UNIVERSIDAD DE CONCEPCIÓN

Centro de Investigación en Ingeniería Matemática (CI^{2}MA)

Coupled mixed finite element and finite volume methods for a solid velocity-based model of multidimensional settling

Julio Careaga, Gabriel N. Gatica

PREPRINT 2023-01

SERIE DE PRE-PUBLICACIONES

Coupled mixed finite element and finite volume methods for a solid velocity-based model of multidimensional settling*

Julio Careaga† Gabriel N. Gatica‡

Dedicated to Professor Stefan Diehl on the occasion of his 60th birthday

Abstract

In this paper we introduce and analyze a model of sedimentation based on a solid velocity formulation. A particular feature of the governing equations is given by the fact that the velocity field is non-divergence free. We introduce extra variables such as the pseudostress tensor relating the velocity gradient with the pressure, thus leading to a mixed variational formulation consisting of two systems of equations coupled through their source terms. A result of existence and uniqueness of solutions is shown by means of a fixed-point strategy and the help of the Babuška-Brezzi theory and Banach theorem. Additionally, we employ suitable finite dimensional subspaces to approximate both systems of equations via associated mixed finite element methods. The well-posedness of the resulting coupled scheme is also treated via a fixed-point approach, and hence the discrete version of the existence and uniqueness result is derived analogously to the continuous case. The above is then combined with a finite volume method for the transport equation. Finally, several numerical results illustrating the performance of the proposed model and the full numerical scheme, and confirming the theoretical rates of convergence, are presented.

Key words: solid phase velocity, multidimensional settling, mixed finite elements, fixed-point

Mathematics Subject Classifications (2020): 35Q70, 65M08, 65N12, 65N15, 65N30, 70-08, 76T20

1 Introduction

The phenomenon of sedimentation of suspended solid particles immersed in a fluid under the influence of gravity is present in a wide range of engineering applications. This physical process is used in several industries such as food, biomedical, brewing, paper manufacturing, petroleum, oil, mineral processing and wastewater treatment industry. Most of the models of sedimentation are derived from the theory of continuum mechanics and involve coupled systems of partial differential equations (PDEs). Such models are typically obtained from the equations of conservation of mass and momentum for the solid and fluid phase. If the viscous, gravity and solid-fluid interaction forces of both phases are taken into account, the number of unknowns becomes considerable, where constitutive assumptions and a reduction of the model equations and variables are desired. With the help of dimensional analysis many models are reduced to a scalar transport-flow equation for the local concentration of solid particles coupled with a Stokes-type system for the velocity-pressure pair.

*This work was partially supported by ANID-Chile through the projects Centro de Modelamiento Matemático (FB210005), and Anillo of Computational Mathematics for Desalination Processes (ACT210087); and by Centro de Investigación en Ingeniería Matemática (CI^{2}MA)

†CI^{2}MA and Departamento de Ingeniería Matemática, Universidad de Concepción, Casilla 160-C, Concepción, Chile, email: juliocareaga@udec.cl

‡CI^{2}MA and Departamento de Ingeniería Matemática, Universidad de Concepción, Casilla 160-C, Concepción, Chile, email: ggatica@ci2ma.udec.cl.

Inspired by the work done by Gustavsson and Opelstrup [45], we study a model for multidimensional sedimentation based on the solid phase velocity field $\boldsymbol{u}$, pressure p and volume fraction of the solid particles ϕ. The model equations, obtained by combining the ideas of [45] and [23], are essentially a variant of the model presented in [45], where the velocity field has the peculiarity of not being divergence-free. The system is strongly coupled containing nonlinear terms, and therefore the construction of numerical schemes for its approximation are involved and need an special treatment. For an introduction to the theory of sedimentation based on the postulates of continuum mechanics for mixtures of two superimposed continuous media we refer to [24, 35, 62]. One-dimensional PDE models of sedimentation have been studied widely in the literature, from the development of numerical approximations for batch and continuous settling to the study of steady state solutions, see e.g. [18, 19, 20, 31, 32, 33, 36, 41]. The special case of quasi-one-dimensional models where the influence of the domain geometry is reflected in coefficient terms of a one-dimensional conservation law have been studied in [8, 15, 16, 17, 25].

Several mathematical models and numerical approximations of multidimensional sedimentation have been reported in the literature. Bürger, Concha and Wendland [23] presented a complete derivation of a multidimensional model of sedimentation starting from the conservation equations (mass and momentum) for the solid and fluid phases. After a dimensional analysis the model is reduced to a Stokes type system coupled to a strongly degenerate convection-diffusion equation, consisting of three unknowns and whose velocity field is the so called "volume average velocity of the mixture" denoted by $\boldsymbol{q}$. Borhan and Acrivos [14] introduce a two-dimensional model for the batch sedimentation in inclined channels, with a single momentum equation for the mixture which varies in its definition depending on the fluid-suspension interfase. In [49], the authors provide a detailed discussion about the viscouss stress-tensors of the solid phase and mixture based in the rheological complexity of colloidal suspensions. They also show numerical simulations of their approach in cilindrical coordinates for the case of a "deep-cone" thickener. Rao et. al. [54] presented a two-phase model based on the so called "mass-average velocity" (denoted by $\boldsymbol{v}$) for the simulation of batch sedimentation and shear between concentric rotating cylinders. The velocity field, does not satisfy the condition of divergence free, and the continuity equation is no longer solenoidal since the density varies with the local volume fraction. The batch model derived by Gustavsson and Oppelstrup in [45] is based on the solid phase velocity field $\boldsymbol{u}$ and contemplates the possibility of a horizontal movement of the bottom of the vessel by having Dirichlet boundary conditions on $\boldsymbol{u}$. One of the main features of this model that differs from typical $\boldsymbol{q}$-based ones (see [23, 22]) is that like the model in [54], the divergence of the velocity field is non-zero but varies with the pressure and local volume fraction. A finite differences scheme is introduced in [45] to simulate the model equations.

A variety of advanced numerical schemes have been used to compute numerical approximations for multidimensional models of sedimentation [21, 22, 46, 52, 54, 61]. In [46], the authors develop a finite element numerical scheme for a $\boldsymbol{q}$-based model coupled with the "κ-ϵ equations" for turbulent flows. They also treat the time approximations with a fractional-step scheme in order to have second order accuracy and preserve physical oscillations. Bürger et. al. [21], introduced a multiresultion finite volume scheme for the two-dimensional batch-sedimentation model given in [23], and presented numerical examples of the process in channels with inclined walls. In [22], the authors worked on an advanced numerical method combining a stabilized finite element scheme with finite volumes in staggered grids for the continuous sedimentation of the model [23]. Numerical simulations of the continuous sedimentation applied to wastewater treatment plants are also included. Another paper dedicated to the approximation of such models using a combined finite element-finite volume schemes can be found in [52]. In [61], a CFD approach for a $\boldsymbol{v}$-based model coupled to κ-ε equations is considered. On the other hand, and regarding the coupled flow and transport problem driven by a scalar non-linear reaction-diffusion equation interacting with the Stokes or Brinkman equations, which models the transport of species within a viscous flow, we stress that diverse combinations of primal

and mixed finite element methods have been developed in recent years for its numerical solution (see, e.g., [3, 4, 5, 6, 11, 12, 27]). Mathematical modeling based on CFD simulations of two-dimensional sedimentation can be found in [26, 29, 37, 47, 48]. Experiments of sedimentation in vessels with inclined walls can be found in [56, 59], and studies on the Boycott effect produced in these vessels were reported in [57, 58]. Sedimentation models coupled to a energy balance and a temperature equation are studiend in [53, 55]. Applications of sedimentation models can be found in mineral processing [30, 34, 38, 39], wastewater treatment [1, 40, 61], petroleum industry [50, 51], magma chambers [13, 60] and oil-in-water emulsions [28].

According to the previous discussion, the first purpose of the paper is to analyze the existence and uniqueness of weak solutions for the decoupled velocity-pressure system. Secondly, we seek to propose reliable finite element methods for its numerical solution, and finally we aim to approximate the transport equation by a conservative finite volume scheme. The paper is organized as follows. The rest of this section collects some useful notations to be employed along the paper. Then, in Section 2 we derive the model equations. Next, in Section 3 we introduce some auxiliary unknowns to derive a mixed variational formulation for the aforementioned system, and employ a fixed-point strategy along with the Babuška-Brezzi theory and Banach theorem to address its solvability. In Section 4, we consider suitable finite dimensional subspaces to define the associated mixed finite element scheme, and adopt the discrete analogue of the analysis from Section 3 to prove its well-posedness. A priori error estimates and corresponding rates of convergence are also provided here. Then, in Section 5 we propose a finite volume discretization for the transport equation and couple the resulting procedure with the velocity-pressure mixed finite element formulation from Section 4. Finally, illustrative numerical results are shown in Section 6 and concluding remarks are presented in Section 7.

1.1 Further notations

In this section we recall some standard notation used in the rest of the paper. For any vector field $\boldsymbol{w} = (w_i)_{i=1,d}$ in $\mathbb{R}^d$ we let $\nabla \boldsymbol{w}$ and $\mathrm{div}(\boldsymbol{w})$ be the gradient and divergence of $\boldsymbol{w}$, respectively. In addition, for any tensor $\boldsymbol{\tau} = (\tau_{ij})_{i,j=1,d}$ we let $\mathbf{div}(\boldsymbol{\tau})$ be the divergence operator div acting along the rows of $\boldsymbol{\tau}$, and $\boldsymbol{\tau}^{\mathrm{t}}$ is the transposed matrix of $\boldsymbol{\tau}$. Besides, the trace of $\boldsymbol{\tau}$, the sum of the elements on the main diagonal, is denoted by $\mathrm{tr}(\boldsymbol{\tau})$ and its deviator is defined by $\boldsymbol{\tau}^{\mathrm{d}} := \boldsymbol{\tau} - \frac{1}{d}\mathrm{tr}(\boldsymbol{\tau})\mathbb{I}$, where $\mathbb{I}$ is the identity matrix of order d. Note that the deviator is defined such that $\mathrm{tr}(\boldsymbol{\tau}^{\mathrm{d}}) = 0$. The inner product between the two matrix functions $\boldsymbol{\tau}$, $\boldsymbol{\eta}$ is given by $\boldsymbol{\tau} : \boldsymbol{\eta} = \mathrm{tr}(\boldsymbol{\eta}^{\mathrm{t}}\boldsymbol{\tau})$, and it follows from the definition that $\boldsymbol{\tau}^{\mathrm{d}} : \boldsymbol{\eta} = \boldsymbol{\tau}^{\mathrm{d}} : \boldsymbol{\eta}^{\mathrm{d}} = \boldsymbol{\tau} : \boldsymbol{\eta}^{\mathrm{d}}$. We consider domains $\Omega \subset \mathbb{R}^d$ for $d \in \{2, 3\}$ with polyhedral boundary $\Gamma := \partial\Omega$ and outward unit normal vector $\boldsymbol{n}$. In terms of notation, for any generic scalar functional space M, we denote its vectorial extension by $\mathbf{M}$ and its matrix counterpart by $\mathbb{M}$. Furthermore, we employ the standard notation for the space of square-integrable measurable scalar functions $L^2(\Omega)$, along with its norm $\|\cdot\|_{0;\Omega}$, and the Sobolev spaces $H^k(\Omega)$ for an integer k. We also introduce the spaces:

$$\mathbf{H}(\mathrm{div};\Omega) := \left\{ \boldsymbol{w} \in \mathbf{L}^2(\Omega) : \quad \mathrm{div}(\boldsymbol{w}) \in L^2(\Omega) \right\},$$
$$\mathbb{H}(\mathbf{div};\Omega) := \left\{ \boldsymbol{\tau} \in \mathbb{L}^2(\Omega) : \quad \mathbf{div}(\boldsymbol{\tau}) \in \mathbf{L}^2(\Omega) \right\},$$

with their norms $\|\cdot\|_{\mathrm{div};\Omega}$ and $\|\cdot\|_{\mathbf{div};\Omega}$, respectively. The norm $\|\cdot\|$, with no subscripts, shall be used for any element or operator when it does not cause confusion about the space to which it refers.

2 The model problem

In this section we present a concise derivation of a model of sedimentation based on the solid phase velocity field. Starting from the equations of conservation of mass and momentum of each phase, the

derivation of the model equations is mainly based on the ideas of [45] combined with some of the constitutive assumptions made in [23].

We begin by assuming that the sedimentation process occurs on a fixed domain $\Omega \subset \mathbb{R}^d$ for $d \in \{2, 3\}$, away from sources. Let the volume fraction of solid particles $\phi := \phi(\boldsymbol{x}, t)$ be a function of the position $\boldsymbol{x} \in \Omega$ and time $t \in (0, T]$, where $T \in \mathbb{R}^+$ is the end time. The function ϕ is a scalar quantity that takes values between zero and $\phi_{\max}$, the maximum volume fraction. The equations for the conservation of mass of both phases, solid and fluid, are given by

$$\partial_t \phi + \operatorname{div}(\phi \boldsymbol{u}) = 0, \tag{2.1}$$

$$\partial_t (1 - \phi) + \operatorname{div}((1 - \phi)\boldsymbol{u}_{\mathrm{f}}) = 0, \tag{2.2}$$

where $\boldsymbol{u} := \boldsymbol{u}(\boldsymbol{x}, t)$ and $\boldsymbol{u}_{\mathrm{f}} := \boldsymbol{u}_{\mathrm{f}}(\boldsymbol{x}, t)$, both vector quantities, are the solid and fluid velocity fields, respectively. For the conservation of momentum of each phase we consider the gravity, viscous and interaction forces, and pressure gradients. Then, neglecting the material derivatives of each phase, the momentum equations are determined by

$$-\mathbf{div}(\boldsymbol{T}_{\mathrm{s}}) + \nabla p_{\mathrm{s}} = -\rho_{\mathrm{s}} g \phi \boldsymbol{k} + \boldsymbol{m}, \tag{2.3}$$

$$-\mathbf{div}(\boldsymbol{T}_{\mathrm{f}}) + \nabla p_{\mathrm{f}} = -\rho_{\mathrm{f}} g \,(1 - \phi)\, \boldsymbol{k} - \boldsymbol{m}, \tag{2.4}$$

where $p_{\mathrm{s}} := p_{\mathrm{s}}(\boldsymbol{x}, t)$ and $p_{\mathrm{f}} := p_{\mathrm{f}}(\boldsymbol{x}, t)$ are the pressures of the solid and fluid phase, respectively, $\boldsymbol{T}_{\mathrm{s}}$ and $\boldsymbol{T}_{\mathrm{f}}$ are the solid and fluid viscous stress tensors, respectively, g is the acceleration of gravity and $\boldsymbol{k} := (0, 1)^{\mathrm{t}}$ is the upward unit vector. The term $\boldsymbol{m}$ present in both equations is the solid-fluid interaction force per unit volume, to be defined later. Adding the momentum equations (2.3) and (2.4) we get the momentum equation of the mixture:

$$-\mathbf{div}(\boldsymbol{T}_{\mathrm{mix}}) + \nabla(p_{\mathrm{f}} + p_{\mathrm{s}}) = -g\rho(\phi)\boldsymbol{k}, \tag{2.5}$$

where $\boldsymbol{T}_{\mathrm{mix}} := \boldsymbol{T}_{\mathrm{s}} + \boldsymbol{T}_{\mathrm{f}}$ is the viscous stress tensor of the mixture and $\rho := \rho(\phi) = \rho_{\mathrm{s}}\phi + \rho_{\mathrm{f}}(1 - \phi)$. Now, in order to reduce the number of unknowns we define the excess pore pressure p, the hidrostatic pressure p_{h} and the effective solid stress function $\sigma_{\mathrm{e}} := \sigma_{\mathrm{e}}(\phi)$, which is a given function that satisfies

$$\sigma_{\mathrm{e}}'(\phi) \begin{cases} = 0, & \text{if } \phi \le \phi_{\mathrm{c}}, \\ \ge 0, & \text{if } \phi_{\mathrm{c}} < \phi \le \phi_{\max}, \end{cases}$$

where $\phi_{\mathrm{c}} \in (0, \phi_{\max}]$ is the critical concentration, see [23]. Having these new variables, the pressures of both phases are then given by

$$p_{\mathrm{s}} = \phi(p + p_{\mathrm{h}}) + \sigma_{\mathrm{e}} \quad \text{and} \quad p_{\mathrm{f}} = (1 - \phi)(p + p_{\mathrm{h}}).$$

Moreover, computing the gradient $\nabla p_{\mathrm{h}} = -\rho_{\mathrm{f}} g \boldsymbol{k}$, we have

$$\nabla p_{\mathrm{f}} = (1 - \phi)\nabla p - (p + p_{\mathrm{h}})\nabla\phi - \rho_{\mathrm{f}} g (1 - \phi)\boldsymbol{k} \quad \text{and} \tag{2.6}$$

$$\nabla(p_{\mathrm{s}} + p_{\mathrm{f}}) = \nabla p + \nabla\sigma_{\mathrm{e}} - \rho_{\mathrm{f}} g \boldsymbol{k}. \tag{2.7}$$

The solid-fluid interaction force $\boldsymbol{m}$ after neglecting the material derivative of its dynamic part is modeled by a function depending on the relative velocity $\boldsymbol{v}_{\mathrm{r}} := \boldsymbol{u} - \boldsymbol{u}_{\mathrm{f}}$, ϕ, p and p_{h}, that is

$$\boldsymbol{m} := \boldsymbol{m}(\phi, \nabla\phi, \boldsymbol{v}_{\mathrm{r}}, p, p_{\mathrm{h}}) = -\frac{1 - \phi}{D(\phi)}\boldsymbol{v}_{\mathrm{r}} + (p + p_{\mathrm{h}})\nabla\phi, \tag{2.8}$$

where $D := D(\phi)$ is the Darcy function, which depends on the drag related to the permeability of the material. The solid-fluid interaction force (2.8) is the result of combining both approaches

[45] (dynamic part) and [23] (hydrostatic part). Now, to obtain a constitutive relation for the relative velocity $\boldsymbol{v}_{\rm r}$, we follow the same idea as in [45]. Replacing $\boldsymbol{m}$ and (2.6) into equation (2.4) and neglecting the tensor $\boldsymbol{T}_{\rm f}$ (see [45] for further details), one arrives to the constitutive relation for $\boldsymbol{v}_{\rm r}$:

$$\boldsymbol{v}_{\rm r} = D(\phi)\nabla p.$$

Defining the volume-average velocity by

$$\boldsymbol{q} := \phi\boldsymbol{u} + (1-\phi)\boldsymbol{u}_{\rm f},$$

we obtain $\mathrm{div}(\boldsymbol{q}) = 0$, from which, replacing $\boldsymbol{u}_{\rm f} = \boldsymbol{u} - \boldsymbol{v}_{\rm r} = \boldsymbol{u} - D(\phi)\nabla p$, it follows that

$$\mathrm{div}(\boldsymbol{u} - \xi(\phi)\nabla p) = 0, \tag{2.9}$$

where $\xi(\phi) := (1-\phi)D(\phi)$. Finally, since $\boldsymbol{T}_{\rm f}$ was neglected, we have that $\boldsymbol{T}_{\rm mix} = \boldsymbol{T}_{\rm s}$, and we assume that $\boldsymbol{T}_{\rm s} := \boldsymbol{T}_{\rm s}(\phi, \boldsymbol{u})$ is given by

$$\boldsymbol{T}_{\rm s} := \mu(\phi)\boldsymbol{e}(\boldsymbol{u}), \quad \text{where} \quad \boldsymbol{e}(\boldsymbol{u}) := \tfrac{1}{2}(\nabla\boldsymbol{u} + \nabla\boldsymbol{u}^{\rm t}),$$

and $\mu := \mu(\phi)$ is the viscosity function of the mixture which is assumed to be a positive bounded function, i.e., there exist $\mu_{\rm min}$, $\mu_{\rm max} \in \mathbb{R}^+$, with $\mu_{\rm min} < \mu_{\rm max}$, such that

$$\mu_{\rm min} \leq \mu(\phi) \leq \mu_{\rm max} \qquad \forall\, \phi \in [0,1]\,. \tag{2.10}$$

Having all the ingredients described here, the main system of equations is

$$\partial_t\phi + \mathrm{div}(\phi\boldsymbol{u}) = 0 \quad \text{in } \Omega, \tag{2.11a}$$
$$-\mathbf{div}(\mu(\phi)\boldsymbol{e}(\boldsymbol{u})) + \nabla p = \mathbf{g}(\phi) \quad \text{in } \Omega, \tag{2.11b}$$
$$\mathrm{div}\,(\boldsymbol{u} - \xi(\phi)\nabla p) = 0 \quad \text{in } \Omega, \tag{2.11c}$$

where $\mathbf{g}(\phi) := -\rho_{\rm d} g\phi\boldsymbol{k} - \nabla\sigma_{\rm e}(\phi)$ and $\rho_{\rm d} = \rho_{\rm s} - \rho_{\rm f}$. The first equation is the conservation of mass of the solid phase (2.1), Equation (2.11b) is the momentum equation of the mixture (2.5) after replacing the variables and constitutive relations described before, and the third equation corresponds to (2.9). The unknowns of the system are the volume fraction of solid particles ϕ, the solid phase velocity $\boldsymbol{u}$ and the excess pore pressure p. This PDE system is complemented with suitable initial and boundary conditions.

3 The continuous formulation

In this section we derive a mixed variational formulation for the system (2.11b)–(2.11c) solved for the pair $\boldsymbol{u}$ and p, for a given function $\phi \in L^2(\Omega)$, and then apply a fixed-point approach along with the Babuška-Brezzi theory and Banach theorem to address its solvability.

3.1 Derivation of the mixed formulation

We begin by introducing some further spaces and related results to be employed in what follows. In particular, we will make use of the decompositions:

$$\mathbb{L}^2(\Omega) = \mathbb{L}^2_{\rm tr}(\Omega) \oplus \mathcal{P}_0(\Omega)\mathbb{I} \quad \text{and} \quad \mathbb{H}(\mathbf{div};\Omega) = \mathbb{H}_{\mathbf{tr}}(\mathbf{div};\Omega) \oplus \mathcal{P}_0(\Omega)\mathbb{I}, \tag{3.1}$$

where $\mathcal{P}_0(\Omega)$ is the set of constant polinomial functions over Ω,

$$\mathbb{L}^2_{\mathrm{tr}}(\Omega) := \left\{ \boldsymbol{\tau} \in \mathbb{L}^2 : \quad \mathrm{tr}(\boldsymbol{\tau}) = 0 \right\}, \quad \text{and}$$
$$\mathbb{H}_{\mathbf{tr}}(\mathbf{div};\Omega) := \left\{ \boldsymbol{\tau} \in \mathbb{H}(\mathbf{div};\Omega) : \quad \int_\Omega \mathrm{tr}(\boldsymbol{\tau}) = 0 \right\}.$$

Note that in light of the second decomposition in (3.1), any element $\boldsymbol{\tau} \in \mathbb{H}(\mathbf{div};\Omega)$ can be uniquely decomposed as $\boldsymbol{\tau} = \boldsymbol{\tau}_0 + c\mathbb{I}$, with $\boldsymbol{\tau} \in \mathbb{H}_{\mathbf{tr}}(\mathbf{div};\Omega)$ and $c \in \mathbb{R}$ computed by

$$c = \frac{1}{d|\Omega|} \int_\Omega \mathrm{tr}(\boldsymbol{\tau}), \tag{3.2}$$

where $|\Omega|$ is the measure of the domain Ω. We also define the following spaces:

$$L^2_0(\Omega) := \left\{ q \in L^2(\Omega) : \quad \int_\Omega q = 0 \right\} \quad \text{and}$$
$$\mathbb{L}^2_{\mathtt{skew}}(\Omega) := \left\{ \boldsymbol{\tau} \in \mathbb{L}^2(\Omega) : \quad \boldsymbol{\tau} = -\boldsymbol{\tau}^{\mathrm{t}} \right\}.$$

Given a function $\phi \in L^2(\Omega)$, the decoupled model equations (2.11b)-(2.11c) is combined with no-slip boundary conditions for $\boldsymbol{u}$ and Neumann boundary conditions for p, thus yielding

$$-\mathbf{div}(\mu(\phi)\boldsymbol{e}(\boldsymbol{u})) + \nabla p = \mathbf{g}(\phi) \quad \text{in } \Omega, \tag{3.3a}$$
$$\mathrm{div}(\boldsymbol{u} - \xi(\phi)\nabla p) = 0 \quad \text{in } \Omega, \tag{3.3b}$$
$$\boldsymbol{u} = \mathbf{0}, \quad \nabla p \cdot \boldsymbol{n} = 0 \quad \text{on } \Gamma. \tag{3.3c}$$

We are now going to treat (3.3a) and (3.3b) separately as PDE systems themselves, introducing extra new variables, but coupled through their source terms. Starting from Equation (3.3a), we introduce the pseudostress tensor $\widehat{\boldsymbol{\sigma}} := \mu(\phi)\boldsymbol{e}(\boldsymbol{u}) - p\mathbb{I}$ in Ω, which we seek to belong to $\mathbb{H}(\mathbf{div};\Omega)$. Considering the space decomposition (3.1), there exists $\boldsymbol{\sigma} \in \mathbb{H}_{\mathbf{tr}}(\mathbf{div};\Omega)$, such that $\widehat{\boldsymbol{\sigma}} = \boldsymbol{\sigma} + c_0\mathbb{I}$, where c_0 is determined by (3.2), that is

$$c_0 = \frac{1}{d|\Omega|} \int \mathrm{tr}(\widehat{\boldsymbol{\sigma}}).$$

Then, instead of searching for the pseudostress tensor we focus on the new variable $\boldsymbol{\sigma} = \widehat{\boldsymbol{\sigma}} - c_0\mathbb{I}$ which, according to (3.3a), satisfies

$$\mathbf{div}(\boldsymbol{\sigma}) = \mathbf{div}(\widehat{\boldsymbol{\sigma}}) = -\mathbf{g}(\phi). \tag{3.4}$$

Next, we define two new variables

$$\boldsymbol{t} := \left(\boldsymbol{e}(\boldsymbol{u})\right)^{\mathrm{d}} \quad \text{and the vorticity} \quad \boldsymbol{\gamma} := \frac{1}{2}\left(\nabla\boldsymbol{u} - \nabla\boldsymbol{u}^{\mathrm{t}}\right),$$

which are related as follows

$$\boldsymbol{t} = \nabla\boldsymbol{u} - \boldsymbol{\gamma} - \frac{1}{d}\mathrm{tr}(\boldsymbol{e}(\boldsymbol{u}))\mathbb{I}, \quad \text{and} \quad \boldsymbol{e}(\boldsymbol{u}) = \nabla\boldsymbol{u} - \boldsymbol{\gamma} = \boldsymbol{t} + \frac{1}{d}\mathrm{tr}(\boldsymbol{e}(\boldsymbol{u}))\mathbb{I}.$$

Note that $\mathrm{tr}(\boldsymbol{t}) = \mathrm{tr}(\boldsymbol{\gamma}) = 0$, whence $\boldsymbol{t} \in \mathbb{L}^2_{\mathrm{tr}}(\Omega)$. Then, equation (3.4) plus the relations between the new variables $\boldsymbol{t}$, $\boldsymbol{\gamma}$ and $\boldsymbol{\sigma}$ yield the following system of equations:

$$\boldsymbol{\sigma} = \mu(\phi)\left(\boldsymbol{t} + \frac{1}{d}\mathrm{tr}(\boldsymbol{e}(\boldsymbol{u}))\mathbb{I}\right) - \widetilde{p}\mathbb{I} \quad \text{in } \Omega, \tag{3.5a}$$
$$\boldsymbol{t} = \nabla\boldsymbol{u} - \boldsymbol{\gamma} - \frac{1}{d}\mathrm{tr}(\boldsymbol{e}(\boldsymbol{u}))\mathbb{I} \quad \text{in } \Omega, \tag{3.5b}$$
$$\mathbf{div}(\boldsymbol{\sigma}) = -\mathbf{g}(\phi) \quad \text{in } \Omega\,, \tag{3.5c}$$

where $\widetilde{p} := p + c_0$ is the new unknown pressure. The above system is complemented with an incompressibility hypothesis, which serves as uniqueness condition for $\widetilde{p}$, namely

$$\int_\Omega \widetilde{p} = 0. \tag{3.6}$$

System (3.5) needs to be solved together with the respective boundary condition from (3.3c), that is $\boldsymbol{u} = \boldsymbol{0}$ on Γ, and condition (3.6). Note that the pressure $\widetilde{p}$ can be computed in terms of the new variables $\boldsymbol{t}$ and $\boldsymbol{\sigma}$. In fact, taking trace on both sides of (3.5a), we get

$$\mathrm{tr}(\boldsymbol{\sigma}) = \mu(\phi)\mathrm{tr}(\boldsymbol{e}(\boldsymbol{u})) - d\widetilde{p}, \tag{3.7}$$

which yields

$$\widetilde{p} = \frac{1}{d}\mathrm{tr}\Big(\mu(\phi)\boldsymbol{e}(\boldsymbol{u}) - \boldsymbol{\sigma}\Big). \tag{3.8}$$

Furthermore, combining the obtained expression for $\widetilde{p}$ and (3.6) we obtain

$$\int_\Omega \widetilde{p} = \int_\Omega \frac{1}{d}\mathrm{tr}\Big(\mu(\phi)\boldsymbol{e}(\boldsymbol{u}) - \boldsymbol{\sigma}\Big) = 0,$$

which, along with the fact that $\boldsymbol{\sigma} \in \mathbb{H}_{\mathbf{tr}}(\mathbf{div};\Omega)$, implies

$$\int_\Omega \mu(\phi)\mathrm{tr}(\boldsymbol{e}(\boldsymbol{u})) = 0.$$

To derive the weak formulation of (3.5), we test (3.5a) with a tensor $\boldsymbol{s} \in \mathbb{L}^2_{\mathrm{tr}}(\Omega)$, which gives

$$\int_\Omega \mu(\phi)\boldsymbol{t} : \boldsymbol{s} - \int_\Omega \boldsymbol{\sigma}^{\mathrm{d}} : \boldsymbol{s} = 0 \qquad \forall \boldsymbol{s} \in \mathbb{L}^2_{\mathrm{tr}}(\Omega). \tag{3.9}$$

Observe that since $\boldsymbol{s} = \boldsymbol{s}^{\mathrm{d}}$, there holds $\boldsymbol{\sigma} : \boldsymbol{s} = \boldsymbol{\sigma} : \boldsymbol{s}^{\mathrm{d}} = \boldsymbol{\sigma}^{\mathrm{d}} : \boldsymbol{s}$. On the other hand, equation (3.5b) can be multiplied by a test function $\boldsymbol{\tau} \in \mathbb{H}_{\mathbf{tr}}(\mathbf{div};\Omega)$ and integrated over Ω to obtain

$$\begin{aligned}\int_\Omega \boldsymbol{\tau} : \boldsymbol{t} &= \int_\Omega \boldsymbol{\tau} : \nabla \boldsymbol{u} - \int_\Omega \boldsymbol{\tau} : \boldsymbol{\gamma} - \int_\Omega \frac{1}{d}\boldsymbol{\tau} : \mathrm{tr}(\boldsymbol{e}(\boldsymbol{u}))\mathbb{I} \\ &= -\int_\Omega \boldsymbol{u} \cdot \mathbf{div}(\boldsymbol{\tau}) + \langle \boldsymbol{\tau}\boldsymbol{n}, \boldsymbol{u}\rangle_\Gamma - \int_\Omega \boldsymbol{\tau} : \boldsymbol{\gamma} - \int_\Omega \frac{1}{d}\mathrm{tr}(\boldsymbol{\tau})\mathrm{tr}(\boldsymbol{e}(\boldsymbol{u})),\end{aligned}$$

where $\langle \cdot, \cdot \rangle_\Gamma$ denotes the duality pairing between $\mathbf{H}^{-1/2}(\Gamma)$ and $\mathbf{H}^{1/2}(\Gamma)$. For the last term on the right hand side of the above equation, we notice that from (3.8)

$$\mathrm{tr}(\boldsymbol{e}(\boldsymbol{u})) = (\mathrm{tr}(\boldsymbol{\sigma}) + d\widetilde{p})/(\mu(\phi)),$$

so that we get

$$\int_\Omega \frac{1}{d}\mathrm{tr}(\boldsymbol{\tau})\mathrm{tr}(\boldsymbol{e}(\boldsymbol{u})) = \int_\Omega \frac{\mathrm{tr}(\boldsymbol{\tau})\mathrm{tr}(\boldsymbol{\sigma})}{d\mu(\phi)} + \int_\Omega \frac{\widetilde{p}\mathrm{tr}(\boldsymbol{\tau})}{\mu(\phi)}.$$

Then, using the boundary condition $\boldsymbol{u} = \boldsymbol{0}$ on Γ, we arrive at the weak form of (3.5b):

$$-\int_\Omega \boldsymbol{\tau} : \boldsymbol{t} - \int_\Omega \boldsymbol{u} \cdot \mathbf{div}(\boldsymbol{\tau}) - \int_\Omega \boldsymbol{\gamma} : \boldsymbol{\tau} - \int_\Omega \frac{\mathrm{tr}(\boldsymbol{\sigma})\,\mathrm{tr}(\boldsymbol{\tau})}{d\mu(\phi)} = \int_\Omega \frac{\widetilde{p}\,\mathrm{tr}(\boldsymbol{\tau})}{\mu(\phi)} \quad \forall \boldsymbol{\tau} \in \mathbb{H}_{\mathbf{tr}}(\mathbf{div};\Omega). \tag{3.10}$$

The third equation (3.5c) is simply tested against $\boldsymbol{w} \in \mathbf{L}^2(\Omega)$, that is

$$-\int_\Omega \boldsymbol{w} \cdot \mathbf{div}(\boldsymbol{\sigma}) = \int_\Omega \mathbf{g}(\phi) \cdot \boldsymbol{w}, \qquad \forall \boldsymbol{w} \in \mathbf{L}^2(\Omega). \tag{3.11}$$

In addition, because of the antisymmetry of $\boldsymbol{\gamma}$, this function is searched in $\mathbb{L}^2_{\mathrm{skew}}(\Omega)$ and the symmetry of $\boldsymbol{\sigma}$ is weakly enforced by

$$\int_\Omega \boldsymbol{\sigma} : \boldsymbol{\eta} = 0 \qquad \forall \boldsymbol{\eta} \in \mathbb{L}^2_{\mathrm{skew}}(\Omega), \tag{3.12}$$

Finally, gathering (3.12) and (3.11), we can write

$$-\int_\Omega \boldsymbol{w} \cdot \mathbf{div}(\boldsymbol{\sigma}) - \int_\Omega \boldsymbol{\eta} : \boldsymbol{\sigma} = \int_\Omega \mathbf{g}(\phi) \cdot \boldsymbol{w} \qquad \forall \boldsymbol{w} \in \mathbf{L}^2(\Omega), \quad \forall \boldsymbol{\eta} \in \mathbb{L}^2_{\mathrm{skew}}(\Omega). \tag{3.13}$$

The equations (3.9), (3.10) and (3.13) constitute the mixed variational formulation of system (3.5), and if we assume that $\widetilde{p}$ is known, the unknowns of this system are $\boldsymbol{t}$, $\boldsymbol{\sigma}$, $\boldsymbol{u}$ and $\boldsymbol{\gamma}$, which will be searched by pairs on the spaces:

$$\mathbb{X} := \mathbb{L}^2_{\mathrm{tr}}(\Omega) \times \mathbb{H}_{\mathbf{tr}}(\mathbf{div}; \Omega) \quad \text{and} \quad \mathbb{V} := \mathbf{L}^2(\Omega) \times \mathbb{L}^2_{\mathrm{skew}}(\Omega).$$

In this way, given $\widetilde{p} \in \mathbb{L}^2_0(\Omega)$ and $\phi \in L^2(\Omega)$, the first problem to be solved is: Find $((\boldsymbol{t}, \boldsymbol{\sigma}), (\boldsymbol{u}, \boldsymbol{\gamma})) \in \mathbb{X} \times \mathbb{V}$ such that

$$\begin{aligned}
\mathbf{a}_\phi(\boldsymbol{t}, \boldsymbol{s}) + \mathbf{b}_1(\boldsymbol{s}, \boldsymbol{\sigma}) &= 0 && \forall \boldsymbol{s} \in \mathbb{L}^2_{\mathrm{tr}}(\Omega), && \text{(3.14a)}\\
\mathbf{b}_1(\boldsymbol{t}, \boldsymbol{\tau}) - \mathbf{c}_\phi(\boldsymbol{\sigma}, \boldsymbol{\tau}) + \mathbf{b}_2(\boldsymbol{\tau}, (\boldsymbol{u}, \boldsymbol{\gamma})) &= \boldsymbol{F}_{\phi,\widetilde{p}}(\boldsymbol{\tau}) && \forall \boldsymbol{\tau} \in \mathbb{H}_{\mathbf{tr}}(\mathbf{div}; \Omega), && \text{(3.14b)}\\
\mathbf{b}_2(\boldsymbol{\sigma}, (\boldsymbol{w}, \boldsymbol{\eta})) &= \boldsymbol{G}_\phi(\boldsymbol{w}, \boldsymbol{\eta}) && \forall (\boldsymbol{w}, \boldsymbol{\eta}) \in \mathbb{V}, && \text{(3.14c)}
\end{aligned}$$

where the bilinear forms and functionals

$$\begin{aligned}
&\mathbf{a}_\phi : \mathbb{L}^2_{\mathrm{tr}}(\Omega) \times \mathbb{L}^2_{\mathrm{tr}}(\Omega) \to \mathbb{R}, && \mathbf{b}_1 : \mathbb{L}^2_{\mathrm{tr}}(\Omega) \times \mathbb{H}_{\mathbf{tr}}(\mathbf{div}; \Omega) \to \mathbb{R},\\
&\mathbf{c}_\phi : \mathbb{H}_{\mathbf{tr}}(\mathbf{div}; \Omega) \times \mathbb{H}_{\mathbf{tr}}(\mathbf{div}; \Omega) \to \mathbb{R}, && \mathbf{b}_2 : \mathbb{H}_{\mathbf{tr}}(\mathbf{div}; \Omega) \times \mathbb{V} \to \mathbb{R},\\
&\boldsymbol{F}_{\phi,\widetilde{p}} : \mathbb{H}_{\mathbf{tr}}(\mathbf{div}; \Omega) \to \mathbb{R}, && \boldsymbol{G}_\phi : \mathbb{V} \to \mathbb{R},
\end{aligned}$$

are defined by

$$\begin{aligned}
\mathbf{a}_\phi(\boldsymbol{t}, \boldsymbol{s}) &:= \int_\Omega \mu(\phi) \boldsymbol{t} : \boldsymbol{s}, \quad \mathbf{b}_1(\boldsymbol{s}, \boldsymbol{\sigma}) := -\int_\Omega \boldsymbol{s} : \boldsymbol{\sigma}^{\mathrm{d}}, \quad \mathbf{c}_\phi(\boldsymbol{\sigma}, \boldsymbol{\tau}) := \int_\Omega \frac{\mathrm{tr}(\boldsymbol{\sigma})\,\mathrm{tr}(\boldsymbol{\tau})}{d\mu(\phi)},\\
\mathbf{b}_2(\boldsymbol{\sigma}, (\boldsymbol{w}, \boldsymbol{\eta})) &:= -\int_\Omega \boldsymbol{w} \cdot \mathbf{div}(\boldsymbol{\sigma}) - \boldsymbol{\eta} : \boldsymbol{\sigma}, \quad \boldsymbol{F}_{\phi,\widetilde{p}}(\boldsymbol{\tau}) := \int_\Omega \frac{\widetilde{p}\,\mathrm{tr}(\boldsymbol{\tau})}{\mu(\phi)},\\
\boldsymbol{G}_\phi(\boldsymbol{w}, \boldsymbol{\eta}) &:= \int_\Omega \mathbf{g}(\phi) \cdot \boldsymbol{w}\,.
\end{aligned} \tag{3.15}$$

On the other hand, for the weak formulation of (3.3b) we proceed in a similar fashion as for (3.3a) by defining an extra variable $\widetilde{\boldsymbol{\sigma}} := \xi(\phi)\nabla\widetilde{p} - \boldsymbol{u}$, with which we have the system of equations

$$\begin{aligned}
\widetilde{\boldsymbol{\sigma}} &= \xi(\phi)\nabla\widetilde{p} - \boldsymbol{u} && \text{in } \Omega, && \text{(3.16a)}\\
\mathrm{div}(\widetilde{\boldsymbol{\sigma}}) &= 0, && \text{in } \Omega. && \text{(3.16b)}
\end{aligned}$$

This system is complemented with the boundary condition $\widetilde{\boldsymbol{\sigma}} \cdot \boldsymbol{n} = (\xi(\phi)\nabla\widetilde{p} - \boldsymbol{u}) \cdot \boldsymbol{n} = 0$ on Γ which follows from the second equation in (3.3c). Dividing (3.16a) by $\xi(\phi)$, multiplying by a test function $\widetilde{\boldsymbol{\tau}} \in \mathbf{H}(\mathrm{div}; \Omega)$, and integrating we get

$$\int_\Omega \frac{1}{\xi(\phi)} \widetilde{\boldsymbol{\sigma}} \cdot \widetilde{\boldsymbol{\tau}} = -\int_\Omega \widetilde{p}\,\mathrm{div}(\widetilde{\boldsymbol{\tau}}) + \langle \widetilde{\boldsymbol{\tau}} \cdot \boldsymbol{n}, \widetilde{p} \rangle - \int_\Omega \frac{1}{\xi(\phi)} \boldsymbol{u} \cdot \widetilde{\boldsymbol{\tau}}.$$

Because of the above boundary condition for $\widetilde{\boldsymbol{\sigma}}$, we impose $\widetilde{\boldsymbol{\tau}} \cdot \boldsymbol{n} = 0$ on Γ so that the second term in the right hand side vanishes, as well the space of test functions becomes

$$\mathbf{H}_0(\mathrm{div};\Omega) = \left\{ \widetilde{\boldsymbol{\tau}} \in \mathbf{H}(\mathrm{div};\Omega) : \quad \widetilde{\boldsymbol{\tau}} \cdot \boldsymbol{n} = 0 \text{ on } \Gamma \right\},$$

and the vectorial spaces to search the unknowns are

$$\mathbf{M} := \mathbf{H}_0(\mathrm{div};\Omega), \qquad Q := L^2_0(\Omega).$$

Hence, given $\boldsymbol{u} \in \mathbf{L}^2(\Omega)$ and $\phi \in L^2(\Omega)$, the second problem reduces to: Find $(\widetilde{\boldsymbol{\sigma}}, \widetilde{p}) \in \mathbf{M} \times Q$ such that

$$\widetilde{\mathbf{a}}_\phi(\widetilde{\boldsymbol{\sigma}}, \widetilde{\boldsymbol{\tau}}) + \widetilde{\mathbf{b}}(\widetilde{p}, \widetilde{\boldsymbol{\tau}}) = \widetilde{\boldsymbol{F}}_{\phi,\boldsymbol{u}}(\widetilde{\boldsymbol{\tau}}) \quad \forall\, \widetilde{\boldsymbol{\tau}} \in \mathbf{M}, \tag{3.17a}$$

$$\widetilde{\mathbf{b}}(q, \widetilde{\boldsymbol{\sigma}}) = 0 \qquad\qquad \forall\, q \in Q, \tag{3.17b}$$

where $\widetilde{\mathbf{a}}_\phi : \mathbf{M} \times \mathbf{M} \to \mathbb{R}$, $\widetilde{\mathbf{b}} : Q \times \mathbf{M} \to \mathbb{R}$ and $\widetilde{\boldsymbol{F}}_{\phi,\boldsymbol{u}} : \mathbf{M} \to \mathbb{R}$ are defined by

$$\widetilde{\mathbf{a}}_\phi(\widetilde{\boldsymbol{\sigma}}, \widetilde{\boldsymbol{\tau}}) = \int_\Omega \frac{1}{\xi(\phi)} \widetilde{\boldsymbol{\sigma}} \cdot \widetilde{\boldsymbol{\tau}}, \qquad \widetilde{\mathbf{b}}(\widetilde{p}, \widetilde{\boldsymbol{\tau}}) := \int_\Omega \widetilde{p}\, \mathrm{div}(\widetilde{\boldsymbol{\tau}}), \qquad \widetilde{\boldsymbol{F}}_{\phi,\boldsymbol{u}}(\widetilde{\boldsymbol{\tau}}) = -\int_\Omega \frac{1}{\xi(\phi)} \boldsymbol{u} \cdot \widetilde{\boldsymbol{\tau}}.$$

3.2 The fixed-point estrategy

Observe that the system (3.14) needs a given pressure $\widetilde{p} \in L^2_0(\Omega)$, which can be obtained from (3.17), while the system (3.17) needs a given velocity $\boldsymbol{u} \in \boldsymbol{L}^2(\Omega)$, which can be found by solving system (3.14). This coupled nature between the two systems of equations suggests the fixed-point approach to be explained next. In fact, we first let $\mathbf{S}_\phi : Q \to \mathbf{L}^2(\Omega)$ be the operator defined for each fixed function $\phi \in L^2(\Omega)$ by

$$\mathbf{S}_\phi(\widetilde{p}) := \boldsymbol{u} \quad \forall \widetilde{p} \in Q, \tag{3.18}$$

where $((\boldsymbol{t}, \boldsymbol{\sigma}), (\boldsymbol{u}, \boldsymbol{\gamma})) \in \mathbb{X} \times \mathbb{V}$ is the solution of (3.14) for the given $\widetilde{p} \in Q$ and $\phi \in L^2(\Omega)$. In turn, we let $\widetilde{\mathbf{S}}_\phi : \mathbf{L}^2(\Omega) \to Q$ be the operator defined for each fixed function $\phi \in L^2(\Omega)$ by

$$\widetilde{\mathbf{S}}_\phi(\boldsymbol{u}) := \widetilde{p} \quad \forall \boldsymbol{u} \in \mathbf{L}^2(\Omega),$$

where $(\widetilde{\boldsymbol{\sigma}}, \widetilde{p}) \in \mathbf{M} \times Q$ is the solution of (3.17) for the given $\boldsymbol{u} \in \mathbf{L}^2(\Omega)$ and $\phi \in L^2(\Omega)$. The well-posedness of $\mathbf{S}_\phi$ and $\widetilde{\mathbf{S}}_\phi$, or equivalently the uniqueness of solutions of both systems (3.14) and (3.17), respectively, is shown next in Section 3.3. Now, defining the operator $\mathbf{T}_\phi = \widetilde{\mathbf{S}}_\phi \circ \mathbf{S}_\phi : Q \to Q$, we readily realize that solving the coupled problem (3.14)–(3.17) is equivalent to finding a fixed point of $\mathbf{T}_\phi$, that is: Find $\widetilde{p} \in Q$ such that

$$\mathbf{T}_\phi(\widetilde{p}) = \widetilde{p}. \tag{3.19}$$

We end this section by remarking that the linearity of the operators $\mathbf{S}_\phi$ and $\widetilde{\mathbf{S}}_\phi$, which follows from the corresponding properties of $\boldsymbol{F}_{\phi,\widetilde{p}}$ and $\widetilde{\boldsymbol{F}}_{\phi,\boldsymbol{u}}$, guarantees that $\mathbf{T}_\phi$ is linear as well.

3.3 Solvability of the decoupled problems

In this section we analyze the existence and uniqueness of solutions for the systems (3.14) and (3.17), respectively. We begin with (3.14) by letting $\boldsymbol{\mathcal{A}}_\phi : \mathbb{L}^2_{\mathrm{tr}}(\Omega) \to \mathbb{L}^2_{\mathrm{tr}}(\Omega)'$, $\boldsymbol{\mathcal{B}}_1 : \mathbb{L}^2_{\mathrm{tr}}(\Omega) \to \mathbb{H}_{\mathbf{tr}}(\mathbf{div};\Omega)'$, $\boldsymbol{\mathcal{B}}_2 : \mathbb{H}_{\mathbf{tr}}(\mathbf{div};\Omega) \to \mathbb{V}'$, and $\boldsymbol{\mathcal{C}}_\phi : \mathbb{H}_{\mathbf{tr}}(\mathbf{div};\Omega) \to \mathbb{H}_{\mathbf{tr}}(\mathbf{div};\Omega)'$ be the operators induced by the bilinear forms

$\mathbf{a}_\phi$, $\mathbf{b}_1$, $\mathbf{b}_2$ and $\mathbf{c}_\phi$, respectively, and denote by $\boldsymbol{\mathcal{B}}_1^* : \mathbb{H}_{\mathbf{tr}}(\mathbf{div};\Omega) \to \mathbb{L}^2_{\mathrm{tr}}(\Omega)'$ and $\boldsymbol{\mathcal{B}}_2^* : \mathbb{V} \to \mathbb{H}_{\mathbf{tr}}(\mathbf{div};\Omega)'$ the corresponding "adjoints" of $\boldsymbol{\mathcal{B}}_1$ and $\boldsymbol{\mathcal{B}}_2$. Then, (3.14) can be rewritten, equivalently, as: Find $((\boldsymbol{t},\boldsymbol{\sigma}),(\boldsymbol{u},\boldsymbol{\gamma})) \in \mathbb{X} \times \mathbb{V}$ such that

$$\begin{aligned} [\boldsymbol{\mathcal{A}}_\phi(\boldsymbol{t}),\boldsymbol{s}] + [\boldsymbol{\mathcal{B}}_1^*(\boldsymbol{\sigma}),\boldsymbol{s}] &= 0 && \forall \boldsymbol{s} \in \mathbb{L}^2_{\mathrm{tr}}(\Omega), \\ [\boldsymbol{\mathcal{B}}_1(\boldsymbol{t}),\boldsymbol{\tau}] - [\boldsymbol{\mathcal{C}}_\phi(\boldsymbol{\sigma}),\boldsymbol{\tau}] + [\boldsymbol{\mathcal{B}}_2^*(\boldsymbol{u},\boldsymbol{\gamma}),\boldsymbol{\tau}] &= [\boldsymbol{F}_{\phi,\widetilde{p}},\boldsymbol{\tau}] && \forall \boldsymbol{\tau} \in \mathbb{H}_{\mathbf{tr}}(\mathbf{div};\Omega), \\ [\boldsymbol{\mathcal{B}}_2(\boldsymbol{\sigma}),(\boldsymbol{w},\boldsymbol{\eta})] &= [\boldsymbol{G}_\phi,(\boldsymbol{w},\boldsymbol{\eta})] && \forall (\boldsymbol{w},\boldsymbol{\eta}) \in \mathbb{V}. \end{aligned} \tag{3.20}$$

where $[\cdot\,,\cdot]$ denotes the duality pairing induced by the operators and functionals used in each case. Note that (3.20) has the same structure of [43, eq. (1.2)] with $\boldsymbol{\mathcal{A}}_\phi$ linear, and hence in what follows we apply the theory developed there to address its solvability. More precisely, we have the following result.

Lemma 3.1. *Given $\widetilde{p} \in Q$ and $\phi \in L^2(\Omega)$, there exists a unique $((\boldsymbol{t},\boldsymbol{\sigma}),(\boldsymbol{u},\boldsymbol{\gamma})) \in \mathbb{X} \times \mathbb{V}$ solution of (3.14), and hence one can define $\mathbf{S}_\phi(\widetilde{p}) := \boldsymbol{u}$. Moreover, there exists a positive constant $\beta > 0$, depending on d, $\mu_{\min}$, and $\mu_{\max}$, such that*

$$\|\mathbf{S}_\phi(\widetilde{p})\|_{0;\Omega} = \|\boldsymbol{u}\|_{0;\Omega} \,\leq\, \|((\boldsymbol{t},\boldsymbol{\sigma}),(\boldsymbol{u},\boldsymbol{\gamma}))\|_{\mathbb{X}\times\mathbb{V}} \,\leq\, \beta \left\{ \|(\mu(\phi))^{-1}\widetilde{p}\|_{0;\Omega} + \|\mathbf{g}(\phi)\|_{0;\Omega} \right\}. \tag{3.21}$$

Proof. It reduces to prove the hypotheses of the abstract result given by [43, Theorem 2.1]. To do so, we begin by noticing from (2.10) that

$$\|\boldsymbol{\mathcal{A}}_\phi\| \leq \mu_{\max}, \quad \|\boldsymbol{\mathcal{B}}_1\| \leq 1, \quad \|\boldsymbol{\mathcal{C}}_\phi\| \leq \frac{1}{\mu_{\min}},$$

and since $\boldsymbol{\mathcal{B}}_2(\boldsymbol{\tau}) := (\mathbf{div}(\boldsymbol{\tau}), \frac{1}{2}(\boldsymbol{\tau} - \boldsymbol{\tau}^{\mathrm{t}}))$, it follows that $\|\boldsymbol{\mathcal{B}}_2\| \leq 1$ and

$$\mathrm{kernel}(\boldsymbol{\mathcal{B}}_2) := \left\{ \boldsymbol{\tau} \in \mathbb{H}_{\mathbf{tr}}(\mathbf{div};\Omega) : \quad \mathbf{div}(\boldsymbol{\tau}) = \mathbf{0}, \quad \boldsymbol{\tau} = \boldsymbol{\tau}^{\mathrm{t}} \right\}.$$

Now we show the properties of the operators $\boldsymbol{\mathcal{A}}_\phi$, $\boldsymbol{\mathcal{B}}_1$, $\boldsymbol{\mathcal{C}}_\phi$ and $\boldsymbol{\mathcal{B}}_2$ needed to apply [43, Theorem 2.1]. The boundedness and linearity of $\boldsymbol{\mathcal{A}}_\phi$ certainly implies its Lipschitz continuity. Furthermore, $\mu(\phi) > \mu_{\min}$ (cf. (2.10)) yields

$$\mathbf{a}_\phi(\boldsymbol{s},\boldsymbol{s}) \,\geq\, \mu_{\min}\, \|\boldsymbol{s}\|^2_{0,\Omega} \qquad \forall\, \boldsymbol{s} \in \mathbb{L}^2_{\mathrm{tr}}(\Omega)\,,$$

which says that $\mathbf{a}_\phi$ is $\mathbb{L}^2_{\mathrm{tr}}(\Omega)$-elliptic, and hence $\boldsymbol{\mathcal{A}}_\phi$ is clearly strongly monotone, that is

$$[\boldsymbol{\mathcal{A}}_\phi(\boldsymbol{s}) - \boldsymbol{\mathcal{A}}_\phi(\widetilde{\boldsymbol{s}}), \boldsymbol{s} - \widetilde{\boldsymbol{s}}] = \mathbf{a}_\phi(\boldsymbol{s} - \widetilde{\boldsymbol{s}}, \boldsymbol{s} - \widetilde{\boldsymbol{s}}) \,\geq\, \mu_{\min}\, \|\boldsymbol{s} - \widetilde{\boldsymbol{s}}\|^2_{0;\Omega} \quad \forall\, \boldsymbol{s},\, \widetilde{\boldsymbol{s}} \in \mathbb{L}^2_{\mathrm{tr}}(\Omega)\,.$$

On the other hand, we have

$$[\boldsymbol{\mathcal{C}}_\phi(\boldsymbol{\tau}),\boldsymbol{\tau}] = \int_\Omega \frac{(\mathrm{tr}(\boldsymbol{\tau}))^2}{d\mu(\phi)} \geq 0 \qquad \forall \boldsymbol{\tau} \in \mathbb{H}(\mathbf{div};\Omega),$$

which shows that $\boldsymbol{\mathcal{C}}_\phi$ is positive semi-definite in $\mathbb{H}(\mathbf{div};\Omega)$ and in particular in $\mathrm{kernel}(\boldsymbol{\mathcal{B}}_2)$. In addition, the operator $\boldsymbol{\mathcal{B}}_1$ satisfies the following inf-sup condition

$$\sup_{\substack{\boldsymbol{s} \in \mathbb{L}^2_{\mathrm{tr}}(\Omega) \\ \boldsymbol{s} \neq \mathbf{0}}} \frac{[\boldsymbol{\mathcal{B}}_1(\boldsymbol{s}),\boldsymbol{\tau}]}{\|\boldsymbol{s}\|_{0;\Omega}} \geq \beta_1 \|\boldsymbol{\tau}\|_{\mathbf{div};\Omega}, \quad \forall \boldsymbol{\tau} \in \mathrm{kernel}(\boldsymbol{\mathcal{B}}_2)\,,$$

which follows from the inequality $\|\boldsymbol{\tau}^{\mathrm{d}}\|_{0;\Omega} \geq \beta_0 \|\boldsymbol{\tau}\|_{0;\Omega} = \beta_0 \|\boldsymbol{\tau}\|_{\mathbf{div};\Omega}$ for $\boldsymbol{\tau} \in \mathbb{H}_{\mathbf{tr}}(\mathbf{div};\Omega)$ such that $\mathbf{div}(\boldsymbol{\tau}) = \mathbf{0}$, see [42, Lemma 2.3]. Finally, to show the inf-sup condition of $\boldsymbol{\mathcal{B}}_2$ we use the same proof presented in [42]. More precisely, given $(\boldsymbol{w},\boldsymbol{\eta}) \in \mathbb{V}$, we consider the auxiliary boundary value problem:

$$\mathbf{div}(\boldsymbol{e}(\boldsymbol{z}) + \boldsymbol{\eta}) = \boldsymbol{w} \quad \text{in} \quad \Omega, \quad \boldsymbol{z} = \mathbf{0} \quad \text{on} \quad \Gamma\,,$$

which has a unique solution $\boldsymbol{z} \in \mathbf{H}^1_0(\Omega)$ that satisfies

$$|\boldsymbol{z}|_{0;\Omega} \,\leq\, C\Big\{\|\boldsymbol{w}\|_{0;\Omega} + \|\boldsymbol{\eta}\|_{0;\Omega}\Big\}$$

for a positive real constant C. Next, we define $\widehat{\boldsymbol{\tau}} := \boldsymbol{e}(\boldsymbol{z}) + \boldsymbol{\eta}$, so that $\mathbf{div}(\widehat{\boldsymbol{\tau}}) = \boldsymbol{w}$ and thus $\widehat{\boldsymbol{\tau}} \in \mathbb{H}(\mathbf{div};\Omega)$. Then, if $\boldsymbol{\tau}_0$ is the $\mathbb{H}_{\mathbf{tr}}(\mathbf{div};\Omega)$-component of $\widehat{\boldsymbol{\tau}}$, we deduce that $[\boldsymbol{\mathcal{B}}_2(\boldsymbol{\tau}_0),(\boldsymbol{w},\boldsymbol{\eta})] = \|(\boldsymbol{w},\boldsymbol{\eta})\|^2$, and

$$\|\boldsymbol{\tau}_0\|^2_{\mathbf{div};\Omega} \leq \|\widehat{\boldsymbol{\tau}}\|^2_{\mathbf{div};\Omega} = \|\boldsymbol{e}(\boldsymbol{z})\|^2_{0;\Omega} + \|\boldsymbol{w}\|^2_{0;\Omega} + \|\boldsymbol{\eta}\|^2_{0;\Omega} \leq (1+2C^2)\|(\boldsymbol{w},\boldsymbol{\eta})\|^2_{\mathbb{V}}.$$

Hence, for all $(\boldsymbol{w},\boldsymbol{\eta}) \in \mathbb{V}$ we have

$$\sup_{\substack{\boldsymbol{\tau}\in\mathbb{H}_{\mathbf{tr}}(\mathbf{div};\Omega)\\ \boldsymbol{\tau}\neq\mathbf{0}}} \frac{[\boldsymbol{\mathcal{B}}_2(\boldsymbol{\tau}),(\boldsymbol{w},\boldsymbol{\eta})]}{\|\boldsymbol{\tau}\|_{\mathbf{div};\Omega}} \geq \frac{[\boldsymbol{\mathcal{B}}_2(\boldsymbol{\tau}_0),(\boldsymbol{w},\boldsymbol{\eta})]}{\|\boldsymbol{\tau}_0\|_{\mathbf{div};\Omega}} = \frac{\|(\boldsymbol{w},\boldsymbol{\eta})\|^2_{\mathbb{V}}}{\|\boldsymbol{\tau}_0\|_{\mathbf{div};\Omega}} \geq \beta_2\|(\boldsymbol{w},\boldsymbol{\eta})\|_{\mathbb{V}},$$

where $\beta_2 = 1/(1+2C^2)^{1/2}$. Hence, having the system (3.20) satisfied the hypotheses of [43, Theorem 2.1], we conclude that (3.14), equivalently (3.18), has a unique solution $((\boldsymbol{t},\boldsymbol{\sigma}),(\boldsymbol{u},\boldsymbol{\gamma})) \in \mathbb{X}\times\mathbb{V}$, and there exists a positive constant $\bar{\beta}$ that depends on the norm of $\boldsymbol{\mathcal{A}}_\phi$, $\boldsymbol{\mathcal{C}}_\phi$, $\boldsymbol{F}_{\phi,\widetilde{p}}$, β_1 and β_2, that is on β_1, β_2, $\mu_{\min}$, and $\mu_{\max}$, such that

$$\|((\boldsymbol{t},\boldsymbol{\sigma}),(\boldsymbol{u},\boldsymbol{\gamma}))\| \,\leq\, \bar{\beta}\Big\{\|\boldsymbol{F}_{\phi,\widetilde{p}}\| + \|\boldsymbol{G}_\phi\|\Big\}.$$

Finally, using from the definitions of $\boldsymbol{F}_{\phi,\widetilde{p}}$ and $\boldsymbol{G}_\phi$ (cf. (3.15)) that

$$\|\boldsymbol{F}_{\phi,\widetilde{p}}\| \,\leq\, \sqrt{d}\,\|(\mu(\phi))^{-1}\,\widetilde{p}\|_{0,\Omega} \quad\text{and}\quad \|\boldsymbol{G}_\phi\| \,\leq\, \|\mathbf{g}(\phi)\|_{0;\Omega}\,, \tag{3.22}$$

we arrive at (3.21) and finish the proof. □

From now on for the second system (3.17), we assume that $\phi_{\max} < 1$ and that there exist $\xi_{\min}$, $\xi_{\max} \in \mathbb{R}^+$ such that for $\phi < \phi_{\max}$

$$\xi_{\min} \,\leq\, \xi(\phi) \,\leq\, \xi_{\max}\,. \tag{3.23}$$

The following lemma employs the classical Babuška-Brezzi theory to show the existence and uniqueness of solutions for (3.17) in the space $\mathbf{M}\times Q$.

Lemma 3.2. *Given $\boldsymbol{u} \in \mathbf{L}^2(\Omega)$ and $\phi \in L^2(\Omega)$ such that $\phi \leq \phi_{\max}$ in Ω, there exists a unique $(\widetilde{\boldsymbol{\sigma}},\widetilde{p}) \in \mathbf{M}\times Q$ solution of (3.17), and hence one can define $\widetilde{\mathbf{S}}_\phi(\boldsymbol{u}) := \widetilde{p}$. Moreover, there exists $\widetilde{\beta} > 0$, depending on $\xi_{\min}$ and $\xi_{\max}$, such that*

$$\|\widetilde{\mathbf{S}}_\phi(\boldsymbol{u})\| = \|\widetilde{p}\| \leq \|(\widetilde{\boldsymbol{\sigma}},\widetilde{p})\|_{\mathbf{M}\times Q} \leq \widetilde{\beta}\|(\xi(\phi))^{-1}\boldsymbol{u}\|_{0;\Omega}.$$

Proof. We begin by observing that both bilinear forms $\widetilde{\mathbf{a}}_\phi$ and $\widetilde{\mathbf{b}}$, and $\widetilde{\boldsymbol{F}}_{\phi,\boldsymbol{u}}$ are bounded, namely:

$$\begin{aligned}
|\widetilde{\mathbf{a}}_\phi(\widetilde{\boldsymbol{\tau}},\widetilde{\boldsymbol{\sigma}})| &\leq \|(\xi(\phi))^{-1}\,\widetilde{\boldsymbol{\tau}}\|_{\mathrm{div};\Omega}\|\widetilde{\boldsymbol{\sigma}}\|_{\mathrm{div};\Omega} && \forall\,\widetilde{\boldsymbol{\tau}},\widetilde{\boldsymbol{\sigma}} \in \mathbf{M}\,,\\
|\widetilde{\mathbf{b}}(q,\widetilde{\boldsymbol{\tau}})| &\leq \|q\|_{0;\Omega}\|\widetilde{\boldsymbol{\tau}}\|_{\mathrm{div};\Omega} && \forall\,(q,\widetilde{\boldsymbol{\tau}}) \in Q\times\mathbf{M}\,,\quad\text{and}\\
|\widetilde{\boldsymbol{F}}_{\phi,\boldsymbol{u}}(\widetilde{\boldsymbol{\tau}})| &\leq \|(\xi(\phi))^{-1}\,\boldsymbol{u}\|_{0;\Omega}\,\|\widetilde{\boldsymbol{\tau}}\|_{\mathrm{div};\Omega} && \forall\,\widetilde{\boldsymbol{\tau}} \in \mathbf{M}\,.
\end{aligned} \tag{3.24}$$

Next, we find that $\mathrm{kernel}(\widetilde{\mathbf{b}}) = \big\{\widetilde{\boldsymbol{\tau}} \in M : \quad \mathrm{div}(\widetilde{\boldsymbol{\tau}}) = 0\big\}$, so that, using (3.23), for each $\widetilde{\boldsymbol{\tau}} \in \mathrm{kernel}(\widetilde{\boldsymbol{b}})$ there holds

$$\widetilde{\mathbf{a}}_\phi(\widetilde{\boldsymbol{\tau}},\widetilde{\boldsymbol{\tau}}) \geq \frac{1}{\xi_{\max}}\|\widetilde{\boldsymbol{\tau}}\|^2_{0;\Omega} = \frac{1}{\xi_{\max}}\|\widetilde{\boldsymbol{\tau}}\|^2_{\mathrm{div};\Omega},$$

which shows that $\widetilde{\mathbf{a}}_\phi$ is elliptic in kernel$(\widetilde{\mathbf{b}})$. On the other hand, the bilinear form $\widetilde{\boldsymbol{b}}$ satisfies the following inf-sup condition:

$$\sup_{\substack{\widetilde{\boldsymbol{\tau}}\in\mathbf{M}\\ \widetilde{\boldsymbol{\tau}}\neq\mathbf{0}}} \frac{\widetilde{\boldsymbol{b}}(q,\widetilde{\boldsymbol{\tau}})}{\|\widetilde{\boldsymbol{\tau}}\|_{\mathrm{div};\Omega}} \geq \beta_3 \|q\|_{0;\Omega}, \quad \forall q \in Q,$$

with $\beta_3 > 0$, see [42] for further details. Then, a straightforward application of [42, Theorem 2.3], and the fact that $\|\widetilde{\boldsymbol{F}}_{\phi,\boldsymbol{u}}\|_{\mathbf{M}'} \leq \|(\xi(\phi))^{-1}\boldsymbol{u}\|_{0;\Omega}$ (cf. (3.24)) conclude the proof. □

Note that Lemmas 3.1 and 3.2 imply the well posedness of the operators $\mathbf{S}_\phi$ and $\widetilde{\mathbf{S}}_\phi$, respectively, and therefore their composition, the operator $\mathbf{T}_\phi$, is well defined.

3.4 Well posedness of the fixed point equation

The aim of this section is to apply the Banach fixed point theorem to show that (3.19) has a unique solution. To this end, for a given $r > 0$, we define the ball $\mathcal{B}_r := \left\{ q \in Q : \quad \|q\|_{0;\Omega} \leq r \right\}$. The next lemma shows that the operator $\mathbf{T}_\phi$ maps $\mathcal{B}_r$ into itself for a sufficiently small data.

Lemma 3.3. *Given $r > 0$ and $\phi \in L^2(\Omega)$ such that $\phi < \phi_{\max}$ on Ω, we assume*

$$\frac{\beta\widetilde{\beta}}{\xi_{\min}} \left\{ \frac{r}{\mu_{\min}} + \|\mathbf{g}(\phi)\|_{0;\Omega} \right\} \leq r, \tag{3.25}$$

where β and $\widetilde{\beta}$ are the constants determined by Lemmas 3.1 and 3.2, respectively. Then, the inclusion $\mathbf{T}_\phi(\mathcal{B}_r) \subseteq \mathcal{B}_r$ holds.

Proof. Given $\widetilde{p} \in \mathcal{B}_r$, it follows from Lemma 3.1 that

$$\|\mathbf{S}_\phi(\widetilde{p})\|_{0;\Omega} \leq \beta \left\{ \|(\mu(\phi))^{-1}\widetilde{p}\|_{0;\Omega} + \|\mathbf{g}(\phi)\|_{0;\Omega} \right\} \leq \beta \left\{ \frac{r}{\mu_{\min}} + \|\mathbf{g}(\phi)\|_{0;\Omega} \right\},$$

which, along with Lemma 3.2, yields

$$\|\mathbf{T}_\phi(\widetilde{p})\| = \|\widetilde{\mathbf{S}}_\phi(\mathbf{S}_\phi(\widetilde{p}))\|_{0;\Omega} \leq \frac{\widetilde{\beta}}{\xi_{\min}} \|\mathbf{S}_\phi(\widetilde{p})\|_{0;\Omega} \leq \frac{\widetilde{\beta}\beta}{\xi_{\min}} \left\{ \frac{r}{\mu_{\min}} + \|\mathbf{g}(\phi)\|_{0;\Omega} \right\}.$$

The foregoing inequality and (3.25) imply $\|\mathbf{T}_\phi(\widetilde{p})\|_{0;\Omega} \leq r$, thus completing the proof. □

Continuing, we now show that both operators $\mathbf{S}_\phi$ and $\widetilde{\mathbf{S}}_\phi$ are Lipschitz-continuous.

Lemma 3.4. *Let $\phi \in L^2(\Omega)$ such that $\phi < \phi_{\max}$ on Ω. Then, the following inequalities hold*

$$\begin{aligned} \|\mathbf{S}_\phi(\widetilde{p}_1) - \mathbf{S}_\phi(\widetilde{p}_2)\|_{0;\Omega} &\leq \frac{\beta}{\mu_{\min}} \|\widetilde{p}_1 - \widetilde{p}_2\|_{0;\Omega} \qquad \forall\, \widetilde{p}_1,\, \widetilde{p}_2 \in Q\,, \\ \|\widetilde{\mathbf{S}}_\phi(\boldsymbol{u}_1) - \widetilde{\mathbf{S}}_\phi(\boldsymbol{u}_2)\|_{0;\Omega} &\leq \frac{\widetilde{\beta}}{\xi_{\min}} \|\boldsymbol{u}_1 - \boldsymbol{u}_2\|_{0;\Omega} \qquad \forall\, \boldsymbol{u}_1,\, \boldsymbol{u}_2 \in \mathbf{L}^2(\Omega)\,. \end{aligned}$$

Proof. To prove the first inequality, we make use of the linearity of $\mathbf{S}_\phi$. Indeed, given $\widetilde{p}_1, \widetilde{p}_2 \in L^2_0(\Omega)$, $\mathbf{S}_\phi(\widetilde{p}_1) - \mathbf{S}_\phi(\widetilde{p}_2) = \mathbf{S}_\phi(\widetilde{p}_1 - \widetilde{p}_2)$ returns the solution of system (3.14) whose right hand side of (3.14b)

is evaluated at $\widetilde{p}_1 - \widetilde{p}_2$, and the right hand side of (3.14c) is zero. Then, replacing $\boldsymbol{G}_\phi = \boldsymbol{0}$ in Lemma 3.1, we get

$$\|\mathbf{S}_\phi(\widetilde{p}_1) - \mathbf{S}_\phi(\widetilde{p}_2)\|_{0;\Omega} \leq \frac{\beta}{\mu_{\min}} \|\widetilde{p}_1 - \widetilde{p}_2\|_{0;\Omega}.$$

Analogously, using similar arguments but employing Lemma 3.2 instead, we find that

$$\|\widetilde{\mathbf{S}}_\phi(\boldsymbol{u}_1) - \widetilde{\mathbf{S}}_\phi(\boldsymbol{u}_2)\|_{0;\Omega} = \|\widetilde{\mathbf{S}}_\phi(\boldsymbol{u}_1 - \boldsymbol{u}_2)\|_{0;\Omega} \leq \frac{\widetilde{\beta}}{\xi_{\min}} \|\boldsymbol{u}_1 - \boldsymbol{u}_2\|_{0;\Omega}\,.$$

□

We are now in conditions to introduce the theorem that states the existence and uniqueness of weak solutions of the coupled system (3.14)–(3.17), and hence the ones for (3.3).

Theorem 3.5. *Let $\phi \in L^2(\Omega)$ such that $\phi < \phi_{\max}$ in Ω, and let $r > 0$. In addition, let β, $\widetilde{\beta}$ be the constants determined by Lemmas 3.1 and 3.2, respectively, and assume that the data satisfies (3.25) and*

$$\frac{\beta\widetilde{\beta}}{\xi_{\min}\,\mu_{\min}} < 1\,.$$

Then, the coupled system (3.14)-(3.17) has a unique solution $((\boldsymbol{t},\boldsymbol{\sigma}),(\boldsymbol{u},\boldsymbol{\gamma})) \in \mathbb{X}\times\mathbb{V}$ and $(\widetilde{\boldsymbol{\sigma}},\widetilde{p}) \in \mathbf{M}\times Q$, with $\widetilde{p} \in \mathcal{B}_r$, and there hold

$$\|((\boldsymbol{t},\boldsymbol{\sigma}),(\boldsymbol{u},\boldsymbol{\gamma}))\|_{\mathbb{X}\times\mathbb{Y}} \leq \beta\left\{\|(\mu(\phi))^{-1}\widetilde{p}\|_{0;\Omega} + \|\mathbf{g}(\phi)\|_{0;\Omega}\right\},$$
$$\|(\widetilde{\boldsymbol{\sigma}},\widetilde{p})\|_{\mathbf{M}\times Q} \leq \widetilde{\beta}\,\|(\xi(\phi))^{-1}\boldsymbol{u}\|_{0;\Omega}\,.$$

Proof. From the assumption (3.25), Lemma 3.3 ensures that $\mathbf{T}_\phi(\mathcal{B}_r) \subseteq \mathcal{B}_r$. Then, making use of Lemma 3.4 and the linearity of $\mathbf{T}_\phi$, for each $\widetilde{p}_1, \widetilde{p}_2 \in Q$ we have

$$\|\mathbf{T}_\phi(\widetilde{p}_1) - \mathbf{T}_\phi(\widetilde{p}_2)\|_{0;\Omega} = \|\mathbf{T}_\phi(\widetilde{p}_1 - \widetilde{p}_2)\|_{0;\Omega} = \|\widetilde{\mathbf{S}}_\phi(\mathbf{S}_\phi(\widetilde{p}_1 - \widetilde{p}_2))\|_{0;\Omega} \leq \frac{\beta\widetilde{\beta}}{\xi_{\min}\,\mu_{\min}} \|\widetilde{p}_1 - \widetilde{p}_2\|_{0;\Omega}\,,$$

and since $\mu_{\min}\,\xi_{\min} > \widetilde{\beta}\beta$, we get that $\mathbf{T}_\phi$ is a contraction. The uniqueness of solution for the equation $\mathbf{T}_\phi(\widetilde{p}) = \widetilde{p}$, or equivalently of the coupled system (3.14)-(3.17) with $\widetilde{p} \in \mathcal{B}_r$, is a result of the application of the Banach fixed-point theorem. The a prori estimates follow from Lemmas 3.1 and 3.2. □

4 Discrete formulation

In this section we introduce the Galerkin scheme associated with the coupled system given by (3.14) and (3.17), and proceed analogously to the analysis developed in Section 3 to derive its well-posedness.

4.1 Preliminaries and setting of the scheme

We begin by letting $\{\mathcal{T}_h\}_{h>0}$ be a regular family of triangulations of Ω made up of triangles K (when $d = 2$) or tetrahedra K (when $d = 3$) of diameter h_K, where h stands for both a sub-index of each triangulation $\mathcal{T}_h$, as well as for the largest diameter, that is $h := \max\left\{h_K : \quad K \in \mathcal{T}_h\right\}$. Next, given an integer $\ell \geq 0$ and a subset Λ of $\mathbb{R}^d$, we denote by $\mathcal{P}_\ell(\Lambda)$ the space of polynomials of total degree

less than or equal to ℓ defined on Λ, $\boldsymbol{\mathcal{P}}_\ell(\Lambda) := [\mathcal{P}_\ell(\Lambda)]^d$ and $\mathbb{P}_\ell(\Lambda) := [\mathcal{P}_\ell(\Lambda)]^{d\times d}$. We also define $\bar{\mathcal{P}}_\ell$ the set of polynomials of degree equal to ℓ. Then, denoting by $\boldsymbol{x}$ a generic vector of $\mathbb{R}^d$, the local Raviart–Thomas space of order ℓ is defined for each $K \in \mathcal{T}_h$ by

$$\mathbf{RT}_\ell(K) := \boldsymbol{\mathcal{P}}_\ell(K) \oplus \bar{\mathcal{P}}_\ell(K)\,\boldsymbol{x}\,.$$

In turn, the local bubble space is defined by

$$\mathbf{B}_\ell(K) = \begin{cases} \mathrm{curl}(b_K \mathcal{P}_\ell(K)) & \text{if } d = 2, \\ \mathrm{curl}(b_K \boldsymbol{\mathcal{P}}_\ell(K)) & \text{if } d = 3, \end{cases}$$

where b_K is the bubble function on K, which is defined as the product of its $d+1$ barycentric coordinates, $\mathrm{curl}(u) = \left(\frac{\partial u}{\partial y}, -\frac{\partial u}{\partial x}\right)$ for $d = 2$ and $u : K \to \mathbb{R}$, and $\mathrm{curl}(\boldsymbol{u}) = \nabla \times \boldsymbol{u}$ for $d = 3$ and $\boldsymbol{u} : K \to \mathbb{R}^3$. Furthermore, we set the global spaces:

$$\begin{aligned}
\boldsymbol{\mathcal{P}}_\ell(\Omega) &:= \left\{ \boldsymbol{w}_h \in \mathbf{L}^2(\Omega) : \quad \boldsymbol{w}_h|_K \in \boldsymbol{\mathcal{P}}_\ell(K), \quad \forall\, K \in \mathcal{T}_h \right\}, \\
\mathbf{RT}_\ell(\Omega) &:= \left\{ \boldsymbol{w}_h \in \mathbf{H}(\mathrm{div};\Omega) : \quad \boldsymbol{w}_h|_K \in \mathbf{RT}_\ell(K), \quad \forall\, K \in \mathcal{T}_h \right\}, \\
\mathbf{B}_\ell(\Omega) &:= \left\{ \boldsymbol{w}_h \in \mathbf{H}(\mathrm{div};\Omega) : \quad \boldsymbol{w}_h|_K \in \mathbf{B}_\ell(K), \quad \forall\, K \in \mathcal{T}_h \right\},
\end{aligned}$$

and their corresponding tensor versions

$$\begin{aligned}
\mathbb{P}_\ell(\Omega) &:= \left\{ \boldsymbol{\tau}_h \in \mathbb{L}^2(\Omega) : \quad \boldsymbol{\tau}_h|_K \in \mathbb{P}_\ell(K), \quad \forall\, K \in \mathcal{T}_h \right\}, \\
\mathbb{RT}_\ell(\Omega) &:= \left\{ \boldsymbol{\tau}_h \in \mathbb{H}(\mathbf{div};\Omega) : \quad \boldsymbol{\tau}_{h,i} \in \mathbf{RT}_\ell(\Omega), \quad \forall\, i = \{1,\dots,d\} \right\}, \\
\mathbb{B}_\ell(\Omega) &:= \left\{ \boldsymbol{\tau}_h \in \mathbb{H}(\mathbf{div};\Omega) : \quad \boldsymbol{\tau}_{h,i} \in \mathbf{B}_\ell(\Omega), \quad \forall\, i = \{1,\dots,d\} \right\},
\end{aligned}$$

where $\boldsymbol{\tau}_{h,i}$ stands for the i-th row of the matrix $\boldsymbol{\tau}_h$. In what follows, we let $\mathbb{H}_h^{\boldsymbol{t}}$, $\mathbb{M}_h$, $\mathbf{H}_h^{\boldsymbol{u}}$, and $\mathbb{H}_h^{\gamma}$ be finite dimensional subspaces of $\mathbb{L}^2_{\mathrm{tr}}(\Omega)$, $\mathbb{H}(\mathbf{div};\Omega)$, $\mathbf{L}^2(\Omega)$, and $\mathbb{L}^2_{\mathrm{skew}}(\Omega)$, respectively, and we let $\mathbb{H}_h^{\boldsymbol{\sigma}} := \mathbb{H}_{\mathbf{tr}}(\mathbf{div};\Omega) \cap \mathbb{M}_h$ be the finite element subspace for the pseudostress.

Suitable subspaces to approximate (3.14) are the PEERS finite elements introduced by Arnold et. al. [9] for elasticity problems, which, letting $\mathbb{M}_h = \mathbb{RT}_\ell(\Omega) \oplus \mathbb{B}_\ell(\Omega)$, are defined as

$$\begin{gathered}
\mathbb{H}_h^{\boldsymbol{t}} := \mathbb{L}^2_{\mathrm{tr}}(\Omega) \cap \mathbb{P}_{\ell+d}(\Omega), \quad \mathbb{H}_h^{\boldsymbol{\sigma}} := \mathbb{H}_{\mathbf{tr}}(\mathbf{div};\Omega) \cap \mathbb{M}_h, \quad \mathbf{H}_h^{\boldsymbol{u}} := \boldsymbol{\mathcal{P}}_\ell(\Omega), \\
\mathbb{H}_h^{\gamma} := \mathbb{C}(\bar{\Omega}) \cap \mathbb{L}^2_{\mathrm{skew}}(\Omega) \cap \mathbb{P}_{\ell+1}(\Omega).
\end{gathered} \tag{4.1}$$

Hence the discrete subspaces for $\mathbb{X} \times \mathbb{V}$ is $\mathbb{X}_h := \mathbb{H}_h^{\boldsymbol{t}} \times \mathbb{M}_h^{\boldsymbol{\sigma}}$ and $\mathbb{V}_h := \mathbf{H}_h^{\boldsymbol{u}} \times \mathbb{H}_h^{\gamma}$. We observe that $\mathbb{M}_h$ contains the multiples of $\mathbb{I}$, and in consequence we can equivalently define $\mathbb{H}_h^{\boldsymbol{\sigma}}$ by

$$\mathbb{H}_h^{\boldsymbol{\sigma}} = \left\{ \boldsymbol{\tau}_h - \left(\frac{1}{d|\Omega|} \int_\Omega \mathrm{tr}(\boldsymbol{\tau}_h) \right) \mathbb{I} : \quad \boldsymbol{\tau}_h \in \mathbb{M}_h \right\}.$$

Moreover, in light of the fact that $\mathbf{div}(\mathbb{M}_h) \subseteq \mathbf{H}_h^{\boldsymbol{u}}$, the discrete kernel of the operator $\boldsymbol{\mathcal{B}}_2$ reduces to

$$V_{0,h} := \left\{ \boldsymbol{\tau}_h \in \mathbb{H}_h^{\boldsymbol{\sigma}} : \quad \int_\Omega \boldsymbol{\tau}_h : \boldsymbol{\eta}_h = 0 \quad \forall \boldsymbol{\eta}_h \in \mathbb{H}_h^{\gamma}, \quad \mathbf{div}(\boldsymbol{\tau}_h) = \mathbf{0} \right\}.$$

The subspaces for the approximation of the second system of equations are chosen by

$$\mathbf{M}_h := \mathbf{H}_0(\mathrm{div};\Omega) \cap \mathbf{RT}_\ell(\Omega), \qquad Q_h := L^2_0(\Omega) \cap \mathcal{P}_\ell(\Omega). \tag{4.2}$$

Hence, given $\widetilde{p}_h \in Q_h$ and $\phi \in L^2(\Omega)$, the discrete formulation of (3.14) reads: Find $((\boldsymbol{t}_h, \boldsymbol{\sigma}_h), (\boldsymbol{u}_h, \boldsymbol{\gamma}_h)) \in \mathbb{X}_h \times \mathbb{V}_h$ such that

$$\begin{aligned}
\mathbf{a}_\phi(\boldsymbol{t}_h, \boldsymbol{s}_h) + \mathbf{b}_1(\boldsymbol{s}_h, \boldsymbol{\sigma}_h) &= 0 && \forall \boldsymbol{s}_h \in \mathbb{H}_h^{\boldsymbol{t}}, && \text{(4.3a)}\\
\mathbf{b}_1(\boldsymbol{t}_h, \boldsymbol{\tau}_h) - \mathbf{c}_\phi(\boldsymbol{\sigma}_h, \boldsymbol{\tau}_h) + \mathbf{b}_2(\boldsymbol{\tau}_h, (\boldsymbol{u}_h, \boldsymbol{\gamma}_h)) &= \boldsymbol{F}_{\phi,\widetilde{p}_h}(\boldsymbol{\tau}_h) && \forall \boldsymbol{\tau}_h \in \mathbb{H}_h^{\boldsymbol{\sigma}}, && \text{(4.3b)}\\
\mathbf{b}_2(\boldsymbol{\sigma}_h, (\boldsymbol{w}_h, \boldsymbol{\eta}_h)) &= \boldsymbol{G}_\phi(\boldsymbol{w}_h, \boldsymbol{\eta}_h) && \forall (\boldsymbol{w}_h, \boldsymbol{\eta}_h) \in \mathbb{V}_h. && \text{(4.3c)}
\end{aligned}$$

In turn, given $\boldsymbol{u}_h \in \mathbf{H}_h^{\boldsymbol{u}}$ and $\phi \in L^2(\Omega)$, the discrete formulation of system (3.17) is given by: Find $(\widetilde{\boldsymbol{\sigma}}_h, \widetilde{p}_h) \in \mathbf{M}_h \times Q_h$ such that

$$\begin{aligned}
\widetilde{\mathbf{a}}_\phi(\widetilde{\boldsymbol{\sigma}}_h, \widetilde{\boldsymbol{\tau}}_h) + \widetilde{\mathbf{b}}(\widetilde{p}_h, \widetilde{\boldsymbol{\tau}}_h) &= \widetilde{\boldsymbol{F}}_{\phi,\boldsymbol{u}_h}(\widetilde{\boldsymbol{\tau}}_h) && \forall\, \widetilde{\boldsymbol{\tau}}_h \in \mathbf{M}_h, && \text{(4.4a)}\\
\widetilde{\mathbf{b}}(q_h, \widetilde{\boldsymbol{\sigma}}_h) &= 0 && \forall\, q_h \in Q_h. && \text{(4.4b)}
\end{aligned}$$

To state the discrete version of (3.19) we first define the discrete counterparts of the operators $\mathbf{S}_\phi$, $\widetilde{\mathbf{S}}_\phi$, and $\mathbf{T}_\phi$, for which we consider here that $\phi \in L^2(\Omega)$ is a given function. Let $\mathbf{S}_{\phi,h} : Q_h \to \mathbb{H}_h^{\boldsymbol{u}}$ be the operator defined by

$$\mathbf{S}_{\phi,h}(\widetilde{p}_h) := \boldsymbol{u}_h \quad \forall\, \widetilde{p}_h \in Q_h,$$

where $((\boldsymbol{t}_h, \boldsymbol{\sigma}_h), (\boldsymbol{u}_h, \boldsymbol{\gamma}_h)) \in \mathbb{X}_h \times \mathbb{V}_h$ is the unique solution of (4.3) for the given $\widetilde{p}_h \in Q_h$ and $\phi \in L^2(\Omega)$. Similarly, we let $\widetilde{\mathbf{S}}_{\phi,h} : \mathbb{H}_h^{\boldsymbol{u}} \to Q_h$ be the operator defined by

$$\widetilde{\mathbf{S}}_{\phi,h}(\boldsymbol{u}_h) := \widetilde{p}_h \quad \forall\, \boldsymbol{u}_h \in \mathbb{H}_h^{\boldsymbol{u}},$$

where $(\widetilde{\boldsymbol{\sigma}}_h, \widetilde{p}_h) \in \mathbf{M}_h \times Q_h$ is the unique solution of (4.4) for the given $\boldsymbol{u}_h \in \mathbf{H}_h^{\boldsymbol{u}}$ and $\phi \in L^2(\Omega)$. The well-posedness of $\mathbf{S}_{\phi,h}$ and $\widetilde{\mathbf{S}}_{\phi,h}$, which is equivalent to proving the uniqueness of solutions of (4.3) and (4.4), respectively, is shown in Section 4.2. Finally, we define the operator $\mathbf{T}_{\phi,h} := \widetilde{\mathbf{S}}_{\phi,h} \circ \mathbf{S}_{\phi,h} : Q \to Q$, and realize that the coupled problem (4.3)–(4.4) is equivalent to finding a fixed point of $\mathbf{T}_{\phi,h}$, that is $\widetilde{p}_h \in Q$ such that

$$\mathbf{T}_{\phi,h}(\widetilde{p}_h) = \widetilde{p}_h. \tag{4.5}$$

Analogously as in the continuous case, $\mathbf{S}_{\phi,h}$, $\widetilde{\mathbf{S}}_{\phi,h}$, and $\mathbf{T}_{\phi,h}$ are linear operators.

4.2 Solvability of the discrete problem

The solvability analyses of the discrete problems (4.3) and (4.4), and of the fixed point equation (4.5), are carried out using analogous arguments to those used for the continuous case. In particular, the next two results are the discrete versions of Lemmas 3.1 and 3.2, respectively.

Lemma 4.1. *Given $\widetilde{p}_h \in Q_h$ and $\phi \in L^2(\Omega)$, there exists a unique $((\boldsymbol{t}_h, \boldsymbol{\sigma}_h), (\boldsymbol{u}_h, \boldsymbol{\gamma}_h)) \in \mathbb{X}_h \times \mathbb{V}_h$ solution of (4.3), and hence one can define $\mathbf{S}_{\phi,h}(\widetilde{p}_h) := \boldsymbol{u}_h$. Moreover, there exists a positive constant $\beta > 0$, depending on d, $\mu_{\min}$, and $\mu_{\max}$, and thus independent of h, such that*

$$\|\mathbf{S}_{\phi,h}(\widetilde{p}_h)\|_{0;\Omega} = \|\boldsymbol{u}_h\|_{0;\Omega} \leq \|((\boldsymbol{t}_h, \boldsymbol{\sigma}_h), (\boldsymbol{u}_h, \boldsymbol{\gamma}_h))\|_{\mathbb{X}\times\mathbb{Y}} \leq \beta \left\{ \|(\mu(\phi))^{-1}\widetilde{p}_h\|_{0;\Omega} + \|\mathbf{g}(\phi)\|_{0;\Omega} \right\}.$$

Proof. The Lipschitz continuity and strong monotonicity of $\boldsymbol{\mathcal{A}}_\phi$ was shown in the proof of Lemma 3.1. On the other hand, since $[\boldsymbol{\mathcal{C}}_\phi(\boldsymbol{\tau}_h), \boldsymbol{\tau}_h] \geq 0$ for all $\boldsymbol{\tau}_h \in \mathbb{M}_h^{\boldsymbol{\sigma}}$, $\boldsymbol{\mathcal{C}}_\phi$ is clearly positive semi-definite in $V_{h,0}$, the discrete kernel of $\boldsymbol{\mathcal{B}}_2$. In addition, given $\boldsymbol{\tau}_h \in V_{0,h}$, we have that $\mathbf{div}(\boldsymbol{\tau}_h) = \mathbf{0}$ and $\|\boldsymbol{\tau}_h^{\mathrm{d}}\|_{0;\Omega} \geq \beta_0 \|\boldsymbol{\tau}_h\|_{\mathbf{div};\Omega}$ [42, Lemma 2.3], and since $\boldsymbol{\tau}_h^{\mathrm{d}} \in \mathbb{H}_h^{\boldsymbol{t}}$, we find that

$$\sup_{\substack{\boldsymbol{s}_h \in \mathbb{H}_h^{\boldsymbol{t}} \\ \boldsymbol{s} \neq \mathbf{0}}} \frac{[\boldsymbol{\mathcal{B}}_1(\boldsymbol{s}_h), \boldsymbol{\tau}_h]}{\|\boldsymbol{s}_h\|_{0;\Omega}} \geq \frac{[\boldsymbol{\mathcal{B}}_1(-\boldsymbol{\tau}_h^{\mathrm{d}}), \boldsymbol{\tau}_h]}{\|\boldsymbol{\tau}_h^{\mathrm{d}}\|_{0;\Omega}} = \|\boldsymbol{\tau}_h^{\mathrm{d}}\|_{0;\Omega} \geq \beta_0 \|\boldsymbol{\tau}_h\|_{\mathbf{div};\Omega},$$

which shows that the operator $\boldsymbol{\mathcal{B}}_1$ satisfies the discrete inf-sup condition. In turn, the same condition for $\boldsymbol{\mathcal{B}}_2$ has been shown in previous works, e.g. [44]. Consequently, having the scheme satisfied the hypotheses of [43, Theorem 3.4], the proof is concluded. □

Lemma 4.2. *Given $\boldsymbol{u}_h \in \mathbf{H}_h^{\boldsymbol{u}}$ and $\phi \in L^2(\Omega)$ such that $\phi < \phi_{\max}$ on Ω, there exists a unique $(\widetilde{\boldsymbol{\sigma}}_h, \widetilde{p}_h) \in \mathbf{M}_h \times Q_h$ solution of (3.17), and hence one can define $\widetilde{\mathbf{S}}_{\phi,h}(\boldsymbol{u}_h) := \widetilde{p}_h$. Moreover, there exists $\widetilde{\beta} > 0$, depending on $\xi_{\min}$ and $\xi_{\max}$, and thus independent of h, such that*

$$\|\widetilde{\mathbf{S}}_{\phi,h}(\boldsymbol{u}_h)\| = \|\widetilde{p}_h\| \leq \|(\widetilde{\boldsymbol{\sigma}}_h, \widetilde{p}_h)\|_{\mathbf{M}\times Q} \leq \widetilde{\beta}\|(\xi(\phi))^{-1}\boldsymbol{u}_h\|_{0;\Omega}.$$

Proof. Since $\mathrm{div}(\mathbf{M}_h) \subseteq Q_h$, the proof of this lemma is straightforward and analogous to that of Lemma 3.1. □

As in the continuous case, given $r > 0$, we define the ball $\mathcal{B}_{r,h} := \{q_h \in Q_h : \quad \|q_h\|_{0;\Omega} \leq r\}$. The analogous versions of Lemmas 3.3 and 3.4 are provided next.

Lemma 4.3. *Given $r > 0$ and $\phi \in L^2(\Omega)$ such that $\phi < \phi_{\max}$ on Ω, we assume that the data, β, and $\widetilde{\beta}$ satisfy (3.25), where β and $\widetilde{\beta}$ are the constants determined by Lemmas 4.1 and 4.2, respectively. Then, the inclusion $\mathbf{T}_{\phi,h}(\mathcal{B}_{r,h}) \subseteq \mathcal{B}_{r,h}$ holds.*

Proof. Given $\widetilde{p}_h \in \mathcal{B}_{r,h}$, it follows from Lemma 4.1 that

$$\|\mathbf{S}_{\phi,h}(\widetilde{p}_h)\|_{0;\Omega} \leq \beta\left\{\|(\mu(\phi))^{-1}\widetilde{p}_h\|_{0;\Omega} + \|\mathbf{g}(\phi)\|_{0;\Omega}\right\} \leq \beta\left\{\frac{r}{\mu_{\min}} + \|\mathbf{g}(\phi)\|_{0;\Omega}\right\},$$

which along with Lemma 4.2 yields

$$\|\widetilde{\mathbf{S}}_{\phi,h}(\mathbf{S}_{\phi,h}(\widetilde{p}_h))\|_{0;\Omega} \leq \frac{\widetilde{\beta}}{\xi_{\min}}\|\mathbf{S}_{\phi,h}(\widetilde{p}_h)\|_{0;\Omega} \leq \frac{\beta\widetilde{\beta}}{\xi_{\min}}\left\{\frac{r}{\mu_{\min}} + \|\mathbf{g}(\phi)\|_{0;\Omega}\right\}.$$

The above inequality and (3.25) imply $\|\mathbf{T}_{\phi,h}(\widetilde{p}_h)\|_{0;\Omega} \leq r$, which ends the proof. □

Lemma 4.4. *Let $\phi \in L^2(\Omega)$ such that $\phi < \phi_{\max}$ in Ω. Then, the following inequalities hold*

$$\|\mathbf{S}_{\phi,h}(\widetilde{p}_1) - \mathbf{S}_{\phi,h}(\widetilde{p}_2)\|_{0;\Omega} \leq \frac{\beta}{\mu_{\min}}\|\widetilde{p}_1 - \widetilde{p}_2\|_{0;\Omega} \quad \forall \widetilde{p}_1, \widetilde{p}_2 \in Q_h,$$

$$\|\widetilde{\mathbf{S}}_{\phi,h}(\boldsymbol{u}_1) - \widetilde{\mathbf{S}}_{\phi,h}(\boldsymbol{u}_2)\|_{0;\Omega} \leq \frac{\widetilde{\beta}}{\xi_{\min}}\|\boldsymbol{u}_1 - \boldsymbol{u}_2\|_{0;\Omega} \quad \forall \boldsymbol{u}_1, \boldsymbol{u}_2 \in \mathbf{H}_h^{\boldsymbol{u}}.$$

Proof. To prove the first inequality, we make use of the linearity of $\mathbf{S}_{\phi,h}$. Indeed, given $\widetilde{p}_1, \widetilde{p}_2 \in Q_h$, $\mathbf{S}_{\phi,h}(\widetilde{p}_1) - \mathbf{S}_{\phi,h}(\widetilde{p}_2) = \mathbf{S}_{\phi,h}(\widetilde{p}_1 - \widetilde{p}_2)$ returns the solution of system (4.3) whose right hand side of (4.3b) is evaluated in $\widetilde{p}_1 - \widetilde{p}_2$ and the right hand side of (4.3c) is zero. Then, replacing $\boldsymbol{G}_\phi = \mathbf{0}$ in Lemma 4.1, we have

$$\|\mathbf{S}_{\phi,h}(\widetilde{p}_1) - \mathbf{S}_{\phi,h}(\widetilde{p}_2)\|_{0;\Omega} \leq \frac{\beta}{\mu_{\min}}\|\widetilde{p}_1 - \widetilde{p}_2\|_{0;\Omega}.$$

Analogously, using now Lemma 4.2, we obtain for each $\boldsymbol{u}_1$, $\boldsymbol{u}_2 \in \mathbf{H}_h^{\boldsymbol{u}}$

$$\|\widetilde{\mathbf{S}}_{\phi,h}(\boldsymbol{u}_1) - \widetilde{\mathbf{S}}_{\phi,h}(\boldsymbol{u}_2)\|_{0;\Omega} = \|\widetilde{\mathbf{S}}_{\phi,h}(\boldsymbol{u}_1 - \boldsymbol{u}_2)\|_{0;\Omega} \leq \frac{\widetilde{\beta}}{\xi_{\min}}\|\boldsymbol{u}_1 - \boldsymbol{u}_2\|_{0;\Omega}.$$

□

We end this section by establishing the existence and uniqueness of solutions of the discrete coupled problem (4.3)–(4.4) through the fixed point equation (4.5).

Theorem 4.5. *Let $\phi \in L^2(\Omega)$ such that $\phi < \phi_{\max}$ on Ω, and let $r > 0$. In addition, let β, $\widetilde{\beta}$ be the constants determined by Lemmas 4.1 and 4.2, respectively, and assume that the data satisfy (3.25) and*

$$\frac{\beta\widetilde{\beta}}{\mu_{\min}\,\xi_{\min}} < 1\,.$$

Then, the coupled system (4.3)-(4.4) has a unique solution $((\boldsymbol{t}_h,\boldsymbol{\sigma}_h),(\boldsymbol{u}_h,\boldsymbol{\gamma}_h)) \in \mathbb{X}_h\times\mathbb{V}_h$ and $(\widetilde{\boldsymbol{\sigma}}_h,\widetilde{p}_h) \in \mathbf{M}_h \times Q_h$, with $\widetilde{p}_h \in \mathcal{B}_{r,h}$, and there holds

$$\|((\boldsymbol{t}_h,\boldsymbol{\sigma}_h),(\boldsymbol{u}_h,\boldsymbol{\gamma}_h))\|_{\mathbb{X}\times\mathbb{Y}} \leq \beta\left\{\|(\mu(\phi))^{-1}\widetilde{p}_h\|_{0;\Omega} + \|\mathbf{g}(\phi)\|_{0;\Omega}\right\},$$
$$\|(\widetilde{\boldsymbol{\sigma}}_h,\widetilde{p}_h)\|_{\mathbf{M}\times Q} \leq \widetilde{\beta}\,\|(\xi(\phi))^{-1}\boldsymbol{u}_h\|_{0;\Omega}\,.$$

Proof. From the assumption (3.25), Lemma 4.3 ensures that $\mathbf{T}_{\phi,h}(\mathcal{B}_{r,h}) \subseteq \mathcal{B}_{r,h}$. Then, making use of Lemma 4.4 and the linearity of $\mathbf{T}_{\phi,h}$, for each $\widetilde{p}_1,\widetilde{p}_2 \in Q_h$ we have

$$\|\mathbf{T}_{\phi,h}(\widetilde{p}_1) - \mathbf{T}_{\phi,h}(\widetilde{p}_2)\|_{0;\Omega} = \|\widetilde{\mathbf{S}}_{\phi,h}(\mathbf{S}_{\phi,h}(\widetilde{p}_1-\widetilde{p}_2))\|_{0;\Omega} \leq \frac{\widetilde{\beta}\beta}{\mu_{\min}\,\xi_{\min}}\,\|\widetilde{p}_1-\widetilde{p}_2\|_{0;\Omega}\,,$$

and since $\mu_{\min}\,\xi_{\min} > \widetilde{\beta}\beta$, we get that $\mathbf{T}_{\phi,h}$ is a contraction. The uniqueness of solution of the equation $\mathbf{T}_{\phi,h}(\widetilde{p}_h) = \widetilde{p}_h$, or equivalently of the coupled system (4.3)-(4.4), with $\widetilde{p}_h \in \mathcal{B}_{r,h}$, is a result of the application of the Banach fixed-point theorem. The a priori estimates are then consequence of Lemmas 4.1 and 4.2. □

4.3 A priori error analysis

Let $\mathbf{A}_\phi : \big(\mathbb{X}\times\mathbb{V}\big)\times\big(\mathbb{X}\times\mathbb{V}\big) \to \mathbb{R}$ and $\widetilde{\mathbf{A}}_\phi : \big(\mathbf{M}\times Q\big)\times\big(\mathbf{M}\times Q\big) \to \mathbb{R}$ be the bilinear forms arising from adding the left hand sides of (3.14) and (3.17), respectively. Then, introducing the generic notations

$$(\vec{\boldsymbol{s}},\vec{\boldsymbol{w}}) := ((\boldsymbol{s},\boldsymbol{\tau}),(\boldsymbol{w},\boldsymbol{\eta})) \in \mathbb{X}\times\mathbb{V}\,,\quad (\vec{\boldsymbol{s}}_h,\vec{\boldsymbol{w}}_h) := ((\boldsymbol{s}_h,\boldsymbol{\tau}_h),(\boldsymbol{w}_h,\boldsymbol{\eta}_h)) \in \mathbb{X}_h\times\mathbb{V}_h\,,$$

the pairs of continuous and discrete schemes (3.14)–(4.3) and (3.17)–(4.4) can be rewritten as

$$\begin{aligned}
&\text{Find } (\vec{\boldsymbol{t}},\vec{\boldsymbol{u}}) := ((\boldsymbol{t},\boldsymbol{\sigma}),(\boldsymbol{u},\boldsymbol{\gamma})) \in \mathbb{X}\times\mathbb{V} \text{ such that}\\
&\mathbf{A}_\phi((\vec{\boldsymbol{t}},\vec{\boldsymbol{u}}),(\vec{\boldsymbol{s}},\vec{\boldsymbol{w}})) = \boldsymbol{F}_{\phi,\widetilde{p}}(\boldsymbol{\tau}) + \boldsymbol{G}_\phi(\boldsymbol{w},\boldsymbol{\eta}) \qquad \forall\,(\vec{\boldsymbol{s}},\vec{\boldsymbol{w}}) \in \mathbb{X}\times\mathbb{V}\,,\\
&\text{Find } (\vec{\boldsymbol{t}}_h,\vec{\boldsymbol{u}}_h) := ((\boldsymbol{t}_h,\boldsymbol{\sigma}_h),(\boldsymbol{u}_h,\boldsymbol{\gamma}_h)) \in \mathbb{X}_h\times\mathbb{V}_h \text{ such that}\\
&\mathbf{A}_\phi((\vec{\boldsymbol{t}}_h,\vec{\boldsymbol{u}}_h),(\vec{\boldsymbol{s}}_h,\vec{\boldsymbol{w}}_h)) = \boldsymbol{F}_{\phi,\widetilde{p}_h}(\boldsymbol{\tau}_h) + \boldsymbol{G}_\phi(\boldsymbol{w}_h,\boldsymbol{\eta}_h) \qquad \forall\,(\vec{\boldsymbol{s}}_h,\vec{\boldsymbol{w}}_h) \in \mathbb{X}_h\times\mathbb{V}_h\,,
\end{aligned} \tag{4.6}$$

and

$$\begin{aligned}
&\text{Find } (\widetilde{\boldsymbol{\sigma}},\widetilde{p}) \in \mathbf{M}\times Q \text{ such that}\\
&\widetilde{\mathbf{A}}_\phi((\widetilde{\boldsymbol{\sigma}},\widetilde{p}),(\widetilde{\boldsymbol{\tau}},q)) = \widetilde{\boldsymbol{F}}_{\phi,\boldsymbol{u}}(\widetilde{\boldsymbol{\tau}}) \qquad \forall\,(\widetilde{\boldsymbol{\tau}},q) \in \mathbf{M}\times Q\,,\\
&\text{Find } (\widetilde{\boldsymbol{\sigma}}_h,\widetilde{p}_h) \in \mathbf{M}_h\times Q_h \text{ such that}\\
&\widetilde{\mathbf{A}}_\phi((\widetilde{\boldsymbol{\sigma}}_h,\widetilde{p}_h),(\widetilde{\boldsymbol{\tau}}_h,q_h)) = \widetilde{\boldsymbol{F}}_{\phi,\boldsymbol{u}_h}(\widetilde{\boldsymbol{\tau}}_h) \qquad \forall\,(\widetilde{\boldsymbol{\tau}}_h,q_h) \in \mathbf{M}_h\times Q_h\,,
\end{aligned} \tag{4.7}$$

respectively. Thus, as a consequence of Lemmas 4.1 and 4.2, we deduce the existence of positive constants $\alpha_{\mathbf{A}_\phi}$ and $\alpha_{\widetilde{\mathbf{A}}_\phi}$, independent of h, with which $\mathbf{A}_\phi$ and $\widetilde{\mathbf{A}}_\phi$ satisfy global discrete inf-sup

conditions, namely

$$\sup_{\substack{(\vec{\boldsymbol{s}}_h,\vec{\boldsymbol{w}}_h)\in\mathbb{X}_h\times\mathbb{V}_h\\(\vec{\boldsymbol{s}}_h,\vec{\boldsymbol{w}}_h)\neq(\mathbf{0},\mathbf{0})}}\frac{\mathbf{A}_\phi((\vec{\boldsymbol{r}}_h,\vec{\boldsymbol{v}}_h),(\vec{\boldsymbol{s}}_h,\vec{\boldsymbol{w}}_h))}{\|(\vec{\boldsymbol{s}}_h,\vec{\boldsymbol{w}}_h)\|}\geq\alpha_{\mathbf{A}_\phi}\|(\vec{\boldsymbol{r}}_h,\vec{\boldsymbol{v}}_h)\|\qquad\forall\,(\vec{\boldsymbol{r}}_h,\vec{\boldsymbol{v}}_h)\in\mathbb{X}_h\times\mathbb{V}_h\,,\tag{4.8}$$

and

$$\sup_{\substack{(\widetilde{\boldsymbol{\tau}}_h,q_h)\in\mathbf{M}_h\times Q\\(\widetilde{\boldsymbol{\tau}}_h,q_h)\neq(\mathbf{0},\mathbf{0})}}\frac{\widetilde{\mathbf{A}}_\phi((\widetilde{\boldsymbol{\zeta}}_h,\varphi_h),(\widetilde{\boldsymbol{\tau}}_h,q_h))}{\|(\widetilde{\boldsymbol{\tau}}_h,q_h)\|}\geq\alpha_{\widetilde{\mathbf{A}}_\phi}\|(\widetilde{\boldsymbol{\zeta}}_h,\varphi_h)\|\qquad\forall\,(\widetilde{\boldsymbol{\zeta}}_h,\varphi_h)\in\mathbf{M}_h\times Q_h\,.\tag{4.9}$$

Hereafter, given a subspace X_h of a generic normed space $\big(X,\|\cdot\|_X\big)$, we set for each $x\in X$

$$\mathrm{dist}(x,X_h):=\inf_{x_h\in X_h}\|x-x_h\|_X\,.$$

The following lemma states the Céa estimate for our coupled problem.

Lemma 4.6. *Let $\phi\in L^2(\Omega)$ such that $\phi<\phi_{\max}$ in Ω, and assume that the data satisfies*

$$\frac{\sqrt{d}}{\mu_{\min}\,\alpha_{\mathbf{A}_\phi}}+\frac{1}{\xi_{\min}\,\alpha_{\widetilde{\mathbf{A}}_\phi}}\leq\frac{1}{2}\,.\tag{4.10}$$

Then, there exist positive constants C_1, C_2 ,depending on $\|\mathbf{A}_\phi\|$, $\alpha_{\mathbf{A}_\phi}$, and $\|\widetilde{\mathbf{A}}_\phi\|,\alpha_{\widetilde{\mathbf{A}}_\phi}$, respectively, and thus independent of h, such that

$$\begin{aligned}&\|((\vec{\boldsymbol{t}},\vec{\boldsymbol{u}}),(\widetilde{\boldsymbol{\sigma}},\widetilde{p}))-((\vec{\boldsymbol{t}}_h,\vec{\boldsymbol{u}}_h),(\widetilde{\boldsymbol{\sigma}}_h,\widetilde{p}_h))\|\\&\quad\leq C_1\,\mathrm{dist}\big((\vec{\boldsymbol{t}},\vec{\boldsymbol{u}}),\mathbb{X}_h\times\mathbb{V}_h\big)+C_2\,\mathrm{dist}\big((\widetilde{\boldsymbol{\sigma}},\widetilde{p}),\mathbf{M}_h\times Q_h\big)\,.\end{aligned}\tag{4.11}$$

Proof. It proceeds similarly as in [44]. In fact, bearing in mind (4.8) and (4.9), straightforward applications of the classical Strang estimate to the pairs (4.6) and (4.7) yield

$$\|(\vec{\boldsymbol{t}},\vec{\boldsymbol{u}})-(\vec{\boldsymbol{t}}_h,\vec{\boldsymbol{u}}_h)\|_{\mathbb{X}\times\mathbb{V}}\leq\left(1+\frac{\|\mathbf{A}_\phi\|}{\alpha_{\mathbf{A}_\phi}}\right)\mathrm{dist}\,\big((\vec{\boldsymbol{t}},\vec{\boldsymbol{u}}),\mathbb{X}_h\times\mathbb{V}_h\big)+\frac{1}{\alpha_{\mathbf{A}_\phi}}\|\boldsymbol{F}_{\phi,\widetilde{p}}-\boldsymbol{F}_{\phi,\widetilde{p}_h}\|\,,\tag{4.12}$$

and

$$\|(\widetilde{\boldsymbol{\sigma}},\widetilde{p})-(\widetilde{\boldsymbol{\sigma}}_h,\widetilde{p}_h)\|_{\mathbf{M}\times Q}\leq\left(1+\frac{\|\widetilde{\mathbf{A}}_\phi\|}{\alpha_{\widetilde{\mathbf{A}}_\phi}}\right)\mathrm{dist}\,\big((\widetilde{\boldsymbol{\sigma}},\widetilde{p}),\mathbf{M}_h\times Q_h\big)+\frac{1}{\alpha_{\widetilde{\mathbf{A}}_\phi}}\|\widetilde{\boldsymbol{F}}_{\phi,\boldsymbol{u}}-\widetilde{\boldsymbol{F}}_{\phi,\boldsymbol{u}_h}\|\,,\tag{4.13}$$

respectively. In turn, it follows from (3.22) and (2.10) that

$$\|\boldsymbol{F}_{\phi,\widetilde{p}}-\boldsymbol{F}_{\phi,\widetilde{p}_h}\|=\|\boldsymbol{F}_{\phi,\widetilde{p}-\widetilde{p}_h}\|\leq\sqrt{d}\,\|(\mu(\phi))^{-1}\,(\widetilde{p}-\widetilde{p}_h)\|_{0,\Omega}\leq\frac{\sqrt{d}}{\mu_{\min}}\|\widetilde{p}-\widetilde{p}_h\|_{0,\Omega}\,,\tag{4.14}$$

whereas the bound for $\|\widetilde{\boldsymbol{F}}_{\phi,\boldsymbol{u}}\|$ (cf. (3.24)) and (3.23) give

$$\|\widetilde{\boldsymbol{F}}_{\phi,\boldsymbol{u}}-\widetilde{\boldsymbol{F}}_{\phi,\boldsymbol{u}_h}\|=\|\widetilde{\boldsymbol{F}}_{\phi,\boldsymbol{u}-\boldsymbol{u}_h}\|\leq\frac{1}{\xi_{\min}}\|\boldsymbol{u}-\boldsymbol{u}_h\|_{0,\Omega}\,.\tag{4.15}$$

In this way, employing (4.14) and (4.15) back into (4.12) and (4.13), respectively, and then adding the resulting inequalities, and using (4.10), we obtain (4.11) with

$$C_1:=2\left(1+\frac{\|\mathbf{A}_\phi\|}{\alpha_{\mathbf{A}_\phi}}\right)\quad\text{and}\quad C_2:=2\left(1+\frac{\|\widetilde{\mathbf{A}}_\phi\|}{\alpha_{\widetilde{\mathbf{A}}_\phi}}\right).$$

□

We can now state the rates of convergence of the Galerkin scheme (4.3)–(4.4) with the finite element subspaces (4.1) and (4.2).

Theorem 4.7. *Given $\phi \in L^2(\Omega)$ such that $\phi < \phi_{\max}$ on Ω, let $(\vec{\boldsymbol{t}}, \vec{\boldsymbol{u}}) \in \mathbb{X} \times \mathbb{V}$ and $(\widetilde{\boldsymbol{\sigma}}, \widetilde{p}) \in \mathbf{M} \times Q$ be the unique solution of the coupled system (3.14)–(3.17) from Lemma 3.5 and let $(\vec{\boldsymbol{t}}_h, \vec{\boldsymbol{u}}_h) \in \mathbb{X}_h \times \mathbb{V}_h$, $(\widetilde{\boldsymbol{\sigma}}_h, \widetilde{p}_h) \in \mathbf{M}_h \times Q_h$ be the unique solution of its discrete counterpart from Lemma 3.5. Assume that the data satisfies the inequality (4.10). Given an integer $\ell \geq 0$, assume that there exists $\kappa \in (0, \ell+1]$ such that $\boldsymbol{t} \in \mathbb{H}^\kappa(\Omega) \cap \mathbb{L}^2_{\mathrm{tr}}$, $\boldsymbol{\sigma} \in \mathbb{H}^\kappa(\Omega) \cap \mathbb{H}_{\mathbf{tr}}(\mathbf{div}; \Omega)$, $\mathbf{div}(\boldsymbol{\sigma}) \in \mathbf{H}^\kappa(\Omega)$, $\boldsymbol{u} \in \mathbf{H}^\kappa(\Omega) \cap \mathbf{L}^2(\Omega)$, $\boldsymbol{\gamma} \in \mathbb{H}^\kappa(\Omega) \cap \mathbb{L}^2_{\mathrm{skew}}(\Omega)$, $\widetilde{\boldsymbol{\sigma}} \in \mathbf{H}^\kappa(\Omega) \cap \mathbf{M}$, $\mathrm{div}(\widetilde{\boldsymbol{\sigma}}) \in H^\kappa(\Omega)$ and $\widetilde{p} \in H^\kappa(\Omega) \cap Q$. Then, there exists a positive constant C, independent of h, such that*

$$\|((\vec{\boldsymbol{t}}, \vec{\boldsymbol{u}}), (\widetilde{\boldsymbol{\sigma}}, \widetilde{p})) - ((\vec{\boldsymbol{t}}_h, \vec{\boldsymbol{u}}_h), (\widetilde{\boldsymbol{\sigma}}_h, \widetilde{p}_h))\| \leq C\, h^\kappa \Big\{ \|\boldsymbol{t}\|_{\kappa,\Omega} + \|\boldsymbol{\sigma}\|_{\kappa,\Omega} + \|\mathbf{div}(\boldsymbol{\sigma})\|_{\kappa,\Omega} + \|\boldsymbol{u}\|_{\kappa,\Omega} + \|\boldsymbol{\gamma}\|_{\kappa,\Omega} + \|\widetilde{\boldsymbol{\sigma}}\|_{\kappa,\Omega} + \|\mathrm{div}(\widetilde{\boldsymbol{\sigma}})\|_{\kappa,\Omega} + \|\widetilde{p}\|_{\kappa,\Omega} \Big\}.$$

Proof. The proof follows from (4.11) and the approximation properties of the finite element subspaces $\mathbb{H}^{\boldsymbol{t}}_h$, $\mathbb{H}^{\boldsymbol{\sigma}}_h$, $\mathbf{H}^{\boldsymbol{u}}_h$, $\mathbb{H}^{\boldsymbol{\gamma}}_h$, $\mathbf{M}_h$ and Q_h. □

5 Fully coupled scheme

In this section we address the approximation of the transport equation (2.11a) for the volume fraction of solid particles ϕ, and the coupling to the systems (4.3) and (4.4). Since (2.11a) is of the hyperbolic type, we propose a vertex centered finite volume scheme based on an upwind numerical flux for its approximation. In what follows, we let $\mathcal{N}_h$ be the set of nodes of the triangulation $\mathcal{T}_h$ (primal mesh). We introduce the so-called dual mesh $\mathcal{T}_h^\star$ which consists of the non-overlapping control volumes $K_j^\star$ surrounding the node $s_j \in \mathcal{N}_h$. The control volumes on the dual mesh $\mathcal{T}_h^*$ are built by connecting the barycenters b_K of each triangle $K \in \mathcal{T}_h$ with the middle point of each edge of K, see Figure 5.1. Note that each node in the primal mesh is related to one control volume in the dual mesh. We denote by $\mathcal{E}_h(K^\star)$ the set of inner edges of $K^\star \in \mathcal{T}_h^\star$, so that $e \in \mathcal{E}_h(K^\star)$ if $e \subset \partial K^\star$ and there exists $K \in \mathcal{T}_h$ such that e is the piece of segment connecting b_K and the middle point of an edge of K (see the blue dash-dotted segments in Figure 5.1). In turn, we define the space of piecewise constant functions over the dual mesh

$$S_h := \Big\{ \varphi \in L^2(\Omega) : \quad \varphi|_{K^\star} \in \mathcal{P}_0(K^\star), \quad \forall\, K^\star \in \mathcal{T}_h^\star \Big\}.$$

To introduce the finite volume method, we consider the integral form of equation (2.11a) over each control volume in the dual mesh $\mathcal{T}_h^\star$, this is

$$\int_{K^\star} \partial_t \phi = -\int_{\partial K^\star} \phi \boldsymbol{u} \cdot \boldsymbol{n}, \qquad \forall K^\star \in \mathcal{T}_h^\star. \tag{5.1}$$

Then, the sought unknown is approximated by $\phi_h(t) \in S_h$, i.e., for each $t \in (0, T]$, ϕ is a piecewise constant function on $\mathcal{T}_h^\star$. Furthermore, given a time step $\Delta t > 0$ and $n \in \mathbb{N}$, we make use of the superscript n to denote the evaluation of each function at the time $t^n = n\Delta t$, e.g., $\phi_h^n = \phi_h(t^n)$, $\boldsymbol{u}_h^n = \boldsymbol{u}_h(t^n)$, and the analogous notation is used for other functions. Next, approximating the time derivative in (5.1) by forward Euler and the numerical flux by an upwind approach we obtain

$$\int_{K^\star} \frac{\phi_{h,K^\star}^{n+1} - \phi_{h,K^\star}^n}{\Delta t} = -\sum_{e \in \mathcal{E}_h(K^\star)} \int_e \mathrm{Upw}(\boldsymbol{u}_h^n \cdot \boldsymbol{n}_e; \phi_{h,K^\star}^n, \phi_{h,K_e^\star}^n), \qquad \forall K^\star \in \mathcal{T}_h^\star, \tag{5.2}$$

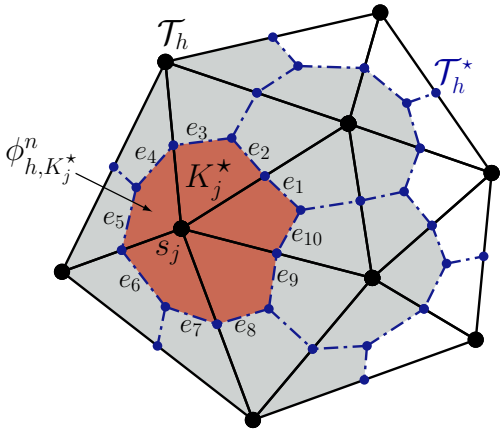

Figure 5.1: Schematic of the primal mesh $\mathcal{T}_h$ (black triangles) and dual mesh $\mathcal{T}_h^\star$ defined by the control volumes $K_j^\star$ delimited by the blue dash-dotted lines and boundary Γ. The boundary of $K_j^\star$ is given by the union of the edges e_i with $i = 1, \dots, 10$.

where $\phi^n_{h,K^\star} = \phi^n_h|_{K^\star}$, $\boldsymbol{n}_e$ is the normal of the edge $e \in \mathcal{E}_h(K^\star)$, $K_e^\star \in \mathcal{T}_h^\star$ denotes the adjacent control volume such that $e \subseteq \bar{K}^\star \cap \bar{K}_e^\star$ and $K^\star \neq K_e^\star$, and the upwind operator is defined by

$$\mathrm{Upw}(a; b, c) := \max\{a, 0\}b + \min\{a, 0\}c, \qquad a, b, c \in \mathbb{R}.$$

The velocity function $\boldsymbol{u}_h^n$ is determined by solving (4.4)–(4.3) (the fixed point equation (4.5)) at the time t^n with ϕ given by the piecewise constant function that for each $K \in \mathcal{T}_h$ is computed by

$$\phi(\boldsymbol{x}) = \frac{1}{|K|}\int_K \phi_h^n, \qquad \boldsymbol{x} \in K.$$

Note the difference between ϕ_h^n and ϕ, the first one is in S_h (dual mesh) while the second one is in $\mathcal{P}_0(\Omega)$ (primal mesh). For the case of a piecewise linear (or constant) approximation $\boldsymbol{u}_h^n$, the integral over the edge in (5.2) can be computed with the Gauss quadrature rule to obtain

$$\phi^{n+1}_{h,K^\star} = \phi^n_{h,K^\star} - \frac{\Delta t}{|K^\star|}\sum_{e\in\mathcal{E}_h(K^\star)} |e|\mathrm{Upw}(\boldsymbol{u}_h^n(\bar{\boldsymbol{x}}_e)\cdot\boldsymbol{n}_e; \phi^n_{h,K^\star}, \phi^n_{h,K_e^\star}), \qquad \forall K^\star \in \mathcal{T}_h^\star, \tag{5.3}$$

where $|e|$ and $|K^\star|$ are the measures of the edge e and control volume $K^\star$, respectively, and $\bar{\boldsymbol{x}}_e$ is the middle point of e. The discrete approximation of (5.3) needs to be solved for a sufficiently small time step Δt.

6 Numerical examples

For the numerical examples presented in this section, the mixed finite element formulation (4.3)–(4.4) and the discrete equation (4.5) have been implemented using the open source finite element library `FEniCS` [2], whereas the discrete transport equation (5.3) is implemented in python by own made routines. All the examples of this section are made in 2D. However, the model, numerical scheme and analysis are also valid in 3D. For the first example, the approximation of the fixed-point equation (4.5) is executed until the iterative process reaches the stopping criterion

$$\frac{|\mathbf{coeff}^{(m+1)} - \mathbf{coeff}^{(m)}|}{|\mathbf{coeff}^{(m+1)}|} < \mathtt{tol},$$

$\texttt{dof}_1$	h	$\texttt{e}(\boldsymbol{t})$	$\texttt{r}(\boldsymbol{t})$	$\texttt{e}(\boldsymbol{\sigma})$	$\texttt{r}(\boldsymbol{\sigma})$	$\texttt{e}(\boldsymbol{u})$	$\texttt{r}(\boldsymbol{u})$
421	0.5	0.4130	-	1.8442	-	0.0694	-
1657	0.25	0.2074	0.9940	0.9491	0.9583	0.0344	1.0106
6577	0.125	0.1016	1.0292	0.4743	1.0009	0.0170	1.0173
26209	0.0625	0.0501	1.0205	0.2361	1.0064	0.0085	1.0042
104641	0.0312	0.0248	1.0112	0.1176	1.0049	0.0042	1.0008
$\texttt{dof}_2$	Iter	$\texttt{e}(\gamma)$	$\texttt{r}(\gamma)$	$\texttt{e}(\widetilde{\boldsymbol{\sigma}})$	$\texttt{r}(\widetilde{\boldsymbol{\sigma}})$	$\texttt{e}(\widetilde{p})$	$\texttt{r}(\widetilde{p})$
45	6	0.3603	-	2.2614	-	0.2574	-
169	6	0.1841	0.9687	1.2430	0.8634	0.1303	0.9818
657	6	0.0733	1.3278	0.6287	0.9834	0.0654	0.9952
2593	6	0.0275	1.4146	0.3152	0.9961	0.0327	0.9988
10305	6	0.0100	1.4567	0.1577	0.9990	0.0164	0.9997

Table 6.1: Example 1. Error history for the mixed finite element approximations of the systems (4.3) and (4.4), respectively, computed with $\ell = 0$. The degrees of freedom of the spaces $\mathbb{X}_h \times \mathbb{V}_h$ ($\texttt{dof}_1$) and $\mathbf{M}_h \times Q_h$ ($\texttt{dof}_2$), meshsizes and fixed-point iterations (Iter) are informed on each mesh refinement.

where $\mathbf{coeff}^{(m)}$ is the vector of coefficients of $\widetilde{p}_h$ at the iteration $m \in \mathbb{N}$, $|\cdot|$ is the euclidean norm of $\mathbb{R}^{\dim(Q_h)}$, and $\texttt{tol}$ is a positive number. For the rest of the examples, equation (4.5) is solved by the Newton-Raphson algorithm with null initial guess, with absolute and relative tolerance of 10^{-10}, and the solution of tangent systems resulting from the linearization is carried out with the multifrontal massively parallel sparse direct solver `MUMPS` [7]. In addition, the integral condition (3.6) is added to the system (4.4) by means of a Lagrange multiplier. The following constitutive functions have been used for the numerical examples:

$$\mu(\phi) = \mu_0(1-\phi/\phi_{\mathrm{m}})^{-2.5}, \qquad \mu_0 = 100\,[\mathrm{kg/(m\,s)}], \quad \phi_{\mathrm{m}} = 1.2,$$
$$\xi(\phi) = \xi_0(1-\phi)\exp(-4.5\phi), \quad \xi_0 = 1/360\,[\mathrm{m^2/(Pa\,s)}].$$

Note that the Darcy function used is the decreasing function $D(\phi) = \xi_0\exp(-4.5\phi)$, where $D(0) = \xi_0$. Other parameters are $g = 9.81\,[\mathrm{m/s^2}]$, $\rho_{\mathrm{f}} = 950\,[\mathrm{kg/m^3}]$, $\rho_{\mathrm{s}} = 1000\,[\mathrm{kg/m^3}]$, and $\phi_{\max} = 0.95$. For all examples, we have set $\sigma_{\mathrm{e}} \equiv 0$ and time step used in all the examples is $\Delta t = 0.02\,\mathrm{s}$.

Example 1. Accuracy verification

We test the order of accuracy of the numerical scheme used to approximate the decoupled system (4.3)–(4.4) in the unit square $\Omega = (0,1)^2$ for a given smooth volume fraction $\phi < \phi_{\max}$. For this purpose, we set $\mu_0 = 1$ and $\xi_0 = 1$, and consider

$$\phi(\boldsymbol{x}) = 1 - 0.9\exp(x_1x_2(x_1-1)(x_2-1)) \quad \forall\,\boldsymbol{x} \in \Omega.$$

Furthermore, we use the manufactured exact solution:

$$\boldsymbol{u}(\boldsymbol{x}) = \begin{pmatrix} \sin(\pi x_1)x_2(x_2-1) \\ -x_1(x_1-1)\sin(\pi x_2) \end{pmatrix},$$
$$\widetilde{p}(\boldsymbol{x}) = \cos(\pi x_1) + \cos(\pi x_2),$$

for which the right hand sides of (4.3c) and (4.4b) have to be modified by adding the terms

$$\mathbf{f}_1(\boldsymbol{w}_h,\boldsymbol{\eta}_h) := \int_\Omega \Big(-\mathbf{div}(\mu(\phi)\boldsymbol{e}(\boldsymbol{u})) + \nabla\widetilde{p} - \mathbf{g}(\phi)\Big)\cdot \boldsymbol{w}_h \qquad \forall\,(\boldsymbol{w}_h,\boldsymbol{\eta}_h)\in \mathbb{V}_h,$$
$$\mathbf{f}_2(q_h) := -\int_\Omega \mathrm{div}(\boldsymbol{u} - \xi(\phi)\nabla p)q_h \qquad \forall\, q_h \in Q_h,$$

respectively. The errors between each component of the numerical solution $\boldsymbol{t}_h$, $\boldsymbol{\sigma}_h$, $\boldsymbol{u}_h$, $\boldsymbol{\gamma}_h$, $\widetilde{\boldsymbol{\sigma}}_h$, $\widetilde{p}_h$, computed with a meshsize h, and the exact manufactured solution, are given by

$$\mathtt{e}(\boldsymbol{t}) := \|\boldsymbol{t}-\boldsymbol{t}_h\|_{0;\Omega}, \qquad \mathtt{e}(\boldsymbol{\sigma}) := \|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{\mathbf{div};\Omega}, \qquad \mathtt{e}(\boldsymbol{u}) := \|\boldsymbol{u}-\boldsymbol{u}_h\|_{0;\Omega},$$
$$\mathtt{e}(\boldsymbol{\gamma}) := \|\boldsymbol{\gamma}-\boldsymbol{\gamma}_h\|_{0;\Omega}, \qquad \mathtt{e}(\widetilde{\boldsymbol{\sigma}}) := \|\widetilde{\boldsymbol{\sigma}}-\widetilde{\boldsymbol{\sigma}}_h\|_{\mathrm{div};\Omega}, \qquad \mathtt{e}(\widetilde{p}) := \|\widetilde{p}-\widetilde{p}_h\|_{0;\Omega}.$$

Then, the rates of convergence between two numerical errors $\mathtt{e}$ and $\mathtt{e}'$ made with two consecutive meshsizes h and h', respectively, are

$$\mathtt{r}(\boldsymbol{t}) := \frac{\log(\mathtt{e}(\boldsymbol{t})/\mathtt{e}'(\boldsymbol{t}))}{\log(h/h')}, \quad \mathtt{r}(\boldsymbol{\sigma}) := \frac{\log(\mathtt{e}(\boldsymbol{\sigma})/\mathtt{e}'(\boldsymbol{\sigma}))}{\log(h/h')}, \quad \mathtt{r}(\boldsymbol{u}) := \frac{\log(\mathtt{e}(\boldsymbol{u})/\mathtt{e}'(\boldsymbol{u}))}{\log(h/h')},$$
$$\mathtt{r}(\boldsymbol{\gamma}) := \frac{\log(\mathtt{e}(\boldsymbol{\gamma})/\mathtt{e}'(\boldsymbol{\gamma}))}{\log(h/h')}, \quad \mathtt{r}(\widetilde{\boldsymbol{\sigma}}) := \frac{\log(\mathtt{e}(\widetilde{\boldsymbol{\sigma}})/\mathtt{e}'(\widetilde{\boldsymbol{\sigma}}))}{\log(h/h')}, \quad \mathtt{r}(\widetilde{p}) := \frac{\log(\mathtt{e}(\widetilde{p})/\mathtt{e}'(\widetilde{p}))}{\log(h/h')}.$$

The degrees of freedom for the system (4.3) and (4.4), denoted by `dof1` and `dof2`, respectively, and the errors and rates of convergence produced by successive mesh refinements, are reported in Table 6.1 (for $\ell = 0$) and 6.2 (for $\ell = 1$). The number of iterations made by the fixed-point algorithm with a tolerance of $\mathtt{tol} = 10^{-6}$ was 6 for all the examples, for $\ell = 0$ and $\ell = 1$. From Table 6.1, we observe that for all variables the error decreases as the mesh size is smaller and the rates of convergence tend to 1, with the exception of $\boldsymbol{\gamma}$ whose rate of convergence is greater than 1. For the case of polynomial degrees $\ell = 1$ in Table 6.2, the errors behave as expected with rates of convergence approaching 2 as Theorem 4.7 states. Hence, both examples show that for $\ell = 0$ and $\ell = 1$ the errors produced with the scheme (4.3) and (4.4), and the fixed-point strategy, are of order $\mathcal{O}(h^{\ell+1})$, the optimum established in Theorem 4.7. Additionally, in Figure 6.2 we show the numerical solution of the six unknowns produced in the accurcy test.

Examples 2 and 3

For the second and third example we simulate the transient case, the two-dimensional batch sedimentation of flocculated particles in a channel with inclined walls. The domain considered in both simulations is a truncated cone of hight 2 meters and 6 meters wide, and represents the vertical cross-section of a channel. In Example 2 we start the simulation with an initial concentration

$$\phi^0(\boldsymbol{x}) = \begin{cases} 0.1 & \text{if } x_2 < 1.5\,\mathrm{m}, \\ 0 & \text{if } x_2 \geq 1.5\,\mathrm{m}, \end{cases}$$

where below the level $x_2 = 1.5$ meters the concentration is homogeneous and above it there is only clear water. For the Example 3 we consider an initial condition whose concentration of solid particles is located above the line $3x_2 - 2x_1 = 3$, at the left top corner of the domain

$$\phi^0(\boldsymbol{x}) = \begin{cases} 0.1 & \text{if } x_2 \geq \frac{2}{3}(x_1+3), \\ 0 & \text{otherwise.} \end{cases}$$

In Figure 6.3, left column, we show the simulation of Example 2 at four time points $t = 4$, 6, 8 and 10 seconds. We observe that as soon as the time evolves, the sedimentation process takes place, and the

$\mathtt{dof}_1$	h	$\mathtt{e}(\boldsymbol{t})$	$\mathtt{r}(\boldsymbol{t})$	$\mathtt{e}(\boldsymbol{\sigma})$	$\mathtt{r}(\boldsymbol{\sigma})$	$\mathtt{e}(\boldsymbol{u})$	$\mathtt{r}(\boldsymbol{u})$
889	0.5	0.0538	-	0.2887	-	0.0134	-
3505	0.25	0.0136	1.9824	0.0744	1.957	0.0034	1.9879
13921	0.125	0.0035	1.9538	0.0188	1.9826	0.0008	1.9989
55489	0.0625	0.0009	1.9738	0.0047	1.9907	0.0002	2.0003
221569	0.0312	0.0002	1.9874	0.0012	1.9953	0.0001	2.0002
$\mathtt{dof}_2$	Iter	$\mathtt{e}(\boldsymbol{\gamma})$	$\mathtt{r}(\boldsymbol{\gamma})$	$\mathtt{e}(\widetilde{\boldsymbol{\sigma}})$	$\mathtt{r}(\widetilde{\boldsymbol{\sigma}})$	$\mathtt{e}(\widetilde{p})$	$\mathtt{r}(\widetilde{p})$
137	6	0.0438	-	0.5606	-	0.0391	-
529	6	0.0122	1.8458	0.1345	2.059	0.0099	1.9797
2081	6	0.0035	1.7882	0.0338	1.9918	0.0025	1.9946
8257	6	0.001	1.8826	0.0085	1.9978	0.0006	1.9986
32897	6	0.0002	1.9415	0.0021	1.9995	0.0002	1.9997

Table 6.2: Example 1. Error history for the mixed finite element approximations of the systems (4.3) and (4.4), respectively, computed with $\ell = 1$. The degrees of freedom of the spaces $\mathbb{X}_h \times \mathbb{V}_h$ ($\mathtt{dof}_1$) and $\mathbf{M}_h \times Q_h$ ($\mathtt{dof}_2$), meshsizes and fixed-point iterations (Iter) are informed on each mesh refinement.

suspended solid particles move downward concentrating at the bottom of the channel, as expected. In accordance with the no-slip boundary condition, the solid-fluid interface moves faster at the middle of the channel and the concentration of solid particles increases towards the side walls and bottom. Furthermore, on each time iteration, the conservation of mass is fulfilled. In addition, in the second column of Figure 6.3, we show the simulation of Example 3 at the time points $t = 5$, 15, 20 and 35 seconds. We observe the downward falling of the solid particles which tend to concentrate at the boundaries and move towards the bottom. A layer of concentrated particles is clearly divided from a region of low concentration that shows the path that particles have followed in the sedimentation process. The movement of the solid particles is not only vertical but also moves to the right due accumulation of particles at the left side wall. In this simulation the conservation of mass is also satisfied.

Examples 4 and 5

In the fourth and fifth example, we perform the batch sedimentation on an inclined channel of rectangular cross-section with an angle of inclination of 30° degrees with respect to the vertical. The length and width of the rectangular domain are 10 and 2 meters, respectively. In Example 4 the initial condition chosen is

$$\phi^0(\boldsymbol{x}) = \begin{cases} 0.1 & \text{if } x_2 \leq 8, \\ 0 & \text{if } x_2 > 8, \end{cases}$$

and for Example 5 we consider an initial concentration varying with the width of the rectangle given by

$$\phi^0(\boldsymbol{x}) = \begin{cases} 0.025 + 0.0433x_1 - 0.025x_2 & \text{if } x_2 \leq 8, \\ 0 & \text{if } x_2 > 8. \end{cases}$$

In Figure 6.3, first row, we show the simulation of Example 4 at the time points $t = 6$, 9 and 12 seconds. We observe that the solid particles tend to settle towards the down side wall (right wall) and

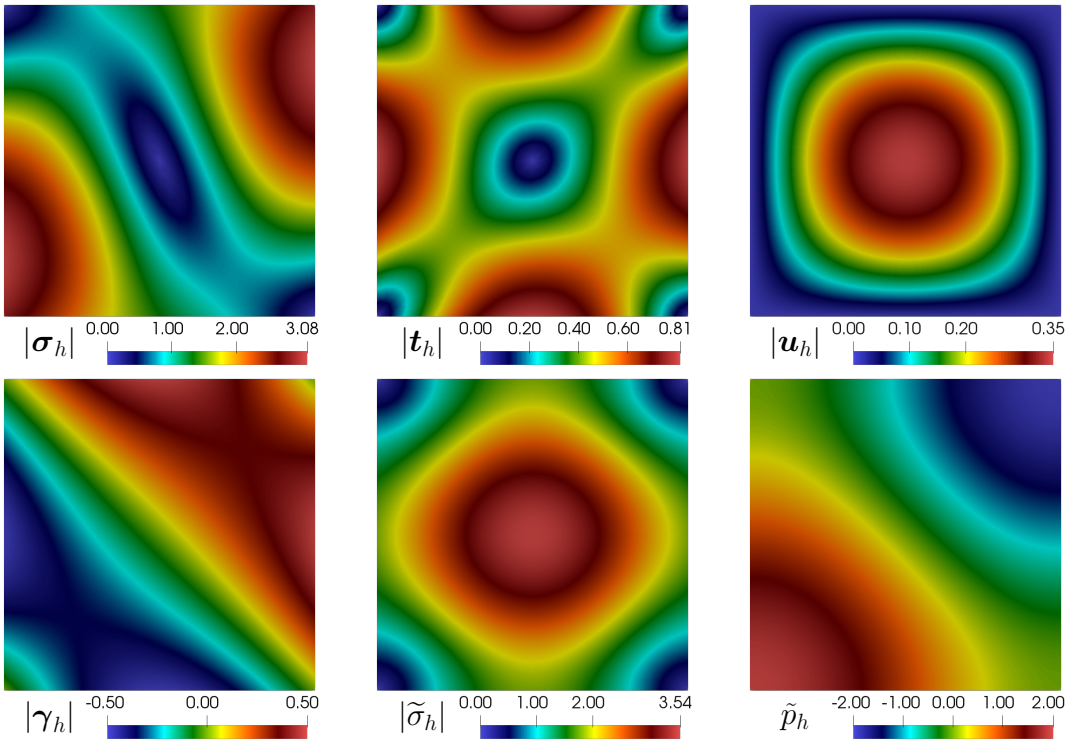

Figure 6.2: Example 1: Magnitud plots of the tensorial and vector components of the numerical solution of the coupled system (4.3)–(4.4) and plot of $\widetilde{p}$ for $h = 0.03125$.

bottom as expected. However, at the top side wall (left wall) near the line $x_2 = 8$ the solid particles do not separate from the clear water immediately, but begin to settle at a slower pace, which can be explained as a consequence of the no-slip boundary condition. Example 5 is shown in the bottom row of Figure 6.3, for the time points $t = 6$, 15 and 20 seconds. In this case the initial concentration is zero at the left side wall and increases continuously up to 0.1 at the right side wall. As in the previous example the concentration increases towards the right side and bottom wall but due to the initial concentration, the particles tend to separate from the left side wall. Already at $t = 20\,\mathrm{s}$, the solid-liquid is located near the left side wall and only a small amount of particles at the middle of the domain remain to settle. In both examples, the conservation of mass is fulfilled.

7 Discussion

We have introduced a model for multidimensional sedimentation based on the solid phase velocity which can be seen as a simplified version of [45]. From the model consisting of three unknowns and equations, we have studied the solutions of the decoupled velocity-pressure equations by using mixed finite element approaches. The uniqueness of the decoupled system is shown depending on the assumption that the function ξ is bounded and away from zero. Despite of the fact that the analysis of existence and uniqueness of solutions relies on this condition, the case when $\xi(\phi) = 0$ reduces to the classical Stokes system which has been widely analyzed in the literature [42]. One of the main advantages of having a mixed formulation is that extra variables, such as the pseudostress and vorticity tensors, can be recovered as a component of the solution of the problem. Such variables may have practical interest in the study of, for instance, the viscosity of mixtures [49]. Although the finite element subspaces for system (4.3) are based on PEERS finite elements [9], other suitable election

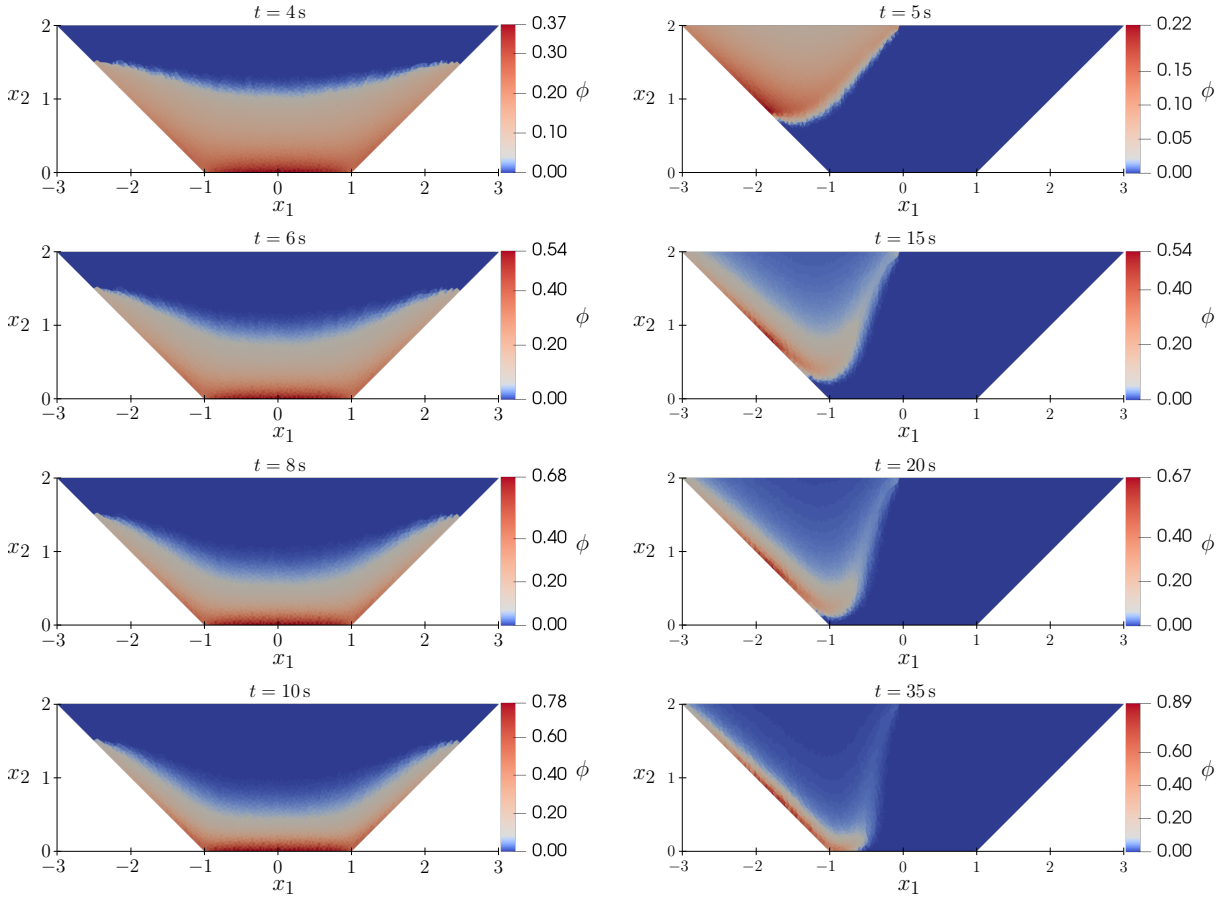

Figure 6.3: Examples 2 (left column) and 3 (right column): Batch sedimentation in a vessel with inclined walls at four different time points. Both examples were computed with the same mesh with $h = 0.0803$.

could be the Arnold-Falk-Winther subspaces [10], which are also stable.

A finite-volume method making use of an upwind flux approximation over a dual mesh is proposed to get conservation of mass for the solids volume fraction when coupling to the mixed finite element scheme for the velocity, pressure and additional variables. From the numerical examples we conclude that the coupled numerical scheme performs well, it is estable under the sufficiently small time step and gives physically relevant numerical solutions (cf., e.g., [14, 21]). Despite of the simple geometries considered in the five examples, the scheme can be implemented for domains with more involved geometries and even in three-dimensions. Another advantage of our model, as compared with, for example, the classical $\boldsymbol{q}$-model [23], is that the transport equation is linear in ϕ and a simple upwind scheme is sufficient to obtain physically relevant approximate solutions.

The model equations, analysis and numerical scheme can be extended to include, for example, a more general viscous stress tensor (eventually non-linear in $\boldsymbol{u}$), a non-zero boundary condition for the velocity field or to consider constitutive functions μ and ξ being able to assume zero values. Further studies can be done to compare the model presented here with related systems, and also with problems such as the one in [23].

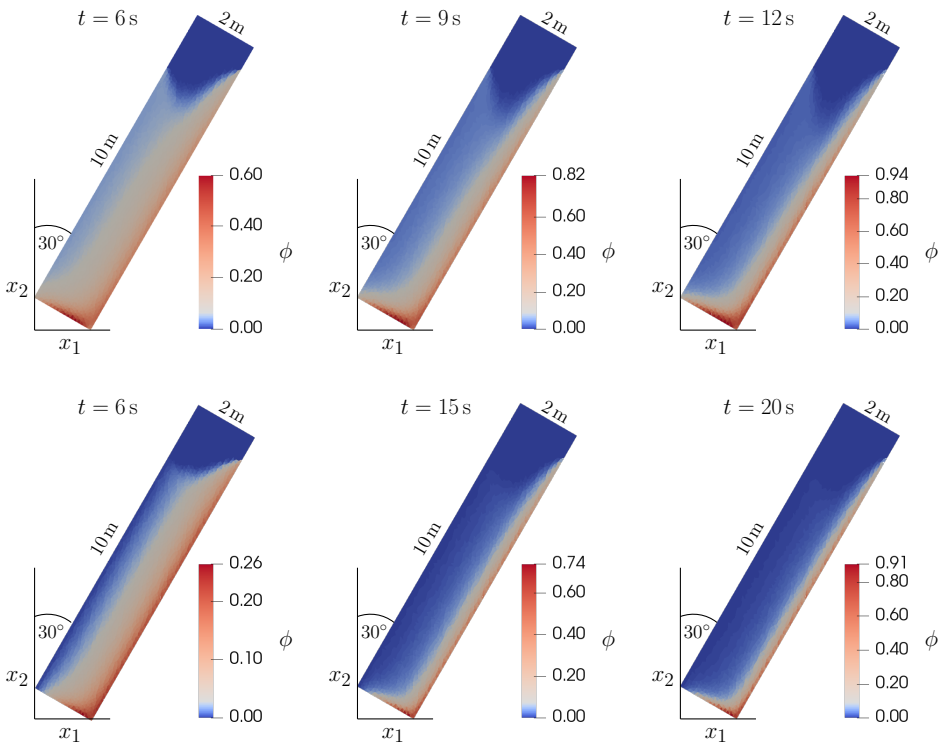

Figure 6.4: Examples 4 (top row) and 5 (bottom row): Sedimentation in inclined channel at four different time points. Both examples were computed with the same mesh with $h = 0.1569$.

References

[1] K. Abood, T. Das, D.R. Lester, S.P. Usher, A.D. Stickland, C. Rees, N. Eshtiaghi, and D.J. Batstone, *Characterising sedimentation velocity of primary waste water solids and effluents*. Water Research 219 (2022), 118555.

[2] M.S Alnæs, J. Blechta, J. Hake, A. Johansson, B. Kehlet, A. Logg, C. Richardson, J. Ring, M.E Rognes, and G.N Wells, *The fenics project version 1.5*. Arch. Numer. Softw. 3 (2015), no. 100, 9–23.

[3] M. Alvarez, G.N. Gatica, and R. Ruiz-Baier, *A mixed-primal finite element approximation of a sedimentation–consolidation system*. Math. Models Methods Appl. Sci. 26 (2016), no. 05, 867–900.

[4] M. Alvarez, G.N. Gatica, and R. Ruiz-Baier, *A vorticity-based fully-mixed formulation for the 3D Brinkman–Darcy problem*. Comput. Methods in Appl. Mech. Eng. 307 (2016), 68–95.

[5] M. Alvarez, G.N. Gatica, and R. Ruiz-Baier, *A mixed-primal finite element method for the coupling of brinkman–darcy flow and nonlinear transport*. IMA J. Numer. Anal. 41 (2020), no. 1, 381–411.

[6] M. Alvarez, G.N. Gatica, and R. Ruiz–Baier, *An augmented mixed-primal finite element method for a coupled flow-transport problem*. ESAIM: Math. Model Numer. Anal. 49 (2015), no. 5, 1399–1427.

[7] P.R. Amestoy, I.S. Duff, and J.-Y. L-Excellent, *Multifrontal parallel distributed symmetric and unsymmetric solvers*. Comput. Methods Appl. Mech. Engrg. 184 (2000), no. 2-4, 501–520.

[8] G. Anestis, Eine eindimensionale Theorie der Sedimentation in Absetzbehältern veränderlichen Querschnitts und in Zentrifugen. PhD thesis, TU Vienna, Austria, 1981.

[9] D.N. Arnold, F. Brezzi, and J. Douglas, *PEERS: A new mixed finite element for plane elasticity*. Jpn. J. Appl. Math. 1 (1984), 347–367.

[10] D.N. Arnold, R.S. Falk, and R. Winther, *Mixed finite element methods for linear elasticity with weakly imposed symmetry*. Math. Comput. 76 (2007), no. 260, 1699–1724.

[11] G.A. Benavides, S. Caucao, G.N. Gatica, and A.A. Hopper, *A Banach spaces-based analysis of a new mixed-primal finite element method for a coupled flow-transport problem*. Comput. Methods Appl. Mech. Engrg. 371 (2020), 113285.

[12] G.A. Benavides, S. Caucao, G.N. Gatica, and A.A. Hopper, *A new non-augmented and momentum-conserving fully-mixed finite element method for a coupled flow-transport problem*. Calcolo 59 (2022), no. 1, Paper No. 6.

[13] F. Blanchette, T. Peacock, and J.W.M. Bush, *The boycott effect in magma chambers*. Res. Lett. 31 (2004), L05611.

[14] A. Borhan and A. Acrivos, *The sedimentation of nondilute suspensions in inclined settlers*. Phys. Fluids 31 (1988), no. 12, 3488.

[15] R. Bürger, J. Careaga, and S. Diehl, *Entropy solutions of a scalar conservation law modeling sedimentation in vessels with varying cross-sectional area*. SIAM J. Appl. Math. 77 (2017), no 2, 789–811.

[16] R. Bürger, J. Careaga, and S. Diehl, *A simulation model for settling tanks with varying cross-sectional area*. Chem. Eng. Commun. 204 (2017), no. 11, 1270–1281.

[17] R. Bürger, J.J.R. Damasceno, and K.H. Karlsen, *A mathematical model for batch and continuous thickening of flocculated suspensions in vessels with varying cross-section*. Int. J. Miner. Process. 73 (2004), 183–208.

[18] R. Bürger, S. Diehl, S. Farås, and I. Nopens, *On reliable and unreliable numerical methods for the simulation of secondary settling tanks in wastewater treatment*. Computers Chem. Eng. 41 (2012), 93–105.

[19] R. Bürger, S. Evje, K.H. Karlsen, and K.-A. Lie, *Numerical methods for the simulation of the settling of flocculated suspensions*. Chem. Eng. J. 80 (2000), 91–104.

[20] R. Bürger, K.H. Karlsen, and J.D. Towers, *A model of continuous sedimentation of flocculated suspensions in clarifier-thickener units*. SIAM J. Appl. Math. 65 (2005), 882–940.

[21] R. Bürger, R. Ruiz-Baier, K. Schneider, and H. Torres, *A multiresolution method for the simulation of sedimentation in inclined channels*. Int. J. Numer. Anal. Model. 9 (2012), no. 3, 479–504.

[22] R. Bürger, R. Ruiz-Baier, and H. Torres, *A stabilized finite volume element formulation for sedimentation-consolidation processes*. SIAM J. Sci. Comput. 34 (2012), B265–B289.

[23] R.Bürger, W.L. Wendland, and F. Concha, *Model equations for gravitational sedimentation-consolidation processes*. ZAMM Z. Angew. Math. Mech. 80 (2000), 79–92.

[24] M.C. Bustos, F. Concha, R. Bürger, and E.M. Tory, Sedimentation and Thickening: Phenomenological Foundation and Mathematical Theory, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1999.

[25] J.-Ph. Chancelier, M. Cohen de Lara, and F. Pacard, *Analysis of a conservation PDE with discontinuous flux: a model of settler*. SIAM J. Appl. Math. 54 (1994), no. 4, 954–995.

[26] W. Chun-Liang and Z. Jie-Min, *Eulerian simulation of sedimentation flows in vertical and inclined vessels*. Chinese Phys. 14 (2005), no. 3, 620–627.

[27] E. Colmenares, G.N. Gatica, and J.C. Rojas, *A Banach spaces-based mixed-primal finite element method for the coupling of brinkman flow and nonlinear transport*. Calcolo 59 (2022), Paper No. 51.

[28] T. Kanti Deb, N. Lebaz, M. Sinan Ozdemir, R. Govoreanu, A. Mhamdi, G. Sin, and N. Sheibat-Othman, *Monitoring and modeling of creaming in oil-in-water emulsions*. Ind. Eng. Chem. Res. 61 (2022), no. 13, 4638–4647.

[29] A. Deininger, E. Holthausen, and P.A. Wilderer, *Velocity and solids distribution in circular secondary clarifiers: Full scale measurements and numerical modelling*. Water Res. 32 (1998), no. 10, 2951–2958.

[30] J.E. Dickinson and K.P. Galvin, *Fluidized bed desliming in fine particle flotation – part I*. Chem. Eng. Sci. 108 (2014), 283–298.

[31] S. Diehl, *Dynamic and steady-state behavior of continuous sedimentation*. SIAM J. Appl. Math. 57 (1997), no. 4, 991–1018.

[32] S. Diehl, *A regulator for continuous sedimentation in ideal clarifier-thickener units*. J. Eng. Math. 60 (2008), 265–291.

[33] S. Diehl, *The solids-flux theory – confirmation and extension by using partial differential equations*. Water Res. 42 (2008), no. 20, 4976–4988.

[34] E. Doroodchi, D.F. Fletcher, and K.P Galvin, *Influence of inclined plates on the expansion behaviour of particulate suspensions in a liquid fluidised bed*. Chem. Eng. Sci. 59 (2004), no. 17, 3559–3567.

[35] D.A. Drew and S.L. Passman, Theory of Multicomponent Fluids, vol. 135, Springer-Verlag, New York, 1999.

[36] G.A. Ekama, J.L. Barnard, F.W. Günthert, P. Krebs, J.A. McCorquodale, D.S. Parker, and E.J. Wahlberg, Secondary Settling Tanks: Theory, Modelling, Design and Operation, IAWQ scientific and technical report no. 6, International Association on Water Quality, England, 1997.

[37] G.A. Ekama and P. Marais, *Assessing the applicability of the 1D flux theory to full-scale secondary settling tank design with a 2D hydrodynamic model*. Water Res. 38 (2004), 495–506.

[38] K.P. Galvin, N.G. Harvey, and J.E. Dickinson, *Fluidized bed desliming in fine particle flotation – part III flotation of difficult to clean coal*. Minerals Eng. 66-68 (2014), 94–101.

[39] K.P. Galvin, A. Callen, J. Zhou, and E. Doroodchi, *Performance of the reflux classifier for gravity separation at full scale*. Minerals Eng. 18 (2005), no. 1, 19–24.

[40] H. Gao and M.K. Stenstrom, *The influence of wind in secondary settling tanks for wastewater treatment – a computational fluid dynamics study. part II: Rectangular secondary settling tanks*. Water Environ. Res. 92 (2020), no. 4, 551–561.

[41] P. Garrido, R. Burgos, F. Concha, and R. Bürger, *Settling velocities of particulate systems: 13. A simulator for batch and continuous sedimentation of flocculated suspensions*. Int. J. Miner. Process. 73 (2004), 131–144.

[42] G.N. Gatica, A Simple Introduction to the Mixed Finite Element Method, Springer International Publishing, 2014.

[43] G.N. Gatica, N. Heuer, and S. Meddahi, *On the numerical analysis of nonlinear twofold saddle point problems*. IMA J. Numer. Anal. 23 (2003), no. 2, 301–330.

[44] G.N. Gatica, R. Oyarzúa, R. Ruiz-Baier, and Y.D. Sobral, *Banach spaces-based analysis of a fully-mixed finite element method for the steady-state model of fluidized beds*. Comput. Math. Appl. 84 (2021), 244–276.

[45] K. Gustavsson and J. Oppelstrup, *Consolidation of concentrated suspensions – numerical simulations using a two-phase fluid model*. Comput. Visual. Sci. 3 (2000), no. 1–2, 39–45.

[46] D. Kleine and B.D. Reddy, *Finite element analysis of flows in secondary settling tanks*. Int. J. Numer Methods Eng. 64 (2005), no. 7, 849–876.

[47] M. Latsa, D. Assimacopoulos, A. Stamou, and N. Markatos, *Two-phase modeling of batch sedimentation*. Appl. Math. Model. 23 (1999), no. 12, 881–897.

[48] H. Laux and T. Ytrehus., *Computer simulation and experiments on two-phase flow in an inclined sedimentation vessel*. Powder Technol. 94 (1997), no. 1, 35–49.

[49] D.R. Lester, M. Rudman, and P.J. Scales, *Macroscopic dynamics of flocculated colloidal suspensions*. Chem. Eng. Sci. 65 (2010), no. 24, 6362–6378.

[50] D.N. Madge, J. Romero, and W.L. Strand, *Process reagents for the enhanced removal of solids and water from oil sand froth*. Minerals Eng. 18 (2005), no. 2, 159–169.

[51] S.J. McCaffery, L. Elliott, and D.B. Ingham, *Two-dimensional enhanced sedimentation in inclined fracture channels*. Math. Engrg. 7 (1998), 97–125.

[52] M.S. Nigam, *Numerical simulation of buoyant mixture flows*. Int. J. Multiph. Flow 29 (2003), no. 6, 983–1015.

[53] K.V. Parchevsky, *Numerical simulation of sedimentation in the presence of 2D compressible convection and reconstruction of the particle-radius distribution function*. J. Eng. Math. 41 (2001), 203–219.

[54] R. Rao, L. Mondy, A. Sun, and S. Altobelli, *A numerical and experimental study of batch sedimentation and viscous resuspension*. Int. J. Numer. Methods Fluids 39 (2002), no. 6, 465–483.

[55] C. Reyes, C.F. Ihle, F. Apaz, and L.A. Cisternas, *Heat-assisted batch settling of mineral suspensions in inclined containers*. Minerals 9 (2019), no. 4, 228.

[56] U. Schaflinger, *Experiments on sedimentation beneath downward-facing inclined walls*. Int. J. Multiph. Flow 11 (1985), no. 2, 189–199.

[57] E.S.G. Shaqfeh and A. Acrivos, *The effects of inertia on the buoyancy-driven convection flow in settling vessels having inclined walls*. Phys. Fluids 29 (1986), no. 12, 3935.

[58] E.S.G. Shaqfeh and A. Acrivos, *The effects of inertia on the stability of the convective flow in inclined particle settlers*. Phys. Fluids 30 (1987), no. 4, 960.

[59] E.S.G. Shaqfeh and A. Acrivos, *Enhanced sedimentation in vessels with inclined walls: Experimental observations*. Phys. Fluids 30 (1987), no. 7, 1905.

[60] R. Stephen Sparks, H.E. Huppert, T. Koyaguchi, and M.A. Hallworth, *Origin of modal and rhythmic igneous layering by sedimentation in a convecting magma chamber*. Nature 361 (1993), no. 6409, 246–249.

[61] J. Su, L. Wang, Y. Zhang, and Z. Gu, *A numerical study on influent flow rate variations in a secondary settling tank*. Processes 7 (2019), no. 12, 884.

[62] M. Ungarish, Hydrodynamics of Suspensions, Springer Berlin Heidelberg, 1993.

Centro de Investigación en Ingeniería Matemática (CI^{2}MA)

PRE-PUBLICACIONES 2022 - 2023

2022-27 Rodolfo Araya, Cristian Cárcamo, Abner Poza: *A stabilized finite element method for the Stokes–Temperature coupled problem*

2022-28 Jorge Albella, Rodolfo Rodríguez, Pablo Venegas: *Numerical approximation of a potentials formulation for the elasticity vibration problem*

2022-29 Liliana Camargo, Bibiana López-Rodríguez, Mauricio Osorio, Manuel Solano: *An adaptive and quasi-periodic HDG method for Maxwell's equations in heterogeneous media*

2022-30 Gabriel N. Gatica, Nicolas Nuñez, Ricardo Ruiz-Baier: *Mixed-primal methods for natural convection driven phase change with Navier-Stokes-Brinkman equations*

2022-31 Diego Paredes, Frederic Valentin, Henrique M. Versieux: *Revisiting the robustness of the multiscale hybrid-mixed method: the face-based strategy*

2022-32 Franz Chouly, Tom Gustafsson, Patrick Hild: *A Nitsche method for the elastoplastic torsion problem*

2022-33 Olusegun Adebayo, Stéphane P. A. Bordas, Franz Chouly, Raluca Eftimie, Gwenaël Rolin, Stéphane Urcun: *Modelling Keloids Dynamics: A Brief Review and New Mathematical Perspectives*

2022-34 Nicolás Carro, David Mora, Jesús Vellojín: *A finite element model for concentration polarization and osmotic effects in a membrane channel*

2022-35 Claudio I. Correa, Gabriel N. Gatica, Esteban Henriquez, Ricardo Ruiz-Baier, Manuel Solano: *Banach spaces-based mixed finite element methods for the coupled Navier–Stokes and Poisson–Nernst–Planck equations*

2022-36 Gabriel R. Barrenechea, Antonio Tadeu A. Gomes, Diego Paredes: *A Multiscale Hybrid Method*

2022-37 Rodrigo Abarca Del Rio, Fernando Campos, Dieter Issler, Mauricio Sepúlveda: *Study of Avalanche Models Using Well Balanced Finite Volume Schemes*

2023-01 Julio Careaga, Gabriel N. Gatica: *Coupled mixed finite element and finite volume methods for a solid velocity-based model of multidimensional settling*

Para obtener copias de las Pre-Publicaciones, escribir o llamar a: Director, Centro de Investigación en Ingeniería Matemática, Universidad de Concepción, Casilla 160-C, Concepción, Chile, Tel.: 41-2661324, o bien, visitar la página web del centro: `http://www.ci2ma.udec.cl`







Centro de Investigación en
Ingeniería Matemática (CI²MA)
Universidad de Concepción

Casilla 160-C, Concepción, Chile
Tel.: 56-41-2661324/2661554/2661316
http://www.ci2ma.udec.cl