

# **Higher order Discontinuous Galerkin methods for the Laplace-Beltrami problem on unfitted smooth surfaces**

**Bachelor's thesis**

Institute for Numerical and Applied Mathematics  
University of Göttingen

September 2018

Supervisor: Jun-Prof. Dr. Christoph Lehrenfeld  
Second Assessor: Prof. Dr. Gert Lube  
Author: Fabian Heimann

## Abstract

The Poisson equation is a partial differential equation of high significance because it can describe physical problems in electrostatics, gravitational sciences, and elsewhere. One of its generalisations is the Laplace-Beltrami equation, in which the Laplace operator is simply replaced by the respective differential operator on a curved surface. If we want to solve this problem on a simple surface with a suitable right-hand side, we might be able to find an analytic solution. However, on more complicated geometries and with unsuitable right-hand sides, that solution method has its limitations. Instead, numerical methods become relevant.

In this thesis, we will motivate, present, and compare several such numerical methods. Obviously, we cannot give a comprehensive overview but rather focus on a specific subclass of methods often called unfitted Discontinuous Galerkin finite element methods. Roughly, we will discuss four such methods: Firstly, we will consider a second order accurate method presented recently by Burman et al. in [4]. Afterwards, we will slightly modify this method by a technique called hybridisation. From a computational point of view, that means at best to reduce the complexity of the arising linear algebra problem at the expense of more degrees of freedom in the discrete vectors representing the approximate solution. Thirdly, we will apply some ideas presented by Lehrenfeld in [12, 13, 15] to Burman et al.'s method to arrive at a method of higher order accuracy. Namely, we will be able to choose a polynomial order  $k$  such that the numerical error scales with  $h^k$  in the  $H^1$ -norm and  $h^{k+1}$  in the  $L^2$  norm, where  $h$  is the meshsize of the discretisation. Lastly, we will also briefly consider a hybrid variant of this higher order method.

For each method, we will give some motivation, proof—to a different extent—relevant properties, and eventually present some numerical results.

# Contents

<b>1</b>	<b>Introduction</b>	<b>4</b>
1.1	The Laplace-Beltrami problem . . . . .	5
<b>2</b>	<b>A low order method by Burman et al.</b>	<b>8</b>
2.1	Discontinuous Galerkin in the plane . . . . .	8
2.2	Discontinuous Galerkin on an unfitted surface . . . . .	11
2.3	Analysis of the method . . . . .	15
2.4	Numerical examples . . . . .	36
2.5	Hybrid DG in the plane . . . . .	38
2.6	Hybrid DG on an unfitted surface . . . . .	40
<b>3</b>	<b>Higher order methods</b>	<b>46</b>
3.1	The isoparametric mapping . . . . .	47
3.2	Higher order DG method . . . . .	50
3.3	Elements of an analysis of the method . . . . .	53
3.4	Numerical examples . . . . .	55
3.4.1	2D . . . . .	55
3.4.2	3D . . . . .	56
3.5	Higher order Hybrid DG method . . . . .	57
3.6	Numerical examples . . . . .	60
<b>4</b>	<b>Conclusion and outlook</b>	<b>65</b>

# 1 Introduction

The Poisson equation is a very influential example of an elliptic partial differential equation. It is often posed on domains like the unit square  $[0, 1]^2$  in two dimensions or the unit cube  $[0, 1]^3$  in 3D. It reads (with zero Dirichlet boundary conditions)

$$\begin{aligned} -\operatorname{div}(\nabla u) &= f && \text{on } \Omega, \\ u &= 0 && \text{on } \partial\Omega, \end{aligned}$$

where  $\Omega$  could be  $[0, 1]^n$  as mentioned above.

When we regard for example the unit square as a subset of  $\mathbb{R}^3$ , it can be seen as a flat or uncurved surface. This raises the question whether the problem can be generalized to cases of a curved surface  $\Gamma$ . For example, we could try to pose the problem on the sphere as a two-dimensional surface, as the replacement for the unit square. Now the issue arises how we can make sense of the notion of the respective differential operators on the surface. Fortunately, both the gradient as well as the divergence can be introduced with regard to our sufficiently regular surface  $\Gamma$ , symbolically  $\nabla_\Gamma$  and  $\operatorname{div}_\Gamma$ . In terms of these differential operators we can now pose the generalized problem, which is called the Laplace-Beltrami equation

$$-\operatorname{div}_\Gamma(\nabla_\Gamma u) = f \quad \text{on } \Gamma.$$

If  $\Gamma$  is a surface with boundary, we should accordingly add boundary conditions like in the Poisson case. However, in this thesis we will be mostly interested in closed surfaces  $\Gamma$ . There, one additionally requires that the average of the solution should be zero in order to obtain a well-posed problem.

In general, there are (at least) two classes of methods to solve such a problem numerically. One class—which we will call *fitted*—directly meshes the relevant surface. An example of such an approach for the case of a sphere is illustrated in Fig. 1 on the left-hand side.

The other class of methods—which we will call *unfitted*—does not take into consideration the geometry of the surface  $\Gamma$  when the mesh is generated. This means that a sphere would be regarded as a surface in a three dimensional domain which is meshed, like  $\Omega = [0, 1]^3$ . The solution is then approximated on the function space induced by the background mesh. An example of such an approach for the circle embedded in a two-dimension domain is depicted on the right-hand side in Fig. 1.

In this thesis, we will present and compare different such unfitted methods. Most importantly, we will refer to a method introduced by Burman et al. in 2016 in the paper [4]. The

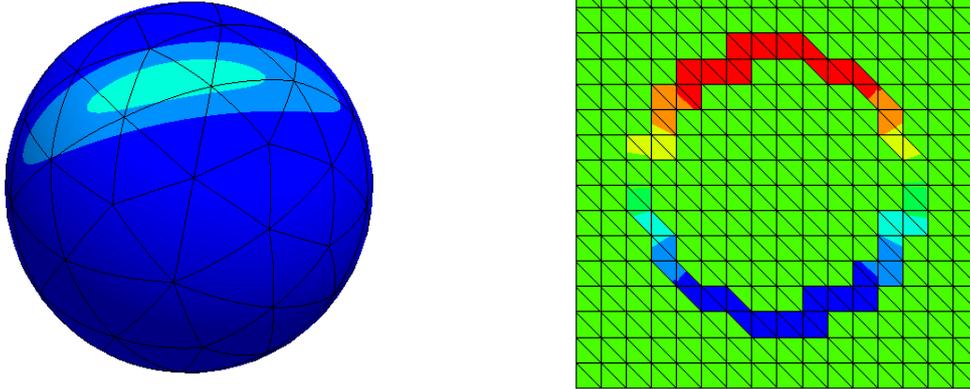


Figure 1: A fitted surface approach (left hand side) applied to a sphere compared with an unfitted (right hand side) approach at a problem on a circle. The coloring of the surface corresponds to a solution field on it.

section following this introduction will motivate and illustrate that method. As it will turn out, the method converges with second order in the  $L^2$  norm. This means that the numerical error in the mentioned norm will asymptotically decrease by a factor  $2^2 = 4$  if the mesh size is halved.

The latter part of this thesis will deal with the question how one could improve on that result. The two main ingredients for doing so will be a higher order approximation of the surface  $\Gamma$ , and a higher order trial and test function space. The main issue is the first point. There we will make use of the isoparametric mapping presented by Lehrenfeld, e.g., in [12, 13, 15]. It also relies on the polygonal approximation of the surface employed by Burman et al. but improves its approximation quality with a specific deformation of the mesh. Furthermore, it will be necessary to add an additional stabilisation in the volume of the active mesh for higher order methods.

Another development that is mostly relevant for the higher order case is the hybridisation of the Discontinuous Galerkin methods. The idea is to decrease the number of couplings between degrees of freedom at the cost of increasing the total number of degrees of freedom: An additional function space on the facets is used to decouple unknowns from directly neighbouring elements of the trial and test functions at the boundaries of each element. We will discuss those methods in the respective subsections.

## 1.1 The Laplace-Beltrami problem

Let us now in some more detail introduce the problem we aim to solve in this thesis, namely the Laplace-Beltrami problem. The exposition will follow [4] and to some extent fit the needs

of the forthcoming analysis.

Let  $\Gamma$  denote a compact and oriented  $C^k$  hypersurface (for  $k \geq 2$ ) without boundary. It should be embedded in  $\mathbb{R}^d$  and equipped with a normal field  $n: \Gamma \rightarrow \mathbb{R}^d$  of class  $C^{k-1}$ . We define neighbourhoods of  $\Gamma$  with  $\delta$  as a variable as

$$U_\delta(\Gamma) = \{x \in \mathbb{R}^d \mid \text{dist}(x, \Gamma) < \delta\}.$$

Then it can be shown that there exists a unique signed distance function  $\rho \in C^k(U_{\delta_0}(\Gamma))$  induced by the normal field  $n$  (see [6, 4]). On the neighbourhood, there furthermore exists a closest point projection  $p: U_{\delta_0} \rightarrow \Gamma$  determined by

$$p(x) = x - \rho(x)n(p(x)),$$

such that a point  $x \in U_{\delta_0}$  is mapped to the unique point  $p(x) \in \Gamma$  satisfying  $|p(x) - x| = \text{dist}(x, \Gamma)$  (see Fig. 2). With the help of  $p$ , we can extend any function on  $\Gamma$  to the neighbourhood  $U_{\delta_0}$  by setting

$$u^e(x) = u \circ p(x).$$

This can especially be done with the  $n$ , whose extension  $n^e$  we will also write as  $n_\Gamma$ .

If there exists a  $\tilde{\Gamma} \subset U_{\delta_0}(\Gamma)$  such that  $p|_{\tilde{\Gamma}}: \tilde{\Gamma} \rightarrow \Gamma$  is a bijection, then we also have the inverse function  $p|_{\tilde{\Gamma}}^{-1}: \Gamma \rightarrow \tilde{\Gamma}$ . Given a function  $w$  defined on  $\tilde{\Gamma}$  we can lift it to  $\Gamma$ , or construct a function  $w^l$  defined on  $\Gamma$ :

$$w^l := w \circ p|_{\tilde{\Gamma}}^{-1}.$$

This definition then satisfies

$$(w^l(x))^e = w^l \circ p = w \quad \text{on} \quad \tilde{\Gamma}.$$

For a function space  $V$  of functions on  $\tilde{\Gamma}$ , we set

$$V^l = \{w^l \mid w \in V\}.$$

Also, for a function space  $V$  of functions on  $\Gamma$ , we define

$$V^e = \{w^e \mid w \in V\}.$$

Now we come to the differential operator  $\nabla_\Gamma$ . A function  $u: \Gamma \rightarrow \mathbb{R}$  is of class  $C^l(\Gamma)$ ,  $l \leq k$ ,

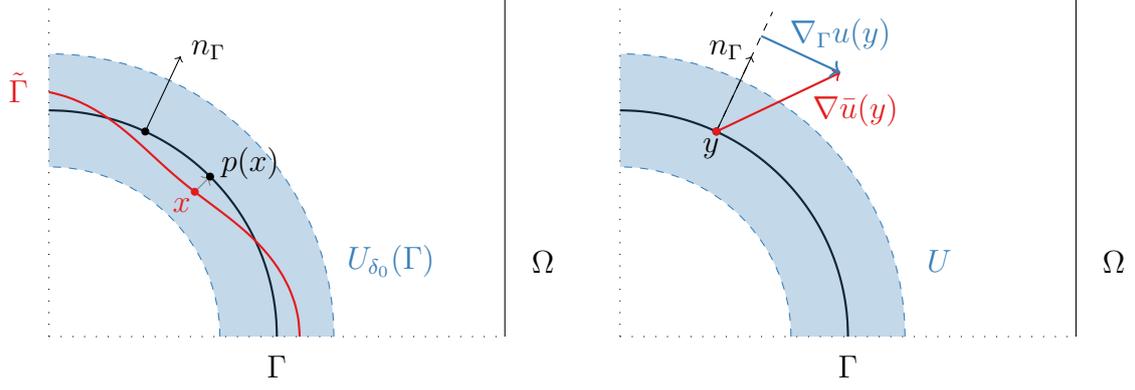


Figure 2: Sketch of the domain  $\Omega$ , the surface  $\Gamma$ ,  $U_{\delta_0}(\Gamma)$ , and several other entities. For convenience, only a part of the whole region  $\Omega$  is depicted (therefore the dotted boundary).

if there is an extension  $\bar{u} \in C^l(U)$  with  $\bar{u}|_{\Gamma} = u$  for a neighbourhood  $U$  of  $\Gamma$ . We can take the usual  $\mathbb{R}^d$  gradient of  $\bar{u}$  and exploit that to define

$$\nabla_{\Gamma} u = P_{\Gamma} \nabla \bar{u}, \quad P_{\Gamma} = I - n_{\Gamma} \otimes n_{\Gamma}.$$

It can be shown that this defined quantity (evaluated on  $\Gamma$ ) is independent of the neighbourhood  $U$  and independent of the extension  $\bar{u}$  of  $u$ . For convenience, we could for example consider  $\bar{u}(x) := u \circ p(x)$ .

Given this differential operator, we introduce Sobolev spaces on  $\Gamma$  by dint of integrals on  $\Gamma$ ,  $\|w\|_{\Gamma}^2 = (w, w)_{\Gamma}$  ( $(v, w)_{\Gamma} = \int_{\Gamma} v w$ ). The space  $H^m(\Gamma)$  is the subspace of  $L^2$  functions on  $\Gamma$  with norm

$$\|w\|_{m,\Gamma}^2 = \sum_{k=0}^m \|D_{\Gamma}^{P,k} w\|_{\Gamma}^2$$

finite, where  $D_{\Gamma}^{0,k} w = w$ ,  $D_{\Gamma}^{1,k} w = \nabla_{\Gamma} w$ , etc.

The Laplace-Beltrami problem in its strong form is then defined as follows: Find  $u: \Gamma \rightarrow \mathbb{R}$  such that

$$-\Delta_{\Gamma} u = f \quad \text{on } \Gamma \tag{1}$$

for a  $f \in L^2(\Gamma)$  with zero average and  $\Delta_{\Gamma} := \text{div}_{\Gamma} \nabla_{\Gamma}$ . Note that we would have to explain the meaning of the divergence with regard to  $\Gamma$  here.<sup>1</sup> However, we mostly use the strong form of the problem to motivate the following weak form: Find  $u \in H^1(\Gamma)/\mathbb{R}$  such that

$$a(u, v) = (\nabla_{\Gamma} u, \nabla_{\Gamma} v)_{\Gamma} = l(v) = (f, v)_{\Gamma} \quad \forall v \in H^1(\Gamma)/\mathbb{R}. \tag{2}$$

<sup>1</sup> $\text{div}_{\Gamma} = \text{tr}(\nabla_{\Gamma} \cdot)$  where the  $i$ -th row of  $\nabla_{\Gamma}$  corresponds to the surface gradient of the  $i$ -th component

The unique solvability of this problem follows as usual from a Poincaré-type inequality and the Lax-Milgram Lemma.<sup>2</sup> Furthermore, by the Lax-Milgram lemma, for smooth surfaces our solution  $u$  (measured in terms of the  $H^2$ -norm) can be bound by the surface norm of the right-hand side  $f$ , namely

$$\|u\|_{2,\Gamma} \lesssim \|f\|_{\Gamma}. \quad [4, \text{Eq. (2.11)}] \quad (3)$$

## 2 A low order method by Burman et al.

This section is devoted to the method presented in the paper [4]. Because of the tight connection to this paper, we will generally follow Burman et al. in naming the lemmata/theorems etc. and add a B. in front to remind the reader of the origin of the lemma/theorem etc. Sometimes, we want to prove intermediate steps a little more carefully, and introduce the names ending with  $*i$  for  $i \in \mathbb{N}$  for those results. In naming constants, we will use the name of the lemma and give constants not associated to a specific lemma a name with the number 0 at that place.

### 2.1 Discontinuous Galerkin in the plane

Let us first of all motivate the method of Burman et al. by reviewing several influential methods which we illustrate at the Poisson equation. So—to be concrete—imagine we want to solve the following problem

$$\begin{aligned} -\Delta u &= f && \text{on } \Omega, \\ u &= 0 && \text{on } \partial\Omega \end{aligned}$$

for some sufficiently regular region  $\Omega$  (a concrete example we will often use is  $\Omega = [0, 1]^d$ ,  $d = 2, 3$ ) with the boundary  $\partial\Omega$ . Furthermore,  $f$  should be a right-hand side sufficiently smooth.<sup>3</sup>

The standard Finite Element method to solve such a problem—the continuous Galerkin method (CG)—goes as follows: First of all, we restrict our attention to a small subspace of the infinite dimensional Sobolev space in which the actual solution to the problem can be found, as standard analysis results yield. Namely, we only consider the subspace of functions which are

<sup>2</sup>That of course is a bit brief here. We will discuss those techniques in detail for the discrete case later on.

<sup>3</sup>We restrict our attention here to zero Dirichlet boundary conditions. However, that is not a severe restriction since the problem with arbitrary Dirichlet boundary conditions can be reduced to our problem with a standard procedure. See, e.g., [14, p. 46].

polynomials on each element  $T \in \mathcal{T}_h$ , where  $\mathcal{T}_h$  is a triangulation of  $\Omega$ , i.e.,  $\Omega = \cup\{T \in \mathcal{T}_h\}$ . Technically, we can write

$$V_h^{\text{cont}} = \{u \in H^1(\Omega) \mid u|_T \in \mathcal{P}_k(T) \forall T \in \mathcal{T}_h\} \subset H^1(\Omega).$$

This means that a function from this solution space is continuous along an edge of two adjacent  $T \in \mathcal{T}_h$ . Imposing the Dirichlet boundary conditions leads to

$$V_{h,0}^{\text{cont}} = \{u \in H^1(\Omega) \mid u|_{\partial\Omega} = 0 \text{ and } u|_T \in \mathcal{P}_k(T) \forall T \in \mathcal{T}_h\} \subset H^1(\Omega).$$

After choosing that solution space, we set up the weak problem by multiplying the strong form of the PDE with a test function  $v$  (which ultimately should come from  $V_{h,0}^{\text{cont}}$  as well) and integrating over the domain  $\Omega$ . Then we have

$$-\int_{\Omega} \text{div}(\nabla u)v \, \text{dr} = -\int_{\partial\Omega} v \frac{\partial u}{\partial n} \, \text{dS} + \int_{\Omega} \nabla u \cdot \nabla v \, \text{dr}$$

due to Green's first theorem. This leads to the following weak form of the PDE problem:

Find  $u \in V = V_{h,0}^{\text{cont}}$  such that

$$a(u, v) = f(v) \quad \text{holds for all } v \in V,$$

where

$$a(u, v) := \int_{\Omega} \nabla u \cdot \nabla v \, \text{dr} - \int_{\partial\Omega} \underbrace{v}_{=0} \frac{\partial u}{\partial n} \, \text{dS} = \int_{\Omega} \nabla u \cdot \nabla v \, \text{dr}, \quad f(v) := \int_{\Omega} f \cdot v \, \text{dr}.$$

From the standard theory of those methods, we obtain the well-posedness of this problem and the result that the approximate solution will indeed converge against the actual solution of the problem (see, e.g., [14, p. 45, 64]).

**From CG to DG** One might also consider a different choice for the finite dimensional space of test and solution functions. Namely, one might drop the requirement that the functions should be continuous along different  $T \in \mathcal{T}_h$ , leading to

$$V_h^{\text{discont}} = \{u \in L^2(\Omega) \mid u|_T \in \mathcal{P}_k(T) \forall T \in \mathcal{T}_h\}.$$

On the first glance, this might seem a bit like a waste of degrees of freedom, since the actual solution will be continuous, and therefore can't be approximated significantly better with the

discontinuous function space. However, the so called discontinuous Galerkin methods, which one obtains from such a discretisation, have several potential advantages compared to CG methods, e.g., in the context of convection-dominated problems, when one tries to parallelise the numerical calculations, or for simultaneous mesh and polynomial order refinements. For an overview of several Discontinuous Galerkin methods and their (dis)advantages, I would like to refer the reader to [1, 5, 9].

When we again apply the finite element machinery, i.e., multiply with a test function and integrate over  $\Omega$ , we now arrive at a different situation in the following sense: Since the functions  $u$  and  $v$  now are not continuous any more we can only apply Green's first theorem on each element  $T$  on its own. This leads to

$$-\int_{\Omega} \operatorname{div}(\nabla u)v \, \mathrm{d}\mathbf{r} = -\sum_{T \in \mathcal{T}_h} \int_T \operatorname{div}(\nabla u)v \, \mathrm{d}\mathbf{r} = -\sum_{T \in \mathcal{T}_h} \int_{\partial T} v \frac{\partial u}{\partial n} \, \mathrm{d}S - \int_T \nabla u \cdot \nabla v \, \mathrm{d}\mathbf{r}.$$

Each boundary of one element  $T$ ,  $\partial T$  now consists of several facets  $F_1, F_2, \dots$  such that  $\partial T = F_1 \cup F_2 \cup \dots$ . When we collect those facets of the mesh in sets  $\mathcal{F}^{int}$  and  $\mathcal{F}^{ext}$  corresponding to facets located in the interior and on the boundary (exterior), respectively, we can rewrite the integral over  $\sum_T \int_{\partial T}$  as follows

$$\sum_T \int_{\partial T} v \frac{\partial u}{\partial n} \, \mathrm{d}S = \sum_{F \in \mathcal{F}^{ext}} \int_F v \frac{\partial u}{\partial n} \, \mathrm{d}S + \sum_{F \in \mathcal{F}^{int}} \int_F v_{left} \frac{\partial u_{left}}{\partial n_{left}} + v_{right} \frac{\partial u_{right}}{\partial n_{right}},$$

where we have exploited the fact that each inner facet appears twice in the sum over all  $\partial T$  and all outer facet only once. With  $f_{left}/f_{right}$  for some function  $f$  we refer to the function on the element left/right to  $F$  for a fixed orientation of  $F$ .

It is  $n_{left} = -n_{right}$ , so we can write  $\frac{\partial u_{right}}{\partial n_{right}}$  as  $-\frac{\partial u_{right}}{\partial n_{left}}$ . This leads to

$$\sum_T \int_{\partial T} v \frac{\partial u}{\partial n} \, \mathrm{d}S = \sum_{F \in \mathcal{F}^{ext}} \int_F v \frac{\partial u}{\partial n} \, \mathrm{d}S + \sum_{F \in \mathcal{F}^{int}} \int_F v_{left} \frac{\partial u_{left}}{\partial n_{left}} - v_{right} \frac{\partial u_{right}}{\partial n_{left}}$$

Since the actual solution is continuously differentiable along surfaces, we replace both  $\frac{\partial u_{left}}{\partial n_{left}}$  and  $\frac{\partial u_{right}}{\partial n_{left}}$  with an average, namely

$$\left\{ \frac{\partial u}{\partial n} \right\} = \frac{1}{2} \left( \frac{\partial u_{left}}{\partial n_{left}} + \frac{\partial u_{right}}{\partial n_{left}} \right),$$

which together with the abbreviation  $[v] = v_{left} - v_{right}$  leads to

$$\sum_T \int_{\partial T} v \frac{\partial u}{\partial n} dS = \sum_{F \in \mathcal{F}^{ext}} \int_F v \frac{\partial u}{\partial n} dS + \sum_{F \in \mathcal{F}^{int}} \int_F \left\{ \frac{\partial u}{\partial n} \right\} [v] dS.$$

Since  $[u]$  vanishes for the exact solution, we add two consistent terms for symmetry and stability to the bilinear form, leading to

$$\sum_T \int_{\partial T} v \frac{\partial u}{\partial n} dS = \sum_{F \in \mathcal{F}^{ext}} \int_F v \frac{\partial u}{\partial n} dS + \sum_{F \in \mathcal{F}^{int}} \int_F \left\{ \frac{\partial u}{\partial n} \right\} [v] + \left\{ \frac{\partial v}{\partial n} \right\} [u] - \frac{\alpha p^2}{h} [u][v] dS.$$

Regarding the exterior facets, we want to impose zero Dirichlet boundary condition on the solution  $u$ . Therefore we can replace the first sum analogously with

$$\sum_{F \in \mathcal{F}^{ext}} \int_F v \frac{\partial u}{\partial n} + u \frac{\partial v}{\partial n} + \frac{\alpha p^2}{h} uv dS.$$

Putting all this together, we arrive at

$$\begin{aligned} a(u, v) &= \sum_{T \in \mathcal{T}_h} \int_T \nabla u \cdot \nabla v dx + \sum_{F \in \mathcal{F}^{ext}} \left( - \int_F v \frac{\partial u}{\partial n} dS - \int_F u \frac{\partial v}{\partial n} dS + \frac{\alpha p^2}{h} \int_F uv dS \right) \\ &+ \sum_{F \in \mathcal{F}^{int}} \left( \int_F - \left\{ \frac{\partial u}{\partial n} \right\} [v] dS + \int_F - \left\{ \frac{\partial v}{\partial n} \right\} [u] dS + \frac{\alpha k^2}{h} \int_F [u][v] dS \right). \end{aligned} \quad (4)$$

The motivation of adding the symmetry and stability terms can be seen in the analysis of the method with the respective bilinear form: The stability term will be used in the coercivity proof to absorb the contribution from the other facet summands. An example of such an estimate will appear in the proof of Proposition B.5.5. The symmetry term on the other hand ensures adjoint consistency, which is important in the  $L^2$ -norm error estimates. We will later on skip that part of the proof and refer the reader to [1] for an example of such a proof.

This is a very common discretisation of the Poisson problem called (symmetric) interior penalty discontinuous Galerkin method. The method by Burman et al. can be regarded as an application of this idea to the case of a surface integration region within the domain  $\Omega$ . This is what we will elaborate on next.

## 2.2 Discontinuous Galerkin on an unfitted surface

In order to be able to transfer the method introduced in the last subsection to the unfitted surface, we first of all need to introduce the geometric entities and function spaces.

Let  $\Omega \subset \mathbb{R}^d$  (in this thesis, we consider  $d = 2, 3$ ) be a background domain with regard to the surface  $\Gamma$ ,  $\Omega \supset U_{\delta_0}(\Gamma)$ . Furthermore, assume we are given a quasi-uniform triangulation (into triangles in 2D and tetrahedra in 3D)  $\tilde{\mathcal{T}}_h$  of  $\Omega$ , such that  $\Omega = \bigcup\{T \mid T \in \tilde{\mathcal{T}}_h\}$ . We assume that the triangulation is such that the meshsize  $h$  at each point only deviates by a threshold factor like 0.5 from  $h_{max}$ . Throughout the analysis, we will sometimes operate with the function  $h$  as if it was globally constant. That would be strictly speaking only allowed for a uniform meshsize, but the derivation could be easily repaired by inserting the threshold factor, though at the cost of readability.

Let us follow the convention that the boundaries of the elements are actually contained in each  $T \in \tilde{\mathcal{T}}_h$  such that  $T_1 \cap T_2 \neq \emptyset$  for adjacent  $T_1 \neq T_2 \in \tilde{\mathcal{T}}_h$ . The discrete surface is then defined in terms of the piecewise linear interpolation of the signed distance function  $\rho$ ,  $\rho_h$ , namely as

$$\Gamma_h = \{x \in \Omega \mid \rho_h(x) = 0\}.$$

Note that due to the fact that  $\rho_h$  was introduced as an element-wise linear interpolation,  $\Gamma_h$  is a polygonal. For that polygonal, a normal vector can be obtained with the help of the interpolant of the signed distance function  $\rho_h$ :

$$n_h = \frac{\nabla \rho_h}{\|\nabla \rho_h\|}. \quad (5)$$

For the method to be well-posed, we have to make the following assumptions, which indeed are satisfied if  $\rho_h$  is chosen as we did and  $h$  is sufficiently small.

**Assumption B.A** The following properties hold:

- (1)  $\Gamma_h \subset U_{\delta_0}(\Gamma)$ .
- (2) The closest point projection  $p: \Gamma_h \rightarrow \Gamma$  is a bijection for  $0 < h \leq h_0$ .
- (3) It holds

$$\|\rho\|_{L^\infty(\Gamma_h)} \leq C^{B.2.0.1} \cdot h^2, \quad \|n^e - n_h\|_{L^\infty(\Gamma_h)} \leq C^{B.2.0.2} h.$$

In terms of the discrete surface we can now define the discrete integration regions, which will replace the elements and facets of the discontinuous Galerkin method in the plane.

First of all, let us introduce the active background mesh  $\mathcal{T}_h$  as

$$\mathcal{T}_h := \{T \in \tilde{\mathcal{T}}_h \mid T \cap \Gamma_h \neq \emptyset\}.$$

The interior facets of these simplices are then

$$\mathcal{F}_h = \{F = T^+ \cap T^- \mid T^+, T^- \in \mathcal{T}_h, T^+ \neq T^-, \text{ and } \text{meas}_{d-1}(T^+ \cap T^-) > 0\}.$$

Those facets come with the usual normal vectors  $n_F^+, n_F^-$  (corresponding to the neighboring elements  $T_\pm$ ) of the background mesh. Especially, it holds  $n_F^+ = -n_F^-$ . Next, we introduce the approximation of the surface as

$$\mathcal{K}_h = \{K = \Gamma_h \cap T \mid T \in \mathcal{T}_h\},$$

and the corresponding set of edges as

$$\mathcal{E}_h = \{E = F \cap \Gamma_h \mid F \in \mathcal{F}_h\}.$$

We use the notation  $\partial\mathcal{K}_h$  to refer to  $\{\partial K \mid K \in \mathcal{K}_h\}$ .

For the  $K \in \mathcal{K}_h$ , the associated normal vector is the  $n_h$  as defined in (5). That discrete normal vector induces a modified projection operator  $P_{\Gamma_h}$  and a gradient  $\nabla_{\Gamma_h}$ , analogously to  $P_\Gamma$  and  $\nabla_\Gamma$ :

$$\nabla_{\Gamma_h} u = P_{\Gamma_h} \nabla \bar{u} \quad \text{and} \quad P_{\Gamma_h} = I - n_h \otimes n_h.$$

Note that we are not strict regarding the type of functions to which  $\nabla_{\Gamma_h}$  can be applied (or more precisely, regarding their domain of definition).<sup>4</sup>

In the case of  $E \in \mathcal{E}_h$ , the choice of normal vectors is not canonical, since we are dealing with a subset of  $\Omega$  of codimension 2 (and not 1 as previously). We choose to assign to each edge  $E \in \mathcal{E}_h$  two normals  $n_E^+$  and  $n_E^-$  (with  $n_E^+ \neq n_E^-$  typically) which are coplanar to the respective  $K$  and orthogonal to  $E$  (the latter criterion is trivial for  $d = 2$ ). An illustration of the different normal and conormal vectors is given in Fig. 3.

The active background mesh induces a discrete neighbourhood of  $\Gamma_h$ , namely

$$N_h = \bigcup \{T \mid T \in \mathcal{T}_h\}. \tag{6}$$

For the reader's convenience, all those regions are depicted in an example in Fig. 4.

On those geometric entities, we can now introduce the relevant approximation function

---

<sup>4</sup>“Officially” the definition is one for functions  $u$  with domain of definition  $\Gamma_h$ , which are then extended (as in the discussion before the definition of  $\nabla_\Gamma$ ) to functions  $\bar{u}$  on a neighbourhood. However, we will also apply  $\nabla_{\Gamma_h}$  to functions  $u$ , whose domain of definition already is a neighbourhood of  $\Gamma_h$ . There  $\nabla_{\Gamma_h} u = P_{\Gamma_h} \nabla u$ .

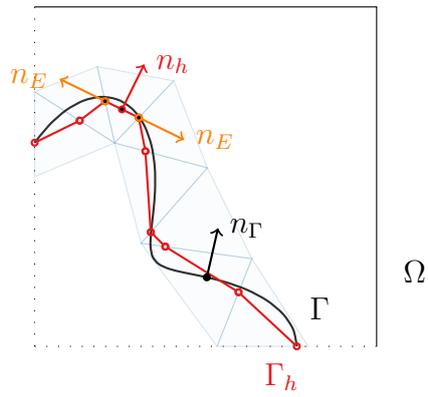


Figure 3: Illustration of the different discrete and continuous (co)normals.

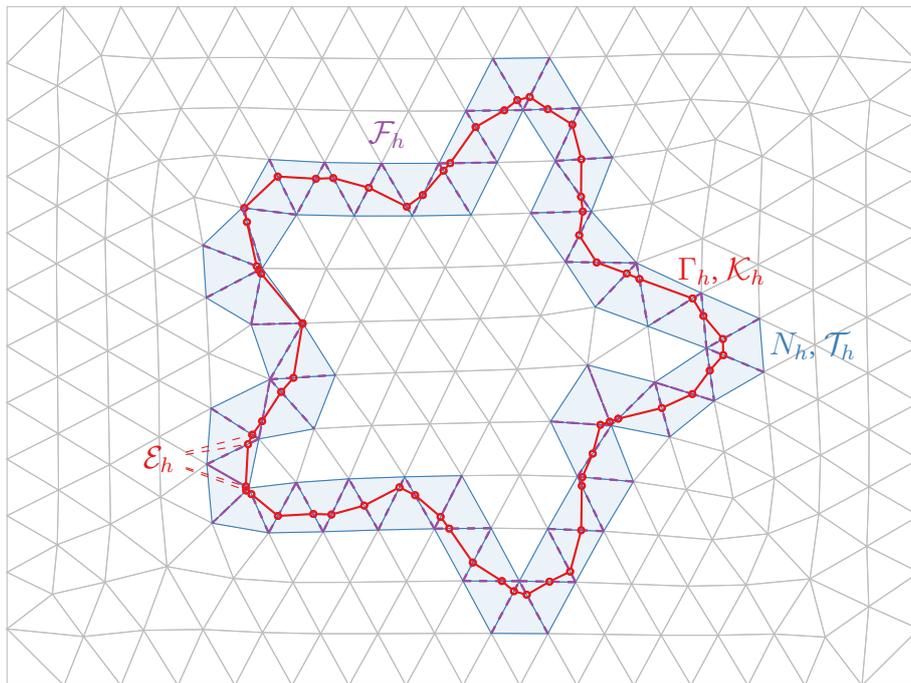


Figure 4: Illustration of the different discrete regions.

space. We define

$$V_h = \bigoplus_{T \in \mathcal{T}_h} P_1(T)$$

and consider the subspace of functions with zero average

$$V_{h,0} = \left\{ v \in V_h \mid \int_{\Gamma_h} v = 0 \right\}.$$

Note that this definition is an example of a discontinuous function space, since no continuity constraint is imposed.

Now we are able to transfer the DG method in the plane to the discrete version of the surface,  $\Gamma_h$ . The “volume” elements regarding  $\Gamma_h$  now are the  $K \in \mathcal{K}_h$ , while the edges  $E \in \mathcal{E}$  play the role of the facets in the standard interior penalty DG method. This motivates the following bilinear form

$$\begin{aligned} a_h(u, v) &= \sum_{K \in \mathcal{K}_h} \int_K \nabla_{\Gamma_h} u \cdot \nabla_{\Gamma_h} v - \sum_{E \in \mathcal{E}_h} \int_E \{n_E \cdot \nabla u\} [v] - \sum_{E \in \mathcal{E}_h} \int_E \{n_E \cdot \nabla v\} [u] \\ &\quad + \frac{\beta_E}{h} \sum_{E \in \mathcal{E}_h} \int_E [u][v]. \end{aligned}$$

Note that the explicit dependence on the polynomial order in  $\beta_E$  was dropped.

In order to ensure stability of the method also for ill-posed cut configurations, which are configurations where  $|\Gamma \cap T|$  is tiny compared to  $|T|$ , a further stabilisation needs to be added, namely

$$j_h(u, v) = \sum_{F \in \mathcal{F}_h} \int_F \frac{\beta_F}{h^2} [u][v] + \int_F \gamma (n_F \cdot [\nabla u]) \cdot (n_F \cdot [\nabla v])$$

This is similar to the Ghost-penalty stabilisation used in unfitted Nitsche methods for enforcing boundary conditions. [2] The right hand side is

$$l_h(v) = \int_{\Gamma_h} f^e \cdot v.$$

Then the discrete problem takes the form: Find  $u_h \in V_{h,0}$  such that for all  $v_h \in V_{h,0}$

$$A_h(u_h, v_h) = l_h(v_h) \quad \text{with} \quad A_h(u_h, v_h) = a_h(u_h, v_h) + j_h(u_h, v_h). \quad (7)$$

## 2.3 Analysis of the method

The analysis of the method proceeds in several steps. I cannot give all the details of the paper here, but would like to summarise what to me seem to be the most important points.

The starting point for the analysis of the method are the suiting norms. First of all, let us define the following discrete energy norm according to the bilinear form in (7):

$$\|v\|_h^2 := \sum_{K \in \mathcal{K}_h} \|\nabla_{\Gamma_h} v\|_K^2 + \sum_{E \in \mathcal{E}_h} \|h^{-1/2}[v]\|_E^2 + \sum_{F \in \mathcal{F}_h} \|h^{-1}[v]\|_F^2 + \|n_F \cdot [\nabla v]\|_F^2. \quad (8)$$

Furthermore, for a function  $v \in (H^2(\Gamma))^e$  or  $v \in V_h$ , we define

$$\|v\|_{*,h}^2 := \|v\|_h^2 + \sum_{K \in \mathcal{K}_h} \|h^{1/2} \nabla_{\Gamma_h} v\|_{\partial K}^2. \quad (9)$$

Note that by construction the sum over all integrals over  $\partial K$  for all  $K \in \mathcal{K}_h$  is just two times the integral over all  $E \in \mathcal{E}_h$ .

**Remark** It would be possible to insert the edge-normal into the last part of the  $*, h$ -norm, leading to

$$\|v\|_h^2 + \sum_{K \in \mathcal{K}_h} \|h^{1/2} n_E \nabla_{\Gamma_h} v\|_{\partial K}^2.$$

That would make the proof of Proposition B.5.5 a bit shorter, namely in Equation (13). However, we want to follow the definition of Burman et al. here. In the high order case however, we will go this alternative way.

The facet part of the discrete norm is abbreviated as

$$\|v\|_{\mathcal{F}_h}^2 := \sum_{F \in \mathcal{F}_h} \|h^{-1}[v]\|_F^2 + \|n_F \cdot [\nabla v]\|_F^2.$$

Another important ingredient of the analysis of the method is the technique of lifting the bilinear form and the norms. In order to do so, we first of all lift the discrete approximations of the surface with the closest point projection, namely

$$\mathcal{K}_h^l := \{p(K) \mid K \in \mathcal{K}_h\}$$

This is a meaningful definition because by construction  $K \subset \Gamma_h$  and by assumption  $\Gamma_h \subset U_{\delta_0}$ , so  $K \subset U_{\delta_0}$ . The closest point projection  $p$  was just defined on  $U_{\delta_0}$ . By assumption the image of  $\Gamma_h$  under  $p$  is  $\Gamma$ , and therefore  $\Gamma = \bigcup_{K^l \in \mathcal{K}_h^l} K^l$ . The set  $\mathcal{E}_h^l$  is defined accordingly:

$$\mathcal{E}_h^l := \{p(E) \mid E \in \mathcal{E}_h\}$$

Then we can also define our bilinear form with respect to those geometric entities (and the

according gradient), leading to

$$a_h^l(u, v) = \sum_{K^l \in \mathcal{K}_h^l} \int_{K^l} \nabla_{\Gamma} u \cdot \nabla_{\Gamma} v - \sum_{E^l \in \mathcal{E}_h^l} \int_{E^l} \{n_{E^l} \cdot \nabla u\} [v] - \sum_{E^l \in \mathcal{E}_h} \int_{E^l} \{n_{E^l} \cdot \nabla v\} [u] \\ + \frac{\beta_E}{h} \sum_{E^l \in \mathcal{E}_h} \int_{E^l} [u][v].$$

For functions  $v \in H^2(\Gamma)$  or  $v \in V_h^l$ , the following norm can be introduced

$$\|v\|_*^2 := \sum_{K^l \in \mathcal{K}_h^l} \|\nabla_{\Gamma} v\|_{K^l}^2 + \sum_{E^l \in \mathcal{E}_h^l} \|h^{-1/2}[v]\|_{E^l}^2 + \sum_{K^l \in \mathcal{K}_h^l} \|h^{1/2}\nabla_{\Gamma} v\|_{\partial K^l}^2.$$

In the following, we will sometimes abbreviate certain sums of norms on entities which are collected in a set like  $\mathcal{K}_h$  or  $\mathcal{E}_h$ . Namely, we will just write

$$\|\dots\|_{\mathcal{X}}^2 \quad \text{for} \quad \sum_{X \in \mathcal{X}} \|\dots\|_X^2 \quad \text{for some set of regions } \mathcal{X}.$$

The solution to the strong PDE problem  $u$  fulfills an identity with the lifted bilinear form, as one can show with a calculation which is essentially the derivation of the weak problem the other way round:

$$a_h^l(u, v^l) = l(v^l) \quad \forall v \in V_h|_{\Gamma_h}. \quad (10)$$

The forthcoming analysis of the method will now proceed as follows: As a first important result, we want to show coercivity and continuity with respect to the norms  $\|\dots\|_h$  and  $\|\dots\|_{h,*}$ <sup>5</sup>. It will turn out that the norm  $\|\dots\|_h$  is suitable for the coercivity estimate and the  $\|\dots\|_{h,*}$  norm for continuity. We invite the reader to already have a look at Proposition B.5.5 and its proof to know the motivation for the technical estimates which directly follow this remark. The motivation can be briefly summarised as follows: The proof of Prop. B.5.5 will require the equivalence of the two norms, which is stated in lemma B.5.\*4. For that we will now prepare and proof a series of estimates in the lemmata B.5.1 to B.5.4. We collect the estimates and preparatory results in the following paragraphs, whose names refer to [4].

**Section 3: Domain Perturbation Related Estimates** The fact that we numerically consider only a polygonal approximation  $\Gamma_h$  of the actual surface  $\Gamma$  leads to a numerical error. In order to control it, certain estimates about integration on those regions are needed.

First of all, the respective gradients have to be related to each other. Since the numerical

---

<sup>5</sup>Or to be more precise: With respect to one of them; but the mentioned proposition is the key intermediate step.

surface  $\Gamma_h$  comes with its own normal vector, the gradient of an extended function  $\bar{u}$  is projected to a slightly shifted tangential space. But fortunately, the gradients are related to each other in terms of an operator  $B$ , which depends on the signed distance function  $\rho$  and its Hessian  $H$ :

$$\begin{aligned} B &= P_\Gamma(I - \rho H)P_{\Gamma_h}, \\ \nabla_{\Gamma_h} v^e &= B^T \nabla_\Gamma v \\ \nabla_\Gamma w^l &= B^{-T} \nabla_{\Gamma_h} w. \end{aligned}$$

Those representations allow one to show the boundedness of  $B$ , its inverse, and  $P_\Gamma - BB^T$  by exploiting the estimates from the Assumption B.A. (see [4, Lemma 3.1, Appendix] for the details)

Furthermore, it can be shown that the change in the integration measure if we go from a  $K \in \mathcal{K}_h$  to a  $K^l$ , or from an  $E \in \mathcal{E}_h$  to an  $E^l$ , is bounded. (see [4, Lemma 3.2, Appendix])

The most important consequence of these two observations for the forthcoming analysis are that the norms on the discrete and lifted geometric entities are equivalent. This is summarised in the following lemma.

**Lemma B.3.3** For  $v \in L^2(\mathcal{K}_h)$  and  $w \in L^2(\mathcal{K}_h)$  it holds<sup>6</sup>

$$C^{B.3.3.1} \|v\|_{\mathcal{K}_h} \leq \|v^l\|_{\mathcal{K}_h^l} \leq C^{B.3.3.2} \|v\|_{\mathcal{K}_h}, \quad C^{B.3.3.3} \|w^e\|_{\mathcal{K}_h} \leq \|w\|_{\mathcal{K}_h^l} \leq C^{B.3.3.4} \|w^e\|_{\mathcal{K}_h}.$$

For  $v \in L^2(\mathcal{E}_h)$ ,  $w \in L^2(\mathcal{E}_h^l)$  it holds

$$C^{B.3.3.5} \|v\|_{\mathcal{E}_h} \leq \|v^l\|_{\mathcal{E}_h^l} \leq C^{B.3.3.6} \|v\|_{\mathcal{E}_h}, \quad C^{B.3.3.7} \|w^e\|_{\mathcal{E}_h} \leq \|w\|_{\mathcal{E}_h^l} \leq C^{B.3.3.8} \|w^e\|_{\mathcal{E}_h}.$$

For  $v \in H^1(\mathcal{K}_h)$ ,  $w \in H^1(\mathcal{K}_h^l)$  we have

$$\begin{aligned} C^{B.3.3.9} \|\nabla_{\Gamma_h} v\|_{\mathcal{K}_h} &\leq \|\nabla_\Gamma v^l\|_{\mathcal{K}_h^l} \leq C^{B.3.3.10} \|\nabla_{\Gamma_h} v\|_{\mathcal{K}_h}, \\ C^{B.3.3.11} \|\nabla_{\Gamma_h} w^e\|_{\mathcal{K}_h} &\leq \|\nabla_\Gamma w\|_{\mathcal{K}_h^l} \leq C^{B.3.3.12} \|\nabla_{\Gamma_h} w^e\|_{\mathcal{K}_h}. \end{aligned}$$

For  $v \in H^2(\Gamma)^e$  or  $v \in V_h$  it holds

$$C^{B.3.3.13} \|\nabla_{\Gamma_h} v\|_{\partial\mathcal{K}_h} \leq \|\nabla_\Gamma v^l\|_{\partial\mathcal{K}_h^l} \leq C^{B.3.3.14} \|\nabla_{\Gamma_h} v\|_{\partial\mathcal{K}_h}.$$

---

<sup>6</sup> $v \in L^2(\mathcal{K}_h)$  is a shorthand for  $v: X \rightarrow \mathbb{R}$  with  $X$  such that  $X \supset K \forall K \in \mathcal{K}_h$  and  $v|_K \in L^2(K) \forall K \in \mathcal{K}_h$ . It will be also applied to e.g.  $H^1(\dots)$  later on.

This lemma can be applied to the norm definitions to obtain the next lemma.

**Lemma B.3.4** For a function  $v \in H^2(\Gamma)^e$  or  $v \in V_h$ , we have

$$\|v^l\|_* \leq C^{B.3.4} \|v\|_{*,h}.$$

**Proof** Let us begin with writing down the left hand side of the inequality squared:

$$\|v^l\|_*^2 = \underbrace{\|\nabla_{\Gamma} v^l\|_{\mathcal{K}_h^l}^2}_{\leq (C^{3.3.10})^2 \|\nabla_{\Gamma_h} v\|_{\mathcal{K}_h}^2} + h^{-1} \cdot \underbrace{\|[v^l]\|_{\mathcal{E}_h^l}^2}_{\leq (C^{B.3.3.6})^2 \|[v]\|_{\mathcal{E}_h}^2} + h \cdot \underbrace{\|\nabla_{\Gamma} v^l\|_{\partial \mathcal{K}_h^l}^2}_{\leq (C^{3.3.14})^2 \|\nabla_{\Gamma_h} v\|_{\partial \mathcal{K}_h}^2}.$$

This implies that the result holds with

$$C^{B.3.4} = \max\{C^{3.3.10}, C^{3.3.6}, C^{3.3.14}\}. \quad \square$$

**Section 4** In that section, firstly several useful inequalities are introduced. We start with a trace inequality for  $v \in H^1(\mathcal{T}_h)$ :<sup>7</sup>

$$\|v\|_{\partial T} \leq C^{B.4.0.1} \left( h^{-1/2} \|v\|_T + h^{1/2} \|\nabla v\|_T \right) \quad \forall T \in \mathcal{T}_h.$$

A similar estimate does also hold with respect to the surface  $\Gamma$  restricted to the element instead of the boundary

$$\|v\|_{\Gamma \cap T} \leq C^{B.4.0.2} \left( h^{-1/2} \|v\|_T + h^{1/2} \|\nabla v\|_T \right) \quad \forall T \in \mathcal{T}_h,$$

under the assumption that the surface is reasonably resolved. The latter is always provided in the relevant lemmata by the assumption of a sufficiently small mesh size. (See the reference in [4] for details)

Correspondingly, it holds for the skeleton  $\mathcal{F}_h$ :

$$\|v\|_{E \cap F} \leq C^{B.4.0.3} \left( h^{-1/2} \|v\|_F + h^{1/2} \|\nabla v\|_F \right) \quad \forall E \in \mathcal{E}_h, F \in \mathcal{F}_h.$$

For a facet  $\hat{F}$  of an element  $\hat{T}$ ,  $\hat{F} \subset \partial \hat{T}$  and therefore  $\|\dots\|_{\hat{F}} \leq \|\dots\|_{\partial \hat{T}}$ . This provides us with

$$\|v\|_{E \cap F} \leq C^{B.4.0.4} \left( h^{-1} \|v\|_T + \|\nabla v\|_T + h \|\nabla \otimes \nabla v\|_T \right) \quad \forall E \in \mathcal{E}_h, F \in \mathcal{F}_h,$$

---

<sup>7</sup>See Fn. 6 for notation.

where  $C^{B.4.0.4} = 2 \cdot C^{B.4.0.1} \cdot C^{B.4.0.3}$ .

For a function from the discrete space  $V_h$ ,  $v_h \in V_h$ , we have simplified similar estimates, namely

$$\|\nabla v_h\|_T \leq C^{B.4.0.5} h^{-1} \|v_h\|_T \quad \forall T \in \mathcal{T}_h$$

Furthermore, if we want to relate the integral over  $\partial T$  with the one over  $T$ , we have

$$\|v_h\|_{\partial T} \leq C^{B.4.0.6} h^{-1/2} \|v_h\|_T, \quad \|\nabla v_h\|_{\partial T} \leq C^{B.4.0.7} h^{-1/2} \|\nabla v_h\|_T \quad \forall T \in \mathcal{T}_h.$$

Again, instead of the boundary  $\partial T$ , we can also use the surface:

$$\begin{aligned} \|v_h\|_{K \cap T} &\leq C^{B.4.0.7} h^{-1/2} \|v_h\|_T, \quad \|\nabla v_h\|_{K \cap T} \leq C^{B.4.0.8} h^{-1/2} \|\nabla v_h\|_T \quad \forall T \in \mathcal{T}_h, K \in \mathcal{K}_h \\ \|v_h\|_{E \cap F} &\leq C^{B.4.0.9} h^{-1/2} \|v_h\|_F, \quad \|\nabla v_h\|_{E \cap F} \leq C^{B.4.0.10} h^{-1/2} \|\nabla v_h\|_F \quad \forall F \in \mathcal{F}_h, E \in \mathcal{E}_h. \end{aligned}$$

For the parts of the analysis after Lemma B.6.1 there are some estimates on interpolation operators required which could be conveniently introduced here. However, the detailed estimates are relevant only for details we will skip in this thesis. Therefore let us plainly introduce the notation  $I_h: L^2(N_h) \rightarrow V_h$  for the Scott-Zhang interpolation operator mapping a  $L^2$  function on  $N_h$  into our discrete function space.<sup>8</sup> The reader might want to consider [4] and the references therein for further details.

**Section 5** In section 5, Burman et al. introduce the concept of fat intersection coverings. Those are necessary for the analysis since in an unfitted setting it can happen that a particular  $K \in \mathcal{K}_h$  is only a small cut compared to the background triangle  $T \in \mathcal{T}_h$  (see, e.g., Fig. 5). The fat intersection property states that for such elements we can find a nearby element which has a significant cut. To state this more accurately, we start with some definitions for a point  $x \in \Gamma$ . Let  $B_\delta(x) = \{y \in \mathbb{R}^d \mid |x - y| < \delta\}$ , where  $|\dots|$  is the usual euclidean norm, and  $D_\delta$  the intersection of this neighbourhood with  $\Gamma$ ,  $D_\delta = B_\delta(x) \cap \Gamma$ . Then we collect all elements of  $\mathcal{K}_h$  which have an intersection with that  $D_\delta$  when they are lifted,

$$\mathcal{K}_{\delta,x} := \{K \in \mathcal{K}_h \mid K^l \cap D_\delta(x) \neq \emptyset\}.$$

---

<sup>8</sup>As Christoph Lehrenfeld pointed out to me, the Clément interpolator would also suffice here, since there are no strong boundary conditions to impose.

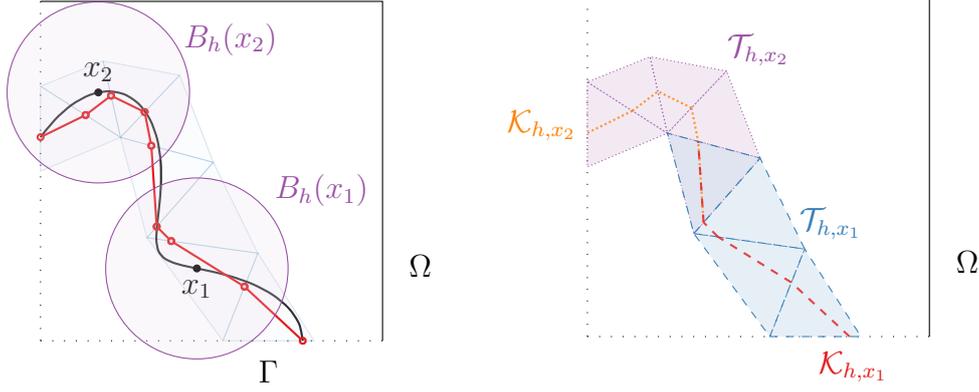


Figure 5: Fat intersection property helper definitions illustration.

Furthermore, the corresponding elements are

$$\mathcal{T}_{\delta,x} := \{T \in \mathcal{T}_h \mid T \cap \Gamma_h \in \mathcal{K}_{\delta,x}\}.$$

Those regions are illustrated for  $\delta = h$  for a simple example (a part of the starfish we will introduce in detail later) in Fig. 5. The fat intersection property then is: For every triangulation  $\tilde{\mathcal{T}}_h$ , resulting from the maximal mesh size parameter  $h \in (0, h_0]$  for some reasonably small  $h_0$ , there is a set  $\mathcal{X}_h \subset \Gamma$  such that:

- (1) The set  $\{\mathcal{K}_{h,x} \mid x \in \mathcal{X}_h\}$  covers  $\mathcal{K}_h$ , i.e.,

$$\mathcal{K}_h = \bigcup_{x \in \mathcal{X}_h} \mathcal{K}_{h,x}.$$

- (2) The set  $\{\mathcal{T}_{h,x} \mid x \in \mathcal{X}_h\}$  covers  $\mathcal{T}_h$ , i.e.,

$$\mathcal{T}_h = \bigcup_{x \in \mathcal{X}_h} \mathcal{T}_{h,x}.$$

- (3) For every point  $y \in \mathbb{R}^d$  we have

$$\#\{x \in \mathcal{X}_h \mid y \in \mathcal{T}_{h,x}\} \leq C^{B.5.0.1}.$$

- (4) For every  $x \in \mathcal{X}_h$ , we have

$$\#\mathcal{T}_{h,x} \leq C^{B.5.0.2}.$$

- (5) For every  $x \in \mathcal{X}_h$ : Each element  $T \in \mathcal{T}_{h,x}$  shares at least one facet with another

$$T' \in \mathcal{T}_{h,x}.$$

(6) For every  $x \in \mathcal{X}_h$  we can find a  $T_x \in \mathcal{T}_{h,x}$  such that

$$C^{B.5.0.3}|T_x| \leq h \cdot |T_x \cap \Gamma_h| = h \cdot |K_x| \leq C^{B.5.0.4}|T_x|.$$

Some explanations might be convenient now: Property (1) states that we have to choose enough points for  $\mathcal{X}_h$  such that the whole discrete surface  $\Gamma_h$  is covered with the  $\mathcal{K}_{h,x}$  induced by the  $x \in \mathcal{X}_h$ . Property (2) just mirrors that with regard to the active background mesh and the  $\mathcal{T}_{h,x}$ . Property (3) says that only  $C^{B.5.0.1}$  of the sets  $\mathcal{T}_{h,x}$  should intersect with each other. All the constants do not depend on  $h$ . In the example in Fig. 5, we have  $C^{B.5.0.1} = 2$ . The properties (4)-(6) now focus on a specific  $\mathcal{T}_{h,x}$  (for some  $x \in \mathcal{X}_h$ ). Firstly, such a region should consist at most out of  $C^{B.5.0.2}$  elements. In the example, this would be 5, due to  $\mathcal{T}_{h,x_2}$ . Secondly, for each element in  $\mathcal{T}_{h,x}$  we can find a neighbour within that set. An example of an element with only exactly one such other element is the one double coloured in Fig. 5 (the double colouring in the volume is unfortunately hard to recognise. The reader might want to consider the correspondingly coloured boundaries as a guide). Lastly, our specific  $\mathcal{T}_{h,x}$  has at least one element with a “fat intersection”.

It has been shown that this property holds for regular meshes in [3].

From (2) and (3) we can conclude the following:<sup>9</sup>

**Lemma B.5.\*1** For  $v \in L^2(N_h)$ ,  $w \in L^2(\mathcal{K}_h)$  and  $\mathcal{X}_h, \mathcal{T}_{h,x}$  as above, it holds

$$\begin{aligned} \|v\|_{\mathcal{T}_h}^2 &\leq \sum_{x \in \mathcal{X}_h} \|v\|_{\mathcal{T}_{h,x}}^2 \leq C^{B.5.*1} \|v\|_{\mathcal{T}_h}^2, \\ \|w\|_{\mathcal{K}_h}^2 &\leq \sum_{x \in \mathcal{X}_h} \|w\|_{\mathcal{K}_{h,x}}^2 \leq C^{B.5.*1} \|w\|_{\mathcal{K}_h}^2. \end{aligned}$$

Furthermore, we can define a local inner facet space for each  $x \in \mathcal{X}_h, \mathcal{T}_{h,x}$  as

$$\mathcal{F}_{h,x} := \{T^+ \cap T^- \mid T^+, T^- \in \mathcal{T}_{h,x}, T^+ \neq T^-\},$$

which also satisfies

$$\|v\|_{\mathcal{F}_h}^2 \leq \sum_{x \in \mathcal{X}_h} \|v\|_{\mathcal{F}_{h,x}}^2 \leq C^{B.5.*1} \|v\|_{\mathcal{F}_h}^2$$

for  $v \in L^2(\mathcal{F}_h)$ .

---

<sup>9</sup>This is our first helper lemma, see the first paragraph of this section for a remark on the admittedly a bit complicated numbering policy.

**Proof** Let us begin with the left  $\leq$  in the first line. Inserting the definitions gives that we need to show

$$\|v\|_{\mathcal{T}_h}^2 = \sum_{T \in \mathcal{T}_h} \|v\|_T^2 \leq \sum_{x \in \mathcal{X}_h} \sum_{T \in \mathcal{T}_{h,x}} \|v\|_T^2 = \sum_{x \in \mathcal{X}_h} \|v\|_{\mathcal{T}_{h,x}}^2.$$

Since each  $T \in \mathcal{T}_h$  appears in one  $\mathcal{T}_{h,x}$ , which is (2), the result holds.

For the right hand side  $\leq$ , we can rearrange the double sum to obtain

$$\sum_{x \in \mathcal{X}_h} \|v\|_{\mathcal{T}_{h,x}}^2 = \sum_{T \in \mathcal{T}_h} \alpha_T \|v\|_T^2, \quad \text{where } \alpha_T := \#\{x \in \mathcal{X}_h \mid T \in \mathcal{T}_{h,x}\}.$$

By (3) we have  $\alpha_T \leq C^{B.5.0.1}$ , since otherwise we could pick an arbitrary point  $\hat{y}$  within the  $\hat{T}$  in question to arrive at a contradiction. Therefore,

$$\sum_{x \in \mathcal{X}_h} \|v\|_{\mathcal{T}_{h,x}}^2 \leq \underbrace{C^{B.5.0.1}}_{=: C^{B.5.*1}} \cdot \sum_{T \in \mathcal{T}_h} \|v\|_T^2.$$

For the second and the third line a similar argument can be put forward regarding  $\mathcal{K}_h$  and  $\mathcal{F}_h$ .  $\square$

To make use of the fat intersection property, another lemma is needed. This will allow us to relate the norms of functions on adjacent elements.

**Lemma B.5.1** For a function  $v$  which is discontinuous but element-wise linear on a quasi-uniform mesh  $\mathcal{T}_h$ , consider two simplicial elements  $T_1, T_2$  sharing a face  $F$ . There we have

$$\begin{aligned} \|v\|_{T_1}^2 &\leq C^{B.5.1.1} \left( \|v\|_{T_2}^2 + h \|[v]\|_F^2 + h^3 \|n_F \cdot [\nabla v]\|_F^2 \right), \\ \|\nabla v\|_{T_1}^2 &\leq C^{B.5.1.2} \left( \|\nabla v\|_{T_2}^2 + h^{-1} \|[v]\|_F^2 + h \|n_F \cdot [\nabla v]\|_F^2 \right). \end{aligned}$$

For a proof see [4, Lemma 5.1] and the reference therein. In the proof the following inequality is exploited:

$$\|[ \nabla v ]\|_F^2 \leq C^{B.5.0.5} \left( \|n_F \cdot [\nabla v]\|_F^2 + h^{-2} \|[v]\|_F^2 \right). \quad (11)$$

We continue with some further intermediate results

**Lemma B.5.\*2** Let  $x \in \mathcal{X}_h$ ,  $\mathcal{T}_{h,x}$  be given as in the definition of the fat intersection property, and  $v \in V_h$ . Furthermore let  $\mathcal{F}_{h,x}$  be given as in Lemma B.5.\*1. Then it holds

$$\|v\|_{\mathcal{T}_{h,x}}^2 \leq C^{B.5.*2} \left( \|v\|_{T_x}^2 + h\|v\|_{\mathcal{F}_{h,x}}^2 + h^3\|n_F[\nabla v]\|_{\mathcal{F}_{h,x}} \right).$$

**Proof** We start by writing out the sum on the left hand side:

$$\|v\|_{\mathcal{T}_{h,x}}^2 = \sum_{T \in \mathcal{T}_{h,x}} \|v\|_T^2.$$

By fat intersection property item (4), the sum goes at most over  $C^{B.5.0.2}$  entries and we are now going to show that each is bounded. Eventually, we will arrive at the product of this element-wise bound and  $C^{B.5.0.2}$  as the total upper bound.

So let us assume we are at an arbitrary  $T \in \mathcal{T}_{h,x}$  and want to consider  $\|v\|_T^2$ . We know that  $\mathcal{T}_{h,x}$  is connected by construction and that there is some “fat”  $T_x \in \mathcal{T}_{h,x}$ . Therefore we can find  $T_0, \dots, T_n$  such that  $T_0 = T$  and  $T_n = T_x$ , and  $n \leq C^{B.5.0.2}$ . And for each  $i = 1, \dots, n$ ,  $T_i$  and  $T_{i-1}$  share a face, which we would like to call  $\hat{F}_{i-1/2}$ . Then we can apply Lemma B.5.1 to  $T_0, T_1$ , and  $\hat{F}_{1/2}$  to obtain

$$\|v\|_{T_0}^2 \leq C^{B.5.1.1} \left( \|v\|_{T_1}^2 + h\|v\|_{\hat{F}_{1/2}}^2 + h^3\|n_F[\nabla v]\|_{\hat{F}_{1/2}}^2 \right).$$

Now we continue with  $T_1, T_2, F_{1+1/2}$ :

$$\begin{aligned} \|v\|_{T_0}^2 &\leq C^{B.5.1.1} \left( C^{B.5.1.1} \left[ \|v\|_{T_2}^2 + h\|v\|_{\hat{F}_{1+1/2}}^2 + h^3\|n_F[\nabla v]\|_{\hat{F}_{1+1/2}}^2 \right] \right. \\ &\quad \left. + h\|v\|_{\hat{F}_{1/2}}^2 + h^3\|n_F[\nabla v]\|_{\hat{F}_{1/2}}^2 \right), \end{aligned}$$

which together with  $\tilde{C}^{B.5.1.1} = \max\{1, C^{B.5.1.1}\}$  leads to

$$\|v\|_{T_0}^2 \leq (\tilde{C}^{B.5.1.1})^2 \left( \|v\|_{T_2}^2 + \sum_{i=1}^2 h\|v\|_{\hat{F}_{i-1/2}}^2 + h^3\|n_F[\nabla v]\|_{\hat{F}_{i-1/2}}^2 \right).$$

This can be iterated to

$$\|v\|_{T_0}^2 \leq (\tilde{C}^{B.5.1.1})^n \left( \|v\|_{T_n}^2 + \sum_{i=1}^n h\|v\|_{\hat{F}_{i-1/2}}^2 + h^3\|n_F[\nabla v]\|_{\hat{F}_{i-1/2}}^2 \right).$$

By definition  $T_n$  was  $T_x$ , and furthermore we note  $\{F_{1/2}, F_{1+1/2}, \dots, F_{n-1/2}\} \subset \mathcal{F}_{h,x}$ , such

that we arrive at

$$\|v\|_T^2 \leq (\tilde{C}^{B.5.1.1})^{C^{B.5.0.2}} \left( \|v\|_{T_x}^2 + h \| [v] \|_{\mathcal{F}_{h,x}}^2 + h^3 \| n_F [\nabla v] \|_{\mathcal{F}_{h,x}}^2 \right).$$

That gives us the estimate to be proven for

$$C^{B.5.*2} = C^{B.5.0.2} \cdot (\max\{C^{B.5.1.1}, 1\})^{C^{B.5.0.2}}. \quad \square$$

**Lemma B.5.\*3** For each  $x \in \mathcal{X}_h$ , with  $\mathcal{X}_h$  satisfying the fat intersection property, we have for  $v$  elementwise constant (this is  $v|_T = \alpha_T$ )

$$\|v\|_{T_x} \leq C^{B.5.*3} \cdot h \cdot \|v\|_{K_x}.$$

**Proof** We observe that by property (6) of the fat intersection property

$$\|v\|_{T_x} = \alpha_{T_x} \cdot |T_x| \leq \alpha_{T_x} \frac{h}{C^{B.5.0.3}} |K_x| = \frac{h}{C^{B.5.0.3}} \|v\|_{K_x},$$

such that the lemma holds with  $C^{B.5.*3} = (C^{B.5.0.3})^{-1}$ . □

Now we are ready to state the first norm estimate

**Lemma B.5.2** For  $v \in V_h$

$$h^{-1} \|v - \lambda_{\Gamma_h}(v)\|_{N_h}^2 \leq C^{B.5.2} \left( \|\nabla_{\Gamma_h} v\|_{\mathcal{K}_h}^2 + h \|\nabla v\|_{N_h}^2 + h^{-2} \|[v]\|_{\mathcal{F}_h}^2 \right).$$

holds for  $h \in (0, h_0]$  for some  $h_0 < 1$  small enough. Here  $\lambda_{\Gamma_h}(v)$  denotes the mean value of  $v$  over  $\Gamma_h$ .

**Proof** We assume  $\lambda_{\Gamma_h}(v) = 0$ . For the case  $\lambda_{\Gamma_h}(v) = \alpha \neq 0$  we can consider the function  $\hat{v} = v - \alpha \in V_h$  to show the estimate, since the terms on the right hand are not sensible to the addition of a constant.

To estimate  $\|v\|_{N_h}^2$  we first of all note that  $\|\dots\|_{N_h}$  and  $\|\dots\|_{\mathcal{T}_h}$  can be used interchangeably

by the definition of  $N_h$ , (6):

$$\begin{aligned} \|v\|_{N_h}^2 &= \|v\|_{\mathcal{T}_h}^2 \stackrel{5.*1}{\leq} \sum_{x \in \mathcal{X}_h} \|v\|_{\mathcal{T}_{h,x}}^2 \stackrel{5.*2}{\leq} C^{B.5.*2} \left( \sum_{x \in \mathcal{X}_h} \|v\|_{T_x}^2 + h \|[v]\|_{\mathcal{F}_{h,x}}^2 + h^3 \|n_F[\nabla v]\|_{\mathcal{F}_{h,x}}^2 \right) \\ &\stackrel{5.*1}{\leq} C^{B.5.*2} \left( \sum_{x \in \mathcal{X}_h} \|v\|_{T_x}^2 \right) + \underbrace{C^{B.5.*1} C^{B.5.*2}}_{:=\pi_1} h \|[v]\|_{\mathcal{F}_h}^2 + C^{B.5.*1} C^{B.5.*2} h^3 \|n_F[\nabla v]\|_{\mathcal{F}_h}^2. \end{aligned}$$

Now we have to consider the last summand in more detail. With the abbreviation  $CS + Y$  for the application of Cauchy-Schwarz and Young's inequality<sup>10</sup> we have

$$\begin{aligned} \|n_F[\nabla v]\|_{\mathcal{F}_h}^2 &\leq \sum_{F \in \mathcal{F}_h} \int_F \underbrace{|n_F|^2}_{=1} |\nabla v_{left} - \nabla v_{right}|^2 \stackrel{CS+Y}{\leq} 2 \sum_{F \in \mathcal{F}_h} \int_F |\nabla v_{left}|^2 + |\nabla v_{right}|^2 \\ &\stackrel{\mathcal{F}_h \subset \partial \mathcal{T}_h}{\leq} 2 \|\nabla v\|_{\partial \mathcal{T}_h}^2 \leq 2(C^{B.4.0.7})^2 h^{-1} \|\nabla v\|_{\mathcal{T}_h}^2. \end{aligned}$$

Together with  $h < h^{-1}$  (since  $0 < h \leq h_0 < 1$ ) we arrive at

$$h^{-1} \|v\|_{N_h}^2 \leq \frac{C^{B.5.*2}}{h} \left( \sum_{x \in \mathcal{X}_h} \|v\|_{T_x}^2 \right) + \pi_1 h^{-2} \|[v]\|_{\mathcal{F}_h}^2 + \pi_1 h (C^{B.4.0.7})^2 \|\nabla v\|_{\mathcal{T}_h}^2.$$

The two second summands already look promising, so we continue with the first. We now introduce a scalar element-wise constant average function  $\bar{v}$  of  $v$ , which satisfies  $\bar{v}|_T = \frac{1}{|T|} \int_T v$ . This function satisfies

$$\|v - \bar{v}\|_T \leq h C^{B.5.0.6} \cdot \|\nabla v\|_T.$$

That could be shown by an explicit calculation on the reference triangle and then mapped to  $T$ , or by an  $L^2$  interpolation property for the approximation with piecewise constants. Then, we have by using Cauchy-Schwarz and Young's inequality again

$$\begin{aligned} \sum_{x \in \mathcal{X}_h} \|v\|_{T_x}^2 &\leq 2 \sum_{x \in \mathcal{X}_h} \|v - \bar{v}\|_{T_x}^2 + 2 \sum_{x \in \mathcal{X}_h} \|\bar{v}\|_{T_x}^2 \\ &\leq 2C^{B.5.*1} \sum_{T \in \mathcal{T}_h} \|v - \bar{v}\|_T^2 + \sum_{x \in \mathcal{X}_h} 2C^{B.5.*3} h \|\bar{v}\|_{K_x}^2 \\ &\leq 2C^{B.5.*1} (C^{B.5.0.6})^2 h^2 \|\nabla v\|_{\mathcal{T}_h}^2 + 2C^{B.5.*1} C^{B.5.*3} h \|\bar{v}\|_{\mathcal{K}_h}^2. \end{aligned}$$

---

<sup>10</sup> $\|a \pm b\|^2 = \|a\|^2 + \|b\|^2 \pm 2(a, b) \stackrel{CS}{\leq} \|a\|^2 + \|b\|^2 + 2\|a\|\|b\| \stackrel{Y}{\leq} \|a\|^2 + \|b\|^2 + 2\left(\frac{\|a\|^2 + \|b\|^2}{2}\right) = 2\|a\|^2 + 2\|b\|^2$

To exchange  $v$  and  $\bar{v}$  in the last summand, we observe

$$\begin{aligned} h\|\bar{v}\|_{\mathcal{K}_h}^2 &\leq h\|v\|_{\mathcal{K}_h}^2 + h\|\bar{v} - v\|_{\mathcal{K}_h}^2 \leq h\|v\|_{\mathcal{K}_h}^2 + (C^{B.4.0.7})^2\|v - \bar{v}\|_{\mathcal{T}_h}^2 \\ &\leq h\|v\|_{\mathcal{K}_h}^2 + (C^{B.4.0.7})^2 h^2 (C^{B.5.0.6})^2 \|\nabla v\|_{\mathcal{T}_h}^2 \end{aligned}$$

The last step in the proof now is to estimate the term  $h\|v\|_{\mathcal{K}_h}^2$ . We will skip this detail and just take from [4, (5.24)] that

$$h\|v\|_{\mathcal{K}_h}^2 \leq C \left( h^{-1}\|[v]\|_{\mathcal{F}_h}^2 + h^4\|v\|_{N_h}^2 + h\|\nabla_{\Gamma_h} v\|_{\mathcal{K}_h}^2 \right).$$

Inserting this back again gives us

$$\begin{aligned} \sum_{x \in \mathcal{X}_h} \|v\|_{T_x}^2 &\leq h^2 \|\nabla v\|_{\mathcal{T}_h}^2 \overbrace{2C^{B.5.*1} (C^{B.5.0.6})^2 (1 + C^{B.5.*3} (C^{B.4.0.7})^2)}^{:=\pi_2} \\ &\quad + \underbrace{2C^{B.5.*1} C^{B.5.*3} C}_{:=\pi_3} \left( h^{-1}\|[v]\|_{\mathcal{F}_h}^2 + h^4\|v\|_{N_h}^2 + h\|\nabla_{\Gamma_h} v\|_{\mathcal{K}_h}^2 \right), \end{aligned}$$

such that

$$\begin{aligned} \frac{C^{B.5.*2}}{h} \left( \sum_{x \in \mathcal{X}_h} \|v\|_{T_x}^2 \right) &\leq C^{B.5.*2} h \|\nabla v\|_{\mathcal{T}_h}^2 \pi_2 \\ &\quad + C^{B.5.*2} \pi_3 \left( h^{-2}\|[v]\|_{\mathcal{F}_h}^2 + h^3\|v\|_{N_h}^2 + \|\nabla_{\Gamma_h} v\|_{\mathcal{K}_h}^2 \right) \end{aligned}$$

and

$$\begin{aligned} h^{-1}\|v\|_{N_h}^2 &\leq C^{B.5.*2} h \|\nabla v\|_{\mathcal{T}_h}^2 \pi_2 + C^{B.5.*2} \pi_3 \left( h^{-2}\|[v]\|_{\mathcal{F}_h}^2 + h^3\|v\|_{N_h}^2 + \|\nabla_{\Gamma_h} v\|_{\mathcal{K}_h}^2 \right) \\ &\quad + \pi_1 h^{-2}\|[v]\|_{\mathcal{F}_h}^2 + \pi_1 h (C^{B.4.0.7})^2 \|\nabla v\|_{\mathcal{T}_h}^2 \\ &\leq \|\nabla_{\Gamma_h} v\|_{\mathcal{K}_h}^2 \underbrace{[C^{B.5.*2} \pi_3]}_{:=\pi_4} + h \|\nabla v\|_{N_h}^2 \underbrace{[\pi_2 C^{B.5.*2} \pi_1 (C^{B.4.0.7})^2]}_{:=\pi_5} \\ &\quad + h^{-2}\|[v]\|_{\mathcal{F}_h}^2 \underbrace{[\pi_3 C^{B.5.*2} + \pi_1]}_{:=\pi_6} + h^3\|v\|_{N_h}^2 \underbrace{[\pi_3 C^{B.5.*2}]}_{:=\pi_7}. \end{aligned}$$

For something like  $h$  sufficiently small such that  $h^4 \cdot \pi_7 \leq \frac{1}{2}$ , we obtain the estimate of the lemma with  $C^{B.5.2} = 2 \max\{\pi_4, \pi_5, \pi_6\}$ .  $\square$

In the next lemma, we estimate the gradient of  $v$  on  $N_h$  against the discrete gradient of the discrete surface and the facet triple norm:

**Lemma B.5.3** For  $v \in V_h$  and  $h \in (0, h_0]$  we have

$$h\|\nabla v\|_{N_h}^2 \leq C^{B.5.3} \left( h^2 \|\nabla_{\Gamma_h} v\|_{\mathcal{K}_h}^2 + \|v\|_{\mathcal{F}_h}^2 \right),$$

for  $h_0$  sufficiently small, where  $a = \lambda_{\Gamma_h}(\nabla v)$ .

**Proof** First of all note that for  $v \in V_h$  the gradient  $\nabla v$  lies in  $V_h \times V_h \times V_h$ , since differentiation only reduces the polynomial order. Furthermore note that we can derive a vector valued version of Lemma B.5.2 by simply applying the actual Lemma B.5.2 on each component. Starting with (6) and Cauchy-Schwarz and Young's inequality, we obtain

$$h\|\nabla v\|_{N_h}^2 = h\|\nabla v\|_{\mathcal{T}_h}^2 \leq 2h\|a\|_{\mathcal{T}_h}^2 + 2h\|\nabla v - a\|_{\mathcal{T}_h}^2.$$

Now we observe that higher derivatives of  $v$  vanish, by construction of  $V_h$ . So Lemma B.5.2. gives us

$$h\|\nabla v\|_{N_h}^2 \leq 2h\|a\|_{\mathcal{T}_h}^2 + 2C^{B.5.2}\|[\nabla v]\|_{\mathcal{F}_h}^2. \quad (12)$$

Regarding the first term we arrive at

$$h\|a\|_{\mathcal{T}_h}^2 = ha^2 \underbrace{\left( \sum_{T \in \mathcal{T}_h} |T| \right)}_{|N_h|} \leq C^{sr} h^2 a^2 \underbrace{\left( \sum_{K \in \mathcal{K}_h} |K| \right)}_{|\Gamma_h|} = C^{sr} h^2 \|a\|_{\mathcal{K}_h}^2,$$

applying the assumption of a shape regular triangulation in the estimate in the middle.

From differential geometry arguments (see [4]) we can argue that  $a \mapsto \|P_{\Gamma} a\|_{\Gamma}$  for  $a \in \mathbb{R}^d$  defines a norm on  $\mathbb{R}^d$ . On a finite dimensional vector space all norms are equivalent, so we have  $\|a\|_{\Gamma} \leq C^{\sim|\dots|} \|P_{\Gamma} a\|_{\Gamma}$ . With that we can estimate

$$\begin{aligned} \|a\|_{\mathcal{K}_h} &\leq (C^{B.3.3.1})^{-1} \|a\|_{\mathcal{K}_h^i} = (C^{B.3.3.1})^{-1} \|a\|_{\Gamma} \leq \frac{C^{\sim|\dots|}}{C^{B.3.3.1}} \|P_{\Gamma} a\|_{\Gamma} \\ &\leq \frac{C^{\sim|\dots|}}{C^{B.3.3.1}} C^{B.3.3.2} \|P_{\Gamma} a\|_{\mathcal{K}_h} \stackrel{\Delta\text{-ineq}}{\leq} \frac{C^{\sim|\dots|}}{C^{B.3.3.1}} C^{B.3.3.2} (\|P_{\Gamma_h} a\|_{\mathcal{K}_h} + \|(P_{\Gamma_h} - P_{\Gamma})a\|_{\mathcal{K}_h}). \end{aligned}$$

From Assumption B.A and inserting the definitions of  $P_{\Gamma}$  and  $P_{\Gamma_h}$  we obtain  $\|(P_{\Gamma_h} - P_{\Gamma})a\|_{\mathcal{K}_h} \leq Ch$ ,

where  $C$  is proportional to  $C^{B.2.0.2}$ . This leads to

$$\|a\|_{\mathcal{K}_h} \leq \frac{C^{\sim\|\dots\|}}{C^{B.3.3.1}} C^{B.3.3.2} (\|P_{\Gamma_h} a\|_{\mathcal{K}_h} + Ch\|a\|_{\mathcal{K}_h}) \leq 2 \underbrace{\frac{C^{\sim\|\dots\|}}{C^{B.3.3.1}} C^{B.3.3.2}}_{\pi_1} \|P_{\Gamma_h} a\|_{\mathcal{K}_h},$$

for  $h$  so small that  $hC \frac{C^{\sim\|\dots\|}}{C^{B.3.3.1}} C^{B.3.3.2} < \frac{1}{2}$ . Hence,  $h^2\|a\|_{\mathcal{K}_h}^2 \leq h^2\pi_1^2\|P_{\Gamma_h} a\|_{\mathcal{K}_h}^2$ .

We continue to estimate the latter term:

$$\begin{aligned} h^2\|P_{\Gamma_h} a\|_{\mathcal{K}_h}^2 &\stackrel{CS+Y}{\leq} 2h^2\|\underbrace{P_{\Gamma_h} \nabla v}_{\nabla_{\Gamma_h}}\|_{\mathcal{K}_h}^2 + 2h^2\|P_{\Gamma_h}(a - \nabla v)\|_{\mathcal{K}_h}^2 \\ &\leq 2h^2\|\nabla_{\Gamma_h} v\|_{\mathcal{K}_h}^2 + 2h^2\|P_{\Gamma_h}\|^2\|(a - \nabla v)\|_{\mathcal{K}_h}^2 \\ &\leq 2h^2\|\nabla_{\Gamma_h} v\|_{\mathcal{K}_h}^2 + 2h(C^{B.4.0.7})^2\|P_{\Gamma_h}\|^2\|(a - \nabla v)\|_{\mathcal{F}_h}^2 \\ &\leq 2h^2\|\nabla_{\Gamma_h} v\|_{\mathcal{K}_h}^2 + 2(C^{B.4.0.7})^2 C^{B.5.2}\|P_{\Gamma_h}\|^2\|[\nabla v]\|_{\mathcal{F}_h}^2. \end{aligned}$$

Putting those results together leads to

$$h\|a\|_{\mathcal{F}_h}^2 \leq \underbrace{2C^{sr}\pi_1^2}_{=: \pi_2} h^2\|\nabla_{\Gamma_h} v\|_{\mathcal{K}_h}^2 + \underbrace{2C^{sr}\pi_1^2(C^{B.4.0.7})^2 C^{B.5.2}\|P_{\Gamma_h}\|^2}_{=: \pi_3} \|[\nabla v]\|_{\mathcal{F}_h}^2,$$

where we can apply (11) now, to arrive at

$$h\|a\|_{\mathcal{F}_h}^2 \leq \pi_2 h^2\|\nabla_{\Gamma_h} v\|_{\mathcal{K}_h}^2 + \pi_3 C^{B.5.0.5} (\|n_F \cdot [\nabla v]\|_{\mathcal{F}_h}^2 + h^{-2}\|[v]\|_{\mathcal{F}_h}^2).$$

This can be inserted into (12), to result in

$$\begin{aligned} h\|\nabla v\|_{N_h}^2 &\leq 2\pi_2 h^2\|\nabla_{\Gamma_h} v\|_{\mathcal{K}_h}^2 + 2\pi_3 C^{B.5.0.5} \left( \|n_F \cdot [\nabla v]\|_{\mathcal{F}_h}^2 + \frac{\|[v]\|_{\mathcal{F}_h}^2}{h^2} \right) + 2C^{B.5.2}\|[\nabla v]\|_{\mathcal{F}_h}^2 \\ &= 2\pi_2 h^2\|\nabla_{\Gamma_h} v\|_{\mathcal{K}_h}^2 + (2\pi_3 C^{B.5.0.5} + 2C^{B.5.2})\|n_F \cdot [\nabla v]\|_{\mathcal{F}_h}^2 + 2\pi_3 C^{B.5.0.5} h^{-2}\|[v]\|_{\mathcal{F}_h}^2 \end{aligned}$$

Comparing this to the result we want to show, namely

$$h\|\nabla v\|_{N_h}^2 \leq C^{B.5.3} \left( h^2\|\nabla_{\Gamma_h} v\|_{\mathcal{K}_h}^2 + h^{-2}\|[v]\|_{\mathcal{F}_h}^2 + \|n_F \cdot [\nabla v]\|_{\mathcal{F}_h}^2 \right),$$

we can finish the proof by setting

$$C^{B.5.3} = \max\{2\pi_2, 2\pi_3 C^{B.5.0.5} + 2C^{B.5.2}\}. \quad \square$$

Let us now come to the next estimate.

**Lemma B.5.4** For  $v \in V_h$ ,  $h \in (0, h_0]$  for some  $h_0$  sufficiently small,

$$h \|\nabla_{\Gamma_h} v\|_{\partial\mathcal{K}_h}^2 \leq C^{B.5.4} \left( \|\nabla_{\Gamma_h} v\|_{\Gamma_h}^2 + \|v\|_{\mathcal{F}_h}^2 \right).$$

**Proof** First of all,  $\|\dots\|_{\partial\mathcal{K}_h} = 2 \cdot \|\dots\|_{\mathcal{E}_h}$  by construction. With that we have

$$\begin{aligned} h \|\nabla_{\Gamma_h} v\|_{\partial\mathcal{K}_h}^2 &= 2h \|\nabla_{\Gamma_h} v\|_{\mathcal{E}_h}^2 \leq 2h^0 C^{B.4.0.9} \|\nabla_{\Gamma_h} v\|_{\mathcal{F}_h}^2 \\ &\leq 2h^{-1} C^{B.4.0.9} C^{B.4.0.6} \|\nabla_{\Gamma_h} v\|_{\mathcal{T}_h}^2 \\ &\stackrel{B.5.*1}{\leq} 2h^{-1} C^{B.4.0.9} C^{B.4.0.6} \sum_{x \in \mathcal{X}_h} \|\nabla_{\Gamma_h} v\|_{\mathcal{T}_{h,x}}^2. \end{aligned}$$

Note that since  $n_h$  and  $\nabla v$  are element-wise constant, also  $\nabla_{\Gamma_h} v$  is element-wise constant. So when we apply B.5.\*2, the right-hand side term vanishes:

$$\begin{aligned} h \|\nabla_{\Gamma_h} v\|_{\partial\mathcal{K}_h}^2 &\leq h^{-1} \underbrace{2C^{B.4.0.9} C^{B.4.0.6} C^{B.5.*2}}_{=:\pi_1} \cdot \left( \sum_{x \in \mathcal{X}_h} \|\nabla_{\Gamma_h} v\|_{\mathcal{T}_x}^2 + h \|[\nabla_{\Gamma_h} v]\|_{\mathcal{F}_{h,x}}^2 \right) \\ &\leq \pi_1 C^{B.5.*3} \left( \sum_{x \in \mathcal{X}_h} \|\nabla_{\Gamma_h} v\|_{\mathcal{K}_x}^2 \right) + \pi_1 \left( \sum_{x \in \mathcal{X}_h} \|[\nabla_{\Gamma_h} v]\|_{\mathcal{F}_{h,x}}^2 \right) \\ &\leq \underbrace{\pi_1 C^{B.5.*3} C^{B.5.0.1}}_{=:\pi_2} \|\nabla_{\Gamma_h} v\|_{\mathcal{K}_h}^2 + \underbrace{\pi_1 C^{B.5.*1}}_{=:\pi_3} \|[\nabla_{\Gamma_h} v]\|_{\mathcal{F}_h}^2. \end{aligned}$$

On the first summand we can apply the fact that  $\Gamma_h = \bigcup \mathcal{K}_h$ , such that

$$\|\nabla_{\Gamma_h} v\|_{\mathcal{K}_h}^2 = \|\nabla_{\Gamma_h} v\|_{\Gamma_h}^2.$$

Regarding the second term one can show with the definition  $\nabla_{\Gamma_h} \dots = P_{\Gamma_h} \nabla \dots$ , an identity about jump-operators, and an estimate on the difference between  $P_{\Gamma_h}$  and  $P_{\Gamma}$  based on Assumption B.A that [4, (5.43)]

$$\begin{aligned} \|[\nabla_{\Gamma_h} v]\|_{\mathcal{F}_h}^2 &\leq Ch^2 \|\nabla v\|_{\partial\mathcal{T}_h}^2 + D \|[\nabla v]\|_{\mathcal{F}_h}^2 \leq Ch C^{B.4.0.7} \|\nabla v\|_{N_h}^2 + D \|[\nabla v]\|_{\mathcal{F}_h}^2 \\ &\leq \underbrace{C C^{B.4.0.7} C^{B.5.3}}_{=:\pi_4} \left( h^2 \|\nabla_{\Gamma_h} v\|_{\mathcal{K}_h}^2 + \|v\|_{\mathcal{F}_h}^2 \right) + D \|[\nabla v]\|_{\mathcal{F}_h}^2 \end{aligned}$$

Putting together those estimates gives

$$\begin{aligned} h\|\nabla_{\Gamma_h} v\|_{\partial\mathcal{K}_h}^2 &\leq \pi_2\|\nabla_{\Gamma_h} v\|_{\Gamma_h}^2 + \pi_3\pi_4 h^2\|\nabla_{\Gamma_h} v\|_{\Gamma_h}^2 + \pi_3\pi_4\|v\|_{\mathcal{F}_h}^2 + \pi_3 D\|[\nabla v]\|_{\mathcal{F}_h}^2 \\ &\leq (\pi_2 + \pi_3\pi_4)\|\nabla_{\Gamma_h} v\|_{\Gamma_h}^2 + (\pi_3\pi_4 + \pi_3 D)\|v\|_{\mathcal{F}_h}^2. \end{aligned}$$

So the estimate holds with  $C^{B.5.4} = \max\{\pi_2 + \pi_3\pi_4, \pi_3\pi_4 + \pi_3 D\}$ .  $\square$

Now we are able to pose a helper lemma which states the equivalence of the norms  $\|\cdot\|_{*,h}$  and  $\|\cdot\|_h$ .

**Lemma B.5.\*4** For all  $v \in V_h$  we have

$$\|v\|_h \leq \|v\|_{*,h} \leq C^{B.5.*4}\|v\|_h$$

**Proof** First of all, let us review the definitions of the respective norms:

$$\begin{aligned} \|v\|_h^2 &= \|\nabla_{\Gamma_h} v\|_{\mathcal{K}_h}^2 + \|h^{-1/2}[v]\|_{\mathcal{E}_h}^2 + \|h^{-1}[v]\|_{\mathcal{F}_h}^2 + \|n_F \cdot [\nabla v]\|_{\mathcal{F}_h}^2 \\ \|v\|_{*,h}^2 &= \|v\|_h^2 + \|h^{1/2}\nabla_{\Gamma_h} v\|_{\partial\mathcal{K}_h}^2 \end{aligned}$$

Hence, the left hand side inequality follows directly by definition. For the right hand side inequality we insert Lemma B.5.4 to obtain

$$\|v\|_{*,h}^2 \leq \|v\|_h^2 + C^{B.5.4} \left( \|\nabla_{\Gamma_h} v\|_{\Gamma_h}^2 + \|v\|_{\mathcal{F}_h}^2 \right)$$

Since the norm on  $\mathcal{K}_h$  is the same as the norm  $\Gamma_h$ , the estimate holds with  $C^{B.5.*4} = (1 + C^{B.5.4})^{1/2}$ .  $\square$

Now we can show the important fact that our bilinear form is both coercive and continuous:

**Proposition B.5.5 (Proposition 5.1 in [4])** The discrete bilinear form  $A_h$  satisfies

$$A_h(v, v) \geq C^{B.5.5.1}\|v\|_h^2 \quad \forall v \in V_h \quad (\text{coercivity})$$

and

$$A_h(u, v) \leq C^{B.5.5.2}\|u\|_h \cdot \|v\|_h \quad \forall u, v \in V_h, \quad (\text{continuity})$$

for  $\beta_E, \beta_F, \gamma$  large enough and  $h \in (0, h_0]$  for  $h_0$  small enough.

**Proof** Regarding coercivity we start with inserting  $v$  two times in  $A_h$  to obtain

$$\begin{aligned} A_h(v, v) &= \|\nabla_{\Gamma_h} v\|_{\mathcal{K}_h}^2 - 2(\{n_E \cdot \nabla v\}, [v])_{\mathcal{E}_h} + \frac{\beta_E}{h} \|[v]\|_{\mathcal{E}_h}^2 + \frac{\beta_F}{h^2} \|[v]\|_{\mathcal{F}_h}^2 + \gamma \|n_F \cdot [\nabla v]\|_{\mathcal{F}_h}^2 \\ &\geq \|\nabla_{\Gamma_h} v\|_{\mathcal{K}_h}^2 - 2(\{n_E \cdot \nabla v\}, [v])_{\mathcal{E}_h} + \frac{\beta_E}{h} \|[v]\|_{\mathcal{E}_h}^2 + \min\{\beta_F, \gamma\} \|v\|_{\mathcal{F}_h}^2 \end{aligned}$$

Now we apply Cauchy's inequality with  $\epsilon$ , namely  $(a, b) \leq \epsilon \cdot (a, a) + \frac{(b, b)}{4\epsilon}$

$$\begin{aligned} 2(\{n_E \cdot \nabla v\}, [v])_{\mathcal{E}_h} &= 2(h^{1/2} \{n_E \cdot \nabla v\}, h^{-1/2}[v])_{\mathcal{E}_h} \\ &\leq 2\epsilon \|h^{1/2} \{n_E \cdot \nabla v\}\|_{\mathcal{E}_h}^2 + \frac{1}{2\epsilon} \|h^{-1/2}[v]\|_{\mathcal{E}_h}^2, \end{aligned}$$

such that

$$\begin{aligned} A_h(v, v) &\geq \|\nabla_{\Gamma_h} v\|_{\mathcal{K}_h}^2 - 2\epsilon \|h^{1/2} \{n_E \cdot \nabla v\}\|_{\mathcal{E}_h}^2 \\ &\quad + \left(\beta_E - \frac{1}{2\epsilon}\right) \|h^{-1/2}[v]\|_{\mathcal{E}_h}^2 + \min\{\beta_F, \gamma\} \|v\|_{\mathcal{F}_h}^2. \end{aligned}$$

Now we exploit an interesting identity stemming from the orthogonality of  $n_h$  and  $n_E$ , namely

$$n_E \cdot \nabla_{\Gamma_h} v = n_E \cdot (\nabla v + n_h \cdot (n_h \cdot \nabla v)) = n_E \cdot \nabla v + \underbrace{n_E n_h}_{=0} \cdot (n_h \cdot \nabla v) = n_E \cdot \nabla v,$$

to obtain

$$2\epsilon \|h^{1/2} \{n_E \cdot \nabla v\}\|_{\mathcal{E}_h}^2 \leq \epsilon h \|n_E \cdot \nabla_{\Gamma_h} v\|_{\partial \mathcal{K}_h}^2 \leq \epsilon C^{B.5.4} \left( \|\nabla_{\Gamma_h} v\|_{\Gamma_h}^2 + \|v\|_{\mathcal{F}_h}^2 \right)$$

Putting this together yields

$$\begin{aligned} A_h(v, v) &\geq (1 - \epsilon C^{B.5.4}) \|\nabla_{\Gamma_h} v\|_{\mathcal{K}_h}^2 \\ &\quad + \left(\beta_E - \frac{1}{2\epsilon}\right) \|h^{-1/2}[v]\|_{\mathcal{E}_h}^2 + (\min\{\beta_F, \gamma\} - \epsilon C^{B.5.4}) \|v\|_{\mathcal{F}_h}^2. \end{aligned}$$

We come to the conclusion that

$$C^{B.5.5.1} = \min \left\{ (1 - \epsilon C^{B.5.4}), \left(\beta_E - \frac{1}{2\epsilon}\right), (\min\{\beta_F, \gamma\} - \epsilon C^{B.5.4}) \right\}.$$

For  $\epsilon$  small and  $\beta_E, \beta_F$ , and  $\gamma$  large this term indeed is positive.

Regarding continuity we begin with writing down  $A_h(u, v)$  with an inner product notation

parallel to our  $\|\dots\|_{\mathcal{K}_h}$  notation:

$$\begin{aligned} A_h(u, v) &= (\nabla_{\Gamma_h} u, \nabla_{\Gamma_h} v)_{\mathcal{K}_h} - (\{n_E \cdot \nabla u\}, [v])_{\mathcal{E}_h} - (\{n_E \cdot \nabla v\}, [u])_{\mathcal{E}_h} \\ &\quad + \frac{\beta_E}{h} ([u], [v])_{\mathcal{E}_h} + \frac{\beta_F}{h^2} ([u], [v])_{\mathcal{F}_h} + \gamma (n_F \cdot [\nabla u], n_F \cdot [\nabla v])_{\mathcal{F}_h} \end{aligned}$$

Applying now Cauchy-Schwarz inequality to all summands leads to

$$\begin{aligned} A_h(u, v) &\leq \|\nabla_{\Gamma_h} u\|_{\mathcal{K}_h} \cdot \|\nabla_{\Gamma_h} v\|_{\mathcal{K}_h} + \|\{n_E \cdot \nabla u\}\|_{\mathcal{E}_h} \cdot \|[v]\|_{\mathcal{E}_h} + \|\{n_E \cdot \nabla v\}\|_{\mathcal{E}_h} \cdot \|[u]\|_{\mathcal{E}_h} \\ &\quad + \frac{\beta_E}{h} \|[u]\|_{\mathcal{E}_h} \cdot \|[v]\|_{\mathcal{E}_h} + \frac{\beta_F}{h^2} \|[u]\|_{\mathcal{F}_h} \cdot \|[v]\|_{\mathcal{F}_h} + \gamma \|n_F \cdot [\nabla u]\|_{\mathcal{F}_h} \cdot \|n_F \cdot [\nabla v]\|_{\mathcal{F}_h}. \end{aligned}$$

Now we can apply a variant of Cauchy-Schwarz inequality about sums of products of real numbers, namely

$$\sum_{i=1}^n u_i v_i \leq \left( \sum_{i=1}^n u_i^2 \right)^{1/2} \cdot \left( \sum_{i=1}^n v_i^2 \right)^{1/2},$$

with  $u_1 = \|\nabla_{\Gamma_h} u\|_{\mathcal{K}_h}$ ,  $u_2 = \sqrt{h} \|\{n_E \cdot \nabla u\}\|_{\mathcal{E}_h}$ ,  $u_3 = \sqrt{1/h} \|[u]\|_{\mathcal{E}_h}$ ,  $u_4 = \sqrt{\beta_E/h} \|[u]\|_{\mathcal{E}_h}$ ,  $u_5 = \sqrt{\beta_F} h^{-1} \|[u]\|_{\mathcal{F}_h}$ ,  $u_6 = \sqrt{\gamma} \|n_F \cdot [\nabla u]\|_{\mathcal{F}_h}$ ,  $v_1 = \|\nabla_{\Gamma_h} v\|_{\mathcal{K}_h}$ ,  $v_2 = \sqrt{h} \|\{n_E \cdot \nabla v\}\|_{\mathcal{E}_h}$ ,  $v_3 = \sqrt{1/h} \|[v]\|_{\mathcal{E}_h}$ ,  $v_4 = \sqrt{\beta_E/h} \|[v]\|_{\mathcal{E}_h}$ ,  $v_5 = \sqrt{\beta_F} h^{-1} \|[v]\|_{\mathcal{F}_h}$ ,  $v_6 = \sqrt{\gamma} \|n_F \cdot [\nabla v]\|_{\mathcal{F}_h}$ .

Inserting this yields

$$\begin{aligned} A_h(u, v) &\leq \\ &\left( \|\nabla_{\Gamma_h} u\|_{\mathcal{K}_h}^2 + h \|\{n_E \cdot \nabla u\}\|_{\mathcal{E}_h}^2 + \frac{1 + \beta_E}{h} \|[u]\|_{\mathcal{E}_h}^2 + \frac{\beta_F}{h^2} \|[u]\|_{\mathcal{F}_h}^2 + \gamma \|n_F \cdot [\nabla u]\|_{\mathcal{F}_h}^2 \right)^{\frac{1}{2}} \\ &\cdot \left( \|\nabla_{\Gamma_h} v\|_{\mathcal{K}_h}^2 + h \|\{n_E \cdot \nabla v\}\|_{\mathcal{E}_h}^2 + \frac{1 + \beta_E}{h} \|[v]\|_{\mathcal{E}_h}^2 + \frac{\beta_F}{h^2} \|[v]\|_{\mathcal{F}_h}^2 + \gamma \|n_F \cdot [\nabla v]\|_{\mathcal{F}_h}^2 \right)^{\frac{1}{2}}, \end{aligned}$$

while

$$\|v\|_{*,h} = \sqrt{\|\nabla_{\Gamma_h} v\|_{\mathcal{K}_h}^2 + \frac{1}{h} \|[v]\|_{\mathcal{E}_h}^2 + \frac{1}{h^2} \|[v]\|_{\mathcal{F}_h}^2 + \|n_F \cdot [\nabla v]\|_{\mathcal{F}_h}^2 + h \|\nabla_{\Gamma_h} v\|_{\partial \mathcal{K}_h}^2}.$$

Regarding the first  $\mathcal{E}_h$  term in the upper bound for  $A_h$  we can estimate

$$h \|\{n_E \cdot \nabla v\}\|_{\mathcal{E}_h}^2 \leq \frac{h}{2} \|n_E \cdot \nabla v\|_{\partial \mathcal{K}_h}^2 = \frac{h}{2} \|n_E \cdot \nabla_{\Gamma_h} v\|_{\partial \mathcal{K}_h}^2 \leq \frac{h}{2} \|\nabla_{\Gamma_h} v\|_{\partial \mathcal{K}_h}^2. \quad (13)$$

Combining those results, we end up with

$$A_h(u, v) \leq \underbrace{\max\left\{1, \frac{1}{2}, 1 + \beta_E, \beta_F, \gamma\right\}}_{=\max\{1+\beta_E, \beta_F, \gamma\}} \cdot \|u\|_{*,h} \cdot \|u\|_{*,h},$$

such that  $C^{B.5.5.2} = \max\{1 + \beta_E, \beta_F, \gamma\} \cdot (C^{B.5.*4})^2$ .  $\square$

The result that the bilinear form is coercive and continuous on the discrete function space implies that our discrete problem has a unique solution by the Lax-Milgram lemma. It reads (see, e.g., [14, 2.28]):

**Lemma (Lax-Milgram)** Given a Hilbert space  $V$ , a coercive and continuous bilinear form  $A(\dots, \dots)$  and a continuous linear form  $f(\dots)$ . Then there exists a unique solution  $u \in V$  solving

$$A(u, v) = f(v) \quad \forall v \in V.$$

The solution satisfies the stability estimate  $\|u\|_V \leq \alpha_1^{-1} \|f\|_V$ , where  $\alpha_1$  is the coercivity constant of  $A$ .

Now that we have proven that there is a unique solution to our discrete problem, we can investigate how far it deviates from the continuous solution. That is typically done in a Strang-type lemma like

**Lemma B.6.1** With  $u \in H^2(\Gamma)$  the solution of (1) and  $u_h$  the solution of (7), it holds

$$\begin{aligned} \|u^e - u_h\|_{*,h} \leq & C^{B.6.1} (\|u^e - I_h u^e\|_{*,h} + \sup_{v \in V_h} \|v\|_h^{-1} (a_h(I_h u^e, v) - a_h^l((I_h u^e)^l, v^l))) \\ & + \sup_{v \in V_h} \|v\|_h^{-1} (l(v^l) - l_h(v)). \end{aligned}$$

**Proof** Writing  $e_h = I_h u^e - u_h$ , we can start with the triangle inequality to obtain

$$\|u^e - u_h\|_{*,h} \leq \|u^e - I_h u^e\|_{*,h} + \|e_h\|_{*,h}.$$

By definition  $e_h \in V_h$ , since  $V_h$  is a vector space. With the norm equivalence of B.5.\*4 in mind, we now investigate  $\|e_h\|_h$ :

$$\begin{aligned}
\|e_h\|_h^2 &\leq (C^{B.5.5.1})^{-1} A_h(I_h u^e - u_h, e_h) \stackrel{(7)}{=} (C^{B.5.5.1})^{-1} (A_h(I_h u^e, e_h) - l_h(e_h)) \\
&\stackrel{(10)}{=} (C^{B.5.5.1})^{-1} (A_h(I_h u^e, e_h) - a_h^l(u, e_h^l) + l(e_h^l) - l_h(e_h)) \\
&\stackrel{(7)}{=} (C^{B.5.5.1})^{-1} ((a_h(I_h u^e, e_h) - a_h^l((I_h u^e)^l, e_h^l)) + (l(e_h^l) - l_h(e_h)) \\
&\quad - a_h^l(u - (I_h u^e)^l, e_h^l) + j_h(I_h u^e, e_h)) \\
&= (C^{B.5.5.1})^{-1} (I + II + III + IV).
\end{aligned}$$

Regarding the first term we can of course consider the supremum over all  $v \in V_h$  instead of the particular  $e_h$  and end up with an upper bound:

$$I \leq \|e_h\|_h \sup_{v \in V_h} \|v\|_h^{-1} (a_h(I_h u^e, v) - a_h^l((I_h u^e)^l, v^l)).$$

The same point can be made about II, leading to

$$II \leq \|e_h\|_h \sup_{v \in V_h} \|v\|_h^{-1} (l(v^l) - l_h(v)).$$

III: Apply Cauchy-Schwarz on  $a_h^l$  respectively the norm  $\|\cdot\|_*$  to obtain

$$\begin{aligned}
III &\leq \|u - (I_h u^e)^l\|_* \cdot \|e_h^l\|_* \stackrel{B.3.4}{\leq} (C^{B.3.4})^2 \|u^e - (I_h u^e)\|_{*,h} \cdot \|e_h\|_{*,h} \\
&\leq (C^{B.3.4})^2 C^{B.5.*4} \|u^e - (I_h u^e)\|_{*,h} \cdot \|e_h\|_h.
\end{aligned}$$

To bound IV, we note that  $j_h$  is the bilinear form associated to the facet norm. That allows to apply Cauchy-Schwarz again. Furthermore, for  $u^e$  the jump terms vanish since it is the continuous solution. This leads to

$$IV = j_h(I_h u^e - u^e, e_h) \leq \|u^e - I_h u^e\|_{\mathcal{F}_h} \cdot \|e_h\|_{\mathcal{F}_h} \leq \|u^e - I_h u^e\|_{*,h} \cdot \|e_h\|_h.$$

Collecting these results, we end up with

$$C^{B.6.1} = 1 + \frac{C^{B.5.*4}}{C^{B.5.5.1}} \left( (C^{B.3.4})^2 C^{B.5.*4} + 1 \right). \quad \square$$

The analysis in the Burman et al. paper now goes on to estimate the second and third term

in the last lemma as follows: For all  $u, v \in V_h$

$$|a_h^l(u^l, v^l) - a_h(u, v)| \lesssim h^2 \|u\|_h \|v\|_h \quad (14)$$

$$|l(u^l) - l_h(u)| \lesssim h^2 \|v\|_h \quad (15)$$

Eventually, we arrive at

**Theorem B.6.1** It holds

$$\|u^e - u_h\|_{*,h} \lesssim h \|f\|_\Gamma, \quad \|u^e - u_h\|_{\mathcal{K}_h} \lesssim h^2 \|f\|_\Gamma.$$

**Comments on the proof** For the first part, we have to combine (14) and (15) with the Strang lemma B.6.1. Then an interpolation bound (we skipped that in section 4) is needed as well as Equation (3), and the estimate follows. For the second estimate, the error bound in the  $L^2$  norm, an argument exploiting the dual problem is sufficient. The reader might want to consider [4, Theorem 6.1] for more details.

## 2.4 Numerical examples

Let us now illustrate and support the theoretical findings by some numerical results. In this subsection and in all further numerical considerations, we will follow Burman et al. in slightly modifying the solved partial differential equation in order to facilitate the implementation. Namely, we consider the problem

$$-\Delta_\Gamma u + u = f \quad \text{on } \Gamma.$$

The continuous bilinear form then takes the form

$$a(u, v) = (\nabla_\Gamma u, \nabla_\Gamma v)_\Gamma + (u, v)_\Gamma.$$

Compared to Equation (2), the tweaked bilinear form has the constant functions not in its kernel and therefore we do not have to consider the solution space  $H^1(\Gamma)/\mathbb{R}$ . This theoretical benefit is mirrored on the implementational side by the fact that we do not have to ensure that our solution has zero average.

Furthermore, we will not only consider test problems for which a signed distance function can be given easily. That will only be the case for the circle and the sphere. Otherwise, we will only have a levelset function “approximating” a signed distance function. But the theoretical

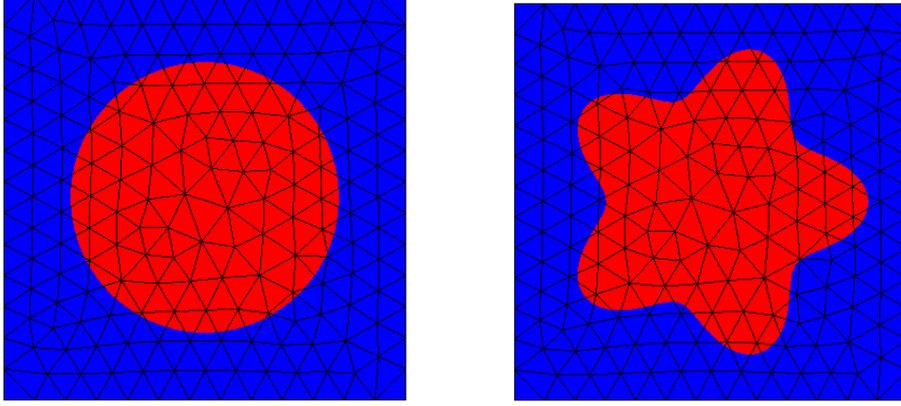


Figure 6: Circle and starfish geometry stemming from  $\phi_1$  and  $\phi_2$

assumption of a signed distance function was stronger than what would actually be needed for the method to work, as the fact that Burman et al. themselves consider a problem with a mere levelset function in [4, Section 8.1] mirrors.

In their paper, Burman et al. present two numerical examples in three dimensions. To complement those demonstrations, I want to investigate two examples in two dimensions here, namely the circle and the starfish geometry. They are given by the following levelset functions

$$\phi_1 = \sqrt{x^2 + y^2} - R, \quad \phi_2 = \sqrt{x^2 + y^2} - [r_0 + 0.2 \cdot \sin(5 \cdot \arctan2(x, y))],$$

where  $\arctan2(x, y)$  is the function which extends  $\arctan(y/x)$  to the singular cases and all four quadrants.

The resulting geometries are shown in Fig. 6. The corresponding right hand side is calculated such that the solution  $u$  turns out to be  $u = \sin(y)$ . For that the following representation of the Laplace-Beltrami operator is implemented in Sympy (see [4, p.20])<sup>11</sup>

$$\Delta_\Gamma = \Delta u - n_\Gamma \cdot \nabla \otimes \nabla u n_\Gamma - \text{tr}(\nabla n_\Gamma) \nabla u \cdot n_\Gamma.$$

The right hand side is—especially in the starfish case—to long to be shown here. The calculations are performed on an unstructured background mesh on the region  $[-1.5, 1.5]^2$ . The maximal allowed mesh size is controlled by a parameter  $n$  as  $0.5^{n+2}$ . The stabilisation is set to  $\beta_E = \beta_F = 50$ . We observe the error between the analytic solution and the numerical approximation in the  $L^2$ - and  $H^1$ -norm on  $\Gamma$ . The results are shown in Fig. 7 and 8 and

<sup>11</sup>André Massing provided me with an implementation from a different context and I adapted it to our case.

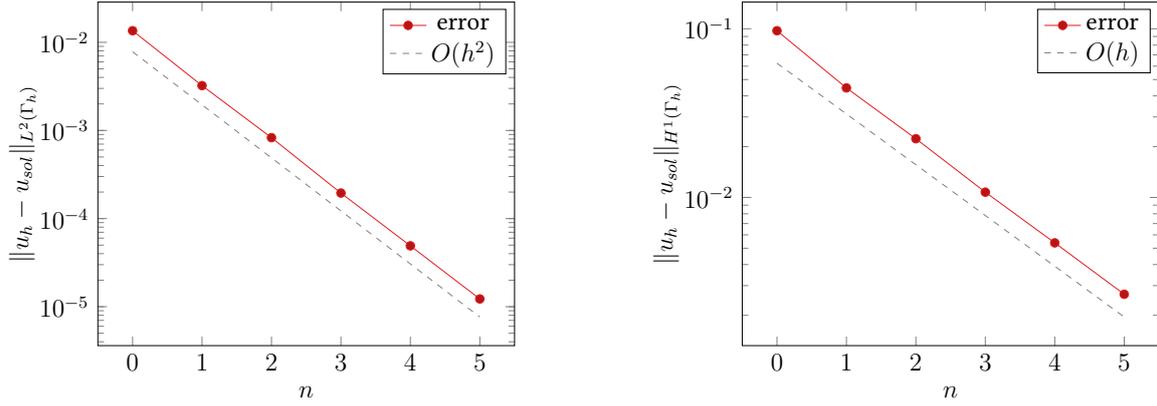


Figure 7: Convergence behaviour of Burman et al.'s method for the circle case.  $h_{\max} = 0.5^{n+2}$ .

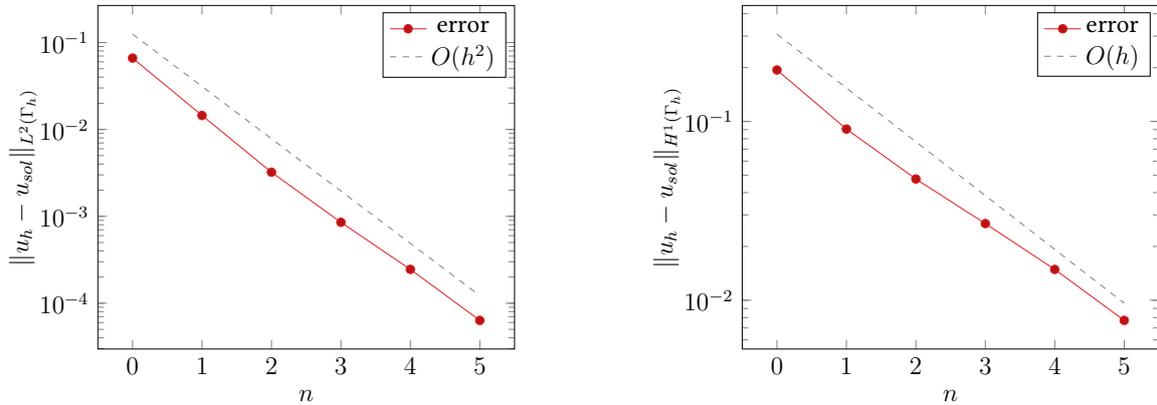


Figure 8: Convergence behaviour of Burman et al.'s method for the starfish case.  $h_{\max} = 0.5^{n+2}$ .

exemplify the theoretically suggested convergence rates of 2 in the  $L^2$  norm and 1 in the  $H^1$  norm. The numerical experiments are performed using `ngsxfem`<sup>12</sup>, which is an additional package for the finite element software `ngsolve`. [17]<sup>13</sup>

## 2.5 Hybrid DG in the plane

In our background example, the Poisson equation, it is possible to consider another variant of the presented Discontinuous Galerkin method (Subsection 2.1), namely a hybrid DG variant. The underlying idea goes as follows: If we consider how many degrees of freedom on adjacent elements are coupled with each other (which eventually lead to non-zero entries in the matrix of the discrete problem), we observe that all do that. To reduce those number of couplings

<sup>12</sup>Available at <https://github.com/ngsxfem/ngsxfem>

<sup>13</sup>Available at <https://ngsolve.org/>

one might introduce an additional polynomial function space. To be a little more precise, let  $\mathcal{T}_h$  denote the triangulation of the region  $\Omega$  and  $\mathcal{F}_h$  the facets of the corresponding mesh. Then the discrete function spaces are defined as

$$\begin{aligned} V_A &= \{u \in L^2(\Omega) \mid u|_T \in \mathcal{P}_k(T) \forall T \in \mathcal{T}_h\}, \\ V_B &= \{u \in L^2(S) \mid u|_{\partial\Omega} = 0 \text{ and } u|_F \in \mathcal{P}_k(F) \forall F \in \mathcal{F}_h\}, \\ V_h &= V_A \times V_B, \end{aligned}$$

where the skeleton  $S$  should be defined as

$$S = \bigcup_{F \in \mathcal{F}_h} F.$$

The discrete bilinear form corresponding to the interior penalty Discontinuous Galerkin method of Subsection 2.1, the hybrid(ised) interior penalty method, takes the form

$$\begin{aligned} A(u, \hat{u}; v, \hat{v}) &= \sum_{T \in \mathcal{T}_h} \int_T \nabla u \nabla v - \sum_{T \in \mathcal{T}_h} \int_{\partial T} n \nabla u (v - \hat{v}) - \sum_{T \in \mathcal{T}_h} \int_{\partial T} n \nabla v (u - \hat{u}) \\ &\quad + \frac{\alpha k^2}{h} \sum_{T \in \mathcal{T}_h} \int_{\partial T} (u - \hat{u})(v - \hat{v}). \end{aligned}$$

If we compare this to Equation (4), we observe first of all that the volume contribution remained unchanged. The facet integrals are now sorted differently; we are not any more summing over  $F \in \mathcal{F}_h$ , but instead over the boundaries of the  $T \in \mathcal{T}_h$ . Apart from that, we still have the consistency term scaled with  $\alpha k^2/h$  and the jump times normal gradient summands. Regarding the latter, we do not average but just take “our” element’s value, in some sense because evaluating the average again would make it impossible to evaluate the contributions from each element separately.

When this method is implemented and compared to the plain Discontinuous Galerkin method in the plane, the convergence results will be very similar. The difference is rather on the computational side: The number of non-zero entries in the final matrix scales differently for both variants; for the hybridised method we often start with more non-zero entries for first order polynomial approximation spaces. But when we increase the order, the number of non-zero entries will not rise as fast as in the non-hybridised case. We will demonstrate such a behaviour in a later subsection and keep it with this general remark here. Furthermore, the fact that in assembling the bilinear form of the hybrid problem only degrees of freedom on each element are needed locally, can be an advantage for parallelisation. But that of course

depends on the general implementation framework chosen.

For further details on this method, also about the analysis, I would like to refer the reader to [10].

## 2.6 Hybrid DG on an unfitted surface

Let us now transfer that procedure to the unfitted surface. First of all, we would like to keep our regions  $\mathcal{T}_h$ ,  $\mathcal{K}_h$ ,  $\mathcal{F}_h$ , and  $\mathcal{E}_h$  as well as the discrete interface  $\Gamma_h$  as they are. We start by again considering the same discontinuous polynomial function space for the volume part:

$$V_{h,A} = \bigoplus_{T \in \mathcal{T}_h} P_1(T) \quad \text{and} \quad V_{h,A,0} = \left\{ v \in V_{h,A} \mid \int_{\Gamma_h} v = 0 \right\}.$$

Furthermore, we introduce a function space on the facets:

$$V_{h,B} = \bigoplus_{F \in \mathcal{F}_h} P_1(F).$$

Since we have not only the boundary jump terms in the bilinear form  $a_h$ , but also the normal gradient jump terms in the stabilisation  $j_h$ , we now use a new function space to introduce a scalar variable for them:

$$V_{h,C} = \bigoplus_{F \in \mathcal{F}_h} P_0(F).$$

It is of one polynomial order less, since in the bilinear form it is the counterpart of the gradient of a function of polynomial order 1.

This eventually leads to the following definition:

$$V_{h,hyb} := V_{h,A,0} \times V_{h,B} \times V_{h,C}.$$

Now we are able to transfer the discrete bilinear form  $A_h$  of Equation (7) along the lines of the bilinear form for (H)DG in the plane:

$$\begin{aligned} A_h^{hyb}(u, \hat{u}, \hat{\sigma}; v, \hat{v}, \hat{\tau}) &= \sum_{K \in \mathcal{K}_h} \int_K \nabla_{\Gamma_h} u \cdot \nabla_{\Gamma_h} v - \sum_{K \in \mathcal{K}_h} \int_{\partial K} (n_E \cdot \nabla u)(v - \hat{v}) \\ &\quad - \sum_{K \in \mathcal{K}_h} \int_{\partial K} (n_E \cdot \nabla v)(u - \hat{u}) + \frac{\beta_E}{h} \sum_{K \in \mathcal{K}_h} \int_{\partial K} (u - \hat{u})(v - \hat{v}) \\ &\quad + \sum_{T \in \mathcal{T}_h} \int_{\partial T \cap S} \frac{\beta_F}{h^2} (u - \hat{u})(v - \hat{v}) + \int_{\partial T \cap S} \gamma (n_F \cdot \nabla u - \hat{\sigma}) \cdot (n_F \cdot \nabla v - \hat{\tau}), \end{aligned}$$

where again  $S = \bigcup_{F \in \mathcal{F}_h} F$ . But in this context (by the definition of  $\mathcal{F}_h$ ) this only selects the boundary facets which are adjacent to another element in the active mesh  $\mathcal{T}_h$ , see Fig. 4. One could also stabilise on the whole  $\partial\mathcal{T}_h$  with the latter terms and we actually implemented that “extended” method. However, it turns out that this extension is not as competitive to the DG method in terms of non-zero entries of the resulting matrix as the variant we chose. This is why we consider the restricted variant here.

If we now take into consideration the norms to show continuity and coercivity for the Burman et al. method (Equations (8) and (9)) and those of the Hybrid DG in the plane method ([10, Eq. (1.2.40), (1.2.41)]), we want to introduce the following norms according to that bilinear form here:

$$\| (u, \hat{u}, \hat{\sigma}) \|_{h,hyb}^2 := \|\nabla_{\Gamma_h} u\|_{\mathcal{K}_h}^2 + \frac{1}{h} \|u - \hat{u}\|_{\partial\mathcal{K}_h}^2 + \underbrace{\frac{1}{h^2} \|u - \hat{u}\|_{\partial\mathcal{T}_h \cap S}^2 + \|n_F \cdot \nabla u - \hat{\sigma}\|_{\partial\mathcal{T}_h \cap S}^2}_{=:\| (u, \hat{u}, \hat{\sigma}) \|_{\mathcal{F}_h,hyb}^2}$$

$$\begin{aligned} \| (u, \hat{u}, \hat{\sigma}) \|_{*,h,hyb}^2 &:= \|\nabla_{\Gamma_h} u\|_{\mathcal{K}_h}^2 + \frac{1}{h} \|u - \hat{u}\|_{\partial\mathcal{K}_h}^2 + \frac{1}{h^2} \|u - \hat{u}\|_{\partial\mathcal{T}_h \cap S}^2 + \|n_F \cdot \nabla u - \hat{\sigma}\|_{\partial\mathcal{T}_h \cap S}^2 \\ &+ h \|\nabla_{\Gamma_h} u\|_{\partial\mathcal{K}_h}^2. \end{aligned}$$

Now we can show—as in the subsection about the method of Burman et al.—the following lemmata:

**Lemma B<sub>hyb</sub>.1** The inequality  $\sqrt{2} \| (u, \hat{u}, \hat{\sigma}) \|_{\mathcal{F}_h,hyb} \geq \|u\|_{\mathcal{F}_h}$  holds.

**Proof** The norm  $\|u\|_{\mathcal{F}_h}$  was defined in terms of a sum over integrals over  $\mathcal{F}_h$ . To be able to compare it to the hybrid norm, we start by writing down the latter in a similar manner:

$$\begin{aligned} \| (u, \hat{u}, \hat{\sigma}) \|_{\mathcal{F}_h,hyb}^2 &= \sum_{T \in \mathcal{T}_h} \frac{1}{h^2} \|u - \hat{u}\|_{\partial T \cap S}^2 + \|n_F \cdot \nabla u - \hat{\sigma}\|_{\partial T \cap S}^2 \\ &= \sum_{F \in \mathcal{F}_h} \frac{1}{h^2} \|u^{left} - \hat{u}\|_F^2 + \|n_F \cdot \nabla u^{left} - \hat{\sigma}\|_F^2 \\ &\quad + \frac{1}{h^2} \|u^{right} - \hat{u}\|_F^2 + \|n_F \cdot \nabla u^{right} - \hat{\sigma}\|_F^2 \end{aligned}$$

If we start to write down  $\|u\|_{\mathcal{F}_h}^2$ , we observe

$$\|u\|_{\mathcal{F}_h}^2 = \sum_{F \in \mathcal{F}_h} \frac{1}{h^2} \|u^{left} - u^{right}\|_F^2 + \|n_F \nabla u^{left} - n_F \nabla u^{right}\|_F^2,$$

where we can apply Cauchy-Schwarz and Young's inequality:

$$\begin{aligned} \|u\|_{\mathcal{F}_h}^2 &\leq \sum_{F \in \mathcal{F}_h} \frac{2}{h^2} \cdot \left( \|u^{left} - \hat{u}\|_F^2 + \|u^{right} - \hat{u}\|_F^2 \right) + 2 \|n_F \nabla u^{left} - \hat{\sigma}\|_F^2 \\ &\quad + 2 \|n_F \nabla u^{right} - \hat{\sigma}\|_F^2 = 2 \| (u, \hat{u}, \hat{\sigma}) \|_{\mathcal{F}_h, \text{hyb}}^2. \quad \square \end{aligned}$$

**Corollary B<sub>hyb</sub>.2** (Hybrid variant of Lemma B.5.3) For  $(u, \hat{u}, \hat{\sigma}) \in V_h$ ,

$$h \|\nabla u\|_{N_h}^2 \lesssim h^2 \|\nabla_{\Gamma_h} u\|_{\mathcal{K}_h}^2 + \| (u, \hat{u}, \hat{\sigma}) \|_{\mathcal{F}_h, \text{hyb}}^2$$

**Proof** The statement follows immediately from the previous lemma and B.5.3.

**Corollary B<sub>hyb</sub>.3** (Hybrid variant of Corollary B.5.1) For  $h \in (0, h_0]$  with  $h_0$  sufficiently small, it is

$$h^{-1} \|u - \lambda_{\Gamma_h}(u)\|_{N_h}^2 \lesssim \|\nabla_{\Gamma_h} u\|_{\Gamma_h}^2 + \| (u, \hat{u}, \hat{\sigma}) \|_{\mathcal{F}_h, \text{hyb}}^2$$

**Proof** Follows immediately from the previous lemma and Corollary B.5.1.

**Lemma B<sub>hyb</sub>.4**  $\| (u, \hat{u}, \hat{\sigma}) \|_{h, \text{hyb}}$  defines a norm on  $V_{h, \text{hyb}}$ .

**Proof** There are three statements to show:

(1) If  $\| (u, \hat{u}, \hat{\sigma}) \|_{h, \text{hyb}} = 0$ , then  $0 = u = \hat{u} = \hat{\sigma}$ .

(2) For  $\alpha \in \mathbb{R}$  we have

$$\| (\alpha u, \alpha \hat{u}, \alpha \hat{\sigma}) \|_{h, \text{hyb}} = |\alpha| \cdot \| (u, \hat{u}, \hat{\sigma}) \|_{h, \text{hyb}}.$$

(3) For  $(u, \hat{u}, \hat{\sigma}), (v, \hat{v}, \hat{\tau}) \in V_{h, \text{hyb}}$ ,

$$\| (u + v, \hat{u} + \hat{v}, \hat{\sigma} + \hat{\tau}) \|_{h, \text{hyb}} \leq \| (u, \hat{u}, \hat{\sigma}) \|_{h, \text{hyb}} + \| (v, \hat{v}, \hat{\tau}) \|_{h, \text{hyb}}.$$

The properties (2) and (3) follow from the fact that each summand of our norm is a semi-norm. For example we have

$$\|\alpha u - \alpha \hat{u}\|_{\partial T \cap S} = |\alpha| \|u - \hat{u}\|_{\partial T \cap S},$$

or

$$\|u + v - \hat{u} - \hat{v}\|_{\partial T \cap S} \leq \|u - \hat{u}\|_{\partial T \cap S} + \|v - \hat{v}\|_{\partial T \cap S}.$$

So the most interesting point is actually (1). Assume

$$0 = \|\nabla_{\Gamma_h} u\|_{\mathcal{K}_h}^2 + \frac{1}{h} \|u - \hat{u}\|_{\partial \mathcal{K}_h}^2 + \|(u, \hat{u}, \hat{\sigma})\|_{\mathcal{F}_h, h_{yb}}^2.$$

This implies that each summand must vanish. From the previous lemma and the fact that the left- and right-hand side summands vanish, we arrive at

$$0 = \|u - \lambda_{\Gamma_h}(u)\|_{N_h}^2,$$

which implies  $0 = u - \lambda_{\Gamma_h}(u)$  a.e., since the  $L^2$ -norm actually is a norm, which implies  $0 = u$  by construction of  $V_{h, h_{yb}}$ . From

$$\begin{aligned} 0 = \|(u, \hat{u}, \hat{\sigma})\|_{\mathcal{F}_h, h_{yb}}^2 &= \sum_{F \in \mathcal{F}_h} \frac{1}{h^2} \|u^{left} - \hat{u}\|_F^2 + \|n_F \cdot \nabla u^{left} - \hat{\sigma}\|_F^2 \\ &\quad + \frac{1}{h^2} \|u^{right} - \hat{u}\|_F^2 + \|n_F \cdot \nabla u^{right} - \hat{\sigma}\|_F^2 \end{aligned}$$

in turn we can conclude  $0 = \hat{u} = \hat{\sigma}$ , since the  $L^2$ -norm on each facet is a norm, and we already observed that  $u^{left}, u^{right}$  and their gradients vanish.  $\square$

**Corollary B<sub>hyb</sub>.5**  $\|(u, \hat{u}, \hat{\sigma})\|_{*, h, h_{yb}}$  defines a norm on  $V_{h, h_{yb}}$ .

**Proof** For the reasons given in the previous proof, the important thing to show is that the purported norm is definite. But here, we are equipped with even stronger assumptions than in the previous proof.  $\square$

**Lemma B<sub>hyb</sub>.6** (Hybrid variant of Lemma B.5.4) For  $h \in (0, h_0]$  for  $h_0$  small enough it holds

$$h \|\nabla_{\Gamma_h} u\|_{\partial \mathcal{K}_h}^2 \leq C^{B_{hyb}.6} \left( \|\nabla_{\Gamma_h} u\|_{\Gamma_h}^2 + \|(u)\|_{\mathcal{F}_h, h_{yb}}^2 \right).$$

**Proof** Lemma B.5.3 together with the fact that the hybrid facet triple norm can be used as an upper bound for the non-hybrid facet triple norm.  $\square$

**Lemma B<sub>hyb</sub>.7** The norms  $\| (u, \hat{u}, \hat{\sigma}) \|_{h,hyb}$  and  $\| (u, \hat{u}, \hat{\sigma}) \|_{*,h,hyb}$  are equivalent on  $V_{h,hyb}$ .

**Proof** The previous Lemma can be applied here in the same manner as Lemma B.5.4 in the non-hybrid proof of Lemma B.5.\*4.

**Proposition B<sub>hyb</sub>.8**  $A_h^{hyb}$  is coercive and stable regarding the norm  $\| (u, \hat{u}, \hat{\sigma}) \|_{h,hyb}$ .

**Proof** We begin with coercivity. That means, it has to be shown that

$$A_h^{hyb}(u, \hat{u}, \hat{\sigma}; u, \hat{u}, \hat{\sigma}) \geq C^{B_{hyb}.8.1} \| (u, \hat{u}, \hat{\sigma}) \|_{h,hyb}^2 \quad \forall (u, \hat{u}, \hat{\sigma}) \in V_{h,hyb}.$$

To make everything a bit more convenient, we use the abbreviation  $\mathbf{u} = (u, \hat{u}, \hat{\sigma})$ . With that, we firstly just apply the definition of the bilinear form and identity the “facet” tripple norm:

$$A_h^{hyb}(\mathbf{u}, \mathbf{u}) \geq \|\nabla_{\Gamma_h} u\|_{\mathcal{K}_h}^2 - 2(n_E \nabla u, u - \hat{u})_{\partial\mathcal{K}_h}^2 + \frac{\beta_E}{h} \|u - \hat{u}\|_{\partial\mathcal{K}_h}^2 + \min\{\beta_F, \gamma\} \|\mathbf{u}\|_{\mathcal{F}_h,hyb}^2.$$

On the second summand, we apply again the Cauchy-Schwarz inequality with  $\epsilon$ . Everything works the same for  $\| \dots \|_{\partial\mathcal{K}_h}$  as it did for  $\| \dots \|_{\mathcal{E}_h}$  in the non-hybrid case, leading to

$$2(n_E \nabla u, u - \hat{u})_{\partial\mathcal{K}_h}^2 \leq 2\epsilon \|h^{1/2} n_E \nabla u\|_{\partial\mathcal{K}_h}^2 + \frac{1}{2\epsilon} \|h^{-1/2}(u - \hat{u})\|_{\partial\mathcal{K}_h}^2.$$

On the first summand we can again apply the parallel argumentation:

$$\begin{aligned} 2\epsilon \|h^{1/2} n_E \nabla u\|_{\partial\mathcal{K}_h}^2 &= 2\epsilon \|h^{1/2} n_E \nabla_{\Gamma_h} u\|_{\partial\mathcal{K}_h}^2 \leq 2\epsilon h \|\nabla_{\Gamma_h} u\|_{\partial\mathcal{K}_h}^2 \\ &\stackrel{B_{hyb}.6}{\leq} 2\epsilon C^{B_{hyb}.6} \left( \|\nabla_{\Gamma_h} u\|_{\Gamma_h}^2 + \|u\|_{\mathcal{F}_h,hyb}^2 \right). \end{aligned}$$

Putting all these parts together, we arrive at the statement with

$$C^{B_{hyb}.8.1} = \min \left\{ (1 - 2\epsilon C^{B_{hyb}.6}), \left( \beta_E - \frac{1}{2\epsilon} \right), (\min\{\beta_F, \gamma\} - 2\epsilon C^{B_{hyb}.6}) \right\}.$$

Regarding continuity we follow the same general strategy as in the non-hybrid case, namely showing boundedness regarding the  $*$ -norm and then applying the norm equivalence. Starting

with a summand-wise application of Cauchy-Schwarz in the bilinear form, we arrive at

$$\begin{aligned}
A_h^{hyb}(\mathbf{u}, \mathbf{v}) &\leq \|\nabla_{\Gamma_h} u\|_{\mathcal{K}_h} \cdot \|\nabla_{\Gamma_h} v\|_{\mathcal{K}_h} + \|n_E \nabla u\|_{\partial \mathcal{K}_h} \cdot \|v - \hat{v}\|_{\partial \mathcal{K}_h} \\
&+ \|n_E \nabla v\|_{\partial \mathcal{K}_h} \cdot \|u - \hat{u}\|_{\partial \mathcal{K}_h} + \frac{\beta_E}{h} \|u - \hat{u}\|_{\partial \mathcal{K}_h} \cdot \|v - \hat{v}\|_{\partial \mathcal{K}_h} \\
&+ \frac{\beta_F}{h^2} \|u - \hat{u}\|_{\partial \mathcal{T}_h \cap S} \cdot \|v - \hat{v}\|_{\partial \mathcal{T}_h \cap S} + \gamma \|n_F \nabla u - \hat{\sigma}\|_{\partial \mathcal{T}_h \cap S} \cdot \|n_F \nabla v - \hat{\tau}\|_{\partial \mathcal{T}_h \cap S}.
\end{aligned}$$

Here we again insert an  $1 = h^0 = h^{1/2} \cdot h^{-1/2}$  and apply the variant of Cauchy Schwarz about sums of products:

$$\begin{aligned}
A_h^{hyb}(\mathbf{u}, \mathbf{v}) &\leq \left( \|\nabla_{\Gamma_h} u\|_{\mathcal{K}_h}^2 + h \|n_E \nabla u\|_{\partial \mathcal{K}_h}^2 + \frac{1 + \beta_E}{h} \|u - \hat{u}\|_{\partial \mathcal{K}_h}^2 \right. \\
&\quad \left. + \frac{\beta_F}{h^2} \|u - \hat{u}\|_{\partial \mathcal{T}_h \cap S}^2 + \gamma \|n_F \nabla u - \hat{\sigma}\|_{\partial \mathcal{T}_h \cap S}^2 \right)^{1/2} \\
&\quad \left( \|\nabla_{\Gamma_h} v\|_{\mathcal{K}_h}^2 + h \|n_E \nabla v\|_{\partial \mathcal{K}_h}^2 + \frac{1 + \beta_E}{h} \|v - \hat{v}\|_{\partial \mathcal{K}_h}^2 \right. \\
&\quad \left. + \frac{\beta_F}{h^2} \|v - \hat{v}\|_{\partial \mathcal{T}_h \cap S}^2 + \gamma \|n_F \nabla v - \hat{\tau}\|_{\partial \mathcal{T}_h \cap S}^2 \right)^{1/2}
\end{aligned}$$

For each second term, we also apply the entity about  $n_E \nabla_{(\Gamma_h)} u$  to obtain

$$h \|n_E \cdot \nabla u\|_{\partial \mathcal{K}_h}^2 = h \|n_E \cdot \nabla_{\Gamma_h} u\|_{\partial \mathcal{K}_h}^2 \leq h \|\nabla_{\Gamma_h} u\|_{\partial \mathcal{K}_h}^2.$$

Together with the definition of the norm  $\|\cdot\|_{*,h,hyb}$ , we arrive at

$$A_h^{hyb}(\mathbf{u}, \mathbf{v}) \leq \max\{1 + \beta_E, \beta_F, \gamma\} \cdot \|\mathbf{u}\|_{*,h,hyb} \cdot \|\mathbf{v}\|_{*,h,hyb}.$$

Together with Lemma B<sub>hyb</sub>.7 the result holds.  $\square$

From those estimates one could go on to show that also the hybrid method converges with second order. However, again technical details of the interpolation operator—besides other things—would be required. Therefore we will skip a detailed proof and confirm numerically that the hybrid method actually converges with second order like the method by Burman et al.

**Numerical Example** Let us for a change consider a new geometry in 2D to test the hybrid method. Note that the remarks on the tweaked numerical problem and on the signed distance function assumption also apply here, as in all following subsections with numerical results.

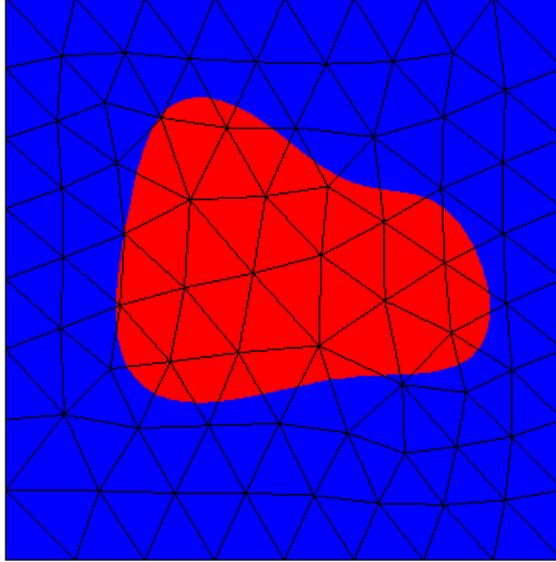


Figure 9: The potato geometry

The new geometry stems from the polynomial levelset function

$$\phi = (x - 1)^4 + (y - 1)^4 + 12 \cdot y + 1.4 \cdot (x - 2)^3 - 15$$

and mimics the shape of a potato when combined with  $\Omega = [-4, 4]^2$ . In Fig. 9 the resulting regions are shown. But of course, the method also works with the old geometries. In Fig. 10 the resulting numerical errors measured in the  $L^2$ - and  $H^1$ -norm on the interface are displayed. The convergence rates are—as expected—second order in the  $L^2$ -norm and first order in the  $H^1$ -norm. The absolute errors are of the same order of magnitude as those of the DG method. The stabilisation parameters have been chosen  $\beta_E = 10$ ,  $\beta_F = 100$ ,  $\gamma = 1$ . These choices were made on a heuristic basis, which is a drawback we will briefly discuss in the last section.

### 3 Higher order methods

In the previous section, we discussed a second order accurate method to solve the Laplace-Beltrami problem in two variants (plain Discontinuous Galerkin and Hybrid DG). It was bound to that order of convergence because of two reasons: Firstly, the geometry was approximated only with a polygonal and secondly, the discrete function space was only chosen to contain piecewise linear functions. This section is devoted to improving on both issues. First of all, we

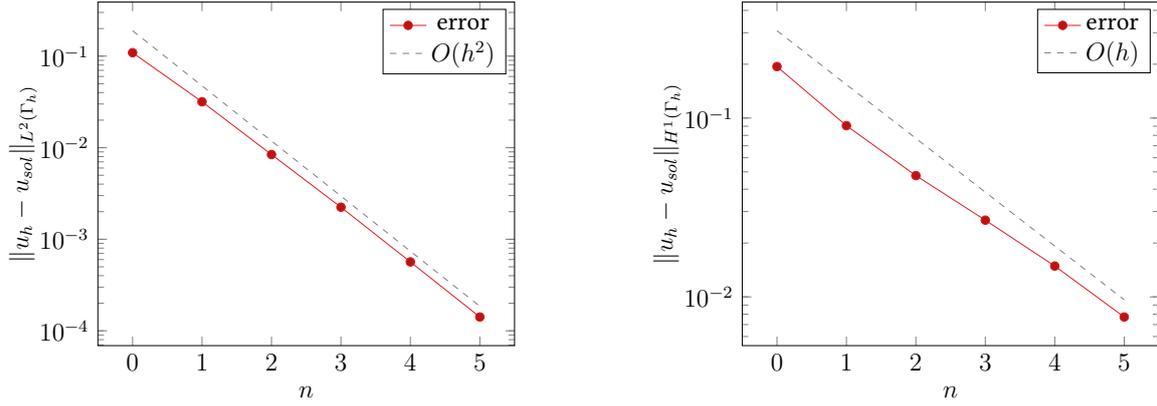


Figure 10: Convergence behaviour of the hybrid DG method for the potato geometry. The mesh is an unstructured simplicial mesh with  $h_{\max} = \frac{2}{3} \cdot 0.5^n$ .

will introduce the idea of an isoparametric mapping to enhance the approximation properties of our polygonal  $\Gamma_h$ . This will be discussed in the first subsection. In the following subsection, that improved approximated geometry will be exploited to reformulate the Discontinuous Galerkin method from the prior section. Afterwards, we will sketch how the resulting method can be shown to indeed feature higher order convergence. Finally, numerical examples illustrate the results and we will show how to transfer the ideas to a hybrid variant of the method.

### 3.1 The isoparametric mapping

The fact that the polygonal approximation of the surface  $\Gamma$ ,  $\Gamma_h$ , consisted of a straight line on each element facilitates the required numerical integration on this region severely. If we simply project the signed distance function  $\rho$  into a space of polynomials of higher order, there is in general no explicit representation of the surface (neither in total nor on each element), and a more complicated method would be needed to perform the numerical integration.<sup>14</sup> In this approach, we would like to exploit that desirable property of the polygonal approximation of the surface  $\Gamma_h$  and therefore use it as our starting point. Its approximation properties should then be improved with a bijective mapping  $\Psi: \Omega \rightarrow \Omega$  such that  $\Psi(\Gamma_h)$  is a higher order approximation of  $\Gamma$ .

What would be a desirable property of such a mapping  $\Psi$ ? In order to perfectly map  $\Gamma_h$  onto  $\Gamma$ , all points on  $\Gamma_h$  should be mapped to points on  $\Gamma$ . By definition of those regions, this

<sup>14</sup>This is not an unsolvable problem. See, e.g., [16] for an example of such an approach for quadrilaterals. But of course, these methods come with their own limitations and are more involved than the method arising from the explicit representation of the polygonal first-order approximation of the interface. For details on this procedure, see [11].

means that points with  $\rho_h(x) = 0$  should be mapped to  $\Psi(x)$  where  $\rho(\Psi(x)) = 0$ . That is just a special case of the following condition: Given a point  $x \in T \in \mathcal{T}_h$ , find a point  $y \in \Omega$  such that

$$\rho_h(x) = \rho(y).$$

This  $y$  should then be  $\Psi(x)$ . But obviously, this is not enough to define a unique image of the point  $x$ . Because if we are for instance on  $\Gamma_h$ ,  $x \in \Gamma_h$ , then any point  $y$  on the exact surface would satisfy our condition. Therefore we have to add a search direction. A natural choice would be to take the normal direction of the surface,  $G = \nabla\rho/\|\nabla\rho\|$ . Then the required condition reads: Find

$$\Psi(x) := y = x + d(x)G(x) \tag{16}$$

such that  $d(x)$  is the smallest (in absolute) value such that  $\rho_h(x) = \rho(y)$ .

Also with that improved condition, there remain some issues to solve. One of them is the observation that it sometimes will happen that the point we selected with the improved criterion lies in another element than  $T \ni x$ . In this case we would have to evaluate  $\rho$  on a point out of our local element  $T$ , which is undesired from a computational point of view. To circumvent this, we replace  $\rho$  with  $\mathcal{E}_T\rho_{h,*}$ , which should be the polynomial extension of  $\rho_{h,*}|_T$  to  $\mathbb{R}^d$ , and  $\rho_{h,*}$  should be a higher order polynomial approximation of  $\rho$ . Then the condition for the mapping  $\Psi_h$  reads: Given  $x \in T \in \mathcal{T}_h$ , find

$$\Psi_h(x) := y = x + d_h(x) \frac{\nabla\rho_{h,*}}{\|\nabla\rho_{h,*}\|},$$

where  $d_h(x)$  is the smallest (in absolute) number such that

$$\rho_h(x) = \mathcal{E}_T\rho_{h,*}(\Psi_h(x)).$$

The resulting transformation  $\Psi_h$  can now shown to be element-wise smooth, but will be discontinuous across element boundaries. But the jumps will be of higher order so that we can project  $\Psi_h$  into the space  $(V_h^{\text{cont}})^d$  of continuous element-wise polynomials of order  $k$  in each dimension. In order to do so, we apply a projection,  $P_h = P_h^1$ ,<sup>15</sup>

$$\Theta_h := P_h\Psi_h.$$

---

<sup>15</sup>For a volume problem, it is necessary to add another projection  $P_h^2$  here (see [13]). Therefore the seemingly unnecessary index 1 of the projection.  $P_h^2$  would be of type  $V_h^{\text{cont}}|_{N_h} \rightarrow V_h^{\text{cont}}$  and extend the transformation to the whole domain.

Here,  $P_h^1: C(\mathcal{T}_h) \rightarrow V_h^{\text{cont}}|_{N_h}$  projects (for each component) a function which is continuous on every element of the active background mesh into a function on the discrete space  $V_h^{\text{cont}}$  on the neighbourhood  $N_h$ . Note that the latter was by definition just the union of all elements of the active background mesh. In the next paragraph, the construction of  $P_h^1$  will be discussed in more detail.

**The projection  $P_h^1$**  Assume we are given a function  $v$  defined on  $N_h$  with  $v|_T \in C(T)$  for every  $T \in \mathcal{T}_h$ .<sup>16</sup> Further assume that we are equipped with a local interpolation operator which gives us the coefficients of the polynomial approximation on each element for some  $v|_T$ . Then the remaining problem is that we would arrive at conflicting values for degrees of freedom shared by adjacent elements. In order to resolve that, we simply take the average of those values. That idea—which is often called an Oswald type interpolation—could be formalised as follows: Let  $\{\phi_i\}_{i=1,\dots,N}$  be a basis of the polynomial space  $V_h^{\text{cont}}|_{N_h}$ . Then for each  $T \in \mathcal{T}_h$ , we have a set of degrees of freedom associated to the element  $T$ :

$$S_T := \{i \in \{1, \dots, N\} \mid \text{supp}(\phi_i) \cap \overset{\circ}{T} \neq \emptyset\}.$$

Then with the local interpolation operator we can find for each  $T \in \mathcal{T}_h$  coefficients  $c_{i,T}$  such that

$$\sum_{i \in S_T} c_{i,T} \phi_i|_T$$

approximates  $v|_T$ . The total discontinuous approximation of  $v$  can then be written as the sum of those local functions

$$\sum_{T \in \mathcal{T}_h} \sum_{i \in S_T} c_{i,T} \phi_i|_T.$$

But we can also collect the degrees of freedom differently, by defining for all degrees of freedom  $i$  the associated elements  $T$ ,

$$S_i := \{T \in \mathcal{T}_h \mid \text{supp}(\phi_i) \cap \overset{\circ}{T} \neq \emptyset\},$$

and then rearranging

$$\sum_{T \in \mathcal{T}_h} \sum_{i \in S_T} c_{i,T} \phi_i|_T = \sum_{i=1}^N \sum_{T \in S_i} c_{i,T} \phi_i|_T.$$

---

<sup>16</sup>The previously used notation  $C(\mathcal{T}_h)$  is a shorthand for that. See also Fn. 6f.

Now we simply take the average of those  $c_{i,T}$  for each  $i$  to arrive at

$$P_h^1 v := \sum_{i=1}^N \frac{\sum_{T \in \mathcal{S}_i} c_{i,T}}{\sum_{T \in \mathcal{S}_i} 1} \cdot \phi_i|_{N_h}.$$

The constructed mapping and the improved approximation of the surface,

$$\Gamma_{h,*} := \Theta_h(\Gamma_h)$$

satisfy several desirable properties, as it is shown, e.g., in [15]. Let us now collect some of these results.

**Lemma H.1** For  $h \in (0, h_0]$  for some small  $h_0$ , with  $\Theta_h$  being the mapping constructed in this subsection and  $p$  the closest point projection, it holds

$$\begin{aligned} \Theta_h(x) &= x \quad \text{for } x \text{ vertex in } \mathcal{T}_h \\ \|\Theta_h - \text{id}\|_\infty &\lesssim h^2, \quad \|D\Theta_h - I\|_\infty \lesssim h, \\ \|\Theta_h - p\|_{\infty, N_h} + h\|D(\Theta_h - p)\|_{\infty, N_h} &\lesssim h^{k+1} \quad \Rightarrow \quad \text{dist}(\Gamma, \Gamma_{h,*}) \lesssim h^{k+1}. \end{aligned}$$

Note that we deviate here a bit in the style of presentation from [13]. There, the Lemma is stated with regard to  $\Psi$ , the “optimal” mapping of Equation (16), while we used the closest point projection. That has the reason that [13] avoids the assumption of a signed distance function, while it was already made in the introduction in [4]. Therefore we also take such a signed distance function for granted.

A proof of those properties can be found in [15], Lemmata 3.4, 3.6, and 3.7. Briefly, the first equation states that the mapping  $\Theta_h$  is the identity on vertices. The last two equations quantify the deviation of the mapping from the identity and the closest point projection (and the deviations of the respective derivatives). Compared to the identity,  $\Theta_h$  decreases with  $h^2$ , since the transformation “repairs” the approximation error of  $\Gamma_h$ , which is also of second order. Furthermore,  $\Theta_h$  approximates the optimal mapping  $p$  with higher polynomial order. This can be seen as a consequence of the procedure of projecting  $\rho$  to  $\rho_h$  and that in turn into the polynomial space  $(V_h^{\text{cont}})^d$ .

## 3.2 Higher order DG method

We are now in a position to state the higher order Discontinuous Galerkin method. As in the low order case, the first interesting question is with regard to what geometric entities the

method will be introduced. Here, we also rely on the low order constructions and improve on them by means of the isoparametric mapping, i.e., we define

$$\begin{aligned}\mathcal{T}_{h,*} &= \{\Theta_h(T) \mid T \in \mathcal{T}_h\} & \mathcal{F}_{h,*} &= \{\Theta_h(F) \mid F \in \mathcal{F}_h\} \\ \mathcal{K}_{h,*} &= \{\Theta_h(K) \mid K \in \mathcal{K}_h\} & \mathcal{E}_{h,*} &= \{\Theta_h(E) \mid E \in \mathcal{E}_h\}.\end{aligned}$$

Associated to those regions are also normals which can be expressed in terms of the mapping  $\Theta_h$ . For example, the discrete surface normal  $n_h$  becomes

$$n_{h,*} := \frac{D\Theta_h^{-T} n_h}{\|D\Theta_h^{-T} n_h\|}.$$

The same transformation is applied to the facet and edge normals  $n_F$ , leading to

$$n_{F,*} := \frac{D\Theta_h^{-T} n_F}{\|D\Theta_h^{-T} n_F\|}, \quad n_{E,*} := (I - n_{h,*} \otimes n_{h,*}) n_{F,*}.$$

The improved discrete normal of course also induces an improved gradient operator, namely

$$\nabla_{\Gamma_{h,*}} u := (I - n_{h,*} \otimes n_{h,*}) \nabla u.$$

Furthermore, we define two tightly related projection operators:

$$P_{\Gamma_{h,*}} = (I - n_{h,*} \otimes n_{h,*}), \quad Q_{\Gamma_{h,*}} = n_{h,*} \otimes n_{h,*}.$$

Also, we introduce inner products on those regions as we did in the low order case. For instance

$$(u, v)_{\mathcal{T}_{h,*}} := \sum_{T \in \mathcal{T}_{h,*}} \int_T uv \, dx.$$

Note that a nearby way to actually calculate such an integral would consist in integrating over  $T \in \mathcal{T}_h$  and applying the transformation formula.

As we did with the inner products, we transfer our notation for norms, such that, e.g.,

$$\|u\|_{\mathcal{T}_{h,*}} := \sqrt{(u, u)_{\mathcal{T}_{h,*}}}, \quad \text{and so on.}$$

That improved geometric accuracy is the first ingredient to achieve higher order convergence.

The second is to consider higher order finite element spaces. Hence, we consider now

$$V_{h,prelim} := \bigoplus_{T \in \mathcal{T}_h} P_k(T),$$

and also map this function space to the improved discrete elements:

$$V_{h,*} := \{v \circ \Theta_h^{-1} \mid v \in V_{h,prelim}\} =: V_{h,prelim} \circ \Theta_h^{-1}.$$

That equation should mean that  $V_{h,*}$  is defined as the term in the middle and the term on the right-hand side introduces a shorthand notation we will exploit later on.

To assure the unique solvability of the problem, again we consider the subspace of functions with zero average:

$$V_{h,*,0} := \{v \in V_{h,*} \mid \lambda_{\Gamma_{h,*}}(v) = 0\}.$$

Now we can state the bilinear form of the higher order method. The jump and average notations are understood in terms of the transformed normals and regions as one would expect. Then we can basically rephrase

$$\begin{aligned} a_{h,*}(u, v) &= (\nabla_{\Gamma_{h,*}} u, \nabla_{\Gamma_{h,*}} v)_{\mathcal{K}_{h,*}} - (\{n_{E,*} \nabla u\}, [v])_{\mathcal{E}_{h,*}} - (\{n_{E,*} \nabla v\}, [u])_{\mathcal{E}_{h,*}} \\ &\quad + \frac{\gamma_E}{h} ([u], [v])_{\mathcal{E}_{h,*}}. \end{aligned}$$

Regarding the stabilisation we will directly transfer the two terms and add a third which operates on the whole  $N_{h,*} = \bigcup_{T \in \mathcal{T}_{h,*}} T$ :

$$j_{h,*}(u, v) = \frac{\gamma_{F,0}}{h^2} ([u], [v])_{\mathcal{F}_{h,*}} + \gamma_{F,1} (n_{F,*} [\nabla u], n_{F,*} [\nabla v])_{\mathcal{F}_{h,*}} + \frac{\gamma_n}{h} (n_{h,*} \nabla u, n_{h,*} \nabla v)_{\mathcal{T}_{h,*}}.$$

Note that the new stabilisation term can be regarded as an additional normal diffusion, which can be seen as a discrete normal extension to control functions on the active mesh. For that we note that a smooth function that vanishes on the surface can not be controlled without this additional stabilization. The right hand side of the problem is plainly

$$l_{h,*} = (f^e, v)_{\Gamma_{h,*}},$$

and the discrete problem becomes: Find  $u_h \in V_{h,*,0}$  such that

$$A_{h,*}(u_h, v_h) = l_{h,*}(v_h) \quad \forall v_h \in V_{h,*,0}, \quad \text{where} \quad A_{h,*}(u, v) := a_{h,*}(u, v) + j_{h,*}(u, v).$$

### 3.3 Elements of an analysis of the method

Proving the higher order convergence property of the introduced method proceeds similar to the proofs in the low order case from a distant standpoint: Two norms are introduced to bound the bilinear form from above and below. The equivalence of those norms for the discrete function space is shown. Then the unique solvability of the discrete problem follows from the Lax-Milgram Lemma.

Let us therefore again start by introducing suitable norms. The norm which takes the role of the norm introduced in (8) will be

$$\begin{aligned} \|v\|_{h,*}^2 := & \|\nabla_{\Gamma_{h,*}} v\|_{\mathcal{K}_{h,*}}^2 + \|h^{-1/2}[v]\|_{\mathcal{E}_{h,*}}^2 + \|h^{-1}[v]\|_{\mathcal{F}_{h,*}}^2 + \|n_{F,*}[\nabla v]\|_{\mathcal{F}_{h,*}}^2 \\ & + \|h^{-1/2}n_{h,*}\nabla v\|_{\mathcal{T}_{h,*}}^2. \end{aligned}$$

Note that the first four summands are the corresponding summands from the low order case (just with improved integration regions, normals, and gradient), and the fifth summand stems from the new stabilisation on  $N_{h,*}$ . Again, to facilitate the continuity proof another norm is introduced:

$$\|v\|_{*,h,*}^2 := \|v\|_{h,*}^2 + \|h^{1/2}n_{E,*} \cdot \nabla_{\Gamma_{h,*}} v\|_{\partial\mathcal{K}_{h,*}}^2.$$

Note that we inserted here an  $n_{E,*}$  in the new contribution, as mentioned in the remark on the respective low order norm.

Let us again collect the terms in the discrete energy norm associated to the stabilisation<sup>17</sup>:

$$\|v\|_{\mathcal{F}_{h,*}} := \|h^{-1}[v]\|_{\mathcal{F}_{h,*}}^2 + \|n_{F,*}[\nabla v]\|_{\mathcal{F}_{h,*}}^2 + \|h^{-1/2}n_{h,*}\nabla v\|_{\mathcal{T}_{h,*}}^2.$$

One important ingredient for the analysis will be the fact that the norms on the transformed and untransformed geometric entities of the discretisation are equivalent.

**Lemma H.2** For a function  $v \in H^1(N_h)$  we have

$$\begin{aligned} \|v\|_{\mathcal{T}_h} &\sim \|v \circ \Theta_h^{-1}\|_{\mathcal{T}_{h,*}} & \|v\|_{\mathcal{F}_h} &\sim \|v \circ \Theta_h^{-1}\|_{\mathcal{F}_{h,*}} \\ \|v\|_{\mathcal{K}_h} &\sim \|v \circ \Theta_h^{-1}\|_{\mathcal{K}_{h,*}} & \|v\|_{\mathcal{E}_h} &\sim \|v \circ \Theta_h^{-1}\|_{\mathcal{E}_{h,*}} \\ \|\nabla(v \circ \Theta_h^{-1})\|_2 &\sim \|\nabla v\|_2 & \text{a.e. in } & N_h. \end{aligned}$$

<sup>17</sup>Of course, the name  $\|\cdot\|_{\mathcal{F}_{h,*}}$  now really is an abuse of notation, as it already was in the hybrid low order case to a certain extent. But we would like to highlight the tight connection between the respective norms for the different methods.

For a proof of the left-hand side statements (which uses results we summarised in Lemma H.1) see [15, Lemma 3.18].

Then, the coercivity / continuity proof roughly goes as follows: A property similar to the Lemma B.5.2 in the low order case is taken from the literature for higher order continuous element-wise polynomials. It is extended to the case of discontinuous element-wise polynomials by exploiting a useful property of the aforementioned Oswald projection. Those results can be combined in a lemma which takes the role of Lemma B.5.4 in the low order case. Afterwards, the known estimation techniques are applied to show coercivity and continuity in the new norms. Throughout this proof, similar clues as in the low-order case are exploited. We note that especially the discrete inequalities mentioned in the Paragraph Section 4 in the analysis subsection of the low order method can be generalised and will be helpful.

To be concrete, the proof goes along the following lemmata:

**Lemma H.3** The Oswald projection, regarded as an operator  $\mathcal{O}_h: \mathcal{P}_{dc}^k(\mathcal{T}_h) \rightarrow \mathcal{P}^k(\mathcal{T}_h)$  with  $\mathcal{P}_{dc}^k$  being the space of discontinuous elementwise polynomials of order  $k$  and  $\mathcal{P}^k$  the space of continuous elementwise polynomials, satisfies for all  $v \in \mathcal{P}_{dc}^k(\mathcal{T}_h)$

$$\|v - \mathcal{O}_h(v)\|_T^2 \lesssim \sum_{F \in \mathcal{F}_h(T)} h \| [v] \|_F^2,$$

where  $\mathcal{F}_h(T)$  denotes the set of faces  $F \in \mathcal{F}_h$  with  $F \cap T \neq \emptyset$ .

**Lemma H.4** For  $v \in \mathcal{P}^k(\mathcal{T}_h)$

$$\begin{aligned} h^{-1} \|v\|_{\mathcal{T}_h}^2 &\lesssim \|v\|_{\mathcal{K}_h}^2 + h \|Q_{\Gamma_h} \nabla v\|_{\mathcal{T}_h}^2 \\ h^{-1} \|v - \lambda_h(v)\|_{\mathcal{T}_h}^2 &\lesssim \|\nabla_{\Gamma_h} v\|_{\mathcal{K}_h}^2 + h \|Q_{\Gamma_h} \nabla v\|_{\mathcal{T}_h}^2. \end{aligned}$$

**Lemma H.5** For  $v \in \mathcal{P}_{dc}^k(\mathcal{T}_h)$

$$\begin{aligned} h^{-1} \|v\|_{\mathcal{T}_h}^2 &\lesssim \|v\|_{\mathcal{K}_h}^2 + \|[v]\|_{\mathcal{F}_h}^2 + h \|Q_{\Gamma_h} \nabla v\|_{\mathcal{T}_h}^2 \\ h^{-1} \|v - \lambda_h(v)\|_{\mathcal{T}_h}^2 &\lesssim \|\nabla_{\Gamma_h} v\|_{\mathcal{K}_h}^2 + h^\beta \|[v]\|_{\mathcal{F}_h}^2 + h \|Q_{\Gamma_h} \nabla v\|_{\mathcal{T}_h}^2. \end{aligned}$$

**Lemma H.6** For  $v \in \mathcal{P}_{dc}^k(\mathcal{T}_h)$  we have

$$h \|n_E \cdot \nabla_{\Gamma_h} v\|_{\mathcal{E}_h}^2 \lesssim \|\nabla_{\Gamma_h} v\|_{\mathcal{K}_h}^2 + \|[v]\|_{\mathcal{F}_h, * }^2$$

Note that we are now able to transfer these results to the mapped integration regions with

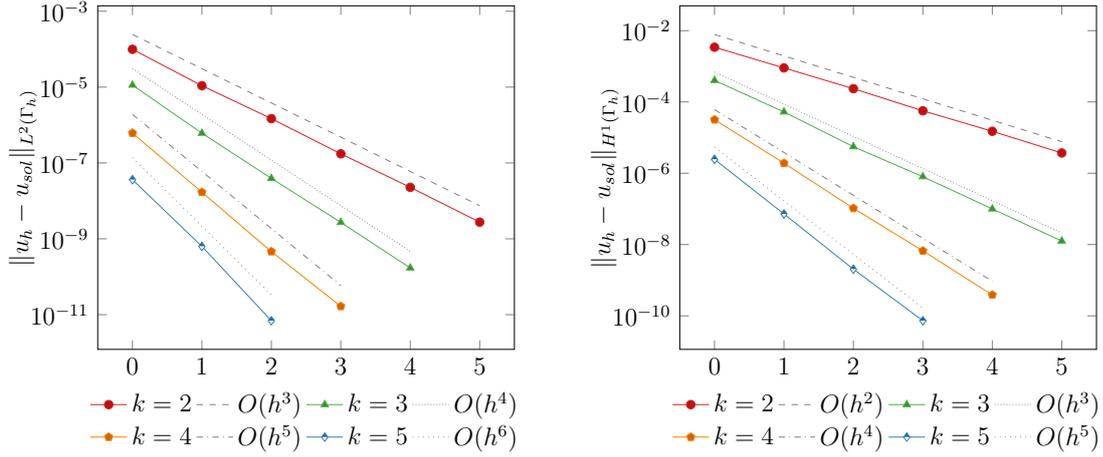


Figure 11: Convergence plots of the higher order method for the circle test case. Error measured in the  $L^2$ -/  $H^1$ -norm on the left / right hand side.

the help of Lemma H.2. There, we then can apply the usual techniques for the coercivity and continuity estimate.

Afterwards, the analysis needs to be continued in the direction of a Strang-type lemma and discrete error estimates. The details of that are an interesting open question for future research, as it is the question which value  $\beta$  has to be inserted in Lemma H.5.

### 3.4 Numerical examples

Let us now illustrate the method at some numerical investigations in 2D and 3D. Fortunately, the isoparametric mapping is already implemented in the software package `ngsxfem`, which will be also used in this subsection.

#### 3.4.1 2D

In two dimensions we again reconsider our examples from the previous section, i.e., the circle and the starfish. The respective results for the higher order method are shown in Figures 11 and 12 for  $k = 2, \dots, 5$ . One observes the higher order convergence property, namely that the error in the  $L^2$  decays like  $h^{k+1}$ , while in the  $H^1$  we have  $h^k$ . The stabilisation parameter were chosen as  $\gamma_E = 4(k+1)^2$  (scales with the polynomial degree  $k$ ),  $\gamma_{F,0} = 10$ ,  $\gamma_{F,1} = 0.001$ ,  $\gamma_n = h + 1/h$ . The threshold of the isoparametric mapping was 1. The maximal allowed meshsize was  $h_{max} = 0.5^{n+2}$ .

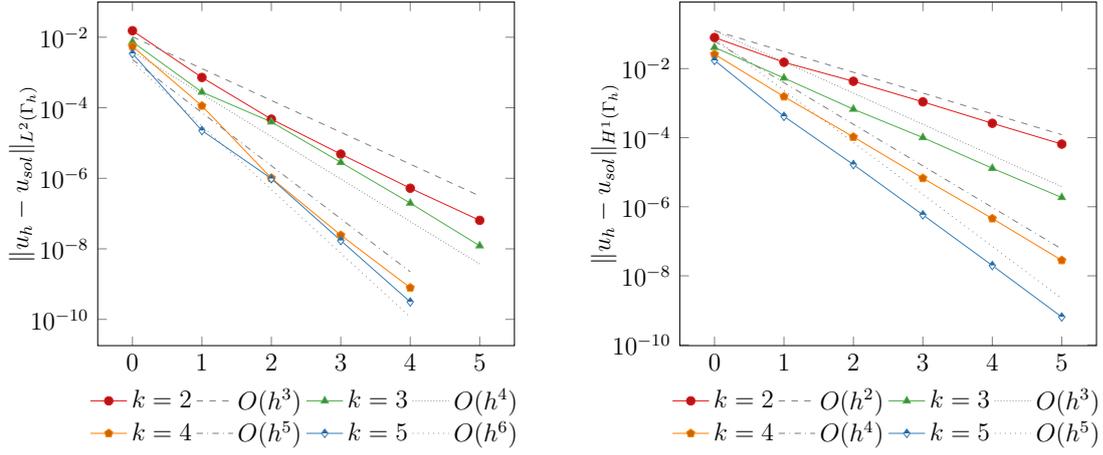


Figure 12: Convergence plots of the higher order method for the starfish test case. Error measured in the  $L^2$ -/  $H^1$ -norm on the left / right hand side.

### 3.4.2 3D

Let us now also consider some examples in three dimensions. We selected two geometries from the Burman et al. paper [4], namely the circle and the cheese geometry. They stem from the levelset functions

$$\begin{aligned}\phi_1 &= \sqrt{x^2 + y^2 + z^2} - R, \quad \text{and} \\ \phi_3 &= (x^2 - 1)^2 + (y^2 - 1)^2 + (z^2 - 1)^2 + (x^2 + y^2 - 4)^2 \\ &\quad + (x^2 + z^2 - 4)^2 + (y^2 + z^2 - 4)^2 - 16.\end{aligned}$$

Furthermore, we will additionally consider a torus geometry which can be generated from the levelset function

$$\phi_2 = \sqrt{z^2 + (\sqrt{x^2 + y^2} - R)^2} - r.$$

As the background region  $\Omega$  we consider  $[-l, l]^3$ , where  $l = 2$  for  $\phi_1$  and  $\phi_2$ , and  $l = 3$  for  $\phi_3$ . The right hand side is chosen such that it results in the following solutions  $u_1, u_2, u_3$  for the respective geometries:

$$\begin{aligned}u_1 &= \sin(\pi z) \quad u_2 = \sin\left(\frac{\pi x}{2}\right) \cdot \sin\left(\frac{\pi y}{2}\right) \cdot \sin\left(\frac{\pi z}{2}\right) \\ u_3 &= x \cdot y - 5y + z + x \cdot z\end{aligned}$$

Again, those functions are given to the program and the right-hand side is calculated by an automatic symbolic computation.

The geometries as well as some numerical solutions on them are shown in Fig. 13.

The observed numerical errors are depicted in Figures 14, 15, 16. As one would expect, the method converges with order  $h^k$  in the  $H^1$ -norm and  $h^{k+1}$  in the  $L^2$  norm. There are some deviations from this behaviour on coarse grids, but we are mostly interested in the asymptotical behaviour for  $h \rightarrow 0$ . Note that the maximal allowed meshsize was chosen to be  $0.5^{n+1}$ , where  $n$  is the refinement level. The stabilisation parameters were chosen  $\gamma_E = 4(k+1)^2$  (scales with the polynomial degree  $k$ ),  $\gamma_{F,0} = 10$ ,  $\gamma_{F,1} = 0.1$ ,  $\gamma_n = h + 1/h$ . The threshold of the isoparametric mapping was 1. Note that this latter threshold only influences the pre-asymptotic behaviour like described by Lehrenfeld in [12].

### 3.5 Higher order Hybrid DG method

Obviously, it is also possible to apply the idea of hybridisation—which we already know from the previous section—to the discontinuous method presented in the last subsection. Doing so will be the topic of this subsection.

To set up the discrete problem, we begin by introducing the discrete function spaces. First of all, our volume part remains unchanged:

$$V_{A,pre} := \bigoplus_{T \in \mathcal{T}_h} P_k(T), \quad V_{A,pre,0} = \{v \in V_{A,pre} \mid \lambda_{\Gamma_h}(v) = 0\}, \quad V_{A,*} = V_{A,pre,0} \circ \Theta_h^{-1}.$$

Correspondingly, the facet space also uses higher order polynomials which are then concatenated with the inverse isoparametric mapping:

$$V_{B,pre} := \bigoplus_{F \in \mathcal{F}_h} P_k(F), \quad V_{B,*} = V_{B,pre} \circ \Theta_h^{-1}.$$

For the facet normal gradient jump term again the gradient reduces the polynomial order by one and we can therefore use only one polynomial order less, leading to

$$V_{C,pre} := \bigoplus_{F \in \mathcal{F}_h} P_{k-1}(F), \quad V_{C,*} = V_{C,pre} \circ \Theta_h^{-1}.$$

The whole discrete function then again is plainly the product of those components:

$$V_{h,hyp,*} = V_{A,*} \times V_{B,*} \times V_{C,*}.$$

Now we come to the bilinear form of the discrete problem. Having in mind that the only “new” term in the higher order variant of the DG method was the normal gradient stabilisation

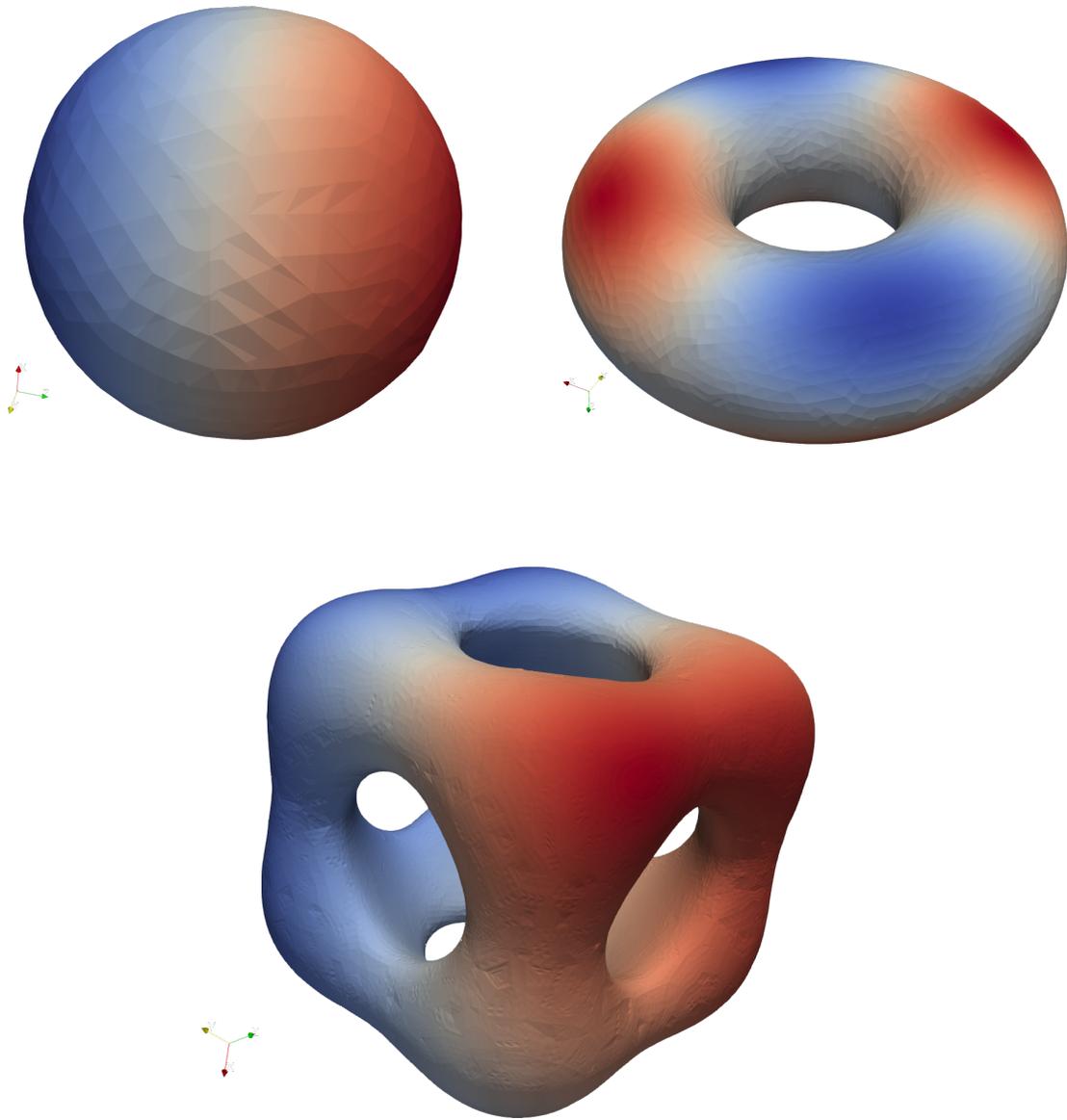


Figure 13: The three geometries in 3D with the respective solutions  $u$  plotted as colourmap. Note that the geometric approximation of the paraview isosurface was improved by the deformation from the isoparametric mapping.

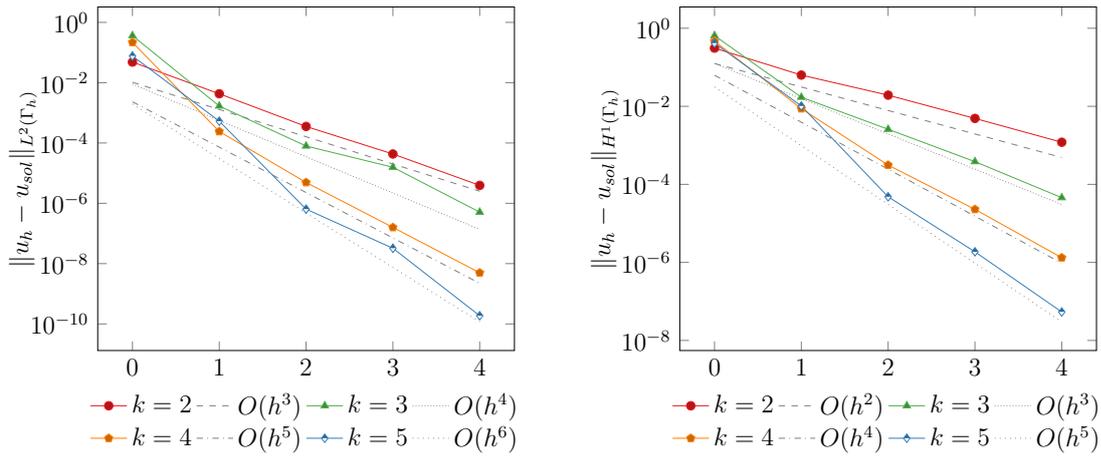


Figure 14: Numerical error for the higher order dG method at the sphere in 3D.

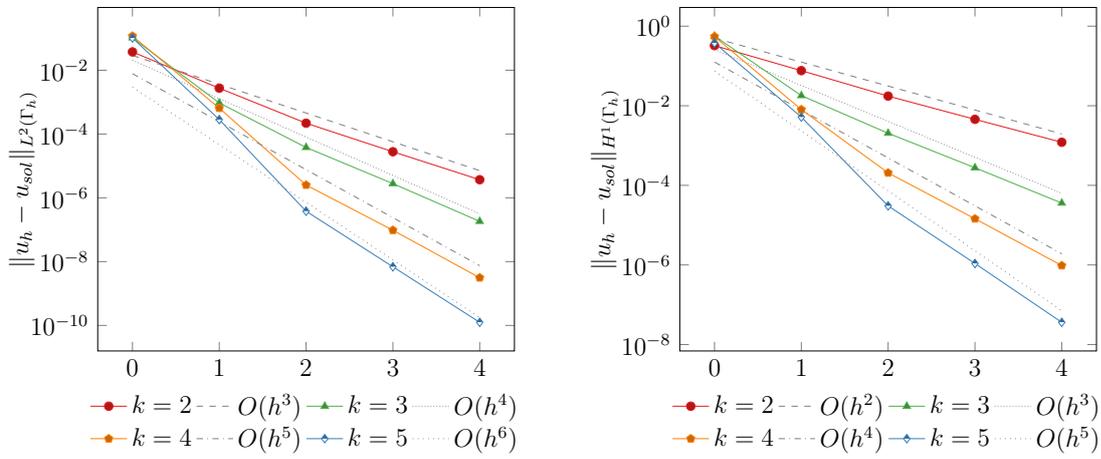


Figure 15: Numerical error for the higher order dG method at the torus in 3D.

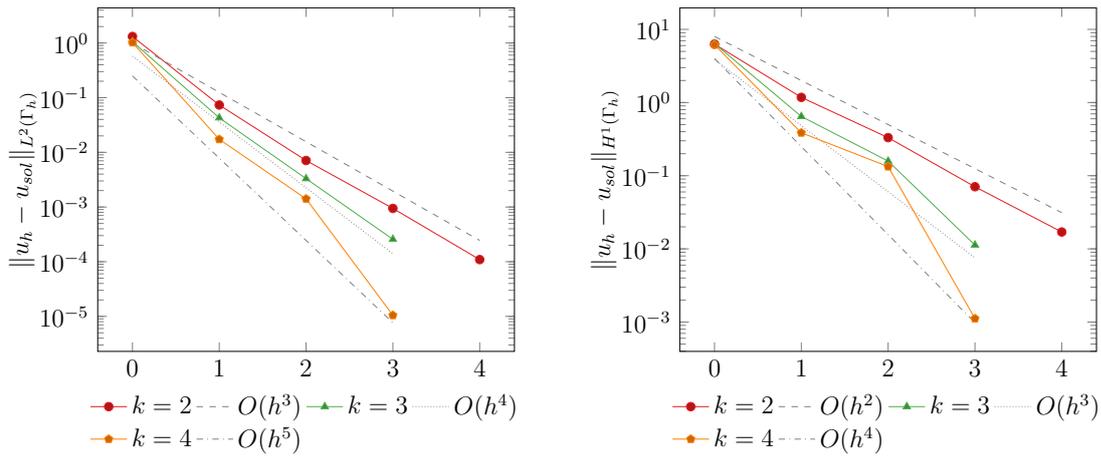


Figure 16: Numerical error for the higher order dG method at the cheese in 3D.

on the volume, we realise that the bilinear form of the low order hybrid method only needs minor adjustments:

$$\begin{aligned}
A_{h,*}^{hyb}(u, \hat{u}, \hat{\sigma}; v, \hat{v}, \hat{\tau}) &= \sum_{K \in \mathcal{K}_{h,*}} \int_K \nabla_{\Gamma_{h,*}} u \cdot \nabla_{\Gamma_{h,*}} v - \sum_{K \in \mathcal{K}_{h,*}} \int_{\partial K} (n_{E,*} \cdot \nabla u)(v - \hat{v}) \\
&- \sum_{K \in \mathcal{K}_{h,*}} \int_{\partial K} (n_{E,*} \cdot \nabla v)(u - \hat{u}) + \frac{\gamma_E}{h} \sum_{K \in \mathcal{K}_{h,*}} \int_{\partial K} (u - \hat{u})(v - \hat{v}) \\
&+ \sum_{T \in \mathcal{T}_{h,*}} \int_{\partial T \cap S_*} \frac{\gamma_{F,0}}{h^2} (u - \hat{u})(v - \hat{v}) + \int_{\partial T \cap S_*} \gamma_{F,1} (n_{F,*} \cdot \nabla u - \hat{\sigma}) \cdot (n_{F,*} \cdot \nabla v - \hat{\tau}) \\
&+ \sum_{T \in \mathcal{T}_{h,*}} \int_T \frac{\gamma_n}{h} n_{h,*} \cdot \nabla u \cdot n_{h,*} \cdot \nabla v.
\end{aligned}$$

The discrete variational problem then goes as follows: Find  $(u, \hat{u}, \hat{\sigma}) \in V_{h,hyb,*}$  such that

$$A_{h,*}^{hyb}(u, \hat{u}, \hat{\sigma}; v, \hat{v}, \hat{\tau}) = l_{h,*}(v) \quad \forall (v, \hat{v}, \hat{\tau}) \in V_{h,hyb,*}.$$

The analysis of this method should in principle not face major issues which do not also occur in the analysis of the hybrid low order/ non-hybrid high order method. However, we will not go into the details here and instead end this section with a numerical investigation of the convergence properties of the method.

### 3.6 Numerical examples

Let us now illustrate the method suggested in the previous subsection by some numerical examples. Again, all the calculations are performed with the xfem package.

Starting in two dimensions, we again reconsider the potato geometry. The maximal allowed meshsize was again chosen as  $h_{max} = \frac{2}{3}0.5^n$ . The stabilisation parameters were  $\gamma_E = 4(k+1)^2$  for  $k = 2, 3$  and  $\gamma_E = 0.1 \cdot (k+1)^2$  for  $k = 4, 5$ ,  $\gamma_{F,0} = 100$ ,  $\gamma_{F,1} = 1$ ,  $\gamma_n = \frac{0.1}{h} + 0.1 \cdot h$ . The threshold for the isoparametric mapping was 10.25. The numerical errors are plotted in Fig. 17. One can observe the order  $k$  in the  $H^1$ -norm and  $k+1$  in the  $L^2$ -norm as expected.

Coming to 3D, we again reconsider the torus test case. The stabilisations are chosen  $\gamma_E = 4 \cdot (k+1)^2$ ,  $\gamma_{F,0} = 100$ ,  $\gamma_{F,1} = 1$ ,  $\gamma_n = 0.1h + \frac{0.1}{h}$ . The threshold for the isoparametric mapping is 10.25. The results, which are given in Fig. 18 exemplify the expected convergence orders. Also, the absolute error of the hybrid method is comparable to the non-hybrid higher order method. Note that, however, for a detailed comparison of the numerical errors for the respective methods we refer the reader to Tab. 1, 2 since there both methods were applied on exactly the same problem with exactly the same parameters.

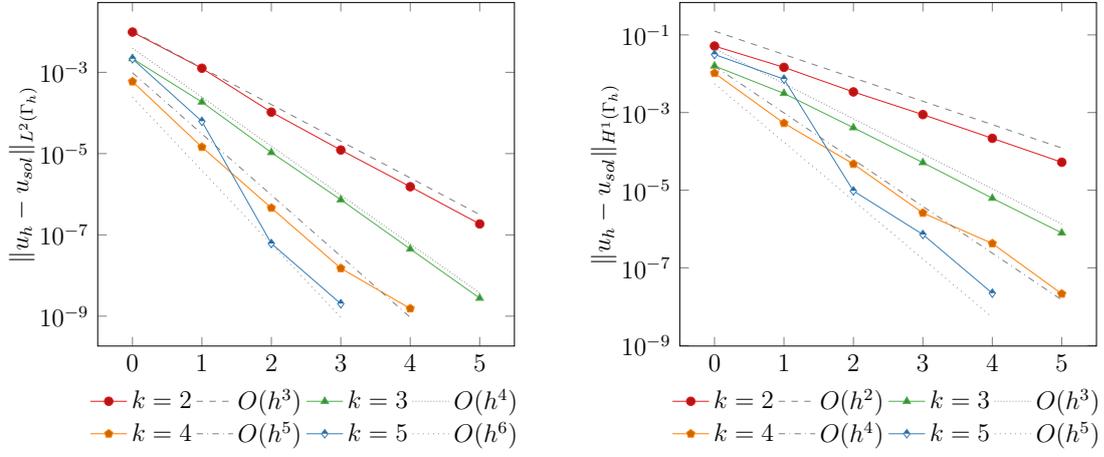


Figure 17: Observed numerical error of the higher order hybrid dG method with the potato geometry.

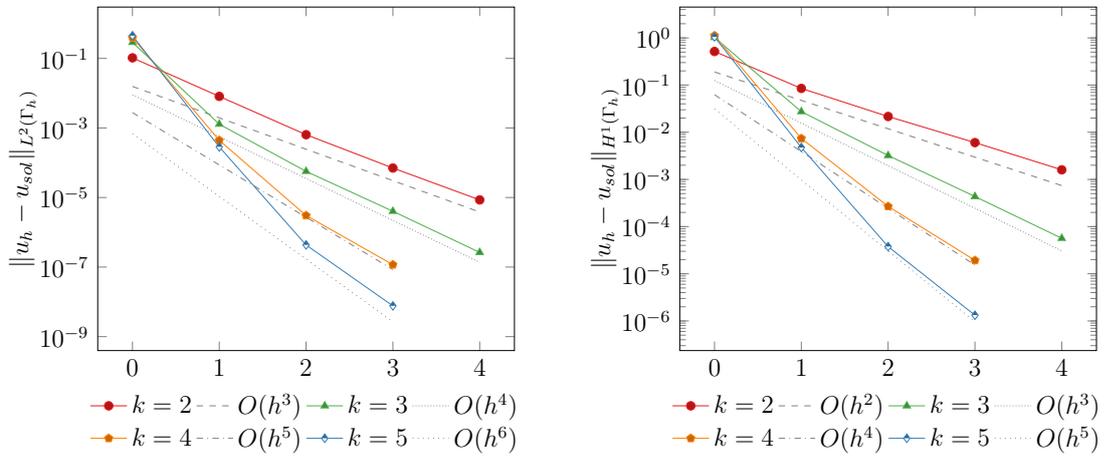


Figure 18: Observed numerical error of the higher order hybrid dG method with the torus geometry.

Let us now come to a comparison of the computational effort of the hybrid and non-hybrid higher order DG methods. A relevant part of the computational effort in terms of CPU time as well as memory consumption is in both methods the solution of the linear algebra problem. Therefore, a salient measure for the comparison of the computational effort is the number of non-zero entries in the resulting matrix. Furthermore, we also want to ensure that we are comparing respectively methods with the same accuracy. Therefore, we compare non-zero entries in the matrix,  $L^2$ -error and  $H^1$ -error for each of the methods.

There is a further computational clue to reduce the effort involved in a hybrid DG method, namely static condensation. This technique relies on the observation that the “volume” part of the matrix of the hybrid bilinear form is block-diagonal. Therefore we can further simplify the problem by applying a Schur-complement reduction. For details on this technique, the reader might want to consider [10]. From the computation perspective it is fully implemented in `ngsolve` and can be adopted without significant effort to our application.<sup>18</sup>

In Tables 1, 2 the described quantities are given for the three mentioned methods (DG, hybrid DG without condensation, hybrid with condensation). In 2D we consider again the potato test case and in 3D the sphere. First of all, note that the observations we will summarise in a moment do not depend on the mesh size  $h$ . However, to be sure about that, all results are given for two refinement levels. Regarding the numerical error measured in the  $L^2$ - and  $H^1$ -norm, we observe that the hybrid methods seem to result in slightly better results. However, the difference in absolute numbers is not very large and decreases for higher polynomial order. So, roughly, both methods lead to the same results in terms of the numerical error.

When we consider the number of matrix entries which are unequal zero, we observe the following result: The dG method always leads to less non-zero entries than the hybrid method without static condensation. The factor lies somewhere between one and two. But with the static condensation, the hybrid method can result in less non-zero entries. That generally is the case in 2D. In 3D, it can be only observed at  $k = 5$ . However, the clear trend is that an increase in the polynomial order is in favour of the hybrid method. For example, at  $k = 3$  in 2D, the hybrid method is a factor of two better, while we are at factor three to four at  $k = 5$ . Therefore, we expect that the hybrid method with condensation also in 3D at  $k > 5$  fares better. So it is especially an interesting method for (very) high orders.

To summarise these observations, we note that—measured in terms of the non-zero entries of the matrix—both the higher order dG method and its hybrid variant have ranges of applications where they perform better than the other variant. For dG, those are especially low to medium

---

<sup>18</sup>As long as the bilinear form actually is formulated in terms of an element-boundary integrator. But those integrators are a feature of `ngsxfem` the author implemented while working on this thesis.

	$h_{max} = 1/3$			$h_{max} = 1/6$		
	DG	HDG - C	HDG + C	DG	HDG - C	HDG + C
$k = 2$						
$L^2$ error	$1.5 \cdot 10^{-3}$	$1.3 \cdot 10^{-3}$		$1.2 \cdot 10^{-4}$	$1 \cdot 10^{-4}$	
$H^1$ error	$1.7 \cdot 10^{-2}$	$1.4 \cdot 10^{-2}$		$3.8 \cdot 10^{-3}$	$3.4 \cdot 10^{-3}$	
nze	12k	18k	8k	23k	34k	16k
$k = 3$						
$L^2$ error	$2 \cdot 10^{-4}$	$1.9 \cdot 10^{-4}$		$1.1 \cdot 10^{-5}$	$1.1 \cdot 10^{-5}$	
$H^1$ error	$3.4 \cdot 10^{-3}$	$3.2 \cdot 10^{-3}$		$4.5 \cdot 10^{-4}$	$4.1 \cdot 10^{-4}$	
nze	33k	43k	16k	64k	84k	31k
$k = 4$						
$L^2$ error	$1.4 \cdot 10^{-5}$	$1.4 \cdot 10^{-5}$		$3.9 \cdot 10^{-7}$	$3.8 \cdot 10^{-7}$	
$H^1$ error	$3.4 \cdot 10^{-4}$	$3.3 \cdot 10^{-4}$		$2.4 \cdot 10^{-5}$	$2.4 \cdot 10^{-5}$	
nze	74k	87k	27k	144k	169k	52k
$k = 5$						
$L^2$ error	$3.3 \cdot 10^{-6}$	$3.3 \cdot 10^{-6}$		$4.6 \cdot 10^{-8}$	$4.6 \cdot 10^{-8}$	
$H^1$ error	$6.6 \cdot 10^{-5}$	$6.6 \cdot 10^{-5}$		$3.1 \cdot 10^{-6}$	$3.1 \cdot 10^{-6}$	
nze	146k	155k	40k	283k	301k	78k

Table 1: Computational comparison of the dG, hybrid dG without static condensation, and hybrid dG with static condensation higher order methods at the potato in 2D. Displayed are respectively the numerical error and the number of non-zero entries in the arising matrix (nze).

	$h_{max} = 1/4$			$h_{max} = 1/8$		
	DG	HDG - C	HDG + C	DG	HDG - C	HDG + C
$k = 2$						
$L^2$ error	$7.3 \cdot 10^{-3}$	$7.0 \cdot 10^{-3}$		$8 \cdot 10^{-4}$	$5.9 \cdot 10^{-4}$	
$H^1$ error	$9.4 \cdot 10^{-2}$	$7.6 \cdot 10^{-2}$		$2.7 \cdot 10^{-2}$	$2.1 \cdot 10^{-2}$	
nze	115k	200k	201k	433k	754k	767k
$k = 3$						
$L^2$ error	$2.5 \cdot 10^{-3}$	$2.4 \cdot 10^{-3}$		$8.1 \cdot 10^{-5}$	$6.6 \cdot 10^{-5}$	
$H^1$ error	$3.5 \cdot 10^{-2}$	$3.0 \cdot 10^{-2}$		$4.2 \cdot 10^{-3}$	$3.6 \cdot 10^{-3}$	
nze	461k	722k	635k	1.7M	2.72M	2.4M
$k = 4$						
$L^2$ error	$5.7 \cdot 10^{-4}$	$5.7 \cdot 10^{-4}$		$5.9 \cdot 10^{-6}$	$5.8 \cdot 10^{-6}$	
$H^1$ error	$5.0 \cdot 10^{-3}$	$4.9 \cdot 10^{-3}$		$3.9 \cdot 10^{-4}$	$3.7 \cdot 10^{-4}$	
nze	1.58M	2.24M	1.76M	5.16M	7.3M	5.73M
$k = 5$						
$L^2$ error	$3.5 \cdot 10^{-5}$	$3.5 \cdot 10^{-5}$		$7.4 \cdot 10^{-7}$	$7.2 \cdot 10^{-7}$	
$H^1$ error	$1.4 \cdot 10^{-3}$	$1.4 \cdot 10^{-3}$		$6.1 \cdot 10^{-5}$	$5.7 \cdot 10^{-5}$	
nze	4.05M	5.24M	3.64M	13.2M	17.1M	11.9M

Table 2: Computational comparison of the dG, hybrid dG without static condensation, and hybrid dG with static condensation higher order methods at the sphere in 3D. Displayed are respectively the numerical error and the number of non-zero entries in the arising matrix (nze).

order calculations in 3D, while for hybrid dG the preferred applications are in 2D and of (very) high order in 3D. Of course, it would be interesting to consider more examples and see at which polynomial degree the hybrid method with condensation overtakes the dG method.

## 4 Conclusion and outlook

In this thesis, we surveyed different unfitted discontinuous Galerkin methods for solving the Laplace-Beltrami problem on a surface. To be able to properly introduce those methods, we began with a formulation of the Laplace-Beltrami problem in terms of mathematical concepts which are suited to our ends. Afterwards, we presented a low order accurate method by Burman et al. It was motivated as one instance of a larger class of methods—symmetric discontinuous Galerkin interior penalty methods—known from the planar case. The introduction of the discrete problem of this method was followed by an analysis of certain desirable properties it features. We mostly followed the paper [4] in doing so but chose a more detailed style of presentation for the coercivity and continuity proof. In that manner, we were able to shed some light on details which are at best only implicitly mentioned in the original paper. Regarding the following error estimates we only briefly sketched the proofs. Afterwards, we complemented the numerical demonstrations in the paper with two examples in 2D. They suggested that the estimate of second-order convergence in the  $L^2$ -norm and first order in the  $H^1$ -norm indeed is sharp. We finished the part of this thesis about low order methods with a hybrid variant of the method by Burman et al. Roughly, the bottom line there was that the general technique of hybridisation, as known from the planar case, can be also applied here, both in terms of analysis as of implementation.

Afterwards, we addressed higher order methods. As a crucial ingredient, we first introduced the isoparametric mapping by Lehrenfeld. It provides one with a higher order approximation of the surface, which then is the basis for the discrete function spaces. Here, higher order polynomials were chosen. After we applied those concepts to our method, some elements of an analysis were given. Finally, we considered several numerical examples in 2D and 3D, which exemplified the expected higher order of convergence. Also for this method, we considered a hybrid variant. However, we only focused on numerical investigations and left proofs of the salient statements for future investigations. The most important observation on the numerical side was that there are applications where the hybrid and the non-hybrid method respectively end up being computationally less expensive than their counterparts. In terms of the numerical error, the methods are comparable and lead to almost identical numbers.

Let us now eventually mention some open issues which go beyond the scope of this thesis and are therefore left for future research. Most notably, and as already mentioned, there are several things one could add to the respective analysis subsections. In general, the outline given in the analysis of the Burman paper, symbolically

$$\text{Coercivity} + \text{Continuity} \rightsquigarrow \text{Strang-type Lemma} \rightsquigarrow \text{Error bounds},$$

could be filled up, for all methods, to a level of detail which we were only able to achieve for the first step for the first method. Apart from that, the following open points seem most relevant:

**Necessity of the gradient stabilisation for higher order methods** In the transition from low-order to high-order methods, we introduced an additional stabilisation term, namely  $\gamma_n/h(n_h \nabla u, n_h \nabla w)_{\mathcal{T}_h}$ . In the analysis, a corresponding summand is added to the norm tailored for the coercivity and continuity proofs. This raises the question whether that new stabilisation makes the gradient jump stabilisation term  $\gamma_{F,1}(n_F[\nabla u], n_F[\nabla v])_{\mathcal{F}_h}$  unnecessary. There are different ways to approach this question, one being to simply disable the stabilisation in a numerical program. The author did this for a test case to trigger ill-conditioned cut configurations taken from [4, Section 8.2]: A sphere ( $R = 0.4$ ) is placed in a triangulated cube somewhat larger than the sphere ( $[-1.6, 1.6]^3$ ). The mesh is an unstructured tetrahedral mesh with  $h_{max} = 0.5^{n_{ref}}$ . Then the sphere is moved through the cube along the diagonal: To position the sphere around the point  $(\delta, \delta, \delta)$ , we choose the levelset function

$$\phi = \sqrt{(x - \delta)^2 + (y - \delta)^2 + (z - \delta)^2} - R.$$

Then the Laplace-Beltrami problem presented in the previous sections is solved for a parameter  $\delta \in [-1, 1]$ . The numerical error in the  $L^2$ - and  $H^1$ -norm of the discrete interface are considered for different mesh refinements levels. In the paper [4], the condition number of the matrix was plotted. However, the stabilisation with constant  $\gamma_{F,1}$  is not only important for the boundedness of the condition number, but also for the coercivity and therefore the well-posedness of the discrete problem. Therefore we can also investigate the necessity of the  $\gamma_{F,1} > 0$  by considering the numerical error.<sup>19</sup> The results for this test are depicted in Figures 19, 20, 21. We observe that the numerical error seems relatively stable and no problematic cases for  $\gamma_{F,1} = 0$  can be triggered with this test. This raises the question whether the stabilisation term can be also dropped in different cases, which could be investigated numerically, or even

---

<sup>19</sup>Although it of course would also be an interesting further investigation to consider the condition number.

in general, which could be shown by a corresponding theoretical estimate. If it would turn out that the term is actually not needed, that would be beneficial for the hybrid method, since then only one additional facet function space would be required. That is one reason why this open question is an interesting one.

The plots also illustrate another issue discussed in the second to last of these paragraphs, namely the problem on finding suitable values for the stabilisation constants. Regarding  $\gamma_{F,1}$ , we can see that choosing the parameter too large can result in an additional numerical error. Therefore it would clearly not be advisable to choose  $\gamma_{F,1}$  plainly as something like  $10^5$  to ensure that the parameter is “sufficiently large”, as required by the analysis.

**Quadrilateral Finite Elements** Because of a flexible implementation we presented elsewhere (see [8]) it is possible to migrate from the simplicial meshes employed throughout this thesis to quadrilateral /hexahedral meshes by changing one line in the respective python scripts. That allows for a brief outlook on the question how the methods of this thesis work on quadrilateral /hexahedral meshes. As an example, we consider the two dimensional higher order DG method with  $k = 4$  for the circle geometry. We consider structured and unstructured meshes. The results are given in Fig. 22.

We can observe that in the case of structured meshes the method roughly converges with the right order. However, there is one refinement level which deteriorates the numerical solution, which is surprising. In the case of an unstructured quadrilateral mesh, no clear convergence trend can be observed. It is an interesting open question to investigate which improvements of the stabilisation etc. are needed to transfer the methods to quadrilateral meshes. Note that not much effort yet went into testing the relevant code, such that future research should also include the possibility of implementation mistakes.

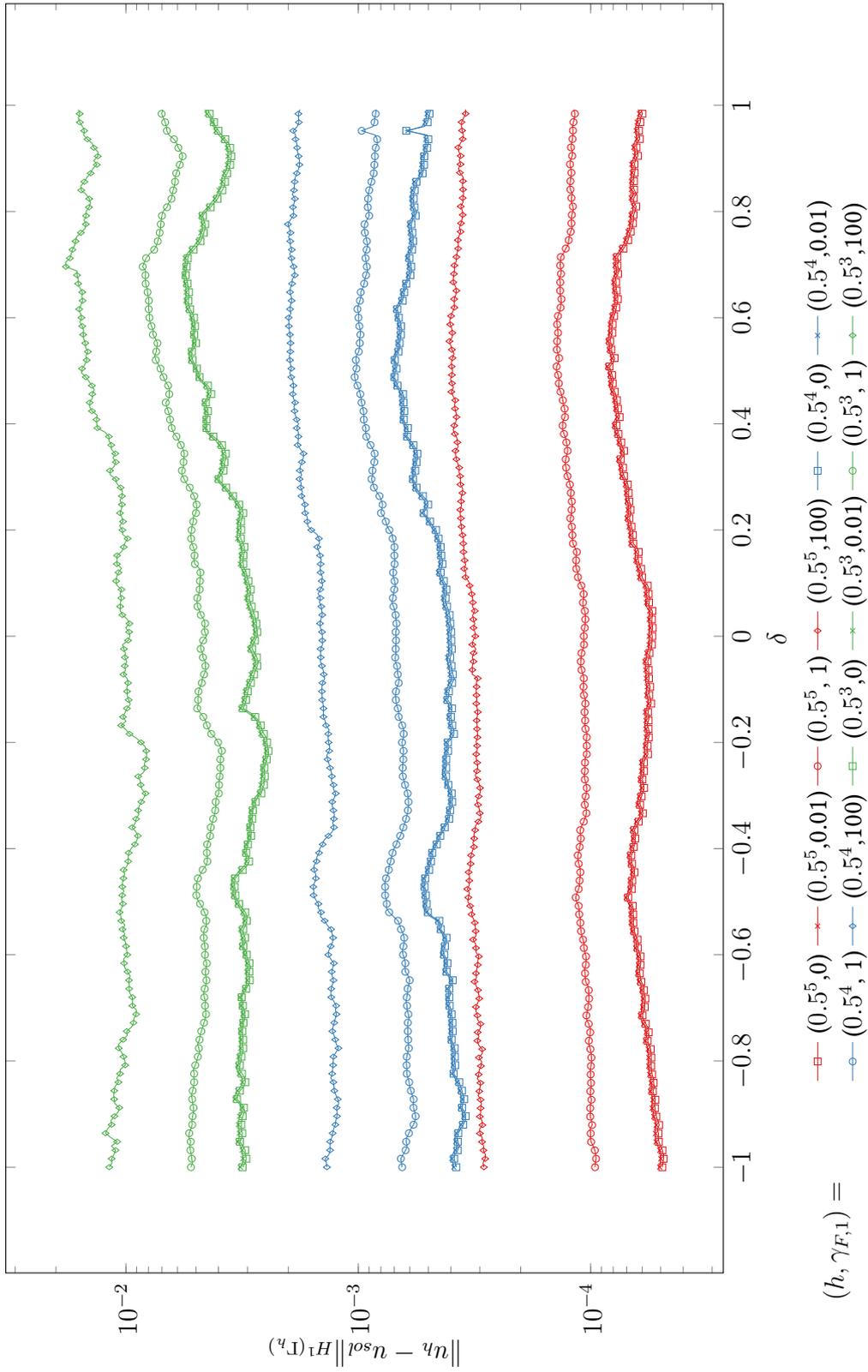


Figure 19: Results of the moving sphere test case for  $k = 3$  in the  $H^1$ -norm. Note that to increase the readability of this plot, only each second data point is shown.

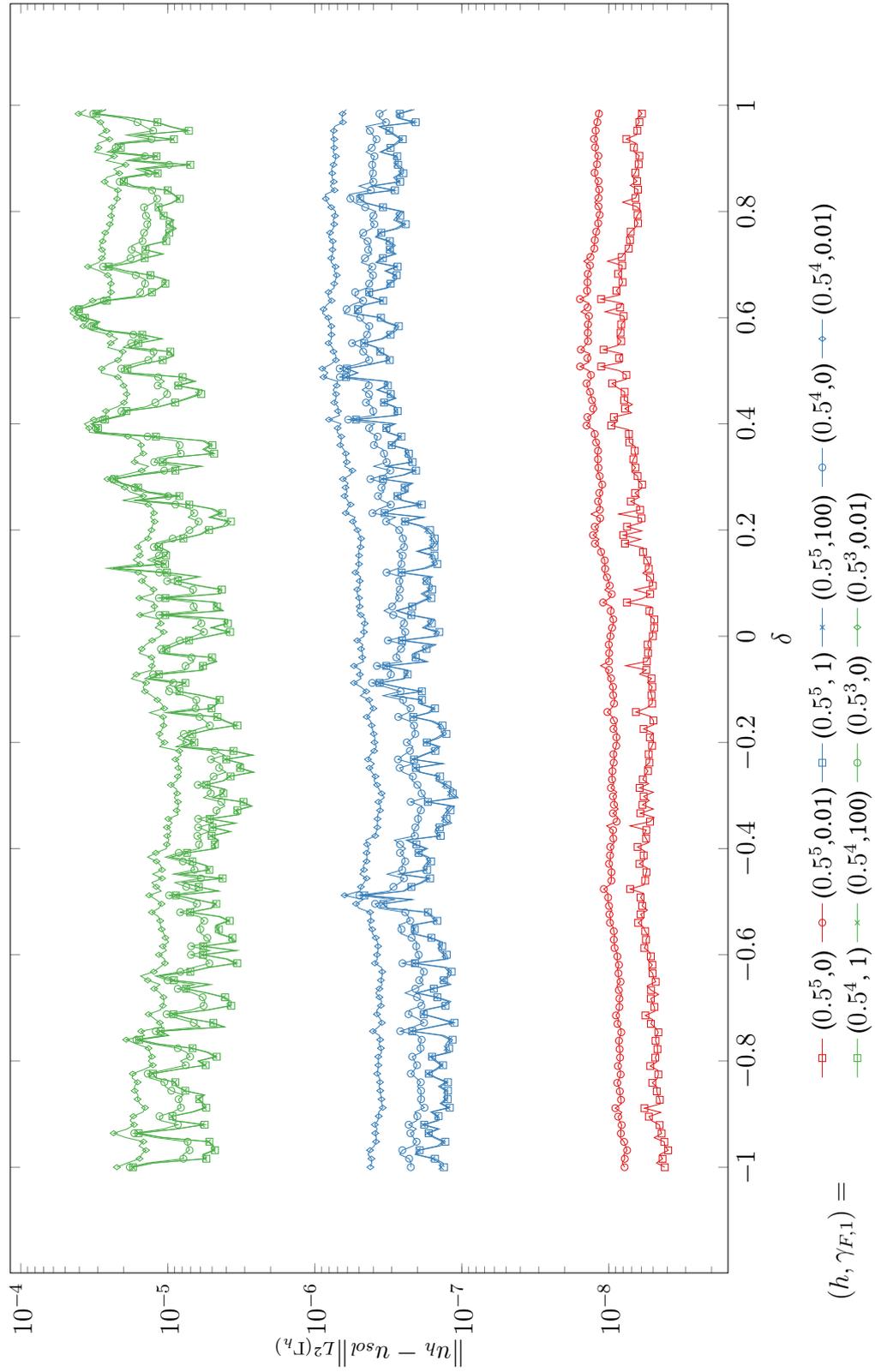


Figure 20: Results of the moving sphere test case for  $k = 4$  in the  $L^2$ -norm. Note that to increase the readability of this plot, only each second data point is shown.

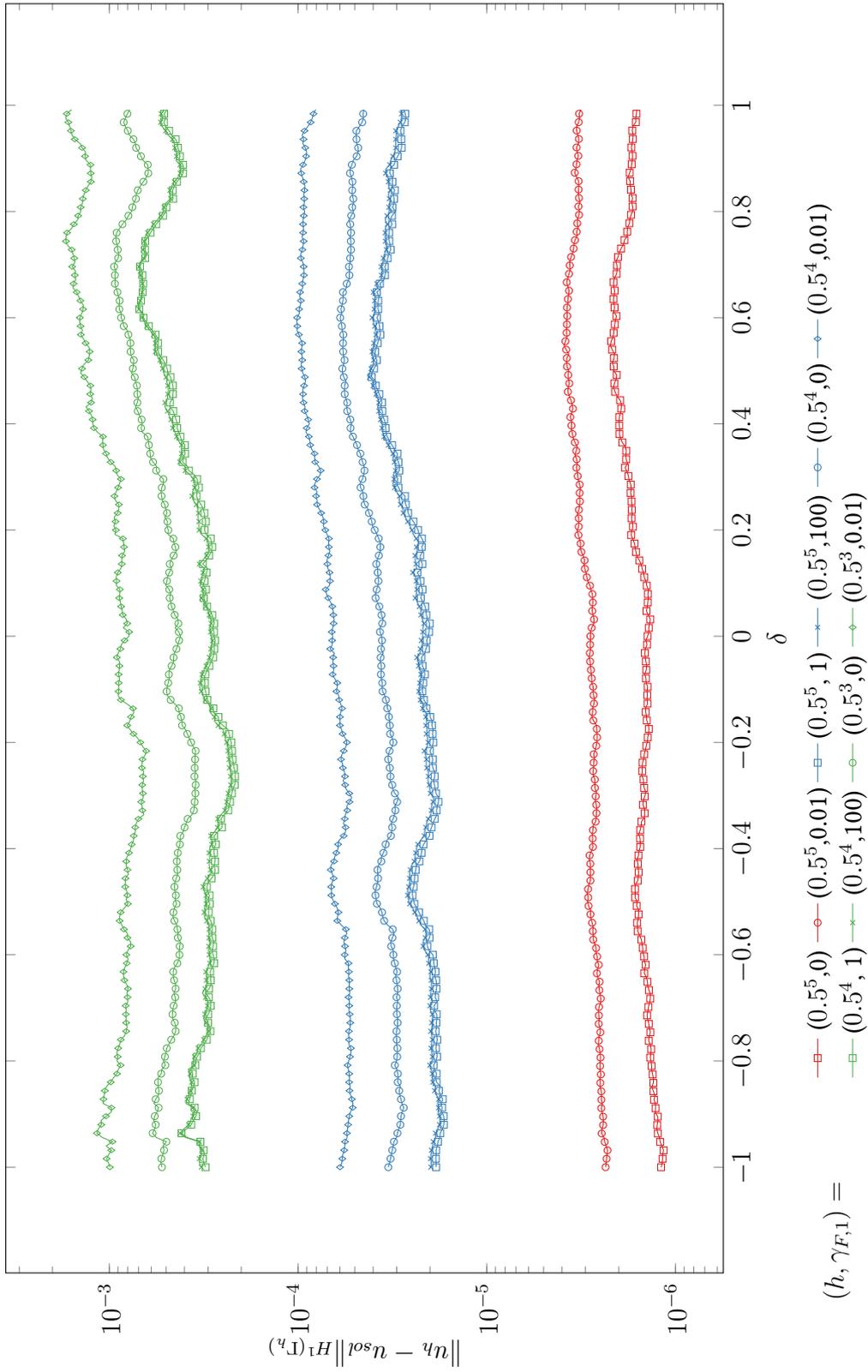


Figure 21: Results of the moving sphere test case for  $k = 4$  in the  $H^1$ -norm. Note that to increase the readability of this plot, only each second data point is shown.

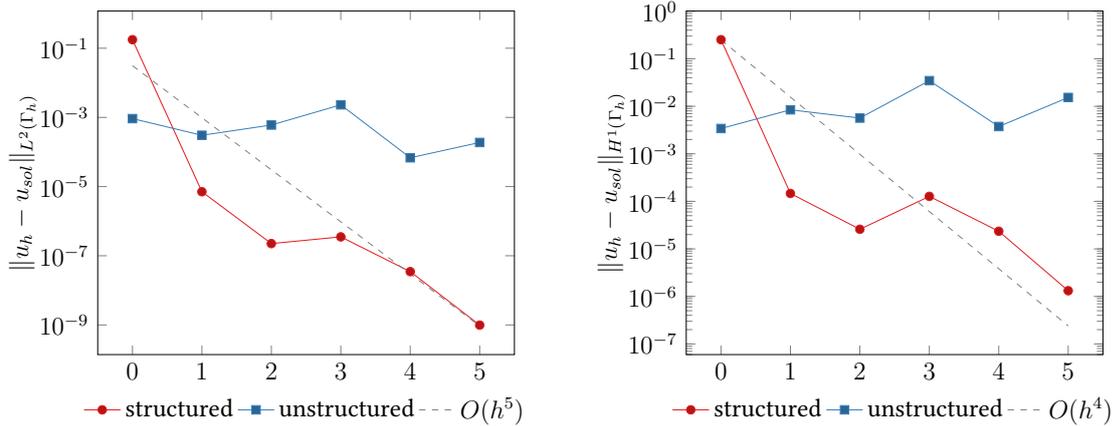


Figure 22: The circle geometry with the high order DG method ( $k = 4$ ) on a structured (red) and unstructured (blue) quadrilateral mesh.

**Iterative Solvers and Preconditioning** All the numerical results within this thesis were obtained with a direct solver. That has the benefit of not having to consider the issue of setting up an iterative solver with an appropriate preconditioner. But on the other hand, it has the downside that only systems up to a certain difficulty or number of elements or degrees of freedom can be considered because of finite (memory) resources. For example, the test with the cheese geometry could be only executed for a limited polynomial order respectively number of refinements. To go further there, it would be interesting to try out iterative solvers and preconditioners on the presented methods.

**Stabilisation constants** In the different presented methods, we often made use of stabilisation terms, which scale with constants like  $\beta_E$ . Choosing these constants right is sometimes a complicated issue since there are no clear predictions from the analysis which numerical values should be chosen. On the other hand, the method can severely fail to converge if, e.g., a crucial parameter is too small. That motivates the question whether one could replace those stabilisation parameters with techniques like lifting, known from DG discretizations in the plane. It is however not obvious how these lifting strategies can be carried over.

**Vector valued problems** With the Laplace-Beltrami equation, we considered a simple scalar valued problem. However, many parts of the analysis and the method itself could be extended to vector valued problems. Interesting research in that direction has been done, e.g., in [7]. A salient open question would now be how the results of our investigations could be transferred to that case.

## References

- [1] Douglas N. Arnold, Franco Brezzi, Bernardo Cockburn, and L. Donatella Marini. Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM Journal on Numerical Analysis*, 39(5):1749–1779, 2002.
- [2] Erik Burman and Peter Hansbo. Fictitious domain finite element methods using cut elements: II. A stabilized Nitsche method. *Applied Numerical Mathematics*, 62(4):328–341, 2012.
- [3] Erik Burman, Peter Hansbo, and Mats G. Larson. A stabilized cut finite element method for partial differential equations on surfaces: The laplace–beltrami operator. *Computer Methods in Applied Mechanics and Engineering*, 285:188–207, 2015.
- [4] Erik Burman, Peter Hansbo, Mats G. Larson, and André Massing. A cut discontinuous Galerkin method for the Laplace–Beltrami operator. *IMA Journal of Numerical Analysis*, 37(1):138–169, 2017.
- [5] Daniele Antonio Di Pietro and Alexandre Ern. *Mathematical aspects of discontinuous Galerkin methods*, volume 69. Springer Science & Business Media, 2011.
- [6] David Gilbarg and Neil S. Trudinger. Classics in mathematics. *Elliptic partial differential equations of second order*, 2001.
- [7] Sven Gross, Thomas Jankuhn, Maxim A. Olshanskii, and Arnold Reusken. A trace finite element method for vector-laplacians on surfaces. *SIAM Journal on Numerical Analysis*, 56(4):2406–2429, 2018.
- [8] Fabian Heimann and Christoph Lehrenfeld. Numerical Integration on Hyperrectangles in Isoparametric Unfitted Finite Elements. *forthcoming in Proceedings of ENUMATH 2017*, 2018.
- [9] Jan S. Hesthaven and Tim Warburton. *Nodal discontinuous Galerkin methods: algorithms, analysis, and applications*. Springer Science & Business Media, 2007.
- [10] Christoph Lehrenfeld. Hybrid discontinuous Galerkin methods for solving incompressible flow problems. *Diploma Thesis, MathCCES/IGPM, Rheinisch-Westfälische Technische Hochschule Aachen*, 2010.

- [11] Christoph Lehrenfeld. *On a Space-Time Extended Finite Element Method for the Solution of a Class of Two-Phase Mass Transport Problems*. PhD thesis, RWTH Aachen, February 2015.
- [12] Christoph Lehrenfeld. High order unfitted finite element methods on level set domains using isoparametric mappings. *Computer Methods in Applied Mechanics and Engineering*, 300:716–733, 2016.
- [13] Christoph Lehrenfeld. A Higher Order Isoparametric Fictitious Domain Method for Level Set Domains. In *Geometrically Unfitted Finite Element Methods and Applications*, pages 65–92. Springer, 2017.
- [14] Christoph Lehrenfeld. Numerics of partial differential equations: Lecture notes (summer semester 2017). July 10, 2017.
- [15] Christoph Lehrenfeld and Arnold Reusken. Analysis of a high-order unfitted finite element method for elliptic interface problems. *IMA Journal of Numerical Analysis*, 38(3):1351–1387, 2018.
- [16] RI Saye. High-order quadrature methods for implicitly defined surfaces and volumes in hyperrectangles. *SIAM Journal on Scientific Computing*, 37(2):A993–A1019, 2015.
- [17] Joachim Schöberl. C++ 11 Implementation of Finite Elements in NGSolve. *ASC Report*, (30), 2014.

## Acknowledgements

The work on this thesis was conducted in a supportive and motivating environment. I am grateful for that and would like to thank everybody who contributed to it.

Most importantly, I would like to thank Christoph Lehrenfeld for his great support in advising this thesis. His passion for the topics at stake not only manifested in him answering my questions more or less at any time on each day of the week but also extended to my own working process. Apart from him suggesting the topic in general, I benefited from several discussions of details of this text and more implementation-related issues.

Furthermore, I thank Prof. Gert Lube for agreeing to assess this thesis, as well as him and Christoph for inviting me to give a talk about it in their Oberseminar. Also, thanks for the suggestions of all the other participants there.

Some weeks earlier, André Massing also gave a tightly related talk in the Oberseminar, and I am grateful for interesting discussions throughout his visit in Göttingen.

Fabian Chlumsky-Hartmann provided me with a multitude of language and typography improvements, which I thankfully took into consideration.

Last but not least, I thank the development team of Ngsolve, most notably Joachim Schöberl, for providing such a helpful software package. I am sure that most of the numerical experiments of this thesis would have been far beyond of its scope without Ngsolve.

## **Selbstständigkeitserklärung**

Hiermit versichere ich, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Göttingen, am 3.9.2018