A SUPERCONVERGENT LDG-HYBRIDIZABLE GALERKIN METHOD FOR SECOND-ORDER ELLIPTIC PROBLEMS

BERNARDO COCKBURN, BO DONG, AND JOHNNY GUZMÁN

ABSTRACT. We identify and study an LDG-hybridizable Galerkin method, which is not an LDG method, for second-order elliptic problems in several space dimensions with remarkable convergence properties. Unlike all other known discontinuous Galerkin methods using polynomials of degree $k \geq 0$ for both the potential as well as the flux, the order of convergence in L^2 of both unknowns is $k + 1$. Moreover, both the approximate potential as well as its numerical trace superconverge in L^2-like norms, to suitably chosen projections of the potential, with order $k + 2$. This allows the application of element-by-element postprocessing of the approximate solution which provides an approximation of the potential converging with order $k + 2$ in L^2. The method can be thought to be *in between* the hybridized version of the Raviart-Thomas and that of the Brezzi-Douglas-Marini mixed methods.

1. INTRODUCTION

In this paper, we consider the LDG-hybridizable (LDG-H) Galerkin methods recently introduced in [11] and show how to define their numerical traces in order to achieve the optimal order of convergence for the approximation to the flux, and to obtain superconvergence properties similar to those of the hybridized mixed methods of Raviart-Thomas (RT) [16] and the Brezzi-Douglas-Marini (BDM) [4] methods; see also [10].

For the sake of simplicity of the exposition, we carry this out in the setting of the model second-order elliptic problem

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

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

$$(1.1c) \qquad\qquad u = g \qquad \text{on } \partial\Omega_D,$$

$$(1.1d) \qquad\qquad \boldsymbol{q} \cdot \boldsymbol{n} = \mathsf{q_N} \qquad \text{on } \partial\Omega_N,$$

where $\Omega \subset \mathbb{R}^d$ is a polyhedral domain ($d \geq 2$), $f \in L^2(\Omega)$, and $\boldsymbol{c} = \boldsymbol{c}(\boldsymbol{x})$ is a symmetric $d \times d$ matrix function that is uniformly positive definite on Ω with

components in $L^\infty(\Omega)$. As usual, we assume that the $(d-1)$-Lebesgue measure of $\partial\Omega_D$ is not zero, that $\partial\Omega = \overline{\partial\Omega_D} \cup \overline{\partial\Omega_N}$ and that $\partial\Omega_D \cap \partial\Omega_N = \emptyset$.

To describe our results, we need to introduce what we will call hybridized Galerkin methods; they are one of the methods studied in [11]. To do that, let us introduce some notation. We denote by $\Omega_h = \{K\}$ a triangulation of the domain Ω of shape-regular simplexes K and set $\partial\Omega_h := \{\partial K : K \in \Omega_h\}$. We associate to this triangulation the set of interior faces $\mathscr{E}_h^i$ and the set of boundary faces $\mathscr{E}_h^\partial$. We say that $e \in \mathscr{E}_h^i$ if there are two simplexes K^+ and K^- in Ω_h such that $e = \partial K^+ \cap \partial K^-$, and we say that $e \in \mathscr{E}_h^\partial$ if there is a simplex in Ω_h such that $e = \partial K \cap \partial\Omega$. We set $\mathscr{E}_h := \mathscr{E}_h^i \cup \mathscr{E}_h^\partial$.

The hybridized Galerkin methods [11] are dual-mixed hybrid methods (see the definition in [8] and an early example in [15]) which seek an approximation to the exact solution $(\boldsymbol{q}|_\Omega, u|_\Omega, u|_{\mathscr{E}_h \setminus \partial\Omega_N})$, $(\boldsymbol{q}_h, u_h, \lambda_h)$, in a finite-dimensional space $\boldsymbol{V}_h \times W_h \times M_h$ of the form

$$(1.2a) \qquad \boldsymbol{V}_h := \{\boldsymbol{v} \in \boldsymbol{L}^2(\Omega) : \boldsymbol{v}|_K \in \boldsymbol{V}(K) \quad \forall K \in \Omega_h\},$$

$$(1.2b) \qquad W_h := \{\omega \in L^2(\Omega) : \omega|_K \in W(K) \quad \forall K \in \Omega_h\},$$

$$(1.2c) \qquad M_h := \{\mathsf{m} \in L^2(\partial\Omega_h) : \mathsf{m}|_e \in M(e) \quad \forall e \in \mathscr{E}_h, \quad \mathsf{m}|_{\partial\Omega_D} = 0\},$$

and determines it by requiring that

$$(1.3a) \qquad (\boldsymbol{c}\,\boldsymbol{q}_h, \boldsymbol{v})_{\Omega_h} - (u_h, \nabla \cdot \boldsymbol{v})_{\Omega_h} + \langle \widehat{u}_h, \boldsymbol{v} \cdot \boldsymbol{n} \rangle_{\partial\Omega_h} = 0,$$

$$(1.3b) \qquad -(\boldsymbol{q}_h, \nabla\omega)_{\Omega_h} + \langle \widehat{\boldsymbol{q}}_h \cdot \boldsymbol{n}, \omega \rangle_{\partial\Omega_h} = (f, \omega)_{\Omega_h},$$

$$(1.3c) \qquad \langle \widehat{\boldsymbol{q}}_h \cdot \boldsymbol{n}, \mu \rangle_{\partial\Omega_h} = \langle \mathsf{q_N}, \mu \rangle_{\partial\Omega_N},$$

for all $(\boldsymbol{v}, \omega, \mu) \in \boldsymbol{V}_h \times W_h \times M_h$. Here, we have used the notation

$$(\boldsymbol{\sigma}, \boldsymbol{v})_{\Omega_h} := \sum_{K \in \Omega_h} \int_K \boldsymbol{\sigma}(x) \cdot \boldsymbol{v}(x)\,dx,$$

$$(\zeta, \omega)_{\Omega_h} := \sum_{K \in \Omega_h} \int_K \zeta(x)\,\omega(x)\,dx,$$

$$\langle \zeta, \boldsymbol{v} \cdot \boldsymbol{n} \rangle_{\partial\Omega_h} := \sum_{K \in \Omega_h} \int_{\partial K} \zeta(\gamma)\,\boldsymbol{v}(\gamma) \cdot \boldsymbol{n}\,d\gamma,$$

for any functions $\boldsymbol{\sigma}, \boldsymbol{v}$ in $\boldsymbol{H}^1(\Omega_h) := [H^1(\Omega_h)]^d$ and ζ, ω in the space $H^1(\Omega_h) = \{v \in L^2(\Omega) : v|_K \in H^1(K) \quad \forall K \in \Omega_h\}$. The outward normal unit vector to ∂K is denoted by $\boldsymbol{n}$.

To complete the description of the hybridized Galerkin methods, the definition of numerical traces $(\widehat{\boldsymbol{q}}_h, \widehat{u}_h)$ on the faces of the triangulation $\mathscr{E}_h$ has to be provided. The choice which is relevant here is

$$(1.4a) \qquad \widehat{u}_h = \begin{cases} \mathsf{P}_\partial\,g & \text{on } \mathscr{E}_h \cap \partial\Omega_D, \\ \lambda_h & \text{on } \mathscr{E}_h \setminus \partial\Omega_D, \end{cases}$$

$$(1.4b) \qquad \widehat{\boldsymbol{q}}_h = \boldsymbol{q}_h + \tau\,(u_h - \widehat{u}_h)\boldsymbol{n} \quad \text{on } \mathscr{E}_h,$$

where P_∂ denotes an L^2-projection defined as follows. Given any function $\zeta \in L^2(\mathscr{E}_h)$ and an arbitrary face $e \in \mathscr{E}_h$, the restriction of $\mathsf{P}_\partial\zeta$ to e is defined as the element of $\mathcal{P}^k(e)$ that satisfies

$$(1.5) \qquad \langle \mathsf{P}_\partial\zeta - \zeta, \omega \rangle_e = 0, \qquad \forall\,\omega \in \mathcal{P}^k(e).$$

Note that by suitably choosing the *local spaces* $\boldsymbol{V}(K)$, $W(K)$, and $M(e)$, and the values of the *local stabilization parameters* τ, we can obtain the hybridized RT_k, the hybridized BDM_k and the $\mathrm{LDG\text{-}H}_k$ methods; see Tables 1 and 2. In Table 1 and in the remainder of this paper, we denote the space of polynomials of degree at most $k \geq 0$ defined on D by $\mathcal{P}^k(D)$, and set $\boldsymbol{\mathcal{P}}^k(D) := [\mathcal{P}^k(D)]^d$. Since all these methods can be implemented in the same way and can be used in different elements while being automatically coupled, what is relevant, as argued in [11], is to find out which method should be used in what *element* in order to optimize the computational effort. It is thus important from this perspective to develop DG methods as accurate and efficient as mixed methods so that they can be used in situations in which mixed method cannot. The LDG-H methods we uncover in this paper are the first example of those methods.

TABLE 1. The local spaces

method	$\boldsymbol{V}(K)$	$W(K)$	$M(e)$
RT_k	$\boldsymbol{\mathcal{P}}^k(K) \oplus \boldsymbol{x}\,\mathcal{P}^k(K)$	$\mathcal{P}^k(K)$	$\mathcal{P}^k(e)$
$\mathrm{LDG\text{-}H}_k$	$\boldsymbol{\mathcal{P}}^k(K)$	$\mathcal{P}^k(K)$	$\mathcal{P}^k(e)$
BDM_k	$\boldsymbol{\mathcal{P}}^k(K)$	$\mathcal{P}^{k-1}(K)$	$\mathcal{P}^k(e)$

TABLE 2. The local stabilization parameters τ

method	$\tau\lvert_{\partial K}$
RT_k	$\equiv 0$
$\mathrm{LDG\text{-}H}_k$	$\geq 0, \not\equiv 0$
BDM_k	$\equiv 0$

It is well known that the RT_k and BDM_k methods provide an approximation $\boldsymbol{q}_h$ to the flux which converges in L^2 with order $k+1$, that u_h and λ_h superconverge in L^2-like norms to suitably chosen projections of the potential u with order $k+2$, and that, as a consequence, it is possible to postprocess the approximate solution to obtain another approximation u_h^* converging in L^2 with order $k+2$; see [1] and [4], and also [10]. In this paper, we use an extension of the postprocessing proposed in [17, 18] and [14]. Given the similarities between these two mixed methods and the $\mathrm{LDG\text{-}H}_k$ method, it is natural to ask if it is possible to choose the local penalization parameters τ as to obtain similar convergence and superconvergence results. The main contribution of this paper is to show that this is actually possible.

Indeed, we show that this happens if we take, on each simplex $K \in \Omega_h$,

$$(1.6) \qquad \tau = \begin{cases} 0, & \text{on } \partial K \setminus e_K^\tau, \\ \tau_K, & \text{on } e_K^\tau, \end{cases}$$

where e_K^τ is an arbitrary but fixed face of K and τ_K is a strictly positive real number. Since the local penalization parameter τ is nonzero only on a single face of each simplex, we call this LDG-H method the *single-face hybridizable* DG method; for

simplicity, we are going to refer to the method under consideration by the SF-H$_k$ method. It is interesting to note two of the *minimal dissipation* DG methods considered in [6], in the framework of a study of superconvergence properties of DG methods for one-dimensional steady-state convection-diffusion problems, happen to be an SF-H method. The first is called the md-DG method (see Table 1 in [6]) and is obtained, in our notation, by taking on each interior node x_i,

$$\tau(x_i^+) = 0 \qquad \text{and} \qquad \tau(x_i^-) = k/h_i,$$

where h_i is the size of the interval to the left of the node x_i. The second is called the md-LDG method, and is obtained by taking the above choice of parameters τ formally letting $\tau(x_i^-)$ go to infinity. The authors are not aware of any other instance of SF-H methods. In particular, let us emphasize that SF-H *methods are not* LDG *methods* whenever the stabilization parameters τ are finite; see the discussion about LDG-H methods in [11].

   In Table 3, we compare the orders of convergence for the flux of this method and the above-mentioned mixed methods. We have also included the order of convergence for the general LDG-H methods; it can be deduced from their characterization [11] and the study of DG methods carried out in [5]. Finally, in Table 4, we display the orders of convergence of the postprocessed approximation u_h^* to the potential.

TABLE 3. The orders of convergence in h of the L^2-errors

method	$\| \boldsymbol{q} - \boldsymbol{q}_h \|_{L^2(\Omega_h)}$	$\| u - u_h \|_{L^2(\Omega_h)}$	condition
RT$_k$	$k+1$	$k+1$	$k \geq 0$
LDG-H$_k$ [5]	k	$k+1$	$k \geq 1$ and $\tau = \mathcal{O}(1/h)$
LDG-H$_k$ [5]	$k+1/2$	$k+1$	$k \geq 0$ and $\tau = \mathcal{O}(1)$
SF-H $_k$	$k+1$	$k+1$	$k \geq 0$
BDM$_k$	$k+1$	k	$k \geq 1$

TABLE 4. The orders of convergence in h of $\| u - u_h^* \|_{L^2(\Omega_h)}$

method	order	condition
RT$_k$	$k+2$	$k \geq 0$
SF-H $_k$	$k+2$	$k \geq 1$
SF-H $_k$	$k+2$	$k = 0,\, f = 0$
BDM$_k$	$k+2$	$k \geq 2$
BDM$_k$	$k+1$	$k = 1$

   We also uncover new relations between these three methods. One of the main features of the hybridizable Galerkin methods proposed in [11] is that the only degrees of freedom that turn out to be globally coupled are those of the so-called Lagrange multiplier λ_h. This implies, in particular, that the LDG-H methods can be *more efficiently* implemented than the LDG methods introduced in [12]. In fact, see the discussion in [11], they can be implemented as efficiently as the hybridized

RT_k; see [7] for the case $k = 0$ in two dimensions, and BDM_k mixed methods, see [10] for the case $k \geq 0$ in multi-dimensions for both of these methods. Here we show that the stiffness matrix of the Lagrange multiplier for the RT_k, BDM_k and SF-H$_k$ methods is actually *identical* and that, when $f|_K \in \mathcal{P}^{k-1}(K)$ for all $K \in \Omega_h$, these methods provide the *same* approximation $(\boldsymbol{q}_h, \lambda_h)$.

Next, let us briefly comment on the approach taken to carry out the a priori error analysis of the SF-H$_k$ methods. We did not take the approach used in [5] to analyze DG and LDG methods, or that used in the unified analysis of DG methods [2]. Instead, we exploited the unifying framework introduced in [11] to render the analysis of the SF-H$_k$ methods as close as possible to those of the hybridized RT_k and BDM_k methods. Since a key ingredient in those analyzes is the existence of a projection $(\boldsymbol{\Pi}, \mathbb{P})$ satisfying the so-called *commutativity* property

$$\nabla \cdot \boldsymbol{\Pi}\boldsymbol{\sigma} = \mathbb{P}^t \, \nabla \cdot \boldsymbol{\sigma},$$

for all $\boldsymbol{\sigma} \in \boldsymbol{H}(div, \Omega)$, the crucial step in the analysis was to find a similar projection. Unlike the above-mentioned mixed methods, the space of fluxes $\boldsymbol{V}_h$ of the SF-H$_k$ methods is not included in $\boldsymbol{H}(div, \Omega)$ and, as a consequence, the above commutativity property can only be satisfied in a weak sense. We found a *new* projection satisfying the following weak version of the commutativity property:

$$-(\nabla \zeta, \boldsymbol{\Pi}\boldsymbol{\sigma})_{\Omega_h} = (\mathbb{P}\zeta, \nabla \cdot \boldsymbol{\sigma})_{\Omega_h}$$

for all $(\boldsymbol{\sigma}, \zeta) \in \boldsymbol{H}^1(\Omega_h) \times H^1(\Omega_h)$ such that $\zeta|_{\partial\Omega} = 0$. Just as the local spaces of the SF-H$_k$ methods are, roughly speaking, "in between" the local spaces of the RT_k and BDM_k methods, this projection can also be considered to be "in between" the corresponding projections of those mixed methods. The construction of this projection, which is intimately linked to the definition of the numerical traces of the method $\widehat{u}_h$ and $\widehat{\boldsymbol{q}}_h$ and to the choice of the local spaces, is certainly the most interesting aspect of the analysis of the SF-H$_k$ methods. The first component of the projection, $\boldsymbol{\Pi}$, was used in the error analysis of the *minimal dissipation* LDG method in [9].

The paper is organized as follows. In Section 2, we state and discuss our main results and then prove them in Section 3. In Section 4, we display numerical experiments validating the theoretical results. Finally, in Section 5, we end with some concluding remarks.

2. THE MAIN RESULTS

2.1. **The projection** $(\boldsymbol{\Pi}, \mathbb{P})$. In this subsection, we define the projection

$$(\boldsymbol{\Pi}, \mathbb{P}) : \boldsymbol{H}^1(\Omega_h) \times H^1(\Omega_h) \to \boldsymbol{V}_h \times W_h,$$

and gather its main properties.

Given a function $\boldsymbol{\sigma} \in \boldsymbol{H}^1(\Omega_h)$ and an arbitrary simplex $K \in \Omega_h$, the restriction of $\boldsymbol{\Pi}\boldsymbol{\sigma}$ to K is defined as the element of $\boldsymbol{\mathcal{P}}^k(K)$ that satisfies

$$(2.1a) \qquad (\boldsymbol{\Pi}\boldsymbol{\sigma} - \boldsymbol{\sigma}, \boldsymbol{v})_K = 0, \qquad \forall \boldsymbol{v} \in \boldsymbol{\mathcal{P}}^{k-1}(K), \text{ if } k \geq 1,$$

$$(2.1b) \qquad \langle (\boldsymbol{\Pi}\boldsymbol{\sigma} - \boldsymbol{\sigma}) \cdot \boldsymbol{n}, \omega \rangle_e = 0, \qquad \forall \omega \in \mathcal{P}^k(e) \text{ and } \forall e \in \partial K, e \neq e_K^\tau.$$

Similarly, given a function $\zeta \in H^1(\Omega_h)$ and an arbitrary simplex $K \in \Omega_h$, the restriction of $\mathbb{P}\zeta$ to K is defined as the element of $\mathcal{P}^k(K)$ that satisfies

$$(2.2a) \qquad (\mathbb{P}\zeta - \zeta, \mathsf{w})_K = 0, \qquad \forall\, \mathsf{w} \in \mathcal{P}^{k-1}(K), \text{ if } k \geq 1,$$

$$(2.2b) \qquad \langle \mathbb{P}\zeta - \zeta, \omega \rangle_{e_K^\tau} = 0, \qquad \forall\, \omega \in \mathcal{P}^k(e_K^\tau).$$

We gather the main properties of this projection in the following result. To state it, we need to recall the definition of some classical projections. Given a function $\boldsymbol{\sigma} \in \boldsymbol{H}^1(\Omega_h)$ and an arbitrary simplex $K \in \Omega_h$, the restriction of $\boldsymbol{\Pi}^{\mathrm{RT}}\boldsymbol{\sigma}$ to K is defined as the element of $\boldsymbol{\mathcal{P}}^k(K) \oplus \boldsymbol{x}\, \mathcal{P}(K)^k$ that satisfies

$$(2.3a) \qquad (\boldsymbol{\Pi}^{\mathrm{RT}}\boldsymbol{\sigma} - \boldsymbol{\sigma}, \boldsymbol{v})_K = 0, \qquad \forall\, \boldsymbol{v} \in \boldsymbol{\mathcal{P}}^{k-1}(K), \text{ if } k \geq 1,$$

$$(2.3b) \qquad \langle (\boldsymbol{\Pi}^{\mathrm{RT}}\boldsymbol{\sigma} - \boldsymbol{\sigma}) \cdot \boldsymbol{n}, \omega \rangle_e = 0, \qquad \forall\, \omega \in \mathcal{P}^k(e), \text{ for all faces of } K.$$

Given a function $\zeta \in H^1(\Omega_h)$ and an arbitrary simplex $K \in \Omega_h$, the restriction of $\mathsf{P}^\ell \zeta$ to K is defined as the element of $\mathcal{P}^\ell(K)$ that satisfies

$$(2.4) \qquad (\mathsf{P}^\ell \zeta - \zeta, \omega)_K = 0, \qquad \forall\, \omega \in \mathcal{P}^\ell(K).$$

To simplify the notation, we are going to write P instead of P^k. Note that $(\boldsymbol{\Pi}^{\mathrm{RT}}, \mathsf{P})$ is nothing but the projection for the RT_k method. We are now ready to state our result.

Proposition 2.1. *The projection $(\boldsymbol{\Pi}, \mathbb{P})$ given by (2.1) and (2.2) is well defined. Moreover, on each simplex $K \in \Omega_h$, it satisfies the* orthogonality *properties*

$$\begin{aligned}
&\text{(i)} \quad (\zeta - \mathbb{P}\zeta, \nabla \cdot \boldsymbol{v})_K = 0, \\
&\text{(ii)} \quad (\boldsymbol{\sigma} - \boldsymbol{\Pi}\boldsymbol{\sigma}, \nabla\omega)_K = 0, \\
&\text{(iii)} \quad \langle \mathbb{P}\zeta - \mathsf{P}_\partial \zeta, \boldsymbol{\Pi}\boldsymbol{\sigma} \cdot \boldsymbol{n} - \mathsf{P}_\partial \boldsymbol{\sigma} \cdot \boldsymbol{n} \rangle_e = 0 \quad \text{for all faces } e \text{ of } K,
\end{aligned}$$

for all $(\boldsymbol{v}, \omega) \in \boldsymbol{\mathcal{P}}^k(K) \times \mathcal{P}^k(K)$, and the weak commutativity *property*

$$\text{(iv)} \quad -(\nabla\zeta, \boldsymbol{\Pi}\boldsymbol{\sigma})_K = (\mathbb{P}\zeta, \nabla \cdot \boldsymbol{\sigma})_K - \langle \mathsf{P}_\partial \zeta, \mathsf{P}_\partial \boldsymbol{\sigma} \cdot \boldsymbol{n} \rangle_{\partial K},$$

for all $(\boldsymbol{\sigma}, \zeta) \in \boldsymbol{H}^1(\Omega_h) \times H^1(\Omega_h)$. Finally, we have the following approximation estimates

$$\begin{aligned}
&\text{(v)} \quad \| \boldsymbol{\Pi}\boldsymbol{\sigma} \cdot \boldsymbol{n} - \mathsf{P}_\partial \boldsymbol{\sigma} \cdot \boldsymbol{n} \|_{L^2(e_K^\tau)} \leq C\, h_K^{r+1/2} \, |\mathsf{P}\nabla \cdot \boldsymbol{\sigma}|_{H^r(K)}, \\
&\text{(vi)} \quad \| \boldsymbol{\Pi}\boldsymbol{\sigma} - \boldsymbol{\Pi}^{\mathrm{RT}}\boldsymbol{\sigma} \|_{L^2(K)} \leq C\, h_K^{r+1} \, |\mathsf{P}\nabla \cdot \boldsymbol{\sigma}|_{H^r(K)}, \\
&\text{(vii)} \quad \| \mathbb{P}\zeta - \mathsf{P}\zeta \|_{L^2(K)} \leq C\, h_K^{s+1} \, |\nabla\zeta|_{\boldsymbol{H}^s(K)},
\end{aligned}$$

where $r, s \in [0, k]$, h_K is the diameter of K, and C depends only on k and the shape-regularity parameters of the simplex K, for any $(\boldsymbol{\sigma}, \zeta) \in \boldsymbol{H}^1(K) \times H^1(K)$.

We are going to show that the three orthogonality properties imply all the others; they are thus the crucial properties for the analysis. Note also that, by simply adding the identity (iv) over all $K \in \Omega_h$, we obtain the weak commutativity property discussed in the introduction.

2.2. **Characterization of the approximate solution.** Next we give a characterization of the approximate solution provided by the SF-H$_k$ method. We begin by characterizing the difference between the numerical traces and the traces of the approximate solutions on each simplex.

Proposition 2.2. *For each simplex $K \in \Omega_h$, we have that,*

$$(\widehat{\boldsymbol{q}}_h - \boldsymbol{q}_h) \cdot \boldsymbol{n} = \tau \, (u_h - \widehat{u}_h) = \mathsf{P}_{\partial} \boldsymbol{q} \cdot \boldsymbol{n} - \boldsymbol{\Pi} \boldsymbol{q} \cdot \boldsymbol{n} \qquad on \ \partial K.$$

We see that the jump $(\widehat{\boldsymbol{q}}_h - \boldsymbol{q}_h) \cdot \boldsymbol{n}$ is *independent* of the value of τ whereas the jump $\widehat{u}_h - u_h$ is inversely proportional to τ. Moreover, by the estimate (v) of Proposition 2.1, we have that,

$$\| \, (\widehat{\boldsymbol{q}}_h - \boldsymbol{q}_h) \cdot \boldsymbol{n} \, \|_{L^2(e_K^\tau)} \leq \, C \, h_K^{r+1/2} \, | \, \mathsf{P} f \, |_{H^r(K)},$$

for any $r \in [0, k]$, and we see that the size of jump under consideration depends solely on the smoothness of $f|_K$. For example, if $\mathsf{P} f|_K$ is a polynomial of degree $k - 1$, then $(\widehat{\boldsymbol{q}}_h - \boldsymbol{q}_h) \cdot \boldsymbol{n} = 0$ on e_K^τ. This implies that $(\widehat{\boldsymbol{q}}_h - \boldsymbol{q}_h) \cdot \boldsymbol{n} = 0$ on ∂K for every $K \in \Omega_h$ and, as a consequence, that $\boldsymbol{q}_h \in \boldsymbol{H}(div, \Omega)$. Now, if $f \in H^r(K)$, for some $r \in [0, k]$, then we have that

$$\| \, (\widehat{\boldsymbol{q}}_h - \boldsymbol{q}_h) \cdot \boldsymbol{n} \, \|_{L^2(e_K^\tau)} \leq \, C \, h_K^{r+1/2} \, | \, f \, |_{H^r(K)},$$

by well-known approximation properties of the projection P.

Next, we give a characterization of the approximate solution which follows from a similar result for more general methods obtained in [11]. To state it, we need to introduce the *local solvers* associated with the method. The first local solver is defined on the simplex $K \in \Omega_h$ as the mapping $\mathsf{m} \in L^2(\partial K) \rightarrow (\mathcal{Q}\mathsf{m}, \mathcal{U}\mathsf{m}) \in \boldsymbol{\mathcal{P}}^k(K) \times \mathcal{P}^k(K)$ where

$$(2.5a) \qquad\qquad (\boldsymbol{c}\,\mathcal{Q}\mathsf{m}, \boldsymbol{v})_K - (\mathcal{U}\mathsf{m}, \nabla \cdot \boldsymbol{v})_K = -\langle \mathsf{m}, \boldsymbol{v} \cdot \boldsymbol{n} \rangle_{\partial K},$$

$$(2.5b) \qquad\qquad -(\nabla w, \mathcal{Q}\mathsf{m})_K + \langle w, \widehat{\mathcal{Q}}\mathsf{m} \cdot \boldsymbol{n} \rangle_{\partial K} = 0,$$

for all $(\boldsymbol{v}, w) \in \boldsymbol{\mathcal{P}}^k(K) \times \mathcal{P}^k(K)$, where

$$(2.5c) \qquad\qquad \widehat{\mathcal{Q}}\mathsf{m} = \mathcal{Q}\mathsf{m} + \tau(\mathcal{U}\mathsf{m} - \mathsf{P}_{\partial}\mathsf{m})\boldsymbol{n}.$$

The other local solver is defined on the simplex $K \in \Omega_h$ as the mapping $f \in L^2(K) \rightarrow (\mathcal{Q}f, \mathcal{U}f) \in \boldsymbol{\mathcal{P}}^k(K) \times \mathcal{P}^k(K)$ where

$$(2.6a) \qquad\qquad (\boldsymbol{c}\,\mathcal{Q}f, \boldsymbol{v})_K - (\mathcal{U}f, \nabla \cdot \boldsymbol{v})_K = 0,$$

$$(2.6b) \qquad\qquad -(\nabla w, \mathcal{Q}f)_K + \langle w, \widehat{\mathcal{Q}}f \cdot \boldsymbol{n} \rangle_{\partial K} = (f, w)_K,$$

for all $(\boldsymbol{v}, w) \in \boldsymbol{\mathcal{P}}^k(K) \times \mathcal{P}^k(K)$, where

$$(2.6c) \qquad\qquad \widehat{\mathcal{Q}}f = \mathcal{Q}f + \tau \,\mathcal{U}f\boldsymbol{n}.$$

We can now state our characterization result.

Theorem 2.3. *The approximate solution $(\boldsymbol{q}_h, u_h, \lambda_h) \in \boldsymbol{V}_h \times W_h \times M_h$ given by the SF-H$_k$ method is well defined. Moreover, we have that*

$$(\boldsymbol{q}_h, u_h) = (\mathcal{Q}\lambda_h, \mathcal{U}\lambda_h) + (\mathcal{Q}g, \mathcal{U}g) + (\mathcal{Q}f, \mathcal{U}f),$$

where λ_h can be characterized as the function in M_h satisfying

$$a_h(\lambda_h, \mu) = b_h(\mu) \qquad \forall \mu \in M_h,$$

where

$$a_h(\eta, \mu) := (\boldsymbol{c}\,\mathfrak{Q}\eta, \mathfrak{Q}\mu)_{\Omega_h},$$

$$b_h(\mu) := \langle g, \mathfrak{Q}\mu \cdot \boldsymbol{n}\rangle_{\partial\Omega_D} + (f, \mathfrak{U}\mu)_{\Omega_h} - \langle \mu, \mathsf{q}_{\mathsf{N}}\rangle_{\partial\Omega_N},$$

for all η and $\mu \in M_h$.

This result allows us to shed light on the effect of local stabilization parameters τ on the approximate solution. It will also allow us to compare the RT_k, the BDM_k and the $\mathrm{SF}\text{-}\mathrm{H}_k$ methods, see [10] for a comparison of the hybridized version of the RT_k and the BDM_k methods. These results are gathered in the following theorem. To state it, we use the projection P^{k-1} which is defined by (2.4) for $k \geq 1$ and which we take to be identically zero when $k = 0$. We keep this convention in the remainder of the paper.

Theorem 2.4. *We have that*
(i) *The function $(\boldsymbol{q}_h, \mathsf{P}^{k-1}u_h, \lambda_h)$, is independent of the values of the local stabilization parameters τ. Moreover, changes in the local stabilization parameters τ_K only affect the function $\mathfrak{U}f|_{e_K^\tau}$.*
(ii) *If $\mathsf{P}f|_K \in \mathcal{P}^{k-1}(K)$ for all simplexes $K \in \Omega_h$, then $(\boldsymbol{q}_h, \mathsf{P}^{k-1}u_h, \lambda_h)$ is the same for the RT_k, the BDM_k (if $k \geq 1$) and the $\mathrm{SF}\text{-}\mathrm{H}_k$ methods. Moreover, $u_h|_{e_K^\tau} = \widehat{u}_h|_{e_K^\tau}$ for all $K \in \Omega_h$.*
(iii) *The bilinear form $a_h(\cdot, \cdot)$ is always the same for the RT_k, the BDM_k (if $k \geq 1$) and the $\mathrm{SF}\text{-}\mathrm{H}_k$ methods.*

2.3. **A priori error estimates.** In this subsection, we obtain a priori error estimates for the error of the approximation $(\boldsymbol{q}_h, u_h, \lambda_h) \in \boldsymbol{V}_h \times W_h \times M_h$ given by the $\mathrm{SF}\text{-}\mathrm{H}_k$ and the numerical trace $\widehat{u}_h$ defined by (1.4a). To state them, we need to introduce new notation.

For any real-valued function ζ in $H^l(\Omega_h)$, we set

$$|\zeta|_{H^l(\Omega_h)} := \Big(\sum_{K \in \Omega_h} |\zeta|_{H^l(K)}^2 \Big)^{\frac{1}{2}}.$$

For a vector-valued function $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_d) \in \boldsymbol{H}^l(\Omega_h)$ we set

$$|\boldsymbol{\sigma}|_{\boldsymbol{H}^l(\Omega_h)} := \Big(\sum_{i=1}^d |\sigma_i|_{H^l(\Omega_h)}^2 \Big)^{\frac{1}{2}}.$$

We can now state our results.

We begin by measuring the error in the approximation of the flux $\boldsymbol{q}$ in the norm $\|\boldsymbol{\sigma}\|_{L^2(\Omega_h;\boldsymbol{c})} = (\boldsymbol{c}\,\boldsymbol{\sigma}, \boldsymbol{\sigma})_{\Omega_h}^{1/2}$.

Theorem 2.5. *Suppose that the exact flux $\boldsymbol{q}$ belongs to $\boldsymbol{H}^{r+1}(\Omega_h)$ for some $r \in [0, k]$. Then*

$$\|\boldsymbol{q} - \boldsymbol{q}_h\|_{\boldsymbol{L}^2(\Omega_h;\boldsymbol{c})} \leq \|\boldsymbol{q} - \boldsymbol{\Pi}\boldsymbol{q}\|_{\boldsymbol{L}^2(\Omega_h;\boldsymbol{c})} \leq C\, h^{r+1}\, |\boldsymbol{q}|_{\boldsymbol{H}^{r+1}(\Omega_h)},$$

for some constant C independent of h and the exact solution $(\boldsymbol{q}, u)$.

Note that the upper bound of the error is independent of the local stabilization parameters τ, in complete agreement with the characterization of the approximate solution given by Theorem 2.3. It is interesting to realize that the first estimate also holds for the RT_k and BDM_k methods when the projection $\boldsymbol{\Pi}$ is suitably chosen; see [13] and [4]. Such an estimate is obtained by using the commutativity property and

the fact that the image of their projections is in $\boldsymbol{H}(div, \Omega)$. Since our projection $(\boldsymbol{\Pi}, \boldsymbol{p})$ only satisfies a weak version of the commutativity property, a much more delicate analysis has to be carried out to obtain it.

In [5], it was shown that for general LDG methods with penalization parameters of order $1/h$, the order of convergence of the approximations for flux $\boldsymbol{q}$ using polynomials of degree k is only k; this order is sharp because it is actually attained for some LDG methods. It was also shown that, for DG methods with both penalization parameters of order one, the order of convergence of the approximations for the flux using polynomials of degree k is $(k + 1/2)$. Here, we obtain an order of convergence of $(k + 1)$. *No other DG method has this property.*

Next, we present several estimates for the error in the approximation of the potential u. The first is a superconvergence result. To state it, we need to introduce the adjoint equations

$$(2.7a) \qquad\qquad \boldsymbol{c}\,\boldsymbol{\psi} + \nabla\varphi = 0 \quad \text{in } \Omega,$$

$$(2.7b) \qquad\qquad \nabla \cdot \boldsymbol{\psi} = \theta \quad \text{in } \Omega,$$

$$(2.7c) \qquad\qquad \varphi = 0 \quad \text{on } \partial\Omega_D$$

$$(2.7d) \qquad\qquad \boldsymbol{\psi} \cdot \boldsymbol{n} = 0 \quad \text{on } \partial\Omega_N.$$

We also need to assume that the following elliptic regularity result holds

$$(2.8) \qquad \|\boldsymbol{\psi}\|_{\boldsymbol{H}^{s+1}(\Omega_h)} + \|\nabla\varphi\|_{\boldsymbol{H}^{s+1}(\Omega_h)} \leq \mathsf{C}_{er}\,\|\theta\|_{H^s(\Omega)}$$

for $s \in [0, k]$. Note that, since we are working with domains that can be triangulated by using straight-faced simplexes, the above result only holds if such a domain is convex and $s = 0$. However, we want to write this assumption in such a generality since the method will be extended to domains with smooth curved boundaries in a forthcoming paper.

Theorem 2.6. *Suppose that the exact flux $\boldsymbol{q}$ belongs to $\boldsymbol{H}^{r+1}(\Omega_h)$ for $r \in [0, k]$. Set $\kappa := \max_{K \in \Omega_h} \frac{1}{h_K\,\tau_K}$. Then,*

$$\|\mathbb{P}u - u_h\|_{H^{-s}(\Omega_h)} \leq C\,\mathfrak{C}_{\kappa}^{r,s}(\boldsymbol{q})\,h^{r+s+2},$$

where

$$\mathfrak{C}_{\kappa}^{r,s}(\boldsymbol{q}) = \begin{cases} \mathsf{C}_{er}\,|\boldsymbol{q}|_{\boldsymbol{H}^{r+1}(\Omega_h)} + \kappa\,|\nabla \cdot \boldsymbol{q}|_{H^r(\Omega_h)}, & \text{for } s \in [0, k-1],\, k \geq 1, \\ \mathsf{C}_{er}|\boldsymbol{q}|_{\boldsymbol{H}^1(\Omega_h)} + \mathsf{C}_{er}|\mathsf{q}_{\mathsf{N}}|_{H^1(\partial\Omega_N)}, & \text{for } r = s = k = 0 \text{ and } f = 0. \end{cases}$$

Moreover, for $k = 0$ and general $f \in L^2(\Omega_h)$,

$$\|\mathbb{P}u - u_h\|_{L^2(\Omega_h)} \leq C\,\mathsf{C}_{\kappa}(\boldsymbol{q})\,h,$$

where $\mathsf{C}_{\kappa}(\boldsymbol{q}) = (1 + h)\mathsf{C}_{er}\,|\boldsymbol{q}|_{\boldsymbol{H}^1(\Omega_h)} + h\,\kappa\,|\nabla \cdot \boldsymbol{q}|_{L^2(\Omega_h)}$.

It is interesting to note that the above superconvergence result holds for any choice of local stabilization parameters τ_K such that κ is uniformly bounded, that is, such that $1/(h_K\,\tau_K)$ is uniformly bounded with respect to h. This shows that τ_K cannot be too small for superconvergence to take place.

A straightforward consequence of this theorem is the following result.

Corollary 2.7. *Suppose that the exact flux $(\boldsymbol{q}, u)$ belongs to $\boldsymbol{H}^{r+1}(\Omega_h) \times H^{r+1}(\Omega_h)$ for $r \in [0, k]$. Then*

$$\|u - u_h\|_{L^2(\Omega_h)} \leq C\,h^{r+1}\,\left(\mathsf{C}_{\kappa}(\boldsymbol{q}) + |u|_{H^{r+1}(\Omega_h)}\right),$$

where

$$\mathsf{C}_\kappa(\boldsymbol{q}) = \begin{cases} \min\{\mathfrak{C}_\kappa^{r-1,0}(\boldsymbol{q}),\ h\,\mathfrak{C}_\kappa^{r,0}(\boldsymbol{q})\} & \text{if } k \geq 1, \\ \mathfrak{C}_\kappa(\boldsymbol{q}) & \text{if } k = 0. \end{cases}$$

Note that the above result shows that if $1/\tau_K$ is uniformly bounded for quasi-uniform triangulations, the convergence of u_h is still optimal, provided $\boldsymbol{q}$ is smoother than required, that is, provided $\boldsymbol{q} \in \boldsymbol{H}^{r+1}(\Omega_h)$ instead of just $\boldsymbol{q} \in \boldsymbol{H}^r(\Omega_h)$. Of course, in this case, the superconvergence of u_h to $\mathbb{P}u$ is lost.

The next result is a superconvergence result for the Lagrange multiplier λ_h. To state it, we use the following norm:

$$\| \mathsf{P}_\partial u - \widehat{u}_h \|_{L^2(\mathscr{E}_h;h)} = \big(\sum_{K \in \Omega_h} h_K \| \mathsf{P}_\partial u - \widehat{u}_h \|_{L^2(\partial K)}^2 \big)^{1/2}.$$

Theorem 2.8. *Suppose that the exact solution $(u, \boldsymbol{q})$ of (1.1) belongs to $H^{r+1}(\Omega_h) \times \boldsymbol{H}^{r+1}(\Omega_h)$ for some $r \in [0, k]$. Then,*

$$\|\mathsf{P}_\partial u - \widehat{u}_h\|_{L^2(\mathscr{E}_h;h)} \leq C \left(\mathfrak{C}_0^{r,0}(\boldsymbol{q}) + |\boldsymbol{q}|_{\boldsymbol{H}^{r+1}(\Omega_h)} \right) h^{r+2},$$

if $k \geq 1$, or if $k = 0$ and $f = 0$.

There are no results of this type for any other DG method. However, the RT_k and the BDM_k methods have both similar results. Here we exploited the similarity of the SF-H$_k$ methods with the RT_k and BDM_k methods to obtain these superconvergence results.

2.4. Postprocessing. We end this section by showing how to exploit the superconvergence results to postprocess u_h, $\boldsymbol{q}_h$ and $\hat{u}_h$ to get a better approximation to u defined as follows.

On the simplex K, we define the new approximation of u, $u_h^\star$, as the function of $\mathcal{P}^{k+1}(K)$ given by

$$(2.9\text{a}) \qquad\qquad u_h^\star = \overline{u}_h + \tilde{u}_h,$$

where

$$(2.9\text{b}) \qquad\qquad \overline{u}_h = \begin{cases} \frac{1}{d} \sum_{e \in \partial K} \widehat{u}_h|_e & \text{if } k = 0, \\ \frac{1}{|K|} \int_K u_h\, dx & \text{if } k > 0, \end{cases}$$

and $\tilde{u}_h$ is the polynomial in $\mathcal{P}_0^{k+1}(K)$ satisfying

$$(2.9\text{c}) \qquad (\boldsymbol{a}\,\nabla\tilde{u}_h, \nabla w)_K = (f, w)_K - \langle w, \widehat{\boldsymbol{q}}_h \cdot \boldsymbol{n}\rangle_{\partial K} \quad \forall w \in \mathcal{P}^{k+1}(K).$$

Here $\boldsymbol{a} = \boldsymbol{c}^{-1}$ and $\mathcal{P}_0^{k+1}(K)$ is the collection of functions in $\mathcal{P}^{k+1}(K)$ with mean zero. The postprocessing technique just introduced is a slight modification of a postprocessing proposed in [17, 18] and [14]; it consists of using the numerical trace $\widehat{\boldsymbol{q}}_h$ instead of $\boldsymbol{q}_h$.

It is easy to see that this postprocessing is associated to a locally conservative method. Indeed, the scheme satisfied by $u_h^\star$ on each simplex $K \in \Omega_h$ is

$$(\boldsymbol{a}\,\nabla u_h^\star, \nabla w)_K + \langle w, \widehat{\boldsymbol{q}}_h \cdot \boldsymbol{n}\rangle_{\partial K} = (f, w)_K \quad \forall w \in \mathcal{P}^{k+1}(K).$$

As a consequence, if we take D_h to be the union of an arbitrary set of simplexes $K \in \Omega_h$, we get that

$$\langle 1, \widehat{\boldsymbol{q}}_h \cdot \boldsymbol{n}\rangle_{\partial D_h} = (f, 1)_{D_h},$$

which is nothing but the property of local conservativity.

Note that $\tilde{u}_h$ is well defined. Indeed, if we take $w = 1$ in equation (2.9c), the right-hand side is also equal to zero thanks to equation (1.3b). The fact that it provides a better approximation to the potential u than u_h is contained in the following result.

Theorem 2.9. *Suppose that the exact solution $(u, \boldsymbol{q})$ belongs to $H^{r+2}(\Omega_h) \times \boldsymbol{H}^{r+1}(\Omega_h)$ for $r \in [0, k]$. Then, if $k \geq 1$,*

$$\|u - u_h^\star\|_{L^2(\Omega_h)} \leq C\, h^{r+2} \left(\mathfrak{C}_0^{r,0}(\boldsymbol{q}) + |\boldsymbol{q}|_{\boldsymbol{H}^{r+1}(\Omega_h)} + |u|_{H^{r+2}(\Omega_h)} \right),$$

and if $k = 0$ and $f \equiv 0$,

$$\|u - u_h^\star\|_{L^2(\Omega_h)} \leq C\, h^2 \left(\mathfrak{C}_0(\boldsymbol{q}) + |u|_{H^2(\Omega_h)} \right).$$

Note that when $\mathsf{P}f|_K \in \mathcal{P}^{k-1}(K)$ for all $K \in \Omega_h$, by Theorem 2.4 we have that the function $(\widehat{\boldsymbol{q}}_h \cdot \boldsymbol{n}, \mathsf{P}^{k-1}u_h, \lambda_h)$ is the same for the RT_k, BDM_k ($k \geq 1$) and $\mathrm{SF\text{-}H}_k$ methods. As a consequence, the postprocessed approximation $u_h^\star$ is also the same for all these methods.

Note also that in [3], a general postprocessing which is solely based on approximation results was obtained. When applied to the $\mathrm{SF\text{-}H}_k$ method for $k \geq 1$, it gives rise to an approximation of u which converges with the same orders as ours. However, unlike such postprocessing, our postprocessed solution $u_h^\star$ is associated to a locally conservative scheme; it is also easier to compute.

Let us end this section by noting that all the error estimates for $k \geq 1$ hold if in the equation (1.3b), we replace f by any function $\mathfrak{I}_h f$ such that $\mathfrak{I}_h f|_K \in \mathcal{P}^{k-1}(K)$ for all $K \in \Omega_h$ and such that

$$\| f - \mathfrak{I}_h f \|_{H^{-1}(\Omega)} \leq C\, h^{r+1} |f|_{H^r(\Omega_h)}.$$

Moreover, by statement (ii) of Theorem 2.4, the function $(\boldsymbol{q}_h, \mathsf{P}^{k-1}u_h, \lambda_h)$ provided by the RT_k, the BDM_k and the $\mathrm{SF\text{-}H}_k$ method is the same; in particular, we have that $\boldsymbol{q}_h \in \boldsymbol{H}(div, \Omega)$. The postprocessed approximation $u_h^\star$ is also the same for those three methods.

3. Proofs

In this section, we present detailed proofs of all our results.

3.1. Proof of Proposition 2.1: The properties of $(\boldsymbol{\Pi}, \mathbb{P})$.

3.1.1. *Two key auxiliary results about polynomials.* To prove Proposition 2.1, we begin by stating and proving two lemmas whose use is crucial in our analysis.

Lemma 3.1. *Given the face e of the simplex K and a function $z \in \mathcal{P}^k(e)$, there is a unique function $Z \in \mathcal{P}^k(K)$ such that*

$$\text{(i)} \quad \langle Z, \omega \rangle_e = \langle z, \omega \rangle_e \qquad \forall\, \omega \in \mathcal{P}^k(e),$$

$$\text{(ii)} \quad (Z, \mathsf{w})_K = 0 \qquad \forall\, \mathsf{w} \in \mathcal{P}^{k-1}(K).$$

Moreover,

$$\text{(iii)} \quad \| Z \|_{L^2(K)} \leq C\, h_K^{1/2} \| z \|_{L^2(e)},$$

where h_K is the diameter of the simplex K and C depends solely on k and the shape-regularity parameters of the simplex K.

Lemma 3.2. *Given the face e of the simplex K and the function z such that for all faces e' of K different from e, $z|_{e'} \in \mathcal{P}^k(e')$, there is a unique function $\boldsymbol{Z} \in \boldsymbol{\mathcal{P}}^k(K)$ such that*

$$\text{(i)} \quad \langle \boldsymbol{Z} \cdot \boldsymbol{n}, \omega \rangle_{e'} = \langle z, \omega \rangle_{e'} \qquad \forall\, \omega \in \mathcal{P}^k(e'),\ e' \neq e,$$

$$\text{(ii)} \quad (\boldsymbol{Z}, \boldsymbol{v})_K = 0 \qquad \forall\, \boldsymbol{v} \in \boldsymbol{\mathcal{P}}^{k-1}(K).$$

Moreover,

$$\text{(iii)} \quad \|\,\boldsymbol{Z}\,\|_{\boldsymbol{L}^2(K)} \leq C\, h_K^{1/2} \|\, z\,\|_{L^2(\partial K \setminus e)},$$

where h_K is the diameter of the simplex K and C depends solely on k and the shape-regularity parameters of the simplex K.

We are only going to give a detailed proof of Lemma 3.2 since the proof of Lemma 3.1 is similar and simpler.

Proof of Lemma 3.2. Let us begin by proving that the function $\boldsymbol{\sigma} \in \boldsymbol{\mathcal{P}}^k(K)$ satisfying (i) and (ii) exists and is unique. Since the linear system determined by equations (i) and (ii) is square, indeed,

$$\dim(\boldsymbol{\mathcal{P}}^{k-1}(K)) = \binom{k-1+d}{d} \times d,$$

$$\sum_{e' \in \partial K, e' \neq e} \dim(\mathcal{P}^k(e')) = \binom{k+d-1}{d-1} \times d,$$

$$\dim(\boldsymbol{\mathcal{P}}^k(K)) = \binom{k+d}{d} \times d,$$

and $\binom{k-1+d}{d} + \binom{k+d-1}{d-1} = \binom{k+d}{d}$, we only need to show that if $\boldsymbol{\sigma} \in \boldsymbol{\mathcal{P}}^k(K)$ satisfies

$$(\boldsymbol{Z}, \boldsymbol{v})_K = 0 \qquad \forall\, \boldsymbol{v} \in \boldsymbol{\mathcal{P}}^{k-1}(K),$$

$$\langle \boldsymbol{Z} \cdot \boldsymbol{n}, \omega \rangle_{e'} = 0 \qquad \forall\, \omega \in \mathcal{P}^k(e'), e' \in \partial K, e' \neq e,$$

then $\boldsymbol{Z} = \boldsymbol{0}$ on K.

Let T be the affine mapping that transforms the element K to the reference simplex $\widehat{K}$. Moreover, let us denote by e_i, $i = 1, \ldots, d+1$, the faces of K where $e := e_{d+1}$. Assume that the mapping T is such that $\widehat{e}_i := T(e_i)$ is the face of $\widehat{K}$ lying on the plane $\widehat{x}_i = 0$, and set $\widehat{\boldsymbol{Z}}(\widehat{x}) := \boldsymbol{Z}(T^{-1}(\widehat{x}))$. Then the above equations become

$$(\widehat{\boldsymbol{Z}}, \widehat{\boldsymbol{v}})_{\widehat{K}} = 0 \qquad \forall\, \widehat{\boldsymbol{v}} \in \boldsymbol{\mathcal{P}}^{k-1}(\widehat{K}),$$

$$\langle \widehat{\boldsymbol{Z}} \cdot \boldsymbol{n}_i, \widehat{\omega} \rangle_{\widehat{e}_i} = 0 \qquad \forall\, \widehat{\omega} \in \mathcal{P}^k(\widehat{e}_i),\ i = 1, ..., d,$$

since spaces of polynomials of a given total degree are invariant under affine transformations. Now, let $\{\widehat{\boldsymbol{n}}_j\}_{j=1}^d$ be the basis of $\mathbb{R}^d$ dual to $\{\boldsymbol{n}_i\}_{i=1}^d$, that is, $\boldsymbol{n}_i \cdot \widehat{\boldsymbol{n}}_j = \delta_{ij}$. Then we can write $\widehat{\boldsymbol{Z}} = \sum_{j=1}^d \widehat{p}_j \widehat{\boldsymbol{n}}_j$, where $\widehat{p}_j \in \mathcal{P}^k(\widehat{K})$, and obtain that

$$\sum_{j=1}^d (\widehat{p}_j, \widehat{\boldsymbol{n}}_j \cdot \widehat{\boldsymbol{v}})_{\widehat{K}} = 0 \qquad \forall\, \widehat{\boldsymbol{v}} \in \boldsymbol{\mathcal{P}}^{k-1}(\widehat{K}),$$

$$\langle \widehat{p}_i, \widehat{\omega} \rangle_{\widehat{e}_i} = 0 \qquad \forall\, \widehat{\omega} \in \mathcal{P}^k(\widehat{e}_i),\ i = 1, ..., d.$$

The last equation implies that, for any $i = 1, ..., d$, $\widehat{p}_i|_{\widehat{e}_i} = 0$ and hence that $\widehat{p}_i = \widehat{x}_i \widehat{\mathsf{p}}_i$ for some polynomial $\widehat{\mathsf{p}}_i$ in $\mathcal{P}^{k-1}(\widehat{K})$. Taking $\widehat{\boldsymbol{v}} = \widehat{\mathsf{p}}_i \, \boldsymbol{n}_i$, we get

$$(\widehat{p}_i, \widehat{\mathsf{p}}_i)_{\widehat{K}} = (\widehat{x}_i, \widehat{\mathsf{p}}_i^2)_{\widehat{K}} = 0,$$

and, since $\widehat{x}_i > 0$ on $\widehat{K}$, we conclude that $\widehat{\mathsf{p}}_i = 0$. This implies that $\boldsymbol{Z} = \boldsymbol{0}$ on K. This proves the existence and uniqueness of $\boldsymbol{Z}$ satisfying the conditions (i) and (ii).

The estimate (iii) follows now from a simple scaling argument. This completes the proof. $\qquad\square$

3.1.2. *Proof of the orthogonality properties.*
It is not difficult to see that the fact that $(\boldsymbol{\Pi}, \mathbb{P})$ is well defined is a direct corollary of Lemmas 3.1 and 3.2.

Now, let us prove the orthogonality properties. The property (i) follows from the property (2.2a) defining $\mathbb{P}$ and the orthogonality property (ii) follows from the property (2.1a) defining $\boldsymbol{\Pi}$. The orthogonality property (iii) follows from the properties (2.2b) and (2.1b) defining $\mathbb{P}$ and $\boldsymbol{\Pi}$, and from the definition of the projection P_∂, (1.5). In fact, it follows from the fact that on each face e of any simplex K, we have that either $\mathbb{P}\zeta = \mathsf{P}_\partial\zeta$ or $\boldsymbol{\Pi}\boldsymbol{\sigma} \cdot \boldsymbol{n} = \mathsf{P}_\partial\boldsymbol{\sigma} \cdot \boldsymbol{n}$.

3.1.3. *Proof of the weak commutativity property.*
The weak commutativity property (iv) is a direct consequence of the three orthogonality properties we just proved. Indeed, we have that

$$
\begin{aligned}
-(\nabla\zeta, \boldsymbol{\Pi}\boldsymbol{\sigma})_K &= (\zeta, \nabla \cdot \boldsymbol{\Pi}\boldsymbol{\sigma})_K - \langle\zeta, \boldsymbol{\Pi}\boldsymbol{\sigma} \cdot \boldsymbol{n}\rangle_{\partial K} \\
&= (\mathbb{P}\zeta, \nabla \cdot \boldsymbol{\Pi}\boldsymbol{\sigma})_K - \langle\zeta, \boldsymbol{\Pi}\boldsymbol{\sigma} \cdot \boldsymbol{n}\rangle_{\partial K} && \text{by (i)}, \\
&= -(\nabla\mathbb{P}\zeta, \boldsymbol{\Pi}\boldsymbol{\sigma})_K + \langle\mathbb{P}\zeta - \zeta, \boldsymbol{\Pi}\boldsymbol{\sigma} \cdot \boldsymbol{n}\rangle_{\partial K} \\
&= -(\nabla\mathbb{P}\zeta, \boldsymbol{\sigma})_K + \langle\mathbb{P}\zeta - \zeta, \boldsymbol{\Pi}\boldsymbol{\sigma} \cdot \boldsymbol{n}\rangle_{\partial K} && \text{by (ii)}, \\
&= -(\nabla\mathbb{P}\zeta, \boldsymbol{\sigma})_K + \langle\mathbb{P}\zeta - \mathsf{P}_\partial\zeta, \boldsymbol{\sigma} \cdot \boldsymbol{n}\rangle_{\partial K} && \text{by (iii)}, \\
&= (\mathbb{P}\zeta, \nabla \cdot \boldsymbol{\sigma})_K - \langle\mathsf{P}_\partial\zeta, \boldsymbol{\sigma} \cdot \boldsymbol{n}\rangle_{\partial K}, \\
&= (\mathbb{P}\zeta, \nabla \cdot \boldsymbol{\sigma})_K - \langle\mathsf{P}_\partial\zeta, \mathsf{P}_\partial\boldsymbol{\sigma} \cdot \boldsymbol{n}\rangle_{\partial K},
\end{aligned}
$$

by the definition of the projection P_∂, (1.5). This completes the proof of (iv).

3.1.4. *Proof of the estimates* (v) *and* (vi).
Note that, by the definition of the projections $\boldsymbol{\Pi}$, (2.1), and $\boldsymbol{\Pi}^{\mathrm{RT}}$, (2.3), we have that

$$
\begin{aligned}
(\boldsymbol{\Pi}^{\mathrm{RT}}\boldsymbol{\sigma} - \boldsymbol{\Pi}\boldsymbol{\sigma}, \boldsymbol{v})_K &= 0, \\
\langle(\boldsymbol{\Pi}^{\mathrm{RT}}\boldsymbol{\sigma} - \boldsymbol{\Pi}\boldsymbol{\sigma}) \cdot \boldsymbol{n}, \omega\rangle_e &= \langle\mathsf{P}_\partial\boldsymbol{\sigma} \cdot \boldsymbol{n} - \boldsymbol{\Pi}\boldsymbol{\sigma} \cdot \boldsymbol{n}, \omega\rangle_{e \cap e_K^\tau}.
\end{aligned}
$$

for all $\boldsymbol{v} \in \boldsymbol{\mathcal{P}}^{k-1}(K)$, if $k > 1$, and for all $\omega \in \mathcal{P}^k(e)$ and all faces e of K. By a well-known scaling argument, we immediately obtain that

$$\|\boldsymbol{\Pi}^{\mathrm{RT}}\boldsymbol{\sigma} - \boldsymbol{\Pi}\boldsymbol{\sigma}\|_{\boldsymbol{L}^2(K)} \leq C\,h_K^{1/2}\,\|\mathsf{P}_\partial\boldsymbol{\sigma} \cdot \boldsymbol{n} - \boldsymbol{\Pi}\boldsymbol{\sigma} \cdot \boldsymbol{n}\|_{L^2(e_K^\tau)}.$$

It remains to estimate the above right-hand side. To do that, we note that, for any ω in $\mathcal{P}^k(K)$, we have that

$$\langle\omega, \boldsymbol{\Pi}\boldsymbol{\sigma} \cdot \boldsymbol{n} - \mathsf{P}_\partial\boldsymbol{\sigma} \cdot \boldsymbol{n}\rangle_{e_K^\tau} = \langle\omega, (\boldsymbol{\Pi}\boldsymbol{\sigma} - \boldsymbol{\sigma}) \cdot \boldsymbol{n}\rangle_{\partial K} = (\omega, \nabla \cdot (\boldsymbol{\Pi}\boldsymbol{\sigma} - \boldsymbol{\sigma}))_K,$$

by the definition of the projections $\mathbf{\Pi}$, (2.1). Taking $\omega = Z$, where Z is given by Lemma 3.1 with $z := \mathbf{\Pi}\boldsymbol{\sigma} \cdot \boldsymbol{n} - \mathsf{P}_\partial \boldsymbol{\sigma} \cdot \boldsymbol{n}$, we get that, for any p in $\mathcal{P}^{k-1}(K)$,

$$
\begin{aligned}
\| z \|_{L^2(e_K^\tau)}^2 &= \| \mathbf{\Pi}\boldsymbol{\sigma} \cdot \boldsymbol{n} - \mathsf{P}_\partial \boldsymbol{\sigma} \cdot \boldsymbol{n} \|_{L^2(e_K^\tau)}^2 \\
&= -(\mathsf{P}\nabla \cdot \boldsymbol{\sigma} - p, Z)_K \\
&\leq C\, h_K^{1/2} \| \mathsf{P}\nabla \cdot \boldsymbol{\sigma} - p \|_{L^2(K)} \| z \|_{L^2(e_K^\tau)},
\end{aligned}
$$

and, after a direct application of the Bramble-Hilbert lemma, we get

$$
\| \mathsf{P}_\partial \boldsymbol{\sigma} \cdot \boldsymbol{n} - \mathbf{\Pi}\boldsymbol{\sigma} \cdot \boldsymbol{n} \|_{L^2(e_K^\tau)} \leq C\, h_K^{r+1/2} \, | \mathsf{P}\nabla \cdot \boldsymbol{\sigma} |_{H^r(K)},
$$

where $r \in [0, k]$. This completes the proof of the estimates (v) and (vi).

3.1.5. *Proof of the estimate* (vii). Note that, by the definition of the projections $\mathbb{P}$, (2.2), and P, (2.4), we have that

$$
\begin{aligned}
(\mathbb{P}\zeta - \mathsf{P}\zeta, \mathsf{w})_K &= 0, \\
\langle \mathbb{P}\zeta - \mathsf{P}\zeta, \omega \rangle_{e_K^\tau} &= \langle \mathsf{P}_\partial \zeta - \mathsf{P}\zeta, \omega \rangle_{e_K^\tau},
\end{aligned}
$$

for all $\mathsf{w} \in \mathcal{P}^{k-1}(K)$, if $k \geq 1$, and for all $\omega \in \mathcal{P}^k(e_K^\tau)$. This implies that Lemma 3.1 holds with $z := \mathsf{P}_\partial \zeta - \mathsf{P}\zeta$ and $Z = \mathbb{P}\zeta - \mathsf{P}\zeta$. As a consequence,

$$
\| \mathbb{P}\zeta - \mathsf{P}\zeta \|_{L^2(K)} \leq C\, h_K^{1/2} \| \mathsf{P}_\partial \zeta - \mathsf{P}\zeta \|_{L^2(e_K^\tau)}^2.
$$

It remains to estimate the above right-hand side.

To do that, we note that, for any $\boldsymbol{v}$ in $\boldsymbol{\mathcal{P}}^k(K) \oplus \boldsymbol{x}\, \mathcal{P}^k(K)$, we have that

$$
(3.1) \qquad \langle \mathsf{P}\zeta - \mathsf{P}_\partial \zeta, \boldsymbol{v} \cdot \boldsymbol{n} \rangle_{\partial K} = \langle \mathsf{P}\zeta - \zeta, \boldsymbol{v} \cdot \boldsymbol{n} \rangle_{\partial K} = (\nabla \mathsf{P}\zeta - \nabla \zeta, \boldsymbol{v})_K,
$$

by the definition of the projection P, (2.4) and that of the projection P_∂, (1.5). A well-known scaling argument states that given any function z such that its restriction to each face e of k belongs to $\mathcal{P}^k(e)$, there is a function $\boldsymbol{Z}$ in $\boldsymbol{\mathcal{P}}^k(K) \oplus \boldsymbol{x}\, \mathcal{P}^k(K)$ such that

$$
\begin{aligned}
&\text{(i)} \quad \langle \boldsymbol{Z} \cdot \boldsymbol{n}, \omega \rangle_e = \langle z, \omega \rangle_e \qquad \forall\, \omega \in \mathcal{P}^k(e), \\
&\text{(ii)} \quad (\boldsymbol{Z}, \boldsymbol{v})_K = 0 \qquad \forall\, \boldsymbol{v} \in \boldsymbol{\mathcal{P}}^{k-1}(K). \\
&\text{(iii)} \quad \| \boldsymbol{Z} \|_{\boldsymbol{L}^2(K)} \leq C\, h_K^{1/2} \| z \|_{L^2(\partial K)},
\end{aligned}
$$

where h_K is the diameter of the simplex K and C depends solely on the shape-regularity constants of the simplex K. Taking $\boldsymbol{v} := \boldsymbol{Z}$ with $z = \mathsf{P}\zeta - \mathsf{P}_\partial \zeta$ in equation (3.1), we obtain that

$$
\begin{aligned}
\| z \|_{L^2(\partial K)}^2 &= (\nabla \mathsf{P}\zeta - \nabla \zeta, \boldsymbol{Z})_K \\
&= (\boldsymbol{p} - \nabla \zeta, \boldsymbol{Z})_K \\
&\leq C\, h_K^{1/2} \| \nabla \zeta - \boldsymbol{p} \|_{L^2(K)} \| z \|_{L^2(\partial K)},
\end{aligned}
$$

for any $\boldsymbol{p} \in \boldsymbol{\mathcal{P}}^{k-1}(K)$. Thus, after a direct application of the Bramble-Hilbert lemma, we get

$$
\| \mathsf{P}\zeta - \mathsf{P}_\partial \zeta \|_{L^2(\partial K)} \leq C\, h_K^{r+1/2} \, | \nabla \zeta |_{\boldsymbol{H}^r(K)},
$$

where $r \in [0, k]$. This completes the proof of estimate (vii).

This completes the proof of Proposition 2.1.

3.2. **Characterization of the approximate solution.** To prove the results of the characterization of the approximate solution of the SF-H method, we begin by proving two auxiliary results concerning key properties of the local solvers.

3.2.1. *Two auxiliary results about the local solvers.* To state the first auxiliary result, we need to introduce the following decomposition of our local spaces:

$$\boldsymbol{\mathcal{P}}^k(K) = \boldsymbol{\mathcal{V}}(K) \oplus \boldsymbol{\mathcal{V}}^\perp(K),$$
$$\mathcal{P}^k(K) = \mathcal{W}(K) \oplus \mathcal{W}^\perp(K),$$

where

$$\boldsymbol{\mathcal{V}}(K) := \{\boldsymbol{v} \in \boldsymbol{\mathcal{P}}^k(K) : \ \nabla \cdot \boldsymbol{v} = 0\},$$
$$\boldsymbol{\mathcal{V}}^\perp(K) := \{\boldsymbol{v} \in \boldsymbol{\mathcal{P}}^k(K) : \ (\boldsymbol{c}\,\boldsymbol{v}, \boldsymbol{\sigma})_K = 0 \ \forall\, \boldsymbol{\sigma} \in \boldsymbol{\mathcal{V}}(K)\},$$

and

$$\mathcal{W}(K) := \mathcal{P}^{k-1}(K),$$
$$\mathcal{W}^\perp(K) := \{\omega \in \mathcal{P}^k(K) : \ (\omega, \zeta)_K = 0 \ \forall\, \zeta \in \mathcal{W}(K)\},$$

Lemma 3.3. *Let K be any simplex of the triangulation Ω_h. Then the local mapping $(\boldsymbol{\mathcal{Q}}\mathsf{m}, \mathcal{U}\mathsf{m})$ given by equations (2.5) can be obtained as follows:*

(i) *Set*
$$\mathcal{U}\mathsf{m}|_{e_K^\tau} = \mathsf{P}_\partial \mathsf{m}|_{e_K^\tau}.$$

(ii) *Compute $\boldsymbol{\mathcal{Q}}\mathsf{m} \in \boldsymbol{\mathcal{V}}(K)$ by solving*
$$(\boldsymbol{c}\,\boldsymbol{\mathcal{Q}}\mathsf{m}, \boldsymbol{v})_K = -\langle \mathsf{m}, \boldsymbol{v} \cdot \boldsymbol{n}\rangle_{\partial K} \quad \forall\, \boldsymbol{v} \in \boldsymbol{\mathcal{V}}(K).$$

(iii) *Compute $\mathsf{P}^{k-1}\,\mathcal{U}\mathsf{m}$ by solving*
$$(\mathsf{P}^{k-1}\,\mathcal{U}\mathsf{m}, \nabla \cdot \boldsymbol{v})_K = \langle \mathsf{m}, \boldsymbol{v} \cdot \boldsymbol{n}\rangle_{\partial K} \quad \forall\, \boldsymbol{v} \in \boldsymbol{\mathcal{V}}^\perp(K).$$

Similarly, the local mapping $(\boldsymbol{\mathcal{Q}}f, \mathcal{U}f)$ given by equations (2.6) can be obtained as follows:

(α) *Compute $\mathcal{U}f|_{e_K^\tau}$ by solving*
$$\langle \omega, \tau\,\mathcal{U}f\rangle_{e_K^\tau} = (f, \omega)_K \quad \forall\, \omega \in \mathcal{W}^\perp(K).$$

(β) *Compute $\boldsymbol{\mathcal{Q}}f \in \boldsymbol{\mathcal{V}}^\perp(K)$ by solving*
$$(\omega, \nabla \cdot \boldsymbol{\mathcal{Q}}f)_K = -\langle \omega, \tau\,\mathcal{U}f\rangle_{e_K^\tau} + (f, \omega)_K \quad \forall\, \omega \in \mathcal{W}(K).$$

(γ) *Compute $\mathsf{P}^{k-1}\,\mathcal{U}f$ by solving*
$$(\mathsf{P}^{k-1}\,\mathcal{U}f, \nabla \cdot \boldsymbol{v})_K = (\boldsymbol{c}\,\boldsymbol{\mathcal{Q}}f, \boldsymbol{v})_K \quad \forall\, \boldsymbol{v} \in \boldsymbol{\mathcal{V}}^\perp(K).$$

Proof. Let us begin by proving the properties of the first local mapping. Thus, integrating by parts in the equation (2.5b), we obtain

$$(\omega, \nabla \cdot \boldsymbol{\mathcal{Q}}\mathsf{m})_K = \langle \omega, (\widehat{\boldsymbol{\mathcal{Q}}}\mathsf{m} - \boldsymbol{\mathcal{Q}}\mathsf{m}) \cdot \boldsymbol{n}\rangle_{\partial K} = \langle \omega, \tau(\,\mathcal{U}\mathsf{m} - \mathsf{m})\rangle_{e_K^\tau}$$

for all $\omega \in \mathcal{P}^k(K)$, by the definition of the numerical trace $\widehat{\boldsymbol{\mathcal{Q}}}\mathsf{m}$, (1.4b) and (1.6). Taking $\omega \in \mathcal{W}^\perp(K)$, we see that

$$\langle \omega, \tau(\,\mathcal{U}\mathsf{m} - \mathsf{m})\rangle_{e_K^\tau} = 0 \quad \forall\, \omega \in \mathcal{W}^\perp(K).$$

Using the fact that $\mathcal{W}^\perp(K)|_{e_K^\tau} = \mathcal{P}^k(e_K^\tau)$ which follows by a simple application of Lemma 3.1, we have that (i) holds. As a consequence, we see that

$$(\omega, \nabla \cdot \mathbf{Q}\mathsf{m})_K = 0 \quad \forall\, \omega \in \mathcal{P}^k(K),$$

and hence that $\mathbf{Q}\mathsf{m} \in \mathbf{\mathcal{V}}(K)$. The property (ii) can now be obtained by restricting the test functions $\boldsymbol{v}$ to the space $\mathbf{\mathcal{V}}(K)$ in the equation (2.5a). Now that we know $\mathbf{Q}\mathsf{m}$, we obtain the formulation (iii) for $\mathsf{P}^{k-1}\,\mathcal{U}\mathsf{m}$ by restricting the test functions $\boldsymbol{v}$ to the space $\mathbf{\mathcal{V}}^\perp(K)$ in the equation (2.5a). It remains to show that $\mathsf{P}^{k-1}\,\mathcal{U}\mathsf{m}$ is uniquely defined by those equations. But this follows from the fact that the system of equations is square and that $\nabla \cdot \mathbf{\mathcal{V}}^\perp(K) = \mathcal{P}^{k-1}(K) = \mathcal{W}(K)$, which in turn follows from the fact that $\nabla \cdot \mathbf{\mathcal{V}}(K) = \{0\}$ and $\nabla \cdot \mathbf{\mathcal{P}}^k(K) = \mathcal{P}^{k-1}(K)$. This completes the proof of the properties of the first local lifting.

The proof the properties (α) and (γ) of the second local mapping is similar to the proof of the properties (i) and (iii) of the first local mapping, respectively. Let us prove property (β). If we take $\boldsymbol{v} \in \mathbf{\mathcal{V}}$ in the equation (2.6a), we see that $\mathbf{Q}f \in \mathbf{\mathcal{V}}^\perp(K)$. Since the equation in (β) is obtained from (2.6b) by restricting the tests functions ω to $\mathcal{W}(K)$, we only have to prove that $\mathbf{Q}f$ given by (β) is well defined. But this follows from the fact that the system is a square system and $\nabla \cdot \mathbf{\mathcal{V}}^\perp(K) = \mathcal{P}^{k-1}(K) = \mathcal{W}(K)$. This completes the proof. $\qquad\square$

The second auxiliary result concerns the jumps of the local solvers.

Lemma 3.4. *For each simplex $K \in \Omega_h$, we have that, on ∂K,*

$$(\widehat{\mathbf{Q}}\mathsf{m} - \mathbf{Q}\mathsf{m}) \cdot \boldsymbol{n} = \tau\,(\mathcal{U}\mathsf{m} - \mathsf{P}_\partial \mathsf{m}) = 0,$$

$$(\widehat{\mathbf{Q}}f - \mathbf{Q}f) \cdot \boldsymbol{n} = \tau\,\mathcal{U}f = \mathsf{P}_\partial \boldsymbol{q} \cdot \boldsymbol{n} - \mathbf{\Pi}\boldsymbol{q} \cdot \boldsymbol{n}.$$

Proof. Let us begin by proving the second identity since its proof is more involved. Taking $\omega = Z$ in the identity (β) of Lemma 3.3, where Z is given by Lemma 3.1 with $e = e_K^\tau$, we obtain that

$$
\begin{aligned}
\langle z, (\widehat{\mathbf{Q}}f - \mathbf{Q}f) \cdot \boldsymbol{n} \rangle_{e_K^\tau} &= \langle z, \tau\,\mathcal{U}f \rangle_{e_K^\tau} \\
&= (Z, \nabla \cdot (\boldsymbol{q} - \mathbf{\Pi}\boldsymbol{q}))_K \\
&= -(\nabla Z, \boldsymbol{q} - \mathbf{\Pi}\boldsymbol{q})_K + \langle Z, (\boldsymbol{q} - \mathbf{\Pi}\boldsymbol{q}) \cdot \boldsymbol{n} \rangle_{\partial K} \\
&= \langle z, \mathsf{P}_\partial \boldsymbol{q} \cdot \boldsymbol{n} - \mathbf{\Pi}\boldsymbol{q} \cdot \boldsymbol{n} \rangle_{e_K^\tau},
\end{aligned}
$$

by the properties of the projection $\mathbf{\Pi}$, (2.1). As a consequence, we immediately obtain that, on e_K^τ,

$$(\widehat{\mathbf{Q}}f - \mathbf{Q}f) \cdot \boldsymbol{n} = \tau_K\,\mathcal{U}f = \mathsf{P}_\partial \boldsymbol{q} \cdot \boldsymbol{n} - \mathbf{\Pi}\boldsymbol{q} \cdot \boldsymbol{n}.$$

A similar argument gives that, on e_K^τ,

$$(\widehat{\mathbf{Q}}\mathsf{m} - \mathbf{Q}\mathsf{m}) \cdot \boldsymbol{n} = \tau_K\,(\mathcal{U}\mathsf{m} - \mathsf{P}_\partial \mathsf{m}) = 0.$$

This completes the proof of Lemma 3.4. $\qquad\square$

3.2.2. *Proof of Theorem 2.3: Characterization of the approximate solution.* The following result is a particular case of a general result proven in [11].

Theorem 3.5. *The approximate solution $(\boldsymbol{q}_h, u_h, \lambda_h) \in \boldsymbol{V}_h \times W_h \times M_h$ given by the* SF-H$_k$ *method is well defined. Moreover, we have that*

$$(\boldsymbol{q}_h, u_h) = (\mathbf{Q}\lambda_h, \mathcal{U}\lambda_h) + (\mathbf{Q}g, \mathcal{U}g) + (\mathbf{Q}f, \mathcal{U}f),$$

where λ_h can be characterized as the function in M_h satisfying

$$a_h(\lambda_h, \mu) = b_h(\mu) \qquad \forall \mu \in M_h,$$

where

$$a_h(\eta, \mu) = (c\,\mathbf{Q}\eta, \mathbf{Q}\mu)_{\Omega_h} \quad - \langle \mu - \mathcal{U}\mu, (\widehat{\mathbf{Q}}\eta - \mathbf{Q}\eta) \cdot \boldsymbol{n} \rangle_{\partial\Omega_h},$$

$$b_h(\mu) = \langle g, \widehat{\mathbf{Q}}\mu \cdot \boldsymbol{n} \rangle_{\partial\Omega_D} + (f, \mathcal{U}\mu)_{\Omega_h} - \langle \mu, \mathsf{q_N} \rangle_{\partial\Omega_N}$$
$$- \langle \mu - \mathcal{U}\mu, (\widehat{\mathbf{Q}}f - \mathbf{Q}f) \cdot \boldsymbol{n} \rangle_{\partial\Omega_h}$$
$$+ \langle -\mathcal{U}f, (\widehat{\mathbf{Q}}\mu - \mathbf{Q}\mu) \cdot \boldsymbol{n} \rangle_{\partial\Omega_h}$$
$$+ \langle \mu - \mathcal{U}\mu, (\widehat{\mathbf{Q}}g - \mathbf{Q}g) \cdot \boldsymbol{n} \rangle_{\partial\Omega_h}$$
$$- \langle g - \mathcal{U}g, (\widehat{\mathbf{Q}}\mu - \mathbf{Q}\mu) \cdot \boldsymbol{n} \rangle_{\partial\Omega_h},$$

for all η and $\mu \in M_h$.

Theorem 2.3 follows from this result if we show that on $\mathscr{E}_h$,

$$(\widehat{\mathbf{Q}}\mu - \mathbf{Q}\mu) \cdot \boldsymbol{n} = 0 \quad \text{and} \quad (\mu - \mathcal{U}\mu)(\widehat{\mathbf{Q}}f - \mathbf{Q}f) = 0,$$

for all $\mu \in M_h$. Since this is a straightforward consequence of Lemma 3.4, this completes the proof of Theorem 2.3.

3.2.3. *Proof of Proposition* 2.2: *Characterization of the jumps.* By the definition of the numerical traces (1.4) and (1.6), we have that

$$(\widehat{\boldsymbol{q}}_h - \boldsymbol{q}_h) \cdot \boldsymbol{n} = \begin{cases} \tau_K (u_h - \widehat{u}_h) & \text{on } e_K^\tau, \\ 0 & \text{otherwise} \end{cases}$$
$$= \begin{cases} \tau_K (u_h - \widehat{u}_h) & \text{on } e_K^\tau, \\ \mathsf{P}_\partial \boldsymbol{q} \cdot \boldsymbol{n} - \boldsymbol{\Pi}\boldsymbol{q} \cdot \boldsymbol{n} & \text{otherwise,} \end{cases}$$

by the definition of the projection $\boldsymbol{\Pi}$, (2.1b), and that of the projection P_∂, (1.5). So, we only have to prove that

$$(\widehat{\boldsymbol{q}}_h - \boldsymbol{q}_h) \cdot \boldsymbol{n} = \mathsf{P}_\partial \boldsymbol{q} \cdot \boldsymbol{n} - \boldsymbol{\Pi}\boldsymbol{q} \cdot \boldsymbol{n} \qquad \text{on } e_K^\tau.$$

But, by Theorem 2.3, we have that

$$(\widehat{\boldsymbol{q}}_h - \boldsymbol{q}_h) \cdot \boldsymbol{n} = (\widehat{\mathbf{Q}}\lambda_h - \mathbf{Q}\lambda_h) \cdot \boldsymbol{n} + (\widehat{\mathbf{Q}}g - \mathbf{Q}g) \cdot \boldsymbol{n} + (\widehat{\mathbf{Q}}f - \mathbf{Q}f) \cdot \boldsymbol{n}$$
$$= \mathsf{P}_\partial \boldsymbol{q} \cdot \boldsymbol{n} - \boldsymbol{\Pi}\boldsymbol{q} \cdot \boldsymbol{n}$$

on the face e_K^τ. This completes the proof of Proposition 2.2.

3.2.4. *Proof of Theorem* 2.4. The statement (i) of Theorem 2.4 follows directly from Theorem 2.3 and from Lemma 3.3.

To prove the remaining statements, we are going to use the fact that the RT_k, BDM_k and $\mathrm{SF\text{-}H}_k$ methods have exactly the same structure and satisfy the characterization Theorem 2.3; see [11]. The only difference between these methods is the choice of local spaces (see Table 1) and the choice of the local stabilization parameters τ; see Table 2. Thus, to prove statement (ii) we only have to show that the functions $(\mathbf{Q}m, \mathsf{P}^{k-1}\,\mathcal{U}m)$ and $(\mathbf{Q}f, \mathsf{P}^{k-1}\,\mathcal{U}f)$ are the same for all these methods whenever $f|_K \in \mathcal{P}^{k-1}(K)$ for all $K \in \Omega_h$. Similarly, to prove statement (iii), we only have to show that $\mathbf{Q}m$ is the same for all these methods.

To do that, we begin by noting that we have, by Lemma 3.3, that the function $(\boldsymbol{\mathsf{Q}}\mathsf{m}, \mathsf{P}^{k-1}\,\mathcal{U}\mathsf{m}) \in \boldsymbol{\mathcal{V}}(K) \times \mathcal{W}(K)$ is determined by

$$(\boldsymbol{c}\,\boldsymbol{\mathsf{Q}}\mathsf{m}, \boldsymbol{v})_K = - \langle \mathsf{m}, \boldsymbol{v} \cdot \boldsymbol{n} \rangle_{\partial K} \qquad\qquad \forall\, \boldsymbol{v} \in \boldsymbol{\mathcal{V}}(K),$$

$$(\mathsf{P}^{k-1}\,\mathcal{U}\mathsf{m}, \nabla \cdot \boldsymbol{v})_K = (\boldsymbol{c}\,\boldsymbol{\mathsf{Q}}\mathsf{m}, \boldsymbol{v})_K + \langle \mathsf{m}, \boldsymbol{v} \cdot \boldsymbol{n} \rangle_{\partial K} \quad \forall\, \boldsymbol{v} \in \boldsymbol{\mathcal{V}}^{\perp}(K).$$

and that the function $(\boldsymbol{\mathsf{Q}}f, \mathsf{P}^{k-1}\,\mathcal{U}f) \in \boldsymbol{\mathcal{V}}^{\perp}(K) \times \mathcal{W}(K)$ is determined by the equations

$$(\omega, \nabla \cdot \boldsymbol{\mathsf{Q}}f)_K = - \langle \omega, \tau\,\mathcal{U}f \rangle_{e_K^{\tau}} + (f, \omega)_K \quad \forall\, \omega \in \mathcal{W}(K),$$

$$(\mathsf{P}^{k-1}\,\mathcal{U}f, \nabla \cdot \boldsymbol{v})_K = (\boldsymbol{c}\,\boldsymbol{\mathsf{Q}}\mathsf{m}, \boldsymbol{v})_K \qquad\qquad \forall\, \boldsymbol{v} \in \boldsymbol{\mathcal{V}}^{\perp}(K),$$

where $\mathcal{U}f|_{e_K^{\tau}} = 0$, by (α) of Lemma 3.3, if $f|_K \in \mathcal{P}^{k-1}(K)$. Since the four equations above also hold (the third whenever $f|_K \in \mathcal{P}^{k-1}(K)$) for the BDM_k method, we conclude that the statements (ii) and (iii) hold if we exclude the RT_k method.

To show that these statements also hold if we include it, we note that the above equations hold for the RT_k method if we modify the definition of the spaces $\boldsymbol{\mathcal{V}}(K)$ and $\boldsymbol{\mathcal{V}}(K)$ by

$$\boldsymbol{\mathcal{V}}_{\mathrm{RT}}(K) := \{\boldsymbol{v} \in \boldsymbol{\mathcal{P}}^k(K) \oplus \boldsymbol{x}\,\mathcal{P}^k(K): \ \nabla \cdot \boldsymbol{v} = 0\},$$

$$\boldsymbol{\mathcal{V}}_{\mathrm{RT}}^{\perp}(K) := \{\boldsymbol{v} \in \boldsymbol{\mathcal{P}}^k(K) \oplus \boldsymbol{x}\,\mathcal{P}^k(K): \ (\boldsymbol{c}\,\boldsymbol{v}, \boldsymbol{\sigma})_K = 0 \ \forall\, \boldsymbol{\sigma} \in \boldsymbol{\mathcal{V}}_{\mathrm{RT}}(K)\},$$

and if we replace the third equation by

$$\nabla \cdot \boldsymbol{\mathsf{Q}}f = \mathsf{P}f.$$

Thus, the result follows from the fact that

$$\boldsymbol{\mathcal{V}}_{\mathrm{RT}}(K) = \{\boldsymbol{v} \in \boldsymbol{\mathcal{P}}^k(K): \ \nabla \cdot \boldsymbol{v} = 0\} = \boldsymbol{\mathcal{V}}(K),$$

and from the fact that, if $\boldsymbol{\mathsf{Q}}f \in \boldsymbol{\mathcal{V}}_{\mathrm{RT}}^{\perp}(K)$ and $\nabla \cdot \boldsymbol{\mathsf{Q}}f = \mathsf{P}f \in \mathcal{P}^{k-1}(K)$, then $\boldsymbol{\mathsf{Q}}f$ belongs to the space $\{\boldsymbol{v} \in \boldsymbol{\mathcal{P}}^k(K): \ (\boldsymbol{c}\,\boldsymbol{v}, \boldsymbol{\sigma})_K = 0 \ \forall\, \boldsymbol{\sigma} \in \boldsymbol{\mathcal{V}}(K)\} = \boldsymbol{\mathcal{V}}^{\perp}(K)$. This completes the proof of Theorem 2.4.

3.3. Proof of the error estimates.

The proof of the error estimates is based on the error equations and the properties of the projection $(\boldsymbol{\Pi}, \mathbb{P})$ gathered in Proposition 2.1. The error equations are

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

(3.2b) $\qquad (\omega, \nabla \cdot (\boldsymbol{q} - \boldsymbol{q}_h))_{\Omega_h} - \langle \omega, (\widehat{\boldsymbol{q}}_h - \boldsymbol{q}_h) \cdot \boldsymbol{n} \rangle_{\partial \Omega_h} = 0,$

(3.2c) $\qquad u - \widehat{u}_h = g - \mathsf{P}_{\partial}g \qquad$ on $\partial \Omega_D,$

(3.2d) $\qquad (\boldsymbol{q} - \widehat{\boldsymbol{q}}_h) \cdot \boldsymbol{n} = \mathsf{q}_\mathsf{N} - \mathsf{P}_{\partial}\mathsf{q}_\mathsf{N} \qquad$ on $\partial \Omega_N,$

for all $(\boldsymbol{v}, \omega) \in \boldsymbol{V}_h \times W_h$.

A direct consequence of the weak commutativity identity (iv) of Proposition 2.1 that we find convenient to use in our analysis is contained in the following result.

Corollary 3.6. *For all $(\boldsymbol{\sigma}, \zeta) \in \boldsymbol{H}^1(\Omega_h) \times H^1(\Omega_h)$, we have*

$(\alpha) \quad (\zeta, \nabla \cdot \boldsymbol{\Pi}\boldsymbol{\sigma})_{\Omega_h} = (\mathbb{P}\zeta, \nabla \cdot \boldsymbol{\sigma})_{\Omega_h} + \langle \mathsf{P}_{\partial}\zeta, \boldsymbol{\Pi}\boldsymbol{\sigma} \cdot \boldsymbol{n} - \mathsf{P}_{\partial}\boldsymbol{\sigma} \cdot \boldsymbol{n} \rangle_{\partial \Omega_h},$

$(\beta) \quad \| \boldsymbol{\Pi}\boldsymbol{\sigma} - \boldsymbol{\sigma} \|_{\boldsymbol{L}^2(\Omega_h)} \leq C\,h^{r+1}\,|\boldsymbol{\sigma}|_{\boldsymbol{H}^{r+1}(\Omega_h)},$

$(\gamma) \quad \| \mathbb{P}\zeta - \zeta \|_{L^2(\Omega_h)} \leq C\,h^{s+1}\,|\zeta|_{H^{s+1}(\Omega_h)},$

where $r, s \in [0, k]$ and C depends only on k and the shape-regularity parameters of the simplexes $K \in \Omega_h$.

3.3.1. *Proof of Theorem* 2.5: *The error in the flux.* Theorem 2.5 follows immediately from the following auxiliary result.

Lemma 3.7. *We have* $(\boldsymbol{c}\,(\boldsymbol{q} - \boldsymbol{q}_h), \boldsymbol{\Pi q} - \boldsymbol{q}_h)_{\Omega_h} = 0$.

Indeed, this implies that

$$(\boldsymbol{c}\,(\boldsymbol{q} - \boldsymbol{q}_h), \boldsymbol{q} - \boldsymbol{q}_h)_{\Omega_h} = (\boldsymbol{c}\,(\boldsymbol{q} - \boldsymbol{q}_h), \boldsymbol{q} - \boldsymbol{\Pi q})_{\Omega_h},$$

and hence, that

$$\|\boldsymbol{q} - \boldsymbol{q}_h\|_{\boldsymbol{L}^2(\Omega_h, \boldsymbol{c})} \le \|\boldsymbol{\Pi q} - \boldsymbol{q}\|_{\boldsymbol{L}^2(\Omega_h, \boldsymbol{c})} \le C\,h^{s+1}\,|\boldsymbol{q}|_{\boldsymbol{H}^{s+1}(\Omega_h)},$$

for some $s \in [0, k]$, by the estimate (vi) of Proposition 2.1. This proves Theorem 2.5.

Let us prove Lemma 3.7.

Proof. By the error equation (3.2a) with $\boldsymbol{v} := \boldsymbol{\Pi q} - \boldsymbol{q}_h$, we have

$$
\begin{aligned}
(\boldsymbol{c}\,(\boldsymbol{q} - \boldsymbol{q}_h), \boldsymbol{\Pi q} - \boldsymbol{q}_h)_{\Omega_h} = {} & (u - u_h, \nabla \cdot (\boldsymbol{\Pi q} - \boldsymbol{q}_h))_{\Omega_h} \\
& - \langle u - \widehat{u}_h, (\boldsymbol{\Pi q} - \boldsymbol{q}_h) \cdot \boldsymbol{n}\rangle_{\partial\Omega_h}.
\end{aligned}
$$

By the identity (α) of Corollary 3.6 with $(\boldsymbol{\sigma}, \zeta) := (\boldsymbol{q} - \boldsymbol{q}_h, u - u_h)$, we get that

$$
\begin{aligned}
(\boldsymbol{c}\,(\boldsymbol{q} - \boldsymbol{q}_h), \boldsymbol{\Pi q} - \boldsymbol{q}_h)_{\Omega_h} = {} & (\mathbb{P}u - u_h, \nabla \cdot (\boldsymbol{q} - \boldsymbol{q}_h))_{\Omega_h} \\
& + \langle \mathsf{P}_\partial u - u_h, (\boldsymbol{\Pi q} - \boldsymbol{q}) \cdot \boldsymbol{n}\rangle_{\partial\Omega_h} \\
& - \langle u - \widehat{u}_h, (\boldsymbol{\Pi q} - \boldsymbol{q}_h) \cdot \boldsymbol{n}\rangle_{\partial\Omega_h},
\end{aligned}
$$

and by the error equation (3.2b) with $\omega := \mathbb{P}u - u_h$,

$$
\begin{aligned}
(\boldsymbol{c}\,(\boldsymbol{q} - \boldsymbol{q}_h), \boldsymbol{\Pi q} - \boldsymbol{q}_h)_{\Omega_h} = {} & \langle \mathbb{P}u - u_h, (\widehat{\boldsymbol{q}}_h - \boldsymbol{q}_h) \cdot \boldsymbol{n}\rangle_{\partial\Omega_h} \\
& + \langle \mathsf{P}_\partial u - u_h, (\boldsymbol{\Pi q} - \boldsymbol{q}) \cdot \boldsymbol{n}\rangle_{\partial\Omega_h} \\
& - \langle u - \widehat{u}_h, (\boldsymbol{\Pi q} - \boldsymbol{q}_h) \cdot \boldsymbol{n}\rangle_{\partial\Omega_h}.
\end{aligned}
$$

It we denote by T the right-hand side of the above equations, it is not difficult to see that, after a few simple algebraic manipulations, we have that $T = \sum_{i=1}^{5} T_i$, where

$$
\begin{aligned}
T_1 &:= \langle \widehat{u}_h - u_h, (\widehat{\boldsymbol{q}}_h - \boldsymbol{q}_h) \cdot \boldsymbol{n}\rangle_{\partial\Omega_h}, \\
T_2 &:= \langle \widehat{u}_h - u_h, (\boldsymbol{\Pi q} - \boldsymbol{q}) \cdot \boldsymbol{n}\rangle_{\partial\Omega_h}, \\
T_3 &:= \langle \widehat{u}_h - u, (\boldsymbol{q} - \widehat{\boldsymbol{q}}_h) \cdot \boldsymbol{n}\rangle_{\partial\Omega_h}, \\
T_4 &:= \langle \mathsf{P}_\partial u - u, (\boldsymbol{\Pi q} - \boldsymbol{q}) \cdot \boldsymbol{n}\rangle_{\partial\Omega_h}, \\
T_5 &:= \langle \mathbb{P}u - u, (\widehat{\boldsymbol{q}}_h - \boldsymbol{q}_h) \cdot \boldsymbol{n}\rangle_{\partial\Omega_h}.
\end{aligned}
$$

We are going to show that $T = 0$.

We begin by noting that,

$$
\begin{aligned}
T_1 &= \sum_{K \in \Omega_h} \langle \widehat{u}_h - u_h, (\widehat{\boldsymbol{q}}_h - \boldsymbol{q}_h) \cdot \boldsymbol{n}\rangle_{\partial K} \\
&= \sum_{K \in \Omega_h} \langle \widehat{u}_h - u_h, \mathsf{P}_\partial \boldsymbol{q} \cdot \boldsymbol{n} - \boldsymbol{\Pi q} \cdot \boldsymbol{n}\rangle_{\partial K},
\end{aligned}
$$

by Proposition 2.2. By the definition of the projection P_∂, (1.5),

$$T_1 = \sum_{K \in \Omega_h} \langle \widehat{u}_h - u_h, (\boldsymbol{q} - \boldsymbol{\Pi q}) \cdot \boldsymbol{n} \rangle_{\partial K}$$
$$= - T_2,$$

by the definition of the projection $\boldsymbol{\Pi}$, (2.1b). Thus, $T_1 + T_2 = 0$.

Next, let us show that $T_3 + T_4 = 0$. By the fact that the numerical trace $\widehat{u}_h$ and the normal component of the numerical trace $\widehat{\boldsymbol{q}}_h$ are single-valued on the interior faces, by definition of $\widehat{u}_h$, (1.4a), and the equation (1.3c) satisfied by $\widehat{\boldsymbol{q}}_h$, we have that

$$T_3 = \langle \widehat{u}_h - u, (\boldsymbol{q} - \widehat{\boldsymbol{q}}_h) \cdot \boldsymbol{n} \rangle_{\partial \Omega},$$
$$= \langle \mathsf{P}_\partial g - g, (\boldsymbol{q} - \widehat{\boldsymbol{q}}_h) \cdot \boldsymbol{n} \rangle_{\partial \Omega_D} + \langle \lambda_h - u, \mathsf{q_N} - \mathsf{P}_\partial \mathsf{q_N} \rangle_{\partial \Omega_N},$$

by the definition of the numerical traces at the boundary. By using the definition of the projection P_∂, (1.5), we get

$$T_3 = \langle \mathsf{P}_\partial u - u, \boldsymbol{q} \cdot \boldsymbol{n} - \mathsf{P}_\partial \boldsymbol{q} \cdot \boldsymbol{n} \rangle_{\partial \Omega_D} + \langle \mathsf{P}_\partial u - u, \boldsymbol{q} \cdot \boldsymbol{n} - \mathsf{P}_\partial \boldsymbol{q} \cdot \boldsymbol{n} \rangle_{\partial \Omega_N}$$
$$= \langle \mathsf{P}_\partial u - u, \boldsymbol{q} \cdot \boldsymbol{n} - \mathsf{P}_\partial \boldsymbol{q} \cdot \boldsymbol{n} \rangle_{\partial \Omega}$$
$$= \langle \mathsf{P}_\partial u - u, \boldsymbol{q} \cdot \boldsymbol{n} - \mathsf{P}_\partial \boldsymbol{q} \cdot \boldsymbol{n} \rangle_{\partial \Omega_h}$$
$$= \langle \mathsf{P}_\partial u - u, \boldsymbol{q} \cdot \boldsymbol{n} - \boldsymbol{\Pi q} \cdot \boldsymbol{n} \rangle_{\partial \Omega_h}$$
$$= - T_4,$$

Finally, let us show that $T_5 = 0$. By the definition of the numerical trace $\widehat{\boldsymbol{q}}_h$, (1.4b),

$$T_5 = \sum_{K \in \Omega_h} \langle \mathbb{P}u - u, (\widehat{\boldsymbol{q}}_h - \boldsymbol{q}_h) \cdot \boldsymbol{n} \rangle_{\partial K}$$
$$= \sum_{K \in \Omega_h} \langle \mathbb{P}u - u, (\widehat{\boldsymbol{q}}_h - \boldsymbol{q}_h) \cdot \boldsymbol{n} \rangle_{e_K^\tau}$$
$$= 0,$$

by the definition of the projection $\mathbb{P}$, (2.2b). This completes the proof. $\qquad \square$

3.3.2. *Proof of Theorem* 2.6: *Superconvergence of u_h.* Since

$$\| \mathbb{P}u - u_h \|_{H^{-s}(\Omega_h)} = \sup_{\theta \in \mathcal{C}_0^\infty(\Omega)} \frac{(\mathbb{P}u - u_h, \theta)_\Omega}{\| \theta \|_{H^s(\Omega)}},$$

we need to estimate the number $(\mathbb{P}u - u_h, \theta)_\Omega$. It is expressed in a suitable way in the following auxiliary result. Let us recall that P^{k-1} is defined by (2.4) for $k \geq 1$, and is $\mathsf{P}^{k-1} \equiv 0$ for $k = 0$.

Lemma 3.8. *We have*

$$(\mathbb{P}u - u_h, \theta)_{\Omega_h} = (\boldsymbol{c} \, (\boldsymbol{q} - \boldsymbol{q}_h), \boldsymbol{\Pi \psi} - \boldsymbol{\psi})_{\Omega_h} + (\mathsf{P}^{k-1} \nabla \varphi - \nabla \varphi, \boldsymbol{q} - \boldsymbol{\Pi q})_{\Omega_h}$$
$$- \sum_{K \in \Omega_h} \tau^{-1} \langle \boldsymbol{\Pi q} \cdot \boldsymbol{n} - \mathsf{P}_\partial \boldsymbol{q} \cdot \boldsymbol{n}, \boldsymbol{\Pi \psi} \cdot \boldsymbol{n} - \mathsf{P}_\partial \boldsymbol{\psi} \cdot \boldsymbol{n} \rangle_{e_K^\tau}.$$

Assume that $k \geq 1$. Then, applying the Cauchy-Schwarz inequality and using the estimate of $\boldsymbol{q} - \boldsymbol{q}_h$ in Theorem 2.5, and the approximation properties of the

projections P^{k-1} and $\mathbb{P}$, (v) in Proposition 2.1 and (β) in Corollary 3.6, we readily obtain

$$
\begin{aligned}
(\mathbb{P}u - u_h, \theta)_{\Omega_h} \leq & C\, h^{r+1}\, |\,\boldsymbol{q}\,|_{\boldsymbol{H}^{r+1}(\Omega_h)}\, h^{s+1}\, |\,\boldsymbol{\psi}\,|_{\boldsymbol{H}^{s+1}(\Omega_h)} \\
& + C\, h^{r+1}\, |\,\boldsymbol{q}\,|_{\boldsymbol{H}^{r+1}(\Omega_h)}\, h^{s+1}\, |\,\nabla\varphi\,|_{\boldsymbol{H}^{s+1}(\Omega_h)} \\
& + \max_{K \in \Omega_h} \frac{1}{h_K\, \tau_K}\, C\, h^{r+1}\, |\,f\,|_{H^r(\Omega_h)}\, h^{s+1}\, |\,\theta\,|_{H^s(\Omega_h)},
\end{aligned}
$$

where $r, s \in [0, k-1]$. Since $\kappa := \max_{K \in \Omega_h} \frac{1}{h_K\, \tau_K}$ and using the elliptic regularity assumption (2.8), we get

$$
(\mathbb{P}u - u_h, \theta)_{\Omega_h} \leq C\, \mathfrak{C}_\kappa^{r,s}(\boldsymbol{q})\, h^{r+s+2}\, |\,\theta\,|_{H^s(\Omega_h)}.
$$

This completes the proof of Theorem 2.6 for $k \geq 1$.

In the case $k = 0$, we have that

$$
\begin{aligned}
(\mathbb{P}u - u_h, \theta)_{\Omega_h} \leq & C\, h\, |\,\boldsymbol{q}\,|_{\boldsymbol{H}^1(\Omega_h)}\, h\, |\,\boldsymbol{\psi}\,|_{\boldsymbol{H}^1(\Omega_h)} \\
& + C\, h\, |\,\boldsymbol{q}\,|_{\boldsymbol{H}^1(\Omega_h)}\, |\,\nabla\varphi\,|_{\boldsymbol{H}^1(\Omega_h)} \\
& + \kappa\, C\, h\, |\,f\,|_{L^2(\Omega_h)}\, h\, |\,\theta\,|_{L^2(\Omega_h)},
\end{aligned}
$$

and, after using the elliptic regularity assumption (2.8), we get

$$
(\mathbb{P}u - u_h, \theta)_{\Omega_h} \leq C\, \mathcal{C}_\kappa(\boldsymbol{q})\, h\, |\,\theta\,|_{L^2(\Omega_h)}.
$$

Finally, let us consider the case $k = 0$ and $f = 0$. By the identity (v) of Proposition 2.1 we have that $\boldsymbol{\Pi}\boldsymbol{\sigma} \cdot \boldsymbol{n} = \mathsf{P}_\partial \boldsymbol{\sigma} \cdot \boldsymbol{n}$, and by the identity (vi) of Proposition 2.1 we have that $\boldsymbol{\Pi}\boldsymbol{q} = \boldsymbol{\Pi}^{\mathrm{RT}}\boldsymbol{q}$. This implies that

$$
\begin{aligned}
(\mathbb{P}u - u_h, \theta)_{\Omega_h} = & (\boldsymbol{c}\,(\boldsymbol{q} - \boldsymbol{q}_h), \boldsymbol{\Pi}\boldsymbol{\psi} - \boldsymbol{\psi})_{\Omega_h} + \langle \varphi, (\boldsymbol{q} - \boldsymbol{\Pi}\boldsymbol{q}) \cdot \boldsymbol{n} \rangle_{\partial\Omega} \\
= & (\boldsymbol{c}\,(\boldsymbol{q} - \boldsymbol{q}_h), \boldsymbol{\Pi}^{\mathrm{RT}}\boldsymbol{\psi} - \boldsymbol{\psi})_{\Omega_h} + \langle \varphi - \mathsf{P}_\partial\varphi, \mathsf{q}_{\mathsf{N}} - \mathsf{P}_\partial \mathsf{q}_{\mathsf{N}} \rangle_{\partial\Omega_N}
\end{aligned}
$$

by the adjoint equation (2.7c) and the boundary condition (1.1d). As a consequence, we get

$$
\begin{aligned}
(\mathbb{P}u - u_h, \theta)_{\Omega_h} \leq & C\, h\, |\,\boldsymbol{q}\,|_{\boldsymbol{H}^1(\Omega_h)}\, h\, |\,\boldsymbol{\psi}\,|_{\boldsymbol{H}^1(\Omega_h)} \\
& + C\, h\, |\,\mathsf{q}_{\mathsf{N}}\,|_{H^1(\partial\Omega_N)}\, h\, |\,\varphi\,|_{H^1(\partial\Omega_N)},
\end{aligned}
$$

and since

$$
|\,\varphi\,|_{H^1(\partial\Omega_N)} \leq C\, |\,\varphi\,|_{H^2(\Omega)},
$$

by the elliptic regularity assumption (2.8), we get

$$
(\mathbb{P}u - u_h, \theta)_{\Omega_h} \leq C\, \mathfrak{C}_\kappa^{0,0}(\boldsymbol{q})\, h^2\, |\,\theta\,|_{L^2(\Omega)}.
$$

It remains to prove Lemma 3.8.

Proof. By the adjoint equation (2.7b), we have that

$$
\begin{aligned}
(\mathbb{P}u - u_h, \theta)_{\Omega_h} = & (\mathbb{P}u - u_h, \nabla \cdot \boldsymbol{\psi})_{\Omega_h} \\
= & (u - u_h, \nabla \cdot \boldsymbol{\Pi}\boldsymbol{\psi})_{\Omega_h} - \langle \mathsf{P}_\partial u - u_h, (\boldsymbol{\Pi}\boldsymbol{\psi} - \boldsymbol{\psi}) \cdot \boldsymbol{n} \rangle_{\partial\Omega_h},
\end{aligned}
$$

by the identity (α) of Corollary 3.6 with $(\boldsymbol{\sigma}, \zeta) := (\boldsymbol{\psi}, u - u_h)$. By the error equation (3.2a) with $\boldsymbol{v} := \boldsymbol{\Pi}\boldsymbol{\psi}$, we get

$$
\begin{aligned}
(\mathbb{P}u - u_h, \theta)_{\Omega_h} =& (\boldsymbol{c}\,(\boldsymbol{q} - \boldsymbol{q}_h), \boldsymbol{\Pi}\boldsymbol{\psi})_{\Omega_h} + \langle u - \widehat{u}_h, \boldsymbol{\Pi}\boldsymbol{\psi} \cdot \boldsymbol{n}\rangle_{\partial\Omega_h} \\
& - \langle \mathsf{P}_\partial u - u_h, (\boldsymbol{\Pi}\boldsymbol{\psi} - \boldsymbol{\psi}) \cdot \boldsymbol{n}\rangle_{\partial\Omega_h} \\
=& (\boldsymbol{c}\,(\boldsymbol{q} - \boldsymbol{q}_h), \boldsymbol{\Pi}\boldsymbol{\psi} - \boldsymbol{\psi})_{\Omega_h} + (\boldsymbol{c}\,(\boldsymbol{q} - \boldsymbol{q}_h), \boldsymbol{\psi})_{\Omega_h} \\
& + \langle u - \widehat{u}_h, \boldsymbol{\Pi}\boldsymbol{\psi} \cdot \boldsymbol{n}\rangle_{\partial\Omega_h} - \langle \mathsf{P}_\partial u - u_h, (\boldsymbol{\Pi}\boldsymbol{\psi} - \boldsymbol{\psi}) \cdot \boldsymbol{n}\rangle_{\partial\Omega_h},
\end{aligned}
$$

and, by the adjoint equation (2.7a),

$$
\begin{aligned}
(\mathbb{P}u - u_h, \theta)_{\Omega_h} =& (\boldsymbol{c}\,(\boldsymbol{q} - \boldsymbol{q}_h), \boldsymbol{\Pi}\boldsymbol{\psi} - \boldsymbol{\psi})_{\Omega_h} - (\boldsymbol{q} - \boldsymbol{q}_h, \nabla\varphi)_{\Omega_h} \\
& + \langle u - \widehat{u}_h, \boldsymbol{\Pi}\boldsymbol{\psi} \cdot \boldsymbol{n}\rangle_{\partial\Omega_h} - \langle \mathsf{P}_\partial u - u_h, (\boldsymbol{\Pi}\boldsymbol{\psi} - \boldsymbol{\psi}) \cdot \boldsymbol{n}\rangle_{\partial\Omega_h}.
\end{aligned}
$$

By the orthogonality property (ii) of Proposition 2.1, we get that

$$
\begin{aligned}
(\mathbb{P}u - u_h, \theta)_{\Omega_h} =& (\boldsymbol{c}\,(\boldsymbol{q} - \boldsymbol{q}_h), \boldsymbol{\Pi}\boldsymbol{\psi} - \boldsymbol{\psi})_{\Omega_h} + (\boldsymbol{q} - \boldsymbol{\Pi}\boldsymbol{q}, \mathsf{P}^{k-1}\nabla\varphi - \nabla\varphi)_{\Omega_h} \\
& - (\boldsymbol{\Pi}\boldsymbol{q} - \boldsymbol{q}_h, \nabla\varphi)_{\Omega_h} + \langle u - \widehat{u}_h, \boldsymbol{\Pi}\boldsymbol{\psi} \cdot \boldsymbol{n}\rangle_{\partial\Omega_h} \\
& - \langle \mathsf{P}_\partial u - u_h, (\boldsymbol{\Pi}\boldsymbol{\psi} - \boldsymbol{\psi}) \cdot \boldsymbol{n}\rangle_{\partial\Omega_h}.
\end{aligned}
$$

If we denote by T the last three terms of the above right-hand side, we see that, after some simple algebraic manipulations, we can write $T = \sum_{i=1}^{4} T_i$, where

$$
\begin{aligned}
T_1 &= - \langle \widehat{u}_h - u_h, (\boldsymbol{\Pi}\boldsymbol{\psi} - \boldsymbol{\psi}) \cdot \boldsymbol{n}\rangle_{\partial\Omega_h}, \\
T_2 &= - \langle \mathsf{P}_\partial u - u, (\boldsymbol{\Pi}\boldsymbol{\psi} - \boldsymbol{\psi}) \cdot \boldsymbol{n}\rangle_{\partial\Omega_h}, \\
T_3 &= - \langle \widehat{u}_h - u, \boldsymbol{\psi} \cdot \boldsymbol{n}\rangle_{\partial\Omega_h}, \\
T_4 &= - (\boldsymbol{\Pi}\boldsymbol{q} - \boldsymbol{q}_h, \nabla\varphi)_{\Omega_h}.
\end{aligned}
$$

By the definition of the numerical trace $\widehat{u}_h$, (1.4a) and (1.6), we have that

$$
\begin{aligned}
T_1 &= - \sum_{K \in \Omega_h} \langle \widehat{u}_h - u_h, \boldsymbol{\Pi}\boldsymbol{\psi} \cdot \boldsymbol{n} - \mathsf{P}_\partial\boldsymbol{\psi} \cdot \boldsymbol{n}\rangle_{e_K^\tau} \\
&= - \sum_{K \in \Omega_h} \tau^{-1} \langle \boldsymbol{\Pi}\boldsymbol{q} \cdot \boldsymbol{n} - \mathsf{P}_\partial\boldsymbol{q} \cdot \boldsymbol{n}, \boldsymbol{\Pi}\boldsymbol{\psi} \cdot \boldsymbol{n} - \mathsf{P}_\partial\boldsymbol{\psi} \cdot \boldsymbol{n}\rangle_{e_K^\tau},
\end{aligned}
$$

by Proposition 2.2.

It remains to show that $T_2 + T_3 + T_4 = 0$. By the definition of the projection P_∂, (1.5),

$$
\begin{aligned}
T_2 &= - \langle \mathsf{P}_\partial u - u, -\boldsymbol{\psi} \cdot \boldsymbol{n}\rangle_{\partial\Omega_h} \\
&= - \langle \mathsf{P}_\partial u - u, -\boldsymbol{\psi} \cdot \boldsymbol{n}\rangle_{\partial\Omega_D} \\
&= \langle \mathsf{P}_\partial u - u, \boldsymbol{\psi} \cdot \boldsymbol{n}\rangle_{\partial\Omega_D}.
\end{aligned}
$$

By the definition of the numerical trace $\widehat{u}_h$, (1.4a),

$$
\begin{aligned}
T_3 &= - \langle \widehat{u}_h - u, \boldsymbol{\psi} \cdot \boldsymbol{n}\rangle_{\partial\Omega} \\
&= - \langle \widehat{u}_h - u, \boldsymbol{\psi} \cdot \boldsymbol{n}\rangle_{\partial\Omega_D} \\
&= - \langle \mathsf{P}_\partial u - u, \boldsymbol{\psi} \cdot \boldsymbol{n}\rangle_{\partial\Omega_D} \\
&= - T_2.
\end{aligned}
$$

Next, we show that $T_4 = 0$. Integrating by parts, we obtain

$$T_4 = (\nabla \cdot (\mathbf{\Pi q} - \boldsymbol{q}_h), \varphi)_{\Omega_h} - \langle (\mathbf{\Pi q} - \boldsymbol{q}_h) \cdot \boldsymbol{n}, \varphi \rangle_{\partial \Omega_h}$$
$$= (\nabla \cdot (\boldsymbol{q} - \boldsymbol{q}_h), \mathbb{P} \varphi)_{\Omega_h} + \langle (\mathbf{\Pi q} - \boldsymbol{q}) \cdot \boldsymbol{n}, \mathsf{P}_\partial \varphi \rangle_{\partial \Omega_h}$$
$$- \langle (\mathbf{\Pi q} - \boldsymbol{q}_h) \cdot \boldsymbol{n}, \varphi \rangle_{\partial \Omega_h}$$

by the identity (α) of Corollary 3.6 with $(\boldsymbol{\sigma}, \zeta) := (\boldsymbol{q} - \boldsymbol{q}_h, \varphi)$. By the error equation (3.2b) with $\omega := \mathbb{P} \varphi$,

$$T_4 = \langle \mathbb{P} \varphi, (\widehat{\boldsymbol{q}}_h - \boldsymbol{q}_h) \cdot \boldsymbol{n} \rangle_{\partial \Omega_h} + \langle (\mathbf{\Pi q} - \boldsymbol{q}) \cdot \boldsymbol{n}, \mathsf{P}_\partial \varphi \rangle_{\partial \Omega_h}$$
$$- \langle (\mathbf{\Pi q} - \boldsymbol{q}_h) \cdot \boldsymbol{n}, \varphi \rangle_{\partial \Omega_h}$$
$$= \langle \mathsf{P}_\partial \varphi, (\widehat{\boldsymbol{q}}_h - \boldsymbol{q}_h) \cdot \boldsymbol{n} \rangle_{\partial \Omega_h} + \langle (\mathbf{\Pi q} - \boldsymbol{q}) \cdot \boldsymbol{n}, \mathsf{P}_\partial \varphi \rangle_{\partial \Omega_h}$$
$$- \langle (\mathbf{\Pi q} - \boldsymbol{q}_h) \cdot \boldsymbol{n}, \mathsf{P}_\partial \varphi \rangle_{\partial \Omega_h},$$

by the definition of the projection P_∂, (1.5), the definition of the projection $\mathbb{P}$, (2.2b), and the definition of the numerical trace $\widehat{\boldsymbol{q}}_h$, (1.4b). Hence

$$T_4 = \langle \mathsf{P}_\partial \varphi, (\widehat{\boldsymbol{q}}_h - \boldsymbol{q}) \cdot \boldsymbol{n} \rangle_{\partial \Omega_h} = \langle \mathsf{P}_\partial \varphi, (\widehat{\boldsymbol{q}}_h - \boldsymbol{q}) \cdot \boldsymbol{n} \rangle_{\partial \Omega} = 0,$$

by the adjoint equation (2.7c) and the equation (1.3c) for $\widehat{\boldsymbol{q}}_h$.

This completes the proof. □

3.3.3. *Proof of Theorem* 2.8: *Superconvergence of* $\widehat{u}_h$. To prove this theorem, let us begin by estimating $\| \mathsf{P}_\partial u - \widehat{u}_h \|^2_{L^2(e)}$ for each face e of each simplex K. For the face e_K^τ, we have that, by definition of the projection $\mathbb{P}$, (2.2),

$$\| \mathsf{P}_\partial u - \widehat{u}_h \|_{L^2(e_K^\tau)} = \| \mathbb{P} u - \widehat{u}_h \|_{L^2(e_K^\tau)}$$
$$\leq \| \mathbb{P} u - u_h \|_{L^2(e_K^\tau)} + \| u_h - \widehat{u}_h \|_{L^2(e_K^\tau)}$$
$$\leq \| \mathbb{P} u - u_h \|_{L^2(e_K^\tau)} + C \frac{h_K^{r+1/2}}{\tau_K} | f |_{H^r(K)}$$

by Proposition 2.2 and the identity (v) of Proposition 2.1. By using a classical inverse inequality, we can conclude that

$$h_K^{1/2} \| \mathsf{P}_\partial u - \widehat{u}_h \|_{L^2(e_K^\tau)} \leq C \left(\| \mathbb{P} u - u_h \|_{L^2(K)} + \frac{h_K^{r+1}}{\tau_K} | f |_{H^r(K)} \right).$$

Now we consider the error in the faces e of K which are different from the face e_K^τ. By the error equation (3.2a), we have that, for all $\boldsymbol{v} \in \boldsymbol{\mathcal{P}}^k(K)$,

$$\langle \widehat{u}_h - \mathsf{P}_\partial u, \boldsymbol{v} \cdot \boldsymbol{n} \rangle_{\partial K \backslash e_K^\tau} = (\boldsymbol{c} (\boldsymbol{q} - \boldsymbol{q}_h), \boldsymbol{v})_K - (\mathbb{P} u - u_h, \nabla \cdot \boldsymbol{v})_K$$
$$- \langle \widehat{u}_h - \mathsf{P}_\partial u, \boldsymbol{v} \cdot \boldsymbol{n} \rangle_{e_K^\tau}.$$

Taking $\boldsymbol{v} := \boldsymbol{Z}$ given by Lemma 3.2 with $z = \widehat{u}_h - \mathsf{P}_\partial u$, we obtain that

$$\| \widehat{u}_h - \mathsf{P}_\partial u \|_{L^2(\partial K \backslash e_K^\tau)} \leq C \, (h_K^{1/2} \| \boldsymbol{q} - \boldsymbol{q}_h \|_{\boldsymbol{L}^2(K)} + h_K^{-1/2} \| \mathbb{P} u - u_h \|_{L^2(K)}$$
$$+ \| \widehat{u}_h - \mathsf{P}_\partial u \|_{L^2(e_K^\tau)}),$$

and using the estimate for the error in e_K^τ,

$$h_K^{1/2} \| \widehat{u}_h - \mathsf{P}_\partial u \|_{L^2(\partial K \backslash e_K^\tau)} \leq C \, (\| \mathbb{P} u - u_h \|_{L^2(K)} + h_K \| \boldsymbol{q} - \boldsymbol{q}_h \|^2_{\boldsymbol{L}^2(K)}$$
$$+ \frac{h_K^{r+1}}{\tau_K} | f |_{H^r(K)}).$$

As a consequence, we get

$$\| P_\partial u - \widehat{u}_h \|_{L^2(\mathscr{E}_h;h)} \leq C \left(\| \mathbb{P}u - u_h \|_{L^2(\Omega_h)} + h \| \boldsymbol{q} - \boldsymbol{q}_h \|_{\boldsymbol{L}^2(\Omega_h)} \right.$$
$$\left. + \kappa \, h^{r+2} \, | f |_{H^r(\Omega_h)} \right),$$

where $\kappa := \max_{K \in \Omega_h} 1/(\tau_K \, h_K)$. The result now follows from Theorems 2.6, 2.5 and 2.4 (i). This completes the proof of Theorem 2.8.

3.3.4. *Proof of Theorem* 2.9: *The error estimate for* $u_h^\star$. By the definition of $u_h^\star$, (2.9a), we have that

$$\| u - u_h^\star \|_{L^2(K)} \leq \| \overline{u} - \overline{u}_h \|_{L^2(K)} + \| \tilde{u} - \tilde{u}_h \|_{L^2(K)},$$

where $\overline{u}$ is defined in (2.9b) and $\tilde{u} = u - \overline{u}$. We estimate each of the two terms of the right-hand side separately.

We begin by estimating the second term. Since, by Poincaré's inequality, we have

$$\| \tilde{u} - \tilde{u}_h \|_{L^2(K)} \leq C \, h_K \| \nabla(\tilde{u} - \tilde{u}_h) \|_{L^2(K)},$$

it is enough to estimate the error in the gradient. To do that, we note that, by the definition of $\tilde{u}_h$, (2.9c), we have

$$(\boldsymbol{a} \, \nabla(\tilde{u} - \tilde{u}_h), \nabla w)_K = -\langle w, (\boldsymbol{q} - \widehat{\boldsymbol{q}}_h) \cdot \boldsymbol{n} \rangle_{\partial K} \qquad \forall w \in \mathcal{P}_0^{k+1}(K).$$

Then

$$\| \nabla(\mathsf{P}^{k+1}\tilde{u} - \tilde{u}_h) \|^2_{L^2(K;\boldsymbol{a})} = (\boldsymbol{a} \, \nabla(\tilde{u} - \tilde{u}_h), \nabla(\mathsf{P}^{k+1}\tilde{u} - \tilde{u}_h))_K$$
$$+ (\boldsymbol{a} \, \nabla(\mathsf{P}^{k+1}\tilde{u} - \tilde{u}), \nabla(\mathsf{P}^{k+1}\tilde{u} - \tilde{u}_h))_K$$
$$= - \langle \mathsf{P}^{k+1}\tilde{u} - \tilde{u}_h, (\boldsymbol{q} - \widehat{\boldsymbol{q}}_h) \cdot \boldsymbol{n} \rangle_{\partial K}$$
$$+ (\boldsymbol{a}\nabla(\mathsf{P}^{k+1}\tilde{u} - \tilde{u}), \nabla(\mathsf{P}^{k+1}\tilde{u} - \tilde{u}_h))_K.$$

Let us estimate the first term of the right-hand side. For any arbitrary $\omega \in \mathcal{P}_0^{k+1}(K)$, we have

$$\langle \omega, (\boldsymbol{q} - \widehat{\boldsymbol{q}}_h) \cdot \boldsymbol{n} \rangle_{\partial K} = \langle \omega, (\boldsymbol{q} - \boldsymbol{q}_h) \cdot \boldsymbol{n} \rangle_{\partial K} + \langle \omega, (\boldsymbol{q}_h - \widehat{\boldsymbol{q}}_h) \cdot \boldsymbol{n} \rangle_{\partial K}$$
$$= \sum_{i=1}^{3} T_i,$$

where

$$T_1 = (\nabla \omega, \boldsymbol{q} - \boldsymbol{q}_h)_K,$$
$$T_2 = (\omega, \nabla \cdot (\boldsymbol{q} - \boldsymbol{q}_h))_K,$$
$$T_3 = \langle \omega, (\boldsymbol{q}_h - \widehat{\boldsymbol{q}}_h) \cdot \boldsymbol{n} \rangle_{\partial K}.$$

By using Cauchy-Schwarz inequality, we get that

$$T_1 \leq \| \nabla \omega \|_{L^2(K;\boldsymbol{a})} \| \boldsymbol{q} - \boldsymbol{q}_h \|_{L^2(K;\boldsymbol{c})}.$$

By using the definition of the Raviart-Thomas projection $\boldsymbol{\Pi}^{\mathrm{RT}}$, (2.3), and by using its commutativity property, we get that, for any $r \in [0, k]$,

$$T_2 = (\omega, f - \mathsf{P}f)_K + (\omega, \nabla \cdot (\boldsymbol{\Pi}^{\mathrm{RT}}\boldsymbol{q} - \boldsymbol{q}_h))_K$$
$$\leq \| \omega \|_{L^2(K)} \left(h_K^r \, | f |_{H^r(K)} + h_K^{-1} \| \boldsymbol{q} - \boldsymbol{q}_h \|_{L^2(K;\boldsymbol{c})} \right)$$
$$\leq \| \nabla \omega \|_{L^2(K;\boldsymbol{a})} \left(h_K^{r+1} \, | f |_{H^r(K)} + \| \boldsymbol{q} - \boldsymbol{q}_h \|_{L^2(K;\boldsymbol{c})} \right)$$

by Poincaré's inequality. Finally, by the definition of the numerical trace $\widehat{\boldsymbol{q}}_h$, (1.4b),

$$
\begin{aligned}
T_3 &= \|\,\omega\,\|_{L^2(e_K^\tau)}\,\|\,(\boldsymbol{q}_h - \widehat{\boldsymbol{q}}_h)\cdot\boldsymbol{n}\,\|_{L^2(e_K^\tau)} \\
&\leq C\,\|\,\omega\,\|_{L^2(e_K^\tau)}\,h_K^{r+1/2}\,|\,f\,|_{H^r(K)},
\end{aligned}
$$

by Proposition 2.2 and identity (v) of Proposition 2.1. Applying a simple inverse inequality, we get

$$
T_3 \leq C\,\|\,\nabla\omega\,\|_{L^2(K;\boldsymbol{a})}\,h_K^{r+1}\,|\,f\,|_{H^r(K)}.
$$

As a consequence,

$$
\langle\omega,(\boldsymbol{q}-\widehat{\boldsymbol{q}}_h)\cdot\boldsymbol{n}\rangle_{\partial K} \leq C\,\|\,\nabla\omega\,\|_{L^2(K;\boldsymbol{a})}\left(\|\,\boldsymbol{q}-\boldsymbol{q}_h\,\|_{L^2(K;\boldsymbol{c})} + h_K^{r+1}\,|\,f\,|_{H^r(K)}\right).
$$

This implies that

$$
\begin{aligned}
\|\,\nabla(\mathsf{P}^{k+1}\tilde{u}-\tilde{u}_h)\,\|_{L^2(K;\boldsymbol{a})} &\leq \|\,\nabla(\mathsf{P}^{k+1}\tilde{u}-\tilde{u})\,\|_{L^2(K)} \\
&\quad + C\left(\|\,\boldsymbol{q}-\boldsymbol{q}_h\,\|_{L^2(K;\boldsymbol{c})} + h_K^{r+1}\,|\,f\,|_{H^r(K)}\right),
\end{aligned}
$$

and so,

$$
\begin{aligned}
\|\,\mathsf{P}^{k+1}\tilde{u}-\tilde{u}_h\,\|_{L^2(\Omega_h;\boldsymbol{a})} &\leq Ch\,\|\,\nabla(\mathsf{P}^{k+1}\tilde{u}-\tilde{u})\,\|_{L^2(\Omega_h)} \\
&\quad + C\left(h\,\|\,\boldsymbol{q}-\boldsymbol{q}_h\,\|_{L^2(\Omega_h;\boldsymbol{c})} + h^{r+2}\,|\,f\,|_{H^r(\Omega_h)}\right) \\
&\leq C\,h^{r+2}\left(|\,\boldsymbol{q}\,|_{\boldsymbol{H}^{r+1}(\Omega_h)} + |\,u\,|_{H^{r+2}(\Omega_h)}\right),
\end{aligned}
$$

by Theorem 2.5 and the well-known approximation properties of P^{k+1}.

Let us now estimate the error $\overline{u}-\overline{u}_h$. We begin by considering the case $k \geq 1$. In this case, since $\overline{u}-\overline{u}_h = \overline{\mathbb{P}(u-u_h)}$, we get

$$
\|\,\overline{u}-\overline{u}_h\,\|_{L^2(K)} \leq \|\,\mathbb{P}u-u_h\,\|_{L^2(K)} \leq C\,\mathfrak{C}_\kappa^{r,0}(\boldsymbol{q})\,h^{r+2}
$$

by Theorem 2.6. Note that by Theorem 2.4, $\mathsf{P}^{k-1}u_h$ is independent of the value of the local stabilization parameters τ. This implies that the same is true for $\overline{u}_h$ and so, we get that

$$
\|\,\overline{u}-\overline{u}_h\,\|_{L^2(K)} \leq C\,\mathfrak{C}_0^{r,0}(\boldsymbol{q})\,h^{r+2}.
$$

It remains to consider the case $k=0$ and $f=0$. We have that

$$
\begin{aligned}
\overline{u}-\overline{u}_h &= \frac{1}{|K|}\int_K u(x)\,dx - \frac{1}{d}\sum_{e\in\partial K}\widehat{u}_h|_e \\
&= \frac{1}{|K|}\int_K (u-\mathsf{P}^1 u)(x)\,dx \\
&\quad + \frac{1}{|K|}\int_K \mathsf{P}^1 u(x)\,dx - \frac{1}{d}\sum_{e\in\partial K}\frac{1}{|e|}\int_e \mathsf{P}_\partial(\mathsf{P}^1 u)\,d\gamma \\
&\quad - \frac{1}{d}\sum_{e\in\partial K}\frac{1}{|e|}\int_e (\mathsf{P}_\partial(u-\mathsf{P}^1 u))\,d\gamma \\
&\quad - \frac{1}{d}\sum_{e\in\partial K}\frac{1}{|e|}\int_e (\widehat{u}_h-\mathsf{P}_\partial u)\,d\gamma.
\end{aligned}
$$

Since, for any function $\omega \in \mathcal{P}^1(K)$, we have that

$$\frac{1}{|K|} \int_K \omega(x)\, dx - \frac{1}{d} \sum_{e \in \partial K} \frac{1}{|e|} \int_e \mathsf{P}_\partial \omega \, d\gamma = 0,$$

we readily obtain that

$$\begin{aligned}
\|\,\overline{u} - \overline{u}_h\,\|_{L^2(K)} \leq &C \,\|\, u - \mathsf{P}^1 u\,\|_{L^2(K)} + C\, h_K \,|\, u - \mathsf{P}^1 u\,|_{H^1(K)} \\
&+ C\, h_k \,\|\, \widehat{u}_h - \mathsf{P}_\partial u\,\|_{L^2(\partial K)},
\end{aligned}$$

and so,

$$\|\,\overline{u} - \overline{u}_h\,\|_{L^2(\Omega_h)} \leq C\, h^2 \left(\mathcal{C}_\kappa(\boldsymbol{q}) + |\, u\,|_{H^2(\Omega_h)} \right).$$

Since, by Theorem 2.4, λ_h is independent of the value of the local stabilization parameter τ, so is $\overline{u}_h$ and so

$$\|\,\overline{u} - \overline{u}_h\,\|_{L^2(K)} \leq C\, h^2 \left(\mathcal{C}_0(\boldsymbol{q}) + |\, u\,|_{H^2(\Omega_h)} \right).$$

This completes the proof of Theorem 2.9.

4. NUMERICAL EXPERIMENTS

In this section, we carry out numerical experiments to validate the theoretical convergence properties of the SF-H$_k$ method.

To do that, we use uniform meshes obtained by discretizing $\Omega = (-\frac{1}{2}, \frac{1}{2}) \times (-\frac{1}{2}, \frac{1}{2})$ with squares of side 2^{-l} which are then divided into two triangles as indicated in Figure 1; the resulting mesh is denoted by "mesh=l".
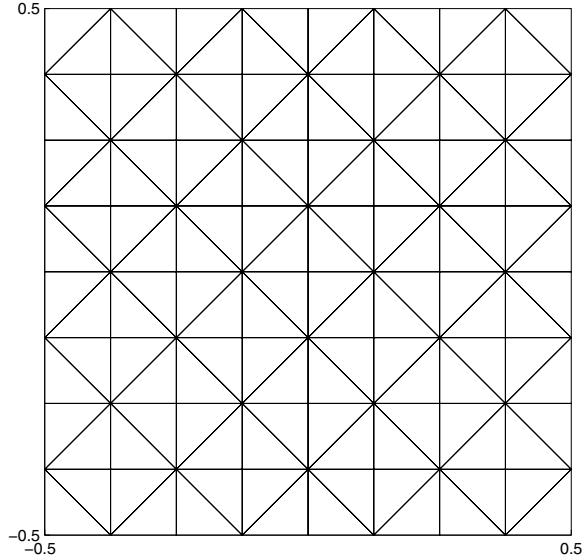

FIGURE 1. Example of a mesh with $h = 1/2^3$.

The test problem is obtained by taking $\partial\Omega_N = \emptyset$, $\boldsymbol{c} = \boldsymbol{I}$ and choosing g and f so that the exact solution is $u(x, y) = \cos(\pi x)\cos(\pi y)$ on the domain Ω. The history of convergence of the SF-H method with

$$\tau_K = 1/h = 2^l,$$

on the "mesh=l", is displayed in Table 5 for polynomials of degree $k = 0$, $k = 1$ and $k = 2$. We observe optimal convergence rates of the quantities $\|u - u_h\|_{L^2(\Omega)}$ and $\|q - q_h\|_{L^2(\Omega)}$ for $k = 0, 1, 2$ as predicted by Theorems 2.5 and 2.7. We also see that $\|P_\partial u - \lambda_h\|_{L^2(\mathcal{E}_h;h)}$ and $\|u - u_h^*\|_{L^2(\Omega)}$ superconverges with rate $O(h^{k+2})$ for $k = 1, 2$ just as predicted by Theorems 2.8 and 2.9. These results do not guarantee that these quantities are superconvergent if $k = 0$ and $f \not\equiv 0$. Since we do not observe superconvergence, we can conclude that the theoretical results for such a case are actually sharp.

TABLE 5. History of convergence of the SF-H$_k$ method.

k	mesh ℓ	$\|u - u_h\|_{L^2(\Omega)}$ error	order	$\|q - q_h\|_{L^2(\Omega)}$ error	order	$\|P_\partial u - \lambda_h\|_{L^2(\mathcal{E}_h;h)}$ error	order	$\|u - u_h^\star\|_{L^2(\Omega)}$ error	order
	1	.11e+1	-	.17e+1	-	.28e-0	-	.22e-0	-
	2	.36e-0	1.54	.78e-0	1.12	.92e-1	1.61	.57e-1	1.96
0	3	.12e-0	1.50	.41e-0	0.94	.35e-1	1.37	.19e-1	1.57
	4	.53e-1	1.29	.21e-0	0.97	.14e-1	1.21	.79e-2	1.27
	5	.24e-1	1.13	.10e-0	0.98	.69e-2	1.15	.36e-2	1.14
	6	.12e-1	1.05	.53e-1	0.99	.32e-2	1.12	.17e-2	1.10
	1	.21e-0	-	.23e-0	-	.31e-1	-	.21e-1	-
	2	.43e-1	2.27	.12e-0	0.94	.75e-2	2.02	.40e-2	2.38
1	3	.78e-2	2.47	.31e-1	1.94	.10e-2	2.96	.53e-3	2.91
	4	.17e-2	2.19	.79e-2	1.99	.12e-3	3.00	.68e-4	2.98
	5	.42e-3	2.05	.20e-2	2.00	.15e-4	3.01	.85e-5	2.99
	6	.10e-3	2.01	.50e-3	2.00	.19e-5	3.00	.11e-6	2.99
	1	.68e-1	-	.89e-1	-	.72e-2	-	.81e-2	-
	2	.38e-2	4.12	.91e-2	3.29	.41e-3	4.12	.40e-3	4.35
2	3	.32e-3	3.58	.12e-2	2.96	.27e-4	3.93	.25e-4	3.97
	4	.32e-4	3.31	.15e-3	2.98	.18e-5	3.96	.16e-5	3.99
	5	.37e-5	3.12	.19e-4	2.99	.11e-6	3.98	.10e-6	4.00
	6	.45e-7	3.01	.23e-5	3.00	.70e-8	3.99	.63e-8	4.00

Next we explore the effect of the size of τ_K on the quality of the approximation. In Table 6, we see that as τ diminishes the quality of the approximation to u deteriorates. However, the effect of taking $\tau = 1/h^2$ or $\tau = 1/h$ is almost negligible especially when the grids are not coarse. We also see that the order of convergence is $k + 1$ for $\tau = 1/h^2, \tau = 1/h$ and $\tau = 1$, but it is only k for $\tau = h$. This is in perfect agreement with Corollary 2.7.

We end with an example where the exact solution is harmonic, that is, $\boldsymbol{c} = \boldsymbol{I}, f \equiv 0$, and display the convergence rates for $k = 0$ in Table 7. We take $\partial\Omega_N = \emptyset$ and choose g and so that $u(x, y) = e^x \sin(y)$ is the solution. We see that the quantities $\|P_\partial u - \lambda_h\|_{L^2(\mathcal{E}_h;h)}$ and $\|u - u_h^*\|_{L^2(\Omega)}$ superconverge with the rate $O(h^2)$ as our theoretical results predict.

TABLE 6. Effect of τ on the convergence of $\| u - u_h \|_{L^2(\Omega)}$.

k	mesh ℓ	$\tau = 1/h^2$ error	order	$\tau = 1/h$ error	order	$\tau = 1$ error	order	$\tau = h$ error	order
0	1	.61e+0	-	.11e+1	-	.21e+1	-	.40e+1	-
	2	.21e+0	1.50	.36e-0	1.54	.11e+1	0.88	.42e+1	-0.06
	3	.95e-1	1.17	.12e-0	1.50	.57e-0	0.97	.43e+1	-0.02
	4	.46e-1	1.04	.53e-1	1.29	.28e-0	1.00	.43e+1	-0.00
	5	.23e-1	1.02	.24e-1	1.13	.14e-0	1.01	.43e+1	0.00
	6	.11e-1	1.01	.12e-1	1.05	.70e-1	1.01	.43e+1	0.00
1	1	.15e+0	-	.21e-0	-	.35e-0	-	.67e-0	-
	2	.27e-1	2.47	.43e-1	2.27	.14e-0	1.34	.54e-0	0.29
	3	.66e-2	2.06	.78e-2	2.47	.35e-1	1.98	.28e-0	0.97
	4	.16e-2	2.00	.17e-2	2.19	.88e-2	2.00	.14e-0	0.99
	5	.41e-3	2.00	.42e-3	2.05	.22e-2	2.00	.67e-1	1.00
	6	.10e-3	2.00	.10e-3	2.01	.60e-3	2.00	.35e-1	1.00
2	1	.33e-1	-	.68e-1	-	.14e-0	-	.28e-0	-
	2	.19e-2	4.10	.38e-2	4.12	.14e-1	3.31	.56e-1	2.33
	3	.23e-3	3.07	.32e-3	3.58	.18e-2	2.96	.14e-1	1.98
	4	.28e-4	3.00	.32e-4	3.31	.23e-3	2.99	.35e-2	2.00
	5	.36e-5	3.00	.37e-5	3.12	.28e-4	3.00	.88e-3	2.00
	6	.45e-6	3.00	.45e-7	3.01	.35e-5	3.00	.22e-3	2.00

TABLE 7. History of convergence for a harmonic exact solution.

k	mesh ℓ	$\|u - u_h\|_{L^2(\Omega)}$ error	order	$\|q - q_h\|_{L^2(\Omega)}$ error	order	$\|P_\partial u - \lambda_h\|_{L^2(\mathcal{E}_h;h)}$ error	order	$\|u - u_h^\star\|_{L^2(\Omega)}$ error	order
0	1	.17e-0	-	.22e-0	-	.29e-1	-	.23e-1	-
	2	.87e-1	0.94	.11e-0	0.96	.79e-2	1.84	.62e-2	1.87
	3	.44e-1	0.99	.57e-1	0.98	.21e-2	1.90	.16e-2	1.93
	4	.22e-1	0.99	.29e-1	0.99	.55e-3	1.96	.41e-3	1.97
	5	.11e-1	1.00	.14e-1	1.00	.10e-3	1.98	.10e-3	1.99
	6	.55e-2	1.00	.72e-2	1.00	.35e-4	2.00	.26e-4	2.00

5. CONCLUDING REMARKS

The error analysis carried out here for the SF-H$_k$ method also holds for the hybridized versions of the RT$_K$ and the BDM$_k$ methods. We simply have to replace the local space $\boldsymbol{\mathcal{P}}^k(K) \times \mathcal{P}^k(K)$ by the local space $\boldsymbol{V}(K) \times W(K)$ given by Table 1, use the definition of the local stabilization parameter τ given in Table 2, and suitably define the projection $(\boldsymbol{\Pi}, \mathbb{P})$. Indeed, with such changes, the first four properties of Proposition 2.1, on which the whole analysis is based, hold. For this reason, we can consider this analysis to be a *unifying* analysis of these three methods.

A study of the optimal way to choose the local stabilization parameter τ falls beyond the scope of this paper and will be carried out elsewhere. Extensions of these results to more general second-order elliptic equations and other boundary conditions are straightforward. The extension of these results to the case of hanging nodes, variable-degree approximations and curved domains constitute the subject of ongoing work.

References

[1] D. N. Arnold and F. Brezzi, *Mixed and nonconforming finite element methods: implementation, postprocessing and error estimates*, RAIRO Modél. Math. Anal. Numér. **19** (1985), 7–32. MR813687 (87g:65126)

[2] D. N. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini, *Unified analysis of discontinuous Galerkin methods for elliptic problems*, SIAM J. Numer. Anal. **39** (2002), 1749–1779. MR1885715 (2002k:65183)

[3] J. H. Bramble and J. Xu, *A local post-processing technique for improving the accuracy in mixed finite-element approximations*, SIAM J. Numer. Anal. **26** (1989), no. 6, 1267–1275. MR1025087 (90m:65193)

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

[5] P. Castillo, B. Cockburn, I. Perugia, and D. Schötzau, *An a priori error analysis of the local discontinuous Galerkin method for elliptic problems*, SIAM J. Numer. Anal. **38** (2000), 1676–1706. MR1813251 (2002k:65175)

[6] F. Celiker and B. Cockburn, *Superconvergence of the numerical traces of discontinuous Galerkin and hybridized mixed methods for convection-diffusion problems in one space dimension*, Math. Comp. **76** (2007), 67–96. MR2261012

[7] Z. Chen, *Equivalence between and multigrid algorithms for nonconforming and mixed methods for second-order elliptic problems*, East-West J. Numer. Math. **4** (1996), 1–33. MR1393063 (98c:65184)

[8] P. Ciarlet, *The finite element method for elliptic problems*, North-Holland, Armsterdam, 1978. MR0520174 (58:25001)

[9] B. Cockburn and B. Dong, *An analysis of the minimal dissipation local discontinuous Galerkin method for convection-diffusion problems*, J. Sci. Comput. **32** (2007), 233–262. MR2320571

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

[11] B. Cockburn, J. Gopalakrishnan, and R. Lazarov, *Unified hybridization of discontinuous Galerkin, mixed and continuous Galerkin methods for second order elliptic problems*, Submitted.

[12] B. Cockburn and C.-W. Shu, *The local discontinuous Galerkin method for time-dependent convection-diffusion systems*, SIAM J. Numer. Anal. **35** (1998), 2440–2463. MR1655854 (99j:65163)

[13] J. Douglas, Jr. and J. E. Roberts, *Global estimates for mixed methods for second order elliptic equations*, Math. Comp. **44** (1985), 39–52. MR771029 (86b:65122)

[14] L. Gastaldi and R. H. Nochetto, *Sharp maximum norm error estimates for general mixed finite element approximations to second order elliptic equations*, RAIRO Modél. Math. Anal. Numér. **23** (1989), 103–128. MR1015921 (91b:65125)

[15] J. T. Oden and J. K. Lee, *Dual-mixed hybrid finite element method for second-order elliptic problems*, Mathematical aspects of finite element methods (Proc. Conf., Consiglio Naz. delle Ricerche (C.N.R.), Rome, 1975), Springer, Berlin, 1977, pp. 275–291. Lecture Notes in Math., Vol. 606. MR0520341 (58:25010)

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

[17] Rolf Stenberg, *A family of mixed finite elements for the elasticity problem*, Numer. Math **53** (1988), 513–538. MR954768 (89h:65192)

[18] Rolf Stenberg, *Postprocessing schemes for some mixed finite elements*, RAIRO Modél. Math. Anal. Numér. **25** (1991), 151–167. MR1086845 (92a:65303)

School of Mathematics, 127 Vincent Hall, University of Minnesota, Minneapolis, Minnesota 55455

E-mail address: cockburn@math.umn.edu

Division of Applied Mathematics, Brown University, Providence, Rhode Island 02912

E-mail address: bdong@dam.brown.edu

School of Mathematics, 127 Vincent Hall, University of Minnesota, Minneapolis, Minnesota 55455

E-mail address: guzma033@umn.edu