# A POSTERIORI ERROR ESTIMATES FOR THE LAPLACE-BELTRAMI OPERATOR ON PARAMETRIC $C^2$ SURFACES

ANDREA BONITO* AND ALAN DEMLOW†

**Key words.** Laplace-Beltrami operator, surface finite element methods, a posteriori error estimates, adaptive finite element methods

**AMS subject classification.** 58J32, 65N15, 65N30

**Abstract.** We prove new a posteriori error estimates for surface finite element methods (SFEM). Surface FEM approximate solutions to PDE posed on surfaces. Prototypical examples are elliptic PDE involving the Laplace-Beltrami operator. Typically the surface is approximated by a polyhedral or higher-order polynomial approximation. The resulting FEM exhibits both a geometric consistency error due to the surface approximation and a standard Galerkin error. A posteriori estimates for SFEM require practical access to geometric information about the surface in order to computably bound the geometric error. It is thus advantageous to allow for maximum flexibility in representing surfaces in practical codes when proving a posteriori error estimates for SFEM. However, previous a posteriori estimates using general parametric surface representations are suboptimal by one order on $C^2$ surfaces. Proofs of error estimates optimally reflecting the geometric error instead employ the closest point projection, which is defined using the signed distance function. Because the closest point projection is often unavailable or inconvenient to use computationally, a posteriori estimates using the signed distance function have notable practical limitations. We merge these two perspectives by assuming *practical* access only to a general parametric representation of the surface, but using the distance function as a *theoretical* tool. This allows us to derive sharper geometric estimators which exhibit improved experimentally observed decay rates when implemented in adaptive surface finite element algorithms.

**1. Introduction.** The Laplace-Beltrami operator (or surface Laplacian) has received a great deal of attention recently in part due to its ubiquity in geometric PDEs and in particular in applications involving surfaces that evolve and are the domain of an underlying PDE. Typical examples are mean curvature flow and surface diffusion appearing in materials science modeling [34] or Willmore flow as a prototype for equilibrium shapes of membranes governed by bending energy [33].

In this paper we consider finite element approximation of solutions to the Laplace-Beltrami problem

$$-\Delta_\gamma u = f \text{ on } \gamma. \tag{1.1}$$

Here $\gamma \subset \mathbb{R}^{n+1}$ is an orientable, $C^2$ hypersurface, and $\Delta_\gamma$ is the Laplace-Beltrami operator on $\gamma$. We mainly focus on the cases where $\gamma$ is closed so that the compatibility condition $\int_\gamma f = 0$ must be assumed to guarantee existence of a solution, and we additionally impose $\int_\gamma u = 0$ in order to fix a unique solution.

Dziuk defined a canonical piecewise linear finite element method for approximating solutions to (1.1) in [27]. This method proceeds by approximating $\gamma$ by a polyhedral surface $\Gamma$ having triangular faces which serve as the finite element mesh. Finite element shape functions are then defined on this mesh and used to approximately solve (1.1). The corresponding stiffness matrix matches the cotangent formula [36]

for the approximation of the surface Laplacian on a polyhedral surface. This procedure was extended to higher-degree finite element spaces and surface approximations in [24]. A variational crime is committed in this method due to the approximation of $\gamma$ by $\Gamma$, and the resulting consistency error is often called a "geometric error". If $\Gamma$ is a degree $k$ surface approximation on a quasi-uniform mesh of width $h$ and a degree $r$ finite element space is used to construct the finite element approximation $U$, a priori error analysis yields

$$\|u - U\|_{H^1(\gamma)} \leq O(h^r) + O(h^{k+1}) \tag{1.2}$$

under the assumption that $\gamma$ and $u$ are sufficiently smooth.

The assumption that $\gamma$ is $C^2$ is fundamental in this error analysis. When $\gamma$ is $C^2$ it may be represented *implicitly* as the 0 level set of a signed distance function $d$, and there is also a uniquely defined closest-point projection $\mathbf{P}_d$ mapping a tubular neighborhood of $\gamma$ onto $\gamma$. The properties of $\mathbf{P}_d$ are used integrally in proving the error estimate (1.2). If on the other hand one were to represent $\gamma$ *parametrically* via some *arbitrary* smooth map $\mathbf{P} : \Gamma \to \gamma$, then the corresponding natural finite element error analysis would only yield

$$\|u - U\|_{H^1(\gamma)} \leq O(h^r) + O(h^k). \tag{1.3}$$

The properties of the closest point projection thus lead in effect to a "geometric superconvergence" result in which the consistency error is of higher order than one might expect based on generic considerations. We emphasize that obtaining the superior a priori convergence rate in the geometric error seen in (1.2) does not generally required *practical computational access* to the closest point projection. Rather, theoretical use in proofs generally suffices.

A posteriori error estimates were proved for the case $r = k = 1$ in [25] under the assumption that $\gamma$ is $C^2$. These estimates have the form

$$\|u - U\|_{H^1(\gamma)} \leq \eta + \mathcal{G}_{\mathbf{P}_d}, \tag{1.4}$$

where $\eta$ is a standard residual-type error estimator for controlling energy errors, and $\mathcal{G}_{\mathbf{P}_d}$ controls the geometric consistency error a posteriori. Notably, $\mathcal{G}_{\mathbf{P}_d}$ heuristically retains the "superconvergent" a priori order $h^{k+1} = h^2$ that is seen in (1.2). There are two significant drawbacks to the approach to a posteriori error estimation for (1.1) taken in [25]. First, the assumption that $\gamma$ is $C^2$ may not hold in practice. Secondly, in contrast to the case of a priori error analysis, the a posteriori estimates of [25] assume *practical computational access* to the closest point projection $\mathbf{P}_d$. This assumption may be unrealistic. Closed-form analytical expressions for $\mathbf{P}_d$ exist only in the very restricted event that $\gamma$ is a sphere or a torus. If $\gamma$ is computationally represented as the zero level set of *some* sufficiently smooth function, then it is possible to approximate $\mathbf{P}_d$ by for example using a Newton-type algorithm [25, 30]. Practical experience however indicates that this procedure can add significant expense to the code. Finally, it may be that $\gamma$ is given as a parametric representation. A posteriori error estimates in which the surface representation follows the framework of [25] have also been proved for discontinuous Galerkin [23] and cut [26] surface finite element methods along with $L_2$ and $L_\infty$ estimates for Dziuk's method [17].

An alternate approach to representing $\gamma$ in the context of a posteriori error estimation and adaptivity is given in [12, 13]. In these works $\gamma$ is only assumed to be globally

Lipschitz and elementwise $C^{1,\alpha}$. In addition, $\gamma$ is computationally represented via a parametrization $\mathbf{P} : \Gamma \to \gamma$. The estimators then have the form

$$\|u - U\|_{H^1(\gamma)} \leq \eta + \mathcal{G}_{\mathbf{P}}, \tag{1.5}$$

where $\eta$ is a standard residual error estimator as above, and $\mathcal{G}_{\mathbf{P}}$ bounds the geometric consistency error by using information from the parametrization $\mathbf{P}$. This framework avoids the two main flaws of the approach of [25]: It allows for surfaces less regular than $C^2$, and allows for a much more flexible surface representation that may take the form of an implicit representation if it is available, but does not require computational access to the distance map $\mathbf{P}_d$. The price that is paid for these advantages is that the geometric consistency estimator $\mathcal{G}_{\mathbf{P}}$ heuristically only retains the reduced a priori convergence order $O(h^k)$ seen in (1.3), even if $\gamma$ is $C^2$. In the latter case, adaptive algorithms based on $\mathcal{G}_{\mathbf{P}}$ generally resolve the geometry much more than is necessary to reach a given error tolerance. Quasi-optimal error decay for adaptive finite element approximations of $u$ lying in certain regularity classes is derived in [12, 13]. However, these regularity classes are artificially restricted by the surface approximation when $\gamma$ is $C^2$. We demonstrate computationally below that over-resolution of the geometry considerably affects the efficiency of the adaptive algorithm.

Our goal in this paper is to produce a posteriori error estimates for finite element approximations to (1.1) which combine the major advantages of the parametric and implicit approaches to representing $C^2$ surfaces $\gamma$. More precisely, we assume that our code has *practical* access only to *some* reasonable parametric representation $\mathbf{P}$ of $\gamma$, as in [12, 13, 31]. On the other hand, we know in this case that the closest point projection $\mathbf{P}_d$ exists. We make *theoretical* use of its properties to produce computable a posteriori error estimators that require information only from $\mathbf{P}$, but which heuristically retain the "superconvergent" geometric order $h^{k+1}$ seen in (1.2) and $\mathcal{G}_{\mathbf{P}_d}$. Our proofs that these estimators are reliable and efficient require a number of sometimes technical steps, but underlying them is the simple observation that the closest point parametrization $\mathbf{P}_d : \Gamma \to \gamma$ is optimal in $L_\infty$. That is, for any other parametrization $\mathbf{P} : \Gamma \to \gamma$,

$$|\mathbf{x} - \mathbf{P}_d(\mathbf{x})| \leq |\mathbf{x} - \mathbf{P}(\mathbf{x})|, \quad \mathbf{x} \in \Gamma. \tag{1.6}$$

Note that we do not consider surfaces with less than $C^2$ regularity as many critical properties of the closest point projection do not hold beneath that threshold.

Finally, we point out that the type of parametric finite element methods considered here are used to approximate time dependent problems such as the mean curvature flow [28], capillary surfaces [3], surface diffusion [5, 6], Willmore flow [6, 7, 9, 14, 29, 32], fluid biomembranes [8, 15], and fluid membranes with orientational order [10, 11]. The analysis of these methods is largely open, but we refer to [4, 7, 8, 19, 20, 21, 22] as well as the survey [22] for some of the early work including level set and phase field approaches.

The paper is outlined as follows. In Section 2 we define approximations of surfaces and lay out assumptions which must be placed on the resolution of $\gamma$ by its discrete approximations in order for our a posteriori estimates to hold. In Section 3 we prove our a posteriori error estimates. Section 4 contains numerical tests illustrating the advantages of our estimates. Finally Section 5 contains some concluding remarks and discussion of possible future research directions.

## 2. Preliminaries.

**2.1. Representation of Parametric Surfaces.** We assume that the surface $\gamma$ is described as the deformation of a $d$ dimensional polyhedral surface $\overline{\Gamma}$ by a globally bi-Lipschitz *homeomorphism* $\mathbf{P} : \overline{\Gamma} \to \gamma \subset \mathbb{R}^{n+1}$. The overline notation is to emphasize that $\overline{\Gamma}$ is piecewise affine. Thus there is $\overline{L} > 0$ such that for all $\overline{\mathbf{x}}, \overline{\mathbf{y}} \in \overline{\Gamma}$

$$\overline{L}^{-1}|\overline{\mathbf{x}} - \overline{\mathbf{y}}| \leq |\tilde{\mathbf{x}} - \tilde{\mathbf{y}}| \leq \overline{L}|\overline{\mathbf{x}} - \overline{\mathbf{y}}|, \qquad \text{where } \tilde{\mathbf{x}} = \mathbf{P}(\overline{\mathbf{x}}), \ \tilde{\mathbf{y}} = \mathbf{P}(\overline{\mathbf{y}}). \qquad (2.1)$$

The (closed) facets of $\overline{\Gamma}$ are denoted $\overline{T}$, and form the collection $\overline{\mathcal{T}} = \{\overline{T}\}$. We assume that these facets are all simplices. Extension to other element shapes such as quadrilaterals and to nonconforming discretizations is possible under reasonable assumptions with minor modifications. We let $\mathbf{P}_T : \overline{T} \to \mathbb{R}^{n+1}$ be the restriction of $\mathbf{P}$ to $\overline{T}$. This partition of $\overline{\Gamma}$ induces the partition $\widetilde{\mathcal{T}} = \{\widetilde{T}\}_{\overline{T} \in \overline{\mathcal{T}}}$ of $\gamma$ upon setting

$$\widetilde{T} := \mathbf{P}_T(\overline{T}), \quad \overline{T} \in \overline{\mathcal{T}}.$$

Note that this *non-overlapping* parametrization allows for not necessarily globally $C^2$ parameterizations of $\gamma$. We additionally define *macro patches*

$$\overline{\omega}_T = \cup_{\overline{T}', \overline{T}' \cap \overline{T} \neq \emptyset} \overline{T}'$$

and

$$\widetilde{\omega}_T = \mathbf{P}(\overline{\omega}_T).$$

Finally, we let $h_T = |\overline{T}|^{\frac{1}{n}}$.

Let $\widehat{T}$ be the unit reference simplex, which we sometimes refer to as the *universal parametric domain*. We denote by $\overline{\mathbf{X}}_T : \mathbb{R}^n \to \mathbb{R}^{n+1}$ the affine map such that $\overline{T} = \overline{\mathbf{X}}_T(\widehat{T})$ and we let $\boldsymbol{\chi}_T := \mathbf{P} \circ \overline{\mathbf{X}}_T : \widehat{T} \to \widetilde{T}$ be the corresponding local parametrization of $\widetilde{T}$. We extend this property by assuming that there are patches $\widehat{\omega}_T, \overline{T} \in \overline{\mathcal{T}}$, consisting of the universal parametric domain $\widehat{T}$ and other shape-regular simplices of unit size such that $\overline{\mathbf{X}}_T$ may be extended as a continuous, piecewise-affine bijection such that $\overline{\omega}_T = \overline{\mathbf{X}}_T(\widehat{\omega}_T)$. Because $\gamma$ is closed, the domains $\widehat{\omega}_T$ may be constructed so that they are convex, and we assume that this is the case. The patchwise parametric maps may be easily constructed by sewing together their elementwise counterparts except in a few pathological cases involving very course meshes such as a 6-triangle triangulation of the sphere consisting of two "stacked" tetrahedra. In that case, the neighbors of any element consist of the entire set of elements which cannot be "flattened out" (plane and sphere are not homotopic).

We follow [16] and define the shape regularity constant of the subdivision $\overline{\mathcal{T}} := \{\overline{T}\}_{i=0}^M$ as the smallest constant $\overline{c}$ such that

$$\overline{c}^{-1}\mathrm{diam}(\overline{T})|\mathbf{w}| \leq |D\overline{\mathbf{X}}_T(\hat{x})\mathbf{w}| \leq \overline{c}\,\mathrm{diam}(\overline{T})|\mathbf{w}|, \qquad \mathbf{w} \in \mathbb{R}^n, \qquad \overline{T} \in \overline{\mathcal{T}}, \qquad (2.2)$$

and assume that $\overline{c} < \infty$. In the following, we omit to mention the dependency on $\overline{c}$ of the constants appearing in our argumentation. We additionally assume that the number of elements in each patch $\widehat{\omega}_T$ is uniformly bounded. This assumption automatically follows from shape regularity for triangulations of Euclidean domains, but the situation is more subtle for surface triangulations as illustrated in Figure 2.1. Such a bound does for example hold if $\overline{\Gamma}$ is derived by systematic refinement of an initial surface mesh with a uniform bound on the number of elements in a patch [25], or more generally using adaptive refinement strategies [12, 13]. In addition, this implies that

all elements in $\overline{\omega}_T$ have uniformly equivalent diameters, as for shape regular triangulations on Euclidean domains. Finally, because each vertex has uniformly bounded valence the number of parametric patches $\widehat{\omega}_T$ needed may be less than $\#\mathcal{T}$, that is, $\widehat{\omega}_T = \widehat{\omega}_{T'}$ may hold for $\overline{T} \neq \overline{T}'$. We thus assume that there are a finite collection of such parametric patches independent of the discretization, and that properties of these patches such as constants in extension operators are uniform across the collection.



FIG. 2.1. *Two different configurations when $n = 2$ illustrating that the number of elements sharing the same vertex could be arbitrarily large even when using triangles satisfying* (2.2) *with a uniform constant $\overline{c}$.*

We also extend $\boldsymbol{\chi}_T = \mathbf{P} \circ \overline{\mathbf{X}}_T : \widehat{\omega}_T \to \widetilde{\omega}_T$ to be a local parametrization of $\gamma$ which is bi-Lipschitz with element scaling. Namely, there exists a universal constant $L \geq 1$ such that for each fixed $\overline{T} \in \overline{\mathcal{T}}$ and for all $\tilde{\mathbf{x}}, \tilde{\mathbf{y}} \in \widetilde{\omega}_T$,

$$h_T L^{-1} |\hat{\mathbf{x}} - \hat{\mathbf{y}}| \leq |\tilde{\mathbf{x}} - \tilde{\mathbf{y}}| \leq L h_T |\hat{\mathbf{x}} - \hat{\mathbf{y}}|, \quad \text{where } \tilde{\mathbf{x}} = \boldsymbol{\chi}_T(\hat{\mathbf{x}}), \ \tilde{\mathbf{y}} = \boldsymbol{\chi}_T(\hat{\mathbf{y}}). \qquad (2.3)$$

The collection of these parametrizations is denoted $\boldsymbol{\chi}$, i.e. $\boldsymbol{\chi} = \{\boldsymbol{\chi}_T\}_{\overline{T} \in \overline{\mathcal{T}}}$. We further assume that $\mathbf{P}(\mathbf{v}) = \mathbf{v}$ for all vertices $\mathbf{v}$ of $\overline{\Gamma}$, so that $\overline{\mathbf{X}}_T$ is the nodal interpolant of $\boldsymbol{\chi}_T$ into linears.

Finally, we note that a function $v_T : \widetilde{T} \to \mathbb{R}$ defines uniquely two functions $\hat{v}_T : \widehat{T} \to \mathbb{R}$ and $\overline{v}_T : \overline{T} \to \mathbb{R}$ via the maps $\boldsymbol{\chi}_T$ and $\mathbf{P}$, namely

$$\hat{v}_T(\hat{\mathbf{x}}) := v_T(\boldsymbol{\chi}_T(\hat{\mathbf{x}})) \quad \forall \, \hat{\mathbf{x}} \in \widehat{T} \qquad \text{and} \qquad \overline{v}_T(\bar{\mathbf{x}}) := v_T(\mathbf{P}(\bar{\mathbf{x}})) \quad \forall \, \bar{\mathbf{x}} \in \overline{T}. \qquad (2.4)$$

Conversely, a function $\hat{v}_T : \widehat{T} \to \mathbb{R}$ (respectively, $\overline{v}_T : \overline{T} \to \mathbb{R}$) defines uniquely the two functions $v_T : \widetilde{T} \to \mathbb{R}$ and $\overline{v}_T : \overline{T} \to \mathbb{R}$ (respectively, $v_T : \widetilde{T} \to \mathbb{R}$ and $\hat{v}_T : \widehat{T} \to \mathbb{R}$). When no confusion is possible, we will always denote by $v$ the two lifts $\overline{v}$ or $\hat{v}$ of $v : \widetilde{T} \to \mathbb{R}$ and set $\tilde{\mathbf{x}}_T := \boldsymbol{\chi}_T(\hat{\mathbf{x}})$ for all $\hat{\mathbf{x}} \in \widehat{T}$.

Before proceeding further, we note that as a general rule, we use hat symbols to denote quantities related to $\widehat{T}$, an overline to refer to quantities on $\overline{\Gamma}$, tilde to characterize quantities in $\gamma$ and bold to indicate vector quantities.

**2.2. Interpolation of Parametric Surfaces and Finite Element Spaces.**

**2.2.1. Finite Element Spaces and Surface Approximations.** For $s \geq 1$, let $\mathcal{P}_s$ be the space of polynomials of degree at most $s$ and let $I_s : C^0(\widehat{T}) \to \mathcal{P}_s$ be the corresponding Lagrange interpolation operator. Note that we will use the same notation for the componentwise Lagrange interpolation operator $I_s : C^0(\widehat{T})^d \to (\mathcal{P}_s)^d$, and also for the naturally defined Lagrange interpolant on the faces of $\overline{\Gamma}$.

We fix $k \geq 1$ and define $\mathbf{X}_T := I_k \boldsymbol{\chi}_T$ to be the (component wise) interpolant of degree at most $k$ of $\boldsymbol{\chi}_T$. Then $T := T_k := \mathbf{X}_T(\widehat{T})$ is the piecewise polynomial interpolation of $\widetilde{T}$ with associated subdivision $\mathcal{T} = \{\mathbf{X}_T(\widehat{T})\}_{\overline{T} \in \overline{\mathcal{T}}}$. The global quantities are defined by

$$\mathbf{X} := \left\{ \mathbf{X}_T \right\}_{\overline{T} \in \overline{\mathcal{T}}}, \qquad \Gamma := \cup_{\overline{T} \in \overline{\mathcal{T}}} T$$

and we denote by $S_{\mathcal{T}}$ the set of interior faces of $\mathcal{T}$. Note that may equivalently write $T = (I_k \mathbf{P})(\overline{T})$, $\overline{T} \in \overline{\mathcal{T}}$, where $I_k$ is the Lagrange interpolant naturally defined on the faces of $\gamma$. In order to define a parametric map from $\Gamma$ to $\gamma$, we let

$$\mathbf{P} : \Gamma \to \gamma, \qquad \mathbf{P}(\mathbf{x}) = \mathbf{P} \circ (I_k \mathbf{P})^{-1}(\mathbf{x}), \ \ \mathbf{x} \in \Gamma.$$

Note that use the same notation for $\mathbf{P} : \overline{\Gamma} \to \gamma$ and $\mathbf{P} : \Gamma \to \gamma$. The definitions are consistent, and the domain should be clear from the context. Thus this ambiguity in notation should cause no confusion.

For $T \in \mathcal{T}$, we define its diameter as the diameter of the corresponding element in $\overline{\mathcal{T}}$, i.e. $h_T = |\overline{T}|^{\frac{1}{n}}$ with $\overline{T} = (\overline{\mathbf{X}} \circ \mathbf{X}^{-1})(T)$. As for $\overline{\mathcal{T}}$, we assume that there exists $c > 0$ such that

$$c^{-1} \mathrm{diam}(\overline{T})|\mathbf{w}| \leq |D\mathbf{X}_T(\hat{x})\mathbf{w}| \leq c \, \mathrm{diam}(\overline{T})|\mathbf{w}|, \qquad \mathbf{w} \in \mathbb{R}^n, \tag{A1}$$

i.e. the shape regularity constant $\mathcal{T}$ is strictly positive. This can be guaranteed from (2.2) and assuming that the resolution of $\gamma$ by $\Gamma$ is fine enough [12, 13, 16]. However in this work the focus is on deriving a-posteriori error estimators and we assume (A1) directly. Moreover, we do not explicitly mention the dependency on $c$ of the constants appearing in our argumentation below. Finally, we denote by $\omega_T$ the element patch given by $\omega_T = \mathbf{X}_T(\widehat{\omega}_T)$.

We fix $r \geq 1$ and define the finite element space on the approximate surface $\Gamma$ by

$$\mathbb{V}(\mathcal{T}) := \left\{ V \in C^0(\Gamma) \mid V|_T = \widehat{V} \circ \mathbf{X}^{-1} \text{ for some } \widehat{V} \in \mathcal{P}_r, \ T \in \mathcal{T}, \ \int_\Gamma V = 0 \right\}. \tag{2.5}$$

At this point we emphasize that the integer $k$ denotes the polynomial degree for the approximation of $\gamma$ by $\Gamma$ while $r$ denotes the polynomial degree for the finite element approximation of $u$. In general, they do not need to be equal.

**2.2.2. Geometric Estimators.** We define in this section an estimator for the *geometric error* due to the approximation of $\gamma$ by a piecewise polynomial surface $\Gamma$. For $T \in \mathcal{T}$, we define the *geometric element indicator*

$$\lambda_T := \|\nabla(\mathbf{P} - I_k \mathbf{P})\|_{L_\infty(\overline{T})}, \tag{2.6}$$

and the corresponding *geometric estimator*

$$\lambda := \max_{T \in \mathcal{T}} \lambda_T. \tag{2.7}$$

In [12] an estimator equivalent to $\lambda$ is used in order to control the geometric consistency error. However, $\lambda$ has a heuristic a priori order of $h^k$ for smooth surfaces, and the corresponding a priori error estimate is (1.3). In view of (1.2), the geometric error should instead be $O(h^{k+1})$ for smooth surfaces, and we realize that the geometry approximation is overestimated by $\lambda$ when considering $C^2$ surfaces. Let

$$\beta_T := \|\mathbf{P} - I_k \mathbf{P}\|_{L_\infty(\overline{T})}, \ \ \beta := \max_{T \in \mathcal{T}} \beta_T. \tag{2.8}$$

As we shall see, in this context the appropriate estimator is

$$\mu_T := \beta_T + (\lambda_T)^2, \qquad \mu := \max_{T \in \mathcal{T}} \mu_T. \tag{2.9}$$

When $\gamma$ is sufficiently smooth, $\mu \lesssim h^{k+1} + h^{2k} \lesssim h^{k+1}$, so estimating the geometry using $\mu$ preserves the "superconvergent" geometric consistency error observed in (1.2).

## 2.3. Basic differential geometry.

### 2.3.1. The distance function map.
The structure of the map $\mathbf{P}$ depends on the application. The most popular is when $\widetilde{T}$ is described on $T$ via the *distance function* $d : \mathcal{N} \to \gamma$. $d$ is well defined on a local tubular neighborhood $\mathcal{N}$ given the smoothness assumption on $\gamma$. Let

$$\widetilde{T} \ni \tilde{\mathbf{x}} = \mathbf{x} - d(\mathbf{x})\nabla d(\mathbf{x}) =: \mathbf{P}_d(\mathbf{x}) \qquad \forall\, \mathbf{x} \in \mathcal{N}. \tag{2.10}$$

Notice that we introduced the notation $\mathbf{P}_d$ for the lift given by the distance function, i.e. the closest point in $\gamma$, as it will appear multiple times in our analysis below. We assume from now on that $\overline{\Gamma} \subset \mathcal{N}$. Notice that for $\overline{\mathbf{x}} \in \overline{\Gamma}$, $\mathbf{P}(\overline{\mathbf{x}})$ (defined by a generic $\boldsymbol{\chi}$) and $\mathbf{P}_d(\overline{\mathbf{x}})$ are both on $\gamma$ but are not necessarily the same points. We postpone the discussion of this point until Section 2.5.1, see (1.6).

In this work and as in [12, 13, 16], we do not assume that the distance function is available to the user. However, the $C^2$ regularity assumed on $\gamma$ will allow us to use the distance function as a theoretical tool to improve upon the geometric estimators provided in [12, 13].

We further describe some basic geometric notions. Given $\mathbf{x} \in \mathcal{N}$, $\boldsymbol{\nu}(\mathbf{x}) := \nabla d(\mathbf{P}_d(\mathbf{x})) = \nabla d(\mathbf{x})$ is the normal to $\gamma$ at the closest point $\mathbf{P}_d(\mathbf{x}) \in \gamma$, and $\mathbf{W}(\mathbf{x}) := D^2 d(\mathbf{x})$ is the Weingarten map. Let also $\kappa_i(\mathbf{x})$, $i = 1, .., n$, be the eigenvalues of $\mathbf{W}(\mathbf{x})$. These are the principal curvatures of $\gamma$ if $\mathbf{x} \in \gamma$ and of parallel surfaces if $\mathbf{x} \notin \gamma$. In addition, we have the relationship

$$\kappa_i(\mathbf{x}) = \frac{\kappa_i(\mathbf{P}_d(\mathbf{x}))}{1 + d\kappa_i(\mathbf{P}_d(\mathbf{x}))}. \tag{2.11}$$

For the sake of convenience we also define the maximum principal curvature

$$K(x) = \max_{1 \le i \le n} |\kappa_i(x)|, \quad x \in \mathcal{N}. \tag{2.12}$$

We may now more precisely write

$$\mathcal{N} = \{\mathbf{x} \in \mathbb{R}^{n+1} : \mathrm{dist}(\mathbf{x}, \gamma) < \|K\|_{L_\infty(\gamma)}^{-1}\}. \tag{2.13}$$

### 2.3.2. Differential operators and area elements.
In this subsection we recall basic differential geometry notations and definitions and refer to [13] for details.

Let $\mathbf{G} \in \mathbb{R}^{(n+1) \times n}$ be the matrix

$$\mathbf{G} := \mathbf{G}_\gamma := [\widehat{\partial}_1 \boldsymbol{\chi}, \dots, \widehat{\partial}_d \boldsymbol{\chi}],$$

whose $j$-th column $\widehat{\partial}_j \boldsymbol{\chi} \in \mathbb{R}^{n+1}$ is the vector of partial derivatives of $\boldsymbol{\chi}$ with respect to the $j^{th}$ coordinate of $\widehat{T}$. Note that because $\boldsymbol{\chi}$ is a diffeomorphism, the set $\{\widehat{\partial}_j \boldsymbol{\chi}\}_{j=1}^d$ consists of tangent vectors to $\gamma$ which are linearly independent and forms a basis of the tangent plane of $\gamma$. The *first fundamental form* of $\gamma$ is the symmetric and positive definite matrix $\mathbf{g} \in \mathbb{R}^{n \times n}$ defined by

$$\mathbf{g} = (g_{ij})_{1 \le i,j \le d} := (\widehat{\partial}_i \boldsymbol{\chi}^T \widehat{\partial}_j \boldsymbol{\chi})_{1 \le i,j \le d} = \mathbf{G}^T \mathbf{G}. \tag{2.14}$$

We will also need $\widetilde{\mathbf{D}} = \widetilde{\mathbf{G}}^{-1}$, where

$$\widetilde{\mathbf{G}} := [\mathbf{G}, \boldsymbol{\nu}^T] = [\widehat{\partial}_1 \boldsymbol{\chi}, \dots, \widehat{\partial}_d \boldsymbol{\chi}, \boldsymbol{\nu}^T],$$

and $\mathbf{D} \in \mathbb{R}^{n \times (n+1)}$ resulting from $\widetilde{\mathbf{D}}$ by suppressing its last row. With this notation, we have for $v : \gamma \to \mathbb{R}$

$$\nabla \hat{v} = \nabla_\gamma v \mathbf{G} \qquad \nabla_\gamma v = \nabla \hat{v}\, \mathbf{D}.$$

With the surface area element defined as

$$q := \sqrt{\det \mathbf{g}}, \tag{2.15}$$

the Laplace-Beltrami operator on $\gamma$ is given by

$$\Delta_\gamma v = \frac{1}{q} \mathrm{div}\big(q \nabla \hat{v} \mathbf{g}^{-1}\big). \tag{2.16}$$

We could alternately compute the elementary area on $\gamma$ via the map $\mathbf{P}_d \circ \mathbf{X}$ instead of via $\boldsymbol{\chi} = \mathbf{P} \circ \mathbf{X}$, and we denote by $q_d$ the area element obtained by doing so.

The discussion above applies as well to the piecewise polynomial surface $\Gamma$ (recall that we dropped the index specifying the considered patch). The corresponding matrix quantities are indexed by $\Gamma$, so that

$$\nabla \hat{v} = \nabla_\Gamma v \mathbf{G}_\Gamma, \qquad \nabla_\Gamma v = \nabla \hat{v}\, \mathbf{D}_\Gamma. \tag{2.17}$$

The first fundamental form and the elementary area on $\Gamma$ are denoted $\mathbf{g}_\Gamma$ and $q_\Gamma$. Finally, the Laplace-Beltrami operator on $\Gamma$ reads

$$\Delta_\Gamma v = \frac{1}{q_\Gamma} \mathrm{div}\big(q_\Gamma \nabla \hat{v} \mathbf{g}_\Gamma^{-1}\big). \tag{2.18}$$

In addition, we recall that for $T \in \mathcal{T}$ and $S$ a side of $T$, the outside pointing unit co-normals $\mathbf{n}$ of $T$ and $\widehat{\mathbf{n}}$ of $\widehat{T}$ are related by the following expressions

$$\widehat{\mathbf{n}} = \mathbf{n} \mathbf{G}_\Gamma \frac{r_\Gamma}{q_\Gamma} \qquad \text{and} \qquad \mathbf{n} = \frac{q_\Gamma}{r_\Gamma} \widehat{\mathbf{n}} \mathbf{D}_\Gamma, \tag{2.19}$$

where $r_\Gamma$ is the elementary area associated with the subsimplex $\widehat{S} := \boldsymbol{\chi}^{-1}(S)$. Hence, the tangential gradient of $v$ in the direction $\mathbf{n}$ on $S$ is given by

$$\nabla_\Gamma v \cdot \mathbf{n} = \frac{q_\Gamma}{r_\Gamma} \nabla v \mathbf{g}_\Gamma^{-1}|_{\widehat{S}} \cdot \widehat{\mathbf{n}}. \tag{2.20}$$

**2.4. Constants and inequalities.** We use different sets of constants and notation depending on the situation. We employ a constant $C$ (or $C_j$) which may depend on the space dimension $n$, the polynomial degrees $r$ and $k$, and shape regularity constants, but is independent of $L$, the discretization parameter $h$, and all other geometric information about $\gamma$. This notation is used mainly §2.5. The notation $a \lesssim b$ means that $a \leq Db$ with $D$ depending possibly on the same quantities as $C$ and also on $L$, but independent of other geometric information about $\gamma$ and the discretization parameter $h$. Finally, by $a \preceq b$ we mean $a \leq Gb$ with $G$ independent of the discretization parameter $h$ but possibly depending on geometric information about $\gamma$ and the right hand side $f$ in addition to the quantities hidden in $\lesssim$.

These different levels of precision with respect to constants reflect three different situations: In the next subsection we are concerned about verifying assumptions which assure the validity of our estimates and thus are as precise as possible concerning constants, including global constants such as the Lipschitz constant $L$. Global

constants such as $L$ are typically hidden as we do in $\lesssim$ when proving residual-type estimates, but it is often desirable to retain relevant *local* geometric information such as curvature variation in our estimates so that this information is taken into account when driving refinement in adaptive FEM. Finally, it becomes highly technical to reflect local geometric information when proving convergence of AFEM as in [12], so it is sometimes convenient to hide such information as we do in $\preceq$.

**2.5. Geometric resolution assumptions.** In order to prove a posteriori error estimates for surface finite element methods and analyze corresponding adaptive algorithms, it is necessary to assume some resolution of the continuous surface $\gamma$ by its discrete approximation. We thus make three main assumptions concerning the resolution of $\gamma$ by $\Gamma$ and the discrepancy between $\mathbf{P}$ and $\mathbf{P}_d$. These assumptions require more resolution of $\gamma$ than are needed to prove corresponding a posteriori estimates in [12] but are necessary to take advantage of the superior properties of the closest point projection $\mathbf{P}_d$.

**2.5.1. Restriction on distance between $\Gamma$ and $\gamma$.** We start by noting that the closest point property (1.6) satisfied by the distance function projection implies

$$|\mathbf{P}(\mathbf{x}) - \mathbf{P}_d(\mathbf{x})| \le |\mathbf{P}(\mathbf{x}) - \mathbf{x} + \mathbf{x} - \mathbf{P}_d(\mathbf{x})| \le 2|\mathbf{x} - \mathbf{P}(\mathbf{x})| \le 2|\mathbf{X}_T(\hat{\mathbf{x}}) - \boldsymbol{\chi}_T(\hat{\mathbf{x}})|$$

for $\mathbf{x} \in T$ and $\hat{\mathbf{x}} = \mathbf{X}^{-1}(\mathbf{x})$. As a consequence and using that $\mathbf{X}_T(\hat{\mathbf{x}}) - \boldsymbol{\chi}_T(\hat{\mathbf{x}}) = (I_k\mathbf{P} - \mathbf{P}) \circ \overline{\mathbf{X}}_T$, we have that for $\mathbf{x} \in T$ the discrepancy between the two lifts is estimated by the geometric estimator

$$|\mathbf{P}(\mathbf{x}) - \mathbf{P}_d(\mathbf{x})| \le 2\beta_T. \tag{2.21}$$

Recalling (2.13) and (2.12), we define for $0 < \alpha \le 1$

$$\mathcal{N}_\alpha := \{\mathbf{x} \in \mathbb{R}^{d+1} : \operatorname{dist}(\mathbf{x}, \gamma) \le \alpha \|K\|_{L_\infty(\gamma)}^{-1}\}. \tag{2.22}$$

and we make the assumption that

$$\Gamma \subset \mathcal{N}_{1/2}. \tag{A2}$$

This assumption is sufficient to guarantee that all points $\mathbf{x}$ appearing in our arguments below remain in $\mathcal{N}$, and that

$$\kappa_i(\mathbf{x}) \simeq \kappa_i(\mathbf{P}_d(\mathbf{x})) \tag{2.23}$$

by (2.11). Combining (1.6) with the observation that $|\mathbf{x} - \mathbf{P}(x)| \le \beta$, $\mathbf{x} \in \Gamma$, we see that (A2) is implied by

$$\beta \le \frac{1}{2}\|K\|_{L_\infty(\gamma)}^{-1}, \tag{A3}$$

and we shall assume (A3) throughout. This assumption is verifiable a posteriori. Recall that $K$ is the maximum over $\gamma$ of the principal curvatures $\kappa_i$, which are in turn the eigenvalues of the Weingarten map $\mathbf{W} = D\boldsymbol{\nu}$. In the most practically relevant case $n = 2$, the normal $\boldsymbol{\nu}$ may in turn be computed from the parametric maps $\boldsymbol{\chi}$ by taking a cross product of the columns of the matrix $\mathbf{G}$ and normalizing the result.

While we do not analyze mesh refinement schemes here, the assumption $\Gamma \subset \mathcal{N}_\alpha$ for $\alpha$ sufficiently small guarantees that mesh refinement carried out using the framework of [12] will result in a *sequence* of meshes lying in $\mathcal{N}_{\alpha'}$ with $\alpha < \alpha' < \frac{1}{2}$ sufficiently small. The threshold value of $\alpha$ depends only on $\|I_k\|_{L_\infty \to L_\infty}$ and is in principal computable. This property will be necessary to prove convergence of adaptive FEM using our estimators, which we hope to carry out in future work.

**2.5.2. Restriction on the mismatch between P and $\mathbf{P}_d$.** Our second assumption concerns the mismatch between the parametric map $\mathbf{P}$ used to describe $\gamma$ in our code and the distance function map $\mathbf{P}_d$. We assume that

$$\mathbf{P}_d \circ \mathbf{P}^{-1}(\widetilde{T}) \subset \widetilde{\omega}_T, T \in \mathcal{T}. \tag{A4}$$

In view of (2.8), (2.21), and the fact that $\gamma$ is a $C^2$ surface, we have that $|\mathbf{x} - \mathbf{P}_d \circ \mathbf{P}^{-1}(\mathbf{x})| \lesssim h^2$ independent of $k$. Because we expect that $\mathrm{dist}(\widetilde{T}, \partial \widetilde{\omega}_T) \simeq h_T$, (A4) is satisfied for sufficiently fine meshes.

As we prove in the following proposition, it is possible to check (A4) computationally without accessing $\mathbf{P}_d$ if the reference patches $\widehat{\omega}_T$ and constant $L$ in (2.3) are known. Before stating our proposition, we note that because $\gamma$ is a closed surface, there is a constant $c_{\widehat{T}}$ such that

$$\mathrm{dist}(\widehat{T}, \partial \widehat{\omega}_T) \geq c_{\widehat{T}}, \ T \in \mathcal{T}. \tag{2.24}$$

PROPOSITION 2.1. *Assume that* (A2) *is satisfied. Then* (A4) *is satisfied if*

$$3\beta_T \leq h_T L^{-1} c_{\widehat{T}}, \quad T \in \mathcal{T}. \tag{2.25}$$

*Proof.* Let $\mathrm{dist}_g$ denote the geodesic distance on $\gamma$ and $\mathrm{dist}_E$ the Euclidean distance. Then for $\mathbf{x}, \mathbf{y} \in \gamma$, $|\mathbf{x} - \mathbf{y}| = \mathrm{dist}_E(\mathbf{x}, \mathbf{y}) \leq \mathrm{dist}_g(\mathbf{x}, \mathbf{y})$. Also, if $\hat{x} \in \widehat{T}$ and $\hat{y} \in \partial \widehat{\omega}_T$, we have from (2.3) that $|\chi_T(\hat{\mathbf{x}}) - \chi_T(\hat{\mathbf{y}})| \geq h_T L^{-1} |\hat{\mathbf{x}} - \hat{\mathbf{y}}| \geq h_T L^{-1} c_{\widehat{T}}$. Thus

$$\mathrm{dist}_g(\widetilde{T}, \partial \widetilde{\omega}_T) \geq \mathrm{dist}_E(\widetilde{T}, \partial \widetilde{\omega}_T) \geq h_T L^{-1} c_{\widehat{T}}. \tag{2.26}$$

Consider now $\mathbf{x} \in T \subset \Gamma$. We wish to control $\mathrm{dist}_g(\mathbf{P}(\mathbf{x}), \mathbf{P}_d(\mathbf{x}))$. Note that the corresponding Euclidean distance is bounded in (2.21). Let $\tilde{\rho} : [0, 1] \to \mathbb{R}^{n+1}$ be given by $\tilde{\rho}(s) = s\mathbf{P}(\mathbf{x}) + (1-s)\mathbf{P}_d(\mathbf{x})$. A short computation and (2.21) imply that for every $s \in [0, 1]$

$$\mathrm{dist}_E(\tilde{\rho}(s), \gamma) \leq \frac{1}{2}|\mathbf{P}(\mathbf{x}) - \mathbf{P}_d(\mathbf{x})| \leq \beta_T, \tag{2.27}$$

so assumption (A3) implies that $\tilde{\rho}(s) \in \mathcal{N}_{1/2}$, $0 \leq s \leq 1$. We also define $\rho : [0, 1] \to \gamma$ by $\rho(s) = \mathbf{P}_d(\tilde{\rho}(s))$, which is well defined since $\tilde{\rho}(s) \in \mathcal{N}$. Thus by (2.22), $|d(\tilde{\rho}(s))\kappa_i(\tilde{\rho}(s))| \leq \frac{1}{2}$, $0 \leq s \leq 1$. Let $\rho_j$, $1 \leq j \leq n+1$ be the components of $\rho$. Then the length of the curve $\rho$ is given by

$$\mathrm{length}(\rho) = \int_0^1 \left[ \sum_{j=1}^{n+1} \left( \frac{d\rho_j}{ds} \right)^2 \right]^{1/2} ds. \tag{2.28}$$

But $\frac{d\rho_j}{ds} = [D\mathbf{P}_d(\tilde{\rho}(s))(\mathbf{P}(\mathbf{x}) - \mathbf{P}_d(\mathbf{x}))]_j$, so by (2.21) we have

$$\begin{aligned}
\mathrm{length}(\rho) &= \int_0^1 |D\mathbf{P}_d(\tilde{\rho}(s))(\mathbf{P}(\mathbf{x}) - \mathbf{P}_d(\mathbf{x}))| \, ds \\
&\leq \||D\mathbf{P}_d \circ \tilde{\rho}\|_{\ell_2 \to \ell_2}\|_{L_\infty(\tilde{\rho}[0,1])} 2\beta_T.
\end{aligned} \tag{2.29}$$

Because $\mathbf{W}\nabla d = 0$, there is an orthonormal basis $\{\nabla d, \mathbf{e}_1, ..., \mathbf{e}_n\}$ of $\mathbb{R}^{n+1}$, and $\mathbf{W} = \sum_{\ell=1}^{n} \kappa_\ell \mathbf{e}_\ell \otimes \mathbf{e}_\ell$. From (2.10), we thus have that $D\mathbf{P}_d = I - \nabla d \otimes \nabla d - d\mathbf{W} = \sum_{\ell=1}^{n}(1 - d\kappa_\ell)\mathbf{e}_\ell \otimes \mathbf{e}_\ell$. For $\mathbf{x} \in \mathcal{N}_{1/2}$, (2.22) yields

$$\|D\mathbf{P}_d(\mathbf{x})\|_{\ell_2 \to \ell_2} \leq \max_{1 \leq \ell \leq n} |1 - d(\mathbf{x})\kappa_\ell(\mathbf{x})| \leq 1 + \frac{1}{2} = \frac{3}{2},$$

which when combined with the observation that $\tilde{\rho}([0,1]) \subset \mathcal{N}_{1/2}$ and inserted into (2.29) yields

$$\text{length}(\rho) \leq 3\beta_T. \tag{2.30}$$

In particular, $\text{dist}_g(\mathbf{P}(\mathbf{x}), \mathbf{P}_d(\mathbf{x})) \leq \text{length}(\rho) \leq 3\beta_T$, which when combined with (2.26) yields (2.25). □

**2.5.3. Nondegeneracy of area elements.** The Lipschitz property (2.3) implies that on $\widetilde{\omega}_T$,

$$L^{-n}(h_T)^n \lesssim q \lesssim L^n(h_T)^n \tag{2.31}$$

We additionally assume that $q_\Gamma$ and $q_d$ are equivalent to $q$ in the sense that

$$C_1 \leq \frac{q}{q_\Gamma}, \frac{q_d}{q_\Gamma} \leq C_2. \tag{A5}$$

As above, we discuss the possibility of computationally checking the assumption (A5). From [12, Lemma 4.1] and [13, Lemma 5.5]), we have the following.

LEMMA 2.2. *Assume that $\lambda$ is sufficiently small with respect to $L$ and other nonessential constants. Then*

$$C_1 \leq \frac{q}{q_\Gamma} \leq C_2, \quad \|\boldsymbol{\nu} \circ \mathbf{P} - \boldsymbol{\nu}_\Gamma\|_{L_\infty(T)} \leq CL^{2n-1}\lambda_T, \ T \in \mathcal{T}. \tag{2.32}$$

Thus we shall assume throughout that

$$\lambda \text{ is sufficiently small that (2.32) holds.} \tag{A6}$$

We do not further discuss the precise threshold in (A6), but intead refer to [12].

In order to provide a condition for the second assumption in (A5), we first bound $|\boldsymbol{\nu} - \boldsymbol{\nu}_\Gamma| = |\boldsymbol{\nu} \circ \mathbf{P}_d - \boldsymbol{\nu}_\Gamma|$. We will also use these estimates in later sections.

LEMMA 2.3. *Assume that $T \in \mathcal{T}$, and that the assumptions (A2), (A4), and (A6) hold. Then*

$$\|\boldsymbol{\nu} - \boldsymbol{\nu}_\Gamma\|_{L_\infty(T)} \leq C[L^{2n-1}\lambda_T + L^2\|K\|_{L_\infty(\omega_T)}\beta_T], \tag{2.33}$$

$$\|\boldsymbol{\nu} - \boldsymbol{\nu}_\Gamma\|^2_{L_\infty(T)} \lesssim (\lambda_T)^2 + \|K\|_{L_\infty(\omega_T)}\beta_T. \tag{2.34}$$

*Proof.* For $\mathbf{x} \in T$,

$$\boldsymbol{\nu}(\mathbf{x}) - \boldsymbol{\nu}_\Gamma(\mathbf{x}) = \boldsymbol{\nu}(\mathbf{P}_d(\mathbf{x})) - \boldsymbol{\nu}(\mathbf{P}(\mathbf{x})) + \boldsymbol{\nu}(\mathbf{P}(\mathbf{x})) - \boldsymbol{\nu}_\Gamma(\mathbf{x}).$$

Employing (2.32) directly yields

$$\|\boldsymbol{\nu} \circ \mathbf{P} - \boldsymbol{\nu}_\Gamma\|_{L_\infty(T)} \leq CL^{2n-1}\lambda_T. \tag{2.35}$$

For the first difference, the assumption (A4) implies that $\mathbf{P}(\mathbf{x})$ and $\mathbf{P}_d(\mathbf{x})$ lie in the same element patch $\widetilde{\omega}_T$. Let $\widehat{\omega}_T$ be the reference patch corresponding to $T$, and let $\hat{\mathbf{x}} = \chi^{-1}(\mathbf{P}(\mathbf{x}))$ and $\hat{\mathbf{x}}_d = \chi^{-1}(\mathbf{P}_d(\mathbf{x}))$. $\widehat{\omega}_T$ is convex, so the line segment between $\hat{\mathbf{x}}$ and $\hat{\mathbf{x}}_d$ lies in $\widehat{\omega}_T$ also, and by (2.3) we have $|\hat{\mathbf{x}} - \hat{\mathbf{x}}_d| \leq L(h_T)^{-1}|\mathbf{P}(\mathbf{x}) - \mathbf{P}_d(\mathbf{x})|$. Using the Lipschitz character of $\chi$, we map this line segment to $\gamma$ and thus obtain a Lipschitz curve $\xi$ lying in $\widetilde{\omega}_T$. Using the arc length formula (cf. (2.28)) and the Lipschitz bound (2.3), we find that

$$|\xi| \leq C \int_{\hat{\mathbf{x}}}^{\hat{\mathbf{x}}_d} L h_T \leq CL^2|\mathbf{P}(\mathbf{x}) - \mathbf{P}_d(\mathbf{x})|.$$

Letting $D\xi$ be the unit tangent vector along the curve $\xi$, we have that

$$
\begin{aligned}
|\boldsymbol{\nu}(\mathbf{P}_d(\mathbf{x})) - \boldsymbol{\nu}(\mathbf{P}(\mathbf{x}))| &= \left| \int_\xi \mathbf{W} D\xi \right| \\
&\leq |\xi| \|\|\mathbf{W}\|_{\ell_2 \to \ell_2}\|_{L_\infty(\xi)} \leq CL^2|\mathbf{P}(\mathbf{x}) - \mathbf{P}_d(\mathbf{x})|\|K\|_{L_\infty(\omega_T)}.
\end{aligned}
\tag{2.36}
$$

Whence, in view of (2.21), we get

$$\|\boldsymbol{\nu} \circ \mathbf{P}_d - \boldsymbol{\nu} \circ \mathbf{P}\|_{L_\infty(T)} \leq CL^2\|K\|_{L_\infty(\omega_T)}\beta_T$$

which along with (2.35) yields (2.33). Noting that $\|\boldsymbol{\nu} \circ \mathbf{P} - \boldsymbol{\nu} \circ \mathbf{P}_d\|_{L_\infty(T)} \leq 2$ also trivially yields (2.34). □

The following lemma finally yields a checkable (up to a nonessential constant independent of geometric quantities) condition guaranteeing that the second relationship in (A5) holds.

LEMMA 2.4. *Under the assumptions of Lemma 2.3, there is a constant $C$ such that*

$$
\begin{aligned}
\left\| 1 - \frac{q_\Gamma}{q_d} \right\|_{L_\infty(T)} \leq{}& \frac{C[L^{4n-2}(\lambda_T)^2 + L^4(\|K\|_{L_\infty(\omega_T)}\beta_T)^2]}{1 - C[L^{4n-2}(\lambda_T)^2 + L^4(\|K\|_{L_\infty(\omega_T)}\beta_T)^2]} \\
&+ C\|K\|_{L_\infty(\omega_T)}\beta_T.
\end{aligned}
\tag{2.37}
$$

*Thus for $[L^{4n-2}(\lambda_T)^2 + L^4(\|K\|_{L_\infty(\omega_T)}\beta_T)^2]$ and $\|K\|_{L_\infty(\omega_T)}\beta_T$ sufficiently small, the second relationship in (A5) holds, and*

$$\left\| 1 - \frac{q_\Gamma}{q_d} \right\|_{L_\infty(T)} \lesssim (\lambda_T)^2 + \|K\|_{L_\infty(\omega_T)}\beta_T. \tag{2.38}$$

*Proof.* By Proposition 2.1 of [25], $|\boldsymbol{\nu}| = |\boldsymbol{\nu}_\Gamma| = 1$, and (A2), we have

$$\left| \frac{q_d}{q_\Gamma} \right| = |\boldsymbol{\nu} \cdot \boldsymbol{\nu}_\Gamma \Pi_{i=1}^n (1 - d(x)\kappa_i(x))| \lesssim 1. \tag{2.39}$$

This is the desired upper bound in (A5).

Again employing Proposition 2.1 of [25], then using (2.11), $|\boldsymbol{\nu}| = |\boldsymbol{\nu}_\Gamma| = 1$, $|1 -$

$\boldsymbol{\nu} \cdot \boldsymbol{\nu}_\Gamma| = \frac{1}{2}|\boldsymbol{\nu} - \boldsymbol{\nu}_\Gamma|^2$, and $d\kappa_i \leq \frac{1}{2} < 1$, we have for $x \in \Gamma$ that

$$
\begin{aligned}
\left|1 - \frac{q_\Gamma}{q_d}\right| &= \left|1 - \frac{1}{\boldsymbol{\nu} \cdot \boldsymbol{\nu}_\Gamma \Pi_{i=1}^n (1 - d(x)\kappa_i(x))}\right| \\
&= \left|(1 - \frac{1}{\boldsymbol{\nu} \cdot \boldsymbol{\nu}_\Gamma}) \frac{1}{\Pi_{i=1}^n (1 + d(x)\kappa_i(\mathbf{P}_d(x)))} - \frac{1}{\Pi_{i=1}^n (1 + d(x)\kappa_i(\mathbf{P}_d(x)))} + 1\right| \\
&\leq C \left|\frac{\boldsymbol{\nu} \cdot \boldsymbol{\nu}_\Gamma - 1}{\boldsymbol{\nu} \cdot \boldsymbol{\nu}_\Gamma - 1 + 1}\right| + C|d(x)K(\mathbf{P}_d(x))| \\
&\leq C \frac{\frac{1}{2}|\boldsymbol{\nu} - \boldsymbol{\nu}_\Gamma|^2}{1 - \frac{1}{2}|\boldsymbol{\nu} - \boldsymbol{\nu}_\Gamma|^2} + C|d(x)|K(\mathbf{P}_d(x)).
\end{aligned}
\tag{2.40}
$$

We next estimate $\|d\|_{L_\infty(T)}$. Recalling the closest point property (1.6), we have for $\hat{\mathbf{x}} \in \widehat{T}$ and $\mathbf{x} = \mathbf{X}_T(\hat{\mathbf{x}}) \in T$

$$
|d(\mathbf{x})| = |\mathbf{P}_d(\mathbf{x}) - \mathbf{x}| \leq |\mathbf{P}(\mathbf{x}) - \mathbf{x}| = |\boldsymbol{\chi}_T(\hat{\mathbf{x}}) - \mathbf{X}_T(\widehat{\mathbf{x}})|
$$

and so

$$
\|d\|_{L_\infty(T)} \leq \|\boldsymbol{\chi}_T - \mathbf{X}_T\|_{L_\infty(\widehat{T})} = \beta_T.
\tag{2.41}
$$

Inserting the square of (2.33) and (2.41) into (2.40) and using assumption (A4) to obtain $K(\mathbf{P}_d(\mathbf{x})) \leq \|K\|_{L_\infty(\omega_T)}$ completes the proof of (2.37). Instead employing (2.34) in the numerator yields (2.38). □

We conclude by stating our final assumption:

$$
L^{4n-2}(\lambda_T)^2 + L^4(\|K\|_{L_\infty(\omega_T)}\beta_T)^2 + \|K\|_{L_\infty(\omega_T)}\beta_T
\tag{A7}
$$
$$
\text{is small enough that (2.38) holds.}
$$

**3. Surface finite element method and a posteriori estimates.** In this section we first derive a weak formulation of $-\Delta_\gamma u = f$ as well as its finite element counterpart. We then prove a posteriori estimates under the assumption that $\gamma$ is $C^2$.

**3.1. Variational Formulation and Galerkin Method.** The space of square integrable functions on $\gamma$, with vanishing mean value is

$$
L_{2,\#}(\gamma) := \left\{v \in L_2(\gamma) \mid \int_\gamma v = 0\right\}
$$

and its subspace containing square integrable weak derivatives is

$$
H_\#^1(\gamma) := \left\{v \in L_{2,\#}(\gamma) \mid \nabla_\gamma v \in [L_2(\gamma)]^{n+1}\right\}.
$$

The weak formulation of $-\Delta_\gamma u = f$ is: For $f \in L_{2,\#}(\gamma)$, seek $u \in H_\#^1(\gamma)$ satisfying

$$
\int_\gamma \nabla_\gamma u \cdot \nabla_\gamma \varphi = \int_\gamma f \varphi, \qquad \forall \varphi \in H_\#^1(\gamma).
\tag{3.1}
$$

Existence and uniqueness of a solution $u \in H_\#^1(\gamma)$ is a consequence of the Lax-Milgram theorem.

On the surface approximation $\mathcal{T}$, the fully discrete problem associated with (3.1) consists of finding $U \in \mathbb{V}(\mathcal{T})$ satisfying

$$\int_\Gamma \nabla_\Gamma U \cdot \nabla_\Gamma V = \int_\Gamma F_\Gamma \, V \qquad \forall \, V \in \mathbb{V}(\mathcal{T}), \tag{3.2}$$

where $F_\Gamma \in L_{2,\#}(\Gamma)$ a suitable approximation of $f$. The Lax-Milgram theorem again ensures that (3.2) admits a unique solution $U$.

Following [13, 12], we take

$$F_\Gamma := f \frac{q}{q_\Gamma}, \tag{3.3}$$

which satisfies $\int_\Gamma F_\Gamma = \int_\gamma f = 0$. However, in [12, 13, 25], this choice also leads to no geometric consistency error in the approximation of the right hand side. In our setting, we recall all the quantities in (3.2) are defined using the practical lift $\mathbf{P}$. However, in order to take advantage of the super approximation properties of the distance function, they will have to be related to the corresponding quantities defined via the distance lift $\mathbf{P}_d$. As a consequence, geometric inconsistency will have to analyzed even for the right hand side.

**3.2. A Posteriori Residual Error Estimators and Oscillations.** In view of the considerations developed in Section 2.3.2, we define as in [12] the PDE error indicator for any $V \in \mathbb{V}(\mathcal{T})$ by

$$\eta_\mathcal{T}(V, F_\Gamma, T)^2 := h_T^2 \|F_\Gamma + \Delta_\Gamma U\|_{L_2(T)}^2 + h_T \|\mathcal{J}(V)\|_{L_2(\partial T)}^2 \qquad \forall \, T \in \mathcal{T},$$

where for an inter-element face $S \in S_\mathcal{T}$ ($S = T_+ \cap T_-$ with $T_+, T_- \in \mathcal{T}$)

$$\mathcal{J}(V)|_S := \nabla_\Gamma V|_{T_+} \cdot \mathbf{n}^+ + \nabla_\Gamma V|_{T_-} \cdot \mathbf{n}^-$$

and $\mathbf{n}^\pm$ are the outward pointing unit co-normal of $T^\pm$. The global estimator is the $l_2$ sum of the local quantities, i.e.

$$\eta_\mathcal{T}(V, F_\Gamma) := \left( \sum_{T \in \mathcal{T}} \eta_\mathcal{T}(V, F_\Gamma, T)^2 \right)^{1/2}.$$

We also introduce the *oscillation* for any integer $m \geq 1$, $m' \geq 0$, $V \in \mathbb{V}(\mathcal{T})$ and $T \in \mathcal{T}$

$$\begin{aligned}
\mathrm{osc}_\mathcal{T}(V, f, T)^2 &:= h_T^2 \left\| (\mathrm{id} - \Pi_m^2) \left( fq + \mathrm{div}\big( q_\Gamma \nabla V \mathbf{G}_\Gamma^{-1} \big) \right) \right\|_{L_2(\widehat{T})}^2 \\
&+ h_T \left\| (\mathrm{id} - \Pi_{m'}^2) \left( q_\Gamma^+ \nabla V^+ (\mathbf{G}_\Gamma^+)^{-1} \widehat{\mathbf{n}}^+ + q_\Gamma^- \nabla V^- (\mathbf{G}_\Gamma^-)^{-1} \widehat{\mathbf{n}}^- \right) \right\|_{L_2(\partial \widehat{T})}^2,
\end{aligned} \tag{3.4}$$

where $\widehat{\mathbf{n}}^\pm$ is defined according to (2.19), $\mathbf{g}_\Gamma^\pm$ and $q_\Gamma^\pm = \sqrt{\det \mathbf{g}_\Gamma^\pm}$ are the first fundamental form and area element associated to $T^\pm$, and $\Pi_m^p$ denotes the best $L_p$-approximation operator onto the space $\mathbb{P}^m$ of polynomials of degree $\leq m$; the domain is implicit from the context. As for the estimator, we define the global quantity

$$\mathrm{osc}_\mathcal{T}(V, f) = \left( \sum_{T \in \mathcal{T}} \mathrm{osc}_\mathcal{T}(V, f, T)^2 \right)^{1/2}.$$

In [12], $m = 2r - 2$ and $m' = 2r - 1$ is advocated to guarantee that the oscillations decay faster provided the surface parametrization has appropriate piecewise Besov regularity. In turn, this Besov regularity matches the regularity needed for the adaptive algorithm to deliver optimal rate of convergence when using $\lambda$ as geometric estimator. In contrast, since our new geometric estimator asymptotically scales like $\mu \approx \lambda^2 + \beta$, less Besov regularity might be required of the parametrization to deliver the same rate of convergence. In any event, we leave the choice of $m$ and $m'$ open for further studies on the optimality of the proposed algorithm but note that the constants appearing below might depend on these parameters.

**3.3. Geometric Inconsistencies.** The approximation of $u$ in (3.1) by $U$ in (3.2) depends on the approximation of $\gamma$ by $\Gamma$ (geometric approximation) and the approximation of $u \circ \mathbf{P}_d$ by $U$ on $\Gamma$ (or equivalently of $u$ by $U \circ \mathbf{P}_d^{-1}$). This is reflected in the error equation we propose to derive now. We start with the error

$$\|\nabla_\gamma(u - U \circ \mathbf{P}_d^{-1})\|_{L_2(\gamma)} = \sup_{v \in H^1_\#(\gamma),\ \|\nabla_\gamma v\|_{L_2(\gamma)}=1} \int_\gamma \nabla_\gamma(u - U \circ \mathbf{P}_d^{-1}) \cdot \nabla_\gamma v \quad (3.5)$$

and work on the integral term. Using the relation (3.1) defining $u$, we write

$$\int_\gamma \nabla_\gamma(u - U \circ \mathbf{P}_d^{-1}) \cdot \nabla_\gamma v = \int_\gamma fv - \int_\gamma \nabla_\gamma(U \circ \mathbf{P}_d^{-1}) \cdot \nabla_\gamma v.$$

In order to use the relation (3.2) satisfied by $U$, one needs to write the above two integrals on $\Gamma$ using the change of variables induced by $\mathbf{P}_d$. For the first integral this leads to

$$\int_\gamma fv = \int_\Gamma f_d v_d \frac{q_d}{q_\Gamma},$$

where $v_d = v \circ \mathbf{P}_d$, $f_d = f \circ \mathbf{P}_d$. The change of variables in the second integral is more involved, reading

$$\int_\gamma \nabla_\gamma(U \circ \mathbf{P}_d^{-1}) \cdot \nabla_\gamma v = \int_\Gamma \nabla_\Gamma U \cdot \nabla_\Gamma v_d - \int_\gamma \nabla(U \circ \mathbf{P}_d^{-1})^T \mathbf{E}_\Gamma^d \nabla_\gamma v$$

with

$$\mathbf{E}_\Gamma^d := \frac{q_\Gamma}{q_d} \mathbf{A}_\gamma \left(I - d\mathbf{W}\right) \mathbf{A}_\Gamma \left(I - d\mathbf{W}\right) \mathbf{A}_\gamma - \mathbf{A}_\gamma. \quad (3.6)$$

Here $\mathbf{A}_\gamma = (I - \boldsymbol{\nu} \otimes \boldsymbol{\nu})$, $\mathbf{A}_\Gamma = (I - \boldsymbol{\nu}_\Gamma \otimes \boldsymbol{\nu}_\Gamma)$, and the other relevant geometric quantities were defined in Section 2.3.1. Also, the quantities $d$ and $\mathbf{W}$ appearing above are evaluated on $\Gamma$. Adding and subtracting $\int_\Gamma F v_d$ from the integral term in (3.5), we get

$$\int_\gamma \nabla_\gamma(u - U \circ \mathbf{P}_d^{-1}) \cdot \nabla_\gamma v = \underbrace{\int_\Gamma F v_d - \int_\Gamma \nabla_\Gamma U \cdot \nabla_\Gamma v_d}_{=:I(v_d)} + \underbrace{\int_\Gamma f_d v_d \frac{q_d}{q_\Gamma} - \int_\Gamma F v_d}_{=:II(v_d)}$$
$$+ \underbrace{\int_\gamma \nabla(U \circ \mathbf{P}_d^{-1})^T \mathbf{E}_\Gamma^d \nabla_\gamma v}_{=:III(v)}. \quad (3.7)$$

We can now state our main result.

THEOREM 3.1 (Efficiency and Reliability). *Let $\gamma$ be a closed, compact and $C^2$ hypersurface in $\mathbb{R}^{n+1}$. Let the assumptions (A1)-(A7) of §2.2 and §2.5 hold. Then*

$$\|\nabla_\gamma(u - U \circ \mathbf{P}_d^{-1})\|_{L_2(\gamma)}$$

$$\lesssim \eta_{\mathcal{T}}(U, F_\Gamma) + \left( \sum_{T \in \mathcal{T}} \left[ \left( \lambda_T^2 + \beta_T \|K\|_{L_\infty(\widetilde{\omega}_T)} \right) \|\nabla_\Gamma U\|_{L_2(T)} + \beta_T \|f\|_{L_2(T)} \right]^2 \right)^{1/2}$$

$$\preceq \eta_{\mathcal{T}}(U, F_\Gamma) + \mu$$

*and*

$$\eta_{\mathcal{T}}(U, F_\Gamma) \preceq \|\nabla_\gamma(u - U \circ \mathbf{P}_d^{-1})\|_{L_2(\gamma)} + \mathrm{osc}_{\mathcal{T}}(U, f) + \mu.$$

Its proof directly follows upon estimating $I(v_d)$, $II(v_d)$ and $III(v_d)$. This is the subject of Sections 3.4, 3.5 and 3.6 below.

**3.4. Term I: A-posteriori Estimators.** The first term in the right hand side of (3.7) is estimated by the estimator $\eta_{\mathcal{T}}$ and is the focus of this section. We first take advantage of the relation (3.2) satisfied by the finite element solution $U$ to write

$$I(v_d) = \int_\Gamma F(v_d - V) - \int_\Gamma \nabla_\Gamma U \cdot \nabla_\Gamma(v_d - V) \tag{3.8}$$

for any $V \in \mathbb{V}(\mathcal{T})$. (3.8) is free of geometric approximation and standard arguments [1, 35] to derive upper and lower bounds for the energy error on flat domains can be extended to this case; see [12, 13, 25, 31]. With the notations introduced in Section 3.2, we have the following a posteriori error estimation result. Its proof is omitted as it is in essence Lemma 4.5 in [12].

LEMMA 3.2 (A posteriori upper and lower bounds).

$$\sup_{v \in H^1_\#(\gamma), \ \|\nabla_\gamma v\|_{L_2(\gamma)} = 1} I(v \circ \mathbf{P}_d) \lesssim \eta_{\mathcal{T}}(U, F) \tag{3.9}$$

*and*

$$\eta_{\mathcal{T}}(U, F_\Gamma) \lesssim \sup_{v \in H^1_\#(\gamma), \ \|\nabla v\|_{L_2(\gamma)} = 1} I(v \circ \mathbf{P}_d) + \mathrm{osc}_{\mathcal{T}}(U, f). \tag{3.10}$$

**3.5. Term III: Geometric Inconsistency in the Dirichlet form.** We discuss in this section the geometric inconsistencies appearing when approximating the PDE (3.1) on an approximated surface $\Gamma$. As indicated by the error equation (3.7), we need to analyze the geometric consistency for the distance lift, which is provided in [25]. Note that critical to taking advantage of the superconvergent properties of the distance lift is that the consistency error induced by using the practical lift does not appear.

Following (3.7), appropriately bounding Term III reduces to bounding $\|\mathbf{E}_\Gamma^d\|_{L_\infty}$ defined by (3.6). We carry this out in the following lemma.

LEMMA 3.3 (Geometric Consistency For the Distance Lift). *If the assumptions (A1)-(A7) of §2.2 and §2.5 hold, then for $T = \mathbf{X}_T(\widehat{T})$ we have*

$$\|\mathbf{E}_\Gamma^d\|_{L_\infty(T)} \lesssim (\lambda_T)^2 + \|K\|_{L_\infty(\omega_T)} \beta_T \preceq \mu_T, \tag{3.11}$$

*and*

$$\sup_{v\in H^1_\#(\gamma),\ \|\nabla v\|_{L_2(\gamma)}=1} III(v) \lesssim \left(\sum_{T\in\mathcal{T}} \left(\left[(\lambda_T)^2 + \|K\|_{L_\infty(\omega_T)}\beta_T\right]\|\nabla_\Gamma U\|_{L_2(T)}\right)^2\right)^{1/2}$$
$$\preceq \mu.$$

*Proof.* Because $\mathbf{A}_\gamma$ and $\mathbf{A}_\Gamma$ are projections, $\|\mathbf{A}_\gamma\|_{\ell_2\to\ell_2} \leq 1$ and $\|\mathbf{A}_\Gamma\|_{\ell_2\to\ell_2} \leq 1$. In addition, $\|\mathbf{W}\|_{\ell_2\to\ell_2} \leq K$. By (2.13) $\|d\mathbf{W}\|_{\ell_2\to\ell_2} \leq dK \leq 1$ and $\|I - d\mathbf{W}\|_{\ell_2\to\ell_2} \lesssim 1$. Using these facts easily yields that for $\mathbf{x}\in\Gamma$,

$$\|\mathbf{A}_\gamma(I-d\mathbf{W})\mathbf{A}_\Gamma(I-d\mathbf{W})\mathbf{A}_\gamma\|_{\ell_2\to\ell_2} \lesssim 1. \tag{3.12}$$

Elementary calculations and the assumption (A2) also yield that

$$\begin{aligned}
&\|\mathbf{A}_\gamma - \mathbf{A}_\gamma(I-d\mathbf{W})\mathbf{A}_\Gamma(I-d\mathbf{W})\mathbf{A}_\gamma\|_{\ell_2\to\ell_2}\\
&\qquad\lesssim \|\mathbf{A}_\gamma - \mathbf{A}_\gamma\mathbf{A}_\Gamma\mathbf{A}_\gamma\|_{\ell_2\to\ell_2} + |d(\mathbf{x})|K(\mathbf{x})\\
&\qquad= \|(\boldsymbol{\nu}_\Gamma - (\boldsymbol{\nu}_\Gamma\cdot\boldsymbol{\nu})\boldsymbol{\nu})\otimes(\boldsymbol{\nu}_\Gamma - (\boldsymbol{\nu}_\Gamma\cdot\boldsymbol{\nu})\boldsymbol{\nu})\|_{\ell_2\to\ell_2} + |d(\mathbf{x})|K(\mathbf{x})\\
&\qquad\lesssim |\boldsymbol{\nu}-\boldsymbol{\nu}_\Gamma|^2 + |d(\mathbf{x})|K(\mathbf{P}_d(\mathbf{x})),
\end{aligned} \tag{3.13}$$

where we used the property $|1-\boldsymbol{\nu}\cdot\boldsymbol{\nu}_\Gamma| = \frac{1}{2}|\boldsymbol{\nu}-\boldsymbol{\nu}_\Gamma|^2$ and (2.23) to derive the last inequality.

Recalling the definition (3.6) yields

$$\begin{aligned}
\|\mathbf{E}^d_\Gamma\|_{\ell_2\to\ell_2} \leq &\left|1 - \frac{q_\Gamma}{q_d}\right|\|\mathbf{A}_\gamma(I-d\mathbf{W})\mathbf{A}_\Gamma(I-d\mathbf{W})\mathbf{A}_\gamma\|_{\ell_2\to\ell_2}\\
&+ \|\mathbf{A}_\gamma - \mathbf{A}_\gamma(I-d\mathbf{W})\mathbf{A}_\Gamma(I-d\mathbf{W})\mathbf{A}_\gamma\|_{\ell_2\to\ell_2}.
\end{aligned}$$

Inserting (2.34) into (3.13), applying (A4) to obtain $K(\mathbf{P}_d(\mathbf{x})) \leq \|K\|_{L_\infty(\widetilde{\omega}_T)}$, and then gathering the result along with (3.12) and (2.37) into the above inequality yields (3.11). The second assertion follows from the definition of $III(v)$ in (3.7) and a Cauchy Schwarz inequality. □

**3.6. Term II: Data Inconsistencies.** We now focus on the second term in (3.7):

$$II(v_d) = \int_\Gamma f_d v_d \frac{q_d}{q_\Gamma} - \int_\Gamma F v_d.$$

where $F = (f\circ\mathbf{P})\frac{q}{q_\Gamma}$.

Recall that we do not assume that the distance function, and therefore $\mathbf{P}_d$, is accessible to the user. Otherwise, the choice $\mathbf{P} = \mathbf{P}_d$ would lead to $II = 0$. The mismatch between $\mathbf{P}(\mathbf{x})$ and $\mathbf{P}_d(\mathbf{x})$ is reflected in a non vanishing term $II$ which must be accounted for.

Lemma 3.4 provides an estimate for the effect of the discrepancy $\tilde{v} - \tilde{v}\circ\mathbf{P}_d\circ\mathbf{P}^{-1}$ for $\tilde{v}\in H^1(\gamma)$ in $L_2(\gamma)$. Its proof uses standard properties of mollifiers. Given $v\in H^1(\mathbb{R}^n)$, there exists a function $v^\epsilon$ such that for every $U\subset\mathbb{R}^n$

$$\|v - v^\epsilon\|_{L_2(U)} \lesssim \epsilon|v|_{H^1(U+B(0,\epsilon))} \tag{3.14}$$

$$\text{and}\qquad |v^\epsilon|_{W^1_\infty(U)} \lesssim \epsilon^{-n/2}|v^\epsilon|_{H^1(U+B(0,\epsilon))}. \tag{3.15}$$

We are now in position to estimate in $L_2$ the effect of the mismatch of $\mathbf{P}_d \neq \mathbf{P}$ on a $H^1$ function.

LEMMA 3.4 (Geometric Error in Function Evaluation). *Assume that assumptions* (A1)-(A7) *of Section §2.2 hold. Then for* $\tilde{w} \in H^1(\gamma)$ *and* $T \in \mathcal{T}$ *we have*

$$\|\tilde{w} - \tilde{w} \circ \mathbf{P}_d \circ \mathbf{P}^{-1}\|_{L_2(\widehat{T})} \lesssim \beta_T \|\tilde{w}\|_{H^1(\tilde{\omega}_T)}.$$

*Proof.* For notational ease, let $\psi = \mathbf{P}_d \circ \mathbf{P}^{-1}$. Fix $T \in \mathcal{T}$, let $\hat{w}(\hat{x}) = \tilde{w} \circ \chi_T(\hat{x})$ for $\hat{x} \in \widehat{\omega}_T$, and let $\hat{\psi} = (\chi_T)^{-1} \circ \psi \circ \chi_T$. Then by change of variables and (2.31),

$$\|\tilde{w} - \tilde{w} \circ \psi\|_{L_2(\widetilde{T})} \lesssim h_T^{n/2} \|\hat{w} - \hat{w} \circ \hat{\psi}\|_{L_2(\widehat{T})}.$$

The assumption $\mathbf{P}_d \circ \mathbf{P}^{-1}(\widetilde{T}) \subset \tilde{\omega}_T$ given in (A4) is equivalent to $\hat{\psi}(\widehat{T}) \subset \widehat{\omega}_T$ and is sufficient to ensure that the quantity on the right hand side is well-defined.

Note that $\hat{w}$ is defined on the reference patch $\widehat{\omega}_T$. There is a universal extension operator mapping $H^1(\widehat{\omega}_T)$ to $H^1(\mathbb{R}^n)$ which is bounded both in $L_2$ and in the $H^1$-seminorm. We thus may assume that $\hat{w}$ is defined and bounded in $H^1$ on all of $\mathbb{R}^n$ and that

$$|\hat{w}|_{H^1(\mathbb{R}^n)} \lesssim |\hat{w}|_{H^1(\widehat{\omega}_T)}.$$

Given $\epsilon > 0$, we denote by $\hat{w}^\epsilon$ the standard mollification of $\hat{w}$ with radius $\epsilon$. We may then apply (3.14) and (3.15) to $\hat{w}$ without restriction on $\epsilon$, and

$$\|\hat{w} - \hat{w} \circ \hat{\psi}\|_{L_2(\widehat{T})} \lesssim \|\hat{w} - \hat{w}^\epsilon\|_{L_2(\widehat{T})} + \|\hat{w}^\epsilon - \hat{w}^\epsilon \circ \hat{\psi}\|_{L_2(\widehat{T})} + \|\hat{w}^\epsilon \circ \hat{\psi} - \hat{w} \circ \hat{\psi}\|_{L_2(\widehat{T})}.$$

We first compute using (3.14) that

$$\|\hat{w} - \hat{w}^\epsilon\|_{L_2(\widehat{T})} \lesssim \epsilon |\hat{w}|_{H^1(\mathbb{R}^n)} \lesssim |\hat{w}|_{H^1(\widehat{\omega}_T)}.$$

Similarly, after applying a change of variables formula using (2.31) and (A5) and applying the restriction $\hat{\psi}(\widehat{T}) \subset \widehat{\omega}_T$ we find that

$$\|(\hat{w}^\epsilon - \hat{w}) \circ \hat{\psi}\|_{L_2(\widehat{T})} \lesssim \|\hat{w}^\epsilon - \hat{w}\|_{L_2(\widehat{\omega}_T)} \lesssim \epsilon |\hat{w}|_{H^1(\widehat{\omega}_T)}.$$

We finally bound the term $\|\hat{w}^\epsilon - \hat{w}^\epsilon \circ \hat{\psi}\|_{L_2(\widehat{T})}$. Let $\{x_\ell\}$ be a lattice on $\mathbb{R}^n$ with minimum distance between $x_\ell$ and $x_j$ $(\ell \neq j)$ equivalent to $\epsilon$ and such that $\{B_\epsilon(x_\ell)\}$ covers $\mathbb{R}^n$. The set $\{B_{M\epsilon}(x_\ell)\}$ then has finite overlap for any $M > 0$, with the maximum cardinality of the overlap depending on $M$. Let also $\epsilon = \|I - \hat{\psi}\|_{L_\infty(\widehat{T})}$. Then applying (3.15), we find that

$$\begin{aligned}
\|\hat{w}^\epsilon - \hat{w}^\epsilon \circ \hat{\psi}\|_{L_2(\widehat{T})}^2 &\leq \sum_\ell \|\hat{w}^\epsilon - \hat{w}^\epsilon \circ \hat{\psi}\|_{L_2(B_\epsilon(x_\ell) \cap \widehat{T})}^2 \\
&\lesssim \epsilon^n \sum_\ell \|\hat{w}^\epsilon - \hat{w}^\epsilon \circ \hat{\psi}\|_{L_\infty(B_\epsilon(x_\ell) \cap \widehat{T})}^2 \lesssim \epsilon^{n+2} \sum_\ell |\hat{w}^\epsilon|_{W_\infty^1(B_{2\epsilon}(x_\ell))}^2 \qquad (3.16) \\
&\lesssim \epsilon^2 \sum_\ell |\hat{w}|_{H^1(B_{3\epsilon}(x_\ell))}^2 \lesssim \epsilon^2 |\hat{w}|_{H^1(\mathbb{R}^n)}^2 \lesssim \epsilon^2 |\hat{w}|_{H^1(\widehat{\omega}_T)}^2.
\end{aligned}$$

We finally compute using the bi-lipschitz character of $\chi$ that

$$\epsilon = \|I - \hat{\psi}\|_{L_\infty(\widehat{T})} = \|(\chi_T)^{-1} \circ (I - \psi) \circ \chi_T\|_{L_\infty(\widehat{T})}$$
$$\leq Lh_T^{-1}\|(I - \psi) \circ \chi_T\|_{L_\infty(\widehat{T})} = Lh_T^{-1}\|I - \mathbf{P}_d \circ \mathbf{P}^{-1}\|_{L_\infty(\widetilde{T})}. \tag{3.17}$$

Recalling (2.21) then yields

$$\epsilon \lesssim h_T^{-1}\beta_T.$$

Again performing a change of variables yields

$$\|\tilde{w} - \tilde{w} \circ \psi\|_{L_2(\widetilde{T})}^2 \lesssim h_T^n\|\hat{w} - \hat{w} \circ \hat{\psi}\|_{L_2(\widehat{T})}^2 \lesssim h_T^n\epsilon^2|\hat{w}|_{H^1(\hat{\omega}_T)}^2$$
$$\lesssim h_T^n h_T^{-2}\beta_T^2 h_T^{2-n}|\tilde{w}|_{H^1(\widetilde{\omega}_T)}^2 = \beta_T^2|\tilde{w}|_{H^1(\widetilde{\omega}_T)}^2.$$

□

We now return to term II and estimate the discrepancy between the theoretical quantity

$$\int_\Gamma (fv) \circ \mathbf{P}_d \frac{q_d}{q_\Gamma}$$

and the practical quantity

$$\int_\Gamma (f \circ \mathbf{P})(v \circ \mathbf{P}_d)\frac{q}{q_\Gamma}.$$

This is the purpose of the next lemma.

LEMMA 3.5 (Data Geometric Consistency Error). *For every $v \in H^1(\gamma)$ and $f \in L_2(\gamma)$ we have*

$$\left|\int_\Gamma (fv) \circ \mathbf{P}_d\frac{q_d}{q_\Gamma} - \int_\Gamma (f \circ \mathbf{P})(v \circ \mathbf{P}_d)\frac{q}{q_\Gamma}\right| \lesssim \sum_{T \in \mathcal{T}} \beta_T\|f\|_{L_2(T)}\|v\|_{H^1(\omega_T)}$$
$$\lesssim \left(\sum_{T \in \mathcal{T}}(\beta_T)^2\|f\|_{L_2(T)}^2\right)^{1/2}\|v\|_{H^1(\gamma)} \tag{3.18}$$
$$\lesssim \|f\|_{L_2(\gamma)}\|v\|_{H^1(\gamma)}\beta.$$

*Thus*

$$\sup_{v \in H_\#^1(\gamma),\ \|\nabla v\|_{L_2(\gamma)}=1} II(v \circ \mathbf{P}_d) \lesssim \left(\sum_{T \in \mathcal{T}}\left[\|f\|_{L_2(\tilde{T})}\beta_T\right]^2\right)^{1/2} \preceq \beta.$$

*Proof.* We write both integrals on $\gamma$ using the lifts $\mathbf{P}_d$ and $\mathbf{P}$ respectively:

$$\int_\Gamma (fv) \circ \mathbf{P}_d\frac{q_d}{q_\Gamma} = \int_\gamma fv, \qquad \int_\Gamma (f \circ \mathbf{P})(v \circ \mathbf{P}_d)\frac{q}{q_\Gamma} = \int_\gamma f(v \circ \mathbf{P}_d \circ \mathbf{P}^{-1})$$

so that

$$\left|\int_\Gamma (fv) \circ \mathbf{P}_d\frac{q_d}{q_\Gamma} - \int_\Gamma (f \circ \mathbf{P})(v \circ \mathbf{P}_d)\frac{q}{q_\Gamma}\right| \lesssim \sum_{T \in \mathcal{T}} \|f\|_{L_2(T)}\|v - v \circ \mathbf{P}_d \circ \mathbf{P}^{-1}\|_{L_2(T)}.$$

The first "$\lesssim$" in (3.18) follows upon invoking Lemma 3.4, the second upon using Cauchy-Schwarz, and the third upon applying finite overlap of $\{\widetilde{\omega}_T\}_{T \in \mathcal{T}}$. The second assertion in Lemma 3.5 follows from the definition of $II(v_d)$ in (3.7) and the application of a Cauchy-Schwarz inequality. □

**4. Numerical Illustrations.** We present several examples illustrating the benefits when using the proposed geometric estimator $\mu$ instead of $\lambda$. We consider the adaptive algorithm proposed in [12, 13] and recalled now:

**AFEM:** Given an initial subdivision $\mathcal{T}_0$ and parameters $\varepsilon_0 > 0$, $0 < \rho < 1$, and $\omega > 0$, set $k = 0$.
   1. $\mathcal{T}_k^+ = \mathsf{ADAPT\_SURFACE}(\mathcal{T}_k, \omega\varepsilon_k)$
   2. $[U_{k+1}, \mathcal{T}_{k+1}] = \mathsf{ADAPT\_PDE}(\mathcal{T}_k^+, \varepsilon_k)$
   3. $\varepsilon_{k+1} = \rho\varepsilon_k$; $k = k + 1$
   4. go to 1.

The procedure $\mathsf{ADAPT\_SURFACE}(\mathcal{T}_k, \omega\varepsilon_k)$ targets a better approximation of the surface and produces a finer subdivision $\mathcal{T}_k^+$ such that the geometric estimator $\lambda_{\widehat{\mathcal{T}}_k^+}$ reaches a value below $\omega\varepsilon_k$. This is performed using a greedy algorithm, i.e. the elements of the subdivisions are successively refined until the geometric estimator on each element is below the targeted tolerance $\varepsilon$:

$\mathcal{T}_* = \mathsf{ADAPT\_SURFACE}(\mathcal{T}, \varepsilon)$
   1. if $\mathcal{M} := \{T \in \mathcal{T} : \lambda_{\mathcal{T}}(\gamma, T) > \varepsilon\} = \emptyset$
          return($\mathcal{T}$) and exit
   2. $\mathcal{T} = \mathsf{REFINE}(\mathcal{T}, \mathcal{M})$
   3. go to 1.

The procedure $\mathsf{ADAPT\_PDE}(\mathcal{T}_k^+, \varepsilon_k)$ instead consists of several iterations of the adaptive loop

$$\mathsf{SOLVE} \rightarrow \mathsf{ESTIMATE} \rightarrow \mathsf{MARK} \rightarrow \mathsf{REFINE} \tag{4.1}$$

aimed towards reducing the PDE error until the finer subdivision / approximate surface $\mathcal{T}_{k+1}/\Gamma_{k+1}$ are such that the associated finite element solution $U_{k+1}$ satisfies $\eta_{\mathcal{T}_{k+1}}(U_{k+1}, F_{\Gamma_{k+1}}) \leq \varepsilon_k$.

In both cases, the resulting subdivision is conforming upon refining additional elements without increasing the overall complexity. We refer to [12, 13] for additional details.

While our theory above is presented for subdivision made of simplices, our implementation is based on the deal.ii library [2] and uses quadrilaterals and finite element spaces $Q_s$ (instead of $\mathcal{P}_s$) based on polynomials of degree at most $s$ in each direction on the reference hypercube. In addition, in both of our examples below the surface is not closed, that is, $\partial\gamma \neq \emptyset$. Our arguments extend to these different situations under some further assumptions, most notably that $\mathbf{P}(\Gamma) = \mathbf{P}_d(\Gamma) = \gamma$, which is indeed the case for the two examples below.

**4.1. PDE error driven on a half sphere.** The first example consists of a case where the geometry, i.e. the surface, is smooth and therefore the dominant term estimating the error (see Theorem 3.1) is the PDE estimator $\eta$.

We consider the half sphere $\gamma = \mathbb{S}^2 \cap \{\mathbf{x} \geq 0\}$. We also choose $\omega = 1$ and $\rho = \frac{1}{2}$. The exact solution is given in spherical coordinates by $u = \sin(\phi)\sin(\theta)^{3/5}$, with $f = -\Delta_\gamma u$. Thus $u$ has a singularity at the north pole similar to a corner singularity on a nonconvex polygonal domain, and a graded mesh is necessary to recover optimal convergence. For the approximation parameters, we take $r = 2$ and $k = 1$, which asymptotically yields decays $\eta \sim N^{-1}$, $\lambda \sim N^{-\frac{1}{2}}$ and $\mu \sim N^{-1}$ when using $N$ degrees of freedom in a mesh. This is optimal for the given quantity. The striking difference in the performance of the adaptive routine when using the BCMMN estimator and the proposed one is illustrated in Figure 4.1 and Table 4.1. The meshes indicate that
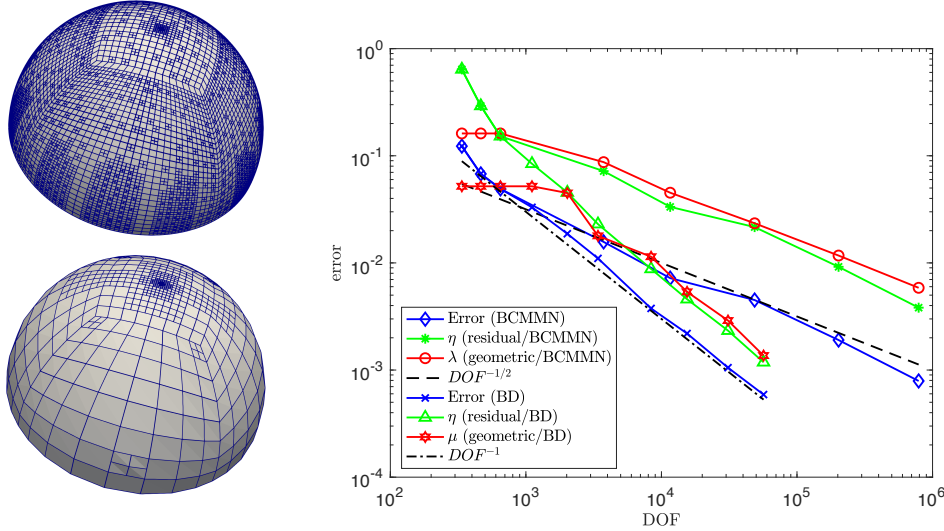
FIG. 4.1. *Smooth surface: Comparison between the algorithm based on the geometric estimators $\lambda$ (BCMMN) and $\mu$ (BD). The subdivisions constructed to achieve an error tolerance $\varepsilon = 3/2^8$ are provided in the left column for (top) BCMMN and (bottom) BD. When $\gamma$ is the upper half sphere, the adaptive algorithm based on $\mu$ produces a geometric refinement drastically less pronounced in order to achieve the targeted tolerance. In fact, the BCMMN algorithm yields an overrefinement for the geometry in this context. The right plot provide the evolution of the errors and estimators during the adaptive algorithms. The BD algorithm yields optimal error decay $DOF^{-1}$ unlike the BCMMN version which exhibits a convergence order of $DOF^{-1/2}$.*

|                                                                      | BCMMN | BD   |
|----------------------------------------------------------------------|-------|------|
| Total no of PDE refinement loops (4.1)                               | 9     | 10   |
| Total no of Geometric refinement loop in ADAPT_SURFACE               | 1388  | 7    |
| Total no of degree of freedom to achieve tolerance $\varepsilon = 3/2^8$ | 48689 | 3403 |

TABLE 4.1

*Comparison of the BCMMN [12, 13] and the proposed estimator when used in an adaptive finite element loop.*

the BCMMN estimator equidistributes far more degrees of freedom across the sphere in order to resolve geometry, while the BD estimator concentrates refinement at the north pole in order to control the singularity in the PDE solution. The BCMMN estimator quantitatively yields suboptimal convergence $N^{-1/2}$ of the overall error due to overrefinement based on the geometric estimator. In contrast, the proposed BD algorithm yields optimal $N^{-1}$ convergence.

**4.2. PDE error driven on a $C^{2,\alpha}$ surface.** We now turn our attention to a case where the geometric error plays an important role due to limited regularity of $\gamma$. For this, we consider the surface $\gamma$ given as the graph of

$$z(x,y) = \left(\frac{3}{4} - x^2 - y^2\right)_+^{2+\alpha}$$

for $\alpha = \frac{2}{5}$ and where $(x,y) \in (0,1)^2$ which implies $\gamma \in C^{2,\alpha}$. Here we also denote by $(v)_+$ the positive portion of $v$. In this case again, the parametrization of $\gamma$ is given by $z$. We set $f = 1$ and impose vanishing Dirichlet boundary conditions on $\partial\gamma$.

We execute both algorithms up to a final tolerance $\varepsilon = 5 \times 10^{-7}$ and again with $\omega = 1$ and $\rho = \frac{1}{2}$. For the BCMMN algorithm we set $r = k = 3$ while $r = 3$, $k = 2$ is chosen for the BD algorithm. The exact solution $u$ is unknown but we expect that the PDE estimator will behave like $N^{-3/2}$ when using $N$ degrees of freedom and optimal meshes. This is indeed the case for the BD algorithm as illustrated in Figure 4.2, where in addition to the values of the estimator for different values of $N$, the subdivisions constructed in both cases to guarantee and error smaller that $\varepsilon$ are reported. Note that the BCMMN algorithm seems to exhibit a slightly suboptimal error decay, likely due to inability to resolve the geometric error notion controlled by $\lambda$ with rate $N^{-3/2}$.

We finally note that the convergence order of $DOF^{-1}$ observed under uniform refinement above indicates that roughly $u \in H^3(\gamma)$. To see this, let $\boldsymbol{\chi} = (x, y, z(x,y))$ be the lift from $\Omega = (0,1)^2$ to $\gamma$, and let $\hat{u} = u \circ \boldsymbol{\chi}$. According to (2.16), $\hat{u}$ solves the elliptic PDE $-\mathrm{div}(q\nabla\hat{u}\mathbf{g}^{-1}) = q$ in $\Omega$. Note that $q = 1$ and the coefficient $q\mathbf{g}^{-1} = I$ outside of the circle $x^2 + y^2 = 1$, and these quantities vary inside of the circle. At the three corners of $\Omega$ where $z(x,y) = 0$ we are thus solving $-\Delta\hat{u} = 1$, and standard regularity theory for polygonal domains indicates that in those regions $u \in H^s$ for any $s < 3$ (cf. [18, Theorem 14.6]). At the origin (the corner of $\Omega$ at which $z(x,y)$ varies; cf. the illustration in Figure 4.2) $\hat{u}$ solves an elliptic PDE with smooth coefficient. Establishing the regularity of $\hat{u}$ near this corner is more complicated due to the presence of the coefficient, but the coefficient is smooth and isotropic at the origin and so it is plausible that the regularity is the same here as at the other corners. In addition, away from the boundary $\hat{u}$ solves a PDE with smooth coefficient. In particular, it is possible to calculate using the definitions in Section 2.3.2 that at a minimum the area element $q \in C^1(\Omega)$ (and thus also $q \in H^1(\Omega)$) and that the coefficient $q\mathbf{g}^{-1} \in C^2(\Omega)$. In fact, each appears to possess at least fractionally more smoothness than this. Standard elliptic regularity theory thus yields also that at the least $\hat{u} \in H^3(\mathrm{int}(\Omega))$. Thus elliptic corner singularities and not the regularity of $\gamma$ appear to place the heaviest restriction on the regularity of $\hat{u}$ and thus on convergence rates under quasiuniform refinement. On the other hand, these corner singularities have infinite smoothness in the context of adaptive refinement and are not responsible for the limited convergence rate observed when using BCMMN refinement.

**5. Perspectives.** In this section we briefly discuss two further questions raised by our work. The first is the question of convergence of adaptive FEM naturally generated from our a posteriori estimators, as were employed in the numerical tests above. The recent paper [12] from which we drew important elements of our technical structure proved convergence and optimality of an AFEM for the Laplace-Beltrami operator in which the geometric contribution is measured only by $\lambda$. We hope to prove similar results for our AFEM employing $\mu$.

Another fundamental question concerns the regularity of $\gamma$. Our results above strongly use the assumption that $\gamma$ is globally $C^2$, while [12] requires substantially less surface regularity. If $\gamma$ is $C^2$, our results yield a geometric error of heuristic a priori order $h^2$. On the other hand, if $\gamma$ is only $C^{1,1}$, our results have not been proven to apply, while the framework of [12] is still valid. That work however gives a geometric error contribution of heuristic a priori order $h$. Thus there is a jump in provable geometric error order from $h$ to $h^2$ when moving from $C^{1,1}$ to $C^2$ surfaces. To our knowledge it is a completely open question whether this is an artifact of proof, and if so, how to provide a unified theory of a priori and a posteriori error estimation for surfaces of varying regularities.
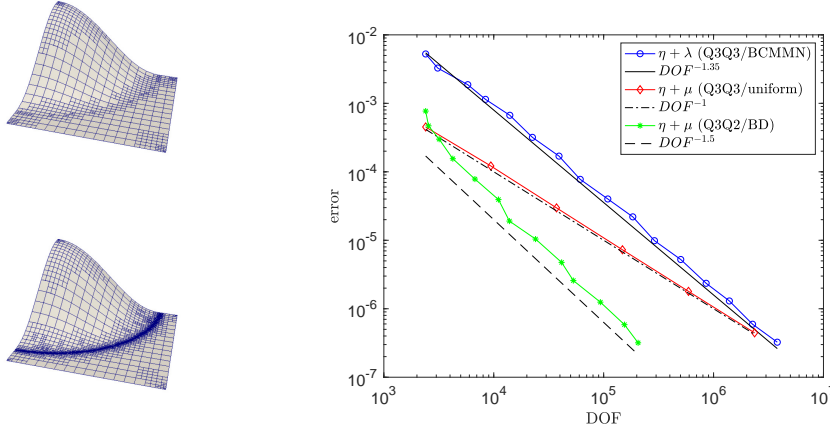
FIG. 4.2. *Smooth surface: Comparison between the algorithm based on the geometric estimators $\lambda$ (BCMN) with $r = k = 3$ and $\mu$ (BD) with $r = 3$, $k = 2$. The subdivisions constructed to achieve an error tolerance $\varepsilon = 5 \times 10^{-7}$ are provided in the left column for (top) BCMMN and (bottom) BD. In this case $\gamma$ is $C^{2,\frac{2}{5}}$ and the geometry approximation influences the rate of convergence. Again, the adaptive algorithm based on $\mu$ produces a geometric refinement drastically less pronounced in order to achieve the targeted tolerance. The right plot provide the evolution of the error estimators during the adaptive algorithms. With a one degree lower for the polynomial approximation of the geometry, the adaptive algorithm based on the BD estimator exhibits an optimal error decay $DOF^{-3/2}$, while the BCMMN algorithm appears to exhibit a slightly suboptimal error decay. For comparison, we also provide the outcome of a sequence of uniform refinement with $r = k = 3$ leading to an error decay of $DOF^{-1}$.*

## REFERENCES

[1] M. AINSWORTH AND J. T. ODEN, *A posteriori error estimation in finite element analysis*, Pure and Applied Mathematics (New York), Wiley-Interscience [John Wiley & Sons], New York, 2000.

[2] W. BANGERTH, R. HARTMANN, AND G. KANSCHAT, *deal.II—a general-purpose object-oriented finite element library*, ACM Trans. Math. Software, 33 (2007), pp. Art. 24, 27.

[3] E. BÄNSCH, *Finite element discretization of the Navier-Stokes equations with a free capillary surface*, Numer. Math., 88 (2001), pp. 203–235.

[4] E. BÄNSCH, P. MORIN, AND R.H. NOCHETTO, *Surface diffusion of graphs: variational formulation, error analysis, and simulation*, SIAM J. Numer. Anal., 42 (2004), pp. 773–799 (electronic).

[5] E. BÄNSCH, P. MORIN, AND R. H. NOCHETTO, *A finite element method for surface diffusion: the parametric case*, J. Comput. Phys., 203 (2005), pp. 321–343.

[6] J.W. BARRETT, H. GARCKE, AND R. NÜRNBERG, *Parametric approximation of Willmore flow and related geometric evolution equations*, SIAM J. Sci. Comput., 31 (2008), pp. 225–253.

[7] JOHN W. BARRETT, HARALD GARCKE, AND ROBERT NÜRNBERG, *Computational parametric Willmore flow with spontaneous curvature and area difference elasticity effects*, SIAM J. Numer. Anal., 54 (2016), pp. 1732–1762.

[8] ——, *Finite element approximation for the dynamics of asymmetric fluidic biomembranes*, Math. Comp., 86 (2017), pp. 1037–1069.

[9] ——, *Stable variational approximations of boundary value problems for Willmore flow with Gaussian curvature*, IMA J. Numer. Anal., 37 (2017), pp. 1657–1709.

[10] S. BARTELS, G. DOLZMANN, AND R. H. NOCHETTO, *A finite element scheme for the evolution of orientation order in fluid membranes*, M2AN Math. Model. Numer. Anal., 44 (2010), pp. 1–31.

[11] ——, *Finite element methods for director fields on flexible surfaces*, Interfaces Free Bound., ((to appear)).

[12] ANDREA BONITO, J. MANUEL CASCON, KHAMRON MEKCHAY, PEDRO MORIN, AND RICARDO H.

NOCHETTO, *High-order AFEM for the Laplace-Beltrami operator: Convergence rates.*, Found. Comput. Math., 16 (2016), pp. 1473–1539.

[13] A. BONITO, J. M. CASCÓN, P. MORIN, AND R. H. NOCHETTO, *AFEM for geometric PDE: the Laplace-Beltrami operator*, in Analysis and numerics of partial differential equations, vol. 4 of Springer INdAM Ser., Springer, Milan, 2013, pp. 257–306.

[14] A. BONITO, R.H. NOCHETTO, AND M.S. PAULETTI, *Parametric FEM for geometric biomembranes*, J. Comput. Phys., 229 (2010), pp. 3171–3188.

[15] A. BONITO, R. H. NOCHETTO, AND M. S. PAULETTI, *Dynamics of biomembranes: effect of the bulk fluid*, Math. Model. Nat. Phenom., 6 (2011), pp. 25–43.

[16] A. BONITO AND J.E. PASCIAK, *Convergence analysis of variational and non-variational multigrid algorithm for the Laplace-Beltrami operator*, Math. Comp., 81 (2012), pp. 1263–1288.

[17] FERNANDO CAMACHO AND ALAN DEMLOW, *$L_2$ and pointwise a posteriori error estimates for FEM for elliptic PDEs on surfaces*, IMA J. Numer. Anal., 35 (2015), pp. 1199–1227.

[18] MONIQUE DAUGE, *Elliptic boundary value problems on corner domains*, vol. 1341 of Lecture Notes in Mathematics, Springer-Verlag, Berlin, 1988.

[19] K. DECKELNICK AND G. DZIUK, *Discrete anisotropic curvature flow of graphs*, M2AN Math. Model. Numer. Anal., 33 (1999), pp. 1203–1222.

[20] ———, *Error estimates for a semi-implicit fully discrete finite element scheme for the mean curvature flow of graphs*, Interfaces Free Bound., 2 (2000), pp. 341–359.

[21] ———, *Error analysis of a finite element method for the Willmore flow of graphs*, Interfaces Free Bound., 8 (2006), pp. 21–46.

[22] K. DECKELNICK, G. DZIUK, AND C.M. ELLIOTT, *Fully discrete finite element approximation for anisotropic surface diffusion of graphs*, SIAM J. Numer. Anal., 43 (2005), pp. 1112–1138 (electronic).

[23] ANDREAS DEDNER AND PRAVIN MADHAVAN, *Adaptive discontinuous Galerkin methods on surfaces*, Numer. Math., 132 (2016), pp. 369–398.

[24] A. DEMLOW, *Higher-order finite element methods and pointwise error estimates for elliptic problems on surfaces*, SIAM J. Numer. Anal., 47 (2009), pp. 805–827.

[25] A. DEMLOW AND G. DZIUK, *An adaptive finite element method for the Laplace-Beltrami operator on implicitly defined surfaces*, SIAM J. Numer. Anal., 45 (2007), pp. 421–442 (electronic).

[26] ALAN DEMLOW AND MAXIM OLSHANSKII, *An adaptive surface finite element method based on volume meshes*, SIAM J. Numer. Anal., 50 (2012), pp. 1624–1647.

[27] G. DZIUK, *Finite elements for the Beltrami operator on arbitrary surfaces*, in Partial differential equations and calculus of variations, vol. 1357 of Lecture Notes in Math., Springer, Berlin, 1988, pp. 142–155.

[28] ———, *An algorithm for evolutionary surfaces*, Numer. Math., 58 (1991), pp. 603–611.

[29] ———, *Computational parametric willmore flow*, Numer. Math., 111 (2008), pp. 55–80.

[30] JÖRG GRANDE, *Analysis of highly accurate finite element based algorithms for computing distances to level sets*, SIAM J. Numer. Anal., 55 (2017), pp. 376–399.

[31] K. MEKCHAY, P. MORIN, AND R. H. NOCHETTO, *AFEM for the Laplace-Beltrami operator on graphs: design and conditional contraction property*, Math. Comp., 80 (2011), pp. 625–648.

[32] R.E. RUSU, *An algorithm for the elastic flow of surfaces*, Interfaces Free Bound., 7 (2005), pp. 229–239.

[33] UDO SEIFERT, *Configurations of fluid membranes and vesicles*, Advances in physics, 46 (1997), pp. 13–137.

[34] JEAN E. TAYLOR, *Some mathematical challenges in materials science*, Bull. Amer. Math. Soc. (N.S.), 40 (2003), pp. 69–87. Mathematical challenges of the 21st century (Los Angeles, CA, 2000).

[35] R. VERFÜRTH, *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Technique*, Wiley-Teubner, Chichester, 1996.

[36] MAX WARDETZKY, *Convergence of the cotangent formula: An overview*, in Discrete differential geometry, Springer, 2008, pp. 275–286.