

## Stabilization of high order cut finite element methods on surfaces

MATS G. LARSON

Department of Mathematics and Mathematical Statistics, Umeå University, SE-90187 Umeå, Sweden  
mats.larson@math.umu.se

AND

SARA ZAHEDI\*

Department of Mathematics, KTH Royal Institute of Technology, SE-100 44 Stockholm, Sweden

\*Corresponding author: sara.zahedi@math.kth.se

[Received on 9 October 2017; revised on 20 December 2018]

We develop and analyse a stabilization term for cut finite element approximations of an elliptic second-order partial differential equation on a surface embedded in  $\mathbb{R}^d$ . The new stabilization term combines properly scaled normal derivatives at the surface together with control of the jump in the normal derivatives across faces, and provides control of the variation of the finite element solution on the active three-dimensional elements that intersect the surface. We show that the condition number of the stiffness matrix is  $O(h^{-2})$ , where  $h$  is the mesh parameter. The stabilization term works for linear as well as for higher-order elements and the derivation of its stabilizing properties is quite straightforward, which we illustrate by discussing the extension of the analysis to general  $n$ -dimensional smooth manifolds embedded in  $\mathbb{R}^d$ , with codimension  $d - n$ . We also state the properties of a general stabilization term that are sufficient to prove optimal scaling of the condition number and optimal error estimates in energy- and  $L^2$ -norm. We finally present numerical studies confirming our theoretical results.

**Keywords:** CutFEM; high order discretization; PDEs on surfaces; properties of stabilization; condition number; error estimates.

### 1. Introduction

**Cut finite element method for surface partial differential equations.** CutFEM (Cut Finite Element Method) provides a new high order finite element method for solution of partial differential equations (PDEs) on surfaces. The main approach is to embed the surface into a three-dimensional mesh, and to use the restriction of the finite element functions to the surface as trial and test functions. More precisely, let  $\Gamma$  be a smooth closed surface embedded in  $\mathbb{R}^3$  with exterior unit normal  $n$  and consider the Laplace–Beltrami problem: find  $u \in H^1(\Gamma)$  such that

$$a(u, v) = l(v) \quad \forall v \in H^1(\Gamma). \quad (1.1)$$

The forms are defined by

$$a(w, v) = \int_{\Gamma} \nabla_{\Gamma} w \cdot \nabla_{\Gamma} v, \quad l(v) = \int_{\Gamma} f v, \quad (1.2)$$

where  $\nabla_{\Gamma}$  is the surface gradient and  $f \in L^2(\Gamma)$  a given function with average zero. Let now  $\Gamma$  be embedded into a polygonal domain  $\Omega$  equipped with a quasi-uniform mesh  $\mathcal{T}_{h,0}$  and let  $\Gamma_h$  be a family

of discrete surfaces that converges to  $\Gamma$  in an appropriate manner. Let the active mesh  $\mathcal{T}_h$  consist of all elements that intersect  $\Gamma_h$ . Let  $V_h$  be the space of continuous piecewise polynomial functions on  $\mathcal{T}_h$  with average zero on  $\Gamma_h$ . The basic CutFEM takes the form; find  $u_h \in V_h$  such that

$$a_h(u_h, v) = l_h(v) \quad \forall v \in V_h, \quad (1.3)$$

where

$$a_h(w, v) = \int_{\Gamma_h} \nabla_\Gamma w \cdot \nabla_\Gamma v, \quad l_h(v) = \int_{\Gamma_h} f_h v, \quad (1.4)$$

and  $f_h$  is a suitable representation of  $f$  on  $\Gamma_h$ . In the context of surface partial differential equations this approach, also called trace finite elements, was first introduced by Olshanskii *et al.* (2009) and has then been developed in different directions including high order approximations (see, e.g., Reusken, 2015), discontinuous Galerkin methods (see Burman *et al.*, 2017), transport problems (see, e.g., Olshanskii *et al.*, 2014b), embedded membranes (see, e.g., Cenovic *et al.*, 2016), coupled bulk-surface problems (see, e.g., Gross *et al.*, 2015; Burman *et al.*, 2016c), minimal surface problems (see, e.g., Cenovic *et al.*, 2015) and time-dependent problems on evolving surfaces (see, e.g., Olshanskii & Reusken, 2014; Olshanskii *et al.*, 2014a; Hansbo *et al.*, 2016; Zahedi, 2018). We also refer to the overview article by Burman *et al.* (2015a) and the references therein.

**Previous work on preconditioning and stabilization.** The basic CutFEM method (1.3) manufactures a potentially ill-conditioned linear system of equations and stabilization or preconditioning is therefore needed. It has been shown by Olshanskii & Reusken (2010) that diagonal preconditioning works for piecewise linear approximation and by Reusken (2015) that also quadratic elements can be preconditioned if the full gradient is used instead of the tangential gradient in the bilinear form  $a_h$ , see also the work by Deckelnick *et al.* (2014) where this formulation was proposed.

A stabilized version of (1.3) was introduced by Burman *et al.* (2015b). The method takes the form

$$a_h(u_h, v) + s_h(u_h, v) = l(v) \quad \forall v \in V_h, \quad (1.5)$$

where  $s_h$  is a stabilization form, which is added in order to ensure stability of the method. More precisely, one can show that there is a constant such that for all  $v \in V_h$ ,

$$\|v\|_{L^2(\mathcal{T}_h)}^2 \lesssim h(a_h(v, v) + s_h(v, v)). \quad (1.6)$$

Using (1.6) and the properties of  $a_h$ , it can be shown that the condition number  $\kappa$  of the stiffness matrix satisfies the optimal bound

$$\kappa \lesssim h^{-2}. \quad (1.7)$$

In the work by Burman *et al.* (2015b) only piecewise linear finite elements were considered and the stabilization term was defined by

$$s_h(w, v) = c_F([D_{n_F}^1 w], [D_{n_F}^1 v])_{\mathcal{F}_h}, \quad (1.8)$$

where  $[D_{n_F}^1 v]$  is the jump in the normal derivative of  $v$  at the face  $F$ ,  $\mathcal{F}_h$  is the set of internal faces (i.e., faces with two neighbours) in the active mesh  $\mathcal{T}_h$  and  $c_F > 0$  is a stabilization parameter. Optimal

order bounds on the error and condition number were established. Furthermore, for linear elements the so-called full gradient stabilization

$$s_h(v, w) = ch(\nabla v, \nabla w)_{\mathcal{T}_h} \quad (1.9)$$

was proposed and analysed by [Burman \*et al.\* \(2016b\)](#) as a simpler alternative, and for higher-order elements the normal gradient stabilization

$$s_h(w, v) = c \mathcal{J} h^\alpha (n_h \cdot \nabla w, n_h \cdot \nabla v)_{\mathcal{T}_h}, \quad \alpha \in [-1, 1], \quad (1.10)$$

was proposed and analysed independently by [Burman \*et al.\* \(2018\)](#) and [Grande \*et al.\* \(2018\)](#). The restrictions on  $\alpha$  essentially comes from a lower bound, which implies that optimal order *a priori* error estimates hold, and an upper bound, which implies that the desired stability estimate (1.6) holds. Note that stronger control than (1.6) is sometimes demanded, for instance in cut discontinuous Galerkin methods (see [Burman \*et al.\*, 2017](#)), which may lead to stronger upper restrictions on  $\alpha$ .

The last two stabilization terms, (1.9) and (1.10), can however not be used in the discretization of bulk problems without destroying the convergence order while the ghost penalty stabilization term (see [Burman, 2010](#)), which as (1.8) acts on the element faces, can be used. Thus, in many applications where both bulk and surface problems occur, face stabilization is still needed when discretizing the problem with CutFEM. It should be noted that the stabilization term (1.8) has been used also for other reasons than controlling the condition number. It is used for example by [Burman \*et al.\* \(2019\)](#) to stabilize for convection and by [Hansbo \*et al.\* \(2015\)](#) to improve on the accuracy of computing the mean curvature vector of a surface. We study the problem of computing the mean curvature vector of a surface based on the Laplace–Beltrami operator in Section 9.3.

**New contributions.** In this work we consider stabilization for linear as well as higher-order elements using a combination of face stabilization and normal derivative stabilization at the actual surface. More precisely, the stabilization term takes the form

$$s_h(v, w) = s_{h,F}(v, w) + s_{h,\Gamma}(v, w) \quad (1.11)$$

with

$$s_{h,F}(v, w) = \sum_{j=1}^p c_{F,j} h^{2(j-1+\gamma)} ([D_{n_F}^j v], [D_{n_F}^j w])_{\mathcal{F}_h}, \quad (1.12)$$

$$s_{h,\Gamma}(v, w) = \sum_{j=1}^p c_{\Gamma,j} h^{2(j-1+\gamma)} (D_{n_h}^j v, D_{n_h}^j w)_{\Gamma_h}, \quad (1.13)$$

where  $\gamma \in [0, 1]$  is a parameter,  $[D_{n_F}^j v]$  is the jump in the normal derivative of order  $j$  across the face  $F$  and  $c_{F,j} > 0$ ,  $c_{\Gamma,j} > 0$ , are stabilization constants. This stabilization term with  $\gamma = 1$  has been used with a high order space-time CutFEM for PDEs on evolving surfaces by [Zahedi \(2018\)](#). We note that also for this stabilization term there is a certain range of scaling with the mesh parameter  $h$ , and in fact the following stronger version of (1.6) holds for  $v \in V_h$

$$\|v\|_{L^2(\mathcal{T}_h)}^2 + h^{2\gamma} \|\nabla v\|_{L^2(\mathcal{T}_h)}^2 \lesssim h(a_h(v, v) + s_h(v, v)), \quad (1.14)$$

where we get stronger control of the gradient for smaller  $\gamma \in [0, 1]$ , which corresponds to stronger stabilization.

Using the combination of the face and surface stabilization terms, the proof of (1.6) consists of two steps:

- Using the face terms we can estimate the  $L^2$  norm of a finite element function at an element in terms of the  $L^2$  norm at a neighbouring element, which share a face and the stabilization term on the shared face. Repeating this procedure we can pass from any element to an element, which has a large intersection,  $|T \cap \Gamma_h| \gtrsim h^2$ , with  $\Gamma_h$ .
- Using the stabilization of normal derivatives at the surface, we can control a finite element function on elements that have large intersection with  $\Gamma_h$ .

Using shape regularity we can show that one can pass from any element in  $\mathcal{T}_h$  to an element with a large intersection in a uniformly bounded number of steps (see Demlow & Olshanskii, 2012; Burman *et al.*, 2015b). The analysis is quite straightforward and may in fact also be extended to the case on an  $n$ -dimensional smooth manifold embedded in  $\mathbb{R}^d$  for general codimension  $cd = d - n$ . An important special case is a curve embedded in  $\mathbb{R}^3$ , which may for instance be an intersection of two surfaces (see also Burman *et al.*, 2018) for generalizations to so-called mixed dimensional problems. For clarity, we consider a two-dimensional surface embedded in  $\mathbb{R}^3$  throughout the paper and discuss the minor modifications necessary in a separate section. We also refer to the work by Burman *et al.* (2018) for general background on the analysis of problems with higher codimension embeddings.

**Outline.** In Section 2 we introduce the model problem, in Section 3 we formulate the finite element method, in Section 4 we collect some basic preliminary results, in Section 5 we formulate the properties of the stabilization term, in Section 6 we establish an optimal order condition number estimate, in Section 7 we discuss the extension to problems on  $n$ -dimensional smooth manifolds embedded in  $\mathbb{R}^d$ , in Section 8 we prove optimal order *a priori* error estimates, and in Section 9 we present numerical results that support our theoretical findings. In the last section, Section 10, we conclude and discuss the advantages with the proposed stabilization.

## 2. The Laplace–Beltrami problem on a surface

**The surface.** Let  $\Gamma$  be a smooth, closed, simply connected surface in  $\mathbb{R}^3$  with exterior unit normal  $n$  embedded in a polygonal domain  $\Omega \subset \mathbb{R}^3$ . Let  $U_\delta(\Gamma)$  be the open tubular neighbourhood of  $\Gamma$  with thickness  $\delta > 0$ ,

$$U_\delta(\Gamma) = \{x \in \mathbb{R}^3 : |d(x)| < \delta\}, \quad (2.1)$$

where  $d(x)$  is the signed distance function of  $\Gamma$  with  $d < 0$  in the interior of  $\Gamma$  and  $d > 0$  in the exterior. Note that  $n = \nabla d$  is the outward-pointing unit normal on  $\Gamma$ . Then there is  $\delta_0 > 0$  such that for each  $x \in U_{\delta_0}(\Gamma)$  the closest point projection:

$$p(x) = x - d(x)n(p(x)) \quad (2.2)$$

maps  $x$  to a unique closest point on  $\Gamma$ . In particular, we require

$$\delta_0 \max_{i=1,2} \|\kappa_i\|_{L^\infty(\Gamma)} < 1, \quad (2.3)$$

where for  $x \in \Gamma$ ,  $\kappa_i$  are the principal curvatures. We also define the extension of a function  $u$  on  $\Gamma$  to  $U_{\delta_0}(\Gamma)$  by

$$u^e(x) = u(p(x)), \quad x \in U_{\delta_0}(\Gamma). \quad (2.4)$$

**The Problem.** We consider the problem: find  $u \in H_0^1(\Gamma) = \{v \in H^1(\Gamma) \mid \int_{\Gamma} v \, ds = 0\}$  such that

$$a(u, v) = l(v) \quad \forall v \in H_0^1(\Gamma), \quad (2.5)$$

where

$$a(u, v) = \int_{\Gamma} \nabla_{\Gamma} u \cdot \nabla_{\Gamma} v \, ds, \quad l(v) = \int_{\Gamma} f v \, ds. \quad (2.6)$$

Here  $\nabla_{\Gamma}$  is the tangential gradient on  $\Gamma$  defined by

$$\nabla_{\Gamma} u = P_{\Gamma} \nabla u^e, \quad (2.7)$$

where

$$P_{\Gamma} = I - n \otimes n. \quad (2.8)$$

We used the extension of  $u$ , which is constant in the normal direction to  $\Gamma$  to formally define  $\nabla_{\Gamma} u$ . However, the tangential derivative depends only on the values of  $u$  on  $\Gamma$  and does not depend on the particular choice of extension.

Using Lax–Milgram’s lemma we conclude that there is a unique solution to (2.5), and we also have the elliptic regularity estimate

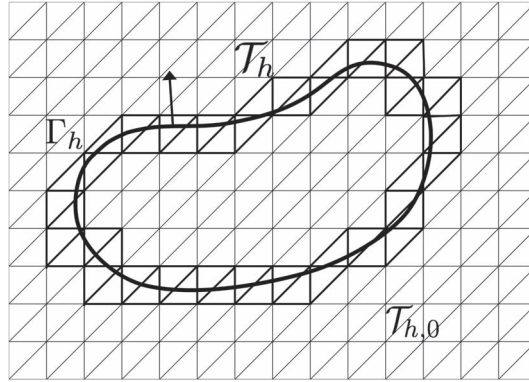
$$\|u\|_{H^{s+2}(\Gamma)} \lesssim \|f\|_{H^s(\Gamma)}, \quad s \in \mathbb{Z}, \, s \geq -1. \quad (2.9)$$

### 3. The method

#### 3.1 The mesh and the finite element space

- *The background mesh and space:* Let  $\mathcal{T}_{h,0}$  be a quasi-uniform partition of  $\Omega$  with mesh parameter  $h \in (0, h_0]$  into shape regular tetrahedra. On  $\mathcal{T}_{h,0}$  we define  $\tilde{V}_h^p$  to be the space of continuous piecewise polynomials of degree  $p \geq 1$ .
- *The discrete surfaces:* Let  $\Gamma_h$  be a sequence of surfaces that converges to  $\Gamma$ . We specify the precise assumptions on the convergence below.
- *The active mesh and the induced surface mesh:* We denote by  $\mathcal{T}_h$  the set consisting of elements in the background mesh that are cut by the discrete geometry  $\Gamma_h$ . These elements form the so-called active mesh. See Fig. 1 for an illustration. The restriction of the active mesh to the discrete surface  $\Gamma_h$  manufactures an induced cut surface mesh that we denote by

$$\mathcal{K}_h = \{K = T \cap \Gamma_h : T \in \mathcal{T}_h\}. \quad (3.1)$$


 FIG. 1. The discrete surface  $\Gamma_h$ , the background mesh  $\mathcal{T}_{h,0}$  and the active mesh  $\mathcal{T}_h$ .

- *The finite element space:* The restriction of the background space to the active mesh together with the condition  $\int_{\Gamma_h} v \, ds_h = 0$  defines our finite element space  $V_h^p$ ,

$$V_h^p = \left\{ v_h \in \tilde{V}_h^p|_{\mathcal{T}_h} : \int_{\Gamma_h} v \, ds_h = 0 \right\}. \quad (3.2)$$

### 3.2 The finite element method

**Method.** Find  $u_h \in V_h^p$  such that

$$A_h(u_h, v) := a_h(u_h, v) + s_h(u_h, v) = l_h(v) \quad \forall v_h \in V_h^p. \quad (3.3)$$

To enforce the condition  $\int_{\Gamma_h} v \, ds_h = 0$  one can use a Lagrange multiplier.

**Forms.**

- The forms  $a_h$  and  $l_h$  are defined by

$$a_h(w, v) = (\nabla_{\Gamma_h} w, \nabla_{\Gamma_h} v)_{\Gamma_h}, \quad l_h(v) = (f_h, v)_{\Gamma_h}, \quad (3.4)$$

where the tangential gradient  $\nabla_{\Gamma_h}$  is defined by

$$\nabla_{\Gamma_h} u_h = P_{\Gamma_h} \nabla u_h \quad (3.5)$$

with

$$P_{\Gamma_h} = I - n_h \otimes n_h. \quad (3.6)$$

- The stabilization form  $s_h$  is defined by

$$s_h(w, v) = s_{h,F}(w, v) + s_{h,\Gamma}(w, v), \quad (3.7)$$

where

$$s_{h,F}(w, v) = \sum_{j=1}^p c_{F,j} h^{2(j-1+\gamma)} ([D_{n_F}^j w], [D_{n_F}^j v])_{\mathcal{F}_h}, \quad (3.8)$$

$$s_{h,\Gamma}(w, v) = \sum_{j=1}^p c_{\Gamma,j} h^{2(j-1+\gamma)} (D_{n_h}^j w, D_{n_h}^j v)_{\mathcal{K}_h}. \quad (3.9)$$

Here  $(D_a^j v)|_x$  is the  $j$ th directional derivative of  $v$  along the line defined by the unit vector  $a(x)$ ,  $[w]|_F$  denotes the jump of  $w$  over the face  $F$ ,  $\mathcal{F}_h$  is the set of internal faces (i.e., faces with two neighbours) in the active mesh  $\mathcal{T}_h$ ,  $c_{\Gamma,j} > 0$ ,  $c_{F,j} > 0$  are stabilization parameters and  $\gamma \in [0, 1]$  is a parameter.

REMARK 3.1 In the special case  $p = 1$  we obtain

$$s_h(w, v) = c_F h^{2\gamma} ([D_{n_F}^1 w], [D_{n_F}^1 v])_{\mathcal{F}_h} + c_\Gamma h^{2\gamma} (D_{n_h}^1 w, D_{n_h}^1 v)_{\mathcal{K}_h}, \quad (3.10)$$

where  $\gamma \in [0, 1]$ , which can be compared with the pure face penalty term

$$s_h(w, v) = c_F ([D_{n_F}^1 w], [D_{n_F}^1 v])_{\mathcal{F}_h} \quad (3.11)$$

developed by [Burman et al. \(2015b\)](#). Note that the scaling with the mesh parameter  $h$  is different in the two terms and that for  $\gamma > 0$  the face penalty term is weaker in the stabilization term we introduce here.

### 3.3 Assumptions

**A1.** We assume that the closed, simply connected approximation  $\Gamma_h$  and its unit normal  $n_h$  are such that for all  $h \in (0, h_0]$ ,

$$\|d\|_{L^\infty(\Gamma_h)} \lesssim h^{p+1}, \quad (3.12)$$

$$\|n - n_h\|_{L^\infty(\Gamma_h)} \lesssim h^p, \quad (3.13)$$

that  $n \cdot n_h > 0$  on  $\Gamma_h$  and that we have a uniform bound on the curvature  $\kappa_h$ .

**A2.** We assume that the mesh is sufficiently small so that there is a constant  $c$  independent of the mesh size  $h$  such that  $\mathcal{T}_h \subset U_\delta(\Gamma)$  with  $\delta = ch \leq \delta_0$  and that the closest point mapping  $p(x) : \Gamma_h \rightarrow \Gamma$  is a bijection.

**A3.** We assume that the data approximation  $f_h$  is such that  $\| |B| f^e - f_h \|_{L^2(\Gamma_h)} \lesssim h^{p+1} \|f\|_{L^2(\Gamma)}$ . Here  $ds = |B| ds_h$  (see Section 4 for details).

REMARK 3.2 There exist different strategies for how to approximate a geometry  $\Gamma$ . In CutFEM the approximation of the geometry and the approximation of the PDE are separated and any strategy that yields an accurate approximation of the geometry, i.e., so that assumption A1 is satisfied, can be used together with the proposed method. Thus, the presented method is not coupled to a specific interface

representation technique and implicit representation techniques as well as explicit representation techniques can be used. When  $\Gamma_h$  is explicitly given standard quadrature schemes can be applied, and it is therefore obvious how to compute the integrals in the bilinear forms in equation (3.3) with high accuracy. In the numerical examples in Sections 9.1 and 9.2 we use an explicit representation of  $\Gamma_h$ , see Section 9. However, when  $\Gamma_h$  is implicitly defined, for example via a level set function, a strategy for approximating integrals on  $\Gamma_h$  is necessary. Several such strategies exist, including explicit reconstruction of the geometry from the implicit representation (see, e.g., Johansson & Larson, 2013; Müller *et al.*, 2013; Saye, 2015; Fries & Omerović, 2016; Kublik & Tsai, 2016). The trace finite element method by Grande *et al.* (2018) uses a level set method and a different strategy, presented by Lehrenfeld (2016), where a transformation from a piecewise linear representation of the interface onto the zero level set of a high order approximation of the level set function is used. Thus, in the work by Grande *et al.* (2018), the analysis of the error in the energy norm takes this transformation into account.

## 4. Preliminary results

### 4.1 Norms

Given a positive semidefinite bilinear form  $b$  we let  $\|v\|_b^2 = b(v, v)$  denote the associated seminorm. In particular, we will use the following norms on  $V_h^p + H^{p+1}(\mathcal{T}_h)$ ,

$$\|v\|_{a_h}^2 = a_h(v, v), \quad \|v\|_{s_h}^2 = \|v\|_{s_h, F}^2 + \|v\|_{s_h, \Gamma}^2, \quad \|v\|_{A_h}^2 = \|v\|_{a_h}^2 + \|v\|_{s_h}^2. \quad (4.1)$$

We let  $H^s(\omega)$ , with  $\omega \subset \mathbb{R}^d$ , denote the standard Sobolev spaces of order  $s$ . For  $\omega \subset \Gamma$  or  $\omega \subset \Gamma_h$ ,  $H^s(\omega)$  denotes the surface Sobolev space, which is based on tangential derivatives.

### 4.2 Some inequalities

- Under A1 and A2 there is a constant such that the following Poincaré inequality holds for  $h \in (0, h_0]$ ,

$$\|v\|_{L^2(\Gamma_h)} \lesssim \|v\|_{a_h} \quad \forall v \in V_h^p \quad (4.2)$$

(see Burman *et al.*, 2015b, Lemma 4.1).

- For  $T \in \mathcal{T}_h$  we have the trace inequalities

$$\|v\|_{L^2(\partial T)}^2 \lesssim h^{-1} \|v\|_{L^2(T)}^2 + h \|\nabla v\|_{L^2(T)}^2 \quad v \in H^1(T), \quad (4.3)$$

$$\|v\|_{L^2(T \cap \Gamma_h)}^2 \lesssim h^{-1} \|v\|_{L^2(T)}^2 + h \|\nabla v\|_{L^2(T)}^2 \quad v \in H^1(T), \quad (4.4)$$

where the first inequality is a standard trace inequality (see, e.g., Brenner & Scott, 2008), and the second inequality is proven by Hansbo *et al.* (2003), and the constant is independent of how  $\Gamma_h$  intersects  $T$  and of  $h \in (0, h_0]$ .

- For  $T \in \mathcal{T}_h$  we have the inverse inequality

$$|v|_{H^j(T)}^2 \lesssim h^{-2(j-s)} \|v\|_{H^s(T)}^2 \quad 0 \leq s \leq j, \quad v \in V_h^p \quad (4.5)$$

(see Brenner & Scott, 2008).

### 4.3 Interpolation

Let  $\pi_h^p : L^2(\mathcal{T}_h) \rightarrow V_h^p$  be the Cl  ment interpolation operator. For all  $v \in H^{p+1}(\mathcal{T}_h)$  and  $T \in \mathcal{T}_h$  the following standard estimate holds

$$\|v - \pi_h^p v\|_{H^m(T)} \lesssim h^{s-m} \|v\|_{H^s(\mathcal{N}_h(T))} \quad m \leq s \leq p+1, \quad m = 0, \dots, p+1, \quad (4.6)$$

where  $\mathcal{N}_h(T) \subset \mathcal{T}_h$  is the union of the elements in  $\mathcal{T}_h$ , which share a node with  $T$ . In particular, we have the stability estimate

$$\|\pi_h^p v\|_{H^m(\mathcal{T}_h)} \lesssim \|v\|_{H^m(\mathcal{T}_h)}. \quad (4.7)$$

### 4.4 Some results for extended and lifted functions

**Extension.** Recall that we define the extension of a function  $u$  on  $\Gamma$  to  $U_{\delta_0}(\Gamma)$  by  $u^e(x) = u(p(x))$ . For  $u \in H^s(\Gamma)$ ,  $s \geq 0$  the following holds

$$\|u^e\|_{H^s(U_{\delta_0}(\Gamma))} \lesssim \delta^{1/2} \|u\|_{H^s(\Gamma)}, \quad (4.8)$$

where  $0 < \delta \leq \delta_0$  and for  $h$  small enough we may take  $\delta \sim h$  (See [Reusken, 2015](#), Lemma 3.1).

Using the definition of the extension of a function  $u$  on  $\Gamma$ , the definition of the closest point projection (2.2), the chain rule and the identity  $n = \nabla d$ , we obtain

$$\nabla u^e = (P_\Gamma - dH) \nabla u \quad \text{in } U_{\delta_0}(\Gamma), \quad (4.9)$$

where  $H(x) = D^2 d(x)$ . For  $x \in U_{\delta_0}(\Gamma)$ , the eigenvalues of  $H(x)$  are  $\kappa_1(x)$ ,  $\kappa_2(x)$  and 0, with corresponding orthonormal eigenvectors  $a_1(x)$ ,  $a_2(x)$  and  $n^e(x)$ . We have the identities

$$\kappa_i(x) = \frac{\kappa_i^e(x)}{1 + d(x)\kappa_i^e(x)}, \quad a_i(x) = a_i^e(x), \quad i = 1, 2 \quad (4.10)$$

and  $H(x)$  may be expressed in the form

$$H(x) = \sum_{i=1}^2 \frac{\kappa_i^e(x)}{1 + d(x)\kappa_i^e(x)} a_i^e(x) \otimes a_i^e(x) \quad (4.11)$$

(see [Gilbarg & Trudinger, 2001](#), Lemma 14.7).

Using the fact that  $H$  is tangential,  $P_\Gamma H = HP_\Gamma = H$ , in the identity (4.9) we obtain

$$\nabla u^e = (I - dH) \nabla_\Gamma u. \quad (4.12)$$

Thus,

$$\nabla_{\Gamma_h} u^e = B^T \nabla_\Gamma u \quad (4.13)$$

with

$$B = P_\Gamma (I - dH) P_{\Gamma_h}. \quad (4.14)$$

**Lifting.** Define the lifting  $v^l$  of a function  $v$  on  $\Gamma_h$  to  $\Gamma$  as

$$(v^l)^e = v^l \circ p = v \quad \text{on } \Gamma_h. \quad (4.15)$$

Using equation (4.13) it follows that

$$\nabla_{\Gamma_h} v = \nabla_{\Gamma_h} (v^l)^e = B^T \nabla_{\Gamma} (v^l) \quad (4.16)$$

and thus we have

$$\nabla_{\Gamma} (v^l) = B^{-T} \nabla_{\Gamma_h} v. \quad (4.17)$$

Here

$$B^{-1} = \left( I - \frac{n \otimes n_h}{n \cdot n_h} \right) (I - dH)^{-1} \quad (4.18)$$

(see Demlow & Dziuk, 2007), and using the fact that  $n$  and  $n_h$  are unit normal vectors it is easy to show that

$$BB^{-1} = P_{\Gamma}, \quad B^{-1}B = P_{\Gamma_h}. \quad (4.19)$$

**Estimates related to  $B$ .** It can be shown (see, e.g., Burman *et al.*, 2015b) that the following estimates hold

$$\|B\|_{L^\infty(\Gamma_h)} \lesssim 1, \quad \|B^{-1}\|_{L^\infty(\Gamma)} \lesssim 1, \quad \|BB^T - I\|_{L^\infty(\Gamma_h)} \lesssim h^{p+1}. \quad (4.20)$$

**Surface measures.** For the surface measure on  $\Gamma_h$ ,  $ds_h(x)$ , and the surface measure on  $\Gamma$ ,  $ds(p(x))$ , we have the following identity

$$ds(p(x)) = |B(x)| ds_h(x), \quad x \in \Gamma_h, \quad (4.21)$$

where

$$|B(x)| = n \cdot n_h (1 - d(x)\kappa_1(x))(1 - d(x)\kappa_2(x)) \quad (4.22)$$

(see Demlow & Dziuk, 2007, Proposition A.1). Using the identity  $\|n(x) - n_h(x)\|_{\mathbb{R}^3}^2 = 2(1 - n(x) \cdot n_h(x))$  and our assumptions, we obtain the following estimates:

$$\|1 - |B|\|_{L^\infty(\Gamma_h)} \lesssim h^{p+1}, \quad \|B\|_{L^\infty(\Gamma_h)} \lesssim 1, \quad \||B|^{-1}\|_{L^\infty(\Gamma_h)} \lesssim 1. \quad (4.23)$$

**Norm equivalences.** Using the identities (4.13) and (4.17) and the bounds in equations (4.20) and (4.23), we obtain the following equivalences:

$$\|v^l\|_{L^2(\Gamma)} \sim \|v\|_{L^2(\Gamma_h)}, \quad \|u\|_{L^2(\Gamma)} \sim \|u^e\|_{L^2(\Gamma_h)}, \quad (4.24)$$

and

$$\|\nabla_{\Gamma} v^l\|_{L^2(\Gamma)} \sim \|\nabla_{\Gamma_h} v\|_{L^2(\Gamma_h)}, \quad \|\nabla_{\Gamma} u\|_{L^2(\Gamma)} \sim \|\nabla_{\Gamma_h} u^e\|_{L^2(\Gamma_h)}. \quad (4.25)$$

We will also use the following bound:

$$\|D^j u^e\|_{L^2(\Gamma_h)} \lesssim \|u\|_{H^m(\Gamma)} \quad u \in H^m(\Gamma), \quad j \leq m. \quad (4.26)$$

The constant in (4.26) depends on the higher-order derivatives of the geometry.

### 5. Properties of the stabilization term

Here we formulate properties of a general stabilization term that are sufficient to prove that the resulting linear system of equations have an optimal scaling of the condition number with the mesh parameter, and that the convergence of the proposed method is of optimal order.

**Properties.** Let  $s_h$  be a semipositive definite bilinear form such that

**P1.** For  $v \in H^{p+1}(\Gamma) \cap H_0^1(\Gamma)$ ,

$$\|v^e - \pi_h^p v^e\|_{s_h} \lesssim h^p \|v\|_{H^{p+1}(\Gamma)}. \quad (5.1)$$

**P2.** For  $v \in H^{p+1}(\Gamma) \cap H_0^1(\Gamma)$ ,

$$\|v^e\|_{s_h} \lesssim h^p \|v\|_{H^{p+1}(\Gamma)}. \quad (5.2)$$

**P3.** For  $v \in H^2(\Gamma)$ ,

$$\|\pi_h^p v^e\|_{s_h} \lesssim h \|v\|_{H^2(\Gamma)}. \quad (5.3)$$

**P4.** For  $v \in V_h^p$ ,

$$\|v\|_{s_h}^2 \lesssim h^{-3} \|v\|_{L^2(\mathcal{T}_h)}^2. \quad (5.4)$$

**P5.** For  $v \in V_h^p$ ,

$$\|v\|_{L^2(\mathcal{T}_h)}^2 \lesssim h(\|v\|_{a_h}^2 + \|v\|_{s_h}^2). \quad (5.5)$$

**REMARK 5.1** The first three properties are used to establish the *a priori* error estimates: P1 is needed to show optimal interpolation error estimates, P2 is used to estimate the consistency error and P3 is used to show optimal  $L^2$ -error estimates. In P3 the assumption that  $v \in H^2(\Gamma)$  reflects the fact that the solution to the dual problem resides in  $H^2(\Gamma)$ . The two last properties are used to show the condition number estimate: P4 is the inverse inequality and P5 is the Poincaré inequality. For the extension of these properties to the more general case of an  $n$ -dimensional smooth manifold embedded in  $\mathbb{R}^d$ , with codimension  $cd = d - n$  see Section 7.

**Verification of P1–P5.** We now show that the stabilization term defined in (3.7) satisfies P1–P5.

**P1.** Using a standard trace inequality on the faces (4.3) and the trace inequality (4.4) for the contributions on  $\Gamma_h$ , we obtain

$$\|w\|_{s_h}^2 = \sum_{j=1}^p \sum_{F \in \mathcal{F}_h} c_{F,j} h^{2(j-1+\gamma)} \| [D_{n_F}^j w] \|_{L^2(F)}^2 \quad (5.6)$$

$$+ \sum_{j=1}^p \sum_{K \in \mathcal{K}_h} c_{K,j} h^{2(j-1+\gamma)} \| D_{n_K}^j w \|_{L^2(K)}^2$$

$$\lesssim \sum_{j=1}^p \sum_{T \in \mathcal{T}_h} h^{2(j-1+\gamma)} \left( h^{-1} \|D^j w\|_{L^2(T)}^2 + h \|D^{j+1} w\|_{L^2(T)}^2 \right) \quad (5.7)$$

$$\lesssim \sum_{j=0}^p \sum_{T \in \mathcal{T}_h} h^{2(j+\gamma)-1} \|D^{j+1}w\|_{L^2(T)}^2 \quad (5.8)$$

$$= \sum_{j=0}^p h^{2(j+\gamma)-1} \|D^{j+1}w\|_{L^2(\mathcal{T}_h)}^2. \quad (5.9)$$

Next, setting  $w = v^e - \pi_h^p v^e$ , using the interpolation error estimate (4.6) and the stability (4.8) of the extension operator with  $\delta \sim h$ , we obtain

$$\|v^e - \pi_h^p v^e\|_{s_h}^2 \lesssim \sum_{j=0}^p h^{2(j+\gamma)-1} \|D^{j+1}(v^e - \pi_h^p v^e)\|_{L^2(\mathcal{T}_h)}^2 \quad (5.10)$$

$$\lesssim \sum_{j=0}^p h^{2(j+\gamma)-1} h^{2(p-j)} \|D^{j+1}v^e\|_{H^{p+1}(\mathcal{T}_h)}^2 \quad (5.11)$$

$$\lesssim h^{2(p+\gamma)} \|v\|_{H^{p+1}(\Gamma)}^2. \quad (5.12)$$

Finally,  $h^{2(p+\gamma)} \lesssim h^{2p}$  for  $\gamma \in [0, 1]$  since  $h \in (0, h_0]$ , and thus P1 follows.

**P2.** Note first that for  $v \in H^{p+1}(\Gamma)$  we have  $\|v^e\|_{s_{h,F}} = 0$  and thus  $\|v^e\|_{s_h} = \|v^e\|_{s_{h,\Gamma}}$ . Next, subtracting  $D_n^j v^e = 0$ , using Assumption A1, and inequality (4.26), we obtain

$$\|v^e\|_{s_h}^2 = \sum_{j=1}^p c_{\Gamma,j} h^{2(j-1+\gamma)} \|D_{n_h}^j v^e\|_{\mathcal{K}_h}^2 \quad (5.13)$$

$$\lesssim \sum_{j=1}^p h^{2(j-1+\gamma)} \|(D_{n_h}^j - D_n^j)v^e\|_{L^2(\mathcal{K}_h)}^2 \quad (5.14)$$

$$\lesssim \sum_{j=1}^p h^{2(j-1+\gamma)} \|n_h - n\|_{L^\infty(\mathcal{K}_h)}^2 \|D^j v^e\|_{L^2(\mathcal{K}_h)}^2 \quad (5.15)$$

$$\lesssim h^{2(p+\gamma)} \|v\|_{H^p(\Gamma)}^2, \quad (5.16)$$

where we used the estimate  $\|n_h - n\|_{L^\infty(\mathcal{K}_h)} \lesssim h^p$ . We conclude that P2 holds for  $\gamma \in [0, 1]$  since  $h \in (0, h_0]$ .

**P3.** For  $w \in V_h^p$  we have the estimate

$$\|w\|_{s_h}^2 \lesssim h^{2\gamma} \|[D_{n_F}^1 w]\|_{L^2(\mathcal{F}_h)}^2 + h^{2\gamma} \|D_{n_h}^1 w\|_{L^2(\mathcal{K}_h)}^2 + h^{2(1+\gamma)-1} \|D^2 w\|_{L^2(\mathcal{T}_h)}^2. \quad (5.17)$$

**Verification of (5.17).** We start from the definition

$$\|w\|_{s_h}^2 = \|w\|_{s_{h,F}}^2 + \|w\|_{s_{h,\Gamma}}^2 \quad (5.18)$$

and estimate the two terms on the right-hand side as follows: first

$$\|w\|_{s_h, F}^2 \lesssim h^{2\gamma} \| [D_{n_F}^1 w] \|_{L^2(\mathcal{F}_h)}^2 + \sum_{j=2}^p h^{2(j-1+\gamma)} \| [D_{n_F}^j w] \|_{L^2(\mathcal{F}_h)}^2 \quad (5.19)$$

$$\lesssim h^{2\gamma} \| [D_{n_F}^1 w] \|_{L^2(\mathcal{F}_h)}^2 + h^{2(1+\gamma)-1} \| [D^2 w] \|_{L^2(\mathcal{F}_h)}^2, \quad (5.20)$$

where for each  $j = 2, \dots, p$  and each face  $F \in \mathcal{F}_h$  with neighbouring elements  $T_1$  and  $T_2$  we used the inverse estimate

$$\| [D_{n_F}^j w] \|_{L^2(F)}^2 \lesssim h^{-1} \| D^j w \|_{L^2(T_1 \cup T_2)}^2 \lesssim h^{-1} h^{2(2-j)} \| D^2 w \|_{L^2(T_1 \cup T_2)}^2. \quad (5.21)$$

Secondly, using a similar approach

$$\|w\|_{s_h, \Gamma}^2 \lesssim h^{2\gamma} \| D_{n_h}^1 w \|_{L^2(\mathcal{K}_h)}^2 + \sum_{j=2}^p h^{2(j-1+\gamma)} \| D_{n_h}^j w \|_{L^2(\mathcal{K}_h)}^2 \quad (5.22)$$

$$\lesssim h^{2\gamma} \| D_{n_h}^1 w \|_{L^2(\mathcal{K}_h)}^2 + \sum_{j=2}^p \underbrace{h^{2(j-1+\gamma)} h^{-1} h^{2(2-j)}}_{=h^{2(1+\gamma)-1}} \| D^2 w \|_{L^2(\mathcal{F}_h)}^2 \quad (5.23)$$

$$\lesssim h^{2\gamma} \| D_{n_h}^1 w \|_{L^2(\mathcal{K}_h)}^2 + h^{2(1+\gamma)-1} \| D^2 w \|_{L^2(\mathcal{F}_h)}^2, \quad (5.24)$$

where we used the inverse estimate

$$\| D_{n_h}^j w \|_{L^2(K)}^2 \leq \| D^j w \|_{L^2(K)}^2 \lesssim h^{-1} \| D^j w \|_{L^2(T)}^2 \lesssim h^{-1} h^{2(2-j)} \| D^2 w \|_{L^2(T)}^2 \quad (5.25)$$

for  $j = 2, \dots, p$ . Thus, (5.17) holds.

Setting  $w = \pi_h^p v^e$  in (5.17) we obtain

$$\| \pi_h^p v^e \|_{s_h}^2 \lesssim \underbrace{h^{2\gamma} \| [D_{n_F}^1 \pi_h^p v^e] \|_{L^2(\mathcal{F}_h)}^2}_I + \underbrace{h^{2\gamma} \| D_{n_h}^1 \pi_h^p v^e \|_{L^2(\mathcal{K}_h)}^2}_{II} \quad (5.26)$$

$$+ h^{2(1+\gamma)-1} \| D^2 \pi_h^p v^e \|_{L^2(\mathcal{F}_h)}^2 \\ = I + II + h^{2(1+\gamma)} \| v \|_{H^2(\Gamma)}^2, \quad (5.27)$$

where we used the stability of the interpolation operator (4.7) and the stability of the extension operator (4.8). Next we have the following bounds.

**Term I.** Using the fact that  $[D_{n_F}^1 v^e] = 0$  for  $v \in H^2(\Gamma)$ , we obtain

$$\|[D_{n_F}^1 \pi_h^p v^e]\|_{L^2(\mathcal{T}_h)}^2 = \|[D_{n_F}^1 (\pi_h^p v^e - v^e)]\|_{L^2(\mathcal{T}_h)}^2 \quad (5.28)$$

$$\lesssim h^{-1} \|[D(\pi_h^p v^e - v^e)]\|_{L^2(\mathcal{T}_h)}^2 + h \|[D^2(\pi_h^p v^e - v^e)]\|_{L^2(\mathcal{T}_h)}^2 \quad (5.29)$$

$$\lesssim h \|D^2 v^e\|_{L^2(\mathcal{T}_h)}^2 \quad (5.30)$$

$$\lesssim h^2 \|v\|_{H^2(\Gamma)}^2, \quad (5.31)$$

where we used the trace inequality (4.3), the interpolation estimate (4.6) and finally the stability (4.8) of the extension operator. Thus, we have

$$I \lesssim h^{2(1+\gamma)} \|v\|_{H^2(\Gamma)}^2. \quad (5.32)$$

**Term II.** Using the fact that  $D_{n^e}^1 v^e = 0$ , we obtain

$$\|D_{n_h}^1 \pi_h^p v^e\|_{L^2(\mathcal{K}_h)}^2 \lesssim \|D_{n^e}^1 \pi_h^p v^e\|_{L^2(\mathcal{K}_h)}^2 + \|n^e - n_h\|_{L^\infty(\mathcal{K}_h)}^2 \|D \pi_h^p v^e\|_{L^2(\mathcal{K}_h)}^2 \quad (5.33)$$

$$\lesssim \|D_{n^e}^1 (\pi_h^p v^e - v^e)\|_{L^2(\mathcal{K}_h)}^2 + h^{2p} \|D \pi_h^p v^e\|_{L^2(\mathcal{T}_h)}^2 \quad (5.34)$$

$$\lesssim h^2 \|v\|_{H^2(\Gamma)}^2 + h^{2p} \|v\|_{H^1(\Gamma)}^2 \quad (5.35)$$

and we arrive at

$$II \lesssim h^{2(1+\gamma)} \|v\|_{H^2(\Gamma)}^2 \quad (5.36)$$

since  $p \geq 1$  and  $h \in (0, h_0]$ .

Collecting the bounds we obtain

$$\|\pi_h^p v^e\|_{s_h} \lesssim h^{1+\gamma} \|v\|_{H^2(\Gamma)} \quad (5.37)$$

and thus P3 holds.

**Simplified proof of P3 in the case  $\gamma=1$ .** Using estimate (5.9) with  $w \in V_h^p$  and inverse inequality (4.5), we obtain

$$\|w\|_{s_h}^2 \lesssim \sum_{j=0}^p h^{2j+1} \|D^{j+1} w\|_{L^2(\mathcal{T}_h)}^2 \quad (5.38)$$

$$\lesssim h \|Dw\|_{L^2(\mathcal{T}_h)}^2 + \underbrace{\sum_{j=1}^p h^{2j+1} \|D^{j+1} w\|_{L^2(\mathcal{T}_h)}^2}_{\lesssim h^3 \|D^2 w\|_{L^2(\mathcal{T}_h)}^2} \quad (5.39)$$

$$\lesssim h \|Dw\|_{L^2(\mathcal{T}_h)}^2 + h^3 \|D^2 w\|_{L^2(\mathcal{T}_h)}^2. \quad (5.40)$$

Setting  $w = \pi_h^p v^e$  and using the stability of the interpolation operator (4.7) and the stability of the extension operator (4.8), we obtain

$$\|\pi_h^p v^e\|_{s_h}^2 \lesssim h \|D\pi_h^p v^e\|_{L^2(\mathcal{T}_h)}^2 + h^3 \|D^2\pi_h^p v^e\|_{L^2(\mathcal{T}_h)}^2 \quad (5.41)$$

$$\lesssim h^2 \|v\|_{H^1(\Gamma)}^2 + h^4 \|v\|_{H^2(\Gamma)}^2. \quad (5.42)$$

**P4.** Starting from (5.9) with  $w \in V_h^p$  and using the inverse inequality, we get

$$\|w\|_{s_h}^2 \lesssim \sum_{j=0}^p h^{2j+2\gamma-1} \|D^{j+1}w\|_{L^2(\mathcal{T}_h)}^2 \lesssim h^{2\gamma-3} \|w\|_{L^2(\mathcal{T}_h)}^2. \quad (5.43)$$

Using that  $h \in (0, h_0]$  we have  $h^{2\gamma-3} \lesssim h^{-3}$  for  $\gamma \geq 0$  and thus Property P4 holds.

**P5.** See the proof of Lemma 6.5 below.

**REMARK 5.2** We note that the normal gradient stabilization (1.10) satisfies  $P_1, P_2, P_4$  and  $P_5$  (see Burman *et al.*, 2018; Grande *et al.*, 2018). To verify  $P_3$  (for optimal error estimates in  $L^2$ ) we use that  $n^e \cdot \nabla v^e = 0$ , interpolation error estimates, the  $H^1$ -stability of the interpolant, assumption A2 (the inclusion  $\mathcal{T}_h \subset U_\delta(\Gamma)$  with  $\delta \sim h$ ), the stability of the extension operator (4.8) and that  $\|n_h - n^e\|_{L^\infty(\mathcal{T}_h)}^2 \lesssim h^{2p}$  to conclude that

$$\|\pi_h^p v^e\|_{s_h}^2 \lesssim h^\alpha \left( \|n^e \cdot \nabla \pi_h^p v^e\|_{L^2(\mathcal{T}_h)}^2 + \|n_h - n^e\|_{L^\infty(\mathcal{T}_h)}^2 \|\nabla \pi_h^p v^e\|_{L^2(\mathcal{T}_h)}^2 \right) \quad (5.44)$$

$$\lesssim h^\alpha \left( \|n^e \cdot \nabla (\pi_h^p v^e - v^e)\|_{L^2(\mathcal{T}_h)}^2 + h^{2p} \delta \|v\|_{H^1(\Gamma)}^2 \right) \lesssim h^\alpha h^2 \delta \|v\|_{H^2(\Gamma)}^2 \quad (5.45)$$

$$\lesssim h^{3+\alpha} \|v\|_{H^2(\Gamma)}^2. \quad (5.46)$$

We also used that  $p \geq 1$ . Thus, we conclude that  $P_3$  holds for  $\alpha \geq -1$ .

## 6. Condition number estimate

We shall show that the spectral condition number of the resulting stiffness matrix scales as  $h^{-2}$ , independent of the position of the geometry relative to the background mesh. In particular, we show that the stabilization term defined in (3.7) has Property P5, and controls the condition number for linear as well as for higher-order elements. We will see that due to the condition  $(v, 1)_{\Gamma_h} = 0$  we are naturally led to studying the spectral condition number associated with the generalized eigenvalue problem.

Let  $\{\varphi_i\}_{i=1}^N$  be the standard basis in  $\tilde{V}_h^p|_{\mathcal{T}_h}$  and for  $v \in \tilde{V}_h^p|_{\mathcal{T}_h}$  let  $\mathbf{v} = [v_1, \dots, v_N]^T \in \mathbb{R}^N$  be its representation with respect to this basis, i.e.,

$$v = \sum_{i=1}^N v_i \varphi_i. \quad (6.1)$$

Recalling that  $V_h^p = \{v \in \tilde{V}_h^p|_{\mathcal{T}_h} : (v, 1)_{\Gamma_h} = 0\} \subset \tilde{V}_h^p|_{\mathcal{T}_h}$  we define the corresponding subspace  $\widehat{\mathbb{R}}^N$  of  $\mathbb{R}^N$ ,

$$\widehat{\mathbb{R}}^N = \{\mathbf{v} \in \mathbb{R}^N : \langle \mathbf{v}, \mathbf{b} \rangle = 0\}, \quad (6.2)$$

where  $\mathbf{b} = [b_1, \dots, b_N]^T$  with  $b_i = \int_{\Gamma_h} \varphi_i$ , and  $\langle \cdot, \cdot \rangle$  is the standard inner product in  $\mathbb{R}^N$  with corresponding norm  $\|\mathbf{x}\|_{\mathbb{R}^N} = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$ .

The stiffness matrix  $\mathbf{A} \in \mathbb{R}^{N \times N}$  associated with the form  $A_h$  on  $\tilde{V}_h^p|_{\mathcal{T}_h}$  is defined by

$$\langle \mathbf{A}\mathbf{w}, \mathbf{v} \rangle = A_h(w, v) \quad \forall v, w \in \tilde{V}_h^p|_{\mathcal{T}_h}. \quad (6.3)$$

The matrix is symmetric positive semidefinite and the kernel of  $\mathbf{A}$  is one-dimensional, and given by

$$\ker(\mathbf{A}) = \text{span}(\boldsymbol{\xi}_0) \subset \mathbb{R}^N, \quad \boldsymbol{\xi}_0 = [1, \dots, 1]^T \in \mathbb{R}^N, \quad (6.4)$$

which corresponds to the subspace of constant functions  $\tilde{V}_h^0|_{\mathcal{T}_h}$  in  $\tilde{V}_h^p|_{\mathcal{T}_h}$ .

Next defining the mass matrix  $\mathbf{M}$ ,

$$\langle \mathbf{M}\mathbf{w}, \mathbf{v} \rangle = M_h(w, v) \quad \forall v, w \in \tilde{V}_h^p|_{\mathcal{T}_h}, \quad (6.5)$$

where

$$M_h(v, w) = (v, w)_{\Gamma_h} + s_{h,M}(v, w) \quad (6.6)$$

and  $s_{h,M}$  is a stabilization form, we note that  $\mathbf{b} = \mathbf{M}\boldsymbol{\xi}_0$ , and thus the orthogonality condition  $\langle \mathbf{v}, \mathbf{b} \rangle = 0$  is equivalent to  $\langle \mathbf{v}, \mathbf{M}\boldsymbol{\xi}_0 \rangle = 0$ . Vectors  $\mathbf{v} \in \widehat{\mathbb{R}}^N$  are thus  $\mathbf{M}$  orthogonal to  $\boldsymbol{\xi}_0$  and we therefore consider the generalized eigenvalue problem

$$\mathbf{A}\mathbf{w} = \lambda \mathbf{M}\mathbf{w}, \quad \lambda \in \mathbb{R}, \mathbf{w} \in \widehat{\mathbb{R}}^N, \quad (6.7)$$

in  $\widehat{\mathbb{R}}^N$  with eigenvalues  $0 < \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_{N-1}$ . We define the spectral condition number by

$$\kappa(\mathbf{A}) = \frac{\lambda_{\max}}{\lambda_{\min}}, \quad (6.8)$$

with the maximal and minimal eigenvalues of the generalized eigenvalue problem (6.7), i.e.,  $\lambda_{\min} = \lambda_1$  and  $\lambda_{\max} = \lambda_{N-1}$ . The eigenvalues  $\lambda_{\max}$  and  $\lambda_{\min}$  are given by

$$\lambda_{\max} = \max_{\mathbf{v} \in \widehat{\mathbb{R}}^N \setminus \{0\}} \frac{\langle \mathbf{A}\mathbf{v}, \mathbf{v} \rangle}{\langle \mathbf{M}\mathbf{v}, \mathbf{v} \rangle}, \quad \lambda_{\min} = \min_{\mathbf{v} \in \widehat{\mathbb{R}}^N \setminus \{0\}} \frac{\langle \mathbf{A}\mathbf{v}, \mathbf{v} \rangle}{\langle \mathbf{M}\mathbf{v}, \mathbf{v} \rangle}. \quad (6.9)$$

Let  $\|v\|_{M_h}^2 = \langle \mathbf{M}\mathbf{v}, \mathbf{v} \rangle = M_h(v, v)$  and  $\|v\|_{s_{h,M}}^2 = s_{h,M}(v, v)$ . The stabilization  $s_{h,M}$  in (6.6) should have the following properties:

**M1.** For  $v \in \tilde{V}_h^p$ ,

$$\|v\|_{s_{h,M}}^2 \lesssim h^{-1} \|v\|_{L^2(\mathcal{T}_h)}^2. \quad (6.10)$$

**M2.** For  $v \in \tilde{V}_h^p$ ,

$$\|v\|_{L^2(\mathcal{T}_h)}^2 \lesssim h(\|v\|_{L^2(\Gamma_h)}^2 + \|v\|_{s_{h,M}}^2) = h\|v\|_{M_h}^2. \quad (6.11)$$

These properties correspond to properties P4 and P5, but are for the stabilization of the mass matrix and result in the following equivalence: for  $v \in \tilde{V}_h^p$

$$\|v\|_{L^2(\mathcal{T}_h)}^2 \sim h\|v\|_{M_h}^2. \quad (6.12)$$

Since  $\mathcal{T}_h$  is quasi-uniform we also have the equivalence

$$\|v\|_{L^2(\mathcal{T}_h)} \sim h^{d/2} \|\mathbf{v}\|_{\mathbb{R}^N}, \quad (6.13)$$

where  $d$  is the dimension of the embedding space  $\mathbb{R}^d$ . Hence,

$$h^d \|\mathbf{v}\|_{\mathbb{R}^N}^2 \sim \|v\|_{L^2(\mathcal{T}_h)}^2 \sim h\|v\|_{M_h}^2. \quad (6.14)$$

The proposed stabilization with  $\gamma = 1$  has the desired properties M1 and M2. Hence, we choose  $s_{h,M}$  as  $s_h$  in (3.7) with  $\gamma = 1$ . To verify that we have the first property note that from equation (5.43) we have that

$$\|v\|_{s_h}^2 \lesssim h^{2\gamma-3} \|v\|_{L^2(\mathcal{T}_h)}^2. \quad (6.15)$$

Thus, for  $\gamma = 1$  the proposed stabilization satisfies M1. We prove the second property in Lemma 6.5.

We now state the main theorem of this section.

**THEOREM 6.1** If there are constants, independent of the mesh size  $h$  and of how the surface cuts the background mesh, such that the following hold:

- $A_h$  is continuous:  $A_h(w, v) \lesssim \|w\|_{A_h} \|v\|_{A_h} \quad \forall v, w \in V_h^p + H^{p+1}(\mathcal{T}_h)$ ,
- $A_h$  is coercive:  $\|v\|_{A_h}^2 \lesssim A_h(v, v) \quad \forall v \in V_h^p$ ,
- The inverse inequality:  $\|v\|_{A_h}^2 \lesssim h^{-3} \|v\|_{L^2(\mathcal{T}_h)}^2 \quad v \in V_h^p$ ,
- The Poincaré inequality:  $\|v\|_{L^2(\mathcal{T}_h)}^2 \lesssim h\|v\|_{A_h}^2 \quad v \in V_h^p$ ,
- Equivalence between the norms:  $\|v\|_{L^2(\mathcal{T}_h)}^2 \sim h\|v\|_{M_h}^2 = \langle \mathbf{M}\mathbf{v}, \mathbf{v} \rangle \quad v \in V_h^p, \mathbf{v} \in \widehat{\mathbb{R}}^N$ ,

then the spectral condition number  $\kappa(\mathbf{A})$ , defined in (6.8), satisfies

$$\kappa(\mathbf{A}) \lesssim h^{-2}. \quad (6.16)$$

*Proof.* For  $v \in V_h^p$  and  $\mathbf{v} \in \widehat{\mathbb{R}}^N$  it follows from the continuity of the form  $A_h$ , the inverse inequality and equivalence between the norms that

$$\langle \mathbf{A}\mathbf{v}, \mathbf{v} \rangle = A_h(v, v) \lesssim \|v\|_{A_h}^2 \lesssim h^{-3} \|v\|_{L^2(\mathcal{T}_h)}^2 \lesssim h^{-2} \langle \mathbf{M}\mathbf{v}, \mathbf{v} \rangle, \quad (6.17)$$

and thus  $\lambda_{\max} \lesssim h^{-2}$ . From coercivity, the Poincaré inequality and equivalence between the norms, we obtain

$$\langle \mathbf{A}\mathbf{v}, \mathbf{v} \rangle = A_h(v, v) \gtrsim \|v\|_{A_h}^2 \gtrsim h^{-1} \|v\|_{L^2(\mathcal{T}_h)}^2 \gtrsim \langle \mathbf{M}\mathbf{v}, \mathbf{v} \rangle, \quad (6.18)$$

and thus  $\lambda_{\min} \gtrsim 1$ . This concludes the proof since from the definition of the spectral condition number in (6.8) and (6.9), we have

$$\kappa(\mathbf{A}) = \frac{\lambda_{\max}}{\lambda_{\min}} \lesssim h^{-2}. \quad (6.19)$$

□

REMARK 6.2 We may use a Lagrange multiplier approach to enforce the condition  $\int_{\Gamma_h} v \, ds_h = 0$ . This leads to a linear system on the form

$$\underbrace{\begin{bmatrix} \mathbf{A} & \mathbf{b} \\ \mathbf{b}^T & 0 \end{bmatrix}}_{\mathbf{B}} \begin{bmatrix} \mathbf{u} \\ \eta \end{bmatrix} = \begin{bmatrix} \mathbf{f} \\ 0 \end{bmatrix}, \quad (6.20)$$

where  $\eta \in \mathbb{R}$  is the Lagrange multiplier. The spectrum of the generalized eigenvalue problem

$$\begin{bmatrix} \mathbf{A} & \mathbf{b} \\ \mathbf{b}^T & 0 \end{bmatrix} \begin{bmatrix} \mathbf{w} \\ w \end{bmatrix} = \lambda \begin{bmatrix} \mathbf{M} & 0 \\ 0 & m \end{bmatrix} \begin{bmatrix} \mathbf{w} \\ w \end{bmatrix}, \quad (6.21)$$

where  $m \in \mathbb{R}_+$  is a positive parameter,  $\lambda \in \mathbb{R}$  and  $[\mathbf{w}, w]^T \in \mathbb{R}^{N+1}$ , consists of the eigenvalues  $0 < \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_{N-1}$  of the generalized eigenvalue problem (6.7) together with  $\pm \lambda_0/m^{1/2}$ , where  $\lambda_0 = \sqrt{\langle \xi_0, \mathbf{b} \rangle} = \sqrt{\langle \xi_0, \mathbf{M}\xi_0 \rangle}$ . To verify this fact let  $\xi_i \in \widehat{\mathbb{R}}^N$  and  $\lambda_i \in \mathbb{R}_+$  satisfy  $\mathbf{A}\xi_i = \lambda_i \mathbf{M}\xi_i$ ,  $i = 1, \dots, N-1$ , then we have

$$\begin{bmatrix} \mathbf{A} & \mathbf{b} \\ \mathbf{b}^T & 0 \end{bmatrix} \begin{bmatrix} \xi_i \\ 0 \end{bmatrix} = \begin{bmatrix} \mathbf{A}\xi_i \\ \mathbf{b}^T \xi_i \end{bmatrix} = \begin{bmatrix} \lambda_i \mathbf{M}\xi_i \\ 0 \end{bmatrix} = \lambda_i \begin{bmatrix} \mathbf{M} & 0 \\ 0 & m \end{bmatrix} \begin{bmatrix} \xi_i \\ 0 \end{bmatrix}, \quad (6.22)$$

where we used the fact that  $\mathbf{b}^T \xi_i = \xi_0^T \mathbf{M}\xi_i = 0$ . Thus, we conclude that  $\lambda_i$  is an eigenvalue of (6.21). To show that  $\pm \lambda_0/\sqrt{m}$  are eigenvalues of (6.21) we proceed as follows:

$$\begin{bmatrix} \mathbf{A} & \mathbf{b} \\ \mathbf{b}^T & 0 \end{bmatrix} \begin{bmatrix} \sqrt{m}\xi_0 \\ \pm \lambda_0 \end{bmatrix} = \begin{bmatrix} \sqrt{m}\mathbf{A}\xi_0 \pm \lambda_0 \mathbf{b} \\ \sqrt{m}\mathbf{b}^T \xi_0 \end{bmatrix} = \begin{bmatrix} \lambda_0 \mathbf{M}\xi_0 \\ \sqrt{m}\lambda_0^2 \end{bmatrix} = \pm \frac{\lambda_0}{\sqrt{m}} \begin{bmatrix} \mathbf{M} & 0 \\ 0 & m \end{bmatrix} \begin{bmatrix} \sqrt{m}\xi_0 \\ \pm \lambda_0 \end{bmatrix}, \quad (6.23)$$

where we used the identity  $\mathbf{b}^T \xi_0 = \xi_0^T \mathbf{M}\xi_0 = \lambda_0^2$ . The spectral condition number of  $\mathbf{B}$  is then

$$\kappa(\mathbf{B}) = \frac{\max\{\lambda_0/\sqrt{m}, \lambda_{N-1}\}}{\min\{\lambda_0/\sqrt{m}, \lambda_1\}}. \quad (6.24)$$

If  $\lambda_1 \leq \lambda_0/\sqrt{m} \leq \lambda_{N-1}$  we have

$$\kappa(\mathbf{B}) = \kappa(\mathbf{A}). \quad (6.25)$$

This is often the case and, in particular, in all the numerical examples considered in this paper. If this is not the case one may scale  $\mathbf{b}$  so that  $\lambda_1 \leq \lambda_0 \leq \lambda_{N-1}$ . For instance, taking  $m \sim h^s$  leads to  $\lambda_0/\sqrt{m} \sim h^{-s/2}$ , since  $\lambda_0 = \langle \xi_0, \mathbf{M}\xi_0 \rangle = \int_{\Gamma_h} 1 \approx \int_{\Gamma} 1$  is approximately constant. Thus, for  $0 < s < 2$  we have  $\lambda_1 \leq \lambda_0/\sqrt{m} \leq \lambda_{N-1}$  for sufficiently small  $h$ , since  $\lambda_1 \sim 1$  and  $\lambda_{N-1} \sim h^{-2}$ . We finally note that the maximal and minimal eigenvalues to

$$\begin{bmatrix} \mathbf{A} & \mathbf{b} \\ \mathbf{b}^T & 0 \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{w}} \\ \tilde{w} \end{bmatrix} = \tilde{\lambda} \begin{bmatrix} \mathbf{I} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{w}} \\ \tilde{w} \end{bmatrix} \quad (6.26)$$

is related to the corresponding eigenvalues of (6.21) as follows:

$$\frac{1}{\max(h^{d-1}, m)} \tilde{\lambda}_{\max} \lesssim \lambda_{\max} \lesssim \frac{1}{\min(h^{d-1}, m)} \tilde{\lambda}_{\max} \quad (6.27)$$

and

$$\frac{1}{\max(h^{d-1}, m)} \tilde{\lambda}_{\min} \lesssim \lambda_{\min} \lesssim \frac{1}{\min(h^{d-1}, m)} \tilde{\lambda}_{\min} \quad (6.28)$$

since

$$\min(h^{d-1}, m)(\langle \mathbf{w}, \mathbf{w} \rangle + w^2) \lesssim \langle \mathbf{w}, \mathbf{M}\mathbf{w} \rangle + mw^2 \lesssim \max(h^{d-1}, m)(\langle \mathbf{w}, \mathbf{w} \rangle + w^2). \quad (6.29)$$

In particular, taking  $m \sim h^{d-1}$  we obtain

$$\lambda_{\max} \sim h^{1-d} \tilde{\lambda}_{\max}, \quad \lambda_{\min} \sim h^{1-d} \tilde{\lambda}_{\min} \quad (6.30)$$

and thus

$$\tilde{\kappa}(\mathbf{B}) = \frac{\tilde{\lambda}_{\max}}{\tilde{\lambda}_{\min}} \sim \frac{\lambda_{\max}}{\lambda_{\min}} = \kappa(\mathbf{B}). \quad (6.31)$$

### 6.1 Proofs of the conditions in the main theorem

In the following lemmas we show that all the conditions in Theorem 6.1 are satisfied.

**LEMMA 6.3** (Continuity and inf-sup condition). There is a constant independent of the mesh size  $h$  and of how the surface cuts the background mesh, such that  $A_h$  is continuous:

$$A_h(w, v) \leq \|w\|_{A_h} \|v\|_{A_h}, \quad \forall v, w \in V_h^p + H^{p+1}(\mathcal{T}_h) \quad (6.32)$$

and  $A_h$  satisfies the inf-sup condition

$$\|w\|_{A_h} \lesssim \sup_{v \in V_h^p \setminus \{0\}} \frac{A_h(w, v)}{\|v\|_{A_h}}, \quad \forall w \in V_h^p. \quad (6.33)$$

*Proof.* The continuity follows directly from the Cauchy–Schwarz inequality. The bilinear form  $A_h$  is coercive  $A_h(v, v) = \|v\|_{A_h}^2$  by definition and the inf-sup condition (6.33) follows from coercivity.  $\square$

LEMMA 6.4 (Inverse inequality). There are constants, independent of the mesh size  $h$  and of how the surface cuts the background mesh, such that the following inverse inequalities hold

$$\|v\|_{A_h}^2 \lesssim h^{-3} \|v\|_{L^2(\mathcal{T}_h)}^2 \quad v \in V_h^p, \quad (6.34)$$

$$\|v\|_{M_h}^2 \lesssim h^{-1} \|v\|_{L^2(\mathcal{T}_h)}^2 \quad v \in V_h^p. \quad (6.35)$$

*Proof.* Using the element wise trace inequality (4.4) and the inverse inequality (4.5), we obtain

$$\begin{aligned} \|v\|_{a_h}^2 &\lesssim h^{-1} \|v\|_{H^1(\mathcal{T}_h)}^2 + h \|v\|_{H^2(\mathcal{T}_h)}^2 \lesssim h^{-3} \|v\|_{L^2(\mathcal{T}_h)}^2, \\ \|v\|_{L^2(\Gamma_h)}^2 &\lesssim h^{-1} \|v\|_{L^2(\mathcal{T}_h)}^2. \end{aligned} \quad (6.36)$$

The result follows from recalling the definition of  $\|v\|_{A_h}^2$  and  $\|v\|_{M_h}^2$ , using the above estimates and Property P4 and M1 of the stabilization terms.  $\square$

LEMMA 6.5 (Poincaré inequality). There are constants, independent of the mesh size  $h$  and of how the surface cuts the background mesh, such that the following Poincaré inequalities hold

$$\|v\|_{L^2(\mathcal{T}_h)}^2 \lesssim h \|v\|_{A_h}^2 \quad v \in V_h^p, \quad (6.37)$$

$$\|v\|_{L^2(\mathcal{T}_h)}^2 \lesssim h \|v\|_{M_h}^2 \quad v \in V_h^p. \quad (6.38)$$

*Proof.* To prove the inequalities we will proceed in two main steps: 1. We use the face penalty to reach an element that has a large intersection with  $\Gamma_h$ . Here we employ a covering of  $\mathcal{T}_h$ , where each covering set contains elements that have a so-called large intersection property (see [Burman et al., 2015b](#)). 2. For elements with a large intersection the normal derivative control on  $\Gamma_h$  is used to control the  $L^2$  norm on the element.

To establish the inequality (6.37) and (6.38) we shall prove that

$$\|v\|_{L^2(\mathcal{T}_h)}^2 \lesssim \sum_{j=0}^p h^{2j+1} \|D_{n_h}^j v\|_{L^2(\Gamma_h)}^2 + \sum_{j=1}^p h^{2j+1} \|[D_{n_F}^j v]\|_{L^2(\mathcal{T}_h)}^2. \quad (6.39)$$

Here we note that the right-hand side of (6.39) satisfies the estimate

$$\begin{aligned} &\sum_{j=0}^p h^{2j+1} \|D_{n_h}^j v\|_{L^2(\Gamma_h)}^2 + \sum_{j=1}^p h^{2j+1} \|[D_{n_F}^j v]\|_{L^2(\mathcal{T}_h)}^2 \\ &\lesssim h (\|v\|_{\Gamma_h}^2 + h^{2(1-\gamma)} \|v\|_{s_h}^2), \end{aligned} \quad (6.40)$$

where we used the definition (3.7) of  $s_h$ . Thus, (6.38) follows from (6.39) using (6.40), the definition of  $\|v\|_{M_h}^2$  and the fact that  $\gamma = 1$ . The Poincaré inequality (4.2) and the fact that  $\gamma \in [0, 1]$  yields

$$h (\|v\|_{\Gamma_h}^2 + h^{2(1-\gamma)} \|v\|_{s_h}^2) \lesssim h (\|v\|_{a_h}^2 + h^{2(1-\gamma)} \|v\|_{s_h}^2) \lesssim h \|v\|_{A_h}^2.$$

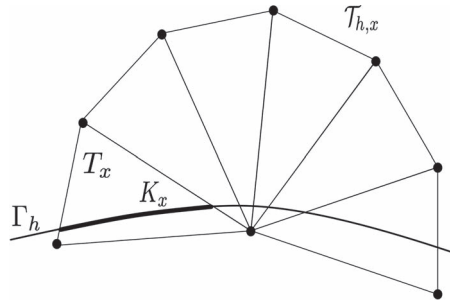


FIG. 2. A set of elements  $\mathcal{T}_{h,x}$  in the cover of  $\mathcal{T}_h$ , an element  $T_x$  with large intersection  $K_x = T_x \cap \Gamma_h$ .

Thus, we see that also (6.37) follows from (6.39). Next we show that (6.39) holds.

**Large intersection coverings of  $\mathcal{T}_h$ .** There is a covering  $\{\mathcal{T}_{h,x} : x \in \mathcal{X}_h\}$  of  $\mathcal{T}_h$ , where  $\mathcal{X}_h$  is an index set, such that (1) each set  $\mathcal{T}_{h,x}$  contains a uniformly bounded number of elements; (2) each element in  $\mathcal{T}_{h,x}$  shares at least one face with another element in  $\mathcal{T}_{h,x}$ ; (3) in each set  $\mathcal{T}_{h,x}$  there is one element  $T_x \in \mathcal{T}_{h,x}$ , which has a large intersection with  $\Gamma_h$ ,

$$h^{d-1} \lesssim |T_x \cap \Gamma_h| = |K_x|, \quad (6.41)$$

where  $d$  is the dimension. See Fig. 2 for an illustration; (4) the number of sets  $\mathcal{T}_{h,x}$  to which an element  $T$  belongs is uniformly bounded for all  $T \in \mathcal{T}_h$ . See Burman *et al.* (2015b) for the construction of the covering.

We prove (6.39) by considering a set  $\mathcal{T}_{h,x}$  in the covering and the following steps.

**Step 1.** We shall show that

$$\|v\|_{L^2(\mathcal{T}_{h,x})}^2 \lesssim \|v\|_{L^2(T_x)}^2 + \sum_{j=1}^p h^{2j+1} ([D_{n_F}^j v], [D_{n_F}^j w])_{\mathcal{T}_{h,x}}, \quad (6.42)$$

where  $\mathcal{T}_{h,x}$  is the set of interior faces in  $\mathcal{T}_{h,x}$ . To prove (6.42) we recall that for two elements  $T_1$  and  $T_2$  that share a face  $F$  it holds

$$\|v\|_{L^2(T_1)}^2 \lesssim \|v\|_{L^2(T_2)}^2 + \sum_{j=1}^p h^{2j+1} \| [D_{n_F}^j v] \|^2_{L^2(F)} \quad (6.43)$$

(see Hansbo *et al.*, 2017), for instance. Now let  $\mathcal{T}_{h,x}^0 = \{T_x\}$  and for  $j = 1, 2, \dots$  let  $\mathcal{T}_{h,x}^j$  be the set of all elements that share a face with an element in  $\mathcal{T}_{h,x}^{j-1}$ . Then there is a uniform constant  $J$  such that

$\mathcal{T}_{h,x}^j = \mathcal{T}_{h,x}^{j-1}$  for  $j \geq J$  and we also have the estimate

$$\|v\|_{\mathcal{T}_{h,x}^j} \lesssim \|v\|_{\mathcal{T}_{h,x}^{j-1}} + \sum_{F \subset \mathcal{T}_{h,x}^j \setminus \mathcal{T}_{h,x}^{j-1}} \sum_{j=1}^p h^{2j+1} \| [D_{n_F}^j v] \|_{L^2(F)}^2, \quad (6.44)$$

where  $\mathcal{T}_{h,x}^j$  is the set of interior faces in  $\mathcal{T}_{h,x}^j$ . Iterating (6.44) we obtain (6.42).

**Step 2.** We shall show that there is a constant such that for all  $T_x \in \mathcal{T}_h$ , which have the large intersection property (6.41),

$$\|v\|_{L^2(T_x)}^2 \lesssim \sum_{j=0}^p h^{2j+1} \|D_{n_h}^j v\|_{L^2(K_x)}^2, \quad (6.45)$$

where  $K_x = \Gamma_h \cap T_x$ . To verify (6.45) we define the cylinder

$$\text{Cyl}_\delta(K_x) = \{x \in \mathbb{R}^d : x = y + tn_h, y \in K_x, |t| \leq \delta\} \quad (6.46)$$

and using Taylor's formula for a polynomial  $v \in P_p(\text{Cyl}_\delta(K_x))$ , we obtain the bound

$$\|v\|_{L^2(\text{Cyl}_\delta(K_x))}^2 \lesssim \sum_{j=0}^p \delta^{2j+1} \|D_{n_h}^j v\|_{L^2(K_x)}^2. \quad (6.47)$$

Using the following estimate, which we verify below, there is a constant such that for all  $v \in P_p(T_x)$ ,

$$\|v\|_{L^2(T_x)} \lesssim \|v\|_{L^2(\text{Cyl}_\delta(K_x))}, \quad (6.48)$$

the desired estimate (6.45) follows for  $\delta \sim h$ .

**Verification of (6.48).** To employ a scaling argument we will construct two regular cylinders with circular cross section and the same centre line that may be mapped to a reference configuration, one containing  $T_x$  and one contained in  $\text{Cyl}_\delta(K_x)$ . See Fig. 3.

Let  $F_x$  be a plane with unit normal  $\bar{n}_x$ , which is tangent to  $K_x$  at an interior point of  $K_x$ . Then we have the bound

$$\|\bar{n}_h - n_h\|_{L^\infty(K_x)} \lesssim h \quad (6.49)$$

and with  $\bar{K}_x$  the closest point projection of  $K_x$  onto  $\bar{F}_x$  we conclude using shape regularity and a uniform bound on the curvature  $\kappa_h$  (Assumption A1), that there is a ball  $\bar{B}_{\bar{r}_1, x} \subset \bar{K}_x \subset \bar{F}_x$  with radius  $\bar{r}_1 \sim h$  and a  $\bar{\delta}_1 \sim h$  such that

$$\text{Cyl}_{\bar{\delta}_1}(\bar{K}_x) = \{x \in \mathbb{R}^d : x = y + t\bar{m}_h, y \in \bar{K}_x, |t| \leq \bar{\delta}_1\} \subset \text{Cyl}_\delta(K_x). \quad (6.50)$$

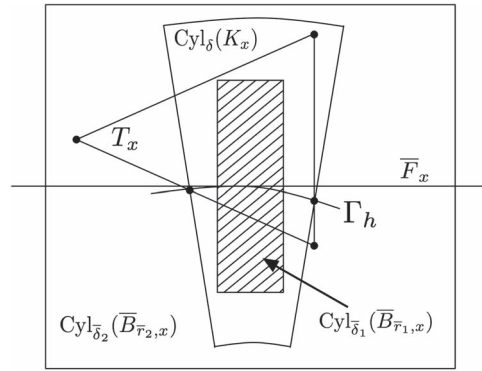


FIG. 3. Schematic figure illustrating a typical configuration of an element  $T$ , the curved intersection with  $\Gamma_h$ , the plane  $\bar{F}_x$ , which is tangent to  $K_x$ , the cylinders  $\text{Cyl}_\delta(K_x)$ ,  $\text{Cyl}_{\delta_1}(\bar{B}_{\bar{r}_1,x})$  and  $\text{Cyl}_{\delta_2}(\bar{B}_{\bar{r}_2,x})$ . Note that we have the inclusions  $\bar{B}_{\bar{r}_1,x} \subset \bar{K}_x \subset \bar{T}_x \subset \bar{B}_{\bar{r}_2,x}$ ,  $T_x \subset \text{Cyl}_{\delta_2}(\bar{B}_{\bar{r}_2,x})$  and  $\text{Cyl}_{\delta_1}(\bar{B}_{\bar{r}_1,x}) \subset \text{Cyl}_\delta(K_x)$ .

Next using shape regularity there is a larger ball  $\bar{B}_{\bar{r}_2,x} \subset \bar{F}_x$  with the same centre as  $\bar{B}_{\bar{r}_1,x}$  such that

$$\bar{T}_x \subset \bar{B}_{\bar{r}_2,x} \subset \bar{F}_x, \quad (6.51)$$

where  $\bar{T}_x$  is the closest point projection of  $T_x$  onto  $\bar{F}_x$ , and a  $\bar{\delta}_2 \sim h$  such that

$$T_x \subset \text{Cyl}_{\bar{\delta}_2}(\bar{B}_{\bar{r}_2,x}). \quad (6.52)$$

Clearly,  $\text{Cyl}_{\bar{\delta}_1}(\bar{B}_{\bar{r}_1,x}) \subset \text{Cyl}_{\bar{\delta}_2}(\bar{B}_{\bar{r}_2,x})$  and using a mapping to a reference configuration, we conclude that there is a constant such that for all polynomials  $v \in P_p(\text{Cyl}_{\bar{\delta}_2}(\bar{B}_{\bar{r}_2,x}))$ ,

$$\|v\|_{\text{Cyl}_{\bar{\delta}_2}(\bar{B}_{\bar{r}_2,x})} \lesssim \|v\|_{\text{Cyl}_{\bar{\delta}_1}(\bar{B}_{\bar{r}_1,x})}, \quad (6.53)$$

which in view of the inclusions (6.50) and (6.52) concludes the proof of (6.48).

**Step 3.** Combining equation (6.42) and (6.45) and using that  $\{\mathcal{T}_{h,x} : x \in \mathcal{X}_h\}$  is a cover of  $\mathcal{T}_h$  completes the proof of (6.39) and the lemma.  $\square$

**REMARK 6.6** Using the same technique as in the proof of Lemma 6.5 we may show the stronger estimate

$$\|v\|_{\mathcal{T}_h}^2 + h^{2\gamma} \|\nabla v\|_{\mathcal{T}_h}^2 \lesssim h \|v\|_{A_h}^2, \quad (6.54)$$

where  $\gamma \in [0, 1]$  is the scaling parameter in  $s_h$ , see (3.8) and (3.9). The modifications are as follows: in Step 1 we show that

$$\|\nabla v\|_{L^2(\mathcal{T}_{h,x})}^2 \lesssim \|\nabla v\|_{L^2(T_x)}^2 + \sum_{j=1}^p h^{2j-1} \| [D_{n_F}^j v] \|_{L^2(F)}^2, \quad (6.55)$$

which after multiplication by  $h^{2\gamma}$  corresponds to

$$h^{2\gamma} \|\nabla v\|_{L^2(\mathcal{T}_{h,x})}^2 \lesssim h^{2\gamma} \|\nabla v\|_{L^2(T_x)}^2 + h \left( \sum_{j=1}^p h^{2(j-1+\gamma)} \| [D_{n_F}^j v] \|_{L^2(F)}^2 \right). \quad (6.56)$$

To show (6.55) we use the estimate

$$\|\nabla v\|_{L^2(T_1)}^2 \lesssim \|\nabla v\|_{L^2(T_2)}^2 + \sum_{j=1}^p h^{2j-1} \| [D_{n_F}^j v] \|_{L^2(F)}^2. \quad (6.57)$$

In Step 2 we show that

$$\|\nabla v\|_{L^2(T_x)}^2 \lesssim \|\nabla_{\Gamma_h} v\|_{K_x}^2 + \sum_{j=1}^p h^{2j-1} \|D_{n_h}^j v\|_{L^2(K_x)}^2, \quad (6.58)$$

which after multiplication by  $h^{2\gamma}$  corresponds to

$$h^{2\gamma} \|\nabla v\|_{L^2(T_x)}^2 \lesssim \underbrace{h^{2\gamma} \|\nabla_{\Gamma_h} v\|_{K_x}^2}_{\lesssim \|\nabla_{\Gamma_h} v\|_{K_x}^2} + h \left( \sum_{j=1}^p h^{2(j-1+\gamma)} \|D_{n_h}^j v\|_{L^2(K_x)}^2 \right). \quad (6.59)$$

Combining (6.56) and (6.59), summing over the cover and using Property P4 of the cover, we obtain (6.54).

**REMARK 6.7** Note the following:

- From (5.43) we see that the scaling of the stabilization term corresponds to the mass matrix. This scaling is the weakest, which guarantees the Poincaré inequality (6.37) (Property P4 of the stabilization term).
- For the operator

$$L = \alpha \Delta_\Gamma + \beta, \quad (6.60)$$

where  $\alpha$  and  $\beta$  are constants, we would have the inverse inequality

$$\|v\|_{A_h} \lesssim (\alpha h^{-3/2} + \beta h^{-1/2}) \|v\|_{\mathcal{T}_h} \quad v \in V_h^p \quad (6.61)$$

instead of (6.34), and the resulting condition number estimate is then

$$\kappa(\mathbf{A}) \lesssim \alpha h^{-2} + \beta. \quad (6.62)$$

- In the case when

$$\|v\|_{a_h} \lesssim (\alpha h^{-3/2} + \beta h^{-1/2}) \|v\|_{\mathcal{T}_h} \quad (6.63)$$

we note that the desired inverse inequality (6.61) holds for

$$\tilde{s}_h = (\alpha h^{-\tau} + \beta) s_h, \quad (6.64)$$

with  $\tau$  in the interval  $0 \leq \tau \leq 2$ , since

$$\|v\|_{\tilde{s}_h}^2 = (\alpha h^{-\tau} + \beta) \|v\|_{s_h}^2 \lesssim (\alpha h^{-\tau} + \beta) h^{-1} \|v\|_{\mathcal{T}_h}^2 \lesssim (\alpha h^{-(1+\tau)} + \beta h^{-1}) \|v\|_{\mathcal{T}_h}^2. \quad (6.65)$$

## 7. Extension to problems on manifolds with general codimension embeddings

The stabilization term (3.7) may be extended to the more general case of an  $n$ -dimensional smooth manifold embedded in  $\mathbb{R}^d$ , with codimension  $cd = d - n > 1$ , by scaling the face penalty term (3.8) in such a way that the stabilization terms associated with the faces and surface scale in the same way,

$$s_{h,F}(w, v) = \sum_{j=1}^p c_{F,j} h^{2(j-1+\gamma)} h^{1-cd} ([D_{n_F}^j w], [D_{n_F}^j v])_{\mathcal{F}_h}, \quad (7.1)$$

$$s_{h,\Gamma}(w, v) = \sum_{j=1}^p c_{\Gamma,j} h^{2(j-1+\gamma)} (D_{n_h}^j w, D_{n_h}^j v)_{\mathcal{K}_h}. \quad (7.2)$$

This is the same scaling as is used for the face stabilization term in the case of piecewise linear elements in general codimension (see Burman *et al.*, 2018, Table 1). With this definition P1–P3 remain the same and may be verified using the results by Burman *et al.* (2018). P4 and P5 take the more general form

**P4.** For  $v \in V_h^p$ ,

$$\|v\|_{s_h}^2 \lesssim h^{-(cd+2)} \|v\|_{L^2(\mathcal{T}_h)}^2. \quad (7.3)$$

**P5.** For  $v \in V_h^p$ ,

$$\|v\|_{L^2(\mathcal{T}_h)}^2 \lesssim h^{cd} (\|v\|_{a_h}^2 + \|v\|_{s_h}^2). \quad (7.4)$$

P4 is verified as in equation (5.43) and below we comment on the minor modifications in the proof of Lemma 6.5 necessary to verify P5.

We next extend the proof of Lemma 6.5, to the case of a general embedding of an  $n$ -dimensional surface in  $\mathbb{R}^d$ . To that end we first note that Step 1 can be directly carried out for  $d$ -dimensional simplices and we obtain

$$\|v\|_{L^2(\mathcal{T}_{h,x})}^2 \lesssim \|v\|_{L^2(T_x)}^2 + h^{cd} \left( \sum_{j=1}^p h^{2j} h^{1-cd} ([D_{n_F}^j v], [D_{n_F}^j w])_{\mathcal{F}_{h,x}} \right). \quad (7.5)$$

For Step 2, the only difference is that we will have  $cd = d - n$  orthonormal normal directions  $\{n_{h,i}\}_{i=1}^{cd}$ , which we may choose to vary smoothly over  $K_x$  (see Burman *et al.*, 2018) for details, and therefore the

definition of the cylinder over  $K_x$  takes the form

$$\text{Cyl}_\delta(K_x) = \left\{ x \in \mathbb{R}^d : x = y + \sum_{i=1}^{cd} t_i n_{h,i}, y \in K_x, |t_i| \leq \delta, i = 1, \dots, cd \right\}. \quad (7.6)$$

Using Taylor's formula in several dimensions and integrating over the cylinder (7.6), we obtain the following generalization of (6.47),

$$\|v\|_{L^2(\text{Cyl}_\delta(K_x))}^2 \lesssim \sum_{j=0}^p \delta^{2j+cd} \|D_{n_h}^j v\|_{L^2(K_x)}^2, \quad (7.7)$$

where we used the fact that, for a monomial  $\Pi_{i=1}^{cd} x_i^{2l_i}$ , with  $\sum_{i=1}^{cd} l_i = j$ , it holds

$$\int_{[-\delta, \delta]^{cd}} \Pi_{i=1}^{cd} x_i^{2l_i} = \Pi_{i=1}^{cd} \int_{-\delta}^{\delta} x_i^{2l_i} dx_i \sim \Pi_{i=1}^{cd} \delta^{2l_i+1} = \delta^{2j+cd}. \quad (7.8)$$

We finally note that (6.48) holds in any dimension  $d$  and the desired estimate in Step 2 follows for  $\delta \sim h$ , and Step 3 is just a combination of Steps 1 and 2. Thus, the proof of Lemma 6.5 easily extends to the case of general codimension embeddings and we conclude that (7.4) holds.

## 8. A priori error estimates

In this section we prove optimal error estimates in the energy norm and in the  $L^2$  norm.

### 8.1 Strang lemma

Using that  $A_h$  is continuous and satisfies the inf-sup condition, we first show a Strang lemma, which connects the error in the energy norm to the interpolation error and the consistency error.

**LEMMA 8.1** Let  $u \in H_0^1(\Gamma)$  be the unique solution of (2.5) and  $u_h \in V_h^p$  the finite element approximation defined by (3.3). Then the following discretization error bound holds

$$\|u^e - u_h\|_{A_h} \lesssim \|u^e - \pi_h^p u^e\|_{A_h} + \sup_{v \in V_h^p \setminus \{0\}} \frac{|A_h(u^e, v) - l_h(v)|}{\|v\|_{A_h}}. \quad (8.1)$$

*Proof.* Adding and subtracting an interpolant and using the triangle inequality, we get

$$\|u^e - u_h\|_{A_h} \leq \|u^e - \pi_h^p u^e\|_{A_h} + \|\pi_h^p u^e - u_h\|_{A_h}. \quad (8.2)$$

Using the inf-sup condition (6.33) for  $A_h$  we have

$$\|\pi_h^p u^e - u_h\|_{A_h} \lesssim \sup_{v \in V_h^p \setminus \{0\}} \frac{A_h(\pi_h^p u^e - u_h, v)}{\|v\|_{A_h}}. \quad (8.3)$$

Adding and subtracting  $A_h(u^e, v)$  and using the weak formulation (3.3) yields

$$A_h(\pi_h^p u^e - u_h, v) = A_h(\pi_h^p u^e - u^e, v) + A_h(u^e - u_h, v) \quad (8.4)$$

$$= A_h(\pi_h^p u^e - u^e, v) + A_h(u^e, v) - l_h(v). \quad (8.5)$$

Finally, using the continuity of  $A_h$  we obtain

$$A_h(\pi_h^p u^e - u^e, v) \leq \|\pi_h^p u^e - u^e\|_{A_h} \|v\|_{A_h}. \quad (8.6)$$

Collecting the estimates we end up with the desired bound

$$\|u^e - u_h\|_{A_h} \lesssim \|u^e - \pi_h^p u^e\|_{A_h} + \sup_{v \in V_h^p \setminus \{0\}} \frac{|A_h(u^e, v) - l_h(v)|}{\|v\|_{A_h}}. \quad (8.7)$$

□

## 8.2 The interpolation error

Next we prove optimal interpolation error estimates using Property P1 of the stabilization term.

LEMMA 8.2 For all  $u \in H^{p+1}(\Gamma)$  we have the following estimate:

$$\|u^e - \pi_h^p u^e\|_{A_h} \lesssim h^p \|u\|_{H^{p+1}(\Gamma)}. \quad (8.8)$$

*Proof.* From the definition of the energy norm  $\|\cdot\|_{A_h}$ , equation (4.1), we have

$$\|u^e - \pi_h^p u^e\|_{A_h}^2 = \underbrace{\|u^e - \pi_h^p u^e\|_{a_h}^2}_I + \underbrace{\|u^e - \pi_h^p u^e\|_{s_h}^2}_H. \quad (8.9)$$

**Term I.** Using the element-wise trace inequality (4.4), standard interpolation estimates (4.6) on elements  $T \in \mathcal{T}_h$  and the stability estimate (4.8) for the extension operator with  $\delta \sim h$ , we obtain

$$I = \|u^e - \pi_h^p u^e\|_{a_h}^2 \quad (8.10)$$

$$\lesssim \sum_{T \in \mathcal{T}_h} \left( h^{-1} \|u^e - \pi_h^p u^e\|_{H^1(T)}^2 + h \|u^e - \pi_h^p u^e\|_{H^2(T)}^2 \right) \quad (8.11)$$

$$\lesssim h^{2p} \|u\|_{H^{p+1}(\Gamma)}^2. \quad (8.12)$$

**Term II.** From Property P1 of the stabilization term, we have that

$$H \lesssim h^{2p} \|u\|_{H^{p+1}(\Gamma)}^2. \quad (8.13)$$

Combining the two estimates (8.10) and (8.13) the result follows. □

### 8.3 The consistency error

The approximation of the geometry  $\Gamma$  by  $\Gamma_h$  and the approximation of the data  $f$  by  $f_h$  leads to a consistency error that we estimate in the next lemma.

**LEMMA 8.3** Let  $u \in H^{p+1}(\Gamma) \cap H_0^1(\Gamma)$  be the solution of (2.5), then the following bound holds

$$\sup_{v \in V_h^p \setminus \{0\}} \frac{|A_h(u^e, v) - l_h(v)|}{\|v\|_{A_h}} \lesssim h^p \|u\|_{H^{p+1}(\Gamma)} + h^{p+1} \|f\|_{L^2(\Gamma)}. \quad (8.14)$$

*Proof.* We have the identity

$$|A_h(u^e, v) - l_h(v)| \leq \underbrace{|a_h(u^e, v) - l_h(v)|}_I + \underbrace{|s_h(u^e, v)|}_{II}. \quad (8.15)$$

**Term I.** Adding  $-a(u, v^l) + l(v^l) = 0$  and using the triangle inequality

$$I \leq \underbrace{|a_h(u^e, v) - a(u, v^l)|}_{I_I} + \underbrace{|l(v^l) - l_h(v)|}_{I_{II}}. \quad (8.16)$$

**Term  $I_I$ .** Adding and subtracting  $\int_{\Gamma_h} \nabla_{\Gamma_h} u^e \cdot \nabla_{\Gamma_h} v |B| \, ds_h$ , using that  $B^T B^{-T} = P_{\Gamma_h}$ , changing the domain of integration from  $\Gamma$  to  $\Gamma_h$ , using the triangle inequality, norm equivalences in (4.25), estimates (4.20) and (4.23), we obtain

$$\begin{aligned} |a_h(u^e, v) - a(u, v^l)| &= \left| \int_{\Gamma_h} (1 - |B|) \nabla_{\Gamma_h} u^e \cdot \nabla_{\Gamma_h} v \, ds_h \right. \\ &\quad + \int_{\Gamma_h} B^T (B^{-T} \nabla_{\Gamma_h} u^e) \cdot B^T (B^{-T} \nabla_{\Gamma_h} v) |B| \, ds_h \\ &\quad \left. - \int_{\Gamma_h} (B^{-T} \nabla_{\Gamma_h} u^e) \cdot (B^{-T} \nabla_{\Gamma_h} v) |B| \, ds_h \right| \\ &\lesssim \left( \|1 - |B|\|_{L^\infty(\Gamma_h)} + \|BB^T - I\|_{L^\infty(\Gamma_h)} \right) \|\nabla_{\Gamma} u\|_{L^2(\Gamma)} \|\nabla_{\Gamma_h} v\|_{L^2(\Gamma_h)} \\ &\lesssim h^{p+1} \|\nabla_{\Gamma} u\|_{L^2(\Gamma)} \|v\|_{a_h}. \end{aligned} \quad (8.17)$$

**Term  $I_{II}$ .** Changing the domain of integration we get

$$|l(v^l) - l_h(v)| = \left| \int_{\Gamma} f v^l \, ds - \int_{\Gamma_h} f_h v \, ds_h \right| \quad (8.18)$$

$$= \left| \int_{\Gamma_h} f^e v |B| \, ds_h - \int_{\Gamma_h} f_h v \, ds_h \right| \quad (8.19)$$

$$\lesssim \| |B| f^e - f_h \|_{L^2(\Gamma_h)} \|v\|_{L^2(\Gamma_h)} \quad (8.20)$$

$$\lesssim h^{p+1} \|f\|_{L^2(\Gamma)} \|v\|_{L^2(\Gamma)}, \quad (8.21)$$

where we at last used Assumption A3 on the data approximation  $f_h$ . Together, the bounds of  $I_I$  and  $I_{II}$ , and the Poincaré inequality (4.2), imply

$$I \lesssim h^{p+1} \|\nabla_\Gamma u\|_{L^2(\Gamma)} \|v\|_{a_h} + h^{p+1} \|f\|_{L^2(\Gamma)} \|v\|_{L^2(\Gamma_h)} \quad (8.22)$$

$$\lesssim (\|\nabla_\Gamma u\|_{L^2(\Gamma)} + \|f\|_{L^2(\Gamma)}) h^{p+1} \|v\|_{a_h}. \quad (8.23)$$

Finally, using that  $u$  is the solution of (2.5) we have the stability estimate  $\|\nabla_\Gamma u\|_{L^2(\Gamma)} \lesssim \|f\|_{L^2(\Gamma)}$ , and we obtain

$$I \lesssim h^{p+1} \|f\|_{L^2(\Gamma)} \|v\|_{a_h}. \quad (8.24)$$

**Term II.** Using the Cauchy–Schwarz inequality and Property P2 of the stabilization term, see Section 5, we obtain

$$|s_h(u^e, v)| \lesssim \|u^e\|_{s_h} \|v\|_{s_h} \lesssim h^p \|u\|_{H^{p+1}(\Gamma)} \|v\|_{s_h}. \quad (8.25)$$

Combining the estimates of  $I$  and  $II$  we obtain the desired estimate

$$|A_h(u^e, v) - I_h(v)| \lesssim h^p \|u\|_{H^{p+1}(\Gamma)} \|v\|_{A_h} + h^{p+1} \|f\|_{L^2(\Gamma)} \|v\|_{a_h}. \quad (8.26)$$

□

#### 8.4 Error estimates

We first prove optimal error estimates in the energy norm and then apply a duality argument to obtain  $L^2$ -error estimates.

**THEOREM 8.4** Let  $u \in H^{p+1}(\Gamma) \cap H_0^1(\Gamma)$  be the solution of (2.5) and  $u_h \in V_h^p$  the finite element approximation defined by (3.3). If assumptions A1–A3 hold and the stabilization term has properties P1–P5, then there is a constant independent of the mesh size  $h$  such that the following error bound holds

$$\|u^e - u_h\|_{A_h} \lesssim h^p \|u\|_{H^{p+1}(\Gamma)} + h^{p+1} \|f\|_{L^2(\Gamma)}. \quad (8.27)$$

*Proof.* Using the Strang lemma followed by the bounds on the interpolation error (8.8) and the consistency error (8.14), we get the desired bound. □

**THEOREM 8.5** Let  $u \in H^{p+1}(\Gamma) \cap H_0^1(\Gamma)$  be the solution of (2.5) and  $u_h \in V_h^p$  the finite element approximation defined by (3.3). If assumptions A1–A3 hold and the stabilization term has properties P1–P5, then there is a constant independent of the mesh size  $h$  such that the following error bound holds

$$\|u^e - u_h\|_{L^2(\Gamma_h)} \lesssim h^{p+1} \|u\|_{H^{p+1}(\Gamma)} + h^{p+1} \|f\|_{L^2(\Gamma)}. \quad (8.28)$$

*Proof.* The proof is similar to the proof of the  $L^2$ -error estimate by Burman *et al.* (2015b) for  $p = 1$ . Let  $e_h = u^e - u_h|_{\Gamma_h}$  and its lift on  $\Gamma$  be  $e_h^l = u - u_h^l$ . Recall that  $\int_\Gamma u \, ds = 0$  and  $\int_{\Gamma_h} u_h \, ds_h = 0$ . We now define  $\tilde{u}_h \in V_h^p$  as

$$\tilde{u}_h = u_h - |\Gamma|^{-1} \int_\Gamma u_h^l \, ds \quad (8.29)$$

so that we have  $\int_{\Gamma} \tilde{u}_h^l = 0$ . By adding and subtracting  $\tilde{u}_h^l$  and using the triangle inequality, we obtain

$$\|e_h^l\|_{L^2(\Gamma)} \leq \underbrace{\|u - \tilde{u}_h^l\|_{L^2(\Gamma)}}_I + \underbrace{\|\tilde{u}_h^l - u_h^l\|_{L^2(\Gamma)}}_{II}. \quad (8.30)$$

**Term I.** Consider the dual problem:

$$a(v, \phi) = (\psi, v)_{\Gamma}, \quad \psi \in L^2(\Gamma) \setminus \mathbb{R}. \quad (8.31)$$

It follows from the Lax–Milgram lemma that there exists a unique solution in  $H^1(\Gamma) \setminus \mathbb{R}$  and we also have the elliptic regularity estimate

$$\|\phi\|_{H^2(\Gamma)} \lesssim \|\psi\|_{L^2(\Gamma)}. \quad (8.32)$$

Setting  $v = u - \tilde{u}_h^l$  in (8.31), adding and subtracting an interpolant and using the weak formulations in (2.5) and (3.3), we obtain

$$(u - \tilde{u}_h^l, \psi)_{\Gamma} = a(u - \tilde{u}_h^l, \phi) \quad (8.33)$$

$$= a(u - \tilde{u}_h^l, \phi - (\pi_h^p \phi^e)^l) + a(u - \tilde{u}_h^l, (\pi_h^p \phi^e)^l) \quad (8.34)$$

$$= \underbrace{a(u - \tilde{u}_h^l, \phi - (\pi_h^p \phi^e)^l)}_{I_I} + \underbrace{l((\pi_h^p \phi^e)^l) - l_h(\pi_h^p \phi^e)}_{I_{II}} \quad (8.35)$$

$$+ \underbrace{a_h(\tilde{u}_h, \pi_h^p \phi^e) - a(\tilde{u}_h^l, (\pi_h^p \phi^e)^l)}_{I_{III}} + \underbrace{s_h(\tilde{u}_h, \pi_h^p \phi^e)}_{I_{IV}}. \quad (8.36)$$

**Term  $I_I$ .** Using the Cauchy–Schwarz inequality, the norm equivalences (see Section 4), and the energy norm estimate (Theorem 8.4), we obtain

$$I_I \lesssim \|\nabla_{\Gamma}(u - \tilde{u}_h^l)\|_{L^2(\Gamma)} \|\nabla_{\Gamma}(\phi - (\pi_h^p \phi^e)^l)\|_{L^2(\Gamma)} \quad (8.37)$$

$$\lesssim \|u^e - u_h\|_{a_h} \|\phi^e - \pi_h^p \phi^e\|_{a_h} \quad (8.38)$$

$$\lesssim (h^p \|u\|_{H^{p+1}(\Gamma)} + h^{p+1} \|f\|_{L^2(\Gamma)}) \|\phi^e - \pi_h^p \phi^e\|_{a_h}. \quad (8.39)$$

The interpolation error estimate (8.10) with  $s = 1$  yields

$$\|\phi^e - \pi_h^p \phi^e\|_{a_h} \lesssim h \|\phi\|_{H^2(\Gamma)} \quad (8.40)$$

and finally, using the elliptic regularity estimate (8.32), we get

$$I_I \lesssim (h^{p+1} \|u\|_{H^{p+1}(\Gamma)} + h^{p+2} \|f\|_{L^2(\Gamma)}) \|\psi\|_{L^2(\Gamma)}. \quad (8.41)$$

**Term  $I_{II} - I_{III}$ .** We use the estimate of Term  $I$  in the proof of Lemma 8.3, together with the following stability estimate

$$\|\nabla_{\Gamma} \tilde{u}_h^l\|_{L^2(\Gamma)} = \|\nabla_{\Gamma} u_h^l\|_{L^2(\Gamma)} \lesssim \|u_h\|_{a_h} \lesssim \|f_h\|_{L^2(\Gamma_h)} \lesssim \|f\|_{L^2(\Gamma)} \quad (8.42)$$

to get

$$|I_{II} + I_{III}| \lesssim (h^{p+1} \|\nabla_{\Gamma} \tilde{u}_h^l\|_{L^2(\Gamma)} + h^{p+1} \|f\|_{L^2(\Gamma)}) \|\pi_h^p \phi^e\|_{a_h} \lesssim h^{p+1} \|f\|_{L^2(\Gamma)} \|\pi_h^p \phi^e\|_{a_h}. \quad (8.43)$$

Adding and subtracting an interpolant, using the interpolation error estimate (8.40) followed by the elliptic regularity estimate (8.32) yield

$$\|\pi_h^p \phi^e\|_{a_h} \lesssim \|\pi_h^p \phi^e - \phi^e\|_{a_h} + \|\phi^e\|_{a_h} \lesssim \|\phi\|_{H^2(\Gamma)} \lesssim \|\psi\|_{L^2(\Gamma)}. \quad (8.44)$$

Thus,

$$|I_{II} + I_{III}| \lesssim h^{p+1} \|f\|_{L^2(\Gamma)} \|\psi\|_{L^2(\Gamma)}. \quad (8.45)$$

**Term  $I_{IV}$ .** We have

$$|I_{IV}| = |s_h(u_h, \pi_h^p \phi^e)| \lesssim \|u_h\|_{s_h} \|\pi_h^p \phi^e\|_{s_h}. \quad (8.46)$$

Using the energy norm error estimate, Theorem 8.4 and Property P2 of the stabilization, we obtain

$$\|u_h\|_{s_h} \lesssim \|u_h - u^e\|_{s_h} + \|u^e\|_{s_h} \quad (8.47)$$

$$\lesssim h^p \|u\|_{H^{p+1}(\Gamma)} + h^{p+1} \|f\|_{L^2(\Gamma)} + h^p \|u\|_{H^{p+1}(\Gamma)} \quad (8.48)$$

and using Property P3 and the elliptic regularity estimate (8.32), we obtain

$$\|\pi_h^p \phi^e\|_{s_h} \lesssim h \|\phi\|_{H^2(\Gamma)} \lesssim h \|\psi\|_{L^2(\Gamma)}. \quad (8.49)$$

We thus have

$$|I_{IV}| \lesssim (h^{p+1} \|u\|_{H^{p+1}(\Gamma_h)} + h^{p+2} \|f\|_{L^2(\Gamma)}) \|\psi\|_{L^2(\Gamma)}. \quad (8.50)$$

**Final estimate of Term  $I$ .** Collecting the estimates of Terms  $I_I - I_{IV}$  and taking  $\psi = u - \tilde{u}_h^l / \|u - \tilde{u}_h^l\|_{L^2(\Gamma)}$ , we obtain

$$I = \|u - \tilde{u}_h^l\|_{L^2(\Gamma)} \lesssim h^{p+1} \|u\|_{H^{p+1}(\Gamma_h)} + h^{p+1} \|f\|_{L^2(\Gamma)}. \quad (8.51)$$

**Term  $II$ .** Using the definition (8.29) of  $\tilde{u}_h$ , adding  $\int_{\Gamma_h} u_h \, ds_h = 0$ , changing the domain of integration and using related bounds, employing the Poincaré inequality (4.2), and the stability estimate (8.42), we

obtain

$$\|\tilde{u}_h^l - u_h^l\|_{L^2(\Gamma)} = \left| \int_{\Gamma} u_h^l ds - \int_{\Gamma_h} u_h ds_h \right| \quad (8.52)$$

$$\lesssim \left| \int_{\Gamma_h} u_h (|B| - 1) ds_h \right| \quad (8.53)$$

$$\lesssim h^{p+1} \|u_h\|_{L^2(\Gamma_h)} \quad (8.54)$$

$$\lesssim h^{p+1} \|u_h\|_{a_h} \quad (8.55)$$

$$\lesssim h^{p+1} \|f\|_{L^2(\Gamma)}. \quad (8.56)$$

**Conclusion of the proof.** Using the estimates of Terms *I* and *II* in equation (8.30) and the equivalence  $\|e_h^l\|_{L^2(\Gamma)} \sim \|e_h\|_{L^2(\Gamma_h)}$  conclude the proof.  $\square$

## 9. Numerical examples

We consider three different problems. The mesh on the domain where the interface is embedded is always generated independently of the position of the interface  $\Gamma$  and we use Lagrange basis functions of degree  $p = 1, 2$  or  $3$ . For the first two problems we use an explicit representation of  $\Gamma$  by using a cubic spline parametrization (see Zahedi, 2018). The normal vector of  $\Gamma_h$  is directly computed from the parametrization. We compare the proposed stabilization with the stabilization term (1.10), and in the computation of the latter stabilization term normal vectors at quadrature points inside elements cut by the interface are needed. These normal vectors are not immediately available with our explicit representation of  $\Gamma_h$ , but for each quadrature point inside an element we find the closest point on  $\Gamma_h$  and use the normal vector at that point on  $\Gamma_h$ . In the last example, for the computation of the mean curvature vector, we use the level set method. We use Lagrange basis functions of degree  $p = 1$ . To evaluate the integrals on  $\Gamma_h$  we find the interface explicitly as the zero level set of the piecewise linear approximation of the level set function. The approximated geometry is then planar on each element and consists of triangles and quadrilaterals, which can be subdivided into triangles. In all examples, the approximation of the geometry is such that assumption A1 is satisfied.

### 9.1 The Laplace–Beltrami problem

Let the interface  $\Gamma$  be a circle of radius 1 centred at the origin. We generate a uniform triangular mesh with  $h = h_{x_1} = h_{x_2}$  on the computational domain:  $[-1.5, 1.5] \times [-1.5, 1.5]$ . A right-hand side  $f$  to equation (2.5) is calculated so that

$$u(x) = \frac{x_1^3 x_2^3}{(x_1^2 + x_2^2)^3} \quad (9.1)$$

is the exact solution. We discretize (2.5) using the proposed CutFEM with Lagrange basis functions of degree  $p = 1, 2$ , and  $3$ . We use a Lagrange multiplier approach to enforce the condition  $\int_{\Gamma_h} v ds_h = 0$ . This leads to a linear system with system matrix  $\mathbf{B}$  defined as in (6.20). We compute the spectral condition number as the ratio between the largest in magnitude and the smallest in magnitude eigenvalue of the system matrix.

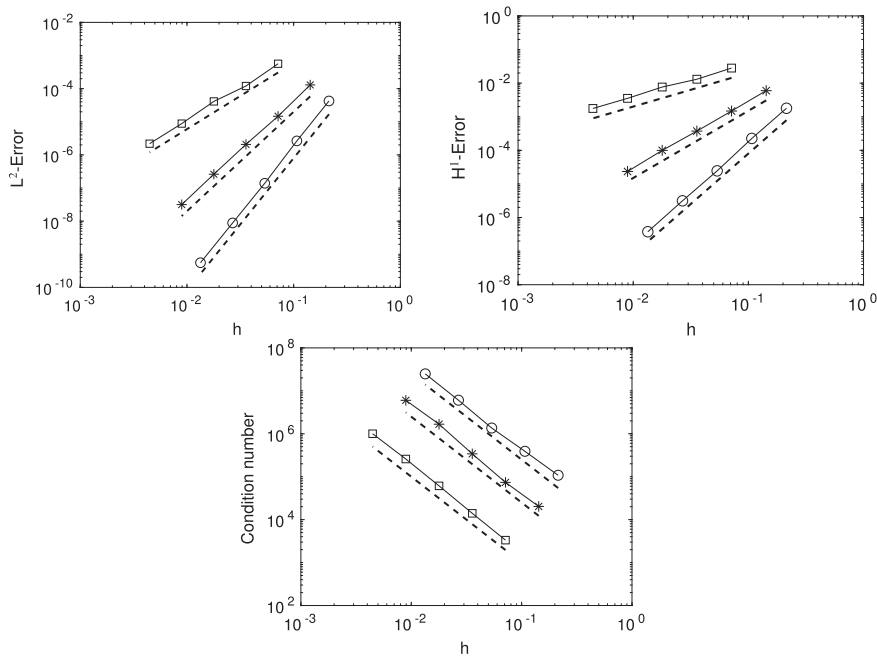


FIG. 4. The Laplace–Beltrami problem: the error and condition number vs mesh size  $h$  for different degrees of polynomials,  $p$ , in the discretization. Squares:  $p = 1$ . Stars:  $p = 2$ . Circles:  $p = 3$ . Top left: the error measured in the  $L^2$ -norm vs mesh size  $h$ . The dashed lines are indicating the expected rate of convergence and are proportional to  $h^{p+1}$ . Top right: the error measured in the  $H^1$ -norm vs mesh size  $h$ . The dashed lines are indicating the expected rate of convergence and are proportional to  $h^p$ . Bottom: the condition number vs mesh size  $h$ . The dashed lines are all proportional to  $h^{-2}$ .

In Fig. 4 we show the error and the spectral condition number of the linear systems as a function of the mesh size  $h$ . The error measured in the  $L^2$ -norm behaves as  $\mathcal{O}(h^{p+1})$  and in the  $H^1$ -norm as  $\mathcal{O}(h^p)$ , as expected. The condition number behaves as  $\mathcal{O}(h^{-2})$  independent of the polynomial degree and how the interface cuts the background mesh. The magnitude of the error and also the condition number depends on the stabilization constants in (3.8) and (3.9) and also on the choice of basis functions. For the results in Fig. 4 we have used  $\gamma = 1$ ,  $c_{F,j} = c_{\Gamma,j} = 2.5 \cdot 10^{-j}$ ,  $j = 1, \dots, p$ . We have not optimized this choice of parameters, and other choices may give better results for example smaller condition numbers with almost the same error. We show this for example in Figs 6 and 8. Other stabilization constants or basis functions can also give better scaling of the condition number with respect to the polynomial degree.

In Figs 5 and 6 we show the effect of the stabilization on the error and the condition number for linear ( $p = 1$ ) and cubic elements ( $p = 3$ ), respectively. We show the convergence of the error with respect to the  $L^2$ -norm. The results are similar when the error is measured in the  $H^1$ -norm. We compare the proposed stabilization with the stabilization terms (1.8) and (1.10) and also show the condition number after diagonal scaling.

We see in Fig. 5 that for linear elements the errors, measured in the  $L^2$ -norm, using the different stabilization terms and also no stabilization are very similar, but we clearly see that when no stabilization is used, the condition number can be extremely large depending on the position of the interface relative the background mesh. We have used  $c_{F,1} = c_{\Gamma,1} = 10^{-1}$  and  $\gamma = 1$  in the proposed stabilization. Note that the stabilization term (1.8) developed in Burman *et al.* (2015b) can be written in the form of

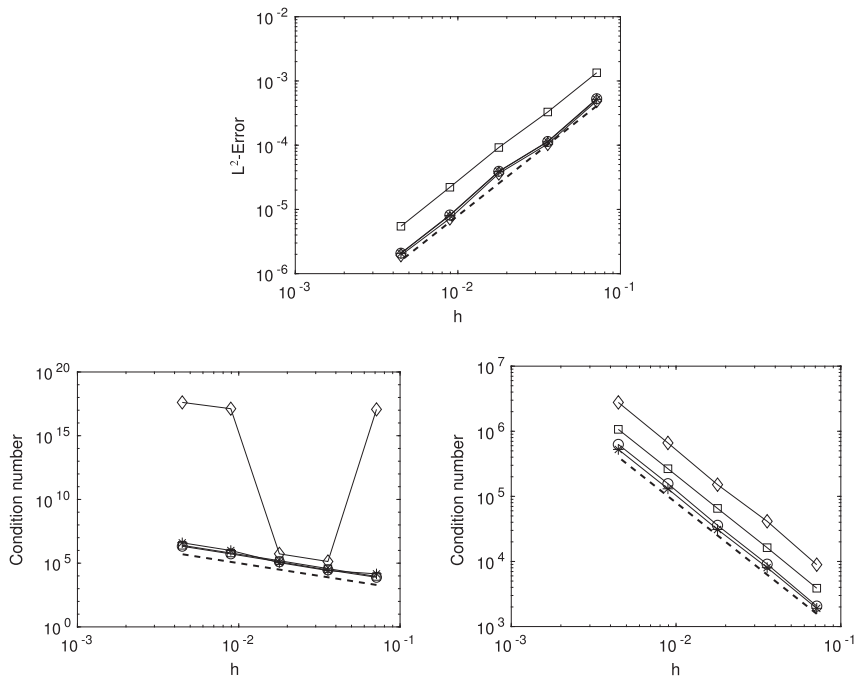


FIG. 5. The Laplace–Beltrami problem: the error and condition number vs mesh size  $h$  using different stabilization terms. Linear elements are used, i.e.,  $p = 1$ . Diamonds: no stabilization is added. Circles: the proposed stabilization with  $\gamma = 1$ . Squares: the pure face stabilization (1.8). Stars: the normal gradient stabilization (1.10) with  $\alpha = 1$ . Top: The error measured in the  $L^2$ -norm vs mesh size  $h$ . The dashed line is indicating the expected rate of convergence and is proportional to  $h^2$ . Bottom left: the spectral condition number versus mesh size  $h$ . The dashed line is proportional to  $h^{-2}$ . Bottom right: the spectral condition number after diagonal scaling versus mesh size  $h$ . The dashed line is proportional to  $h^{-2}$ .

the proposed stabilization, see (3.7)–(3.9), but with  $c_{\Gamma,j} = 0$  for all  $j$  and  $\gamma = 0$ . For this stabilization we have chosen  $c_{F,1} = 10^{-1}$  but, since  $\gamma = 0$ , we get slightly larger errors than the other alternatives. Decreasing the stabilization constant  $c_{F,1}$  will decrease the error. In the stabilization term (1.10) we have chosen  $c_{\mathcal{T}} = 10^{-1}$  and  $\alpha = 1$ . We see in Fig. 5 that an alternative to adding stabilization terms is to do a simple diagonal scaling.

For higher-order elements than linear, the resulting linear systems from unstabilized CutFEM are severely ill-conditioned, and neither a simple diagonal scaling alone nor a pure face stabilization term such as (1.8) improves the condition number significantly, see Fig. 6. When the condition number becomes too large the convergence fails; the error is dominated by roundoff errors and does not decrease with mesh refinement. In general, when the stabilization is not enough, the behaviour of the error depends on how the geometry cuts the background mesh. Therefore, if the approximation of the interface is slightly changed, the magnitude of the error and the convergence shown in Fig. 6, where cubic elements are used, may change a lot for the unstabilized method and the method with only face stabilization. With the proposed stabilization and also with the stabilization term (1.10), we obtain as expected optimal convergence orders, and the condition number of the linear systems scales as  $\mathcal{O}(h^{-2})$  independent of how the interface cuts the background mesh. In Fig. 6 we also show results

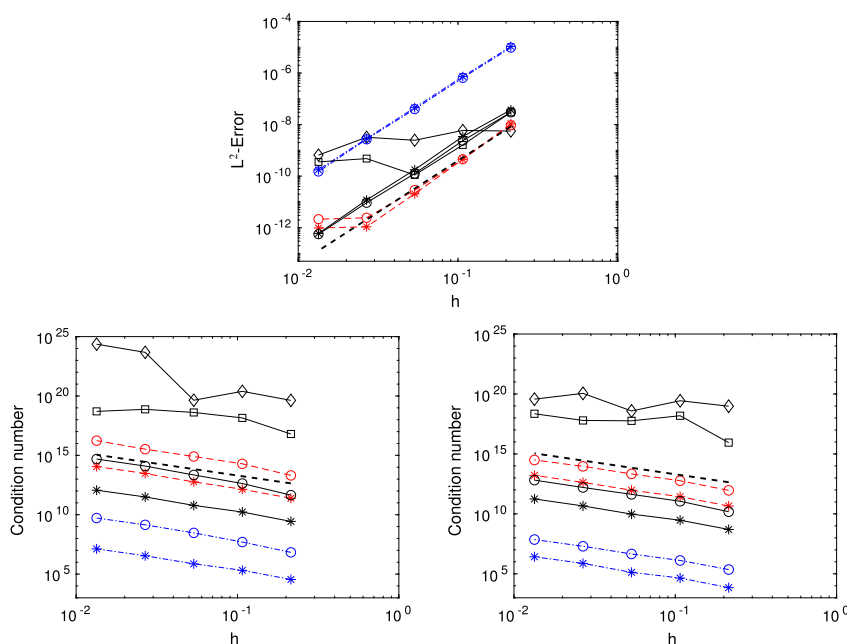


FIG. 6. The Laplace–Beltrami problem: the error and condition number vs mesh size  $h$  using different stabilization terms. Cubic elements are used, i.e.,  $p = 3$ . Diamonds: no stabilization is added. Circles: the proposed stabilization with  $\gamma = 1$ . Squares: pure face stabilization. Stars: the stabilization term (1.10) with  $\alpha = 1$ . Top: the error measured in the  $L^2$ -norm vs mesh size  $h$ . The dashed line is indicating the expected rate of convergence and is proportional to  $h^4$ . Bottom left: The spectral condition number versus mesh size  $h$ . The dashed line is proportional to  $h^{-2}$ . Bottom right: the spectral condition number after diagonal scaling versus mesh size  $h$ . The dashed line is proportional to  $h^{-2}$ . Symbols connected with dashed dotted lines have larger constants in the stabilization terms than symbols connected with solid lines and symbols connected with dashed lines, which have the smallest constants in the stabilization terms.

with different constants in the stabilization terms. We see that the magnitude of the error decreases as the stabilization term gets weaker. However, when the stabilization is too weak the condition number is large and roundoff errors dominate.

In Fig. 6 we see that the proposed stabilization and the stabilization term (1.10) resulted in very similar errors; however, the condition number is smaller when the stabilization term (1.10) is added to the weak form. The difference in the condition number decreases after a diagonal scaling of both matrices. We emphasize that all these results depend on the choice of parameters in the stabilization terms and the basis functions used. We have not optimized the parameters in the stabilization terms. For example, using cubic elements and the proposed stabilization term there are six constants, and the parameter  $\gamma$  that one can modify to optimize the magnitude of the error and the condition number.

Next we vary both  $\gamma$  and the constants  $c_{F,j}$  and  $c_{\Gamma,j}$  in the proposed stabilization term, and study the error and the condition number in Figs 7 and 8 for linear and cubic elements, respectively. By decreasing the constants  $c_{F,j}$  and  $c_{\Gamma,j}$  and/or increasing  $\gamma$  we weaken the stabilization. In general, the error decreases while the condition number increases as the stabilization is weakened. The condition number also increases when the stabilization becomes too strong. We see this for the linear elements as well as the cubic elements. A strengthening of the stabilization sometimes decreases

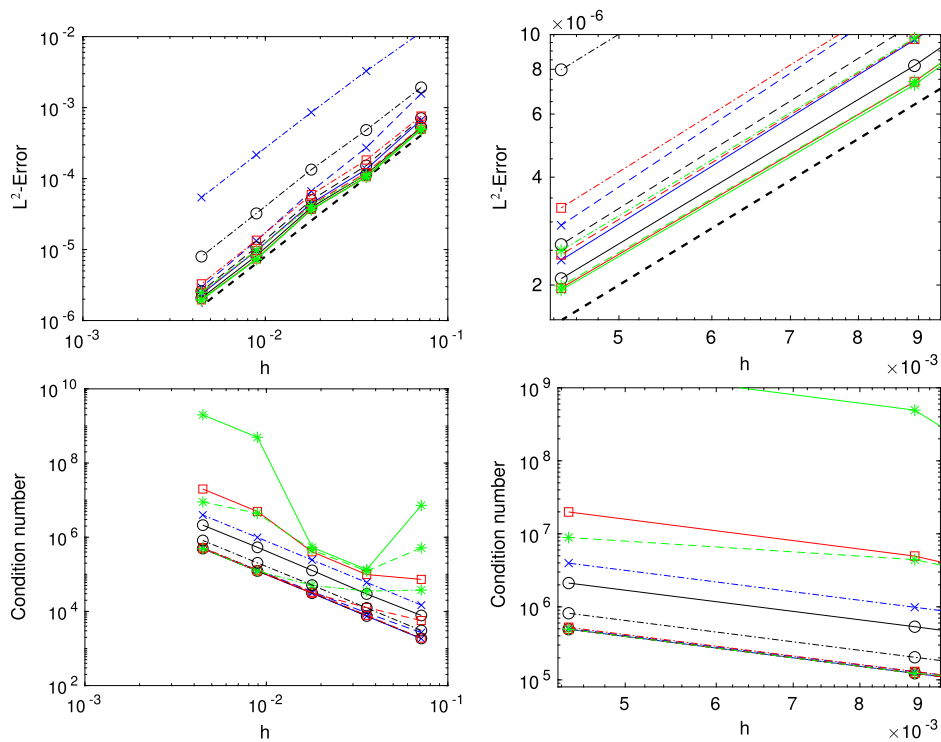


FIG. 7. The Laplace–Beltrami problem: the error and the condition number vs mesh size  $h$  varying the stabilization constants and the parameter  $\gamma$  in (3.7)–(3.9). Linear elements are used, i.e.,  $p = 1$ . Solid lines:  $\gamma = 1$ . Dashed lines:  $\gamma = 1/2$ . Dashed dotted lines:  $\gamma = 0$ .  $c_{F,1} = c_{\Gamma,1} = c$ .  $\times$ :  $c = 1$ .  $\circ$ :  $c = 10^{-1}$ .  $\square$ :  $c = 10^{-2}$ .  $*$ :  $c = 10^{-4}$ . The figures to the right are close ups on part of the figures to the left so that the different curves can be distinguished.

the condition number significantly, but only increases the error slightly. Compare for example the black circles connected with a dashed dotted line,  $\gamma = 0$  in Fig 8, with the black circles connected with a solid line  $\gamma = 1$  in the same figure. By lowering  $\gamma$  the error increases with a factor of 2.5 in this case, but the condition number decreases with a factor 850. Comparing the results in Figs 7 and 8 we see that the choice of parameters in the stabilization term have larger impact on the cubic elements than the linear elements. We also see that, with other choices of parameters than the ones used for the results in Fig. 4, the condition number can scale much better with respect to the polynomial degree.

## 9.2 The mass matrix

We now consider the problem of finding  $u_h \in V_h^p$  such that

$$\int_{\Gamma_h} u_h v_h \, ds_h + s_h(u_h, v_h) = \int_{\Gamma_h} f_h v_h \, ds_h, \quad (9.2)$$

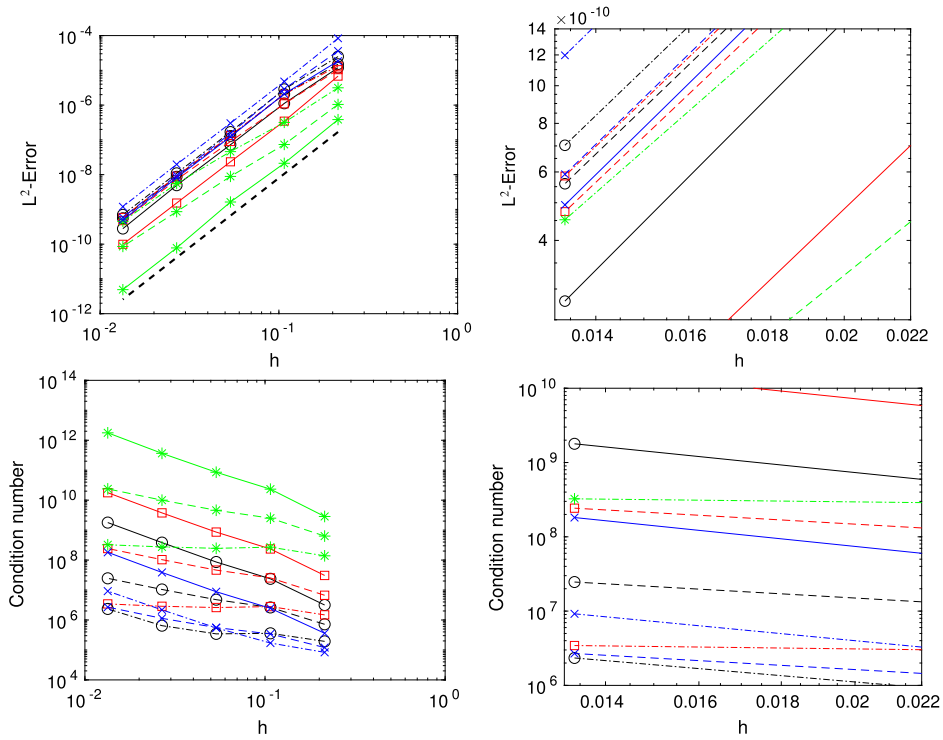


FIG. 8. The Laplace–Beltrami problem: the error and the condition number vs mesh size  $h$  varying the stabilization constants and the parameter  $\gamma$  in (3.7)–(3.9). Cubic elements are used, i.e.,  $p = 3$ . Solid lines:  $\gamma = 1$ . Dashed lines:  $\gamma = 1/2$ . Dashed dotted lines:  $\gamma = 0$ .  $c_{F,1} = c_{\Gamma,1} = c \cdot 0.07$ ,  $c_{F,2} = c_{\Gamma,2} = c \cdot 0.004$ ,  $c_{F,3} = c_{\Gamma,3} = c \cdot 2 \cdot 10^{-4}$ .  $\times$ :  $c = 1$ .  $\circ$ :  $c = 10^{-1}$ .  $\square$ :  $c = 10^{-2}$ .  $*$ :  $c = 10^{-4}$ . The figures to the right are close-ups on part of the figures to the left so that the different curves can be distinguished.

where  $s_h$  is the stabilization term defined in (3.7). The interface  $\Gamma$  is a circle of radius 1 centred at the origin, and we generate a uniform triangular mesh with  $h = h_{x_1} = h_{x_2}$  on the computational domain:  $[-1.5, 1.5] \times [-1.5, 1.5]$ . We let  $f_h$  be

$$f_h = \frac{-(6x_1x_2(x_1^4 - 4x_1^2x_2^2 + x_2^4))}{(x^2 + y^2)^4}. \quad (9.3)$$

We study the condition number of the mass matrix using different degrees of polynomials,  $p$ , and different stabilization terms. Figure 9 shows that the stabilization term we propose results in optimal convergence rates in the  $L^2$ -norm and optimal scaling of the condition number ( $\mathcal{O}(1)$ ) of the mass matrix, as it did for the Laplace–Beltrami operator. As before, the parameters used in (3.7)–(3.9) influence the magnitude of the error and the condition number. For the results shown in Fig. 9 we have used  $c_{F,j} = c_{\Gamma,j} = 0.03 \cdot 20^{-j}$ ,  $j = 1, \dots, p$ . Other choices could give smaller condition numbers without increasing the error significantly, and also give better scaling with respect to the polynomial degree of the condition number.

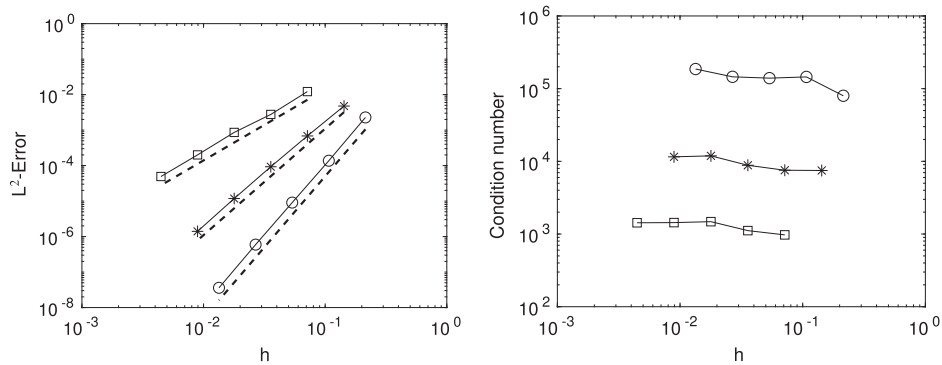


FIG. 9. The mass matrix: the error and condition number versus mesh size  $h$  for different degrees of polynomials,  $p$ , in the space discretization. Squares:  $p = 1$ . Stars:  $p = 2$ . Circles:  $p = 3$ . Left: the error measured in the  $L^2$ -norm vs mesh size  $h$ . The dashed lines are indicating the expected rate of convergence and are proportional to  $h^{p+1}$ . Right: the condition number vs mesh size  $h$ .

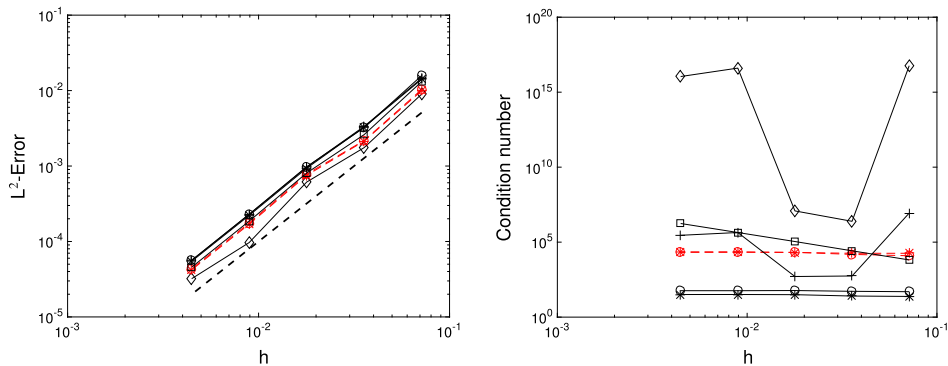


FIG. 10. The mass matrix: the error and condition number vs mesh size  $h$  using different stabilization terms. Linear elements are used, i.e.,  $p = 1$ . Diamonds: no stabilization is added. Circles: the proposed stabilization with  $\gamma = 1$ . Squares: the pure face stabilization. Pluses: pure interface stabilization. Stars: the stabilization term (1.10) with  $\alpha = 1$ . Left: the error measured in the  $L^2$ -norm versus mesh size  $h$ . The dashed line is indicating the expected rate of convergence and is proportional to  $h^2$ . Right: The condition number versus mesh size  $h$ . Symbols connected with solid lines:  $c_{F,j} = 10^{-1}$ ,  $c_{\Gamma,j} = 10^{-1}$  and  $c_{\mathcal{T}} = 10^{-1}$  in (1.10). Symbols connected with dashed lines:  $c_{F,j} = c_{\Gamma,j} = 10^{-4}$  and  $c_{\mathcal{T}} = 10^{-4}$ .

In Figs 10 and 11 we compare different stabilization terms for linear and cubic elements, respectively. The proposed stabilization with  $\gamma = 1$  is compared with using no stabilization, i.e.,  $c_{F,j} = c_{\Gamma,j} = 0$ , a pure face stabilization, i.e.,  $c_{\Gamma,j} = 0$  with  $\gamma = 1$ , a pure interface stabilization, i.e.,  $c_{F,j} = 0$  with  $\gamma = 1$  and the stabilization term (1.10) with  $\alpha = 1$ . In Fig. 10 we also show results with two different constants in the stabilization terms. We see that only the proposed stabilization and the stabilization term (1.10) yield optimal scaling of the condition number ( $\mathcal{O}(1)$ ), and both the error and the condition number are very similar for these two stabilization terms.

For cubic elements we show results with three different constants in the stabilization terms, see Fig. 11. Decreasing the constants in the stabilization leads to smaller error, but larger condition number. When the stabilization is too weak the convergence rate is not optimal.

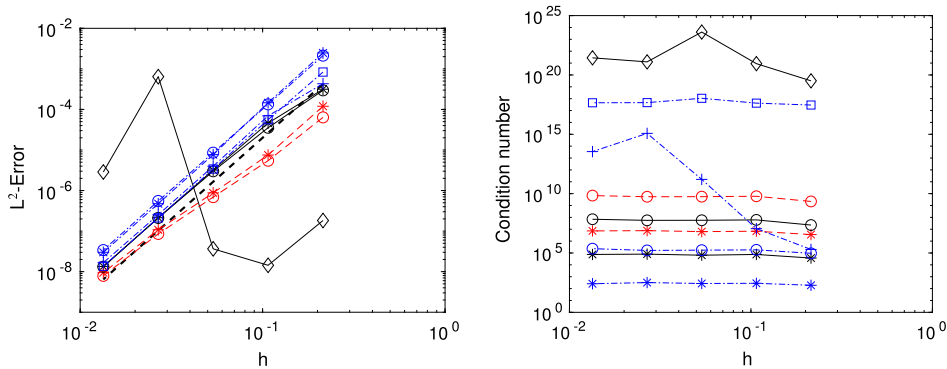


FIG. 11. The mass matrix: the error and condition number vs mesh size  $h$  using different stabilization parameters. Cubic elements are used, i.e.,  $p = 3$ . Diamonds: no stabilization is added. Circles: the proposed stabilization with  $\gamma = 1$ . Squares: the pure face stabilization. Plusses: pure interface stabilization. Stars: the stabilization term (1.10) with  $\alpha = 1$ . Left: the error measured in the  $L^2$ -norm vs mesh size  $h$ . The dashed line is indicating the expected rate of convergence and is proportional to  $h^4$ . Right: the condition number vs mesh size  $h$ . Symbols connected with dashed dotted lines (---) have larger constants in the stabilization terms than symbols connected with solid lines and symbols connected with dashed lines (---) which have the smallest constants in the stabilization terms.

Our observations are similar to what we observed for the Laplace–Beltrami operator. For higher-order elements the linear systems are severely ill-conditioned without stabilization and the behaviour of the error is oscillatory, and the convergence rate will depend on how the geometry cuts the background mesh. We also see that neither the pure face stabilization nor a pure interface stabilization give enough control of the condition number and that we need, as we propose, the combination of them.

### 9.3 The mean curvature vector

We consider the interfaces given by the zero contour of the following level set functions:

$$\phi = \frac{x^2}{0.64} + y^2 - 0.25 \quad \text{Ellipse} \quad (9.4)$$

$$\phi = (z^2 + ((x^2 + y^2)^{1/2} - 1)^2)^{1/2} - 0.5 \quad \text{Torus} \quad (9.5)$$

$$\begin{aligned} \phi = & 0.02 - ((x^2 + y^2 - 0.75)^2 + (z^2 - 1)^2)((y^2 + z^2 - 0.75)^2 \\ & + (x^2 - 1)^2)((z^2 + x^2 - 0.75)^2 + (y^2 - 1)^2). \end{aligned} \quad \text{Deco-cube} \quad (9.6)$$

We generated meshes independently of the position of the given interface. We define the mesh size by  $h = 1/N^{1/d}$  where  $N$  denotes the total number of nodes. We construct an approximate level set function  $\phi_h$  using the nodal interpolant  $\pi_h^1 \phi$  on the background mesh and let the interface be the zero level set of  $\phi_h$ . We use linear Lagrange basis functions. Given the discrete coordinate map  $x_{\Gamma_h} : \Gamma_h \ni x \mapsto x \in \mathbf{R}^d$

we want to find the stabilized discrete mean curvature vector  $H_h \in [V_h^1]^d$  such that

$$(H_h, v_h)_{\Gamma_h} + s_h(H_h, v_h) = (\nabla_{\Gamma_h} x_{\Gamma_h}, \nabla_{\Gamma_h} v_h)_{\Gamma_h}, \quad (9.7)$$

where  $s_h$  is a stabilization term. The torus and the deco-cube are examples from [Hansbo \*et al.\* \(2015\)](#). In [Hansbo \*et al.\* \(2015\)](#) we proved that, using the ghost penalty stabilization (1.8), the method (9.7) is a convergent first-order method with respect to the  $L^2$ -norm. Therefore, we also expect to obtain first-order accurate approximations of the mean curvature vector using the proposed stabilization term.

We compare the approximation from (9.7) with the vector  $H = -(\nabla \cdot n)n$ , where  $n = \frac{\nabla \phi}{|\nabla \phi|}$ . In Fig. 12 we show the error in the  $L^2$ -norm using the proposed stabilization with  $\gamma = 0$ , and the stabilization term (1.10) developed by [Burman \*et al.\* \(2018\)](#) and [Grande \*et al.\* \(2018\)](#) with  $\alpha = -1$ . We emphasize that the analysis by [Hansbo \*et al.\* \(2015\)](#) requires control over the jumps in the normal derivatives across edges on  $\Gamma_h$ , which can be controlled using (1.8), and without appropriate stabilization we don't expect any convergence in  $L^2$ . Note also that we do not expect the element normal gradient stabilization to provide enough stabilization to ensure convergence of the mean curvature vector since some additional tangential control is required, and thus this is a situation where face stabilization comes in naturally. In practice, we note that, depending on the stabilization parameter, we see that we are able to obtain first-order convergence with both stabilization terms. However, we see in the first example, in case of an ellipse, where we have been able to run for finer meshes that the convergence using the stabilization term (1.10) stops after a while for the smaller stabilization parameters and we expect this to happen also for the larger stabilization parameters, but for smaller mesh sizes. The magnitude of the errors are always smaller using the proposed stabilization compared to using the stabilization term (1.10). The magnitude of the error using the pure face stabilization is similar to the results with the proposed stabilization and not shown here. Furthermore, the condition number is better using the proposed stabilization term. We emphasize that, when higher-order elements than linear are used, it is of vital importance to use the proposed stabilization term to control the condition number since the pure face stabilization does not control the condition number, and therefore the error may be dominated by roundoff errors. The proposed stabilization term has been used to compute high order approximations of the mean curvature vector in order to obtain accurate approximations of the surface tension force (see [Frachon & Zahedi, 2019](#)).

In this example, the face stabilization is of importance, and in such cases the proposed stabilization is preferable compared to the stabilization term (1.10).

## 10. Discussion

A strategy to control the condition number of the system matrix resulting from cut finite element discretizations, independently of how elements in the background mesh are cut by the geometry, is to add a stabilization term to the weak formulation. For higher-order elements than linear, the linear systems resulting from cut finite element discretizations without stabilization are extremely ill-conditioned, and it is not obvious how to precondition such linear systems. Therefore, the strategy of adding a stabilization term to the weak form has become a popular alternative to cure this ill-conditioning. The stabilization term most frequently used together with CutFEM is the ghost penalty stabilization developed by [Burman \(2010\)](#), which acts only on the element faces. We have seen in this work that, when higher-order

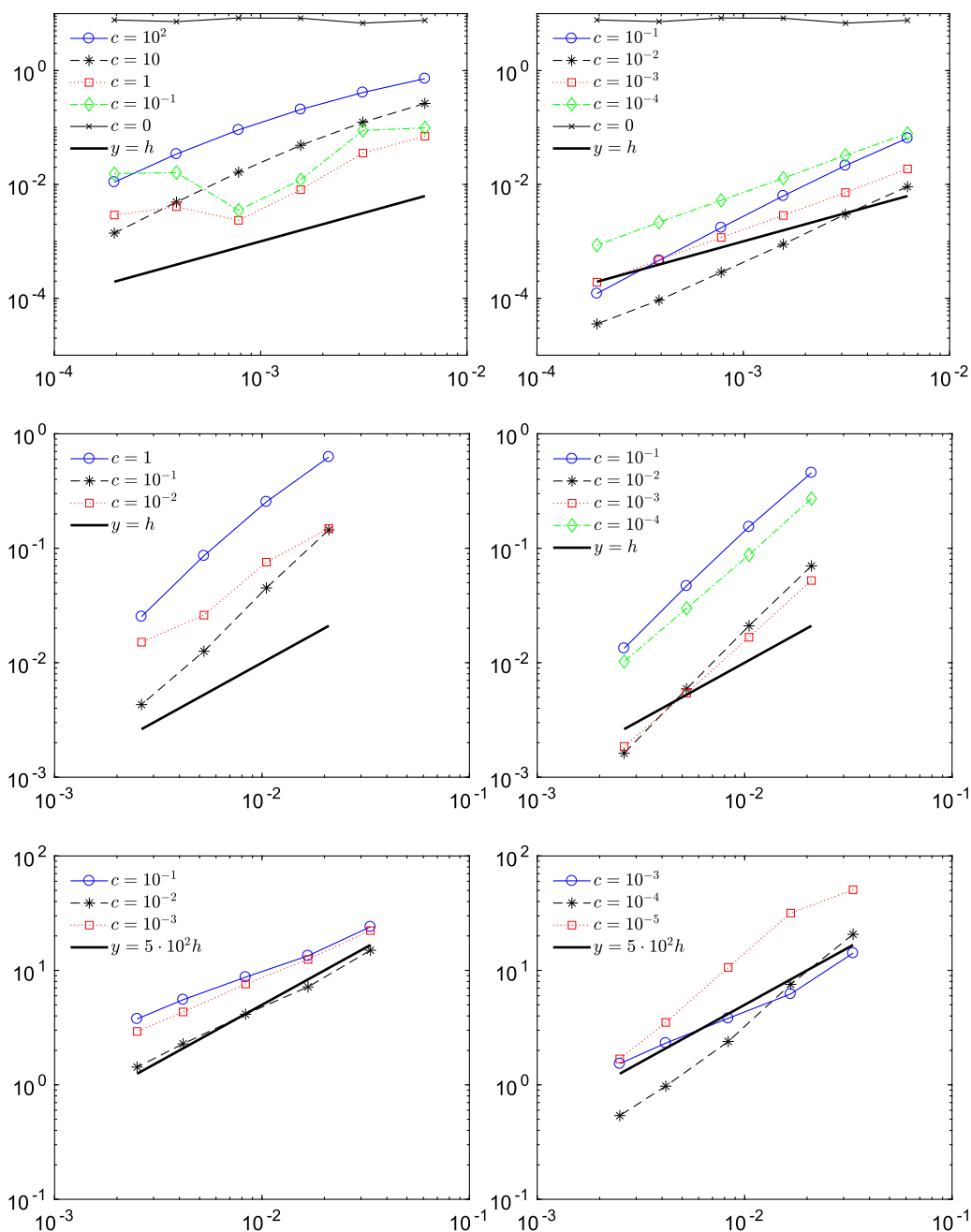


FIG. 12. The mean curvature vector: the error measured in the  $L^2$ -norm versus mesh size  $h$ . Linear elements are used, i.e.,  $p = 1$ . Left: the stabilization term (1.10) developed by Burman *et al.* (2018) and Grande *et al.* (2018) with  $\alpha = -1$  and  $c_F = c$ . Right: the proposed stabilization with  $\gamma = 0$  and  $c_F = c_F = c$ . Top: Ellipse. Middle: Torus. Bottom: Deco-cube.

elements than linear are used in the cut finite element discretizations of surface PDEs, the condition number cannot be controlled by only adding such a face stabilization term. We have proposed a remedy where we keep the face stabilization, but also add a stabilization term acting on the interface/surface, and prove that this new stabilization term controls the condition number for both linear as well as higher-order elements.

As discussed in the introduction and in Section 9 a different stabilization term, the normal gradient stabilization (1.10), which is evaluated on each cut element, has recently been proposed by [Burman \*et al.\* \(2018\)](#) and [Grande \*et al.\* \(2018\)](#). For surface PDEs this stabilization also controls the condition number for linear as well as higher-order elements. This term results in less nonzero entries in the stiffness matrix than the stabilization term we propose, since it is local while the face stabilization term involves neighbouring elements. The implementation of (1.10) is also easier than the stabilization we propose. However, there are a couple of reasons why we choose to keep the ghost penalty term and propose a stabilization term based on this term.

- The element stabilization (1.10) cannot be used in the discretization of bulk problems since it would destroy the convergence order, while the ghost penalty stabilization could be used without destroying the convergence order and with control of the condition number of the linear systems. Many applications include coupled bulk-surface problems and, for those problems, the ghost penalty stabilization is needed. In such applications, the effect of the proposed stabilization on the number of nonzero elements is very small and, since the face stabilization is needed for the bulk problem, the implementation of the proposed stabilization is also easy, as only the stabilization term acting on the surface has to be added.
- The ghost penalty stabilization has shown to be of importance also for other reasons than controlling the condition number. [Burman \*et al.\* \(2018b\)](#) have shown that by adding such a face stabilization a stable discretization for convection dominated problems on surfaces can be obtained, and no other stabilization term such as a streamline diffusion term is needed. [Hansbo \*et al.\* \(2015\)](#) have shown that, with the same face stabilization, a method for computing the mean curvature vector for piecewise linear surfaces based on the Laplace–Beltrami operator enjoys first-order convergence in the  $L^2$ -norm while, in general, without the stabilization convergence, cannot be expected. In Section 9.3 we computed the mean curvature vector of three surfaces, and showed that we obtain better accuracy with the proposed stabilization than with the stabilization (1.10). Our method for computing the mean curvature vector can also be extended to give high order approximations (see [Frachon & Zahedi, 2019](#)). For time-dependent problems on surfaces and coupled-bulk surface problems, adding the ghost penalty stabilization enables to directly approximate space-time integrals in space-time cut finite element formulations using quadrature rules for the integrals over time (see, e.g., [Hansbo \*et al.\*, 2016](#); [Zahedi, 2018](#)). This makes the implementation of space-time CutFEM convenient when higher-order elements are used, in particular, for coupled bulk-surface problems. The proposed stabilization thus also has all these advantages and, in addition, gives control of the condition number for linear as well as higher-order elements.
- The proof of the condition number estimate is simple when the proposed stabilization term is used. The main step in the proof is to show that one covers a nonempty set inside an element by going in the normal direction of the surface from the part of the surface that intersects the element. For the proposed stabilization, one only needs to show that such a nonempty set exist for elements that have a large intersection with the surface. However, to obtain the desired Poincaré inequality with the stabilization term (1.10), one needs to show that such a set exist in each element and, in

particular, that this set is large enough (see, e.g., [Burman \*et al.\*, 2018](#); [Grande \*et al.\*, 2018](#)). We use the simplified structure of the proof to extend our results to the case of general codimension embeddings, see Section 7 where we present the minor modifications of the proof to Lemma 6.5 to verify the generalized version (7.4) of the Poincaré inequality P5.

- We also note that the stabilization term (1.10) requires the computation of the normal vector inside an element. When the interface is implicitly represented by a higher-dimensional function, for example a distance function, this normal vector can be taken to be the normal vector of the level sets of the function defining the interface implicitly. However, when an explicit representation is used for the surface, a normal vector inside an element is not immediately available, but normal vectors on the surface, used in the proposed stabilization, are available.

## Funding

Swedish Foundation for Strategic Research (AM13-0029); Swedish Research Council (2013-4708, 2014-4804); Swedish Research Programme Essence.

## REFERENCES

- BRENNER, S. C. & SCOTT, L. R. (2008) *The Mathematical Theory of Finite Element Methods*. New York: Springer.
- BURMAN, E. (2010) Ghost penalty. *C. R. Acad. Sci. Paris, Ser. I*, **348**, 1217–1220.
- BURMAN, E., CLAUS, S., HANSBO, P., LARSON, M. G. & MASSING, A. (2015a) CutFEM: discretizing geometry and partial differential equations. *Int. J. Numer. Methods Eng.*, **104**, 472–501.
- BURMAN, E., HANSBO, P. & LARSON, M. G. (2015b) A stabilized cut finite element method for partial differential equations on surfaces: the Laplace–Beltrami operator. *Comput. Methods Appl. Mech. Eng.*, **285**, 188–207.
- BURMAN, E., HANSBO, P., LARSON, M. G. & MASSING, A. (2018) Cut finite element methods for partial differential equations on embedded manifolds of arbitrary codimensions. *ESAIM Math. Model. Numer. Anal.*, **52**, 2247–2282.
- BURMAN, E., HANSBO, P., LARSON, M. G., MASSING, A. & ZAHEDI, S. (2016b) Full gradient stabilized cut finite element methods for surface partial differential equations. *Comput. Methods Appl. Mech. Eng.*, **310**, 278–296.
- BURMAN, E., HANSBO, P., LARSON, M. G. & ZAHEDI, S. (2016c) Cut finite element methods for coupled bulk-surface problems. *Numer. Math.*, **133**, 203–231.
- BURMAN, E., HANSBO, P., LARSON, M. G. & MASSING, A. (2017) A cut discontinuous Galerkin method for the Laplace–Beltrami operator. *IMA J. Numer. Anal.*, **37**, 138–169.
- BURMAN, E., HANSBO, P., LARSON, M. G. & LARSSON, K. (2018) Cut finite elements for convection in fractured domains. *Computers & Fluids*.
- BURMAN, E., HANSBO, P., LARSON, M. G. & ZAHEDI, S. (2019) Stabilized CutFEM for the convection problem on surfaces. *Numer. Math.*, **141**, 103–139.
- CENANOVIC, M., HANSBO, P. & LARSON, M. G. (2015) Minimal surface computation using a finite element method on an embedded surface. *Int. J. Numer. Methods Eng.*, **104**, 502–512.
- CENANOVIC, M., HANSBO, P. & LARSON, M. G. (2016) Cut finite element modeling of linear membranes. *Comput. Methods Appl. Mech. Eng.*, **310**, 98–111.
- DECKELNICK, K., ELLIOTT, C. M. & RANNER, T. (2014) Unfitted finite element methods using bulk meshes for surface partial differential equations. *SIAM J. Numer. Anal.*, **52**, 2137–2162.
- DEMLOW, A. & DZIUK, G. (2007) An adaptive finite element method for the Laplace–Beltrami operator on implicitly defined surfaces. *SIAM J. Numer. Anal.*, **45**, 421–442. Electronic.
- DEMLOW, A. & OLSHANSKII, M. A. (2012) An adaptive surface finite element method based on volume meshes. *SIAM J. Numer. Anal.*, **50**, 1624–1647.

- FRACHON, T. & ZAHEDI, S. (2019) A cut finite element method for incompressible two-phase Navier–Stokes flows. *J. Comput. Phys.*, **384**, 77–98.
- FRIES, T.-P. & OMERVIĆ, S. (2016) Higher-order accurate integration of implicit geometries. *Int. J. Numer. Meth. Eng.*, **106**, 323–371.
- GILBARG, D. & TRUDINGER, N. S. (2001) *Elliptic partial differential equations of second order*. Classics in Mathematics. Berlin: Springer, pp. xiv+517. Reprint of the 1998 edition.
- GRANDE, J., LEHRENFELD, C. & REUSKEN, A. (2018) Analysis of a high order trace finite element method for PDEs on level set surfaces. *SIAM J. Numer. Anal.*, **56**, 228–255.
- GROSS, S., OLSHANSKII, M. A. & REUSKEN, A. (2015) A trace finite element method for a class of coupled bulk-interface transport problems. *ESAIM Math. Model. Numer. Anal.*, **49**, 1303–1330.
- HANSBO, A., HANSBO, P. & LARSON, M. G. (2003) A finite element method on composite grids based on Nitsche’s method. *M2AN Math. Model. Numer. Anal.*, **37**, 495–514.
- HANSBO, P., LARSON, M. G. & ZAHEDI, S. (2015) Stabilized finite element approximation of the mean curvature vector on closed surfaces. *SIAM J. Numer. Anal.*, **53**, 1806–1832.
- HANSBO, P., LARSON, M. G. & ZAHEDI, S. (2016) A cut finite element method for coupled bulk-surface problems on time-dependent domains. *Comput. Methods Appl. Mech. Eng.*, **307**, 96–116.
- HANSBO, P., LARSON, M. G. & LARSSON, K. (2017) Cut finite element methods for linear elasticity problems. *Geometrically Unfitted Finite Element Methods and Applications*, vol. 121. Lecture Notes in Computational Science and Engineering. Cham: Springer, pp. 25–63.
- JOHANSSON, A. & LARSON, M. G. (2013) A high order discontinuous Galerkin Nitsche method for elliptic problems with fictitious boundary. *Numer. Math.*, **123**, 607–628.
- KUBLIK, C. & TSAI, R. (2016) Integration over curves and surfaces defined by the closest point mapping. *Res. Math. Sci.*, **3**, 1–17.
- LEHRENFELD, C. (2016) High order unfitted finite element methods on level set domains using isoparametric mappings. *Comput. Methods Appl. Mech. Eng.*, **300**, 716–733.
- MÜLLER, B., KUMMER, F. & OBERLACK, M. (2013) Highly accurate surface and volume integration on implicit domains by means of moment-fitting. *Int. J. Numer. Meth. Eng.*, **96**, 512–528.
- OLSHANSKII, M. A. & REUSKEN, A. (2010) A finite element method for surface PDEs: matrix properties. *Numer. Math.*, **114**, 491–520.
- OLSHANSKII, M. A. & REUSKEN, A. (2014) Error analysis of a space-time finite element method for solving PDEs on evolving surfaces. *SIAM J. Numer. Anal.*, **52**, 2092–2120.
- OLSHANSKII, M. A., REUSKEN, A. & GRANDE, J. (2009) A finite element method for elliptic equations on surfaces. *SIAM J. Numer. Anal.*, **47**, 3339–3358.
- OLSHANSKII, M. A., REUSKEN, A. & XU, X. (2014a) An Eulerian space-time finite element method for diffusion problems on evolving surfaces. *SIAM J. Numer. Anal.*, **52**, 1354–1377.
- OLSHANSKII, M. A., REUSKEN, A. & XU, X. (2014b) A stabilized finite element method for advection-diffusion equations on surfaces. *IMA J. Numer. Anal.*, **34**, 732–758.
- REUSKEN, A. (2015) Analysis of trace finite element methods for surface partial differential equations. *IMA J. Numer. Anal.*, **35**, 1568–1590.
- SAYE, R. (2015) High-order quadrature methods for implicitly defined surfaces and volumes in hyperrectangles. *SIAM J. Sci. Comput.*, **37**, A993–A1019.
- ZAHEDI, S. (2018) A space-time cut finite element method with quadrature in time. *Geometrically Unfitted Finite Element Methods and Applications*, vol. 121. Lecture Notes in Computational Science and Engineering. Cham: Springer, pp. 281–306.