IMA Journal of Numerical Analysis (2016) **36**, 947–983
doi:10.1093/imanum/drv015
Advance Access publication on April 22, 2015

A priori and *a posteriori* error analysis of a pseudostress-based mixed formulation of the Stokes problem with varying density

Sergio Caucao
Departamento de Matemática, Facultad de Ciencias, Universidad del Bío-Bío, Casilla 5-C, Concepción, Chile
scaucao@ubiobio.cl

and

David Mora and Ricardo Oyarzúa*
GIMNAP-Departamento de Matemática, Facultad de Ciencias, Universidad del Bío-Bío, Casilla 5-C, Concepción, Chile and Centro de Investigación en Ingeniería Matemática (CI²MA), Universidad de Concepción, Concepción, Chile
*Corresponding author: royarzua@ubiobio.cl dmora@ubiobio.cl

[Received on 14 August 2014; revised on 16 March 2015]

We propose and analyse a mixed finite element method for the nonstandard pseudostress–velocity formulation of the Stokes problem with varying density ρ in $\mathbb{R}^d$, $d \in \{2, 3\}$. Since the resulting variational formulation does not have the standard dual-mixed structure, we reformulate the continuous problem as an equivalent fixed-point problem. Then, we apply the classical Babuška–Brezzi theory to prove that the associated mapping $\mathbb{T}$ is well defined, and assuming that $\|\nabla\rho/\rho\|_{\mathbf{L}^\infty(\Omega)}$ is sufficiently small, we show that $\mathbb{T}$ is a contraction mapping, which implies that the variational formulation is well posed. Under the same hypothesis on ρ we prove stability of the continuous problem. Next, adapting the arguments of the continuous analysis to the discrete case, we establish suitable hypotheses on the finite element subspaces ensuring that the associated Galerkin scheme becomes well posed. A feasible choice of subspaces is given by Raviart–Thomas elements of order $k \geqslant 0$ for the pseudostress and polynomials of degree k for the velocity. In addition, we derive a reliable and efficient residual-based *a posteriori* error estimator for the problem. The proof of reliability makes use of the global inf–sup condition, Helmholtz decompositions, and local approximation properties of the Clément interpolant and Raviart–Thomas operator. On the other hand, inverse inequalities, the localization technique based on element-bubble and edge-bubble functions, approximation properties of the L^2-orthogonal projector and known results from previous works are the main tools for proving the efficiency of the estimator. Finally, several numerical results illustrating the performance of the mixed finite element method, confirming the theoretical rate of convergence and the theoretical properties of the estimator, and showing the behaviour of the associated adaptive algorithms are reported.

Keywords: Stokes problem; varying density; pseudostress–velocity formulation; mixed finite elements; *a priori* error analysis; efficiency; reliability; *a posteriori* error analysis.

1. Introduction

The numerical simulation of incompressible fluid flow problems, modelled by the Stokes equations, has been widely studied during the last decades. Different formulations (velocity–pressure, vorticity–velocity–pressure and pseudostress–velocity, among others) and different numerical methods (conforming and nonconforming methods) have been introduced and analysed, all of them with different advantages and disadvantages. In particular, the study of numerical methods for the stress- and pseudostress-based formulations for the Stokes problem has become a very active research area during the last decade (see, e.g., Figueroa *et al.*, 2008a,b; Cai *et al.*, 2009; Gatica *et al.*, 2010, 2011a,b, 2012, 2014a), motivated by the fact that they provide a direct approximation of the stress or pseudostress tensor (besides the approximation of the velocity and/or pressure). In particular, in the case of the pseudostress–velocity formulation, it is possible to compute the other physical quantities of interest such as the pressure, velocity gradient, stress and vorticity, in terms of the pseudostress, and can all be approximated with the same accuracy as the pseudostress, applying a simple post-processing procedure. Moreover, these kinds of formulations have a natural applicability to non-Newtonian flows. Indeed, since in this case the constitutive equation is nonlinear, the stress cannot be eliminated, and hence it becomes an unavoidable unknown in the corresponding solvability analysis. Actually, the main advantage of this formulation is that it allows for a unified analysis for linear and nonlinear flows. Moreover, these kinds of formulations have also been extended to the Navier–Stokes equations and multiphysics problems, such as the coupling of fluid flow with porous media flow modelled by the Stokes–Darcy coupled problem (see, e.g., Gatica *et al.*, 2004, 2011c; Cai & Wang, 2010; Cai & Zhang, 2012).

Now, concerning the fluid flow problem studied in this paper, the first work in studying conforming finite element methods for the Stokes problem with varying density is Bernardi *et al.* (1992), where the authors propose and analyse two variational formulations to solve the fluid flow problem. The first one is a velocity–pressure formulation which yields a nonsymmetric saddle-point formulation, whereas the second one is a momentum–pressure formulation which yields a standard saddle-point formulation. Well-posedness of the velocity–pressure formulation is analysed by using a generalization of the Babuška–Brezzi theory introduced in Nicolaides (1982) (see also Bernardi *et al.*, 1988), whereas the classical Babuška–Brezzi theory is applied to prove well-posedness of the momentum–pressure formulation. It is important to note that, in both cases, existence and uniqueness of solution of the continuous and discrete problems are attained by assuming that the variation of the density is not too large. Under similar assumptions, in Ern (1998) the well-posedness of the vorticity–velocity formulation of the Stokes problem with varying density and viscosity has been analysed and the equivalence of the vorticity–velocity and velocity–pressure formulations in appropriate functional spaces has been proved.

In this paper we adapt the results in Gatica *et al.* (2012), and introduce and analyse a pseudostress–velocity formulation for the Stokes problem with varying density which was analysed in Bernardi *et al.* (1992). Since the resulting variational formulation does not have the standard dual-mixed structure, we reformulate the continuous problem as an equivalent fixed-point problem. Then, we apply the classical Babuška–Brezzi theory to prove that the associated mapping $\mathbb{T}$ is well defined, and assuming that $\|\nabla\rho/\rho\|_{\mathbf{L}^\infty(\Omega)}$ is sufficiently small, we show that $\mathbb{T}$ is a contraction mapping, which implies that the variational formulation is well posed. We observe that this assumption is consistent with the approach in Bernardi *et al.* (1992). Next, we adapt the theory developed for the continuous problem to the discrete case, and derive sufficient conditions on the finite element subspaces ensuring that the associated Galerkin scheme becomes well posed.

Next, we derive a reliable and efficient residual-based *a posteriori* error estimator for the mixed problem. We observe here that it is well known that in order to guarantee good convergence behaviour of most finite element solutions, especially under the eventual presence of singularities, one usually needs to apply an adaptive algorithm based on *a posteriori* error estimates. These are represented by global quantities Θ that are expressed in terms of local indicators Θ_T defined on each element T of a given triangulation $\mathcal{T}_h$. The estimator Θ is said to be efficient (respectively, reliable) if there exists

$C_{\text{eff}} > 0$ (respectively, $C_{\text{rel}} > 0$), independent of the mesh sizes, such that

$$C_{\text{eff}}\Theta + \text{h.o.t.} \leqslant \|\text{error}\| \leqslant C_{\text{rel}}\Theta + \text{h.o.t.},$$

where h.o.t. is a generic expression denoting one or several terms of higher order.

The rest of this work is organized as follows. In Section 2 we introduce the model problem and derive the mixed variational formulation. In Section 3 we analyse the well-posedness of the continuous problem. For the existence and uniqueness of solution we introduce an equivalent fixed-point problem and we prove, assuming that $\|\nabla\rho/\rho\|_{\mathbf{L}^\infty(\Omega)}$ is sufficiently small, that it has a unique solution. Under a similar assumption we prove that the solution is stable. Next, in Section 4 we define the Galerkin scheme and derive general hypotheses on the finite element subspaces ensuring that, on the one hand, the discrete scheme becomes well posed, and on the other hand, it satisfies a Céa's estimate. Specific choices of finite element subspaces satisfying these assumptions are introduced in Section 5. In Section 6 we develop the *a posteriori* error analysis. We employ the global continuous inf–sup condition, Helmholtz decomposition, the local approximation properties of the Clément and Raviart–Thomas operators, and assume that $\|\nabla\rho/\rho\|_{0,\Omega}$ is sufficiently small to derive a reliable residual-based *a posteriori* error estimator. On the other hand, in Section 6.2 we apply inverse inequalities, the localization technique based on element-bubble and edge-bubble functions, and approximation properties of the L^2-orthogonal projector to prove the efficiency of the estimator. Finally, several numerical results, illustrating the performance of the proposed mixed finite element method, confirming the reliability and efficiency of the *a posteriori* estimators, and showing the good behaviour of the associated adaptive algorithms, are provided in Section 7.

2. Continuous problem

In this section we introduce and analyse a weak dual-mixed formulation for the Stokes problem with varying density analysed in Bernardi *et al.* (1992). In particular, we discuss existence, uniqueness and stability of solution. We start by introducing some definitions and fixing some notation.

2.1 *Preliminaries*

Given a vector field $\mathbf{v} := (v_1, \ldots, v_d)$ and a tensor field $\boldsymbol{\tau} := (\tau_{ij})_{i,j=1,\ldots,d}$, with $d = 2, 3$, we define the operators

$$\nabla\mathbf{v} = \left(\frac{\partial v_i}{\partial x_j}\right), \quad \text{and} \quad \mathbf{div}\,\boldsymbol{\tau} = (\text{div}\,(\tau_{i1}, \ldots, \tau_{id})),$$

where div is the usual divergence operator acting on vector fields.

Now, let $\mathcal{O}$ be a domain in $\mathbf{R}^d$, with Lipschitz boundary Γ. For $r \geqslant 0$ and $p \in [1, \infty]$, we denote by $L^p(\mathcal{O})$ and $W^{r,p}(\mathcal{O})$ the usual Lebesgue and Sobolev spaces endowed with the norms $\|\cdot\|_{L^p(\mathcal{O})}$ and $\|\cdot\|_{W^{r,p}(\mathcal{O})}$, respectively.

Note that $W^{0,p}(\mathcal{O}) = L^p(\mathcal{O})$. If $p = 2$, we write $H^r(\mathcal{O})$ in place of $W^{r,2}(\mathcal{O})$, and denote the corresponding Lebesgue and Sobolev norms by $\|\cdot\|_{0,\mathcal{O}}$ and $\|\cdot\|_{r,\mathcal{O}}$, respectively. We define

$$\mathbf{H}^r(\mathcal{O}) := [H^r(\mathcal{O})]^d, \quad \mathbb{H}^r(\mathcal{O}) := [H^r(\mathcal{O})]^{d\times d}.$$

Also, we shall make use of the Hilbert space

$$\mathbf{H}(\text{div}\,; \mathcal{O}) := \{\mathbf{w} \in \mathbf{L}^2(\mathcal{O}) : \text{div}\,\mathbf{w} \in L^2(\mathcal{O})\},$$

which is standard in the realm of mixed problems (see Boffi *et al.*, 2013, Section 1.2 or Girault & Raviart, 1986, Chapter 1, Section 2.2 for instance). This space is endowed with the norm

$$\|\mathbf{w}\|^2_{\mathrm{div},\mathcal{O}} = \|\mathbf{w}\|^2_{0,\mathcal{O}} + \|\mathrm{div}\,\mathbf{w}\|^2_{0,\mathcal{O}}.$$

The space of matrix-valued functions whose rows belong to $\mathbf{H}(\mathrm{div}\,;\mathcal{O})$ will be denoted by $\mathbb{H}(\mathbf{div};\mathcal{O})$ and endowed with the norm $\|\cdot\|_{\mathbf{div},\mathcal{O}}$, which can be characterized as

$$\mathbb{H}(\mathbf{div};\mathcal{O}) := \{\boldsymbol{\tau} \in \mathbb{L}^2(\mathcal{O}) : \mathbf{c}^t\boldsymbol{\tau} \in \mathbf{H}(\mathrm{div}\,;\mathcal{O})\ \forall \mathbf{c} \in \mathbf{R}^d\}.$$

Note that if $\boldsymbol{\tau} \in \mathbb{H}(\mathbf{div};\mathcal{O})$, then $\mathbf{div}\,\boldsymbol{\tau} \in \mathbf{L}^2(\mathcal{O})$.

Next, for the sake of simplicity, we will also use the notation

$$(u,v)_\Omega := \int_\Omega uv, \quad (\mathbf{u},\mathbf{v})_\Omega := \int_\Omega \mathbf{u}\cdot\mathbf{v}, \quad (\boldsymbol{\sigma},\boldsymbol{\tau})_\Omega := \int_\Omega \boldsymbol{\sigma}:\boldsymbol{\tau},$$

where $\boldsymbol{\sigma}:\boldsymbol{\tau} = \mathrm{tr}(\boldsymbol{\sigma}^t\boldsymbol{\tau}) = \sum_{i,j=1}^d \sigma_{ij}\tau_{ij}$, with $\boldsymbol{\tau}^t = (\tau_{ji})$ and $\mathrm{tr}\,\boldsymbol{\tau} = \sum_{i=1}^d \tau_{ii}$, for any tensor $\boldsymbol{\sigma} = (\sigma_{ij})$ and $\boldsymbol{\tau} = (\tau_{ij})$. In addition, we denote by

$$\boldsymbol{\tau}^{\mathrm{D}} := \boldsymbol{\tau} - \frac{1}{d}\,\mathrm{tr}(\boldsymbol{\tau})I$$

the deviatoric part of the tensor $\boldsymbol{\tau}$, where I is the identity matrix in $\mathbf{R}^{d\times d}$. It is not difficult to see that $\mathrm{tr}(\boldsymbol{\tau}^{\mathrm{D}}) = 0$, which implies

$$(\boldsymbol{\sigma}^{\mathrm{D}},\boldsymbol{\tau})_\Omega = (\boldsymbol{\sigma}^{\mathrm{D}},\boldsymbol{\tau}^{\mathrm{D}})_\Omega, \tag{2.1}$$

for any tensors $\boldsymbol{\sigma}$ and $\boldsymbol{\tau}$. Also, there hold

$$\|\boldsymbol{\tau}^{\mathrm{D}}\|^2_{0,\Omega} = \|\boldsymbol{\tau}\|^2_{0,\Omega} - \frac{1}{d}\|\mathrm{tr}\,\boldsymbol{\tau}\|^2_{0,\Omega} \quad \text{and} \quad \|\mathrm{tr}\,\boldsymbol{\tau}\|_{0,\Omega} \leqslant \sqrt{d}\,\|\boldsymbol{\tau}\|_{0,\Omega}. \tag{2.2}$$

Furthermore, given a non-negative integer k, we denote by $P_k(\mathcal{O})$ the space of polynomials defined in $\mathcal{O}$ of degree $\leqslant k$.

In addition, it is easy to see that there holds

$$\mathbb{H}(\mathbf{div};\mathcal{O}) = \mathbb{H}_0(\mathbf{div};\mathcal{O}) \oplus P_0(\mathcal{O})I, \tag{2.3}$$

where

$$\mathbb{H}_0(\mathbf{div};\mathcal{O}) := \left\{\boldsymbol{\tau} \in \mathbb{H}(\mathbf{div};\mathcal{O}) : \int_\mathcal{O} \mathrm{tr}\,\boldsymbol{\tau} = 0\right\}. \tag{2.4}$$

More precisely, each $\boldsymbol{\tau} \in \mathbb{H}(\mathbf{div};\mathcal{O})$ can be decomposed uniquely as

$$\boldsymbol{\tau} = \boldsymbol{\tau}_0 + cI, \ \text{ with } \boldsymbol{\tau}_0 \in \mathbb{H}_0(\mathbf{div};\mathcal{O}) \quad \text{and} \quad c := \frac{1}{d|\mathcal{O}|}\int_\mathcal{O} \mathrm{tr}\,\boldsymbol{\tau} \in \mathbb{R}. \tag{2.5}$$

This decomposition will be utilized below to analyse the weak formulation of our problem.

We end this section by mentioning that, throughout the rest of the paper, we shall frequently use the notation C and c, with or without subscripts, bars, tildes or hats, to denote generic positive constants independent of the discretization parameters.

2.2 *Model problem*

In this paper we shall consider a viscous fluid occupying a bounded domain Ω in $\mathbf{R}^d$, $d = 2, 3$, with Lipschitz-continuous boundary $\Gamma = \partial\Omega$, governed by the Stokes equations with varying density:

$$\begin{aligned} \boldsymbol{\sigma} = \nu(\rho\nabla\mathbf{u}) - pI \ \text{ in } \Omega, \quad -\mathbf{div}\,\boldsymbol{\sigma} = \mathbf{f} \ \text{ in } \Omega, \\ \operatorname{div}(\rho\mathbf{u}) = 0 \ \text{ in } \Omega, \quad \mathbf{u} = 0 \ \text{ on } \Gamma, \quad (p, 1)_\Omega = 0. \end{aligned} \tag{2.6}$$

Here, the unknowns are the pseudostress tensor $\boldsymbol{\sigma}$, the fluid velocity $\mathbf{u}$ and the pressure p. The given data are the external force per unit mass $\mathbf{f} \in \mathbf{L}^2(\Omega)$, the viscosity $\nu > 0$, which is assumed to be constant, and the density function $\rho \in H^1(\Omega) \cap W^{1,\infty}(\Omega)$, satisfying

$$\frac{\nabla\rho}{\rho} \in \mathbf{L}^\infty(\Omega) \quad \text{and} \quad 0 < \rho_0 < \rho(x) < \rho_1, \ \text{a.e. in } \Omega, \tag{2.7}$$

where ρ_0 and ρ_1 are positive constants.

The model in (2.6), which is derived from the full steady Navier–Stokes equations for viscous fluids, is well justified if we make the following assumptions.

(i) Only the laminar case is considered and the second-order diffusion term in the viscous stress tensor is neglected.

(ii) The Mach number is small enough, which implies that the coupling between the pressure and the temperature can be neglected.

In particular, (ii) implies that the state law can be chosen as a simple equation linking the density and the temperature, in which the temperature is approximated by a reference one. This model has been applied in several applications in engineering, such as laminar combustion and vapour-phase epitaxy (for details, see Ern *et al.*, 1995, 1996, and the references therein).

Now, in order to rewrite equations (2.6) as a pseudostress–velocity formulation, we first observe that identity $\operatorname{div}(\rho\mathbf{u}) = 0$ in Ω implies

$$\rho \operatorname{div}\mathbf{u} = -\mathbf{u}\cdot\nabla\rho \quad \text{in } \Omega. \tag{2.8}$$

Then, observing that $\operatorname{tr}\boldsymbol{\sigma} = \nu\rho\operatorname{div}\mathbf{u} - dp$, (2.8) implies that the pressure can be written in terms of the pseudostress and the velocity as follows:

$$p = -\frac{1}{d}\left(\nu\mathbf{u}\cdot\nabla\rho + \operatorname{tr}\boldsymbol{\sigma}\right) \quad \text{in } \Omega. \tag{2.9}$$

In this way, we eliminate the pressure from (2.6) and obtain the following equivalent system of equations:

$$\begin{aligned} \frac{\nu^{-1}}{\rho}\boldsymbol{\sigma}^{\mathrm{D}} = \nabla\mathbf{u} + \frac{1}{d}\left(\mathbf{u}\cdot\frac{\nabla\rho}{\rho}\right)I \ \text{ in } \Omega, \quad -\mathbf{div}\,\boldsymbol{\sigma} = \mathbf{f} \ \text{ in } \Omega, \\ \mathbf{u} = 0 \quad \text{on } \Gamma, \quad (\operatorname{tr}\boldsymbol{\sigma}, 1)_\Omega = -\nu(\mathbf{u}\cdot\nabla\rho, 1)_\Omega. \end{aligned} \tag{2.10}$$

2.3 *Dual-mixed variational formulation*

Now we introduce the variational formulation of the model problem (2.10). To do that, we test equations (2.10) by suitable test functions, integrate by parts and use the homogeneous boundary condition and identity (2.1), to obtain the variational problem: find $(\boldsymbol{\sigma}, \mathbf{u}) \in \mathbb{H}(\mathbf{div};\Omega) \times \mathbf{L}^2(\Omega)$ such that $(\operatorname{tr}\boldsymbol{\sigma} + \nu\mathbf{u}\cdot\nabla\rho, 1)_\Omega = 0$ and

$$\begin{aligned} \nu^{-1}\left(\frac{1}{\rho}\boldsymbol{\sigma}^{\mathrm{D}}, \boldsymbol{\tau}^{\mathrm{D}}\right)_\Omega + (\mathbf{div}\,\boldsymbol{\tau}, \mathbf{u})_\Omega - \frac{1}{d}\left(\mathbf{u}\cdot\frac{\nabla\rho}{\rho}, \operatorname{tr}\boldsymbol{\tau}\right)_\Omega &= 0, \\ (\mathbf{div}\,\boldsymbol{\sigma}, \mathbf{v})_\Omega &= -(\mathbf{f}, \mathbf{v})_\Omega, \end{aligned} \tag{2.11}$$

for all $(\boldsymbol{\tau}, \mathbf{v}) \in \mathbb{H}(\mathbf{div};\Omega) \times \mathbf{L}^2(\Omega)$.

Let us now define the tensor

$$\boldsymbol{\sigma}_0 := \boldsymbol{\sigma} + \frac{\nu}{d|\Omega|}(\mathbf{u}\cdot\nabla\rho, 1)_\Omega I. \tag{2.12}$$

It is clear that

$$\boldsymbol{\sigma}_0 \in \mathbb{H}_0(\mathbf{div};\Omega) \quad \text{if and only if} \quad (\operatorname{tr}\boldsymbol{\sigma} + \nu\mathbf{u}\cdot\nabla\rho, 1)_\Omega = 0. \tag{2.13}$$

In this way, owing to (2.12) and (2.5), problem (2.11) can be reformulated equivalently as follows: find $(\boldsymbol{\sigma}_0, \mathbf{u}) \in \mathbb{H}_0(\mathbf{div};\Omega) \times \mathbf{L}^2(\Omega)$ such that

$$\begin{aligned} \nu^{-1}\left(\frac{1}{\rho}\boldsymbol{\sigma}_0^{\mathrm{D}}, \boldsymbol{\tau}^{\mathrm{D}}\right)_\Omega + (\mathbf{div}\,\boldsymbol{\tau}, \mathbf{u})_\Omega - \frac{1}{d}\left(\mathbf{u}\cdot\frac{\nabla\rho}{\rho}, \operatorname{tr}\boldsymbol{\tau}\right)_\Omega &= 0, \\ (\mathbf{div}\,\boldsymbol{\sigma}_0, \mathbf{v})_\Omega &= -(\mathbf{f}, \mathbf{v})_\Omega, \end{aligned} \tag{2.14}$$

for all $(\boldsymbol{\tau}, \mathbf{v}) \in \mathbb{H}_0(\mathbf{div};\Omega) \times \mathbf{L}^2(\Omega)$.

The following lemma establishes that problems (2.11) and (2.14) are in fact equivalent.

LEMMA 2.1 If $(\boldsymbol{\sigma}, \mathbf{u})$ is a solution of (2.11), then $(\boldsymbol{\sigma}_0, \mathbf{u}) := (\boldsymbol{\sigma} + (\nu/d|\Omega|)(\mathbf{u}\cdot\nabla\rho, 1)_\Omega I, \mathbf{u})$ is a solution of (2.14). Conversely, if $(\boldsymbol{\sigma}_0, \mathbf{u})$ is a solution of (2.14), then $(\boldsymbol{\sigma}, \mathbf{u}) := (\boldsymbol{\sigma}_0 - (\nu/d|\Omega|)(\mathbf{u}\cdot\nabla\rho, 1)_\Omega I, \mathbf{u})$ is a solution of (2.11).

Proof. The first assertion is evident. On the other hand, by testing the first equation of (2.14) with $\boldsymbol{\tau} := (\rho - (\rho, 1)_\Omega/|\Omega|)I \in \mathbb{H}_0(\mathbf{div};\Omega)$, it follows that $(\mathbf{u}\cdot\nabla\rho/\rho, 1)_\Omega = 0$, which implies the second assertion. □

As a consequence of the above, in what follows we focus on analysing problem (2.14).

3. Analysis of the continuous problem

In this section we analyse the well-posedness of problem (2.14), that is, we establish stability, existence and uniqueness of solution. In order to do that, we start by writing our problem in the classical variational setting and state the main properties of the bilinear forms involved.

3.1 *Variational formulation*

First, let us define the spaces

$$\mathbb{H} := \mathbb{H}(\mathbf{div};\Omega), \quad \mathbb{H}_0 := \mathbb{H}_0(\mathbf{div};\Omega), \quad \mathbf{Q} := \mathbf{L}^2(\Omega),$$

and the product norm

$$\|(\boldsymbol{\tau},\mathbf{v})\|_{\mathbb{H}\times\mathbf{Q}} := (\|\boldsymbol{\tau}\|^2_{\mathbf{div},\Omega} + \|\mathbf{v}\|^2_{0,\Omega})^{1/2}.$$

Then, defining the bilinear forms $a(\cdot,\cdot) : \mathbb{H}\times\mathbb{H}\to\mathbf{R}$, $b(\cdot,\cdot) : \mathbb{H}\times\mathbf{Q}\to\mathbf{R}$ and $c(\cdot,\cdot) : \mathbb{H}\times\mathbf{Q}\to\mathbf{R}$ as

$$a(\boldsymbol{\sigma},\boldsymbol{\tau}) := \nu^{-1}\left(\frac{1}{\rho}\boldsymbol{\sigma}^{\mathrm{D}},\boldsymbol{\tau}^{\mathrm{D}}\right)_{\Omega}, \quad b(\boldsymbol{\tau},\mathbf{v}) := (\mathbf{div}\,\boldsymbol{\tau},\mathbf{v})_{\Omega}, \quad c(\boldsymbol{\tau},\mathbf{v}) := \frac{1}{d}\left(\mathbf{v}\cdot\frac{\nabla\rho}{\rho},\mathrm{tr}\,\boldsymbol{\tau}\right)_{\Omega}, \tag{3.1}$$

the variational formulation (2.14) reads as follows: find $(\boldsymbol{\sigma}_0,\mathbf{u})\in\mathbb{H}_0\times\mathbf{Q}$ such that

$$\begin{aligned} a(\boldsymbol{\sigma}_0,\boldsymbol{\tau}) + b(\boldsymbol{\tau},\mathbf{u}) - c(\boldsymbol{\tau},\mathbf{u}) &= 0,\\ b(\boldsymbol{\sigma}_0,\mathbf{v}) &= -(\mathbf{f},\mathbf{v})_{\Omega}, \end{aligned} \tag{3.2}$$

for all $(\boldsymbol{\tau},\mathbf{v})\in\mathbb{H}_0\times\mathbf{Q}$.

It is clear that assumption (2.7), Hölder's inequality and (2.2) imply the continuity of these bilinear forms:

$$\begin{aligned} |a(\boldsymbol{\sigma},\boldsymbol{\tau})| &\leqslant \frac{1}{\nu\rho_0}\|\boldsymbol{\sigma}\|_{\mathbf{div},\Omega}\|\boldsymbol{\tau}\|_{\mathbf{div},\Omega}, \quad \boldsymbol{\sigma},\boldsymbol{\tau}\in\mathbb{H},\\ |b(\boldsymbol{\tau},\mathbf{v})| &\leqslant \|\boldsymbol{\tau}\|_{\mathbf{div},\Omega}\|\mathbf{v}\|_{0,\Omega}, \quad \boldsymbol{\tau}\in\mathbb{H},\ \mathbf{v}\in\mathbf{Q},\\ |c(\boldsymbol{\tau},\mathbf{v})| &\leqslant \frac{1}{\sqrt{d}}\left\|\frac{\nabla\rho}{\rho}\right\|_{\mathbf{L}^\infty(\Omega)}\|\boldsymbol{\tau}\|_{\mathbf{div},\Omega}\|\mathbf{v}\|_{0,\Omega}, \quad \boldsymbol{\tau}\in\mathbb{H},\ \mathbf{v}\in\mathbf{Q}. \end{aligned} \tag{3.3}$$

Furthermore, owing to the surjectivity of the divergence operator (see, for instance, Gatica, 2014, Section 2.4.1 or Boffi *et al.*, 2013, Section 4.2.5), it is well known that the bilinear form b satisfies the inf–sup condition:

$$\sup_{\boldsymbol{\tau}\in\mathbb{H}_0\setminus\mathbf{0}}\frac{b(\boldsymbol{\tau},\mathbf{v})}{\|\boldsymbol{\tau}\|_{\mathbf{div},\Omega}} \geqslant \beta\|\mathbf{v}\|_{0,\Omega} \quad \forall\mathbf{v}\in\mathbf{Q}. \tag{3.4}$$

Finally, the following inequality holds (see, for instance, Arnold *et al.*, 1984, Lemma 3.1 or Boffi *et al.*, 2013, Proposition 9.1.1):

$$C_a\|\boldsymbol{\tau}\|^2_{0,\Omega} \leqslant \|\boldsymbol{\tau}^{\mathrm{D}}\|^2_{0,\Omega} + \|\mathbf{div}\,\boldsymbol{\tau}\|^2_{0,\Omega} \quad \forall\boldsymbol{\tau}\in\mathbb{H}_0, \tag{3.5}$$

with C_a depending only on Ω. This inequality and assumption (2.7) imply the ellipticity of the bilinear form $a(\cdot,\cdot)$ on the subspace

$$\mathbb{K}_0 := \{\boldsymbol{\tau}\in\mathbb{H}_0 : \mathbf{div}\,\boldsymbol{\tau} = \mathbf{0} \text{ in } \Omega\},$$

that is,

$$a(\boldsymbol{\tau},\boldsymbol{\tau}) \geqslant \frac{C_a}{\nu\rho_1}\|\boldsymbol{\tau}\|^2_{\mathbf{div},\Omega} \quad \forall\boldsymbol{\tau}\in\mathbb{K}_0. \tag{3.6}$$

3.2 *Stability*

Now we establish the stability of (3.2).

LEMMA 3.1 Let $(\boldsymbol{\sigma}_0, \mathbf{u}) \in \mathbb{H}_0 \times \mathbf{Q}$ be a solution to (3.2). Assume that

$$C_{\text{dep}} \left\| \frac{\nabla \rho}{\rho} \right\|_{\mathbf{L}^\infty(\Omega)} \leqslant \frac{1}{2}, \tag{3.7}$$

with

$$C_{\text{dep}} := \frac{1}{\beta\sqrt{d}} \left(1 + 2\frac{\rho_1}{C_a \rho_0} \right).$$

Then, there exist constants $C_{\boldsymbol{\sigma}}$ and $C_{\mathbf{u}}$, depending only on the stability constants in (3.3)–(3.5), such that

$$\|\boldsymbol{\sigma}_0\|_{\mathbf{div},\Omega} \leqslant C_{\boldsymbol{\sigma}} \|\mathbf{f}\|_{0,\Omega} \quad \text{and} \quad \|\mathbf{u}\|_{0,\Omega} \leqslant C_{\mathbf{u}} \|\mathbf{f}\|_{0,\Omega}. \tag{3.8}$$

(Explicit expressions for $C_{\boldsymbol{\sigma}}$ and $C_{\mathbf{u}}$ can be found in (3.13) and (3.14).)

Proof. Let $(\boldsymbol{\sigma}_0, \mathbf{u}) \in \mathbb{H}_0 \times \mathbf{Q}$ be a solution to (3.2). First, we observe that from the second equation of (3.2), it is easy to conclude that $\mathbf{div}\, \boldsymbol{\sigma}_0 = -\mathbf{f}$, which implies

$$\|\mathbf{div}\, \boldsymbol{\sigma}_0\|_{0,\Omega} = \|\mathbf{f}\|_{0,\Omega}. \tag{3.9}$$

Now, from the inf–sup condition in (3.4), the first equation of (3.2), Hölder's inequality, the inequality in (2.2), and the continuity of the bilinear forms a and c in (3.3), we observe that

$$\begin{aligned} \|\mathbf{u}\|_{0,\Omega} &\leqslant \frac{1}{\beta} \sup_{\boldsymbol{\tau} \in \mathbb{H}_0 \setminus \mathbf{0}} \frac{b(\boldsymbol{\tau}, \mathbf{u})}{\|\boldsymbol{\tau}\|_{\mathbf{div},\Omega}} = \frac{1}{\beta} \sup_{\boldsymbol{\tau} \in \mathbb{H}_0 \setminus \mathbf{0}} \frac{-a(\boldsymbol{\sigma}_0, \boldsymbol{\tau}) + c(\boldsymbol{\tau}, \mathbf{u})}{\|\boldsymbol{\tau}\|_{\mathbf{div},\Omega}} \\ &\leqslant \frac{1}{\nu\rho_0\beta} \|\boldsymbol{\sigma}_0\|_{\mathbf{div},\Omega} + \frac{1}{\beta\sqrt{d}} \left\| \frac{\nabla \rho}{\rho} \right\|_{\mathbf{L}^\infty(\Omega)} \|\mathbf{u}\|_{0,\Omega}. \end{aligned} \tag{3.10}$$

Then, owing to assumption (3.7), we obtain

$$\|\mathbf{u}\|_{0,\Omega} \leqslant \frac{2}{\nu\rho_0\beta} \|\boldsymbol{\sigma}_0\|_{\mathbf{div},\Omega}. \tag{3.11}$$

On the other hand, from the first equation of (3.2) with $\boldsymbol{\tau} = \boldsymbol{\sigma}_0$, there holds

$$a(\boldsymbol{\sigma}_0, \boldsymbol{\sigma}_0) = -b(\boldsymbol{\sigma}_0, \mathbf{u}) + c(\boldsymbol{\sigma}_0, \mathbf{u}) = (\mathbf{f}, \mathbf{u})_\Omega + c(\boldsymbol{\sigma}_0, \mathbf{u}),$$

which, together with assumption (2.7), the continuity of the bilinear form $c(\cdot,\cdot)$ in (3.3) and Hölder's inequality, implies

$$\|\boldsymbol{\sigma}_0^{\mathrm{D}}\|_{0,\Omega}^2 \leqslant \nu\rho_1 \|\mathbf{u}\|_{0,\Omega} \|\mathbf{f}\|_{0,\Omega} + \frac{\nu\rho_1}{\sqrt{d}} \|\mathbf{u}\|_{0,\Omega} \left\| \frac{\nabla \rho}{\rho} \right\|_{\mathbf{L}^\infty(\Omega)} \|\boldsymbol{\sigma}_0\|_{\mathbf{div},\Omega}. \tag{3.12}$$

Hence, adding $(1 + C_a)\|\mathbf{div}\,\boldsymbol{\sigma}_0\|^2_{0,\Omega}$ to both sides of (3.12), and using (3.5), (3.9) and (3.11), we obtain

$$\|\boldsymbol{\sigma}_0\|^2_{\mathbf{div},\Omega} \leqslant \frac{\nu\rho_1}{C_a}\|\mathbf{u}\|_{0,\Omega}\left(\|\mathbf{f}\|_{0,\Omega} + \frac{1}{\sqrt{d}}\left\|\frac{\nabla\rho}{\rho}\right\|_{\mathbf{L}^\infty(\Omega)}\|\boldsymbol{\sigma}_0\|_{\mathbf{div},\Omega}\right) + \frac{(1+C_a)}{C_a}\|\boldsymbol{\sigma}_0\|_{\mathbf{div},\Omega}\|\mathbf{f}\|_{0,\Omega}$$
$$\leqslant \left(\frac{2\rho_1}{C_a\rho_0\beta} + \frac{1+C_a}{C_a}\right)\|\boldsymbol{\sigma}_0\|_{\mathbf{div},\Omega}\|\mathbf{f}\|_{0,\Omega} + \frac{2\rho_1}{C_a\rho_0\beta\sqrt{d}}\left\|\frac{\nabla\rho}{\rho}\right\|_{\mathbf{L}^\infty(\Omega)}\|\boldsymbol{\sigma}_0\|^2_{\mathbf{div},\Omega}.$$

In this way, from assumption (3.7) it follows that

$$\|\boldsymbol{\sigma}_0\|_{\mathbf{div},\Omega} \leqslant 2\left(\frac{2\rho_1}{C_a\rho_0\beta} + \frac{1+C_a}{C_a}\right)\|\mathbf{f}\|_{0,\Omega}, \tag{3.13}$$

which together with (3.11) implies

$$\|\mathbf{u}\|_{0,\Omega} \leqslant \frac{4}{\nu\rho_0\beta}\left(\frac{2\rho_1}{C_a\rho_0\beta} + \frac{1+C_a}{C_a}\right)\|\mathbf{f}\|_{0,\Omega}, \tag{3.14}$$

which completes the proof. □

3.3 *Existence and uniqueness of solution*

As mentioned before, in order to prove the existence and uniqueness of solution, we now introduce the linear mapping

$$\mathbb{T} : (\boldsymbol{\xi}, \mathbf{z}) \in \mathbb{H}_0 \times \mathbf{Q} \to (\boldsymbol{\sigma}_0, \mathbf{u}) \in \mathbb{H}_0 \times \mathbf{Q}$$

as the solution to the following variation of problem (3.2): find $(\boldsymbol{\sigma}_0, \mathbf{u}) \in \mathbb{H}_0 \times \mathbf{Q}$ such that

$$\begin{aligned} a(\boldsymbol{\sigma}_0, \boldsymbol{\tau}) + b(\boldsymbol{\tau}, \mathbf{u}) &= c(\boldsymbol{\tau}, \mathbf{z}), \\ b(\boldsymbol{\sigma}_0, \mathbf{v}) &= -(\mathbf{f}, \mathbf{v})_\Omega, \end{aligned} \tag{3.15}$$

for all $(\boldsymbol{\tau}, \mathbf{v}) \in \mathbb{H}_0 \times \mathbf{Q}$. With the stability properties in Section 3.1, it is not difficult to see that problem (3.15) is uniquely solvable, and hence the operator $\mathbb{T}$ is well defined (see Gatica *et al.*, 2012, Theorem 2.1).

The following lemma establishes that $\mathbb{T}$ is a contraction mapping and hence, according to the Banach fixed-point theorem, it has a unique fixed point in $\mathbb{H}_0 \times \mathbf{Q}$.

LEMMA 3.2 Assume that

$$C_{\mathbb{T}}\left\|\frac{\nabla\rho}{\rho}\right\|_{\mathbf{L}^\infty(\Omega)} < 1, \tag{3.16}$$

with

$$C_{\mathbb{T}} := \frac{1}{\beta\sqrt{d}}\left(1 + \frac{\rho_1}{C_a\rho_0}\right) + \frac{\rho_1\nu}{C_a\sqrt{d}}. \tag{3.17}$$

Then, $\mathbb{T}$ is a contraction mapping in $\mathbb{H}_0 \times \mathbf{Q}$.

Proof. Let $(\boldsymbol{\sigma}_1, \mathbf{u}_1)$, $(\boldsymbol{\sigma}_2, \mathbf{u}_2)$, $(\boldsymbol{\xi}_1, \mathbf{z}_1)$, $(\boldsymbol{\xi}_2, \mathbf{z}_2)$ in $\mathbb{H}_0 \times \mathbf{Q}$, such that

$$\mathbb{T}(\boldsymbol{\xi}_1, \mathbf{z}_1) = (\boldsymbol{\sigma}_1, \mathbf{u}_1) \quad \text{and} \quad \mathbb{T}(\boldsymbol{\xi}_2, \mathbf{z}_2) = (\boldsymbol{\sigma}_2, \mathbf{u}_2).$$

From the definition of $\mathbb{T}$ in (3.15), it follows that

$$\begin{aligned} a(\boldsymbol{\sigma}_1 - \boldsymbol{\sigma}_2, \boldsymbol{\tau}) + b(\boldsymbol{\tau}, \mathbf{u}_1 - \mathbf{u}_2) &= c(\boldsymbol{\tau}, \mathbf{z}_1 - \mathbf{z}_2), \\ b(\boldsymbol{\sigma}_1 - \boldsymbol{\sigma}_2, \mathbf{v}) &= 0, \end{aligned} \tag{3.18}$$

for all $(\boldsymbol{\tau}, \mathbf{v})$ in $\mathbb{H}_0 \times \mathbf{Q}$, which implies

$$\mathbf{div}(\boldsymbol{\sigma}_1 - \boldsymbol{\sigma}_2) = 0, \tag{3.19}$$

and

$$a(\boldsymbol{\sigma}_1 - \boldsymbol{\sigma}_2, \boldsymbol{\sigma}_1 - \boldsymbol{\sigma}_2) = c(\boldsymbol{\sigma}_1 - \boldsymbol{\sigma}_2, \mathbf{z}_1 - \mathbf{z}_2). \tag{3.20}$$

Then, from (3.19), (3.20), the ellipticity of $a(\cdot,\cdot)$ on $\mathbb{K}_0$ in (3.6) and the continuity of c in (3.3), there holds

$$\|\boldsymbol{\sigma}_1 - \boldsymbol{\sigma}_2\|_{\mathbf{div},\Omega} \leqslant \frac{\rho_1 \nu}{C_a \sqrt{d}} \left\| \frac{\nabla \rho}{\rho} \right\|_{\mathbf{L}^\infty(\Omega)} \|\mathbf{z}_1 - \mathbf{z}_2\|_{0,\Omega}. \tag{3.21}$$

Now, from (3.3), (3.4) and the first equation of (3.18), we obtain

$$\begin{aligned} \|\mathbf{u}_1 - \mathbf{u}_2\|_{0,\Omega} &\leqslant \frac{1}{\beta} \sup_{\boldsymbol{\tau} \in \mathbb{H}_0 \setminus \mathbf{0}} \frac{|b(\boldsymbol{\tau}, \mathbf{u}_1 - \mathbf{u}_2)|}{\|\boldsymbol{\tau}\|_{\mathbf{div},\Omega}} \\ &= \frac{1}{\beta} \sup_{\boldsymbol{\tau} \in \mathbb{H}_0 \setminus \mathbf{0}} \frac{|c(\boldsymbol{\tau}, \mathbf{z}_1 - \mathbf{z}_2) - a(\boldsymbol{\sigma}_1 - \boldsymbol{\sigma}_1, \boldsymbol{\tau})|}{\|\boldsymbol{\tau}\|_{\mathbf{div},\Omega}} \\ &\leqslant \frac{1}{\beta \sqrt{d}} \left\| \frac{\nabla \rho}{\rho} \right\|_{\mathbf{L}^\infty(\Omega)} \|\mathbf{z}_1 - \mathbf{z}_2\|_{0,\Omega} + \frac{1}{\nu \rho_0 \beta} \|\boldsymbol{\sigma}_1 - \boldsymbol{\sigma}_2\|_{\mathbf{div},\Omega}, \end{aligned}$$

which together with (3.21) implies

$$\|\mathbf{u}_1 - \mathbf{u}_2\|_{0,\Omega} \leqslant \frac{1}{\beta \sqrt{d}} \left(1 + \frac{\rho_1}{C_a \rho_0}\right) \left\| \frac{\nabla \rho}{\rho} \right\|_{\mathbf{L}^\infty(\Omega)} \|\mathbf{z}_1 - \mathbf{z}_2\|_{0,\Omega}. \tag{3.22}$$

In this way, from (3.21) and (3.22), there holds

$$\begin{aligned} \|\mathbb{T}(\boldsymbol{\xi}_1, \mathbf{z}_1) - \mathbb{T}(\boldsymbol{\xi}_2, \mathbf{z}_2)\|_{\mathbb{H} \times \mathbf{Q}} &\leqslant \|\boldsymbol{\sigma}_1 - \boldsymbol{\sigma}_2\|_{\mathbf{div},\Omega} + \|\mathbf{u}_1 - \mathbf{u}_2\|_{0,\Omega} \\ &\leqslant C_{\mathbb{T}} \left\| \frac{\nabla \rho}{\rho} \right\|_{\mathbf{L}^\infty(\Omega)} \|\mathbf{z}_1 - \mathbf{z}_2\|_{0,\Omega} \\ &\leqslant C_{\mathbb{T}} \left\| \frac{\nabla \rho}{\rho} \right\|_{\mathbf{L}^\infty(\Omega)} \|(\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2, \mathbf{z}_1 - \mathbf{z}_2)\|_{\mathbb{H} \times \mathbf{Q}}. \end{aligned}$$

Therefore, according to assumption (3.16), we obtain that $\mathbb{T}$ is a contraction mapping, which concludes the proof. □

Now we establish the main result of this section.

THEOREM 3.3 Assume that

$$C_{WP}\left\|\frac{\nabla\rho}{\rho}\right\|_{\mathbf{L}^\infty(\Omega)} \leqslant \frac{1}{2}, \tag{3.23}$$

with

$$C_{WP} := \frac{1}{\beta\sqrt{d}}\left(1+2\frac{\rho_1}{C_a\rho_0}\right)+\frac{\rho_1\nu}{C_a\sqrt{d}}. \tag{3.24}$$

Then, there exists a unique $(\boldsymbol{\sigma}_0,\mathbf{u})\in\mathbb{H}_0\times\mathbf{Q}$ satisfying (3.2). Moreover, the solution is stable in the sense that it satisfies inequalities (3.8).

Proof. It is clear that $(\boldsymbol{\sigma}_0,\mathbf{u})\in\mathbb{H}_0\times\mathbf{Q}$ is the unique solution of problem (3.2) if and only if it is the unique fixed point of the mapping $\mathbb{T}$. Then, noting that $C_{\mathbb{T}}\leqslant C_{WP}$, from Lemma 3.2 and the classical Banach fixed-point theorem, it follows that $\mathbb{T}$ has a unique fixed point in $\mathbb{H}_0\times\mathbf{Q}$, which implies the first assertion.

In turn, since $C_{\mathrm{dep}}\leqslant C_{WP}$, the stability of $(\boldsymbol{\sigma}_0,\mathbf{u})$ follows from Lemma 3.1. □

REMARK 3.4 Observe that problem (3.2) can be rewritten alternatively as a nonsymmetric mixed problem: find $(\boldsymbol{\sigma}_0,\mathbf{u})\in\mathbb{H}_0\times\mathbf{Q}$ such that

$$a(\boldsymbol{\sigma}_0,\boldsymbol{\tau})+b_1(\boldsymbol{\tau},\mathbf{u})=0,$$
$$b_2(\boldsymbol{\sigma}_0,\mathbf{v})=-(\mathbf{f},\mathbf{v})_\Omega,$$

for all $(\boldsymbol{\tau},\mathbf{v})\in\mathbb{H}_0\times\mathbf{Q}$, where

$$b_1(\boldsymbol{\tau},\mathbf{v}):=b(\boldsymbol{\tau},\mathbf{v})-c(\boldsymbol{\tau},\mathbf{v})\quad\text{and}\quad b_2(\boldsymbol{\tau},\mathbf{v}):=b(\boldsymbol{\tau},\mathbf{v}).$$

Then, as in Bernardi *et al.* (1992), one could try to prove the well-posedness of problem (3.2) by applying the results in Nicolaides (1982) and Bernardi *et al.* (1988). Nevertheless, in this case, the assumptions that the bilinear forms a, b_1 and b_2 must satisfy are not easily verifiable. Alternatively, one could also try to adapt the results provided in Demkowicz (2006) and prove the well-posedness of problem (3.2) by using the global Babuška inf–sup condition, but again, due to the nonsymmetric nature of the mixed problem, its verification is not trivial.

We now provide the converse of the derivation of (2.11). More precisely, the following theorem establishes that the unique solution of (2.11) solves the original problem described by (2.10). We remark that there are no extra regularity assumptions on the data; only $\mathbf{f}\in\mathbf{L}^2(\Omega)$ is required here.

THEOREM 3.5 Let $(\boldsymbol{\sigma},\mathbf{u})\in\mathbb{H}(\mathbf{div};\Omega)\times\mathbf{L}^2(\Omega)$ be the unique solution of the variational formulation (2.11), such that $(\operatorname{tr}\boldsymbol{\sigma}+\nu\mathbf{u}\cdot\nabla\rho,1)_\Omega=0$. Then $(\nu^{-1}/\rho)\boldsymbol{\sigma}^{\mathrm{D}}=\nabla\mathbf{u}+\frac{1}{2}(\mathbf{u}\cdot(\nabla\rho/\rho))I$ in Ω, $-\mathbf{div}\,\boldsymbol{\sigma}=\mathbf{f}$ in Ω, $\mathbf{u}=0$ on Γ, $(\operatorname{tr}\boldsymbol{\sigma},1)_\Omega=-\nu(\mathbf{u}\cdot\nabla\rho,1)_\Omega$, and therefore $\mathbf{u}\in\mathbf{H}_0^1(\Omega)$.

Proof. It basically follows by applying integration by parts backwards in (2.11) and using suitable test functions. We omit further details. □

REMARK 3.6 It is easy to see from Theorem 3.5 and equation (2.12) that $\boldsymbol{\sigma}_0\in\mathbb{H}_0(\mathbf{div};\Omega)$ satisfies $(\nu^{-1}/\rho)\boldsymbol{\sigma}_0^{\mathrm{D}}=\nabla\mathbf{u}+\frac{1}{2}(\mathbf{u}\cdot(\nabla\rho/\rho))I$ in Ω, and $-\mathbf{div}\,\boldsymbol{\sigma}_0=\mathbf{f}$ in Ω.

4. The mixed finite element scheme

In this section we introduce the Galerkin scheme of problem (3.2) and analyse its well-posedness by establishing suitable assumptions on the finite element subspaces involved. Then, we provide specific examples for these subspaces, satisfying the required hypotheses.

4.1 *Preliminaries*

We start by selecting the following arbitrary discrete spaces:

$$\mathbf{H}_h \subseteq \mathbf{H}(\mathbf{div};\Omega), \quad Q_h \subseteq L^2(\Omega). \tag{4.1}$$

Then, we define the subspaces

$$\begin{aligned} \mathbb{H}_h &:= \{\boldsymbol{\tau} \in \mathbb{H}(\mathbf{div};\Omega) \,:\, \mathbf{c}^t\boldsymbol{\tau} \in \mathbf{H}_h \quad \forall \mathbf{c} \in \mathbf{R}^d\},\\ \mathbb{H}_{h,0} &:= \mathbb{H}_h \cap \mathbb{H}_0(\mathbf{div};\Omega),\\ \mathbf{Q}_h &:= [Q_h]^d. \end{aligned} \tag{4.2}$$

In this way, the Galerkin scheme for (3.2) reduces to the following: find $(\boldsymbol{\sigma}_{h,0}, \mathbf{u}_h) \in \mathbb{H}_{h,0} \times \mathbf{Q}_h$ such that

$$\begin{aligned} a(\boldsymbol{\sigma}_{h,0}, \boldsymbol{\tau}_h) + b(\boldsymbol{\tau}_h, \mathbf{u}_h) - c(\boldsymbol{\tau}_h, \mathbf{u}_h) &= 0,\\ b(\boldsymbol{\sigma}_{h,0}, \mathbf{v}_h) &= -(\mathbf{f}, \mathbf{v}_h)_\Omega, \end{aligned} \tag{4.3}$$

for all $(\boldsymbol{\tau}_h, \mathbf{v}_h) \in \mathbb{H}_{h,0} \times \mathbf{Q}_h$.

Now we establish general hypotheses on the finite element subspaces (4.2), ensuring later on the well-posedness of (4.3). We start by observing that in order to have a meaningful space $\mathbb{H}_{h,0}$, we need to be able to eliminate multiples of the identity matrix from $\mathbb{H}_h$. This request is certainly satisfied if we assume the following.

(H.0) $[P_0(\Omega)]^{d\times d} \subseteq \mathbb{H}_h$.

Then, it follows that $I \in \mathbb{H}_h$ for all h, and hence there holds the decomposition

$$\mathbb{H}_h = \mathbb{H}_{h,0} \oplus P_0(\Omega)I.$$

Now we look at the discrete kernel on b, which is defined by

$$\mathbb{K}_{h,0} := \{\boldsymbol{\tau}_h \in \mathbb{H}_{h,0} \,:\, b(\boldsymbol{\tau}_h, \mathbf{v}_h) = 0 \ \ \forall \mathbf{v}_h \in \mathbf{Q}_h\}.$$

In order to have a more explicit definition of $\mathbb{K}_{h,0}$ we introduce the following assumption.

(H.1) $\operatorname{div} \mathbf{H}_h \subseteq Q_h$.

Then, it follows from the definition of b that

$$\mathbb{K}_{h,0} := \{\boldsymbol{\tau}_h \in \mathbb{H}_{h,0} \,:\, \mathbf{div}\,\boldsymbol{\tau}_h = \mathbf{0} \text{ in } \Omega\} \subseteq \mathbb{K}_0.$$

Next, we assume that the discrete version of (3.4) holds.
(**H.2**) There exists $\hat{\beta} > 0$, independent of h, such that

$$\sup_{\boldsymbol{\tau}_h \in \mathbb{H}_{h,0} \setminus \mathbf{0}} \frac{b(\boldsymbol{\tau}_h, \mathbf{v}_h)}{\|\boldsymbol{\tau}_h\|_{\mathbf{div},\Omega}} \geqslant \hat{\beta} \|\mathbf{v}_h\|_{0,\Omega} \quad \forall \mathbf{v}_h \in \mathbf{Q}_h. \tag{4.4}$$

4.2 *Well-posedness of the discrete problem*

In this section we adapt the analysis from Section 3 to the discrete case to prove the well-posedness of (4.3). First, we observe that, since we are considering conforming finite element subspaces, the continuity of the bilinear forms a, b and c (cf. (3.3)) are inherited from the continuous case, with the exact same constants. Moreover, since $\mathbb{K}_{h,0} \subseteq \mathbb{K}_0$, we deduce that the ellipticity of a on $\mathbb{K}_{h,0}$ holds:

$$a(\boldsymbol{\tau}_h, \boldsymbol{\tau}_h) \geqslant \frac{C_a}{\nu \rho_1} \|\boldsymbol{\tau}_h\|^2_{\mathbf{div},\Omega} \quad \forall \boldsymbol{\tau}_h \in \mathbb{K}_{h,0}. \tag{4.5}$$

In this way, according to (3.3), (4.4), (4.5) and the classical Babuška–Brezzi theory, and similarly to the analysis of the continuous problem, we are able to introduce the well-defined linear mapping

$$\hat{\mathbb{T}} : (\boldsymbol{\xi}_h, \mathbf{z}_h) \in \mathbb{H}_{h,0} \times \mathbf{Q}_h \to (\boldsymbol{\sigma}_{h,0}, \mathbf{u}_h) \in \mathbb{H}_{h,0} \times \mathbf{Q}_h$$

as the solution to the following problem: find $(\boldsymbol{\sigma}_{h,0}, \mathbf{u}_h) \in \mathbb{H}_{h,0} \times \mathbf{Q}_h$ such that

$$\begin{aligned} a(\boldsymbol{\sigma}_{h,0}, \boldsymbol{\tau}_h) + b(\boldsymbol{\tau}_h, \mathbf{u}_h) &= c(\boldsymbol{\tau}_h, \mathbf{z}_h), \\ b(\boldsymbol{\sigma}_{h,0}, \mathbf{v}_h) &= -(\mathbf{f}, \mathbf{v}_h)_\Omega, \end{aligned} \tag{4.6}$$

for all $(\boldsymbol{\tau}_h, \mathbf{v}_h) \in \mathbb{H}_{h,0} \times \mathbf{Q}_h$.

REMARK 4.1 It is easy to see that $(\boldsymbol{\sigma}_{h,0}, \mathbf{u}_h)$ is the solution of (4.3) if and only if $\hat{\mathbb{T}}(\boldsymbol{\sigma}_{h,0}, \mathbf{u}_h) = (\boldsymbol{\sigma}_{h,0}, \mathbf{u}_h)$. In this way, in order to prove that (4.3) is well posed, we proceed analogously to Section 3.3, and prove that $\hat{\mathbb{T}}$ has a unique fixed point in $\mathbb{H}_{h,0} \times \mathbf{Q}_h$.

THEOREM 4.2 Assume that hypotheses **(H.0)**, **(H.1)** and **(H.2)** hold. In addition, assume that

$$\hat{C}_{WP} \left\| \frac{\nabla \rho}{\rho} \right\|_{\mathbf{L}^\infty(\Omega)} \leqslant \frac{1}{2}, \tag{4.7}$$

with

$$\hat{C}_{WP} := \frac{1}{\hat{\beta}\sqrt{d}} \left(1 + 2\frac{\rho_1}{C_a \rho_0}\right) + \frac{\rho_1 \nu}{C_a \sqrt{d}}. \tag{4.8}$$

Then, there exists a unique $(\boldsymbol{\sigma}_{h,0}, \mathbf{u}_h) \in \mathbb{H}_{h,0} \times \mathbf{Q}_h$ satisfying (4.3). Moreover, there exist positive constants $\hat{C}_{\boldsymbol{\sigma}}$ and $\hat{C}_{\mathbf{u}}$, depending only on the stability constants in (3.3), (4.4) and (4.5), such that

$$\|\boldsymbol{\sigma}_{h,0}\|_{\mathbf{div},\Omega} \leqslant \hat{C}_{\boldsymbol{\sigma}} \|\mathbf{f}\|_{0,\Omega} \quad \text{and} \quad \|\mathbf{u}_h\|_{0,\Omega} \leqslant \hat{C}_{\mathbf{u}} \|\mathbf{f}\|_{0,\Omega}. \tag{4.9}$$

(Explicit expressions for $\hat{C}_{\boldsymbol{\sigma}}$ and $\hat{C}_{\mathbf{u}}$ can be found in (4.10) and (4.11).)

Proof. Let

$$\hat{C}_{\hat{\mathbb{T}}} := \frac{1}{\hat{\beta}\sqrt{d}}\left(1 + \frac{\rho_1}{C_a\rho_0}\right) + \frac{\rho_1 \nu}{C_a\sqrt{d}}.$$

It is clear that $\hat{C}_{\hat{\mathbb{T}}} \leqslant \hat{C}_{WP}$. Then, we proceed analogously to Lemma 3.2, to prove that the mapping $\hat{\mathbb{T}}$ has a unique fixed point $(\boldsymbol{\sigma}_{h,0}, \mathbf{u}_h) \in \mathbb{H}_{h,0} \times \mathbf{Q}_h$ which, according to Remark 4.1, is also the unique solution of (4.3).

Next, we let

$$\hat{C}_{\text{dep}} := \frac{1}{\hat{\beta}\sqrt{d}}\left(1 + 2\frac{\rho_1}{C_a\rho_0}\right),$$

and observe that $\hat{C}_{\text{dep}} \leqslant \hat{C}_{WP}$. Then, noting that from the second equation of (4.3) there holds

$$\|\mathbf{div}\,\boldsymbol{\sigma}_{h,0}\|_{0,\Omega} \leqslant \|\mathbf{f}\|_{0,\Omega},$$

we proceed as in the proof of Lemma 3.1 to obtain that

$$\|\boldsymbol{\sigma}_{h,0}\|_{\mathbf{div},\Omega} \leqslant 2\left(\frac{2\rho_1}{C_a\rho_0\hat{\beta}} + \frac{1+C_a}{C_a}\right)\|\mathbf{f}\|_{0,\Omega}, \tag{4.10}$$

and

$$\|\mathbf{u}_h\|_{0,\Omega} \leqslant \frac{4}{\nu\rho_0\hat{\beta}}\left(\frac{2\rho_1}{C_a\rho_0\hat{\beta}} + \frac{1+C_a}{C_a}\right)\|\mathbf{f}\|_{0,\Omega}, \tag{4.11}$$

which concludes the proof. □

4.3 *A priori error estimate*

Now we establish the corresponding Céa *a priori* error estimate. To that end, we first introduce some notation and state some previous results. We begin by defining the set

$$\mathbb{H}_h^{\mathbf{f}} := \{\boldsymbol{\tau}_h \in \mathbb{H}_{h,0} : b(\boldsymbol{\tau}_h, \mathbf{v}_h) = -(\mathbf{f}, \mathbf{v}_h)_\Omega \ \forall \mathbf{v}_h \in \mathbf{Q}_h\},$$

which is clearly nonempty, since (4.4) holds. Also, it is not difficult to see that, due to the inf–sup condition (4.4), the following inequality holds (see, for instance, Gatica, 2014, Theorem 2.6):

$$\inf_{\boldsymbol{\tau}_h \in \mathbb{H}_h^{\mathbf{f}}} \|\boldsymbol{\sigma}_0 - \boldsymbol{\tau}_h\|_{\mathbf{div},\Omega} \leqslant \left(1 + \frac{1}{\hat{\beta}}\right) \inf_{\boldsymbol{\tau}_h \in \mathbb{H}_{h,0}} \|\boldsymbol{\sigma}_0 - \boldsymbol{\tau}_h\|_{\mathbf{div},\Omega}. \tag{4.12}$$

In turn, in order to simplify the subsequent analysis, we write $\mathbf{e}_{\sigma} = \boldsymbol{\sigma}_0 - \boldsymbol{\sigma}_{h,0}$ and $\mathbf{e}_{\mathbf{u}} = \mathbf{u} - \mathbf{u}_h$. As usual, for a given $(\hat{\boldsymbol{\tau}}_h, \hat{\mathbf{v}}_h) \in \mathbb{H}_h^{\mathbf{f}} \times \mathbf{Q}_h$, we shall then decompose these errors into

$$\mathbf{e}_{\sigma} = \boldsymbol{\xi}_{\sigma} + \boldsymbol{\chi}_{\sigma} \quad \text{and} \quad \mathbf{e}_{\mathbf{u}} = \boldsymbol{\xi}_{\mathbf{u}} + \boldsymbol{\chi}_{\mathbf{u}}, \tag{4.13}$$

with

$$\begin{aligned} \boldsymbol{\xi}_{\sigma} &:= \boldsymbol{\sigma}_0 - \hat{\boldsymbol{\tau}}_h \in \mathbb{H}_0, \quad \boldsymbol{\chi}_{\sigma} := \hat{\boldsymbol{\tau}}_h - \boldsymbol{\sigma}_{h,0} \in \mathbb{H}_{h,0}, \\ \boldsymbol{\xi}_{\mathbf{u}} &:= \mathbf{u} - \hat{\mathbf{v}}_h \in \mathbf{Q}, \quad \boldsymbol{\chi}_{\mathbf{u}} := \hat{\mathbf{v}}_h - \mathbf{u}_h \in \mathbf{Q}_h. \end{aligned} \tag{4.14}$$

Finally, we observe that Galerkin orthogonality holds:

$$\begin{aligned} a(\mathbf{e}_{\boldsymbol{\sigma}},\boldsymbol{\tau}_h)+b(\boldsymbol{\tau}_h,\mathbf{e}_{\mathbf{u}})-c(\boldsymbol{\tau}_h,\mathbf{e}_{\mathbf{u}})&=0,\\ b(\mathbf{e}_{\boldsymbol{\sigma}},\mathbf{v}_h)&=0, \end{aligned} \tag{4.15}$$

for all $(\boldsymbol{\tau}_h,\mathbf{v}_h)\in\mathbb{H}_{h,0}\times\mathbf{Q}_h$.

We now establish the main result of this section.

THEOREM 4.3 Assume that hypotheses **(H.0)**, **(H.1)** and **(H.2)** hold. In addition, assume that

$$\max\{C_{WP},\hat{C}_{WP}\}\left\|\frac{\nabla\rho}{\rho}\right\|_{\mathbf{L}^\infty(\Omega)}\leqslant\frac{1}{2}, \tag{4.16}$$

with C_{WP} and $\hat{C}_{WP}$ defined in (3.24) and (4.8), respectively. Let $(\boldsymbol{\sigma}_0,\mathbf{u})\in\mathbb{H}_0\times\mathbf{Q}$ and $(\boldsymbol{\sigma}_{h,0},\mathbf{u}_h)\in\mathbb{H}_{h,0}\times\mathbf{Q}_h$ be the unique solutions of the continuous and discrete problems (3.2) and (4.3), respectively. Then, there exists $C_{\text{cea}}>0$, independent of h, such that

$$\|\boldsymbol{\sigma}_0-\boldsymbol{\sigma}_{h,0}\|_{\mathbf{div},\Omega}+\|\mathbf{u}-\mathbf{u}_h\|_{0,\Omega}\leqslant C_{\text{cea}}\left\{\inf_{\boldsymbol{\tau}_h\in\mathbb{H}_{h,0}}\|\boldsymbol{\sigma}_0-\boldsymbol{\tau}_h\|_{\mathbf{div},\Omega}+\inf_{\mathbf{v}_h\in\mathbf{Q}_h}\|\mathbf{u}-\mathbf{v}_h\|_{0,\Omega}\right\}. \tag{4.17}$$

Proof. Let $(\hat{\boldsymbol{\tau}}_h,\hat{\mathbf{v}}_h)\in\mathbb{H}_h^{\mathbf{f}}\times\mathbf{Q}_h$, and define $\boldsymbol{\xi}_{\boldsymbol{\sigma}}$, $\boldsymbol{\xi}_{\mathbf{u}}$, $\boldsymbol{\chi}_{\boldsymbol{\sigma}}$ and $\boldsymbol{\chi}_{\mathbf{u}}$, as in (4.14). It is easy to see that the first equation of (4.15) can be rewritten as

$$a(\boldsymbol{\chi}_{\boldsymbol{\sigma}},\boldsymbol{\tau}_h)+b(\boldsymbol{\tau}_h,\boldsymbol{\chi}_{\mathbf{u}})-c(\boldsymbol{\tau}_h,\boldsymbol{\chi}_{\mathbf{u}})=-a(\boldsymbol{\xi}_{\boldsymbol{\sigma}},\boldsymbol{\tau}_h)-b(\boldsymbol{\tau}_h,\boldsymbol{\xi}_{\mathbf{u}})+c(\boldsymbol{\tau}_h,\boldsymbol{\xi}_{\mathbf{u}})\quad\forall\boldsymbol{\tau}_h\in\mathbb{H}_{h,0}. \tag{4.18}$$

Then, from the inf–sup condition (4.4), (4.18) and the continuity of a, b and c in (3.3), it follows that

$$\begin{aligned} \|\boldsymbol{\chi}_{\mathbf{u}}\|_{0,\Omega}&\leqslant\frac{1}{\hat{\beta}}\sup_{\boldsymbol{\tau}_h\in\mathbb{H}_{h,0}\setminus\mathbf{0}}\frac{b(\boldsymbol{\tau}_h,\boldsymbol{\chi}_{\mathbf{u}})}{\|\boldsymbol{\tau}_h\|_{\mathbf{div},\Omega}}\\ &=\frac{1}{\hat{\beta}}\sup_{\boldsymbol{\tau}_h\in\mathbb{H}_{h,0}\setminus\mathbf{0}}\frac{-a(\boldsymbol{\chi}_{\boldsymbol{\sigma}},\boldsymbol{\tau}_h)-a(\boldsymbol{\xi}_{\boldsymbol{\sigma}},\boldsymbol{\tau}_h)-b(\boldsymbol{\tau}_h,\boldsymbol{\xi}_{\mathbf{u}})+c(\boldsymbol{\tau}_h,\boldsymbol{\chi}_{\mathbf{u}})+c(\boldsymbol{\tau}_h,\boldsymbol{\xi}_{\mathbf{u}})}{\|\boldsymbol{\tau}_h\|_{\mathbf{div},\Omega}}\\ &\leqslant\frac{1}{\nu\rho_0\hat{\beta}}(\|\boldsymbol{\xi}_{\boldsymbol{\sigma}}\|_{\mathbf{div},\Omega}+\|\boldsymbol{\chi}_{\boldsymbol{\sigma}}\|_{\mathbf{div},\Omega})+\frac{1}{\hat{\beta}}\left(1+\frac{1}{\sqrt{d}}\left\|\frac{\nabla\rho}{\rho}\right\|_{\mathbf{L}^\infty(\Omega)}\right)\|\boldsymbol{\xi}_{\mathbf{u}}\|_{0,\Omega}\\ &\quad+\frac{1}{\hat{\beta}\sqrt{d}}\left\|\frac{\nabla\rho}{\rho}\right\|_{\mathbf{L}^\infty(\Omega)}\|\boldsymbol{\chi}_{\mathbf{u}}\|_{0,\Omega}, \end{aligned}$$

which together with assumption (4.16) implies

$$\|\boldsymbol{\chi}_{\mathbf{u}}\|_{0,\Omega}\leqslant\frac{2}{\nu\rho_0\hat{\beta}}(\|\boldsymbol{\xi}_{\boldsymbol{\sigma}}\|_{\mathbf{div},\Omega}+\|\boldsymbol{\chi}_{\boldsymbol{\sigma}}\|_{\mathbf{div},\Omega})+\frac{2}{\hat{\beta}}\left(1+\frac{1}{\sqrt{d}}\left\|\frac{\nabla\rho}{\rho}\right\|_{\mathbf{L}^\infty(\Omega)}\right)\|\boldsymbol{\xi}_{\mathbf{u}}\|_{0,\Omega}. \tag{4.19}$$

In turn, since $\hat{\boldsymbol{\tau}}_h\in\mathbb{H}_h^{\mathbf{f}}$, we observe that $\boldsymbol{\chi}_{\boldsymbol{\sigma}}\in\mathbb{K}_{h,0}$, and then, from (4.18) with $\boldsymbol{\tau}_h=\boldsymbol{\chi}_{\boldsymbol{\sigma}}$, we obtain

$$a(\boldsymbol{\chi}_{\boldsymbol{\sigma}},\boldsymbol{\chi}_{\boldsymbol{\sigma}})=-a(\boldsymbol{\xi}_{\boldsymbol{\sigma}},\boldsymbol{\chi}_{\boldsymbol{\sigma}})+c(\boldsymbol{\chi}_{\boldsymbol{\sigma}},\boldsymbol{\xi}_{\mathbf{u}})+c(\boldsymbol{\chi}_{\boldsymbol{\sigma}},\boldsymbol{\chi}_{\mathbf{u}}),$$

and using the continuity of a and c in (3.3), and the ellipticity of a in (4.5), we obtain

$$\|\boldsymbol{\chi}_{\boldsymbol{\sigma}}\|_{\mathbf{div},\Omega} \leqslant \frac{\rho_1}{C_a \rho_0}\|\boldsymbol{\xi}_{\boldsymbol{\sigma}}\|_{\mathbf{div},\Omega} + \frac{\nu\rho_1}{C_a\sqrt{d}}\left\|\frac{\nabla\rho}{\rho}\right\|_{\mathbf{L}^\infty(\Omega)}\|\boldsymbol{\xi}_{\mathbf{u}}\|_{0,\Omega} + \frac{\nu\rho_1}{C_a\sqrt{d}}\left\|\frac{\nabla\rho}{\rho}\right\|_{\mathbf{L}^\infty(\Omega)}\|\boldsymbol{\chi}_{\mathbf{u}}\|_{0,\Omega}. \tag{4.20}$$

In this way, combining (4.19) and (4.20) it follows that

$$\|\boldsymbol{\chi}_{\boldsymbol{\sigma}}\|_{\mathbf{div},\Omega} \leqslant \frac{k_1}{2}\|\boldsymbol{\xi}_{\boldsymbol{\sigma}}\|_{\mathbf{div},\Omega} + \frac{k_2}{2}\|\boldsymbol{\xi}_{\mathbf{u}}\|_{0,\Omega} + \frac{2\rho_1}{\rho_0 C_a\hat{\beta}\sqrt{d}}\left\|\frac{\nabla\rho}{\rho}\right\|_{\mathbf{L}^\infty(\Omega)}\|\boldsymbol{\chi}_{\boldsymbol{\sigma}}\|_{\mathbf{div},\Omega},$$

which together with assumption (4.16) yields

$$\|\boldsymbol{\chi}_{\boldsymbol{\sigma}}\|_{\mathbf{div},\Omega} \leqslant k_1\|\boldsymbol{\xi}_{\boldsymbol{\sigma}}\|_{\mathbf{div},\Omega} + k_2\|\boldsymbol{\xi}_{\mathbf{u}}\|_{0,\Omega}, \tag{4.21}$$

with

$$k_1 := \frac{\rho_1}{C_a\rho_0}\left(1 + \frac{2}{\hat{\beta}\sqrt{d}}\left\|\frac{\nabla\rho}{\rho}\right\|_{\mathbf{L}^\infty(\Omega)}\right),$$

$$k_2 := \frac{\nu\rho_1}{C_a\sqrt{d}}\left(1 + \frac{2}{\hat{\beta}} + \frac{2}{\hat{\beta}\sqrt{d}}\left\|\frac{\nabla\rho}{\rho}\right\|_{\mathbf{L}^\infty(\Omega)}\right)\left\|\frac{\nabla\rho}{\rho}\right\|_{\mathbf{L}^\infty(\Omega)}.$$

As a consequence, combining (4.19) and (4.21), we obtain

$$\|\boldsymbol{\chi}_{\mathbf{u}}\|_{0,\Omega} \leqslant k_3\|\boldsymbol{\xi}_{\boldsymbol{\sigma}}\|_{\mathbf{div},\Omega} + k_4\|\boldsymbol{\xi}_{\mathbf{u}}\|_{0,\Omega}, \tag{4.22}$$

with

$$k_3 := \frac{2}{\nu\rho_0\hat{\beta}}(1 + k_1),$$

$$k_4 := \frac{2}{\hat{\beta}}\left(1 + \frac{1}{\sqrt{d}}\left\|\frac{\nabla\rho}{\rho}\right\|_{\mathbf{L}^\infty(\Omega)} + \frac{k_2}{\nu\rho_0}\right).$$

Therefore, according to the triangle inequality, from (4.21) and (4.22), we obtain

$$\|\mathbf{e}_\sigma\|_{\mathbf{div},\Omega} + \|\mathbf{e}_{\mathbf{u}}\|_{0,\Omega} \leqslant (1 + k_1 + k_3)\|\boldsymbol{\xi}_{\boldsymbol{\sigma}}\|_{\mathbf{div},\Omega} + (1 + k_2 + k_4)\|\boldsymbol{\xi}_{\mathbf{u}}\|_{0,\Omega},$$

and since $(\hat{\boldsymbol{\tau}}_h, \hat{\mathbf{v}}_h) \in \mathbb{H}_h^{\mathbf{f}} \times \mathbf{Q}_h$ is arbitrary, we obtain

$$\|\mathbf{e}_\sigma\|_{\mathbf{div},\Omega} + \|\mathbf{e}_{\mathbf{u}}\|_{0,\Omega} \leqslant (1 + k_1 + k_3)\inf_{\boldsymbol{\tau}_h\in\mathbb{H}_h^{\mathbf{f}}}\|\boldsymbol{\sigma}_0 - \boldsymbol{\tau}_h\|_{\mathbf{div},\Omega} + (1 + k_2 + k_4)\inf_{\mathbf{v}_h\in\mathbf{Q}_h}\|\mathbf{u} - \mathbf{v}_h\|_{0,\Omega},$$

which together with (4.12), concludes the proof. □

Remark 4.4 An alternative proof for the Céa's estimate (4.17) can be obtained by adapting the proof of Gatica *et al.* (2014b, Theorem 4.2) to our case, where the main tools are the superposition principle and a Strang-type error estimate.

4.4 *Approximating the pressure and the original pseudostress*

First, we propose a post-processing procedure to approximate the pressure. To do that, we observe that if $(\boldsymbol{\sigma}_0, \mathbf{u}) \in \mathbb{H}_0 \times \mathbf{Q}$ is the unique solution of (3.2), then, according to (2.9) and (2.12), it is possible to recover the pressure $p \in L^2_0(\Omega) := \{q \in L^2(\Omega) : (q, 1)_\Omega = 0\}$ from the identity

$$p = -\frac{\nu}{d}\left(\mathbf{u}\cdot\nabla\rho - \frac{1}{|\Omega|}(\mathbf{u}, \nabla\rho)_\Omega\right) - \frac{1}{d}\operatorname{tr}\boldsymbol{\sigma}_0. \tag{4.23}$$

In this way, if $(\boldsymbol{\sigma}_{h,0}, \mathbf{u}_h) \in \mathbb{H}_{h,0} \times \mathbf{Q}_h$ is the unique solution of (4.3), it is reasonable to think that the function

$$p_h =: -\frac{\nu}{d}\left(\mathbf{u}_h\cdot\nabla\rho - \frac{1}{|\Omega|}(\mathbf{u}_h, \nabla\rho)_\Omega\right) - \frac{1}{d}\operatorname{tr}\boldsymbol{\sigma}_{h,0} \tag{4.24}$$

is a good approximation for the pressure. This result is established in the following corollary.

COROLLARY 4.5 Assume that the hypotheses of Theorem 4.3 hold. Let $(\boldsymbol{\sigma}_0, \mathbf{u}) \in \mathbb{H}_0 \times \mathbf{Q}$ and $(\boldsymbol{\sigma}_{h,0}, \mathbf{u}_h) \in \mathbb{H}_{h,0} \times \mathbf{Q}_h$ be the unique solutions of the continuous and discrete problems (3.2) and (4.3), respectively. Then, there exists $C > 0$, independent of h, such that

$$\|p - p_h\|_{0,\Omega} \leqslant C\left\{\inf_{\boldsymbol{\tau}_h\in\mathbb{H}_{h,0}} \|\boldsymbol{\sigma}_0 - \boldsymbol{\tau}_h\|_{\mathbf{div},\Omega} + \inf_{\mathbf{v}_h\in\mathbf{Q}_h}\|\mathbf{u} - \mathbf{v}_h\|_{0,\Omega}\right\}.$$

Proof. From (4.23) and (4.24), Hölder's and the triangle inequalities, it follows that

$$\begin{aligned}\|p - p_h\|_{0,\Omega} &\leqslant \frac{\nu}{d}\|(\mathbf{u} - \mathbf{u}_h)\cdot\nabla\rho\|_{0,\Omega} + \frac{\nu}{d|\Omega|^{1/2}}|(\mathbf{u} - \mathbf{u}_h, \nabla\rho)_\Omega| + \frac{1}{d}\|\operatorname{tr}(\boldsymbol{\sigma}_0 - \boldsymbol{\sigma}_{h,0})\|_{0,\Omega}\\ &\leqslant \frac{\nu}{d}\|\rho\|_{W^{1,\infty}(\Omega)}\|\mathbf{u} - \mathbf{u}_h\|_{0,\Omega} + \frac{\nu}{d|\Omega|^{1/2}}\|\rho\|_{1,\Omega}\|\mathbf{u} - \mathbf{u}_h\|_{0,\Omega} + \frac{1}{d}\|\operatorname{tr}(\boldsymbol{\sigma}_0 - \boldsymbol{\sigma}_{h,0})\|_{0,\Omega}.\end{aligned}$$

Then, the result follows from Theorem 4.3. □

Now, in order to approximate the original pseudostress in (2.12), let us recall that in Section 2, Lemma 2.1, we proved that formulations (2.11) and (2.14) are equivalent. That is, we proved that $(\boldsymbol{\sigma}, \mathbf{u}) \in \mathbb{H} \times \mathbf{Q}$ is the unique solution of (2.11) if and only if $(\boldsymbol{\sigma}_0, \mathbf{u}) \in \mathbb{H}_0 \times \mathbf{Q}$ is the unique solution of (2.14), where $\boldsymbol{\sigma}_0$ and $\boldsymbol{\sigma}$ are related by

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}_0 - \frac{\nu}{d|\Omega|}(\mathbf{u}, \nabla\rho)_\Omega I. \tag{4.25}$$

In turn, in this section we proposed a mixed finite element method to approximate the solution of (2.14) (or equivalently (3.2)).

As a result, if $(\boldsymbol{\sigma}_{h,0}, \mathbf{u}_h) \in \mathbb{H}_{h,0} \times \mathbf{Q}_h$ is the unique solution of (4.3), it is easy to see that the tensor

$$\boldsymbol{\sigma}_h := \boldsymbol{\sigma}_{h,0} - \frac{\nu}{d|\Omega|}(\mathbf{u}_h, \nabla\rho)_\Omega I \tag{4.26}$$

approximates $\boldsymbol{\sigma} \in \mathbb{H}$ in (4.25). This result is established in the following corollary.

COROLLARY 4.6 Assume that the hypotheses of Theorem 4.3 hold. Let $(\boldsymbol{\sigma}_0, \mathbf{u}) \in \mathbb{H}_0 \times \mathbf{Q}$ and $(\boldsymbol{\sigma}_{h,0}, \mathbf{u}_h) \in \mathbb{H}_{h,0} \times \mathbf{Q}_h$ be the unique solutions of the continuous and discrete problems (3.2) and (4.3), respectively. Then, there exists $C > 0$, independent of h, such that

$$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\mathbf{div},\Omega} \leqslant C \left\{ \inf_{\boldsymbol{\tau}_h \in \mathbb{H}_{h,0}} \|\boldsymbol{\sigma}_0 - \boldsymbol{\tau}_h\|_{\mathbf{div},\Omega} + \inf_{\mathbf{v}_h \in \mathbf{Q}_h} \|\mathbf{u} - \mathbf{v}_h\|_{0,\Omega} \right\}. \tag{4.27}$$

Proof. First, from (4.25), (4.26) and the triangle inequality, it is easy to see that

$$\begin{aligned}
\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\mathbf{div},\Omega} &= \left\| \boldsymbol{\sigma}_0 - \boldsymbol{\sigma}_{h,0} - \frac{\nu}{d|\Omega|}(\mathbf{u} - \mathbf{u}_h, \nabla\rho)_\Omega I \right\|_{\mathbf{div},\Omega} \\
&\leqslant \|\boldsymbol{\sigma}_0 - \boldsymbol{\sigma}_{h,0}\|_{\mathbf{div},\Omega} + \frac{\nu}{d^{1/2}|\Omega|^{1/2}} |(\mathbf{u} - \mathbf{u}_h, \nabla\rho)_\Omega| \\
&\leqslant \|\boldsymbol{\sigma}_0 - \boldsymbol{\sigma}_{h,0}\|_{\mathbf{div},\Omega} + \frac{\nu\|\nabla\rho\|_{0,\Omega}}{d^{1/2}|\Omega|^{1/2}} \|\mathbf{u} - \mathbf{u}_h\|_{0,\Omega}.
\end{aligned}$$

Then, the result is a direct application of Theorem 4.3. □

5. Particular choices of discrete spaces

We now specify examples of finite element subspaces satisfying the hypotheses (**H.0**), (**H.1**) and (**H.2**). To this end, we let $\mathcal{T}_h$ be a regular family of triangulations of the polygonal region $\bar{\Omega}$ by triangles T of diameter h_T such that $\bar{\Omega} = \bigcup\{T : T \in \mathcal{T}_h\}$ and define $h := \max\{h_T : T \in \mathcal{T}_h\}$. Now, given an integer $l \geqslant 0$ and a subset S of $\mathbf{R}^d$, we denote by $P_l(S)$ the space of polynomials of total degree at most l defined on S.

5.1 *The Raviart–Thomas element*

For each integer $k \geqslant 0$ and for each $T \in \mathcal{T}_h$, we define the local Raviart–Thomas space of order k (see, for instance, Boffi *et al.*, 2013, Section 2.3.1):

$$\mathbf{RT}_k(T) := [P_k(T)]^d \oplus P_k(T)\mathbf{x},$$

where $\mathbf{x} := (x_1, \ldots, x_d)^t$ is a generic vector of $\mathbf{R}^d$. Then, we specify the discrete spaces in (4.2) by defining

$$\begin{aligned}
\mathbf{H}_h &:= \{\tau \in \mathbf{H}(\mathrm{div}\,; \Omega) : \ \tau|_T \in \mathbf{RT}_k(T) \ \forall T \in \mathcal{T}_h\}, \\
Q_h &:= \{v \in L^2(\Omega) : \ v|_T \in P_k(T) \ \forall T \in \mathcal{T}_h\}.
\end{aligned} \tag{5.1}$$

It is well known that these subspaces satisfy the following approximation properties (see, e.g., Hiptmair, 2002, Theorem 3.16).

For each $s \in (0, k+1]$ and for each $\tau \in \mathbf{H}^s(\Omega)$, with $\mathrm{div}\,\tau \in H^s(\Omega)$, there exists $\tau_h \in \mathbf{H}_h$, such that

$$\|\tau - \tau_h\|_{\mathrm{div}\,,\Omega} \leqslant Ch^s\{\|\tau\|_{s,\Omega} + \|\mathrm{div}\,\tau\|_{s,\Omega}\}. \tag{5.2}$$

For each $s \in [0, k+1]$ and for each $v \in H^s(\Omega)$ there exists $v_h \in Q_h$ such that

$$\|v - v_h\|_{0,\Omega} \leqslant Ch^s\|v\|_{s,\Omega}. \tag{5.3}$$

Moreover, it is easy to see that the corresponding discrete spaces $\mathbb{H}_h$ and $\mathbf{Q}_h$ satisfy assumptions (**H.0**), (**H.1**) and (**H.2**). In particular, the proof of the inf–sup condition (4.4) can be found in Gatica *et al.* (2012, Lemma 2.4).

According to the above, and Theorem 4.3, we are able to establish the convergence of the Galerkin scheme (4.3) for this particular choice of spaces.

THEOREM 5.1 Assume that

$$\max\{C_{WP}, \hat{C}_{WP}\} \left\| \frac{\nabla \rho}{\rho} \right\|_{\mathbf{L}^\infty(\Omega)} \leqslant \frac{1}{2}, \tag{5.4}$$

with C_{WP} and $\hat{C}_{WP}$ defined in (3.24) and (4.8), respectively. In addition, let $\mathbb{H}_h$ and $\mathbf{Q}_h$ be the finite element subspaces defined by (4.2) in terms of the specific discrete spaces given by (5.1). Then, the Galerkin scheme (4.3) has a unique solution $(\boldsymbol{\sigma}_{h,0}, \mathbf{u}_h) \in \mathbb{H}_{h,0} \times \mathbf{Q}_h$ and there exists $C_1 > 0$, independent of h, such that

$$\|(\boldsymbol{\sigma}_{h,0}, \mathbf{u}_h)\|_{\mathbb{H}\times\mathbf{Q}} \leqslant C_1 \|\mathbf{f}\|_{0,\Omega}.$$

Moreover, let $(\boldsymbol{\sigma}_0, \mathbf{u}) \in \mathbb{H}_0 \times \mathbf{Q}$ be the unique solution of the continuous problem (3.2) and assume that $\boldsymbol{\sigma}_0 \in \mathbb{H}^s(\Omega)$, $\mathbf{div}\,\boldsymbol{\sigma}_0 \in \mathbf{H}^s(\Omega)$ and $\mathbf{u} \in \mathbf{H}^s(\Omega)$ for some $s \in (0, k+1]$. Then, there exists $C_2 > 0$, independent of h, such that

$$\|\boldsymbol{\sigma}_0 - \boldsymbol{\sigma}_{h,0}\|_{\mathbf{div},\Omega} + \|\mathbf{u} - \mathbf{u}_h\|_{0,\Omega} \leqslant C_2 h^s \{\|\boldsymbol{\sigma}_0\|_{s,\Omega} + \|\mathbf{div}\,\boldsymbol{\sigma}_0\|_{s,\Omega} + \|\mathbf{u}\|_{s,\Omega}\}.$$

Proof. Since the finite element subspaces $\mathbb{H}_h$ and $\mathbf{Q}_h$ satisfy hypotheses **(H.0)**, **(H.1)** and **(H.2)**, then the proof is a straightforward application of Theorems 4.2 and 4.3, and properties (5.2) and (5.3). □

Finally, from Corollary 4.5 and Theorem 5.1 we obtain the optimal convergence of the post-processed pressure introduced in (4.24).

COROLLARY 5.2 Let $(\boldsymbol{\sigma}_0, \mathbf{u}) \in \mathbb{H}_0 \times \mathbf{Q}$ be the unique solution of the continuous problem (3.2), and $p \in L^2_0(\Omega)$ given by (4.23). In addition, let p_h be the discrete pressure computed by the post-processing formula (4.24). Assume that hypotheses of Theorem 5.1 hold. Then, there exists $C > 0$, independent of h, such that

$$\|p - p_h\|_{0,\Omega} \leqslant C h^s \{\|\boldsymbol{\sigma}_0\|_{s,\Omega} + \|\mathbf{div}\,\boldsymbol{\sigma}_0\|_{s,\Omega} + \|\mathbf{u}\|_{s,\Omega}\}.$$

5.2 *The Brezzi–Douglas–Marini element*

Now, for each integer $k \geqslant 0$, we introduce the following discrete spaces in (4.2):

$$\begin{aligned} \mathbf{H}_h &:= \{\tau \in \mathbf{H}(\mathrm{div}\,; \Omega) : \ \tau|_T \in [P_{k+1}(T)]^d \ \ \forall T \in \mathcal{T}_h\}, \\ Q_h &:= \{v \in L^2(\Omega) : \ v|_T \in P_k(T) \ \ \forall T \in \mathcal{T}_h\}. \end{aligned} \tag{5.5}$$

We remark that the product space $\mathbf{H}_h \times Q_h$ constitutes the finite element approximation for the mixed problem introduced by Brezzi, Douglas and Marini (BDM) (see, e.g., Boffi *et al.*, 2013, Section 2.3.1).

Again, it is well known that these subspaces satisfy the following approximation properties (see, e.g., Hiptmair, 2002, Theorem 3.16).

For each $s \in (0, k+1]$ and for each $\tau \in \mathbf{H}^s(\Omega)$, with $\operatorname{div} \tau \in H^s(\Omega)$, there exists $\tau_h \in \mathbf{H}_h$, such that

$$\|\tau - \tau_h\|_{\operatorname{div},\Omega} \leqslant C h^s \{\|\tau\|_{s,\Omega} + \|\operatorname{div} \tau\|_{s,\Omega}\}. \tag{5.6}$$

For each $s \in [0, k+1]$ and for each $v \in H^s(\Omega)$ there exists $v_h \in Q_h$ such that

$$\|v - v_h\|_{0,\Omega} \leqslant C h^s \|v\|_{s,\Omega}. \tag{5.7}$$

Moreover, the corresponding discrete spaces $\mathbb{H}_h$ and $\mathbf{Q}_h$ satisfy assumptions (**H.0**), (**H.1**) and (**H.2**). For the proof of the inf–sup condition (4.4) in (**H.2**), we just comment that it follows analogously to the Raviart–Thomas case (see again Gatica *et al.*, 2012, Lemma 2.4), recalling that it is also possible to construct a Fortin operator by using the BDM-projection.

Now we establish the convergence of the Galerkin scheme (4.3) for this particular choice of spaces.

Theorem 5.3 Assume that

$$\max\{C_{WP}, \hat{C}_{WP}\} \left\| \frac{\nabla \rho}{\rho} \right\|_{\mathbf{L}^\infty(\Omega)} \leqslant \frac{1}{2}, \tag{5.8}$$

with C_{WP} and $\hat{C}_{WP}$ defined in (3.24) and (4.8), respectively. In addition, let $\mathbb{H}_{h,0}$ and $\mathbf{Q}_h$ be the finite element subspaces defined by (4.2) in terms of the specific discrete spaces given by (5.5). Then, the Galerkin scheme (4.3) has a unique solution $(\boldsymbol{\sigma}_{h,0}, \mathbf{u}_h) \in \mathbb{H}_{h,0} \times \mathbf{Q}_h$ and there exists $C_1 > 0$, independent of h, such that

$$\|(\boldsymbol{\sigma}_{h,0}, \mathbf{u}_h)\|_{\mathbb{H}\times\mathbf{Q}} \leqslant C_1 \|\mathbf{f}\|_{0,\Omega}.$$

Moreover, let $(\boldsymbol{\sigma}_0, \mathbf{u}) \in \mathbb{H}_0 \times \mathbf{Q}$ be the unique solution of the continuous problem (3.2) and assume that $\boldsymbol{\sigma}_0 \in \mathbb{H}^s(\Omega)$, $\mathbf{div}\, \boldsymbol{\sigma}_0 \in \mathbf{H}^s(\Omega)$ and $\mathbf{u} \in \mathbf{H}^s(\Omega)$ for some $s \in (0, k+1]$. Then, there exists $C_2 > 0$, independent of h, such that

$$\|\boldsymbol{\sigma}_0 - \boldsymbol{\sigma}_{h,0}\|_{\mathbf{div},\Omega} + \|\mathbf{u} - \mathbf{u}_h\|_{0,\Omega} \leqslant C_2 h^s \{\|\boldsymbol{\sigma}_0\|_{s,\Omega} + \|\mathbf{div}\, \boldsymbol{\sigma}_0\|_{s,\Omega} + \|\mathbf{u}\|_{s,\Omega}\}.$$

Proof. Since the finite element subspaces $\mathbb{H}_{h,0}$ and $\mathbf{Q}_h$ satisfy hypotheses **(H.0)**, **(H.1)** and **(H.2)**, then the proof is a straightforward application of Theorem 4.2 and 4.3, and properties (5.6) and (5.7). □

We end this section by establishing the rate of convergence of the post-processed pressure computed by formula (4.24). Its proof follows from Corollary 4.5 and Theorem 5.3.

Corollary 5.4 Let $(\boldsymbol{\sigma}_0, \mathbf{u}) \in \mathbb{H}_0 \times \mathbf{Q}$ be the unique solution of the continuous problem (3.2), and $p \in L_0^2(\Omega)$ given by (4.23). In addition, let p_h be the discrete pressure computed by the post-processing formula (4.24). Assume that the hypotheses of Theorem 5.3 hold. Then, there exists $C > 0$, independent of h, such that

$$\|p - p_h\|_{0,\Omega} \leqslant C h^s \{\|\boldsymbol{\sigma}_0\|_{s,\Omega} + \|\mathbf{div}\, \boldsymbol{\sigma}_0\|_{s,\Omega} + \|\mathbf{u}\|_{s,\Omega}\}.$$

6. A residual-based *a posteriori* error estimator

In this section we restrict ourselves to the two-dimensional case, and derive a reliable and efficient residual-based *a posteriori* error estimate for our mixed method (4.3), with the discrete spaces introduced in Section 5.1. The extension to three dimensions should be quite straightforward.

We first introduce some notation. For each $T \in \mathcal{T}_h$ we let $\mathcal{E}(T)$ be the set of edges of T, and we denote by $\mathcal{E}_h$ the set of all edges of $\mathcal{T}_h$, subdivided as follows:

$$\mathcal{E}_h = \mathcal{E}_h(\Gamma) \cup \mathcal{E}_h(\Omega),$$

where $\mathcal{E}_h(\Gamma) := \{e \in \mathcal{E}_h : e \subseteq \Gamma\}$ and $\mathcal{E}_h(\Omega) := \{e \in \mathcal{E}_h : e \subseteq \Omega\}$. In what follows, h_e stands for the diameter of a given edge e. Also, for each edge $e \in \mathcal{E}_h$ we fix a unit normal vector $\mathbf{n}_e := (n_1, n_2)^{\mathrm{T}}$ to the edge e (its particular orientation is not relevant) and let $\mathbf{t}_e := (-n_2, n_1)^{\mathrm{T}}$ be the corresponding fixed unit tangential vector along e. Hence, given $\mathbf{v} \in \mathbf{L}^2(\Omega)$ and $\boldsymbol{\tau} \in \mathbb{L}^2(\Omega)$ such that $\mathbf{v}|_T \in [C(T)]^2$ and $\boldsymbol{\tau}|_T \in [C(T)]^{2\times 2}$, for each $T \in \mathcal{T}_h$, we let $[\mathbf{v} \cdot \mathbf{t}_e]$ and $[\boldsymbol{\tau}\mathbf{t}_e]$ be the tangential jumps across e of $\mathbf{v}$ and $\boldsymbol{\tau}$, respectively, that is, $[\mathbf{v} \cdot \mathbf{t}_e] := \{(\mathbf{v}|_{T'})|_e - (\mathbf{v}|_{T''})|_e\} \cdot \mathbf{t}_e$ and $[\boldsymbol{\tau}\mathbf{t}_e] := \{(\boldsymbol{\tau}|_{T'})|_e - (\boldsymbol{\tau}|_{T''})|_e\}\mathbf{t}_e$, where T' and T'' are the triangles of $\mathcal{T}_h$ having e as an edge. From now on, when no confusion arises, we will simply write $\mathbf{t}$ and $\mathbf{n}$ instead of $\mathbf{t}_e$ and $\mathbf{n}_e$, respectively. Finally, for sufficiently smooth scalar, vector and tensor fields q, $\mathbf{v} := (v_1, v_2)^{\mathrm{T}}$ and $\boldsymbol{\tau} := (\tau_{ij})_{2\times 2}$, respectively, we let

$$\mathbf{curl}\,\mathbf{v} := \begin{pmatrix} \dfrac{\partial v_1}{\partial x_2} & -\dfrac{\partial v_1}{\partial x_1} \\ \dfrac{\partial v_2}{\partial x_2} & -\dfrac{\partial v_2}{\partial x_1} \end{pmatrix}, \quad \mathrm{curl}\, q := \left(\frac{\partial q}{\partial x_2}, -\frac{\partial q}{\partial x_1} \right)^{\mathrm{T}},$$

$$\mathrm{rot}\,\mathbf{v} := \frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2} \quad \text{and} \quad \mathbf{rot}\,\boldsymbol{\tau} := \left(\frac{\partial \boldsymbol{\tau}_{12}}{\partial x_1} - \frac{\partial \boldsymbol{\tau}_{11}}{\partial x_2}, \frac{\partial \boldsymbol{\tau}_{22}}{\partial x_1} - \frac{\partial \boldsymbol{\tau}_{21}}{\partial x_2} \right)^{\mathrm{T}}.$$

Now, let $(\boldsymbol{\sigma}_0, \mathbf{u}) \in \mathbb{H}_0 \times \mathbf{Q}$ and $(\boldsymbol{\sigma}_{h,0}, \mathbf{u}_h) \in \mathbb{H}_{h,0} \times \mathbf{Q}_h$ be the unique solutions of the continuous and discrete formulations (3.2) and (4.3), respectively. Then, we introduce the global *a posteriori* error estimator

$$\Theta := \left\{ \sum_{T \in \mathcal{T}_h} \Theta_T^2 \right\}^{1/2}, \tag{6.1}$$

where, for each $T \in \mathcal{T}_h$,

$$\begin{aligned} \Theta_T^2 := \|\mathbf{f} + \mathbf{div}\,\boldsymbol{\sigma}_{h,0}\|_{0,T}^2 &+ h_T^2 \left\| \mathbf{rot} \left(\frac{\nu^{-1}}{\rho} \boldsymbol{\sigma}_{h,0}^{\mathrm{D}} - \frac{1}{2} \left(\mathbf{u}_h \cdot \frac{\nabla \rho}{\rho} \right) I \right) \right\|_{0,T}^2 \\ &+ h_T^2 \left\| \nabla \mathbf{u}_h - \left(\frac{\nu^{-1}}{\rho} \boldsymbol{\sigma}_{h,0}^{\mathrm{D}} - \frac{1}{2} \left(\mathbf{u}_h \cdot \frac{\nabla \rho}{\rho} \right) I \right) \right\|_{0,T}^2 \\ &+ \sum_{e \in \mathcal{E}(T)} h_e \left\| \left[\left(\frac{\nu^{-1}}{\rho} \boldsymbol{\sigma}_{h,0}^{\mathrm{D}} - \frac{1}{2} \left(\mathbf{u}_h \cdot \frac{\nabla \rho}{\rho} \right) I \right) \mathbf{t} \right] \right\|_{0,e}^2. \end{aligned}$$

6.1 *Reliability of the a posteriori error estimator*

The main result of this section is stated in the following theorem.

THEOREM 6.1 Assume that

$$C_{\mathrm{ap}} C_{\mathrm{glob}} \left\| \frac{\nabla \rho}{\rho} \right\|_{0,\Omega} \leqslant \frac{1}{2}. \tag{6.2}$$

Then, there exists $C_{\rm rel} > 0$, independent of h, such that

$$\|\boldsymbol{\sigma}_0 - \boldsymbol{\sigma}_{h,0}\|_{\mathbf{div},\Omega} + \|\mathbf{u} - \mathbf{u}_h\|_{0,\Omega} \leqslant C_{\rm rel}\Theta \tag{6.3}$$

(and an explicit expression for $C_{\rm ap}$ is given in Lemma 6.5).

We begin the derivation of (6.3) by recalling that the continuous dependence result given by Lemma 3.1 is equivalent to the global inf–sup condition for the continuous formulation (3.2). Then, applying this estimate to the error $(\boldsymbol{\sigma}_0 - \boldsymbol{\sigma}_{h,0}, \mathbf{u} - \mathbf{u}_h) \in \mathbb{H}_0 \times \mathbf{Q}$, we obtain

$$||(\boldsymbol{\sigma}_0 - \boldsymbol{\sigma}_{h,0}, \mathbf{u} - \mathbf{u}_h)||_{\mathbb{H}\times\mathbf{Q}} \leqslant C_{\rm glob} \sup_{(\boldsymbol{\tau},\mathbf{v})\in\mathbb{H}_0\times\mathbf{Q}} \frac{|\mathcal{R}(\boldsymbol{\tau},\mathbf{v})|}{||(\boldsymbol{\tau},\mathbf{v})||_{\mathbb{H}\times\mathbf{Q}}}, \tag{6.4}$$

where $\mathcal{R}: \mathbb{H}_0 \times \mathbf{Q} \to \mathbb{R}$ is the residual functional

$$\mathcal{R}(\boldsymbol{\tau},\mathbf{v}) := a(\boldsymbol{\sigma}_0 - \boldsymbol{\sigma}_{h,0}, \boldsymbol{\tau}) + b(\boldsymbol{\tau}, \mathbf{u} - \mathbf{u}_h) - c(\boldsymbol{\tau}, \mathbf{u} - \mathbf{u}_h) + b(\boldsymbol{\sigma}_0 - \boldsymbol{\sigma}_{h,0}, \mathbf{v}) \quad \forall (\boldsymbol{\tau},\mathbf{v}) \in \mathbb{H}_0 \times \mathbf{Q}.$$

More precisely, according to (3.2), (4.3) and the definition of the bilinear forms a, b and c, we find that, for any $(\boldsymbol{\tau},\mathbf{v}) \in \mathbb{H}_0 \times \mathbf{Q}$, there holds

$$\mathcal{R}(\boldsymbol{\tau},\mathbf{v}) := \mathcal{R}_1(\boldsymbol{\tau}) + \mathcal{R}_2(\mathbf{v}),$$

where

$$\mathcal{R}_1(\boldsymbol{\tau}) = -(\mathbf{u}_h, \mathbf{div}\,\boldsymbol{\tau})_\Omega - \left(\frac{\nu^{-1}}{\rho}\boldsymbol{\sigma}_{h,0}^{\rm D} - \frac{1}{2}\left(\mathbf{u}_h \cdot \frac{\nabla\rho}{\rho}\right) I, \boldsymbol{\tau}\right)_\Omega \quad \text{and} \quad \mathcal{R}_2(\mathbf{v}) = -(\mathbf{f} + \mathbf{div}\,\boldsymbol{\sigma}_{h,0}, \mathbf{v})_\Omega.$$

Hence, the supremum in (6.4) can be bounded in terms of $\mathcal{R}_1$ and $\mathcal{R}_2$ as follows:

$$\|(\boldsymbol{\sigma}_0 - \boldsymbol{\sigma}_{h,0}, \mathbf{u} - \mathbf{u}_h)\|_{\mathbb{H}\times\mathbf{Q}} \leqslant C_{\rm glob}\{\|\mathcal{R}_1\|_{\mathbb{H}_0'} + \|\mathcal{R}_2\|_{\mathbf{Q}'}\}. \tag{6.5}$$

In this way, we have transformed (6.4) into an estimate involving global inf–sup conditions on $\mathbb{H}_0$ and $\mathbf{Q}$, separately.

Throughout the rest of this section, we provide suitable upper bounds for $\mathcal{R}_1$ and $\mathcal{R}_2$. We begin by establishing the corresponding estimate for $\mathcal{R}_2$, whose proof follows from a straightforward application of the Cauchy–Schwarz inequality.

LEMMA 6.2 There holds

$$\|\mathcal{R}_2\|_{\mathbf{Q}'} = ||\mathbf{f} + \mathbf{div}\,\boldsymbol{\sigma}_{h,0}||_{0,\Omega} = \left\{\sum_{T\in\mathcal{T}_h} \|\mathbf{f} + \mathbf{div}\,\boldsymbol{\sigma}_{h,0}\|_{0,T}^2\right\}^{1/2}.$$

Our next goal is to bound the remaining term $\|\mathcal{R}_1\|_{\mathbb{H}_0'}$, for which we need some preliminary results. We begin with the following lemma showing the existence of a stable Helmholtz decomposition for $\mathbb{H}(\mathbf{div};\Omega)$. For its proof we refer the reader to Gatica *et al.* (2011d, Lemma 3.3).

LEMMA 6.3 There exists $C_{\rm hel} > 0$, such that every $\boldsymbol{\tau} \in \mathbb{H}(\mathbf{div};\Omega)$ can be decomposed as $\boldsymbol{\tau} = \boldsymbol{\eta} + \mathbf{curl}\,\boldsymbol{\chi}$, where $\boldsymbol{\eta} \in \mathbb{H}^1(\Omega)$, $\boldsymbol{\chi} \in \mathbf{H}^1(\Omega)$ and

$$\|\boldsymbol{\eta}\|_{1,\Omega} + \|\boldsymbol{\chi}\|_{1,\Omega} \leqslant C_{\rm hel}\|\boldsymbol{\tau}\|_{\mathbf{div},\Omega}.$$

We now recall three well-known approximation operators: the orthogonal projector from $\mathbf{L}^2(\Omega)$ into Q_h (see Di Pietro & Ern, 2012, Lemma 1.58), the Raviart–Thomas interpolator (see Boffi *et al.*, 2013, Section 2.3.1 or Gatica, 2014, Section 3.4.4) and the Clément operator onto the space of continuous piecewise linear functions (see Clément, 1975).

The orthogonal projector $\mathcal{P}_h^k : L^2(\Omega) \to Q_h$ is characterized by the following identity:

$$(\mathcal{P}_h^k(v), z_h)_\Omega = (v, z_h)_\Omega \quad \forall z_h \in Q_h. \tag{6.6}$$

In addition, it is well known that, for each $v \in H^s(\Omega)$, with $s \in \{0, \ldots, k+1\}$, there holds

$$|v - \mathcal{P}_h^k(v)|_{m,T} \leqslant C h_T^{s-m} |v|_{s,T} \quad \forall T \in \mathcal{T}_h,\ \forall m \in \{0, \ldots, s\}. \tag{6.7}$$

The Raviart–Thomas interpolation operator $\Pi_h^k : \mathbf{H}^1(\Omega) \to \mathbf{H}_h$ (recall the discrete spaces in Section 5.1), given $\tau \in \mathbf{H}^1(\Omega)$, is characterized by the following identities:

$$\int_e (\Pi_h^k \tau \cdot \mathbf{n}) r = \int_e (\tau \cdot \mathbf{n}) r \ \ \forall \text{ edges } e \text{ of } \mathcal{T}_h,\ \forall r \in P_k(e), \quad \text{when } k \geqslant 0, \tag{6.8}$$

and

$$\int_T \Pi_h^k \tau \cdot \mathbf{r} = \int_T \tau \cdot \mathbf{r} \ \ \forall T \in \mathcal{T}_h,\ \forall \mathbf{r} \in [P_{k-1}(T)]^2, \quad \text{when } k \geqslant 1. \tag{6.9}$$

As a consequence of (6.8) and (6.9), there holds

$$\operatorname{div}(\Pi_h^k \tau) = \mathcal{P}_h^k(\operatorname{div} \tau). \tag{6.10}$$

In addition, the operator Π_h^k satisfies the following approximation properties (see, for instance, Boffi *et al.*, 2013, Proposition 2.5.4 or Gatica, 2014, Lemma 3.17, Lemma 3.18):

$$\|\tau - \Pi_h^k(\tau)\|_{0,T} \leqslant c_1 h_T^m |\tau|_{m,T} \quad \forall\, T \in \mathcal{T}_h, \tag{6.11}$$

for each $\tau \in \mathbf{H}^m(\Omega)$, with $m \in \{1, \ldots, k+1\}$,

$$\|\operatorname{div}(\tau - \Pi_h^k(\tau))\|_{0,T} \leqslant c_2 h_T^m |\operatorname{div} \tau|_{m,T} \quad \forall\, T \in \mathcal{T}_h, \tag{6.12}$$

for each $\tau \in \mathbf{H}^m(\Omega)$, such that $\operatorname{div} \tau \in H^m(\Omega)$, with $m \in \{0, \ldots, k+1\}$ and

$$\|\tau \cdot \mathbf{n} - \Pi_h^k(\tau) \cdot \mathbf{n}\|_{0,e} \leqslant c_3 h_e^{1/2} \|\tau\|_{1,T_e} \quad \forall \text{ edges } e \in \mathcal{T}_h, \tag{6.13}$$

for each $\tau \in \mathbf{H}^1(\Omega)$, where $T_e \in \mathcal{T}_h$ contains e on its boundary.

The Clément operator $I_h : H^1(\Omega) \to X_h$ approximates optimally nonsmooth functions by continuous piecewise linear functions, where

$$X_h := \{v \in C(\bar{\Omega}) :\ v|_T \in P_1(T)\ \forall T \in \mathcal{T}_h\}.$$

Moreover, the operator I_h satisfies the following approximation properties (see Clément, 1975):

$$\|v - I_h v\|_{k,T} \leqslant c_3 h_T^{1-k} \|v\|_{1,\Delta T} \ \ \forall T \in \mathcal{T}_h, k = 0, 1, \quad \text{and} \quad \|v - I_h v\|_{0,e} \leqslant c_4 h_e^{1/2} \|v\|_{1,\Delta e} \ \ \forall e \in \mathcal{E}_h, \tag{6.14}$$

for all $v \in H^1(\Omega)$, where ΔT and Δe are the unions of all elements intersecting with T and e, respectively.

At this point, we recall that each operator defined above is uniformly bounded, that is, there exist positive constants $C_{\mathcal{P}}$, C_{Π} and C_I, independent of h, such that

$$\|\mathcal{P}_h^k(v)\|_{0,\Omega} \leqslant C_{\mathcal{P}}\|v\|_{0,\Omega}, \quad \|\Pi_h^k(\tau)\|_{\mathrm{div},\Omega} \leqslant C_{\Pi}\|\tau\|_{1,\Omega}, \quad \|I_h^k(z)\|_{1,\Omega} \leqslant C_I\|z\|_{1,\Omega}, \tag{6.15}$$

for all $v \in L^2(\Omega)$, $\tau \in \mathbf{H}^1(\Omega)$ and $z \in H^1(\Omega)$.

We conclude the description of the interpolation operators by mentioning that, in what follows, we will make use of the vector version of $\mathcal{P}_h^k$ and I_h, say $\mathbf{P}_h^k : \mathbf{L}^2(\Omega) \to \mathbf{Q}_h$ and $\mathbf{I}_h : \mathbf{H}^1(\Omega) \to \mathbf{X}_h := X_h \times X_h$, respectively, each of them defined component-wise by $\mathcal{P}_h^k$ and I_h, respectively. In addition, we will make use of the tensor version of Π_h^k, say $\boldsymbol{\Pi}_h^k : \mathbb{H}^1(\Omega) \to \mathbb{H}_h$, defined row-wise by Π_h^k, and the tensor version of $\mathcal{P}_h^k$, say $\mathbb{P}_h^k$, defined component-wise by $\mathcal{P}_h^k$. Clearly, $\mathbf{P}_h^k$, $\mathbb{P}_h^k$, $\mathbf{I}_h$ and $\boldsymbol{\Pi}_h^k$ inherit the same approximation properties stated above.

The next lemma establishes a technical result, which is required to estimate $\|\mathcal{R}_1\|_{\mathbb{H}_0'}$.

LEMMA 6.4 Let $\boldsymbol{\eta} \in \mathbb{H}^1(\Omega)$, $\boldsymbol{\chi} \in \mathbf{H}^1(\Omega)$ and $g = (1/2|\Omega|)(\mathrm{tr}(\boldsymbol{\Pi}_h^k\boldsymbol{\eta} + \mathbf{curl}\,(\mathbf{I}_h\boldsymbol{\chi})), 1)_\Omega$. Then, there hold

$$|\mathcal{R}_1(\boldsymbol{\eta} - \boldsymbol{\Pi}_h^k\boldsymbol{\eta})| \leqslant c_1 \sum_{T\in\mathcal{T}_h} h_T \left\|\nabla\mathbf{u}_h - \left(\frac{\nu^{-1}}{\rho}\boldsymbol{\sigma}_{h,0}^{\mathrm{D}} - \frac{1}{2}\left(\mathbf{u}_h \cdot \frac{\nabla\rho}{\rho}\right) I\right)\right\|_{0,T} \|\boldsymbol{\eta}\|_{1,T}, \tag{6.16}$$

$$\begin{aligned}|\mathcal{R}_1(\mathbf{curl}\,(\boldsymbol{\chi} - \mathbf{I}_h\boldsymbol{\chi}))| \leqslant\ & c_2 \sum_{T\in\mathcal{T}_h} h_T \left\|\mathbf{rot}\,\left(\frac{\nu^{-1}}{\rho}\boldsymbol{\sigma}_{h,0}^{\mathrm{D}} - \frac{1}{2}\left(\mathbf{u}_h \cdot \frac{\nabla\rho}{\rho}\right) I\right)\right\|_{0,T} \|\boldsymbol{\chi}\|_{1,\Delta T}\\ &+ c_3 \sum_{e\in\mathcal{E}_h(\Omega)} h_e^{1/2} \left\|\left[\left(\frac{\nu^{-1}}{\rho}\boldsymbol{\sigma}_{h,0}^{\mathrm{D}} - \frac{1}{2}\left(\mathbf{u}_h \cdot \frac{\nabla\rho}{\rho}\right) I\right)\mathbf{t}\right]\right\|_{0,e} \|\boldsymbol{\chi}\|_{1,\Delta e}\end{aligned} \tag{6.17}$$

and

$$|\mathcal{R}_1(gI)| \leqslant \frac{\max\{C_I, C_\Pi\}}{\sqrt{2|\Omega|}} \left\|\frac{\nabla\rho}{\rho}\right\|_{0,\Omega} \|\mathbf{u} - \mathbf{u}_h\|_{0,\Omega} \left(\|\boldsymbol{\eta}\|_{1,\Omega} + \|\boldsymbol{\chi}\|_{1,\Omega}\right). \tag{6.18}$$

Proof. In what follows, we proceed similarly to the proof of Gatica *et al.* (2011d, Lemma 3.6). To do that, and for the sake of simplicity, we first introduce the following notation:

$$\boldsymbol{\zeta} := \left(\frac{\nu^{-1}}{\rho}\boldsymbol{\sigma}_{h,0}^{\mathrm{D}} - \frac{1}{2}\left(\mathbf{u}_h \cdot \frac{\nabla\rho}{\rho}\right) I\right).$$

Since $\nabla\mathbf{u}_h|_T \in [P_{k-1}(T)]^{2\times2}$ for all $T \in \mathcal{T}_h$, from (6.9) we obtain

$$\int_T \nabla\mathbf{u}_h : (\boldsymbol{\eta} - \boldsymbol{\Pi}_h^k\boldsymbol{\eta}) = 0 \quad \forall T \in \mathcal{T}_h,$$

and according to (6.10) and the definition of $\mathcal{R}_1$, we deduce that

$$\mathcal{R}_1(\boldsymbol{\eta} - \boldsymbol{\Pi}_h^k\boldsymbol{\eta}) = \sum_{T\in\mathcal{T}_h} \int_T (\nabla\mathbf{u}_h - \boldsymbol{\zeta}) : (\boldsymbol{\eta} - \boldsymbol{\Pi}_h^k\boldsymbol{\eta}),$$

which together with (6.11) yields (6.16).

Next, using that div $(\mathbf{curl}\,(\boldsymbol{\chi} - \mathbf{I}_h\boldsymbol{\chi})) = 0$, and integrating by parts on each $T \in \mathcal{T}_h$, we obtain

$$\begin{aligned}\mathcal{R}_1(\mathbf{curl}\,(\boldsymbol{\chi} - \mathbf{I}_h\boldsymbol{\chi})) &= \sum_{T\in\mathcal{T}_h} - \int_T \boldsymbol{\zeta} : \mathbf{curl}\,(\boldsymbol{\chi} - \mathbf{I}_h\boldsymbol{\chi}) \\ &= \sum_{T\in\mathcal{T}_h} \left\{ - \int_T (\boldsymbol{\chi} - \mathbf{I}_h\boldsymbol{\chi}) \cdot \mathbf{rot}\,(\boldsymbol{\zeta}) + \int_{\partial T} (\boldsymbol{\zeta}\mathbf{t}) \cdot (\boldsymbol{\chi} - \mathbf{I}_h\boldsymbol{\chi}) \right\} \\ &= - \sum_{T\in\mathcal{T}_h} \int_T (\boldsymbol{\chi} - \mathbf{I}_h\boldsymbol{\chi}) \cdot \mathbf{rot}\,(\boldsymbol{\zeta}) + \sum_{e\in\mathcal{E}_h} \int_e [\boldsymbol{\zeta}\mathbf{t}] \cdot (\boldsymbol{\chi} - \mathbf{I}_h\boldsymbol{\chi}),\end{aligned}$$

and then, from the approximation properties of $\mathbf{I}_h^k$ in (6.14), and applying Hölder's and the triangle inequalities, we obtain (6.17).

Finally, recalling that $(\mathbf{u} \cdot (\nabla\rho/\rho), 1)_\Omega = 0$ (see Lemma 2.1), and using the continuity of $\boldsymbol{\Pi}_h^k$ and $\mathbf{I}_h$, it is easy to obtain

$$\begin{aligned}|\mathcal{R}_1(gI)| &= \left| \frac{1}{2|\Omega|} \int_\Omega \mathrm{tr}\, \left(\boldsymbol{\Pi}_h^k\boldsymbol{\eta} + \mathbf{curl}\,(\mathbf{I}_h\boldsymbol{\chi})\right) \int_\Omega (\mathbf{u}_h - \mathbf{u}) \cdot \frac{\nabla\rho}{\rho} \right| \\ &\leqslant \frac{1}{\sqrt{2|\Omega|}} \left\| \frac{\nabla\rho}{\rho} \right\|_{0,\Omega} \|\mathbf{u} - \mathbf{u}_h\|_{0,\Omega} (\|\boldsymbol{\Pi}_h^k\boldsymbol{\eta}\|_{0,\Omega} + \|\mathbf{curl}\,(\mathbf{I}_h\boldsymbol{\chi})\|_{0,\Omega}) \\ &\leqslant \frac{1}{\sqrt{2|\Omega|}} \left\| \frac{\nabla\rho}{\rho} \right\|_{0,\Omega} \|\mathbf{u} - \mathbf{u}_h\|_{0,\Omega} (C_\Pi \|\boldsymbol{\eta}\|_{1,\Omega} + C_I \|\boldsymbol{\chi}\|_{1,\Omega}),\end{aligned}$$

which yields (6.18) and completes the proof. □

The following lemma establishes the estimate for $\mathcal{R}_1$.

LEMMA 6.5 There exist $C > 0$, independent of h, such that

$$\|\mathcal{R}_1\|_{\mathbb{H}_0'} \leqslant C \left\{ \sum_{T\in\mathcal{T}_h} \hat{\Theta}_T^2 \right\}^{1/2} + C_{\mathrm{ap}} \left\| \frac{\nabla\rho}{\rho} \right\|_{0,\Omega} \|\mathbf{u} - \mathbf{u}_h\|_{0,\Omega},$$

where $C_{\mathrm{ap}} = C_{\mathrm{hel}} \max\{C_I, C_\Pi\}/\sqrt{2|\Omega|}$ and, for each $T \in \mathcal{T}_h$,

$$\begin{aligned}\hat{\Theta}_T^2 := h_T^2 \left\| \mathbf{rot}\, \left(\frac{\nu^{-1}}{\rho} \boldsymbol{\sigma}_{h,0}^{\mathrm{D}} - \frac{1}{2} \left(\mathbf{u}_h \cdot \frac{\nabla\rho}{\rho} \right) I \right) \right\|_{0,T}^2 + h_T^2 \left\| \nabla\mathbf{u}_h - \left(\frac{\nu^{-1}}{\rho} \boldsymbol{\sigma}_{h,0}^{\mathrm{D}} - \frac{1}{2} \left(\mathbf{u}_h \cdot \frac{\nabla\rho}{\rho} \right) I \right) \right\|_{0,T}^2 \\ + \sum_{e\in\mathcal{E}(T)} h_e \left\| \left[\left(\frac{\nu^{-1}}{\rho} \boldsymbol{\sigma}_{h,0}^{\mathrm{D}} - \frac{1}{2} \left(\mathbf{u}_h \cdot \frac{\nabla\rho}{\rho} \right) I \right) \mathbf{t} \right] \right\|_{0,e}^2 .\end{aligned}$$

Proof. Let $\boldsymbol{\tau} \in \mathbb{H}_0(\mathbf{div}, \Omega)$. It follows from Lemma 6.3 that there exist $\boldsymbol{\eta} \in \mathbb{H}^1(\Omega)$ and $\boldsymbol{\chi} \in \mathbf{H}^1(\Omega)$ such that $\boldsymbol{\tau} = \boldsymbol{\eta} + \mathbf{curl}\,\boldsymbol{\chi}$ in Ω and

$$\|\boldsymbol{\eta}\|_{1,\Omega} + \|\boldsymbol{\chi}\|_{1,\Omega} \leqslant C_{\mathrm{hel}} \|\boldsymbol{\tau}\|_{\mathbf{div},\Omega}. \tag{6.19}$$

Then, since $\mathcal{R}_1(\boldsymbol{\tau}_h) = 0$ for all $\boldsymbol{\tau}_h \in \mathbb{H}_{h,0}(\Omega)$, which follows from the first equation of the Galerkin scheme (4.3), we obtain

$$\mathcal{R}_1(\boldsymbol{\tau}) = \mathcal{R}_1(\boldsymbol{\tau} - \boldsymbol{\tau}_h) \quad \forall \boldsymbol{\tau}_h \in \mathbb{H}_{h,0}(\Omega).$$

In particular, for $\boldsymbol{\tau}_h := \Pi_h^k \boldsymbol{\eta} + \mathbf{curl}\,(\mathbf{I}_h \boldsymbol{\chi}) - gI$, with $g = (1/2|\Omega|)\int_\Omega \mathrm{tr}(\boldsymbol{\Pi}_h^k \boldsymbol{\eta} + \mathbf{curl}\,(\mathbf{I}_h \boldsymbol{\chi}))$, we obtain

$$\mathcal{R}_1(\boldsymbol{\tau}) = \mathcal{R}_1(\boldsymbol{\eta} - \boldsymbol{\Pi}_h^k \boldsymbol{\eta}) + \mathcal{R}_1(\mathbf{curl}\,(\boldsymbol{\chi} - \mathbf{I}_h \boldsymbol{\chi})) + \mathcal{R}_1(gI).$$

Hence, the proof follows from Lemma 6.4, estimate (6.19) and the fact that the numbers of triangles in $\#\Delta T$ and $\#\Delta e$ are bounded. □

We end this section by observing that the reliability estimate (6.3) is a direct consequence of Lemmas 6.2 and 6.5 and assumption (6.2).

6.2 *Efficiency of the a posteriori error estimator*

The main result of this section is stated as follows.

THEOREM 6.6 There exists $C_{\mathrm{eff}} > 0$, independent of h, such that

$$C_{\mathrm{eff}} \Theta \leqslant \|\boldsymbol{\sigma}_0 - \boldsymbol{\sigma}_{h,0}\|_{\mathbf{div},\Omega} + \|\mathbf{u} - \mathbf{u}_h\|_{0,\Omega} + \text{h.o.t.}, \tag{6.20}$$

where h.o.t. stands for higher-order terms.

We remark in advance that the proof of (6.20) makes frequent use of the identities provided by Theorem 3.5 and Remark 3.6. We begin with the estimate for the zero-order term appearing in the definition of Θ_T.

LEMMA 6.7 There holds

$$\|\mathbf{f} + \mathbf{div}\,\boldsymbol{\sigma}_{h,0}\|_{0,T} \leqslant \|\boldsymbol{\sigma}_0 - \boldsymbol{\sigma}_{h,0}\|_{\mathbf{div},T} \quad \forall T \in \mathcal{T}_h.$$

Proof. It suffices to recall, as established by Remark 3.6, that $\mathbf{f} = -\mathbf{div}\,\boldsymbol{\sigma}_0$ in Ω. □

In order to derive the upper bounds for the remaining terms defining the global *a posteriori* error estimator Θ (cf. (6.1)), we use results from Carstensen (1997), inverse inequalities, and the localization technique based on element-bubble and edge-bubble functions. To this end, we now introduce further notation and preliminary results. Given $T \in \mathcal{T}_h$ and $e \in \mathcal{E}(T)$, we let ϕ_T and ϕ_e be the usual element-bubble and edge-bubble functions, respectively (see Verfürth, 1996 for details). In particular, ϕ_T satisfies $\phi_T \in P_3(T)$, $\mathrm{supp}\,\phi_T \subseteq T$, $\phi_T = 0$ on ∂T and $0 \leqslant \phi_T \leqslant 1$ in T. Similarly, $\phi_e|_T \in P_2(T)$, $\mathrm{supp}\,\phi_e \subseteq w_e := \bigcup\{T' \in \mathcal{T} : e \in \mathcal{E}(T')\}$, $\phi_e = 0$ on $\partial T \backslash e$ and $0 \leqslant \phi_e \leqslant 1$ in w_e. We also recall from Verfürth (1994) that, given $k \in \mathbb{N} \cup \{0\}$, there exists an extension operator $L : C(e) \to C(w_e)$ that satisfies $L(p) \in P_k(T)$ and $L(p)|_e = p$ for all $p \in P_k(e)$. A corresponding vector version of L, that is, the component-wise application of L, is denoted by $\mathbf{L}$. Additional properties of ϕ_T, ϕ_e and L are collected in the following lemma. See Verfürth (1994, Lemma 1.3) for its proof.

LEMMA 6.8 Given $k \in \mathbb{N} \cup \{0\}$, there exist positive constants c_1, c_2, c_3 and c_4, depending only on k and the shape regularity of the triangulations (minimum angle condition), such that, for each triangle T and

$e \in \mathcal{E}(T)$, there hold

$$\|\phi_T q\|_{0,T}^2 \leqslant \|q\|_{0,T}^2 \leqslant c_1 \|\phi_T^{1/2} q\|_{0,T}^2 \quad \forall q \in P_k(T), \tag{6.21}$$

$$\|\phi_e L(p)\|_{0,e}^2 \leqslant \|p\|_{0,e}^2 \leqslant c_2 \|\phi_e^{1/2} p\|_{0,e}^2 \quad \forall p \in P_k(e) \tag{6.22}$$

and

$$c_3 h_e^{1/2} \|p\|_{0,e} \leqslant \|\phi_e^{1/2} L(p)\|_{0,T} \leqslant c_4 h_e^{1/2} \|p\|_{0,e} \quad \forall p \in P_k(e). \tag{6.23}$$

The following inverse estimate will be also used. We refer the reader to Ciarlet (1978, Theorem 3.2.6) for its proof.

LEMMA 6.9 Let $k, l, m \in \mathbb{N} \cup \{0\}$ such that $l \leqslant m$. Then, there exists $c > 0$, depending only on k, l, m and the shape regularity of the triangulations, such that, for each triangle T, there holds

$$|q|_{m,t} \leqslant c h_T^{l-m} |q|_{l,T} \quad \forall q \in P_k(T). \tag{6.24}$$

In addition, we shall make use of the following estimate for smooth functions (see, for instance, Di Pietro & Ern, 2012):

$$\|v\|_{0,e}^2 \leqslant C(h_e^{-1} \|v\|_{0,T}^2 + h_e |v|_{1,T}^2) \quad \forall v \in H^1(T), \tag{6.25}$$

where T is a generic triangle having e as an edge, and C is a constant depending only on the minimum angle of T.

Finally, in order to simplify the notation, we define

$$\boldsymbol{\zeta} := \frac{\nu^{-1}}{\rho} \boldsymbol{\sigma}_{h,0}^{\mathrm{D}} - \frac{1}{2} \left(\mathbf{u}_h \cdot \frac{\nabla \rho}{\rho} \right) I \tag{6.26}$$

and

$$\mathcal{M} := \frac{\nu^{-1}}{\rho} (\boldsymbol{\sigma}_{h,0}^{\mathrm{D}} - \boldsymbol{\sigma}_0^{\mathrm{D}}) - \frac{1}{2} \left((\mathbf{u}_h - \mathbf{u}) \cdot \frac{\nabla \rho}{\rho} \right) I. \tag{6.27}$$

Observe that

$$\|\mathcal{M}\|_{0,T} \leqslant C\{\|\boldsymbol{\sigma}_0 - \boldsymbol{\sigma}_{h,0}\|_{\mathbf{div},T} + \|\mathbf{u} - \mathbf{u}_h\|_{0,T}\} \quad \forall T \in \mathcal{T}_h. \tag{6.28}$$

In the sequel, we assume that ρ^{-1} and $\rho^{-1} \nabla \rho$ are at least in $H^{k+2}(T)$ and $[H^{k+2}(T)]^2$, respectively, for all $T \in \mathcal{T}_h$.

Now we estimate the rest of the terms defining the *a posteriori* error estimator Θ_T, separately.

LEMMA 6.10 There exists $C > 0$, independent of h, such that

$$h_T \left\| \mathbf{rot} \left(\frac{\nu^{-1}}{\rho} \boldsymbol{\sigma}_{h,0}^{\mathrm{D}} - \frac{1}{2} \left(\mathbf{u}_h \cdot \frac{\nabla \rho}{\rho} \right) I \right) \right\|_{0,T} \leqslant C\{\|\boldsymbol{\sigma}_0 - \boldsymbol{\sigma}_{h,0}\|_{\mathbf{div},T} + \|\mathbf{u} - \mathbf{u}_h\|_{0,T} + \text{h.o.t.}\}$$

for all $T \in \mathcal{T}_h$.

Proof. First, adding and subtracting $\mathbf{P}_h^r(\boldsymbol{\zeta})$, with $r \geqslant k+1$, and using the triangle inequality, we obtain

$$\begin{aligned}\|\mathbf{rot}\,(\boldsymbol{\zeta})\|_{0,T} &\leqslant \|\mathbf{rot}\,(\boldsymbol{\zeta}-\mathbf{P}_h^r\boldsymbol{\zeta})\|_{0,T} + \|\mathbf{rot}\,(\mathbf{P}_h^r\boldsymbol{\zeta})\|_{0,T}\\ &\leqslant |\boldsymbol{\zeta}-\mathbf{P}_h^r\boldsymbol{\zeta}|_{1,T} + \|\mathbf{rot}\,(\mathbf{P}_h^r\boldsymbol{\zeta})\|_{0,T}.\end{aligned} \tag{6.29}$$

Then, in what follows we proceed as in the proof of Carstensen (1997, Lemma 6.1) to estimate $\|\mathbf{rot}\,(\mathbf{P}_h^r\boldsymbol{\zeta})\|_{0,T}$. In fact, since $\nabla\mathbf{u} = -\boldsymbol{\zeta} + \mathcal{M} = (\nu^{-1}/\rho)\boldsymbol{\sigma}_0^{\mathrm{D}} - \frac{1}{2}(\mathbf{u}\cdot(\nabla\rho/\rho))I$ in Ω, from (6.21), and integrating by parts, we find that

$$\begin{aligned}\|\mathbf{rot}\,(\mathbf{P}_h^r\boldsymbol{\zeta})\|_{0,T}^2 &\leqslant c_1\|\phi_T^{1/2}\mathbf{rot}\,(\mathbf{P}_h^r\boldsymbol{\zeta})\|_{0,T}^2\\ &= c_1\int_T \phi_T\mathbf{rot}\,(\mathbf{P}_h^r\boldsymbol{\zeta})\mathbf{rot}\,(\mathbf{P}_h^r\boldsymbol{\zeta}-\boldsymbol{\zeta}+\mathcal{M})\\ &= c_1\int_T (\mathbf{P}_h^r\boldsymbol{\zeta}-\boldsymbol{\zeta}+\mathcal{M}) : \mathbf{curl}\,(\phi_T\mathbf{rot}\,(\mathbf{P}_h^r\boldsymbol{\zeta})).\end{aligned}$$

Therefore, from (6.21), (6.24), (6.28), we obtain

$$\|\mathbf{rot}\,(\mathbf{P}_h^r\boldsymbol{\zeta})\|_{0,T} \leqslant Ch_T^{-1}\{\|\boldsymbol{\zeta}-\mathbf{P}_h^r\boldsymbol{\zeta}\|_{0,T} + \|\boldsymbol{\sigma}_0-\boldsymbol{\sigma}_{h,0}\|_{\mathbf{div},T} + \|\mathbf{u}-\mathbf{u}_h\|_{0,T}\}, \tag{6.30}$$

which together with (6.7) and (6.29) implies

$$\begin{aligned}h_T\|\mathbf{rot}\,(\boldsymbol{\zeta})\|_{0,T} &\leqslant C\{h_T|\boldsymbol{\zeta}-\mathbf{P}_h^r\boldsymbol{\zeta}|_{1,T} + \|\boldsymbol{\zeta}-\mathbf{P}_h^r\boldsymbol{\zeta}\|_{0,T} + \|\boldsymbol{\sigma}_0-\boldsymbol{\sigma}_{h,0}\|_{\mathbf{div},T} + \|\mathbf{u}-\mathbf{u}_h\|_{0,T}\}\\ &\leqslant C\{\|\boldsymbol{\sigma}_0-\boldsymbol{\sigma}_{h,0}\|_{\mathbf{div},T} + \|\mathbf{u}-\mathbf{u}_h\|_{0,T} + h_T^{r+1}|\boldsymbol{\zeta}|_{r+1,T}\},\end{aligned}$$

which concludes the proof. □

LEMMA 6.11 There exists $C>0$, independent of h, such that

$$h_T\|\nabla\mathbf{u}_h-\boldsymbol{\zeta}\|_{0,T} \leqslant C\{h_T\|\boldsymbol{\sigma}_0-\boldsymbol{\sigma}_{h,0}\|_{\mathbf{div},T} + (h_T+1)\|\mathbf{u}-\mathbf{u}_h\|_{0,T} + \text{h.o.t.}\}$$

for all $T\in\mathcal{T}_h$.

Proof. Given $r \geqslant k+1$, we add and subtract $\mathbf{P}_h^r\boldsymbol{\zeta}$, and use the triangle inequality to obtain

$$\|\nabla\mathbf{u}_h-\boldsymbol{\zeta}\|_{0,T} \leqslant \|\boldsymbol{\zeta}-\mathbf{P}_h^r\boldsymbol{\zeta}\|_{0,T} + \|\nabla\mathbf{u}_h-\mathbf{P}_h^r\boldsymbol{\zeta}\|_{0,T}. \tag{6.31}$$

Then, proceeding similarly to the proof of Carstensen (1997, Lemma 6.3), noting that $\boldsymbol{\zeta}-\mathcal{M}-\nabla\mathbf{u}=0$ in Ω, integrating by parts and using (6.21), we find that

$$\begin{aligned}\|\nabla\mathbf{u}_h-\mathbf{P}_h^r\boldsymbol{\zeta}\|_{0,T}^2 &\leqslant c_1\|\phi_T^{1/2}(\nabla\mathbf{u}_h-\mathbf{P}_h^r\boldsymbol{\zeta})\|_{0,T}^2\\ &= c_1\int_T \phi_T(\nabla\mathbf{u}_h-\mathbf{P}_h^r\boldsymbol{\zeta})(\boldsymbol{\zeta}-\mathcal{M}+\nabla(\mathbf{u}_h-\mathbf{u})-\mathbf{P}_h^r))\\ &= c_1\left\{\int_T(\boldsymbol{\zeta}-\mathbf{P}_h^r\boldsymbol{\zeta}-\mathcal{M})\phi_T(\nabla\mathbf{u}_h-\mathbf{P}_h^r\boldsymbol{\zeta}) - \int_T(\mathbf{u}_h-\mathbf{u})\cdot\mathbf{div}(\phi_T(\nabla\mathbf{u}_h-\mathbf{P}_h^r\boldsymbol{\zeta}))\right\}.\end{aligned}$$

Therefore, applying the Cauchy–Schwarz and the triangle inequalities, from (6.21), (6.24) and (6.28), we obtain

$$\|\nabla \mathbf{u}_h - \mathbf{P}_h^r \boldsymbol{\zeta}\|_{0,T} \leqslant C\{\|\boldsymbol{\zeta} - \mathbf{P}_h^r \boldsymbol{\zeta}\|_{0,T} + \|\boldsymbol{\sigma}_0 - \boldsymbol{\sigma}_{h,0}\|_{\mathbf{div},T} + (1 + h_T^{-1})\|\mathbf{u} - \mathbf{u}_h\|_{0,T}\},$$

which together with (6.7) and (6.31) implies

$$h_T\|\nabla \mathbf{u}_h - \boldsymbol{\zeta}\|_{0,T} \leqslant C\{h_T\|\boldsymbol{\sigma}_0 - \boldsymbol{\sigma}_{h,0}\|_{\mathbf{div},T} + (h_T + 1)\|\mathbf{u} - \mathbf{u}_h\|_{0,T} + (h_T + 1)h_T^{r+1}|\boldsymbol{\zeta}|_{r+1,T}\},$$

which concludes the proof. □

LEMMA 6.12 There exists $C > 0$, independent of h, such that

$$h_e^{1/2}\left\|\left[\left(\frac{\nu^{-1}}{\rho}\boldsymbol{\sigma}_{h,0}^{\mathrm{D}} - \frac{1}{2}\left(\mathbf{u}_h \cdot \frac{\nabla\rho}{\rho}\right) I\right)\mathbf{t}\right]\right\|_{0,e} \leqslant C\sum_{T\subseteq w_e}(\|\boldsymbol{\sigma}_0 - \boldsymbol{\sigma}_{h,0}\|_{\mathbf{div},T} + \|\mathbf{u} - \mathbf{u}_h\|_{0,T}) + \text{h.o.t.} \quad \forall e \in \mathcal{E}_h.$$

Proof. Given $e \in \mathcal{E}_h$ and $l \geqslant k+1$, we add and subtract $\mathbf{P}_h^l\boldsymbol{\zeta}$, and utilize the triangle inequality to obtain

$$\|[\boldsymbol{\zeta}\mathbf{t}]\|_{0,e} \leqslant \|[(\boldsymbol{\zeta} - \mathbf{P}_h^l\boldsymbol{\zeta})\mathbf{t}]\|_{0,e} + \|[\mathbf{P}_h^l\boldsymbol{\zeta}\mathbf{t}]\|_{0,e}. \tag{6.32}$$

Observe that, according to inequality (6.25), and the fact that $h_e \leqslant h_T$ for $T \subseteq w_e$, the following inequality holds:

$$\begin{aligned} h_e^{1/2}\|[(\boldsymbol{\zeta} - \mathbf{P}_h^l\boldsymbol{\zeta})\mathbf{t}]\|_{0,e} &\leqslant Ch_e^{1/2}\left\{(h_e^{-1/2}\sum_{T\subseteq w_e}\|\boldsymbol{\zeta} - \mathbf{P}_h^l\boldsymbol{\zeta}\|_{0,T} + h_e^{1/2}\sum_{T\subseteq w_e}|\boldsymbol{\zeta} - \mathbf{P}_h^l\boldsymbol{\zeta}|_{1,T}\right\} \\ &\leqslant C\sum_{T\subseteq w_e} h_T^{l+1}|\boldsymbol{\zeta}|_{l+1,T}. \end{aligned} \tag{6.33}$$

In this way, in the sequel, we proceed analogously to the proof of Carstensen (1997, Lemma 6.2) to bound $\|[\mathbf{P}_h^l\boldsymbol{\zeta}\mathbf{t}]\|_{0,e}$.

First, to simplify the notation, we let $\boldsymbol{\kappa} := [\mathbf{P}_h^l\boldsymbol{\zeta}\mathbf{t}]$. Then, we use (6.22) to obtain

$$\|\boldsymbol{\kappa}\|_{0,e}^2 \leqslant c_2\|\phi_e^{1/2}\boldsymbol{\kappa}\|_{0,e}^2 = c_2\int_e(\phi_e\mathbf{L}(\boldsymbol{\kappa}))\boldsymbol{\kappa}, \tag{6.34}$$

where $\mathbf{L} : [C(e)]^2 \to [C(w_e)]^2$ is the extension operator defined above.

Now, integrating by parts on each $T \in w_e$, and using that $-\boldsymbol{\zeta} + \mathcal{M} = \nabla\mathbf{u} = (\nu^{-1}/\rho)\boldsymbol{\sigma}_0^{\mathrm{D}} - \frac{1}{2}(\mathbf{u}\cdot(\nabla\rho/\rho))I$ in Ω, we find that

$$\int_e(\phi_e\mathbf{L}(\boldsymbol{\kappa}))\boldsymbol{\kappa} = \sum_{T\subseteq w_e}\left\{\int_T(\mathbf{P}_h^l\boldsymbol{\zeta} - \boldsymbol{\zeta} + \mathcal{M}) : \mathbf{curl}\,(\phi_e\mathbf{L}(\boldsymbol{\kappa})) + \int_T(\phi_e\mathbf{L}(\boldsymbol{\kappa}))\mathbf{rot}\,(\mathbf{P}_h^l\boldsymbol{\zeta})\right\}. \tag{6.35}$$

On the other hand, using estimates (6.23) and (6.24) and the fact that $0 \leqslant \phi_e \leqslant 1$, we obtain

$$\|\phi_e\mathbf{L}(\boldsymbol{\kappa})\|_{0,T} \leqslant \|\phi_e^{1/2}\mathbf{L}(\boldsymbol{\kappa})\|_{0,T} \leqslant ch_e^{1/2}\|\boldsymbol{\kappa}\|_{0,e}. \tag{6.36}$$

In this way, from (6.24), (6.34) and (6.35), we deduce that

$$\|\boldsymbol{\kappa}\|_{0,e}^2 \leqslant C \sum_{T\subseteq w_e} \{h_T^{-1}(\|\boldsymbol{\zeta} - \mathbf{P}_h^l\boldsymbol{\zeta}\|_{0,T} + \|\mathcal{M}\|_{0,T}) + \|\mathbf{rot}\,(\mathbf{P}_h^l\boldsymbol{\zeta})\|_{0,T}\}\|\phi_e \mathbf{L}(\boldsymbol{\kappa})\|_{0,T},$$

which together with (6.28), (6.30) and (6.36) implies

$$\|\boldsymbol{\kappa}\|_{0,e} \leqslant C h_e^{1/2} \sum_{T\subseteq w_e} h_T^{-1}\{\|\boldsymbol{\zeta} - \mathbf{P}_h^l\boldsymbol{\zeta}\|_{0,T} + \|\boldsymbol{\sigma}_0 - \boldsymbol{\sigma}_{h,0}\|_{\mathbf{div},T} + \|\mathbf{u} - \mathbf{u}_h\|_{0,T}\}. \tag{6.37}$$

Therefore, from (6.32) (6.37) and the fact that $h_e \leqslant h_T$ for $T \subseteq w_e$, we obtain

$$h_e^{1/2}\|[\boldsymbol{\zeta}\cdot\mathbf{t}]\|_{0,e} \leqslant C\left\{\sum_{T\subseteq w_e} \|\boldsymbol{\sigma}_0 - \boldsymbol{\sigma}_{h,0}\|_{\mathbf{div},T} + \|\mathbf{u} - \mathbf{u}_h\|_{0,T} + h_T^{l+1}|\boldsymbol{\zeta}|_{l+1,T}\right\},$$

which concludes the proof. □

We end this section by observing that the efficiency estimate (6.20) follows straightforwardly from Lemmas 6.10, 6.11 and 6.12.

7. Numerical results

In this section we present two numerical examples in $\mathbf{R}^2$, illustrating the performance of the mixed finite element scheme (4.3), confirming the reliability and efficiency of the *a posteriori* error estimator Θ derived in Section 6, and showing the behaviour of the associated adaptive algorithm. Here we consider the specific finite element subspaces $\mathbb{H}_{h,0}$ and $\mathbf{Q}_h$ defined in terms of the specific discrete spaces given by (5.1) with $k = 0$. In addition, the zero integral mean condition for tensors in the space $\mathbb{H}_{h,0}$ is imposed via a real Lagrange multiplier. In what follows, N stands for the total number of degrees of freedom defining $\mathbb{H}_{h,0} \times \mathbf{Q}_h$. Denoting by $(\boldsymbol{\sigma}_0, \mathbf{u}) \in \mathbb{H}_0 \times \mathbf{Q}$ and $(\boldsymbol{\sigma}_{h,0}, \mathbf{u}_h) \in \mathbb{H}_{h,0} \times \mathbf{Q}_h$, the solutions of (3.2) and (4.3), respectively, the individual errors are defined by

$$\mathrm{e}(\boldsymbol{\sigma}_0) := \|\boldsymbol{\sigma}_0 - \boldsymbol{\sigma}_{h,0}\|_{\mathbf{div},\Omega}, \quad \mathrm{e}(\mathbf{u}) := \|\mathbf{u} - \mathbf{u}_h\|_{0,\Omega}, \quad \mathrm{e}(p) := \|p - p_h\|_{0,\Omega}$$

and

$$\mathrm{e}(\boldsymbol{\sigma}_0, \mathbf{u}) := \{(\mathrm{e}(\boldsymbol{\sigma}_0))^2 + (\mathrm{e}(\mathbf{u}))^2\}^{1/2},$$

where the approximate pressure p_h is computed by the post-processing formula (4.24). The effectivity index with respect to Θ is given by

$$\mathrm{eff}(\Theta) := \mathrm{e}(\boldsymbol{\sigma}_0, \mathbf{u})/\Theta.$$

Furthermore, we define the experimental rates of convergence

$$r(\boldsymbol{\sigma}_0) := \frac{\log(\mathrm{e}(\boldsymbol{\sigma}_0)/\mathrm{e}'(\boldsymbol{\sigma}_0))}{\log(h/h')}, \quad r(\mathbf{u}) := \frac{\log(\mathrm{e}(\mathbf{u})/\mathrm{e}'(\mathbf{u}))}{\log(h/h')},$$

$$r(p) := \frac{\log(\mathrm{e}(p)/\mathrm{e}'(p))}{\log(h/h')}, \quad r(\boldsymbol{\sigma}_0, \mathbf{u}) := \frac{\log(\mathrm{e}(\boldsymbol{\sigma}_0, \mathbf{u})/\mathrm{e}'(\boldsymbol{\sigma}_0, \mathbf{u}))}{\log(h/h')},$$

where h and h' are two consecutive mesh sizes with errors e and e', respectively. However, when the adaptive algorithm is applied (see details below), the expression $\log(h/h')$ appearing in the computation of the above rates is replaced by $-\frac{1}{2}\log(N/N')$, where N and N' denote the corresponding degrees of freedom of each triangulation.

The examples to be considered in this section are described next. For the two of them we choose $\nu = 1$. Example 7.1 is used to illustrate the performance of the mixed finite element scheme (4.3) and to corroborate the reliability and efficiency of the *a posteriori* error estimator Θ. Example 7.2 is utilized to illustrate the behaviour of the associated adaptive algorithm, which applies the following procedure from Verfürth (1996).

(1) Start with a coarse mesh $\mathcal{T}_h$.

(2) Solve the discrete problem (4.3) for the current mesh $\mathcal{T}_h$.

(3) Compute $\Theta_T := \Theta$ for each triangle $T \in \mathcal{T}_h$.

(4) Check the stopping criterion and decide whether to finish or go to the next step.

(5) Use *blue–green* refinement on those $T' \in \mathcal{T}_h$ whose indicator $\Theta_{T'}$ satisfies

$$\Theta_{T'} \geqslant \frac{1}{2}\max_{T\in\mathcal{T}_h}\{\Theta_T : \quad T \in \mathcal{T}_h\}.$$

(6) Define the resulting mesh as the actual mesh $\mathcal{T}_h$ and go to step (2).

EXAMPLE 7.1 We consider the region $\Omega := (-1,1)\times(-1,1)$ and define the density function

$$\rho(x_1,x_2) := \exp(\mu(x_1+x_2)) \quad \forall (x_1,x_2)\in\Omega,$$

where μ is a parameter in $\mathbf{R}$. We note that

$$\left\|\frac{\nabla\rho}{\rho}\right\|_{\mathbf{L}^\infty(\Omega)} = |\mu|, \tag{7.1}$$

and then, as we shall see in Fig. 1, and as predicted in (4.16), the performance of our method depends strongly on the choice of μ.

In turn, we choose the datum $\mathbf{f}$ so that the exact solution is given by the smooth functions

$$\mathbf{u}(x_1,x_2) = \frac{\mathrm{curl}\,(\sin^2(\pi x_1)\sin^2(\pi x_2))}{\rho(x_1,x_2)} \quad \forall (x_1,x_2)\in\Omega,$$
$$p(x_1,x_2) = x_1\sin(x_2) \quad \forall (x_1,x_2)\in\Omega,$$

where $\mathrm{curl}\,\varphi := (\partial\varphi/\partial x_2, -\partial\varphi/\partial x_1)^{\mathrm{T}}$ for any sufficiently smooth function φ.

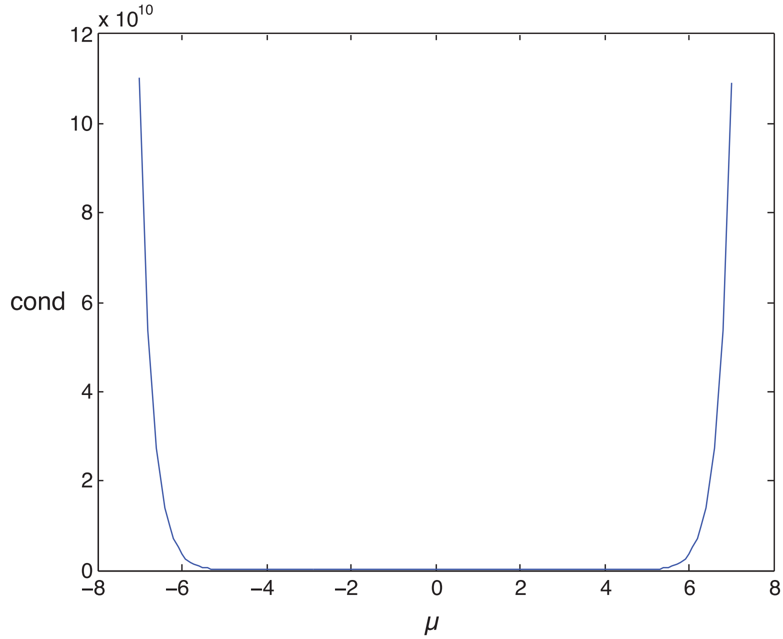

FIG. 1. Example 7.1, μ vs. condition number for $h = \frac{1}{4}$.

EXAMPLE 7.2 We consider the L-shaped domain given by $\Omega := (-1,1)^2 \backslash [0,1]^2$. Then, we choose the density

$$\rho(x_1,x_2) = (x_1 - 0.01)^2 + (x_2 - 0.01)^2 \quad \forall (x_1,x_2) \in \Omega,$$

and the datum $\mathbf{f}$ so that the exact solution is given by

$$\mathbf{u}(x_1,x_2) = \frac{1}{\rho(x_1,x_2)} \mathbf{curl}\,(x_1^2 x_2^2 (x_1^2 - 1)^2 (x_2^2 - 1)^2) \quad \forall (x_1,x_2) \in \Omega,$$

$$p(x_1,x_2) = \frac{x_1 - 0.01}{(x_1 - 0.01)^2 + (x_2 - 0.01)^2} + p_0 \quad \forall (x_1,x_2) \in \Omega, \quad p_0 = 0.4153036413.$$

Note that the fluid velocity $\mathbf{u}$ and the fluid pressure p have high gradients around the origin.

The numerical results shown below were obtained using a MATLAB code. In Table 1 we summarize the convergence history of our mixed finite element scheme (4.3), with $\mu = 2$ and for a set of shape-regular triangulations of the computational domain Ω. We observe there that, looking at the experimental rates of convergence, the $\mathcal{O}(h)$ predicted by Theorem 5.1, with $s = 1$, is attained in all the unknowns. In addition, we note that the effectivity index $\mathrm{eff}(\Theta)$ remains always in a neighbourhood of 0.1, which illustrates the reliability and efficiency of Θ in the case of a regular solution.

Now, having in mind assumption (4.16), in Fig. 1 we display the relation between μ (cf. (7.1)) and the condition number of the global matrix given by the left-hand side of (4.3) computed with the command *condest* in MATLAB, considering a fixed mesh of size $h = \frac{1}{4}$. We observe here that the condition number is stable for $|\mu| \leqslant 6$ and blows up for $|\mu| > 6$. This phenomenon shows that assumption (4.16),

TABLE 1 *Example* 7.1, *uniform scheme*

N	h	$\mathrm{e}(\boldsymbol{\sigma}_0)$	$r(\boldsymbol{\sigma}_0)$	$\mathrm{e}(\mathbf{u})$	$r(\mathbf{u})$	$\mathrm{e}(p)$	$r(p)$
337	$\frac{1}{2}$	3.869e+01	—	1.777e+01	—	3.353e+00	—
1313	$\frac{1}{4}$	1.351e+01	1.548	1.058e+01	0.763	1.798e+00	0.916
5185	$\frac{1}{8}$	5.882e+00	1.211	5.477e+00	0.959	9.026e−01	1.004
20609	$\frac{1}{16}$	2.812e+00	1.070	2.757e+00	0.995	4.493e−01	1.011
82177	$\frac{1}{32}$	1.389e+00	1.020	1.380e+00	1.000	2.242e−01	1.005
328193	$\frac{1}{64}$	6.924e−01	1.006	6.904e−01	1.001	1.121e−01	1.002

N	h	$\mathrm{e}(\boldsymbol{\sigma}_0, \mathbf{u})$	$r(\boldsymbol{\sigma}_0, \mathbf{u})$	Θ	$\mathrm{eff}(\Theta)$
337	$\frac{1}{2}$	4.258e+01	—	3.500e+02	0.122
1313	$\frac{1}{4}$	1.716e+01	1.337	1.682e+02	0.102
5185	$\frac{1}{8}$	8.037e+00	1.104	8.206e+01	0.098
20609	$\frac{1}{16}$	3.938e+00	1.034	3.971e+01	0.099
82177	$\frac{1}{32}$	1.958e+00	1.010	1.939e+01	0.101
328193	$\frac{1}{64}$	9.778e−01	1.003	9.567e+00	0.102

TABLE 2 *Example* 7.2, *uniform scheme*

N	$\mathrm{e}(\boldsymbol{\sigma}_0)$	$\mathrm{e}(\mathbf{u})$	$\mathrm{e}(\boldsymbol{\sigma}_0, \mathbf{u})$	$r(\boldsymbol{\sigma}_0, \mathbf{u})$	Θ	$\mathrm{eff}(\Theta)$
257	2.417e+00	1.955e+00	3.109e+00	—	9.418e+01	0.033
993	3.217e+00	3.074e+00	4.450e+00	−0.530	2.480e+02	0.018
3905	3.369e+00	3.974e+00	5.210e+00	−0.230	5.600e+02	0.009
15489	1.336e+00	2.108e+00	2.496e+00	1.068	3.739e+02	0.007
61697	7.222e−01	7.696e−01	1.055e+00	1.246	1.880e+02	0.006
246273	2.752e−01	2.576e−01	3.769e−01	1.488	8.918e+01	0.004

beyond being just a theoretical hypothesis, in practice ensures the performance of the numerical method for small values of $\|\nabla\rho/\rho\|_{\mathbf{L}^\infty(\Omega)}$.

Next, in Tables 2 and 3 we provide the convergence history of the uniform and adaptive schemes, as applied to Example 7.2. We observe that the errors of the adaptive procedure decrease faster than those obtained by the uniform one, which is confirmed by the global experimental rates of convergence

TABLE 3 *Example* 7.2, *adaptive scheme*

N	$\mathrm{e}(\boldsymbol{\sigma}_0)$	$\mathrm{e}(\mathbf{u})$	$\mathrm{e}(\boldsymbol{\sigma}_0,\mathbf{u})$	$r(\boldsymbol{\sigma}_0,\mathbf{u})$	Θ	$\mathrm{eff}(\Theta)$
257	2.417e+00	1.955e+00	3.109e+00	—	9.418e+01	0.033
429	4.304e+00	3.037e+00	5.268e+00	−2.058	2.486e+02	0.021
557	3.876e+00	4.012e+00	5.578e+00	−0.439	5.836e+02	0.010
725	1.691e+00	2.018e+00	2.633e+00	5.697	3.734e+02	0.007
777	1.611e+00	1.580e+00	2.257e+00	4.448	2.385e+02	0.009
985	1.322e+00	8.983e−01	1.599e+00	2.908	1.667e+02	0.010
1325	1.113e+00	5.064e−01	1.223e+00	1.808	1.017e+02	0.012
1733	1.051e+00	4.166e−01	1.130e+00	0.585	6.965e+01	0.016
2167	1.018e+00	3.713e−01	1.084e+00	0.379	5.378e+01	0.020
3745	9.669e−01	3.107e−01	1.016e+00	0.237	3.706e+01	0.027
6269	8.605e−01	2.596e−01	8.988e−01	0.474	2.687e+01	0.033
9603	8.190e−01	2.181e−01	8.475e−01	0.276	2.089e+01	0.041
19569	5.081e−01	1.354e−01	5.258e−01	1.341	1.467e+01	0.036
35055	4.158e−01	1.039e−01	4.286e−01	0.701	1.059e+01	0.040
68769	2.662e−01	5.823e−02	2.725e−01	1.344	7.537e+00	0.036
99177	2.191e−01	4.864e−02	2.244e−01	1.062	6.150e+00	0.036
207481	1.502e−01	3.037e−02	1.533e−01	1.033	4.289e+00	0.036
311277	1.209e−01	2.371e−02	1.233e−01	1.075	3.445e+00	0.036

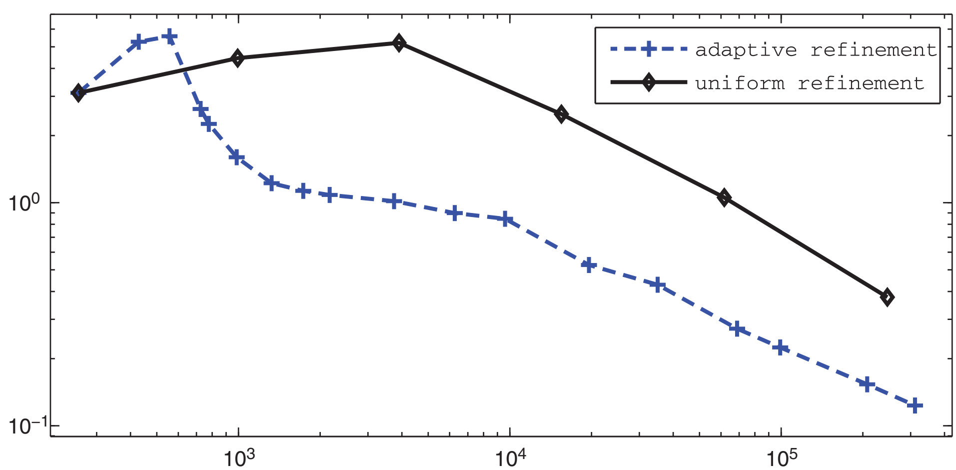

FIG. 2. Example 7.2, $\mathrm{e}(\boldsymbol{\sigma},\mathbf{u})$ vs. N for uniform/adaptive schemes.

provided there. This fact is also illustrated in Fig. 2, where we display the total errors $\mathrm{e}(\boldsymbol{\sigma}_0,\mathbf{u})$ vs. the number of degrees of freedom N for both refinements. As shown by the values of $r(\boldsymbol{\sigma}_0,\mathbf{u})$, the adaptive method is able to keep the quasi-optimal rate of convergence $\mathcal{O}(h)$ for the total error. Furthermore, the effectivity indexes remain bounded from above and below, which confirms the reliability and efficiency

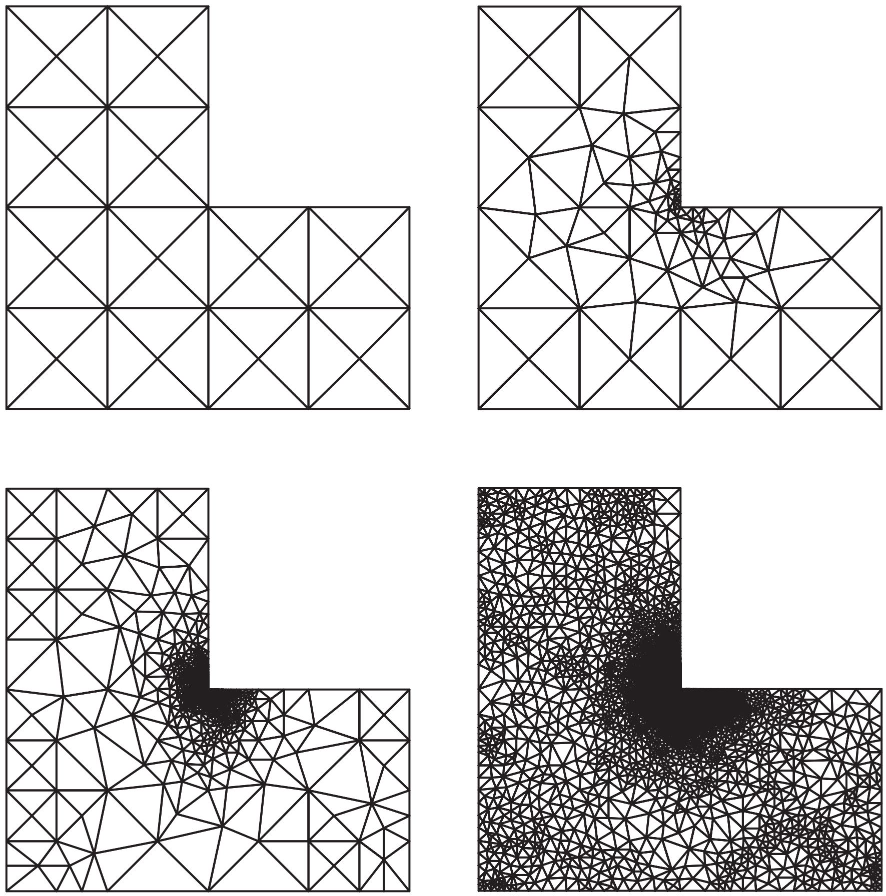

FIG. 3. Example 7.2, adapted meshes with 257, 777, 19569 and 311277 degrees of freedom.

of Θ in this case of a nonsmooth solution. Intermediate meshes obtained with the adaptive refinements are displayed in Fig. 3. Note that the method is able to recognize the region with high gradients.

Acknowledgements

The authors are deeply grateful to Prof. Gabriel N. Gatica (Universidad de Concepción) for his helpful advice and comments that helped to improve the paper.

Funding

This work was partially supported by CONICYT-Chile (Anillo project ACT 1118 (ANANUM) to S.C.); CONICYT-Chile (FONDECYT project 1140791, Anillo project ACT 1118 (ANANUM) to D.M.); CONICYT-Chile (FONDECYT project 11121347, Anillo project ACT 1118 (ANANUM) to R.O.).

REFERENCES

ARNOLD, D. N., DOUGLAS, J. & GUPTA, Ch. P. (1984) A family of higher order mixed finite element methods for plane elasticity. *Numer. Math.*, **45**, 1–22.

BERNARDI, C., CANUTO, C. & MADAY, Y. (1988) Generalized inf–sup conditions for Chebyshev spectral approximation of the Stokes problem. *SIAM J. Numer. Anal.*, **25**, 1237–1271.

BERNARDI, C., LAVAL, F., MÉTIVET, B. & PERNAUD-THOMAS, B. (1992) Finite element approximation of viscous flows with varying density. *SIAM J. Numer. Anal.*, **29**, 1203–1243.

BOFFI, D., BREZZI, F. & FORTIN, M. (2013) *Mixed Finite Element Methods and Applications*. Springer Series in Computational Mathematics, vol. 44. Heidelberg: Springer.

CAI, Z., TONG, C., VASSILEVSKI, P. S. & WANG, C. (2009) Mixed finite element methods for incompressible flow: stationary Stokes equations. *Numer. Methods Partial Differential Equations*, **26**, 957–978.

CAI, Z. & WANG, Y. (2010) Pseudostress–velocity formulation for incompressible Navier–Stokes equations. *Internat. J. Numer. Methods Fluids*, **63**, 341–356.

CAI, Z. & ZHANG, S. (2012) Mixed methods for stationary Navier–Stokes equations based on pseudostress–pressure–velocity formulation. *Math. Comp.*, **81**, 1903–1927.

CARSTENSEN, C. (1997) *A posteriori* error estimate for the mixed finite element method. *Math. Comp.*, **66**, 465–476.

CIARLET, P. G. (1978) *The Finite Element Method for Elliptic Problems*. Amsterdam, New York, Oxford: North-Holland.

CLÉMENT, P. (1975) Approximation by finite element functions using local regularisation. *RAIRO Modélisation Mathématique et Analyse Numérique*, **9**, 77–84.

DEMKOWICZ, L. (2006) *Babuška ⇔ Brezzi ??* ICES Reports 0608, UT Austin.

DI PIETRO, D. A. & ERN, A. (2012) *Mathematical Aspects of Discontinuous Galerkin Methods*. Berlin: Springer.

ERN, A. (1998) Vorticity–velocity formulation of the Stokes problem with variable density and viscosity. *Math. Models Methods Appl. Sci.*, **8**, 203–218.

ERN, A., DOUGLAS, C. C. & SMOOKE, M. D. (1995) Detailed chemistry modelling of laminar diffusion flames on parallel computers. *Internat. J. Supercomput. Appl. High Perf. Comput.*, **9**, 167–186.

ERN, A., GIOVANGIGLI, V. & SMOOKE, M. D. (1996) Numerical study of a three-dimensional chemical vapor deposition reactor with detailed chemistry. *J. Comput. Phys.*, **126**, 21–39.

FIGUEROA, L., GATICA, G. N. & HEUER, N. (2008a) *A priori* and *a posteriori* error analysis of an augmented mixed finite element method for incompressible fluid flows. *Comput. Methods Appl. Mech. Engrg.*, **198**, 280–291.

FIGUEROA, L., GATICA, G. N. & MÁRQUEZ, A. (2008b) Augmented mixed finite element methods for the stationary Stokes equations. *SIAM J. Sci. Comput.*, **31**, 1082–1119.

GATICA, G. N. (2006) Analysis of a new augmented mixed finite element method for linear elasticity allowing $\mathbb{RT}_0 - \mathbb{P}_1 - \mathbb{P}_0$ approximations. *ESAIM: Math. Model. Numer. Anal.*, **40**, 1–28.

GATICA, G. N. (2014) *A Simple Introduction to the Mixed Finite Element Method. Theory and Applications*. Springer Briefs in Mathematics. Cham, Heidelberg, New York, Dordrecht, London: Springer.

GATICA, G. N., GATICA, L. F. & MÁRQUEZ, A. (2011a) Augmented mixed finite element methods for a vorticity-based velocity–pressure–stress formulation of the Stokes problem in 2D. *Internat. J. Numer. Methods Fluids*, **67**, 450–477.

GATICA, G. N., GATICA, L. F. & MÁRQUEZ, A. (2014a) Analysis of a pseudostress-based mixed finite element method for the Brinkman model of porous media flow. *Numer. Math.*, **126**, 635-–677.

GATICA, G. N., GONZÁLEZ, M. & MEDDAHI, S. (2004) A low-order mixed finite element method for a class of quasi-Newtonian Stokes flows. I: *a priori* error analysis. *Comput. Methods Appl. Mech. Engrg.*, **193**, 881–892.

GATICA, G. N., MÁRQUEZ, A. & SÁNCHEZ, M. A. (2010) Analysis of a velocity–pressure–pseudostress formulation for incompressible flow. *Comput. Methods Appl. Mech. Engrg.*, **199**, 1064–1079.

GATICA, G. N., MÁRQUEZ, A. & SÁNCHEZ, M. A. (2011b) *A priori* and *a posteriori* error analyses of a velocity-pseudostress formulation for a class of quasi-Newtonian Stokes flows. *Comput. Methods Appl. Mech. Engrg.*, **200**, 1619–1636.

Gatica, G. N., Márquez, A. & Sánchez, M. A. (2012) Pseudostress-based mixed finite element methods for the Stokes problem in R^n with Dirichlet boundary conditions. I: *a priori* error analysis. *Commun. Comput. Phys.*, **12**, 109–134.

Gatica, G. N., Oyarzúa, R. & Sayas, F. J. (2011c) Analysis of fully-mixed finite element methods for the Stokes–Darcy coupled problem. *Math. Comp.*, **80**, 1911–1948.

Gatica, G. N., Oyarzúa, R. & Sayas, F. J. (2011d) A residual-based *a posteriori* error estimator for a fully-mixed formulation of the Stokes–Darcy coupled problem. *Comput. Methods Appl. Mech. Engrg.*, **200**, 1877–1891.

Gatica, G. N., Ruiz-Baier, R. & Tierra, G. (2014b) A mixed finite element method for the Darcy equations with pressure dependent porosity. Preprint 2014-06, Centro de Investigación en Ingenieria Matemática (CI2MA), UDEC.

Girault, V. & Raviart, P.-A. (1986) *Finite Element Methods for Navier–Stokes Equations. Theory and Algorithms*. Springer Series in Computational Mathematics vol. 5. Berlin: Springer.

Hiptmair, R. (2002) Finite elements in computational electromagnetism. *Acta Numer.*, **11**, 237–339.

Nicolaides, R. A. (1982) Existence, uniqueness and approximation for generalized saddle point problems. *SIAM J. Numer. Anal.*, **19**, 349–357.

Verfürth, R. (1994) *A posteriori* error estimation and adaptive mesh-refinement techniques. *J. Comput. Appl. Math.*, **50**, 67–83.

Verfürth, R. (1996) *A Review of a Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*. Chichester: Wiley Teubner.