MATHEMATICS OF COMPUTATION
Volume 83, Number 285, January 2014, Pages 87–111
S 0025-5718(2013)02711-1
Article electronically published on May 16, 2013

ANALYSIS OF VARIABLE-DEGREE HDG METHODS FOR CONVECTION-DIFFUSION EQUATIONS. PART II: SEMIMATCHING NONCONFORMING MESHES

YANLAI CHEN AND BERNARDO COCKBURN

ABSTRACT. In this paper, we provide a projection-based analysis of the h-version of the hybridizable discontinuous Galerkin methods for convection-diffusion equations on semimatching nonconforming meshes made of simplexes; the degrees of the piecewise polynomials are allowed to vary from element to element. We show that, for approximations of degree k on all elements, the order of convergence of the error in the diffusive flux is $k+1$ and that of a projection of the error in the scalar unknown is 1 for $k=0$ and $k+2$ for $k>0$. We also show that, for the variable-degree case, the projection of the error in the scalar variable is h times the projection of the error in the vector variable, provided a simple condition is satisfied for the choice of the degree of the approximation on the elements with hanging nodes. These results hold for any (bounded) irregularity index of the nonconformity of the mesh. Moreover, our analysis can be extended to hypercubes.

1. INTRODUCTION

In this paper, we continue our work [6] on a priori error analysis of variable-degree hybridizable discontinuous Galerkin (HDG) methods on nonconforming meshes for the model convection-diffusion problem

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

$$\nabla \cdot (\mathbf{q} + u\,\mathbf{v}) = f \qquad \text{in } \Omega, \tag{1.1b}$$

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

Here Ω is a polyhedral domain in $\mathbb{R}^d$, $f \in L^2(\Omega)$, and $c = c(x)$ is a symmetric $d \times d$ matrix function that is uniformly positive definite on Ω with components in $L^\infty(\Omega)$. We take the velocity $\mathbf{v} \in \boldsymbol{W}^{1,\infty}(\Omega)$ to be *divergence-free.*

Whereas in [6] we showed that the nonconformity of the meshes could degrade both the optimal convergence of the vector variable and the superconvergence of the projection of the error in the scalar variable (without degrading the optimality of its convergence), here we show that if the so-called *semimatching* nonconforming meshes are used and the polynomial degrees are suitably chosen on the elements having hanging nodes, both the *optimal* convergence of the vector variable as well as the *superconvergence* of the projection of the error in the scalar variable are *not* degraded. This holds for semimatching nonconforming meshes with any (bounded)

Received by the editor July 5, 2011 and, in revised form, December 21, 2011, and May 4, 2012.
2010 *Mathematics Subject Classification.* Primary 65M60, 65N30.
The research of the second author was partially supported by the National Science Foundation (grant DMS-0712955).

©2013 American Mathematical Society

irregularity index (the maximum difference of refinement levels between adjacent elements in the mesh).

Let us put our result in proper context. The HDG methods were introduced in the framework of diffusion problems in [10] in order to ensure that the only globally-coupled degrees of freedom are those of the approximate trace of the scalar variable on the boundaries of the elements. This distinctive feature renders them as efficiently implementable as the hybridized version of the traditional mixed methods [18, 3] of the corresponding index. Moreover, they can be more accurate than all DG methods considered in the unified analysis performed in [1]. This is the case of the so-called local DG-hybridizable (LDG-H) method which, for simplicity, will be referred to as the HDG method.

Indeed, these HDG methods were shown [7, 11, 12] to provide optimally convergent approximate fluxes and to have superconvergent approximations for the scalar variable similar to those exhibited by the classical mixed methods [18, 3, 4]. This was established for the first time for a special HDG method in [7] and then extended to a wide class of HDG methods in [12]. More recently, a projection-based approach to their analysis was proposed in [11] for simplicial elements and in [15] for much more general elements, which also provides optimally convergent approximate fluxes and superconvergent approximations for the scalar variable. Thus, for conforming meshes of simplexes and if polynomials of degree $k \geq 0$ are used for all the components of the flux and for the scalar variable, the HDG methods converge with order $k+1$ in the flux and superconverge with order $k+1+\min\{1,k\}$ in a projection of the scalar variable, whereas all the DG methods considered in [1] converge only with order k in the flux and with order $k+1$ in the scalar variable. These orders of convergence cannot improve when convection is added; see for example [2, 16]. However, for a special LDG method for pure diffusion defined on Cartesian meshes, the order of convergence of the flux does increase to $k+1/2$; see [13].

We are interested in extending the above-mentioned projection-based analysis to explore how the convergence properties of the HDG methods are affected when we (i) include convection, (ii) allow the polynomial degree to vary from element to element, and (iii) allow the use of nonconforming meshes. This has not been done so far. Note that, since the superconvergence of the scalar variable does *not* take place in the convection-dominated regime (see the numerical experiments in [17, 8]), we are only interested in the diffusion-dominated regime. The motivation for the incorporation of the last two properties comes from the fact that DG methods and, in particular, HDG methods, are extremely well-suited for adaptivity.

In [6] we showed that, when going from conforming meshes to general nonconforming meshes, the order of convergence of the flux can degrade from $k+1$ to $k+1/2$, and that the order of superconvergence of a projection of the scalar variable can decay from $k+1+\min\{k,1\}$ to $k+1$. Here we show that, when nonconforming semimatching meshes are used (and the degree of the approximation on the elements with hanging nodes is suitably chosen) this degradation does not take place. In the case of uniform-degree, we summarize the orders of convergence given by our main result in Table 1.

The paper is organized as follows. In Section 2, we describe the semimatching nonconforming meshes we are going to consider, define the HDG method and state and discuss our main result, Theorem 2.1. In Section 3, we display the main steps of its proof and defer to Section 4 the technical and long proof of its key estimate.

TABLE 1. Comparison of the convergence orders in the $L^2(\Omega)$-norm for DG and HDG methods with uniform degree $k \geq 0$ in terms of the conformity properties of the meshes.

method	conformity of the meshes $\mathcal{T}_h$	order (flux)	order (scalar)
DG pure diffusion	conforming [1]	k	$k+1$
LDG pure diffusion	conforming [13] Cartesian meshes	$k+1/2$	$k+1$
LDG pure diffusion	nonconforming [5]	k	$k+1$
HDG pure diffusion	conforming [11]	$k+1$	$k+1+\min\{k,1\}$ projection of the scalar variable
HDG	nonconforming [6]	$k+1/2$	$k+1$
HDG	nonconforming semimatching	$k+1$	$k+1+\min\{k,1\}$ projection of the scalar variable

We end in Section 5 by sketching the extension of our results to HDG methods defined on squares and cubes, and (for the purely diffusive case) to mixed methods.

2. Main Results

In this section, we state and discuss our main results. First, we describe in detail the structure of the nonconforming meshes we are going to consider. We then define the HDG method. Finally, we provide upper bounds of the L^2-norm of a suitably defined projection of the errors.

2.1. **The semimatching nonconforming meshes.** The semimatching nonconforming meshes $\mathcal{T}_h$ of Ω we are going to consider here are defined in terms of any family of sequentially refined *conforming* triangulations of Ω, $\{\mathsf{T}_h^\ell\}_{\ell\geq 1}$. We assume that they are such that, for every level index $\ell \geq 1$,

(2.1a) T_h^ℓ is made of simplexes K such that $\dfrac{h_\mathsf{K}}{\rho_\mathsf{K}} \leq \sigma$,

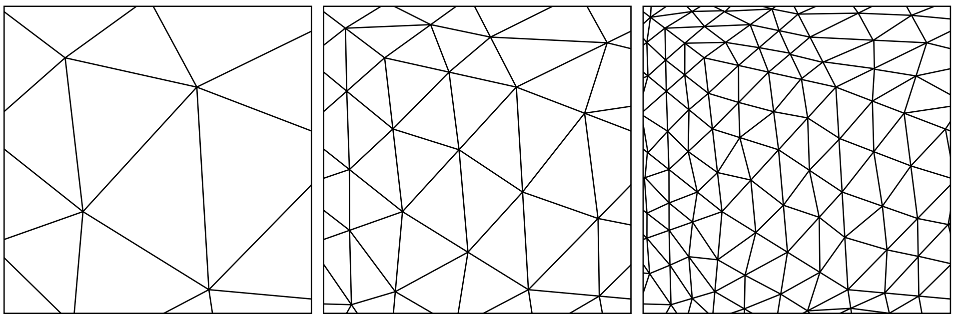
(2.1b) $\mathsf{T}_h^{\ell+1}$ is a refinement of T_h^ℓ such that no element of T_h^ℓ is unrefined,

(2.1c) $$\forall\, \mathsf{K} \in \mathsf{T}_h^\ell : \max_{\mathsf{K}'\in\mathsf{T}_h^{\ell+1}:\mathsf{K}'\subset\mathsf{K}} h_{\mathsf{K}'} \leq \kappa \min_{\mathsf{K}'\in\mathsf{T}_h^{\ell+1}:\mathsf{K}'\subset\mathsf{K}} h_{\mathsf{K}'},$$

(2.1d) $$\forall\, \mathsf{K} \in \mathsf{T}_h^\ell : \max_{\mathsf{K}'\in\mathsf{T}_h^{\ell+n}:\mathsf{K}'\subset\mathsf{K}} h_{\mathsf{K}'} \leq \mathsf{c}\,\eta^n\, h_\mathsf{K}.$$

The simplexes K and their faces F are assumed to be *open* sets in $\mathbb{R}^d$ and $\mathbb{R}^{d-1}$, respectively. As usual, h_K denotes the diameter of K and ρ_K the radius of the biggest ball included in K. Note that the first property states that the simplexes of each of the triangulations T_h^ℓ are (uniformy) shape-regular. The third property states that their successive refinements are locally uniform. The last property states that their nth refinement is made of simplexes whose diameters are of the order of η^n times the original diameter; here we assume that $\eta \in (0,1)$. The shape-regularity constant σ, the local uniformity constant κ, and local refinement factor η will appear in our error estimate.

Let us give examples of families of meshes with different refinement factors η. Take Ω to be the unit square and take T_h^1 to be the triangulation obtained by dividing Ω into two identical triangles. Now define $\mathsf{T}_h^{\ell+1}$ as the refinement of T_h^ℓ obtained by *bisecting* its triangles into two right-angled triangles; in this case, $\eta = 1/\sqrt{2}$. If the refinement is obtained by four congruent triangles, we have that $\eta = 1/2$. See another example in Figure 1.



T_h^1 T_h^2 T_h^3

FIGURE 1. An example of a family of triangulations $\{\mathsf{T}_h^\ell\}_{\ell\geq 1}$ for which $\eta = 1/2$.

We can now define the triangulations we are going to consider. We say that $\mathcal{T}_h$ is a semimatching nonconforming mesh if it is a collection of simplexes $\{K\}$ from $\{\mathsf{T}_h^\ell\}_{\ell\geq 1}$. That is, for each element $K \in \mathcal{T}_h$ there is a set $\{\mathsf{K}_K^\ell\}_{\ell=1}^{\ell_K}$ such that

$$\mathsf{K}_K^\ell \in \mathsf{T}_h^\ell, \text{ for } \ell = 1, \cdots, \ell_K, \tag{2.2a}$$

$$\mathsf{K}_K^\ell \supset K, \text{ for } \ell = 1, \cdots, \ell_K, \tag{2.2b}$$

$$\mathsf{K}_K^{\ell_K} = K. \tag{2.2c}$$

We call K_K^ℓ the ℓ^{th} ancestor of K, with its ℓ_K^{th} ancestor being itself. See an example in Figure 2.

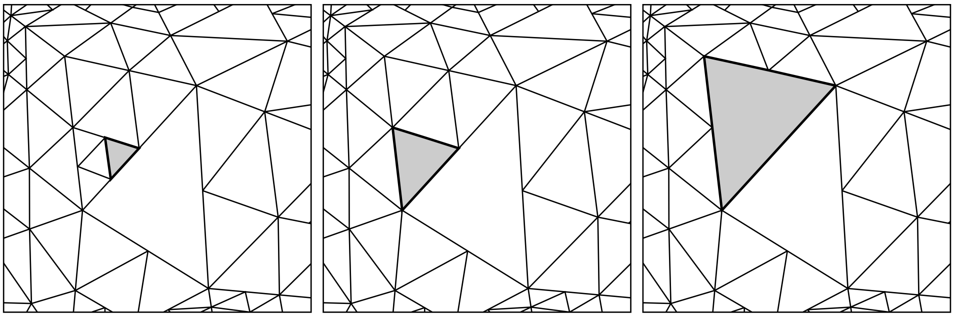
The notation associated to the meshes is the following. We set $\partial\mathcal{T}_h := \{\partial\mathsf{K} : \mathsf{K} \in \mathcal{T}_h\}$. The set of faces of the simplex K is denoted by $\mathcal{F}(K)$ and the set of faces F of the triangulation $\mathcal{T}_h$ by $\mathcal{E}_h$. Since hanging nodes are permitted, $\mathcal{E}_h$ should be understood as the set containing the smallest common $(d-1)$-dimensional interfaces of neighboring elements of $\mathcal{T}_h$. A face $F \in \mathcal{E}_h$ is called a boundary face if it is the face of a simplex $K \in \mathcal{T}_h$ lying on the boundary of Ω. Every other face $F \in \mathcal{E}_h$ is called an interior face and is of the form $F = \partial K^+ \cap \partial K^-$ where $K^\pm \in \mathcal{T}_h$. In addition, we call ℓ_K the refinement level for K.

2.2. **The HDG method.** We can now define the HDG method we are going to study. Given a mesh $\mathcal{T}_h$, we define the following finite element spaces:

$$\mathbf{V}_h = \{\mathbf{r} \in \mathbf{L}^2(\mathcal{T}_h) : \quad \mathbf{r}|_K \in \mathbf{P}_{k(K)}(K) \quad \forall\, K \in \mathcal{T}_h\}, \tag{2.3a}$$

$$W_h = \{w \in L^2(\mathcal{T}_h) : \quad w|_K \in P_{k(K)}(K) \quad \forall\, K \in \mathcal{T}_h\}, \tag{2.3b}$$

$$M_h = \{\mu \in L^2(\mathcal{E}_h) : \quad \mu|_F \in P_{k(F)}(F) \quad \forall\, F \in \mathcal{E}_h, \tag{2.3c}$$



$K = \mathsf{K}_K^3 \in \mathsf{T}_h^3, \ell_K = 3$ $\mathsf{K}_K^2 \in \mathsf{T}_h^2$ $\mathsf{K}_K^1 \in \mathcal{T}_h^1$

FIGURE 2. An example of 2-irregular nonconforming mesh $\mathcal{T}_h$ (left) and of the set $\{\mathsf{K}_K^\ell\}_{\ell=1}^{\ell_K}$ (in gray).

where $\mathbf{P}_k(D) = [P_k(D)]^d$ and $P_k(D)$ is the space of polynomials on the domain D of total degree at most k. Here, following [10], we take

$$k(F) = k(K) \qquad \text{if } F = \partial K \cap \partial\Omega, \tag{2.3d}$$

$$k(F) = \max\{k(K^+), k(K^-)\} \quad \text{if } F = \partial K^+ \cap \partial K^-. \tag{2.3e}$$

However, we *further* require that

$$k(K^+) \geq k(K^-) \text{ whenever } \ell_{K^+} \geq \ell_{K^-}. \tag{2.3f}$$

Note that the last condition is *not* necessary for the HDG method to be well defined; indeed, it is not required in [10]. We are adding it here because it seems to be needed for technical reasons in order to ensure optimal convergence orders in the approximate flux and superconvergence of the projection of the error in the scalar variable. Examples of triangulations satisfying and not satisfying this last condition are given in Figure 3.

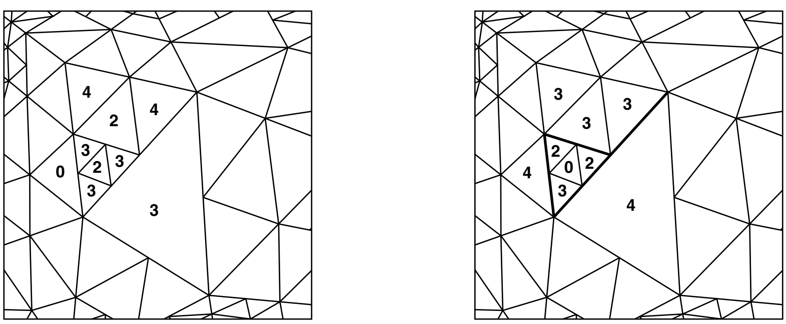


FIGURE 3. Illustration of the condition (2.3f) on the relation between local polynomial degrees $k(K)$ (the numbers in the triangles) and the level of refinement ℓ_K. The triangulation in the left does satisfy the condition whereas the triangulation in the right does not. Therein, the thick solid line points to the neighboring triangles for which the condition is not satisfied.

The HDG method seeks the approximations u_h in W_h, $\mathbf{q}_h$ in $\mathbf{V}_h$, and $\widehat{u}_h$ in M_h by requiring that

$$(c\,\mathbf{q}_h, \mathbf{r})_{\mathcal{T}_h} - (u_h, \nabla\cdot\mathbf{r})_{\mathcal{T}_h} + \langle \widehat{u}_h, \mathbf{r}\cdot\mathbf{n}\rangle_{\partial\mathcal{T}_h} = 0, \tag{2.4a}$$

$$-(\mathbf{q}_h + u_h\mathbf{v}, \nabla w)_{\mathcal{T}_h} + \langle (\widehat{\mathbf{q}_h + u_h\mathbf{v}})\cdot\mathbf{n}, w\rangle_{\partial\mathcal{T}_h} = (f, w)_{\mathcal{T}_h}, \tag{2.4b}$$

$$\langle \mu, \widehat{u}_h\rangle_{\partial\Omega} = \langle \mu, g\rangle_{\partial\Omega}, \tag{2.4c}$$

$$\langle \mu, (\widehat{\mathbf{q}}_h + \widehat{u_h\mathbf{v}})\cdot\mathbf{n}\rangle_{\partial\mathcal{T}_h\setminus\partial\Omega} = 0, \tag{2.4d}$$

hold for all $\mathbf{r} \in \mathbf{V}_h$, $w \in W_h$ and $\mu \in M_h$, where the numerical trace for the total flux $\widehat{\mathbf{q}}_h + \widehat{u_h\mathbf{v}}$ is given by

$$\mathbf{q}_h + \widehat{u_h\mathbf{v}} = \widehat{\mathbf{q}_h + \widehat{u}_h\mathbf{v}} + \tau(u_h - \widehat{u}_h)\mathbf{n} \qquad \text{on } \partial\mathcal{T}_h. \tag{2.4e}$$

For the method to be well defined, the stabilization function τ has to be suitably defined. As shown in [6], the following assumptions are sufficient to guarantee this:

A1. For any simplex $K \in \mathcal{T}_h$, $\min(\tau - \frac{1}{2}\mathbf{v}\cdot\mathbf{n})|_{\partial K} \geq \gamma_0 > 0$.
A2. On any face $F \in \mathcal{E}_h$, τ is a constant.

Here, $\mathbf{n}$ denotes the unit outward normal to K.

Finally, we introduce a new approximation to u, u_h^*, defined on each element K as the element of $P_{k(K)+1}(K)$ that satisfies

$$\begin{aligned}(\nabla u_h^*, \nabla\omega)_K &= -\,(c\,\boldsymbol{q}_h, \nabla\omega)_K \quad \forall\, \omega \in P_{k(K)+1}(K),\\ (u_h^*, 1)_K &= (u_h, 1)_K.\end{aligned}$$

2.3. **The a priori error estimates.** The estimates we obtain are upper bounds on L^2-norms of a projection of the errors in both the vector variable and the scalar variable. We begin by introducing such a projection. We also introduce its dual which is required since we are using a duality argument to estimate the error in the scalar variable.

2.3.1. *The projection.* The projection

$$\mathit{\Pi}_h(\mathbf{q}, u) = (\mathit{\Pi}_{\mathbf{V}}\mathbf{q}, \mathit{\Pi}_W u) \in \mathbf{V}_h \times W_h$$

is used to render the structure of the equations for the *projection of the errors*,

$$\epsilon_h^u := \mathit{\Pi}_W u - u_h, \tag{2.5a}$$

$$\boldsymbol{\epsilon}_h^{\mathbf{q}} := \mathit{\Pi}_{\mathbf{V}}\mathbf{q} - \mathbf{q}_h, \tag{2.5b}$$

$$\epsilon_h^{\widehat{u}} := P_M u - \widehat{u}_h, \tag{2.5c}$$

$$(\epsilon_h^{\widehat{\mathbf{q}}} + \boldsymbol{\epsilon}_h^{\widehat{u\mathbf{v}}})\cdot\mathbf{n} := P_M\left((\mathbf{q} + u\mathbf{v})\cdot\mathbf{n}\right) - (\widehat{\mathbf{q}}_h + \widehat{u_h\mathbf{v}})\cdot\mathbf{n}, \tag{2.5d}$$

as close as possible as those defining the HDG method itself. Here, P_M is the L^2-projection from $L^2(\partial\mathcal{T}_h)$ into the finite element space

$$\mathbb{M}_h := \{m \in L^2(\partial\mathcal{T}_h) : m|_F \in P_{k(F)}(F) \quad \text{for all faces } F \text{ of } K \in \mathcal{T}_h\}. \tag{2.6}$$

Note that, in general, $P_M\eta$ is double-valued on the internal faces. However, if the function η is single-valued on internal faces, $P_M\eta$ is also single-valued therein and we are going to assume, by slightly abusing the notation, that $P_M\eta$ lies on M_h.

It is defined as follows [6]. On K, the projection $\mathit{\Pi}_h(\mathbf{q},u) = (\mathit{\Pi}_{\mathbf{V}}\mathbf{q}, \mathit{\Pi}_W u)$ is the element of $\mathbf{P}_{k(K)}(K) \times P_{k(K)}(K)$ which solves the equations

$$((\mathit{\Pi}_{\mathbf{V}}\mathbf{q} - \mathbf{q}) + (\mathit{\Pi}_W u - u)\,\mathbf{v}, \mathbf{r})_K = 0 \;\; \forall\, \mathbf{r} \in \mathbf{P}_{k(K)-1}(K), \tag{2.7a}$$

$$(\mathit{\Pi}_W u - u, w)_K = 0 \;\; \forall\, w \in P_{k(K)-1}(K), \tag{2.7b}$$

$$\langle(\mathit{\Pi}_{\mathbf{V}}\mathbf{q} - \mathbf{q} + (P_M u - u)\mathbf{v})\cdot\mathbf{n} + \tau(\mathit{\Pi}_W u - u), \mu\rangle_F = 0 \;\; \forall\, \mu \in P_{k(K)}(F), \tag{2.7c}$$

for all faces F of the simplex K.

2.3.2. *The dual problem and the dual projection.* To obtain the error in the scalar variable, we use the so-called dual problem which we introduce next. For any given $\Theta \in L^2(\Omega)$, we define $(\mathbf{\Phi}, \Psi)$ to be the solution of

$$c\,\mathbf{\Phi} + \nabla\Psi = 0 \quad \text{in} \quad \Omega, \tag{2.8a}$$

$$\nabla\cdot(\mathbf{\Phi} - \Psi\,\mathbf{v}) = \Theta \quad \text{in} \quad \Omega, \tag{2.8b}$$

$$\Psi = 0 \quad \text{on} \quad \partial\Omega. \tag{2.8c}$$

As usual, we assume that the solution of the dual problem satisfies the following regularity estimate:

$$\|\mathbf{\Phi}\|_{H^1(\Omega)} + \|\Psi\|_{H^2(\Omega)} \le C_{\mathrm{R}}\|\Theta\|_\Omega. \tag{2.9}$$

The projection $\mathit{\Pi}_h^*$ associated to the dual problem is defined as follows [6]. On any simplex $K \in \mathcal{T}_h$, $\mathit{\Pi}_h^*(\mathbf{\Phi},\Psi) = (\mathit{\Pi}_{\mathbf{V}}^*\mathbf{\Phi}, \mathit{\Pi}_W^*\Psi)$ is the element of $\mathbf{P}_{k(K)}(K) \times P_{k(K)}(K)$ determined by requiring that

$$((\mathit{\Pi}_{\mathbf{V}}^*\mathbf{\Phi} - \mathbf{\Phi}) - (\mathit{\Pi}_W^*\Psi - \Psi)\mathbf{v}, \mathbf{r})_K = 0 \;\; \forall\, \mathbf{r} \in \mathbf{P}_{k(K)-1}(K), \tag{2.10a}$$

$$(\mathit{\Pi}_W^*\Psi - \Psi, w)_K = 0 \;\; \forall\, w \in P_{k(K)-1}(K), \tag{2.10b}$$

$$\langle(\mathit{\Pi}_{\mathbf{V}}^*\mathbf{\Phi} - \mathbf{\Phi})\cdot\mathbf{n} + (\tau - \mathbf{v}\cdot\mathbf{n})(\mathit{\Pi}_W^*\Psi - \Psi), \mu\rangle_F = 0 \;\; \forall\, \mu \in P_{k(K)}(F), \tag{2.10c}$$

for all faces F of the simplex K. An analysis of the approximation properties of this projection can be found in [6].

2.3.3. *Estimates of the projection of the errors.* The estimates of the projection of the errors are expressed in terms of norms we define next. The $L^2(D)$-norm is denoted by $\|\cdot\|_D$. We also set

$$\|\mu\|^2_{\partial\mathcal{T}_h,\mathsf{w}} := \sum_{K\in\mathcal{T}_h}\sum_{F\in\mathcal{F}(K)} \mathsf{w}_F\,\|\mu\|^2_{F\setminus\partial\Omega}, \tag{2.11a}$$

for any $\mu \in L^2(\partial\mathcal{T}_h)$. The positive function w is a function on $\partial\mathcal{T}_h \setminus \partial\Omega$ defined as follows:

$$\mathsf{w} := h_F \text{ on } F \in \mathcal{F}(K) \text{ for } K \in \mathcal{T}_h, \tag{2.11b}$$

where h_F denotes the diameter of the face F.

Let us emphasize that the variability from element to element of the polynomial degree as well as the nonconformity of the mesh are going to be captured by a single term involving the difference $(P_M - P_{\mathcal{M}})$ on $\partial\mathcal{T}_h$. The operator $P_{\mathcal{M}}$ is the L^2-projection into the following space:

$$\mathcal{M}_h := \{w \in L^2(\partial\mathcal{T}_h) : w|_F \in P_{k(K)}(F) \text{ for all faces } F \text{ of } K \in \mathcal{T}_h\}. \tag{2.12}$$

Note that the functions in $\mathcal{M}_h$ are usually double-valued on the internal faces and that $P_M = P_{\mathcal{M}}$ whenever $k(K) = k$ for all simplexes K of the conforming mesh $\mathcal{T}_h$.

Theorem 2.1. *Suppose that the assumptions* **A1** *and* **A2** *on the stabilization function τ are satisfied. Then*

$$\|\boldsymbol{\epsilon}_h^{\mathbf{q}}\| \le C_{\mathbf{q}}\,(\|\mathbf{q} - \Pi_{\mathbf{V}}\mathbf{q}\| + \|Z_P\|_{\partial\mathcal{T}_h,\mathsf{w}}),$$

where $Z_P := (P_M - P_{\mathcal{M}})(\boldsymbol{Z}\cdot\boldsymbol{n} + Z)$, where $\boldsymbol{Z} := \mathbf{q} + (u - P_M u)\,\mathbf{v}$ and $Z := \tau\, u$. Moreover, if the elliptic regularity inequality (2.9) *holds, then*

$$\|\epsilon_h^u\| \le C_u\, h^{\min_{K\in\mathcal{T}_h}\{1,k(K)\}}\,\big(\|\mathbf{q} - \Pi_{\mathbf{V}}\mathbf{q}\| + |\mathbf{v}|_{W^{1,\infty}(\Omega)}\,\|\Pi_W u - u\| + \|Z_P\|_{\partial\mathcal{T}_h,\mathsf{w}}\big).$$

The constants $C_{\mathbf{q}}$ and C_u depend on c, on the constants c, κ, η, σ describing the properties of the triangulations (2.1), *and on $\ell_{\max} := \max\{\ell_K : K \in \mathcal{T}_h\}$ and $\ell_{\min} := \min\{\ell_K : K \in \mathcal{T}_h\}$. The constant C_u also depends on C_{R}, γ_0, $|\mathbf{v}|_{\boldsymbol{W}^{1,\infty}(\Omega)}$ and remains bounded when $\max_{K\in\mathcal{T}_h}\tau_K\, h_K$ remains bounded.*

Let us briefly discuss this result. First, note that, in the case of conforming meshes and uniform-degree ($P_M - P_{\mathcal{M}} = 0$) and for zero (or constant) convective velocities $\mathbf{v}$, we immediately recover the error estimates obtained in [11]; see also [15]. Thus, when we take the stabilization function τ in such a way that the errors of the projection converge optimally, for example, when τ and τ^{-1} are uniformly bounded (see [11, 15]), we get an order of convergence of $k+1$ for the error in the vector variable and an order of convergence of $k+1+\min\{k,1\}$ for the projection of the error in the scalar variable.

Furthermore, when we take nonconforming semimatching meshes and nonconstant convective velocities $\mathbf{v}$, the above orders of convergence remain the *same*. Indeed, we have (see the Appendix) that

$$\|Z_P\|_{\partial K,\mathsf{w}} \le C\, h_K^{k(K)+1}\,\mathcal{D}_K(\mathbf{q},u), \tag{2.13}$$

where $\mathcal{D}_K(\mathbf{q},u) := |\mathbf{q}|_{H^{k+1}(K)} + (h\,|\mathbf{v}|_{W^{1,\infty}(K)} + \|\tau\|_{L^\infty(\partial K)})\,|u|_{H^{k+1}(K)}$. In other words, the optimality of the convergence of the vector variable and the superconvergence of the projection of the error of the scalar variable *remain unchanged* with the use of the semimatching nonconforming meshes $\mathcal{T}_h$.

Finally, it is not difficult to get that, when $k(K) \ge 1$ for all $K \in \mathcal{T}_h$, we have

$$\|u - u_h^*\| \le \|\epsilon_h^u\| + C\,h\,(\|\mathbf{q} - \mathbf{q}_h\| + \inf_{\omega\in W_h^*}\|\nabla(u-\omega)\|)$$

where C only depends on the shape-regularity constant σ and

$$W_h^* := \{w \in L^2(\mathcal{T}_h) : \quad w|_K \in P_{k(K)+1}(K) \quad \forall\, K \in \mathcal{T}_h\}.$$

So, as a consequence of the above result, we can write that

$$\begin{aligned}\|u - u_h^*\| \le C\,h\,\big(&\|\mathbf{q} - \Pi_{\mathbf{V}}\mathbf{q}\| + |\mathbf{v}|_{W^{1,\infty}(\Omega)}\,\|\Pi_W u - u\| \\ &+ \|Z_P\|_{\partial\mathcal{T}_h,\mathsf{w}} + \inf_{\omega\in W_h^*}\|\nabla(u-\omega)\|\big),\end{aligned}$$

and so, in the uniform degree case, u_h^* converges to u with order $k+2$ for smooth enough solutions.

3. Main steps of the proof of Theorem 2.1

In this section, we give a sketch of the proof of Theorem 2.1 in order to make as clear as possible the main underlying ideas. The detailed proof of the key estimate is very technical and long, and is thus provided in the next section.

Step 1: Preliminaries. We begin by gathering several simple results already proven in [6].

Proposition 3.1 (The equations of the projection of the errors)**.** *Suppose that the assumption* **A2** *on the stabilization function* τ *is satisfied. Then we have*

$$(c\,\boldsymbol{\epsilon}_h^{\mathbf{q}}, \mathbf{r})_{\mathcal{T}_h} - (\epsilon_h^u, \nabla\cdot\mathbf{r})_{\mathcal{T}_h} + \langle \epsilon_h^{\widehat{u}}, \mathbf{r}\cdot\mathbf{n}\rangle_{\partial\mathcal{T}_h} = (c\,(\mathit{\Pi}_{\mathbf{V}}\mathbf{q} - \mathbf{q}), \mathbf{r})_{\mathcal{T}_h}, \tag{3.1a}$$

$$-(\boldsymbol{\epsilon}_h^{\mathbf{q}} + \epsilon_h^u\mathbf{v}, \nabla w)_{\mathcal{T}_h} + \langle(\boldsymbol{\epsilon}_h^{\mathbf{q}} + \epsilon_h^{\widehat{u}}\mathbf{v})\cdot\mathbf{n} + \tau\,(\epsilon_h^u - \epsilon_h^{\widehat{u}}), w\rangle_{\partial\mathcal{T}_h} = 0, \tag{3.1b}$$

$$\langle \mu, \epsilon_h^{\widehat{u}}\rangle_{\partial\Omega} = 0, \tag{3.1c}$$

$$\langle \mu, (\boldsymbol{\epsilon}_h^{\mathbf{q}} + \epsilon_h^{\widehat{u}}\,\mathbf{v})\cdot\mathbf{n} + \tau\,(\epsilon_h^u - \epsilon_h^{\widehat{u}})\rangle_{\partial\mathcal{T}_h\setminus\partial\Omega} = \langle \mu, Z_P\rangle_{\partial\mathcal{T}_h\setminus\partial\Omega}, \tag{3.1d}$$

for all $\mathbf{r}\in\mathbf{V}_h$*,* $w\in W_h$*, and* $\mu\in M_h$*. Here* Z_P *is defined in Theorem* 2.1*.*

Using the above equations, we obtain the following results by using standard energy and duality arguments.

Lemma 3.2 (The energy identity)**.** *Suppose that the assumption* **A2** *on the stabilization function* τ *is satisfied. Then we have*

$$(c\,\boldsymbol{\epsilon}_h^{\mathbf{q}}, \boldsymbol{\epsilon}_h^{\mathbf{q}})_{\mathcal{T}_h} + \|\sqrt{\gamma}\,(\epsilon_h^u - \epsilon_h^{\widehat{u}})\|^2_{\partial\mathcal{T}_h} = (c\,(\mathit{\Pi}_{\mathbf{V}}\mathbf{q} - \mathbf{q}), \boldsymbol{\epsilon}_h^{\mathbf{q}})_{\mathcal{T}_h} + \Lambda_P(\boldsymbol{Z}\cdot\boldsymbol{n} + Z).$$

Here $\gamma := \tau - \frac{1}{2}\mathbf{v}\cdot\mathbf{n}$ *and*

$$\Lambda_P(z) := \langle \epsilon_h^u - \epsilon_h^{\widehat{u}}, (P_M - P_{\mathcal{M}})z\rangle_{\partial\mathcal{T}_h} \quad \forall\, z\in L^2(\partial\mathcal{T}_h),$$

where $\boldsymbol{Z}$ *and* Z *are defined in Theorem* 2.1*.*

Lemma 3.3 (The duality argument)**.** *Suppose that the assumption* **A2** *on the stabilization function* τ *is satisfied. Then, for any function* Θ *in* $L^2(\Omega)$*, we have*

$$(\epsilon_h^u, \Theta)_{\mathcal{T}_h} = T_{\mathbf{q}} + T_{\mathbf{v}} + T_P^1 + T_P^2 + T_P^3,$$

where

$$\begin{aligned}
T_{\mathbf{q}} &:= (c(\mathbf{q}-\mathbf{q}_h), \mathit{\Pi}_{\mathbf{V}}^*\boldsymbol{\Phi} - \boldsymbol{\Phi})_{\mathcal{T}_h} + (\mathbf{q} - \mathit{\Pi}_{\mathbf{V}}\mathbf{q}, \nabla(\Psi_h - \Psi))_{\mathcal{T}_h},\\
T_{\mathbf{v}} &:= (\boldsymbol{\delta}\mathbf{v}(u - \mathit{\Pi}_W u), \nabla\Psi_h),\\
T_P^1 &:= \langle (P_{\mathcal{M}} - P_M)\Psi, Z_P\rangle_{\partial\mathcal{T}_h},\\
T_P^2 &:= \Lambda_P(\zeta),\\
T_P^3 &:= \Lambda_P(\boldsymbol{\zeta}\cdot\boldsymbol{n}),
\end{aligned}$$

for any $\Psi_h\in W_h$*. Here* $\boldsymbol{\zeta} := \boldsymbol{\Phi}$ *and* $\zeta := (\tau - \mathbf{v}\cdot\mathbf{n})(\Psi - \mathit{\Pi}_W^*\Psi)$*.*

We have introduced the auxiliary function $\boldsymbol{\delta}\mathbf{v} := \mathbf{v} - \mathbf{v}_0$, where $\mathbf{v}\in\boldsymbol{W}^{1,\infty}(\Omega)$ is the divergence-free convective velocity. The function $\mathbf{v}_0$ is defined on the simplex $K\in\mathcal{T}_h$ as the element of $\mathbf{P}_0(K)$ such that

$$\langle(\mathbf{v} - \mathbf{v}_0)\cdot\mathbf{n}, 1\rangle_F = 0, \tag{3.2}$$

for all faces F of K. Note that $\mathbf{v}_0$ is well defined because $\mathbf{v}$ is divergence-free.

Step 2: Estimate of the functional $\Lambda(z)$. It is important to emphasize that the *only* difference between the analysis we are carrying out here and that of the HDG method for general nonconforming meshes done in Part I, resides in the manner we estimate the terms involving the linear functional Λ_P, namely, the terms $\Lambda_P(\boldsymbol{Z}\cdot\boldsymbol{n}+Z)$, $T_P^2=\Lambda_P(\zeta)$ and $T_P^3=\Lambda_P(\boldsymbol{\zeta}\cdot\boldsymbol{n})$.

In fact, in [6], we estimated $\Lambda_P(z)$ by simply applying a weighted Cauchy-Schwarz inequality:

$$|\,\Lambda_P(z)\,|\leq \|r\,(\epsilon_h^u-\epsilon_h^{\widehat{u}})\|_{\partial\mathcal{T}_h}\,\|r^{-1}\,(P_M-P_{\mathcal{M}})z\|_{\partial\mathcal{T}_h},$$

where the weight function r was taken to be equal to $\sqrt{\gamma}$ for the estimate of the term $\Lambda_P(\boldsymbol{Z}\cdot\boldsymbol{n}+Z)$, and equal to one for the terms $T_P^2=\Lambda_P(\zeta)$ and $T_P^3=\Lambda_P(\boldsymbol{\zeta}\cdot\boldsymbol{n})$. However, by taking advantage of the structure of the semimatching meshes and of the definition of the polynomial degrees on the elements having hanging nodes, we can improve the above estimate and can eliminate the potentially ensuing loss of convergence for general nonconforming meshes uncovered in [6]. This is the main contribution of this paper. The new estimate is contained in the following result.

Lemma 3.4. *For all $z\in L^2(\partial\mathcal{T}_h)$, we have*

$$|\,\Lambda_P(z)\,|\leq C\,\|c\|_{L^\infty(\Omega)}\,C_\Lambda\,\|\,\mathbf{q}-\mathbf{q}_h\,\|_\Omega\,\|\,(P_M-P_{\mathcal{M}})z\|_{\partial\mathcal{T}_h,\mathsf{w}}.$$

Here C is a constant depending on the shape-regularity constant σ, and

$$C_\Lambda^2:=\left(\frac{\kappa\,\mathsf{c}}{1-\eta}+1\right)\sigma\,(\ell_{\max}-\ell_{\min}+1)\,(d+1),$$

where c, κ, η, σ are the constants describing the properties of the triangulations (2.1),

$$\ell_{\max}:=\max\{\ell_K:K\in\mathcal{T}_h\}\quad\textit{and}\quad\ell_{\min}:=\min\{\ell_K:K\in\mathcal{T}_h\}.$$

The next section is devoted to a detailed proof of this result. Here, let us just note that this result states, roughly speaking, that we can take the weighting function r to be $\mathsf{w}^{-1/2}\sim h^{-1/2}$ which is what prevents the loss of convergence properties. Indeed, in the general case treated in [6], the weighting functions r could only be taken to be of order one.

Step 3: Estimate of $\boldsymbol{\epsilon}_h^{\mathbf{q}}$. By using the energy identity of Lemma 3.2 and the above result, we can easily obtain the estimate for $\boldsymbol{\epsilon}_h^{\mathbf{q}}$ in Theorem 2.1.

Lemma 3.5. *Suppose that the assumptions* **A1** *and* **A2** *on the stabilization function τ are satisfied. Then, we have*

$$\|\boldsymbol{\epsilon}_h^{\mathbf{q}}\|_\Omega\leq\sqrt{3}\,\|c\|_{L^\infty(\Omega)}\,\|c^{-1}\|_{L^\infty(\Omega)}\,(\|\mathbf{q}-\mathit{\Pi}_{\mathbf{V}}\mathbf{q}\|_\Omega+C\,C_\Lambda\,\|Z_P\|_{\partial\mathcal{T}_h,\mathsf{w}}),$$

where the constant C depends only on the shape-regularity constant σ.

Proof. If we apply the estimate of Lemma 3.4 to the right-hand side of the energy identity of Lemma 3.2, we obtain

$$\begin{aligned}(c\,\boldsymbol{\epsilon}_h^{\mathbf{q}},\boldsymbol{\epsilon}_h^{\mathbf{q}})_{\mathcal{T}_h}+\|\sqrt{\gamma}\,(\epsilon_h^u-\epsilon_h^{\widehat{u}})\|_{\partial\mathcal{T}_h}^2&\leq(c\,(\mathit{\Pi}_{\mathbf{V}}\mathbf{q}-\mathbf{q}),\boldsymbol{\epsilon}_h^{\mathbf{q}})_{\mathcal{T}_h}\\&\quad+C\,\|c\|_{L^\infty(\Omega)}\,C_\Lambda\,\|\mathbf{q}-\mathbf{q}_h\,\|_\Omega\,\|\,Z_P\,\|_{\partial\mathcal{T}_h,\mathsf{w}}.\end{aligned}$$

The estimate now follows after a few simple manipulations. ☐

Step 4: A first estimate of ϵ_h^u. Next, we begin to prove the estimate for the projection of the error of the scalar variable, ϵ_h^u.

A first estimate of ϵ_h^u is the following.

Lemma 3.6. *Suppose that the assumptions* **A1** *and* **A2** *on the stabilization function τ are satisfied. Then, we have*

$$\|\epsilon_h^u\|_\Omega \le C_1\,\|\mathbf{q} - \Pi_{\mathbf{V}}\mathbf{q}\|_\Omega + C_2\,\|u - \Pi_W u\|_\Omega + C_3\,\|Z_P\|_{\partial\mathcal{T}_h,\mathsf{w}}.$$

Here

$$\begin{aligned}
C_1 &:= C\,(H_{\mathbf{q}} + C_\Lambda\,H_P^2 + C_\Lambda\,H_P^3),\\
C_2 &:= H_{\mathbf{v}},\\
C_3 &:= H_P^1 + C\,C_\Lambda\,(H_{\mathbf{q}} + C_\Lambda\,H_P^2 + C_\Lambda\,H_P^3),
\end{aligned}$$

where the constant C depends only on the shape-regularity constant σ and

$$\begin{aligned}
H_{\mathbf{q}} &:= \max\left\{\|c\|_{L^\infty(\Omega)} \sup_{\Theta\in L^2(\Omega)\setminus\{0\}} \frac{\|\Pi_{\mathbf{V}}^*\mathbf{\Phi} - \mathbf{\Phi}\|_\Omega}{\|\Theta\|_\Omega}, \sup_{\Theta\in L^2(\Omega)\setminus\{0\}} \frac{\|\nabla(\Psi_h - \Psi)\|_{\mathcal{T}_h}}{\|\Theta\|_\Omega}\right\},\\
H_{\mathbf{v}} &:= \|\delta\mathbf{v}\|_{L^\infty(\Omega)} \sup_{\Theta\in L^2(\Omega)\setminus\{0\}} \frac{\|\nabla\Psi_h\|_{\mathcal{T}_h}}{\|\Theta\|_\Omega},\\
H_P^1 &:= \sup_{\Theta\in L^2(\Omega)\setminus\{0\}} \frac{\|(P_{\mathcal{M}} - P_M)\Psi\|_{\partial\mathcal{T}_h,\mathsf{w}^{-1}}}{\|\Theta\|_\Omega},\\
H_P^2 &:= \|c\|_{L^\infty(\Omega)} \sup_{\Theta\in L^2(\Omega)\setminus\{0\}} \frac{\|(P_M - P_{\mathcal{M}})\zeta\|_{\partial\mathcal{T}_h,\mathsf{w}}}{\|\Theta\|_\Omega},\\
H_P^3 &:= \|c\|_{L^\infty(\Omega)} \sup_{\Theta\in L^2(\Omega)\setminus\{0\}} \frac{\|(P_M - P_{\mathcal{M}})\boldsymbol{\zeta}\cdot\mathbf{n}\|_{\partial\mathcal{T}_h,\mathsf{w}}}{\|\Theta\|_\Omega},
\end{aligned}$$

for any $\Psi_h \in W_h$.

Proof. To prove this result, we only have to estimate each of the five terms of the right-hand side of the identity of Lemma 3.3. Thus, after applying the Cauchy-Schwarz inequality and Lemma 3.4, it is not difficult to see that we get

$$\begin{aligned}
|T_{\mathbf{q}}| &\le H_{\mathbf{q}}\,\|\Theta\|_\Omega\,(\|\mathbf{q} - \mathbf{q}_h\|_\Omega + \|\mathbf{q} - \Pi_{\mathbf{V}}\mathbf{q}\|_\Omega),\\
|T_{\mathbf{v}}| &\le H_{\mathbf{v}}\,\|\Theta\|_\Omega\,\|u - \Pi_W u\|_\Omega,\\
|T_P^1| &\le H_P^1\,\|\Theta\|_\Omega\,\|Z_P\|_{\partial\mathcal{T}_h,\mathsf{w}},\\
|T_P^2| &\le C_\Lambda\,H_P^2\,\|\Theta\|_\Omega\,\|\mathbf{q} - \mathbf{q}_h\|_\Omega,\\
|T_P^3| &\le C_\Lambda\,H_P^3\,\|\Theta\|_\Omega\,\|\mathbf{q} - \mathbf{q}_h\|_\Omega.
\end{aligned}$$

We then obtain that

$$\begin{aligned}
\|\epsilon_h^u\| \le\ & H_{\mathbf{q}}\,(\|\mathbf{q} - \mathbf{q}_h\|_\Omega + \|\mathbf{q} - \Pi_{\mathbf{V}}\mathbf{q}\|_\Omega) + H_{\mathbf{v}}\,\|u - \Pi_W u\|_\Omega\\
&+ H_P^1\,\|Z_P\|_{\partial\mathcal{T}_h,\mathsf{w}} + C_\Lambda(H_P^2 + H_P^3)\,\|\mathbf{q} - \mathbf{q}_h\|_\Omega,
\end{aligned}$$

and the estimate easily follows after using the first estimate of Theorem 2.1. The proof is complete. $\square$

Step 5: Estimate of the quantities H. Next, we estimate the auxiliary quantities $H_{\mathbf{q}}$, $H_{\mathbf{v}}$ H_P^1, H_P^2, and H_P^3. To describe the estimates, we need to introduce the constants

$$\Upsilon_K^{\max} := \max\{\min\left(\tau - \frac{1}{2}\mathbf{v}\cdot\mathbf{n}\right)|_F : F \text{ is a face of } K\}$$
$$(\tau\mathbf{v})_K^{\max} := \max\big|(\tau - \mathbf{v}\cdot\mathbf{n})|_{\partial K}\big|, \quad (\tau\mathbf{v})_K^{*} := \max\big|(\tau - \mathbf{v}\cdot\mathbf{n})|_{\partial K\setminus F^*}\big|,$$

where F^* is a face of K at which $|(\tau - \mathbf{v}\cdot\mathbf{n})|_{\partial K}|$ attains its maximum. We also need the quantities $C^{i,*}_{\mathbf{V},K}$:

$$C^{1,*}_{\mathbf{V},K} := 1 + \frac{h|\mathbf{v}|_{W^{1,\infty}(K)}}{\Upsilon_K^{\max}},$$
$$C^{2,*}_{\mathbf{V},K} := (\tau\mathbf{v})_K^{*} + \frac{h|\mathbf{v}|_{W^{1,\infty}(K)}((\tau\mathbf{v})_K^{\max} + |\mathbf{v}|_{W^{1,\infty}(K)}\, h)}{\Upsilon_K^{\max}}.$$

Finally, we are also going to need the set

$$\mathcal{T}_{P,h} := \{K \in \mathcal{T}_h : (P_{\mathcal{M}} - P_M) \neq 0 \text{ on some face } F \text{ of } K\},$$

which is nothing but the set of elements $K \in \mathcal{T}_h$ that either have a hanging node in one of their faces or have a neighboring element with higher polynomial degree.

Lemma 3.7. *Suppose that the assumptions* **A1** *and* **A2** *on the stabilization function τ are satisfied. Suppose also that the elliptic regularity inequality* (2.9) *holds. Then we have that*

$$H_{\mathbf{q}} \le C\,\|c\|_{L^\infty(\Omega)}(1 + \max_{K\in\mathcal{T}_h}(C^{1,*}_{\mathbf{V},K} + C^{2,*}_{\mathbf{V},K}\, h_K))\, C_{\mathrm{R}} \max_{K\in\mathcal{T}_h} h_K^{\min\{1,k(K)\}},$$
$$H_{\mathbf{v}} \le C\,|\mathbf{v}|_{W^{1,\infty}(\Omega)}\, h,$$
$$H_P^1 \le C\, C_{\mathrm{R}} \max_{K\in\mathcal{T}_{P,h}} h_K^{\min\{1,k(K)\}},$$
$$H_P^2 \le C\, C_{\mathrm{R}} \max_{K\in\mathcal{T}_{P,h}} \tau_K\, h_K^{1+\min\{1,k(K)\}},$$
$$H_P^3 \le C\, C_{\mathrm{R}} \max_{K\in\mathcal{T}_{P,h}} h_K,$$

where the constant C depends only on the shape-regularity constant σ.

The proof of this lemma is essentially the same as in [6] (Proposition 2.3) and thus omitted.

Step 6: The final estimate of ϵ_h^u. To obtain the estimate of ϵ_h^u of Theorem 2.1, we only need to insert the estimates of Lemma 3.7 into the estimate of Lemma 3.6. This concludes the proof of Theorem 2.1.

4. Proof of the key estimate, Lemma 3.4

In this section, we provide a detailed proof of Lemma 3.4 which contains the *key* estimate of the linear functional $\Lambda_P(z)$. To do that, we proceed in several steps. In the first two steps, we express $\Lambda_P(z)$ solely in terms of z and $\epsilon_h^{\widehat{u}}$. In the third step, we exploit the structure of the semimatching meshes and the definition of the polynomial degrees on the elements having hanging nodes, to express the term involving $\epsilon_h^{\widehat{u}}$ in terms of suitably defined integrals of $\epsilon_h^{\widehat{u}}$. In the fourth step, we establish a relation between those linear functionals of $\epsilon_h^{\widehat{u}}$ and the error in the flux $\mathbf{q}-\mathbf{q}_h$. We do that by using the first error equation, namely,

$$\langle \epsilon_h^{\widehat{u}}, \mathbf{r}\cdot\mathbf{n}\rangle_{\partial\mathcal{T}_h} = (c\,(\mathbf{q}-\mathbf{q}_h), \mathbf{r})_{\mathcal{T}_h} + (\epsilon_h^u, \nabla\cdot\mathbf{r})_{\mathcal{T}_h} \qquad \forall \mathbf{r}\in\mathbf{V}_h,$$

and construct special *divergence-free* test functions $\mathbf{r}$ in order to *eliminate* the projection of the error ϵ_h^u. In the last step, we conclude.

Step 1: A first estimate of $\Lambda_P(z)$. We start by estimating $\Lambda_P(z)$ solely in terms of z and $\epsilon_h^{\widehat{u}}$.

Lemma 4.1. *For all $z\in L^2(\partial\mathcal{T}_h)$, we have that*

$$|\Lambda_P(z)| \le \|(P_M - P_{\mathcal{M}})\epsilon_h^{\widehat{u}}\|_{\partial\mathcal{T}_h,\mathsf{w}^{-1}}\, \|(P_M - P_{\mathcal{M}})z\|_{\partial\mathcal{T}_h,\mathsf{w}}.$$

Proof. We have

$$\begin{aligned}
\Lambda_P(z) &= \langle \epsilon_h^u - \epsilon_h^{\widehat{u}}, (P_M - P_{\mathcal{M}})z\rangle_{\partial\mathcal{T}_h}\\
&= \sum_{K\in\mathcal{T}_h}\sum_{F\in\mathcal{F}(K)} \langle \epsilon_h^u - \epsilon_h^{\widehat{u}}, (P_M - P_{\mathcal{M}})z\rangle_F\\
&= -\sum_{K\in\mathcal{T}_h}\sum_{F\in\mathcal{F}(K)} \langle (P_M - P_{\mathcal{M}})\epsilon_h^{\widehat{u}}, (P_M - P_{\mathcal{M}})z\rangle_F,
\end{aligned}$$

since $(P_M - P_{\mathcal{M}})\epsilon_h^u = 0$ on F. This implies that

$$\begin{aligned}
|\Lambda_P(z)| &\le \sum_{K\in\mathcal{T}_h}\sum_{F\in\mathcal{F}(K)} |\langle (P_M - P_{\mathcal{M}})\epsilon_h^{\widehat{u}}, (P_M - P_{\mathcal{M}})z\rangle_F|\\
&\le \sum_{K\in\mathcal{T}_h}\sum_{F\in\mathcal{F}(K)} \|(P_M - P_{\mathcal{M}})\epsilon_h^{\widehat{u}}\|_{F,\mathsf{w}^{-1}}\, \|(P_M - P_{\mathcal{M}})z\|_{F,\mathsf{w}},
\end{aligned}$$

and the result follows. This completes the proof. ∎

In an effort to eventually bound $\|(P_M - P_{\mathcal{M}})\epsilon_h^{\widehat{u}}\|_{\partial\mathcal{T}_h,\mathsf{w}^{-1}}$, in the next three steps, we focus on the study of the term $\|(P_M - P_{\mathcal{M}})\epsilon_h^{\widehat{u}}\|_{F,\mathsf{w}^{-1}}$ for $F\in\mathcal{F}(K)$ with a particular $K\in\mathcal{T}_h$. For simplicity, we slightly abuse the notation and use $P_M - P_{\mathcal{M}}$ to denote $(P_M - P_{\mathcal{M}})\,|_F$.

Step 2: An expression for $\|(P_M - P_{\mathcal{M}})\epsilon_h^{\widehat{u}}\|^2_{F,\mathsf{w}^{-1}}$. Now, let us rewrite the expression involving the error in the numerical trace $\epsilon_h^{\widehat{u}}$ in a more suitable manner.

Lemma 4.2. *For $F\in\mathcal{F}(K)$ where $K\in\mathcal{T}_h$, we have*

$$\|(P_M - P_{\mathcal{M}})\epsilon_h^{\widehat{u}}\|^2_{F,\mathsf{w}^{-1}} = \mathsf{w}_F^{-1}\,\langle \epsilon_h^{\widehat{u}}, (P_M - P_{\mathcal{M}})\epsilon_h^{\widehat{u}}\rangle_{F\setminus\partial\Omega}.$$

Proof. To prove this result, let us begin by recalling that $P_{\mathcal{M}}$ is the L^2-projection into the space of polynomials of degree $k(K)$ defined on F, $P_{k(K)}(F)$; see (2.12). Recall also that the restriction of P_M on F is the L^2-projection into the space $\{\mu|_F : \mu\in M_h\}$. Thus, we have that

$$\begin{aligned}
\|(P_M - P_{\mathcal{M}})\epsilon_h^{\widehat{u}}\|^2_{F,\mathsf{w}^{-1}} &= \langle \mathsf{w}^{-1}\,(P_M - P_{\mathcal{M}})\epsilon_h^{\widehat{u}}, (P_M - P_{\mathcal{M}})\epsilon_h^{\widehat{u}}\rangle_F\\
&= \mathsf{w}_F^{-1}\,\langle (P_M - P_{\mathcal{M}})\epsilon_h^{\widehat{u}}, (P_M - P_{\mathcal{M}})\epsilon_h^{\widehat{u}}\rangle_F
\end{aligned}$$

by definition of the weight function w, (2.11). Now, by definition of the projections P_M and $P_{\mathcal{M}}$, we get

$$\begin{aligned}\|(P_M - P_{\mathcal{M}})\epsilon_h^{\widehat{u}}\|^2_{F,\mathsf{w}^{-1}} &= \mathsf{w}_F^{-1}\,\langle \epsilon_h^{\widehat{u}}, (P_M - P_{\mathcal{M}})^2 \epsilon_h^{\widehat{u}}\rangle_F \\ &= \mathsf{w}_F^{-1}\,\langle \epsilon_h^{\widehat{u}}, (P_M - P_{\mathcal{M}}) \epsilon_h^{\widehat{u}}\rangle_{F\setminus\partial\Omega},\end{aligned}$$

since $(P_M - P_{\mathcal{M}})^2 = (P_M - P_{\mathcal{M}})$ because $\{\mu|_F : \mu \in M_h\} \supset P_{k(K)}(F)$ by (2.3e), and since $(P_M - P_{\mathcal{M}}) = 0$ when $F \subset \partial\Omega$ because in that case $\{\mu|_F : \mu \in M_h\} = P_{k(K)}(F)$ by (2.3d). This completes the proof. $\square$

Step 3: A representation of $\langle \epsilon_h^{\widehat{u}}, (P_M - P_{\mathcal{M}})\epsilon_h^{\widehat{u}}\rangle_{F\setminus\partial\Omega}$. We see that we have to estimate the term $|\langle \epsilon_h^{\widehat{u}}, \delta\rangle_{F\setminus\partial\Omega}|$ for δ of the form of $(P_M - P_{\mathcal{M}})\epsilon_h^{\widehat{u}}$. The idea is to use the first error equation (3.1a) to relate $\langle \epsilon_h^{\widehat{u}}, \delta\rangle_{F\setminus\partial\Omega}$ with the error in $\boldsymbol{q}$. For this purpose, we need an admissible $\boldsymbol{r}$ that is divergence-free and coincides with δ on F. This is a highly nontrivial task due to the nonconformity of the meshes. Next, we are going to rewrite the quantity $\langle \epsilon_h^{\widehat{u}}, (P_M - P_{\mathcal{M}})\epsilon_h^{\widehat{u}}\rangle_{F\setminus\partial\Omega}$ to facilitate this process. From now on, until the last step of this proof, we *fix* the simplex $K \in \mathcal{T}_h$ and its face $F \in \mathcal{F}(K) \setminus \partial\Omega$, and set $\ell := \ell_K$.

To state this rewriting, we need to introduce some notation. We first define the set of elements adjacent to K through F, $\mathcal{T}_h(F,K)$, the set of their ancestors (including themselves), and the set of ancestors of ancestors (excluding themselves). To be exact, we set

$$\mathcal{T}_h(F,K) := \{K' \in \mathcal{T}_h : K \cap K' = \emptyset, F \cap \partial K' \neq \emptyset\}, \tag{4.1a}$$

$$\mathbb{T}_h^m(F,K) := \{\mathsf{K}^m_{K'} : K' \in \mathcal{T}_h(F,K)\}, \tag{4.1b}$$

$$\widetilde{\mathbb{T}}_h^{m-1}(F,K) := \{\mathsf{K}^{m-1}_{\mathbb{K}} : \mathbb{K} \in \mathbb{T}_h^m(F,K)\}. \tag{4.1c}$$

We also define the set of all elements, on level m of the conforming family of meshes, that are adjacent to K through F.

$$\mathsf{T}_h^m(F,K) := \{\mathsf{K} \in \mathsf{T}_h^m : K \cap \mathsf{K} = \emptyset, F \cap \partial\mathsf{K} \neq \emptyset\}, \tag{4.1d}$$

Note that there is only one simplex in $\mathsf{T}_h^\ell(F,K)$ which we are going to denote by K_F. The above sets are illustrated in Figure 4.

Finally, we set $\delta\ell_F := \max\{\ell_{K'} :\ K' \in \mathcal{T}_h(F,K)\} - \ell$. We know that $\mathcal{T}_h(F,K)$ is not the empty set since $F \in \mathcal{F}(K) \setminus \partial\Omega$, as noted at the beginning of this step. If $\delta\ell_F \geq 0$, for any given function $\delta \in L^2(F)$, we define the function $\mathrm{P}^m\delta$ on F as follows. For every $\mathbb{K} \in \mathbb{T}_h^m(F,K)$, we set $\mathrm{P}^m\delta$ on $F \cap \partial\mathbb{K}$ to be the average of δ therein; elsewhere, we set $\mathrm{P}^m(\delta)$ to be zero. That is,

$$\mathrm{P}^m\delta := \begin{cases} \frac{1}{|F\cap\partial\mathbb{K}|}\int_{F\cap}\partial\mathbb{K}\delta & \text{in } F\cap\partial\mathbb{K}, \text{when } \mathbb{K} \in \mathbb{T}_h^m(F,K), \\ 0 & \text{elsewhere in } F. \end{cases}$$

With this notation, we have the following representation result. It decomposes $(P_M - P_{\mathcal{M}})\epsilon_h^{\widehat{u}}$ into a sequence of piecewise constant functions on F and one high-order residual term in such a way that it is possible to find the test function $\boldsymbol{r}$ with the properties stated at the beginning of this step.

Lemma 4.3. *We have, for $\delta = (P_M - P_{\mathcal{M}})\widetilde{\delta}$ with $\widetilde{\delta} \in L^2(F)$, that*

$$\langle \epsilon_h^{\widehat{u}}, \delta\rangle_F = \begin{cases} 0 & \textit{if } \delta\ell_F < 0, \\ \sum_{m=\ell+1}^{\ell+\delta\ell_F} \langle \epsilon_h^{\widehat{u}}, \delta^{0,m}\rangle_F + \langle \epsilon_h^{\widehat{u}}, \delta^{\perp}\rangle_F, & \textit{otherwise}, \end{cases}$$

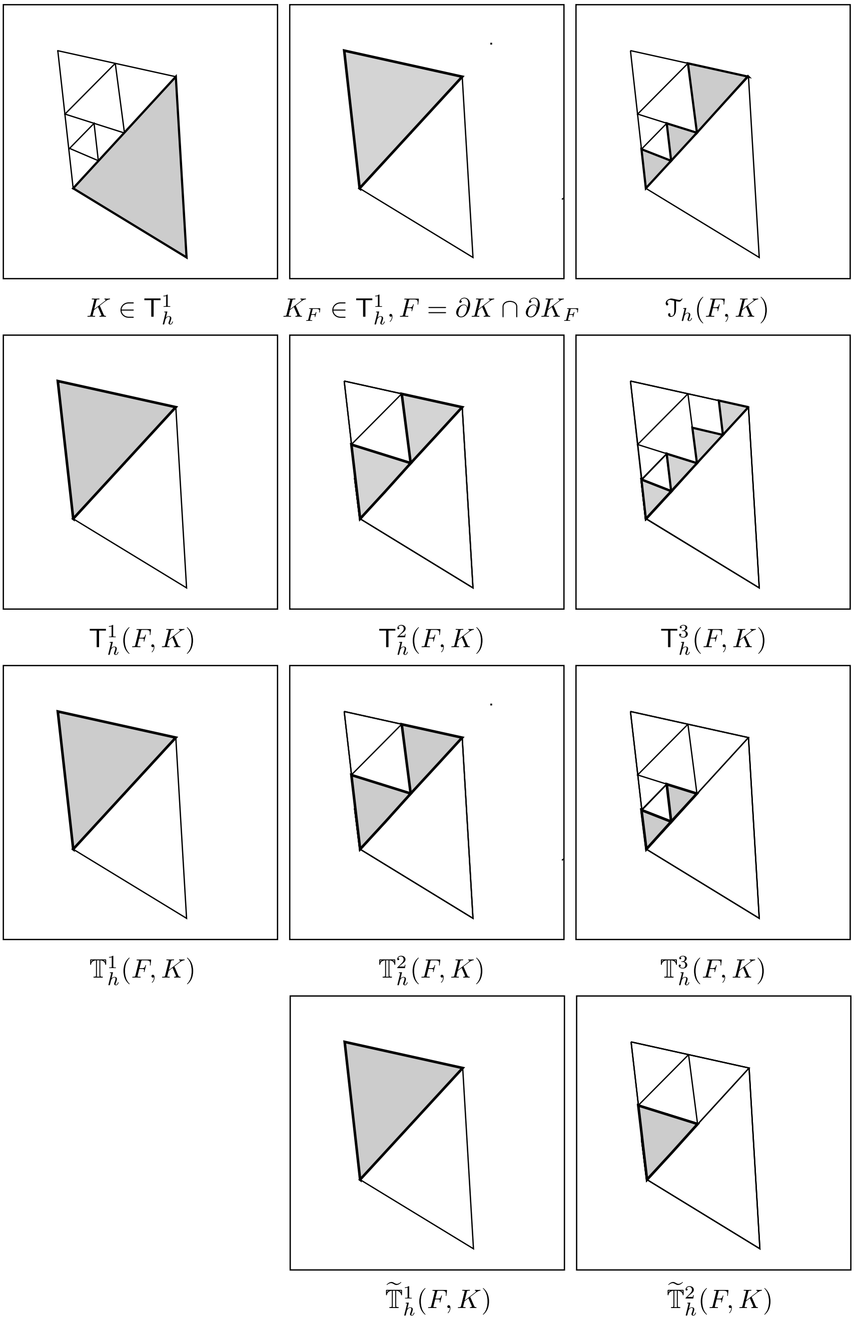


FIGURE 4. Example of the sets $\mathcal{T}_h(F, K)$,$\mathsf{T}_h^m(F, K)$, $\mathbb{T}_h^m(F, K)$ and $\widetilde{\mathbb{T}}_h^{m-1}(F, K)$ (in gray).

where

$$
\begin{aligned}
\delta^{0,\ell} &:= \mathrm{P}^\ell \, \delta, \\
\delta^{0,m} &:= \mathrm{P}^m \, (\mathrm{Id} - \mathrm{P}^{m-1}) \dots (\mathrm{Id} - \mathrm{P}^\ell)\delta, \quad m = \ell + 1, \dots, \ell + \delta\ell_F, \\
\delta^\perp &:= \delta^{\perp, \ell + \delta\ell_F} \text{ with } \delta^{\perp, m} := (\mathrm{Id} - \mathrm{P}^m)(\mathrm{Id} - \mathrm{P}^{m-1}) \dots (\mathrm{Id} - \mathrm{P}^\ell)\delta.
\end{aligned}
$$

Moreover, on $\widetilde{\mathbb{F}} := F \cap \partial\widetilde{\mathbb{K}}$, *where* $\widetilde{\mathbb{K}} \in \widetilde{\mathbb{T}}_h^{m-1}(F, K)$, *we have that*

$$\delta^{0,m}|_{\widetilde{\mathbb{F}}\cap\partial\mathsf{K}} \in P_0(\widetilde{\mathbb{F}} \cap \partial\mathsf{K}) \quad \forall\, \mathsf{K} \in \mathsf{T}_h^m(F, K) : \ \mathsf{K} \subset \widetilde{\mathbb{K}}, \tag{4.2a}$$

$$\langle \delta^{0,m}, 1\rangle_{\widetilde{\mathbb{F}}} = 0. \tag{4.2b}$$

Elsewhere, $\delta^{0,m} = 0$. *Also, for any* $K' \in \mathcal{T}_h(F, K)$, *we have, with* $F' := F \cap \partial K'$, *that*

$$\delta^{\perp}|_{F'} \in P_{k(K')}(F'), \tag{4.3a}$$

$$\langle \delta^{\perp}, 1\rangle_{F'} = 0. \tag{4.3b}$$

Finally, we have that

$$\|\delta\|_F^2 = \sum_{m=\ell+1}^{\ell+\delta\ell_F} \|\delta^{0,m}\|_F^2 + \|\delta^{\perp}\|_F^2.$$

Proof. Let us begin by showing that, if $\delta\ell_F < 0$, then $\delta = 0$ on F. We know there is a simplex K' in $\mathcal{T}_h \cap \mathsf{T}_h^{\ell+\delta\ell_F}$ such that $F = \partial K \cap \partial K'$ by the construction of the mesh $\mathcal{T}_h$ (2.1). By definition of the space M_h (2.3c), we have that, by (2.3f), $k(K) \geq k(K')$ because $\ell \geq \ell + \delta\ell_F$. This implies that $P_M - P_{\mathcal{M}} = 0$ on F and so, that $\delta = 0$ on F.

If $\delta\ell_F \geq 0$, we proceed as follows. For any function $\delta \in L^2(F)$, we can write

$$\begin{aligned}
\delta &= \mathrm{P}^{\ell}\delta + (\mathrm{Id} - \mathrm{P}^{\ell})\delta \\
&= \delta^{0,\ell} + \delta^{\perp,\ell} \\
&= \delta^{0,\ell} + \mathrm{P}^{\ell+1}\delta^{\perp,\ell} + (\mathrm{Id} - \mathrm{P}^{\ell+1})\delta^{\perp,\ell} \\
&= \delta^{0,\ell} + \delta^{0,\ell+1} + \delta^{\perp,\ell+1} \\
&= \sum_{m=\ell}^{\ell+\delta\ell_F} \delta^{0,m} + \delta^{\perp}.
\end{aligned}$$

This $L^2(F)$-orthogonal decomposition immediately leads to:

$$\|\delta\|_F^2 = \sum_{m=\ell}^{\ell+\delta\ell_F} \|\delta^{0,m}\|_F^2 + \|\delta^{\perp}\|_F^2.$$

It remains to show that $\delta^{0,\ell} = 0$ in order to prove the final identity of the lemma. But, by the definition of the projections P_M and $P_{\mathcal{M}}$, we have that

$$\delta^{0,\ell} = \frac{1}{|F|}\langle (P_M - P_{\mathcal{M}})\epsilon_h^{\widehat{u}}, 1\rangle_F = \frac{1}{|F|}\langle \epsilon_h^{\widehat{u}}, (P_M - P_{\mathcal{M}})1\rangle_F = 0.$$

Let us now prove the properties of $\delta^{0,m}$ for $m > \ell$. By the definition of P^m, we only have to show the integral identity. So

$$\langle \delta^{0,m}, 1\rangle_{\widetilde{\mathbb{F}}} = \langle \delta^{\perp,m-2}, (\mathrm{Id} - \mathrm{P}^{m-1})\mathrm{P}^m 1\rangle_{\widetilde{\mathbb{F}}} = \langle \delta^{\perp,m-2}, (\mathrm{Id} - \mathrm{P}^{m-1})1\rangle_{\widetilde{\mathbb{F}}} = 0.$$

Note that $\delta^{\perp,m-2}$ is well defined since $m > \ell$.

It remains to prove the properties of $\delta^{\perp}$. Pick any $K' \in \mathcal{T}_h(F, K)$. Then $K' \in \mathsf{T}_h^m(F, K)$ and this implies that $\mathsf{P}^n = 0$ on F' for $n > m$. As a consequence, we have that $\delta^{\perp} = (\mathrm{Id} - \mathrm{P}^m)\cdots(\mathrm{Id} - \mathrm{P}^{\ell})\delta$ and the result follows.

This completes the proof. $\square$

Step 4: Estimate of $\langle \epsilon_h^{\widehat{u}}, \delta^{0,m}\rangle_F$ and $\langle \epsilon_h^{\widehat{u}}, \delta^{\perp}\rangle_F$. As suggested by the representation result we just proved, we next obtain estimates of $\langle \epsilon_h^{\widehat{u}}, \delta^{0,m}\rangle_F$ and $\langle \epsilon_h^{\widehat{u}}, \delta^{\perp}\rangle_F$.

This crucial result is given by the next lemma.

Lemma 4.4. *We have that*

$$\begin{aligned} |\,\langle \epsilon_h^{\widehat{u}}, \delta^{0,m}\rangle_F\,| &\le C\,\|c\|_{L^\infty(\Omega)}\,(\kappa\,\mathsf{c}\,\eta^{m-1-\ell}\,h_{K_F})^{1/2}\,\|\mathbf{q}_h-\mathbf{q}\|_{K_F}\,\|\delta^{0,m}\|_F,\\ |\,\langle \epsilon_h^{\widehat{u}}, \delta^{\perp}\rangle_F\,| &\le C\,\|c\|_{L^\infty(\Omega)}\,h_{K_F}^{1/2}\|\mathbf{q}_h-\mathbf{q}\|_{K_F}\,\|\delta^{\perp}\|_F. \end{aligned}$$

The constant C depends only on the shape-regularity constant σ.

In order to facilitate the reading of rather technical proof, we postpone the proof of this result. Thus in Step 6, we provide a detailed proof of the first estimate, and in Step 7, a proof of the second estimate.

Step 5: Conclusion. We are now ready to conclude the proof of Lemma 3.4. We begin by noting that, as stated at the beginning of Step 3, the function δ in the analysis carried out in Steps 3 and 4 is of the form $(P_M - P_{\mathcal{M}})\epsilon_h^{\widehat{u}}$. Thus, by Lemma 4.2 and by the representation Lemma 4.3, we have

$$\|\delta\|_F^2 = |\,\langle \epsilon_h^{\widehat{u}}, \delta\rangle_F\,| \le \left(\sum_{m=\ell+1}^{\ell+\delta\ell_F} |\,\langle \epsilon_h^{\widehat{u}}, \delta^{0,m}\rangle_F\,| + |\,\langle \epsilon_h^{\widehat{u}}, \delta^{\perp}\rangle_F\,|\right).$$

If we now use the estimates of Lemma 4.4, we get that

$$\|\delta\|_F^2 \le C\,\|c\|_{L^\infty(\Omega)}\left(\sum_{m=\ell+1}^{\ell+\delta\ell_F}(\kappa\,\mathsf{c}\,\eta^{m-1-\ell})^{1/2}\,\|\delta^{m,0}\|_F + \|\delta^{\perp}\|_F\right)\sqrt{h_{K_F}}\,\|\mathbf{q}_h-\mathbf{q}\|_{K_F}.$$

Next, we multiply both sides by $\mathsf{w}_F^{-1/2}$, apply the Cauchy-Schwarz inequality and use the last identity of the representation Lemma 4.3 to obtain

$$\begin{aligned} \mathsf{w}_F^{-1/2}\|\delta\|_F^2 &\le C\,\|c\|_{L^\infty(\Omega)}\left(\sum_{m=\ell+1}^{\ell+\delta\ell_F}\kappa\,\mathsf{c}\,\eta^{m-1-\ell}+1\right)^{1/2}\left(\frac{h_{K_F}}{\mathsf{w}_F}\right)^{1/2}\|\mathbf{q}_h-\mathbf{q}\|_{K_F}\,\|\delta\|_F\\ &\le C\,\|c\|_{L^\infty(\Omega)}\left(\frac{\kappa\,\mathsf{c}}{1-\eta}+1\right)^{1/2}\left(\frac{h_{K_F}}{\mathsf{w}_F}\right)^{1/2}\|\mathbf{q}_h-\mathbf{q}\|_{K_F}\,\|\delta\|_F, \end{aligned}$$

by the fact that we are assuming $\eta \in (0,1)$. Next, we note that, by definition of w_F, (2.11), $h_{K_F}/\mathsf{w}_F = h_{K_F}/h_F \le \sigma$, by the shape-regularity assumption (2.1a). This implies that

$$\mathsf{w}_F^{-1/2}\|\delta\|_F \le C\,\|c\|_{L^\infty(\Omega)}\left(\frac{\kappa\,\mathsf{c}}{1-\eta}+1\right)^{1/2}\sigma^{1/2}\,\|\mathbf{q}_h-\mathbf{q}\|_{K_F},$$

and so,

$$\begin{aligned}
\Theta &:= \|(P_M - P_{\mathcal{M}})\epsilon_h^{\widehat{u}}\|^2_{\partial\mathcal{T}_h,\mathsf{w}^{-1}} \\
&= \sum_{\ell=\ell_{\min}}^{\ell_{\max}} \sum_{K\in\mathcal{T}_h\cap\mathsf{T}_h^\ell} \sum_{F\in\mathcal{F}(K)} \mathsf{w}_F^{-1}\|\delta\|_F^2 \\
&\le C\,\|c\|^2_{L^\infty(\Omega)}\left(\frac{\kappa\,\mathsf{c}}{1-\eta}+1\right)\sigma \sum_{\ell=\ell_{\min}}^{\ell_{\max}} \sum_{K\in\mathcal{T}_h\cap\mathsf{T}_h^\ell} \sum_{F\in\mathcal{F}(K)} \|\mathbf{q}_h-\mathbf{q}\|^2_{K_F} \\
&\le C\,\|c\|^2_{L^\infty(\Omega)}\left(\frac{\kappa\,\mathsf{c}}{1-\eta}+1\right)\sigma\,(\ell_{\max}-\ell_{\min}+1)\,(d+1)\,\|\mathbf{q}_h-\mathbf{q}\|^2_\Omega \\
&= C\,\|c\|^2_{L^\infty(\Omega)}\,C_\Lambda^2\|\mathbf{q}_h-\mathbf{q}\|^2_\Omega.
\end{aligned}$$

The estimate of $|\Lambda_P(z)|$ now follows by inserting the above estimate in the estimate of Lemma 4.1, namely, in the inequality

$$|\Lambda_P(z)| \le \|(P_M - P_{\mathcal{M}})\epsilon_h^{\widehat{u}}\|_{\partial\mathcal{T}_h,\mathsf{w}^{-1}}\,\|(P_M - P_{\mathcal{M}})z\|_{\partial\mathcal{T}_h,\mathsf{w}}.$$

Thus, to complete the proof of Lemma 3.4, we only have to prove Lemma 4.4. We do that in the remainder of the section.

Step 6: Proof of Lemma 4.4, estimate of $\langle\epsilon_h^{\widehat{u}},\delta^{0,m}\rangle_F$. To obtain the estimate of this quantity, we are going to use the following auxiliary result.

Lemma 4.5. *Let $\widetilde{\mathbb{K}}$ be any element in $\widetilde{\mathbb{T}}_h^{m-1}(F,K)$; set $\widetilde{\mathbb{F}} := F\cap\partial\widetilde{\mathbb{K}}$. Then there is a function $\boldsymbol{r}^{0,m}:\widetilde{\mathbb{K}}\to\mathbb{R}^d$ such that*

$$\begin{aligned}
&\text{(i)}\quad \boldsymbol{r}^{0,m}|_{\mathsf{K}} \in \boldsymbol{V}(\mathsf{K}) \quad \forall \mathsf{K}\in\mathsf{T}_h^m : \mathsf{K}\subset\widetilde{\mathbb{K}},\\
&\text{(ii)}\quad \boldsymbol{r}^{0,m}\in\boldsymbol{H}(div,\widetilde{\mathbb{K}}),\\
&\text{(iii)}\quad \boldsymbol{r}^{0,m}\cdot\boldsymbol{n} = \delta^{0,m}\,\chi_{\widetilde{\mathbb{F}}} \ \textit{on}\ \partial\widetilde{\mathbb{K}},\\
&\text{(iv)}\quad \nabla\cdot\boldsymbol{r}^{0,m} = 0\ \textit{on}\ \widetilde{\mathbb{K}},\\
&\text{(v)}\quad \|\boldsymbol{r}^{0,m}\|_{\widetilde{\mathbb{K}}} \le C\,\kappa^{1/2}\,h_{\widetilde{\mathbb{K}}}^{1/2}\,\|\delta^{0,m}\|_{\widetilde{\mathbb{F}}}.
\end{aligned}$$

Here κ is the local uniformity constant of assumption (2.1c) *on the meshes and the constant C depends only on the shape-regularity constant σ.*

Proof. To prove this result, we proceed as follows. First, note that $\widetilde{\mathbb{K}}$ is partitioned into a set of elements $\mathsf{K}\in\mathsf{T}_h^m$. For every such $\mathsf{K}\in\mathsf{T}_h^m$, we define $\boldsymbol{r}|_{\mathsf{K}}$ to be the element of $RT_0(\mathsf{K}) := \boldsymbol{P}_0(\mathsf{K}) + \boldsymbol{x}\,P_0(\mathsf{K})$ such that $\boldsymbol{r}\cdot\boldsymbol{n} = \delta^{0,m}\,\chi_{\widetilde{\mathbb{F}}\cap\partial\mathsf{K}}$ on $\partial\mathsf{K}$. This can be done because, by property (4.2a) of the representation result Lemma 4.3, $\delta^{0,m}$ is constant on $\widetilde{\mathbb{F}}\cap\partial\mathsf{K}$.

Next, we introduce the solution ϕ of the auxiliary problem

$$-\Delta\phi = \nabla\cdot\boldsymbol{r} \text{ in } \widetilde{\mathbb{K}},\quad \boldsymbol{n}\cdot\nabla\phi = 0 \text{ on } \partial\widetilde{\mathbb{K}},\quad \text{and}\quad (\phi,1)_{\widetilde{\mathbb{K}}} = 0.$$

Note that ϕ is well defined because of property (4.2b) of the representation result Lemma 4.3. Finally, we set $\boldsymbol{r}^{0,m} := \boldsymbol{r} + \Pi_0^{\mathrm{RT}}\nabla\phi$, where Π_0^{RT} is the Raviart-Thomas projection [4] onto the space $RT_0(\mathsf{K})$ on each element $\mathsf{K}\in\mathsf{T}_h^m$ such that $\mathsf{K}\subset\widetilde{\mathbb{K}}$.

Let us prove that this is the function we seek. Property (iii) holds by construction. Property (ii) also holds by construction since both $\boldsymbol{r}$ and $\Pi_0^{\mathrm{RT}}\nabla\phi$ lie on $\boldsymbol{H}(div,\widetilde{\mathbb{K}})$. Property (iv) follows from the definition of ϕ and from the so-called *commutativity* property of the Raviart-Thomas projection. Property (i) follows from the simple fact that a divergence free function in $RT_0(\mathsf{K})$ must be constant, thus in $\boldsymbol{V}(\mathsf{K})$: on each element $\mathsf{K} \in \mathsf{T}_h^m : \mathsf{K} \subset \widetilde{\mathbb{K}}$, $\boldsymbol{r}^{0,m}$ is divergence free (property (iv)) and is in $RT_0(\mathsf{K})$ since both $\boldsymbol{r}$ and $\Pi_0^{\mathrm{RT}} \nabla\phi$ lie in $RT_0(\mathsf{K})$.

It remains to prove property (v). We have that

$$\begin{aligned}\|\boldsymbol{r}^{0,m}\|_{\widetilde{\mathbb{K}}} &\le \|\boldsymbol{r}\|_{\widetilde{\mathbb{K}}} + \|(\mathrm{Id} - \Pi_0^{\mathrm{RT}})\nabla\phi\|_{\widetilde{\mathbb{K}}} + \|\nabla\phi\|_{\widetilde{\mathbb{K}}} \\ &\le \|\boldsymbol{r}\|_{\widetilde{\mathbb{K}}} + C\,(\max_{\mathsf{K}\in\mathsf{T}_h^m:\mathsf{K}\subset\widetilde{\mathbb{K}}} h_{\mathsf{K}})\,|\nabla\phi|_{\boldsymbol{H}^1(\widetilde{\mathbb{K}})} + \|\nabla\phi\|_{\widetilde{\mathbb{K}}},\end{aligned}$$

by the approximation properties of the Raviart-Thomas projection.

Now, from the definition of ϕ, we get that

$$\begin{aligned}\|\nabla\phi\|^2_{\widetilde{\mathbb{K}}} &= -\,(\boldsymbol{r},\nabla\phi)_{\widetilde{\mathbb{K}}} + \langle\phi, \boldsymbol{r}\cdot\boldsymbol{n}\rangle_{\partial\widetilde{\mathbb{K}}} \\ &= -\,(\boldsymbol{r},\nabla\phi)_{\widetilde{\mathbb{K}}} + \langle\phi, \delta^{0,m}\rangle_{\widetilde{\mathbb{F}}} && \text{by property (iii),} \\ &\le \|\boldsymbol{r}\|_{\widetilde{\mathbb{K}}}\|\nabla\phi\|_{\widetilde{\mathbb{K}}} + \|\phi\|_{\partial\widetilde{\mathbb{K}}}\,\|\delta^{0,m}\|_{\widetilde{\mathbb{F}}} \\ &\le \|\boldsymbol{r}\|_{\widetilde{\mathbb{K}}}\|\nabla\phi\|_{\widetilde{\mathbb{K}}} + C\,h^{1/2}_{\widetilde{\mathbb{K}}}\,\|\nabla\phi\|_{\widetilde{\mathbb{K}}}\|\delta^{0,m}\|_{\widetilde{\mathbb{F}}},\end{aligned}$$

by the trace inequality and the fact that the average of ϕ of $\widetilde{\mathbb{K}}$ is equal to zero. This implies that

$$\|\nabla\phi\|_{\widetilde{\mathbb{K}}} \le \|\boldsymbol{r}\|_{\widetilde{\mathbb{K}}} + C\,h^{1/2}_{\widetilde{\mathbb{K}}}\,\|\delta^{0,m}\|_{\widetilde{\mathbb{F}}}.$$

Also, note that, since the normal component of $\nabla\phi$ is zero on $\partial\widetilde{\mathbb{K}}$, we have that

$$|\nabla\phi|_{\boldsymbol{H}^1(\widetilde{\mathbb{K}})} = \|\Delta\phi\|_{\widetilde{\mathbb{K}}} = \|\nabla\cdot\boldsymbol{r}\|_{\widetilde{\mathbb{K}}},$$

by definition of ϕ. We thus get that

$$\|\boldsymbol{r}^{0,m}\|_{\widetilde{\mathbb{K}}} \le 2\,\|\boldsymbol{r}\|_{\widetilde{\mathbb{K}}} + C\,(\max_{\mathsf{K}\in\mathsf{T}_h^m:\mathsf{K}\subset\widetilde{\mathbb{K}}} h_{\mathsf{K}})\,\|\nabla\cdot\boldsymbol{r}\|_{\widetilde{\mathbb{K}}} + C\,h^{1/2}_{\widetilde{\mathbb{K}}}\,\|\delta^{0,m}\|_{\widetilde{\mathbb{F}}}.$$

Finally, a simple computation gives that

$$\begin{aligned}\|\boldsymbol{r}\|_{\widetilde{\mathbb{K}}} &\le C\,(\max_{\mathsf{K}\in\mathsf{T}_h^m:\mathsf{K}\subset\widetilde{\mathbb{K}}} h_{\mathsf{K}}^{1/2})\|\delta^{0,m}\|_{\widetilde{\mathbb{F}}} \\ &\le C\,h^{1/2}_{\widetilde{\mathbb{K}}}\|\delta^{0,m}\|_{\widetilde{\mathbb{F}}},\end{aligned}$$

and that

$$\begin{aligned}\|\nabla\cdot\boldsymbol{r}\|_{\widetilde{\mathbb{K}}} &\le C\,(\max_{\mathsf{K}\in\mathsf{T}_h^m:\mathsf{K}\subset\widetilde{\mathbb{K}}} h_{\mathsf{K}}^{-1/2})\|\delta^{0,m}\|_{\widetilde{\mathbb{F}}} \\ &\le C\,(\max_{\mathsf{K}\in\mathsf{T}_h^m:\mathsf{K}\subset\widetilde{\mathbb{K}}} h_{\mathsf{K}})^{-1}\,\kappa^{1/2}\,h^{1/2}_{\widetilde{\mathbb{K}}}\,\|\delta^{0,m}\|_{\widetilde{\mathbb{F}}},\end{aligned}$$

by the local uniformity assumption on the meshes (2.1c). This implies that

$$\|\boldsymbol{r}^{0,m}\|_{\widetilde{\mathsf{K}}} \le C\,\kappa^{1/2}\,h^{1/2}_{\widetilde{\mathbb{K}}}\,\|\delta^{0,m}\|_{\widetilde{\mathbb{F}}}.$$

This completes the proof. $\square$

The estimate we seek is the following.

Lemma 4.6. *Let $\widetilde{\mathbb{K}} \in \widetilde{\mathbb{T}}_h^{m-1}(F,K)$ and $\widetilde{\mathbb{F}} := F \cap \partial\widetilde{\mathbb{K}}$. We have that*

$$|\,\langle\epsilon_h^{\widehat{u}}, \delta^{0,m}\rangle_{\widetilde{\mathbb{F}}}\,| \le C\,\|c\|_{L^\infty(\widetilde{\mathbb{K}})}\,\kappa^{1/2}\,\|\mathbf{q}_h - \mathbf{q}\|_{\widetilde{\mathbb{K}}}\,h^{1/2}_{\widetilde{\mathbb{K}}}\,\|\delta^{0,m}\|_{\widetilde{\mathbb{F}}}.$$

The constant C depends only on the shape-regularity constant σ.

Proof. Let $\boldsymbol{r}^{0,m}$ be the auxiliary function of the previous lemma. Then

$$\begin{aligned}\langle \epsilon_h^{\widehat{u}}, \delta^{0,m}\rangle_{\widetilde{\mathbb{F}}} &= \langle \epsilon_h^{\widehat{u}}, \boldsymbol{r}^{0,m}\cdot \mathbf{n}\rangle_{\partial\widetilde{\mathbb{K}}} && \text{by property (iii) of Lemma 4.5,}\\ &=(c\,(\mathbf{q}_h-\mathbf{q}), \boldsymbol{r}^{0,m})_{\widetilde{\mathbb{K}}} + (\epsilon_h^{u}, \nabla\cdot \boldsymbol{r}^{0,m})_{\widetilde{\mathbb{K}}},\end{aligned}$$

by the first error equation (3.1a) with $\boldsymbol{r} := \boldsymbol{r}^{0,m}$, which is an allowed choice of test function $\boldsymbol{r}$ by properties (i) and (ii) of Lemma 4.5. By property (iv) of Lemma 4.5, this implies that

$$\langle \epsilon_h^{\widehat{u}}, \delta^{0,m}\rangle_{\widetilde{\mathbb{F}}} = (c\,(\mathbf{q}_h-\mathbf{q}), \boldsymbol{r}^{0,m})_{\widetilde{\mathbb{K}}},$$

and the result follows by property (v) of Lemma 4.5. This completes the proof. □

Corollary 4.7. *We have that*

$$|\,\langle \epsilon_h^{\widehat{u}}, \delta^{0,m}\rangle_F\,| \le C\,\|c\|_{L^\infty(\Omega)}\,(\kappa\,\mathsf{c}\,\eta^{m-1-\ell}\,h_{K_F})^{1/2}\,\|\mathbf{q}_h-\mathbf{q}\|_{K_F}\,\|\delta^{0,m}\|_F.$$

The constant C depends only on the shape-regularity constant σ.

Proof. By the fact that F can be covered by the set of $F\cap\partial\widetilde{\mathbb{K}}$, with $\widetilde{\mathbb{K}}\in\widetilde{\mathbb{T}}_h^{m-1}(F,K)$,

$$|\,\langle \epsilon_h^{\widehat{u}}, \delta^{0,m}\rangle_F\,| \le \sum_{\widetilde{\mathbb{K}}\in\widetilde{\mathbb{T}}_h^{m-1}(F,K)} |\,\langle \epsilon_h^{\widehat{u}}, \delta^{0,m}\rangle_{F\cap\partial\widetilde{\mathbb{K}}}\,|,$$

and by the previous lemma

$$\begin{aligned}|\,\langle \epsilon_h^{\widehat{u}}, \delta^{0,m}\rangle_F\,| &\le C\,\|c\|_{L^\infty(\Omega)} \sum_{\widetilde{\mathbb{K}}\in\widetilde{\mathbb{T}}_h^{m-1}(F,K)} (\kappa\,h_{\widetilde{\mathbb{K}}})^{1/2}\|\mathbf{q}_h-\mathbf{q}\|_{\widetilde{\mathbb{K}}}\,\|\delta^{0,m}\|_{\widetilde{\mathbb{F}}}\\ &\le C\,\|c\|_{L^\infty(\Omega)} \max_{\widetilde{\mathbb{K}}\in\widetilde{\mathbb{T}}_h^{m-1}(F,K)} (\kappa\,h_{\widetilde{\mathbb{K}}})^{1/2}\,\|\mathbf{q}_h-\mathbf{q}\|_{K_F}\,\|\delta^{0,m}\|_F\\ &\le C\,\|c\|_{L^\infty(\Omega)}\,(\kappa\,\mathsf{c}\,\eta^{m-1-\ell}\,h_{K_F})^{1/2}\,\|\mathbf{q}_h-\mathbf{q}\|_{K_F}\,\|\delta^{0,m}\|_F,\end{aligned}$$

by the local uniformity assumption (2.1c) and the local refinement assumption (2.1d) on the meshes. This completes the proof. □

Step 7: Proof of Lemma 4.4, estimate of $\langle \epsilon_h^{\widehat{u}}, \delta^{\perp}\rangle_F$. To estimate this quantity, we proceed as in the previous step. So, we begin by obtaining the following auxiliary result.

Lemma 4.8. *Let K' be any element in $\mathcal{T}_h(F,K)$; set $F' := F\cap\partial K'$. Then there is a function $\boldsymbol{r}^\perp : K'\to\mathbb{R}^d$ such that*

$$\begin{aligned}&\text{(i)}\quad \boldsymbol{r}^\perp\in \boldsymbol{V}(K'),\\ &\text{(ii)}\quad \boldsymbol{r}^\perp\cdot\boldsymbol{n} = \delta^\perp\,\chi_{F'}\ \text{ on } \partial K',\\ &\text{(iii)}\quad \nabla\cdot\boldsymbol{r}^\perp = 0\ \text{ on } K',\\ &\text{(iv)}\quad \|\boldsymbol{r}^\perp\|_{K'}\le C\,h_{K'}^{1/2}\,\|\delta^\perp\|_{F'}.\end{aligned}$$

The constant C depends only on the shape-regularity constant σ.

Proof. If $k(K')=0$, then $\delta^\perp=0$ by properties (4.3a) and (4.3b), and the result holds with $\boldsymbol{r}^\perp := \mathbf{0}$. Let us now assume that $k(K')\ge 1$. By the definition of the BDM projection (see [4]) there is a unique function $\boldsymbol{r}^\perp\in\boldsymbol{P}_{k(K')}(K')$ such that

$$\begin{aligned}\langle \boldsymbol{r}^\perp\cdot\boldsymbol{n}, \mu\rangle_{F''} &= \langle \delta^\perp, \mu\rangle_{F''\cap F'} && \forall\,\mu\in P_{k(K')}(F'')\quad\text{for all faces } F'' \text{ of } K',\\ (\boldsymbol{r}^\perp, \nabla w)_{K'} &= \langle \delta^\perp, w-\overline{w}\rangle_{F'} && \forall\, w\in P_{k(K')-1}(K'),\\ (\boldsymbol{r}^\perp, \boldsymbol{\eta})_{K'} &= 0 && \forall\,\boldsymbol{\eta}\in\boldsymbol{\Phi}_{k(K')},\end{aligned}$$

where $\overline{w}$ denotes the average of w over K', and $\mathbf{\Phi}_{k(K')}$ is the subset of divergence-free functions in $\boldsymbol{P}_{k(K')}(K')$ that have zero normal component on $\partial K'$. Let us show that this function satisfies (i) to (iv).

It certainly satisfies property (i), by construction. It satisfies property (ii) by the first equation defining $\boldsymbol{r}^{\perp}$ and by property (4.3a) of the representation Lemma 4.3. To see that it satisfies property (iii), we note that, for any $w \in P_{k(K')-1}(K')$,

$$\begin{aligned}(\nabla\cdot\boldsymbol{r}^{\perp}, w)_{K'} &= -\,(\boldsymbol{r}^{\perp}, \nabla w)_{K'} + \langle \boldsymbol{r}^{\perp}\cdot\boldsymbol{n}, w\rangle_{\partial K'}\\ &= -\,(\boldsymbol{r}^{\perp}, \nabla w)_{K'} + \langle \delta^{\perp}, w\rangle_{F'} \qquad \text{by property (ii)},\\ &= \langle \delta^{\perp}, \overline{w}\rangle_{F'},\end{aligned}$$

by the second equation defining $\boldsymbol{r}^{\perp}$. Property (iii) now follows from assumption (4.3b) of the representation Lemma 4.3. Finally, property (iv) follows from the definition of $\boldsymbol{r}^{\perp}$ by a scaling argument. This completes the proof. $\square$

The estimate we seek is the following.

Lemma 4.9. *Let K' be any simplex in $\mathcal{T}_h(F, K)$; set $F' := F \cap \partial K'$. Then we have*

$$|\,\langle \epsilon_h^{\widehat{u}}, \delta^{\perp}\rangle_{F'}\,| \le C\,\|c\|_{L^\infty(K')}\,|\|\mathbf{q}_h - \mathbf{q}\|_{K'}\,h_{K'}^{1/2}\,\|\delta^{\perp}\|_{F'}.$$

The constant C depends only on the shape-regularity constant σ.

Proof. Let $\boldsymbol{r}^{\perp}$ be the auxiliary function of the previous lemma. Then, by property (ii) of Lemma 4.8,

$$\begin{aligned}\langle \epsilon_h^{\widehat{u}}, \delta^{\perp}\rangle_{F'} &= \langle \epsilon_h^{\widehat{u}}, \boldsymbol{r}^{\perp}\cdot\mathbf{n}\rangle_{\partial K'}\\ &= (c\,(\mathbf{q}_h - \mathbf{q}), \boldsymbol{r}^{\perp})_{K'} + (\epsilon_h^{u}, \nabla\cdot\boldsymbol{r}^{\perp})_{K'}\end{aligned}$$

by the first error equation (3.1a) with $\boldsymbol{r} := \boldsymbol{r}^{\perp}$, which we can do by property (i) of Lemma 4.8. By property (iii) of Lemma 4.8, this implies that

$$\langle \epsilon_h^{\widehat{u}}, \delta^{\perp}\rangle_{F'} = (c\,(\mathbf{q}_h - \mathbf{q}), \boldsymbol{r}^{\perp})_{K'},$$

and the result follows by property (iv) of Lemma 4.8. This completes the proof. $\square$

Corollary 4.10. *We have that*

$$|\,\langle \epsilon_h^{\widehat{u}}, \delta^{\perp}\rangle_{F}\,| \le C\,\|c\|_{L^\infty(\Omega)}\,h_{K_F}^{1/2}\|\mathbf{q}_h - \mathbf{q}\|_{K_F}\,\|\delta^{\perp}\|_{F}.$$

The constant C depends only on the shape-regularity constant σ.

Proof. By the fact that F can be covered by the set of $F\cap\partial K'$, with $K' \in \mathcal{T}_h(F, K)$,

$$|\,\langle \epsilon_h^{\widehat{u}}, \delta^{\perp}\rangle_{F}\,| \le \sum_{K'\in\mathcal{T}_h(F,K)} |\,\langle \epsilon_h^{\widehat{u}}, \delta^{\perp}\rangle_{F\cap\partial K'}\,|,$$

and by the previous lemma,

$$\begin{aligned}|\,\langle \epsilon_h^{\widehat{u}}, \delta^{\perp}\rangle_{F}\,| &\le C\,\|c\|_{L^\infty(\Omega)} \sum_{K'\in\mathcal{T}_h(F,K)} h_{K'}^{1/2}\,\|\mathbf{q}_h - \mathbf{q}\|_{K'}\,\|\delta^{\perp}\|_{F'}\\ &\le C\,\|c\|_{L^\infty(\Omega)}\,h_{K_F}^{1/2}\|\mathbf{q}_h - \mathbf{q}\|_{K_F}\,\|\delta^{\perp}\|_{F}.\end{aligned}$$

This completes the proof. $\square$

The proof of Lemma 3.4 is now complete.

5. Extensions and concluding remarks

Note that the same techniques presented in this paper can be used to analyze the hybridized Raviart-Thomas (RT) and the Brezzi-Douglas-Marini mixed methods for pure diffusion equations on the semimatching nonconforming meshes under consideration. A priori error estimates for the RT method were obtained in [9] for the variable-degree version of this method but only for conforming meshes.

Our analysis can also be easily extended to include the six HDG methods (and mixed methods) on d-dimensional rectangles (for $d = 2, 3$) considered in [15]. In fact, the only step that needs to be modified when the elements are d-dimensional rectangles is the key Lemma 4.4. Therefore, to prove this extension, we only need to ensure that the key Lemma 4.4 also holds in the case of d-dimensional rectangles. This is done by slightly modifying the proofs of Lemmas 4.5 and 4.8, as we are going to show below.

For the sake of completeness, let us describe the above-mentioned methods. We use $\boldsymbol{V}(K)$, $W(K)$ and $M(F)$ to denote the local spaces. That is, the finite element spaces defined by (2.3) becomes

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

We use $P_k(K)$ to denote the space of polynomials of total degree k on K, $\widetilde{P}_k(K)$ the space of homogeneous polynomials of degree k. $\boldsymbol{P}_k(K)$ denotes the space $[P_k(K)]^d$ and $\boldsymbol{\Phi}_k(K)$ denotes the subspace of functions in $\boldsymbol{P}_k(K)$ which are divergence-free and whose normal component on ∂K is zero. We are ready to display three methods in Table 2 that have $M(F) = P_k(F)$. The mixed methods $\mathbf{BDFM}_{[k+1]}$ and $\mathbf{BDM}_{[k]}$ are well known and $\mathbf{HDG}^P_{[k]}$ was newly discovered in [15].

When $M(F) = Q_k(F)$, we display three methods in Table 3. The mixed method $\mathbf{RT}_{[k]}$ is well known, $\mathbf{TNT}_{[k]}$ and $\mathbf{HDG}^Q_{[k]}$ were newly discovered in [15]. Here, $P_{\ell_1,\ell_2}(K)$ for $d = 2$ and $P_{\ell_1,\ell_2,\ell_3}(K)$ for $d = 3$ denote the space of polynomials of degree ℓ_i on the ith variable for $i = 1, \dots, d$. The space of polynomials of degree k in each variable is denoted by $Q_k(K)$ and its corresponding vector space by $\mathbf{Q}_k(K)$. Moreover, we set

$$\begin{aligned} \mathbf{H}^k(K) :=& \{((x^2-x)x^{k-1}(aL_k(y)+b), (y^2-y)y^{k-1}(cL_k(z)+d), \\ & (z^2-z)z^{k-1}(eL_k(x)+f)) : (a,b,c,d,e,f) \in \mathbb{R}^6\}, \\ \mathbf{H}^k_M(K) :=& \mathbf{H}^k(K) \oplus \{((x^2-x)x^{k-1}L_k(y)L_k(z), 0, 0)\}, \end{aligned}$$

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

Let us now show how to modify the proofs of Lemmas 4.5 and 4.8 in the case in which the elements are d-dimensional rectangles.

Modification of the proof of Lemma 4.5. To do this, we only need to replace $RT_0(\mathsf{K})$ by

$$RT_{[0]}(\mathsf{K}) = \begin{cases} P_{1,0} \times P_{0,1} & \text{for } d = 2, \\ P_{1,0,0} \times P_{0,1,0} \times P_{0,0,1} & \text{for } d = 3. \end{cases}$$

TABLE 2. Methods for which $M(F) = P_k(F), k \geq 0$.

Method	K is a square		K is a cube	
	$\boldsymbol{V}(K)$	$W(K)$	$\boldsymbol{V}(K)$	$W(K)$
$\mathbf{BDFM}_{[k+1]}$	$P_{k+1}(K)\backslash\{y^{k+1}\}$ $\times\left(P_{k+1}(K)\backslash\{x^{k+1}\}\right)$	$P_k(K)$	$P_{k+1}(K)\backslash\{\widetilde{P}^{k+1}(y,z)\}$ $\times\left(P_{k+1}(K)\backslash\{\widetilde{P}^{k+1}(x,z)\}\right)$ $\times\left(P_{k+1}(K)\backslash\{\widetilde{P}^{k+1}(x,y)\}\right)$	$P_k(K)$
$\mathbf{HDG}^P_{[k]}$	$\boldsymbol{P}_k(K)$ $\bigoplus\nabla\times(xy\widetilde{P}_k(K))$	$P_k(K)$	$\boldsymbol{P}_k(K)$ $\bigoplus\nabla\times(yz\widetilde{P}_k(K),0,0)$ $\bigoplus\nabla\times(0,zx\widetilde{P}_k(K),0)$	$P_k(K)$
$\mathbf{BDM}_{[k]}$ $k\geq 1$	$\boldsymbol{P}_k(K)$ $\bigoplus\nabla\times(xyx^k)$ $\bigoplus\nabla\times(xyy^k)$	$P_{k-1}(K)$	$\boldsymbol{P}_k(K)$ $\bigoplus\nabla\times(0,0,xy\widetilde{P}_k(y,z))$ $\bigoplus\nabla\times(0,xz\widetilde{P}_k(x,y),0)$ $\bigoplus\nabla\times(yz\widetilde{P}_k(x,z),0,0)$	$P_{k-1}(K)$

TABLE 3. Methods for which $M(F) = Q_k(F), k \geq 0$.

Method	K is a square		K is a cube	
	$\boldsymbol{V}(K)$	$W(K)$	$\boldsymbol{V}(K)$	$W(K)$
$\mathbf{RT}_{[k]}$	$P_{k+1,k}(K)$ $\times P_{k,k+1}(K)$	$Q_k(K)$	$P_{k+1,k,k}(K)$ $\times P_{k,k+1,k}(K)$ $\times P_{k,k,k+1}(K)$	$Q_k(K)$
$\mathbf{TNT}_{[k]}$	$\boldsymbol{Q}_k(K)$ $\bigoplus\{(x^{k+1},0),(0,y^{k+1})\}$ $\bigoplus\{(x^{k+1}y^k,0)\}$	$Q_k(K)$	$\mathbf{Q}_k(K)\bigoplus\mathbf{H}^k_M(K)$	$Q_k(K)$
$\mathbf{HDG}^Q_{[k]}$	$\boldsymbol{Q}_k(K)$ $\bigoplus\{(x^{k+1},0),(0,y^{k+1})\}$	$Q_k(K)$	$\mathbf{Q}_k(K)\bigoplus\mathbf{H}^k(K)$	$Q_k(K)$

and Π_0^{RT} by $\Pi_{[0]}^{\mathrm{RT}}$. To verify that $\boldsymbol{r}^{0,m}$ is the function that we seek, the only different (and tedious) part is to check the fact that divergence free functions in $RT_{[0]}(\mathsf{K})$ are in $\boldsymbol{V}(\mathsf{K})$ for all the elements listed in Table 2 and Table 3. This is indeed the case.

Modification of the proof of Lemma 4.8. For the three methods in Table 2 using $M(F) = P_k(F)$, no change is needed since $\boldsymbol{P}_{k(K')}(K') \subset \boldsymbol{V}(K')$. For the other three methods in Table 3 using $M(F) = Q_k(F)$, we simply need to change the projection above defining $\boldsymbol{r}^\perp$ to the the following so-called **TNT** projection, see [14, 15]. That is, $\boldsymbol{r}^\perp$ is defined to be the unique function in $\boldsymbol{Q}_{k(K')}(K')$ such that

$$\begin{aligned}
\langle \boldsymbol{r}^\perp\cdot\boldsymbol{n},\mu\rangle_{F''} &= \langle\delta^\perp,\mu\rangle_{F''\cap F'} && \forall\,\mu\in Q_{k(K')}(F'') \quad \text{for all faces } F'' \text{ of } K',\\
(\boldsymbol{r}^\perp,\nabla w)_{K'} &= \langle\delta^\perp, w-\overline{w}\rangle_{F'} && \forall\, w\in Q_{k(K')}(K'),\\
(\boldsymbol{r}^\perp,\boldsymbol{\eta})_{K'} &= 0 && \forall\,\boldsymbol{\eta}\in\boldsymbol{\Phi}^Q_{k(K')},
\end{aligned}$$

where $\boldsymbol{\Phi}^Q_{k(K')}$ is the subset of divergence-free functions in $\boldsymbol{Q}_{k(K')}(K')$ that have zero normal component on $\partial K'$.

Appendix A. Estimate of Z_P

Here, we prove the estimate (2.13). To do that, we begin by noting that, by definition (see Theorem 2.1) we have that $Z_P := (P_M - P_{\mathcal{M}})(\boldsymbol{Z}\cdot\boldsymbol{n} + Z)$, $\boldsymbol{Z} := \mathbf{q} + (u - P_M u)\,\mathbf{v}$, and $Z := \tau\, u$, and we easily get that

$$\begin{aligned} Z_P &= (P_M - P_{\mathcal{M}})(\mathbf{q}\cdot\mathbf{n} + \tau u + (u - P_M u)\,\mathbf{v}\cdot\mathbf{n}) \\ &= (P_M - P_{\mathcal{M}})(\mathbf{q}\cdot\mathbf{n}) + \tau(P_M - P_{\mathcal{M}})(u) + (P_M - P_{\mathcal{M}})((u - P_M u)\,\mathbf{v}\cdot\mathbf{n}), \end{aligned}$$

by the assumption **A2** on τ. Now, pick any $K \in \mathcal{T}_{P,h}$ and let $\mathit{\Pi}_k$ and $\boldsymbol{\mathit{\Pi}}_k$ be the $L^2(K)$-projections into $P_k(K)$ and $\boldsymbol{P}_k$, respectively. Then we have, for $k := k(K) \le k(F)$ for any $F \in \mathcal{F}(\mathcal{K})$, that

$$\begin{aligned} Z_P = {} & (P_M - P_{\mathcal{M}})((\mathbf{q} - \boldsymbol{\mathit{\Pi}}_k\mathbf{q})\cdot\mathbf{n}) + \tau(P_M - P_{\mathcal{M}})(u - \mathit{\Pi}_k u) \\ & + (P_M - P_{\mathcal{M}})((u - P_M u)\,\mathbf{v}\cdot\mathbf{n}), \end{aligned}$$

since we have that $k = k(K) \le k(F)$ for any $F \subset \partial K$. Finally,

$$\begin{aligned} Z_P = {} & (P_M - P_{\mathcal{M}})((\mathbf{q} - \boldsymbol{\mathit{\Pi}}_k\mathbf{q})\cdot\mathbf{n}) + \tau(P_M - P_{\mathcal{M}})(u - \mathit{\Pi}_k u) \\ & + (P_M - P_{\mathcal{M}})((u - P_M u)\,(\mathbf{v} - \mathbf{v_0})\cdot\mathbf{n}), \end{aligned}$$

since $(P_M - P_{\mathcal{M}})(u - P_M u) = 0$ and since $\mathbf{v}_0$ is a constant on each face $F \in \mathcal{F}(\mathcal{K})$; see (3.2). This readily implies that

$$\begin{aligned} \|Z_P\|_{\partial K} &\le 2\|\boldsymbol{\mathit{\Pi}}_k\mathbf{q} - \mathbf{q}\|_{\partial K} + 2(\|\tau\|_{L^\infty(\partial K)} + h\,|\mathbf{v}|_{W^{1,\infty}(K)})\|\mathit{\Pi}_k u - u\|_{\partial K} \\ &\le C\, h_K^{k(K)+\frac{1}{2}}\,\mathcal{D}_K(\mathbf{q}, u), \end{aligned}$$

by Lemma 3.9 of [16]. Finally, invoking the definition of w, (2.11b), gives

$$\|Z_P\|_{\partial K,\mathsf{w}} \le C\, h_K^{k(K)+1}\,\mathcal{D}_K(\mathbf{q}, u).$$

This completes the proof of the estimate (2.13).

Acknowledgements

The authors would like to thank the referees for constructive criticism that led to an improved presentation of the material in this paper.

References

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

[2] B. Ayuso and L. D. Marini, Discontinuous Galerkin methods for advection-diffusion-reaction problems. *SIAM J. Numer. Anal.*, 47(2):1391–1420, 2009. MR2485457 (2010d:65312)

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

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

[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., 38:1676–1706, 2000. MR1813251 (2002k:65175)

[6] Y. Chen and B. Cockburn. Analysis of variable-degree HDG methods for convection-diffusion equations. Part I: General nonconforming meshes, *IMA J. Num. Anal.*, 32(4):1267–1293, 2012, doi:10.1093/imanum/drr058. MR2991828

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

[8] B. Cockburn, B. Dong, J. Guzmán, M. Restelli, and R. Sacco. Superconvergent and optimally convergent LDG-hybridizable discontinuous Galerkin methods for convection-diffusion-reaction problems. *SIAM J. Sci. Comput.*, 31:3827–3846, 2009. MR2556564 (2010m:65216)

[9] B. Cockburn and J. Gopalakrishnan. Error analysis of variable degree mixed methods for elliptic problems via hybridization. *Math. Comp.*, 74:1653–1677, 2005. MR2164091 (2006e:65215)

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

[11] B. Cockburn, J. Gopalakrishnan, and F.-J. Sayas. A projection-based error analysis of HDG methods. *Math. Comp.*, 79:1351–1367, 2010. MR2629996 (2011d:65354)

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

[13] B. Cockburn, G. Kanschat, I. Perugia, and D. Schötzau, Superconvergence of the local discontinuous Galerkin method for elliptic problems on Cartesian grids, *SIAM J. Numer. Anal.*, 39:264–285, 2001. MR1860725 (2002g:65140)

[14] B. Cockburn and W. Qiu. Commuting diagrams for the TNT elements. *Math. Comp.* (to appear).

[15] B. Cockburn, W. Qiu and K. Shi. Conditions for superconvergence of HDG methods for second-order elliptic problems. *Math. Comp.*, 81:1355–1368, 2012.

[16] P. Houston, C. Schwab, and E. Süli. Discontinuous *hp*-finite element methods for advection-diffusion-reaction problems. *SIAM J. Numer. Anal.*, 39:2133–2163 (electronic), 2002. MR1897953 (2003d:65108)

[17] N. C. Nguyen, J. Peraire, and B. Cockburn. An implicit high-order hybridizable discontinuous Galerkin method for linear convection-diffusion equations. *J. Comput. Phys.*, 228:3232–3254, 2009. MR2513831 (2010h:65232)

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

Department of Mathematics, University of Massachusetts Dartmouth, 285 Old Westport Road, North Dartmouth, Massachusetts 02747

E-mail address: `yanlai.chen@umassd.edu`

School of Mathematics, University of Minnesota, Minneapolis, Minnesota 55455

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