HYBRIDIZABLE DISCONTINUOUS GALERKIN AND MIXED
FINITE ELEMENT METHODS FOR ELLIPTIC PROBLEMS ON
SURFACES

BERNARDO COCKBURN\textsuperscript{1} AND ALAN DEMLOW\textsuperscript{2}

Abstract. We define and analyze hybridizable discontinuous Galerkin methods for the Laplace-Beltrami problem on implicitly defined surfaces. We show that the methods can retain the same convergence and superconvergence properties they enjoy in the case of flat surfaces. Special attention is paid to the relative effect of approximation of the surface and that introduced by discretizing the equations. In particular, we show that when the geometry is approximated by polynomials of the same degree of those used to approximate the solution, although the optimality of the approximations is preserved, the superconvergence is lost. To recover it, the surface has to be approximated by polynomials of one additional degree. We also consider mixed surface finite element methods as a natural part of our presentation. Numerical experiments verifying and complementing our theoretical results are shown.

1. Introduction

We extend the definition of the so-called hybridizable discontinuous Galerkin (HDG) methods to second-order elliptic problems on implicitly defined surfaces and carry out their a priori error analysis. We prove that the resulting methods can be implemented as efficiently as they are for the case of flat surfaces, and that the fact that the surface is not flat does not degrade their remarkable optimal convergence and superconvergence properties so long as surface approximations of a sufficiently high degree are employed.

The numerical solution of PDEs defined on surfaces arises naturally in many applications. We consider the following model problem in order to focus on basic issues arising in the definition and analysis of such numerical methods. Let $\Gamma$ be a smooth 2-dimensional surface embedded in $\mathbb{R}^3$ without boundary. Let $f$ be given data satisfying $\int_{\Gamma} f \, d\sigma = 0$ where $d\sigma$ is surface measure, and let $u$ solve

\begin{equation}
- \Delta_{\Gamma} u = f \text{ on } \Gamma.
\end{equation}

Here $\Delta_{\Gamma}$ is the Laplace-Beltrami operator on $\Gamma$. We require $\int_{\Gamma} u \, d\sigma = 0$ in order to guarantee uniqueness. Surfaces with boundary can be handled in much the same way, as we discuss in our numerical examples below.

\begin{itemize}
\item[1991 Mathematics Subject Classification.] 58J32, 65N15, 65N30.
\item[Key words and phrases.] Laplace-Beltrami operator, surface finite element methods, a priori error estimates, boundary value problems on surfaces, discontinuous Galerkin methods, hybridizable finite element methods, mixed finite element methods.
\item[\textsuperscript{1}] Partially supported by NSF grant DMS-1115331.
\item[\textsuperscript{2}] Partially supported by NSF grant DMS-1318652.
\end{itemize}
Our results are part of an ongoing effort to devise high-order accurate finite element method for surface problems. Let us briefly review it to place them into historical perspective. A canonical continuous Galerkin surface finite element method for solving (1.1) was defined by Dziuk in [26]. There a polyhedral approximation $\Gamma_h$ to $\Gamma$ whose triangular faces form the finite element triangulation was employed, along with continuous piecewise linear finite element spaces. Surface FEM are now an active field of study. Recent works include both basic numerical analysis such as development and analysis of higher-order methods, adaptive algorithms, and multigrid methods [23, 24, 36, 9] and development and analysis of algorithms for various applications.

A few very recent works have been devoted to the numerical analysis of surface discontinuous Galerkin methods [22, 34, 35, 1]. This is a natural development as the practical use of DG methods is now widespread, including in applications such as two-phase flow problems in which surface PDE appear. Discontinuous Galerkin (including HDG) methods have well-known advantages when solving certain types of elliptic problems, such as advection-dominated problems. Moreover, HDG methods have recently been shown to be computationally more efficient than the classical FEM for high polynomial degree, even in the purely diffusive case; see [33]. For our purposes the most closely related work on surface DG methods appears in [22, 1], which provide convergence analysis of interior penalty and other classical DG methods for (1.1). [1] contains the most extensive analysis, essentially combining the unified DG analysis framework of [4] and the surface FEM framework of [26, 23] to obtain high-order DG methods on high-order surface approximations. Optimal-order $L_2$ and DG-energy error estimates are obtained. In comparison with [1], we employ the framework of hybridizable methods instead of that of [4], naturally consider mixed methods as part of our analysis [15], and also allow more general choices of element geometries including quadrilaterals. Surface FEM employing quadrilateral elements have increasing importance in practice as they have for example been recently implemented in the widely used library deal.ii [6, 7], but such elements have generally not been included in surface FEM error analyses.

The HDG methods were originally defined in [15] for the Laplace equation (on flat surfaces). They employ three approximation spaces and corresponding discrete approximations. These are a scalar approximation $u_h \in W_h$ to $u$, a vector approximation $q_h \in V_h$ to $q = -\nabla u$, and an interelement multiplier $\tilde{u}_h \in M_h$ which approximates traces of $u$ on the mesh skeleton. In the context of triangular meshes on Euclidean domains, we for the time being think of $W_h$ as discontinuous polynomials of a given degree $r$, $V_h$ as discontinuous $[P_r]^2$ functions, and $M_h$ as discontinuous $P_r$ functions living on the mesh skeleton. The HDG methods share many of the positive characteristics exhibited by traditional DG methods, such as good stability properties for singularly perturbed problems. An additional advantage of HDG methods is that the finite element system is globally coupled only through the multiplier space $M_h$, so that at least for sufficiently high polynomial degree the global linear system of equations that must be solved is relatively small in comparison with traditional DG methods; see also [33] for a comparison with the traditional FEM. In addition, HDG methods achieve simultaneously optimal ($O(h^{r+1})$) convergence in both the $L_2$ and energy norms (even for $r = 0$) and admit locally-defined superconvergent postprocessed approximations to $u$ of order $r + 2$, whenever $r \geq 1$.
Standard surface FEM have two error sources [26, 23, 32]. These are the Galerkin error resulting from replacing infinite-dimensional spaces by finite-dimensional spaces in the variational problem, and a geometric error resulting from the approximation of $\Gamma$ by $\Gamma_h$. Discontinuous Galerkin (or more broadly nonconforming) methods typically have additional error sources arising from the interelement jumps of both the scalar and vector variable. The size of these jumps has to be properly balanced by a suitable choice of the stabilization function in order to achieve the above-mentioned convergence and superconvergence properties; see [16]. Thus it is especially important to understand the relationship between the Galerkin and geometric errors, as developing higher-order surface FEM naturally requires both higher-degree finite element spaces and higher-degree surface approximations. Employing a projection-based error analysis of the type now typical for HDG methods, we show that properly-defined surface HDG methods achieve optimal convergence in the $L^2$-norm of the scalar and the vector approximations for isoparametric elements, that is, when the degree of the surface approximation is the same as the degree $r$ of the finite element spaces $V_h$, $W_h$, and $M_h$. In contrast, in both standard continuous Galerkin and traditional DG methods optimal convergence is achieved in the energy norm when the polynomial degree of the surface approximation is one less than that of the finite element space, while optimal approximation in $L^2$ requires that the degree of the surface approximation be at least that of the finite element space (cf. [23, 22]). Note, however, that for traditional DG methods, the convergence of the approximation of the vector variable is suboptimal by one power.

A well-known feature of both mixed and HDG methods is that they admit superconvergence estimates (for unstructured meshes) which for HDG methods have the form $\|\Pi u - u_h\| \lesssim h^{r+2}$ for the scalar variable; here $\Pi$ is an appropriate local projection and $r$ is the degree of the finite element space. If such estimates hold, then a variety of simple elementwise postprocessing procedures yield $O(h^{r+2})$-accurate approximations to $u$. An important conclusion of our analysis and numerical tests is that the order of the geometric error is the same for $\|\Pi u - u_h\|$ as for all other error quantities, i.e., $h^{k+1}$ if a degree-$k$ surface approximation is employed. Thus the ability to obtain superconvergent postprocessed solutions is limited by the surface approximation. In particular, if a natural isoparametric HDG method ($k = r$) is used, then postprocessing does not yield an improved approximation.

We now discuss some notable features of our analysis. First, our presentation shares some common features with the analysis of HDG methods on isoparametric Euclidean elements given in [19]. The use of Piola transforms in particular is a common theme in both works. Use of this canonical transformation to map the vector finite element space $V_h$ is natural as $V_h$ essentially serves as a DG approximation to $H(\text{div})$, but this fact was not observed in the original papers defining and analyzing HDG methods. Both [19] and the current work involve curved elements and thus non-affine reference maps, which makes the use of proper canonical transformations more important in the analysis. The similarities between the case of surface FEM and parametric FEM on Euclidean domains may not however be obvious at first glance, since the reference maps arise in somewhat different forms in the two situations. In addition to taking advantage of the analytical advantages of the Piola transform already observed in [19], we also demonstrate in Section 5 below that it significantly simplifies the implementation of HDG methods.
The geometric errors caused by the approximation of $\Gamma$ by $\Gamma_h$ require other nontrivial modifications of the arguments used in the analysis of Euclidean HDG methods. In particular, they result in a coupling of the vector and scalar variables so that a duality argument is needed even to prove energy-norm estimates. Geometric errors may more broadly viewed as a type of variational crime [32]. While we do not fully develop a variational crimes framework here, similar coupling of the variables could be expected when analyzing (for example) quadrature errors in the calculation of the load vector in Euclidean HDG methods. Our approach to geometric errors also differs in an important detail from the analysis previously given for other types of surface FEM including standard and modified continuous Galerkin (CG) methods [26, 23, 38], DG methods [22], and mixed surface FEM [32]. The general approach of all of these works is to define a finite element approximation on the discrete surface $\Gamma_h$, pull it back to the continuous surface $\Gamma$ via a canonical transformation, and there compare it with $u$. We instead transform $u$ to $\Gamma_h$ and carry out the error analysis there. The end result is essentially equivalent, but our analysis circumvents the need for a discrete stability result (cf. the proofs of [23, Corollary 4.2] and [32, Theorem 3.10]) and is thus more direct.

The implementation and some analytical aspects of HDG and of traditional mixed finite element methods have much in common [15], so we also consider mixed surface FEM below. Numerical analysis of mixed surface FEM has previously been considered from two different perspectives. Surface Raviart-Thomas spaces appeared quite early in the literature in the context of boundary element methods; see especially Nédélec [37] and Bendali [8] along with later works such as [31, 11]. While solution of the Laplace-Beltrami problem is not the primary goal in these papers, there are parallels between certain aspects of their presentation and our framework. Bendali [8] in particular employs polynomial surface approximations and analyzes the resulting geometric errors in boundary element methods employing surface Raviart-Thomas spaces. Mixed finite element methods on surfaces have also been recently considered by Holst and Stern in [32] from the perspective of finite element exterior calculus (FEEC). There a “variational crimes” framework is used to extend the FEEC mixed methods of [5] to generate approximations to the Hodge Laplace problem on surfaces. Their error analysis includes estimates in the natural $H(\text{div}; \Gamma) \times L^2$ variational norm for Raviart-Thomas and BDM triangular mixed surface approximations to (1.1) defined on piecewise polynomial approximations to $\Gamma$. We include standard individual error estimates for the flux and scalar variables in the mixed method (cf. [28]) as a natural part of our analysis, although the scalar finite element space that we employ differs from that in [8, 32] except when the element reference map is affine.

The structure of surface HDG methods also renders their design and analysis simpler in certain ways than other types of surface DG methods. As is pointed out in [22], an important issue arising in surface DG methods is the appropriate choice of conormal vectors on element edges. In contrast to the situation on Euclidean domains, the outward-pointing conormals on edges shared by adjacent elements on $\Gamma_h$ do not generally point in opposite directions. [22] investigates several possible choices of conormals for use in interior penalty-DG methods and concludes that two of the choices (either employing the actual element conormals or averaging them) are roughly equivalent in practice. In contrast, the Piola transform that we employ to define surface HDG methods dictates that we use actual conormals...
as they are appropriately preserved under Piola transformations. Secondly, fewer geometric error terms arise in HDG methods than in the IPDG methods considered in [22]. Roughly speaking, geometric errors in surface CG methods result from approximation of the Dirichlet form \( \int_{\Gamma} \nabla_u \cdot \nabla_v \, d\sigma \), approximation of \( f \) by some function \( f_h \) defined on \( \Gamma_h \), and enforcement of the condition \( \int_{\gamma} u \, d\sigma = 0 \) on \( \Gamma_h \) instead of \( \Gamma \). In the IPDG analysis in [22] there also arise what might be called “geometric-DG” error terms involving typical DG-related quantities such as interelement jumps, traces, penalty parameters, etc. By properly defining penalty parameters and employing Piola transforms, we avoid any such geometric-DG error terms in our analysis of HDG methods. Instead, there arise essentially only the same geometric error terms as in CG methods. As we discuss in §5 below, while the linear algebraic structure of HDG methods is more complicated than CG methods, implementation of HDG methods also requires metric information about \( \Gamma \) only in the same three terms as in CG methods.

The outline of the paper is as follows. In Section 2, we give all the information needed for setting up the problem. In Section 3, we describe the family of HDG and mixed methods under consideration. In Section 4, we obtain their convergence (and superconvergence) properties. In Section 5, we consider the case in which there is no error in the approximation of the surface and demonstrate that use of Piola transforms simplifies implementation. In Section 6, we present numerical experiments validating our theoretical estimates. We end in Section 7 with concluding remarks.

2. Preliminaries

In this section we record a number of preliminaries concerning geometry, transformations of functions between the continuous and discrete surfaces \( \Gamma \) and \( \Gamma_h \), and analytical results.

2.1. Geometric and analytic preliminaries on \( \Gamma \). We assume throughout that \( \Gamma \) is a compact, oriented, \( C^\infty \), two-dimensional surface without boundary which is embedded in \( \mathbb{R}^3 \). Extension to higher-dimensional surfaces of codimension 1 is for the most part immediate.

Let \( d \) be the oriented distance function for \( \Gamma \) satisfying \( d < 0 \) on the interior of \( \Gamma \) and \( d > 0 \) on the exterior. The vector \( \vec{\nu} = \nabla d \) is the outward-pointing unit normal, and \( \mathbf{H} = \nabla \vec{\nu} = \nabla^2 d \) is the Weingarten map. These quantities are all defined on a tubular neighborhood \( U \) of \( \Gamma \) (defined below). Here we express these quantities in the coordinates of the embedding space \( \mathbb{R}^3 \). For \( x \in \Gamma \), the eigenvalues \( \kappa_1, \kappa_2 \) of \( \mathbf{H} \) corresponding to eigenvectors perpendicular to \( \vec{\nu} \) are the principal curvatures at \( x \). Let \( U \subset \mathbb{R}^3 \) be a strip of width \( \delta \) about \( \Gamma \). We assume that \( \delta < \min_{i=1,2} 1/\|\kappa_i\|_{L_{\infty}(\Gamma)} \), in order to ensure that the orthogonal projection

\[
a(x) = x - d(x)\vec{\nu}(x)
\]

onto \( \Gamma \) is unique [29, 24]. Let \( \mathbf{P} = \mathbf{I} - \vec{\nu} \otimes \vec{\nu} \) be the projection onto the tangent plane to \( \Gamma \) at \( x \), where \( \otimes \) is the outer product defined by \( (\vec{a} \otimes \vec{b})\vec{c} = \vec{a}(\vec{b} \cdot \vec{c}) \). Then \( \nabla_{\Gamma} = \mathbf{P} \nabla \) is the tangential gradient, \( \text{div}_{\Gamma} = \nabla_{\Gamma} \cdot \) is the tangential divergence, and \( \Delta_{\Gamma} = \text{div}_{\Gamma} \nabla_{\Gamma} \) is the Laplace-Beltrami operator. We use standard notation \( (H^1(\Gamma), W^1_0(\Gamma), \text{etc.}) \) for Sobolev spaces and norms of functions with \( j \) tangential derivatives in \( L_p(\Gamma) \).
Finally, given the conditions \( f \in L^2(\Gamma) \), \( \int_T f \, d\sigma = 0 \), and \( \Gamma \) is a closed and smooth surface, the problem (1.1) has a unique solution \( u \in H^2(\Gamma) \) satisfying \( \int_T u \, d\sigma = 0 \), and there holds the regularity estimate
\[
\|u\|_{H^2(\Gamma)} \leq C_T \|f\|_{L^2(\Gamma)}.
\]

2.2. Discrete surface approximations. Typically surface finite element methods have been defined over triangular meshes or their higher-order generalizations [26, 23], but we wish to allow for more general element geometry including especially quadrilateral elements. We denote by \( \Gamma_h \subset U \) our “base” surface approximation which in the case of triangular elements is a polyhedron having triangular faces. More generally, \( \Gamma_h \) is a generalized “polyhedron” consisting of a set of linear edges interpolated by faces which may not be planar. We assume that each of the faces \( T \) of \( \Gamma_h \) is of the form \( A_T K \) for some reference element \( K \) and continuous map \( A_T : K \to T \). \( A_T \) may for example be an affine map from a unit triangle or a bilinear map from a unit square; in the latter case \( A_T(K) \) is not generally planar. We assume that \( a|_{\Gamma_h} : \Gamma_h \to \Gamma \) is a bijection, and that \( \|d|\|_{L^\infty(\Gamma_h)} \leq C h^2 \). The nodes of \( \Gamma_h \) are often assumed to lie on \( \Gamma \), but this is not necessary. Although we could easily consider many reference elements, for example, \( \Gamma_h \) could have a mix of quadrilateral and triangular faces, for simplicity, we restrict ourselves to a single reference element. Finally, let \( T_h \) be the set of faces of \( \Gamma_h \) and \( E_h \) the set of edges of the elements in \( T_h \). Let \( h_T = \text{diam}(T) \) denote the local mesh size, \( DA_T \) and \( D(A_T^{-1}) \) the tangent maps of \( A_T \) and \( A_T^{-1} \) respectively, \( JA_T \) the Jacobian of \( A_T \), and \( D^2A_T \) the Hessian matrix of \( A_T \). We assume that
\[
\begin{align*}
(2.2a) & \quad \|D A_T\|_{L^\infty(\Gamma)} \lesssim h_T, \\
(2.2b) & \quad \|D(A_T^{-1})\|_{L^\infty(\Gamma)} \lesssim h_T^{-1}, \\
(2.2c) & \quad JA_T \simeq h_T^2, \\
(2.2d) & \quad \|D^2 A_T\|_{L^\infty(\Gamma)} \lesssim h_T^2.
\end{align*}
\]
Here and in what follows we write \( a \lesssim b \) whenever \( a \leq Cb \) for a constant \( C \) not depending on essential quantities, and \( a \simeq b \) when \( a \lesssim b \lesssim a \). For a shape regular triangular mesh the first three relationships in (2.2) are satisfied, and \( D^2 A_T = 0 \). For quadrilateral meshes (2.2d) is satisfied for affine but not for arbitrary bilinear reference mappings, even those preserving shape regularity. However, (2.2) does hold for “asymptotically parallelogram” sequences of meshes generated by asymptotically affine mappings from the unit square. Such meshes may be obtained from regular refinement of any initial shape-regular quadrilateral surface \( \Gamma_h \) even when projection of newly created nodes onto \( \Gamma \) is accounted for; cf. [2] for discussion of the Euclidean case. General bilinear reference mappings may lead to degeneration of approximation properties [2, 3], but this problem may be circumvented by use of appropriately shaped reference elements.

We denote by \( \vec{n}_h \) the outward unit normal on \( \Gamma_h \). We assume that \( \min_{x \in \Gamma_h} \vec{n}_h(x) \cdot \vec{v}(x) \geq \gamma > 0 \) for some constant \( \gamma > 0 \). This condition is guaranteed for \( h \) sufficiently small by the assumption that the vertices of \( \Gamma_h \) are \( O(h^2) \) from \( \Gamma \), since this implies that \( |\vec{v} - \vec{v}_h| = O(h) \). Finally, given an edge \( e \in E_h \) of an element \( T \in T_h \), we let \( \vec{n}_h \) be the outward-pointing conormal, that is, the unit vector which at \( x \in e \) lies in the tangent plane of \( T \) at \( x \) and which is outward-pointing and normal to \( e \). \( \vec{n}_h \) is edgewise constant on triangular approximations \( \Gamma_h \) but not in general.
We next describe a family of discrete surfaces $\Gamma^k_h$ ($k \geq 1$) which in the simplest case are piecewise polynomial approximations to $\Gamma$. Let $a_k : \Gamma_h \to \mathbb{R}^3$ be a continuous map whose range is a closed surface which we denote by

$$\Gamma^k_h = \{a_k(x) : x \in \Gamma_h\}.$$ 

Let $\bar{\mu}^k_h$ be the (piecewise smooth) uniform normal on $\Gamma^k_h$. We additionally assume that

$$\|a - a_k\|_{L_\infty(\Gamma_h)} + \|d\|_{L_\infty(\Gamma_h^k)} + h\|\tilde{\nu} - \bar{\mu}^k_h\|_{L_\infty(\Gamma^k_h)} + h^2|a - a_k|_{W^2_2(\Gamma_h)} \lesssim h^{k+1}. \tag{2.3}$$

In the case of triangular meshes a canonical definition (cf. [8, 23]) is to take each component of $a_k$ to be the Lagrange interpolant of the corresponding component of the projection $a$ restricted to $\Gamma_h$. For quadrilateral meshes, $a_k$ may be taken as a $Q_k$ interpolant, i.e., for $T \in T_h a_k|_T = q_{k,T} \circ A^{-1}_T$ for some $Q_k$ function $q_{k,T}$ defined on $K$, we then set $T^k_h$ to be the image under $a_k$ of $T_h$, i.e., for $K \in T^k_h$, $K = a_k(T)$ for some $T \in T_h$; cf. Figure 1.

![Figure 1](image-url)

**Figure 1.** The relationship between $\Gamma$, $\Gamma_h$, and $\Gamma^k_h$.

We also define the skeleton of the triangulation $T^k_h$ as

$$E^k_h = \{a_k(e) : e \in E_h\}. \tag{2.4}$$

Given an edge $e \in E^k_h$ of an element $K \in T^k_h$ and $x \in e$, we let $\bar{n}_h(x)$ be the outward-pointing conormal. We similarly define $\bar{n}$ on $\Gamma$.

### 2.3. The correspondence between $\Gamma_h$, $\Gamma^k_h$, and $\Gamma$: Geometry and Sobolev spaces.

Our analysis requires a number of relationships between functions defined on $\Gamma$ and $\Gamma^k_h$, as in [26] and [24]. We first establish relationships between functions defined on the continuous surface $\Gamma$ and the discrete surfaces $\Gamma^k_h$. For $v \in H^1(\Gamma)$, we define its extension to $U$ by $v'(x) = v(a(x))$ for all $x \in U$. For $v_h \in H^1(\Gamma^k_h)$, we define the lift $\bar{v}_h \in H^1(\Gamma)$ by $\bar{v}_h(a(\hat{x})) = v_h(\hat{x})$ for all $\hat{x} \in \Gamma^k_h$. Finally, for any $v_h \in H^1(\Gamma^k_h)$, we define its extension to $U$ by $v'_h(x) = \bar{v}_h(a(x))$ for all $x \in U$.

We also introduce notation concerning surface and edge measures. For $x \in \Gamma^k_h$, let $\mu_{hk}(x)$ satisfy $\mu_{hk}(x) \, ds_{hk}(x) = ds(a(x))$, where $ds$ and $ds_{hk}$ are surface measure on $\Gamma$ and $\Gamma^k_h$, respectively. Also, for $x_e \in e \in E^k_h$, let $\mu_{hk}^e(x_e)$ satisfy $\mu_{hk}^e(x_e) \, ds_{hk}(x_e) = ds(a(x_e))$, where $ds$ and $ds_{hk}$ are the measures on $a(e)$ and $e$, respectively.

**Proposition 2.1.** Let $x \in K \in T^k_h$, let $x_e \in e \in E^k_h$, and let $t_{hk}$ be the unit tangent vector to $e$. Then

\begin{align}
\mu_{hk}(x) &= \tilde{\nu}(x) \cdot \tilde{\nu}^k_h(x) \Pi_{i=1}^2 \Pi_{j=1}^3 (1 - d(x) z_i(x)), \tag{2.5a} \\
\mu_{hk}^e(x_e) &= |[\vec{P} - d\vec{H}]| \, t_{hk}(x_e)|, \tag{2.5b} \\
|1 - \mu_{hk}(x)| + |1 - \mu_{hk}^e(x_e)| &\lesssim h^{k+1}. \tag{2.5c}
\end{align}
Proof. (2.5a) and the first half of (2.5c) are proved in [23]. The inequality (2.5b) is an elementary consequence of the change of variables formula for line integrals, and the second half of (2.5c) follows from (2.5b) after a short calculation.

We also shall need to compare Sobolev norms of functions defined on $\Gamma$ and $\Gamma^k_h$. Let $v \in W^j_p(\Gamma)$ with $j \geq 0$ and $1 \leq p \leq \infty$. Then (cf. [26], [23])

\[
\|v^i\|_{W^j_p(\Gamma^k_h)} \lesssim |v|_{W^j_p(\Gamma)}, \quad i = 0, 1,
\]

\[
\|D_{1,h}^j v^i\|_{L^p(\Gamma^k_h)} \lesssim \sum_{\min(j,1) \leq m \leq j} \|D_{1,h}^m v\|_{L^p(\Gamma)}.
\]

Since $D_{1,h}^j$ acts elementwise, the same inequalities hold elementwise. Similar relationships also hold for functions defined on $\Gamma_h$ and $\Gamma^k_h$.

Combining (2.2), (2.3), and (2.5c), we see that the properties (2.2) hold with the base reference mapping $A_T$ replaced by either the high-degree reference mapping $a_k \circ A_T$ or by $a \circ A_T$, since smoothness of $\Gamma$ guarantees that the derivatives of $a$ are uniformly bounded independent of $h$.

2.4. Piola transforms on surfaces. Following [19], we employ Piola transforms as a canonical mapping for the vector variable $q$. Piola transforms are standard tools for mixed finite element methods on Euclidean domains, and their surface incarnation has similar features [8, 25, 31, 40]. We however list their properties in some detail as their usage on surfaces may be slightly less familiar.

Following [40], let $S_0, S_1$ be surfaces, let $\Phi : S_0 \to S_1$, and let $\varphi$ be its inverse mapping. Let also $F$ and $f$ be the corresponding tangent maps, i.e., $F : T_0 \to T_1$ and $f : T_1 \to T_0$, where $T_i$ is the tangent space of $S_i$. Finally, let $\mu$ satisfy $\mu \, d\sigma_0 = d\sigma_1$, where $d\sigma_i$ is surface measure on $S_i$. The surface Piola transform is given by

\[
q_1 = P_{\varphi}q_0 = \frac{1}{\mu}Fq_0, \quad q_0 \in H(\text{div}; S_0), \quad q_0 = P_{\varphi}q_1 = \mu f q_1, \quad q_1 \in H(\text{div}; S_1).
\]

Here we use the convention that $q, \tilde{q}$ are column vectors. The standard identity

\[
\text{div}_{S_0} q_0 = \mu \text{div}_{S_1} q_1
\]

holds provided $q_0$ and $q_1$ are tangent vectors; a slightly more complicated relationship holds if not. The above identity, the surface divergence theorem, and change of variables formulas yield for $v_0 \in H^1(S_0)$, $v_1 \in H^1(S_1)$ with $v_0(x) = v_1(\Phi(x))$ that

\[
\int_{S_0} v_0 \text{div}_{S_0} q_0 \, d\sigma_0 = \int_{S_1} v_1 \text{div}_{S_1} q_1 \, d\sigma_1,
\]

\[
\int_{S_0} q_0 \cdot \nabla_{S_0} v_0 \, d\sigma_0 = \int_{S_1} q_1 \cdot \nabla_{S_1} v_1 \, d\sigma_1,
\]

\[
\int_{\partial S_0} q_0 \cdot n_0 v_0 \, ds_0 = \int_{\partial S_1} q_1 \cdot n_1 v_1 \, ds_1.
\]

In the last equation we have assumed that $S_i$ ($i = 0, 1$) are manifolds with boundaries having outward-pointing unit conormals $\vec{n}_i$, $i = 0, 1$ on $\partial S_i$.

Let now $K \in T^k_h$ with image $a(K) \subset \Gamma$. Letting $a$ correspond to $\Phi$ above, we have $F = \nabla a = P - dH$ and $\mu = \mu_{hh}$. $\varphi$ corresponds to the inverse of $a$ viewed
as a mapping $\Gamma_h^k \to \Gamma$, and it can easily be verified that $f = [\mathbf{I} - \frac{\bar{\nu} \otimes \bar{\nu}_h^k}{\bar{\nu} \cdot \bar{\nu}_h^k}] [\mathbf{I} - d\mathbf{H}]^{-1}$.

Given $q \in H(\text{div}; \Gamma)$, we define $\tilde{q} \in H(\text{div}; \Gamma_h^k)$ by

\begin{equation}
(2.10) \quad \tilde{q}(x) = \mathcal{P}_{a^{-1}} q = \mu_{hh} \left[ \mathbf{I} - \frac{\bar{\nu} \otimes \bar{\nu}_h^k}{\bar{\nu} \cdot \bar{\nu}_h^k} \right] [\mathbf{I} - d\mathbf{H}]^{-1} q(a(x)), \quad x \in \Gamma_h^k.
\end{equation}

Similarly,

\begin{equation}
(2.11) \quad q(a(x)) = \mathcal{P}_a \tilde{q} = \frac{1}{\mu_{hh}} [\mathbf{P} - d\mathbf{H}] \tilde{q}(x), \quad a(x) \in \Gamma.
\end{equation}

We emphasize that the surface Piola transform preserves surface $H(\text{div})$ spaces, and that the relationships (2.9a), (2.9b), and (2.9c) hold with $S_0 = K$, $S_1 = a(K)$.

In our analysis we compare tangential gradients of extensions ($\nabla_{\Gamma_h^k} v^\ell$) with Piola transforms of gradients ($\mathcal{P}_{a^{-1}} \nabla_{\Gamma} v$). We summarize this comparison as follows.

**Proposition 2.2.** Let $v \in H^1(\Gamma)$, and let $q = -\nabla_{\Gamma} v$. Then

\begin{equation}
(2.12) \quad |\nabla_{\Gamma_h^k} v^\ell + \mathcal{P}_{a^{-1}} q| \lesssim h^{k+1} |q^\ell|.
\end{equation}

**Proof.** The identity (2.2.15) of [24] yields $\nabla_{\Gamma_h^k} v^\ell = \mathcal{P}_h [\mathbf{I} - d\mathbf{H}] \mathcal{P} \nabla_{\Gamma} v$. Combining this with (2.10) while noting that we may premultiply the matrix on the right hand side of (2.10) by $\mathbf{P}_h$ and postmultiply by $\mathbf{P}$ without changing the expression yields

\[ \nabla_{\Gamma_h^k} v^\ell + \mathcal{P}_{a^{-1}} q = \mathcal{P}_h \left[ (\mathbf{I} - d\mathbf{H}) - \mu_{hh} \left[ \mathbf{I} - \frac{\bar{\nu} \otimes \bar{\nu}_h^k}{\bar{\nu} \cdot \bar{\nu}_h^k} \right] \right] [\mathbf{I} - d\mathbf{H}]^{-1} \mathcal{P}_h v^\ell. \]

The inequality (2.3) implies that both $\mathbf{I} - d\mathbf{H}$ and $[\mathbf{I} - d\mathbf{H}]^{-1}$ are $O(h^{k+1})$ perturbations of $\mathbf{I}$. Using this observation along with (2.5c), we have

\[ |\nabla_{\Gamma_h^k} v^\ell + \mathcal{P}_{a^{-1}} q| \lesssim |\mathcal{P}_h \frac{\bar{\nu} \otimes \bar{\nu}_h^k}{\bar{\nu} \cdot \bar{\nu}_h^k} \mathcal{P}_h v^\ell| + h^{k+1} |q^\ell|. \]

An elementary computation and the inequality (2.3) leads to

\[ |\mathcal{P}_h \frac{\bar{\nu} \otimes \bar{\nu}_h^k}{\bar{\nu} \cdot \bar{\nu}_h^k} \mathcal{P}_h v^\ell| = \left| \left( \frac{\bar{\nu} \cdot \bar{\nu}_h^k}{\bar{\nu} \cdot \bar{\nu}_h^k} v^\ell - \bar{\nu} \right) \otimes \left( \frac{\bar{\nu} \cdot \bar{\nu}_h^k}{\bar{\nu} \cdot \bar{\nu}_h^k} \bar{\nu} \right) \right| \lesssim h^{2k}, \]

which completes the proof. \hfill \Box

### 3. A Hybridizable Discontinuous Galerkin Method

In this section we define our hybridizable discontinuous Galerkin method.

#### 3.1. Finite element spaces

We start by introducing spaces on the reference element. Let $\hat{e}$ be the unit interval, and let $M(\hat{e})$ be a finite-dimensional space on $\hat{e}$. Let $K$ be the reference element and $e$ any of its edges with associated affine bijection $A_e : \hat{e} \to e$, and consider finite-dimensional spaces

\begin{align}
(3.1a) & \quad V \subset H(\text{div}; K), \\
(3.1b) & \quad \mathcal{W} \subset H^1(K), \\
(3.1c) & \quad M(e) = M(\hat{e}) \circ A_e^{-1} \subset L_2(e).
\end{align}

Next we define surface finite element spaces as follows. On $\Gamma_h$ we define

\begin{align}
(3.2a) & \quad \hat{V}_h = \{ \hat{v} \in [L_2(\Gamma_h)]^3 : \hat{v} \in \mathcal{P}_{A_e} V, \quad T \in \mathcal{T}_h \}, \\
(3.2b) & \quad \hat{W}_h = \{ \hat{w} \in L_2(\Gamma_h) : \hat{w}|_T = \hat{w} \circ A^{-1} \text{ for some } \hat{w} \in \mathcal{W}, \quad T \in \mathcal{T}_h \}, \\
(3.2c) & \quad \hat{M}_h = \{ \hat{\eta} \in L_2(\hat{E}_h) : \hat{\eta}|_e = \eta \circ A^{-1} \text{ for some } \eta \in M(e), \quad e \in \mathcal{E}_h \}.
\end{align}
We can now define the spaces on $\Gamma^k_h$ as

\begin{align}
(3.3a) & \quad V_h = \mathcal{P}_{a_k} \tilde{V}_h, \\
(3.3b) & \quad W_h = \{ w \in L_2(\Gamma^k_h) : w = \tilde{w} \circ (a_k)^{-1} \text{ for some } \tilde{w} \in \tilde{W}_h \}, \\
(3.3c) & \quad W^0_h = W_h \cap L_2^0(\Gamma^k_h), \\
(3.3d) & \quad M_h = \{ \eta \in L_2(\mathcal{E}^k_h) : \eta = \tilde{\eta} \circ (a_k)^{-1} \text{ for some } \tilde{\eta} \in \tilde{M}_h \}.
\end{align}

Here $L_2^0(\Gamma^k_h) = \{ w \in L_2(\Gamma^k_h) : (w, 1)_{\Gamma^k_h} = 0 \}$. We also use the auxiliary space $\mathcal{E}^k_h$. This choice ensures that a standard commuting diagram property holds [10]. The hybridized mixed methods

The identity (2.8) implies that, unless the reference map $a_k \circ A$ is affine, $\nabla \Gamma^k_h \cdot V_h \not\subset W_h$ even if $\nabla \cdot V_h \subset W_h$. For mixed FEM on surfaces, the canonical choice of the scalar space is $W_{h,mixed} = \{ w \in L_2(\Gamma^k_h) : w = J(a_k \circ A)^{-1} \tilde{w} \circ (a_k)^{-1}, \text{ some } \tilde{w} \in W_h \}$, where $J(a_k \circ A)$ is the Jacobian of $a_k \circ A$ [8, 32]. This choice ensures that a standard commuting diagram property holds [10]. The hybridized mixed methods that we consider thus are not conforming unless the reference map is affine. We however still obtain optimal error estimates, and our choice of $W_h$ significantly simplifies implementation and analysis as compared with the choice $W_{h,mixed}$.

The space $V_h$ will roughly speaking serve as an approximating space for $q = -\nabla \Gamma^k_h \cdot u$, which naturally lies in $H(\text{div}; \Gamma)$. Given $q \in H(\text{div}; \Gamma)$, we thus define $\hat{q} \in H(\text{div}; \Gamma^k_h)$ to be the Piola transform (2.10) of $q$ to $\Gamma^k_h$. We may now more precisely state that $V_h$ will serve as an approximating space for $\hat{q}$. Similarly, the spaces $W_h$ and $M_h$ serve as approximating spaces to $u$, or more precisely, to the extension $u^\circ$ restricted to $\Gamma^k_h$ and $\mathcal{E}^k_h$, respectively.

3.2. The HDG method. The hybridizable discontinuous Galerkin method consists of finding $(q_h, u_h, \tilde{n}_h) \in V_h \times W_h \times M_h$ such that

\begin{align}
(3.5a) & \quad (q_h, v)_{\Gamma^k_h} - (u_h, \nabla \Gamma^k_h \cdot v)_{\Gamma^k_h} + \langle \tilde{n}_h, \bar{n}_h \cdot v \rangle_{\partial \Gamma^k_h} = 0, \\
(3.5b) & \quad -(q_h, \nabla \Gamma^k_h w)_{\Gamma^k_h} + \langle \tilde{n}_h, \bar{n}_h \cdot w \rangle_{\partial \Gamma^k_h} = (f_h, w)_{\Gamma^k_h}, \\
(3.5c) & \quad \text{on } \partial \Gamma^k_h, \quad \bar{q}_h \cdot \bar{n}_h = q_h \cdot \bar{n}_h + \tau (u_h - \tilde{u}_h), \\
(3.5d) & \quad -\langle \tilde{q}_h, \bar{n}_h \rangle_{\partial \Gamma^k_h} = 0
\end{align}

for all $(v, w, \eta) \in V_h \times W_h \times M_h$. Here $\tau \geq 0$ is a scalar function on $\partial \Gamma^k_h$ which may be double-valued on $\mathcal{E}^k_h$ and $f_h \in L_2^0(\Gamma^k_h)$ is an approximation to $f$. A convenient choice for theoretical purposes is $f_h = \mu_h f$, while for practical purposes $f_h = f^\circ - |\Gamma^k_h|^{-1} \int_{\Gamma^k_h} f^\circ d\sigma_{hk}$ may be more convenient; cf. [26, p. 149]. Also, $(\cdot, \cdot)_{\partial \Gamma^k_h}$ denotes the $L_2$-inner product computed elementwise, and similarly $(v, w)_{\partial \Gamma^k_h} = \sum_{K \in \mathcal{T}_h} \int_{\partial K} vw \, ds$. We here use the notation $(\cdot, \cdot)_{\partial \Gamma^k_h}$ instead of $(\cdot, \cdot)_{\mathcal{E}^k_h}$ since not all of the quantities so integrated are single-valued on $\mathcal{E}^k_h$. Finally, note that if (3.5a)-(3.5d) are satisfied for $(v, w, \eta) \in V_h \times W_h \times M_h$, then they are also satisfied...
for \((v, w, \mu) \in V_h \times W_h \times M_h\). We use the more general definition for the sake of convenience below.

Finally, note that, as it is typical of HDG methods, given \(\hat{u}_h\), we may compute \((q_h, u_h)\) on the element \(K\) and \(\hat{q}_h \cdot \hat{n}_h\) on \(\partial K\) by solving:

\[
(q_h, v)_K - (u_h, \nabla \tau^h \cdot v)_K = - (\hat{u}_h, \hat{n}_h \cdot v)_{\partial K},
\]

\[
-(q_h, \nabla \tau^h w)_K + (\hat{q}_h \cdot \hat{n}_h, w)_{\partial K} = (f_h, w)_K,
\]

for all \((v, w) \in V_h(K) \times W_h(K)\). In order to determine \(\hat{u}_h \in M_h\), we solve the equation (3.5d) with the restriction \((u_h, 1)_{\Gamma^h} = 0\). Recall that the static condensation just described yields a globally coupled system for \(\hat{u}_h\) only. In practice one may enforce the condition \((u_h, 1)_{\Gamma^h} = 0\) by for example fixing \(\hat{u}_h\) at an arbitrary node, solving for \(\hat{u}_h\) and then \(u_h\), and subtracting \(\frac{1}{|\Gamma^h|} \int_{\Gamma^h} u_h \, \sigma_{hhk}\) from the resulting \(u_h\) and \(\hat{u}_h\). This procedure is equivalent to solving (3.5).

3.3. Existence and uniqueness of solutions. We now establish existence and uniqueness of the HDG solution for a very simple choice of local spaces and stabilization function. The below lemma does not cover all of the methods considered in our error analysis, but does demonstrate how numerical enforcement of the condition \(\int_T u \, \sigma = 0\) affects the existence and uniqueness proof.

**Lemma 3.1.** If the stabilization function \(\tau\) is strictly positive-valued, and \(\nabla W \subseteq V\), then the equations (3.5) defining the HDG method are uniquely solvable.

Note that the shape of the elements of these HDG methods is arbitrary. Note also that the local spaces allowed are quite general since only a simple inclusion condition between the local spaces \(W\) and \(V\) is required, and that no condition on the reference element \(K\) is needed. Moreover, there is no condition on the space of traces \(M(e)\). As we see in the proof below, the fact that the stabilization function \(\tau\) is strictly positive plays a crucial role in guaranteeing these properties.

**Proof.** We set \(f_h = 0\), and then show that we must have \(q_h = 0\), \(u_h = 0\), and \(\hat{u}_h = 0\). Letting \(v = q_h\), \(w = u_h\), and \(\eta = \hat{u}_h\) in (3.5a)–(3.5d), adding together the first two and last of these, integrating the second term by parts elementwise, and finally employing (3.5c), we find that

\[
0 = (q_h, q_h)_{\Gamma^h} - (u_h, \nabla \tau^h \cdot q_h)_{\Gamma^h} + (\hat{q}_h \cdot \hat{n}_h \cdot q_h)_{\partial \Gamma^h} - (q_h, \nabla \tau^h u_h)_{\Gamma^h}
\]

\[
+ (\hat{q}_h \cdot \hat{n}_h, u_h)_{\partial \Gamma^h} - (\hat{q}_h \cdot \hat{n}_h, \hat{u}_h)_{\partial \Gamma^h}
\]

\[
= (q_h, q_h)_{\Gamma^h} + (\nabla \tau^h u_h, q_h)_{\Gamma^h} - (u_h, q_h \cdot \hat{n}_h)_{\partial \Gamma^h}
\]

\[
+ (\hat{q}_h \cdot \hat{n}_h \cdot q_h)_{\partial \Gamma^h} - (q_h, \nabla \tau^h u_h)_{\Gamma^h} + (\hat{q}_h \cdot \hat{n}_h, u_h - \hat{u}_h)_{\partial \Gamma^h}
\]

\[
= (q_h, q_h)_{\Gamma^h} + (\hat{q}_h \cdot \hat{n}_h - q_h \cdot \hat{n}_h, u_h - \hat{u}_h)_{\partial \Gamma^h}
\]

\[
= (q_h, q_h)_{\Gamma^h} + \langle \tau (u_h - \hat{u}_h), u_h - \hat{u}_h \rangle_{\partial \Gamma^h}.
\]

Since \(\tau \geq 0\), we thus have \(q_h = 0\) and since \(\tau > 0\), we have \(u_h = \hat{u}_h\) on \(\partial \Gamma^h\).

Taking a test function \(v \in V_h\) with support on the element \(K \in \mathcal{T}^h\) in the first equation defining the HDG method, (3.5a), and integrating the second term by parts and canceling boundary terms then yields

\[\langle \nabla \tau^h u_h, v \rangle_K = 0\] for all \(v \in V_h(K)\).
By the definition of the space $W_h$, (3.2b) and (3.3b), we have that $u_h = u \circ A^{-1} \circ a_k^{-1}$ for some $u \in W$. Similarly, by the definition of the space $V_h$, we have that $v = P_a h P a_k$ for some $v \in V$. Then, by the identity (2.9b) with $S_0 := K$, $v_0 := u_h$ and $q_0 := v$, and $S_1 := K$, $v_1 := u$ and $q_1 := v$, we get that

$$\langle \nabla u, v \rangle_K = 0$$

for all $v \in V_h$. Since $\nabla W$ is included in $V_h$, we can take $v = \nabla \hat{u}$ so that we then have $\nabla \hat{u} = 0$. This implies that $u_h$ is a constant on $K \in T_k h$. But $u_h = \hat{u}_h$ is single-valued on $\partial T_k h$, so that $u_h$ is continuous and thus also constant. Finally, $u_h \in W^0_h$ implies that $\int_{T_k h} u_h = 0$, so that we obtain $u_h = 0$. This completes the proof.

3.4. Superconvergent HDG and mixed methods. The equations (3.5a)–(3.5d) are also uniquely solvable under weaker positivity conditions on the stabilization function $\tau$. For example, for the so-called single-face HDG method on triangles proposed in [14], it is possible to take $\tau$ equal to zero on all but one of the edges of each triangle. However, the method obtained by taking $\tau$ identically zero on $\partial T_h$ is not well defined. If we instead take the local spaces to be those defining (for example) the BDM or RT mixed methods and let $\tau = 0$, then (3.5a)–(3.5d) reduce to their hybridized version and are again uniquely solvable; the fact that our mixed methods are not conforming does not affect their solvability properties.

Although choosing the stabilization functions $\tau$ to be strictly positive on $\partial T_k h$ allows us to consider elements of fairly arbitrary shape (like nonlinear hexagons) and general local spaces (tailored, for example, to singularities of the exact solution), the resulting HDG methods are not necessarily guaranteed to converge with optimal orders of convergence in both the scalar and vector approximations. Here we are interested in considering HDG methods for which this actually happens for two reasons. The first is that the analysis is significantly more difficult. The second is that the error analysis immediately applies to the hybridized version of most known mixed methods, since these methods fit the general framework of the HDG methods with the stabilization function $\tau$ identically equal to zero.

Our error analysis applies to a whole family of HDG methods we define next. Its definition follows the general approach of superconvergent HDG and mixed methods for curved elements in flat surfaces considered in [19], with some modifications. One such change is our definition of the so-called reference stabilization function $\tilde{\tau}$, which in [19] is defined uniquely on the reference element $K$ and then transformed to each physical element. The resulting definition of $\tau$ is generally dual-valued on each edge. For implementation purposes, it may be more convenient to define a stabilization function $\hat{\tau}$ on the base mesh skeleton $\mathcal{E}_h$; we use $\hat{\tau} = 1$ in our numerical experiments below. The reference stabilization function $\tilde{\tau}$ then varies depending on the physical element, which is excluded in [19]. This presents no problem in the analysis as dependence of the necessary approximation bounds on $\tilde{\tau}$ has been well-analyzed [17, 18].

We suppose that the local spaces and the local so-called stabilization function $\tau$ satisfy the following set of assumptions.

**Assumptions A**: The following assumptions are used to obtain an error estimate of the vector approximation.
• Orthogonality properties of $\Pi$. On the reference element $K$, there exists a projection $\Pi(q, u) = (\Pi_V q, \Pi_W u) \in V \times W$ satisfying the following properties:

\[(A.1) \quad (\Pi_V q, v)_{K} = (q, v)_{K} \quad \text{for all } v \in \nabla W,\]
\[(A.2) \quad (\Pi_W u, w)_{K} = (u, w)_{K} \quad \text{for all } w \in \nabla \cdot V,\]
\[(A.3) \quad \langle \Pi_V q \cdot n + \tau \Pi_W u, \mu \rangle_{e} = \langle q \cdot n + \nabla P_M u, \mu \rangle_{e} \quad \text{for all } \mu \in M(e).\]

Here $\tau$ is a reference stabilization function which may vary with the particular physical element under consideration, but which on each edge $e$ satisfies either $\tau = 0$ for all occurrences or $c_e \leq \tau \leq C_e$ for some fixed constants $c_e, C_e > 0$. Also, the local projection $P_M$ is the standard $L^2$-orthogonal projection from $L^2(e)$ onto $M(e)$.

• Properties of the traces of the spaces. For the reference element $K$, and for any of its edges $e$,

\[(A.4) \quad V \cdot n | e \subset M(e),\]
\[(A.5) \quad W | e \subset M(e).\]

• The semi-positivity property of $\tau$. For the reference element $K$ and any of its edges $e$,

\[(A.6) \quad \langle \tau \mu, \mu \rangle_{e} \geq 0 \quad \text{for all } \mu \in M(e).\]

Assumptions $B$: The following assumptions are used to obtain a superconvergence result for the scalar approximation.

• For any $(q, u) \in [H^1(K)]^n \times H^1(K)$,

\[(B.1) \quad \|\Pi_V q - q\|_{K} \leq C_{app} (\|u\|_{1, K} + \|q\|_{1, K}).\]

• The reference space $W$ is not too small. On the reference element $K$, we have

\[(B.2) \quad \|P^0(K)\|^2 \subset \nabla W.\]

Note that our definition of $\Pi$ does not assume that the pair $(q, u)$ satisfies $q = -\nabla u$, although this relationship often holds in practice when the projection is employed. Also note that the right hand side of (B.1) involves both $q$ and $u$ even though the left hand side appears to only involve $q$. This is due to (A.3) in which the definition of $\Pi_V$ is assumed to employ information about $u$ unless $\tau = 0$ (as occurs in standard mixed methods).

$\Pi$ depends on the reference stabilization function $\tau$, and so the optimal choice of $C_{app}$ in (B.1) may vary with $\tau$. However, we require the existence of an absolute constant $C_{app}$ (depending on $c_e$ and $C_e$) which gives a uniform bound on $C_{app}$ independent of the particular choice of $\tau$ meeting our assumptions. This assumption is met for the element choices given below.
In the next section we also list additional Assumptions G which ensure that numerical enforcement of the condition $\int_\Gamma u \, d\sigma = 0$ does not degrade convergence rates. We delay stating them as one of them does not involve only the reference space and elements as do Assumptions A and B.

In Table 1, taken from [18], we display the main examples of HDG and mixed methods which satisfy all of the above assumptions. They include several superconvergent HDG methods as well as the well-known mixed methods of Raviart-Thomas, $\text{RT}_r$, of Brezzi-Douglas-Marini, $\text{BDM}_r$, and of Brezzi-Douglas-Fortin-Marini, $\text{BDFM}_{r+1}$ defined on triangles; the mixed methods $\text{BDFM}_{[r+1]}$, $\text{BDM}_{[r]}$ on squares; and the mixed methods using tensor product spaces $\text{RT}_{[r]}$ and $\text{TNT}_{[r]}$ defined also on squares. Our analysis also includes the lowest-order ($r = 0$) versions of the below spaces as appropriate, though in these cases superconvergence is not obtained.

<table>
<thead>
<tr>
<th>method</th>
<th>$V$</th>
<th>$W$</th>
<th>$\tau$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{BDFM}_{r+1}$</td>
<td>$P^{r+1}(K)$ \setminus { $x^{r+1}_1 }$</td>
<td>$P^r(K)$</td>
<td>0</td>
</tr>
<tr>
<td>$\text{RT}_r$</td>
<td>$P^r(K)^2 \oplus \mathbb{P}^{r}(\partial K)$</td>
<td>$P^r(K)$</td>
<td>0</td>
</tr>
<tr>
<td>$\text{HDG}_r$</td>
<td>$P^r(K)^2$</td>
<td>$P^r(K)$</td>
<td>$&gt; 0$, $O(1)$</td>
</tr>
<tr>
<td>$\text{BDM}_r$</td>
<td>$P^r(K)^2$</td>
<td>$P^{r-1}(K)$</td>
<td>0</td>
</tr>
</tbody>
</table>

$K$ is the reference triangle.

<table>
<thead>
<tr>
<th>method</th>
<th>$V$</th>
<th>$W$</th>
<th>$\tau$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{BDFM}_{[r+1]}$</td>
<td>$P^{r+1}(K) \setminus { x^{r+1}_1 }$ \times $P^r(K)^2 \setminus { x^{r+1}_1 }$</td>
<td>$P^r(K)$</td>
<td>0</td>
</tr>
<tr>
<td>$\text{HDG}_{[r]}$</td>
<td>$P^r(K)^2 \oplus \mathbb{x}_1 \mathbb{P}^{r}(\partial K)$</td>
<td>$P^r(K)$</td>
<td>$&gt; 0$, $O(1)$</td>
</tr>
<tr>
<td>$\text{BDM}_{[r]}$</td>
<td>$P^r(K)^2 \oplus \mathbb{x}_1 \mathbb{P}^{r}(\partial K)$</td>
<td>$P^{r-1}(K)$</td>
<td>0</td>
</tr>
</tbody>
</table>

$K$ is the unit square.

<table>
<thead>
<tr>
<th>method</th>
<th>$V$</th>
<th>$W$</th>
<th>$\tau$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{RT}_{[r]}$</td>
<td>$P^{r+1,r}(K)$ \times $P^{r,r+1}(K)$</td>
<td>$Q^r(K)$</td>
<td>0</td>
</tr>
<tr>
<td>$\text{TNT}_{[r]}$</td>
<td>$Q^r(K)^2 \oplus \mathbb{H}_3(K)$</td>
<td>$Q^r(K)$</td>
<td>0</td>
</tr>
<tr>
<td>$\text{HDG}_{[r]}$</td>
<td>$Q^r(K)^2 \oplus H_3(K)$</td>
<td>$Q^r(K) \oplus H_2(K)$</td>
<td>$&gt; 0$, $O(1)$</td>
</tr>
</tbody>
</table>

$K$ is the unit square.

Let us recall that by $\mathbb{P}^r(D)$, we denote the space of polynomials of total degree $r$ defined on $D$. By $\mathbb{P}^r(D)$, we denote the space of homogeneous polynomials of degree $r$ defined on $D$. By $\mathbb{P}^r(\partial K)$, we denote the functions whose restriction to each edge $e$ of $K$ belong to $\mathbb{P}^r(e)$. By $\mathbb{P}^{l_1,l_2}(D)$, we denote the space of polynomials
of degree \( \ell_i \) on the \( i-th \) variable, \( i = 1, 2 \). By \( \mathbb{Q}^r(D) \), we denote the space of polynomials of degree \( r \) in each variable defined on \( D \). Finally, the spaces \( H_2 \) and \( H_3 \) of \( H(\text{div}) \)-conforming bubble functions are given by

\[
H_2(K) := \text{span}\{(\hat{B}_{r+1}(\mathcal{E}_1), 0), (0, \hat{B}_{r+1}(\mathcal{E}_2))\},
\]

\[
H_3(K) := H_2 \oplus \text{span}\{(\hat{B}_{r+1}(\mathcal{E}_1) \tilde{P}_r(\mathcal{E}_2), \tilde{P}_r(\mathcal{E}_1) \hat{B}_{r+1}(\mathcal{E}_2))\},
\]

where \( \hat{B}_{r+1} := (\hat{P}_{r+1} - \hat{P}_{r-1})/(4\ell + 2) \) for \( \ell > 0 \), and \( \hat{P}_r(z) := P_r(-1 + 2z) \), where \( P_r \) is the Legendre polynomial of degree \( r \). Note that, since \( \hat{B}_{r+1}(0) = \hat{B}_{r+1}(1) = 0 \), the normal trace of the basis functions of \( H_2(K) \) and \( H_3(K) \) are zero on the border of the unit square \( K \).

4. Error estimates

In this section, we prove error estimates for the surface HDG method. Following [19], we first rewrite the Assumptions on the local spaces and stability function on the reference element for the physical elements. We then find the equations satisfied by the projections of the errors, carry out an “energy” argument, and finally employ a duality argument in order to complete our analysis.

4.1. The assumptions on the physical elements. We begin our analysis by rewriting Assumptions A and B on the elements of \( T_k^h \). Below we write the assumptions for \( u^e \) and \( \bar{q} = P_{a^{-1}}^e \) for \( u \in H^1(\Gamma) \), \( q \in H(\text{div} \Gamma) \) since this is how we shall apply \( \Pi \) the assumptions, but they could equivalently be written for any \( (w, v) \in H^1(\Gamma)^r \times H(\text{div} \Gamma)^r \). We need to introduce two sets of projections. The first is the following. For any edge \( e \in \mathcal{E}_h \), we define the projections \( P_M \) and \( P_N \) from \( L^2(e) \) onto \( M_h(e) \) and \( N_h(e) \), respectively, by

\[
(4.1a) \quad \langle u^e, \lambda \rangle_e = \langle P_M u^e, \lambda \rangle_e \quad \forall \lambda \in N_h(e),
\]

\[
(4.1b) \quad \langle \bar{q} \cdot n, \mu \rangle_e = \langle P_N(\bar{q} \cdot n), \mu \rangle_e \quad \forall \mu \in M_h(e).
\]

We now make the assumption that

\[
(4.2) \quad \tau = h_k^{-1} \tilde{\tau} \circ A^{-1} \circ (a_k)^{-1} / J_e
\]

for some allowable reference stabilization function \( \tilde{\tau} \). Here \( J_e = \mu_k^e |\mathcal{e}|/|\mathcal{e}| \), where \( \mu_k^e \) is defined in (3.4) and \( \mathcal{e} \in \mathcal{E}_h \) corresponds to \( e \in \mathcal{E}_h^k \). Shape regularity guarantees that \( h_k^{-1}/J_e \simeq 1 \), so we may for example choose \( \tau = 1/\mu_k^e \) (which is convenient for implementation) while preserving the uniform bound \( \tilde{\tau} \simeq 1 \) which is needed for analysis. Then

\[
(4.3) \quad \tau P_M u^e = P_N(\tau u^e).
\]

The second set of projections is the following. For any \( q \in H(\text{div} \Gamma) \) and \( u \in H^1(\Gamma) \), we set

\[
(4.4a) \quad \Pi q := h_k^{n-1} P_{a_k}^e P_A \prod h_k^{1-n} P_{A^{-1}} \Pi (a_k)^{-1} \bar{q},
\]

\[
(4.4b) \quad \Pi u^e := C_{a_k} C_A \prod C_{A^{-1}} C_{(a_k)^{-1}} u^e.
\]

where \( C_B \) is the composition operator \( C_B z := z \circ B \).

With this notation, we can rewrite Assumptions A as follows.
Lemma 4.1 (Assumptions A). If Assumptions A on the reference elements hold, then we have for any $K \in T_h^k$

\begin{align}
\langle \Pi \tilde{q}, v \rangle_K &= (\tilde{q}, v)_K \quad \forall v \in \nabla_h W_h(K), \\
\langle \Pi u^\ell, w \rangle_K &= (u^\ell, w)_K \quad \forall w \in \nabla_h V(K), \\
\langle \Pi q \cdot \tilde{n}_h + \tau \Pi u^\ell, \mu \rangle_e &= (P_N(\tilde{q} \cdot \tilde{n}_h) + \tau P_M u^\ell, \mu)_e \quad \forall \mu \in M_h(e)
\end{align}

for all edges $e$ of $K$, provided (4.2) holds, in addition and

\[ P_M u^\ell = C_{a_k} C_A P_M C_{(a_k)^{-1}} u^\ell. \]

Moreover, we have

\begin{align}
V(K) \cdot \tilde{n}_h|_e &\subset N_h(e) \quad \text{for all edges } e \text{ of } K, \\
W_h(K)|_e &\subset M_h(e) \quad \text{for all edges } e \text{ of } K.
\end{align}

Finally, for any edge $e \in E_h^k$, we have

\[ \langle \tau \mu, \mu \rangle_e \geq 0 \quad \forall \mu \in M_h(e). \]

The proof is just a slight variation of the corresponding result in [19] thanks to the Piola identities (2.9).

We next rewrite the consequences of Assumption (B.1) and Assumption (B.2) we are going to use.

Lemma 4.2. If Assumption (B.1) holds, then for each element $K$ and any $(\tilde{q}, u^\ell) \in H^1(K) \times H^1(K)$,

\[ \| \Pi \tilde{q} - \tilde{q} \|_K \leq C_{\text{app}} h_K (\| u^\ell \|_{1,K} + \| \tilde{q} \|_{1,K}). \]

Remark 4.3. The corresponding Lemma 4.2 in [19] assumes that the meshes under consideration are “1-regular”, i.e., that the first three properties of (2.2) are satisfied. However, the fourth property (2.2d) is in fact also needed (the mesh must be “2-regular”) due to the Piola transform. For this reason, and also because the proof is slightly more complex in the case of surfaces, we sketch the proof of Lemma 4.2.

Proof. Fix $K \in T_h^k$, and let $B = a_k \circ A_K$. As is remarked at the end of §2.3, (2.2) hold with $B$ replacing $A$ in all quantities. We may compute precisely as in the corresponding proof of Lemma 4.2 in [19] that

\[ \| \Pi \tilde{q} - \tilde{q} \|_K \lesssim h_K^{n/2} \| q \|_{1,K} + h_K \| u^\ell \|_{1,K}, \]

where $\tilde{q} = h^{n-1} P_B \tilde{q} = h^{n-1} DB \tilde{q} / JB$. Regarding $q$ as a column $n$-vector (with $n = 2$) and $DB$ as a $(n+1) \times n$ matrix, we have from a short calculation that $q = h^{1-n} JB(DB^T DB)^{-1} DB^T \tilde{q} = h^{1-n} \frac{1}{J^2} \text{cof}(DB^T DB) DB^T \tilde{q}$, where \text{cof}(DB^T DB) is the cofactor matrix of $DB^T DB$. Thus

\[ |D\tilde{q}| \lesssim h^{1-n} \left( |D \left[ \frac{1}{JB} \text{cof}(DB^T DB) DB^T \right] \tilde{q}| \right), \]

(4.9)

where the last bound follows by combining an elementary calculation with (2.2). Inserting this inequality into (4.8) and transforming the integral from $K$ to $K$ completes the proof. \qed
Lemma 4.4. If Assumption (B.2) holds, then for any function \( z \in H^2(K) \), there exists an \( z_0 \in W^0_h \) such that, such that

\[
\| \nabla_{\Gamma_h^k}(z - z_0) \|_K \leq Ch_K \| z \|_{2,K}.
\]

Remark 4.5. The corresponding lemma, also labeled Lemma 4.3, in [18] gives a result for approximating functions \( z \in [H^1(K)]^2 \) by functions \( z_0 \in V_0 \), where \( V_0 \) consists of functions in \( V_h \) satisfying \((q - \Pi q, v)_K = 0, q \in [H^1(K)]^2\). However, \( V_0 \) may be empty even when Assumption (B.2) holds. To see this, note that the orthogonality relationship (4.5a) holds for functions in \( \nabla_{\Gamma_h^k} W_h(K) \). Such functions are obtained by employing the chain rule under the mapping \( a_k \circ A \), whereas functions in \( V_h \) and thus \( V_0 \) are obtained by Piola transform. For non-affine reference transformations, this generally implies \( \nabla_{\Gamma_h^k} W_h(K) \cap V_h(K) = \emptyset \) even if \( \nabla W \cap V \neq \emptyset \). The proof in [18] may easily be corrected as we do here.

We now state our Assumptions G, which are necessary to ensure that approximate enforcement of the condition \( \int_{\Gamma_h^k} u \, d\sigma = 0 \) does not degrade convergence rates (in the case of Assumption G.1) or superconvergence rates (in the case of Assumption G.2). Assumptions G should not be necessary were a similar analysis carried out for a coercive operator, e.g., for a Dirichlet problem on a non-closed surface.

Assumptions G:

- The surface \( \Gamma_h^k \) sufficiently resolves \( \Gamma \).

\[ (G.1) \quad \| \frac{1}{\mu_{hk}} \|_{L^\infty(\Gamma_h^k)}^{1/2} \| \mu_{hk} \|_{L^\infty(\Gamma_h^k)}^{1/2} \| 1 - \frac{1}{\mu_{hk}} \|_{L^\infty(\Gamma_h^k)} \leq \frac{1}{2}. \]

- The reference space \( V \) is not too small.

\[ (G.2) \quad \mathbb{P}^0 \subset \nabla \cdot V. \]

Assumption G.1 holds asymptotically by (2.5c). However, we instead view G.1 as a geometric condition requiring that \( \Gamma_h^k \) sufficiently resolve \( \Gamma \). This condition requires more than our previous assumption \( \Gamma_h^k \subset U \), but it is relatively weak and can be explicitly checked if enough information about \( \Gamma \) is available. The consequences of Assumption G.2 are given in the following lemma.

Lemma 4.6. If Assumption G.2 holds, then for \( K \in \mathcal{T}_h^k \)

\[ (4.10) \quad \inf_{\lambda \in \nabla_{\Gamma_h^k} V(K)} \| 1 - \lambda \|_{L^\infty(K)} \lesssim h. \]

Proof. Given \( K \in \mathcal{T}_h^k \), let \( B = a_k \circ A_K \). Because \( V_h \) is the Piola transform of \( V \) to \( \Gamma_h^k \), (2.8) implies that \( \lambda \in \nabla_{\Gamma_h^k} \cdot V_h \) has the form

\[ (4.11) \quad \frac{1}{JB^A \lambda} \]
where $\lambda \in \nabla \cdot V$ and $JB$ is the Jacobian of $B$. Using $1 \in V$, we choose $\lambda = \frac{1}{|K|} \int_K JB \, dx$ below to compute

$$
\inf_{\lambda \in \nabla \cdot V_h} \|1 - \lambda\|_{L^\infty(K)} \lesssim \inf_{\Delta \in V} \|J B^{-1}\|_{L^\infty(K)} \|J B - \Delta\|_{L^\infty(K)}
$$

(4.12)

$$
\lesssim h^{-n} \|D(J B)\|_{L^\infty(K)}.
$$

In the last inequality we have employed the observation at the end of Section 2.3 that (2.2) holds for $B$. A short calculation involving (2.2) also yields $\|D(J B)\|_{L^\infty(K)} \lesssim h^{n+1}$, completing the proof. \hfill \square

4.2. Error equations. We now develop error equations for our method. First we define “projections” of the error:

$$
\epsilon_q = \Pi \hat{q} - q_h,
$$

(4.13a)

$$
\epsilon_u = \Pi u^\ell - u_h,
$$

(4.13b)

$$
\epsilon_{\bar{u}} = P_M u^\ell - \tilde{u}_h,
$$

(4.13c)

$$
\bar{n}_h \cdot \epsilon_q = P_N (\hat{q} \cdot \bar{n}_h) - \tilde{q}_h \cdot \bar{n}_h.
$$

(4.13d)

We compute using (4.5c), (3.5c), and (4.3) that

$$
\bar{n}_h \cdot \epsilon_q = P_N (\hat{q} \cdot \bar{n}_h) - \tilde{q}_h \cdot \bar{n}_h
$$

$$
= P_N (\hat{q} \cdot \bar{n}_h + \tau u^\ell) - \tau P_M (u^\ell) - \tilde{q}_h \cdot \bar{n}_h
$$

$$
= \Pi \hat{q} \cdot \bar{n}_h + \tau (\Pi u^\ell - P_M u^\ell) - \tilde{q}_h \cdot \bar{n}_h
$$

$$
= (\Pi \hat{q} \cdot \bar{n}_h - q_h \cdot \bar{n}_h) + \tau ((\Pi u^\ell - u_h) - (P_M u^\ell - \tilde{u}_h))
$$

$$
= \epsilon_q \cdot \bar{n}_h + \tau (\epsilon_u - \epsilon_{\bar{u}}).
$$

(4.14)

Using (3.5a), (4.5b), (4.6a), and integrating by parts, we compute for $v \in V_h$ that

$$
(\epsilon_q, v)_{\Gamma_h^k} - (\epsilon_u, \nabla \Gamma_h^k \cdot v)_{\Gamma_h^k} + (\epsilon_{\bar{u}}, v \cdot \bar{n}_h)_{\partial \Gamma_h^k}
$$

$$
= (\Pi \hat{q}, v)_{\Gamma_h^k} - (\Pi u^\ell, \nabla \Gamma_h^k \cdot v)_{\Gamma_h^k} + (P_M u^\ell, v \cdot \bar{n}_h)_{\partial \Gamma_h^k}
$$

$$
= (\Pi \hat{q}, v)_{\Gamma_h^k} - (u^\ell, \nabla \Gamma_h^k \cdot v)_{\Gamma_h^k} + (u^\ell, v \cdot \bar{n}_h)_{\partial \Gamma_h^k}
$$

$$
= (\Pi \hat{q} + \nabla \Gamma_h^k u^\ell, v)_{\Gamma_h^k}.
$$

(4.15)

Using (3.5b), (4.5a), (4.1b), (2.9a), and integrating by parts, we find for $w \in W_h$ that

$$
-(\epsilon_q, \nabla \Gamma_h^k w)_{\Gamma_h^k} + (\epsilon_q \cdot \bar{n}_h, w)_{\partial \Gamma_h^k}
$$

$$
= -(\Pi \hat{q}, \nabla \Gamma_h^k w)_{\Gamma_h^k} + (P_N (\hat{q} \cdot \bar{n}_h), w)_{\partial \Gamma_h^k} - (f_h, w)_{\Gamma_h^k}
$$

$$
= - (\tilde{q}, \nabla \Gamma_h^k w)_{\Gamma_h^k} + (\tilde{q} \cdot \bar{n}_h, w)_{\partial \Gamma_h^k} - (f_h, w)_{\Gamma_h^k}
$$

$$
= (\nabla \Gamma_h^k \cdot \tilde{q}, w)_{\Gamma_h^k} - (f_h, w)_{\Gamma_h^k}
$$

$$
= (\nabla \Gamma \cdot q, w^\ell)_{\Gamma} - (f_h, w)_{\Gamma_h^k}
$$

$$
= (f, w^\ell)_{\Gamma} - (f_h, w)_{\Gamma_h^k}
$$

$$
= (\mu_h, f^\ell - f_h, w)_{\Gamma_h^k}.
$$

(4.16)
Finally, (3.5d), the property (2.9c) of the Piola transform, and continuity of the normal traces of \( q \) on \( \Gamma \) yield for \( \eta \in M_h \) that

\[
-\langle \epsilon q \cdot \hat{n}_h, \eta \rangle_{\partial \Gamma^k_h} = -\langle P_N(\hat{q} \cdot \hat{n}_h), \eta \rangle_{\partial \Gamma^k_h} \\
(4.17)
\]

\[
= -\langle \hat{q} \cdot \hat{n}_h, \eta \rangle_{\partial \Gamma^k_h} \\
= -\langle q \cdot n, \eta_v \rangle_{a(\partial \Gamma^k_h)} = 0.
\]

In the corresponding Euclidean HDG error equations the right-hand-sides of (4.15) and (4.16) vanish. These nonzero right-hand-sides here encode geometric errors.

4.3. “Energy” argument. We first insert (4.14) into (4.16) to obtain

\[
-(\epsilon_q, \nabla_t \ell_w)^{\Gamma^k_h} + (\epsilon_q \cdot \hat{n}_h + \tau(\epsilon_u - \epsilon_{\hat{u}}), w)^{\partial \Gamma^k_h} \\
(4.18)
\]

\[
= (\mu_{hk} f^\ell - f_h, w)^{\Gamma^k_h}, \ w \in W_h.
\]

Similarly, from (4.14) and (4.17) we obtain

\[
-(\epsilon_q \cdot \hat{n}_h + \tau(\epsilon_u - \epsilon_{\hat{u}}), \eta)^{\partial \Gamma^k_h} = 0, \ \eta \in M_h.
\]

We set \( v = \epsilon_q, \ w = \epsilon_u, \) and \( \eta = \epsilon_{\hat{u}} \) in (4.15), (4.18), and (4.19), respectively, add these three equations together, and then integrate by parts in order to obtain

\[
(\Pi \hat{q} + \nabla_t \ell_w)^{\Gamma^k_h} + (\mu_{hk} f^\ell - f_h, \epsilon_u)^{\Gamma^k_h} = (\epsilon_q, \epsilon_q)^{\Gamma^k_h} - (\epsilon_q, \nabla_t \ell_w)^{\Gamma^k_h} \\
(4.20)
\]

\[
+ (\epsilon_q \cdot \hat{n}_h + \tau(\epsilon_u - \epsilon_{\hat{u}}), \epsilon_u)^{\partial \Gamma^k_h} \\
- (\epsilon_q \cdot \hat{n}_h + \tau(\epsilon_u - \epsilon_{\hat{u}}), \epsilon_u)^{\partial \Gamma^k_h} \\
= (\epsilon_q, \epsilon_q)^{\Gamma^k_h} + (\tau(\epsilon_u - \epsilon_{\hat{u}}), \epsilon_u - \epsilon_{\hat{u}})^{\partial \Gamma^k_h}.
\]

Writing \( \Pi \hat{q} + \nabla_t \ell_w = (\Pi \hat{q} - \hat{q}) + (\hat{q} + \nabla_t \ell_w) \), combining (4.20) with Proposition 2.2, noting that we have used only the Assumptions A, and recalling that \( \tau \geq 0 \), we obtain the following lemma.

Lemma 4.7. Let Assumptions A hold. Then

\[
(\epsilon_q, \epsilon_q)^{\Gamma^k_h} + (\tau(\epsilon_u - \epsilon_{\hat{u}}), \epsilon_u - \epsilon_{\hat{u}})^{\partial \Gamma^k_h} \\
(4.21)
\]

\[
= (\Pi \hat{q} + \nabla_t \ell_w, \epsilon_q)^{\Gamma^k_h} + (\mu_{hk} f^\ell - f_h, \epsilon_u)^{\Gamma^k_h}.
\]

Thus

\[
\|\Pi \hat{q} - q_h\| \lesssim \|\Pi \hat{q} - \hat{q}\|_{L^2(\Gamma^k_h)} + Ch^k \|q^\ell\|_{L^2(\Gamma^k_h)} \\
(4.22)
\]

\[
+ \sqrt{\|\mu_{hk} f^\ell - f_h\|_{L^2(\Gamma^k_h)} \|\epsilon_u\|_{L^2(\Gamma^k_h)}},
\]

and if the choice \( f_h = \mu_{hk} f^\ell \) is made, there holds

\[
\|\Pi \hat{q} - q_h\|_{L^2(\Gamma^k_h)} \leq \|\Pi \hat{q} - \hat{q}\|_{L^2(\Gamma^k_h)} + Ch^k \|q^\ell\|_{L^2(\Gamma^k_h)}.
\]

Unlike the corresponding analysis for Euclidean domains given in [17, 19], (4.21) and (4.22) indicates a coupling between \( \epsilon_q \) and \( \epsilon_u \) and thus the need for a duality argument in order to obtain a final bound for \( \epsilon_q \). Choosing the discrete data \( f_h = \mu_{hk} f^\ell \) yields the pure “energy” bound (4.23) containing only approximation and geometric error terms. While convenient for theoretical purposes, this choice
of discrete data may be expensive or impossible to compute in practice, depending on the amount of information about the continuous surface \( \Gamma \) that is available.

4.4. **Duality argument.** We next bound \( \epsilon_\ast \). Let \((\Phi, \Psi)\) solve the system

\[
\Phi + \nabla_\Gamma \Psi = 0 \text{ on } \Gamma, \\
\nabla_\Gamma \cdot \Phi = \Theta \text{ on } \Gamma,
\]

(4.24)

\[
\int_\Gamma \Theta \, d\sigma = \int_\Gamma \Psi \, d\sigma = 0.
\]

Following [17], we then define

\[
H(\Theta) = \sup_{\|\Theta\|_{L^2(\Gamma)} = 1} (\|\Pi \Phi - \bar{\Phi}\|_{L^2(\Gamma^k_h)} + \inf_{\Psi_h \in W^2_h} \|\nabla_\Gamma \Psi (\Psi^\ell - \Psi_h)\|_{L^2(\Gamma^k_h)}),
\]

(4.25)

where \(\Theta, \Phi, \text{ and } \Psi\) are related by (4.24). (2.1) implies that

\[
\|\Phi\|_{H^1(\Gamma)} + \|\Psi\|_{H^2(\Gamma)} \lesssim \|\Theta\|_{L^2(\Gamma)}.
\]

By Lemma 4.2,

(4.27)

Assumption B.1 \(\Rightarrow H(\Theta) \lesssim 1\).

Lemmas 4.2 and 4.4 combined also yield

(4.28)

Assumptions B.1 and B.2 \(\Rightarrow H(\Theta) \lesssim h\).

In addition, when Assumption B.1 holds then (4.26) implies that

\[
\|\hat{\Phi}\|_{L^2(\Gamma^k_h)} + \|\Pi \hat{\Phi}\|_{L^2(\Gamma^k_h)} + \|\Psi^\ell\|_{L^2(\Gamma^k_h)} + \|\Pi \Psi^\ell\|_{L^2(\Gamma^k_h)} \lesssim 1.
\]

We now state our basic duality lemma.

**Lemma 4.8.** Let Assumptions A and G.1 hold along with (2.1). Then

\[
\|\epsilon_\ast\|_{L^2(\Gamma^k_h)} \lesssim (H(\Theta) + h)(\|\epsilon_\ast\|_{L^2(\Gamma^k_h)} + \|\bar{q} - \Pi \bar{q}\|_{L^2(\Gamma^k_h)})
\]

\[
+ h^{k+1} \|q^\ell\|_{L^2(\Gamma^k_h)} + \|\mu_k f^\ell - f_h\|_{L^2(\Gamma^k_h)}
\]

\[
+ \|u^\ell - \Pi u^\ell\|_{L^2(\Gamma^k_h)} \inf_{\lambda \in \nabla_\Gamma \cdot V_h} \|1 - \lambda\|_{L^2(\Gamma^k_h)}.
\]

(4.30)

**Proof.** Taking \(\Theta \in L^2(\Gamma)\) with \(\|\Theta\|_{L^2(\Gamma)} = 1\) in (4.24), letting \(\tilde{\Phi}\) be the Piola transform of the resulting solution to \(\Gamma^k_h\) and then employing (2.9b), integrating by parts, using (4.5a) and (4.5c), and finally again integrating by parts, we find that

\[
\langle \epsilon_\ast^\ell, \Theta \rangle_{\Gamma^k_h} = (\epsilon_{\ast, \nabla_\Gamma \cdot \Phi})_{\Gamma^k_h} = (\epsilon_{\ast, \nabla_\Gamma \cdot \Phi})_{\Gamma^k_h}
\]

\[
= - (\nabla_\Gamma \epsilon_{\ast, \tilde{\Phi}})_{\Gamma^k_h} + \langle \epsilon_{\ast, \tilde{\Phi} \cdot \bar{\Phi}} \rangle_{\partial \Gamma^k_h}
\]

\[
= - (\nabla_\Gamma \epsilon_{\ast, \tilde{\Phi}^\ell})_{\Gamma^k_h} + \langle \epsilon_{\ast, \tilde{\Phi} \cdot \bar{\Phi}} \rangle_{\partial \Gamma^k_h} + (\epsilon_{\ast, \tau (\Pi \Psi^\ell - \Psi^\ell)})_{\partial \Gamma^k_h}
\]

\[
= (\epsilon_{\ast, \nabla_\Gamma \cdot \Pi \tilde{\Phi}})_{\Gamma^k_h} + \langle \epsilon_{\ast, \tau (\Pi \Psi^\ell - \Psi^\ell)} \rangle_{\partial \Gamma^k_h}. \tag{4.31}
\]

Employing (4.15) and collecting terms while recalling that \(\epsilon_\ast\) is single-valued on \(E^k_h\) so that \(\langle \epsilon_\ast, \Phi \cdot \bar{\Phi} \rangle_{\partial \Gamma^k_h} = 0\), we find that

\[
(\epsilon_{\ast, \nabla_\Gamma \cdot \Pi \tilde{\Phi}})_{\Gamma^k_h} = (\epsilon_{\ast, \Pi \tilde{\Phi}})_{\Gamma^k_h} + (\epsilon_{\ast, \Pi \tilde{\Phi} \cdot \bar{\Phi}})_{\partial \Gamma^k_h}
\]

\[
- (\Pi \bar{q} + \nabla_\Gamma u^\ell, \Pi \tilde{\Phi})_{\Gamma^k_h}
\]

\[
= (-\nabla_\Gamma u^\ell - q_h, \Pi \tilde{\Phi})_{\Gamma^k_h} + \langle \epsilon_\ast, (\Pi \tilde{\Phi} - \tilde{\Phi}) \cdot \bar{\Phi} \rangle_{\partial \Gamma^k_h}. \tag{4.32}
\]

\[
(\epsilon_{\ast, \nabla_\Gamma \cdot \Pi \tilde{\Phi}})_{\Gamma^k_h} = (\epsilon_{\ast, \Pi \tilde{\Phi}})_{\Gamma^k_h} + (\epsilon_{\ast, \Pi \tilde{\Phi} \cdot \bar{\Phi}})_{\partial \Gamma^k_h}
\]
Collecting (4.31) and (4.32) yields
\[
(\epsilon_u, \Theta) = (-\nabla_{\Gamma_h} u^\ell - q_h, \Pi \hat{\Phi})_{\Gamma_h}^k + (\epsilon_u, (\Pi \hat{\Phi} - \tilde{\Phi}) \cdot \tilde{n}_h)_{\partial \Gamma_h^k} + (\epsilon_u, \tau (\Pi \Psi^\ell - \Psi^\ell))_{\partial \Gamma_h^k}.
\]
(4.33)

Employing (4.5c), the definition of $P_M$, (4.14), (4.16) along with (4.17), integrating by parts, employing (4.5b) and then finally integrating by parts again yields
\[
(\epsilon_u, (\Pi \hat{\Phi} - \tilde{\Phi}) \cdot \tilde{n}_h)_{\partial \Gamma_h^k} + (\epsilon_u, \tau (\Pi \Psi^\ell - \Psi^\ell))_{\partial \Gamma_h^k}
\]
\[
= (\epsilon_u - \epsilon_u, \tau (\Pi \Psi^\ell - \Psi^\ell))_{\Gamma_h^k}
\]
\[
= (\epsilon_u - \epsilon_u, \tau (\Pi \Psi^\ell - \Psi^\ell))_{\Gamma_h^k} - (\epsilon_u - \epsilon_u, \tau P_M \Psi^\ell)_{\Gamma_h^k}
\]
\[
= (\epsilon_u - \epsilon_u, \tau (\Pi \Psi^\ell - \Psi^\ell))_{\Gamma_h^k} - (\epsilon_u - \epsilon_u, \tau \tilde{\Phi} P_M \Psi^\ell)_{\Gamma_h^k}.
\]
(4.34)

Combining (4.33) and (4.34), carrying out a series of elementary manipulations, and then employing (4.5a) to insert an arbitrary $\Psi_h \in W_h^0$ yields
\[
(\epsilon^\ell_u, \Theta) = (-\nabla_{\Gamma_h} u^\ell - q_h, \Pi \hat{\Phi})_{\Gamma_h^k} + (\epsilon^\ell_u, \nabla_{\Gamma_h} \Psi^\ell)_{\Gamma_h^k} + (\mu_{hk} f^\ell - f_h, \Pi \Psi^\ell)_{\Gamma_h^k}
\]
\[
= (-\nabla_{\Gamma_h} u^\ell - \tilde{q}, \Pi \hat{\Phi})_{\Gamma_h^k} + (\tilde{q} - q_h, \Pi \hat{\Phi} - \tilde{\Phi})_{\Gamma_h^k} + (\tilde{q} - q_h, \tilde{\Phi})_{\Gamma_h^k}
\]
\[
+ (\Pi \tilde{q} - q_h, \nabla_{\Gamma_h} \Psi^\ell)_{\Gamma_h^k} + (\mu_{hk} f^\ell - f_h, \Pi \Psi^\ell)_{\Gamma_h^k}
\]
\[
= (-\nabla_{\Gamma_h} u^\ell - \tilde{q}, \Pi \hat{\Phi})_{\Gamma_h^k} + (\tilde{q} - q_h, \Pi \hat{\Phi} - \tilde{\Phi})_{\Gamma_h^k}
\]
\[
+ (\tilde{q} - q_h, \tilde{\Phi} + \nabla_{\Gamma_h} \Psi^\ell)_{\Gamma_h^k} + (\Pi \tilde{q} - q_h, \nabla_{\Gamma_h} (\Psi^\ell - \Psi_h))_{\Gamma_h^k}
\]
\[
+ (\mu_{hk} f^\ell - f_h, \Pi \Psi^\ell)_{\Gamma_h^k}.
\]
(4.35)

Applying (2.12) to (4.35) and employing (4.25) and (4.29) yields
\[
|\langle \epsilon^\ell_u, \Theta \rangle | \leq \| \tilde{q} - q_h \|_{L^2(\Gamma_h^k)} \| \Pi \hat{\Phi} - \tilde{\Phi} \|_{L^2(\Gamma_h^k)}
\]
\[
+ \| \tilde{q} - \Pi \tilde{q} \|_{L^2(\Gamma_h^k)} \| \nabla_{\Gamma_h} (\Psi^\ell - \Psi_h) \|_{L^2(\Gamma_h^k)}
\]
\[
+ h^{k+1} \| q^\ell \|_{L^2(\Gamma_h^k)} \| \Pi \hat{\Phi} \|_{L^2(\Gamma_h^k)} + \| \tilde{q} - q_h \|_{L^2(\Gamma_h^k)} \| \Psi^\ell \|_{L^2(\Gamma_h^k)}
\]
\[
+ \| \mu_{hk} f^\ell - f_h \|_{L^2(\Gamma_h^k)} \| \Pi \Psi^\ell \|_{L^2(\Gamma_h^k)}
\]
\[
\leq [H(\Theta) + h] (\| \epsilon_\ell_u \|_{L^2(\Gamma_h^k)} + \| \tilde{q} - \Pi \tilde{q} \|_{L^2(\Gamma_h^k)})
\]
\[
+ h^{k+1} \| q^\ell \|_{L^2(\Gamma_h^k)} + \| \mu_{hk} f^\ell - f_h \|_{L^2(\Gamma_h^k)}.
\]
(4.36)

Let $\epsilon^\ell_u = \frac{1}{|\Gamma|} \int_{\Gamma} \epsilon^\ell_u \, d\sigma$. We next observe that
\[
\| \epsilon_u \|_{L^2(\Gamma_h^k)} \leq \| \frac{1}{\mu_{hk}} L^\infty(\Gamma_h^k) \| \epsilon^\ell_u \|_{L^2(\Gamma)}
\]
\[
\leq \| \frac{1}{\mu_{hk}} L^\infty(\Gamma_h^k) \| \| \epsilon^\ell_u - \epsilon_u \|_{L^2(\Gamma)} + \| \epsilon^\ell_u \|_{L^2(\Gamma)}.
\]
Recalling that $\int_{\Gamma} u \, d\sigma = 0 = \int_{\Gamma_h} u_h \, d\sigma_h$ and employing (4.5b) yields for any $\lambda \in \nabla \Gamma_h \cdot V_h$

$$
\int_{\Gamma} \epsilon_u^\ell \, d\sigma = \int_{\Gamma} (1 - \frac{1}{\mu_{hh}}) \epsilon_u^\ell \, d\sigma_h + \int_{\Gamma_h} (\Pi u^\ell - u^\ell) \, d\sigma_h \\
+ \int_{\Gamma_h} u^\ell \, d\sigma_h - \int_{\Gamma} u \, d\sigma \\
= \int_{\Gamma} (1 - \frac{1}{\mu_{hh}}) \epsilon_u^\ell \, d\sigma_h + \int_{\Gamma_h} (\Pi u^\ell - u^\ell)(1 - \lambda) \, d\sigma_h \\
+ \int_{\Gamma} \frac{1}{\mu_{hh}} - 1 u \, d\sigma \\
\leq \|1 - \frac{1}{\mu_{hh}} \|_{L_\infty(V_h)} \|\Gamma\|^{1/2} (\|\epsilon_u^\ell\|_{L_2(\Gamma)} + \|u\|_{L_2(\Gamma)}) \\
+ \|u^\ell - \Pi u^\ell\|_{L_2(V_h)} \inf_{q \in V_h} \|1 - \lambda\|_{L_2(V_h)}.
$$

(4.38)

Combining the preceding two inequalities while canceling factors of $\Gamma$ and observing that $\|\epsilon_u^\ell\|_{L_2(\Gamma)} \leq \|\mu_{hh}\|_{L_\infty(V_h)} \|\epsilon_u\|_{L_2(V_h)}$ yields

$$
\|\epsilon_u\|_{L_2(V_h)} \leq \|\frac{1}{\mu_{hh}}\|_{L_\infty(V_h)} \|\mu_{hh}\|_{L_\infty(V_h)} \|1 - \frac{1}{\mu_{hh}}\|_{L_\infty(V_h)} \|\epsilon_u\|_{L_2(V_h)} \\
+ C (\|\epsilon_u^\ell - \epsilon_u\|_{L_2(\Gamma)} + \|\Pi u\|_{L_2(\Gamma)}) \\
+ \|u^\ell - \Pi u^\ell\|_{L_2(V_h)} \inf_{q \in V_h} \|1 - \lambda\|_{L_2(V_h)}.
$$

(4.39)

Observe that $\|\epsilon_u^\ell - \epsilon_u\|_{L_2(\Gamma)} = \sup_{\theta \in L_2(\Gamma)} (\epsilon_u^\ell - \epsilon_u, \Theta)_{\Gamma} = \sup_{\theta \in L_2(\Gamma)} (\epsilon_u^\ell, \Theta)_{\Gamma}$. Combining this observation with Assumption G.1, (4.29), (4.36), and (4.39), applying the Poincaré inequality $\|u\|_{L_2(\Gamma)} \lesssim \|q\|_{L_2(\Gamma)}$, and reabsorbing the term $\|\epsilon_u\|_{L_2(V_h)}$ completes the proof of (4.30).

4.5. Summary of error estimates. We finally collate our energy and duality estimates into error bounds.

**Theorem 4.9.** Assume that $\Gamma$ admits the $H^2$ regularity estimate (2.1), and that Assumptions A, B.1, and G.1 hold. Then

$$
\|\epsilon_u\|_{L_2(V_h)} + \|\epsilon_u\|_{L_2(V_h)} \lesssim \|\hat{q} - \Pi \hat{q}\|_{L_2(V_h)} + \|u^\ell - \Pi u^\ell\| \\
+ h^{k+1} \|q^\ell\|_{L_2(V_h)} + \|\mu_h f^\ell - f_h\|_{L_2(V_h)}.
$$

(4.40)

If in addition Assumption G.2 holds

$$
\|\epsilon_u\|_{L_2(V_h)} + \|\epsilon_u\|_{L_2(V_h)} \lesssim \|\hat{q} - \Pi \hat{q}\|_{L_2(V_h)} + \|u^\ell - \Pi u^\ell\| \\
+ h^{k+1} \|q^\ell\|_{L_2(V_h)} + \|\mu_h f^\ell - f_h\|_{L_2(V_h)}.
$$

(4.41)

If in addition Assumption B.2 holds, then we obtain the superconvergence estimate

$$
\|\epsilon_u\|_{L_2(V_h)} \lesssim h(\|\hat{q} - \Pi \hat{q}\|_{L_2(V_h)} + \|u^\ell - \Pi u^\ell\|_{L_2(V_h)}) \\
+ h^{k+1} \|q^\ell\|_{L_2(V_h)} + \|\mu_h f^\ell - f_h\|_{L_2(V_h)}.
$$

(4.42)
Proof. Assumption B.1 implies that $H(\Theta) \lesssim 1$. We thus insert (4.22) into (4.30) and perform elementary manipulations in order to bound $\|\epsilon_u\|_{L^2(\Gamma_k^h)}$ by the right hand side of (4.40), and finally insert the result into (4.22) in order to similarly bound $\|\epsilon_q\|_{L^2(\Gamma_k^h)}$. Similar manipulations yield (4.41) when Assumption G.2 holds. If additionally Assumption B.2 holds, then we have instead that $H(\Theta) \lesssim h$, which by a similar argument yields (4.42). □

In the following corollary we summarize the convergence rates obtained from Theorem 4.9 when elements in Table 1 are used. We omit the proof, which requires only checking of assumptions and tedious but elementary calculation of constants arising in reference transformations. Error estimates for superconvergent postprocessing schemes may also be obtained from (4.42); we do not give proofs but instead discuss this in the context of numerical examples in Section 6 below.

Corollary 4.10. Assume that $\Gamma \in C^{\max(k+1, r+3)}$, that $f \in H^r(\Gamma)$, that (2.2) and (2.3) hold, and that $|a_k\|_{W_j^0(\tau_k^h)} \lesssim |a|_{W_j^0(\tau_k^h)}$, $0 \leq j \leq r+2$. Assume that any of the elements in Table 1 are used with $r \geq 1$ in the case of BDM elements and $r \geq 0$ otherwise. Assume in addition that either $f_h = f^\ell - |\Gamma_k^h|^{-1} \int_{\Gamma_k^h} f^\ell \, d\sigma_k$ or $f_h = \mu_{kh} f^\ell$. Then

$$\|\epsilon_q\|_{L^2(\Gamma_k^h)} + \|\epsilon_u\|_{L^2(\Gamma_k^h)} \lesssim h^{r+1}\|f\|_{H^r(\Gamma)} + h^{k+1}\|f\|_{L^2(\Gamma)}.$$  

(4.43)

If any of the elements in Table 1 with the exception of the BDM elements are used with $r \geq 1$, then

$$\|\epsilon_u\| \lesssim h^{r+2}\|f\|_{H^r(\Gamma)} + h^{k+1}\|f\|_{L^2(\Gamma)}.$$  

(4.44)

5. Parametric methods and notes on efficient implementation

(3.5a)–(3.5d) describe the construction of surface finite element methods using finite element spaces defined on polynomial approximations to $\Gamma$. The error estimates summarized in Theorem 4.9 indicate that isoparametric HDG approximations in which the degree $r$ of the finite element spaces match the degree $k$ of the surface approximation lead to optimal-order methods with balanced geometric and Galerkin errors. If only low-order approximation is desired, this approach is entirely satisfactory, since in this case a computationally efficient pairing of a piecewise (bi)linear surface approximation with piecewise (bi)linear finite element spaces may be made. When a higher-order approximation is desired, it might be preferable to use parametric finite element spaces defined on $\Gamma$ (instead of $\Gamma_k^h$) by transformations from the linear surface approximation $\Gamma_h$. Standard continuous Galerkin methods of this type were defined and analyzed in [23], and similar constructions have been described for Raviart-Thomas spaces in the boundary element literature [8, 31].

Recalling the definitions (3.2a)–(3.2c) we define

$$V_{h,\Gamma} = \{ \mathcal{P}_a \tilde{v} : \tilde{v} \in \tilde{V}_h \}, \quad W_{h,\Gamma} = \{ \tilde{w}^\ell : \tilde{w} \in \tilde{W}_h \},$$  

$$W_{h,\Gamma}^0 = W_{h,\Gamma} \cap L^0_2(\Gamma), \quad M_{h,\Gamma} = \{ \tilde{\eta}^\ell : \tilde{\eta} \in \tilde{M}_h \}.$$  

(5.1)
A parametric surface HDG method is then given by finding \((q_h, u_h, \hat{u}_h) \in V_{h,\Gamma} \times W_{h,\Gamma}^0 \times M_{h,\Gamma}\) such that for all \((v, w, \eta) \in V_{h,\Gamma} \times W_{h,\Gamma} \times M_{h,\Gamma}\)

\begin{align}
(5.2) & \quad (q_h, v)_{a(\Gamma_h)} - (u_h, \nabla_{\Gamma} \cdot v)_{a(\Gamma_h)} + \langle \hat{u}_h, v \cdot n \rangle_{\partial a(\Gamma_h)} = 0, \\
(5.3) & \quad (q_h, \nabla_{\Gamma}^+ w)_{a(\Gamma_h)} + (q_h \cdot n, w)_{\partial a(\Gamma_h)} + \langle \tau(u_h - \hat{u}_h), w \rangle_{\partial a(\Gamma_h)} = \langle f, w \rangle_{a(\Gamma_h)}, \\
(5.4) & \quad - (q_h \cdot n, \eta)_{\partial a(\Gamma_h)} - \langle \tau(u_h - \hat{u}_h), \eta \rangle_{\partial a(\Gamma_h)} = 0, \\
(5.5) & \quad (u_h, 1)_{a(\Gamma_h)} = 0.
\end{align}

If all system integrals are computed exactly then error behavior is as in the Euclidean case, i.e., the results of Theorem 4.9 hold but with none of the geometric error terms depending on \(k\) or the error in representing \(f\) present. In practice there will necessarily be quadrature errors, however. In particular, computation of the boxed expressions in (5.2)–(5.5) requires the use of formulas (2.5a) and (2.5b) along with inexact quadrature. Most of the remaining expressions may be computed exactly by Piola transforming to reference elements and using appropriate quadrature. The only exceptions are the two expressions involving \(\tau\), which may be computed exactly on the linear skeleton \(E_h\) assuming that the stabilization function \(\tilde{\tau}\) on \(E_h\) is piecewise polynomial and \(\tau = \tilde{\tau}/\mu_{h1}\). The same applies to the method (3.5) on polynomial surface approximations to \(\Gamma\), that is, in both cases geometric information about \(\Gamma\) is needed only to compute the three terms in (3.5) corresponding to the boxed terms above. Such metric information is also needed when computing essentially the same three terms as above—the stiffness matrix, load vector, and mean-value-zero condition—when implementing higher-order standard (CG) finite element methods. In this regard surface HDG methods may be implemented as efficiently as surface CG methods, even though they involve many more terms.

Using either higher-order polynomial surface approximations or employing parametric finite element spaces as described here may add significant expense when assembling system matrices. This is due in part to the necessity of repeatedly computing the distance function \(d\), the projection \(a\), and associated geometric information. The latter issue can become particularly important when \(a\) and \(d\) must be approximated numerically, which is the case in all but a small (though relatively important) set of examples which essentially includes only spheres and tori; cf. [24] for a more detailed discussion of this issue. The amount of geometric information about \(\Gamma\) available also varies dramatically. In some applications the exact surface is not known at all, for example because it is a free boundary which is itself approximated in the computation [27, 30] or because information about the surface is obtained from a scan of a physical object [39]. In these cases only approximate surface representations are possible. In other applications \(\Gamma\) is exactly represented as the level set of some smooth function. Here polynomial approximations to \(\Gamma\) have the advantage that computation of the exact projection \(a\) is isolated to computing \(\Gamma_h^k\); all other geometric information is then obtained from \(\Gamma_h^k\). Analysis of geometric errors is also transparent in this situation. Parametric methods defined on \(\Gamma\), on the other hand, are convenient because quadrature errors can typically be reduced simply by using higher-order quadrature formulas. Construction and analysis of \(p-\) or \(hp\)-methods on surfaces may thus prove more convenient using exact lifting of spaces as in (5.1).
6. Numerical examples

In this section we describe numerical tests that both illustrate and complement our theoretical results. Topics we consider include the interaction between Galerkin and geometric errors, the effects of domain regularity on convergence rates, and superconvergent postprocessed approximations to \( u \). Our tests also go beyond our theoretical results in that we consider Dirichlet boundary conditions and problems where the \( H^2 \) regularity result (4.26) does not hold.

We generated a simple testbed of sample problems as follows. Let \( S^2 \) be the unit sphere with angular spherical coordinates \((\phi, \theta)\), where \( 0 \leq \phi < 2\pi \) is the azimuthal angle in the \( xy \)-plane and \( 0 \leq \theta = \cos^{-1} z \leq \pi \) is the polar angle from the \( z \)-axis. We fix an angle \( \omega \) and let \( \Gamma \) consist of points in \( S^2 \) such that \( 0 \leq \phi \leq \omega \); cf. [24] for an illustration of the domain when \( \omega = \frac{5\pi}{4} \). Setting \( \lambda = \frac{\omega}{2} \), we let \( u(\phi, \theta) = (\sin \theta)^{\lambda} \sin (\lambda \phi) \) be our test solution and may compute that \( f = \lambda (\lambda + 1) u \).

We also have \( u = 0 \) on \( \partial \Gamma \). \( u \) is then singular at the poles for most choices of \( \omega \), the main exception coming when \( \omega = \pi \) so that \( \Gamma \) is a half-sphere, \( \lambda = 1 \), and \( u \) is infinitely differentiable. Otherwise we may expect \( u \) (and any solution of \(-\Delta v = g\) with \( g \) sufficiently smooth) to satisfy \( u \in H^{1+\lambda-\epsilon}(\Gamma) \) for any \( \epsilon > 0 \). Thus \( u \) will have the same regularity as the solution to a Dirichlet boundary value problem on a polygon having vertex opening angle \( \omega \).

Computation of geometric quantities is straightforward: \( d(x) = |x| - 1, \nu(x) = \frac{x}{|x|} \), \( P(x) = I - \nu \otimes \nu \), \( H(x) = \frac{1}{2} P(x) \), and \( \kappa_1(x) = \kappa_2(x) = \frac{1}{|x|} \). Many previous works on approximation of elliptic problems on surfaces have considered more complex surfaces for which a distance function is not explicitly available [26, 24, 23, 22], but the issues associated with doing so are much the same here as for other methods.

We also present results for surfaces with boundary instead of closed surfaces as in our analysis because computations that we carried out on the closed sphere yielded no additional insight into the behavior of the method. Finally, our experiments were carried out using a MATLAB code built on top of the iFEM package [12].

6.1. Test 1: Low-order approximations on linear surface approximations.
In this section we display results for the above test problem with \( \omega = \pi \) (so that \( u \) is smooth) and \( \omega = \frac{31\pi}{16} \) (so that we expect only \( u \in H^{1+\lambda-\epsilon}(\Gamma) \) with \( \lambda = \frac{16}{3} \approx 0.516 \)). In both cases we take \( f_h(x) = f(a(x)) \) in our finite element method, and a piecewise linear surface approximation is used so that \( k = 1 \). Error decrease is shown for the case \( \omega = \pi \) in the left plot of Figure 2 and for the case \( \omega = \frac{31\pi}{16} \) in the right plot. For the case \( \omega = \pi \), we show results for polynomial degree \( r = 0, 1, 2 \). Theorem 4.9 predicts that both \( \|u - u_h\|_{L_2(P_3^h)} \) and \( \|q - q_h\|_{L_2(P_1^h)} \) will behave as \( O(h^{r+1} + h^{k+1}) \) in this case (it is easy to compute that here \( \|\mu_{hk} f^\ell - f_h\|_{L_2(V_h)} \lesssim h^2 \)), which translates to optimal-order convergence for \( r = 0 \) and \( r = 1 \) but suboptimal order \( O(h^2) \) convergence when \( r = 2 \). This is precisely what is observed in Figure 2.

When computing on a spherical subdomain with sufficiently nonsmooth boundary it can be expected that lack of domain regularity rather than the accuracy of the surface approximation will constrain the order of convergence. When \( \omega = \frac{31\pi}{16} \), both the Laplace-Beltrami problem and its dual have solutions lying only in \( H^{1+\lambda-\epsilon}(\Gamma) \). Although we do not assume sufficient regularity to directly apply Theorem 4.9, we may still predict from it that \( \|q - q_h\|_{L_2(P_1^h)} \lesssim h^\lambda \) and \( \|u^\ell - u_h\|_{L_2(\Gamma)} \lesssim h^{2\lambda} \). These
asymptotic rates of convergences are those observed in continuous Galerkin methods on similar polygonal domains with reentrant corners, and we indeed observe these rates for surface HDG methods in Figure 2.

Figure 2. Error decrease with respect to $h$. Left: $\omega = \pi$. Right: $\omega = 31\pi/16$.

6.2. Test 2: Parametric approximation. Next we describe tests using the exact parametric method described in §5. Letting $r = 2$, we see in Figure 3 that optimal order $O(h^3)$ is achieved. As an intermediate case, we also let $f_h = f^\ell \mu_h$ with surface approximation of degree $k = 1$, which eliminates the error term in (4.40) and (4.42) due to approximation of $f$ but not the geometric error term $h^{k+1}\|q^\ell\|_{L_2(\Gamma_h^k)}$ due to the approximation of $\Gamma$ by $\Gamma_h^k$. As shown in Figure 3, in this case the order of convergence is $O(h^2)$ as expected from theory, but there is significant reduction of five to ten times in the error as compared with the case $f_h = f^\ell$ (also shown for comparison). Thus significant error reduction, though not improved convergence order, may be possible by only improving approximation of $f$.

6.3. Test 3: Superconvergent postprocessing. One advantage of HDG and mixed methods on Euclidean domains is the availability of simple elementwise postprocessing schemes which yield superconvergent ($O(h^{r+2})$) approximations $u_h^\ell$ to $u$. Such schemes are also of potential use in the context of surface FEM, although the presence of geometric errors lessens their utility. A number of such schemes exist in the literature; cf. [18, 17, 41] among many others. In [19], a postprocessing scheme for isoparametric Euclidean elements is defined and proved to be superconvergent. There the postprocessing scheme is defined on a reference element and the result transformed to the physical elements to obtain the final postprocessed approximation $u_h^*$. It is possible to similarly analyze this scheme for surface HDG methods, with the additional complication of geometric errors. We do not give the details, but if (4.42) holds along with all assumptions A, B, and G, then we obtain

\[
\|u^\ell - u_h^*\|_{L_2(\Gamma_h^k)} \lesssim h^{r+2} + h^{k+1}.
\]
This estimate indicates that any geometric error present in the original approximation is still present in the postprocessing error, that is, obtaining a better approximation to \( u \) via postprocessing depends on the geometric error in the original approximation being of sufficiently high order. This is true even if geometric errors are entirely eliminated in the definition of the postprocessing scheme itself, as occurs for example when an isoparametric method is used in the definition of \( u_h \) and \( q_h \) but a parametric method over \( \Gamma \) is used in the definition of \( u_h^* \).

For our numerical computations we use a different postprocessing scheme in order to illustrate that a general pattern holds for the analysis of such methods. Assume that \((q_h, u_h)\) solves either (3.5a)–(3.5d) or the parametric method introduced in §5. Let also \( K \in \mathcal{T}^k_h \) be an element in the computational mesh on which \((q_h, u_h)\) is defined. Let \( W_{h,K} \) be a finite-dimensional space on the curved triangle \( a(K) \subset \Gamma \); we may think of \( W_{h,K} \) as the lift of \( P_{r+1} \) to \( \Gamma \) in the case \( k = 1 \) or when the parametric method is used. Also, let \( W^0_{h,K} \perp 1 \) on \( a(K) \) be the subspace of \( W_{h,K} \) having 0 mean. We then seek \( u_h^* \in W_{h,K} \) such that

\[
\begin{align*}
\int_{a(K)} \nabla \Gamma u_h^* \cdot \nabla \Gamma v_h \, d\sigma &= \int_{a(K)} f v_h \, d\sigma - \int_{\partial a(K)} P_a q_h \cdot n v_h \, d\sigma, \quad v_h \in W^0_{h,K}, \\
\int_{a(K)} u_h^* \, d\sigma &= \int_{a(K)} u_h \, d\sigma.
\end{align*}
\]

The scheme (6.2) is discussed in the context of HDG methods in [17] and analyzed for mixed methods in [41]. The post processing scheme analyzed in [19] essentially replaces the right hand side in the first line of (6.2) with \( \int_{a(K)} -P_a q_h \cdot \nabla \Gamma v_h \, d\sigma \).

The analysis of (6.2) requires trace error estimates which we have not given above, but we still expect (6.1) to hold for \( r \geq 1 \).

In Figure 3 we illustrate that any geometric errors present in the definition of \( u_h \) and \( q_h \) are still present in \( u_h^* \). When a parametric method is used with \( r = 1 \) or \( r = 2 \) so that no geometric error is present, the postprocessed solution \( u_h^* \) defined in (6.2) indeed superconverges with order \( h^{r+2} \), as can be expected from (6.1) with geometric terms eliminated. When \( k = 1 \), the geometric error is of order \( h^2 \), and the postprocessed solution does not superconverge with order \( h^3 \) as in the Euclidean and parametric surface cases. In order to achieve a superconvergent postprocessed solution it would be necessary to take \( k = r + 1 \) in the original approximation.

We conclude this subsection by noting that a number of other definitions of postprocessed solutions are possible (cf. [18, 17, 41] among others). All can be expected to follow the same pattern: Superconvergence as in the Euclidean case may be expected only if the geometric error \( O(h^{k+1}) \) is of sufficiently high order.

7. Concluding remarks

Our results can be extended to the case of more involved stabilization functions \( \tau \), and to the use of non-conforming meshes [13]. The extension of the a posteriori error estimates for HDG methods on flat surfaces obtained in [20, 21] constitutes the subject of ongoing research.

References

\[ h^2 \|q - q_h\|_{L^2(f)} = h^3 \|u - u_h^r\|_{L^2(r)} \]

**Figure 3.** Left: Error decrease with respect to \( h \) for parametric approximations with \( r = 2 \) and for approximations with \( r = 2 \) and \( k = 1 \). Right: Decrease of the postprocessing error \( \|u - u_h^r\|_{L^2(r)} \) with respect to \( h \) for parametric approximations with \( r = 1 \) and \( r = 2 \), and for an approximate surface HDG method with \( r = 1 \), \( k = 1 \).


School of Mathematics, University of Minnesota, Minnesota, USA, and Department of Mathematics and Statistics, King Fahd University of Petroleum and Minerals, Dahr-an, Saudi Arabia

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

Department of Mathematics, Texas A&M University, College Station, TX 77843–3368

E-mail address: demlow@math.tamu.edu