{I} = \sum_{F\neq F^{\star}_K} (\nu \mathrm{P}_F \mathrm{L} - P_F p \mathrm{I})(\boldsymbol{n}_{F,F'} \otimes \boldsymbol{n}^{\star}_{F,F^{\star}_K}),$$

by the definition of the stabilization tensor S_{SF}. The remaining estimates of Theorem 2.2 now follow as in the previous case. This completes the proof of Theorem 2.2.

Step 8: Estimates for $\mathrm{S} = \mathrm{S}_{nt}$. The estimate of the error in the velocity gradient and in the pressure of Theorem 2.3 follows from the special form of the stabilization tensor S_{nt}.

We start with an estimate for the velocity gradient. To do that, we need an auxiliary result which we state next. In it, we use the following notation. For each face F of K, we denote by $\mathcal{B}_F$ an orthogonal basis of the vectors orthogonal to $\boldsymbol{n}_F$.

Lemma 4.8. *The set*

$$\mathsf{B} := \{\mathrm{I}\} \cup \{\boldsymbol{t} \otimes \boldsymbol{n}_F : \ F \textit{ face of } K, \quad \boldsymbol{t} \in \mathcal{B}_F\}$$

is a basis of the space of square matrices of order n.

Proof. We only have to prove that if A is a square matrix of order n such that $\mathrm{A} : \mathrm{W} = 0$ for all $\mathrm{W} \in \mathsf{B}$, then $\mathrm{A} = 0$. So, assume that A is such a matrix. Since

$$0 = \mathrm{A} : (\boldsymbol{t} \otimes \boldsymbol{n}_F) = (\mathrm{A}\boldsymbol{n}_F) \cdot \boldsymbol{t}, \qquad \forall \boldsymbol{t} \in \mathcal{B}_F,$$

then $\mathrm{A}\boldsymbol{n}_F = \lambda_F\, \boldsymbol{n}_F$. Using the fact that $\sum_F |F| \boldsymbol{n}_F = \boldsymbol{0}$ and choosing any face F', we get that

$$0 = \mathrm{A}(\sum_F |F| \boldsymbol{n}_F) = \sum_F |F| \lambda_F \boldsymbol{n}_F = \sum_{F \neq F'} |F|(\lambda_F - \lambda_{F'})\boldsymbol{n}_F$$

and therefore $\lambda_F = \lambda$ for all faces F of K. This implies that $\mathrm{A} = \lambda\,\mathrm{I}$. But the condition $\mathrm{A} : \mathrm{I} = 0$ implies that $\lambda = 0$. This shows that $\mathrm{A} = 0$ and completes the proof. □

The dual basis to B can be separated as $\frac{1}{n}\mathrm{I}$ and a set of matrices indexed with F and $\boldsymbol{t} \in \mathcal{B}_F$: $\{\frac{1}{n}\mathrm{I}\} \cup \{\mathrm{W}_{F,\boldsymbol{t}} \ : \ F \text{ face of } K, \boldsymbol{t} \in \mathcal{B}_F\}$. The estimate for the velocity gradient follows from the next result.

Proposition 4.9. *We have that*

$$\Pi\,\mathrm{L} - \mathrm{L} = \sum_F \sum_{\boldsymbol{t} \in \mathcal{B}_{F,t}} \Big((\mathrm{P}_F \mathrm{L} - \mathrm{L}) : (\boldsymbol{t} \otimes \boldsymbol{n}_F) \ + \tau_t(\boldsymbol{\Pi u} - \boldsymbol{P}_F \boldsymbol{u}) \cdot \boldsymbol{t} \Big) \mathrm{W}_{F,\boldsymbol{t}}.$$

Proof. For any square matrix A of order n, we can write that

$$\mathrm{A} = \sum_F \sum_{\boldsymbol{t} \in \mathcal{B}_{F,t}} \big(\mathrm{A} : (\boldsymbol{t} \otimes \boldsymbol{n}_F)\big)\, \mathrm{W}_{F,\boldsymbol{t}} + \frac{\mathrm{tr}\, A}{n} \mathrm{I}.$$

Hence, recalling that $\mathrm{A} : (\boldsymbol{t} \otimes \boldsymbol{n}_F) = \mathrm{A}\boldsymbol{n}_F \cdot \boldsymbol{t}$, for $\mathrm{A} := \Pi\,\mathrm{L} - \mathrm{L}$, we get that

$$\Pi\,\mathrm{L} - \mathrm{L} = \sum_F \sum_{\boldsymbol{t} \in \mathcal{B}_{F,t}} \big((\Pi\,\mathrm{L} - \mathrm{L})\boldsymbol{n}_F \cdot \boldsymbol{t}\big)\, \mathrm{W}_{F,\boldsymbol{t}},$$

since $\mathrm{tr}\,(\Pi\,\mathrm{L} - \mathrm{L}) = 0$ by (2.10d). The identity now follows after a straightforward application of Proposition 4.6. This completes the proof. □

The estimate of the pressure follows from that of L and the the following result.

Proposition 4.10. *If $\nabla \cdot \boldsymbol{u} = 0$, we have*

$$\mathit{\Pi} p = \sum_F \frac{|F|}{|\partial K|} \left(P_F p + \nu\big((\Pi\,\mathrm{L} - \mathrm{P}_F \mathrm{L})\boldsymbol{n}_F\big) \cdot \boldsymbol{n}_F\right).$$

Proof. If we multiply the identity of Proposition 4.6 by $\frac{|F|}{|\partial K|}$, take the dot product with $\boldsymbol{n}_F$ and add over all the faces of K, we obtain

$$\Pi p = \sum_F \frac{|F|}{|\partial K|}\Big(P_F p + \nu\big((\Pi\,\mathrm{L} - \mathrm{P}_F\mathrm{L})\boldsymbol{n}_F\big)\cdot\boldsymbol{n}_F - \nu\tau_n(\boldsymbol{\Pi u} - \boldsymbol{P}_F\boldsymbol{u})\cdot\boldsymbol{n}_F\Big),$$

and the identity follows if we show that $\sum_F |F|(\boldsymbol{\Pi u} - \boldsymbol{P}_F\boldsymbol{u})\cdot\boldsymbol{n}_F = 0$.

To do that, we begin by noting that, since $\sum_F |F|\boldsymbol{n}_F = \mathbf{0}$, then

$$\sum_F |F|\,\boldsymbol{\Pi u}\cdot\boldsymbol{n}_F = 0 = \sum_F |F|\,\mathbf{\Pi}^{\mathrm{RT}}\boldsymbol{u}\cdot\boldsymbol{n}_F,$$

where $\mathbf{\Pi}^{\mathrm{RT}}$ is the Raviart-Thomas projection of order k. We are now going to show that for every face, the polynomial $\delta := (\mathbf{\Pi}^{\mathrm{RT}}\boldsymbol{u} - \boldsymbol{P}_F\boldsymbol{u})\cdot\boldsymbol{n}_F$ vanishes. This will prove the result.

Then pick an arbitrary face F of K and take δ as above. Since $\nabla\cdot\boldsymbol{u} = 0$, $\mathbf{\Pi}^{\mathrm{RT}}\boldsymbol{u} \in \boldsymbol{\mathcal{P}}_k(K)$, and so $\delta \in \mathcal{P}_k(K)$. Moreover, it satisfies

$$\begin{aligned}(\delta, q)_K &= 0 \qquad \forall\, q \in \mathcal{P}_{k-1}(K),\\ \langle\delta, \mu\rangle_F &= 0 \qquad \forall\, \mu \in \mathcal{P}_k(F),\end{aligned}$$

by the definition of $\boldsymbol{P}_F$ (3.3) and that of $\mathbf{\Pi}^{\mathrm{RT}}$. This implies that $\delta = 0$ and completes the proof. □

The proof of the estimates for L and p in Theorem 2.3 are now a direct consequence of Propositions 4.9 and 4.10. Note that the matrices $\mathrm{W}_{F,\boldsymbol{t}}$ in Proposition 4.9 can be bounded uniformly in terms of geometric quantities that depend only on the shape regularity of the mesh.

5. Numerical results

In this section, we carry out numerical experiments devised to verify the theoretical orders of convergence of the approximations provided by the HDG method given by Theorem 2.4, and those of the postprocessed velocity given by Theorem 2.5.

As a test problem, we take the flow uncovered by Kovasznay [25]. We consider the Stokes problem whose exact solution coincides with the analytical solution of the incompressible Navier-Stokes equations obtained by Kovasznay, namely,

$$\begin{aligned}u_1 &= 1 - \exp(\lambda x_1)\cos(2\pi x_2),\\ u_2 &= \frac{\lambda}{2\pi}\exp(\lambda x_1)\sin(2\pi x_2),\\ p &= \frac{1}{2}\exp(2\lambda x_1),\end{aligned}$$

where $\lambda = \frac{Re}{2} - \sqrt{\frac{Re^2}{4} + 4\pi^2}$ and $Re = \frac{1}{\nu}$ is the Reynolds number. The Kovasznay flow is also a solution of the Stokes problem with the source term $\boldsymbol{f} = -(\boldsymbol{u}\cdot\nabla)\boldsymbol{u}$. We take Dirichlet boundary conditions for the velocity as the restriction of the exact solution to the domain boundary. Here the computational domain is $\Omega = (-0.5, 1.5)\times(0,2)$ and $\nu = 0.1$ so that the Reynolds number is $Re = 10$.

In our experiments, we consider meshes that are refinements of a uniform mesh of 32 congruent triangles. Each refinement is obtained by subdividing each triangle into four congruent triangles. We say that the mesh has level ℓ ($h := 2/2^{\ell+2}$) if it is obtained from the original mesh by ℓ of these refinements. On these meshes, we consider polynomials of degree k to represent all the approximate variables using a nodal basis within each element, with the nodes uniformly distributed.

The results for the SFH method are listed in Table 5.2. Note that the orders of convergence agree with those predicted by our theoretical analysis summarized in Tables 2.1 and 2.3. Indeed, note that the errors in the velocity, the pressure, the velocity gradient are of order $k+1$, for $k \geq 1$, and that the error in the postprocessed velocity is of order $k+2$, also for $k \geq 1$. For $k = 0$, the errors in the approximation of all these variables converge with order one, except for the velocity when $\tau = h$ in which there is no convergence. Note also that the errors for the pressure and the velocity gradient are independent of the value of the parameter τ defining the stabilization tensor, in full agreement with our analysis.

The results for several choices of τ_n and τ_t are listed in Tables 5.3, 5.4, and 5.5. Again, the numerical results agree with the theoretical analysis summarized in Table 5.1. Indeed, for $\tau_t \neq 1/h$ the errors in the velocity, the pressure, the velocity gradient are of order $k + 1$, for $k \geq 1$, and that the error in the postprocessed velocity is of order $k + 2$, also for $k \geq 1$. For $\tau_t = 1/h$, the order of convergence of the errors in the velocity gradient and the postprocessed velocity are reduced by one. Remarkably enough, the pressure seems to converge with the optimal order of $k + 1$ (unless $\tau_n = 1/h$); this is the only case not predicted by our theoretical results. Finally, for $k = 0$, the errors in the approximation of all these variables converge with order one, except for the velocity when $\tau_n = \tau_t = h$ in which there is no convergence.

TABLE 5.1. Expected optimal convergence for the method with $\mathrm{S} = \mathrm{S}_{nt}$ for the quantities $\eta_{\boldsymbol{u}} := \|\boldsymbol{u} - \boldsymbol{u}_h\|_\Omega$, $\eta_p := \|p - p_h\|_\Omega$, $\eta_{\mathrm{L}} := \|\mathrm{L} - \mathrm{L}_h\|_\Omega$, $\eta_{\boldsymbol{u}^\star} := \|\boldsymbol{u} - \boldsymbol{u}_h^\star\|_\Omega$. We take constant values of τ_t and any of $\tau_n \in \{h, 1, 1/h\}$. The quantities with an asterisk are not valid for the case $\tau_t = \tau_n = h$. In that case, the quantity has to be lowered by one unit.

	$\tau_t \in \{h, 1\}$			$\tau_t = 1/h$		
	$\eta_{\boldsymbol{u}}$	$\eta_p, \eta_{\mathrm{L}}$	$\eta_{\boldsymbol{u}^\star}$	$\eta_{\boldsymbol{u}}$	$\eta_p, \eta_{\mathrm{L}}$	$\eta_{\boldsymbol{u}^\star}$
$k = 0$	1^*	1	1	0	0	0
$k \geq 1$	$(k+1)^*$	$k+1$	$k+2$	$k+1$	k	$k+1$

Finally, let us point out that the postprocessed velocity is $\boldsymbol{H}(\mathrm{div})$-conforming and exactly divergence-free. Indeed, it is numerically verified that the divergence of the postprocessed velocity is zero within machine precision and that its normal component is continuous across interior faces.

Table 5.2. History of convergence of the SFH method.

degree k	mesh $2h^{-1}$	$\lVert \boldsymbol{u} - \boldsymbol{u}_h \rVert_{\mathcal{T}_h}$ error	order	$\lVert p - p_h \rVert_{\mathcal{T}_h}$ error	order	$\lVert \mathrm{L} - \mathrm{L}_h \rVert_{\mathcal{T}_h}$ error	order	$\lVert \boldsymbol{u} - \boldsymbol{u}_h^\star \rVert_{\mathcal{T}_h}$ error	order
$\tau = h$									
	4	1.18e+1	−−	1.88e-0	−−	3.22e+1	−−	5.02e-0	−−
	8	1.22e+1	−0.05	1.19e-0	0.66	2.06e+1	0.64	2.03e-0	1.31
0	16	1.21e+1	0.01	6.02e-1	0.99	1.13e+1	0.87	6.10e-1	1.73
	32	1.22e+1	0.00	3.06e-1	0.97	5.99e-0	0.92	2.11e-1	1.53
	64	1.22e+1	0.00	1.60e-1	0.93	3.19e-0	0.91	8.73e-2	1.27
	4	4.14e-0	−−	1.22e-0	−−	1.22e+1	−−	1.06e-0	−−
	8	2.07e-0	1.00	3.50e-1	1.80	3.71e-0	1.72	1.61e-1	2.73
1	16	1.08e-0	0.94	9.12e-2	1.94	1.01e-0	1.88	2.32e-2	2.79
	32	5.47e-1	0.98	2.26e-2	2.01	2.63e-1	1.94	3.14e-3	2.89
	64	2.75e-1	0.99	5.59e-3	2.02	6.70e-2	1.97	4.09e-4	2.94
	4	8.92e-1	−−	2.64e-1	−−	2.74e-0	−−	1.67e-1	−−
	8	2.47e-1	1.85	4.17e-2	2.66	4.39e-1	2.64	1.52e-2	3.46
2	16	6.67e-2	1.89	5.57e-3	2.90	6.11e-2	2.84	1.07e-3	3.83
	32	1.70e-2	1.97	7.02e-4	2.99	7.90e-3	2.95	7.03e-5	3.93
	64	4.29e-3	1.99	8.75e-5	3.00	1.00e-3	2.98	4.48e-6	3.97
$\tau = 1$									
	4	8.48e-0	−−	1.88e-0	−−	3.22e+1	−−	5.02e-0	−−
	8	4.50e-0	0.92	1.19e-0	0.66	2.06e+1	0.64	2.03e-0	1.31
0	16	1.99e-0	1.17	6.02e-1	0.99	1.13e+1	0.87	6.10e-1	1.73
	32	9.08e-1	1.14	3.06e-1	0.97	5.99e-0	0.92	2.11e-1	1.53
	64	4.29e-1	1.08	1.60e-1	0.93	3.19e-0	0.91	8.73e-2	1.27
	4	2.54e-0	−−	1.22e-0	−−	1.22e+1	−−	1.06e-0	−−
	8	6.45e-1	1.98	3.50e-1	1.80	3.71e-0	1.72	1.61e-1	2.73
1	16	1.61e-1	2.01	9.12e-2	1.94	1.01e-0	1.88	2.32e-2	2.79
	32	3.91e-2	2.04	2.26e-2	2.01	2.63e-1	1.94	3.14e-3	2.89
	64	9.58e-3	2.03	5.59e-3	2.02	6.70e-2	1.97	4.09e-4	2.94
	4	5.31e-1	−−	2.64e-1	−−	2.74e-0	−−	1.67e-1	−−
	8	7.61e-2	2.8	4.17e-2	2.66	4.39e-1	2.64	1.52e-2	3.46
	16	9.70e-3	2.97	5.57e-3	2.90	6.11e-2	2.84	1.07e-3	3.83
2	32	1.21e-3	3.00	7.02e-4	2.99	7.90e-3	2.95	7.03e-5	3.93
	64	1.50e-4	3.01	8.75e-5	3.00	1.00e-3	2.98	4.48e-6	3.97
$\tau = 1/h$									
	4	6.94e-0	−−	1.88e-0	−−	3.22e+1	−−	5.02e-0	−−
	8	2.78e-0	1.32	1.19e-0	0.66	2.06e+1	0.64	2.03e-0	1.31
0	16	8.92e-1	1.64	6.02e-1	0.99	1.13e+1	0.87	6.10e-1	1.73
	32	3.32e-1	1.42	3.06e-1	0.97	5.99e-0	0.92	2.11e-1	1.53
	64	1.48e-1	1.17	1.6e-1	0.93	3.19e-0	0.91	8.73e-2	1.27
	4	1.82e-0	−−	1.22e-0	−−	1.22e+1	−−	1.06e-0	−−
	8	3.42e-1	2.41	3.5e-1	1.80	3.71e-0	1.72	1.61e-1	2.73
1	16	7.26e-2	2.24	9.12e-2	1.94	1.01e-0	1.88	2.32e-2	2.79
	32	1.65e-2	2.13	2.26e-2	2.01	2.63e-1	1.94	3.14e-3	2.89
	64	4.02e-3	2.04	5.59e-3	2.02	6.70e-2	1.97	4.09e-4	2.94
	4	3.78e-1	−−	2.64e-1	−−	2.74e-0	−−	1.67e-1	−−
	8	4.38e-2	3.11	4.17e-2	2.66	4.39e-1	2.64	1.52e-2	3.46
2	16	4.66e-3	3.23	5.57e-3	2.90	6.11e-2	2.84	1.07e-3	3.83
	32	5.59e-4	3.06	7.02e-4	2.99	7.90e-3	2.95	7.03e-5	3.93
	64	6.92e-5	3.01	8.75e-5	3.00	1.00e-3	2.98	4.48e-6	3.97

TABLE 5.3. History of convergence of the HDG method for $\tau_n = h$.

degree k	mesh $2h^{-1}$	$\Vert \boldsymbol{u}-\boldsymbol{u}_h\Vert_{\mathcal{T}_h}$ error	order	$\Vert p-p_h\Vert_{\mathcal{T}_h}$ error	order	$\Vert \mathrm{L}-\mathrm{L}_h\Vert_{\mathcal{T}_h}$ error	order	$\Vert \boldsymbol{u}-\boldsymbol{u}_h^\star\Vert_{\mathcal{T}_h}$ error	order
$\tau_t = h$									
	4	3.40e-0	−−	1.16e-0	−−	1.70e+1	−−	2.55e-0	−−
	8	4.10e-0	−0.27	5.42e-1	1.10	1.33e+1	0.35	1.42e-0	0.85
0	16	3.94e-0	0.06	3.41e-1	0.67	8.24e-0	0.70	4.88e-1	1.54
	32	3.87e-0	0.03	1.94e-1	0.81	4.60e-0	0.84	1.87e-1	1.38
	64	3.85e-0	0.01	1.11e-1	0.80	2.54e-0	0.85	8.23e-2	1.18
	4	1.45e-0	−−	8.75e-1	−−	7.33e-0	−−	5.80e-1	−−
	8	6.84e-1	1.08	2.69e-1	1.70	2.57e-0	1.51	1.06e-1	2.45
1	16	3.45e-1	0.99	7.34e-2	1.87	7.34e-1	1.81	1.62e-2	2.71
	32	1.73e-1	1.00	1.86e-2	1.98	1.95e-1	1.92	2.23e-3	2.86
	64	8.65e-2	1.00	4.65e-3	2.00	4.99e-2	1.96	2.93e-4	2.93
	4	3.20e-1	−−	2.07e-1	−−	1.97e-0	−−	1.02e-1	−−
	8	8.23e-2	1.96	3.38e-2	2.62	3.13e-1	2.65	1.07e-2	3.26
2	16	2.13e-2	1.95	4.59e-3	2.88	4.38e-2	2.84	7.20e-4	3.89
	32	5.38e-3	1.98	5.86e-4	2.97	5.66e-3	2.95	4.67e-5	3.95
	64	1.35e-3	2.00	7.34e-5	3.00	7.15e-4	2.98	2.96e-6	3.98
$\tau_t = 1$									
	4	3.11e-0	−−	1.22e-0	−−	1.75e+1	−−	2.62e-0	−−
	8	2.99e-0	0.06	5.93e-1	1.04	1.66e+1	0.07	1.73e-0	0.6
0	16	1.62e-0	0.88	5.06e-1	0.23	1.23e+1	0.44	6.67e-1	1.37
	32	7.82e-1	1.05	2.82e-1	0.84	7.47e-0	0.72	2.41e-1	1.47
	64	3.70e-1	1.08	1.46e-1	0.94	4.14e-0	0.85	9.51e-2	1.34
	4	1.22e-0	−−	9.19e-1	−−	8.15e-0	−−	6.31e-1	−−
	8	4.14e-1	1.56	2.85e-1	1.69	3.34e-0	1.29	1.39e-1	2.18
1	16	1.20e-1	1.78	7.80e-2	1.87	1.13e-0	1.56	2.46e-2	2.50
	32	3.15e-2	1.93	1.99e-2	1.97	3.26e-1	1.80	3.62e-3	2.76
	64	8.00e-3	1.98	4.96e-3	2.00	8.69e-2	1.91	4.90e-4	2.89
	4	2.74e-1	−−	2.21e-1	−−	2.27e-0	−−	1.16e-1	−−
	8	4.84e-2	2.50	3.63e-2	2.61	4.28e-1	2.41	1.42e-2	3.03
2	16	7.14e-3	2.76	4.97e-3	2.87	6.86e-2	2.64	1.11e-3	3.68
	32	9.66e-4	2.89	6.42e-4	2.95	9.72e-3	2.82	7.84e-5	3.83
	64	1.25e-4	2.95	8.09e-5	2.99	1.29e-3	2.91	5.18e-6	3.92
$\tau_t = 1/h$									
	4	2.90e-0	−−	1.28e-0	−−	1.75e+1	−−	2.63e-0	−−
	8	2.65e-0	0.13	6.13e-1	1.07	1.82e+1	−0.06	1.90e-0	0.47
0	16	1.18e-0	1.17	9.44e-1	−0.62	1.41e+1	0.37	8.14e-1	1.22
	32	4.89e-1	1.27	1.05e-0	−0.15	8.49e-0	0.73	3.75e-1	1.12
	64	2.86e-1	0.77	1.12e-0	−0.10	4.79e-0	0.83	2.56e-1	0.55
	4	1.08e-0	−−	9.50e-1	−−	8.79e-0	−−	6.72e-1	−−
	8	3.18e-1	1.76	3.01e-1	1.66	3.89e-0	1.18	1.64e-1	2.03
1	16	7.30e-2	2.12	8.4e-2	1.84	1.46e-0	1.42	3.14e-2	2.39
	32	1.63e-2	2.16	2.19e-2	1.94	5.15e-1	1.50	5.56e-3	2.50
	64	3.89e-3	2.07	5.70e-3	1.94	2.12e-1	1.28	1.14e-3	2.29
	4	2.51e-1	−−	2.38e-1	−−	2.58e-0	−−	1.30e-1	−−
	8	3.86e-2	2.70	3.89e-2	2.61	5.51e-1	2.23	1.81e-2	2.85
2	16	4.51e-3	3.10	5.45e-3	2.84	9.79e-2	2.49	1.58e-3	3.52
	32	5.43e-4	3.05	7.82e-4	2.80	1.88e-2	2.38	1.53e-4	3.37
	64	6.70e-5	3.02	1.27e-4	2.62	4.18e-3	2.17	1.71e-5	3.16

TABLE 5.4. History of convergence of the HDG method for $\tau_n = 1$.

degree k	mesh $2h^{-1}$	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{\mathcal{T}_h}$ error	order	$\|p-p_h\|_{\mathcal{T}_h}$ error	order	$\|\mathrm{L}-\mathrm{L}_h\|_{\mathcal{T}_h}$ error	order	$\|\boldsymbol{u}-\boldsymbol{u}_h^\star\|_{\mathcal{T}_h}$ error	order
$\tau_t = h$									
0	4	2.18e-0	−−	1.22e-0	−−	1.44e+1	−−	2.22e-0	−−
	8	1.78e-0	0.29	5.07e-1	1.26	8.97e-0	0.68	1.06e-0	1.06
	16	1.06e-0	0.75	2.76e-1	0.88	4.82e-0	0.90	3.66e-1	1.54
	32	5.93e-1	0.84	1.62e-1	0.77	2.82e-0	0.77	1.65e-1	1.15
	64	3.16e-1	0.91	9.96e-2	0.70	1.75e-0	0.69	7.93e-2	1.06
1	4	1.09e-0	−−	8.7e-1	−−	6.29e-0	−−	4.94e-1	−−
	8	3.34e-1	1.71	2.58e-1	1.75	1.85e-0	1.77	7.84e-2	2.65
	16	1.04e-1	1.68	7.03e-2	1.87	4.67e-1	1.98	1.11e-2	2.83
	32	2.89e-2	1.85	1.81e-2	1.96	1.19e-1	1.98	1.45e-3	2.93
	64	7.63e-3	1.92	4.58e-3	1.98	3.02e-2	1.97	1.88e-4	2.95
2	4	2.57e-1	−−	2.06e-1	−−	1.81e-0	−−	9.13e-2	−−
	8	4.60e-2	2.48	3.27e-2	2.66	2.42e-1	2.90	8.77e-3	3.38
	16	6.71e-3	2.78	4.42e-3	2.88	3.11e-2	2.96	5.39e-4	4.02
	32	9.32e-4	2.85	5.7e-4	2.96	3.93e-3	2.98	3.52e-5	3.94
	64	1.23e-4	2.92	7.22e-5	2.98	5.01e-4	2.97	2.28e-6	3.95
$\tau_t = 1$									
0	4	2.06e-0	−−	1.35e-0	−−	1.47e+1	−−	2.25e-0	−−
	8	1.56e-0	0.40	5.75e-1	1.23	1.05e+1	0.48	1.23e-0	0.87
	16	7.19e-1	1.12	4.82e-1	0.25	6.75e-0	0.64	4.61e-1	1.42
	32	3.34e-1	1.10	2.66e-1	0.86	4.14e-0	0.71	2.00e-1	1.20
	64	1.58e-1	1.08	1.44e-1	0.89	2.45e-0	0.76	9.38e-2	1.09
1	4	9.55e-1	−−	9.36e-1	−−	6.97e-0	−−	5.25e-1	−−
	8	2.51e-1	1.93	2.87e-1	1.71	2.34e-0	1.57	1.01e-1	2.38
	16	6.61e-2	1.93	7.85e-2	1.87	7.48e-1	1.65	1.68e-2	2.59
	32	1.62e-2	2.03	2.01e-2	1.97	2.08e-1	1.85	2.39e-3	2.81
	64	3.98e-3	2.02	5.04e-3	1.99	5.51e-2	1.92	3.21e-4	2.89
2	4	2.31e-1	−−	2.27e-1	−−	2.12e-0	−−	1.03e-1	−−
	8	3.47e-2	2.74	3.77e-2	2.59	3.50e-1	2.60	1.19e-2	3.11
	16	4.21e-3	3.04	5.1e-3	2.89	4.89e-2	2.84	8.20e-4	3.86
	32	5.26e-4	3.00	6.50e-4	2.97	6.56e-3	2.90	5.56e-5	3.88
	64	6.54e-5	3.01	8.14e-5	3.00	8.49e-4	2.95	3.62e-6	3.94
$\tau_t = 1/h$									
0	4	1.96e-0	−−	1.54e-0	−−	1.48e+1	−−	2.25e-0	−−
	8	1.62e-0	0.27	7.23e-1	1.09	1.17e+1	0.34	1.39e-0	0.69
	16	8.02e-1	1.02	1.27e-0	−0.81	8.78e-0	0.41	6.48e-1	1.10
	32	4.47e-1	0.84	1.25e-0	0.02	6.21e-0	0.50	3.84e-1	0.75
	64	3.06e-1	0.55	1.23e-0	0.02	4.14e-0	0.58	2.85e-1	0.43
1	4	8.71e-1	−−	9.94e-1	−−	7.58e-0	−−	5.55e-1	−−
	8	2.33e-1	1.90	3.35e-1	1.57	2.97e-0	1.35	1.31e-1	2.09
	16	5.84e-2	1.99	9.84e-2	1.77	1.30e-0	1.19	2.79e-2	2.23
	32	1.44e-2	2.02	2.69e-2	1.87	5.18e-1	1.33	5.52e-3	2.34
	64	3.64e-3	1.98	7.14e-3	1.91	2.21e-1	1.23	1.17e-3	2.24
2	4	2.20e-1	−−	2.58e-1	−−	2.48e-0	−−	1.17e-1	−−
	8	3.21e-2	2.77	4.83e-2	2.41	5.30e-1	2.23	1.74e-2	2.75
	16	3.85e-3	3.06	6.83e-3	2.82	9.81e-2	2.43	1.60e-3	3.44
	32	4.91e-4	2.97	9.78e-4	2.80	2.01e-2	2.29	1.66e-4	3.27
	64	6.32e-5	2.96	1.52e-4	2.68	4.43e-3	2.18	1.84e-5	3.17

TABLE 5.5. History of convergence of the HDG method for $\tau_n = 1/h$.

degree k	mesh $2h^{-1}$	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{\mathcal{T}_h}$ error	order	$\|p-p_h\|_{\mathcal{T}_h}$ error	order	$\|\mathrm{L}-\mathrm{L}_h\|_{\mathcal{T}_h}$ error	order	$\|\boldsymbol{u}-\boldsymbol{u}_h^\star\|_{\mathcal{T}_h}$ error	order
$\tau_t = h$									
0	4	1.58e-0	$--$	1.25e-0	$--$	1.31e+1	$--$	2.1e-0	$--$
	8	9.72e-1	0.7	5.23e-1	1.25	8.00e-0	0.71	9.79e-1	1.10
	16	4.67e-1	1.06	2.84e-1	0.88	4.57e-0	0.81	3.55e-1	1.46
	32	2.43e-1	0.94	1.66e-1	0.77	2.82e-0	0.70	1.65e-1	1.11
	64	1.25e-1	0.96	1.01e-1	0.72	1.76e-0	0.68	7.93e-2	1.06
1	4	8.93e-1	$--$	8.78e-1	$--$	5.82e-0	$--$	4.57e-1	$--$
	8	2.21e-1	2.02	2.59e-1	1.76	1.68e-0	1.79	7.4e-2	2.63
	16	6.24e-2	1.82	7.11e-2	1.87	4.57e-1	1.88	1.11e-2	2.73
	32	1.56e-2	2.00	1.83e-2	1.96	1.20e-1	1.93	1.49e-3	2.90
	64	3.90e-3	2.00	4.61e-3	1.99	3.07e-2	1.97	1.93e-4	2.95
2	4	2.28e-1	$--$	2.07e-1	$--$	1.74e-0	$--$	9.00e-2	$--$
	8	3.64e-2	2.64	3.30e-2	2.65	2.33e-1	2.90	8.77e-3	3.36
	16	4.35e-3	3.06	4.52e-3	2.87	3.14e-2	2.89	5.59e-4	3.97
	32	5.48e-4	2.99	5.80e-4	2.96	4.03e-3	2.96	3.66e-5	3.93
	64	6.86e-5	3.00	7.30e-5	2.99	5.10e-4	2.98	2.34e-6	3.97
$\tau_t = 1$									
0	4	1.55e-0	$--$	1.45e-0	$--$	1.32e+1	$--$	2.10e-0	$--$
	8	9.84e-1	0.66	6.11e-1	1.24	8.29e-0	0.67	1.04e-0	1.01
	16	4.23e-1	1.22	5.68e-1	0.11	4.84e-0	0.78	3.76e-1	1.47
	32	2.23e-1	0.92	3.34e-1	0.76	3.20e-0	0.60	1.72e-1	1.13
	64	1.17e-1	0.92	1.91e-1	0.81	2.06e-0	0.64	8.08e-2	1.09
1	4	8.16e-1	$--$	9.77e-1	$--$	6.43e-0	$--$	4.85e-1	$--$
	8	1.99e-1	2.03	3.13e-1	1.64	2.02e-0	1.67	9.63e-2	2.33
	16	5.70e-2	1.81	8.93e-2	1.81	7.08e-1	1.52	1.72e-2	2.48
	32	1.46e-2	1.97	2.44e-2	1.87	2.08e-1	1.76	2.57e-3	2.75
	64	3.74e-3	1.96	6.40e-3	1.93	5.80e-2	1.85	3.60e-4	2.83
2	4	2.15e-1	$--$	2.35e-1	$--$	2.06e-0	$--$	1.00e-1	$--$
	8	3.31e-2	2.70	4.12e-2	2.51	3.46e-1	2.58	1.21e-2	3.05
	16	3.95e-3	3.07	6.21e-3	2.73	5.10e-2	2.76	9.31e-4	3.70
	32	5.12e-4	2.95	8.37e-4	2.89	7.38e-3	2.79	7.02e-5	3.73
	64	6.58e-5	2.96	1.09e-4	2.94	1.01e-3	2.86	4.93e-6	3.83
$\tau_t = 1/h$									
0	4	1.51e-0	$--$	1.77e-0	$--$	1.32e+1	$--$	2.10e-0	$--$
	8	1.11e-0	0.44	1.08e-0	0.71	8.78e-0	0.59	1.16e-0	0.85
	16	5.30e-1	1.07	2.32e-0	-1.11	5.72e-0	0.62	5.45e-1	1.09
	32	3.88e-1	0.45	2.53e-0	-0.12	4.76e-0	0.27	3.93e-1	0.47
	64	3.44e-1	0.18	2.64e-0	-0.06	4.39e-0	0.11	3.45e-1	0.19
1	4	7.66e-1	$--$	1.09e-0	$--$	7.08e-0	$--$	5.19e-1	$--$
	8	2.04e-1	1.91	4.75e-1	1.19	2.88e-0	1.30	1.42e-1	1.87
	16	5.67e-2	1.84	2.00e-1	1.25	1.75e-0	0.72	3.92e-2	1.85
	32	1.43e-2	1.99	9.25e-2	1.11	8.85e-1	0.98	9.89e-3	1.99
	64	3.60e-3	1.99	4.39e-2	1.08	4.46e-1	0.99	2.50e-3	1.99
2	4	2.10e-1	$--$	2.84e-1	$--$	2.5e-0	$--$	1.18e-1	$--$
	8	3.23e-2	2.70	7.20e-2	1.98	6.13e-1	2.03	2.03e-2	2.53
	16	3.82e-3	3.08	1.62e-2	2.15	1.40e-1	2.13	2.44e-3	3.06
	32	4.85e-4	2.98	3.68e-3	2.14	3.56e-2	1.97	3.17e-4	2.94
	64	6.12e-5	2.99	8.65e-4	2.09	9.02e-3	1.98	4.07e-5	2.96

6. Concluding remarks

We end the paper with some comments on the different effects of the stabilization parameters τ_n and τ_t, on alternative postprocessings of the velocity and on the expected similarity of our results with HDG methods for linear isotropic elasticity.

The role of τ_n and τ_t. It is interesting to note that the HDG method with the stabilization S_{nt} displayed very different convergence properties with respect to the size of τ_n and τ_t. It kept its optimal convergence and superconvergence properties whenever τ_t remained of order one. This suggests that HDG-like methods that use approximate velocities whose normal components are continuous, that is, methods for which $\boldsymbol{V}_h \subset \boldsymbol{H}(\mathrm{div}, \Omega)$, should also have the above-mentioned convergence properties. This constitutes the subject of ongoing research.

On the other hand, by taking $\tau_t = 1/h$, the loss of the optimality in the convergence of the velocity gradient, but not of the pressure when $\tau_n \neq 1/h$, was observed; this might be a two-dimensional phenomenon. In any case, the above-mentioned loss implies that the postprocessed velocity no longer converges faster than the original approximation. This suggests that this might also be true for HDG-like methods using velocity spaces whose tangential components are continuous, that is, methods for which $\boldsymbol{V}_h \subset \boldsymbol{H}(\boldsymbol{curl}, \Omega)$.

Alternative postprocessings for the case $k = 0$. Let us emphasize that the HDG method for $k = 0$ can be considered to be a finite volume method. As we have seen, if the stabilization tensor S is suitably chosen, not only the velocity but also its gradient and pressure converge with order one.

In this case, since the postprocessed velocity $\boldsymbol{u}_h^\star$ can at most converge with order one, it is reasonable to consider simpler ways to compute postprocessings. Here, we want to briefly discuss the following alternatives:

$$\boldsymbol{u}_h^{\mathrm{RT},*}(\boldsymbol{x}) := \sum_F \frac{|F|}{n|K|}(\boldsymbol{x} - \boldsymbol{x}_F)\, \widehat{\boldsymbol{u}}_h|_F \cdot \boldsymbol{n}_F,$$
$$\boldsymbol{u}_h^{\mathrm{CR},*}(\boldsymbol{x}) := \sum_F (1 - n\,\lambda_F(\boldsymbol{x}))\, \widehat{\boldsymbol{u}}_h|_F.$$

Here $\boldsymbol{x}_F$ denotes the vertex opposite to the face F and $\lambda_F(\cdot)$ is the corresponding barycentric coordinate—not to be confused with the eigenvalues of the stabilization tensor in (2.3) and (2.4).

The function $\boldsymbol{u}_h^{\mathrm{RT},*}$ is nothing but the modification of the lowest-order Raviart-Thomas projection; see [2, 16]. It is in $\boldsymbol{H}(\mathrm{div}, \Omega)$ and is divergence-free by the equation (2.2c) defining the HDG method; as a consequence $\boldsymbol{u}_h^{\mathrm{RT},*}$ is constant on each simplex. The function $\boldsymbol{u}_h^{\mathrm{CR},*}$ is a Crouzeix-Raviart-like projection. It is not in $\boldsymbol{H}(\mathrm{div}, \Omega)$ but its divergence is equal to zero inside each element thanks to equation (2.2c). On the other hand, its average on each face of the triangulation is uniquely defined.

We present the error and order convergence of $\|\boldsymbol{u} - \boldsymbol{u}_h^{\mathrm{RT},\star}\|_{\mathcal{T}_h}$ and $\|\boldsymbol{u} - \boldsymbol{u}_h^{\mathrm{CR},\star}\|_{\mathcal{T}_h}$ in Table 6.1 for the SFH method and in Table 6.2 for the HDG method for several choices of τ_n and τ_t.

For the SFH method, we observe that the approximation $\boldsymbol{u}_h^*$ is better than $\boldsymbol{u}_h^{\mathrm{CR},\star}$ which in turn is better than the approximation $\boldsymbol{u}_h^{\mathrm{RT},\star}$. A similar result holds for the other choice of the stabilization parameter S. Note that for the SFH method, the errors are independent of the value of the parameter τ defining the stabilization tensor.

TABLE 6.1. History of convergence of $\|\boldsymbol{u} - \boldsymbol{u}_h^{\mathrm{RT},\star}\|_{\mathcal{T}_h}$ and $\|\boldsymbol{u} - \boldsymbol{u}_h^{\mathrm{CR},\star}\|_{\mathcal{T}_h}$ for the special case $k = 0$ for the SFH method.

	$\\|\boldsymbol{u} - \boldsymbol{u}_h^{\mathrm{RT},\star}\\|_{\mathcal{T}_h}$						$\\|\boldsymbol{u} - \boldsymbol{u}_h^{\mathrm{CR},\star}\\|_{\mathcal{T}_h}$					
mesh	$\tau = h$		$\tau = 1$		$\tau = 1/h$		$\tau = h$		$\tau = 1$		$\tau = 1/h$	
$2h^{-1}$	error	order	error	order	error	order	error	order	error	order	error	order
4	6.55e-0	−−	6.55e-0	−−	6.55e-0	−−	13.4e-0	−−	13.4e-0	−−	13.4e-0	−−
8	3.49e-0	0.91	3.49e-0	0.91	3.49e-0	0.91	5.22e-0	1.36	5.22e-0	1.36	5.22e-0	1.36
16	1.68e-0	1.06	1.68e-0	1.06	1.68e-0	1.06	1.56e-0	1.74	1.56e-0	1.74	1.56e-0	1.74
32	8.45e-1	0.99	8.45e-1	0.99	8.45e-1	0.99	5.29e-1	1.56	5.29e-1	1.56	5.29e-1	1.56
64	4.26e-1	0.99	4.26e-1	0.99	4.26e-1	0.99	2.16e-1	1.29	2.16e-1	1.29	2.16e-1	1.29

Another postprocessing of the velocity for $k \geq 1$. Let us note that, if we are *not* interested in obtaining an $H(\mathrm{div})$-conforming postprocessed velocity $\boldsymbol{u}^\circ$, we can use the projection by Stenberg [31], or the following slight modification:

$$\begin{aligned}(\boldsymbol{\nabla}\, \boldsymbol{u}_h^\circ - \mathrm{L}_h, \boldsymbol{\nabla}\, \boldsymbol{v})_K &= 0 \quad \forall\, \boldsymbol{v} \in \boldsymbol{\mathcal{P}}_{k+1}(K),\\ (\boldsymbol{u}_h^\circ - \boldsymbol{u}_h, \boldsymbol{v})_K &= 0 \quad \forall\, \boldsymbol{v} \in \boldsymbol{\mathcal{P}}_0(K).\end{aligned}$$

These projections have convergence properties similar to those of the projection introduced in this paper and are simpler to implement.

Extension to the isotropic linear elasticity equations. An HDG method for the equations linear three-dimensional isotropic elasticity can be easily defined which should have convergence properties similar to the ones displayed by the method just analyzed. Indeed, note that the corresponding equations (with prescribed displacement at the border) can be written as

$$\begin{aligned}\mathrm{L} - \nabla\, \boldsymbol{u} &= 0 && \text{on } \Omega,\\ -\, \boldsymbol{\nabla\cdot}(\mu \mathrm{L}) + \boldsymbol{\nabla}\, p &= \boldsymbol{f} && \text{on } \Omega,\\ \varepsilon\, p + \nabla\cdot \boldsymbol{u} &= 0 && \text{on } \Omega,\\ \boldsymbol{u} &= \boldsymbol{g} && \text{on } \partial\Omega,\end{aligned}$$

where $\varepsilon = (1 - 2\,\nu)(1 + \nu)/E$. Here E is the Young modulus and $\nu \in (0, 1/2]$ is the Poisson ratio. The advantage of this formulation is that is holds for both compressible ($\nu \in (0, 1/2)$) and incompressible ($\nu = 1/2$) materials. Examples of methods using this formulation are the Galerkin least-squares method introduced and studied in [22] and the LDG method considered in [19].

Clearly, for incompressible materials, this system of equations is essentially the same as the Stokes system (1.1). It is thus reasonable to expect that the HDG methods for this formulation will have the same convergence properties.

TABLE 6.2. History of convergence of $\|\boldsymbol{u}-\boldsymbol{u}_h^{\mathrm{RT},\star}\|_{\mathcal{T}_h}$ and $\|\boldsymbol{u}-\boldsymbol{u}_h^{\mathrm{CR},\star}\|_{\mathcal{T}_h}$ for the special case $k=0$ for the HDG method.

	$\|\boldsymbol{u}-\boldsymbol{u}_h^{\mathrm{RT},\star}\|_{\mathcal{T}_h}$						$\|\boldsymbol{u}-\boldsymbol{u}_h^{\mathrm{CR},\star}\|_{\mathcal{T}_h}$					
mesh	$\tau_t=h$		$\tau_t=1$		$\tau_t=1/h$		$\tau_t=h$		$\tau_t=1$		$\tau_t=1/h$	
$2h^{-1}$	error	order	error	order	error	order	error	order	error	order	error	order
$\tau_n=h$												
4	6.32e-0	−−	6.32e-0	−−	6.32e-0	−−	6.97e-0	−−	7.14e-0	−−	7.17e-0	−−
8	3.23e-0	0.97	3.42e-0	0.89	3.61e-0	0.81	3.62e-0	0.95	4.41e-0	0.69	4.84e-0	0.57
16	1.60e-0	1.01	1.65e-0	1.05	1.83e-0	0.98	1.24e-0	1.55	1.70e-0	1.37	2.06e-0	1.23
32	8.30e-1	0.95	8.46e-1	0.97	1.06e-0	0.79	4.66e-1	1.41	6.08e-1	1.49	9.30e-1	1.15
64	4.22e-1	0.98	4.28e-1	0.98	7.17e-1	0.57	2.03e-1	1.20	2.36e-1	1.37	6.29e-1	0.56
$\tau_n=1$												
4	6.34e-0	−−	6.34e-0	−−	6.33e-0	−−	6.06e-0	−−	6.14e-0	−−	6.15e-0	−−
8	3.17e-0	1.00	3.31e-0	0.94	3.57e-0	0.83	2.69e-0	1.17	3.11e-0	0.98	3.50e-0	0.81
16	1.59e-0	0.99	1.63e-0	1.02	1.87e-0	0.93	9.19e-1	1.55	1.16e-0	1.43	1.62e-0	1.11
32	8.27e-1	0.94	8.48e-1	0.94	1.13e-0	0.72	4.08e-1	1.17	4.96e-1	1.22	9.46e-1	0.77
64	4.21e-1	0.97	4.34e-1	0.97	7.83e-1	0.53	1.95e-1	1.07	2.30e-1	1.11	6.99e-1	0.44
$\tau_n=1/h$												
4	6.36e-0	−−	6.36e-0	−−	6.36e-0	−−	5.72e-0	−−	5.74e-0	−−	5.73e-0	−−
8	3.15e-0	1.01	3.23e-0	0.98	3.46e-0	0.88	2.48e-0	1.21	2.63e-0	1.12	2.91e-0	0.98
16	1.59e-0	0.99	1.59e-0	1.02	1.84e-0	0.91	8.91e-1	1.47	9.42e-1	1.48	1.35e-0	1.11
32	8.26e-1	0.94	8.22e-1	0.95	1.18e-0	0.64	4.07e-1	1.13	4.25e-1	1.15	9.65e-1	0.48
64	4.21e-1	0.97	4.18e-1	0.97	9.18e-1	0.37	1.95e-1	1.06	1.99e-1	1.10	8.46e-1	0.19

APPENDIX A. A CHARACTERIZATION OF THE BDM PROJECTION IN 3D

It is well known that the BDM projection of a sufficiently smooth function $\boldsymbol{u}$ into $\boldsymbol{\mathcal{P}}_{k+1}(K)$, $k \geq 0$, is determined by the following equations:

$$\begin{aligned} \langle(\boldsymbol{\Pi}^{\mathrm{BDM}}\boldsymbol{u}-\boldsymbol{u})\cdot\boldsymbol{n},\mu\rangle_F &= 0 \quad \forall\, \mu\in\mathcal{P}_{k+1}(F) \quad \forall F \text{ face of } K,\\ (\boldsymbol{\Pi}^{\mathrm{BDM}}\boldsymbol{u}-\boldsymbol{u},\boldsymbol{\nabla}\, w)_K &= 0 \quad \forall\, w\in\boldsymbol{\mathcal{P}}_k(K),\\ (\boldsymbol{\Pi}^{\mathrm{BDM}}\boldsymbol{u}-\boldsymbol{u},\boldsymbol{p})_K &= 0 \quad \forall\, \boldsymbol{p}\in\boldsymbol{\Phi}_{k+1}(K), \end{aligned}$$

where $\boldsymbol{\Phi}_{k+1}(K)$ is the set of polynomials in $\boldsymbol{\mathcal{P}}_{k+1}(K)$ that are divergence-free and whose normal component is zero on ∂K; see [4].

Next, we give an alternative characterization of this projection motivated by the need to define our post-processed approximate velocity. The result itself is a consequence of a characterization of $\boldsymbol{\Phi}_{k+1}(K)$ that can be derived from [12, Lemma 2.4].

Proposition A.1. *For any $k\geq 0$, we have that the* BDM*-projection onto $\boldsymbol{\mathcal{P}}_{k+1}(K)$ is characterized by the equations*

$$\begin{aligned} \langle(\boldsymbol{\Pi}^{\mathrm{BDM}}\boldsymbol{u}-\boldsymbol{u})\cdot\boldsymbol{n},\mu\rangle_F &= 0 \quad \forall\, \mu\in\mathcal{P}_{k+1}(F) \quad \forall F \text{ face of } K,\\ (\boldsymbol{\Pi}^{\mathrm{BDM}}\boldsymbol{u}-\boldsymbol{u},\boldsymbol{\nabla}\, w)_K &= 0 \quad \forall\, w\in\boldsymbol{\mathcal{P}}_k(K),\\ (\boldsymbol{\nabla}\times(\boldsymbol{\Pi}^{\mathrm{BDM}}\boldsymbol{u}-\boldsymbol{u}),(\boldsymbol{\nabla}\times\boldsymbol{v})\,\mathrm{B}_K)_K &= 0 \quad \forall\, \boldsymbol{v}\in\boldsymbol{\mathcal{S}}_k(K). \end{aligned}$$

Proof. The result follows if we prove that

$$\boldsymbol{\Phi}_{k+1}(K)=\{\boldsymbol{\nabla}\times((\boldsymbol{\nabla}\times\boldsymbol{v})\mathrm{B}_K)\,:\,\boldsymbol{v}\in\boldsymbol{\mathcal{S}}_k(K)\} \tag{A.1}$$

and that the tangential components of $(\boldsymbol{\nabla}\times\boldsymbol{v})\mathrm{B}_{\mathrm{K}}$ on ∂K are zero.

Let us first prove the second assertion. If we take any vector-valued function $\boldsymbol{a}$, by the definition of the symmetric matrix bubble B_K, we have that, on the face $\lambda_i=0$,

$$\boldsymbol{a}\mathrm{B}_K=\lambda_{i-3}\lambda_{i-2}\lambda_{i-1}(\boldsymbol{a}\cdot\boldsymbol{\nabla}\,\lambda_i)\,\boldsymbol{\nabla}\,\lambda_i,$$

and we immediately see that the tangential component is identically equal to zero.

The proof of (A.1) is done in several steps. First, it is clear that functions of the form $\boldsymbol{\nabla}\times(\boldsymbol{a}\mathrm{B}_K)$ are divergence free. As a second step we are going to prove that all of these functions have zero normal component on ∂K. By symmetry, it is enough to prove that

$$\nabla\lambda_3 \cdot (\boldsymbol{\nabla}\times(\boldsymbol{a}\mathrm{B}_K))\Big|_{\lambda_3=0} = 0. \tag{A.2}$$

To prove this, note first that by the chain rule, we can write

$$\nabla\lambda_3 \cdot (\boldsymbol{\nabla}\times \boldsymbol{w}) = \sum_{m=0}^{2} \frac{\partial}{\partial\lambda_m}\Big(\nabla\lambda_3 \cdot (\nabla\lambda_m \times \boldsymbol{w})\Big).$$

If we apply this formula to

$$\boldsymbol{w} = \boldsymbol{a}\mathrm{B}_{\mathrm{K}} = \sum_{\ell=0}^{3} \lambda_{\ell-3}\lambda_{\ell-2}\lambda_{\ell-1}(\boldsymbol{a}\cdot\nabla\lambda_\ell)\nabla\lambda_\ell$$

the term for $\ell = 3$ vanishes as well as those with $\ell = m$. We therefore have to prove that the functions

$$\sum_{\ell=0}^{2}\left(\sum_{\substack{m=0\\ m\neq\ell}}^{2}\Big(\nabla\lambda_3\cdot(\nabla\lambda_m\times\nabla\lambda_\ell)\Big)\frac{\partial}{\partial\lambda_m}\Big(\lambda_{\ell-3}\lambda_{\ell-2}\lambda_{\ell-1}(\boldsymbol{a}\cdot\nabla\lambda_\ell)\Big)\right) =: \sum_{\ell=0}^{2} f_\ell$$

vanish at $\lambda_3 = 0$. This is actually true for each f_ℓ separately. We show in detail the case $\ell = 0$ (the other two follow by symmetry). Simple arguments show that

$$\begin{aligned}
f_0\Big|_{\lambda_3=0} &= \sum_{m=1}^{2}\big(\nabla\lambda_3\cdot(\nabla\lambda_m\times\nabla\lambda_0)\big)\frac{\partial}{\partial\lambda_m}\Big(\lambda_1\lambda_2\lambda_3(\boldsymbol{a}\cdot\nabla\lambda_0)\Big)\Big|_{\lambda_3=0}\\
&=\Big(\lambda_1\lambda_2\,(\boldsymbol{a}\cdot\nabla\lambda_0)\sum_{m=1}^{2}\big(\nabla\lambda_3\cdot(\nabla\lambda_m\times\nabla\lambda_0)\big)\frac{\partial\lambda_3}{\partial\lambda_m}\Big)\Big|_{\lambda_3=0}\\
&= -\Big(\sum_{m=1}^{2}\big(\nabla\lambda_3\cdot(\nabla\lambda_m\times\nabla\lambda_0)\big)\Big)\Big(\lambda_1\lambda_2(\boldsymbol{a}\cdot\nabla\lambda_0)\Big)\Big|_{\lambda_3=0}.
\end{aligned}$$

However,

$$\nabla\lambda_3\cdot(\nabla(\lambda_1+\lambda_2)\times\nabla\lambda_0) = -\nabla\lambda_3\cdot(\nabla(\lambda_0+\lambda_3-1)\times\nabla\lambda_0) = 0$$

which proves that $f_0 = 0$ on $\lambda_3 = 0$. This finishes the proof of (A.2).

Consider now the map

$$\begin{array}{rccc} T: & \boldsymbol{\mathcal{S}}_k(K) & \longrightarrow & \boldsymbol{\Phi}_{k+1}(K)\\ & \boldsymbol{v} & \longmapsto & \boldsymbol{\nabla}\times((\boldsymbol{\nabla}\times\boldsymbol{v})\mathrm{B}_K). \end{array}$$

We have just shown that functions of the form $\boldsymbol{\nabla}\times((\boldsymbol{\nabla}\times\boldsymbol{v})\mathrm{B}_K)$ are divergence free and have zero normal component on ∂K. It is simple to see that they are polynomials of degree no greater than $k+1$ when $\boldsymbol{v}$ is a vector-valued polynomial function of degree not greater than k. This means that T is well defined. By [27, Lemma 2], $\dim\boldsymbol{\mathcal{S}}_k(K) = \dim\boldsymbol{\Phi}_{k+1}(K)$, so if we prove that T is one–to–one, then it will be onto and (A.1) will have been proven.

If $T\boldsymbol{v} = \mathbf{0}$, then $0 = (\boldsymbol{v}, T\boldsymbol{v})_K = (\boldsymbol{\nabla}\times \boldsymbol{v}, (\boldsymbol{\nabla}\times \boldsymbol{v})\mathrm{B}_K)_K$. Since

$$\boldsymbol{a} \cdot (\boldsymbol{a}\mathrm{B}_K) = \sum_{\ell=0}^{2} \lambda_{\ell-1}\lambda_{\ell-2}\lambda_{\ell-3}|\boldsymbol{a} \cdot \nabla\lambda_\ell|^2,$$

we easily see that if $T\boldsymbol{v} = \mathbf{0}$, then $\boldsymbol{\nabla}\times \boldsymbol{v} = \mathbf{0}$ in K. Thus $\boldsymbol{v} = \boldsymbol{\nabla}\,\psi$ for some $\psi \in \mathcal{P}_{k+1}(K)$. However, since $\boldsymbol{\mathcal{S}}_k(K)$ is orthogonal to $\boldsymbol{\nabla}\,\mathcal{P}_{k+1}(K)$, we find that $\boldsymbol{v} = \mathbf{0}$, and the proof is complete. □

The set equality (A.1) is also true if we take $\boldsymbol{v}$ in the larger spaces $\boldsymbol{\mathcal{P}}_k(K)$ [12, Lemma 2.4] and $\boldsymbol{N}_k(K)$ [23, Lemma 2.3]. Actually, the spaces $\boldsymbol{\mathcal{S}}_k(K)$, $\boldsymbol{N}_k(K)$, and $\boldsymbol{\mathcal{P}}_k(K)$ differ only by gradients [27], which are in the null space of curl and therefore in the null space of the operator T in our proof. Our self-contained proof gives the smallest space which makes (A.1) true.

References

[1] 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** (2002), 1749–1779. MR1885715 (2002k:65183)

[2] P. Bastian and B. Rivière, *Superconvergence and $H(\mathrm{div})$ projection for discontinuous Galerkin methods*, Internat. J. Numer. Methods Fluids **42** (2003), 1043–1057. MR1991232 (2004f:65177)

[3] F. Brezzi, J. Douglas, Jr., and L. D. Marini, *Two families of mixed finite elements for second order elliptic problems*, Numer. Math. **47** (1985), 217–235. MR799685 (87g:65133)

[4] F. Brezzi and M. Fortin, *Mixed and hybrid finite element methods*, Springer-Verlag, 1991. MR1115205 (92d:65187)

[5] J. Carrero, B. Cockburn, and D. Schötzau, *Hybridized, globally divergence-free LDG methods. Part I: The Stokes problem*, Math. Comp. **75** (2006), 533–563. MR2196980 (2006m:76040)

[6] P. Castillo, *Performance of discontinuous Galerkin methods for elliptic PDE's*, SIAM J. Sci. Comput. **24** (2002), 524–547. MR1951054 (2003m:65200)

[7] B. Cockburn, B. Dong, and J. Guzmán, *A superconvergent LDG-hybridizable Galerkin method for second-order elliptic problems*, Math. Comp. **77** (2008), 1887–1916. MR2429868 (2009d:65166)

[8] B. Cockburn and J. Gopalakrishnan, *A characterization of hybridized mixed methods for second order elliptic problems*, SIAM J. Numer. Anal. **42** (2004), 283–301. MR2051067 (2005e:65183)

[9] ———, *Incompressible finite elements via hybridization. Part I: The Stokes system in two space dimensions*, SIAM J. Numer. Anal. **43** (2005), 1627–1650. MR2182142 (2006m:65262)

[10] ———, *Incompressible finite elements via hybridization. Part II: The Stokes system in three space dimensions*, SIAM J. Numer. Anal. **43** (2005), 1651–1672. MR2182143 (2006m:65263)

[11] ———, *The derivation of hybridizable discontinuous Galerkin methods for Stokes flow*, SIAM J. Numer. Anal. **47** (2009), 1092–1125. MR2485446

[12] B. Cockburn, J. Gopalakrishnan, and J. Guzmán, *A new elasticity element made for enforcing weak stress symmetry*, Math. Comp. **79** (2010), 1331–1349.

[13] B. Cockburn, J. Gopalakrishnan, and R. Lazarov, *Unified hybridization of discontinuous Galerkin, mixed and continuous Galerkin methods for second order elliptic problems*, SIAM J. Numer. Anal. **47** (2009), 1319–1365. MR2485455 (2010b:65251)

[14] B. Cockburn, J. Gopalakrishnan, and F.-J. Sayas, *A projection-based error analysis of HDG methods*, Math. Comp. **79** (2010), 1351–1367. MR2629996

[15] B. Cockburn, J. Guzmán, and H. Wang, *Superconvergent discontinuous Galerkin methods for second-order elliptic problems*, Math. Comp. **78** (2009), 1–24. MR2448694 (2009i:65213)

[16] B. Cockburn, G. Kanschat, and D. Schötzau, *A locally conservative LDG method for the incompressible Navier-Stokes equations*, Math. Comp. **74** (2005), 1067–1095. MR2136994 (2006a:65157)

[17] ———, *A note on discontinuous Galerkin divergence-free solutions of the Navier-Stokes equations*, J. Sci. Comput. **31** (2007), 61–73. MR2304270 (2008f:76109)

[18] B. Cockburn, G. Kanschat, D. Schötzau, and C. Schwab, *Local discontinuous Galerkin methods for the Stokes system*, SIAM J. Numer. Anal. **40** (2002), 319–343. MR1921922 (2003g:65141)

[19] B. Cockburn, D. Schötzau, and J. Wang, *Discontinuous Galerkin methods for incompressible elastic materials*, Comput. Methods Appl. Mech. Engrg. **195** (2006), 3184–3204, C. Dawson, Ed. MR2220915 (2006m:74052)

[20] M. Dauge, *Stationary Stokes and Navier-Stokes systems on two- or three-dimensional domains with corners. I. Linearized equations*, SIAM J. Math. Anal. **20** (1989), no. 1, 74–97. MR977489 (90b:35191)

[21] M. Farhloul, *Mixed and nonconforming finite element methods for the Stokes problem*, Canad. Appl. Math. Quart. **3** (1995), 399–418. MR1372793 (97c:76036)

[22] L. P. Franca and R. Stenberg, *Error analysis of some Galerkin least squares methods for the elasticity equations*, SIAM J. Numer. Anal. **28** (1991), 1680–1697. MR1135761 (92k:73066)

[23] J. Gopalakrishnan and J. Guzmán, *A second elasticity element using the matrix bubble with tightened stress symmetry*, Submitted, (2009).

[24] R. B. Kellogg and J. E. Osborn, *A regularity result for the Stokes problem in a convex polygon*, J. Functional Analysis **21** (1976), no. 4, 397–431. MR0404849 (53:8649)

[25] L. I. G. Kovasznay, *Laminar flow behind a two-dimensional grid*, Proc. Camb. Philos. Soc. **44** (1948), 58–62. MR0024282 (9:476d)

[26] J.-C. Nédélec, *Mixed finite elements in* $\mathbf{R}^3$, Numer. Math. **35** (1980), 315–341. MR592160 (81k:65125)

[27] J.-C. Nédélec, *A new family of mixed finite elements in* $\mathbf{R}^3$, Numer. Math. **50** (1986), 57–81. MR864305 (88e:65145)

[28] N.C. Nguyen, J. Peraire, and B. Cockburn, *A hybridizable discontinuous Galerkin method for Stokes flow*, Comput. Methods Appl. Mech. Engrg. **199** (2010), 582–597.

[29] P. A. Raviart and J. M. Thomas, *A mixed finite element method for second order elliptic problems*, Mathematical Aspects of Finite Element Method, Lecture Notes in Math. 606 (I. Galligani and E. Magenes, eds.), Springer-Verlag, New York, 1977, pp. 292–315. MR0483555 (58:3547)

[30] S. J. Sherwin, R. M. Kirby, J. Peiró, R. L. Taylor, and O. C. Zienkiewicz, *On 2D elliptic discontinuous Galerkin methods*, Internat. J. Numer. Methods Engrg. **65** (2006), no. 5, 752–784. MR2195978 (2007b:65127)

[31] R. Stenberg, *Some new families of finite elements for the Stokes equations*, Numer. Math. **56** (1990), 827–838. MR1035181 (91d:65176)

School of Mathematics, University of Minnesota, Minneapolis, Minnesota 55455
E-mail address: `cockburn@math.umn.edu`

Department of Mathematics, University of Florida, Gainesville, Florida 32611–8105
E-mail address: `jayg@math.ufl.edu`

Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge Massachusetts 02139
E-mail address: `cuongng@mit.edu`

Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge Massachusetts 02139
E-mail address: `peraire@mit.edu`

Departamento de Matemática Aplicada, Universidad de Zaragoza, 50009 Zaragoza, Spain
E-mail address: `sayas002@umn.edu`