Aachen Institute for Advanced Study in Computational Engineering Science

Preprint: AICES-2011/04-2

27/April/2011

hp-Analysis of a Hybrid DG Method for Stokes Flow

H. Egger and C. Waluga

Financial support from the Deutsche Forschungsgemeinschaft (German Research Association) through grant GSC 111 is gratefully acknowledged.

List of AICES technical reports: http://www.aices.rwth-aachen.de/preprints

hp-ANALYSIS OF A HYBRID DG METHOD FOR STOKES FLOW

HERBERT EGGER[1] AND CHRISTIAN WALUGA[2]

ABSTRACT. This manuscript deals with the hp error analysis of a hybrid discontinuous Galerkin method for incompressible flow. Besides the usual coercivity and boundedness estimates, we establish the inf-sup stability for the discrete incompressibility constraint with a constant, which is only slightly sub-optimal with respect to the polynomial degree. This result holds on irregular and hybrid meshes in two and three spatial dimensions, and its proof is based on a new stability estimate for the L^2 projection on simplex elements. The sharp estimate for the inf-sup constants in turn allows to derive a-priori estimates, which are optimal with respect to the meshsize and only slightly sub-optimal with respect to the polynomial degree. In addition to the a-priori results, we also present a rigorous hp analysis of a residual-type a-posteriori error estimator. Reliability and efficiency are proven and the explicit dependence of the estimates on the polynomial approximation order is elaborated. The theoretical results are illustrated with numerical tests.

Keywords: Stokes equation, discontinuous Galerkin methods, hybridization, a-posteriori error estimation, hp-analysis

AMS subject classification: 65N12, 65N15, 65N30, 65N35

1. Introduction

Discontinuous Galerkin (DG) methods have been introduced and since then been widely used for the solution of hyperbolic problems, e.g. in neutron transport [54, 44, 39] or compressible flow [24, 38, 9]. On the other hand, these methods can be applied as well for the discretization of second order elliptic problems, cf. e.g. [3, 5]. A combination of both approaches allows to treat problems in purely hyperbolic, convection dominated, and elliptic regimes in a unified manner. For a systematic introduction and various applications of DG methods, let us refer to [5, 55] and [22], and to the references given there.

For elliptic problems, DG methods are closely related to mixed methods [14]; in fact, they can be understood as stabilized mixed methods, and hybridization [4] can be employed for an efficient implementation. The generalization of this idea to discontinuous Galerkin and domain decomposition methods has been investigated for instance in [8, 19, 26].

Discontinuous Galerkin methods are well-suited also for the numerical solution of incompressible flow problems, and several schemes have been proposed and analyzed for this purpose, see e.g. [61, 21, 57, 34, 55]. In this paper, we consider a hybrid DG method for Stokes flow, which is closely related to the DG methods investigated in [61, 57, 34]. The term *hybrid* refers

[1]Center of Mathematics, Technical University of Munich, Boltzmannstraße 3, 85748 Garching, Germany. Email: herbert.egger@ma.tum.de

[2]Aachen Institute for Advanced Study in Computational Engineering Science, RWTH Aachen University, Schinkelstraße 2, 52062 Aachen, Germany. Email: waluga@aices.rwth-aachen.de

to the fact, that, like for the hybrid mixed methods [4], additional variables are used for the approximation of fields on the skeleton of the mesh. Let us also mention the work [50, 43, 20] on other hybrid DG formulations for incompressible flow, which make explicit use of velocity gradients, use different approximation spaces, or a vorticity formulation. A comparison of several methods has been given in [23].

The method considered in this paper can be seen as a hybrid version of a symmetric interior penalty Galerkin method included in the analysis of [57]. The hybrid method only involves approximations for the velocity and pressure on the elements, and additional hybrid variables for the velocities at the skeleton. In contrast to related DG methods, the hybrid formulation allows an element-wise assembling process, and the elimination of most degrees of freedom already on the element level.

The focus of the current presentation is a detailed *hp* error analysis for the hybrid DG method on irregular meshes made up of simplices and hypercubes. With similar tools as in [34, 55], the extension of the method to the Ossen and Navier-Stokes equations is possible, but such generalizations are not in the scope of the present work.

The basic result of the first part of our presentation, which deals with the a-priori analysis, will be the proof of an inf-sup stability condition for the discrete divergence constraint. The presence of the hybrid variables allows a particularly simple construction of a *Fortin operator* based on local L^2 projections. The *hp* bounds for the Fortin operator, and consequently also for the inf-sup stability constant, then follow from stability and approximation estimates for the local projections. Based on recent results of [18], we are able to derive sharp estimates also for simplex elements. Our analysis thus applies to irregular hybrid meshes containing simplices and hypercubes. Let us mention [61] and [57], where similar *hp* stability estimates for related DG methods on quadrilateral and hexahedral meshes were derived with different arguments. *hp* estimates for the inf-sup constant are also known for conforming finite elements [59] and spectral methods [12]; see Remark 6.11 for a short discussion.

In the second part of this manuscript, we investigate the a-posteriori error estimation based on the framework of [1, 41] for non-conforming finite element methods; see also [40, 17] and [36, 15] for related work on DG methods, and [46] for the *hp* a-posteriori estimates of conforming finite element methods. We consider in detail a residual-type error estimator, derive efficiency and reliability estimates, and particularly investigate the explicit dependence of the bounds on the polynomial degree. For the proof of efficiency, we utilize an averaging operator of Oswald-type [52, 40], which allows to construct a conforming approximation for the finite element solution by local postprocessing. A detailed *hp* analysis of this averaging procedure is given in the appendix. Similar results for quadrilateral and hexahedral meshes or for the two dimensional case can be found in [15] and [36]; see also [7, 45] for the related construction of *hp* approximations.

The outline of our presentation is as follows: In Section 2, we shortly discuss the continuous Stokes problem and recall some basic results for later reference. After introducing the basic setting and our main notation in Section 3, we present the hybrid DG method in Section 4 and highlight some of its basic properties in Section 5. The a-priori and a-posteriori error analysis of the method is given in Sections 6 and 7, and for illustration of the theoretical findings, we report on numerical tests in Section 8.

2. The Stokes problem

Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain in $d = 2$ or 3 space dimensions. As a model for the flow of an incompressible viscous fluid confined in Ω, we consider the stationary Stokes problem with homogenous Dirichlet boundary conditions, given by

$$\begin{cases} -\Delta\,\boldsymbol{u} + \boldsymbol{\nabla} p &= \boldsymbol{f} \quad \text{in } \Omega, \\ \operatorname{div} \boldsymbol{u} &= 0 \quad \text{in } \Omega, \\ \boldsymbol{u} &= \boldsymbol{0} \quad \text{on } \partial\Omega. \end{cases} \tag{1}$$

Throughout the presentation, we assume that $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega)$, and we require that p has zero average to guarantee uniqueness of the pressure. The standard weak formulation of (1) then is the following: Find $\boldsymbol{u} \in \boldsymbol{H}_0^1(\Omega)$ and $p \in L_0^2(\Omega)$ such that

$$\begin{cases} \mathbf{a}(\boldsymbol{u},\boldsymbol{v}) + \mathbf{b}(\boldsymbol{v},p) &= (\boldsymbol{f},\boldsymbol{v})_\Omega \quad \text{for all } \boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega), \\ \mathbf{b}(\boldsymbol{u},q) &= 0 \qquad\quad\;\, \text{for all } q \in L_0^2(\Omega), \end{cases} \tag{2}$$

with bilinear forms $\mathbf{a}$ and $\mathbf{b}$ defined by

$$\mathbf{a}(\boldsymbol{u},\boldsymbol{v}) := (\boldsymbol{\nabla}\boldsymbol{u}, \boldsymbol{\nabla}\boldsymbol{v})_\Omega \qquad \text{and} \qquad \mathbf{b}(\boldsymbol{v},q) := -(\operatorname{div} \boldsymbol{v}, q)_\Omega.$$

Here, $\boldsymbol{\nabla}\boldsymbol{u}$ denotes the velocity gradient tensor $[\boldsymbol{\nabla}\boldsymbol{u}]_{ij} = \partial_j \boldsymbol{u}_i$, and $(a,b)_\Omega := \int_\Omega a\, b\, dx$ is the scalar product of square integrable functions. Corresponding definitions are used for vector and tensor valued functions, which are denoted with bold symbols. The spaces

$$\boldsymbol{H}_0^1(\Omega) := \{\boldsymbol{v} \in [H^1(\Omega)]^d : \boldsymbol{v} = 0 \text{ on } \partial\Omega\} \quad \text{and} \quad L_0^2(\Omega) := \{q \in L^2(\Omega) : \int_\Omega q\, dx = 0\},$$

are the natural choice for the variational treatment of the Stokes problem [33, 14]. It is well-known, that under our assumptions on the domain and the data, problem (2) has a unique solution $(\boldsymbol{u}, p) \in \boldsymbol{H}_0^1(\Omega) \times L_0^2(\Omega)$. The essential ingredient for the analysis is the following inf-sup condition [14]: There exists a constant β_Ω depending only on the domain Ω such that

$$\sup_{\boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega)} \frac{\mathbf{b}(\boldsymbol{v},q)}{\|\boldsymbol{\nabla}\boldsymbol{v}\|_{L^2(\Omega)}} \geq \beta_\Omega \|q\|_{L^2(\Omega)}, \qquad \text{for all } q \in L_0^2(\Omega). \tag{3}$$

Here and below, functions in the denominator are tacitly assumed to be non-zero. We refer to [42, 49, 33] for a proof of condition (3), which is equivalent to the surjectivity of the divergence operator $\operatorname{div} : \boldsymbol{H}_0^1(\Omega) \to L_0^2(\Omega)$. The mixed problem (2) is equivalent to the variational problem: Find $(\boldsymbol{u}, p) \in \boldsymbol{H}_0^1(\Omega) \times L_0^2(\Omega)$ such that

$$\mathcal{B}(\boldsymbol{u}, p; \boldsymbol{v}, q) = (\boldsymbol{f}, \boldsymbol{v})_\Omega \qquad \text{for all } (\boldsymbol{v}, q) \in \boldsymbol{H}_0^1(\Omega) \times L_0^2(\Omega), \tag{4}$$

where the combined bilinear form $\mathcal{B}$ for the Stokes problem is defined by

$$\mathcal{B}(\boldsymbol{u}, p; \boldsymbol{v}, q) := \mathbf{a}(\boldsymbol{u},\boldsymbol{v}) + \mathbf{b}(\boldsymbol{v},p) + \mathbf{b}(\boldsymbol{u},q). \tag{5}$$

We will switch between the two formulations (2) and (4) as convenient. Based on the unique solvability of (2), one easily verifies that also the combined bilinear form $\mathcal{B}$ satisfies an inf-sup condition, namely, for all $\boldsymbol{u} \in \boldsymbol{H}_0^1(\Omega)$ and $p \in L_0^2(\Omega)$ there holds

$$\sup_{(\boldsymbol{v},q) \in \boldsymbol{H}_0^1 \times L_0^2} \frac{\mathcal{B}(\boldsymbol{u}, p; \boldsymbol{v}, q)}{\|\boldsymbol{\nabla}\boldsymbol{v}\|_{L^2(\Omega)} + \|q\|_{L^2(\Omega)}} \geq c_\Omega \left(\|\boldsymbol{\nabla}\boldsymbol{u}\|_{L^2(\Omega)} + \|p\|_{L^2(\Omega)}\right), \tag{6}$$

with stability constant $c_\Omega \sim \beta_\Omega^{-2}$; see [65] for sharp estimates. The unique solvability of (4), and hence also of (2), then follows also directly from the stability condition (6) by application of the Babuška-Aziz lemma [6, 49].

At the end of this section, let us shortly recall some well-known regularity results for solutions of the Stokes problem, which may serve as a justification for the regularity assumptions used in the derivation of our method and the a-priori error analysis below.

Remark 2.1. On domains with $C^{1,1}$ boundary, any solution of the variational problem (2) is regular, i.e., $\boldsymbol{u} \in \boldsymbol{H}^2(\Omega)$ and $p \in H^1(\Omega)$, whenever $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega)$; in this case, the weak solution is actually a strong solution of the Stokes problem (1). The same regularity also holds for convex polygonal domains in $\mathbb{R}^2$. For proofs of these statements, and further regularity results (including the case of L^p spaces), we refer to [35] and the references given there.

3. Notation and preliminaries

Before we turn to the discretization of the Stokes problem, let us shortly introduce some relevant notation. For ease of presentation, we primarily use the terminology of the two dimensional case, but all our results hold also in three dimensions.

3.1. **Assumptions on the mesh.** Let $\mathcal{T}_h = \{T\}$ denote a partition of the domain Ω into affine families of triangles and quadrilaterals (respectively tetrahedra and hexahedra in three dimensions). As usual, h_T and ρ_T denote the diameter of the element T and of the largest ball inscribed in T. The ratio $\gamma_T := h_T/\rho_T$ serves as a measure for the local shape regularity. By $\partial\mathcal{T}_h := \{\partial T : T \in \mathcal{T}_h\}$ and $\mathcal{E}_h := \{E_{ij} = \partial T_i \cap \partial T_j) : i > j\} \cup \{E_{i,0} : \partial T_i \cap \partial\Omega\}$ we denote the set of element boundaries respectively of edges (faces) between adjacent elements and on the boundary. h_E again refers to the diameter of E, and the union $\mathcal{E} := \bigcup_{E\in\mathcal{E}_h} E$ of all edges (faces) is called *skeleton* or *interface.*

We assume that the mesh $(\mathcal{T}_h, \mathcal{E}_h)$ is uniformly shape-regular and locally-quasi-uniform in the following sense: there exists a constant $\gamma > 0$ (independent of the partition) such that

$$\gamma_T \geq \gamma \qquad \text{and} \qquad h_E \geq \gamma h_T \quad \text{if} \quad E \subset \partial T \qquad \text{for all } T \in \mathcal{T}_h,\ E \in \mathcal{E}_h. \tag{7}$$

Note that (7) also implies $h_T \geq \gamma h_{T'}$ for all T, $T' \in \mathcal{T}_h$ which share a common edge (face). To each element $T \in \mathcal{T}_h$ and any facet $E \in \mathcal{E}_h$, we also associate a polynomial degree of approximation k_T or $k_E \in \mathbb{N}$, and we set $k_{\max} := \max\{k_T : T \in \mathcal{T}_h\}$. We further assume that

$$k_E \geq k_T \qquad \text{for all } T \in \mathcal{T}_h,\ E \in \mathcal{E}_h \text{ with } E \subset \partial T, \tag{8}$$

and we require that also the polynomial degree distribution is locally quasi-uniform, by which we mean: there exists a constant $\eta > 0$ such that

$$k_T \leq \eta\, k_{T'} \qquad \text{for all } T, T' \in \mathcal{T}_h \text{ such that } \partial T \cap \partial T' \neq \emptyset. \tag{9}$$

Remark 3.1. Our setting covers irregular hybrid meshes, i.e., non-conforming partitions that contain different element types; a typical valid situation is depicted in Figure 1. Note that the interface mesh $\mathcal{E}_h$ satisfies a certain *geometric compatibility condition*, i.e., intersections between adjacent elements are resolved by the partition of the skeleton. Condition (8) is a second *algebraic compatibility condition.* Both compatibility conditions will be used explicitly in the a-priori as well as in the a-posteriori error analysis below.

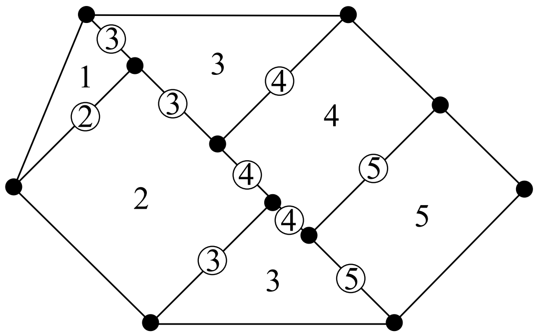

FIGURE 1. A typical hybrid irregular mesh covered by the analysis of this paper. The circles denote the endpoints of interface edges. Hanging nodes are allowed, but the partition of the skeleton is required to resolve the interfaces (*geometric compatibility*). The polynomial degree associated to an edge is at least as large as that of the adjacent elements (*algebraic compatibility*).

3.2. **Broken Sobolev spaces.** For $s \geq 0$, we define the broken Sobolev spaces

$$H^s(\mathcal{T}_h) := \{v \in L^2(\Omega) : v|_T \in H^s(T) \quad \text{for all } T \in \mathcal{T}_h\}.$$

Functions $u \in H^1(\mathcal{T}_h)$ have well-defined piecewise derivatives, which are denoted with the standard symbols. The scalar product of piecewise defined functions is given by

$$(u,v)_{\mathcal{T}_h} := \sum\nolimits_{T\in\mathcal{T}_h} (u,v)_T \qquad \text{where} \qquad (u,v)_T := \int_T u\, v\, dx,$$

and the corresponding norm is $\|u\|_{0,\mathcal{T}_h} := \sqrt{(u,u)_{\mathcal{T}_h}}$. Similar definitions can be made for arbitrary $s \geq 0$ and also for functions on the interface partition or the element boundaries. The space $L^2(\partial\mathcal{T}_h) = \prod_T L^2(\partial T)$ is equipped with the scalar product

$$\langle u,v\rangle_{\partial\mathcal{T}_h} := \sum\nolimits_{T\in\mathcal{T}_h} \langle u,v\rangle_{\partial T} \qquad \text{where} \qquad \langle u,v\rangle_{\partial T} := \int_{\partial T} u\, v\, ds,$$

and the corresponding norm is denoted by $|\cdot|_{0,\partial\mathcal{T}_h}$. For convenience of the reader, we use different symbols for scalar products and norms over elements and element boundaries, respectively. Note that for $s \neq$ integer, the functions in $H^s(\partial\mathcal{T}_h)$ are just the traces of functions in $H^{s+1/2}(\mathcal{T}_h)$, and that any function $v \in L^2(\mathcal{E})$ can be identified with a function in $L^2(\partial\mathcal{T}_h)$ by doubling its values at the interfaces.

All definitions naturally extend to vector and tensor valued functions, which are denoted with bold symbols throughout the text.

4. A hybrid discontinuous Galerkin method

Let us now introduce the hybrid DG method which is the subject of our analysis. For the approximation of velocity and pressure, we utilize finite element spaces of discontinuous piecewise polynomial functions, namely

$$\boldsymbol{V}_h := \{\boldsymbol{v}_h \in \boldsymbol{L}^2(\mathcal{T}_h) : \boldsymbol{v}_h|_T \in \boldsymbol{\mathcal{P}}_{k_T}(T) \quad \text{for all } T \in \mathcal{T}_h\},$$
$$Q_h := \{q_h \in L^2_0(\Omega) : q_h|_T \in \mathcal{P}_{k_T-1}(T) \quad \text{for all } T \in \mathcal{T}_h\}.$$

As usual, $\mathcal{P}_k$ is the space of polynomials of total degree $\le k$, and according to our convention, $\boldsymbol{\mathcal{P}}_k := [\mathcal{P}_k]^d$ are the vector valued polynomials. For our hybrid method, we additionally require a space for approximating the velocity on the skeleton. Here, we choose

$$\widehat{\boldsymbol{V}}_h := \{\hat{\boldsymbol{v}}_h \in \boldsymbol{L}^2(\mathcal{E}_h) : \hat{\boldsymbol{v}}_h|_E \in [\mathcal{P}_{k_E}(E)]^d \quad \text{for all } E \in \mathcal{E}_h, \quad \hat{\boldsymbol{v}}_h|_{\partial\Omega} = 0\}.$$

Remark 4.1. Dirichlet boundary conditions are explicitly included in the definition of the hybrid space $\widehat{\boldsymbol{V}}_h$, and no conditions are imposed on functions in $\boldsymbol{V}_h$. Also note that we utilize the same polynomial spaces for simplicial and rectangular elements, and also for the interface elements, which, in three dimensions, may have a complicated shape.

As discretization for the variational form (2) of the Stokes problem, we then consider

Method 4.1. *Given $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega)$, find $\boldsymbol{u}_h \in \boldsymbol{V}_h$, $\hat{\boldsymbol{u}}_h \in \widehat{\boldsymbol{V}}_h$, and $p_h \in Q_h$, such that*

$$\left\{ \begin{array}{ll} \mathbf{a}_h(\boldsymbol{u}_h, \hat{\boldsymbol{u}}_h; \boldsymbol{v}_h, \hat{\boldsymbol{v}}_h) + \mathbf{b}_h(\boldsymbol{v}_h, \hat{\boldsymbol{v}}_h; p_h) & = (\boldsymbol{f}, \boldsymbol{v}_h)_{\mathcal{T}_h}, \\ \mathbf{b}_h(\boldsymbol{u}_h, \hat{\boldsymbol{u}}_h; q_h) & = 0, \end{array} \right. \tag{10}$$

for all $\boldsymbol{v}_h \in \boldsymbol{V}_h$, $\hat{\boldsymbol{v}}_h \in \widehat{\boldsymbol{V}}_h$, and $q_h \in Q_h$, with bilinear forms defined by

$$\mathbf{a}_h(\boldsymbol{u}, \hat{\boldsymbol{u}}; \boldsymbol{v}, \hat{\boldsymbol{v}}) := (\boldsymbol{\nabla} \boldsymbol{u}, \boldsymbol{\nabla} \boldsymbol{v})_{\mathcal{T}_h} - \langle \boldsymbol{\partial}_n \boldsymbol{u}, \boldsymbol{v} - \hat{\boldsymbol{v}} \rangle_{\partial\mathcal{T}_h} - \langle \boldsymbol{u} - \hat{\boldsymbol{u}}, \boldsymbol{\partial}_n \boldsymbol{v} \rangle_{\partial\mathcal{T}_h} + \langle \alpha(\boldsymbol{u} - \hat{\boldsymbol{u}}), \boldsymbol{v} - \hat{\boldsymbol{v}} \rangle_{\partial\mathcal{T}_h},$$
$$\mathbf{b}_h(\boldsymbol{v}, \hat{\boldsymbol{v}}; q) := -(\operatorname{div} \boldsymbol{v}, q)_{\mathcal{T}_h} + \langle \boldsymbol{v} - \hat{\boldsymbol{v}}, q\boldsymbol{n} \rangle_{\partial\mathcal{T}_h}.$$

The element-wise constant stabilization parameter $\alpha > 0$ will be specified below.

Remark 4.2. Method 4.1 is a special case of the hybrid mortar methods studied in [27], and it is closely related to the discontinuous Galerkin methods investigated in [30, 34, 55]. This can be seen by replacing $\hat{\boldsymbol{u}}$ and $\hat{\boldsymbol{v}}$ with $\{\boldsymbol{u}\} = \frac{1}{2}(\boldsymbol{u} + \boldsymbol{u}')$ and $\{\boldsymbol{v}\} := \frac{1}{2}(\boldsymbol{v} + \boldsymbol{v}')$ denoting the averages of the function values on neighboring elements T, T' at the interface $\partial T \cap \partial T'$.

Like on the continuous level, we define a combined bilinear form

$$\mathcal{B}_h(\boldsymbol{u}_h, \hat{\boldsymbol{u}}_h, p_h; \boldsymbol{v}_h, \hat{\boldsymbol{v}}_h, q_h) := \mathbf{a}_h(\boldsymbol{u}_h, \hat{\boldsymbol{u}}_h; \boldsymbol{v}_h, \hat{\boldsymbol{v}}_h) + \mathbf{b}_h(\boldsymbol{v}_h, \hat{\boldsymbol{v}}_h; ph) + \mathbf{b}_h(\boldsymbol{u}_h, \hat{\boldsymbol{u}}_h; q_h), \tag{11}$$

which allows us to rewrite (10) in compact form as

$$\mathcal{B}_h(\boldsymbol{u}_h, \hat{\boldsymbol{u}}_h, p_h; \boldsymbol{v}_h, \hat{\boldsymbol{v}}_h, q_h) = (\boldsymbol{f}, \boldsymbol{v}_h)_{\mathcal{T}_h}. \tag{12}$$

We will use both formulations (10) and (12) in our analysis as convenient.

5. Basic properties of the hybrid method

Let us shortly mention some basic properties of the hybrid method, which follow almost immediately from its definition. As the first statement shows, the variational problem (10) is indeed a reasonable approximation for the Stokes problem.

Proposition 5.1 (Consistency). *Any solution $\boldsymbol{u} \in \boldsymbol{H}^2(\mathcal{T}_h) \cap \boldsymbol{H}^1_0(\Omega)$ and $p \in H^1(\mathcal{T}_h) \cap L^2_0(\Omega)$ of the Stokes problem* (2)*, also satisfies the discrete variational principle* (10)*, i.e.,*

$$\mathcal{B}_h(\boldsymbol{u}, \boldsymbol{u}, p; \boldsymbol{v}_h, \hat{\boldsymbol{v}}_h, q_h) = (\boldsymbol{f}, \boldsymbol{v}_h)_{\partial\mathcal{T}_h} \qquad \text{for all } \boldsymbol{v}_h \in \boldsymbol{V}_h,\ \hat{\boldsymbol{v}}_h \in \widehat{\boldsymbol{V}}_h,\ \text{and } \ q_h \in Q_h.$$

Proof. Testing (10) with arbitrary $\boldsymbol{v}_h$, $\hat{\boldsymbol{v}}_h$ and $q_h = 0$, integrating by parts on each element, and using (1), we obtain that

$$\mathbf{a}_h(\boldsymbol{u}, \boldsymbol{u}; \boldsymbol{v}_h, \hat{\boldsymbol{v}}_h) + \mathbf{b}_h(\boldsymbol{v}_h, \hat{\boldsymbol{v}}_h; p) = (\boldsymbol{f}, \boldsymbol{v}_h)_{\mathcal{T}_h} + \langle \boldsymbol{\partial}_n \boldsymbol{u} - p\boldsymbol{n}, \hat{\boldsymbol{v}}_h \rangle_{\partial\mathcal{T}_h}$$

for all $(\boldsymbol{v}_h, \hat{\boldsymbol{v}}_h)$ in the test space. Due to the regularity assumption on the solution, the normal flux $\boldsymbol{\partial}_n \boldsymbol{u} - p\boldsymbol{n}$ is continuous across element interfaces. Since $\hat{\boldsymbol{v}}_h$ is single valued on each interface, it follows by summation that $\langle \boldsymbol{\partial}_n \boldsymbol{u} - p\boldsymbol{n}, \hat{\boldsymbol{v}}_h \rangle_{\partial\mathcal{T}_h}$ vanishes for all $\hat{\boldsymbol{v}}_h \in L^2(\mathcal{E})$. Moreover, since $\operatorname{div} \boldsymbol{u} = 0$ and $\boldsymbol{u}$ is continuous across element interfaces, we further obtain that $\mathbf{b}_h(\boldsymbol{u}, \boldsymbol{u}; q_h) = -(\operatorname{div} \boldsymbol{u}, q_h)_{\mathcal{T}_h} + \langle \boldsymbol{u} - \boldsymbol{u}, q_h \boldsymbol{n} \rangle_{\partial\mathcal{T}_h} = 0$, for any $q_h \in Q_h$, which completes the proof. □

Remark 5.2. For our argument, we required that the normal flux $\boldsymbol{\sigma} \cdot \boldsymbol{n} := \boldsymbol{\partial}_n \boldsymbol{u} - p\boldsymbol{n}$ is well-defined on element interfaces $E = \partial T \cap \partial T'$. To ensure this, it would suffice to have $\boldsymbol{u} \in \boldsymbol{H}^s(\mathcal{T}_h)$ for some $s > 3/2$; cf. [34] and Remark 2.1, and also [14] for other conditions. The additional regularity assumption can also be avoided by using lifting operators to define an equivalent discrete problem; see [53] and [57] for details.

The second result of this section states the continuity of numerical fluxes, which is a property of many DG methods, and yields local conservation on the discrete level. We make explicit use of this continuity in the derivation of a-posteriori error estimates in Section 7.

Proposition 5.3 (Flux continuity). *Let $(\boldsymbol{u}_h, \hat{\boldsymbol{u}}_h, p_h)$ denote a solution of Method 4.1. Then the* numerical fluxes *defined element-wise by*

$$\boldsymbol{\sigma}_h \cdot \boldsymbol{n}|_{\partial T} := \boldsymbol{\partial}_n \boldsymbol{u}_h - p_h \boldsymbol{n} + \alpha_T(\boldsymbol{u}_h - \hat{\boldsymbol{u}}_h), \qquad T \in \mathcal{T}_h,$$

are continuous, i.e., $\boldsymbol{\sigma}_h \cdot \boldsymbol{n}|_{\partial T \cap E} + \boldsymbol{\sigma}_h \cdot \boldsymbol{n}|_{\partial T' \cap E} = 0$ on $E = \partial T \cap \partial T'$.

Proof. We utilize the geometric and algebraic compatibility condition of the approximation: By construction of the mesh, $E = \partial T \cap \partial T'$ is resolved by the interface mesh; cf Remark 3.1. Moreover, by definition of the approximation spaces and the algebraic compatibility condition (8), we have $\boldsymbol{\sigma}_h \cdot \boldsymbol{n}|_{\partial T \cap E} \in \boldsymbol{\mathcal{P}}_{k_E}(E)$. The result then follows by testing (10) with functions $\hat{\boldsymbol{v}}_h \in \widehat{\boldsymbol{V}}_h$ that vanish outside E and such that $\hat{\boldsymbol{v}}_h|_E$ span $\boldsymbol{\mathcal{P}}_k(E)$. □

Testing the discrete variational principle with piecewise constant functions, we obtain the following local conservation properties, which resemble the conservation of mass and momentum on the continuous level.

Corollary 5.4 (Local conservation). *Let $(\boldsymbol{u}_h, \hat{\boldsymbol{u}}_h, p_h)$ be a solution of Method 4.1. Then*

$$\langle \boldsymbol{u}_h \cdot \boldsymbol{n}, 1 \rangle_{\partial T} = 0 \qquad \text{and} \qquad (\boldsymbol{f}, \mathbf{1}_i)_T = \langle \boldsymbol{\sigma}_h \boldsymbol{n}, \mathbf{1}_i \rangle_{\partial T}$$

for all elements $T \in \mathcal{T}_h$, i.e., the hybrid DG method is locally conservative.

Here, $\mathbf{1}_i$ is the vector function with component i equal to one and the others zero.

6. A priori analysis

In this section, we establish the unique solvability of the discrete problem, and derive the basic a-priori error estimates. We particularly investigate the dependence of the estimates on the polynomial degree of approximation. For a more traditional h analysis of related DG methods for incompressible flow problems, see [34, 55], and also [61, 57] for hp estimates.

Many of the arguments utilized in the following, e.g. the use of mesh dependent norms, are standard tools for the analysis of DG methods, see e.g. [3, 5] and the references therein.

6.1. **Preliminaries.** We start with introducing some further notation and recalling a few basic results on polynomial approximation, which we require for our analysis.

Lemma 6.1. *For any polynomial $v_k \in P_k(T)$ with $T \in \mathcal{T}_h$, there holds*

$$|v_k|^2_{0,\partial T} \le C_T \tfrac{k_T^2}{h_T} \|v_k\|^2_{0,T}, \qquad \|\nabla v_k\|^2_{0,T} \le C_T \tfrac{k_T^4}{h_T^2} \|v_k\|^2_{0,T}, \tag{13}$$

with constant C_T depending only on the shape of the element.

Remark 6.2. Assuming uniform shape regularity of the mesh, the estimates hold uniformly for all elements with a constant $C_T = C_\gamma$ depending only on the shape regularity of the mesh. Note that on simple elements, optimal constants C_T can be computed explicitly [63, 15, 32].

Let us next recall some polynomial approximation results, which will be used for the stability analysis and for characterizing the approximation properties of the finite element spaces.

Lemma 6.3. *For any element $T \in \mathcal{T}_h$ and every function $u \in H^s(T)$ with $s \ge 0$, there exists a polynomial function $u_k \in \mathcal{P}_{k_T}(T)$ such that*

$$\tfrac{k_T}{h_T}\|u-u_k\|_{0,T} + \sqrt{\tfrac{k_T}{h_T}}|u-u_k|_{0,\partial T} + \|\nabla(u-u_k)\|_{0,T} \le C_s h_T^{\min\{s,k_T\}} k_T^{-s} |u|_{H^{s+1}(T)},$$

with a constant C_s depending only on s and the shape of the element T. Here, $|\cdot|_{H^s}$ denotes the seminorm of H^s. If $s_T > 1/2$, then the same estimate holds for $\sqrt{\frac{h_T}{k_T}}|\partial_n(u-u_k)|_{0,\partial T}$.

Proof. The proof of the two-dimensional result [7] carries over also to three dimensions. □

For any element $T \in \mathcal{T}_h$ and any edge $E \in \mathcal{E}_h$, we denote by

$$\Pi_T^k : H^1(T) \to \mathcal{P}_k(T) \qquad \text{and} \qquad \Pi_E^k : H^1(T) \to \mathcal{P}_k(E)$$

the L^2 orthogonal projections onto the local polynomial spaces. We use again bold symbols for the vector valued analogues. Let us summarize the basic stability and approximation properties of these operators.

Lemma 6.4. *For any $v \in H^1(T)$, and every edge (face) $E \subset \partial T$, there holds*

$$(i) \quad \|\Pi_T^k v - v\|_{0,T} \le C \tfrac{h_T}{k_T} \|\nabla v\|_{0,T}, \qquad (ii) \quad \|\nabla \Pi_T^k v\|_{0,T} \le C\sqrt{k_T}\,\|\nabla v\|_{0,T},$$

$$(iii) \quad |v - \Pi_T^k v|_{0,\partial T} \le C\sqrt{\tfrac{h_T}{k_T}}\,\|\nabla v\|_{0,T}, \qquad (iv) \quad |v - \Pi_E^k v|_{0,E} \le C\sqrt{\tfrac{h_T}{k_T}}\,\|\nabla v\|_{0,T},$$

with a constant C depending only on the shape of the element T.

Proof. According to Lemma 6.3, there exists a function $v_k \in \mathcal{P}_k(T)$ with good approximation properties. Due to the optimality of the L^2 projection, we obtain

$$\|\Pi_T^k v - v\|_{0,T} \le \|v_k - v\|_{0,T} \le C_1 \tfrac{h_T}{k_T}\,\|\nabla v\|_{0,T},$$

which yields the first estimate. The third estimate is proven in [37] for quadrilateral and hexahedral elements, and in [18] for simplices. The fourth estimate follows from the third one by optimality of the L^2 projection on the edge. To show the second estimate, we take again $v_k \in \mathcal{P}_k(T)$ as in Lemma 6.3. By the triangle inequality, we have

$$\|\nabla \Pi_T^k v\|_T \le \|\nabla v_k\|_T + \|\nabla(\Pi_T^k v - v_k)\|_T.$$

According to Lemma 6.3, the first term is bounded by a multiple of $\|\nabla v\|_T$. Let us denote $w_k = \Pi_T^k v - v_k$. Then the second term can be further estimated by

$$\begin{aligned}\|\nabla(\Pi_T^k v - v_k)\|_T^2 &= -(\Pi_T^k v - v_k, \Delta w_k)_T + \langle \Pi_T^k v - v_k, \partial_n w_k\rangle_{\partial T}\\ &= -(v - v_k, \Delta w_k)_T + \langle \Pi_T^k v - v_k, \partial_n w_k\rangle_{\partial T}\\ &= (\nabla(v - v_k), \nabla w_k)_T + \langle \Pi_T^k v - v, \partial_n w_k\rangle_{\partial T}.\end{aligned}$$

The first term here can be bounded by $C\|\nabla v\|_T\|\nabla w_k\|_T$, and using the (iii) and the discrete trace inequality (13), we obtain

$$\langle \Pi_T^k v - v, \partial_n w_k\rangle_{\partial T} \le C'\sqrt{\tfrac{h_T}{k_T}}\|\nabla v\|_T\sqrt{\tfrac{k_T^2}{h_T}}\|\nabla w_k\|_T = C'\sqrt{k_T}\|\nabla v\|_T\|\nabla w_k\|_T.$$

The result now follows by dividing through $\|\nabla w_k\|_T = \|\nabla(\Pi_T^k v - v_k)\|_T$. □

Remark 6.5. To the best of our knowledge, the sharp stability estimate (ii) for the L^2 projection has not been shown before for simplicial elements. Corresponding results for quadrilateral and hexahedral elements can however be found in the literature, e.g., in [58, 37].

6.2. **Ellipticity and boundedness.** In order to guarantee stability of the bilinear form $\mathbf{a}_h$, we choose the piecewise constant stabilization parameter α element-wise such that

$$4C_T k_T^2/h_T \le \alpha_T \le C k_T^2/h_T, \tag{14}$$

where C is some positive constant and C_T is taken from Lemma 6.1. The second condition is only required to keep the approximation error and the condition number of the global system reasonably small. We then define a pair of scaled norms on $L^2(\partial\mathcal{T}_h)$ by

$$|v|_{\pm 1/2,h} := \big(\textstyle\sum_{\mathcal{T}\in\mathcal{T}_h} |v|^2_{\pm 1/2,\partial T}\big)^{1/2}, \quad \text{where} \quad |v|_{\pm 1/2,\partial T} := \alpha_T^{\pm 1/2}|v|_{\partial T}.$$

Similar norms are frequently used for the analysis of non-conforming methods. Note, that by definition of the norms and the Cauchy-Schwarz inequality, we have

$$\langle u, v\rangle_{\partial\mathcal{T}_h} \le |u|_{1/2,\partial\mathcal{T}_h}|v|_{-1/2,h} \qquad \text{for all } u, v \in L^2(\partial\mathcal{T}_h),$$

which mimics the inequality for the duality product on the continuous level. By application of the first estimate in (13), we further obtain that

$$|q_h|_{-1/2,\partial\mathcal{T}_h} \le \tfrac{1}{2}\|q_h\|_{0,\mathcal{T}_h} \qquad \text{and} \qquad |\boldsymbol{\partial}_n \boldsymbol{v}_h|_{-1/2,\partial\mathcal{T}_h} \le \tfrac{1}{2}\|\boldsymbol{\nabla}\boldsymbol{v}_h\|_{0,\mathcal{T}_h}, \tag{15}$$

for all $q_h \in Q_h$ and $\boldsymbol{v}_h \in \boldsymbol{V}_h$. The inverse inequalities (15) will be used at several places.

The analysis of Method 4.1 is most naturally carried out with respect to the following mesh-dependent norms: Discrete stability is verified with respect to

$$\|(\boldsymbol{v},\hat{\boldsymbol{v}})\|_{1,\mathcal{T}_h} := \Big(\|\boldsymbol{\nabla}\boldsymbol{v}\|^2_{0,\mathcal{T}_h} + |\boldsymbol{v}-\hat{\boldsymbol{v}}|^2_{1/2,\partial\mathcal{T}_h}\Big)^{1/2} \quad \text{and} \quad \|q\|_{0,\mathcal{T}_h} := (q,q)^{1/2}_{\mathcal{T}_h},$$

and we use a second pair of stronger norms defined by

$$|\!|\!|(\boldsymbol{v},\hat{\boldsymbol{v}})|\!|\!|_{1,\mathcal{T}_h} := \Big(\|(\boldsymbol{v},\hat{\boldsymbol{v}})\|^2_{1,\mathcal{T}_h} + |\boldsymbol{\partial}_n\boldsymbol{v}|^2_{-1/2,\partial\mathcal{T}_h}\Big)^{1/2} \quad \text{and} \quad |\!|\!|q|\!|\!|_{0,\mathcal{T}_h} := \Big(\|q\|^2_{0,\mathcal{T}_h} + |q|^2_{-1/2,\partial\mathcal{T}_h}\Big)^{1/2},$$

to bound the bilinear forms also on infinite dimensional spaces. On the discrete level, the two sets of energy norms are equivalent, i.e., for all $\boldsymbol{v}_h \in \boldsymbol{V}_h$, $\hat{\boldsymbol{v}}_h \in \widehat{\boldsymbol{V}}_h$, and $q_h \in Q_h$, there holds

$$\|(\boldsymbol{v}_h, \hat{\boldsymbol{v}}_h)\|_{1,\mathcal{T}_h} \le |||(\boldsymbol{v}_h, \hat{\boldsymbol{v}}_h)|||_{1,\mathcal{T}_h} \le \sqrt{5/4}\|(\boldsymbol{v}_h, \hat{\boldsymbol{v}}_h)\|_{1,\mathcal{T}_h},$$

$$\text{and} \qquad \|q_h\|_{0,\mathcal{T}_h} \le |||q_h|||_{0,\mathcal{T}_h} \le \sqrt{5/4}\|q_h\|_{0,\mathcal{T}_h}.$$

These estimates are a direct consequence of (13) and the definition of the norms.

The ellipticity of the bilinear form $\mathbf{a}_h$ now follows with standard arguments, cf. e.g., [3, 26]. We provide a sketch of the proof in order to keep track of the constants.

Proposition 6.6 (Ellipticity)**.** *The bilinear form $\mathbf{a}_h$ is uniformly elliptic, i.e., there holds*

$$\mathbf{a}_h(\boldsymbol{u}_h, \hat{\boldsymbol{u}}_h; \boldsymbol{u}_h, \hat{\boldsymbol{u}}_h) \ge \tfrac{1}{2}\|(\boldsymbol{u}_h, \hat{\boldsymbol{u}}_h)\|^2_{1,\mathcal{T}_h}, \qquad \text{for all } \boldsymbol{u}_h \in \boldsymbol{V}_h \text{ and } \hat{\boldsymbol{u}}_h \in \widehat{\boldsymbol{V}}_h.$$

Proof. We use Cauchy-Schwarz and Young's inequality to bound the non-symmetric terms

$$-2\langle \boldsymbol{\partial}_n \boldsymbol{u}_h, \boldsymbol{u}_h - \hat{\boldsymbol{u}}_h\rangle_{\partial\mathcal{T}_h} \ge -2|\boldsymbol{\partial}_n \boldsymbol{u}_h|^2_{-1/2,\partial\mathcal{T}_h} - \tfrac{1}{2}|\boldsymbol{u}_h - \hat{\boldsymbol{u}}_h|^2_{1/2,\partial\mathcal{T}_h} \ge -\tfrac{1}{2}\|(\boldsymbol{u}_h, \hat{\boldsymbol{u}}_h)\|^2_{1,\mathcal{T}_h}.$$

Here we applied (15) for the last step. Adding the remaining terms, yields the result. □

Using the two pairs of energy norms, we obtain the following bounds.

Proposition 6.7 (Boundedness)**.** *For any element $\boldsymbol{u}_h$, $\boldsymbol{v}_h \in \boldsymbol{V}_h$, every $\hat{\boldsymbol{u}}_h$, $\hat{\boldsymbol{v}}_h \in \widehat{\boldsymbol{V}}_h$, and all functions $\boldsymbol{u} \in \boldsymbol{H}^1_0(\Omega) \cap \boldsymbol{H}^2(\mathcal{T}_h)$, the bilinear form $\mathbf{a}_h$ is bounded by*

$$\mathbf{a}_h(\boldsymbol{u}_h - \boldsymbol{u}, \hat{\boldsymbol{u}}_h - \boldsymbol{u}; \boldsymbol{v}_h, \hat{\boldsymbol{v}}_h) \le C_a |||(\boldsymbol{u}_h - \boldsymbol{u}, \hat{\boldsymbol{u}}_h - \boldsymbol{u})|||_{1,\mathcal{T}_h} \|(\boldsymbol{v}_h, \hat{\boldsymbol{v}}_h)\|_{1,\mathcal{T}_h},$$

with constant $C_a = \sqrt{5/4}$. Similarly, with $C_b = \sqrt{5d/4}$, there holds

$$\mathbf{b}_h(\boldsymbol{u}_h - \boldsymbol{u}, \hat{\boldsymbol{u}}_h - \boldsymbol{u}; p_h) \le C_b |||(\boldsymbol{u}_h - \boldsymbol{u}, \hat{\boldsymbol{u}}_h - \boldsymbol{u})|||_{1,\mathcal{T}_h} \|p_h\|_{0,\mathcal{T}_h},$$

$$\mathbf{b}_h(\boldsymbol{u}_h, \hat{\boldsymbol{u}}_h; p_h - p) \le C_b \, \|(\boldsymbol{u}_h, \hat{\boldsymbol{u}}_h)\|_{1,\mathcal{T}_h} |||p_h - p|||_{0,\mathcal{T}_h},$$

for any $\boldsymbol{u}_h \in \boldsymbol{V}_h$, $\hat{\boldsymbol{u}}_h \in \widehat{\boldsymbol{V}}_h$, $p_h \in Q_h$, and every $\boldsymbol{u} \in \boldsymbol{H}^1_0(\Omega) \cap \boldsymbol{H}^2(\mathcal{T}_h)$, $p \in L^2_0(\Omega) \cap H^1(\mathcal{T}_h)$.

Proof. By the Cauchy-Schwarz inequality, and the discrete trace inequality, we obtain

$$\langle \boldsymbol{\partial}_n \boldsymbol{v}_h, \boldsymbol{u}_h - \hat{\boldsymbol{u}}_h\rangle_{\partial\mathcal{T}_h} \le \tfrac{1}{2}|\boldsymbol{\partial}_n \boldsymbol{v}_h|_{-1/2,\partial\mathcal{T}_h}|\boldsymbol{u}_h - \hat{\boldsymbol{u}}_h|_{1/2,\partial\mathcal{T}_h} \le \|\boldsymbol{\nabla} \boldsymbol{v}_h\|_{0,\mathcal{T}_h}|\boldsymbol{u}_h - \hat{\boldsymbol{u}}_h|_{1/2,\partial\mathcal{T}_h}.$$

The first bound then follows by applying the Cauchy-Schwarz inequality to the remaining terms. The bounds for $\mathbf{b}_h$ follow similarly by noting that $|\operatorname{div} \boldsymbol{v}(x)| \le \sqrt{d}|\boldsymbol{\nabla} \boldsymbol{v}(x)|$. □

6.3. **A Fortin operator.** Next, we explicitly construct a Fortin operator [31], which is used for the proof of the discrete inf-sup condition for the bilinear form $\mathbf{b}_h$ in the next section. Let us define global projection operators $\boldsymbol{\Pi}_h : \boldsymbol{H}^1_0(\Omega) \to \boldsymbol{V}_h$ and $\widehat{\boldsymbol{\Pi}}_h : \boldsymbol{H}^1_0(\Omega) \to \widehat{\boldsymbol{V}}_h$ by

$$(\boldsymbol{\Pi}_h \boldsymbol{v})|_T := \boldsymbol{\Pi}^k_T \boldsymbol{v}|_T \qquad \text{and} \qquad (\widehat{\boldsymbol{\Pi}}_h \boldsymbol{v})|_E := \boldsymbol{\Pi}^k_E \boldsymbol{v}|_E$$

for all $T \in \mathcal{T}_h$ and $E \in \mathcal{E}_h$. An appropriate operator for our problem is then given by

$$(\boldsymbol{\Pi}_h, \widehat{\boldsymbol{\Pi}}_h) : \boldsymbol{H}^1_0(\Omega) \mapsto \boldsymbol{V}_h \times \widehat{\boldsymbol{V}}_h, \qquad \boldsymbol{u} \mapsto (\boldsymbol{\Pi}_h \boldsymbol{u}, \widehat{\boldsymbol{\Pi}}_h \boldsymbol{u}) \tag{16}$$

The following lemma summarizes the basic properties of this construction.

Lemma 6.8. *For any function* $\boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega)$ *there holds*

$$\mathbf{b}_h(\boldsymbol{\Pi}_h \boldsymbol{v}, \widehat{\boldsymbol{\Pi}}_h \boldsymbol{v}; q_h) = \mathbf{b}(\boldsymbol{v}, q_h) \qquad \text{for all } q_h \in Q_h; \tag{17}$$

we say that the operator $(\boldsymbol{\Pi}_h, \widehat{\boldsymbol{\Pi}}_h)$ *satisfies the* Fortin property. *Moreover, there holds*

$$\|(\boldsymbol{\Pi}_h \boldsymbol{v}, \widehat{\boldsymbol{\Pi}}_h \boldsymbol{v})\|_{1,\mathcal{T}_h} \le C_\Pi\, k_{\max}^{1/2} \|\boldsymbol{\nabla} \boldsymbol{v}\|_\Omega \qquad \text{for all } \boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega), \tag{18}$$

with constant C_Π *independent of the meshsize and the polynomial degree distribution.*

Proof. First, observe that for any $\boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega)$ and every element $T \in \mathcal{T}_h$ there holds

$$\begin{aligned} &-(\operatorname{div} \boldsymbol{\Pi}_h \boldsymbol{v}, q_h)_T + \langle \boldsymbol{\Pi}_h \boldsymbol{v} - \widehat{\boldsymbol{\Pi}}_h \boldsymbol{v}, q_h \boldsymbol{n} \rangle_{\partial T} = (\boldsymbol{\Pi}_h \boldsymbol{v}, \boldsymbol{\nabla} q_h)_T - \langle \widehat{\boldsymbol{\Pi}}_h \boldsymbol{v}, q_h \boldsymbol{n} \rangle_{\partial T} \\ &\quad = (\boldsymbol{v}, \boldsymbol{\nabla} q_h)_T - \langle \boldsymbol{v}, q_h \boldsymbol{n} \rangle_{\partial T} = -(\operatorname{div} \boldsymbol{v}, q_h)_T. \end{aligned}$$

Summing over all elements, and using the definition of the global operators, yields (17). In order to verify the bound (18), we use the estimates of Lemma 6.4, which yield

$$\begin{aligned} &\|\boldsymbol{\nabla} \boldsymbol{\Pi}_h \boldsymbol{v}\|_{0,T}^2 + |\boldsymbol{\Pi}_h \boldsymbol{v} - \widehat{\boldsymbol{\Pi}}_h \boldsymbol{v}|_{1/2,\partial T}^2 \\ &\quad = \|\boldsymbol{\nabla} \boldsymbol{\Pi}_h \boldsymbol{v}\|_{0,T}^2 + |\boldsymbol{\Pi}_h \boldsymbol{v} - \boldsymbol{v}|_{1/2,\partial T}^2 + |\boldsymbol{v} - \widehat{\boldsymbol{\Pi}}_h \boldsymbol{v}|_{1/2,\partial T}^2 \le C k_T \|\boldsymbol{\nabla} \boldsymbol{v}\|_{0,T}^2, \end{aligned}$$

with constant C independent of h_T and k_T. Using the element-wise definition of the Fortin operator, bounding k_T by $k_{\max}$, and summing over all elements yields the result. □

Remark 6.9. The simple proof of the previous result reveals that one could increase the polynomial degree for the pressure by one order, and thus obtain a Fortin operator also for *equal order* discretizations; see Section 8 for numerical results in this direction.

6.4. **The inf-sup stability.** The following stability result for the bilinear form $\mathbf{b}_h$ is the central ingredient for the a-priori error analysis of the hybrid DG method.

Proposition 6.10 (Inf-sup stability)**.** *There exists a constant* $\beta > 0$*, which is independent of the meshsize and the polynomial degree distribution, such that*

$$\sup_{(\boldsymbol{v}_h, \hat{\boldsymbol{v}}_h) \in \boldsymbol{V}_h \times \widehat{\boldsymbol{V}}_h} \frac{\mathbf{b}_h(\boldsymbol{v}_h, \hat{\boldsymbol{v}}_h; q_h)}{\|(\boldsymbol{v}_h, \hat{\boldsymbol{v}}_h)\|_{1,\mathcal{T}_h}} \ge \beta k_{\max}^{-1/2} \|q_h\|_{0,\mathcal{T}_h} \qquad \text{for all } q_h \in Q_h.$$

Proof. By the continuous inf-sup condition (3), there exists a function $\boldsymbol{u} \in \boldsymbol{H}_0^1(\Omega)$ such that

$$\beta_\Omega \|q_h\|_{0,\mathcal{T}_h} \|\boldsymbol{\nabla} \boldsymbol{u}\|_{0,\mathcal{T}_h} \le \mathbf{b}(\boldsymbol{u}, q_h) = \|q_h\|_{0,\mathcal{T}_h}^2.$$

We then use the Fortin projector $(\boldsymbol{\Pi}_h, \widehat{\boldsymbol{\Pi}}_h)$ to define a discrete pair of functions $(\boldsymbol{v}_h, \hat{\boldsymbol{v}}_h)$ that satisfies the required bound. By the properties stated in Lemma 6.8, we conclude that

$$\begin{aligned} \mathbf{b}_h(\boldsymbol{\Pi}_h \boldsymbol{u}, \widehat{\boldsymbol{\Pi}}_h \boldsymbol{u}; q_h) &= \mathbf{b}(\boldsymbol{u}, q_h) \ge \beta_\Omega \|q_h\|_{0,\mathcal{T}_h} \|\boldsymbol{\nabla} \boldsymbol{u}\|_{0,\mathcal{T}_h} \\ &\ge \beta_\Omega / C_\Pi\, k_{\max}^{-1/2} \|q_h\|_{0,\mathcal{T}_h} \|(\boldsymbol{\Pi}_h \boldsymbol{u}, \widehat{\boldsymbol{\Pi}}_h \boldsymbol{u})\|_{1,\mathcal{T}_h}, \end{aligned}$$

which proofs the assertion with constant $\beta = \beta_\Omega / C_\Pi$. □

Remark 6.11. A dependence of the inf-sup condition on the polynomial approximation order has been observed for various high order methods: Stability with a constant $\beta(k) = \beta k^{-(d-1)/2}$ has been shown to hold for conforming discretizations on quadrilaterals [59], for spectral methods [12], and for a non-symmetric discontinuous Galerkin method using $P_k - P_{k-2}$ elements on irregular quadrilateral and hexahedral meshes [61]. For some of the methods, this dependence can even be shown to be optimal [12]. According to the numerical evidence reported in [61], the dependence may be improved for discontinuous Galerkin methods. In fact, Proposition 6.10 yields an inf-sup stability constant $\beta(k) = \beta k^{-1/2}$ also in three dimensions.

Remark 6.12. An *hp* stability estimate similar to that of Proposition 6.10 has been derived in [57] for a class of related discontinuous Galerkin methods on quadrilateral and hexahedral meshes. There, the authors employ the Raviart-Thomas projector for the construction of a Fortin operator, and the resulting stability estimate is one order sub-optimal in k. Here, we use the L^2 projection to construct the Fortin operator, which yields sub-optimality only by one half order in k. Moreover, our result applies to more general meshes.

A combination of the previous results yields the well-definedness of the hybrid DG method.

Corollary 6.13. *Method 4.1 has a unique solution* $\boldsymbol{u}_h \in \boldsymbol{V}_h$, $\hat{\boldsymbol{u}}_h \in \widehat{\boldsymbol{V}}_h$, *and* $p_h \in Q_h$.

Proof. The result follows from Propositions 6.6 and 6.10 by Brezzi's theorem [14]. □

6.5. **A-priori error estimates.** As a final ingredient for the a-priori error analysis, let us characterize the approximation properties of the finite element spaces in Method 4.1 with respect to the energy norms used in our analysis.

Proposition 6.14 (Approximation)**.** *Let* $\boldsymbol{u} \in \boldsymbol{H}_0^1(\Omega)$ *be such that* $\boldsymbol{u}|_T \in \boldsymbol{H}^{s_T+1}(T)$ *for some* $s_T \geq 1$ *for any* $T \in \mathcal{T}_h$. *Then there holds*

$$\inf_{(\boldsymbol{v}_h \hat{\boldsymbol{v}}_h) \in \boldsymbol{V}_h \times \widehat{\boldsymbol{V}}_h} \|\!|\!|(\boldsymbol{u} - \boldsymbol{v}_h, \boldsymbol{u} - \hat{\boldsymbol{v}}_h)|\!|\!|_{1,\mathcal{T}_h} \leq C\Big(\sum\nolimits_T \frac{h_T^{2\min\{k_T,s_T\}}}{k_T^{2s_T-1}} \|\boldsymbol{u}\|^2_{H^{s_T+1}(T)}\Big)^{1/2}.$$

If $p \in L_0^2(\Omega)$ *such that* $p|_T \in H^{s_T}(T)$ *for some* $s_T \geq 1$ *and all* $T \in \mathcal{T}_h$, *then*

$$\inf_{q_h \in Q_h} \|\!|\!|p - q_h|\!|\!|_{0,\mathcal{T}_h} \leq C\Big(\sum\nolimits_T \frac{h_T^{2\min\{k_T,s_T\}}}{k_T^{2s_T}} \|p\|^2_{H^{s_T}(T)}\Big)^{1/2}.$$

The constant C *of both estimates may depend on the regularity of the functions, but is otherwise independent of the functions, the meshsize, and the polynomial degree distribution.*

Proof. The estimates are a direct consequence of the results stated in Lemma 6.3. □

By combining the consistency of the method, the properties of the bilinear forms, and the approximation result, we now obtain the basic a-priori error estimate in the energy norm.

Theorem 6.15 (A-priori estimate)**.** *Assume that the solution* $(\boldsymbol{u}, p)$ *of problem* (2) *is regular, i.e.* $\boldsymbol{u}|_T \in \boldsymbol{H}^{s_T+1}(T)$ *and* $p|_T \in \boldsymbol{H}^{s_T}(T)$ *with some* $s_T \geq 1$ *on any* $T \in \mathcal{T}_h$. *Moreover, let* $(\boldsymbol{u}_h, \hat{\boldsymbol{u}}_h, p_h)$ *denote the solution of Method 4.1. Then, the error estimate*

$$\begin{aligned} &\|(\boldsymbol{u} - \boldsymbol{u}_h, \boldsymbol{u} - \hat{\boldsymbol{u}}_h)\|_{1,\mathcal{T}_h} + k_{\max}^{-1/2}\|p - p_h\|_{0,\mathcal{T}_h} \\ &\qquad \leq C k_{\max}^{1/2} \left(\sum\nolimits_T \frac{h_T^{2\min\{s_T,k_T\}}}{k_T^{2s_T-1}} \|\boldsymbol{u}\|^2_{H^{s_T+1}(T)} + \frac{h_T^{2\min\{s_T,k_T\}}}{k_T^{2s_T}} \|p\|^2_{H^{s_T}(T)}\right)^{1/2}, \end{aligned}$$

holds with a constant C *independent of the meshsize and the polynomial degree distribution.*

Proof. The result follows with standard arguments, see e.g. [14, 13, 29] and also [57]. We present a detailed proof only to keep track of the dependence on the polynomial degree:
Step 1: Let $\boldsymbol{z}_h \in \boldsymbol{V}_h$, $\hat{\boldsymbol{z}}_h \in \widehat{\boldsymbol{V}}_h$ be such that $\mathbf{b}_h(\boldsymbol{z}_h, \hat{\boldsymbol{z}}_h; q_h) = 0$ for all $q_h \in Q_h$. Due to the coercivity of the bilinear form $\mathbf{a}_h$ and the consistency of the method, there holds

$$\begin{aligned}
&\tfrac{1}{2}\|(\boldsymbol{u}_h - \boldsymbol{z}_h, \hat{\boldsymbol{u}}_h - \hat{\boldsymbol{z}}_h)\|^2_{1,\mathcal{T}_h} \\
&\quad \le \mathbf{a}_h(\boldsymbol{u}_h - \boldsymbol{z}_h, \hat{\boldsymbol{u}}_h - \hat{\boldsymbol{z}}_h; \boldsymbol{u}_h - \boldsymbol{z}_h, \hat{\boldsymbol{u}}_h - \hat{\boldsymbol{z}}_h) \\
&\quad = \mathbf{a}_h(\boldsymbol{u} - \boldsymbol{z}_h, \boldsymbol{u} - \hat{\boldsymbol{z}}_h; \boldsymbol{u}_h - \boldsymbol{z}_h, \hat{\boldsymbol{u}}_h - \hat{\boldsymbol{z}}_h) + \mathbf{b}_h(\boldsymbol{u}_h - \boldsymbol{z}_h, \hat{\boldsymbol{u}}_h - \hat{\boldsymbol{z}}_h; p - p_h).
\end{aligned}$$

Moreover, we have $\mathbf{b}_h(\boldsymbol{u}_h - \boldsymbol{z}_h, \hat{\boldsymbol{u}}_h - \hat{\boldsymbol{z}}_h, p_h) = \mathbf{b}_h(\boldsymbol{u}_h - \boldsymbol{z}_h, \hat{\boldsymbol{u}}_h - \hat{\boldsymbol{z}}_h, q_h) = 0$ for all $q_h \in Q_h$. Replacing the last term, and using the boundedness of the bilinear forms, we obtain

$$\begin{aligned}
&\tfrac{1}{2}\|(\boldsymbol{u}_h - \boldsymbol{z}_h, \hat{\boldsymbol{u}}_h - \hat{\boldsymbol{z}}_h)\|^2_{1,\mathcal{T}_h} \\
&\quad \le \big(|\!|\!|(\boldsymbol{u} - \boldsymbol{z}_h, \boldsymbol{u} - \hat{\boldsymbol{z}}_h)|\!|\!|_{1,\mathcal{T}_h} + C_b |\!|\!|p - q_h|\!|\!|_{0,\mathcal{T}_h}\big)\|(\boldsymbol{u}_h - \boldsymbol{z}_h, \hat{\boldsymbol{u}}_h - \hat{\boldsymbol{z}}_h)\|_{1,\mathcal{T}_h},
\end{aligned}$$

which holds uniformly for all appropriate functions $\boldsymbol{z}_h$, $\hat{\boldsymbol{z}}_h$.
Step 2: Following [13, Rem. 4.10], we show next that functions $(\boldsymbol{z}_h, \hat{\boldsymbol{z}}_h)$ satisfying the discrete constraint have sufficient approximation properties: Let $(\boldsymbol{v}_h, \hat{\boldsymbol{v}}_h) \in \boldsymbol{V}_h \times \widehat{\boldsymbol{V}}_h$ be given, and $\boldsymbol{u}$ denote the solution of the Stokes problem. Due to the discrete stability and boundedness of $\mathbf{b}_h$, a pair of functions $(\boldsymbol{w}_h, \hat{\boldsymbol{w}}_h) \in \boldsymbol{V}_h \times \widehat{\boldsymbol{V}}_h$ exists, such that

$$\mathbf{b}_h(\boldsymbol{w}_h, \hat{\boldsymbol{w}}_h; q_h) = \mathbf{b}_h(\boldsymbol{u} - \boldsymbol{v}_h, \boldsymbol{u} - \hat{\boldsymbol{v}}_h; q_h) \qquad \text{for all } q_h \in Q_h,$$

with the additional property that

$$\|(\boldsymbol{w}_h, \hat{\boldsymbol{w}}_h)\|_{1,\mathcal{T}_h} \le (C_b/\beta)\, k_{\max}^{1/2}\, |\!|\!|(\boldsymbol{u} - \boldsymbol{v}_h, \boldsymbol{u} - \hat{\boldsymbol{v}}_h)|\!|\!|_{1,\mathcal{T}_h},$$

and β from Proposition 6.10. Due to the equivalence of the norms on the finite dimensional spaces, we also have $|\!|\!|(\boldsymbol{w}_h, \hat{\boldsymbol{w}}_h)|\!|\!|_{1,\mathcal{T}_h} \le c k_{\max}^{1/2} |\!|\!|(\boldsymbol{u} - \boldsymbol{v}_h, \boldsymbol{u} - \hat{\boldsymbol{v}}_h)|\!|\!|_{1,\mathcal{T}_h}$. Hence, the functions $\boldsymbol{z}_h = \boldsymbol{v}_h + \boldsymbol{w}_h$ and $\hat{\boldsymbol{z}}_h = \hat{\boldsymbol{v}}_h + \hat{\boldsymbol{w}}_h$ satisfy

$$\mathbf{b}_h(\boldsymbol{z}_h, \hat{\boldsymbol{z}}_h; q_h) = \mathbf{b}_h(\boldsymbol{v}_h, \hat{\boldsymbol{v}}_h; q_h) + \mathbf{b}_h(\boldsymbol{u} - \boldsymbol{v}_h, \boldsymbol{u} - \hat{\boldsymbol{v}}_h; q_h) = 0,$$

where the last equality is due to the consistency of the method. Moreover,

$$|\!|\!|(\boldsymbol{u} - \boldsymbol{z}_h, \boldsymbol{u} - \hat{\boldsymbol{z}}_h)|\!|\!|_{1,\mathcal{T}_h} \le (1 + c k_{\max}^{1/2}) |\!|\!|(\boldsymbol{u} - \boldsymbol{v}_h, \boldsymbol{u} - \boldsymbol{v}_h)|\!|\!|_{1,\mathcal{T}_h},$$

which shows that $(\boldsymbol{z}_h, \hat{\boldsymbol{z}}_h)$ is as good approximating as arbitrary functions $(\boldsymbol{v}_h, \hat{\boldsymbol{v}}_h) \in \boldsymbol{V}_h \times \widehat{\boldsymbol{V}}_h$, and we obtain the a-priori estimate for the functions $(\boldsymbol{u}_h,\ \hat{\boldsymbol{u}}_h)$ by application of the approximation result of Lemma 6.3.
Step 3: For estimating the error in p_h, we again use the discrete inf-sup stability, to obtain

$$\begin{aligned}
\beta k_{\max}^{-1/2}\, \|q_h - p_h\|_{0,\mathcal{T}_h} &\le \sup_{(\boldsymbol{v}_h, \hat{\boldsymbol{v}}_h)} \frac{\mathbf{b}_h(\boldsymbol{v}_h, \hat{\boldsymbol{v}}_h; q_h - p_h)}{\|(\boldsymbol{v}_h, \hat{\boldsymbol{v}}_h)\|_{1,\mathcal{T}_h}} \\
&\le \sup_{(\boldsymbol{v}_h, \hat{\boldsymbol{v}}_h)} \frac{\mathbf{b}_h(\boldsymbol{v}_h, \hat{\boldsymbol{v}}_h; q_h - p) + \mathbf{a}_h(\boldsymbol{u}_h - \boldsymbol{u}, \hat{\boldsymbol{u}}_h - \boldsymbol{u}; \boldsymbol{v}_h, \hat{\boldsymbol{v}}_h)}{\|(\boldsymbol{v}_h, \hat{\boldsymbol{v}}_h)\|_{1,\mathcal{T}_h}} \\
&\le C_b\, |\!|\!|q_h - p|\!|\!|_{0,\mathcal{T}_h} + |\!|\!|(\boldsymbol{u} - \boldsymbol{u}_h, \boldsymbol{u} - \hat{\boldsymbol{u}}_h)|\!|\!|_{1,\mathcal{T}_h}.
\end{aligned}$$

The estimate for the pressure then follows from Lemma 6.3 and the result for the velocity. □

Remark 6.16. Using the standard duality argument of Aubin-Nitsche, one can obtain corresponding estimates for the error in the L^2 norm, which are again optimal with respect to the meshsize, and slightly sub-optimal with respect to the polynomial degree. For this argument, we utilize the symmetry of the formulation; see the discussion on adjoint consistency in [5].

After having established the basic a-priori error estimates, we now turn to the construction and analysis of suitable a-posteriori error estimators, which may be of interest for applications.

7. A posteriori error estimates

The subseqent analysis relies on ideas from thje a-posteriori error estimation of non-conforming, mixed, and discontinuous Galerkin methods for elliptic problems [11, 16, 56, 64, 17]; see also [36, 46] for the *hp* a-posteriori error analysis of continuous and discontinuous Galerkin methods. Our presentation closely follows the approach outlined in [41].

7.1. **A residual error estimator.** An a-posteriori error estimator for the hybrid DG method has to measure the residuals in the two equations in (1), and also the discontinuity of the discrete solution. In the following, we consider the *residual error estimator* of the form

$$\eta_R := \Big(\sum\nolimits_T \eta_{c,T}^2 + \eta_{nc,T}^2\Big)^{1/2}$$

with local contributions defined by

$$\begin{aligned} \eta_{c,T}^2 &:= \tfrac{h_T^2}{k_T^2}\|\boldsymbol{f} + \Delta \boldsymbol{u}_h - \boldsymbol{\nabla} p_h\|_{0,T}^2 + \|\operatorname{div} \boldsymbol{u}_h\|_{0,T}^2 + k_T|\boldsymbol{u}_h - \hat{\boldsymbol{u}}_h|_{1/2,\partial T}^2, \\ \eta_{nc,T}^2 &:= \|\boldsymbol{\nabla}(\boldsymbol{u}_h - \widetilde{\boldsymbol{u}}_h)\|_{0,T}^2, \end{aligned}$$

where $\widetilde{\boldsymbol{u}}_h \in \boldsymbol{H}_0^1(\Omega)$ is some conforming approximation for the discrete solution $\boldsymbol{u}_h$. As we will show, the choice of $\widetilde{\boldsymbol{u}}_h$ only affects the efficiency, but not the reliability of the estimator. A particular construction that can be realized by local post-processing is discussed in Section 7.3.

Remark 7.1. Using the continuity of the numerical fluxes $\boldsymbol{\sigma}_h \cdot \boldsymbol{n}$ stated in Proposition 5.3, one can show that $|\boldsymbol{u}_h - \hat{\boldsymbol{u}}_h|_{1/2,\partial T}$ is of the same order as the jump $|[\boldsymbol{\partial}_n \boldsymbol{u}_h - p_h \boldsymbol{n}]|_{-1/2,\partial\mathcal{T}_h}$ of the normal flux, which arises in residual error estimates of conforming discretizations [62, 2].

7.2. **Reliability.** As in [41], we start by defining a projection of the finite element solution onto the continuous solution space: Let $\widetilde{\boldsymbol{u}} \in \boldsymbol{H}_0^1(\Omega)$ and $\widetilde{p} \in L_0^2(\Omega)$ be the solution of

$$\mathcal{B}(\widetilde{\boldsymbol{u}}, \widetilde{p}; \boldsymbol{v}, q) = \mathcal{B}(\boldsymbol{u}_h, p_h; \boldsymbol{v}, q) \qquad \text{for all } \boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega) \text{ and } q \in L_0^2(\Omega), \tag{19}$$

with $\mathcal{B}$ denoting the combined bilinear form of the Stokes problem on the continuous level; cf. Section 2. Due to the inf-sup stability condition (5) for the bilinear form $\mathcal{B}$, existence of a unique solution $(\widetilde{\boldsymbol{u}}, \widetilde{p})$ is granted by the Babuška-Aziz lemma [6]. By means of the triangle inequality, the error can be decomposed into a *conforming* and a *non-conforming* part

$$\begin{aligned} &\|(\boldsymbol{u} - \boldsymbol{u}_h, \boldsymbol{u} - \hat{\boldsymbol{u}}_h)\|_{1,\mathcal{T}_h} + \|p - p_h\|_{0,\mathcal{T}_h} \\ &\qquad \le \underbrace{\|\boldsymbol{\nabla}(\boldsymbol{u} - \widetilde{\boldsymbol{u}})\|_{0,\mathcal{T}_h} + \|p - \widetilde{p}\|_{0,\mathcal{T}_h}}_{\text{conforming}} + \underbrace{\|(\widetilde{\boldsymbol{u}} - \boldsymbol{u}_h, \widetilde{\boldsymbol{u}} - \hat{\boldsymbol{u}}_h)\|_{1,\mathcal{T}_h} + \|\widetilde{p} - p_h\|_{0,\mathcal{T}_h}}_{\text{non-conforming}}, \end{aligned}$$

where we used that $\|(\boldsymbol{v}, \boldsymbol{v}|_{\mathcal{E}})\|_{1,\mathcal{T}_h} = \|\boldsymbol{\nabla} \boldsymbol{v}\|_{0,\mathcal{T}_h}$ for $\boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega)$, since the jump terms vanish for functions which are continuous across interfaces. The two error components can now be estimated separately.

Proposition 7.2 (Conforming error)**.** *The conforming error component is bounded by*

$$\|\boldsymbol{\nabla}(\boldsymbol{u} - \widetilde{\boldsymbol{u}})\|_{0,\mathcal{T}_h} + \|p - \widetilde{p}\|_{0,\mathcal{T}_h} \le C\big(\sum\nolimits_T \eta_{c,T}^2\big)^{1/2},$$

with a constant C independent of the meshsize and the polynomial degree distribution.

Proof. The starting point for the following arguments is the continuous inf-sup condition (5), which implies that

$$c_\Omega(\|\boldsymbol{\nabla}(\boldsymbol{u}-\widetilde{\boldsymbol{u}})\|_{1,\mathcal{T}_h}+\|p-\widetilde{p}\|_{0,\mathcal{T}_h}) \le \sup_{(\boldsymbol{v},q)\in \boldsymbol{H}_0^1(\Omega)\times L_0^2(\Omega)} \frac{\mathcal{B}(\boldsymbol{u}-\boldsymbol{u}_h,p-p_h;\boldsymbol{v},q)}{\|\boldsymbol{\nabla}\boldsymbol{v}\|_{0,\mathcal{T}_h}+\|q\|_{0,\mathcal{T}_h}}.$$

Let us further estimate the numerator: By the flux continuity of Proposition 5.3, we have

$$\langle \boldsymbol{\partial}_n\boldsymbol{u}_h - p_h\boldsymbol{n} - \alpha_T(\boldsymbol{u}_h-\hat{\boldsymbol{u}}_h),\boldsymbol{v}\rangle_{\partial\mathcal{T}_h} = 0 \qquad \text{for all } \boldsymbol{v}\in \boldsymbol{H}_0^1(\Omega).$$

Note that we make use of the two compatibility conditions at this point; cf Remark 3.1. Adding this term to $\mathcal{B}(\boldsymbol{u}-\boldsymbol{u}_h,p-p_h;\boldsymbol{v},q)$, we obtain

$$\begin{aligned}\mathcal{B}(\boldsymbol{u}-\boldsymbol{u}_h,p-p_h,\boldsymbol{v},q) &= (\boldsymbol{f},\boldsymbol{v})_{\mathcal{T}_h} - (\boldsymbol{\nabla}\boldsymbol{u}_h - p_h I,\boldsymbol{\nabla}\boldsymbol{v})_{\mathcal{T}_h} + (\operatorname{div}\boldsymbol{u}_h,q)_{\mathcal{T}_h}\\ &\quad + \langle \boldsymbol{\partial}_n\boldsymbol{u}_h - p_h\boldsymbol{n} - \alpha(\boldsymbol{u}_h-\hat{\boldsymbol{u}}_h),\boldsymbol{v}\rangle_{\partial\mathcal{T}_h} =: (*).\end{aligned}$$

Since $(\boldsymbol{u}_h,\hat{\boldsymbol{u}}_h)$ is the discrete solution, we also know that for any $\boldsymbol{v}_h\in\boldsymbol{V}_h$

$$\begin{aligned}0 &= (\boldsymbol{f},\boldsymbol{v}_h)_{\mathcal{T}_h} - \mathcal{B}_h(\boldsymbol{u}_h,\hat{\boldsymbol{u}}_h,p_h;\boldsymbol{v}_h,0,0)\\ &= (\boldsymbol{f},\boldsymbol{v}_h)_{\mathcal{T}_h} - (\boldsymbol{\nabla}\boldsymbol{u}_h - p_hI,\boldsymbol{\nabla}\boldsymbol{v}_h)_{\mathcal{T}_h} + \langle\boldsymbol{\partial}_n\boldsymbol{v}_h,\boldsymbol{u}_h-\hat{\boldsymbol{u}}_h\rangle_{\partial\mathcal{T}_h}\\ &\qquad + \langle \boldsymbol{\partial}_n\boldsymbol{u}_h - p_h\boldsymbol{n} - \alpha(\boldsymbol{u}_h-\hat{\boldsymbol{u}}_h),\boldsymbol{v}_h\rangle_{\partial\mathcal{T}_h}.\end{aligned}$$

Subtracting this term from $(*)$, and integrating by parts, further yields

$$\begin{aligned}(*) &= (\boldsymbol{f}+\Delta\boldsymbol{u}_h - \boldsymbol{\nabla}p_h,\boldsymbol{v}-\boldsymbol{v}_h)_{\mathcal{T}_h} + (\operatorname{div}\boldsymbol{u}_h,q)_{\partial\mathcal{T}_h}\\ &\qquad - \langle\alpha(\boldsymbol{u}_h-\hat{\boldsymbol{u}}_h),\boldsymbol{v}-\boldsymbol{v}_h\rangle_{\partial\mathcal{T}_h} - \langle\boldsymbol{\partial}_n\boldsymbol{v}_h,\boldsymbol{u}_h-\hat{\boldsymbol{u}}_h\rangle_{\partial\mathcal{T}_h}.\end{aligned}$$

The individual terms can further be estimated by the Cauchy-Schwarz inequality, which yields

$$\begin{aligned}(*) &\le \sum\nolimits_T \|\boldsymbol{f}+\Delta\boldsymbol{u}_h-\boldsymbol{\nabla}p_h\|_{0,T}\|\boldsymbol{v}-\boldsymbol{v}_h\|_{0,T} + \sum\nolimits_T |\boldsymbol{u}_h-\hat{\boldsymbol{u}}_h|_{1/2,\partial T}|\boldsymbol{v}-\boldsymbol{v}_h|_{1/2,\partial T}\\ &\qquad + \|\operatorname{div}\boldsymbol{u}_h\|_{0,\mathcal{T}_h}\|q\|_{0,\mathcal{T}_h} + |\boldsymbol{u}_h-\hat{\boldsymbol{u}}_h|_{1/2,\partial\mathcal{T}_h}|\boldsymbol{\partial}_n\boldsymbol{v}_h|_{-1/2,\partial\mathcal{T}_h}\end{aligned}$$

for all test functions $\boldsymbol{v}_h\in V_h$. By Lemma 6.3, we can define $\boldsymbol{v}_h$ element-wise, such that

$$\|\boldsymbol{v}-\boldsymbol{v}_h\|_{0,T}\le C\tfrac{h_T}{k_T}\|\boldsymbol{\nabla}\boldsymbol{v}\|_T \quad \text{and} \quad |\boldsymbol{v}-\boldsymbol{v}_h|_{1/2,\partial T}\le C\sqrt{k_T}\|\boldsymbol{\nabla}\boldsymbol{v}\|_T$$

Moreover, due to the discrete trace inequality (15), we have $|\boldsymbol{\partial}_n\boldsymbol{v}_h|_{-1/2,h}\le\frac{1}{2}\|\boldsymbol{\nabla}\boldsymbol{v}_h\|_{0,\mathcal{T}_h}$. The proof is completed by inserting these estimates, summing over all elements, and using the Cauchy-Schwarz inequality and the previous estimates. □

We can also obtain an upper bound for the non-conforming error, which is similar to that for second order elliptic problems given in [41].

Proposition 7.3 (Non-conforming error)**.** *The non-conforming error is bounded by*

$$\|\boldsymbol{\nabla}(\widetilde{\boldsymbol{u}}-\boldsymbol{u}_h)\|_{0,\mathcal{T}_h}+\|\widetilde{p}-p_h\|_{0,\mathcal{T}_h}\le C\inf_{\boldsymbol{v}\in\boldsymbol{H}_0^1(\Omega)}\|\boldsymbol{\nabla}(\boldsymbol{u}_h-\boldsymbol{v})\|_{0,\mathcal{T}_h},$$

with a constant C that is independent of the meshsize and the approximation order.

Proof. By the continuous inf-sup condition (3) for the divergence operator, and the definition of the functions $\widetilde{\boldsymbol{u}}$ and $\widetilde{p}$ in (19), we conclude that

$$\beta_\Omega\|\widetilde{p}-p_h\|_{0,\mathcal{T}_h}\le\sup_{\boldsymbol{v}\in\boldsymbol{H}_0^1(\Omega)}\frac{\mathbf{b}(\widetilde{p}-p_h,\boldsymbol{v})}{\|\boldsymbol{\nabla}\boldsymbol{v}\|_{0,\mathcal{T}_h}}=\sup_{\boldsymbol{v}\in\boldsymbol{H}_0^1(\Omega)}\frac{\mathbf{a}(\boldsymbol{u}_h-\widetilde{\boldsymbol{u}},\boldsymbol{v})}{\|\boldsymbol{\nabla}\boldsymbol{v}\|_{0,\mathcal{T}_h}}\le\|\boldsymbol{\nabla}(\boldsymbol{u}_h-\widetilde{\boldsymbol{u}})\|_{0,\mathcal{T}_h}.$$

To estimate the non-conforming error, it remains to bound the error in the velocity. From (19), we obtain that $b(\widetilde{\boldsymbol{u}}, q) = \mathbf{b}(\boldsymbol{u}_h, q)$ for all $q \in L^2_0$, and that for all $\boldsymbol{v} \in \boldsymbol{H}^1_0(\Omega)$ there holds

$$\begin{aligned}\|\boldsymbol{\nabla}(\boldsymbol{u}_h - \widetilde{\boldsymbol{u}})\|^2_{0,\mathcal{T}_h} &= \mathbf{a}(\boldsymbol{u}_h - \widetilde{\boldsymbol{u}}, \boldsymbol{u}_h - \widetilde{\boldsymbol{u}}) = \mathbf{a}(\boldsymbol{u}_h - \widetilde{\boldsymbol{u}}, \boldsymbol{u}_h - \boldsymbol{v}) + \mathbf{b}(\boldsymbol{u}_h - \boldsymbol{v}, p_h - \widetilde{p})\\ &\le \big(\|\boldsymbol{\nabla}(\boldsymbol{u}_h - \widetilde{\boldsymbol{u}})\|_{0,\mathcal{T}_h} + C_b\|\widetilde{p} - p_h\|_{0,\mathcal{T}_h}\big)\|\boldsymbol{\nabla}(\boldsymbol{u}_h - \boldsymbol{v})\|_{0,\mathcal{T}_h}.\end{aligned}$$

Together with the estimate for $\|\widetilde{p} - p_h\|_{0,\mathcal{T}_h}$, and by division with $\|\boldsymbol{\nabla}(\boldsymbol{u}_h - \widetilde{\boldsymbol{u}})\|_{0,\mathcal{T}_h}$, we obtain

$$\|\boldsymbol{\nabla}(\boldsymbol{u}_h - \widetilde{\boldsymbol{u}})\|_{0,\mathcal{T}_h} \le (1 + C_b\beta_\Omega^{-1})\|\boldsymbol{\nabla}(\boldsymbol{u}_h - \boldsymbol{v})\|_{0,\mathcal{T}_h} \qquad \text{for all } \boldsymbol{v} \in \boldsymbol{H}^1_0(\Omega).$$

The conforming error component can then be bounded by the infimum over all $\boldsymbol{v} \in \boldsymbol{H}^1_0$. □

The reliability of the residual estimator now follows directly from Propositions 7.2 and 7.3.

Theorem 7.4 (Reliability). *The residual estimator η_R is reliable, i.e., there holds*

$$\|(\boldsymbol{u} - \boldsymbol{u}_h, \boldsymbol{u} - \hat{\boldsymbol{u}}_h)\|_{1,\mathcal{T}_h} + \|p - p_h\|_{0,\mathcal{T}_h} \le C_R\, \eta_R$$

with a constant C_R independent of the meshsize and the polynomial degree distribution.

7.3. **Efficiency.** In order to establish the efficiency of the residual-estimator, we will show in the following, that the individual terms appearing in the definition of the element contributions can be bounded by the jump terms, which are already part of the energy norm.

Lemma 7.5. *For any $T \in \mathcal{T}_h$, the element residuals can be bounded from above by*

$$\tfrac{h_T}{k_T}\|\boldsymbol{f} + \Delta\boldsymbol{u}_h - \boldsymbol{\nabla} p_h\|_{0,T} \le Ck_T|\boldsymbol{u}_h - \hat{\boldsymbol{u}}_h|_{1/2,\partial T} + \mathrm{osc}_T(\boldsymbol{f}),$$

with a constant C only depending on the shape of the element.

As usual, the local and global data oscillation terms are defined by

$$\mathrm{osc}_T(\boldsymbol{f}) := \tfrac{h_T}{k_T}\|\boldsymbol{\Pi}^k_T\boldsymbol{f} - \boldsymbol{f}\|_{0,T} \qquad \text{respectively} \qquad \mathrm{osc}(\boldsymbol{f}) := \Big(\sum\nolimits_T \mathrm{osc}_T(\boldsymbol{f})^2\Big)^{1/2}.$$

Proof. The residuals are first split into a discrete part and a data oscillation term, namely

$$\|\boldsymbol{f} + \Delta\boldsymbol{u}_h - \boldsymbol{\nabla} p_h\|_{0,T} \le \|\boldsymbol{\Pi}^k_T\boldsymbol{f} + \Delta\boldsymbol{u}_h - \boldsymbol{\nabla} p_h\|_{0,T} + \|\boldsymbol{f} - \boldsymbol{\Pi}^k_T\boldsymbol{f}\|_{0,T} = (i) + (ii).$$

Since $(\boldsymbol{u}_h, p_h)$ is the discrete solution, we obtain for any $\boldsymbol{v}_h \in \boldsymbol{\mathcal{P}}_k(T)$ that

$$\begin{aligned}(\boldsymbol{\Pi}_T\boldsymbol{f} + \Delta\boldsymbol{u}_h - \boldsymbol{\nabla} p_h, \boldsymbol{v}_h)_T &= \alpha_T\langle \boldsymbol{u}_h - \hat{\boldsymbol{u}}_h, \boldsymbol{v}_h\rangle_{\partial T} - \langle \boldsymbol{\partial}_n\boldsymbol{v}_h, \boldsymbol{u}_h - \hat{\boldsymbol{u}}_h\rangle_{\partial T}\\ &\le |\boldsymbol{u}_h - \hat{\boldsymbol{u}}_h|_{1/2,\partial T}\big(|\boldsymbol{v}_h|_{1/2,\partial T} + |\boldsymbol{\partial}_n\boldsymbol{v}_h|_{-1/2,\partial T}\big).\end{aligned}$$

Using the discrete trace inequalities (13) and (15) for the last term, yields

$$(i) = \sup\nolimits_{\boldsymbol{v}_h \in \boldsymbol{\mathcal{P}}_k(T)}(\boldsymbol{\Pi}^k_T\boldsymbol{f} + \Delta\boldsymbol{u}_h - \boldsymbol{\nabla} p_h, \boldsymbol{v}_h)_T/\|\boldsymbol{v}_h\|_{0,T} \le C\tfrac{k_T^2}{h_T}|\boldsymbol{u}_h - \hat{\boldsymbol{u}}_h|_{1/2,\partial T},$$

and scaling by $\frac{h_T}{k_T}$ yields the required bound. □

As a next step, let us derive the corresponding estimate for the divergence residual.

Lemma 7.6. *There holds $\|\operatorname{div}\boldsymbol{u}_h\|_{0,\mathcal{T}_h} \le |\boldsymbol{u}_h - \hat{\boldsymbol{u}}_h|_{1/2,\partial\mathcal{T}_h}$.*

Proof. Since $(\boldsymbol{u}_h, \hat{\boldsymbol{u}}_h)$ is the solution of the discrete problem (10), we obtain

$$(\operatorname{div} \boldsymbol{u}_h, q_h)_{\mathcal{T}_h} = \langle \boldsymbol{u}_h - \hat{\boldsymbol{u}}_h, q_h \boldsymbol{n} \rangle_{\partial\mathcal{T}_h} \leq |\boldsymbol{u}_h - \hat{\boldsymbol{u}}_h|_{1/2,\partial\mathcal{T}_h} \|q_h\|_{0,\mathcal{T}_h} \qquad \text{for all } q_h \in Q_h.$$

Using the Gauß-Green formula, we can further estimate the integral of the divergence by

$$(\operatorname{div} \boldsymbol{u}_h, 1)_{\mathcal{T}_h} = \langle \boldsymbol{u}_h, 1\boldsymbol{n} \rangle_{\partial\mathcal{T}_h} = \langle \boldsymbol{u}_h - \hat{\boldsymbol{u}}_h, 1\boldsymbol{n} \rangle_{\partial\mathcal{T}_h} \leq |\boldsymbol{u}_h - \hat{\boldsymbol{u}}_h|_{1/2,\partial\mathcal{T}_h} \|1\|_{0,\mathcal{T}_h},$$

which implies that the average $\overline{\operatorname{div} \boldsymbol{u}_h} = |\Omega|^{-1} \int_\Omega \operatorname{div} \boldsymbol{u}_h \, dx$ of the divergence can be estimated by the jump terms as well. A combination of the two estimates yields the result. □

In order to establish a bound for the non-conforming error, we have to specify the function $\widetilde{\boldsymbol{u}}_h$ used in the definition of $\eta_{nc,T}$. To estimate the non-conforming error, we require that $\mathcal{T}_h$ can be turned into a conforming simplicial mesh $\overline{\mathcal{T}_h}$ without drastically changing the size or shape regularity of the elements. Such a mesh $\mathcal{T}_h$ is then called *κ-regular-closable*. For details, see Section A.3 and Remark A.11.

Lemma 7.7. *Let $\mathcal{T}_h$ be κ-regular-closable in the sense of Definition A.10. Then there exists an averaging operator $\Pi_{av} : \boldsymbol{V}_h \to \boldsymbol{H}_0^1(\Omega)$ that satisfies*

$$\|\boldsymbol{\nabla}(\boldsymbol{v}_h - \boldsymbol{\Pi}_{av}\boldsymbol{v}_h)\|_{0,\mathcal{T}_h} \leq C_\kappa \Big(\sum\nolimits_T k_T^2 |\boldsymbol{v}_h - \hat{\boldsymbol{v}}|^2_{1/2,\partial T} \Big)^{1/2} \tag{20}$$

for all $\boldsymbol{v}_h \in \boldsymbol{V}_h$ and $\hat{\boldsymbol{v}} \in \boldsymbol{L}^2(\mathcal{E})$ with a constant C_κ that is independent of the meshsize and the polynomial approximation order.

Proof. A detailed proof is given in the appendix. □

Remark 7.8. Similar averaging operators have been proposed and analyzed in [40, 52]; see also [15, 36] for corresponding *hp* estimates. Assuming a uniform polynomial degree $k_T = k$ for all elements, and choosing $\hat{\boldsymbol{v}} = \hat{\boldsymbol{v}}_h \in \widehat{\boldsymbol{V}}_h$, the statement simplifies to

$$\|\boldsymbol{\nabla}(\boldsymbol{v}_h - \Pi_{av}\boldsymbol{v}_h)\|_{0,\mathcal{T}_h} \leq C_\kappa k |\boldsymbol{v}_h - \hat{\boldsymbol{v}}_h|_{1/2,\partial\mathcal{T}_h},$$

which is sub-optimal by one order of k. In two dimensions and for 1-irregular hexahedral meshes, the dependence in k can be improved by one order, and the sub-optimality vanishes; cf. e.g. [15, 66] and Remark A.7.

As a consequence of the previous estimates, we obtain the following statement.

Theorem 7.9 (Efficiency)**.** *Let $\mathcal{T}_h$ be κ-regular-closable. Then the residual error estimator is efficient, i.e., there holds*

$$c_R \, \eta_R \leq \Big(\sum\nolimits_T k_T^2 |\boldsymbol{u}_h - \hat{\boldsymbol{u}}_h|^2_{1/2,\partial T} + \operatorname{osc}(\boldsymbol{f})^2 \Big)^{1/2}. \tag{21}$$

with a constant c_R independent of the meshsize and the polynomial degree distribution.

Proof. The result follows directly by combining Lemmas 7.5–7.7. □

Remark 7.10. Combining the reliability and efficiency estimates, and bounding k_T by $k_{\max}$, we obtain that, up to data oscillation terms,

$$k_{\max}^{-1} \eta_R \preceq \|\boldsymbol{u} - \boldsymbol{u}_h\|_{1,\mathcal{T}_h} + \|p - p_h\|_{0,\mathcal{T}_h} \preceq \eta_R.$$

Thus we observe a discrepancy in the upper and lower bounds for the estimator by one order of k. A similar sub-optimality of a-posteriori error estimates with respect to the polynomial degree has been observed by several authors, e.g. [46, 36] and the references given there.

7.4. **A simple jump estimator.** For the efficiency estimate of the residual estimator η_R, we verified that that the element residuals, the divergence term, and the non-conforming error component can be estimated from above by the jump terms. This motivates the following, much simpler construction, which we call the *jump-estimator*:

$$\eta_J := \big(\sum_T \eta_{J,T}^2\big)^{1/2} \qquad \text{with} \qquad \eta_{J,T} = |\boldsymbol{u}_h - \hat{\boldsymbol{u}}_h|_{1/2,\partial T}. \tag{22}$$

Efficiency and reliability for the jump-estimator directly follow from the results for the residual-estimator. Note that this estimator is particularly convenient from an implementation point of view, since it only involves terms which are already present in the discrete variational problem. Some numerical results for the jump-estimator are given in the following.

At the end of this section, let us note that both estimators are assembled from local contributions They are therefore well suited to guide adaptive mesh refinement [62, 51, 47].

8. Numerical results

For illustration of the theoretical results, we consider two test problems from [28]. We conduct convergence studies demonstrating the a-priori error estimates, and a series of numerical experiments with adaptive refinement illustrating the performance of the a-posteriori error estimator. All results are obtained with a self-made finite element code based on the Dune framework [10].

8.1. **Convergence studies.** The first test problem is a simple *colliding-flow* on the unit square $\Omega = (-1, 1)^2$. Boundary conditions are chosen, such that the exact solution is

$$\boldsymbol{u} = \big(20xy^3, 5x^4 - 5y^4\big)\,, \quad p = 60x^2y - 20y^3.$$

The hybrid discontinuous Galerkin method is used for the numerical solution on a sequence of uniformly refined meshes. Discretization errors in the L^2-norm and the energy norm are computed using the analytic solution.

In Table 1, we list the numerical errors obtained with different orders of approximation. The same polynomial degree is used for the velocities on the elements and the skeleton. We also include results for an equal approximation.

The observed errors and convergence rates are in good agreement with the theoretical predictions. Note that, due to the consistency of the method, and since the exact solution is a polynomial, the numerical error is zero for polynomial order $k = 4$.

8.2. **Backward facing step.** As a second test case, we consider a *backward-facing-step flow* on the geometry depicted in Figure 2. At the in- and outflow boundaries, we impose parabolic velocity profiles

$$\boldsymbol{u}(-2, y) := (\,8\,(1 - y)\,(y - 0.5)\,, 0\,) \quad \text{and} \quad \boldsymbol{u}(10, y) = (\,y(1 - y), 0\,)\,,$$

and we apply a *no-slip* condition on the rest of the boundary. The boundary velocities are chosen such that the compatibility condition $\int_{\partial\Omega} \boldsymbol{u} \cdot \boldsymbol{n}\, dx = 0$ is satisfied, which is due to the incompressibility of the fluid.

Table 8.2 compares the performance of the hybrid DG method studied in this paper on uniformly and adaptively refined meshes. As refinement indicators for the adaptive algorithm,

$\mathcal{P}_1 - \mathcal{P}_0$	L^2 norm	rate	energy norm	rate	η_J estimate	rate
0	$9.9944 \cdot 10^{0}$	–	$2.6287 \cdot 10^{1}$	–	$1.8425 \cdot 10^{1}$	–
1	$2.4845 \cdot 10^{0}$	2.01	$1.3692 \cdot 10^{1}$	0.94	$9.9243 \cdot 10^{0}$	0.89
2	$6.0697 \cdot 10^{-1}$	2.03	$6.9134 \cdot 10^{0}$	0.99	$5.0865 \cdot 10^{0}$	0.96
3	$1.4954 \cdot 10^{-1}$	2.02	$3.4625 \cdot 10^{0}$	1.00	$2.5622 \cdot 10^{0}$	0.99
$\mathcal{P}_2 - \mathcal{P}_1$	L^2 norm	rate	energy norm	rate	η_J estimate	rate
0	$1.8346 \cdot 10^{0}$	–	$4.2024 \cdot 10^{0}$	–	$4.2715 \cdot 10^{0}$	–
1	$2.3046 \cdot 10^{-1}$	2.99	$1.0615 \cdot 10^{0}$	1.99	$1.0775 \cdot 10^{0}$	1.99
2	$2.8307 \cdot 10^{-2}$	3.03	$2.6386 \cdot 10^{-1}$	2.01	$2.6742 \cdot 10^{-1}$	2.01
3	$3.4890 \cdot 10^{-3}$	3.02	$6.5568 \cdot 10^{-2}$	2.01	$6.6362 \cdot 10^{-2}$	2.01
$\mathcal{P}_3 - \mathcal{P}_2$	L^2 norm	rate	energy norm	rate	η_J estimate	rate
0	$1.7166 \cdot 10^{-1}$	–	$3.2031 \cdot 10^{-1}$	–	$3.6025 \cdot 10^{-1}$	–
1	$1.0375 \cdot 10^{-2}$	4.05	$3.9232 \cdot 10^{-2}$	3.03	$4.4069 \cdot 10^{-2}$	3.03
2	$6.4246 \cdot 10^{-4}$	4.01	$4.8699 \cdot 10^{-3}$	3.01	$5.4627 \cdot 10^{-3}$	3.01
3	$4.9786 \cdot 10^{-5}$	3.69	$6.3819 \cdot 10^{-4}$	2.93	$6.8020 \cdot 10^{-4}$	3.01
$\mathcal{P}_2 - \mathcal{P}_2$	L^2 norm	rate	energy norm	rate	η_J estimate	rate
0	$2.5644 \cdot 10^{0}$	–	$4.5101 \cdot 10^{0}$	–	$3.9772 \cdot 10^{0}$	–
1	$2.9708 \cdot 10^{-1}$	3.11	$1.0895 \cdot 10^{0}$	2.05	$9.6245 \cdot 10^{-1}$	2.05
2	$3.5398 \cdot 10^{-2}$	3.06	$2.6556 \cdot 10^{-1}$	2.04	$2.3413 \cdot 10^{-1}$	2.04
3	$4.3072 \cdot 10^{-3}$	3.04	$6.5374 \cdot 10^{-2}$	2.02	$5.7530 \cdot 10^{-2}$	2.02

TABLE 1. (Colliding flow) Errors of the numerical solution for different orders of inf-sup stable finite element approximations on a sequence of uniformly refined meshes. The right-most columns list the error predictions of the jump-estimator.

we use the local contributions of the jump-estimator (22), and we employ a Dörfler marking strategy, where the 5% of the elements with the largest error contribution are refined.

A sequence of computational meshes for a third order ($\mathcal{P}_3 - \mathcal{P}_2$) approximation is depicted in Figure 2. As expected, the mesh is mostly refined towards the re-entrant corner. No refinement is made near the inflow and outflow regions, where the velocity profile is quadratic and therefore can be represented perfectly in the approximation spaces.

9. CONCLUSION

In this article, we investigated a hybrid discontinuous Galerkin method for the Stokes problem.

We proved discrete stability and boundedness results, and showed that the inf-sup constant depends only weakly on the polynomial degree of approximation. Our result $\beta(k) = \beta k^{-1/2}$ is sharper than previous estimates for related discontinuous Galerkin methods [57]. Our analysis applies to irregular and hybrid meshes in two and three space dimensions. The established a-priori error estimates for the hybrid method are optimal with respect to the meshsize and only slightly sub-optimal with respect to the polynomial degree.

In the second part of the manuscript, we presented a rigorous *hp* analysis of two a-posteriori error estimators. As part of our results, we constructed and analyzed an averaging operator of

$k = 1$						
uniform refinement			adaptive refinement			
level	η_J estimate	rate	level	step	η_J estimate	rate
0	$1.49451 \cdot 10^0$	–	0.00	0	$1.49451 \cdot 10^0$	–
1	$1.03083 \cdot 10^0$	0.54	0.98	32	$7.97898 \cdot 10^{-1}$	0.93
2	$5.89461 \cdot 10^{-1}$	0.81	1.97	68	$4.16091 \cdot 10^{-1}$	0.95
3	$3.18865 \cdot 10^{-1}$	0.89	2.97	105	$2.07307 \cdot 10^{-1}$	1.01
$k = 2$						
uniform refinement			adaptive refinement			
level	η_J estimate	rate	level	step	η_J estimate	rate
0	$4.17980 \cdot 10^{-1}$	–	0.00	0	$4.17980 \cdot 10^{-1}$	–
1	$2.45551 \cdot 10^{-1}$	0.77	1.00	53	$3.33849 \cdot 10^{-2}$	3.65
2	$1.63682 \cdot 10^{-1}$	0.59	2.01	110	$6.61105 \cdot 10^{-3}$	2.32
3	$1.11671 \cdot 10^{-1}$	0.55	3.02	164	$1.53205 \cdot 10^{-3}$	2.09

TABLE 2. (Backward-facing-step flow) Comparison of convergence rates for uniform vs. adaptive refinements. The 'level' for the adaptive results is computed as $\log_2(\sqrt{N_i/N_0})$, where N_i denotes the number of degrees of freedom in refinement step i.

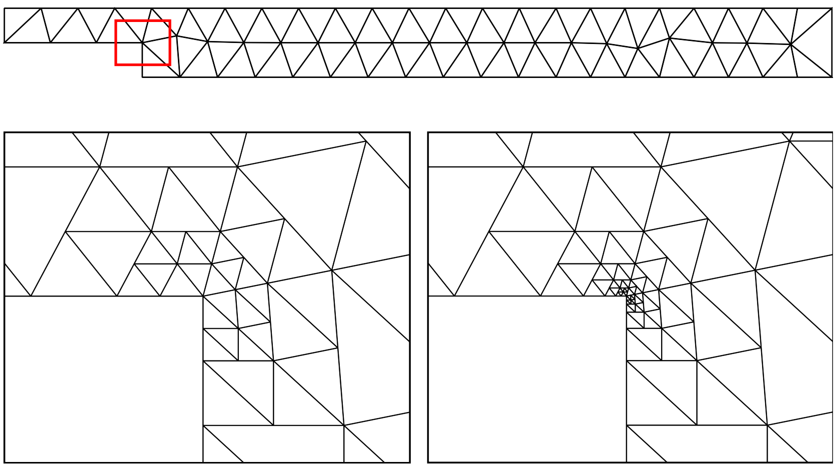

FIGURE 2. Triangulation of the backward facing step domain before (top) and after 16 (bottom left) and 32 (bottom right) non-conforming adaptive refinement steps for a third order approximation.

Oswald-type [52, 40], which allows to obtain conforming approximations for the finite element solution by local post-processing. The sharp estimates for this averaging procedure seem to be new for tetrahedral meshes.

Let us finally mention that by similar techniques as in [34, 55], the proposed method can be extended to the Oseen and Navier-Stokes equations. A detailed analysis for these problems is subject of ongoing research.

Appendix

The goal of the appendix is provide a proof of Lemma 7.7. All results are concerned with the discrete setting, and for convenience, we therefore skip the subscript h referring to discrete functions in other parts of the manuscript. We will proceed as follows:

(i) If $\mathcal{T}_h$ is a conforming simplicial mesh, and V_h is a space of piecewise polynomials on $\mathcal{T}_h$, then for any $u \in V_h$ we construct a continuous, piecewise polynomial approximation $\tilde{u}$ which satisfies the estimates of Lemma 7.7. This construction is a variant of a similar result presented in [40]; see also [15, 36] for *hp*-estimates on hexahedral and two-dimensional meshes.

(ii) More general meshes $\mathcal{T}_h$ can be treated by turning them into a conforming simplicial mesh $\widetilde{\mathcal{T}_h}$ through a refinement step. We require that this refinement does not substantially alter the local shape regularity and meshsize of the elements. On the conforming mesh, we then utilize the construction of (i).

A.1. **Auxiliary results.** Let us start by recalling some preliminary results, that will be needed for the analysis of the averaging process of step (i). Similar arguments are frequently used for the construction of conforming *hp* interpolation operators; see e.g. [48, 25, 45].

We only consider the three dimensional case in detail; corresponding (and even sharper) results for the two-dimensional case can be obtained in a similar manner. Therefore, throughout this section, T will denote a tetrahedron, and f, e, and ν denote faces, edges and vertices. Some of our statments are only given for elements of unit size. The results for general elements then follow by the usual scaling arguments.

The first statement is a slight improvement of [45, Lem. B.2]

Lemma A.1. *Let T be a tetrahedron of unit size. Then for each vertex ν of T, and for each $k \in \mathbb{N}$ there exists a function $\phi_\nu^T \in \mathcal{P}_k(T)$ such that $\phi_\nu^T(\nu) = 1$ and ϕ_ν^T vanishes on the face opposite to ν. Moreover*

$$\|\phi_\nu^T\|_{L^2(T)} \preceq k^{-5/2}, \qquad \|\phi_\nu^T\|_{H^1(T)} \preceq k^{-1/2}, \tag{23}$$

and for any edge e and face f adjacent to ν there holds

$$\|\phi_\nu^T\|_{L^2(e)} \preceq k^{-1}, \qquad \|\phi_\nu^T\|_{L^2(f)} \preceq k^{-2}, \tag{24}$$

Here and below $a \preceq b$ is used to denote $a \leq Cb$ with a constant C that may depend on the shape of T, but which is independent of a, b, the polynomial degree k and the meshsize h.

Proof. Without loss of generality, we may assume that $T = \{(x, y, z) : 0 \leq x, y, z, x+y+z \leq 1\}$ is the reference tetrahedron, and that $f = \{(x, y, 0) : 0 \leq x, y, x + y \leq 1\}$, $e = \{(x, 0, 0) : 0 \leq x \leq 1\}$, and $\nu = (1, 0, 0)$. We define $\phi_\nu^T = (1 + x)J_{k-1}^{2,2}(x)w_{k-1}$ where $J_k^{\alpha,\beta}$ denotes the Jacobi polynomial of order k with weights α and β, cf e.g. [60], and $w_k = [k(k+1)]^{-1} \approx k^{-2}$. It can be verified by elementary calculations that ϕ_ν has the required L^2 estimates. The result in the H^1 norm follows from the L^2 estimate by the inverse inequality (13). □

Remark A.2. Writing ϕ_ν^T in terms of barycentric coordinates, one sees that the values of ϕ_ν^T on the faces f of the element are uniquely determined by the geometry of the face and the polynomial order. Therefore, the element-wise definition allows to extend ϕ_ν to a continuous function on patches of elements.

The following results are taken almost verbatim from [45, App.B]. We present somewhat simplified statements, that suffice for our purposes.

Lemma A.3. *Let T be a tetrahedron of unit size. Then for each edge e of T, there exists a polynomial preserving extension operator $\pi_e^T : H_0^1(e) \to H^1(T)$ such that $\pi_e^T u_e|_e = u_e$ and the extension $\pi_e^T u_e$ vanishes on the faces $f \subset \partial T$ with $e \not\subset \partial f$. Moreover, $\pi_e^T u_k \in P_k(T)$ for all $u_k \in P_k(T) \cap H_0^1(e)$, and there holds*

$$\|\pi_e^T u_k\|_{H^1(T)} \preceq k^{1/2}\|u_k\|_{L^2(e)} \qquad \text{and} \qquad \|\pi_e^T u_k\|_{L^2(f)} \preceq \|u_k\|_{L^2(e)}. \tag{25}$$

For each face $f \subset \partial T$ there exists an operator $\pi_f^T : H_0^1(f) \to H^1(T)$ such that $\pi_f^T u_f|_f = u_f$ and $\pi_f^T u_f = 0$ on $\partial T \setminus f$. Moreover, for $u_k \in P_k(f) \cap H_0^1(f)$ there holds $\pi_f^T u_k \in P_k(T)$ and

$$\|\pi_f^T u_f\|_{H^1(T)} \preceq k\|u_f\|_{L^2(f)}. \tag{26}$$

Remark A.4. The extensions π_e^T and π_f^T can be used to define global extension operators π_e and π_f, which extend functions from an edge or a face into a patch of elements adjacent to e or f, respectively. Here, one utilizes, that the construction of π_e^T is done in two steps: first, the function u_e is extended from the edge e to adjacent faces f in a way that only depends on the geometry of the corresponding face f. In a second step, the face functions are extended into the interior of the domains. For details, see [48, 25] and [45, App.B].

Let us also recall the following well-known inequalities

$$\|u_e\|_{L^\infty(e)} \preceq k\|u_e\|_{L^2(e)} \quad \text{and} \quad \|u_f\|_{L^\infty(f)} \leq k^2\|u_f\|_{L^2(f)}, \tag{27}$$

which hold for polynomials $u_e \in P_k(e)$ respectively $u_f \in P_k(f)$ and follow readily from interpolation arguments and inverse inequalities [58, 32].

A.2. **The averaging operator on conforming meshes.** We now turn to the construction and the analysis of an averaging operator, which maps discontinuous piecewise polynomials defined on a conforming simplicial mesh to continuous piecewise polynomials over the same mesh. A similar operator has been considered in [36] for two dimensions; see also [40, 15] for related constructions.

A.2.1. *Preliminaries.* Let $\mathcal{T}_h$ denote a conforming γ-shape-regular tetrahedral mesh, and define a space of discontinuous piecewise polynomial functions

$$V_h := \{v_h \in L^2(\Omega) : v_h|_T \in P_{k_T}(T) \text{ for all } T \in \mathcal{T}_h\}.$$

For functions $u \in V_h$, we define on each face f in $\mathcal{T}_h$ the jump over the face by [5]

$$[u] = \begin{cases} u_T n_T + u_{T'} n_{T'} & \text{if } \overline{f} = \partial T \cap \partial T', \\ u_T n_T & \text{if } f \subset \partial\Omega. \end{cases}$$

Here, n_T is the unit normal vector pointing to the exterior of the element T, and $u_T = u|_T$ denotes the restriction of u to the element T. Note that the jump of a function $[u]$ is always defined with respect to a face f. To each face f of the mesh, we associate a polynomial degree and a meshsize parameter by

$$k_f := \max\{k_T : f \subset \partial T\} \qquad \text{and} \qquad h_f := \min\{h_T : f \subset \partial T\}.$$

Moreover, we define for each geometric entity of the mesh another polynomial degree by

$$\tilde{k}_\nu := \max\{k_T : \nu \in \overline{T}\} \quad \text{and} \quad \tilde{k}_g := \max\{\tilde{k}_\nu : \nu \in \overline{g}\}, \quad g \in \{e, f, T\},$$

which will be used for the averaging process. Note that due to the local-uniformity of the polynomial degree distribution (9), we have $\tilde{k}_T \leq \eta\, k_T$ for all elements $T \in \mathcal{T}_h$.

A.2.2. *The basic construction.* Any polynomial $u_T \in \mathcal{P}_{k_T}(T)$ can be decomposed as

$$u_T = u_T^n + u_T^e + u_T^f + u_T^i$$

into a nodal, edge, face, and inner part, which are defined, respectively, by

$$u_T^n := \sum\nolimits_{\nu} u_T(\nu)\phi_\nu^T, \qquad u_T^e := \sum\nolimits_e \pi_e^T(u_T - u_T^n)|_e,$$
$$u_T^f := \sum\nolimits_f \pi_f^T(u_T - u_T^n - u_T^e)|_f, \qquad u_T^i := u_T - u_T^n - u_T^e - u_T^f.$$

The function ϕ_ν^T always carries the polynomial degree $\tilde{k}_\nu$ of the vertex ν. As a next step, we define averages for the nodal, edge and face parts as follows:

$$\bar{u}_\nu := \tfrac{1}{|\mathcal{T}(\nu)|}\sum\nolimits_{T\in\mathcal{T}(\nu)} u_T(\nu), \quad \bar{u}_e := \tfrac{1}{|\mathcal{T}(e)|}\sum\nolimits_{T\in\mathcal{T}(e)} (u_T - u_T^n)\big|_e,$$
$$\text{and} \qquad \bar{u}_f := \tfrac{1}{|\mathcal{T}(f)|}\sum\nolimits_{T\in\mathcal{T}(f)} (u_T - u_T^n - u_T^e)\big|_f.$$

Here, $\mathcal{T}(g) = \{T : g \cap \overline{T} \neq \emptyset\}$, $g \in \{\nu, e, f\}$ denotes the set of elements adjacent to the entity g, and $|\mathcal{T}(g)|$ is the cardinality of the set $\mathcal{T}(g)$. Using these averages, we can construct a continuous approximation $\tilde{u}$ for $u \in V_h$ element-wise by

$$\tilde{u}_T = \tilde{u}_T^n + \tilde{u}_T^e + \tilde{u}_T^f + \tilde{u}_T^i$$

with nodal, edge, face, and inner parts defined by

$$\tilde{u}_T^n := \sum\nolimits_\nu \bar{u}_\nu \phi_\nu^T, \qquad \tilde{u}_T^e := \sum\nolimits_e \pi_e^T \bar{u}_e, \qquad \tilde{u}_T^f := \sum\nolimits_f \pi_f^T \bar{u}_f, \qquad \tilde{u}_T^i := u_T^i.$$

Here, the summation is done only over entities $\nu, e, f \subset \partial T$ that do not lie on the boundary $\partial\Omega$. As a direct consequence of the construction, we then obtain that $\tilde{u}$ lies in the space

$$\widetilde{V_h} := \{\boldsymbol{v}_h \in H_0^1(\Omega) : v_h|_T \in P_{\tilde{k}_T}(T) \text{ for all } T \in \mathcal{T}_h\},$$

were we have used that the element-wise definitions of ϕ_ν^T and the extension operators π_e^T, π_f^T can be extended in a continuous way to pathces of elements; see Remark A.2 and A.4.

A.2.3. *Estimates for the averaging operator.* The following estimates follow by simple geometric arguments; see [40, Sec. 2] for similar results.

Lemma A.5. *Let $u \in V_h$ be given, and denote by*

$$\mathcal{F}(g) := \{f : g \subset \overline{f}\}, \quad g \in \{\nu, e\} \qquad \textit{and} \qquad \mathcal{F}(T) := \{f : \overline{f} \cap \overline{T} \neq \emptyset\},$$

the set of faces that contain g or are adjacent to T.

(i) For any element $T \in \mathcal{T}_h$, there holds

$$\sum\nolimits_{\nu\in\overline{T}} |u_T^n(\nu) - \tilde{u}_T^n(\nu)| \preceq \sum\nolimits_{f\in\mathcal{F}(\nu)} |[u(\nu)]|$$

(ii) For any edge e of the mesh, there holds

$$\|u_T^e - \tilde{u}_T^e\|_{L^2(e)} \preceq \sum\nolimits_{f\in\mathcal{F}(e)} \|[u^e]\|_{L^2(e)}.$$

(iii) For any face f of the mesh, there holds

$$\|u_T^f - \tilde{u}_T^f\|_{L^2(f)} \preceq \|[u^f]\|_{L^2(f)}.$$

Proof. By definition of $\tilde{u}$ and the average $\bar{u}_\nu$, we have

$$|u_T^n(\nu) - \tilde{u}_T^n(\nu)| \leq \tfrac{1}{|\mathcal{T}(\nu)|} \sum\nolimits_{T' \in \mathcal{T}(\nu)} |u_T(\nu) - u_{T'}(\nu)|.$$

By walking through the elements around the node ν, we obtain similar as in [40, Sec. 2]

$$|u_T(\nu) - u_{T'}(\nu)| \leq \sum\nolimits_{f \in \mathcal{F}(\nu)} |[u](\nu)|,$$

which yields the first result. The other statements follow in a similar way. □

We are now in the position to prove the following approximation result.

Proposition A.6. *Let $u \in V_h$ be given and $\tilde{u} \in \widetilde{V_h}$ be defined as above. Then*

$$\|\nabla(u - \tilde{u})\|_{0,\mathcal{T}_h} \leq C_{av} \Big(\sum\nolimits_f k_f^4 h_f^{-1} |[u_h]|^2_{L^2(f)} \Big)^{1/2},$$

with a constant C_{av} independent of the meshsize and the polynomial degree distributions.

Proof. We use the triangle inequality twice to estimate

$$\|\nabla(u - \tilde{u})\|_{0,\mathcal{T}_h} \leq \sum\nolimits_T \|u_T^n - \tilde{u}_T^n\|_{H^1(T)} + \|u_T^e - \tilde{u}_T^e\|_{H^1(T)} + \|u_T^f - \tilde{u}_T^f\|_{H^1(T)}.$$

The inner terms vanish, since $\tilde{u}_T^i = u^i$ by definition. It remains to estimate the nodal, edge and face terms on the individual elements. For ease of notation, we skip the subindex T in the following, and assume that the elements are of unit size. The dependence on the meshsize is obtained by the usual scaling argument. Since the polynomial degree is locally uniform, it suffices to use one symbol k for all local estimates.

Step 1: Let us first estimate the nodal parts. By Lemma A.5(i) and Lemma A.1, there holds

$$\begin{aligned}\|u^n - \tilde{u}^n\|_{H^1(T)} &\leq \sum\nolimits_{\nu \in \overline{T}} |u(\nu) - \bar{u}_\nu| \|\phi_\nu\|_{H^1(T)} \preceq \sum\nolimits_{f \in \mathcal{F}(T)} \|[u]\|_{L^\infty(f)} k^{-1/2} \\ &\preceq \sum\nolimits_{f \in \mathcal{F}(T)} k^{3/2} \|[u]\|_{L^2(f)},\end{aligned}$$

where we used (27) for the last inequality.

Step 2: For the edge parts, we obtain by Lemma A.3 and Lemma A.5(ii), that

$$\|u^e - \tilde{u}^e\|_{H^1(T)} \preceq \sum\nolimits_{e \subset \partial T} k^{1/2} \|u^e - \tilde{u}^e\|_{L^2(e)} \preceq \sum\nolimits_{e \subset \partial T} \sum\nolimits_{f \in \mathcal{F}(e)} k^{1/2} \|[u^e]\|_{L^2(e)}.$$

Recall that the jump $[u^e]$ is defined with respect to a face f. Now, by definition of u^e, there holds $[u^e]|_e = [u - u^n]|_e = [u]|_e - [u^n]|_e$ on each face f with $e \subset \partial f$. This yields

$$\begin{aligned}k^{1/2} \|[u^e]\|_{L^2(e)} &\leq k^{1/2} \|[u]\|_{L^2(e)} + k^{1/2} \|[u^n]\|_{L^2(e)} \\ &\preceq k^{3/2} \|[u]\|_{L^2(f)} + k^{1/2} k^2 \|[u]\|_{L^2(f)} k^{-1} \preceq k^{3/2} \|[u]\|_{L^2(f)}.\end{aligned}$$

where we used a discrete trace inequality (13), and an estimate of the nodal part similar as in Step 1. Summing up, we obtain

$$\|u^e - \tilde{u}^e\|_{H^1(T)} \preceq \sum\nolimits_{e \subset \partial T} \sum\nolimits_{f \in \mathcal{F}(e)} k^{3/2} \|[u]\|_{L^2(f)} \preceq \sum\nolimits_{f \in \mathcal{F}(T)} k^{3/2} \|[u]\|_{L^2(f)}.$$

Step 3: By Lemma A.3 and Lemma A.5(iii), we obtain for the face parts

$$\|u^f - \tilde{u}^f\|_{H^1(T)} \preceq \sum\nolimits_{f\in\mathcal{F}(T)} k\|u^f - \tilde{u}^f\|_{L^2(f)} \preceq \sum\nolimits_{f\in\mathcal{F}(T)} k\|[u^f]\|_{L^2(f)}.$$

The face terms can be split into $[u^f] = [u] - [u^n] - [u^e]$, which yields

$$k\|[u^f]\|_{L^2(f)} \leq k\|[u]\|_{L^2(f)} + k\|[u^n]\|_{L^2(f)} + k\|[u^e]\|_{L^2(f)}.$$

In a similar manner to Step 1, we can bound the nodal part by

$$k\|[u^n]\|_{L^2(f)} \preceq k\|[u]\|_{L^\infty(f)}k^{-2} \leq k\|[u]\|_{L^2(f)},$$

where we used Lemma A.1 to estimate $\|\phi_\nu\|_{L^2(f)}$ by k^{-2}. For the edge part, we apply the second estimate of Lemma A.3, to get

$$k\|[u^e]\|_{L^2(f)} \preceq \sum\nolimits_{e'\subset\overline{f}} k\|[u^e]\|_{L^2(e')},$$

which can be further estimated as in Step 2. Combination of the estimates for the individual terms then yields

$$\|u^f - \tilde{u}^f\|_{H^1(T)} \preceq \sum\nolimits_{f\in\mathcal{F}(T)} k^2\|[u]\|_{L^2(f)}.$$

The proof is completed by scaling, summation over all elements, and using the finite overlap of patches $\mathcal{F}(T)$. □

Remark A.7. For two dimensional triangulations, the estimate of Proposition A.6 can be sharpened, i.e., there holds

$$\|\nabla(u-\tilde{u})\|_{0,\mathcal{T}_h} \leq C_{av}\Big(\sum\nolimits_e k^2h^{-1}\|[u]\|^2_{L^2(e)}\Big)^{1/2}.$$

with a constant C_{av} independent of u, the polynomial degree distribution, and the meshsize. The proof of this result follows in the same way as that of Proposition A.6. This result has been announces in [36] for two dimensional hybrid meshes, and estimates with the same dependence on the polynomial degree have been obtained for the Oswald interpolation operator on quadrilateral and hexahedral meshes; cf. [15, 66].

A.3. **Refinement of irregular meshes.** We now turn to generalization of the results of the previous section to more general meshes. Let us introduce some notation to characterize the appropriate cases.

Definition A.8 (regular simplicial refinement)**.** *A conforming simplicial partition $T_h := \{T_i\}$ of an element T into simplices $T_i \subset T$ with $h_T \leq \kappa h_{T_i}$ and $\gamma_T \leq \kappa\gamma_{T_i}$ for all $T_i \in T_h$ is called* κ-regular simplicial refinement *of T.*

Remark A.9. Let us mention some examples: Every element is the 1-regular refinement of itself. A uniform refinement of a triangle or quadrilateral into 4 elements is 1/2-regular. Similarly, a division of a quadrilateral into two triangles is κ-regular with some $\kappa < 1$ depending on the shape of the quadrilateral.

Definition A.10. *The mesh $\mathcal{T}_h$ is called* κ-regular-closable*, if it can be turned into a conforming simplicial mesh $\overline{\mathcal{T}_h}$ (the* κ-regular closure *of $\mathcal{T}_h$) by κ-regular simplicial refinements of its elements.*

Remark A.11. The typical 1-irregular meshes arising in discontinuous Galerkin methods are κ-regular-closable. In a first step, every quadrilateral (or hexahedral) element is decomposed into simplices, and in a second step, the hanging nodes of the simplicial mesh are removed; see Figure 3 for some illustration. Note that the elimination is possible in a finite number of steps and without introducing new nodes.

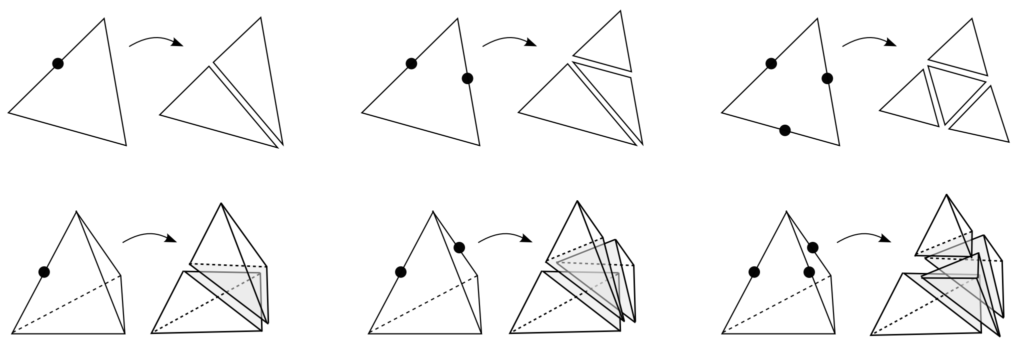

FIGURE 3. Elimination of hanging nodes in a triangle and a tetrahedron.

If $\overline{\mathcal{T}_h}$ is the κ-regular-closure of $\mathcal{T}_h$, then

$$\gamma_{\overline{T}} \geq \kappa\, \gamma_T \qquad \text{and} \qquad h_{\overline{T}} \geq \kappa h_T$$

for all $T \in \mathcal{T}_h$ and $\overline{T} \in T_h$. Hence, we can identify the size and shape parameter of refined elements $\overline{T}$ with those of the parent element T. This will be used explicitly for the final result.

A.4. **Proof of Lemma 7.7.**

Step 1: We start by turning the mesh $\mathcal{T}_h$ into a conforming simplicial mesh $\overline{\mathcal{T}_h}$, and for each element $\overline{T} \in T_h$, $T \in \mathcal{T}_h$, we set $k_{\overline{T}} := k_T$. We further define a space

$$\overline{V_h} := \{v \in L^2(\Omega) : v|_{\overline{T}} \in P_{k_{\overline{T}}} \quad \text{for all } \overline{T} \in \mathcal{T}_h\}.$$

Note that by definition, $V_h \subset \overline{V_h}$, such that every function $u_h \in V_h$ can be identified with a function in $\overline{u}_h \in \overline{V_h}$. Moreover,

$$\sum\nolimits_{\overline{f}} k_{\overline{f}}^4 h_{\overline{f}}^{-1} \|[\overline{u}_h]\|^2_{L^2(\overline{f})} \leq c_\kappa \sum\nolimits_f k_f^4 h_f^{-1} \|[u_h]\|^2_{L^2(f)}, \tag{28}$$

where summation is done over all faces $\overline{f}$ of the refined mesh $\overline{\mathcal{T}_h}$ or the faces f of the coarse mesh $\mathcal{T}_h$, respectively. This holds, because $[u_h]$ vanishes on faces $\overline{f} \not\subset \mathcal{E}$, and since $\sum_{\overline{f}\subset f} k_{\overline{f}}^4 h_{\overline{f}}^{-1} \|[\overline{u}_h]\|^2_{L^2(\overline{f})} \approx k_f^4 h_f^{-1} \|[u_h]\|^2_{L^2(f)}$ due to the κ-regularity of the refinement and the definition of the local-uniformity of the polynomial degree distribution.

Step 2: Using (28) and Proposition A.6 with V_h and u replaced by $\overline{V_h}$ and u_h, we obtain

$$\|\nabla(u_h - \tilde{u})\|_{0,\mathcal{T}_h} = \|\nabla(u_h - \tilde{u})\|_{0,\overline{\mathcal{T}}_h} \preceq \sum\nolimits_{\overline{f}} k_{\overline{f}}^4 h_{\overline{f}}^{-1} \|[\overline{u}]\|^2_{L^2(\overline{f})} \preceq \sum\nolimits_f k_f^4 h_f^{-1} \|[u]\|^2_{L^2(f)}, \tag{29}$$

which is the generalization of Proposition A.6 to κ-regular-closable meshes.

Step 3: We can split the interface terms by the triangle inequality into

$$\|[u]\|_{L^2(f)} = \|u_T - u_{T'}\|_{L^2(f)} \leq \|u_T - \hat{u}\|_{L^2(f\cap\partial T)} + \|u_{T'} - \hat{u}\|_{L^2(f\cap\partial T')},$$

where $\hat{u} \in L^2(f)$ is an arbitrary function. A similar arguments holds for boundary faces, where we require that $\hat{u} = 0$ on $\partial\Omega$. By the local quasi-uniformity (8) of the polynomial degree distribution, and the shape regularity of the mesh, we further have

$$k_f \sim k_T \quad \text{and} \quad h_f \sim h_T \qquad \text{for all } T \in \mathcal{T}_h \text{ and } f \subset \partial T,$$

and we can replace h_f and k_f by h_T and k_T. Lemma 7.7 then follows by summation of all element contributions.

Acknowledgments

Financial support by the German Research Association (DFG) through grant GSC 111 is gratefully acknowledged. Parts of the manuscript were written during the stay of the first author at the University of Graz and the Chemnitz University of Technology.

References

[1] M. Ainsworth, *Robust a posteriori error estimation for nonconforming finite element approximation*, SIAM J. Numer. Anal., 42 (2005), pp. 2320–2341.

[2] M. Ainsworth and J. T. Oden, *A posteriori error estimation in finite element analysis*, Teubner, Jan 2000.

[3] D. N. Arnold, *An interior penalty finite element method with discontinuous elements*, SIAM J. Numer. Anal., 19 (1982), pp. 742–760.

[4] D. N. Arnold and F. Brezzi, *Mixed and nonconforming finite element methods: Implementation, postprocessing and error estimates*, Math. Model. Numer. Anal., 19 (1985), pp. 7–32.

[5] D. N. Arnold, F. Brezzi, B. Cockburn, and D. Marini, *Unified analysis of discontinuous Galerkin methods for elliptic problems*, SIAM J. Numer. Anal., 39 (2002), pp. 1749–1779.

[6] I. Babuška, *Error-bounds for finite element method*, Numer. Math., 16 (1971), pp. 322–333.

[7] I. Babuška and M. Suri, *The hp version of the finite element method with quasiuniform meshes*, M2AN, 21 (1987), pp. 199–238.

[8] C. Baiocchi, F. Brezzi, and L. D. Marini, *Stabilization of Galerkin Methods and Applications to Domain Decomposition*, Lecture Notes in Computer Science, Springer, 1992, pp. 343–355.

[9] F. Bassi and S. Rebay, *A high-order accurate discontinuous finite element method for the numerical solution of the compressible Navier-Stokes equations*, J. Comput. Phys., 131 (1997), pp. 267–279.

[10] P. Bastian, M. Blatt, A. Dedner, C. Engwer, R. Klöfkorn, M. Ohlberger, and O. Sander, *A generic grid interface for parallel and adaptive scientific computing. part I: Abstract framework*, Computing, 82 (2008), pp. 103–119.

[11] R. Becker, P. Hansbo, and M. G. Larson, *Energy norm a posteriori error estimation for discontinuous Galerkin methods*, Comput. Meth. Appl. Mech. Engrg., 192 (2003), pp. 723–733.

[12] C. Bernardi and Y. Maday, *Spectral methods*, vol. 5, North-Holland, Amsterdam, 1997, pp. 209–485.

[13] D. Braess, *Finite Elements. Theory, Fast Solvers and Applications in Solid Mechanics*, Cambridge University Press, 3rd ed., 2007.

[14] F. Brezzi and M. Fortin, *Mixed and hybrid finite element methods*, Springer-Verlag, New York, 1991.

[15] E. Burman and A. Ern, *Continuous interior penalty hp-finite element methods for advection and advection-diffusion equations*, Math. Comp., 76 (2007), p. 1119.

[16] R. Bustinza, G. N. Gatica, and B. Cockburn, *An a posteriori error estimate for the local discontinuous Galerkin method applied to linear and nonlinear diffusion problems*, J. Sci. Comput., 22–23 (2005), pp. 147–185.

[17] C. Carstensen and J. Hu, *A unifying theory of a posteriori error control for nonconforming finite element methods*, Numer. Math., 107 (2007), pp. 473–502.

[18] A. Chernov, *Optimal convergence estimates for the trace of the polynomial L^2-projection operator on a simplex*, INS Preprint, (2010), pp. 1–26.

[19] B. Cockburn, J. Gopalakrishnan, and R. Lazarov, *Unified hybridization of discontinuous Galerkin, mixed and continuous Galerkin methods for second order elliptic problems*, SIAM J. Numer. Anal., 47 (2009), pp. 1319–1365.

[20] B. Cockburn, J. Gopalakrishnan, N. C. Nguyen, J. Peraire, and F.-J. Sayas, *Analysis of HDG methods for Stokes flow*, Math. Comp., 80 (2011), pp. 723–760.

[21] B. Cockburn, G. Kanschat, and D. Schötzau, *Local discontinuous Galerkin methods for the Stokes system*, Siam J. Numer. Anal., (2003).

[22] B. Cockburn, G. E. Karniadakis, and C.-W. Shu, eds., *Discontinuous Galerkin Methods: Theory, Computation and Applications*, Berlin Heidelberg, 2000, Springer-Verlag.

[23] B. Cockburn, N. Nguyen, and J. Peraire, *A comparison of HDG methods for Stokes flow*, J. Sci. Comput., (2010), pp. 1–23.

[24] B. Cockburn and C.-W. Shu, *The Runge-Kutta discontinuous Galerkin method for conservation laws V: Multidimensional systems*, Comput. Phys., 141 (1998), pp. 199–224.

[25] L. Demkowicz, J. Gopalakrishnan, and J. Schöberl, *Polynomial extension operators. i*, SIAM J. Numer. Anal., 46 (2008), pp. 3006–3031.

[26] H. Egger and J. Schöberl, *A mixed-hybrid-discontinuous Galerkin finite element method for convection-diffusion problems*, IMA J. Numer. Anal., (2009).

[27] H. Egger and C. Waluga, *A hybrid mortar method for incompressible flow*, Tech. Rep. AICES-2011-04-01, RWTH Aachen, 2011.

[28] H. C. Elman, D. J. Silvester, and A. J. Wathern, *Finite Elements and Fast Iterative Solvers: with Applications in Incompressible Fluid Dynamics*, Oxford University Press, New York, 2005.

[29] A. Ern and J.-L. Guermond, *Theory and Practice of Finite Elements*, vol. 159 of Applied Mathematical Sciences, Springer, Ney York, 2004.

[30] L. Filippini and A. Toselli, *hp finite element approximations on non-matching grids for the Stokes problem*, Tech. Rep. 2002-22, ETH Zürich, 2002.

[31] M. Fortin, *An analysis of the convergence of mixed finite element methods*, RAIRO, 11 (1977), pp. 341–354.

[32] E. Georgoulis, *Inverse-type estimates on hp-finite element spaces and applications*, Math. Comp., 77 (2008), pp. 201–219.

[33] V. Girault and P. Raviart, *Finite Element Approximation of the Navier-Stokes Equations: Theory and Algorithms*, Springer Series in Computational Mathematics, Springer, 1986.

[34] V. Girault, B. Rivière, and M. F. Wheeler, *A discontinuous Galerkin method with non-overlapping domain decomposition for the Stokes and Navier-Stokes problems*, Math. Comp., 74 (2005), pp. 53–84.

[35] P. Grisvard, *Elliptic problems in nonsmooth domains*, no. 24 in Pitman Monographs and Studies in Mathematics, Pitman, Boston, MA, 1985.

[36] P. Houston, D. Schötzau, and T. P. Wihler, *Energy norm a posteriori error estimation of hp-adaptive discontinuous Galerkin methods for elliptic problems*, M3AS, 17 (2007), pp. 33–62.

[37] P. Houston, C. Schwab, and E. Süli, *Discontinuous hp-finite element methods for advection-diffusion-reaction problems*, Siam J. Numer. Anal., (2002).

[38] J. Jaffré, C. Johnson, and A. Szepessy, *Convergence of the discontinuous Galerkin finite element method for hyperbolic conservation laws*, Math. Models Methods Appl. Sci., 5 (1995), pp. 367–386.

[39] C. Johnson and J. Pitkäranta, *An analysis of the discontinuous Galerkin method for a scalar hyperbolic equation*, Math. Comput., 46 (1986), pp. 1–26.

[40] O. A. Karakashian and F. Pascal, *A posteriori error estimates for a discontinuous Galerkin approximation of second-order elliptic problems*, SIAM J. Numer. Anal., 41 (2003), pp. 2374–2399.

[41] K. Y. Kim, *A posteriori error estimators for locally conservative methods of nonlinear elliptic problems*, Applied numerical mathematics, 57 (2007), pp. 1065–1080.

[42] O. Ladyshenskaya, *The mathematical theory of viscous incompressible flow*, Gordon & Breach, 1969.

[43] C. Lehrenfeld, *Hybrid discontinuous Galerkin methods for incompressible flow problems*, Master's thesis, RWTH Aachen, 2010.

[44] P. LeSaint and P. A. Raviart, *On a finite element method for solving the neutron transport equation*, in Mathematical Aspects of Finite Elements in Partial Differential Equations, C. de Boor, ed., New York, 1974, Academic Press, pp. 89–123.

[45] J. M. Melenk and S. Sauter, *Convergence analysis for finite element discretizations of the Helmholtz equation with Dirichlet-to-Neumann boundary conditions*, Math. Comp., 79 (2010), pp. 1871–1914.

[46] J. M. Melenk and B. Wohlmuth, *On residual-based aposteriori error estimation in hp-FEM*, Adv. Comput. Math., 15 (2001), pp. 311–331.
[47] W. F. Mitchell and M. A. McClain, *A survey of hp-adaptive strategies for elliptic partial differential equations*, to appear in Annals of the European Academy of Sciences, (2011).
[48] R. Munoz-Sola, *Polynomial liftings on a tetrahedron and applications to the hp-version of the finite element method in three dimensions*, SIAM J. Numer. Anal., 34 (1997), pp. 282–314.
[49] J. Nečas, *Les méthodes directes en théorie des équations elliptiques*, Masson, Paris, 1967.
[50] N. C. Nguyen, J. Peraire, and B. Cockburn, *A hybridizable discontinuous Galerkin method for Stokes flow*, Comput. Methods Appl. Mech. Engrg., 199 (2010), pp. 582–597.
[51] R. H. Nochetto, K. G. Siebert, and A. Veeser, *Theory of adaptive finite element methods: An introduction*, in Multiscale, Nonlinear and Adaptive Approximation, R. A. DeVore and A. Kunoth, eds., Springer, 2009, pp. 409–542.
[52] P. Oswald, *On a BPX-preconditioner for P1 elements*, Computing, (1993).
[53] I. Perugia and D. Schötzau, *An hp-analysis of the local discontinuous Galerkin method for diffusion problems*, J. Sci. Comput., 17 (2002), pp. 561–571.
[54] W. H. Reed and T. R. Hill, *Triangular mesh methods for the neutron transport equation*, Tech. Rep. LA-UR-73-479, Los Alamos Scientific Laboratory, Los Alamos NM, 1973.
[55] B. Riviére, *Discontinuous Galerkin Methods for Solving Elliptic and Parabolic Equations: Theory and Implementation*, vol. 35 of Frontiers in Mathematics, SIAM, 2008.
[56] B. Rivière and M. F. Wheeler, *A posteriori error estimates for a discontinuous Galerkin method applied to elliptic problems*, Computers and Mathematics with Applications, 46 (2003), pp. 141–163.
[57] D. Schötzau, C. Schwab, and A. Toselli, *Mixed hp-DGFEM for incompressible flows*, SIAM J. Numer. Anal., (2003), pp. 2171–2194.
[58] C. Schwab, *p- and hp- finite element methods: theory and applications in solid and fluid mechanics*, Oxford University Press, Jan 1998.
[59] R. Stenberg and M. Suri, *Mixed hp-finite element methods for problems in elasticity and Stokes flow*, Numer. Math., 72 (1996), pp. 367–389.
[60] G. Szegö, *Orthogonal Polynomials*, vol. 23 of American Mathematical Society Colloquium Publications, American Mathematical Society, New York, 1939.
[61] A. Toselli, *hp discontinuous Galerkin approximations for the Stokes problem*, Mathematical Models and Methods in Applied Sciences, 12 (2002), pp. 1565–1598.
[62] R. Verfürth, *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*, Teubner, Apr 1996.
[63] T. Warburton and J. S. Hesthaven, *On the constants in hp-finite element trace inverse inequalities*, Comp. Meth. Appl. Mech. Engrg., 192 (2003), pp. 2765–2773.
[64] B. Wohlmuth, *A residual based error estimator for mortar finite element discretizations*, Numer. Math., 84 (1999), pp. 143–171.
[65] J. Xu and L. Zikatanov, *Some observations on Babuška and Brezzi theories*, Numer. Math., 94 (2003), pp. 195–202.
[66] L. Zhu, S. Giani, P. Houston, and D. Schötzau, *Energy norm a-posteriori error estimation for hp-adaptive discontinuous Galerkin methods for elliptic problems in three dimensions*, Mathematical Models and Methods in Applied Sciences, (2009).