AFEM for Geometric PDE: The Laplace-Beltrami Operator

In memory of Enrico Magenes

Andrea Bonito * J. Manuel Cascón † Pedro Morin ‡ Ricardo H. Nochetto §

December 17, 2012

Abstract

We present several applications governed by geometric PDE, and their parametric finite element discretization, which might yield singular behavior. The success of such discretization hinges on an adequate variational formulation of the Laplace-Beltrami operator, which we describe in detail for polynomial degree 1. We next present a complete a posteriori error analysis which accounts for the usual PDE error as well as the geometric error induced by interpolation of the surface. This leads to an adaptive finite element method (AFEM) and its convergence. We discuss a contraction property of AFEM and show its quasi-optimal cardinality.

1 Introduction

Besides its intrinsic interest in differential geometry [31, 30, 55], the Laplace-Beltrami operator (or surface Laplacian) has received a great deal of attention also in the applied and numerical communities. Basic geometric partial differential equations (PDE) such as the mean curvature flow and surface diffusion appear naturally in materials science modeling [53], whereas Willmore flow is a building block in the dynamics of membranes governed by bending energy [34]. This article is about applications, formulation, Galerkin approximation, and adaptivity for a PDE on a surface γ governed by the Laplace-Beltrami operator Δ_γ, such as

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

One of the major goals of this paper is the design and analysis of *parametric* adaptive finite element methods (AFEM) for (1.1) of polynomial degree 1. Our discussion is based on [12].

The first FEM for the Laplace-Beltrami operator on parametric surfaces is due to G. Dziuk [35], who also developed an optimal *a priori* error analysis accounting for the approximation of the surface and PDE by piecewise linear polynomials. This seminal work was followed by parametric

*Department of Mathematics, Texas A&M University, 3368 TAMU, College Station, TX 77843-3368, USA (bonito@math.tamu.edu). The work of this author was partially supported by NSF Grant DMS-0914977.

†Departamento de Economía e Historia Económica, Universidad de Salamanca, Salamanca 37008, SPAIN. casbar@usal.es The work of this author was partially supported by Secretaría de Estado de Investigación, Desarrollo e Innovación through grant: CGL2011-29396-C03-02 (Spain)

‡Instituto de Matemática Aplicada del Litoral (IMAL), Güemes 3450 and Departamento de Matemática, Facultad de Ingeniería Química, Universidad Nacional del Litoral, 3000 Santa Fe, Argentina (pmorin@santafe-conicet.gov.ar). Partially supported by CONICET through grant PIP 112-200801-02182, Universidad Nacional del Litoral through grants CAI+D 062-312, 062-309, and Agencia Nacional de Promoción Científica y Tecnológica, through grant PICT-2008-0622 (Argentina).

§Department of Mathematics and Institute for Physical Science and Technology, University of Maryland, College Park, MD 20742, USA (rhn@math.umd.edu). The work of this author was partially supported by NSF Grant DMS-1109325, and the General Research Board of the University of Maryland.

FEM for time dependent problems such as the mean curvature flow [36], capillary surfaces [2], surface diffusion [4, 7], Willmore flow [7, 17, 37, 50], fluid biomembranes [16], and fluid membranes with orientational order [9, 10]. The analysis of these methods is largely open, except for graphs [5, 24, 25, 26, 27]. We refer to the survey by K. Deckelnick, G. Dziuk, and Ch. Elliott [27] for some of the early work, including level set and phase field approaches.

A. Demlow and G. Dziuk gave the first *a posteriori* error analysis for piecewise linear polynomials [29], and later A. Demlow extended it to higher polynomial degree [28]. This extension is important in light of applications in fluid dynamics [2] and biomembrane dynamics [17, 16]. O. Lakkis and R.H. Nochetto formulated an a posteriori error analysis for the mean curvature flow of graphs in [44].

Even though *adaptivity theory* for linear elliptic PDE on flat domains in any dimensions and the energy norm is now mature [21, 48, 47, 51], much less is known for elliptic problems on manifolds; we refer to the survey [49] for the state of the art of AFEM on flat domains. For the Laplace-Beltrami operator on graphs we mention the convergence theory of K. Mekchay, P. Morin and R.H. Nochetto [46], whereas for general parametric surfaces and polynomial degree we are only aware of [12]. We expose here results from [12] and restrict them to the particular case of polynomial degree 1 for the sake of clarity.

The purpose of this paper is threefold. We first discuss in §2 several applications of the Laplace-Beltrami operator we have recently developed. This serves as a motivation for the rest of the paper as well as illustration of the significance of adequate formulations and discretizations of rather complex problems which look seemingly untractable. We next discuss parametric FEM for (1.1) on piecewise C^1 surfaces which are merely globally Lipschitz. This is inspired by singularities observed in geometric flows, such as pinching [4, 6, 7, 17], point defects [9], and line tension [38]. This in turn makes it unfeasible to use the signed distance function as in [35, 29, 28]. Our approach, developed in §3 and §4, allows for kinks aligned with the initial mesh, and yields optimal convergence rates even for surfaces which are not piecewise C^2. Our third goal is to present a rather complete discussion of adaptivity theory for AFEM on surfaces. The algorithm reads

AFEM: Given an initial surface-mesh pair $(\Gamma_0, \mathcal{T}_0)$, and parameters $\varepsilon_0 > 0$, $0 < \rho < 1$, and $\omega > 0$, set $k = 0$ and iterate

$[\mathcal{T}_k^+, \Gamma_k^+] = \mathsf{ADAPT\_SURFACE}\,(\mathcal{T}_k, \omega\varepsilon_k)$
$[\mathcal{T}_{k+1}, \Gamma_{k+1}] = \mathsf{ADAPT\_PDE}\,(\mathcal{T}_k^+, \varepsilon_k)$
$\varepsilon_{k+1} = \rho\varepsilon_k$; $k = k + 1$.

AFEM consists of two main modules: ADAPT_PDE is the usual adaptive cycle for flat domains driven by the a posteriori PDE error estimator, whereas ADAPT_SURFACE is a new module that accounts for and controls surface interpolation error. In §5 we discuss the a posteriori error analysis for (1.1) on parametric surfaces, with emphasis on C^1 parametric representations $\mathcal{X} : \Omega \to \mathbb{R}^{d+1}$ of γ and their piecewise linear interpolants $\mathcal{F}_\mathcal{T} : \Omega \to \mathbb{R}^{d+1}$, which describes the polyhedral counterpart $\Gamma = \mathcal{F}_\mathcal{T}(\Omega)$ of γ; hereafter $\Omega \subset \mathbb{R}^d$ is the parametric domain. This interpolation is governed by the *geometric error estimator*

$$\lambda_\Gamma := \|\nabla(\mathcal{X} - \mathcal{F}_\mathcal{T})\|_{L^\infty(\Omega)}. \tag{1.2}$$

The module ADAPT_SURFACE guarantees that its output satisfies $\lambda_{\mathcal{T}_k^+} \le \omega\varepsilon_k$, with ω a parameter small relative to 1. This is critical for ADAPT_PDE to contract, a fundamental property of AFEM shown in §7. We embark on the study of cardinality of AFEM in §8: we first prove that AFEM delivers the best asymptotic convergence rate possible for the given regularity of data γ, f and solution u (Theorem 8.3), and secondly we construct a greedy algorithm that realizes ADAPT_SURFACE (Proposition 8.4). The role of ω is crucial for the theory of §7 and §8. We conclude in §9 with a computational investigation showing that ω must be small indeed to achieve optimal performance of AFEM.

2 Motivation: Geometric PDE

The Laplace-Beltrami operator is ubiquitous in applications involving surfaces that evolve and/or are the domain of an underlying PDE. In order to motivate the study of this operator we mention a few applications where it appears naturally.

2.1 Biomembranes: Modeling and Simulations

Predicting the shape of a cell bounded by a lipid bilayer membrane has inspired a significant body of research in the past forty years ranging from purely mechanical descriptions to advanced mathematical analysis. We consider the Helfrich model for geometric biomembranes [41], which associates to a closed surface γ, describing the biomembrane, the *bending (or Willmore)* energy

$$J(\gamma) = \frac{1}{2}\int_\gamma (H - H_0)^2. \tag{2.1}$$

Hereafter H stands for the *mean curvature* of γ and H_0 is the *spontaneous curvature* induced by the surrounding medium.

Fluid Membranes. We start with $H_0 = 0$. The first variation (or shape derivative) of $J(\gamma)$, subject to volume and area constraints, is given in strong form by [33, 55]

$$\delta_\gamma J(\gamma) = \Big(\Delta_\gamma H + \frac{1}{2}H^3 - 2\kappa H\Big)\boldsymbol{\nu} + \Big(\lambda H\boldsymbol{\nu} + p\boldsymbol{\nu}\Big), \tag{2.2}$$

where κ is the Gaussian curvature of γ, and λ, p are the Lagrange multipliers for the area and volume constraints, respectively. It is important to notice that $\delta_\gamma J(\gamma)$ is a vector field perpendicular to γ because $\boldsymbol{\nu}$ is the unit normal to γ. A (geometric) gradient flow consists of deforming γ in the direction opposite to the shape gradient, namely prescribing a vector velocity $\mathbf{v}$ to γ according to

$$\mathbf{v} = -\delta_\gamma J(\gamma). \tag{2.3}$$

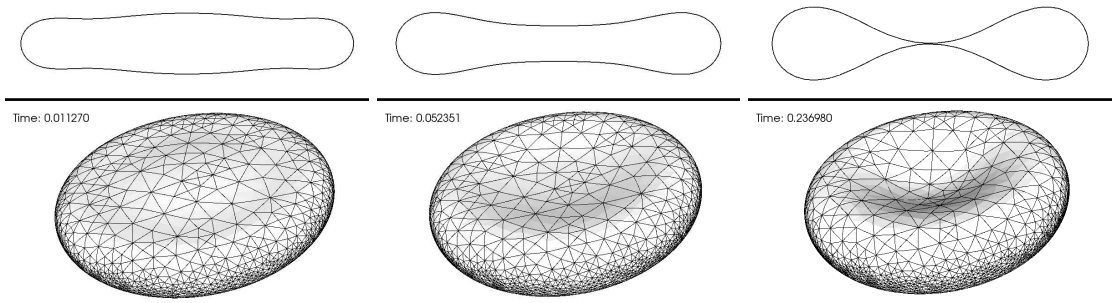


Figure 1: Evolution of an initial axisymmetric ellipsoid of aspect ratio 5x5x1. For each frame the picture on the bottom is a 3D view of the surface mesh and that on the top is a 2D cut through a symmetry plane. The equilibrium is characterized by the formation of an extreme depression of the center to the point of almost pinching (red blood cell). During the evolution the thickening of the outer circular edge occurs faster than the motion on the center, producing a depressed circular ring in between the outer edge and the center (first frame). This in turn is responsible for the appearance of a center bump instead of a depression. Later the evolution continues to squeeze this bump to a depression at the expense of more thickening and rounding of the outer circular edge.

This flow decreases the energy $J(\gamma)$ while keeping area and volume constant, and thus leads to equilibrium configurations such as that in Figure 1, which mimics a *red blood cell*. The simulations

in Figure 1 were performed with the finite element method of A. Bonito, R.H. Nochetto, and M.S. Pauletti [17], which replaces H in (2.2) by the vector curvature $\mathbf{H} = H\boldsymbol{\nu}$ (see also §2.2).

We now consider the more physically realistic model that couples the membrane with a fluid. In order to do this, we assume the simplest situation in which the fluid is Newtonian, and thus is governed by the Navier-Stokes equation for incompressible fluids in the deformable domain Ω_t

$$\begin{aligned} \rho D_t \mathbf{v} - \operatorname{div}(-pI + \mu D(\mathbf{v})) = 0 \qquad & \text{in } \Omega_t, \\ \operatorname{div} \mathbf{v} = 0 \qquad & \text{in } \Omega_t, \end{aligned} \tag{2.4}$$

where $D(\mathbf{v}) = \frac{1}{2}(\nabla \mathbf{v} + \nabla \mathbf{v}^T)$ is the symmetric part of the gradient and $\Sigma = -pI + \mu D(\mathbf{v})$ is the Cauchy stress tensor. The membrane interacts with the fluid only through the boundary condition, which represents a balance of forces at the interface $\gamma = \gamma_t = \partial\Omega_t$:

$$\Sigma \boldsymbol{\nu} = k\, \delta_\gamma J(\gamma), \tag{2.5}$$

where k is the membrane bending rigidity coefficient. In [16] A. Bonito, R.H. Nochetto, and M.S. Pauletti couple the FEM of [17] with a Taylor-Hood discretization of (2.4) in an ALE framework involving a semi-implicit Euler method in time. Figure 2 displays the complex behavior of the fluid membrane and quite noticeable inertial effects, which lead to a more singular pinching than in Figure 1. We give a comparison in Figure 3.

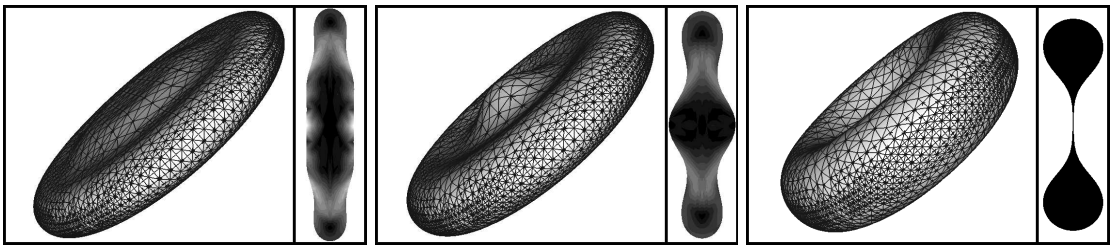


Figure 2: Evolution of a fluid membrane with initial axisymmetric ellipsoidal shape of aspect ratio $5\times5\times1$ and final shape similar to a red blood cell. Each frame shows the membrane mesh and a symmetry cut along a big axis. The fluid flow is quite complex, creating first a bump in the middle and next moving towards the circumference and producing a depresion in the center with flat pinching profile. The inertial effects are due to unrealistic physical parameters.

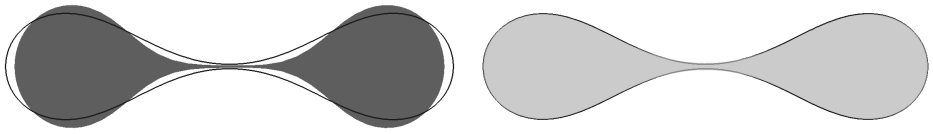


Figure 3: Comparison of final configuration of the geometric biomembrane of Figure 1 and the fluid biomembrane of Figure 2 with unrealistic (left) and realistic (right) physical parameters. For the latter the inertial effects are not significant and the purely geometric evolution is similar to the fluid driven one. The pinching on the left occurs with a much flatter and thinner neck in the center and thicker torus outside.

Director Fields on Flexible Surfaces. The orientation of the bilipids is about 32^o relative to the unit normal to γ for living cells. In order to describe this situation we consider the simple model introduced by S. Bartels, G. Dolzmann, and R.H. Nochetto [9], which is in turn inspired on the model by M. Laradji and O.G. Mouritsen [45] for flat membranes. The starting point is to modify the energy (2.1) to incorporate the effect of a director field $\mathbf{n}$ so that

$$J(\gamma, \mathbf{n}) = \frac{1}{2}\int_\gamma |\operatorname{div}_\gamma \boldsymbol{\nu} - \delta \operatorname{div}_\gamma \mathbf{n}|^2 + \frac{\lambda}{2}\int_\gamma |\nabla_\gamma \mathbf{n}|^2 + \frac{1}{2\varepsilon}\int_\gamma f(\mathbf{n}\cdot\boldsymbol{\nu}), \tag{2.6}$$

with $|\mathbf{n}| = 1$ everywhere in γ. Here div_γ,∇_γ stand for the tangential divergence and gradient to γ, $H = -\mathrm{div}_\gamma \boldsymbol{\nu}$, and $\delta, \lambda > 0$. We thus see that $H_0 = -\delta \mathrm{div}_\gamma \mathbf{n}$ acts as a spontaneous curvature term induced by the director field $\mathbf{n}$. The function $f(x) := (x^2 - \xi_0^2)^2$ in the last term of (2.6) penalizes the deviation of the angle between $\mathbf{n}$ and $\boldsymbol{\nu}$ from $\arccos \xi_0$. It is worth stressing now that if this angle were constant everywhere on γ, then the projection of $\mathbf{n}$ on γ would have a constant length, which in turn would lead to the creation of defects (or singularities) of $\mathbf{n}$. This is due to the topological obstruction that there cannot be a smooth tangential vector field with nonzero constant length defined on a closed surface. Therefore the study of defects and their influence on membrane shape becomes an intriguing matter.

This is precisely what has been accomplished in [9], via an L^2-gradient flow (or relaxation dynamics) for $J(\gamma, \mathbf{n})$:

$$\mathbf{v} = -\delta_\gamma J(\gamma, \mathbf{n}), \qquad \partial_t \mathbf{n} = -\delta_n J(\gamma, \mathbf{n}), \tag{2.7}$$

where $\mathbf{v}$ is the velocity of γ. The expression of $\delta_\gamma J(\gamma, \mathbf{n})$, the first variation of J with respect to γ (or shape derivative) is now much more involved than (2.2), whereas $\delta_n J(\gamma, \mathbf{n})$ is rather simple; we refer to [9] for details. This dynamics involves again the Laplace-Beltrami operator Δ_γ.

We display in Figure 4 the evolution of a sphere γ (first row) along with the director field $\mathbf{n}$ on a plane cutting through north and south poles. The initial director field $\mathbf{n}_0$ has a couple of defects $\pm e^{i\theta}$ of degree $+1$, which persist through the evolution and lead to the formation of cone-like singularities at the poles, one pointing inwards (north pole) and the other outwards (south pole). This configuration shows some analogies to echinocyte shapes observed in lab experiments [42]. We refer to [9] for other examples and discussion, including defects of degree ± 1.

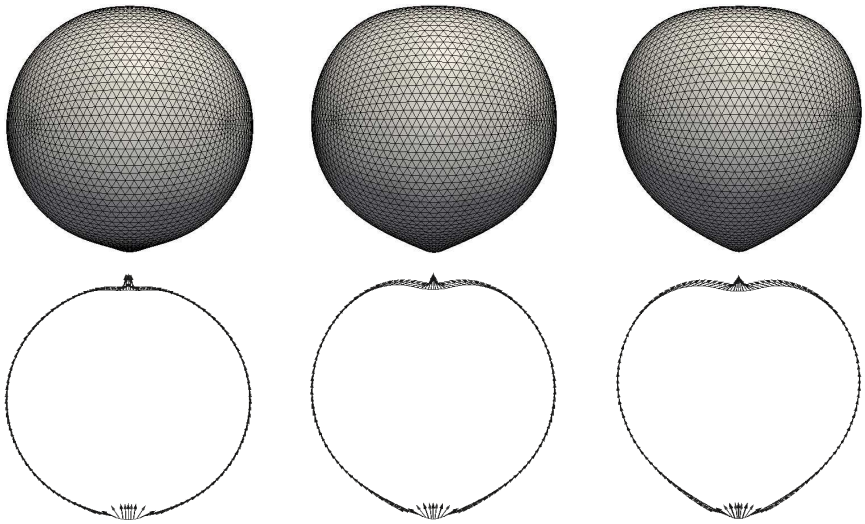


Figure 4: Biomembrane case with inward and outward pointing defects of positive degree one: Snapshots of the surface and the director field along a (deformed) geodesic through the north and south pole after $n = 50, 500, 1400$ time steps. The surface develops inward and outward cones at the poles while the director field remains nearly unchanged during the evolution.

2.2 The Laplace-Beltrami Operator and Curvature

The Laplace-Beltrami operator makes yet another fundamental appearance in the definition and calculation of *curvature*. If $\mathbf{x}$ is the identity on γ, then the following relation for the vector curvature $\mathbf{H} = H\boldsymbol{\nu}$ is well known in differential geometry [30, 31]:

$$\mathbf{H} = -\Delta_\gamma \mathbf{x}. \tag{2.8}$$

This crucial formula was first used for computation by G. Dziuk [35] with piecewise linear finite elements. In the context of geometric evolution of §2.1 we advance in time from t_n to t_{n+1} via a semi-implicit Euler method $\mathbf{x}_{n+1} = \mathbf{x}_n + \tau_n \mathbf{v}_{n+1}$, which keeps the geometry explicit,

$$\int_{\gamma_n} \mathbf{H}_{n+1} \cdot \mathbf{\Psi} - \tau_n \int_{\gamma_n} \nabla_{\gamma_n} \mathbf{v}_{n+1} : \nabla_{\gamma_n} \mathbf{\Psi} = \int_{\gamma_n} \nabla_{\gamma_n} \mathbf{x}_n : \nabla_{\gamma_n} \mathbf{\Psi}. \tag{2.9}$$

This equation for $\mathbf{H}_{n+1}$ is coupled with the equation for velocity $\mathbf{v}_{n+1}$, which comes from the gradient flows (2.3) or (2.7), or the Navier-Stokes equations (2.4)-(2.5). Getting separate equations for $\mathbf{H}_{n+1}$ and $\mathbf{v}_{n+1}$ is effectively an operator splitting technique, introduced by G. Dziuk [36], which has been used in a number of papers; see e.g. [2, 3, 5, 6, 7, 8, 9, 16, 17, 32, 37, 38, 50].

The *mean curvature flow* of a surface γ is governed by $V = -H$, with V being the scalar normal velocity of γ. On the basis of (2.8), this geometric PDE can be reformulated as a heat equation for the position $\mathbf{x}$ on γ, following a seminal idea of G. Dziuk [36]:

$$\partial_t \mathbf{x} = V\boldsymbol{\nu} = -\mathbf{H} = \Delta_\gamma \mathbf{x}.$$

This allows for a simple and efficient finite element discretization [36]. The analysis of the resulting FEM is still open, except for the case of graphs [24, 25, 27].

Expression (2.8) is also a crucial building block in the approach of E. Bänsch to Navier-Stokes equations with *free capillary surfaces* [2]. On the free surface γ, the Cauchy stress tensor Σ satisfies the Young-Laplace equation

$$\boldsymbol{\nu}\Sigma = \mathbf{H},$$

which allows for the following simple and elegant weak formulation of the boundary term

$$\int_\gamma \boldsymbol{\nu}\Sigma\mathbf{w}^T = \int_\gamma \mathbf{H}\mathbf{w}^T = -\int_\gamma \Delta_\gamma \mathbf{x}\mathbf{w}^T = \int_\gamma \nabla_\gamma \mathbf{x} : \nabla_\gamma \mathbf{w}, \tag{2.10}$$

where $\mathbf{w}$ is a suitable test function. This leads again to a simple and efficient FEM [2].

2.3 Surface Diffusion and Epitaxial Films

Surface diffusion is a 4th order geometric driven motion of a surface with normal velocity proportional to the surface Laplacian of mean curvature:

$$V = \Delta_\gamma H. \tag{2.11}$$

This PDE corresponds to the H^{-1} gradient flow of the area functional $J(\gamma) = \int_\gamma 1$, and has been studied by J. Cahn and J. Taylor [20] among others. E. Bänsch, P. Morin, and R.H. Nochetto proposed a parametric FEM upon combining (2.9) and (2.10) [4]. Other related schemes have been developed by J. Barrett, H. Garcke and R. Nürnberg [6, 7]. The analysis of this problem is still open, except for the graph case [5, 27].

Surface diffusion may lead to singularity formation in finite time, depending on the initial configuration. This is depicted in Figures 5-6 which display the evolution of an initial $8 \times 1 \times 1$ prism [4]. This simulation shows that adaptivity is essential to approximate singular situations produced by the flow.

Modeling the deformation of the free surface γ of stressed epitaxial films leads to a variant of (2.11). The stress accounts for the misfit between the crystalline structure of the substrate and epitaxial film, and causes a plastic deformation of γ. This morphological instability of the free surface may eventually lead to crack formation and fracture, an issue of paramount importance in Materials Science. The dynamics of γ is governed by

$$V = \Delta_\gamma(H + \varepsilon),$$

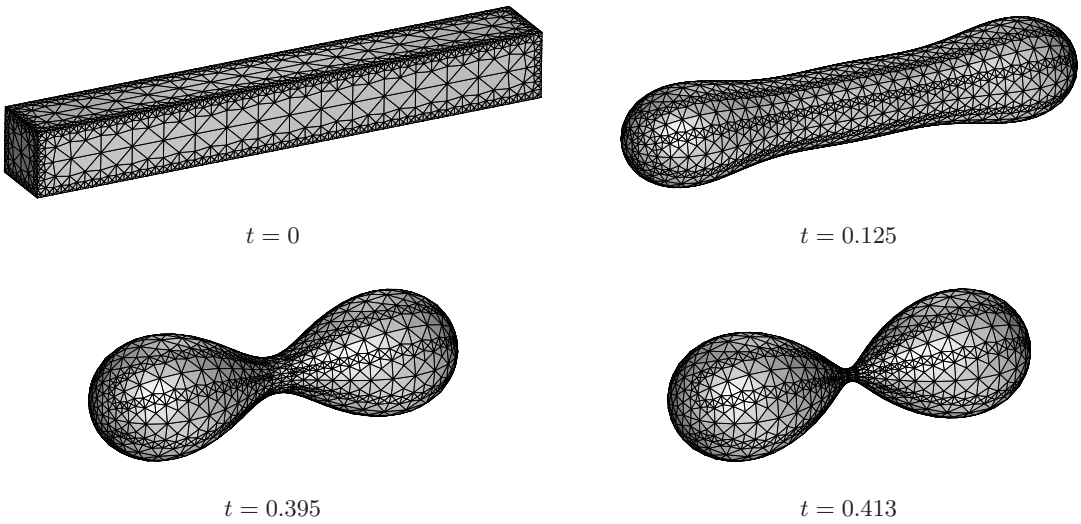


Figure 5: Pinch-off in finite time. Evolution by surface diffusion of an $8 \times 1 \times 1$ prism at various time instants leading to a dumbbell and cusp formation.

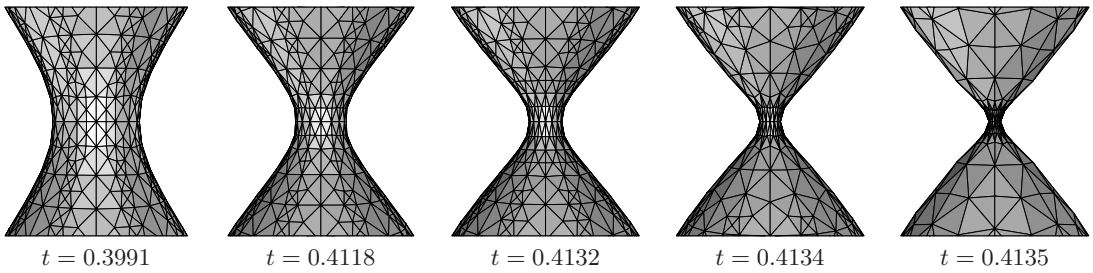


Figure 6: Detailed view of the pinch-off produced by surface diffusion of the $8 \times 1 \times 1$ prism. Adaptivity becomes essential when approaching the pinch-off configuration.

ε being the elastic energy density of the bulk enclosed by γ (see [4, 5, 32] and the references therein). Applications to material science are given in [3, 8].

Consider now a simplified situation where elasticity is replaced by the Laplace operator in the bulk Ω enclosed between the free surface γ and the substrate Σ (see Figure 7). We let $\varepsilon = |\nabla u|^2$, where u solves the problem

$$-\Delta u = 0 \quad \text{in } \Omega, \qquad \partial_\nu u = 0 \quad \text{on } \gamma,$$

and $u = x$ on the bottom Σ and lateral boundary. This yields interesting configurations including mushroom-like formations, thereby leading to defects in materials such as inclusions [32].

2.4 Geometrically Consistent Accuracy Preserving Algorithm

The chief geometric identity (2.8) turns out to play an important role when performing mesh modifications (refinement/coarsening/smoothing) on manifolds with incomplete information on their geometry, yet preserving position and curvature accuracy. This is typically the case when the surface γ is unknown as in the examples provided in Sections 2.1 and 2.3: γ is known only through its approximation Γ and the approximation of its vector curvature $\mathbf{H}$, still labeled $\mathbf{H}$.

The naive approach when performing mesh modification consists of (i) apply the mesh modification to Γ; (ii) compute the corresponding curvature $\mathbf{H}$ according to a discrete version of (2.8)

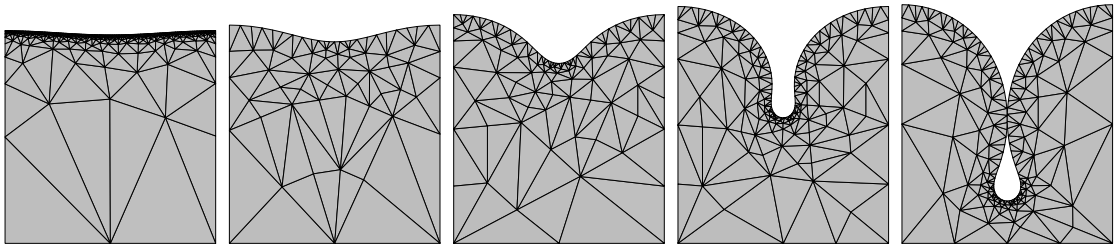


Figure 7: Domain dynamics governed by coupling surface diffusion with the Laplace operator in the bulk. This leads to a mushroom-like free surface that gives rise to an inclusion in finite time.

ensuring geometric consistency (here $\mathbf{X}$ is the identity on Γ). It turns out that the last step yields loss of accuracy on the approximation of the curvature regardless of polynomial degree, which is inherent to computing two derivatives numerically – an unstable process.

To circumvent this issue, A. Bonito, R.H. Nochetto, and S.M. Pauletti [15] propose a Geometrically Consistent Accuracy Preserving Algorithm (GCAP) which reverses the above process: $\mathbf{X}$ is dissociated from Γ itself in that it is no longer the identity on Γ. In essence, the GCAP algorithm proceeds as follows: (i) the mesh modifications are performed on Γ to give the new surface Γ_*; (ii) the new approximation $\mathbf{H}_*$ of vector curvature is obtained *projecting* the existing one $\mathbf{H}$ on Γ_*; (iii) the approximation $\mathbf{X}_*$ of the identity vector on Γ_* is obtained by *solving* the Laplace-Beltrami equation (2.8) discretely with the curvature $\mathbf{H}_*$ given in (ii). We stress that the concatenation of projection and inversion of (2.8) is numerically stable.

To compare the naive and GCAP algorithms, Figure 8 depicts the effect of a global refinement on a square approximation Γ of a circle γ; here $\Gamma_* = \Gamma$. We refer to [15] for similar results for two dimensional surfaces, higher polynomial approximations, and coarsening as well as mesh smoothing.

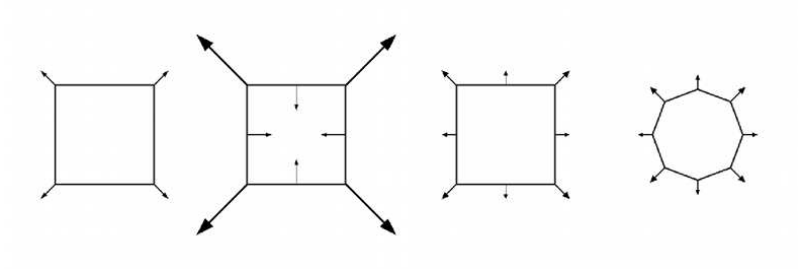


Figure 8: Refinement procedures on a uniform partition of the unit circle using piecewise linear curves. The arrows on the piecewise linear curve represent the approximation of the curvature $\mathbf{H}$, all scaled down by the same multiplicative factor 0.3. We depict the starting approximation of the curvature (first), that after one global bisection of the surface approximation using the naive approach (second), and that with the GCAP method (third). In contrast with them GCAP algorithm, the standard algorithm does not preserve the accuracy of the geometric approximations. The last picture (fourth) depicts the new approximation of curvature over the surface parametrized by the vector $\mathbf{X}_*$ obtained in step (iii) of the GCAP algorithm.

3 Parametric Surfaces

In this section we discuss both how to represent and interpolate a parametric surface. This is instrumental for the design, analysis, and implementation of AFEM on parametric surfaces.

3.1 Representation of Parametric Surfaces

We assume that the surface γ is described as the deformation of a d dimensional polyhedral surface Γ_0 by a globally Lipschitz *homeomorphism* $P_0 : \Gamma_0 \to \gamma \subset \mathbb{R}^{d+1}$. If $\Gamma_0 = \bigcup_{i=1}^{I} \Gamma_0^i$ is made up of I (closed) faces $\Gamma_0^i, i = 1, \ldots, I$, we denote by $P_0^i : \Gamma_0^i \to \mathbb{R}^{d+1}$ the restriction of P_0 to Γ_0^i. We refer to Γ_0^i as a *macro-element* which induces the partition $\{\gamma^i\}_{i=1}^{I}$ of γ upon setting

$$\gamma^i := P_0^i(\Gamma_0^i).$$

In order to avoid technicalities, we assume that all the macro-elements are simplices, i.e. there is a (closed) reference simplex $\Omega \subset \mathbb{R}^d$, from now on called the *parametric domain*, and an affine map $\mathcal{F}_0^i : \mathbb{R}^d \to \mathbb{R}^{d+1}$ such that $\Gamma_0^i = \mathcal{F}_0^i(\Omega)$; Figure 9 sketches the situation when $d = 2$. We thus let $\mathcal{X}^i := P_0^i \circ \mathcal{F}_0^i : \Omega \to \gamma^i$ be a local parametrization of γ which is globally bi-Lipschitz, namely there exists a universal constant $L \geq 1$ such that for all $1 \leq i \leq I$

$$L^{-1}|\hat{x} - \hat{y}| \leq |\mathcal{X}^i(\hat{x}) - \mathcal{X}^i(\hat{y})| \leq L|\hat{x} - \hat{y}|, \qquad \forall \hat{x}, \hat{y} \in \Omega. \tag{3.1}$$

This *minimal regularity* of γ, to be soon strengthened out locally in each macro-element, implies the more familiar condition, valid for a.e. $\hat{x} \in \Omega$,

$$L^{-1}|w| \leq |\widehat{\nabla}\mathcal{X}^i(\hat{x})w| \leq L|w| \qquad \forall\ w \in \mathbb{R}^d; \tag{3.2}$$

hence $L \geq 1$ is the Lipschitz constant of $\mathcal{X}^i$ and so of γ^i. We further assume that $P_0(\mathbf{v}) = \mathbf{v}$ for all vertices $\mathbf{v}$ of Γ_0, so that $\mathcal{F}_0^i$ is the nodal interpolant of $\mathcal{X}^i$ into linear polynomials.

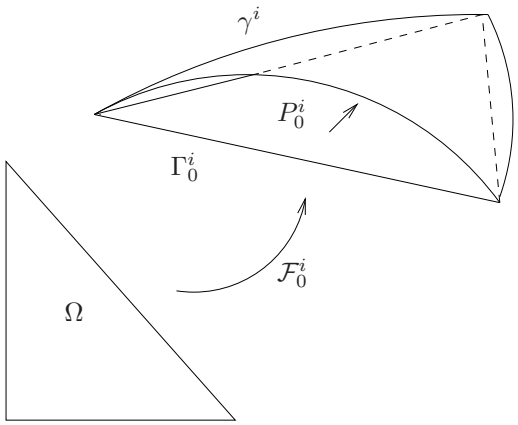


Figure 9: Representation of each component γ^i when $d = 2$ as a parametrization from a flat triangle $\Gamma_0^i \subset \mathbb{R}^3$ as well as from the master triangle $\Omega \subset \mathbb{R}^2$. The map $\mathcal{F}_0^i : \Omega \to \Gamma_0^i$ is affine.

The structure of the map P_0 depends on the application. For instance, if γ^i is described on Γ_0^i via the *distance function* $\mathrm{dist}(x)$ to γ, then

$$\gamma^i \ni \tilde{x} = x - \mathrm{dist}(x)\nabla\,\mathrm{dist}(x) = P_0(x) \qquad \forall\ x \in \Gamma_0^i,$$

provided $\mathrm{dist}(x)$ is sufficiently small so that the distance is uniquely defined. If, instead, γ^i is the *zero level set* $\phi(x) = 0$ of a function ϕ, then

$$\Gamma_0^i \ni x = \tilde{x} + \frac{\nabla\phi(\tilde{x})}{|\nabla\phi(\tilde{x})|}\,|x - \tilde{x}| = P_0^{-1}(\tilde{x}), \qquad \forall\ \tilde{x} \in \gamma^i,$$

is the inverse map of P_0. In both cases, dist and ϕ must be C^2 for P_0 to be $C^1(\Gamma_0^i)$. Yet another option is to view γ^i as a graph on Γ_0^i, in which case P_0^i is a lift in the normal direction to Γ_0^i and

P_0 is $C^1(\Gamma_0^i)$ if and only if γ^i is; we refer to [46]. Notice that the inverse mapping theorem implies $(P_0^i)^{-1} \in C^1(\gamma^i)$.

The *regularity* of γ is expressed in terms of the regularity of the maps $\mathcal{X}^i$. If $s \geq 0$, $2 \leq p \leq \infty$, we say that γ is piecewise W_p^s, and write $\gamma \in W_p^s(\Gamma_0)$, whenever $\mathcal{X}^i \in [W_p^s(\Omega)]^{d+1}$, $i = 1, \ldots, I$. We denote the associated semi-norm by

$$|\gamma|_{W_p^s(\Gamma_0)} := \left(\sum_{i=1}^{I} |\mathcal{X}^i|^p_{W_p^s(\Omega)} \right)^{1/p}.$$

Note that this *non-overlapping* parametrization allows for piecewise smooth surfaces γ with possible kinks matched by the decomposition $\{\gamma^i\}_{i=1}^I$. Similarly, we say that $\gamma \in C^{1,\alpha}(\Gamma_0)$, $0 \leq \alpha \leq 1$, whenever $\mathcal{X}^i \in [C^{1,\alpha}(\Omega)]^{d+1}$, $i = 1, \ldots, I$ and define

$$|\gamma|_{C^{1,\alpha}(\Gamma_0)} := \max_{i=1,..,I} |\mathcal{X}^i|_{C^{1,\alpha}(\Omega)}.$$

Finally, we note that a function $v : \gamma^i \to \mathbb{R}$ defines uniquely two functions $\hat{v} : \Omega \to \mathbb{R}$ and $\bar{v} : \Gamma_0^i \to \mathbb{R}$ via the maps $\mathcal{X}^i$ and P_0, namely

$$\hat{v}(\hat{x}) := v(\mathcal{X}^i(\hat{x})) \quad \forall\, \hat{x} \in \Omega \qquad \text{and} \qquad \bar{v}(\bar{x}) := v(P_0(\bar{x})) \quad \forall\, \bar{x} \in \Gamma_0^i; \tag{3.3}$$

we set $\tilde{x} = \mathcal{X}^i(\hat{x})$ for all $\hat{x} \in \Omega$. Conversely, a function $\hat{v} : \Omega \to \mathbb{R}$ (respectively, $\bar{v} : \Gamma_0^i \to \mathbb{R}$) defines uniquely the two functions $v : \gamma^i \to \mathbb{R}$ and $\bar{v} : \Gamma_0^i \to \mathbb{R}$ (respectively, $v : \gamma^i \to \mathbb{R}$ and $\hat{v} : \Omega \to \mathbb{R}$). We will always denote by v the two lifts $\tilde{v}$ or $\hat{v}$ of $v : \gamma^i \to \mathbb{R}$.

3.2 Interpolation of Parametric Surfaces

The initial partition of Γ_0 in macro-elements (or faces) induces a conforming triangulation $\mathcal{T}_0$ of Γ_0. We only discuss the class of conforming meshes $\mathbb{T}(\mathcal{T}_0)$ created by successive bisections of this initial mesh $\mathcal{T}_0$. However, our results remain valid for any refinement strategy satisfying Conditions 3, 4 and 6 in [14]. In particular, successive bisections, quad-refinement and red-refinement all with hanging nodes are admissible refinement strategies. For more details, we refer to [14, Section 6].

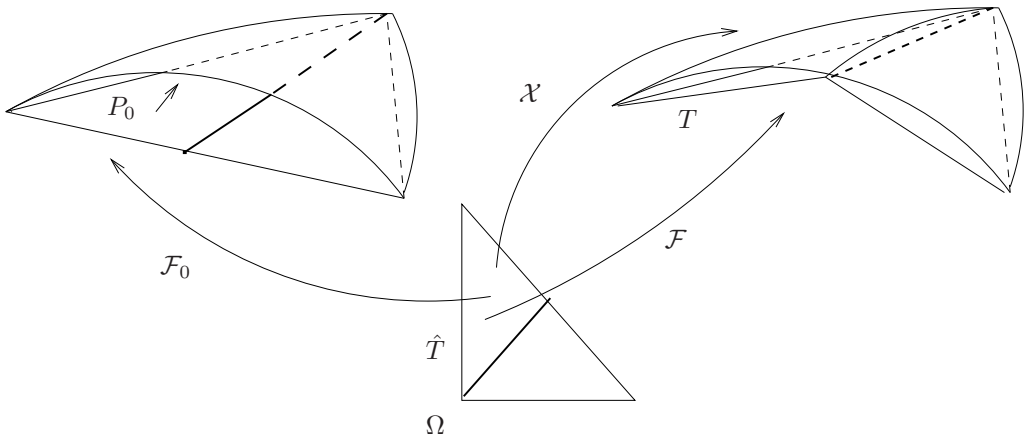


Figure 10: Effect of one bisection of the macro-element $\mathcal{F}_0(\Omega)$ when $d = 2$ (left). The parametric domain Ω is split into two triangles in $\mathbb{R}^2$ via the affine map $\mathcal{F}_0^{-1}$ (bottom), whereas γ is interpolated by a new piecewise linear surface $\Gamma = \mathcal{F}(\Omega)$ (right), with $\mathcal{F} = \mathcal{I}_{\mathcal{T}}\mathcal{X}$ the piecewise linear interpolant of the parametrization $\mathcal{X}$ defined in Ω. The superscript i is omitted for simplicity.

Given $\mathcal{T}_0$, we define a shape regular forest $\mathbb{T}(\mathcal{T}_0)$, and for each $\mathcal{T} \in \mathbb{T}(\mathcal{T}_0)$, a *piecewise affine* approximation $\Gamma = \Gamma(\mathcal{T})$ of γ, and a finite element space $\mathbb{V}(\mathcal{T})$ on Γ as follows. Note first that

conforming graded bisections of each macro-element Γ_0^i induce a family of shape regular partitions $\mathcal{T}^i(\Omega)$ of the parametric domain $\Omega \subset \mathbb{R}^d$. Let $\mathbb{V}(\mathcal{T}^i(\Omega))$ be the finite element space of C^0 piecewise linear polynomials on $\mathcal{T}^i(\Omega)$, and let $\mathcal{I}_{\mathcal{T}^i} : C^0(\Omega) \to \mathbb{V}(\mathcal{T}^i(\Omega))$ be the corresponding Lagrange interpolation operator. Let $\mathcal{F}_{\mathcal{T}^i} = \mathcal{I}_{\mathcal{T}^i}\mathcal{X}^i$ be the interpolant of $\mathcal{X}^i$ in $\mathbb{V}(\mathcal{T}^i(\Omega))$, $\Gamma^i := \mathcal{F}_{\mathcal{T}^i}(\Omega)$ and

$$\mathcal{T}^i := \left\{T = \mathcal{F}_{\mathcal{T}^i}(\widehat{T}) \quad | \quad \widehat{T} \in \mathcal{T}^i(\Omega)\right\};$$

the set Γ^i is a piecewise affine interpolation of γ^i. The global mesh $\mathcal{T}$, piecewise affine surface Γ, and parametrization $\mathcal{F}_{\mathcal{T}}$ of Γ are given by

$$\mathcal{T} := \cup_{i=1}^I \mathcal{T}^i, \qquad \Gamma := \cup_{i=1}^I \Gamma^i, \qquad \mathcal{F}_{\mathcal{T}} := \{\mathcal{F}_{\mathcal{T}^i}\}_{i=1}^I .$$

We need a few properties before discussing shape regularity of $\mathbb{T}(\mathcal{T}_0) = \{\mathcal{T}\}$. We define

$$\mathbb{V}(\mathcal{T}) := \Big\{ V \in C^0(\Gamma) \;\Big|\; V|_{\Gamma^i} \text{ is the lift of some } V \in \mathbb{V}(\mathcal{T}^i(\Omega)) \text{ via } \mathcal{F}_{\mathcal{T}^i},$$
$$V = 0 \text{ on } \partial\Gamma, \text{ or } \int_\Gamma V = 0 \text{ if } \partial\Gamma = \emptyset \Big\},$$

and note that $\mathbb{V}(\mathcal{T})$ is not a subspace of $\mathbb{V}(\mathcal{T}_0)$, which is a lack of consistency we must account for. Since most properties discussed below are valid independently of the superscript i, we omit it from now on. Figure 10 depicts one bisection refinement for $d = 2$.

If $\widehat{T} \in \mathcal{T}(\Omega)$ and $T = \mathcal{F}_{\mathcal{T}}(\widehat{T}) \in \mathcal{T}$, we define the *geometric element indicator*

$$\lambda_\Gamma(T) := \left\| \widehat{\nabla}(\mathcal{X} - \mathcal{F}_{\mathcal{T}}) \right\|_{L^\infty(\widehat{T})}, \tag{3.4}$$

and the corresponding *geometric estimator*

$$\lambda_\Gamma := \max_{T \in \mathcal{T}} \lambda_\Gamma(T). \tag{3.5}$$

Note that two different meshes giving rise to the same surface Γ yield the same λ_Γ, which is thus of pure geometric nature; this explains the subscript Γ. Moreover, $\lambda_\Gamma(T)$ is evaluated in $\widehat{T}$, which belongs to the parametric domain Ω instead of the polyhedral surface Γ. The geometric estimator may not decrease upon each refinement, especially in the pre-asymptotic regime, but the following *quasi-monotonicity* property is valid instead: there exists a constant $\Lambda_0 \geq 1$, depending on $\mathcal{T}_0$, and dimension d, such that

$$\lambda_{\Gamma_*} \leq \Lambda_0 \lambda_\Gamma \tag{3.6}$$

for all conforming refinements $\mathcal{T}_*$ of $\mathcal{T}$ [12, Lemma 3.1]. This result is also valid elementwise.

We recall that $\mathbb{T}(\mathcal{T}_0)$ is the forest of all conforming refinements $\mathcal{T}$ of $\mathcal{T}_0$, denoted $\mathcal{T} \geq \mathcal{T}_0$, obtained by the aforementioned bisection procedure. We say that $\mathbb{T}(\mathcal{T}_0)$ is *shape regular* if there is a constant C_0 only depending on $\mathcal{T}_0$, such that for all $\widehat{T} \in \mathcal{T}(\Omega)$

$$C_0^{-1}|\hat{x} - \hat{y}| \leq |\mathcal{F}_{\mathcal{T}}(\hat{x}) - \mathcal{F}_{\mathcal{T}}(\hat{y})| \leq C_0|\hat{x} - \hat{y}| \qquad \forall\ \hat{x}, \hat{y} \in \widehat{T}. \tag{3.7}$$

Since the forest induced by bisection on the flat parametric domain Ω is shape regular [11, 49, 52], we observe that (3.7) states that the deformation of $\widehat{T} \in \mathcal{T}(\Omega)$ leading to $T \in \mathcal{T}$ does not degenerate. We also point out that (3.7) implies the usual condition on the Jacobian $\widehat{\nabla}\mathcal{F}_{\mathcal{T}}$, valid for a.e. $\hat{x} \in \Omega$

$$C_0^{-1}|w| \leq |\widehat{\nabla}\mathcal{F}_{\mathcal{T}}(\hat{x})w| \leq C_0|w| \qquad \forall w \in \mathbb{R}^d, \tag{3.8}$$

and that $\widehat{\nabla}\mathcal{F}_{\mathcal{T}}$ happens to be constant on $\widehat{T}$ for an affine map $\mathcal{F}_{\mathcal{T}}$ [22].

We stress that a bi-Lipschitz parametrization satisfying (3.1) does *not* guarantee that $\mathbb{T}(\mathcal{T}_0)$ is shape regular. This pathological situation is depicted in Figure 11. This issue has been tackled

by A. Bonito and J. Pasciak [18] assuming that the surface γ is W^2_∞ and $\mathcal{T}_0$ is sufficiently fine. We now present a similar result, invoking piecewise C^1-regularity of γ, which hinges on (3.6): the forest $\mathbb{T}(\mathcal{T}_0)$ is shape-regular with $C_0 = 2L$ provided

$$\lambda_{\Gamma_0} \leq \frac{1}{2\Lambda_0 L}, \tag{3.9}$$

where $L > 1$ is the constant in (3.1) [12, Lemma 3.2]. Figure 11 illustrates an intermediate degenerate situation in which a triangle $T \in \mathcal{T}$ is split into two triangles $T_1, T_2 \in \mathcal{T}_*$ with $\lambda_{\Gamma_*}(T_1) > (2\Lambda_0 L)^{-1}$ and (3.9) being violated. This thereby leads to refinement of T_1, which opens up and gives rise to nondegenerate descendents eventually satisfying (3.9).

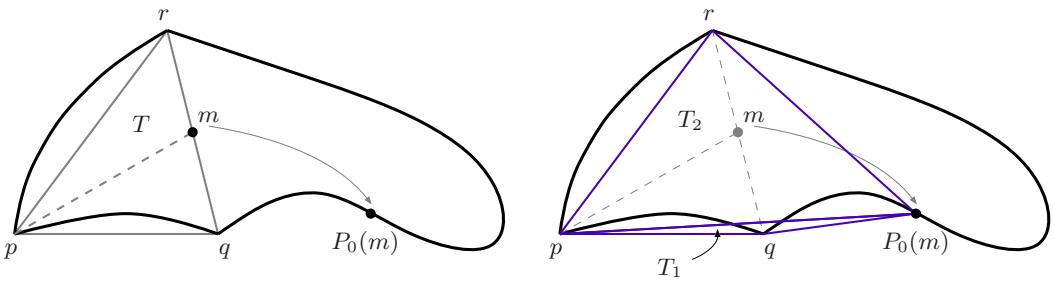


Figure 11: Smooth surface leading to a degenerate triangle. The point $P_0(m)$ is (almost) aligned with p and q. When the triangle $T = pqr \in \mathcal{T}$ is split into pqm and rpm, the new elements of $\mathcal{T}_*$ are $T_1 = pqP_0(m)$, $T_2 = rpP_0(m)$. The triangle T_1 is degenerate and $\lambda_{\Gamma_*}(\widehat{T}_1) > (2\Lambda_0 L)^{-1}$, thus violating (3.9). This forces ADAPT_SURFACE to refine further, which in turn opens up T_1 leading eventually to nondegenerate descendents of T_1.

4 The Laplace-Beltrami Operator

4.1 Basic Differential Geometry

In this subsection we give a matrix formulation of some basic differential geometry facts. We assume γ to be piecewise C^1, namely $\gamma^i \in C^1(\Gamma_0^i)$ for all $1 \leq i \leq I$, and Γ to be piecewise affine.

Our first task is to relate the gradient $\widehat{\nabla}$ in the parametric domain Ω with the tangential gradient ∇_γ on γ. To this end, let $\mathbf{T} \in \mathbb{R}^{(d+1)\times d}$ be the matrix

$$\mathbf{T} := \mathbf{T}_\gamma := [\widehat{\partial}_1 \mathcal{X}, \ldots, \widehat{\partial}_d \mathcal{X}],$$

whose i-th column $\widehat{\partial}_i \mathcal{X} \in \mathbb{R}^{d+1}$ is the vector of partial derivatives of $\mathcal{X}$ with respect to the i^{th} coordinate of Ω. Since $\mathcal{X}$ is a diffeomorphism, the set $\{\widehat{\partial}_i \mathcal{X}\}_{i=1}^d$ of tangent vectors to γ is well defined, linearly independent, and expands the tangent hyperplane to each γ^j at interior points for all $1 \leq j \leq I$. The *first fundamental form* of γ is the symmetric and positive definite matrix $\mathbf{g} \in \mathbb{R}^{d\times d}$ defined by

$$\mathbf{g} = \left(g_{\gamma,ij}\right)_{1\leq i,j\leq d} := \left(\widehat{\partial}_i \mathcal{X}^T \widehat{\partial}_j \mathcal{X}\right)_{1\leq i,j\leq d} = \mathbf{T}^T \mathbf{T}. \tag{4.1}$$

Given $\hat{v}(\hat{x}) = v(\tilde{x})$, the tangent gradient $\nabla_\gamma v(\tilde{x}) = \sum_{i=1}^d \alpha_i(\hat{x}) \widehat{\partial}_i \mathcal{X}(\hat{x})$ satisfies the relation

$$\widehat{\partial}_i \hat{v}(\hat{x}) = \nabla_\gamma v(\tilde{x})\ \widehat{\partial}_i \mathcal{X}(\hat{x}) \quad \text{for } 1 \leq i \leq d,$$

whence

$$\widehat{\nabla} \hat{v} = \nabla_\gamma v\, \mathbf{T} \tag{4.2}$$

and $(\alpha_i)_{i=1}^d = \mathbf{g}^{-1}(\widehat{\partial}_i\hat{v})_{i=1}^d$. To get the reverse relation, we augment $\mathbf{T}$ to the matrix $\widetilde{\mathbf{T}} \in \mathbb{R}^{(d+1)\times(d+1)}$ by adding the (outer) unit normal $\boldsymbol{\nu} = (\nu_1, \cdots, \nu_{d+1}) \in \mathbb{R}^{(d+1)}$ to the tangent hyperplane $\mathrm{span}\{\widehat{\partial}\mathcal{X}_i\}_{i=1}^d$ to γ as the last column, namely

$$\widetilde{\mathbf{T}} := \big[\mathbf{T}, \boldsymbol{\nu}^T\big] = \big[\widehat{\partial}_1\mathcal{X}, \ldots, \widehat{\partial}_d\mathcal{X}, \boldsymbol{\nu}^T\big].$$

Since $\widetilde{\mathbf{T}}$ is invertible, we let $\widetilde{\mathbf{D}} = \widetilde{\mathbf{T}}^{-1}$. We thus realize that

$$\nabla_\gamma v = \nabla_\gamma v\, \widetilde{\mathbf{T}}\, \widetilde{\mathbf{D}} = \big[\widehat{\nabla}\hat{v}, 0\big]\widetilde{\mathbf{D}} = \widehat{\nabla}\hat{v}\ \mathbf{D}, \tag{4.3}$$

where $\mathbf{D} \in \mathbb{R}^{d\times(d+1)}$ results from $\widetilde{\mathbf{D}}$ by cutting off its last row. Moreover, writing

$$\mathbf{I}_{(d+1)\times(d+1)} = \widetilde{\mathbf{T}}^{-1}\widetilde{\mathbf{T}} = \begin{bmatrix} \mathbf{D} \\ \mathbf{v} \end{bmatrix} \begin{bmatrix} \mathbf{T} & \boldsymbol{\nu}^T \end{bmatrix} = \begin{bmatrix} \mathbf{DT} & \mathbf{D}\boldsymbol{\nu}^T \\ \mathbf{vT} & \mathbf{v}\boldsymbol{\nu}^T \end{bmatrix}$$

with $\mathbf{v} \in \mathbb{R}^{d+1}$, we deduce $\mathbf{DT} = \mathbf{I}_{d\times d}$ and $\mathbf{vT} = 0$ whence $\mathbf{v}$ is parallel to $\boldsymbol{\nu}$ and $\mathbf{v} = \boldsymbol{\nu}$ because $\mathbf{v}\boldsymbol{\nu}^T = 1$. Reverting the order of multiplication, we also infer that

$$\mathbf{I}_{(d+1)\times(d+1)} = \widetilde{\mathbf{T}}\widetilde{\mathbf{T}}^{-1} = \begin{bmatrix} \mathbf{T} & \boldsymbol{\nu}^T \end{bmatrix} \begin{bmatrix} \mathbf{D} \\ \boldsymbol{\nu} \end{bmatrix} = \mathbf{TD} + \boldsymbol{\nu}^T\boldsymbol{\nu},$$

and $\mathbf{TD} = \mathbf{I}_{(d+1)\times(d+1)} - \boldsymbol{\nu}^T\boldsymbol{\nu}$. This shows that $\mathbf{TD}$ is symmetric and

$$\mathbf{TDD}^T\mathbf{T}^T = \mathbf{TDTD} = \mathbf{TD} = \mathbf{I}_{(d+1)\times(d+1)} - \boldsymbol{\nu}^T\boldsymbol{\nu}, \tag{4.4}$$

as well as

$$\mathbf{DD}^T\mathbf{T}^T\mathbf{T} = \mathbf{DTDT} = \mathbf{I}_{d\times d}.$$

Therefore, the first fundamental form $\mathbf{g}$ has inverse $\mathbf{g}^{-1} = \mathbf{DD}^T$. We let

$$q := \sqrt{\det \mathbf{g}} \tag{4.5}$$

be the elementary area of γ and point out the change of variables formula for $\omega \subset \Omega$

$$\int_\omega \hat{v}q = \int_{\mathcal{X}(\omega)} v. \tag{4.6}$$

The discussion above applies as well to the piecewise affine surface Γ. We denote the corresponding matrices $\mathbf{T}_\Gamma = \widehat{\nabla}\mathcal{F}_\mathcal{T}$ and $\mathbf{D}_\Gamma$ associated with $\mathcal{F}_\mathcal{T} : \Omega \to \Gamma$, and get

$$\nabla_\Gamma v = \widehat{\nabla}\hat{v}\, \mathbf{D}_\Gamma. \tag{4.7}$$

The first fundamental form $\mathbf{G}_\Gamma$ of Γ and its elementary area Q_Γ are defined by

$$\mathbf{G}_\Gamma := \mathbf{T}_\Gamma^T\, \mathbf{T}_\Gamma, \qquad Q_\Gamma := \sqrt{\det \mathbf{G}_\Gamma}. \tag{4.8}$$

It is worth noticing that, since $\mathcal{F}_\mathcal{T}$ is affine, $\mathbf{G}_\Gamma$ and Q_Γ are constant on each $\widehat{T} \in \mathcal{T}(\Omega)$ ($T \in \mathcal{T}$).

4.2 Variational Formulation and Galerkin Method

We now introduce basic Lebesgue and Sobolev spaces on the surface γ. Let

$$L^2_\#(\gamma) := \Big\{ v \in L^2(\gamma) \;\Big|\; \int_\gamma v = 0 \quad \text{if} \quad \partial\gamma = \emptyset \Big\}$$

be the space of L^2 functions, with vanishing meanvalue whenever the surface γ is closed, and

$$H^1_\#(\gamma) := \Big\{ v \in L^2_\#(\gamma) \;\Big|\; \nabla_\gamma v|_{\gamma^i} \in [L^2(\gamma^i)]^{d+1},\ v|_{\gamma^i} = v|_{\gamma^j} \text{ on } \gamma^i \cap \gamma^j\ 1 \le i,j \le I,\ v = 0 \text{ on } \partial\gamma \Big\},$$

where ∇_γ and traces are well defined in each component γ^i due to (4.3). We define the weak form of the *Laplace-Beltrami operator* $\Delta_\gamma v$ for any function $v \in H^1_\#(\gamma)$ to be

$$\langle -\Delta_\gamma v, \varphi \rangle := \sum_{i=1}^{I} \int_{\gamma^i} \nabla_\gamma v \, \nabla_\gamma^T \varphi \qquad \forall \varphi \in H^1_\#(\gamma), \tag{4.9}$$

where $\langle \cdot, \cdot \rangle$ denotes the $(H^1_\#(\gamma))^*$-$H^1_\#(\gamma)$ duality product. In order to derive a strong form of Δ_γ, we now assume that $\mathcal{X}^i$ is C^2 and $v \in H^2(\gamma^i)$ for each $1 \le i \le d$. In view of (4.3), integrating by parts in Ω we obtain

$$\int_{\gamma^i} \nabla_\gamma v \, \nabla_\gamma^T \varphi = \int_\Omega \widehat{\nabla} \hat{v} \mathbf{D} \, \mathbf{D}^T \widehat{\nabla} \hat{\varphi}^T q = \int_\Omega -\frac{1}{q} \widehat{\mathrm{div}}\big(q \widehat{\nabla} \hat{v} \mathbf{g}^{-1}\big) \hat{\varphi} q + \int_{\partial\Omega} q \widehat{\nabla} \hat{v} \mathbf{g}^{-1} \widehat{\mathbf{n}}^T \hat{\varphi},$$

where $\widehat{\mathbf{n}}$ is the unit outer normal to Ω. We thus discover that inside γ^i the following expression for the Laplace-Beltrami operator holds

$$\Delta_\gamma v = \frac{1}{q} \widehat{\mathrm{div}}\big(q \widehat{\nabla} \hat{v} \mathbf{g}^{-1}\big). \tag{4.10}$$

The boundary term instead leads to jumps across the boundary $\partial\gamma^i$ with other pieces γ^j of γ and can be equivalently written as

$$\int_{\partial\Omega} q \widehat{\nabla} \hat{v} \mathbf{g}^{-1} \widehat{\mathbf{n}}^T \hat{\varphi} = \int_{\partial\gamma^i} \nabla_\gamma v \mathbf{n}^T \varphi, \tag{4.11}$$

where $\mathbf{n}$ in the unit outer normal to γ^i in the tangent plane to γ^i. Combining (4.10) with (4.11) yields

$$\int_{\gamma^i} \nabla_\gamma v \nabla_\gamma^T \varphi = \int_{\gamma^i} -\Delta_\gamma v \, \varphi + \int_{\partial\gamma^i} \nabla_\gamma v \, \mathbf{n}^T \varphi, \tag{4.12}$$

which is the Gauss-Green formula for C^2 surfaces.

Expression (4.11) is not obvious and, since it is quite important for the subsequent discussion, we prove it now. Recall that $\Omega \subset \mathbb{R}^d$ is the canonical unit simplex and notice that a change of variables in Ω dictated by a rotation leaves the left-hand side of (4.11) unchanged. We exploit this property to assume, for convenience, that an arbitrary $\hat{x} \in \partial\Omega$ belongs to the $(d-1)$-subsimplex $\widehat{S}$ with outer normal given by $\widehat{\mathbf{n}} = [-1, 0, \cdots, 0]$. We observe that the affine function $\widehat{\phi}(\hat{x}) = \hat{x}\widehat{\mathbf{n}}^T$ vanishes on $\widehat{S}$ and $\widehat{\nabla}\widehat{\phi} = \widehat{\mathbf{n}} = \nabla_\gamma \phi \, \mathbf{T}$, according to (4.2), whence

$$\nabla_\gamma \phi \, \widehat{\partial}_1 \mathcal{X} = -1, \quad \nabla_\gamma \phi \, \widehat{\partial}_i \mathcal{X} = 0 \quad 2 \le i \le d;$$

moreover, $\nabla_\gamma \phi = |\nabla_\gamma \phi| \mathbf{n}$. We now introduce the auxiliary matrix $\mathbf{S} \in \mathbb{R}^{(d+1)\times(d-1)}$

$$\mathbf{S} = [\widehat{\partial}_2 \mathcal{X}, \cdots, \widehat{\partial}_d \mathcal{X}], \qquad r = \sqrt{\det(\mathbf{S}^T \mathbf{S})},$$

and point out that the quantity r is the elementary area associated with the subsimplex $\widehat{S}$ at $\hat{x}$. Since the $(d-1)$-dimensional space $\mathrm{span}\{\widehat{\partial}_i \mathcal{X}\}_{i=2}^d$ is tangent to the curvilinear simplex $\widetilde{S} = \mathcal{X}(\widehat{S})$, we can decompose $\widehat{\partial}_1 \mathcal{X}$ orthogonally as follows

$$\widehat{\partial}_1 \mathcal{X} = \alpha \mathbf{n} + \mathbf{S}\mathbf{b}, \qquad \alpha \in \mathbb{R}, \ \mathbf{b} \in \mathbb{R}^{d-2},$$

where $\mathbf{b}$ is the least squares solution $\mathbf{b} = (\mathbf{S}^T \mathbf{S})^{-1} \mathbf{S}^T \widehat{\partial}_1 \mathcal{X}^T$ and $|\widehat{\partial}_1 \mathcal{X}|^2 = \alpha^2 + |\mathbf{S}\mathbf{b}|^2$. Hence,

$$-1 = \nabla_\gamma \phi \, \widehat{\partial}_1 \mathcal{X} = \alpha |\nabla_\gamma \phi| \quad \Rightarrow \quad \alpha = -|\nabla_\gamma \phi|^{-1}.$$

We compute $q^2 = \det \mathbf{g}$ using the expression for block matrices

$$\mathbf{g} = \begin{bmatrix} |\widehat{\partial}_1 \mathcal{X}|^2 & \widehat{\partial}_1 \mathcal{X}\, \mathbf{S} \\ \mathbf{S}^T \widehat{\partial}_1 \mathcal{X}^T & \mathbf{S}^T \mathbf{S} \end{bmatrix} \quad \Longrightarrow \quad \det \mathbf{g} = \det(\mathbf{S}^T \mathbf{S})\, \big(|\widehat{\partial}_1 \mathcal{X}|^2 - \widehat{\partial}_1 \mathcal{X}\, \mathbf{S} (\mathbf{S}^T \mathbf{S})^{-1}\, \mathbf{S}^T \widehat{\partial}_1 \mathcal{X}^T \big),$$

to infer that

$$q^2 = r^2 \alpha^2 \quad \Longrightarrow \quad |\nabla_\gamma \phi| = \frac{r}{q} \quad \Longrightarrow \quad \widehat{\mathbf{n}} = \mathbf{n}\, \mathbf{T} \frac{r}{q}.$$

To finally derive (4.11), we recall that $\mathbf{TD} = \mathbf{I}_{(d+1)\times(d+1)} - \boldsymbol{\nu}^T \boldsymbol{\nu}$ and observe that

$$\int_{\widehat{S}} q \widehat{\nabla} \hat{v} \mathbf{g}^{-1} \widehat{\mathbf{n}}^T \hat{\varphi} = \int_{\widehat{S}} q \widehat{\nabla} \hat{v} \mathbf{D} \mathbf{D}^T \widehat{\mathbf{n}}^T \hat{\varphi} = \int_{\widehat{S}} r \nabla_\gamma v \mathbf{D}^T \mathbf{T}^T \mathbf{n}^T \varphi = \int_{\widetilde{S}} \nabla_\gamma v\, \mathbf{n}^T \varphi.$$

We now build on (4.9) and write the weak formulation of $-\Delta_\gamma u = f$ as follows: given $f \in L^2_\#(\gamma)$, we seek $u \in H^1_\#(\gamma)$ satisfying

$$\sum_{i=1}^{I} \int_{\gamma^i} \nabla_\gamma u\, \nabla_\gamma^T \varphi = \int_\gamma f\, \varphi, \qquad \forall\, \varphi \in H^1_\#(\gamma). \tag{4.13}$$

Existence and uniqueness of a solution $u \in H^1_\#(\gamma)$ is a consequence of the Lax-Milgram theorem provided γ is Lipschitz. Combining (4.13) with (4.12) and (4.11) yields for each component γ^i,

$$-\Delta_{\gamma^i} u = f \quad 1 \le i \le I, \tag{4.14}$$

together with vanishing jump conditions at the interfaces $\gamma^i \cap \gamma^j$

$$\mathcal{J}(u)|_{\gamma^i \cap \gamma^j} = \nabla_{\gamma^i} u\, \mathbf{n}^i + \nabla_{\gamma^j} u\, \mathbf{n}^j = 0 \qquad \forall\, 1 \le i, j \le I, \tag{4.15}$$

because $f \in L^2_\#(\gamma)$ cannot balance this singular term otherwise.

We next formulate an approximation to the Laplace-Beltrami operator on a piecewise affine approximation Γ of γ supported by a mesh $\mathcal{T} \in \mathbb{T}(\mathcal{T}_0)$. If $F_\Gamma \in L^2_\#(\Gamma)$ is a suitable approximation of f, then the finite element solution $U : \Gamma \to \mathbb{R}$ solves

$$U \in \mathbb{V}(\mathcal{T}) : \quad \int_\Gamma \nabla_\Gamma U \nabla_\Gamma^T V = \int_\Gamma F_\Gamma\, V \qquad \forall\, V \in \mathbb{V}(\mathcal{T}). \tag{4.16}$$

To this end we choose F_Γ to be

$$F_\Gamma := f \frac{q}{Q_\Gamma}, \tag{4.17}$$

because this specific choice of F_Γ satisfies the compatibility property

$$\int_\Gamma F_\Gamma = \int_\gamma f = 0, \tag{4.18}$$

whenever γ is closed, and allows us to handle separately the approximation of surface γ and forcing f. In particular, (4.16) admits a unique solution U as a consequence of the Lax-Milgram theorem. Since Γ is piecewise affine, the quantities $\widehat{\nabla} \widehat{U}, \mathbf{G}_\Gamma, Q_\Gamma$ are piecewise constant, whence

$$\Delta_\Gamma U|_T = 0 \qquad \forall\, T \in \mathcal{T}. \tag{4.19}$$

We refer to [12] where we account for piecewise polynomial Γ and the fact that $\Delta_\Gamma U|_T \neq 0$. The formula (4.12) extends to every element $T \in \mathcal{T}$:

$$\int_T \nabla_\Gamma U \nabla_\Gamma^T V = \int_T -\Delta_\Gamma U\, V + \int_{\partial T} \nabla_\Gamma U\, \mathbf{n}_T^T V \qquad \forall\, V \in \mathbb{V}(\mathcal{T}). \tag{4.20}$$

5 A Posteriori Error Analysis

In order to study the discrepancy between u and U we need to agree on comparing them in a common domain, say γ. Our goal is thus to obtain a posteriori error estimates for the energy error $\|\nabla_\gamma(u-U)\|_{L^2(\gamma)}$. This requires developing an a priori error analysis for the interpolation error committed in replacing γ by Γ in (4.16), which is a sort of consistency error, and its impact on the PDE error. We are concerned with these issues in this section and refer to [28, 29, 46].

5.1 Geometric Error and Estimator

We now quantify the error arising from approximating γ, the so-called *geometric error*. To this end we resort to the matrix formulation of §4.1 to relate the geometric error with the geometric estimator λ_Γ of (3.4).

Given $T \in \mathcal{T}$, we will deal with the regions $\widehat{T} \in \mathcal{T}(\Omega)$ and $\widetilde{T} \in \mathcal{T}(\gamma)$ given by

$$\widehat{T} := \left\{\mathcal{F}_T^{-1}(x) \mid x \in T\right\} \qquad \text{and} \qquad \widetilde{T} := \left\{\mathcal{X}(\hat{x}) \mid \hat{x} \in \widehat{T}\right\}. \tag{5.1}$$

On mapping back and forth to $\widehat{T}$, and using (4.6), we easily see that

$$\int_T v = \int_{\widetilde{T}} v \frac{Q_\Gamma}{q}. \tag{5.2}$$

We are now able to quantify the consistency error alluded to at the beginning of this section.

Lemma 5.1 (consistency error)**.** *For all $v, w \in H^1(\gamma)$ there holds*

$$\int_\Gamma \nabla_\Gamma v \nabla_\Gamma^T w - \int_\gamma \nabla_\gamma v \nabla_\gamma^T w = \int_\gamma \nabla_\gamma v \mathbf{E}_\Gamma \nabla_\gamma^T w,$$

where $\mathbf{E}_\Gamma \in \mathbb{R}^{(d+1)\times(d+1)}$ stands for the following error matrix

$$\mathbf{E}_\Gamma := \frac{1}{q}\mathbf{T}(Q_\Gamma \mathbf{G}_\Gamma^{-1} - q\mathbf{g}^{-1})\mathbf{T}^T. \tag{5.3}$$

Proof. We first note that combining (4.2) with (4.3), we get

$$\nabla_\gamma v = \nabla_\Gamma v \ \mathbf{T}_\Gamma \mathbf{D} \quad \text{and} \quad \nabla_\Gamma v = \nabla_\gamma v \ \mathbf{T}\mathbf{D}_\Gamma, \tag{5.4}$$

which together with (5.2) gives

$$\int_\Gamma \nabla_\Gamma v \ \nabla_\Gamma^T w = \int_\gamma \nabla_\gamma v \ \mathbf{T}\mathbf{D}_\Gamma \mathbf{D}_\Gamma^T \mathbf{T}^T \ \nabla_\gamma^T w \frac{Q_\Gamma}{q} \qquad \forall v, w \in H^1(\gamma). \tag{5.5}$$

Since (4.4) allows us to write

$$\int_\gamma \nabla_\gamma v \ \nabla_\gamma^T w = \int_\gamma \nabla_\gamma v \ \mathbf{T}\mathbf{D}\mathbf{D}^T \mathbf{T}^T \ \nabla_\gamma^T w \qquad \forall v, w \in H^1(\gamma), \tag{5.6}$$

which is a counterpart of (5.5), the assertion follows immediately from (4.1) and (4.8). □

Our next task is to estimate $\mathbf{E}_\Gamma$ in (5.3), which entails dealing with $\mathbf{g}, \mathbf{G}_\Gamma$ and q, Q_Γ.

Lemma 5.2 (properties of $\mathbf{G}_\Gamma$ and Q_Γ)**.** *The matrices $\mathbf{g}$ and $\mathbf{G}_\Gamma$ have eigenvalues in the interval $[L^{-2}, L^2]$ and $[\frac{1}{2}L^{-2}, \frac{3}{2}L^2]$, respectively, provided the initial mesh $\mathcal{T}_0$ satisfies*

$$\lambda_{\Gamma_0} \leq \frac{1}{6\Lambda_0 L^3}. \tag{5.7}$$

Moreover, the forest $\mathbb{T}(\mathcal{T}_0)$ is shape regular, $L^{-d} \lesssim q, Q_\Gamma \lesssim L^d$, and for all $\mathcal{T} \in \mathbb{T}(\mathcal{T}_0)$

$$\|q - Q_\Gamma\|_{L^\infty(\gamma)} + \|\mathbf{g} - \mathbf{G}_\Gamma\|_{L^\infty(\gamma)} \lesssim \lambda_\Gamma. \tag{5.8}$$

Proof. Since $L \geq 1$, (5.7) yields (3.9), which in turn gives shape regularity of the forest $\mathbb{T}(\mathcal{T}_0)$ and (3.8) with constant $C_0 = 2L$. Hence, using the definitions of $\mathbf{g}$ and $\mathbf{G}_\Gamma$, we deduce $\|\mathbf{g}-\mathbf{G}_\Gamma\|_{L^\infty(\gamma)} \leq 3L\lambda_\Gamma$. On the other hand, invoking (3.2) we see that $\boldsymbol{\xi}^T\mathbf{g}\boldsymbol{\xi} = \left|\nabla\mathcal{X}\boldsymbol{\xi}\right|^2$ for all $\boldsymbol{\xi} \in \mathbb{R}^d$, whence

$$L^{-2}|\boldsymbol{\xi}|^2 \leq \boldsymbol{\xi}^T\mathbf{g}\boldsymbol{\xi} \leq L^2|\boldsymbol{\xi}|^2.$$

Since $\lambda_\Gamma \leq \frac{1}{6L^3}$, due to (5.7) and (3.6), then the previous estimates readily imply

$$\frac{1}{2}L^{-2}|\boldsymbol{\xi}|^2 \leq \big(L^{-2} - 3L\lambda_\Gamma\big)|\boldsymbol{\xi}|^2 \leq \boldsymbol{\xi}^T\mathbf{G}_\Gamma\boldsymbol{\xi} \leq \big(L^2 + 3L\lambda_\Gamma\big)|\boldsymbol{\xi}|^2 \leq \frac{3}{2}L^2|\boldsymbol{\xi}|^2,$$

as well as $L^{-d} \lesssim q, Q_\Gamma \lesssim L^d$ because $q^2 = \det\mathbf{g}, Q_\Gamma^2 = \det\mathbf{G}_\Gamma$ are products of the d eigenvalues of $\mathbf{g}, \mathbf{G}_\Gamma$. Moreover, since

$$q - Q_\Gamma = \frac{\det\mathbf{g} - \det\mathbf{G}_\Gamma}{q + Q_\Gamma},$$

it only remains to obtain an estimate for the numerator. The definition of determinant readily yields $\big|\det\mathbf{g} - \det\mathbf{G}\big| \lesssim L^{2d-1}\lambda_\Gamma$, and completes the proof. □

We stress that if $\mathcal{T}_0$ does not satisfy (5.7) but $\varepsilon_0 \leq (6\Lambda_0 L^3\omega)^{-1}$, then the algorithm AFEM of §1 will first refine $\mathcal{T}_0$ to make it comply with (5.7) without ever solving the PDE. In this sense, (5.7) is not a serious restriction for AFEM, although necessary for the subsequent theory.

Corollary 5.3 (estimate of $\mathbf{E}_\Gamma$)**.** *If λ_{Γ_0} satisfies* (5.7)*, then we have for all $\mathcal{T} \in \mathbb{T}(\mathcal{T}_0)$ and corresponding Γ*

$$\|\mathbf{E}_\Gamma\|_{L^\infty(\widehat{T})} \lesssim \lambda_\Gamma(T) \qquad \forall\, T \in \mathcal{T},$$

where the hidden constant depends on $\mathcal{T}_0$ and the Lipschitz constant L of γ.

Proof. According to (5.3), and $\|\mathbf{T}\|_{L^\infty(\gamma)} = \|\mathbf{T}^T\|_{L^\infty(\gamma)} \leq L$, we infer that

$$\|\mathbf{E}_\Gamma\|_{L^\infty(\widehat{T})} \lesssim \big\|Q_\Gamma\mathbf{G}_\Gamma^{-1} - q\mathbf{g}^{-1}\big\|_{L^\infty(\widehat{T})}.$$

The lower bounds on the eigenvalues of $\mathbf{g}$ and $\mathbf{G}_\Gamma$ imply $\|\mathbf{g}^{-1}\|_{L^\infty(\widehat{T})}, \|\mathbf{G}_\Gamma^{-1}\|_{L^\infty(\widehat{T})} \lesssim L^2$, which together with the expression

$$Q_\Gamma\mathbf{G}_\Gamma^{-1} - q\mathbf{g}^{-1} = (Q_\Gamma - q)\mathbf{G}_\Gamma^{-1} + q\mathbf{G}_\Gamma^{-1}(\mathbf{g} - \mathbf{G}_\Gamma)\mathbf{g}^{-1}$$

and (5.8) gives the asserted estimate. □

We now give a constructive expression for unit normals in $\mathbb{R}^{d+1}$, thereby generalizing the usual vector product in $\mathbb{R}^3$, and next use it to derive an error estimate for $\mathbf{D}_\Gamma$.

Lemma 5.4 (unit normal)**.** *Let $\{\mathbf{e}_j\}_{j=1}^{d+1}$ be the canonical unit vectors of $\mathbb{R}^{d+1}$. For each $\widehat{x} \in \Omega$, and $x = \mathcal{X}(\widehat{x}) \in \gamma$, let $\mathbf{N}(\widehat{x}) = \sum_{j=1}^{d+1} A_j(\widehat{x})\mathbf{e}_j$, where A_j stands for the determinant*

$$A_j(\widehat{x}) := \det\big(\mathbf{e}_j, \widehat{\partial}_1\mathcal{X}(\widehat{x}), \cdots, \widehat{\partial}_d\mathcal{X}(\widehat{x})\big)$$

We then have $|\mathbf{N}(\widehat{x})| = q(\widehat{x})$ and the unit normal vector $\boldsymbol{\nu}(x)$ to γ at x is given by $\boldsymbol{\nu}(x) = \mathbf{N}(\widehat{x})/\big|\mathbf{N}(\widehat{x})\big|$. Moreover, a similar result holds true also for Γ, upon replacing $\mathcal{X}$ by $\mathcal{F}_\mathcal{T}$, provided λ_{Γ_0} satisfies (5.7)*, i. e., $|\mathbf{N}_\Gamma(\widehat{x})| = Q_\Gamma(\widehat{x})$ and $\boldsymbol{\nu}_\Gamma(x) = \mathbf{N}_\Gamma(\widehat{x})/|\mathbf{N}_\Gamma(\widehat{x})|$.*

Proof. We fix $\widehat{x} \in \Omega$ and drop it from the notation. Since $\mathbf{T}$ is full rank, some A_j must be non-zero whence $\mathbf{N} \neq 0$. Moreover, the vector $\mathbf{N}$ is orthogonal to the tangent hyperplane to γ at x because

$$\mathbf{N} \cdot \widehat{\partial}_i\mathcal{X} = \sum_{j=1}^{d+1} A_j\mathbf{e}_j \cdot \widehat{\partial}_i\mathcal{X} = \det\big(\widehat{\partial}_i\mathcal{X}, \widehat{\partial}_1\mathcal{X}, \cdots, \widehat{\partial}_i\mathcal{X}, \cdots, \widehat{\partial}_d\mathcal{X}\big) = 0.$$

Hence, $\boldsymbol{\nu} = \mathbf{N}/|\mathbf{N}|$ is well defined. To prove that $|\mathbf{N}| = q$ recall that $\mathbf{T} = \left[\widehat{\partial}_1\mathcal{X}, \cdots, \widehat{\partial}_d\mathcal{X}\right]$ to write

$$|\mathbf{N}|^2 = \sum_{j=1}^{d+1} A_j^2 = \sum_{j=1}^{d+1} A_j \det\left(\mathbf{e}_j, \widehat{\partial}_1\mathcal{X}, \cdots, \widehat{\partial}_d\mathcal{X}\right) = \det\left(\mathbf{N}, \widehat{\partial}_1\mathcal{X}, \cdots, \widehat{\partial}_d\mathcal{X}\right)$$
$$= \left\{\det\left([\mathbf{N},\mathbf{T}]^T[\mathbf{N},\mathbf{T}]\right)\right\}^{1/2} = \left\{\det\begin{bmatrix}\mathbf{N}^T\mathbf{N} & 0\\ 0 & \mathbf{T}^T\mathbf{T}\end{bmatrix}\right\}^{1/2} = |\mathbf{N}|\, q.$$

This implies $|\mathbf{N}| = q$ because $|\mathbf{N}| \neq 0$. The same argument applies to Γ. □

Lemma 5.5 (error of $\boldsymbol{\nu}$ and $\mathbf{D}$)**.** *If* (5.7) *holds for the initial mesh* $\mathcal{T}_0$*, then for all* $\mathcal{T} \in \mathbb{T}(\mathcal{T}_0)$

$$\|\boldsymbol{\nu} - \boldsymbol{\nu}_\Gamma\|_{L^\infty(\gamma)} + \|\mathbf{D} - \mathbf{D}_\Gamma\|_{L^\infty(\gamma)} \lesssim \lambda_\Gamma. \tag{5.9}$$

Proof. Lemmas 5.2 and 5.4 imply $L^{-d} \lesssim |\mathbf{N}(\widehat{x})|, |\mathbf{N}_\Gamma(\widehat{x})| \lesssim L^d$ for all $\widehat{x} \in \Omega$, whence

$$\boldsymbol{\nu} - \boldsymbol{\nu}_\Gamma = \frac{\mathbf{N}}{|\mathbf{N}|} - \frac{\mathbf{N}_\Gamma}{|\mathbf{N}_\Gamma|} = \frac{1}{|\mathbf{N}|}(\mathbf{N} - \mathbf{N}_\Gamma) + \left(\frac{1}{|\mathbf{N}|} - \frac{1}{|\mathbf{N}_\Gamma|}\right)\mathbf{N}_\Gamma \quad \Rightarrow \quad \big|\boldsymbol{\nu} - \boldsymbol{\nu}_\Gamma\big| \lesssim L^d\big|\mathbf{N} - \mathbf{N}_\Gamma\big|.$$

To estimate $\mathbf{N} - \mathbf{N}_\Gamma = \sum_{j=1}^{d+1}\big(A_j - A_{\Gamma,j}\big)\mathbf{e}_j$, we observe that each A_j (resp. $A_{\Gamma,j}$) is a sum of factors of the form $\widehat{\partial}_i\mathcal{X}\cdot\mathbf{e}_m$ (resp. $\widehat{\partial}_i\mathcal{F}_\mathcal{T}\cdot\mathbf{e}_m$), whence

$$\big|A_j - A_{\Gamma,j}\big| \lesssim L^{d-1}\lambda_\Gamma \qquad \Rightarrow \quad |\boldsymbol{\nu} - \boldsymbol{\nu}_\Gamma| \lesssim L^{2d-1}\lambda_\Gamma.$$

For the remaining estimate for $\mathbf{D} - \mathbf{D}_\Gamma$ we recall the definition $\widetilde{\mathbf{T}} = [\mathbf{T}, \boldsymbol{\nu}^T]$ to infer that

$$\|\widetilde{\mathbf{T}} - \widetilde{\mathbf{T}}_\Gamma\|_{L^\infty(\gamma)} \leq \|\mathbf{T} - \mathbf{T}_\Gamma\|_{L^\infty(\gamma)} + \|\boldsymbol{\nu} - \boldsymbol{\nu}_\Gamma\|_{L^\infty(\gamma)} \lesssim \lambda_\Gamma.$$

We now show that $\widetilde{\mathbf{D}} = \widetilde{\mathbf{T}}^{-1}$ is uniformly bounded. To see this, we write $\widetilde{\mathbf{T}}\tilde{\mathbf{w}} = \mathbf{T}\mathbf{w} + w_{d+1}\boldsymbol{\nu}^T$ for $\tilde{\mathbf{w}} = (\mathbf{w}, w_{d+1}) \in \mathbb{R}^{d+1}$ and recall (3.2) to get

$$L^{-2}|\tilde{\mathbf{w}}|^2 \leq L^{-2}|\mathbf{w}|^2 + |w_{d+1}|^2 \leq \big|\widetilde{\mathbf{T}}\tilde{\mathbf{w}}\big|^2 \leq L^2|\mathbf{w}|^2 + |w_{d+1}|^2 \leq L^2|\tilde{\mathbf{w}}|^2,$$

as well as $\|\widetilde{\mathbf{T}}^{-1}\|_{L^\infty(\Omega)}, \|\widetilde{\mathbf{T}}_\Gamma^{-1}\|_{L^\infty(\Omega)} \lesssim L^2$. Since $\widetilde{\mathbf{D}} - \widetilde{\mathbf{D}}_\Gamma = \widetilde{\mathbf{T}}^{-1}\big(\widetilde{\mathbf{T}}_\Gamma - \widetilde{\mathbf{T}}\big)\widetilde{\mathbf{T}}_\Gamma^{-1}$, the desired estimate follows immediately from the previous one for $\widetilde{\mathbf{T}} - \widetilde{\mathbf{T}}_\Gamma$. □

We finally point out the equivalence of norms on γ and Γ provided (5.7) is valid [12, Lemma 5.6]

$$\|v\|_{L^2(\widetilde{T})} \approx \|v\|_{L^2(T)}, \quad |v|_{L^2(\widetilde{T})} \approx |v|_{L^2(T)} \quad \forall\, T \in \mathcal{T}. \tag{5.10}$$

5.2 Upper and Lower Bounds for the Energy Error

We now derive an error representantion formula leading to lower and upper bounds for the energy error. Given $\mathcal{T} \in \mathbb{T}(\mathcal{T}_0)$, we let the usual interior and jump residual for $V \in \mathbb{V}(\mathcal{T})$ be

$$\mathcal{R}_T(V) := F_\Gamma|_T + \Delta_\Gamma V|_T = F_\Gamma|_T \qquad \forall\, T \in \mathcal{T},$$
$$\mathcal{J}_S(V) := \nabla_\Gamma V^+|_S \cdot \mathbf{n}_S^+ + \nabla_\Gamma V^-|_S \cdot \mathbf{n}_S^- \qquad \forall\, S \in \mathcal{S},$$

where n_S^+ and n_S^- are outward unit normals to S with respect to T^+ and T^-, on the supporting planes containing T^+ and T^- respectively; T^+ and T^- are elements in $\mathcal{T}$ that share the side $S \in \mathcal{S}$ where $\mathcal{S}$ denotes the set of interior faces of $T \in \mathcal{T}$. We stress that, in contrast to flat domains, $\mathbf{n}_S^+ \neq \mathbf{n}_S^-$ because the vector may have different supporting hyperplanes. Similarly,

$\nabla_\Gamma V^+|_S = \widehat{\nabla} V^+ \mathbf{D}_\Gamma|_{\widehat{S}}$ and $\nabla_\Gamma V^-|_S = \widehat{\nabla} V^- \mathbf{D}_\Gamma|_{\widehat{S}}$ are tangential gradients of V on T^+ and T^- restricted to S, respectively. Note that, according to (4.14),

$$\Delta_\Gamma V|_T = Q_\Gamma^{-1} \widehat{\mathrm{div}}\big(Q_\Gamma \widehat{\nabla} \widehat{V} \mathbf{G}_\Gamma^{-1}\big)|_{\widehat{T}} = 0 \qquad \forall\, T \in \mathcal{T},$$

provided V and Γ are piecewise linear. We refer to [12, 29] for the case $\Delta_\Gamma V|_T \neq 0$.

Subtracting the weak formulations (4.13) and (4.16), and employing (4.12) to integrate by parts elementwise, we obtain for all $v \in H^1(\gamma)$:

$$\int_\gamma \nabla_\gamma (u - U) \cdot \nabla_\gamma v = I_1 + I_2 + I_3, \tag{5.11}$$

with

$$\begin{aligned}
I_1 &:= \sum_{T \in \mathcal{T}} \int_T F_\Gamma (v - V) - \sum_{S \in \mathcal{S}} \int_S \mathcal{J}_S(U)(v - V), \\
I_2 &:= \int_\Gamma \nabla_\Gamma U \cdot \nabla_\Gamma v - \int_\gamma \nabla_\gamma U \cdot \nabla_\gamma v = \int_\gamma \nabla_\gamma U \mathbf{E}_\Gamma \nabla_\gamma^T v, \\
I_3 &:= \int_\gamma f v - \int_\Gamma F_\Gamma v.
\end{aligned}$$

The choice $F_\Gamma = \frac{q}{Q_\Gamma} f$ of(4.17) implies $I_3 = 0$ so that only I_1 and I_2 need to be estimated. Observe that I_1 is the usual residual term, whereas I_2 is the geometry consistency term studied in §5.1 which accounts for the discrepancy between γ and Γ.

We focus now on I_1. The PDE error indicator is defined as follows for any $V \in \mathbb{V}(\mathcal{T})$

$$\eta_\mathcal{T}(V, T)^2 := h_T^2 \, \|F_\Gamma\|_{L^2(T)}^2 + \frac{1}{2} \sum_{S \subset \partial T} h_T \, \|\mathcal{J}_S(V)\|_{L^2(S)}^2 \qquad \forall\, T \in \mathcal{T},$$

where $h_T := |T_0|^{\frac{1}{d}}$ and T_0 is the preimage of T in the initial triangulation $\mathcal{T}_0$, ie. $T_0 = \mathcal{F}_0 \circ \mathcal{F}_\mathcal{T}^{-1}(T)$. This definition of h_T guarantees the strict reduction property

$$h_{T'} \leq 2^{-b/d} h_T \tag{5.12}$$

for all T' obtained from T after b bisections. We also introduce the *data oscillation*

$$\mathrm{osc}_\mathcal{T}(f, T) := h_T \|F_\Gamma - \overline{F}_\Gamma\|_{L^2(T)} \qquad \forall\, T \in \mathcal{T}, \tag{5.13}$$

where $\overline{F}_\Gamma$ stands for the meanvalue of F_Γ on $T \in \mathcal{T}$. Finally, for any subset $\tau \subset \mathcal{T}$ we set

$$\eta_\mathcal{T}(V, \tau)^2 := \sum_{T \in \tau} \eta_\mathcal{T}(V, T)^2, \qquad \text{and} \qquad \mathrm{osc}_\mathcal{T}(f, \tau)^2 := \sum_{T \in \tau} \mathrm{osc}_\mathcal{T}(f, T)^2,$$

and simply write $\eta_\mathcal{T}(V)$ and $\mathrm{osc}_\mathcal{T}(f)$ whenever $\tau = \mathcal{T}$.

Standard arguments [1, 54] to derive upper and lower bounds for the energy error on flat domains can be extended to this case; see [29, 46, 12]. We thus sketch the proof.

Lemma 5.6 (a posteriori upper and lower bounds)**.** *Assume that λ_{Γ_0} satisfies* (5.7)*. Let $u \in H^1(\gamma)$ be the solution of* (4.13)*, $(\Gamma, \mathcal{T})$ an approximating surface-mesh pair, and $U \in \mathbb{V}(\mathcal{T})$ be the Galerkin solution of* (4.16)*. Then there exist constants C_1, C_2 and Λ_1 depending only on $\mathcal{T}_0$, the Lipschitz constant of γ, and $\|f\|_{L^2(\gamma)}$, such that*

$$\|\nabla_\gamma (u - U)\|_{L^2(\gamma)}^2 \leq C_1 \eta_\mathcal{T}(U)^2 + \Lambda_1 \lambda_\Gamma^2, \tag{5.14}$$

$$C_2 \eta_\mathcal{T}(U)^2 \leq \|\nabla_\gamma (u - U)\|_{L^2(\gamma)}^2 + \mathrm{osc}_\mathcal{T}(f)^2 + \Lambda_1 \lambda_\Gamma^2. \tag{5.15}$$

Proof. Our departing point is (5.11) with $v \in H^1_\#(\gamma)$ arbitrary and $V \in \mathbb{V}(\mathcal{T})$ its Scott-Zhang interpolant, built over the parametric domain Ω [19]. Using interpolation estimates and (5.10) yields

$$|I_1| \lesssim \eta_{\mathcal{T}}(U)\|\nabla_\gamma v\|_{L^2(\gamma)}.$$

Since $\|\nabla_\Gamma U\|_{L^2(\gamma)} \lesssim \|f\|_{L^2(\gamma)}$, invoking Corollary 5.3 gives

$$|I_2| \lesssim \lambda_\Gamma \|\nabla_\gamma v\|_{L^2(\gamma)}.$$

Since $I_3 = 0$ we obtain the upper bound (5.14). To prove (5.15) we resort to a local argument due to R. Verfürth [54]. Let $T \in \mathcal{T}$ and b_T be corresponding cubic bubble. If $v = \overline{F}_\Gamma b_T \in H^1_0(T)$, then

$$\|\nabla_\gamma v\|_{L^2(T)} \lesssim h_T^{-1}\|\overline{F}_\Gamma\|_{L^2(T)}.$$

Therefore, inserting v into (5.11) and taking $V = 0$ leads to

$$\|\overline{F}_\Gamma\|^2_{L^2(T)} \lesssim \int_T \overline{F}_\Gamma v \lesssim h_T^{-1}\Big(\|\nabla_\gamma(u-U)\|_{L^2(T)} + \lambda_\Gamma(T)\Big)\|\overline{F}_\Gamma\|_{L^2(T)}.$$

This combined with the triangle inequality gives part of (5.15). It remains to deal with the jump, for which we select an arbitrary side $S \in \mathcal{S}$ with adjacent elements $T^\pm$. Let b_S be a piecewise quadratic bubble with value 1 at the midpoint of S and 0 at any other quadratic node. Let $v = \mathcal{J}_S(U)b_S \in H^1_0(\omega_S)$ where $\omega_S = T^+ \cup T^-$. Replacing v into (5.11) and taking $V = 0$ yields

$$\|\mathcal{J}_S(U)\|^2_{L^2(S)} \lesssim \int_S \mathcal{J}_S(U)v \le \Big(\|\nabla_\gamma(u-U)\|_{L^2(\omega_S)} + h_S\|F_\Gamma\|_{L^2(\omega_S)} + \lambda_\Gamma(\omega_S)\Big)\|\nabla_\gamma v\|_{L^2(\omega_S)}.$$

To conclude the proof we invoke the property $\|\nabla_\gamma v\|_{L^2(\omega_S)} \lesssim h_S^{-1/2}\|\mathcal{J}_S(U)\|_{L^2(S)}$ along with the previous estimate for $h_S\|F_\Gamma\|_{L^2(\omega_S)}$. □

To prove optimality of AFEM we need a localized upper bound for the distance between two discrete solutions. This bound measures $\|\nabla_\gamma(U_* - U)\|_{L^2(\gamma)}$ in terms of the PDE estimator restricted to the refined set and geometric estimator [12, Lemma 4.13].

Lemma 5.7 (localized upper bound)**.** *Assume that λ_{Γ_0} satisfies* (5.7)*. For $(\mathcal{T},\Gamma)$, $(\mathcal{T}_*,\Gamma_*)$ pairs of mesh-surface approximations with $\mathcal{T} \le \mathcal{T}_*$, let $\mathcal{R} := \mathcal{R}_{\mathcal{T}\to\mathcal{T}^*} \subset \mathcal{T}$ be the set of elements refined in $\mathcal{T}$ to obtain $\mathcal{T}^*$. Let $U \in \mathbb{V}(\mathcal{T})$ and $U \in \mathbb{V}(\mathcal{T}_*)$ be the corresponding discrete solutions of* (4.16) *on Γ and Γ_*, respectively. Then the following localized upper bound is valid*

$$\|\nabla_\gamma(U_* - U)\|^2_{L^2(\gamma)} \le C_1\eta_{\mathcal{T}}(U,\mathcal{R})^2 + \Lambda_1\lambda_\Gamma(\mathcal{R})^2, \tag{5.16}$$

with constants C_1, Λ_1 as in Lemma 5.6.

Proof. We start from the error representation formula (5.11) by replacing γ by Γ_* and taking as a test function $v = E_* := U - U_* \in H^1_\#(\gamma)$

$$\|\nabla_\gamma(U_* - U)\|^2_{L^2(\gamma)} \simeq \int_{\Gamma_*} \nabla_{\Gamma_*}(U_* - U)\cdot\nabla_{\Gamma_*}E_* = I_1 + I_2 + I_3.$$

To estimate I_1, we proceed as in the flat case [21, 49, 51]. We first construct an approximation $V \in \mathbb{V}(\mathcal{T})$ of $E_* \in \mathbb{V}(\mathcal{T}_*)$. Let ω be the union of elements of $\mathcal{T}$ which are refined in $\mathcal{T}_*$, and denote by ω_i one of the connected components of its interior. Let $\mathcal{T}_i$ be the subset of $\mathcal{T}$ contained in ω_i and let $\mathbb{V}(\mathcal{T}_i)$ be the restriction of $\mathbb{V}(\mathcal{T})$ to ω_i. We now can construct the Scott-Zhang operator on the corresponding flat domains $\widehat{\omega}_i = \mathcal{F}_{\mathcal{T}}^{-1}(\omega)$ and then lift them to Γ via $\mathcal{F}_{\mathcal{T}}$. We denote these lifts by $P_i : H^1(\widehat{\omega}_i) \to \mathbb{V}(\mathcal{T}_i)$. Let $V \in \mathbb{V}(\mathcal{T})$ be the following approximation of the error $E_* \in \mathbb{V}(\mathcal{T}_*)$:

$$V := P_iE_* \quad \text{in } \omega_i, \qquad V := E_* \quad \text{elsewhere.}$$

By construction, V has conforming boundary values on $\partial\omega_i$, is continuous in Γ, i.e $V \in \mathbb{V}(\mathcal{T})$ and is an H^1-stable approximation to E_*. Since $V = E_*$ in $\Gamma\backslash\omega$ we obtain by standard argument

$$|I_1| \leq C_1 \eta_{\mathcal{T}}(U, \mathcal{R}) \|\nabla_\Gamma E_*\|_{L^2(\Gamma)}.$$

To estimate I_2, we first note that $I_2|_{\Gamma\backslash\omega} = 0$ because Γ and Γ_* coincide in the unrefined region $\Gamma\backslash\omega$. Adding and subtracting $\int_{\widetilde{\omega}} \nabla_\gamma U \nabla_\gamma E_*$, with $\widetilde{\omega} = \mathcal{X} \circ \mathcal{F}_{\mathcal{T}}^{-1}(\omega)$, we obtain

$$I_2 = \int_{\widetilde{\omega}} \nabla_\gamma U \mathbf{E}_\Gamma \nabla_\gamma^T E_* - \int_{\widetilde{\omega}} \nabla_\gamma U \mathbf{E}_{\Gamma_*} \nabla_\gamma^T E_*.$$

Combining Corollary 5.3 with (5.10) and (3.6), in its elementwise form, we obtain

$$|I_2| \lesssim (\lambda_\Gamma(\mathcal{R}) + \lambda_{\Gamma^*}(\mathcal{R})) \, \|\nabla_\Gamma E_*\|_{L^2(\gamma)} \lesssim (1 + \Lambda_0^2) \|f\|_{L^2(\gamma)} \lambda_\Gamma(\mathcal{R}).$$

We note that the choice (4.17) of discrete forcing terms F_{Γ_*} and F_Γ implies $I_3 = 0$. Finally, collecting the estimates above we conclude (5.16). □

5.3 Properties of the PDE Estimator and Data Oscillation

As indicated in (5.14)-(5.15), we have access to the energy error $\|\nabla_\gamma(u-U)\|_{L^2(\gamma)}$ only through the PDE estimator $\eta_{\mathcal{T}}(U)$, the geometric estimator λ_Γ, and data oscillation $\mathrm{osc}_{\mathcal{T}}(f)$. As is customary for flat domains, (5.13) guarantees that $\mathrm{osc}_{\mathcal{T}}(f)$ is dominated by $\eta_{\mathcal{T}}(U)$ locally:

$$\mathrm{osc}_{\mathcal{T}}(f, T) \leq \eta_{\mathcal{T}}(U, T) \qquad \forall\, T \in \mathcal{T}. \tag{5.17}$$

The main novelty in (5.14)-(5.16) with respect to flat domains, which is also the chief challenge of the present analysis, is the presence of λ_Γ. In this respect, we show now the equivalence of $\eta_{\mathcal{T}}(U)$ and the *total error*

$$\mathcal{E}_{\mathcal{T}}(U, f) := \Big(\|\nabla_\gamma(u - U)\|_{L^2(\gamma)}^2 + \mathrm{osc}_{\mathcal{T}}(f)^2 \Big)^{\frac{1}{2}} \tag{5.18}$$

provided λ_Γ is small relative to $\eta_{\mathcal{T}}(U)$. We refer to [21] for a similar result for flat domains.

Lemma 5.8 (equivalence of estimator and total error)**.** *Let C_1, C_2, Λ_1 be given in Lemma 5.6. If*

$$\lambda_\Gamma^2 \leq \frac{C_2}{2\Lambda_1} \eta_{\mathcal{T}}(U)^2, \tag{5.19}$$

then there exist explicit constants $C_3 \geq C_4 > 0$, depending on C_1, C_2, such that

$$C_4 \eta_{\mathcal{T}}(U) \leq \mathcal{E}_{\mathcal{T}}(U, f) \leq C_3 \eta_{\mathcal{T}}(U). \tag{5.20}$$

Proof. Combining (5.14) with (5.19), we infer that

$$\|\nabla_\gamma(u - U)\|_{L^2(\gamma)}^2 \leq \Big(C_1 + \frac{C_2}{2} \Big) \eta_{\mathcal{T}}(U)^2. \tag{5.21}$$

This, together with (5.17), gives the upper bound in (5.20). We next resort to (5.15) and (5.19) to obtain

$$C_2 \eta_{\mathcal{T}}(U)^2 \leq \|\nabla_\gamma(u - U)\|_{L^2(\gamma)}^2 + \mathrm{osc}_{\mathcal{T}}(f)^2 + \frac{C_2}{2} \eta_{\mathcal{T}}(U)^2,$$

which implies the lower bound in (5.20) and concludes the proof. □

It turns out that the usual reduction property of $\eta_{\mathcal{T}}(U)$ [21, Corollary 3.4], which is instrumental to prove a contraction property of AFEM, is also polluted by the presence of λ_Γ as stated below. The following result is proved in [46, Lemma 4.2] for any polynomial degree.

Lemma 5.9 (reduction of residual error estimator)**.** *Let λ_{Γ_0} satisfy* (5.7)*. Given a mesh-surface pair $(\mathcal{T},\Gamma)$, let $\mathcal{M}\subset\mathcal{T}$ be a subset of elements bisected at least $b\geq 1$ times in refining $\mathcal{T}$ to obtain $\mathcal{T}_*\geq\mathcal{T}$. If $\xi := 1-2^{-\frac{b}{d}}$, then there exist constants Λ_2 and Λ_3, solely depending on the shape regularity of $\mathcal{T}_0$, the Lipschitz constant L of γ, and $\|f\|_{L^2(\gamma)}$, such that for any $\delta>0$*

$$\begin{aligned}\eta_{\mathcal{T}_*}(U_*)^2 \leq (1+\delta)\big(\eta_{\mathcal{T}}(U)^2 - \xi\eta_{\mathcal{T}}(U,\mathcal{M})^2\big)\\ + (1+\delta^{-1})\big(\Lambda_3\|\nabla_\gamma(U_*-U)\|^2_{L^2(\gamma)} + \Lambda_2\lambda_\Gamma^2\big).\end{aligned} \tag{5.22}$$

Proof. Let $S\in\mathcal{S}_*$ be an interior side and $T^+,T^-\in\mathcal{T}^*$ be two elements sharing S. The component of $\nabla_{\Gamma_*}U_*$ tangential to S does not jump, because U_* is continuous across S, whence

$$\big|\mathcal{J}_S(U_*)\big| = \big|\nabla_{\Gamma_*}U_*^+ - \nabla_{\Gamma_*}U_*^-\big|.$$

where $U_*^\pm = U_*|_{T^\pm}$. Therefore

$$\begin{aligned}\big|\mathcal{J}_S(U_*)-\mathcal{J}_S(U)\big| \leq \big|\nabla_{\Gamma_*}(U_*^+-U^+)\big| + \big|\nabla_{\Gamma_*}(U_*^--U^-)\big|\\ + \big|\nabla_{\Gamma_*}U^+ - \nabla_\Gamma U^+\big| + \big|\nabla_{\Gamma_*}U^- - \nabla_\Gamma U^-\big|.\end{aligned}$$

Employing an inverse estimate together with (5.10), the first two terms can be bounded as follows:

$$h_S\|\nabla_{\Gamma_*}(U_*^\pm - U^\pm)\|^2_{L^2(S)} \lesssim \|\nabla_{\Gamma_*}(U_*^\pm-U^\pm)\|^2_{L^2(T^\pm)} \approx \|\nabla_\gamma(U_*^\pm-U^\pm)\|^2_{L^2(\widetilde{T}^\pm)}.$$

For the next two terms we use (4.7), in conjunction with (5.9) and (3.6), to write

$$h_S\|\nabla_{\Gamma_*}U^\pm - \nabla_\Gamma U^\pm\|_{L^2(S)} \lesssim \|\widehat{\nabla}\widehat{U}^\pm\big(\mathbf{D}_{\Gamma_*}-\mathbf{D}_\Gamma\big)\|_{L^2(\widehat{T}^\pm)} \lesssim \lambda_\Gamma,$$

where the hidden constant depends on $\|f\|_{L^2(\gamma)}$.

We now turn our attention to the interior residual. Let $T_*\in\mathcal{T}_*$ and $T=\mathcal{F}_\mathcal{T}\circ\mathcal{F}_{\mathcal{T}_*}^{-1}(T_*), \widehat{T} = \mathcal{F}_{\mathcal{T}_*}^{-1}(T_*)$ be the corresponding sets in Γ and Ω. Since $F_\Gamma = \frac{q}{Q_\Gamma}f$ we infer that

$$\Big|\int_{T_*}|F_{\Gamma_*}|^2 - \int_T|F_\Gamma|^2\Big| = \int_{\widehat{T}}|qf|^2\frac{\big|Q_\Gamma - Q_{\Gamma_*}\big|}{Q_\Gamma Q_{\Gamma_*}} \lesssim \lambda_\Gamma\|f\|^2_{L^2(\widetilde{T})},$$

because of (5.8) and the lower bounds for Q_Γ and Q_{Γ_*}, as well as (3.6).

Collecting the estimates above, we realize that we have derived the bound

$$\eta_{\mathcal{T}_*}(U_*)^2 \leq (1+\delta)\eta_\mathcal{T}(U,\mathcal{T}_*)^2 + (1+\delta^{-1})\Big(\Lambda_3\|\nabla_\gamma(U_*-U)\|^2_{L^2(\gamma)} + \Lambda_2\lambda_\Gamma^2\Big).$$

It remains to deal with the set $\mathcal{M}$, namely to prove

$$\eta_\mathcal{T}(U,\mathcal{T}_*)^2 \leq \eta_\mathcal{T}(U)^2 - \xi\eta_\mathcal{T}(U,\mathcal{M})^2.$$

This is exactly the same argument as for flat domains because of the definition of meshsize h_T and (5.12) [21, Corollary 3.4]. This concludes the proof. ☐

Another difference with the theory of adaptivity for flat domains is the behavior of data oscillation under refinement. The usual situation is that $\operatorname{osc}_\mathcal{T}(f)$ does not increase upon refinement from $\mathcal{T}$ to $\mathcal{T}_*$ [48]. This is no longer true because $\operatorname{osc}_\mathcal{T}(f)$ and $\operatorname{osc}_{\mathcal{T}_*}(f)$ are defined on different domains Γ and Γ_*. Instead, we have the following substitute.

Lemma 5.10 (quasi-monotonicity of data oscillation)**.** *Let λ_{Γ_0} satisfy* (5.7)*. Let $(\mathcal{T},\Gamma)$, $(\mathcal{T}_*,\Gamma_*)$ be mesh-surface pairs with $\mathcal{T}\leq\mathcal{T}_*$ and discrete forcing functions defined according to* (4.17)*. Then, there exists a constant $C_5\geq 1$, depending only on $\mathcal{T}_0$ and the Lipschitz constant L of γ, such that*

$$\operatorname{osc}_{\mathcal{T}_*}(f) \leq C_5\operatorname{osc}_\mathcal{T}(f). \tag{5.23}$$

Proof. Let $T_* \in \mathcal{T}_*$ and so in Γ_*, and let $T = \mathcal{F}_\Gamma \circ \mathcal{F}_{\Gamma_*}^{-1}(T_*)$ be the corresponding set in Γ, but perhaps not in $\mathcal{T}$. Using (4.17) and the fact that Q_Γ is piecewise constant, we realize that

$$\int_{T_*} \left|F_{\Gamma_*} - \overline{F}_{\Gamma_*}\right|^2 \leq \int_T \left|f\frac{q}{Q_{\Gamma_*}} - \overline{F}_\Gamma \frac{Q_\Gamma}{Q_{\Gamma_*}}\right|^2 \frac{Q_{\Gamma_*}}{Q_\Gamma} = \int_T \left|F_\Gamma - \overline{F}_\Gamma\right|^2 \frac{Q_\Gamma}{Q_{\Gamma_*}} \leq C_5^2 \int_T \left|F_\Gamma - \overline{F}_\Gamma\right|^2,$$

where C_5^2 is the maximum of the ratios $Q_\Gamma/Q_{\Gamma_*}|_{T_*}$ for all $T_* \in \mathcal{T}_*$ and is bounded by L^{2d}. □

6 AFEM: Design and Properties

Since λ_Γ and $\eta_\mathcal{T}(U)$ account for quite different effects, the algorithm AFEM is designed to handle them separately via the modules ADAPT_SURFACE and ADAPT_PDE:

AFEM: Given Γ_0, $\mathcal{T}_0$, and parameters $\varepsilon_0 > 0$, $0 < \rho < 1$, and $\omega > 0$, set $k = 0$.

1. $[\mathcal{T}_k^+, \Gamma_k^+] = \mathsf{ADAPT\_SURFACE}(\mathcal{T}_k, \omega\varepsilon_k)$
2. $[\mathcal{T}_{k+1}, \Gamma_{k+1}] = \mathsf{ADAPT\_PDE}(\mathcal{T}_k^+, \varepsilon_k)$
3. $\varepsilon_{k+1} = \rho\varepsilon_k$; $k = k+1$
4. Goto 1.

We notice the presence of the factor ω, which is employed to make the geometric error small relative to the current tolerance ε_k. This turns out to be essential for both contraction and optimality of AFEM, and is further discussed in §§7–9.

6.1 Module ADAPT_SURFACE

Given a tolerance $\tau > 0$ and admissible subdivision $\mathcal{T}$, $[\mathcal{T}^+, \Gamma^+] = \mathsf{ADAPT\_SURFACE}(\mathcal{T}, \Gamma, \tau)$ improves the surface resolution until

$$\lambda_{\Gamma^+} \leq \tau \tag{6.1}$$

where λ_Γ is the *geometric estimator* introduced in (3.4). This module is based on a *greedy* algorithm

$[\mathcal{T}^+, \Gamma^+] = \mathsf{ADAPT\_SURFACE}(\mathcal{T}, \Gamma, \tau)$
  while $\mathcal{M} := \{T \in \mathcal{T} \,|\, \lambda_\mathcal{T}(T) > \tau\} \neq \emptyset$
    $\mathcal{T} := \mathsf{REFINE}(\mathcal{T}, \mathcal{M})$
    $\Gamma := \mathcal{F}_\mathcal{T}(\Omega)$
  end while
  return($\mathcal{T}$,Γ)

where $\mathsf{REFINE}(\mathcal{T}, \mathcal{M})$ refines all elements in the marked set $\mathcal{M}$ and keeps conformity; more details are given in §6.2. To derive convergence rates for AFEM, we require that ADAPT_SURFACE is *t-optimal*, i.e. there exists a constant C such that the set $\mathcal{M}^+$ of all the elements marked for refinement in a call to $\mathsf{ADAPT\_SURFACE}(\mathcal{T}, \Gamma, \tau)$ satisfies

$$\#\mathcal{M}^+ \leq C\tau^{-1/t}, \tag{6.2}$$

whenever γ belongs to a suitable approximation class, $\mathbb{B}_t$ with $0 < t \leq 1/d$ (see §8.1). In §8.3 we show that this assumption is satisfied provided that $\gamma \in W_p^{1+td}(\Gamma_0)$ for some $tp > 1$.

6.2 Module ADAPT_PDE

Given a tolerance $\varepsilon > 0$ and admissible subdivision $\mathcal{T}^+$, $[\mathcal{T}, U] = \mathsf{ADAPT\_PDE}(\mathcal{T}^+, \varepsilon)$ outputs a refinement $\mathcal{T} \geq \mathcal{T}^+$ and the associated finite element solution $U \in \mathbb{V}(\mathcal{T})$ such that

$$\eta_\mathcal{T}(U) \leq \varepsilon. \tag{6.3}$$

The module ADAPT_PDE is the standard adaptive sequence:

$[\mathcal{T}, \Gamma] = \mathsf{ADAPT\_PDE}(\mathcal{T}, \varepsilon)$
  $U = \mathsf{SOLVE}(\mathcal{T})$
  $\{\eta_{\mathcal{T}}(U,T)\}_{T\in\mathcal{T}} = \mathsf{ESTIMATE}(\mathcal{T}, U)$
  while $\eta_{\mathcal{T}}(U) > \varepsilon$
    $\mathcal{M} := \mathsf{MARK}(\mathcal{T}, \{\eta_{\mathcal{T}}(U,T)\}_{T\in\mathcal{T}})$
    $\mathcal{T} := \mathsf{REFINE}(\mathcal{T}, \mathcal{M})$
    $\Gamma := \mathcal{F}_{\mathcal{T}}(\Omega)$
    $U = \mathsf{SOLVE}(\mathcal{T})$
    $\{\eta_{\mathcal{T}}(U,T)\}_{T\in\mathcal{T}} = \mathsf{ESTIMATE}(\mathcal{T}, U)$
  end while
  return($\mathcal{T}$,Γ)

We describe below the modules SOLVE, ESTIMATE, MARK and REFINE separately.

Procedure SOLVE. This procedure solves the SPD linear system resulting for (4.16). For simplicity we assume that the linear system is solved exactly. In this context, the approximate solution of the discrete problem can be handled as in [51]. We refer to [43] for a hierarchical basis multigrid preconditioner and to [18] for standard variational and non-variational multigrid algorithms.

Procedure ESTIMATE. Given the Galerkin solution $U \in \mathbb{V}(\mathcal{T})$ of (4.16) ESTIMATE computes the PDE error indicators $\{\eta_{\mathcal{T}}(U,T)\}_{T\in\mathcal{T}}$. We emphasize that this procedure does not compute the oscillation terms, which are only needed to carry out the analysis.

The equivalence stated in Lemma 5.8 is critical to deduce that the ADAPT_PDE strategy based on the reduction of the error indicators $\eta_{\mathcal{T}}(U)$ is successful in reducing the total error $\|\nabla_\gamma(u - U)\|_{L^2(\gamma)} + \mathrm{osc}_{\mathcal{T}}(f)$. To see this we impose the constraint on the parameter ω

$$\omega \le \omega_1 := \sqrt{\frac{C_2}{2\Lambda_0^2\Lambda_1}}, \tag{6.4}$$

and observe that the input $\mathcal{T}^+$ to ADAPT_PDE as well as all inner iterates satisfy, in view of (3.6),

$$\lambda_\Gamma^2 \le \Lambda_0^2\lambda_{\Gamma^+}^2 \le \frac{C_2}{2\Lambda_1}\varepsilon_k^2.$$

Since $\eta_{\mathcal{T}}(U) > \varepsilon_k$, we deduce the validity of (5.19) whence that of (5.20) within ADAPT_PDE.

Procedure MARK. We rely on an optimal *Dörfler's* marking strategy for the selection of elements. Given the set of indicators $\{\eta_{\mathcal{T}}(U,T)\}_{T\in\mathcal{T}}$ and a marking parameter $\theta \in (0,1]$, MARK outputs a subset of marked elements $\mathcal{M} \subset \mathcal{T}$ such that

$$\eta_{\mathcal{T}}(U,\mathcal{M}) \ge \theta\eta_{\mathcal{T}}(U). \tag{6.5}$$

In contrast to [46], MARK only employs the error indicators and does not use the oscillation nor surface indicators. We will see that quasi-optimality of AFEM requires that $\mathcal{M}$ be *minimal* and θ sufficiently small.

Procedure REFINE. Given a triangulation $\mathcal{T}$ and a subset $\mathcal{M}$ of marked elements, the call $\mathcal{T}_* = \mathsf{REFINE}(\mathcal{T}, \mathcal{M})$ bisects all elements in $\mathcal{M}$ at least $b \ge 1$ times while maintaining mesh conformity, to obtain a new mesh $\mathcal{T}_*$. The new surface Γ_* is obtained by piecewise linear interpolation of the parametrization $\mathcal{X}$ via $\mathcal{F}_{\mathcal{T}_*} = \mathcal{I}_{\mathcal{T}_*}\mathcal{X}$, namely, $\Gamma_* = \mathcal{F}_{\mathcal{T}_*}(\Omega)$.

To ensure conformity of $\mathcal{T}_*$ some additional elements of $\mathcal{T}\backslash\mathcal{M}$ need to be refined. The complexity of the overall refinement algorithm is controlled in a cumulative way, as was proved by P. Binev, W. Dahmen, and R. DeVore for $d = 2$ [11] and R. Stevenson [52] for $d > 2$; see also the survey [49]. The precise statement of this result is in the following lemma.

Lemma 6.1 (Complexity of REFINE)**.** *Assume that $\mathcal{T}_0$ is suitably labeled (condition (b) of §4 in [52]). Let $\{\mathcal{T}_k\}_{k\geq 0}$ be any sequence of meshes produced by succesive calls $\mathcal{T}_{k+1} = \mathsf{REFINE}(\mathcal{T}_k, \mathcal{M}_k)$. Then, there exists a constant C_6 solely depending on $\mathcal{T}_0$ and the refinement depth b such that*

$$\#\mathcal{T}_k - \#\mathcal{T}_0 \leq C_6 \sum_{j=0}^{k-1} \#\mathcal{M}_j, \qquad \forall\, k \geq 1. \tag{6.6}$$

It is worth noticing that the user parameter $b \geq 1$ only entails a minimal refinement, which does not force an interior node property [48, 47] or an extra refinement to improve the surface approximation [46].

Remark 6.2 (alternative subdivision strategies)**.** For simplicity we only discuss the refinement strategy based on simplex bisection. However, all the results obtained can be extended to any strategy satisfying Conditions 3, 4 and 6 in [14], such as quadrilaterals with hanging nodes.

7 Conditional Contraction Property

The procedure ADAPT_PDE is known to yield a contraction property in the "flat" case. In the present context, however, the surface approximation is responsible for lack of consistency in that the sequence of finite element spaces is no longer nested. This in turn leads to failure of a key orthogonality property between discrete solutions, the Pythagoras property. We have, instead, a perturbation result referred to as *quasi-orthogonality* below. Its proof follows the steps of that for graphs [46, Lemma 4.4]. In this section, we use the notation

$$\begin{aligned} e^j &:= \|\nabla_\gamma(u - U^j)\|_{L^2(\gamma)}, \quad E^j := \|\nabla_\gamma(U^{j+1} - U^j)\|_{L^2(\gamma)}, \\ \eta^j &:= \eta_{\mathcal{T}^j}(U^j), \quad \eta^j(\mathcal{M}^j) := \eta_{\mathcal{T}^j}(U^j, \mathcal{M}^j), \quad \lambda^j := \lambda_{\Gamma^j}, \end{aligned}$$

where $\mathcal{T}^j$ are meshes obtained after each inner iteration of ADAPT_PDE, starting with $\mathcal{T}^0 = \mathcal{T}^+$, and Γ^j, U^j are the corresponding discrete surfaces and Galerkin solutions.

Lemma 7.1 (Quasi-orthogonality)**.** *Let $\Lambda_2 > 0$ be the constant of Lemma 5.9, which solely depends on the Lipschitz constant L of γ and $\|f\|_{L^2(\gamma)}$. Then, for $i = j, j+1$ with $j \geq 0$, we have*

$$(e^j)^2 - \frac{3}{2}(E^j)^2 - \Lambda_2(\lambda^i)^2 \leq (e^{j+1})^2 \leq (e^j)^2 - \frac{1}{2}(E^j)^2 + \Lambda_2(\lambda^i)^2. \tag{7.1}$$

Proof. Since the symmetry of the Dirichlet form implies

$$(e^j)^2 = (e^{j+1})^2 + (E^j)^2 + 2\int_\gamma \nabla_\gamma(u - U^{j+1})\nabla_\gamma^T(U^{j+1} - U^j),$$

we just have to examine the last term. Combining (4.13), (4.16), and (4.17) with Lemma 5.1 yields

$$\Big|\int_\gamma \nabla_\gamma(u - U^{j+1})\nabla_\gamma^T(U^{j+1} - U^j)\Big| \lesssim \|f\|_{L^2(\gamma)}\lambda^j E^j,$$

which gives (7.1) after applying Young's inequality. □

Remark 7.2 (validity of (7.1))**.** Relation (7.1) is also true for any pair of triangulations $(\mathcal{T}, \mathcal{T}_*)$, with $\mathcal{T}_* \geq \mathcal{T}$, discrete solution $U_* \in \mathbb{V}(\mathcal{T}_*)$ on the finer space, and any discrete function $V \in \mathbb{V}(\mathcal{T})$.

Theorem 7.3 (Conditional Contraction Property)**.** *Let $\theta \in (0, 1]$ be the marking parameter of MARK and let $\{\mathcal{T}^j, \Gamma^j, U^j\}_{j\geq 0}^J$ be a sequence of meshes, piecewise affine surfaces and discrete*

solutions generated by the procedure ADAPT_PDE$(\mathcal{T}^0, \varepsilon)$ *within AFEM with tolerance* ε*, i.e.* $\lambda^0 \leq \omega\varepsilon$*. Assume that the AFEM parameter* ω *satisfies*

$$\omega \leq \omega_2 := \frac{\xi\theta^2}{\Lambda_0\sqrt{32\Lambda_2(2\Lambda_3+1)}}, \tag{7.2}$$

where $\xi = 1 - 2^{-b/d}$ *is defined in Lemma 5.9. There exist constants* $0 < \alpha < 1$ *and* $\beta > 0$ *such that*

$$(e^{j+1})^2 + \beta(\eta^{j+1})^2 \leq \alpha^2\Big((e^j)^2 + \beta(\eta^j)^2\Big) \qquad \forall\, 0 \leq j < J. \tag{7.3}$$

Moreover, the number of inner iterates J *of* ADAPT_PDE *is uniformly bounded.*

Proof. [1] Let $\beta > 0$ be a scaling parameter to be found later. We combine (7.1) and (5.22) to write

$$\begin{aligned}(e^{j+1})^2 + \beta(\eta^{j+1})^2 \leq (e^j)^2 &+ \Big(-\frac{1}{2} + \beta(1+\delta^{-1})\Lambda_3\Big)(E^j)^2 \\ &+ \Lambda_2\Big(1 + \beta(1+\delta^{-1})\Big)(\lambda^j)^2 + \beta(1+\delta)\Big((\eta^j)^2 - \xi\eta^j(\mathcal{M}^j)^2\Big).\end{aligned}$$

Here $\mathcal{M}^j$ is the set of elements in $\mathcal{T}^j$ marked for refinement at the j-th subiteration. To remove the factor of E^j we now choose β dependent on δ, to be

$$\beta(1+\delta^{-1})\Lambda_3 = \frac{1}{2} \quad \Rightarrow \quad \beta(1+\delta) = \frac{\delta}{2\Lambda_3}, \tag{7.4}$$

and thereby obtain

$$(e^{j+1})^2 + \beta(\eta^{j+1})^2 \leq (e^j)^2 + \Lambda_2\Big(1 + \beta(1+\delta^{-1})\Big)(\lambda^j)^2 + \beta(1+\delta)\Big((\eta^j)^2 - \xi\eta^j(\mathcal{M}^j)^2\Big).$$

[2] Invoking Dörfler marking (6.5), we deduce

$$(\eta^j)^2 - \xi\eta^j(\mathcal{M}^j)^2 \leq (1-\xi\theta^2)(\eta^j)^2.$$

Since the initial mesh $\mathcal{T}^0$ comes from ADAPT_SURFACE we know that $\lambda^0 \leq \omega\varepsilon \leq \omega\eta^j$ for all inner iterations $1 \leq j \leq J$ of ADAPT_PDE. Using (3.6) yields $\lambda^j \leq \Lambda_0\omega\eta^j$, whence

$$\begin{aligned}(e^{j+1})^2 + \beta(\eta^{j+1})^2 \leq &(e^j)^2 - \beta(1+\delta)\frac{\xi\theta^2}{2}(\eta^j)^2 \\ &+ \beta\Big((1+\delta)\Big(1 - \frac{\xi\theta^2}{2}\Big) + \Lambda_2\Big(1 + \frac{1}{2\Lambda_3}\Big)\frac{\Lambda_0^2\omega^2}{\beta}\Big)(\eta^j)^2.\end{aligned}$$

Applying the simpler upper bound (5.21), which is valid for the inner iterates of ADAPT_PDE, and replacing β according to (7.4), we obtain

$$(e^{j+1})^2 + \beta(\eta^{j+1})^2 \leq \alpha_1(\delta)(e^j)^2 + \alpha_2(\delta)\beta(\eta^j)^2$$

with

$$\alpha_1(\delta)^2 := 1 - \delta\frac{\xi\theta^2}{4\Lambda_3 C_3}, \qquad \alpha_2(\delta)^2 := (1+\delta)\left(1 - \frac{\xi\theta^2}{2}\right) + \Lambda_2\left(1 + \frac{1}{2\Lambda_3}\right)\frac{\Lambda_0^2\omega^2}{\beta}.$$

It remains to prove that δ can be chosen so that $\alpha_2(\delta)^2 < 1$. We then fix the parameter δ so that

$$(1+\delta)\Big(1 - \frac{\xi\theta^2}{2}\Big) = 1 - \frac{\xi\theta^2}{4} \quad \Rightarrow \quad \delta = \frac{\xi\theta^2}{4 - 2\xi\theta^2},$$

Now, according to (7.4), we obtain $\beta = \frac{\xi\theta^2}{2\Lambda_3(4-\xi\theta^2)} \geq \frac{\xi\theta^2}{8\Lambda_3}$ and since $\omega \leq \omega_2$ we infer that

$$\Lambda_2\left(1+\frac{1}{2\Lambda_3}\right)\frac{\Lambda_0^2\omega^2}{\beta} \leq \frac{4\Lambda_2(2\Lambda_3+1)}{\xi\theta^2}\Lambda_0^2\omega^2 \leq \frac{\xi\theta^2}{8}.$$

Hence $\alpha_2^2 \leq 1 - \frac{\xi\theta^2}{8} < 1$, and the choice $\alpha := \max\{\alpha_1, \alpha_2\} < 1$ yields the desired estimate (7.3).
3 The contraction property (7.3) guarantees that ADAPT_PDE stops in a finite number of iterations J. To show that J is independent of the iteration counter k of AFEM, take $k \geq 1$ and note that before the call ADAPT_PDE$(\mathcal{T}_k^+, \varepsilon_k)$ we have

$$\eta_k = \eta_{\mathcal{T}_k}(U_k) \leq \varepsilon_{k-1} = \frac{\varepsilon_k}{\rho}, \qquad \lambda_k = \lambda_{\Gamma_k} \leq \Lambda_0\lambda_{\Gamma_{k-1}^+} \leq \frac{\Lambda_0\omega}{\rho}\varepsilon_k.$$

We next combine (5.22), with $\delta = 1$, and (7.1) to get

$$\eta_{\mathcal{T}_k^+}(U_k^+)^2 \lesssim \eta_k^2 + \lambda_k^2 + \|\nabla_\gamma(U_k^+ - U_k)\|_{L^2(\gamma)}^2 \lesssim \eta_k^2 + \lambda_k^2 + \|\nabla_\gamma(u - U_k)\|_{L^2(\gamma)}^2,$$

where the hidden constants depend on Λ_2, Λ_3. The bounds on η_k, λ_k, together with (5.14), yield

$$(\eta^0)^2 = \eta_{\mathcal{T}_k^+}(U_k^+)^2 \lesssim \eta_k^2 + \lambda_k^2 \lesssim \varepsilon_k^2.$$

Since the stopping condition of ADAPT_PDE is $\eta^J \leq \varepsilon_k$, (7.3) implies that J is bounded independently of k, as asserted. □

That J is uniformly bounded dictates the complexity of ADAPT_PDE because the most expensive module SOLVE is run just J times. However, this property is not required for the study of cardinality of §8.

8 Optimal Cardinality

In this section we study the cardinality of AFEM, which is dictated by the regularity of u, f and γ. We first discuss in §8.1 the best approximation error achievable with piecewise linear polynomials for both surface and PDE solution. We show next in §8.2 that AFEM delivers the best convergence rate provided the procedure ADAPT_SURFACE is t-optimal, namely it satisfies (6.2). We conclude in §8.3 with a greedy algorithm for ADAPT_SURFACE that is t-optimal.

8.1 Approximation Classes

We define classes of functions and surfaces in terms of decay rate of the approximation error as a function of the number of degrees of freedom N. Let $\mathbb{T}_N \subset \mathbb{T} := \mathbb{T}(\mathcal{T}_0)$ be the set of all possible conforming triangulations, generated on γ with at most N elements more than $\mathcal{T}_0$ by successive bisection of $\mathcal{T}_0$:

$$\mathbb{T}_N := \big\{\mathcal{T} \in \mathbb{T} \mid \#\mathcal{T} - \#\mathcal{T}_0 \leq N\big\}.$$

Given $v \in H^1_\#(\gamma), f \in L^2(\gamma)$, the notion of *total error* $\mathcal{E}_\mathcal{T}(V,f)^2 = \|\nabla_\gamma(v-V)\|_{L^2(\gamma)}^2 + \mathrm{osc}_\mathcal{T}(f)^2$ is defined in (5.18). In view of Lemma 5.8 and the fact that AFEM is driven by $\eta_\mathcal{T}(U)$ and λ_Γ, we assess the quality of the best approximation (v,f) with N degrees of freedom in terms of

$$\sigma(N; v, f, \gamma) := \inf_{\mathcal{T}\in\mathbb{T}_N}\inf_{V\in\mathbb{V}(\mathcal{T})}\mathcal{E}_\mathcal{T}(V,f).$$

This is consistent with the approach taken for flat domains in [21, 49]. For $s > 0$, we define the nonlinear (algebraic) approximation class $\mathbb{A}_s(\gamma)$ to be

$$\mathbb{A}_s(\gamma) := \Big\{(v,f) \mid |v,f|_{\mathbb{A}_s} := \sup_{N\geq 1}\big(N^s\,\sigma(N; v, f, \gamma)\big) < \infty\Big\}.$$

We emphasize that the approximability of the surface γ only appears implicitly by measuring the errors on γ. In fact, the definition of data oscillation (5.13), and in particular the specific choice of F_Γ, implies that $\mathrm{osc}^2_{\mathcal{T}}(f)$ entails the approximation of f by piecewise constants on γ but does not include the approximation of γ by Γ; this is quite different for higher order approximations of γ, as is shown in [12]. On the other hand, the generic range of s is dictated by polynomial degree, namely $0 < s \leq 1/d$.

An alternative and useful definition to $(u, f) \in \mathbb{A}_s(\gamma)$ is as follows: given $\varepsilon > 0$, there exists a mesh $\mathcal{T}_\varepsilon \in \mathbb{T}(\mathcal{T}_0)$ with $\mathcal{T}_\varepsilon \geq \mathcal{T}_0$ and a discrete function $V_\varepsilon \in \mathbb{V}(\mathcal{T}_\varepsilon)$ so that

$$\|\nabla_\gamma(u - V_\varepsilon)\|^2_{L^2(\gamma)} + \mathrm{osc}_{\mathcal{T}_\varepsilon}(f)^2 \leq \varepsilon^2, \qquad \#\mathcal{T}_\varepsilon - \#\mathcal{T}_0 \leq |u, f|^{\frac{1}{s}}_{\mathbb{A}_s} \varepsilon^{-\frac{1}{s}}; \tag{8.1}$$

$\Gamma_\varepsilon = \mathcal{F}_{\mathcal{T}_\varepsilon}(\Omega)$ might not be a good approximaton of γ. In fact, approximations of (u, f) and γ are handled separately. The characterization of $\mathbb{A}_s(\gamma)$ in terms of Besov regularity is an open issue.

Similarly, for surfaces and $t > 0$, we define the approximation class as follows:

$$\mathbb{B}_t := \Big\{\gamma \in W^1_\infty \mid |\gamma|_{\mathbb{B}_t} := \sup_{N\geq 1} N^t \inf_{\mathcal{T}\in\mathbb{T}_N} \lambda_\Gamma < \infty\Big\}.$$

This means that surfaces in $\mathbb{B}_t$ are parametrized by Lipschitz maps $\mathcal{X} : \Omega \to \mathbb{R}^{d+1}$ which can in turn be approximated with rate N^{-t} in W^1_∞ over Ω with N degrees of freedom. In section 8.3, we give a constructive greedy algorithm that realizes this rate provided γ belong to a suitable Sobolev space in the nonlinear scale of W^1_∞. The generic range of exponents t for linear elements, or equivalently polyhedral surfaces Γ, is $0 < t \leq 1/d$.

8.2 Convergence Rates

We now prove that AFEM achieves the asymptotic decay rate $\min\{s, t\}$, dictated by the classes $\mathbb{A}_s(\gamma)$ and $\mathbb{B}_t$, but without ever using either s or t in its formulation. We establish the link between the performance of AFEM and the best possible error by adapting a clever idea of R. Stevenson [51] for the Laplace operator, further extended by J.M. Cascón et al [21] to general elliptic PDE, in flat domains; we refer to the survey [49] for a thorough discussion. The insight is that

any marking strategy that reduces the total error relative to its current value must contain a substantial portion of the error estimator, and so it can be related to Dörfler Marking. (8.2)

Exploiting next the minimality of Dörfler marking enables us to compare meshes generated by AFEM with the best meshes within $\mathbb{T}$. The approach of [21, 49, 51] does not apply directly in the present context because of the consistency error due to surface interpolation. We account for this discrepancy below upon making the parameter ω of ADAPT_SURFACE sufficiently small. Let

$$\omega_3 := \frac{C_4}{\Lambda_0\sqrt{3\Lambda_1 + 2\Lambda_2}}, \qquad \omega_4 := \frac{C_4}{2\Lambda_0}\sqrt{\Big(1 - \frac{\theta^2}{\theta^2_*}\Big)\frac{1}{\Lambda_2}} \tag{8.3}$$

be two thresholds for ω to be used next and θ_* be a threshold for the Dörfler parameter θ

$$\theta_* := \frac{C_4}{\sqrt{2 + 3C_1}}; \tag{8.4}$$

since $C_4 = \sqrt{C_2/2}$ and $C_2 \leq C_1$, we see that $\theta_* < 1$.

Lemma 8.1 (Dörfler marking). *Let λ_{Γ_0} satisfy* (5.7)*, and the parameters θ and ω satisfy*

$$0 < \theta < \theta_*, \qquad 0 < \omega \leq \min\{\omega_1, \omega_3\}, \tag{8.5}$$

where θ_*, ω_3 are defined in (8.3), (8.4), *and ω_1 in* (6.4). *Let $\mu = \frac{1}{2}\big(1 - \frac{\theta^2}{\theta_*^2}\big)$ and $(\Gamma, \mathcal{T}, U)$ be the approximate surface, mesh and discrete solution produced by an inner iterate of* ADAPT_PDE. *If $(\Gamma_*, \mathcal{T}_*, U_*)$ is a surface-mesh-solution triple with $\mathcal{T}_* \geq \mathcal{T}$, such that the total error satisfies*

$$\mathcal{E}_{\mathcal{T}_*}(U_*, f) \leq \mu\, \mathcal{E}_{\mathcal{T}}(U, f), \tag{8.6}$$

then the refined set $\mathcal{R} := \mathcal{R}_{\mathcal{T}\to\mathcal{T}_*}$ satisfies Dörfler property with parameter θ, namely

$$\eta_{\mathcal{T}}(U, \mathcal{R}) \geq \theta \eta_{\mathcal{T}}(U). \tag{8.7}$$

Proof. We proceed as in [21, Lemma 5.9] using the notation $e(U) = \|\nabla_\gamma(u - U)\|_{L^2(\gamma)}$. Since $\omega \leq \omega_1$, we combine the lower bound of (5.20) with (8.6) to write

$$\begin{aligned}(1-\mu)C_4^2\eta_{\mathcal{T}}(U)^2 &\leq (1-\mu)\big(e(U)^2 + \mathrm{osc}_{\mathcal{T}}(f)^2\big)\\ &\leq e(U)^2 - e(U_*)^2 + \mathrm{osc}_{\mathcal{T}}(f)^2 - \mathrm{osc}_{\mathcal{T}_*}(f)^2.\end{aligned}$$

We now estimate separately error and oscilation terms. According to (7.1) and (5.16), we obtain

$$e(U)^2 - e(U_*)^2 \leq \frac{3}{2}\|\nabla_\gamma(U_* - U)\|^2_{L^2(\gamma)} + \Lambda_2\lambda_{\mathcal{T}}^2 \leq \frac{3}{2}C_1\eta_{\mathcal{T}}(U,\mathcal{R})^2 + \Big(\frac{3}{2}\Lambda_1 + \Lambda_2\Big)\lambda_\Gamma^2.$$

On the other hand, the data oscillation terms verify

$$\mathrm{osc}_{\mathcal{T}}(f)^2 - \mathrm{osc}_{\mathcal{T}_*}(f)^2 \leq \mathrm{osc}_{\mathcal{T}}(f,\mathcal{R})^2 \leq \eta_{\mathcal{T}}(U,\mathcal{R})^2$$

because they coincide over $\mathcal{T}\backslash\mathcal{R}$ and the estimator dominates the oscillation locally (see (5.17)). Since $(\Gamma, \mathcal{T})$ is produced within ADAPT_PDE, we have $\eta_{\mathcal{T}}(U) > \varepsilon$ and $\lambda_{\Gamma^+} \leq \omega\varepsilon$, whence

$$\lambda_\Gamma \leq \Lambda_0\lambda_{\Gamma^+} \leq \Lambda_0\omega\varepsilon \leq \Lambda_0\omega\eta_{\mathcal{T}}(U).$$

Collecting these three estimates, and using that $\omega \leq \omega_3$, we infer that

$$\left(1 + \frac{3}{2}C_1\right)\eta_{\mathcal{T}}(U,\mathcal{R})^2 \geq \left((1-\mu)C_4^2 - \Lambda_0^2\omega^2\left(\frac{3}{2}\Lambda_1 + \Lambda_2\right)\right)\eta_{\mathcal{T}}(U)^2 \geq (1-2\mu)\frac{C_4^2}{2}\eta_{\mathcal{T}}(U)^2.$$

Finally, the asserted estimate (8.7) is a consequence of the definition of θ_*, μ and $\theta < \theta_*$. □

Lemma 8.2 (cardinality of $\mathcal{M}$)**.** *Let λ_{Γ_0} satisfy* (5.7) *and the procedure* MARK *select a set $\mathcal{M}$ with minimal cardinality. Let the parameters θ and ω satisfy*

$$0 < \theta < \theta_*, \qquad 0 < \omega \leq \min\{\omega_1, \omega_4\} \tag{8.8}$$

with $\theta_*, \omega_4, \omega_1$ given in (8.4), (8.3), *and* (6.4), *respectively. Let u be the solution of* (4.13), *and let $(\Gamma, \mathcal{T}, U)$ be produced within* ADAPT_PDE. *If $(u, f) \in \mathbb{A}_s(\gamma)$, then*

$$\#\mathcal{M} \lesssim |u, f|_s^{\frac{1}{s}}\, \mathcal{E}_{\mathcal{T}}(U, f)^{-\frac{1}{s}}.$$

Proof. We set

$$\delta^2 = \hat{\mu}\,\mathcal{E}_{\mathcal{T}}(U,f)^2 = \hat{\mu}\Big(e(U)^2 + \mathrm{osc}_{\mathcal{T}}(f)^2\Big),$$

for $0 < \hat{\mu} < \mu = \frac{1}{2}\big(1 - \frac{\theta^2}{\theta_*^2}\big) < 1$ sufficiently small to be determined later. Since $(u, f) \in \mathbb{A}_s(\gamma)$, there exists a pair $(\Gamma_\delta, \mathcal{T}_\delta)$ with $\mathcal{T}_\delta \geq \mathcal{T}_0$ and a $V_\delta \in \mathbb{V}(\mathcal{T}_\delta)$ such that

$$\#\mathcal{T}_\delta - \#\mathcal{T}_0 \lesssim |u, f|_s^{\frac{1}{s}}\delta^{-\frac{1}{s}}, \qquad e(V_\delta)^2 + \mathrm{osc}_{\mathcal{T}_\delta}(f)^2 \leq \delta^2. \tag{8.9}$$

Let $\mathcal{T}_* = \mathcal{T} \oplus \mathcal{T}_\delta$ be the overlay of $\mathcal{T}$ and $\mathcal{T}_\delta$, which satisfies [21, Lemma 3.7], [49]

$$\#\mathcal{T}_* \leq \#\mathcal{T} + \#\mathcal{T}_\delta - \#\mathcal{T}_0. \tag{8.10}$$

Let $U_* \in \mathbb{V}(\mathcal{T}_*)$ be the corresponding Galerkin solution. We observe that $\mathcal{T}_* \geq \mathcal{T}_\delta, \mathcal{T}$, and invoke the upper bound of (7.1) in conjunction with (5.23) to write

$$e(U_*)^2 + \operatorname{osc}_{\mathcal{T}_*}(f)^2 \leq e(V_\delta)^2 + \Lambda_2 \lambda_\Gamma^2 + C_5^2 \operatorname{osc}_{\mathcal{T}_\delta}(f)^2.$$

We recall that $\lambda_\Gamma \leq \Lambda_0 \lambda_{\Gamma^+} < \Lambda_0 \omega \varepsilon$, in view of (3.6) and (6.1), and $\eta_{\mathcal{T}}(U) > \varepsilon$ because of (6.3). Combining this with (5.20), we arrive at

$$\lambda_\Gamma^2 \leq \frac{\Lambda_0^2 \omega^2}{C_4^2} \Big(e(U)^2 + \operatorname{osc}_{\mathcal{T}}(f)^2 \Big) = \frac{\Lambda_0^2 \omega^2}{\hat{\mu} C_4^2} \delta^2.$$

Using the fact that $C_5 \geq 1$ and $\omega \leq \omega_4$, we choose $\hat{\mu} = \frac{\mu}{2C_5^2}$ to end up with

$$e(U_*)^2 + \operatorname{osc}_{\mathcal{T}_*}(f)^2 \leq \Big(C_5^2 + \frac{\Lambda_0^2 \Lambda_2 \omega^2}{\hat{\mu} C_4^2} \Big) \delta^2 = \mu \Big(e(U)^2 + \operatorname{osc}_{\mathcal{T}}^2(f) \Big).$$

We thus deduce from Lemma 8.1 that the subset $\mathcal{R} := \mathcal{R}_{\mathcal{T} \to \mathcal{T}_*} \subset \mathcal{T}$ satisfies Dörfler property (8.7). Since the set $\mathcal{M} \subset \mathcal{T}$ also satisfies this property, but with minimal cardinality, we infer that

$$\#\mathcal{M} \leq \#\mathcal{R} \leq \#\mathcal{T}_* - \#\mathcal{T} \leq \#\mathcal{T}_\delta - \#\mathcal{T}_0 \lesssim |u, f|_{\mathbb{A}_s}^{\frac{1}{s}} \delta^{-\frac{1}{s}}.$$

The asserted estimate finally follows upon using the definition of δ. □

The quasi-optimal cardinality of AFEM is a direct consequence of Lemma 8.2 and Theorem 7.3. We prove this next.

Theorem 8.3 (convergence rate of AFEM). *Let $\gamma \in \mathbb{B}_t$ and $(u, f) \in \mathbb{A}_s(\gamma)$ for some $0 < t, s \leq n/d$. Let $\varepsilon_0 \leq (6\omega\Lambda_0 L^3)^{-1}$ be the initial tolerance, and the parameters θ, ω satisfy*

$$0 < \theta \leq \theta_*, \qquad 0 < \omega \leq \omega_* := \min\{\omega_1, \omega_2, \omega_3, \omega_4\}, \tag{8.11}$$

where $\theta_*, \omega_1, \cdots, \omega_4$ are given in (8.4), (6.4), (7.2), and (8.3), respectively. Let the procedure MARK select sets with minimal cardinality, and the procedure ADAPT_SURFACE be t-optimal on the surface γ. Let u be the solution of (4.13) and $\{\Gamma_k, \mathcal{T}_k, U_k\}_{k \geq 0}$ a sequence of approximate surfaces, meshes and discrete solution generated by the outer loop of AFEM.

Then there exists a constant C, depending on the Lipschitz constant L of γ, $\|f\|_{L^2(\gamma)}$, the refinement depth b, the initial triangulation $\mathcal{T}_0$, and AFEM parameters (θ, ω, ρ) such that

$$e(U_k) + \operatorname{osc}_{\mathcal{T}_k}(f) + \omega^{-1} \lambda_{\Gamma_k} \leq C \big(|u, f|_{\mathbb{A}_s}^{\frac{r}{s}} + \omega^{-r} |\gamma|_{\mathbb{B}_t}^{\frac{r}{t}} \big) \big(\#\mathcal{T}_k - \#\mathcal{T}_0 \big)^{-r}, \tag{8.12}$$

with $r = \min\{s, t\}$.

Proof. We start by noting that since $\omega\varepsilon_0 \leq \frac{1}{6\Lambda_0 L^3}$ the first output of ADAPT_SURFACE fulfills $\lambda_{\Gamma_0^+} \leq \frac{1}{6\Lambda_0}$ which is (5.7) and implies that $\mathbb{T}(\mathcal{T}_0^+)$ is shape regular.

There are two instances where elements are added, inside ADAPT_SURFACE and ADAPT_PDE. For ADAPT_SURFACE we make the assumption (6.2) of t-optimality:

$$\#\mathcal{M}_k^+ \lesssim \omega^{-\frac{1}{t}} |\gamma|_{\mathbb{B}_t}^{\frac{1}{t}} \varepsilon_k^{-\frac{1}{t}}.$$

For ADAPT_PDE, Lemma 8.2 yields

$$\#\mathcal{M}_k^j \lesssim |u, f|_{\mathbb{A}_s}^{\frac{1}{s}} \Big(e(U_k^j) + \operatorname{osc}_{\mathcal{T}_k^j}(f) \Big)^{-\frac{1}{s}} \qquad 0 \leq j < J,$$

with j denoting the inner loop iteration counter. Since the inner iterates of ADAPT_PDE satisfy Theorem 7.3 and

$$e(U_k^j) + \operatorname{osc}_{T_k^j}(f) \approx e(U_k^j) + \eta_{\mathcal{T}_k^j}(U_k^j),$$

we deduce that

$$\Big(e(U_k^j) + \operatorname{osc}_{\mathcal{T}_k^j}(f)\Big)^{-\frac{1}{s}} \lesssim \alpha^{\frac{J-j-1}{s}} \Big(e(U_k^{J-1}) + \eta_{\mathcal{T}_k^{J-1}}(f)\Big)^{-\frac{1}{s}} \leq \alpha^{\frac{J-j-1}{s}} \varepsilon_k^{-\frac{1}{s}}.$$

This implies

$$\sum_{j=0}^{J-1} \#\mathcal{M}_k^j \lesssim |u,f|_{\mathbb{A}_s}^{\frac{1}{s}} \varepsilon_k^{-\frac{1}{s}} \sum_{j=0}^{J-1} \alpha^{\frac{J-j-1}{s}} \lesssim |u,f|_{\mathbb{A}_s}^{\frac{1}{s}} \varepsilon_k^{-\frac{1}{s}}.$$

To do a full counting argument, we resort to the crucial estimate (6.6), which combined with the estimates above and the relation $\varepsilon_{k+1} = \rho\varepsilon_k$ of step 3 of AFEM gives

$$\#\mathcal{T}_k - \#\mathcal{T}_0 \leq C_6 \sum_{i=0}^{k-1} \Big(\#\mathcal{M}_i^+ + \sum_{j=0}^{J-1} \#\mathcal{M}_i^j\Big) \lesssim C_6 \Big(\omega^{-\frac{1}{t}} |\gamma|_{\mathbb{B}_t}^{\frac{1}{t}} + |u,f|_{\mathbb{A}_s}^{\frac{1}{s}}\Big) \sum_{i=0}^{k-1} \varepsilon_i^{-\frac{1}{r}},$$

where $r = \min\{s,t\}$. Since $\rho < 1$, we obtain $\sum_{i=0}^{k-1} \varepsilon_i^{-\frac{1}{r}} = \varepsilon_{k-1}^{-\frac{1}{r}} \sum_{i=0}^{k-1} \rho^{\frac{i}{r}} \lesssim \varepsilon_k^{-\frac{1}{r}}$, whence

$$\#\mathcal{T}_k - \#\mathcal{T}_0 \lesssim C_6 \Big(\omega^{-\frac{1}{t}} |\gamma|_{\mathbb{B}_t}^{\frac{1}{t}} + |u,f|_{\mathbb{A}_s}^{\frac{1}{s}}\Big) \varepsilon_k^{-\frac{1}{r}}.$$

Moreover, the stopping criteria (6.1) and (6.3) guarantee that

$$e(U_k) + \operatorname{osc}_{\mathcal{T}_k}(f) + \omega^{-1}\lambda_{\Gamma_k} \leq C\varepsilon_k,$$

which implies the desired estimate (8.12). □

Besides the condition $\omega \leq \omega_*$ in (8.11), the right-hand side of (8.12) suggests that ω should not be too small to optimize this bound. An optimal choice of ω for the case $s = t$, which unfortunately is not computable, appears to be

$$\omega = \min\big\{\omega_*, |u,f|_{\mathbb{A}_s} |\gamma|_{\mathbb{B}_s}^{-1}\big\}.$$

8.3 Greedy Algorithm

To conclude we show that ADAPT_SURFACE is t-optimal provided γ belongs to W_p^{1+td}, which is just above the nonlinear Sobolev scale of W_∞^1 for polynomial degree 1 in dimensions d:

$$\operatorname{sob}(W_\infty^1) = 1 - \frac{d}{\infty} = 1 < \operatorname{sob}(W_p^{1+td}) = 1 + td - \frac{d}{p} \quad \Rightarrow \quad tp > 1.$$

Proposition 8.4 (greedy algorithm)**.** *Let γ be piecewise of class $W_p^{1+td}(\Gamma_0)$, with $tp > 1$, $t \leq 1/d$, and globally of class W_∞^1. Then module $[\mathcal{T}^+, \Gamma^+] =$ ADAPT_SURFACE$(\mathcal{T}, \Gamma, \tau)$ terminates in a finite number of steps and the set $\mathcal{M}^+$ of marked elements satisfies*

$$\#\mathcal{M}^+ \leq C |\gamma|_{W_p^{1+td}(\Gamma_0)}^{1/t} \tau^{-1/t},$$

where $|\gamma|_{W_p^{1+td}(\Gamma_0)} = \Big(\sum_{i=1}^I |\mathcal{X}^i|_{W_p^{1+tp}(\Omega)}^p\Big)^{1/p}$. Moreover, $\gamma \in \mathbb{B}_t$ and

$$|\gamma|_{\mathbb{B}_t} \lesssim |\gamma|_{W_p^{1+td}(\Gamma_0)}.$$

Proof. We first observe that $W_p^{1+td} \subset W_\infty^1 \subset C^0$ so that the Lagrange interpolation operator $\mathcal{I}_\mathcal{T}$ is well defined. In addition, for an approximation pair $(\Gamma, \mathcal{T})$ local interpolation estimates give

$$\lambda_\Gamma(T) \lesssim h_T^r |\mathcal{X}|_{W_p^{1+td}(\widehat{T})}, \qquad \forall T \in \mathcal{T} = \mathcal{T}(\Gamma), \tag{8.13}$$

for $r = \mathrm{sob}(W_p^{1+td}) - \mathrm{sob}(W_\infty^1) = td - \frac{d}{p} > 0$. This shows that ADAPT_SURFACE terminates in finite number of steps, say m.

To prove that ADAPT_SURFACE is t-optimal, namely to show (6.2), let $\mathcal{M}^+ = \mathcal{M}_0 \cup \cdots \cup \mathcal{M}_{m-1}$ be the set of marked elements. We organize the elements in $\mathcal{M}^+$ by size in such a way that allows for a counting argument. Let $\mathcal{P}_j$ be the set of elements T of $\mathcal{M}^+$ with size

$$2^{-(j+1)} \le |T| < 2^{-j} \quad \Rightarrow \quad 2^{-(j+1)/d} \le h_T < 2^{-j/d}.$$

We recall that $|T|$ is the measure of $\overline{T}$, the preimage of T in the initial triangulation $\mathcal{T}_0$. We proceed in several steps.

1 We first observe that all T's in $\mathcal{P}_j$ are *disjoint*. This is because if T_1, $T_2 \in \mathcal{P}_j$ and $\mathring{T}_1 \cap \mathring{T}_2 \neq \emptyset$, then one of them is contained in the other, say $T_1 \subset T_2$, due to the bisection procedure. Thus $|T_1| \le \frac{1}{2}|T_2|$, contradicting the definition of $\mathcal{P}_j$. This implies

$$2^{-(j+1)} \, \#\mathcal{P}_j \le |\Gamma_0| \quad \Rightarrow \quad \#\mathcal{P}_j \le |\Gamma_0| \, 2^{j+1}. \tag{8.14}$$

2 In light of (8.13), we have for $T \in \mathcal{P}_j$

$$\tau \le \lambda_\Gamma(T) \lesssim 2^{-(j/d)r} |\mathcal{X}|_{W_p^{1+td}(\widehat{T})}.$$

Therefore $\tau^p \, \#\mathcal{P}_j \lesssim 2^{-(j/d)rp} \sum_{T \in \mathcal{P}_j} |\mathcal{X}|^p_{W_p^{1+td}(\widehat{T})} \le 2^{-(j/d)rp} \, |\gamma|^p_{W_p^{1+td}(\Gamma_0)}$, whence

$$\#\mathcal{P}_j \lesssim \tau^{-p} \, 2^{-(j/d)rp} \, |\gamma|^p_{W_p^{1+td}(\Gamma_0)}. \tag{8.15}$$

3 The two bounds for $\#\mathcal{P}_j$ in (8.14) and (8.15) are complementary. The first is good for j small whereas the second is suitable for j large (think of $\tau \ll 1$). The crossover takes place for j_0 such that

$$2^{j_0+1}|\Gamma_0| = \tau^{-p} \, 2^{-j_0 rp/d} |\mathcal{X}|^p_{W_p^{1+td}(\Omega)} \quad \Rightarrow \quad 2^{j_0} \approx \tau^{-1/t} \, \frac{|\gamma|^{1/t}_{W_p^{1+td}(\Gamma_0)}}{|\Gamma_0|^{1/tp}}.$$

4 We now compute

$$\#\mathcal{M}^+ = \sum_j \#\mathcal{P}_j \lesssim \sum_{j \le j_0} 2^j |\Gamma_0| + \tau^{-p} \, |\gamma|^p_{W_p^{1+t}(\Gamma_0)} \sum_{j > j_0} (2^{-rp/d})^j.$$

Since $\sum_{j \le j_0} 2^j \approx 2^{j_0}$, $\sum_{j > j_0} (2^{-rp/d})^j \lesssim 2^{-(rp/d)j_0} = 2^{(1-tp)j_0}$ we can write

$$\#\mathcal{M}^+ \lesssim \left(\tau^{-1/t} + \tau^{-p}\tau^{-1/t+p}\right) |\Gamma_0|^{1-1/tp} \, |\gamma|^{1/t}_{W_p^{1+td}(\Gamma_0)} \approx \tau^{-1/t} \, |\Gamma_0|^{1-1/tp} \, |\gamma|^{1/t}_{W_p^{1+td}(\Gamma_0)}.$$

5 Upon termination, $\lambda_{\Gamma^+} \le \tau$ and $\#\mathcal{M}^+ \lesssim \tau^{-1/t} |\gamma|^{1/t}_{W_p^{1+td}(\Gamma_0)}$, which is valid regardless of the input $\mathcal{T}$ of ADAPT_SURFACE. If we take $\mathcal{T} = \mathcal{T}_0$ and invoke Lemma 6.1 we deduce that

$$\#\mathcal{T}^+ - \#\mathcal{T}_0 \lesssim \tau^{-1/t} |\gamma|^{1/t}_{W_p^{1+td}(\Gamma_0)} \quad \Rightarrow \quad \lambda_{\Gamma^+} \left(\#\mathcal{T}^+ - \#\mathcal{T}_0\right)^t \lesssim |\gamma|_{W_p^{1+td}(\Gamma_0)},$$

and that $\gamma \in \mathbb{B}_t$. This concludes the proof. ∎

9 Asymptotics: Role of ω

In order to analyze the role of ω in the convergence rate of AFEM and its performance, we solve the problem

$$-\Delta_\gamma u = 1, \quad \text{in } \gamma, \qquad u = 0, \quad \text{on } \partial\gamma,$$

where γ is the graph of class $C^{1,\alpha}$ given by

$$z(x,y) = \left(0.75 - x^2 - y^2\right)_+^{1+\alpha},$$

over the flat domain $\Omega = (0,1)^2$, and consider two cases $\alpha = 3/5$ and $\alpha = 2/5$.

It turns out that $z \in W_p^{1+2t}(\Omega)$ for $t < (\alpha + \frac{1}{p})/2$. Moreover, to enforce the gap $\mathrm{sob}(W_p^{1+2t}) - \mathrm{sob}(W_\infty^1) > 0$ with $\mathrm{sob}(W_p^{1+2t}) = 1 + 2t - \frac{2}{p}$ and $\mathrm{sob}(W_\infty^1) = 1$ we need $tp > 1$. These conditions can be achieved provided (see §8.3)

$$\begin{aligned} \alpha = 3/5: &\qquad t = 1/2, \quad p > 2, &&\Rightarrow \qquad z \in \mathbb{B}_{\frac{1}{2}} \\ \alpha = 2/5: &\qquad t < 2/5, \quad p > 5/2, &&\Rightarrow \qquad z \in \mathbb{B}_t, \quad \forall t < 2/5. \end{aligned}$$

On the other hand $(u,f) \in \mathbb{A}_{\frac{1}{2}}$ in both cases. This is a consequence of the fact that $\Delta_\gamma u = f$ can be written in the parameter domain Ω as $\frac{1}{q}\widehat{\mathrm{div}}(q\mathbf{g}^{-1}\widehat{\nabla}^T u) = f$, with coefficient matrix $\mathbf{A} = q\mathbf{g}^{-1} \in C^\alpha(\bar\Omega) \cap W_p^1(\Omega)$ and $1 < p < \frac{1}{1-\alpha}$; see (5.5). Extending u and qf by odd reflection and $\mathbf{A}$ by even reflection to the unit squares around Ω, u is a solution to $-\mathrm{div}\left(\mathbf{A}\nabla^T u\right) = qf$ on the ball B centered at $(1/2,1/2)$ and radius 1, with coefficient $\mathbf{A} \in C^\alpha$ and right-hand side in L^∞. By Theorem 3.13 in [40] this implies $\nabla u \in C^\alpha(\bar B)$, and thus

$$\mathbf{A} : D^2 u = f + \mathrm{div}\mathbf{A} \cdot \nabla u \in L^p(B).$$

Applying Calderón-Zygmund theory we obtain $u \in W_p^2(\Omega)$ [39, Theorem 9.11], whence $(u,f) \in \mathbb{A}_{\frac{1}{2}}$ [49, §5.4].

In the following, we use the notation $\eta_k := \eta_{\mathcal{T}_k}(U_k)$, and $\lambda_k := \lambda_{\Gamma_k}$.

9.1 Case $\alpha = 3/5$

We recall that in this case $\gamma \in \mathbb{B}_{1/2}$. Since the pair $(u,f) \in \mathbb{A}_{1/2}$ we expect a decay of $\eta_k + \lambda_k/\omega$ proportional to $N_k^{-1/2}$, where $N_k = \#\mathcal{T}_k - \#\mathcal{T}_0$. In Figure 12 we plot $\eta_k + \lambda_k/\omega$ (left) and $\eta_k + \lambda_k$ (right) versus the number of elements in logarithmic scale for $\omega = 0.1, 1, 10$, and observe that in the three cases both notions of error decay (asymptotically) as $N^{-1/2}$.

In Figure 13 we show the behavior of the different indicators η_k, λ_k/ω and their sum, for the three values of ω considered above. We observe the following:

$\omega = 0.1$**.** At the beginning $\eta_k \ll \lambda_k/\omega$, thus λ_k in ADAPT_SURFACE guides the refinement initially, and η_k decreases very slowly because ADAPT_PDE exits without refining. The indicators η_k and λ_k/ω are of comparable size when the number of elements is around $2 \cdot 10^5$, when the refinement starts to occur due to both λ_k and η_k and both quantities decrease as $N^{-0.5}$.

$\omega = 1$**.** At the beginning $\eta_k < \lambda_k/\omega$ and the behavior is similar to the case $\omega = 0.1$. When the meshes have about 10^3 elements the curves for η_k and λ_k/ω meet and they both start to decrease at the optimal rate $N^{-0.5}$.

$\omega = 10$**.** At the beginning $\lambda_k/\omega < \eta_k$, and the situation is opposite to the case of ω small. The refinement is initially guided by η_k in ADAPT_PDE, and λ_k decreases very slowly because ADAPT_SURFACE exits without refining. The two curves for η_k and λ_k/ω meet when the meshes have about 10^4 elements, and they both start to decrease at the optimal rate $N^{-0.5}$.

In Figure 14 we show three meshes after 10, 20 and 30 refinements have been performed, with 192, 1216 and 5564 elements, respectively.

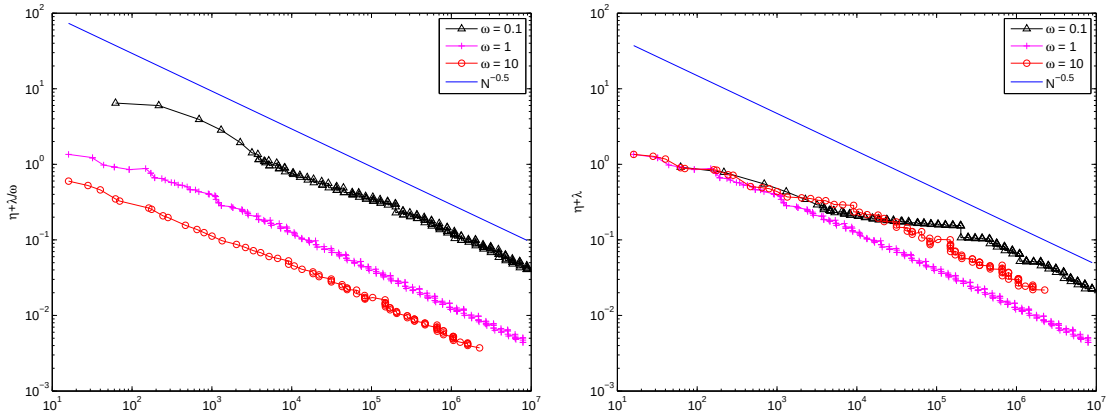
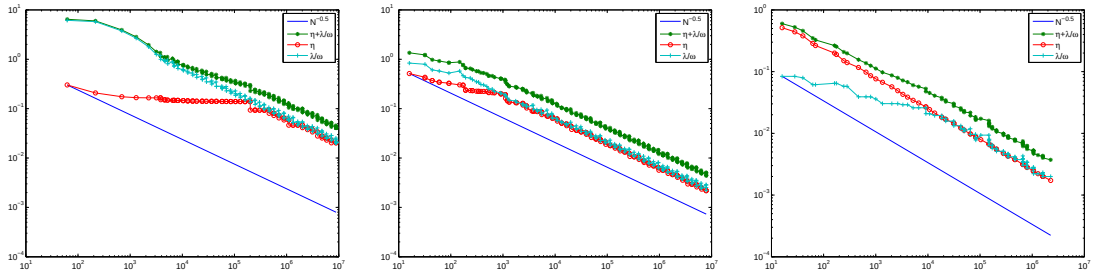


Figure 12: $\eta_k + \lambda_k/\omega$ (left) and $\eta_k + \lambda_k$ (right) versus the number of elements in logarithmic scale for $\omega = 0.1, 1, 10$. We observe that $\eta_k + \lambda_k/\omega$ decays as $N^{-0.5}$ right from the beginning, whereas $\eta_k + \lambda_k$ shows the same decay after the meshes have some refinement, depending on the value of ω. Our theory predicts the decay of $N^{-0.5}$ for both notions of total error if ω is sufficiently small, but the best relation between the error $\eta_k + \lambda_k$ and #DOFs is obtained for $w = 1$, which is *not so small*.



$\omega = 0.1$ $\omega = 1$ $\omega = 10$

Figure 13: η_k, λ_k/ω and $\eta_k + \lambda_k/\omega$ for $\omega = 0.1$ (left) $\omega = 1$ (middle) and $\omega = 10$ (right).

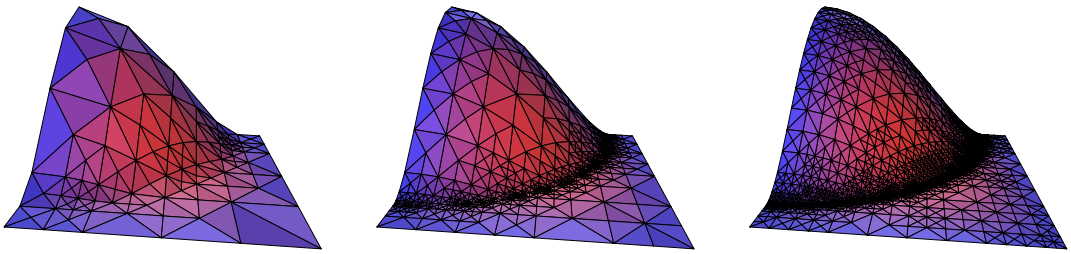


Figure 14: Meshes after 10, 20 and 30 refinements have been performed, $C^{1.6}$-surface, with $\omega = 1$. They are composed of 192, 1216 and 5564 elements, respectively.

9.2 Case $\alpha = 2/5$

We recall that in this case $\gamma \in \mathbb{B}_{0.4}$, whereas the pair $(u, f) \in \mathbb{A}_{1/2}$. We thus expect a decay of $\eta_k + \lambda_k/\omega$ proportional to $N^{-0.4}$. In Figure 15 we plot $\eta_k + \lambda_k/\omega$ (left) and $\eta_k + \lambda_k$ (right) versus the number of elements in logarithmic scale for $\omega = 0.1, 1, 10$, and observe that in the three cases both notions of error decay (asymptotically) as $N^{-0.4}$.

In Figure 16 we show the behavior of η_k, λ_k/ω and their sum, for the same values of ω. We observe the following:

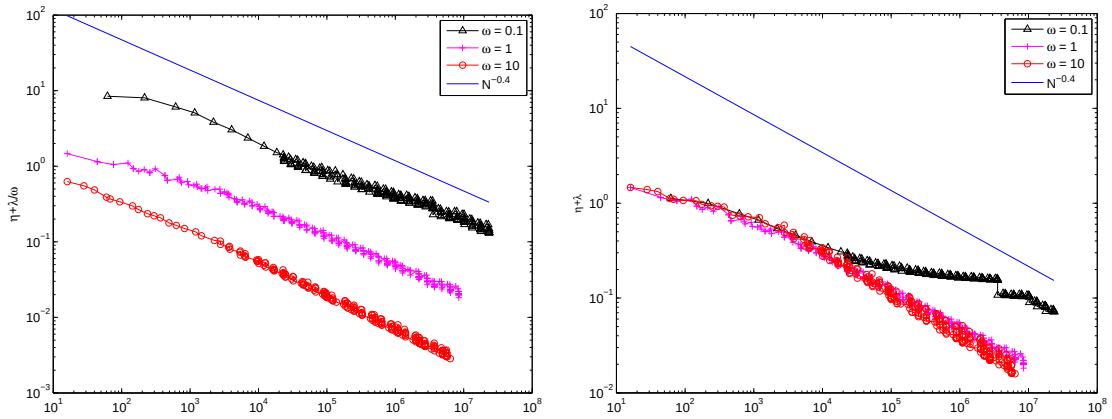
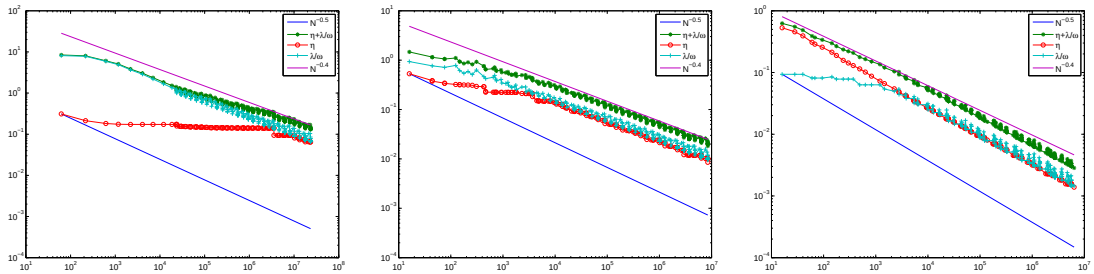


Figure 15: $\eta_k + \lambda_k/\omega$ (left) and $\eta_k + \lambda_k$ (right) versus the number of elements in logarithmic scale for $\omega = 0.1, 1, 10$. We observe that $\eta_k + \lambda_k/\omega$ decays as $N^{-0.4}$ right from the beginning, whereas $\eta_k + \lambda_k$ shows the same decay after the meshes have some refinement, depending on the value of ω. Our theory predicts the decay of $N^{-0.4}$ for both notions of total error if ω is sufficiently small. The best relation between the error $\eta_k + \lambda_k$ seems to occur for $\omega = 1$ and $\omega = 10$.



$\omega = 0.1$ $\omega = 1$ $\omega = 10$

Figure 16: η_k, λ_k/ω and $\eta_k + \lambda_k/\omega$ for $\omega = 0.1$ (left) $\omega = 1$ (middle) and $\omega = 10$ (right).

$\omega = 0.1$**.** At the beginning $\eta_k \ll \lambda_k/\omega$, thus λ_k in ADAPT_SURFACE guides the refinement initially, and η_k decreases very slowly because ADAPT_PDE exits without refining. The asymptotic regime starts when both indicators have a comparable magnitude, and both quantities decrease as $N^{-0.4}$. This instance is reached when the meshes have more than 10^6 elements, because λ_k/ω decreases more slowly than in the previous example, and takes longer to reach the initial value of η_k.

$\omega = 1$**.** This case is similar to the previous one, with the change of behavior occuring when the meshes have 10^4 elements.

$\omega = 10$**.** The situation now is opposite to the previous cases of ω small. At the beginning the refinement is initially guided by η_k in ADAPT_PDE, and λ_k decreases very slowly because ADAPT_SURFACE exits without refining. It is interesting to notice that η_k decreases as $N^{-0.5}$ in this transient initial phase. When the meshes have about 10^3 elements both indicators are of comparable size, and the overall rate seems to be a little bit better than $N^{-0.4}$. This happens because λ_k is divided by 10 and its effect is not so visible in the picture. In the long run the decay cannot be better than $N^{-0.4}$.

It is also interesting to notice that λ_k does not decrease monotonically, mainly because the strongly curved part of γ is not aligned to the grid. This behavior is consistent with (3.6) and, in fact, shows that we cannot expect monotonicity of $\lambda_\Gamma(\mathcal{T})$ upon refinement, thereby justifying (3.6).

10 Conclusions and Comments

We finish the paper with the following remarks about this and related work.

- **Coupling PDE-Geometry:** This is a new feature in adaptivity and leads to separate handling of geometry and PDE resolution with specific relative tolerances. The current algorithm is different from that for graphs [46] studied by K. Mekchay, P. Morin, and R.H. Nochetto. The present paper studies polynomial degree 1, but the theory for parametric surfaces extends to higher polynomial degree [12].

- **Convergence rates:** We show optimal convergence rates in the energy norm

$$\|\nabla(u - U_k)\|_{L^2(\gamma)} \lesssim (\#\mathcal{T}_k - \#\mathcal{T}_0)^{-s}$$

  provided this is the rate of the best approximation of u in H^1 and that of γ in W^1_∞. This optimal result is consistent with that derived for flat domains by R. Stevenson [51] for the Poisson equation with data in H^{-1} and by J.M. Cascón et al [21] for elliptic PDE with variable coefficients. None of them involve coarsening as the seminal paper [11] by P. Binev, W. Dahmen and R. DeVore. The present estimates extend those in [21] to the Laplace-Beltrami operator.

- **Weaker conditions on f:** We refer to A. Cohen, R. DeVore, and R.H. Nochetto [23] for convergence rates of elliptic PDE in flat domains with $f \in H^{-1}$ and A piecewise constant:

$$\operatorname{div}(A\nabla u) = f. \tag{10.1}$$

  Paper [23] shows that approximability of u is sufficient for a complete theory. Whether this is true for the Laplace-Beltrami operator is still an open question.

- **Weaker conditions on γ:** We assume γ is W^2_p with $p > d$, which implies γ is C^1. In the flat case, this corresponds to piecewise continuous A. We refer to A. Bonito, R. DeVore, and R.H. Nochetto [13] for optimal convergence rates of AFEM for (10.1) with weaker regularity assumptions on A. This could be especially relevant to perform adaptivity on problems where the singularity location is not known beforehand, such as those in §2.

References

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

[2] E. Bänsch. Finite element discretization of the Navier-Stokes equations with a free capillary surface. *Numer. Math.*, 88(2):203–235, 2001.

[3] E. Bänsch, F. Haußer, O. Lakkis, B. Li, and A. Voigt. Finite element method for epitaxial growth with attachment-detachment kinetics. *J. Comput. Phys.*, 194(2):409–434, 2004.

[4] E. Bänsch, P. Morin, and R. H. Nochetto. A finite element method for surface diffusion: the parametric case. *J. Comput. Phys.*, 203(1):321–343, 2005.

[5] E. Bänsch, P. Morin, and R.H. Nochetto. Surface diffusion of graphs: variational formulation, error analysis, and simulation. *SIAM J. Numer. Anal.*, 42(2):773–799 (electronic), 2004.

[6] J.W. Barrett, H. Garcke, and R. Nürnberg. A parametric finite element method for fourth order geometric evolution equations. *J. Comput. Phys.*, 222(1):441–462, 2007.

[7] J.W. Barrett, H. Garcke, and R. Nürnberg. Parametric approximation of Willmore flow and related geometric evolution equations. *SIAM J. Sci. Comput.*, 31(1):225–253, 2008.

[8] J.W. Barrett, H. Garcke, and R. Nürnberg. Finite-element approximation of coupled surface and grain boundary motion with applications to thermal grooving and sintering. *European J. Appl. Math.*, 21(6):519–556, 2010.

[9] S. Bartels, G. Dolzmann, and R. H. Nochetto. A finite element scheme for the evolution of orientation order in fluid membranes. *M2AN Math. Model. Numer. Anal.*, 44(1):1–31, 2010.

[10] S. Bartels, G. Dolzmann, and R. H. Nochetto. Finite element methods for director fields on flexible surfaces. *Interfaces Free Bound.*, (to appear).

[11] P. Binev, W. Dahmen, and R. DeVore. Adaptive finite element methods with convergence rates. *Numer. Math.*, 97(2):219–268, 2004.

[12] A. Bonito, J.M. Cascón, K. Mekchay, P. Morin, and R. H. Nochetto. AFEM for the Laplace-Beltrami operator on parametric surfaces: convergence rates. in preparation.

[13] A Bonito, R DeVore, and R.H. Nochetto. Adaptive finite element methods for elliptic problems with discontinuous coefficients. in preparation.

[14] A. Bonito and R. H. Nochetto. Quasi-optimal convergence rate of an adaptive discontinuous Galerkin method. *SIAM J. Numer. Anal.*, 48(2):734–771, 2010.

[15] A. Bonito, R. H. Nochetto, and M. S. Pauletti. Geometrically consistent mesh modification. *SIAM J. Numer. Anal.*, 48(5):1877–1899, 2010.

[16] A. Bonito, R. H. Nochetto, and M. S. Pauletti. Dynamics of biomembranes: effect of the bulk fluid. *Math. Model. Nat. Phenom.*, 6(5):25–43, 2011.

[17] A. Bonito, R.H. Nochetto, and M.S. Pauletti. Parametric FEM for geometric biomembranes. *J. Comput. Phys.*, 229(9):3171–3188, 2010.

[18] A. Bonito and J.E. Pasciak. Convergence analysis of variational and non-variational multigrid algorithm for the laplace-beltrami operator. *Math. Comp.*, 81:1263–1288, 2012.

[19] S.C. Brenner and L. R. Scott. *The mathematical theory of finite element methods*, volume 15 of *Texts in Applied Mathematics*. Springer-Verlag, New York, second edition, 2002.

[20] J. Cahn and J.E. Taylor. Surface motion by surface diffusion. *Acta Metall. Mater.*, 42:1045–1063, 1994.

[21] J. M. Cascón, C. Kreuzer, R. H. Nochetto, and K. G. Siebert. Quasi-optimal convergence rate for an adaptive finite element method. *SIAM J. Numer. Anal.*, 46(5):2524–2550, 2008.

[22] P. G. Ciarlet. *The finite element method for elliptic problems*. North-Holland Publishing Co., Amsterdam, 1978. Studies in Mathematics and its Applications, Vol. 4.

[23] A Cohen, R DeVore, and R.H. Nochetto. Convergence rates for afem with h^{-1} data. *Found. Comput. Math.* in preparation.

[24] K. Deckelnick and G. Dziuk. Discrete anisotropic curvature flow of graphs. *M2AN Math. Model. Numer. Anal.*, 33(6):1203–1222, 1999.

[25] K. Deckelnick and G. Dziuk. Error estimates for a semi-implicit fully discrete finite element scheme for the mean curvature flow of graphs. *Interfaces Free Bound.*, 2(4):341–359, 2000.

[26] K. Deckelnick and G. Dziuk. Error analysis of a finite element method for the Willmore flow of graphs. *Interfaces Free Bound.*, 8(1):21–46, 2006.

[27] K. Deckelnick, G. Dziuk, and C.M. Elliott. Fully discrete finite element approximation for anisotropic surface diffusion of graphs. *SIAM J. Numer. Anal.*, 43(3):1112–1138 (electronic), 2005.

[28] 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.

[29] 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.

[30] U. Dierkes, S. Hildebrandt, A. Küster, and O. Wohlrab. *Minimal surfaces. I*, volume 295 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1992. Boundary value problems.

[31] M. P. do Carmo. *Differential geometry of curves and surfaces*. Prentice-Hall Inc., Englewood Cliffs, N.J., 1976. Translated from the Portuguese.

[32] G. Doğan, P. Morin, R.H. Nochetto, and M. Verani. Discrete gradient flows for shape optimization and applications. *Comput. Methods Appl. Mech. Engrg.*, 196(37-40):3898–3914, 2007.

[33] G. Doğan and R. H. Nochetto. First variation of the general curvature-dependent surface energy. *ESAIM Math. Model. Numer. Anal.*, 46(1):59–79, 2012.

[34] Q. Du, C. Liu, and X. Wang. Simulating the deformation of vesicle membranes under elastic bending energy in three dimensions. *J. Comput. Phys.*, 212(2):757–777, 2006.

[35] 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.

[36] G. Dziuk. An algorithm for evolutionary surfaces. *Numer. Math.*, 58(6):603–611, 1991.

[37] G. Dziuk. Computational parametric willmore flow. *Numer. Math.*, 111(1):55–80, 2008.

[38] Ch. M. Elliott and B. Stinner. Modeling and computation of two phase geometric biomembranes using surface finite elements. *J. Comput. Phys.*, 229(18):6585–6612, 2010.

[39] 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.

[40] Q. Han and F. Lin. *Elliptic partial differential equations*, volume 1 of *Courant Lecture Notes in Mathematics*. Courant Institute of Mathematical Sciences, New York, second edition, 2011.

[41] W. Helfrich. Elastic properties of lipid bilayers - theory and possible experiments. *Zeitschrift Fur Naturforschung C-A Journal Of Biosciences*, 28:693, 1973.

[42] K. Khairy, J. Foo, and J. Howard. Shapes of red blood cells: Comparison of 3d confocal images with the bilayer-couple model. *Cellular and Molecular Bioengineering*, 1:173–181, 2008. 10.1007/s12195-008-0019-5.

[43] R. Kornhuber and H. Yserentant. Multigrid methods for discrete elliptic problems on triangular surfaces. *Comput. Vis. Sci.*, 11(4-6):251–257, 2008.

[44] O. Lakkis and R.H. Nochetto. A posteriori error analysis for the mean curvature flow of graphs. *SIAM J. Numer. Anal.*, 42(5):1875–1898 (electronic), 2005.

[45] M. Laradji and O. G. Mouritsen. Elastic properties of surfactant monolayers at liquid-liquid interfaces: A molecular dynamics study. *J. Chem. Phys.*, 112(19):8621–8630, 2000.

[46] K. Mekchay, P. Morin, and R. H. Nochetto. AFEM for the Laplace-Beltrami operator on graphs: design and conditional contraction property. *Math. Comp.*, 80(274):625–648, 2011.

[47] P. Morin, R. H. Nochetto, and K. G. Siebert. Convergence of adaptive finite element methods. *SIAM Rev.*, 44(4):631–658 (electronic) (2003), 2002. Revised reprint of "Data oscillation and convergence of adaptive FEM" [SIAM J. Numer. Anal. **3**8 (2000), no. 2, 466–488 (electronic); MR1770058 (2001g:65157)].

[48] P. Morin, R.H. Nochetto, and K. G. Siebert. Data oscillation and convergence of adaptive FEM. *SIAM J. Numer. Anal.*, 38(2):466–488 (electronic), 2000.

[49] R. H. Nochetto, K. G. Siebert, and A. Veeser. Theory of adaptive finite element methods: an introduction. In *Multiscale, nonlinear and adaptive approximation*, pages 409–542. Springer, Berlin, 2009.

[50] R.E. Rusu. An algorithm for the elastic flow of surfaces. *Interfaces Free Bound.*, 7(3):229–239, 2005.

[51] R. Stevenson. Optimality of a standard adaptive finite element method. *Found. Comput. Math.*, 7(2):245–269, 2007.

[52] R. Stevenson. The completion of locally refined simplicial partitions created by bisection. *Math. Comp.*, 77(261):227–241 (electronic), 2008.

[53] J.E. Taylor. Some mathematical challenges in materials science. *Bull. Amer. Math. Soc. (N.S.)*, 40(1):69–87 (electronic), 2003. Mathematical challenges of the 21st century (Los Angeles, CA, 2000).

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

[55] T. J. Willmore. *Riemannian geometry*. Oxford Science Publications. The Clarendon Press Oxford University Press, New York, 1993.