J Sci Comput (2013) 55:392–431
DOI 10.1007/s10915-012-9639-y

Analysis of HDG Methods for Oseen Equations

Aycil Cesmelioglu · Bernardo Cockburn ·
Ngoc Cuong Nguyen · Jaume Peraire

Received: 29 January 2012 / Revised: 16 July 2012 / Accepted: 28 August 2012 /
Published online: 8 September 2012

Abstract We propose a hybridizable discontinuous Galerkin (HDG) method to numerically solve the Oseen equations which can be seen as the linearized version of the incompressible Navier-Stokes equations. We use same polynomial degree to approximate the velocity, its gradient and the pressure. With a special projection and postprocessing, we obtain optimal convergence for the velocity gradient and pressure and superconvergence for the velocity. Numerical results supporting our theoretical results are provided.

Keywords Oseen equations · Discontinuous Galerkin methods · Hybridizable · Postprocessing · Superconvergence

1 Introduction

This paper presents an a priori error analysis of a hybridizable discontinuous Galerkin (HDG) method for the Oseen problem. Oseen equations can be thought of as a linearized

B. Cockburn supported in part by the National Science Foundation (Grant DMS-0712955).

N.C. Nguyen, J. Peraire supported in part by the Singapore-MIT Alliance.

A. Cesmelioglu (✉)
Department of Mathematics and Statistics, Oakland University, Rochester, MI 48309, USA
e-mail: cesmelio@oakland.edu

B. Cockburn
School of Mathematics, University of Minnesota, Minneapolis, MN 55455, USA
e-mail: cockburn@math.umn.edu

N.C. Nguyen · J. Peraire
Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

N.C. Nguyen
e-mail: cuongng@mit.edu

J. Peraire
e-mail: peraire@mit.edu

version of the incompressible Navier-Stokes equations. In fact, the only difference between these two equations is in the convection term. A common practice to approximate the Navier-Stokes flow is to use Picard iterations which results in a sequence of Oseen equations. Hence, this paper is a natural extension of [7, 17] towards the goal of analyzing the HDG approximation to the Navier-Stokes equations.

Let us briefly discuss the HDG methods. The DG methods for elliptic problem [2], when compared to mixed methods, have been criticized [6] to have suboptimal convergence orders and that they give higher number of globally coupled degrees of freedom for the same size mesh and yield lower sparsity for the stiffness matrices. The HDG methods were introduced in [14] to address those criticisms. What makes HDG methods attractive is that they significantly reduce these problems while keeping the superconvergence properties that are observed with classical mixed methods. Roughly, the unknowns are functions and their derivatives on the domain and traces of these functions along the interior boundaries of a given mesh. Then, the globally coupled degree of freedoms are only those of the numerical traces. This makes the implementation efficient by giving a sparser form for the related stiffness matrix. This makes the HDG methods competitive with BDM [4] and RT [21] mixed methods. The reader may refer to [14] for a more elaborate discussion.

Oseen equations have been numerically analyzed in [11] using LDG method and in [5] using continuous interior penalty finite element method. The first method yields optimal a priori estimates for the errors in the pressure and velocity when using polynomials of degree k for the velocity and $k-1$ for the pressure. In [5], only quasi optimal error estimates are given when equal degree polynomials are used for the velocity and pressure. In [17], HDG approximation applied to the Stokes problem has been analyzed by using a special projection designed to fit the structure of the numerical traces. Then, [10] simplifies the definition of this projection by using L^2-projection for the pressure. In [7], HDG method applied to the convection-diffusion problem has been analyzed. In the Oseen case, if we let the convective velocity to be zero, then we get the Stokes problem whereas when we eliminate the pressure term we get a convection-diffusion problem. Hence, the projection we use in this paper is a hybrid version of the projections used in [7, 10, 17].

The organization of the rest of paper is as follows. We continue this section by defining the incompressible Oseen problem in velocity gradient-velocity-pressure form and introducing general notation. In the next section, we formulate our HDG approximation. There we introduce the stabilization tensor S and the numerical traces. Then, a projection defined specifically for this problem is constructed and its approximation properties are stated. Next, we define the postprocessing which enables us to get superconvergence in the velocity approximation. We end the results section by stating our main theorem that the errors are optimal and that we have superconvergence of the velocity after the postprocessing step. In Sect. 3, we give detailed proofs for all the results stated in the previous section. We would like to note that our method is not robust with respect to the Reynolds number. This can be seen in the numerical results section where we use the well-known Kovasznay flow, an analytical solution for the Navier-Stokes equations. We apply our method with two different Reynolds numbers and elaborate on the agreement of the numerical results with our theoretical analysis. We also discuss the choice of the stabilization parameters on the accuracy of the scheme. Finally, we conclude by a brief summary and possible extensions. The appendix is dedicated to the approximation properties of the auxiliary projection introduced in Sect. 2.

1.1 Oseen Equations

Let $\Omega \subset \mathbb{R}^n$ be a polygonal domain for $n = 2$ or a Lipschitz polyhedral domain for $n = 3$. The incompressible Oseen equations are given by the following set of equations and boundary condition:

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

$$-\nu \nabla \cdot \mathrm{L} + (\boldsymbol{\beta} \cdot \nabla)\boldsymbol{u} + \nabla p = \boldsymbol{f} \quad \text{in } \Omega, \tag{1.1b}$$

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

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

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

where $\boldsymbol{u}$ is the velocity, p is the pressure, ν is the kinematic viscosity and $\boldsymbol{f} \in L^2(\Omega)^n$ is the external body force. The Dirichlet boundary datum $\boldsymbol{g} \in H^{1/2}(\partial\Omega)^n$ is assumed to satisfy

$$\int_{\partial\Omega} \boldsymbol{g} \cdot \boldsymbol{n} = 0,$$

for compatibility. Finally the convective velocity $\boldsymbol{\beta}$ is assumed to be divergence free for simplicity and we further suppose that $\boldsymbol{\beta} \in W^{1,\infty}(\Omega)^n$.

The rest of this section is devoted for the notation to be used throughout the paper. For a given function space X, we denote $X^{n\times n}$ by X and X^n by $\boldsymbol{X}$. We use the usual definitions [1] for the Sobolev spaces $W^{k,p}(D)$ for a given domain D with norm

$$\|\phi\|_{k,p,D} = \left(\sum_{|\alpha|\le k} \left\| D^\alpha \phi \right\|_{0,p,D}^p \right)^{1/p}.$$

For vector and matrix valued functions $\boldsymbol{\phi}$ and Φ, we use

$$\|\boldsymbol{\phi}\|_{k,p,D} = \sum_{i=1}^n \|\boldsymbol{\phi}_i\|_{k,p,D}, \qquad \|\Phi\|_{k,p,D} = \sum_{i,j=1}^n \|\Phi\|_{k,p,D}.$$

Specifically, when $p = 2$ and $k < \infty$, we denote $W^{k,2}(D)$ by $H^k(D)$ and $\|\cdot\|_{k,2,D}$, by $\|\cdot\|_{k,D}$, when $k = 0$ and $p = 2$, we denote $W^{0,2}(D)$ by $L^2(D)$ and the norm by $\|\cdot\|_D$. The space of polynomials of degree at most k defined on a simplex K is denoted by $P_k(K)$. Furthermore, $P_k^\perp(K)$ denotes the space of polynomials on K of degree at most k which are perpendicular to the space of polynomials of order at most $k-1$, i.e.,

$$P_k^\perp(K) := \left\{ p \in P_k(K) : (p, q)_K = 0,\ \forall q \in P_{k-1}(K) \right\}.$$

Similarly, on a face F,

$$P_k^\perp(F) := \left\{ p \in P_k(F) : \langle p, q \rangle_F = 0,\ \forall q \in P_{k-1}(F) \right\}.$$

For a given matrix M, the transpose of M is given by M^T.

2 Results

This section introduces the HDG approximation for the Oseen problem and states the results.

2.1 HDG Method for Oseen Problem

The triangulation $\mathcal{T}_h$ of Ω is assumed to be shape regular and conforming. We denote the set of all faces F of all elements $K \in \mathcal{T}_h$ by $\mathcal{E}_h$, all the interior faces by $\mathcal{E}_h^0$ and all the boundary faces by $\mathcal{E}_h^\partial$. We equip these spaces with

$$(\phi, \psi)_{\mathcal{T}_h} := \sum_{K \in \mathcal{T}_h} (\phi, \psi)_K, \qquad \langle \phi, \psi \rangle_{\partial \mathcal{T}_h} := \sum_{K \in \mathcal{T}_h} \langle \phi, \psi \rangle_{\partial K}$$

for scalar functions, with

$$(\boldsymbol{\phi}, \boldsymbol{\psi})_{\mathcal{T}_h} := \sum_{i=1}^{n} (\phi_i, \psi_i)_{\mathcal{T}_h}, \qquad \langle \boldsymbol{\phi}, \boldsymbol{\psi} \rangle_{\partial \mathcal{T}_h} := \sum_{i=1}^{n} \langle \phi_i, \psi_i \rangle_{\partial \mathcal{T}_h}$$

for vector valued functions and with

$$(\Phi, \Psi)_{\mathcal{T}_h} := \sum_{i,j=1}^{n} (\Phi_{ij}, \Psi_{ij})_{\mathcal{T}_h}, \qquad \langle \Phi, \Psi \rangle_{\partial \mathcal{T}_h} := \sum_{i,j=1}^{n} \langle \Phi_{ij}, \Psi_{ij} \rangle_{\partial \mathcal{T}_h}$$

for matrix valued functions. We denote the norm deduced from $(\cdot, \cdot)_{\mathcal{T}_h}$ by $\| \cdot \|_{\mathcal{T}_h}$ and from $(\cdot, \cdot)_{\partial \mathcal{T}_h}$ by $\| \cdot \|_{\partial \mathcal{T}_h}$. We will also use the following broken Sobolev spaces:

$$H^1(\mathcal{T}_h) := \prod_{K \in \mathcal{T}_h} H^1(K), \qquad \mathrm{H}^1(\mathcal{T}_h) = H^1(\mathcal{T}_h)^{d \times d}, \qquad \boldsymbol{H}^1(\mathcal{T}_h) = H^1(\mathcal{T}_h)^d.$$

Let us first introduce the approximation spaces for the velocity gradient L, velocity $\boldsymbol{u}$, pressure p and the velocity trace $\boldsymbol{u}|_{\mathcal{E}_h}$.

$$\begin{aligned}
\mathrm{G}_h &:= \left\{ \mathrm{G} \in \mathrm{L}^2(\mathcal{T}_h) : \mathrm{G}|_K \in \mathrm{P}_k(K),\ \forall K \in \mathcal{T}_h \right\}, \\
\boldsymbol{V}_h &:= \left\{ \boldsymbol{v} \in \boldsymbol{L}^2(\mathcal{T}_h) : \boldsymbol{v}|_K \in \boldsymbol{P}_k(K),\ \forall K \in \mathcal{T}_h \right\}, \\
P_h &:= \left\{ p \in L^2(\mathcal{T}_h) : p|_K \in P_k(K),\ \forall K \in \mathcal{T}_h \right\}, \\
\boldsymbol{M}_h &:= \left\{ \boldsymbol{\mu} \in \boldsymbol{L}^2(\mathcal{E}_h) : \boldsymbol{\mu}|_K \in \boldsymbol{P}_k(F),\ \forall F \in \mathcal{E}_h \right\}.
\end{aligned}$$

Because we are dealing with discontinuous polynomials, for an interior face $F = \partial K^- \cap \partial K^+$ in $\mathcal{E}_h^0$, we define jump and average by

$$\{\!\{\mathrm{G}\}\!\} = \frac{1}{2}\big(\mathrm{G}^+ + \mathrm{G}^-\big), \qquad [\![\boldsymbol{\phi} \otimes \boldsymbol{n}]\!] = \boldsymbol{\phi}^+ \otimes \boldsymbol{n}^+ + \boldsymbol{\phi}^- \otimes \boldsymbol{n}^-, \qquad [\![\mathrm{G}\boldsymbol{n}]\!] = \mathrm{G}^+\boldsymbol{n}^+ + \mathrm{G}^-\boldsymbol{n}^-,$$

where $^+$ and $^-$ denote the trace of the function from inside of K^+ and K^-, respectively. By convention, we extend these definitions to the boundary faces $F \in \mathcal{E}_h^\partial$ as follows:

$$\{\!\{\mathrm{G}\}\!\} = \mathrm{G}, \qquad [\![\boldsymbol{\phi} \otimes \boldsymbol{n}]\!] = \boldsymbol{\phi}, \qquad [\![\mathrm{G}\boldsymbol{n}]\!] = \mathrm{G}\boldsymbol{n}.$$

With the definitions of the approximation spaces, the method can be stated as follows:

Find $(\mathrm{L}_h, \boldsymbol{u}_h, p_h, \widehat{\boldsymbol{u}}_h) \in \mathrm{G}_h \times \boldsymbol{V}_h \times P_h \times \boldsymbol{M}_h$ such that

$$(\mathrm{L}_h, \mathrm{G})_{\mathcal{T}_h} + (\boldsymbol{u}_h, \nabla \cdot \mathrm{G})_{\mathcal{T}_h} - \langle \widehat{\boldsymbol{u}}_h, \mathrm{G}\boldsymbol{n} \rangle_{\partial \mathcal{T}_h} = 0, \tag{2.1a}$$

$$\begin{aligned}(\nu \mathrm{L}_h, \nabla \boldsymbol{v})_{\mathcal{T}_h} - (\boldsymbol{u}_h \otimes \boldsymbol{\beta}, \nabla \boldsymbol{v})_{\mathcal{T}_h} - (p_h, \nabla \cdot \boldsymbol{v})_{\mathcal{T}_h} \\ - \left\langle \nu \widehat{\mathrm{L}}_h \boldsymbol{n} - \widehat{p}_h \boldsymbol{n} - (\widehat{\boldsymbol{u}_h \otimes \boldsymbol{\beta}}) \boldsymbol{n}, \boldsymbol{v} \right\rangle_{\partial \mathcal{T}_h} = (\boldsymbol{f}, \boldsymbol{v})_{\mathcal{T}_h},\end{aligned} \tag{2.1b}$$

$$-(\boldsymbol{u}_h, \nabla q)_{\mathcal{T}_h} + \langle \widehat{\boldsymbol{u}}_h \cdot \boldsymbol{n}, q \rangle_{\partial \mathcal{T}_h} = 0, \tag{2.1c}$$

$$\langle \widehat{\boldsymbol{u}}_h, \boldsymbol{\mu} \rangle_{\partial \Omega} = \langle \boldsymbol{g}, \boldsymbol{\mu} \rangle_{\partial \Omega}, \tag{2.1d}$$

$$\left\langle \nu \widehat{\mathrm{L}}_h \boldsymbol{n} - \widehat{p}_h \boldsymbol{n} - (\widehat{\boldsymbol{u}_h \otimes \boldsymbol{\beta}}) \boldsymbol{n}, \boldsymbol{\mu} \right\rangle_{\partial \mathcal{T}_h \backslash \partial \Omega} = 0, \tag{2.1e}$$

$$(p_h, 1)_{\Omega} = 0, \tag{2.1f}$$

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 the trace is defined to be

$$\nu \widehat{\mathrm{L}}_h \boldsymbol{n} - \widehat{p}_h \boldsymbol{n} - (\widehat{\boldsymbol{u}_h \otimes \boldsymbol{\beta}}) \boldsymbol{n} = \nu \mathrm{L}_h \boldsymbol{n} - p_h \boldsymbol{n} - (\widehat{\boldsymbol{u}}_h \otimes \boldsymbol{\beta}) \boldsymbol{n} - \mathrm{S}(\boldsymbol{u}_h - \widehat{\boldsymbol{u}}_h), \quad \text{on } \partial \mathcal{T}_h, \tag{2.1g}$$

where S, to be defined in the next section, ensures the stability of the scheme. To simplify the notation we define the following linear mapping $\boldsymbol{F}$ on $\mathrm{G}_h \times \boldsymbol{V}_h \times P_h \times \boldsymbol{M}_h$ for the trace of the flux :

$$\boldsymbol{F}(\mathrm{L}, \boldsymbol{u}, p, \widehat{\boldsymbol{u}}) := \nu \mathrm{L}\boldsymbol{n} - p\boldsymbol{n} - (\widehat{\boldsymbol{u}} \otimes \boldsymbol{\beta})\boldsymbol{n} - \mathrm{S}(\boldsymbol{u} - \widehat{\boldsymbol{u}}).$$

With this notation, (2.1b) and (2.1e) can be written as

$$\begin{aligned}(\nu \mathrm{L}_h, \nabla \boldsymbol{v})_{\mathcal{T}_h} - (\boldsymbol{u}_h \otimes \boldsymbol{\beta}, \nabla \boldsymbol{v})_{\mathcal{T}_h} - (p_h, \nabla \cdot \boldsymbol{v})_{\mathcal{T}_h} - \left\langle \boldsymbol{F}(\mathrm{L}_h, \boldsymbol{u}_h, p_h, \widehat{\boldsymbol{u}}_h), \boldsymbol{v} \right\rangle_{\partial \mathcal{T}_h} \\ = (\boldsymbol{f}, \boldsymbol{v})_{\mathcal{T}_h} \quad \forall \boldsymbol{v} \in \boldsymbol{V}_h\end{aligned} \tag{2.2}$$

$$\left\langle \boldsymbol{F}(\mathrm{L}_h, \boldsymbol{u}_h, p_h, \widehat{\boldsymbol{u}}_h), \boldsymbol{\mu} \right\rangle_{\partial \mathcal{T}_h \backslash \partial \Omega} = 0, \quad \forall \boldsymbol{\mu} \in \boldsymbol{M}_h. \tag{2.3}$$

Remark 2.1 If we pick $\mathrm{G} = q\mathrm{I}$, $q \in P_k(K)$ in (2.1a) and use (2.1c), we can deduce that $\operatorname{tr} \mathrm{L}_h = 0$. This means that we have a divergence-free approximate solution when tested with the pressure space. This reflects the divergence-free condition on the exact velocity.

2.2 Stabilization Tensor S

Let us define the stabilization tensor S that shows up in the definition of the trace (2.1g).

$$\mathrm{S} := \nu \tau_n \boldsymbol{n} \otimes \boldsymbol{n} + \nu \tau_t (\mathrm{I} - \boldsymbol{n} \otimes \boldsymbol{n}) \quad \text{on } \partial K, \tag{2.4}$$

where the parameters τ_n and τ_t, which control the normal and tangential components of the jumps in the approximate velocity, are assumed to be nonnegative constants on ∂K. The matrix-valued function S is obviously symmetric and constant on each face of K. It is also positive semidefinite as

$$\langle \mathrm{S}\boldsymbol{v}, \boldsymbol{v} \rangle_{\partial K} = \nu \tau_n \|\boldsymbol{n} \cdot \boldsymbol{v}\|_{\partial K}^2 + \nu \tau_t \|\boldsymbol{\tau} \cdot \boldsymbol{v}\|_{\partial K}^2 \geq 0, \quad \boldsymbol{v} \in \mathbb{R}^n.$$

Fix an edge F of K. Then, $\mathrm{S}|_F$ has two eigenvalues; $\nu\tau_n$, the eigenvalue with scalar multiples of the unit normal vector $\boldsymbol{n}_F$ as eigenvectors and $\nu\tau_t$, the eigenvalue with scalar multiples of any unit tangent vector $\boldsymbol{\tau}_F$ of K as eigenvectors.

As we will see in the error analysis, we will need to make the following assumption:

$$\min\left(\nu\tau_n - \frac{1}{2}(\boldsymbol{\beta}\cdot\boldsymbol{n})\right)\bigg|_{\partial K} \geq \gamma > 0, \qquad \min\left(\nu\tau_t - \frac{1}{2}(\boldsymbol{\beta}\cdot\boldsymbol{n})\right)\bigg|_{\partial K} \geq \gamma > 0, \tag{2.5}$$

for some positive constant γ. Note that this assumption is equivalent to $\mathrm{S}_{\boldsymbol{\beta}} := (\mathrm{S} - \frac{1}{2}(\boldsymbol{\beta}\cdot\boldsymbol{n})\mathrm{I})|_{\partial K}$ being positive definite.

Remark 2.2 The assumption that τ_t and τ_n are constant on ∂K is taken only for the sake of simplicity and it can easily be relaxed.

The eigenspaces of S are exactly the same for $\mathrm{S}_{\boldsymbol{\beta}}$ and indeed λ is an eigenvalue for S if and only if $\lambda - \frac{1}{2}(\boldsymbol{\beta}\cdot\boldsymbol{n})$ is an eigenvalue of $\mathrm{S}_{\boldsymbol{\beta}}$ with the same eigenvectors. Hence when we assume (2.5), we are assuming that the eigenvalues of $\mathrm{S}_{\boldsymbol{\beta}}$ on ∂K are positive.

Let us label all the faces of K by $\mathcal{F} := \{F_i, i = 1, \ldots, n+1\}$. As τ_t and τ_n are constants on ∂K, if we denote the maximum eigenvalue of $\mathrm{S}|_F$ over all $F \subset \mathcal{F}$ by $\Lambda_K^{\max}$, then $\Lambda_K^{\max} = \nu\max(\tau_n, \tau_t)$.

For $\tau_n \geq \tau_t$, we pick the unit normal vectors corresponding to n of the faces, say $\mathcal{B}_K = \{\boldsymbol{n}_{F_i}, i = 1, \ldots, n\}$ to be the basis of $\mathbb{R}^n$. Let us also denote the dual basis by $\mathcal{B}_K^* = \{\boldsymbol{n}_{F_i}^*, i = 1, \ldots, n\}$. The elements of the dual basis are the rows of the inverse of the matrix whose columns are the elements of $\mathcal{B}_K$. Similarly, if $\tau_t \geq \tau_n$, we pick a vertex and choose as $\mathcal{B}_K$ the unit vectors originating from this vertex along its associated edges. The dual basis $\mathcal{B}_K^*$ can be picked in a similar fashion. Let us denote by $\lambda_K^{\min}$ the minimum eigenvalue of S corresponding to the vectors in $\mathcal{B}_K$. In the case of $\tau_n \geq \tau_t$,

$$\lambda_K^{\min} = \nu\tau_n = \Lambda_K^{\max},$$

whereas when $\tau_t \geq \tau_n$,

$$\lambda_K^{\min} = \nu\tau_t = \Lambda_K^{\max}.$$

By our assumptions on τ_n and τ_t,

$$\Gamma_K^{\min} := \lambda_K^{\min} - \frac{1}{2}\max_{i=1,\ldots,n}\left(\max(\boldsymbol{\beta}\cdot\boldsymbol{n})|_{F_i}\right) > 0. \tag{2.6}$$

As the stabilization tensor S is constant on each face F of K, we can write the numerical traces explicitly in terms of $(\mathrm{L}_h, \boldsymbol{u}_h, p_h)$ as in [17]. The numerical trace for the velocity is

$$\widehat{\boldsymbol{u}}_h = -\mathrm{A}[\![\nu\mathrm{L}_h\boldsymbol{n} - p_h\boldsymbol{n}]\!] + \mathrm{A}\mathrm{S}^+\boldsymbol{u}_h^+ + \mathrm{A}\mathrm{S}^-\boldsymbol{u}_h^-,$$

where $\mathrm{A} := (\mathrm{S}^+ + \mathrm{S}^-)^{-1}$ and the numerical trace for $\nu\mathrm{L} - p\mathrm{I}$ is

$$\begin{aligned}\nu\widehat{\mathrm{L}}_h - \widehat{p}_h\mathrm{I} - (\widehat{\boldsymbol{u}_h\otimes\boldsymbol{\beta}}) &= \mathrm{S}^-\mathrm{A}\left(\nu\mathrm{L}_h^+ - p_h^+\mathrm{I}\right) + \mathrm{S}^+\mathrm{A}\left(\nu\mathrm{L}_h^- - p_h^-\mathrm{I}\right)\\ &\quad - \mathrm{S}^+\mathrm{A}\mathrm{S}^-[\![\boldsymbol{u}_h\otimes\boldsymbol{n}]\!] - \widehat{\boldsymbol{u}}_h\otimes\boldsymbol{\beta}.\end{aligned}$$

For the case where $\tau_t = \tau_n = \tau$ is constant on $\partial\mathcal{T}_h$, we can compare the HDG method with the LDG method applied to the Oseen equations [11], we see that the only difference is in the definition of the numerical traces. In the case of the LDG discretization, the velocity numerical trace does not depend on neither the velocity gradient nor the pressure.

2.3 A Special Projection Π_h

The error analysis of the HDG method relies on defining a special projection that is designed specifically for the numerical traces that appear in the HDG method. We define the following new projection inspired by the structure of the numerical trace of $(\nu\mathrm{L} - p\mathrm{I} + \boldsymbol{u}\otimes\boldsymbol{\beta})\boldsymbol{n}$, (2.1g).

For $(\mathrm{L}, \boldsymbol{u}, p) \in \mathrm{H}^1(\mathcal{T}_h) \times \boldsymbol{H}^1(\mathcal{T}_h) \times H^1(\mathcal{T}_h)$, define $\Pi_h(\mathrm{L}, \boldsymbol{u}, p) = (\Pi\mathrm{L}, \boldsymbol{\Pi u}, \Pi p) \in \mathrm{G}_h \times \boldsymbol{V}_h \times P_h$ such that for any $K \in \mathcal{T}_h$,

$$\big(\nu(\Pi\mathrm{L} - \mathrm{L}), \mathrm{G}\big)_K - \big((\boldsymbol{\Pi u} - \boldsymbol{u})\otimes\boldsymbol{\beta}, \mathrm{G}\big)_K = 0, \quad \forall \mathrm{G} \in \mathrm{P}_{k-1}(K), \tag{2.7a}$$

$$(\boldsymbol{\Pi u} - \boldsymbol{u}, \boldsymbol{v})_K = 0, \quad \forall \boldsymbol{v} \in \boldsymbol{P}_{k-1}(K), \tag{2.7b}$$

$$(\Pi p - p, q)_K = 0, \quad \forall q \in P_k(K), \tag{2.7c}$$

$$\big\langle \boldsymbol{F}(\Pi\mathrm{L} - \mathrm{L}, \boldsymbol{\Pi u} - \boldsymbol{u}, \Pi p - p, \boldsymbol{P}_M\boldsymbol{u} - \boldsymbol{u}), \boldsymbol{\mu}\big\rangle_F = 0, \quad \forall \boldsymbol{\mu} \in \boldsymbol{P}_k(F) \tag{2.7d}$$

for all faces F of K where $\boldsymbol{P}_M$ is the L^2-projection into $\boldsymbol{M}_h$. Even though it is not obvious from our notation, which is chosen for simplicity, $\Pi\mathrm{L}$ and $\boldsymbol{\Pi u}$ depend on L, $\boldsymbol{u}$ and p. Further, Π_h depends on $\boldsymbol{\beta}$, ν and S. Observe also that Πp is just the L^2-projection of p to the space $P_k(K)$. The next theorem shows that the projection Π_h defined above is well-posedness and satisfies good approximation properties. From now on C will denote a generic constant independent of h, L, $\boldsymbol{u}$ and p.

Theorem 2.3 (Well-posedness and the approximation properties of the projection Π_h) *Assume that τ_n and τ_t are nonnegative constants on ∂K which satisfy (2.5). Then, the problem (2.7a)–(2.7d) has a unique solution $\Pi_h(\mathrm{L}, \boldsymbol{u}, p)$. Moreover, for each $K \in \mathcal{T}_h$, this solution satisfies*

$$\|\Pi p - p\|_K \le C h_K^{k_p+1} |p|_{k_p+1,K}, \tag{2.8}$$

$$\|\boldsymbol{\Pi u} - \boldsymbol{u}\|_K \le C\left(\frac{\Lambda_K^{\max} + h_K|\boldsymbol{\beta}|_{\boldsymbol{W}^{1,\infty}(K)}}{\Gamma_K^{\min}}\right) h_K^{k_{\boldsymbol{u}}+1}|\boldsymbol{u}|_{k_{\boldsymbol{u}}+1,K} + \frac{C}{\Gamma_K^{\min}} h_K^{k_\sigma+1}\big|\nabla\cdot(\nu\mathrm{L} - p\mathrm{I})\big|_{k_\sigma,K}. \tag{2.9}$$

In addition, if $\operatorname{tr}\mathrm{L} = 0$, *then we have*

$$\nu\|\Pi\mathrm{L} - \mathrm{L}\|_K \le C\nu h_K^{k_\mathrm{L}+1}|\mathrm{L}|_{k_\mathrm{L}+1,K} + C\big(\nu\tau_t + (1+\nu)|\boldsymbol{\beta}|_{\boldsymbol{W}^{1,\infty}(K)}h_K\big)\|\boldsymbol{\Pi u} - \boldsymbol{u}\|_K + C\big(\nu\tau_t + h_K|\boldsymbol{\beta}|_{\boldsymbol{W}^{1,\infty}(K)}\big)h_K^{k_{\boldsymbol{u}}+1}|\boldsymbol{u}|_{k_{\boldsymbol{u}}+1,K} + \|\Pi p - p\|_K, \tag{2.10}$$

where $k_\mathrm{L}, k_{\boldsymbol{u}}, k_\sigma, k_p \in [0, k]$.

Remark 2.4 Let us say a few words on the approximation errors. First, as Πp is the L^2-projection, the pressure error is optimal regardless of what τ_n, τ_t and $\boldsymbol{\beta}$ are. As we see $\Gamma_K^{\min}$ in the denominator for the bound of $\|\boldsymbol{\Pi u} - \boldsymbol{u}\|_K$, it has an effect on the convergence properties. Consider $k_\mathrm{L} = k_{\boldsymbol{u}} = k_p = k \ge 0$. If $\Gamma_K^{\min} = \mathcal{O}(h_K)$ or $\frac{\Lambda_K^{\max} + h_K|\boldsymbol{\beta}|_{\boldsymbol{W}^{1,\infty}(K)}}{\Gamma_K^{\min}} = \mathcal{O}(h_K^{-1})$, then we can only have an order of k for $\|\boldsymbol{\Pi u} - \boldsymbol{u}\|_K$. In this case, if $\tau_t = \mathcal{O}(h_K)$, we can still have an order of $k+1$ for $\|\Pi\mathrm{L} - \mathrm{L}\|_K$, otherwise we do not have optimality. Likewise, if

$\Gamma_K^{\min} = \mathcal{O}(1)$ and $\frac{\Lambda_K^{\max} + h_K |\boldsymbol{\beta}|_{W^{1,\infty}(K)}}{\Gamma_K^{\min}} = \mathcal{O}(1)$, then $\|\boldsymbol{\Pi u} - \boldsymbol{u}\|_K$ has optimal order of convergence $k+1$. In this case, if $\tau_t = \mathcal{O}(h_K^{-1})$, then we lose an order for $\|\Pi \mathrm{L} - \mathrm{L}\|_K$ unless we have $\nu < C h_K |\boldsymbol{\beta}|_{W^{1,\infty}(K)}$, otherwise, we have optimal order.

Finally, note that the above convergence results do hold as the stabilization function τ_n tends to infinity. For the Stokes problem ($\boldsymbol{\beta} = 0$), the reader may refer to [17, Table 2.2].

2.4 A Priori Estimates

Let us define the projection of the approximation errors $\mathrm{E}^{\mathrm{L}} := \Pi \mathrm{L} - \mathrm{L}_h$, $\boldsymbol{E}^{\boldsymbol{u}} := \boldsymbol{\Pi u} - \boldsymbol{u}_h$, $E^p := \Pi p - p_h$ and $\boldsymbol{E}^{\widehat{\boldsymbol{u}}} = \boldsymbol{P}_M \boldsymbol{u} - \widehat{\boldsymbol{u}}_h$. Next we introduce the following dual problem to get the optimal L^2-error estimate for the velocity. Given $\boldsymbol{\theta} \in \boldsymbol{L}^2(\Omega)$, define

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

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

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

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

Assume that the solution to the dual problem satisfies the following regularity estimate:

$$\nu \|\Phi\|_{1,\Omega} + \nu \|\boldsymbol{\phi}\|_{2,\Omega} + \|\phi\|_{1,\Omega} \le C_r \|\boldsymbol{\theta}\|_{\Omega}, \tag{2.12}$$

for all $\boldsymbol{\theta} \in \boldsymbol{L}^2(\Omega)$. In the following, we define another projection $\Pi_h^*(\Phi, \boldsymbol{\phi}, \phi) = (\Pi^* \Phi, \boldsymbol{\Pi}^* \boldsymbol{\phi}, \Pi^* \phi) \in \mathrm{P}_k(K) \times \boldsymbol{P}_k(K) \times P_k(K)$ associated to the above dual problem. The definition closely follows the definition of the projection Π_h. For any $K \in \mathcal{T}_h$,

$$\nu\big(\Pi^*\Phi - \Phi, \mathrm{G}\big)_K + \big(\big(\boldsymbol{\Pi}^*\boldsymbol{\phi} - \boldsymbol{\phi}\big) \otimes \boldsymbol{\beta}, \mathrm{G}\big)_K = 0, \quad \forall \mathrm{G} \in \mathrm{P}_{k-1}(K), \tag{2.13a}$$

$$\big(\boldsymbol{\Pi}^*\boldsymbol{\phi} - \boldsymbol{\phi}, \boldsymbol{v}\big)_K = 0, \quad \forall \boldsymbol{v} \in \boldsymbol{P}_{k-1}(K), \tag{2.13b}$$

$$\big(\Pi^*\phi - \phi, q\big)_K = 0, \quad \forall q \in P_k(K), \tag{2.13c}$$

$$\big\langle \boldsymbol{F}^*\big(\Pi^*\Phi - \Phi, \boldsymbol{\Pi}^*\boldsymbol{\phi} - \boldsymbol{\phi}, \Pi^*\phi - \phi, \boldsymbol{P}_M\boldsymbol{\phi} - \boldsymbol{\phi}\big), \boldsymbol{\mu}\big\rangle_F = 0, \quad \forall \boldsymbol{\mu} \in \boldsymbol{P}_k(F), \tag{2.13d}$$

for all faces F of K where

$$\boldsymbol{F}^*(\Phi, \boldsymbol{\phi}, \phi, \widehat{\boldsymbol{\phi}}) := \nu \Phi \boldsymbol{n} + \phi \boldsymbol{n} + (\boldsymbol{\phi} \otimes \boldsymbol{\beta})\boldsymbol{n} - \mathrm{S}(\boldsymbol{\phi} - \widehat{\boldsymbol{\phi}}).$$

This $\boldsymbol{F}^*$ stands for the transposed flux compatible with the structure of the dual problem. Now we state the approximation properties of Π_h^*. The proof is very similar to the proof of the same results corresponding to the projection Π_h and is provided in Appendix 5.1.2. Let $\widetilde{\mathrm{S}}_{\boldsymbol{\beta}} := \mathrm{S} - (\boldsymbol{\beta} \cdot \boldsymbol{n})\mathrm{I}$. We set $\widetilde{\Lambda}_K^{\max}$ to be the maximum eigenvalue of $\widetilde{\mathrm{S}}_{\boldsymbol{\beta}}$ over all faces of K.

Theorem 2.5 (Well-posedness and approximation properties of Π_h^*) *Suppose that the same assumptions on τ_n and τ_t as in Theorem 2.3 hold. Then, the problem (2.13a)-(2.13d) has a unique solution $\Pi_h^*(\Phi, \boldsymbol{\phi}, \phi)$. Moreover, for each $K \in \mathcal{T}_h$, this solution satisfies*

$$\big\|\Pi^*\phi - \phi\big\|_K \le C h_K^{k_\phi + 1} |\phi|_{k_\phi+1,K},$$

$$\left\|\boldsymbol{\Pi}^*\boldsymbol{\phi}-\boldsymbol{\phi}\right\|_K \le C\left(\frac{\widetilde{\Lambda}_K^{\max}+h_K|\boldsymbol{\beta}|_{W^{1,\infty}(K)}}{\Gamma_K^{\min}}\right)h_K^{k_{\boldsymbol{\phi}}+1}|\boldsymbol{\phi}|_{k_{\boldsymbol{\phi}}+1,K} + \frac{C}{\Gamma_K^{\min}}h_K^{k_\sigma+1}\left|\nabla\cdot(\nu\Phi+\phi\mathrm{I})\right|_{k_\sigma,K}.$$

Furthermore, if $\operatorname{tr}\Phi=0$, *then*

$$\nu\left\|\Pi^*\Phi-\Phi\right\|_K \le C\nu h_K^{k_\Phi+1}|\Phi|_{k_\Phi+1,K} + C\big(\nu\tau_t+(1+\nu)|\boldsymbol{\beta}|_{W^{1,\infty}(K)}h_K\big)\left\|\boldsymbol{\Pi}^*\boldsymbol{\phi}-\boldsymbol{\phi}\right\|_K + C\big(\nu\tau_t+|\boldsymbol{\beta}|_{W^{1,\infty}(K)}h_K\big)h_K^{k_{\boldsymbol{\phi}}+1}|\boldsymbol{\phi}|_{k_{\boldsymbol{\phi}}+1,K} + \left\|\Pi^*\phi-\phi\right\|_K,$$

where $k_\Phi, k_{\boldsymbol{\phi}}, k_\sigma, k_\phi \in [0,k]$.

2.5 Postprocessing

We can obtain a better convergence result, namely an additional order for the approximate velocity $\boldsymbol{u}_h$, by postprocessing. In addition, this new approximation is divergence free and H(div) conforming. This postprocessing is identical to the one in [17] but for the sake of completeness, we present it here. Its definition is motivated by the Brezzi-Douglas-Marini (BDM) projection [4]. First we need to introduce more notation. Let B_K be a symmetric bubble matrix first defined in [15] based on the barycentric coordinates λ_i of the tetrahedron K, i.e.,

$$\mathrm{B}_K := \sum_{i=0}^{3}\lambda_{i-3}\lambda_{i-2}\lambda_{i-1}\nabla\lambda_i\otimes\nabla\lambda_i$$

with the subindices of barycentric coordinates calculated mod 4. This matrix does not vanish on ∂K but its rows have zero tangential components. Next, we introduce the necessary spaces. We denote by $\widetilde{P}_k$ the space of homogeneous polynomials of degree at most k. Let

$$\boldsymbol{S}_k := \{\boldsymbol{v}\in\widetilde{\boldsymbol{P}}_k : \boldsymbol{v}\cdot\boldsymbol{x}=0\},$$

and define the Nédélec space of the first kind [20] as follows:

$$\boldsymbol{N}_k(K) := \boldsymbol{P}_{k-1}(K)\oplus \mathrm{S}_k.$$

Finally, we set

$$\boldsymbol{S}_k(K) := \big\{\boldsymbol{v}\in\boldsymbol{N}_k(K) : (\boldsymbol{v},\nabla\phi)_K=0,\ \forall\phi\in P_{k+1}(K)\big\}.$$

Now we can define the postprocessed velocity $\boldsymbol{u}_h^*$ on $K\in\mathcal{T}_h$. Let $\boldsymbol{u}_h^*\in\boldsymbol{P}_{k+1}(K)$ be such that

$$\left\langle\big(\boldsymbol{u}_h^*-\widehat{\boldsymbol{u}}_h\big)\cdot\boldsymbol{n},\mu\right\rangle_F=0,\quad \forall\mu\in P_k(F),$$
$$\left\langle(\boldsymbol{n}\times\nabla)\big(\boldsymbol{u}_h^*\cdot\boldsymbol{n}\big)-\boldsymbol{n}\times\big(\{\!\{\mathrm{L}_h^T\}\!\}\boldsymbol{n}\big),(\boldsymbol{n}\times\nabla)\mu\right\rangle_F=0,\quad \forall\mu\in P_{k+1}(F)^\perp,$$

for all faces of K and

$$\big(\boldsymbol{u}_h^*-\boldsymbol{u},\nabla w\big)_K=0,\quad \forall w\in P_k(K),$$
$$\big(\nabla\times\boldsymbol{u}_h^*-\boldsymbol{w}_h,(\nabla\times\boldsymbol{v})\mathrm{B}_K\big)_K=0,\quad \forall\boldsymbol{v}\in\boldsymbol{S}_k(K),$$

where

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

approximates the vorticity of $\boldsymbol{u}_h^*$. Note that the first two equations for a fixed face F determine the value of $\boldsymbol{u}_h^* \cdot \boldsymbol{n} \in P_{k+1}(F)$. The last two conditions guarantee the well-posedness of $\boldsymbol{u}_h^*$ in a similar way to the BDM projection. The details can be found in [17, Proposition A.1]. Note that for 2-D, $\boldsymbol{n} \times \nabla$ is defined to be the tangential derivative $n_2\partial_1 + n_1\partial_2$, $\boldsymbol{n} \times \boldsymbol{w}$ corresponds to $n_1 w_2 - n_2 w_1$ and $\nabla \times \boldsymbol{u}$ is replaced by $\partial_1 u_2 - \partial_2 u_1$. Also, the bubble matrix reduces to the scalar bubble function $b_K := \lambda_0\lambda_1\lambda_2$, $\boldsymbol{w}_h$ reduces to $w_h := \mathrm{L}_{21}^h - \mathrm{L}_{12}^h$ and we replace the last equation by $(\nabla \times \boldsymbol{u}_h^* - w_h, w b_K)_K = 0$, for all $w \in P_{k-1}(K)$.

2.6 Main Theorem

Let us define a new norm for the projection of the approximation error of the numerical trace.

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

Now let $\boldsymbol{\beta}_0 \in \boldsymbol{P}_0(K)$ be a function such that

$$\big\langle (\boldsymbol{\beta} - \boldsymbol{\beta}_0) \cdot \boldsymbol{n}, 1 \big\rangle_F = 0, \quad \text{for all faces } F \text{ of } K,$$

for all $K \in \mathcal{T}_h$. Define $\boldsymbol{\delta\beta} := \boldsymbol{\beta} - \boldsymbol{\beta}_0$. This $\boldsymbol{\beta}_0$ exists as $\nabla \cdot \boldsymbol{\beta} = 0$. Indeed, it corresponds to the lowest order Raviart-Thomas projection of $\boldsymbol{\beta}$ and $\boldsymbol{\delta\beta}$ satisfies [3]

$$\|\boldsymbol{\delta\beta}\|_K \le C h_K \|\nabla \boldsymbol{\beta}\|_K, \quad \forall K \in \mathcal{T}_h.$$

The next theorem provides the main result of this paper which not only gives the convergence orders of the HDG method but also of the postprocessed velocity.

Theorem 2.6 *Suppose that the assumptions in Theorem 2.3 hold. Then,*

$$\left\|\mathrm{E}^{\mathrm{L}}\right\|_{\Omega} \le \|\Pi \mathrm{L} - \mathrm{L}\|_{\Omega}, \tag{2.14}$$

$$\left\|E^p\right\|_{\Omega} \le C_p^1 \|\Pi \mathrm{L} - \mathrm{L}\|_{\Omega} + C_p^2 \left\|\boldsymbol{E}^{\boldsymbol{u}}\right\|_{\Omega}, \tag{2.15}$$

where $C_p^1 := C \max(\nu, \|\boldsymbol{\beta}_0\|_{\boldsymbol{L}^\infty(\Omega)}, h|\boldsymbol{\beta}|_{\boldsymbol{W}^{1,\infty}(\Omega)})$ *and* $C_p^2 := C|\boldsymbol{\beta}|_{\boldsymbol{W}^{1,\infty}(\Omega)}$. *In addition, if the regularity estimate (2.12) holds,*

$$\begin{aligned} \left\|\boldsymbol{E}^{\boldsymbol{u}}\right\|_{\Omega} + \left|\boldsymbol{E}^{\widehat{\boldsymbol{u}}}\right|_h \le{}& C C_r \nu h^{\min(k,1)} \max_{K \in \mathcal{T}_h} \left(C_{\mathrm{HL}}^K\right) \|\Pi \mathrm{L} - \mathrm{L}\|_{\Omega} \\ &+ C_r h^{\min(k,1)} |\boldsymbol{\beta}|_{\boldsymbol{W}^{1,\infty}(\Omega)} \|\boldsymbol{\Pi u} - \boldsymbol{u}\|_{\Omega}, \end{aligned} \tag{2.16}$$

where

$$C_{\mathrm{HL}}^K := 1 + \left(\tau_t + \frac{1+\nu}{\nu} h_K |\boldsymbol{\beta}|_{\boldsymbol{W}^{1,\infty}(K)}\right)\left(\frac{(\widetilde{\Lambda}_K^{\max} + h_K|\boldsymbol{\beta}|_{\boldsymbol{W}^{1,\infty}(K)}) h_K + \nu}{\Gamma_K^{\min}} + h_K\right).$$

Finally, the postprocessed velocity $\boldsymbol{u}_h^* \in \boldsymbol{H}(\mathrm{div}, \Omega)$ *is divergence-free and satisfies*

$$\left\|\boldsymbol{u}_h^* - \boldsymbol{u}\right\|_{\Omega} \le C h^{k_{\boldsymbol{u}}+2} |\boldsymbol{u}|_{k_{\boldsymbol{u}}+2,\Omega} + C\left(\left\|\boldsymbol{E}^{\widehat{\boldsymbol{u}}}\right\|_h + \left\|\boldsymbol{E}^{\boldsymbol{u}}\right\|_{\Omega} + h^{\min(k,1)} \|\Pi \mathrm{L} - \mathrm{L}\|_{\Omega}\right), \tag{2.17}$$

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

Remark 2.7 Let us elaborate on the implications of the above theorem on the convergence properties of the numerical solution. Assume that all the constants of the above inequalities are uniformly bounded. Then, for $k=0$, $\|\mathrm{E}^{\mathrm{L}}\|_{\Omega}$ converges with the same order as $\|\Pi\mathrm{L}-\mathrm{L}\|_{\Omega}$ and the rest converge like $\|\Pi\mathrm{L}-\mathrm{L}\|_{\Omega}+\|\boldsymbol{\Pi u}-\boldsymbol{u}\|_{\Omega}$. Thus, there is no superconvergence property that we can exploit. For $k\geq 1$, $\|\mathrm{E}^{\mathrm{L}}\|_{\Omega}$ and $\|E^{p}\|_{\Omega}$ converges with the same order of $\|\Pi\mathrm{L}-\mathrm{L}\|_{\Omega}$. But this time, the rest converges with an additional order. This enables us to come up with a $\boldsymbol{u}_h^*$ that converges with an additional order. When $\boldsymbol{\beta}=\mathbf{0}$, we recover the results in [17]. See Tables 2.3 and 2.4 therein. What makes a difference is the knowledge of when the constants are uniformly bounded. If there is no uniform bound then the convergence properties gets complicated depending on τ_n, τ_t, ν and their relation with $\boldsymbol{\beta}$ through the assumption (2.5), the constants in the above equations and the ones given in Theorem 2.3 for the projection estimates.

3 Proofs

This section provides the proofs for the results stated in the previous section. However, we will refer the reader to [17] for the superconvergence result (2.17) of the postprocessed velocity as its proof follows exactly. The order of the proofs is as follows. Initially, we assume that the projections Π_h and Π_h^* are well-posed and satisfy the approximation properties stated in the previous section. Then, depending on this we prove the error estimates using an energy argument for the velocity gradient, and a duality argument for the velocity. Also, the estimate for the velocity trace is presented and an inf-sup condition is used to obtain the estimate for the pressure. The last section establishes the well-posedness and approximation properties of the projection Π_h.

3.1 Error Equations

Lemma 3.1 *The projection of the approximation errors* E^{L}, $\boldsymbol{E}^{\boldsymbol{u}}$, E^{p} *and* $\boldsymbol{E}^{\widehat{\boldsymbol{u}}}$ *satisfy*

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

$$\begin{aligned}&-\left(\nu\nabla\cdot\mathrm{E}^{\mathrm{L}},\boldsymbol{v}\right)_{\mathcal{T}_h}+\left(\nabla\cdot\left(\boldsymbol{E}^{\boldsymbol{u}}\otimes\boldsymbol{\beta}\right),\boldsymbol{v}\right)_{\mathcal{T}_h}+\left(\nabla E^{p},\boldsymbol{v}\right)_{\mathcal{T}_h}\\&\quad-\left\langle\boldsymbol{F}\left(0,\boldsymbol{E}^{\boldsymbol{u}},0,\boldsymbol{E}^{\widehat{\boldsymbol{u}}}\right),\boldsymbol{v}\right\rangle_{\partial\mathcal{T}_h}=0,\end{aligned} \tag{3.1b}$$

$$-\left(\boldsymbol{E}^{\boldsymbol{u}},\nabla q\right)_{\mathcal{T}_h}+\left\langle\boldsymbol{E}^{\widehat{\boldsymbol{u}}}\cdot\boldsymbol{n},q\right\rangle_{\partial\mathcal{T}_h}=0, \tag{3.1c}$$

$$\left\langle\boldsymbol{E}^{\widehat{\boldsymbol{u}}},\boldsymbol{\mu}\right\rangle_{\partial\Omega}=0, \tag{3.1d}$$

$$\left\langle\boldsymbol{F}\left(\mathrm{E}^{\mathrm{L}},\boldsymbol{E}^{\boldsymbol{u}},E^{p},\boldsymbol{E}^{\widehat{\boldsymbol{u}}}\right),\boldsymbol{\mu}\right\rangle_{\partial\mathcal{T}_h\setminus\partial\Omega}=0, \tag{3.1e}$$

$$\left(E^{p},1\right)_{\Omega}=0, \tag{3.1f}$$

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 If we insert (2.1g) in (2.1a)–(2.1e) and apply integration by parts, we get

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

$$- (\nu\nabla \cdot \mathrm{L}_h, \boldsymbol{v})_{\mathcal{T}_h} + \big(\nabla \cdot (\boldsymbol{u}_h \otimes \boldsymbol{\beta}), \boldsymbol{v}\big)_{\mathcal{T}_h} + (\nabla p_h, \boldsymbol{v})_{\mathcal{T}_h}$$
$$- \big\langle \boldsymbol{F}(0, \boldsymbol{u}_h, 0, \widehat{\boldsymbol{u}}), \boldsymbol{v}\big\rangle_{\partial\mathcal{T}_h} = (\boldsymbol{f}, \boldsymbol{v})_{\mathcal{T}_h}, \tag{3.2b}$$

$$-(\boldsymbol{u}_h, \nabla q)_{\mathcal{T}_h} + \langle \widehat{\boldsymbol{u}}_h \cdot \boldsymbol{n}, q\rangle_{\partial\mathcal{T}_h} = 0, \tag{3.2c}$$
$$\langle \widehat{\boldsymbol{u}}_h, \boldsymbol{\mu}\rangle_{\partial\Omega} = \langle \boldsymbol{g}, \boldsymbol{\mu}\rangle_{\partial\Omega}, \tag{3.2d}$$
$$\big\langle \boldsymbol{F}(\mathrm{L}_h, \boldsymbol{u}_h, p_h, \widehat{\boldsymbol{u}}_h), \boldsymbol{\mu}\big\rangle_{\partial\mathcal{T}_h\setminus\partial\Omega} = 0, \tag{3.2e}$$
$$(p_h, 1)_\Omega = 0, \tag{3.2f}$$

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

Observe that the exact solution satisfies

$$(\mathrm{L}, \mathrm{G})_{\mathcal{T}_h} + (\boldsymbol{\Pi}\boldsymbol{u}, \nabla \cdot \mathrm{G})_{\mathcal{T}_h} - \langle \boldsymbol{P}_M\boldsymbol{u}, \mathrm{G}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} = 0, \tag{3.3}$$

$$(\nu\Pi\mathrm{L}, \nabla\boldsymbol{v})_{\mathcal{T}_h} - (\boldsymbol{\Pi}\boldsymbol{u} \otimes \boldsymbol{\beta}, \nabla\boldsymbol{v})_{\mathcal{T}_h} - (\Pi p, \nabla \cdot \boldsymbol{v})_{\mathcal{T}_h}$$
$$- \big\langle \boldsymbol{F}(\Pi\mathrm{L}, \boldsymbol{\Pi}\boldsymbol{u}, \Pi p, \boldsymbol{P}_M\boldsymbol{u}), \boldsymbol{v}\big\rangle_{\partial\mathcal{T}_h} = (\boldsymbol{f}, \boldsymbol{v})_{\mathcal{T}_h}, \tag{3.4}$$

$$-(\boldsymbol{\Pi}\boldsymbol{u}, \nabla q)_{\mathcal{T}_h} + \langle \boldsymbol{P}_M\boldsymbol{u} \cdot \boldsymbol{n}, q\rangle_{\partial\mathcal{T}_h} = 0, \tag{3.5}$$
$$\langle \boldsymbol{P}_M\boldsymbol{u}, \boldsymbol{\mu}\rangle_{\partial\Omega} = \langle \boldsymbol{g}, \boldsymbol{\mu}\rangle_{\partial\Omega}, \tag{3.6}$$
$$\big\langle \boldsymbol{F}(\Pi\mathrm{L}, \boldsymbol{\Pi}\boldsymbol{u}, \Pi p, \boldsymbol{P}_M\boldsymbol{u}), \boldsymbol{\mu}\big\rangle_{\partial\mathcal{T}_h\setminus\partial\Omega} = 0, \tag{3.7}$$
$$(p, 1)_\Omega = 0 \tag{3.8}$$

using the projection defined by (2.7a)–(2.7d) and the fact that $\langle \mathrm{S}(\boldsymbol{P}_M\boldsymbol{u} - \boldsymbol{u}), \boldsymbol{\mu}\rangle_{\partial\mathcal{T}_h} = 0$ for all $\boldsymbol{\mu} \in \boldsymbol{M}_h$ (as τ_n and τ_t and hence S is assumed to be constant on each ∂K). Integration by parts in the second equation yields

$$- (\nu\nabla \cdot \Pi\mathrm{L}, \boldsymbol{v})_{\mathcal{T}_h} - (\boldsymbol{\Pi}\boldsymbol{u} \otimes \boldsymbol{\beta}, \nabla\boldsymbol{v})_{\mathcal{T}_h} + (\nabla\Pi p, \boldsymbol{v})_{\mathcal{T}_h}$$
$$- \big\langle \boldsymbol{F}(0, \boldsymbol{\Pi}\boldsymbol{u}, 0, \boldsymbol{P}_M\boldsymbol{u}), \boldsymbol{v}\big\rangle_{\partial\mathcal{T}_h} = (\boldsymbol{f}, \boldsymbol{v})_{\mathcal{T}_h}. \tag{3.9}$$

Then, we subtract the equations defined by (3.2a)–(3.2f) from (3.3), (3.5)–(3.9). Observe that in the above equations we cannot replace $(\mathrm{L}, \mathrm{G})_{\mathcal{T}_h}$ with $(\Pi\mathrm{L}, \mathrm{G})_{\mathcal{T}_h}$ as $\mathrm{G} \notin \mathrm{P}_{k-1}(K)$. So we write the difference of $(\mathrm{L}, \mathrm{G})_{\mathcal{T}_h}$ and $(\mathrm{L}_h, \mathrm{G})_{\mathcal{T}_h}$ as $(\mathrm{E}^{\mathrm{L}}, \mathrm{G})_{\mathcal{T}_h} - (\Pi\mathrm{L} - \mathrm{L}, \mathrm{G})_{\mathcal{T}_h}$. Similarly we write $(p, 1)_\Omega - (p_h, 1)_\Omega = (\Pi p - p_h, 1)_\Omega - (\Pi p - p, 1)_\Omega = (\Pi p - p_h, 1)_\Omega$ by (2.7c). This completes the proof. □

3.2 Estimates

Lemma 3.2 (An identity for the velocity gradient)

$$\nu\big\|\mathrm{E}^{\mathrm{L}}\big\|^2_{\mathcal{T}_h} + \big\langle \mathrm{S}_{\boldsymbol{\beta}}\big(\boldsymbol{E}^{\boldsymbol{u}} - \boldsymbol{E}^{\widehat{\boldsymbol{u}}}\big), \boldsymbol{E}^{\boldsymbol{u}} - \boldsymbol{E}^{\widehat{\boldsymbol{u}}}\big\rangle_{\partial\mathcal{T}_h} = \big(\Pi\mathrm{L} - \mathrm{L}, \nu\mathrm{E}^{\mathrm{L}}\big)_{\mathcal{T}_h}.$$

Proof Let $\mathrm{G} = \nu \mathrm{E}^{\mathrm{L}}$, $\boldsymbol{v} = \boldsymbol{E}^{\boldsymbol{u}}$, $q = E^p$ and $\boldsymbol{\mu} = \boldsymbol{E}^{\widehat{\boldsymbol{u}}}$ in (3.1a)–(3.1c). Adding the resulting equations, cancelling and rearranging terms, we obtain

$$\begin{aligned} \nu \|\mathrm{E}^{\mathrm{L}}\|^2_{\mathcal{T}_h} &- \langle \boldsymbol{E}^{\widehat{\boldsymbol{u}}}, \nu \mathrm{E}^{\mathrm{L}}\boldsymbol{n} - E^p \boldsymbol{n}\rangle_{\partial\mathcal{T}_h} + (\nabla\cdot(\boldsymbol{E}^{\boldsymbol{u}}\otimes\boldsymbol{\beta}), \boldsymbol{E}^{\boldsymbol{u}})_{\mathcal{T}_h} \\ &- \langle \boldsymbol{F}(0, \boldsymbol{E}^{\boldsymbol{u}}, 0, \boldsymbol{E}^{\widehat{\boldsymbol{u}}}), \boldsymbol{E}^{\boldsymbol{u}}\rangle_{\partial\mathcal{T}_h} = (\Pi \mathrm{L} - \mathrm{L}, \nu \mathrm{E}^{\mathrm{L}})_{\mathcal{T}_h}. \end{aligned}$$

Now let $\boldsymbol{\mu} = \boldsymbol{E}^{\widehat{\boldsymbol{u}}}$ in (3.1d) and (3.1e). This yields

$$\langle \boldsymbol{F}(\mathrm{E}^{\mathrm{L}}, \boldsymbol{E}^{\boldsymbol{u}}, E^p, \boldsymbol{E}_{\widehat{\boldsymbol{u}}}), \boldsymbol{E}^{\boldsymbol{u}}\rangle_{\partial\mathcal{T}_h} = 0$$

which implies

$$-\langle \nu \mathrm{E}^{\mathrm{L}}\boldsymbol{n} - E^p\boldsymbol{n}, \boldsymbol{E}^{\widehat{\boldsymbol{u}}}\rangle_{\partial\mathcal{T}_h} = \langle \boldsymbol{F}(0, \boldsymbol{E}^{\boldsymbol{u}}, 0, \boldsymbol{E}_{\widehat{\boldsymbol{u}}}), \boldsymbol{E}^{\boldsymbol{u}}\rangle_{\partial\mathcal{T}_h}.$$

Using these in the previous equation gives

$$\nu \|\mathrm{E}^{\mathrm{L}}\|^2_{\mathcal{T}_h} - (\boldsymbol{E}^{\boldsymbol{u}}\otimes\boldsymbol{\beta}, \nabla \boldsymbol{E}^{\boldsymbol{u}})_{\mathcal{T}_h} + \langle \boldsymbol{F}(0, \boldsymbol{E}^{\boldsymbol{u}}, 0, \boldsymbol{E}_{\widehat{\boldsymbol{u}}}), \boldsymbol{E}^{\boldsymbol{u}} - \boldsymbol{E}^{\widehat{\boldsymbol{u}}}\rangle_{\partial\mathcal{T}_h} = (\Pi \mathrm{L} - \mathrm{L}, \nu \mathrm{E}^{\mathrm{L}})_{\mathcal{T}_h}.$$

Observe that for all $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{w}$,

$$(\boldsymbol{u}\otimes\boldsymbol{v}, \nabla\boldsymbol{w}) = -(\nabla\boldsymbol{u}, \boldsymbol{w}\otimes\boldsymbol{v}) - (\boldsymbol{u}\cdot\boldsymbol{w}, \nabla\cdot\boldsymbol{v}) + \langle \boldsymbol{u}, (\boldsymbol{v}\cdot\boldsymbol{n})\boldsymbol{w}\rangle.$$

This means for $\nabla\cdot\boldsymbol{v} = 0$, $(\nabla\cdot(\boldsymbol{w}\otimes\boldsymbol{v}), \boldsymbol{w})_{\mathcal{T}_h} = \frac{1}{2}\langle \boldsymbol{w}, (\boldsymbol{v}\cdot\boldsymbol{n})\boldsymbol{w}\rangle_{\partial\mathcal{T}_h}$. Thus,

$$\nu \|\mathrm{E}^{\mathrm{L}}\|^2_{\mathcal{T}_h} - \frac{1}{2}\langle \boldsymbol{E}^{\boldsymbol{u}}, (\boldsymbol{\beta}\cdot\boldsymbol{n})\boldsymbol{E}^{\boldsymbol{u}}\rangle_{\partial\mathcal{T}_h} - \langle \boldsymbol{F}(0, \boldsymbol{E}^{\boldsymbol{u}}, 0, \boldsymbol{E}_{\widehat{\boldsymbol{u}}}), \boldsymbol{E}^{\boldsymbol{u}} - \boldsymbol{E}^{\widehat{\boldsymbol{u}}}\rangle_{\partial\mathcal{T}_h} = (\Pi \mathrm{L} - \mathrm{L}, \nu \mathrm{E}^{\mathrm{L}})_{\mathcal{T}_h}.$$

Using $\langle (\boldsymbol{v}\otimes\boldsymbol{\beta})\boldsymbol{n}, \boldsymbol{u}\rangle_{\partial\mathcal{T}_h} = \langle (\boldsymbol{\beta}\cdot\boldsymbol{n})\boldsymbol{v}, \boldsymbol{u}\rangle_{\partial\mathcal{T}_h}$ and rearranging terms, it is easy to show that

$$\begin{aligned} &-\frac{1}{2}\langle (\boldsymbol{\beta}\cdot\boldsymbol{n})\boldsymbol{E}^{\boldsymbol{u}}, \boldsymbol{E}^{\boldsymbol{u}}\rangle_{\partial\mathcal{T}_h} - \langle \boldsymbol{F}(0, \boldsymbol{E}^{\boldsymbol{u}}, 0, \boldsymbol{E}_{\widehat{\boldsymbol{u}}}), \boldsymbol{E}^{\boldsymbol{u}} - \boldsymbol{E}^{\widehat{\boldsymbol{u}}}\rangle_{\partial\mathcal{T}_h} \\ &\quad = -\frac{1}{2}\langle (\boldsymbol{\beta}\cdot\boldsymbol{n})\boldsymbol{E}^{\widehat{\boldsymbol{u}}}, \boldsymbol{E}^{\widehat{\boldsymbol{u}}}\rangle_{\partial\mathcal{T}_h} + \left\langle \left(\mathrm{S} - \frac{1}{2}(\boldsymbol{\beta}\cdot\boldsymbol{n})\mathrm{I}\right)(\boldsymbol{E}^{\boldsymbol{u}} - \boldsymbol{E}^{\widehat{\boldsymbol{u}}}), \boldsymbol{E}^{\boldsymbol{u}} - \boldsymbol{E}^{\widehat{\boldsymbol{u}}}\right\rangle_{\partial\mathcal{T}_h} \\ &\quad = \left\langle \left(\mathrm{S} - \frac{1}{2}(\boldsymbol{\beta}\cdot\boldsymbol{n})\mathrm{I}\right)(\boldsymbol{E}^{\boldsymbol{u}} - \boldsymbol{E}^{\widehat{\boldsymbol{u}}}), \boldsymbol{E}^{\boldsymbol{u}} - \boldsymbol{E}^{\widehat{\boldsymbol{u}}}\right\rangle_{\partial\mathcal{T}_h}, \end{aligned}$$

where in the second equality we used the facts $\boldsymbol{\beta} \in \boldsymbol{H}(\mathrm{div}, \Omega)$, $\boldsymbol{E}^{\widehat{\boldsymbol{u}}}$ is single valued and vanishes on $\partial\Omega$. Therefore,

$$\nu \|\mathrm{E}^{\mathrm{L}}\|^2_{\mathcal{T}_h} + \left\langle \left(\mathrm{S} - \frac{1}{2}(\boldsymbol{\beta}\cdot\boldsymbol{n})\mathrm{I}\right)(\boldsymbol{E}^{\boldsymbol{u}} - \boldsymbol{E}^{\widehat{\boldsymbol{u}}}), \boldsymbol{E}^{\boldsymbol{u}} - \boldsymbol{E}^{\widehat{\boldsymbol{u}}}\right\rangle_{\partial\mathcal{T}_h} = (\Pi \mathrm{L} - \mathrm{L}, \nu \mathrm{E}^{\mathrm{L}})_{\mathcal{T}_h}.$$

□

Define the following seminorm on $\partial\mathcal{T}_h$ based on a semi-positive definite matrix M and α such that

$$|\boldsymbol{v}|_{\alpha\mathrm{M}} := \left\{ \sum_{K\in\mathcal{T}_h} \alpha \langle \mathrm{M}\boldsymbol{v}, \boldsymbol{v}\rangle_{\partial K} \right\}^{1/2},$$

where $\alpha > 0$ is a constant. An immediate consequence of the previous lemma is the next corollary.

Corollary 3.3 (An estimate for the velocity gradient)

$$\|\mathrm{E}^{\mathrm{L}}\|^2_{\mathcal{T}_h} + \left|\boldsymbol{E}^{\boldsymbol{u}} - \boldsymbol{E}^{\widehat{\boldsymbol{u}}}\right|^2_{\mathrm{S}_{\boldsymbol{\beta}/\nu}} \leq \|\Pi\mathrm{L} - \mathrm{L}\|^2_{\mathcal{T}_h}. \tag{3.10}$$

In the proof of next lemma we use a duality argument.

Lemma 3.4 (An identity for the velocity)

$$\left(\boldsymbol{E}^{\boldsymbol{u}}, \boldsymbol{\theta}\right)_{\mathcal{T}_h} = T_{\mathrm{L}} + T_{\boldsymbol{\beta}},$$

where $T_{\boldsymbol{\beta}}$ and T_{L} are defined as

$$T_{\mathrm{L}} := \left(\mathrm{L} - \mathrm{L}_h, \nu\Pi^*\Phi - \nu\Phi\right)_{\mathcal{T}_h} - \left(\nu\mathrm{L} - \nu\Pi\mathrm{L}, \nabla(\boldsymbol{\phi}_h - \boldsymbol{\phi})\right)_{\mathcal{T}_h} \quad \textit{and}$$

$$T_{\boldsymbol{\beta}} := \left((\boldsymbol{u} - \boldsymbol{\Pi}\boldsymbol{u}) \otimes \delta\boldsymbol{\beta}, \nabla\boldsymbol{\phi}_h\right)_{\mathcal{T}_h}.$$

Proof Recall the dual problem defined by (2.11a)–(2.11d). In particular by (2.11b), we have

$$\left(\boldsymbol{E}^{\boldsymbol{u}}, \boldsymbol{\theta}\right)_{\mathcal{T}_h} = -\left(\boldsymbol{E}^{\boldsymbol{u}}, \nu\nabla\cdot\Phi\right)_{\mathcal{T}_h} - \left(\boldsymbol{E}^{\boldsymbol{u}}, \nabla\cdot(\boldsymbol{\phi}\otimes\boldsymbol{\beta})\right)_{\mathcal{T}_h} - \left(\boldsymbol{E}^{\boldsymbol{u}}, \nabla\phi\right)_{\mathcal{T}_h}.$$

Using integration by parts and the properties (2.13a) and (2.13c) of the projection Π_h^*,

$$\begin{aligned}\left(\boldsymbol{E}^{\boldsymbol{u}}, \boldsymbol{\theta}\right)_{\mathcal{T}_h} &= \left(\nabla\boldsymbol{E}^{\boldsymbol{u}}, \nu\Phi + \boldsymbol{\phi}\otimes\boldsymbol{\beta}\right)_{\mathcal{T}_h} + \left(\nabla\cdot\boldsymbol{E}^{\boldsymbol{u}}, \phi\right)_{\mathcal{T}_h} - \left\langle\boldsymbol{E}^{\boldsymbol{u}}, \boldsymbol{F}^*(\Phi, \boldsymbol{\phi}, \phi, 0)\right\rangle_{\partial\mathcal{T}_h}\\ &= \left(\nabla\boldsymbol{E}^{\boldsymbol{u}}, \nu\Pi^*\Phi + \left(\boldsymbol{\Pi}^*\boldsymbol{\phi}\right)\otimes\boldsymbol{\beta}\right)_{\mathcal{T}_h} + \left(\nabla\cdot\boldsymbol{E}^{\boldsymbol{u}}, \Pi^*\phi\right)_{\mathcal{T}_h} - \left\langle\boldsymbol{E}^{\boldsymbol{u}}, \boldsymbol{F}^*(\Phi, \boldsymbol{\phi}, \phi, 0)\right\rangle_{\partial\mathcal{T}_h}.\end{aligned}$$

Applying integration by parts once again yields

$$\begin{aligned}\left(\boldsymbol{E}^{\boldsymbol{u}}, \boldsymbol{\theta}\right)_{\mathcal{T}_h} = &-\left(\boldsymbol{E}^{\boldsymbol{u}}, \nabla\cdot\left(\nu\Pi^*\Phi + \left(\boldsymbol{\Pi}^*\boldsymbol{\phi}\right)\otimes\boldsymbol{\beta}\right)\right)_{\mathcal{T}_h} - \left(\boldsymbol{E}^{\boldsymbol{u}}, \nabla\Pi^*\phi\right)_{\mathcal{T}_h}\\ &- \left\langle\boldsymbol{E}^{\boldsymbol{u}}, \boldsymbol{F}^*\left(\Phi - \Pi^*\Phi, \boldsymbol{\phi} - \boldsymbol{\Pi}^*\boldsymbol{\phi}, \phi - \Pi^*\phi, 0\right)\right\rangle_{\partial\mathcal{T}_h}.\end{aligned}$$

Now apply (3.1a) with $\mathrm{G} = \nu\Pi^*\Phi$,

$$\begin{aligned}\left(\boldsymbol{E}^{\boldsymbol{u}}, \boldsymbol{\theta}\right)_{\mathcal{T}_h} = &\left(\mathrm{L} - \mathrm{L}_h, \nu\Pi^*\Phi\right)_{\mathcal{T}_h} - \left\langle\boldsymbol{E}^{\widehat{\boldsymbol{u}}}, \nu\Pi^*\Phi\boldsymbol{n}\right\rangle_{\partial\mathcal{T}_h}\\ &- \left(\boldsymbol{E}^{\boldsymbol{u}}, \nabla\cdot\left(\left(\boldsymbol{\Pi}^*\boldsymbol{\phi}\right)\otimes\boldsymbol{\beta}\right)\right)_{\mathcal{T}_h} - \left(\boldsymbol{E}^{\boldsymbol{u}}, \nabla\Pi^*\phi\right)_{\mathcal{T}_h}\\ &- \left\langle\boldsymbol{E}^{\boldsymbol{u}}, \boldsymbol{F}^*\left(\Phi - \Pi^*\Phi, \boldsymbol{\phi} - \boldsymbol{\Pi}^*\boldsymbol{\phi}, \phi - \Pi^*\phi, 0\right)\right\rangle_{\partial\mathcal{T}_h}.\end{aligned} \tag{3.11}$$

Let us work with the first four right hand side terms in (3.11). Pick an arbitrary $\boldsymbol{\phi}_h \in \boldsymbol{V}_h$. We add and subtract $(\mathrm{L} - \mathrm{L}_h, \nu\Phi)_{\mathcal{T}_h}$ and $\nu(\mathrm{L} - \Pi\mathrm{L}, \nabla(\boldsymbol{\phi}_h - \boldsymbol{\phi}))_{\mathcal{T}_h}$ and obtain

$$\begin{aligned}&\left(\mathrm{L} - \mathrm{L}_h, \nu\Pi^*\Phi\right)_{\mathcal{T}_h} - \left\langle\boldsymbol{E}^{\widehat{\boldsymbol{u}}}, \nu\Pi^*\Phi\boldsymbol{n}\right\rangle_{\partial\mathcal{T}_h} - \left(\boldsymbol{E}^{\boldsymbol{u}}, \nabla\cdot\left(\left(\boldsymbol{\Pi}^*\boldsymbol{\phi}\right)\otimes\boldsymbol{\beta}\right)\right)_{\mathcal{T}_h} - \left(\boldsymbol{E}^{\boldsymbol{u}}, \nabla\Pi^*\phi\right)_{\mathcal{T}_h}\\ &= T_{\mathrm{L}} + (\mathrm{L} - \mathrm{L}_h, \nu\Phi)_{\mathcal{T}_h} + \nu\left(\mathrm{L} - \Pi\mathrm{L}, \nabla(\boldsymbol{\phi}_h - \boldsymbol{\phi})\right)_{\mathcal{T}_h} - \left\langle\boldsymbol{E}^{\widehat{\boldsymbol{u}}}, \nu\Pi^*\Phi\boldsymbol{n}\right\rangle_{\partial\mathcal{T}_h}\\ &\quad - \left(\boldsymbol{E}^{\boldsymbol{u}}, \nabla\cdot\left(\left(\boldsymbol{\Pi}^*\boldsymbol{\phi}\right)\otimes\boldsymbol{\beta}\right)\right)_{\mathcal{T}_h} - \left(\boldsymbol{E}^{\boldsymbol{u}}, \nabla\Pi^*\phi\right)_{\mathcal{T}_h}.\end{aligned} \tag{3.12}$$

Observe at this point that by the virtue of (2.11a) (applied twice) we can rewrite the second and third terms on the right hand side of (3.12) as

$$\begin{aligned}&(\mathrm{L}-\mathrm{L}_h,\nu\Phi)_{\mathcal{T}_h}+\nu\big(\mathrm{L}-\Pi\mathrm{L},\nabla(\boldsymbol{\phi}_h-\boldsymbol{\phi})\big)_{\mathcal{T}_h}\\&\quad=-(\mathrm{L}_h,\nu\Phi)_{\mathcal{T}_h}+\nu(\mathrm{L}-\Pi\mathrm{L},\nabla\boldsymbol{\phi}_h)_{\mathcal{T}_h}+\nu(\Pi\mathrm{L},\nabla\boldsymbol{\phi})_{\mathcal{T}_h}\\&\quad=-(\mathrm{L}_h,\nu\nabla\boldsymbol{\phi})_{\mathcal{T}_h}+\nu(\mathrm{L}-\Pi\mathrm{L},\nabla\boldsymbol{\phi}_h)_{\mathcal{T}_h}+\nu(\Pi\mathrm{L},\nabla\boldsymbol{\phi})_{\mathcal{T}_h}\\&\quad=\big(\mathrm{E}^{\mathrm{L}},\nu\nabla\boldsymbol{\phi}\big)_{\mathcal{T}_h}+\nu(\mathrm{L}-\Pi\mathrm{L},\nabla\boldsymbol{\phi}_h)_{\mathcal{T}_h}.\end{aligned}$$

Thus,

$$\begin{aligned}\big(\boldsymbol{E}^{\boldsymbol{u}},\boldsymbol{\theta}\big)_{\mathcal{T}_h}&=T_{\mathrm{L}}+\big(\mathrm{E}^{\mathrm{L}},\nu\nabla\boldsymbol{\phi}\big)_{\mathcal{T}_h}+\nu(\mathrm{L}-\Pi\mathrm{L},\nabla\boldsymbol{\phi}_h)_{\mathcal{T}_h}-\big\langle\boldsymbol{E}^{\widehat{\boldsymbol{u}}},\nu\Pi^*\Phi\boldsymbol{n}\big\rangle_{\partial\mathcal{T}_h}\\&\quad-\big(\boldsymbol{E}^{\boldsymbol{u}},\nabla\cdot\big(\big(\boldsymbol{\Pi}^*\boldsymbol{\phi}\big)\otimes\boldsymbol{\beta}\big)\big)_{\mathcal{T}_h}-\big(\boldsymbol{E}^{\boldsymbol{u}},\nabla\Pi^*\phi\big)_{\mathcal{T}_h}\\&\quad-\big\langle\boldsymbol{E}^{\boldsymbol{u}},\boldsymbol{F}^*\big(\Phi-\Pi^*\Phi,\boldsymbol{\phi}-\boldsymbol{\Pi}^*\boldsymbol{\phi},\phi-\Pi^*\phi,0\big)\big\rangle_{\partial\mathcal{T}_h}.\end{aligned}$$

Recall that we defined $\boldsymbol{\delta\beta}:=\boldsymbol{\beta}-\boldsymbol{\beta}_0$, where $\boldsymbol{\beta}_0$ is the lowest order Raviart-Thomas projection of $\boldsymbol{\beta}$. Also, observe from (2.7b) that

$$\big((\boldsymbol{\Pi}\boldsymbol{u}-\boldsymbol{u})\otimes\boldsymbol{\beta}_0,\mathrm{G}\big)_{\mathcal{T}_h}=(\boldsymbol{\Pi}\boldsymbol{u}-\boldsymbol{u},\mathrm{G}\boldsymbol{\beta}_0)_{\mathcal{T}_h}=0,\quad\forall\mathrm{G}\in\mathrm{P}_{k-1}(K)$$

as $\boldsymbol{\beta}_0$ is constant on each K. Hence by (2.7a), we can replace $\nu(\mathrm{L}-\Pi\mathrm{L},\nabla\boldsymbol{\phi}_h)_{\mathcal{T}_h}$ with $T_{\boldsymbol{\beta}}:=((\boldsymbol{u}-\boldsymbol{\Pi}\boldsymbol{u})\otimes\boldsymbol{\delta\beta},\nabla\boldsymbol{\phi}_h)_{\mathcal{T}_h}$. Therefore,

$$\big(\boldsymbol{E}^{\boldsymbol{u}},\boldsymbol{\theta}\big)_{\mathcal{T}_h}=T_{\mathrm{L}}+T_{\boldsymbol{\beta}}+R,$$

where

$$\begin{aligned}R:=&\big(\mathrm{E}^{\mathrm{L}},\nu\nabla\boldsymbol{\phi}\big)_{\mathcal{T}_h}-\big\langle\boldsymbol{E}^{\widehat{\boldsymbol{u}}},\nu\Pi^*\Phi\boldsymbol{n}\big\rangle_{\partial\mathcal{T}_h}-\big(\boldsymbol{E}^{\boldsymbol{u}},\nabla\cdot\big(\big(\boldsymbol{\Pi}^*\boldsymbol{\phi}\big)\otimes\boldsymbol{\beta}\big)\big)_{\mathcal{T}_h}-\big(\boldsymbol{E}^{\boldsymbol{u}},\nabla\Pi^*\phi\big)_{\mathcal{T}_h}\\&-\big\langle\boldsymbol{E}^{\boldsymbol{u}},\boldsymbol{F}^*\big(\Phi-\Pi^*\Phi,\boldsymbol{\phi}-\boldsymbol{\Pi}^*\boldsymbol{\phi},\phi-\Pi^*\phi,0\big)\big\rangle_{\partial\mathcal{T}_h}.\end{aligned}$$

In (3.1b), taking $\boldsymbol{v}=\boldsymbol{\Pi}^*\boldsymbol{\phi}$ we can rewrite the third term in the definition of R as

$$\begin{aligned}&\big(\boldsymbol{E}^{\boldsymbol{u}},\nabla\cdot\big(\big(\boldsymbol{\Pi}^*\boldsymbol{\phi}\big)\otimes\boldsymbol{\beta}\big)\big)_{\mathcal{T}_h}\\&\quad=-\big(\nu\nabla\cdot\mathrm{E}^{\mathrm{L}},\boldsymbol{\Pi}^*\boldsymbol{\phi}\big)_{\mathcal{T}_h}+\big(\nabla E^p,\boldsymbol{\Pi}^*\boldsymbol{\phi}\big)_{\mathcal{T}_h}-\big\langle\boldsymbol{F}\big(0,\boldsymbol{E}^{\boldsymbol{u}},0,\boldsymbol{E}^{\widehat{\boldsymbol{u}}}\big),\boldsymbol{\Pi}^*\boldsymbol{\phi}\big\rangle_{\partial\mathcal{T}_h}\\&\quad=\big(\nu\mathrm{E}^{\mathrm{L}},\nabla\boldsymbol{\Pi}^*\boldsymbol{\phi}\big)_{\mathcal{T}_h}-\big(E^p,\nabla\cdot\boldsymbol{\Pi}^*\boldsymbol{\phi}\big)_{\mathcal{T}_h}-\big\langle\boldsymbol{F}\big(\mathrm{E}^{\mathrm{L}},\boldsymbol{E}^{\boldsymbol{u}},E^p,\boldsymbol{E}^{\widehat{\boldsymbol{u}}}\big),\boldsymbol{\Pi}^*\boldsymbol{\phi}\big\rangle_{\partial\mathcal{T}_h}.\end{aligned}$$

Using (2.11c) and rearranging terms yield

$$\begin{aligned}R=&\big(\nu\mathrm{E}^{\mathrm{L}},\nabla\big(\boldsymbol{\phi}-\boldsymbol{\Pi}^*\boldsymbol{\phi}\big)\big)_{\mathcal{T}_h}-\big\langle\boldsymbol{E}^{\widehat{\boldsymbol{u}}},\nu\Pi^*\Phi\boldsymbol{n}\big\rangle_{\partial\mathcal{T}_h}\\&+\big\langle\boldsymbol{F}\big(\mathrm{E}^{\mathrm{L}},\boldsymbol{E}^{\boldsymbol{u}},E^p,\boldsymbol{E}^{\widehat{\boldsymbol{u}}}\big),\boldsymbol{\Pi}^*\boldsymbol{\phi}\big\rangle_{\partial\mathcal{T}_h}+\big(E^p,\nabla\cdot\big(\boldsymbol{\Pi}^*\boldsymbol{\phi}-\boldsymbol{\phi}\big)\big)_{\mathcal{T}_h}\\&-\big(\boldsymbol{E}^{\boldsymbol{u}},\nabla\Pi^*\phi\big)_{\mathcal{T}_h}-\big\langle\boldsymbol{E}^{\boldsymbol{u}},\boldsymbol{F}^*\big(\Phi-\Pi^*\Phi,\boldsymbol{\phi}-\boldsymbol{\Pi}^*\boldsymbol{\phi},\phi-\Pi^*\phi,0\big)\big\rangle_{\partial\mathcal{T}_h}.\end{aligned}$$

Integrating by parts once more, using (2.13b), (3.1e) with $\boldsymbol{\mu}=\boldsymbol{P}_M\boldsymbol{\phi}$ and recalling (2.11d) imply

$$\begin{aligned}R=&\big\langle\nu\mathrm{E}^{\mathrm{L}}\boldsymbol{n},\boldsymbol{\phi}-\boldsymbol{\Pi}^*\boldsymbol{\phi}\big\rangle_{\partial\mathcal{T}_h}-\big\langle\boldsymbol{E}^{\widehat{\boldsymbol{u}}},\nu\Pi^*\Phi\boldsymbol{n}\big\rangle_{\partial\mathcal{T}_h}+\big\langle\boldsymbol{F}\big(\mathrm{E}^{\mathrm{L}},\boldsymbol{E}^{\boldsymbol{u}},E^p,\boldsymbol{E}^{\widehat{\boldsymbol{u}}}\big),\boldsymbol{\Pi}^*\boldsymbol{\phi}-\boldsymbol{P}_M\boldsymbol{\phi}\big\rangle_{\partial\mathcal{T}_h}\\&+\big\langle E^p,\big(\boldsymbol{\Pi}^*\boldsymbol{\phi}-\boldsymbol{\phi}\big)\boldsymbol{n}\big\rangle_{\partial\mathcal{T}_h}-\big(\boldsymbol{E}^{\boldsymbol{u}},\nabla\Pi^*\phi\big)_{\mathcal{T}_h}\\&-\big\langle\boldsymbol{E}^{\boldsymbol{u}},\boldsymbol{F}^*\big(\Phi-\Pi^*\Phi,\boldsymbol{\phi}-\boldsymbol{\Pi}^*\boldsymbol{\phi},\phi-\Pi^*\phi,0\big)\big\rangle_{\partial\mathcal{T}_h}.\end{aligned}$$

Rearranging terms and adding and subtracting $\langle \boldsymbol{E}^{\widehat{\boldsymbol{u}}}, \nu \Phi \boldsymbol{n}\rangle_{\partial\mathcal{T}_h}$ we get

$$\begin{aligned}R &= \big\langle \nu \mathrm{E}^{\mathrm{L}}\boldsymbol{n}, \boldsymbol{\phi} - \boldsymbol{P}_M\boldsymbol{\phi}\big\rangle_{\partial\mathcal{T}_h} - \big\langle \boldsymbol{E}^{\widehat{\boldsymbol{u}}}, \nu\big(\Pi^*\Phi - \Phi\big)\boldsymbol{n}\big\rangle_{\partial\mathcal{T}_h} - \big\langle \boldsymbol{E}^{\widehat{\boldsymbol{u}}}, \nu\Phi\boldsymbol{n}\big\rangle \\ &\quad + \big\langle \boldsymbol{F}^*\big(0, \boldsymbol{E}^{\boldsymbol{u}}, E^p, \boldsymbol{E}^{\widehat{\boldsymbol{u}}}\big), \boldsymbol{\Pi}^*\boldsymbol{\phi} - \boldsymbol{P}_M\boldsymbol{\phi}\big\rangle_{\partial\mathcal{T}_h} \\ &\quad + \big\langle E^p, \big(\boldsymbol{\Pi}^*\boldsymbol{\phi} - \boldsymbol{\phi}\big)\boldsymbol{n}\big\rangle_{\partial\mathcal{T}_h} - \big(\boldsymbol{E}^{\boldsymbol{u}}, \nabla\Pi^*\phi\big)_{\mathcal{T}_h} \\ &\quad - \big\langle \boldsymbol{E}^{\boldsymbol{u}}, \boldsymbol{F}^*\big(\Phi - \Pi^*\Phi, \boldsymbol{\phi} - \boldsymbol{\Pi}^*\boldsymbol{\phi}, \phi - \Pi^*\phi, 0\big)\big\rangle_{\partial\mathcal{T}_h}.\end{aligned}$$

The first term vanishes by the definition of $\boldsymbol{P}_M$. Thus, the above equation can be rewritten as

$$\begin{aligned}R &= -\big\langle \boldsymbol{E}^{\boldsymbol{u}} - \boldsymbol{E}^{\widehat{\boldsymbol{u}}}, \nu\big(\Phi - \Pi^*\Phi\big)\boldsymbol{n}\big\rangle_{\partial\mathcal{T}_h} - \big\langle \boldsymbol{E}^{\widehat{\boldsymbol{u}}}, \nu\Phi\boldsymbol{n}\big\rangle_{\partial\mathcal{T}_h} - \big\langle \boldsymbol{E}^{\widehat{\boldsymbol{u}}}, (\boldsymbol{\beta}\cdot\boldsymbol{n})\big(\boldsymbol{\Pi}^*\boldsymbol{\phi} - \boldsymbol{P}_M\boldsymbol{\phi}\big)\big\rangle_{\partial\mathcal{T}_h} \\ &\quad - \big\langle \boldsymbol{E}^{\boldsymbol{u}}, (\boldsymbol{\beta}\cdot\boldsymbol{n})\big(\boldsymbol{\phi} - \boldsymbol{\Pi}^*\boldsymbol{\phi}\big)\big\rangle_{\partial\mathcal{T}_h} - \big\langle \mathrm{S}\big(\boldsymbol{E}^{\boldsymbol{u}} - \boldsymbol{E}^{\widehat{\boldsymbol{u}}}\big), \boldsymbol{\Pi}^*\boldsymbol{\phi} - \boldsymbol{P}_M\boldsymbol{\phi}\big\rangle_{\partial\mathcal{T}_h} \\ &\quad + \big\langle E^p, (\boldsymbol{P}_M\boldsymbol{\phi} - \boldsymbol{\phi})\boldsymbol{n}\big\rangle_{\partial\mathcal{T}_h} - \big(\boldsymbol{E}^{\boldsymbol{u}}, \nabla\Pi^*\phi\big)_{\mathcal{T}_h} - \big\langle \boldsymbol{E}^{\boldsymbol{u}}, \big(\phi - \Pi^*\phi\big)\boldsymbol{n}\big\rangle_{\partial\mathcal{T}_h}.\end{aligned}$$

Note in the above we used $((\boldsymbol{u}\otimes\boldsymbol{\beta})\boldsymbol{n}, \boldsymbol{v})_{\mathcal{T}_h} = ((\boldsymbol{\beta}\cdot\boldsymbol{n})\boldsymbol{u}, \boldsymbol{v})_{\mathcal{T}_h}$. Now we observe that the sixth term drops because of the definition of $\boldsymbol{P}_M$. Then, adding and subtracting $(\boldsymbol{E}^{\widehat{\boldsymbol{u}}}, (\boldsymbol{\beta}\cdot\boldsymbol{n})\boldsymbol{\phi})_{\mathcal{T}_h}$ yield

$$\begin{aligned}R &= -\big\langle \boldsymbol{E}^{\boldsymbol{u}} - \boldsymbol{E}^{\widehat{\boldsymbol{u}}}, \nu\big(\Phi - \Pi^*\Phi\big)\boldsymbol{n}\big\rangle_{\partial\mathcal{T}_h} - \big\langle \boldsymbol{E}^{\boldsymbol{u}} - \boldsymbol{E}^{\widehat{\boldsymbol{u}}}, (\boldsymbol{\beta}\cdot\boldsymbol{n})\big(\boldsymbol{\phi} - \boldsymbol{\Pi}^*\boldsymbol{\phi}\big)\big\rangle_{\partial\mathcal{T}_h} \\ &\quad - \big\langle \mathrm{S}\big(\boldsymbol{E}^{\boldsymbol{u}} - \boldsymbol{E}^{\widehat{\boldsymbol{u}}}\big), \boldsymbol{\Pi}^*\boldsymbol{\phi} - P_M\boldsymbol{\phi}\big\rangle_{\partial\mathcal{T}_h} \\ &\quad - \big(\boldsymbol{E}^{\boldsymbol{u}}, \nabla\Pi^*\phi\big)_{\mathcal{T}_h} - \big\langle \boldsymbol{E}^{\boldsymbol{u}}, \big(\phi - \Pi^*\phi\big)\boldsymbol{n}\big\rangle_{\partial\mathcal{T}_h} - \big\langle \boldsymbol{E}^{\widehat{\boldsymbol{u}}}, \nu\Phi\boldsymbol{n} + (\boldsymbol{\beta}\cdot\boldsymbol{n})(\boldsymbol{\phi} - \boldsymbol{P}_M\boldsymbol{\phi})\big\rangle_{\partial\mathcal{T}_h}.\end{aligned}$$

Combining terms further

$$\begin{aligned}R &= -\big\langle \boldsymbol{E}^{\boldsymbol{u}} - \boldsymbol{E}^{\widehat{\boldsymbol{u}}}, \boldsymbol{F}^*\big(\Phi - \Pi^*\Phi, \boldsymbol{\phi} - \boldsymbol{\Pi}^*\boldsymbol{\phi}, 0, 0\big)\big\rangle_{\partial\mathcal{T}_h} \\ &\quad - \big\langle \boldsymbol{E}^{\widehat{\boldsymbol{u}}}, \nu\Phi\boldsymbol{n} + (\boldsymbol{\beta}\cdot\boldsymbol{n})(\boldsymbol{\phi} - \boldsymbol{P}_M\boldsymbol{\phi})\big\rangle_{\partial\mathcal{T}_h} - \big\langle \boldsymbol{E}^{\widehat{\boldsymbol{u}}}, \Pi^*\phi\boldsymbol{n}\big\rangle_{\partial\mathcal{T}_h} - \big\langle \boldsymbol{E}^{\boldsymbol{u}}, \big(\phi - \Pi^*\phi\big)\boldsymbol{n}\big\rangle_{\partial\mathcal{T}_h},\end{aligned}$$

where we used (3.1c) to replace $(\boldsymbol{E}^{\boldsymbol{u}}, \nabla\Pi^*\phi)_{\mathcal{T}_h}$ by $\langle \boldsymbol{E}^{\widehat{\boldsymbol{u}}}, \Pi^*\phi\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}$. Consider the second term on the right hand side. We have

$$-\big\langle \boldsymbol{E}^{\widehat{\boldsymbol{u}}}, \nu\Phi\boldsymbol{n} + (\boldsymbol{\beta}\cdot\boldsymbol{n})(\boldsymbol{\phi} - \boldsymbol{P}_M\boldsymbol{\phi})\big\rangle_{\partial\mathcal{T}_h} = -\big\langle \boldsymbol{E}^{\widehat{\boldsymbol{u}}}, \nu\Phi\boldsymbol{n} + (\boldsymbol{\beta}\cdot\boldsymbol{n})(\boldsymbol{\phi} - \boldsymbol{P}_M\boldsymbol{\phi})\big\rangle_{\partial\Omega} = 0,$$

where the first equality holds as $\boldsymbol{E}^{\widehat{\boldsymbol{u}}}$ is single valued on $\mathcal{E}_h^0$ and $\nu\Phi\boldsymbol{n} + (\boldsymbol{\beta}\cdot\boldsymbol{n})(\boldsymbol{\phi} - \boldsymbol{P}_M\boldsymbol{\phi})$ is continuous on $\mathcal{E}_h^0$, and the second equality follows by the error equation (3.1d). Furthermore, we can add $\langle \boldsymbol{E}^{\widehat{\boldsymbol{u}}}, \phi\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}$ by (3.1d) and the facts that $\boldsymbol{E}^{\widehat{\boldsymbol{u}}}$ is single valued on $\mathcal{E}_h^0$ and $\phi\boldsymbol{n}$ is continuous. Finally,

$$R = -\big\langle \boldsymbol{E}^{\boldsymbol{u}} - \boldsymbol{E}^{\widehat{\boldsymbol{u}}}, \boldsymbol{F}^*\big(\Phi - \Pi^*\Phi, \boldsymbol{\phi} - \boldsymbol{\Pi}^*\boldsymbol{\phi}, \phi - \Pi^*\phi, 0\big)\big\rangle_{\partial\mathcal{T}_h} = 0$$

by (2.13d) with the choice $\boldsymbol{\mu} = \boldsymbol{E}^{\boldsymbol{u}} - \boldsymbol{E}^{\widehat{\boldsymbol{u}}}$. In other words,

$$\big(\boldsymbol{E}^{\boldsymbol{u}}, \boldsymbol{\theta}\big)_{\mathcal{T}_h} = T_{\mathrm{L}} + T_{\boldsymbol{\beta}}. \qquad \square$$

Next we use this identity to bound $\|\boldsymbol{E}^{\boldsymbol{u}}\|_{\Omega}$. Let $\boldsymbol{\phi}_h \in \boldsymbol{P}_k(K)$ be such that

(i) $(\nabla(\boldsymbol{\phi}_h - \boldsymbol{\phi}), \nabla \boldsymbol{w})_K = 0, \ \forall \boldsymbol{w} \in \boldsymbol{P}_k(K)$, and
(ii) $(\boldsymbol{\phi}_h - \boldsymbol{\phi}, 1)_K = 0$.

Note that (i) implies

$$\|\nabla \boldsymbol{\phi}_h\|_K \le \|\nabla \boldsymbol{\phi}\|_K. \tag{3.13}$$

Theorem 3.5 (Estimate for the velocity)

$$\|\boldsymbol{E}^{\boldsymbol{u}}\|_{\Omega} \le 2H_{\mathrm{L}}\|\Pi \mathrm{L} - \mathrm{L}\|_{\mathcal{T}_h} + H_{\boldsymbol{\beta}}\|\boldsymbol{\Pi u} - \boldsymbol{u}\|_{\mathcal{T}_h}, \tag{3.14}$$

where H_{L} *and* $H_{\boldsymbol{\beta}}$ *are defined by*

$$H_{\mathrm{L}} := \nu \max\left\{ \sup_{\boldsymbol{\theta}\in \boldsymbol{L}^2(\Omega)\setminus\{\boldsymbol{0}\}} \frac{\|\Pi^*\Phi - \Phi\|_{\mathcal{T}_h}}{\|\boldsymbol{\theta}\|_{\Omega}}, \sup_{\boldsymbol{\theta}\in \boldsymbol{L}^2(\Omega)\setminus\{0\}} \frac{\|\nabla(\boldsymbol{\phi}_h - \boldsymbol{\phi})\|_{\mathcal{T}_h}}{\|\boldsymbol{\theta}\|_{\Omega}} \right\},$$

$$H_{\boldsymbol{\beta}} := \|\delta\boldsymbol{\beta}\|_{\boldsymbol{L}^{\infty}(\Omega)} \sup_{\boldsymbol{\theta}\in L^2(\Omega)\setminus\{\boldsymbol{0}\}} \frac{\|\nabla\boldsymbol{\phi}_h\|_{\mathcal{T}_h}}{\|\boldsymbol{\theta}\|_{\Omega}}.$$

Proof Recall that

$$\|\boldsymbol{E}^{\boldsymbol{u}}\|_{\Omega} = \sup_{\boldsymbol{\theta}\in \boldsymbol{L}^2(\Omega)\setminus\{\boldsymbol{0}\}} \frac{(\boldsymbol{E}^{\boldsymbol{u}}, \boldsymbol{\theta})_{\mathcal{T}_h}}{\|\boldsymbol{\theta}\|_{\Omega}}.$$

Based on the identity of the previous lemma, it suffices to bound the terms T_{L} and $T_{\boldsymbol{\beta}}$. Therefore, the result is just a simple consequence of (3.10), (3.5) and the definitions of H_{L} and $H_{\boldsymbol{\beta}}$. □

Now, we need to estimate H_{L} and $H_{\boldsymbol{\beta}}$.

Proposition 3.6 *Assume that* $\mathrm{tr}\,\mathrm{L} = 0$. *Then*, *there holds*

$$\mathrm{H}_{\mathrm{L}} \le C C_r \nu \max_{K\in\mathcal{T}_h}\left(C^K_{\mathrm{H_L}}\right) h^{\min(k,1)},$$

where

$$C^K_{\mathrm{H_L}} := 1 + \left(\tau_t + \frac{1+\nu}{\nu} h_K |\boldsymbol{\beta}|_{\boldsymbol{W}^{1,\infty}(K)}\right)\left(\frac{(\widetilde{\Lambda}^{\max}_K + h_K|\boldsymbol{\beta}|_{W^{1,\infty}(K)})h_K + \nu}{\Gamma^{\min}_K} + h_K\right).$$

Proof We will bound the first term using the approximation property of the projection Π^*_h and the regularity assumption (2.12). Let $k_\sigma = k_\Phi = k_\phi = 0$ and $k_{\boldsymbol{\phi}} = \min(k, 1)$ in Theorem 2.5. Then,

$$\begin{aligned}
\nu\|\Pi^*\Phi - \Phi\|_K &\le C h_K \nu|\Phi|_{1,K} + C\widetilde{C}_2\nu\|\boldsymbol{\Pi}^*\boldsymbol{\phi} - \boldsymbol{\phi}\|_K \\
&\quad + C\widetilde{C}_1 h_K^{\min(k,1)+1}\nu|\boldsymbol{\phi}|_{\min(k,1)+1,K} + \|\Pi^*\phi - \phi\|_K \\
&\le C h_K \nu|\Phi|_{1,K} + C\widetilde{C}_2\left(\frac{\widetilde{\Lambda}^{\max}_K + h_K|\boldsymbol{\beta}|_{W^{1,\infty}(K)}}{\Gamma^{\min}_K}\right) h_K^{\min(k,1)+1}\nu|\boldsymbol{\phi}|_{\min(k,1)+1,K} \\
&\quad + C\frac{\nu\widetilde{C}_2}{\Gamma^{\min}_K} h_K \left|\nabla\cdot(\nu\Phi + \phi\mathrm{I})\right|_K
\end{aligned}$$

$$+ C\widetilde{C}_1 h_K^{\min(k,1)+1} \nu |\boldsymbol{\phi}|_{\min(k,1)+1,K} + C h_K |\phi|_{1,K},$$

where $\widetilde{C}_1 := \tau_t + \frac{1}{\nu} h_K |\boldsymbol{\beta}|_{\boldsymbol{W}^{1,\infty}(K)} \le \tau_t + \frac{1+\nu}{\nu} h_K |\boldsymbol{\beta}|_{\boldsymbol{W}^{1,\infty}(K)} =: \widetilde{C}_2$. Now using the regularity estimate (2.12), as $\min(k,1) \le 1$, $h_K \le h_K^{\min(k,1)}$ and $\widetilde{C}_1 \le \widetilde{C}_2$,

$$\nu \left\| \Pi^* \Phi - \Phi \right\|_K \le C C_{\mathrm{HL}}^K C_r h_K^{\min(k,1)} \|\boldsymbol{\theta}\|_K,$$

For the second term, observe that by the regularity estimate,

$$\left\| \nabla(\boldsymbol{\phi}_h - \boldsymbol{\phi}) \right\|_K \le C h_K |\boldsymbol{\phi}|_{2,K} \le C C_r h_K \|\boldsymbol{\theta}\|_K, \tag{3.15}$$

for $k \ge 1$. The result follows from (3.13), (3.15) and $h_K \le h_K^{\min(k,1)} \le h^{\min(k,1)}$. □

Proposition 3.7 *There holds*

$$H_{\boldsymbol{\beta}} \le C_r h |\boldsymbol{\beta}|_{W^{1,\infty}(\Omega)}.$$

Proof We bound $H_{\boldsymbol{\beta}} = \|\boldsymbol{\delta\beta}\|_{\boldsymbol{L}^\infty(\Omega)} \sup_{\boldsymbol{\theta} \in L^2(\Omega)\setminus\{\boldsymbol{0}\}} \frac{\|\nabla \boldsymbol{\phi}_h\|_{\mathcal{T}_h}}{\|\boldsymbol{\theta}\|_\Omega}$ the same way as we bound the second term in the definition of H_{L}. From (3.13),

$$\|\nabla \boldsymbol{\phi}_h\|_K \le \|\nabla \boldsymbol{\phi}\|_K \le C_r \|\boldsymbol{\theta}\|_\Omega.$$

So the result follows. □

Proposition 3.8 (Estimate for the velocity trace)

$$\left\| E^{\widehat{\boldsymbol{u}}} \right\|_h \le C\big(h \|\Pi \mathrm{L} - \mathrm{L}\|_\Omega + \left\| E^{\boldsymbol{u}} \right\|_\Omega\big). \tag{3.16}$$

Proof The proof is exactly the same as in [17]. In summary, it uses the error equation (3.1a) and a scaling argument. □

In the proof of the pressure result we need the following space and its orthogonal decomposition. Define

$$R_k(\partial K) := \big\{\delta \in L^2(\partial K) : \delta|_F \in P_k(F),\ F \in \mathcal{E}(K)\big\}.$$

Lemma 3.9 *The following decomposition is orthogonal in* $L^2(\partial K)$:

$$R_k(\partial K) = \big\{\boldsymbol{v} \cdot \boldsymbol{n}|_{\partial K} : \boldsymbol{v} \in \boldsymbol{P}_k(K)^\perp\big\} \oplus \big\{q|_{\partial K} : q \in P_k(K)^\perp\big\}.$$

Proof [9, Lemma 4.5] □

At this point let us recall the well-known inf-sup condition [18] that we use to prove the estimate on the pressure.

$$\kappa \le \inf_{q \in L_0^2(\Omega)} \sup_{\boldsymbol{w} \in \boldsymbol{H}_0^1(\Omega)\setminus\{\boldsymbol{0}\}} \frac{(q, \nabla \cdot \boldsymbol{w})_\Omega}{\|q\|_\Omega \|\boldsymbol{w}\|_{1,\Omega}}, \tag{3.17}$$

where $\kappa > 0$ is independent of q and $\boldsymbol{w}$. Here $L_0^2(\Omega)$ is a subspace of $L^2(\Omega)$ such that

$$L_0^2(\Omega) = \left\{ q \in L^2(\Omega) : \int_\Omega q = 0 \right\}.$$

We also introduce a projection $\boldsymbol{P} : \boldsymbol{H}^1(\mathcal{T}_h) \to \boldsymbol{V}_h$ satisfying

$$(\boldsymbol{P}\boldsymbol{w} - \boldsymbol{w}, \boldsymbol{v})_K = 0, \quad \forall \boldsymbol{v} \in \boldsymbol{P}_{k-1}(K), \tag{3.18}$$

$$\big\langle (\boldsymbol{P}\boldsymbol{w} - \boldsymbol{w}) \cdot \boldsymbol{n}, \boldsymbol{\mu} \cdot \boldsymbol{n} \big\rangle_{\partial K} = 0, \quad \forall \boldsymbol{\mu} \in \boldsymbol{P}_k(K)^{\perp}, \tag{3.19}$$

for all $K \in \mathcal{T}_h$.

Lemma 3.10 (Estimate for the pressure) *The following estimate holds*

$$\|E^p\|_{\Omega} \le \frac{1}{\kappa}\big(H_p^1 \|\Pi \mathrm{L} - \mathrm{L}\|_{\Omega} + H_p^2 \|\boldsymbol{E}^{\boldsymbol{u}}\|_{\Omega}\big).$$

where

$$H_p^1 = C \max\left(\nu,\ \|\boldsymbol{\beta}_0\|_{L^\infty(\Omega)} \sup_{\boldsymbol{w}\in \boldsymbol{H}_0^1(\Omega)\setminus\{\boldsymbol{0}\}} \frac{\|\boldsymbol{P}\boldsymbol{w}\|_{\mathcal{T}_h}}{\|\boldsymbol{w}\|_{1,\Omega}},\ h|\boldsymbol{\beta}|_{\boldsymbol{W}^{1,\infty}(\Omega)} \sup_{\boldsymbol{w}\in \boldsymbol{H}_0^1(\Omega)\setminus\{\boldsymbol{0}\}} \frac{\|\boldsymbol{P}\boldsymbol{w}\|_{\mathcal{T}_h}}{\|\boldsymbol{w}\|_{1,\Omega}}\right),$$

$$H_p^2 = C|\boldsymbol{\beta}|_{\boldsymbol{W}^{1,\infty}(\Omega)} \max\left(h \sup_{\boldsymbol{w}\in \boldsymbol{H}_0^1(\Omega)\setminus\{0\}} \frac{\|\nabla \boldsymbol{P}\boldsymbol{w}\|_{\mathcal{T}_h}}{\|\boldsymbol{w}\|_{1,\Omega}},\ \sup_{\boldsymbol{w}\in \boldsymbol{H}_0^1(\Omega)\setminus\{0\}} \frac{\|\boldsymbol{P}\boldsymbol{w}\|_{\mathcal{T}_h}}{\|\boldsymbol{w}\|_{1,\Omega}}\right).$$

Proof Observe from (3.1f), $q := E^p \in L_0^2(\Omega)$. Hence by the inf-sup condition (3.17),

$$\|E^p\|_{\Omega} \le \frac{1}{\kappa} \sup_{\boldsymbol{w}\in \boldsymbol{H}_0^1(\Omega)\setminus\{\boldsymbol{0}\}} \frac{(E^p, \nabla\cdot \boldsymbol{w})_{\Omega}}{\|\boldsymbol{w}\|_{1,\Omega}}. \tag{3.20}$$

Let us rewrite the numerator. First integrate by parts and use the projection $\boldsymbol{P}$.

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

Then, using (3.1b) with $\boldsymbol{v} = \boldsymbol{P}\boldsymbol{w}$,

$$\begin{aligned}\big(E^p, \nabla\cdot\boldsymbol{w}\big)_{\Omega} &= -\big(\nu\nabla\cdot \mathrm{E}^{\mathrm{L}}, \boldsymbol{P}\boldsymbol{w}\big)_{\mathcal{T}_h} - \big(\boldsymbol{E}^{\boldsymbol{u}} \otimes \boldsymbol{\beta}, \nabla\boldsymbol{P}\boldsymbol{w}\big)_{\mathcal{T}_h}\\ &\quad + \big\langle \big(\boldsymbol{E}^{\widehat{\boldsymbol{u}}} \otimes \boldsymbol{\beta}\big)\boldsymbol{n} + \mathrm{S}\big(\boldsymbol{E}^{\boldsymbol{u}} - \boldsymbol{E}^{\widehat{\boldsymbol{u}}}\big), \boldsymbol{P}\boldsymbol{w}\big\rangle_{\partial\mathcal{T}_h} + \big\langle E^p\boldsymbol{n}, \boldsymbol{w}\big\rangle_{\partial\mathcal{T}_h}.\end{aligned}$$

Integrating by parts again and using the definitions of $\boldsymbol{P}$ and $\boldsymbol{P}_M$,

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

Finally, by (3.1e) with $\boldsymbol{\mu} = \boldsymbol{P}_M\boldsymbol{w}$ and as $\boldsymbol{w} = 0$ on $\partial\Omega$,

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

Observe that $\langle \mathrm{S}(\boldsymbol{E}^{\boldsymbol{u}} - \boldsymbol{E}^{\widehat{\boldsymbol{u}}}), \boldsymbol{P}\boldsymbol{w} - \boldsymbol{P}_M\boldsymbol{w}\rangle_{\partial\mathcal{T}_h} = 0$. Indeed, (3.19) and the orthogonal decomposition given in Lemma 3.9 guarantees the existence of $q \in P_k(K)^{\perp}$ such that

$$(\boldsymbol{P}\boldsymbol{w} - \boldsymbol{P}_M\boldsymbol{w}) \cdot \boldsymbol{n} = q|_{\partial K}.$$

Then,

$$\begin{aligned}\big\langle \mathrm{S}\big(\boldsymbol{E}^{\boldsymbol{u}} - \boldsymbol{E}^{\widehat{\boldsymbol{u}}}\big)\cdot\boldsymbol{n}, (\boldsymbol{P}\boldsymbol{w} - \boldsymbol{P}_M\boldsymbol{w})\cdot\boldsymbol{n}\big\rangle_{\partial K} &= \tau_n\big\langle\big(\boldsymbol{E}^{\boldsymbol{u}} - \boldsymbol{E}^{\widehat{\boldsymbol{u}}}\big)\cdot\boldsymbol{n}, (\boldsymbol{P}\boldsymbol{w} - \boldsymbol{P}_M\boldsymbol{w})\cdot\boldsymbol{n}\big\rangle_{\partial K}\\ &= \tau_n\big\langle\big(\boldsymbol{E}^{\boldsymbol{u}} - \boldsymbol{E}^{\widehat{\boldsymbol{u}}}\big)\cdot\boldsymbol{n}, q\big\rangle_{\partial K}\\ &= \tau_n\big(\nabla\cdot\boldsymbol{E}^{\boldsymbol{u}}, q\big)_K = 0,\end{aligned}$$

where we used (3.1c) and the fact that $q \in P_k(K)^\perp$ in the last equality. Before bounding the right hand side of this equation let us first rewrite the third term. We have,

$$\begin{aligned}\big\langle\big(\boldsymbol{E}^{\widehat{\boldsymbol{u}}}\otimes\boldsymbol{\beta}\big)\boldsymbol{n}, \boldsymbol{P}\boldsymbol{w}\big\rangle_{\partial\mathcal{T}_h} &= \big\langle\boldsymbol{E}^{\widehat{\boldsymbol{u}}}, (\boldsymbol{P}\boldsymbol{w}\otimes\boldsymbol{\beta})\boldsymbol{n}\big\rangle_{\partial\mathcal{T}_h}\\ &= \big\langle\boldsymbol{E}^{\widehat{\boldsymbol{u}}}, (\boldsymbol{P}\boldsymbol{w}\otimes\boldsymbol{\beta}_0)\boldsymbol{n}\big\rangle_{\partial\mathcal{T}_h} + \big\langle\boldsymbol{E}^{\widehat{\boldsymbol{u}}}, (\boldsymbol{P}\boldsymbol{w}\otimes\boldsymbol{\delta\beta})\boldsymbol{n}\big\rangle_{\partial\mathcal{T}_h}.\end{aligned}$$

Then, picking $\mathrm{G} = \boldsymbol{P}\boldsymbol{w}\otimes\boldsymbol{\beta}_0$ in (3.1a),

$$\big\langle\big(\boldsymbol{E}^{\widehat{\boldsymbol{u}}}\otimes\boldsymbol{\beta}\big)\boldsymbol{n}, \boldsymbol{P}\boldsymbol{w}\big\rangle_{\partial\mathcal{T}_h} = (\mathrm{L}-\mathrm{L}_h, \boldsymbol{P}\boldsymbol{w}\otimes\boldsymbol{\beta}_0)_{\mathcal{T}_h} + \big(\boldsymbol{E}^{\boldsymbol{u}}, \boldsymbol{\beta}_0\cdot\nabla\boldsymbol{P}\boldsymbol{w}\big)_{\mathcal{T}_h} + \big\langle\boldsymbol{E}^{\widehat{\boldsymbol{u}}}, (\boldsymbol{P}\boldsymbol{w}\otimes\boldsymbol{\delta\beta})\boldsymbol{n}\big\rangle_{\partial\mathcal{T}_h}.$$

Hence,

$$\begin{aligned}\big(E^p, \nabla\cdot\boldsymbol{w}\big)_\Omega &= \big(\nu\mathrm{E}^{\mathrm{L}}, \nabla\boldsymbol{w}\big)_{\mathcal{T}_h} - \big(\boldsymbol{E}^{\boldsymbol{u}}\otimes\boldsymbol{\delta\beta}, \nabla\boldsymbol{P}\boldsymbol{w}\big)_{\mathcal{T}_h}\\ &\quad + (\mathrm{L}-\mathrm{L}_h, \boldsymbol{P}\boldsymbol{w}\otimes\boldsymbol{\beta}_0)_{\mathcal{T}_h} + \big\langle\boldsymbol{E}^{\widehat{\boldsymbol{u}}}, (\boldsymbol{P}\boldsymbol{w}\otimes\boldsymbol{\delta\beta})\boldsymbol{n}\big\rangle_{\partial\mathcal{T}_h}.\end{aligned}$$

Let us consider the term $\langle\boldsymbol{E}^{\widehat{\boldsymbol{u}}}, (\boldsymbol{P}\boldsymbol{w}\otimes\boldsymbol{\delta\beta})\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}$ first. By the Cauchy-Schwarz inequality and the trace inequality,

$$\begin{aligned}\big\langle\boldsymbol{E}^{\widehat{\boldsymbol{u}}}, (\boldsymbol{P}\boldsymbol{w}\otimes\boldsymbol{\delta\beta})\boldsymbol{n}\big\rangle_{\partial K} &\le \big\|\boldsymbol{E}^{\widehat{\boldsymbol{u}}}\big\|_{\partial K} h_K^{-1/2}\|\boldsymbol{P}\boldsymbol{w}\|_K h_K|\boldsymbol{\beta}|_{\boldsymbol{W}^{1,\infty}(K)}\\ &= h_K^{1/2}\big\|\boldsymbol{E}^{\widehat{\boldsymbol{u}}}\big\|_{\partial K}\|\boldsymbol{P}\boldsymbol{w}\|_K|\boldsymbol{\beta}|_{\boldsymbol{W}^{1,\infty}(K)}.\end{aligned}$$

Therefore summing up over all $K \in \mathcal{T}_h$,

$$\big\langle\boldsymbol{E}^{\widehat{\boldsymbol{u}}}, (\boldsymbol{P}\boldsymbol{w}\otimes\boldsymbol{\delta\beta})\boldsymbol{n}\big\rangle_{\partial\mathcal{T}_h} \le \big\|\boldsymbol{E}^{\widehat{\boldsymbol{u}}}\big\|_h\|\boldsymbol{P}\boldsymbol{w}\|_{\mathcal{T}_h}|\boldsymbol{\beta}|_{\boldsymbol{W}^{1,\infty}(\Omega)}.$$

Hence, again by the Cauchy-Schwarz inequality,

$$\begin{aligned}\big|\big(E^p, \nabla\cdot\boldsymbol{w}\big)_\Omega\big| &\le \nu\big\|\mathrm{E}^{\mathrm{L}}\big\|_{\mathcal{T}_h}\|\nabla\boldsymbol{w}\|_{\mathcal{T}_h} + h|\boldsymbol{\beta}|_{\boldsymbol{W}^{1,\infty}(\Omega)}\big\|\boldsymbol{E}^{\boldsymbol{u}}\big\|_{\mathcal{T}_h}\|\nabla\boldsymbol{P}\boldsymbol{w}\|_{\mathcal{T}_h}\\ &\quad + \|\boldsymbol{\beta}_0\|_{\boldsymbol{L}^\infty(\Omega)}\big(\|\Pi\mathrm{L}-\mathrm{L}\|_{\mathcal{T}_h} + \big\|\mathrm{E}^{\mathrm{L}}\big\|_{\mathcal{T}_h}\big)\|\boldsymbol{P}\boldsymbol{w}\|_{\mathcal{T}_h}\\ &\quad + \big\|\boldsymbol{E}^{\widehat{\boldsymbol{u}}}\big\|_h\|\boldsymbol{P}\boldsymbol{w}\|_{\mathcal{T}_h}|\boldsymbol{\beta}|_{\boldsymbol{W}^{1,\infty}(\Omega)}.\end{aligned}$$

Then, from Corollary 3.3 and Proposition 3.8,

$$\big|\big(E^p, \nabla\cdot\boldsymbol{w}\big)_\Omega\big| \le \big(H_p^1\|\Pi\mathrm{L}-\mathrm{L}\|_{\mathcal{T}_h} + H_p^2\big\|\boldsymbol{E}^{\boldsymbol{u}}\big\|_\Omega\big)\|\boldsymbol{w}\|_{1,\Omega},$$

and by (3.20),

$$\big\|E^p\big\|_\Omega \le \frac{1}{\kappa}\big(H_p^1\|\Pi\mathrm{L}-\mathrm{L}\|_{\mathcal{T}_h} + H_p^2\big\|\boldsymbol{E}^{\boldsymbol{u}}\big\|_\Omega\big). \qquad \square$$

Now let us bound the terms H_p^1 and H_p^2 that show up in the estimate for the pressure. We will need the following projection, the approximation properties of which can be found in [13, Proposition 2.1], for the proof:

For any $z \in H^1(K)$ and for any given face F of K, define $P_F z \in P_k(K)$ by

$$(P_F z, w)_K = (z, w), \quad \forall w \in P_{k-1}(K),$$
$$\langle P_F z, \mu\rangle_F = \langle z, \mu\rangle_F, \quad \forall \mu \in P_k(F).$$

We can also define vector valued and matrix valued projections $\boldsymbol{P}_F$ and P_F by considering P_F applied to each component separately.

Proposition 3.11

$$H_p^1 \le C \max\bigl(\nu, |\boldsymbol{\beta}_0|, h|\boldsymbol{\beta}|_{W^{1,\infty}(\Omega)}\bigr), \tag{3.21}$$

$$H_p^2 \le C|\boldsymbol{\beta}|_{W^{1,\infty}(\Omega)}. \tag{3.22}$$

Proof We will basically show that H_p^1 and H_p^2 are $\mathcal{O}(1)$. Observe that from an inverse inequality,

$$\|\nabla \boldsymbol{P w}\|_{\mathcal{T}_h} = \bigl\|\nabla(\boldsymbol{P w} - c)\bigr\|_{\mathcal{T}_h} \le h^{-1}\|\boldsymbol{P w} - c\|_{\mathcal{T}_h} \le C\|\nabla \boldsymbol{w}\|_{\mathcal{T}_h} \le C\|\boldsymbol{w}\|_{1,\Omega},$$

where c is a piecewise constant. Therefore,

$$\sup_{\boldsymbol{w}\in \boldsymbol{H}_0^1(\Omega)\setminus\{0\}} \frac{\|\nabla \boldsymbol{P w}\|_{\mathcal{T}_h}}{\|\boldsymbol{w}\|_{1,\Omega}} \le C.$$

Lastly, as a consequence of the approximation properties of the projection $\boldsymbol{P}_F$,

$$\|\boldsymbol{P w}\|_{\mathcal{T}_h} \le C\|\boldsymbol{w}\|_{1,\Omega}$$

which yields (3.21). The bound (3.21) follows trivially from this. □

Equation (2.15) is a trivial consequence of Lemma 3.10 and Proposition 3.11.

3.3 Characterization and Approximation Properties of Π_h

Next we prove Theorem 2.3. In order to do that we need auxiliary results that we state next. We omit the proofs and refer the reader to [17, Lemma A.1], [16, Lemma 4.8, Proposition 4.9]. The first lemma is necessary to prove (2.9).

Lemma 3.12 *For all $\boldsymbol{v} \in \boldsymbol{P}_k(K)^\perp$ and for any face F of K, we have*

$$\|\boldsymbol{v}\|_K \le C h_K^{1/2}\|\boldsymbol{v}\|_F.$$

The next two results will be used in the proofs of (2.10) and (2.8).

Lemma 3.13 *The set $\mathcal{B} := \{\mathrm{I}_n\} \cup \{\boldsymbol{t} \otimes \boldsymbol{n}_F : F \text{ is a face of } K, \boldsymbol{t} \in \mathcal{B}_F\}$ is a basis of the space of $n \times n$ matrices where $\mathcal{B}_F$ is an orthogonal basis of the vectors orthogonal to $\boldsymbol{n}_F$ for each face F of K. Here I_n denotes the $n \times n$ identity matrix.*

The dual basis of $\mathcal{B}$ is of the form

$$\mathcal{B}^* := \left\{\frac{1}{n}\mathrm{I}\right\} \cup \{\mathrm{W}_{F,\boldsymbol{t}} : F \text{ is a face of } K,\ \boldsymbol{t} \in \mathcal{B}_F\}. \tag{3.23}$$

Note that $\mathrm{W}_{F,\boldsymbol{t}}$ is uniformly bounded with respect to F and $\boldsymbol{t}$ and the bound depends only on the shape regularity parameter of the mesh. As a consequence we have the following representation for a square matrix.

Lemma 3.14 *Let* A *be an* $n \times n$ *matrix. Then,*

$$\mathrm{A} = \sum_F \sum_{\boldsymbol{t}\in\mathcal{B}_F} \big(\mathrm{A} : (\boldsymbol{t} \otimes \boldsymbol{n}_F)\big) W_{F,\boldsymbol{t}} + \frac{\operatorname{tr}\mathrm{A}}{n}\mathrm{I}_n, \tag{3.24}$$

where the first sum is over all faces F *of* K.

In order to have a simpler analysis for the a priori error estimation, we find suitable characterizations for the projection Π_h. To prove (2.9) we characterize $\boldsymbol{\Pi u}$ independent of $\Pi\mathrm{L}$ and Πp in the next theorem.

Proof of the well-posedness of Π_h We want to show that there exists a unique $(\Pi\mathrm{L}, \boldsymbol{\Pi u}, \Pi p)$ which is the solution to the problem defined by (2.7a)–(2.7d). It is sufficient to show uniqueness because the number of unknowns exactly matches the number of equations of the problem. Indeed, the number of equations is computed by

$$\#_E = \underbrace{\overbrace{d^2 \dim P_{k-1}(K)}^{\text{from (2.7a)}} + \overbrace{d \dim P_{k-1}(K)}^{\text{from (2.7b)}} + \overbrace{\dim P_k(K)}^{\text{from (2.7c)}} + \overbrace{d(d+1)\dim P_k(F)}^{\text{from (2.7d)}}}_{=(d^2+d)(\dim P_{k-1}(K)+\dim P_k(F))+\dim P_k(K)},$$

whereas the number of unknows is simply

$$\#_U = \underbrace{\overbrace{d^2 \dim P_k(K)}^{\text{for } \Pi\mathrm{L}} + \overbrace{d \dim P_k(K)}^{\text{for } \boldsymbol{\Pi u}} + \overbrace{\dim P_k(K)}^{\text{for } \Pi p}}_{=(d^2+d)\dim P_k(K)+\dim P_k(K)}.$$

Therefore, $\#_E - \#_U = 0$. To show uniqueness, we set $\mathrm{L} = 0$, $\boldsymbol{u} = \boldsymbol{0}$ and $p = 0$. We need to prove that $\Pi\mathrm{L} = 0$, $\boldsymbol{\Pi u} = \boldsymbol{0}$, $\Pi p = 0$. But this is trivial once we prove the approximation results (2.8)–(2.10) for Π_h. Hence, the rest of the proof is postponed to the part where we obtain the approximation results. □

Theorem 3.15 (Characterization of $\boldsymbol{\Pi u}$) *For any* $K \in \mathcal{T}_h$,

$$(\boldsymbol{\Pi u} - \boldsymbol{u}, \boldsymbol{v})_K = 0, \quad \forall \boldsymbol{v} \in \boldsymbol{P}_{k-1}(K), \tag{3.25a}$$

$$\begin{aligned}&\big\langle \mathrm{S}(\boldsymbol{\Pi u} - \boldsymbol{u}), \boldsymbol{v}\big\rangle_{\partial K} - \big((\boldsymbol{\Pi u} - \boldsymbol{u}) \otimes \boldsymbol{\beta}, \nabla \boldsymbol{v}\big)_K \\ &\quad = -\big(\nabla \cdot (\nu\mathrm{L} - p\mathrm{I}), \boldsymbol{v}\big)_K - \big\langle \big((\boldsymbol{P}_M \boldsymbol{u} - \boldsymbol{u}) \otimes \boldsymbol{\delta\beta}\big)\boldsymbol{n}, \boldsymbol{v}\big\rangle_{\partial K}, \quad \forall \boldsymbol{v} \in \boldsymbol{P}_k(K)^{\perp}.\end{aligned} \tag{3.25b}$$

Proof The first equation (3.25a) is exactly the same as (2.7b). To obtain (3.25b), let $\boldsymbol{\mu} = \boldsymbol{v} \in \boldsymbol{P}_k(K)^\perp$ in (2.7d). Integration by parts yields

$$\begin{aligned}
\langle \mathrm{S}(\boldsymbol{\Pi u} - \boldsymbol{u}), \boldsymbol{v}\rangle_{\partial K} &= \langle \nu(\Pi \mathrm{L} - \mathrm{L})\boldsymbol{n} - (\Pi p - p)\boldsymbol{n} - \big((\boldsymbol{P}_M\boldsymbol{u} - \boldsymbol{u}) \otimes \boldsymbol{\beta}\big)\boldsymbol{n}, \boldsymbol{v}\rangle_{\partial K} \\
&= \big(\nabla \cdot \nu(\Pi \mathrm{L} - \mathrm{L}), \boldsymbol{v}\big)_K + \nu(\Pi \mathrm{L} - \mathrm{L}, \nabla \boldsymbol{v})_K \\
&\quad - \big(\nabla(\Pi p - p), \boldsymbol{v}\big)_K - (\Pi p - p, \nabla \cdot \boldsymbol{v})_K - \langle (\boldsymbol{P}_M\boldsymbol{u} - \boldsymbol{u}) \otimes \boldsymbol{\beta})\boldsymbol{n}, \boldsymbol{v}\rangle_{\partial K} \\
&= \big(\nabla \cdot \nu(\Pi \mathrm{L} - \mathrm{L}), \boldsymbol{v}\big)_K + \big((\boldsymbol{\Pi u} - \boldsymbol{u}) \otimes \boldsymbol{\beta}, \nabla \boldsymbol{v}\big)_K - \big(\nabla(\Pi p - p), \boldsymbol{v}\big)_K \\
&\quad - \langle (\boldsymbol{P}_M\boldsymbol{u} - \boldsymbol{u}) \otimes \boldsymbol{\beta})\boldsymbol{n}, \boldsymbol{v}\rangle_{\partial K},
\end{aligned}$$

by using (2.7a) and (2.7c) for the second equality. Note that we can replace $\langle((\boldsymbol{P}_M\boldsymbol{u} - \boldsymbol{u}) \otimes \boldsymbol{\beta})\boldsymbol{n}, \boldsymbol{v}\rangle_{\partial K}$ by $\langle((\boldsymbol{P}_M\boldsymbol{u} - \boldsymbol{u}) \otimes \boldsymbol{\delta\beta})\boldsymbol{n}, \boldsymbol{v}\rangle_{\partial K}$. Therefore as $\boldsymbol{v} \in \boldsymbol{P}_k(K)^\perp$,

$$\langle \mathrm{S}(\boldsymbol{\Pi u} - \boldsymbol{u}), \boldsymbol{v}\rangle_{\partial K} = -(\nabla \cdot \nu \mathrm{L} - p\mathrm{I}, \boldsymbol{v})_K + \big((\boldsymbol{\Pi u} - \boldsymbol{u}) \otimes \boldsymbol{\beta}, \nabla \boldsymbol{v}\big)_K - \big((\boldsymbol{P}_M\boldsymbol{u} - \boldsymbol{u}) \otimes \boldsymbol{\delta\beta}\big)\boldsymbol{n}, \boldsymbol{v}\rangle_{\partial K}.$$

Hence, (3.25b) follows. □

Now using this characterization we can prove the estimate for $\boldsymbol{\Pi u} - \boldsymbol{u}$.

Proof of (2.9) Let $\boldsymbol{\delta}^{\boldsymbol{u}} := \boldsymbol{\Pi u} - \boldsymbol{u}_k$, where $\boldsymbol{u}_k$ is the L^2-projection of $\boldsymbol{u}$ into $\boldsymbol{P}_k(K)$. Then, we split up the error by the triangle inequality as follows:

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

It is enough to estimate the first term as the second term can be easily estimated using the properties of the L^2-projection. Observe that we can replace $\boldsymbol{u}$ in (3.25a) by $\boldsymbol{u}_k$. This implies that $\boldsymbol{\delta}^{\boldsymbol{u}} \in \boldsymbol{P}_k(K)^\perp$. Also, from (3.25b),

$$\begin{aligned}
&\langle \mathrm{S}\boldsymbol{\delta}^{\boldsymbol{u}}, \boldsymbol{v}\rangle_{\partial K} - \big(\boldsymbol{\delta}^{\boldsymbol{u}} \otimes \boldsymbol{\delta\beta}, \nabla \boldsymbol{v}\big)_K \\
&\quad = \underbrace{\langle \mathrm{S}(\boldsymbol{u} - \boldsymbol{u}_k), \boldsymbol{v}\rangle_{\partial K} - \big((\boldsymbol{u} - \boldsymbol{u}_k) \otimes \boldsymbol{\delta\beta}, \nabla \boldsymbol{v}\big)_K}_{=:b_{\boldsymbol{u}}(\boldsymbol{v})} \underbrace{-\big(\nabla \cdot (\nu \mathrm{L} - p\mathrm{I}), \boldsymbol{v}\big)_K}_{=:b_\sigma(\boldsymbol{v})} \\
&\qquad \underbrace{-\langle (\boldsymbol{P}_M\boldsymbol{u} - \boldsymbol{u}) \otimes \boldsymbol{\delta\beta})\boldsymbol{n}, \boldsymbol{v}\rangle_{\partial K}}_{=:b_p(\boldsymbol{v})},
\end{aligned}$$

for all $\boldsymbol{v} \in \boldsymbol{P}_k(K)^\perp$. Let $\boldsymbol{v} = \boldsymbol{\delta}^{\boldsymbol{u}}$ above.

$$\langle \mathrm{S}\boldsymbol{\delta}^{\boldsymbol{u}}, \boldsymbol{\delta}^{\boldsymbol{u}}\rangle_{\partial K} - \big(\boldsymbol{\delta}^{\boldsymbol{u}} \otimes \boldsymbol{\delta\beta}, \nabla \boldsymbol{\delta}^{\boldsymbol{u}}\big)_K = b_{\boldsymbol{u}}\big(\boldsymbol{\delta}^{\boldsymbol{u}}\big) + b_\sigma\big(\boldsymbol{\delta}^{\boldsymbol{u}}\big) + b_p\big(\boldsymbol{\delta}^{\boldsymbol{u}}\big).$$

Recall that we can write $(\boldsymbol{\delta}^{\boldsymbol{u}} \otimes \boldsymbol{\delta\beta}, \nabla \boldsymbol{\delta}^{\boldsymbol{u}})_K = \frac{1}{2}\langle(\boldsymbol{\beta} \cdot \boldsymbol{n})\boldsymbol{\delta}^{\boldsymbol{u}}, \boldsymbol{\delta}^{\boldsymbol{u}}\rangle_{\partial K}$. Therefore,

$$\langle \mathrm{S}_{\boldsymbol{\beta}}\boldsymbol{\delta}^{\boldsymbol{u}}, \boldsymbol{\delta}^{\boldsymbol{u}}\rangle_{\partial K} = b_{\boldsymbol{u}}\big(\boldsymbol{\delta}^{\boldsymbol{u}}\big) + b_\sigma\big(\boldsymbol{\delta}^{\boldsymbol{u}}\big) + b_p\big(\boldsymbol{\delta}^{\boldsymbol{u}}\big).$$

As we have the assumption (2.6) and $\mathrm{S}_{\boldsymbol{\beta}}$ is positive definite, using Lemma 3.12 with $\boldsymbol{v} = \boldsymbol{\delta}^{\boldsymbol{u}}$,

$$\langle \mathrm{S}_{\boldsymbol{\beta}}\boldsymbol{\delta}^{\boldsymbol{u}}, \boldsymbol{\delta}^{\boldsymbol{u}}\rangle_{\partial K} \ge \Gamma_K^{\min} \sum_F \sum_{i=1}^n \|\boldsymbol{\delta}^{\boldsymbol{u}} \boldsymbol{n}_{F_i}\|_{F_i}^2 \ge C\Gamma_K^{\min} h_K^{-1} \|\boldsymbol{\delta}^{\boldsymbol{u}}\|_K^2.$$

Therefore,

$$\left\|\boldsymbol{\delta}^{\boldsymbol{u}}\right\|_K^2 \leq C\frac{h_K}{\Gamma_K^{\min}}\langle \mathrm{S}_\beta \boldsymbol{\delta}^{\boldsymbol{u}}, \boldsymbol{\delta}^{\boldsymbol{u}}\rangle_{\partial K} \leq C\frac{h_K}{\Gamma_K^{\min}}\big(\|b_{\boldsymbol{u}}\| + \|b_\sigma\| + \|b_p\|\big)\left\|\boldsymbol{\delta}^{\boldsymbol{u}}\right\|_K,$$

which implies

$$\left\|\boldsymbol{\delta}^{\boldsymbol{u}}\right\|_K \leq C\frac{h_K}{\Gamma_K^{\min}}\big(\|b_{\boldsymbol{u}}\| + \|b_\sigma\| + \|b_p\|\big).$$

The last step is to estimate the norms on the right hand side of the above inequality. Let us begin with the norm of $b_{\boldsymbol{u}}$. By a scaling argument and the trace inequality,

$$\begin{aligned}\left|b_{\boldsymbol{u}}(\boldsymbol{v})\right| &\leq \Lambda_K^{\max}\|\boldsymbol{u}-\boldsymbol{u}_k\|_{\partial K}\|\boldsymbol{v}\|_{\partial K} + \|\boldsymbol{u}-\boldsymbol{u}_k\|_K\|\boldsymbol{\delta\beta}\|_{L^\infty(K)}\|\nabla \boldsymbol{v}\|_K \\ &\leq C\Lambda_K^{\max} h_K^{-1}\big(\|\boldsymbol{u}-\boldsymbol{u}_k\|_K + h_K\left\|\nabla(\boldsymbol{u}-\boldsymbol{u}_k)\right\|_K\big)\|\boldsymbol{v}\|_K \\ &\quad + C\|\boldsymbol{u}-\boldsymbol{u}_k\|_K|\boldsymbol{\beta}|_{W^{1,\infty}(K)}\|\boldsymbol{v}\|_K.\end{aligned}$$

Then, from the approximation properties of the L^2-projection,

$$\|b_{\boldsymbol{u}}\| \leq C\big(\Lambda_K^{\max} + h_K|\boldsymbol{\beta}|_{W^{1,\infty}(K)}\big)h_K^{k_{\boldsymbol{u}}}|\boldsymbol{u}|_{k_{\boldsymbol{u}}+1,K}.$$

For the norm of b_σ, noting that $\boldsymbol{v} \in \boldsymbol{P}_k(K)^\perp$, we rewrite $b_\sigma(\boldsymbol{v})$ as follows:

$$b_\sigma(\boldsymbol{v}) = -\big(\nabla\cdot(\nu\mathrm{L}-p\mathrm{I}), \boldsymbol{v}\big)_K = -\big(\nabla\cdot(\nu\mathrm{L}-p\mathrm{I}) - \big(\nabla\cdot(\nu\mathrm{L}-p\mathrm{I})\big)_{k-1}, \boldsymbol{v}\big)_K,$$

where $(\cdot)_{k-1}$ denotes the L^2-projection into $\boldsymbol{P}_{k-1}(K)$ such that $(\nabla\cdot(\nu\mathrm{L}-p\mathrm{I}))_{k-1} \equiv 0$ for $k=0$. Then, this implies

$$\|b_\sigma\| \leq \left\|\nabla\cdot(\nu\mathrm{L}-p\mathrm{I}) - \nabla\cdot(\nu\mathrm{L}-p\mathrm{I})_{k-1}\right\|_K \leq Ch_K^{k_\sigma}\left|\nabla\cdot(\nu\mathrm{L}-p\mathrm{I})\right|_{k_\sigma,K}, \quad 0\leq k_\sigma \leq k,$$

by the approximation property of the L^2-projection. Lastly, by a scaling argument, the trace inequality and the approximation properties of $\boldsymbol{P}_M$,

$$\begin{aligned}\left|b_p(\boldsymbol{v})\right| &\leq \|\boldsymbol{P}_M\boldsymbol{u}-\boldsymbol{u}\|_{\partial K}\|\boldsymbol{\delta\beta}\|_{L^\infty(\partial K)}\|\boldsymbol{v}\|_{\partial K} \\ &\leq C\big(\|\boldsymbol{P}_M\boldsymbol{u}-\boldsymbol{u}\|_K + h_K\left\|\nabla(\boldsymbol{P}_M\boldsymbol{u}-\boldsymbol{u})\right\|_K\big)|\boldsymbol{\beta}|_{W^{1,\infty}(K)}\|\boldsymbol{v}\|_K \\ &\leq Ch_K^{k_{\boldsymbol{u}}+1}|\boldsymbol{\beta}|_{W^{1,\infty}(K)}|\boldsymbol{u}|_{k_{\boldsymbol{u}}+1,K}\|\boldsymbol{v}\|_K.\end{aligned}$$

Therefore,

$$\|b_p\| \leq Ch_K^{k_{\boldsymbol{u}}+1}|\boldsymbol{\beta}|_{W^{1,\infty}(K)}|\boldsymbol{u}|_{k_{\boldsymbol{u}}+1,K}.$$

Combining these,

$$\left\|\boldsymbol{\delta}^{\boldsymbol{u}}\right\|_K \leq C\left(\frac{\Lambda_K^{\max} + h_K|\boldsymbol{\beta}|_{W^{1,\infty}(K)}}{\Gamma_K^{\min}}\right)h_K^{k_{\boldsymbol{u}}+1}|\boldsymbol{u}|_{k_{\boldsymbol{u}}+1,K} + \frac{C}{\Gamma_K^{\min}}h_K^{k_\sigma+1}\left|\nabla\cdot(\nu\mathrm{L}-p\mathrm{I})\right|_{k_\sigma,K}.$$

This and the approximation of the L^2-projection yield (2.9). □

We need additional projections for the proof of (2.10) and (2.8). Given a $\mathrm{L}\in\mathrm{H}^1(K)$, we define a projection P^1 as follows: $\mathrm{P}^1\mathrm{L}\in\mathrm{P}_k(K)$ such that

$$\big(\mathrm{P}^1\mathrm{L},\mathrm{G}\big)_K=(\mathrm{L},\mathrm{G})_K,\quad \forall \mathrm{G}\in \mathrm{P}_{k-1}(K),$$
$$\big\langle \mathrm{P}^1\mathrm{L}\boldsymbol{n}_F,\mathrm{G}\big\rangle_F=\langle \mathrm{L}\boldsymbol{n}_F,\boldsymbol{\mu}\rangle_F,\quad \forall \boldsymbol{\mu}\in \boldsymbol{P}_k(F),$$

for all faces F of the simplex K except for an arbitrarily chosen one.

To take care of the convection term, we introduce another projection P^2. $\mathrm{P}^2\mathrm{L}\in \mathrm{P}_k(K)$ is defined such that

$$\big(\nu\mathrm{P}^2\mathrm{L},\mathrm{G}\big)_K=(\nu\mathrm{L},\mathrm{G})_K+\big((\boldsymbol{\Pi u}-\boldsymbol{u})\otimes\boldsymbol{\delta\beta},\mathrm{G}\big),\quad \forall \mathrm{G}\in \mathrm{P}_{k-1}(K),$$
$$\big\langle \nu\mathrm{P}^2\mathrm{L}\boldsymbol{n}_F,\boldsymbol{\mu}\big\rangle_F=\langle \nu\mathrm{L}\boldsymbol{n}_F,\boldsymbol{\mu}\rangle_F+\big\langle\big((\boldsymbol{P}_M\boldsymbol{u}-\boldsymbol{u})\otimes\boldsymbol{\delta\beta}\big)\boldsymbol{n}_F,\boldsymbol{\mu}\big\rangle_F,\quad \forall \boldsymbol{\mu}\in \boldsymbol{P}_k(F),$$

for all faces of the simplex K except for an arbitrary one. The reader may refer to [8, 17] for the well-posedness of these projections.

Theorem 3.16 (Characterization of $\nu\Pi\mathrm{L}-\Pi p\mathrm{I}$) *For any* $K\in\mathcal{T}_h$,

$$\big(\nu(\Pi\mathrm{L}-\mathrm{L})-(\Pi p-p)\mathrm{I},\mathrm{G}\big)_K=\big((\boldsymbol{\Pi u}-\boldsymbol{u})\otimes\boldsymbol{\delta\beta},\mathrm{G}\big)_K,\quad \forall \mathrm{G}\in \mathrm{P}_{k-1}(K),\qquad(3.26\text{a})$$
$$\begin{aligned}\big\langle \nu(\Pi\mathrm{L}-\mathrm{L})\boldsymbol{n}_F-(\Pi p-p)\boldsymbol{n}_F,\boldsymbol{\mu}\big\rangle_F&=\big\langle\big((\boldsymbol{P}_M\boldsymbol{u}-\boldsymbol{u})\otimes\boldsymbol{\delta\beta}\big)\boldsymbol{n}_F,\boldsymbol{\mu}\big\rangle_F\\&\quad+\big\langle \mathrm{S}(\boldsymbol{\Pi u}-\boldsymbol{u}),\boldsymbol{\mu}\big\rangle_F,\quad \forall \boldsymbol{\mu}\in \boldsymbol{P}_k(F),\qquad(3.26\text{b})\end{aligned}$$

for all faces F *of* K.

Proof The first equation follows from (2.7a) and (2.7c) as (2.7b) implies

$$\big(\big((\boldsymbol{u}-\boldsymbol{\Pi u})\otimes\boldsymbol{\beta}\big),\mathrm{G}\big)_K=\big(\big((\boldsymbol{u}-\boldsymbol{\Pi u})\otimes\boldsymbol{\delta\beta}\big),\mathrm{G}\big)_K.$$

For the second equation pick an arbitrary face F of K and let $\boldsymbol{\mu}\in \boldsymbol{P}_k(F)$. Then, there exists $\boldsymbol{w}\in \boldsymbol{P}_k^{\perp}(K)$ such that $\boldsymbol{w}=\boldsymbol{\mu}$ on F. This implies that

$$\begin{aligned}&\big\langle \nu(\Pi\mathrm{L}-\mathrm{L})\boldsymbol{n}_F-(\Pi p-p)\boldsymbol{n}_F,\boldsymbol{w}\big\rangle_F\\&\quad=\big\langle \nu(\Pi\mathrm{L}-\mathrm{L})\boldsymbol{n}-(\Pi p-p)\boldsymbol{n},\boldsymbol{w}\big\rangle_{\partial K}-\big\langle \nu(\Pi\mathrm{L}-\mathrm{L})\boldsymbol{n}-(\Pi p-p)\boldsymbol{n},\boldsymbol{w}\big\rangle_{\partial K\setminus F}.\end{aligned}$$

Equation (2.7d) with $\boldsymbol{\mu}=\boldsymbol{w}$ gives,

$$\big\langle \nu(\Pi\mathrm{L}-\mathrm{L})\boldsymbol{n}-(\Pi p-p)\boldsymbol{n},\boldsymbol{w}\big\rangle_{\partial K\setminus F}=\big\langle (\boldsymbol{P}_M\boldsymbol{u}-\boldsymbol{u})\otimes\boldsymbol{\delta\beta})\boldsymbol{n},\boldsymbol{w}\big\rangle_{\partial K\setminus F}+\big\langle \mathrm{S}(\boldsymbol{\Pi u}-\boldsymbol{u}),\boldsymbol{w}\big\rangle_{\partial K\setminus F}.$$

Therefore,

$$\big\langle \nu(\Pi\mathrm{L}-\mathrm{L})\boldsymbol{n}-(\Pi p-p)\boldsymbol{n},\boldsymbol{\mu}\big\rangle_F=\big\langle\big((\boldsymbol{P}_M\boldsymbol{u}-\boldsymbol{u})\otimes\boldsymbol{\delta\beta}\big)\boldsymbol{n},\boldsymbol{w}\big\rangle_F+\big\langle \mathrm{S}(\boldsymbol{\Pi u}-\boldsymbol{u}),\boldsymbol{w}\big\rangle_F+T,$$

where $T:=\langle \nu(\Pi\mathrm{L}-\mathrm{L})\boldsymbol{n}-(\Pi p-p)\boldsymbol{n},\boldsymbol{w}\rangle_{\partial K}-\langle((\boldsymbol{P}_M\boldsymbol{u}-\boldsymbol{u})\otimes\boldsymbol{\delta\beta})\boldsymbol{n},\boldsymbol{w}\rangle_{\partial K}-\langle \mathrm{S}(\boldsymbol{\Pi u}-\boldsymbol{u}),\boldsymbol{w}\rangle_{\partial K}$. From (3.25b),

$$T=\big\langle \nu(\Pi\mathrm{L}-\mathrm{L})\boldsymbol{n}-(\Pi p-p)\boldsymbol{n},\boldsymbol{w}\big\rangle_{\partial K}-\big((\boldsymbol{\Pi u}-\boldsymbol{u})\otimes\boldsymbol{\beta},\nabla\boldsymbol{w}\big)_K+\big(\nabla\cdot(\nu\mathrm{L}-p\mathrm{I}),\boldsymbol{w}\big)_K.$$

Integrating by parts and cancelling terms,

$$\begin{aligned}T&=\big(\nabla\cdot\big(\nu(\Pi\mathrm{L}-\mathrm{L})-(\Pi p-p)\mathrm{I}\big),\boldsymbol{w}\big)_K+\big(\nu(\Pi\mathrm{L}-\mathrm{L})-(\Pi p-p)\mathrm{I},\nabla\boldsymbol{w}\big)_K\\&\quad-\big((\boldsymbol{\Pi u}-\boldsymbol{u})\otimes\boldsymbol{\beta},\nabla\boldsymbol{w}\big)_K+\big(\nabla\cdot(\nu\mathrm{L}-p\mathrm{I}),\boldsymbol{w}\big)_K\end{aligned}$$

$$= \big(\nabla\cdot(\nu \varPi \mathrm{L} - \varPi p \mathrm{I}), \boldsymbol{w}\big)_K + \big(\nu(\varPi \mathrm{L} - \mathrm{L}) - (\varPi p - p)\mathrm{I}, \nabla \boldsymbol{w}\big)_K$$
$$- \big((\boldsymbol{\varPi u} - \boldsymbol{u}) \otimes \boldsymbol{\beta}, \nabla \boldsymbol{w}\big)_K.$$

The first term vanishes as $\boldsymbol{w} \in \boldsymbol{P}_k^\perp(K)$ and the sum of the second and third terms vanishes by (3.26a). Since F was arbitrary the result follows. □

The rest of the proof of Theorem 2.3 comes next.

Proof of (2.10) From the above result and Theorem 3.16, we will first represent $\varPi\mathrm{L} - \mathrm{L}$ in terms of the dual basis $\mathcal{B}^*$ that we defined in (3.23). If we take $\mathrm{A} = \nu(\varPi\mathrm{L} - \mathrm{L})$ in (3.24) and use the fact that $\boldsymbol{M} : (\boldsymbol{t} \otimes \boldsymbol{n}_F) = (\boldsymbol{M}\boldsymbol{n}_F)\cdot \boldsymbol{t}$ for any matrix $\boldsymbol{M}$, we get

$$\nu(\varPi\mathrm{L} - \mathrm{L}) = \sum_F \sum_{\boldsymbol{t}\in\mathcal{B}_F} \big(\nu(\varPi\mathrm{L} - \mathrm{L})\boldsymbol{n}_F\cdot \boldsymbol{t}\big)\mathrm{W}_{F,\boldsymbol{t}} + \nu\frac{\mathrm{tr}(\varPi\mathrm{L} - \mathrm{L})}{n}\mathrm{I}_n. \tag{3.27}$$

Thus, our aim is to find estimates for $\|\nu(\varPi\mathrm{L} - \mathrm{L})\boldsymbol{n}_F\cdot\boldsymbol{t}\|_K$ for each face F of K and $\|\nu\frac{\mathrm{tr}(\varPi\mathrm{L}-\mathrm{L})}{n}\mathrm{I}_n\|_K$. Let us begin with $\|\nu(\varPi\mathrm{L} - \mathrm{L})\boldsymbol{n}_F\cdot\boldsymbol{t}\|_K$ and rewrite $\nu(\varPi\mathrm{L} - \mathrm{L})\boldsymbol{n}_F$ as follows:

$$\nu(\varPi\mathrm{L} - \mathrm{L})\boldsymbol{n}_F = \nu\big(\varPi\mathrm{L} - \mathrm{P}^2\mathrm{L}\big)\boldsymbol{n}_F + \nu\big(\mathrm{P}^2\mathrm{L} - \mathrm{P}^1\mathrm{L}\big)\boldsymbol{n}_F + \nu\big(\mathrm{P}^1\mathrm{L} - \mathrm{L}\big)\boldsymbol{n}_F. \tag{3.28}$$

The identities in Theorem 3.16 together with the fact that S is constant on each face F of K imply that

$$\nu\big(\varPi\mathrm{L} - \mathrm{P}^2\mathrm{L}\big)\boldsymbol{n}_F = (\varPi p - P_F p)\boldsymbol{n}_F + \mathrm{S}|_F(\boldsymbol{\varPi u} - P_F\boldsymbol{u}), \tag{3.29}$$

for any K. To prove this let $\boldsymbol{\delta} := \nu(\varPi\mathrm{L} - \mathrm{P}^2\mathrm{L})\boldsymbol{n}_F - (\varPi p - P_F p)\boldsymbol{n}_F - \mathrm{S}|_F(\boldsymbol{\varPi u} - \boldsymbol{P}_F\boldsymbol{u})$. It is enough to show by Lemma 3.12 that $\boldsymbol{\delta} \in \boldsymbol{P}_k^\perp(K)$ and $\boldsymbol{\delta} = 0$ on F. This is equivalent to proving

(a) $(\boldsymbol{\delta}, \boldsymbol{w})_K = 0, \forall \boldsymbol{w} \in \boldsymbol{P}_{k-1}(K)$,
(b) $\langle\boldsymbol{\delta}, \boldsymbol{\mu}\rangle_F = 0, \forall \boldsymbol{\mu} \in \boldsymbol{P}_k(F)$.

To prove (a), let $\boldsymbol{w} \in \boldsymbol{P}_{k-1}(K)$. As $\mathrm{S}|_F$ is symmetric and constant, by the definition of P^2 and P_F,

$$\begin{aligned}(\boldsymbol{\delta}, \boldsymbol{w})_K &= \big(\nu\big(\varPi\mathrm{L} - \mathrm{P}^2\mathrm{L}\big) - (\varPi p - P_F p)\mathrm{I}, \boldsymbol{w}\otimes\boldsymbol{n}_F\big)_K - (\boldsymbol{\varPi u} - \boldsymbol{P}_F\boldsymbol{u}, \mathrm{S}|_F\boldsymbol{w})_K \\ &= \big(\nu(\varPi\mathrm{L} - \mathrm{L}) - (\varPi p - p)\mathrm{I}, \boldsymbol{w}\otimes\boldsymbol{n}_F\big)_K - \big((\boldsymbol{\varPi u} - \boldsymbol{u})\otimes\boldsymbol{\delta\beta}, \boldsymbol{w}\otimes\boldsymbol{n}_F\big)_K \\ &\quad - (\boldsymbol{\varPi u} - \boldsymbol{u}, \mathrm{S}|_F\boldsymbol{w})_K.\end{aligned}$$

Thus, by (3.26a) and the property of $\varPi_h$,

$$(\boldsymbol{\delta}, \boldsymbol{w})_K = 0.$$

To prove (b), let $\boldsymbol{\mu} \in \boldsymbol{P}_k(F)$. Then, similar to before, by the properties of P^2 and P_F and (3.26b),

$$\begin{aligned}\langle\boldsymbol{\delta}, \boldsymbol{\mu}\rangle_F &= \big\langle\nu\big(\varPi\mathrm{L} - \mathrm{P}^2\mathrm{L}\big) - (\varPi p - P_F p)\mathrm{I}, \boldsymbol{\mu}\otimes\boldsymbol{n}_F\big\rangle_F - \big\langle\mathrm{S}|_F(\boldsymbol{\varPi u} - \boldsymbol{P}_F\boldsymbol{u}), \boldsymbol{\mu}\big\rangle_F \\ &= \big\langle\nu(\varPi\mathrm{L} - \mathrm{L}) - (\varPi p - p)\mathrm{I}, \boldsymbol{\mu}\otimes\boldsymbol{n}_F\big\rangle_F - \big\langle(\boldsymbol{P}_M\boldsymbol{u} - \boldsymbol{u})\otimes\boldsymbol{\delta\beta})\boldsymbol{n}_F, \boldsymbol{\mu}\big\rangle_F \\ &\quad - \big\langle\mathrm{S}|_F(\boldsymbol{\varPi u} - \boldsymbol{u}), \boldsymbol{\mu}\big\rangle_F = 0.\end{aligned}$$

Therefore (3.29) holds. With (3.28), (3.29) and the fact that $\boldsymbol{n}_F \perp \boldsymbol{t}$,

$$\nu(\Pi \mathrm{L}-\mathrm{L})\boldsymbol{n}_F\cdot\boldsymbol{t} = \big(\nu\tau_t(\boldsymbol{\Pi u}-\boldsymbol{P}_F\boldsymbol{u}) + \nu\big(\mathrm{P}^2\mathrm{L}-\mathrm{P}^1\mathrm{L}\big)\boldsymbol{n}_F + \nu\big(\mathrm{P}^1\mathrm{L}-\mathrm{L}\big)\boldsymbol{n}_F\big)\cdot\boldsymbol{t}.$$

Observe that by the definition of P^1 and P^2, $\delta^{\mathrm{L}} := (\mathrm{P}^2\mathrm{L}-\mathrm{P}^1\mathrm{L})$ satisfy

$$\big(\delta^{\mathrm{L}},\mathrm{G}\big)_K = \big((\boldsymbol{\Pi u}-\boldsymbol{u})\otimes\boldsymbol{\delta\beta},\mathrm{G}\big)_K,\quad \forall \mathrm{G}\in \mathrm{P}_{k-1}(K),$$
$$\big\langle\delta^{\mathrm{L}}\boldsymbol{n}_F,\boldsymbol{\mu}\big\rangle_F = \big\langle(\boldsymbol{P}_M\boldsymbol{u}-\boldsymbol{u})\otimes\boldsymbol{\delta\beta})\boldsymbol{n}_F,\boldsymbol{\mu}\big\rangle_F,\quad \forall\boldsymbol{\mu}\in\boldsymbol{P}_k(F),$$

for all faces of K except for an arbitrary one. Thus, the second term is bounded by

$$\big\|\big(\mathrm{P}^2\mathrm{L}-\mathrm{P}^1\mathrm{L}\big)\boldsymbol{n}_F\big\|_K \le C|\boldsymbol{\beta}|_{W^{1,\infty}(K)}h_K\|\boldsymbol{\Pi u}-\boldsymbol{u}\|_K.$$

We bound the first term by the approximation property of $\boldsymbol{P}_F$ and the triangle inequality,

$$\|\boldsymbol{\Pi u}-\boldsymbol{P}_F\boldsymbol{u}\|_K \le \|\boldsymbol{\Pi u}-\boldsymbol{u}\|_K + \|\boldsymbol{u}-\boldsymbol{P}_F\boldsymbol{u}\|_K \le \|\boldsymbol{\Pi u}-\boldsymbol{u}\|_K + Ch_K^{k_{\boldsymbol{u}}+1}|\boldsymbol{u}|_{k_{\boldsymbol{u}}+1,K}.$$

The last term is bounded again by the approximation property of P^1 as

$$\big\|\big(\mathrm{P}^1\mathrm{L}-\mathrm{L}\big)\boldsymbol{n}_F\big\|_K \le Ch_K^{k_{\mathrm{L}}+1}|\mathrm{L}|_{k_{\mathrm{L}}+1,K}.$$

Now as $W_{F,\boldsymbol{t}}$ can be bounded uniformly with constants that depends only on the shape regularity of the mesh, we have

$$\begin{aligned}&\Big\|\sum_F\sum_{\boldsymbol{t}\in\mathcal{B}_F}\big(\nu(\Pi\mathrm{L}-\mathrm{L})\boldsymbol{n}_F\cdot\boldsymbol{t}\big)\mathrm{W}_{F,\boldsymbol{t}}\Big\|_K\\ &\quad\le C\nu h_K^{k_{\mathrm{L}}+1}|\mathrm{L}|_{k_{\mathrm{L}}+1,K}\\ &\qquad + C\nu\big(\tau_t+|\boldsymbol{\beta}|_{W^{1,\infty}(K)}h_K\big)\|\boldsymbol{\Pi u}-\boldsymbol{u}\|_K + C\nu\tau_t h_K^{k_{\boldsymbol{u}}+1}|\boldsymbol{u}|_{k_{\boldsymbol{u}}+1,K}.\end{aligned}\tag{3.30}$$

Now let us now bound $\|\nu\frac{\mathrm{tr}(\Pi\mathrm{L}-\mathrm{L})}{n}\mathrm{I}_n\|_K = \nu\|\,\mathrm{tr}(\Pi\mathrm{L}-\mathrm{L})\|_K$. First observe that if we take $\mathrm{G}=q\mathrm{I}$, $q\in P_{k-1}(K)$ in (2.7a), we obtain

$$\big(\nu\,\mathrm{tr}(\Pi\mathrm{L}-\mathrm{L}),q\big)_K = \big(\mathrm{tr}\big((\boldsymbol{\Pi u}-\boldsymbol{u})\otimes\boldsymbol{\delta\beta}\big),q\big)_k,\quad\forall q\in P_{k-1}(K).\tag{3.31}$$

Now using (3.27) in (3.26b) and picking $\mu = w\boldsymbol{n}_{\partial K}$, $w\in P_k^{\perp}(K)$ yields

$$\big\langle\nu\,\mathrm{tr}(\Pi\mathrm{L}-\mathrm{L}),w\big\rangle_{\partial K} = \langle\zeta,w\rangle_{\partial K},\quad\forall w\in P_k^{\perp}(K),\tag{3.32}$$

where

$$\begin{aligned}\zeta := -\Big(\Big(\sum_F\sum_{t\in\mathcal{B}_F}\big(\nu(\Pi\mathrm{L}-\mathrm{L})\boldsymbol{n}_F\cdot\boldsymbol{t}\big)\mathrm{W}_{F,t}\Big)\boldsymbol{n}_{\partial K}\Big)\cdot\boldsymbol{n}_{\partial K} + (\Pi p - p)\\ + \big((\boldsymbol{P}_M\boldsymbol{u}-\boldsymbol{u})\otimes\boldsymbol{\delta\beta}\big)\boldsymbol{n}_{\partial K}\cdot\boldsymbol{n}_{\partial K} + \nu\tau_n(\boldsymbol{\Pi u}-\boldsymbol{u})\cdot\boldsymbol{n}_{\partial K}.\end{aligned}$$

Observe that we can write the last term as

$$\big\langle\nu\tau_n(\boldsymbol{\Pi u}-\boldsymbol{u})\cdot\boldsymbol{n}_{\partial K},w\big\rangle_{\partial K} = \nu\tau_n\big(\big(\nabla\cdot(\boldsymbol{\Pi u}-\boldsymbol{u}),w\big)_K + (\boldsymbol{\Pi u}-\boldsymbol{u},\nabla w)_K\big).$$

In this equation the first term vanishes as $w \in P_k^{\perp}(K)$ and the fact that $\boldsymbol{u}$ is divergence free. The second term also vanishes by (2.7b). Hence,

$$\zeta := -\left(\left(\sum_F \sum_{t\in\mathcal{B}_F} \big(\nu(\Pi \mathrm{L} - \mathrm{L})\boldsymbol{n}_F \cdot \boldsymbol{t}\big)\mathrm{W}_{F,t}\right)\boldsymbol{n}_{\partial K}\right) \cdot \boldsymbol{n}_{\partial K} + (\Pi p - p)$$
$$+ \big((\boldsymbol{P}_M\boldsymbol{u} - \boldsymbol{u}) \otimes \boldsymbol{\delta\beta}\big)\boldsymbol{n}_{\partial K} \cdot \boldsymbol{n}_{\partial K}.$$

Assuming $\operatorname{tr}\mathrm{L} = 0$ and by (3.31) and (3.32) we have

$$\begin{aligned}\big\|\nu \operatorname{tr}(\Pi \mathrm{L} - \mathrm{L})\big\|_K = \big\|P_F \nu \operatorname{tr}(\Pi \mathrm{L} - \mathrm{L})\big\|_K &\le \big\|P_F \operatorname{tr}\big((\boldsymbol{\Pi u} - \boldsymbol{u}) \otimes \boldsymbol{\delta\beta}\big)\big\|_K + Ch_K^{1/2}\|P_F\zeta\|_F \\ &\le \big\|P_F \operatorname{tr}\big((\boldsymbol{\Pi u} - \boldsymbol{u}) \otimes \boldsymbol{\delta\beta}\big)\big\|_K + C\|P_F\zeta\|_K \\ &\le \big\|\operatorname{tr} P_F\big((\boldsymbol{\Pi u} - \boldsymbol{u}) \otimes \boldsymbol{\delta\beta}\big)\big\|_K + C\|\zeta\|_K.\end{aligned}$$

We bound the second term by using the first part of the proof and the property of the L^2-projection as follows:

$$\begin{aligned}\|\zeta\|_K \le C\nu h_K^{k_\mathrm{L}+1}|\mathrm{L}|_{k_\mathrm{L}+1,K} + C\nu\big(\tau_t + |\boldsymbol{\beta}|_{W^{1,\infty}(K)}h_K\big)\|\boldsymbol{\Pi u} - \boldsymbol{u}\|_K \\ + C\big(\nu\tau_t + h_K|\boldsymbol{\beta}|_{\boldsymbol{W}^{1,\infty}(K)}\big)h_K^{k_{\boldsymbol{u}}+1}|\boldsymbol{u}|_{k_{\boldsymbol{u}}+1,K} + \|\Pi p - p\|_K.\end{aligned} \tag{3.33}$$

Finally, as the trace operator is continuous on finite dimensional spaces,

$$\begin{aligned}\big\|\operatorname{tr} P_F\big((\boldsymbol{\Pi u} - \boldsymbol{u}) \otimes \boldsymbol{\delta\beta}\big)\big\|_K \le C\big\|P_F\big((\boldsymbol{\Pi u} - \boldsymbol{u}) \otimes \boldsymbol{\delta\beta}\big)\big\|_K \le C\big\|(\boldsymbol{\Pi u} - \boldsymbol{u}) \otimes \boldsymbol{\delta\beta}\big\|_K \\ \le Ch_K|\boldsymbol{\beta}|_{W^{1,\infty}(K)}\|\boldsymbol{\Pi u} - \boldsymbol{u}\|_K.\end{aligned}$$

Therefore, combining everything we obtain

$$\begin{aligned}\|\Pi \mathrm{L} - \mathrm{L}\|_K \le C\nu h_K^{k_\mathrm{L}+1}|\mathrm{L}|_{k_\mathrm{L}+1,K} + C\big(\nu\tau_t + (1+\nu)|\boldsymbol{\beta}|_{W^{1,\infty}(K)}h_K\big)\|\boldsymbol{\Pi u} - \boldsymbol{u}\|_K \\ + C\big(\nu\tau_t + h_K|\boldsymbol{\beta}|_{\boldsymbol{W}^{1,\infty}(K)}\big)h_K^{k_{\boldsymbol{u}}+1}|\boldsymbol{u}|_{k_{\boldsymbol{u}}+1,K} + \|\Pi p - p\|_K.\end{aligned} \tag{3.34}$$

□

4 Numerical Experiments

In this section, we carry out numerical experiments to verify the theoretical orders of convergence of the approximations provided by the HDG method and of the postprocessed velocity given by Theorem 2.6.

As a test problem, we consider the analytical solution of the incompressible Navier-Stokes equations obtained by Kovasznay [19], that is,

$$u_1(x_1, x_2) = 1 - \exp(\lambda x_1)\cos(2\pi x_2),$$
$$u_2(x_1, x_2) = \frac{\lambda}{2\pi}\exp(\lambda x_1)\sin(2\pi x_2),$$
$$p(x_1, x_2) = -\frac{1}{2}\exp(2\lambda x_1) + C,$$

where $\lambda = \frac{1}{2\nu} - \sqrt{\frac{1}{4\nu^2} + 4\pi^2}$. Note that the Reynolds number is inversely proportional to the viscosity ν. The constant C in the definition of the exact solution p is chosen so that the normalization condition (1.1e) is satisfied. Note that the Kovasznay flow is also a solution of the Oseen problem with $\boldsymbol{\beta} = \boldsymbol{u}$. We take the Dirichlet boundary condition (1.1d) for the velocity as the restriction of the exact solution to the domain boundary. Here the computational domain is $\Omega = (0, 2) \times (-0.5, 1.5)$. The viscosity ν is allowed to take two values; 0.1 and 0.001. These choices are made to see the effect of the stabilization parameters τ_n and τ_t on the accuracy of our scheme.

In our experiments, we consider meshes that are refinements of a uniform mesh of 32 triangles. Each refinement is obtained by subdividing each triangle into four 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.

The results, in terms of the L^2-norm, for several choices of τ_n and τ_t and several choices for the polynomial degree k are listed for $\nu = 0.1$ in Tables 1, 2, 3 and 4. For $\nu = 0.001$, Table 5 provides the results for different polynomial degrees for specifically chosen τ_n and τ_t. The numerical results seem to agree with the theoretical analysis summarized in Theorem 2.6. We observe that in the first case when $\nu = 0.1$ (small Reynolds number), the HDG method suffers from poor convergence and accuracy for $\tau_n = \tau_t = h$ since our stabilization parameters $\tau_n = \tau_t$ are significantly smaller than the magnitude of the convective velocity which results eventually in the violation of assumption (2.5). One observation is that if $\tau_t = \frac{1}{h}$, the orders of convergence of the errors in the velocity gradient and the postprocessed velocity are reduced by one. Indeed, we can only obtain an order of k for $\mathrm{E^L}$ unless $\nu < Ch_K|\boldsymbol{\beta}|_{\boldsymbol{W}^{1,\infty}(K)}$, i.e., unless we have locally high Reynolds number. This is because $\|\mathrm{E^L}\|_\Omega$ behaves like $\|\Pi\mathrm{L} - \mathrm{L}\|_\Omega$ which loses an order when $\tau_t = \frac{1}{h}$. See (2.10). Consequently, our postprocessed velocity can only get an order of $k + 1$. Remarkably enough, the pressure seems to converge with the optimal order of $k + 1$, unless $\tau_n = 1/h$. An explanation for this might be the Reynolds number through the constant C_p^1. Whenever we have $\nu, |\boldsymbol{\beta}_0| < Ch_K|\boldsymbol{\beta}|_{\boldsymbol{W}^{1,\infty}(K)}$, even if we lose an order for $\|\Pi\mathrm{L}-\mathrm{L}\|_\Omega$, we can still have an order of $k+1$. Motivated by the accuracy issues related to the violation of the assumption (2.5), as the next test parameters, we pick our τ_n and τ_t as follows:

$$\tau_n = \tau_t = \frac{1}{2\nu} \max_{\boldsymbol{x}\in\mathcal{T}_h} \boldsymbol{\beta}(\boldsymbol{x}) \cdot \boldsymbol{n} + 1. \tag{4.1}$$

Note that this choice guarantees (2.5). Table 4 demonstrates this particular choice of stabilization parameters which yield optimal (order $k + 1$) errors in the velocity, the pressure, the velocity gradient, for $k \geq 1$, and superconvergence (order $k + 2$) of the postprocessed velocity. For $k = 0$, the errors in the approximation of all these variables converge with order one.

Finally, let us emphasize 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.

For our second test case $\nu = 0.001$, a much higher Reynolds number, we again use the special τ_n, τ_t defined in (4.1). In this case, we lose superconvergence, that is, the velocity gradient converges only with order k, thus the postprocessed velocity converges only with order $k + 1$. This may be because of the convection-dominated effect. Indeed, there are constants in Theorem 2.3 and Theorem 2.6 which depend on h, ν, τ_n, τ_t and $\boldsymbol{\beta}$ on each

Table 1 History of convergence of the HDG method for $\tau_n = h$ for $\nu = 0.1$

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
$\tau_t = h$									
1	4	3.32e−0	–	1.91e−0	–	1.19e−9	–	1.11e−0	–
	8	2.86e−0	0.21	9.13e−1	1.07	5.32e−0	1.16	2.61e−1	2.09
	16	3.10e−0	−0.11	5.79e−1	0.66	3.50e−0	0.61	1.47e−1	0.83
	32	3.60e−0	−0.22	4.37e−1	0.41	2.50e−0	0.49	1.10e−1	0.42
	64	6.83e−0	−0.93	3.49e−1	0.32	2.33e−0	0.10	4.86e−2	1.18
2	4	7.98e−1	–	6.75e−1	–	3.65e−0	–	2.56e−1	–
	8	1.35e−0	−0.76	3.36e−1	1.01	3.28e−0	0.16	1.57e−1	0.70
	16	1.11e−0	0.28	1.62e−1	1.05	2.01e−0	0.71	5.74e−2	1.45
	32	3.70e−1	1.59	2.38e−2	2.77	1.76e−1	3.51	1.65e−3	5.12
	64	2.68e−2	3.79	1.06e−3	4.49	1.35e−2	3.70	8.39e−5	4.30
3	4	2.89e−1	–	1.96e−1	–	1.03e−0	–	6.22e−2	–
	8	4.88e−1	−0.76	1.58e−1	0.31	1.09e−0	−0.07	5.42e−2	0.20
	16	2.06e−2	4.57	2.57e−3	5.94	2.90e−2	5.23	4.08e−4	7.05
	32	4.03e−3	2.35	2.62e−4	3.30	2.77e−3	3.39	1.84e−5	4.47
	64	2.94e−2	−2.87	1.05e−3	−2.00	1.09e−2	−1.97	3.27e−5	−0.83
$\tau_t = 1$									
1	4	2.02e−0	–	1.42e−0	–	1.02e−9	–	9.04e−1	–
	8	5.86e−1	1.79	5.16e−1	1.46	3.54e−0	1.53	1.67e−1	2.44
	16	1.74e−1	1.75	1.28e−1	2.01	1.31e−0	1.44	3.03e−2	2.46
	32	4.90e−2	1.83	2.99e−2	2.10	4.39e−1	1.57	5.06e−3	2.58
	64	1.28e−2	1.93	7.21e−3	2.05	1.26e−1	1.80	7.32e−4	2.79
2	4	4.24e−1	–	5.48e−1	–	3.01e−0	–	1.75e−1	–
	8	9.61e−2	2.14	5.92e−2	3.21	7.14e−1	2.07	2.25e−2	2.96
	16	1.13e−2	3.09	6.12e−3	3.27	1.00e−1	2.83	1.49e−3	3.92
	32	1.34e−3	3.07	6.72e−4	3.19	1.33e−2	2.91	9.25e−5	4.01
	64	1.67e−4	3.00	7.91e−5	3.09	1.74e−3	2.94	5.88e−6	3.98
3	4	1.19e−1	–	1.04e−1	–	8.43e−1	–	4.61e−2	–
	8	1.23e−2	3.27	7.59e−3	3.77	9.36e−2	3.17	2.59e−3	4.15
	16	5.43e−4	4.50	3.04e−4	4.64	5.47e−3	4.10	7.28e−5	5.15
	32	3.15e−5	4.11	1.68e−5	4.17	3.51e−4	3.96	2.33e−6	4.96
	64	1.96e−6	4.01	1.02e−6	4.05	2.25e−5	3.96	7.46e−8	4.97
$\tau_t = 1/h$									
1	4	1.46e−0	–	1.26e−0	–	9.39e−0	–	8.08e−1	–
	8	3.20e−1	2.19	5.52e−1	1.19	3.72e−0	1.33	1.74e−1	2.22
	16	7.77e−2	2.04	1.33e−1	2.05	1.59e−0	1.22	3.57e−2	2.28
	32	1.75e−2	2.15	3.23e−2	2.05	5.92e−1	1.43	6.71e−3	2.41
	64	4.10e−3	2.09	8.09e−3	2.00	2.29e−1	1.37	1.49e−3	2.17

Table 1 (*Continued*)

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
2	4	3.24e−1	–	5.23e−1	–	3.02e−0	–	1.49e−1	–
	8	4.48e−2	2.86	4.91e−2	3.41	7.10e−1	2.09	1.89e−2	2.97
	16	4.75e−3	3.24	5.75e−3	3.09	1.18e−1	2.59	1.63e−3	3.54
	32	5.53e−4	3.10	7.58e−4	2.92	2.08e−2	2.50	1.49e−4	3.44
	64	6.75e−5	3.03	1.18e−4	2.69	4.33e−3	2.26	1.58e−5	3.24
3	4	7.90e−2	–	8.23e−2	–	8.33e−1	–	4.26e−2	–
	8	4.18e−3	4.24	4.99e−3	4.04	7.90e−2	3.40	1.99e−3	4.42
	16	2.37e−4	4.14	3.06e−4	4.03	6.38e−3	3.63	7.70e−5	4.69
	32	1.43e−5	4.05	2.11e−5	3.86	5.72e−4	3.48	3.20e−6	4.59
	64	8.84e−7	4.02	1.68e−6	3.65	6.12e−5	3.23	1.60e−7	4.32

element. These constants affect the convergence behavior of the velocity gradient. Hence in order to see superconvergence results for high Reynolds number, we need to use very fine grids dictated by these constants.

5 Conclusion

In this paper, we have analyzed an HDG method to solve the Oseen problem. For the numerical trace, we have used a special type of stabilization tensor S which allows us to control the normal and tangential components of the interelement jumps of the approximate velocity by means of two parameters τ_n and τ_t. We proved that these parameters and $\boldsymbol{\beta}\cdot\boldsymbol{n}$ has an effect on the convergence properties. Indeed, the assumption (2.5) gives us a way to choose the stabilization parameters so that we always have optimal results for any Reynolds number. Our numerical experiments further validate our theoretical results. In summary, for low Reynolds numbers, we have optimal results in all variables. In addition, after a postprocessing we obtain a new velocity which superconverges. For high Reynolds numbers, we still have the optimality but we lose superconvergence for the postprocessed velocity due to convection dominated effects.

5.1 Extensions

One can consider a general stabilization tensor S and the analysis can also be adapted to general nonconforming meshes and variable degree approximations as has been done in [7, 17]. A couple of nontrivial extensions are discussed next.

5.1.1 Divergence-Conforming HDG Methods for Oseen Flow

Note that our error estimates do not contain τ_n in the numerator. This allows us to let τ_n go to ∞ in the solution of (2.1a)–(2.1f) which yields nothing but the following method.

Find $(\mathrm{L}_h, \boldsymbol{u}_h, p_h, \widehat{\boldsymbol{u}}_h, \delta_h) \in \mathrm{G}_h \times \boldsymbol{V}_h \times P_h \times \boldsymbol{M}_h \times M_h^\partial$ such that

Table 2 History of convergence of the HDG method for $\tau_n = 1$ for $\nu = 0.1$

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
$\tau_t = h$									
1	4	1.39e−0	–	1.03e−0	–	7.48e−0	–	6.83e−1	–
	8	5.90e−1	1.24	4.47e−1	1.21	3.44e−0	1.12	1.71e−1	2.00
	16	1.75e−1	1.75	1.10e−1	2.02	1.07e−0	1.69	2.49e−2	2.78
	32	4.60e−2	1.93	2.69e−2	2.03	3.06e−1	1.81	3.55e−3	2.81
	64	1.18e−2	1.96	6.68e−3	2.01	8.20e−2	1.90	4.83e−4	2.88
2	4	4.24e−1	–	4.65e−1	–	2.66e−0	–	1.78e−1	–
	8	6.43e−2	2.72	4.64e−2	3.33	3.65e−1	2.87	1.29e−2	3.78
	16	9.05e−3	2.83	5.89e−3	2.98	5.42e−2	2.75	1.09e−3	3.57
	32	1.18e−3	2.94	7.04e−4	3.07	7.56e−3	2.84	8.30e−5	3.71
	64	1.51e−4	2.96	8.63e−5	3.03	1.02e−3	2.90	5.81e−6	3.84
3	4	9.27e−2	–	7.81e−2	–	6.20e−1	–	3.32e−2	–
	8	9.89e−3	3.23	5.52e−3	3.82	5.90e−2	3.39	1.72e−3	4.28
	16	5.44e−4	4.18	2.73e−4	4.34	3.25e−3	4.18	5.34e−5	5.01
	32	3.18e−5	4.10	1.58e−5	4.11	2.03e−4	4.00	1.82e−6	4.87
	64	1.96e−6	4.02	9.64e−7	4.04	1.29e−5	3.98	5.99e−8	4.93
$\tau_t = 1$									
1	4	1.15e−0	–	1.03e−0	–	7.11e−0	–	6.38e−1	–
	8	3.41e−1	1.75	4.53e−1	1.18	3.01e−0	1.24	1.46e−1	2.13
	16	8.67e−2	1.98	1.16e−1	1.97	9.55e−1	1.66	2.18e−2	2.74
	32	2.20e−2	1.98	2.80e−2	2.05	2.79e−1	1.77	3.14e−3	2.80
	64	5.50e−3	2.00	6.87e−3	2.03	7.53e−2	1.89	4.21e−4	2.90
2	4	3.04e−1	–	4.62e−1	–	2.55e−0	–	1.36e−1	–
	8	4.08e−2	2.90	4.60e−2	3.33	4.30e−1	2.57	1.30e−2	3.39
	16	4.99e−3	3.03	5.81e−3	2.99	6.24e−2	2.79	9.71e−4	3.74
	32	6.13e−4	3.02	7.03e−4	3.05	8.48e−3	2.88	6.73e−5	3.85
	64	7.61e−5	3.01	8.59e−5	3.03	1.11e−3	2.94	4.43e−6	3.92
3	4	7.02e−2	–	6.63e−2	–	6.38e−1	–	3.27e−2	–
	8	4.44e−3	3.98	4.50e−3	3.88	5.39e−2	3.57	1.42e−3	4.52
	16	2.56e−4	4.11	2.67e−4	4.08	3.68e−3	3.87	5.10e−5	4.80
	32	1.56e−5	4.04	1.65e−5	4.02	2.45e−4	3.91	1.73e−6	4.88
	64	9.59e−7	4.02	1.02e−6	4.01	1.58e−5	3.95	5.60e−8	4.95
$\tau_t = 1/h$									
1	4	9.77e−1	–	1.06e−0	–	6.83e−0	–	5.97e−1	–
	8	2.56e−1	1.93	5.20e−1	1.02	3.28e−0	1.06	1.55e−1	1.95
	16	6.15e−2	2.06	1.33e−1	1.96	1.38e−0	1.25	3.02e−2	2.35
	32	1.50e−2	2.04	3.31e−2	2.01	5.59e−1	1.31	5.94e−3	2.35
	64	3.76e−3	2.00	8.42e−3	1.97	2.31e−1	1.27	1.34e−3	2.15

Table 2 (*Continued*)

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
2	4	2.60e−1	–	4.75e−1	–	2.76e−0	–	1.26e−1	–
	8	3.38e−2	2.94	5.04e−2	3.24	5.73e−1	2.27	1.61e−2	2.97
	16	4.00e−3	3.08	6.56e−3	2.94	1.10e−1	2.39	1.56e−3	3.36
	32	4.96e−4	3.01	8.82e−4	2.89	2.13e−2	2.36	1.53e−4	3.35
	64	6.34e−5	2.97	1.34e−4	2.72	4.51e−3	2.24	1.64e−5	3.23
3	4	5.85e−2	–	6.42e−2	–	6.85e−1	–	3.32e−2	–
	8	3.37e−3	4.12	5.04e−3	3.67	6.89e−2	3.31	1.64e−3	4.34
	16	2.04e−4	4.05	3.57e−4	3.82	6.23e−3	3.47	7.04e−5	4.54
	32	1.30e−5	3.97	2.59e−5	3.78	5.97e−4	3.38	3.12e−6	4.50
	64	8.35e−7	3.96	2.02e−6	3.68	6.38e−5	3.23	1.58e−7	4.30

$$(\mathrm{L}_h, \mathrm{G})_{\mathcal{T}_h} + (\boldsymbol{u}_h, \nabla\cdot\mathrm{G})_{\mathcal{T}_h} - \langle \widehat{\boldsymbol{u}}_h, \mathrm{G}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} = 0, \tag{5.1a}$$

$$\begin{aligned}&\big(-\nabla\cdot(\nu\mathrm{L}_h) + \nabla\cdot(\boldsymbol{u}_h\otimes\boldsymbol{\beta}) + \nabla p_h, \cdot\boldsymbol{v}\big)_{\mathcal{T}_h}\\ &\quad + \big\langle \nu\tau_t(\boldsymbol{u}_h - \widehat{\boldsymbol{u}}_h)_t + \nu\delta_h\boldsymbol{n}, \boldsymbol{v}\big\rangle_{\partial\mathcal{T}_h} = (\boldsymbol{f}, \boldsymbol{v})_{\mathcal{T}_h},\end{aligned} \tag{5.1b}$$

$$(\nabla\cdot\boldsymbol{u}_h, q)_{\mathcal{T}_h} = 0, \tag{5.1c}$$

$$\big\langle \delta_h, q^\perp\big\rangle_{\partial\mathcal{T}_h} = 0 \tag{5.1d}$$

$$\langle \widehat{\boldsymbol{u}}_h, \boldsymbol{\mu}\rangle_{\partial\Omega} = \langle \boldsymbol{g}, \boldsymbol{\mu}\rangle_{\partial\Omega}, \tag{5.1e}$$

$$\big\langle \nu\mathrm{L}_h\boldsymbol{n} - p_h\boldsymbol{n} - (\widehat{\boldsymbol{u}}_h\otimes\boldsymbol{\beta})\boldsymbol{n} - \nu\tau_t(\boldsymbol{u}_h - \widehat{\boldsymbol{u}}_h)_t - \nu\delta_h\boldsymbol{n}, \boldsymbol{\mu}\big\rangle_{\partial\mathcal{T}_h\backslash\partial\Omega} = 0, \tag{5.1f}$$

$$\big\langle (\boldsymbol{u}_h - \widehat{\boldsymbol{u}}_h)\cdot\boldsymbol{n}, \eta\big\rangle_{\partial\mathcal{T}_h} = 0 \tag{5.1g}$$

$$(p_h, 1)_\Omega = 0, \tag{5.1h}$$

for all $(\mathrm{G}, \boldsymbol{v}, q, q^\perp, \boldsymbol{\mu}, \eta) \in \mathrm{G}_h \times \boldsymbol{V}_h \times P_h^{k-1} \times P_h^\perp \times \boldsymbol{M}_h \times M_h^\partial$. For the above we define the additional approximation space

$$M_h^\partial := \big\{\mu \in L^2(\partial\mathcal{T}_h) : \mu|_{\partial K} \in R_k(\partial K),\ \forall K \in \mathcal{T}_h\big\}$$

where $R_k(\partial K)$ is defined in Lemma 3.9. The test spaces for the method employs the following orthogonal decomposition of the pressure space P_h:

$$P_h = P_h^{k-1} \oplus P_h^\perp,$$

where

$$\begin{aligned}P_h^{k-1} &:= \big\{q \in L^2(\mathcal{T}_h) : q|_K \in P_{k-1}(K),\ \forall K \in \mathcal{T}_h\big\},\\ P_h^\perp &:= \big\{q \in L^2(\mathcal{T}_h) : q|_K \in P_k(K)^\perp,\ \forall K \in \mathcal{T}_h\big\}.\end{aligned}$$

Also, $(\boldsymbol{u}_h - \widehat{\boldsymbol{u}}_h)_t$ denotes the tangential component of $\boldsymbol{u}_h - \widehat{\boldsymbol{u}}_h$. This problem has been analyzed for the Stokes problem in [9].

Table 3 History of convergence of the HDG method for $\tau_n = 1/h$ for $\nu = 0.1$

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
$\tau_t = h$									
1	4	1.03e−0	–	1.09e−0	–	7.21e−0	–	6.56e−1	–
	8	3.30e−1	1.64	4.09e−1	1.41	3.41e−0	1.08	1.78e−1	1.88
	16	7.29e−2	2.18	1.08e−1	1.92	1.11e−0	1.62	2.63e−2	2.76
	32	1.66e−2	2.14	2.70e−2	2.00	3.23e−1	1.78	3.79e−3	2.79
	64	3.97e−3	2.06	6.72e−3	2.01	8.54e−2	1.92	5.01e−4	2.92
2	4	3.16e−1	–	4.10e−1	–	2.41e−0	–	1.46e−1	–
	8	3.65e−2	3.11	4.51e−2	3.18	3.41e−1	2.82	1.21e−2	3.58
	16	4.52e−3	3.02	5.87e−3	2.94	5.73e−2	2.57	1.20e−3	3.33
	32	5.55e−4	3.02	7.03e−4	3.06	8.16e−3	2.81	9.17e−5	3.71
	64	6.89e−5	3.01	8.63e−5	3.03	1.07e−3	2.93	6.21e−6	3.88
3	4	6.4e−2	–	5.58e−2	–	5.57e−1	–	2.85e−2	–
	8	3.88e−3	4.04	4.11e−3	3.76	4.25e−2	3.71	1.18e−3	4.60
	16	2.36e−4	4.04	2.52e−4	4.03	2.87e−3	3.89	4.83e−5	4.60
	32	1.46e−5	4.01	1.56e−5	4.01	1.97e−4	3.87	1.80e−6	4.75
	64	9.10e−7	4.01	9.65e−7	4.02	1.28e−5	3.94	6.04e−8	4.89
$\tau_t = 1$									
1	4	9.14e−1	–	1.11e−0	–	6.79e−0	–	6.23e−1	–
	8	2.66e−1	1.78	4.28e−1	1.37	3.04e−0	1.16	1.56e−1	2.00
	16	6.13e−2	2.12	1.20e−1	1.84	9.83e−1	1.63	2.19e−2	2.83
	32	1.52e−2	2.02	3.15e−2	1.93	2.95e−1	1.74	3.27e−3	2.75
	64	3.81e−3	1.99	8.13e−3	1.95	8.32e−2	1.83	4.69e−4	2.80
2	4	2.61e−1	–	4.21e−1	–	2.44e−0	–	1.23e−1	–
	8	3.15e−2	3.05	5.08e−2	3.05	3.76e−1	2.70	1.24e−2	3.31
	16	4.02e−3	2.97	6.95e−3	2.87	6.02e−2	2.64	1.10e−3	3.50
	32	5.16e−4	2.96	8.90e−4	2.97	8.99e−3	2.74	8.69e−5	3.66
	64	6.61e−5	2.96	1.13e−4	2.97	1.26e−3	2.83	6.31e−6	3.78
3	4	5.62e−2	–	5.69e−2	–	5.87e−1	–	2.82e−2	–
	8	3.35e−3	4.07	4.73e−3	3.59	4.67e−2	3.65	1.18e−3	4.58
	16	2.11e−4	3.99	3.25e−4	3.87	3.38e−3	3.79	4.64e−5	4.66
	32	1.36e−5	3.95	2.15e−5	3.92	2.42e−4	3.81	1.78e−6	4.71
	64	8.74e−7	3.96	1.38e−6	3.96	1.66e−5	3.87	6.35e−8	4.81
$\tau_t = 1/h$									
1	4	8.24e−1	–	1.18e−0	–	6.46e−0	–	5.86e−1	–
	8	2.33e−1	1.82	5.83e−1	1.02	3.49e−0	0.89	1.65e−1	1.83
	16	5.32e−2	2.13	2.34e−1	1.32	1.66e−0	1.07	3.32e−2	2.31
	32	1.32e−2	2.01	1.07e−1	1.12	8.13e−1	1.03	8.10e−3	2.04
	64	3.30e−3	2.00	5.24e−2	1.04	4.03e−1	1.01	2.04e−3	1.99

Table 3 (*Continued*)

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
2	4	2.36e−1	–	4.47e−1	–	2.74e−0	–	1.22e−1	–
	8	3.06e−2	2.95	7.21e−2	2.63	5.76e−1	2.25	1.84e−2	2.73
	16	3.89e−3	2.98	1.47e−2	2.30	1.40e−1	2.04	2.41e−3	2.93
	32	4.85e−4	3.00	3.26e−3	2.17	3.52e−2	1.99	3.02e−4	3.00
	64	6.03e−5	3.01	7.71e−4	2.08	8.86e−3	1.99	3.75e−5	3.01
3	4	5.15e−2	–	6.23e−2	–	6.54e−1	–	2.94e−2	–
	8	3.10e−3	4.06	7.25e−3	3.10	7.08e−2	3.21	1.62e−3	4.18
	16	1.91e−4	4.02	8.40e−4	3.11	8.33e−3	3.09	9.29e−5	4.12
	32	1.19e−5	4.00	9.96e−5	3.08	1.03e−3	3.01	5.74e−6	4.02
	64	7.49e−7	4.00	1.20e−5	3.05	1.29e−4	3.00	3.61e−7	3.99

Table 4 History of convergence of the HDG method for $\tau_n=\tau_t=\frac{1}{2\nu}\max_{\boldsymbol{x}\in\mathcal{T}_h}\boldsymbol{\beta}(\boldsymbol{x})\cdot\boldsymbol{n}+1$ for $\nu=0.1$

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
1	4	1.07e−0	–	6.37e−0	–	8.88e−0	–	9.71e−1	–
	8	2.76e−1	1.95	1.41e−0	2.17	4.87e−0	0.87	2.4e−1	2.02
	16	5.93e−2	2.22	3.23e−1	2.13	1.98e−0	1.3	4.48e−2	2.42
	32	1.33e−2	2.16	7.75e−2	2.06	7.23e−1	1.45	7.89e−3	2.5
	64	3.08e−3	2.11	1.89e−2	2.04	2.39e−1	1.59	1.3e−3	2.6
2	4	3.64e−1	–	1.2e−0	–	4.79e−0	–	3.35e−1	–
	8	4.47e−2	3.03	1.87e−1	2.67	9.99e−1	2.26	3.97e−2	3.07
	16	4.41e−3	3.34	2.29e−2	3.03	1.75e−1	2.51	3.29e−3	3.59
	32	4.61e−4	3.26	2.79e−3	3.04	2.91e−2	2.59	2.49e−4	3.72
	64	5.27e−5	3.13	3.46e−4	3.01	4.46e−3	2.71	1.8e−5	3.79
3	4	6.69e−2	–	2.37e−1	–	1.32e−0	–	5.68e−2	–
	8	3.93e−3	4.09	2.17e−2	3.45	1.28e−1	3.36	3.1e−3	4.2
	16	2.04e−4	4.27	1.33e−3	4.03	1.09e−2	3.56	1.23e−4	4.65
	32	1.16e−5	4.14	8.17e−5	4.02	8.62e−4	3.66	4.79e−6	4.69
	64	6.86e−7	4.07	5.09e−6	4	6.3e−5	3.78	1.75e−7	4.78

5.1.2 HDG Method for Incompressible Navier-Stokes Equations

As suggested in the introduction, the main extension of this paper is to the incompressible Navier-Stokes equations where we have a nonlinear velocity rather than the known convective velocity $\boldsymbol{\beta}$.

$$3\mathrm{L}-\nabla\boldsymbol{u}=0 \quad \text{in } \Omega,$$
$$-\nu\nabla\cdot\mathrm{L}+(\boldsymbol{u}\cdot\nabla)\boldsymbol{u}+\nabla p=\boldsymbol{f} \quad \text{in } \Omega,$$

Table 5 History of convergence of the HDG method for $\tau_n = \tau_t = \frac{1}{2\nu} \max_{\boldsymbol{x}\in\mathcal{T}_h} \boldsymbol{\beta}(\boldsymbol{x}) \cdot \boldsymbol{n} + 1$ for $\nu = 0.001$

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
1	4	2.09e−1	–	3.23e−9	–	4.19e−0	–	2.59e−1	–
	8	2.81e−1	−0.43	9.4e−0	1.78	4.39e−0	−0.07	2.81e−1	−0.11
	16	5.02e−2	2.48	5.46e−1	4.11	2.03e−0	1.11	5.02e−2	2.48
	32	9.87e−3	2.35	4.66e−2	3.55	9.95e−1	1.03	9.86e−3	2.35
	64	2.33e−3	2.09	7.53e−3	2.63	4.93e−1	1.01	2.32e−3	2.09
2	4	2.11e−1	–	1.57e−9	–	4.27e−0	–	2.09e−1	–
	8	6.35e−2	1.73	6.06e−1	4.7	1.3e−0	1.72	6.35e−2	1.72
	16	6.67e−3	3.25	4.65e−2	3.71	2.75e−1	2.24	6.67e−3	3.25
	32	7.59e−4	3.13	4.14e−3	3.49	5.82e−2	2.24	7.59e−4	3.13
	64	7.99e−5	3.25	3.74e−4	3.47	1.32e−2	2.14	7.98e−5	3.25
3	4	1.9e−2	–	4.03e−1	–	3.7e−1	–	1.89e−2	–
	8	3.48e−3	2.45	4.09e−2	3.3	1.31e−1	1.49	3.48e−3	2.45
	16	1.68e−4	4.37	7.11e−4	5.85	1.38e−2	3.25	1.68e−4	4.37
	32	1.07e−5	3.97	1.66e−5	5.42	1.67e−3	3.05	1.07e−5	3.97
	64	6.73e−7	3.99	6.76e−7	4.62	2.06e−4	3.02	6.7e−7	3.99

$$\begin{aligned} \nabla \cdot \boldsymbol{u} = 0 \quad & \text{in } \Omega, \\ \boldsymbol{u} = \boldsymbol{g} \quad & \text{on } \partial\Omega, \\ \int_\Omega p = 0, & \end{aligned}$$

where $\boldsymbol{u}$ is the velocity, p is the pressure, ν is the kinematic viscosity and $\boldsymbol{f} \in L^2(\Omega)^n$ is the external body force. As before we assume

$$\int_{\partial\Omega} \boldsymbol{g} \cdot \boldsymbol{n} = 0.$$

The above set of equations can be solved iteratively by a Picard iteration. This means that we solve an Oseen problem where the convective velocity comes as an input from the last iteration. The application of this approach to the devising of LDG methods for the Navier-Stokes was carried out in [12]. Its application to HDG methods constitutes the subject of ongoing research.

Appendix: Approximation Properties of the Auxiliary Projection $\boldsymbol{\Pi}_h^*$

The proofs in this appendix are quite similar to the proofs for the approximation properties of the projection Π_h given in Sect. 3.3. We only need to recall that $\widetilde{\mathrm{S}}_{\boldsymbol{\beta}} := \mathrm{S} - (\boldsymbol{\beta} \cdot \boldsymbol{n})\mathrm{I}$ and that $\widetilde{\Lambda}_K^{\max}$ is defined to be the maximum eigenvalue of $\widetilde{\mathrm{S}}_{\boldsymbol{\beta}}$ over all faces of K.

A.1 Approximation Properties of $\Pi^*\boldsymbol{\phi}$

Proposition A.1 (Characterization of $\Pi^*\boldsymbol{\phi}$)

$$\left(\Pi^*\boldsymbol{\phi} - \boldsymbol{\phi}, \boldsymbol{v}\right)_K = 0, \quad \forall \boldsymbol{v} \in \boldsymbol{P}_{k-1}(K), \tag{A.1a}$$

$$\begin{aligned}&\left\langle \widetilde{\mathrm{S}}_{\boldsymbol{\beta}}\left(\Pi^*\boldsymbol{\phi} - \boldsymbol{\phi}\right), \boldsymbol{\mu}\right\rangle_{\partial K} + \left(\left(\boldsymbol{\Pi}^*\boldsymbol{\phi} - \boldsymbol{\phi}\right) \otimes \boldsymbol{\beta}, \nabla \boldsymbol{\mu}\right)_K = -\left(\nabla \cdot (\nu \Phi + \phi \mathrm{I}), \boldsymbol{\mu}\right)_K, \\ &\quad \forall \boldsymbol{\mu} \in \boldsymbol{P}_k(K)^{\perp}.\end{aligned} \tag{A.1b}$$

Proof First equation follows from the definition of Π_h^*. The second follows from (2.13d). Indeed if we take $\boldsymbol{\mu} \in P_k(K)^{\perp}$ and using integration by parts,

$$\left\langle \mathrm{S}\left(\boldsymbol{\Pi}^*\boldsymbol{\phi} - \boldsymbol{\phi}\right), \boldsymbol{\mu}\right\rangle_{\partial K} - \left(\left(\boldsymbol{\Pi}^*\boldsymbol{\phi} - \boldsymbol{\phi}\right) \otimes \boldsymbol{\beta}\right)\boldsymbol{n}, \boldsymbol{\mu}\rangle_{\partial K} = \left\langle \left(\nu \Pi^*\Phi - \Phi\right)\boldsymbol{n} + \left(\Pi^*\phi - \phi\right)\boldsymbol{n}, \boldsymbol{\mu}\right\rangle_{\partial K}.$$

Then, using integration by parts on the right hand side,

$$\begin{aligned}\left\langle \widetilde{\mathrm{S}}_{\boldsymbol{\beta}}\left(\boldsymbol{\Pi}^*\boldsymbol{\phi} - \boldsymbol{\phi}\right), \boldsymbol{\mu}\right\rangle_{\partial K} &= \left(\nu \nabla \cdot \left(\Pi^*\Phi - \Phi\right), \boldsymbol{\mu}\right)_K + \left(\nu\left(\Pi^*\Phi - \Phi\right), \nabla \boldsymbol{\mu}\right)_K \\ &\quad + \left(\nabla\left(\Pi^*\phi - \phi\right), \boldsymbol{\mu}\right)_K + \left(\Pi^*\phi - \phi, \nabla \cdot \boldsymbol{\mu}\right)_K \\ &= -\left(\nabla \cdot (\nu \Phi + \phi \mathrm{I}), \boldsymbol{\mu}\right)_K - \left(\left(\boldsymbol{\Pi}^*\boldsymbol{\phi} - \boldsymbol{\phi}\right) \otimes \boldsymbol{\beta}, \nabla \boldsymbol{\mu}\right)_K,\end{aligned}$$

where the last equality holds by (2.13a) and (2.13c) and as $\boldsymbol{w} \in \boldsymbol{P}_k^{\perp}$. □

Let $\boldsymbol{\delta}^{\boldsymbol{\phi}} = \boldsymbol{\Pi}^*\boldsymbol{\phi} - \boldsymbol{\phi}_k$ where $\boldsymbol{\phi}_k$ is the L^2-projection onto $\boldsymbol{P}_k(K)$. Then, from (A.1a), $\boldsymbol{\delta}^{\boldsymbol{\phi}} \in \boldsymbol{P}_k(K)^{\perp}$. Then, from the second characterization,

$$\begin{aligned}\left\langle (\widetilde{\mathrm{S}}_{\boldsymbol{\beta}}\boldsymbol{\delta}^{\boldsymbol{\phi}}, \boldsymbol{\mu}\right\rangle_{\partial K} + \left(\boldsymbol{\delta}^{\boldsymbol{\phi}} \otimes \boldsymbol{\delta\beta}, \nabla \boldsymbol{\mu}\right)_K &= \underbrace{\left(\widetilde{\mathrm{S}}_{\boldsymbol{\beta}}(\boldsymbol{\phi} - \boldsymbol{\phi}_k), \boldsymbol{\mu}\right)_K + \left((\boldsymbol{\phi} - \boldsymbol{\phi}_k) \otimes \boldsymbol{\beta}, \nabla \boldsymbol{w}\right)_K}_{b_{\boldsymbol{\phi}}(\boldsymbol{\mu})} \\ &\quad \underbrace{-\left(\nabla \cdot (\nu \Phi + \phi \mathrm{I}), \boldsymbol{\mu}\right)_K}_{b_{\Phi}(\boldsymbol{\mu})}.\end{aligned}$$

Now let $\boldsymbol{\mu} = \boldsymbol{\delta}^{\boldsymbol{\phi}}$. Then, as $(\boldsymbol{\delta}^{\boldsymbol{\phi}} \otimes \boldsymbol{\delta\beta}, \nabla \boldsymbol{\delta}^{\boldsymbol{\phi}})_K = \frac{1}{2}\langle(\boldsymbol{\beta} \cdot \boldsymbol{n})\boldsymbol{\delta}^{\boldsymbol{\phi}}, \boldsymbol{\delta}^{\boldsymbol{\phi}}\rangle_{\partial K}$, we have,

$$\left\langle \mathrm{S}_{\boldsymbol{\beta}}\boldsymbol{\delta}^{\boldsymbol{\phi}}, \boldsymbol{\delta}^{\boldsymbol{\phi}}\right\rangle_{\partial K} = b_{\boldsymbol{\phi}}\left(\boldsymbol{\delta}^{\boldsymbol{\phi}}\right) + b_{\Phi}\left(\boldsymbol{\delta}^{\boldsymbol{\phi}}\right).$$

From Lemma 3.12,

$$\left\|\boldsymbol{\delta}^{\boldsymbol{\phi}}\right\|_{0,K} \le C h_K^{1/2}\left\|\boldsymbol{\delta}^{\boldsymbol{\phi}}\right\|_{0,F}.$$

Then, as $\mathrm{S}_{\boldsymbol{\beta}}$ is positive definite, for any face F of K,

$$\left\langle \mathrm{S}_{\boldsymbol{\beta}}\boldsymbol{\delta}^{\boldsymbol{\phi}}, \boldsymbol{\delta}^{\boldsymbol{\phi}}\right\rangle_{\partial K} \ge \sum_F \sum_{i=1}^{d} \left\langle \mathrm{S}_{\boldsymbol{\beta}}\boldsymbol{\delta}^{\boldsymbol{\phi}}, \boldsymbol{\delta}^{\boldsymbol{\phi}}\right\rangle_{F_i} \ge C h_K^{-1} \Gamma_K^{\min}\left\|\boldsymbol{\delta}^{\boldsymbol{\phi}}\right\|_K^2.$$

Therefore,

$$\left\|\boldsymbol{\delta}^{\boldsymbol{\phi}}\right\|_K \le C\frac{h_K}{\Gamma_K^{\min}}\left(\|b_{\boldsymbol{\phi}}\| + \|b_{\Phi}\|\right).$$

$\|b_{\boldsymbol{\phi}}\|$ is bounded exactly as in Sect. 3.3, with one difference in the outcome. Rather than $\Lambda_K^{\max}$, we have $\widetilde{\Lambda}_K^{\max}$ and $\|b_\Phi\|$ is also bounded the same way except that we have $|\nabla\cdot(\nu\Phi+\phi\mathrm{I})|_{k_\sigma}$ rather than $|\nabla\cdot(\nu\Phi-\phi\mathrm{I})|_{k_\sigma}$.

A.2 Approximation Properties of $\nu\Pi^*\Phi+\Pi^*\phi\mathrm{I}$

As in Sect. 3.3, we need two additional projections. We introduce a projection $\widetilde{\mathrm{P}}^1$ similar to P^1 as defined in Sect. 3.3 and we define $\widetilde{\mathrm{P}}^2$ to suit to the form of the projection $\Pi_h^\star$. Let $\widetilde{\mathrm{P}}^1\Phi\in\mathrm{P}_k(K)$ be such that

$$\big(\widetilde{\mathrm{P}}^1\Phi,\mathrm{G}\big)_K=(\Phi,\mathrm{G})_K,\quad \forall\mathrm{G}\in\mathrm{P}_{k-1}(K),$$
$$\big\langle\widetilde{\mathrm{P}}^1\Phi\boldsymbol{n}_F,\boldsymbol{\mu}\big\rangle_F=\langle\Phi\boldsymbol{n}_F,\boldsymbol{\mu}\rangle_F,\quad \forall\boldsymbol{\mu}\in\boldsymbol{P}_k(F),$$

for all faces F of the simplex K except for an arbitrary one and let $\widetilde{\mathrm{P}}^2\Phi\in\mathrm{P}_k(K)$ be such that

$$\big(\widetilde{\mathrm{P}}^2\Phi,\mathrm{G}\big)_K=(\Phi,\mathrm{G})_K-\big(\big(\boldsymbol{\Pi}^*\boldsymbol{\phi}-\boldsymbol{\phi}\big)\otimes\boldsymbol{\delta\beta},\mathrm{G}\big),\quad \forall\mathrm{G}\in\mathrm{P}_{k-1}(K),$$
$$\big\langle\widetilde{\mathrm{P}}^2\Phi\boldsymbol{n}_F,\boldsymbol{\mu}\big\rangle_F=\langle\Phi\boldsymbol{n}_F,\boldsymbol{\mu}\rangle_F,\quad \forall\boldsymbol{\mu}\in\boldsymbol{P}_k(F),$$

for all faces F of the simplex K except for an arbitrary one.

Proposition A.2 (Characterization of $\nu\Pi^*\Phi+\Pi^*\phi\mathrm{I}$)

$$\big(\nu\big(\Pi^*\Phi-\Phi\big)+\big(\Pi^*\phi-\phi\big)\mathrm{I},\mathrm{G}\big)_K$$
$$=-\big(\big(\boldsymbol{\Pi}^*\boldsymbol{\phi}-\boldsymbol{\phi}\big)\otimes\boldsymbol{\delta\beta},\mathrm{G}\big)_K,\quad \forall\mathrm{G}\in\mathrm{P}_{k-1}(K),\tag{A.2a}$$
$$\big\langle\nu\big(\Pi^*\Phi-\Phi\big)\boldsymbol{n}_F+\big(\Pi^*\phi-\phi\big)\boldsymbol{n}_F,\boldsymbol{\mu}\big\rangle_F$$
$$=\big\langle\widetilde{\mathrm{S}}_{\boldsymbol{\beta}}\big(\boldsymbol{\Pi}^*\boldsymbol{\phi}-\boldsymbol{\phi}\big),\boldsymbol{\mu}\big\rangle_F\quad \forall\boldsymbol{\mu}\in\boldsymbol{P}_k(F),\tag{A.2b}$$

for all faces F of K.

Proof The first equation follows directly from (2.13a)–(2.13c). For the second pick an arbitrary face F of K and let $\boldsymbol{w}\in\boldsymbol{P}_k(F)$. Then, there exists $\boldsymbol{\mu}\in\boldsymbol{P}_k(K)^\perp$ such that $\boldsymbol{\mu}=\boldsymbol{w}$ on F. Therefore, in a similar fashion to the proof of the result for $(\Pi\mathrm{L}-\mathrm{L})-(\Pi p-p)\mathrm{I}$, splitting the integral over ∂K to F and $\partial K\backslash F$ and using (2.13d) with this $\boldsymbol{\mu}$,

$$\begin{aligned}&\big\langle\nu\big(\Pi^*\Phi-\Phi\big)\boldsymbol{n}_F+\big(\Pi^*\phi-\phi\big)\boldsymbol{n}_F,\boldsymbol{\mu}\big\rangle_F\\&\quad=-\big\langle\big(\Pi^*\Phi-\Phi\big)\otimes\boldsymbol{\beta}\big)\boldsymbol{n}_F,\boldsymbol{\mu}\big\rangle_F+\big\langle\mathrm{S}\big(\boldsymbol{\Pi}^*\boldsymbol{\phi}-\boldsymbol{\phi}\big),\boldsymbol{\mu}\big\rangle_F+T\\&\quad=\big\langle\widetilde{\mathrm{S}}_{\boldsymbol{\beta}}\big(\boldsymbol{\Pi}^*\boldsymbol{\phi}-\boldsymbol{\phi}\big),\boldsymbol{\mu}\big\rangle_F+T,\end{aligned}$$

where

$$\begin{aligned}T:=&\big\langle\nu\big(\Pi^*\Phi-\Phi\big)\boldsymbol{n}_F+\big(\Pi^*\phi-\phi\big)\boldsymbol{n}_F,\boldsymbol{\mu}\big\rangle_{\partial K}+\big\langle\big(\Pi^*\Phi-\Phi\big)\otimes\boldsymbol{\beta}\big)\boldsymbol{n}_F,\boldsymbol{\mu}\big\rangle_{\partial K}\\&-\big\langle\mathrm{S}\big(\boldsymbol{\Pi}^*\boldsymbol{\phi}-\boldsymbol{\phi}\big),\boldsymbol{\mu}\big\rangle_{\partial K}\\=&\big\langle\nu\big(\Pi^*\Phi-\Phi\big)\boldsymbol{n}_F+\big(\Pi^*\phi-\phi\big)\boldsymbol{n}_F,\boldsymbol{\mu}\big\rangle_{\partial K}-\big\langle\widetilde{\mathrm{S}}_{\boldsymbol{\beta}}\big(\boldsymbol{\Pi}^*\boldsymbol{\phi}-\boldsymbol{\phi}\big),\boldsymbol{\mu}\big\rangle_{\partial K}.\end{aligned}$$

But by (A.1b) and integration by parts,

$$T = \big(\nabla \cdot \nu\big(\Pi^*\Phi - \Phi\big), \boldsymbol{\mu}\big)_K + \big(\nu\big(\Pi^*\Phi - \Phi\big), \nabla\boldsymbol{\mu}\big)_K + \big(\nabla\big(\Pi^*\phi - \phi\big), \boldsymbol{\mu}\big)_K$$
$$+ \big(\Pi^*\phi - \phi, \nabla \cdot \boldsymbol{\mu}\big)_K + \big(\big(\boldsymbol{\Pi}^*\boldsymbol{\phi} - \boldsymbol{\phi}\big) \otimes \boldsymbol{\beta}, \nabla\boldsymbol{\mu}\big)_K + \big(\nabla \cdot (\nu\Phi + \phi\mathrm{I}), \boldsymbol{\mu}\big)_K.$$

The first, third and the last terms vanish by cancellation and from the fact that $\boldsymbol{w} \in \boldsymbol{P}_k(K)^\perp$. The second, fourth and fifth terms vanish by (A.2a). □

The proof of the estimate for $\nu(\Pi^*\Phi - \Phi)$ is very similar to the one for $\nu(\Pi\mathrm{L} - \mathrm{L})$. In short, from the representation of $\nu(\Pi^*\Phi - \Phi)$ using (3.24), it boils down to bounding $\nu(\Pi^*\Phi - \Phi)\boldsymbol{n}_F \cdot \boldsymbol{t}$ for all faces F of K and for all $\boldsymbol{t} \in \mathcal{B}_F$. Using the projections we write $(\Pi^*\Phi - \Phi)\boldsymbol{n}_F = (\Pi^*\Phi - \widetilde{\mathrm{P}}^2\Phi)\boldsymbol{n}_F + (\widetilde{\mathrm{P}}^2\Phi - \widetilde{\mathrm{P}}^1\Phi)\boldsymbol{n}_F + (\widetilde{\mathrm{P}}^1\Phi - \Phi)\boldsymbol{n}_F$. It is easy to show from the properties of $\widetilde{\mathrm{P}}^2_F$, P_F and (A.2b) that

$$\nu\big(\Pi^*\Phi - \widetilde{\mathrm{P}}^2\Phi\big)\boldsymbol{n}_F = -\big(\Pi^*\phi - \widetilde{\mathrm{P}}^1\phi\big)\boldsymbol{n}_F + \mathrm{S}|_F\big(\boldsymbol{\Pi}^*\boldsymbol{\phi} - \widetilde{P}^1\boldsymbol{\phi}\big).$$

Therefore,

$$\big\|\nu\big(\Pi^*\Phi - \Phi\big)\boldsymbol{n}_F \cdot \boldsymbol{t}\big\|_{0,K} \le \big\|\nu\tau_t\big(\boldsymbol{\Pi}^*\boldsymbol{\phi} - \widetilde{\mathrm{P}}^1\boldsymbol{\phi}\big)\big\|_{0,K} + \big\|\big(\widetilde{\mathrm{P}}^2\Phi - \widetilde{\mathrm{P}}^1\Phi\big)\boldsymbol{n}_F\big\|_{0,K}$$
$$+ \big\|\big(\widetilde{\mathrm{P}}^1\Phi - \Phi\big)\boldsymbol{n}_F)\big\|_{0,K}.$$

The terms on the right hand side are bounded exactly the same way with the only difference being the equations defining $\delta^\Phi := (\widetilde{\mathrm{P}}^2\Phi - \widetilde{\mathrm{P}}^1\Phi)\boldsymbol{n}_F$. Now they are given by

$$\big(\delta^\Phi, \mathrm{G}\big)_K = -\big(\big(\Pi^*\boldsymbol{\phi} - \boldsymbol{\phi}\big) \otimes \boldsymbol{\delta\beta}, \mathrm{G}\big)_K, \quad \forall \mathrm{G} \in \mathrm{P}_{k-1}(K),$$
$$\big\langle\delta^\Phi, \boldsymbol{\mu}\big\rangle_F = 0, \quad \forall \boldsymbol{\mu} \in \boldsymbol{P}_k(F),$$

for all faces F of K except for an arbitrarily chosen one.

References

1. Adams, R.A.: Sobolev Spaces. Pure and Applied Mathematics, vol. 65. Academic Press, San Diego (1975) [A subsidiary of Harcourt Brace Jovanovich, Publishers]
2. Arnold, D.N., Brezzi, F., Cockburn, B., Marini, L.D.: Unified analysis of discontinuous Galerkin methods for elliptic problems. SIAM J. Numer. Anal. **39**(5), 1749–1779 (2001/02) (electronic)
3. Brezzi, F., Fortin, M.: Mixed and Hybrid Finite Element Methods. Springer Series in Computational Mathematics, vol. 15. Springer, New York (1991)
4. Brezzi, F., Douglas, J. Jr., Marini, L.D.: Two families of mixed finite elements for second order elliptic problems. Numer. Math. **47**(2), 217–235 (1985)
5. Burman, E., Fernández, M.A., Hansbo, P.: Continuous interior penalty finite element method for Oseen's equations. SIAM J. Numer. Anal. **44**(3), 1248–1274 (2006). (electronic)
6. Castillo, P.: Performance of discontinuous Galerkin methods for elliptic PDEs. SIAM J. Sci. Comput. **24**(2), 524–547 (2002)
7. Chen, Y., Cockburn, B.: Analysis of variable-degree HDG methods for convection-diffusion equations. Part I: general nonconforming meshes. IMA J. Numer. Anal. (2012). doi:10.1093/imanum/drr058
8. Cockburn, B., Dong, B.: An analysis of the minimal dissipation local discontinuous Galerkin method for convection-diffusion problems. J. Sci. Comput. **32**(2), 233–262 (2007)
9. Cockburn, B., Sayas, F.J.: Divergence-conforming {HDG} methods for Stokes flow (2012, submitted)
10. Cockburn, B., Shi, K.: Conditions for superconvergence of HDG methods for Stokes flow. Math. Comput. (2012, accepted)

11. Cockburn, B., Kanschat, G., Schötzau, D.: The local discontinuous Galerkin method for the Oseen equations. Math. Comput. **73**(246), 569–593 (2004). (electronic)
12. Cockburn, B., Kanschat, G., Schotzau, D.: A locally conservative LDG method for the incompressible Navier-Stokes equations. Math. Comput. **74**(251), 1067–1095 (2005). (electronic)
13. Cockburn, B., Dong, B., Guzmán, J.: A superconvergent LDG-hybridizable Galerkin method for second-order elliptic problems. Math. Comput. **77**(264), 1887–1916 (2008)
14. Cockburn, B., Gopalakrishnan, J., Lazarov, R.: Unified hybridization of discontinuous Galerkin, mixed, and continuous Galerkin methods for second order elliptic problems. SIAM J. Numer. Anal. **47**(2), 1319–1365 (2009)
15. Cockburn, B., Gopalakrishnan, J., Guzmán, J.: A new elasticity element made for enforcing weak stress symmetry. Math. Comput. **79**(271), 1331–1349 (2010)
16. Cockburn, B., Gopalakrishnan, J., Sayas, F.-J.: A projection based error analysis of HDG methods. Math. Comput. **79**, 1351–1367 (2010)
17. Cockburn, B., Gopalakrishnan, J., Nguyen, N.C., Peraire, J., Sayas, F.-J.: Analysis of an HDG methods for Stokes flow. Math. Comput. **80**, 723–760 (2011)
18. Girault, V., Raviart, P.-A.: Finite Element Methods for Navier-Stokes Equations. Springer Series in Computational Mathematics, vol. 5. Springer, Berlin (1986). Theory and algorithms
19. Kovasznay, L.I.G.: Laminar flow behind two-dimensional grid. Proc. Camb. Philos. Soc. **44**, 58–62 (1948)
20. Nédélec, J.-C.: Mixed finite elements in $\mathbf{R}^3$. Numer. Math. **35**(3), 315–341 (1980)
21. Raviart, P.-A., Thomas, J.M.: A mixed finite element method for 2nd order elliptic problems. In: Mathematical Aspects of Finite Element Methods (Proc. Conf., Consiglio Naz. delle Ricerche (C.N.R.), Rome, 1975). Lecture Notes in Math., vol. 606, pp. 292–315. Springer, Berlin (1977)