

## CONVERGENCE ANALYSIS OF VARIATIONAL AND NON-VARIATIONAL MULTIGRID ALGORITHMS FOR THE LAPLACE-BELTRAMI OPERATOR

ANDREA BONITO AND JOSEPH E. PASCIAK

**ABSTRACT.** We design and analyze variational and non-variational multigrid algorithms for the Laplace-Beltrami operator on a smooth and closed surface. In both cases, a uniform convergence for the  $V$ -cycle algorithm is obtained provided the surface geometry is captured well enough by the coarsest grid. The main argument hinges on a perturbation analysis from an auxiliary variational algorithm defined directly on the smooth surface. In addition, the vanishing mean value constraint is imposed on each level, thereby avoiding singular quadratic forms without adding additional computational cost.

Numerical results supporting our analysis are reported. In particular, the algorithms perform well even when applied to surfaces with a large aspect ratio.

### 1. INTRODUCTION

Geometric differential equations have received increasing interest in recent years. They play a crucial role in many applications with impact in bio-technologies [10, 19], image processing [2, 14] and fluid flows [4] for instance. These applications all have in common the presence of the Laplace-Beltrami operator [29] (or surface Laplacian). The latter is also a fundamental tool for preserving geometric consistency when performing mesh modification such as refinement or smoothing on surfaces [5].

We consider variational and non-variational multigrid algorithms for the Laplace-Beltrami problem (written in weak form)

$$(1) \quad \tilde{u} \in H^1(\gamma) : a(\tilde{u}, \tilde{v}) := \int_{\gamma} \nabla_{\gamma} \tilde{u} \cdot \nabla_{\gamma} \tilde{v} = \langle \tilde{f}, \tilde{v} \rangle, \quad \forall \tilde{v} \in H^1(\gamma),$$

Here  $\gamma$  is a given smooth closed hyper-surface,  $\nabla_{\gamma}$  denotes the tangential gradient,  $\langle \cdot, \cdot \rangle$  is the  $H^{-1} - H^1$  dual product and  $\tilde{f} \in H^{-1}(\gamma)$  has vanishing mean value, i.e.  $\langle \tilde{f}, 1 \rangle = 0$ . Precise conditions on  $\gamma$  are provided below. The existence and uniqueness of a solution  $\tilde{u}$  to (1) is guaranteed by the Lax-Milgram Lemma provided that the additional condition  $\int_{\gamma} \tilde{u} = 0$  is prescribed.

In this paper, we consider multigrid algorithms applied to the solution of the linear systems which result from finite element approximation of (1). Our goal will be to develop and analyze certain natural multigrid algorithms. In all cases, we start with a sequence of triangular approximations to  $\gamma$  which are developed in a

---

Received by the editor August 23, 2009 and, in revised form, March 23, 2011.  
2010 *Mathematics Subject Classification.* Primary 65N30, 65N55.

multilevel fashion, i.e., triangles on a given level are partitioned into a finer mesh and the new nodes (no longer on  $\gamma$ ) are moved to  $\gamma$  using a “lift” function. This results in a sequence of approximate surfaces  $\Gamma_k$ ,  $k = 1, \dots, J$ . For approximation, we use piecewise linear (continuous) finite elements with respect to the triangles making up  $\Gamma_k$ ,  $k = 1, \dots, J$  and use quadratic forms coming from the Laplace-Beltrami operator on the approximate surfaces. This results in a non-variational multigrid algorithm. The multilevel structure also gives rise to a sequence of nested subspaces on  $\Gamma_J$  and we also consider the corresponding variational algorithm (using the Laplace-Beltrami operator on  $\Gamma_J$  to define the quadratic forms on every subspace). We note that this variational algorithm is closely related to the hierarchical algorithm studied in [21] as the coarse to fine interpolation operators in both algorithms can be thought of as a lifting to the surface approximation  $\Gamma_J$  of the standard interpolation operators on refinements of a reference polygonal triangulation.

The variational multigrid algorithms were first proposed and implemented by Holst [20]. A related hierarchical basis preconditioner for finite element approximations of (1) was studied by Kornhuber and Yserentant [21]. Also, a BPX-preconditioner was studied in [23] for the specific case of the sphere with non-local approximation spaces.

Kornhuber and Yserentant studied a hierarchical basis algorithm for these types of problems. Specifically, they proved polylogarithmic convergence rates (typical of hierarchical basis preconditioning methods) under the assumption that the triangles in the sequence of approximating manifolds  $\Gamma_k$  were quasi-uniform (see, (6) below). One of the results of our paper (Lemma 2.1) shows that mesh quasi-uniformity holds for piecewise smooth maps provided that the coarse grid is sufficiently fine and thus verifies the hypothesis needed for the application of the results of [21]. Some sort of condition on the coarse grid seems necessary for a general theorem. This is illustrated by Remark 2.2 which shows that even lifts which are smooth diffeomorphisms may fail to lead to mesh quasi-uniformity and (6)<sup>1</sup>.

Other than the above-mentioned papers, there has not been much work on iterative solvers for discretizations of (1). In contrast, the finite element approximation properties for this problem have been extensively developed; see e.g. [15, 16, 13, 12, 25, 26, 27].

Our goal will be to develop and analyze the variational and variational multigrid algorithms. In both cases, we show that the V-cycle algorithm converges uniformly (independent of  $J$ ) provided that the coarsest grid approximates  $\gamma$  sufficiently well. The non-variational result is surprising in that it shows uniform V-cycle convergence even in cases where  $\Gamma_k$  only approximates  $\gamma$  with order  $h_k$ , a condition that is easily satisfied in practice.

For the purposes of analysis, we introduce a third multigrid algorithm living entirely on  $\gamma$ . Even though this algorithm is essentially not implementable, we use it as a starting point for our analysis. The algorithm on  $\gamma$  can be analyzed using classical multigrid techniques; see, e.g., [9]. We then analyze both of our multigrid algorithms by perturbation from the algorithm on  $\gamma$ . We note that perturbation analysis for pairs of variational algorithms has been given in [7]. The analysis there

---

<sup>1</sup>In any event, the question of how one comes up with a suitable lift for a general surface is a difficult one. Such a lift should result in a good coarse grid approximation. Note that the related problem of getting a good piecewise linear approximation to a surface using a minimal number of nodes has been well studied and has been shown to be NP-complete [1].

is based on the product representation of the multigrid error for both algorithms. The perturbation analysis given below for the non-variational algorithm is new (as far as we know) and it applies when only one of the algorithms is variational, i.e., the product representation does not hold on the other. An alternative approach for the analysis of the variational algorithm could have been based on the classical multigrid theory; however, this theory would not lead to uniform convergence of the V-cycle for the non-variational algorithm.

An additional issue that we are required to deal with is the vanishing mean value constraint. In [22], problems with a compatibility condition are thought of as singular and solutions are sought in a larger space (not satisfying the compatibility condition). Our approach seems natural and involves imposing the mean value condition on each level avoiding all singular quadratic forms. In implementation, we modify our basis functions by subtracting an appropriate constant so that they satisfy this condition as well. One might think that this is troublesome as the basis functions are technically no longer local. However, the underlying quadratic form does not see constants and the resulting stiffness matrix is still sparse leading to an efficient multigrid implementation.

Although we verify theoretical uniform convergence, it is still possible to tune the multigrid algorithms in practice. In Section 5, we provide numerical results which illustrate the convergence behavior of the multigrid algorithms on a model problem. One of the most important requirements is the generation of meshes which are as uniform as possible. Triangles with bad aspect ratios slow down the multigrid convergence. In contrast, multigrid algorithms can be made to work quite well even on domains with large aspect ratios provided that the triangles on all the grid levels remain reasonably quasi-uniform.

The remainder of the paper is organized as follows. We start by describing the notion of approximating surface and provide instrumental tools to link them with  $\gamma$  in Section 2. The descriptions of the different multigrid solvers are provided in Section 3 along with the convergence of the non-implementable one defined on the smooth surface  $\gamma$ . In Section 4, we provide a perturbation analysis leading to the convergence of the variational and non-variational solvers. The latter convergence holds provided conditions on the smoother are satisfied. Those conditions are proved to hold for the standard Jacobi and Gauss-Seidel smoothers. We end this paper by providing numerical experiments in Section 5.

## 2. PARAMETRIC SURFACES AND APPROXIMATIONS

This section is devoted to the description of the parameterized surface  $\gamma \subset \mathbb{R}^3$ , assumed to be compact and at least  $C^2$ . The construction of approximating surfaces  $\Gamma$  is provided along with crucial geometric estimates linking quantities defined over  $\gamma$  to those defined over  $\Gamma$ .

**2.1. Differential calculus operators on surfaces.** Let  $\tilde{f}$  and  $\tilde{v}$  be  $C^1$  scalar and vector-valued functions defined on  $\gamma$ . Let  $\mathcal{E}(\tilde{f})$  be an extension of  $\tilde{f}$  to a tubular neighborhood of  $\gamma$ . Its existence is guaranteed by the smoothness of  $\gamma$ . Although the following definitions of the surface operator involve  $\mathcal{E}$ , they are independent of the specific extension chosen (see, for example, [17, 28]).

We can define the tangential gradient of  $\tilde{f}$  as the orthogonal projection of  $\nabla\mathcal{E}(\tilde{f})$  to the tangent hyperplane of  $\gamma$ ,

$$\nabla_\gamma \tilde{f} := \nabla\mathcal{E}(\tilde{f})|_\gamma - \nabla\mathcal{E}(\tilde{f})|_\gamma \cdot \boldsymbol{\nu} \boldsymbol{\nu},$$

where  $\boldsymbol{\nu}$  is the outward-pointing unit normal of  $\gamma$ . The tangential gradient of  $\tilde{\mathbf{v}}$  is the unique tensor  $\nabla_\gamma \tilde{\mathbf{v}}$  such that  $(\nabla_\gamma \tilde{\mathbf{v}})^t \mathbf{a} = \nabla_\gamma(\tilde{\mathbf{v}} \cdot \mathbf{a})$  for all  $\mathbf{a}$  in  $\mathbb{R}^d$ . The tangential divergence of  $\tilde{\mathbf{v}}$  is defined by  $\operatorname{div}_\gamma \tilde{\mathbf{v}} := \operatorname{Trace}(\nabla_\gamma \tilde{\mathbf{v}})$ . Finally, for  $C^2(\gamma)$  functions or vectors, the surface Laplacian or Laplace-Beltrami operator is defined by  $\Delta_\gamma \cdot := \operatorname{div}_\gamma(\nabla_\gamma \cdot)$ .

For any smooth manifold  $\mathcal{M}$ , we denote by  $H^i(\mathcal{M})$  the standard Hilbert spaces with the convention  $L^2(\mathcal{M}) = H^0(\mathcal{M})$ ; see [18]. We denote by

$$\|\cdot\|_{i,\mathcal{M}} := \|\cdot\|_{H^i(\mathcal{M})}$$

the standard associated norms and  $|\cdot|_{i,\mathcal{M}}$  the  $H^i(\mathcal{M})$  semi-norms. Typically,  $\mathcal{M}$  will be  $\gamma$  or one of its continuous and polygonal approximations. In the case of the latter, the norm in  $H^1(\Gamma)$  is still well defined (see, e.g., [15, 17]) and given by

$$\|v\|_{1,\Gamma}^2 = \|v\|_{0,\Gamma}^2 + \sum_K \int_K |\nabla_K v|^2$$

where the sum is over the polygonal faces making up  $\Gamma$ .

**2.2. Reference polygonal surface and parameterized surface.** In order to describe the surface, we follow [24, 21, 25]. We assume that we are given a reference polygonal surface  $\widehat{\Gamma}_*$  consisting of triangular faces and a lift  $\mathbf{l}$  mapping  $\widehat{\Gamma}_*$  bijectively onto  $\gamma$ . We denote by  $\widehat{\mathcal{T}}_*$  the set of faces  $\widehat{K}$  composing  $\widehat{\Gamma}_*$ , i.e., such that  $\widehat{\Gamma}_* = \cup_{\widehat{K} \in \widehat{\mathcal{T}}_*} \widehat{K}$ . We assume that  $\mathbf{l}$  is continuous and piecewise regular. Specifically, for all  $\widehat{K} \in \widehat{\mathcal{T}}_*$ , we assume that

$$(2) \quad \mathbf{l} \in W^{2,\infty}(\widehat{K}).$$

As the surface  $\widehat{\Gamma}_*$  is made up of triangular faces, we can think of the lift as a piecewise parameterization of  $\gamma$ . In fact,  $\{(\widehat{K}, \mathbf{l}|_{\widehat{K}})\}$  is a local coordinated system parameterizing  $\gamma$ . Accordingly,  $D\mathbf{l}(x)$  is locally a  $3 \times 2$  matrix for each  $\hat{x} \in \widehat{K}$  and  $\widehat{K} \in \widehat{\mathcal{T}}_*$ . We assume that the lift is non-degenerate in the sense that there are positive constants  $c$  and  $C$  uniform for  $\hat{x} \in \widehat{\Gamma}_*$  satisfying

$$(3) \quad c|w| \leq |D\mathbf{l}(x)w| \leq C|w|, \quad \forall w \in \mathbb{R}^2.$$

It is not difficult to see that (3) implies

$$(4) \quad \tilde{c} \leq |\det((D\mathbf{l})^t D\mathbf{l}(x))| \leq \tilde{C}$$

with  $\tilde{c}$  and  $\tilde{C}$  uniform for  $\hat{x} \in \widehat{\Gamma}_*$ .

These conditions ensure a uniform norm equivalence between Sobolev spaces defined over  $\gamma$  and subsequently defined approximate surfaces (see Lemma 2.2).

When the reference polygon  $\widehat{\Gamma}_*$  is close enough to  $\gamma$ , the oriented distance function  $d$  for  $\gamma$  can be used to construct a lift from  $\widehat{\Gamma}_*$  to  $\gamma$ . Indeed, such lift is given by

$$(5) \quad \mathbf{l}(\mathbf{x}) := \mathbf{x} - d(\mathbf{x})\nabla d(\mathbf{x}).$$

This lift is generally higher order. However, we do not restrict our consideration to this lift since it might not always be accessible in practice.

**2.3. Uniform surface approximation.** Let  $\widehat{\mathcal{T}}_h$  be a refinement of  $\widehat{\mathcal{T}}_*$  consisting of shape regular triangles of maximum diameter  $h$ . We define

$$\widehat{V}_h = \left\{ \widehat{v} \in C^0(\widehat{\Gamma}_*; \mathbb{R}) \mid \widehat{v}|_{\widehat{K}} \in \mathbb{P}^1(\widehat{K}), \forall \widehat{K} \in \widehat{\mathcal{T}}_h \right\}.$$

Associated with the above space, there are natural nodal values and we denote  $i_h$  to be the corresponding interpolant [11]. Let  $\mathbf{F}_h \equiv i_h(\mathbf{l})$ . This gives rise to the approximate surface  $\Gamma_h = \mathbf{F}_h(\widehat{\Gamma}_*)$ . Clearly,  $\mathbf{F}_h(x_i)$  is on  $\gamma$  for each node  $x_i$  of  $\widehat{\mathcal{T}}_h$ . The above construction leads to uniform norm equivalences (see Lemma 2.2) and the following lemma is the first step in this direction.

**Lemma 2.1.** *If the lift  $\mathbf{l}$  satisfies the regularity assumptions (2) and (3), then there exists  $\bar{h}$ ,  $c$  and  $C$  positive constants uniform in  $h$  and  $x \in \widehat{\Gamma}_*$  such that if  $h \leq \bar{h}$ ,  $\mathbf{F}_h$  is one-to-one and*

$$(7) \quad c|w| \leq |D\mathbf{F}_h(x)w| \leq C|w|, \quad \forall w \in \mathbb{R}^2.$$

*Proof.* Let  $c$  and  $C$ , with or without subscript, be generic positive constants uniform in  $h$  and let  $x$  be in  $\widehat{\Gamma}_*$ . We note that because of (2) and well-known approximation properties of the interpolant [11],

$$(7) \quad \|\mathbf{F}_h - \mathbf{l}\|_{W^{1,\infty}(\widehat{K})} \leq ch\|\mathbf{l}\|_{W^{2,\infty}(\widehat{K})}.$$

Let  $G = (D\mathbf{F}_h)^t D\mathbf{F}_h$  and  $\bar{G} = (D\mathbf{l})^t D\mathbf{l}$ . Note that (3) is the same as

$$(8) \quad c^2|w|^2 \leq (\bar{G}w) \cdot w \leq C^2|w|^2, \quad \forall w \in \mathbb{R}^2$$

and, to prove the lemma, we need only show that

$$(9) \quad c_1^2|w|^2 \leq (Gw) \cdot w \leq C_1^2|w|^2, \quad \forall w \in \mathbb{R}^2$$

for appropriate  $c_1$  and  $C_1$ .

We clearly have

$$(10) \quad G - \bar{G} = ((D\mathbf{F}_h)^t - (D\mathbf{l})^t)((D\mathbf{F}_h - D\mathbf{l}) + D\mathbf{l}) + (D\mathbf{l})^t(D\mathbf{F}_h - D\mathbf{l}).$$

Obvious manipulations using this, (7) and (3) give

$$|(Gw) \cdot w - (\bar{G}w) \cdot w| \leq ch|w|^2, \quad \forall w \in \mathbb{R}^2$$

from which (9) follows for  $h < \bar{h}$  and  $\bar{h}$  sufficiently small.

The mapping  $\mathbf{F}_h$  is piecewise linear and maps triangles onto triangles. It is one-to-one provided that the image of each triangle of  $\widehat{\mathcal{T}}_*$  is a non-degenerate triangle. This non-degeneracy is a consequence of (6). This completes the proof of the lemma.  $\square$

*Remark 2.1* (Quasi-uniform surface triangulations). It follows from (6) that all of the mapped triangles defining  $\Gamma_h$  are uniformly quasi-uniform with constants only depending on those appearing in (6) and the shape regularity of the sequence of triangulation  $\{\widehat{\mathcal{T}}_h\}$ .

*Remark 2.2* (Regular lifts and mesh quasi-uniformity). We provide an example showing that lifts  $\mathbf{l}$  which are smooth diffeomorphisms do not necessarily lead

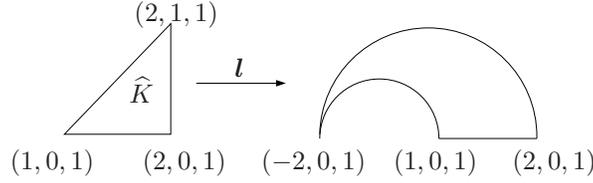


FIGURE 1. A smooth invertible map (11) resulting in a degenerate triangle.

to non-degenerate partitions  $\mathcal{T}_h$ . Consider the element  $\widehat{K}$  defined by its vertices  $(1, 0, 0)$ ,  $(2, 0, 0)$ , and  $(2, 1, 0)$  and the lift  $\mathbf{l} : \widehat{K} \rightarrow \mathbb{R}^3$  defined by

$$(11) \quad \mathbf{l}(\hat{x}_1, \hat{x}_2, \hat{x}_3) = \begin{pmatrix} \hat{x}_1 \cos(\pi \hat{x}_2) \\ \hat{x}_1 \sin(\pi \hat{x}_2) \\ 1 \end{pmatrix}$$

(see Figure 1). Although  $\mathbf{l}$  is smooth and uniformly invertible on  $\widehat{K}$ , it maps its three vertices onto three points on a line and so the resulting triangle is degenerate (has zero area). Accordingly, the triangles of the resulting approximate manifold are not quasi-uniform and (6) cannot hold.

The approximations  $\Gamma_h$  of  $\gamma$  induce approximations  $a_h(\cdot, \cdot) : H^1(\Gamma_h) \times H^1(\Gamma_h) \rightarrow \mathbb{R}$  of the continuous bilinear form  $a(\cdot, \cdot)$  defined by

$$a_h(u, v) := \int_{\Gamma_h} \nabla_{\Gamma_h} u \cdot \nabla_{\Gamma_h} v.$$

**2.4. The link between the surface  $\gamma$  and its approximations.** It will be useful to consider over  $\gamma$  quantities defined over  $\Gamma_h$  and vice versa. Any  $\tilde{v} : \gamma \rightarrow \mathbb{R}$  defines a function  $v = \tilde{v} \circ \mathbf{l} \circ \mathbf{F}_h^{-1} : \Gamma_h \rightarrow \mathbb{R}$ . Similarly, each function  $v : \Gamma_h \rightarrow \mathbb{R}$  defines a function  $\tilde{v} = v \circ \mathbf{F}_h \circ \mathbf{l}^{-1} : \gamma \rightarrow \mathbb{R}$ . The composition of lifts of the above will sometimes be omitted for clarity of notation in later arguments.

The spaces  $L^2(\gamma)$  and  $H^1(\gamma)$  are isomorphic to  $L^2(\Gamma_h)$  and  $H^1(\Gamma_h)$  as given by the next lemma. The proof follows [24, 25].

**Lemma 2.2** (Equivalence of norms). *Assume that the lift  $\mathbf{l}$  satisfies (2)-(3) and let  $\bar{h}$  be the restriction on the mesh size provided by Lemma 2.1. There exist  $c, C$  and  $\bar{h} \leq \bar{h}$  depending on derivatives of  $\mathbf{l}$  such that for all approximations  $\Gamma_h$  satisfying  $h < \bar{h}$  and for all  $\tilde{v} \in H^i(\gamma)$ ,  $i = 0, 1$ , we have*

$$(12) \quad c \|\tilde{v}\|_{i,\gamma} \leq \|v\|_{i,\Gamma_h} \leq C \|\tilde{v}\|_{i,\gamma}.$$

Here  $v$  and  $\tilde{v}$  are related as above.

*Proof.* In this proof,  $c$  is a generic constant uniform in  $h$ . We first provide the proof for the  $H^1$  semi-norms.

For each  $\widehat{K} \in \widehat{\mathcal{T}}_h$ , set  $K = \mathbf{F}_h(\widehat{K})$  and  $\tilde{K} = \mathbf{l}(\widehat{K})$  so that  $(\widehat{K}, \mathbf{F}_h)$  and  $(\widehat{K}, \mathbf{l})$  are local parameterizations of  $\Gamma_h$  and  $\gamma$ , respectively. We define  $q := \sqrt{\det(G)}$  and  $\bar{q} := \sqrt{\det(\bar{G})}$ . For  $v, w$  in  $H^1(\Gamma_h)$ , the change of variable formula yields

$$(13) \quad \int_K (\nabla_{\Gamma_h} v)^t \nabla_{\Gamma_h} w = \int_{\tilde{K}} (\nabla_{\gamma} \tilde{v})^t \nabla_{\gamma} \tilde{w} + \int_{\tilde{K}} (\nabla_{\gamma} \tilde{v})^t A \nabla_{\gamma} \tilde{w},$$

where

$$(14) \quad A = \frac{1}{q} (D\mathbf{l}(G^{-1})(D\mathbf{l})^t q - \bar{q}I).$$

We first observe that  $I$  in (14) can be replaced by the projector  $P \equiv D\mathbf{l}(\bar{G}^{-1})(D\mathbf{l})^t$  since  $P$  is the identity on  $\text{Range}(D\mathbf{l})(x)$  and  $\nabla_\gamma v(x)$  is in  $\text{Range}(D\mathbf{l})$  for any  $v$ . Thus, it suffices to bound

$$(15) \quad \begin{aligned} \|A\|_{L^\infty(\tilde{K})} &\equiv \left\| \frac{1}{q} \left( D\mathbf{l}(G^{-1})(D\mathbf{l})^t q - \bar{q}D\mathbf{l}\bar{G}^{-1}(D\mathbf{l})^t \right) \right\|_{L^\infty(\tilde{K})} \\ &\leq \left| 1 - \frac{q}{\bar{q}} \right| \|D\mathbf{l}(G^{-1})(D\mathbf{l})^t\|_{L^\infty(\tilde{K})} \\ &\quad + \|D\mathbf{l}(\bar{G}^{-1} - G^{-1})(D\mathbf{l})^t\|_{L^\infty(\tilde{K})}. \end{aligned}$$

We start by estimating the first term on the right-hand side above. It easily follows from (2) and (7) that  $|q - \bar{q}| \leq ch$  and hence

$$(16) \quad \left| 1 - \frac{q}{\bar{q}} \right| \leq ch$$

follows from (4). In addition, (3) and (2) imply that

$$\|D\mathbf{l}(G^{-1})(D\mathbf{l})^t\|_{L^\infty(\tilde{K})} \leq C.$$

We bound the second term on the right side of (15) as follows. First,

$$\|D\mathbf{l}(G^{-1} - \bar{G}^{-1})(D\mathbf{l})^t\|_{L^\infty(\tilde{K})} = \|D\mathbf{l}(G^{-1})(\bar{G} - G)(\bar{G}^{-1})(D\mathbf{l})^t\|_{L^\infty(\tilde{K})}.$$

Because of (3) and (7), the matrices  $D\mathbf{l}(G^{-1})$  and  $\bar{G}^{-1}(D\mathbf{l})^t$  are uniformly bounded. Applying (10) and (7) gives

$$\|\bar{G} - G\|_{L^\infty(\tilde{K})} \leq ch.$$

Combining the above results shows that

$$(17) \quad \|A\|_{L^\infty(\tilde{K})} \leq ch.$$

This completes the proof in the case of the  $H^1$  semi-norms.

For  $L^2$  norms, we note that

$$\int_K \mathbf{v}\mathbf{w} = \int_{\tilde{K}} \mathbf{v}\mathbf{w} + \int_{\tilde{K}} \left( \frac{q}{\bar{q}} - 1 \right) \mathbf{v}\mathbf{w}$$

and the desired results follows from (16). This completes the proof of the lemma.  $\square$

The above proof gives the following geometric error estimate.

**Lemma 2.3** (Geometric error). *Assume that the lift  $\mathbf{l}$  satisfies (2)-(3) and let  $\bar{h}$  be the mesh size restriction provided by Lemma 2.1. There exists a constant  $C$  depending on derivatives of  $\mathbf{l}$  so that for all approximation  $\Gamma_h$  satisfying  $h < \bar{h}$  and for all  $\tilde{v}, \tilde{w}$  in  $H^1(\tilde{K})$ , we have*

$$(18) \quad \left| \int_{\tilde{K}} \nabla_\gamma \tilde{u} \cdot \nabla_\gamma \tilde{v} - \int_K \nabla_{\Gamma_h} u \cdot \nabla_{\Gamma_h} v \right| \leq Ch |\tilde{u}|_{1,\tilde{K}} |\tilde{v}|_{1,\tilde{K}}.$$

In the above,  $K$  is a triangle of  $\Gamma_h$  and  $\tilde{K}$  is its lift to  $\gamma$ .

*Proof.* Simply combines (13) and (17). □

*Remark 2.3* (The signed distance function and geometric error). Surprisingly, due to the orthogonality inherent to the projection constructed using the signed distance function  $d$  (see (5)), an extra order of approximation is obtained provided that  $\gamma$  is a  $C^2$  surface. Namely, there exists  $C$  such that for  $h$  sufficiently small

$$(19) \quad \left| \int_{\gamma} \nabla_{\gamma} \tilde{u} \cdot \nabla_{\gamma} \tilde{v} - \int_{\Gamma_h} \nabla_{\Gamma_h} u \cdot \nabla_{\Gamma_h} v \right| \leq Ch^2 |\tilde{u}|_{1,\gamma} |\tilde{v}|_{1,\gamma}$$

(cf. [15]).

### 3. MULTIGRID ALGORITHM FOR THE LAPLACE-BELTRAMI OPERATOR

We introduce the multigrid solver for problem (1). We start by setting up two different finite element spaces leading to variational and non-variational multigrid algorithms. In order to study their convergence, a third auxiliary algorithm defined on  $\gamma$  is introduced along with the discussion of its convergence. The latter is not implementable but is instrumental in the study of the other two algorithms.

**3.1. Multilevel finite element spaces.** To help the reader, we reserve upper case for discrete functions. Also, here and in the remainder of this paper, we shall have to deal with a sequence of meshes and we shall use subscripts  $k$  in place of  $h_k$ .

Let  $\{\widehat{\Gamma}_k\}_{k=1}^J$  be a sequence of nested refinements [11] of  $\widehat{\Gamma}_*$  satisfying the following. First,  $h_1 < \bar{h}$ , where  $\bar{h}$  is the restriction on the mesh size provided in Lemma 2.2. Subsequently, finer meshes are nested and shape regular with mesh size  $h_k$  satisfying

$$(20) \quad c_l \beta^{k-1} h_1 \leq h_k \leq c_u \beta^{k-1} h_1$$

for some  $\beta < 1$  with constants  $c_l$  and  $c_u$  not depending on  $k$ . Typically, the meshes are defined by geometric refinement and  $\beta = 1/2$ . Possible instances are bisection or Red-Green refinement strategies. We then define

$$\widehat{\mathbb{M}}_k = \left\{ \widehat{V} \in C^0(\widehat{\Gamma}_k; \mathbb{R}) \mid \widehat{V}|_{\widehat{K}} \in \mathbb{P}^1(\widehat{K}) \ \forall \widehat{K} \in \widehat{\mathcal{T}}_k \right\}.$$

We also define for  $k = 1, 2, \dots, J$ ,

$$\mathbb{W}_k = \left\{ V : \Gamma_k \rightarrow \mathbb{R} \mid V \circ \mathbf{F}_k \in \widehat{\mathbb{M}}_k \right\} \subset H^1(\Gamma_k).$$

Our first multigrid algorithm will utilize the above sequence of spaces and the Laplace-Beltrami operator on the corresponding manifold  $\Gamma_k$ . This will give rise to a non-variational multigrid algorithm.

The second algorithm is variational and is based on the Laplace-Beltrami operator on the finest grid only. To this end, we consider the family of *nested* finite element spaces over  $\Gamma_J$  by

$$\mathbb{V}_k = \left\{ V : \Gamma_J \rightarrow \mathbb{R} \mid V \circ \mathbf{F}_J \in \widehat{\mathbb{M}}_k \right\} \subset H^1(\Gamma_J).$$

This same set of spaces could alternatively been arrived at by

$$\mathbb{V}_k = \left\{ V : \Gamma_J \rightarrow \mathbb{R} \mid V \circ \mathbf{F}_J \circ \mathbf{F}_k^{-1} \in \mathbb{W}_k \right\} \subset H^1(\Gamma_J).$$

It is natural to consider a third multigrid algorithm which is defined by lifting our finite element spaces to  $\gamma$  and using the Laplace-Beltrami operator there. Of course, in general, the resulting forms are not computable without the introduction

of additional quadrature errors. However, we shall use this algorithm as a basis for our analysis. Specifically, we consider the spaces  $\widetilde{\mathbb{M}}_k$  on  $\gamma$  given by

$$\widetilde{\mathbb{M}}_k = \left\{ \widetilde{V} : \gamma \rightarrow \mathbb{R} \mid \widetilde{V} \circ \mathbf{l} \in \widehat{\mathbb{M}}_k \right\} \subset H^1(\gamma).$$

The corresponding subspaces whose lifts to  $\gamma$  have vanishing mean value are denoted by  $\mathbb{V}_k^0$ ,  $\mathbb{W}_k^0$  and  $\mathbb{M}_k^0$ . In the following,  $\mathbb{M}_k$  stands for either  $\mathbb{V}_k$  or  $\mathbb{W}_k$  and  $\mathbb{M}_k^0$  stands, respectively, for either  $\mathbb{V}_k^0$  or  $\mathbb{W}_k^0$ . Note that this notation is consistent with the fact that the lifts to  $\gamma$  of  $\mathbb{V}_k$  and  $\mathbb{W}_k$  coincide with  $\widetilde{\mathbb{M}}_k$ .

We end this section by providing interpolation estimates. Over  $\widehat{\mathbb{M}}_k$ , we denote the nodal interpolation operator by  $\widehat{\pi}_k : H^1(\widehat{\Gamma}_*) \rightarrow \widehat{\mathbb{M}}_k$ . We shall denote this operator lifted to  $H^1(\gamma)$  by  $\widetilde{\pi}_k$ , i.e.,

$$\widetilde{\pi}_k \widetilde{v} := \widehat{\pi}_k(\widetilde{v} \circ \mathbf{l}).$$

The following interpolation estimate holds:

**Lemma 3.1** (Interpolation Estimate). *Assume the lift  $\mathbf{l}$  satisfies (2)-(3). Then, there exists a constant  $C$  uniform in  $k$  such that for  $\widetilde{v} \in H^2(\gamma)$ ,*

$$\|\widetilde{v} - \widetilde{\pi}_k \widetilde{v}\|_{1,\gamma} \leq Ch_k \|\widetilde{v}\|_{2,\gamma}.$$

*Proof.* This is a direct consequence of standard interpolation estimates [11] and assumptions (2)-(3) on the lift. □

**3.2. The V-cycle multigrid algorithm.** The multigrid algorithm is an iterative procedure for solving the problem: find  $U \in \mathbb{M}_J^0$  satisfying

$$a_J(U, \Theta) = \langle F, \Theta \rangle, \quad \Theta \in \mathbb{M}_J^0.$$

Here  $F \in (\mathbb{M}_J^0)^*$  (the space of linear functionals on  $\mathbb{M}_J^0$ ) and  $\langle \cdot, \cdot \rangle$  denotes the duality pairing. We shall define the algorithm abstractly in terms of prolongation and smoothing operators.

At each level  $k$ , we start with a bilinear form  $A_k(\cdot, \cdot)$  over  $\mathbb{M}_k^0 \times \mathbb{M}_k^0$  approximating  $a(\cdot, \cdot)$ ; see Section 2.3. Recall that  $\mathbb{M}_k^0$  stands for  $\mathbb{V}_k^0$  or  $\mathbb{W}_k^0$  and we define  $A_k(\cdot, \cdot) = a_J(\cdot, \cdot)$  and  $A_k(\cdot, \cdot) = a_k(\cdot, \cdot)$ , respectively. Let  $\mathfrak{A}_k : \mathbb{M}_k^0 \rightarrow (\mathbb{M}_k^0)^*$  be defined by

$$(21) \quad \langle \mathfrak{A}_k V, W \rangle = A_k(V, W), \quad V, W \in \mathbb{M}_k^0.$$

The spaces  $\mathbb{M}_k$ ,  $k = 1, \dots, J$  are connected through prolongation operators  $I_k : \mathbb{M}_{k-1} \rightarrow \mathbb{M}_k$ . In the case of  $\mathbb{V}_k$ , the spaces are nested and  $I_k$  is the natural embedding (and is the lift of the natural embedding of  $\widehat{M}_{k-1} \subset \widehat{M}_k$ ). For  $\mathbb{W}_k$ , we use the natural embedding on  $\gamma$ , specifically, for  $U \in \mathbb{W}_{k-1}$ , we first lift  $U$  to  $\gamma$  (a function in  $\widetilde{\mathbb{M}}_{k-1} \subset \widetilde{\mathbb{M}}_k$ ). The resulting function in  $\widetilde{\mathbb{M}}_k$  is then lifted back to  $\Gamma_k$ . This procedure works for  $\mathbb{V}_k$  and coincides with the natural embedding. We note that  $I_k$  maps  $\mathbb{M}_{k-1}^0$  into  $\mathbb{M}_k^0$  by construction. An important consequence of this construction is that the lift of  $V$  belongs to  $\widetilde{\mathbb{M}}_{k-1}$  and  $I_k V$  to  $\gamma$  coincide. Finally, we would get the same  $I_k$  by lifting back to piecewise linear spaces on  $\widehat{\Gamma}_*$  and using the embedding there. The last definition is that used in the implementation.

The adjoint,  $I_k^t : (\mathbb{M}_k^0)^* \rightarrow (\mathbb{M}_{k-1}^0)^*$ , is defined by

$$\langle I_k^t F, V \rangle = \langle F, I_k V \rangle, \quad F \in (\mathbb{M}_k^0)^*, V \in \mathbb{M}_{k-1}^0.$$

Finally, we require smoothing operators  $R_k : (\mathbb{M}_k^0)^* \rightarrow \mathbb{M}_k^0$ . To take into account the possible “lack of symmetry” of  $R_k$  we introduce its adjoint  $R_k^t : (\mathbb{M}_k^0)^* \rightarrow \mathbb{M}_k^0$  defined by

$$\langle F, R_k^t G \rangle = \langle G, R_k F \rangle, \quad F, G \in (\mathbb{M}_k^0)^*.$$

---

**Algorithm 3.1** Multigrid Algorithm:=  $Mg_k : \mathbb{M}_k^0 \times (\mathbb{M}_k^0)^* \rightarrow \mathbb{M}_k^0$

---

For  $k = 1$  set  $Mg_1(V, F) = \mathfrak{A}_1^{-1} F$ .  
 For  $k = 2, 3, \dots, J$  define  $Mg_k(W, F)$  from  $Mg_{k-1}(\cdot, \cdot)$  by first setting  $V = W$  and performing the following steps:  
 $V \leftarrow V + R_k(F - \mathfrak{A}_k V)$  (Pre-smoothing).  
 $U \leftarrow Mg_{k-1}(0, I_k^t(F - \mathfrak{A}_k V))$  (Correction).  
 Set  $V \leftarrow V + I_k U$ .  
 $V \leftarrow V + R_k^t(F - \mathfrak{A}_k V)$  (Post-smoothing).  
 Set  $Mg_k(W, F) = V$ .

---

Algorithm 3.1 gives the  $V$ -cycle multigrid algorithm which we shall analyze. Set  $E_k : \mathbb{M}_k^0 \rightarrow \mathbb{M}_k^0$  to be

$$(22) \quad E_k(U) = Mg_k(U, 0), \quad k = 1, \dots, J.$$

Note that  $E_k, k = 1, 2, \dots, J$  is symmetric in the  $A_k(\cdot, \cdot)$  inner product for  $k = 1, \dots, J$ . The following recurrence relation is satisfied, (see, e.g., [9]).

$$(23) \quad E_k = K_k^* [I - I_k P_{k-1} + I_k E_{k-1}^p P_{k-1}] K_k.$$

Here

$$(24) \quad K_k = I - R_k \mathfrak{A}_k,$$

$$(25) \quad K_k^* = I - R_k^t \mathfrak{A}_k,$$

and  $P_{k-1} : \mathbb{M}_k^0 \rightarrow \mathbb{M}_{k-1}^0$  is defined by

$$(26) \quad A_{k-1}(P_{k-1} V, W) = A_k(V, I_k W), \quad V \in \mathbb{M}_k^0, W \in \mathbb{M}_{k-1}^0.$$

The operator pairs  $(K_k, K_k^*)$  and  $(P_{k-1}, I_k)$  are adjoint with respect to the  $A_k(\cdot, \cdot)$  inner products. In addition, in the variational case,  $P_{k-1}$  is a projector.

*Remark 3.1* (Functional Representation and Implementation). The functional presentation of the multigrid algorithm is quite natural from the implementation point of view. Indeed, we use a basis  $\{\theta_k^i\}$  for representing functions in  $\mathbb{M}_k^0$  and the dual basis,  $\{\langle F, \theta_k^i \rangle\}$  for representing functionals in  $(\mathbb{M}_k^0)^*$ . Then the application of  $\mathfrak{A}_k$  reduces to multiplication by the stiffness matrix and the action of  $R_k$  can be taken to be, e.g., a Jacobi or Gauss-Seidel sweep based on the stiffness matrix.

**3.3. Auxiliary multigrid algorithm on  $\tilde{\mathbb{M}}_k^0$ .** We now consider a multigrid algorithm defined on  $\tilde{\mathbb{M}}_k^0$ , the lift to  $\gamma$  of functions in  $\hat{\mathbb{M}}_k^0$ . This algorithm will be instrumental in the perturbation argument of Section 4.1 used to prove the convergence of the multigrid algorithm on  $\mathbb{M}_k^0$ . We provide the conditions on the smoother which will ensure the convergence of our multigrid algorithms on  $\tilde{\mathbb{M}}_k^0$  and state theorems which guarantee that these conditions are fulfilled for our application.

The multigrid algorithm based on  $\tilde{\mathbb{M}}_k^0$  is defined as in Section 3.2 for  $\mathbb{M}_k^0$  but with the choices  $A_k(\cdot, \cdot) = a(\cdot, \cdot)$  and  $I_k$  as the natural embedding (recall that the spaces  $\tilde{\mathbb{M}}_k^0$  are nested). We denote by  $\tilde{\mathcal{E}}_k, \tilde{\mathcal{K}}_k, \tilde{\mathcal{K}}_k^*$ , and  $\tilde{\mathcal{P}}_k$ , the operators corresponding to

(22), (24), (25) and (26), respectively. Moreover, for convenience of notation, we introduce the following definitions for  $V \in \mathbb{M}_k^0$ ,

$$\mathcal{E}_k V = (\tilde{\mathcal{E}}_k \tilde{V}) \circ \mathbf{l} \circ \mathbf{F}_k^{-1}, \quad \mathcal{P}_k V = (\tilde{\mathcal{P}}_k \tilde{V}) \circ \mathbf{l} \circ \mathbf{F}_k^{-1}, \quad \mathcal{K}_k V = (\tilde{\mathcal{K}}_k \tilde{V}) \circ \mathbf{l} \circ \mathbf{F}_k^{-1}$$

and similarly for  $\tilde{V} \in \tilde{\mathbb{M}}_k^0$ ,

$$\tilde{E}_k \tilde{V} = \widetilde{E_k V}, \quad \tilde{P}_k \tilde{V} = \widetilde{P_k V}, \quad \tilde{K}_k \tilde{V} = \widetilde{K_k V}.$$

The multigrid analysis starts with the introduction of a base inner product. In our case, we use

$$(\tilde{U}, \tilde{V})_k = \int_{\gamma} \tilde{U} \tilde{V}, \quad \tilde{U}, \tilde{V} \in \tilde{\mathbb{M}}_k^0,$$

and we define operators  $\tilde{A}_k : \tilde{\mathbb{M}}_k^0 \rightarrow \tilde{\mathbb{M}}_k^0$  by

$$(27) \quad \left( \tilde{A}_k \tilde{W}_k, \tilde{V}_k \right)_k = a(\tilde{W}_k, \tilde{V}_k), \quad \forall \tilde{W}_k, \tilde{V}_k \in \tilde{\mathbb{M}}_k^0.$$

We start with the following regularity condition.

**Condition 3.1** (Regularity condition). For some  $0 < \alpha \leq 1$ , there is a constant  $C_P$  not depending on  $k$  such that

$$(28) \quad |a((I - \tilde{\mathcal{P}}_{k-1})\tilde{V}_k, \tilde{V}_k)| \leq C_P^{2\alpha} h_k^{2\alpha} \|\tilde{A}_k \tilde{V}_k\|_{0,k}^{2\alpha} a(\tilde{V}_k, \tilde{V}_k)^{1-\alpha}, \quad \forall \tilde{V}_k \in \tilde{\mathbb{M}}_k^0.$$

The following theorem shows that the regularity condition holds for the multigrid algorithm on  $\gamma$ .

**Theorem 3.2** (Regularity condition). *Let  $\gamma$  be a compact and  $C^2$  hyper-surface of  $\mathbb{R}^3$ . Let  $\hat{\Gamma}_*$  be a reference triangular surface and assume that the lift  $\mathbf{l}$  satisfies (2)-(3). Let  $\{\hat{\Gamma}_k\}$  be a shape regular sequence of approximations satisfying (20). Then (28) holds with  $\alpha = 1$  provided that  $h < \bar{h}$ , where  $\bar{h}$  is the restriction on the mesh size provided by Lemma 2.2.*

*Proof.* The proof is classical and we sketch it for completeness. Let  $H_0^1(\gamma)$  denote the functions in  $H^1(\gamma)$  with zero mean value. We consider the variational problem: Find  $u \in H_0^1(\gamma)$  satisfying

$$a(u, \phi) = (f, \phi), \quad \forall \phi \in H_0^1(\gamma)$$

with  $f \in L^2(\gamma)$ . Here we used  $(\cdot, \cdot)$  to denote the  $L^2(\gamma)$ -inner product. The smoothness of the surface implies, by classical interior regularity estimates [3], that  $u$  is in  $H^2(\gamma)$  and there is a constant  $c$  not depending on  $f$  satisfying  $\|u\|_{2,\gamma} \leq c\|f\|_{L^2(\gamma)}$ . The finite element duality argument and Lemma 3.1 immediately imply that for  $\tilde{U} \in \tilde{\mathbb{V}}_k$ ,

$$\|\tilde{U} - \tilde{\mathcal{P}}_{k-1} \tilde{U}\|_{L^2(\gamma)} \leq Ch_k |\tilde{U} - \tilde{\mathcal{P}}_{k-1} \tilde{U}|_{1,\gamma}.$$

Thus,

$$(29) \quad \begin{aligned} a((I - \tilde{\mathcal{P}}_{k-1})\tilde{U}, \tilde{U}) &\leq \|(I - \tilde{\mathcal{P}}_{k-1})\tilde{U}\|_{L^2(\gamma)} \|\tilde{A}_k \tilde{U}\|_{L^2(\gamma)} \\ &\leq Ch_k a((I - \tilde{\mathcal{P}}_{k-1})\tilde{U}, \tilde{U})^{1/2} \|\tilde{A}_k \tilde{U}\|_{L^2(\gamma)}. \end{aligned}$$

Here we used the fact that  $\tilde{\mathcal{P}}_{k-1}$  is an orthogonal projector with respect to the  $a(\cdot, \cdot)$ -inner product. The inequality (28) with  $\alpha = 1$  follows immediately.  $\square$

Next, we introduce the smoother conditions. We shall only need these for multilevel spaces on  $\gamma$  although they hold for appropriate Gauss-Seidel and Jacobi smoothers on all three of our multilevel sequences.

**Condition 3.2** (Smoother conditions). There exists  $\omega > 0$  not depending on  $J$  such that

$$\omega h_k^2 \|\tilde{V}_k\|_{L^2(\gamma)}^2 \leq \left\langle (\tilde{\mathcal{R}}_k + \tilde{\mathcal{R}}_k^t - \tilde{\mathcal{R}}_k^t \tilde{A}_k \tilde{\mathcal{R}}_k) \tilde{V}_k, \tilde{V}_k \right\rangle, \quad \tilde{V}_k \in \tilde{\mathbb{M}}_k^0, \quad k = 2, \dots, J.$$

In addition, the smoother is stable, i.e.,

$$|(I - \tilde{\mathcal{R}}_k \tilde{A}_k)|_{1,\gamma} \leq 1.$$

In Section 4.2, we observe that the above condition is satisfied for the appropriate Jacobi and Gauss-Seidel smoothers. The following corollary is standard [6] (see also, e.g., [9],[22]) and guarantees uniform convergence of the multigrid algorithm on  $\gamma$ .

**Corollary 3.3** (Convergence of the multigrid algorithm using  $\tilde{\mathbb{M}}_k^0$ ). *Assume that we are given smoothers satisfying the above smoother conditions. Then for  $h < \bar{h}$ , there is a  $\delta$  in  $[0, 1)$  satisfying*

$$|\tilde{\mathcal{E}}_k|_{1,\gamma} \leq \delta.$$

4. CONVERGENCE OF THE MULTIGRID ALGORITHMS ON  $\mathbb{M}_k^0$

Our goal will be to use Corollary 3.3 and a perturbation argument to show that  $\tilde{\mathcal{E}}_J$  and  $E_J$  are close. We note that perturbation algorithms for pairs of variational spaces have been given in [7]. The analysis there is based on the product representation of the multigrid error for both algorithms. The perturbation analysis given below is new (as far as we know) and it applies when only one of the algorithms is variational, i.e., the product representation does not hold on the other. This analysis relies on properties of the smoother used. We discuss the case of Jacobi and Gauss-Seidel smoothers and ensure their applicability.

**4.1. The Perturbation analysis.** We first observe that it follows from the geometric estimate that  $I_K$  is uniformly bounded independent of  $k$  in the norms associated with the  $A_k(\cdot, \cdot)$  and  $A_{k-1}(\cdot, \cdot)$  inner products. It follows that  $|\tilde{P}_k|_{1,\gamma}$  is also uniformly bounded.

We first prove the following lemma relying on the geometric estimate (18):

**Lemma 4.1** (Perturbation of the projection operator). *Let  $\{\hat{\Gamma}_k\}$  be a shape regular sequence of meshes satisfying (20) and assume that the lift  $\mathbf{l}$  satisfies (2)-(3). Then,*

$$(30) \quad |\tilde{\mathcal{P}}_{k-1} - \tilde{P}_{k-1}|_{1,\gamma} \leq ch_k.$$

*Proof.* Let  $\tilde{V}, \tilde{W}$  be in  $\tilde{\mathbb{M}}_k^0$ . Then,

$$\begin{aligned} |a((\tilde{\mathcal{P}}_{k-1} - \tilde{P}_{k-1})\tilde{V}, \tilde{W})| &= |A_k((I - I_k P_{k-1})V, I_k \mathcal{P}_{k-1}W) \\ &\quad + a((I - \tilde{P}_{k-1})\tilde{V}, \tilde{\mathcal{P}}_{k-1}\tilde{W}) - A_k((I - I_k P_{k-1})V, I_k \mathcal{P}_{k-1}W)| \\ &\leq |A_k((I - I_k P_{k-1})V, I_k \mathcal{P}_{k-1}W)| + ch_k |\tilde{V}|_{1,\gamma} |\tilde{W}|_{1,\gamma} \end{aligned}$$

where we used the geometric estimates for the last inequality. In a similar way, but using (20),

$$\begin{aligned} |A_k((I - I_k P_{k-1})V, I_k \mathcal{P}_{k-1}W)| &= |A_{k-1}(P_{k-1}V, \mathcal{P}_{k-1}W) - a(\widetilde{P_{k-1}V}, \widetilde{\mathcal{P}_{k-1}W}) \\ &\quad + a(\widetilde{P_{k-1}V}, \widetilde{\mathcal{P}_{k-1}W}) - A_k(I_k P_{k-1}V, I_k \mathcal{P}_{k-1}W)| \\ &\leq ch_k |\tilde{V}|_{1,\gamma} |\tilde{W}|_{1,\gamma}. \end{aligned}$$

Combining the above estimates completes the proof of the lemma. □

The above lemma coupled with assumption on the smoother (see (31)) yield the convergence of the multigrid solver on  $\mathbb{M}_k^0$  as stated in the next theorem.

**Theorem 4.2** (Convergence of the V-cycle on  $\mathbb{M}_k^0$ ). *Let  $\gamma$  be a compact and  $C^3$  hyper-surface of  $\mathbb{R}^3$ . Let  $\widehat{\Gamma}_*$  be a given reference triangular surface and assume that the lift  $\mathbf{l}$  satisfies (2)-(3). Let  $\{\widehat{\Gamma}_k\}$  be a shape regular sequence of approximations satisfying (20). Consider the V-cycle with  $m_k = 1$  and suppose that*

$$(31) \quad |\widetilde{K}_k - \widetilde{\mathcal{K}}_k|_{1,\gamma} \leq ch_k,$$

where  $\widetilde{\mathcal{K}}_k$  satisfies the Condition 3.2. Then, if  $h_1$  is sufficiently small, there is a  $\delta < 1$  not depending on  $J$  and satisfying

$$(32) \quad |\widetilde{E}_k|_{1,\gamma} \leq \delta.$$

*Proof.* We prove the theorem by induction on  $k$ . Specifically, we shall show that for  $h \leq h_1$ ,

- (i)  $|\widetilde{E}_\ell|_{1,\gamma} \leq \delta < 1$ .
- (ii)  $|\widetilde{E}_\ell - \widetilde{\mathcal{E}}_\ell|_{1,\gamma} \leq |\widetilde{E}_{\ell-1} - \widetilde{\mathcal{E}}_{\ell-1}|_{1,\gamma} + c_1 h_\ell$ .

The value of  $c_1$  will become apparent at the end of the proof. The first assumption is clearly valid for  $\ell = 1$  as  $\widetilde{E}_1 = 0$ . The second assumption is defined to be vacuous in the case of  $\ell = 1$ .

We first note that (i) follows from (ii). Indeed, Corollary 3.3 implies that there is a  $\tilde{\delta} < 1$  satisfying  $|\widetilde{\mathcal{E}}_k|_{1,\gamma} \leq \tilde{\delta}$ . Then (ii) implies

$$(33) \quad |\widetilde{E}_k|_{1,\gamma} \leq \tilde{\delta} + c_1 \sum_{\ell=2}^k h_\ell \leq \tilde{\delta} + c_1 c_u h_1 (1 - \beta)^{-1}.$$

This shows that (i) follows from (ii) for sufficiently small  $h_1$ . In fact, a slightly smaller  $h_1$  and (33) implies the theorem once (i) and (ii) have been verified.

Suppose that (i) and (ii) hold for  $\ell = k - 1$  and  $k \geq 2$ . First we note that geometric estimates imply

$$(34) \quad \begin{aligned} |a(\widetilde{E}_k \widetilde{U}, \widetilde{V}) - a(\widetilde{\mathcal{E}}_k \widetilde{U}, \widetilde{V})| &\leq |a(\widetilde{E}_k \widetilde{U}, \widetilde{V}) - A_k(E_k U, V)| \\ &\quad + |A_k(E_k U, V) - a(\widetilde{\mathcal{E}}_k \widetilde{U}, \widetilde{V})| \\ &\leq ch_k |\widetilde{E}_k \widetilde{U}|_{1,\gamma} |\widetilde{V}|_{1,\gamma} + |A_k(E_k U, V) - a(\widetilde{\mathcal{E}}_k \widetilde{U}, \widetilde{V})|. \end{aligned}$$

The recurrence relation for  $\widetilde{\mathcal{E}}_k$  (compare to (23)) is given by

$$\widetilde{\mathcal{E}}_k = \widetilde{\mathcal{K}}_k^* \left[ I - \widetilde{\mathcal{P}}_{k-1} + \widetilde{\mathcal{E}}_{k-1} \widetilde{\mathcal{P}}_{k-1} \right] \widetilde{\mathcal{K}}_k.$$

Thus, it follows that for  $U, V \in \mathbb{M}_k^0$ ,

$$(35) \quad \begin{aligned} A_k(E_k U, V) - a(\widetilde{\mathcal{E}}_k \widetilde{U}, \widetilde{V}) &= A_k([I - I_k P_{k-1} + I_k E_{k-1} P_{k-1}] K_k U, K_k V) \\ &\quad - a([I - \widetilde{\mathcal{P}}_{k-1} + \widetilde{\mathcal{E}}_{k-1} \widetilde{\mathcal{P}}_{k-1}] \widetilde{\mathcal{K}}_k \widetilde{U}, \widetilde{\mathcal{K}}_k \widetilde{V}). \end{aligned}$$

We shall bound the difference above by breaking it into pieces. For the first part, the geometric estimate and (30) give

$$\begin{aligned}
 & |A_k((I - I_k P_{k-1})K_k U, K_k V) - a((I - \tilde{P}_{k-1})\tilde{\mathcal{K}}_k \tilde{U}, \tilde{\mathcal{K}}_k \tilde{V})| \\
 & \leq |A_k((I - I_k P_{k-1})K_k U, K_k V) - a((I - \tilde{P}_{k-1})\tilde{K}_k \tilde{U}, \tilde{K}_k \tilde{V})| \\
 (36) \quad & + |a((I - \tilde{P}_{k-1})(\tilde{\mathcal{K}}_k - \tilde{K}_k)\tilde{U}, \tilde{K}_k \tilde{V})| \\
 & + |a((I - \tilde{P}_{k-1})\tilde{\mathcal{K}}_k \tilde{U}, (\tilde{\mathcal{K}}_k - \tilde{K}_k)\tilde{V})| \\
 & + |a((\tilde{P}_{k-1} - \tilde{P}_{k-1})\tilde{\mathcal{K}}_k \tilde{U}, \tilde{\mathcal{K}}_k \tilde{V})| \\
 & \leq ch_k |\tilde{U}|_{1,\gamma} |\tilde{V}|_{1,\gamma}.
 \end{aligned}$$

When  $k = 2$ ,  $\tilde{E}_1 = \tilde{\mathcal{E}}_1 = 0$ . For  $k > 2$ , the extra terms in (35) are bounded by

$$\begin{aligned}
 & |A_k(I_k E_{k-1} P_{k-1} K_k U, K_k V) - a(\tilde{\mathcal{E}}_{k-1} \tilde{\mathcal{P}}_{k-1} \tilde{\mathcal{K}}_k \tilde{U}, \tilde{\mathcal{K}}_k \tilde{V})| \\
 & \leq |A_k(I_k E_{k-1} P_{k-1} K_k U, K_k V) - a(\tilde{E}_{k-1} \tilde{P}_{k-1} \tilde{K}_k \tilde{U}, \tilde{P}_{k-1} \tilde{K}_k \tilde{V})| \\
 & + |a(\tilde{E}_{k-1} \tilde{P}_{k-1} \tilde{K}_k \tilde{U}, \tilde{P}_{k-1} \tilde{K}_k \tilde{V}) - a(\tilde{\mathcal{E}}_{k-1} \tilde{\mathcal{P}}_{k-1} \tilde{\mathcal{K}}_k \tilde{U}, \tilde{\mathcal{K}}_k \tilde{V})| \\
 & \leq ch_{k-1} |\tilde{U}|_{1,\gamma} |\tilde{V}|_{1,\gamma} + |a(\tilde{E}_{k-1} \tilde{P}_{k-1} \tilde{K}_k \tilde{U}, \tilde{P}_{k-1} \tilde{K}_k \tilde{V}) \\
 & - a(\tilde{\mathcal{E}}_{k-1} \tilde{\mathcal{P}}_{k-1} \tilde{\mathcal{K}}_k \tilde{U}, \tilde{\mathcal{K}}_k \tilde{V})|.
 \end{aligned}$$

We used the geometric estimate and

$$A_k(I_k E_{k-1} P_{k-1} K_k U, K_k V) = A_{k-1}(E_{k-1} P_{k-1} K_k U, P_{k-1} K_k V)$$

above. Hence, (20) gives

$$\begin{aligned}
 & |A_k(I_k E_{k-1} P_{k-1} K_k U, K_k V) - a(\tilde{\mathcal{E}}_{k-1} \tilde{\mathcal{P}}_{k-1} \tilde{\mathcal{K}}_k \tilde{U}, \tilde{\mathcal{K}}_k \tilde{V})| \\
 & \leq |a((\tilde{E}_{k-1} - \tilde{\mathcal{E}}_{k-1})\tilde{\mathcal{P}}_{k-1} \tilde{\mathcal{K}}_k \tilde{U}, \tilde{\mathcal{P}}_{k-1} \tilde{\mathcal{K}}_k \tilde{V})| \\
 (37) \quad & + |a(\tilde{E}_{k-1}(\tilde{P}_{k-1} \tilde{K}_k - \tilde{\mathcal{P}}_{k-1} \tilde{\mathcal{K}}_k)\tilde{U}, \tilde{\mathcal{P}}_{k-1} \tilde{\mathcal{K}}_k \tilde{V})| \\
 & + |a(\tilde{E}_{k-1} \tilde{P}_{k-1} \tilde{K}_k \tilde{U}, (\tilde{P}_{k-1} \tilde{K}_k - \tilde{\mathcal{P}}_{k-1} \tilde{\mathcal{K}}_k)\tilde{V})| \\
 & + ch_k |\tilde{U}|_{1,\gamma} |\tilde{V}|_{1,\gamma}.
 \end{aligned}$$

Notice that the first term on the right-hand side of (37) is bounded by

$$|a((\tilde{E}_{k-1} - \tilde{\mathcal{E}}_{k-1})\tilde{\mathcal{P}}_{k-1} \tilde{\mathcal{K}}_k \tilde{U}, \tilde{\mathcal{P}}_{k-1} \tilde{\mathcal{K}}_k \tilde{V})| \leq |\tilde{E}_{k-1} - \tilde{\mathcal{E}}_{k-1}|_{1,\gamma} |\tilde{U}|_{1,\gamma} |\tilde{V}|_{1,\gamma}.$$

The remaining terms are bounded in the obvious way using (30), (31) and

$$|\tilde{P}_{k-1} \tilde{K}_k - \tilde{\mathcal{P}}_{k-1} \tilde{\mathcal{K}}_k|_{1,\gamma} \leq |(\tilde{P}_{k-1} - \tilde{\mathcal{P}}_{k-1})\tilde{K}_k|_{1,\gamma} + |\tilde{\mathcal{P}}_{k-1}(\tilde{K}_k - \tilde{\mathcal{K}}_k)|_{1,\gamma} \leq ch_k.$$

Gathering the above estimates, we obtain

$$\begin{aligned}
 (38) \quad & |a(\tilde{E}_k \tilde{U}, \tilde{V}) - a(\tilde{\mathcal{E}}_k \tilde{U}, \tilde{V})| \leq |\tilde{E}_{k-1} - \tilde{\mathcal{E}}_{k-1}|_{1,\gamma} |\tilde{U}|_{1,\gamma} |\tilde{V}|_{1,\gamma} \\
 & + ch_k (|\tilde{U}|_{1,\gamma} + |\tilde{E}_k \tilde{U}|_{1,\gamma}) |\tilde{V}|_{1,\gamma}.
 \end{aligned}$$

Now, it immediately follows from (38) and (33) that

$$|\tilde{E}_k|_{1,\gamma} \leq \tilde{\delta} + c_1 c_u (1 - \beta)^{-1} h_1 + ch_k (1 + |\tilde{E}_k|_{1,\gamma}).$$

Choosing  $h_1$  so small that  $ch_1 \leq 1/2$  implies that

$$|\tilde{E}_k|_{1,\gamma} \leq 3 + 2c_1 c_u (1 - \beta)^{-1} h_1.$$

This and (38) imply that

$$|\tilde{E}_k - \tilde{\mathcal{E}}_k|_{1,\gamma} \leq |\tilde{E}_{k-1} - \tilde{\mathcal{E}}_{k-1}|_{1,\gamma} + ch_k(4 + 2c_1c_u(1 - \beta)^{-1}h_1).$$

Choosing  $h_1$  so small that  $2cc_u(1 - \beta)^{-1}h_1 \leq 1/2$  and setting  $c_1 = 8c$  we find that  $c(4 + 2c_1c_u(1 - \beta)^{-1}h_1) \leq c_1$ . This implies (ii) for  $k$  and completes the proof of the theorem.  $\square$

**Corollary 4.3** (Convergence of the V-cycle on  $\mathbb{M}_k^0$ ). *Under the assumptions of Theorem 4.2, then for  $h$  sufficiently small, there is a  $\delta \in [0, 1)$  satisfying*

$$|E_k|_{1,\Gamma_J} \leq \delta, \quad k = 1, \dots, J,$$

for the variational algorithm and

$$|E_k|_{1,\Gamma_k} \leq \delta, \quad k = 1, \dots, J,$$

for the non-variational algorithm.

*Proof.* Simply combine Theorem 4.2 to the geometric estimate (18).  $\square$

*Remark 4.1* (Smallness conditions on the coarse grid approximation). The smallness condition on the coarse grid appears in two places in our analysis. The first is in Lemma 2.1 and gives rise to quasi-uniform triangulations. We note that without further assumption, one can apply the results of [21] and obtain preconditioning results for their hierarchical method with logarithmic condition number bounds. Alternatively, one can recover a uniform preconditioner based on the variable V-cycle non-variational algorithm (without any further restrictions on the coarse grid). This result is not hard with the machinery of this paper (some of the analysis is already contained in (30)). We avoided inclusion of this result and its proof for brevity. Our analysis suggests that one might have to further refine the coarse grid for the V-cycle results of this paper although the necessity of further refinement is not observed in our numerical experiments.

It is often convenient to use multigrid as a preconditioner in the conjugate gradient iteration. In this case, we are required to estimate the spectral conditioning of the operator  $B_J A_J \equiv I - E_J$  on  $\mathbb{M}_k^0$ .

**Corollary 4.4** (Multigrid as a preconditioner). *Let  $\delta$  be the number appearing in Corollary 4.3. The operator  $B_J A_J$  is symmetric and positive definite (in the  $A_J(\cdot, \cdot)$  or  $B_J^{-1}(\cdot, \cdot)$  inner products) with condition number bounded by*

$$K(B_J A_J) = (1 + \delta)/(1 - \delta).$$

*Proof.* It is an immediate consequence of the convergence rates provided in Corollary 4.3 that

$$-\delta A_J(V, V) \leq A_J(E_J V, V) \leq \delta A_J(V, V), \quad \forall V \in \mathbb{M}_k^0$$

from which the estimate of the corollary follows.  $\square$

*Remark 4.2* (Jacobi and Gauss-Seidel smoothers). We shall show that the smoother condition (31) holds for appropriate Jacobi and Gauss-Seidel smoothers (see Lemmas 4.7 and 4.8). Thus, we have shown uniform convergence for the V-cycle algorithm with one smoothing per level for a case where the multigrid algorithm is not variational, i.e., for  $\{\mathbb{W}_k^0\}$ . As far as we know, this is the first such example in the literature. The common approach to deal with a non-variational algorithm requires either the W-cycle or variable V-cycle algorithm (see, e.g., [8, 9]).

*Remark 4.3* (Direct analysis of the multigrid algorithm). The analysis for the variational algorithm on  $\{\mathbb{V}_k^0\}$  could be done more directly. Indeed, it is possible to combine the geometric error estimates and the finite element duality argument on  $\gamma$  to show full regularity and approximation ( $\alpha = 1$  in (28)) in this case. In contrast, this approach only leads to  $\alpha = 1/2$  in the non-variational case although  $\alpha = 1$  results from using the signed distance function on smoother manifolds. In the case of the signed distance lift, it is possible to put together a different analysis to guarantee uniform convergence of the V-cycle algorithm with spaces  $\{\mathbb{W}_k^0\}$ . This is not as satisfactory as the above analysis which works for less restrictive lifts.

**4.2. Jacobi and Gauss-Seidel smoothers.** In this section, we study the Jacobi and Gauss-Seidel smoothers and show that they satisfy the conditions assumed in Theorem 4.2.

Let  $\varphi_k^i, i = 0, \dots, N_k$  be the Lagrange basis associated to the finite element space  $\mathbb{M}_k$  (in the case of  $\mathbb{V}_k, \varphi_k^i$  is the lift to  $\Gamma_J$  of the Lagrange basis function from  $\widehat{\mathbb{M}}_k$ ). The corresponding functions (the lifts of  $\varphi_k^i$ ) in  $\widetilde{\mathbb{M}}_k$  are denoted by  $\widetilde{\varphi}_k^i$ . We denote by  $\theta_k^i := \varphi_k^i - |\gamma|^{-1} \int_\gamma \widetilde{\varphi}_k^i$  where  $|\gamma|$  denotes the measure of  $\gamma$ . Similarly, we set  $\widetilde{\theta}_k^i := \widetilde{\varphi}_k^i - |\gamma|^{-1} \int_\gamma \widetilde{\varphi}_k^i$ .

*Remark 4.4* (Implementation of the Vanishing Mean Value). Our multigrid algorithms are implemented in either  $\{\mathbb{W}_k^0\}$  or  $\{\mathbb{V}_k^0\}$ . Conceptually, we use  $\{\theta_k^i\} i = 1, \dots, N_k$  as our computational basis. Technically, this means that our basis functions no longer have compact support. However, because the form  $a_k(\cdot, \cdot)$  kills constants, the stiffness matrix is still sparse. The action of the smoother is more or less local (as discussed below).

*Remark 4.5* (Implementation of the prolongation operator). One must be able to implement the prolongation operator in the multigrid algorithm. For  $U \in \mathbb{M}_{k-1}^0$ , we write

$$U = \sum_{i=1}^{N_{k-1}} c_i \theta_{k-1}^i \quad \text{and set} \quad W = \sum_{i=1}^{N_{k-1}} c_i \varphi_{k-1}^i.$$

Given  $v$  we define  $\bar{v} = |\gamma|^{-1} \int_\gamma v$ . Then  $U = W - \bar{W}$  so  $I_k U = I_k W - \bar{W}$ . As  $W$  vanishes on the zeroth node,

$$I_k W = \sum_{i=1}^{N_k} d_i \varphi_k^i$$

and

$$I_k U = \sum_{i=1}^{N_k} d_i \theta_k^i + \overline{I_k W} - \bar{W} = \sum_{i=1}^{N_k} d_i \theta_k^i.$$

Thus, implementing  $I_k$  with our basis functions is no more difficult than implementing  $I_k$  using the usual basis functions (i.e., computing  $\{d_i\}$ ).

As our multigrid analysis is done on  $\gamma$ , we consider  $\widetilde{\mathbb{M}}_k^i = \text{span}(\widetilde{\theta}_k^i)$ . It is clear that  $\widetilde{V} \in \widetilde{\mathbb{M}}_k^0$  can be written uniquely as

$$\widetilde{V} = \sum_{i=0}^{N_k} \widetilde{V}^i \widetilde{\varphi}_k^i.$$

As the operation of removing the mean value is linear,

$$(39) \quad \tilde{V} = \sum_{i=0}^{N_k} \tilde{V}^i \tilde{\theta}_k^i.$$

The above summation is not direct since  $0 = \sum_{i=0}^{N_k} \tilde{\theta}_k^i$ . The following result shows the  $L^2(\gamma)$  stability for decompositions into these spaces.

**Lemma 4.5** ( $L^2$  Stability of the decomposition of  $\tilde{\mathbb{M}}_k^0$ ). *Assume that the lift  $\mathbf{l}$  satisfies (2)-(3) and let  $\bar{h}$  be the restriction on the mesh size provided by Lemma 2.2. Let  $\tilde{V} \in \tilde{\mathbb{M}}_k^0$  be decomposed as in (39). For  $h_k < \bar{h}$ , there is a constant  $c$  not depending on  $k$  and satisfying*

$$\sum_{i=0}^{N_k} \|\tilde{V}^i \tilde{\theta}_k^i\|_{L^2(\gamma)}^2 \leq c \|\tilde{V}\|_{L^2(\gamma)}^2.$$

*Proof.* In this proof,  $c$  denotes a generic constant uniform in  $k$  and we consider  $\{\varphi_k^i\}$  coming from  $\Gamma_k$ . The triangles of  $\Gamma_k$  are quasi-uniform by Lemma 2.1. As a result, the set of basis functions  $\{\varphi_k^i\}$  generates a stable  $L^2(\Gamma)$  decomposition

$$\sum_{i=0}^{N_k} \|\tilde{V}_k^i \varphi_k^i\|_{L^2(\Gamma)}^2 \leq c \left\| \sum_{i=0}^{N_k} \tilde{V}_k^i \varphi_k^i \right\|_{L^2(\Gamma)}^2.$$

Thus, Lemma 2.2 implies

$$(40) \quad \begin{aligned} \sum_{i=0}^{N_k} \|\tilde{V}_k^i \tilde{\theta}_k^i\|_{L^2(\gamma)}^2 &\leq \sum_{i=0}^{N_k} \|\tilde{V}_k^i \tilde{\varphi}_k^i\|_{L^2(\gamma)}^2 \leq c \left\| \sum_{i=0}^{N_k} \tilde{V}_k^i \tilde{\varphi}_k^i \right\|_{L^2(\gamma)}^2 \\ &= c \left\| \sum_{i=0}^{N_k} \tilde{V}_k^i \tilde{\theta}_k^i \right\|_{L^2(\gamma)}^2. \quad \square \end{aligned}$$

Even though the decomposition is not local in  $L^2(\gamma)$ , it is local in  $a(\cdot, \cdot)$ . This combined with the  $L^2$  stability above is sufficient to ensure Condition 3.2 in the case of Jacobi or Gauss-Seidel smoothers (see, e.g., Theorems 8.1 and 8.2 of [9]).

**Corollary 4.6** (Condition 3.2). *Assume that the lift  $\mathbf{l}$  satisfies (2)-(3) and let  $\bar{h}$  be the restriction on the mesh size provided by Lemma 2.2. The Jacobi and Gauss-Seidel smoothers on  $\tilde{\mathbb{M}}_k^0$  satisfy Condition 3.2.*

*Remark 4.6* (Zero point smoothing). It is an interesting question as to whether the zero node correction is, indeed, necessary. First, it is not difficult to show that stability of the decomposition fails to hold without the  $i = 0$  term. Without  $i = 0$ , the decompositions are, of course, direct and writing

$$\theta_k^0 = - \sum_{i=1}^{N_k} \theta_k^i,$$

we find

$$\sum_{i=1}^{N_k} \|\theta_k^i\|_k^2 \geq ch_k^{-2} \|\theta_k^0\|_k^2.$$

Thus the decomposition inequality fails in a significant way. Computational results given later provide additional evidence that adding the zero node smoothing results in a better algorithm.

We now focus on (31) and first consider the Jacobi smoother. Its implementation is discussed in Remark 4.7 below. The corresponding operators  $\tilde{\mathcal{K}}_k$  and  $K_k$  are given by

$$\tilde{\mathcal{K}}_k = I - \beta \sum_{i=0}^{N_k} \tilde{\mathcal{P}}_k^i \quad \text{and} \quad K_k = I - \beta \sum_{i=0}^{N_k} P_k^i.$$

The operators  $\tilde{\mathcal{P}}_k^i$  and  $P_k^i$  are orthogonal projectors defined on  $\tilde{\mathbb{M}}_k^0$  and  $\mathbb{M}_k^0$  with respect to  $a(\cdot, \cdot)$  and  $A_k(\cdot, \cdot)$  onto the subspaces spanned by  $\tilde{\theta}_k^i$  and  $\theta_k^i$ , respectively. The positive constant  $\beta$  is chosen so that

$$\beta \left| \sum_{i=0}^{N_k} \tilde{\mathcal{P}}_k^i \right|_{1,\gamma} \leq 1.$$

Note that we include the term  $i = 0$  in the smoother. This subspace does not correspond to a basis function. We then have the estimate (31).

**Lemma 4.7** (Jacobi smoother). *Assume that the lift  $\mathbf{l}$  satisfies (2)-(3). Then, equation (31) holds for the Jacobi smoother provided  $h_1$  is sufficiently small.*

*Proof.* This proof follows Theorem 5.5 of [7] (see, also, Theorem 11.3 of [9]). Clearly, for  $\tilde{U}, \tilde{V} \in \tilde{\mathbb{M}}_k$ ,

$$a((\tilde{K}_k - \tilde{\mathcal{K}}_k)\tilde{U}, \tilde{V}) = \beta \sum_{i=0}^{N_k} a((\tilde{\mathcal{P}}_k^i - \tilde{P}_k^i)\tilde{U}, \tilde{V}).$$

Let  $\gamma_k^i$  denote the support of  $\tilde{\varphi}_k^i$  and  $\tilde{W}$  be in  $\tilde{\mathbb{M}}_k^0$ . It is immediate that

$$|\tilde{\mathcal{P}}_k^i \tilde{W}|_{1,\gamma} \leq |\tilde{W}|_{1,\gamma_k^i} \quad \text{and} \quad A_k(P_k^i W, P_k^i W) \leq \int_{\Gamma_k^i} (\nabla_{\Gamma_k^i} W)^t \nabla_{\Gamma_k^i} W$$

where  $\Gamma_k^i$  is the lift of  $\gamma_k^i$  to  $\Gamma_k$  or  $\Gamma_J$  (depending on the definition of  $\mathbb{M}_k^0$ ). It follows from Lemma 2.3 and the second inequality above imply that for  $h_1$  sufficiently small,

$$|\tilde{P}_k^i \tilde{W}|_{1,\gamma} \leq C |\tilde{W}|_{1,\gamma_k^i}.$$

The geometric estimate (18) then gives

$$\begin{aligned} |a((\tilde{\mathcal{P}}_k^i - \tilde{P}_k^i)\tilde{U}, \tilde{V})| &= |a((I - P_k^i)\tilde{U}, \tilde{\mathcal{P}}_k^i \tilde{V})| \\ (41) \quad &= |a(\tilde{U}, \tilde{\mathcal{P}}_k^i \tilde{V}) - A_k(U, \mathcal{P}_k^i \tilde{V}) + A_k(P_k^i U, \mathcal{P}_k^i V) - a(\tilde{P}_k^i \tilde{U}, \tilde{\mathcal{P}}_k^i \tilde{V})| \\ &\leq ch |\tilde{U}|_{1,\gamma_k^i} |\tilde{V}|_{1,\gamma_k^i}. \end{aligned}$$

Summing and applying the Schwartz inequality gives

$$|a((\tilde{K}_k - \tilde{\mathcal{K}}_k)\tilde{U}, \tilde{V})| \leq ch \left( \sum_{i=0}^{N_k} |\tilde{U}|_{1,\gamma_k^i}^2 \right)^{1/2} \left( \sum_{i=0}^{N_k} |\tilde{V}|_{1,\gamma_k^i}^2 \right)^{1/2} \leq ch |\tilde{U}|_{1,\gamma} |\tilde{V}|_{1,\gamma}$$

where we used the fact that the image of any triangle  $K \in \Gamma_k$  on  $\gamma$  is overlapped by the supports of at most three functions in  $\{\tilde{\varphi}_k^i\}$ . This completes the proof.  $\square$

We now consider the Gauss-Seidel smoother. In this case, the operators  $\tilde{K}_k$  and  $K_k$  are defined by

$$\tilde{K}_k = \prod_{j=1}^l (I - \tilde{P}_k^j) \quad \text{and} \quad K_k = \prod_{j=1}^l (I - P_k^j)$$

and satisfy estimate (31) as stated in the next lemma.

**Lemma 4.8** (Gauss-Seidel smoother). *Assume that the lift  $l$  satisfies (2)-(3). Then (31) holds for the Gauss-Seidel smoother provided  $h_1$  is sufficiently small.*

*Proof.* The lemma follows from (41) and the proof of Theorem 11.4 in [9]. □

*Remark 4.7* (Implementation of the Jacobi and Gauss-Seidel smoothers). The Jacobi smoother is defined by

$$R_k(F) = \sum_{i=0}^{N_k} F(\theta_k^i) A_k(\theta_k^i, \theta_k^i)^{-1} \theta_k^i.$$

As  $F(\theta_k^i)$ ,  $i = 1, \dots, N_k$  are just the coefficients of  $F$  in the dual basis (and are given as data), multiplying the  $N_k$ -dimensional coefficient vector for  $F$  by the inverse of the diagonal of the  $(N_k \times N_k)$  stiffness matrix gives the basis expansion for the sum defining  $R_k$  excluding the first term. In implementation, it is natural to build the full matrix  $A_k(\theta_k^i, \theta_k^j) = A_k(\varphi_k^i, \varphi_k^j)$ ,  $i, j = 0, \dots, N_k$  and hence  $A_k(\theta_k^0, \theta_k^0)$  is readily available. The increment from the zeroth term is then given by

$$(42) \quad F(\theta_k^0) A_k(\theta_k^0, \theta_k^0)^{-1} \theta_k^0 = \left[ \left( - \sum_{i=1}^{N_k} F(\theta_k^i) \right) \right] A_k(\theta_k^0, \theta_k^0)^{-1} \left( - \sum_{i=1}^{N_k} \theta_k^i \right).$$

Although this adds some computational effort, it does not change the asymptotic run time. The implementation of the Gauss-Seidel smoother involves a Gauss-Seidel sweep avoiding the zeroth node followed by a zeroth node fix up similar to (42).

### 5. NUMERICAL EXPERIMENTS

In this section, we give the results of numerical experiments which illustrate the theory developed earlier. Our example takes  $\gamma$  to be the rotation about the  $z$  axis of an ellipse in the  $xz$  plane with major axis 10 (along  $x = 0$ ) and minor axis 1. This produces an anisotropic surface. As we shall see, this anisotropic aspect does not adversely affect the convergence of the multigrid algorithm as long as the triangles remain quasi-uniform on all levels.

The first major problem that one faces is to somehow determine the placement of the computational nodes on the surface. As discussed earlier, we need a reference surface  $\hat{\Gamma}_*$  and a lift  $l$  mapping  $\hat{\Gamma}_*$  one-to-one and onto  $\gamma$ . We defined  $\hat{\Gamma}_*$  to be the boundary of the region  $[-1, -1]^2 \times [-z_m, z_m]$ . The parameter  $z_m$  was introduced to help generate a shape regular surface mesh. Specifically, we set  $z_m = 10 \sin(70^\circ) \approx 9.4$  unless otherwise noted.

To define  $l$ , the six faces of  $\hat{\Gamma}_*$  are mapped to  $\gamma$  as follows. First we consider the four faces of varying  $z$ . The intersection of these faces with a plane of constant  $z$  is the boundary of the square with corner coordinates  $(\pm 1, \pm 1)$ . The lift map points on this boundary to the same  $z$  value (and hence to a circle of radius  $b_z = (1 - (z/10)^2)^{1/2}$ ). Points on the square are moved to the circle along the line emanating from the origin (see Figure 2). The bottom and top faces of  $\hat{\Gamma}_*$  are then mapped

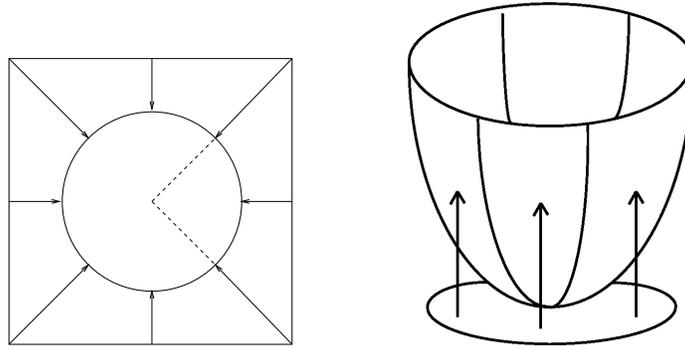


FIGURE 2. Construction of the lift from a box to an ellipsoid. The lift from the four faces of varying  $z$  is the radial projection on each constant  $z$  plane (left). The lifts from the bottom and top parts are obtained by an additional projection along the  $z$  direction (right).

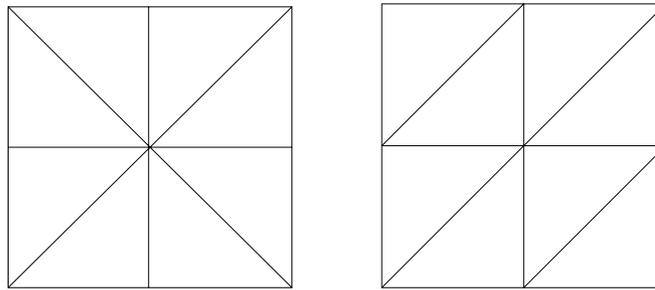


FIGURE 3. Good and bad triangulation of the top and bottom faces.

to the remainder (bottom and top) parts of the rotated ellipse. We first map the boundary of the square  $[-1, 1]^2$  to the circle of radius  $r_m = \sqrt{1 - z_m^2}/100$  by collapsing any boundary point to the boundary of the circle along the ray emanating from the origin (as in Figure 2). Interior points in  $[-1, 1]^2$  along this ray are collapsed proportionally into the interior of the circle. The result of this is a map  $(x, y) \rightarrow (\tilde{x}, \tilde{y})$ . The lift is then defined by  $(x, y, z_m) \rightarrow (\tilde{x}, \tilde{y}, 10\sqrt{1 - \tilde{x}^2 - \tilde{y}^2})$ . Because of the corners, this lift is singular along the lines  $x = \pm y$ .

Because of the above-mentioned singularity, the above map does not satisfy (2) globally, i.e., with  $K$  replaced by  $\hat{\Gamma}_*$ . This can lead to meshes which fail to satisfy (3). Indeed, the two lines meeting at the corner map to a circle. Thus, a triangle of size  $h$  with the corner as one of its vertices and the two remaining vertices on the boundary of the face will flatten out as  $h$  becomes small as all three vertices get mapped to the circle (which is approximately a straight line in the limit of small  $h$ ).

The above problem can be avoided with a judicious choice of coarse triangulation. To get a triangulation of  $\hat{\Gamma}_*$  satisfying (2), we need to have the edges of the triangulation follow the singularities of  $\mathbf{l}$ . These are along the top and bottom faces on the lines  $x = \pm y$ . Thus, our coarse grid triangulation of the top and bottom

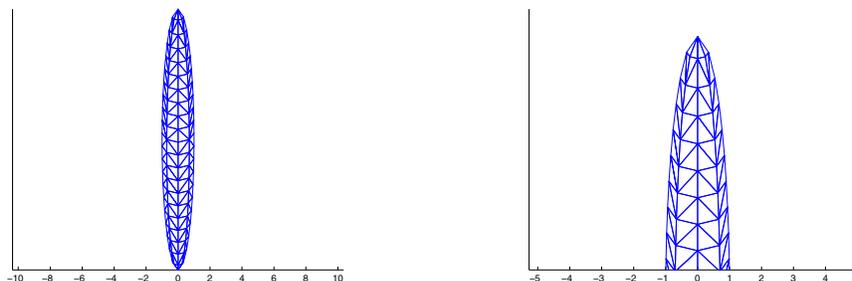


FIGURE 4. The mesh  $\Gamma_1$ .

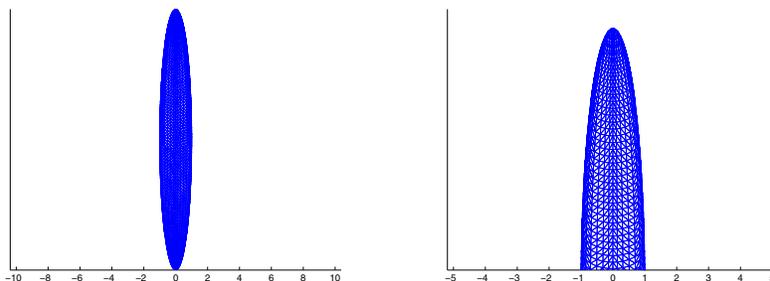


FIGURE 5. The mesh  $\Gamma_3$ .

faces was taken as in the left triangulation of Figure 3. The triangulation of the remaining faces does not introduce other singularities as long as they respect the boundary of the faces, i.e., the support of a triangle is always contained in one of the faces.

*Remark 5.1* (Triangulation satisfying (2) and (3)). In contrast, the coarse triangulation of the top and bottom faces illustrated in the right triangulation of Figure 3 fails to satisfy (2) and (3) and leads to multigrid algorithms with much slower convergence (compared with those reported below).

The multigrid algorithm behaves poorly in the presence of anisotropic elements. To get a reasonably quasi-uniform coarse grid mesh, we put a  $2 \times 20$  regular mesh on each  $z$ -varying face and a  $2 \times 2$  regular mesh on the top and bottom faces. The resulting approximation  $\Gamma_1$  of  $\gamma$  is illustrated in Figure 4.

The results reported below correspond to uniformly refined grids (on  $\widehat{\Gamma}_*$ ), where, to define the finer grid, each triangle in a coarser grid is broken into four by connecting the midpoints of the edges. For the initial results, these grids give rise to the approximation surfaces  $\{\Gamma_k\}$  defined using  $\mathbf{l}$  as discussed earlier. A finer approximation to  $\gamma$  is illustrated in Figure 5. We use the Gauss-Seidel smoother which includes the 0-subspace and the V-cycle algorithm with one pre- and post-smoothing.

The tables below report the convergence rates associated with the preconditioned operator  $B_J A_J$  (see Corollary 4.4). In the case of the non-variational algorithm

( $M_k^0 = \mathbb{W}_k^0$ ), we also report the largest  $\lambda_N$  and smallest  $\lambda_1$  eigenvalue of  $B_J A_J$ . Note that for the variational algorithm  $\lambda_N = 1$ .

Table 1 reports the results for the non-variational algorithm. We include the multilevel case (starting from the coarsest mesh) as well as the two level case. As expected the two level case leads to smaller condition numbers (a faster convergent algorithm). The multilevel algorithm does very well, almost as good as the two level. Although there is some deterioration in the smallest eigenvalue as the number of levels increase, this deterioration appears to be slowing down for finer discretization which is in agreement with a uniform bound as the number of levels becomes large guaranteed by the theory. The largest eigenvalue remains fairly close to one. It is larger in the multilevel case as there is apparently a multiplicative effect. In contrast, the largest eigenvalue for the two level algorithm is more closely related to the fine grid approximation to the surface and so converges to one as the mesh size decreases. The presence of question marks in the largest eigenvalue results indicates our uncertainty in the last digits of the result. This is because there are a large number of eigenvectors with eigenvalues very close to one and it is very difficult for us to obtain the largest with high accuracy.

	Multilevel			Two-level			
$J$	$\lambda_1$	$\lambda_N$	$\rho$	$\lambda_1$	$\lambda_N$	$\rho$	#unknowns.
2	.615	1.02	.385	.615	1.02	.385	673
3	.464	1.03	.536	.480	1.007	.520	2,689
4	.348	1.03	.652	.361	1.002	.639	10,753
5	.269	1.03	.731	.275	1.0009?	.725	43,009
6	.216	1.03	.784	.220	1.0097?	.780	172,033
7	.181	1.03	.819	.187	1.0003?	.813	688,129
8	.160	1.03	.840	.163	1.00015?	.837	2,752,514

TABLE 1. Eigenvalues of  $B_J A_J$  and convergence rates  $\rho$  for the non-variational algorithm.

Table 2 gives the corresponding results for the variational algorithm. Comparing Table 2 with Table 1, we see very little difference between the convergence rates of the variational and non-variational algorithms.

$J$	Multilevel	Two-level
2	.384	.384
3	.534	.518
4	.652	.638
5	.731	.724
6	.784	.780
7	.816	.813
8	.838	.837

TABLE 2. Convergence rates for the variational algorithm.

Table 3 illustrates what happens if zeroth degree of freedom is left out of the smoother. One could think of dealing with the indefiniteness by pinning one point, i.e., look for a solution which vanishes on the zeroth node. If we do this on all grids and the zeroth node is common to all, then the interpolation will preserve this condition. Assuming that we apply Gauss-Seidel smoothing, the resulting algorithm is algebraically the same as our multigrid algorithm without the zeroth node in the smoother (see Remark 4.6). The theory fails in this case and the resulting algorithms behaves somewhat worse (comparing Table 3 with Table 2). It is somewhat strange, though, that the two level algorithms (with or without the zeroth point in the smoother) appear to be performing the same for finer grid problems.

J	Multilevel	Two-level
2	.488	.488
3	.658	.585
4	.759	.657
5	.825	.722
6	.869	.778
7	.899	.813
8	.921	.835

TABLE 3. Convergence rates for the variational algorithm without the zeroth point in the smoother.

Table 4 illustrates what happens when finer grids  $\{\Gamma_k\}$  are generated using the oriented distance function. Specifically, grids are generated recursively as follows. First, if  $x_j$  is a node of  $\Gamma_k$ , then  $x_j$  is on  $\gamma$  and becomes a node of  $\Gamma_{k+1}$ . Next, for each midpoint  $x_j$  of an edge of a triangle of  $\Gamma_k$  we define  $\tilde{x}_j$  to be the point on  $\gamma$  which is closest to  $x_j$ . The collection of nodes  $\{\tilde{x}_j\}$  are the new nodes of  $\Gamma_{k+1}$ . Comparing Table 4 and Table 2, we see that the use of the distance function leads to faster converging multigrid algorithms. Apparently, this is because the resulting grids are more shape regular. It is interesting to note that for grids generated using the distance function, the two level algorithm actually performs worse than the multilevel algorithm. This is a consequence of the way that the meshes were generated. The multilevel algorithm uses the distance function starting from the coarsest grid. In contrast, the two level algorithm was implemented using the original lift on level  $J - 1$  and the distance function only to generate the  $J$ th grid. The triangles on the  $J$ 'th grid generated in the multilevel case ended up being more shape regular than those generated by the two level algorithm. Clearly, the multilevel algorithm with grids based on the distance function gives the best results of all multigrid algorithms tested.

As a final comparison, we consider the surfaces resulting from three different ellipses with major axis, 1, 3 and 10. In all cases, we took a coarse mesh mapped from  $2 \times 2$  grids on the top and bottom and  $2 \times 2$ ,  $2 \times 6$  and  $2 \times 20$  grids (depending on the major axis) in the  $z$ -varying faces. We used  $z_m = \sin(45^\circ)$ ,  $z_m = 3 \sin(55^\circ)$ , and  $z_m = 10 \sin(65^\circ)$ , respectively. The choice of  $65^\circ$  was made to improve the mesh quality in the last case (in contrast to  $z_m = 10 \sin(70^\circ)$  used in the earlier examples).

J	Multilevel	Two-level
2	.364	.366
3	.437	.496
4	.488	.630
5	.538	.722
6	.564	.777
7	.583	.813
8	.597	.837

TABLE 4. Convergence rates for the variational algorithm and grids generated using the distance function.

J	Major Axis		
	1	3	10
2	.100	.124	.271
3	.127	.166	.345
4	.157	.192	.395
5	.195	.233	.422
6	.231	.256	.451
7	.256	.273	.478
8	.273	.287	.495

TABLE 5. Convergence rates for the variational algorithm for different major axis.

Table 5 reports the convergence rates for the variational algorithm (including the zero point smoothing and finer grids based on the distance function). The results for the sphere (major axis 1) are marginally better than the other two cases, with the largest major axis being the worst. It is our belief that this behavior results from the mesh quality. If one defines the aspect of a triangle to be the ratio of the triangle's area to its diameter squared (an equilateral triangle has aspect .433) and defines the aspect of a grid to be the minimum of the aspects of all its triangles, then our coarse grids have aspect .41, .37 and .30, respectively.<sup>2</sup> The numerical results suggest that one can expect a lower aspect grid to result in a multigrid algorithm with a larger condition number.

#### ACKNOWLEDGMENTS

This work was supported in part by award number KUS-C1-016-04 made by King Abdulla University of Science and Technology (KAUST). The first author was also supported in part by the National Science Foundation through Grant DMS-0914977 while the second was also supported in part by the National Science Foundation through Grant DMS-0609544.

<sup>2</sup>The aspect of the earlier runs ( $z_m = 10 \sin(70^\circ)$ ) was .24.

## REFERENCES

- [1] Pankaj K. Agarwal and Subhash Suri. Surface approximation and geometric partitions. *Proceedings of the Fifth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 24–33, 1994. Arlington, Virginia, United States. MR1285148 (95h:68181)
- [2] G. Aubert and P. Kornprobst. *Mathematical problems in image processing*, volume 147 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 2002. Partial differential equations and the calculus of variations, With a foreword by Olivier Faugeras. MR1865346 (2002m:49001)
- [3] Th. Aubin. *Nonlinear analysis on manifolds. Monge-Ampère equations*, volume 252 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, New York, 1982. MR681859 (85j:58002)
- [4] E. Bänsch. Finite element discretization of the Navier-Stokes equations with a free capillary surface. *Numer. Math.*, 88(2):203–235, 2001. MR1826211 (2002h:76079)
- [5] A. Bonito, R.H. Nochetto, and M.S. Pauletti. Geometrically consistent mesh modifications. *SIAM J. Numer. Anal.*, 48 (2010), no. 5, 1877–1899. MR2733102 (2011j:65263)
- [6] D. Braess and W. Hackbusch. A new convergence proof for the multigrid method including the V-cycle. *SIAM J. Numer. Anal.*, 20(5):967–975, 1983. MR714691 (85h:65233)
- [7] J.H. Bramble, D.Y. Kwak, and J.E. Pasciak. Uniform convergence of multigrid V-cycle iterations for indefinite and nonsymmetric problems. *SIAM J. Numer. Anal.*, 31(6):1746–1763, 1994. MR1302683 (95i:65170)
- [8] J.H. Bramble, J.E. Pasciak, and J. Xu. The analysis of multigrid algorithms with nonnested spaces or noninherited quadratic forms. *Math. Comp.*, 56(193):1–34, 1991. MR1052086 (91h:65159)
- [9] J.H. Bramble and X. Zhang. The analysis of multigrid methods. In *Handbook of numerical analysis, Vol. VII*, Handb. Numer. Anal., VII, pages 173–415. North-Holland, Amsterdam, 2000. MR1804746 (2001m:65183)
- [10] P.B. Canham. The minimum energy of bending as a possible explanation of the biconcave shape of the human red blood cell. *Journal of Theoretical Biology*, 26(1):61–81, January 1970.
- [11] P. G. Ciarlet and J.-L. Lions, editors. *Handbook of numerical analysis. Vol. II*. Handbook of Numerical Analysis, II. North-Holland, Amsterdam, 1991. Finite element methods. Part 1. MR1115235 (92f:65001)
- [12] A. Demlow. Higher-order finite element methods and pointwise error estimates for elliptic problems on surfaces. *SIAM J. Numer. Anal.*, 47(2):805–827, 2009. MR2485433 (2010a:65233)
- [13] A. Demlow and G. Dziuk. An adaptive finite element method for the Laplace-Beltrami operator on implicitly defined surfaces. *SIAM J. Numer. Anal.*, 45(1):421–442 (electronic), 2007. MR2285862 (2008c:65320)
- [14] G. Doğan, P. Morin, and R.H. Nochetto. A variational shape optimization approach for image segmentation with a Mumford-Shah functional. *SIAM J. Sci. Comput.*, 30(6):3028–3049, 2008. MR2452377 (2009j:68184)
- [15] G. Dziuk. Finite elements for the Beltrami operator on arbitrary surfaces. In *Partial differential equations and calculus of variations*, volume 1357 of *Lecture Notes in Math.*, pages 142–155. Springer, Berlin, 1988. MR976234 (90i:65194)
- [16] G. Dziuk and C. M. Elliott. Finite elements on evolving surfaces. *IMA J. Numer. Anal.*, 27(2):262–292, 2007. MR2317005 (2008c:65253)
- [17] D. Gilbarg and N.S. Trudinger. *Elliptic partial differential equations of second order*, volume 224 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition, 1983. MR737190 (86c:35035)
- [18] E. Hebey. *Nonlinear analysis on manifolds: Sobolev spaces and inequalities*, volume 5 of *Courant Lecture Notes in Mathematics*. New York University, Courant Institute of Mathematical Sciences, New York, 1999.
- [19] W. Helfrich. Elastic properties of lipid bilayers - theory and possible experiments. *Zeitschrift Fur Naturforschung C-A Journal Of Biosciences*, 28:693, 1973.
- [20] M. Holst. Adaptive numerical treatment of elliptic systems on manifolds. *Adv. Comput. Math.*, 15(1-4):139–191 (2002), 2001. A posteriori error estimation and adaptive computational methods. MR1887732 (2003a:65108)

- [21] R. Kornhuber and H. Yserentant. Multigrid methods for discrete elliptic problems on triangular surfaces. *Comput. Vis. Sci.*, 11(4-6):251–257, 2008. MR2425494 (2009i:65236)
- [22] Y.-J. Lee, J. Wu, J. Xu, and L. Zikatanov. A sharp convergence estimate for the method of subspace corrections for singular systems of equations. *Math. Comp.*, 77(262):831–850, 2008. MR2373182 (2008k:65106)
- [23] J. Maes, A. Kunoth, and A. Bultheel. BPX-type preconditioners for second and fourth order elliptic problems on the sphere. *SIAM J. Numer. Anal.*, 45(1):206–222 (electronic), 2007. MR2285851 (2008c:65301)
- [24] K. Mekchay. *Convergence of Adaptive Finite Element Methods*. PhD thesis, University of Maryland, 2005. MR2708247
- [25] 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), no. 274, 625–648. MR2772090
- [26] Maxim A. Olshanskii and Arnold Reusken. A finite element method for surface PDEs: matrix properties. *Numer. Math.*, 114(3):491–520, 2010. MR2570076 (2010m:65289)
- [27] Maxim A. Olshanskii, Arnold Reusken, and Jörg Grande. A finite element method for elliptic equations on surfaces. *SIAM J. Numer. Anal.*, 47(5):3339–3358, 2009. MR2551197 (2010k:65265)
- [28] J. Sokołowski and J.-P. Zolésio. *Introduction to shape optimization*, volume 16 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 1992. Shape sensitivity analysis. MR1215733 (94d:49002)
- [29] T. J. Willmore. *Riemannian geometry*. Oxford Science Publications. The Clarendon Press Oxford University Press, New York, 1993. MR1261641 (95e:53002)

DEPARTMENT OF MATHEMATICS, TEXAS A&M UNIVERSITY, COLLEGE STATION, TEXAS 77843-3368

*E-mail address:* bonito@math.tamu.edu

DEPARTMENT OF MATHEMATICS, TEXAS A&M UNIVERSITY, COLLEGE STATION, TEXAS 77843-3368

*E-mail address:* pasciak@math.tamu.edu