A novel view on computations of steady flows of Bingham fluids using implicit constitutive relations

Jaroslav Hron[1]*, Josef Málek[1]*, Jan Stebel[2]*, Kryštof Touška[1]

[1] *Charles University, Faculty of Mathematics and Physics, Sokolovská 83, 18675 Prague 8, Czech Republic*
[2] *Technical University of Liberec, Studentská 1402/2, 46117 Liberec, Czech Republic*

SUMMARY

Within the framework of implicit constitutive relations we investigate steady flows of stress–power-law and Bingham fluids. This leads naturally to the setting of mixed formulations for the corresponding boundary value problems. We propose and compare several such formulations, their stable approximations and particular examples of stable finite element spaces resulting in regular (well posed) linearized problems. This systematic theoretical setting is then analyzed computationally and for the Bingham fluid we identify certain combinations of the implicit function formulation and mixed finite element approximations which result in problem solvable by the Newton method with mesh-refinement independent number of iterations.

Received . . .

KEY WORDS: incompressible fluid; non-Newtonian fluid; Bingham fluid; power-law fluid; stress–power-law fluid; implicit constitutive relation; finite element method; mixed formulation

1. INTRODUCTION

There are many flows of various fluids, such as polymeric liquids, powders, food materials, etc., that exhibit the formation of "dead-zones" - these are subdomains in which the fluid is merely rotating and translating, and in fact no real flow takes place inside such parts of the flow container. Such behavior of materials is usually described by the following dichotomy. If the shear stress is below certain (given) critical value in a part of the fluid domain, then this part moves as a rigid body. On the other hand, if the shear stress exceeds this critical value, the fluid behaves as a Navier–Stokes fluid or a power-law fluid, depending on the response characterized by a specific constitutive relation. The *critical shear stress* whose value plays a key role in the total response of the material is called the yield stress, and such material behavior is named *the presence of yield stress in a simple shear flow*, or more generally the presence of an activation criterion (in a simple shear flow). Since such response differs from the behavior of a Navier–Stokes (Newtonian) fluid, the presence of yield stress belongs among (significant) non-Newtonian phenomena. In fact, as a material particle can either be located in the dead-zone or flow according to the generalized Navier–Stokes equations, the above described fluid response is an interesting example of mixing. The importance of investigating such

*Correspondence to: Jaroslav Hron jaroslav.hron@mff.cuni.cz; Jan Stebel jan.stebel@tul.cz; Josef Málek josef.malek@mff.cuni.cz;

Contract/grant sponsor: J. Hron and J. Málek acknowledge the support of the project LL1202 in the programme ERC-CZ funded by the Ministry of Education, Youth and Sports of the Czech Republic. The work of J. Stebel was supported by the Ministry of Education, Youth and Sports under the project LO1201 in the framework of the targeted support of the "National Programme for Sustainability I".

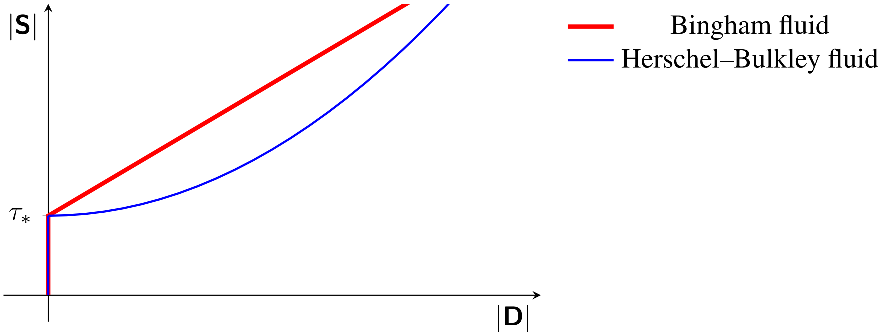


Figure 1. Yield stress response. Bingham and Herschel–Bulkley fluids.

activated materials is expressed for example in the survey [1] or in the collection [2] which was published on the 100 years anniversary of original investigation by E.C. Bingham [3].

The response of incompressible fluid is usually encoded in the deviatoric (traceless) part[†] $\mathbf{S}$ of the Cauchy stress tensor $\mathbf{T}$, i.e.

$$\mathbf{T} = m\mathbf{I} + \mathbf{S} \quad \text{where} \quad m := \frac{\operatorname{tr}\mathbf{T}}{\operatorname{tr}\mathbf{I}} \quad \text{and} \quad \mathbf{S} := \mathbf{T}^{\delta}. \tag{1}$$

With this notation, the above described fluid behavior is written in a full three-dimensional setting as follows, see [4]:

$$\begin{aligned} |\mathbf{S}| \le \tau_* &\Leftrightarrow |\mathbf{D}| = 0, \\ |\mathbf{S}| > \tau_* &\Leftrightarrow \mathbf{S} = \left(2\nu + \frac{\tau_*}{|\mathbf{D}|}\right)\mathbf{D}. \end{aligned} \tag{2}$$

Here, $\tau_* > 0$ is the yield stress (the critical value in which the activation takes place), $\mathbf{D}$ is the symmetric part of the velocity gradient $\nabla \boldsymbol{v}$, $\boldsymbol{v}$ being the velocity. The symbols $\mathbf{0}$ and $\mathbf{I}$ represent the zero and the identity tensors. Finally, the specific form of the generalized viscosity $\nu := \mathbb{R}_0^+ \mapsto \mathbb{R}_0^+$ distinguishes between the fluid of a Bingham or a Herschel–Bulkley type. If $\nu > 0$ is constant then the response described by (2) is associated with a Bingham fluid. If $\nu > 0$ for $\mathbf{D} \neq \mathbf{0}$ and depends on $|\mathbf{D}|^2 := \frac{1}{2}\operatorname{tr}\mathbf{D}^2$ polynomially, one talks about a Herschel–Bulkley fluid. More complex forms for the generalized viscosity can be however considered, see for example the list given in [5, part 6].

Other equivalent forms of (2) can be used, see for example [6, ch. 5, p. 171]:

$$\begin{aligned} |\mathbf{S}| \le \tau_* \quad &\Leftrightarrow \quad \mathbf{D} = \mathbf{0}, \\ |\mathbf{S}| > \tau_* \quad &\Leftrightarrow \quad \mathbf{D} = \left(1 - \frac{\tau_*}{|\mathbf{S}|}\right)\frac{\mathbf{S}}{2\nu}; \end{aligned} \tag{3}$$

alternatively,

$$\begin{aligned} |\mathbf{D}| = 0 \quad &\Leftrightarrow \quad |\mathbf{S}| \le \tau_*, \\ |\mathbf{D}| > 0 \quad &\Leftrightarrow \quad \mathbf{S} = \left(2\nu + \frac{\tau_*}{|\mathbf{D}|}\right)\mathbf{D}. \end{aligned} \tag{4}$$

Recently, Rajagopal and Srinivasa [7] (see also [8]) and Bulíček et. al. [9] observed that the response (2) (sketched in Figure 1) can be equivalently described through

$$2\nu\left(\tau_* + (|\mathbf{S}| - \tau_*)^+\right)\mathbf{D} = (|\mathbf{S}| - \tau_*)^+\,\mathbf{S}\,, \tag{5}$$

where $z^+ := \max\{z, 0\}$ for $z \in \mathbb{R}$. Thus, setting

$$\mathbf{G}_1(\mathbf{S}, \mathbf{D}) := 2\nu\left(\tau_* + (|\mathbf{S}| - \tau_*)^+\right)\mathbf{D} - (|\mathbf{S}| - \tau_*)^+\mathbf{S}, \tag{Bi-1}$$

[†] If $\mathbf{A}$ is a second order tensor, then its deviatoric part is defined through $\mathbf{A}^{\delta} := \mathbf{A} - \frac{\operatorname{tr}\mathbf{A}}{\operatorname{tr}\mathbf{I}}\mathbf{I}$, and $\operatorname{tr}\mathbf{A} := \sum_{i=1}^{d} A_{ii}$.

we can capture the material behavior described in (2) by implicit relation $\mathbf{G}_1(\mathbf{S},\mathbf{D}) = \mathbf{0}$. Using the equivalences (3) and (4) this can be reformulated as:

$$\mathbf{G}_2(\mathbf{S},\mathbf{D}) = \mathbf{0} \quad \text{with} \quad \mathbf{G}_2(\mathbf{S},\mathbf{D}) := 2\nu(\tau_* + |2\nu\mathbf{D}|)\mathbf{D} - |2\nu\mathbf{D}|\mathbf{S}, \tag{Bi-2}$$

$$\mathbf{G}_3(\mathbf{S},\mathbf{D}) = \mathbf{0} \quad \text{with} \quad \mathbf{G}_3(\mathbf{S},\mathbf{D}) := 2\nu(\tau_* + |2\nu\mathbf{D}|)\mathbf{D} - (|\mathbf{S}| - \tau_*)^+\mathbf{S}, \tag{Bi-3}$$

$$\mathbf{G}_4(\mathbf{S},\mathbf{D}) = \mathbf{0} \quad \text{with} \quad \mathbf{G}_4(\mathbf{S},\mathbf{D}) := 2\nu|\mathbf{S}|\mathbf{D} - (|\mathbf{S}| - \tau_*)^+\mathbf{S}. \tag{Bi-4}$$

We conclude that the Bingham and Herschel–Bulkley fluids are special cases of incompressible fluids described through an implicit constitutive equation of the form

$$\mathbf{G}(\mathbf{S},\mathbf{D}) = \mathbf{0}. \tag{6}$$

One of the main aims of this study is *to investigate problems associated with Bingham and Herschel–Bulkley fluids using the formulations* (Bi-1), (Bi-2), (Bi-3) *or* (Bi-4) *rather than* (2), see below for more comments concerning theoretical results for such problems.

Starting from the seminal work of Rajagopal [10], it has been observed that the framework of implicitly constituted materials provides an appropriate theoretical basis for *derivation/justification* of many models that have been used, mostly successfully, in various areas (engineering, chemistry, food processing) but that have been proposed in an ad hoc manner. It concerns both fluid and solid mechanics. To be more specific, we give a few examples.

In fluid mechanics, the relation $\mathbf{G}(\mathbf{T},\mathbf{D}) = \mathbf{0}$ is rich enough to include models where the generalized viscosity is a function of both the pressure (mean normal stress) and the shear rate. Recall that although the fact that the viscosity of many liquids changes dramatically with the pressure is known since the works of Barus [11] and Bridgman [12], the model cannot be deduced if one starts with the assumption that $\mathbf{T} = \tilde{\mathbf{T}}(\mathbf{D})$, see [10], [13].

In solid mechanics, more specifically in the theory of elasticity, the equation $\mathbf{G}(\mathbf{T},\mathbf{B}) = \mathbf{0}$ with $\mathbf{B} = \mathbf{F}\mathbf{F}^T$, where $\mathbf{F}$ stands for the deformation gradient, is an elegant way out for non-linear relations between the stress and the linearized strain $\boldsymbol{\varepsilon} := \frac{1}{2}(\nabla\boldsymbol{u} + (\nabla\boldsymbol{u})^T)$. Note that the traditional approaches in the non-linear elasticity, stemming from the assumption that $\mathbf{T} = \tilde{\mathbf{T}}(\mathbf{B})$ and investigating consequences that come from an additional assumption that the displacement gradient is small, end-up always with *linear* relations between the stress and the linearized strain $\boldsymbol{\varepsilon}$, see for example Rajagopal [14].

In continuum thermodynamics, the implicit constitutive theory helped naturally to equalize the role of thermodynamical fluxes (such as the heat flux or the dissipative part of the the Cauchy stress) and thermodynamical affinities (such as temperature gradient or $\mathbf{D}$). Such a viewpoint leads to the development of a thermodynamical basis for complete fluid models of Korteweg, Cahn–Hilliard and Allen–Cahn type, see [15], [16] and [17].

As shown above, see (5), the framework of implicitly constituted fluids is also a powerful framework for involving responses with activation or deactivation criteria (that have been studied earlier using calculus of multi-valued functions, theory of variational inequalities, etc.)

It is also worth mentioning that not only does the framework (6) include classical explicit constitutive models

$$\mathbf{S} = \tilde{\mathbf{S}}(\mathbf{D})\,, \tag{7}$$

but it also contains a new class of explicit non-linear models, namely those characterized by

$$\mathbf{D} = \tilde{\mathbf{D}}(\mathbf{S})\,. \tag{8}$$

This has interesting consequences as well. To give an example, let us first observe that a Navier–Stokes fluid characterized by

$$\mathbf{T} = m\mathbf{I} + 2\nu_*\mathbf{D} \qquad \text{with } \nu_* > 0 \tag{9}$$

can be written in the form (7) through $\mathbf{S} = 2\nu_*\mathbf{D}$ or in the form (8) as $\mathbf{D} = \frac{1}{2\nu_*}\mathbf{S}$. More interestingly, the standard power-law model

$$\mathbf{T} = m\mathbf{I} + 2\nu_*|\mathbf{D}|^{r-2}\mathbf{D} \qquad \text{with } \nu_* > 0 \text{ and } r \in (1,\infty) \tag{10}$$

can also be written in both forms, namely,

$$\mathbf{S} = 2\nu_*|\mathbf{D}|^{r-2}\mathbf{D} \Leftrightarrow \mathbf{D} = \left(\tfrac{1}{2\nu_*}\right)^{\frac{1}{r-1}} |\mathbf{S}|^{\frac{2-r}{r-1}}\mathbf{S} = \left(\tfrac{1}{2\nu_*}\right)^{\frac{1}{r-1}} |\mathbf{S}|^{r'-2}\mathbf{S}\,. \tag{11}$$

Here, $r' = r/(r-1)$ denotes the exponent dual to r.

This equivalence suggests to consider and study the following generalizations of power-law fluids, namely,

$$\mathbf{S} = 2\nu_*(1 + |\mathbf{D}|^2)^{\frac{r-2}{2}}\mathbf{D}\,, \tag{12}$$

versus

$$\mathbf{D} = \left(1 + \left(\frac{|\mathbf{S}|}{2\nu_*}\right)^2\right)^{\frac{r'-2}{2}} \frac{\mathbf{S}}{2\nu_*}\,. \tag{13}$$

Although these models are asymptotically (for $|\mathbf{S}|$ and $|\mathbf{D}|$ large) equivalent, (13) is not inversion of (12). Thus, (13) provides a new formula that is capable of capturing the same experimental data as (12). The advantages of this extended framework of power-law models have been investigated recently in [18] for various flows in special geometries. Thus, another aim of this study is *to compare the analytic solutions derived in [18] with the results of numerical simulations that are not based on the* ansatzes *used in derivation of analytical solutions*.

We also notice that the above equations (Bi-1)–(Bi-4), (12) and (13) belong to a *subclass* of fully implicit relations $\mathbf{G}(\mathbf{S},\mathbf{D}) = \mathbf{0}$; this subclass is characterized by the relation

$$\alpha(|\mathbf{D}|^2, |\mathbf{S}|^2)\mathbf{D} = \beta(|\mathbf{D}|^2, |\mathbf{S}|^2)\mathbf{S} \tag{14}$$

with α being positive and β non-negative. In the remaining part of the paper we restrict ourselves to implicit constitutive equations given by (14), i.e., $\mathbf{G}$ in (6) is of the form

$$\mathbf{G}(\mathbf{S},\mathbf{D}) = \alpha(|\mathbf{D}|^2, |\mathbf{S}|^2)\mathbf{D} - \beta(|\mathbf{D}|^2, |\mathbf{S}|^2)\mathbf{S}. \tag{15}$$

In what follows we neither consider the thermal effects nor inhomogeneity of the incompressible fluids. Hence the density and the temperature are constant and are equal to ϱ_* and θ_*. Thus, for a given $T > 0$ the governing equations in $(0,T) \times \Omega$, $\Omega \subset \mathbb{R}^d$ being an open bounded set, take the form

$$\begin{gathered} \varrho_* \left(\frac{\partial \boldsymbol{v}}{\partial t} + \operatorname{div}(\boldsymbol{v} \otimes \boldsymbol{v})\right) = \operatorname{div}\mathbf{S} + \nabla m + \varrho_* \boldsymbol{f}\,, \\ \operatorname{div}\boldsymbol{v} = 0\,, \\ \mathbf{G}(\mathbf{S},\mathbf{D}) = \mathbf{0}. \end{gathered} \tag{16}$$

Besides the existing success of implicit constitutive theory in theoretical foundation of continuum mechanics and thermodynamics, general implicit relations (6) have inspired mathematical analysts to generalize the traditional setting so that its renovation is very suitable for large-data theoretical results concerning existence of weak solution and its properties. More precisely, introducing

$$\mathcal{A} := \{(\mathbf{S},\mathbf{D}) \in \mathbb{R}^{3\times 3}_{\text{sym}} \times \mathbb{R}^{3\times 3}_{\text{sym}};\, \mathbf{G}(\mathbf{S},\mathbf{D}) = \mathbf{0}\}\,, \tag{17}$$

where $\mathbb{R}^{d\times d}_{\text{sym}}$ stands for the set of all symmetric matrices of order d, all the considered examples (Navier–Stokes fluids (9), power-law fluids (10) and (12), stress–power-law fluids (13), Bingham and Herschel–Bulkley fluids (6)) can be put into the following set of requirements:

(A1) *$\mathcal{A}$ contains the origin:* $(\mathbf{0},\mathbf{0}) \in \mathcal{A}$.

(A2) *$\mathcal{A}$ is a monotone graph:*

$$(\mathbf{S}_1 - \mathbf{S}_2)\cdot(\mathbf{D}_1 - \mathbf{D}_2) \geq 0 \qquad \text{for all } (\mathbf{S}_1,\mathbf{D}_1), (\mathbf{S}_2,\mathbf{D}_2) \in \mathcal{A}\,.$$

(A3) *$\mathcal{A}$ is a maximal monotone graph:* Let $(\mathbf{S},\mathbf{D}) \in \mathbb{R}^{3\times 3}_{\text{sym}} \times \mathbb{R}^{3\times 3}_{\text{sym}}$.

$$\text{If } (\bar{\mathbf{S}} - \mathbf{S})\cdot(\bar{\mathbf{D}} - \mathbf{D}) \geq 0 \quad \text{for all } (\bar{\mathbf{S}},\bar{\mathbf{D}}) \in \mathcal{A} \text{ then } (\mathbf{S},\mathbf{D}) \in \mathcal{A}.$$

(A4) *$\mathcal{A}$ is a r graph:* There is $c_* > 0$, $\alpha_* > 0$ and $r > 1$ such that

$$\mathbf{S}\cdot\mathbf{D} \geq -c_* + \alpha_*(|\mathbf{D}|^r + |\mathbf{S}|^{r'}) \qquad \text{for all } (\mathbf{S},\mathbf{D}) \in \mathcal{A}\,.$$

We refer the reader to Lemma 1.1 in [9] for the proof of the statement saying that the Bingham and Herschley–Bulkley fluids given by (5) with $\nu(|\mathbf{D}|) = \nu_0(1+|\mathbf{D}|^2)^{\frac{r-2}{2}}$ fulfill all the structural assumptions **(A1)–(A4)**.

The setting characterized by **(A1)–(A4)** is very transparent and makes the analysis presented recently in [19], [9], [20] very straightforward if one deals with the concept of weak solution. The following text should underline these statements.

Let us, for simplicity of next discussion, assume that $\boldsymbol{v} = \mathbf{0}$ on $(0,T)\times\partial\Omega$. Multiplying $(16)_1$ by $\boldsymbol{v}$, integrating the result over Ω and using the integration by parts, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \varrho_* \frac{|\boldsymbol{v}|^2}{2}\,\mathrm{d}x + \int_\Omega \mathbf{S}\cdot\mathbf{D}\,\mathrm{d}x = \int_\Omega \varrho_* \boldsymbol{f}\cdot\boldsymbol{v}\,\mathrm{d}x\,.$$

Using **(A4)** it implies that

$$\begin{aligned}\sup_t \int_\Omega |\boldsymbol{v}|^2\,\mathrm{d}x + \alpha_*\int_0^T\int_\Omega |\mathbf{D}|^r\,\mathrm{d}x\,\mathrm{d}t + \alpha_*\int_0^T\int_\Omega |\mathbf{S}|^{r'}\,\mathrm{d}x\,\mathrm{d}t \\ \leq C\left(\int_\Omega |\boldsymbol{v}_0|_2^2\,\mathrm{d}x + \int_0^T\int_\Omega |\boldsymbol{f}|^2\,\mathrm{d}x\,\mathrm{d}t + c_*\right).\end{aligned}\tag{18}$$

In [19], the authors considered the framework characterized by **(A1)–(A4)** with one important difference: instead of **(A2)** they require a certain kind of strict monotone property of the graph given by $\mathbf{G}$. This drawback has been removed and the most general results concerning large data existence of weak solution are established in [9] for evolutionary case[‡] and in [20] for systems without inertia or for steady flows. Without going to details (that are not necessary for the main part of the paper) the results can be characterized as follows:

> For arbitrary set of data involving $\Omega\subset\mathbb{R}^d$ with smooth boundary $\partial\Omega$, $T\in(0,\infty)$, $\boldsymbol{v}_0 \in L^2(\Omega)^d$, $\operatorname{div}\boldsymbol{v}_0 = 0$ in Ω and $\boldsymbol{v}_0\cdot\boldsymbol{n} = 0$ on $\partial\Omega$, and $\boldsymbol{f}\in L^2(0,T;L^2(\Omega)^d)$, there is long-time and large-data weak solution to (16) satisfying $\boldsymbol{v} = \mathbf{0}$ on $(0,T)\times\Omega$ and $\boldsymbol{v}(0,.) = \boldsymbol{v}_0$ in Ω provided that the graph $\mathcal{A}$ generated by $\mathbf{G}$ via the identification (17) fulfills the assumptions **(A1)–(A4)** and the function spaces generated by (18) and the Equation (16) are compactly embedded into $L^2(0,T;L^2(\Omega)^d)$, which happens if r appearing in **(A4)** satisfies $r > \frac{2d}{d+2}$.

The bound $r > \frac{2d}{d+2}$ is needed to treat (to take the limit in) the convective term and is required also for analysis of steady flows; for steady Stokes-like system the result holds for $r > 1$, see [20, section 2.5]. Also, in the evolutionary case to obtain the pressure integrable over $(0,T)\times\Omega$ one needs a Navier-slip type boundary conditions, see [9].

We wish to emphasize that the existence results established in [9] and [20] successfully use the two main advantages of the assumptions **(A1)–(A4)**: first, the quantities $\mathbf{S}$ and $\mathbf{D}$ occur in these assumptions in a symmetric way; and second, these assumptions are strongly related to tools developed in the theory of weak solutions. The approach developed in these studies starts from Galerkin finite-dimensional approximations for a suitably regularized problem. These Galerkin approximations are generated by the eigenvalues of a suitable elliptic operator. Recently, Diening, Kreuzer and Süli in [21, 22] considered steady flows and proved the existence result via the convergence of finite-element discretizations (the authors use strict monotone property in **(A2)**). We also refer to [23] and [24] for the proof of the above result in the case of classical power-law fluids (10). Regarding the analysis specifically focused on Bingham fluids we refer to [25, 26].

[‡] In [9], the functional space setting generated by **(A4)** is also extended to the Orlicz functions.

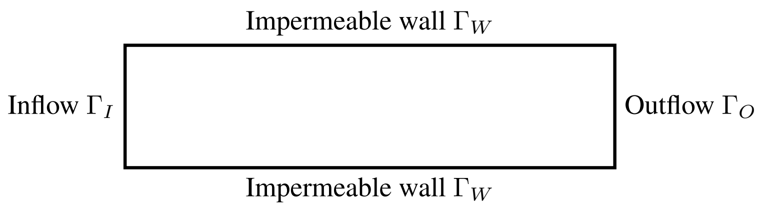


Figure 2. General setting of the boundary notation for the computations.

In the literature one can find a wide range of numerical approaches to solving flows of fluids of a Bingham type (sometimes called visco-plastic flows), see [27, 28, 29, 30, 31, 32, 33] for example.

A further aim of this study is *to analyze computationally several different discrete formulations and identify those that are stable*. We will include different combinations of the implicit constitutive form (see (Bi-1)–(Bi-4) above), different weak formulations and specific discretizations tested either on a problem with known analytical solution and on an established benchmark type configuration from existing literature. This will be discussed in the next sections.

2. WEAK FORMULATIONS

In this section, in order to focus on the structure of the problem given by the implicit constitutive relation, we neglect the inertia and then the system of governing equations reduces to the following problem: to find a triplet $(\mathbf{S}, \boldsymbol{v}, m)$ satisfying

$$-\operatorname{div}\mathbf{S} - \nabla m = \varrho_* \boldsymbol{f} \quad \text{and} \quad \mathbf{G}(\mathbf{S}, \mathbf{D}) = \mathbf{0} \text{ in } \Omega, \tag{19}$$

where $\Omega \subset \mathbb{R}^d$, $d \in \{2, 3\}$, is a bounded domain with Lipschitz boundary, $\varrho_* > 0$ is the density (let us assume for simplicity that $\varrho_* = 1$), $\boldsymbol{f} \in L^2(\Omega; \mathbb{R}^d)$ is the body force, and $\mathbf{G} : \mathbb{R}^{d\times d}_{\text{sym}} \times \mathbb{R}^{d\times d}_{\text{sym}} \to \mathbb{R}^{d\times d}_{\text{sym}}$ is a continuous tensor function, of the form (15), which automatically includes $\operatorname{div}\boldsymbol{v} = \operatorname{tr}\mathbf{D} = 0$; this incompressibility condition can be recovered by taking the trace of the equation $\mathbf{G}(\mathbf{S}, \mathbf{D}) = \mathbf{0}$, under the assumption that the coefficient $\alpha(|\mathbf{D}|, |\mathbf{S}|)$ in (15) is nonzero.

We assume that the boundary $\partial\Omega$ (see Figure 2) is divided into three mutually disjoint parts Γ_I, Γ_W and Γ_O on which we prescribe the nonhomogeneous Dirichlet, homogeneous Dirichlet and the outflow boundary conditions, respectively:

$$\boldsymbol{v} = \boldsymbol{a} \qquad \text{on } \Gamma_I, \tag{20a}$$

$$\boldsymbol{v} = \mathbf{0} \qquad \text{on } \Gamma_W, \tag{20b}$$

$$\boldsymbol{v}_\tau = \mathbf{0} \qquad \text{on } \Gamma_O, \tag{20c}$$

$$m + \mathbf{S}\boldsymbol{n} \cdot \boldsymbol{n} = 0 \qquad \text{on } \Gamma_O. \tag{20d}$$

We present several weak formulations of problem (19)–(20) in which the following function spaces are used:

$$\begin{aligned}
L^q &:= L^q(\Omega),\\
\boldsymbol{W}^{1,q} &:= W^{1,q}(\Omega; \mathbb{R}^d),\\
\boldsymbol{W}^{1,q}_{\text{bc}} &:= \left\{\boldsymbol{\varphi} \in \boldsymbol{W}^{1,q};\ \boldsymbol{\varphi}|_{\Gamma_I} = \mathbf{0},\ \boldsymbol{\varphi}|_{\Gamma_W} = \mathbf{0},\ \boldsymbol{\varphi}_\tau|_{\Gamma_O} = \mathbf{0}\right\},\\
\mathbf{L}^q &:= L^q(\Omega; \mathbb{R}^{d\times d}_{\text{sym}}),\\
\mathbf{L}^q_0 &:= \{\boldsymbol{\xi} \in \mathbf{L}^q;\ \operatorname{tr}\boldsymbol{\xi} = 0 \text{ a.e. in } \Omega\}.
\end{aligned}$$

For scalar functions f, g such that $fg \in L^1(D)$ we define $(f, g)_D := \int_D fg$, similarly for vector and tensor-valued functions. In case $D = \Omega$ we omit the index and write only (f, g).

We primarily consider the following three weak formulations of the problem (19)–(20). We will often work with $\mathbf{D}$ as the symmetric part of the velocity gradient, requiring that $\mathbf{D} := \frac{1}{2}\left(\nabla\boldsymbol{v} + (\nabla\boldsymbol{v})^{\mathrm{T}}\right)$. In order to distinguish clearly what we mean, we use $\mathbf{D}(\boldsymbol{z}) := \frac{1}{2}\left(\nabla\boldsymbol{z} + (\nabla\boldsymbol{z})^{\mathrm{T}}\right)$ where apropriate. First we work with the $(\mathbf{S}, \boldsymbol{v}, m)$ setting and include the divergence-less condition into the formulation, hence we have:

Problem (A) Find $(\mathbf{S}, \boldsymbol{v}, m) \in \mathbf{L}_0^{r'} \times \boldsymbol{W}^{1,r} \times L^{r'}$ such that $\boldsymbol{v} - \boldsymbol{a} \in \boldsymbol{W}^{1,r}_{\mathrm{bc}}$, $\boldsymbol{a} \in \boldsymbol{W}^{1,r}$ and

$$\begin{aligned} (\mathbf{G}(\mathbf{S}, \mathbf{D}(\boldsymbol{v})), \boldsymbol{\xi}) &= 0 && \forall \boldsymbol{\xi} \in \mathbf{L}_0^{r'}, \\ (\mathbf{S}, \mathbf{D}(\boldsymbol{\varphi})) + (m, \operatorname{div} \boldsymbol{\varphi}) &= (\boldsymbol{f}, \boldsymbol{\varphi}) && \forall \boldsymbol{\varphi} \in \boldsymbol{W}^{1,r}_{\mathrm{bc}}, \\ (\operatorname{div} \boldsymbol{v}, \psi) &= 0 && \forall \psi \in L^{r'}. \end{aligned}$$

If one considers the full Cauchy stress $\mathbf{T}$ as an unknown and defines the deviatoric part as $\mathbf{T}^{\delta} := \mathbf{T} - \frac{\operatorname{tr}\mathbf{T}}{\operatorname{tr}\mathbf{I}}\mathbf{I}$, then it is possible to eliminate the pressure (and also the divergence-less condition) as it can be seen from the next formulation:

Problem (B) Find $(\mathbf{T}, \boldsymbol{v}) \in \mathbf{L}^{r'} \times \boldsymbol{W}^{1,r}$ such that $\boldsymbol{v} - \boldsymbol{a} \in \boldsymbol{W}^{1,r}_{\mathrm{bc}}$, $\boldsymbol{a} \in \boldsymbol{W}^{1,r}$ and

$$\begin{aligned} \left(\mathbf{G}(\mathbf{T}^{\delta}, \mathbf{D}(\boldsymbol{v})), \boldsymbol{\xi}\right) &= 0 && \forall \boldsymbol{\xi} \in \mathbf{L}^{r'}, \\ (\mathbf{T}, \mathbf{D}(\boldsymbol{\varphi})) &= (\boldsymbol{f}, \boldsymbol{\varphi}) && \forall \boldsymbol{\varphi} \in \boldsymbol{W}^{1,r}_{\mathrm{bc}}. \end{aligned}$$

We can also relax the constitutive relation by adding the tensor $\mathbf{D}$ to the list of unknowns, which is then required to be equal to the symmetric part of the velocity gradient. This leads to:

Problem (C) Find $(\mathbf{D}, \mathbf{S}, \boldsymbol{v}, m) \in \mathbf{L}_0^{r} \times \mathbf{L}_0^{r'} \times \boldsymbol{W}^{1,r} \times L^{r'}$ such that $\boldsymbol{v} - \boldsymbol{a} \in \boldsymbol{W}^{1,r}_{\mathrm{bc}}$, $\boldsymbol{a} \in \boldsymbol{W}^{1,r}$ and

$$\begin{aligned} (\mathbf{G}(\mathbf{S}, \mathbf{D}), \boldsymbol{\xi}) &= 0 && \forall \boldsymbol{\xi} \in \mathbf{L}_0^{r'}, \\ (\mathbf{D} - \mathbf{D}(\boldsymbol{v}), \boldsymbol{\zeta}) &= 0 && \forall \boldsymbol{\zeta} \in \mathbf{L}_0^{r'}, \\ (\mathbf{S}, \mathbf{D}(\boldsymbol{\varphi})) + (m, \operatorname{div} \boldsymbol{\varphi}) &= (\boldsymbol{f}, \boldsymbol{\varphi}) && \forall \boldsymbol{\varphi} \in \boldsymbol{W}^{1,r}_{\mathrm{bc}}, \\ (\operatorname{div} \boldsymbol{v}, \psi) &= 0 && \forall \psi \in L^{r'}. \end{aligned}$$

Note that in case of internal flow ($\Gamma_O = \emptyset$ and $\boldsymbol{v} \cdot \boldsymbol{n} = 0$ on $\partial\Omega$), an extra condition fixing the value of m has to be incorporated. Typicaly we demand zero integral mean of $m = \frac{\operatorname{tr}\mathbf{T}}{\operatorname{tr}\mathbf{I}}$.

One can of course derive other weak formulations. For example, it is possible to use the space

$$\mathbf{H}^{q}_{\operatorname{div},n} := \{\boldsymbol{\xi} \in \mathbf{L}^{q};\ \operatorname{div} \boldsymbol{\xi} \in L^{q}(\Omega; \mathbb{R}^d),\ \boldsymbol{\xi}\boldsymbol{n} \cdot \boldsymbol{n}|_{\Gamma_O} = 0\}$$

for $\mathbf{T}$ and formulate the following problem:

Problem (D) Find $(\mathbf{T}, \boldsymbol{v}, \mathbf{D}) \in \mathbf{H}^{r'}_{\operatorname{div},n} \times L^{r}(\Omega; \mathbb{R}^d) \times \mathbf{L}_0^{r}$ such that

$$\begin{aligned} \left(\mathbf{G}(\mathbf{T}^{\delta}, \mathbf{D}), \boldsymbol{\eta}\right) &= 0 && \forall \boldsymbol{\eta} \in \mathbf{L}_0^{r}, \\ (-\operatorname{div} \mathbf{T}, \boldsymbol{\varphi}) &= (\boldsymbol{f}, \boldsymbol{\varphi}) && \forall \boldsymbol{\varphi} \in L^{r}(\Omega; \mathbb{R}^d), \\ (\mathbf{D}, \boldsymbol{\xi}) + (\boldsymbol{v}, \operatorname{div} \boldsymbol{\xi}) &= (\boldsymbol{a}, \boldsymbol{\xi}\boldsymbol{n})_{\Gamma_I} && \forall \boldsymbol{\xi} \in \mathbf{H}^{r'}_{\operatorname{div},n}. \end{aligned}$$

While the weak formulations **(A)**–**(D)** are equivalent to each other, a discrepancy can arise between their discretizations. In particular, the approximation of Problem **(D)** requires a completely different choice of finite element spaces (see [34]), which is one reason why we do not examine it in this work. Another reason for avoiding Problem **(D)** in this study is due to our intention to develop tools for evolutionary problems where the additional information concerning the integrability of $\operatorname{div} \mathbf{T}$ is not available.

3. NUMERICAL ANALYSIS OF LINEARIZED SCHEMES

In this section we shall study the finite element approximations of Problems **(A)**–**(C)**.

Let $\mathcal{T}_h$ be a partition of Ω into simplices with the norm $h := \max_{K \in \mathcal{T}_h} \operatorname{diam} K$. Further let $\mathbf{L}_h \subset \mathbf{L}^{r'}$, $\mathbf{L}_{0h} \subset \mathbf{L}_0^{r'} \cap \mathbf{L}_0^{r}$, $\boldsymbol{W}_h \subset \boldsymbol{W}^{1,r}_{\mathrm{bc}}$ and $L_h \subset L^{r'}$ be finite-dimensional spaces built upon $\mathcal{T}_h$, and $\{\boldsymbol{\eta}_i\}_{i=1}^{N_{\mathbf{T}}}$, $\{\boldsymbol{\xi}_i\}_{i=1}^{N_{\mathbf{S}}}$, $\{\boldsymbol{\varphi}_i\}_{i=1}^{N_{\boldsymbol{v}}}$, $\{\psi_i\}_{i=1}^{N_m}$ be their respective bases. At the moment, we do not require any relationships between the discrete spaces—some restrictions will arise later.

The discrete counterparts of Problems **(A)**–**(C)**, obtained by the Galerkin method, i.e. by replacing the function spaces by their finite-dimensional counterparts, will be denoted by Problems **(A**$_h$**)**–**(C**$_h$**)**.

The resulting discrete systems need to be linearized in a suitable way. For example, the linearization via the Newton-Raphson method produces the sequence $\{\mathbf{S}^k, \mathbf{D}^k\}$ of approximate deviatoric stresses and symmetric velocity gradients obtained by solving the linearized constitutive equation:

$$\begin{aligned}\left[\frac{\partial \mathbf{G}}{\partial \mathbf{S}}(\mathbf{S}^k, \mathbf{D}^k)\right] \mathbf{S}^k_\delta + \left[\frac{\partial \mathbf{G}}{\partial \mathbf{D}}(\mathbf{S}^k, \mathbf{D}^k)\right] \mathbf{D}^k_\delta &= -\mathbf{G}(\mathbf{S}^k, \mathbf{D}^k), \\ \left(\mathbf{S}^{k+1}, \mathbf{D}^{k+1}\right) &= \left(\mathbf{S}^k + \mathbf{S}^k_\delta, \mathbf{D}^k + \mathbf{D}^k_\delta\right),\end{aligned} \tag{25}$$

which arises upon using property **(A1)** above. One can use also other linearization techniques, especially in the case when $\mathbf{G}$ is non-differentiable. However in practical computation most often such iteration is applied to a suitable regularized $\mathbf{G}_\varepsilon$. Since we will follow this path, we next focus on the analysis of the linearized schemes having the following form:

Problem (A$_{h,\mathbf{lin}}$**)** Given $\mathbf{G_S}, \mathbf{G_D} \in L^\infty(\Omega; \mathbb{R}^{d^4})$ and $\mathbf{F} \in L^\infty(\Omega; \mathbb{R}^{d\times d})$, find $(\mathbf{S}_h, \boldsymbol{v}_h, m_h) \in \mathbf{L}_{0h} \times \boldsymbol{W}_h \times L_h$ such that

$$\begin{aligned}(\mathbf{G_S}\mathbf{S}_h, \boldsymbol{\xi}_i) + (\mathbf{G_D}\mathbf{D}(\boldsymbol{v}_h), \boldsymbol{\xi}_i) &= (\mathbf{F}, \boldsymbol{\xi}_i) & i &= 1, \dots, N_{\mathbf{S}}, \\ (\mathbf{S}_h, \mathbf{D}(\boldsymbol{\varphi}_i)) + (m_h, \operatorname{div} \boldsymbol{\varphi}_i) &= (\boldsymbol{f}, \boldsymbol{\varphi}_i) & i &= 1, \dots, N_{\boldsymbol{v}}, \\ (\psi_i, \operatorname{div} \boldsymbol{v}_h) &= 0 & i &= 1, \dots, N_m.\end{aligned}$$

Problem (B$_{h,\mathbf{lin}}$**)** Given $\mathbf{G_S}, \mathbf{G_D} \in L^\infty(\Omega; \mathbb{R}^{d^4})$ and $\mathbf{F} \in L^\infty(\Omega; \mathbb{R}^{d\times d})$, find $(\mathbf{T}_h, \boldsymbol{v}_h) \in \mathbf{L}_h \times \boldsymbol{W}_h$ such that

$$\begin{aligned}\left(\mathbf{G_S}\mathbf{T}^\delta_h, \boldsymbol{\eta}_i\right) + (\mathbf{G_D}\mathbf{D}(\boldsymbol{v}_h), \boldsymbol{\eta}_i) &= (\mathbf{F}, \boldsymbol{\eta}_i) & i &= 1, \dots, N_{\mathbf{T}}, \\ (\mathbf{T}_h, \mathbf{D}(\boldsymbol{\varphi}_i)) &= (\boldsymbol{f}, \boldsymbol{\varphi}_i) & i &= 1, \dots, N_{\boldsymbol{v}}.\end{aligned}$$

Problem (C$_{h,\mathbf{lin}}$**)** Given $\mathbf{G_S}, \mathbf{G_D} \in L^\infty(\Omega; \mathbb{R}^{d^4})$ and $\mathbf{F} \in L^\infty(\Omega; \mathbb{R}^{d\times d})$, find $(\mathbf{D}_h, \mathbf{S}_h, \boldsymbol{v}_h, m_h) \in \mathbf{L}_{0h} \times \mathbf{L}_{0h} \times \boldsymbol{W}_h \times L_h$ such that

$$\begin{aligned}(\mathbf{G_S}\mathbf{S}_h, \boldsymbol{\xi}_i) + (\mathbf{G_D}\mathbf{D}_h, \boldsymbol{\xi}_i) &= (\mathbf{F}, \boldsymbol{\xi}_i) & i &= 1, \dots, N_{\mathbf{S}}, \\ (\mathbf{D}_h - \mathbf{D}(\boldsymbol{v}_h), \boldsymbol{\zeta}_i) &= 0 & i &= 1, \dots, N_{\mathbf{S}}, \\ (\mathbf{S}_h, \mathbf{D}(\boldsymbol{\varphi}_i)) + (m_h, \operatorname{div} \boldsymbol{\varphi}_i) &= (\boldsymbol{f}, \boldsymbol{\varphi}_i) & i &= 1, \dots, N_{\boldsymbol{v}}, \\ (\psi_i, \operatorname{div} \boldsymbol{v}_h) &= 0 & i &= 1, \dots, N_m.\end{aligned}$$

In the above identities, the product of fourth and second order tensor is defined by

$$(\mathbf{G_S}\mathbf{S}_h)_{ij} = \sum_{k,l=1}^{d} (\mathbf{G_S})_{ijkl} (\mathbf{S}_h)_{kl}.$$

The elements $\mathbf{G_S}$, $\mathbf{G_D}$, $\mathbf{F}$ represent linearization of the constitutive equation; e.g. in the case of the Newton-Raphson method (25) one has

$$\mathbf{G_S} = \frac{\partial \mathbf{G}}{\partial \mathbf{S}}(\mathbf{S}^k, \mathbf{D}^k), \quad \mathbf{G_D} = \frac{\partial \mathbf{G}}{\partial \mathbf{D}}(\mathbf{S}^k, \mathbf{D}^k), \quad \mathbf{F} = -\mathbf{G}(\mathbf{S}^k, \mathbf{D}^k),$$

where $\mathbf{S}^k$, $\mathbf{D}^k$ is the approximation of the deviatoric part of the stress and symmetric part of the velocity gradient, taken from the previous iteration.

Recall that $\mathbf{A} \in \mathbb{R}^{d^4}$ is positive definite if there is a constant $c > 0$ such that

$$\text{for all } \mathbf{B} \in \mathbb{R}^{d\times d}: \qquad (\mathbf{AB}) \cdot \mathbf{B} \geq c|\mathbf{B}|^2.$$

Our next goal is to find conditions for solvability of the linearized Problems **($\mathbf{A}_{h,\mathbf{lin}}$)-($\mathbf{C}_{h,\mathbf{lin}}$)**. Due to their saddle-point structure, the choice of the finite element spaces is restricted by the requirement that appropriate inf-sup conditions, which are in finite dimensions equivalent to the full rank property of certain matrices, are satisfied.

The main results of this section follow.

Theorem 1 (On well-posedness of Problems **($\mathbf{A}_{h,\mathbf{lin}}$)** and **($\mathbf{C}_{h,\mathbf{lin}}$)**.)
Let $(-\mathbf{G}_\mathbf{S})$ and $\mathbf{G}_\mathbf{D}$ be uniformly positive definite a.e. in Ω and the spaces $\mathbf{L}_{0h}$, L_h, $\boldsymbol{W}_h$ satisfy the following conditions:

(i) There exists $c > 0$ such that

$$\sup_{\boldsymbol{\varphi} \in \boldsymbol{W}_h} \frac{(p, \operatorname{div} \boldsymbol{\varphi})}{\|\boldsymbol{\varphi}\|_{1,2}} \geq c\,\|p\|_2 \qquad \text{for all } p \in L_h;$$

(ii) $\{0\} \neq \{(\mathbf{D}(\boldsymbol{\varphi}))^\delta;\ \boldsymbol{\varphi} \in \boldsymbol{W}_h\} \subset \mathbf{L}_{0h}$.

Then for every $\mathbf{F} \in L^\infty(\Omega; \mathbb{R}^{d\times d})$, Problem **($\mathbf{A}_{h,\mathbf{lin}}$)** has a unique solution $(\mathbf{S}_h, \boldsymbol{v}_h, m_h) \in \mathbf{L}_{0h} \times \boldsymbol{W}_h \times L_h$ and Problem **($\mathbf{C}_{h,\mathbf{lin}}$)** has a unique solution $(\mathbf{S}_h, \boldsymbol{v}_h, m_h, \mathbf{D}_h) \in \mathbf{L}_{0h} \times \boldsymbol{W}_h \times L_h \times \mathbf{L}_{0h}$.

We will prove Theorem 1 in Sections 3.2 and 3.4. The assumptions of Theorem 1 concerning $\mathbf{G}_\mathbf{S}$ and $\mathbf{G}_\mathbf{D}$ can be verified for particular constitutive relations: Consider the class of models of the form

$$\mathbf{G}(\mathbf{S}, \mathbf{D}) = \alpha(|\mathbf{D}|^2)\mathbf{D} - \beta(|\mathbf{S}|^2)\mathbf{S}, \quad \alpha, \beta \geq 0. \tag{26}$$

This class includes for example the standard and generalized power-law fluid as well as the stress–power-law fluids characterized by (12) or (13). A direct computation implies that

$$\mathbf{G}_\mathbf{S}\mathbf{B} \cdot \mathbf{B} = \sum_{i,j,k,l=1} d\frac{\partial \mathbf{G}_{ij}}{\partial \mathbf{S}_{kl}}(\mathbf{S}, \mathbf{D})\mathbf{B}_{ij}\mathbf{B}_{kl} = -2\beta'|\mathbf{S} \cdot \mathbf{B}|^2 - \beta|\mathbf{B}|^2,$$
$$\mathbf{G}_\mathbf{D}\mathbf{B} \cdot \mathbf{B} = \sum_{i,j,k,l=1} d\frac{\partial \mathbf{G}_{ij}}{\partial \mathbf{D}_{kl}}(\mathbf{S}, \mathbf{D})\mathbf{B}_{ij}\mathbf{B}_{kl} = 2\alpha'|\mathbf{D} \cdot \mathbf{B}|^2 + \alpha|\mathbf{B}|^2,$$

where α', β' denote derivative of α and β. Then it follows that $(-\mathbf{G}_\mathbf{S})$ and $\mathbf{G}_\mathbf{D}$ are positive definite for the stress–power-law models (12) or (13) with $r > 1$ and $r' > 1$. This also applies to the variant (Bi-3) of Bingham model. The other formulations of Bingham fluid (Bi-1), (Bi-2), and (Bi-4) that are of the more general form (15)

$$\mathbf{G}(\mathbf{S}, \mathbf{D}) = \alpha(|\mathbf{D}|^2, |\mathbf{S}|^2)\mathbf{D} - \beta(|\mathbf{D}|^2, |\mathbf{S}|^2)\mathbf{S}, \quad \alpha, \beta \geq 0 \tag{27}$$

lead to

$$\mathbf{G}_\mathbf{S}\mathbf{B} \cdot \mathbf{B} = \sum_{i,j,k,l=1} d\frac{\partial \mathbf{G}_{ij}}{\partial \mathbf{S}_{kl}}\mathbf{B}_{ij}\mathbf{B}_{kl} = -2\beta_s|\mathbf{S} \cdot \mathbf{B}|^2 - \beta|\mathbf{B}|^2 + 2\alpha_s(\mathbf{S} \cdot \mathbf{B})(\mathbf{D} \cdot \mathbf{B}),$$
$$\mathbf{G}_\mathbf{D}\mathbf{B} \cdot \mathbf{B} = \sum_{i,j,k,l=1} d\frac{\partial \mathbf{G}_{ij}}{\partial \mathbf{D}_{kl}}\mathbf{B}_{ij}\mathbf{B}_{kl} = 2\alpha_d|\mathbf{D} \cdot \mathbf{B}|^2 + \alpha|\mathbf{B}|^2 - 2\beta_d(\mathbf{S} \cdot \mathbf{B})(\mathbf{D} \cdot \mathbf{B}),$$

where $\alpha_d = \frac{\partial\alpha(d,s)}{\partial d}, \alpha_s = \frac{\partial\alpha(d,s)}{\partial s}$ and $\beta_d = \frac{\partial\beta(d,s)}{\partial d}, \beta_s = \frac{\partial\beta(d,s)}{\partial s}$. We then also require that α_d and $\beta_s \geq 0$. Consequently, the positive definiteness of $(-\mathbf{G}_\mathbf{S})$ and $\mathbf{G}_\mathbf{D}$ is obtained under additional assumption, for example if $\mathbf{D}$ is almost parallel to $\mathbf{S}$ (with appropriate or small coefficients α_s or β_d).

Sufficient conditions for the well-posedness of the Problem **(B$_{h,\text{lin}}$)** are given in the following theorem. Here we shall use the orthogonal decomposition of the space $\mathbf{L}_h$ into the spaces of deviatoric and spherical stresses:

$$\mathbf{L}_h := \mathbf{L}_{0h} \oplus L_h\mathbf{I},$$

where $\mathbf{L}_{0h} := \{\boldsymbol{\eta}^\delta;\ \boldsymbol{\eta} \in \mathbf{L}_h\}$, $L_h := \{\frac{\operatorname{tr}\boldsymbol{\eta}}{\operatorname{tr}\mathbf{I}};\ \boldsymbol{\eta} \in \mathbf{L}_h\}$ and $L_h\mathbf{I} := \{(\psi)\mathbf{I};\ \psi \in L_h\}$.

Theorem 2 (On well-posedness of the Problem **(B$_{h,\text{lin}}$)**.)
Let $\mathbf{G_S}$, $\mathbf{G_D}$ and $\mathbf{L}_h$, $\boldsymbol{W}_h$ satisfy the following conditions:

(i) $(-\mathbf{G_S})$ and $\mathbf{G_D}$ are uniformly positive definite a.e. in Ω;
(ii) $\{0\} \neq \{(\mathbf{D}(\boldsymbol{\varphi}))^\delta;\ \boldsymbol{\varphi} \in \boldsymbol{W}_h\} \subset \mathbf{L}_{0h}$;
(iii) There exists $c_1 > 0$ such that

$$\sup_{\boldsymbol{\varphi}\in\boldsymbol{W}_h} \frac{(\operatorname{tr}\mathbf{T}, \operatorname{div}\boldsymbol{\varphi})}{\|\boldsymbol{\varphi}\|_{1,2}} \geq c_1 \|\operatorname{tr}\mathbf{T}\|_2 \qquad \text{for all } \mathbf{T} \in \mathbf{L}_h;$$

(iv) For every $\boldsymbol{\varphi} \in \boldsymbol{W}_h$ the following equivalence holds:

$$(\operatorname{div}\varphi, p) = 0 \quad \text{for all } p \in L_h \quad\Leftrightarrow\quad (\operatorname{tr}(\mathbf{G}_D\mathbf{D}(\varphi)), p) = 0 \quad \text{for all } p \in L_h;$$

(v) For all $\mathbf{B} \in \mathbb{R}^{d\times d}_{\text{sym}}$ satisfying $\operatorname{tr}\mathbf{B} = 0$, $\operatorname{tr}(\mathbf{G_S}\mathbf{B}) = 0$;
(vi) $(\mathbf{G_D}\mathbf{I})^\delta = 0$.

Then for every $\mathbf{F} \in L^\infty(\Omega; \mathbb{R}^{d\times d})$ Problem **(B$_{h,\text{lin}}$)** has a unique solution $(\mathbf{T}_h, \boldsymbol{v}_h) \in \mathbf{L}_h \times \boldsymbol{W}_h$.

We will prove Theorem 2 in Section 3.3. In contrast to Problem **(A$_{h,\text{lin}}$)**, here the approximation of the mean normal stress is determined by the finite element space $\mathbf{L}_h$. This introduces additional assumptions $(iv) - (vi)$ in Theorem 2. In order to see these conditions in a specific case, let us again consider the constitutive relation of the form (26). Then

$$\begin{aligned}
\operatorname{tr}(\mathbf{G_D}\mathbf{B}) &= 2(\alpha'\mathbf{D}\cdot\mathbf{B})\operatorname{tr}\mathbf{D} + \alpha\operatorname{tr}\mathbf{B},\\
\operatorname{tr}(\mathbf{G_S}\mathbf{B}) &= -2(\beta'\mathbf{S}\cdot\mathbf{B})\operatorname{tr}\mathbf{S} - \beta\operatorname{tr}\mathbf{B},\\
(\mathbf{G_D}\mathbf{I})^\delta &= 2(\alpha'\operatorname{tr}\mathbf{D})\mathbf{D}^\delta.
\end{aligned}$$

Consequently, the additional assumptions $(iv) - (vi)$ of Theorem 2 are satisfied if $\alpha' = 0$, i.e.,

$$\mathbf{G}(\mathbf{S},\mathbf{D}) = \mathbf{D} - \beta(\operatorname{tr}\mathbf{S}^2)\mathbf{S}, \quad \beta \text{ positive}, \tag{28}$$

in particular this holds for the (generalized) stress–power-law (13).

For the more general form (27) we have

$$\begin{aligned}
\operatorname{tr}(\mathbf{G_D}\mathbf{B}) &= 2(\alpha_d\mathbf{D}\cdot\mathbf{B})\operatorname{tr}\mathbf{D} + \alpha\operatorname{tr}\mathbf{B} - 2(\beta_d\mathbf{D}\cdot\mathbf{B})\operatorname{tr}\mathbf{S},\\
\operatorname{tr}(\mathbf{G_S}\mathbf{B}) &= -2(\beta_s\mathbf{S}\cdot\mathbf{B})\operatorname{tr}\mathbf{S} - \beta\operatorname{tr}\mathbf{B} + 2(\alpha_s\mathbf{S}\cdot\mathbf{B})\operatorname{tr}\mathbf{D},\\
(\mathbf{G_D}\mathbf{I})^\delta &= 2(\alpha_d\operatorname{tr}\mathbf{D})\mathbf{D}^\delta - 2(\beta_d\operatorname{tr}\mathbf{D})\mathbf{S}^\delta,
\end{aligned}$$

and the assumptions $(iv) - (vi)$ of Theorem 2 are at least satisfied if we are at the solution, i.e. $\operatorname{tr}\mathbf{D} = 0$. In the following subsection we prove an abstract result which will be then applied to the proof of Theorems 1 and 2.

3.1. Generalized multi-level saddle point problems in finite dimension

Let $\mathbf{A} \in \mathbb{R}^{m_0\times m_0}$, $\mathbf{B}_i \in \mathbb{R}^{m_i\times m_{i-1}}$, $\mathbf{C}_i \in \mathbb{R}^{m_i\times m_{i-1}}$, $i = 1,\dots,n$, be given matrices for some positive integers $n, m_0, \dots, m_n$. We shall study the properties of the following block-tridiagonal matrix:

$$\mathbf{M}_n := \begin{bmatrix} \mathbf{A} & \mathbf{B}_1^\top & & \mathbf{0} \\ \mathbf{C}_1 & \mathbf{0} & \ddots & \\ & \ddots & \ddots & \mathbf{B}_n^\top \\ \mathbf{0} & & \mathbf{C}_n & \mathbf{0} \end{bmatrix}; \tag{29}$$

in particular, our interest is to analyze its invertibility. We observe that if $\mathbf{M}_0 := A$ then for any $n \geq 1$,

$$\mathbf{M}_n = \begin{bmatrix} \mathbf{M}_{n-1} & \widehat{\mathbf{B}}_n^\top \\ \widehat{\mathbf{C}}_n & \mathbf{0} \end{bmatrix}, \text{ where } \widehat{\mathbf{B}}_n := \begin{bmatrix} \mathbf{0} & \dots & \mathbf{0} & \mathbf{B}_n \end{bmatrix} \text{ and } \widehat{\mathbf{C}}_n := \begin{bmatrix} \mathbf{0} & \dots & \mathbf{0} & \mathbf{C}_n \end{bmatrix}.$$

The existence of the inverse matrix $\mathbf{M}_n^{-1}$ is equivalent to the invertibility of the Schur complement of $\mathbf{M}_{n-1}$ in $\mathbf{S}_n$, which is, due to the structure of $\mathbf{M}_n$, defined by:

$$\mathbf{S}_n := -\widehat{\mathbf{C}}_n \mathbf{M}_{n-1}^{-1} \widehat{\mathbf{B}}_n^\top. \tag{30}$$

In the next lemma we show a relation for the Schur complement $\mathbf{S}_n$ which will be used to prove that $\mathbf{M}_n$ is nonsingular.

Lemma 1
Suppose that the matrix $\mathbf{M}_n$ defined in (29) is nonsingular. Then

$$\begin{aligned} \mathbf{S}_n &= \begin{cases} -\mathbf{C}_1 \mathbf{A}^{-1} \mathbf{B}_1^\top & \text{for } n = 1, \\ -\mathbf{C}_n \mathbf{S}_{n-1}^{-1} \mathbf{B}_n^\top & \text{for } n > 1 \end{cases} \\ &= (-1)^n \mathbf{C}_n \left(\mathbf{C}_{n-1} \left(\dots \left(\mathbf{C}_1 \mathbf{A}^{-1} \mathbf{B}_1^\top \right)^{-1} \dots \right)^{-1} \mathbf{B}_{n-1}^\top \right)^{-1} \mathbf{B}_n^\top. \end{aligned} \tag{31}$$

Proof
The case $n = 1$ is trivial. For $n > 1$, it is not difficult to verify that

$$\mathbf{M}_n^{-1} = \begin{bmatrix} \mathbf{M}_{n-1}^{-1} + \mathbf{M}_{n-1}^{-1} \widehat{\mathbf{B}}_n^\top \mathbf{S}_{n-1}^{-1} \widehat{\mathbf{C}}_n \mathbf{M}_{n-1}^{-1} & -\mathbf{M}_{n-1}^{-1} \widehat{\mathbf{B}}_n^\top \mathbf{S}_{n-1}^{-1} \\ -\mathbf{S}_{n-1}^{-1} \widehat{\mathbf{C}}_n \mathbf{M}_{n-1}^{-1} & \mathbf{S}_{n-1}^{-1} \end{bmatrix}.$$

Then from (30) we obtain:

$$\mathbf{S}_n = -\mathbf{C}_n \mathbf{S}_{n-1}^{-1} \mathbf{B}_n^\top.$$

Applying this formula recursively yields the last identity in (31). □

Corollary 1
Let $\mathbf{A}$, $\mathbf{B}_i$, $\mathbf{C}_i$, $i = 1, \dots, n$, satisfy the following conditions:

(C1) C1 $\mathbf{A}$ is positive or negative definite;
(C2) C2 For all $i = 1, \dots, n$: $\mathbf{B}_i$ has full row rank;
(C3) C3 For all $i = 1, \dots, n-1$ there exists $\mathbf{H}_i \in \mathbb{R}^{m_{i-1} \times m_{i-1}}$ positive or negative definite such that $\mathbf{C}_i = \mathbf{B}_i \mathbf{H}_i$;
(C4) C4 There exist $\mathbf{H}_n \in \mathbb{R}^{m_{n-1} \times m_{n-1}}$ positive or negative definite and $\mathbf{K}_n \in \mathbb{R}^{m_n \times m_n}$ nonsingular such that $\mathbf{C}_n = \mathbf{K}_n \mathbf{B}_n \mathbf{H}_n$.

Then $\mathbf{M}_n$ is nonsingular.

Proof
Without loss of generality we can assume that $\mathbf{A}$, $\mathbf{H}_1, \dots, \mathbf{H}_n$ are positive definite. Consequently, $\mathbf{H}_1 \mathbf{A}^{-1}$ is also positive definite. Using (31) and (C2), for any nonzero vector $\boldsymbol{x} \in \mathbb{R}^{m_1}$ we have:

$$\boldsymbol{x} \cdot \mathbf{S}_1 \boldsymbol{x} = -\boldsymbol{x} \cdot \mathbf{B}_1 \mathbf{H}_1 \mathbf{A}^{-1} \mathbf{B}_1^\top \boldsymbol{x} = -\mathbf{B}_1^\top \boldsymbol{x} \cdot \mathbf{H}_1 \mathbf{A}^{-1} \mathbf{B}_1^\top \boldsymbol{x} < 0,$$

i.e. $\mathbf{S}_1$ is negative definite. By the same argument it is possible to successively show that $\mathbf{S}_2, \dots, \mathbf{S}_{n-1}$ is either positive or negative definite. Finally, since

$$\mathbf{S}_n = (-\mathbf{K}_n)(\mathbf{B}_n \mathbf{H}_n \mathbf{S}_{n-1}^{-1} \mathbf{B}_n^\top)$$

is a product of two nonsingular matrices, we have that $\mathbf{S}_n$ is nonsingular. □

Remark 1
The definiteness of $\mathbf{A}$ in the previous corollary is essential. In particular, to assume that $\mathbf{A}$ nonsingular is not sufficient for nonsingularity of $\mathbf{M}_n$, as can be seen from the following example:

$$\mathbf{A} := \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} (= \mathbf{A}^{-1}), \quad \mathbf{B}_1 = \mathbf{C}_1 = \begin{bmatrix} 1 & 0 \end{bmatrix} \quad \Rightarrow \quad \mathbf{S}_1 = \mathbf{C}_1 \mathbf{A}^{-1} \mathbf{B}_1^\top = \mathbf{0}.$$

3.2. Well-posedness of Problem $(\boldsymbol{A}_{h,\mathbf{lin}})$

Proof of the first part of Theorem 1
We define the matrices $\mathbf{A}, \mathbf{B}_1, \mathbf{B}_2 \in \mathbb{R}^{N_{\mathsf{S}} \times N_{\mathsf{S}}}$, $\mathbf{C} \in \mathbb{R}^{N_m \times N_{\boldsymbol{v}}}$ and the vectors of the right hand sides $\boldsymbol{F} \in \mathbb{R}^{N_{\mathsf{S}}}$, $\boldsymbol{R} \in \mathbb{R}^{N_{\boldsymbol{v}}}$ as follows:

$$\mathbf{A}_{ij} := \left(\mathbf{G}_{\mathsf{S}} \boldsymbol{\xi}_j, \boldsymbol{\xi}_i\right), \quad (\mathbf{B}_1)_{ij} := \left(\boldsymbol{\xi}_j, \mathbf{G}_{\mathsf{D}} \boldsymbol{\xi}_i\right), \quad (\mathbf{B}_2)_{ij} := \left(\boldsymbol{\xi}_j, \boldsymbol{\xi}_i\right),$$

$$\mathbf{C}_{ij} := \left(\operatorname{div} \boldsymbol{\varphi}_j, \psi_i\right), \quad \boldsymbol{F}_i := (\mathbf{F}, \boldsymbol{\xi}_i), \quad \boldsymbol{R}_i := (\boldsymbol{f}, \boldsymbol{\varphi}_i).$$

Assumption (ii) implies that there exists a matrix $\mathbf{E} \in \mathbb{R}^{N_{\boldsymbol{v}} \times N_{\mathsf{S}}}$, defined by the following equations:

$$\mathbf{D}(\boldsymbol{\varphi}_i)^\delta = \sum_{j=1}^{N_{\mathsf{S}}} \mathbf{E}_{ij} \boldsymbol{\xi}_j, \; i = 1, \dots, N_{\boldsymbol{v}}.$$

With the help of the above notation, the algebraic representation of $(\mathbf{A}_{h,\mathbf{lin}})$ reads:

$$\begin{bmatrix} \mathbf{A} & \mathbf{B}_1^\top \mathbf{E}^\top & \mathbf{0} \\ \mathbf{E}\mathbf{B}_2 & \mathbf{0} & \mathbf{C}^\top \\ \mathbf{0} & \mathbf{C} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{S} \\ \boldsymbol{V} \\ \boldsymbol{M} \end{bmatrix} = \begin{bmatrix} \boldsymbol{F} \\ \boldsymbol{R} \\ \mathbf{0} \end{bmatrix}, \tag{32}$$

where the vectors $\boldsymbol{S}$, $\boldsymbol{V}$, $\boldsymbol{M}$ hold the degrees of freedom of the solution, i.e.

$$\mathbf{S}_h = \sum_{i=1}^{N_{\mathsf{S}}} S_i \boldsymbol{\xi}_i, \quad \boldsymbol{v}_h = \sum_{i=1}^{N_{\boldsymbol{v}}} V_i \boldsymbol{\varphi}_i, \quad m_h = \sum_{i=1}^{N_m} M_i \psi_i.$$

Since the matrix from (32) is block tridiagonal, we shall use Corollary 1 to show that it is nonsingular. Indeed, from the assumption on $\mathbf{G}_{\mathsf{S}}$ it follows that $\mathbf{A}$ is negative definite, hence (C1) is satisfied. Similarly, $\mathbf{B}_1$ and $\mathbf{B}_2$ are positive definite. Further,

$$\mathbf{E}\mathbf{B}_2 = \mathbf{E}\mathbf{B}_1(\mathbf{B}_1^{-1}\mathbf{B}_2),$$

where $\mathbf{B}_1^{-1}\mathbf{B}_2$ is positive definite, from which (C3) follows. Korn's inequality implies that the rows of $\mathbf{E}$ are linearly independent, i.e. $\mathbf{E}$ has full row rank, and so do $\mathbf{E}\mathbf{B}_1$ and $\mathbf{E}\mathbf{B}_2$. From (i) it follows that

$$\left\|\mathbf{C}^\top \boldsymbol{q}_h\right\| = \sup_{\boldsymbol{V} \in \mathbb{R}^{N_{\boldsymbol{v}}} \setminus \{0\}} \frac{\mathbf{C}^\top \boldsymbol{q}_h \cdot \boldsymbol{V}}{\|\boldsymbol{V}\|} = \sup_{\boldsymbol{v} \in \boldsymbol{W}_h} \frac{\sum_{i=1}^{N_m} q_i \left(\psi_i, \operatorname{div} \boldsymbol{v}\right)}{\|\boldsymbol{v}\|_{1,2}} \geq c \left\|\boldsymbol{q}_h\right\|,$$

i.e. $\mathbf{C}$ has full row rank. Hence, (C2) and (C4) hold true. Since all assumptions of Corollary 1 are satisfied, the system (32) is nonsingular. □

3.3. Well-posedness of Problem ($\boldsymbol{B}_{h,\mathbf{lin}}$)

Proof of Theorem 2
We split the stress tensor $\mathbf{T}_h$ into the spherical and deviatoric parts:

$$\mathbf{T}_h = \mathbf{T}_h^\delta + \frac{\operatorname{tr}\mathbf{T}_h}{\operatorname{tr}\mathbf{I}}\mathbf{I} =: \mathbf{S}_h + m_h\mathbf{I},\ \mathbf{S}_h \in \mathbf{L}_{0h},\ m_h \in L_h.$$

Due to (v) and (vi), we have that

$$(\operatorname{tr}(\mathbf{G}_\mathbf{S}\mathbf{S}_h), \psi_i) = 0,\ i = 1, \dots, N_m,$$

$$(\mathbf{G}_\mathbf{D}\mathbf{D}\boldsymbol{v}_h, \boldsymbol{\xi}_i) = \left(\mathbf{G}_\mathbf{D}\mathbf{D}^\delta\boldsymbol{v}_h, \boldsymbol{\xi}_i\right),\ i = 1, \dots, N_\mathbf{S}.$$

Consequently, ($\mathbf{B}_{h,\mathbf{lin}}$) can be rewritten in the following way:

$$\begin{aligned}
(\mathbf{G}_\mathbf{S}\mathbf{S}_h, \boldsymbol{\xi}_i) + \left(\mathbf{G}_\mathbf{D}\mathbf{D}^\delta\boldsymbol{v}_h, \boldsymbol{\xi}_i\right) &= (\mathbf{F}, \boldsymbol{\xi}_i)\,, && i = 1, \dots, N_\mathbf{S},\\
\left(\mathbf{S}_h, \mathbf{D}^\delta\boldsymbol{\varphi}_i\right) + (m_h, \operatorname{div}\boldsymbol{\varphi}_i) &= (\boldsymbol{f}, \boldsymbol{\varphi}_i)\,, && i = 1, \dots, N_{\boldsymbol{v}},\\
(\operatorname{tr}(\mathbf{G}_\mathbf{D}\mathbf{D}\boldsymbol{v}_h), \psi_i) &= (\operatorname{tr}\mathbf{F}, \psi_i)\,, && i = 1, \dots, N_m.
\end{aligned}$$

We introduce matrices $\mathbf{A}, \mathbf{B}_1, \mathbf{B}_2 \in \mathbb{R}^{N_\mathbf{S}\times N_\mathbf{S}}$, $\mathbf{C}_1, \mathbf{C}_2 \in \mathbb{R}^{N_m\times N_{\boldsymbol{v}}}$ and the vectors of right-hand sides $\boldsymbol{F} \in \mathbb{R}^{N_\mathbf{S}}$, $\boldsymbol{R} \in \mathbb{R}^{N_{\boldsymbol{v}}}$, $\boldsymbol{H} \in \mathbb{R}^{N_m}$ as follows:

$$\mathbf{A}_{ij} := \left(\mathbf{G}_\mathbf{S}\boldsymbol{\xi}_j, \boldsymbol{\xi}_i\right), \quad (\mathbf{B}_1)_{ij} := \left(\boldsymbol{\xi}_j, \mathbf{G}_\mathbf{D}\boldsymbol{\xi}_i\right), \quad (\mathbf{B}_2)_{ij} := \left(\boldsymbol{\xi}_j, \boldsymbol{\xi}_i\right),$$

$$(\mathbf{C}_1)_{ij} := \left(\psi_i, \operatorname{div}\boldsymbol{\varphi}_j\right), \quad (\mathbf{C}_2)_{ij} := \left(\psi_i, \operatorname{tr}(\mathbf{G}_\mathbf{D}\mathbf{D}\boldsymbol{\varphi}_j)\right),$$

$$\boldsymbol{F}_i := (\mathbf{F}, \boldsymbol{\xi}_i)\,, \quad \boldsymbol{R}_i := (\boldsymbol{f}, \boldsymbol{\varphi}_i)\,, \quad \boldsymbol{H}_i := (\operatorname{tr}\mathbf{F}, \psi_i)\,.$$

It is evident that $(-\mathbf{A}), \mathbf{B}_1, \mathbf{B}_2$ are positive definite. Assumption (ii) yields that there exists a matrix $\mathbf{E} \in \mathbb{R}^{N_{\boldsymbol{v}}\times N_\mathbf{S}}$, defined by the following equations:

$$\mathbf{D}(\boldsymbol{\varphi}_i)^\delta = \sum_{j=1}^{N_\mathbf{S}} \mathbf{E}_{ij}\boldsymbol{\xi}_j,\ i = 1, \dots, N_{\boldsymbol{v}}.$$

Due to Korn's inequality, $\mathbf{E}$ has full row rank. Similarly, (iii) implies that $\mathbf{C}_1$ has full row rank. From (iv) it follows that $\ker\mathbf{C}_1 = \ker\mathbf{C}_2$, which means that $\operatorname{Im}\mathbf{C}_1^\top = \operatorname{Im}\mathbf{C}_2^\top$ and hence there is a nonsingular matrix $\mathbf{H} \in \mathbb{R}^{N_m\times N_m}$ such that

$$\mathbf{C}_2 = \mathbf{H}\mathbf{C}_1.$$

With the help of the above notation, the algebraic representation of ($\mathbf{B}_{h,\mathbf{lin}}$) reads:

$$\begin{bmatrix} \mathbf{A} & \mathbf{B}_1^\top\mathbf{E}^\top & \mathbf{0} \\ \mathbf{E}\mathbf{B}_2 & \mathbf{0} & \mathbf{C}_1^\top \\ \mathbf{0} & \mathbf{H}\mathbf{C}_1 & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{S} \\ \boldsymbol{V} \\ \boldsymbol{M} \end{bmatrix} = \begin{bmatrix} \boldsymbol{F} \\ \boldsymbol{R} \\ \boldsymbol{H} \end{bmatrix}. \tag{33}$$

The properties of the matrices $\mathbf{A}$, $\mathbf{B}_1$, $\mathbf{B}_2$, $\mathbf{C}_1$, $\mathbf{E}$, $\mathbf{H}$ imply the assumptions (C1)–(C4) of Corollary 1, hence (33) is well-posed. □

3.4. Well-posedness of Problem ($\boldsymbol{C}_{h,\mathbf{lin}}$)

Proof of the second part of Theorem 1
We define the matrices $\mathbf{A}, \mathbf{B}_1, \mathbf{B}_2 \in \mathbb{R}^{N_\mathbf{S}\times N_\mathbf{S}}$, $\mathbf{C} \in \mathbb{R}^{N_m\times N_{\boldsymbol{v}}}$ and the vectors of right hand sides $\boldsymbol{F} \in \mathbb{R}^{N_\mathbf{S}}$, $\boldsymbol{R} \in \mathbb{R}^{N_{\boldsymbol{v}}}$ as follows:

$$\mathbf{A}_{ij} := \left(\mathbf{G}_\mathbf{D}\boldsymbol{\xi}_j, \boldsymbol{\xi}_i\right), \quad (\mathbf{B}_1)_{ij} := \left(\boldsymbol{\xi}_j, \mathbf{G}_\mathbf{S}\boldsymbol{\xi}_i\right), \quad (\mathbf{B}_2)_{ij} := \left(\boldsymbol{\xi}_j, \boldsymbol{\xi}_i\right),$$

$$\mathbf{\mathsf{C}}_{ij} := -\left(\operatorname{div}\boldsymbol{\varphi}_j, \psi_i\right), \quad \boldsymbol{F}_i := (\mathbf{\mathsf{F}}, \boldsymbol{\xi}_i), \quad \boldsymbol{R}_i := -(\boldsymbol{f}, \boldsymbol{\varphi}_i).$$

Assumption (ii) implies that there exists a matrix $\mathbf{\mathsf{E}} \in \mathbb{R}^{N_{\boldsymbol{v}} \times N_{\mathsf{S}}}$, defined by the following equations:

$$\mathbf{\mathsf{D}}(\boldsymbol{\varphi}_i)^{\delta} = -\sum_{j=1}^{N_{\mathsf{S}}} \mathbf{\mathsf{E}}_{ij}\boldsymbol{\xi}_j, \; i = 1, \ldots, N_{\boldsymbol{v}}.$$

With the help of the above notation, the algebraic representation of problem $(\mathbf{C}_{h,\mathbf{lin}})$ reads:

$$\begin{bmatrix} \mathbf{\mathsf{A}} & \mathbf{\mathsf{B}}_1^{\top} & \mathbf{0} & \mathbf{0} \\ \mathbf{\mathsf{B}}_2 & \mathbf{0} & \mathbf{\mathsf{B}}_2^{\top}\mathbf{\mathsf{E}}^{\top} & \mathbf{0} \\ \mathbf{0} & \mathbf{\mathsf{E}}\mathbf{\mathsf{B}}_2 & 0 & \mathbf{\mathsf{C}}^{\top} \\ \mathbf{0} & \mathbf{0} & \mathbf{\mathsf{C}} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{D} \\ \boldsymbol{S} \\ \boldsymbol{V} \\ \boldsymbol{M} \end{bmatrix} = \begin{bmatrix} \boldsymbol{F} \\ \mathbf{0} \\ \boldsymbol{R} \\ \mathbf{0} \end{bmatrix}. \tag{34}$$

From the assumptions on $\mathbf{\mathsf{G}}_{\mathsf{S}}$ and $\mathbf{\mathsf{G}}_{\mathsf{D}}$ it follows that $\mathbf{\mathsf{A}}$ is negative definite and $\mathbf{\mathsf{B}}_1$ is positive definite. Obviously, $\mathbf{\mathsf{B}}_2$ is positive definite. Korn's inequality implies that the rows of $\mathbf{\mathsf{E}}$ are linearly independent, i.e. $\mathbf{\mathsf{E}}$ has full row rank. From (i) it follows that $\mathbf{\mathsf{C}}$ has full row rank. Hence, one can verify that (C1)–(C4) are satisfied and thus (34) is well-posed. □

4. NUMERICAL RESULTS

Now we will solve the discrete Problems $(\mathbf{A}_h)$–$(\mathbf{C}_h)$ using the conforming finite element method and using the Newton-Raphson method (25) for linearization. The numerical method has been implemented in the FEniCS library [35, 36] on simplical meshes. The Newton method with linesearch from the PETSc library [37] is used together with automatic differentiation to construct the entries of the linearized matrix. The resulting linear algebraic systems are solved by the sparse direct solver MUMPS [38].

4.1. Choice of finite element spaces

It is clear that the finite element spaces have to be chosen in a specific way. The characteristic structure of the discrete problems can be deduced from the linearized systems stemming from the Newton-Raphson method. The algebraic representation of the linearized systems reads as follows:

Problem $(\mathbf{A}_{h,\mathbf{lin}})$ Given $(\bar{\boldsymbol{S}}, \bar{\boldsymbol{V}}) \in \mathbb{R}^{N_{\mathsf{S}}} \times \mathbb{R}^{N_{\boldsymbol{v}}}$, find $(\boldsymbol{S}, \boldsymbol{V}, \boldsymbol{M}) \in \mathbb{R}^{N_{\mathsf{S}}} \times \mathbb{R}^{N_{\boldsymbol{v}}} \times \mathbb{R}^{N_m}$ such that

$$\begin{bmatrix} \mathbf{\mathsf{A}}(\bar{\boldsymbol{S}}, \bar{\boldsymbol{V}}) & \mathbf{\mathsf{B}}_1^{\top}(\bar{\boldsymbol{S}}, \bar{\boldsymbol{V}}) & 0 \\ \mathbf{\mathsf{B}}_2 & \mathbf{0} & \mathbf{\mathsf{C}}^{\top} \\ \mathbf{0} & \mathbf{\mathsf{C}} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{S} \\ \boldsymbol{V} \\ \boldsymbol{M} \end{bmatrix} = \begin{bmatrix} \boldsymbol{F} \\ \boldsymbol{R} \\ \mathbf{0} \end{bmatrix}. \tag{35}$$

Problem $(\mathbf{B}_{h,\mathbf{lin}})$ Given $(\bar{\boldsymbol{T}}, \bar{\boldsymbol{V}}) \in \mathbb{R}^{N_{\mathsf{T}}} \times \mathbb{R}^{N_{\boldsymbol{v}}}$, find $(\boldsymbol{T}, \boldsymbol{V}) \in \mathbb{R}^{N_{\mathsf{T}}} \times \mathbb{R}^{N_{\boldsymbol{v}}}$ such that

$$\begin{bmatrix} \mathbf{\mathsf{A}}'(\bar{\boldsymbol{T}}, \bar{\boldsymbol{V}}) & \mathbf{\mathsf{B}}_1'^{\top}(\bar{\boldsymbol{T}}, \bar{\boldsymbol{V}}) \\ \mathbf{\mathsf{B}}_2' & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{T} \\ \boldsymbol{V} \end{bmatrix} = \begin{bmatrix} \boldsymbol{F} \\ \boldsymbol{R} \end{bmatrix}. \tag{36}$$

Problem $(\mathbf{C}_{h,\mathbf{lin}})$ Given $(\bar{\boldsymbol{S}}, \bar{\boldsymbol{D}}) \in \mathbb{R}^{N_{\mathsf{S}}} \times \mathbb{R}^{N_{\mathsf{D}}}$, find $(\boldsymbol{D}, \boldsymbol{S}, \boldsymbol{V}, \boldsymbol{M}) \in \mathbb{R}^{N_{\mathsf{D}}} \times \mathbb{R}^{N_{\mathsf{S}}} \times \mathbb{R}^{N_{\boldsymbol{v}}} \times \mathbb{R}^{N_m}$ such that

$$\begin{bmatrix} \mathbf{\mathsf{A}}''(\bar{\boldsymbol{S}}, \bar{\boldsymbol{D}}) & \mathbf{\mathsf{B}}_1''^{\top}(\bar{\boldsymbol{S}}, \bar{\boldsymbol{D}}) & 0 & 0 \\ \mathbf{\mathsf{B}}_2'' & \mathbf{0} & \mathbf{\mathsf{B}}_2'^{\top} & \mathbf{0} \\ \mathbf{0} & \mathbf{\mathsf{B}}_2' & \mathbf{0} & \mathbf{\mathsf{C}}^{\top} \\ \mathbf{0} & \mathbf{0} & \mathbf{\mathsf{C}} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{D} \\ \boldsymbol{S} \\ \boldsymbol{V} \\ \boldsymbol{M} \end{bmatrix} = \begin{bmatrix} \boldsymbol{F} \\ \mathbf{0} \\ \boldsymbol{R} \\ \mathbf{0} \end{bmatrix}. \tag{37}$$

From (35)–(37) it is evident that $(\mathbf{A}_{h,\mathbf{lin}})$–$(\mathbf{C}_{h,\mathbf{lin}})$ are saddle-point problems. In fact, all of them have nested saddle point structure, see Section 3.1, that gives a hint for the choice of the finite element spaces. Following Theorems 1 and 2 the following two conditions are sufficient for a stable discrete problem:

- The space of velocities and the space of pressures (mean normal stresses) satisfy the Babuška–Brezzi condition (see for Theorem 1 (i));
- The space of stresses contains all $\mathbf{D}^{\delta}$.

For simplicial meshes there are several typical choices of the finite elements satisfying the above requirements.

Finite elements for Problem $(\mathbf{A}_h)$. Let $k \geq 1$ be an integer. A particular choice is based on the $\mathcal{P}_{k+1}/\mathcal{P}_k$ Taylor–Hood velocity–pressure elements, where $\mathcal{P}_k(K)$ denotes the set of polynomials on K of order at most k. We set

$$
\begin{aligned}
\mathbf{L}_{0h} &:= \{\mathbf{S} \in \mathbf{L}_0^{r'};\ \mathbf{S}_{|K} \in \mathcal{P}_k(K)^{d\times d}\ \forall K \in \mathcal{T}_h\},\\
L_h &:= \{p \in L^{r'} \cap C(\overline{\Omega});\ p_{|K} \in \mathcal{P}_k(K)\ \forall K \in \mathcal{T}_h\},\\
\boldsymbol{W}_h &:= \{\boldsymbol{v} \in \boldsymbol{W}^{1,r}_{\mathsf{bc}};\ \boldsymbol{v}_{|K} \in \mathcal{P}_{k+1}(K)^d\ \forall K \in \mathcal{T}_h\},
\end{aligned}
\tag{$\mathbf{A}_1$}
$$

Alternatively we can use the stable pair $\mathcal{P}_{k+1}/\mathcal{P}_{k-1}^{\text{discontinous}}$, namely its lowest order variant $\mathcal{P}_2/\mathcal{P}_0$. Then we set

$$
\begin{aligned}
\mathbf{L}_{0h} &:= \{\mathbf{S} \in \mathbf{L}_0^{r'};\ \mathbf{S}_{|K} \in \mathcal{P}_1(K)^{d\times d}\ \forall K \in \mathcal{T}_h\},\\
L_h &:= \{p \in L^{r'};\ p_{|K} \in \mathcal{P}_0(K)\ \forall K \in \mathcal{T}_h\},\\
\boldsymbol{W}_h &:= \{\boldsymbol{v} \in \boldsymbol{W}^{1,r}_{\mathsf{bc}};\ \boldsymbol{v}_{|K} \in \mathcal{P}_2(K)^d\ \forall K \in \mathcal{T}_h\}.
\end{aligned}
\tag{$\mathbf{A}_0$}
$$

Finite elements for Problem $(\mathbf{B}_h)$. The first possibility is based on the $\mathcal{P}_{k+1} +$ bubble$/\mathcal{P}_k^{\text{discontinuous}}$ velocity-pressure approximation. We set for $k = 1$

$$
\begin{aligned}
\mathbf{L}_h &:= \{\mathbf{T} \in \mathbf{L}^{r'};\ \mathbf{T}_{|K} \in \mathcal{P}_1(K)^{d\times d}\ \forall K \in \mathcal{T}_h\},\\
\boldsymbol{W}_h &:= \{\boldsymbol{v} \in \boldsymbol{W}^{1,r}_{\mathsf{bc}};\ \boldsymbol{v}_{|K} \in \mathcal{P}_2(K)^d \oplus \mathcal{B}_{d+1}(K)^d\ \forall K \in \mathcal{T}_h\},
\end{aligned}
\tag{$\mathbf{B}_0$}
$$

where $\mathcal{B}_{d+1}(K)$ denotes the space of interior bubble functions of degree at most $d+1$ on K. However this choice does not fulfil the condition (ii) of Theorem 2, nevertheless we will consider it in our numerical tests.

Other approximation is based on the interiror penalty stabilization of $\operatorname{tr}\mathbf{T}$ and $\operatorname{div}\boldsymbol{v}$. Let $\mathcal{E}_h$ denote the set of edges in $\mathcal{T}_h$ and $[\![f]\!]$ the jump of f across a given edge. We set

$$
\begin{aligned}
\mathbf{L}_h &:= \{\mathbf{T} \in \mathbf{L}^{r'};\ \mathbf{T}_{|K} \in \mathcal{P}_k(K)^{d\times d}\ \forall K \in \mathcal{T}_h\},\\
\boldsymbol{W}_h &:= \{\boldsymbol{v} \in \boldsymbol{W}^{1,r}_{\mathsf{bc}};\ \boldsymbol{v}_{|K} \in \mathcal{P}_{k+1}(K)^d\ \forall K \in \mathcal{T}_h\}
\end{aligned}
\tag{$\mathbf{B}_1$}
$$

and define the stabilization operators

$$
\mathbf{C}_1 : \mathbb{R}^l \to \mathbb{R}^l, \quad (\mathbf{C}_1)_{ij} := \sum_{E\in\mathcal{E}_h} ([\![\operatorname{tr}\eta_i]\!], [\![\operatorname{tr}\eta_j]\!])_E\,,
$$

$$
\mathbf{C}_2 : \mathbb{R}^m \to \mathbb{R}^m, \quad (\mathbf{C}_2)_{ij} := \sum_{E\in\mathcal{E}_h} \big([\![\operatorname{div}\boldsymbol{\varphi}_i]\!], [\![\operatorname{div}\boldsymbol{\varphi}_j]\!]\big)_E\,.
$$

The stabilized variant of problem $(\mathbf{B}_{h,\mathbf{lin}})$ then reads:

$$
\begin{bmatrix} \mathbf{A}'(\bar{\boldsymbol{T}}, \bar{\boldsymbol{V}}) + \gamma\mathbf{C}_1 & \mathbf{B}_1'^{\top}(\bar{\boldsymbol{T}}, \bar{\boldsymbol{V}}) \\ B_2' & \gamma\mathbf{C}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{T} \\ \boldsymbol{V} \end{bmatrix} = \begin{bmatrix} \boldsymbol{F} \\ \boldsymbol{R} \end{bmatrix},
$$

where γ is arbitrary positive constant (a typical choice is $\gamma \approx h$).

Another case with a discontinuous pressure space, which we do not consider in this study, could be the $\mathcal{P}_{k+1}/\mathcal{P}_k^{\text{discontinuous}}$ Scott-Vogelius element, which in general leads to unstable approximations, however under some additional assumptions on the topology of $\mathcal{T}_h$ the stability can been established, see [39, 40, 41].

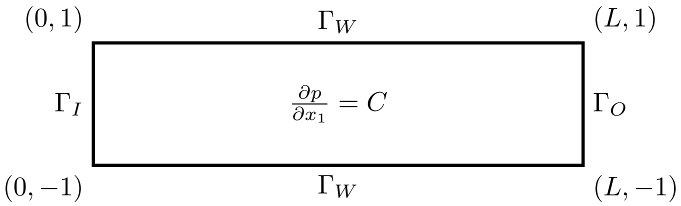


Figure 3. Poiseuille flow.

Finite elements for Problem ($\mathbf{C}_h$). The case of Problem ($\mathbf{C}_h$) can be discretized by directly extending the suggested discretizations of Problem ($\mathbf{A}_h$). For the particular choice of the finite-dimensional spaces, one can use the examples of $\mathbf{L}_{0h}$, $\boldsymbol{W}_h$ from the definitions ($\mathbf{A}_0$) and ($\mathbf{A}_1$) in the previous paragraph and extend the mixed finite element space by $\mathbf{D}_h := \{\mathbf{T}^\delta;\ \mathbf{T} \in \mathbf{L}_h\}$. Such extension will be detoted as ($\mathbf{CA}_0$) and ($\mathbf{CA}_1$).

In the rest of the text we compare the behavior of these different discretizations in several numerical tests. These comparisons are done for the stress–power-law model and for the Bingham fluid model. In both cases we test a simple case of Poiseuille flow, where the analytical solutions for these models are known, and then we solve the flow around cylinder and the driven cavity problem respectively.

4.2. *Stress–power-law model, analytical solution for Poiseuille flow*

Let us consider the stress–power-law model (13). In [18], the analytical solution to the plane Poiseuille flow problem for this model has been derived. The domain for the numerical solution is shown in Figure 3 and the following boundary conditions are prescribed on respective boundary parts:

$$\begin{aligned} \boldsymbol{v}_\tau &= 0 && \text{on } \Gamma_I := \{0\} \times (-1,1) \text{ and } \Gamma_O := \{L\} \times (-1,1) \\ \boldsymbol{v} &= \mathbf{0} && \text{on } \Gamma_W := (0,L) \times \{-1,1\}, \\ \mathbf{T}\boldsymbol{n} \cdot \boldsymbol{n} = -p + \mathbf{S}\boldsymbol{n} \cdot \boldsymbol{n} &= LC && \text{on } \Gamma_I, \\ \mathbf{T}\boldsymbol{n} \cdot \boldsymbol{n} = -p + \mathbf{S}\boldsymbol{n} \cdot \boldsymbol{n} &= 0 && \text{on } \Gamma_O, \end{aligned} \tag{38}$$

where L is the length of the domain and C is the prescribed normal stress, which reduces to the pressure in this case. If we use the stress–power-law model in the form

$$\operatorname{div} \boldsymbol{v} = 0, \qquad -\operatorname{div} \mathbf{T} = \mathbf{0}, \qquad \mathbf{D} = \frac{1}{2\nu_*}\left(1 + \frac{\beta_*}{(2\nu_*)^2}|\mathbf{T}^\delta|^2\right)^n \mathbf{T}^\delta, \qquad n = \frac{2-r}{2(r-1)},$$

with $\nu_* = \frac{1}{2}$ and $\beta_* = 1$, then the exact solution is given by

$$\begin{aligned} \boldsymbol{v}(x_1,x_2) &= \left(\frac{(1+2(Cx_2)^2)^{n+1} - (1+2C^2)^{n+1}}{2C(n+1)}, 0\right), \\ p(x_1,x_2) &= C(x_1 - L). \end{aligned} \tag{39}$$

For $r = 2.0$ the numerical solution is exact up to the floating point precision for the discretizations where the exact solution is a polynomial included in the finite element spaces. For other values of r this is not the case and we observe expected convergence rates, see [42, 43]. In Figure 4 we show the results for the power index $r = 1.4$, in Figure 5 we show the convergence plots for $r = 6.0$ as example.

4.3. *Stress–power-law model, flow around cylinder benchmark*

In the first example we compare the classical power-law (12) and stress–power-law (13) models on the benchmark problem of steady flow around a cylinder [44]. The geometry is depicted in Figure 6

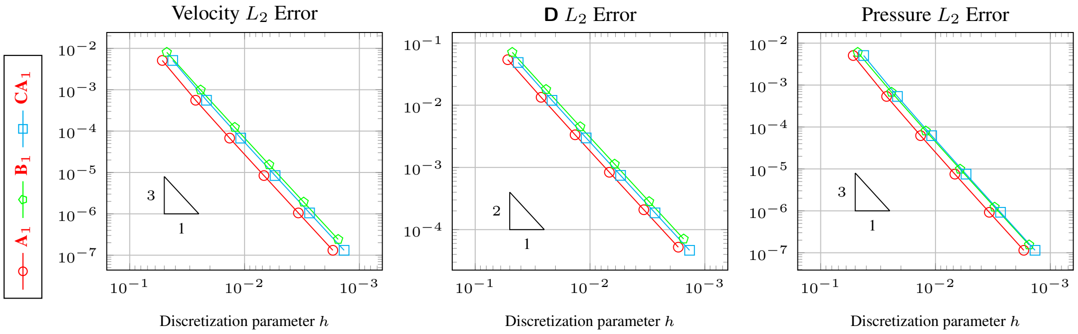


Figure 4. Convergence of the numerical solution to the analytical solution for the Poiseuille flow for the stress–power-law model with $r = 1.4$

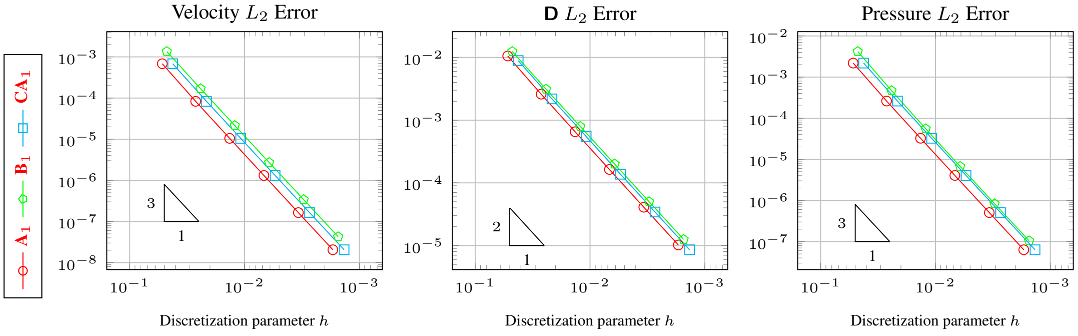


Figure 5. Convergence of the numerical solution to the analytical solution for the Poiseuille flow for stress–power-law model with $r = 6.0$

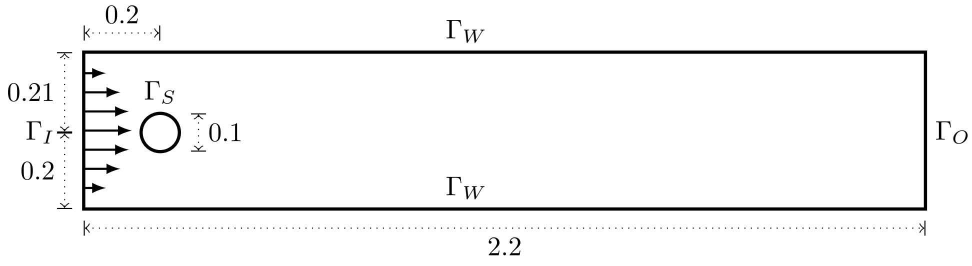


Figure 6. Flow around a cylinder.

and the following boundary conditions are prescribed:

$$
\begin{aligned}
\boldsymbol{v} &= \left(0.3\frac{4x_2(0.41-x_2)}{0.41^2}, 0\right) && \text{on } \Gamma_I := \{0\} \times (0, 0.41),\\
\boldsymbol{v} &= \boldsymbol{0} && \text{on } \Gamma_S \text{ and } \Gamma_W := (0, 2.2) \times \{0, 0.41\},\\
\mathbf{T}\boldsymbol{n}\cdot\boldsymbol{n} = -p + \mathbf{S}\boldsymbol{n}\cdot\boldsymbol{n} &= 0 && \text{on } \Gamma_O := \{2.2\} \times (0, 0.41),\\
\boldsymbol{v}_\tau &= 0 && \text{on } \Gamma_O.
\end{aligned}
$$

The reference power-law model was solved using the usual velocity-pressure formulation and $\mathcal{P}_2/\mathcal{P}_1$ Taylor–Hood elements. For the stress–power-law model

$$
\operatorname{div}\boldsymbol{v} = 0, \qquad \varrho_*(\boldsymbol{v}\cdot\nabla)\boldsymbol{v} - \operatorname{div}\mathbf{T} = \boldsymbol{0}, \qquad \mathbf{D} = \frac{1}{2\nu_*}\left(1 + \frac{\beta_*}{(2\nu_*)^2}|\mathbf{T}^\delta|^2\right)^n \mathbf{T}^\delta, \; n = \frac{2-r}{2(r-1)}
$$

we used the formulations $(\mathbf{A}_h)$-$(\mathbf{C}_h)$ with the selected finite element spaces $(\mathbf{A}_1)$, $(\mathbf{B}_1)$, and $(\mathbf{CA}_1)$. The material parameters are set to be $\varrho_* = \beta_* = 1$, $\nu_* = 10^{-3}$. Due to the small value of ν_* it is

	Δp		C_D		C_L	
r	$\mathbf{S}(\mathbf{D})$	$\mathbf{D}(\mathbf{S})$	$\mathbf{S}(\mathbf{D})$	$\mathbf{D}(\mathbf{S})$	$\mathbf{S}(\mathbf{D})$	$\mathbf{D}(\mathbf{S})$
1.4	$9.274{\cdot}10^{-2}$	$9.201{\cdot}10^{-2}$	$2.855{\cdot}10^{0}$	$2.797{\cdot}10^{0}$	$-1.075{\cdot}10^{-2}$	$-4.260{\cdot}10^{-3}$
1.6	$9.954{\cdot}10^{-2}$	$9.925{\cdot}10^{-2}$	$3.669{\cdot}10^{0}$	$3.645{\cdot}10^{0}$	$-1.040{\cdot}10^{-2}$	$-9.991{\cdot}10^{-3}$
1.8	$1.076{\cdot}10^{-1}$	$1.075{\cdot}10^{-1}$	$4.572{\cdot}10^{0}$	$4.565{\cdot}10^{0}$	$-2.695{\cdot}10^{-3}$	$-2.707{\cdot}10^{-3}$
2	$1.173{\cdot}10^{-1}$	$1.172{\cdot}10^{-1}$	$5.579{\cdot}10^{0}$	$5.579{\cdot}10^{0}$	$1.064{\cdot}10^{-2}$	$1.066{\cdot}10^{-2}$
3	$2.128{\cdot}10^{-1}$	$2.072{\cdot}10^{-1}$	$1.423{\cdot}10^{1}$	$1.387{\cdot}10^{1}$	$3.694{\cdot}10^{-1}$	$3.544{\cdot}10^{-1}$
4	$5.844{\cdot}10^{-1}$	$5.367{\cdot}10^{-1}$	$4.656{\cdot}10^{1}$	$4.330{\cdot}10^{1}$	$1.216{\cdot}10^{0}$	$1.180{\cdot}10^{0}$
6	$7.685{\cdot}10^{0}$	$6.556{\cdot}10^{0}$	$7.044{\cdot}10^{2}$	$6.136{\cdot}10^{2}$	$4.841{\cdot}10^{0}$	$4.142{\cdot}10^{0}$

Table I. Comparison of power-law ($\mathbf{S}(\mathbf{D})$) and stress–power-law ($\mathbf{D}(\mathbf{S})$). The reference values for $r = 2.0$ are $\Delta p = 0.118$, $C_D = 5.579$ and $C_L = 1.061 \cdot 10^{-2}$, see for example [44].

convenient to substitute $\tilde{\mathbf{T}} := \frac{1}{2\nu_*}\mathbf{T}$ which leads to the system

$$\operatorname{div} \boldsymbol{v} = 0, \qquad -2\nu_* \operatorname{div} \tilde{\mathbf{T}} = \mathbf{0}, \qquad \mathbf{D} = (1 + \beta_* |\tilde{\mathbf{T}}^\delta|^2)^n \tilde{\mathbf{T}}^\delta.$$

We evaluate the following quantities:

- pressure drop $\qquad \Delta p := p(A) - p(B)$,
- drag coefficient $\qquad C_D := 500 \displaystyle\int_{\Gamma_S} \mathbf{T}\boldsymbol{n} \cdot (1,0)^\top$,
- lift coefficient $\qquad C_L := 500 \displaystyle\int_{\Gamma_S} \mathbf{T}\boldsymbol{n} \cdot (0,1)^\top$,

where $A = (0.15, 0.2)$, $B = (0.25, 0.2)$, and Γ_S is the surface of the cylinder.

We observed that on a sufficiently fine mesh, all discrete formulations lead to very similar solutions with negligible differences. The results obtained on such a sufficiently refined mesh are presented in Figure 7. For $r = 2$ the results are comparable to the reference solution computed in [44] for the Navier–Stokes system, see Table I. In this case the results for the discretization ($\mathbf{CA}_1$) is identical to the corresponding smaller one ($\mathbf{A}_1$).

4.4. *Bingham model, analytical solution for Poiseuille flow*

Next we consider the Poiseuille flow problem, with the same setting as in Figure 3, with the Bingham fluid model with yield stress of type (2). As noted in the Section 1 the description (4) can be captured by implicit relation of the form (Bi-1), (Bi-2), (Bi-3), or (Bi-4). Now we want to investigate if any combination of such models with the finite element formulations ($\mathbf{A}_0$), ($\mathbf{A}_1$), ($\mathbf{B}_0$), ($\mathbf{B}_1$), ($\mathbf{CA}_0$) or ($\mathbf{CA}_1$) can lead to some advantage in the process of numerical solution.

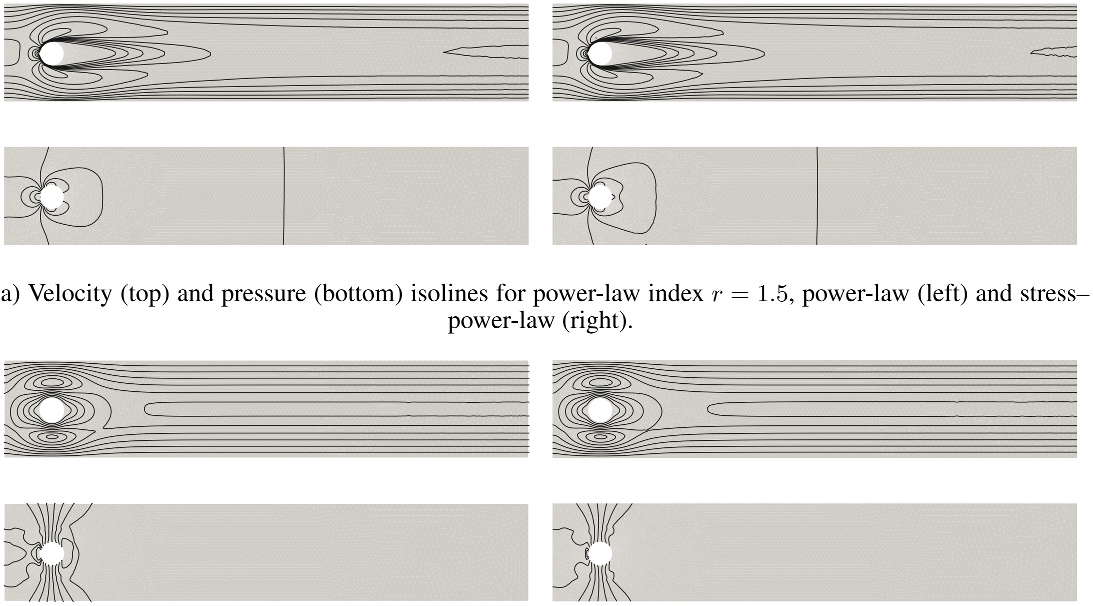
For the computations, the constitutive relations (Bi-1), (Bi-2), (Bi-3), or (Bi-4) are regularized by replacing the positive part $(x)^+ := \frac{1}{2}(|x| + x)$ by $(x)^+_\varepsilon := \frac{1}{2}(\sqrt{x^2 + \varepsilon^2} + x)$, where $\varepsilon \geq 0$ is the regularization parameter. The absolute value $|\cdot|$ is replaced by the regularized version $|x|^2_\varepsilon := x^2 + \varepsilon^2$. The regularized implicit functions are plotted for scalar variables in Figure 8 with large regularization $\varepsilon = 1$ and in the Figure 9 with smaller regularization parameter $\varepsilon = 1 \cdot 10^{-2}$.

The setup for the Poiseuille flow is the same as in Section 4.2. We solve the following system

$$\operatorname{div} \boldsymbol{v} = 0, \qquad -\operatorname{div} \mathbf{T} = \mathbf{0}, \qquad \mathbf{G}^\varepsilon_i(\mathbf{D}, \mathbf{T}^\delta) = \mathbf{0},$$

with $\nu = 1$, given value of τ_*, and equipped with the boundary conditions (38).

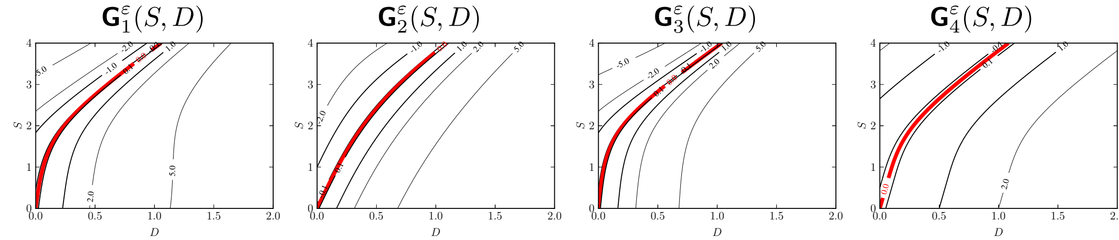
We start the solution process by setting the regularization parameter ε to sufficiently high value ($\varepsilon \approx 1$), which is decreased by factor 2 after solving the regularized problem by the Newton method to prescribed precision (L_2 norm of the residual vector less then $1 \cdot 10^{-10}$) until the regularization parameter ε is less then a prescribed value of $1 \cdot 10^{-8}$. More sophisticated algorithms for controlling the convergence with respect to the regularization can be found in [28, 45] or [46].



(a) Velocity (top) and pressure (bottom) isolines for power-law index $r = 1.5$, power-law (left) and stress–power-law (right).

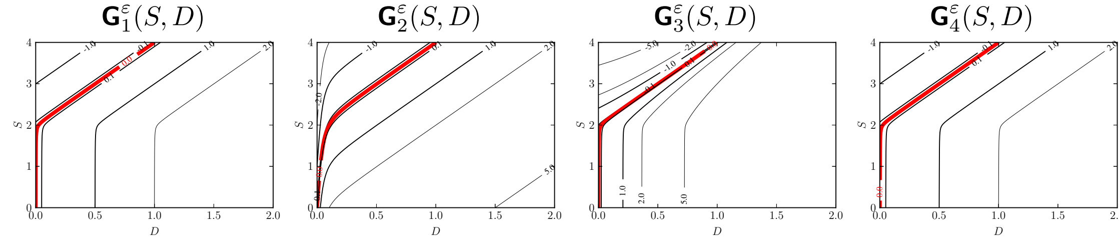
(b) Velocity (top) and pressure (bottom) isolines for power-law index $r = 5$, power-law (left) and stress–power-law (right).

Figure 7. Flow of stress–power-law fluid around cylinder. Comparison of classical power-law (left) with stress–power-law (right).



$\mathbf{G}_1^\varepsilon(S, D)$ $\mathbf{G}_2^\varepsilon(S, D)$ $\mathbf{G}_3^\varepsilon(S, D)$ $\mathbf{G}_4^\varepsilon(S, D)$

Figure 8. The implicit constitutive formulas for the Bingham model - regularized by $\varepsilon = 1$. Shown are the isolines with the 0-levelset marked in red.



$\mathbf{G}_1^\varepsilon(S, D)$ $\mathbf{G}_2^\varepsilon(S, D)$ $\mathbf{G}_3^\varepsilon(S, D)$ $\mathbf{G}_4^\varepsilon(S, D)$

Figure 9. The implicit constitutive formulas for the Bingham model - regularized by $\varepsilon = 1 \cdot 10^{-1}$. Shown are the isolines with the 0-levelset marked in red.

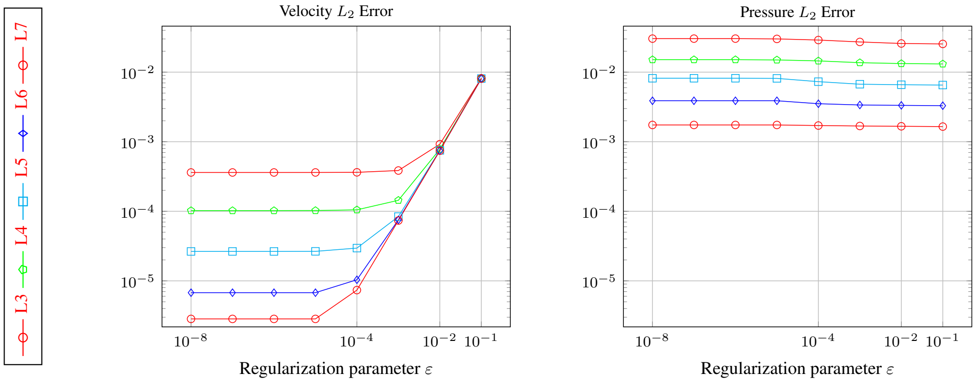


Figure 10. Errors for the Poiseuille flow depending on regularization parameter for discretization (**CA**$_0$) and model (Bi-2) for sequence of the mesh refinement levels.

Corresponding analytical solution for the Poiseuille flow in the case of Bingham type fluid can be derived, see for example [31]. Denoting by C the prescribed pressure gradient, the analytical solution reads:

$$\begin{aligned} \boldsymbol{v}(x_1,x_2) &= \begin{cases} (\frac{C}{2}(-x_2^2+1)-\tau_*(x_2+1),0) & -1<x_2<-\frac{\tau_*}{C}, \\ (\frac{C}{2}(-(\frac{\tau_*}{C})^2+1)-\tau_*(-\frac{\tau_*}{C}+1),0) & -\frac{\tau_*}{C}<x_2<\frac{\tau_*}{C}, \\ (\frac{C}{2}(-x_2^2+1)-\tau_*(-x_2+1),0) & \frac{\tau_*}{C}<x_2<1, \end{cases} \\ p(x_1,x_2) &= -C(x_1-L). \end{aligned} \tag{40}$$

In Figures 11 and 12 we show the convergence of particular discretizations and variants of the implicit constitutive formula to the analytical solution. It can be observed that most of the combinations converge in velocity as expected, while apart of the discretization (**CA**$_0$), the pressure error starts to decrease only on sufficiently fine meshes.

Over all, irrespective of the chosen discretization, the implicit form (Bi-4) does lead to discrete problems where either the number of nonlinear iterations needed to attain the desired precision increases with the mesh refinement or does not converge at all. The other observation is that the discretization (**B**$_0$) itself diverges for any of the implicit forms. This can be attributed to the fact that this choice of finite elements does not fulfil the condition (ii) of Theorem 2 and so is not covered by the Theorem 2.

The combination of discretizations (**A**$_0$), (**A**$_1$) and (**B**$_1$) together with models (Bi-1) (Bi-2) and (Bi-3) exhibit convergence in both velocity and pressure with expected order of convergence. The same is true for the extended discretizations (**CA**$_0$) and (**CA**$_1$) as can be seen in the Figure 12.

Figure 10 demonstrates that for given mesh refinement level decreasing the regularization parameter ε below certain value does not decrease the total error. Clearly at some point the discretization error becomes dominant and it would be of great interest to find adaptively such value of the regularization parameter ε which is sufficient to bring the regularization error bellow the discretization error on given mesh.

4.5. Bingham model, driven cavity flow benchmark

We compute the classical problem of lid driven cavity flow. The domain is the unit square $\Omega = (0,1)\times(0,1)$ and the material parameters are $\nu_* = 1$ and $\tau_* = 5, 50, 500$. At the top wall of the domain the velocity is prescribed to be $\boldsymbol{v}_D = (1,0)$ and no-slip boundary condition is prescribed on the remaining parts of the boundary. The boundary conditions for the driven cavity are:

$$\begin{aligned} \boldsymbol{v} &= (1,0) && \text{on } \Gamma_T, \\ \boldsymbol{v} &= \boldsymbol{0} && \text{on } \Gamma_W, \end{aligned}$$

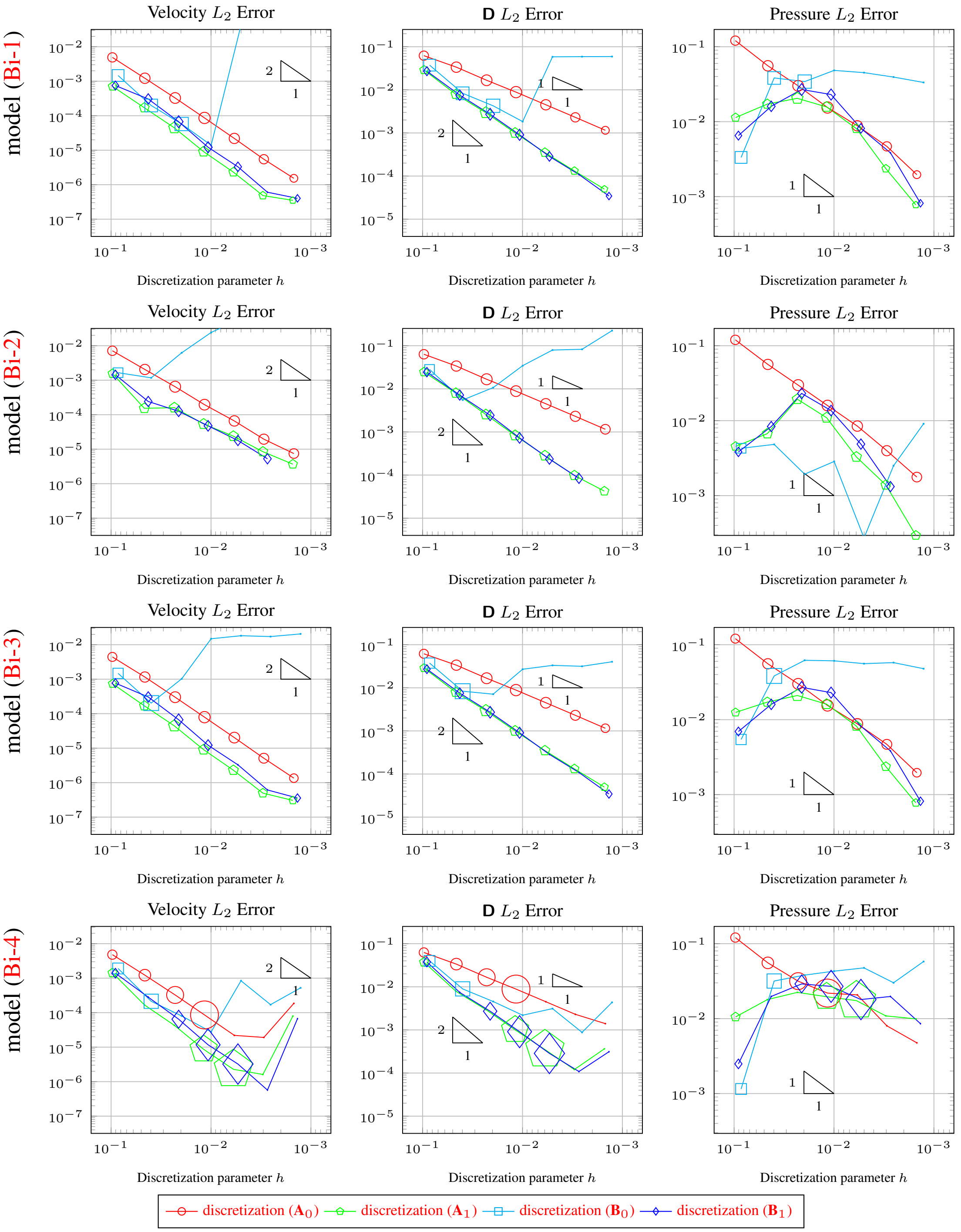


Figure 11. Convergence with mesh refinement for the Poiseuille flow using different Bingham models with $\tau_* = 0.2$ and discretization variants ($\mathbf{A}_0$), ($\mathbf{A}_1$), ($\mathbf{B}_0$) and ($\mathbf{B}_1$); (regularization $\varepsilon = 1 \cdot 10^{-8}$); Size of each markers is proportional to the number of nonlinear iterations needed to converge below prescribed precision and is missing if the nonlinear solver did not converge below the prescribed precision in less than 2000 iterations.

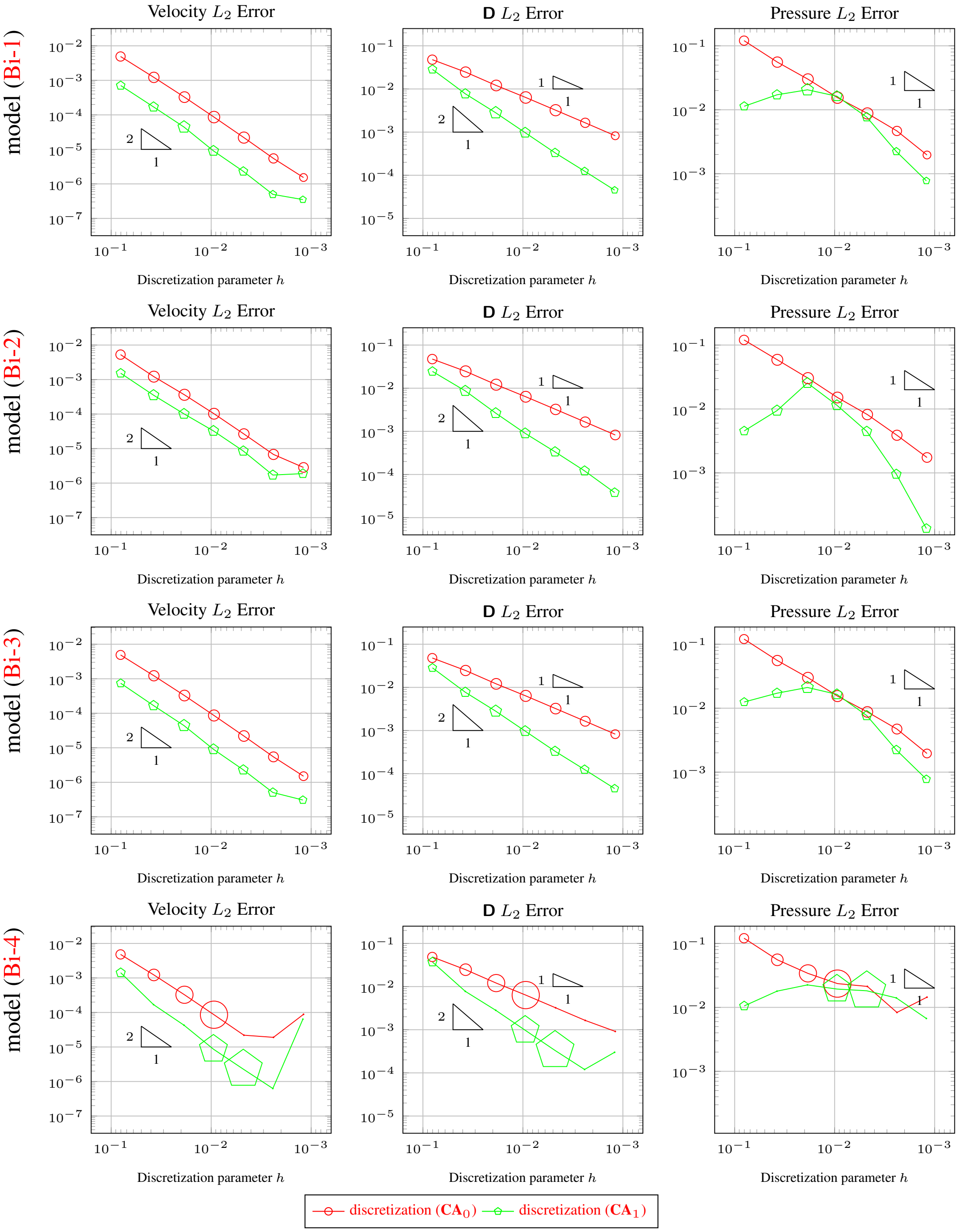
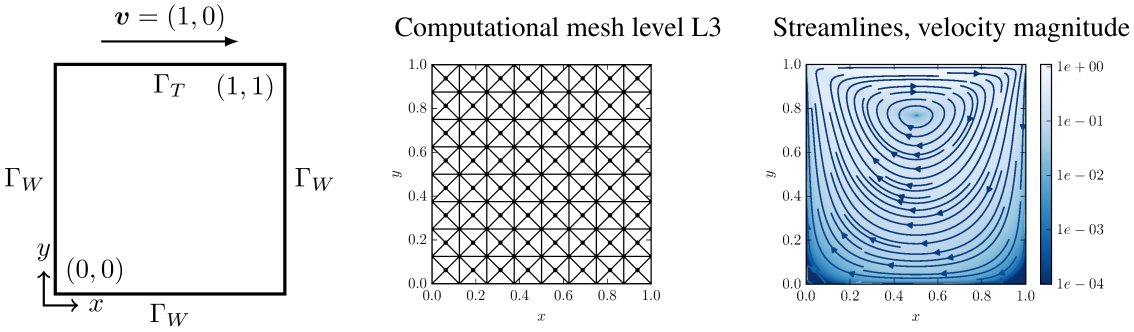


Figure 12. Convergence with mesh refinement for the Poiseuille flow using different Bingham models with $\tau_* = 0.2$ and discretization variants (**CA**$_0$) and (**CA**$_1$); (regularization $\varepsilon = 1 \cdot 10^{-8}$); Size of each markers is proportional to the number of nonlinear iterations needed to converge below prescribed precision and is missing if the nonlinear solver did not converge below the prescribed precision in less than 2000 iterations.



Computational mesh level L3

Streamlines, velocity magnitude

Figure 13. Boundary conditions for the driven cavity problem, computational coarse mesh, and the solution for the Stokes equation with constant viscosity. Shown are the streamlines and the color represents the velocity magnitude.

mesh refinement level	number of elements	number of dofs for discretization (**CA**$_0$)
L1	16	290
L2	64	1 122
L3	256	4 418
L4	1 024	17 538
L5	4 096	69 890
L6	16 384	279 042
L7	65 536	1 115 138

Table II. Mesh refinement levels for the driven cavity problem.

where Γ_T and Γ_W are defined in Figure 13. To obtain unique pressure we include the following additional condition

$$\int_\Omega \operatorname{tr} \mathbf{T} \, \mathrm{d}x = 0,$$

by means of Lagrange multiplier. This leads to the system:

$$\operatorname{div} \boldsymbol{v} = 0, \quad -\operatorname{div} \mathbf{T} = \mathbf{0}, \quad \mathbf{G}^{\varepsilon}(\mathbf{D}, \mathbf{T}^{\delta}) = \mathbf{0}, \quad \int_\Omega \operatorname{tr} \mathbf{T} \, \mathrm{d}x = 0.$$

We use the same procedure to solve the nonlinear problem as in the previous section. We start by setting the regularization parameter ε to sufficiently high value ($\varepsilon \approx 10$), which is then decreased by factor 2 whenever the Newton method converges to prescribed precision (L_2 norm of the residual vector less then $1 \cdot 10^{-10}$). This is repeated until the regularization parameter ε is less then a prescribed value of $1 \cdot 10^{-8}$.

We observe that this procedure applied to the driven cavity problem converges independently of the mesh refinement only for the extended discretizations (**CA**$_0$) and (**CA**$_1$). In all the other cases the convergence deteriorates with the mesh refinement. The same effect is observed with respect to the choice of the implicit relation. The models (Bi-1), (Bi-3) and (Bi-4) do not lead to a successful convergence of our simple nonlinear solution procedure, once the spatial discretization is fine enough, as can be seen in Figure 14 for the mesh L6.

On the other hand the combination of model (Bi-2) and the discretization (**CA**$_0$) can be solved with the number of nonlinear iterations almost independent on the mesh refinement, as can be seen in Figure 15, and also with very mild dependence of the number of nonlinear iterations with respect to the yield stress parameter τ_* as demonstrated in Figure 16. The resulting flow fields are shown in Figure 17 for yield stress thresholds $\tau_* = 5, 50, 500$ and in Figure 18 the solution for $\tau_* = 50$ is compared for range of mesh refinements as shown in Table II.

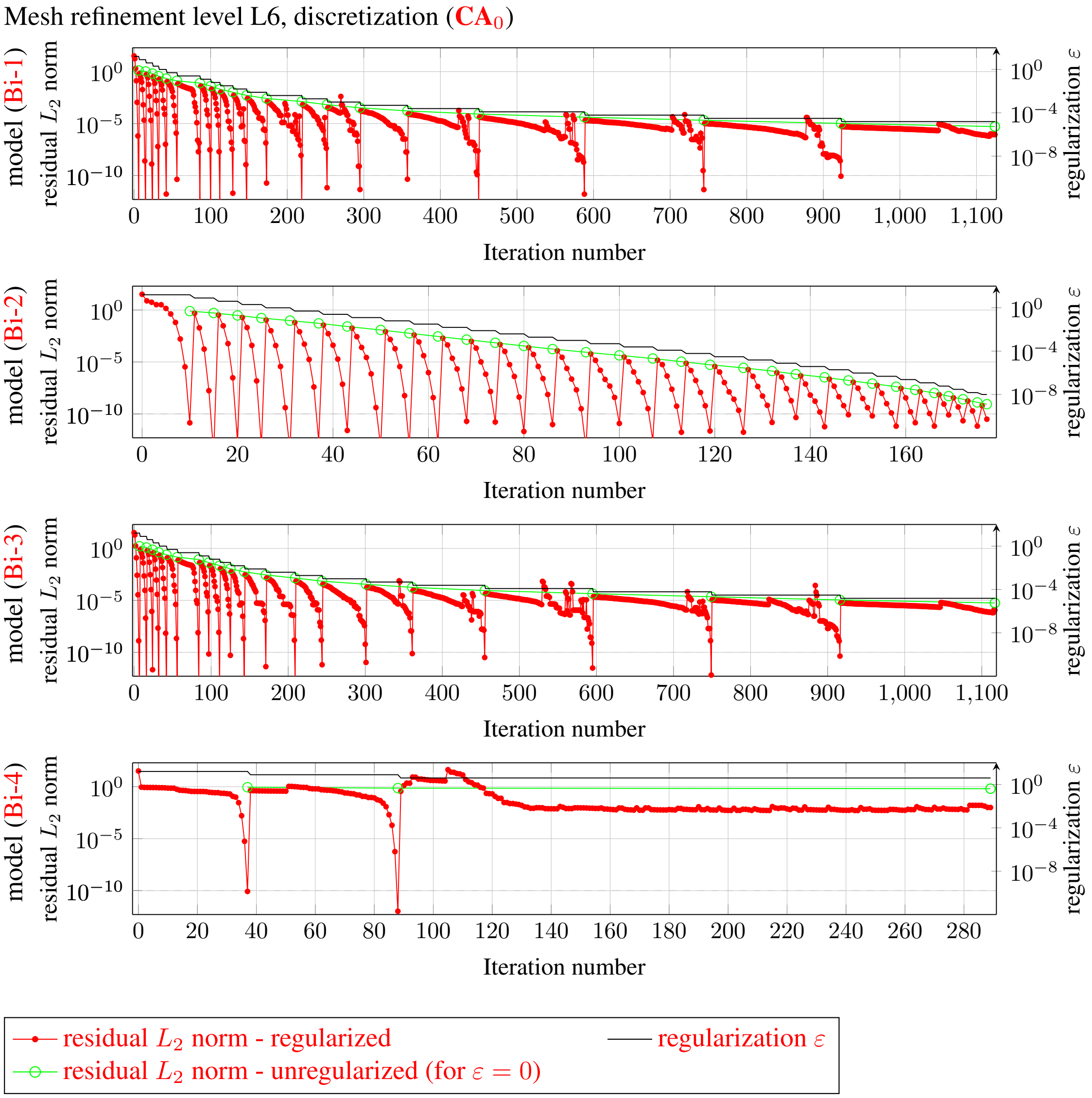


Figure 14. Convergence process for the driven cavity problem and Bingham model, $\tau_* = 50$. Iteration number counts number of the Newton method iterations (i.e. each iteration represents one linear problem solve). The black line represents the value of the regularization parameter ε for the given iteration. The red line shows the residuum of the regularized problem at each iteration, the value of ε is decreased by factor 2 whenever this residuum decreases below $1 \cdot 10^{-10}$. The green line shows the residuum of the unregularized problem at the given iteration (i.e. for $\varepsilon = 0$).

5. CONCLUSION

In this study we followed the three aims as stated in the introduction of this paper. The novel approach is based on formulating the models in the framework of implicit constitutive equations. First we investigate problems associated with the Bingham and stress–power-law fluids using the implicit formulation of the type (5) rather than (2). We introduced three variants of weak formulations and showed well-posedness of the resulting linearized discrete problems for conforming mixed finite element discretization satisfying additional requirement for stability, see Theorems 1 and 2.

Then the selected discretizations were tested on Poiseuille flow problem where the analytic solution is known for both fluids: stress–power-law and Bingham type constitutive model. Finally a

Model (Bi-2), discretization ($\mathbf{CA}_0$)

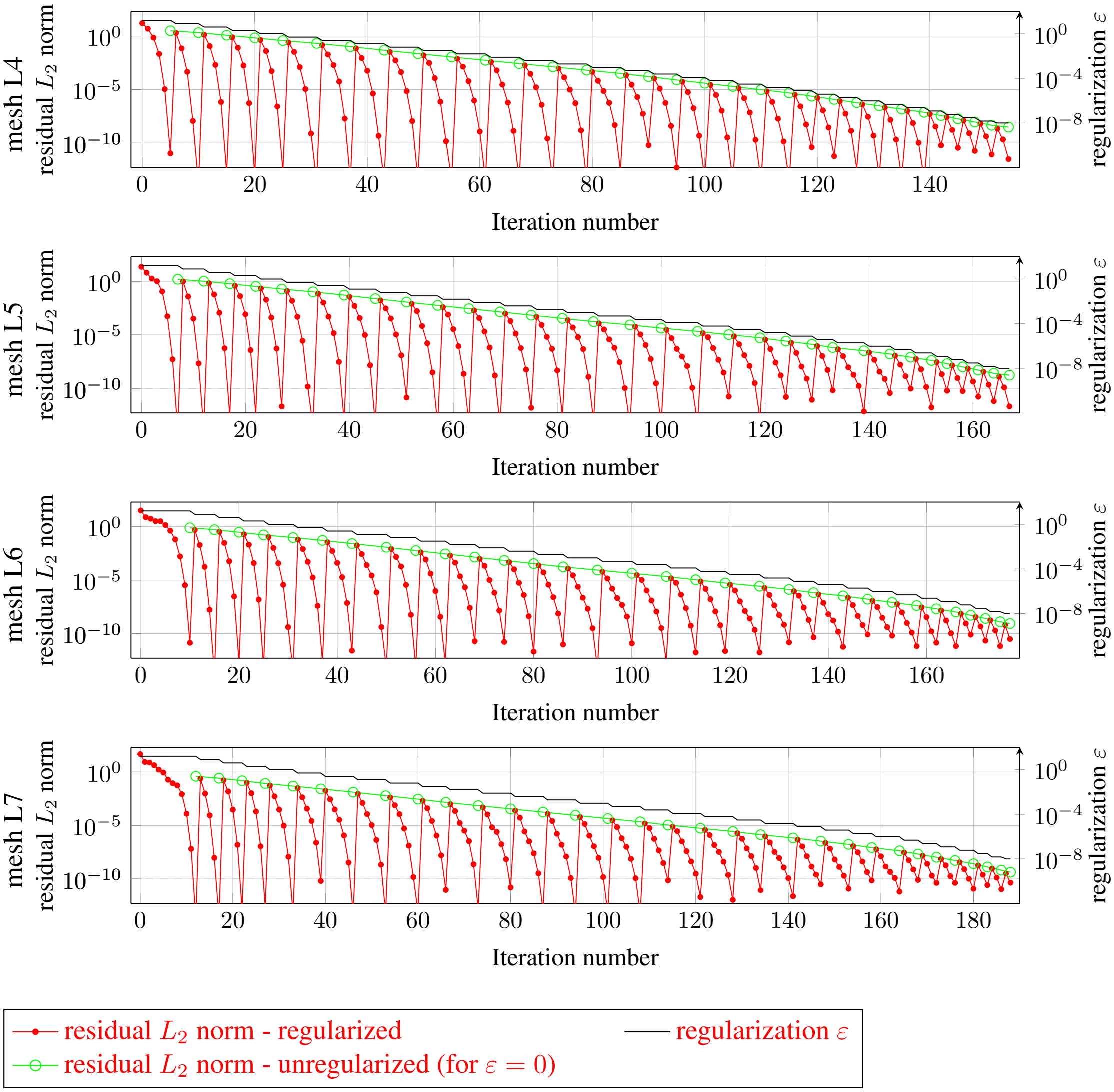


Figure 15. Convergence process for the driven cavity problem and Bingham model, $\tau_* = 5, 50, 500$. See Figure 14 for detailed explanation.

nontrivial benchmark test for each of the fluid models is presented: the flow around cylinder for the stress–power-law fluid, and the driven cavity problem is computed for the Bingham fluid.

The observations for the stress–power-law model is that all of the selected conforming discretizations do perform equally well in wide range of the power-law index r as demonstrated in Figures 4 and 5 for the Poiseuille flow and in Section 4.3 for the flow around cylinder.

For the Bingham model, additional regularization is introduced in order to use a simple Newton solver as described in Section 4.4. The results for the Poiseuille flow are summarized in Figures 11 and 12 and clearly demonstrate the difficulty of this problem. From all the tested combinations of implicit forms of the Bingham models and selected discretizations, we can identify that for the Poiseuille flow the implicit relations (Bi-1), (Bi-2) and (Bi-3) combined with the discretizations ($\mathbf{A}_0$), ($\mathbf{A}_1$), ($\mathbf{CA}_0$), ($\mathbf{CA}_1$) and ($\mathbf{B}_1$) show expected convergence with mesh refinement and are solvable by Newton method with nonlinear iteration count independent of the mesh refinement. However, for the driven cavity problem Figure 14 shows taht only the implicit model in the form (Bi-2) together with the discretization ($\mathbf{CA}_0$) was solvable by the simple iterative procedure based on the Newton method with iteration count independent on the mesh refinement (see Figure 15) and

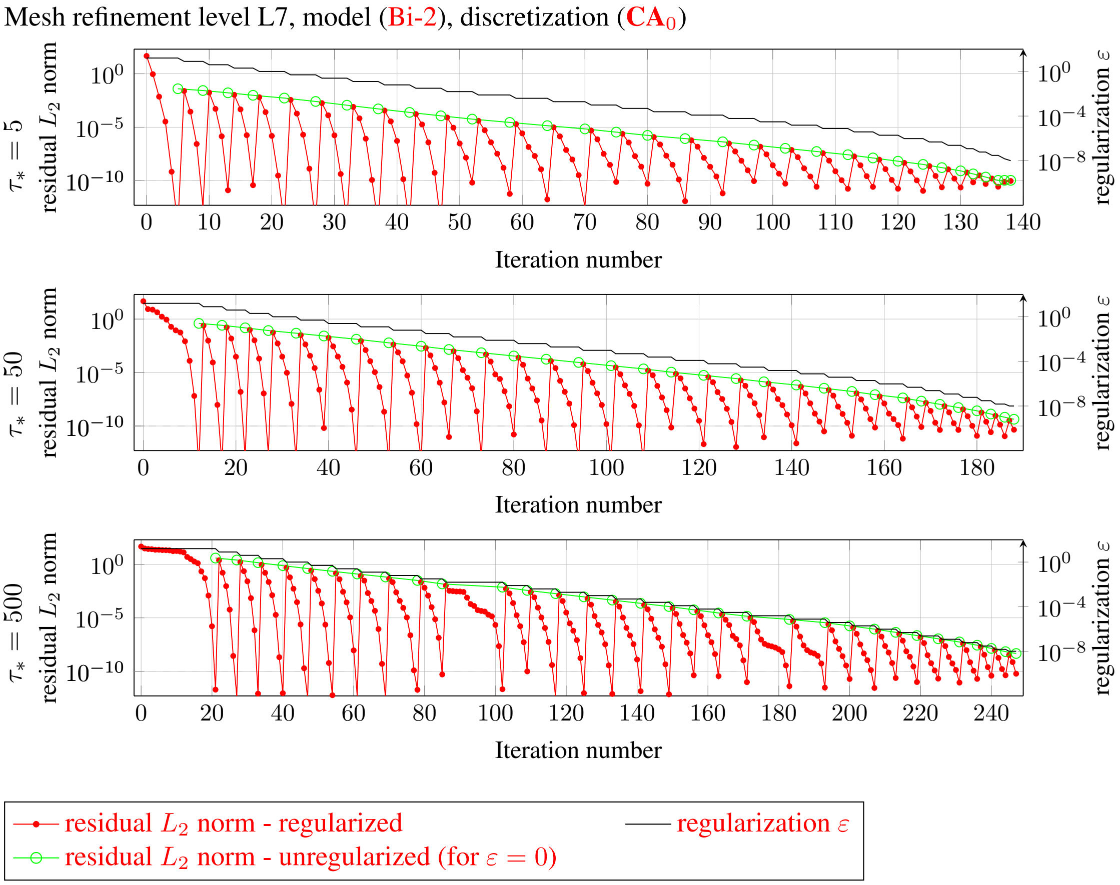


Figure 16. Convergence process for the driven cavity problem and Bingham model, $\tau_* = 5, 50, 500$. See Figure 14 for detailed explanation.

with arbitrary small regularization parameter ε. Moreover for this specific combination the number of nonlinear iterations is almost independent on the yield parameter τ_* for the driven cavity problem, see Figure 16. This fulfils the remaining aims of this study. In all of the computations performed in this study we tried to ensure that the resulting linear problems are solvable; however we used a direct solver to solve the linearized system. It remains for further investigation to identify a suitable iterative procedure and preconditioning to solve effectively such mixed systems as attempted for example in [47, 48].

REFERENCES

[1] Balmforth NJ, Frigaard IA, Ovarlez G. Yielding to stress: Recent developments in viscoplastic fluid mechanics. *Annual Review of Fluid Mechanics* 2014; **46**(1):121–146, doi:10.1146/annurev-fluid-010313-141424.

[2] de Bruyn JR, Moyers-Gonzalez M, Frigaard I. Viscoplastic Fluids from Theory to Application: 10 Years On. *Journal of Non Newtonian Fluid Mechanics* 2016; **238**:1–5, doi:10.1016/j.jnnfm.2016.11.008.

[3] Bingham EC. Plastic flow. *Journal of the Franklin Institute* 1916; **181**(6):845–848, doi:10.1016/S0016-0032(16)90156-X.

[4] Duvaut G, Lions JL. *Inequalities in mechanics and physics*. Springer: Berlin, 1976.

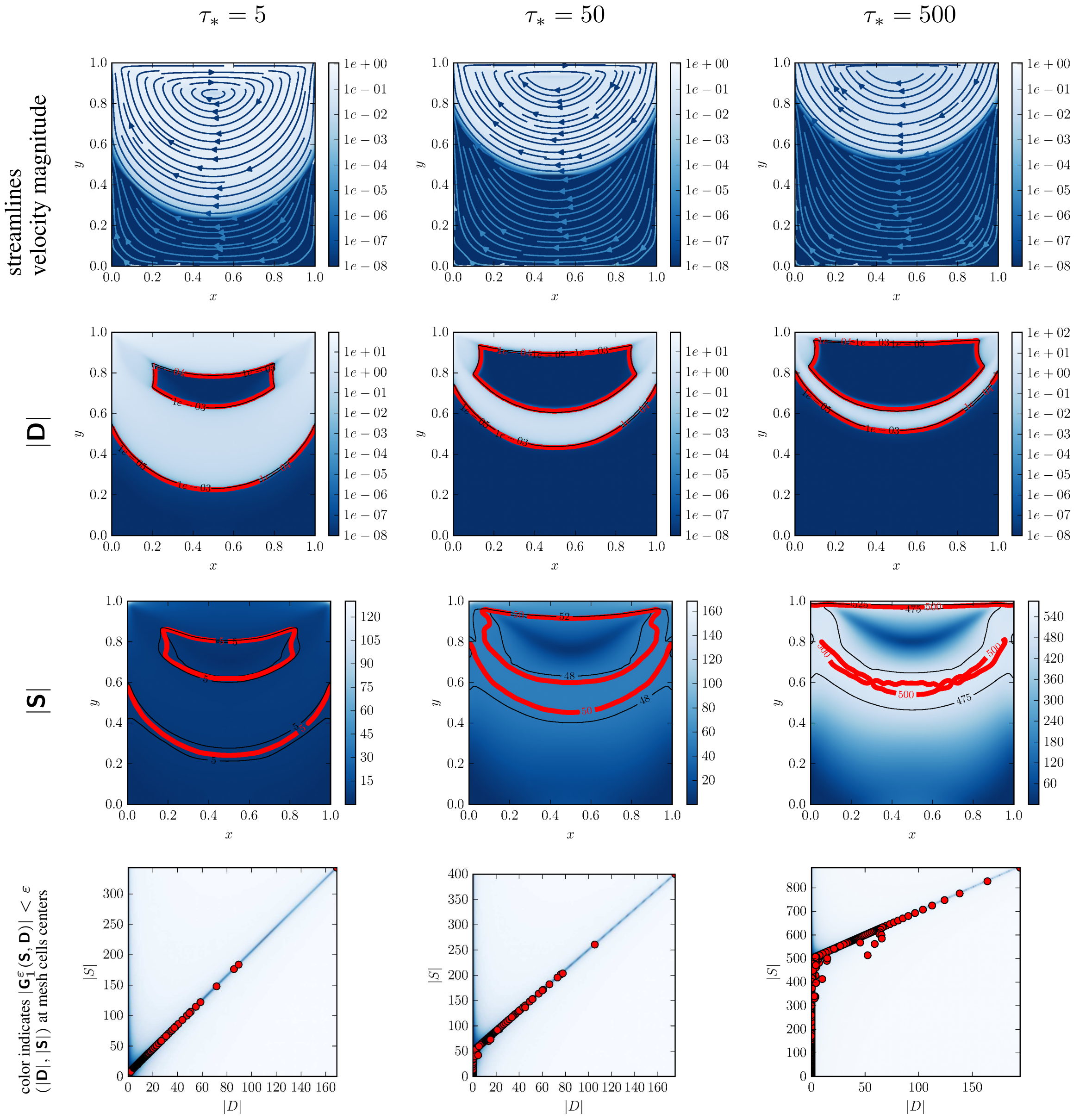


Figure 17. The computed flow field for $\tau_* = 5, 50, 500$. The red isoline in the $|\mathbf{D}|$ and $|\mathbf{S}|$ plots indicates the yield surface. Regularized by $\varepsilon = 1 \cdot 10^{-8}$.

[5] Málek J, Průša V. Derivation of equations for continuum mechanics and thermodynamics of fluids. *Handbook of Mathematical Analysis in Mechanics of Viscous Fluids*, Giga Y, Novotny A (eds.). chap. 6, Springer-Verlag GmbH, 2016.

[6] Ionescu IR, Sofonea M. *Functional and Numerical Methods in Viscoplasticity*. Clarendon Press Oxford Mathematical Monographs, 1993.

[7] Rajagopal KR, Srinivasa AR. On the thermodynamics of fluids defined by implicit constitutive relations. *Zeitschrift für Angewandte Mathematik und Physik* 2008; **59**(4):715–729.

[8] Málek J, Rajagopal KR. Compressible generalized Newtonian fluids. *Zeitschrift fuer Angewandte Mathematik und Physik (ZAMP)* 2010; **61**:1097–1110.

[9] Bulíček M, Gwiazda P, Málek J, Świerczewska-Gwiazda A. On unsteady flows of implicitly constituted incompressible fluids. *SIAM J. Math. Anal.* 2012; **44**(4):2756–2801.

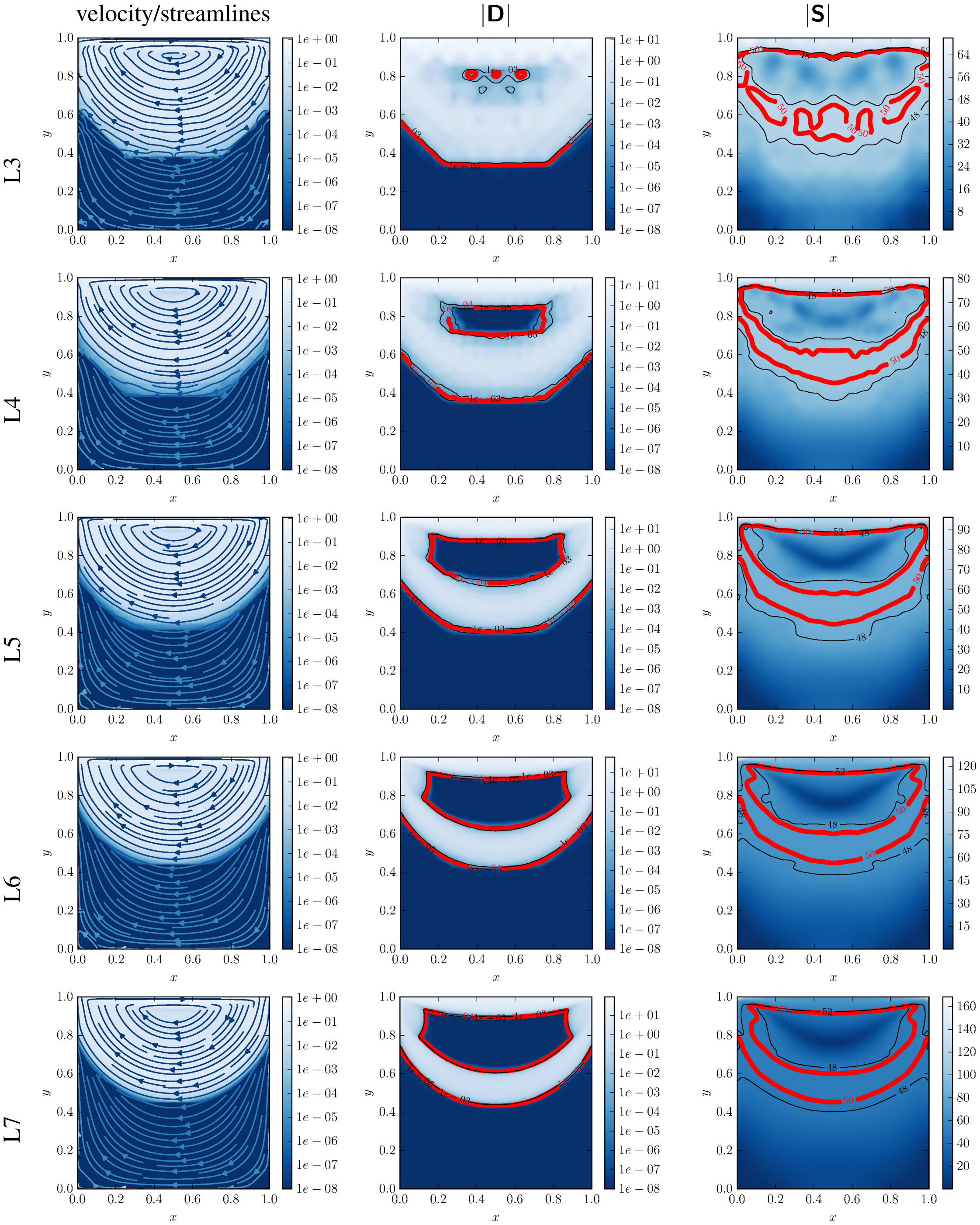


Figure 18. Comparison of the solution (velocity, |**D**| and |**S**|) on recursively refined meshes L3, L4, L5, L6 and L7 (see Table II); $\tau_* = 50$; The red isoline in the |**D**| and |**S**| plots indicates the yield surface. Regularized by $\varepsilon = 1 \cdot 10^{-8}$.

[10] Rajagopal KR. On implicit constitutive theories. *Appl. Math.* 2003; **48**(4):279–319.

[11] Barus C. Isothermals, isopiestics and isometrics relative to viscosity. *American Jour. Sci.* 1893; **45**:87–96.

[12] Bridgman PW. *The physics of high pressure*. the MacMillan Company: New York, 1931.

[13] Rajagopal KR. On implicit constitutive theories for fluids. *J. Fluid Mech.* 2006; **550**:243–249.

[14] Rajagopal KR. Conspectus of concepts of elasticity. *Mathematics and Mechanics of Solids* 2011; **16**(5):536–562.

[15] Heida M, Málek J. On compressible Korteweg fluid-like materials. *Internat. J. Engrg. Sci.* 2010; **48**(11):1313–1324.

[16] Heida M, Málek J, Rajagopal KR. On the development and generalizations of Cahn-Hilliard equations within a thermodynamic framework. *Z. Angew. Math. Phys.* 2012; **63**(1):145–169.

[17] Heida M, Málek J, Rajagopal KR. On the development and generalizations of Allen-Cahn and Stefan equations within a thermodynamic framework. *Z. Angew. Math. Phys.* 2012; **63**(1):759–776.

[18] Málek J, Průša V, Rajagopal KR. Generalizations of the Navier-Stokes fluid from a new perspective. *International Journal of Engineering Science* 2010; **48**(12):1907 – 1924.

[19] Bulíček M, Gwiazda P, Málek J, Świerczewska-Gwiazda A. On steady flows of incompressible fluids with implicit power-law-like rheology. *Adv. Calc. Var.* 2009; **2**:109–136.

[20] Bulíček M, Gwiazda P, Málek J, Rajagopal KR, Świerczewska-Gwiazda A. On flows of fluids described by an implicit constitutive equation characterized by a maximal monotone graph. *Partial differential equations and fluid mechanics*, *London Math. Soc. Lecture Note Ser.*, vol. 402, Robinson JC, Rodrigo JL, Sadowski W (eds.). Cambridge Univ. Press: Cambridge, 2012; 23–51, doi:10.1017/CBO9781139235792.003.

[21] Diening L, Kreuzer C, Süli E. Finite element approximation of steady flows of incompressible fluids with implicit power-law rheology. *SIAM J. Numer. Anal.* 2012; **51**(2):984–1015.

[22] Kreuzer C, Süli E. Adaptive finite element approximation of steady flows of incompressible fluids with implicit power-law-like rheology. *ESAIM: Mathematical Modelling and Numerical Analysis* jul 2016; **50**(5):1333–1369, doi:10.1051/m2an/2015085.

[23] Diening L, Růžička M, Wolf J. Existence of weak solutions for unsteady motions of generalized newtonian fluids. *ANNALI DELLA SCUOLA NORMALE SUPERIORE - CLASSE DI SCIENZE* 2010; (1):1–46, doi:10.2422/2036-2145.2010.1.01.

[24] Breit D, Diening L, Schwarzacher S. Solenoidal Lipschitz truncation for parabolic PDEs. *Mathematical Models and Methods in Applied Sciences* 2013; **23**(14):2671–2700, doi:10.1142/s0218202513500437.

[25] Bulíček M, Málek J. On unsteady internal flows of Bingham fluids subject to threshold slip on the impermeable boundary. *Recent Developments of Mathematical Fluid Mechanics*. Springer Nature, 2016; 135–156, doi:10.1007/978-3-0348-0939-9_8.

[26] Chupin L, Mathe J. Existence theorem for homogeneous incompressible Navier–Stokes equation with variable rheology. *European Journal of Mechanics - B/Fluids* 2017; **61**:135–143, doi:10.1016/j.euromechflu.2016.09.020.

[27] Saramito P. A damped Newton algorithm for computing viscoplastic fluid flows. *Journal of Non Newtonian Fluid Mechanics* 2016; **238**:6–15, doi:10.1016/j.jnnfm.2016.05.007.

[28] De los Reyes JC, Andrade SG. Numerical simulation of two-dimensional Bingham fluid flow by semismooth Newton methods. *Journal of Computational and Applied Mathematics* 2010; **235**(1):11–32.

[29] Faria CO, Karam Filho J. A regularized–stabilized mixed finite element formulation for viscoplasticity of Bingham type. *Computers and Mathematics with Applications* 2013; **66**(6):975–995.

[30] dos Santos DDO, Frey S, Naccache MF, de Souza Mendes PR. Numerical approximations for flow of viscoplastic fluids in a lid-driven cavity. *Journal of Non Newtonian Fluid Mechanics* 2011; **166**(12-13):667–679.

[31] Aposporidis A, Haber E, Olshanskii MA, Veneziani A. A mixed formulation of the Bingham fluid flow problem: Analysis and numerical solution. *Computer Methods in Applied Mechanics and Engineering* 2011; **200**(29-32):2434–2446.

[32] Dean E, Glowinski R, Guidoboni G. On the numerical simulation of Bingham visco-plastic flow: Old and new results. *Journal of Non Newtonian Fluid Mechanics* 2007; **142**(1-3):36–62.

[33] Muravleva EA, Olshanskii MA. Two finite-difference schemes for calculation of Bingham fluid flows in a cavity. *Russian Journal of Numerical Analysis and Mathematical Modelling* 2008; **23**(6):615–634.

[34] Howell J, Walkington N. Inf–sup conditions for twofold saddle point problems. *Numerische Mathematik* 2011; **118**(4):663–693.

[35] Logg A, Mardal KA, Wells GN ((eds.)). *Automated Solution of Differential Equations by the Finite Element Method*. Springer Berlin Heidelberg, 2012.

[36] Alnæs MS, Blechta J, Hake J, Johansson A, Kehlet B, Logg A, Richardson C, Ring J, Rognes ME, Wells GN. The FEniCS project version 1.5. *Archive of Numerical Software* 2015; **3**(100), doi: 10.11588/ans.2015.100.20553.

[37] Balay S, Abhyankar S, Adams MF, Brown J, Brune P, Buschelman K, Dalcin L, Eijkhout V, Gropp WD, Kaushik D, *et al.*. PETSc users manual. *Technical Report ANL-95/11 - Revision 3.6*, Argonne National Laboratory 2015. URL http://www.mcs.anl.gov/petsc.

[38] Amestoy P, Duff I, L'Excellent JY. Multifrontal parallel distributed symmetric and unsymmetric solvers. *Computer Methods in Applied Mechanics and Engineering* 2000; **184**(2–4):501 – 520, doi: http://dx.doi.org/10.1016/S0045-7825(99)00242-X.

[39] Scott L, Vogelius M. Norm estimates for a maximal right inverse of the divergence operator in spaces of piecewise polynomials. *Modélisation mathématique et analyse numérique* 1985; **19**(1):111–143.

[40] Qin J. On the convergence of some low order mixed finite elements for incompressible fluids. PhD Thesis, The Pennsylvania State University 1994.

[41] Zhang S. A new family of stable mixed finite elements for the 3D Stokes equations. *Mathematics of computation* 2005; **74**(250):543–554.

[42] Brezzi F, Fortin M ((eds.)). *Mixed and Hybrid Finite Element Methods*. Springer New York, 1991, doi:10.1007/978-1-4612-3172-1.

[43] Boffi D, Brezzi F, Fortin M. *Mixed Finite Element Methods and Applications*. Springer Berlin Heidelberg, 2013, doi:10.1007/978-3-642-36519-5.

[44] John V, Matthies G. Higher-order finite element discretizations in a benchmark problem for incompressible flows. *International Journal for Numerical Methods in Fluids* 2001; **37**(8):885–903.

[45] Mandal S, Ouazzi A, Turek S. Modified Newton solver for yield stress fluids. *Lecture Notes in Computational Science and Engineering*. Springer Nature, 2016; 481–490, doi:10.1007/978-3-319-39929-4_46.

[46] De los Reyes JC, González S. Path following methods for steady laminar Bingham flow in cylindrical pipes. *Mathematical Modelling and Numerical Analysis* 2008; **43**(1):81–117, doi:10.1051/m2an/2008039.

[47] Grinevich PP, Olshanskii MA. An iterative method for the Stokes-type problem with variable viscosity. *SIAM Journal on Scientific Computing* 2009; **31**(5):3959–3978, doi:10.1137/08744803.

[48] He X, Neytcheva M, Vuik C. On preconditioning of incompressible non-Newtonian flow problems. *Journal of Computational Mathematics* 2014; **33**(1):33–58, doi:10.4208/jcm.1407-m4486.