AN ANALYSIS OF THE MINIMAL DISSIPATION LOCAL DISCONTINUOUS GALERKIN METHOD FOR CONVECTION-DIFFUSION PROBLEMS

By

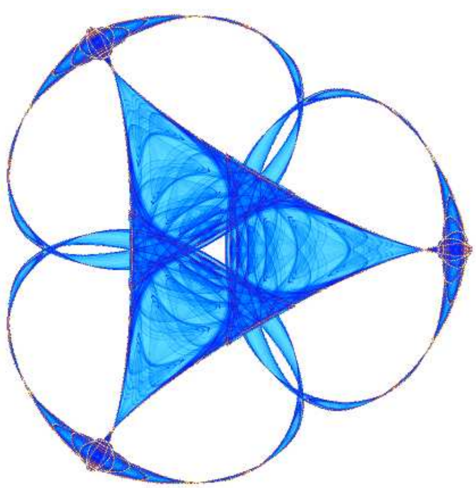
Bernardo Cockburn

and

Bo Dong

IMA Preprint Series # 2146

(November 2006)



INSTITUTE FOR MATHEMATICS AND ITS APPLICATIONS

UNIVERSITY OF MINNESOTA

400 Lind Hall

207 Church Street S.E.

Minneapolis, Minnesota 55455–0436

Phone: 612-624-6066 Fax: 612-626-7370

URL: http://www.ima.umn.edu

AN ANALYSIS OF THE MINIMAL DISSIPATION LOCAL DISCONTINUOUS GALERKIN METHOD FOR CONVECTION-DIFFUSION PROBLEMS

BERNARDO COCKBURN AND BO DONG

ABSTRACT. We analyze the so-called the minimal dissipation local discontinuous Galerkin method for convection-diffusion or diffusion problems. The distinctive feature of this method is that the stabilization parameters associated with the numerical trace of the flux are identically equal to zero in the interior of the domain; this is why its dissipation is said to be minimal. We show that the orders of convergence of the approximations for the potential and the flux using polynomials of degree k are the same as those of all known discontinuous Galerkin methods for both unknowns, namely, $(k+1)$ and k, respectively. Our numerical results verify that these orders of convergence are sharp. The novelty of the analysis is that it bypasses a seemingly indispensable condition, namely, the positivity of the above mentioned stabilization parameters, by using a new, carefully defined projection tailored to the very definition of the numerical traces.

1. INTRODUCTION

In this paper, we analyze the so-called minimal dissipation local discontinuous Galerkin method (MD-LDG) and show that it has the same convergence properties as all other known DG methods even though its stabilization parameters are identically equal to zero in the interior of the domain. To carry out the analysis, we consider the model convection-diffusion problem

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

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

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

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

where $\Omega \subset \mathbb{R}^d$ is a polyhedral domain $(d \geq 2)$, $\partial\Omega_D$ is nonempty, $f \in L^2(\Omega)$ and the diffusion coefficient $\boldsymbol{c} = \boldsymbol{c}(\boldsymbol{x})$ is a symmetric positive definite $d \times d$ matrix function such that

$$(\boldsymbol{c}\,\boldsymbol{x}, \boldsymbol{x}) \leq \eta^2(\boldsymbol{x}, \boldsymbol{x}) \quad \text{and} \quad (\boldsymbol{c}^{-1}\,\boldsymbol{x}, \boldsymbol{x}) \leq \gamma^2(\boldsymbol{x}, \boldsymbol{x}) \quad \forall \boldsymbol{x} \in \mathbb{R}^d,$$

for some positive constants η and γ. The convective velocity $\mathsf{v} = \mathsf{v}(\boldsymbol{x})$ is assumed to be divergence-free with components in $L^\infty(\Omega)$. We assume that $\mathsf{v} \cdot \boldsymbol{n} \geq 0$ on $\partial\Omega_N$. Note that the convective velocity v can be taken to be identically zero.

Key words and phrases. minimal dissipation local discontinuous Galerkin method, convection-diffusion equation.

The first author was supported in part by the National Science Foundation (Grant DMS-0411254) and by the University of Minnesota Supercomputing Institute.

Let us put our results in perspective. In [2] a unified analysis of DG methods for second-order elliptic equations was carried out; it included the interior penalty (IP) method [3, 1], the NIPG method [15], the method of Brezzi et al. [7], the method of Bassi et al. [4], and the local discontinuous Galerkin (LDG) method introduced by Cockburn and Shu in [14]. It was shown that if the stabilization parameters are positive (and suitably chosen), the L^2-norm of the error of the flux $\boldsymbol{q}$ and the potential u are of order k and $k+1$, respectively, when polynomials of degree k are used. In [10], it was shown that when the stabilization parameters associated to the numerical trace of the flux are taken to be $O(1)$, the orders of convergence of the errors in the flux and the potential given by the LDG method are k and $k+\frac{1}{2}$, respectively; when the stabilization parameters are taken to be $O(h^{-1})$, they are k and $k+1$. All these analyses depend on a seemingly indispensable condition, namely, the positivity of the stabilization parameters on all the faces. Since the MD-LDG method we consider is a special LDG method for which the stabilization parameter is taken to be identically zero on all interior faces, a new technique is required to carry out its analysis.

The MD-LDG method has been analyzed in the one-dimensional case in the framework of time-dependent convection-diffusion problems in [11], where optimal error estimates of the hp-version of the method were obtained. It was also analyzed in the framework of steady-state convection-diffusion problems in [12] where superconvergence results of the numerical traces were obtained for the h-version of the method. In both cases, it was shown that the orders of convergence of the approximations to the flux and to the potential are $k+1$ when polynomials of degree k are used. The only analysis of a MD-LDG method in several-space dimensions was carried out in [13] for time-dependent convection-diffusion problems. Therein, the numerical traces associated with the diffusion term were chosen to be the averages; the stabilization parameters were thus taken to be zero. By taking advantage of the time-dependent nature of the problem and of the stabilization effect of the convection, error estimates were obtained; the order of convergence of the approximation to the flux was shown to be k. In our case, we do not rely on the stabilization effect of the convection to compensate for the lack of stabilization in the interior of the domain; our results hold even when the convective term vanishes. Moreover, for our steady state problem, the above mentioned choice of numerical fluxes for the diffusive term does *not* give a well-defined method if the convection terms vanish.

Although no analysis of the MD-LDG method is available for elliptic problems in several space dimensions, the MD-LDG method seems to work well. Indeed, it was successfully implemented by Siddarth and Carrero, in the case of piecewise linear approximations, for the system of linear elasticity; their numerical experiments confirmed first order accuracy for the stress and second order for the displacement, see [16]. In this paper, we present the first analysis of the MD-LDG method on multidimensional steady state convection-diffusion problems with variable coefficients. The novelty of our analysis is the introduction of a projection which allows us to compensate the lack of stabilization of the jumps at the interior borders of the elements. We prove that the MD-LDG method using polynomials of degree k converges with order k for the flux and with order $k+1$ for the potential.

The organization of the paper is as follows. In section 2, we describe the MD-LDG method, and state our main results, namely, the results on existence and uniqueness of the approximation, the error estimates and the conditioning of the stiffness matrix. In section 3, we introduce the new projection and give detailed proofs of all our results. We display numerical experiments that verify the theoretical results in section 4. In section 5, we extend the results to more general numerical traces and curved boundary domains in two dimensions. We end in section 6 with some concluding remarks.

2. The Method and Main Results

In this section we describe the minimal dissipation local discontinuous Galerkin method and state our main theorems.

2.1. **The MD-LDG method.** To describe the MD-LDG method, we begin by introducing the finite element spaces associated to the triangulation $\Omega_h = \{K\}$ of the domain Ω of shape-regular tetrahedra K. We set

$$\boldsymbol{V}_h := \{\boldsymbol{v} \in \boldsymbol{L}^2(\Omega) \,:\, \boldsymbol{v}|_K \in \boldsymbol{\mathcal{P}}^k(K) \quad \forall K \in \Omega_h\}, \tag{2.2}$$

$$W_h := \{\omega \in L^2(\Omega) \,:\, \omega|_K \in \mathcal{P}^k(K) \quad \forall K \in \Omega_h\}, \tag{2.3}$$

where $\mathcal{P}^k(K)$ is the space of polynomial functions of degree at most $k \geq 1$ on K, and $\boldsymbol{\mathcal{P}}^k(K) = [\mathcal{P}^k(K)]^d$.

Next, we obtain the weak formulation of the exact solution. Multiplying the first two equations of (1.1) by test functions $\boldsymbol{v}$ and ω, respectively, integrating formally on each elements K, and adding over all $K \in \Omega_h$, we get the weak formulation that is satisfied by the exact solution $(\boldsymbol{q}, u)$:

$$(\boldsymbol{c}\,\boldsymbol{q}, \boldsymbol{v})_{\Omega_h} - (u, \nabla\cdot\boldsymbol{v})_{\Omega_h} + \langle u, \boldsymbol{v}\cdot\boldsymbol{n}\rangle_{\partial\Omega_h} = 0, \tag{2.4a}$$

$$-(\boldsymbol{q} + \mathbf{v}\,u, \nabla\omega)_{\Omega_h} + \langle(\boldsymbol{q} + \mathbf{v}\,u)\cdot\boldsymbol{n}, \omega\rangle_{\partial\Omega_h} = (f, \omega)_{\Omega_h}, \tag{2.4b}$$

for all $(\boldsymbol{v}, \omega) \in \boldsymbol{H}^1(\Omega_h) \times H^1(\Omega_h)$. Here $\boldsymbol{n}$ is the outward normal unit vector to ∂K, $\partial\Omega_h := \{\partial K : K \in \Omega_h\}$, and we have used the notation

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

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

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

The approximate solution $(\boldsymbol{q}_h, u_h)$ given by the MD-LDG method is defined by a discrete version of the mixed formulation (2.4). It is defined as the only element of $\boldsymbol{V}_h \times W_h$ satisfying

$$(\boldsymbol{c}\,\boldsymbol{q}_h, \boldsymbol{v})_{\Omega_h} - (u_h, \nabla\cdot\boldsymbol{v})_{\Omega_h} + \langle\widehat{u}_h^{\mathbf{v}_0}, \boldsymbol{v}\cdot\boldsymbol{n}\rangle_{\partial\Omega_h} = 0, \tag{2.5a}$$

$$-(\boldsymbol{q}_h + \mathbf{v}\,u_h, \nabla\omega)_{\Omega_h} + \langle(\widehat{\boldsymbol{q}}_h + \mathbf{v}\,\widehat{u}_h^{\mathbf{v}})\cdot\boldsymbol{n}, \omega\rangle_{\partial\Omega_h} = (f, \omega)_{\Omega_h}, \tag{2.5b}$$

for all $(\boldsymbol{v}, \omega) \in \boldsymbol{V}_h \times W_h$. Here $\widehat{u}_h^{\mathbf{v}_0}$ and $\widehat{\boldsymbol{q}}_h$ are the numerical traces associated with diffusion, and $\widehat{u}_h^{\mathbf{v}}$ is the numerical trace associated with convection. To complete the definition of the method, we need to define these numerical traces.

To do that, we need to introduce some notation. We denote by $\mathscr{E}_h^i$ all the interior faces, and Γ all the boundary faces. 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 $e \in \Gamma$ if there is a simplex in Ω_h such that $e = \partial K \cap \partial\Omega$. We set $\mathscr{E}_h := \mathscr{E}_h^i \cup \Gamma$. Now let e be an interior face shared by elements K_1 and K_2, and define the unit normal vectors $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$ on e pointing exterior to K_1 and K_2, respectively. The *average* and the *jump* of a scalar-valued function ζ on $e \in \mathscr{E}_h^i$ are given by

$$\{\!\{\zeta\}\!\} := \frac{1}{2}(\zeta_1 + \zeta_2), \qquad [\![\zeta\, \boldsymbol{n}]\!] := \zeta_1 \boldsymbol{n}_1 + \zeta_2 \boldsymbol{n}_2,$$

where $\zeta_i := \zeta|_{\partial K_i}$. For a vector-valued function $\boldsymbol{\sigma}$, we define $\boldsymbol{\sigma}_1$ and $\boldsymbol{\sigma}_2$ analogously and set

$$\{\!\{\boldsymbol{\sigma}\}\!\} := \frac{1}{2}(\boldsymbol{\sigma}_1 + \boldsymbol{\sigma}_2), \qquad [\![\boldsymbol{\sigma} \cdot \boldsymbol{n}]\!] := \boldsymbol{\sigma}_1 \cdot \boldsymbol{n}_1 + \boldsymbol{\sigma}_2 \cdot \boldsymbol{n}_2 \qquad \text{on } e \in \mathscr{E}_h^i.$$

For $e \in \Gamma$, each ζ and $\boldsymbol{\sigma}$ has a uniquely defined restriction on e; we set

$$[\![\zeta\, \boldsymbol{n}]\!] = \zeta \boldsymbol{n}_e, \qquad \{\!\{\boldsymbol{\sigma}\}\!\} = \boldsymbol{\sigma} \qquad \text{on } e \in \Gamma.$$

We don't require either of the quantities $\{\!\{\zeta\}\!\}$ or $[\![\boldsymbol{\sigma} \cdot \boldsymbol{n}]\!]$ on boundary faces, and leave them undefined.

The numerical trace for the potential associated with the convective term, $\widehat{u}_h^{\mathbf{v}}$, is nothing but the classical upwinding trace. In our notation, it can be expressed as

$$\widehat{u}_h^{\mathbf{v}} = \begin{cases} \{\!\{u_h\}\!\} + \boldsymbol{\beta} \cdot [\![u_h\, \boldsymbol{n}]\!] & \text{if } e \in \mathscr{E}_h^i, \\ u_h & \text{if } e \in \Gamma_{\mathbf{v}}^+, \\ g & \text{if } e \in \Gamma_{\mathbf{v}}^-, \end{cases} \tag{2.6a}$$

where $\boldsymbol{\beta}$ is any function on $\mathscr{E}_h$ such that, for $\boldsymbol{x} \in \partial K \cap \mathscr{E}_h$,

$$\boldsymbol{\beta} \cdot \boldsymbol{n}_K(\boldsymbol{x}) = \frac{1}{2}\operatorname{sign}\big(\mathbf{v}(\boldsymbol{x}) \cdot \boldsymbol{n}_K(\boldsymbol{x})\big), \tag{2.6b}$$

where $\boldsymbol{n}_K(\boldsymbol{x})$ is the outward unit normal of K at $\boldsymbol{x}$, and

$$\Gamma_{\mathbf{v}}^- := \{e \in \Gamma \,:\, \mathbf{v} \cdot \boldsymbol{n}_e < 0\}, \qquad \Gamma_{\mathbf{v}}^+ := \Gamma \setminus \Gamma_{\mathbf{v}}^-. \tag{2.6c}$$

Here, $\boldsymbol{n}_e$ is the outward unit normal of e. As pointed out in the introduction, we are assuming that $\Gamma_{\mathbf{v}}^- \subset \partial\Omega_D$. We also assume that our triangulation Ω_h is such that if the face $e \in \Gamma$, then $\mathbf{v} \cdot \boldsymbol{n}_e$ does not change sign on e.

The numerical trace for the potential associated with the diffusive term, $\widehat{u}_h^{\mathbf{v}_0}$, has a similar definition, namely,

$$\widehat{u}_h^{\mathbf{v}_0} = \begin{cases} \{\!\{u_h\}\!\} + \boldsymbol{\beta}_0 \cdot [\![u_h\, \boldsymbol{n}]\!] & \text{if } e \in \mathscr{E}_h^i, \\ u_h & \text{if } e \in \partial\Omega_N, \\ g & \text{if } e \in \partial\Omega_D, \end{cases} \tag{2.7a}$$

where $\boldsymbol{\beta}_0$ is any function on $\mathscr{E}_h$ such that, for $e \in \partial K \cap \mathscr{E}_h$,

$$\boldsymbol{\beta}_0 \cdot \boldsymbol{n}_K(e) = \frac{1}{2}\operatorname{sign}\big(\mathbf{v}_0 \cdot \boldsymbol{n}_K(e)\big), \tag{2.7b}$$

where $\mathbf{v}_0$ is *any* nonzero piecewise constant vector in $\boldsymbol{H}(div, \Omega)$ and

$$\Gamma_{\mathbf{v}_0}^- = \{e \in \Gamma \,:\, \mathbf{v}_0 \cdot \boldsymbol{n}_e < 0\}, \qquad \Gamma_{\mathbf{v}_0}^+ = \Gamma \setminus \Gamma_{\mathbf{v}_0}^- \tag{2.7c}$$

Finally, the numerical trace for the flux is given by

$$\widehat{\boldsymbol{q}}_h = \begin{cases} \{\!\{\boldsymbol{q}_h\}\!\} - \boldsymbol{\beta}_0 \,[\![\boldsymbol{q}_h \cdot \boldsymbol{n}]\!] & \text{if } e \in \mathscr{E}_h^i, \\ \boldsymbol{q}_h, & \text{if } e \in \partial\Omega_D \cap \Gamma^-_{\mathbf{v}_0}, \\ \boldsymbol{q}_h + \alpha(u_h - g)\boldsymbol{n} & \text{if } e \in \partial\Omega_D \cap \Gamma^+_{\mathbf{v}_0}, \\ \mathsf{q}_\mathsf{N}\boldsymbol{n} & \text{if } e \in \partial\Omega_N, \end{cases} \tag{2.8}$$

where α is a positive parameter.

Note that the auxiliary function $\mathbf{v}_0$ does not need to be associated with the convective velocity $\mathbf{v}$; it can be chosen completely independently without changing the convergence properties of the method. However, it is natural to link these two functions (whenever $\mathbf{v}$ is not identically equal to zero in any domain containing an element $K \in \Omega_h$) by taking, for example,

$$\mathbf{v}_0 = \boldsymbol{\Pi}_0^{\mathrm{RT}}\mathbf{v},$$

where $\boldsymbol{\Pi}_0^{\mathrm{RT}}|_K$ is the Raviart-Thomas projection onto $\boldsymbol{\mathcal{P}}^0(K) \oplus \boldsymbol{x}\,\mathcal{P}^0(K)$; see, [6]. Note that $\boldsymbol{\Pi}_0^{\mathrm{RT}}\mathbf{v}$ is a constant vector on each element K given that $\mathbf{v}$ is divergence-free.

This completes the definition of the MD-LDG method. Let us point out that the numerical trace for the flux in the LDG methods considered in [10] and later in [2], is of the form

$$\widehat{\boldsymbol{q}}_h = \begin{cases} \{\!\{\boldsymbol{q}_h\}\!\} - \mathbf{b}\,[\![\boldsymbol{q}_h \cdot \boldsymbol{n}]\!] + \alpha\,[\![u_h\,\boldsymbol{n}]\!] & \text{if } e \in \mathscr{E}_h^i, \\ \boldsymbol{q}_h + \alpha(u_h - g)\boldsymbol{n} & \text{if } e \in \partial\Omega_D, \\ \mathsf{q}_\mathsf{N}\boldsymbol{n} & \text{if } e \in \partial\Omega_N, \end{cases}.$$

Since α is taken to be different from zero in all the faces in $\mathscr{E}_h^i \cup \partial\Omega_D$, each of the jumps in the approximate potential induce a loss of energy. In the the MD-LDG method, this dissipative effect is minimized by setting the stabilization parameter α to zero on all $\mathscr{E}_h$ except on $\partial\Omega_D \cap \Gamma^+_{\mathbf{v}_0}$, where its positivity is essential to guarantee that the method is well defined. Of course, if $\partial\Omega_D \cap \Gamma^+_{\mathbf{v}_0} = \emptyset$, the stabilization parameter α is identically zero in $\mathscr{E}_h$!

2.2. **A priori error estimates.** First we give a theorem which guarantees the existence and uniqueness of the solution defined by the MD-LDG method.

Theorem 2.1. *The MD-LDG method defined by the weak formulation* (2.5) *and the numerical traces* (2.6), (2.7), *and* (2.8) *has a unique solution.*

Now we introduce new notation and state our approximation results. We let $H^\ell(\Omega_h)$ be the space of functions on Ω whose restriction to each element K belongs to the Sobolev space $H^\ell(K)$, and set $\boldsymbol{H}^\ell(\Omega_h) := [H^\ell(\Omega_h)]^d$. Similarly, we define $H^\ell(\partial\Omega_h)$ and $\boldsymbol{H}^\ell(\partial\Omega_h)$. For any real-valued function ζ in $H^l(\Omega_h)$, we set

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

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

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

For each $K \in \Omega_h$, we denote by h_K the diameter of K and we set $h := \max_{K\in\Omega_h} h_K$. We can now state our results.

We begin by measuring the error in the approximation of the flux $\boldsymbol{q}$ in the norm

$$\|\,\boldsymbol{\sigma}\,\|_{L^2(\Omega_h;\boldsymbol{c})} = (\boldsymbol{c}\,\boldsymbol{\sigma},\boldsymbol{\sigma})^{1/2}_{\Omega_h}.$$

Theorem 2.2. *Suppose that the exact solution $(\boldsymbol{q}, u)$ of (1.1) belongs to $\boldsymbol{H}^r(\Omega_h) \times H^{r+1}(\Omega_h)$ for some $r \in [1, k]$. Let $(\boldsymbol{q}_h, u_h) \in \boldsymbol{V}_h \times W_h$ be the approximate solution given by the MD-LDG method with $\alpha = O(h^{-1})$, then we have*

$$\|\boldsymbol{q} - \boldsymbol{q}_h\|_{\boldsymbol{L}^2(\Omega_h;\boldsymbol{c})} \le \mathsf{C}_1(\boldsymbol{q}, u)\, h^r,$$

where

$$\mathsf{C}_1(\boldsymbol{q}, u) = C\,[(1+\gamma)|u|_{H^{r+1}(\Omega_h)} + (1+\eta)|\boldsymbol{q}|_{\boldsymbol{H}^r(\Omega_h)}]$$

and C is a constant independent of h.

In the one-dimensional case, the error estimate of $\boldsymbol{q}$ given by MD-LDG method was shown to be of order $k+1$ which is optimal; see [12] and [11]. However, the best possible order of convergence of $\boldsymbol{q}$ given by the above result is only of order k. Our numerical results confirm that this order is actually sharp.

Next, we give an a priori estimate of the error in the approximation of the potential u.

Theorem 2.3. *Suppose that Ω is convex and that the exact solution $(\boldsymbol{q}, u)$ of (1.1) belongs to $\boldsymbol{H}^r(\Omega_h) \times H^{r+1}(\Omega_h)$ for some $r \in [1, k]$. Let $(\boldsymbol{q}_h, u_h) \in \boldsymbol{V}_h \times W_h$ be the approximate solution given by the MD-LDG method with $\alpha = O(h^{-1})$, then we have*

$$\|e_u\|_{L^2(\Omega_h)} \le \mathsf{C}_2(\boldsymbol{q}, u)\, h^{r+1},$$

where

$$\mathsf{C}_2(\boldsymbol{q}, u) = (1+\gamma+\eta)\mathsf{C}_1(\boldsymbol{q}, u)$$

Note that the hypothesis of convexity of the polyhedral domain is needed in order to be able to use elliptic regularity results for the problem under consideration. Note also that the above results do hold in the purely elliptic case, that is, when the convective velocity $\mathbf{v}$ is identically equal to zero.

Finally, let us emphasize that the elimination of dissipativity effects, as implemented in the multidimensional MD-LDG method, does not lead to an improvement of the order of convergence of the flux, as it did in the one-dimensional case. Indeed, for the standard LDG methods, for which the stabilization parameter α is $O(h^{-1})$ on all faces in $\mathscr{E}_h \setminus \partial\Omega_N$, the order of convergence of the approximations for the potential and the flux using polynomials of degree k is $(k+1)$ and k, respectively. For the MD-LDG method, if we choose the stabilization parameter α to be $O(h^{-1})$ only on $\partial\Omega_D \cap \Gamma^+_{\mathbf{v}_0}$ and zero elsewhere, the orders of convergence remain unchanged.

2.3. **The condition number of the Schur complement matrix.** When the convective velocity $\mathbf{v}$ is zero, it is easy to see that the matrix equations associated with the formulation (2.5) are of the form

$$\begin{pmatrix} M & B \\ -B^t & C \end{pmatrix} \begin{pmatrix} [\boldsymbol{q}_h] \\ [u_h] \end{pmatrix} = \begin{pmatrix} E \\ F \end{pmatrix}$$

where $[\boldsymbol{q}_h]$ and $[u_h]$ are the degrees of freedom of $\boldsymbol{q}_h$ and u_h, respectively. Since the method under consideration is an LDG method, the matrix M is block diagonal and the unknown $[\boldsymbol{q}_h]$ can be easily eliminated from the equations to obtain

$$(B^t M^{-1} B + C)\,[u_h] = F - B_2 M^{-1} F. \tag{2.9}$$

The Schur-complement matrix $B^t M^{-1} B + C$ can be easily seen to be symmetric and positive definite. An upper bound for its condition number κ is given in the following result.

Theorem 2.4. *Assume that the triangulation $\Omega_h = \{K\}$ of shape-regular tetrahedra K is quasi-uniform. Then, if α is of order h^{-1},*

$$\kappa \le C\,h^{-2},$$

where C is independent of h.

Let us compare our estimate with a similar result obtained in [9]. Therein it was proved that the condition number for the Schur-complement matrix of the LDG method is bounded by a quantity of order

$$\left(\alpha\,h + \frac{1}{\alpha\,h}\right)\,h^{-2}.$$

The sharpness of the above bound was suggested by strong numerically evidence, see Figs. 3 and 4 in [9]. Note, however, that to deduce this upper bound, the stabilization parameter α was taken to be constant over all the faces of the triangulation. This is not true in our case since we have taken α to be constant over $\partial\Omega_D \cap \Gamma^+_{\mathbf{v}_0}$. Another difference is that in the above mentioned numerical results, the parameter $\boldsymbol{\beta}$ of the LDG method was taken to be equal to zero [8]. This is not true in our case since $\boldsymbol{\beta}$ is given by (2.6b).

3. Proofs

3.1. **A key projection.** We begin by introducing a projection which plays an important role in our analysis. For any 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 to be an element of $\boldsymbol{\mathcal{P}}^k(K)$ that satisfies

$$(\boldsymbol{\Pi}\boldsymbol{\sigma} - \boldsymbol{\sigma}, \boldsymbol{v})_K = 0, \qquad \forall\, \boldsymbol{v} \in \boldsymbol{\mathcal{P}}^{k-1}(K), \tag{3.10a}$$

$$\langle (\boldsymbol{\Pi}\boldsymbol{\sigma} - \boldsymbol{\sigma}) \cdot \boldsymbol{n}_i, \omega \rangle_{e_i} = 0, \qquad \forall\, \omega \in \mathcal{P}^k(e_i),\ i = 1, ..., d \tag{3.10b}$$

where e_i is a face of K, and $\boldsymbol{n}_i$ is the outward normal unit vector of $e_i, i = 1, ..., d$.

Let us recall that the Raviart-Thomas projection $\boldsymbol{\Pi}^{\rm RT}\boldsymbol{\sigma}$ restricted to the simplex K is defined as the element of $\boldsymbol{\mathcal{P}}^k(K) \oplus \boldsymbol{x}\,\mathcal{P}(K)^k$ that satisfies

$$(\boldsymbol{\Pi}^{\rm RT}\boldsymbol{\sigma} - \boldsymbol{\sigma}, \boldsymbol{v})_K = 0, \qquad \forall\, \boldsymbol{v} \in \boldsymbol{\mathcal{P}}^{k-1}(K), \tag{3.11a}$$

$$\langle (\boldsymbol{\Pi}^{\rm RT}\boldsymbol{\sigma} - \boldsymbol{\sigma}) \cdot \boldsymbol{n}, \omega \rangle_e = 0, \qquad \forall\, \omega \in \mathcal{P}^k(e), \text{ for all faces e of } K. \tag{3.11b}$$

for any given function $\boldsymbol{\sigma} \in \boldsymbol{H}^1(\Omega_h)$ and an arbitrary simplex $K \in \Omega_h$. It is thus easy to see that the new projection $\boldsymbol{\Pi}$ is defined in a similar way to the Raviart-Thomas projection $\boldsymbol{\Pi}^{\rm RT}$. The difference between the two projections is that the local space of the projection $\boldsymbol{\Pi}$ has less degrees of freedom and, accordingly, less constraints on ∂K.

The projection $\boldsymbol{\Pi}$ is a well defined operator, as the next result states.

Lemma 3.1. *The projection $\boldsymbol{\Pi}$ is well-defined.*

To prove this lemma, we are going to rely on the following auxiliary result.

Lemma 3.2. *Given the faces $e_1, ..., e_d$ of the simplex K and functions $\boldsymbol{\sigma} \in \boldsymbol{L}^2(K)$ and $\zeta_i \in L^2(e_i), i = 1, ..., d$, there is a unique function $\boldsymbol{Z} \in \boldsymbol{\mathcal{P}}^k(K)$ such that,*

$$(\boldsymbol{Z} - \boldsymbol{\sigma}, \boldsymbol{v})_K = 0, \qquad \forall\, \boldsymbol{v} \in \boldsymbol{\mathcal{P}}^{k-1}(K), \tag{3.12a}$$

$$\langle \boldsymbol{Z} \cdot \boldsymbol{n}_i - \zeta_i, \omega\rangle_{e_i} = 0, \qquad \forall\, \omega \in \mathcal{P}^k(e_i),\ i = 1, ..., d. \tag{3.12b}$$

where $\boldsymbol{n}_i$ is the outward normal unit vector of e_i. Moreover,

$$\|\boldsymbol{Z}\|_{\boldsymbol{L}^2(K)} \le C(\|\boldsymbol{\sigma}\|_{\boldsymbol{L}^2(K)} + h_K^{1/2} \sum_{i=1}^{d} \|\zeta_i\|_{L^2(e_i)}) \tag{3.13}$$

where C depends only on d, k and the shape regular constant.

Proof of Lemma 3.1. If we set $\zeta_i := \boldsymbol{\sigma} \cdot \boldsymbol{n}|_{e_i}, i = 1, ..., d$, we can see that the Lemma immediately follows from Lemma 3.2. □

The following lemma shows that the projection $\boldsymbol{\Pi}$ has optimal approximation order.

Lemma 3.3. *For any $\boldsymbol{\sigma} \in \boldsymbol{H}^r(K)$, $1 \le r \le k+1, 0 \le s \le r$, we have*

$$\|\boldsymbol{\Pi\sigma} - \boldsymbol{\sigma}\|_{\boldsymbol{H}^s(K)} \le C h_K^{r-s} |\boldsymbol{\sigma}|_{\boldsymbol{H}^r(K)}.$$

Proof. The estimate follows from the fact that $\boldsymbol{\Pi}$ leaves invariant functions in $\boldsymbol{\mathcal{P}}^k(K)$ and from (3.13). □

It remains to prove Lemma 3.2.

Proof. Since the linear system determined by equations (3.12a) and (3.12b) is square, indeed, since

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

$$\sum_{i=1}^{d} \dim(\mathcal{P}^k(e_i)) = \binom{k+d-1}{d-1} \times d,$$

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

and

$$\binom{k-1+d}{d} + \binom{k+d-1}{d-1} = \binom{k+d}{d},$$

we only need to show that if $\boldsymbol{Z} \in \boldsymbol{\mathcal{P}}^k(K)$ satisfies

$$\begin{aligned} (\boldsymbol{Z}, \boldsymbol{v})_K &= 0, \qquad \forall\, \boldsymbol{v} \in \boldsymbol{\mathcal{P}}^{k-1}(K), \\ \langle \boldsymbol{Z} \cdot \boldsymbol{n}_i, \omega\rangle_{e_i} &= 0, \qquad \forall\, \omega \in \mathcal{P}^k(e_i),\ i = 1, 2, ..., d, \end{aligned} \tag{3.14}$$

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

Let T be an affine transformation that transforms the element K to the reference simplex $\tilde{K}$. We denote by $e_i, i = 1, 2, ..., d$, any d faces of K and assume that the transformation T is such that $\tilde{e}_i = T(e_i)$ is the face of $\tilde{K}$ lying on the plane $\tilde{x}_i = 0$. The outward normal unit vectors $\boldsymbol{n}_i$ of e_i are linearly independent. We let $\tilde{\boldsymbol{n}}_i, i = 1, ..., d$, be the unit vectors so that

$$\tilde{\boldsymbol{n}}_i \cdot \boldsymbol{n}_j = \delta_{ij}$$

Suppose $\boldsymbol{Z}$ is transformed to $\widetilde{\boldsymbol{Z}}$ which can be decomposed as

$$\widetilde{\boldsymbol{Z}} = \sum_{i=1}^{d} \tilde{p}_i \tilde{\boldsymbol{n}}_i, \qquad \text{where} \quad \tilde{p}_i \in \mathcal{P}^k(\tilde{K}).$$

From (3.14) we get

$$(\sum_{i=1}^{d} \tilde{p}_i \tilde{\boldsymbol{n}}_i, \tilde{\boldsymbol{v}})_{\tilde{K}} = 0, \qquad \forall\, \tilde{\boldsymbol{v}} \in \boldsymbol{\mathcal{P}}^{k-1}(\tilde{K}), \tag{3.15a}$$

$$\langle \sum_{i=1}^{d} \tilde{p}_i \tilde{\boldsymbol{n}}_i \cdot \boldsymbol{n}_j, \tilde{\omega} \rangle_{\tilde{e}_j} = 0, \qquad \forall\, \tilde{\omega} \in \mathcal{P}^k(\tilde{e}_j),\ i = 1, ..., d. \tag{3.15b}$$

From (3.15b) we have $\tilde{p}_j|_{\tilde{e}_j} = 0$ for $j = 1, ..., d$. Thus, $\tilde{p}_j = \tilde{x}_j \tilde{\mathsf{p}}_j$ for some polynomial $\tilde{\mathsf{p}}_j \in \mathcal{P}^{k-1}(\tilde{K})$.

Taking $\boldsymbol{v} = \tilde{\mathsf{p}}_j\, \boldsymbol{n}_j$ in (3.15a), we get

$$\int_{\tilde{K}} \tilde{p}_j \tilde{\mathsf{p}}_j dx = \int_{\tilde{K}} \tilde{x}_j \tilde{\mathsf{p}}_j^2 dx = 0. \tag{3.16}$$

Since $\tilde{x}_j > 0$ on $\tilde{K}$, we conclude that $\tilde{\mathsf{p}}_j = 0, j = 1, ..., d$, which implies that $\boldsymbol{Z} \equiv \boldsymbol{0}$.

The estimate (3.13) following from a simple scaling argument. This completes the proof of the lemma. □

3.2. **The MD-LDG is well defined: Proof of Theorem 2.1.** Due to the linearity of the problem it is enough to show that, when $f = 0$, $g = 0$ and $\mathsf{q_N} = 0$, the only solution $(\boldsymbol{q}_h, u_h) \in \boldsymbol{V}_h \times W_h$ of (2.5) with the numerical traces defined by (2.6), (2.7), and (2.8), is the trivial solution.

Taking $(\boldsymbol{v}, w) := (\boldsymbol{q}_h, u_h)$ in (2.5), adding the equations and performing some simple algebraic manipulations, see [10], we easily get that

$$(\boldsymbol{c}\boldsymbol{q}_h, \boldsymbol{q}_h)_{\Omega_h} + \frac{1}{2}\langle |\mathbf{v}\cdot\boldsymbol{n}|, [\![u_h \boldsymbol{n}]\!]^2\rangle_{\mathscr{E}_h} + \langle \alpha, u_h^2\rangle_{\partial\Omega_D \cap \Gamma^+_{\mathbf{v}_0}} = 0.$$

Since $\boldsymbol{c}$ is uniformly positive definite and α is positive, we conclude that $\boldsymbol{q}_h = \boldsymbol{0}$ and $u_h|_{\partial\Omega_D \cap \Gamma^+_{\mathbf{v}_0}} = 0$.

Since $\boldsymbol{q}_h = \boldsymbol{0}$, we can rewrite the equation (2.5a) as

$$-(u_h, \nabla\cdot\boldsymbol{v})_{\Omega_h} + \langle \widehat{u}_h^{\mathbf{v}_0},\, \boldsymbol{v}\cdot\boldsymbol{n}\rangle_{\partial\Omega_h} = 0 \qquad \forall \boldsymbol{v} \in \boldsymbol{V}_h,$$

or, after a simple integration by parts, as

$$(\nabla u_h, \boldsymbol{v})_{\Omega_h} - \langle u_h - \widehat{u}_h^{\mathbf{v}_0},\, \boldsymbol{v}\cdot\boldsymbol{n}\rangle_{\partial\Omega_h} = 0 \ \ \forall \boldsymbol{v} \in \boldsymbol{V}_h.$$

By the definition of the numerical trace $\widehat{u}_h^{\mathbf{v}_0}$, (2.7a), and since we are taking $g = 0$, we have

$$(u_h - \widehat{u}_h^{\mathbf{v}_0})|_{\partial K} = \begin{cases} (\frac{1}{2} - \boldsymbol{\beta}_0\cdot\boldsymbol{n})\ [\![u_h\,\boldsymbol{n}]\!]\cdot\boldsymbol{n} & \text{on } \partial K \cap \mathscr{E}_h^i, \\ 0 & \text{on } \partial K \cap \partial\Omega_N, \\ u_h & \text{on } \partial K \cap \partial\Omega_D, \end{cases}$$

and since $u_h|_{\partial\Omega_D \cap \Gamma^+_{\mathbf{v}_0}} = 0$,

$$(u_h - \widehat{u}_h^{\mathbf{v}_0})|_{\partial K} = \begin{cases} (\frac{1}{2} - \boldsymbol{\beta}_0\cdot\boldsymbol{n})\ [\![u_h\,\boldsymbol{n}]\!]\cdot\boldsymbol{n} & \text{on } \partial K \setminus \partial\Omega_N, \\ 0 & \text{on } \partial K \cap \partial\Omega_N. \end{cases} \tag{3.17}$$

This implies that

$$(3.18)\qquad (\nabla u_h, \boldsymbol{v})_{\Omega_h} - \langle (\frac{1}{2} - \boldsymbol{\beta}_0 \cdot \boldsymbol{n})\ [\![u_h\, \boldsymbol{n}]\!] \cdot \boldsymbol{n},\, \boldsymbol{v} \cdot \boldsymbol{n} \rangle_{\partial\Omega_h \setminus \partial\Omega_N} = 0\ \ \forall \boldsymbol{v} \in \boldsymbol{V}_h.$$

we only have to find a suitably chosen test function $\boldsymbol{v} \in \boldsymbol{V}_h$ to conclude the proof. Such a choice is based on the observation that, since $\mathbf{v}_0$ is divergence-free, on each $K \in \Omega_h$, there is at least one face e of K on which $(\frac{1}{2} - \boldsymbol{\beta}_0 \cdot \boldsymbol{n}_e) = 0$. Let us denote by e_K one of those faces. Then, on the simplex $K \in \Omega_h$, we take $\boldsymbol{v} := \boldsymbol{Z}$ given by Lemma 3.2 with

$$\boldsymbol{\sigma} := \nabla u_h \quad \text{ and } \quad \zeta_i := -(\frac{1}{2} - \boldsymbol{\beta}_0 \cdot \boldsymbol{n}_{e_i})\ [\![u_h\, \boldsymbol{n}]\!] \cdot \boldsymbol{n}_{e_i},$$

where $e_i \neq e_K$ for $i = 1, \ldots, d$, to obtain that

$$(\nabla u_h, \nabla u_h)_{\Omega_h} + \langle 1, (\frac{1}{2} - \boldsymbol{\beta}_0 \cdot \boldsymbol{n})^2\ [\![u_h\, \boldsymbol{n}]\!]^2 \rangle_{\partial\Omega_h \setminus \partial\Omega_N} = 0.$$

This implies that, on each simplex $K \in \Omega_h$, u_h is a constant and that

$$[\![u_h\, \boldsymbol{n}]\!]\,|_{\mathscr{E}_h^i} = \boldsymbol{0} \quad \text{ and } u_h|_{\partial\Omega_D \cap \Gamma^-_{\mathbf{v}_0}} = 0.$$

Since $u_h|_{\partial\Omega_D \cap \Gamma^+_{\mathbf{v}_0}} = 0$, we have that $u_h|_{\Omega_h}$ is a constant and $u_h|_{\partial\Omega_D} = 0$, and hence that u_h is equal to zero. This completes the proof of Theorem 2.1.

3.3. **The error estimates: Proof of Theorem 2.2 and 2.3.**

3.3.1. *Preliminaries.* In what follows, we are going to use the following inverse and trace inequalities.

Lemma 3.4. *For any $\boldsymbol{\sigma} \in \boldsymbol{\mathcal{P}}^k(K)$ and $\zeta \in \mathcal{P}^k(K)$ there exist positive constants C_1 and C_2 such that*

$$(3.19a)\qquad \|\boldsymbol{\sigma} \cdot \boldsymbol{n}\|^2_{L^2(e)} \le C_1 h_K^{-1} \|\boldsymbol{\sigma}\|^2_{\boldsymbol{L}^2(K)},$$

$$(3.19b)\qquad \|\zeta\|^2_{L^2(e)} \le C_1 h_K^{-1} \|\zeta\|^2_{L^2(K)},$$

$$(3.19c)\qquad \|\nabla\zeta\|^2_{L^2(K)} \le C_2 h_K^{-2} \|\zeta\|^2_{L^2(K)};$$

where e is an face of ∂K, and C_1 and C_2 depend only on k and the shape regularity constant of the triangulation.

Lemma 3.5. *For any $\boldsymbol{\psi} \in \boldsymbol{H}^1(K)$ and $\varphi \in H^1(K)$ there exist positive constants C_0 such that*

$$(3.20a)\qquad \|\boldsymbol{\psi} \cdot \boldsymbol{n}\|^2_{L^2(e)} \le C_0 \|\boldsymbol{\psi}\|_{\boldsymbol{L}^2(K)} \|\boldsymbol{\psi}\|_{\boldsymbol{H}^1(K)},$$

$$(3.20b)\qquad \|\varphi\|^2_{L^2(e)} \le C_0 \|\varphi\|_{L^2(K)} \|\varphi\|_{H^1(K)},$$

where e is an face of ∂K, and C_0 is independent of mesh size h.

To prove Theorem 2.2 and 2.3, we proceed as follows. We first get an intermediate error estimate in the approximation of the flux $\boldsymbol{q}$ which involves the error in the approximation of the potential u. Then we use a duality argument to obtain the error estimate of u in terms of the estimate of the error in the flux; we then insert the intermediate error estimate obtained in the first step and obtain the error estimate in the approximation of the potential. Finally, we insert the error estimate of u into the intermediate estimate and get the final error estimate of $\boldsymbol{q}$.

To facilitate the analysis of the method, especially in the first step, we rewrite it in compact form as follows. From the weak formulation (2.5a) and (2.5b), we get

$$B(\boldsymbol{q}_h, u_h; \boldsymbol{v}, \omega) = \mathcal{F}(\boldsymbol{v}, \omega) \qquad \text{for all } (\boldsymbol{v}, \omega) \in \boldsymbol{V}_h \times W_h, \tag{3.21a}$$

where

$$\begin{aligned} B(\boldsymbol{\sigma}, \zeta; \boldsymbol{v}, \omega) =& (\boldsymbol{c}\,\boldsymbol{\sigma}, \boldsymbol{v})_{\Omega_h} - (\zeta, \nabla\cdot\boldsymbol{v})_{\Omega_h} + \langle \widehat{\zeta}^{\mathsf{v}_0}, \boldsymbol{v}\cdot\boldsymbol{n}\rangle_{\partial\Omega_h\setminus\partial\Omega_D} \\ & - (\boldsymbol{\sigma} + \mathbf{v}\,\zeta, \nabla\omega)_{\Omega_h} + \langle \widehat{\boldsymbol{\sigma}}\cdot\boldsymbol{n}, \omega\rangle_{\partial\Omega_h\setminus\Gamma} + \langle \boldsymbol{\sigma}, \omega\boldsymbol{n}\rangle_{\partial\Omega_D} \\ & + \langle \mathbf{v}\cdot\boldsymbol{n}\,\widehat{\zeta}^{\mathbf{v}}, \omega\rangle_{\partial\Omega_h\setminus\Gamma^-_{\mathbf{v}}} + \langle \alpha\zeta, \omega\rangle_{\partial\Omega_D\cap\Gamma^+_{\mathbf{v}_0}}, \end{aligned} \tag{3.21b}$$

and

$$\begin{aligned} \mathcal{F}(\boldsymbol{v}, \omega) =& (f, \omega)_{\Omega_h} + \langle g, \boldsymbol{v}\cdot\boldsymbol{n}\rangle_{\partial\Omega_D} - \langle \mathsf{q}_N, \omega\rangle_{\partial\Omega_N} \\ & - \langle \mathbf{v}\cdot\boldsymbol{n}\,g, \omega\rangle_{\Gamma^-_{\mathbf{v}}} + \langle g, \omega\rangle_{\partial\Omega_D\cap\Gamma^+_{\mathbf{v}_0}}. \end{aligned} \tag{3.21c}$$

3.3.2. *Step 1: Intermediate estimate of the error in* $\boldsymbol{q}$. In this step, we prove an estimate of the error in the approximation to $\boldsymbol{q}$ which is expressed in terms of the error in u. To state it, we need to introduce some notation. We estimate the quantity $|||(\boldsymbol{\sigma},\zeta)|||$, where

$$|||(\boldsymbol{\sigma},\zeta)|||^2 := B(\boldsymbol{\sigma},\zeta;\boldsymbol{\sigma},\zeta) = (\boldsymbol{c}\,\boldsymbol{\sigma},\boldsymbol{\sigma})_{\Omega_h} + \frac{1}{2}\langle |\mathbf{v}\cdot\boldsymbol{n}|, [\![\zeta\boldsymbol{n}]\!]^2\rangle_{\mathscr{E}_h} + \langle \alpha, \zeta^2\rangle_{\partial\Omega_D\cap\Gamma^+_{\mathbf{v}_0}}.$$

We also need to introduce two projections. For any simplex $K \in \Omega_h$, let $\mathsf{P}|_K$ be the L^2-projection onto $\mathcal{P}^k(K)$, and let $\boldsymbol{\Pi}|_K$ be the projection defined in (3.10) with $\{e_i, i = 1, ..., d\} \supseteq \{e \in \partial K : \boldsymbol{n}_e \cdot \boldsymbol{\beta}_0(e) \le 0\}$.

Lemma 3.6. *Under the same assumption as Theorem 2.2 we have*

$$|||(\boldsymbol{\Pi q} - \boldsymbol{q}_h, \mathsf{P}u - u_h)||| \le C(h^r\Theta + \sqrt{\Phi})$$

where

$$\Theta = (\eta + \alpha^{-1/2}h^{-1/2})|\boldsymbol{q}|_{\boldsymbol{H}^r(\Omega_h)} + (\gamma + \alpha^{1/2}h^{1/2} + h^{1/2}\|\mathbf{v}\|^{1/2}_{\boldsymbol{L}^\infty(\Omega_h)})|u|_{H^{r+1}(\Omega_h)}$$

and

$$\Phi = \|\mathbf{v}\|_{\boldsymbol{L}^\infty(\Omega_h)} h^r |u|_{H^{r+1}(\Omega_h)} \|\mathsf{P}e_u\|_{L^2(\Omega_h)}.$$

Proof. Set $\boldsymbol{e_q} = \boldsymbol{q} - \boldsymbol{q}_h$ and $e_u = u - u_h$, and note that, since $\boldsymbol{\Pi q}_h = \boldsymbol{q}_h$ and $\mathsf{P}u_h = u_h$, we have $\boldsymbol{\Pi q} - \boldsymbol{q}_h = \boldsymbol{\Pi e_q}$ and $\mathsf{P}u - u_h = \mathsf{P}e_u$. Hence, by Galerkin orthogonality,

$$B(\boldsymbol{\Pi e_q}, \mathsf{P}e_u; \boldsymbol{v}, \omega) = B(\boldsymbol{\Pi q} - \boldsymbol{q}, \mathsf{P}u - u; \boldsymbol{v}, \omega) \qquad \forall(\boldsymbol{v}, \omega) \in \boldsymbol{V}_h \times W_h,$$

and so

$$\begin{aligned} |||(\boldsymbol{\Pi e_q}, \mathsf{P}e_u)|||^2 =& B(\boldsymbol{\Pi e_q}, \mathsf{P}e_u; \boldsymbol{\Pi e_q}, \mathsf{P}e_u) \\ =& B(\boldsymbol{\Pi q} - \boldsymbol{q}, \mathsf{P}u - u; \boldsymbol{\Pi e_q}, \mathsf{P}e_u) =: \sum_{i=1}^{6} T_i + \sum_{i=1}^{2} T_i^{\mathbf{v}}, \end{aligned}$$

where

$$
\begin{aligned}
T_1 =&(\boldsymbol{c}\,(\boldsymbol{\Pi q}-\boldsymbol{q}),\boldsymbol{\Pi e_q})_{\Omega_h},\\
T_2 =&-(\mathsf{P}u-u,\nabla\cdot\boldsymbol{\Pi e_q})_{\Omega_h},\\
T_3 =&\langle\widehat{\mathsf{P}u}^{\mathbf{v}_0}-u,\boldsymbol{\Pi e_q}\cdot\boldsymbol{n}\rangle_{\partial\Omega_h\setminus\partial\Omega_D},\\
T_4 =&-(\boldsymbol{\Pi q}-\boldsymbol{q},\nabla\mathsf{P}e_u)_{\Omega_h},\\
T_5 =&\langle(\widehat{\boldsymbol{\Pi q}}-\boldsymbol{q})\cdot\boldsymbol{n},\mathsf{P}e_u\rangle_{\partial\Omega_h\setminus\Gamma}+\langle(\boldsymbol{\Pi q}-\boldsymbol{q})\cdot\boldsymbol{n},\mathsf{P}e_u\rangle_{\partial\Omega_D},\\
T_6 =&\langle\alpha(\mathsf{P}u-u),\mathsf{P}e_u\rangle_{\partial\Omega_D\cap\Gamma^+_{\mathbf{v}_0}},
\end{aligned}
$$

and

$$
\begin{aligned}
T_1^{\mathbf{v}} =&-(\mathbf{v}\,(\mathsf{P}u-u),\nabla\mathsf{P}e_u)_{\Omega_h},\\
T_2^{\mathbf{v}} =&\langle\mathbf{v}\cdot\boldsymbol{n}\,(\widehat{\mathsf{P}u}^{\mathbf{v}}-u),\mathsf{P}e_u\rangle_{\partial\Omega_h\setminus\Gamma^-_{\mathbf{v}}}.
\end{aligned}
$$

Note that the terms $T_i, i=1,...,6$, are associated with diffusion, and $T_i^{\mathbf{v}}, i=1,2$, are associated with convection.

Let us estimate each of these terms. By Cauchy-Schwarz inequality and the approximation result of the projection $\boldsymbol{\Pi}$, Lemma 3.3, we have

$$
\begin{aligned}
T_1 &\le \eta\|\boldsymbol{q}-\boldsymbol{\Pi q}\|_{\boldsymbol{L}^2(\Omega_h)}\|\boldsymbol{\Pi e_q}\|_{\boldsymbol{L}^2(\Omega_h,\boldsymbol{c})}\\
&\le C\eta h^r|\boldsymbol{q}|_{\boldsymbol{H}^r(\Omega_h)}\|\boldsymbol{\Pi e_q}\|_{\boldsymbol{L}^2(\Omega_h,\boldsymbol{c})}.
\end{aligned}
$$

From the orthogonality property of the L^2-projection P, we have that

$$T_2=0.$$

Using trace inequality (3.20b) and inverse inequality (3.19a), we get

$$
\begin{aligned}
T_3 &\le C\|\mathsf{P}u-u\|^{1/2}_{L^2(\Omega_h)}\|\mathsf{P}u-u\|^{1/2}_{H^1(\Omega_h)}\gamma h^{-1/2}\|\boldsymbol{\Pi e_q}\|_{\boldsymbol{L}^2(\Omega_h;\boldsymbol{c})}\\
&\le C\gamma\, h^r|u|_{H^{r+1}(\Omega_h)}\|\boldsymbol{\Pi e_q}\|_{\boldsymbol{L}^2(\Omega_h;\boldsymbol{c})}.
\end{aligned}
$$

Note that better orders of convergence of other terms can be obtained, if we ask for more regularity of $\boldsymbol{q}$ or $\mathbf{v}$, or pick a different α. However, the best order of convergence of the term T_3 is of order at most h^r. Thus, it is this term the one preventing us from obtaining better convergence order for $\boldsymbol{q}$.

Next, from the definition of the projection $\boldsymbol{\Pi}$, (3.10a), we have that

$$T_4=0.$$

By using the trace inequality (3.20a), we have

$$
\begin{aligned}
T_6 &\le \|\mathsf{P}u-u\|_{L^2(\partial\Omega_D\cap\Gamma^+_{\mathbf{v}_0})}\ \|\alpha\mathsf{P}e_u\|_{L^2(\partial\Omega_D\cap\Gamma^+_{\mathbf{v}_0})}\\
&\le C\alpha^{1/2}h^{r+1/2}|u|_{H^{r+1}(\Omega_h)}\ \langle\alpha,(\mathsf{P}e_u)^2\rangle^{1/2}_{\partial\Omega_D\cap\Gamma^+_{\mathbf{v}_0}}.
\end{aligned}
$$

Let now us consider the term T_5. Using the definition of $\boldsymbol{\beta}_0$, (2.7b), we rewrite T_5 as

$$T_5=T_5'+\langle(\boldsymbol{\Pi q}-\boldsymbol{q})\cdot\boldsymbol{n},\mathsf{P}e_u\rangle_{\partial\Omega_D\cap\Gamma^+_{\mathbf{v}_0}},$$

where

$$\begin{aligned}T_5' =&\langle(\widehat{\mathbf{\Pi} \boldsymbol{q}}-\boldsymbol{q})\cdot\boldsymbol{n}, \mathsf{P}e_u\rangle_{\partial\Omega_h\setminus(\partial\Omega_N\cup\Gamma^+_{\mathbf{v}_0})}\\ =&\sum_{K\in\Omega_h}\sum_{\substack{e\in\partial K\setminus\partial\Omega_N\\ \boldsymbol{n}_e\cdot\boldsymbol{\beta}_0\le 0}}\langle\mathbf{\Pi}\boldsymbol{q}-\boldsymbol{q},\, [\![\mathsf{P}e_u\boldsymbol{n}]\!]\rangle_e\\ =&0,\end{aligned}$$

by the definition of the projection $\mathbf{\Pi}$, (3.10b). Let us empahsize the fact that the projection $\mathbf{\Pi}$ was constructed as to ensure the very last equality; Its definition of projection $\mathbf{\Pi}$ is thus strongly related to those of the numerical traces. So, using the trace inequality (3.20a), we obtain that

$$\begin{aligned}T_5 &= \langle(\mathbf{\Pi}\boldsymbol{q}-\boldsymbol{q})\cdot\boldsymbol{n}, \mathsf{P}e_u\rangle_{\partial\Omega_D\cap\Gamma^+_{\mathbf{v}_0}}\\ &\le \|(\mathbf{\Pi}\boldsymbol{q}-\boldsymbol{q})\cdot\boldsymbol{n}\|_{L^2(\partial\Omega_D\cap\Gamma^+_{\mathbf{v}_0})}\,\|\mathsf{P}e_u\|_{L^2(\partial\Omega_D\cap\Gamma^+_{\mathbf{v}_0})}\\ &\le C\alpha^{-1/2}h^{r-1/2}|\boldsymbol{q}|_{\boldsymbol{H}^r(\Omega_h)}\,\langle\alpha,\,(\mathsf{P}e_u)^2\,\rangle^{1/2}_{\partial\Omega_D\cap\Gamma^+_{\mathbf{v}_0}}.\end{aligned}$$

Here we want to point out that if we use L^2-projection instead of the projection $\mathbf{\Pi}$, then T_5' is no longer zero; indeed, after using a trace and an inverse inequalities, the upper bound of this term is $Ch^{r-1}\|\mathsf{P}e_u\|_{L^2(\Omega_h)}$. This would make the order of error estimate of the potential u not optimal unless we ask for additional regularity of the flux $\boldsymbol{q}$. The use of the projection $\mathbf{\Pi}$ allows us ot avoid having to do that.

We have finished estimating the terms associated with the diffusion. If there is no convection, namely, $\mathbf{v}=0$, then we get

$$|\!|\!|(\mathbf{\Pi}e_{\boldsymbol{q}}, \mathsf{P}e_u)|\!|\!|^2 \le \sum_{i=1}^{6} T_i \le Ch^r\Theta|\!|\!|(\mathbf{\Pi}\boldsymbol{e}_{\boldsymbol{q}}, \mathsf{P}e_u)|\!|\!|.$$

which implies that

$$|\!|\!|(\mathbf{\Pi}e_{\boldsymbol{q}}, \mathsf{P}e_u)|\!|\!| \le Ch^r\Theta.$$

If the convection velocity $\mathbf{v}\neq 0$, we need to estimate the terms $T_1^{\mathbf{v}}$ and $T_2^{\mathbf{v}}$. Using Cauchy-Schwarz inequality and inverse inequality (3.19c), we get that

$$\begin{aligned}T_1^{\mathbf{v}} &\le \|\mathbf{v}\|_{\boldsymbol{L}^\infty(\Omega_h)}\|\mathsf{P}u-u\|_{L^2(\Omega_h)}\|\mathsf{P}e_u\|_{H^1(\Omega_h)}\\ &\le C\|\mathbf{v}\|_{\boldsymbol{L}^\infty(\Omega_h)}h^r\|u\|_{H^{r+1}(\Omega_h)}\|\mathsf{P}e_u\|_{L^2(\Omega_h)}.\end{aligned}$$

Using Cauchy-Schwarz inequality and trace inequality (3.20b), we get

$$\begin{aligned}T_2^{\mathbf{v}} &\le C\|\mathsf{P}u-u\|_{L^2(\partial\Omega_h\setminus\Gamma^-_{\mathbf{v}})}\,\|(\mathbf{v}\cdot\boldsymbol{n})\,[\![\mathsf{P}e_u\,\boldsymbol{n}]\!]\|_{L^2(\mathscr{E}_h\setminus\Gamma^-_{\mathbf{v}})}\\ &\le C\|\mathbf{v}\|^{1/2}_{\boldsymbol{L}^\infty(\partial\Omega_h)}h^{r+1/2}|u|_{H^{r+1}(\Omega_h)}\,\langle|\mathbf{v}\cdot\boldsymbol{n}|,\,[\![\mathsf{P}e_u\,\boldsymbol{n}]\!]^2\rangle^{1/2}_{\mathscr{E}_h}.\end{aligned}$$

Hence, we have

$$|\!|\!|(\mathbf{\Pi}e_{\boldsymbol{q}}, \mathsf{P}e_u)|\!|\!|^2 \le \sum_{i=1}^{6} T_i + \sum_{i=1}^{2} T_i^{\mathbf{v}} \le C(h^r\Theta|\!|\!|(\mathbf{\Pi}\boldsymbol{e}_{\boldsymbol{q}}, \mathsf{P}e_u)|\!|\!| + \Phi).$$

where Θ and Φ are given by Theorem 2.2, and the result follows. This completes the proof of Lemma 3.6. □

3.3.3. *Step 2: Estimate of the error in u.* In this step, we estimate the error in u. Thanks to the identity

$$\|e_u\|_{L^2(\Omega_h)} = \sup_{\theta \in \mathcal{C}_0^\infty(\Omega)} \frac{(e_u, \theta)_{\Omega_h}}{\|\theta\|_{L^2(\Omega)}},$$

we only need to estimate the term $(e_u, \theta)_{\Omega_h}$. Since we do that by using a duality argument, we have to introduce the corresponding adjoint problem, namely,

$$\begin{aligned} \boldsymbol{c}\boldsymbol{\psi} + \nabla\varphi &= 0 && \text{in } \Omega,\\ \nabla \cdot (\boldsymbol{\psi} - \mathsf{v}\varphi) &= \theta && \text{in } \Omega,\\ \varphi &= 0 && \text{on } \partial\Omega_D,\\ (\boldsymbol{\psi} - \mathsf{v}\varphi) \cdot \boldsymbol{n} &= 0 && \text{on } \partial\Omega_N. \end{aligned}$$

Multiplying both sides of the first two equations by test functions $\boldsymbol{v}$ and ω respectively, and integrating by parts, we get the weak formulation

$$(\boldsymbol{c}\boldsymbol{\psi}, \boldsymbol{v})_{\Omega_h} + (\nabla\varphi, \boldsymbol{v})_{\Omega_h} = 0, \tag{3.22a}$$

$$(\nabla \cdot (\boldsymbol{\psi} - \mathsf{v}\varphi), \omega)_{\Omega_h} = (\theta, \omega)_{\Omega_h}. \tag{3.22b}$$

These equations are going to be combined with the error equations

$$(\boldsymbol{c}\boldsymbol{e_q}, \boldsymbol{v})_{\Omega_h} - (e_u, \nabla \cdot \boldsymbol{v})_{\Omega_h} + \langle u - \widehat{u}_h^{\mathsf{v}_0}, \boldsymbol{v} \cdot \boldsymbol{n}\rangle_{\partial\Omega_h} = 0, \tag{3.23a}$$

$$-(\boldsymbol{e_q} + \mathsf{v}e_u, \nabla\omega)_{\Omega_h} + \langle(\boldsymbol{q} - \widehat{\boldsymbol{q}}_h) + \mathsf{v}(u - \widehat{u}_h^{\mathsf{v}}), \omega\boldsymbol{n}\rangle_{\partial\Omega_h} = 0, \tag{3.23b}$$

where the numerical traces $\widehat{u}_h^{\mathsf{v}_0}, \widehat{u}_h^{\mathsf{v}}$ and $\widehat{\boldsymbol{q}}_h$ are defined by (2.6), (2.7), and (2.8), respectively.

Let P and $\boldsymbol{\Pi}$ be the projections used in Lemma 3.6. Taking $\omega = e_u$ in the weak formulation of the adjoint problem, (3.22b), and integrating by parts, we get

$$\begin{aligned} (\theta, e_u)_{\Omega_h} =& \big(\nabla \cdot (\boldsymbol{\psi} - \mathsf{v}\varphi), e_u\big)_{\Omega_h}\\ =& \big(\nabla \cdot (\boldsymbol{\psi} - \boldsymbol{\Pi}\boldsymbol{\psi}), e_u\big)_{\Omega_h} + (\nabla \cdot \boldsymbol{\Pi}\boldsymbol{\psi}, e_u)_{\Omega_h} - (\mathsf{v} \cdot \nabla\varphi, e_u)_{\Omega_h}\\ =& -\big(\boldsymbol{\psi} - \boldsymbol{\Pi}\boldsymbol{\psi}, \nabla e_u\big)_{\Omega_h} + \langle e_u, (\boldsymbol{\psi} - \boldsymbol{\Pi}\boldsymbol{\psi}) \cdot \boldsymbol{n}\rangle_{\partial\Omega_h}\\ &+ (\nabla \cdot \boldsymbol{\Pi}\boldsymbol{\psi}, e_u)_{\Omega_h} - (\mathsf{v} \cdot \nabla\varphi, e_u)_{\Omega_h}. \end{aligned}$$

Using the orthogonality property of the projection $\boldsymbol{\Pi}$, (3.10a),

$$\begin{aligned} (\theta, e_u)_{\Omega_h} =& -\big(\boldsymbol{\psi} - \boldsymbol{\Pi}\boldsymbol{\psi}, \nabla(u - \mathsf{P}u)\big)_{\Omega_h} + \langle e_u, (\boldsymbol{\psi} - \boldsymbol{\Pi}\boldsymbol{\psi}) \cdot \boldsymbol{n}\rangle_{\partial\Omega_h}\\ &- (\mathsf{v} \cdot \nabla\varphi, e_u)_{\Omega_h} + (\nabla \cdot \boldsymbol{\Pi}\boldsymbol{\psi}, e_u)_{\Omega_h}. \end{aligned} \tag{3.24}$$

Now let us we rewrite the last term on the right hand side. From the error equations (3.23a), we get

$$(\nabla \cdot \boldsymbol{\Pi}\boldsymbol{\psi}, e_u)_{\Omega_h} = (\boldsymbol{c}\boldsymbol{e_q}, \boldsymbol{\Pi}\boldsymbol{\psi})_{\Omega_h} + \langle u - \widehat{u}_h^{\mathsf{v}_0}, \boldsymbol{\Pi}\boldsymbol{\psi} \cdot \boldsymbol{n}\rangle_{\partial\Omega_h}.$$

Using the weak formulation of the adjoint problem, (3.22a), we get

$$\begin{aligned} (\nabla \cdot \boldsymbol{\Pi}\boldsymbol{\psi}, e_u)_{\Omega_h} =& (\boldsymbol{c}\boldsymbol{e_q}, \boldsymbol{\Pi}\boldsymbol{\psi} - \boldsymbol{\psi})_{\Omega_h} - (\boldsymbol{e_q}, \nabla(\varphi - \mathsf{P}\varphi))_{\Omega_h}\\ &- (\boldsymbol{e_q}, \nabla\mathsf{P}\varphi)_{\Omega_h} + \langle u - \widehat{u}_h^{\mathsf{v}_0}, \boldsymbol{\Pi}\boldsymbol{\psi} \cdot \boldsymbol{n}\rangle_{\partial\Omega_h}. \end{aligned}$$

Then using the error equation (3.23b), we get

$$\begin{aligned}(\nabla\cdot\boldsymbol{\Pi\psi}, e_u)_{\Omega_h} =&(\boldsymbol{c}\boldsymbol{e}_{\boldsymbol{q}}, \boldsymbol{\Pi\psi}-\boldsymbol{\psi})_{\Omega_h} - (\boldsymbol{e}_{\boldsymbol{q}}, \nabla(\varphi-\mathsf{P}\varphi))_{\Omega_h}\\ &+(\mathsf{v}e_u, \nabla\mathsf{P}\varphi)_{\Omega_h} - \langle(\boldsymbol{q}-\widehat{\boldsymbol{q}}_h)+\mathsf{v}(u-\widehat{u}_h^{\mathsf{v}}), \mathsf{P}\varphi\,\boldsymbol{n}\rangle_{\partial\Omega_h}\\ &+\langle u-\widehat{u}_h^{\mathsf{v}_0}, \boldsymbol{\Pi\psi}\cdot\boldsymbol{n}\rangle_{\partial\Omega_h}.\end{aligned}$$

Noting that

$$\langle\mathsf{v}(u-\widehat{u}_h^{\mathsf{v}}), \varphi\boldsymbol{n}\rangle_{\partial\Omega_h\setminus\partial\Omega_N} = \langle u-\widehat{u}_h^{\mathsf{v}_0}, \boldsymbol{\psi}\cdot\boldsymbol{n}\rangle_{\partial\Omega_h\setminus\partial\Omega_N} = 0$$

and

$$\langle\mathsf{v}(u-\widehat{u}_h^{\mathsf{v}}), \varphi\boldsymbol{n}\rangle_{\partial\Omega_N} - \langle u-\widehat{u}_h^{\mathsf{v}_0}, \boldsymbol{\psi}\cdot\boldsymbol{n}\rangle_{\partial\Omega_N} = 0,$$

we obtain that

$$\begin{aligned}(\nabla\cdot\boldsymbol{\Pi\psi}, e_u)_{\Omega_h} =&(\boldsymbol{c}\boldsymbol{e}_{\boldsymbol{q}}, \boldsymbol{\Pi\psi}-\boldsymbol{\psi})_{\Omega_h} - (\boldsymbol{e}_{\boldsymbol{q}}, \nabla(\varphi-\mathsf{P}\varphi))_{\Omega_h}\\ &+(\mathsf{v}e_u, \nabla\mathsf{P}\varphi)_{\Omega_h} - \langle\boldsymbol{q}-\widehat{\boldsymbol{q}}_h, \mathsf{P}\varphi\,\boldsymbol{n}\rangle_{\partial\Omega_h}\\ &+\langle\mathsf{v}\cdot\boldsymbol{n}(u-\widehat{u}_h^{\mathsf{v}}), \varphi-\mathsf{P}\varphi\rangle_{\partial\Omega_h} + \langle u-\widehat{u}_h^{\mathsf{v}_0}, (\boldsymbol{\Pi\psi}-\boldsymbol{\psi})\cdot\boldsymbol{n}\rangle_{\partial\Omega_h}.\end{aligned}$$

Hence, from above equality and (3.24) we have

$$(\theta, e_u)_{\Omega_h} = \sum_{i=1}^{4} S_i + S^{\mathsf{v}},$$

where

$$\begin{aligned}S_1 =& -\big(\boldsymbol{\psi}-\boldsymbol{\Pi\psi}, \nabla(u-\mathsf{P}u)\big)_{\Omega_h},\\ S_2 =&(\boldsymbol{c}\boldsymbol{e}_{\boldsymbol{q}}, \boldsymbol{\Pi\psi}-\boldsymbol{\psi})_{\Omega_h} - (\boldsymbol{e}_{\boldsymbol{q}}, \nabla(\varphi-\mathsf{P}\varphi))_{\Omega_h},\\ S_3 =&\langle(\boldsymbol{q}-\widehat{\boldsymbol{q}}_h), \mathsf{P}\varphi\,\boldsymbol{n}\rangle_{\partial\Omega_h},\\ S_4 =&\langle e_u, (\boldsymbol{\psi}-\boldsymbol{\Pi\psi})\cdot\boldsymbol{n}\rangle_{\partial\Omega_h} + \langle u-\widehat{u}_h^{\mathsf{v}_0}, (\boldsymbol{\Pi\psi}-\boldsymbol{\psi})\cdot\boldsymbol{n}\rangle_{\partial\Omega_h},\\ S^{\mathsf{v}} =&\langle\mathsf{v}\cdot\boldsymbol{n}(u-\widehat{u}_h^{\mathsf{v}}), \varphi-\mathsf{P}\varphi\rangle_{\partial\Omega_h} + (\mathsf{v}e_u, \nabla(\mathsf{P}\varphi-\varphi))_{\Omega_h}.\end{aligned}$$

The terms $S_i, i = 1, \cdots, 4$, are associated with diffusion and the term S^{v} is associated with convection. Now we only need to estimate these terms. Since Ω is a convex polyhedral domain, we have the elliptic regularity

$$\|\boldsymbol{\psi}\|_{\boldsymbol{H}^1(\Omega_h)} + \|\varphi\|_{H^2(\Omega_h)} \le C\|\theta\|_{L^2(\Omega_h)}.$$

By the approximation properties of projections $\boldsymbol{\Pi}$ and P, we have

$$\begin{aligned}S_1 \le&\|\boldsymbol{\psi}-\boldsymbol{\Pi\psi}\|_{\boldsymbol{L}^2(\Omega_h)}\|u-\mathsf{P}u\|_{H^1(\Omega_h)}\\ \le&Ch^{r+1}|\boldsymbol{\psi}|_{\boldsymbol{H}^1(\Omega_h)}|u|_{H^{r+1}(\Omega_h)}\\ \le&Ch^{r+1}|u|_{H^{k+1}(\Omega_h)}\|\theta\|_{L^2(\Omega_h)},\end{aligned}$$

and

$$\begin{aligned}S_2 \le&\|\boldsymbol{e}_{\boldsymbol{q}}\|_{\boldsymbol{L}^2(\Omega_h;\boldsymbol{c})}(\|\boldsymbol{\Pi\psi}-\boldsymbol{\psi}\|_{\boldsymbol{L}^2(\Omega_h;\boldsymbol{c})} + \|\nabla(\varphi-\mathsf{P}\varphi)\|_{\boldsymbol{L}^2(\Omega_h;\boldsymbol{c}^{-1})})\\ \le&C(\|\boldsymbol{q}-\boldsymbol{\Pi q}\|_{\boldsymbol{L}^2(\Omega_h;\boldsymbol{c})} + \|\boldsymbol{\Pi}\boldsymbol{e}_{\boldsymbol{q}}\|_{\boldsymbol{L}^2(\Omega_h;\boldsymbol{c})})\,h(\eta|\boldsymbol{\psi}|_{\boldsymbol{H}^1(\Omega_h)} + \gamma|\varphi|_{H^2(\Omega_h)})\\ \le&Ch(\eta+\gamma)(\eta h^r|\boldsymbol{q}|_{\boldsymbol{H}^r(\Omega_h)} + |\!|\!|(\boldsymbol{\Pi}\boldsymbol{e}_{\boldsymbol{q}}, \mathsf{P}e_u)|\!|\!|)\|\theta\|_{L^2(\Omega_h)},\end{aligned}$$

Because $\langle (\boldsymbol{q}-\widehat{\boldsymbol{q}}_h), \varphi\, \boldsymbol{n}\rangle_{\partial\Omega_h} = 0$, we have

$$\begin{aligned} S_3 =&\langle (\boldsymbol{q}-\widehat{\boldsymbol{q}}_h), \mathsf{P}\varphi - \varphi\, \boldsymbol{n}\rangle_{\partial\Omega_h}\\ \le&\|(\boldsymbol{q}-\widehat{\boldsymbol{q}}_h)\cdot \boldsymbol{n}\|_{L^2(\partial\Omega_h)}\|\mathsf{P}\varphi-\varphi\|_{L^2(\partial\Omega_h)}\\ \le&C(\|\boldsymbol{q}-\boldsymbol{q}_h\|_{L^2(\partial\Omega_h)} + \|\alpha e_u\|_{L^2(\partial\Omega_D\cap\Gamma^+_{\mathsf{v}_0})})\|\mathsf{P}\varphi-\varphi\|_{L^2(\partial\Omega_h)}. \end{aligned}$$

Using the trace inequality (3.20a) and the inverse inequality (3.19a), we get

$$\begin{aligned} \|\boldsymbol{q}-\boldsymbol{q}_h\|_{L^2(\partial\Omega_h)} \le& \|\boldsymbol{q}-\boldsymbol{\Pi q}\|_{L^2(\partial\Omega_h)} + \|\boldsymbol{\Pi e_q}\|_{\boldsymbol{L}^2(\partial\Omega_h)}\\ &\le Ch^{r-1/2}\|\boldsymbol{q}\|_{\boldsymbol{H}^r(\Omega_h)} + C\gamma h^{-1/2}\|\boldsymbol{\Pi e_q}\|_{L^2(\Omega_h;\boldsymbol{c})}, \end{aligned}$$

Using trace inequality (3.20b), we get

$$\begin{aligned} \|\alpha\, e_u\|_{L^2(\partial\Omega_D\cap\Gamma^+_{\mathsf{v}_0})} \le& \|\alpha(u-\mathsf{P}u)\|_{L^2(\partial\Omega_h\cap\Gamma^+_{\mathsf{v}_0})} + \|\alpha \mathsf{P}e_u\|_{L^2(\partial\Omega_h\cap\Gamma^+_{\mathsf{v}_0})}\\ &\le \alpha h^{r+1/2}\|u\|_{H^{r+1}(\Omega_h)} + \alpha^{1/2}\|\alpha(\mathsf{P}e_u)^2\|^{1/2}_{L^2(\partial\Omega_D\cap\Gamma^+_{\mathsf{v}_0})}, \end{aligned}$$

and

$$\|\mathsf{P}\varphi-\varphi\|_{L^2(\partial\Omega_h)} \le \|\mathsf{P}\varphi-\varphi\|^{1/2}_{L^2(\Omega_h)}\|\mathsf{P}\varphi-\varphi\|^{1/2}_{H^1(\Omega_h)} \le Ch^{3/2}\|\theta\|_{L^2(\Omega_h)}.$$

Hence

$$\begin{aligned} S_3 \le&C(h^{r+1}\|\boldsymbol{q}\|_{\boldsymbol{H}^r(\Omega_h)} + \alpha h^{r+2}\|u\|_{H^{r+1}(\Omega_h)})\|\theta\|_{L^2(\Omega_h)}\\ &+ Ch(\gamma+\alpha^{1/2}h^{1/2})\|\!|(\boldsymbol{\Pi e_q}, \mathsf{P}e_u)|\!\|\,\|\theta\|_{L^2(\Omega_h)}. \end{aligned}$$

Using (3.17), we obtain

$$\begin{aligned} S_4 =&\langle \widehat{u}_h^{\mathsf{v}_0} - u_h, (\boldsymbol{\psi}-\boldsymbol{\Pi\psi})\cdot\boldsymbol{n}\rangle_{\partial\Omega_h}\\ =&S_4' + \langle \mathsf{P}e_u, (\boldsymbol{\psi}-\boldsymbol{\Pi\psi})\cdot\boldsymbol{n}\rangle_{\partial\Omega_D\cap\Gamma^+_{\mathsf{v}_0}} + \langle u-\mathsf{P}u, (\boldsymbol{\psi}-\boldsymbol{\Pi\psi})\cdot\boldsymbol{n}\rangle_{\partial\Omega_D}. \end{aligned}$$

where

$$\begin{aligned} S_4' =&\langle (\boldsymbol{\beta}_0\cdot\boldsymbol{n} - \frac{1}{2})\,[\![u_h\boldsymbol{n}]\!], \boldsymbol{\psi}-\boldsymbol{\Pi\psi}\rangle_{\partial\Omega_h\setminus\Gamma} + \langle \mathsf{P}e_u, (\boldsymbol{\psi}-\boldsymbol{\Pi\psi})\cdot\boldsymbol{n}\rangle_{\partial\Omega_D\cap\Gamma^-_{\mathsf{v}_0}}\\ =&0, \end{aligned}$$

by the definition of the projection $\boldsymbol{\Pi}$, (3.10b). So by the trace inequalities (3.20a) and (3.20b), we get

$$\begin{aligned} S_4 \le&\alpha^{-1/2}\langle \alpha\mathsf{P}e_u, \mathsf{P}e_u\rangle^{1/2}_{\partial\Omega_D\cap\Gamma^+_{\mathsf{v}_0}}\|(\boldsymbol{\psi}-\boldsymbol{\Pi\psi})\cdot\boldsymbol{n}\|_{L^2(\partial\Omega_D\cap\Gamma^+_{\mathsf{v}_0})}\\ &+ \|u-\mathsf{P}u\|_{L^2(\partial\Omega_D)}\|(\boldsymbol{\psi}-\boldsymbol{\Pi\psi})\cdot\boldsymbol{n}\|_{L^2(\partial\Omega_D)}\\ \le&C\big(\alpha^{-1/2}h^{1/2}\|\!|(\boldsymbol{\Pi e_q}, \mathsf{P}e_u)|\!\| + h^{r+1}\|u\|_{H^{r+1}(\Omega_h)}\big)\|\theta\|_{L^2(\Omega_h)}. \end{aligned}$$

Note that if we use L^2-projection instead of the projection $\boldsymbol{\Pi}$, the term S_4' is nonzero and is bounded by $C\|e_u\|_{L^2(\Omega_h)}\|\theta\|_{L^2(\Omega_h)}$. This will make us *unable* to get a error estimate of u. So we see that it is here that the use the projection $\boldsymbol{\Pi}$ is *necessary*, in this approach. We can see again that, the definition of the projection is tailored to the definition of the numerical traces.

We have finished estimating the first four terms associated with diffusion. Now we only need to estimate the last term associated with convection. By using the trace inequality (3.20b), we get

$$S^{\mathsf{v}} \leq Ch\|\mathsf{v}\|_{\boldsymbol{L}^\infty(\Omega_h)}\|e_u\|_{L^2(\Omega_h)}\|\theta\|_{L^2(\Omega_h)}.$$

Therefore, we have that

$$\begin{aligned}\|e_u\|_{L^2(\Omega_h)} \leq& (\sum_{i=1}^{4} S_i + S^{\mathsf{v}})/\|\theta\|_{L^2(\Omega_h)}\\ \leq& Ch^{r+1}(1+\alpha h)|u|_{H^{r+1}(\Omega_h)} + Ch^{r+1}\eta(\eta+\gamma)|\boldsymbol{q}|_{\boldsymbol{H}^r(\Omega_h)}\\ &+ Ch(\eta+\gamma+\alpha^{1/2}h^{1/2}+\alpha^{-1/2}h^{-1/2})\|\!|(\boldsymbol{\Pi e_q}, \mathsf{P}e_u)\|\!|\\ &+ Ch\|\mathsf{v}\|_{\boldsymbol{L}^\infty(\Omega_h)}\|e_u\|_{L^2(\Omega_h)}.\end{aligned}$$

We let $C_{h,\alpha} = \eta+\gamma+h^{1/2}\alpha^{1/2}+\alpha^{-1/2}h^{-1/2}$. Using the intermediate error estimate of $\boldsymbol{q}$, Lemma 3.6, we get

$$\begin{aligned}\|e_u\|_{L^2(\Omega_h)} \leq& Ch^{r+1}\, C_{h,\alpha}(\gamma+\alpha^{1/2}h^{1/2}+h^{1/2}\|\mathsf{v}\|^{1/2}_{L^\infty(\Omega_h)}))|u|_{H^{r+1}(\Omega_h)}\\ &+ Ch^{r+1}C_{h,\alpha}(\eta+\alpha^{-1/2}h^{-1/2})|\boldsymbol{q}|_{\boldsymbol{H}^r(\Omega_h)}\\ &+ Ch^{(r+2)/2}C_{h,\alpha}\|\mathsf{v}\|^{1/2}_{\boldsymbol{L}^\infty(\Omega_h)}|u|^{1/2}_{H^{r+1}(\Omega_h)}\|e_u\|^{1/2}_{L^2(\Omega_h)}\\ &+ Ch\|\mathsf{v}\|_{\boldsymbol{L}^\infty(\Omega_h)}\|e_u\|_{L^2(\Omega_h)}.\end{aligned}$$

If there is no convection, namely, $\mathsf{v} = 0$, then the last two terms on the right hand side of the inequality are zero, and the error estimate of u is obtained. If the convection velocity v is not zero, then for h small enough we have

$$\begin{aligned}\|e_u\|_{L^2(\Omega_h)} \leq& Ch^{r+1}\, C_{h,\alpha}(\gamma+\alpha^{1/2}h^{1/2}+h^{1/2}\|\mathsf{v}\|^{1/2}_{L^\infty(\Omega_h)}))|u|_{H^{r+1}(\Omega_h)}\\ &+ Ch^{r+1}C_{h,\alpha}(\eta+\alpha^{-1/2}h^{-1/2})|\boldsymbol{q}|_{\boldsymbol{H}^r(\Omega_h)}\\ &+ Ch^{r+2}C^2_{h,\alpha}\|\mathsf{v}\|^{1/2}_{\boldsymbol{L}^\infty(\Omega_h)}|u|_{H^{r+1}(\Omega_h)}.\end{aligned}$$

This completes the proof of Theorem 2.3.

3.3.4. *Step 3: Error estimate in $\boldsymbol{q}$.* To end the proof of Theorem 2.2, we only need to insert the result in Theorem 2.3 into the estimate of Lemma 3.6 and use Lemma 3.3.

3.4. **The condition number of the Schur-complement matrix: Proof of Theorem 2.4.** To prove the estimate of the condition number of the Schur-complement matrix

$$A := (B^t\, M^{-1}\, B + C)$$

given by (2.9), we begin by proving the following simple result.

Lemma 3.7. *We have that, for all $\zeta \in W_h$,*

$$[\zeta]^t A[\zeta] = \mathcal{A}(\zeta,\zeta) := (\boldsymbol{c}\boldsymbol{\sigma}(\zeta), \boldsymbol{\sigma}(\zeta))_{\Omega_h} + \langle \alpha, \zeta^2\rangle_{\partial\Omega_D\cap\Gamma^+_{\mathsf{v}_0}},$$

where $\boldsymbol{\sigma}(\zeta)$ is the element of $\boldsymbol{V}_h$ such that

$$(\boldsymbol{c}\,\boldsymbol{\sigma}(\zeta), \boldsymbol{v})_{\Omega_h} = (\zeta, \nabla\cdot\boldsymbol{v})_{\Omega_h} - \langle \widehat{\zeta}^{\mathsf{v}_0}, \boldsymbol{v}\cdot\boldsymbol{n}\rangle_{\partial\Omega_h\setminus\partial\Omega_D} \tag{3.25}$$

for all $\boldsymbol{v} \in \boldsymbol{V}_h$.

Proof. By the definition of $\mathcal{A}(\zeta,\zeta)$, we have

$$\mathcal{A}(\zeta,\zeta) = [\boldsymbol{\sigma}]^t M[\boldsymbol{\sigma}] + [\zeta]^t C[\zeta].$$

It is easy to see that (3.25) is equivalent to the following matrix expression,

$$M[\boldsymbol{\sigma}] = -B[\zeta]. \tag{3.26}$$

Using the definition of A and (3.26), we get

$$\begin{aligned}[\zeta]^t A[\zeta] =&[\zeta]^t(B^t M^{-1} B + C)[\zeta]\\ =&[\zeta]^t B^t M^{-1} B\,[\zeta] + [\zeta]^t C[\zeta]\\ =&(-M[\boldsymbol{\sigma}])^t\, M^{-1}\,(-M[\boldsymbol{\sigma}]) + [\zeta]^t C[\zeta]\\ =&[\boldsymbol{\sigma}]^t M[\boldsymbol{\sigma}] + [\zeta]^t C[\zeta].\end{aligned}$$

Hence, $\mathcal{A}(\zeta,\zeta) = [\zeta]^t A[\zeta]$ for all $\zeta \in W_h$. This completes the proof. □

We are now ready to estimate the condition number of the matrix A, κ. Since A is symmetric and positive definite, we immediately have that

$$\kappa \le \frac{C_l}{C_u},$$

where

$$C_l\,[\zeta]^2_{L^2(\Omega_h)} \le \mathcal{A}(\zeta,\zeta) \le C_u\,[\zeta]^2_{L^2(\Omega_h)} \quad \forall \zeta \in W_h.$$

Thus, we only have to find expressions for C_l and C_u.

Let us begin by finding C_u. Integrating by parts in the formulation defining $\boldsymbol{\sigma} := \boldsymbol{\sigma}(\zeta)$, (3.25), we get

$$(\boldsymbol{c}\boldsymbol{\sigma}, \boldsymbol{v})_{\Omega_h} = -\,(\nabla\zeta, \boldsymbol{v})_{\Omega_h} + \langle \zeta - \widehat{\zeta}^{\mathbf{v}_0}, \boldsymbol{v}\cdot\boldsymbol{n}\rangle_{\partial\Omega_h},$$

and inserting the expression of the numerical trace given by (3.17), we obtain

$$(\boldsymbol{c}\boldsymbol{\sigma}, \boldsymbol{v})_{\Omega_h} = -\,(\nabla\zeta, \boldsymbol{v})_{\Omega_h} + \langle(\frac{1}{2} - \boldsymbol{\beta}_0\cdot\boldsymbol{n})\,[\![\zeta\boldsymbol{n}]\!], \boldsymbol{v}\rangle_{\partial\Omega_h\setminus\Gamma} + \langle\zeta, \boldsymbol{v}\cdot\boldsymbol{n}\rangle_{\partial\Omega_D}. \tag{3.27}$$

Taking $\boldsymbol{v} = \boldsymbol{\sigma}$, we obtain, after a simple application of the Cauchy-Schwarz inequality and the trace inequality of Lemma 3.7,

$$(\boldsymbol{c}\boldsymbol{\sigma}, \boldsymbol{\sigma})_{\Omega_h} \le \gamma(1 + C_1^2)^{1/2}\,\|\,\zeta\,\|_{1,h}\,(\boldsymbol{c}\boldsymbol{\sigma}, \boldsymbol{\sigma})^{1/2}_{\Omega_h},$$

where

$$\|\,\zeta\,\|^2_{1,h} := \|\nabla\zeta\|^2_{L^2(\Omega_h)} + \langle h^{-1},\, [\![\zeta\boldsymbol{n}]\!]^2\rangle_{\mathscr{E}_h^i\cup\partial\Omega_D},$$

This implies that

$$(\boldsymbol{c}\boldsymbol{\sigma}, \boldsymbol{\sigma})_{\Omega_h} \le \gamma^2(1 + C_1^2)\,\|\,\zeta\,\|^2_{1,h},$$

and so,

$$\begin{aligned}\mathcal{A}(\zeta,\zeta) \le& C_3\left(h^{-2} + \alpha\, h^{-1}\right)\,\|\zeta\|^2_{L^2(\Omega_h)},\\ \le& C_3\left(h^{-2} + \alpha\, h^{-1}\right)\,C_{qu}\,h^d\,[\,\zeta\,]^2,\end{aligned}$$

where C_3 depends on γ^2 and an inverse inequality constant; C_{qu} is a constant given the assumption of quasi-uniformity of the triangulation. This implies that we can take

$$C_u = C_3\left(h^{-2} + \alpha\, h^{-1}\right)\,C_{qu}\,h^d.$$

Let us now find C_l. To do that we take, on each $K \in \Omega_h$, $\boldsymbol{v} := \boldsymbol{Z}$ as given in Lemma 3.2 with

$$(\boldsymbol{Z}, \nabla\zeta)_K = -(\nabla\zeta, \nabla\zeta)_K,$$

$$\langle \boldsymbol{Z}, [\![\zeta \boldsymbol{n}]\!]\rangle_{e_i} = \begin{cases} \langle h^{-1} [\![\zeta \boldsymbol{n}]\!], [\![\zeta \boldsymbol{n}]\!]\rangle_{e_i} & \text{if } e_i \nsubseteq \partial\Omega_N, \\ 0 & \text{if } e_i \subseteq \partial\Omega_N, \end{cases} \quad i = 1, ..., d,$$

where $\{e_1, ..., e_d\} \supseteq \{e \in \partial K : \boldsymbol{\beta}_0 \cdot \boldsymbol{n}_e \neq \frac{1}{2}\}$, and insert it in (3.27) to get

$$(\boldsymbol{c\sigma}, \boldsymbol{Z})_{\Omega_h} = (\nabla\zeta, \nabla\zeta)_{\Omega_h} + \langle h^{-1}, [\![\zeta \boldsymbol{n}]\!]^2\rangle_{\mathscr{E}_h^i \cup (\partial\Omega_D \cap \Gamma^-_{\mathbf{v}_0})} + \langle [\![\zeta \boldsymbol{n}]\!], \boldsymbol{Z}\rangle_{\partial\Omega_D \cap \Gamma^+_{\mathbf{v}_0}}.$$

This implies that

$$\begin{aligned}(\nabla\zeta, \nabla\zeta)_{\Omega_h} + \langle h^{-1}, [\![\zeta \boldsymbol{n}]\!]^2\rangle_{\mathscr{E}_h^i \cup (\partial\Omega_D \cap \Gamma^-_{\mathbf{v}_0})} =& (\boldsymbol{c\sigma}, \boldsymbol{Z})_{\Omega_h} - \langle [\![\zeta \boldsymbol{n}]\!], \boldsymbol{Z}\rangle_{\partial\Omega_D \cap \Gamma^+_{\mathbf{v}_0}} \\ \le& \eta (\boldsymbol{c\sigma}, \boldsymbol{\sigma})^{1/2}_{\Omega_h} \|\boldsymbol{Z}\|_{\boldsymbol{L}^2(\Omega_h)} \\ &+ C_1 \langle h^{-1}, [\![\zeta \boldsymbol{n}]\!]^2\rangle^{1/2}_{\partial\Omega_D \cap \Gamma^+_{\mathbf{v}_0}} \|\boldsymbol{Z}\|_{L^2(\Omega_h)}.\end{aligned}$$

So by (3.13),

$$(\nabla\zeta, \nabla\zeta)_{\Omega_h} + \langle h^{-1}, [\![\zeta \boldsymbol{n}]\!]^2\rangle_{\mathscr{E}_h^i \cup (\partial\Omega_D \cap \Gamma^-_{\mathbf{v}_0})} \le C_4(\eta(\boldsymbol{c\sigma}, \boldsymbol{\sigma})_{\Omega_h} + C_1 \langle h^{-1}, [\![\zeta \boldsymbol{n}]\!]^2\rangle_{\partial\Omega_D \cap \Gamma^+_{\mathbf{v}_0}})$$

where C_4 depends only on d, k and the shape regularity constant. Then

$$\|\zeta\|^2_{1,h} \le \max\{C_4\eta, (1 + C_4C_1)h^{-1}\alpha^{-1}\}\mathcal{A}(\zeta, \zeta).$$

Since, by the inequality (1.8) in [5], we have that

$$\|\zeta\|^2_{L^2(\Omega_h)} \le C_5 \|\zeta\|^2_{1,h}$$

where C_5 depends only on the shape regularity constant, we obtain that

$$C_l\,[\,\zeta\,]^2 \le (C_6 \max\{1, h^{-1}\alpha^{-1}\})^{-1} \|\zeta\|^2_{L^2(\Omega_h)} \le \mathcal{A}(\zeta, \zeta)$$

where

$$C_l = C'_{qu}\, h^d\, (C_6 \max\{1, h^{-1}\alpha^{-1}\})^{-1},$$

C_6 depend on η, d, k, an inverse inequality constant and the shape regularity constant, and C'_{qu} is a constant by the assumption on the quasi-uniformity of the mesh.

We then have that

$$\kappa \le \frac{C_u}{C_l} \le C\, h^{-2},$$

whenever α is of order h^{-1}. This completes the proof of Theorem 2.4.

4. Numerical experiment

In this section, we numerically verify the sharpness of our theoretical results. We take the diffusion coefficient $\boldsymbol{c}$ to be the identity matrix , the convection velocity $\mathbf{v}$ to be zero, and $\partial\Omega_N$ to be empty. We take the domain to be the unit square $[0, 1] \times [0, 1]$. The boundary data u_D is chosen so that the exact solution is $u(\boldsymbol{x}) = \frac{1}{2}\ln((x + 0.1)^2 + (y + 0.1)^2)$; as a consequence, $f = 0$.

We take the stabilization parameter α in the numerical trace of $\boldsymbol{q}$ to be $\alpha = h^{-1}$. Here we use uniform meshes, and "mesh=i" means that we used a uniform mesh with $2(4)^i$ evenly distributed elements.

The history of convergence of our methods is displayed in table 1 and 2. In both tables, the first column shows the polynomial degree k we used to approximate the unknown $\boldsymbol{q}$ and u; the second column in the tables displays the mesh number; the third column is the L^2-norm of error of the approximation, and the last column is the order of convergence. The approximate order of convergence, ri, is defined by

$$ri := \frac{\ln(\frac{e(i-1)}{e(i)})}{\ln 2}, \quad i > 1,$$

where $e(i)$ is the error of the approximation computed on the mesh i.

TABLE 1. History of convergence for $\boldsymbol{q}$.

k	mesh	$\|\boldsymbol{q} - \boldsymbol{q}_h\|_{L^2(\Omega_h)}$	convergence order
	1	0.18E+00	–
	2	0.91E-01	0.99
1	3	0.45E-01	1.00
	4	0.23E-01	0.99
	5	0.12E-01	0.99
	1	0.46E-01	–
	2	0.14E-01	1.73
2	3	0.36E-02	1.95
	4	0.88E-03	2.02
	5	0.22E-03	1.98

In Table 1 we can see that when polynomials of degree $k = 1, 2$ are used, the L^2-norms of the error in $\boldsymbol{q}$ are of order k. So our error estimate in Theorem 2.2 is sharp.

TABLE 2. History of convergence for u.

k	mesh	$\|u - u_h\|_{L^2(\Omega_h)}$	convergence order
	1	0.80E-02	–
	2	0.21E-02	1.92
1	3	0.53E-03	2.02
	4	0.13E-03	2.03
	5	0.32E-04	2.02
	1	0.20E-02	–
	2	0.33E-03	2.59
2	3	0.45E-04	2.87
	4	0.57E-05	2.98
	5	0.72E-06	2.99

In Table 2 we can see that when polynomials of degree $k = 1, 2$ are used, the L^2-norms of the error in u converge with optimal order $k + 1$.

5. Some extensions

5.1. **Extension to more general numerical traces.** In section 2.1, the functions $\boldsymbol{\beta}$ and $\boldsymbol{\beta}_0$ in numerical traces are defined by (2.6b) and (2.7b). Here we show that our results actually hold for any functions $\boldsymbol{\beta} \in H(div, \Omega_h)$ and $\boldsymbol{\beta}_0 \in H(div, \Omega_h)$ which satisfy that

$$\text{sign}\big(\boldsymbol{\beta} \cdot \boldsymbol{n}_K(\boldsymbol{x})\big) = \text{sign}\,\big(\mathsf{v}(\boldsymbol{x}) \cdot \boldsymbol{n}_K(\boldsymbol{x})\big), \tag{5.1a}$$

$$\text{sign}\big(\boldsymbol{\beta}_0 \cdot \boldsymbol{n}_K(e)\big) = \text{sign}\,\big(\mathsf{v}_0 \cdot \boldsymbol{n}_K(e)\big). \tag{5.1b}$$

The proof of the existence and uniqueness of the solution is almost the same as what we showed in section 3.2. The only difference is that in the equation (3.18), $(\frac{1}{2} - \boldsymbol{\beta}_0 \cdot \boldsymbol{n})$ might not be zero on any faces. But we can take $\boldsymbol{v} = \mathsf{v}_0 u_h$ in (3.18) and get that $[\![u_h \boldsymbol{n}]\!] = 0$ on at least one face e_K for each simplex K, then follow the original argument.

Since our error analysis relies on a suitable definition projection $\boldsymbol{\Pi}$, closely related to that of the numerical traces, we need to modify it. For the generalized numerical traces, (5.1), we define the projection $\boldsymbol{\Pi}$ as follows. The restriction of $\boldsymbol{\Pi}$ to K is such that for any given $\boldsymbol{\sigma} \in \boldsymbol{H}^1(\Omega_h)$, $\boldsymbol{\Pi\sigma} \in \boldsymbol{\mathcal{P}}(K)$ satisfies

$$(\boldsymbol{\Pi\sigma} - \boldsymbol{\sigma}, \boldsymbol{v})_K = 0 \quad \forall\, \boldsymbol{v} \in \boldsymbol{\mathcal{P}}^{k-1}(K), \tag{5.2a}$$

$$\langle (\boldsymbol{\Pi\sigma} - \boldsymbol{\sigma}) \cdot \boldsymbol{n}, \omega \rangle_{e_i} = 0 \quad \forall \omega \in \mathcal{P}^k(e_i), \text{ if } \boldsymbol{\beta}_0 \cdot \boldsymbol{n}_{e_i} \geq 0 \text{ or } e_i \in \Gamma, \tag{5.2b}$$

$$\langle (\widehat{\boldsymbol{\Pi\sigma}} - \boldsymbol{\sigma}) \cdot \boldsymbol{n}, \omega \rangle_{e_i} = 0 \quad \forall \omega \in \mathcal{P}^k(e_i), \text{ if } \boldsymbol{\beta}_0 \cdot \boldsymbol{n}_{e_i} < 0 \text{ and } e_i \notin \Gamma, \tag{5.2c}$$

where $e_i, i = 1, ..., d$ are d faces of the simplex K, and $\{e_i, i = 1, ..., d\} \supseteq \{e \in \partial K : \boldsymbol{n}_e \cdot \boldsymbol{\beta}_0(e) \leq 0\}$. The well-definiteness and the approximation order of the projection can be proved in a same way as in section 3.1.

In the error estimates, the only steps associating the definition of the projection $\boldsymbol{\Pi}$ with the numerical traces are the eliminations of T_5' and S_4'. So it is enough to show that for the numerical traces with $\boldsymbol{\beta}$ and $\boldsymbol{\beta_0}$ defined in (5.1), the projection defined above in (5.2) can still make T_5' and S_4' zero. Since

$$T_5' = \langle (\widehat{\boldsymbol{\Pi q}} - \boldsymbol{q}) \cdot \boldsymbol{n}, \mathsf{P}e_u \rangle_{\partial\Omega_h \setminus (\partial\Omega_N \cup \Gamma^+_{\mathsf{v}_0})} = \langle \widehat{\boldsymbol{\Pi q}} - \boldsymbol{q}, [\![\mathsf{P}e_u \boldsymbol{n}]\!] \rangle_{\mathscr{E}_h^i \cup (\partial\Omega_D \cap \Gamma^-_{\mathsf{v}_0})}$$

and S_4' can be rewrite as

$$S_4' = \langle \boldsymbol{\psi} - \widehat{\boldsymbol{\Pi\psi}}, [\![u_h \boldsymbol{n}]\!] \rangle_{\mathscr{E}_h^i} + \langle (\boldsymbol{\psi} - \widehat{\boldsymbol{\Pi\psi}}) \cdot \boldsymbol{n}, \mathsf{P}e_u \rangle_{\partial\Omega_D \cap \Gamma^-_{\mathsf{v}_0}},$$

we only need to show that if $e \in \mathscr{E}_h^i \cup (\partial\Omega_D \cap \Gamma^-_{\mathsf{v}_0})$, then

$$\langle (\boldsymbol{\sigma} - \widehat{\boldsymbol{\Pi\sigma}}) \cdot \boldsymbol{n}, \omega \rangle_e = 0 \qquad \forall\, \omega \in \mathcal{P}^k(e) \tag{5.3}$$

for any $\boldsymbol{\sigma} \in \boldsymbol{H}^1(\Omega_h)$. This is true because for any $e \in \mathscr{E}_h^i \cup (\partial\Omega_D \cap \Gamma^-_{\mathsf{v}_0})$, there is a simplex K such that $e \in \partial K$ and $\boldsymbol{\beta}_0 \cdot \boldsymbol{n}_K(e) \leq 0$. If $\boldsymbol{\beta}_0 \cdot \boldsymbol{n}_K(e) < 0$, then (5.3) follows from (5.2c). If $\boldsymbol{\beta}_0 \cdot \boldsymbol{n}_K(e) = 0$, (5.3) follows from (5.2b).

5.2. **Extension to curved domain.** At the beginning of the paper, we assumed Ω to be a polyhedral domain in $\mathbb{R}^d$ $(d \geq 2)$. Here we show that when $d = 2$ we can extend our results to curved boundary domains.

We assume that $\mathbf{v}_0 \cdot \boldsymbol{n}$ does not change sign on any face $e \in \Gamma$, and we modify the numerical trace $\widehat{\boldsymbol{q}}_h$ to be

$$\widehat{\boldsymbol{q}}_h = \begin{cases} \{\!\{\boldsymbol{q}_h\}\!\} - \boldsymbol{\beta}_0 [\![\boldsymbol{q}_h \cdot \boldsymbol{n}]\!], & \text{if } e \in \mathscr{E}_h^i, \\ \boldsymbol{q}_h + \alpha(u_h - g)\boldsymbol{n}, & \text{if } e \in \partial\Omega_D, \\ \mathsf{q}_{\mathsf{N}} \boldsymbol{n}, & \text{if } e \in \partial\Omega_N. \end{cases}$$

Note that in the case of a polyhedral domain, the stabilization parameters α is nonzero only on $\partial\Omega_D \cap \Gamma^+_{\mathbf{v}_0}$. In the curved boundary domain, we take α different from zero on $\partial\Omega_D$.

The proof of the existence and uniqueness of the solution for the curved boundary domain is almost the same as that for the polyhedral convex domain.

In the analysis of the error estimates, if K has no curved faces, the restriction of $\boldsymbol{\Pi}$ to K is defined in the same way as for the polygonal domain, namely, $\{e_1, e_2\} \supseteq \{e \in \partial K : \boldsymbol{n}_e \cdot \boldsymbol{\beta}_0(e) \le 0\}$. If K has a curved face, the restriction of $\boldsymbol{\Pi}$ to K is defined with e_1, e_2 are the straight faces.

For the intermediate estimate of $\boldsymbol{q}$, the proof is almost the same as that for the polyhedral domain. The only difference is that the terms in form of $\langle(\boldsymbol{\Pi}\boldsymbol{q} - \boldsymbol{q}) \cdot \boldsymbol{n}, \mathsf{P}e_u\rangle_e$ do not vanish when $e \in \partial\Omega_D \cap \Gamma^-_{\mathbf{v}_0}$. But the stabilization parameters α is nonzero on $\partial\Omega_D \cap \Gamma^-_{\mathbf{v}_0}$, so these terms can be estimated.

For the estimate of u, we use the projection $\boldsymbol{\Pi}$ defined above and the rest part of the proof is the same. So the orders of convergence for $\boldsymbol{q}$ and u are the also same as in the polyhedral domain case.

6. Conclusions

In this paper, we investigated the convergence properties of the MD-LDG method for two or higher dimensional convection-diffusion equations with variable coefficients. We showed that, even though the stabilization parameters on the interior faces are identically equal to zero, the L^2-norm error estimates of the auxiliary variable $\boldsymbol{q}$ and the primary variable u are of order k and $(k+1)$, respectively, when polynomials of degree k are used. Our numerical experiments indicate the sharpness of the L^2-error estimates in $\mathbb{R}^2$.

References

[1] D. N. Arnold, *An interior penalty finite element method with discontinuous elements*, SIAM J. Numer. Anal. **19** (1982), 742–760.

[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.

[3] G. A. Baker, *Finite element methods for elliptic equations using nonconforming elements*, Math. Comp. **31** (1977), 45–59.

[4] F. Bassi, S. Rebay, G. Mariotti, S. Pedinotti, and M. Savini, *A high-order accurate discontinuous finite element method for inviscid and viscous turbomachinery flows*, 2nd European Conference on Turbomachinery Fluid Dynamics and Thermodynamics (Antwerpen, Belgium) (R. Decuypere and G. Dibelius, eds.), Technologisch Instituut, March 5–7 1997, pp. 99–108.

[5] S. Brenner, *Poincaré-Friedrichs inequalities for piecewise H^1 functions*, SIAM J. Numer. Anal. **41** (2003), 306–324.

[6] F. Brezzi and M. Fortin, *Mixed and hybrid finite element methods*, Springer Verlag, 1991.

[7] F. Brezzi, G. Manzini, L. D. Marini, P. Pietra, and A. Russo, *Discontinuous finite elements for diffusion problems*, in Atti Convegno in onore di F. Brioschi (Milan 1997), Istituto Lombardo, Accademia di Scienze e Lettere, 1999, pp. 197–217.

[8] P. Castillo, private communication.

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

[10] 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.

[11] P. Castillo, B. Cockburn, D. Schötzau, and C. Schwab, *Optimal a priori error estimates for the hp-version of the local discontinuous Galerkin method for convection-diffusion problems*, Math. Comp. **71** (2002), 455–478.

[12] 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.

[13] B. Cockburn and C. Dawson, *Some extensions of the local discontinuous Galerkin method for convection-diffusion equations in multidimensions*, The Proceedings of the Conference on the Mathematics of Finite Elements and Applications: MAFELAP X (J.R. Whiteman, ed.), Elsevier, 2000, pp. 225–238.

[14] 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.

[15] B. Rivière, M. F. Wheeler, and V. Girault, *Improved energy estimates for interior penalty, constrained and discontinuous Galerkin methods for elliptic problems. Part I*, Comp. Geosci. **3** (1999), 337–360.

[16] S. Siddarth, J. Carrero, B. Cockburn, K.K. Tamma, and R. Kanapady, *The local discontinuous Galerin method and component design integration for 3D elasticity*, Proceedings of the Third M.I.T. Conference on Computational Fluid and Solid Mechanics, Cambridge, MA, USA (K.J. Bathe, ed.), Elsevier, 2005, pp. 492–494.

(B. Cockburn) 127 Vincent Hall, University of Minnesota, Minneapolis, MN 55455, USA

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

(B. Dong) 127 Vincent Hall, University of Minnesota, Minneapolis, MN 55455, USA
E-mail address: `bdong@math.umn.edu`