

A PROJECTION-BASED ERROR ANALYSIS OF HDG METHODS

BERNARDO COCKBURN, JAYADEEP GOPALAKRISHNAN,
AND FRANCISCO-JAVIER SAYAS

ABSTRACT. We introduce a new technique for the error analysis of hybridizable discontinuous Galerkin (HDG) methods. The technique relies on the use of a new projection whose design is inspired by the form of the numerical traces of the methods. This renders the analysis of the projections of the discretization errors simple and concise. By showing that these projections of the errors are bounded in terms of the distance between the solution and its projection, our studies of influence of the stabilization parameter are reduced to local analyses of approximation by the projection. We illustrate the technique on a specific HDG method applied to a model second-order elliptic problem.

1. INTRODUCTION

This paper is dedicated to presenting a new technique for the error analysis of an emerging class of numerical methods called hybridizable discontinuous Galerkin (HDG) methods [11]. The main idea is to devise a projection that matches the form of the numerical traces of the HDG method. The analysis of the projection of the error then becomes simple, concise and independent of the particular choice of the stabilization parameters.

We do not aim for maximal generality, but rather to convey the advantages of our technique compared to other DG analyses. Hence we only consider one specific HDG method, namely, the LDG-hybridizable (LDG-H) method [11] for the model problem

$$\begin{aligned} (1.1a) \quad & c \mathbf{q} + \nabla u = 0 && \text{in } \Omega, \\ (1.1b) \quad & \nabla \cdot \mathbf{q} = f && \text{in } \Omega, \\ (1.1c) \quad & u = g && \text{on } \partial\Omega, \end{aligned}$$

Received by the editor December 22, 2008 and, in revised form, April 9, 2009.

2010 *Mathematics Subject Classification.* Primary 65M60, 65N30, 35L65.

Key words and phrases. Discontinuous Galerkin methods, hybridization, superconvergence, postprocessing.

The first author was supported in part by the National Science Foundation (Grant DMS-0712955) and by the University of Minnesota Supercomputing Institute.

The second author was supported in part by the National Science Foundation under grants DMS-0713833 and SCREMS-0619080.

The third author was partially supported by MEC/FEDER Project MTM2007–63204, Gobierno de Aragón (Grupo PDIE) and was a Visiting Professor of the School of Mathematics, University of Minnesota, during the development of this work.

©2010 American Mathematical Society
Reverts to public domain 28 years from publication

where Ω is a Lipschitz polyhedral domain in \mathbb{R}^n ($n \geq 2$). Here $c : \Omega \mapsto \mathbb{R}^{n \times n}$ is a variable matrix-valued coefficient, which we assume to be symmetric and uniformly positive definite, f is in $L^2(\Omega)$ and g in $H^{1/2}(\partial\Omega)$.

Let us put our result in historical perspective. To do that, we begin by describing the HDG method. Consider a partitioning of the domain Ω into elements K forming a mesh \mathcal{T}_h satisfying the standard finite element conditions [7]; the faces of K are going to be denoted by F . This method yields a scalar approximation u_h to u , a vector approximation \mathbf{q}_h to \mathbf{q} , and a scalar approximation \hat{u}_h to the trace of u on element boundaries, in spaces of the form

$$(1.2a) \quad \mathbf{V}_h = \{\mathbf{v} : \text{for all mesh elements } K, \mathbf{v}|_K \in \mathbf{V}(K)\},$$

$$(1.2b) \quad W_h = \{w : \text{for all mesh elements } K, w|_K \in W(K)\},$$

$$(1.2c) \quad M_h = \{\mu : \text{for all mesh faces } F, \mu|_F \in M(F)\},$$

respectively, where $\mathbf{V}(K)$, $W(K)$, and $M(F)$ are finite dimensional spaces. The HDG approximations u_h in W_h , \mathbf{q}_h in \mathbf{V}_h , and the numerical trace \hat{u}_h in M_h , are determined by requiring that

$$(1.3a) \quad (c \mathbf{q}_h, \mathbf{r})_{\mathcal{T}_h} - (u_h, \nabla \cdot \mathbf{r})_{\mathcal{T}_h} + \langle \hat{u}_h, \mathbf{r} \cdot \mathbf{n} \rangle_{\partial\mathcal{T}_h} = 0,$$

$$(1.3b) \quad -(\mathbf{q}_h, \nabla w)_{\mathcal{T}_h} + \langle \hat{\mathbf{q}}_h \cdot \mathbf{n}, w \rangle_{\partial\mathcal{T}_h} = (f, w)_{\mathcal{T}_h},$$

$$(1.3c) \quad \langle \hat{u}_h, \mu \rangle_{\partial\Omega} = \langle g, \mu \rangle_{\partial\Omega},$$

$$(1.3d) \quad \langle \hat{\mathbf{q}}_h \cdot \mathbf{n}, \mu \rangle_{\partial\mathcal{T}_h \setminus \partial\Omega} = 0,$$

hold for all \mathbf{r} in \mathbf{V}_h , $w \in W_h$, and $\mu \in M_h$, with a specific $\hat{\mathbf{q}}_h$ defined on $\partial\mathcal{T}_h = \{\partial K : K \in \mathcal{T}_h\}$; see [11]. Above and throughout, we use the notation

$$(v, w)_{\mathcal{T}_h} = \sum_{K \in \mathcal{T}_h} (v, w)_K \quad \text{and} \quad \langle v, w \rangle_{\partial\mathcal{T}_h} = \sum_{K \in \mathcal{T}_h} \langle v, w \rangle_{\partial K},$$

where we write $(u, v)_D = \int_D uv \, dx$ whenever D is a domain of \mathbb{R}^n , and $\langle u, v \rangle_D = \int_D uv \, dx$ whenever D is a domain of \mathbb{R}^{n-1} . For vector functions \mathbf{v} and \mathbf{w} , the notations are similarly defined with the integrand being the dot product $\mathbf{v} \cdot \mathbf{w}$. For HDG methods \hat{u}_h is taken to be an *unknown* in M_h , while the numerical trace $\hat{\mathbf{q}}_h$ is prescribed on $\partial\mathcal{T}_h$ in such a way that both \mathbf{q}_h and u_h can be eliminated from the above equations to give rise to a single equation for \hat{u}_h ; see [11]. Thus the often made criticism that DG methods have too many unknowns does not apply to HDG methods. Additional advantages of HDG methods include the ability to postprocess to get higher order solutions as we shall see in Section 5.

In contrast, the DG methods of the last century for second-order elliptic problems use only the first two equations (1.3a)–(1.3b), together with specific prescriptions of both \hat{u}_h and $\hat{\mathbf{q}}_h$ to define their approximations. For *most* such methods, the above mentioned elimination is not feasible. Moreover, their analysis is also very different from ours, as it does not require the use of any special projection. This can be seen in [2], where most of the then known DG methods were analyzed in a single unifying framework. All of these methods were shown to converge with order $k+1$ in the scalar variable u and with order k in the flux \mathbf{q} . The order of convergence in u is optimal and that in \mathbf{q} *sub-optimal* since the methods use as local spaces $\mathbf{V}(K)$ and $W(K)$ the sets $\mathcal{P}_k(K)$ and $\mathcal{P}_k(K)$, respectively, where $\mathcal{P}_k(K) := [\mathcal{P}_k(K)]^n$ and $\mathcal{P}_k(K)$ is the space of polynomials of total degree at most k .

Two particular DG methods of the local discontinuous Galerkin (LDG) type, fitting in the above-mentioned unifying framework, deserve special mention as they provide approximations to the flux which converge with better orders of convergence. Their analyses do use special projections. The first is defined in the one-dimensional case and provides an order of convergence for the approximate flux of $k + 1$; see [6]. The second is an extension of that method to multiple dimensions. It uses Cartesian grids and takes as local spaces $V(K)$ and $W(K)$ the sets $\mathcal{Q}_k(K)$ and $\mathcal{Q}_k(K)$, respectively, where $\mathcal{Q}_k(K) := [\mathcal{Q}_k(K)]^n$ and $\mathcal{Q}_k(K)$ is the space of polynomials of degree at most k in each variable. The order of convergence for the approximate flux can be proven to be $k + 1/2$; see [13]. In both cases, the special projections used in the analysis were carefully devised to capture the structure of the numerical fluxes \hat{u}_h and \hat{q}_h .

Another example of a method requiring a special projection to carry out its analysis is the so-called minimal dissipation LDG method. This is a special LDG method whose penalization parameter is identically zero in all the interior faces; as a consequence, it *cannot* be analyzed as in [2]. However, by using a suitably defined projection, the lack of stabilization can be overcome and sharp error estimates can be obtained. The flux can be shown to converge with the *sub-optimal* but sharp order of k ; see [8].

Now, let us consider the HDG methods for which

$$(1.4) \quad \hat{q}_h = q_h + \tau (u_h - \hat{u}_h) \mathbf{n}, \quad \text{on } \partial\mathcal{T}_h,$$

for some nonnegative penalty function τ defined on $\partial\mathcal{T}_h$ which we assume to be constant on each face of the triangulation. As explained in [11], these methods are called the LDG-hybridizable (LDG-H) methods because the above numerical trace is that of the LDG method applied separately on each mesh element K . The well-known hybridized versions of the Raviart-Thomas (RT) method [18] and the Brezzi-Douglas-Marini (BDM) method [5] can be considered to be HDG methods for which $\tau = 0$; their analyses are carried out by using the celebrated RT and BDM projections. The analysis of HDG methods for which τ is not identically equal to zero, has been carried out in [12] also by using special projections. All of the above methods were proven to provide approximations to the flux converging with the *optimal* order of $k + 1$; the first HDG method with this property was introduced in [9].

In this paper, we provide a new approach to the error analysis of HDG methods. It is an alternative to the techniques in [12] for general HDG methods, and those in [9] for the particular method treated therein. We recover all the results of [9, 12] and obtain new superconvergence results for the projection of the scalar variable with our new approach. The novelty of our analysis is the use of new projections Π_V, Π_W which are fitted to the structure of the numerical traces of the scheme in the sense that

$$P_M(\mathbf{q} \cdot \mathbf{n}) = \Pi_V \mathbf{q} \cdot \mathbf{n} + \tau(\Pi_W u - P_M u) \quad \text{on } \partial\mathcal{T}_h,$$

where P_M is the L^2 -projection into M_h ; cf. the expression for $\hat{q}_h \cdot \mathbf{n}$ from (1.4).

Projections Π_V, Π_W tailored to the numerical traces have actually been widely used. Indeed, beginning with the simplest example, notice that the projections used to analyze both the RT and BDM methods capture the structure of its numerical

fluxes by satisfying

$$P_M(\mathbf{q} \cdot \mathbf{n}) = \Pi_V \mathbf{q} \cdot \mathbf{n} \quad \text{on } \partial\mathcal{T}_h,$$

This should be compared with the definition of their numerical trace [11]

$$\widehat{\mathbf{q}}_h \cdot \mathbf{n} = \mathbf{q}_h \cdot \mathbf{n} \quad \text{on } \partial\mathcal{T}_h.$$

Besides the projections used to carry out the analysis of the DG methods in [8], [6] and [13], we have the projections used in [9] to analyze an HDG type method called single-face hybridizable (SFH) method. The name arises due to the fact that for every simplex K , the penalty function τ is nonzero just on a single face, say F_K . The projections Π_V and Π_W used there are such that for each simplex $K \in \mathcal{T}_h$,

$$\begin{aligned} P_M(\mathbf{q} \cdot \mathbf{n}) &= \Pi_V \mathbf{q} \cdot \mathbf{n} && \text{on } \partial K \setminus F_K, \\ P_M u &= \Pi_W u && \text{on } F_K. \end{aligned}$$

This is in accordance with the numerical traces of the SFH method given by

$$\begin{aligned} \widehat{\mathbf{q}}_h \cdot \mathbf{n} &= \mathbf{q}_h \cdot \mathbf{n} && \text{on } \partial K \setminus F_K, \\ \widehat{\mathbf{q}}_h \cdot \mathbf{n} &= \mathbf{q}_h \cdot \mathbf{n} + \tau(u_h - \widehat{u}_h) && \text{on } F_K, \end{aligned}$$

The SFH method is also an HDG method of the LDG-H type, so our projection can also be used to analyze the SFH method. It is interesting to point out that the component $\Pi_V \mathbf{q}$ of the SFH projection introduced in [9] and ours do *coincide* when τ is as above, whereas the component $\Pi_W u$ *does not*, except when $\nabla \cdot \mathbf{q}$ is a polynomial of degree $k - 1$ on each simplex $K \in \mathcal{T}_h$. The final estimates for the SFH method are the same with both approaches.

The remainder of this paper is organized as follows. In Section 2, we define the new projection Π_h and state its key properties. In Section 3, we use an energy argument to obtain an optimal error estimate for the approximate flux \mathbf{q}_h and its numerical trace $\widehat{\mathbf{q}}_h$. In Section 4, we use a duality argument to obtain a superconvergence estimate of the projection of the error in u_h and its numerical trace \widehat{u}_h . In Section 5 we discuss previously known element-by-element postprocessing of the flux and the scalar variable. They will result in further approximations \mathbf{q}_h^* and u_h^* with interesting properties. Finally, we conclude in Section 6 by summarizing our results and relating them to the previous work [12]. Proof of the properties of the projection have been gathered in the Appendix.

2. THE PROJECTION

The main ingredient of our error analysis is a new projection Π_h into the product space $\mathbf{V}_h \times W_h$. In this section, we introduce it and establish its properties. From now on, we are going to use the following local spaces:

$$\mathbf{V}(K) = \mathcal{P}_k(K), \quad W(K) = \mathcal{P}_k(K) \quad \text{and} \quad M(F) = \mathcal{P}_k(F).$$

Also, we consider the stabilization function $\tau : \partial\mathcal{T}_h \rightarrow \mathbb{R}$ to be constant on each face.

The projected function is denoted by $\Pi_h(\mathbf{q}, u)$ or by $(\Pi_V \mathbf{q}, \Pi_W u)$ where $\Pi_V \mathbf{q}$ and $\Pi_W u$ are the components of the projection in \mathbf{V}_h and W_h , respectively. The values

of the projection on any simplex K are fixed by requiring that the components satisfy the equations

$$(2.1a) \quad (\mathbf{\Pi}_V \mathbf{q}, \mathbf{v})_K = (\mathbf{q}, \mathbf{v})_K \quad \text{for all } \mathbf{v} \in \mathcal{P}_{k-1}(K),$$

$$(2.1b) \quad (\Pi_W u, w)_K = (u, w)_K \quad \text{for all } w \in \mathcal{P}_{k-1}(K),$$

$$(2.1c) \quad \langle \mathbf{\Pi}_V \mathbf{q} \cdot \mathbf{n} + \tau \Pi_W u, \mu \rangle_F = \langle \mathbf{q} \cdot \mathbf{n} + \tau u, \mu \rangle_F \quad \text{for all } \mu \in \mathcal{P}_k(F),$$

for all faces F of the simplex K . If $k = 0$, then (2.1a) and (2.1b) are vacuous and Π_h is defined solely by (2.1c). Note that although we denoted the first component of the projection by $\mathbf{\Pi}_V \mathbf{q}$, it depends not just on \mathbf{q} , but rather on both \mathbf{q} and u , as we see from (2.1). The same is true for $\Pi_W u$. Hence the notation $(\mathbf{\Pi}_V \mathbf{q}, \Pi_W u)$ for $\Pi_h(\mathbf{q}, u)$ is somewhat misleading, but its convenience outweighs this disadvantage.

The domain of Π_h is a subspace of $L^2(\Omega)^n \times L^2(\Omega)$ on which the right-hand sides of (2.1) are well defined. Indeed, all functions \mathbf{q} and u that are regular enough for their traces $\mathbf{q} \cdot \mathbf{n}$ and u to be in $L^2(\partial K)$ are in the domain of Π_h , for example, $(\mathbf{q}, u) \in \mathbf{H}^1(\mathcal{T}_h) \times H^1(\mathcal{T}_h)$, where

$$H^1(\mathcal{T}_h) = \prod_{K \in \mathcal{T}_h} H^1(K), \quad \mathbf{H}^1(\mathcal{T}_h) = H^1(\mathcal{T}_h)^n.$$

That the left-hand sides of (2.1) uniquely determine $\mathbf{\Pi}_V \mathbf{q}$ and $\Pi_W u$ is proved as part of the next theorem.

To state the theorem, we need to introduce additional notation and conventions. We use $\|\cdot\|_D$ to denote the $L^2(D)$ -norm for any D . More generally, we denote the norm and seminorm on any Sobolev space X by $\|\cdot\|_X$ and $|\cdot|_X$, respectively. Also, as is usual in finite element analysis, we restrict elements to have shape regularity, i.e., all mesh elements K under consideration satisfy $h_K/\rho_K \leq \gamma$, where $h_K = \text{diam}(K)$, ρ_K is the diameter of the largest ball contained in K , and γ is a fixed constant. We use C , with or without subscripts, to denote a generic constant, independent of the elements and functions involved in our inequalities (but it may depend on γ). The value of C at different occurrences may differ. While we absorb dependencies on c into C , the dependencies on τ will always be explicitly mentioned.

Our first result states that the projection Π_h is well defined and has reasonable approximation properties. See Appendix A for a detailed proof.

Theorem 2.1. *Suppose $k \geq 0$, $\tau|_{\partial K}$ is nonnegative and $\tau_K^{\max} := \max \tau|_{\partial K} > 0$. Then the system (2.1) is uniquely solvable for $\mathbf{\Pi}_V \mathbf{q}_h$ and $\Pi_W u$. Furthermore, there is a constant C_Π independent of K and τ such that*

$$\begin{aligned} \|\mathbf{\Pi}_V \mathbf{q} - \mathbf{q}\|_K &\leq C h_K^{\ell_q+1} |\mathbf{q}|_{\mathbf{H}^{\ell_q+1}(K)} + C h_K^{\ell_u+1} \tau_K^* |u|_{H^{\ell_u+1}(K)}, \\ \|\Pi_W u - u\|_K &\leq C h_K^{\ell_u+1} |u|_{H^{\ell_u+1}(K)} + C \frac{h_K^{\ell_q+1}}{\tau_K^{\max}} |\nabla \cdot \mathbf{q}|_{H^{\ell_q}(K)}, \end{aligned}$$

for ℓ_u, ℓ_q in $[0, k]$. Here $\tau_K^* := \max \tau|_{\partial K \setminus F^*}$, where F^* is a face of K at which $\tau|_{\partial K}$ is maximum.

Note that if τ is of unit order, both approximation errors converge with the optimal order of $k + 1$, when the functions \mathbf{q} and u are smooth enough. If $\tau \equiv 1$, or if τ is bounded above and below uniformly by fixed constants, then a standard Bramble-Hilbert argument proves the order $k + 1$ approximation estimates of the theorem. However, to track the dependence of the constants on τ , and to compare with previous works, we need to perform a more careful analysis. This is done in

Appendix A in elaborate detail. To make an immediate comparison, note that for the τ used in the SFH method [9], $\tau_K^* = 0$ and the approximation properties of $\Pi_V \mathbf{q}$ are independent of the stabilization function τ .

We conclude this section with a property of the projection that we use critically in the error analysis of the method.

Proposition 2.1 (A weak commutativity property). *For any w in W_h and any (Φ, Ψ) in the domain of Π_h , we have*

$$(w, \nabla \cdot \Phi)_K = (w, \nabla \cdot \Pi_V \Phi)_K + \langle w, \tau(\Pi_W \Psi - \Psi) \rangle_{\partial K}.$$

Proof. For any $w \in W_h$,

$$\begin{aligned} (w, \nabla \cdot \Phi)_K &= -(\nabla w, \Phi)_K + \langle w, \Phi \cdot \mathbf{n} \rangle_{\partial K} \\ &= -(\nabla w, \Pi_V \Phi)_K + \langle w, \Phi \cdot \mathbf{n} \rangle_{\partial K} && \text{by (2.1a),} \\ &= -(\nabla w, \Pi_V \Phi)_K + \langle w, \Pi_V \Phi \cdot \mathbf{n} + \tau(\Pi_W \Psi - \Psi) \rangle_{\partial K} && \text{by (2.1c),} \\ &= (w, \nabla \cdot \Pi_V \Phi)_K + \langle w, \tau(\Pi_W \Psi - \Psi) \rangle_{\partial K}. \end{aligned}$$

□

3. FLUX ERROR ESTIMATES BY AN ENERGY ARGUMENT

The purpose of this section is to give error estimates under minimal regularity assumptions on the solution. Since the projection Π_h is designed to fit the structure of the numerical trace, the equations satisfied by the projection of the errors have a form amenable to simple analysis, as we see next. We will need the L^2 -orthogonal projection onto M_h , which we denote by P_M . Note that because τ is piecewise constant,

$$\langle \tau(P_M u - u), \mu \rangle_{\partial \mathcal{T}_h} = 0 \quad \text{for all } \mu \in M_h,$$

a fact that we will use repeatedly without explicit mention. The projection of the errors satisfy the following.

Lemma 3.1. *Let $\varepsilon_h^q = \Pi_V \mathbf{q} - \mathbf{q}_h$, $\varepsilon_h^u = \Pi_W u - u_h$, and $\varepsilon_h^{\hat{u}} = P_M u - \hat{u}_h$. Then*

$$(3.1a) \quad (c \varepsilon_h^q, \mathbf{r})_{\mathcal{T}_h} - (\varepsilon_h^u, \nabla \cdot \mathbf{r})_{\mathcal{T}_h} + \langle \varepsilon_h^{\hat{u}}, \mathbf{r} \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h} = (c(\Pi_V \mathbf{q} - \mathbf{q}), \mathbf{r})_{\mathcal{T}_h},$$

$$(3.1b) \quad -(\varepsilon_h^q, \nabla w)_{\mathcal{T}_h} + \langle \hat{\varepsilon}_h \cdot \mathbf{n}, w \rangle_{\partial \mathcal{T}_h} = 0,$$

$$(3.1c) \quad \langle \varepsilon_h^{\hat{u}}, \mu \rangle_{\partial \Omega} = 0,$$

$$(3.1d) \quad \langle \hat{\varepsilon}_h \cdot \mathbf{n}, \mu \rangle_{\partial \mathcal{T}_h \setminus \partial \Omega} = 0,$$

for all $\mathbf{r} \in \mathbf{V}_h$, $w \in W_h$, and $\mu \in M_h$, where

$$(3.1e) \quad \hat{\varepsilon}_h \cdot \mathbf{n} := \varepsilon_h^q \cdot \mathbf{n} + \tau(\varepsilon_h^u - \varepsilon_h^{\hat{u}}) = P_M(\mathbf{q} \cdot \mathbf{n}) - \hat{\mathbf{q}}_h \cdot \mathbf{n} \quad \text{on } \partial \mathcal{T}_h \setminus \partial \Omega.$$

Proof. Let us begin by noting that the exact solution \mathbf{q} and u satisfies

$$\begin{aligned} (c \mathbf{q}, \mathbf{r})_{\mathcal{T}_h} - (u, \nabla \cdot \mathbf{r})_{\mathcal{T}_h} + \langle u, \mathbf{r} \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h} &= 0, \\ -(\mathbf{q}, \nabla w)_{\mathcal{T}_h} + \langle \mathbf{q} \cdot \mathbf{n}, w \rangle_{\partial \mathcal{T}_h} &= (f, w)_{\mathcal{T}_h}, \end{aligned}$$

for all $\mathbf{r} \in \mathbf{V}_h$ and $w \in W_h$. By the definition of Π_h and P_M , the above implies

$$\begin{aligned} (c \Pi_V \mathbf{q}, \mathbf{r})_{\mathcal{T}_h} - (\Pi_W u, \nabla \cdot \mathbf{r})_{\mathcal{T}_h} + \langle P_M u, \mathbf{r} \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h} &= (c(\Pi_V \mathbf{q} - \mathbf{q}), \mathbf{r})_{\mathcal{T}_h}, \\ -(\Pi_V \mathbf{q}, \nabla w)_{\mathcal{T}_h} + \langle \Pi_V \mathbf{q} \cdot \mathbf{n} - \tau(u - \Pi_W u), w \rangle_{\partial \mathcal{T}_h} &= (f, w)_{\mathcal{T}_h}, \end{aligned}$$

for all $\mathbf{r} \in \mathbf{V}_h$ and $w \in W_h$. Subtracting (1.3a) and (1.3b) from the above two equations, respectively, we obtain (3.1a) and (3.1b). The equation (3.1c) follows directly from the boundary condition (1.3c). To prove (3.1d) we proceed as follows:

$$\begin{aligned} \langle \mu, \widehat{\boldsymbol{\varepsilon}}_h \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h \setminus \partial \Omega} &= \langle (\boldsymbol{\Pi}_V \mathbf{q} - \mathbf{q}_h) \cdot \mathbf{n} + \tau(\Pi_W u - u_h - P_M u + \widehat{u}_h), \mu \rangle_{\partial \mathcal{T}_h \setminus \partial \Omega} \\ &= \langle (\mathbf{q} - \mathbf{q}_h) \cdot \mathbf{n} + \tau(u - u_h - u + \widehat{u}_h), \mu \rangle_{\partial \mathcal{T}_h \setminus \partial \Omega} \\ &= \langle \mathbf{q} \cdot \mathbf{n}, \mu \rangle_{\partial \mathcal{T}_h \setminus \partial \Omega} - \langle \widehat{\mathbf{q}}_h \cdot \mathbf{n}, \mu \rangle_{\partial \mathcal{T}_h \setminus \partial \Omega}, \end{aligned}$$

where we have used the definition of Π_h . Note that this proves the identity (3.1e) because $\widehat{\mathbf{q}}_h \cdot \mathbf{n} \in M_h$. Since \mathbf{q} is in $H(\text{div}, \Omega)$ and since $\widehat{\mathbf{q}}_h$ satisfies (1.3d), both terms above are zero. This completes the proof. \square

Lemma 3.2. *We have*

$$(c \boldsymbol{\varepsilon}_h^q, \boldsymbol{\varepsilon}_h^q)_{\mathcal{T}_h} + \langle \tau(\varepsilon_h^u - \varepsilon_h^{\widehat{u}}), (\varepsilon_h^u - \varepsilon_h^{\widehat{u}}) \rangle_{\partial \mathcal{T}_h} = (c(\boldsymbol{\Pi}_V \mathbf{q} - \mathbf{q}), \boldsymbol{\varepsilon}_h^q)_{\mathcal{T}_h}.$$

Proof. Taking $\mathbf{r} := \boldsymbol{\varepsilon}_h^q$ in (3.1a), $w = \varepsilon_h^u$ in (3.1b), $\mu = -\widehat{\boldsymbol{\varepsilon}}_h \cdot \mathbf{n}$ in (3.1c) and $\mu = -\varepsilon_h^{\widehat{u}}$ in (3.1d), and adding the resulting four equations, we get

$$(c \boldsymbol{\varepsilon}_h^q, \boldsymbol{\varepsilon}_h^q)_{\mathcal{T}_h} + \Theta_h = (c(\boldsymbol{\Pi}_V \mathbf{q} - \mathbf{q}), \boldsymbol{\varepsilon}_h^q)_{\mathcal{T}_h},$$

where

$$\begin{aligned} \Theta_h &= -(\varepsilon_h^u, \nabla \cdot \boldsymbol{\varepsilon}_h^q)_{\mathcal{T}_h} + \langle \varepsilon_h^{\widehat{u}}, \boldsymbol{\varepsilon}_h^q \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h} - (\boldsymbol{\varepsilon}_h^q, \nabla \varepsilon_h^u)_{\mathcal{T}_h} \\ &\quad + \langle \widehat{\boldsymbol{\varepsilon}}_h \cdot \mathbf{n}, \varepsilon_h^u \rangle_{\partial \mathcal{T}_h} - \langle \varepsilon_h^{\widehat{u}}, \widehat{\boldsymbol{\varepsilon}}_h \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h}. \end{aligned}$$

But, by integration by parts, we get

$$\begin{aligned} \Theta_h &= -\langle \varepsilon_h^u, \boldsymbol{\varepsilon}_h^q \cdot \mathbf{n} \rangle_{\mathcal{T}_h} + \langle \varepsilon_h^{\widehat{u}}, \boldsymbol{\varepsilon}_h^q \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h} + \langle \widehat{\boldsymbol{\varepsilon}}_h \cdot \mathbf{n}, \varepsilon_h^u \rangle_{\partial \mathcal{T}_h} - \langle \varepsilon_h^{\widehat{u}}, \widehat{\boldsymbol{\varepsilon}}_h \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h} \\ &= -\langle \boldsymbol{\varepsilon}_h^q \cdot \mathbf{n}, \varepsilon_h^u - \varepsilon_h^{\widehat{u}} \rangle_{\mathcal{T}_h} + \langle \widehat{\boldsymbol{\varepsilon}}_h \cdot \mathbf{n}, \varepsilon_h^u - \varepsilon_h^{\widehat{u}} \rangle_{\partial \mathcal{T}_h} \\ &= \langle \tau(\varepsilon_h^u - \varepsilon_h^{\widehat{u}}), (\varepsilon_h^u - \varepsilon_h^{\widehat{u}}) \rangle_{\partial \mathcal{T}_h}, \end{aligned}$$

by the definition of $\widehat{\boldsymbol{\varepsilon}}_h$, namely (3.1e). This completes the proof. \square

While Lemma 3.2 identifies an “energy” norm, Lemma 3.1 gives “consistency” relations. These are enough to immediately prove an error estimate for the flux. To state it, we need the norm $\|\cdot\|_h$ defined by $\|\mu\|_h^2 = \sum_{K \in \mathcal{T}_h} h_K \|\mu\|_{\partial K}^2$ for any function $\mu \in L^2(\partial \mathcal{T}_h) := \prod_{K \in \mathcal{T}_h} L^2(\partial K)$, and the c -weighted $L^2(\Omega)$ -norm $\|\mathbf{q}\|_c := (c \mathbf{q}, \mathbf{q})_{\mathcal{T}_h}^{1/2}$.

Theorem 3.1 (Flux error estimates). *Let \mathbf{q}_h , u_h , and \widehat{u}_h solve the HDG equations (1.3) and let the exact solution \mathbf{q}, u be in the domain of Π_h . Then, for $k \geq 0$,*

$$(3.2) \quad \|\boldsymbol{\Pi}_V \mathbf{q} - \mathbf{q}_h\|_c \leq \|\boldsymbol{\Pi}_V \mathbf{q} - \mathbf{q}\|_c,$$

$$(3.3) \quad \|P_M(\mathbf{q} \cdot \mathbf{n}) - \widehat{\mathbf{q}}_h \cdot \mathbf{n}\|_h \leq C_{1,\tau} \|\boldsymbol{\Pi}_V \mathbf{q} - \mathbf{q}\|_c,$$

where $C_{1,\tau} = C \max\{1, (h_K \tau_K^{\max})^{1/2} : K \in \mathcal{T}_h\}$.

Proof. By the Cauchy-Schwarz inequality applied to the identity of Lemma 3.2, we get

$$(3.4) \quad (c \boldsymbol{\varepsilon}_h^q, \boldsymbol{\varepsilon}_h^q)_{\mathcal{T}_h} + \langle \tau(\varepsilon_h^u - \varepsilon_h^{\widehat{u}}), (\varepsilon_h^u - \varepsilon_h^{\widehat{u}}) \rangle_{\partial \mathcal{T}_h} \leq \|\boldsymbol{\Pi}_V \mathbf{q} - \mathbf{q}\|_c \|\boldsymbol{\varepsilon}_h^q\|_c,$$

and (3.2) follows since $\tau \geq 0$. For (3.3), we start by using the identity (3.1e) to get

$$\begin{aligned} \|P_M(\mathbf{q} \cdot \mathbf{n}) - \widehat{\mathbf{q}}_h \cdot \mathbf{n}\|_h &\leq \|\boldsymbol{\varepsilon}_h^q \cdot \mathbf{n}\|_h + \|\tau(\varepsilon_h^u - \widehat{\varepsilon}_h^u)\|_h \\ &\leq C \|\boldsymbol{\varepsilon}_h^q\|_c + \max_{K \in \mathcal{T}_h} (h_K \tau_K^{\max})^{1/2} \langle \tau(\varepsilon_h^u - \widehat{\varepsilon}_h^u), (\varepsilon_h^u - \widehat{\varepsilon}_h^u) \rangle_{\partial \mathcal{T}_h}^{1/2} \end{aligned}$$

where we used an inverse inequality and the fact that c^{-1} is uniformly bounded. Then (3.3) follows from (3.4). \square

4. SUPERCONVERGENCE OF THE SCALAR VARIABLE BY A DUALITY ARGUMENT

In the previous section, we established an error estimate for \mathbf{q}_h that only requires that the solution be as regular as required for the application of the projection. On domains permitting higher regularity estimates, we can perform an analogue of the Aubin-Nitsche duality argument [3, 17] to get higher rates of convergence. In particular, such arguments will give us error estimates for u_h and \widehat{u}_h .

We thus begin by introducing the dual problem for any given Θ in $L^2(\Omega)$:

$$\begin{aligned} (4.1a) \quad &c \boldsymbol{\Phi} - \nabla \Psi = 0 && \text{on } \Omega, \\ (4.1b) \quad &\nabla \cdot \boldsymbol{\Phi} = \Theta && \text{on } \Omega, \\ (4.1c) \quad &\Psi = 0 && \text{on } \partial\Omega. \end{aligned}$$

We assume that this boundary value problem admits the regularity estimate

$$(4.2) \quad \|\boldsymbol{\Phi}\|_{H^1(\Omega)} + \|\Psi\|_{H^2(\Omega)} \leq C_{\text{reg}} \|\Theta\|_{\Omega}$$

for all Θ in $L^2(\Omega)$. This is well known to hold in several cases, e.g., if $c \equiv 1$ and Ω is a convex polygon [16]. Recall that we have been tacitly assuming that (\mathbf{q}, u) is in the domain of Π_h . By (4.2), $(\boldsymbol{\Phi}, \Psi)$ is also regular enough to apply Π_h , so we have the following lemma.

Lemma 4.1 (The duality argument). *For any $\Psi_h \in W_h$, we have*

$$(\varepsilon_h^u, \Theta)_{\mathcal{T}_h} = (c(\mathbf{q} - \mathbf{q}_h), \Pi_V \boldsymbol{\Phi} - \boldsymbol{\Phi})_{\mathcal{T}_h} + (\mathbf{q} - \Pi_V \mathbf{q}, \nabla \Psi - \nabla \Psi_h)_{\mathcal{T}_h}.$$

Consequently,

$$\|\varepsilon_h^u\|_{\mathcal{T}_h} \leq C H(\Theta) \|\Pi_V \mathbf{q} - \mathbf{q}\|_{\mathcal{T}_h},$$

where

$$H(\Theta) := \sup_{\Theta \in L^2(\Omega) \setminus \{0\}} \frac{\|\Pi_V \boldsymbol{\Phi} - \boldsymbol{\Phi}\|_{\mathcal{T}_h}}{\|\Theta\|_{\mathcal{T}_h}} + \sup_{\Theta \in L^2(\Omega) \setminus \{0\}} \inf_{\Psi_h \in W_h} \frac{\|\nabla \Psi - \nabla \Psi_h\|_{\mathcal{T}_h}}{\|\Theta\|_{\mathcal{T}_h}}.$$

Proof. We have

$$\begin{aligned} (\varepsilon_h^u, \Theta)_{\mathcal{T}_h} &= (\varepsilon_h^u, \nabla \cdot \boldsymbol{\Phi})_{\mathcal{T}_h} && \text{by (4.1b),} \\ &= (\varepsilon_h^u, \nabla \cdot \Pi_V \boldsymbol{\Phi}) + \langle \varepsilon_h^u, \tau(\Pi_W \Psi - \Psi) \rangle_{\partial \mathcal{T}_h} && \text{by Prop. 2.1,} \\ &= (c \boldsymbol{\varepsilon}_h^q, \Pi_V \boldsymbol{\Phi})_{\mathcal{T}_h} + \langle \varepsilon_h^u, \Pi_V \boldsymbol{\Phi} \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h} \\ &\quad - (c(\Pi_V \mathbf{q} - \mathbf{q}), \Pi_V \boldsymbol{\Phi})_{\mathcal{T}_h} + \langle \varepsilon_h^u, \tau(\Pi_W \Psi - \Psi) \rangle_{\partial \mathcal{T}_h} && \text{by (3.1a),} \\ &= (c(\mathbf{q} - \mathbf{q}_h), \Pi_V \boldsymbol{\Phi})_{\mathcal{T}_h} + \langle \varepsilon_h^u, (\Pi_V \boldsymbol{\Phi} - \boldsymbol{\Phi}) \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h} \\ &\quad + \langle \varepsilon_h^u, \tau(\Pi_W \Psi - \Psi) \rangle_{\partial \mathcal{T}_h}, \end{aligned}$$

by the continuity of $\Phi \cdot \mathbf{n}$ and the fact that $\varepsilon_h^{\hat{u}} = 0$ on $\partial\Omega$ by (3.1c). Then

$$\begin{aligned} (\varepsilon_h^u, \Theta)_{\mathcal{T}_h} &= (c(\mathbf{q} - \mathbf{q}_h), \Pi_V \Phi)_{\mathcal{T}_h} + \langle \varepsilon_h^u - \varepsilon_h^{\hat{u}}, \tau(\Pi_W \Psi - \Psi) \rangle_{\partial\mathcal{T}_h} \quad \text{by (2.1c),} \\ &= (c(\mathbf{q} - \mathbf{q}_h), \Pi_V \Phi)_{\mathcal{T}_h} + \langle \tau(\varepsilon_h^u - \varepsilon_h^{\hat{u}}), \Pi_W \Psi \rangle_{\partial\mathcal{T}_h} \\ &\quad - \langle \varepsilon_h^u - \varepsilon_h^{\hat{u}}, \tau P_M \Psi \rangle_{\partial\mathcal{T}_h} \\ &= (c(\mathbf{q} - \mathbf{q}_h), \Pi_V \Phi)_{\mathcal{T}_h} + \langle \tau(\varepsilon_h^u - \varepsilon_h^{\hat{u}}), \Pi_W \Psi \rangle_{\partial\mathcal{T}_h} \\ &\quad - \langle \varepsilon_h^q \cdot \mathbf{n}, P_M \Psi \rangle_{\partial\mathcal{T}_h} \quad \text{by (3.1d),} \\ &= (c(\mathbf{q} - \mathbf{q}_h), \Pi_V \Phi)_{\mathcal{T}_h} + \langle \tau(\varepsilon_h^u - \varepsilon_h^{\hat{u}}), \Pi_W \Psi \rangle_{\partial\mathcal{T}_h} \\ &\quad - \langle \varepsilon_h^q \cdot \mathbf{n}, \Psi \rangle_{\partial\mathcal{T}_h}. \end{aligned}$$

Moreover,

$$\begin{aligned} (\varepsilon_h^u, \Theta)_{\mathcal{T}_h} &= (c(\mathbf{q} - \mathbf{q}_h), \Pi_V \Phi)_{\mathcal{T}_h} + (\nabla \cdot \varepsilon_h^q, \Pi_W \Psi)_{\mathcal{T}_h} - \langle \varepsilon_h^q \cdot \mathbf{n}, \Psi \rangle_{\partial\mathcal{T}_h} \quad \text{by (3.1b),} \\ &= (c(\mathbf{q} - \mathbf{q}_h), \Pi_V \Phi)_{\mathcal{T}_h} + (\nabla \cdot \varepsilon_h^q, \Psi)_{\mathcal{T}_h} - \langle \varepsilon_h^q \cdot \mathbf{n}, \Psi \rangle_{\partial\mathcal{T}_h} \quad \text{by (2.1b),} \\ &= (c(\mathbf{q} - \mathbf{q}_h), \Pi_V \Phi)_{\mathcal{T}_h} - (\varepsilon_h^q, \nabla \Psi)_{\mathcal{T}_h}. \end{aligned}$$

To bring this into the needed form, we continue:

$$\begin{aligned} (\varepsilon_h^u, \Theta)_{\mathcal{T}_h} &= (c(\mathbf{q} - \mathbf{q}_h), \Pi_V \Phi - \Phi)_{\mathcal{T}_h} + (c(\mathbf{q} - \mathbf{q}_h), \Phi)_{\mathcal{T}_h} \\ &\quad - (\Pi_V \mathbf{q} - \mathbf{q}_h, \nabla \Psi)_{\mathcal{T}_h} \\ &= (c(\mathbf{q} - \mathbf{q}_h), \Pi_V \Phi - \Phi)_{\mathcal{T}_h} + (\mathbf{q} - \Pi_V \mathbf{q}, \nabla \Psi)_{\mathcal{T}_h} \quad \text{by (4.1a),} \\ &= (c(\mathbf{q} - \mathbf{q}_h), \Pi_V \Phi - \Phi)_{\mathcal{T}_h} + (\mathbf{q} - \Pi_V \mathbf{q}, \nabla \Psi - \nabla \Psi_h)_{\mathcal{T}_h} \quad \text{by (2.1a).} \end{aligned}$$

Finally, the inequality of the lemma follows by applying the Cauchy-Schwarz inequality to the identity and using the first estimate of Theorem 3.1. \square

Theorem 4.1. *Suppose the regularity assumption (4.2) holds. Then*

$$\begin{aligned} \|\Pi_W u - u_h\|_{\mathcal{T}_h} &\leq C_{2,\tau} h^{\min\{k,1\}} \|\Pi_V \mathbf{q} - \mathbf{q}\|_{\mathcal{T}_h} \quad \text{for } k \geq 0, \\ \|P_M u - \hat{u}_h\|_h &\leq C_{2,\tau} h \|\Pi_V \mathbf{q} - \mathbf{q}\|_{\mathcal{T}_h} \quad \text{for } k \geq 1, \end{aligned}$$

where $h = \max\{h_K : K \in \mathcal{T}_h\}$ and $C_{2,\tau} = C \max\{1, h_K \tau_K^* : K \in \mathcal{T}_h\}$.

Proof. Let us prove the first inequality. By the first estimate of Theorem 2.1 with ℓ_q set to 0 and $\ell_u = \min\{k, 1\}$, we get that

$$\begin{aligned} H(\Theta) &\leq C h^{\min\{k,1\}} C_{2,\tau} \sup_{\Theta \in L^2(\Omega) \setminus 0} \frac{\|\Phi\|_{H^1(\Omega)} + \|\Psi\|_{H^{\ell_u+1}(\Omega)}}{\|\Theta\|_{\Omega}} \\ &\leq C h^{\min\{k,1\}} C_{2,\tau} C_{\text{reg}}. \end{aligned}$$

by the regularity estimate (4.2), and the first estimate follows.

The second estimate follows from the first from the same local argument used in [5] to obtain a similar estimate for the BDM method. Indeed, when $k \geq 1$, we can select a function $\mathbf{r} \in \mathcal{P}_k(K)$ such that $\mathbf{r} \cdot \mathbf{n} = \varepsilon_h^{\hat{u}}$ on ∂K and $\|\mathbf{r}\|_K \leq Ch_K^{1/2} \|\varepsilon_h^{\hat{u}}\|_{\partial K}$. Using $h_K \mathbf{r}$ as the test function in (3.1a), and applying an inverse inequality, we find that

$$\begin{aligned} h_K \|\varepsilon_h^{\hat{u}}\|_{\partial K}^2 &= h_K (c(\Pi_V \mathbf{q} - \mathbf{q}), \mathbf{r})_K + h_K (\varepsilon_h^u, \nabla \cdot \mathbf{r})_K - h_K (c\varepsilon_h^q, \mathbf{r})_K \\ &\leq Ch_K \|\mathbf{r}\|_K (\|\Pi_V \mathbf{q} - \mathbf{q}\|_K + \|\varepsilon_h^q\|_K) + C \|\mathbf{r}\|_K \|\varepsilon_h^u\|_K. \end{aligned}$$

Applying the first estimate of the theorem for ε_h^u and the estimate in Theorem 3.1 for ε_h^q , we get the second inequality of the theorem. This completes the proof. \square

5. ENHANCED ACCURACY BY POSTPROCESSING

In this section we describe a few techniques to postprocess the approximate solution and flux.

5.1. Flux postprocessing. We can obtain a postprocessed flux \mathbf{q}_h^* with better conservation properties. Although \mathbf{q}_h^* converges at the same order as \mathbf{q}_h , it is in $H(\operatorname{div}, \Omega)$ and its divergence converges at one higher order than \mathbf{q}_h .

To define \mathbf{q}_h^* , we use a slight modification of the Raviart-Thomas projection [18], as used in the framework of Darcy flows [4], or for the Navier-Stokes equations [14], or for the postprocessing of superconvergent DG methods for second-order elliptic problems [12]. We follow [12]. Thus, on each simplex $K \in \mathcal{T}_h$, we take $\mathbf{q}_h^* := \mathbf{q}_h + \boldsymbol{\eta}_h$ where $\boldsymbol{\eta}_h$ is the only element of $\mathcal{P}_k(K) + \mathbf{x} \mathcal{P}_k(K)$ satisfying

$$\begin{aligned} (\boldsymbol{\eta}_h, \mathbf{v})_K &= 0 && \text{for all } \mathbf{v} \in \mathcal{P}_{k-1}(K), \\ \langle \boldsymbol{\eta}_h \cdot \mathbf{n}, \mu \rangle_F &= \langle (\widehat{\mathbf{q}}_h - \mathbf{q}_h) \cdot \mathbf{n}, \mu \rangle_F && \text{for all } \mu \in \mathcal{P}_k(F) \text{ and all faces } F \text{ of } K. \end{aligned}$$

If $L^2(K)$ -orthogonal hierarchical basis functions are used, computation of $\boldsymbol{\eta}_h$ reduces to solving, on each element, a linear system of size $n+1$ times the dimension of the space of homogeneous polynomials of degree k . It is instructive to compare the above definition to that of the Raviart-Thomas projection, namely on each simplex $K \in \mathcal{T}_h$, we set $\boldsymbol{\Pi}_k^{\text{RT}} \mathbf{q}$ as the only element of $\mathcal{P}_k(K) + \mathbf{x} \mathcal{P}_k(K)$ satisfying

$$\begin{aligned} (\boldsymbol{\Pi}_k^{\text{RT}} \mathbf{q} - \mathbf{q}, \mathbf{v})_K &= 0 && \text{for all } \mathbf{v} \in \mathcal{P}_{k-1}(K), \\ \langle (\boldsymbol{\Pi}_k^{\text{RT}} \mathbf{q} - \mathbf{q}) \cdot \mathbf{n}, \mu \rangle_F &= 0, && \text{for all } \mu \in \mathcal{P}_k(F) \text{ and all faces } F \text{ of } K. \end{aligned}$$

This comparison yields the following theorem. For a proof, see [12].

Theorem 5.1. *For any $k \geq 0$, we have that $\mathbf{q}_h^* \in H(\operatorname{div}, \Omega)$. Moreover,*

$$\begin{aligned} \|\mathbf{q} - \mathbf{q}_h^*\|_\Omega &\leq C \|\mathbf{q} - \boldsymbol{\Pi}_k^{\text{RT}} \mathbf{q}\|_\Omega + C_{1,\tau} \|\boldsymbol{\Pi}_v \mathbf{q} - \mathbf{q}\|_{\mathcal{T}_h}, \\ \|\nabla \cdot (\mathbf{q} - \mathbf{q}_h^*)\|_\Omega &= \inf_{f_h \in W_h} \|f - f_h\|_{\mathcal{T}_h}. \end{aligned}$$

5.2. Postprocessing the approximate scalar variable. There are a few well-known ways to postprocess to obtain a new approximation u_h^* of enhanced accuracy.

As the first postprocessing method, we define $u_{h,1}^*$ in $\mathcal{P}_{k+1}(K)$ satisfying

$$(5.1a) \quad (c^{-1} \nabla u_{h,1}^*, \nabla w)_K = (f, w)_K - \langle \mathbf{q}_h \cdot \mathbf{n}, w \rangle_{\partial K}, \quad \text{for all } w \in \mathcal{P}_{k+1}^0(K),$$

$$(5.1b) \quad m_K(u_h^*) = m_K(u_h),$$

where $m_K(v) = \operatorname{meas}(K)^{-1}(v, 1)_K$ and $\mathcal{P}_{k+1}^0(K) = \{p \in \mathcal{P}_{k+1}(K) : m_K(p) = 0\}$. Clearly, $u_{h,1}^*$ satisfies a discrete Neumann problem on each element with the computed approximate solution as data. This method was introduced in [15, 19, 20] in the framework of mixed methods.

A variation of this postprocessing was proposed in [9, 12] for DG methods. It is obtained simply by substituting $\widehat{\mathbf{q}}_h$ in place of \mathbf{q}_h on the right-hand side of (5.1a). This yields our second postprocessing alternative. The solution so obtained is denoted by $u_{h,2}^*$.

As a third alternative, we define $u_{h,3}^*$ obtained by following [19, 20]. Let $\mathcal{W}_{k+1}(K)$ denote the $L^2(K)$ -orthogonal complement of $\mathcal{P}_{k-1}(K)$ in $\mathcal{P}_{k+1}(K)$. The solution $u_{h,3}^*$ is of the form $u_h + \eta_h$ where η_h is the unique function in $\mathcal{W}_{k+1}(K)$ satisfying

$$(5.2) \quad (\nabla \eta_h, \nabla w)_K = -(\nabla u_h + c \mathbf{q}_h, \nabla w)_K \quad \text{for all } w \in \mathcal{W}_{k+1}(K).$$

As in the case of the flux postprocessing, if we are using an $L^2(K)$ -orthogonal hierarchical basis to find η_h , we need only invert a symmetric, positive definite matrix whose order is the dimension of $\mathcal{W}_{k+1}(K)$. Note also that to evaluate the right-hand side of (5.2), we need only use $n - 1$ -dimensional quadratures, as

$$-(\nabla u_h + c \mathbf{q}_h, \nabla w)_{\mathcal{T}_h} = \langle \hat{u}_h - u_h, \nabla w \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h \setminus \partial \Omega} + \langle g - u_h, \nabla w \cdot \mathbf{n} \rangle_{\partial \Omega}.$$

This follows by setting $\mathbf{r} = \nabla w$ in (1.3a), the first equation of the HDG method.

All the above postprocessed solutions converge at a higher rate than u_h (whenever $k \geq 1$) as stated in the next theorem. It can be proved using the superconvergence estimate of Theorem 4.1 in a standard way (see [20]) so we omit it.

Theorem 5.2. *Under the same assumption as Theorem 4.1, the result $u_{h,i}^*$ of any of the three above mentioned postprocessings satisfy*

$$\|u - u_{h,i}^*\|_{\mathcal{T}_h} \leq (C_{2,\tau} + \delta_{i2} C_{1,\tau}) h^{\min\{k,1\}} \|\Pi_V \mathbf{q} - \mathbf{q}\|_{\mathcal{T}_h} + C h^{\ell+2} \|u\|_{H^{\ell+2}(\mathcal{T}_h)},$$

for any $k \geq 0$ and $i = 1, 2, 3$, and any $\ell \in [0, k]$.

6. CONCLUDING REMARKS

We have presented a technique for error analysis of HDG methods that is remarkable for its brevity, especially in comparison with previous DG analyses [2, 9, 12]. We achieved this through the use of a new projection Π_h . While brevity and elegance is traditionally achieved in the analysis of mixed methods, like RT and BDM methods, via the use of projections with commutativity properties, in the case of DG methods, commutativity seems not to be of paramount importance. Rather, what seems important is a projection tailored to fit the structure of the DG method, such as our Π_h .

Because Π_h is adapted to the structure of our numerical traces, we found it easy to estimate the projection of the errors. To summarize, we proved,

$$\begin{aligned} \|\Pi_V \mathbf{q} - \mathbf{q}_h\|_c &\leq \|\Pi_V \mathbf{q} - \mathbf{q}\|_c, \\ \|P_M(\mathbf{q} \cdot \mathbf{n}) - \hat{\mathbf{q}}_h \cdot \mathbf{n}\|_h &\leq C_{1,\tau} \|\Pi_V \mathbf{q} - \mathbf{q}\|_c, \\ \|\Pi_W u - u_h\|_{\mathcal{T}_h} &\leq C_{2,\tau} h^{\min\{k,1\}} \|\Pi_V \mathbf{q} - \mathbf{q}\|_{\mathcal{T}_h}, \\ \|P_M u - \hat{u}_h\|_h &\leq C_{2,\tau} h \|\Pi_V \mathbf{q} - \mathbf{q}\|_{\mathcal{T}_h}, \end{aligned}$$

where the first three estimates hold for all $k \geq 0$ and the last for $k \geq 1$. Thus, by the approximation properties of the projection Π_h (Theorem 2.1), if the penalty function τ is such that τ_K^{\max} is of order one on each $K \in \mathcal{T}_h$, we obtain the *optimal* order of convergence of $k + 1$ for the approximate flux and its numerical trace. Of course, by the triangle inequality, the above estimates imply that the error of these variables converges to zero at the optimal order. If $k \geq 1$, the projection of the errors for u and its trace superconverge at order $k + 2$. This can be exploited to get locally postprocessed solutions of enhanced accuracy.

To end, note that the above estimates imply that

$$\|(\hat{\mathbf{q}}_h - \mathbf{q}_h) \cdot \mathbf{n}\|_h = O(h^{k+1}).$$

According to the main result in [12], such an inequality implies optimal order of convergence for the numerical flux \mathbf{q}_h and its postprocessing \mathbf{q}_h^* for $k \geq 0$. It also implies, for $k \geq 1$, the superconvergence of the orthogonal projection of the error in u into $\mathcal{P}_{k-1}(K)$, and furthermore the superconvergence of \hat{u}_h . This is in perfect agreement with our results.

The extension of our approach to other equations of practical interest appearing in, for example, fluid flow and solid mechanics, can prove to be useful not only to analyze already existing HDG methods but also to devise new ones. This constitutes the objective of ongoing research.

APPENDIX A. PROOF OF THEOREM 2.1

We begin by observing that once we prove the approximation estimates of the theorem, then the unisolvency of the equations defining the projection follows as a corollary. This is because the number of equations and unknowns in (2.1), namely

$$\dim(\mathcal{P}_{k-1}(K)) + \dim(\mathcal{P}_{k-1}(K)) + (n+1)\dim(\mathcal{P}_k(F))$$

and

$$\dim(\mathcal{P}_k(K)) + \dim(\mathcal{P}_k(K)),$$

respectively, coincide, i.e. (2.1) is a square linear system. Hence setting $u = 0$ and $\mathbf{q} = 0$ in the approximation estimates we find that the projection must vanish.

In view of this, in the remainder of this section, we develop estimates for *any* $\Pi_V \mathbf{q}$ and $\Pi_W u$ satisfying (2.1) without assuming uniqueness *a priori* (although it will follow *a posteriori*). But first, we begin with two auxiliary lemmas.

A.1. Two estimates involving orthogonal polynomials. Let

$$\mathcal{P}_k^\perp(K) := \{w \in \mathcal{P}_k(K) : (w, \zeta)_K = 0 \forall \zeta \in \mathcal{P}_{k-1}(K)\}.$$

Lemma A.1. *Let F be any face of a simplex K . The trace map*

$$\gamma_F : \mathcal{P}_k^\perp(K) \mapsto \mathcal{P}_k(F) \quad \text{defined by} \quad \gamma_F(p) = p|_F$$

is a bijection. Moreover,

$$\|p\|_K \leq Ch_K^{1/2} \|p\|_F \quad \text{for all } p \in \mathcal{P}_k^\perp(K).$$

Proof. We first prove that γ_F is injective. Suppose $\gamma_F(p) = 0$ for some $p \in \mathcal{P}_k^\perp(K)$. Then we can write $p = \lambda_F q$, where λ_F denotes the barycentric coordinate function of K that vanishes on F and q is some function in $\mathcal{P}_{k-1}(K)$. But since $(p, w)_K = 0$ for all $w \in \mathcal{P}_{k-1}(K)$, we then have $(\lambda_F q, q)_K = 0$, so $q = 0$ and hence $p = 0$. The surjectivity of γ_F now follows by counting dimensions and using the injectivity.

Finally, the estimate of the lemma follows from the injectivity and a standard scaling argument. \square

Lemma A.2. *Let η be a nonnegative function on ∂K , constant on each face of K , and such that $\eta^{\max} := \max \eta > 0$. Let $p \in \mathcal{P}_k^\perp(K)$ satisfy the equation*

$$\langle \eta p, w \rangle_{\partial K} = b(w) \quad \text{for all } w \in \mathcal{P}_k^\perp(K),$$

where $b : \mathcal{P}_k^\perp(K) \rightarrow \mathbb{R}$ is linear. Then

$$\|p\|_K \leq C \frac{h_K}{\eta^{\max}} \|b\|,$$

where $\|b\| := \sup_{w \in \mathcal{P}_k^\perp(K) \setminus \{0\}} b(w) / \|w\|_K$.

Proof. Let F be a face of K at which $\eta = \eta^{\max}$. Then, by Lemma A.1,

$$\begin{aligned} \|p\|_K^2 &\leq C h_K \|p\|_F^2 = C \frac{h_K}{\eta^{\max}} \langle \eta p, p \rangle_F \leq C \frac{h_K}{\eta^{\max}} \langle \eta p, p \rangle_{\partial K}, \quad \text{since } \eta \geq 0, \\ &= C \frac{h_K}{\eta^{\max}} b(p) \leq C \frac{h_K}{\eta^{\max}} \|b\| \|p\|_K, \end{aligned}$$

and the wanted estimate follows. □

A.2. Decoupling the projection component $\Pi_W u$. Now we characterize the second (scalar) component of the projection $\Pi_h(\mathbf{q}, u) \equiv (\mathbf{\Pi}_V \mathbf{q}, \Pi_W u)$, namely $\Pi_W u$, independently of the first, and prove its approximation properties.

Proposition A.1. *On each element $K \in \mathcal{T}_h$, the component $\Pi_W u$ satisfies*

(A.1a) $\quad (\Pi_W u, v)_K = (u, v)_K \quad \text{for all } v \in \mathcal{P}_{k-1}(K),$

(A.1b) $\quad \langle \tau \Pi_W u, w \rangle_{\partial K} = \langle \nabla \cdot \mathbf{q}, w \rangle_K + \langle \tau u, w \rangle_{\partial K} \quad \text{for all } w \in \mathcal{P}_k^\perp(K).$

Proof. The first equation (A.1a) is the same as an equation defining the projection (2.1b). For the second, note that (2.1c) implies

$$\langle \tau \Pi_W u, w \rangle_{\partial K} = \langle (\mathbf{q} - \mathbf{\Pi}_V \mathbf{q}) \cdot \mathbf{n} + \tau u, w \rangle_{\partial K} \quad \text{for all } w \in \mathcal{P}_k^\perp(K).$$

Simplifying the right-hand side using

$$\begin{aligned} \langle (\mathbf{q} - \mathbf{\Pi}_V \mathbf{q}) \cdot \mathbf{n}, w \rangle_{\partial K} &= \langle \nabla \cdot (\mathbf{q} - \mathbf{\Pi}_V \mathbf{q}), w \rangle_K + \langle \mathbf{q} - \mathbf{\Pi}_V \mathbf{q}, \nabla w \rangle_K \\ &= \langle \nabla \cdot (\mathbf{q} - \mathbf{\Pi}_V \mathbf{q}), w \rangle_K && \text{by (2.1a),} \\ &= \langle \nabla \cdot \mathbf{q}, w \rangle_K && \text{as } w \in \mathcal{P}_k^\perp(K), \end{aligned}$$

we finish the proof. □

Proposition A.1 permits comparison with the SFH method. Suppose τ is selected as in the SFH method and suppose $\nabla \cdot \mathbf{q} \in \mathcal{P}_{k-1}(K)$. Then the system (A.1) becomes

$$\begin{aligned} (\Pi_W u, v)_K &= (u, v)_K \quad \text{for all } v \in \mathcal{P}_{k-1}(K), \\ \langle \tau \Pi_W u, \mu \rangle_{F_K} &= \langle \tau u, \mu \rangle_{F_K} \quad \text{for all } \mu \in \mathcal{P}_k(F_K). \end{aligned}$$

Note that to obtain the last equation, we used the surjectivity of γ_F (Lemma A.1). Thus, in this case, Π_W coincides with the projection used in the analysis of the SFH method [9] (denoted there by \mathbb{P}).

We are now ready to obtain the estimate of $\Pi_W u - u$ in Theorem 2.1.

Proposition A.2. *Suppose the assumptions on τ in Theorem 2.1 hold. Then,*

$$\|\Pi_W u - u\|_K \leq C h_K^{\ell_u + 1} |u|_{H^{\ell_u + 1}(K)} + C \frac{h_K^{\ell_q + 1}}{\tau_K^{\max}} |\nabla \cdot \mathbf{q}|_{H^{\ell_q}(K)},$$

for ℓ_u, ℓ_q in $[0, k]$ (and consequently, $\Pi_W u$ is uniquely determined by (A.1)).

Proof. To prove the result, we set $\delta^u := \Pi_W u - u_k$, where u_k is the $L^2(K)$ -orthogonal projection of u into $\mathcal{P}_k(K)$, and note that

$$\|\Pi_W u - u\|_K \leq \|u - u_k\|_K + \|\delta^u\|_K.$$

The first term can be readily estimated by using the standard approximation properties of the L^2 -projection. Let us estimate the second term.

Equation (A.1a) shows that δ^u belongs to $\mathcal{P}_k^\perp(K)$, and (A.1b) implies

$$\langle \tau \delta^u, w \rangle_{\partial K} = b_{\mathbf{q}}(w) + b_u(w) \quad \text{for all } w \in \mathcal{P}_k^\perp(K),$$

where $b_{\mathbf{q}}(w) := (\nabla \cdot \mathbf{q}, w)_K$ and $b_u(w) := \langle \tau(u - u_k), w \rangle_{\partial K}$. By Lemma A.2 with $\eta := \tau$, $p := \delta^u$ and $b = b_{\mathbf{q}} + b_u$, this implies that

$$\|\delta^u\|_K \leq C \frac{h_K}{\tau_K^{\max}} (\|b_{\mathbf{q}}\| + \|b_u\|).$$

Let us estimate $\|b_{\mathbf{q}}\|$. Since $w \in \mathcal{P}_k^\perp(K)$, we have

$$b_{\mathbf{q}}(w) = (\nabla \cdot \mathbf{q} - (\nabla \cdot \mathbf{q})_{k-1}, w)_K,$$

where $(\nabla \cdot \mathbf{q})_{k-1}$ is the $L^2(K)$ -projection of $\nabla \cdot \mathbf{q}$ into $\mathcal{P}_{k-1}(K)$ when $k \geq 1$. If $k = 0$, we set $(\nabla \cdot \mathbf{q})_{-1} \equiv 0$. Hence,

$$\|b_{\mathbf{q}}\| \leq \|\nabla \cdot \mathbf{q} - (\nabla \cdot \mathbf{q})_{k-1}\|_K \leq C h_K^{\ell_{\mathbf{q}}} |\nabla \cdot \mathbf{q}|_{H^{\ell_{\mathbf{q}}}(K)},$$

for $\ell_{\mathbf{q}}$ in $[0, k]$, by the approximation properties of the L^2 -projection.

Finally, let us estimate $\|b_u\|$. By a scaling argument,

$$b_u(w) \leq \tau_K^{\max} \|u - u_k\|_{\partial K} \|w\|_{\partial K} \leq C h_K^{-1/2} \tau_K^{\max} \|u - u_k\|_{\partial K} \|w\|_K,$$

A trace inequality and the approximation properties of the L^2 -projection imply

$$\|b_u\| \leq C \tau_K^{\max} h_K^{-1} (\|u - u_k\|_K + h_K |u - u_k|_{H^1(K)}) \leq C \tau_K^{\max} h_K^{\ell_u} |u|_{H^{\ell_u+1}(K)},$$

for any ℓ_u in $[0, k]$. This completes the proof. \square

A.3. Properties of the flux component $P_V \mathbf{q}$. To study the flux component, we recall another projection \mathbf{B}_V introduced and studied in [8] and later used in [9, 12]. Let F^* be a face of K at which $\tau|_{\partial K}$ is a maximum. For any function \mathbf{q} in the domain of $\mathbf{\Pi}_V$, the restriction of $\mathbf{B}_V \mathbf{q}$ to K is defined to be the unique element of $\mathcal{P}^k(K)$ satisfying

$$(A.2a) \quad (\mathbf{B}_V \mathbf{q}, \mathbf{v})_K = (\mathbf{q}, \mathbf{v})_K \quad \forall \mathbf{v} \in \mathcal{P}^{k-1}(K),$$

$$(A.2b) \quad \langle \mathbf{B}_V \mathbf{q} \cdot \mathbf{n}, \mu \rangle_F = \langle \mathbf{q} \cdot \mathbf{n}, \mu \rangle_F \quad \forall \mu \in \mathcal{P}^k(F),$$

for all faces F of K different from F^* . We use it to prove the following lemma. Clearly, the proof of Theorem 2.1 would be complete once we prove the lemma.

Proposition A.3. *Suppose the assumptions on τ in Theorem 2.1 hold. Then,*

$$\|\mathbf{\Pi}_V \mathbf{q} - \mathbf{q}\|_K \leq C h_K^{\ell_{\mathbf{q}}+1} |\mathbf{q}|_{H^{\ell_{\mathbf{q}}+1}(K)} + C h_K^{\ell_u+1} \tau_K^* |u|_{H^{\ell_u+1}(K)},$$

for $\ell_u, \ell_{\mathbf{q}}$ in $[0, k]$.

Proof. A basis for \mathbb{R}^n is furnished by the set of unit normals \mathbf{n}_F for the n faces $F \neq F^*$ of K . Letting $\{\tilde{\mathbf{n}}_F : F \neq F^*\}$ denote its dual basis, that is, $\tilde{\mathbf{n}}_F \cdot \mathbf{n}_{F'} = \delta_{FF'}$, we can write $\mathbf{\Pi}_V \mathbf{q} - \mathbf{q} = \sum_{F \neq F^*} ((\mathbf{\Pi}_V \mathbf{q} - \mathbf{q}) \cdot \mathbf{n}_F) \tilde{\mathbf{n}}_F$. Hence, it is enough to estimate each of the functions $(\mathbf{\Pi}_V \mathbf{q} - \mathbf{q}) \cdot \mathbf{n}_F$. Moreover, since we have

$$\|(\mathbf{\Pi}_V \mathbf{q} - \mathbf{q}) \cdot \mathbf{n}_F\|_K \leq \|(\mathbf{B}_V \mathbf{q} - \mathbf{q}) \cdot \mathbf{n}_F\|_K + \|(\mathbf{\Pi}_V \mathbf{q} - \mathbf{B}_V \mathbf{q}) \cdot \mathbf{n}_F\|_K,$$

by the approximation properties of \mathbf{B}_V established in [8], it is enough to estimate the function $\delta_F^{\mathbf{q}} := (\mathbf{\Pi}_V \mathbf{q} - \mathbf{B}_V \mathbf{q}) \cdot \mathbf{n}_F$.

From equations (2.1a) and (A.2a), we conclude that $\delta_F^{\mathbf{q}}$ is in $\mathcal{P}_k^\perp(K)$. Subtracting (A.2b) from (2.1c), we obtain

$$(A.3) \quad \langle \delta_F^{\mathbf{q}}, \mu \rangle_F = \langle \tau(u - \mathbf{\Pi}_V u), \mu \rangle_F \quad \text{for all } \mu \in \mathcal{P}_k(F), \text{ and any } F \neq F^*.$$

By Lemma A.1, $w = \gamma_F^{-1}(\mu)$ is in $\mathcal{P}_k^\perp(K)$ and $w = \mu$ on F . Hence,

$$\langle \eta \delta_F^{\mathbf{q}}, w \rangle_{\partial K} = b(w) \quad \text{for all } w \in \mathcal{P}_k^\perp(K),$$

where η is the characteristic function of F and $b(w) := \langle \tau \eta (u - \Pi_w u), w \rangle_{\partial K}$. By Lemma A.2, this implies that

$$\|\delta_F^{\mathbf{q}}\|_K \leq C h_K \|b\|.$$

It only remains to estimate $\|b\|$. But

$$\begin{aligned} |b(w)| &\leq C \tau_K^* h_K^{-1/2} \|u - \Pi_w u\|_{\partial K} \|w\|_K \\ &\leq C \tau_K^* h_K^{-1/2} \|u - u_k\|_{\partial K} \|w\|_K + C \tau_K^* h_K^{-1} \|u_k - \Pi_w u\|_K \|w\|_K, \end{aligned}$$

where u_k is the L^2 -projection of u into $\mathcal{P}_k(K)$. Hence, since $\tau_K^* \leq \tau_K^{\max}$,

$$\|b\| \leq C h_K^{\ell_q} |\nabla \cdot \mathbf{q}|_{H^{\ell_q}(K)} + C \tau_K^* h_K^{\ell_u} \|u\|_{H^{\ell_u+1}(K)},$$

for ℓ_u, ℓ_q in $[0, k]$, by the approximation properties of the L^2 -projection and the estimates of Proposition A.2. This proves the required estimate for $\delta_F^{\mathbf{q}}$ and completes the proof. \square

We have thus finished the proof of Theorem 2.1. Before closing, let us state one more result that characterizes $\Pi_V \mathbf{q}$ and compare with the case of the SFH method.

Proposition A.4. *Suppose that the assumptions on τ of Theorem 2.1 hold. Then, on each simplex $K \in \mathcal{T}_h$, $\Pi_V \mathbf{q}$ is the only element of $\mathcal{P}_k(K)$ such that*

$$(A.4a) \quad (\Pi_V \mathbf{q}, \mathbf{v})_K = (\mathbf{q}, \mathbf{v})_K \quad \text{for all } \mathbf{v} \in \mathcal{P}_{k-1}(K),$$

$$(A.4b) \quad \langle \Pi_V \mathbf{q} \cdot \mathbf{n}, \mu \rangle_F = \langle \mathbf{q} \cdot \mathbf{n} + \tau (u - \Pi_w u), \mu \rangle_F \quad \text{for all } \mu \in \mathcal{P}_k(F),$$

and all faces F of K except one arbitrarily chosen.

Note that for the SFH method, taking all faces F where $\tau = 0$, we immediately see that Π_V coincides with the projection in the analysis of that method in [9]. It was denoted therein by Π_V and is nothing but the projection \mathbf{B}_V given by (A.2).

Proof. The component $\Pi_V \mathbf{q}$ obviously satisfies (A.4) as these equations are identical to (2.1a) and (2.1c). Since the system (A.4) is square, we need only show uniqueness. So, let us set the right-hand side equal to zero. For any face F , consider the function $\delta_F := \Pi_V \mathbf{q} \cdot \mathbf{n}_F$. Then by equation (A.4a) with $\mathbf{v} := \mathbf{n}_F v$ with $v \in \mathcal{P}_{k-1}(K)$, equation (A.4b), and Lemma A.1, we get that $\delta_F = 0$. Since this can be done for all faces but one, we readily obtain that $\Pi_V \mathbf{q} = \mathbf{0}$. This completes the proof. \square

ACKNOWLEDGEMENTS

The authors would like to thank the referees for constructive criticism leading to a better version of this paper.

REFERENCES

- [1] D. N. ARNOLD AND F. BREZZI, *Mixed and nonconforming finite element methods: implementation, postprocessing and error estimates*, RAIRO Modél. Math. Anal. Numér. **19** (1985), 7–32. MR813687 (87g:65126)
- [2] D. N. ARNOLD, F. BREZZI, B. COCKBURN, AND L. D. MARINI, *Unified analysis of discontinuous Galerkin methods for elliptic problems*, SIAM J. Numer. Anal. **39** (2002), 1749–1779. MR1885715 (2002k:65183)
- [3] J.-P. AUBIN, *Behavior of the error of the approximate solutions of boundary value problems for linear elliptic operators by Galerkin's and finite difference methods*, Ann. Scuola Norm. Sup. Pisa (3) **21** (1967), 599–637. MR0233068 (38:1391)
- [4] P. BASTIAN AND B. RIVIÈRE, *Superconvergence and $H(\text{div})$ projection for discontinuous Galerkin methods*, Internat. J. Numer. Methods Fluids **42** (2003), 1043–1057. MR1991232 (2004f:65177)
- [5] F. BREZZI, J. DOUGLAS, JR., AND L. D. MARINI, *Two families of mixed finite elements for second order elliptic problems*, Numer. Math., 47 (1985), pp. 217–235. MR799685 (87g:65133)
- [6] F. CELIKER AND B. COCKBURN, *Superconvergence of the numerical traces of discontinuous Galerkin and hybridized mixed methods for convection-diffusion problems in one space dimension*, Math. Comp. **76** (2007), 67–96. MR2261012 (2008e:65225)
- [7] P. CIARLET, *The finite element method for elliptic problems*, North-Holland, Amsterdam, 1978. MR0520174 (58:25001)
- [8] B. COCKBURN AND B. DONG, *An analysis of the minimal dissipation local discontinuous Galerkin method for convection-diffusion problems*, J. Sci. Comput. **32** (2007), 233–262. MR2320571 (2008e:65345)
- [9] B. COCKBURN, B. DONG, AND J. GUZMÁN, *A superconvergent LDG-hybridizable Galerkin method for second-order elliptic problems*, Math. Comp. **77** (2008), 1887–1916. MR2429868 (2009d:65166)
- [10] B. COCKBURN, J. GOPALAKRISHNAN, *Error analysis of variable degree mixed methods for elliptic problems via hybridization*, Math. Comp. **74** (2005), 1653–1677. MR2164091 (2006e:65215)
- [11] B. COCKBURN, J. GOPALAKRISHNAN, AND R. LAZAROV, *Unified hybridization of discontinuous Galerkin, mixed, and conforming Galerkin methods for second order elliptic problems*, SIAM J. Numer. Anal. **47** (2009), 1319–1365. MR2485455
- [12] B. COCKBURN, J. GUZMÁN, AND H. WANG, *Superconvergent discontinuous Galerkin methods for second-order elliptic problems*, Math. Comp. **78** (2009), 1–24. 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** (2001), 264–285. MR1860725 (2002g:65140)
- [14] B. COCKBURN, G. KANSCHAT, AND D. SCHÖTZAU, *A locally conservative LDG method for the incompressible Navier-Stokes equations*, Math. Comp. **74** (2005), 1067–1095. MR2136994 (2006a:65157)
- [15] L. GASTALDI AND R. H. NOCHETTO, *Sharp maximum norm error estimates for general mixed finite element approximations to second order elliptic equations*, RAIRO Modél. Math. Anal. Numér. **23** (1989), 103–128. MR1015921 (91b:65125)
- [16] P. GRISVARD, *Elliptic Problems in Nonsmooth Domains*, no. 24 in Monographs and Studies in Mathematics, Pitman Advanced Publishing Program, Marshfield, Massachusetts, 1985. MR775683 (86m:35044)
- [17] J. NITSCHKE, *Ein Kriterium für die Quasi-Optimalität des Ritzschen Verfahrens*, Numer. Math., 11 (1968), pp. 346–348. MR0233502 (38:1823)
- [18] P.-A. RAVIART AND J. M. THOMAS, *A mixed finite element method for second order elliptic problems*, Mathematical Aspects of Finite Element Method (I. Galligani and E. Magenes, eds.), Lecture Notes in Math. 606, Springer-Verlag, New York, 1977, pp. 292–315. MR0483555 (58:3547)
- [19] R. STENBERG, *A family of mixed finite elements for the elasticity problem*, Numer. Math **53** (1988), 513–538. MR954768 (89h:65192)
- [20] R. STENBERG, *Postprocessing schemes for some mixed finite elements*, RAIRO Modél. Math. Anal. Numér. **25** (1991), 151–167. MR1086845 (92a:65303)

SCHOOL OF MATHEMATICS, UNIVERSITY OF MINNESOTA, VINCENT HALL, MINNEAPOLIS, MINNESOTA 55455

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

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF FLORIDA, GAINESVILLE, FLORIDA 32611-8105

E-mail address: `jayg@math.ufl.edu`

DEPARTAMENTO DE MATEMÁTICA APLICADA, CPS, UNIVERSIDAD DE ZARAGOZA, 50018 ZARAGOZA, SPAIN

E-mail address: `sayas002@umn.edu`