GEOMETRIC TRANSFORMATION
OF FINITE ELEMENT METHODS:
THEORY AND APPLICATIONS

MICHAEL HOLST AND MARTIN LICHT

ABSTRACT. We present a new technique to apply finite element methods to partial differential equations over curved domains. A change of variables along a coordinate transformation satisfying only low regularity assumptions can translate a Poisson problem over a curved physical domain to a Poisson problem over a polyhedral parametric domain. This greatly simplifies both the geometric setting and the practical implementation, at the cost of having globally rough non-trivial coefficients and data in the parametric Poisson problem. Our main result is that a recently developed broken Bramble-Hilbert lemma is key in harnessing regularity in the physical problem to prove higher-order finite element convergence rates for the parametric problem. Numerical experiments are given which confirm the predictions of our theory.

1. INTRODUCTION

The computational theory of partial differential equations has been in a paradoxical situation from its very inception: partial differential equations over domains with curved boundaries are of theoretical and practical interest, but numerical methods are generally conceived only for polyhedral domains. Overcoming this geometric gap continues to inspire much ongoing research in computational mathematics for treating partial differential equations with geometric features. Computational methods for partial differential equations over curved domains commonly adhere to the philosophy of approximating the *physical domain* of the partial differential equation by a *parametric domain*. We mention isoparametric finite element methods [2], surface finite element methods [10, 8, 3], or isogeometric analysis [14] as examples, which describe the parametric domain by a polyhedral mesh whose cells are piecewise distorted to approximate the physical domain closely or exactly.

In this contribution we approach the topic from a different point-of-view. Our technique assumes to explicitly know a transformation of the physical domain, on which the original partial differential equation is stated, onto a polyhedral parametric domain. For example, the unit ball is a domain with curved boundary that is homeomorphic to the unit cube. Under mild regularity assumptions on that transformation, the partial differential equation on the curved physical domain can be

transformed to an equivalent partial differential equation over the polyhedral parametric domain. It remains then to numerically solve a partial differential equation over the parametric domain. We believe that the simplification of the geometry is of practical appeal; the trade-off is the coordinate transformation contributing low regularity terms to the parametric coefficients and parametric right-hand side. In a finite element method, the effect of using merely approximate problem data can be controlled easily by Strang's lemma. We emphasize that the transformation from the physical problem to the parametric problem takes places prior to any numerical analysis.

In this article, we give a thorough exposition of this technique and present exemplary numerical computations. A mathematical challenge is the regularity of the geometric transformation: in practice, the transformation is a diffeomorphism *locally* on each cell but of rather low regularity *globally* on the entire domain, typically no more than bi-Lipschitz. Thus it is not immediately evident how to leverage any higher regularity of the original physical problem for quasi-optimal error estimates in the finite element method for the practical parametric problem. Our main finding is how to overcome that obstacle via a broken Bramble-Hilbert lemma that has risen to prominence only recently [22, 5]; we believe the recent nature of the result to be the reason why our ostensibly simple technique, at its core only involving a change of variables, has not been received earlier by theoretical numerical analysists. Thus a second purpose of our article is to advertise the broken Bramble-Hilbert lemma to a broader audience. In a separate manuscript [13], we develop a generalization of the broken Bramble-Hilbert lemma suitable for use in finite element exterior calculus; an application of this result appears in [12].

As evidence that the approach has substantial potential for applications, we note that it has been fairly easy to implement our ideas in the finite element software library `FEniCS` [1], so we believe that practitioners can easily adopt this article's technique. Our numerical experiments confirm the theoretically predicted convergence rates.

We now finish this brief introduction by outlining the larger context and pointing out some further possibilities of our research. This work is a stepping stone towards developing *intrinsic* finite element methods for partial differential equations over manifolds, where it may be inconvenient or infeasible to describe the manifold extrinsically using a larger embedding manifold, so that one must work with an intrinsic description. If the manifold is computationally represented by a collection of coordinate charts onto parametric domains, then this article lays the foundation for the a priori error analysis of a finite element method. This research agenda will also touch upon finite element methods over embedded surfaces [5].

We are not aware of a prior discussion of this transformation technique in the literature of theoretical numerical analysis, but concrete applications are well-established in computational physics. An example is what is known as *cubed sphere* in atmospheric and seismic modeling [19, 18]. We hope that our work helps to connect those developments in practical computational physics with the numerical analysis of finite element methods.

For our numerical experiments we have calculated the parametric coefficients manually, which is feasible in applications such as atmospheric modeling with a fixed geometry of interest. However, these calculations can be automated when the

transformations are restricted to more specific classes, which seems to be more conforming to the demands on numerical methods in engineering. For example, our contribution complements the *parametric finite elements* [23] that have recently been formalized in computational engineering, albeit without a formal error analysis. Moreover, our results enable a priori error estimates for *NURBS-enhanced finite element methods* [21], where the physical geometry is represented over a reference geometry in terms of non-uniform rational B-splines (NURBS, [14]). Another area of application that we envision are rigorous error estimates for simplified computational models in physical modeling over unstructured meshes.

The remainder of this work is structured as follows. In Section 2 we introduce our model problem and review the relevant aspects of Galerkin theory. In Section 3 we prove the broken Bramble-Hilbert lemma and in Section 4 we elaborate on the a priori error analysis. Finally, we discuss numerical results in Section 5.

2. Model Problem and Abstract Galerkin Theory

As a model problem throughout this article we consider a variant of the Poisson equation with a diffusion tensor of low regularity. We then outline an abstract Galerkin theory that includes variational crimes.

2.1. **Function Spaces.** Let $\Omega \subseteq \mathbb{R}^n$ be a domain. For $p \in [1, \infty]$ we let $L^p(\Omega)$ be the Banach space of p-integrable functions over Ω. Moreover, for $p \in [1, \infty]$ and $s \in \mathbb{R}_0^+$ we let $W^{s,p}(\Omega)$ denote the Sobolev-Slobodeckij space over Ω with regularity index s and integrability index p. We write $\| \cdot \|_{W^{s,p}(\Omega)}$ for the norm of $W^{s,p}(\Omega)$, and we write $| \cdot |_{W^{s,p}(\Omega)}$ for the associated seminorm. In the case $p = 2$ the space $W^{s,p}(\Omega)$ carries a Hilbert space structure, and we denote by $\langle \cdot, \cdot \rangle_{W^{s,2}(\Omega)}$ the Sobolev-Slobodeckij semiscalar product of order s.

Whenever $\Gamma \subseteq \partial\Omega$ is a closed subset of the domain boundary, we define the space $W^{s,p}(\Omega, \Gamma)$ as the closure of the smooth functions in $W^{s,p}(\Omega)$ that vanish in an open neighborhood of Γ. For every boundary part $\Gamma \subseteq \partial\Omega$ we let Γ^c denote the *complementary boundary part*, which we define as the closure of $\partial\Omega \setminus \Gamma$. Then we write $W^{-s,p}(\Omega, \Gamma^c) := W^{s,p}(\Omega, \Gamma)^*$ for the dual space of $W^{s,p}(\Omega, \Gamma)$. This is a Banach space in its own right.

2.2. **Physical Model Problem.** We now introduce the physical model problem. Let $\hat{\Omega} \subseteq \mathbb{R}^n$ be a Lipschitz domain and $\hat{\Gamma}_D \subseteq \partial\hat{\Omega}$ be closed and non-empty. We let $\hat{\Gamma}_N \subseteq \partial\hat{\Omega}$ be the complementary boundary part. For simplicity we write

$$W_D^{s,p}(\hat{\Omega}) := W^{s,p}(\hat{\Omega}, \hat{\Gamma}_D), \quad W_N^{-s,p}(\hat{\Omega}) := W^{s,p}(\hat{\Omega}, \hat{\Gamma}_D)^*.$$

We assume that $\hat{A} \in L^\infty(\hat{\Omega})^{n \times n}$ is an essentially symmetric matrix field over $\hat{\Omega}$ that is invertible almost everywhere with $\hat{A}^{-1} \in L^\infty(\hat{\Omega})^{n \times n}$. We write $\| \cdot \|_{L^2(\hat{\Omega}, \hat{A})}$ for the associated weighted L^2 norm on the Hilbert space $L^2(\hat{\Omega})^n$, which is equivalent to the usual L^2 norm.

We introduce the symmetric bilinear form of the Poisson problem:

$$(1) \qquad \hat{B} : W_D^{1,2}(\hat{\Omega}) \times W_D^{1,2}(\hat{\Omega}) \to \mathbb{R}, \quad (\hat{u}, \hat{v}) \mapsto \int_{\hat{\Omega}} \nabla\hat{u} \cdot \hat{A}\nabla\hat{v} \, d\hat{x}.$$

Given a functional $\hat{F} \in W_{\mathrm{N}}^{-1,2}(\hat{\Omega})$, the model problem is to find $\hat{u} \in W_{\mathrm{D}}^{1,2}(\hat{\Omega})$ with

$$(2) \qquad\qquad \hat{B}(\hat{u}, \hat{v}) = \hat{F}(\hat{v}), \quad \hat{v} \in W_{\mathrm{D}}^{1,2}(\hat{\Omega}).$$

We recall that there exists $\hat{c}_P > 0$, depending only on $\hat{\Omega}$, $\hat{\Gamma}_{\mathrm{D}}$, and $\hat{A}$, such that

$$\hat{c}_P \|\hat{v}\|_{W^{1,2}(\hat{\Omega})}^2 \le \hat{B}(\hat{v}, \hat{v}), \quad \hat{v} \in W_{\mathrm{D}}^{1,2}(\hat{\Omega}).$$

The Lax-Milgram lemma [4] thus implies that the model problem (2) has a unique solution $\hat{u} \in W_{\mathrm{D}}^{1,2}(\hat{\Omega})$ satisfying the stability estimate

$$(3) \qquad\qquad \hat{c}_P \|\hat{u}\|_{W_{\mathrm{D}}^{1,2}(\hat{\Omega})} \le \|\hat{F}\|_{W_{\mathrm{N}}^{-1,2}(\hat{\Omega})}.$$

2.3. **Domain Transformation.** We henceforth call the domain $\hat{\Omega} \subseteq \mathbb{R}^n$ the *physical domain*. Additionally we now assume to be given another domain $\check{\Omega} \subseteq \mathbb{R}^n$, henceforth called *parametric domain*, and a homeomorphism

$$(4) \qquad\qquad \Phi : \check{\Omega} \to \hat{\Omega}$$

from the parametric domain onto the physical domain. As a minimal regularity assumption on this homeomorphism we assume that

$$(5) \qquad\qquad \Phi_i \in W^{1,\infty}(\check{\Omega}), \quad \Phi_i^{-1} \in W^{1,\infty}(\hat{\Omega})$$

for each coordinate index $1 \le i \le n$. This regularity assumption is satisfied, for example, in the special case that Φ is bi-Lipschitz. We write

$$\check{\Gamma}_{\mathrm{D}} = \Phi^{-1}(\hat{\Gamma}_{\mathrm{D}}), \quad \check{\Gamma}_{\mathrm{N}} = \Phi^{-1}(\hat{\Gamma}_{\mathrm{N}})$$

for the corresponding boundary patches along the parametric domain. On the parametric domain, too, we introduce the short-hand notation

$$W_{\mathrm{D}}^{s,p}(\check{\Omega}) := W^{s,p}(\check{\Omega}, \check{\Gamma}_{\mathrm{D}}), \quad W_{\mathrm{N}}^{-s,p}(\check{\Omega}) := W^{s,p}(\check{\Omega}, \check{\Gamma})^*.$$

The homeomorphism Φ and its inverse Φ^{-*} define isomorphisms between Sobolev spaces on the parametric domain and the physical domain:

$$(6a) \qquad\qquad \Phi^* : W_{\mathrm{D}}^{1,2}(\hat{\Omega}) \to W_{\mathrm{D}}^{1,2}(\check{\Omega}), \quad \hat{v} \mapsto \hat{v} \circ \Phi,$$

$$(6b) \qquad\qquad \Phi^{-*} : W_{\mathrm{D}}^{1,2}(\check{\Omega}) \to W_{\mathrm{D}}^{1,2}(\hat{\Omega}), \quad \check{v} \mapsto \check{v} \circ \Phi^{-1}.$$

2.4. **Parametric and Physical Model Problem.** The model problem over the physical domain is equivalent to a variational problem of the same class over the parametric domain. To begin with, we call the matrix field $\hat{A} : \hat{\Omega} \to \mathbb{R}^{n \times n}$ the *physical coefficient* and introduce the corresponding *parametric coefficient* as

$$(7) \qquad \check{A} : \check{\Omega} \to \mathbb{R}^{n \times n}, \quad \check{x} \mapsto |\det \mathrm{D}\Phi|_{|\check{x}} \cdot \mathrm{D}\Phi_{|\Phi(\check{x})}^{-1} \hat{A}_{|\Phi(\check{x})} \mathrm{D}\Phi_{|\Phi(\check{x})}^{-t}.$$

Next we define the *parametric bilinear form*

$$(8) \qquad \check{B} : W_{\mathrm{D}}^{1,2}(\check{\Omega}) \times W_{\mathrm{D}}^{1,2}(\check{\Omega}) \to \mathbb{R}, \quad (\check{u}, \check{v}) \mapsto \int_{\check{\Omega}} \nabla \check{u} \cdot \check{A} \nabla \check{v},$$

and the *parametric right-hand side* $\check{F} \in W_{\mathrm{N}}^{-1,2}(\check{\Omega})$ via

$$(9) \qquad\qquad \check{F}(\check{v}) := \hat{F}(\Phi^{-*}\check{v}), \quad \check{v} \in W_{\mathrm{D}}^{1,2}(\check{\Omega}).$$

Note the relation

$$\check{B}(\Phi^*\hat{u}, \Phi^*\hat{v}) = \hat{B}(\hat{u}, \hat{v}), \quad \hat{u}, \hat{v} \in W_{\mathrm{D}}^{1,2}(\hat{\Omega}).$$

There exists a constant $\hat{c}_P > 0$, which we call *parametric coercivity constant* and which depends only on $\hat{\Omega}$, $\hat{\Gamma}_D$, and $\hat{A}$, that satisfies

$$(10) \qquad \hat{c}_P \|\hat{v}\|^2_{W^{1,2}(\hat{\Omega})} \leq \hat{B}(\hat{v}, \hat{v}), \quad \hat{v} \in W^{1,2}_D(\hat{\Omega}).$$

Such a constant can also be bounded in terms of $\hat{c}_P$ and the derivatives of Φ and Φ^{-1} up to first order:

$$\hat{c}_P \leq \|\det D\Phi\|_{L^\infty(\hat{\Omega})} \|D\Phi^{-1}\|^2_{L^\infty(\hat{\Omega})} \hat{c}_P$$

The parametric model problem is finding $\hat{u} \in W^{-1,2}_N(\hat{\Omega})$ such that

$$(11) \qquad \hat{B}(\hat{u}, \hat{v}) = \hat{F}(\hat{v}), \quad \hat{v} \in W^{1,2}_D(\hat{\Omega}).$$

The unique solution $\hat{u} \in W^{1,2}_D(\hat{\Omega})$ of the physical model problem (2) and the unique solution $\hat{u} \in W^{1,2}_D(\hat{\Omega})$ of the parametric model problem (11) satisfy $\hat{u} = \Phi^* \hat{u}$. We henceforth call (2) the *physical model problem* and (11) the *parametric model problem*. We have

$$(12) \qquad \hat{c}_P \|\hat{u}\|_{W^{1,2}(\hat{\Omega})} \leq \|\hat{F}\|_{W^{-1,2}_N(\hat{\Omega})}.$$

We consider the transformation of the physical right-hand side $\hat{F}$ in more detail. The physical right-hand side $\hat{F}$ can be represented by a scalar function $\hat{f} \in L^2(\hat{\Omega})$ and a vector field $\hat{\mathbf{g}} \in L^2(\hat{\Omega})^n$ such that

$$(13) \qquad \hat{F}(\hat{v}) = \int_{\hat{\Omega}} \hat{f} \hat{v} \, d\hat{x} + \int_{\hat{\Omega}} \hat{\mathbf{g}} \cdot \nabla \hat{v} \, d\hat{x}, \quad \hat{v} \in W^{1,2}_D(\hat{\Omega}).$$

The parametric right-hand side $\hat{F}$ is then represented as follows: for every $\hat{v} \in W^{1,2}_D(\hat{\Omega})$ we have

$$(14) \qquad \hat{F}(\hat{v}) = \int_{\hat{\Omega}} |\det D\Phi| (\hat{f} \circ \Phi) \hat{v} \, d\hat{x} + \int_{\hat{\Omega}} |\det D\Phi| \left(D\Phi^{-1}_{|\Phi}(\hat{\mathbf{g}}_{|\Phi}) \right) \nabla \hat{v} \, d\hat{x}.$$

Remark 2.1. *Representation (13) is not only a theoretical consequence of the Riesz representation theorem; it appears practically when encoding boundary conditions in the problem data. Suppose that we search $\hat{u}_* \in W^{1,2}(\hat{\Omega})$ such that distributionally $-\operatorname{div}(\hat{A}\nabla \hat{u}_*) = \hat{f}$, where $\hat{f} \in L^2(\hat{\Omega})$, and satisfying the following mixed boundary conditions: along the Dirichlet boundary part $\hat{\Gamma}_D$, we want $\hat{u}_*$ to have the same boundary traces as some function $\hat{w} \in W^{1,2}(\hat{\Omega})$, and along the Neumann boundary part $\hat{\Gamma}_N$, we want $\hat{u}_*$ to the same normal boundary trace as some vector field $\hat{\mathbf{g}} \in L^2(\hat{\Omega})$ with $\operatorname{div} \hat{\mathbf{g}} \in L^2(\hat{\Omega})$. It is easily seen that if $\hat{u} \in W^{1,2}_D(\hat{\Omega})$ satisfies the model problem with right-hand side*

$$(15) \qquad \int_{\hat{\Omega}} \hat{f} \hat{v} \, d\hat{x} - \int_{\hat{\Omega}} \nabla \hat{w} \cdot \hat{A} \nabla \hat{v} \, d\hat{x} + \int_{\hat{\Omega}} (\operatorname{div} \hat{\mathbf{g}}) \hat{v} + \hat{\mathbf{g}} \cdot \nabla \hat{v} \, d\hat{x}, \quad \hat{v} \in W^{1,2}(\hat{\Omega}),$$

then such $\hat{u}_*$ is found by $\hat{u}_* = \hat{w} + \hat{u}$. The corresponding right-hand side of the parametric model problem is

$$(16) \qquad \int_{\hat{\Omega}} \hat{f} \hat{v} \, d\hat{x} - \int_{\hat{\Omega}} \nabla \hat{w} \cdot \hat{A} \nabla \hat{v} \, d\hat{x} + \int_{\hat{\Omega}} (\operatorname{div} \hat{\mathbf{g}}) \hat{v} + \hat{\mathbf{g}} \cdot \nabla \hat{v} \, d\hat{x}, \quad \hat{v} \in W^{1,2}(\hat{\Omega}),$$

where

$$(17) \qquad \hat{f} = |\det D\Phi| (\hat{f} \circ \Phi), \quad \hat{w} = \hat{w} \circ \Phi, \quad \hat{\mathbf{g}} = |\det D\Phi| \cdot D\Phi^{-1}_{|\Phi}(\hat{\mathbf{g}}_{|\Phi}).$$

We note that $\hat{\mathbf{g}}$ is the Piola transformation of the vector field $\hat{\mathbf{g}}$, which is known to preserve the class of divergence-conforming square-integrable vector fields. We note that the parametric right-hand side has the same structure as the physical right-hand side, encoding the same type of boundary conditions.

2.5. **Galerkin Theory.** We review conforming and non-conforming Galerkin approximation theories for the parametric model problem. We assume that we have a closed subspace $\hat{V}_h \subseteq W_{\mathrm{D}}^{1,2}(\hat{\Omega})$. A conforming Galerkin approximation for the model problem (2) seeks a solution $\hat{u}_h \in \hat{V}_h$ to

$$(18) \qquad \hat{B}(\hat{u}_h, \hat{v}_h) = \hat{F}(\hat{v}_h), \quad \hat{v}_h \in \hat{V}_h.$$

As in the case of the original problem, the Lax-Milgram lemma gives a unique solution $\hat{u}_h \in \hat{V}_h$ to the discrete problem (18), and we have

$$\hat{c}_P \|\hat{u}_h\|_{W_{\mathrm{D}}^{1,2}(\hat{\Omega})} \le \|\hat{F}\|_{W_{\mathrm{N}}^{-1,2}(\hat{\Omega})}.$$

In many applications, the bilinear form of the model problem or the right-hand side functional cannot be evaluated exactly but merely approximately over the Galerkin space. Formalizing, we assume to have another bounded bilinear form

$$(19) \qquad \hat{B}_h : \hat{V}_h \times \hat{V}_h \to \mathbb{R},$$

which is ought to approximate the original bilinear form $\hat{B}$ over the Galerkin space $\hat{V}_h$. We consider the following problem: given an approximate right-hand side functional $\hat{F}_h \in \hat{V}_h^*$, we seek a solution $\underline{\hat{u}}_h \in \hat{V}_h$ of

$$(20) \qquad \hat{B}_h(\underline{\hat{u}}_h, \hat{v}_h) = \hat{F}_h(\hat{v}_h), \quad \hat{v}_h \in \hat{V}_h.$$

It is practically reasonable to assume the existence of a *discrete parametric coercivity constant* $\hat{c}_{P,h} > 0$ such that

$$(21) \qquad \hat{c}_{P,h} \|\hat{v}_h\|_{W^{1,2}(\hat{\Omega})} \le \hat{B}_h(\hat{v}_h, \hat{v}_h), \quad \hat{v}_h \in \hat{V}_h.$$

Under this assumption, we can again apply the Lax-Milgram lemma to establish the well-posedness and stability of the Galerkin method. There exists a unique solution $\underline{\hat{u}}_h \in \hat{V}_h$ to the non-conforming discrete problem such that

$$(22) \qquad \hat{c}_{P,h} \|\underline{\hat{u}}_h\|_{\hat{V}_h} \le \|\hat{F}_h\|_{\hat{V}_h^*}.$$

We discuss a priori error estimates after an excursion into finite element approximation theory in the next section.

Remark 2.2. *Any conforming Galerkin method for the parametric model problem over $\hat{V}_h \subseteq W_{\mathrm{D}}^{1,2}(\hat{\Omega})$ translates into a conforming Galerkin method for the physical model problem over $\hat{V}_h \subseteq W_{\mathrm{D}}^{1,2}(\hat{\Omega})$. The subspaces are related through $\hat{V}_h = \Phi^* \hat{V}_h$.*

3. Finite Element Spaces and Error Estimates

This section introduces finite element spaces for our model problem and proves an approximation result central to our numerical approach. We continue the geometric setup of the preceding section but introduce a triangulation of the parametric domain as additional structure.

3.1. **Simplices and Triangulations.** We commence with gathering a few definitions concerning simplices and triangulations that will be used below.

A non-empty set $T \subseteq \mathbb{R}^n$ is a *d-dimensional simplex* if it is the convex closure of $d + 1$ affinely independent points $x_0, \ldots, x_d$, which are called the *vertices* of the simplex. For any d-dimensional simplex T we let $\mathcal{F}(T)$ denote the set of its $d + 1$ facets, where a facet is to be understood as a subsimplex of T whose vertices are all but one of the $d + 1$ vertices of T.

For the purpose of this article, a simplicial complex is a collection $\mathcal{T}$ of simplices such that for all $T \in \mathcal{T}$ and all $S \in \mathcal{F}(T)$ we have $S \in \mathcal{T}$ and such that for all $T_1, T_2 \in \mathcal{T}$ the intersection $T_1 \cap T_2$ is either empty or a simplex whose vertices are vertices of both T_1 and T_2.

For any simplex T of positive dimension d we let h_T be its diameter, and we call $\mu(T) := \operatorname{diam}(T)^d / \operatorname{vol}^d(T)$ the *shape measure* of T. The shape measure of a simplicial complex $\mathcal{T}$ is the maximum of the shape measures of its simplices and is denoted by $\mu(\mathcal{T})$. We also let $h_\mathcal{T}$ be the maximum diameter of any simplex of $\mathcal{T}$.

Following [22], we call a finite simplicial complex $\mathcal{T}$ *face-connected* whenever the following condition is true: for all n-dimensional simplices $S, T \in \mathcal{T}$ with $S \cap T \neq \emptyset$, there exists a sequence $T_0, T_1, \ldots, T_N$ of n-dimensional simplices of $\mathcal{T}$ such that $T_0 = S$ and $T_N = T$ and such that for all $1 \leq i \leq N$ we have that $F_i := T_i \cap T_{i-1}$ satisfies $F_i \in \mathcal{F}(T_{i-1}) \cap \mathcal{F}(T_i)$ and $S \cap T \subseteq F_i$. In other words, whenever two n-dimensional simplices share a common subsimplex, then we can traverse from the first to the second simplex by crossing facets of adjacent n-dimensional simplices and such that every simplex during the traversal will contain the intersection of the original two simplices as a subset.

3.2. **Polynomial Approximation over a Simplex.** Whenever T is a simplex of dimension d, we let $\mathcal{P}_r(T)$ denote the polynomials over T of degree at most $r \in \mathbb{N}_0$. We study interpolation and projection operators onto the polynomials of a simplex.

We first recall the definition of the Lagrange points over the simplex T. Letting $x_0, \ldots, x_d$ denote the vertices of the simplex T, we define the set of degree r *Lagrange points* by

$$\mathfrak{L}_r(T) := \left\{ \, (\alpha_0 x_0 + \cdots + \alpha_d x_d) / r \, \Big| \, \alpha = (\alpha_0, \ldots, \alpha_d) \in \mathbb{N}_0^{d+1}, \ |\alpha| = r \, \right\}.$$

We note that $\mathfrak{L}_r(F) \subseteq \mathfrak{L}_r(T)$ for every $F \in \mathcal{F}(T)$. We distinguish inner and outer Lagrange points: we let $\partial \mathfrak{L}_r(T) \subseteq \mathfrak{L}_r(T)$ be the set of Lagrange points of T that lie on the boundary of T and let $\mathring{\mathfrak{L}}_r(T) = \mathfrak{L}_r(T) \setminus \partial \mathfrak{L}_r(T)$.

For every $x \in \mathfrak{L}_r(T)$ we let δ_x^T denote the Dirac delta associated to that Lagrange point, which is an element of the dual space of $\mathcal{P}_r(T)$. The Dirac deltas associated to the Lagrange points constitute a basis for the dual space of $\mathcal{P}_r(T)$. The Lagrange polynomials are the associated predual basis: for every $x \in \mathfrak{L}_r(T)$ we let the polynomial $\Phi_{r,x}^T \in \mathcal{P}_r(T)$ be defined uniquely by

$$\Phi_{r,x}^T(y) = \delta_y^T \Phi_{r,x}^T = \delta_{xy}, \quad y \in \mathfrak{L}_r(T),$$

where δ_{xy} denotes the Kronecker delta. Obviously,

$$(23) \qquad\qquad v = \sum_{x \in \mathfrak{L}_r(T)} (\delta_x^T v) \Phi_{r,x}^T, \quad v \in \mathcal{P}_r(T).$$

It is worthwhile to extend the domain of the Dirac Deltas onto $L^p(T)$. For each Lagrange node $x \in \mathfrak{L}_r(T)$ we uniquely define $\Psi_{r,x}^T \in \mathcal{P}_r(T)$ through the condition

$$\int_T \Psi_{r,x}^T v \, \mathrm{d}x = \delta_x^T v, \quad v \in \mathcal{P}_r(T),$$

or equivalently,

$$\int_T \Psi_{r,x}^T \Phi_{r,y}^T \, \mathrm{d}x = \delta_{xy}, \quad y \in \mathfrak{L}_r(T).$$

The following estimates derive from a scaling argument and is stated without proof:

Lemma 3.1. *Let T be a d-dimensional simplex. Let $p \in [1, \infty)$, $s \in \mathbb{R}$, and $r \in \mathbb{N}_0$. Then there exists $C_{\mu,d,r,s,p} > 0$, depending only on d, r, s, p, and $\mu(T)$, such that*

$$(24) \qquad \left| \Phi_{r,x}^T \right|_{W^{s,p}(T)} \leq C_{\mu,d,r,s,p} h_T^{\frac{d}{p}-s}, \quad \left| \Psi_{r,x}^T \right|_{W^{s,p}(T)^*} \leq C_{\mu,d,r,s,p} h_T^{-\frac{d}{p}+s}.$$

We fix quasi-optimal interpolation operators over each triangle. Specifically, for each n-dimensional $T \in \mathcal{T}$ we assume to have an idempotent linear mapping

$$P_{T,r,s,p} : W^{s,p}(T) \to \mathcal{P}_r(T) \subseteq W^{s,p}(T)$$

that satisfies

$$\int_T P_{T,r,s,p} v \, \mathrm{d}x = \int_T v \, \mathrm{d}x, \quad v \in W^{s,p}(T),$$

and such that for some $C_{d,s,r,p,\mu}^{\mathrm{I}} > 0$, depending only on n, s, p, r, and $\mu(\mathcal{T})$, we have

$$|u - P_{T,r,s,p}u|_{W^{s,p}(T)} \leq C_{d,s,r,p,\mu}^{\mathrm{I}} \inf_{v \in W^{s,p}(T)} |u - v|_{W^{s,p}(T)}, \quad v \in W^{s,p}(T).$$

This existence of such a mapping follows from a scaling argument.

The following trace inequality will be used.

Lemma 3.2. *Let T be a d-dimensional simplex and let F be a facet of T. Let $p \in [1, \infty)$ and $s \in (1/p, 1]$. Then there exists a constant $C_{p,s,d,\mu}^{\mathrm{Tr}} > 0$, depending only on p, s, d, and $\mu(T)$, such that*

$$\|v - P_{T,r,s,p}v\|_{L^p(F)} \leq C_{p,s,d,\mu}^{\mathrm{Tr}} h_T^{s-\frac{1}{p}} |v - P_{T,r,s,p}v|_{W^{s,p}(T)}, \quad v \in W^{s,p}(T).$$

Proof. We note that $v - P_{T,r,s,p}v$ has zero average over T. The lemma now follows by combining the trace inequality that is Lemma 7.2 of [11] and the Poincaré inequality that is Lemma 7.1 of [11]. □

3.3. **Polynomial Approximation over Triangulations.** We now extend our discussion to piecewise polynomial approximation spaces over entire triangulations. We fix a triangulation $\mathcal{T}$ of the parametric domain $\hat{\Omega}$, i.e., a simplicial complex the union of whose simplices is the closure of the parametric domain. In order to formally handle boundary conditions, we assume that the boundary part $\hat{\Gamma}_\mathrm{D}$ is the union of simplices in $\mathcal{T}$.

We first introduce the *broken* or *non-conforming* Lagrange space

$$\mathcal{P}_{r,-1}(\mathcal{T}) := \left\{ u \in L^1(\hat{\Omega}) \ \middle|\ \forall T \in \mathcal{T} : u_{|T} \in \mathcal{P}_r(T) \right\}.$$

The *conforming* Lagrange spaces without and with boundary conditions are

$$\mathcal{P}_r(\mathcal{T}) := \mathcal{P}_{r,-1}(\mathcal{T}) \cap W^{1,2}(\hat{\Omega}), \qquad \mathcal{P}_{r,\mathrm{D}}(\mathcal{T}) := \mathcal{P}_{r,-1}(\mathcal{T}) \cap W_\mathrm{D}^{1,2}(\hat{\Omega}).$$

We construct a basis for the global finite element space from the Lagrange basis functions over single simplices. We first introduce the Lagrange points of the triangulation,

$$\mathfrak{L}_r(\mathcal{T}) := \bigcup_{T \in \mathcal{T}} \mathfrak{L}_r(T).$$

We note that Lagrange points can be shared between distinct simplices. Extending the notion of Lagrange polynomials to the case of triangulations, for each $x \in \mathfrak{L}_r(\mathcal{T})$ we define the function $\Phi_{r,x}^{\mathcal{T}} \in \mathcal{P}_r(\mathcal{T})$ on each simplex $T \in \mathcal{T}$ via

$$\Phi_{r,x|T}^{\mathcal{T}} := \begin{cases} \Phi_{r,x}^T & \text{if } x \in \mathfrak{L}_r(T), \\ 0 & \text{if } x \notin \mathfrak{L}_r(T). \end{cases}$$

For the degrees of freedom we assume that $r \in \mathbb{N}$, $p \in [1, \infty)$ and $s \in (1/p, 1]$ and apply the following construction. Whenever $x \in \mathring{\mathfrak{L}}_r(T)$ is an internal Lagrange point of a full-dimensional simplex $T \in \mathcal{T}$, then we define

$$J_{r,s,p,x}^{\mathcal{T}} : W^{s,p}(\hat{\Omega}) \to \mathbb{R}, \quad v \mapsto \delta_x^T P_{T,r,s,p} v.$$

Whenever $x \in \mathfrak{L}_r(\mathcal{T})$ is not an internal Lagrange point of any full-dimensional simplex of the triangulation, then we first fix a facet $F_x \in \mathcal{T}$ of codimension one for which $x \in \mathfrak{L}_r(F)$ holds; moreover, if $x \in \hat{\Gamma}_D$, then we require that $F_x \subseteq \hat{\Gamma}_D$. It is easily seen that this condition can always be satisfied. Now we define

$$J_{r,s,p,x}^{\mathcal{T}} : W^{s,p}(\hat{\Omega}) \to \mathbb{R}, \quad \hat{v} \mapsto \int_{F_x} \Psi_{r,x}^F \operatorname{tr}_F(\hat{v}) \, \mathrm{d}x.$$

The global projection is

$$\Pi : W^{s,p}(\hat{\Omega}) \to \mathcal{P}_r(\mathcal{T}), \quad \hat{v} \mapsto \sum_{x \in \mathfrak{L}_r(\mathcal{T})} J_{r,s,p,x}^{\mathcal{T}}(\hat{v}) \Phi_{r,x}^{\mathcal{T}}.$$

This operator is idempotent and bounded. Furthermore, the interpolant preserves boundary conditions: $\Pi v \in \mathcal{P}_{r,D}(\mathcal{T})$ whenever $v \in W^{s,p}(\hat{\Omega}, \hat{\Gamma})$.

This completes the construction of a Scott-Zhang-type interpolant. We now discuss a general error estimate for this approximation operator. A special case of the following result is due to Veeser [22], and a slightly different version is due to Camacho and Demlow [5].

Theorem 3.3. *Assume that $\mathcal{T}$ is face-connected and let $p \in [1, \infty)$ and $s \in \mathbb{R}$ with $s > 1/p$. Then there exist $C_\Pi > 0$ such that for all $r \in \mathbb{N}$, all full-dimensional $T \in \mathcal{T}$, and all $\hat{v} \in W^{s,p}(\hat{\Omega})$ we have*

$$\left| \hat{v} - \Pi\hat{v} \right|_{W^{s,p}(T)} \leq \left| \hat{v} - P_{T,r,s,p}\hat{v} \right|_{W^{s,p}(T)} + C_\Pi \sum_{\substack{T' \in \mathcal{T} \\ T \cap T' \neq \emptyset}} \left| \hat{v} - P_{T',r,s,p}\hat{v} \right|_{W^{s,p}(T')}.$$

The constant C_Π depends only on p, d, r, s, and $\mu(\mathcal{T})$.

Proof. We first observe via the triangle inequality that

$$\left| \hat{v} - \Pi\hat{v} \right|_{W^{s,p}(T)} \leq \left| \hat{v} - P_{T,r,s,p}\hat{v} \right|_{W^{s,p}(T)} + \left| P_{T,r,s,p}\hat{v} - \Pi\hat{v} \right|_{W^{s,p}(T)}.$$

The polynomial identity (23) and Lemma 3.1 give

$$\left| P_{T,r,s,p}\hat{v} - \Pi\hat{v} \right|_{W^{s,p}(T)} \leq \sum_{x \in \mathfrak{L}_r(T)} \left| \delta_x^T P_{T,r,s,p}\hat{v} - \delta_x^T \Pi\hat{v} \right| \cdot \left| \Phi_{r,x}^T \right|_{W^{s,p}(T)}$$

$$\leq C_{\mu,d,r,s,p} h_T^{\frac{n}{p}-s} \sum_{x \in \mathfrak{L}_r(T)} \left| \delta_x^T P_{T,r,s,p}\hat{v} - \delta_x^T \Pi\hat{v} \right|.$$

Suppose that $x \in \mathring{\mathfrak{L}}_r(T)$ is an internal Lagrange point. Then definitions imply

$$\delta_x^T P_{T,r,s,p}\hat{v} = \delta_x^T \Pi\hat{v}.$$

If instead $x \in \partial\mathfrak{L}_r(T)$ is a Lagrange point in the boundary, the estimate is more intricate. Recall that the facet F_x associated to the Lagrange point x is a facet of some simplex $S \in \mathcal{T}$ in the triangulation. Since $\mathcal{T}$ is face-connected, there exists a sequence $T_0, T_1, \ldots, T_N$ of n-dimensional simplices of $\mathcal{T}$ such that $T_0 = S$ and $T_N = T$ and such that for all $1 \leq i \leq N$ we have that $F_i := T_i \cap T_{i-1}$ satisfies $F_i \in \mathcal{F}(T_{i-1}) \cap \mathcal{F}(T_i)$ and $T \cap S \subseteq F_i$. Furthermore, we may assume that $F_0 = F_x$.

Unrolling definitions and using a telescopic expansion, we see

$$\delta_x^T P_{T,r,s,p}\hat{v} - \delta_x^T \Pi\hat{v} \, \mathrm{d}x$$

$$= \delta_x^T P_{T,r,s,p}\hat{v} - \int_{F_x} \Psi_{r,x}^{F_x} \operatorname{tr}_{F_x} \hat{v} \, \mathrm{d}x$$

$$= \int_{F_N} \Psi_{r,x}^{F_N} \operatorname{tr}_{F_N} P_{T,r,s,p}\hat{v} \, \mathrm{d}x - \int_{F_0} \Psi_{r,x}^{F_0} \operatorname{tr}_{F_0} \hat{v} \, \mathrm{d}x$$

$$= \int_{F_0} \Psi_{r,x}^{F_0} \operatorname{tr}_{F_0} (P_{T_0,r,s,p}\hat{v} - \hat{v}) \, \mathrm{d}x$$

$$+ \sum_{i=1}^{N} \int_{F_i} \Psi_{r,x}^{F_i} \operatorname{tr}_{F_i} (P_{T_i,r,s,p}\hat{v} - P_{T_{i-1},r,s,p}\hat{v}) \, \mathrm{d}x.$$

Since $F_0 \subset T_0$ and $P_{T_0,r,s,p}\hat{v} - \hat{v}$ has by construction vanishing mean over T, we use Hölder's inequality and Lemma 3.1 to conclude that

$$\left| \int_{F_0} \Psi_{r,x}^{F_0} \operatorname{tr}_{F_0} (P_{T_0,r,s,p}\hat{v} - \hat{v}) \, \mathrm{d}x \right| \leq C_{\mu,d,r,s,p} h_{T_0}^{\frac{1-n}{p}} \left\| \operatorname{tr}_{F_0} (P_{T_0,r,s,p}\hat{v} - \hat{v}) \right\|_{L^p(F_0)}.$$

Similarly, for $1 \leq i \leq N$ we observe $F_i \subseteq T_i \cap T_{i-1}$ and thus get

$$\left| \int_{F_i} \Psi_{r,x}^{F_i} \operatorname{tr}_{F_i} (P_{T_i,r,s,p}\hat{v} - P_{T_{i-1},r,s,p}\hat{v}) \, \mathrm{d}x \right|$$

$$\leq C_{\mu,d,r,s,p} h_T^{\frac{1-n}{p}} \left\| \operatorname{tr}_{F_i} (P_{T_i,r,s,p}\hat{v} - \hat{v}) \right\|_{L^p(F_i)}$$

$$+ C_{\mu,d,r,s,p} h_T^{\frac{1-n}{p}} \left\| \operatorname{tr}_{F_i} (P_{T_{i-1},r,s,p}\hat{v} - \hat{v}) \right\|_{L^p(F_i)}.$$

Here we have used Lemma 3.1 and the fact that $\hat{v} \in W^{s,p}(\hat{\Omega})$ has well-defined traces along simplex facets. Recall that Lemma 3.2 gives

$$\| \operatorname{tr}_F (P_{T,r,s,p}\hat{v} - \hat{v}) \|_{L^p(F)} \leq C_{p,s,d,\mu}^{\mathrm{Tr}} h_T^{s-\frac{1}{p}} \left| P_{T,r,s,p}\hat{v} - \hat{v} \right|_{W^{s,p}(T)}$$

whenever $F \subseteq T$ is a facet of a simplex $T \in \mathcal{T}$. Note that the shape measure of $\mathcal{T}$ bounds the number of simplices of the triangulation adjacent to T (see [16,

Lemma II.4.1]). Hence there exists $C_\Pi > 0$, depending only on s, r, d, and $\mu(\mathcal{T})$, such that

$$\left|P_{T,r,s,p}\hat{v} - \Pi\hat{v}\right|_{W^{s,p}(T)} \leq C_\Pi h_T^{\frac{n}{p}-s} \sum_{\substack{T' \in \mathcal{T} \\ T \cap T' \neq \emptyset}} h_T^{\frac{1-n}{p}} h_T^{s-\frac{1}{p}} \left|P_{T',r,s,p}\hat{v} - \hat{v}\right|_{W^{s,p}(T')}.$$

This completes the proof. $\square$

3.4. **Applications in Approximation Theory.** One major application of the approximation theorem is to compare the approximation quality of conforming and non-conforming finite element spaces. One the one hand, for every $\hat{v} \in W^{1,p}(\hat{\Omega})$ we obviously have

$$(25) \qquad \inf_{\hat{w}_h \in \mathcal{P}_{r,-1}(\mathcal{T})} |\hat{v} - \hat{w}_h|_{W^{1,p}(\hat{\Omega})} \leq \inf_{\hat{w}_h \in \mathcal{P}_{r,\mathrm{D}}(\mathcal{T})} |\hat{v} - \hat{w}_h|_{W^{1,p}(\hat{\Omega})}.$$

Surprisingly, a converse inequality holds when assuming $\hat{v}$ to satisfy a minor amount of additional global regularity. For every $\hat{v} \in W^{1,p}(\hat{\Omega})$, we find by Theorem 3.3 that

$$|\hat{v} - \Pi\hat{v}|_{W^{1,p}(\hat{\Omega})} \leq (1 + C_\Pi) \sum_{T \in \mathcal{T}} \left|\hat{v} - P_{T,r,1,p}\hat{v}\right|_{W^{1,p}(T)}$$

$$\leq (1 + C_\Pi) C_{d,1,r,p,\mu}^{\mathrm{I}} \sum_{T \in \mathcal{T}} \inf_{w_T \in \mathcal{P}_r(T)} |\hat{v} - w_T|_{W^{1,p}(T)}.$$

The last inequality uses the best approximation property of the local interpolator. We now have obtained a converse to Inequality (25): for every $\hat{v} \in W^{1,p}(\hat{\Omega})$ we have

$$(26) \quad \begin{aligned} &\inf_{\hat{w}_h \in \mathcal{P}_{r,\mathrm{D}}(\mathcal{T})} |\hat{v} - \hat{w}_h|_{W^{1,p}(\hat{\Omega})} \\ &\qquad\qquad \leq (1 + C_\Pi) C_{d,1,r,p,\mu}^{\mathrm{I}} \inf_{\hat{w}_h \in \mathcal{P}_{r,-1}(\mathcal{T})} |\hat{v} - \hat{w}_h|_{W^{1,p}(\hat{\Omega})}. \end{aligned}$$

Remark 3.4. *The above error estimate seems to have appeared first in* [22] *in the special case of the H^1-seminorm. A generalization to general integer Sobolev regularity and Lebesgue exponents was published later in* [5]. *The intended use in the later publication was similar to ours, namely to facilitate error estimates in the presence of variational crimes in surface finite element methods.*

Notably, the broken Bramble-Hilbert lemma seems to close gaps in earlier proofs of error estimates for isoparametric finite element methods. For example, the last inequality on p.1224 of [2] *is a Bramble-Hilbert-type estimate of a function on a triangle patch with first-order Sobolev regularity but piecewise higher regularity; the proof in that reference requires a piecewise Bramble-Hilbert lemma which was not in circulation at the point of publication.*

4. A Priori Error Estimates

In this section we attend to a thorough discussion of a priori error estimates for a finite element method of the parametric model problem and the critical role of the finite element approximation result in Theorem 3.3 in facilitating optimal convergence rates. We continue the discussion of a Galerkin method as in Section 2 with the concrete example of $\hat{V}_h = \mathcal{P}_{r,\mathrm{D}}(\mathcal{T})$ as a Galerkin space.

We consider a parametric right-hand side of the form

$$(27) \qquad \hat{F}(\hat{v}) = \int_{\hat{\Omega}} \hat{f}\hat{v}\,\mathrm{d}\hat{x} + \int_{\hat{\Omega}} \hat{\mathbf{g}} \cdot \nabla\hat{v}\,\mathrm{d}\hat{x}, \quad \hat{v} \in W_{\mathrm{D}}^{1,2}(\hat{\Omega}).$$

We let $\hat{u}$ be the unique solution of

$$\int_{\hat{\Omega}} \nabla\hat{u} \cdot \hat{A}\nabla\hat{v}\,\mathrm{d}\hat{x} = \hat{F}(\hat{v}), \quad \hat{v} \in W_{\mathrm{D}}^{1,2}(\hat{\Omega}).$$

We also let $\hat{u}_h$ be the solution of the conforming Galerkin problem

$$\int_{\hat{\Omega}} \nabla\hat{u}_h \cdot \hat{A}\nabla\hat{v}_h\,\mathrm{d}\hat{x} = \hat{F}(\hat{v}_h), \quad \hat{v}_h \in \mathcal{P}_{r,\mathrm{D}}(\mathcal{T}).$$

In practice, the parametric coefficient and right-hand side can only be approximated. Suppose that $\hat{A}_h \in L^{\infty}(\hat{\Omega})^{n \times n}$ is an approximate parametric coefficient and that we have an approximate right-hand side

$$(28) \qquad \hat{F}_h(\hat{v}_h) = \int_{\hat{\Omega}} \hat{f}_h\hat{v}_h\,\mathrm{d}\hat{x} + \int_{\hat{\Omega}} \hat{\mathbf{g}}_h \cdot \nabla\hat{v}_h\,\mathrm{d}\hat{x}, \quad \hat{v}_h \in \hat{V}_h,$$

of a form analogous to the original right-hand side. We assume that there exists $\hat{c}_{P,h} > 0$ satisfying the discrete coercivity estimate (21); see Remark 4.3 below. Practically, we compute the solution $\underline{\hat{u}}_h$ of the approximate Galerkin problem

$$\int_{\hat{\Omega}} \nabla\underline{\hat{u}}_h \cdot \hat{A}_h\nabla\hat{v}_h\,\mathrm{d}\hat{x} = \hat{F}_h(\hat{v}_h), \quad \hat{v}_h \in \mathcal{P}_{r,\mathrm{D}}(\mathcal{T}).$$

Our goal is estimate the error $\hat{u} - \underline{\hat{u}}_h$ in different norms. We now recall some standard results in the Galerkin theory of elliptic problems.

The conforming Galerkin approximation is the optimal approximation in $\hat{V}_h$ with respect to the $\hat{A}$-weighted norms of the gradient:

$$(29) \qquad \|\nabla\hat{u} - \nabla\hat{u}_h\|_{L^2(\hat{\Omega},\hat{A})} = \inf_{\hat{v}_h \in \hat{V}_h} \|\nabla\hat{u} - \nabla\hat{v}_h\|_{L^2(\hat{\Omega},\hat{A})}.$$

Céa's Lemma estimates the approximation in the full norm of $W^{1,2}(\hat{\Omega})$:

$$(30) \qquad \sqrt{\hat{c}_P}\|\hat{u} - \hat{u}_h\|_{W^{1,2}(\hat{\Omega})} \leq \inf_{\hat{v}_h \in V_h} \|\nabla\hat{u} - \nabla\hat{v}_h\|_{L^2(\hat{\Omega},\hat{A})}.$$

The error of the conforming Galerkin method in the L^2 norm is estimated by an Aubin-Nitsche-type argument, which requires the theoretical discussion of an auxiliary problem. We let $\hat{z} \in W_{\mathrm{D}}^{1,2}(\hat{\Omega})$ be the unique solution of

$$(31) \qquad \hat{B}(\hat{v}, \hat{z}) = \langle \hat{u} - \hat{u}_h, \hat{v} \rangle_{L^2(\hat{\Omega})}, \quad \hat{v} \in W_{\mathrm{D}}^{1,2}(\hat{\Omega}).$$

We then find for $\hat{z}_h \in \hat{V}_h$ arbitrary that

$$\|\hat{u} - \hat{u}_h\|_{L^2(\hat{\Omega})}^2 = \langle \hat{u} - \hat{u}_h, \hat{u} - \hat{u}_h \rangle_{L^2(\hat{\Omega})} = \hat{B}(\hat{u} - \hat{u}_h, \hat{z}) = \hat{B}(\hat{u} - \hat{u}_h, \hat{z} - \hat{z}_h),$$

and consequently

$$(32) \qquad \|\hat{u} - \hat{u}_h\|_{L^2(\hat{\Omega})}^2 \leq \|\nabla\hat{u} - \nabla\hat{u}_h\|_{L^2(\hat{\Omega},A)}\|\nabla\hat{z} - \nabla\hat{z}_h\|_{L^2(\hat{\Omega},A)}.$$

Hence we expect the L^2 error of the conforming Galerkin method to converge generally faster than the error in the $W^{1,2}$ norm, an intuition made rigorous whenever estimates for $\nabla\hat{z} - \nabla\hat{z}_h$ are available.

The corresponding error estimates for the non-conforming Galerkin approximation reduce to estimates for $\hat{u}_h - \hat{\underline{u}}_h$, that is, we compare the conforming with the non-conforming approximation. The triangle inequality in conjunction with (10) gives

$$(33) \qquad \|\nabla\hat{u} - \nabla\hat{\underline{u}}_h\|_{L^2(\hat{\Omega})} \leq \|\nabla\hat{u} - \nabla\hat{u}_h\|_{L^2(\hat{\Omega})} + \hat{c}_P^{-\frac{1}{2}}\|\nabla\hat{u}_h - \nabla\hat{\underline{u}}_h\|_{L^2(\hat{\Omega},\hat{A})},$$

$$(34) \qquad \|\hat{u} - \hat{\underline{u}}_h\|_{L^2(\hat{\Omega})} \leq \|\hat{u} - \hat{u}_h\|_{L^2(\hat{\Omega})} + \hat{c}_P^{-\frac{1}{2}}\|\nabla\hat{u}_h - \nabla\hat{\underline{u}}_h\|_{L^2(\hat{\Omega},\hat{A})}.$$

We use definitions to get

$$\|\nabla\hat{u}_h - \nabla\hat{\underline{u}}_h\|^2_{L^2(\hat{\Omega},\hat{A})}$$
$$= \hat{B}(\hat{u}_h - \hat{\underline{u}}_h, \hat{u}_h - \hat{\underline{u}}_h)$$
$$= \hat{F}(\hat{u}_h - \hat{\underline{u}}_h) - \hat{B}(\hat{\underline{u}}_h, \hat{u}_h - \hat{\underline{u}}_h)$$
$$= \hat{F}(\hat{u}_h - \hat{\underline{u}}_h) - \hat{F}_h(\hat{u}_h - \hat{\underline{u}}_h) + \hat{B}_h(\hat{\underline{u}}_h, \hat{u}_h - \hat{\underline{u}}_h) - \hat{B}(\hat{\underline{u}}_h, \hat{u}_h - \hat{\underline{u}}_h).$$

Consequently, we derive

$$\|\hat{u}_h - \hat{\underline{u}}_h\|_{W^{1,2}(\hat{\Omega},\hat{A})}$$

$$(35) \qquad \leq \sup_{\hat{w}_h \in \mathcal{P}_{r,\mathrm{D}}(\mathcal{T})} \frac{(\hat{F} - \hat{F}_h)(\hat{w}_h) + \hat{B}_h(\hat{\underline{u}}_h, \hat{w}_h) - \hat{B}(\hat{\underline{u}}_h, \hat{w}_h)}{\|\hat{w}_h\|_{W^{1,2}(\hat{\Omega},\hat{A})}}.$$

Under natural regularity assumptions on the coefficients, (35) leads to optimal error estimates. In particular, in the special case that the conforming and non-conforming methods coincide, (30) is recovered.

We need further tools to derive actual convergence rates from these abstract estimates. We recall a general polynomial approximation estimate that follows from [9, Theorem 3.1, Proposition 6.1] and a scaling argument.

Lemma 4.1. *Let $T \in \mathcal{T}$ be a d-simplex, $p \in [1,\infty]$, and $r \in \mathbb{N}_0$. Let $s \in \mathbb{R}$ with $r < s \leq r+1$. Then there exists $C^{\mathrm{BH}}_{d,\mu,s} > 0$ that depends only on d, s, and $\mu(\mathcal{T})$ such that for all $m \in \mathbb{N}_0$ with $m \leq r$ we get*

$$(36) \qquad \inf_{\hat{v}_h \in \mathcal{P}_r(T)} |\hat{v} - \hat{v}_h|_{W^{m,p}(T)} \leq C^{\mathrm{BH}}_{d,\mu,s} h_T^{s-m} |\hat{v}|_{W^{s,p}(T)}, \quad \hat{v} \in W^{s,p}(T).$$

We use the polynomial approximation result to derive quantitative error estimates in terms of the mesh size. We write $\underline{u}_h := \Phi^{-*}\hat{\underline{u}}_h$ for the transformation of the non-conforming Galerkin solution onto the physical domain. It is then easily verified that

$$\|\hat{u} - \hat{\underline{u}}_h\|_{L^2(\hat{\Omega})} \leq \|\det \mathrm{D}\Phi^{-1}\|_{L^\infty(\hat{\Omega})}\|\hat{u} - \hat{\underline{u}}_h\|_{L^2(\hat{\Omega})},$$

$$\|\nabla\hat{u} - \nabla\hat{\underline{u}}_h\|_{L^2(\hat{\Omega})} \leq \|\det \mathrm{D}\Phi^{-1}\|_{L^\infty(\hat{\Omega})}\|\mathrm{D}\Phi^{-1}\|_{L^\infty(\hat{\Omega})}\|\nabla\hat{u} - \nabla\hat{\underline{u}}_h\|_{L^2(\hat{\Omega})}.$$

Due to (33) and (34), it remains to analyze the error $\hat{u} - \hat{u}_h$ of the conforming Galerkin method (over the parametric domain), and the non-conformity error terms in (35).

The physical setting captures the relevant regularity features of the problem. Local regularity features of the physical solution translate to local regularity features of the parametric solution because the transformation Φ is piecewise smooth by assumption. Given $s \in \mathbb{R}_0^+$, there exists $C^\Phi_{n,s} > 0$, bounded in terms of n, s,

$\|\Phi\|_{W^{l+1,\infty}(\hat{\Omega})}$ and $\|\Phi^{-1}\|_{W^{l+1,\infty}(\hat{\Omega})}$, such that whenever we have $\hat{u} \in W^{s,2}(\hat{\Omega})$, we get on every n-simplex $T \in \mathcal{T}$

$$|\hat{u}|_{W^{s,2}(T)} \le C_{n,s}^{\Phi}\|\hat{u}\|_{W^{s,2}(\Phi(T))}.$$

So suppose that $\hat{u} \in W^{s,2}(\Phi(T))$ for some $s \ge 1$, and let $r \in \mathbb{N}$ the largest integer with $r < s$. In conjunction with the Galerkin optimality, Theorem 3.3, and Lemma 4.1, we then obtain the estimate

$$
\begin{aligned}
|\hat{u} - \hat{u}_h|_{W^{1,2}(\hat{\Omega})} &= \inf_{\hat{v}_h \in \mathcal{P}_{r,\mathrm{D}}(\mathcal{T})} |\hat{u} - \hat{v}_h|_{W^{1,2}(\hat{\Omega})} \\
&= (1 + C_\Pi) \sum_{T \in \mathcal{T}} \left|\hat{v} - P_{T,r,1}\hat{v}\right|_{W^{1,2}(T)} \\
&\le (1 + C_\Pi) C_{d,1,r,2,\mu}^{\mathrm{I}} C_{d,\mu,s}^{\mathrm{BH}} \sum_{T \in \mathcal{T}} h_T^{s-1} |\hat{u}|_{W^{s,2}(T)}.
\end{aligned}
$$

Recall the definition of $\hat{z}$ as the solution of the parametric model problem with right-hand side $\hat{u} - \hat{u}_h$. It then follows that $\check{z} := \Phi^{-*}\hat{z}$ is the solution of the physical model problem with right-hand side $\Phi^{-*}(\hat{u} - \hat{u}_h)$. Consequently, whenever $\hat{z} \in W^{t,2}(\hat{\Omega})$ for some $t \in \mathbb{R}_0^+$, then for all n-simplices $T \in \mathcal{T}$ it follows that $|\hat{u}|_{W^{t,2}(T)} \le C_{n,t}^{\Phi}\|\hat{u}\|_{W^{t,2}(\Phi(T))}$. Similar to the arguments used above, we then see

$$(37) \qquad |\hat{z} - \hat{z}_h|_{W^{1,2}(\hat{\Omega})} \le (1 + C_\Pi) C_{d,1,r,\mu}^{\mathrm{I}} C_{d,\mu,t}^{\mathrm{BH}} \sum_{T \in \mathcal{T}} h_T^{t-1} |\hat{z}|_{W^{t,2}(T)}.$$

This provides convergence rates for the conforming Galerkin method.

For error estimates for the non-conforming problem, it remains to analyze the difference $\nabla\hat{u}_h - \nabla\underline{\hat{u}}_h$ using (35). Via (10) and (22) we get

$$
\begin{aligned}
&\|\hat{u}_h - \underline{\hat{u}}_h\|_{W^{1,2}(\hat{\Omega},\hat{A})} \\
&\qquad \le \frac{\hat{c}_{P,h}}{\hat{c}_P}\left(\|\hat{F} - \hat{F}_h\|_{W_{\mathrm{N}}^{-1,2}(\hat{\Omega})} + \|\hat{A} - \hat{A}_h\|_{L^\infty(\hat{\Omega})}|\underline{\hat{u}}_h|_{W^{1,2}(\hat{\Omega})}\right).
\end{aligned}
$$

Note that

$$\hat{c}_{P,h}|\underline{\hat{u}}_h|_{W^{1,2}(\hat{\Omega})} \le \|\hat{F}_h\|_{W_{\mathrm{N}}^{-1,2}(\hat{\Omega})} \le \|\hat{f}_h\|_{L^2(\hat{\Omega})} + \|\hat{\mathbf{g}}_h\|_{L^2(\hat{\Omega})}$$

As for the approximate right-hand side F_h as in (28), we get

$$\|\hat{F} - \hat{F}_h\|_{W_{\mathrm{N}}^{-1,2}(\hat{\Omega})} \le \|\hat{f} - \hat{f}_h\|_{L^2(\hat{\Omega})} + \|\hat{\mathbf{g}} - \hat{\mathbf{g}}_h\|_{L^2(\hat{\Omega})}.$$

Consequently, it remains to bound the errors of the approximate right-hand side and the approximate coefficient. Let us make the regularity assumptions

$$\hat{A} \in W^{l,\infty}(\hat{\Omega}), \quad \hat{f} \in W^{l,2}(\hat{\Omega}), \quad \hat{\mathbf{g}} \in W^{l,2}(\hat{\Omega}).$$

For each n-simplex $T \in \mathcal{T}$ it then follows that

$$(38) \qquad |\hat{f}|_{W^{l,2}(T)} \le C_{n,l}^{\Phi}\|\hat{f}\|_{W^{l,2}(\Phi(T))},$$

$$(39) \qquad |\hat{\mathbf{g}}|_{W^{l,2}(T)} \le C_{n,l}^{\Phi}\|\hat{\mathbf{g}}\|_{W^{l,2}(\Phi(T))},$$

$$(40) \qquad |\hat{A}|_{W^{l,\infty}(T)} \le C_{n,l}^{\Phi}\|\hat{A}\|_{W^{l,\infty}(\Phi(T))}.$$

In other words, the parametric data and coefficients inherit the piecewise regularity of their physical counterparts. Let $k \in \mathbb{N}_0$ be the largest integer with $k < l$; we use k as the polynomial degree of the data approximation. Suppose that we specifically

choose the approximate parametric coefficient $\hat{A}_h$ in each component as the piece-wise best L^2 approximation of $\hat{A}$ by polynomials of degree at most k. Similarly, suppose we choose $\hat{f}_h$ and $\hat{\mathbf{g}}_h$ as piecewise polynomial best L^2 approximations of degree at most k. Then Lemma 4.1 yields

$$\|\hat{A} - \hat{A}_h\|_{L^\infty(T)} \le C^{\mathrm{BH}}_{d,\mu,l} h_T^l |\hat{A}|_{W^{l,\infty}(T)},$$

$$\|\hat{f} - \hat{f}_h\|_{L^2(T)} \le C^{\mathrm{BH}}_{d,\mu,l} h_T^l |\hat{f}|_{W^{l,2}(T)},$$

$$\|\hat{\mathbf{g}} - \hat{\mathbf{g}}_h\|_{L^2(T)} \le C^{\mathrm{BH}}_{d,\mu,l} h_T^l |\hat{\mathbf{g}}|_{W^{l,2}(T)}$$

for each n-simplex $T \in \mathcal{T}$. This completes the desired estimates.

Remark 4.2. *We compare our approach to a classical finite element method for curved domains, namely isoparametric finite element methods. The latter assume an affine mesh $\mathcal{T}_h$ of a polyhedral parametric domain $\hat{\Omega}_h$, and a piecewise polynomial coordinate transformation Φ_h whose image is a curved polyhedral domain that "approximates" (in whatever sense) the physical domain. The finite element method over the image of Φ can be pulled back to a finite element method over the polyhedral parametric domain $\hat{\Omega}_h$. Since each Φ_h is piecewise polynomial, the pullback introduces only polynomial terms that can be evaluated exactly. We may thus interpret the approximate isoparametric meshes as a tool to develop approximate coefficients on the underlying affine mesh.*

Isoparametric finite element methods use an approximate coordinate transformation whose coefficients can be determined exactly. Our approach in this article is strictly different in that we use an exact coordinate transformation whose coefficients are then approximated. It seems that both procedures lead to similar results in practice.

One of the first error estimates for the isoparametric finite element method is due to Ciarlet and Raviart [6], who transferred the canonical interpolant to the isoparametric setting. The idea is to take the nodal values of the canonical interpolant within the physical domain. This idea can be replicated in our setting without essential difficulty. However, a severe restriction is the requirement of higher regularity for the solution. A more refined error analysis for the case of conforming geometry was established by Clement [7], Scott and Zhang [20], and recently by Ern and Guermond [11] Lenoir's contribution [15] focuses on the construction of curved triangulations for a given domain and gives error estimates again only for the case that the physical solution is continuous (see Lemma 7 in the reference).

To our best knowledge, however, it is only with Inequality (26) that rigorous a priori error estimates for higher order isoparametric FEM are available. Furthermore, our method is more flexible whenever the (possibly non-polynomial) coordinate transformation is actually known.

Remark 4.3. *The discrete coercivity condition (21) holds for sufficiently fine approximation of the parametric coefficient. The feasibility of quasi-optimal positivity preserving interpolation is not studied in this article; we refer to the literature for results on positivity preserving interpolation of functions [17].*

Remark 4.4. *Let $r \in \mathbb{N}$ and $k \in \mathbb{N}_0$. Suppose that $\hat{u} \in W^{r+1,2}(\hat{\Omega})$ and that $\hat{A}$, $\hat{f}$, and $\hat{\mathbf{g}}$ have regularity $W^{k+1,2}(\hat{\Omega})$. Let us also assume that full elliptic regularity holds for the physical model problem, so that $\hat{z} \in W^{2,2}(\hat{\Omega})$. Letting $C > 0$ denote a generic constant and letting $h > 0$ denote the maximum diameter of a simplex in*

the triangulation, we have

$$|\hat{u} - \underline{\hat{u}}_h|_{W^{1,2}(\hat{\Omega})} \leq Ch^r \|\hat{u}\|_{W^{r+1,2}(\hat{\Omega})} + Ch^{k+1} \|\hat{u}\|_{W^{1,2}(\hat{\Omega})},$$

$$\|\hat{u} - \underline{\hat{u}}_h\|_{L^2(\hat{\Omega})} \leq Ch^{r+1} \|\hat{u}\|_{W^{r+1,2}(\hat{\Omega})} + Ch^{k+1} \|\hat{u}\|_{W^{1,2}(\hat{\Omega})}.$$

Analogous estimates are known for surface finite element methods, where k signifies the degree of geometric approximation. Our a priori error estimate is the sum of a classical Galerkin approximation error and additional error terms due to coefficient and data approximation. This is reminiscent of what has been called "almost best-approximation error" and "geometric error" in the analysis of surface finite element methods [8, p.2].

5. Examples and Computational Experiments

This articles finishes with a few examples of coordinate transformations from polyhedral parametric domains onto curved physical domains and illustrative numerical computations within those geometric setups.

We have conducted our numerical experiments with the `FEniCS` Library. We list the observed errors for polynomial order $1 \leq r \leq 4$ and several levels of uniform refinement. The polynomial degree of the interpolated coefficients and data has always been equal to the degree of the finite element space. All linear system of equations have been solved with the conjugate gradient method and `FEniCS`'s built-in `amg` preconditioner, the absolute and relative error tolerance set to `1e-20` each. The theoretically predicted convergence rates are achieved in practice.

5.1. **Anulus.** Models in geophysics and climate science assume static homogeneous conditions over a large interior part of the Earth and pose partial differential equations only over a thin outer volume of the planet. The equations reign over an n-dimensional anulus, that is, an n-ball with an internal n-ball removed.

We introduce the parametric anulus $\hat{\mathcal{A}}$ and the physical anulus $\hat{\mathcal{A}}$ by

$$\hat{\mathcal{A}} := \left\{ \hat{x} \in \mathbb{R}^n \;\middle|\; \frac{1}{2} < \|\hat{x}\|_1 < 1 \right\}, \quad \hat{\mathcal{A}} := \left\{ \hat{x} \in \mathbb{R}^n \;\middle|\; \frac{1}{2} < \|\hat{x}\|_2 < 1 \right\}.$$

We consider a coordinate transformations from the parametric anulus onto the physical anulus,

$$\Phi : \hat{\mathcal{A}} \to \hat{\mathcal{A}}, \quad \hat{x} = \|\hat{x}\|_1 \|\hat{x}\|_2^{-1} \hat{x},$$

which is easily seen to be invertible and bi-Lipschitz. Furthermore, over the intersections of $\hat{\mathcal{A}}$ with any of the 2^n coordinate quadrants, the transformation Φ is a diffeomorphism with derivatives of all orders pointwise bounded. In particular, it is easy to find a (coarse) initial triangulation of $\hat{\mathcal{A}}$ such that the coordinate transformation Φ is piecewise smooth.

This construction allows us to transport partial differential equations over an *Euclidean anulus* $\hat{\mathcal{A}}$ to partial differential equations over the polyhedral *Manhattan-metric anulus* $\hat{\mathcal{A}}$. Suppose that we want to solve the Poisson problem

$$(41) \qquad \int_{\hat{\mathcal{A}}} \nabla \hat{u} \cdot \nabla \hat{v} \, \mathrm{d}\hat{x} = \int_{\hat{\mathcal{A}}} \hat{f} \cdot \hat{v} \, \mathrm{d}\hat{x} + \int_{\hat{\mathcal{A}}} \hat{\mathbf{g}} \cdot \nabla \hat{v} \, \mathrm{d}\hat{x}, \quad \hat{v} \in H_0^1(\hat{\mathcal{A}}),$$
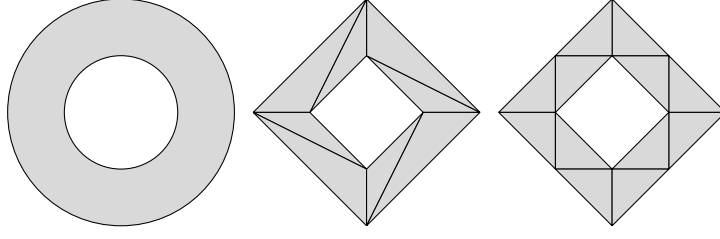
FIGURE 1. From left to right: physical anulus $\hat{\mathcal{A}}$, and parametric anulus $\hat{\mathcal{A}}$ with two possible triangulations for which Φ is a piecewise diffeomorphism. Our computations use the first triangulation.

over the Euclidean anulus. Along the transformation Φ we can translate this into an equivalent Poisson problem over the parametric anulus $\hat{\mathcal{A}}$ of the form

$$
(42) \quad \int_{\hat{\mathcal{A}}} |\det \mathrm{D}\Phi|\, \nabla \hat{u} \cdot \mathrm{D}\Phi_{|\Phi}^{-1}\, \mathrm{D}\Phi_{|\Phi}^{-t} \nabla \hat{v}\, \mathrm{d}\hat{x}
$$
$$
= \int_{\hat{\mathcal{A}}} |\det \mathrm{D}\Phi|\, (\hat{f} \circ \Phi) \hat{v}\, \mathrm{d}\hat{x} + \int_{\hat{\Omega}} |\det \mathrm{D}\Phi| \left(\mathrm{D}\Phi_{|\Phi}^{-1} (\hat{\mathbf{g}} \circ \Phi) \right) \nabla \hat{v}\, \mathrm{d}\hat{x}
$$

for any test function $\hat{v} \in H_0^1(\hat{\mathcal{A}})$. Thus the parametric problem can be solved with textbook methods. However, it is a stronger result and a consequence of Theorem 3.3 that the piecewise regularity of the parametric solution $\hat{u}$, which is inherited from the physical solution $\hat{u}$ on each cell, leads to the same convergence rates that the corresponding global regularity of $\hat{u}$ would suggest.

Remark 5.1. *For illustration, consider the radially symmetric Dirichlet problem*

$$
(43) \quad -\Delta \hat{u} = 1, \quad \hat{u}_{|\partial \hat{\mathcal{A}}} = 0,
$$

over the physical anulus $\hat{\mathcal{A}}$ with $n = 2$. The solution is the function

$$
(44) \quad \hat{u}(x, y) = \frac{1}{4} + \frac{3 \ln(x^2 + y^2)}{32 \ln(2)} - \frac{x^2 + y^2}{4}.
$$

The results of the computational experiments (for the corresponding parametric) problems are summarized in Table 1. We have used the first triangulation in Figure 1 The expected convergence behavior is clearly visible for all polynomial orders.

5.2. **Quadrant of Unit Ball.** Our second example geometry considers the positive quadrant of the Euclidean unit ball. We transport differential equations over that domain to differential equations over the positive quadrant of the Manhatten unit ball. The homeomorphism is the identity near the origin. We concretely define

$$
\hat{\mathcal{B}} := \left\{ \hat{x} \in (\mathbb{R}_0^+)^n \,\middle|\, \|\hat{x}\|_1 < 1 \right\}, \quad \hat{\mathcal{B}} := \left\{ \hat{x} \in (\mathbb{R}_0^+)^n \,\middle|\, \|\hat{x}\|_2 < 1 \right\}.
$$

Consider the transformation

$$
\Psi : \hat{\mathcal{B}} \to \hat{\mathcal{B}}, \quad \hat{x} \mapsto \begin{cases} \hat{x} & \text{if } \|\hat{x}\|_1 \le \frac{1}{2}, \\ \left(\|\hat{x}\|_1^{-1} - \|\hat{x}\|_2^{-1} + 2\frac{\|\hat{x}\|_1}{\|\hat{x}\|_2} - 1 \right) \hat{x} & \text{if } \frac{1}{2} < \|\hat{x}\|_1 \le 1. \end{cases}
$$

This mapping is the identity over the set of points with Manhatten distance at most $\frac{1}{2}$ from the origin. It is easily verified that both Ψ and Ψ^{-1} are bi-Lipschitz. If a
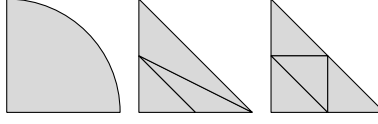
FIGURE 2. From left to right: physical domain $\hat{\mathcal{B}}$, and parametric domain $\hat{\mathcal{B}}$ with two possible triangulations for which Ψ is a piecewise diffeomorphism. Our computations use the first triangulation.

TABLE 1. Convergence table for example problem over the anulus.

(A) H^1 seminorm of error e and convergence rate ρ

	$r = 1$		$r = 2$		$r = 3$		$r = 4$	
L	$|e|_1$	ρ	$|e|_1$	ρ	$|e|_1$	ρ	$|e|_1$	ρ
0	0.25098	–	0.022662	–	0.0092381	–	0.0044456	–
1	0.17023	0.56	0.011661	0.95	0.0019261	2.26	0.00038776	3.51
2	0.12360	0.46	0.0059485	0.97	6.1275e-04	1.65	7.0433e-05	2.46
3	0.06808	0.86	0.0016833	1.82	8.9297e-05	2.77	5.5085e-06	3.67
4	0.035349	0.94	4.3925e-04	1.93	1.1607e-05	2.94	3.7089e-07	3.89
5	0.017975	0.97	1.1169e-04	1.97	1.4614e-06	2.98	2.3773e-08	3.96
6	0.0090598	0.98	2.8134e-05	1.98	1.8270e-07	2.99	1.5005e-09	3.98

(B) Convergence in L^2 norm and convergence rate ρ

	$r = 1$		$r = 2$		$r = 3$		$r = 4$	
L	$\|e\|_2$	ρ	$\|e\|_2$	ρ	$\|e\|_2$	ρ	$\|e\|_2$	ρ
0	0.028212	–	0.0014496	–	6.6475e-04	–	2.3259e-04	–
1	0.015748	0.84	6.4554e-04	1.16	9.1570e-05	2.85	1.5631e-05	3.89
2	0.0089178	0.82	2.0229e-04	1.67	1.4589e-05	2.64	1.4000e-06	3.48
3	0.0026699	1.73	2.5682e-05	2.97	9.7970e-07	3.89	5.1223e-08	4.77
4	7.0914e-04	1.91	3.1727e-06	3.01	5.9079e-08	4.05	1.7106e-09	4.90
5	1.8158e-04	1.96	3.9619e-07	3.00	3.5422e-09	4.05	5.4605e-11	4.96
6	4.5877e-05	1.98	4.9728e-08	2.99	2.1552e-10	4.03	1.7219e-12	4.98

triangulation of $\hat{\mathcal{B}}$ accommodates the case distinction in the definition, then both Ψ and Ψ^{-1} have bounded derivatives of all orders over each cell.

Remark 5.2. *For the purpose of demonstration, we let $n = 2$ and solve the Poisson problem over $\hat{\mathcal{B}}$ with homogeneous Dirichlet boundary conditions:*

$$(45) \qquad -\Delta \hat{u} = 24x^2y^2 - 2(x^2 + y^2) + 2(x^4 + y^4), \quad \hat{u}_{|\partial \hat{\mathcal{B}}} = 0.$$

The solution is the polynomial

$$(46) \qquad \hat{u} = x^2 y^2 (1 - x^2 - y^2).$$

Though the function $\hat{u}$ is even a polynomial, its parametric counterpart $\hat{u}$ is not. The results of the computational experiments (for the corresponding parametric problem) are summarized in Table 2. We have used the first triangulation in Figure 2 The theoretically predicted convergence behavior emerges.

TABLE 2. Convergence table for example problem over the positive ball quadrant.

(A) H^1 seminorm of error e and convergence rate ρ

	$r = 1$		$r = 2$		$r = 3$		$r = 4$	
L	$|e|_1$	ρ	$|e|_1$	ρ	$|e|_1$	ρ	$|e|_1$	ρ
0	0.10633	–	0.11316	–	0.054669	–	0.036351	–
1	0.10436	0.02	0.060710	0.89	0.020062	1.44	0.0051925	2.80
2	0.083847	0.31	0.029172	1.05	0.0055015	1.86	6.5194e-04	2.99
3	0.052386	0.67	0.0093838	1.63	8.6523e-04	2.66	5.1075e-05	3.67
4	0.030721	0.76	0.0027669	1.76	1.2138e-04	2.83	3.4510e-06	3.88
5	0.016719	0.87	7.5126e-04	1.88	1.6042e-05	2.91	2.2429e-07	3.94
6	0.0087313	0.93	1.9570e-04	1.94	2.0609e-06	2.96	1.4286e-08	3.97

(B) Convergence in L^2 norm and convergence rate ρ

	$r = 1$		$r = 2$		$r = 3$		$r = 4$	
L	$\|e\|_2$	ρ	$\|e\|_2$	ρ	$\|e\|_2$	ρ	$\|e\|_2$	ρ
0	0.0088039	–	0.0066160	–	0.0034871	–	0.0017926	–
1	0.0067425	0.38	0.0032217	1.03	6.9095e-04	2.33	1.5377e-04	3.54
2	0.0044515	0.59	8.9753e-04	1.84	1.148e-04	2.58	1.1642e-05	3.72
3	0.0017281	1.36	1.3976e-04	2.68	9.2456e-06	3.63	4.6968e-07	4.63
4	5.7065e-04	1.59	2.0850e-05	2.74	6.2609e-07	3.88	1.6004e-08	4.87
5	1.6469e-04	1.79	2.8474e-06	2.87	4.0472e-08	3.95	5.2234e-10	4.93
6	4.4342e-05	1.89	3.7210e-07	2.93	2.5686e-09	3.97	1.6668e-11	4.96

REFERENCES

[1] M. S. ALNÆS, J. BLECHTA, J. HAKE, A. JOHANSSON, B. KEHLET, A. LOGG, C. RICHARDSON, J. RING, M. E. ROGNES, AND G. N. WELLS, *The FEniCS Project Version 1.5*, Archive of Numerical Software, 3 (2015).

[2] C. BERNARDI, *Optimal finite-element interpolation on curved domains*, SIAM Journal on Numerical Analysis, 26 (1989), pp. 1212–1240.

[3] 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, Springer, 2013, pp. 257–306.

[4] D. BRAESS, *Finite Elements: Theory, Fast Solvers, and Applications in Solid Mechanics*, Cambridge University Press, 2007.

[5] F. CAMACHO AND A. DEMLOW, L^2 *and pointwise a posteriori error estimates for FEM for elliptic PDEs on surfaces*, IMA Journal of Numerical Analysis, 35 (2014), pp. 1199–1227.

[6] P. G. CIARLET AND P.-A. RAVIART, *Interpolation theory over curved elements, with applications to finite element methods*, Computer Methods in Applied Mechanics and Engineering, 1 (1972), pp. 217–249.

[7] P. CLÉMENT, *Approximation by finite element functions using local regularization*, Revue française d'automatique, informatique, recherche opérationnelle. Analyse numérique, 9 (1975), pp. 77–84.

[8] A. DEMLOW, *Higher-order finite element methods and pointwise error estimates for elliptic problems on surfaces*, SIAM Journal on Numerical Analysis, 47 (2009), pp. 805–827.

[9] T. DUPONT AND R. SCOTT, *Polynomial approximation of functions in Sobolev spaces*, Mathematics of Computation, 34 (1980), pp. 441–463.

[10] G. Dziuk and C. M. Elliott, *Finite elements on evolving surfaces*, IMA Journal of Numerical Analysis, 27 (2007), pp. 262–292.

[11] A. Ern and J.-L. Guermond, *Finite element quasi-interpolation and best approximation*, ESAIM: Mathematical Modelling and Numerical Analysis, 51 (2017), pp. 1367–1385.

[12] E. Gawlik and M. Holst, *Finite element exterior calculus for parabolic problems on evolving surfaces*. Preprint.

[13] E. Gawlik, M. Holst, and M. Licht, *A Scott-Zhang interpolant and piecewise Bramble-Hilbert lemma for finite element exterior calculus*. Preprint.

[14] T. J. Hughes, J. A. Cottrell, and Y. Bazilevs, *Isogeometric analysis: CAD, finite elements, NURBS, exact geometry and mesh refinement*, Computer methods in applied mechanics and engineering, 194 (2005), pp. 4135–4195.

[15] M. Lenoir, *Optimal isoparametric finite elements and error estimates for domains involving curved boundaries*, SIAM Journal on Numerical Analysis, 23 (1986), pp. 562–580.

[16] M. W. Licht, *On the A Priori and A Posteriori Error Analysis in Finite Element Exterior Calculus*, PhD thesis, Dissertation, Department of Mathematics, University of Oslo, Norway, 2017.

[17] R. Nochetto and L. Wahlbin, *Positivity preserving finite element approximation*, Mathematics of computation, 71 (2002), pp. 1405–1419.

[18] M. Rančić, R. Purser, and F. Mesinger, *A global shallow-water model using an expanded spherical cube: Gnomonic versus conformal coordinates*, Quarterly Journal of the Royal Meteorological Society, 122 (1996), pp. 959–982.

[19] C. Ronchi, R. Iacono, and P. S. Paolucci, *The "cubed sphere": a new method for the solution of partial differential equations in spherical geometry*, Journal of Computational Physics, 124 (1996), pp. 93–114.

[20] L. R. Scott and S. Zhang, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Mathematics of Computation, 54 (1990), pp. 483–493.

[21] R. Sevilla, S. Fernández-Méndez, and A. Huerta, *NURBS-enhanced finite element method (NEFEM)*, International Journal for Numerical Methods in Engineering, 76 (2008), pp. 56–83.

[22] A. Veeser, *Approximating gradients with continuous piecewise polynomial functions*, Foundations of Computational Mathematics, 16 (2016), pp. 723–750.

[23] P. Zulian, T. Schneider, K. Hormann, and R. Krause, *Parametric finite elements with bijective mappings*, BIT Numerical Mathematics, 57 (2017), pp. 1185–1203.

UCSD Department of Mathematics, 9500 Gilman Drive MC0112, La Jolla, CA 92093-0112

E-mail address: mholst@ucsd.edu

UCSD Department of Mathematics, 9500 Gilman Drive MC0112, La Jolla, CA 92093-0112

E-mail address: mlicht@ucsd.edu