MATHEMATICS OF COMPUTATION
Volume 86, Number 306, July 2017, Pages 1643–1670
http://dx.doi.org/10.1090/mcom/3195
Article electronically published on November 28, 2016

ANALYSIS OF A HYBRIDIZABLE DISCONTINUOUS GALERKIN METHOD FOR THE STEADY-STATE INCOMPRESSIBLE NAVIER-STOKES EQUATIONS

AYCIL CESMELIOGLU, BERNARDO COCKBURN, AND WEIFENG QIU

Abstract. We present the first a priori error analysis of the hybridizable discontinuous Galerkin method for the approximation of the Navier-Stokes equations proposed in J. Comput. Phys. vol. 230 (2011), pp. 1147-1170. The method is defined on conforming meshes made of simplexes and provides piecewise polynomial approximations of fixed degree k to each of the components of the velocity gradient, velocity and pressure. For the stationary case, and under the usual smallness condition for the source term, we prove that the method is well defined and that the global L^2-norm of the error in each of the above-mentioned variables converges with the optimal order of $k+1$ for $k \geq 0$. We also prove a superconvergence property of the velocity which allows us to obtain an elementwise postprocessed approximate velocity, $H(\mathrm{div})$-conforming and divergence-free, which converges with order $k+2$ for $k \geq 1$. In addition, we show that these results only depend on the inverse of the stabilization parameter of the jump of the normal component of the velocity. Thus, if we superpenalize those jumps, these converegence results do hold by assuming that the pressure lies in $H^1(\Omega)$ only. Moreover, by letting such stabilization parameters go to infinity, we obtain new $H(\mathrm{div})$-conforming methods with the above-mentioned convergence properties.

1. Introduction

In this paper, we provide the first a priori error analysis of the hybridizable discontinuous Galerkin (HDG) method proposed in [28] for the stationary incompressible Navier-Stokes equations, namely,

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

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

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

$$\boldsymbol{u} = \boldsymbol{0} \quad \text{on } \partial\Omega, \tag{1.1d}$$

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

Received by the editor February 8, 2015 and, in revised form, January 31, 2016.
2010 *Mathematics Subject Classification.* Primary 65N30.
Key words and phrases. Discontinuous Galerkin, hybridization, Navier-Stokes equations, post-processing, superconvergence.
The second author was supported in part by the National Science Foundation through DMS Grant 1115331.
The third author was partially supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. CityU 11302014). The third author is the corresponding author.

where $\boldsymbol{u}$ is the velocity, p is the pressure, ν is the kinematic viscosity and $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega)$ is the external body force. The domain $\Omega \subset \mathbb{R}^d$ is polygonal ($d = 2$) or polyhedral ($d = 3$).

To discuss our results, let us introduce the HDG method under consideration [28]. We consider conforming meshes $\mathcal{T}_h$ of Ω made of shape-regular simplexes K. We denote the set of faces F of the element K by $\mathcal{F}(K)$, by $\mathcal{E}_h$ the set of all faces F of all elements $K \in \mathcal{T}_h$ and set $\partial\mathcal{T}_h := \{\partial K : K \in \mathcal{T}_h\}$. For scalar-valued functions ϕ and ψ, we write

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

Here $(\cdot, \cdot)_D$ denotes the integral over the domain $D \subset \mathbb{R}^d$, and $\langle\cdot, \cdot\rangle_D$ denotes the integral over $D \subset \mathbb{R}^{d-1}$. For vector-valued and matrix-valued functions, a similar notation is taken. For example, we write $(\boldsymbol{\phi}, \boldsymbol{\psi})_{\mathcal{T}_h} := \sum_{i=1}^n (\phi_i, \psi_i)_{\mathcal{T}_h}$ for vector-valued functions and $(\phi, \psi)_{\mathcal{T}_h} := \sum_{1\le i,j\le n}(\phi_{ij}, \psi_{ij})_{\mathcal{T}_h}$ for matrix-valued functions.

The HDG method provides an approximation $(\mathrm{L}_h, \boldsymbol{u}_h, p_h, \widehat{\boldsymbol{u}}_h) \in \mathrm{G}_h \times \boldsymbol{V}_h \times Q_h \times \boldsymbol{M}_h^0$ to the exact solution $(\mathrm{L}|_{\mathcal{T}_h}, \boldsymbol{u}|_{\mathcal{T}_h}, p|_{\mathcal{T}_h}, \boldsymbol{u}|_{\mathcal{E}_h})$ in the finite dimensional space

$$\begin{aligned}
\mathrm{G}_h &:= \{\mathrm{G} \in \mathrm{L}^2(\Omega) : \quad \mathrm{G}|_K \in \mathrm{P}_k(K),\ \forall K \in \mathcal{T}_h\},\\
\boldsymbol{V}_h &:= \{\boldsymbol{v} \in \boldsymbol{L}^2(\Omega) : \quad \boldsymbol{v}|_K \in \boldsymbol{P}_k(K),\ \forall K \in \mathcal{T}_h\},\\
Q_h &:= \{p \in L_0^2(\Omega) : \quad p|_K \in P_k(K),\ \ \forall K \in \mathcal{T}_h\},\\
\boldsymbol{M}_h &:= \{\boldsymbol{\mu} \in \boldsymbol{L}^2(\mathcal{E}_h) : \boldsymbol{\mu}|_F \in \boldsymbol{P}_k(F),\ \ \forall F \in \mathcal{E}_h\},\\
\boldsymbol{M}_h^0 &:= \{\boldsymbol{\mu} \in \boldsymbol{M}_h : \qquad \boldsymbol{\mu}|_{\partial\Omega} = 0\}.
\end{aligned}$$

Here $P_k(D)$ denotes the set of polynomials of total degree at most $k \ge 0$ defined on D, $\boldsymbol{P}_k(D)$ denotes the set of vector-valued functions whose d components lie in $P_k(D)$, $\mathrm{P}_k(K)$ denotes the set of square matrix-valued functions whose $d\times d$ entries also lie in $P_k(D)$, and $L_0^2(\Omega) = \{p \in L^2(\Omega) : \int_\Omega p = 0\}$.

The method determines the approximate solution by requiring that it solves the following weak formulation:

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

$$\begin{aligned}(\nu\mathrm{L}_h, \nabla\boldsymbol{v})_{\mathcal{T}_h} - (\boldsymbol{u}_h \otimes \boldsymbol{\beta}, \nabla\boldsymbol{v})_{\mathcal{T}_h} - (p_h, \nabla\cdot\boldsymbol{v})_{\mathcal{T}_h}&\\
-\langle\nu\,\widehat{\mathrm{L}}_h\boldsymbol{n} - \widehat{p}_h\boldsymbol{n} - (\widehat{\boldsymbol{u}}_h \otimes \boldsymbol{\beta})\boldsymbol{n}, \boldsymbol{v}\rangle_{\partial\mathcal{T}_h} &= (\boldsymbol{f}, \boldsymbol{v})_{\mathcal{T}_h},\end{aligned} \tag{1.2b}$$

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

$$\langle\nu\,\widehat{\mathrm{L}}_h\boldsymbol{n} - \widehat{p}_h\boldsymbol{n} - (\widehat{\boldsymbol{u}}_h \otimes \boldsymbol{\beta})\boldsymbol{n}, \boldsymbol{\mu}\rangle_{\partial\mathcal{T}_h} = 0, \tag{1.2d}$$

for all $(\mathrm{G}, \boldsymbol{v}, q, \boldsymbol{\mu}) \in \mathrm{G}_h \times \boldsymbol{V}_h \times Q_h \times \boldsymbol{M}_h^0$. Here,

$$(\nu\,\widehat{\mathrm{L}}_h - \widehat{p}_h)\boldsymbol{n} := \nu\mathrm{L}_h\boldsymbol{n} - p_h\boldsymbol{n} - \mathrm{S}(\boldsymbol{u}_h - \widehat{\boldsymbol{u}}_h) \quad \text{on } \partial\mathcal{T}_h, \tag{1.2e}$$

$$\mathrm{S} := \mathrm{S}_{\boldsymbol{\beta}} + \mathrm{S}_{\boldsymbol{n}}, \tag{1.2f}$$

$$\mathrm{S}_{\boldsymbol{\beta}} := \max(\boldsymbol{\beta}\cdot\boldsymbol{n}, 0)\,\mathrm{Id}, \tag{1.2g}$$

$$\mathrm{S}_{\boldsymbol{n}} := \zeta_n\, h_K^{-1}\boldsymbol{n}\otimes\boldsymbol{n} \qquad \text{on } \partial\mathcal{T}_h, \tag{1.2h}$$

where the stabilization parameter ζ_n in (1.2h) is chosen to be a positive, and

$$\boldsymbol{\beta} = \mathbb{P}(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h). \tag{1.2i}$$

The operator $\mathbb{P}$ from $\boldsymbol{H}^1(\mathcal{T}_h) \times \boldsymbol{L}^2(\mathcal{E}_h)$ into

$$\{\boldsymbol{v} \in H(\text{div};\Omega) : \boldsymbol{v}|_K \in RT_k(K) := \boldsymbol{P}_k(K) + \boldsymbol{x}\, P_k(K)\}$$

is defined on the element K by the following equations:

$$(\mathbb{P}(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h) - \boldsymbol{u}_h, \boldsymbol{v})_K = 0 \quad \forall \boldsymbol{v} \in \boldsymbol{P}_{k-1}(K), \tag{1.3a}$$

$$\langle (\mathbb{P}(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h) - \widehat{\boldsymbol{u}}_h) \cdot \boldsymbol{n}, \lambda \rangle_{\partial K} = 0 \quad \forall \lambda \in P_k(F), \text{ for each face } F \text{ of } K. \tag{1.3b}$$

This operator is similar to that proposed back in 2003 by Bastian and Rivière [4] in the framework of Darcy flow. Here we use the Raviart-Thomas spaces instead of the Brezzi-Douglas-Marini spaces used in [4]. Also, here the definition of the numerical trace $\widehat{\boldsymbol{u}}_h$ is not the same.

Note that the stabilization tensor S used in [28, (34)] was defined in a slightly different manner, namely,

$$\mathrm{S} := |\boldsymbol{\beta}|\,\mathrm{Id} + \tau\,\mathrm{Id}. \tag{1.4}$$

For this choice, the stabilization associated to the convective part is stronger. On the other hand, since τ was taken to be an order one quantity, the stabilization on the normal part of the velocity jumps is not that strong, but that of the tangential components is. Note that for the HDG method for the Stokes equations, [9], it was shown that the diffusive part of the above stabilization tensor can be of the form

$$\tau_n\, \boldsymbol{n} \otimes \boldsymbol{n} + \tau_t\, (\mathrm{Id} - \boldsymbol{n} \otimes \boldsymbol{n}),$$

and that the convergence properties of the HDG method remain unchanged when the stabilization of the tangential component τ_t is of order one, or even equal to zero, and the inverse of the stabilization of the normal component τ_n is bounded. Here we take $\tau_t := 0$ and $\tau_n := \zeta_n\, h_K^{-1}$.

The numerical experiments carried out in [28] suggest that the L^2-norm of the error in the velocity, the pressure and even in the velocity gradient converge with the optimal order $k+1$ for any $k \geq 0$, and that an elementwise postprocessed $H(\text{div})$-conforming and divergence-free approximate velocity can be obtained which, for $k \geq 1$, converges with order $k+2$. Let us note that the above-mentioned postprocessed velocity is similar to that proposed by Bastian and Rivière in 2003 [4] in that it uses the same Brezzi-Douglas-Marini spaces of index $k+1$. However, its definition differs considerably from that of Bastian and Rivière as it uses the approximate gradient of the velocity; see [28] and the references therein.

In this paper, we put in firm mathematical grounds of the above-mentioned experimental results. We also show that, just as for the HDG method for the Stokes problem [9], the bounds of the errors depend affinely on $1/\zeta_n$. In particular, if we superpenalize the normal components of the jumps in the velocity by taking ζ_n to be huge, the accuracy of the numerical approximation to the velocity and its gradient not only remain unchanged but only require the pressure to lie in $H^1(\Omega)$. This also means that $H(\text{div})$-conforming HDG methods can immediately be obtained, as was shown in [14] for the Stokes equations.

To the knowledge of the authors, no other known finite element method for the Stokes or the Navier-Stokes equations has these properties. See the classic mixed methods [6, 17, 18], the stabilized methods proposed in [20, 21, 23] and the DG methods [3, 7, 13, 25, 29, 31]. Perhaps the only method with similar convergence properties is the one proposed for the Stokes equations in [30] by using the RT and BDM elements developed for diffusion problems. For the Navier-Stokes equations, the pioneering interior penalty (IP)-like methods proposed by [22] used piecewise-solenoidal approximate velocities but were not locally conservative. So did the method developed in [19]. Locally conservative LDG methods for the Navier-Stokes were introduced and analyzed in [10–12]. More recently, an IP-like method and a compact discontinuous Galerkin (CDG) method were introduced in [26]. These methods use the approximation space for the velocity field as a direct sum of a solenoidal space and an irrotational space, in such a way that their weak forms can be split into two uncoupled problems: one associated with velocities and hybrid pressures, and the other one only concerned with computation of pressure in the interior of the elements. Numerical experiments indicating the optimal convergence order of velocity and the pressure in L^2-norm were reported. Finally, we refer the reader to [24] for an HDG discretization for the time-dependent incompressible Navier-Stokes equations.

Let us briefly comment on the main novelty of our error analysis. Note that in [8] for the Oseen problem and [9, 15] for the Stokes equations, a simple energy argument gives the inequality

$$\nu\|\mathrm{L}_h\|^2_{\mathcal{T}_h} + \Sigma_{K\in\mathcal{T}_h}\|\tau^{1/2}\,(\boldsymbol{u}_h - \widehat{\boldsymbol{u}}_h)\|^2_{\partial K} \le (\boldsymbol{f}, \boldsymbol{u}_h)_{\mathcal{T}_h},$$

where the stabilization tensor S is defined as (1.4). The above inequality cannot lead to the much needed estimate

$$\nu\big(\|\nabla\boldsymbol{u}_h\|^2_{\mathcal{T}_h} + \Sigma_{K\in\mathcal{T}_h} h_K^{-1}\,\|\boldsymbol{u}_h - \widehat{\boldsymbol{u}}_h\|^2_{\partial K}\big) \le C\,\|\boldsymbol{f}\|^2_{\mathcal{T}_h},$$

unless we take $\tau|_{\partial K}$ to be of order νh_K^{-1}. However, in this case, it is impossible to achieve superconvergent approximations; see the numerical experiments for HDG methods for the Stokes equations in [9, 27]. In our case, with our choice of stabilization tensor, the energy argument gives

$$\nu\|\mathrm{L}_h\|^2_{\mathcal{T}_h} + \Sigma_{K\in\mathcal{T}_h}\zeta_n\, h_K^{-1}\,\|(\boldsymbol{u}_h - \widehat{\boldsymbol{u}}_h)\cdot\boldsymbol{n}\|^2_{\partial K} \le (\boldsymbol{f}, \boldsymbol{u}_h)_{\mathcal{T}_h},$$

and it would seem that we would be in an even worse predicament. However, we can prove that

$$\|\nabla\boldsymbol{u}_h\|_{\mathcal{T}_h} + \big(\Sigma_{K\in\mathcal{T}_h} h_K^{-1}\|\boldsymbol{u}_h - \widehat{\boldsymbol{u}}_h\|^2_{\partial K}\big)^{1/2} \le C\big(\|\mathrm{L}_h\|^2_{\mathcal{T}_h} + \Sigma_{K\in\mathcal{T}_h} h_K^{-1}\,\|(\boldsymbol{u}_h - \widehat{\boldsymbol{u}}_h)\cdot\boldsymbol{n}\|^2_{\partial K}\big)^{1/2}.$$

In other words, we prove that the tangential jumps of the velocity are actually controlled by the normal jumps and the approximate velocity gradient. This novel contribution is what makes the analysis work.

The organization of the paper is as follows. In Section 2, we present our main results. The rest of the paper is devoted to proving them. In Section 3, we establish the main properties of the forms defining the HDG methods. In Section 4, we prove the new stability estimate; other, more standard, estimates are gathered in the Appendix. In Section 5, we prove the existence and uniqueness of the approximation. In Section 6, we provide a detailed proof of the error estimates. We end in Section 7, with a brief comment on an $H(\mathrm{div})$-conforming version of the HDG methods we have considered here.

2. Main results

In this section, we present and briefly discuss our main results, namely, the existence and uniqueness of the HDG approximation, Theorem 2.3, and the corresponding error estimates, Theorem 2.4.

2.1. Notation.

We begin by introducing some notation. We use the standard definitions [1] for the Sobolev spaces $W^{\ell,p}(D)$ for a given domain D with norm

$$\|\phi\|_{\ell,p,D} = (\sum_{|\alpha|\leq\ell} \|D^\alpha\phi\|^p_{0,p,D})^{1/p}.$$

For vector- and matrix-valued functions $\boldsymbol{\phi}$ and Φ, we use $\|\boldsymbol{\phi}\|_{\ell,p,D} = \sum_{i=1}^d \|\boldsymbol{\phi}_i\|_{\ell,p,D}$, and $\|\Phi\|_{\ell,p,D} = \sum_{i,j=1}^d \|\Phi_{ij}\|_{\ell,p,D}$. Moreover, when $p=2$ and $\ell<\infty$, we denote $W^{\ell,2}(D)$ by $H^\ell(D)$ and $\|\cdot\|_{\ell,2,D}$, by $\|\cdot\|_{\ell,D}$; when $\ell=0$ and $p=2$, we denote $W^{0,2}(D)$ by $L^2(D)$ and the norm by $\|\cdot\|_D$.

We also introduce the following mesh-dependent norms and seminorms:

$$\begin{aligned}
|||(\boldsymbol{v},\boldsymbol{\mu})|||_{0,h} &:= \Big(\|\boldsymbol{v}\|^2_{\mathcal{T}_h} + \sum_{K\in\mathcal{T}_h} h_K\left(\|\boldsymbol{\mu}\|^2_{\partial K} + \|\boldsymbol{v}-\boldsymbol{\mu}\|^2_{\partial K}\right)\Big)^{\frac{1}{2}}, \\
|||(\boldsymbol{v},\boldsymbol{\mu})|||_{1,h} &:= \Big(\|\nabla\boldsymbol{v}\|^2_{\mathcal{T}_h} + \sum_{K\in\mathcal{T}_h} h_K^{-1}\|\boldsymbol{v}-\boldsymbol{\mu}\|^2_{\partial K}\Big)^{1/2} \qquad \forall\ (\boldsymbol{v},\boldsymbol{\mu})\in\boldsymbol{H}^1(\mathcal{T}_h)\times\boldsymbol{L}^2(\mathcal{E}_h), \\
|||(\boldsymbol{v},\boldsymbol{\mu})|||_{\infty,h} &:= \|\boldsymbol{v}\|_{L^\infty(\Omega)} + \|\boldsymbol{\mu}\|_{L^\infty(\mathcal{E}_h)} \qquad \forall\ (\boldsymbol{v},\boldsymbol{\mu})\in\boldsymbol{L}^\infty(\Omega)\times\boldsymbol{L}^\infty(\mathcal{E}_h),
\end{aligned}$$

and set

$$\|\boldsymbol{v}\|_{0,h} := \|\boldsymbol{v}\|_{L^2(\Omega)}, \quad \|\boldsymbol{v}\|_{1,h} := |||(\boldsymbol{v},\{\!\{\boldsymbol{v}\}\!\})|||_{1,h},$$

where the average of $\boldsymbol{v}$, $\{\!\{\boldsymbol{v}\}\!\}$, is defined as follows: On an interior face $F=\partial K^-\cap\partial K^+$, we have $\{\!\{\boldsymbol{v}\}\!\} := \frac{1}{2}(\boldsymbol{v}^+ + \boldsymbol{v}^-)$, where $\boldsymbol{v}^\pm$ denote the trace of $\boldsymbol{v}$ from the interior of $K^\pm$ and $\boldsymbol{n}^\pm$ is the outward unit normal to $K^\pm$. On a boundary face $F\subset\partial K^-\cap\partial\Omega$, we take $\{\!\{\boldsymbol{v}\}\!\}=\boldsymbol{0}$. Note that $\|\boldsymbol{v}\|_{1,h}$ is nothing but the standard discrete H^1-norm of $\boldsymbol{v}$; see, for example, [5].

Finally, we denote the L^2-orthogonal projections onto G_h, Q_h and $\boldsymbol{M}_h$ by Π_G, Π_Q and Π_M, respectively.

2.2. The convective velocity $\mathbb{P}(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)$.

Next, we must ensure that the convective velocity $\mathbb{P}(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)$ given by (1.3) is actually well defined. In the following result, we gather this and other of its main properties.

Proposition 2.1. *For any $(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)\in\boldsymbol{V}_h\times\boldsymbol{M}_h$, we have that $\mathbb{P}(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)$ is well defined and that*

(i) $\mathbb{P}(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)\in H(\mathrm{div},\Omega)$.
(ii) $\nabla\cdot\mathbb{P}(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)=0$ *in* Ω *if* $(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)$ *satisfies the third equation defining the HDG methods,* (1.2c).
(iii) $\mathbb{P}(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)\in\boldsymbol{V}_h$.
(iv) $\|\mathbb{P}(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)\|_{i,h}\leq C_{\mathrm{stab},i}|||(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)|||_{i,h}, \qquad i=0,1.$
(v) $\mathbb{P}(\boldsymbol{u},\boldsymbol{u}|_{\mathcal{E}_h})=\mathbb{P}(\Pi^{\mathrm{RT}}\boldsymbol{u},\Pi_M\boldsymbol{u})=\Pi^{\mathrm{RT}}\boldsymbol{u}.$

Here, Π^{RT} is the Raviart-Thomas (RT) projection from $\boldsymbol{C}^0(\overline{\Omega})$ into

$$\{\boldsymbol{v}\in H(\mathrm{div},\Omega) : \boldsymbol{v}|_K\in RT_k(K)\}$$

defined on each element K by

$$(\Pi^{\mathrm{RT}}\boldsymbol{u}-\boldsymbol{u},\boldsymbol{v})_K=0 \qquad \forall \boldsymbol{v}\in \boldsymbol{P}_{k-1}(K), \tag{2.1a}$$

$$\langle(\Pi^{\mathrm{RT}}\boldsymbol{u}-\boldsymbol{u})\cdot\boldsymbol{n},\lambda\rangle_F=0 \quad \forall\lambda\in P_k(F),\ \text{for each face } F \text{ of } K. \tag{2.1b}$$

Note that the equations (1.3) defining the projection $\mathbb{P}$ operator are almost identical to those defining the RT projection Π^{RT}. In fact, property (v) of Proposition 2.1 easily follows by simply comparing these equations.

2.3. **A new stability estimate.** As announced in the Introduction, the main novelty of our error analysis is the use of a stability estimate which we present next. It states, roughly speaking, that the weighted jumps of the tangential component of the velocity can be controlled by the gradient L and the normal jumps.

Proposition 2.2. *If* $(\mathrm{L},\boldsymbol{v},\boldsymbol{\mu})\in \mathrm{G}_h\times\boldsymbol{V}_h\times\boldsymbol{M}_h$ *satisfies the first and third equations defining the HDG method,* (1.2a) *and* (1.2c)*, respectively, then*

$$|||(\boldsymbol{v},\boldsymbol{\mu})|||_{1,h}\le C_{\mathrm{HDG}}\Big(\|\mathrm{L}\|^2_{\mathcal{T}_h}+\sum_{K\in\mathcal{T}_h}h_K^{-1}\|(\boldsymbol{v}-\boldsymbol{\mu})\cdot\boldsymbol{n}\|^2_{\partial K}\Big)^{1/2}.$$

We strongly use this estimate to obtain our main results on the HDG method under consideration, namely, Thereom 2.3 on the existence, uniqueness and boundedness of the approximation and Theorem 2.4 on a priori error estimates.

2.4. **Existence and uniqueness.** Here we establish that the HDG methods under consideration define a unique approximate solution under a classic smallness condition on $\boldsymbol{f}$, as we see in the following result.

Theorem 2.3 (Existence, uniqueness and boundedness)**.** *If the quantity* $\nu^{-2}\|\boldsymbol{f}\|_\Omega$ *is small enough, the HDG method* (1.2) *has a unique solution* $(\mathrm{L}_h,\boldsymbol{u}_h,p_h,\widehat{\boldsymbol{u}}_h)\in \mathrm{G}_h\times\boldsymbol{V}_h\times Q_h\times\boldsymbol{M}_h^0$. *Furthermore,*

$$|||(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)|||_{1,h}\le C\nu^{-1}\|\boldsymbol{f}\|_\Omega,$$

for some constant C independent of ν, the discretization parameters and the exact solution.

2.5. **A priori error estimates.** Finally, we give an estimate of the size of the projection of the approximation errors,

$$e^{\mathrm{L}}:=\Pi_G\mathrm{L}-\mathrm{L}_h,\quad e^{\boldsymbol{u}}:=\Pi^{\mathrm{RT}}\boldsymbol{u}-\boldsymbol{u}_h,\quad e^{\widehat{\boldsymbol{u}}}:=\Pi_M(\boldsymbol{u}|_{\mathcal{E}_h})-\widehat{\boldsymbol{u}}_h\quad\text{and}\quad e^p:=\Pi_Q p-p_h.$$

We provide two estimates for the error in the velocity, $e^{\boldsymbol{u}}$. The first is an optimality result and the second a superconvergence result. The latter strongly uses the solution $(\boldsymbol{\phi},\psi)$ of the dual problem

$$\Phi-\nabla\boldsymbol{\phi}=0 \qquad \text{in } \Omega, \tag{2.2a}$$

$$-\nu\nabla\cdot\Phi-\nabla\cdot(\boldsymbol{\phi}\otimes\boldsymbol{u})-\nabla\psi-(\nabla\boldsymbol{\phi})^\top\boldsymbol{u}=\boldsymbol{\theta} \qquad \text{in } \Omega, \tag{2.2b}$$

$$\nabla\cdot\boldsymbol{\phi}=0 \qquad \text{in } \Omega, \tag{2.2c}$$

$$\boldsymbol{\phi}=0 \qquad \text{on } \partial\Omega, \tag{2.2d}$$

for which we assume that we have the following regularity estimate:

$$\|\Phi\|_{1,\Omega}+\|\boldsymbol{\phi}\|_{2,\Omega}+\|\psi\|_{1,\Omega}\le C_r\|\boldsymbol{\theta}\|_\Omega. \tag{2.3}$$

This dual problem was used back in 1998 in the pioneering work on DG methods for the incompressible Navier-Stokes equations presented in [22]; the only difference is that the pressures differ by the quantity $\frac{1}{2}\boldsymbol{\phi}\cdot\boldsymbol{u}$. Therein, the above elliptic regularity estimate was obtained for a domain Ω with a $\mathcal{C}^2(\Omega)$ boundary by assuming that $\|\boldsymbol{u}\|_{H^1(\Omega)}$ is small enough compared with the viscosity coefficient ν. Under a similar condition, for Ω a convex polyhedron, and assuming that $\boldsymbol{u}\in L^\infty(\Omega)$, we can obtain the regularity inequality by using a standard regularity estimate for the Stokes equations; see [18]. We can now state our main result.

Theorem 2.4 (Error estimates)**.** *Assume that $\boldsymbol{u}\in\mathcal{C}^0(\overline{\Omega})$. Then, if $\nu^{-2}\|\boldsymbol{f}\|_\Omega$ is small enough, we have that, for any polynomial degree $k\geq 0$,*

$$\nu^{1/2}\big(\|e^{\mathrm{L}}\|_\Omega+|\!|\!|(e^{\boldsymbol{u}},e^{\widehat{\boldsymbol{u}}})|\!|\!|_{1,h}\big)+\big(\Sigma_{K\in\mathcal{T}_h}\zeta_n h_K^{-1}\|e^{\boldsymbol{u}}-e^{\widehat{\boldsymbol{u}}}\|^2_{\partial K}\big)^{1/2} \leq \big(C_{\mathrm{L}}\, h^{k+1}+C\,\big(\Sigma_{K\in\mathcal{T}_h}\zeta_n^{-1}h_K\|\Pi_Q p-p\|^2_{\partial K}\big)^{1/2}\big), \tag{2.4a}$$

$$\|e^{\boldsymbol{u}}\|_\Omega\leq\nu^{-1/2}\,\big(C_{\boldsymbol{u}}\,h^{k+1}+C\,\big(\Sigma_{K\in\mathcal{T}_h}\zeta_n^{-1}h_K\|\Pi_Q p-p\|^2_{\partial K}\big)^{1/2}\big), \tag{2.4b}$$

$$\|e^p\|_\Omega\leq C_p\,h^{k+1}. \tag{2.4c}$$

Here, the constants $C_{\mathrm{L}}, C_{\boldsymbol{u}}$ depend on $\|\boldsymbol{u}\|_{L^\infty(\Omega)}$, $\|\boldsymbol{u}\|_{k+2,\Omega}$, and k, while the constant C_p, depends on $\|\boldsymbol{u}\|_{L^\infty(\Omega)}$, $\|\boldsymbol{u}\|_{k+2,\Omega}$, $\|p\|_{k+1,\Omega},\nu$ and k.

Furthermore, if and $\nu^{-1}\|\nabla\boldsymbol{u}\|_\Omega$ is small enough, $\boldsymbol{u}\in\boldsymbol{W}^{1,\infty}(\Omega)$ and the regularity estimate (2.3) *holds, then*

$$\|e^{\boldsymbol{u}}\|_\Omega\leq\nu^{-1/2}\,h\big(C_D\,h^{k+1}+\big(\Sigma_{K\in\mathcal{T}_h}\zeta_n^{-1}h_K\|\Pi_Q p-p\|^2_{\partial K}\big)^{1/2}\big),\quad\forall k\geq 1. \tag{2.5}$$

Let $\boldsymbol{u}_h^*\in H(\mathrm{div},\Omega)$ be the postprocessed approximate velocity introduced in [9, (2.9)], *then we have $\nabla\cdot\boldsymbol{u}_h^*=0$ in Ω, and*

$$\|\boldsymbol{u}_h^*-\boldsymbol{u}\|_\Omega\leq C\,\left(\|e^{\boldsymbol{u}}\|_\Omega+h\|e^{\mathrm{L}}\|_\Omega\right)+Ch^{k+2}|\boldsymbol{u}|_{k+2,\Omega},\quad\forall k\geq 1. \tag{2.6}$$

Here, the constant C_D depends on $\|\boldsymbol{u}\|_{W^{1,\infty}(\Omega)}$ and $\|\boldsymbol{u}\|_{k+2,\Omega}$.

When we take $\zeta_n\geq 1$, by the approximation properties of Π_Q, the error estimate (2.4b) gives optimal convergence of the global L^2-norm of the error in velocity, and the estimate (2.6) gives superconvergence of the postprocessed numerical approximation to velocity. However, in order to achieve the high order accuracy of numerical approximation to the velocity mentioned above, the pressure is required to have H^{k+1}-regularity if we choose ζ_n to be of order one. If we superpenalize and take $\zeta_n:=\nu^{-1}h^{-2k}$ in (1.2h), the error estimates (2.4a), (2.4b) are

$$\nu^{1/2}\big(\|e^{\mathrm{L}}\|_\Omega+|\!|\!|(e^{\boldsymbol{u}},e^{\widehat{\boldsymbol{u}}})|\!|\!|_{1,h}\big)+\big(\Sigma_{K\in\mathcal{T}_h}\zeta_n h_K^{-1}\|e^{\boldsymbol{u}}-e^{\widehat{\boldsymbol{u}}}\|^2_{\partial K}\big)^{1/2} \leq\big(C_{\mathrm{L}}\,h^{k+1}+C\,\nu^{1/2}h^{k+1}|p|_{H^1(\Omega)}\big), \tag{2.7a}$$

$$\|e^{\boldsymbol{u}}\|_\Omega\leq\nu^{-1/2}\,\big(C_{\boldsymbol{u}}\,h^{k+1}+\nu^{1/2}h^{k+1}|p|_{H^1(\Omega)}\big), \tag{2.7b}$$

and the estimate (2.6) is

$$\|e^{\boldsymbol{u}}\|_\Omega\leq\nu^{-1/2}\,h\big(C_D\,h^{k+1}+\nu^{1/2}h^{k+1}|p|_{H^1(\Omega)}\big),\quad\forall k\geq 1. \tag{2.8}$$

These estimates show that the numerical approximation to the velocity $\boldsymbol{u}$ has high order accuracy even if the pressure p has H^1-regularity only.

To end, let us briefly mention that we are not using an extension of the auxiliary projection used in [8, (2.7)] to analyze the HDG method applied to an Oseen problem. The reason is that, according to [8, Theorem 2.3], the approximation property of this auxiliary projection depends on the $\boldsymbol{W}^{1,\infty}$-norm of the approximate velocity $\boldsymbol{\beta}$, which is not necessarily bounded in our setting. Instead, we use the standard

RT projection for the velocity on each element and L^2-orthogonal projections for all other unknowns.

3. The forms defining the method and their main properties

In order to simplify our analysis, and make it as close as possible to that proposed in [10], we rewrite the equations defining the HDG method under consideration in terms of several forms for which we then prove the continuity and coercivity properties we are going to use in the analysis.

3.1. **Rewriting the HDG method in compact form.** If we set

$$\begin{aligned}
A_h((\boldsymbol{v},\boldsymbol{\mu}),\mathrm{G}) :=& (\boldsymbol{v},\nabla\cdot\mathrm{G})_{\mathcal{T}_h} - \langle\boldsymbol{\mu},\mathrm{G}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h},\\
B_h((\boldsymbol{v},\boldsymbol{\mu}),q) :=& -(\boldsymbol{v},\nabla q)_{\mathcal{T}_h} + \langle q,\boldsymbol{\mu}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h},\\
J_n((\boldsymbol{v}_1,\boldsymbol{\mu}_1),(\boldsymbol{v}_2,\boldsymbol{\mu}_2)) :=& \sum_{K\in\mathcal{T}_h}\zeta_n h_K^{-1}\langle(\boldsymbol{v}_1-\boldsymbol{\mu}_1)\cdot\boldsymbol{n},(\boldsymbol{v}_2-\boldsymbol{\mu}_2)\cdot\boldsymbol{n}\rangle_{\partial K},\\
\mathcal{O}_h(\boldsymbol{\beta};(\boldsymbol{v}_1,\boldsymbol{\mu}_1),(\boldsymbol{v}_2,\boldsymbol{\mu}_2)) :=& -(\boldsymbol{v}_1\otimes\boldsymbol{\beta},\nabla\boldsymbol{v}_2)_{\mathcal{T}_h}\\
&+\langle(\boldsymbol{\mu}_1\otimes\boldsymbol{\beta})\boldsymbol{n}+S_{\boldsymbol{\beta}}(\boldsymbol{v}_1-\boldsymbol{\mu}_1),\boldsymbol{v}_2-\boldsymbol{\mu}_2\rangle_{\partial\mathcal{T}_h},
\end{aligned}$$

where $\mathrm{S}_{\boldsymbol{\beta}} := \max(\boldsymbol{\beta}\cdot\boldsymbol{n},0)\,\mathrm{Id}$, where the arguments of the above forms are such that all the integrals make sense, the equations defining the HDG method under consideration, (1.2), can be rewritten in compact form as follows:

$$\begin{aligned}
(\mathrm{L}_h,\mathrm{G})_{\mathcal{T}_h} &+ A_h((\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h),\mathrm{G}) - A_h((\boldsymbol{v},\boldsymbol{\mu}),\nu\mathrm{L}_h)\\
&- B_h((\boldsymbol{v},\boldsymbol{\mu}),p_h) + B_h((\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h),q)\\
&+ J_n((\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h),(\boldsymbol{v},\boldsymbol{\mu})) + \mathcal{O}_h(\boldsymbol{\beta};(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h),(\boldsymbol{v},\boldsymbol{\mu})) = (\boldsymbol{f},\boldsymbol{v})_{\mathcal{T}_h},
\end{aligned}\tag{3.1}$$

where $\boldsymbol{\beta} = \mathbb{P}(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)$, for all $(\mathrm{G},\boldsymbol{v},q,\boldsymbol{\mu}) \in \mathrm{G}_h\times\boldsymbol{V}_h\times Q_h\times\boldsymbol{M}_h^0$.

Next, we obtain the main properties of these forms.

3.2. **Properties of the bilinear forms associated to the Stokes operator.**

3.2.1. *Continuity properties of A_h and B_h.* We begin with a result similar to Proposition 4.1 in [10].

Proposition 3.1 (Continuity of A_h and B_h)**.** *There are positive constants C_A and C_B, independent of the mesh-size h, such that*

$$\begin{aligned}
|A_h((\boldsymbol{v},\boldsymbol{\mu}),\mathrm{G})| &\le C_A|||(\boldsymbol{v},\boldsymbol{\mu})|||_{1,h}\|\mathrm{G}\|_\Omega \quad \forall\,(\boldsymbol{v},\boldsymbol{\mu})\in\boldsymbol{V}(h)\times\boldsymbol{M}_h,\quad\forall\,\mathrm{G}\in\mathrm{G}_h,\\
|B_h((\boldsymbol{v},\boldsymbol{\mu}),q)| &\le C_B|||(\boldsymbol{v},\boldsymbol{\mu})|||_{1,h}\|q\|_\Omega \quad \forall\,(\boldsymbol{v},\boldsymbol{\mu})\in\boldsymbol{V}(h)\times\boldsymbol{M}_h,\quad\forall\,q\in Q_h.
\end{aligned}$$

Proof. By using integration by parts, the Cauchy-Schwarz and inverse inequalities, we get

$$\begin{aligned}
A_h((\boldsymbol{v},\boldsymbol{\mu}),\mathrm{G}) &= \sum_{K\in\mathcal{T}_h}\big(-(\nabla\boldsymbol{v},\mathrm{G})_K + \langle\boldsymbol{v}-\boldsymbol{\mu},\mathrm{G}\boldsymbol{n}\rangle_{\partial K}\big)\\
&\le C\sum_{K\in\mathcal{T}_h}(\|\nabla\boldsymbol{v}\|_K + h_K^{-1/2}\|\boldsymbol{v}-\boldsymbol{\mu}\|_{\partial K})\|\mathrm{G}\|_K,
\end{aligned}$$

and the first inequality follows. The second inequality is proven in a similar manner. This completes the proof. □

3.2.2. *Coercivity of* A_h. The next result is similar to the first result of Proposition 4.3 in [10].

Proposition 3.2. *Assume that* $(\mathrm{L}_{\boldsymbol{v},\boldsymbol{\mu}}, \boldsymbol{v}, \boldsymbol{\mu})$ *satisfies the first equation defining the HDG method,* (1.2a). *Then*

$$A_h((\boldsymbol{v},\boldsymbol{\mu}), \mathrm{L}_{\boldsymbol{v},\boldsymbol{\mu}}) = \|\mathrm{L}_{\boldsymbol{v},\boldsymbol{\mu}}\|^2_{\mathcal{T}_h}.$$

Proof. By the definition of A_h,

$$A_h((\boldsymbol{v},\boldsymbol{\mu}), \mathrm{L}_{\boldsymbol{v},\boldsymbol{\mu}}) = (\boldsymbol{v}, \nabla\cdot \mathrm{L}_{\boldsymbol{v},\boldsymbol{\mu}})_{\mathcal{T}_h} - \langle \boldsymbol{\mu}, \mathrm{L}_{\boldsymbol{v},\boldsymbol{\mu}}\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} = (\mathrm{L}_{\boldsymbol{v},\boldsymbol{\mu}}, \mathrm{L}_{\boldsymbol{v},\boldsymbol{\mu}})_{\mathcal{T}_h},$$

by the first equation defining the HDG method, (1.2a). This completes the proof. □

3.2.3. *An* inf-sup *condition.* Next, we have a result similar to Proposition 4.4 in [10].

Proposition 3.3. *For any* $q \in L^2_0(\Omega)$, *we have that*

$$\|q\|_\Omega \le \frac{C_{\boldsymbol{V},\boldsymbol{M}}}{\kappa} \sup_{(\boldsymbol{v},\boldsymbol{\mu})\in \boldsymbol{V}_h\times \boldsymbol{M}_h^0\setminus\{(\boldsymbol{0},\boldsymbol{0})\}} \frac{B_h((\boldsymbol{v},\boldsymbol{\mu}),q)}{\|\!|\!|(\boldsymbol{v},\boldsymbol{\mu})|\!|\!|_{1,h}},$$

where

$$C_{\boldsymbol{V},\boldsymbol{M}} := \sup_{\boldsymbol{w}\in \boldsymbol{H}^1_0(\Omega)\setminus\{\boldsymbol{0}\}} \frac{|\!|\!|(P\boldsymbol{w}, \Pi_M\boldsymbol{w})|\!|\!|_{1,h}}{\|\boldsymbol{w}\|_{1,\Omega}},$$

and the operator P *is defined in Appendix* A.5.

Proof. To prove the inequality, we use a standard inf-sup condition (see [18]), namely, that for any $q \in Q_h \subset L^2_0(\Omega)$, we have that

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

Since $\boldsymbol{V}(K)\times Q(K)\times \boldsymbol{M}(\partial K) := \boldsymbol{P}_k(K)\times P_k(K)\times \Pi_{F\in\mathcal{F}(K)}\,\boldsymbol{P}_k(F)$, we have that $\nabla Q(K)\subset \boldsymbol{V}(K)$ and $\boldsymbol{n}Q(K)|_{\partial K}\subset \boldsymbol{M}(\partial K)$, and so

$$(q, \nabla\cdot\boldsymbol{w})_{\mathcal{T}_h} = -(\nabla q, P\boldsymbol{w})_{\mathcal{T}_h} + \langle q\,\boldsymbol{n}, \Pi_M\boldsymbol{w}\rangle_{\partial\mathcal{T}_h} = B_h((P\boldsymbol{w}, \Pi_M\boldsymbol{w}), q).$$

Then, we get

$$\|q\|_\Omega \le \frac{1}{\kappa} \sup_{\boldsymbol{w}\in \boldsymbol{H}^1_0(\Omega)\setminus\{\boldsymbol{0}\}} \frac{B_h((P\boldsymbol{w}, \Pi_M\boldsymbol{w}), q)}{\|\boldsymbol{w}\|_{1,\Omega}},$$

and the result follows. This completes the proof. □

3.3. **Properties of the trilinear form associated to the convection,** $\mathcal{O}_h$**.** We begin by gathering several continuity properties of the form $\mathcal{O}_h$; their detailed proofs are given in the Appendix. The first one is similar to Proposition 4.2 in [10]. We use the following notation:

$$\boldsymbol{V}(h) := \boldsymbol{H}^1_0(\Omega) + \boldsymbol{V}_h.$$

Proposition 3.4. *There is a positive constant* $C_{\mathcal{O}}$ *such that*

$$|\mathcal{O}_h(\boldsymbol{\beta};(\boldsymbol{u},\widehat{\boldsymbol{u}}),(\boldsymbol{v},\boldsymbol{\mu})) - \mathcal{O}_h(\boldsymbol{\gamma};(\boldsymbol{u},\widehat{\boldsymbol{u}}),(\boldsymbol{v},\boldsymbol{\mu}))| \le C_{\mathcal{O}}\|\boldsymbol{\beta}-\boldsymbol{\gamma}\|_{1,h}|\!|\!|(\boldsymbol{u},\widehat{\boldsymbol{u}})|\!|\!|_{1,h}|\!|\!|(\boldsymbol{v},\boldsymbol{\mu})|\!|\!|_{1,h}$$

for all $\boldsymbol{\beta},\boldsymbol{\gamma}\in \boldsymbol{V}(h)$, *all* $(\boldsymbol{u},\widehat{\boldsymbol{u}})\in \boldsymbol{H}^1(\mathcal{T}_h)\times \boldsymbol{L}^2(\mathcal{E}_h)$, *and all* $(\boldsymbol{v},\boldsymbol{\mu})\in \boldsymbol{V}_h\times \boldsymbol{M}_h^0$.

We are going to use the following continuity results which take advantage of the extra regularity of some of the arguments.

Proposition 3.5. *There are positive constants* $C^{\infty}_{\mathcal{O},1}$ *and* $C^{\infty}_{\mathcal{O},2}$ *such that*

$$\mathcal{O}_h(\boldsymbol{\beta}; (\boldsymbol{u}, \widehat{\boldsymbol{u}}), (\boldsymbol{v}, \boldsymbol{\mu})) \le C^{\infty}_{\mathcal{O},1} \|\boldsymbol{\beta}\|_{L^\infty(\Omega)} \|(\boldsymbol{u}, \widehat{\boldsymbol{u}})\|_{0,h} \,|\!|\!|(\boldsymbol{v}, \boldsymbol{\mu})|\!|\!|_{1,h}$$

for all $\boldsymbol{\beta} \in \boldsymbol{C}^0(\Omega)$ *and for all* $(\boldsymbol{u}, \widehat{\boldsymbol{u}}) \in \boldsymbol{H}^1(\mathcal{T}_h) \times L^2(\mathcal{E}_h)$ *and all* $(\boldsymbol{v}, \boldsymbol{\mu}) \in \boldsymbol{V}_h \times \boldsymbol{M}_h^0$*, and*

$$\begin{aligned}|\mathcal{O}_h(\boldsymbol{\beta}; (\boldsymbol{u}, \widehat{\boldsymbol{u}}), (\boldsymbol{v}, \boldsymbol{\mu})) - \mathcal{O}_h(\boldsymbol{\gamma}; (\boldsymbol{u}, \widehat{\boldsymbol{u}}), (\boldsymbol{v}, \boldsymbol{\mu}))| \\ \le C^{\infty}_{\mathcal{O},2} \|(\boldsymbol{\beta} - \boldsymbol{\gamma}, \boldsymbol{0})\|_{0,h} |\!|\!|(\boldsymbol{u}, \widehat{\boldsymbol{u}})|\!|\!|_{\infty,h} \,|\!|\!|(\boldsymbol{v}, \boldsymbol{\mu})|\!|\!|_{1,h}\end{aligned}$$

for all $\boldsymbol{\beta}, \boldsymbol{\gamma} \in \boldsymbol{V}(h)$*, all* $(\boldsymbol{u}, \widehat{\boldsymbol{u}}) \in \boldsymbol{L}^\infty(\Omega) \times \boldsymbol{L}^\infty(\mathcal{E}_h)$*, and all* $(\boldsymbol{v}, \boldsymbol{\mu}) \in \boldsymbol{V}_h \times \boldsymbol{M}_h^0$*.*

Finally, we present a simple coercivity estimate similar to the one in Proposition 4.3 in [10].

Proposition 3.6. *Let* $\boldsymbol{\beta} \in \{\boldsymbol{v} \in H(\mathrm{div}; \Omega) : \nabla \cdot \boldsymbol{v} = 0,\ \boldsymbol{v}|_K \in \boldsymbol{H}^1(K), \forall K \in \mathcal{T}_h\}$*. Then*

$$\mathcal{O}_h(\boldsymbol{\beta}; (\boldsymbol{v}, \boldsymbol{\mu}), (\boldsymbol{v}, \boldsymbol{\mu})) = \langle (\mathrm{S}_{\boldsymbol{\beta}} - \frac{1}{2}\boldsymbol{\beta} \cdot \boldsymbol{n}\, \mathrm{Id})\, (\boldsymbol{v} - \boldsymbol{\mu}), \boldsymbol{v} - \boldsymbol{\mu}\rangle_{\partial\mathcal{T}_h} \ge 0 \quad \forall (\boldsymbol{v}, \boldsymbol{\mu}) \in \boldsymbol{V}_h \times \boldsymbol{M}_h.$$

Proof. Since $(\nabla \cdot (\boldsymbol{v} \otimes \boldsymbol{\beta}), \boldsymbol{v})_{\mathcal{T}_h} = \frac{1}{2}\langle (\boldsymbol{v} \otimes \boldsymbol{\beta})\boldsymbol{n}, \boldsymbol{v}\rangle_{\partial\mathcal{T}_h}$ for any divergence-free function $\boldsymbol{\beta}$, we have

$$\begin{aligned}&\mathcal{O}_h(\boldsymbol{\beta}; (\boldsymbol{v}, \boldsymbol{\mu}), (\boldsymbol{v}, \boldsymbol{\mu})) = -(\boldsymbol{v} \otimes \boldsymbol{\beta}, \nabla \boldsymbol{v})_{\mathcal{T}_h} + \langle (\boldsymbol{\mu} \otimes \boldsymbol{\beta})\boldsymbol{n} + S_{\boldsymbol{\beta}}(\boldsymbol{v} - \boldsymbol{\mu}), \boldsymbol{v} - \boldsymbol{\mu}\rangle_{\partial\mathcal{T}_h} \\ &= \frac{1}{2}\langle (\boldsymbol{v} \otimes \boldsymbol{\beta})\boldsymbol{n}, \boldsymbol{v}\rangle_{\partial\mathcal{T}_h} - \langle (\boldsymbol{v} \otimes \boldsymbol{\beta})\boldsymbol{n}, \boldsymbol{\mu}\rangle_{\partial\mathcal{T}_h} - \langle ((\boldsymbol{v} - \boldsymbol{\mu}) \otimes \boldsymbol{\beta})\boldsymbol{n} - S_{\boldsymbol{\beta}}(\boldsymbol{v} - \boldsymbol{\mu}), \boldsymbol{v} - \boldsymbol{\mu}\rangle_{\partial\mathcal{T}_h}.\end{aligned}$$

Because $\boldsymbol{\mu}$ is single-valued and $\boldsymbol{\beta} \in H(\mathrm{div}; \Omega)$, $\frac{1}{2}\langle (\boldsymbol{\mu} \otimes \boldsymbol{\beta})\boldsymbol{n}, \boldsymbol{\mu}\rangle_{\partial\mathcal{T}_h} = 0$. Adding this term, we get, after a few rearrangements, that

$$\mathcal{O}_h(\boldsymbol{\beta}; (\boldsymbol{v}, \boldsymbol{\mu}), (\boldsymbol{v}, \boldsymbol{\mu})) = -\frac{1}{2}\langle ((\boldsymbol{v} - \boldsymbol{\mu}) \otimes \boldsymbol{\beta})\boldsymbol{n}, \boldsymbol{v} - \boldsymbol{\mu}\rangle_{\partial\mathcal{T}_h} + S_{\boldsymbol{\beta}}(\boldsymbol{v} - \boldsymbol{\mu}), \boldsymbol{v} - \boldsymbol{\mu}\rangle_{\partial\mathcal{T}_h},$$

and the result follows. This completes the proof. □

4. Proof of the stability estimate of Proposition 2.2

In this section, we prove the new stability estimate of Proposition 2.2. More standard estimates are gathered and proved in the Appendix. We proceed in several steps.

Step 1: A representation of the matrix $\nabla \boldsymbol{v}$. Let us recall the following result on the representation of square matrices [9, Lemma 4.8].

Lemma 4.1. *The set*

$$\mathcal{B}_K := \{\mathrm{Id}\} \cup \{\boldsymbol{t} \otimes \boldsymbol{n}_F : F \text{ is a face of } K, \boldsymbol{t} \in \mathcal{B}_F\},$$

where Id *is the* $n \times n$ *identity matrix, is a basis of the space of* $n \times n$ *matrices where* $\mathcal{B}_F$ *is an orthogonal basis of the vectors orthogonal to* $\boldsymbol{n}_F$ *(tangent vectors) for each face* F *of* K*. Furthermore, the dual basis of* $\mathcal{B}_K$ *is of the form*

$$\mathcal{B}_K^* := \left\{\frac{1}{d}\mathrm{Id}\right\} \cup \{\mathrm{W}_{F,t} : F \text{ is a face of } K, \boldsymbol{t} \in \mathcal{B}_F\}$$

where $\mathrm{W}_{F,t}$ *is uniformly bounded with respect to* F *and* $\boldsymbol{t}$ *and the bound depends only on the shape regularity parameter of the mesh.*

Since $\nabla \boldsymbol{v}$ is a square matrix, we can write

$$\nabla \boldsymbol{v} = \frac{1}{d}\,(\nabla\cdot \boldsymbol{v})\ \mathrm{Id} + \sum_{F\in\partial K}\sum_{\boldsymbol{t}\in\mathcal{B}_F} \alpha_{F,\boldsymbol{t}}(\boldsymbol{t}\otimes \boldsymbol{n}_F),$$

where $\alpha_{F,\boldsymbol{t}} := \nabla \boldsymbol{v} : \mathrm{W}_{F,\boldsymbol{t}}$.

Step 2: Estimate of $\nabla\cdot\boldsymbol{v}$. Next, we estimate the $L^2(\Omega)$-norm of the divergence of $\boldsymbol{v}$.

Lemma 4.2. *Let $(\boldsymbol{v},\boldsymbol{\mu}) \in \boldsymbol{V}_h \times \boldsymbol{M}_h$ satisfy the third equation defining the HDG methods,* (1.2c). *Then, there is a constant $C > 0$ such that*

$$\|\nabla\cdot \boldsymbol{v}\|^2_{\mathcal{T}_h} \le C \sum_{K\in\mathcal{T}_h} h_K^{-1}\|(\boldsymbol{v}-\boldsymbol{\mu})\cdot \boldsymbol{n}\|^2_{\partial K}.$$

Proof. If we integrate by parts in the equation (1.2c), we get

$$\begin{aligned}(\nabla\cdot\boldsymbol{v}, q)_K =& \langle(\boldsymbol{v}-\boldsymbol{\mu})\cdot\boldsymbol{n}, q\rangle_{\partial K}\\ &\le \|(\boldsymbol{v}-\boldsymbol{\mu})\cdot\boldsymbol{n}\|_{\partial K}\|q\|_{\partial K}\\ &\le C h_K^{-1/2}\|(\boldsymbol{v}-\boldsymbol{\mu})\cdot\boldsymbol{n}\|_{\partial K}\|q\|_K,\end{aligned}$$

by a simple inverse inequality; the constant C depends on the shape regularity constant of the mesh and the dimension of the space. Now, taking $q := \nabla\cdot\boldsymbol{v}$, we get $\|\nabla\cdot\boldsymbol{v}\|_K \le C\, h_K^{-1/2}\|(\boldsymbol{v}-\boldsymbol{\mu})\cdot\boldsymbol{n}\|_{\partial K}$, and the result follows. This concludes the proof. ☐

Step 3: Estimate of $\nabla\boldsymbol{v}$. Now, we estimate the $L^2(\Omega)$-norm of the gradient of $\boldsymbol{v}$.

Lemma 4.3. *Let $(\mathrm{L},\boldsymbol{v},\boldsymbol{\mu}) \in \mathrm{G}_h \times \boldsymbol{V}_h \times \boldsymbol{M}_h$ satisfy the first and third equations defining the HDG method,* (1.2a) *and* (1.2c)*, respectively. Then, there is a constant $C > 0$ such that*

$$\|\nabla\boldsymbol{v}\|^2_{\mathcal{T}_h} \le C\big(\|\mathrm{L}\|^2_{\mathcal{T}_h} + \sum_{K\in\mathcal{T}_h} h_K^{-1}\|(\boldsymbol{v}-\boldsymbol{\mu})\cdot\boldsymbol{n}\|^2_{\partial K}\big).$$

Proof. Applying the triangle and the Cauchy-Schwarz inequalities to the last equality of Step 1, we get

$$\|\nabla\boldsymbol{v}\|_K \le \frac{1}{d}\|\nabla\cdot\boldsymbol{v}\|_K + C(\Sigma_{F\in\partial K}\Sigma_{\boldsymbol{t}\in\mathcal{B}_F}\|\alpha_{F,\boldsymbol{t}}\|^2_K)^{\frac{1}{2}},$$

where C depends on the shape regularity constant of the mesh and the dimension of the space. The first term on the right-hand side can be bounded as wanted by using Lemma 4.2.

It remains to bound the second term on the right-hand side. Define

$$\mathrm{G}^* := \sum_{F\in\partial K}\sum_{\boldsymbol{t}\in\mathcal{B}_F}\lambda_F\alpha_{F,\boldsymbol{t}}\mathrm{W}_{F,\boldsymbol{t}}$$

where λ_F stands for the barycentric coordinates associated with the face F so that $\lambda_F = 0$ on F. Observe that, since $\alpha_{F,\boldsymbol{t}} = \nabla\boldsymbol{v} : W_{F,\boldsymbol{t}} \in P_{k-1}(K)$ for any face F of K, we have that $\mathrm{G}^* \in \mathrm{P}_k(K)$. Because we assumed that $(\mathrm{L},\boldsymbol{v},\boldsymbol{\mu})$ satisfies the first equation defining the HDG method, (1.2a), we can take $\mathrm{G} := \mathrm{G}^*$ therein to get

$$\begin{aligned}(\mathrm{L}, \mathrm{G}^*)_K + \langle \boldsymbol{v} - \boldsymbol{\mu}, \mathrm{G}^* \boldsymbol{n}\rangle_{\partial K} &= (\nabla \boldsymbol{v}, \mathrm{G}^*)_K \\ &= \Big(\frac{1}{d}(\nabla\cdot \boldsymbol{v})\mathrm{Id} + \sum_{F\in\partial K}\sum_{\boldsymbol{t}\in\mathcal{B}_F} \alpha_{F,\boldsymbol{t}}(\boldsymbol{t}\otimes \boldsymbol{n}_F), \sum_{F'\in\partial K}\sum_{\boldsymbol{t}'\in\mathcal{B}_{F'}} \lambda_{F'}\alpha_{F',\boldsymbol{t}'}\mathrm{W}_{F',\boldsymbol{t}'}\Big)_K \\ &= \sum_{F\in\partial K}\sum_{\boldsymbol{t}\in\mathcal{B}_F} (\alpha_{F,\boldsymbol{t}}, \lambda_F \alpha_{F,\boldsymbol{t}})_K,\end{aligned}$$

since, by construction, $\mathrm{Id} : \mathrm{W}_{F',\boldsymbol{t}'} = 0$ and $(\boldsymbol{t}\otimes\boldsymbol{n}_F) : \mathrm{W}_{F',\boldsymbol{t}'} = 0$ except when $F = F'$ and $\boldsymbol{t} = \boldsymbol{t}'$ case in which $(\boldsymbol{t}\otimes\boldsymbol{n}_F) : \mathrm{W}_{F',\boldsymbol{t}'} = 1$. Therefore, since $0 < \lambda_F$ on K, there exists $C > 0$ such that

$$\begin{aligned}C\sum_{F\in\partial K}\sum_{\boldsymbol{t}\in\mathcal{B}_F}\|\alpha_{F,\boldsymbol{t}}\|_K^2 \le \sum_{F\in\partial K}\sum_{\boldsymbol{t}\in\mathcal{B}_F}(\alpha_{F,\boldsymbol{t}}, \lambda_F\alpha_{F,\boldsymbol{t}})_K &= (\mathrm{L}, \mathrm{G}^*)_K + \langle \boldsymbol{v} - \boldsymbol{\mu}, \mathrm{G}^*\boldsymbol{n}\rangle_{\partial K} \\ &\le \|\mathrm{L}\|_K\|\mathrm{G}^*\|_K + \langle \boldsymbol{v} - \boldsymbol{\mu}, \mathrm{G}^*\boldsymbol{n}\rangle_{\partial K}.\end{aligned}$$

Now we bound $\langle \boldsymbol{v} - \boldsymbol{\mu}, \mathrm{G}^*\boldsymbol{n}\rangle_{\partial K}$. By the definition of G^*, we get

$$\begin{aligned}\langle \boldsymbol{v} - \boldsymbol{\mu}, \mathrm{G}^*\boldsymbol{n}\rangle_{\partial K} &= \sum_{F'\in\partial K}\sum_{F\in\partial K}\sum_{\boldsymbol{t}\in\mathcal{B}_F}\langle(\boldsymbol{v}-\boldsymbol{\mu})\otimes\boldsymbol{n}_{F'}, \lambda_F\alpha_{F,\boldsymbol{t}}\mathrm{W}_{F,\boldsymbol{t}}\rangle_{F'} \\ &= \sum_{F'\in\partial K}\sum_{F\in\partial K}\sum_{\boldsymbol{t}\in\mathcal{B}_F}\langle((\boldsymbol{v}-\boldsymbol{\mu})\cdot\boldsymbol{n}_{F'})\boldsymbol{n}_{F'}\otimes\boldsymbol{n}_{F'}, \lambda_F\alpha_{F,\boldsymbol{t}}\mathrm{W}_{F,\boldsymbol{t}}\rangle_{F'}\end{aligned}$$

since on F', $\lambda_F = 0$ whenever $F = F'$ and $\boldsymbol{t}_{F'}\otimes\boldsymbol{n}_{F'} : \mathrm{W}_{F,\boldsymbol{t}} = 0$ whenever $F \neq F'$. Since $\lambda_F \le 1$ on K and $\mathrm{W}_{F,\boldsymbol{t}}$ is uniformly bounded, we obtain

$$\langle \boldsymbol{v} - \boldsymbol{\mu}, \mathrm{G}^*\boldsymbol{n}\rangle_{\partial K} \le C h_K^{-1/2}\|(\boldsymbol{v}-\boldsymbol{\mu})\cdot\boldsymbol{n}\|_{\partial K}\Big(\sum_{F\in\partial K}\sum_{\boldsymbol{t}\in\mathcal{B}_F}\|\alpha_{F,\boldsymbol{t}}\|_K^2\Big)^{\frac{1}{2}}.$$

Also, we have that

$$\|\mathrm{G}^*\|_K \le C\sum_{F\in\partial K}\sum_{\boldsymbol{t}\in\mathcal{B}_F}\|\alpha_{F,\boldsymbol{t}}\|_K \le C\Big(\sum_{F\in\partial K}\sum_{\boldsymbol{t}\in\mathcal{B}_F}\|\alpha_{F,\boldsymbol{t}}\|_K^2\Big)^{\frac{1}{2}}$$

since, again, $\lambda_F \le 1$ on K and $\mathrm{W}_{F,t}$ is uniformly bounded. The constant C depends on the shape-regularity constant of the mesh and the dimension of the space. Combining the above results, we get

$$\sum_{F\in\partial K}\sum_{\boldsymbol{t}\in\mathcal{B}_F}\|\alpha_{F,\boldsymbol{t}}\|_K^2 \le C\Big(\|\mathrm{L}\|_K^2 + Ch_K^{-1}\|(\boldsymbol{v}-\boldsymbol{\mu})\cdot\boldsymbol{n}\|_{\partial K}^2\Big)^{1/2}\Big(\sum_{F\in\partial K}\sum_{\boldsymbol{t}\in\mathcal{B}_F}\|\alpha_{F,\boldsymbol{t}}\|_K^2\Big)^{\frac{1}{2}}.$$

This completes the proof of Lemma 4.3. □

Step 4: Estimate of the tangential trace. To conclude the proof of the stability estimate in Proposition 2.2, we bound the tangential component of the jump as indicated in the following result.

Lemma 4.4. *If* $(\mathrm{L}, \boldsymbol{v}, \boldsymbol{\mu}) \in \mathrm{G}_h \times \boldsymbol{V}_h \times \boldsymbol{M}_h$ *satisfies the first and third equations defining the HDG method, that is,* (1.2a) *and* (1.2c)*, respectively. Then, there is a constant* $C > 0$ *such that*

$$\sum_{K\in\mathcal{T}_h} h_K^{-1}\|(\boldsymbol{v}-\boldsymbol{\mu})\times\boldsymbol{n}\|_{\partial K}^2 \le C\big(\|\mathrm{L}\|_{\mathcal{T}_h}^2 + \sum_{K\in\mathcal{T}_h} h_K^{-1}\|(\boldsymbol{v}-\boldsymbol{\mu})\cdot\boldsymbol{n}\|_{\partial K}^2\big).$$

Proof. The proof follows similar arguments as in the proof of Lemma 4.3. Define

$$\mathrm{G}^* := \sum_{F\in\partial K}\sum_{\boldsymbol{t}\in\mathcal{B}_F} \eta_{F,\boldsymbol{t}}\mathrm{W}_{F,\boldsymbol{t}}$$

where $\eta_{F,\boldsymbol{t}}$ is any extension of $(\boldsymbol{v}-\boldsymbol{\mu})\cdot\boldsymbol{t}$ from F to K such that $\eta_{F,\boldsymbol{t}}|_F = (\boldsymbol{v}-\boldsymbol{\mu})\cdot\boldsymbol{t}$ and such that $\|\eta_{F,\boldsymbol{t}}\|_K$ is equivalent to $h_K^{1/2}\|(\boldsymbol{v}-\boldsymbol{\mu})\cdot\boldsymbol{t}\|_F$. Since $(\mathrm{L},\boldsymbol{v},\boldsymbol{\mu})\in \mathrm{G}_h\times\boldsymbol{V}_h\times\boldsymbol{M}_h$ satisfies the first equation defining the HDG method, (1.2a), we set $\mathrm{G} := \mathrm{G}^* \in \mathrm{P}_k(K)$ therein to get

$$(\mathrm{L},\mathrm{G}^*)_K + \langle \boldsymbol{v}-\boldsymbol{\mu}, \mathrm{G}^*\boldsymbol{n}\rangle_{\partial K} = (\nabla\boldsymbol{v},\mathrm{G}^*)_K,$$

or, equivalently,

$$\langle \boldsymbol{n}\times(\boldsymbol{v}-\boldsymbol{\mu})\times\boldsymbol{n}, \mathrm{G}^*\boldsymbol{n}\rangle_{\partial K} = (\nabla\boldsymbol{v},\mathrm{G}^*)_K - (\mathrm{L},\mathrm{G}^*)_K - \langle((\boldsymbol{v}-\boldsymbol{\mu})\cdot\boldsymbol{n})\,\boldsymbol{n}, \mathrm{G}^*\boldsymbol{n}\rangle_{\partial K}.$$

Since we defined the coefficient $\eta_{F,\boldsymbol{t}}$ in such a way that

$$\|(\boldsymbol{v}-\boldsymbol{\mu})\times\boldsymbol{n}\|_{\partial K}^2 = \langle \boldsymbol{n}\times(\boldsymbol{v}-\boldsymbol{\mu})\times\boldsymbol{n}, \mathrm{G}^*\boldsymbol{n}\rangle_{\partial K},$$

we immediately get that, by Lemma 4.3,

$$\|(\boldsymbol{v}-\boldsymbol{\mu})\times\boldsymbol{n}\|_{\partial K}^2 \le C(\|\mathrm{L}\|_K^2 + \sum_{K\in\mathcal{T}_h} h_K^{-1}\|(\boldsymbol{v}-\boldsymbol{\mu})\cdot\boldsymbol{n}\|_{\partial K}^2)^{1/2}\|\mathrm{G}^*\|_K,$$

which implies the result by the definition of G^*. □

Step 5: Conclusion. Proposition 2.2 now follows from Lemmas 4.3 and 4.4.

5. Proof of the existence, uniqueness and boundedness of the approximate solution

In this section, we prove Theorem 2.3 on the existence, uniqueness and boundedness of the approximate solution of the HDG method. The idea is to define mapping $\mathcal{F}$ on

$$\boldsymbol{Z}_h = \{(\boldsymbol{v},\boldsymbol{\mu})\in \boldsymbol{V}_h\times\boldsymbol{M}_h^0 : B_h((\boldsymbol{v},\boldsymbol{\mu}),q) = 0, \quad \forall q\in Q_h\}, \tag{5.1}$$

such that any of its fixed points satisfies the equations defining the HDG method, and prove that, on a certain ball contained in $\boldsymbol{Z}_h$, it is a contraction.

Step 1: Definition of the operator $\mathcal{F}$. We start by defining $\mathcal{F}$. For $(\boldsymbol{w},\widehat{\boldsymbol{w}})\in\boldsymbol{Z}_h$, we take $\mathcal{F}(\boldsymbol{w},\widehat{\boldsymbol{w}})$ to be the component $(\boldsymbol{u},\widehat{\boldsymbol{u}})$ of the solution $(\mathrm{L},\boldsymbol{u},p,\widehat{\boldsymbol{u}})\in \mathrm{G}_h\times\boldsymbol{V}_h\times Q_h\times\boldsymbol{M}_h^0$ of

$$\begin{aligned}(\mathrm{L},\mathrm{G})_{\mathcal{T}_h} &+ A_h((\boldsymbol{u},\widehat{\boldsymbol{u}}),\mathrm{G}) - A_h((\boldsymbol{v},\boldsymbol{\mu}),\nu\mathrm{L}) - B_h((\boldsymbol{v},\boldsymbol{\mu}),p) + B_h((\boldsymbol{u},\widehat{\boldsymbol{u}}),q)\\ &+ J_n((\boldsymbol{u},\widehat{\boldsymbol{u}}),(\boldsymbol{v},\boldsymbol{\mu})) + \mathcal{O}_h(\mathbb{P}(\boldsymbol{w},\widehat{\boldsymbol{w}});(\boldsymbol{u},\widehat{\boldsymbol{u}}),(\boldsymbol{v},\boldsymbol{\mu})) = (\boldsymbol{f},\boldsymbol{v})_{\mathcal{T}_h}\end{aligned} \tag{5.2}$$

for all $(\mathrm{G},\boldsymbol{v},q,\boldsymbol{\mu})\in \mathrm{G}_h\times\boldsymbol{V}_h\times Q_h\times\boldsymbol{M}_h^0$. We clearly see that any of its fixed points satisfy the equations defining the HDG approximate solution; see (3.1).

Step 2: Proof of the upper bound of the approximate solution. Next, we establish the boundedness result of Theorem 2.3 assuming that the solution exists. We have, by Proposition 2.2, that

$$\nu|\!|\!|(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)|\!|\!|_{1,h}^2 \le C_{\rm HDG}^2\,\nu\Big(\|{\rm L}_h\|_{\mathcal{T}_h}^2+\sum_{K\in\mathcal{T}_h} h_K^{-1}\|(\boldsymbol{u}_h-\widehat{\boldsymbol{u}}_h)\cdot\boldsymbol{n}\|_{\partial K}^2\Big).$$

To estimate the right-hand side, we use an energy argument. Thus, we take ${\rm G}:=\nu{\rm L}_h$, $(\boldsymbol{v},\boldsymbol{\mu}):=(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)$ and $q:=p_h$ in the compact formulation of the HDG method, (3.1). By the identity of Proposition 3.6 and by the definition of the stabilization tensor ${\rm S}_{\boldsymbol{\beta}}$, we obtain

$$\nu\,\|{\rm L}_h\|_{\mathcal{T}_h}^2+\langle\frac{1}{2}|\boldsymbol{\beta}\cdot\boldsymbol{n}|(\boldsymbol{u}_h-\widehat{\boldsymbol{u}}_h),\boldsymbol{u}_h-\widehat{\boldsymbol{u}}_h\rangle_{\partial\mathcal{T}_h}+\nu\sum_{K\in\mathcal{T}_h} h_K^{-1}\|(\boldsymbol{u}_h-\widehat{\boldsymbol{u}}_h)\cdot\boldsymbol{n}\|_{\partial K}^2 = (\boldsymbol{f},\boldsymbol{u}_h)_{\mathcal{T}_h}.$$

As a consequence, we get that

$$\nu\,|\!|\!|(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)|\!|\!|_{1,h}^2 \le C_{\rm HDG}^2\,\nu\,\|\boldsymbol{f}\|_\Omega\,\|\boldsymbol{u}_h\|_{\mathcal{T}_h} \le C_2C_{\rm HDG}^2\,\|\boldsymbol{f}\|_\Omega|\!|\!|(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)|\!|\!|_{1,h}$$

by the second (discrete Poincaré) inequality of Proposition A.2 with $q=2$. This proves the stability result of Theorem 2.3.

It also shows that $\mathcal{F}$ maps K_h into K_h, where

$$K_h:=\{(\boldsymbol{v},\boldsymbol{\mu})\in\boldsymbol{Z}_h:|\!|\!|(\boldsymbol{v},\boldsymbol{\mu})|\!|\!|_{1,h}\le C_2C_{\rm HDG}^2\nu^{-1}\|\boldsymbol{f}\|_\Omega\}.$$

Step 3: The operator $\mathcal{F}$ is a contraction on K_h. To prove this, let $(\boldsymbol{w}_1,\widehat{\boldsymbol{w}}_1)$, $(\boldsymbol{w}_2,\widehat{\boldsymbol{w}}_2)\in K_h$ and set $(\boldsymbol{u}_1,\widehat{\boldsymbol{u}}_1):=\mathcal{F}(\boldsymbol{w}_1,\widehat{\boldsymbol{w}}_1)$ and $(\boldsymbol{u}_2,\widehat{\boldsymbol{u}}_2):=\mathcal{F}(\boldsymbol{w}_2,\widehat{\boldsymbol{w}}_2)$. By definition, there exists ${\rm L}_1,{\rm L}_2\in{\rm G}_h$, $p_1,p_2\in Q_h$ such that both $({\rm L}_1,\boldsymbol{u}_1,p_1,\widehat{\boldsymbol{u}}_1)$ and $({\rm L}_2,p_2,\boldsymbol{u}_2,\widehat{\boldsymbol{u}}_2)$ satisfy (5.2).

If we now set $\delta_{\rm L}:={\rm L}_1-{\rm L}_2$, $\delta_{\boldsymbol{u}}:=\boldsymbol{u}_1-\boldsymbol{u}_2$, $\delta_p:=p_1-p_2$ and $\delta_{\widehat{\boldsymbol{u}}}:=\widehat{\boldsymbol{u}}_1-\widehat{\boldsymbol{u}}_2$, we get that

$$\begin{aligned}(\delta_{\rm L},{\rm G})_{\mathcal{T}_h}&+A_h((\delta_{\boldsymbol{u}},\delta_{\widehat{\boldsymbol{u}}}),{\rm G})-A_h((\boldsymbol{v},\boldsymbol{\mu}),\nu\delta_{\rm L})-B_h((\boldsymbol{v},\boldsymbol{\mu}),\delta_p)+B_h((\delta_{\boldsymbol{u}},\delta_{\widehat{\boldsymbol{u}}}),q)\\&+J_n((\delta_{\boldsymbol{u}},\delta_{\widehat{\boldsymbol{u}}}),(\boldsymbol{v}-\boldsymbol{\mu}))+\mathcal{O}_h(\mathbb{P}(\boldsymbol{w}_1,\widehat{\boldsymbol{w}}_1);(\boldsymbol{u}_1,\widehat{\boldsymbol{u}}_1),(\boldsymbol{v},\boldsymbol{\mu}))\\&-\mathcal{O}_h(\mathbb{P}(\boldsymbol{w}_2,\widehat{\boldsymbol{w}}_2);(\boldsymbol{u}_2,\widehat{\boldsymbol{u}}_2),(\boldsymbol{v},\boldsymbol{\mu}))=0\end{aligned}$$

for all $({\rm G},\boldsymbol{v},q,\boldsymbol{\mu})\in{\rm G}_h\times\boldsymbol{V}_h\times Q_h\times\boldsymbol{M}_h^0$. Taking $({\rm G},\boldsymbol{v},q,\boldsymbol{\mu}):=(\nu\delta_{\rm L},\delta_{\boldsymbol{u}},\delta_p,\delta_{\widehat{\boldsymbol{u}}})$, we obtain

$$\begin{aligned}\nu\|\delta_{\rm L}\|_{\mathcal{T}_h}^2&+J_n((\delta_{\boldsymbol{u}},\delta_{\widehat{\boldsymbol{u}}}),(\delta_{\boldsymbol{u}},\delta_{\widehat{\boldsymbol{u}}}))+\mathcal{O}_h(\mathbb{P}(\boldsymbol{w}_1,\widehat{\boldsymbol{w}}_1);(\boldsymbol{u}_1,\widehat{\boldsymbol{u}}_1),(\delta_{\boldsymbol{u}},\delta_{\widehat{\boldsymbol{u}}}))\\&-\mathcal{O}_h(\mathbb{P}(\boldsymbol{w}_2,\widehat{\boldsymbol{w}}_2);(\boldsymbol{u}_2,\widehat{\boldsymbol{u}}_2),(\delta_{\boldsymbol{u}},\delta_{\widehat{\boldsymbol{u}}}))=0.\end{aligned}$$

As a consequence, by Proposition 2.2,

$$\begin{aligned}\nu|\!|\!|(\delta_{\boldsymbol{u}},\delta_{\widehat{\boldsymbol{u}}})|\!|\!|_{1,h}^2\le{}&\nu\|\delta_{\rm L}\|_{\mathcal{T}_h}^2+J_n((\delta_{\boldsymbol{u}},\delta_{\widehat{\boldsymbol{u}}}),(\delta_{\boldsymbol{u}},\delta_{\widehat{\boldsymbol{u}}}))\\&=C_{\rm HDG}^2\big(\mathcal{O}_h(\mathbb{P}(\boldsymbol{w}_2,\widehat{\boldsymbol{w}}_2);(\boldsymbol{u}_2,\widehat{\boldsymbol{u}}_2),(\delta_{\boldsymbol{u}},\delta_{\widehat{\boldsymbol{u}}}))-\mathcal{O}_h(\mathbb{P}(\boldsymbol{w}_1,\widehat{\boldsymbol{w}}_1);(\boldsymbol{u}_1,\widehat{\boldsymbol{u}}_1),(\delta_{\boldsymbol{u}},\delta_{\widehat{\boldsymbol{u}}}))\big)\\&=C_{\rm HDG}^2\big(\mathcal{O}_h(\mathbb{P}(\boldsymbol{w}_2,\widehat{\boldsymbol{w}}_2);(\boldsymbol{u}_1,\widehat{\boldsymbol{u}}_1),(\delta_{\boldsymbol{u}},\delta_{\widehat{\boldsymbol{u}}}))-\mathcal{O}_h(\mathbb{P}(\boldsymbol{w}_1,\widehat{\boldsymbol{w}}_1);(\boldsymbol{u}_1,\widehat{\boldsymbol{u}}_1),(\delta_{\boldsymbol{u}},\delta_{\widehat{\boldsymbol{u}}}))\big)\\&\quad-C_{\rm HDG}^2\mathcal{O}_h(\mathbb{P}(\boldsymbol{w}_2,\widehat{\boldsymbol{w}}_2);(\delta_{\boldsymbol{u}},\delta_{\widehat{\boldsymbol{u}}}),(\delta_{\boldsymbol{u}},\delta_{\widehat{\boldsymbol{u}}}))\\&\le C_{\rm HDG}^2\big(\mathcal{O}_h(\mathbb{P}(\boldsymbol{w}_2,\widehat{\boldsymbol{w}}_2);(\boldsymbol{u}_1,\widehat{\boldsymbol{u}}_1),(\delta_{\boldsymbol{u}},\delta_{\widehat{\boldsymbol{u}}}))-\mathcal{O}_h(\mathbb{P}(\boldsymbol{w}_1,\widehat{\boldsymbol{w}}_1);(\boldsymbol{u}_1,\widehat{\boldsymbol{u}}_1),(\delta_{\boldsymbol{u}},\delta_{\widehat{\boldsymbol{u}}}))\big)\\&\le C_{\rm HDG}^2\big(C_{\mathcal{O}}\|\mathbb{P}(\boldsymbol{w}_2,\widehat{\boldsymbol{w}}_2)-\mathbb{P}(\boldsymbol{w}_1,\widehat{\boldsymbol{w}}_1)\|_{1,h}|\!|\!|(\boldsymbol{u}_1,\widehat{\boldsymbol{u}}_1)|\!|\!|_{1,h}|\!|\!|(\delta_{\boldsymbol{u}},\delta_{\widehat{\boldsymbol{u}}})|\!|\!|_{1,h}\big)\\&\le C_{\rm HDG}^2\big(C_{\mathcal{O}}C_{{\rm stab},1}|\!|\!|(\boldsymbol{w}_2-\boldsymbol{w}_1,\widehat{\boldsymbol{w}}_2-\widehat{\boldsymbol{w}}_1)|\!|\!|_{1,h}|\!|\!|(\boldsymbol{u}_1,\widehat{\boldsymbol{u}}_1)|\!|\!|_{1,h}|\!|\!|(\delta_{\boldsymbol{u}},\delta_{\widehat{\boldsymbol{u}}})|\!|\!|_{1,h}\big)\end{aligned}$$

where the last three inequalities follow by Proposition 3.6, Proposition 3.4 and Proposition 2.1, respectively. Since $(\boldsymbol{u}_1, \widehat{\boldsymbol{u}}_1) \in K_h$, we obtain

$$\nu \| (\delta_{\boldsymbol{u}}, \delta_{\widehat{\boldsymbol{u}}}) \|_{1,h} \leq C_2 C_{\mathrm{HDG}}^4 C_{\mathcal{O}} C_{\mathrm{stab},1} \nu^{-1} \|\boldsymbol{f}\|_\Omega \| (\boldsymbol{w}_2 - \boldsymbol{w}_1, \widehat{\boldsymbol{w}}_2 - \widehat{\boldsymbol{w}}_1) \|_{1,h}.$$

Therefore, $\mathcal{F}$ is a contraction if

$$\nu^{-2} \|\boldsymbol{f}\|_\Omega < \frac{1}{C_2 C_{\mathrm{HDG}}^4 C_{\mathcal{O}} C_{\mathrm{stab},1}},$$

that is, if $\nu^{-2}\|\boldsymbol{f}\|_\Omega$ is small enough.

Step 4: Conclusion. Since $\mathcal{F}$ is a contraction on K_h, it has a unique fixed point $(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h) \in K_h$ which gives the solution $(\mathrm{L}_h, \boldsymbol{u}_h, p_h, \widehat{\boldsymbol{u}}_h)$ to the problem (3.1) satisfying the bound of Theorem 2.3. This concludes the proof of Theorem 2.3.

6. Proof of the error estimates

In this section, we prove the error estimates of Theorem 2.4. To do that, we proceed in several steps.

Step 1: The error equations. We start our error analysis by obtaining the equations satisfied by the projections of the errors.

Lemma 6.1 (Error equations)**.** *The projection of the error* $(e^L, e^{\boldsymbol{u}}, e^p, e^{\widehat{\boldsymbol{u}}})$ *satisfies*

$$\begin{aligned}
&(e^{\mathrm{L}}, \mathrm{G})_{\mathcal{T}_h} + A_h((e^{\boldsymbol{u}}, e^{\widehat{\boldsymbol{u}}}), \mathrm{G}) - A_h((\boldsymbol{v}, \boldsymbol{\mu}), \nu e^{\mathrm{L}}) - B_h((\boldsymbol{v}, \boldsymbol{\mu}), e^p) + B_h((e^{\boldsymbol{u}}, e^{\widehat{\boldsymbol{u}}}), q) \\
&+ J_n((e^{\boldsymbol{u}}, e^{\widehat{\boldsymbol{u}}}), (\boldsymbol{v}, \boldsymbol{\mu})) = \mathcal{O}_h(\mathbb{P}(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h); (\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h), (\boldsymbol{v}, \boldsymbol{\mu})) - \mathcal{O}_h(\boldsymbol{u}; (\boldsymbol{u}, \boldsymbol{u}|_{\mathcal{E}_h}), (\boldsymbol{v}, \boldsymbol{\mu})) \\
&+ \langle \nu(\mathrm{L} - \Pi_G \mathrm{L})\boldsymbol{n} - (p - \Pi_Q p)\boldsymbol{n}, \boldsymbol{v} - \boldsymbol{\mu} \rangle_{\partial \mathcal{T}_h},
\end{aligned}$$

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

Proof. Note that for any $(\mathrm{G}, \boldsymbol{v}, q, \boldsymbol{\mu}) \in \mathrm{G}_h \times \boldsymbol{V}_h \times Q_h \times \boldsymbol{M}_h$, the solution of the equations (1.1) defining the HDG method satisfies

$$\begin{aligned}
&(\mathrm{L}, \mathrm{G})_{\mathcal{T}_h} + A_h((\boldsymbol{u}, \boldsymbol{u}|_{\mathcal{E}_h}), \mathrm{G}) - A_h((\boldsymbol{v}, \boldsymbol{\mu}), \nu \mathrm{L}) - B_h((\boldsymbol{v}, \boldsymbol{\mu}), p) + B_h((\boldsymbol{u}, \boldsymbol{u}|_{\mathcal{E}_h}), q) \\
&+ J_n((\boldsymbol{u}, \boldsymbol{u}|_{\mathcal{E}_h}), (\boldsymbol{v}, \boldsymbol{\mu})) + \mathcal{O}_h(\boldsymbol{u}; (\boldsymbol{u}, \boldsymbol{u}|_{\mathcal{E}_h}), (\boldsymbol{v}, \boldsymbol{\mu})) = (\boldsymbol{f}, \boldsymbol{v})_{\mathcal{T}_h}.
\end{aligned}$$

By the definition of Π^{RT}, (2.1), and the fact that Π_G and Π_Q are simple L^2-projections, we have that for any $(\mathrm{G}, \boldsymbol{v}, q, \boldsymbol{\mu}) \in \mathrm{G}_h \times \boldsymbol{V}_h \times Q_h \times \boldsymbol{M}_h$,

$$\begin{aligned}
&(\Pi_G \mathrm{L}, \mathrm{G})_{\mathcal{T}_h} + A_h((\Pi^{\mathrm{RT}} \boldsymbol{u}, \Pi_M(\boldsymbol{u}|_{\mathcal{E}_h})), \mathrm{G}) - A_h((\boldsymbol{v}, \boldsymbol{\mu}), \nu \Pi_G \mathrm{L}) \\
&- \langle \nu(\mathrm{L} - \Pi_G \mathrm{L})\boldsymbol{n}, \boldsymbol{v} - \boldsymbol{\mu} \rangle_{\partial \mathcal{T}_h} - B_h((\boldsymbol{v}, \boldsymbol{\mu}), \Pi_Q p) + \langle (p - \Pi_Q p)\boldsymbol{n}, \boldsymbol{v} - \boldsymbol{\mu} \rangle_{\partial \mathcal{T}_h} \\
&+ B_h((\Pi^{\mathrm{RT}} \boldsymbol{u}, \Pi_M(\boldsymbol{u}|_{\mathcal{E}_h})), q) + J_n((\Pi^{\mathrm{RT}} \boldsymbol{u}, \Pi_M(\boldsymbol{u}|_{\mathcal{E}_h})), (\boldsymbol{v}, \boldsymbol{\mu})) \\
&+ \mathcal{O}_h(\boldsymbol{u}; (\boldsymbol{u}, \boldsymbol{u}|_{\mathcal{E}_h}), (\boldsymbol{v}, \boldsymbol{\mu})) = (\boldsymbol{f}, \boldsymbol{v})_{\mathcal{T}_h}.
\end{aligned}$$

Subtracting (3.1) from this equation gives the result. □

Step 2: Estimate of the error in the velocity gradient. In this step, we prove the estimate of the error in the velocity gradient (2.4a). The estimate (2.4b) is an immediate consequence of (2.4a) due to the second discrete Poincaré inequality of Proposition A.2.

First of all, note that by the definition of Π^{RT} in (2.1), $(\Pi_G \mathrm{L}, \Pi^{\mathrm{RT}}\boldsymbol{u}, \Pi_M(\boldsymbol{u}|_{\mathcal{E}_h}))$ satisfies (1.2a), (1.2c). So, $(e^{\mathrm{L}}, e^{\boldsymbol{u}}, e^{\widehat{\boldsymbol{u}}})$ also satisfies (1.2a), (1.2c). Second, since $\boldsymbol{u}$ is divergence free, $\nabla \cdot (\Pi^{\mathrm{RT}}\boldsymbol{u}) = 0$ and this implies that $\Pi^{\mathrm{RT}}\boldsymbol{u}, e^{\boldsymbol{u}} \in \boldsymbol{V}_h$. Thus we take $\mathrm{G} := \nu e^{\mathrm{L}}$, $\boldsymbol{v} := e^{\boldsymbol{u}}$, $q := e^p$ and $\boldsymbol{\mu} := e^{\widehat{\boldsymbol{u}}}$ in the error equations given in Lemma 6.1, to obtain

(6.1)
$$\begin{aligned}\nu\|e^{\mathrm{L}}\|^2_{\mathcal{T}_h} + J_n((e^{\boldsymbol{u}} - e^{\widehat{\boldsymbol{u}}}), (e^{\boldsymbol{u}} - e^{\widehat{\boldsymbol{u}}})) &= \mathcal{O}_h(\mathbb{P}(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h); (\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h), (e^{\boldsymbol{u}}, e^{\widehat{\boldsymbol{u}}}))\\ &\quad - \mathcal{O}_h(\boldsymbol{u}; (\boldsymbol{u}, \boldsymbol{u}|_{\mathcal{E}_h}), (e^{\boldsymbol{u}}, e^{\widehat{\boldsymbol{u}}})) + \langle \nu(\mathrm{L} - \Pi_G \mathrm{L})\boldsymbol{n} - (p - \Pi_Q p)\boldsymbol{n}, e^{\boldsymbol{u}} - e^{\widehat{\boldsymbol{u}}}\rangle_{\partial\mathcal{T}_h}\\ &= T_1 + T_2 + T_3 + T_4 + T_5,\end{aligned}$$

where

$$\begin{aligned}T_1 :=& \mathcal{O}_h(\mathbb{P}(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h); (\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h), (e^{\boldsymbol{u}}, e^{\widehat{\boldsymbol{u}}})) - \mathcal{O}_h(\mathbb{P}(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h); (\Pi^{\mathrm{RT}}\boldsymbol{u}, \Pi_M(\boldsymbol{u}|_{\mathcal{E}_h})), (e^{\boldsymbol{u}}, e^{\widehat{\boldsymbol{u}}})),\\ T_2 :=& \mathcal{O}_h(\mathbb{P}(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h); (\Pi^{\mathrm{RT}}\boldsymbol{u}, \Pi_M(\boldsymbol{u}|_{\mathcal{E}_h})), (e^{\boldsymbol{u}}, e^{\widehat{\boldsymbol{u}}}))\\ &- \mathcal{O}_h(\mathbb{P}(\boldsymbol{u}, \boldsymbol{u}|_{\mathcal{E}_h}); (\Pi^{\mathrm{RT}}\boldsymbol{u}, \Pi_M(\boldsymbol{u}|_{\mathcal{E}_h})), (e^{\boldsymbol{u}}, e^{\widehat{\boldsymbol{u}}})),\\ T_3 :=& \mathcal{O}_h(\mathbb{P}(\boldsymbol{u}, \boldsymbol{u}|_{\mathcal{E}_h}); (\Pi^{\mathrm{RT}}\boldsymbol{u}, \Pi_M(\boldsymbol{u}|_{\mathcal{E}_h})), (e^{\boldsymbol{u}}, e^{\widehat{\boldsymbol{u}}})) - \mathcal{O}_h(\Pi^{\mathrm{RT}}\boldsymbol{u}; (\boldsymbol{u}, \boldsymbol{u}|_{\mathcal{E}_h}), (e^{\boldsymbol{u}}, e^{\widehat{\boldsymbol{u}}})),\\ T_4 :=& \mathcal{O}_h(\Pi^{\mathrm{RT}}\boldsymbol{u}; (\boldsymbol{u}, \boldsymbol{u}|_{\mathcal{E}_h}), (e^{\boldsymbol{u}}, e^{\widehat{\boldsymbol{u}}})) - \mathcal{O}_h(\boldsymbol{u}; (\boldsymbol{u}, \boldsymbol{u}|_{\mathcal{E}_h}), (e^{\boldsymbol{u}}, e^{\widehat{\boldsymbol{u}}})),\\ T_5 :=& \langle \nu(\mathrm{L} - \Pi_G \mathrm{L})\boldsymbol{n} - (p - \Pi_Q p)\boldsymbol{n}, e^{\boldsymbol{u}} - e^{\widehat{\boldsymbol{u}}}\rangle_{\partial\mathcal{T}_h}.\end{aligned}$$

Let us bound the terms T_i, $i = 1, \ldots, 5$. We start with T_1. By the definition of the projection of the errors and Proposition 3.6, we have that

$$T_1 = -\,\mathcal{O}_h(\mathbb{P}(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h); (e^{\boldsymbol{u}}, e^{\widehat{\boldsymbol{u}}}), (e^{\boldsymbol{u}}, e^{\widehat{\boldsymbol{u}}})) \le 0.$$

Let us bound T_2. Applying Proposition 3.4, Proposition 2.1 and linearity of the projection $\mathbb{P}$, respectively, we obtain

$$\begin{aligned}T_2 &\le C_{\mathcal{O}}\, \|\mathbb{P}(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h) - \mathbb{P}(\boldsymbol{u}, \boldsymbol{u}|_{\mathcal{E}_h})\|_{1,h} |\!|\!|(\Pi^{\mathrm{RT}}\boldsymbol{u}, \Pi_M(\boldsymbol{u}|_{\mathcal{E}_h}))|\!|\!|_{1,h}\, |\!|\!|(e^{\boldsymbol{u}}, e^{\widehat{\boldsymbol{u}}})|\!|\!|_{1,h},\\ &= C_{\mathcal{O}}\, \|\mathbb{P}(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h) - \mathbb{P}(\Pi^{\mathrm{RT}}\boldsymbol{u}, \Pi_M(\boldsymbol{u}|_{\mathcal{E}_h}))\|_{1,h} |\!|\!|(\Pi^{\mathrm{RT}}\boldsymbol{u}, \Pi_M(\boldsymbol{u}|_{\mathcal{E}_h}))|\!|\!|_{1,h}\, |\!|\!|(e^{\boldsymbol{u}}, e^{\widehat{\boldsymbol{u}}})|\!|\!|_{1,h},\\ &= C_{\mathcal{O}}\, \|\mathbb{P}(e^{\boldsymbol{u}}, e^{\widehat{\boldsymbol{u}}})\|_{1,h} |\!|\!|(\Pi^{\mathrm{RT}}\boldsymbol{u}, \Pi_M(\boldsymbol{u}|_{\mathcal{E}_h}))|\!|\!|_{1,h}\, |\!|\!|(e^{\boldsymbol{u}}, e^{\widehat{\boldsymbol{u}}})|\!|\!|_{1,h}.\end{aligned}$$

Then

$$\begin{aligned}T_2 &\le C_{\mathcal{O}}\, C_{\mathrm{stab},1}\, |\!|\!|(\Pi^{\mathrm{RT}}\boldsymbol{u}, \Pi_M(\boldsymbol{u}|_{\mathcal{E}_h}))|\!|\!|_{1,h}\, |\!|\!|(e^{\boldsymbol{u}}, e^{\widehat{\boldsymbol{u}}})|\!|\!|^2_{1,h}, && \text{by Proposition 2.1,}\\ &\le C_{\mathcal{O}} C_{\mathrm{stab},1} C_{\mathrm{HDG}} \|\nabla u\|_{L^2(\Omega)} |\!|\!|(e^{\boldsymbol{u}}, e^{\widehat{\boldsymbol{u}}})|\!|\!|^2_{1,h}, && \text{by Lemma A.3,}\\ &\le \frac{1}{2}\nu C^{-2}_{\mathrm{HDG}} |\!|\!|(e^{\boldsymbol{u}}, e^{\widehat{\boldsymbol{u}}})|\!|\!|^2_{1,h},\end{aligned}$$

if the smallness condition

$$\nu^{-1}\|\nabla u\|_{L^2(\Omega)} \le \frac{1}{2\, C_{\mathcal{O}} C_{\mathrm{stab},1} C^3_{\mathrm{HDG}}},$$

is satisfied.

Next we bound the term T_3. Since, by Proposition 2.1(v), $\mathbb{P}(\boldsymbol{u}, \boldsymbol{u}|_{\mathcal{E}_h}) = \Pi^{\mathrm{RT}}\boldsymbol{u}$, we have

$$
\begin{aligned}
T_3 &= \mathcal{O}_h(\Pi^{\mathrm{RT}}\boldsymbol{u}; (\Pi^{\mathrm{RT}}\boldsymbol{u} - \boldsymbol{u}, \Pi_M(\boldsymbol{u}|_{\mathcal{E}_h}) - \boldsymbol{u}|_{\mathcal{E}_h}), (e^{\boldsymbol{u}}, e^{\widehat{\boldsymbol{u}}})) \\
&\leq C^{\infty}_{\mathcal{O},1} \, \|\Pi^{\mathrm{RT}}\boldsymbol{u}\|_{L^\infty(\Omega)} |\!|\!|(\Pi^{\mathrm{RT}}\boldsymbol{u} - \boldsymbol{u}, \Pi_M(\boldsymbol{u}|_{\mathcal{E}_h}) - \boldsymbol{u}|_{\mathcal{E}_h})|\!|\!|_{0,h} |\!|\!|(e^{\boldsymbol{u}}, e^{\widehat{\boldsymbol{u}}})|\!|\!|_{1,h}, \\
&\leq C^{\infty}_{\mathcal{O},1} \, C^{\infty}_{\mathrm{stab}} \, \|\boldsymbol{u}\|_{L^\infty(\Omega)} |\!|\!|(\Pi^{\mathrm{RT}}\boldsymbol{u} - \boldsymbol{u}, \Pi_M(\boldsymbol{u}|_{\mathcal{E}_h}) - \boldsymbol{u}|_{\mathcal{E}_h})|\!|\!|_{0,h} |\!|\!|(e^{\boldsymbol{u}}, e^{\widehat{\boldsymbol{u}}})|\!|\!|_{1,h},
\end{aligned}
$$

where to obtain the last two inequalities we applied Proposition 3.5 and Lemma A.3, respectively.

Now, let us bound T_4. We have, by Proposition 3.5,

$$
T_4 \leq C^{\infty}_{\mathcal{O},2} \, \|\boldsymbol{u}\|_{L^\infty(\Omega)} |\!|\!|(\Pi^{\mathrm{RT}}\boldsymbol{u} - \boldsymbol{u}, \boldsymbol{0})|\!|\!|_{0,h} |\!|\!|(e^{\boldsymbol{u}}, e^{\widehat{\boldsymbol{u}}})|\!|\!|_{1,h}.
$$

Finally, let us bound T_5. We have

$$
\begin{aligned}
T_5 \leq & \Big(\sum_{K\in\mathcal{T}_h} h_K \|\nu(\mathrm{L} - \Pi_G \mathrm{L})\boldsymbol{n}\|^2_{\partial K} \Big)^{1/2} \Big(\sum_{K\in\mathcal{T}_h} h_K^{-1} \|e^{\boldsymbol{u}} - e^{\widehat{\boldsymbol{u}}}\|^2_{\partial K} \Big)^{1/2} \\
& + \Big(\sum_{K\in\mathcal{T}_h} \zeta_n^{-1} h_K \|p - \Pi_Q p\|_{\partial K} \Big)^{1/2} \Big(\sum_{K\in\mathcal{T}_h} \zeta_n h_K^{-1} \|(e^{\boldsymbol{u}} - e^{\widehat{\boldsymbol{u}}}) \cdot \boldsymbol{n}\|^2_{\partial K} \Big)^{1/2} \\
\leq & \Big(\sum_{K\in\mathcal{T}_h} h_K \|\nu(\mathrm{L} - \Pi_G \mathrm{L})\boldsymbol{n}\|^2_{\partial K} \Big)^{1/2} |\!|\!|(e^{\boldsymbol{u}}, e^{\widehat{\boldsymbol{u}}})|\!|\!|_{1,h} \\
& + \Big(\sum_{K\in\mathcal{T}_h} \zeta_n^{-1} h_K \|p - \Pi_Q p\|^2_{\partial K} \Big)^{1/2} \Big(J_n((e^{\boldsymbol{u}} - e^{\widehat{\boldsymbol{u}}}), (e^{\boldsymbol{u}} - e^{\widehat{\boldsymbol{u}}})) \Big)^{1/2}.
\end{aligned}
$$

We are now ready to conclude. Indeed, since, by Proposition 2.2,

$$
\nu C_{\mathrm{HDG}}^{-2} |\!|\!|(e^{\boldsymbol{u}}, e^{\widehat{\boldsymbol{u}}})|\!|\!|^2_{1,h} \leq \nu \|e^{\mathrm{L}}\|^2_{\mathcal{T}_h} + J_n((e^{\boldsymbol{u}} - e^{\widehat{\boldsymbol{u}}}), (e^{\boldsymbol{u}} - e^{\widehat{\boldsymbol{u}}})),
$$

the inequality (6.1) implies that

$$
\begin{aligned}
\big(\nu \|e^{\mathrm{L}}\|^2_{\mathcal{T}_h} & + J_n((e^{\boldsymbol{u}} - e^{\widehat{\boldsymbol{u}}}), (e^{\boldsymbol{u}} - e^{\widehat{\boldsymbol{u}}})) + \nu C_{\mathrm{HDG}}^{-2} |\!|\!|(e^{\boldsymbol{u}}, e^{\widehat{\boldsymbol{u}}})|\!|\!|^2_{1,h}\big)^{1/2} \\
& \leq C \big[C^{\infty}_{\mathcal{O},1} \, C^{\infty}_{\mathrm{stab}} \, \|\boldsymbol{u}\|_{L^\infty(\Omega)} |\!|\!|(\Pi^{\mathrm{RT}}\boldsymbol{u} - \boldsymbol{u}, \Pi_M(\boldsymbol{u}|_{\mathcal{E}_h}) - \boldsymbol{u}|_{\mathcal{E}_h})|\!|\!|_{0,h} \\
& \quad + C^{\infty}_{\mathcal{O},2} \, \|\boldsymbol{u}\|_{L^\infty(\Omega)} |\!|\!|(\Pi^{\mathrm{RT}}\boldsymbol{u} - \boldsymbol{u}, \boldsymbol{0})|\!|\!|_{0,h} \\
& \quad + \Big(\sum_{K\in\mathcal{T}_h} h_K \|\nu(\mathrm{L} - \Pi_G \mathrm{L})\boldsymbol{n}\|^2_{\partial K} + \zeta_n^{-1} h_K \|p - \Pi_Q p\|^2_{\partial K} \Big)^{1/2} \big],
\end{aligned}
$$

and the result now follows by the approximation properties of the Raviart-Thomas projection and those of the L^2-projections Π_G and Π_Q.

Step 3: Estimate of the error in the pressure. In this step, we prove the estimate (2.4c) of the error in pressure.

Since $e^p \in Q_h \subset L^2_0(\Omega)$, we can apply Proposition 3.3 to get that

$$
\|e^p\|_\Omega \leq \frac{C_{\mathrm{V,M}}}{\kappa} \sup_{(\boldsymbol{v},\boldsymbol{\mu})\in \boldsymbol{V}_h \times \boldsymbol{M}_h^0 \setminus\{(\boldsymbol{0},\boldsymbol{0})\}} \frac{B_h((\boldsymbol{v},\boldsymbol{\mu}), e^p)}{|\!|\!|(\boldsymbol{v},\boldsymbol{\mu})|\!|\!|_{1,h}}.
$$

Using the error equation of Lemma 6.1 with $\mathrm{G} := 0$, $q := 0$, $\boldsymbol{v} := P\boldsymbol{w}$, $\boldsymbol{\mu} := \Pi_M \boldsymbol{w}$, we obtain

$$\begin{aligned} B_h((P\boldsymbol{w},\Pi_M \boldsymbol{w}), e^p) =\ & -A_h((P\boldsymbol{w},\Pi_M \boldsymbol{w}), \nu e^{\mathrm{L}}) + J_n((e^{\boldsymbol{u}}, e^{\widehat{\boldsymbol{u}}}), (P\boldsymbol{w},\Pi_M \boldsymbol{w})) \\ & - \mathcal{O}_h(\mathbb{P}(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h); (\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h), (P\boldsymbol{w},\Pi_M \boldsymbol{w})) + \mathcal{O}_h(\boldsymbol{u}; (\boldsymbol{u}, \boldsymbol{u}|_{\mathcal{E}_h}), (P\boldsymbol{w},\Pi_M \boldsymbol{w})) \\ & - \langle \nu(\mathrm{L} - \Pi_G \mathrm{L})\boldsymbol{n} - (p - \Pi_Q p)\boldsymbol{n}, P\boldsymbol{w} - \Pi_M \boldsymbol{w}\rangle_{\mathcal{T}_h} \\ =\ & T_1 + T_2 + T_3 + T_4 + T_5, \end{aligned}$$

where

$$\begin{aligned} T_1 :=\ & -A_h((P\boldsymbol{w},\Pi_M \boldsymbol{w}), \nu e^{\mathrm{L}}), \\ T_2 :=\ & J_n((e^{\boldsymbol{u}}, e^{\widehat{\boldsymbol{u}}}), (P\boldsymbol{w},\Pi_M \boldsymbol{w})), \\ T_3 :=\ & -\mathcal{O}_h(\mathbb{P}(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h); (\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h), (P\boldsymbol{w},\Pi_M \boldsymbol{w})) \\ & + \mathcal{O}_h(\mathbb{P}(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h); (\Pi^{\mathrm{RT}}\boldsymbol{u}, \Pi_M \boldsymbol{u}), (P\boldsymbol{w},\Pi_M \boldsymbol{w})), \\ T_4 :=\ & -\mathcal{O}_h(\mathbb{P}(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h); (\Pi^{\mathrm{RT}}\boldsymbol{u}, \Pi_M \boldsymbol{u}), (P\boldsymbol{w},\Pi_M \boldsymbol{w})) \\ & + \mathcal{O}_h(\boldsymbol{u}; (\Pi^{\mathrm{RT}}\boldsymbol{u}, \Pi_M \boldsymbol{u}), (P\boldsymbol{w},\Pi_M \boldsymbol{w})), \\ T_5 :=\ & -\mathcal{O}_h(\boldsymbol{u}; (\Pi^{\mathrm{RT}}\boldsymbol{u}, \Pi_M \boldsymbol{u}), (P\boldsymbol{w},\Pi_M \boldsymbol{w})) + \mathcal{O}_h(\boldsymbol{u}; (\boldsymbol{u}, \boldsymbol{u}|_{\mathcal{E}_h}), (P\boldsymbol{w},\Pi_M \boldsymbol{w})), \\ T_6 :=\ & -\langle \nu(\mathrm{L} - \Pi_G \mathrm{L})\boldsymbol{n} - (p - \Pi_Q p)\boldsymbol{n}, P\boldsymbol{w} - \Pi_M \boldsymbol{w}\rangle_{\mathcal{T}_h}. \end{aligned}$$

Let us estimate the terms T_i, $i = 1, \dots, 6$. In what follows, we set $(\boldsymbol{v}, \boldsymbol{\mu}) := (P\boldsymbol{w}, \Pi_M \boldsymbol{w})$. We start by bounding T_1. By Proposition 3.1, we have that

$$T_1 \le C_A \nu \, \|e^{\mathrm{L}}\|_\Omega \, |||(\boldsymbol{v}, \boldsymbol{\mu})|||_{1,h}.$$

We claim that $T_2 = 0$. To see this, we only need to show that the expression $T_2(K) := \langle (e^{\boldsymbol{u}} - e^{\widehat{\boldsymbol{u}}}) \cdot \boldsymbol{n}, (P\boldsymbol{w} - \Pi_M \boldsymbol{w}) \cdot \boldsymbol{n}\rangle_{\partial K}$ is zero for any element $K \in \mathcal{T}_h$. Indeed, note that, by [14, Lemma 4.1], for each element $K \in \mathcal{T}_h$, there exists $q \in P_k(K)^\perp$ such that $q|_{\partial K} = (P\boldsymbol{w} - \Pi_M \boldsymbol{w}) \cdot \boldsymbol{n}$. Then, by Lemma 6.1 with $\mathrm{G} := 0$, $\boldsymbol{v} := \boldsymbol{0}$, and $\boldsymbol{\mu} := \boldsymbol{0}$, we have that

$$T_2(K) = \langle (e^{\boldsymbol{u}} - e^{\widehat{\boldsymbol{u}}}) \cdot \boldsymbol{n}, q\rangle_{\partial K} = (\nabla \cdot e^{\boldsymbol{u}}, q)_K = 0,$$

since $\nabla e^{\boldsymbol{u}} \in P_{k-1}(K)$ and $q \in P_k(K)^\perp$. Hence, $T_2 = 0$, as claimed.
Next, we bound T_3. We have, by Proposition 3.4,

$$\begin{aligned} T_3 &= \mathcal{O}_h(\mathbb{P}(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h); (e^{\boldsymbol{u}}, e^{\widehat{\boldsymbol{u}}}), (\boldsymbol{v}, \boldsymbol{\mu})) \\ &\le C_{\mathcal{O}} \|\mathbb{P}(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h)\|_{1,h} |||(e^{\boldsymbol{u}}, e^{\widehat{\boldsymbol{u}}})|||_{1,h} |||(\boldsymbol{v}, \boldsymbol{\mu})|||_{1,h} \\ &\le C_{\mathcal{O}} (\|\mathbb{P}(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h) - \boldsymbol{u}\|_{1,h} + \|\nabla \boldsymbol{u}\|_{L^2(\Omega)}) |||(e^{\boldsymbol{u}}, e^{\widehat{\boldsymbol{u}}})|||_{1,h} |||(\boldsymbol{v}, \boldsymbol{\mu})|||_{1,h}. \end{aligned}$$

To bound T_4, we first use Proposition 3.5 and then Lemma A.3 to get

$$\begin{aligned} T_4 &\le C^\infty_{\mathcal{O},2} |||(\mathbb{P}(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h) - \boldsymbol{u}, \boldsymbol{0})|||_{0,h} |||(\Pi^{\mathrm{RT}}\boldsymbol{u}, \Pi_M \boldsymbol{u})|||_{\infty,h} |||(\boldsymbol{v}, \boldsymbol{\mu})|||_{1,h} \\ &\le C^\infty_{\mathcal{O},2} |||(\mathbb{P}(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h) - \boldsymbol{u}, \boldsymbol{0})|||_{0,h} \|\boldsymbol{u}\|_{L^\infty(\Omega)} |||(\boldsymbol{v}, \boldsymbol{\mu})|||_{1,h}. \end{aligned}$$

Let us now bound T_5. We have, by Proposition 3.5

$$T_5 \le C^\infty_{\mathcal{O},1} \, \|\boldsymbol{u}\|_{L^\infty(\Omega)} |||(\Pi^{\mathrm{RT}}\boldsymbol{u} - \boldsymbol{u}, \Pi_M \boldsymbol{u} - \boldsymbol{u})|||_{0,h} |||(\boldsymbol{v}, \boldsymbol{\mu})|||_{1,h}.$$

Finally, we bound T_6 as follows:

$$T_6 \le C \Big(\sum_{K \in \mathcal{T}_h} h_K \|\nu(\mathrm{L} - \Pi_G \mathrm{L})\boldsymbol{n} - (p - \Pi_Q p)\boldsymbol{n}\|^2_{\partial K} \Big)^{1/2} |||(\boldsymbol{v}, \boldsymbol{\mu})|||_{1,h}.$$

We are now ready to conclude. Indeed, gathering the above estimates, we get that

$$\begin{aligned}\|e^p\|_\Omega \le \frac{C\,C_{\mathrm{V,M}}}{\kappa}\Big(C_A\nu\,\|e^{\mathrm{L}}\|_\Omega + \Big(\sum_{K\in\mathcal{T}_h} h_K\|\nu(\mathrm{L}-\Pi_G\mathrm{L})\boldsymbol{n}-(p-\Pi_Q p)\boldsymbol{n}\|^2_{\partial K}\Big)^{1/2} \\ + C_{\mathcal{O}}(\|\mathbb{P}(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)-\boldsymbol{u}\|_{1,h} + \|\nabla\boldsymbol{u}\|_{L^2(\Omega)})\|\!|(e^{\boldsymbol{u}},e^{\widehat{\boldsymbol{u}}})\|\!|_{1,h} \\ + C^\infty_{\mathcal{O},2}\|\!|(\mathbb{P}(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)-\boldsymbol{u},\boldsymbol{0})\|\!|_{0,h}\|\boldsymbol{u}\|_{L^\infty(\Omega)} \\ + C^\infty_{\mathcal{O},1}\,\|\boldsymbol{u}\|_{L^\infty(\Omega)}\|\!|(\Pi^{\mathrm{RT}}\boldsymbol{u}-\boldsymbol{u},\Pi_M\boldsymbol{u}-\boldsymbol{u})\|\!|_{0,h}\Big).\end{aligned}$$

Since, by property (v) of Proposition 2.1, $\mathbb{P}(\Pi^{\mathrm{RT}}\boldsymbol{u},\Pi_M\boldsymbol{u}) = \Pi^{\mathrm{RT}}\boldsymbol{u}$, we can write, for $i=0,1$, that

$$\begin{aligned}\|\!|(\mathbb{P}(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)-\boldsymbol{u},\boldsymbol{0})\|\!|_{i,h} &\le \|\!|(\mathbb{P}(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)-\mathbb{P}(\Pi^{\mathrm{RT}}\boldsymbol{u},\Pi_M\boldsymbol{u}),\boldsymbol{0})\|\!|_{i,h} \\ &\quad + \|\!|(\Pi^{\mathrm{RT}}\boldsymbol{u}-\boldsymbol{u},\boldsymbol{0})\|\!|_{i,h} \\ &= \|\!|(\mathbb{P}(e^{\boldsymbol{u}},e^{\widehat{\boldsymbol{u}}}),\boldsymbol{0})\|\!|_{i,h} + \|\!|(\Pi^{\mathrm{RT}}\boldsymbol{u}-\boldsymbol{u},\boldsymbol{0})\|\!|_{i,h} \\ &\le C_{\mathrm{stab},i}\,\|\!|(e^{\boldsymbol{u}},e^{\widehat{\boldsymbol{u}}})\|\!|_{i,h} + \|\!|(\Pi^{\mathrm{RT}}\boldsymbol{u}-\boldsymbol{u},\boldsymbol{0})\|\!|_{i,h}.\end{aligned}$$

Thus we get that

$$\begin{aligned}\|e^p\|_\Omega \le& \frac{C\,C_{\mathrm{HDG}}}{\kappa}\Big(\nu\,\|e^{\mathrm{L}}\|_\Omega + \Big(\sum_{K\in\mathcal{T}_h} h_K\|\nu(\mathrm{L}-\Pi_G\mathrm{L})\boldsymbol{n}-(p-\Pi_Q p)\boldsymbol{n}\|^2_{\partial K}\Big)^{1/2} \\ &+ C_{\mathcal{O}}(C_{\mathrm{stab},1}\,\|\!|(e^{\boldsymbol{u}},e^{\widehat{\boldsymbol{u}}})\|\!|_{1,h} + \|\!|\Pi^{\mathrm{RT}}\boldsymbol{u}-\boldsymbol{u}\|\!|_{1,h} + \|\nabla\boldsymbol{u}\|_{L^2(\Omega)})\|\!|(e^{\boldsymbol{u}},e^{\widehat{\boldsymbol{u}}})\|\!|_{1,h} \\ &+ C^\infty_{\mathcal{O},2}(C_{\mathrm{stab},0}\,\|\!|(e^{\boldsymbol{u}},e^{\widehat{\boldsymbol{u}}})\|\!|_{0,h} + \|\!|\Pi^{\mathrm{RT}}\boldsymbol{u}-\boldsymbol{u}\|\!|_{0,h})\|\boldsymbol{u}\|_{L^\infty(\Omega)} \\ &+ C^\infty_{\mathcal{O},1}\,\|\boldsymbol{u}\|_{L^\infty(\Omega)}\|\!|(\Pi^{\mathrm{RT}}\boldsymbol{u}-\boldsymbol{u},\Pi_M\boldsymbol{u}-\boldsymbol{u})\|\!|_{0,h}\Big),\end{aligned}$$

and the result follows by the approximation properties of $\Pi_{\mathrm{G}},\Pi_Q,\Pi_M$ and Π^{RT}.

Step 4: A duality argument. To estimate the L^2-error in the velocity, we use a duality argument in which we are going to use the solution of the problem (2.2). We have the following result.

Lemma 6.2. *Assume that the regularity estimate* (2.3) *holds so that* $\boldsymbol{\phi}$ *lies in* $\boldsymbol{H}^2(\Omega)$ *whenever* $\boldsymbol{\theta}\in\boldsymbol{L}^2(\Omega)$. *Then, we have that*

$$(e^{\boldsymbol{u}},\boldsymbol{\theta})_{\mathcal{T}_h} = T_1+\cdots+T_6,$$

where

$$\begin{aligned}T_1 &:= -\langle e^{\boldsymbol{u}}-e^{\widehat{\boldsymbol{u}}},\nu\delta_\Phi\boldsymbol{n}+\delta_\psi\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} \\ T_2 &:= \langle\nu(\mathrm{L}-\Pi_G\mathrm{L})\boldsymbol{n},\Pi^{\mathrm{RT}}\boldsymbol{\phi}-\Pi_M\boldsymbol{\phi}\rangle_{\partial\mathcal{T}_h} \\ T_3 &:= -\big((e^{\boldsymbol{u}},\nabla\cdot(\boldsymbol{\phi}\otimes\boldsymbol{u}))_{\mathcal{T}_h} + \mathcal{O}_h(\boldsymbol{u};(e^{\boldsymbol{u}},e^{\widehat{\boldsymbol{u}}}),(\Pi^{\mathrm{RT}}\boldsymbol{\phi},\Pi_M\boldsymbol{\phi}))\big) \\ T_4 &:= -\mathcal{O}_h(\boldsymbol{u};(\delta_{\boldsymbol{u}},\delta_{\widehat{\boldsymbol{u}}}),(\Pi^{\mathrm{RT}}\boldsymbol{\phi},\Pi_M\boldsymbol{\phi})) \\ T_5 &:= \big(\mathcal{O}_h(\boldsymbol{u};(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h),(\delta_{\boldsymbol{\phi}},\delta_{\widehat{\boldsymbol{u}}})) - \mathcal{O}_h(\mathbb{P}(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h);(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h),(\delta_{\boldsymbol{\phi}},\delta_{\widehat{\boldsymbol{u}}}))\big) \\ T_6 &:= \big(\mathcal{O}_h(\mathbb{P}(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h);(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h),(\boldsymbol{\phi},\boldsymbol{\phi}|_{\mathcal{E}_h})) - \mathcal{O}_h(\boldsymbol{u};(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h),(\boldsymbol{\phi},\boldsymbol{\phi}|_{\mathcal{E}_h})) \\ &\quad - (e^{\boldsymbol{u}},(\nabla\boldsymbol{\phi})^\top\boldsymbol{u})_{\mathcal{T}_h}\big),\end{aligned}$$

and $\delta_\Phi := \Phi-\Pi_G\Phi,\ \delta_{\boldsymbol{\phi}} := \boldsymbol{\phi}-\Pi^{\mathrm{RT}}\boldsymbol{\phi},\ \delta_\psi := \psi-\Pi_Q\psi$, *and* $\delta_{\widehat{\boldsymbol{\phi}}} := \boldsymbol{\phi}-\Pi_M\boldsymbol{\phi}$.

Proof. By the first three equations defining the dual problem, (2.2a), (2.2b) and (2.2c), we have

$$\begin{aligned}(e^{\boldsymbol{u}}, \boldsymbol{\theta})_{\mathcal{T}_h} =& - (e^{\boldsymbol{u}}, \nu\nabla \cdot \Phi)_{\mathcal{T}_h} - (e^{\boldsymbol{u}}, \nabla \cdot (\boldsymbol{\phi} \otimes \boldsymbol{u}))_{\mathcal{T}_h} - (e^{\boldsymbol{u}}, (\nabla \boldsymbol{\phi})^\top \boldsymbol{u})_{\mathcal{T}_h} - (e^{\boldsymbol{u}}, \nabla\psi)_{\mathcal{T}_h}\\ &- (\nu e^{\mathrm{L}}, \Phi)_{\mathcal{T}_h} + (\nu e^{\mathrm{L}}, \nabla\boldsymbol{\phi})_{\mathcal{T}_h}\\ &- (e^p, \nabla\cdot\boldsymbol{\phi})_{\mathcal{T}_h}\\ =& - (e^{\boldsymbol{u}}, \nu\nabla \cdot \Phi)_{\mathcal{T}_h} - (\nu e^{\mathrm{L}}, \Phi)_{\mathcal{T}_h}\\ &- (e^{\boldsymbol{u}}, \nabla\psi)_{\mathcal{T}_h}\\ &+ (\nu e^{\mathrm{L}}, \nabla\boldsymbol{\phi})_{\mathcal{T}_h} - (e^{\boldsymbol{u}}, \nabla \cdot (\boldsymbol{\phi} \otimes \boldsymbol{u}))_{\mathcal{T}_h} - (e^p, \nabla\cdot\boldsymbol{\phi})_{\mathcal{T}_h} - (e^{\boldsymbol{u}}, (\nabla \boldsymbol{\phi})^\top \boldsymbol{u})_{\mathcal{T}_h}\\ =& - (e^{\boldsymbol{u}}, \nu\nabla \cdot \Pi_G\Phi)_{\mathcal{T}_h} - (\nu e^{\mathrm{L}}, \Pi_G\Phi)_{\mathcal{T}_h} - (e^{\boldsymbol{u}}, \nu\nabla\cdot\delta_\Phi)_{\mathcal{T}_h}\\ &- (e^{\boldsymbol{u}}, \nabla\Pi_Q\psi)_{\mathcal{T}_h} - (e^{\boldsymbol{u}}, \nabla\delta_\psi)_{\mathcal{T}_h} + I,\end{aligned}$$

where $I := (\nu e^{\mathrm{L}}, \nabla\boldsymbol{\phi})_{\mathcal{T}_h} - (e^{\boldsymbol{u}}, \nabla \cdot (\boldsymbol{\phi} \otimes \boldsymbol{u}))_{\mathcal{T}_h} - (e^p, \nabla\cdot\boldsymbol{\phi})_{\mathcal{T}_h} - (e^{\boldsymbol{u}}, (\nabla \boldsymbol{\phi})^\top \boldsymbol{u})_{\mathcal{T}_h}$. If we now set $(\mathrm{G}, \boldsymbol{v}, q, \boldsymbol{\mu})$ equal to $(\nu\Pi_G\Phi, \boldsymbol{0}, -\Pi_Q\psi, \boldsymbol{0})$ in the error equation of Lemma 6.1, we see that we now have

$$\begin{aligned}(e^{\boldsymbol{u}}, \boldsymbol{\theta})_{\mathcal{T}_h} =& - \langle e^{\widehat{\boldsymbol{u}}}, \nu\Pi_G\Phi\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - (e^{\boldsymbol{u}}, \nu\nabla\cdot\delta_\Phi)_{\mathcal{T}_h} - \langle e^{\widehat{\boldsymbol{u}}}, \Pi_Q\psi\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - (e^{\boldsymbol{u}}, \nabla\delta_\psi)_{\mathcal{T}_h} + I\\ =& - \langle e^{\widehat{\boldsymbol{u}}}, \nu\Pi_G\Phi\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle e^{\boldsymbol{u}}, \nu\delta_\Phi\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle e^{\widehat{\boldsymbol{u}}}, \Pi_Q\psi\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle e^{\boldsymbol{u}}, \delta_\psi\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} + I\\ =& - \langle e^{\boldsymbol{u}} - e^{\widehat{\boldsymbol{u}}}, \nu\delta_\Phi\boldsymbol{n} + \delta_\psi\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} + I.\end{aligned}$$

Let us now work on the term I. By the definition of the projection Π^{RT}, we have

$$\begin{aligned}I =& (\nu e^{\mathrm{L}}, \nabla\Pi^{\mathrm{RT}}\boldsymbol{\phi})_{\mathcal{T}_h} - (e^p, \nabla\cdot\Pi^{\mathrm{RT}}\boldsymbol{\phi})_{\mathcal{T}_h}\\ &- (e^{\boldsymbol{u}}, \nabla \cdot (\boldsymbol{\phi} \otimes \boldsymbol{u}))_{\mathcal{T}_h} - (e^{\boldsymbol{u}}, (\nabla \boldsymbol{\phi})^\top \boldsymbol{u})_{\mathcal{T}_h} + (\nu e^{\mathrm{L}}, \nabla\delta_{\boldsymbol{\phi}})_{\mathcal{T}_h}.\end{aligned}$$

Now, since $\nabla\cdot\boldsymbol{\phi} = 0$, we have that $\nabla\cdot\Pi^{\mathrm{RT}}\boldsymbol{\phi} = 0$ and so $\Pi^{\mathrm{RT}}\boldsymbol{\phi} \in \boldsymbol{V}_h$. This means that we can take $(\mathrm{G}, \boldsymbol{v}, q, \boldsymbol{\mu}) := (0, \Pi^{\mathrm{RT}}\boldsymbol{\phi}, 0, \Pi_M\boldsymbol{\phi})$ in the error equation of Lemma 6.1, to get

$$\begin{aligned}I =& \langle \nu e^{\mathrm{L}}\boldsymbol{n}, \Pi^{\mathrm{RT}}\boldsymbol{\phi} - \Pi_M\boldsymbol{\phi}\rangle_{\partial\mathcal{T}_h} - \langle e^p, (\Pi^{\mathrm{RT}}\boldsymbol{\phi} - \Pi_M\boldsymbol{\phi})\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}\\ &- \Sigma_{K\in\mathcal{T}_h} h_K^{-1}\langle (e^{\boldsymbol{u}} - e^{\widehat{\boldsymbol{u}}})\cdot\boldsymbol{n}, (\Pi^{\mathrm{RT}}\boldsymbol{\phi} - \Pi_M\boldsymbol{\phi})\cdot\boldsymbol{n}\rangle_{\partial K}\\ &+ \mathcal{O}_h(\mathbb{P}(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h); (\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h), (\Pi^{\mathrm{RT}}\boldsymbol{\phi}, \Pi_M\boldsymbol{\phi})) - \mathcal{O}_h(\boldsymbol{u}; (\boldsymbol{u}, \boldsymbol{u}|_{\mathcal{E}_h}), (\Pi^{\mathrm{RT}}\boldsymbol{\phi}, \Pi_M\boldsymbol{\phi}))\\ &+ \langle \nu(\mathrm{L} - \Pi_G\mathrm{L})\boldsymbol{n} - (p - \Pi_Q p)\boldsymbol{n}, \Pi^{\mathrm{RT}}\boldsymbol{\phi} - \Pi_M\boldsymbol{\phi}\rangle_{\partial\mathcal{T}_h}\\ &- (e^{\boldsymbol{u}}, \nabla \cdot (\boldsymbol{\phi} \otimes \boldsymbol{u}))_{\mathcal{T}_h} - (e^{\boldsymbol{u}}, (\nabla \boldsymbol{\phi})^\top \boldsymbol{u})_{\mathcal{T}_h} + (\nu e^{\mathrm{L}}, \nabla\delta_{\boldsymbol{\phi}})_{\mathcal{T}_h}\\ =& \langle \nu(\mathrm{L} - \Pi_G\mathrm{L})\boldsymbol{n}, \Pi^{\mathrm{RT}}\boldsymbol{\phi} - \Pi_M\boldsymbol{\phi}\rangle_{\partial\mathcal{T}_h}\\ &+ \mathcal{O}_h(\mathbb{P}(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h); (\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h), (\Pi^{\mathrm{RT}}\boldsymbol{\phi}, \Pi_M\boldsymbol{\phi})) - \mathcal{O}_h(\boldsymbol{u}; (\boldsymbol{u}, \boldsymbol{u}|_{\mathcal{E}_h}), (\Pi^{\mathrm{RT}}\boldsymbol{\phi}, \Pi_M\boldsymbol{\phi}))\\ &- (e^{\boldsymbol{u}}, \nabla \cdot (\boldsymbol{\phi} \otimes \boldsymbol{u}))_{\mathcal{T}_h} - (e^{\boldsymbol{u}}, (\nabla \boldsymbol{\phi})^\top \boldsymbol{u})_{\mathcal{T}_h}.\end{aligned}$$

So we have

$$\begin{aligned}(e^{\boldsymbol{u}}, \boldsymbol{\theta})_{\mathcal{T}_h} =& - \langle e^{\boldsymbol{u}} - e^{\widehat{\boldsymbol{u}}}, \nu\delta_\Phi\boldsymbol{n} + \delta_\psi\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} + \langle \nu(\mathrm{L} - \Pi_G\mathrm{L})\boldsymbol{n}, \Pi^{\mathrm{RT}}\boldsymbol{\phi} - \Pi_M\boldsymbol{\phi}\rangle_{\partial\mathcal{T}_h}\\ &+ \mathcal{O}_h(\mathbb{P}(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h); (\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h), (\Pi^{\mathrm{RT}}\boldsymbol{\phi}, \Pi_M\boldsymbol{\phi}))\\ &- \mathcal{O}_h(\boldsymbol{u}; (\boldsymbol{u}, \boldsymbol{u}|_{\mathcal{E}_h}), (\Pi^{\mathrm{RT}}\boldsymbol{\phi}, \Pi_M\boldsymbol{\phi}))\\ &- (e^{\boldsymbol{u}}, \nabla \cdot (\boldsymbol{\phi} \otimes \boldsymbol{u}))_{\mathcal{T}_h} - (e^{\boldsymbol{u}}, (\nabla \boldsymbol{\phi})^\top \boldsymbol{u})_{\mathcal{T}_h},\end{aligned}$$

and the identity follows after simple algebraic manipulations. This completes the proof. $\square$

Step 5: Superconvergence of the velocity. Now, we are ready to prove the superconvergence estimate (2.5). We need to bound each term $T_i, i = 1, \ldots, 6$ in Lemma 6.2. It is easy to see that

$$
\begin{aligned}
T_1 &\leq Ch\|(e^{\boldsymbol{u}}, e^{\widehat{\boldsymbol{u}}})\|_{1,h}\big(\|\Phi\|_{H^1(\Omega)} + \|\psi\|_{H^1(\Omega)}\big) \\
&\leq Ch\|(e^{\boldsymbol{u}}, e^{\widehat{\boldsymbol{u}}})\|_{1,h}\|\boldsymbol{\theta}\|_{\Omega} && \text{by the regularity inequality (2.3),} \\
T_2 &\leq Ch^{k+2}\|\boldsymbol{u}\|_{H^{k+2}(\Omega)}\|\boldsymbol{\phi}\|_{H^2(\Omega)} && \text{by property (iii) of Lemma A.3,} \\
&\leq Ch^{k+2}\|\boldsymbol{u}\|_{H^{k+2}(\Omega)}\|\boldsymbol{\theta}\|_{\Omega} && \text{by the regularity inequality (2.3),} \\
T_3 &= -\big((e^{\boldsymbol{u}}, \nabla\cdot(\boldsymbol{\phi}\otimes\boldsymbol{u}))_{\mathcal{T}_h} + \mathcal{O}_h(\boldsymbol{u}; (e^{\boldsymbol{u}}, e^{\widehat{\boldsymbol{u}}}), (\Pi^{\mathrm{RT}}\boldsymbol{\phi}, \Pi_M\boldsymbol{\phi}))\big) \\
&= (e^{\boldsymbol{u}}\otimes\boldsymbol{u}, \nabla(\Pi^{\mathrm{RT}}\boldsymbol{\phi} - \boldsymbol{\phi}))_{\mathcal{T}_h} - \langle (e^{\widehat{\boldsymbol{u}}}\otimes\boldsymbol{u})\boldsymbol{n}, \Pi^{\mathrm{RT}}\boldsymbol{\phi} - \Pi_M\boldsymbol{\phi}\rangle_{\partial\mathcal{T}_h} \\
&\qquad - \langle \max(\boldsymbol{u}\cdot\boldsymbol{n}, 0)(e^{\boldsymbol{u}} - e^{\widehat{\boldsymbol{u}}}), \Pi^{\mathrm{RT}}\boldsymbol{\phi} - \Pi_M\boldsymbol{\phi}\rangle_{\partial\mathcal{T}_h} \\
&\leq Ch\|\boldsymbol{u}\|_{L^\infty(\Omega)}\|(e^{\boldsymbol{u}}, e^{\widehat{\boldsymbol{u}}})\|_{1,h}\|\boldsymbol{\phi}\|_{H^2(\Omega)} && \text{by property (iii) of Lemma A.3,} \\
&\leq Ch\|\boldsymbol{u}\|_{L^\infty(\Omega)}\|(e^{\boldsymbol{u}}, e^{\widehat{\boldsymbol{u}}})\|_{1,h}\|\boldsymbol{\theta}\|_{\Omega},
\end{aligned}
$$

by the regularity inequality (2.3).

In order to bound T_4, we define

$$
\overline{\boldsymbol{u}}_h|_K := \frac{1}{|K|}(\boldsymbol{u}, 1)_K \quad \forall K \in \mathcal{T}_h. \tag{6.2}
$$

Then, we have

$$
\begin{aligned}
T_4 &= -\mathcal{O}_h(\boldsymbol{u}; (\delta_{\boldsymbol{u}}, \delta_{\widehat{\boldsymbol{u}}}), (\Pi^{\mathrm{RT}}\boldsymbol{\phi}, \Pi_M\boldsymbol{\phi})) \\
&= (\delta_{\boldsymbol{u}}\otimes\boldsymbol{u}, \nabla\Pi^{\mathrm{RT}}\boldsymbol{\phi})_{\mathcal{T}_h} - \langle (\delta_{\widehat{\boldsymbol{u}}}\otimes\boldsymbol{u})\boldsymbol{n}, \Pi^{\mathrm{RT}}\boldsymbol{\phi} - \Pi_M\boldsymbol{\phi}\rangle_{\partial\mathcal{T}_h} \\
&\qquad - \langle \max(\boldsymbol{u}\cdot\boldsymbol{n}, 0)(\delta_{\boldsymbol{u}} - \delta_{\widehat{\boldsymbol{u}}}), \Pi^{\mathrm{RT}}\boldsymbol{\phi} - \Pi_M\boldsymbol{\phi}\rangle_{\partial\mathcal{T}_h} \\
&= (\delta_{\boldsymbol{u}}\otimes(\boldsymbol{u} - \overline{\boldsymbol{u}}_h), \nabla\Pi^{\mathrm{RT}}\boldsymbol{\phi})_{\mathcal{T}_h} - \langle (\delta_{\widehat{\boldsymbol{u}}}\otimes\boldsymbol{u})\boldsymbol{n}, \Pi^{\mathrm{RT}}\boldsymbol{\phi} - \Pi_M\boldsymbol{\phi}\rangle_{\partial\mathcal{T}_h} \\
&\qquad - \langle \max(\boldsymbol{u}\cdot\boldsymbol{n}, 0)(\delta_{\boldsymbol{u}} - \delta_{\widehat{\boldsymbol{u}}}), \Pi^{\mathrm{RT}}\boldsymbol{\phi} - \Pi_M\boldsymbol{\phi}\rangle_{\partial\mathcal{T}_h}
\end{aligned}
$$

since $\Pi^{\mathrm{RT}}\boldsymbol{\phi} \in \boldsymbol{V}_h$ because $\nabla\cdot\boldsymbol{\phi} = 0$. So, we get

$$
\begin{aligned}
T_4 &\leq Ch^{k+2}\|\boldsymbol{u}\|_{W^{1,\infty}(\Omega)}\|\boldsymbol{\phi}\|_{H^2(\Omega)} \quad \text{by Lemma A.3,} \\
&\leq Ch^{k+2}\|\boldsymbol{u}\|_{W^{1,\infty}(\Omega)}\|\boldsymbol{\theta}\|_{\Omega},
\end{aligned}
$$

by the regularity inequality (2.3).

Let us now bound T_5. We begin by rewriting this term as $T_5 = T_{51} + T_{52} + T_{53}$, where

$$
\begin{aligned}
T_{51} &:= \big(\mathcal{O}_h(\boldsymbol{u}; (e^{\boldsymbol{u}}, e^{\widehat{\boldsymbol{u}}}), (\delta_{\boldsymbol{\phi}}, \delta_{\widehat{\boldsymbol{u}}})) - \mathcal{O}_h(\mathbb{P}(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h); (e^{\boldsymbol{u}}, e^{\widehat{\boldsymbol{u}}}), (\delta_{\boldsymbol{\phi}}, \delta_{\widehat{\boldsymbol{u}}}))\big), \\
T_{52} &:= -\big(\mathcal{O}_h(\boldsymbol{u}; (\delta_{\boldsymbol{u}}, \delta_{\widehat{\boldsymbol{u}}}), (\delta_{\boldsymbol{\phi}}, \delta_{\widehat{\boldsymbol{u}}})) - \mathcal{O}_h(\mathbb{P}(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h); (\delta_{\boldsymbol{u}}, \delta_{\widehat{\boldsymbol{u}}}), (\delta_{\boldsymbol{\phi}}, \delta_{\widehat{\boldsymbol{u}}}))\big), \\
T_{53} &:= \big(\mathcal{O}_h(\boldsymbol{u}; (\boldsymbol{u}, \boldsymbol{u}|_{\mathcal{E}_h}), (\delta_{\boldsymbol{\phi}}, \delta_{\widehat{\boldsymbol{u}}})) - \mathcal{O}_h(\mathbb{P}(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h); (\boldsymbol{u}, \boldsymbol{u}|_{\mathcal{E}_h}), (\delta_{\boldsymbol{\phi}}, \delta_{\widehat{\boldsymbol{u}}}))\big).
\end{aligned}
$$

According to Proposition 3.4, we have

$$
\begin{aligned}
T_{51} &\leq C\|\boldsymbol{u} - \mathbb{P}(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h)\|_{1,h}\|(e^{\boldsymbol{u}}, e^{\widehat{\boldsymbol{u}}})\|_{1,h}\|(\delta_{\boldsymbol{\phi}}, \delta_{\widehat{\boldsymbol{u}}})\|_{1,h} \leq Ch^{2k+2}\|\boldsymbol{\theta}\|_{\Omega}, \\
T_{52} &\leq C\|\boldsymbol{u} - \mathbb{P}(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h)\|_{1,h}\|(\delta_{\boldsymbol{u}}, \delta_{\widehat{\boldsymbol{u}}})\|_{1,h}\|(\delta_{\boldsymbol{\phi}}, \delta_{\widehat{\boldsymbol{u}}})\|_{1,h} \leq Ch^{2k+1}\|\boldsymbol{\theta}\|_{\Omega},
\end{aligned}
$$

and

$$\begin{aligned}T_{53} =& - (\boldsymbol{u}\otimes(\boldsymbol{u}-\mathbb{P}(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)),\nabla\delta_{\boldsymbol{\phi}})_{\mathcal{T}_h} + \langle(\boldsymbol{u}\otimes(\boldsymbol{u}-\mathbb{P}(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)))\boldsymbol{n},\delta_{\boldsymbol{u}}-\delta_{\widehat{\boldsymbol{u}}}\rangle_{\partial\mathcal{T}_h}\\ \le&\ C\,h^{k+2}\|\boldsymbol{u}\|_{L^\infty(\Omega)}\|\boldsymbol{\theta}\|_\Omega.\end{aligned}$$

So, we have

$$T_5 \le C(h^{2k+1}+h^{k+2}\|\boldsymbol{u}\|_{L^\infty(\Omega)})\|\boldsymbol{\theta}\|_\Omega \le Ch^{k+2}(1+\|\boldsymbol{u}\|_{L^\infty(\Omega)})\|\boldsymbol{\theta}\|_\Omega, \quad k\ge 1.$$

Finally, let us estimate T_6. Since $(e^{\boldsymbol{u}},(\nabla\boldsymbol{\phi})^\top\boldsymbol{u})_{\mathcal{T}_h} = (\boldsymbol{u}\otimes e^{\boldsymbol{u}},\nabla\boldsymbol{\phi})_{\mathcal{T}_h}$, we can write that $T_6 = T_{61}+T_{62}+T_{63}$, where

$$\begin{aligned}T_{61} :=& (\boldsymbol{u}\otimes(\boldsymbol{u}_h-\mathbb{P}(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h),\nabla\boldsymbol{\phi})_{\mathcal{T}_h}\\ T_{62} :=& (\boldsymbol{u}\otimes(\boldsymbol{u}-\Pi^{\rm RT}\boldsymbol{u}),\nabla\boldsymbol{\phi})_{\mathcal{T}_h}\\ T_{63} :=& ((\boldsymbol{u}-\boldsymbol{u}_h)\otimes(\mathbb{P}(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)-\boldsymbol{u}),\nabla\boldsymbol{\phi})_{\mathcal{T}_h}.\end{aligned}$$

Simple algebraic manipulations give us that

$$\begin{aligned}T_{61} &= (\boldsymbol{u}\otimes(\boldsymbol{u}_h-\mathbb{P}(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h),\nabla\boldsymbol{\phi})_{\mathcal{T}_h}\\ &= ((\boldsymbol{u}-\overline{\boldsymbol{u}}_h)\otimes(\boldsymbol{u}_h-\mathbb{P}(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)),\nabla\boldsymbol{\phi})_{\mathcal{T}_h} + (\overline{\boldsymbol{u}}_h\otimes(\boldsymbol{u}_h-\mathbb{P}(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)),\nabla\boldsymbol{\phi})_{\mathcal{T}_h}\\ &= ((\boldsymbol{u}-\overline{\boldsymbol{u}}_h)\otimes(\boldsymbol{u}_h-\mathbb{P}(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)),\nabla\boldsymbol{\phi})_{\mathcal{T}_h} + (\overline{\boldsymbol{u}}_h\otimes(\boldsymbol{u}_h-\mathbb{P}(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)),\nabla(\boldsymbol{\phi}-\Pi^{\rm RT}\boldsymbol{\phi}))_{\mathcal{T}_h}\\ &\quad + (\overline{\boldsymbol{u}}_h\otimes(\boldsymbol{u}_h-\mathbb{P}(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)),\nabla\Pi^{\rm RT}\boldsymbol{\phi})_{\mathcal{T}_h}\\ &= ((\boldsymbol{u}-\overline{\boldsymbol{u}}_h)\otimes(\boldsymbol{u}_h-\mathbb{P}(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)),\nabla\boldsymbol{\phi})_{\mathcal{T}_h} + (\overline{\boldsymbol{u}}_h\otimes(\boldsymbol{u}_h-\mathbb{P}(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)),\nabla(\boldsymbol{\phi}-\Pi^{\rm RT}\boldsymbol{\phi}))_{\mathcal{T}_h}.\end{aligned}$$

Indeed, since $\nabla\cdot\boldsymbol{u}=0$, then $\nabla\Pi^{\rm RT}\boldsymbol{u}\in\boldsymbol{P}_{k-1}(\mathcal{T}_h)$ and, according to the definition of $\mathbb{P}(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)$ in (1.3), we have that $(\overline{\boldsymbol{u}}_h\otimes(\boldsymbol{u}_h-\mathbb{P}(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h),\nabla\Pi^{\rm RT}\boldsymbol{u})_{\mathcal{T}_h}=0$. By the definition of $\overline{\boldsymbol{u}}_h$ in (6.2), approximation property of $\Pi^{\rm RT}$, Proposition 2.1 and Theorem 2.4, we get

$$T_{61}\le Ch^{k+2}\|\boldsymbol{\theta}\|_\Omega \quad \forall k\ge 1.$$

The following estimate follows easily:

$$T_{62}\le Ch^{k+2}\|\boldsymbol{\theta}\|_\Omega \quad \forall k\ge 1.$$

The estimate of the term T_{63} is more delicate. Indeed, by Proposition A.2, Proposition 2.1 and Theorem 2.4, we get

$$\begin{aligned}T_{63} &= (e^{\boldsymbol{u}}\otimes(\mathbb{P}(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)-\boldsymbol{u}),\nabla\boldsymbol{\phi})_{\mathcal{T}_h} + ((\boldsymbol{u}-\Pi^{\rm RT}\boldsymbol{u})\otimes(\mathbb{P}(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)-\boldsymbol{u}),\nabla\boldsymbol{\phi})_{\mathcal{T}_h}\\ &\le C\big(\|e^{\boldsymbol{u}}\|_{L^4(\Omega)}\|\mathbb{P}(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)-\boldsymbol{u}\|_{L^4(\Omega)} + \|\boldsymbol{u}-\Pi^{\rm RT}\boldsymbol{u}\|_{L^\infty(\Omega)}\|\mathbb{P}(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)-\boldsymbol{u}\|_{\mathcal{T}_h}\big)\|\nabla\boldsymbol{\phi}\|_{\mathcal{T}_h}\\ &\le C\big(\|e^{\boldsymbol{u}}\|_{1,h}\|\mathbb{P}(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)-\boldsymbol{u}\|_{1,h} + \|\boldsymbol{u}-\Pi^{\rm RT}\boldsymbol{u}\|_{L^\infty(\Omega)}\|\mathbb{P}(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)-\boldsymbol{u}\|_{\mathcal{T}_h}\big)\|\nabla\boldsymbol{\phi}\|_{\mathcal{T}_h}\\ &\le C(h^{2k+1}+h^{k+2})\|\boldsymbol{\phi}\|_{1,\Omega}\le Ch^{k+2}\|\boldsymbol{\theta}\|_\Omega \quad \forall k\ge 1.\end{aligned}$$

This implies that

$$T_6\le Ch^{k+2}\|\boldsymbol{\theta}\|_\Omega.$$

With all the above estimates, and after taking $\boldsymbol{\theta}=e^{\boldsymbol{u}}$, we can conclude that the superconvergence estimate (2.5) of Theorem 2.4 does hold. To complete the proof of Theorem 2.4, it remains to prove the estimate for the post processed velocity. For a proof using the post processing operator defined in [9, (2.9)] to obtain the superconvergence result, see [9, Theorem 2.5].

7. Extension to $H(\text{div})$-conforming HDG methods

We have seen that all our error estimates provide the same orders of convergence as the stabilization parameter of the jumps of the normal component of the velocity goes to infinity. This means that we can work with $H(\text{div})$-conforming spaces for the velocity exactly as it was done for the HDG methods for the Stokes equations in [14].

Appendix A. Stability estimates

A.1. An estimate of the trace of the nonlinear term.

Lemma A.1. *For any $\boldsymbol{v}, \boldsymbol{\mu} \in \boldsymbol{L}^2(\partial\mathcal{T}_h)$ and $\boldsymbol{\beta} \in \boldsymbol{L}^4(\partial\mathcal{T}_h)$, there exists a constant $C > 0$ such that*

$$\langle (\boldsymbol{v} \otimes \boldsymbol{\beta})\boldsymbol{n}, \boldsymbol{\mu}\rangle_{\partial\mathcal{T}_h} \leq C\Big(\sum_{K\in\mathcal{T}_h} h_K^{\frac{2-d}{2}}\|\boldsymbol{v}\|_{\partial K}^2\Big)^{\frac{1}{2}}\Big(\sum_{K\in\mathcal{T}_h} h_K\|\boldsymbol{\beta}\|_{L^4(\partial K)}^4\Big)^{\frac{1}{4}}\Big(\sum_{K\in\mathcal{T}_h} h_K^{-1}\|\boldsymbol{\mu}\|_{\partial K}^2\Big)^{\frac{1}{2}}.$$

Proof. By the generalized Hölder's inequality,

$$\begin{aligned}
\langle (\boldsymbol{v} \otimes \boldsymbol{\beta})\boldsymbol{n}, \boldsymbol{\mu}\rangle_{\partial\mathcal{T}_h} &= \sum_{K\in\mathcal{T}_h} \langle (\boldsymbol{v} \otimes \boldsymbol{\beta})\boldsymbol{n}, \boldsymbol{\mu}\rangle_{\partial K} \\
&\leq \sum_{K\in\mathcal{T}_h} \|\boldsymbol{v}\|_{\partial K}\|\boldsymbol{\beta}\|_{L^4(\partial K)}\|\boldsymbol{\mu}\|_{L^4(\partial K)} \\
&\leq \sum_{K\in\mathcal{T}_h} h_K^{\frac{2-d}{4}}\|\boldsymbol{v}\|_{\partial K} h_K^{\frac{1}{4}}\|\boldsymbol{\beta}\|_{L^4(\partial K)} h_K^{\frac{d-3}{4}}\|\boldsymbol{\mu}\|_{L^4(\partial K)} \\
&\leq \Big(\sum_{K\in\mathcal{T}_h} h_K^{\frac{2-d}{2}}\|\boldsymbol{v}\|_{\partial K}^2\Big)^{\frac{1}{2}}\Big(\sum_{K\in\mathcal{T}_h} h_K\|\boldsymbol{\beta}\|_{L^4(\partial K)}^4\Big)^{\frac{1}{4}}\Big(\sum_{K\in\mathcal{T}_h} h_K^{d-3}\|\boldsymbol{\mu}\|_{L^4(\partial K)}^4\Big)^{\frac{1}{4}} \\
&\leq C\Big(\sum_{K\in\mathcal{T}_h} h_K^{\frac{2-d}{2}}\|\boldsymbol{v}\|_{\partial K}^2\Big)^{\frac{1}{2}}\Big(\sum_{K\in\mathcal{T}_h} h_K\|\boldsymbol{\beta}\|_{L^4(\partial K)}^4\Big)^{1/4}\Big(\sum_{K\in\mathcal{T}_h} h_K^{-1}\|\boldsymbol{\mu}\|_{\partial K}^2\Big)^{\frac{1}{2}}.
\end{aligned}$$

Note that in the last inequality, we used $\|\boldsymbol{\mu}\|_{L^4(\partial K)} \lesssim h_K^{\frac{1-d}{4}}\|\boldsymbol{\mu}\|_{\partial K}$ for all $K \in \mathcal{T}_h$ and that $\sum a_i^2 \leq (\sum a_i)^2$. This completes the proof. ☐

A.2. A bound for the the discrete L^q-norm.

Proposition A.2. *For $k \geq 0$ and $1 \leq q < \infty$ for $d = 2$, $1 \leq q \leq 6$ for $d = 3$, there exist positive constants C and C_q such that*

$$\begin{aligned}
\|\boldsymbol{v}\|_{L^q(\Omega)} &\leq C\|\boldsymbol{v}\|_{1,h}, && \forall \boldsymbol{v} \in \boldsymbol{V}(h), \\
\|\boldsymbol{v}\|_{L^q(\Omega)} &\leq C_q|\!|\!|(\boldsymbol{v}, \boldsymbol{\mu})|\!|\!|_{1,h}, && \forall (\boldsymbol{v}, \boldsymbol{\mu}) \in \boldsymbol{V}(h) \times \boldsymbol{M}_h^0,
\end{aligned}$$

where C_q is independent of the mesh-size. Here $\boldsymbol{V}(h) := \boldsymbol{H}_0^1(\Omega) + \boldsymbol{V}_h$.

Note that for $q = 2$, the first inequality is a discrete version of the Poincaré inequality; see [2, 5].

Proof. The first inequality has been obtained in [22, Proposition 4.5]; see also [16, Theorem 5.3]. The second inequality follows from the first and from

$$\|\boldsymbol{v}\|_{1,h} \leq C|\!|\!|(\boldsymbol{v}, \boldsymbol{\mu})|\!|\!|_{1,h} \quad \forall (\boldsymbol{v}, \boldsymbol{\mu}) \in \boldsymbol{V}(h) \times \boldsymbol{M}_h^0.$$

This completes the proof. ☐

A.3. **Proof of the stability properties of the convective form.**

A.3.1. *Proof of Proposition* 3.4. Since $\boldsymbol{\beta} \in \boldsymbol{V}(h)$, after integrating by parts, we get

$$\begin{aligned}\mathcal{O}_h(\boldsymbol{\beta}; (\boldsymbol{u}, \widehat{\boldsymbol{u}}), (\boldsymbol{v}, \boldsymbol{\mu})) =& (\nabla \boldsymbol{u}, \boldsymbol{v} \otimes \boldsymbol{\beta})_{\mathcal{T}_h} - \langle ((\boldsymbol{u} - \widehat{\boldsymbol{u}}) \otimes \boldsymbol{\beta})\boldsymbol{n}, \boldsymbol{v} \rangle_{\partial \mathcal{T}_h} \\ &+ \langle S_{\boldsymbol{\beta}}(\boldsymbol{u} - \widehat{\boldsymbol{u}}), \boldsymbol{v} - \boldsymbol{\mu} \rangle_{\partial \mathcal{T}_h},\end{aligned}$$

and so, we can write

$$\mathcal{O}_h(\boldsymbol{\beta}; (\boldsymbol{u}, \widehat{\boldsymbol{u}}), (\boldsymbol{v}, \boldsymbol{\mu})) - \mathcal{O}_h(\boldsymbol{\gamma}; (\boldsymbol{u}, \widehat{\boldsymbol{u}}), (\boldsymbol{v}, \boldsymbol{\mu})) = O_1 + O_2 + O_3,$$

where

$$\begin{aligned} O_1 &:= (\nabla \boldsymbol{u}, \boldsymbol{v} \otimes (\boldsymbol{\beta} - \boldsymbol{\gamma}))_{\mathcal{T}_h}, \\ O_2 &:= - \langle ((\boldsymbol{u} - \widehat{\boldsymbol{u}}) \otimes (\boldsymbol{\beta} - \boldsymbol{\gamma}))\boldsymbol{n}, \boldsymbol{v} \rangle_{\partial \mathcal{T}_h}, \\ O_3 &:= \langle (S_{\boldsymbol{\beta}} - S_{\boldsymbol{\gamma}})(\boldsymbol{u} - \widehat{\boldsymbol{u}}), \boldsymbol{v} - \boldsymbol{\mu} \rangle_{\partial \mathcal{T}_h}. \end{aligned}$$

Let us estimate the first term O_1. Applying the Cauchy-Schwarz inequality, we get

$$\begin{aligned} O_1 &\leq \|\boldsymbol{\beta} - \boldsymbol{\gamma}\|_{L^4(\mathcal{T}_h)} \|\boldsymbol{u}\|_{1,h} \|\boldsymbol{v}\|_{L^4(\mathcal{T}_h)}, \\ &\leq C \|\boldsymbol{\beta} - \boldsymbol{\gamma}\|_{L^4(\mathcal{T}_h)} \|\boldsymbol{u}\|_{1,h} \|\boldsymbol{v}\|_{1,h}, \qquad \text{by the first inequality of Proposition A.2,} \\ &\leq C \|\boldsymbol{\beta} - \boldsymbol{\gamma}\|_{1,h} |||(\boldsymbol{u}, \widehat{\boldsymbol{u}})|||_{1,h} |||(\boldsymbol{v}, \boldsymbol{\mu})|||_{1,h}, \end{aligned}$$

by the second inequality of Proposition A.2. Next, let us estimate O_2. We have

$$\begin{aligned} O_2 \leq& C \Big(\sum_{K \in \mathcal{T}_h} h_K \|\boldsymbol{\beta} - \boldsymbol{\gamma}\|^4_{L^4(\partial K)} \Big)^{\frac{1}{4}} \Big(\sum_{K \in \mathcal{T}_h} h_K^{-1} \|\boldsymbol{u} - \widehat{\boldsymbol{u}}\|^2_{\partial K} \Big)^{\frac{1}{2}} \Big(\sum_{K \in \mathcal{T}_h} h_K \|\boldsymbol{v}\|^4_{L^4(\partial K)} \Big)^{\frac{1}{4}} \\ \leq& C \|\boldsymbol{\beta} - \boldsymbol{\gamma}\|_{1,h} |||(\boldsymbol{u}, \widehat{\boldsymbol{u}})|||_{1,h} |||(\boldsymbol{v}, \boldsymbol{\mu})|||_{1,h}, \end{aligned}$$

by [22, (7.7)] and the second inequality of Proposition A.2. It remains to bound O_3. Since, by definition, $S_{\boldsymbol{\beta}} = \max(\boldsymbol{\beta} \cdot \boldsymbol{n}, 0)$ and the function $a \mapsto \max(a, 0)$ is Lipschitz, we obtain by Lemma A.1, that

$$\begin{aligned} O_3 &\leq \big(\sum_{K \in \mathcal{T}_h} h_K \|\boldsymbol{\beta} - \boldsymbol{\gamma}\|^4_{L^4(\partial K)} \big) \big(\sum_{K \in \mathcal{T}_h} h_K^{-1} \|\boldsymbol{u} - \widehat{\boldsymbol{u}}\|^2_{\partial K} \big)^{\frac{1}{2}} \big(\sum_{K \in \mathcal{T}_h} h_K^{-1} \|\boldsymbol{v} - \boldsymbol{\mu}\|^2_{\partial K} \big)^{\frac{1}{2}} \\ &\leq C \big(\sum_{K \in \mathcal{T}_h} h_K \|\boldsymbol{\beta} - \boldsymbol{\gamma}\|^4_{L^4(\partial K)} \big) |||(\boldsymbol{u}, \widehat{\boldsymbol{u}})|||_{1,h} |||(\boldsymbol{v}, \boldsymbol{\mu})|||_{1,h} \\ &\leq C \, \|\boldsymbol{\beta} - \boldsymbol{\gamma}\|_{1,h} \, |||(\boldsymbol{u}, \widehat{\boldsymbol{u}})|||_{1,h} |||(\boldsymbol{v}, \boldsymbol{\mu})|||_{1,h}, \end{aligned}$$

by [22, (7.7)] and the second inequality of Proposition A.2 again. This completes the proof of Proposition 3.4.

A.3.2. *Proof of the stability properties of Proposition* 3.5. Let us prove the first estimate. We have

$$\begin{aligned} \Theta :=& \mathcal{O}_h(\boldsymbol{\beta}; (\boldsymbol{u}, \widehat{\boldsymbol{u}}), (\boldsymbol{v}, \boldsymbol{\mu})) \\ =& - (\boldsymbol{u} \otimes \boldsymbol{\beta}, \nabla \boldsymbol{v})_{\mathcal{T}_h} + \langle \widehat{\boldsymbol{u}} \otimes \boldsymbol{\beta} + S_{\boldsymbol{\beta}}(\boldsymbol{u} - \widehat{\boldsymbol{u}}), \boldsymbol{v} - \boldsymbol{\mu} \rangle_{\partial \mathcal{T}_h} \\ \leq& \|\boldsymbol{\beta}\|_{L^\infty(\Omega)} \big(\sum_{K \in \mathcal{T}_h} \|\boldsymbol{u}\|_K \, \|\nabla \boldsymbol{v}\|_K + h_K^{1/2} \, (\|\widehat{\boldsymbol{u}}\|_{\partial K} + \|\boldsymbol{u} - \widehat{\boldsymbol{u}}\|_{\partial K}) \, h_K^{-1/2} \|\boldsymbol{v} - \boldsymbol{\mu}\|_{\partial K} \big), \end{aligned}$$

and the result follows.

Let us now prove the second estimate. We have

$$\begin{aligned}\Theta &:= \mathcal{O}_h(\boldsymbol{\beta};(\boldsymbol{u},\widehat{\boldsymbol{u}}),(\boldsymbol{v},\boldsymbol{\mu})) - \mathcal{O}_h(\boldsymbol{\gamma};(\boldsymbol{u},\widehat{\boldsymbol{u}}),(\boldsymbol{v},\boldsymbol{\mu}))\\ &= -\,(\boldsymbol{u}\otimes(\boldsymbol{\beta}-\boldsymbol{\gamma}),\nabla\boldsymbol{v})_{\mathcal{T}_h} + \langle \widehat{\boldsymbol{u}}\otimes(\boldsymbol{\beta}-\boldsymbol{\gamma}) + (S_{\boldsymbol{\beta}}-S_{\boldsymbol{\gamma}})(\boldsymbol{u}-\widehat{\boldsymbol{u}}),\boldsymbol{v}-\boldsymbol{\mu}\rangle_{\partial\mathcal{T}_h}\\ &\le \|\!|(\boldsymbol{u},\widehat{\boldsymbol{u}})|\!\|_{\infty,h}\Big(\sum_{K\in\mathcal{T}_h}\|\boldsymbol{\beta}-\boldsymbol{\gamma}\|_K\,\|\nabla\boldsymbol{v}\|_K + h_K^{1/2}\|\boldsymbol{\beta}-\boldsymbol{\gamma}\|_{\partial K}\,h_K^{-1/2}\|\boldsymbol{v}-\boldsymbol{\mu}\|_{\partial K}\Big),\end{aligned}$$

and the result follows. This completes the proof of Proposition 3.5.

A.4. **Properties of the convective velocity.** Here, we prove Proposition 2.1. The fact that $\mathbb{P}(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)$ is well defined is a direct consequence of the fact that the Raviart-Thomas projection is well defined; see [6]. Also since $\widehat{\boldsymbol{u}}_h$ is single-valued along the interelement boundaries, the second equation defining $\mathbb{P}(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)$, (1.3b), immediately implies property (i). Property (iii) follows directly form property (ii). Let us prove property (ii). If $(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)$ satisfies (1.2c), then, for all $q\in P_k(K)$,

$$\begin{aligned}0 &= -(\boldsymbol{u}_h,\nabla q)_K + \langle\widehat{\boldsymbol{u}}_h\cdot\boldsymbol{n},q\rangle_{\partial K} = -(\mathbb{P}(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h),\nabla q)_K + \langle\mathbb{P}(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)\cdot\boldsymbol{n},q\rangle_{\partial K}\\ &= (\nabla\cdot\mathbb{P}(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h),q)_K\end{aligned}$$

by the equations defining the convective velocity (1.3). This readily implies property (ii).

To prove property (iv), we observe that the equations defining the convective velocity, (1.3), imply that for any $K\in\mathcal{T}_h$, we have

$$\begin{aligned}(\mathbb{P}(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)-\boldsymbol{u}_h,\boldsymbol{v})_K &= 0 && \forall\boldsymbol{v}\in\boldsymbol{P}_{k-1}(K),\\ \langle(\mathbb{P}(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)-\boldsymbol{u}_h)\cdot\boldsymbol{n},\lambda\rangle_F &= -\langle(\boldsymbol{u}_h-\widehat{\boldsymbol{u}}_h)\cdot\boldsymbol{n},\lambda\rangle_F && \forall\,\lambda\in P_k(F)\text{ for all faces }F\text{ of }K.\end{aligned}$$

Since $(\mathbb{P}(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)-\boldsymbol{u}_h)|_K\in RT_k(K)$, we have

$$\|\mathbb{P}(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)-\boldsymbol{u}_h\|_K\le Ch_K^{1/2}\|\boldsymbol{u}_h-\widehat{\boldsymbol{u}}_h\|_{\partial K},$$

and property (iv) immediately follows by using standard inverse inequalities.

Finally, property (v) immediately follows by comparing the equations defining the operator $\mathbb{P}$, (1.3), with those defining the Raviart-Thomas projection, (2.1). This concludes the proof of Proposition 2.1.

A.5. **Properties of the auxiliary projection P.** Let us recall the projection $P:\boldsymbol{H}^1(\mathcal{T}_h)\to\boldsymbol{V}_h$ [14, Section 6.1]. On each element K, $P\boldsymbol{w}$ is defined as the solution of

$$\begin{aligned}(P\boldsymbol{w}-\boldsymbol{w},\boldsymbol{v})_K &= 0 \quad \forall\boldsymbol{v}\in\boldsymbol{P}_{k-1}(K),\\ \langle(P\boldsymbol{w}-\boldsymbol{w})\cdot\boldsymbol{n},\boldsymbol{v}\cdot\boldsymbol{n}\rangle_{\partial K} &= 0 \quad \forall\boldsymbol{v}\in P_k(K)^{\perp}.\end{aligned}$$

Here $P_k(K)^{\perp} := \{p\in P_k(K):(p,q)_K = 0,\quad\forall q\in P_{k-1}(K)\}$. As shown in [14, Section 6.1], this is a well-defined projection which coincides with the Raviart-Thomas projection Π^{RT}, (2.1), whenever $\boldsymbol{w}$ is divergence free. Moreover,

$$\text{(A.1)}\quad \|P\boldsymbol{w}-\boldsymbol{w}\|_K + h_K^{1/2}\,\|P\boldsymbol{w}-\Pi_M\boldsymbol{w}\|_{\partial K}\le Ch_K^{k+1}|\boldsymbol{w}|_{k+1,K}\quad\forall\boldsymbol{w}\in H^{k+1}(K).$$

A.6. **Stability estimates of the Raviart-Thomas projection.** In the following result we gather some simple properties of the Raviart-Thomas projection.

Lemma A.3. *We have*

(i) $\|\!|(\Pi^{\mathrm{RT}}\boldsymbol{u}, \Pi_M\boldsymbol{u})\|\!|_{1,h} \le C_{\mathrm{HDG}}\|\nabla\boldsymbol{u}\|_{L^2(\Omega)}$ *for any* $\boldsymbol{u} \in \boldsymbol{H}^1(\Omega)$.

(ii) $\|\Pi^{\mathrm{RT}}\boldsymbol{u}\|_{L^\infty(\Omega)} \le C^\infty_{\mathrm{stab}}\|\boldsymbol{u}\|_{L^\infty(\Omega)}$ *for any* $\boldsymbol{u} \in \boldsymbol{C}^0(\overline{\Omega})$.

(iii) $\|\Pi^{\mathrm{RT}}\boldsymbol{\phi} - \Pi_M\boldsymbol{\phi}\|_{\partial K} \le \begin{cases} Ch_K^{3/2}\|\boldsymbol{\phi}\|_{2,K} & \textit{if } \boldsymbol{\phi} \in H^2(K) \textit{ and } k \ge 1, \\ Ch_K^{1/2}\|\boldsymbol{\phi}\|_{1,K} & \textit{if } \boldsymbol{\phi} \in H^1(K). \end{cases}$

Proof. Let us prove property (i). It is easy to see that $(\Pi_G(\nabla\boldsymbol{u}), \Pi^{\mathrm{RT}}\boldsymbol{u}, \Pi_M\boldsymbol{u})$ satisfies the first and third equations defining the HDG methods, namely, (1.2a) and (1.2c), respectively. Then, by Proposition 2.2, we have

$$\begin{aligned}\|\!|(\Pi^{\mathrm{RT}}\boldsymbol{u}, \Pi_M\boldsymbol{u})\|\!|_{1,h} &\le C_{\mathrm{HDG}}\big(\|\Pi_G(\nabla\boldsymbol{u})\|^2_{\mathcal{T}_h} + \sum_{K\in\mathcal{T}_h} h_K^{-1}\|(\Pi^{\mathrm{RT}}\boldsymbol{u} - \Pi_M\boldsymbol{u})\cdot\boldsymbol{n}\|^2_{\partial K}\big)^{1/2} \\ &= C_{\mathrm{HDG}}\|\Pi_G(\nabla\boldsymbol{u})\|_{\mathcal{T}_h} \le C_{\mathrm{HDG}}\|\nabla\boldsymbol{u}\|_{\mathcal{T}_h},\end{aligned}$$

due to the fact that $(\Pi^{\mathrm{RT}}\boldsymbol{u} - \Pi_M\boldsymbol{u})\cdot\boldsymbol{n}|_{\partial\mathcal{T}_h} = 0$ and to the fact that Π_G is the $L^2(\Omega)$-projection into G_h. This proves property (i).

Property (ii) follows by a straightforward scaling argument. It remains to prove property (iii). We have

$$\begin{aligned}\|\Pi^{\mathrm{RT}}\boldsymbol{\phi} - \Pi_M\boldsymbol{\phi}\|_{\partial K} &\le \|\Pi^{\mathrm{RT}}\boldsymbol{\phi} - \boldsymbol{\phi}\|_{\partial K} + \|\boldsymbol{\phi} - \Pi_M\boldsymbol{\phi}\|_{\partial K} \\ &\le Ch_K^{-1/2}\|\Pi^{\mathrm{RT}}\boldsymbol{\phi} - P\boldsymbol{\phi}\|_K + Ch_K^{1/2}\|\nabla(\Pi^{\mathrm{RT}}\boldsymbol{\phi} - P\boldsymbol{\phi})\|_K + \|\Pi_M\boldsymbol{\phi} - \boldsymbol{\phi}\|_{\partial K} \\ &\le Ch_K^{3/2}\|\boldsymbol{\phi}\|_{2,K},\end{aligned}$$

by the approximation properties of the Raviart-Thomas projection [6]. The other inequality can be shown similarly. This completes the proof. □

Acknowledgements

The authors thank Weiwei Sun, from the Department of Mathematics of City University of Hong Kong, for valuable discussions on the dual problem.

References

[1] R. A. Adams, *Sobolev Spaces*, Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, 1975. Pure and Applied Mathematics, Vol. 65. MR0450957

[2] D. N. Arnold, *An interior penalty finite element method with discontinuous elements*, SIAM J. Numer. Anal. **19** (1982), no. 4, 742–760, DOI 10.1137/0719052. MR664882

[3] G. A. Baker, W. N. Jureidini, and O. A. Karakashian, *Piecewise solenoidal vector fields and the Stokes problem*, SIAM J. Numer. Anal. **27** (1990), no. 6, 1466–1485, DOI 10.1137/0727085. MR1080332

[4] P. Bastian and B. Rivière, *Superconvergence and* $H(\mathrm{div})$ *projection for discontinuous Galerkin methods*, Internat. J. Numer. Methods Fluids **42** (2003), no. 10, 1043–1057, DOI 10.1002/fld.562. MR1991232

[5] S. C. Brenner, *Poincaré-Friedrichs inequalities for piecewise* H^1 *functions*, SIAM J. Numer. Anal. **41** (2003), no. 1, 306–324, DOI 10.1137/S0036142902401311. MR1974504

[6] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, Springer Series in Computational Mathematics, vol. 15, Springer-Verlag, New York, 1991. MR1115205

[7] J. Carrero, B. Cockburn, and D. Schötzau, *Hybridized globally divergence-free LDG methods. I. The Stokes problem*, Math. Comp. **75** (2006), no. 254, 533–563, DOI 10.1090/S0025-5718-05-01804-1. MR2196980

[8] A. Cesmelioglu, B. Cockburn, N. C. Nguyen, and J. Peraire, *Analysis of HDG methods for Oseen equations*, J. Sci. Comput. **55** (2013), no. 2, 392–431, DOI 10.1007/s10915-012-9639-y. MR3044180

[9] B. Cockburn, J. Gopalakrishnan, N. C. Nguyen, J. Peraire, and F.-J. Sayas, *Analysis of HDG methods for Stokes flow*, Math. Comp. **80** (2011), no. 274, 723–760, DOI 10.1090/S0025-5718-2010-02410-X. MR2772094

[10] B. Cockburn, G. Kanschat, and D. Schotzau, *A locally conservative LDG method for the incompressible Navier-Stokes equations*, Math. Comp. **74** (2005), no. 251, 1067–1095, DOI 10.1090/S0025-5718-04-01718-1. MR2136994

[11] B. Cockburn, G. Kanschat, and D. Schötzau, *A note on discontinuous Galerkin divergence-free solutions of the Navier-Stokes equations*, J. Sci. Comput. **31** (2007), no. 1-2, 61–73, DOI 10.1007/s10915-006-9107-7. MR2304270

[12] B. Cockburn, G. Kanschat, and D. Schötzau, *An equal-order DG method for the incompressible Navier-Stokes equations*, J. Sci. Comput. **40** (2009), no. 1-3, 188–210, DOI 10.1007/s10915-008-9261-1. MR2511732

[13] B. Cockburn, G. Kanschat, D. Schötzau, and C. Schwab, *Local discontinuous Galerkin methods for the Stokes system*, SIAM J. Numer. Anal. **40** (2002), no. 1, 319–343 (electronic), DOI 10.1137/S0036142900380121. MR1921922

[14] B. Cockburn and F.-J. Sayas, *Divergence-conforming HDG methods for Stokes flows*, Math. Comp. **83** (2014), no. 288, 1571–1598, DOI 10.1090/S0025-5718-2014-02802-0. MR3194122

[15] B. Cockburn and K. Shi, *Conditions for superconvergence of HDG methods for Stokes flow*, Math. Comp. **82** (2013), no. 282, 651–671, DOI 10.1090/S0025-5718-2012-02644-5. MR3008833

[16] D. A. Di Pietro and A. Ern, *Mathematical Aspects of Discontinuous Galerkin Methods*, Mathématiques & Applications (Berlin) [Mathematics & Applications], vol. 69, Springer, Heidelberg, 2012. MR2882148

[17] M. Fortin, *Finite element solution of the Navier-Stokes equations*, Acta Numer., Cambridge Univ. Press, Cambridge, 1993, pp. 239–284, DOI 10.1017/S0962492900002373. MR1224684

[18] V. Girault and P.-A. Raviart, *Finite element approximation of the Navier-Stokes equations*, Lecture Notes in Mathematics, vol. 749, Springer-Verlag, Berlin-New York, 1979. MR548867

[19] V. Girault, B. Rivière, and M. F. Wheeler, *A discontinuous Galerkin method with nonoverlapping domain decomposition for the Stokes and Navier-Stokes problems*, Math. Comp. **74** (2005), no. 249, 53–84, DOI 10.1090/S0025-5718-04-01652-7. MR2085402

[20] T. J. R. Hughes and L. P. Franca, *A new finite element formulation for computational fluid dynamics. VII. The Stokes problem with various well-posed boundary conditions: symmetric formulations that converge for all velocity/pressure spaces*, Comput. Methods Appl. Mech. Engrg. **65** (1987), no. 1, 85–96, DOI 10.1016/0045-7825(87)90184-8. MR914609

[21] T. J. R. Hughes, L. P. Franca, and M. Balestra, *A new finite element formulation for computational fluid dynamics. V. Circumventing the Babuška-Brezzi condition: a stable Petrov-Galerkin formulation of the Stokes problem accommodating equal-order interpolations*, Comput. Methods Appl. Mech. Engrg. **59** (1986), no. 1, 85–99, DOI 10.1016/0045-7825(86)90025-3. MR868143

[22] O. A. Karakashian and W. N. Jureidini, *A nonconforming finite element method for the stationary Navier-Stokes equations*, SIAM J. Numer. Anal. **35** (1998), no. 1, 93–120, DOI 10.1137/S0036142996297199. MR1618436

[23] N. Kechkar and D. Silvester, *Analysis of locally stabilized mixed finite element methods for the Stokes problem*, Math. Comp. **58** (1992), no. 197, 1–10, DOI 10.2307/2153016. MR1106973

[24] C. Lehrenfeld and J. Schöberl, *High order exactly divergence-free hybrid discontinuous Galerkin methods for unsteady incompressible flows*, Comput. Methods Appl. Mech. Engrg. **307** (2016), 339–361, DOI 10.1016/j.cma.2016.04.025. MR3511719

[25] A. Montlaur, S. Fernandez-Mendez, and A. Huerta, *Discontinuous Galerkin methods for the Stokes equations using divergence-free approximations*, Internat. J. Numer. Methods Fluids **57** (2008), no. 9, 1071–1092, DOI 10.1002/fld.1716. MR2435083

[26] A. Montlaur, S. Fernandez-Mendez, J. Peraire, and A. Huerta, *Discontinuous Galerkin methods for the Navier-Stokes equations using solenoidal approximations*, Internat. J. Numer. Methods Fluids **64** (2010), no. 5, 549–564, DOI 10.1002/fld.2161. MR2683650

[27] N. C. Nguyen, J. Peraire, and B. Cockburn, *A hybridizable discontinuous Galerkin method for Stokes flow*, Comput. Methods Appl. Mech. Engrg. **199** (2010), no. 9-12, 582–597, DOI 10.1016/j.cma.2009.10.007. MR2796169

[28] N. C. Nguyen, J. Peraire, and B. Cockburn, *An implicit high-order hybridizable discontinuous Galerkin method for the incompressible Navier-Stokes equations*, J. Comput. Phys. **230** (2011), no. 4, 1147–1170, DOI 10.1016/j.jcp.2010.10.032. MR2753354

[29] D. Schötzau, C. Schwab, and A. Toselli, *Stabilized hp-DGFEM for incompressible flow*, Math. Models Methods Appl. Sci. **13** (2003), no. 10, 1413–1436, DOI 10.1142/S0218202503002970. MR2013492

[30] R. Stenberg, *Some new families of finite elements for the Stokes equations*, Numer. Math. **56** (1990), no. 8, 827–838, DOI 10.1007/BF01405291. MR1035181

[31] A. Toselli, *hp discontinuous Galerkin approximations for the Stokes problem*, Math. Models Methods Appl. Sci. **12** (2002), no. 11, 1565–1597, DOI 10.1142/S0218202502002240. MR1938957

Department of Mathematics and Statistics, Oakland University, Rochester, Michigan 48309
E-mail address: cesmelio@oakland.edu

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

Department of Mathematics, City University of Hong Kong, 83 Tat Chee Avenue, Kowloon, Hong Kong, People's Republic of China
E-mail address: weifeqiu@cityu.edu.hk