Penalty-Free Discontinuous Galerkin Methods for Incompressible Navier-Stokes Equations

Beatrice Riviere · **Shirin Sardar**

Abstract A first-order discontinuous Galerkin method is proposed for solving the steady-state incompressible Navier-Stokes equations. The stability of this penalty-free method is obtained by locally enriching the discrete space with a quadratic polynomial. A priori error estimates are derived. Numerical examples confirm the theoretical convergence.

Keywords incompressible flow · stabilized DG · inf-sup · error estimates

1 Introduction

Various discontinuous Galerkin (DG) methods have been proposed for solving the steady-state incompressible Navier-Stokes equations. However, very few of those methods have been theoretically analyzed. The interior-penalty DG method is analyzed in [11, 19] where the velocity (respectively the pressure) is approximated by fully discontinuous polynomials of degree k (respectively $k-1$). The earlier work [15] analyzes a scheme where the velocity is approximated by pointwise divergence-free discontinuous polynomials and the pressure by continuous polynomials. The local discontinuous Galerkin method has been theoretically investigated in the paper [8], and in a following note [9] that results in a pointwise divergence-free velocity. All these DG methods contain a stabilization term, also called a penalty term, that helps enforcing the weak continuity of the numerical solution. For symmetric discretization of the elliptic operator, the penalty term involves a penalty parameter that is required to be

B. Riviere
6100 Main Street MS 134
Houston, TX 77005
E-mail: riviere@caam.rice.edu

S. Sardar
6100 Main Street MS 134
Houston, TX 77005
E-mail: ss30@rice.edu

large enough in order to have coercivity. In a non-exhaustive list, we point to some papers that formulate various discontinuous Galerkin methods for solving the incompressible steady-state Navier-Stokes equations, but do not address the theoretical convergence: [17, 14, 2]. For the linear problem of Stokes, there is a large literature on the analysis of DG methods: see for instance [13, 22, 21, 7, 11] and the references therein.

All the methods mentioned above include a stabilizing penalty term. The use of the penalty term is crucial in the stability and convergence analysis for most of the DG methods applied to elliptic problems. Removing the penalty term in the non-symmetric interior penalty DG method yields the Oden-Babuška-Baumann method [18], the analysis of which is found in [20] for polynomials of degree at least equal to two. Recently, penalty-free discontinuous Galerkin methods were introduced in [4, 1] as an answer to the interesting theoretical question of proving stability and convergence of the Oden-Babuška-Baumann method in any dimension for polynomial degree equal to one. An inf-sup condition can be proved if the discrete space of discontinuous piecewise linears is enriched with a quadratic polynomial on each mesh element. Using the same type of enrichment, the symmetric penalty-free DG method was then analyzed for elliptic problems in [5], and for the Stokes problem in [6].

In this paper, we apply the enriched low-order penalty-free DG method to the nonlinear Navier-Stokes equations. The approximation of the velocity is in the locally enriched space, and the approximation of the pressure is piecewise constant. Special care is needed to handle the nonlinear convection term. As a consequence of the scheme, the numerical solution satisfies a local momentum equation on each mesh element, that is penalty-free. This paper considers both the symmetric and non-symmetric discretization of the elliptic operator, and we point out in the paper the differences in the analysis between the two.

The outline of the paper is as follows. In the next section, the Navier-Stokes equations and the penalty-free DG scheme are stated. Section 3 focuses on proving existence of the numerical solution. The argument is based on showing existence of a fixed point of a particular map G, so that the fixed-point is the numerical solution. One key ingredient is an inf-sup condition. Section 4 derives an optimal error estimate. The proof of the inf-sup condition is given in Section 5. Continuity of the map G is obtained in Section 6. Finally, numerical examples confirm the theoretical convergence rates.

2 Model Problem and Scheme

We define the Navier-Stokes equations on a convex polygon $\Omega \subset \mathbb{R}^d$ with $d = 2, 3$. Given $\boldsymbol{f}$ in $\boldsymbol{L}^2(\Omega)$ find $(\boldsymbol{u}, p)$ in $\boldsymbol{H}_0^1(\Omega) \times L_0^2(\Omega)$, such that

$$-\mu \Delta \boldsymbol{u} + \boldsymbol{u} \cdot \nabla \boldsymbol{u} + \nabla p = \boldsymbol{f}, \quad \text{in } \Omega, \tag{1}$$

$$\nabla \cdot \boldsymbol{u} = 0, \quad \text{in } \Omega, \tag{2}$$

$$\boldsymbol{u} = 0, \quad \text{on } \partial\Omega. \tag{3}$$

The fluid viscosity μ is a constant $0 < \mu \leqslant 1$ and $\boldsymbol{f}$ is a force acting on the fluid. We first define the discrete spaces. Let $\mathcal{T}_h$ be a shape-regular subdivision of Ω. Let $\mathcal{F}_i$ be the set of interior faces and let $\mathcal{F}_e$ be the set of boundary faces. Also denote $\mathscr{F}_h = \mathcal{F}_i \cup \mathcal{F}_e$ the set of all faces. For an element τ in $\mathcal{T}_h$, the parameter h_τ denotes its diameter and for a face F in $\mathscr{F}_h$, h_F denotes the diameter of F. Set $h = \max_{\tau\in\mathcal{T}_h} h_\tau$ and let $\tilde{h}$ be the function such that $\tilde{h}|_\tau = h_\tau$ and $\tilde{h}|_F = h_F$ for all τ in $\mathcal{T}_h$ and F in $\mathscr{F}_h$. To an interior face F of $\mathcal{F}_i$ shared by two elements τ_F^1 and τ_F^2, we associate a unit normal vector $\boldsymbol{n}_F$ directed from τ_F^1 to τ_F^2 and we define the jump and average of a function v (scalar or vector) by:

$$[v] = (v|_{\tau_F^1}) - (v|_{\tau_F^2}), \quad \{v\} = \frac{1}{2}(v|_{\tau_F^1}) + \frac{1}{2}(v|_{\tau_F^2}).$$

As usual, the choice of direction of $\boldsymbol{n}_F$ is obtained by ordering the elements of $\mathcal{T}_h$. To simplify the writing, we also denote by $\bar{\boldsymbol{v}}$ the spatial average of a function $\boldsymbol{v}$ on elements or faces:

$$\bar{\boldsymbol{v}}|_F = \frac{1}{|F|}\int_F \boldsymbol{v}, \quad \bar{\boldsymbol{v}}|_\tau = \frac{1}{|\tau|}\int_\tau \boldsymbol{v}.$$

From the approximation properties, there is a positive constant C_A independent of h such that

$$\|v - \bar{v}\|_\tau \leqslant C_A h_\tau \|\nabla v\|_\tau, \quad \forall \tau \in \mathcal{T}_h. \tag{4}$$

The L^2 norm on a domain $\mathcal{O}$ is denoted as usual by $\|\cdot\|_{\mathcal{O}}$. Similarly, the L^2 inner-product on elements and faces is denoted by $(\cdot,\cdot)_{\mathcal{T}_h}$ and $(\cdot,\cdot)_{\mathscr{F}_h}$. We will also use the notation:

$$\|\cdot\|_{\mathcal{T}_h} = \left(\sum_{\tau\in\mathcal{T}_h} \|\cdot\|_\tau^2\right)^{1/2}, \quad \|\cdot\|_{\mathscr{F}_t} = \left(\sum_{F\in\mathscr{F}_t} \|\cdot\|_F^2\right)^{1/2}, \quad t = i, h.$$

The velocity space is the piecewise linear space enriched with one quadratic function. Define

$$\boldsymbol{V}_h = \{\boldsymbol{v} \in L^2(\Omega)^d, \quad \boldsymbol{v}|_\tau \in (\mathbb{P}_1(\tau))^d + \boldsymbol{\alpha}\phi, \quad \forall \boldsymbol{\alpha} \in \mathbb{R}^d\},$$

where

$$\phi(\boldsymbol{x}) = \sum_{i=1}^d x_i^2, \quad \forall \boldsymbol{x} = (x_1, \dots, x_d).$$

The space $\boldsymbol{V}_h$ is equipped with the norm:

$$\|\boldsymbol{v}_h\|_{\boldsymbol{V}} = \left(\|\nabla \boldsymbol{v}_h\|_{\mathcal{T}_h}^2 + \|\tilde{h}^{-\frac{1}{2}}[\boldsymbol{v}_h]\|_{\mathscr{F}_h}^2\right)^{1/2}.$$

The pressure space $Q_h \subset L_0^2(\Omega)$ is the space of discontinuous piecewise constants and it is equipped with the norm:

$$\|q_h\|_Q = \|\tilde{h}^{\frac{1}{2}}[q_h]\|_{\mathscr{F}_i}.$$

In the text, we also use the notation:

$$\|(\boldsymbol{v}_h, q_h)\|_{\boldsymbol{V}\times Q} = \left(\|\boldsymbol{v}_h\|_{\boldsymbol{V}}^2 + \|q_h\|_Q^2\right)^{1/2}.$$

In the analysis below we will use an equivalent norm to $\|\cdot\|_{\boldsymbol{V}}$ (see [6]): there exists a constant $C_E > 0$ independent of h such that

$$C_E\|\boldsymbol{v}_h\|_{\boldsymbol{V}} \leqslant \left(\|\nabla \boldsymbol{v}_h\|_{\mathcal{T}_h}^2 + ||\tilde{h}^{-\frac{1}{2}}\overline{[\boldsymbol{v}_h]}||_{\mathscr{F}_h}^2\right)^{1/2} \leqslant \|\boldsymbol{v}_h\|_{\boldsymbol{V}}, \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h. \tag{5}$$

Given the symmetrization parameter ϵ in $\{-1, +1\}$, we define two bilinear forms $a(\cdot,\cdot)$ and $b(\cdot,\cdot)$, that are the DG discretizations of the diffusion and pressure terms respectively. For all $\boldsymbol{u}_h, \boldsymbol{v}_h$ in $\boldsymbol{V}_h$, for all q_h in Q_h, we define

$$a(\boldsymbol{u}_h, \boldsymbol{v}_h) = (\nabla \boldsymbol{u}_h, \nabla \boldsymbol{v}_h)_{\mathcal{T}_h} - (\{\nabla \boldsymbol{u}_h\}\boldsymbol{n}_F, [\boldsymbol{v}_h])_{\mathscr{F}_h} + \epsilon([\boldsymbol{u}_h], \{\nabla \boldsymbol{v}_h\}\boldsymbol{n}_F)_{\mathscr{F}_h}, \tag{6}$$

$$b(q_h, \boldsymbol{v}_h) = -(q_h, \nabla\cdot \boldsymbol{v}_h)_{\mathcal{T}_h} + (\{q_h\}, [\boldsymbol{v}_h]\cdot \boldsymbol{n}_F)_{\mathscr{F}_h}. \tag{7}$$

To discretize the nonlinear convection term we introduce the notation z^{int} respectively z^{ext} to denote the restriction of z to the element τ, respectively the neighboring element to τ. The vector $\boldsymbol{n}_\tau$ is the unit normal outward to τ.

$$c(\boldsymbol{z}_h, \boldsymbol{u}_h, \boldsymbol{v}_h, \boldsymbol{w}_h) = (\boldsymbol{u}_h \cdot \nabla \boldsymbol{v}_h, \boldsymbol{w}_h)_{\mathcal{T}_h} + \frac{1}{2}((\nabla\cdot \boldsymbol{u}_h)\boldsymbol{v}_h, \boldsymbol{w}_h)_{\mathcal{T}_h} - \frac{1}{2}([\boldsymbol{u}_h], \{\boldsymbol{v}_h\cdot \boldsymbol{w}_h\})_{\mathscr{F}_h}$$
$$+ \sum_{\tau\in\mathcal{T}_h}\int_{\partial\tau_-^{\boldsymbol{z}_h}} |\{\boldsymbol{u}_h\}\cdot \boldsymbol{n}_\tau|(\boldsymbol{v}_h^{int} - \boldsymbol{v}_h^{ext})\cdot \boldsymbol{w}_h^{int}, \quad \forall \boldsymbol{z}_h, \boldsymbol{u}_h, \boldsymbol{v}_h, \boldsymbol{w}_h \in \boldsymbol{V}_h, \tag{8}$$

where

$$\partial\tau_-^{\boldsymbol{w}_h} = \{\boldsymbol{x}\in\partial\tau \ : \ \{\boldsymbol{w}_h\}\cdot \boldsymbol{n}_\tau < 0\}. \tag{9}$$

The penalty-free DG method is defined as follows: find $(\boldsymbol{u}_h, p_h) \in \boldsymbol{V}_h \times Q_h$ such that

$$\mu a(\boldsymbol{u}_h, \boldsymbol{v}_h) + b(p_h, \boldsymbol{v}_h) - b(q_h, \boldsymbol{u}_h) + c(\boldsymbol{u}_h, \boldsymbol{u}_h, \boldsymbol{u}_h, \boldsymbol{v}_h) = \ell(\boldsymbol{v}_h, q_h). \tag{10}$$

for all $(\boldsymbol{v}_h, q_h)$ in $\boldsymbol{V}_h \times Q_h$. The form $\ell(\cdot,\cdot)$ is simply $\ell(\boldsymbol{v}_h, q_h) = (\boldsymbol{f}, \boldsymbol{v}_h)_{\mathcal{T}_h}$. The form c has been previously studied in [11,19,12]. We recall the following bounds (see Lemma 6.1 in [11] and Proposition 4.1 in [12]).

Lemma 1 *There is a constant C_0 independent of h such that*

$$|c(\boldsymbol{z}_h, \boldsymbol{u}_h, \boldsymbol{v}_h, \boldsymbol{w}_h)| \leqslant C_0\|\boldsymbol{u}_h\|_{\boldsymbol{V}}\,\|\boldsymbol{v}_h\|_{\boldsymbol{V}}\,\|\boldsymbol{w}_h\|_{\boldsymbol{V}}, \quad \forall \boldsymbol{z}_h, \boldsymbol{u}_h, \boldsymbol{v}_h, \boldsymbol{w}_h \in \boldsymbol{V}_h, \tag{11}$$

$$c(\boldsymbol{v}_h, \boldsymbol{v}_h, \boldsymbol{w}_h, \boldsymbol{w}_h) \geqslant 0 \quad \forall \boldsymbol{v}_h, \boldsymbol{w}_h \in \boldsymbol{V}_h. \tag{12}$$

Problem (10) is nonlinear and the next section proves existence of a solution by applying the Leray-Schauder theorem. We conclude this section by recalling some inverse and trace inequalities needed in the analysis. There exist constants C_I and C_T independent of h such that:

$$\|\tilde{h}\Delta \boldsymbol{v}_h\|_{\mathcal{T}_h} \leqslant C_I\|\nabla \boldsymbol{v}_h\|_{\mathcal{T}_h}, \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{13}$$

$$\|\tilde{h}^{-1/2}[\boldsymbol{v}_h]\|_{\mathscr{F}_h} \leqslant C_T\|\tilde{h}^{-1}\boldsymbol{v}_h\|_{\mathcal{T}_h}, \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{14}$$

$$\|\tilde{h}^{1/2}\{\nabla \boldsymbol{v}_h\}\boldsymbol{n}_F\|_{\mathscr{F}_h} \leqslant C_T\|\nabla \boldsymbol{v}_h\|_{\mathcal{T}_h}, \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{15}$$

$$\|\tilde{h}^{1/2}[\nabla \boldsymbol{v}_h]\boldsymbol{n}_F\|_{\mathscr{F}_i} \leqslant C_T\|\nabla \boldsymbol{v}_h\|_{\mathcal{T}_h}, \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h. \tag{16}$$

3 Existence and Uniqueness of Numerical Solution

In this part we prove the existence of a numerical solution. We construct a map G by linearizing the convection term in problem (10) and show existence of a fixed point of G. Define G as follows:

$$G : \boldsymbol{V}_h \times Q_h \to \boldsymbol{V}_h \times Q_h$$
$$(\tilde{\boldsymbol{u}}_h, \tilde{p}_h) \mapsto (\boldsymbol{u}_h, p_h)$$

where $(\boldsymbol{u}_h, p_h)$ satisfies for all $(\boldsymbol{v}_h, q_h)$ in $\boldsymbol{V}_h \times Q_h$:

$$\mu a(\boldsymbol{u}_h, \boldsymbol{v}_h) + b(p_h, \boldsymbol{v}_h) - b(q_h, \boldsymbol{u}_h) + c(\tilde{\boldsymbol{u}}_h, \tilde{\boldsymbol{u}}_h, \boldsymbol{u}_h, \boldsymbol{v}_h) = \ell(\boldsymbol{v}_h, q_h). \tag{17}$$

We observe that a fixed point of the map G is a solution of (10). We first need to check that the map G is well-defined, i.e. that there exists a unique solution $(\boldsymbol{u}_h, p_h)$ to problem (17). By denoting the left-hand side of (17) by $\mathcal{S}_{\tilde{\boldsymbol{u}}_h}((\boldsymbol{u}_h, p_h), (\boldsymbol{v}_h, q_h))$, we rewrite problem (17) as:

$$\mathcal{S}_{\tilde{\boldsymbol{u}}_h}((\boldsymbol{u}_h, p_h), (\boldsymbol{v}_h, q_h)) = \ell(\boldsymbol{v}_h, q_h), \quad \forall (\boldsymbol{v}_h, q_h) \in \boldsymbol{V}_h \times Q_h. \tag{18}$$

One critical part of the proof is an inf-sup condition for the form $\mathcal{S}_{\tilde{\boldsymbol{u}}_h}$. The inf-sup result holds if $\tilde{\boldsymbol{u}}_h$ is in a ball of radius $\mathcal{C}$ defined below:

$$\mathcal{C} = \begin{cases} \mu \frac{C_E}{4C_0}(\max(C_I^2 + 1, C_T^2, C_T^2(\frac{1}{4} + C_A^2)))^{-1/2} & \text{for } \epsilon = 1 \\ \mu \frac{C_E}{4C_0(1+4C_T^2)}(\max(C_I^2 + 1, C_T^2, C_T^2(\frac{1}{4} + C_A^2)))^{-1/2} & \text{for } \epsilon = -1 \end{cases} \tag{19}$$

We now state the inf-sup condition and postpone its proof to Section 5.

Proposition 1 *Assume $\tilde{\boldsymbol{u}}_h$ satisfies the following bound:*

$$\|\tilde{\boldsymbol{u}}_h\|_{\boldsymbol{V}} \leqslant \mathcal{C}, \tag{20}$$

where $\mathcal{C}$ is defined in (19). There exists a constant $\beta > 0$ independent of h and of $\tilde{\boldsymbol{u}}_h$ such that for all $(\boldsymbol{u}_h, p_h)$ in $\boldsymbol{V}_h \times Q_h$ there holds

$$\beta \|(\boldsymbol{u}_h, p_h)\|_{\boldsymbol{V}\times Q} \leqslant \sup_{(\boldsymbol{v}_h, q_h) \in \boldsymbol{V}_h \times Q_h} \frac{\mathcal{S}_{\tilde{\boldsymbol{u}}_h}((\boldsymbol{u}_h, p_h), (\boldsymbol{v}_h, q_h))}{\|(\boldsymbol{v}_h, q_h)\|_{\boldsymbol{V}\times Q}}. \tag{21}$$

One can easily check that $\mathcal{S}_{\tilde{\boldsymbol{u}}_h}$ is continuous and that ℓ is continuous. In fact,

$$|\ell(\boldsymbol{v}_h, q_h)\| \leqslant C_P \|\boldsymbol{f}\|_{L^2(\Omega)} \|\boldsymbol{v}_h\|_{\boldsymbol{V}}, \quad \forall (\boldsymbol{v}_h, q_h) \in \boldsymbol{V}_h \times Q_h,$$

where C_P is the constant in the generalized Poincaré inequality valid for discontinuous piecewise polynomials (see for instance Lemma 6.2 in [11] or [3]). Therefore, by Lax-Milgram's theorem, there exists a unique $(\boldsymbol{u}_h, p_h)$ in $\boldsymbol{V}_h \times Q_h$ satisfying (18). In addition we have the following bound for $(\boldsymbol{u}_h, p_h)$:

$$\|(\boldsymbol{u}_h, p_h)\|_{\boldsymbol{V}\times Q} \leqslant \frac{C_P}{\beta} \|\boldsymbol{f}\|_{L^2(\Omega)}. \tag{22}$$

Under the small data assumption

$$\|\boldsymbol{f}\|_{L^2(\Omega)} \leqslant \frac{\beta \mathcal{C}}{C_P}, \tag{23}$$

we then conclude that the solution $\boldsymbol{u}_h$ also satisfies the bound

$$\|\boldsymbol{u}_h\|_{\boldsymbol{V}} \leqslant \mathcal{C}.$$

Define the space

$$\boldsymbol{X}_h = \{\boldsymbol{v}_h \in \boldsymbol{V}_h : \quad \|\boldsymbol{v}_h\|_{\boldsymbol{V}} \leqslant \mathcal{C}\}.$$

Thus we showed that the restriction of the map G to the space $\boldsymbol{X}_h \times Q_h$ is a well-defined map onto $\boldsymbol{X}_h \times Q_h$. Second, in order to apply the Leray-Schauder theorem, we need to show that the map G is continuous. The proof of the following proposition is technical and postponed to Section 6.

Proposition 2 *Let $(\tilde{\boldsymbol{u}}_n)_{n\geqslant 0}$ be a sequence in $\boldsymbol{X}_h$ that converges to $\tilde{\boldsymbol{u}}$ in $\boldsymbol{X}_h$. Let $(\tilde{p}_n)_{n\geqslant 0}$ be a sequence in Q_h that converges to $\tilde{p}$ in Q_h. Define $(\boldsymbol{u}_n, p_n) = G(\tilde{\boldsymbol{u}}_n, \tilde{p}_n)$ and $(\boldsymbol{u}, p) = G(\tilde{\boldsymbol{u}}, \tilde{p})$. Then, the sequence $(\boldsymbol{u}_n)_{n\geqslant 0}$ converges to $\boldsymbol{u}$ in $\boldsymbol{X}_h$ and the sequence $(p_n)_{n\geqslant 0}$ converges to p in Q_h.*

Moreover, one can check that G maps bounded sets to bounded sets, and this implies that G is compact. The last assumption in the Leray-Schauder theorem is to show that the set

$$Z = \{(\tilde{\boldsymbol{u}}_h, \tilde{p}_h) \in \boldsymbol{X}_h \times Q_h : \ (\tilde{\boldsymbol{u}}_h, \tilde{p}_h) = \lambda G(\tilde{\boldsymbol{u}}_h, \tilde{p}_h) \text{ for some } 0 \leqslant \lambda \leqslant 1\}$$

is bounded, which is trivial. Therefore we have shown that there exists a fixed point of the map G restricted to $\boldsymbol{X}_h \times Q_h$. We summarize the result in the following theorem.

Theorem 1 *Under the small data assumption (23) there exists $(\boldsymbol{u}_h, p_h)$ in $\boldsymbol{X}_h \times Q_h$ that satisfies*

$$\mu a(\boldsymbol{u}_h, \boldsymbol{v}_h) + b(p_h, \boldsymbol{v}_h) - b(q_h, \boldsymbol{u}_h) + c(\boldsymbol{u}_h, \boldsymbol{u}_h, \boldsymbol{u}_h, \boldsymbol{v}_h) = \ell(\boldsymbol{v}_h, q_h), \forall (\boldsymbol{v}_h, q_h) \in \boldsymbol{V}_h \times Q_h$$

Remark 1 After some algebraic manipulation, since $\mathcal{C}$ and β depend on μ, we can rewrite (23) under the form:

$$\|\boldsymbol{f}\|_{L^2(\Omega)} \leqslant M_1 \frac{\mu^2}{\mu + M_2}, \tag{24}$$

for some constants M_1 and M_2 independent of μ.

We finish this section by proving a local uniqueness result. The proof uses a technical property of the nonlinear part of the c form (see Proposition 4.10 in [10]). There is a positive constant C_1 independent of h such that for any $\boldsymbol{z}_h, \boldsymbol{w}_h, \boldsymbol{v}_h$ in $\boldsymbol{V}_h$, we have

$$|c_{n\ell}(\boldsymbol{z}_h, \boldsymbol{w}_h, \boldsymbol{z}_h, \boldsymbol{v}_h) - c_{n\ell}(\boldsymbol{w}_h, \boldsymbol{w}_h, \boldsymbol{z}_h, \boldsymbol{v}_h)| \leqslant C_1 \|\boldsymbol{z}_h - \boldsymbol{w}_h\|_{\boldsymbol{V}} \|\boldsymbol{v}_h\|_{\boldsymbol{V}} \|\boldsymbol{z}_h\|_{\boldsymbol{V}}, \quad \forall \boldsymbol{z}_h, \boldsymbol{w}_h, \boldsymbol{v}_h \in \boldsymbol{V}_h$$

Theorem 2 *Problem 10 has at most one solution $(\boldsymbol{u}_h, p_h)$ in $\boldsymbol{X}_h \times Q_h$ such that*

$$\|\boldsymbol{u}_h\|_{\boldsymbol{V}} < (C_0 + C_1)\frac{\beta}{\sqrt{2}}. \tag{25}$$

Proof Assume that $(\boldsymbol{u}_1, p_1)$ and $(\boldsymbol{u}_2, p_2)$ are two solutions of Problem (10) and denote $\boldsymbol{w} = \boldsymbol{u}_1 - \boldsymbol{u}_2$ and $r = p_1 - p_2$. Then, we have for all $(\boldsymbol{v}_h, q_h)$ in $\boldsymbol{V}_h \times Q_h$

$$\mu a(\boldsymbol{w}, \boldsymbol{v}_h) + b(r, \boldsymbol{v}_h) - b(q_h, \boldsymbol{w}) + c(\boldsymbol{u}_1, \boldsymbol{u}_1, \boldsymbol{u}_1, \boldsymbol{v}_h) - c(\boldsymbol{u}_2, \boldsymbol{u}_2, \boldsymbol{u}_2, \boldsymbol{v}_h) = 0. \tag{26}$$

We can write:

$$\begin{aligned} c(\boldsymbol{u}_1, \boldsymbol{u}_1, \boldsymbol{u}_1, \boldsymbol{v}_h) - c(\boldsymbol{u}_2, \boldsymbol{u}_2, \boldsymbol{u}_2, \boldsymbol{v}_h) &= c(\boldsymbol{u}_2, \boldsymbol{u}_2, \boldsymbol{w}, \boldsymbol{v}_h) + c(\boldsymbol{u}_1, \boldsymbol{w}, \boldsymbol{u}_1, \boldsymbol{v}_h) \\ &+ c_{n\ell}(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_1, \boldsymbol{v}_h) - c_{n\ell}(\boldsymbol{u}_2, \boldsymbol{u}_2, \boldsymbol{u}_1, \boldsymbol{v}_h). \end{aligned}$$

Using the inf-sup condition (21), we have:

$$\begin{aligned} \beta\|(\boldsymbol{w}, r)\|_{\boldsymbol{V}\times Q} &\leqslant \sup_{(\boldsymbol{v}_h, q_h)\in \boldsymbol{V}_h\times Q_h} \frac{\mu a(\boldsymbol{w}, \boldsymbol{v}_h) + b(r, \boldsymbol{v}_h) - b(q_h, \boldsymbol{w}) + c(\boldsymbol{u}_2, \boldsymbol{u}_2, \boldsymbol{w}, \boldsymbol{v}_h)}{\|(\boldsymbol{v}_h, q_h)\|_{\boldsymbol{V}\times Q}} \\ &\leqslant \sup_{(\boldsymbol{v}_h, q_h)\in \boldsymbol{V}_h\times Q_h} \frac{|c(\boldsymbol{u}_1, \boldsymbol{w}, \boldsymbol{u}_1, \boldsymbol{v}_h)| + |c_{n\ell}(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_1, \boldsymbol{v}_h) - c_{n\ell}(\boldsymbol{u}_2, \boldsymbol{u}_2, \boldsymbol{u}_1, \boldsymbol{v}_h)|}{\|(\boldsymbol{v}_h, q_h)\|_{\boldsymbol{V}\times Q}}. \end{aligned}$$

Using the continuity property (11) and the bound (26), we have

$$\beta\|(\boldsymbol{w}, r)\|_{\boldsymbol{V}\times Q} \leqslant (C_0 + C_1)\|\boldsymbol{u}_1\|_{\boldsymbol{V}}\|\boldsymbol{w}\|_{\boldsymbol{V}},$$

and with assumption (25), this implies that $\boldsymbol{w} = \boldsymbol{0}$ and $r = 0$.

In the next section, optimal error estimates are obtained.

4 Error Estimates

In this section, error estimates are derived. We recall an approximation operator $R_h : \boldsymbol{H}_0^1(\Omega) \mapsto \boldsymbol{V}_h$ satisfying:

$$b(q_h, R_h(\boldsymbol{v}) - \boldsymbol{v}) = 0, \quad \forall \boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega), \quad \forall q_h \in Q_h, \tag{27}$$

$$\|R_h(\boldsymbol{v}) - \boldsymbol{v}\|_{\boldsymbol{V}} \leqslant C_R h^{s-1}|\boldsymbol{v}|_{H^s(\Omega)}, \quad \forall s \in [1, 2], \quad \forall \boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega) \cap \boldsymbol{H}^s(\Omega). \tag{28}$$

More details on the construction of the operator R_h can be found in [21,10]. The following theorem establishes an a priori error estimate for the numerical solution.

Theorem 3 *Assume there is a solution $(\boldsymbol{u}, p)$ to (1)-(3) such that $\boldsymbol{u} \in \boldsymbol{H}^2(\Omega)$, $p \in H^1(\Omega)$ and the following bound holds:*

$$|\boldsymbol{u}|_{\boldsymbol{H}^1(\Omega)} \leqslant \frac{\beta}{2\sqrt{2}C_0 C_R}. \tag{29}$$

Then, there exists a constant M independent of h, $\boldsymbol{u}$ and p such that

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{\boldsymbol{V}} \leqslant Mh\left(|\boldsymbol{u}|_{\boldsymbol{H}^2(\Omega)} + \|p\|_{H^1(\Omega)}\right).$$

Proof We decompose the errors into a numerical error and an approximation error. Denote $\pi p \in Q_h$ an optimal approximation of p; for instance, πp could be an interpolant or a L^2 projection. Define

$$\boldsymbol{\xi_u} = \boldsymbol{u}_h - R_h(\boldsymbol{u}), \quad \xi_p = p_h - \pi p, \quad \eta_u = \boldsymbol{u} - R_h(\boldsymbol{u}), \quad \eta_p = p - \pi p$$

The error equation is:

$$\mu a(\boldsymbol{\xi_u}, \boldsymbol{v}_h) + b(\xi_p, \boldsymbol{v}_h) - b(q_h, \boldsymbol{\xi_u}) + c(\boldsymbol{u}_h, \boldsymbol{u}_h, \boldsymbol{u}_h, \boldsymbol{v}_h) - c(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{u}, \boldsymbol{v}_h)$$
$$= \mu a(\boldsymbol{\eta_u}, \boldsymbol{v}_h) + b(\eta_p, \boldsymbol{v}_h) - b(q_h, \boldsymbol{\eta_u}).$$

Using the fact that $\boldsymbol{u}$ is in $\boldsymbol{H}_0^1(\Omega)$, we write:

$$c(\boldsymbol{u}_h, \boldsymbol{u}_h, \boldsymbol{u}_h, \boldsymbol{v}_h) - c(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{u}, \boldsymbol{v}_h) = c(\boldsymbol{u}_h, \boldsymbol{u}_h, \boldsymbol{u}_h, \boldsymbol{v}_h) - c(\boldsymbol{u}_h, \boldsymbol{u}, \boldsymbol{u}, \boldsymbol{v}_h)$$
$$= c(\boldsymbol{u}_h, \boldsymbol{u}_h, \boldsymbol{\xi_u}, \boldsymbol{v}_h) + c(\boldsymbol{u}_h, \boldsymbol{\xi_u}, R_h(\boldsymbol{u}), \boldsymbol{v}_h) - c(\boldsymbol{u}_h, \boldsymbol{\eta_u}, R_h(\boldsymbol{u}), \boldsymbol{v}_h) - c(\boldsymbol{u}_h, \boldsymbol{u}, \boldsymbol{\eta_u}, \boldsymbol{v}_h).$$

Using the inf-sup condition (21), we obtain:

$$\beta \|(\boldsymbol{\xi_u}, \xi_p)\|_{\boldsymbol{V}\times Q} \leqslant \sup_{(\boldsymbol{v}_h, q_h)\in \boldsymbol{V}_h \times Q_h} \frac{1}{\|(\boldsymbol{v}_h, q_h)\|_{\boldsymbol{V}\times Q}} \big(\mu a(\boldsymbol{\eta_u}, \boldsymbol{v}_h) + b(\eta_p, \boldsymbol{v}_h)$$
$$-b(q_h, \boldsymbol{\eta_u}) - c(\boldsymbol{u}_h, \boldsymbol{\xi_u}, R_h(\boldsymbol{u}), \boldsymbol{v}_h) + c(\boldsymbol{u}_h, \boldsymbol{\eta_u}, R_h(\boldsymbol{u}), \boldsymbol{v}_h) + c(\boldsymbol{u}_h, \boldsymbol{u}, \boldsymbol{\eta_u}, \boldsymbol{v}_h)\big).$$

The term $b(q_h, \boldsymbol{\eta_u})$ vanishes due to (27). From (28), we have

$$\|R_h(\boldsymbol{u})\|_{\boldsymbol{V}} \leqslant C_R |\boldsymbol{u}|_{\boldsymbol{H}^1(\Omega)}$$

With the continuity of c (property (11)), we then obtain

$$c(\boldsymbol{u}_h, \boldsymbol{\xi_u}, R_h(\boldsymbol{u}), \boldsymbol{v}_h) \leqslant C_0 C_R \|\boldsymbol{\xi_u}\|_{\boldsymbol{V}} |\boldsymbol{u}|_{\boldsymbol{H}^1(\Omega)} \|\boldsymbol{v}_h\|_{\boldsymbol{V}}.$$

Using arguments similar to those found in [11], we obtain the following bound:

$$\mu a(\boldsymbol{\eta_u}, \boldsymbol{v}_h) + b(\eta_p, \boldsymbol{v}_h) + c(\boldsymbol{u}_h, \boldsymbol{\eta_u}, R_h(\boldsymbol{u}), \boldsymbol{v}_h) + c(\boldsymbol{u}_h, \boldsymbol{u}, \boldsymbol{\eta_u}, \boldsymbol{v}_h)$$
$$\leqslant Mh(|\boldsymbol{u}|_{\boldsymbol{H}^2(\Omega)} + |p|_{H^1(\Omega)}) \|\boldsymbol{v}_h\|_{\boldsymbol{V}},$$

where M is a constant independent of h, $\boldsymbol{u}$ and p but depends on μ. Combining the bounds above yields

$$\beta \|(\boldsymbol{\xi_u}, \xi_p)\|_{\boldsymbol{V}\times Q} \leqslant Mh(|\boldsymbol{u}|_{\boldsymbol{H}^2(\Omega)} + |p|_{H^1(\Omega)}) + C_0 C_R \|\boldsymbol{\xi_u}\|_{\boldsymbol{V}} |\boldsymbol{u}|_{\boldsymbol{H}^1(\Omega)}.$$

Using assumption (29), we conclude:

$$\frac{\beta}{2\sqrt{2}} \|\boldsymbol{\xi_u}\|_{\boldsymbol{V}} + \beta \|\xi_p\|_Q \leqslant Mh(|\boldsymbol{u}|_{\boldsymbol{H}^2(\Omega)} + |p|_{H^1(\Omega)}).$$

The final result is obtained by using a triangle inequality and the optimal approximations of $R_h(\boldsymbol{u})$ and πp.

5 An inf-sup condition

In this section, we give the proof of Proposition 1. We fix $\tilde{\boldsymbol{u}}_h$ satisfying (20). For a given pair $(\boldsymbol{u}_h, p_h)$, we show that there exist constants β_1 and β_2 independent of h and $\tilde{\boldsymbol{u}}_h$ and that there exist functions $\boldsymbol{v}_h$ and q_h satisfying:

$$\beta_1 \|(\boldsymbol{u}_h, p_h)\|^2_{\boldsymbol{V}\times Q} \leqslant \mathcal{S}_{\tilde{\boldsymbol{u}}_h}((\boldsymbol{u}_h, p_h), (\boldsymbol{v}_h, q_h)), \tag{30}$$

$$\|(\boldsymbol{v}_h, q_h)\|_{\boldsymbol{V}\times Q} \leqslant \beta_2 \|(\boldsymbol{u}_h, p_h)\|_{\boldsymbol{V}\times Q}. \tag{31}$$

The inf-sup condition then follows with $\beta = \beta_1/\beta_2$. We will choose $q_h = p_h$. For the choice of the velocity $\boldsymbol{v}_h$, a linear combination of special functions is chosen. We will make use of a projection result proved in [6].

Lemma 2 *Let $\boldsymbol{a}_h$ be a piecewise constant vector on the mesh elements and let $\boldsymbol{b}_h, \boldsymbol{c}_h$ be piecewise constant vectors on the mesh faces. There exists a unique function $\boldsymbol{\phi}_h \in \boldsymbol{V}_h$ such that*

$$\frac{1}{|\tau|}\int_\tau \boldsymbol{\phi}_h = \boldsymbol{a}_h|_\tau, \quad \forall \tau \in \mathcal{T}_h,$$
$$\{\nabla \boldsymbol{\phi}_h\}|_F \boldsymbol{n}_F = \boldsymbol{b}_h|_F, \quad \forall F \in \mathscr{F}_h,$$
$$\frac{1}{|F|}\int_F \{\boldsymbol{\phi}_h\} = \boldsymbol{c}_h|_F, \quad \forall F \in \mathcal{F}_i.$$

Moreover, $\boldsymbol{\phi}_h$ satisfies the following a priori estimate for a constant M_P independent of h

$$\|\tilde{h}^{-\frac{1}{2}}\boldsymbol{\phi}_h\|^2_{\mathcal{T}_h} + \|\boldsymbol{\phi}_h\|^2_{\boldsymbol{V}} \leqslant M_P\left(\|\tilde{h}^{-1}\boldsymbol{a}_h\|^2_{\mathcal{T}_h} + \|\tilde{h}^{\frac{1}{2}}\boldsymbol{b}_h\|^2_{\mathscr{F}_h} + \|\tilde{h}^{-\frac{1}{2}}\boldsymbol{c}_h\|^2_{\mathcal{F}_i}\right).$$

We note that the constant M_P introduced below only depends on inverse, trace and approximation inequalities. Indeed, one can show that

$$M_P = 4\max\left(C_I^2 + 1, C_T^2, C_T^2(\frac{1}{4} + C_A^2)\right).$$

We now construct two special functions $\boldsymbol{w}_h$ and $\boldsymbol{z}_h$ in $\boldsymbol{V}_h$.

Lemma 3 *Let $\boldsymbol{w}_h \in \boldsymbol{V}_h$ be the projection defined in Lemma 2 with the arguments*

$$\boldsymbol{a}_h = \boldsymbol{0}, \quad \boldsymbol{b}_h|_F = \epsilon\widetilde{h}^{-1}[\overline{\boldsymbol{u}_h}]|_F, \quad \boldsymbol{c}_h = \boldsymbol{0}, \tag{32}$$

and let $\boldsymbol{z}_h$ be the projection defined by the arguments

$$\boldsymbol{a}_h = \boldsymbol{0}, \quad \boldsymbol{b}_h = \boldsymbol{0}, \quad \boldsymbol{c}_h|_F = -\tilde{h}[p_h]|_F. \tag{33}$$

Then, the following equalities and inequalities hold

$$c(\tilde{\boldsymbol{u}}_h, \tilde{\boldsymbol{u}}_h, \boldsymbol{u}_h, \boldsymbol{w}_h) \geqslant -N_0 \|\tilde{\boldsymbol{u}}_h\|_{\boldsymbol{V}} (||\nabla \boldsymbol{u}_h||^2_{\mathcal{T}_h} + ||\tilde{h}^{-\frac{1}{2}} [\overline{\boldsymbol{u}_h}]||^2_{\mathscr{F}_h}), \quad (34)$$

$$c(\tilde{\boldsymbol{u}}_h, \tilde{\boldsymbol{u}}_h, \boldsymbol{u}_h, \boldsymbol{z}_h) \geqslant -\frac{N_1}{2} \|\tilde{\boldsymbol{u}}_h\|^2_{\boldsymbol{V}} (||\nabla \boldsymbol{u}_h||^2_{\mathcal{T}_h} + ||\tilde{h}^{-\frac{1}{2}} [\overline{\boldsymbol{u}_h}]||^2_{\mathscr{F}_h}) - \frac{1}{2} \|\tilde{h}^{\frac{1}{2}} [p_h]\|^2_{\mathscr{F}_i}, \quad (35)$$

$$a(\boldsymbol{u}_h, \boldsymbol{w}_h) = \|\tilde{h}^{-\frac{1}{2}} [\overline{\boldsymbol{u}_h}]\|^2_{\mathscr{F}_h}, \quad (36)$$

$$b(p_h, \boldsymbol{w}_h) = 0, \quad (37)$$

$$a(\boldsymbol{u}_h, \boldsymbol{z}_h) \geqslant -N_2 ||\nabla \boldsymbol{u}_h||^2_{\mathcal{T}_h} - \frac{1}{4} \|\tilde{h}^{1/2} [p_h]\|^2_{\mathscr{F}_i}, \quad (38)$$

$$b(p_h, \boldsymbol{z}_h) = \|\tilde{h}^{1/2} [p_h]\|^2_{\mathscr{F}_i}. \quad (39)$$

The constants N_0, N_1 and N_2 are independent of h and $\tilde{\boldsymbol{u}}_h$. In fact, $N_0 = C_0 M_P^{1/2}/C_E$, $N_1 = C_0^2 M_P/(2C_E^2)$ and $N_2 = C_T^2$.

Proof By Lemma 2, we have

$$\|\boldsymbol{w}_h\|_{\boldsymbol{V}} \leqslant M_P^{1/2} ||\tilde{h}^{-\frac{1}{2}} [\overline{\boldsymbol{u}_h}]||_{\mathscr{F}_h}, \quad \|\boldsymbol{z}_h\|_{\boldsymbol{V}} \leqslant M_P^{1/2} ||\tilde{h}^{\frac{1}{2}} [p_h]||_{\mathscr{F}_i}. \quad (40)$$

These bounds are combined with (11) to yield

$$c(\tilde{\boldsymbol{u}}_h, \tilde{\boldsymbol{u}}_h, \boldsymbol{u}_h, \boldsymbol{w}_h) \leqslant C_0 M_P^{1/2} \|\tilde{\boldsymbol{u}}_h\|_{\boldsymbol{V}} \|\boldsymbol{u}_h\|_{\boldsymbol{V}} ||\tilde{h}^{-\frac{1}{2}} [\overline{\boldsymbol{u}_h}]||_{\mathscr{F}_h}, \quad (41)$$

$$c(\tilde{\boldsymbol{u}}_h, \tilde{\boldsymbol{u}}_h, \boldsymbol{u}_h, \boldsymbol{z}_h) \leqslant C_0 M_P^{1/2} \|\tilde{\boldsymbol{u}}_h\|_{\boldsymbol{V}} \|\boldsymbol{u}_h\|_{\boldsymbol{V}} ||\tilde{h}^{\frac{1}{2}} [p_h]||_{\mathscr{F}_i}. \quad (42)$$

Using the equivalence of norms (5) in (41) we obtain (34) with $N_0 = C_0 M_P^{1/2}/C_E$. Similarly we obtain (35) with $N_1 = C_0^2 M_P/(2C_E^2)$. Next, we rewrite the form $a(\cdot,\cdot)$ by integrating by parts.

$$a(\boldsymbol{u}_h, \boldsymbol{w}_h) = -(\Delta \boldsymbol{u}_h, \boldsymbol{w}_h)_{\mathcal{T}_h} + ([\nabla \boldsymbol{u}_h] \boldsymbol{n}_F, \{\boldsymbol{w}\})_{\mathscr{F}_i} + \epsilon([\boldsymbol{u}_h], \{\nabla \boldsymbol{w}_h\} \boldsymbol{n}_F)_{\mathscr{F}_h}.$$

For any function $\boldsymbol{\chi}$ in $\boldsymbol{V}_h$, it is easy to check that $\Delta \boldsymbol{\chi}$ is a piecewise constant vector on each mesh element, and that $\nabla \boldsymbol{\chi} \boldsymbol{n}_F$ is a piecewise constant vector on each mesh face. Indeed for the latter, if one writes $\boldsymbol{\chi} = (\chi_1, \ldots, \chi_d)$ then, one can check that $\nabla \chi_i$ belongs to the Raviart-Thomas space RT_0, and therefore $\nabla \chi_i \cdot \boldsymbol{n}_F$ is constant on each face. Thus we can write

$$a(\boldsymbol{u}_h, \boldsymbol{w}_h) = -(\Delta \boldsymbol{u}_h, \overline{\boldsymbol{w}_h})_{\mathcal{T}_h} + ([\nabla \boldsymbol{u}_h] \boldsymbol{n}_F, \overline{\{\boldsymbol{w}\}})_{\mathscr{F}_i} + \epsilon(\overline{[\boldsymbol{u}_h]}, \{\nabla \boldsymbol{w}_h\} \boldsymbol{n}_F)_{\mathscr{F}_h}.$$

By construction of $\boldsymbol{w}_h$, this simplifies to

$$a(\boldsymbol{u}_h, \boldsymbol{w}_h) = (\overline{[\boldsymbol{u}_h]}, \tilde{h}^{-1} \overline{[\boldsymbol{u}_h]})_{\mathscr{F}_h},$$

which is (36). Moreover by integration by parts, we obtain:

$$b(p_h, \boldsymbol{w}_h) = (\nabla p_h, \boldsymbol{w}_h)_{\mathcal{T}_h} - ([p_h], \{\boldsymbol{w}_h \cdot \boldsymbol{n}_F\})_{\mathcal{F}_i}.$$

Since p_h is piecewise constant, we can write:

$$b(p_h, \boldsymbol{w}_h) = -([p_h], \{\boldsymbol{w}_h \cdot \boldsymbol{n}_F\})_{\mathcal{F}_i} = -([p_h], \overline{\{\boldsymbol{w}_h\}} \cdot \boldsymbol{n}_F)_{\mathcal{F}_i},$$

which is zero by construction of $\boldsymbol{w}_h$. Similarly, we have

$$b(p_h, \boldsymbol{z}_h) = -([p_h], \overline{\{\boldsymbol{z}_h\}} \cdot \boldsymbol{n}_F)_{\mathcal{F}_i} = \|\tilde{h}^{1/2}[p_h]\|^2_{\mathscr{F}_i},$$

by construction of $\boldsymbol{z}_h$. It remains to bound $a(\boldsymbol{u}_h, \boldsymbol{z}_h)$. We follow the argument above and obtain:

$$a(\boldsymbol{u}_h, \boldsymbol{z}_h) = -(\Delta \boldsymbol{u}_h, \overline{\boldsymbol{z}_h})_{\mathcal{T}_h} + ([\nabla \boldsymbol{u}_h]\boldsymbol{n}_F, \overline{\{\boldsymbol{z}\}})_{\mathscr{F}_i} + \epsilon(\overline{[\boldsymbol{u}_h]}, \{\nabla \boldsymbol{z}_h\}\boldsymbol{n}_F)_{\mathscr{F}_h}.$$

By construction of $\boldsymbol{z}_h$, this simplifies to:

$$a(\boldsymbol{u}_h, \boldsymbol{z}_h) = -([\nabla \boldsymbol{u}_h]\boldsymbol{n}_F, \tilde{h}[p_h])_{\mathscr{F}_i}.$$

Using Cauchy-Schwarz's inequality and the trace inequality (16), we obtain:

$$a(\boldsymbol{u}_h, \boldsymbol{z}_h) \leqslant C_T \|\nabla \boldsymbol{u}_h\|_{\mathcal{T}_h} \|\tilde{h}^{1/2}[p_h]\|_{\mathscr{F}_i}.$$

The final result (38) is obtained by applying Young's inequality.

We now select the test function $\boldsymbol{v}_h$ as:

$$\boldsymbol{v}_h = A\boldsymbol{u}_h + B\boldsymbol{w}_h + 2\boldsymbol{z}_h,$$

with different constants A, B depending on the sign of the DG parameter ϵ. In the case where $\epsilon = 1$, the constants are

$$A = \frac{1}{\mu}\left(1 + 2\mu N_2 + N_1\|\tilde{\boldsymbol{u}}_h\|^2_{\boldsymbol{V}} + BN_0\|\tilde{\boldsymbol{u}}_h\|_{\boldsymbol{V}}\right),$$
$$B = \frac{1 + N_1\|\tilde{\boldsymbol{u}}_h\|^2_{\boldsymbol{V}}}{\mu - N_0\|\tilde{\boldsymbol{u}}_h\|_{\boldsymbol{V}}}.$$

In the case where $\epsilon = -1$, the constants are:

$$A = \frac{1}{\mu}\left(2 + 4\mu N_2 + 2N_1\|\tilde{\boldsymbol{u}}_h\|^2_{\boldsymbol{V}} + 2BN_0\|\tilde{\boldsymbol{u}}_h\|_{\boldsymbol{V}}\right),$$
$$B = \frac{1 + 4C_T^2(1 + 2\mu N_2) + (1 + 4C_T^2)N_1\|\tilde{\boldsymbol{u}}_h\|^2_{\boldsymbol{V}}}{\mu - (1 + 4C_T^2)N_0\|\tilde{\boldsymbol{u}}_h\|_{\boldsymbol{V}}}.$$

With this choice of $\boldsymbol{v}_h$ and the choice $q_h = p_h$, we compute:

$$\begin{aligned}\mathcal{S}_{\tilde{\boldsymbol{u}}_h}((\boldsymbol{u}_h, p_h), (\boldsymbol{v}_h, q_h)) &= \mu A a(\boldsymbol{u}_h, \boldsymbol{u}_h) + \mu B a(\boldsymbol{u}_h, \boldsymbol{w}_h) + 2\mu a(\boldsymbol{u}_h, \boldsymbol{z}_h) \\ &+ Ab(p_h, \boldsymbol{u}_h) + Bb(p_h, \boldsymbol{w}_h) + 2b(p_h, \boldsymbol{z}_h) - b(p_h, \boldsymbol{u}_h) + Ac(\widetilde{\boldsymbol{u}_h}, \widetilde{\boldsymbol{u}_h}, \boldsymbol{u}_h, \boldsymbol{u}_h) \\ &+ Bc(\widetilde{\boldsymbol{u}_h}, \widetilde{\boldsymbol{u}_h}, \boldsymbol{u}_h, \boldsymbol{w}_h) + 2c(\widetilde{\boldsymbol{u}_h}, \widetilde{\boldsymbol{u}_h}, \boldsymbol{u}_h, \boldsymbol{z}_h).\end{aligned}$$

We now note that $b(p_h, \boldsymbol{u}_h) = 0$. For the case $\epsilon = 1$, we also have: $a(\boldsymbol{u}_h, \boldsymbol{u}_h) = \|\boldsymbol{u}_h\|^2_{\mathcal{T}_h}$. In the case of $\epsilon = -1$, we have:

$$a(\boldsymbol{u}_h, \boldsymbol{u}_h) = \|\boldsymbol{u}_h\|^2_{\mathcal{T}_h} - 2(\{\nabla \boldsymbol{u}_h\}\boldsymbol{n}_F, [\boldsymbol{u}_h])_{\mathscr{F}_h}.$$

Using Cauchy-Schwarz's inequality and the trace inequality (15), we obtain:

$$\begin{aligned} 2(\{\nabla \boldsymbol{u}_h\}\boldsymbol{n}_F, [\boldsymbol{u}_h])_{\mathscr{F}_h} &\leqslant 2\|\tilde{h}^{-1/2}\overline{[\boldsymbol{u}_h]}\|_{\mathscr{F}_h}\|\tilde{h}^{1/2}\{\nabla \boldsymbol{u}_h\}\boldsymbol{n}_F\|_{\mathscr{F}_h} \\ &\leqslant 2C_T\|\tilde{h}^{-1/2}\overline{[\boldsymbol{u}_h]}\|_{\mathscr{F}_h}\|\nabla \boldsymbol{u}_h\|_{\mathcal{T}_h} \\ &\leqslant \frac{1}{2}\|\nabla \boldsymbol{u}_h\|^2_{\mathcal{T}_h} + 2C_T^2\|\tilde{h}^{-1/2}\overline{[\boldsymbol{u}_h]}\|^2_{\mathscr{F}_h}. \end{aligned}$$

Thus, we have

$$a(\boldsymbol{u}_h, \boldsymbol{u}_h) \geqslant \frac{1}{2}\|\nabla \boldsymbol{u}_h\|^2_{\mathcal{T}_h} - 2C_T^2\|\tilde{h}^{-1/2}\overline{[\boldsymbol{u}_h]}\|^2_{\mathscr{F}_h}.$$

Using the positivity result (12), Lemma 3, and the lower bounds for $a(\boldsymbol{u}_h, \boldsymbol{u}_h)$, the expression above simplifies to:

$$\mathcal{S}_{\tilde{\boldsymbol{u}}_h}((\boldsymbol{u}_h, p_h), (\boldsymbol{v}_h, q_h)) \geqslant \frac{1}{2}\left(\|\nabla \boldsymbol{u}_h\|^2_{\mathcal{T}_h} + \|\tilde{h}^{-1/2}\overline{[\boldsymbol{u}_h]}\|^2_{\mathscr{F}_h} + \|\tilde{h}^{1/2}[p_h]\|^2_{\mathscr{F}_i}\right).$$

This implies that $\beta_1 = 1/2$ in (30). It remains to check that the constants A and B are positive and bounded, and that $\boldsymbol{v}_h$ is bounded. First, in the non-symmetric case ($\epsilon = 1$), condition (20) is equivalent to

$$\|\tilde{\boldsymbol{u}}_h\| \leqslant \frac{\mu}{2N_0},$$

which implies that $\mu^{-1} \leqslant B \leqslant 2\mu^{-1}(1 + N_1\mathcal{C}^2)$. This means that A is positive, and bounded above independent of h and $\tilde{\boldsymbol{u}}_h$. Second, in the symmetric case ($\epsilon = -1$), condition (20) is equivalent to

$$\|\tilde{\boldsymbol{u}}_h\| \leqslant \frac{\mu}{2N_0(1 + 4C_T^2)},$$

which implies that $\mu^{-1} \leqslant B \leqslant 2\mu^{-1}(1 + 4C_T^2(1 + 2\mu N_2) + (1 + 4C_T^2)N_1\mathcal{C}^2)$. This also means that A is positive and bounded above independent of h and $\tilde{\boldsymbol{u}}_h$. We now find the constant β_2 in (31). By construction of $\boldsymbol{v}_h$, we have

$$\|\boldsymbol{v}_h\|_{\boldsymbol{V}} \leqslant A\|\boldsymbol{u}_h\|_{\boldsymbol{V}} + B\|\boldsymbol{w}_h\|_{\boldsymbol{V}} + 2\|\boldsymbol{z}_h\|_{\boldsymbol{V}}$$

Using (40), this yields

$$\|\boldsymbol{v}_h\|_{\boldsymbol{V}} \leqslant (A + BM_P^{1/2})\|\boldsymbol{u}_h\|_{\boldsymbol{V}} + 2M_P^{1/2}\|p_h\|_{L^2(\Omega)}.$$

Therefore there exists a constant β_2 independent of h such that (31) holds. In fact,

$$\beta_2 = \sqrt{2}\max(A + BM_P^{1/2}, 2M_P^{1/2}).$$

Since the constants A, B and M_P are independent of h and of $\tilde{\boldsymbol{u}}_h$, we can conclude.

6 Continuity of map G

In this section we prove that G is continuous. To simplify the writing, we drop the subscript h in what follows. Let $(\tilde{\boldsymbol{u}}_n)_{n\geqslant 0}$ be a sequence in $\boldsymbol{X}_h$ that converges to $\tilde{\boldsymbol{u}}$ in $\boldsymbol{X}_h$. Let $(\tilde{p}_n)_{n\geqslant 0}$ be a sequence in Q_h that converges to $\tilde{p}$ in Q_h. Define $(\boldsymbol{u}_n, p_n) = G(\tilde{\boldsymbol{u}}_n, \tilde{p}_n)$ and $(\boldsymbol{u}, p) = G(\tilde{\boldsymbol{u}}, \tilde{p})$. We will show that the sequence $(\boldsymbol{u}_n)_{n\geqslant 0}$ converges to $\boldsymbol{u}$ in $\boldsymbol{X}_h$ and the sequence $(p_n)_{n\geqslant 0}$ converges to p in Q_h. By definition we have for all $(\boldsymbol{v}, q)$ in $\boldsymbol{V}_h \times Q_h$.

$$\mu a(\boldsymbol{u}-\boldsymbol{u}_n, \boldsymbol{v}) + b(p-p_n, \boldsymbol{v}) - b(q, \boldsymbol{u}-\boldsymbol{u}_n) + c(\tilde{\boldsymbol{u}}, \tilde{\boldsymbol{u}}, \boldsymbol{u}, \boldsymbol{v}) - c(\tilde{\boldsymbol{u}}_n, \tilde{\boldsymbol{u}}_n, \boldsymbol{u}, \boldsymbol{v}) = 0. \tag{43}$$

We decompose the form c into a trilinear part and a nonlinear part. We write

$$c(\boldsymbol{z}_h, \boldsymbol{u}_h, \boldsymbol{v}_h, \boldsymbol{w}_h) = c_\ell(\boldsymbol{u}_h, \boldsymbol{v}_h, \boldsymbol{w}_h) + c_{n\ell}(\boldsymbol{z}_h, \boldsymbol{u}_h, \boldsymbol{v}_h, \boldsymbol{w}_h),$$

where the form $c_{n\ell}$ is

$$c_{n\ell}(\boldsymbol{z}, \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) = \sum_{\tau\in\mathcal{T}_h} \int_{\partial\tau_-^{\boldsymbol{z}}} |\{\boldsymbol{u}\}\cdot \boldsymbol{n}_\tau| (\boldsymbol{v}^{int} - \boldsymbol{v}^{ext})\cdot \boldsymbol{w}^{int}.$$

We note that

$$c_\ell(\tilde{\boldsymbol{u}}, \boldsymbol{u}, \boldsymbol{v}) - c_\ell(\tilde{\boldsymbol{u}}_n, \boldsymbol{u}_n, \boldsymbol{v}) = c_\ell(\tilde{\boldsymbol{u}}_n, \boldsymbol{u}-\boldsymbol{u}_n, \boldsymbol{v}) + c_\ell(\tilde{\boldsymbol{u}}-\tilde{\boldsymbol{u}}_n, \boldsymbol{u}, \boldsymbol{v}), \tag{44}$$

and

$$\begin{aligned} c_{n\ell}(\tilde{\boldsymbol{u}}, \tilde{\boldsymbol{u}}, \boldsymbol{u}, \boldsymbol{v}) - c_{n\ell}(\tilde{\boldsymbol{u}}_n, \tilde{\boldsymbol{u}}_n, \boldsymbol{u}_n, \boldsymbol{v}) &= c_{n\ell}(\tilde{\boldsymbol{u}}_n, \tilde{\boldsymbol{u}}-\tilde{\boldsymbol{u}}_n, \boldsymbol{u}, \boldsymbol{v}) \\ +c_{n\ell}(\tilde{\boldsymbol{u}}_n, \tilde{\boldsymbol{u}}_n, \boldsymbol{u}-\boldsymbol{u}_n, \boldsymbol{v}) &+ c_{n\ell}(\tilde{\boldsymbol{u}}, \tilde{\boldsymbol{u}}, \boldsymbol{u}, \boldsymbol{v}) - c_{n\ell}(\tilde{\boldsymbol{u}}_n, \tilde{\boldsymbol{u}}, \boldsymbol{u}, \boldsymbol{v}). \end{aligned} \tag{45}$$

Using (44) and (45) in (43), we obtain

$$\begin{aligned} \mathcal{S}_{\tilde{\boldsymbol{u}}_n}((\boldsymbol{u}-\boldsymbol{u}_n, p-p_n), (\boldsymbol{v}, q)) &= -c_\ell(\tilde{\boldsymbol{u}}-\tilde{\boldsymbol{u}}_n, \tilde{\boldsymbol{u}}, \boldsymbol{v}) - c_{n\ell}(\tilde{\boldsymbol{u}}_n, \tilde{\boldsymbol{u}}-\tilde{\boldsymbol{u}}_n, \boldsymbol{u}, \boldsymbol{v}) \\ &\quad -c_{n\ell}(\tilde{\boldsymbol{u}}, \tilde{\boldsymbol{u}}, \boldsymbol{u}, \boldsymbol{v}) + c_{n\ell}(\tilde{\boldsymbol{u}}_n, \tilde{\boldsymbol{u}}, \boldsymbol{u}, \boldsymbol{v}). \end{aligned} \tag{46}$$

Using continuity of c, we have

$$|c_\ell(\tilde{\boldsymbol{u}}-\tilde{\boldsymbol{u}}_n, \tilde{\boldsymbol{u}}, \boldsymbol{v})| \leqslant C_0 \|\tilde{\boldsymbol{u}}-\tilde{\boldsymbol{u}}_n\|_{\boldsymbol{V}} \|\tilde{\boldsymbol{u}}\|_{\boldsymbol{V}} \|\boldsymbol{v}\|_{\boldsymbol{V}}, \tag{47}$$

$$|c_{n\ell}(\tilde{\boldsymbol{u}}_n, \tilde{\boldsymbol{u}}-\tilde{\boldsymbol{u}}_n, \boldsymbol{u}, \boldsymbol{v})| \leqslant C_0 \|\tilde{\boldsymbol{u}}-\tilde{\boldsymbol{u}}_n\|_{\boldsymbol{V}} \|\boldsymbol{u}\|_{\boldsymbol{V}} \|\boldsymbol{v}\|_{\boldsymbol{V}}. \tag{48}$$

Using the inf-sup condition (21), the bounds (47), (48), equation (46) becomes

$$\begin{aligned} \beta\|(\boldsymbol{u}-\boldsymbol{u}_n, p-p_n)\|_{(\boldsymbol{V},Q)} &\leqslant C_0(\|\tilde{\boldsymbol{u}}\|_{\boldsymbol{V}} + \|\boldsymbol{u}\|_{\boldsymbol{V}})\|\tilde{\boldsymbol{u}}-\tilde{\boldsymbol{u}}_n\|_{\boldsymbol{V}} \\ &\quad + \sup_{\boldsymbol{v}\in\boldsymbol{V}_h} \frac{|c_{n\ell}(\tilde{\boldsymbol{u}}_n, \tilde{\boldsymbol{u}}, \boldsymbol{u}, \boldsymbol{v}) - c_{n\ell}(\tilde{\boldsymbol{u}}, \tilde{\boldsymbol{u}}, \boldsymbol{u}, \boldsymbol{v})|}{\|\boldsymbol{v}\|_{\boldsymbol{V}}}. \end{aligned}$$

To conclude that G is continuous, it suffices to show that:

$$\lim_{n\to\infty} \sup_{\boldsymbol{v}\in\boldsymbol{V}_h} \frac{|c_{n\ell}(\tilde{\boldsymbol{u}}_n, \tilde{\boldsymbol{u}}, \boldsymbol{u}, \boldsymbol{v}) - c_{n\ell}(\tilde{\boldsymbol{u}}, \tilde{\boldsymbol{u}}, \boldsymbol{u}, \boldsymbol{v})|}{\|\boldsymbol{v}\|_{\boldsymbol{V}}} = 0.$$

This is proved in the following lemma.

Lemma 4 *Assume that* $(\tilde{\boldsymbol{u}}_n)_n$ *converges to* $\tilde{\boldsymbol{u}}$ *in* $\boldsymbol{V}_{bs}$. *Then for any* $\boldsymbol{u} \in \boldsymbol{V}_{bs}$, *we have*

$$\lim_{n\to\infty} \sup_{\boldsymbol{v}\in\boldsymbol{V}_h} \frac{|c_{n\ell}(\tilde{\boldsymbol{u}}_n, \tilde{\boldsymbol{u}}, \boldsymbol{u}, \boldsymbol{v}) - c_{n\ell}(\tilde{\boldsymbol{u}}, \tilde{\boldsymbol{u}}, \boldsymbol{u}, \boldsymbol{v})|}{\|\boldsymbol{v}\|_{\boldsymbol{V}}} = 0. \tag{49}$$

Proof Denote by θ the integrand of $\ell(\tilde{\boldsymbol{u}}_n, \tilde{\boldsymbol{u}}, \boldsymbol{u}, \boldsymbol{v})$. We have

$$c_{n\ell}(\tilde{\boldsymbol{u}}_n, \tilde{\boldsymbol{u}}, \boldsymbol{u}, \boldsymbol{v}) - c_{n\ell}(\tilde{\boldsymbol{u}}, \tilde{\boldsymbol{u}}, \boldsymbol{u}, \boldsymbol{v}) = \sum_{\tau\in\mathcal{T}_h} \Big(\int_{\partial\tau_-^{\tilde{\boldsymbol{u}}_n}} \theta - \int_{\partial\tau_-^{\tilde{\boldsymbol{u}}}} \theta \Big).$$

We now fix an element τ in $\mathcal{T}_h$ and consider one face F in $\partial\tau$. We recall the definition (9):

$$\partial\tau_-^{\tilde{\boldsymbol{u}}} = \{\boldsymbol{x} \in \partial\tau \ : \ \{\tilde{\boldsymbol{u}}\} \cdot \boldsymbol{n}_\tau < 0\}, \quad \partial\tau_-^{\tilde{\boldsymbol{u}}_n} = \{\boldsymbol{x} \in \partial\tau \ : \ \{\tilde{\boldsymbol{u}}_n\} \cdot \boldsymbol{n}_\tau < 0\}.$$

From the definition of $\boldsymbol{V}_h$ the functions $\{\tilde{\boldsymbol{u}}\} \cdot \boldsymbol{n}_\tau$ and $\{\tilde{\boldsymbol{u}}_n \cdot \boldsymbol{n}_\tau\}$ are quadratic polynomials on F. Define

$$W_F = \int_{F\cap\partial\tau_-^{\tilde{\boldsymbol{u}}_n}} \theta - \int_{F\cap\partial\tau_-^{\tilde{\boldsymbol{u}}}} \theta.$$

From equivalence of norms and the fact that $\tilde{\boldsymbol{u}}_n$ converges to $\tilde{\boldsymbol{u}}$ in $\boldsymbol{V}_h$, we have

$$\lim_{n\to\infty} \|\{\tilde{\boldsymbol{u}}\} \cdot \boldsymbol{n}_F - \{\tilde{\boldsymbol{u}}_n\} \cdot \boldsymbol{n}_F\|_{L^\infty(F)} = 0. \tag{50}$$

Clearly this implies that if the quadratic function $\{\tilde{\boldsymbol{u}}\} \cdot \boldsymbol{n}_F$ never vanishes on $\overline{F}$, then there exists $n_0 \geqslant 0$ so that for all $n \geqslant n_0$ the function $\{\tilde{\boldsymbol{u}}_n\} \cdot \boldsymbol{n}_F$ has the same strict sign than $\{\tilde{\boldsymbol{u}}\} \cdot \boldsymbol{n}_F$. This implies that $W = 0$. In the other case where the quadratic function $\{\tilde{\boldsymbol{u}}\} \cdot \boldsymbol{n}_F$ vanishes to at least one point on $\overline{F}$, we write using the characteristic function:

$$W_F = \int_F \theta(\chi_{F\cap\partial\tau_-^{\tilde{\boldsymbol{u}}_n}} - \chi_{F\cap\partial\tau_-^{\tilde{\boldsymbol{u}}}}).$$

This gives by expanding the definition of θ:

$$|W_F| \leqslant \|\{\tilde{\boldsymbol{u}}\} \cdot \boldsymbol{n}_\tau\|_{L^4(F)} \|[\boldsymbol{u}]\|_{L^4(F)} \|\boldsymbol{v}^{\text{int}}\|_{L^4(F)} \|\chi_{F\cap\partial\tau_-^{\tilde{\boldsymbol{u}}_n}} - \chi_{F\cap\partial\tau_-^{\tilde{\boldsymbol{u}}}}\|_{L^4(F)}.$$

Using trace inequalities, we have for constants M_F^* independent of $\boldsymbol{v}$ and n:

$$\sum_{\tau\in\mathcal{T}_h} \sum_{F\in\partial\tau} |W_F| \leqslant \|\boldsymbol{v}\|_{\boldsymbol{V}} \sum_{\tau\in\mathcal{T}_h} \sum_{F\in\partial\tau} M_F^* \|\chi_{F\cap\partial\tau_-^{\tilde{\boldsymbol{u}}_n}} - \chi_{F\cap\partial\tau_-^{\tilde{\boldsymbol{u}}}}\|_{L^4(F)}.$$

This implies that

$$\sup_{\boldsymbol{v}\in\boldsymbol{V}_h} \frac{\ell(\tilde{\boldsymbol{u}}_n, \tilde{\boldsymbol{u}}, \boldsymbol{u}, \boldsymbol{v}) - \ell(\tilde{\boldsymbol{u}}, \tilde{\boldsymbol{u}}, \boldsymbol{u}, \boldsymbol{v})}{\|\boldsymbol{v}\|_{\boldsymbol{V}}} \leqslant \sum_{\tau\in\mathcal{T}_h} \sum_{F\in\partial\tau} M_F^* \|\chi_{F\cap\partial\tau_-^{\tilde{\boldsymbol{u}}_n}} - \chi_{F\cap\partial\tau_-^{\tilde{\boldsymbol{u}}}}\|_{L^4(F)}.$$

It suffices then to show that for a given face F of a mesh element τ

$$\lim_{n\to\infty} \|\chi_{F\cap\partial\tau_-^{\tilde{\boldsymbol{u}}_n}} - \chi_{F\cap\partial\tau_-^{\tilde{\boldsymbol{u}}}}\|_{L^4(F)} = 0.$$

Table 1 Errors and convergence rates for viscosity $\mu = 1$ and for $\epsilon = 1$.

$\sqrt{2}h$	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{\mathcal{T}_h}$	rate	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{L^2(\Omega)}$	rate	$\|p-p_h\|_Q$	rate
1/2	1.44×10^{0}	–	1.52×10^{-1}	–	1.01×10^{0}	–
1/4	9.10×10^{-1}	0.66	5.41×10^{-2}	1.48	1.00×10^{0}	0.01
1/8	5.12×10^{-1}	0.83	1.82×10^{-2}	1.57	6.95×10^{-1}	0.53
1/16	2.70×10^{-1}	0.92	5.27×10^{-3}	1.81	3.96×10^{-1}	0.81
1/32	1.38×10^{-1}	0.97	1.40×10^{-3}	1.91	2.08×10^{-1}	0.93

The quadratic function $\{\widetilde{\boldsymbol{u}}\}\cdot\boldsymbol{n}_F$ vanishes at one or two points located in $\bar{F}$. Fix $\epsilon > 0$. From (50), we see that there exists $n_0 \geqslant 0$ so that for all $n \geqslant n_0$ the function $\{\widetilde{\boldsymbol{u}}_n\}\cdot\boldsymbol{n}_F$ has the same strict sign than $\{\widetilde{\boldsymbol{u}}\}\cdot\boldsymbol{n}_F$ on F except on one or two (at most) intervals whose total length is bounded by ϵ. So we have

$$\|\chi_{F\cap\partial\mathcal{T}_{-}^{\tilde{u}_n}} - \chi_{F\cap\partial\mathcal{T}_{-}^{\tilde{u}}}\|^4_{L^4(F)} \leqslant \epsilon,$$

which concludes the proof.

7 Numerical Results

We test the proposed method on a unit square for the follwing exact solution (see [16,8]). The velocity and pressure are defined as:

$$\boldsymbol{u}(x,y) = (1 - e^{\lambda x}\cos(2\pi y), \frac{\lambda}{2\pi}e^{\lambda x}\sin(2\pi y)),$$
$$p(x,y) = \frac{1}{2} - \frac{1}{2}e^{2\lambda x},$$

with the parameter

$$\lambda = \frac{-8\pi^2\mu}{1 + (1 + 64\pi^2\mu^2)^{1/2}}.$$

A Picard iteration is used for solving the nonlinear problem (10). The stopping criterion is that the broken gradient norm of the difference between two successive iterates is less than 10^{-10}. The scheme is slightly modified to take into account the fact that the exact velocity does not vanishes on the boundary.
First the viscosity μ is set equal to one. The domain is partitioned into a triangular mesh and the numerical solutions are obtained on a sequence of successive uniformly refined meshes. Table 1 and table 2 show the errors and the convergence rates for the symmetric and non-symmetric a, namely $\epsilon = -1$ and $\epsilon = +1$, respectively. We observe first order convergence rates for $\|\boldsymbol{u}-\boldsymbol{u}_h\|_{\mathcal{T}_h}$ and $\|p-p_h\|_Q$. We also see that the method is second order convergent for the L^2 error $\|\boldsymbol{u}-\boldsymbol{u}_h\|_{L^2(\Omega)}$.
Second we decrease the viscosity and simulate the flow with $\mu = 0.1$ and $\mu = 0.01$. In other words, we increase the Reynolds numbers. Table 3 and table 4 show the errors and convergence rates $\mu = 0.1$ and $\mu = 0.01$ respectively. The non-symmetric form a (with $\epsilon = 1$) is chosen. We observe as expected

Table 2 Errors and convergence rates for viscosity $\mu = 1$ and for $\epsilon = -1$.

$\sqrt{2}h$	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{\mathcal{T}_h}$	rate	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{L^2(\Omega)}$	rate	$\|p-p_h\|_Q$	rate
1/2	1.93×10^{0}	–	2.44×10^{-1}	–	9.52×10^{-1}	–
1/4	9.34×10^{-1}	1.05	5.31×10^{-2}	2.20	1.05×10^{0}	−0.14
1/8	5.15×10^{-1}	0.85	1.78×10^{-2}	1.57	7.10×10^{-1}	0.56
1/16	2.71×10^{-1}	0.93	5.22×10^{-3}	1.77	3.98×10^{-1}	0.83
1/32	1.38×10^{-2}	0.97	1.40×10^{-3}	1.91	2.08×10^{-1}	0.99

Table 3 Errors and convergence rates for viscosity $\mu = 10^{-1}$.

$\sqrt{2}h$	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{\mathcal{T}_h}$	rate	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{L^2(\Omega)}$	rate	$\|p-p_h\|_Q$	rate
1/2	1.43×10^{0}	–	1.06×10^{-1}	–	9.19×10^{-2}	–
1/4	8.64×10^{-1}	0.73	4.77×10^{-2}	1.15	9.07×10^{-2}	0.02
1/8	4.92×10^{-1}	0.81	1.91×10^{-2}	1.32	7.44×10^{-2}	0.28
1/16	2.75×10^{-1}	0.84	6.15×10^{-3}	1.63	4.71×10^{-2}	0.66
1/32	1.47×10^{-1}	0.90	1.74×10^{-3}	1.82	2.60×10^{-2}	0.86

Table 4 Errors and convergence rates for viscosity $\mu = 10^{-2}$.

$\sqrt{2}h$	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{\mathcal{T}_h}$	rate	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{L^2(\Omega)}$	rate	$\|p-p_h\|_Q$	rate
1/2	2.16×10^{0}	–	1.84×10^{-1}	–	1.11×10^{-1}	–
1/4	1.42×10^{0}	0.60	6.95×10^{-2}	1.40	5.52×10^{-2}	1.01
1/8	7.72×10^{-1}	0.88	2.72×10^{-2}	1.35	2.71×10^{-2}	1.02
1/16	4.21×10^{-1}	0.87	9.59×10^{-3}	1.51	1.46×10^{-2}	0.89
1/32	2.44×10^{-1}	0.79	3.09×10^{-3}	1.63	7.62×10^{-3}	0.94

first order convergence of the method. For this example, the method with a symmetric form ($\epsilon = -1$) was not stable. We believe that the lack of stability is due to the tighter small data condition (23) that holds for the symmetric case.

8 Conclusions

In this paper we analyzed a first-order discontinuous Galerkin method for the steady-state incompressible Navier-Stokes equations that is not stabilized by the standard penalty jump term. Stabilization is obtained by locally enriching the velocity space by one quadratic function on each mesh element. The main contribution of the paper is the theoretical analysis of the scheme. It is based on the application of the Leray-Schauder theorem. Convergence of the method is proved under a small data condition. This theoretical result is confirmed by a numerical test problem.

The penalty-free numerical solution satisfies the momentum equation locally on each mesh element. It would be of interest to further investigate the benefits of this property by solving a complex application, such as the coupling of

Navier-Stokes region with a porous medium region via the Beavers-Joseph-Saffman condition.

References

1. P.F. Antonietti, F. Brezzi, and L.D. Marini. Bubble stabilization of discontinuous Galerkin methods. *Computer Methods in Applied Mechanics and Engineering*, 198:1651–1659, 2009.
2. F. Bassi, A. Crivellini, D.A. Di Pietro, and S. Rebay. An artificial compressibility flux for the discontinuous Galerkin solution of the incompressible Navier-Stokes equation. *Journal of Computational Physics*, 218:794–815, 2006.
3. S. Brenner. Poincaré-friedrichs inequalities for piecewise h^1 functions. *SIAM Journal on Numerical Analysis*, 41:306–324, 2003.
4. F. Brezzi and L.D. Marini. Bubble stabilization of discontinuous Galerkin methods. *in: W. Fitzgibbon, R. Hoppe, J. Periaux, O. Pironneau, Y. Vassilevski (Eds.), Proceedings of International Conference on the Occasion of the 60th Birthday of Y.A. Kuznetsov, Institute of Numerical Mathematics of the Russian Academy of Sciences, Moskow*, pages 25–36, 2006.
5. E. Burman and B. Stamm. Low order discontinuous Galerkin methods for second-order elliptic problems. *SIAM Journal on Numerical Analysis*, 47:508–533, 2008.
6. E. Burman and B. Stamm. Bubble stabilized discontinuous Galerkin method for Stokes problem. *Mathematical Models and Methods in Applied Science*, 20:297–313, 2010.
7. B. Cockburn, G. Kanschat, D. Schötzau, and C. Schwab. Local discontinuous Galerkin methods for the Stokes system. *SIAM Journal on Numerical Analysis*, 40:319–343, 2002.
8. B. Cockburn, G. Kanschat, D. Schötzau, and C. Schwab. A locally conservative LDG method for the incompressible Navier-Stokes equations. *Mathematics of Computation*, 74:1067–1095, 2004.
9. B. Cockburn, G. Kanschat, D. Schötzau, and C. Schwab. A note on discontinuous Galerkin divergence-free solutions of the Navier-Stokes equations. *Journal of Scientific Computing*, 31:61–73, 2007.
10. V. Girault and B. Riviere. DG approximation of coupled Navier-Stokes and Darcy equations by Beavers-Joseph-Saffman interface condition. *SIAM Journal on Numerical Analysis*, 47:2052–2089, 2009.
11. V. Girault, B. Riviere, and M.F. Wheeler. A discontinuous Galerkin method with nonoverlapping domain decomposition for the Stokes and Navier-Stokes problems. *Mathematics of Computation*, 74:53–84, 2004.
12. V. Girault, B. Riviere, and M.F. Wheeler. A splitting method using discontinuous Galerkin for the transient incompressible Navier-Stokes problems. *ESAIM: Mathematical Modelling and Numerical Analysis*, 39:1115–1147, 2005.
13. P. Hansbo and M.G. Larson. Discontinuous Galerkin methods for incompressible and nearly incompressible elasticity by Nitsche's method. *Computer Methods in Applied Mechanics and Engineering*, 191:1895–1908, 2002.
14. F. Heimann, C. Engwer, O. Ippisch, and P. Bastian. An unfitted interior penalty discontinuous Galerkin method for incompressible Navier-Stokes two-phase flow. *International Journal for Numerical Methods in Fluids*, 2012. Published online, DOI: 10.1002/fld.3653.
15. O.A. Karakashian and W. Jureidini. A nonconforming finite element method for the stationary Navier-Stokes equations. *SIAM Journal on Numerical Analysis*, 35:93–120, 1998.
16. L.I.G. Kovasnay. Laminar flow behind a two-dimensional grid. *Proc. Camb. Philos. Soc.*, 44:58–62, 1948.
17. A. Montlaur, S. Fernandez-Mendez, J. Peraire, and A. Huerta. Discontinuous Galerkin methods for the Navier-Stokes equations using solenoidal approximations. *International Journal for Numerical Methods in Fluids*, 64:549–564, 2010.

18. J.T. Oden, I. Babuška, and C.E. Baumann. A discontinuous hp finite element method for diffusion problems. *Journal of Computational Physics*, 146:491–519, 1998.
19. B. Riviere and V. Girault. Discontinuous finite element methods for incompressible flows on subdomains with non-matching interfaces. *Computer Methods in Applied Mechanics and Engineering*, 195:3274–3292, 2006.
20. B. Riviere, M.F. Wheeler, and V. Girault. A priori error estimates for finite element methods based on discontinuous approximation spaces for elliptic problems. *SIAM Journal on Numerical Analysis*, 39:902–931, 2001.
21. D. Schötzau, C. Schwab, and A. Toselli. Mixed hp-DGFEM for incompressible flows. *SIAM Journal on Numerical Analysis*, 40:2171–2194, 2002.
22. A. Toselli. Hp discontinuous Galerkin approximations for the Stokes problem. *Mathematical Models and Methods in Applied Science*, 12:1565–1597, 2002.