NUMERICAL ANALYSIS OF UNSTEADY IMPLICITLY CONSTITUTED INCOMPRESSIBLE FLUIDS: THREE-FIELD FORMULATION*

P. E. FARRELL†, P. A. GAZCA-OROZCO‡, AND E. SÜLI§

Abstract. In the classical theory of fluid mechanics a linear relationship between the shear stress and the symmetric velocity gradient tensor is often assumed. Even when a nonlinear relationship is assumed, it is typically formulated in terms of an explicit relation. Implicit constitutive models provide a theoretical framework that generalises this, allowing for general implicit constitutive relations. Since it is generally not possible to solve explicitly for the shear stress in the constitutive relation, a natural approach is to include the shear stress as a fundamental unknown in the formulation of the problem. In this work we present a mixed formulation with this feature, discuss its solvability and approximation using mixed finite element methods, and explore the convergence of the numerical approximations to a weak solution of the model.

Key words. Implicitly constituted models, non–Newtonian fluids, finite element method

AMS subject classifications. 65M60, 65M12, 35Q35, 76A05

1. Implicitly constituted models.

In the classical theory of continuum mechanics the balance laws of momentum, mass, and energy do not determine completely the behaviour of a system. Additional information that captures the specific properties of the material to be studied is needed; this is what is commonly known as a *constitutive relation*. The constitutive law usually expresses the stress tensor in terms of other kinematical quantities (e.g. the symmetric velocity gradient) and, even if it is nonlinear, it is typically formulated by means of an explicit relationship. It has been known for some time that in many cases explicit constitutive relations are not adequate when modeling materials with viscoelastic or inelastic responses (see e.g. [51, 52]), which has led to the introduction of many ad-hoc models that try to fit the experimental data. Implicitly constituted models, introduced in [51], provide a theoretical framework that not only serves to justify these ad-hoc models, but also generalises them. The physical justification of these types of models, including a study of their thermodynamical consistency, is available and will not be discussed here; the interested reader is referred to [53, 52, 54].

If a fluid occupies part of a space represented by a simply-connected open set $\Omega \subset \mathbb{R}^d$, where $d \in \{2,3\}$, then the evolution of the system during a given time interval $[0,T)$, for $T > 0$, is determined by the usual equations of balance of mass, momentum, angular momentum and energy, which in Eulerian coordinates take the

*Submitted to the editors April 2019.

Funding: This research is supported by the Engineering and Physical Sciences Research Council [grant numbers EP/K030930/1, EP/R029423/1], and by the EPSRC Centre For Doctoral Training in Partial Differential Equations: Analysis and Applications [grant number EP/L015811/1]. The second author was partially supported by CONACyT (Scholarship 438269).

†Mathematical Institute, University of Oxford, UK (patrick.farrell@maths.ox.ac.uk).

‡Mathematical Institute, University of Oxford, UK (gazcaorozco@maths.ox.ac.uk).

§Mathematical Institute, University of Oxford, UK (endre.suli@maths.ox.ac.uk).

form:

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \boldsymbol{u}) = 0,$$
$$\frac{\partial(\rho \boldsymbol{u})}{\partial t} + \operatorname{div}(\rho \boldsymbol{u} \otimes \boldsymbol{u}) = \operatorname{div} \boldsymbol{T} + \rho \boldsymbol{f}, \tag{1.1}$$
$$\boldsymbol{T} = \boldsymbol{T}^{\mathrm{T}},$$
$$\frac{\partial(\rho e)}{\partial t} + \operatorname{div}(\rho e \boldsymbol{u}) = \operatorname{div}(\boldsymbol{T}\boldsymbol{u} - \boldsymbol{q}).$$

Here:

- $\boldsymbol{u} : [0,T) \times \overline{\Omega} \to \mathbb{R}^d$ is the velocity field;
- $\rho : [0,T) \times \overline{\Omega} \to \mathbb{R}$ is the density;
- $\boldsymbol{T} : (0,T) \times \overline{\Omega} \to \mathbb{R}^{d\times d}$ is the Cauchy stress;
- $e : [0,T) \times \overline{\Omega} \to \mathbb{R}$ is the internal energy;
- $\boldsymbol{q} : (0,T) \times \overline{\Omega} \to \mathbb{R}^d$ is the heat flux.

The constitutive law relates the Cauchy stress (or some other appropriate measure of the stress) and the heat flux to other kinematical variables such as the shear strain, temperature, etc. In the following we will assume that the material is incompressible, homogeneous and undergoes an isothermal process. This implies that the energy equation decouples from the system and that the Cauchy stress can be split in two components:

$$\boldsymbol{T} = -p\boldsymbol{I} + \boldsymbol{S}, \tag{1.2}$$

where $\boldsymbol{I}$ is the identity matrix, $p\colon (0,T)\times\Omega \to \mathbb{R}$ is the pressure (mean normal stress), and $\boldsymbol{S}\colon (0,T)\times\Omega \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$ is the shear stress (hereafter referred only as "stress"). In this work we will consider constitutive relations of the form

$$\boldsymbol{G}(\cdot, \boldsymbol{S}, \boldsymbol{D}(\boldsymbol{u})) = \boldsymbol{0}, \tag{1.3}$$

where $\boldsymbol{G}\colon Q \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$ and $\boldsymbol{D}(\boldsymbol{u}) := \frac{1}{2}(\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^{\mathrm{T}})$ is the symmetric velocity gradient; here Q is used to denote the parabolic cylinder $(0,T)\times\Omega$. The precise assumptions on this implicit function will be stated in the next section.

For a rigorous mathematical analysis of models of implicitly constituted fluids the reader is referred to [13, 14]. Existence of weak solutions for problems of this type was obtained in [13] and [14] for the steady and unsteady cases, respectively. Some extensions include [15, 46, 50], where additional physical responses are incorporated into the system.

As for the numerical analysis of these systems, very few results have been published so far. In [21] the convergence of a finite element discretisation to a weak solution of the problem was proved for the steady case, and the corresponding a-posteriori analysis was carried out in [43]. More recently, this approach was extended to the time-dependent case in [61]. Also, several finite element discretisations were compared computationally in [41] for problems with Bingham and stress-power-law-like rheology.

Numerical methods for the incompressible Navier–Stokes equations are usually based on a velocity-pressure formulation, and extensive studies have been carried out over the years in relation to this (see e.g. [33, 10]). Such a formulation is possible, because in the case of a Newtonian fluid the explicit constitutive relation $\boldsymbol{S} = 2\mu\boldsymbol{D}(\boldsymbol{u})$ allows one to eliminate the deviatoric stress $\boldsymbol{S}$ from the momentum equation. In

This manuscript is for review purposes only.

contrast, formulations that treat the stress as a fundamental unknown have also been introduced to study problems in elasticity and incompressible flows [1, 4, 27, 28, 2, 26, 29, 30, 39, 40]; the key advantages of these formulations are that they are naturally applicable to nonlinear constitutive models where it is not possible to eliminate the stress, and that they allow the direct computation of the stress without resorting to numerical differentiation. In this work we will consider the mathematical analysis of a mixed formulation that treats the stress as an unknown, and illustrate its performance by means of numerical simulations.

The results here could be considered an extension of the works [21, 61, 41]. One of the advantages of the approach presented here with respect to [21, 61] is that it can handle the constitutive relation in a more natural way, since the stress plays a more prominent role in the weak formulation considered. In addition, in [21, 61] no numerical simulations were presented. On the other hand, while extensive numerical computations with 3-field and 4-field formulations were performed in [41], no convergence analysis of the methods considered was discussed. The work presented here fills this gap.

2. Preliminaries.

2.1. Function spaces.

Throughout this work we will assume that $\Omega \subset \mathbb{R}^d$, with $d \in \{2,3\}$, is a bounded Lipschitz polygonal domain (unless otherwise stated), and use standard notation for Lebesgue, Sobolev and Bochner–Sobolev spaces (e.g. $(W^{k,r}(\Omega), \|\cdot\|_{W^{k,r}(\Omega)})$ and $(L^q(0,T;W^{n,r}(\Omega)), \|\cdot\|_{L^q(0,T;W^{n,r}(\Omega))})$). We will define $W^{k,r}_0(\Omega)$ for $r \in [1,\infty)$ as the closure of the space of smooth functions with compact support $C_0^\infty(\Omega)$ with respect to the norm $\|\cdot\|_{W^{k,r}(\Omega)}$ and we will denote the dual space of $W^{1,r}_0(\Omega)$ by $W^{-1,r'}(\Omega)$. Here r' is used to denote the Hölder conjugate of r, i.e. the number defined by the relation $1/r + 1/r' = 1$. The duality pairing will be written in the usual way using brackets $\langle\cdot,\cdot\rangle$. The space of traces on the boundary of functions in $W^{1,r}(\Omega)$ will be denoted by $W^{1/r',r}(\partial\Omega)$.

If X is a Banach space, $C_w([0,T];X)$ will be used to denote the space of continuous functions in time with respect to the weak topology of X. For $r \in [1,\infty)$ we also define the following useful subspaces:

$$
\begin{aligned}
L^r_0(\Omega) &:= \left\{ q \in L^r(\Omega) \,:\, \int_\Omega q = 0 \right\},\\
L^2_{\mathrm{div}}(\Omega)^d &:= \overline{\{\boldsymbol{v} \in C_0^\infty(\Omega)^d \,:\, \operatorname{div} \boldsymbol{v} = 0\}}^{\|\cdot\|_{L^2(\Omega)}},\\
W^{1,r}_{0,\mathrm{div}}(\Omega)^d &:= \overline{\{\boldsymbol{v} \in C_0^\infty(\Omega)^d \,:\, \operatorname{div} \boldsymbol{v} = 0\}}^{\|\cdot\|_{W^{1,r}(\Omega)}},\\
L^r_{\mathrm{tr}}(Q)^{d\times d} &:= \{\boldsymbol{\tau} \in L^r(Q)^{d\times d} \,:\, \mathrm{tr}(\boldsymbol{\tau}) = 0\},\\
L^r_{\mathrm{sym}}(Q)^{d\times d} &:= \{\boldsymbol{\tau} \in L^r(Q)^{d\times d} \,:\, \boldsymbol{\tau}^{\mathrm{T}} = \boldsymbol{\tau}\}.
\end{aligned}
$$

In the definition of the space $L^r_{\mathrm{tr}}(Q)^{d\times d}$ above, $\mathrm{tr}(\boldsymbol{\tau})$ denotes the usual matrix trace of the $d\times d$ matrix function $\boldsymbol{\tau}$. In the various estimates the letter c will denote a generic positive constant whose exact value could change from line to line, whenever the explicit dependence on the parameters is not important.

2.2. Interpolation inequalities.

The following embeddings will be useful when deriving various estimates. Assume that the Banach spaces (W_1, W_2, W_3) form an interpolation triple in the sense that

$$
\|v\|_{W_2} \leq c\|v\|^{\lambda}_{W_1}\|v\|^{1-\lambda}_{W_3}, \quad \text{for some } \lambda \in (0,1),
$$

and $W_1 \hookrightarrow W_2 \hookrightarrow W_3$. Then (cf. [56]) $L^r(0,T;W_1)\cap L^\infty(0,T;W_3) \hookrightarrow L^{r/\lambda}(0,T;W_2)$, for $r\in[1,\infty)$ and

$$\|v\|_{L^{r/\lambda}(0,T;W_2)} \leq c\|v\|^{1-\lambda}_{L^\infty(0,T;W_3)}\|v\|^{\lambda}_{L^r(0,T;W_1)}. \tag{2.1}$$

An example of an interpolation triple that can be combined with this result is given by the Gagliardo–Nirenberg inequality, which states that for given $p,r\in[1,\infty)$, there is a constant $c_{p,r}>0$ such that [20]:

$$\|v\|_{L^s(\Omega)} \leq c_{p,r}\|\nabla v\|^{\lambda}_{L^r(\Omega)}\|v\|^{1-\lambda}_{L^p(\Omega)} \qquad \forall\, v\in W^{1,r}_0(\Omega)\cap L^p(\Omega), \tag{2.2}$$

provided that $s\in[1,\infty)$ and $\lambda\in(0,1)$ satisfy

$$\lambda = \frac{\frac{1}{p}-\frac{1}{s}}{\frac{1}{d}-\frac{1}{r}+\frac{1}{p}}.$$

A particularly useful example can be obtained if we assume that $r>\frac{2d}{d+2}$ and take $p=2$ and $\lambda=\frac{d}{d+2}$:

$$\|v\|_{L^{\frac{r(d+2)}{d}}(Q)} \leq c\|\nabla v\|^{\lambda}_{L^r(Q)}\|v\|^{1-\lambda}_{L^\infty(0,T;L^2(\Omega))} \qquad \forall\, v\in L^r(0,T;W^{1,r}_0(\Omega))\cap L^\infty(0,T;L^2(\Omega)). \tag{2.3}$$

2.3. Compactness and continuity in time. In this work we will use Simon's compactness lemma (see [60]) instead of the usual Aubin–Lions lemma to extract convergent subsequences when taking the discretisation limit in the time–dependent problem. Assume that X and H are Banach spaces such that the compact embedding $X \hookrightarrow\hookrightarrow H$ holds. Simon's lemma states that if $\mathcal{U}\subset L^p(0,T;H)$, for some $p\in[1,\infty)$, and it satisfies:

- $\mathcal{U}$ is bounded in $L^1_{\mathrm{loc}}(0,T;X)$;
- $\int_0^{T-\epsilon}\|v(t+\epsilon,\cdot)-v(t,\cdot)\|^p_H \to 0$, as $\epsilon\to 0$, uniformly for $v\in\mathcal{U}$;

then $\mathcal{U}$ is relatively compact in $L^p(0,T;H)$.

Let X and V be reflexive Banach spaces such that $X\hookrightarrow V$ densely and let V^* be the dual space of V. The following continuity properties (see [56]) will be important when identifying the initial condition:

$$v\in L^1(0,T;V^*),\ \partial_t v\in L^1(0,T;V^*) \Longrightarrow v\in C([0,T];V^*), \tag{2.4}$$

$$v\in L^\infty(0,T;X)\cap C_w([0,T];V) \Longrightarrow v\in C_w([0,T];X). \tag{2.5}$$

2.4. Implicit constitutive relation and its approximation. In the mathematical analysis of these systems it is more convenient to work not with the function $\boldsymbol{G}$, but with its graph $\mathcal{A}$, which is introduced in the usual way:

$$(\boldsymbol{D},\boldsymbol{S})\in\mathcal{A}(\cdot) \Longleftrightarrow \boldsymbol{G}(\cdot,\boldsymbol{S},\boldsymbol{D})=\mathbf{0}. \tag{2.6}$$

We will assume that $\mathcal{A}$ is a *maximal monotone r-graph* for some $r>1$, which means that the following properties hold for almost every $z\in Q$:

(A1) [*$\mathcal{A}$ includes the origin*] $(\mathbf{0},\mathbf{0})\in\mathcal{A}(z)$.

(A2) [*$\mathcal{A}$ is a monotone graph*] For every $(\mathbf{D}_1,\mathbf{S}_1),(\mathbf{D}_2,\mathbf{S}_2)\in\mathcal{A}(z)$,

$$(\mathbf{S}_1-\mathbf{S}_2):(\mathbf{D}_1-\mathbf{D}_2)\geq 0.$$

(A3) [$\mathcal{A}$ *is maximal monotone*] If $(\mathbf{D},\mathbf{S}) \in \mathbb{R}^{d\times d}_{\text{sym}} \times \mathbb{R}^{d\times d}_{\text{sym}}$ is such that

$$(\hat{\mathbf{S}} - \mathbf{S}) : (\hat{\mathbf{D}} - \mathbf{D}) \geq 0 \quad \text{for all } (\hat{\mathbf{D}}, \hat{\mathbf{S}}) \in \mathcal{A}(z),$$

then $(\mathbf{D},\mathbf{S}) \in \mathcal{A}(z)$.

(A4) [$\mathcal{A}$ *is an r-graph*] There is a non-negative function $m \in L^1(Q)$ and a constant $c > 0$ such that

$$\mathbf{S} : \mathbf{D} \geq -m + c(|\mathbf{D}|^r + |\mathbf{S}|^{r'}) \quad \text{for all } (\boldsymbol{D}, \boldsymbol{S}) \in \mathcal{A}(z).$$

(A5) [*Measurability*] The set-valued map $z \mapsto \mathcal{A}(z)$ is $\mathcal{L}(Q)$–$(\mathcal{B}(\mathbb{R}^{d\times d}_{\text{sym}} \otimes \mathbb{R}^{d\times d}_{\text{sym}}))$ measurable; here $\mathcal{L}(Q)$ denotes the family of Lebesgue measurable subsets of Q and $\mathcal{B}(\mathbb{R}^{d\times d}_{\text{sym}})$ is the family of Borel subsets of $\mathbb{R}^{d\times d}_{\text{sym}}$.

(A6) [*Compatibility*] For any $(\boldsymbol{D}, \boldsymbol{S}) \in \mathcal{A}(z)$ we have that

$$\text{tr}(\boldsymbol{D}) = 0 \Longleftrightarrow \text{tr}(\boldsymbol{S}) = 0.$$

Assumption (A6) was not included in the original works [13, 14, 21], but it is needed for consistency with the physical property that $\boldsymbol{S}$ is traceless if and only if the velocity field is divergence-free (see the discussion in [62]). A very important consequence of Assumption (A5) (see [62]) is the existence of a measurable function (usually called a *selection*) $\boldsymbol{\mathcal{D}} : Q \times \mathbb{R}^{d\times d}_{sym} \to \mathbb{R}^{d\times d}_{sym}$ such that $(\boldsymbol{\mathcal{D}}(z,\boldsymbol{\sigma}),\boldsymbol{\sigma}) \in \mathcal{A}(z)$ for all $\boldsymbol{\sigma} \in \mathbb{R}^{d\times d}_{sym}$.

In the existence results it will be useful to approximate the selection using smooth functions. To that end, let us define the mollification:

$$\boldsymbol{\mathcal{D}}^k(\cdot,\boldsymbol{\sigma}) := \int_{\mathbb{R}^{d\times d}_{\text{sym}}} \boldsymbol{\mathcal{D}}(\cdot,\boldsymbol{\sigma} - \boldsymbol{\tau})\rho^k(\boldsymbol{\tau})\,\mathrm{d}\boldsymbol{\tau}, \tag{2.7}$$

where $\rho^k(\boldsymbol{\tau}) = k^{d^2}\rho(k\boldsymbol{\tau})$, $k \in \mathbb{N}$, and $\rho \in C^\infty_0(\mathbb{R}^{d\times d}_{\text{sym}})$ is a mollification kernel. It is possible to check (see e.g. [62]) that this mollification satisfies analogous monotonicity and coercivity properties to those of the selection $\boldsymbol{\mathcal{D}}$, i.e. we have that

- For every $\boldsymbol{\tau}_1, \boldsymbol{\tau}_2 \in \mathbb{R}^{d\times d}_{\text{sym}}$ and for almost every $z \in Q$ the monotonicity condition

$$(\boldsymbol{\mathcal{D}}^k(z,\boldsymbol{\tau}_1) - \boldsymbol{\mathcal{D}}^k(z,\boldsymbol{\tau}_2)) : (\boldsymbol{\tau}_1 - \boldsymbol{\tau}_2) \geq 0 \tag{2.8}$$

  holds.
- There is a constant $C_* > 0$ and a nonnegative function $g \in L^1(Q)$ such that for all $k \in \mathbb{N}$, for every $\boldsymbol{\tau} \in \mathbb{R}^{d\times d}$, and for almost every $z \in Q$ we have

$$\boldsymbol{\tau} : \boldsymbol{\mathcal{D}}^k(z,\boldsymbol{\tau}) \geq -g(z) + C_*(|\boldsymbol{\tau}|^{r'} + |\boldsymbol{\mathcal{D}}^k(z,\boldsymbol{\tau})|^r). \tag{2.9}$$

- For any sequence $\{\boldsymbol{S}_k\}_{k\in\mathbb{N}}$ bounded in $L^{r'}(Q)^{d\times d}$, we have for arbitrary $\boldsymbol{B} \in \mathbb{R}^{d\times d}_{\text{sym}}$ and $\phi \in C^\infty_0(Q)$ with $\phi \geq 0$:

$$\liminf_{k\to\infty} \int_Q (\boldsymbol{\mathcal{D}}^k(\cdot,\boldsymbol{S}^k) - \boldsymbol{\mathcal{D}}(\cdot,\boldsymbol{B})) : (\boldsymbol{S}^k - \boldsymbol{B})\phi(\cdot) \geq 0. \tag{2.10}$$

It is important to remark that (2.8), (2.9) and (2.10) are the essential properties; the explicit form (2.7) of the approximation to the selection is not very important. There are other ways to achieve the same result; for instance piecewise affine interpolation or a generalised Yosida approximation could also be used (see [61, 62]). The following is a localized version of Minty's lemma that will aid in the identification of the implicit constitutive relation (for a proof see [12]).

LEMMA 2.1. *Let $\mathcal{A}$ be a maximal monotone r-graph satisfying (A1)–(A4) for some $r > 1$. Suppose that $\{\boldsymbol{D}^n\}_{n\in\mathbb{N}}$ and $\{\boldsymbol{S}^n\}_{n\in\mathbb{N}}$ are sequences of functions defined on a measurable set $\hat{Q} \subset Q$, such that:*

$$
\begin{aligned}
&(\boldsymbol{D}^n(\cdot), \boldsymbol{S}^n(\cdot)) \in \mathcal{A}(\cdot) && \text{a.e. in } \hat{Q},\\
&\boldsymbol{D}^n \rightharpoonup \boldsymbol{D}, && \text{weakly in } L^r(\hat{Q})^{d\times d},\\
&\boldsymbol{S}^n \rightharpoonup \boldsymbol{S}, && \text{weakly in } L^{r'}(\hat{Q})^{d\times d},\\
&\limsup_{n\to\infty} \int_{\hat{Q}} \boldsymbol{S}^n : \boldsymbol{D}^n \leq \int_{\hat{Q}} \boldsymbol{S} : \boldsymbol{D}.
\end{aligned}
$$

Then,

$$(\boldsymbol{D}(\cdot), \boldsymbol{S}(\cdot)) \in \mathcal{A}(\cdot) \qquad \text{a.e. in } \hat{Q}.$$

The goal of this work is to prove convergence of a three-field finite element approximation of the following system:

$$
\begin{aligned}
\partial_t \boldsymbol{u} - \operatorname{div}(\boldsymbol{S} - \boldsymbol{u}\otimes\boldsymbol{u}) + \nabla p &= \boldsymbol{f} && \text{in } (0,T)\times\Omega,\\
\operatorname{div} \boldsymbol{u} &= 0 && \text{in } (0,T)\times\Omega,\\
(\boldsymbol{D}(\boldsymbol{u}), \boldsymbol{S}) &\in \mathcal{A}(\cdot) && \text{a.e. in } (0,T)\times\Omega,\\
\boldsymbol{u} &= \boldsymbol{0} && \text{on } (0,T)\times\partial\Omega,\\
\boldsymbol{u}(0,\cdot) &= \boldsymbol{u}_0(\cdot) && \text{in } \Omega,
\end{aligned}
\tag{2.11}
$$

where $\mathcal{A}(\cdot)$ satisfies (A1)–(A6). The next section introduces the notation and tools that will be useful in the analysis of the discrete problem.

2.5. Finite element approximation. In this section, the notation and assumptions regarding the finite element approximation will be presented. Essentially the same arguments would work for any method based on a Galerkin approximation, but here we will focus only on finite element methods. Consider a family of triangulations $\{\mathcal{T}_n\}_{n\in\mathbb{N}}$ of Ω satisfying the following assumptions:

- (Affine equivalence). Given $n \in \mathbb{N}$ and an element $K \in \mathcal{T}_n$, there is an affine invertible mapping $\boldsymbol{F}_K : K \to \hat{K}$, where $\hat{K}$ is the closed standard reference simplex in $\mathbb{R}^d$.
- (Shape-regularity). There is a constant c_τ, independent of n, such that

$$h_K \leq c_\tau \rho_K \qquad \text{for every } K \in \mathcal{T}_n,\ n \in \mathbb{N},$$

  where $h_K := \operatorname{diam}(K)$ and ρ_K is the diameter of the largest inscribed ball.
- The mesh size $h_n := \max_{K\in\mathcal{T}_n} h_K$ tends to zero as $n \to \infty$.

Define the conforming finite element spaces associated with the triangulation $\mathcal{T}_n$:

$$
\begin{aligned}
V^n &:= \left\{ \boldsymbol{v} \in W^{1,\infty}_0(\Omega)^d \,:\, \boldsymbol{v}|_K \circ \boldsymbol{F}_K^{-1} \in \hat{\mathbb{P}}_{\mathbb{V}},\ K \in \mathcal{T}_n,\ \boldsymbol{v}|_{\partial\Omega} = 0 \right\},\\
M^n &:= \left\{ \boldsymbol{q} \in L^\infty(\Omega) \,:\, \boldsymbol{q}|_K \circ \boldsymbol{F}_K^{-1} \in \hat{\mathbb{P}}_{\mathbb{M}},\ K \in \mathcal{T}_n \right\},\\
\Sigma^n &:= \left\{ \boldsymbol{\sigma} \in L^\infty(\Omega)^{d\times d} \,:\, \boldsymbol{\sigma}|_K \circ \boldsymbol{F}_K^{-1} \in \hat{\mathbb{P}}_{\mathbb{S}},\ K \in \mathcal{T}_n \right\},
\end{aligned}
$$

where $\hat{\mathbb{P}}_{\mathbb{V}} \subset W^{1,\infty}(\hat{K})^d$, $\hat{\mathbb{P}}_{\mathbb{M}} \subset L^\infty(\hat{K})$ and $\hat{\mathbb{P}}_{\mathbb{S}} \subset L^\infty(\hat{K})^{d\times d}$ are finite-dimensional polynomial subspaces on the reference simplex $\hat{K}$. Each of these spaces will be assumed to have a finite and locally supported basis. As in the continuous case, it will be useful to introduce the following finite-dimensional subspaces for $r > 1$:

$$M_0^n := M^n \cap L_0^{r'}(\Omega), \quad \Sigma^n_{\mathrm{tr}} := \Sigma^n \cap L^r_{\mathrm{tr}}(\Omega)^{d\times d}, \quad \Sigma^n_{\mathrm{sym}} := \Sigma^n \cap L^r_{\mathrm{sym}}(\Omega)^{d\times d},$$

$$V^n_{\mathrm{div}} := \left\{ \boldsymbol{v} \in V^n \,:\, \int_\Omega q \,\mathrm{div}\boldsymbol{v} = 0, \quad \forall\, q \in M^n \right\},$$

$$\Sigma^n_{\mathrm{div}}(\boldsymbol{f}) := \left\{ \boldsymbol{\sigma} \in \Sigma^n_{\mathrm{sym}} \,:\, \int_\Omega \boldsymbol{\sigma} : \boldsymbol{D}(\boldsymbol{v}) = \langle \boldsymbol{f}, \boldsymbol{v}\rangle, \quad \forall\, \boldsymbol{v} \in V^n_{\mathrm{div}} \right\}.$$

Assumption 2.2 (Approximability). *For every $s \in [1,\infty)$ we have that*

$$\inf_{\overline{\boldsymbol{v}} \in V^n} \|\boldsymbol{v} - \overline{\boldsymbol{v}}\|_{W^{1,s}(\Omega)} \to 0 \quad \text{as } n\to\infty \quad \forall\, \boldsymbol{v} \in W^{1,s}_0(\Omega)^d,$$

$$\inf_{\overline{q} \in M^n} \|q - \overline{q}\|_{L^{s}(\Omega)} \to 0 \quad \text{as } n\to\infty \quad \forall\, q \in L^s(\Omega),$$

$$\inf_{\overline{\boldsymbol{\sigma}} \in \Sigma^n} \|\boldsymbol{\sigma} - \overline{\boldsymbol{\sigma}}\|_{L^{s}(\Omega)} \to 0 \quad \text{as } n\to\infty \quad \forall\, \boldsymbol{\sigma} \in L^s(\Omega)^{d\times d}.$$

Assumption 2.3 (Projector Π^n_Σ). *For each $n \in \mathbb{N}$ there is a linear projector $\Pi^n_\Sigma : L^1_{\mathrm{sym}}(\Omega)^{d\times d} \to \Sigma^n_{\mathrm{sym}}$ such that:*

- *(Preservation of divergence). For every $\boldsymbol{\sigma} \in L^1(\Omega)^{d\times d}$ we have*

$$\int_\Omega \boldsymbol{\sigma} : \boldsymbol{D}(\boldsymbol{v}) = \int_\Omega \Pi^n_\Sigma(\boldsymbol{\sigma}) : \boldsymbol{D}(\boldsymbol{v}) \quad \forall\, \boldsymbol{v} \in V^n_{\mathrm{div}}.$$

- *(L^s–stability). For every $s \in (1,\infty)$ there is a constant $c > 0$, independent of n, such that:*

$$\|\Pi^n_\Sigma \boldsymbol{\sigma}\|_{L^s(\Omega)} \le c\|\boldsymbol{\sigma}\|_{L^s(\Omega)} \qquad \forall\, \boldsymbol{\sigma} \in L^s_{\mathrm{sym}}(\Omega)^{d\times d}.$$

Assumption 2.4 (Projector Π^n_V). *For each $n \in \mathbb{N}$ there is a linear projector $\Pi^n_V : W^{1,1}_0(\Omega)^d \to V^n$ such that the following properties hold:*

- *(Preservation of divergence). For every $\boldsymbol{v} \in W^{1,1}_0(\Omega)^d$ we have*

$$\int_\Omega q \,\mathrm{div}\, \boldsymbol{v} = \int_\Omega q \,\mathrm{div}(\Pi^n_V \boldsymbol{v}) \quad \forall\, q \in M^n.$$

- *($W^{1,s}$–stability). For every $s \in (1,\infty)$ there is a constant $c > 0$, independent of n, such that:*

$$\|\Pi^n_V \boldsymbol{v}\|_{W^{1,s}(\Omega)} \le c\|\boldsymbol{v}\|_{W^{1,s}(\Omega)} \qquad \forall\, \boldsymbol{v} \in W^{1,s}_0(\Omega)^d.$$

Assumption 2.5 (Projector Π^n_M). *For each $n \in \mathbb{N}$ there is a linear projector $\Pi^n_M : L^1(\Omega) \to M^n$ such that for all $s \in (1,\infty)$ there is a constant $c > 0$, independent of n, such that:*

$$\|\Pi^n_M q\|_{L^s(\Omega)} \le c\|q\|_{L^s(\Omega)} \qquad \forall\, q \in L^s(\Omega).$$

It is not difficult to show that the approximability and stability properties imply that for $s \in [1,\infty)$ we have:

$$\begin{aligned} &\|\boldsymbol{\sigma} - \Pi^n_\Sigma \boldsymbol{\sigma}\|_{L^s(\Omega)} \to 0 \quad \text{as } n\to\infty \quad \forall\, \boldsymbol{\sigma} \in L^s_{\mathrm{sym}}(\Omega)^{d\times d},\\ &\|\boldsymbol{v} - \Pi^n_V \boldsymbol{v}\|_{W^{1,s}(\Omega)} \to 0 \quad \text{as } n\to\infty \quad \forall\, \boldsymbol{v} \in W^{1,s}(\Omega)^d,\\ &\|q - \Pi^n_M q\|_{L^s(\Omega)} \to 0 \quad \text{as } n\to\infty \quad \forall\, q \in L^s(\Omega). \end{aligned} \tag{2.12}$$

Remark 2.6. A very important consequence of the previous assumptions is the existence, for every $s \in (1,\infty)$, of two positive constants $\beta_s, \gamma_s > 0$, independent of n, such that the following discrete inf-sup conditions hold:

$$\inf_{q \in M_0^n} \sup_{\boldsymbol{v} \in V^n} \frac{\int_\Omega q \operatorname{div} \boldsymbol{v}}{\|\boldsymbol{v}\|_{W^{1,s}(\Omega)} \|q\|_{L^{s'}(\Omega)}} \geq \beta_s, \tag{2.13}$$

$$\inf_{\boldsymbol{v} \in V_{\mathrm{div}}^n} \sup_{\boldsymbol{\tau} \in \Sigma_{\mathrm{sym}}^n} \frac{\int_\Omega \boldsymbol{\tau} : \boldsymbol{D}(\boldsymbol{v})}{\|\boldsymbol{\tau}\|_{L^{s'}(\Omega)} \|\boldsymbol{v}\|_{W^{1,s}(\Omega)}} \geq \gamma_s. \tag{2.14}$$

Example 2.7. There are several pairs of velocity-pressure spaces known to satisfy the stability Assumptions 2.2 and 2.4. They include the conforming Crouzeix–Raviart element, the MINI element, the $\mathbb{P}_2$–$\mathbb{P}_0$ element and the Taylor–Hood element $\mathbb{P}_k$–$\mathbb{P}_{k-1}$ for $k \geq d$ (see [5, 8, 21, 34, 18]). In addition to stability, the Scott–Vogelius element also satisfies the property that the discretely divergence-free velocities are pointwise divergence-free (the stability can be guaranteed by assuming for example that the mesh has been barycentrically refined, see [59]); another example of a velocity-pressure pair with this property is given by the Guzmán–Neilan element [37, 36]. To satisfy Assumption 2.5, one could use the Clément interpolant [17].

Sometimes it is easier to prove the inf-sup condition directly. For example, if the space of discrete stresses consists of discontinuous $\mathbb{P}_k$ polynomials (with $k \geq 1$):

$$\Sigma^n = \{\boldsymbol{\sigma} \in L^\infty(\Omega)^{d\times d} : \boldsymbol{\sigma}|_K \in \mathbb{P}_k(K)^{d\times d}, \text{ for all } K \in \mathcal{T}_n\},$$

and we have that $\boldsymbol{D}(V^n) \subset \Sigma^n$ (e.g. we could take the Taylor–Hood element $\mathbb{P}_{k+1}$–$\mathbb{P}_k$ for the velocity and the pressure), then the inf-sup condition follows from the fact that for $s \in (1,\infty)$ there is a constant $c > 0$, independent of h, such that for any $\boldsymbol{\sigma} \in \Sigma^n$ there is $\boldsymbol{\tau} \in \Sigma^n$ such that [58]:

$$\int_\Omega \boldsymbol{\tau} : \boldsymbol{\sigma} = \|\boldsymbol{\sigma}\|_{L^s(\Omega)}^s \quad \text{and} \quad \|\boldsymbol{\tau}\|_{L^{s'}(\Omega)} \leq c\|\boldsymbol{\sigma}\|_{L^s(\Omega)}^{s-1}.$$

In case a continuous piecewise polynomial approximation of the stress is preferred, one could use the conforming Crouzeix–Raviart element for the discrete velocity and pressure and the following space for the stress [57] :

$$\Sigma^n = \{\boldsymbol{\sigma} \in C(\overline{\Omega})^{d\times d} : \boldsymbol{\sigma}|_K \in (\mathbb{P}_1(K) \oplus \mathcal{B})^{d\times d}, \text{ for all } K \in \mathcal{T}_n\},$$

where

$$\mathcal{B} := \operatorname{span}\{\lambda_1^2\lambda_2\lambda_3, \lambda_1\lambda_2^2\lambda_3, \lambda_1\lambda_2\lambda_3^2\},$$

and $\{\lambda_j\}_{j=1}^3$ are barycentric coordinates on K.

Remark 2.8. If the discretely divergence-free velocities are in fact exactly divergence free, i.e. if $V_{\mathrm{div}}^n \subset W_{0,\mathrm{div}}^{1,r}(\Omega)^d$, and $\boldsymbol{D}(V^n) \subset \Sigma^n$, then the stress-velocity inf-sup condition also holds for the subspace of traceless stresses. Consequently, fewer degrees of freedom are needed to compute the stress unknowns.

2.6. Time discretisation. In this section we will describe the notation that will be used when performing the time discretisation of the problem. Let $\{\tau_m\}_{m\in\mathbb{N}}$ be a sequence of time steps such that $T/\tau_m \in \mathbb{N}$ and $\tau_m \to 0$, as $m \to \infty$. For each $m \in \mathbb{N}$ we define the equidistant grid:

$$\{t_j^m\}_{j=0}^{T/\tau_m}, \qquad t_j = t_j^m := j\tau_m.$$

This can be used to define the parabolic cylinders $Q_i^j := (t_i, t_j) \times \Omega$, where $0 \le i \le j \le T/\tau_m$. Also, given a set of functions $\{v^j\}_{j=0}^{T/\tau_m}$ belonging to a Banach space X, we can define the piecewise constant interpolant $\overline{v} \in L^\infty(0,T;X)$ as:

$$\overline{v}(t) := v^j, \qquad t \in (t_{j-1}, t_j], \qquad j \in \{1, \dots, T/\tau_m\}, \tag{2.15}$$

and the piecewise linear interpolant $\tilde{v} \in C([0,T];X)$ as:

$$\tilde{v}(t) := \frac{t - t_{j-1}}{\tau_m} v^j + \frac{t_j - t}{\tau_m} v^{j-1}, \qquad t \in [t_{j-1}, t_j], \quad j \in \{1, \dots, T/\tau_m\}. \tag{2.16}$$

For a given function $g \in L^p(0,T;X)$, with $p \in [1,\infty)$, we define the time averages:

$$g_j(\cdot) := \frac{1}{\tau_m} \int_{t_{j-1}}^{t_j} g(t, \cdot) \,\mathrm{d}\, t, \quad j \in \{1, \dots, T/\tau_m\}. \tag{2.17}$$

Then the piecewise constant interpolant $\overline{g}$ defined by (2.15) satisfies [56]:

$$\|\overline{g}\|_{L^p(0,T;X)} \le \|g\|_{L^p(0,T;X)}, \tag{2.18}$$

and

$$\overline{g} \to g \text{ strongly in } L^p(0,T;X), \text{ as } m \to \infty. \tag{2.19}$$

3. Weak formulation. In this section we will present a weak formulation for the problem (2.11), where now we assume that $\boldsymbol{f} \in L^{r'}(0,T;W^{-1,r'}(\Omega)^d)$, $\boldsymbol{u}_0 \in L^2_{\mathrm{div}}(\Omega)^d$ and the graph $\mathcal{A}$ satisfies the assumptions (A1)–(A6) for some $r > \frac{2d}{d+2}$. Similarly to previous works on the analysis of implicitly constituted fluids, a Lipschitz truncation technique will be required when proving that the limit of the sequence of approximate solutions satisfies the constitutive relation. The theory of Lipschitz truncation for time-dependent problems is not as well developed as in the steady case; here it will be necessary to work locally and the equation plays a vital role (several versions of parabolic Lipschitz truncation have appeared in the literature, see e.g. [22, 14, 9, 23]). Since the pressure will not be present in the weak formulation, it will be more convenient to use the construction developed in [9] because it preserves the solenoidality of the velocity. The following lemma states the main properties of this solenoidal Lipschitz truncation.

LEMMA 3.1. ([9, 61]) *Let $p \in (1,\infty)$, $\sigma \in (1, \min(p,p'))$ and let $Q_0 = I_0 \times B_0 \subset \mathbb{R} \times \mathbb{R}^3$ be a parabolic cylinder, where I_0 is an open interval and B_0 is an open ball. Denote by αQ_0, where $\alpha > 0$, the α-scaled version of Q_0 keeping the barycenter the same. Suppose $\{\boldsymbol{e}^l\}_{l\in\mathbb{N}}$ is a sequence of divergence-free functions that is uniformly bounded in $L^\infty(I_0;L^\sigma(B_0)^d)$ and converges to zero weakly in $L^p(I_0;W^{1,p}(B_0)^d)$ and strongly in $L^\sigma(Q_0)^d$. Let $\{\boldsymbol{G}_1^l\}_{l\in\mathbb{N}}$ and $\{\boldsymbol{G}_2^l\}_{l\in\mathbb{N}}$ be sequences that converge to zero weakly in $L^{p'}(Q_0)^{d\times d}$ and strongly in $L^\sigma(Q_0)^{d\times d}$, respectively. Define $\boldsymbol{G}^l := \boldsymbol{G}_1^l + \boldsymbol{G}_2^l$ and suppose that, for any $l \in \mathbb{N}$, the equation*

$$\int_{Q_0} \partial_t \boldsymbol{e}^l \cdot \boldsymbol{w} = \int_{Q_0} \boldsymbol{G}^l : \nabla \boldsymbol{w} \quad \forall\, \boldsymbol{w} \in C^\infty_{0,\mathrm{div}}(Q_0)^d. \tag{3.1}$$

is satisfied. Then there is a number $j_0 \in \mathbb{N}$, a sequence $\{\lambda_{l,j}\}_{l,j\in\mathbb{N}}$ with $2^{2^j} \le \lambda_{l,j} \le 2^{2^{j+1}-1}$, a sequence of functions $\{\boldsymbol{e}^{l,j}\}_{l,j\in\mathbb{N}} \subset L^1(Q_0)^d$, a sequence of open sets $\mathcal{B}_{\lambda_{l,j}} \subset Q_0$, for $l,j \in \mathbb{N}$, and a function $\zeta \in C_0^\infty(\frac{1}{6}Q_0)$ with $\mathbb{1}_{\frac{1}{8}Q_0} \le \zeta \le \mathbb{1}_{\frac{1}{6}Q_0}$ with the following properties:

1. $\boldsymbol{e}^{l,j} \in L^q(\frac{1}{4}I_0; W^{1,q}_{0,\mathrm{div}}(\frac{1}{6}B_0)^d)$ *for any* $q \in [1,\infty)$ *and* $\mathrm{supp}(\boldsymbol{e}^{l,j}) \subset \frac{1}{6}Q_0$, *for any* $j \geq j_0$ *and any* $l \in \mathbb{N}$;
2. $\boldsymbol{e}^{l,j} = \boldsymbol{e}^j$ *on* $\frac{1}{8}Q_0 \setminus \mathcal{B}_{\lambda_{l,j}}$, *for any* $j \geq j_0$ *and any* $l \in \mathbb{N}$;
3. *There is a constant* $c > 0$ *such that*

$$\limsup_{l\to\infty} \lambda^p_{l,j} |\mathcal{B}_{\lambda_{l,j}}| \leq c2^{-j}, \quad \text{for any } j \geq j_0;$$

4. *For* $j \geq j_0$ *fixed, we have as* $l \to \infty$:

$$\begin{aligned} \boldsymbol{e}^{l,j} &\to \boldsymbol{0}, && \text{strongly in } L^\infty(\tfrac{1}{4}Q_0)^d, \\ \nabla \boldsymbol{e}^{l,j} &\rightharpoonup \boldsymbol{0}, && \text{weakly in } L^q(\tfrac{1}{4}Q_0)^{d\times d}, \quad \forall\, q \in [1,\infty); \end{aligned}$$

5. *There is a constant* $c > 0$ *such that:*

$$\limsup_{l\to\infty} \left| \int_{Q_0} \boldsymbol{G}^l : \nabla \boldsymbol{e}^{l,j} \right| \leq c2^{-j}, \quad \text{for any } j \geq j_0;$$

6. *There is a constant* $c > 0$ *such that for any* $\boldsymbol{H} \in L^{p'}(\frac{1}{6}Q_0)^{d\times d}$:

$$\limsup_{l\to\infty} \left| \int_{Q_0} (\boldsymbol{G}^l_1 + \boldsymbol{H}) : \nabla \boldsymbol{e}^{l,j} \zeta \mathbb{1}_{\mathcal{B}^c_{\lambda_{l,j}}} \right| \leq c2^{-j/p}, \quad \text{for any } j \geq j_0.$$

3.1. Mixed formulation and time–space discretisation. Before we present the weak formulation, let us define

$$\check{r} := \min\left\{ \frac{r(d+2)}{2d}, r' \right\}.$$

The weak formulation for (2.11) then reads as follows.

Formulation Ǎ. Find functions

$$\begin{gathered} \boldsymbol{S} \in L^{r'}_{\mathrm{sym}}(Q)^{d\times d} \cap L^{r'}_{\mathrm{tr}}(Q)^{d\times d}, \\ \boldsymbol{u} \in L^r(0,T; W^{1,r}_{0,\mathrm{div}}(\Omega)^d) \cap L^\infty(0,T; L^2_{\mathrm{div}}(\Omega)^d), \\ \partial_t \boldsymbol{u} \in L^{\check{r}}(0,T; (W^{1,\check{r}'}_{0,\mathrm{div}}(\Omega)^d)^*), \end{gathered}$$

such that

$$\langle \partial_t \boldsymbol{u}, \boldsymbol{v}\rangle + \int_\Omega (\boldsymbol{S} - \boldsymbol{u}\otimes\boldsymbol{u}) : \boldsymbol{D}(\boldsymbol{v}) = \langle \boldsymbol{f}, \boldsymbol{v}\rangle \qquad \forall\, \boldsymbol{v} \in W^{1,\check{r}'}_{0,\mathrm{div}}(\Omega)^d, \text{ a.e. } t \in (0,T),$$

$$(\boldsymbol{D}(\boldsymbol{u}), \boldsymbol{S}) \in \mathcal{A}(\cdot), \text{ a.e. in } (0,T)\times\Omega,$$

$$\operatorname*{ess\,lim}_{t\to 0^+} \|\boldsymbol{u}(t,\cdot) - \boldsymbol{u}_0(\cdot)\|_{L^2(\Omega)} = 0.$$

Remark 3.2. In the formulation above all the test-velocities are divergence-free and as a consequence the presure term vanishes. In this section we will carry out the analysis for the velocity and stress variables only. It is known that even in the Newtonian case (i.e. $r = 2$) the pressure is only a distribution in time, when working with a no-slip boundary condition (see e.g. [31]). An integrable pressure can be obtained if Navier's slip boundary condition is used instead [14], but in this work we will confine ourselves to the more common no-slip boundary condition.

Remark 3.3. From (2.4) we have that

$$\boldsymbol{u} \in C([0,T]; (W^{1,\check{r}'}_{0,\mathrm{div}}(\Omega)^d)^*) \hookrightarrow C_w([0,T]; (W^{1,\check{r}'}_{0,\mathrm{div}}(\Omega)^d)^*),$$

and since $\check{r} \leq r'$ we also know that $L^2_{\mathrm{div}}(\Omega)^d \hookrightarrow (W^{1,\check{r}'}_{0,\mathrm{div}}(\Omega)^d)^*$. Combined with (2.5) this yields $\boldsymbol{u} \in C_w([0,T]; L^2_{\mathrm{div}}(\Omega)^d)$ and hence the initial condition only makes sense a priori in this weaker sense. However, for this problem it will be proved that it also holds in the stronger sense described above.

For a given time step τ_m and $j \in \{1, \dots, T/\tau_m\}$, let $\boldsymbol{f}_j \in W^{-1,r'}(\Omega)^d$ and $\boldsymbol{\mathcal{D}}^k_j : \Omega \times \mathbb{R}^{d\times d} \to \mathbb{R}^{d\times d}$ be the time averages associated with $\boldsymbol{f}$ and $\boldsymbol{\mathcal{D}}^k$, respectively (recall (2.17)). The time derivative will be discretised using an implicit Euler scheme; higher order time stepping techniques might not be more advantageous here because higher regularity in time of weak solutions to the problem is not guaranteed a priori. The discrete formulation of the problem can now be introduced.

Formulation $\check{\mathbf{A}}_{\mathbf{k,n,m,l}}$. For $j \in \{1, \dots, T/\tau_m\}$, find functions $\boldsymbol{S}^{k,n,m,l}_j \in \Sigma^n_{\mathrm{sym}}$ and $\boldsymbol{u}^{k,n,m,l}_j \in V^n_{\mathrm{div}}$ such that:

$$\begin{aligned}
&\int_\Omega (\boldsymbol{\mathcal{D}}^k_j(\cdot, \boldsymbol{S}^{k,n,m,l}_j) - \boldsymbol{D}(\boldsymbol{u}^{k,n,m,l}_j)) : \boldsymbol{\tau} = 0 && \forall\, \boldsymbol{\tau} \in \Sigma^n_{\mathrm{sym}},\\
&\frac{1}{\tau_m}\int_\Omega (\boldsymbol{u}^{k,n,m,l}_j - \boldsymbol{u}^{k,n,m,l}_{j-1})\cdot \boldsymbol{v} + \frac{1}{l}\int_\Omega |\boldsymbol{u}^{k,n,m,l}_j|^{2r'-2}\boldsymbol{u}^{k,n,m,l}_j \cdot \boldsymbol{v}\\
&\qquad + \int_\Omega (\boldsymbol{S}^{k,n,m,l}_j : \boldsymbol{D}(\boldsymbol{v}) + \mathcal{B}(\boldsymbol{u}^{k,n,m,l}_j, \boldsymbol{u}^{k,n,m,l}_j, \boldsymbol{v})) = \langle \boldsymbol{f}_j, \boldsymbol{v}\rangle && \forall\, \boldsymbol{v} \in V^n_{\mathrm{div}},\\
&\boldsymbol{u}^{k,n,m,l}_0 = P^n_{\mathrm{div}}\boldsymbol{u}_0.
\end{aligned}$$

Here $P^n_{\mathrm{div}} : L^2(\Omega)^d \to V^n_{\mathrm{div}}$ is simply the L^2–projection defined through

$$\int_\Omega P^n_{\mathrm{div}}\boldsymbol{v}\cdot\boldsymbol{w} = \int_\Omega \boldsymbol{v}\cdot\boldsymbol{w} \qquad \forall\, \boldsymbol{w} \in V^n_{\mathrm{div}}. \tag{3.2}$$

The form $\mathcal{B}$ is meant to represent the convective term and is defined for functions $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in C^\infty_0(\Omega)^d$ as:

$$\mathcal{B}(\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}) := \begin{cases} -\displaystyle\int_\Omega \boldsymbol{u}\otimes\boldsymbol{v} : \boldsymbol{D}(\boldsymbol{w}), & \text{if } V^n_{\mathrm{div}} \subset W^{1,r}_{0,\mathrm{div}}(\Omega)^d,\\ \displaystyle\frac{1}{2}\int_\Omega \boldsymbol{u}\otimes\boldsymbol{w} : \boldsymbol{D}(\boldsymbol{v}) - \boldsymbol{u}\otimes\boldsymbol{v} : \boldsymbol{D}(\boldsymbol{w}), & \text{otherwise.}\end{cases}$$

This definition guarantees that $\mathcal{B}(\boldsymbol{v},\boldsymbol{v},\boldsymbol{v}) = 0$ for every $\boldsymbol{v}$ for which this expression is well defined, regardless of whether $\boldsymbol{v}$ is pointwise divergence-free or not, which is very useful when obtaining a priori estimates; it reduces to the usual weak form of the convective term whenever the velocities are exactly divergence-free. It is now necessary to check that $\mathcal{B}$ can be continuously extended to the spaces involving time. By standard function space interpolation, we have that for almost every $t \in (0,T)$:

$$\begin{aligned}
\int_\Omega |\boldsymbol{u}(t,\cdot)\otimes\boldsymbol{v}(t,\cdot) : \boldsymbol{D}(\boldsymbol{w}(t,\cdot))| &\leq \|\boldsymbol{u}(t,\cdot)\|_{L^{2\check{r}}(\Omega)}\|\boldsymbol{v}(t,\cdot)\|_{L^{2\check{r}}(\Omega)}\|\boldsymbol{D}(\boldsymbol{w}(t,\cdot))\|_{L^{\check{r}'}(\Omega)}\\
&\leq \|\boldsymbol{u}(t,\cdot)\|_{L^{\frac{r(d+2)}{d}}(\Omega)}\|\boldsymbol{v}(t,\cdot)\|_{L^{\frac{r(d+2)}{d}}(\Omega)}\|\boldsymbol{D}(\boldsymbol{w}(t,\cdot))\|_{L^{\check{r}'}(\Omega)}\\
&\leq c\|\boldsymbol{u}(t,\cdot)\|_{W^{1,r}(\Omega)}\|\boldsymbol{v}(t,\cdot)\|_{W^{1,r}(\Omega)}\|\boldsymbol{w}(t,\cdot)\|_{W^{1,\check{r}'}(\Omega)}.
\end{aligned}$$

As in the steady case (cf. [21]), a more restrictive condition is needed in order to bound the additional term in $\mathcal{B}$ whenever the elements are not exactly divergence-free. Namely, if we assume that $r \geq \frac{2(d+1)}{d+2}$ (this is the analogue of the condition $r \geq \frac{2d}{d+1}$ in the steady case) then there is a $q \in (1, \infty]$ such that $\frac{1}{r} + \frac{d}{r(d+2)} + \frac{1}{q} = 1$, and therefore

$$\begin{aligned}\int_\Omega |\boldsymbol{u}(t,\cdot) \otimes \boldsymbol{w}(t,\cdot) : \boldsymbol{D}(\boldsymbol{v}(t,\cdot))| &\leq \|\boldsymbol{u}(t,\cdot)\|_{L^{\frac{r(d+2)}{d}}(\Omega)} \|\boldsymbol{D}(\boldsymbol{v}(t,\cdot))\|_{L^r(\Omega)} \|\boldsymbol{w}(t,\cdot)\|_{L^q(\Omega)} \\ &\leq c\|\boldsymbol{u}(t,\cdot)\|_{W^{1,r}(\Omega)} \|\boldsymbol{v}(t,\cdot)\|_{W^{1,r}(\Omega)} \|\boldsymbol{w}(t,\cdot)\|_{W^{1,\tilde{r}'}(\Omega)}.\end{aligned}$$

On the other hand, using Hölder's inequality we can also obtain the estimate

$$\begin{aligned}\|\mathcal{B}(\boldsymbol{u},\boldsymbol{v},\boldsymbol{w})\|_{L^1(0,T)} \leq& \|\boldsymbol{u}\|_{L^{2r'}(Q)} \|\boldsymbol{v}\|_{L^{2r'}(Q)} \|\boldsymbol{w}\|_{L^r(0,T;W^{1,r}(\Omega))} \\ &+ \|\boldsymbol{u}\|_{L^{2r'}(Q)} \|\boldsymbol{w}\|_{L^{2r'}(Q)} \|\boldsymbol{v}\|_{L^r(0,T;W^{1,r}(\Omega))},\end{aligned}$$

which means that if the $L^{2r'}(Q)^d$ norm of $\boldsymbol{u}$ is finite, then the additional restriction $r \geq \frac{2(d+1)}{d+2}$ is not needed. Moreover, this would also imply that the velocity is an admissible test function, which is useful in the convergence analysis. This motivates the introduction of the penalty term in Formulation $\check{\mathrm{A}}_{\mathrm{k,n,m,l}}$.

Remark 3.4. While Formulation $\check{\mathrm{A}}_{\mathrm{k,n,m,l}}$ does not contain the pressure, in practice the incompressibility condition is enforced through the addition of a Lagrange multiplier $p_j^{k,n,m,l} \in M_0^n$, which could be thought of as the pressure in the system (the reason for the omission of the pressure in the analysis is explained in Remark 3.2). For this reason it is necessary to consider additional assumptions that guarantee inf-sup stability of the spaces V^n and M^n (see Assumptions 2.4 and 2.5). In case the problem does have an integrable pressure p, then it is expected that the sequence of discrete pressures converges to it in $L^1(Q)$.

Remark 3.5. Assumption (A5) also implies the existence of a selection $\boldsymbol{\mathcal{S}} : Q \times \mathbb{R}^{d\times d}_{sym} \to \mathbb{R}^{d\times d}_{sym}$ such that $(\boldsymbol{\tau}, \boldsymbol{\mathcal{S}}(z,\boldsymbol{\tau})) \in \mathcal{A}(z)$ for all $\boldsymbol{\tau} \in \mathbb{R}^{d\times d}_{sym}$, and some models can be written more naturally with a selection of this form; the same analysis as the one presented in this work can be applied to that situation. In fact, in practice it is not necessary to find a selection in order to perform the computations, i.e. in the simulations it is possible to work directly with the implicit function $\boldsymbol{G}$. When performing the analysis though, the function $\boldsymbol{G}$ is not appropriate because many different expressions could lead to the same constitutive relation, but have different mathematical properties.

Remark 3.6. In this work we did not consider a dual formulation, e.g. based on $H(\mathrm{div};\Omega)$, because for the unsteady problem we do not have at our disposal results that guarantee the integrability of $\operatorname{div} \boldsymbol{S}$.

In the next theorem, convergence of the sequence of discrete solutions to a weak solution of the problem is proved. Since the ideas and arguments contained in the proof are similar to the ones presented in the previous sections and follow a similar approach to [61], we will not include here all the details of the calculations unless there is a significant difference.

THEOREM 3.7. *Assume that* $r > \frac{2d}{d+2}$, *let* $\{\Sigma^n, V^n, M^n\}_{n\in\mathbb{N}}$ *be a family of finite element spaces satisfying Assumptions 2.2–2.4. Then for* $k,n,m,l \in \mathbb{N}$ *there exists a sequence* $\{(\boldsymbol{S}_j^{k,n,m,l}, \boldsymbol{u}_j^{k,n,m,l})\}_{j=1}^{T/\tau_m}$ *of solutions of Formulation* $\check{\mathrm{A}}_{\mathrm{k,n,m,l}}$, *and a couple* $(\boldsymbol{S},\boldsymbol{u}) \in L^{r'}_{\mathrm{sym}}(Q)^{d\times d} \cap L^{r'}_{\mathrm{tr}}(Q)^{d\times d} \times L^r(0,T;W^{1,r}_{0,\mathrm{div}}(\Omega)^d) \cap L^\infty(0,T;L^2_{\mathrm{div}}(\Omega)^d)$ *such* *that the corresponding time interpolants (recall (2.15) and (2.16))* $\overline{\boldsymbol{u}}^{k,n,m,l}$, $\tilde{\boldsymbol{u}}^{k,n,m,l}$ *and* $\overline{\boldsymbol{S}}^{k,n,m,l}$ *satisfy (up to a subsequence):*

$$\begin{aligned}
&\overline{\boldsymbol{S}}^{k,n,m,l} \rightharpoonup \boldsymbol{S} && \text{weakly in } L^{r'}(Q)^{d\times d},\\
&\overline{\boldsymbol{u}}^{k,n,m,l} \rightharpoonup \boldsymbol{u} && \text{weakly in } L^{r}(0,T;W^{1,r}_0(\Omega)^d),\\
&\overline{\boldsymbol{u}}^{k,n,m,l}, \tilde{\boldsymbol{u}}^{k,n,m,l} \overset{*}{\rightharpoonup} \boldsymbol{u} && \text{weakly* in } L^{\infty}(0,T;L^2(\Omega)^d),
\end{aligned} \tag{3.3}$$

and $(\boldsymbol{S},\boldsymbol{u})$ *solves Formulation* Ǎ*, with the limits taken in the order* $k\to\infty$, $(n,m)\to\infty$ *and* $l\to\infty$.

Proof. The idea of the proof is common in the analysis of nonlinear PDE: we obtain a priori estimates and use compactness arguments to pass to the limit in the equation. In order to prove the existence of solutions of Formulation $\check{\mathrm{A}}_{\mathrm{k,n,m,l}}$, we need to check that given $(\boldsymbol{S}^{k,n,m,l}_{j-1}, \boldsymbol{u}^{k,n,m,l}_{j-1})$, we can find $(\boldsymbol{S}^{k,n,m,l}_{j}, \boldsymbol{u}^{k,n,m,l}_{j})$, for $j\in\{1,\ldots,T/\tau_m\}$. Testing the equation with $(\boldsymbol{S}^{k,n,m,l}_{j}, \boldsymbol{u}^{k,n,m,l}_{j})$, we see that:

$$\int_\Omega \boldsymbol{\mathcal{D}}^k(\cdot,\boldsymbol{S}^{k,n,m,l}_j):\boldsymbol{S}^{k,n,m,l}_j + \frac{1}{l}\|\boldsymbol{u}^{k,n,m,l}_j\|^{2r'}_{L^{2r'}(\Omega)} \le \langle \boldsymbol{f},\boldsymbol{u}^{k,n,m,l}_j\rangle + \frac{1}{\tau_m}\int_\Omega \boldsymbol{u}^{k,n,m,l}_{j-1}\cdot\boldsymbol{u}^{k,n,m,l}_j. \tag{3.4}$$

On the other hand, since all norms are equivalent in a finite-dimensional normed linear space, there is a constant $C_n>0$ such that:

$$\|\boldsymbol{v}\|_{W^{1,r}(\Omega)} \le C_n\|\boldsymbol{v}\|_{L^{2r'}(\Omega)} \qquad \forall\,\boldsymbol{v}\in V^n_{\mathrm{div}}. \tag{3.5}$$

The constant C_n may blow up as $n\to\infty$, but since n is fixed for now this does not pose a problem. Now, recalling (2.9) and combining (3.4) and (3.5) with a standard corollary of Brouwer's Fixed Point Theorem (cf. [33]) we obtain the existence of solutions of Formulation $\check{\mathrm{A}}_{\mathrm{k,n,m,l}}$. In the first time step (i.e. $j=1$), it is essential to use the fact that the projection P^n_{div} is stable:

$$\|P^n_{\mathrm{div}}\boldsymbol{u}_0\|_{L^2(\Omega)} \le \|\boldsymbol{u}_0\|_{L^2(\Omega)}. \tag{3.6}$$

The estimate (3.5) suffices to guarantee the existence of discrete solutions, but in order to pass to the limit $n\to\infty$, an estimate that does not degenerate as $n\to\infty$ is required. This uniform estimate is a consequence of the discrete inf-sup condition (2.14):

$$\gamma_r\|\boldsymbol{u}^{k,n,m,l}_j\|_{W^{1,r}(\Omega)} \le \|\boldsymbol{\mathcal{D}}^k(\cdot,\boldsymbol{S}^{k,n,m,l}_{j+1})\|_{L^r(\Omega)}. \tag{3.7}$$

Therefore, the following a priori estimate holds:

$$\begin{aligned}
&\sup_{j\in\{1,\ldots,T/\tau_m\}}\|\boldsymbol{u}^{k,n,m,l}_j\|^2_{L^2(\Omega)} + \sum_{j=1}^{T/\tau_m}\|\boldsymbol{u}^{k,n,m,l}_j-\boldsymbol{u}^{k,n,m,l}_{j-1}\|^2_{L^2(\Omega)}\\
&+\tau_m\sum_{j=1}^{T/\tau_m}\|\boldsymbol{S}^{k,n,m,l}_j\|_{L^{r'}(\Omega)} + \tau_m\sum_{j=1}^{T/\tau_m}\|\boldsymbol{u}^{k,n,m,l}_j\|^r_{W^{1,r}(\Omega)}\\
&+\sum_{j=1}^{T/\tau_m}\|\boldsymbol{\mathcal{D}}^k(\cdot,\cdot,\boldsymbol{S}^{k,n,m,l}_j)\|_{L^r(Q^j_{j-1})} + \frac{\tau_m}{l}\sum_{j=1}^{T/\tau_m}\|\boldsymbol{u}^{k,n,m,l}_j\|^{2r'}_{L^{2r'}(\Omega)} \le c,
\end{aligned} \tag{3.8}$$

where c is a positive constant that depends on the data; in particular, c is independent of k,n,m and l. Let $\overline{\boldsymbol{u}}^{k,n,m,l}\in L^\infty(0,T;V^n_{\mathrm{div}})$ and $\tilde{\boldsymbol{u}}^{k,n,m,l}\in C([0,T];V^n_{\mathrm{div}})$

be the piecewise constant and piecewise linear interpolants defined by the sequence $\{\boldsymbol{u}_j^{k,n,m,l}\}_{j=1}^{T/\tau_m}$ (see (2.15) and (2.16)) and let $\overline{\boldsymbol{S}}^{k,n,m,l} \in L^\infty(0,T;\Sigma^n_{\text{sym}})$ be the piecewise constant interpolant defined by the sequence $\{\boldsymbol{S}_j^{k,n,m,l}\}_{j=1}^{T/\tau_m}$. Furthermore, define also the piecewise constant interpolants:

$$\overline{\boldsymbol{f}}(t,\cdot) := \boldsymbol{f}_j(\cdot), \qquad \overline{\boldsymbol{\mathcal{D}}}^k(t,\cdot,\cdot) := \boldsymbol{\mathcal{D}}^k_j(\cdot,\cdot), \quad t \in (t_{j-1},t_j], \quad j \in \{1,\dots,T/\tau_m\}$$

Then the discrete formulation can be rewritten as:

$$\int_\Omega (\overline{\boldsymbol{\mathcal{D}}}^k(t,\cdot,\overline{\boldsymbol{S}}^{k,n,m,l}) - \boldsymbol{D}(\overline{\boldsymbol{u}}^{k,n,m,l})) : \boldsymbol{\tau} = 0 \qquad \forall\, \boldsymbol{\tau} \in \Sigma^n_{\text{sym}},$$

$$\int_\Omega \partial_t \tilde{\boldsymbol{u}}^{k,n,m,l}\cdot \boldsymbol{v} + \frac{1}{l}\int_\Omega |\overline{\boldsymbol{u}}^{k,n,m,l}|^{2r'-2}\overline{\boldsymbol{u}}^{k,n,m,l}\cdot \boldsymbol{v}$$
$$+ \int_\Omega (\overline{\boldsymbol{S}}^{k,n,m,l} : \boldsymbol{D}(\boldsymbol{v}) + \mathcal{B}(\overline{\boldsymbol{u}}^{k,n,m,l},\overline{\boldsymbol{u}}^{k,n,m,l},\boldsymbol{v})) = \langle \overline{\boldsymbol{f}},\boldsymbol{v}\rangle \qquad \forall\, \boldsymbol{v}\in V^n_{\text{div}},$$
$$\tilde{\boldsymbol{u}}^{k,n,m,l}(0,\cdot) = P^n_{\text{div}}\boldsymbol{u}_0(\cdot).$$

The a priori estimate (3.8) can in turn be written as:

$$\|\overline{\boldsymbol{u}}^{k,n,m,l}\|^2_{L^\infty(0,T;L^2(\Omega))} + \tau_m\|\partial_t\tilde{\boldsymbol{u}}^{k,n,m,l}\|^2_{L^2(Q)} + \|\overline{\boldsymbol{S}}^{k,n,m,l}\|^{r'}_{L^{r'}(Q)}$$
$$+ \|\overline{\boldsymbol{u}}^{k,n,m,l}\|^r_{L^r(0,T;W^{1,r}(\Omega))} + \|\boldsymbol{\mathcal{D}}^k(\cdot,\cdot,\overline{\boldsymbol{S}}^{k,n,m,l})\|^r_{L^r(Q)} + \frac{1}{l}\|\overline{\boldsymbol{u}}^{k,n,m,l}\|^{2r'}_{L^{2r'}(Q)} \le c. \tag{3.9}$$

Using the equivalence of norms in finite-dimensional spaces we also obtain

$$\|\partial_t\tilde{\boldsymbol{u}}^{k,n,m,l}\|_{L^\infty(0,T;L^2(\Omega))} \le c(n)\|\partial_t\tilde{\boldsymbol{u}}^{k,n,m,l}\|_{L^2(Q)},$$

and together with the a priori estimate this implies that

$$\|\tilde{\boldsymbol{u}}^{k,n,m,l}\|_{W^{1,\infty}(0,T;L^2(\Omega))} \le c(n,m). \tag{3.10}$$

Therefore, up to subsequences, as $k\to\infty$ we have:

$$\begin{aligned}
\overline{\boldsymbol{u}}^{k,n,m,l} &\to \overline{\boldsymbol{u}}^{n,m,l} && \text{strongly in } L^\infty(0,T;L^2(\Omega)^d),\\
\tilde{\boldsymbol{u}}^{k,n,m,l} &\to \tilde{\boldsymbol{u}}^{n,m,l} && \text{strongly in } W^{1,\infty}(0,T;L^2(\Omega)^d),\\
\overline{\boldsymbol{u}}^{k,n,m,l} &\to \overline{\boldsymbol{u}}^{n,m,l} && \text{strongly in } L^{2r'}(Q)^d,\\
\overline{\boldsymbol{u}}^{k,n,m,l} &\to \overline{\boldsymbol{u}}^{n,m,l} && \text{strongly in } L^r(0,T;W^{1,r}_0(\Omega)^d),\\
\overline{\boldsymbol{S}}^{k,n,m,l} &\to \overline{\boldsymbol{S}}^{n,m,l} && \text{strongly in } L^{r'}(Q)^{d\times d},\\
\boldsymbol{\mathcal{D}}^k(\cdot,\cdot,\overline{\boldsymbol{S}}^{k,n,m,l}) &\rightharpoonup \boldsymbol{D}^{n,m,l} && \text{weakly in } L^r(Q)^{d\times d},\\
\overline{\boldsymbol{\mathcal{D}}}^k(\cdot,\cdot,\overline{\boldsymbol{S}}^{k,n,m,l}) &\rightharpoonup \overline{\boldsymbol{D}}^{n,m,l} && \text{weakly in } L^r(Q)^{d\times d},\\
\boldsymbol{\mathcal{D}}^k_j(\cdot,\boldsymbol{S}^{k,n,m,l}_j) &\rightharpoonup \boldsymbol{D}^{n,m,l}_j && \text{weakly in } L^r(\Omega)^{d\times d}, \text{ for } j\in\{1,\dots,T/\tau_m\}.
\end{aligned}$$

Since the function $\boldsymbol{D}^k_j$ is simply an average in time, the uniqueness of the weak limit implies that

$$\boldsymbol{D}^{n,m,l}_j(\cdot) = \frac{1}{\tau_m}\int_{t_{j-1}}^{t_j} \boldsymbol{D}^{n,m,l}(t,\cdot)\,\mathrm{d}t, \qquad j\in\{1,\dots,T/\tau_m\}, \tag{3.11}$$

and that $\overline{\boldsymbol{D}}^{n,m,l}$ is the piecewise constant interpolant determined by the sequence $\{\boldsymbol{D}^{n,m,l}_j\}_{j=1}^{T/\tau_m}$. Moreover, since the convergence of the velocity and stress sequences is strong, it is straightforward to pass to the limit $k \to \infty$ and thus we obtain

$$\int_\Omega (\overline{\boldsymbol{D}}^{n,m,l} - \boldsymbol{D}(\overline{\boldsymbol{u}}^{n,m,l})) : \boldsymbol{\tau} = 0 \qquad \forall \boldsymbol{\tau} \in \Sigma^n_{\text{sym}},$$

$$\int_\Omega \partial_t \tilde{\boldsymbol{u}}^{n,m,l} \cdot \boldsymbol{v} + \frac{1}{l}\int_\Omega |\overline{\boldsymbol{u}}^{n,m,l}|^{2r'-2}\,\overline{\boldsymbol{u}}^{n,m,l} \cdot \boldsymbol{v}$$
$$+ \int_\Omega (\overline{\boldsymbol{S}}^{n,m,l} : \boldsymbol{D}(\boldsymbol{v}) + \mathcal{B}(\overline{\boldsymbol{u}}^{n,m,l}, \overline{\boldsymbol{u}}^{n,m,l}, \boldsymbol{v})) = \langle \overline{\boldsymbol{f}}, \boldsymbol{v}\rangle \qquad \forall \boldsymbol{v} \in V^n_{\text{div}}.$$

It is also clear that the initial condition $\tilde{\boldsymbol{u}}^{n,m,l}(0,\cdot) = P^n_{\text{div}}\boldsymbol{u}_0(\cdot)$ holds, since the expression on the right-hand side is independent of k. The identification of the constitutive relation can be carried out using (2.10) in exactly the same manner as in [61], which means that (the strong convergence is again essential):

$$(\boldsymbol{D}^{n,m,l}, \overline{\boldsymbol{S}}^{n,m,l}) \in \mathcal{A}(\cdot), \text{ a.e. in } (0,T)\times\Omega. \tag{3.12}$$

The next step is to take the limit in both the time and space discretisations simultaneously. The weak lower semicontinuity of the norms and the estimate (3.9) imply that:

$$\|\overline{\boldsymbol{u}}^{n,m,l}\|^2_{L^\infty(0,T;L^2(\Omega))} + \tau_m\|\partial_t\tilde{\boldsymbol{u}}^{n,m,l}\|^2_{L^2(Q)} + \|\overline{\boldsymbol{S}}^{n,m,l}\|^{r'}_{L^{r'}(Q)}$$
$$+ \|\overline{\boldsymbol{u}}^{n,m,l}\|^r_{L^r(0,T;W^{1,r}(\Omega))} + \|\boldsymbol{D}^{n,m,l}\|^r_{L^r(Q)} + \frac{1}{l}\|\overline{\boldsymbol{u}}^{n,m,l}\|^{2r'}_{L^{2r'}(Q)} \le c, \tag{3.13}$$

and

$$\|\tilde{\boldsymbol{u}}^{n,m,l}\|^2_{L^\infty(0,T;L^2(\Omega))} = \|\overline{\boldsymbol{u}}^{n,m,l}\|^2_{L^\infty(0,T;L^2(\Omega))} \le c, \tag{3.14}$$

where c is a constant, independent of n, m and l. Consequently, there exist (not relabelled) subsequences such that, as $n, m \to \infty$:

$$\overline{\boldsymbol{u}}^{n,m,l} \overset{*}{\rightharpoonup} \boldsymbol{u}^l \qquad \text{weakly* in } L^\infty(0,T;L^2(\Omega)^d),$$
$$\tilde{\boldsymbol{u}}^{n,m,l} \overset{*}{\rightharpoonup} \boldsymbol{u}^l \qquad \text{weakly* in } L^\infty(0,T;L^2(\Omega)^d),$$
$$\overline{\boldsymbol{u}}^{n,m,l} \rightharpoonup \boldsymbol{u}^l \qquad \text{weakly in } L^r(0,T;W^{1,r}_0(\Omega)^d),$$
$$\overline{\boldsymbol{S}}^{n,m,l} \rightharpoonup \boldsymbol{S}^l \qquad \text{weakly in } L^{r'}(Q)^{d\times d},$$
$$\boldsymbol{D}^{n,m,l} \rightharpoonup \boldsymbol{D}^l \qquad \text{weakly in } L^{r}(Q)^{d\times d},$$
$$\overline{\boldsymbol{D}}^{n,m,l} \rightharpoonup \overline{\boldsymbol{D}}^l \qquad \text{weakly in } L^{r}(Q)^{d\times d},$$
$$\frac{1}{l}\int_Q |\overline{\boldsymbol{u}}^{n,m,l}|^{2r'-2}\overline{\boldsymbol{u}}^{n,m,l} \rightharpoonup \frac{1}{l}\int_Q |\boldsymbol{u}^l|^{2r'-2}\boldsymbol{u}^{n,m,l} \qquad \text{weakly in } L^{(2r')'}(Q)^d.$$

At this point it is a standard step to use the Aubin–Lions lemma to obtain strong convergence of subsequences. However, following [61], we will instead use Simon's compactness lemma; this choice is made to avoid the need for stability estimates of P^n_{div} in Sobolev norms, which would require additional assumptions on the mesh. To apply this lemma, it will be more convenient to work with the modified interpolant:

$$\hat{\boldsymbol{u}}^{n,m,l}(t,\cdot) := \begin{cases} \boldsymbol{u}_1^{n,m,l}(\cdot), & \text{if } t\in[0,t_1),\\ \tilde{\boldsymbol{u}}^{n,m,l}(t,\cdot), & \text{if } t \in [t_1,T]. \end{cases}$$

Let $\epsilon > 0$ be such that $s+\epsilon < T$ and let $\boldsymbol{v} \in V^n_{\mathrm{div}}$. Then, using the definition of $\hat{\boldsymbol{u}}^{n,m,l}$ we have

$$\begin{aligned}
&\int_\Omega (\hat{\boldsymbol{u}}^{n,m,l}(s+\epsilon,x) - \hat{\boldsymbol{u}}^{n,m,l}(s+\epsilon,x)) \cdot \boldsymbol{v}(x)\,\mathrm{d}x \\
&= \int_{\max(s,\tau_m)}^{s+\epsilon} \int_\Omega \partial_t \hat{\boldsymbol{u}}^{n,m,l}(t,x)\cdot \boldsymbol{v}(x)\,\mathrm{d}x\,\mathrm{d}t \\
&= \int_{\max(s,\tau_m)}^{s+\epsilon} \int_\Omega \partial_t \tilde{\boldsymbol{u}}^{n,m,l}(t,x)\cdot \boldsymbol{v}(x)\,\mathrm{d}x\,\mathrm{d}t \\
&= \int_{\max(s,\tau_m)}^{s+\epsilon} \Big(-\frac{1}{l}\int_\Omega |\overline{\boldsymbol{u}}^{n,m,l}(t,x)|^{2r'-2}\overline{\boldsymbol{u}}^{n,m,l}(t,x)\cdot \boldsymbol{v}(x)\,\mathrm{d}x \\
&- \int_\Omega (\overline{\boldsymbol{S}}^{n,m,l}(t,x) : \boldsymbol{D}(\boldsymbol{v}(x)) + \mathcal{B}(\overline{\boldsymbol{u}}^{n,m,l}(t,x), \overline{\boldsymbol{u}}^{n,m,l}(t,x), \boldsymbol{v}(x)))\,\mathrm{d}x + \langle \overline{\boldsymbol{f}}(t), \boldsymbol{v}\rangle \Big)\,\mathrm{d}t \\
&\leq c(l) \left(\left(\int_{\max(s,\tau_m)}^{s+\epsilon} \|\boldsymbol{v}\|^r_{W^{1,r}(\Omega)}\,\mathrm{d}t \right)^{1/r} + \left(\int_{\max(s,\tau_m)}^{s+\epsilon} \|\boldsymbol{v}\|^{2r'}_{L^{2r'}(\Omega)}\,\mathrm{d}t \right)^{1/2r'} \right) \\
&\leq c(l)(\epsilon^{1/r} + \epsilon^{1/2r'}) \left(\|\boldsymbol{v}\|_{W^{1,r}(\Omega)} + \|\boldsymbol{v}\|_{L^{2r'}(\Omega)} \right).
\end{aligned}$$

Choosing $\boldsymbol{v} = \hat{\boldsymbol{u}}^{n,m,l}(s+\epsilon,\cdot) - \hat{\boldsymbol{u}}^{n,m,l}(s,\cdot)$ we conclude that

$$\int_0^{T-\epsilon} \|\hat{\boldsymbol{u}}^{n,m,l}(s+\epsilon,\cdot) - \hat{\boldsymbol{u}}^{n,m,l}(s,\cdot)\|^2_{L^2(\Omega)}\,\mathrm{d}s \to 0, \text{ as } \epsilon \to 0.$$

On the other hand, the a priori estimates imply that $\hat{\boldsymbol{u}}^{n,m,l}$ is bounded (uniformly in $n,m \in \mathbb{N}$) in $L^2(Q)^d$ and $L^1(0,T;W^{1,r}_0(\Omega)^d)$. Moreover, since $r > \frac{2d}{d+2}$, the embedding $W^{1,r}(\Omega)^d \hookrightarrow L^2(\Omega)^d$ is compact and thus Simon's compactness lemma guarantees the strong convergence:

$$\hat{\boldsymbol{u}}^{n,m,l} \to \boldsymbol{u}^l \qquad \text{strongly in } L^2(Q)^d. \tag{3.15}$$

Since the interpolants converge to the same limit as $\tau_m \to 0$, using standard function space interpolation (and recalling (2.3)) we also obtain that, as $n,m \to \infty$:

$$\tilde{\boldsymbol{u}}^{n,m,l} \to \boldsymbol{u}^l \qquad \text{strongly in } L^p(0,T;L^2(\Omega)^d), \tag{3.16}$$

$$\overline{\boldsymbol{u}}^{n,m,l} \to \boldsymbol{u}^l \qquad \text{strongly in } L^p(0,T;L^2(\Omega)^d) \cap L^q(Q), \tag{3.17}$$

for $p \in [1,\infty)$ and $q \in [1, \max(2r', \frac{q(d+2)}{d}))$.

Now, using the property (2.12), we can check that $\boldsymbol{u}^l$ is actually divergence-free:

$$0 = \int_0^T \int_\Omega \phi\, \Pi^n_M q \,\mathrm{div} \overline{\boldsymbol{u}}^{n,m,l} \to \int_0^T \int_\Omega \phi\, q\, \mathrm{div} \boldsymbol{u}^l \quad \forall q \in L^{r'}(\Omega),\, \phi \in C^\infty_0(0,T). \tag{3.18}$$

Furthermore, (2.12) also yields convergence of the initial condition, as $n,m \to \infty$:

$$\tilde{\boldsymbol{u}}^{n,m,l}(0,\cdot) = P^n_{\mathrm{div}} \boldsymbol{u}_0 \to \boldsymbol{u}_0 \qquad \text{strongly in } L^2(\Omega)^d. \tag{3.19}$$

The functions $\boldsymbol{D}^l$ and $\overline{\boldsymbol{D}}^l$ can easily be identified using the property (2.19) and the definition of the piecewise constant interpolant (3.11). Indeed, for an arbitrary $\boldsymbol{\sigma} \in C^\infty_0(Q)$ we have, as $n,m \to \infty$:

$$\int_0^T \int_\Omega \overline{\boldsymbol{D}}^{n,m,l} : \boldsymbol{\sigma} = \int_0^T \int_\Omega \boldsymbol{D}^{n,m,l} : \overline{\boldsymbol{\sigma}} \to \int_0^T \int_\Omega \boldsymbol{D}^l : \boldsymbol{\sigma}. \tag{3.20}$$

The uniqueness of the weak limit then implies that $\boldsymbol{D}^l = \overline{\boldsymbol{D}}^l$.

Combining all these properties and using an analogous computation to (3.18) it is possible to prove that the limiting functions are a solution of the following problem:

$$\int_0^T \int_\Omega (\boldsymbol{D}^l - \boldsymbol{D}(\boldsymbol{u}^l)) : \boldsymbol{\tau}\,\varphi = 0 \qquad \forall \boldsymbol{\tau} \in C^\infty_{0,\mathrm{sym}}(\Omega)^{d\times d},\ \varphi \in C^\infty_0(0,T),$$

$$-\int_0^T \int_\Omega \boldsymbol{u}^l \cdot \boldsymbol{v}\,\partial_t\varphi - \int_\Omega \boldsymbol{u}_0 \cdot \boldsymbol{v}\varphi(0) + \int_0^T \int_\Omega (\boldsymbol{S}^l - \boldsymbol{u}^l \otimes \boldsymbol{u}^l) : \boldsymbol{D}(\boldsymbol{v})\,\varphi$$
$$+\frac{1}{l}\int_0^T \int_\Omega |\boldsymbol{u}^l|^{2r'-2}\boldsymbol{u}^l \cdot \boldsymbol{v}\,\varphi = \int_0^T \langle \boldsymbol{f}, \boldsymbol{v}\rangle\,\varphi \qquad \forall \boldsymbol{v} \in C^\infty_{0,\mathrm{div}}(\Omega)^d,\ \varphi \in C^\infty_0(-T,T).$$

From the equation above and the estimate (2.3) we then see that the distributional time derivative belongs to the spaces:

$$(3.21) \qquad \partial_t \boldsymbol{u}^l \in L^{\min(r',(2r')')}(0,T;(W^{1,r}_{0,\mathrm{div}}(\Omega)^d \cap L^{2r'}(\Omega)^d)^*),$$

$$(3.22) \qquad \partial_t \boldsymbol{u}^l \in L^{\min(\check{r},(2r')')}(0,T;(W^{1,\check{r}'}_{0,\mathrm{div}}(\Omega)^d)^*).$$

It is important to note that (3.22) holds uniformly in $l \in \mathbb{N}$, while (3.21) does not. Now, observe that

$$W^{1,r}_{0,\mathrm{div}}(\Omega)^d \cap L^{2r'}(\Omega)^d \hookrightarrow L^2_{\mathrm{div}}(\Omega)^d \hookrightarrow (L^2_{\mathrm{div}}(\Omega)^d)^* \hookrightarrow (W^{1,r}_{0,\mathrm{div}}(\Omega)^d \cap L^{2r'}(\Omega)^d)^*.$$

Combining this with (2.4), (2.5), and the fact that $\boldsymbol{u}^l \in L^\infty(0,T;L^2_{\mathrm{div}}(\Omega)^d)$ guarantees that $\boldsymbol{u}^l \in C_w([0,T],L^2_{\mathrm{div}}(\Omega)^d)$. Let $\boldsymbol{v} \in C^\infty_{0,\mathrm{div}}(\Omega)^d$ and $\varphi \in C^\infty(-T,T)$ be such that $\varphi(0) = 1$; then the following equality holds:

$$(3.23) \qquad \int_0^T \int_\Omega \partial_t(\boldsymbol{u}^l\varphi)\cdot\boldsymbol{v} = -\int_\Omega \boldsymbol{u}^l(0,\cdot)\cdot\boldsymbol{v}\,\varphi(0).$$

On the other hand, using the equation we also have that:

$$(3.24) \qquad \int_0^T \int_\Omega \partial_t(\boldsymbol{u}^l\varphi)\cdot\boldsymbol{v} = \int_0^T \int_\Omega \partial_t\boldsymbol{u}^l\cdot\boldsymbol{v}\,\varphi + \int_0^T \int_\Omega \boldsymbol{u}^l\cdot\boldsymbol{v}\,\partial_t\varphi = -\int_\Omega \boldsymbol{u}_0\cdot\boldsymbol{v}\,\varphi(0).$$

Comparing (3.23) and (3.24) we conclude that $\boldsymbol{u}^l(0,\cdot) = \boldsymbol{u}_0(\cdot)$. This proves that the initial condition is attained in the weak sense expected a priori from the embeddings; however, in this case the stronger condition

$$(3.25) \qquad \operatorname*{ess\,lim}_{t\to 0^+} \|\boldsymbol{u}^l(t,\cdot) - \boldsymbol{u}_0(\cdot)\|_{L^2(\Omega)} = 0$$

holds. To see this, note that (3.16) guarantees that, up to a subsequence, $\tilde{\boldsymbol{u}}^{n,m,l}(t,\cdot) \to \tilde{\boldsymbol{u}}^l(t,\cdot)$ in $L^2(\Omega)^d$ for almost every $t \in [0,T]$, and therefore

$$\begin{aligned}
\|\boldsymbol{u}^l(t,\cdot) - \boldsymbol{u}_0(\cdot)\|^2_{L^2(\Omega)} &= \limsup_{n,m\to\infty} \|\tilde{\boldsymbol{u}}^{n,m,l}(t,\cdot) - \tilde{\boldsymbol{u}}^{n,m,l}(0,\cdot)\|^2_{L^2(\Omega)} \\
&= \limsup_{n,m\to\infty} \Big(\|\tilde{\boldsymbol{u}}^{n,m,l}(t,\cdot)\|^2_{L^2(\Omega)} - \|\tilde{\boldsymbol{u}}^{n,m,l}(0,\cdot)\|^2_{L^2(\Omega)} \\
&\qquad + 2\int_\Omega (\tilde{\boldsymbol{u}}^{n,m,l}(0,\cdot) - \tilde{\boldsymbol{u}}^{n,m,l}(t,\cdot))\cdot\tilde{\boldsymbol{u}}^{n,m,l}(0,\cdot) \Big) \\
&\le \limsup_{n,m\to\infty} \Big(\int_0^t \langle \overline{\boldsymbol{f}}, \overline{\boldsymbol{u}}^{n,m,l}\rangle + 2\int_\Omega (\tilde{\boldsymbol{u}}^{n,m,l}(0,\cdot) - \tilde{\boldsymbol{u}}^{n,m,l}(t,\cdot))\cdot\tilde{\boldsymbol{u}}^{n,m,l}(0,\cdot) \Big) \\
&\le \int_0^t \langle \boldsymbol{f}, \boldsymbol{u}^l\rangle + 2\int_\Omega (\boldsymbol{u}^l(0,\cdot) - \boldsymbol{u}^l(t,\cdot))\cdot\boldsymbol{u}^l(0,\cdot),
\end{aligned}$$

for almost every $t \in [0, T]$. Observe also that the monotonicity of the constitutive relation was used to obtain the next to last inequality. Taking the limit $t \to 0^+$ then yields (3.25).

The identification of the constitutive relation, i.e. proving that $(\boldsymbol{D}^l, \boldsymbol{S}^l) \in \mathcal{A}(\cdot)$ almost everywhere, can be carried out with the help of Lemma 2.1. In order to apply the lemma, the only thing that remains to be proved, since we already know that $(\boldsymbol{D}^{n,m,l}, \overline{\boldsymbol{S}}^{n,m,l}) \in \mathcal{A}(\cdot)$ almost everywhere, is that:

$$\limsup_{n,m\to\infty} \int_0^t \int_\Omega \overline{\boldsymbol{S}}^{n,m,l} : \boldsymbol{D}^{n,m,l} \leq \int_0^t \int_\Omega \boldsymbol{S}^l : \boldsymbol{D}^l, \tag{3.26}$$

for almost every $t \in [0, T]$; then taking $t \to T$ we obtain the result in the whole domain Q. The proof of this fact is essentially the same as in [61] and we will not reproduce it here. Moreover, the following energy identity holds:

$$\frac{1}{2}\|\boldsymbol{u}^l(t,\cdot)\|^2_{L^2(\Omega)} + \int_0^t \int_\Omega \boldsymbol{S}^l : \boldsymbol{D}(\boldsymbol{u}^l) + \frac{1}{l}\int_0^t \|\boldsymbol{u}^l\|^{2r'}_{L^{2r'}(\Omega)} = \int_0^t \langle \boldsymbol{f}, \boldsymbol{u}^l\rangle + \|\boldsymbol{u}_0\|^2_{L^2(\Omega)}, \tag{3.27}$$

In time-dependent problems obtaining an energy identity of this kind is not always possible; in this case the energy equality (3.27) can be proved, since the velocity is an admissible test function in space thanks to the fact that its $L^{2r'}$ norm is under control (some mollification is needed to overcome the low integrability in time, see [62, 44]).

Now, (3.13) and the weak and weak* lower semicontinuity of the norms imply that

$$\|\boldsymbol{u}^l\|^2_{L^\infty(0,T;L^2(\Omega))} + \|\boldsymbol{S}^l\|^{r'}_{L^{r'}(Q)} + \|\boldsymbol{u}^l\|^r_{L^r(0,T;W^{1,r}(\Omega))} + \|\boldsymbol{D}^l\|^r_{L^r(Q)} + \frac{1}{l}\|\boldsymbol{u}^l\|^{2r'}_{L^{2r'}(Q)} \leq c, \tag{3.28}$$

where c is a constant independent of l. From this we see that, up to subsequences, as $l \to \infty$:

$$\begin{aligned}
\boldsymbol{u}^l &\overset{*}{\rightharpoonup} \boldsymbol{u} && \text{weakly* in } L^\infty(0,T;L^2(\Omega)^d),\\
\boldsymbol{u}^l &\rightharpoonup \boldsymbol{u} && \text{weakly in } L^r(0,T;W^{1,r}_0(\Omega)^d),\\
\boldsymbol{S}^l &\rightharpoonup \boldsymbol{S} && \text{weakly in } L^{r'}(Q)^{d\times d},\\
\boldsymbol{D}^l &\rightharpoonup \boldsymbol{D} && \text{weakly in } L^r(Q)^{d\times d},\\
\frac{1}{l}\int_Q |\boldsymbol{u}^l|^{2r'-2}\boldsymbol{u}^l &\to 0 && \text{strongly in } L^1(Q)^d.
\end{aligned} \tag{3.29}$$

Furthermore, since $\check{r} \leq r'$ and $r > \frac{2d}{d+2}$, the embedding $W^{1,\check{r}'}_{0,\mathrm{div}}(\Omega)^d \hookrightarrow L^2_{\mathrm{div}}(\Omega)^d$ is compact and hence by the Aubin–Lions lemma (taking into account (3.22)) we have the strong convergence:

$$\boldsymbol{u}^l \to \boldsymbol{u} \qquad \text{strongly in } L^r(0,T;L^2_{\mathrm{div}}(\Omega)^d). \tag{3.30}$$

With the convergence properties (3.29) and (3.30) it is then possible to pass to the limit and prove that the limiting functions satisfy:

$$\begin{aligned}
\int_\Omega (\boldsymbol{D} - \boldsymbol{D}(\boldsymbol{u})) : \boldsymbol{\tau} = 0 \qquad & \forall \boldsymbol{\tau} \in C^\infty_{0,\mathrm{sym}}(\Omega)^{d\times d}, \text{ a.e. } t \in (0,T),\\
\langle \partial_t \boldsymbol{u}, \boldsymbol{v}\rangle + \int_\Omega (\boldsymbol{S} - \boldsymbol{u}\otimes\boldsymbol{u}) : \boldsymbol{D}(\boldsymbol{v}) = \langle \boldsymbol{f}, \boldsymbol{v}\rangle \qquad & \forall \boldsymbol{v} \in C^\infty_{0,\mathrm{div}}(\Omega)^d, \text{ a.e. } t \in (0,T).
\end{aligned}$$

The same argument used to obtain (3.25) can be used here to prove that the initial condition is attained in the strong sense:

$$\operatorname*{ess\,lim}_{t\to 0^+} \|\boldsymbol{u}(t,\cdot)-\boldsymbol{u}_0(\cdot)\|_{L^2(\Omega)} = 0. \tag{3.31}$$

Moreover, since the penalty term vanishes in the limit $l\to\infty$, we can improve the integrability in time:

$$\partial_t \boldsymbol{u}^l \in L^{\check{r}}(0,T;(W^{1,\check{r}'}_{0,\mathrm{div}}(\Omega)^d)^*). \tag{3.32}$$

To show that $(\boldsymbol{D},\boldsymbol{S})\in\mathcal{A}(\cdot)$, Lemma 2.1 will once again be employed. The main difficulty at this stage, just like in the previous works [21, 61], is that the velocity is no longer an admissible test function (and therefore we do not have an energy equality similar to (3.27)). The idea is now to work with Lipschitz truncations of the error $\boldsymbol{e}^l := \boldsymbol{u}^l - \boldsymbol{u}$; it should be noted however that in the present case we need to verify a number of additional hypotheses before Lemma 3.1 can be applied.

Note that equation (3.1) in Lemma 3.1 is written in divergence form. We then need to make a preliminary step and write the penalty term in this form (see [61]). Let $B_0 \subset\subset \Omega$ be an arbitrary ball compactly contained in Ω and let $q\in[1,(2r')')$. Then from the standard theory of elliptic operators we know that for almost every $t\in[0,T]$ there is a unique $\boldsymbol{g}_3^l(t,\cdot)\in W^{2,q}(B_0)^d\cap W^{1,q}_0(B_0)$ such that:

$$\int_{B_0} \nabla \boldsymbol{g}_3^l(t,\cdot) : \nabla \boldsymbol{v} = \frac{1}{l}\int_{B_0} |\boldsymbol{u}^l(t,\cdot)|^{2r'-2}\boldsymbol{u}^l(t,\cdot)\cdot\boldsymbol{v} \qquad \forall\, \boldsymbol{v}\in C^\infty_{0,\mathrm{div}}(\Omega)^d,$$

$$\|\boldsymbol{g}_3^l(t,\cdot)\|_{W^{2,q}(B_0)} \le c\left\|\frac{1}{l}|\boldsymbol{u}^l(t,\cdot)|^{2r'-2}\boldsymbol{u}^l(t,\cdot)\right\|_{L^q(B_0)}.$$

This means in particular (by (3.29) and standard function space interpolation) that for a fixed time interval $I_0\subset\subset(0,T)$ we have:

$$\boldsymbol{g}_3^l \to \boldsymbol{0} \qquad \text{strongly in } L^q(I_0;W^{1,q}(B_0)^d), \quad \forall\, q\in[1,(2r')'). \tag{3.33}$$

Defining $Q_0 := I_0\times B_0$ and

$$\boldsymbol{G}_1^l := \boldsymbol{S}^l - \boldsymbol{S},$$

$$\boldsymbol{G}_2^l := \boldsymbol{u}^l\otimes\boldsymbol{u}^l - \boldsymbol{u}\otimes\boldsymbol{u} - \nabla\boldsymbol{g}_3^l,$$

we readily see that the error $\boldsymbol{e}^l$ satisfies the equation

$$\int_{Q_0} \partial_t \boldsymbol{e}^l\cdot\boldsymbol{w} = \int_{Q_0} (\boldsymbol{G}_1^l + \boldsymbol{G}_2^l):\nabla\boldsymbol{w} \qquad \forall\,\boldsymbol{w}\in C^\infty_{0,\mathrm{div}}(Q_0)^d. \tag{3.34}$$

Additionally, as a consequence of (3.29), (3.33) and (3.30) we also have that for any $q\in[1,\min(\check{r},(2r')')$, the sequence $\boldsymbol{u}^l$ is bounded in $L^\infty(I_0;W^{1,q}(Q_0)^d)$ and that:

$$\boldsymbol{G}_1^l \rightharpoonup \boldsymbol{0} \qquad \text{weakly in } L^{r'}(Q_0)^{d\times d},$$

$$\boldsymbol{G}_2^l \to \boldsymbol{0} \qquad \text{strongly in } L^{q}(Q_0)^{d\times d},$$

$$\boldsymbol{u}^l \to \boldsymbol{u} \qquad \text{strongly in } L^{q}(Q_0)^{d}.$$

Consequently, the assumptions of Lemma 3.1 are satisfied. It now suffices to prove for an arbitrary $\theta\in(0,1)$ that

$$\limsup_{l\to\infty}\int_{\frac{1}{8}Q_0} [(\boldsymbol{D}(\boldsymbol{u}^l)-\boldsymbol{\mathcal{D}}(\cdot,\boldsymbol{S})):(\boldsymbol{S}^l-\boldsymbol{S})]^\theta \le 0, \tag{3.35}$$

Once this has been shown, Chacon's biting lemma and Vitali's convergence theorem will imply, together with Lemma 2.1, that $(\boldsymbol{D},\boldsymbol{S}) \in \mathcal{A}(\cdot)$ almost everywhere in $\frac{1}{8}Q_0$ (see the details e.g. in [14]). From here then the result follows by observing that Q can be covered by a union of such cylinders (e.g. by using a Whitney covering).

In order to prove (3.35), first let $\mathcal{B}_{\lambda_{l,j}} \subset \Omega$ be the family of open sets and let $\{\boldsymbol{e}^{l,j}\}_{l,j\in\mathbb{N}}$ be the sequence of Lipschitz truncations described in Lemma 3.1. If we define

$$H^l(\cdot) := (\boldsymbol{D}(\boldsymbol{u}^l) - \boldsymbol{\mathcal{D}}(\cdot,\boldsymbol{S})) : (\boldsymbol{S}^l - \boldsymbol{S}) \in L^1(Q), \tag{3.36}$$

then we have by Hölder's inequality that

$$\int_{\frac{1}{8}Q_0} |H^l|^\theta \leq |Q|^{1-\theta}\left(\int_{\frac{1}{8}Q_0\setminus\mathcal{B}_{\lambda_{l,j}}} H^l\right)^\theta + |\mathcal{B}_{\lambda_{l,j}}|^{1-\theta}\left(\int_{\frac{1}{8}Q_0} H^l\right)^\theta.$$

The second term on the right-hand side can be dealt with easily, since H^l is bounded uniformly in $L^1(Q)$ thanks to the a priori estimate (3.28), and the properties described in Lemma 3.1 imply that

$$\limsup_{l\to\infty} |\mathcal{B}_{\lambda_{l,j}}|^{1-\theta} \leq \limsup_{l\to\infty} |\lambda_{l,j}^r \mathcal{B}_{\lambda_{l,j}}|^{1-\theta} \leq c2^{-j(1-\theta)}, \quad \text{for } j \geq j_0, \tag{3.37}$$

where c is a positive constant. For the first term, observe that

$$\begin{aligned}
\int_{\frac{1}{8}Q_0\setminus\mathcal{B}_{\lambda_{l,j}}} H^l &= \int_{\frac{1}{8}Q_0} H^l\,\zeta\,\mathbb{1}_{\mathcal{B}^c_{\lambda_{l,j}}} \\
&= \int_{\frac{1}{8}Q_0} \boldsymbol{D}(\boldsymbol{e}^l) : (\boldsymbol{S}^l - \boldsymbol{S})\,\zeta\,\mathbb{1}_{\mathcal{B}^c_{\lambda_{l,j}}} + \int_{\frac{1}{8}Q_0\setminus\mathcal{B}_{\lambda_{l,j}}} (\boldsymbol{D}(\boldsymbol{u}) - \boldsymbol{\mathcal{D}}(\cdot,\boldsymbol{S})) : (\boldsymbol{S}^l - \boldsymbol{S}) \\
&\leq \left|\int_{\frac{1}{8}Q_0} \boldsymbol{D}(\boldsymbol{e}^{l,j}) : \boldsymbol{G}_1^l\,\zeta\,\mathbb{1}_{\mathcal{B}^c_{\lambda_{l,j}}}\right| + \left|\int_{\frac{1}{8}Q_0} (\boldsymbol{D}(\boldsymbol{u}) - \boldsymbol{\mathcal{D}}(\cdot,\boldsymbol{S})) : (\boldsymbol{S}^l - \boldsymbol{S})\right| \\
&\quad + \left|\int_{\mathcal{B}_{\lambda_{l,j}}} (\boldsymbol{D}(\boldsymbol{u}) - \boldsymbol{\mathcal{D}}(\cdot,\boldsymbol{S})) : (\boldsymbol{S}^l - \boldsymbol{S})\right|,
\end{aligned}$$

where $\zeta \in C^\infty_{0,\mathrm{div}}(\frac{1}{6}Q_0)$ is the function introduced in Lemma 3.1. Taking $\limsup_{l\to\infty}$ the assertion follows by taking $j \to \infty$. In particular, we used for the first term Lemma 3.1 part 6, with $\boldsymbol{H} = \boldsymbol{0}$, for the second term the weak convergence of $\boldsymbol{S}^l$ and for the third term the fact that $\{\boldsymbol{S}^l\}_{l\in\mathbb{N}}$ is bounded, together with (3.37). To conclude the proof, note that the fact that $\boldsymbol{u}$ is divergence-free and Assumption (A6) imply that $\mathrm{tr}(\boldsymbol{S}) = 0$, and so $\boldsymbol{S} \in L^{r'}_{\mathrm{sym}}(\Omega)^{d\times d} \cap L^{r'}_{\mathrm{tr}}(\Omega)^{d\times d}$. ∎

Remark 3.8. Formulation $\check{\mathrm{A}}_{\mathrm{k,n,m,l}}$ is a four-step approximation in which the indices k, n, m, l refer to the approximation of the graph by smooth functions, the finite element discretisation, the discretisation in time, and the penalty term, respectively. The same approach can be used to define a 3-field formulation for the steady problem and the unsteady problem without convection and the proof remains valid with some simplifications; for instance, for the steady system without convective term, only the indices k and n are needed. Furthermore, in those cases the convergence of the sequence of discrete pressures can be guaranteed in the corresponding Lebesgue spaces.

Remark 3.9. The argument used to prove the existence of the discrete solutions is more involved here than in the original works [21, 13], because the coercivity with respect to $\|\boldsymbol{u}_j^{k,n,m,l}\|_{W^{1,r}(\Omega)}$ cannot be deduced from Formulation $\check{\mathrm{A}}_{\mathrm{k,n,m,l}}$ by simply testing with the solution. An alternative approach could be to include in the equation an additional diffusion term of the form:

$$\frac{1}{k}\int_\Omega |\boldsymbol{D}(\boldsymbol{u}_j^{k,n,m,l})|^{r-2}\boldsymbol{D}(\boldsymbol{u}_j^{k,n,m,l}) : \boldsymbol{D}(\boldsymbol{v}),$$

which would be completely acceptable if we only cared about the existence of weak solutions, but is undesirable from the point of view of the computation of the finite element approximations, since it introduces an additional nonlinearity in the discrete problem.

Remark 3.10. In the proof of Theorem 3.7 the limits $k \to \infty$, $(n, m) \to \infty$ and $l \to \infty$ were taken successively. In contrast to the steady case considered in [21], here it is not known whether we can take the limits at once. The result is likely to hold as well, but the proof would require a discrete version of the parabolic Lipschitz truncation, which is not available at the moment.

Remark 3.11. In case the symmetric velocity gradient is a quantity of interest, the approach presented here can be easily extended to a four-field formulation with unknowns $(\boldsymbol{D}, \boldsymbol{S}, \boldsymbol{u}, p)$. The only additional assumption needed in that case would be an inf-sup condition of the form:

$$\inf_{\boldsymbol{\sigma}\in\Sigma^n_{\mathrm{div}}(\boldsymbol{0})} \sup_{\boldsymbol{\tau}\in\Sigma^n_{\mathrm{sym}}} \frac{\int_\Omega \boldsymbol{\sigma} : \boldsymbol{\tau}}{\|\boldsymbol{\sigma}\|_{L^{s'}(\Omega)}\|\boldsymbol{\tau}\|_{L^s(\Omega)}} \geq \delta_s, \tag{3.38}$$

where $\delta_s > 0$ is independent of n.

4. Numerical experiments. According to the analysis carried out in the previous section, the addition of the penalty term is necessary when $r \in (\frac{2d}{d+2}, \frac{3d+2}{d+2}]$. However, in the examples we observed that the method converges regardless of whether the penalty term is present or not. This could be an indication that the requirement to include this penalty term is only a technical obstruction and that there might be a different approach to showing convergence of the numerical method that could avoid its inclusion in the numerical method. On the other hand, it could also be the case that exact solutions with more severe singularities than the ones considered in our numerical experiments are needed to demonstrate pathological behaviour. In any case, it appears that in most applications the penalty term can be safely omitted and for this reason it is not discussed in the numerical examples below.

4.1. Carreau fluid and orders of convergence. The framework presented in this work is so broad that in general it is not possible to guarantee uniqueness of solutions; in particular it is not clear how error estimates could be obtained. However, as this computational example will show, the discrete formulations presented here appear to recover the expected orders of convergence in the cases where these orders are known.

In the first part of this numerical experiment we solved the steady problem without convection with the Carreau constitutive law (as stated in Remark 3.8, the same 3-field approximation can be applied in this setting):

$$\boldsymbol{S}(\boldsymbol{D}) := 2\nu \left(\varepsilon^2 + |\boldsymbol{D}^2|\right)^{\frac{r-2}{2}} \boldsymbol{D}, \tag{4.1}$$

where $r \geq 1$ and $\varepsilon, \nu > 0$. This is one of the most common non-Newtonian models that present a power-law structure (note that for $r = 2$ we recover the Newtonian model), and has the advantage that it is not singular at the origin (i.e. when $\boldsymbol{D} = \boldsymbol{0}$), unlike the usual power-law constitutive relation. Observe that the constitutive relation is smooth, and therefore only the limit $n \to \infty$ is needed in the results from the previous section. The problem was solved on the unit square $\Omega = (0,1)^2$ with a Dirichlet boundary condition for the velocity defined so as to match the value of the exact solution, which was chosen as:

$$\boldsymbol{u}(\boldsymbol{x}) = |\boldsymbol{x}|^{a-1}(x_2, -x_1)^{\mathrm{T}}, \qquad p(\boldsymbol{x}) = |\boldsymbol{x}|^b, \tag{4.2}$$

where a, b are parameters used to control the smoothness of the solutions. Define the auxiliary function $\boldsymbol{F} := \mathbb{R}^{d\times d} \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$ as:

$$\boldsymbol{F}(\boldsymbol{B}) := (\varepsilon + |\boldsymbol{B}^{\mathrm{sym}}|)^{\frac{r-2}{2}} \boldsymbol{B}^{\mathrm{sym}}, \tag{4.3}$$

where $\boldsymbol{B}^{\mathrm{sym}} := \frac{1}{2}(\boldsymbol{B} + \boldsymbol{B}^T)$. In [5, 38] it was proved for systems of the form (4.1) that if $\boldsymbol{F}(\boldsymbol{D}(\boldsymbol{u})) \in W^{1,2}(\Omega)^{d\times d}$ and $p \in W^{1,r'}(\Omega)$ then the following error estimates hold:

$$\|\boldsymbol{F}(\boldsymbol{D}(\boldsymbol{u})) - \boldsymbol{F}(\boldsymbol{D}(\boldsymbol{u}^n))\|_{L^2(\Omega)} \leq c h_n^{\min\{1, \frac{r'}{2}\}},$$
$$\|p - p^n\|_{L^{r'}(\Omega)} \leq c h_n^{\min\{\frac{2}{r'}, \frac{r'}{2}\}}.$$

In our case, the conditions $\boldsymbol{F}(\boldsymbol{D}(\boldsymbol{u})) \in W^{1,2}(\Omega)^{d\times d}$ and $p \in W^{1,r'}(\Omega)$ amount to requiring that $a > 1$ and $b > \frac{2}{r} - 1$. These parameters were then chosen to be $a = 1.01$ and $b = \frac{2}{r} - 0.99$ in order to be close to the regularity threshold. We discretised this problem with the Scott–Vogelius element for the velocity and pressure and discontinuous piecewise polynomials for the stress variables:

$$\Sigma^n = \{\boldsymbol{\sigma} \in L^\infty(\Omega)^{d\times d} : \boldsymbol{\sigma}|_K \in \mathbb{P}_k(K)^{d\times d}, \text{ for all } K \in \mathcal{T}_n\},$$
$$V^n = \{\boldsymbol{w} \in W^{1,r}(\Omega)^d : \boldsymbol{w}|_{\partial\Omega} = \boldsymbol{u}, \boldsymbol{w}|_K \in \mathbb{P}_{k+1}(K)^d \text{ for all } K \in \mathcal{T}_n\},$$
$$M^n = \{q \in L^\infty(\Omega) : q|_k \in \mathbb{P}_k(K) \text{ for all } K \in \mathcal{T}_n\}.$$

The problem was solved using `firedrake` [55] with $\nu = 0.5$, $\varepsilon = 10^{-5}$ and $k = 1$ on a barycentrically refined mesh (obtained using `gmsh` [32]) to guarantee inf-sup stability. The discretised nonlinear problems were linearised using Newton's method with the L^2 line search algorithm of PETSc [3, 11]; the Newton solver was deemed to have converged when the Euclidean norm of the residual fell below 1×10^{-8}. The linear systems were solved with a sparse direct solver from the `umfpack` library [19]. In the implementation, the uniqueness of the pressure was recovered not by using a zero mean condition but rather by orthogonalising against the nullspace of constants. The experimental orders of convergence in the different norms are shown in Tables 1 and 2 (note that the tables do not contain the values of the numerical error, but rather the order of convergence corresponding to the norm indicated in each column).

From Tables 1 and 2 it can be seen that the algorithm recovers the expected orders of convergence. In the case of the stress we obtain the same order as for the pressure, which seems natural from the point of view of the equation. In [38] it is claimed that for $r < 2$ the order of convergence for the velocity should be equal to 1; in our numerical simulations the experimental order of convergence seems to approach $\frac{2}{r}$, which is slightly larger than 1. This difference may be due to the fact that in [38]

Table 1: Experimental order of convergence for the steady problem without convection with $r = 1.5$.

h_n	$\|\boldsymbol{F}(\boldsymbol{D}(\boldsymbol{u}))\|_{L^2(\Omega)}$	$\|\boldsymbol{u}\|_{W^{1,r}(\Omega)}$	$\|p\|_{L^{r'}(\Omega)}$	$\|\boldsymbol{S}\|_{L^{r'}(\Omega)}$
0.5	0.9075	1.0180	0.3647	0.6692
0.25	0.9803	1.2160	0.5396	0.6697
0.125	1.0023	1.2975	0.6565	0.6713
0.0625	1.0062	1.3205	0.6706	0.6716
0.03125	1.0071	1.3319	0.6715	0.6716
Expected	1.0	-	0.667	-

Table 2: Experimental order of convergence for the steady problem without convection with $r = 1.8$.

h_n	$\|\boldsymbol{F}(\boldsymbol{D}(\boldsymbol{u}))\|_{L^2(\Omega)}$	$\|\boldsymbol{u}\|_{W^{1,r}(\Omega)}$	$\|p\|_{L^{r'}(\Omega)}$	$\|\boldsymbol{S}\|_{L^{r'}(\Omega)}$
0.5	0.9132	0.9361	0.4955	0.8434
0.25	0.9826	1.0652	0.7271	0.8822
0.125	1.0040	1.1073	0.8671	0.8948
0.0625	1.0078	1.1167	0.8916	0.8966
0.03125	1.0087	1.1197	0.8959	0.8968
Expected	1.0	-	0.889	-

the author works with piecewise linear elements for the velocity while here quadratic elements were employed.

In the second part of the experiment we employed again the Carreau constitutive law (4.1), but now considering the full system (2.11). The right-hand side, initial condition and boundary condition were chosen so as to match the ones defined by the exact solution:

$$\boldsymbol{u}(t,\boldsymbol{x}) = t|\boldsymbol{x}|^{a-1}(x_2,-x_1)^{\mathrm{T}}, \quad p(t,\boldsymbol{x}) = t^2|\boldsymbol{x}|^b.$$

In [25], the following error estimate for the approximation of time-dependent systems of this form, but without convection, was obtained for $r \in [\frac{2d}{d+2},\infty)$:

$$\|\boldsymbol{u}-\overline{\boldsymbol{u}}^{n,m}\|_{L^\infty(0,T;L^2(\Omega))} + \|\boldsymbol{F}(\boldsymbol{D}(\boldsymbol{u})) - \boldsymbol{F}(\boldsymbol{D}(\overline{\boldsymbol{u}}^{n,m}))\|_{L^2(Q)} \leq c\left(\tau_m + h_n^{\min\{1,\frac{2}{r}\}}\right),$$

assuming that $\boldsymbol{u}_0 \in W^{1,r}_{0,\mathrm{div}}(\Omega)^d$ and that the following additional regularity properties of the solution and the data hold:

$$\|\nabla \boldsymbol{F}(\boldsymbol{D}(\boldsymbol{u}_0))\|_{L^2(\Omega)} + \|\nabla \boldsymbol{S}(\boldsymbol{D}(\boldsymbol{u}_0))\|_{L^2(\Omega)} \leq c,$$
$$\|\boldsymbol{u}\|_{W^{1,2}(0,T;L^2(\Omega))} + \|\boldsymbol{u}\|_{L^2(0,T;W^{2,2}(\Omega))} + \|\boldsymbol{F}(\boldsymbol{D}(\boldsymbol{u}))\|_{L^2(0,T;W^{1,2}(\Omega))} \leq c.$$

The same order of convergence was obtained in [6] for $r \in (\frac{3}{2},2]$ in 3D for a semi-implicit discretisation of the unsteady system with convection assuming that $\boldsymbol{u}_0 \in W^{2,2}_{0,\mathrm{div}}(\Omega)^d$, $\operatorname{div}\boldsymbol{S}(\boldsymbol{D}(\boldsymbol{u}_0)) \in L^2(\Omega)^d$ and that the slightly different regularity assumptions hold:

$$\|\partial_t \boldsymbol{u}\|_{L^\infty(0,T;L^2(\Omega))} + \|\boldsymbol{F}(\boldsymbol{D}(\boldsymbol{u}))\|_{W^{1,2}(Q)} + \|\boldsymbol{F}(\boldsymbol{D}(\boldsymbol{u}))\|_{L^{2((5r-6)/(2-r))}(0,T;W^{1,2}(\Omega))} \leq c.$$

The problem was solved until the final time $T = 0.1$ with the same parameters as above; observe that this choice of parameters guarantees that the required regularity properties are satisfied. Table 3 shows the experimental order of convergence for $r = 1.7$. The order of convergence for the natural norm $\|\boldsymbol{F}(\boldsymbol{D}(\boldsymbol{u}))\|_{L^2(Q)}$ agrees with

Table 3: Experimental order of convergence for the full problem with $r = 1.7$.

h_n	τ_m	$\|\boldsymbol{F}(\boldsymbol{D}(\boldsymbol{u}))\|_{L^2(Q)}$	$\|\boldsymbol{u}\|_{L^\infty(0,T;L^2(\Omega))}$
0.5	0.001	0.9226	1.8703
0.25	0.0005	0.9865	1.9564
0.125	0.00025	1.0057	1.9497
0.0625	0.000125	1.0084	1.9440
0.03125	0.0000625	1.0075	1.9451
Expected		1.0	1.0

the one expected from the theoretical results, while for the velocity we obtain a higher order. This is again likely to be due to the fact that quadratic elements were employed for the velocity variable, while the analysis was performed for linear elements.

4.2. Navier–Stokes/Euler activated fluid. In this section we will consider the classical lid–driven cavity problem with the non–standard constitutive relation:

$$\begin{cases} \begin{cases} \boldsymbol{D} = \delta_s \frac{\boldsymbol{S}}{|\boldsymbol{S}|} + \frac{1}{2\nu}\boldsymbol{S}, & \text{if } |\boldsymbol{D}| \geq \delta_s, \\ \boldsymbol{S} = 0, & \text{if } |\boldsymbol{D}| < \delta_s, \end{cases} & \text{if } (x-\frac{1}{2})^2 + (y-\frac{1}{2})^2 \leq (\frac{3}{8})^2, \\ \boldsymbol{D} = \frac{1}{2\nu}\boldsymbol{S}, & \text{otherwise}, \end{cases} \tag{4.4}$$

where $\nu > 0$ is the viscosity and $\delta_s \geq 0$. This is an example of an activated fluid that in the middle of the domain transitions between a Newtonian fluid (i.e. Navier–Stokes) and an inviscid fluid (i.e. Euler) depending on the magnitude of the symmetric velocity gradient (for a more thorough discussion of activated fluids see [7]). It is analogous to the Bingham constitutive equation for a viscoplastic fluid, but with the roles of the stress and symmetric velocity gradient interchanged; the fact that we can swap the roles of the stress and the symmetric velocity gradient in constitutive relations without any problem is a significant advantage of the framework presented here.

The problem was solved on the unit square $\Omega = (0,1)^2$ with the rest state as the initial condition and with the following boundary conditions:

$$\begin{aligned} &\partial\Omega_1 = (0,1)\times\{1\}, && \partial\Omega_2 := \partial\Omega \setminus \partial\Omega_1, \\ &\boldsymbol{u} = \boldsymbol{0} && \text{on } (0,T)\times\partial\Omega_2, \\ &\boldsymbol{u} = (x^2(1-x)^2 16y^2, 0)^{\mathrm{T}} && \text{on } (0,T)\times\partial\Omega_1. \end{aligned}$$

Although (4.4) has a complicated form, there is a continuous (in $\boldsymbol{D}$) selection available:

$$\boldsymbol{S} = \boldsymbol{\mathcal{S}}(x,y,\boldsymbol{D}) := \begin{cases} 2\nu\left(|\boldsymbol{D}| - \delta_s \mathbb{1}_{B_{3/8}(1/2)}(x,y)\right)^+ \frac{\boldsymbol{D}}{|\boldsymbol{D}|}, & \text{if } |\boldsymbol{D}| \neq 0, \\ \boldsymbol{0}, & \text{if } |\boldsymbol{D}| = 0. \end{cases} \tag{4.5}$$

While the selection stated in (4.5) is already continuous in $\boldsymbol{D}$, Newton's method requires Fréchet-differentiability of $\boldsymbol{\mathcal{S}}$ with respect to $\boldsymbol{D}$ and the constitutive law is not smooth when $|(x-\frac{1}{2}, y-\frac{1}{2})| < \frac{3}{8}$; therefore some regularisation was required for the purpose of applying Newton's method (an alternative would have been to use a non-smooth generalisation such as a semismooth Newton method). For this problem we chose a Papanastasiou-like regularisation (cf. [48]); the Papanastasiou regularisation has been successfully applied to several problems with Bingham rheology [16, 24, 47]. The regularised constitutive relation reads:

$$\boldsymbol{D} = \frac{1}{2\nu}\left(\frac{\delta_s(1-\exp(-M|\boldsymbol{S}|))}{|\boldsymbol{S}|} + 1\right)\boldsymbol{S} \quad \text{for } (x-\tfrac{1}{2})^2 + (y-\tfrac{1}{2})^2 \leq (\tfrac{3}{8})^2, \tag{4.6}$$

where $M > 0$ is the regularisation parameter (as $M \to \infty$ we recover the constitutive relation (4.4), see Figure 1); note that this is not related to the regularisation (2.7), which has the goal of turning the *measurable* selection into a continuous function. For the velocity and pressure we used Scott–Vogelius elements and discontinuous piecewise polynomials were used for the stress (cf. subsection 4.1); the problem was implemented in `firedrake` with $k = 1$, $\nu = \frac{1}{2}$, using the same parameters for the linear and nonlinear solvers described in the previous section, and continuation was employed to reach the values $M = 200$ and $\delta_s = 2.5$; more precisely, the problem was initially solved with $M = 100$ and $\delta_s = 0$ and that solution was used as the Newton guess for the problem with $M+1$ and $\delta_s + 0.05$, repeating the procedure until the desired values were reached. The time step was chosen as $\tau_m = 5 \times 10^{-6}$ and the algorithm was applied until the L^2 norm of the difference of solutions at subsequent time steps was less than 1×10^{-6}.

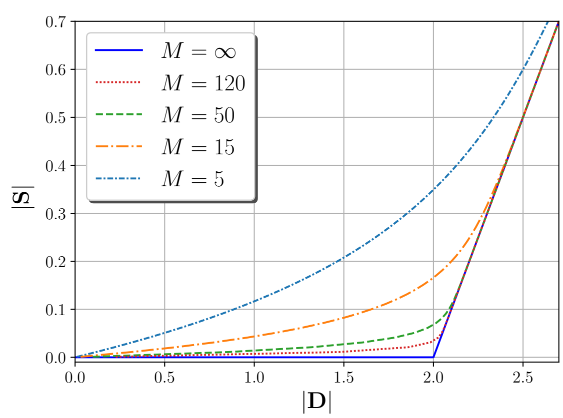

Fig. 1: Regularised constitutive relation for different values of M and $\delta_s = 2$.

Note that when the 'yield strain' parameter δ_s vanishes, we recover the usual Navier–Stokes system. On the other end, if δ_s is taken to be very large this could be taken as an approximation of the incompressible Euler system in the center of the square; notice how in Figure 2 the fluid picks up more speed in the middle of the domain when $\delta_s > 0$ due to the absence of viscosity. This could be an attractive approach to simulating the effects of boundary layers, because it is backed up by a rigorous convergence result; near the boundary the fluid could behave in a Newtonian way and far away δ_s could be taken arbitrarily large so as to make the effects of the viscosity negligible. This is just one of the possibilities that are yet to be explored within this framework of implicitly constituted fluids and mixed formulations and will be studied in more depth in future work.

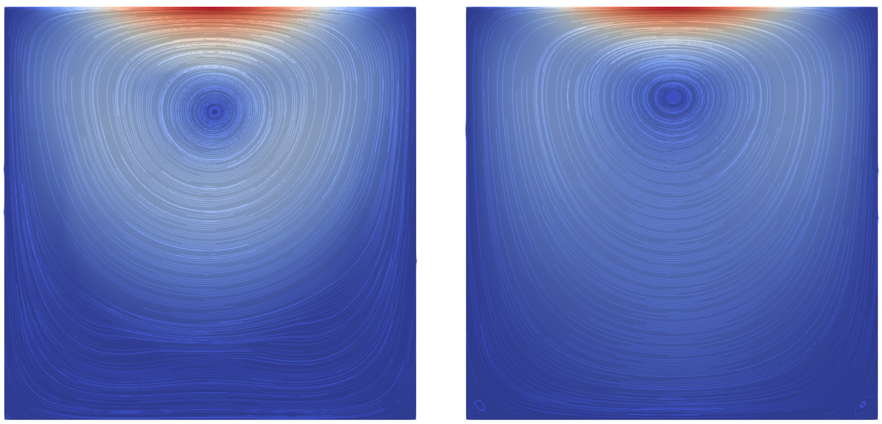

Fig. 2: Streamlines of the steady state for the problem with $\delta_s = 2.5$ (left) and the Newtonian problem (right).

Figure 3 shows the magnitudes of $\boldsymbol{S}$ and $\boldsymbol{D}$ along the line $x = 0.65$ for the steady state of the non-Newtonian problem; it can be clearly seen that the stress is negligibly small for low values of the symmetric velocity gradient in the center of the square and it then suddenly becomes proportional to it. This transition is not the sharpest in the figure because the regularisation parameter M was not taken sufficiently large, but in the limit this would recover the non-smooth relation. In a sense this is similar to solving a Navier–Stokes problem with high Reynolds number, so for high values of M some stabilisation would be required in order to solve this systems efficiently (even more so if the Newtonian fluid outside of the activation region also has a high Reynolds number); this will be the subject of future research.

4.3. Cessation of the Couette flow of a Bingham fluid. The flow between two parallel plates induced by the movement at constant speed of one of the plates receives the name of (plane) Couette flow. It is one of the few examples of a configuration that allows us to find an exact solution for the steady Navier–Stokes equations and it is well known that this solution has a linear profile. In this numerical experiment we will take the Couette flow as the initial condition and investigate the behaviour of the system when the plates stop moving. Physically it is expected that the viscosity and no–slip boundary condition will slow down the flow until it finally stops; it can be seen in [49] that in the Newtonian case the flow does reach the rest state, albeit in infinite time.

This manuscript is for review purposes only.

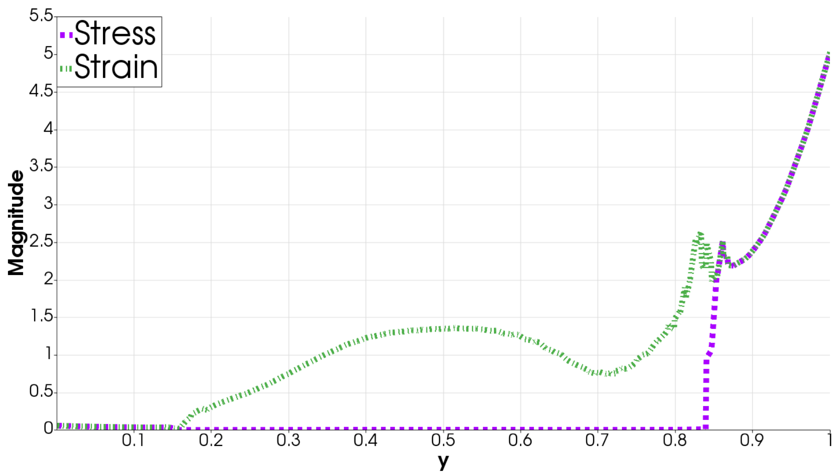

Fig. 3: Magnitude of $\boldsymbol{S}$ and $\boldsymbol{D}$ at $x = 0.65$ for the problem with $\delta_s = 2.5$.

In this section we will solve system (2.11) with the Bingham constitutive relation:

$$\begin{cases} \boldsymbol{S} = \tau_y \frac{\boldsymbol{D}}{|\boldsymbol{D}|} + 2\nu\boldsymbol{D}, & \text{if } |\boldsymbol{S}| \geq \tau_y, \\ \boldsymbol{D} = 0, & \text{if } |\boldsymbol{S}| < \tau_y, \end{cases}$$

where $\nu > 0$ is the viscosity and $\tau_y \geq 0$ is called the yield stress. This is the most common model for a viscoplastic fluid, which is a material that for low stresses (i.e. with a magnitude below the yield stress τ_y) behaves like a solid and like a Newtonian fluid otherwise. Interestingly, viscoplastic fluids in the configuration described above reach the rest state in a finite time and there are theoretical upper bounds for the so called *cessation time* (see [35, 42]), which makes this a good problem to test the numerical algorithm. Just as in the previous section, for this problem there is also a continuous selection available:

$$\boldsymbol{D} = \boldsymbol{\mathcal{D}}(\boldsymbol{S}) := \begin{cases} \frac{1}{2\nu}(|\boldsymbol{S}| - \tau_y)^+ \frac{\boldsymbol{S}}{|\boldsymbol{S}|}, & \text{if } |\boldsymbol{S}| \neq 0, \\ \boldsymbol{0}, & \text{if } |\boldsymbol{S}| = 0. \end{cases} \tag{4.7}$$

For this experiment we again applied the Papanastasiou regularisation to the non-smooth constitutive relation, in order to be able to apply Newton's method. After nondimensionalisation this regularised constitutive law takes the form (compare with (4.6)):

$$\boldsymbol{S}(\boldsymbol{D}) = \left(\frac{Bn}{|\boldsymbol{D}|}(1 - \exp(-M|\boldsymbol{D}|)) + 1 \right) \boldsymbol{D}, \tag{4.8}$$

where $Bn = \frac{\tau_y L}{\nu U}$ is the Bingham number (here U and L are a characteristic velocity and length of the problem, respectively), and $M > 0$ is the regularisation parameter (as $M \to \infty$ we recover the non–smooth relation; compare with Figure 1). The problem was solved on the unit square $\Omega = (0,1)^2$ with the following boundary conditions:

$$\begin{aligned}
\partial\Omega_1 = \{0\} \times (0,1) \cup \{1\} \times (0,1), \qquad \partial\Omega_2 := (0,1) \times \{1\} \cup (0,1) \times \{0\},\\
\boldsymbol{u} = \boldsymbol{0} \qquad \text{on } (0,T) \times \partial\Omega_2,\\
\boldsymbol{u}_\tau = 0 \qquad \text{on } (0,T) \times \partial\Omega_1,\\
-p + \boldsymbol{S}\boldsymbol{n} \cdot \boldsymbol{n} = 0, \quad \text{on } (0,T) \times \partial\Omega_1,
\end{aligned}$$

where $\boldsymbol{u}_\tau$ denotes the component of the velocity tangent to the boundary and $\boldsymbol{n}$ is the unit vector normal to the boundary. The initial condition was taken as a standard Couette flow:

$$\boldsymbol{u}(0, \boldsymbol{x}) = (1 - x_2, 0)^{\mathrm{T}}.$$

For the velocity and pressure we used Taylor–Hood elements and discontinuous piecewise polynomials for the stress. This problem was implemented in FEniCS [45] using the same parameters for the nonlinear and linear solvers described in the previous section, with $k = 1$ and a timestep τ_m between 5×10^{-7} and 1×10^{-6} for the different values of the Bingham number. We quantify the change in the flow through the volumetric flow rate (observe that it is constant in x_1):

$$Q(t) := \int_0^1 (1,0) \cdot \boldsymbol{u}(t, \boldsymbol{x}) \,\mathrm{d}x_2,$$

whose evolution in time is shown in Figure 4 for different values of the Bingham number. An exponential decay of the flow rate is observed in Figure 4, while for positive values of the Bingham number this decay is much faster; these results agree with the ones reported in [42, 16]. In [16] the problem was solved by integrating a one-dimensional equation for u_2; the framework presented here recovers the results obtained there but at the same time has the advantage that it can be applied to a much broader class of problems and geometries.

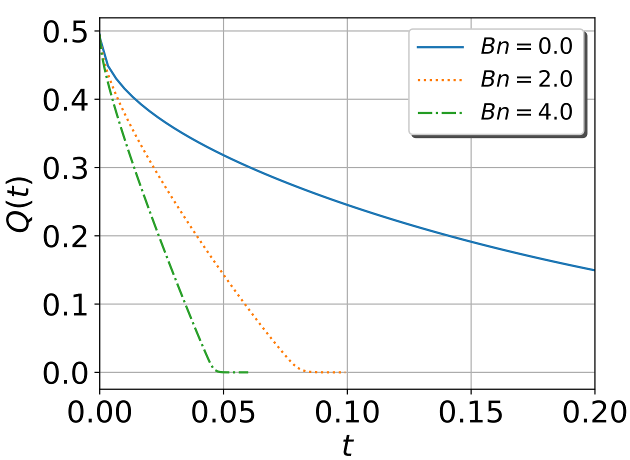

Fig. 4: Evolution of the volumetric flow rate.

5. Conclusions.

In this work we presented a 3-field finite element formulation for the numerical approximation of unsteady implicitly constituted incompressible fluids and identified the necessary conditions that guarantee the convergence of the sequence of numerical approximations to a solution of the continuous problem. Although the convergence analysis was written in terms of a selection $\boldsymbol{\mathcal{D}}$, the finite element formulation presented here can be used in practice with a fully implicit relation; this is in contrast to the works [21, 61], where the algorithms relied on finding an approximate constitutive law expressing the stress $\boldsymbol{S}^k$ in terms of the symmetric velocity gradient $\boldsymbol{D}^k$, which, while always theoretically possible, is not practical for many models. We also presented numerical experiments that showcase the variety of models that the framework of implicitly constituted models can incorporate.

5.1. Acknowledgements.

The second author is grateful for useful discussions with J. Málek, T. Tscherpel, and J. Blechta.

REFERENCES

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

[2] D. N. Arnold and R. Winther, *Mixed finite elements for elasticity*, Numer. Math., 92 (2002), pp. 401–419.

[3] S. Balay, S. Abhyankar, M. F. Adams, J. Brown, P. Brune, K. Buschelman, L. Dalcin, V. Eijhout, W. D. Gropp, D. Kaushik, M. G. Knepley, L. C. McInnes, K. Rupp, B. F. Smith, S. Zampini, H. Zhang, and H. Zhang, PETSc users manual, Tech. Report ANL-95/11-Revision 3.8, Argonne National Laboratory, 2017, http://www.mcs.anl.gov/petsc.

[4] M. A. Behr, L. P. Franca, and T. E. Tezduyar, *Stabilized finite element methods for the velocity-pressure-stress formulation of incompressible flows*, Comput. Methods Appl. Mech. Eng., 104 (1993), pp. 31–48.

[5] L. Belenki, L. Berselli, L. Diening, and M. Růžička, *On the finite element approximation of p-Stokes systems*, SIAM J. Numer. Anal., 50 (2012), pp. 373–397, https://doi.org/10.1137/10080436X.

[6] L. Berselli, L. Diening, and M. Růžička, *Optimal error estimate for semi-implicit space-time discretization for the equations describing incompressible generalized Newtonian fluids*, IMA J. Numer. Anal., 35 (2015), pp. 680–697.

[7] J. Blechta, J. Málek, and K. R. Rajagopal, *On the Classification of Incompressible Fluids and a Mathematical Analysis of the Equations That Govern Their Motion*, (2019), https://arxiv.org/abs/1902.04853.

[8] D. Boffi, F. Brezzi, and M. Fortin, *Mixed Finite Element Methods and Applications*, Springer, 2013.

[9] D. Breit, L. Diening, and S. Schwarzacher, *Solenoidal Lipschitz truncation for parabolic PDEs*, Math. Models Methods Appl. Sci., 23 (2013), pp. 2671–2700.

[10] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, Springer Ser. Comput. Math., 1991.

[11] P. R. Brune, M. G. Knepley, B. F. Smith, and X. Tu, *Composing scalable nonlinear algebraic solvers*, SIAM Rev., 57 (2015), pp. 535–565, https://doi.org/10.1137/130936725

[12] M. Bulíček, P. Gwiazda, J. Malek, K. R. Rajagopal, and A. Świerczewska-Gwiazda, *On flows of fluids described by an implicit constitutive equation characterized by a maximal monotone graph*, vol. 402 of London Math. Soc. Lecture Note Ser, Cambridge Univ. Press: Cambridge, 2012, pp. 23–51.

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

[14] M. Bulíček, P. Gwiazda, J. Málek, and A. Świerczewska-Gwiazda, *On unsteady flows of implicitly constituted incompressible fluids*, SIAM J. Math. Anal., 44 (2012), p. 2756–2801, https://doi.org/10.1137/110830289.

[15] M. Bulíček, J. Málek, V. Průša, and E. Süli, *PDE analysis of a class of thermodynamically compatible viscoelastic rate-type fluids with stress-diffusion*, in Mathematical Analysis in Fluid Mechanics: Selected Recent Results, vol. 710 of AMS Contemporary Mathematics,

2018, pp. 25–51.

[16] M. Chatzimina, G. C. Georgiou, I. Argyropaidas, E. Mitsoulis, and R. R. Huilgol, *Cessation of Couette and Poiseuille flows of a Bingham plastic and finite stopping times*, J. Non-Newtonian Fluid Mech., 129 (2005), pp. 117–127.

[17] P. Clément, *Approximation by finite element functions using local regularization*, RAIRO, Anal. Numér., R2, (1975), pp. 77–84.

[18] M. Crouzeix and P. A. Raviart, *Conforming and nonconforming finite element methods for solving the stationary Stokes equations I*, ESAIM: M2AN, (1973), pp. 33–75.

[19] T. A. Davis, Algorithm 832: UMFPACK V4.3—an unsymmetric-pattern multifrontal method, ACM Trans. Math. Softw., 30(2), pp. 196-199, 2004.

[20] E. DiBenedetto, *Degenerate Parabolic Equations*, Springer Ver., 1993.

[21] L. Diening, D. Kreuzer, and E. Süli, *Finite element approximation of steady flows of incompressible fluids with implicit power-law-like rheology*, SIAM J. Numer. Anal., 51 (2013), pp. 984–1015, https://doi.org/10.1137/120873133.

[22] L. Diening, M. Ružička, and J. Wolf, *Existence of weak solutions for unsteady motions of generalized Newtonian fluids*, Ann. Scuola Norm. Sup. Pisa Cl. Sci., IX (2010), pp. 1–46.

[23] L. Diening, S. Schwarzacher, V. Stroffolini, and A. Verde, *Parabolic Lipschitz truncation and caloric approximation*, Calc. Var., 56:120 (2017).

[24] Y. Dimakopoulos and J. Tsamopoulos, *Transient displacement of a viscoplastic material by air in straight and suddenly constricted tubes*, J. Non-Newtonian Fluid Mech., 112 (2003), pp. 43–75.

[25] S. Eckstein and M. Růžička, *On the full space-time discretization of the generalized Stokes equations: The Dirichlet case*, SIAM J. Numer. Anal., 56 (2018), pp. 2234–2261, https://doi.org/10.1137/16M1099741.

[26] V. J. Ervin, J. S. Howell, and I. Stanculescu, *A dual-mixed approximation method for a three-field model of a nonlinear generalized Stokes problem*, Comput. Methods Appl. Mech. Eng., 197 (2008), pp. 2886–2900.

[27] M. Farhloul and M. Fortin, *New mixed finite element for the Stokes and elasticity problems*, SIAM J. Numer. Anal., 30 (1993), pp. 971–990, https://doi.org/10.1137/0730051.

[28] M. Farhloul and H. Manouzi, *Analysis of non-singular solutions of a mixed Navier-Stokes formulation*, Comput. Methods Appl. Mech. Eng., 129 (1996), pp. 115–131.

[29] M. Farhloul, S. Nicaise, and L. Paquet, *A refined mixed finite-element method for the stationary Navier-Stokes equations with mixed boundary conditions*, IMA J. Numer. Anal., 28 (2008), pp. 25–45.

[30] M. Farhloul, S. Nicaise, and L. Paquet, *A priori and a posteriori error estimations for the dual mixed finite element method of the Navier-Stokes problem*, Numer. Meth. Part. Differ. Equat., 25 (2009), pp. 843–869.

[31] G. P. Galdi, *An Introduction to the Mathematical Theory of the Navier-Stokes Equations: Steady State Problems*, Springer, Second Edition, 2011.

[32] C. Geuzaine and J. F. Remacle, Gmsh: a three-dimensional finite element mesh generator with built-in pre- and post-processing facilities, 79(11), Int. J. Numer. Meth. Eng., pp. 1309-1331, 2009.

[33] V. Girault and P. A. Raviart, *Finite Element Methods for Navier-Stokes Equations: Theory and Algorithms*, Springer Verlag, 1986.

[34] V. Girault and L. R. Scott, *A quasi-local interpolation operator preserving the discrete divergence*, Calcolo, (2003), pp. 1–19.

[35] R. Glowinski, *Numerical Methods for Nonlinear Variational Problems*, Springer–Verlag, New York, 1984.

[36] J. Guzmán and M. Neilan, *Conforming and divergence-free Stokes elements in three dimensions*, IMA J. Numer. Anal., (2014), pp. 1489–1508.

[37] J. Guzmán and M. Neilan, *Conforming and divergence-free Stokes elements on general triangular meshes*, Math. Comput., 83 (2014), pp. 15–36.

[38] A. Hirn, *Approximation of the p-Stokes equations with equal-order finite elements*, J. Math. Fluid Mech, 15 (2013), pp. 65–88.

[39] J. S. Howell, *Dual-mixed finite element approximation of Stokes and nonlinear Stokes problems using trace-free velocity gradients*, J. Comput. Appl. Math, 231 (2009), pp. 780–792.

[40] J. S. Howell and N. J. Walkington, *Dual-mixed finite element methods for the Navier-Stokes equations*, ESAIM: M2AN, 47 (2013), pp. 789–805.

[41] J. Hron, J. Málek, J. Stebel, and K. Touška, *A novel view on computations of steady flows of Bingham fluids using implicit constitutive relations*, Project MORE Preprint, (2017).

[42] R. R. Huilgol, B. Mena, and J. M. Piau, *Finite stopping time problems and rheometry of*

Bingham fluids, J. Non-Newtonian Fluid Mech., 102 (2002), pp. 97–107.

[43] C. Kreuzer and E. Süli, *Adaptive finite element approximation of steady flows of incompressible fluids with implicit power-law-like rheology*, ESAIM: M2AN, 50 (2016), pp. 1333–1369.

[44] J. L. Lions, *Quelques Méthodes De Résolution Des Problèmes Aux Limites Non Linéaires*, Dunod, Paris, 1969.

[45] A. Logg, K. A. Mardal, and G. N. Wells, *FEniCS : Automated Solution of Differential Equations by the Finite Element Method*, Springer, 2011, pp. 337–359, https://doi.org/10.1007/978-3-642-23099-8.

[46] E. Maringová and J. Žabenský, *On a Navier–Stokes–Fourier-like system capturing transitions between viscous and inviscid fluid regimes and between no-slip and perfect-slip boundary conditions*, Nonlinear Anal. Real World Appl., 41 (2018), pp. 152–178.

[47] E. Mitsoulis and R. R. Huilgol, *Entry flows of Bingham plastics in expansions*, J. Non-Newtonian Fluid Mech., 122 (2004), pp. 45–54.

[48] T. C. Papanastasiou, *Flows of materials with yield*, J. Rheol, 31 (1987), pp. 385–404.

[49] T. C. Papanastasiou, G. Georgiou, and A. Alexandrou, *Viscous Fluid Flow*, CRC Press, Boca Raton, 1999.

[50] V. Průša and K. R. Rajagopal, *A new class of models to describe the response of electrorheological and other field dependent fluids*, Adv. Struct. Mater., 89 (2018), pp. 655–673.

[51] K. R. Rajagopal, *On implicit constitutive theories*, Appl. Math., 48 (2003), pp. 279–319.

[52] K. R. Rajagopal, *On implicit constitutive theories for fluids*, J. Fluid Mech., 550 (2006), pp. 243–249.

[53] K. R. Rajagopal, *The elasticity of elasticity*, Z. Angew. Math. Phys, 5807 (2007), pp. 309–317.

[54] K. R. Rajagopal and A. R. Srinivasa, *On the thermodynamics of fluids defined by implicit constitutive relations*, Z. Angew. Math. Phys, 59 (2008), pp. 715–729.

[55] F. Rathgeber, D. A. Ham, L. Mitchell, M. Lange, F. Luporini, A. T. Mcrae, G.-T. Bercea, G. R. Markall, and P. H. Kelly, *Firedrake: automating the finite element method by composing abstractions*, ACM Trans. Math. Softw., 43 (2016).

[56] T. Roubíček, *Nonlinear Partial Differential Equations with Applications*, Birkhäuser, second ed., 2013.

[57] V. Ruas, *An optimal three-field finite element approximation of the Stokes system with continuous extra stresses*, Japan J. Appl. Math., (1994), pp. 113-130.

[58] D. Sandri, *A posteriori estimators for mixed finite element approximations of a fluid obeying the power law*, Comput. Methods Appl. Mech. Eng., (1998), pp. 329–340.

[59] L. R. Scott and M. Vogelius, *Norm estimates for a maximal right inverse of the divergence operator in spaces of piecewise polynomials*, Math. Modelling Numer. Anal., 19 (1985), pp. 111–143.

[60] J. Simon, *Compact sets in the space* $L^p(0,T;B)$, Ann. Mat. Pura Appl., 4 (1987), pp. 65–96.

[61] E. Süli and T. Tscherpel, *Fully discrete finite element approximation of unsteady flows of implicitly constituted incompressible fluids*, IMA J. Numer. Anal., (2018), pp. 1–49.

[62] T. Tscherpel, *FEM for the Unsteady Flow of Implicitly Constituted Incompressible Fluids*, PhD thesis, University of Oxford, 2018.

This manuscript is for review purposes only.