Analysis of a mixed finite element method for the Poisson problem with data in L^p, $\dfrac{2n}{n+2} < p < 2$, $n = 2, 3$ *

Jessica Camaño[†] Cristian Muñoz[‡] Ricardo Oyarzúa[§]

Abstract

In this paper we analyze the numerical approximation of the Poisson problem in mixed form, considering a right-hand side $f \in L^p(\Omega)$, with $p \in (\frac{2n}{n+2}, 2)$, where $n = 2, 3$ is the dimension of Ω. The analysis of the corresponding continuous and discrete problems are carried out by means of the classical Babuška-Brezzi theory, where the associated Galerkin scheme is defined by Raviart–Thomas elements of lowest order combined with piecewise constants. In particular, we prove well-posedness and convergence of the discrete scheme under a quasi-uniformity condition of the mesh. Next, we apply the theory developed for the Poisson problem to a convection-difusion problem, providing well-posedness of the continuous and discrete problems and optimal convergence. Finally, we corroborate the theoretical results with suitable numerical results in two and three dimensions.

Key words: Mixed finite element method, Raviart–Thomas, Lp data, convection-diffusion

Mathematics Subject Classifications (2010): 65N15, 65N12, 65N30, 74S05

1 Introduction

This paper is concerned with the numerical approximation of the Poisson problem:

$$-\Delta u \,=\, f \quad \text{in} \quad \Omega \subseteq \mathbb{R}^n,\, n \in \{2, 3\}, \qquad u \,=\, u_D \quad \text{on} \quad \Gamma := \partial\Omega, \tag{1.1}$$

where, given $p \in (\frac{2n}{n+2}, 2)$, the source data f is a function of $L^p(\Omega)$, $u_D \in \mathrm{H}^{1/2}(\Gamma)$, and Ω is a polyhedral domain. In particular, we are interested in studying the mixed finite element approximation of (1.1).

*This research was partially supported by CONICYT-Chile through project Fondecyt 11121347 and 11140691; project Anillo ACT1118 (ANANUM); project Insercion de Capital Humano Avanzado en la Academia 79130048; and by Universidad del Bío-Bío through DIUBB project GI 151408/VC.

[†]Departamento de Matemática y Física Aplicadas, Universidad Católica de la Santísima Concepción, Casilla 297, Concepción, Chile, and CI^{2}MA, Universidad de Concepción, Casilla 160-C, Concepción, Chile, email: `jecamano@ucsc.cl`

[‡]Departamento de Matemática, Universidad del Bío-Bío, Casilla 5-C, Concepción, Chile, email: `cmunozna@ubiobio.cl`

[§]Departamento de Matemática, Universidad del Bío-Bío, Casilla 5-C, Concepción, Chile and CI^{2}MA, Universidad de Concepción, Casilla 160-C, Concepción, Chile, email: `royarzua@ubiobio.cl`

Our motivation in studying the Galerkin approximation of (1.1) in mixed form arises from the necessity of approximating the flux of certain concentration θ satisfying the convection-diffusion equation:

$$-\Delta\theta + \mathbf{v}\cdot\nabla\theta = g \quad \text{in} \quad \Omega, \qquad \theta = \theta_D \quad \text{on} \quad \Gamma, \tag{1.2}$$

where $\mathbf{v}$ is a given function in $[\mathrm{H}^1(\Omega)]^n$ representing the velocity of a viscous fluid where the concentration is moving, and $g \in L^2(\Omega)$ is an external force. Certainly, the best option for our purposes is to use a mixed method to approximate the solution of (1.2). To that end, we introduce the further unknown $\sigma := \nabla\theta$ in Ω, and proceed as usual to arrive at the mixed variational formulation of (1.2): Find σ and θ in suitable spaces, such that

$$\begin{aligned} \int_\Omega \sigma\cdot\tau + \int_\Omega \theta \,\mathrm{div}\,\tau &= \langle \tau\cdot\nu, \theta_D\rangle_\Gamma\,, \\ \int_\Omega \psi\,\mathrm{div}\,\sigma - \int_\Omega \psi(\mathbf{v}\cdot\sigma) &= -\int_\Omega g\psi, \end{aligned} \tag{1.3}$$

for all τ and ψ. Now, in order to define the spaces for the corresponding unknowns and test functions, we notice that the first term of the first equation of (1.3) is well defined if σ and τ are in $[L^2(\Omega)]^n$. However, if $\sigma \in [L^2(\Omega)]^n$, the second term of the second equation of (1.3) forces the test function ψ to live in a space smaller than $L^2(\Omega)$, and as a consequence, the term $\mathrm{div}\,\sigma$ to live in a space larger than $L^2(\Omega)$. Indeed, by applying Cauchy-Schwarz and Hölder inequalities and then the continuous injection of $\mathrm{H}^1(\Omega)$ into $L^4(\Omega)$ (see e.g. [18, Theorem 1.3.4]), we obtain that there exists $c(\Omega)$, such that

$$\begin{aligned} \left|\int_\Omega \psi(\mathbf{v}\cdot\sigma)\right| &\le \|\psi\mathbf{v}\|_{[L^2(\Omega)]^n}\|\sigma\|_{[L^2(\Omega)]^n} \le \|\psi\|_{L^4(\Omega)}\|\mathbf{v}\|_{[L^4(\Omega)]^n}\|\sigma\|_{[L^2(\Omega)]^n} \\ &\le c(\Omega)\|\psi\|_{L^4(\Omega)}\|\mathbf{v}\|_{[H^1(\Omega)]^n}\|\sigma\|_{[L^2(\Omega)]^n}. \end{aligned} \tag{1.4}$$

According to the above, we obtain that the mixed problem (1.3) can be analyzed if the unknown θ and the test function ψ live both in $L^4(\Omega)$, whereas σ and τ live in $\mathrm{H}(\mathrm{div}_{4/3}, \Omega)$, where

$$\mathrm{H}(\mathrm{div}_{4/3}, \Omega) := \{\tau \in L^2(\Omega) \,:\, \mathrm{div}\,\tau \in L^{4/3}(\Omega)\}.$$

Observe that in (1.4) we could have bounded $\mathbf{v}$ in the L^∞-norm and keep ψ and $\mathrm{div}\,\sigma$ in $L^2(\Omega)$. However, since (1.2) is usually coupled with an equation modelling the velocity $\mathbf{v}$ (see for instance [2, 3, 5, 8, 9, 13, 17] and the references therein), the estimate of $\mathbf{v}$ in the H^1-norm is required to analyze the full system.

Now, concerning the numerical approximation of the model problem (1.1) in mixed form, it is quite surprising to realize that almost no contributions are available in the literature. Among the few works related to this problem, we could mention the article [11] where the author focuses on deriving the a priori error estimate in L^p, with $1 \le p \le \infty$ for the mixed finite element solution of the Poisson problem in $\mathbb{R}^2$, assuming that the solutions are in the standard spaces $\mathrm{H}(\mathrm{div})$ and L^2.The latter is certainly true if the data is in L^s, with $s \ge 2$. Similarly, in [1] the authors focus on proving error estimates in L^p, with $1 \le p \le \infty$, for the 3D Raviart-Thomas approximation of mixed problems.

According to the above discussion, in this paper we analyze the solvability and numerical approximation of the mixed variational formulation of problem (1.1) with data in L^p, with $\frac{2n}{n+2} < p < 2$. As usual for mixed problems, we employ the classical Babuška-Brezzi theory to study the well-posedness of the continuous problem. Since the Lebesgue and Sobolev spaces involved are not standard, the main drawback appears when proving the corresponding inf-sup condition, which is overcome by using suitable auxiliary problems. Similarly, we obtain that the associated Galerkin scheme, defined by Raviart-Thomas elements of lowest order and piecewise constants elements defined on a quiasiuniform mesh, is well posed and convergent. It is pertinent to note that the quasi-uniformity assumption here is required to apply an inverse estimate in the proof of the discrete inf-sup condition. The eventual elimination of this assumption remains an open problem.

The rest of the article is organized as follows. In Section 2 we rewrite the boundary value problem (1.1) as a first-order set of equations, provide the associated mixed variational formulation, and show that it is well posed. Then, the mixed finite element method is defined and analyzed in Section 3. Next, in Section 4 we apply the results derived in the previous sections to the convection-diffusion problem (1.2). Finally, several numerical results illustrating the performance of the mixed method are presented in Section 5.

We end this section by fixing some notations. Throughout the rest of the paper, we utilize the standard terminology for Lebesgue and Sobolev spaces, norms and seminorms, In fact, let $\mathcal{O}$ be a domain in $\mathbb{R}^n$, $n = 2, 3$, with Lipschitz boundary $\partial\mathcal{O}$. For $r \geq 0$ 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 $\mathrm{H}^r(\mathcal{O})$ 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. For $r \geq 0$, we write $| \cdot |_{r,\mathcal{O}}$ for the H^r-seminorm. The space $\mathrm{H}^1_0(\mathcal{O})$ is the space of functions in $\mathrm{H}^1(\mathcal{O})$ with vanishing trace on Γ. Also, the Hilbert space

$$\mathrm{H}(\mathrm{div}\,, \mathcal{O}) := \left\{ \tau \in [L^2(\mathcal{O})]^n : \quad \mathrm{div}\ \tau \in L^2(\mathcal{O}) \right\},$$

is standard in the realm of mixed problems (see [4] or [15] for instance).

In what follows, we employ $\mathbf{0}$ to denote a generic null vector and use C and c, with or without subscripts, bars, tildes or hats, to denote generic positive constants independent of the discretization parameters, which may take different values at different places.

2 The mixed variational formulation

2.1 Preliminaries

Since we are interested in using mixed finite element methods to solve (1.1), we first define the additional unknown $\sigma := \nabla u$ in Ω and rewrite (1.1) as the following first–order set of equations:

$$\sigma = \nabla u \quad \text{in} \quad \Omega, \qquad -\,\mathrm{div}\,\sigma = f \quad \text{in} \quad \Omega, \qquad u = u_D \quad \text{on} \quad \Gamma. \tag{2.1}$$

Hence, to derive the mixed variational formulation, as usual, the first equation of (2.1) is multiplied by a test function τ and integrated by parts, considering in the process the Dirichlet boundary condition $u = u_D$ on Γ. In turn, since $f \in L^p(\Omega)$, the second equation of (2.1) is imposed weakly in the corresponding space. As a result, we arrive at the variational problem: Find $(\sigma, u) \in \mathrm{H}(\mathrm{div}_p\,, \Omega) \times L^q(\Omega)$, such that:

$$\begin{aligned} \int_\Omega \sigma \cdot \tau + \int_\Omega u\,\mathrm{div}\,\tau &= \langle \tau \cdot \nu, u_D \rangle_\Gamma \quad \forall\, \tau \in \mathrm{H}(\mathrm{div}_p\,, \Omega), \\ \int_\Omega v\,\mathrm{div}\,\sigma &= -\int_\Omega f v \quad \forall\, v \in L^q(\Omega), \end{aligned} \tag{2.2}$$

where $\mathrm{H}(\mathrm{div}_p, \Omega)$ is the Sobolev space defined as

$$\mathrm{H}(\mathrm{div}_p, \Omega) := \{\tau \in [L^2(\Omega)]^n \; : \; \mathrm{div}\, \tau \in L^p(\Omega)\}, \tag{2.3}$$

endowed with the norm

$$\|\tau\|_{\mathrm{H}(\mathrm{div}_p, \Omega)} := \Big(\|\tau\|^2_{L^2(\Omega)} + \|\mathrm{div}\, \tau\|^2_{L^p(\Omega)} \Big)^{1/2}, \tag{2.4}$$

$q \in \mathbb{R}$ is the conjugate exponent of p, satisfying $\dfrac{1}{p} + \dfrac{1}{q} = 1$, ν is the unit outward normal to Ω, and $\langle \cdot, \cdot \rangle_\Gamma$ is the duality pairing of $\mathrm{H}^{-1/2}(\Gamma)$, and $\mathrm{H}^{1/2}(\Gamma)$ with respect to the $L^2(\Gamma)$–inner product.

At this point, we recall that given $\tau \in \mathrm{H}(\mathrm{div}, \Omega)$, the normal trace $\tau \cdot \nu$ is defined as the functional in $\mathrm{H}^{-1/2}(\Gamma)$ given by (see e.g. [14, Section 1.3.4])

$$\langle \tau \cdot \nu, \xi \rangle_\Gamma = \int_\Omega \tau \cdot \nabla \tilde{\gamma}_0^{-1}(\xi) + \int_\Omega \tilde{\gamma}_0^{-1}(\xi)\, \mathrm{div}\, \tau \qquad \forall\, \xi \in \mathrm{H}^{1/2}(\Gamma), \tag{2.5}$$

where $\tilde{\gamma}_0^{-1} : \mathrm{H}^{1/2}(\Gamma) \to [\mathrm{H}^1_0(\Omega)]^\perp$ is the right inverse of the well known trace operator $\gamma_0 : \mathrm{H}^1(\Omega) \to \mathrm{H}^{1/2}(\Gamma)$. Then, since $\tilde{\gamma}_0^{-1}(\xi) \in \mathrm{H}^1(\Omega)$, owing to the classical Sobolev embedding $\mathrm{H}^1(\Omega) \subset L^q(\Omega)$, the last term in (2.5) is still well defined if $\mathrm{div}\, \tau \in L^p(\Omega)$. This implies that $\tau \cdot \nu \in \mathrm{H}^{-1/2}(\Gamma)$ for all $\tau \in \mathrm{H}(\mathrm{div}_p, \Omega)$, and as a result, the right-hand side of the first equation of (2.2) is well defined. Moreover, it readily follows that there exists $c(\Omega) > 0$, depending on $|\Omega|$, such that

$$|\, \langle \tau \cdot \nu, \xi \rangle \,| \le c(\Omega) \|\tau\|_{\mathrm{H}(\mathrm{div}_p, \Omega)} \|\xi\|_{1/2,\Gamma}, \qquad \forall\, \tau \in \mathrm{H}(\mathrm{div}_p, \Omega), \; \forall\, \xi \in \mathrm{H}^{1/2}(\Gamma). \tag{2.6}$$

Now, we introduce some notations and previous results that will serve for the forthcoming analysis. We begin by defining the sign function sgn, given by

$$\mathrm{sgn}(v) = \begin{cases} 1 & \text{if} \quad v \ge 0, \\ -1 & \text{if} \quad v < 0, \end{cases}$$

for any scalar function v. It is quite clear that for a given v, there holds

$$v\, \mathrm{sgn}(v) = |v|.$$

In addition, in the sequel we will make use of the well known Hölder, Poincaré and Sobolev inequalities, given respectively by

$$\int_\Omega |fg| \le \|f\|_{L^p(\Omega)} \|g\|_{L^q(\Omega)}, \quad \forall\, f \in L^p(\Omega), \; \forall\, g \in L^q(\Omega), \quad \text{with} \quad \frac{1}{p} + \frac{1}{q} = 1, \tag{2.7}$$

$$\|w\|_{1,\Omega} \le C_P |w|_{1,\Omega} \quad \forall w \in \mathrm{H}^1_0(\Omega) \tag{2.8}$$

and

$$\|w\|_{L^r(\Omega)} \le C_{Sob} \|w\|_{1,\Omega} \quad \forall\, w \in \mathrm{H}^1(\Omega), \quad r \ge 1, \tag{2.9}$$

with $C_P > 0$ and $C_{Sob} > 0$ depending only on $|\Omega|$.

2.2 Analysis of the continuous problem

In what follows, we prove existence and uniqueness of solution of problem (2.2). To that end, and for the sake of simplicity, we now write our problem in the classical variational setting and state the main properties of the bilinear forms involved. We start by defining the spaces

$$\mathrm{H} := \mathrm{H}(\mathrm{div}_p\,,\Omega), \quad \mathrm{Q} := L^q(\Omega),$$

and the product norm

$$\|(\tau, v)\|_{\mathrm{H}\times\mathrm{Q}} := \{\|\tau\|_{\mathrm{H}}^2 + \|v\|_{\mathrm{Q}}^2\}^{1/2},$$

where $\|\cdot\|_{\mathrm{H}} := \|\cdot\|_{\mathrm{H}(\mathrm{div}_p\,,\Omega)}$ and $\|\cdot\|_{\mathrm{Q}} := \|\cdot\|_{L^q(\Omega)}$. Then, defining the bilinear forms $a : \mathrm{H}\times\mathrm{H} \to \mathbb{R}$, $b : \mathrm{H} \times \mathrm{Q} \to \mathbb{R}$ and the functionals $F : \mathrm{H} \to \mathbb{R}$ and $G : \mathrm{Q} \to \mathbb{R}$, given respectively by

$$a(\sigma, \tau) := \int_\Omega \sigma \cdot \tau, \qquad b(\tau, v) := \int_\Omega v \,\mathrm{div}\, \tau, \qquad F(\tau) := \langle \tau \cdot \nu, u_D\rangle_\Gamma\,, \qquad G(v) := -\int_\Omega f v, \tag{2.10}$$

the variational formulation (2.2) reads: Find $(\sigma, u) \in \mathrm{H} \times \mathrm{Q}$, such that

$$\begin{aligned} a(\sigma, \tau) + b(\tau, u) &= F(\tau) \quad \forall\, \tau \in \mathrm{H}, \\ b(\sigma, v) &= G(v) \quad \forall\, v \in \mathrm{Q}. \end{aligned} \tag{2.11}$$

Notice that, owing to the Hölder inequality (2.7), the bilinear forms a and b are bounded:

$$\begin{aligned} |a(\sigma, \tau)| &\le \|\sigma\|_{\mathrm{H}}\|\tau\|_{\mathrm{H}} \quad \forall\, \sigma \in \mathrm{H}, \forall\, \tau \in \mathrm{H}, \\ |b(\tau, v)| &\le \|v\|_{\mathrm{Q}}\|\tau\|_{\mathrm{H}} \quad \forall\, \tau \in \mathrm{H}, \forall\, v \in \mathrm{Q}. \end{aligned} \tag{2.12}$$

In turn, from (2.6) and the Hölder inequality (2.7), it readily follows that F and G are bounded:

$$\begin{aligned} |F(\tau)| &= |\,\langle \tau \cdot \nu, u_D\rangle_\Gamma\,| \le c(\Omega)\|u_D\|_{1/2,\Gamma}\|\tau\|_{\mathrm{H}} \quad \forall\, \tau \in \mathrm{H}, \\ |G(v)| &\le \|f\|_{L^p(\Omega)}\|v\|_{\mathrm{Q}} \quad \forall\, v \in \mathrm{Q}. \end{aligned} \tag{2.13}$$

Throughout the rest of this section we follow the analysis suggested by the classical Babuška-Brezzi theory to conclude that (2.11) is well posed. This requires the inf-sup condition of b and the ellipticity of a on the kernel of b. We start with the following lemma establishing that b satisfies the required inf-sup condition.

Lemma 2.1 *There exists $\beta > 0$, such that*

$$\sup_{\tau\in\mathrm{H}\setminus\mathbf{0}} \frac{b(\tau, v)}{\|\tau\|_{\mathrm{H}}} \ge \beta\|v\|_{\mathrm{Q}} \quad \forall\, v \in \mathrm{Q}.$$

Proof. Given $v \in \mathrm{Q}$, we let $\tilde{\tau} = -\nabla z$, with $z \in \mathrm{H}_0^1(\Omega)$ being the unique solution of the variational problem

$$\int_\Omega \nabla z \cdot \nabla w = \int_\Omega \mathrm{sgn}\,(v)|v|^{q-1} w \quad \forall\, w \in \mathrm{H}_0^1(\Omega). \tag{2.14}$$

Notice that

$$\int_\Omega |\mathrm{sgn}\,(v)|v|^{q-1}|^p = \int_\Omega |v|^{p(q-1)} = \int_\Omega |v|^q < +\infty, \tag{2.15}$$

which implies that $\mathrm{sgn}\,(v)|v|^{q-1} \in L^p(\Omega)$. Then, since $p \in (\frac{2n}{n+2}, 2)$, it is well-known that problem (2.14) is well posed. In turn, from (2.14) it readily follows that $\mathrm{div}\,\tilde{\tau} = \mathrm{sgn}\,(v)|v|^{q-1}$. As a consequence, we obtain that $\tilde{\tau} \in \mathrm{H}$ and

$$\|\mathrm{div}\,\tilde{\tau}\|_{L^p(\Omega)} = \||v|^{q-1}\|_{L^p(\Omega)}. \tag{2.16}$$

On the other hand, utilizing inequalities (2.7), (2.8) and (2.9), from (2.14) with $w = z$, we obtain

$$\begin{aligned} \|\tilde{\tau}\|_{0,\Omega}^2 &\leq \||v|^{q-1}\|_{L^p(\Omega)}\|z\|_{\mathrm{Q}} \leq C_{Sob}\||v|^{q-1}\|_{L^p(\Omega)}\|z\|_{1,\Omega} \\ &\leq C_P C_{Sob}\||v|^{q-1}\|_{L^p(\Omega)}|z|_{1,\Omega} = C_P C_{Sob}\||v|^{q-1}\|_{L^p(\Omega)}\|\tilde{\tau}\|_{0,\Omega}, \end{aligned}$$

from which,

$$\|\tilde{\tau}\|_{0,\Omega} \leq C_P C_{Sob}\||v|^{q-1}\|_{L^p(\Omega)}. \tag{2.17}$$

In this way, from (2.16) and (2.17), it follows that

$$\|\tilde{\tau}\|_{\mathrm{H}} \leq (1 + C_P^2 C_{Sob}^2)^{1/2}\||v|^{q-1}\|_{L^p(\Omega)},$$

which together to the fact that

$$\||v|^{q-1}\|_{L^p(\Omega)} = \left(\int_\Omega (|v|^{q-1})^p\right)^{\frac{1}{p}} = \left(\int_\Omega |v|^q\right)^{\frac{q-1}{q}} = \|v\|_Q^{q-1},$$

implies

$$\|\tilde{\tau}\|_{\mathrm{H}} \leq (1 + C_P^2 C_{Sob}^2)^{1/2}\|v\|_{\mathrm{Q}}^{q-1}. \tag{2.18}$$

Therefore, recalling that $v\,\mathrm{sgn}\,(v) = |v|$, from the definition of $\tilde{\tau}$ and (2.18), we obtain

$$\begin{aligned} \sup_{\tau \in \mathrm{H}\setminus\mathbf{0}} \frac{b(\tau, v)}{\|\tau\|_{\mathrm{H}}} &\geq \frac{b(\tilde{\tau}, v)}{\|\tilde{\tau}\|_{\mathrm{H}}} = \frac{\displaystyle\int_\Omega v\,\mathrm{div}\,\tilde{\tau}}{\|\tilde{\tau}\|_{\mathrm{H}}} \geq (1 + C_P^2 C_{Sob}^2)^{-1/2}\frac{\displaystyle\int_\Omega |v|\,|v|^{q-1}}{\|v\|_{\mathrm{Q}}^{q-1}} \\ &= (1 + C_P^2 C_{Sob}^2)^{-1/2}\frac{\|v\|_{\mathrm{Q}}^q}{\|v\|_{\mathrm{Q}}^{q-1}} = (1 + C_P^2 C_{Sob}^2)^{-1/2}\|v\|_{\mathrm{Q}}, \end{aligned} \tag{2.19}$$

which concludes the proof with $\beta = (1 + C_P^2 C_{Sob}^2)^{-1/2} > 0$. □

We now let V be the kernel of b, that is

$$V := \{\tau \in \mathrm{H} : b(\tau, v) = 0,\ \forall\, v \in \mathrm{Q}\} = \left\{\tau \in \mathrm{H} : \int_\Omega v\,\mathrm{div}\,\tau = 0,\ \forall\, v \in \mathrm{Q}\right\}. \tag{2.20}$$

Observe that if $\tau \in V$, then taking $v = \mathrm{sgn}\,(\mathrm{div}\,\tau)|\mathrm{div}\,\tau|^{p-1}$, which is clearly an element in Q since

$$\int_\Omega |v|^q = \int_\Omega \left|\mathrm{sgn}\,(\mathrm{div}\,\tau)|\mathrm{div}\,\tau|^{p-1}\right|^q = \int_\Omega |\mathrm{div}\,\tau|^{(p-1)q} = \int_\Omega |\mathrm{div}\,\tau|^p < +\infty,$$

it follows that

$$0 = \int_\Omega v\,\mathrm{div}\,\tau = \int_\Omega \mathrm{sgn}\,(\mathrm{div}\,\tau)|\mathrm{div}\,\tau|^{p-1}\mathrm{div}\,\tau = \int_\Omega |\mathrm{div}\,\tau|^p = \|\mathrm{div}\,\tau\|_{L^p(\Omega)}^p,$$

and then $\mathrm{div}\,\tau \equiv 0$ in $L^p(\Omega)$. In this way,

$$V := \{\tau \in\ \mathrm{H} :\ \mathrm{div}\,\tau \equiv 0 \quad \text{in } \Omega\}.$$

The following lemma establishes the ellipticity of a on V.

Lemma 2.2 *There exists $\alpha > 0$, such that*

$$a(\tau,\tau) \geq \alpha \|\tau\|_{\mathrm{H}} \quad \forall\, \tau \in V \tag{2.21}$$

Proof. Given $\tau \in V$, from the definition of V, it readily follows that

$$a(\tau,\tau) = \|\tau\|_{0,\Omega}^2 = \|\tau\|_{\mathrm{H}}^2,$$

which implies (2.21) with $\alpha = 1$. $\square$

The well-posedness of the continuous formulation (2.11) is provided now.

Theorem 2.3 *Let $f \in L^p(\Omega)$, with $p \in \left(\frac{2n}{n+2}, 2\right)$. Then there exists a unique $(\sigma, u) \in \mathrm{H} \times \mathrm{Q}$ solution to* (2.11)*. In addition, there exists $C > 0$, independent of the solution, such that*

$$\|(\sigma, u)\|_{\mathrm{H}\times\mathrm{Q}} \leq C(\|f\|_{L^p(\Omega)} + \|u_D\|_{1/2,\Gamma}). \tag{2.22}$$

Proof. Thanks to Lemmata 2.1 and 2.2 and the fact that the right-hand sides F and G define linear functionals on H and Q, respectively, the proof follows from a straightforward application of the Babuška–Brezzi theory (see e.g. [4, Chapter II] or [14, Chapter 2]). $\square$

3 The mixed finite element scheme

3.1 Preliminaries

Let $\{\mathcal{T}_h\}_{h>0}$ be a regular family of triangulations of Ω by triangles T in $\mathbb{R}^2$ or tetrahedra in $\mathbb{R}^3$ of diameter h_T such that $h := \max\{h_T : T \in \mathcal{T}_h\}$. Then, for each $T \in \mathcal{T}_h$, we let $\mathrm{RT}_0(T)$ be the local Raviart-Thomas of lowest order, i.e.

$$\mathrm{RT}_0(T) := [\mathbb{P}_0(T)]^n \oplus \mathbb{P}_0(T)\mathbf{x},$$

where $\mathbf{x} := (x_1, \ldots, x_n)^t$ is a generic vector of $\mathbb{R}^n$ and $\mathbb{P}_0(T)$ is the space of constant functions on T. In general, given a non–negative integer k and a subset of $\mathbb{R}^n$ S, we let $\mathbb{P}_k(S)$ be the space of polynomials defined on S of degree $\leq k$. Hence, defining the following finite element subspaces to approximate the unknowns $\sigma \in \mathrm{H}$ and $u \in \mathrm{Q}$:

$$\begin{aligned} \mathrm{H}_h &:= \{\tau_h \in \mathrm{H} \,:\, \tau_h|_T \in \mathrm{RT}_0(T), \quad \forall\, T \in \mathcal{T}_h\} \subseteq \mathrm{H}, \\ \mathrm{Q}_h &:= \{v_h \in \mathrm{Q} \,:\, v_h|_T \in \mathbb{P}_0(T), \quad \forall\, T \in \mathcal{T}_h\} \subseteq \mathrm{Q}, \end{aligned} \tag{3.1}$$

the conforming Galerkin scheme for (2.11) reads: Find $(\sigma_h, u_h) \in \mathrm{H}_h \times \mathrm{Q}_h$, such that

$$\begin{aligned} a(\sigma_h, \tau_h) + b(\tau_h, u_h) &= F(\tau_h) \quad \forall\, \tau_h \in \mathrm{H}_h, \\ b(\sigma_h, v_h) &= G(v_h) \quad \forall\, v_h \in \mathrm{Q}_h, \end{aligned} \tag{3.2}$$

where a, b, F and G are the bilinear forms and functionals defined in (2.2).

Next, in Section 3.2 we follow closely the analysis suggested by [14, Section 4.2] to prove that problem (3.2) is well posed by means of the discrete Babuška-Brezzi theory. To that end, we first need to establish some previous results and definitions.

As we will see later in Lemma 3.1, in order to prove the discrete inf-sup condition for b we will require the following inverse inequality

$$\|z_h\|_{0,\Omega} \leq C h^{\frac{n}{2}-\frac{n}{p}} \|z_h\|_{L^p(\Omega)} \qquad \forall\, z_h \in \mathrm{Q}_h, \tag{3.3}$$

which is certainly true for quasi-uniform meshes (see e.g. [12, Corollary 1.141]) . Therefore, from now on we assume that for each $h > 0$, $\mathcal{T}_h$ is quasi-uniform, which means that there exists $\tilde{c} > 0$, independent of h, such that

$$\min_{T \in \mathcal{T}_h} h_T \geq \tilde{c} h \qquad \forall\, h > 0. \tag{3.4}$$

We now introduce the approximation properties of the finite element subspaces introduced above. To that end we first define the space

$$Z := \{\tau \in \mathrm{H}(\mathrm{div}\,, \Omega) \;:\; \tau|_T \in [\mathrm{H}^1(T)]^n, \quad \forall T \in \mathcal{T}_h\}.$$

Then, we let

$$\Pi_h : \mathrm{H}(\mathrm{div}\,, \Omega) \cap Z \to \mathrm{H}_h, \tag{3.5}$$

be the usual Raviart-Thomas interpolator operator, which given $\tau \in \mathrm{H}(\mathrm{div}\,, \Omega) \cap Z$, is characterized by the identity

$$\int_e (\Pi_h(\tau) \cdot \nu)\, \xi = \int_e (\tau \cdot \nu)\, \xi \qquad \forall\, \xi \in \mathbb{P}_0(e),\ \forall \text{ edge } e \text{ of } \mathcal{T}_h. \tag{3.6}$$

In addition, the corresponding commuting diagram property yields

$$\mathrm{div}\,(\Pi_h(\tau)) = \mathcal{P}_h(\mathrm{div}\,\tau) \quad \forall\, \tau \in \mathrm{H}(\mathrm{div}\,, \Omega) \cap Z, \tag{3.7}$$

where $\mathcal{P}_h : L^2(\Omega) \to \mathrm{Q}_h$ is the corresponding orthogonal projection, which satisfies the following error estimate (see [12, Section 1.6.3]): For each $t \in [0,1]$ and for each $w \in \mathrm{H}^t(\Omega)$, there holds

$$\|w - \mathcal{P}_h(w)\|_{L^r(\Omega)} \leq C h^t \|w\|_{W^{1,r}(\Omega)}, \qquad 1 \leq r \leq \infty. \tag{3.8}$$

In turn, given $r > \frac{2n}{n+2}$, it can be proved that there exists $C > 0$, independent of h, such that for each $\tau \in [W^{1,r}(T)]^n$, there holds

$$\|\tau - \Pi_h(\tau)\|_{[L^r(T)]^n} \leq C \frac{h_T^2}{\rho_T} |\tau|_{[W^{1,r}(T)]^n}, \tag{3.9}$$

and for each $\tau \in [W^{1,r}(T)]^n$, with $\mathrm{div}\,\tau \in W^{1,r}(T)$,

$$\|\mathrm{div}\,\tau - \mathrm{div}\,(\Pi_h(\tau))\|_{L^r(T)} \leq C h_T |\mathrm{div}\,\tau|_{W^{1,r}(T)}, \tag{3.10}$$

where ρ_T is the diameter of the largest sphere contained in T (see e.g. [12, Theorem 1.114]). Furthermore, owing to the estimates (3.9) and (3.10), and the fact that the mesh is assumed to be regular, it is not difficult see that the following global estimate holds

$$\|\tau - \Pi_h(\tau)\|_{[L^r(\Omega)]^n} + \|\mathrm{div}\,\tau - \mathrm{div}\,(\Pi_h(\tau))\|_{L^r(\Omega)} \leq c h \Big\{ |\tau|_{[W^{1,r}(\Omega)]^n} + |\mathrm{div}\,\tau|_{W^{1,r}(\Omega)} \Big\}, \tag{3.11}$$

for all $\tau \in [W^{1,r}(\Omega)]^n$, with $\mathrm{div}\,\tau \in W^{1,r}(\Omega)$ and for all $r > \frac{2n}{n+2}$(see e.g. [12, Corollary 1.115]).

3.2 Analysis of the discrete problem

In this section we apply the discrete Babuška-Brezzi theory to prove the well-posedness of the Galerkin scheme (3.2). We start by establishing the discrete inf-sup condition for b.

Lemma 3.1 *Let $p \in (\frac{2n}{n+2}, 2)$, with $n = 2, 3$. Assume that $\mathcal{T}_h$ is a quasi-uniform triangulation of Ω. Then, there exists $\beta^* > 0$, independent of h, such that*

$$\sup_{\substack{\tau_h \in \mathrm{H}_h \\ \tau_h \neq \mathbf{0}}} \frac{b(\tau_h, v_h)}{\|\tau_h\|_{\mathrm{H}}} \geq \beta^* \|v_h\|_{\mathrm{Q}} \quad \forall\, v_h \in \mathrm{Q}_h.$$

Proof. Given $v_h \in \mathrm{Q}_h$, we set

$$\tilde{v}_h = \begin{cases} \operatorname{sgn}(v_h)|v_h|^{q-1} & \text{in} \quad \Omega, \\ 0 & \text{in} \quad B \backslash \overline{\Omega}, \end{cases}$$

where $B \subseteq \mathbb{R}^n$ is an open and bounded convex set containing $\overline{\Omega}$. Notice that $\operatorname{sgn}(v_h)|v_h|^{q-1} \in \mathrm{Q}_h$ since v_h is a piecewise constant function.

Next, let $\hat{\tau} = -\nabla\varphi|_{\Omega}$, with $\varphi \in \mathrm{H}_0^1(B)$ being the unique solution of the variational problem

$$\int_B \nabla\varphi \cdot \nabla w \;=\; \int_B \tilde{v}_h w = \int_\Omega \operatorname{sgn}(v_h)|v_h|^{q-1} w \qquad \forall\, w \in \mathrm{H}_0^1(B). \tag{3.12}$$

Similarly to the proof of Lemma 2.1, we easily obtain

$$\|\hat{\tau}\|_{0,\Omega} \leq C_1 \||v_h|^{q-1}\|_{L^p(\Omega)} \quad \text{and} \quad \operatorname{div}\hat{\tau} = \operatorname{sgn}(v_h)|v_h|^{q-1}. \tag{3.13}$$

In addition, since $\tilde{v}_h \in L^2(B)$ and B is convex, it is well known that $\varphi \in \mathrm{H}^2(B)$, and satisfies the estimate

$$\|\varphi\|_{\mathrm{H}^2(B)} \leq \tilde{C}\|\tilde{v}_h\|_{0,B} \;=\; \tilde{C}\||v_h|^{q-1}\|_{0,\Omega}.$$

It readily follows that $\hat{\tau} \in \mathrm{H}^1(\Omega)$. Moreover, combining the latter estimate with the inverse inequality (3.3), we obtain

$$|\hat{\tau}|_{1,\Omega} \leq \|\varphi\|_{\mathrm{H}^2(B)} \leq \tilde{C}\||v_h|^{q-1}\|_{0,\Omega} \leq C_2 h^{\frac{n}{2}-\frac{n}{p}} \||v_h|^{q-1}\|_{L^p(\Omega)}. \tag{3.14}$$

Now, we let $\hat{\tau}_h = \Pi_h(\hat{\tau})$. It is not difficult to see that the commutative diagram (3.7), the identity on the right-hand side of (3.13), and the fact that $\operatorname{sgn}(v_h)|v_h|^{q-1} \in Q_h$ imply

$$\operatorname{div}\hat{\tau}_h = \mathcal{P}_h(\operatorname{div}\hat{\tau}) = \mathcal{P}_h(\operatorname{sgn}(v_h)|v_h|^{q-1}) = \operatorname{sgn}(v_h)|v_h|^{q-1}. \tag{3.15}$$

In turn, utilizing the triangle inequality, and the estimate (3.9), with $r = 2$, we obtain

$$\|\hat{\tau}_h\|_{0,\Omega} \quad \leq \|\hat{\tau} - \hat{\tau}_h\|_{0,\Omega} + \|\hat{\tau}\|_{0,\Omega} \leq Ch|\hat{\tau}|_{1,\Omega} + \|\hat{\tau}\|_{0,\Omega},$$

which together to the inequalities in (3.13) and (3.14), implies

$$\|\hat{\tau}_h\|_{0,\Omega} \;\leq\; \{CC_2 h^{1+\frac{n}{2}-\frac{n}{p}} + C_1\}\||v_h|^{q-1}\|_{L^p(\Omega)} \tag{3.16}$$

Hence, noting that $1 + \frac{n}{2} - \frac{n}{p} > 0$ and $\||v_h|^{q-1}\|_{L^p(\Omega)} = \|v_h\|_{\mathrm{Q}}^{q-1}$, from (3.15) and (3.16), we get

$$\|\hat{\tau}_h\|_{\mathrm{H}} = \left\{\|\hat{\tau}_h\|_{0,\Omega}^2 + \|\operatorname{div}(\hat{\tau}_h)\|_{L^p(\Omega)}^2\right\}^{1/2} \leq \hat{c}\|v_h\|_{\mathrm{Q}}^{q-1}, \tag{3.17}$$

with $\hat{c} > 0$ independent of h.

Therefore, from (3.15) and (3.17), we obtain

$$\sup_{\substack{\tau_h \in \mathrm{H}_h \\ \tau_h \neq \theta}} \frac{b(\tau_h, v_h)}{\|\tau_h\|_{\mathrm{H}}} \geq \frac{b(\hat{\tau}_h, v_h)}{\|\hat{\tau}_h\|_{\mathrm{H}}} \geq \frac{1}{\hat{c}} \frac{\displaystyle\int_\Omega v_h \operatorname{sgn}(v_h)|v_h|^{q-1}}{\|v_h\|_{\mathrm{Q}}^{q-1}} = \frac{1}{\hat{c}} \frac{\|v_h\|_{\mathrm{Q}}^q}{\|v_h\|_{\mathrm{Q}}^{q-1}} = \frac{1}{\hat{c}} \|v_h\|_{\mathrm{Q}}. \tag{3.18}$$

which concludes the proof with $\beta^* = \frac{1}{\hat{c}}$. □

Remark 3.2 *Observe that in the proof of Lemma 3.1 above we are strongly using the fact that v_h is a piecewise constant function. In fact, if v_h is a piecewise polynomial function of degree k, with $k \geq 1$, $|v_h|^{q-1}$ is not necessarily in the same discrete space where v_h lives and* (3.15) *fails.*

We now look at the discrete kernel of b, which is defined by

$$V_h := \{\tau_h \in \mathrm{H}_h \,:\, b(\tau_h, v_h) = 0, \, \forall \, v_h \in \mathrm{Q}_h\} \,.$$

Since, $\operatorname{div} \mathrm{H}_h \subseteq \mathrm{Q}_h$, it readily follows that

$$V_h = \{\tau_h \in \mathrm{H}_h \,:\, \operatorname{div} \tau_h = 0 \quad \text{in} \quad \Omega\} \,. \tag{3.19}$$

The coerciveness of a in V_h is shown next.

Lemma 3.3 *There exists $\alpha^* > 0$, independent of h, such that*

$$a(\tau_h, \tau_h) \geq \alpha^* \|\tau_h\|_H^2 \quad \forall \, \tau_h \in V_h.$$

Proof. According to the definition of V_h, the proof is straightforward with $\alpha^* = 1$. □

Owing to Lemmata 3.1 and 3.3, we are now in position of establishing the solvability and stability of the Galerkin scheme (3.2), and the corresponding a priori error estimate.

Theorem 3.4 *Let $p \in (\frac{2n}{n+2}, 2)$ and $f \in L^p(\Omega)$. Assume that $\mathcal{T}_h$ is a quasi-uniform triangulation of Ω. Then, there exists a unique $(\sigma_h, u_h) \in \mathrm{H}_h \times \mathrm{Q}_h$ solution to* (3.2)*. In addition, there exist C_1, $C_2 > 0$, independent of h, such that*

$$\|(\sigma_h, u_h)\|_{\mathrm{H} \times \mathrm{Q}} \leq C_1 \big\{ \|f\|_{L^p(\Omega)} \,+\, \|u_D\|_{1/2,\Gamma} \big\} \tag{3.20}$$

$$\|(\sigma - \sigma_h, u - u_h)\|_{\mathrm{H} \times \mathrm{Q}} \leq C_2 \left\{ \inf_{\tau_h \in \mathrm{H}_h} \|\sigma - \tau_h\|_{\mathrm{H}} + \inf_{v_h \in \mathrm{Q}_h} \|u - v_h\|_{\mathrm{Q}} \right\} \tag{3.21}$$

where $(\sigma, u) \in \mathrm{H} \times \mathrm{Q}$ is the unique solution (2.11)*.*

Proof. It follows from Lemmata 3.1 and 3.3, and a direct application of the discrete Babuška-Brezzi theory. □

We now provide the rate of convergence of our mixed finite element method.

Theorem 3.5 *Let $p \in (\frac{2n}{n+2}, 2)$ and $f \in L^p(\Omega)$. Assume that $\mathcal{T}_h$ is a quasi-uniform triangulation of Ω. Let $(\sigma, u) \in \mathrm{H} \times \mathrm{Q}$ and $(\sigma_h, u_h) \in \mathrm{H}_h \times \mathrm{Q}_h$ be the unique solutions of the continuous and discrete mixed formulations (2.11) and (3.2), respectively. Assume that $\sigma \in [\mathrm{H}^1(\Omega)]^n$, $\operatorname{div} \sigma \in W^{1,p}(\Omega)$ and $u \in W^{1,q}(\Omega)$. Then, there exists $C > 0$, independent of h, such that*

$$\|(\sigma, u) - (\sigma_h, u_h)\|_{\mathrm{H} \times \mathrm{Q}} \leq Ch \left\{ \|\sigma\|_{1,\Omega} + \|\operatorname{div} \sigma\|_{W^{1,p}(\Omega)} + \|u\|_{W^{1,q}(\Omega)} \right\} \tag{3.22}$$

Proof. First, from the approximation property (3.8) with $r = q$ and $t = 1$, we easily obtain

$$\|u - \mathcal{P}_h(u)\|_{\mathrm{Q}} \leq Ch\|u\|_{W^{1,q}(\Omega)}. \tag{3.23}$$

Similarly, from (3.9) and (3.10), with $r = 2$ and $r = p$, respectively, we get

$$\|\sigma - \Pi_h(\sigma)\|_{0,\Omega} \leq Ch\|\sigma\|_{1,\Omega} \quad \text{and} \quad \|\operatorname{div} \sigma - \operatorname{div}(\Pi_h(\sigma))\|_{L^p(\Omega)} \leq Ch\|\operatorname{div} \sigma\|_{W^{1,p}(\Omega)}. \tag{3.24}$$

In this way, (3.22) readily follows from (3.23), (3.24) and the Céa estimate (3.21). □

4 Analysis of a convection-difussion problem

In this section we address the solvability and numerical approximation of the convection-diffusion problem (1.3). To that end, we first notice that the mixed structure of (1.3) is not symmetric because of the convective term. Then, in order to analyze the corresponding continuous and discrete problems, as well as the convergence of the associated Galerkin scheme, we will require a suitable generalization of the classical Babuška-Brezzi theory. This generalization is established next.

4.1 A generalization of the Babuška-Brezzi theory

Let H and Q be Hilbert spaces, and let H_h and Q_h be finite dimensional subspaces of H and Q, respectively. Moreover, let $a(\cdot,\cdot) : \mathrm{H} \times \mathrm{H} \to \mathbb{R}$, $b(\cdot,\cdot) : \mathrm{H} \times \mathrm{Q} \to \mathbb{R}$ and $d(\cdot,\cdot) : \mathrm{H} \times \mathrm{Q} \to \mathbb{R}$ be continuous bilinear forms, with continuity constants $\|a\|$, $\|b\|$ and $\|d\|$, and let $F \in \mathrm{H}'$ and $G \in \mathrm{Q}'$. In what follows we establish sufficient conditions to guarantee the well-posedness of the continuous and discrete problems:
Find $(\sigma, u) \in \mathrm{H} \times \mathrm{Q}$, such that

$$\begin{aligned} a(\sigma, \tau) + b(\tau, u) &= F(\tau) \qquad \forall\, \tau \in \mathrm{H}, \\ b(\sigma, v) + d(\sigma, v) &= G(v) \qquad \forall\, v \in \mathrm{Q}, \end{aligned} \tag{4.1}$$

Find $(\sigma_h, u_h) \in \mathrm{H}_h \times \mathrm{Q}_h$, such that

$$\begin{aligned} a(\sigma_h, \tau_h) + b(\tau_h, u_h) &= F(\tau_h) \qquad \forall\, \tau \in \mathrm{H}_h, \\ b(\sigma_h, v_h) + d(\sigma_h, v_h) &= G(v_h) \qquad \forall\, v \in \mathrm{Q}_h. \end{aligned} \tag{4.2}$$

We start by introducing the following hypotheses on a and b:

(**H.1**) There exists $\alpha > 0$, such that

$$a(\tau, \tau) \geq \alpha\|\tau\|_{\mathrm{H}}^2, \qquad \forall\, \tau \in K := \{\tau \in \mathrm{H} : b(\tau, v) = 0, \quad \forall\, v \in \mathrm{Q}\}.$$

($\mathbf{H}.2$) There exists $\beta > 0$, such that

$$\sup_{\tau \in \mathrm{H} \setminus \mathbf{0}} \frac{b(\tau, v)}{\|\tau\|_{\mathrm{H}}} \geq \beta \|v\|_{\mathrm{Q}} \quad \forall\, v \in \mathrm{Q}.$$

($\mathbf{H}.3$) There exists $\alpha^* > 0$, independent of discretization parameter h, such that

$$a(\tau, \tau) \geq \alpha^* \|\tau\|^2_{\mathrm{H}_h}, \qquad \forall\, \tau \in K_h := \{\tau \in \mathrm{H}_h : b(\tau, v) = 0, \quad \forall\, v \in \mathrm{Q}_h\}.$$

($\mathbf{H}.4$) There exists $\beta^* > 0$, independent of of discretization parameter h, such that

$$\sup_{\tau \in \mathrm{H}_h \setminus \mathbf{0}} \frac{b(\tau, v)}{\|\tau\|_{\mathrm{H}_h}} \geq \beta^* \|v\|_{\mathrm{Q}_h} \quad \forall\, v \in \mathrm{Q}_h.$$

Observe that the hypotheses above are nothing but the conditions of the classical Babuška-Brezzi theory.

Under a further assumption on $\|d\|$ it is possible to obtain the well-posedness of (4.1) and (4.2), and the corresponding a priori error estimate. In fact, we have the following theorems. For their proofs we refer to the forthcoming paper [7].

Theorem 4.1 *Assume that a and b satisfy hypotheses ($\mathbf{H}.1$) and ($\mathbf{H}.2$). Assume further that*

$$C_{wp} \|d\| \leq \frac{1}{2}, \tag{4.3}$$

with

$$C_{wp} := \frac{1}{\beta} + \frac{2\|a\|}{\alpha\beta} + \frac{\|a\|}{\beta^2} + \frac{\|a\|^2}{\alpha\beta^2}. \tag{4.4}$$

Then, there exists a unique $(\sigma, u) \in \mathrm{H} \times \mathrm{Q}$ solution to (4.1). In addition, there exists $C > 0$, depending only on $\|a\|$, $\|d\|$, α and β, such that

$$\|\sigma\|_{\mathrm{H}} + \|u\|_{\mathrm{Q}} \leq C(\|F\|_{\mathrm{H}'} + \|G\|_{\mathrm{Q}'}). \tag{4.5}$$

Theorem 4.2 *Assume that a and b satisfy hypotheses ($\mathbf{H}.3$) and ($\mathbf{H}.4$). Assume further that*

$$C^*_{wp} \|d\| \leq \frac{1}{2}, \tag{4.6}$$

with

$$C^*_{wp} := \frac{1}{\beta^*} + \frac{2\|a\|}{\alpha^*\beta^*} + \frac{\|a\|}{{\beta^*}^2} + \frac{\|a\|^2}{\alpha^*{\beta^*}^2}. \tag{4.7}$$

Then, there exists a unique $(\sigma_h, u_h) \in \mathrm{H}_h \times \mathrm{Q}_h$ solution to (4.2) Moreover, there exists $C^* > 0$, depending only on $\|a\|$, $\|d\|$, α^* and β^*, such that

$$\|\sigma_h\|_{\mathrm{H}} + \|u_h\|_{\mathrm{Q}} \leq C^*(\|F|_{\mathrm{H}_h}\|_{\mathrm{H}'_h} + \|G|_{\mathrm{Q}_h}\|_{\mathrm{Q}'_h}) \tag{4.8}$$

Theorem 4.3 *Let $(\sigma, u) \in \mathrm{H} \times \mathrm{Q}$ and $(\sigma_h, u_h) \in \mathrm{H}_h \times \mathrm{Q}_h$ be the unique solutions of problems (4.1) and (4.2), respectively. Assume that*

$$\max\{C_{wp}, C^*_{wp}\} \|d\| \leq \frac{1}{2}, \tag{4.9}$$

with $C_{wp} > 0$ and $C^*_{wp} > 0$ be the constants defined in (4.4) and (4.7), respectively. Then, there exists $C_{cea} > 0$, depending only on $\|a\|$, $\|b\|$, α^* and β^*, such that

$$\|\sigma - \sigma_h\|_{\mathrm{H}} + \|u - u_h\|_{\mathrm{Q}} \leq C_{cea} \left\{ \inf_{\tau_h \in \mathrm{H}_h} \|\sigma - \tau_h\|_{\mathrm{H}} + \inf_{v_h \in \mathrm{Q}_h} \|u - v_h\|_{\mathrm{Q}} \right\}. \tag{4.10}$$

4.2 Analysis of the continuous convection-diffusion problem

Let us consider the bilinear forms

$$a(\sigma,\tau) := \int_\Omega \sigma\cdot\tau, \qquad b(\tau,\psi) := \int_\Omega \psi \operatorname{div}\tau, \qquad d(\tau,\psi) := -\int_\Omega (\mathbf{v}\cdot\tau)\psi,$$

and the functionals

$$F(\tau) := \langle \tau\cdot\nu, \theta_D\rangle_\Gamma \quad \text{and} \quad G(\psi) := -\int_\Omega g\psi.$$

Then, the convection-diffusion (1.3) reads: Find $(\sigma,\theta)\in \mathrm{H}(\operatorname{div}_{4/3},\Omega)\times L^4(\Omega)$, such that

$$\begin{aligned} a(\sigma,\tau)+b(\tau,\theta) &= F(\tau),\\ b(\sigma,\psi)+d(\sigma,\psi) &= G(\psi), \end{aligned} \tag{4.11}$$

for all $(\tau,\psi)\in \mathrm{H}(\operatorname{div}_{4/3},\Omega)\times L^4(\Omega)$.

It is clear that the structure of (4.11) fits into the framework of the theory introduced in Section 4. Therefore, in what follows we apply Theorem 4.1 to prove the well-posedness of (4.11). To that end, we first observe that the forms a, b and d are continuous with continuity constants

$$\|a\| = 1, \qquad \|b\| = 1 \quad \text{and} \quad \|d\| = C_{Sob}\|\mathbf{v}\|_{1,\Omega}, \tag{4.12}$$

where $C_{Sob}>0$ is the positive constant of the Sobolev inequality (2.9). In addition, owing to (2.6) and the fact that $\|\psi\|_{0,\Omega}\le |\Omega|^{1/2}\|\psi\|_{L^4(\Omega)}$, for all $\psi\in L^4(\Omega)$, it is easy to see that F and G are bounded:

$$|F(\tau)|\le c(\Omega)\|\theta_D\|_{1/2,\Gamma}\|\tau\|_{\mathrm{H}(\operatorname{div}_{4/3},\Omega)} \quad \text{and} \quad |G(\psi)|\le |\Omega|^{1/2}\|g\|_{0,\Omega}\|\psi\|_{L^4(\Omega)}$$

In turn, by applying Lemmata 2.1 and 2.2 with $p=4/3$ and $q=4$, we readily obtain that a and b satisfy hypotheses (**H.1**) and (**H.2**) with $\beta>0$ established in (2.19) and ellipticity constant $\alpha=1$.

According to the discussion above, we now can establish the well-posedness of (4.11).

Theorem 4.4 *Assume that*

$$\frac{C_{Sob}(3\beta+2)}{\beta^2}\|\mathbf{v}\|_{1,\Omega}\le \frac{1}{2}. \tag{4.13}$$

Then, there exists a unique $(\sigma,\theta)\in \mathrm{H}(\operatorname{div}_{4/3},\Omega)\times L^4(\Omega)$ solution to (4.11). In addition, there exists $C>0$, independent of the solution, such that

$$\|\sigma\|_{\mathrm{H}(\operatorname{div}_{4/3},\Omega)}+\|\theta\|_{L^4(\Omega)}\le C(\|\theta_D\|_{1/2,\Gamma}+\|g\|_{0,\Omega}). \tag{4.14}$$

Proof. Having verified hypotheses (**H.1**) and (**H.2**) the proof is a straightforward application of Theorem 4.1. □

4.3 Finite element discretization of the convection-diffusion problem

Let $\mathrm{H}_h \subseteq \mathrm{H}(\operatorname{div}_{4/3}, \Omega)$ and $\mathrm{Q}_h \subseteq L^4(\Omega)$ be the finite element spaces defined in (3.1), that is

$$\mathrm{H}_h := \{\tau_h \in \mathrm{H} \,:\, \tau_h|_T \in \mathrm{RT}_0(T), \quad \forall\, T \in \mathcal{T}_h\} \subseteq \mathrm{H},$$

$$\mathrm{Q}_h := \{v_h \in \mathrm{Q} \,:\, v_h|_T \in \mathbb{P}_0(T), \quad \forall\, T \in \mathcal{T}_h\} \subseteq \mathrm{Q},$$

where $\mathcal{T}_h$ is a quiasiuniform mesh. Then, the Galerkin scheme of (4.11) reads: Find $(\sigma_h, \theta_h) \in \mathrm{H}_h \times \mathrm{Q}_h$, such that

$$\begin{aligned} a(\sigma_h, \tau_h) + b(\tau_h, \theta_h) \quad &= F(\tau_h), \\ b(\sigma_h, \psi_h) + d(\sigma_h, \psi_h) \quad &= G(\psi_h), \end{aligned} \tag{4.15}$$

for all $(\tau_h, \psi_h) \in \mathrm{H}_h \times \mathrm{Q}_h$.

Similarly to the continuous case, we observe that owing to Lemmata 3.1 and 3.3 with $p = 4/3$ and $q = 4$, the bilinear forms a and b satisfy hypotheses (**H.3**) and (**H.4**) with $\alpha = 1$ and $\beta^* > 0$ be the constant established in (3.18).

The following theorem establishes the well-posedness of the Galerkin scheme (4.15) and the corresponding a priori estimate.

Theorem 4.5 *Assume that*

$$C_{Sob} \max\left\{ \frac{(3\beta + 2)}{\beta^2}, \frac{(3\beta^* + 2)}{{\beta^*}^2} \right\} \|\mathbf{v}\|_{1,\Omega} \leq \frac{1}{2}. \tag{4.16}$$

Then, there exists a unique $(\sigma_h, \theta_h) \in \mathrm{H}_h \times \mathrm{Q}_h$ *solution to* (4.15). *In addition, there exists* C_1, $C_2 > 0$, *independent of* h, *such that*

$$\|\sigma_h\|_{\mathrm{H}(\operatorname{div}_{4/3},\Omega)} + \|\theta_h\|_{L^4(\Omega)} \leq C_1(\|\theta_D\|_{1/2,\Gamma} + \|g\|_{0,\Omega}) \tag{4.17}$$

and

$$\|\sigma - \sigma_h\|_{\mathrm{H}(\operatorname{div}_{4/3},\Omega)} + \|\theta - \theta_h\|_{L^4(\Omega)} \leq C_2 \left\{ \inf_{\tau_h \in \mathrm{H}_h} \|\sigma - \tau_h\|_{\mathrm{H}(\operatorname{div}_{4/3},\Omega)} + \inf_{\psi_h \in \mathrm{Q}_h} \|\theta - \psi_h\|_{L^4(\Omega)} \right\}. \tag{4.18}$$

Proof. Since a and b satisfy hypotheses (**H.3**) and (**H.4**) the proof follows from a direct application of Theorems 4.2 and 4.3. □

The following theorem provides the theoretical rate of convergence of the Galerkin scheme (4.15), under suitable regularity assumptions on the exact solution.

Theorem 4.6 *Let* $(\sigma, \theta) \in \mathrm{H}(\operatorname{div}_{4/3}, \Omega) \times L^4(\Omega)$ *and* $(\sigma_h, \theta_h) \in \mathrm{H}_h \times \mathrm{Q}_h$ *be the unique solutions of* (4.11) *and* (4.15), *respectively. Assume that* $\sigma \in \mathrm{H}^1(\Omega)$, $\operatorname{div} \sigma \in W^{1,4/3}(\Omega)$ *and* $\theta \in W^{1,4}(\Omega)$. *Then there exists* $C > 0$, *independent of* h, *such that*

$$\|\sigma - \sigma_h\|_{\mathrm{H}(\operatorname{div}_{4/3},\Omega)} + \|\theta - \theta_h\|_{L^4(\Omega)} \leq Ch \left\{ \|\sigma\|_{1,\Omega} + \|\operatorname{div} \sigma\|_{W^{1,4/3}(\Omega)} + \|\theta\|_{W^{1,4}(\Omega)} \right\}. \tag{4.19}$$

Proof. The proof follows from the Céa estimate (4.18) and Theorem 3.5. □

5 Numerical results

In this section we corroborate numerically the theory developed for the Poisson problem as applied to the convection-diffusion problem (1.2). More precisely, in what follows we present two examples illustrating the performance of the Galerkin scheme (4.15) on a set of quasi-uniform triangulations. Our implementation is based on a *FreeFem++* code (see [16]), in conjunction with the direct linear solver UMFPACK (see [10]).

We now introduce some additional notations. The individual errors are denoted by:

$$\mathsf{e}(\sigma) := \|\sigma - \sigma_h\|_{\mathrm{H}(\mathrm{div}_{4/3},\Omega)} \quad \text{and} \quad \mathsf{e}(\theta) := \|\theta - \theta_h\|_{L^4(\Omega)}.$$

Also, we let $r(\sigma)$ and $r(\theta)$ be the experimental rates of convergence given by

$$r(\sigma) := \frac{\log(\mathsf{e}(\sigma)/\mathsf{e}'(\sigma))}{\log(h/h')} \quad \text{and} \quad r(\theta) := \frac{\log(\mathsf{e}(\theta)/\mathsf{e}'(\theta))}{\log(h/h')},$$

where h and h' denote two consecutive meshsizes with errors e and e'.

In Example 1 we verify the theory for the two dimensional case. To that end, we choose the domain $\Omega := (0,1)^2$, the vector field $\mathbf{v}(x_1,x_2) := (e^{x_1}, e^{x_2})^t$ and take g and θ_D so that the exact solution is given by

$$\sigma(x_1,x_2) := \begin{pmatrix} 2x_1\sin(\pi x_2) \\ \\ \pi x_1^2\cos(\pi x_2) \end{pmatrix}, \qquad \theta(x_1,x_2) := x_1^2\sin(\pi x_2).$$

Next, in Example 2 we assess the capability of a 3D implementation of the Galerkin scheme (4.15). Here, we choose the domain $\Omega := (0,1)^3$, the vector field $\mathbf{v}(x_1,x_2,x_3) := (x_1^2, x_2^2, 0)^t$ and take g and θ_D so that the exact solution is given by

$$\sigma(x_1,x_2,x_3) := \begin{pmatrix} x_2(x_3 + e^{x_3+x_1x_2}) \\ x_1(x_3 + e^{x_3+x_1x_2}) \\ e^{x_3+x_1x_2} + x_1x_2 \end{pmatrix} e^{2x_1+x_2}, \qquad \theta(x_1,x_2,x_3) := e^{x_3+x_1x_2} + x_1x_2x_3.$$

In Table 5.1 below, we summarize the convergence history for a sequence of quasi-uniform triangulations. We observe there that the rate of convergence $O(h)$ predicted by Theorem 4.5 is attained in all the cases. Similar results can be seen in Table 5.2 for the 3D case. Next, in figures 5.1 and 5.2 we provide the graphics of the approximate and exact solutions of Example 2. In Figure 5.1 we display the isosurface of θ_h (to the left) and we compare it with its exact counterpart (to the right). In addition, in Figure 5.2 we display the components of the vector field σ_h (top) and we compare them with their exact counterpart (bottom). Here, we display the section of the cube below the plane $x_1 - x_2 + x_3 = 0.5$. All the graphics above were computed with $N = 595968$ degrees of freedom. We observe there that the mixed finite element method provide very accurate approximations to the unknowns. In addition, we notice that the election of $\mathbf{v}$ in both cases leads to a good behaviour of the numerical method. It is pertinent to mention here that the actual influence of assumption (4.16) on the performance of the numerical approximation of (4.11) in both examples is not analyzed in this work since it escapes from the original purposes of this paper and remains an open problem to be addressed in the future. However, there is numerical evidence showing that when having non-symmetric structures as the one presented in (4.1), the associate global matrix of the system becomes ill-posed as $\|d\|$ is too big (see [6, Section 7, Example 1]).

N	h	$\mathsf{e}(\sigma)$	$r(\sigma)$	$\mathsf{e}(\theta)$	$r(\theta)$
1512	0.1074	0.2124	–	0.0287	–
6124	0.0501	0.1042	0.9348	0.0137	0.9691
24273	0.0265	0.0516	1.1012	0.0069	1.0789
98206	0.0131	0.0257	0.9878	0.0034	0.9887
387402	0.0075	0.0128	1.2509	0.0017	1.2793
1541734	0.0039	0.0065	1.0506	0.0009	1.0384

Table 5.1: EXAMPLE 1: Degrees of freedom, meshsizes, errors, rates of convergence $\mathrm{RT}_0 - \mathrm{P}_0$ approximation of the convection-diffusion problem (4.11) in 2D .

N	h	$\mathsf{e}(\sigma)$	$r(\sigma)$	$\mathsf{e}(\theta)$	$r(\theta)$
168	0.7071	1.2336	–	0.8134	–
1248	0.3536	0.6276	0.9749	0.4231	0.9428
9600	0.1768	0.3154	0.9928	0.2137	0.9854
75264	0.0884	0.1579	0.9980	0.1071	0.9963
595968	0.0442	0.0790	0.9995	0.0536	0.9991

Table 5.2: EXAMPLE 2: Degrees of freedom, meshsizes, errors, rates of convergence $\mathrm{RT}_0 - \mathrm{P}_0$ approximation of the convection-diffusion problem (4.11) in 3D .

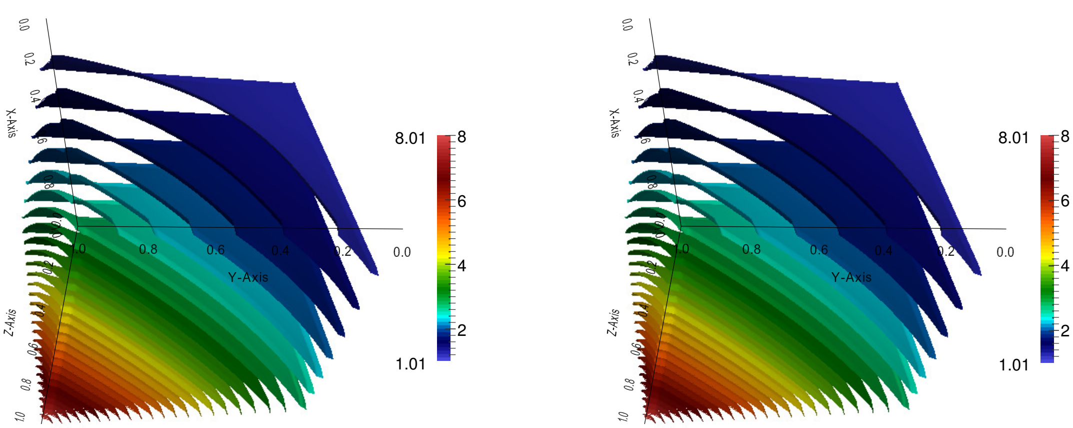

Figure 5.1: Example 2: isosurfaces of θ_h (left) and θ (right), with $N = 595968$.

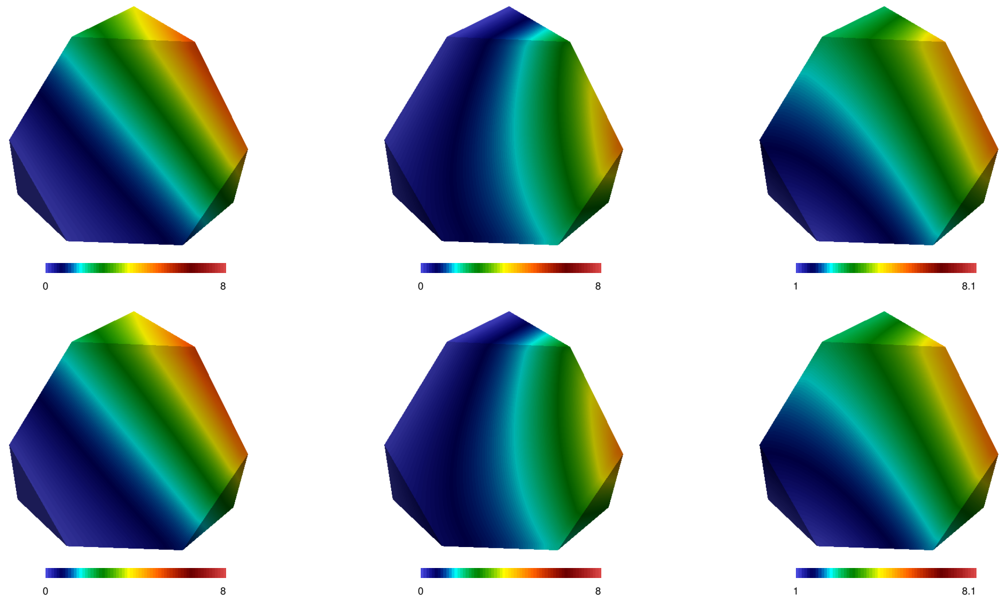

Figure 5.2: Example 2: $\sigma_{1,h}$, $\sigma_{3,h}$, $\sigma_{3,h}$ (from the left to the right, at the top) and σ_1, σ_2, σ_3 (from the left to he right, at the bottom) with $N = 595968$.

References

[1] G. Acosta, T. Apel, R.G. Durán, A.L. Lombardi, *Error estimates for the Raviart-Thomas interpolation of any order on anisotropic tetrahedra* Mathematics of Computation, vol. 80, no. 273, pp. 141-163, (2010).

[2] M. Alvarez, G.N. Gatica, R. Ruiz-Baier, *An augmented mixed-primal finite element method for a coupled flow-transport problem*. ESAIM: Mathematical Modelling and Numerical Analysis, vol. 49, 5, pp. 1399-1427, (2015).

[3] C. Bernardi, B. Métivet, B. Pernaud-Thomas, *Couplage des équations de Navier-Stokes et de la chaleur: le modèle et son approximation par éléments finis*. ESAIM: Mathematical Modelling and Numerical Analysis, vol. 29, no. 7, pp. 871-921, (1995).

[4] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*. Springer Series in Computational Mathematics, 15. Springer-Verlag, New York, 1991.

[5] R. Bürger, R. Ruiz-Baier, H. Torres, *A stabilized finite volume element formulation for sedimentation-consolidation processes*. SIAM Journal on Scientific Computing, vol. 34, 3, pp. B265–B289, (2012).

[6] S. Caucao, D. Mora and R. Oyarzúa, *A priori and a posteriori error analysis of a pseudostress-based mixed formulation of the Stokes problem with varying density*. IMA Journal of Numerical Analysis, DOI: 10.1093/imanum/drv015.

[7] J. Camaño and R. Oyarzúa, *Well-posedness and Galerkin approximation of a nonsymmetric mixed variational formulation*. In preparation.

[8] E. Colmenares, G.N. Gatica and R. Oyarzúa, *Analysis of an augmented mixed–primal formulation for the stationary Boussinesq Problem.* Numerical Methods for Partial Differential Equations, DOI:10.1002/num.22001, (2015).

[9] E. Colmenares, G.N. Gatica and R. Oyarzúa, *An augmented fully-mixed finite element method for the stationary Boussinesq problem.* Preprint 2015-30, Centro de Investigación en Ingeniería Matemática (CI^{2}MA), Universidad de Concepción-Chile, (2015)

[10] T. Davis, *Algorithm 832: UMFPACK V4.3 - an unsymmetric-pattern multifrontal method.* ACM Transactions on Mathematical Software, vol. 30, pp. 196–199, (2004).

[11] R.G. Durán, *Error analysis in L^p, $1 \leq p \leq \infty$, for mixed finite element methods for linear and quasi-linear elliptic problems.* Mathematical Modelling and Numerical Analysis, vol. 22, no. 3, pp. 371–387, (1998).

[12] A. Ern and J.-L. Guermond, *Theory and Practice of Finite Elements.* Applied Mathematical Sciences, Springer-Verlag, New York, 2004., vol. 159, (2004).

[13] M. Farhloul, S. Nicaise and L. Paquet, *A mixed formulation of Boussinesq equations: Analysis of nonsingular solutions.* Mathematics of Computation, vol. 69, no. 231, pp. 965-986, (2000).

[14] G.N. Gatica, *A Simple Introduction to the Mixed Finite Element Method. Theory and Applications.* Springer Briefs in Mathematics, Springer, Cham, 2014. xii+132 pp.

[15] V. Girault and P.-A. Raviart, *Finite Element Approximation of the Navier–Stokes Equations. Lecture Notes in Mathematics*, 749. Springer-Verlag, Berlin-New York, 1979.

[16] F. Hecht, *New development in FreeFem++.* Journal of Numerical Mathematics, vol. 20, 3-4, pp. 251–265, (2012).

[17] R. Oyarzúa, T. Qin and D. Schötzau, *An exactly divergence-free finite element method for a generalized Boussinesq problem.* IMA journal of numerical analysis, vol. 34, no. 3, pp. 1104–1135, (2014).

[18] J.E. Roberts and J.M. Thomas, *Mixed and Hybrid Methods.* In: Handbook of Numerical Analysis, P.G. Ciarlet and J.L. Lions (Eds.), vol. II, Finite Element Methods (Part 1), 1991, North-Holland, Amsterdam.