An Unusual Stabilized Finite Element Method for a Generalized Stokes Problem

Gabriel Raúl Barrenechea, Frédéric Valentin

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

An Unusual Stabilized Finite Element Method for a generalized Stokes problem

Gabriel BARRENECHEA — Frédéric VALENTIN

N° 4173

April 25, 2001

THÈME 4

Rapport de recherche

ISSN 0249-6399 ISRN INRIA/RR--4173--FR+ENG

INRIA
ROCQUENCOURT

An Unusual Stabilized Finite Element Method for a generalized Stokes problem

Gabriel BARRENECHEA* , Frédéric VALENTIN†

Thème 4 — Simulation et optimisation
de systèmes complexes
Projet M3N

Rapport de recherche n° 4173 — April 25, 2001 — 42 pages

Abstract: An unusual stabilized finite element is presented and analyzed herein for a generalized Stokes problem with a dominating zeroth order term. The method consists in subtracting a mesh dependent term from the formulation without compromising consistency. The design of this mesh dependent term, as well as the stabilization parameter involved, are suggested by bubble condensation. Stability is proven for any combination of velocity and pressure spaces, under the hypotheses of continuity for the pressure space. Optimal order error estimates are derived for the velocity and the pressure, using the standard norms for these unknowns. Numerical experiments confirming these theoretical results, and comparisons with previous methods are presented.

Key-words: bubble functions, generalized Stokes problem, stabilized method.

* INRIA Rocquencourt - Projet M3N - e-mail : Gabriel.Barrenechea@inria.fr.
† LNCC, Av. Getúlio Vargas, 333, 25651-070 Petrópolis - RJ, Brazil e-mail: valentin@lncc.br.

Unité de recherche INRIA Rocquencourt
Domaine de Voluceau, Rocquencourt, BP 105, 78153 Le Chesnay Cedex (France)
Téléphone : +33 1 39 63 55 11 — Télécopie : +33 1 39 63 53 30

Une Méthode d'Eléments Finis Stabilisée non-usuelle pour un problème de Stokes generalisé

Résumé : Une nouvelle méthode d'éléments finis stabilisée non-usuelle est proposée et analysée pour un problème de Stokes généralisé avec un terme d' ordre zéro dominant. La méthode consiste à soustraire un terme dépendant du maillage de la formulation sans comprometre la consistence de la méthode. Pour construire ce nouveau terme et aussi pour la construction du paramètre de stabilisation nous nous sommes basés sur la méthode de condensation de la bulle. Nous démontrons la stabilité de la méthode et des estimations d'erreur optimales. Finalement, des experiences numériques confirmant nos résultats théoriques sont montrés.

Mots-clés : fonctions bulle, problème de Stokes généralisé, méthode stabilisée.

Contents

1 Introduction

Numerical solution of Stokes-like systems presents a major difficulty, namely, the need for a compatibility condition (the inf-sup condition, see [13], [5], and the references therein) relating the discrete spaces used to approximate the velocity field $\mathbf{u}$ and the pressure p. It is a well known fact that equal order interpolation spaces for velocity and pressure, which are the most attractive spaces from an implementational point of view, fail to satisfy this condition. To overcome this difficulty, in the mid eighties Hughes et.al. [18], [17] and Brezzi et. al. [7], [6] introduced a new formulation based on a Petrov-Galerkin idea that consists in adding a mesh dependent term to the formulation in order to make stable pairs of spaces that weren't stable. This new formulation received the name of Stabilized Finite Element Methods. An error analysis was performed in all previous references based on a H^1-type seminorm for the pressure, analysis that was improved in [11] where optimal order approximation results were obtained using the L^2-norm of the pressure.

On the other hand, stabilized finite element methods for problems with zeroth order terms have been derived in the last few years. In [15] a stabilized finite element method was proposed for the advective-diffusive equation with a production term. In [21] a Streamline Diffusion Finite Element Method (SD-FEM) was proposed and analyzed for a Stokes problem with a zeroth order term and under the presence of convection (and also the analysis of the method was performed for the full Navier-Stokes equation). On the other hand, in [10] the connection of Stabilized methods with Galerkin methods enriched with bubble functions (connection first pointed out in [19], [3] and [4]) was used to derive a new kind of stabilized finite element method, namely, the Unusual Stabilized Finite Element Method (USFEM). The particularity of such methods is that the mesh-dependent term is now subtracted from the formulation to reduce the impact of the zeroth order term without compromising the consistency of the method. The method from [10] has been recently improved in [12] where an advective-diffusive-reactive equation is treated (see also [16] where the extension to the case of a negative zeroth order term is considered).

The purpose of this work is to derive, analyze and test a new unusual stabilized finite element method, analogous to the one presented in [12], for a generalized Stokes problem, this is, a Stokes-like system with a dominating zeroth order term. This kind of problems arise naturally in the time discretization of a non-steady Stokes problem, or the full Navier-Stokes equations by means of an operator sppliting technique. The derivation of the method is carried out in Section 2 using an idea from [2] to build a special bubble space which reproduces the residual of the discrete equations and then to proceed with the bubble condensation procedure. This calculation suggests us the form of the mesh dependent term to be subtracted from the formulation, as well as the design of the stabilization parameter. The method is then introduced in Section 3 where the stability of the method is proved and an error analysis based on an H^1-type seminorm for the pressure is performed. This estimate is improved in Section 4, where a new stability result is given (based on an inf-sup condition now) and an error estimate is given, this time using the L^2 norm of the pressure. Finally, in Section 5 we report some numerical experiments that confirm our approximation results and the absence of oscillations for the pressure for a wide range of physical parameters.

2 Preliminaries: The effect of bubbles

Let Ω be a bounded open subset of $\mathbb{R}^2$, $\mathbf{f} \in L^2(\Omega)^2$ and σ a positive real number (typically, $\sigma \approx \frac{1}{\Delta t}$ where Δt is the time step in a time discretization procedure). Then, our generalized Stokes problem reads: *Find* $(\mathbf{u}, p) \in H^1(\Omega)^2 \times L^2_0(\Omega)$ *such that:*

$$\begin{aligned} \sigma\mathbf{u} - \nu\Delta\mathbf{u} + \nabla p &= \mathbf{f} && \text{in } \Omega \\ \nabla\cdot\mathbf{u} &= 0 && \text{in } \Omega \\ \mathbf{u} &= 0 && \text{on } \partial\Omega\,, \end{aligned} \tag{2.1}$$

where $L^2_0(\Omega) := \{q \in L^2(\Omega) : (q,1)_\Omega = 0\}$, and $(\cdot,\cdot)_D$ denotes the L^2 inner product in $L^2(D)$ (or in $L^2(D)^2, L^2(D)^{2\times 2}$, when necessary). The weak form of this problem reads: *Find* $(\mathbf{u}, p) \in H^1_0(\Omega)^2 \times L^2_0(\Omega)$ *such that:*

$$a(\mathbf{u},\mathbf{v}) - (p, \nabla\cdot\mathbf{v})_\Omega + (q, \nabla\cdot\mathbf{u})_\Omega = (\mathbf{f},\mathbf{v})_\Omega \quad \forall\,(\mathbf{v},q) \in H^1_0(\Omega)^2 \times L^2_0(\Omega)\,, \tag{2.2}$$

where

$$a(\mathbf{u},\mathbf{v}) := \sigma(\mathbf{u},\mathbf{v})_\Omega + \nu(\nabla\mathbf{u},\nabla\mathbf{v})_\Omega\,. \tag{2.3}$$

Let $\mathbf{V}_h$ and Q_h be finite dimensional subspaces of $H^1_0(\Omega)^2$ and $L^2_0(\Omega)$, respectively. Then, the usual Galerkin scheme reads: *Find* $(\mathbf{u}_h, p_h) \in \mathbf{V}_h \times Q_h$ *such that:*

$$a(\mathbf{u}_h,\mathbf{v}_h) - (p_h, \nabla\cdot\mathbf{v}_h)_\Omega + (q_h, \nabla\cdot\mathbf{u}_h)_\Omega = (\mathbf{f}_h,\mathbf{v}_h)_\Omega \quad \forall\,(\mathbf{v}_h,q_h) \in \mathbf{V}_h \times Q_h\,. \tag{2.4}$$

This problem is well posed only if the interpolation spaces are compatible in the sense of the inf-sup condition (see [13], [5]). The addition of bubble functions to the velocity space has been proved to be a way of building stable pairs of spaces (cf. [1], where the MINI-element is proposed), and this is why we now introduce a suitable velocity space which contains bubbles. Let $\mathcal{T}_h$ be a triangulation of Ω made up of triangles, and let us suppose that this triangulation is shape regular (for the details, see [8]). We denote by $\mathcal{B}_K$ a space of bubble functions to be specified later; let us just say for the moment that this bubble space is a subspace of $H_0^1(K)$. We define the spaces

$$\mathcal{B} := \bigoplus_{K\in\mathcal{T}_h} \mathcal{B}_K \subset H_0^1(\Omega)\,, \tag{2.5}$$

and

$$\mathbf{V}_h := [V_1 + \mathcal{B}]^2, \tag{2.6}$$

where $V_1 := \{v \in \mathcal{C}^0(\overline{\Omega})/v\big|_K \in P^1(K)\,, \forall K \in \mathcal{T}_h\}$, and we seek a solution of (2.2)-(2.3) belonging to $\mathbf{V}_h \times V_1$. Hence, what we look for is a solution of the form

$$\mathbf{u}_h = \mathbf{u}_1 + \sum_{K\in\mathcal{T}_h} \mathbf{u}_B^K, \tag{2.7}$$

where $\mathbf{u}_1$ belongs to $[V_1]^2$ and $\mathbf{u}_B^K$ belongs to $[\mathcal{B}_K]^2$, for all $K \in \mathcal{T}_h$.

What we do now is to give an expression of the bubble part of the solution $\mathbf{u}_B^K$ in terms of its linear part $\mathbf{u}_1$, pressure p, and $\mathbf{f}$, and then plug this representation into the equations to build a new method. This process is called bubble condensation procedure. To this end, we first test (2.4) against a function $(\mathbf{v}_B^K, 0)$ in each K, with $\mathbf{v}_B^K \in [\mathcal{B}_K]^2$, and we arrive to

$$a(\mathbf{u}_1, \mathbf{v}_B^K)_K + a(\mathbf{u}_B^K, \mathbf{v}_B^K)_K - (p, \nabla\cdot \mathbf{v}_B^K)_K = (\mathbf{f}, \mathbf{v}_B^K)_K \quad \forall \mathbf{v}_B^K \in [\mathcal{B}_K]^2\,, \tag{2.8}$$

for all $K \in \mathcal{T}_h$, where the subscript K indicates integration over K. Now, we note that, integrating by parts and using the fact that $\mathbf{u}_1 \in P^1(K)^2$ and $\mathbf{v}_B^K \in H_0^1(K)^2$ we have

$$(\nabla \mathbf{u}_1, \nabla \mathbf{v}_B^K)_K = 0\,, \tag{2.9}$$

$$(p, \nabla\cdot \mathbf{v}_B^K)_K = -(\nabla p, \mathbf{v}_B^K)_K\,. \tag{2.10}$$

Hence, (2.8) reduces to

$$a(\mathbf{u}_B^K, \mathbf{v}_B^K)_K = (\mathbf{f} - \sigma \mathbf{u}_1 - \nabla p, \mathbf{v}_B^K)_K \qquad \forall \mathbf{v}_B^K \in [\mathcal{B}_K]^2\,, \tag{2.11}$$

or, written in another way

$$a(\mathbf{u}_B^K, \mathbf{v}_B^K)_K \;=\; (P_B^K(\mathbf{f} - \sigma\mathbf{u}_1 - \nabla p), \mathbf{v}_B^K)_K \qquad \forall \mathbf{v}_B^K \in [\mathcal{B}_K]^2\,, \tag{2.12}$$

where P_B^K is the L^2-projection from $P^1(K)^2$ onto $[\mathcal{B}_K]^2$. Now, to set our problem in the notations used in [2] we define the operator $S_B^K : [\mathcal{B}_K]^2 \to [\mathcal{B}_K]^2$ as $S_B(\psi) = \mathbf{b}$, where $\mathbf{b}$ is the (unique) solution of

$$a(\mathbf{b}, \mathbf{v}_B^K)_K \;=\; (\psi, \mathbf{v}_B^K)_K \qquad \forall \mathbf{v}_B^K \in [\mathcal{B}_K]^2. \tag{2.13}$$

With this notation we see that the bubble part of $\mathbf{u}_h$ is given by

$$\mathbf{u}_B^K \;=\; S_B P_B^K(\mathbf{f} - \sigma\mathbf{u}_1 - \nabla p) \qquad \forall K \in \mathcal{T}_h. \tag{2.14}$$

What we want is to give an expression for $\mathbf{u}_B^K$ in terms of $\mathbf{u}_1, p$ and $\mathbf{f}$ simpler than the one given in (2.14). For simplicity of the presentation, we will restrict ourselves to the case $\mathbf{f} \in P^1(K)^2\ \forall K \in \mathcal{T}_h$, and hence our residual $\mathbf{f} - \sigma\mathbf{u}_1 - \nabla p$ belongs to $P^1(K)^2$. Moreover, we will consider the construction of each component of the bubble part $\mathbf{u}_B^K$ separately. To this end, we denote $S_B P_B^K = (S_B^1 P_B^1, S_B^2 P_B^2)$ and $\mathbf{u}_B^K = (u_B^1, u_B^2)$.

We are interested in proving that there exists a bubble space $\mathcal{B}_K$ *reproducing the linear operator* $\mu[I_{P^1}]^2$ (where I_{P^1} is the identity operator on $P^1(K)$ and $\mu > 0$ will be specified later); this is, a bubble space $\mathcal{B}_K$ such that $P_{P^1}^K S_B P_B^K = \mu[I_{P^1}]^2$, $P_{P^1}^K$ being the L^2-projection from $[\mathcal{B}_K]^2$ onto $[P^1(K)]^2$. This is done in [2], Sect. 3.1, for a scalar advection diffusion problem. We will show that the same construction works in our case, and hence we will use it to build an alternative formulation for our problem.

We first build an auxiliary space of bubble functions, space that we will use to build explicit bounds for μ. This space is built by decoupling triangle K into four subtriangles generated by joining the midpoints of the edges of K, and denoted by $K_i, i = 0, ..., 3$, K_0 being the central triangle (see Figure 1 below).

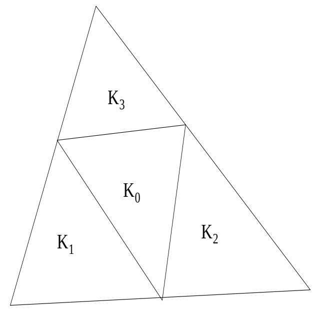

Figure 1: Subdivisions of K.

This auxiliary bubble space is denoted by $\tilde{\mathcal{B}}_K$, and we propose the following base for it: $B_0^K = \{b_K^1, b_K^2, b_K^3\}$, where b_K^j is the cubic bubble on the element K_j. The fact that this is a well suited base for our computations has been proved in [2], § 3.1. Moreover, we observe that B_0^K is an orthogonal base. With this base, we define the set $\{s_1, s_2, s_3\} \subset \tilde{\mathcal{B}}_K$, built by solving the local problems

$$a(s_i, b_K^j)_K \;:=\; \sigma(s_i, b_K^j)_K + \nu(\nabla s_i, \nabla b_K^j)_K \;=\; (z_i, b_K^j)_K \quad j \in \{1,2,3\}\,, \tag{2.15}$$

where $\{z_i : i = 1,2,3\}$ is the orthonormal base of $P^1(K)$ (for the $L^2(K)$ inner product) given by

$$z_i \;:=\; \frac{\varphi_i \sqrt{3}}{|K|^{1/2}} \quad i = 1,2,3, \tag{2.16}$$

where the φ_i's are the only P^1 functions such as $\varphi_i(M_j) = \delta_{ij}$, M_j, $j = 1,2,3$, being the midpoints of the edges of K. Then, we define the matrix $S := (S_{ij})$ by

$$S_{ij} \;=\; \sigma(s_i, s_j)_K + \nu(\nabla s_i, \nabla s_j)_K\,. \tag{2.17}$$

In [2] is proved that, for all μ less or equal than the smallest eigenvalue of S, there exists a virtual bubble space ($\mathcal{B}_K$) reproducing $\mu[I_{P^1}]^2$. This is why we now detail matrix S. By setting $s_i = \Sigma_{r=1}^3 \alpha_i^r b_K^r$, $i = 1,2,3$, we have that, for $i = 1,2,3$, α_i^r satisfies

$$\sigma\,(\Sigma_{r=1}^3 \alpha_i^r b_K^r, b_K^j)_K \;+\; \nu(\nabla \left(\Sigma_{r=1}^3 \alpha_i^r b_K^r\right), \nabla b_K^j)_K \;=\; (z_i, b_K^j)_K\,,$$

or

$$\sigma\alpha_i^j\,||b_K^j||_{0,K_j}^2 \,+\, \nu\alpha_i^j\,||\nabla b_K^j||_{0,K_j}^2 \;=\; (z_i, b_K^j)_K\,. \tag{2.18}$$

Assume now that we have numbered the triangles K_j and the midpoints M_i such that $M_j \notin \overline{K}_j$ $j = 1, 2, 3$. If $\overline{x}_i$ denotes the barycenter of $K_i (i = 1, 2, 3)$, then

$$\varphi_i(\overline{x}_j) \;=\; \begin{cases} \dfrac{2}{3} & \text{if } i \neq j \\[2ex] -\dfrac{1}{3} & \text{if } i = j\,, \end{cases} \tag{2.19}$$

which we write as $\varphi_i(\overline{x}_j) = \frac{1}{3}(2 - 3\delta_{ij})$. We get then

$$\begin{aligned} (z_i, b_K^j)_K &= \frac{9}{20}|K_j|\; z_i(\overline{x}_j) b_K^j(\overline{x}_j) \\ &= \frac{9}{20}\frac{|K|}{4}\frac{\sqrt{3}}{|K|^{1/2}}\,\varphi_i(\overline{x}_j)\,\frac{1}{27} \\ &= \frac{|K|^{1/2}\sqrt{3}}{720}\,(2 - 3\delta_{ij})\,. \end{aligned} \tag{2.20}$$

Hence, from (2.20), (2.18) becomes

$$\alpha_i^j \;=\; \frac{C\,|K|^{1/2}}{\sigma\,||b_K^j||_{0,K_j}^2 + \nu||\nabla b_K^j||_{0,K_j}^2}\,(2 - 3\delta_{ij})\,. \tag{2.21}$$

Moreover, if we denote by l_1, l_2, l_3 the length of the three sides of a triangle K, and $d_K := l_1^2 + l_2^2 + l_3^2$, we have, using the regularity of the triangulation,

$$\int_{K_j} |\nabla b_K^j|^2 \;=\; \frac{d_{K_j}}{720\,|K_j|} \;=\; \frac{d_K}{720\,|K|} \;=\; C_*\,, \tag{2.22}$$

and

$$|K| \;=\; Ch_K^2\,, \tag{2.23}$$

$$||b_K^j||_{0,K_j}^2 \;=\; C^* h_{K_j}^2 \;=\; \frac{C^*}{4} h_K^2\,, \tag{2.24}$$

where h_K denotes the usual element diameter, and C, C^* and C_* are real positive constants depending only on the shape of triangle K. Hence, from (2.22) and (2.24), (2.21) becomes

$$\alpha_i^j = \frac{C\,|K|^{1/2}}{\sigma\, C^* h_K^2 + \nu\, C_*}\,(2 - 3\delta_{ij})\,. \tag{2.25}$$

We are now ready to calculate S_{ij} more explicitly. Indeed, applying (2.17), (2.15), (2.20) and (2.21) we get

$$\begin{aligned} S_{ij} &= a(s_i, s_j)_K \\ &= \sum_{r=1}^{3} \alpha_j^r\, a(s_i, b_K^r)_K \\ &= \sum_{r=1}^{3} \alpha_j^r\, (z_i, b_K^r)_K \\ &= \frac{C\,|K|}{\sigma\, C^* h_K^2 + \nu\, C_*} \sum_{r=1}^{3} (2 - 3\delta_{ir})(2 - 3\delta_{jr}) \\ &= \frac{C\,|K|}{\sigma\, C^* h_K^2 + \nu\, C_*}\,\delta_{ij}\,, \end{aligned} \tag{2.26}$$

and hence S is a scalar matrix. Indeed, using (2.23) we get

$$S = \frac{C_1 h_K^2}{\sigma h_K^2 + C_2 \nu}\, Id^3\,, \tag{2.27}$$

where Id^3 is now the 3×3 identity matrix, and C_1, C_2 are positive constants that depend only on the shape of triangle K.

Now we are ready to apply Theorem 1 (and Remark 4) from [2]. Indeed, since S is a scalar matrix, we see that its smallest eigenvalue is $\mu_0 = \dfrac{C_1 h_K^2}{\sigma h_K^2 + C_2 \nu}$, and hence, for every $\mu \leq \mu_0$ there exists a (virtual) bubble space $(\mathcal{B}_K)$ such that $P_{P^1}^i S_B^i P_B^i = \mu I_{P^1}\,, i = 1, 2$, the constant being the same for both components of $P_{P^1}^K S_B P_B^K$. In particular, for $\mu = \mu_0$ we have that there exists a bubble space such that $P^i_{P^1} S^i_B P^i_B = \dfrac{C_1 h_K^2}{\sigma h_K^2 + C_2 \nu} I_{P^1}, \; i = 1, 2$. Hence, with this choice (2.14) leads to

$$\begin{aligned} P^1_{P^1}(u^1_B) &= \frac{C_1 h_K^2}{\sigma h_K^2 + C_2 \nu}(f_1 - \sigma u_1 - \frac{\partial p}{\partial x_1}) \\ P^2_{P^1}(u^2_B) &= \frac{C_1 h_K^2}{\sigma h_K^2 + C_2 \nu}(f_2 - \sigma u_2 - \frac{\partial p}{\partial x_2}) \,, \end{aligned}$$

and then the bubble part of $\mathbf{u}_h$ satisfies

$$(\mathbf{u}^K_B, \mathbf{v}_1)_K = (P_{P^1}(\mathbf{u}^K_B), \mathbf{v}_1)_K = \frac{C_1 h_K^2}{\sigma h_K^2 + C_2 \nu}(\mathbf{f} - \sigma \mathbf{u}_1 - \nabla p, \mathbf{v}_1)_K \,, \qquad (2.28)$$

for all $\mathbf{v}_1 \in P^1(K)$.

Remark 2.1 *The virtual space whose existence we just proved is not necessarily related to the space spanned by B_0^K. In fact, $\tilde{\mathcal{B}}_K$ was just used to build matrix S and to find an explicit bound for μ_0.* □

Finally, we rewrite (2.4) with this choice of bubble part and using (2.9) we get, for all $\mathbf{v}_1 \in [V_1]^2$

$$\begin{aligned} a(\mathbf{u}_h, \mathbf{v}_1) &= a(\mathbf{u}_1, \mathbf{v}_1) + \sum_{K \in \mathcal{T}_h} a(\mathbf{u}^K_B, \mathbf{v}_1)_K \\ &= a(\mathbf{u}_1, \mathbf{v}_1) + \sum_{K \in \mathcal{T}_h} \sigma(\mathbf{u}^K_B, \mathbf{v}_1)_K \\ &= a(\mathbf{u}_1, \mathbf{v}_1) + \sum_{K \in \mathcal{T}_h} \frac{C_1 h_K^2}{\sigma h_K^2 + C_2 \nu}(\mathbf{f} - \sigma \mathbf{u}_1 - \nabla p, \sigma \mathbf{v}_1)_K \,, (2.29) \end{aligned}$$

and, analogously

$$\begin{aligned} (q, \nabla \cdot \mathbf{u}_h)_\Omega &= (q, \nabla \cdot \mathbf{u}_1)_\Omega + \sum_{K \in \mathcal{T}_h} (q, \nabla \cdot \mathbf{u}^K_B)_K \\ &= (q, \nabla \cdot \mathbf{u}_1)_\Omega - \sum_{K \in \mathcal{T}_h} (\nabla q, \mathbf{u}^K_B)_K \\ &= (q, \nabla \cdot \mathbf{u}_1)_\Omega - \sum_{K \in \mathcal{T}_h} \frac{C_1 h_K^2}{\sigma h_K^2 + C_2 \nu}(\mathbf{f} - \sigma \mathbf{u}_1 - \nabla p, \nabla q)_K (2.30) \end{aligned}$$

Hence, adding (2.29) and (2.30) we arrive to the following "bubble condensed" Galerkin scheme: *Find* $(\mathbf{u}_1, p) \in [V_1]^2 \times V_1$ *such that:*

$$\begin{aligned} a(\mathbf{u}_1, \mathbf{v}_1) \;-\; (p, \nabla\cdot \mathbf{v}_1)_\Omega \;+\; (q, \nabla\cdot \mathbf{u}_1)_\Omega \\ - \sum_{K\in\mathcal{T}_h} \frac{C_1 h_K^2}{\sigma h_K^2 + C_2\nu}(\sigma\mathbf{u}_1 + \nabla p, \sigma\mathbf{v}_1 - \nabla q)_K \\ = \; (\mathbf{f}, \mathbf{v}_1)_\Omega \;-\; \sum_{K\in\mathcal{T}_h} \frac{C_1 h_K^2}{\sigma h_K^2 + C_2\nu}(\mathbf{f}, \sigma\mathbf{v}_1 - \nabla q)_K \,. \end{aligned} \tag{2.31}$$

3 The stabilized finite element method

In this section we present the stabilized finite element method that we are going to use, and state stability and approximation results.

First, based on the calculations made in previous section, we introduce the following unusual stabilized finite element method: *Find* $(\mathbf{u}_h, p_h) \in \mathbf{V}_h \times Q_h$ *such that:*

$$\mathbf{B}((\mathbf{u}_h,p_h),(\mathbf{v},q)) \;=\; \mathbf{F}(\mathbf{v},q) \qquad \forall\,(\mathbf{v},q)\in \mathbf{V}_h\times Q_h\,, \tag{3.1}$$

where $\mathbf{V}_h := [V_k \cap H_0^1(\Omega)]^2$ and $Q_h := V_l \cap L_0^2(\Omega)$, $k,l \geq 1$, and where, for $k \geq 1$, $V_k := \{v \in \mathcal{C}^0(\overline{\Omega})/v\Big|_K \in P^k(K)\,,\,\forall K \in \mathcal{T}_h\}$, $\mathbf{B}$ and $\mathbf{F}$ are given by

$$\begin{aligned}\mathbf{B}((\mathbf{u}_h,p_h),(\mathbf{v},q)) \;&:=\; \sigma(\mathbf{u}_h,\mathbf{v})_\Omega \,+\, \nu(\nabla\mathbf{u}_h,\nabla\mathbf{v})_\Omega \,-\, (p_h,\nabla\cdot\mathbf{v})_\Omega \\ +\,(q,\nabla\cdot\mathbf{u}_h)_\Omega \,-\, &\sum_{K\in\mathcal{T}_h}\tau_K\,(\sigma\mathbf{u}_h-\nu\Delta\mathbf{u}_h+\nabla p_h,\sigma\mathbf{v}-\nu\Delta\mathbf{v}-\nabla q)_K\,,\end{aligned} \tag{3.2}$$

$$\mathbf{F}(\mathbf{v},q) \;:=\; (\mathbf{f},\mathbf{v})_\Omega \,-\, \sum_{K\in\mathcal{T}_h}\tau_K\,(\mathbf{f},\sigma\mathbf{v}-\nu\Delta\mathbf{v}-\nabla q)_K\,. \tag{3.3}$$

Here, the stabilization parameter τ_K is given by

$$\tau_K \;:=\; \frac{h_K^2}{\sigma h_K^2\xi(\lambda_K)+\dfrac{4\nu}{m_k}}\,, \tag{3.4}$$

where

$$\begin{aligned}\lambda_K &= \frac{4\nu}{m_k\sigma h_K^2}\,, && (3.5)\\ m_k &= \min\{\frac{1}{3},C_k\}\,, && (3.6)\\ C_k h_K^2\,\|\Delta v\|_{0,K}^2 &\leq \|\nabla v\|_{0,K}^2 \qquad \forall\, v\in V_k\,, && (3.7)\\ \xi(\lambda) &= \max\{\lambda,1\}\,. && (3.8)\end{aligned}$$

Remark 3.1 *The design of the stabilization parameter has been done in order to have no free constants to set, which is a common fact in Stokes flow stabilization (see, for example [18], and [21]). Indeed, the only non explicit constant in the design of the parameter is m_k, which varies from one kind of element to another one, but these constants have been tabulated in [14]. Also we remark that the method may be implemented using triangle and quadrilateral elements. We also remark that, even if the derivation of the method has been done using equal order interpolation spaces for the velocity and pressure, the method can be used for different order interpolation spaces, as long as we use continuous interpolations for the pressure.* □

Remark 3.2 *If $k = 1$ (i.e., in the case in which we use piecewise linear elements for the velocity), then our method is much simpler, with $m_1 = \frac{1}{3}$. In fact, it reduces to the formulation (2.31) from § 2, with τ_K instead of the quotient used there. On the other hand, in the case of a diffusive dominated flow ($\frac{4\nu}{m_k} \geq \sigma h_K^2$), and of course for the Stokes flow as a limit when $\sigma \to 0$, we recover the GLS method from [17], with the following stabilization parameter*

$$\tau_K = \frac{m_k h_K^2}{8\nu} . \square$$

Remark 3.3 *We recall the formulation of the SDFEM method from [21]: Find $(\mathbf{u}_h, p_h) \in \mathbf{V}_h \times Q_h$ such that:*

$$\mathbf{B}_\delta((\mathbf{u}_h, p_h), (\mathbf{v}_h, q_h)) = \mathbf{F}_\delta(\mathbf{v}_h, q_h) \qquad \forall (\mathbf{v}_h, q_h) \in \mathbf{V}_h \times Q_h , \tag{3.9}$$

where

$$\begin{aligned}
\mathbf{B}_\delta((\mathbf{u}_h, p_h), (\mathbf{v}_h, q_h)) &:= \sigma (\mathbf{u}_h, \mathbf{v}_h)_\Omega + \nu (\nabla \mathbf{u}_h, \nabla \mathbf{v}_h)_\Omega \\
&\quad - (p_h, \nabla\cdot \mathbf{v}_h)_\Omega + (q_h, \nabla\cdot \mathbf{u}_h)_\Omega \\
&\quad + \sum_{K \in \mathcal{T}_h} \frac{\delta h_K^2}{\nu} (\sigma \mathbf{u}_h - \nu \Delta \mathbf{u}_h + \nabla p_h, \nabla q_h)_K , \\
\mathbf{F}_\delta(\mathbf{v}_h, q_h) &:= (\mathbf{f}, \mathbf{v}_h)_\Omega + \sum_{K \in \mathcal{T}_h} \frac{\delta h_K^2}{\nu} (\mathbf{f}, \nabla q_h)_K ,
\end{aligned}$$

and δ is a positive real constant. There are two major differences between the SDFEM method and our method. First, the design of the stabilization *parameter which has a constant to set in the SDFEM method. Moreover, the constant δ for the SDFEM method must depend on σ to have stability. Indeed, in [21] stability is proved under the assumption $\frac{\delta\sigma h_K^2}{\nu} \leq \frac{1}{2}$ and $\delta \leq \frac{C_k}{2}$. The other difference with our method is the absence of $\sigma\mathbf{v}_h - \nu\Delta\mathbf{v}_h$ on the test function term, which is not present in (3.9).* □

Remark 3.4 *From now on, $\|\cdot\|_{l,D}$ and $|\cdot|_{l,D}$ will denote the norm and semi-norm on $H^l(D)$ (or in $H^l(D)^2$ when necessary), respectively, with the usual convention $H^0(D) = L^2(D)$. Also, the following seminorm will be useful later:*

$$|p|_h := \left\{ \sum_{K\in\mathcal{T}_h} \tau_K \, \|\nabla p\|_{0,K}^2 \right\}^{\frac{1}{2}} . \tag{3.10}$$

Now, since we are looking for continuous approximations of the pressure (since we demand $l \geq 1$), the space in which we are looking the discrete pressure for is a subspace of $H^1(\Omega)/\mathbb{R}$, space in which $|\cdot|_h$ is a norm. Further, we state two inverse inequalities which will be useful later on. First, using an inverse estimate analogous to (3.7) (see [8]) we obtain

$$\begin{aligned} |p|_h &= \left\{ \sum_{K\in\mathcal{T}_h} \frac{h_K^2}{\sigma h_K^2 \xi(\lambda_K) + \dfrac{4\nu}{m_k}} \|\nabla p\|_{0,K}^2 \right\}^{1/2} \\ &\leq \left\{ \sum_{K\in\mathcal{T}_h} \frac{m_k}{4\nu} h_K^2 \, \|\nabla p\|_{0,K}^2 \right\}^{1/2} \\ &\leq \left\{ \sum_{K\in\mathcal{T}_h} \frac{m_k}{4\nu C(l)} \|p\|_{0,K}^2 \right\}^{1/2} \\ &\leq C \, \frac{1}{\sqrt{4\nu}} \, \|p\|_{0,\Omega} . \end{aligned} \tag{3.11}$$

Secondly, by the definition of τ_K*, the inverse inequality (3.7) and the fact that* $\frac{m_k}{C_k} \leq 1$*, we obtain*

$$\begin{aligned}
\sum_{K\in\mathcal{T}_h} \nu^2 \tau_K \|\Delta \mathbf{v}\|_{0,K}^2 &= \sum_{K\in\mathcal{T}_h} \frac{\nu^2 h_K^2}{\sigma h_K^2 \xi(\lambda_K) + \dfrac{4\nu}{m_k}} \|\Delta \mathbf{v}\|_{0,K}^2 \\
&\leq \sum_{K\in\mathcal{T}_h} \frac{m_k \nu h_K^2 h_K^{-2}}{4C_k} \|\nabla \mathbf{v}\|_{0,K}^2 \\
&\leq \sum_{K\in\mathcal{T}_h} \nu \|\nabla \mathbf{v}\|_{0,K}^2 . \qquad (3.12)
\end{aligned}$$

Finally, we point out the fact that, since we are interested in an error estimate for k *and* l *fixed,* $C(l)$ *(a constant depending only on* l*) and* m_k *are treated as simple constants. Moreover, we have* $m_k \leq \frac{1}{3}$*.* □

3.1 The stability of the method

The main result concerning numerical stability of our method is stated as follows.

Lemma 3.1 *There exists a constant* $C_\Omega > 0$ *(depending only on* Ω*) such that*

$$\mathbf{B}((\mathbf{v}, q), (\mathbf{v}, q)) \geq C_\Omega \nu \|\mathbf{v}\|_{1,\Omega}^2 + |q|_h^2 ,$$

for all $(\mathbf{v}, q) \in \mathbf{V}_h \times Q_h$*.*

Proof: Let us first remark that, since $\xi(\lambda_K) \geq 1$, we have the following bound for τ_K

$$\tau_K \leq \frac{m_k h_K^2}{m_k \sigma h_K^2 + 4\nu} . \qquad (3.13)$$

Also, by the definition of m_k (cf. (3.6)) we have $m_k \leq C_k$. Now, using Schwarz's inequality and (3.7) we see that (denoting with the subscript K the integration over K)

$$\begin{aligned}
\mathbf{B}((\mathbf{v},q),(\mathbf{v},q))_K &= \sigma\,\|\mathbf{v}\|_{0,K}^2 + \nu\,\|\nabla\mathbf{v}\|_{0,K}^2 \\
&- \tau_K\,(\sigma\mathbf{v}-\nu\Delta\mathbf{v}+\nabla q,\sigma\mathbf{v}-\nu\Delta\mathbf{v}-\nabla q)_K \\
&= \sigma\,\|\mathbf{v}\|_{0,K}^2 + \nu\,\|\nabla\mathbf{v}\|_{0,K}^2 \\
&- \tau_K\{\sigma^2\|\mathbf{v}\|_{0,K}^2 - \|\nabla q\|_{0,K}^2 - 2\sigma\nu\,(\mathbf{v},\Delta\mathbf{v})_K + \nu^2\|\Delta\mathbf{v}\|_{0,K}^2\} \\
&= \sigma\,\|\mathbf{v}\|_{0,K}^2 + \nu\,\|\nabla\mathbf{v}\|_{0,K}^2 + \tau_K\|\nabla q\|_{0,K}^2 \\
&- \sigma^2\tau_K\,\|\mathbf{v}\|_{0,K}^2 + 2\sigma\nu\tau_K\,(\mathbf{v},\Delta\mathbf{v})_K - \tau_K\nu^2\,\|\Delta\mathbf{v}\|_{0,K}^2 \\
&\geq (\sigma-\sigma^2\tau_K)\,\|\mathbf{v}\|_{0,K}^2 + \nu\,\|\nabla\mathbf{v}\|_{0,K}^2 + \tau_K\|\nabla q\|_{0,K}^2 \\
&- 2\sigma\nu\tau_K\,\|\mathbf{v}\|_{0,K}\|\Delta\mathbf{v}\|_{0,K} - \frac{\nu^2\tau_K\,h_K^2}{C_k}\,\|\nabla\mathbf{v}\|_{0,K}^2\,.
\end{aligned}$$

Now, since $2ab \leq \dfrac{1}{\gamma}a^2+\gamma b^2$ with $\gamma>0$, we see that (taking $a=\sqrt{\sigma\nu\tau_K}\|\mathbf{v}\|_{0,K}$ and $b=\sqrt{\sigma\nu\tau_K}\|\Delta\mathbf{v}\|_{0,K}$), applying once more (3.7)

$$\begin{aligned}
\mathbf{B}((\mathbf{v},q),(\mathbf{v},q))_K &\geq (\sigma-\sigma^2\tau_K)\,\|\mathbf{v}\|_{0,K}^2 + \tau_K\|\nabla q\|_{0,K}^2 \\
&+ \left(\nu - \frac{\nu^2\tau_K h_K^{-2}}{C_k}\right)\,\|\nabla\mathbf{v}\|_{0,K}^2 \\
&- \left(\frac{\sigma\nu\tau_K}{\gamma}\,\|\mathbf{v}\|_{0,K}^2 + \frac{\gamma\sigma\nu\tau_K h_K^{-2}}{C_k}\,\|\nabla\mathbf{v}\|_{0,K}^2\right) \\
&= (\sigma-\sigma^2\tau_K - \frac{\sigma\nu\tau_K}{\gamma})\,\|\mathbf{v}\|_{0,K}^2 + \tau_K\|\nabla q\|_{0,K}^2 \\
&+ \left(\nu - \frac{\nu^2\tau_K h_K^{-2}}{C_k} - \frac{\gamma\sigma\nu\tau_K h_K^{-2}}{C_k}\right)\,\|\nabla\mathbf{v}\|_{0,K}^2\,.
\end{aligned}$$

Using (3.13) and the fact that $\dfrac{m_k}{C_k}\leq 1$ we arrive to

$$\begin{aligned}
\mathbf{B}((\mathbf{v},q),(\mathbf{v},q))_K &\geq \left(\sigma - \frac{\sigma^2 m_k h_K^2}{m_k\sigma h_K^2+4\nu} - \frac{\sigma\nu m_k h_K^2}{\gamma(m_k\sigma h_K^2+4\nu)}\right)\,\|\mathbf{v}\|_{0,K}^2 \\
&+ \left(\nu - \frac{m_k\nu^2}{C_k(m_k\sigma h_K^2+4\nu)} - \frac{m_k\sigma\nu\gamma}{C_k(m_k\sigma h_K^2+4\nu)}\right)\,\|\nabla\mathbf{v}\|_{0,K}^2 \\
&+ \tau_K\|\nabla q\|_{0,K}^2
\end{aligned}$$

$$
\begin{aligned}
&\geq \left(\frac{\gamma\sigma^2 m_k h_K^2 + 4\gamma\sigma\nu - \gamma\sigma^2 m_k h_K^2 - \sigma\nu m_k h_K^2}{\gamma(m_k\sigma h_K^2 + 4\nu)}\right) \|\mathbf{v}\|_{0,K}^2 \\
&+ \left(\frac{m_k\sigma\nu h_K^2 + 4\nu^2 - \nu^2 - \gamma\sigma\nu}{(m_k\sigma h_K^2 + 4\nu)}\right) \|\nabla\mathbf{v}\|_{0,K}^2 \\
&+ \tau_K \|\nabla q\|_{0,K}^2 \\
&= \left[\frac{\sigma\nu(4\gamma - m_k h_K^2)}{\gamma(m_k\sigma h_K^2 + 4\nu)} \|\mathbf{v}\|_{0,K}^2 + \frac{3\nu^2 + \sigma\nu(m_k h_K^2 - \gamma)}{m_k\sigma h_K^2 + 4\nu} \|\nabla\mathbf{v}\|_{0,K}^2 \right. \\
&+ \left. \tau_K \|\nabla q\|_{0,K}^2 \right].
\end{aligned}
$$

Finally, it suffices to take $\gamma = \frac{m_k h_K^2}{4}$ to obtain

$$
B((\mathbf{v},q),(\mathbf{v},q))_K \geq \frac{3\nu}{4} \|\nabla\mathbf{v}\|_{0,K}^2 + \tau_K \|\nabla q\|_{0,K}^2 ,
$$

on each K, and the proof is finished by adding over $K \in \mathcal{T}_h$ and applying Poincaré's inequality. □

Remark 3.5 *From the proof of Lemma 3.1, we note that we have not used the fact that* $q \in L_0^2(\Omega)$*, and therefore this lower bound for* $\mathbf{B}$ *is valid also in* $\mathbf{V}_h \times V_l$*. This fact will be exploited in the proof of the error below, in which we will use an interpolate of* p *which does not necessarilly belong to* $L_0^2(\Omega)$*.* □

3.2 Error Analysis

Let k, l be integers with $k, l \geq 1$. We suppose the existence of interpolation operators $\mathbf{I}_h^k : (C^0(\overline{\Omega}))^2 \to [V_k]^2$ and $I_h^l : C^0(\overline{\Omega}) \to V_l$ (the Lagrange interpolate, see [8]) such that the approximation of $\mathbf{u}$ by $\tilde{\mathbf{u}}_h := \mathbf{I}_h^k(\mathbf{u})$ and p by $\tilde{p}_h := I_h^l(p)$ can be estimated as follows

$$
|\eta^{\mathbf{u}}|_{m,K} := |\mathbf{u} - \tilde{\mathbf{u}}_h|_{m,K} \leq C h_K^{s-m} |\mathbf{u}|_{s,K} \quad \forall \mathbf{u} \in H^s(K)^2 , \tag{3.14}
$$

$$
|\eta^{p}|_{n,K} := |p - \tilde{p}_h|_{n,K} \leq C h_K^{t-n} |p|_{t,K} \quad \forall p \in H^t(K) , \tag{3.15}
$$

for all $K \in \mathcal{T}_h$, with $0 \leq m \leq 2$ and $\max\{m, 1\} \leq s \leq k+1$, and $0 \leq n \leq 1$ and $\max\{n, 1\} \leq t \leq l+1$. Further, we define $h := \max\{h_K : K \in \mathcal{T}_h\}$, and from now on, we will suppose that $h \leq 1$.

Now we state the main result of this section, result analogous to the one given in [12] for an advective-reactive-diffusive equation.

Theorem 3.1 *Let us suppose that the solution* $(\mathbf{u}, p)$ *of (2.1) belongs to* $(H^{k+1}(\Omega)\cap H_0^1(\Omega))^2 \times (H^{l+1}(\Omega) \cap L_0^2(\Omega))$. *Then, there exists* $C > 0$*, independent of* h*, such that the error* $(\mathbf{e^u}, e^p) := (\mathbf{u} - \mathbf{u}_h, p - p_h)$ *(where* $(\mathbf{u}_h, p_h)$ *is the solution of (3.1)) satisfies*

$$\|\mathbf{e^u}\|_{1,\Omega} + |e^p|_h \le C\,\frac{\max\{\sigma+1, \nu+1, \frac{1}{\sqrt{4\nu}}\}}{\min\{C_\Omega \nu, 1\}}\left[h^k\,|\mathbf{u}|_{k+1,\Omega} + h^{l+1}\,|p|_{l+1,\Omega}\right].$$

Proof: Let $\mathbf{e}_h^{\mathbf{u}} := \mathbf{u}_h - \tilde{\mathbf{u}}_h$ and $e_h^p := p - \tilde{p}_h$. Then, from Lemma 3.1, the definition of $\mathbf{B}$ and the consistency of the method we have

$$\begin{aligned}
& \min\{C_\Omega\nu, 1\}(\|\mathbf{e}_h^{\mathbf{u}}\|_{1,\Omega}^2 + |e_h^p|_h^2) \\
\le\; & C_\Omega \nu\, \|\mathbf{e}_h^{\mathbf{u}}\|_{1,\Omega}^2 + |e_h^p|_h^2 \\
\le\; & \mathbf{B}((\mathbf{e}_h^{\mathbf{u}}, e_h^p), (\mathbf{e}_h^{\mathbf{u}}, e_h^p)) \\
=\; & \mathbf{B}((\eta^{\mathbf{u}}, \eta^p), (\mathbf{e}_h^{\mathbf{u}}, e_h^p)) \\
=\; & \sigma(\eta^{\mathbf{u}}, \mathbf{e}_h^{\mathbf{u}})_\Omega + \nu(\nabla\eta^{\mathbf{u}}, \nabla\mathbf{e}_h^{\mathbf{u}})_\Omega - (\eta^p, \nabla\cdot\mathbf{e}_h^{\mathbf{u}})_\Omega + (e_h^p, \nabla\cdot\eta^{\mathbf{u}})_\Omega \\
& - \sum_{K\in\mathcal{T}_h} \tau_K\,(\sigma\eta^{\mathbf{u}} - \nu\Delta\eta^{\mathbf{u}} + \nabla\eta^p, \sigma\mathbf{e}_h^{\mathbf{u}} - \nu\Delta\mathbf{e}_h^{\mathbf{u}} - \nabla e_h^p)_K\,,
\end{aligned}$$

but, since $\eta^{\mathbf{u}}$ vanishes on $\partial\Omega$ (since $\mathbf{u}$ belongs to $H_0^1(\Omega)^2$), we have after integration by parts

$$\begin{aligned}
& \min\{C_\Omega\nu, 1\}(\|\mathbf{e}_h^{\mathbf{u}}\|_{1,\Omega}^2 + |e_h^p|_h^2) \\
\le\; & \sigma(\eta^{\mathbf{u}}, \mathbf{e}_h^{\mathbf{u}})_\Omega + \nu(\nabla\eta^{\mathbf{u}}, \nabla\mathbf{e}_h^{\mathbf{u}})_\Omega - (\eta^p, \nabla\cdot\mathbf{e}_h^{\mathbf{u}})_\Omega - (\nabla e_h^p, \eta^{\mathbf{u}})_\Omega \\
& - \sum_{K\in\mathcal{T}_h} \tau_K\,(\sigma\eta^{\mathbf{u}} - \nu\Delta\eta^{\mathbf{u}} + \nabla\eta^p, \sigma\mathbf{e}_h^{\mathbf{u}} - \nu\Delta\mathbf{e}_h^{\mathbf{u}} - \nabla e_h^p)_K\,, \\
\le\; & \left[\sum_{K\in\mathcal{T}_h} \sigma\|\eta^{\mathbf{u}}\|_{0,K}^2 + \nu\,\|\nabla\eta^{\mathbf{u}}\|_{0,K}^2 + \|\eta^p\|_{0,K}^2 + \frac{1}{\tau_K}\,\|\eta^{\mathbf{u}}\|_{0,K}^2\right. \\
& \left. + \tau_K\,\|\sigma\eta^{\mathbf{u}} - \nu\Delta\eta^{\mathbf{u}} + \nabla\eta^p\|_{0,K}^2\right]^{\frac{1}{2}} \\
& \cdot\left[\sum_{K\in\mathcal{T}_h} \sigma\|\mathbf{e}_h^{\mathbf{u}}\|_{0,K}^2 + \nu\,\|\nabla\mathbf{e}_h^{\mathbf{u}}\|_{0,K}^2 + \|\nabla\cdot\mathbf{e}_h^{\mathbf{u}}\|_{0,K}^2 + \tau_K\,\|\nabla e_h^p\|_{0,K}^2\right. \\
& \left. + \tau_K\,\|\sigma\mathbf{e}_h^{\mathbf{u}} - \nu\Delta\mathbf{e}_h^{\mathbf{u}} + \nabla e_h^p\|_{0,K}^2\right]^{\frac{1}{2}}
\end{aligned}$$

$$
\begin{aligned}
\leq\ & C\ \Bigg[\sum_{K\in\mathcal{T}_h}(\sigma+\sigma^2\tau_K+\frac{1}{\tau_K})\,\|\eta^{\mathbf{u}}\|_{0,K}^2\ +\ \nu\,\|\nabla\eta^{\mathbf{u}}\|_{0,K}^2\ +\ \nu^2\tau_K\,\|\Delta\eta^{\mathbf{u}}\|_{0,K}^2 \\
& +\|\eta^p\|_{0,K}^2\ +\ \tau_K\,\|\nabla\eta^p\|_{0,K}^2\Bigg]^{\frac{1}{2}} \\
& \cdot\Bigg[\sum_{K\in\mathcal{T}_h}(\sigma+\sigma^2\tau_K)\,\|\mathbf{e}_h^{\mathbf{u}}\|_{0,K}^2\ +\ (\nu+1)\,\|\nabla\mathbf{e}_h^{\mathbf{u}}\|_{0,K}^2\ +\ \nu^2\tau_K\,\|\Delta\mathbf{e}_h^{\mathbf{u}}\|_{0,K}^2 \\
& +\tau_K\,\|\nabla e_h^p\|_{0,K}^2\Bigg]^{\frac{1}{2}}. \qquad (3.16)
\end{aligned}
$$

Now, we use (3.12), the fact that $\sigma^2\tau_K\leq\sigma$ and $\nu^2\tau_K\leq C\nu\,h_K^2$ and we obtain from (3.16)

$$
\begin{aligned}
& \min\{C_\Omega\nu,1\}(\|\mathbf{e}_h^{\mathbf{u}}\|_{1,\Omega}^2\ +\ |e_h^p|_h^2) \\
\leq\ & C\ \Bigg[\sum_{K\in\mathcal{T}_h}(\sigma+\frac{1}{\tau_K})\,\|\eta^{\mathbf{u}}\|_{0,K}^2\ +\ \nu\,\|\nabla\eta^{\mathbf{u}}\|_{0,K}^2\ +\ \nu h_K^2\,\|\Delta\eta^{\mathbf{u}}\|_{0,K}^2 \\
& +\|\eta^p\|_{0,K}^2\ +\ \frac{h_K^2}{4\nu}\,\|\nabla\eta^p\|_{0,K}^2\Bigg]^{\frac{1}{2}} \\
& \cdot\Bigg[\sum_{K\in\mathcal{T}_h}\sigma\,\|\mathbf{e}_h^{\mathbf{u}}\|_{0,K}^2\ +\ (\nu+1)\,\|\nabla\mathbf{e}_h^{\mathbf{u}}\|_{0,K}^2\ +\ \tau_K\,\|\nabla e_h^p\|_{0,K}^2\Bigg]^{\frac{1}{2}} \\
\leq\ & C\sqrt{\max\{\sigma,\nu+1\}}\ \Bigg[\sum_{K\in\mathcal{T}_h}(\sigma+\frac{1}{\tau_K})\,\|\eta^{\mathbf{u}}\|_{0,K}^2\ +\ \nu\,\|\nabla\eta^{\mathbf{u}}\|_{0,K}^2\ +\ \nu h_K^2\,\|\Delta\eta^{\mathbf{u}}\|_{0,K}^2 \\
& +\|\eta^p\|_{0,K}^2\ +\ \frac{h_K^2}{4\nu}\,\|\nabla\eta^p\|_{0,K}^2\Bigg]^{\frac{1}{2}}\cdot\left[\|\mathbf{e}_h^{\mathbf{u}}\|_{1,\Omega}^2\ +\ |e_h^p|_h^2\right]^{\frac{1}{2}},
\end{aligned}
$$

and hence, dividing by the last term we get

$$
\begin{aligned}
\|\mathbf{e}_h^{\mathbf{u}}\|_{1,\Omega}\ +\ |e_h^p|_h\ \leq\ & C\frac{\sqrt{\max\{\sigma,\nu+1\}}}{\min\{C_\Omega\nu,1\}}\ \Bigg[\sum_{K\in\mathcal{T}_h}(\sigma+\frac{1}{\tau_K})\,\|\eta^{\mathbf{u}}\|_{0,K}^2\ +\ \nu\,\|\nabla\eta^{\mathbf{u}}\|_{0,K}^2 \\
& +\nu h_K^2\,\|\Delta\eta^{\mathbf{u}}\|_{0,K}^2\ +\ \|\eta^p\|_{0,K}^2\ +\ \frac{h_K^2}{4\nu}\,\|\nabla\eta^p\|_{0,K}^2\Bigg]^{\frac{1}{2}}. \qquad (3.17)
\end{aligned}
$$

Finally, since $\mathbf{e^u} = \eta^{\mathbf{u}} - \mathbf{e}_h^{\mathbf{u}}$ and $e^p = \eta^p - e_h^p$, we have thanks to (3.17) and triangular inequality

$$\begin{aligned}
&\|\mathbf{e^u}\|_{1,\Omega} + |e^p|_h \leq \|\mathbf{e}_h^{\mathbf{u}}\|_{1,\Omega} + |e_h^p|_h + \|\eta^{\mathbf{u}}\|_{1,\Omega} + |\eta^p|_h \\
\leq\ & C\frac{\sqrt{\max\{\sigma,\nu+1\}}}{\min\{C_\Omega\nu,1\}}\left[\sum_{K\in\mathcal{T}_h}(\sigma+\frac{1}{\tau_K})\,\|\eta^{\mathbf{u}}\|_{0,K}^2 + \nu\,\|\nabla\eta^{\mathbf{u}}\|_{0,K}^2 + \nu h_K^2\,\|\Delta\eta^{\mathbf{u}}\|_{0,K}^2\right. \\
&\left. + \|\eta^p\|_{0,K}^2 + \frac{h_K^2}{4\nu}\,\|\nabla\eta^p\|_{0,K}^2\right]^{\frac{1}{2}} + \|\eta^{\mathbf{u}}\|_{1,\Omega} + |\eta^p|_h \\
\leq\ & C\frac{\sqrt{\max\{\sigma,\nu+1\}}}{\min\{C_\Omega\nu,1\}}\left[\sum_{K\in\mathcal{T}_h}(\sigma+1+\frac{1}{\tau_K})\,\|\eta^{\mathbf{u}}\|_{0,K}^2 + (\nu+1)\,\|\nabla\eta^{\mathbf{u}}\|_{0,K}^2\right. \\
&\left. + \nu h_K^2\,\|\Delta\eta^{\mathbf{u}}\|_{0,K}^2 + \|\eta^p\|_{0,K}^2 + \frac{h_K^2}{4\nu}\,\|\nabla\eta^p\|_{0,K}^2\right]^{\frac{1}{2}}.
\end{aligned} \tag{3.18}$$

Now, we have to separate two cases:
i).- $\lambda_K < 1$. In this case, we have $4\nu < m_k\sigma h_K^2$ and hence

$$\frac{1}{\tau_K} = \frac{m_k\sigma h_K^2 + 4\nu}{m_k h_K^2} < 2\sigma\,,$$

and then (3.18) becomes

$$\begin{aligned}
&\|\mathbf{e^u}\|_{1,\Omega} + |e^p|_h \\
\leq\ & C\frac{\sqrt{\max\{\sigma,\nu+1\}}}{\min\{C_\Omega\nu,1\}}\left[\sum_{K\in\mathcal{T}_h}(\sigma+1)\,\|\eta^{\mathbf{u}}\|_{0,K}^2 + (\nu+1)\,\|\nabla\eta^{\mathbf{u}}\|_{0,K}^2\right. \\
&\left. + \nu h_K^2\,\|\Delta\eta^{\mathbf{u}}\|_{0,K}^2 + \max\{1,\frac{1}{4\nu}\}(\|\eta^p\|_{0,K}^2 + h_K^2\,\|\nabla\eta^p\|_{0,K}^2)\right]^{\frac{1}{2}} \\
\leq\ & C\,\frac{\max\{\sigma+1,\nu+1\}}{\min\{C_\Omega\nu,1\}}\left[\sum_{K\in\mathcal{T}_h}\|\eta^{\mathbf{u}}\|_{0,K}^2 + \|\nabla\eta^{\mathbf{u}}\|_{0,K}^2 + h_K^2\,\|\Delta\eta^{\mathbf{u}}\|_{0,K}^2\right. \\
&\left. + \max\{1,\frac{1}{4\nu}\}(\|\eta^p\|_{0,K}^2 + h_K^2\,\|\nabla\eta^p\|_{0,K}^2)\right]^{\frac{1}{2}}
\end{aligned}$$

$$\leq \quad C\, \frac{\max\{\sigma+1, \nu+1, \frac{1}{\sqrt{4\nu}}\}}{\min\{C_\Omega \nu, 1\}} \left[h^{2k}\, |\mathbf{u}|^2_{k+1,\Omega} \,+\, h^{2l+2} |p|^2_{l+1,\Omega}\right]^{\frac{1}{2}}$$
$$\leq \quad C\, \frac{\max\{\sigma+1, \nu+1, \frac{1}{\sqrt{4\nu}}\}}{\min\{C_\Omega \nu, 1\}} \left[h^{k}\, |\mathbf{u}|_{k+1,\Omega} \,+\, h^{l+1}\, |p|_{l+1,\Omega}\right].$$

ii).- $\lambda_K \geq 1$. In this case, we have $4\nu \geq m_k \sigma h_K^2$ and hence

$$\frac{1}{\tau_K} \;=\; \frac{m_k \sigma h_K^2 + 4\nu}{m_k h_K^2} \;\leq\; \frac{8\nu}{m_k h_K^2}\,.$$

In this case, (3.18) becomes

$$\begin{aligned}
&\|\mathbf{e}^{\mathbf{u}}\|_{1,\Omega} \,+\, |e^p|_h \\
\leq\;& C\frac{\sqrt{\max\{\sigma, \nu+1\}}}{\min\{C_\Omega\nu, 1\}} \left[\sum_{K\in\mathcal{T}_h} \max\{\sigma+1, \nu\} h_K^{-2}\, \|\eta^{\mathbf{u}}\|^2_{0,K} \,+\, (\nu+1)\, \|\nabla \eta^{\mathbf{u}}\|^2_{0,K}\right. \\
&\left. +\, \nu h_K^2\, \|\Delta \eta^{\mathbf{u}}\|^2_{0,K} \,+\, \max\{1, \frac{1}{4\nu}\}(\|\eta^p\|^2_{0,K} \,+\, h_K^2\, \|\nabla \eta^p\|^2_{0,K})\right]^{\frac{1}{2}} \\
\leq\;& C\, \frac{\max\{\sigma+1, \nu+1\}}{\min\{C_\Omega\nu, 1\}} \left[\sum_{K\in\mathcal{T}_h} h_K^{-2}\, \|\eta^{\mathbf{u}}\|^2_{0,K} \,+\, \|\nabla \eta^{\mathbf{u}}\|^2_{0,K} \,+\, h_K^2\, \|\Delta \eta^{\mathbf{u}}\|^2_{0,K}\right. \\
&\left. +\, \max\{1, \frac{1}{4\nu}\}(\|\eta^p\|^2_{0,K} \,+\, h_K^2\, \|\nabla \eta^p\|^2_{0,K})\right]^{\frac{1}{2}} \\
\leq\;& C\, \frac{\max\{\sigma+1, \nu+1, \frac{1}{\sqrt{4\nu}}\}}{\min\{C_\Omega \nu, 1\}} \left[h^{2k}\, |\mathbf{u}|^2_{k+1,\Omega} \,+\, h^{2l+2} |p|^2_{l+1,\Omega}\right]^{\frac{1}{2}} \\
\leq\;& C\, \frac{\max\{\sigma+1, \nu+1, \frac{1}{\sqrt{4\nu}}\}}{\min\{C_\Omega \nu, 1\}} \left[h^{k}\, |\mathbf{u}|_{k+1,\Omega} \,+\, h^{l+1}\, |p|_{l+1,\Omega}\right].
\end{aligned}$$

Both cases i) and ii) finish the proof. □

4 An improved error estimate

The approximation result presented in previous section has two drawbacks. First, it does not guarantee convergence for the case in which $p \in H^l(\Omega)$, which is an interesting case, specially when $l = 1$. Segundo, it considers only the error in the norm $|\cdot|_h$ for the pressure, norm which is not the natural norm for the pressure space $L^2_0(\Omega)$.

Based on these considerations, in this section we perform an improved error analysis based no longer on the ellipticity of bilinear form $\mathbf{B}$, but on an inf-sup condition, which is valid for the L^2-norm for the pressure.

4.1 A new stability result

Throughout all this section (and the following one), C and $C_i, i > 0$ will denote positive constants independent of h, but who may depend on the physical parameters σ and ν. Moreover, the value of C may vary whenever it is written in two different places.

In the proof of the inf-sup condition below, we will use the following technical result, which is a slight variation of Lemma 3.2 in [11] (see also [20]).

Lemma 4.1 *There exist $C_1, C_2 \in \mathbb{R}^+$ such that*

$$\sup_{\mathbf{0}\neq\mathbf{v}\in\mathbf{V}_h} \frac{(\nabla\cdot\mathbf{v}, p)_\Omega}{\|\mathbf{v}\|_{1,\Omega}} \geq C_1\|p\|_{0,\Omega} - C_2\sqrt{\sigma+4\nu}\,|p|_h\,,$$

for all $p \in Q_h$.

Proof: Since $p \in L^2_0(\Omega)$, there exists (cf. [13]) a nonvanishing $\mathbf{w} \in H^1_0(\Omega)^2$ such that

$$(\nabla\cdot\mathbf{w}, p)_\Omega \geq C_3\,\|p\|_{0,\Omega}\|\mathbf{w}\|_{1,\Omega}\,. \tag{4.1}$$

Furthermore, there exists (cf. [13] pp.109-111) an interpolant (the Clément interpolant, originally proposed in [9]) $\tilde{\mathbf{w}} \in \mathbf{V}_h$ such that

$$\left\{\sum_{K\in\mathcal{T}_h} h_K^{-2}\,\|\mathbf{w}-\tilde{\mathbf{w}}\|^2_{0,K}\right\}^{\frac{1}{2}} \leq C_4\,\|\mathbf{w}\|_{1,\Omega}\,, \tag{4.2}$$

and

$$\|\tilde{\mathbf{w}}\|_{1,\Omega}^2 \;\leq\; C_5\,\|\mathbf{w}\|_{1,\Omega}^2\,. \tag{4.3}$$

Hence, using the fact that $p \in Q_h \subseteq C^0(\overline{\Omega})$ and (4.1) we can integrate by parts to arrive to

$$\begin{aligned}
(\nabla\cdot\tilde{\mathbf{w}}, p)_\Omega &= (\nabla\cdot(\tilde{\mathbf{w}}-\mathbf{w}), p)_\Omega + (\nabla\cdot\mathbf{w}, p)_\Omega \\
&\geq (\nabla\cdot(\tilde{\mathbf{w}}-\mathbf{w}), p)_\Omega + C_3\,\|p\|_{0,\Omega}\|\mathbf{w}\|_{1,\Omega} \\
&= -\sum_{K\in\mathcal{T}_h} (\tilde{\mathbf{w}}-\mathbf{w}, \nabla p)_K + C_3\,\|p\|_{0,\Omega}\|\mathbf{w}\|_{1,\Omega} \\
&\geq -\left(\sum_{K\in\mathcal{T}_h} \frac{1}{\tau_K}\,\|\mathbf{w}-\tilde{\mathbf{w}}\|_{0,K}^2\right)^{\frac{1}{2}} \cdot \left(\sum_{K\in\mathcal{T}_h} \tau_K\,\|\nabla p\|_{0,K}^2\right)^{\frac{1}{2}} \\
&\quad + C_3\,\|p\|_{0,\Omega}\|\mathbf{w}\|_{1,\Omega}\,.
\end{aligned}$$

Now, we have that

$$\frac{1}{\tau_K} = \frac{\sigma h_K^2 \xi(\lambda_K) + \dfrac{4\nu}{m_k}}{h_K^2} \leq \begin{cases} C_6\,\dfrac{4\nu}{h_K^2} & \text{if } \lambda_K \geq 1 \\[2ex] C_6\,\dfrac{\sigma+4\nu}{h_K^2} & \text{if } \lambda_K < 1\,, \end{cases} \tag{4.4}$$

keeping in mind that we have supposed that $h_K \leq 1$. Thus we arrive to

$$\begin{aligned}
(\nabla\cdot\tilde{\mathbf{w}}, p)_\Omega &\geq -C_6\left(\sum_{K\in\mathcal{T}_h} (\sigma+4\nu)h_K^{-2}\,\|\mathbf{w}-\tilde{\mathbf{w}}\|_{0,K}^2\right)^{\frac{1}{2}} \cdot \left(\sum_{K\in\mathcal{T}_h} \tau_K\,\|\nabla p\|_{0,K}^2\right)^{\frac{1}{2}} \\
&\quad + C_3\,\|\mathbf{w}\|_{1,\Omega}\|p\|_{0,\Omega} \\
&\geq C_6\left\{-C_4\sqrt{\sigma+4\nu}\left(\sum_{K\in\mathcal{T}_h} \tau_K\,\|\nabla p\|_{0,K}^2\right)^{\frac{1}{2}} + C_3\|p\|_{0,\Omega}\right\}\|\mathbf{w}\|_{1,\Omega} \\
&= C_6\left\{-C_4\sqrt{\sigma+4\nu}\,|p|_h + C_3\|p\|_{0,\Omega}\right\}\|\mathbf{w}\|_{1,\Omega}\,.
\end{aligned}$$

Finally, using (4.3) we arrive to

$$\frac{(\nabla\cdot\tilde{\mathbf{w}}, p)_\Omega}{\|\tilde{\mathbf{w}}\|_{1,\Omega}} \geq \frac{(\nabla\cdot\tilde{\mathbf{w}}, p)_\Omega}{\sqrt{C_5}\|\mathbf{w}\|_{1,\Omega}} \geq \frac{-C_6C_4}{\sqrt{C_5}}\sqrt{\sigma+4\nu}\,|p|_h + \frac{C_6C_3}{\sqrt{C_5}}\|p\|_{0,\Omega}\,, \tag{4.5}$$

which finishes the proof. $\square$

Now we are ready to prove the second result on stability of the method, namely, the inf-sup condition.

Lemma 4.2 *There exists a constant $C = C(\sigma, \nu) > 0$ such that*

$$\sup_{\mathbf{0}\neq(\mathbf{v},q)\in\mathbf{V}_h\times Q_h} \frac{\mathbf{B}((\mathbf{u},p),(\mathbf{v},q))}{(\|\mathbf{v}\|_{1,\Omega}^2 + \|q\|_{0,\Omega}^2)^{\frac{1}{2}}} \geq C\,(\|\mathbf{u}\|_{1,\Omega}^2 + \|p\|_{0,\Omega}^2)^{\frac{1}{2}}\,,$$

for all $(\mathbf{v}, q) \in \mathbf{V}_h \times Q_h$.

Proof: Let $\mathbf{w} \in \mathbf{V}_h$ be a function for which the supremum in Lemma 4.1 is obtained, and let us assume that $\|\mathbf{w}\|_{1,\Omega} = \|p\|_{0,\Omega}$. Then, using the bilinearity of $\mathbf{B}$, Schwarz's inequality and Lemma 4.1 we arrive to:

$$\begin{aligned}
& \mathbf{B}((\mathbf{u},p),(-\mathbf{w},0)) \\
= \; & -\mathbf{B}((\mathbf{u},0),(\mathbf{w},0)) - \mathbf{B}((\mathbf{0},p),(\mathbf{w},0)) \\
= \; & -\left\{\sigma(\mathbf{u},\mathbf{w})_\Omega + \nu(\nabla\mathbf{u},\nabla\mathbf{w})_\Omega - \sum_{K\in\mathcal{T}_h}\tau_K(\sigma\mathbf{u}-\nu\Delta\mathbf{u},\sigma\mathbf{w}-\nu\Delta\mathbf{w})_K\right\} \\
& + (p,\nabla\cdot\mathbf{w})_\Omega - \sum_{K\in\mathcal{T}_h}\tau_K\,(\nabla p, -\sigma\mathbf{w}+\nu\Delta\mathbf{w})_K \\
\geq \; & -\left\{\sigma\|\mathbf{u}\|_{0,\Omega}^2 + \nu\,\|\nabla\mathbf{u}\|_{0,\Omega}^2 + \sum_{K\in\mathcal{T}_h}\tau_K\|\sigma\mathbf{u}-\nu\Delta\mathbf{u}\|_{0,K}^2\right\}^{\frac{1}{2}} \\
& \cdot\left\{\sigma\|\mathbf{w}\|_{0,\Omega}^2 + \nu\,\|\nabla\mathbf{w}\|_{0,\Omega}^2 + \sum_{K\in\mathcal{T}_h}\tau_K\|\sigma\mathbf{w}-\nu\Delta\mathbf{w}\|_{0,K}^2\right\}^{\frac{1}{2}} \\
& + C_1\,\|p\|_{0,\Omega}\|\mathbf{w}\|_{1,\Omega} - C_2\sqrt{\sigma+4\nu}\,|p|_h\|\mathbf{w}\|_{1,\Omega} \\
& - \left\{\sum_{K\in\mathcal{T}_h}\tau_k\|\nabla p\|_{0,K}^2\right\}^{\frac{1}{2}}\left\{\sum_{K\in\mathcal{T}_h}\tau_k[\sigma^2\|\mathbf{w}\|_{0,K}^2 + \nu^2\|\Delta\mathbf{w}\|_{0,K}^2]\right\}^{\frac{1}{2}}. \qquad (4.6)
\end{aligned}$$

Now, applying (3.12), the fact that $\|\mathbf{w}\|_{1,\Omega} = \|p\|_{0,\Omega}$ and $\sigma\tau_K \leq 1$, (4.6) becomes

$$\begin{aligned}
& \mathbf{B}((\mathbf{u},p),(-\mathbf{w},0)) \\
\geq \; & -C_7\left\{\sigma\|\mathbf{u}\|_{0,\Omega}^2 + \nu\,\|\nabla\mathbf{u}\|_{0,\Omega}^2\right\}^{\frac{1}{2}}\cdot\left\{\sigma\|\mathbf{w}\|_{0,\Omega}^2 + \nu\,\|\nabla\mathbf{w}\|_{0,\Omega}^2\right\}^{\frac{1}{2}}
\end{aligned}$$

$$
\begin{aligned}
& + C_1 \, \|p\|^2_{0,\Omega} \, - \, C_2 \sqrt{\sigma + 4\nu} \, |p|_h \|p\|_{0,\Omega} \\
& - C_8 \, |p|_h \left\{ \sum_{K \in \mathcal{T}_h} [\sigma \, \|\mathbf{w}\|^2_{0,K} \, + \, \nu \, \|\nabla \mathbf{w}\|^2_{0,K}] \right\}^{\frac{1}{2}} \\
\geq \; & -C_7 \, \max\{\sigma, \nu\} \, \|\mathbf{u}\|_{1,\Omega} \|\mathbf{w}\|_{1,\Omega} \, + \, C_1 \, \|p\|^2_{0,\Omega} \, - \, C_2 \sqrt{\sigma + 4\nu} \, |p|_h \|p\|_{0,\Omega} \\
& - C_8 \, |p|_h \left\{ \sum_{K \in \mathcal{T}_h} [\sigma \, \|\mathbf{w}\|^2_{0,K} \, + \, \nu \, \|\nabla \mathbf{w}\|^2_{0,K}] \right\}^{\frac{1}{2}} \\
\geq \; & -C_7 \, \max\{\sigma, \nu\} \, \|\mathbf{u}\|_{1,\Omega} \|p\|_{0,\Omega} \, + \, C_1 \, \|p\|^2_{0,\Omega} \, - \, C_2 \sqrt{\sigma + 4\nu} \, |p|_h \|p\|_{0,\Omega} \\
& - C_8 \, \max\{\sigma, \nu\}^{\frac{1}{2}} \, |p|_h \|p\|_{0,\Omega} \, .
\end{aligned}
$$

Now, using the inequality $ab \leq \frac{1}{2\gamma} a^2 + \frac{\gamma}{2} b^2$, $\gamma > 0$, in all the terms which involve a product of norms above, we arrive to

$$
\begin{aligned}
& \mathbf{B}((\mathbf{u}, p), (-\mathbf{w}, 0)) \\
\geq \; & -\frac{C_7}{2\gamma_1} \max\{\sigma, \nu\} \, \|\mathbf{u}\|^2_{1,\Omega} \\
& + \left(C_1 - \frac{\max\{\sigma, \nu\}}{2} [C_7 \gamma_1 + C_8 \gamma_2] - \frac{C_2 (\sigma + 4\nu)}{2} \gamma_3 \right) \, \|p\|^2_{0,\Omega} \\
& + \left(-\frac{C_2}{2\gamma_3} - \frac{C_8}{2\gamma_2} \right) \, |p|^2_h \\
\geq \; & -C_9 \max\{\sigma, \nu\} \|\mathbf{u}\|^2_{1,\Omega} \, + \, C_{10} \|p\|_{0,\Omega} \, - \, C_{11} |p|^2_h \, ,
\end{aligned}
\tag{4.7}
$$

with $C_9, C_{10}, C_{11} > 0$, if γ_1, γ_2 and γ_3 are chosen small enough.

Finally, we denote $(\mathbf{v}, q) := (\mathbf{u} - \delta \mathbf{w}, p)$, $\delta > 0$. Combining Lemma 3.1 and (4.7) we have

$$
\begin{aligned}
& \mathbf{B}((\mathbf{u}, p), (\mathbf{v}, q)) \\
= \; & \mathbf{B}((\mathbf{u}, p), (\mathbf{u}, p)) \, + \, \delta \, \mathbf{B}((\mathbf{u}, p), (-\mathbf{w}, 0)) \\
\geq \; & C_\Omega \nu \, \|\mathbf{u}\|^2_{1,\Omega} \, + \, |p|^2_h \\
& + \delta \, [-C_9 \max\{\sigma, \nu\} \|\mathbf{u}\|^2_{1,\Omega} \, + \, C_{10} \|p\|^2_{0,\Omega} \, - \, C_{11} |p|^2_h] \\
= \; & (C_\Omega \nu - \delta C_9 \max\{\sigma, \nu\}) \, \|\mathbf{u}\|^2_{1,\Omega} \, + \, \delta C_{10} \|p\|^2_{0,\Omega} \\
& + (1 - \delta C_{11}) \, |p|^2_h \\
\geq \; & C \, (\|\mathbf{u}\|^2_{1,\Omega} \, + \, \|p\|^2_{0,\Omega}) \, ,
\end{aligned}
$$

when choosing $0 < \delta < \min\left\{\frac{C_\Omega \nu}{C_9 \max\{\sigma,\nu\}}, \frac{1}{C_{11}}\right\}$.

On the other hand we have

$$\begin{aligned}\|\mathbf{v}\|^2_{1,\Omega} + \|q\|^2_{0,\Omega} &= \|\mathbf{u}-\delta\mathbf{w}\|^2_{1,\Omega} + \|p\|^2_{0,\Omega} \\ &\leq 2\,\|\mathbf{u}\|^2_{1,\Omega} + (2\delta^2+1)\,\|p\|^2_{0,\Omega} \\ &\leq C\,(\|\mathbf{u}\|^2_{1,\Omega} + \|p\|^2_{0,\Omega})\,,\end{aligned}$$

which shows that

$$\begin{aligned}\mathbf{B}((\mathbf{u},p),(\mathbf{v},q)) &\geq C(\sigma,\nu)\,(\|\mathbf{u}\|^2_{1,\Omega} + \|p\|^2_{0,\Omega}) \\ &\geq C(\sigma,\nu)\,(\|\mathbf{u}\|^2_{1,\Omega} + \|p\|^2_{0,\Omega})^{\frac{1}{2}}(\|\mathbf{v}\|^2_{1,\Omega} + \|q\|^2_{0,\Omega})^{\frac{1}{2}}\,,\end{aligned}$$

and the proof is finished. □

4.2 Error Analysis

Let k,l be integers with $k,l \geq 1$. We use the interpolation operator $\mathbf{I}_h^k : (C^0(\overline{\Omega}))^2 \to [V_k]^2$ used in previous section, whose interpolation error satisfies (3.14). We will denote again $\tilde{\mathbf{u}}_h := \mathbf{I}_h^k(\mathbf{u})$. Now, for the pressure we define $\tilde{p}$ as being the Clément interpolate of p. Denoting now by $\tilde{\eta}^p := p - \tilde{p}$, we have (cf. [13])

$$\|\tilde{\eta}^p\|_{n,\Omega} \leq C h^{t-n}\,\|p\|_{t,\Omega} \quad \forall p \in H^t(\Omega)\,, \tag{4.8}$$

with $0 \leq n \leq 1$ and $\max\{n,1\} \leq t \leq l+1$.

Remark 4.1 *The Clément interpolate of p does not necessarily belong to $L^2_0(\Omega)$, and that poses a problem in the proof of our error estimate below. This is why we consider a modified interpolate to p given by*

$$\overline{p} := \tilde{p} - \frac{1}{|\Omega|}\,(\tilde{p},1)_\Omega\,,$$

which belongs to $L^2_0(\Omega)$, and which satisfies also

$$\begin{aligned}\|q-\overline{q}\|_{0,\Omega} &\leq \|q-\tilde{q}\|_{0,\Omega}\,, \\ \|\nabla(q-\overline{q})\|_{0,K} &= \|\nabla(q-\tilde{q})\|_{0,K} \quad \forall K \in \mathcal{T}_h\,,\end{aligned}$$

for all $q \in H^1(\Omega)$, and thus the interpolation error $\eta^p := p - \overline{p}$ associated to this interpolate satisfies the same approximation properties than $\tilde{\eta}^p$ (cf. (4.8)). $\square$.

The main result concerning approximation is now stated.

Theorem 4.1 *Let us suppose that the solution $(\mathbf{u}, p)$ of (2.1) belongs to $(H^{k+1}(\Omega) \cap H_0^1(\Omega))^2 \times (H^l(\Omega) \cap L_0^2(\Omega))$. Then, there exists $C > 0$, independent of h, such that the error $(\mathbf{e}^{\mathbf{u}}, e^p) := (\mathbf{u} - \mathbf{u}_h, p - p_h)$ satisfies*

$$\|\mathbf{e}^{\mathbf{u}}\|_{1,\Omega} + \|e^p\|_{0,\Omega} \leq C \max\{\sigma, \nu + 1, \frac{1}{4\nu}\} \left(h^k \, |\mathbf{u}|_{k+1,\Omega} + h^l \, \|p\|_{l,\Omega}\right).$$

Proof: Let $\mathbf{e}_h^{\mathbf{u}} := \mathbf{u}_h - \tilde{\mathbf{u}}_h$ and $e_h^p := p_h - \overline{p}$. Since e_h^p belongs to $L_0^2(\Omega)$, we can use the stability result given by Lemma 4.2 to conclude the existence of $(\mathbf{v}, q) \in \mathbf{V}_h \times Q_h$ such that

$$\|\mathbf{v}\|_{1,\Omega}^2 + \|q\|_{0,\Omega}^2 \leq C \,,$$

and

$$\|\mathbf{e}_h^{\mathbf{u}}\|_{1,\Omega} + \|e_h^p\|_{0,\Omega} \leq \mathbf{B}((\mathbf{e}_h^{\mathbf{u}}, e_h^p), (\mathbf{v}, q)) \,. \tag{4.9}$$

Using now the consistency of the scheme we obtain

$$\mathbf{B}((\mathbf{e}_h^{\mathbf{u}}, e_h^p), (\mathbf{v}, q)) = \mathbf{B}((\eta^{\mathbf{u}}, \eta^p), (\mathbf{v}, q)) \,. \tag{4.10}$$

Now, for the right hand side of (4.10) we have

$$\begin{aligned} &\mathbf{B}((\eta^{\mathbf{u}}, \eta^p), (\mathbf{v}, q)) = \\ &\sigma\,(\eta^{\mathbf{u}}, \mathbf{v})_\Omega + \nu\,(\nabla \eta^{\mathbf{u}}, \nabla \mathbf{v})_\Omega - (\eta^p, \nabla\cdot \mathbf{v})_\Omega + (q, \nabla\cdot \eta^{\mathbf{u}})_\Omega \\ &- \sum_{K \in \mathcal{T}_h} \tau_K \,(\sigma \eta^{\mathbf{u}} - \nu \Delta \eta^{\mathbf{u}} + \nabla \eta^p, \sigma \mathbf{v} - \nu \Delta \mathbf{v} - \nabla q)_K \end{aligned}$$

$$
\begin{aligned}
&\leq \; C \left\{ \sum_{K\in\mathcal{T}_h} (\sigma \, \|\eta^{\mathbf{u}}\|_{0,K}^2 \; + \; \nu \, \|\nabla \eta^{\mathbf{u}}\|_{0,K}^2 \; + \; \|\eta^p\|_{0,K}^2 \; + \; \|\nabla\cdot \eta^{\mathbf{u}}\|_{0,K}^2 \right. \\
&\qquad \left. + \, \tau_K \, [\sigma^2 \, \|\eta^{\mathbf{u}}\|_{0,K}^2 + \nu^2 \, \|\Delta \eta^{\mathbf{u}}\|_{0,K}^2 + \|\nabla \eta^p\|_{0,K}^2]) \right\}^{\frac{1}{2}} \\
&\qquad \cdot \left\{ \sum_{K\in\mathcal{T}_h} (\sigma \, \|\mathbf{v}\|_{0,K}^2 \; + \; \nu \, \|\nabla \mathbf{v}\|_{0,K}^2 \; + \; \|\nabla\cdot \mathbf{v}\|_{0,K}^2 \; + \; \|q\|_{0,K}^2 \right. \\
&\qquad \left. + \, \tau_K \, [\sigma^2 \, \|\mathbf{v}\|_{0,K}^2 + \nu^2 \, \|\Delta \mathbf{v}\|_{0,K}^2 + \|\nabla q\|_{0,K}^2]) \right\}^{\frac{1}{2}} \\
&\leq \; C \left\{ \sum_{K\in\mathcal{T}_h} [(\sigma + \sigma^2 \tau_K) \, \|\eta^{\mathbf{u}}\|_{0,K}^2 \; + \; (\nu+1) \, \|\nabla \eta^{\mathbf{u}}\|_{0,K}^2 \; + \; \nu^2 \tau_K \, \|\Delta \eta^{\mathbf{u}}\|_{0,K}^2 \right. \\
&\qquad \left. + \, \|\eta^p\|_{0,K}^2 \; + \; \tau_K \|\nabla \eta^p\|_{0,K}^2] \right\}^{\frac{1}{2}} \\
&\qquad \cdot \left\{ \sum_{K\in\mathcal{T}_h} [(\sigma + \sigma^2 \tau_K) \, \|\mathbf{v}\|_{0,K}^2 \; + \; (\nu+1) \, \|\nabla \mathbf{v}\|_{0,K}^2 \; + \; \nu^2 \tau_K \, \|\Delta \mathbf{v}\|_{0,K}^2 \right. \\
&\qquad \left. + \, \|q\|_{0,K}^2 \; + \; \tau_K \|\nabla q\|_{0,K}^2] \right\}^{\frac{1}{2}} .
\end{aligned}
$$

But, since, $(\mathbf{v}, q) \in \mathbf{V}_h \times Q_h$, we can use (3.11) and (3.12), which, together with the definition of τ_K, lead to

$$
\begin{aligned}
& \mathbf{B}((\eta^{\mathbf{u}}, \eta^p), (\mathbf{v}, q)) \\
&\leq \; C \left\{ \sum_{K\in\mathcal{T}_h} \sigma \|\eta^{\mathbf{u}}\|_{0,K}^2 \; + \; (\nu+1) \, \|\nabla \eta^{\mathbf{u}}\|_{0,K}^2 \; + \; \nu h_K^2 \, \|\Delta \eta^{\mathbf{u}}\|_{0,K}^2 \right. \\
&\qquad \left. + \, \|\eta^p\|_{0,K}^2 \; + \; \frac{1}{4\nu} h_K^2 \, \|\nabla \eta^p\|_{0,K}^2 \right\}^{\frac{1}{2}} \\
&\qquad \cdot \left\{ \sum_{K\in\mathcal{T}_h} \sigma \|\mathbf{v}\|_{0,K}^2 \; + \; (\nu+1) \, \|\nabla \mathbf{v}\|_{0,K}^2 \; + \; \nu \, \|\nabla \mathbf{v}\|_{0,K}^2 \right. \\
&\qquad \left. + \, \|q\|_{0,K}^2 \; + \; \frac{1}{4\nu} \|q\|_{0,K}^2 \right\}^{\frac{1}{2}}
\end{aligned}
$$

$$
\begin{aligned}
&\leq \; C \left\{ \sum_{K \in \mathcal{T}_h} \max\{\sigma, \nu+1\} \, [\|\eta^{\mathbf{u}}\|_{1,K}^2 \, + \, h_K^2 \, \|\Delta \eta^{\mathbf{u}}\|_{0,K}^2] \right. \\
&\qquad \left. + \sum_{K \in \mathcal{T}_h} \max\{1, \frac{1}{4\nu}\} \, [\|\eta^p\|_{0,K}^2 \, + \, h_K^2 \, \|\nabla \eta^p\|_{0,K}^2] \right\}^{\frac{1}{2}} \\
&\qquad \cdot \left\{ \max\{\sigma, \nu+1\} \, \|\mathbf{v}\|_{1,\Omega}^2 \, + \, \max\{1, \frac{1}{4\nu}\} \, \|q\|_{0,\Omega}^2 \right\}^{\frac{1}{2}} \\
&\leq \; C \, \max\{\sigma, \nu+1, \frac{1}{4\nu}\} \left\{ \sum_{K \in \mathcal{T}_h} [\|\eta^{\mathbf{u}}\|_{1,K}^2 \, + \, h_K^2 \, \|\Delta \eta^{\mathbf{u}}\|_{0,K}^2] \right. \\
&\qquad \left. + \sum_{K \in \mathcal{T}_h} [\|\eta^p\|_{0,K}^2 \, + \, h_K^2 \, \|\nabla \eta^p\|_{0,K}^2] \right\}^{\frac{1}{2}} , \qquad (4.11)
\end{aligned}
$$

where we have also used the fact that $\|\mathbf{v}\|_{1,\Omega} + \|q\|_{0,\Omega} \leq C$.

Finally, we have, since $\mathbf{e}^{\mathbf{u}} = \eta^{\mathbf{u}} - \mathbf{e}_h^{\mathbf{u}}$ and $e^p = \eta^p - e_h^p$, and thanks to (4.11) and the triangular inequality

$$
\begin{aligned}
\|\mathbf{e}^{\mathbf{u}}\|_{1,\Omega} + \|e^p\|_{0,\Omega} \;&\leq\; \|\mathbf{e}_h^{\mathbf{u}}\|_{1,\Omega} + \|e_h^p\|_{0,\Omega} + \|\eta^{\mathbf{u}}\|_{1,\Omega} + \|\eta^p\|_{0,\Omega} \\
&\leq\; C \, \max\{\sigma, \nu+1, \frac{1}{4\nu}\} \left\{ \sum_{K \in \mathcal{T}_h} [\|\eta^{\mathbf{u}}\|_{1,K}^2 \, + \, h_K^2 \, \|\Delta \eta^{\mathbf{u}}\|_{0,K}^2] \right. \\
&\qquad \left. + \sum_{K \in \mathcal{T}_h} [\|\eta^p\|_{0,K}^2 \, + \, h_K^2 \, \|\nabla \eta^p\|_{0,K}^2] \right\}^{\frac{1}{2}} \\
&\qquad + C \left\{ \sum_{K \in \mathcal{T}_h} [\|\eta^{\mathbf{u}}\|_{1,K}^2 \, + \, \|\eta^p\|_{0,K}^2] \right\}^{\frac{1}{2}} \\
&\leq\; C \, \max\{\sigma, \nu+1, \frac{1}{4\nu}\} \left\{ \sum_{K \in \mathcal{T}_h} [\|\eta^{\mathbf{u}}\|_{1,K}^2 \, + \, h_K^2 \, \|\Delta \eta^{\mathbf{u}}\|_{0,K}^2] \right. \\
&\qquad \left. + \sum_{K \in \mathcal{T}_h} [\|\eta^p\|_{0,K}^2 \, + \, h_K^2 \, \|\nabla \eta^p\|_{0,K}^2] \right\}^{\frac{1}{2}}
\end{aligned}
$$

$$\leq \; C \, \max\{\sigma, \nu+1, \frac{1}{4\nu}\} \left\{ h^{2k} \, |\mathbf{u}|^2_{k+1,\Omega} \, + \, h^{2l} \, \|p\|^2_{l,\Omega} \right\}^{\frac{1}{2}}$$
$$\leq \; C \, \max\{\sigma, \nu+1, \frac{1}{4\nu}\} \, [h^k \, |\mathbf{u}|_{k+1,\Omega} \, + \, h^l \, \|p\|_{l,\Omega}] \, ,$$

which finishes the proof. □

We end this section by making two remarks concerning the choice of the interpolation for p and a possible improvement on the error estimate.

Remark 4.2 *The choice of the Clément interpolation has been made to consider the case $l = 1$ and $p \in H^1(\Omega)$. Indeed, in this case, the pressure is not necessarily continuous, and then we can not take its Lagrange interpolate as in the proof of Theorem 3.1.* □

Remark 4.3 *The error estimate given by Theorem 4.1 may be improved if we have a more regular pressure. Indeed, if we suppose that $p \in H^{l+1}(\Omega)$, we can apply (4.8) differently and obtain*

$$\sum_{K \in \mathcal{T}_h} [\|\eta^p\|^2_{0,K} \, + \, h^2_K \, \|\nabla \eta^p\|^2_{0,K}] \; \leq \; C h^{2l+2} \, \|p\|^2_{l+1,\Omega} \, ,$$

which leads to the following improved error estimate

$$\|\mathbf{e}^{\mathbf{u}}\|_{1,\Omega} \, + \, \|e^p\|_{0,\Omega} \; \leq \; C \max\{\sigma, \nu+1, \frac{1}{4\nu}\} \, (h^k |\mathbf{u}|_{k+1,\Omega} \, + \, h^{l+1} \|p\|_{l+1,\Omega}) \, .\Box$$

4.3 An alternative symmetric formulation

The purpose of this short section is to propose an alternative method which is symmetric, but in which we loose the positiveness of the associated matrix. More precisely, we look for solutions $(\mathbf{u}_h, p_h) \in \mathbf{V}_h \times Q_h$ of the weak (discrete) problem:

$$\mathbf{B}_s((\mathbf{u}_h, p_h), (\mathbf{v}, q)) \; = \; \mathbf{F}_s(\mathbf{v}, q) \qquad \forall \, (\mathbf{v}, q) \in \mathbf{V}_h \times Q_h \, , \tag{4.12}$$

where

$$\mathbf{B}_s((\mathbf{u}_h, p_h), (\mathbf{v}, q)) \; := \; \sigma(\mathbf{u}_h, \mathbf{v})_\Omega \, + \, \nu(\nabla \mathbf{u}_h, \nabla \mathbf{v})_\Omega \, - \, (p_h, \nabla\cdot \mathbf{v})_\Omega$$
$$- \, (q, \nabla\cdot \mathbf{u}_h)_\Omega \, - \, \sum_{K \in \mathcal{T}_h} \tau_K \, (\sigma \mathbf{u}_h - \nu \Delta \mathbf{u}_h + \nabla p_h, \sigma \mathbf{v} - \nu \Delta \mathbf{v} + \nabla q)_K \, , \tag{4.13}$$

$$\mathbf{F}_s(\mathbf{v},q) := (\mathbf{f},\mathbf{v})_\Omega - \sum_{K\in\mathcal{T}_h} \tau_k\,(\mathbf{f},\sigma\mathbf{v}-\nu\Delta\mathbf{v}+\nabla q)_K\,, \tag{4.14}$$

and where the stabilization parameter τ_K is again given by (3.4) and (3.5)-(3.8).

We remark that this method can be obtained in the same way as the one proposed in previous section simply by subtracting instead of adding (2.29) and (2.30). We also remark that this method is stable. Indeed, what we have the analogue of the stability result presented in § 4.1 by simply noting that

$$\mathbf{B}((\mathbf{v},q),(\mathbf{v},q)) = \mathbf{B}_s((\mathbf{v},q),(\mathbf{v},-q)) \quad \forall\,(\mathbf{v},q)\in\mathbf{V}_h\times Q_h\,, \tag{4.15}$$

and hence all Lemmata presented in previous section is valid, and with the same constants. Moreover, the error estimate given in § 4.2 is also valid, again with the same constants.

The reason why we prefer to use in practice the non-symmetric method presented in § 3 is that the positive-definiteness of the matrix involved, which is no longer valid for this symmetric version, allows us to use conjugate gradient-type methods (such as BICG or BICGSTAB) which are faster that the minimum residual methods available for non-definite problems.

5 Numerical experiments

In this section we report two series of numerical results with the unusual stabilized finite element introduced in Section 3. We illustrate the applicability of the method for different values of the physical coefficients, mainly for problems in which the quotient $\dfrac{\sigma\, h_K^2}{\nu}$ is large.

5.1 An analytical solution

We use as domain the square $(0,1)\times(0,1)$, and we set $\mathbf{f}$ to be such as the exact solution of our problem (2.1) is given by

$$\begin{aligned} u_1(x_1,x_2) &= -256\, x_1^2(x_1-1)^2 x_2(x_2-1)(2x_2-1) \\ u_2(x_1,x_2) &= -u_1(x_2,x_1) \\ p(x_1,x_2) &= 150\,(x_1-0.5)(x_2-0.5)\,. \end{aligned} \tag{5.1}$$

Since all numerical experiences have been performed using conforming P^1 finite elements for velocity and pressure, we have measured the relative finite element errors

$$\varepsilon_{\mathbf{u}}(h) \;:=\; \frac{\|\mathbf{u}-\mathbf{u}_h\|_{1,\Omega}}{h}\,, \tag{5.2}$$

and

$$\varepsilon_p(h) \;:=\; \frac{\|p-p_h\|_{0,\Omega}}{h}\,. \tag{5.3}$$

First, we report the case in which we have considered the case $\sigma = 10^2, \nu = 10^{-3}$. In this case, the results are contained in Tables 1 and 2, and they agree with the theoretical results.

h	mesh	$\|p - p_h\|_{0,\Omega}$	$\varepsilon_p(h)$	$\frac{\|p-p_h\|_{0,\Omega}}{h^2}$
$0.05 \times \sqrt{2}$	20×20	6.510442×10^{-2}	0.920715	13.020885
$0.025 \times \sqrt{2}$	40×40	1.588075×10^{-2}	0.449175	12.704600
$0.01\overline{6} \times \sqrt{2}$	60×60	6.713846×10^{-3}	0.284844	12.084924
$0.0125 \times \sqrt{2}$	80×80	3.560257×10^{-3}	0.201398	11.392822
$0.01 \times \sqrt{2}$	100×100	2.131222×10^{-3}	0.1507001	10.656112

Table 1: Convergence history of $p - p_h$

Now, we show the convergence history of $\|\mathbf{u} - \mathbf{u}_h\|_{1,\Omega}$.

h	mesh	$\|\mathbf{u} - \mathbf{u}_h\|_{1,\Omega}$	$\varepsilon_{\mathbf{u}}(h)$
$0.05 \times \sqrt{2}$	20×20	1.026323	14.514407
$0.025 \times \sqrt{2}$	40×40	0.512346	14.491358
$0.01\overline{6} \times \sqrt{2}$	60×60	0.341159	14.474150
$0.0125 \times \sqrt{2}$	80×80	0.255665	14.462634
$0.01 \times \sqrt{2}$	100×100	0.204416	14.454439

Table 2: Convergence history of $\mathbf{u} - \mathbf{u}_h$

Now, we show an evolution of the finite element errors when σ grows. The interest is to show that the errors are not too much affected by the presence of the zeroth order term. To do this, we consider the solution from (5.1) and a viscosity $\nu = 10^{-3}$. We use uniform meshes of 40×40 and 60×60 nodes (≈ 3200 and ≈ 7200 P^1/P^1 elements, respectively).

σ	40×40 mesh		60×60 mesh	
	$\|\mathbf{u}-\mathbf{u}_h\|_{1,\Omega}$	$\|p-p_h\|_{0,\Omega}$	$\|\mathbf{u}-\mathbf{u}_h\|_{1,\Omega}$	$\|p-p_h\|_{0,\Omega}$
0.1	0.540997	9.884753×10^{-3}	0.344096	4.402761×10^{-3}
1	0.539317	9.972938×10^{-3}	0.344045	4.427834×10^{-3}
10	0.527048	1.233970×10^{-2}	0.343225	4.719979×10^{-3}
100	0.512346	1.588075×10^{-2}	0.341159	6.712846×10^{-3}
10^3	0.511944	1.669916×10^{-2}	0.340951	7.412121×10^{-3}
10^4	0.511943	1.743532×10^{-2}	0.340952	7.622682×10^{-3}

Table 3: Behavior of the Finite Element error when σ grows.

We observe that the error in the velocity remains bounded while σ grows and that the pressure error presents a good behavior even in the very large σ case.

Remark 5.1 *Similar results could be obtained by using SDFEM method, but after a very accurate selection of δ. To illustrate this fact, we consider the viscosity $\nu = 10^{-3}$, use a mesh of 7200 P^1/P^1 elements; we use two values for δ: first, the (inappropriate for this mesh and viscosity) $\delta = \frac{1}{24}$ (in order to recover the same stabilization parameter used in Stokes flow stabilization), and secondly, the value $\delta = 10^{-4}$, correct for this mesh and viscosity. We have the following behavior of the error:*

σ	$\delta = \frac{1}{24}$		$\delta = 10^{-4}$	
	$\|\mathbf{u}-\mathbf{u}_h\|_{1,\Omega}$	$\|p-p_h\|_{0,\Omega}$	$\|\mathbf{u}-\mathbf{u}_h\|_{1,\Omega}$	$\|p-p_h\|_{0,\Omega}$
0.1	0.344192	4.402777×10^{-3}	0.347883	5.55026×10^{-3}
1	0.344380	4.427963×10^{-3}	0.348078	5.53989×10^{-3}
10	0.344339	4.732228×10^{-3}	0.347782	5.640657×10^{-3}
100	0.342894	1.444502×10^{-2}	0.345776	5.960951×10^{-3}
10^3	0.341392	0.125989	0.343405	1.040834×10^{-2}
10^4	0.341022	2.658054	0.342287	2.976455×10^{-2}

Table 4: SDFEM method Finite Element error when σ grows.

We remark that a bad choice of δ in SDFEM method has a direct impact on the behavior of the finite element error, as is shown in Table 4, where the error on the pressure explodes dramatically with σ for the inappropriate case (the error on the velocity remains bounded). We see then that the restriction $\delta \leq \dfrac{9}{25000}$ given by Remark 3.3 (in order fit the stablity condition for all the range of tested values for σ) is not only a theoretical restriction, but a real numerical restriction. $\square$

5.2 The lid-driven cavity problem

In this section we treat the lid driven cavity problem and compare the results from present method with the solution obtained by the SDFEM method. The problem statement is as in Figure 2, using $\mathbf{f} = \mathbf{0}$. First, in Figure 3, we plot pressure contours obtained for the case $\sigma = 10^2, \nu = 10^{-2}$. Pressure contours for $\sigma = 10^3, \nu = 10^{-3}$ and $\sigma = 10^4, \nu = 10^{-4}$ are plotted in Figures 4 and 5, respectively. We see that no oscillations appear on the pressure, even considering a very high quotient $\dfrac{\sigma\, h_K^2}{\nu}$. In all the calculations a uniform mesh of 441 P^1/P^1 elements was used.

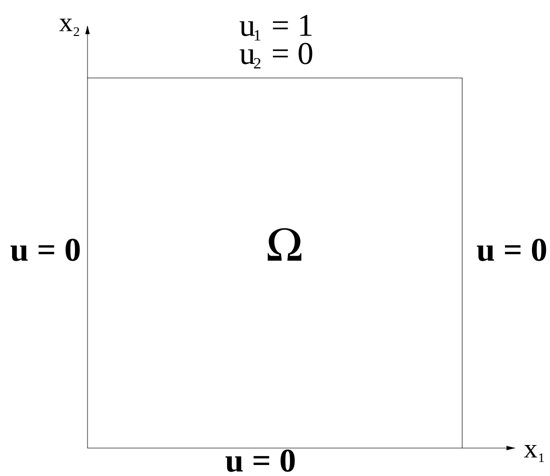

Figure 2: Statement of the problem

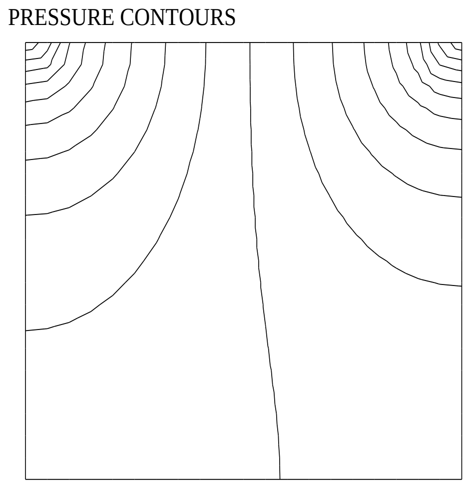

Figure 3: Pressure contours for $\sigma = 10^2, \nu = 10^{-2}$

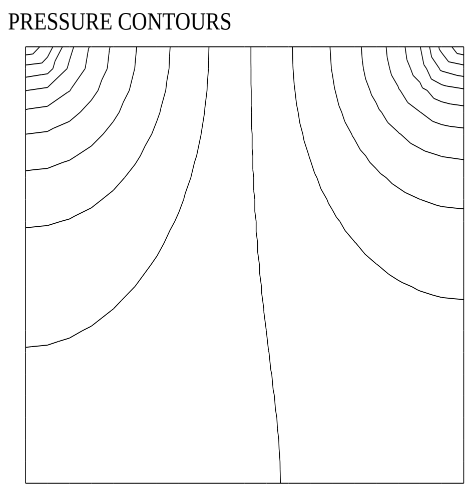

Figure 4: Pressure contours for $\sigma = 10^3, \nu = 10^{-3}$

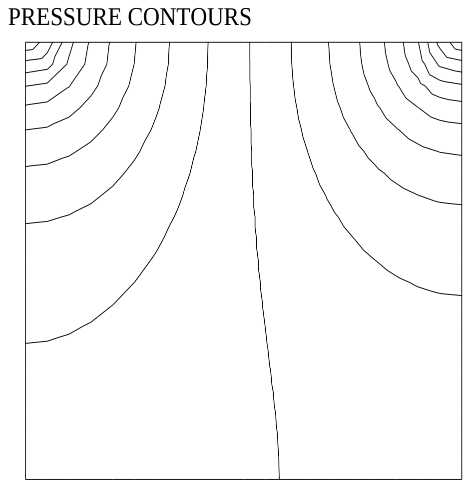

Figure 5: Pressure contours for $\sigma = 10^4, \nu = 10^{-4}$

Finally, in Figure 6 we show a vertical cut of the tangential velocity of the lid driven cavity problem with $\sigma = 10^3$ and $\nu = 10^{-3}$, comparing the solutions of present method and SDFEM method. We observe the presence of a boundary layer on the velocity that is well recovered by our method, and that the SDFEM method fails to capture. The calculation has been made using $\delta = 0.0001$ for the SDFEM method.

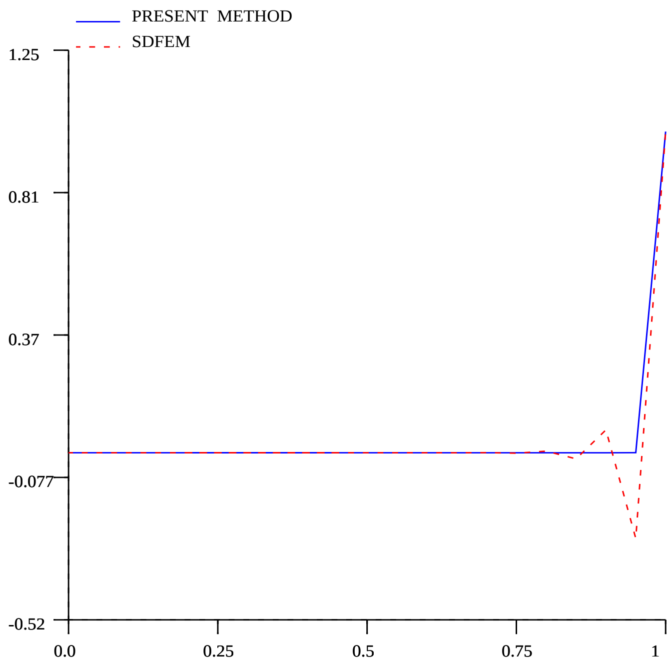

Figure 6: Vertical cut of tangential velocity at $x_1 = 0.5$

Acknowledgement: The authors would like to thank Professor Patrick Le Tallec from Ecole Polytechnique, France, for his helpful discussions and comments.

References

[1] D. Arnold, F. Brezzi, and M. Fortin. A stable finite element for the Stokes equations. *Calcolo*, 21(4):337–344, 1984.

[2] B. Baiocchi, F. Brezzi, and L.P. Franca. Virtual bubbles and Galerkin-least-squares type methods (Ga. L. S.). *Comput. Methods Appl. Mech. Engrg.*, 105:125–141, 1993.

[3] R. Bank and B. Welfert. A comparison between the MINI-Element and the Petrov-Galerkin formulations for the Generalized Stokes problem. *Comput. Methods Appl. Mech. Engrg.*, 83:61–68, 1990.

[4] F. Brezzi, M.O. Bristeau, L. Franca, M. Mallet, and G. Rogé. A relationship between stabilized finite element methods and the Galerkin method with bubble functions. *Comput. Methods Appl. Mech. Engrg.*, 96:117–129, 1992.

[5] F. Brezzi and M. Fortin. *Mixed and Hybrid Finite Element Methods.* Springer-Verlag, 1991.

[6] F. Brezzi and J. Douglas Jr. Stabilized mixed methods for the Stokes problem. *Numer. Math.*, 53:225–235, 1988.

[7] F. Brezzi and J. Pitkaranta. On the stablization of finite element approximations of the Stokes problem. In W. Hackbush, editor, *Efficient Solution of Elliptic Systems*, volume 10 of *Notes on Numerical Fluid Mechanics*, pages 11–19. Vieweg, 1984.

[8] Ph. G. Ciarlet. *The Finite Element Method for Elliptic Problems.* North Holland, 1978.

[9] P. Clément. Approximation by finite element functions using local regularization. *R.A.I.R.O. Anal. Numer.*, 9:77–84, 1975.

[10] L.P. Franca and C. Farhat. Bubble functions prompt unusual stabilized finite element methods. *Comput. Methods Appl. Mech. Engrg.*, 123:299–308, 1995.

[11] L.P. Franca and R. Stenberg. Error analysis of some Galerkin least squares methods for the elasticity equations. *SIAM J. Numer. Anal.*, 28(6):1680–1697, 1991.

[12] L.P. Franca and F. Valentin. On an improved unusual stabilized finite element method for the advective-reactive-diffusive equation. *Comput. Methods Appl. Mech. Engrg.*, 190:1785–1800, 2000.

[13] V. Girault and P.A. Raviart. *Finite Element Methods for the Navier-Stokes Equations.* Springer-Verlag, 1986.

[14] I. Harari and T.J.R. Hughes. What are C and h?: Inequalities for the analysis and design of finite element methods. *Comput. Methods Appl. Mech. Engrg.*, 97(2):157–192, 1992.

[15] I. Harari and T.J.R. Hughes. Stabilized finite element methods for steady advection-diffusion with production. *Comput. Methods Appl. Mech. Engrg.*, 115:165–191, 1994.

[16] G. Hauke. A simple subgrid scale stabilized method for the advection-diffusion-reaction equation. Preprint, 2000.

[17] T.J.R. Hughes and L.P. 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. Methods Appl. Mech. Engrg.*, 65(1):85–96, 1987.

[18] T.J.R. Hughes, L.P. Franca, and M. Balestra. A new finite element formulation for computational fluids dynamics: V. Circumventing the Babuska-Brezzi condition : A stable Petrov-Galerkin formulation of the Stokes problem accommodating equal-order interpolations. *Comput. Methods Appl. Mech. Engrg.*, 59(1):85–99, 1986.

[19] R. Pierre. Simple C^0 approximations for the computation of incompressible flows. *Comput. Methods Appl. Mech. Engrg.*, 68:205–227, 1988.

[20] R. Stenberg. Error analysis of some finite element methods for the Stokes problem. *Math. Comp.*, 54:495–508, 1990.

[21] L. Tobiska and R. Verfürth. Analysis of a streamline diffusion finite element method for the Stokes and Navier-Stokes equations. *SIAM J. Numer. Anal.*, 33(1):107–127, 1996.

Unité de recherche INRIA Rocquencourt
Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38330 Montbonnot-St-Martin (France)
Unité de recherche INRIA Sophia Antipolis : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399