STABILIZED DISCONTINUOUS FINITE ELEMENT APPROXIMATIONS FOR STOKES EQUATIONS

RAYTCHO LAZAROV AND XIU YE

Abstract. In this paper, we derive two stabilized discontinuous finite element formulations, symmetric and nonsymmetric, for the Stokes equations and the equations of the linear elasticity for almost incompressible materials. These methods are derived via stabilization of a saddle point system where the continuity of the normal and tangential components of the velocity/displacements are imposed in a weak sense via Lagrange multipliers. For both methods, almost all reasonable pair of discontinuous finite elements spaces can be used to approximate the velocity and the pressure. Optimal error estimate for the approximation of both the velocity of the symmetric formulation and pressure in L^2 norm are obtained, as well as one in a mesh dependent norm for the velocity in both symmetric and nonsymmetric formulations.

1. Introduction

We consider the Stokes system

$$-\Delta \mathbf{u} + \nabla p = \mathbf{f} \quad \text{in } \Omega, \tag{1.1}$$

$$\nabla \cdot \mathbf{u} + \gamma p = 0 \quad \text{in } \Omega, \tag{1.2}$$

$$\mathbf{u} = 0 \quad \text{on } \partial\Omega, \tag{1.3}$$

where $\mathbf{u} = (u_1, \ldots, u_d)$ is a vector function, Ω is a bounded, open subset in $\mathbf{R}^d$ ($d = 2$ or 3) with Lipschitz boundary $\partial\Omega$. The symbols Δ, ∇, and $\nabla\cdot$ denote the Laplacian, gradient, and divergence operators respectively, and $\mathbf{f}(x)$ is the properly scaled external volumetric force. For $\gamma = 0$ we get the Stokes system (of steady flow of very viscous fluid) for the velocity $\mathbf{u}$ and the pressure p that is rescaled by the viscosity. For $\gamma = 1 - 2\nu$, where $\nu \in (0, \frac{1}{2}]$ is the Poisson's ratio, we get the equations of the linear elasticity (constant coefficients case) for the displacement $\mathbf{u}$ and the pressure p including the incompressible limit $\nu = \frac{1}{2}$.

These two problems are quite similar in regard to their stability. Namely, one can prove the following a priori estimate for the solution of (1.1) – (1.3)

$$\|\mathbf{u}\|_{H^1} + \|p\|_{L^2} \leq C\|\mathbf{f}\|_{H^{-1}} \tag{1.4}$$

with a constant C independent of $\gamma \geq 0$. The above stability relies on the fundamental *inf-sup* condition of Ladyzhenskaya, Babuška, and Brezzi, see e.g. [7], [19], [29].

In order to get stable finite element approximation of this problem we need to have similar property for the finite element spaces for $\mathbf{u}$ and p, correspondingly. Namely, we

Date: April 6, 2005.

1991 *Mathematics Subject Classification.* primary 65N15, 65N30, 76D07; secondary 35B45, 35J50.

Key words and phrases. finite element method, discontinuous Galerkin method, Stokes problem, linear elasticity, almost incompressible material.

need a stable pair of finite element spaces, i.e. that satisfies the *inf-sup* condition. The finite element analysis has produced a large number of stable pair of spaces. However, various simple elements, for example, spaces using P_1 - P_0 polynomials on triangles and Q_1 - P_0 polynomials on quadrilaterals are not stable.

To satisfy the *inf-sup* condition easier (even trivially) various stabilization techniques have been proposed and studied. Perhaps among the very first stabilized finite element formulation for the Stokes equations is the scheme proposed by Hughes, Franca and Balestra [26], which allows use of simpler and more natural finite element spaces. Improvements of the stabilized finite element formulation in [26] were made further by Hughes and Franca [27]. Douglas and Wang [18] proposed an absolutely stabilized finite element formulation for the Stokes problem in which solvability and convergence of the method do not depend on the stability constant. Kechkar and Silvester [28] introduced a local stabilization method in which the jump term of pressure on the All these stabilized finite element methods are designed for conforming elements for the velocity, i.e. velocity is approximated by piecewise polynomial functions in $H^1(\Omega)$.

Discontinuous Galerkin method, though more expensive, allows natural adaptive procedures, constructions that are more flexible, and produce stable approximations. These are some of the reasons that have made the Discontinuous Galerkin method an active research area in recent years (see, e.g. [1, 3, 4, 5, 12, 11, 15, 13, 17, 23, 24]). Relaxing the continuity of approximate functions across element boundary (required in standard finite element methods) gives the discontinuous Galerkin method more localization and flexibility which lead to easier and more natural $h-p$ mesh adaptiation. Recently, Cockburn, Kanschat, Schotzau and Schwab [11] studied a local discontinuous Galerkin method for the Stokes and Navier-Stokes systems in mixed form. They showed that the local discontinuous Galerkin methods in this case can easily handle meshes with hanging nodes, elements of general shapes, local spaces of different types and weakly enforce the conservation of mass element by element. Hansbo and Larson [24] introduced stabilized discontinuous Galerkin method for the equations of the linear elasticity in the incompressible and nearly incompressible case without using pressure variable and proved optimal rate of convergence in certain mesh dependent norm.

The drawback of all discontinuous Galerkin approximations is a substantial increase of the number of degrees of freedom, which leads to a much larger algebraic systems. Attempts to reduce the number of the degrees of freedom has led to constructions using Crouzeix-Raviart nonconforming finite elements (see, e.g. [23] and [21]) or spaces having continuous normal component but discontinuous tangential component across the finite element boundaries (see, e.g. [31]).

In this paper, following the point of view of Douglas and Wang [18] we derive and study discontinuous Galerkin approximations of the Stokes equations and the equations of the linear elasticity in the incompressible limit as stabilization schemes of a certain saddle point problem. First we introduce as new variables the traces on the interfaces of the tangential component of the velocity and the derivative of the normal component of the velocity in normal direction to the finite element faces. To enforce the continuity of the velocity along the interfaces we use Lagrange multipliers. This results to a new saddle point problem which is approximated by finite elements method. This point of view, used exclusively by the mortar finite element method in the context of domain decomposition algorithms, allows to look at the discontinuous Galerkin method as a stabilization technique for approximations of saddle point problems. Then the constructions of Wang and Ye [31], Girault, Rivière and Wheeler [21] are particular choices of the discontinuous spaces. The point here is that to avoid the necessity of choosing stable pairs of spaces we stabilize this saddle point system, which is in general unstable, by adding "small" stabilization terms. These allow to formally eliminate the Lagrange multipliers ending up with a system that involves the velocity and the pressure only. We study two discretizations, one that leads to a symmetric problem and second that produces a nonsymetric linear system. Both methods share the advantages of local discontinuous Galerkin methods in [11, 14] by using discontinuous functions but our method has less unknowns since we do not introduce additional variables. The proposed discretization uses almost arbitrary finite element spaces of discontinuous functions that satisfy only quite mild restriction (4.6) and have optimal convergence rate.

This paper is organized as follows. In Section 2, preliminaries and notations are introduced. In Section 3, we derive two stabilized discontinuous finite element formulations for the Stokes equations. In Section 4 we analize the stability of the discontinuous Galerkin methods and finally in Section 5 we study the error.

2. Preliminaries and Notations

Let D be a bounded domain in $\mathbf{R}^d$. We use standard definitions for the Sobolev spaces $H^s(D)$ and the associated inner products $(\cdot,\cdot)_{s,D}$, norms $\|\cdot\|_{s,D}$, and seminorms $|\cdot|_{s,D}$ for $s \geq 0$. More precisely, for any integer $s \geq 0$, the seminorm $|\cdot|_{s,D}$ and norm $\|\cdot\|_{s,D}$ given by

$$|v|_{s,D} = \left(\sum_{|\alpha|=s} \int_D |\partial^\alpha v|^2 dD \right)^{\frac{1}{2}}, \quad \|v\|_{m,D} = \left(\sum_{s=0}^{m} |v|_{s,D}^2 \right)^{\frac{1}{2}}$$

with the usual notation $\alpha = (\alpha_1, \dots, \alpha_d)$, $|\alpha| = \alpha_1 + \dots + \alpha_d$, $\partial^\alpha = \partial_1^{\alpha_1} \dots \partial_d^{\alpha_d}$. Sobolev spaces of fractional order are defined by real method of interpolation and spaces of negative order by duality.

The space $H^0(D)$ coincides with $L^2(D)$, for which the norm and the inner product are denoted by $\|\cdot\|_D$ and $(\cdot,\cdot)_D$, respectively. When $D = \Omega$ we shall drop the subscript D in the norm and inner product notation. We also use $L_0^2(\Omega)$ to denote the subspace of $L^2(\Omega)$ of functions with mean value zero.

As we mentioned above the boundary value problem (1.1) – (1.3) has unique solution $\mathbf{u} \in H_0^1(\Omega)$ and $p \in L_0^2(\Omega)$ that satisfy the a priori estimate (1.4). For $\gamma \neq 0$ the pressure p will be in the space $L_0^2(\Omega)$ for $\mathbf{u}$ satisfying homogeneous Dirichlet boundary conditions, while for other boundary conditions this might not be valid. If $\gamma = 0$ the pressure is determined up to an additive constant which could be chosen so that $p \in L_0^2(\Omega)$.

Further, we partition Ω into a finite number of open non-overlapping subdomains K such that $\overline{\Omega} = \cup \overline{K}$. The set of all subdomains is denoted by $\mathcal{T}$. The intersection of two subdomains that has positive measure in $\mathbf{R}^{d-1}$ will be called interface (an edge in two dimensions) and is denoted by e. The set of all interfaces (edges) will be denoted by $\mathcal{E}_0$. We add to all such interfaces the intersections of a subdomain $\overline{K}$ and $\partial\Omega$ that have positive measure in $\mathbf{R}^{d-1}$ and denote this set by $\mathcal{E}$. The diameters of K and e are denoted by h_K and h_e, respectively. Further, we use the following notations for functions defined on Ω, possibly discontinuous across the boundaries between two adjacent subdomains:

$$V = \left\{ \mathbf{v} \in L^2(\Omega)^d : \ \mathbf{v}|_K \in H^1(K)^d, \ \Delta \mathbf{v}|_K \in L^2(K)^d, \ K \in \mathcal{T} \right\},$$

$$Q = \{ q \in L_0^2(\Omega) : \ q|_K \in H^1(K), \ K \in \mathcal{T} \}.$$

Multiplying the equations (1.1) and (1.2) by test functions $\mathbf{v} \in V$ and $q \in Q$, respectively, and integrating over a subdomain K by parts we get

$$(\nabla \mathbf{u}, \ \nabla \mathbf{v})_K - (\nabla \cdot \mathbf{v}, \ p)_K - \int_{\partial K} (\nabla \mathbf{u} \ \mathbf{n}) \cdot \mathbf{v} ds + \int_{\partial K} \mathbf{v} \cdot \mathbf{n} p ds \ = \ (\mathbf{f}, \ \mathbf{v})_K, \tag{2.1}$$

$$(\nabla \cdot \mathbf{u} + \gamma p, \ q)_K \ = \ 0. \tag{2.2}$$

Here $\mathbf{n}$ is the unit vector normal to ∂K and pointing outward to K. Note, that in the above formula $\nabla \mathbf{u}$ is a matrix with i-th row ∇u_i and $\nabla \mathbf{u} \ \mathbf{n} = \partial_{\mathbf{n}} \mathbf{u}$ is a vector with i-th component $\nabla u_i \cdot \mathbf{n}$. Since on the boundary, every vector field $\mathbf{v}$ can be decomposed in the form

$$\mathbf{v}|_{\partial K} = (\mathbf{v} \cdot \mathbf{n}) \, \mathbf{n} + \mathbf{n} \times (\mathbf{v} \times \mathbf{n}),$$

we obtain

$$(\nabla \mathbf{u} \ \mathbf{n}) \cdot \mathbf{v} \equiv \partial_{\mathbf{n}} \mathbf{u} \cdot \mathbf{v} = (\partial_{\mathbf{n}} \mathbf{u} \ \cdot \mathbf{n}) \, \mathbf{v} \cdot \mathbf{n} + (\partial_{\mathbf{n}} \mathbf{u} \ \times \mathbf{n}) \ \cdot (\mathbf{v} \times \mathbf{n}).$$

Now we use this identity to give (2.1) an equivalent form:

$$(\nabla \mathbf{u}, \ \nabla \mathbf{v})_K - (\nabla \cdot \mathbf{v}, \ p)_K + \int_{\partial K} (p - \partial_{\mathbf{n}} \mathbf{u} \cdot \mathbf{n}) \, \mathbf{v} \cdot \mathbf{n} \, ds - \int_{\partial K} (\partial_{\mathbf{n}} \mathbf{u} \times \mathbf{n}) \cdot (\mathbf{v} \times \mathbf{n}) \, ds = (\mathbf{f}, \ \mathbf{v})_K.$$

Next, we introduce as Lagrange multipliers the following traces on $e \in \mathcal{E}$

$$\mu = p - \partial_{\mathbf{n}} \mathbf{u} \ \cdot \mathbf{n}, \qquad \boldsymbol{\lambda} = -\partial_{\mathbf{n}} \mathbf{u} \ \times \mathbf{n}. \tag{2.3}$$

Obviously, if $p \in Q$ and $\mathbf{u} \in V$ then the traces exist and we can rewrite the above identity in the form

$$(\nabla \mathbf{u}, \ \nabla \mathbf{v})_K - (\nabla \cdot \mathbf{v}, \ p)_K + \langle \mu, \mathbf{v} \cdot \mathbf{n} \rangle_{\partial K} + \langle \boldsymbol{\lambda}, \mathbf{v} \times \mathbf{n} \rangle_{\partial K} = (\mathbf{f}, \ \mathbf{v})_K. \tag{2.4}$$

We have replaced the integrals $\int_{\partial K} \mu \, \mathbf{v} \cdot \mathbf{n} \, ds$ and $\int_{\partial K} \boldsymbol{\lambda} \cdot (\mathbf{v} \times \mathbf{n}) ds$ by $\langle \mu, \mathbf{v} \cdot \mathbf{n} \rangle_{\partial K}$ and $\langle \boldsymbol{\lambda}, \mathbf{v} \times \mathbf{n} \rangle_{\partial K}$, respectively. Loosely, here $\langle \mu, \mathbf{v} \cdot \mathbf{n} \rangle_{\partial K}$ could be interpreted also as duality pairing between $\mu \in H^{-1/2}(\partial K)$ and $\mathbf{v} \in V$. Similarly, $\langle \boldsymbol{\lambda}, \mathbf{v} \times \mathbf{n} \rangle_{\partial K}$ could be interpreted also as duality pairing for $\boldsymbol{\lambda} \in H^{-1/2}(\partial K)^d$ and $\mathbf{v} \in V$. We shall use these expressions for smooth functions only. For the precise Sobolev spaces that take into account the Dirichlet data on $\partial \Omega$ we refer to the literature on mortar finite element approximations of second order elliptic problems (see, e.g. [6]).

Now summing over all $K \in \mathcal{T}$ we get

$$(\nabla_{\mathcal{T}} \mathbf{u}, \ \nabla_{\mathcal{T}} \mathbf{v}) - (\nabla_{\mathcal{T}} \cdot \mathbf{v}, \ p) + \sum_{K \in \mathcal{T}} \langle \mu, \mathbf{v} \cdot \mathbf{n} \rangle_{\partial K} + \sum_{K \in \mathcal{T}} \langle \boldsymbol{\lambda}, \mathbf{v} \times \mathbf{n} \rangle_{\partial K} = (\mathbf{f}, \ \mathbf{v}). \tag{2.5}$$

Here and further $\nabla_{\mathcal{T}} \mathbf{v}$, $\nabla_{\mathcal{T}} \cdot \mathbf{v}$, and $\Delta_{\mathcal{T}} \mathbf{u}$ are the functions whose restriction to each subdomain $K \in \mathcal{T}$ is equal to $\nabla \mathbf{v}$, $\nabla \cdot \mathbf{v}$, and $\Delta \mathbf{u}$, respectively. Also, we define $h_{\mathcal{T}} \equiv h_{\mathcal{T}}(x) = h_K$ for $x \in K \in \mathcal{T}$ and $h_{\mathcal{E}}(x) \equiv h_{\mathcal{E}} = h_e$ for $x \in e \in \mathcal{E}$.

Let $e = \partial K_1 \cap \partial K_2$ be the common boundary (interface) between two subdomains K_1 and K_2 in $\mathcal{T}$, and $\mathbf{n}_1$ and $\mathbf{n}_2$ be unit normal vectors to e pointing to the exterior of K_1 and K_2, respectively. We define the average $\{\cdot\}$ and the jump $[\cdot]$ on $e \in \mathcal{E}_0$ for scalar q, vector $\mathbf{v}$, and matrix τ, respectively:

$$
\begin{aligned}
\{q\} &= \tfrac{1}{2}(q|_{\partial K_1 \cap e} + q|_{\partial K_2 \cap e}),\\
\{\mathbf{v}\} &= \tfrac{1}{2}(\mathbf{v}|_{\partial K_1 \cap e} + \mathbf{v}|_{\partial K_2 \cap e}),\\
\{\tau\} &= \tfrac{1}{2}(\tau|_{\partial K_1 \cap e} + \tau|_{\partial K_2 \cap e}),\\
[q] &= q|_{\partial K_1 \cap e} - q|_{\partial K_2 \cap e},\\
[\mathbf{v}\cdot\mathbf{n}] &= \mathbf{v}|_{\partial K_1 \cap e}\cdot\mathbf{n}_1 + \mathbf{v}|_{\partial K_2 \cap e}\cdot\mathbf{n}_2,\\
[\mathbf{n}\times(\mathbf{v}\times\mathbf{n})] &= \mathbf{n}_1\times(\mathbf{v}|_{\partial K_1 \cap e}\times\mathbf{n}_1) - \mathbf{n}_2\times(\mathbf{v}|_{\partial K_2 \cap e}\times\mathbf{n}_2),\\
[\tau\ \mathbf{n}] &= \tau|_{\partial K_1 \cap e}\ \mathbf{n}_1 + \tau|_{\partial K_2 \cap e}\ \mathbf{n}_2.
\end{aligned}
\tag{2.6}
$$

If e is part of the boundary $\partial\Omega$ then some of the above quantities are defined in the following way:

$$\{q\} = q|_{\partial\Omega\cap e},\ \{\mathbf{v}\} = \mathbf{v}|_{\partial\Omega\cap e},\ \{\tau\} = \tau|_{\partial\Omega\cap e}$$

and

$$[\mathbf{v}\cdot\mathbf{n}] = \mathbf{v}|_{\partial\Omega\cap e}\cdot\mathbf{n},\ [\mathbf{n}\times(\mathbf{v}\times\mathbf{n})] = \mathbf{n}\times(\mathbf{v}|_{\partial\Omega\cap e}\times\mathbf{n}),\ [\tau\ \mathbf{n}] = \tau|_{\partial\Omega\cap e}\ \mathbf{n}.$$

We rewrite (2.5) so that the solution $\mathbf{u}\in H^2(\Omega)^d,\ p\in L_0^2(\Omega)\cap H^1(\Omega)$ of the problem (1.1) - (1.3) and the Lagrange multipliers μ and $\boldsymbol{\lambda}$ satisfy the following equations for piece-wise smooth functions

$$
\begin{array}{lllllllll}
a^0(\mathbf{u},\mathbf{v}) & + & b(\mathbf{v},p) & + & m(\mu,\mathbf{v}) & + & l(\boldsymbol{\lambda},\mathbf{v}) & = & (\mathbf{f},\mathbf{v}),\ \forall\mathbf{v},\\
b(\mathbf{u},q) & - & c(p,q) & & & & & = & 0,\ \forall q,\\
m(\psi,\mathbf{u}) & & & & & & & = & 0,\ \forall\psi,\\
l(\boldsymbol{\phi},\mathbf{u}) & & & & & & & = & 0,\ \forall\boldsymbol{\phi},
\end{array}
\tag{2.7}
$$

where

$$
\begin{aligned}
a^0(\mathbf{u},\mathbf{v}) &= (\nabla_{\mathcal{T}}\mathbf{u},\ \nabla_{\mathcal{T}}\mathbf{v}),\\
b(\mathbf{v},q) &= -(\nabla_{\mathcal{T}}\cdot\mathbf{v},\ q),\\
c(p,q) &= \gamma(p,q),\\
l(\boldsymbol{\lambda},\mathbf{v}) &= \langle\boldsymbol{\lambda},[\mathbf{v}\times\mathbf{n}]\rangle_{\mathcal{E}},\\
m(\mu,\mathbf{v}) &= \langle\mu,[\mathbf{v}\cdot\mathbf{n}]\rangle_{\mathcal{E}}.
\end{aligned}
\tag{2.8}
$$

Here $\mathbf{v}\in V$ and for μ and q smooth functions, $\langle\mu,q\rangle_{\mathcal{E}}$ denotes the integration over the sum of all interfaces between the subdomains and on the boundary $\partial\Omega$, i.e.

$$\langle\mu,q\rangle_{\mathcal{E}} = \sum_{e\in\mathcal{E}}\int_e \mu q\,ds.$$

This is a typical saddle point problem in which the terms $l(\phi, \mathbf{u})$ and $m(\psi, \mathbf{u})$ impose weakly the continuity of the solution $\mathbf{u}$ across $\mathcal{E}$ and the homogeneous Dirichlet boundary condition on $\partial\Omega$.

Further we shall write (2.7) in the following concise form:

$$a^0(\mathbf{u}, \mathbf{v}) + b(\mathbf{v}, p) - b(\mathbf{u}, q) + l(\boldsymbol{\lambda}, \mathbf{v}) - l(\boldsymbol{\phi}, \mathbf{u}) + m(\mu, \mathbf{v}) - m(\psi, \mathbf{u}) + c(p, q) = (\mathbf{f}, \mathbf{v}). \tag{2.9}$$

Remark 2.1. *The above formulation is quite suitable for other types of boundary conditions. Consider the following two cases (various other boundary conditions could be found in* [19]*, pp. 179–183): on part $\Gamma_1 \subset \partial\Omega$ we specify $\mathbf{u} \times \mathbf{n} = \mathbf{0}$ and $p - \partial_{\mathbf{n}}\mathbf{u} \cdot \mathbf{n} = g_1$ and on another part $\Gamma_2 \subset \partial\Omega$ we have $\mathbf{u} \cdot \mathbf{n} = 0$ and $\partial_{\mathbf{n}}\mathbf{u} \times \mathbf{n} = \mathbf{g}_2$. Then one needs to modify the right hand side of (2.9) by adding the term*

$$\int_{\Gamma_1} g_1 \mathbf{v} \cdot \mathbf{n} \, ds + \int_{\Gamma_2} \mathbf{g}_2 \cdot (\mathbf{v} \times \mathbf{n}) \, ds$$

and to modify the bilinear forms $l(\boldsymbol{\lambda}, \mathbf{v})$ and $m(\mu, \mathbf{v})$ to

$$m(\mu, \mathbf{v}) = \langle \mu, [\mathbf{v} \cdot \mathbf{n}] \rangle_{\mathcal{E}_1}, \qquad l(\boldsymbol{\lambda}, \mathbf{v}) = \langle \boldsymbol{\lambda}, [\mathbf{v} \times \mathbf{n}] \rangle_{\mathcal{E}_2}.$$

Here $\mathcal{E}_1 \subset \mathcal{E}$ goes not include edges (faces) on Γ_1 and $\mathcal{E}_2 \subset \mathcal{E}$ goes not include edges (faces) on Γ_2.

Remark 2.2. *In the case of linear elastic deformations of non homogeneous media the above approach is not very suitable since in this case the equations are written in the form $-\nabla \cdot \sigma(\mathbf{u}) = \mathbf{f}$, where*

$$\sigma(\mathbf{u}) = \frac{\nu E}{(1+\nu)(1-2\nu)} \nabla \cdot \mathbf{u} \, I + \frac{E}{1+\nu} (\nabla \mathbf{u} + (\nabla \mathbf{u})^T)$$

is the stress tensor and the Young's modulus E and the Poisson's ratio ν depend on $x \in \Omega$. A discontinuous Galerkin method for these equations was studied by Hansbo and Larson in [23].

Here we shall avoid the delicate question of exact Sobolev spaces for the solution and test functions $\mathbf{u}, \mathbf{v}$, μ, ψ, and $\boldsymbol{\lambda}, \boldsymbol{\phi}$. As we shall see later, this is not relevant for the method we derive. The interested reader can find quite complete discussion for the case of two dimension in the work by Girault, Rivière, and Wheeler [21]. However, functions that are smooth on each subdomain K the above identities makes sense.

Remark 2.3. *Obviously, if $\mathbf{u} \in H(div; \Omega)$, then $[\mathbf{u} \cdot \mathbf{n}] = 0$ and the above system simplifies, by omitting the terms with $m(\cdot, \cdot)$ in the left hand side of (2.9). Similarly, if $\mathbf{u} \in H(curl; \Omega)$ then we get different simplification by omitting the terms $l(\cdot, \cdot)$ in (2.9).*

3. Finite element discretization

Here we shall introduce finite element discretization of the above saddle point problem using finite element spaces of discontinuous functions.

3.1. **First attempt.** Following [18] we shall derive our discontinuous finite element approximation by stabilizing a discretization that is usually unstable for the whole range of the parameter $\gamma \geq 0$. Assume that $\mathcal{T}$ is a partition of the domain Ω into finite elements K (triangles, tetrahedra, rectangles, bricks, quadrilaterals, etc) with mesh size h_K, so that the partition $\mathcal{T}$ is quasi-uniform, i.e. it is regular and satisfies the inverse assumption (see [10]).

Define the finite element space V_h for the velocity by

$$V_h = \{\mathbf{v} \in L^2(\Omega)^d : \ \mathbf{v}|_K \in \mathcal{V}(K)^d, \quad \forall K \in \mathcal{T}\}$$

and the finite element space Q_h for the pressure by

$$Q_h = \{q \in L^2_0(\Omega) : \ q|_K \in \mathcal{Q}(K), \quad \forall K \in \mathcal{T}\}.$$

Further, the finite elements spaces Λ_h and M_h for the Lagrange multipliers $\boldsymbol{\lambda}$ and μ

$$\Lambda_h = \{\boldsymbol{\lambda} \in L^2(\mathcal{E})^{d-1} : \ \boldsymbol{\lambda}|_e \in \mathcal{L}(e)^{d-1}, \ e \text{ is a common edge of two finite elements }\},$$

$$M_h = \{\mu \in L^2(\mathcal{E}) : \ \mu|_e \in \mathcal{M}(e), \ e \text{ is a common edge of two finite elements}\}.$$

Here the local spaces $\mathcal{V}(K)$, $\mathcal{P}(K)$, $\mathcal{L}(e)$, and $\mathcal{M}(e)$ consist of polynomials that will be specified later. Further, we shall use also the notation $V(h) = V_h + H^{l+1}(\Omega)^d$ and $Q(h) = Q_h + H^{m+1}(\Omega) \cap L^2_0(\Omega)$, where $l \geq 1$ and $m \geq 0$. The finite element discretization of (2.9) is: find $\mathbf{u}_h \in V_h$, $p \in Q_h$, $\boldsymbol{\lambda}_h \in \Lambda_h$, and $\mu_h \in M_h$ such that

$$\begin{aligned} &a^0(\mathbf{u}_h, \mathbf{v}) + b(\mathbf{v}, p_h) - b(\mathbf{u}_h, q) + l(\boldsymbol{\lambda}_h, \mathbf{v}) - l(\phi, \mathbf{u}_h) \\ &\qquad + m(\mu_h, \mathbf{v}) - m(\psi, \mathbf{u}_h) + c(p_h, q) = (\mathbf{f}, \mathbf{v}), \\ &\qquad \forall \mathbf{v} \in V_h, \quad \forall q \in Q_h, \quad \forall \phi \in \Lambda_h, \quad \forall \psi \in M_h. \end{aligned} \tag{3.10}$$

Without proper alignment of the finite dimensional spaces V_h, Q_h, Λ_h, and M_h (they need to satisfy appropriate *inf-sup condition*) the above saddle point problem is in general unstable.

In the past various stabilization procedures for the saddle point problem (3.10) have been proposed and studied (see, e.g. [9], [18], [26], [28], [30]). As we mentioned earlier, if $V_h \subset H^1(\Omega)$, then the forms $m(\cdot,\cdot) \equiv 0$ and $l(\cdot,\cdot) \equiv 0$ so the problem is reduced to $a^0(\mathbf{u}_h, \mathbf{v}) + b(\mathbf{v}, p_h) - b(\mathbf{u}_h, q) = (\mathbf{f}, \mathbf{v})$. For various stabilization procedures of this problem we refer to [9], [18], [26], [27], [28], [30]. In general, the stabilization is achieved by adding "small" term to the left hand side. For example, Brezzi and Pitkäranta [9] add the term $(h^2_{\mathcal{T}} \nabla_{\mathcal{T}} p_h, \ \nabla_{\mathcal{T}} q)$. However, this term does not vanish at the solution so it yield only a first order scheme. Hughes, Franca, and Balestra [26] use a stabilization term $\delta(\nabla_{\mathcal{T}} p_h - \Delta_{\mathcal{T}} \mathbf{u} - \mathbf{f}, h^2_{\mathcal{T}} \ \nabla_{\mathcal{T}} q)$ and prove that for $0 < \delta$ the method is stable and has an optimal order of convergence when the space V_h and Q_h contain piece-wise polynomials of the same order. Finally, for discontinuous pressure spaces Kecher and Sylvester [28] use a stabilization term $\delta\langle h_{\mathcal{E}}[p_h], [q]\rangle_{\mathcal{E}}$, which produces a stable method for $\delta > 0$ that is convergent of first order.

Stabilization of the problem (3.10) for spaces V_h that contain discontinuous functions could be achieved by adding a "small" term $\epsilon\langle \boldsymbol{\lambda}_h, \boldsymbol{\phi}\rangle + \delta\langle [p_h], [q]\rangle_{\mathcal{E}}$. By formally eliminating the Lagrange multiplier $\boldsymbol{\lambda}_h$ we get a system for $\mathbf{u}_h$ and p_h. However, this system is inconsistent, since the added term does not vanish on the exact solution. Such "penalty" formulation, leading to a low order approximation, has been studied in the past (see, e.g. Quarteroni and Valli in [30], p. 312).

3.2. **Motivation for a stabilized approximation.** Now we consider the case of fully discontinuous finite element spaces. The terms $b(\mathbf{u}_h, q)$ $l(\boldsymbol{\phi}, \mathbf{u}_h)$ and $m(\psi, \mathbf{u}_h)$ in (3.10) could be considered as constraints and therefore to have a well-posed problem we need three appropriate inf-sup condition. However, we can stabilize (3.10) by adding three stabilization terms

$$\begin{aligned} &\langle \delta[p_h], [q]\rangle_{\mathcal{E}}, \\ &\langle \epsilon(\boldsymbol{\lambda}_h + \{\partial_{\mathbf{n}}(\mathbf{u}_h \times \mathbf{n})\ \}), \boldsymbol{\phi}\rangle_{\mathcal{E}}, \\ &\langle \epsilon(\mu_h - \{p_h\} + \{\partial_{\mathbf{n}}(\mathbf{u}_h \cdot \mathbf{n})\ \}), \psi\rangle_{\mathcal{E}}. \end{aligned} \tag{3.11}$$

Here $\epsilon > 0$ and $\delta > 0$ are small parameters, in fact these are functions on $\mathcal{E}$, which will be defined later. These three terms are supposed to stabilize the problem so that each is assumed to avoid the necessary inf-sup condition for the form $b(\cdot,\cdot)$, $l(\cdot,\cdot)$, and $m(\cdot,\cdot)$, correspondingly. This allows us to formally eliminate both Lagrange multipliers μ_h and $\boldsymbol{\lambda}_h$ and get

$$a_h^0(\mathbf{u}_h, \mathbf{v}) + b_h(\mathbf{v}, p_h) - b(\mathbf{u}_h, q) + c_h(p_h, q) = (\mathbf{f},\ \mathbf{v}),\ \forall \mathbf{v} \in V_h,\ \forall q \in M_h,$$

where

$$\begin{aligned} a_h^0(\mathbf{u}_h, \mathbf{v}) &= (\nabla_{\mathcal{T}} \mathbf{u}_h,\ \nabla_{\mathcal{T}} \mathbf{v}) + \langle \epsilon^{-1}[\mathbf{u}_h \times \mathbf{n}], [\mathbf{v} \times \mathbf{n}]\rangle_{\mathcal{E}} + \langle \epsilon^{-1}[\mathbf{u}_h \cdot \mathbf{n}], [\mathbf{v} \cdot \mathbf{n}]\rangle_{\mathcal{E}} \\ &\quad - \langle \{\partial_{\mathbf{n}} \mathbf{u}_h\ \times \mathbf{n}\}, [\mathbf{v} \times \mathbf{n}]\rangle_{\mathcal{E}} - \langle \{\partial_{\mathbf{n}} \mathbf{u}_h\ \cdot \mathbf{n}\}, [\mathbf{v} \cdot \mathbf{n}]\rangle_{\mathcal{E}}, \\ b_h(\mathbf{v}, p_h) &= -(\nabla_{\mathcal{T}} \cdot \mathbf{v},\ p_h) + \langle \{p_h\}, [\mathbf{v} \cdot \mathbf{n}]\rangle_{\mathcal{E}}, \\ c_h(p_h, q) &= \gamma(p_h, q) + \langle \delta[p_h], [q]\rangle_{\mathcal{E}_0}. \end{aligned} \tag{3.12}$$

The bilinear form $a_h^0(\cdot,\cdot)$ is nonsymmetric so if we want to get a symmetric form we should use symmetric stabilization terms. namely, we can replace the last two terms of (3.11) by

$$\begin{aligned} &\langle \epsilon(\boldsymbol{\lambda}_h + \{\partial_{\mathbf{n}} \mathbf{u}_h\ \times \mathbf{n}\}), \boldsymbol{\phi} + \{\partial_{\mathbf{n}} \mathbf{v}\ \times \mathbf{n}\}\rangle_{\mathcal{E}}, \\ &\langle \epsilon(\mu_h - \{p_h\} + \{\partial_{\mathbf{n}} \mathbf{u}_h\ \cdot \mathbf{n}\}), \psi - \{q\} + \{\partial_{\mathbf{n}} \mathbf{v}\ \cdot \mathbf{n}\}\rangle_{\mathcal{E}}. \end{aligned} \tag{3.13}$$

Again after formally eliminating $\boldsymbol{\lambda}_h$ and μ_h we get the problem

$$a_h^s(\mathbf{u}_h, \mathbf{v}) + b_h(\mathbf{v}, p_h) - b_h(\mathbf{u}_h, q) + c_h(p_h, q) = (\mathbf{f},\ \mathbf{v}),\ \forall \mathbf{v} \in V_h,\ \forall q \in M_h, \tag{3.14}$$

where

$$\begin{aligned} a_h^s(\mathbf{u}_h, \mathbf{v}) = \ & (\nabla_{\mathcal{T}} \mathbf{u}_h,\ \nabla_{\mathcal{T}} \mathbf{v}) + \langle \epsilon^{-1}[\mathbf{u}_h \times \mathbf{n}], [\mathbf{v} \times \mathbf{n}]\rangle_{\mathcal{E}} + \langle \epsilon^{-1}[\mathbf{u}_h \cdot \mathbf{n}], [\mathbf{v} \cdot \mathbf{n}]\rangle_{\mathcal{E}} \\ & - \langle \{\partial_{\mathbf{n}} \mathbf{u}_h\ \times \mathbf{n}\}, [\mathbf{v} \times \mathbf{n}]\rangle_{\mathcal{E}} - \langle \{\partial_{\mathbf{n}} \mathbf{v}\ \times \mathbf{n}\}, [\mathbf{u}_h \times \mathbf{n}]\rangle_{\mathcal{E}} \\ & - \langle \{\partial_{\mathbf{n}} \mathbf{u}_h\ \cdot \mathbf{n}\}, [\mathbf{v} \cdot \mathbf{n}]\rangle_{\mathcal{E}} - \langle \{\partial_{\mathbf{n}} \mathbf{v}\ \cdot \mathbf{n}\}, [\mathbf{u}_h \cdot \mathbf{n}]\rangle_{\mathcal{E}}. \end{aligned} \tag{3.15}$$

is a symmetric bilinear from on $V_h \times V_h$, which will produce IP (interior penalty) DG method for the Stokes system.

Similarly, we can get a problem with a skew-symmetric nonsymmetric part if we add

$$\langle \epsilon(\boldsymbol{\lambda}_h + \{\partial_{\mathbf{n}} \mathbf{u}_h\ \times \mathbf{n}\}), \boldsymbol{\phi} - \{\partial_{\mathbf{n}} \mathbf{v}\ \times \mathbf{n}\}\rangle_{\mathcal{E}}$$

$$\langle \epsilon(\mu_h - \{p_h\} + \{\partial_{\mathbf{n}} \mathbf{u}_h \cdot \mathbf{n}\}), \psi + \{q\} - \{\partial_{\mathbf{n}} \mathbf{v} \cdot \mathbf{n}\}\rangle_{\mathcal{E}}.$$

Again after elimination of $\boldsymbol{\lambda}_h$ and μ_h we get a problem

$$a_h^{ns}(\mathbf{u}_h, \mathbf{v}) + b_h(\mathbf{v}, p_h) - b_h(\mathbf{u}_h, q) + c_h(p_h, q) = (\mathbf{f},\ \mathbf{v}),\ \forall \mathbf{v} \in V_h,\ \forall q \in M_h, \tag{3.16}$$

where

$$\begin{aligned} a_h^{ns}(\mathbf{u}_h, \mathbf{v}) &= (\nabla_{\mathcal{T}} \mathbf{u}_h,\ \nabla_{\mathcal{T}} \mathbf{v}) + \langle \epsilon^{-1}[\mathbf{u}_h \times \mathbf{n}], [\mathbf{v} \times \mathbf{n}]\rangle_{\mathcal{E}} + \langle \epsilon^{-1}[\mathbf{u}_h \cdot \mathbf{n}], [\mathbf{v} \cdot \mathbf{n}]\rangle_{\mathcal{E}} \\ & -\langle \{\partial_{\mathbf{n}} \mathbf{u}_h \times \mathbf{n}\}, [\mathbf{v} \times \mathbf{n}]\rangle_{\mathcal{E}} + \langle \{\partial_{\mathbf{n}} \mathbf{v} \times \mathbf{n}\}, [\mathbf{u}_h \times \mathbf{n}]\rangle_{\mathcal{E}} \\ & -\langle \{\partial_{\mathbf{n}} \mathbf{u}_h \cdot \mathbf{n}\}, [\mathbf{v} \cdot \mathbf{n}]\rangle_{\mathcal{E}} + \langle \{\partial_{\mathbf{n}} \mathbf{v} \cdot \mathbf{n}\}, [\mathbf{u}_h \cdot \mathbf{n}]\rangle_{\mathcal{E}}. \end{aligned} \tag{3.17}$$

The bilinear from $a_h^{ns}(\cdot, \cdot)$ has a skew-symmetric nonsymmetric part and is coercive on $V_h \times V_h$ for any $\epsilon > 0$. This form in fact will produce the NIP (nonsymmetric interior penalty) discontinuous Galerkin method for the Stokes system.

Remark 3.1. *The case $V_h \subset H(div; \Omega)$ will simplify the bilinear forms a_h^{ns} or a_h^s and b_h so that we get the problem*

$$a_{div}^{\mp}(\mathbf{u}_h, \mathbf{v}) + b(\mathbf{v}, p_h) - b(\mathbf{u}_h, q) + c_h(p_h, q) = (\mathbf{f},\ \mathbf{v}), \forall \mathbf{v} \in V_h, \forall q \in M_h, \tag{3.18}$$

where $a_{div}^{\mp}(\mathbf{u}_h, \mathbf{v})$ is defined by

$$\begin{aligned} a_{div}^{\mp}(\mathbf{u}_h, \mathbf{v}) &= (\nabla_{\mathcal{T}} \mathbf{u}_h,\ \nabla_{\mathcal{T}} \mathbf{v}) + \langle \epsilon^{-1}[\mathbf{u}_h \times \mathbf{n}], [\mathbf{v} \times \mathbf{n}]\rangle_{\mathcal{E}} \\ & -\langle \{\partial_{\mathbf{n}} \mathbf{u}_h \times \mathbf{n}\}, [\mathbf{v} \times \mathbf{n}]\rangle_{\mathcal{E}} \mp \langle [\mathbf{u}_h \times \mathbf{n}], \{\partial_{\mathbf{n}} \mathbf{v} \times \mathbf{n}\}\rangle_{\mathcal{E}}. \end{aligned} \tag{3.19}$$

Both systems were analysed by Wang and Ye in [31] *for $\epsilon = \alpha h_E^{-1}$ and $\delta = 0$ and for appropriate choice of the finite element spaces V_h and M_h. In* [31] *it was shown that the bilinear form $a_{div}^{-}(\mathbf{u}_h, \mathbf{v})$ is symmetric and coercive for sufficiently large α, while the bilinear form $a_{div}^{+}(\mathbf{u}_h, \mathbf{v})$ has skew-symmetric nonsymmetric part and is coercive for any $\alpha > 0$.*

Remark 3.2. *The case $V_h \subset H(curl; \Omega)$ will lead to another simplification of the bilinear forms a_h^{ns}, a_h^s and b_h so that we get the problem*

$$a_{curl}^{\mp}(\mathbf{u}_h, \mathbf{v}) + b_h(\mathbf{v}, p_h) - b_h(\mathbf{u}_h, q) + c_h(p_h, q) = (\mathbf{f},\ \mathbf{v}),\ \forall \mathbf{v} \in V_h, \forall q \in M_h, \tag{3.20}$$

where $a_{curl}^{\mp}(\mathbf{u}_h, \mathbf{v})$ is defined by

$$\begin{aligned} a_{curl}^{\mp}(\mathbf{u}_h, \mathbf{v}) &= (\nabla_{\mathcal{T}} \mathbf{u}_h,\ \nabla_{\mathcal{T}} \mathbf{v}) + \langle \epsilon^{-1}[\mathbf{u}_h \cdot \mathbf{n}], [\mathbf{v} \cdot \mathbf{n}]\rangle_{\mathcal{E}} \\ & -\langle \{\partial_{\mathbf{n}} \mathbf{u}_h \cdot \mathbf{n}\}, [\mathbf{v} \cdot \mathbf{n}]\rangle_{\mathcal{E}} \mp \langle [\mathbf{u}_h \cdot \mathbf{n}], \{\partial_{\mathbf{n}} \mathbf{v} \cdot \mathbf{n}\}\rangle_{\mathcal{E}}. \end{aligned} \tag{3.21}$$

Remark 3.3. *Note that the discrete problems (3.14) or (3.16) do not require $\mathbf{u} = 0$ on the boundary $\partial\Omega$. This boundary condition is imposed in a weak sense by the penalty term*

$$\langle \epsilon^{-1}[\mathbf{u}_h \times \mathbf{n}], [\mathbf{v} \times \mathbf{n}]\rangle_{\mathcal{E}} + \langle \epsilon^{-1}[\mathbf{u}_h \cdot \mathbf{n}], [\mathbf{v} \cdot \mathbf{n}]\rangle_{\mathcal{E}}.$$

Remark 3.4. *If in (3.11) $\delta = 0$, then V_h and Q_h need to satisfy an appropriate inf-sup condition. Such approach for 2-D Stokes equations has been studied by Girault, Rivière, and Wheeler in* [21].

Now we shall get a more concise form of the bilinear forms (3.15) and (3.17) by introducing a new notation. Following [11], for vectors $\mathbf{v}$ and $\mathbf{n}$, let $\mathbf{v} \otimes \mathbf{n}$ denote the matrix whose ij-th element is $v_i n_j$. For a vector $\mathbf{w}$ we define a matrix valued jump $[\![\cdot]\!]$ as

$$[\![\mathbf{w}]\!] = \mathbf{w}|_{\partial K_1} \otimes \mathbf{n}_1 + \mathbf{w}|_{\partial K_2} \otimes \mathbf{n}_2$$

where $e \in \mathcal{E}$ is an edge (face) shared by two adjacent finite elements K and K_j. If $e \in \mathcal{E}$ is an edge on the boundary $\partial\Omega$, define $[\![\mathbf{w}]\!] = \mathbf{w} \otimes \mathbf{n}$. Further, for two matrix valued variable σ and τ we use

$$\sigma : \tau = \sum_{i,j=1}^{d} \sigma_{ij}\tau_{ij}, \;\; \sigma, \;\; \tau \in R^{d\times d}.$$

Using these concise notations we can show that

$$\langle\{\partial_{\mathbf{n}}\mathbf{u}_h \times \mathbf{n}\}, [\mathbf{v}\times\mathbf{n}]\rangle_{\mathcal{E}} + \langle\{\partial_{\mathbf{n}}\mathbf{u}_h \cdot \mathbf{n}\}, [\mathbf{v}\cdot\mathbf{n}]\rangle_{\mathcal{E}} = \langle\{\nabla\mathbf{u}\} : [\![\mathbf{v}]\!]\rangle_{\mathcal{E}}$$

and

$$\langle\epsilon^{-1}[\mathbf{u}_h\times\mathbf{n}], [\mathbf{v}\times\mathbf{n}]\rangle_{\mathcal{E}} + \langle\epsilon^{-1}[\mathbf{u}_h\cdot\mathbf{n}], [\mathbf{v}\cdot\mathbf{n}]\rangle_{\mathcal{E}} = \langle\epsilon^{-1}[\![\mathbf{u}_h]\!] : [\![\mathbf{v}]\!]\rangle_{\mathcal{E}}.$$

Thus, the bilinear forms defined in (3.15) and (3.17) could be written in the form

$$a_h^s(\mathbf{u}_h, \mathbf{v}) = a_h^0(\mathbf{u}_h, \mathbf{v}) - \langle\{\nabla\mathbf{v}\} : [\![\mathbf{u}_h]\!]\rangle_{\mathcal{E}} \tag{3.22}$$

and

$$a_h^{ns}(\mathbf{u}_h, \mathbf{v}) = a_h^0(\mathbf{u}_h, \mathbf{v}) + \langle\{\nabla\mathbf{v}\} : [\![\mathbf{u}_h]\!]\rangle_{\mathcal{E}} \tag{3.23}$$

where

$$a_h^0(\mathbf{u}_h, \mathbf{v}) = (\nabla_{\mathcal{T}}\mathbf{u}_h, \; \nabla_{\mathcal{T}}\mathbf{v}) - \langle\{\nabla\mathbf{u}_h\} : [\![\mathbf{v}]\!]\rangle_{\mathcal{E}} + \langle\epsilon^{-1}[\![\mathbf{u}_h]\!] : [\![\mathbf{v}]\!]\rangle_{\mathcal{E}} \tag{3.24}$$

Using these new notations we can rewrite the problems (3.14) and (3.16) as: find $u_h \in V_h$ and $p_h \in Q_h$ such that

$$A(\mathbf{u}_h, p_h; \; \mathbf{v}, q) = (\mathbf{f}, \; \mathbf{v}) \qquad \forall\mathbf{v}\in V_h, \; \forall q \in Q_h, \tag{3.25}$$

where

$$A(\mathbf{v}, q; \; \mathbf{w}, \; r) = a_h(\mathbf{v}, \; \mathbf{w}) + b_h(\mathbf{w}, \; q) - b_h(\mathbf{v}, \; r) + c_h(q, \; r). \tag{3.26}$$

Here and further $a_h(\mathbf{v}, \; \mathbf{w})$ is either $a_h^s(\mathbf{v}, \; \mathbf{w})$ or $a_h^{ns}(\mathbf{v}, \; \mathbf{w})$. Thus, (3.25) incorporates two methods, one involving the symmetric bilinear form $a_h^s(\mathbf{v}, \mathbf{w})$ and another with a nonsymmetric form $a_h^{ns}(\mathbf{v}, \mathbf{w})$.

The sufficiently smooth solution $\mathbf{u}$, p of the problem (1.1) – (1.3) satisfies the identity

$$A(\mathbf{u}, p; \; \mathbf{v}, \; q) = (\mathbf{f}, \; \mathbf{v}) \qquad \forall\mathbf{v}\in V_h, \; \forall q \in Q_h$$

so that subtracting this from (3.25) we get

$$A(\mathbf{u} - \mathbf{u}_h, p - p_h; \; \mathbf{v}, \; q) = 0 \qquad \forall\mathbf{v}\in V_h, \; \forall q \in Q_h. \tag{3.27}$$

This Galerkin orthogonality condition implies that both methods, the symmetric and the nonsymmetric, are consistent.

4. Stabilized finite element approximation for fully discontinuous spaces

Our goal of this section is to choose the parameters ϵ and δ and the spaces V_h and Q_h in (3.25) so that the method is stable and converges with optimal order.

First, we shall specify the finite element spaces. Let $P_k(K)$ be the set of all polynomials on K of degree less than or equal to k. We will assume that $P_l(K) \subset \mathcal{V}(K)$ and $P_m(K) \subset \mathcal{Q}(K)$ with $l \geq 1$ and $m \geq 0$.

For $K \in \mathcal{T}$ let $\Pi_1^K : H^{l+1}(K)^d \to \mathcal{V}(K)^d$ be the orthogonal L^2 projection. Since $P_l(K) \subset \mathcal{V}(K)$ it is well known that for a $\mathbf{w} \in H^{l+1}(K)^d$

$$|\mathbf{w} - \Pi_1^K \mathbf{w}|_{s,K} \leq C h_K^{l+1-s} \|\mathbf{w}\|_{l+1,K} \quad s = 0, 1, 2, \tag{4.1}$$

$$\|\mathbf{w} - \Pi_1^K \mathbf{w}\|_{\partial K} \leq C h_K^{l+\frac{1}{2}} \|\mathbf{w}\|_{l+1,K}, \tag{4.2}$$

where the constant C depends on l and the minimum angle of the finite element K. In a similar manner we define the orthogonal projection $\Pi_2^K : H^{m+1}(K) \to \mathcal{Q}(K)$ so that

$$|\psi - \Pi_2^K \psi|_{s,K} \leq C h_K^{m+1-s} \|\psi\|_{m+1,K} \quad s = 0, 1, 2 \tag{4.3}$$

$$\|\psi - \Pi_2^K \psi\|_{\partial K} \leq C h_K^{m+\frac{1}{2}} \|\psi\|_{m+1,K}. \tag{4.4}$$

When dealing with fully discontinuous spaces we can define the global L^2 orthogonal projection operators $\Pi_1 : V(h) \to V_h$ and $\Pi_2 : Q(h) \to Q_h$ as

$$\Pi_1 \mathbf{v}(x) = \Pi_1^K \mathbf{v}(x), \quad \Pi_2 q(x) = \Pi_2^K q(x), \; x \in K, \; K \in \mathcal{T}.$$

Next, we specify the "small" parameters ϵ and δ. Namely, we choose

$$\epsilon = \epsilon(x) = \alpha_1^{-1} h_e, \quad \delta = \delta(x) = \alpha_2 h_e \quad \text{for} \quad x \in e \in \mathcal{E} \tag{4.5}$$

where h_e is the length of the edge e for $d = 2$ and the diameter of the face e for $d = 3$. The positive numbers α_1, α_2 will be determined later.

Now we formulate two essential assumptions under which we shall carry our analysis.

Assumption 4.1. *The local polynomial spaces $\mathcal{Q}(K)$ and $\mathcal{V}(K)$ satisfy the inclusion*

$$\nabla \mathcal{Q}(K) \subset \mathcal{V}(K)^d. \tag{4.6}$$

Assumption 4.2. *If $a_h(\cdot,\cdot) = a_h^{ns}(\cdot,\cdot)$ then $\alpha_1 > 0$ is any fixed constant, if $a_h(\cdot,\cdot) = a_h^s(\cdot,\cdot)$ then $\alpha_1 > 0$ is sufficiently large, and α_2 is any positive constant.*

Note, that in order to achieve optimal error approximation the polynomial spaces for the velocity should be one degree higher than the spaces for the pressure. Therefore, assumption 4.1 represents a very mild restriction on the local polynomial spaces. In addition to (4.6) in [11] the inclusion $\nabla \cdot \mathcal{V}(K) \subset \mathcal{Q}(K)$ is required as well. This implies that here we have more freedom in choosing the finite element spaces.

For $\mathbf{v} \in V(h)$ we introduce two norms $|\!|\!| \cdot |\!|\!|_1$ and $|\!|\!| \cdot |\!|\!|_*$

$$|\!|\!|\mathbf{v}|\!|\!|_1^2 = (\nabla_{\mathcal{T}} \mathbf{v}, \nabla_{\mathcal{T}} \mathbf{v}) + \langle h_e^{-1} [\![\mathbf{v}]\!], [\![\mathbf{v}]\!] \rangle_{\mathcal{E}} := \|\nabla_{\mathcal{T}} \mathbf{v}\|^2 + \sum_{e \in \mathcal{E}} h_e^{-1} \|[\![\mathbf{v}]\!]\|_e^2, \tag{4.7}$$

$$|\!|\!|\mathbf{v}|\!|\!|_*^2 = |\!|\!|\mathbf{v}|\!|\!|_1^2 + \sum_{K \in \mathcal{T}} h_K^2 \|\Delta \mathbf{v}\|_K^2 := |\!|\!|\mathbf{v}|\!|\!|_1^2 + (h_{\mathcal{T}} \Delta_{\mathcal{T}} \mathbf{v}, h_{\mathcal{T}} \Delta_{\mathcal{T}} \mathbf{v}) := |\!|\!|\mathbf{v}|\!|\!|_1^2 + \|h_{\mathcal{T}} \Delta_{\mathcal{T}} \mathbf{v}\|^2. \tag{4.8}$$

Also, for $q \in Q(h)$ we shall use the following notation

$$\|\!|q\|\!|_*^2 = \gamma \|q\|^2 + \sum_{e\in\mathcal{E}_0} h_e \|[q]\|_e^2. \tag{4.9}$$

Note that if $\gamma > 0$, then (4.9) is a norm, while if $\gamma = 0$ it is a semi-norm. Since $0 \le \gamma \le C$ it is easy to see that

$$\|\!|q\|\!|_*^2 \le C\|q\|, \qquad \forall q \in Q_h.$$

Let $K \in \mathcal{T}$ has an edge e. It is well-known that there exists a constant C such that for any function $g \in H^1(K)$,

$$\|g\|_e^2 \le C\left(h_K^{-1}\|g\|_K^2 + h_K\|\nabla g\|_K^2\right). \tag{4.10}$$

In particular, for any $\mathbf{v} \in V_h$, we have

$$h_e\|\nabla \mathbf{v}|_K\|_e^2 \le C\left(\|\nabla \mathbf{v}\|_K^2 + h_K^2\|\Delta \mathbf{v}\|_K^2\right)$$

and the standard inverse inequality yields

$$h_K\|\Delta \mathbf{v}\|_K^2 \le C\|\nabla \mathbf{v}\|_K^2 \qquad \forall \mathbf{v} \in V_h.$$

Therefore, there are positive constants c, C independent of h such that

$$c\|\!|\mathbf{v}\|\!|_1 \le \|\!|\mathbf{v}\|\!|_* \le C\|\!|\mathbf{v}\|\!|_1 \quad , \forall \mathbf{v} \in V_h. \tag{4.11}$$

The following lemma provides the estimates for $\|\!|\mathbf{v} - \Pi_1\mathbf{v}\|\!|_1$ and $\|\!|q - \Pi_2 q\|\!|$.

Lemma 4.1. *For any $q \in H^{m+1}(\Omega)$ and $\mathbf{v} \in H^{l+1}(\Omega)^d$, one has*

$$\|\!|\mathbf{v} - \Pi_1\mathbf{v}\|\!|_* \le Ch^l\|\mathbf{v}\|_{l+1} \tag{4.12}$$

$$\|\!|q - \Pi_2 q\|\!|_* \le Ch^{m+1}\|q\|_{m+1}. \tag{4.13}$$

Proof. Using the definition of Π_1, Π_2 and (4.1), (4.2), we first show the estimates

$$|\nabla_{\mathcal{T}}(\mathbf{v} - \Pi_1\mathbf{v})|^2 \le Ch^{2l}\|\mathbf{v}\|_{l+1}^2, \tag{4.14}$$

$$\sum_{e\in\mathcal{E}} h_e^{-1}\|[\![\mathbf{v} - \Pi_1\mathbf{v}]\!]\|_e^2 \le Ch^{2l}\|\mathbf{v}\|_{l+1}^2, \tag{4.15}$$

$$\|q - \Pi_2 q\|^2 \le Ch^{2(m+1)}\|q\|_{m+1}^2, \tag{4.16}$$

$$\sum_{e\in\mathcal{E}_0} h_e\|[q - \Pi_2 q]\|_e^2 \le Ch^{2(m+1)}\|q\|_{m+1}^2. \tag{4.17}$$

Then (4.10) and (4.1) imply

$$\sum_{e\in\mathcal{E}} h_e^{-1}\|[\![\mathbf{v} - \Pi_1\mathbf{v}]\!]\|_e^2 \le C\sum_{K\in\mathcal{T}}(|\mathbf{v} - \Pi_1\mathbf{v}|_{1,K}^2 + h_K^2|\mathbf{v} - \Pi_1\mathbf{v}|_{2,K}^2) \le Ch^{2l}\|\mathbf{v}\|_{l+1}^2. \tag{4.18}$$

We complete the proof by taking into account the definitions of $\|\!|\mathbf{v}\|\!|_*$ and $\|\!|q\|\!|_*$, and the estimates (4.14)-(4.18). □

Theorem 4.1. *Let the assumptions 4.1 and 4.2 hold. Then the bilinear form* $A(\mathbf{v}, q;\ \mathbf{w}, r)$ *satisfies the inf-sup condition*

$$c_0(\|\!|\mathbf{v}|\!\|_1 + \|q\|) \le \sup_{\mathbf{w}\in V_h,\ r\in Q_h} \frac{A(\mathbf{v}, q;\ \mathbf{w}, r)}{\|\!|\mathbf{w}|\!\|_1 + \|r\|}, \quad \forall(\mathbf{v},\ q) \in V_h \times Q_h \tag{4.19}$$

with a constant $c_0 > 0$ *independent of* h *and* γ.

For $(\mathbf{v},\ q) \in V(h) \times Q(h)$ *the bilinear form* $A(\mathbf{v}, q;\ \mathbf{w}, r)$ *is continuous so that*

$$A(\mathbf{v}, q;\ \mathbf{w}, r) \le c_1(\|\!|\mathbf{v}|\!\|_* + \|q\|)(\|\!|\mathbf{w}|\!\|_1 + \|r\|), \quad \forall(\mathbf{w},\ r) \in V_h \times Q_h \tag{4.20}$$

with constants $c_1 > 0$ *independent of* h.

The proof will be based on several lemmas we shall prove below.

First, we show the coercivity of the bilinear form $a_h^s(\cdot,\cdot)$ in $\|\!|\mathbf{v}|\!\|_1$-norm.

Lemma 4.2. *Assume that* $\alpha_1 > 0$. *In addition, if* $a_h(\mathbf{v},\mathbf{v}) = a_h^s(\mathbf{v},\mathbf{v})$, *we assume that* α_1 *is sufficiently large. Then there exists a constant* α_0 *independent of* h, *such that*

$$a_h(\mathbf{v},\mathbf{v}) \ge \alpha_0 \|\!|\mathbf{v}|\!\|_1^2, \quad \forall \mathbf{v} \in V_h. \tag{4.21}$$

Proof. The inequality (4.21) follows immediately for $a_h(\cdot,\cdot) = a_h^{ns}(\cdot,\cdot)$, since

$$a_h^{ns}(\mathbf{v},\mathbf{v}) = (\nabla_{\mathcal{T}}\mathbf{v}, \nabla_{\mathcal{T}}\mathbf{v}) + \alpha_1 \sum_{e\in\mathcal{E}} h_e^{-1} \|[\![\mathbf{v}]\!]\|_e^2 \ge \alpha_0 \|\!|\mathbf{v}|\!\|_1^2 \tag{4.22}$$

with a constant $\alpha_0 = \min\{1, \alpha_1\}$.

Now we consider the case of symmetric form $a_h(\cdot,\cdot)$. It follows from the Cauchy-Schwarz inequality, (4.10), and (4.10) that for an edge (face) $e \in \mathcal{E}_0$ between the elements K_1 and K_2

$$\begin{aligned} \left|\int_e \{\nabla\mathbf{w}\} : [\![\mathbf{v}]\!] ds\right| &\le C\left(h_e\|\nabla\mathbf{w}|_{K_1}\|_e^2 + h_e\|\nabla\mathbf{w}|_{K_2}\|_e^2\right)^{\frac12} h_e^{-\frac12}\|[\![\mathbf{v}]\!]\|_e \\ &\le C\left(\|\nabla_{\mathcal{T}}\mathbf{w}\|_{K_1\cup K_2}^2 + h_e^2\|\nabla_{\mathcal{T}}^2\mathbf{w}\|_{K_1\cup K_2}^2\right)^{\frac12} h_e^{-\frac12}\|[\![\mathbf{v}]\!]\|_e \\ &\le C\|\nabla_{\mathcal{T}}\mathbf{w}\|_{K_1\cup K_2} h_e^{-\frac12}\|[\![\mathbf{v}]\!]\|_e. \end{aligned} \tag{4.23}$$

Similarly, if $e \subset \partial\Omega$ in and edge (face) of K then

$$\begin{aligned} \left|\int_e \{\nabla\mathbf{w}\} : [\![\mathbf{v}]\!] ds\right| &\le C\left(h_e\|\nabla\mathbf{w}|_K\|_e^2\right)^{\frac12} h_e^{-\frac12}\|[\![\mathbf{v}]\!]\|_e \\ &\le C\|\nabla\mathbf{w}\|_K h_e^{-\frac12}\|[\![\mathbf{v}]\!]\|_e. \end{aligned}$$

After summing over $K \in \mathcal{T}$ and taking into account the above inequalities we get

$$\begin{aligned} |\langle\{\nabla\mathbf{w}\} : [\![\mathbf{v}]\!]\rangle_{\mathcal{E}}| &\le C\|\nabla_{\mathcal{T}}\mathbf{w}\| \left(\sum_{e\in\mathcal{E}} h_e^{-1}\|[\![\mathbf{v}]\!]\|_e^2\right)^{\frac12} \\ &\le \frac14\|\nabla_{\mathcal{T}}\mathbf{w}\|^2 + C\sum_{e\in\mathcal{E}} h_e^{-1}\|[\![\mathbf{v}]\!]\|_e^2. \end{aligned}$$

Thus, we obtain

$$\begin{aligned} a_h^s(\mathbf{v},\mathbf{v}) &= (\nabla_{\mathcal{T}}\mathbf{v},\nabla_{\mathcal{T}}\mathbf{v}) + \alpha_1\sum_{e\in\mathcal{E}} h_e^{-1}\|[\![\mathbf{v}]\!]\|_e^2 - 2\sum_{e\in\mathcal{E}}\int_e \{\nabla\mathbf{v}\} : [\![\mathbf{v}]\!]ds \\ &\geq \|\nabla_{\mathcal{T}}\mathbf{v}\|^2 + \alpha_1\sum_{e\in\mathcal{E}} h_e^{-1}\|[\![\mathbf{v}]\!]\|_e^2 - \frac{1}{2}\|\nabla_{\mathcal{T}}\mathbf{v}\|^2 - C\sum_{e\in\mathcal{E}} h_e^{-1}\|[\![\mathbf{v}]\!]\|_e^2 \\ &= \frac{1}{2}\|\nabla_{\mathcal{T}}\mathbf{v}\|^2 + (\alpha_1 - C)\sum_{e\in\mathcal{E}} h_e^{-1}\|[\![\mathbf{v}]\!]\|_e^2 \geq \alpha_0 |\!|\!|\mathbf{v}|\!|\!|_1^2, \end{aligned}$$

with $\alpha_0 = \min(\frac{1}{2}, \alpha_1 - C)$. By choosing α_1 large enough such that $\alpha_1 \geq C + \frac{1}{2}$, we see that the estimate (4.21) holds true with $\alpha_0 = \frac{1}{2}$. □

Next we show that the bilinear forms $a_h^s(\cdot,\cdot)$ and $a_h^{ns}(\cdot,\cdot)$ are bounded in $|\!|\!|\cdot|\!|\!|_*$-norm.

Lemma 4.3. *There exists a constant β_a independent of h such that*

$$|a_h(\mathbf{w},\mathbf{v})| \leq \beta_a |\!|\!|\mathbf{w}|\!|\!|_* |\!|\!|\mathbf{v}|\!|\!|_* \ \forall \mathbf{w},\mathbf{v} \in V(h), \tag{4.24}$$

where $a_h(\cdot,\cdot) = a_h^s(\cdot,\cdot)$ or $a_h(\cdot,\cdot) = a_h^s(\cdot,\cdot)$.

Proof. By the definition of $a_h(\mathbf{w},\mathbf{v})$ and using (4.23) and Schwarz inequality, we see that there exists a constant C such that

$$\begin{aligned} |a_h(\mathbf{w},\mathbf{v})| \leq\ & C\Big\{\|\nabla_{\mathcal{T}}\mathbf{w}\|\|\nabla_{\mathcal{T}}\mathbf{v}\| \\ &+ \left(\|\nabla_{\mathcal{T}}\mathbf{w}\|^2 + \|h_{\mathcal{T}}\Delta_{\mathcal{T}}\mathbf{w}\|^2\right)^{\frac{1}{2}}\left(\sum_{e\in\mathcal{E}} h_e^{-1}\|[\![\mathbf{v}]\!]\|_e^2\right)^{\frac{1}{2}} \\ &+ \left(\|\nabla_{\mathcal{T}}\mathbf{v}\|^2 + \|h_{\mathcal{T}}\Delta_{\mathcal{T}}\mathbf{v}\|^2\right)^{\frac{1}{2}}\left(\sum_{e\in\mathcal{E}} h_e^{-1}\|[\![\mathbf{w}]\!]\|_e^2\right)^{\frac{1}{2}} \\ &+ \left(\sum_{e\in\mathcal{E}} h_e^{-1}\|[\![\mathbf{w}]\!]\|_e^2\right)^{\frac{1}{2}}\left(\sum_{e\in\mathcal{E}} h_e^{-1}\|[\![\mathbf{v}]\!]\|_e^2\right)^{\frac{1}{2}}\Big\} \\ \leq\ & \beta_a |\!|\!|\mathbf{w}|\!|\!|_* |\!|\!|\mathbf{v}|\!|\!|_*, \end{aligned}$$

which proves the desired boundness. □

The following lemma provides a upper bound for the bilinear form $b_h(\cdot,\cdot)$.

Lemma 4.4. *For $(\mathbf{v},q) \in V(h)\times Q(h)$*

$$b_h(\mathbf{v},q) \leq C|\!|\!|\mathbf{v}|\!|\!|_1\left(\|q\| + (\sum_{K\in\mathcal{T}} h_K^2|q|_{1,K}^2)^{\frac{1}{2}}\right). \tag{4.25}$$

Further, if $(\mathbf{v},q) \in V_h\times Q_h$ then

$$|b_h(\mathbf{v},q)| \leq C|\!|\!|\mathbf{v}|\!|\!|_1\|q\|.$$

Proof. By the definition of $b_h(\mathbf{v}, q)$, the Schwarz inequality, and (4.10) we have

$$\begin{aligned} |b_h(\mathbf{v}, q)| &\le C\left\{ \|\nabla_{\mathcal{T}}\mathbf{v}\|\|q\| + \left(\sum_{e\in\mathcal{E}} h_e\|q\|_e^2\right)^{\frac{1}{2}} \left(\sum_{e\in\mathcal{E}} h_e^{-1}\|[\![\mathbf{v}]\!]\|_e^2\right)^{\frac{1}{2}} \right\} \\ &\le C|\!|\!|\mathbf{v}|\!|\!|_1 \left(\|q\| + (\sum_{K\in\mathcal{T}_h} h_K^2|q|_{1,K}^2)^{\frac{1}{2}} \right). \end{aligned}$$

□

Finally, we prove

Lemma 4.5. *There is a constant β_A independent of h such that*

$$A(\mathbf{v}, q; \mathbf{v}, q) \le \beta_A(|\!|\!|\mathbf{v}|\!|\!|_1^2 + |\!|\!|q|\!|\!|_*^2) \;\forall\; (\mathbf{v}, q) \in V_h \times Q_h. \tag{4.26}$$

Proof. First, we note that $A(\mathbf{v}, q; \mathbf{v}, q) = a_h(\mathbf{v}, \mathbf{v}) + c_h(q, q)$. Then the inequality follows immediately from (4.24) with $\beta_A = \max\{1, \alpha_2\}$, if $a_h(\cdot,\cdot)$ is nonsymmetric and from (4.22) with $\beta_A = \max\{\beta_a, \alpha_2\}$, if $a_h(\cdot,\cdot)$ is symmetric. □

Proof. (of Theorem 4.1). Obviously the boundness of the form $A(\cdot;\cdot)$ follows immediately from Lemmas 4.3 and 4.4.

To prove the inequality (4.19) we shall use the fact that the differential problem (1.1) – (1.3) has unique solution that is stable in $H_0^1(\Omega)^d \times L_0^2(\Omega)$. For a given $q \in L_0^2(\Omega)$ let $\boldsymbol{\xi} \in H_0^1(\Omega)^d$ and $\theta \in L_0^2(\Omega)$ be the solution to the problem

$$-\Delta\boldsymbol{\xi} + \nabla\theta = 0,\; \nabla\cdot\boldsymbol{\xi} = -q,\; x \in \Omega.$$

The solution exists and satisfies the *a priori* estimate

$$\|\boldsymbol{\xi}\|_1 + \|\theta\| \le \beta\|q\|, \tag{4.27}$$

where the constant β depend only on the domain Ω. Take $\mathbf{w} = \Pi_1\boldsymbol{\xi} \in V_h$ and use (4.11) and (4.2) to get

$$\begin{aligned} |\!|\!|\Pi_1\boldsymbol{\xi}|\!|\!|_1^2 &= |\!|\!|\mathbf{w}|\!|\!|_1^2 = \|\nabla_{\mathcal{T}}\mathbf{w}\|^2 + \sum_{e\in\mathcal{E}} h_e^{-1}\|[\![\mathbf{w}]\!]\|_e^2 \\ &\le \|\nabla\Pi_1\boldsymbol{\xi}\|^2 + \sum_{e\in\mathcal{E}} h_e^{-1}\|[\![\boldsymbol{\xi} - \Pi_1\boldsymbol{\xi}]\!]\|_e^2 \le C\|\boldsymbol{\xi}\|_1^2. \end{aligned} \tag{4.28}$$

Next,

$$\begin{aligned} A(\mathbf{v}, q; \Pi_1\boldsymbol{\xi}, 0) &= a_h(\mathbf{v}, \Pi_1\boldsymbol{\xi}) + b_h(\Pi_1\boldsymbol{\xi}, q) \\ &= a_h(\mathbf{v}, \Pi_1\boldsymbol{\xi}) + b_h(\boldsymbol{\xi}, q) + b_h(\Pi_1\boldsymbol{\xi} - \boldsymbol{\xi}, q) \\ &= -(q, \nabla\cdot\boldsymbol{\xi}) + a_h(\mathbf{v}, \Pi_1\boldsymbol{\xi}) + b_h(\Pi_1\boldsymbol{\xi} - \boldsymbol{\xi}, q) \\ &\ge \|q\|^2 - C|\!|\!|\mathbf{v}|\!|\!|_1|\!|\!|\boldsymbol{\xi}|\!|\!|_1 + b_h(\Pi_1\boldsymbol{\xi} - \boldsymbol{\xi}, q) \\ &\ge \|q\|^2 - C_1|\!|\!|\mathbf{v}|\!|\!|_1\|q\| + b_h(\Pi_1\boldsymbol{\xi} - \boldsymbol{\xi}, q). \end{aligned} \tag{4.29}$$

Using integration by part, (4.6), (4.2), we transform the term $b_h(\Pi_1\boldsymbol{\xi}-\boldsymbol{\xi},q)$ as follows:

$$b_h(\Pi_1\boldsymbol{\xi}-\boldsymbol{\xi},q)=(\Pi_1\boldsymbol{\xi}-\boldsymbol{\xi},\nabla_{\mathcal{T}}q)-\sum_{K\in\mathcal{T}}\int_{\partial K}(\Pi_1\boldsymbol{\xi}-\boldsymbol{\xi})\cdot\mathbf{n}q\;ds+\langle\{q\},[(\Pi_1\boldsymbol{\xi}-\boldsymbol{\xi})\cdot\mathbf{n}]\rangle_{\mathcal{E}}.$$

Now, using the relation (4.6) and since $\Pi_1\boldsymbol{\xi}$ is a local orthogonal L^2-projection we get that $(\Pi_1\boldsymbol{\xi}-\boldsymbol{\xi},\nabla_{\mathcal{T}}q)=0$. Further, using the identity

$$\sum_{K\in\mathcal{T}}\int_{\partial K}q\mathbf{v}\cdot\mathbf{n}\;ds=\langle[q],\{\mathbf{v}\cdot\mathbf{n}\}\rangle_{\mathcal{E}_0}+\langle\{q\},[\mathbf{v}\cdot\mathbf{n}]\rangle_{\mathcal{E}} \tag{4.30}$$

we transform the last two terms to get

$$\begin{aligned}b_h(\Pi_1\boldsymbol{\xi}-\boldsymbol{\xi},q)&=-\langle[q],\{(\Pi_1\boldsymbol{\xi}-\boldsymbol{\xi})\cdot\mathbf{n}\}\rangle_{\mathcal{E}_0}\\&\geq-\left(\sum_{e\in\mathcal{E}_0}h_e^{-1}\|\{(\Pi_1\boldsymbol{\xi}-\boldsymbol{\xi})\cdot\mathbf{n}\}\|_e^2\right)^{\frac12}\left(\sum_{e\in\mathcal{E}_0}h_e\|[q]\|_e^2\right)^{\frac12}\\&\geq-C\|\boldsymbol{\xi}\|_1|\!|\!|q|\!|\!|_*\geq-C_2\|q\||\!|\!|q|\!|\!|_*\end{aligned} \tag{4.31}$$

Combining all these and choosing $\beta_1=\max\{C_1,\ C_2\}$ we get

$$A(\mathbf{v},q;\Pi_1\boldsymbol{\xi},0)\geq\frac12\|q\|^2-\beta_1(|\!|\!|\mathbf{v}|\!|\!|^2+|\!|\!|q|\!|\!|_*^2).$$

Let $\beta_2=\beta_A/(1+2\beta_1)>0$. Then

$$\begin{aligned}A(\mathbf{v},q;\mathbf{v}+2\beta_2\Pi_1\boldsymbol{\xi},q)&=A(\mathbf{v},q;\mathbf{v},q)+2\beta_2A(\mathbf{v},q;\Pi_1\boldsymbol{\xi},0)\\&\geq\beta_A(|\!|\!|\mathbf{v}|\!|\!|_1^2+|\!|\!|q|\!|\!|_*^2)+\beta_2\|q\|^2-2\beta_2\beta_1(|\!|\!|\mathbf{v}|\!|\!|_1^2+|\!|\!|q|\!|\!|_*^2)\\&=\beta_2(\|q\|^2+|\!|\!|\mathbf{v}|\!|\!|_1^2).\end{aligned}$$

The inequality (4.19) follows easily by using the stability of the projection $\Pi_1\boldsymbol{\xi}$ in $|\!|\!|\cdot|\!|\!|_1$, established in (4.28), and the *a priori* estimate (4.27)

$$\begin{aligned}\sup_{\mathbf{w}\in V_h,r\in Q_h}\frac{A(\mathbf{v},q;\mathbf{w},r)}{|\!|\!|\mathbf{w}|\!|\!|_1+\|r\|}&\geq\frac{A(\mathbf{v},q;\mathbf{v}+2\beta_2\Pi_1\boldsymbol{\xi},q)}{|\!|\!|\mathbf{v}+\beta_2\Pi_1\boldsymbol{\xi}|\!|\!|_1+\|q\|}\\&\geq\frac{\beta_2(\|q\|^2+|\!|\!|\mathbf{v}|\!|\!|_1^2)}{|\!|\!|\mathbf{v}|\!|\!|_1+\beta_2|\!|\!|\Pi_1\boldsymbol{\xi}|\!|\!|_1+\|q\|}\\&\geq\frac{\beta_2(\|q\|^2+|\!|\!|\mathbf{v}|\!|\!|_1^2)}{|\!|\!|\mathbf{v}|\!|\!|_1+C\|q\|}\\&\geq c_0(\|q\|+|\!|\!|\mathbf{v}|\!|\!|_1).\end{aligned}$$

This completes the proof of the Theorem 4.1. □

5. Error estimates

Here we establish an optimal error estimate for the finite element solution. First, we shall establish estimates for $|\!|\!|\mathbf{u}-\mathbf{u}_h|\!|\!|_1$ and $\|p-p_h\|$.

Theorem 5.1. *Let the assumptions 4.1 and 4.2 hold. If* $\mathbf{u} \in H^{l+1}(\Omega)^d$ *and* $p \in h^{m+1}(\Omega)$*, then*

$$\|p - p_h\| + |||\mathbf{u} - \mathbf{u}_h|||_* \le C(h^l \|\mathbf{u}\|_{l+1} + h^{m+1}\|p\|_{m+1}). \tag{5.1}$$

Proof. By first Strang's lemma we have

$$\begin{aligned}
\|p - p_h\| + |||\mathbf{u} - \mathbf{u}_h|||_* &\le \|p - \Pi_2 p\| + |||\mathbf{u} - \Pi_1 \mathbf{u}|||_* \\
&\quad + \frac{1}{c_0} \sup_{\mathbf{v}\in V_h, q \in Q_h} \frac{A(\mathbf{u}_h - \Pi_1\mathbf{u}, p_h - \Pi_2 p; \mathbf{v}, q)}{|||\mathbf{v}|||_1 + \|q\|} \\
&\le \|p - \Pi_2 p\| + |||\mathbf{u} - \Pi_1 \mathbf{u}|||_* \\
&\quad + \frac{1}{c_0} \sup_{\mathbf{v}\in V_h, q \in Q_h} \frac{A(\mathbf{u} - \Pi_1\mathbf{u}, p - \Pi_2 p; \mathbf{v}, q)}{|||\mathbf{v}|||_1 + \|q\|} \\
&\le C(\|p - \Pi_2 p\| + \textstyle\sum_{e\in\mathcal{E}_0} h_e \|[p - \Pi_2 p]\|_e^2 + |||\mathbf{u} - \Pi_1\mathbf{u}|||_*).
\end{aligned}$$

Here we have applied (4.19) and the Galerkin orthogonality (3.27). Then the result follows from the approximation properties of the projections Π_1 and Π_2 established in (4.1) – (4.4). □

The rest of the section is devoted to the error analysis in the L^2-norm for the velocity approximation. We shall use a standard argument by considering the dual problem which seeks $(\mathbf{w}; r) \in H_0^1(\Omega)^d \times L_0^2(\Omega)$ satisfying (in a weak sense)

$$-\Delta \mathbf{w} + \nabla r = \mathbf{u}_h - \mathbf{u}, \quad \nabla \cdot \mathbf{w} = 0 \quad \text{in } \Omega. \tag{5.2}$$

To gain an additional power in h for the error $\|\mathbf{u} - \mathbf{u}_h\|$ we need an assumption regarding the regularity of the solution of (5.2). Namely, we assume that (5.2) has full regularity in the sense that $(\mathbf{w}, r) \in H^2(\Omega)^d \times H^1(\Omega)$ and the following a priori estimate holds true:

$$\|\mathbf{w}\|_2 + \|r\|_1 \le C\|\mathbf{u} - \mathbf{u}_h\|.$$

This assumption is known to hold for Ω a convex polygonal domain in two dimensions. The situation in 3-D is more complicated (for more comments, see [19], p. 185 and the references cited there).

First notice that

$$\begin{aligned}
a_h^s(\mathbf{w},\ \mathbf{v}) &= a_h^s(\mathbf{v},\ \mathbf{w}) \\
a_h^{ns}(\mathbf{w},\ \mathbf{v}) &= a_h^{ns}(\mathbf{v},\ \mathbf{w}) + 2\langle\{\nabla \mathbf{v}\} : [\![\mathbf{w}]\!]\rangle_{\mathcal{E}} - 2\langle\{\nabla \mathbf{w}\} : [\![\mathbf{v}]\!]\rangle_{\mathcal{E}}.
\end{aligned}$$

Then if $a_h(\cdot,\cdot) = a_h^s(\cdot,\cdot)$, we have

$$A(\mathbf{v}, q;\ \mathbf{w}, r) = A(-\mathbf{w}, r;\ -\mathbf{v}, q). \tag{5.3}$$

If $a_h(\cdot,\cdot) = a_n^{ns}(\cdot,\cdot)$, we have

$$A(\mathbf{v}, q;\ \mathbf{w}, r) = A(-\mathbf{w}, r;\ -\mathbf{v}, q) + 2\langle\{\nabla \mathbf{w}\} : [\![\mathbf{v}]\!]\rangle_{\mathcal{E}} - 2\langle\{\nabla \mathbf{v}\} : [\![\mathbf{w}]\!]\rangle_{\mathcal{E}}. \tag{5.4}$$

It is not hard to see that for any $(\mathbf{v}, q) \in V(h) \times Q(h)$, the solution $(\mathbf{w}, r)$ satisfies

$$A(\mathbf{w}, r;\ \mathbf{v}, q) = (\mathbf{u}_h - \mathbf{u},\ \mathbf{v}). \tag{5.5}$$

We assume that the Stokes problem has full regularity in the sense that $(\mathbf{w}, r) \in H^2(\Omega)^d \times H^1(\Omega)$ and the following a priori estimate holds true:

$$\|\mathbf{w}\|_2 + \|r\|_1 \le C\|\mathbf{u}-\mathbf{u}_h\|. \tag{5.6}$$

With the above regularity, it is not hard to establish the following estimate:

$$|\!|\!|\mathbf{w}-\Pi_1\mathbf{w}|\!|\!|_* + \|r-\Pi_2 r\| \le Ch\|\mathbf{u}-\mathbf{u}_h\|. \tag{5.7}$$

Theorem 5.2. *Let $(\mathbf{u}_h,\ p_h) \in V_h \times Q_h$ and $(\mathbf{u},\ p) \in (H^{l+1}(\Omega)\cap H_0^1(\Omega))^d \times H^{m+1}(\Omega)\cap L_0^2(\Omega)$ be the solutions of (3.25) and (1.1)-(1.3), respectively. Then there exists a constant C independent of h such that if $a_h(\cdot,\cdot) = a_h^s(\cdot,\cdot)$, then*

$$\|\mathbf{u}-\mathbf{u}_h\| \le C\left(h^{l+1}\|\mathbf{u}\|_{l+1} + h^{m+2}\|p\|_{m+1}\right), \tag{5.8}$$

and if $a_h(\cdot,\cdot) = a_h^{ns}(\cdot,\cdot)$, then

$$\|\mathbf{u}-\mathbf{u}_h\| \le C\left(h^{l}\|\mathbf{u}\|_{l+1} + h^{m+1}\|p\|_{m+1}\right). \tag{5.9}$$

Proof. Let $\mathbf{v} = \mathbf{u}_h - \mathbf{u}$ and $q = p - p_h$ in (5.5). Using (5.4), (5.3), (3.27) and $[\mathbf{w}] = 0$ on $e \in \mathcal{E}$, we have

$$\begin{aligned}
\|\mathbf{u}-\mathbf{u}_h\|^2 &= A(\mathbf{w}, r;\ \mathbf{u}_h-\mathbf{u}, p-p_h)\\
&= A(\mathbf{u}-\mathbf{u}_h, p-p_h;\ -\mathbf{w}, r) + \theta\sum_{e\in\mathcal{E}}\int_e \{\nabla\mathbf{w}\} : [\![\mathbf{u}-\mathbf{u}_h]\!]ds \qquad (5.10)\\
&= A(\mathbf{u}-\mathbf{u}_h,\ p-p_h;\ -\mathbf{w}+\Pi_1\mathbf{w},\ r-\Pi_2 r) + \theta\sum_{e\in\mathcal{E}}\int_e \{\nabla\mathbf{w}\} : [\![\mathbf{u}-\mathbf{u}_h]\!]ds,
\end{aligned}$$

where $\theta = 0$ if $a_h(\cdot,\cdot) = a_h^s(\cdot,\cdot)$ and $\theta = 2$ if $a_h(\cdot,\cdot) = a_h^{ns}(\cdot,\cdot)$. We will obtain the bound for $\|\mathbf{u}-\mathbf{u}_h\|$ by estimating each term in the right hand side of (5.10). Using (4.24), (5.6) and (5.1) we get

$$\begin{aligned}
|a_h(\mathbf{u}-\mathbf{u}_h,\ \Pi_1\mathbf{w}-\mathbf{w})| &\le C|\!|\!|\mathbf{u}-\mathbf{u}_h|\!|\!|_*|\!|\!|\mathbf{w}-\Pi_1\mathbf{w}|\!|\!|_*\\
&\le Ch\left(h^l\|\mathbf{u}\|_{l+1} + h^{m+1}\|p\|_{m+1}\right)\|\mathbf{u}-\mathbf{u}_h\|.
\end{aligned}$$

Using integration by part, assumption (4.6), (5.1), (5.6) and (4.15), we have

$$\begin{aligned}
|b_h(\Pi_1\mathbf{w}-\mathbf{w},\ p-p_h)| &\le |\sum_K\int_K \nabla(p-p_h)\cdot(\Pi_1\mathbf{w}-\mathbf{w})dx|\\
&\quad +|\sum_{e\in\mathcal{E}_0}\int_e \{(\Pi_1\mathbf{w}-\mathbf{w})\cdot\mathbf{n}\}[p-p_h]ds|\\
&\le |\sum_{K\in\mathcal{T}}\int_K \nabla(p-\Pi_2 p)\cdot(\Pi_1\mathbf{w}-\mathbf{w})dx|\\
&\quad + \left(\sum_{e\in\mathcal{E}_0} h_e^{-1}\|\Pi_1\mathbf{w}-\mathbf{w}\|_e^2\right)^{\frac12}\left(\sum_{e\in\mathcal{E}_0} h_e\|[p-p_h]\|_e^2\right)^{\frac12}\\
&\le Ch\left(h^l|\mathbf{u}|_{l+1} + h^{m+1}\|p\|_{m+1}\right)\|\mathbf{u}-\mathbf{u}_h\|.
\end{aligned}$$

Then (4.25), (5.1) and (5.6) imply

$$\begin{aligned} |b_h(\mathbf{u}-\mathbf{u}_h,\ r-\Pi_2 r)| &= C|\!|\!|\mathbf{u}-\mathbf{u}_h|\!|\!|_* \left(\|r-\Pi_2 r\| + (\sum_{K\in\mathcal{T}} h_K^2 |r-\Pi_2 r|_{1,K}^2)^{\frac{1}{2}} \right) \\ &\leq Ch\left(h^l\|\mathbf{u}\|_{l+1} + h^{m+1}\|p\|_{m+1}\right)\|\mathbf{u}-\mathbf{u}_h\|. \end{aligned}$$

It follows from (5.1), (4.17) and (5.6)

$$\begin{aligned} |c_h(p-p_h,\ r-\Pi_2 r)| &\leq \gamma(p-p_h, r-\Pi_2 r) \\ &\quad +\alpha_2 \left(\sum_{e\in\mathcal{E}_0} h_e \|[p-p_h]\|_e^2\right)^{\frac{1}{2}} \left(\sum_{e\in\mathcal{E}_0} h_e \|[r-\Pi_2 r]\|_e^2\right)^{\frac{1}{2}} \\ &\leq Ch(h^l\|\mathbf{u}\|_{l+1} + h^{m+1}\|p\|_{m+1})\|\mathbf{u}-\mathbf{u}_h\|. \end{aligned}$$

Schwarz inequality, (4.10), (5.1) and (5.6) imply

$$\begin{aligned} \sum_{e\in\mathcal{E}} \int_e \{\nabla \mathbf{w}\} : [\![\mathbf{u}-\mathbf{u}_h]\!] ds &\leq C \left(\sum_{e\in\mathcal{E}} h_e \|\nabla \mathbf{w}\|_e^2\right)^{\frac{1}{2}} \left(\sum_{e\in\mathcal{E}} h_e^{-1} \|[\![\mathbf{u}-\mathbf{u}_h]\!]\|_e^2\right)^{\frac{1}{2}} \\ &\leq C(h^l\|\mathbf{u}\|_{l+1} + h^{m+1}\|p\|_{m+1})\|\mathbf{u}-\mathbf{u}_h\|. \end{aligned}$$

Combining above estimates and (5.10), we prove the theorem. □

Acknowledgment

The authors thank Dr. Jean-Luc Guermond for his valuable comments during the early stage of this work that led to improving the presentation.

References

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

[2] D. Arnold, F. Brezzi and M. Fortin, A stable finite element for the Stokes equations, Calcolo, 21 (1984), 337–344.

[3] I. Babuška and M. Zlamal, Nonconforming elements in the finite element method with penalty, SIAM J. Numer. Anal. 10 (1973), 863–875.

[4] 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), 267–279.

[5] C.E. Baumann and J.T. Oden, A discontinuous *hp* finite element method for convection- diffusion problems, Comput. Methods Appl. Mech. Engrg. 175 (1999), 311–341.

[6] F. Ben Belgacem, The mortar finite element method with Lagrange multipliers, SIAM J. Numer. Anal. 84 (1999), 173–197.

[7] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, New York: Springer 1991.

[8] F. Brezzi, G. Manzini, D. Marini, P. Pietra and A. Russo, Discontinuous Galerkin approximations for elliptic problems, Numerical Methods for Partial Differential Equations, 16 (2000), 365-378.

[9] F. Brezzi and J. Pitkäranta, On the stabilization of finite element approximations of the Stokes equations, in *Efficient Solution of Elliptic System, volume 10, Notes on Numerical Fluid Mechanics*, W. Hackbusch ed., Vieweg & Sohn, Braunschweig, (1984), 11–19.

[10] P.G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, New York, 1978.

[11] B. Cockburn, G. Kanschat, D. Schötzau and C. Schwab, Local discontinuous Galerkin methods for the Stokes system, SIAM J. Numer. Anal. 40 (2002), 319–343.

[12] B. Cockburn, G. Kanschat, and D. Schötzau, The local discontinuous Galerkin methods for incompressible fluid flow: A review, Computer and Fluids (Special Issue: Residual based methods and discontinuous Galerkin schemes)", 34 (2005), 491–506.

[13] B. Cockburn, D. Schötzau and J. Wang, Discontinuous Galerkin methods for incompressible elastic materials, (to appear in Comput. Methods Appl. Mech. Engrg.)

[14] B. Cockburn, G.E. Karniadakis, and C.W. Shu (eds.): *The Discontinuous Galerkin Methods: Theory, Computation and Applications*, Lecture Notes in Computational Science and Engineering, Vol. 11, Springer-Verlag, 2000.

[15] B. Cockburn and C.W. Shu, The local discontinuous Galerkin finite element method for convection-diffusion systems, SIAM J. Numer. Anal. 35 (1998), 2440–2463.

[16] M. Crouzeix and P. A. Raviart, Conforming and nonconforming finite element methods for solving the stationary Stokes equations, R.A.I.R.O. R3, pp. 33-76 (1973), 33–76.

[17] J. Douglas, Jr. and T. Dupont, Interior penalty procedures for elliptic and parabolic Galerkin methods, Lecture Notes in Physics, Vol. 58, Springer-Verlag, Berlin, 1976.

[18] J. Douglas, Jr. and J. Wang, An absolutely stabilized finite element method for the Stokes problem, Mathematics of Computation, 52 (1989) 495-508.

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

[20] V. Girault and P.A. Raviart, *Finite Element Methods for the Navier-Stokes Equations: Theory and Algorithms*, Springer, Berlin, 1986.

[21] V. Girault, B. Rivière, and M.F. Wheeler, A discontinuous Gelerkin method with nonconforming domain decomposition for Stokes and Navier-Stokes problems, Math. Comp., 74 (249) (2004), 53–84.

[22] M.D. Gunzburger, *Finite Element Methods for Viscous Incompressible Flows, A Guide to Theory, Practice and Algorithms*, Academic Press, San Diego, (1989).

[23] P. Hansbo and M.G. Larson, Discontinuous Galerkin methods for incompressible and nearly incompressible elasticity by Nitsche's method, Comput. Meth. Appl. Mech. Engng., 191 (2002) 1895–1908.

[24] P. Hansbo and M.G. Larson, Discontinuous Galerkin and the Crouzeix-Raviart element: Application to elasticity, M2AN Math. Model. Numer. Anal. 37 (2003) 63–72.

[25] P. Hood and C. Taylor, Numerical solution of the Navier-Stokes equations using the finite element technique, Comput. Fluids, 1 (1973) 1–28.

[26] T. Hughes, L. Franca and M. Balestra, A new finite element formulation for computational fluid dynamics: V. Circumventing the Babuška-Brezzi condition: A stable Petrov-Galerkin formulation of the Stokes problem accommodating equal-order interpolation, Comput. Meth. Appl. Mech. Engng., 59 (1986) 85–99.

[27] T. Hughes and L. Franca, A new finite element formulation for computational fluid dynamics: VII. The Stokes problem with various well-posed boundary conditions: symmetric formulations that converge for all velocity/pressure spaces, Comput. Meth. Appl. Mech. Engng., 65 (1987) 85–96.

[28] N. Kechkar and D. Silvester, Analysis of local stabilized mixed finite element methods for the Stokes problem, Mathematics of Computation, 58 (1992), 1–10.

[29] O.A. Ladyzhenskaya, *The Mathematical Theory of Viscous Incompressible Flows*, Gordon and Breach, London, 1969.

[30] A. Quarteroni and X. Valli, *Numerical Approximation of Partial Differential Equations*, Springer Series in Computational Mathematics, v. 23, 1994.

[31] J. Wang, and X. Ye, A new finite element method for Stokes equation by $H(div;\Omega)$ elements, Technical Report 99-09, TICAM, 1999.

Department of Mathematics, Texas A& M University, College Station, TX 77843-3368
E-mail address: lazarov@math.tamu.edu

Department of Mathematics and Statistics, University of Arkansas at Little Rock, Little Rock, AR 72204
E-mail address: xxye@ualr.edu