IMA Journal of Numerical Analysis (2020) **40**, 801–849
doi:10.1093/imanum/dry097
Advance Access publication on 15 February 2019

Fully discrete finite element approximation of unsteady flows of implicitly constituted incompressible fluids

Endre Süli
Mathematical Institute, University of Oxford, Woodstock Road, Oxford, OX2 6GG, UK
endre.suli@maths.ox.ac.uk

and

Tabea Tscherpel*
Institut für Geometrie und Praktische Mathematik, RWTH Aachen, Templergraben 55, 52056 Aachen, Germany
*Corresponding author: tscherpel@igpm.rwth-aachen.de

[Received on 16 April 2018; revised on 6 December 2018]

Implicit constitutive theory provides a very general framework for fluid flow models, including both Newtonian and generalized Newtonian fluids, where the Cauchy stress tensor and the rate of strain tensor are assumed to be related by an implicit relation associated with a maximal monotone graph. For incompressible unsteady flows of such fluids, subject to a homogeneous Dirichlet boundary condition on a Lipschitz polytopal domain $\Omega \subset \mathbb{R}^d$, $d \in \{2, 3\}$, we investigate a fully discrete approximation scheme, using a spatial mixed finite element approximation on general shape-regular simplicial meshes combined with backward Euler time-stepping. We consider the case when the velocity field belongs to the space of solenoidal functions contained in $\mathrm{L}^\infty(0, T; \mathrm{L}^2(\Omega)^d) \cap \mathrm{L}^q(0, T; \mathrm{W}^{1,q}_0(\Omega)^d)$ with $q \in (2d/(d+2), \infty)$, which is the maximal range of q with respect to existence of weak solutions. In order to facilitate passage to the limit with the discretization parameters for the sub-range $q \in (2d/(d+2), (3d+2)/(d+2))$, we introduce a regularization of the momentum equation by means of a penalty term, and first show convergence of a subsequence of approximate solutions to a weak solution of the regularized problem; we then pass to the limit with the regularization parameter. This is achieved by the use of a solenoidal parabolic Lipschitz truncation method, a local Minty-type monotonicity result, and various weak compactness techniques. For $q \geq (3d+2)/(d+2)$ convergence of a subsequence of approximate solutions to a weak solution can be shown directly, without the regularization term.

Keywords: finite element method; time-stepping; implicit constitutive models; convergence; weak compactness; Lipschitz truncation method.

1. Introduction

In the mechanics of viscous incompressible fluids, typical constitutive relations relate the shear stress tensor to the rate of strain tensor through an explicit functional relationship. In the case of a Newtonian fluid the relationship is linear, and in the case of generalized Newtonian fluids it is usually a power-law-like nonlinear, but still explicit, functional relation. Implicit constitutive theory was introduced in order to describe a wide range of non-Newtonian rheology, by admitting implicit and discontinuous constitutive laws, see Rajagopal (2003, 2008). The existence of weak solutions to mathematical models of this kind was explored in Bulíček *et al.* (2009, 2012) for steady and unsteady flows, respectively.

The aim of the present paper is to construct a fully discrete numerical approximation scheme, in the unsteady case, for a class of such implicitly constituted models, where the shear stress and the rate of strain tensors are related through a (possibly discontinuous) maximal monotone graph. The scheme is based on a spatial mixed finite element approximation and a backward Euler discretization with respect to the temporal variable.

We will show weak convergence (up to subsequences) of the sequence of approximate solutions to a weak solution of the regularized problem, and then weak convergence (up to subsequences) of the sequence of weak solutions of the regularized problem to a weak solution of the original problem.

The mathematical ideas contained in the paper are motivated by the existence theory formulated, in the unsteady case, in Bulíček *et al.* (2012), and the convergence theory for finite element approximations of steady implicitly constituted fluid flow models developed in Diening *et al.* (2013).

1.1 *Implicit constitutive law*

Statement of the problem. Let $\Omega \subset \mathbb{R}^d$ with $d \geq 2$ be a bounded Lipschitz domain and denote by $Q = (0,T) \times \Omega$ the parabolic cylinder for a given final time $T \in (0,\infty)$. Furthermore, let $\boldsymbol{f} \colon Q \to \mathbb{R}^d$ be a given external force and let $\boldsymbol{u}_0 \colon \Omega \to \mathbb{R}^d$ be an initial velocity field. We seek a velocity field $\boldsymbol{u} \colon \overline{Q} \to \mathbb{R}^d$, a pressure $\pi \colon Q \to \mathbb{R}$ and a trace-free stress tensor field $\mathbf{S} \colon Q \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$ satisfying the balance law of linear momentum and the incompressibility condition:

$$\begin{aligned} \partial_t \boldsymbol{u} + \operatorname{div}(\boldsymbol{u} \otimes \boldsymbol{u}) - \operatorname{div} \mathbf{S} &= -\nabla \pi + \boldsymbol{f} \qquad && \text{on } Q, \\ \operatorname{div} \boldsymbol{u} &= 0 && \text{on } Q, \end{aligned} \tag{1.1}$$

subject to the following initial condition and no-slip boundary condition:

$$\boldsymbol{u}(0,\cdot) = \boldsymbol{u}_0(\cdot) \qquad \text{in } \Omega, \tag{1.2}$$

$$\boldsymbol{u} = \mathbf{0} \qquad \text{on } (0,T) \times \partial\Omega. \tag{1.3}$$

In order to close the system we need to impose a relation, the so-called *constitutive law*,

$$\mathbf{G}(\cdot, \mathbf{D}\boldsymbol{u}, \mathbf{S}) = \mathbf{0}, \tag{1.4}$$

between the stress tensor $\mathbf{S}$ and the symmetric gradient $\mathbf{D}\boldsymbol{u} = \frac{1}{2}(\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^\top)$, which represents the shear rate of the fluid. In the following we will refer to the problem consisting of (1.1)–(1.4) as **(P)**.

The relation $\mathbf{G}$ may be fully implicit, and we assume that $\mathbf{G}$ can be identified with a maximal monotone graph $\mathcal{A}(z) \subset \mathbb{R}^{d\times d}_{\mathrm{sym}} \times \mathbb{R}^{d\times d}_{\mathrm{sym}}$, for $z \in Q$, as

$$\mathbf{G}\big(z, \mathbf{D}\boldsymbol{u}(z), \mathbf{S}(z)\big) = \mathbf{0} \quad \Leftrightarrow \quad \big(\mathbf{D}\boldsymbol{u}(z), \mathbf{S}(z)\big) \in \mathcal{A}(z),$$

where $\mathcal{A}(\cdot)$ satisfies the following assumption, similarly as in Bulíček *et al.* (2009, p. 110) and Bulíček *et al.* (2012, Sec. 1.2).

ASSUMPTION 1.1 (Properties of $\mathcal{A}(\cdot)$). We assume that the mapping $Q \ni z \mapsto \mathcal{A}(z) \subset \mathbb{R}^{d\times d}_{\mathrm{sym}} \times \mathbb{R}^{d\times d}_{\mathrm{sym}}$ satisfies the following conditions for a.e. $z \in Q$:

(A1) $(\mathbf{0}, \mathbf{0}) \in \mathcal{A}(z)$;

(A2) $\mathcal{A}(z)$ is a monotone graph, i.e., for all $(\mathbf{D}_1, \mathbf{S}_1), (\mathbf{D}_2, \mathbf{S}_2) \in \mathcal{A}(z)$,

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

(A3) $\mathcal{A}(z)$ is a maximal monotone graph, i.e., $(\mathbf{D}, \mathbf{S}) \in \mathbb{R}^{d\times d}_{\mathrm{sym}} \times \mathbb{R}^{d\times d}_{\mathrm{sym}}$ and

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

implies that $(\mathbf{D}, \mathbf{S}) \in \mathcal{A}(z)$;

(A4) There exists a constant $c_* > 0$, a nonnegative function $g \in \mathrm{L}^1(Q)$ and $q \in (1, \infty)$ such that

$$\mathbf{D} : \mathbf{S} \geq -g(z) + c_* \big(|\mathbf{D}|^q + |\mathbf{S}|^{q'}\big) \quad \text{for all } (\mathbf{D}, \mathbf{S}) \in \mathcal{A}(z) \text{ and for a.e. } z \in Q,$$

where q' is the Hölder conjugate of q.

(A5) For any $(\mathbf{D}, \mathbf{S}) \in \mathcal{A}(z)$ we have that

$$\mathrm{tr}(\mathbf{D}) = 0 \;\Leftrightarrow\; \mathrm{tr}(\mathbf{S}) = 0;$$

(A6) $z \mapsto \mathcal{A}(z)$ is $\mathcal{L}(Q) - (\mathcal{B}(\mathbb{R}^{d\times d}_{\mathrm{sym}}) \otimes \mathcal{B}(\mathbb{R}^{d\times d}_{\mathrm{sym}}))$ measurable, where $\mathcal{L}(Q)$ denotes the set of all Lebesgue measurable subsets of Q and $\mathcal{B}(\mathbb{R}^{d\times d}_{\mathrm{sym}})$ denotes the set of all Borel subsets of $\mathbb{R}^{d\times d}_{\mathrm{sym}}$.

REMARK 1.2 (Properties of $\mathcal{A}(\cdot)$).

(i) In Bulíček *et al.* (2012) the authors phrase the condition (A4) in the more general context of Orlicz–Sobolev spaces. Here we will restrict ourselves to the usual Sobolev setting.

(ii) Condition (A5) is added to the list of assumptions contained in Bulíček *et al.* (2009, 2012), Diening *et al.* (2013) and Kreuzer & Süli (2016) to ensure consistency with the thermodynamic framework for incompressible fluids. If the velocity function $\boldsymbol{u}$ is (pointwise) divergence-free, then we have that $\mathrm{tr}(\mathbf{D}\boldsymbol{u}) = \mathrm{div}(\boldsymbol{u}) = 0$. Thus, if $(\mathbf{D}\boldsymbol{u}(z), \mathbf{S}(z)) \in \mathcal{A}(z)$ for a.e. $z \in Q$, then condition (A5) implies that $\mathbf{S}$ is (pointwise) trace-free and this condition need not be imposed separately. In the context of finite element approximations, we cannot simply consider $\mathcal{A}(\cdot)$ as subset of the Cartesian product of the linear space of trace-free $d \times d$ symmetric matrices with itself, since the finite element velocity fields need not be exactly divergence-free.

(iii) A sufficient condition for (A6) to be satisfied is given in Bulíček *et al.* (2009, (A5)(ii), p. 110).

This framework covers explicit relations, including Newtonian fluids, where $q = 2$, and q-fluids describing shear-thinning and shear-thickening behaviour, for $1 < q < 2$ and $q > 2$, respectively. Also, relations, where the stress is a set-valued or discontinuous function of the symmetric gradient, as for Bingham and Herschel–Bulkley fluids, are included, cf. Bulíček *et al.* (2012, Lem. 1.1) for $\mathcal{A}(\cdot)$

restricted to trace-free matrices. Such an $\mathcal{A}(\cdot)$ can be extended to non-trace-free matrices, see Tscherpel (2018). Furthermore, fully implicitly constituted fluids are covered and the constitutive relation is allowed to depend on $z = (t, \boldsymbol{x}) \in Q$.

1.2 *Overview of the context*

The first existence results for explicit constitutive laws were obtained in Ladyženskaja (1969) and Lions (1969) using monotone operator theory for the range of q, for which weak solutions are admissible test functions. Subsequently, the range of q was gradually extended by means of truncations, which made it possible to overcome the admissibility problem for small q caused by the presence of the convective term: first, an L^∞-truncation was developed, see Frehse *et al.* (1997) and Růžička (1997) for the steady case, and Wolf (2007) for the unsteady case; then, a refinement of the Lipschitz truncation method, originating in the work of Acerbi & Fusco (1988), allowed to cover the full range of $q \in \left(\frac{2d}{d+2}, \infty\right)$. The existence of weak solutions for the whole range was proved in Frehse *et al.* (2003), in the steady case, and in Diening *et al.* (2010), in the unsteady case. The restriction on q is required to ensure compactness of the embedding $\mathrm{W}^{1,q}(\Omega) \hookrightarrow \mathrm{L}^2(\Omega)$, which is needed in the convective term for the passage to the limit in the sequence of approximate solutions. Under suitable conditions, existence of strong solutions is available for explicit constitutive laws at least for short intervals of time, see Berselli *et al.* (2010). Based on regularity results, a number of contributions deal with error estimates for strong solutions, a recent one of which is Berselli *et al.* (2015), showing optimal convergence rates for $d = 3$ and $q \in \left(\frac{3}{2}, 2\right]$.

For implicitly constituted fluids, the existence of weak solutions for $q > \frac{2d}{d+2}$ for steady and unsteady flows was proved in Bulíček *et al.* (2009, 2012), generalizing previous results on discontinuous constitutive relations, see Duvaut & Lions (1976), Serëgin (1994), Fuchs & Seregin (2000), Málek *et al.* (2005) and Eberlein & Růžička (2012). In Bulíček *et al.* (2012), a Navier slip boundary condition and $\mathrm{C}^{1,1}$ regularity of $\partial\Omega$ were assumed to avoid technicalities related to lack of regularity of the pressure in the unsteady case. Due to the weak structural assumptions, the existence of short-time strong solutions and uniqueness cannot be expected to hold in general. The proof in Bulíček *et al.* (2012) is constructive and is based on a three-level approximation using finite-dimensional Galerkin subspaces spanned by eigenfunctions of higher order elliptic operators. These Galerkin spaces are not available for practical computations, and therefore we take an alternative route in the construction of a numerical method for the problem and for its convergence analysis. Here we shall consider a mixed finite element approximation under minimal regularity hypotheses; hence, we can only hope for qualitative convergence results rather than quantitative error bounds in terms of the spatial and temporal discretization parameters. The approximation scheme will be constructed for a regularized version of the equations, including a penalty term, and, after passing to the limit with the discretization parameters, we shall pass to the limit with the regularization parameter.

Concerning the numerical analysis of implicitly constituted fluid flow models, to the best of our knowledge, the only results available are those contained in Diening *et al.* (2013) and Kreuzer & Süli (2016), which deal with the steady case under the additional assumption that $\mathcal{A}(\cdot)$ is (generalized) strictly monotone. By means of a discrete Lipschitz truncation method and various weak compactness results, the authors of Diening *et al.* (2013) prove the convergence of a large class of mixed finite element methods for $q > \frac{2d}{d+1}$ for discretely divergence-free finite element functions for the velocity, and for $q > \frac{2d}{d+2}$ for exactly (i.e., pointwise) divergence-free finite element functions for the velocity field. In the case of discretely divergence-free mixed finite element approximations, the more demanding requirement $q > \frac{2d}{d+1}$ arises from the (numerical) modification of the trilinear form associated with the convective term in the weak formulation of the problem. The purpose of this trilinear form is to reinstate the skew-symmetry of the trilinear form, lost in the course of the spatial approximation. In Kreuzer & Süli (2016) an *a posteriori* analysis is performed for implicitly constituted fluid flow models, using discretely divergence-free finite element functions also reproving convergence in this case, but under stronger assumptions on the sequence of graph approximations.

In the unsteady case, no convergence result is available for numerical approximations of implicitly constituted fluid flow models, and even those contributions that are focussed on qualitative convergence results for explicit constitutive laws assume additional restrictions on q. In Carelli *et al.* (2010) convergence to a weak solution of a regularized problem is proved for continuous $q(\boldsymbol{x})$ and, subsequently, the regularization limit is taken under the assumption that q is constant and $q > \frac{2(d+1)}{d+2}$, a limitation that arises from the use of the L^{∞}-truncation developed in Wolf (2007).

For the special cases of Bingham and Herschel–Bulkley fluids, a number of contributions devoted to numerical simulations are available in the literature, see, for example, Bercovier & Engelman (1980), Zhang (2010), Faria & Karam-Filho (2013), Moreno *et al.* (2016), Mahmood *et al.* (2017) and the survey article, Dean *et al.* (2007). We highlight, in particular, the numerical experiments for Bingham fluids in Hron *et al.* (2017) by means of various mixed finite element approximations, motivated by implicit constitutive theory. However, the lack of rigorous numerical analysis in the unsteady case is apparent. The purpose of the paper is therefore to provide a convergence proof for sequences of numerical approximations for a large class of unsteady implicitly constituted fluid flow models, which includes, in particular, the Bingham and Herschel–Bulkley models.

1.3 *Aim and main result*

Our objective is to establish a convergence result for implicitly constituted fluids in the unsteady case for the whole range $q > \frac{2d}{d+2}$. The main challenges concern the implicit, possibly discontinuous, relation between the stress and the shear rate and the lack of admissibility in the convective term for small exponents q.

Hence, additionally to a fully discrete approximation, we introduce two regularizations: the first approximates the potentially multi-valued function, the graph of which is $\mathcal{A}(\cdot)$, by a sequence of single-valued functions; the second improves the integrability of the velocity approximations by means of a penalty term, so that weak solutions of the regularized problem are admissible test functions in the weak form of the convective term. The use of a penalty term is only required for the sub-range $q \in \left(\frac{2d}{d+2}, \frac{3d}{d+2}\right)$. More specifically, we introduce the following three-level approximation:

$k \in \mathbb{N}$: graph approximation;

$l, n \in \mathbb{N}$: discretization in space and in time;

$m \in \mathbb{N}$: regularization by a penalty term in the equation,

the main technical novelty of the paper being in the passage to the discretization limits $l, n \to \infty$.

The main contribution of the paper is the following. Let Ω be a Lipschitz polytopal domain, $q > \frac{2d}{d+2}$, and assume that we have a pair of inf-sup stable finite element spaces for the velocity and the pressure. Also, we assume that a suitable approximation of the graph $\mathcal{A}$ is available, examples of which will be constructed below. Then, a sequence of approximate solutions to the fully discrete problem exists and the corresponding sequence of finite element approximations converges weakly, up to subsequences, to a weak solution of problem **(P)**, when first taking the graph approximation limit, then the spatial and temporal discretization limits, and finally the regularization limit. The precise formulation of this result is contained in Theorem 4.1, and the notion of weak solution is given in Definition 2.1. Important tools in the proof are a local Minty-type convergence result established in Bulíček *et al.* (2012) and Bulíček & Málek (2016), and the solenoidal parabolic Lipschitz truncation constructed in Breit *et al.* (2013) to overcome the admissibility problem for small q.

The paper is structured as follows. Section 2 provides the analytical setting, including the graph approximation and the Lipschitz truncation. Section 3 describes the finite element approximation, the approximation of the convective term and the time-stepping. In Section 4 we first introduce the approximation levels in detail before giving the convergence proof.

2. Analytical preliminaries

By $\mathbb{R}^{d\times d}_{\mathrm{sym}}$ we denote the set of all real-valued symmetric $d \times d$-matrices and we use $:$ for the Frobenius scalar product in $\mathbb{R}^{d\times d}$. For $\omega \subset \mathbb{R}^d$ we denote by $|\omega|$ the d-dimensional Lebesgue measure of ω. By $\mathbb{1}_\omega$ we denote the characteristic function of the set ω. For the (distributional) partial derivatives with respect to time, we use the shorthand notation $\partial_t f := \frac{\partial f}{\partial t}$.

For $\omega \subset \mathbb{R}^d$ open and $p \in [1,\infty)$ let $(\mathrm{L}^p(\omega), \|\cdot\|_{\mathrm{L}^p(\omega)})$ be the standard Lebesgue space of p-integrable functions, and the space of essentially bounded functions when $p = \infty$. For $s \in \mathbb{N}$ let $(\mathrm{W}^{s,p}(\omega), \|\cdot\|_{\mathrm{W}^{s,p}(\omega)})$ be the respective Sobolev spaces. For spaces of vector-valued and tensor-valued functions, we use superscripts d and $d \times d$, respectively (except for in norms). By $\mathrm{L}^p_0(\omega)$ we denote the set of functions in $\mathrm{L}^p(\omega)$ with zero mean integral.

For a general Banach space $(\mathrm{X}, \|\cdot\|_{\mathrm{X}})$, the dual space consisting of all continuous linear functionals on X is denoted by X', and the dual pairing is denoted by $\langle f, g\rangle_{\mathrm{X}',\mathrm{X}}$, if $f \in \mathrm{X}'$ and $g \in \mathrm{X}$. If X is a space of functions defined on ω, then we denote the dual pairing by $\langle f, g\rangle_\omega := \langle f, g\rangle_{\mathrm{X}',\mathrm{X}}$, in case the space X is known from the context. We also use this notation for the integral of the scalar product $f \cdot g$ of two functions f and g, provided that $f \cdot g \in \mathrm{L}^1(\omega)$. Furthermore, if $\omega \subset \mathbb{R}^d$ is measurable and $0 < |\omega| < \infty$, then we denote $⨍_\omega f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} := \frac{1}{|\omega|}\int_\omega f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$.

For a bounded open domain $\Omega \subset \mathbb{R}^d$ and $T \in (0,\infty)$, let $Q = (0,T) \times \Omega$. Denote by $\mathrm{C}^\infty_0(\Omega)$ the set of all smooth and compactly supported functions on Ω and by $\mathrm{C}^\infty_{0,\mathrm{div}}(\Omega)^d$ the set of all functions in $\mathrm{C}^\infty_0(\Omega)^d$ with vanishing divergence. Analogously, define $\mathrm{C}^\infty_0(Q)$ and $\mathrm{C}^\infty_{0,\mathrm{div}}(Q)^d$. We define $\mathrm{W}^{1,p}_0(\Omega) := \overline{\mathrm{C}^\infty_0(\Omega)}^{\|\cdot\|_{\mathrm{W}^{1,p}(\Omega)}}$, for $p \in [1,\infty)$ and $\mathrm{W}^{1,\infty}_0(\Omega) := \mathrm{W}^{1,1}_0(\Omega) \cap \mathrm{W}^{1,\infty}(\Omega)$. For a given $p \in (1,\infty)$, we let the Hölder exponent p' be defined by $\frac{1}{p} + \frac{1}{p'} = 1$. Then, if $p \in (1,\infty)$, $\mathrm{L}^{p'}(\Omega)$ is the dual space of $\mathrm{L}^p(\Omega)$ and $\mathrm{W}^{-1,p'}(\Omega)$ will denote the dual space of $\mathrm{W}^{1,p}_0(\Omega)$. Further, we define the spaces of divergence-free functions: the spaces $\mathrm{L}^2_{\mathrm{div}}(\Omega)^d$ and $\mathrm{W}^{1,p}_{0,\mathrm{div}}(\Omega)^d$, for $p \in [1,\infty)$, are the closures of $\mathrm{C}^\infty_{0,\mathrm{div}}(\Omega)^d$ with respect to the norms $\|\cdot\|_{\mathrm{L}^2(\Omega)}$ and $\|\cdot\|_{\mathrm{W}^{1,p}(\Omega)}$, respectively, and let $\mathrm{W}^{1,\infty}_{0,\mathrm{div}}(\Omega)^d := \mathrm{W}^{1,1}_{0,\mathrm{div}}(\Omega)^d \cap \mathrm{W}^{1,\infty}(\Omega)^d$.

Let $\mathrm{C}(\overline{\Omega})$ be the set of all continuous real-valued functions on $\overline{\Omega}$. With $\mathrm{C}([0,T];\mathrm{X})$ we denote the set of all functions defined on $[0,T]$, taking values in a Banach space X, which are continuous (with respect to the strong topology in X). Similarly, $\mathrm{C}^{0,1}([0,T];\mathrm{X})$ is the space of all Lipschitz continuous functions defined on $[0,T]$, with values in X. Furthermore, we define the space of weakly continuous functions with values in X by

$$\mathrm{C}_w([0,T];\mathrm{X}) := \{\boldsymbol{v}\colon [0,T] \to \mathrm{X}\colon\ t \mapsto \langle \boldsymbol{w}, \boldsymbol{v}(t,\cdot)\rangle_{\mathrm{X}',\mathrm{X}} \in \mathrm{C}([0,T]),\ \forall \boldsymbol{w} \in \mathrm{X}'\}.$$

We denote by $\mathrm{L}^p(0,T;\mathrm{X})$ the standard Bochner space of p-integrable X-valued functions. We use the notation *ess* $lim_{t\to 0_+} f(t)$ to indicate that there exists a zero set $N(f) \subset [0,T]$ such that $t \in (0,T)\backslash N(f)$, when considering the limit of $f(t)$, as $t \to 0_+$.

In the following, $c > 0$ will denote a generic constant, which can change from line to line and depends only on the given data unless specified otherwise.

For the regularized problem we shall require the following function spaces, with associated norms:

$$\mathrm{X}(\Omega) := \mathrm{W}_0^{1,q}(\Omega)^d \cap \mathrm{L}^{2q'}(\Omega)^d, \qquad \|\cdot\|_{\mathrm{X}(\Omega)} := \|\cdot\|_{\mathrm{W}^{1,q}(\Omega)} + \|\cdot\|_{\mathrm{L}^{2q'}(\Omega)}, \tag{2.1}$$

$$\mathrm{X}(Q) := \mathrm{L}^q(0,T;\mathrm{W}_0^{1,q}(\Omega)^d) \cap \mathrm{L}^{2q'}(Q)^d, \quad \|\cdot\|_{\mathrm{X}(Q)} := \|\cdot\|_{\mathrm{L}^q(0,T;\mathrm{W}^{1,q}(\Omega))} + \|\cdot\|_{\mathrm{L}^{2q'}(Q)}, \tag{2.2}$$

and consider their solenoidal subspaces, denoted by $\mathrm{X}_{\mathrm{div}}(\Omega)$ and $\mathrm{X}_{\mathrm{div}}(Q)$, respectively. We note in passing that we shall refer to the sequential version of the Banach–Alaoglu theorem simply as Banach–Alaoglu theorem.

Weak solutions. In what follows, let $\Omega \subset \mathbb{R}^d$, with $d \in \{2,3\}$, be a bounded Lipschitz domain and for $T \in (0,\infty)$ denote $Q = (0,T) \times \Omega$. Furthermore, assume that $q \in (1,\infty)$ is given and let $\mathcal{A}(\cdot) \subset \mathbb{R}^{d\times d}_{\mathrm{sym}} \times \mathbb{R}^{d\times d}_{\mathrm{sym}}$ be a monotone graph satisfying Assumption 1.1 with respect to q.

DEFINITION 2.1 (Weak solution). For a given $\boldsymbol{u}_0 \in \mathrm{L}^2_{\mathrm{div}}(\Omega)^d$ and $\boldsymbol{f} \in \mathrm{L}^{q'}(0,T;\mathrm{W}^{-1,q'}(\Omega)^d)$, we call $(\boldsymbol{u},\mathbf{S})$ a weak solution to problem **(P)** if

$$\boldsymbol{u} \in \mathrm{L}^q\big(0,T;\mathrm{W}^{1,q}_{0,\mathrm{div}}(\Omega)^d\big) \cap \mathrm{L}^\infty\big(0,T;\mathrm{L}^2_{\mathrm{div}}(\Omega)^d\big), \quad \mathbf{S} \in \mathrm{L}^{q'}(Q)^{d\times d},$$

and

$$\begin{aligned} &-\big\langle \boldsymbol{u}, \partial_t\boldsymbol{\xi}\big\rangle_Q - \big\langle \boldsymbol{u}\otimes\boldsymbol{u}, \mathbf{D}\boldsymbol{\xi}\big\rangle_Q + \big\langle \mathbf{S}, \mathbf{D}\boldsymbol{\xi}\big\rangle_Q \\ &\qquad = \langle \boldsymbol{f}, \boldsymbol{\xi}\rangle_Q + \big\langle \boldsymbol{u}_0, \boldsymbol{\xi}(0,\cdot)\big\rangle_\Omega \quad \text{for all } \boldsymbol{\xi} \in \mathrm{C}^\infty_{0,\mathrm{div}}((-T,T)\times\Omega)^d, \end{aligned} \tag{2.3}$$

$$\big(\mathbf{D}\boldsymbol{u}(z), \mathbf{S}(z)\big) \in \mathcal{A}(z) \qquad \text{for a.e. } z \in Q. \tag{2.4}$$

We choose a pressure-free notion of weak solution, because in the unsteady problem subject to homogeneous Dirichlet boundary conditions on Lipschitz domains, one can only expect to establish a distributional (in time) pressure, see Temam (1984, Ch. III, § 3, pp. 307, Rem. 3.5).

2.1 *Implicit constitutive laws*

Approximation of $\mathcal{A}$. The implicit relation encoded by $\mathcal{A}$ can be viewed as a set-valued map. In order to perform the analysis we require a single-valued map, and thus a measurable selection $\mathbf{S}^\star$ of the graph $\mathcal{A}$ is chosen, which may have discontinuities.

LEMMA 2.2 (Measurable selection, (Bulíček *et al.*, 2012, Rem. 1.1, Lem. 2.2)). Let the mapping $Q \ni z \mapsto \mathcal{A} \subset \mathbb{R}^{d\times d}_{\mathrm{sym}} \times \mathbb{R}^{d\times d}_{\mathrm{sym}}$ satisfy Assumption 1.1. Then, there exists a measurable selection $\mathbf{S}^\star \colon Q \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$, i.e.,

$$\big(\mathbf{B}, \mathbf{S}^\star(z,\mathbf{B})\big) \in \mathcal{A}(z) \quad \text{for all } \mathbf{B} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}, \quad \text{for a.e. } z \in Q, \tag{2.5}$$

and $\mathbf{S}^\star$ is $(\mathcal{L}(Q) \otimes \mathcal{B}(\mathbb{R}^{d\times d}_{\mathrm{sym}})) - \mathcal{B}(\mathbb{R}^{d\times d}_{\mathrm{sym}})$-measurable. Furthermore, for a.e. $z \in Q$, one has that

(a1) dom $\mathbf{S}^\star(z,\cdot) = \mathbb{R}^{d\times d}_{\mathrm{sym}}$;

(a2) $\mathbf{S}^\star$ is monotone, i.e., for all $\mathbf{B}_1, \mathbf{B}_2 \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$,

$$(\mathbf{S}^\star(z,\mathbf{B}_1) - \mathbf{S}^\star(z,\mathbf{B}_2)) : (\mathbf{B}_1 - \mathbf{B}_2) \geq 0;$$

(a3) for any $\mathbf{B} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$ one has that

$$\mathbf{B} : \mathbf{S}^\star(z,\mathbf{B}) \geq -g(z) + c_*\big(|\mathbf{B}|^q + |\mathbf{S}^\star(z,\mathbf{B})|^{q'}\big);$$

(a4) Let U be a dense set in $\mathbb{R}^{d\times d}_{\mathrm{sym}}$ and let $(\mathbf{D},\mathbf{S}) \in \mathbb{R}^{d\times d}_{\mathrm{sym}} \times \mathbb{R}^{d\times d}_{\mathrm{sym}}$. The following are equivalent:

  (i) $(\mathbf{S} - \mathbf{S}^\star(z,\mathbf{B})) : (\mathbf{D} - \mathbf{B}) \geq 0 \quad$ for all $\mathbf{B} \in U$;

  (ii) $(\mathbf{D},\mathbf{S}) \in \mathcal{A}(z)$;

(a5) $\mathbf{S}^\star$ is locally bounded, i.e., for a given $r > 0$ there exists a constant $c = c(r)$ such that

$$|\mathbf{S}^\star(z,\mathbf{A})| \leq c \qquad \text{for all } z \in Q \text{ and for all } \mathbf{A} \in B_r(\mathbf{0}) \subset \mathbb{R}^{d\times d}_{\mathrm{sym}}.$$

To show the existence of solutions to the approximate problem considered below, continuity of the (approximate) stress tensor is required. Hence, we introduce the following assumptions on a sequence of approximations of the selection $\mathbf{S}^\star$.

ASSUMPTION 2.3 (Properties of $\mathbf{S}^k$, $k \in \mathbb{N}$). Given the selection $\mathbf{S}^\star \colon Q \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$ in Lemma 2.2, assume that there is a sequence $\{\mathbf{S}^k\}_{k\in\mathbb{N}}$ of Carathéodory functions $\mathbf{S}^k \colon Q \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$ such that:

(α1) $\mathbf{S}^k(z,\cdot)$ is monotone, i.e., for all $\mathbf{A}_1, \mathbf{A}_2 \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$ and for a.e. $z \in Q$, we have

$$\big(\mathbf{S}^k(z,\mathbf{A}_1) - \mathbf{S}^k(z,\mathbf{A}_2)\big) : (\mathbf{A}_1 - \mathbf{A}_2) \geq 0.$$

(α2) There exists a constant $\widetilde{c}_* > 0$ and a nonnegative function $\widetilde{g} \in \mathrm{L}^1(Q)$ such that, for all $k \in \mathbb{N}$, for any $\mathbf{A} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$ and for a.e. $z \in Q$, one has that

$$\mathbf{A} : \mathbf{S}^k(z,\mathbf{A}) \geq -\widetilde{g}(z) + \widetilde{c}_*\left(|\mathbf{A}|^q + \left|\mathbf{S}^k(z,\mathbf{A})\right|^{q'}\right).$$

(α3) Let $U \subset \mathbb{R}^{d\times d}_{\mathrm{sym}}$ be a dense set. For any sequence $\{\mathbf{D}^k\}_{k\in\mathbb{N}}$ bounded in $\mathrm{L}^\infty(Q)^{d\times d}$, for any $\mathbf{B} \in U$ and all $\varphi \in \mathrm{C}_0^\infty(Q)$ such that $\varphi \geq 0$, we have

$$\liminf_{k\to\infty} \int_Q \big(\mathbf{S}^k(\cdot,\mathbf{D}^k) - \mathbf{S}^\star(\cdot,\mathbf{B})\big) : \big(\mathbf{D}^k - \mathbf{B}\big)\varphi \,\mathrm{d}z \geq 0.$$

In the existence proofs in Bulíček *et al.* (2009, 2012) and Diening *et al.* (2013) the approximating sequence $\mathbf{S}^k$ is chosen as the convolution of the selection $\mathbf{S}^\star$ in the second argument with a mollification kernel.

Example 2.4 (Approximation by mollification). Let $\rho \in \mathrm{C}_0^\infty(\mathbb{R}^{d\times d}_{\mathrm{sym}})$ be a mollification kernel, i.e., a nonnegative, radially symmetric function, the support of which is contained in the unit ball $B_1(\mathbf{0}) \subset \mathbb{R}^{d\times d}_{\mathrm{sym}}$ and which satisfies $\int_{\mathbb{R}^{d\times d}_{\mathrm{sym}}} \rho(\mathbf{A})\,\mathrm{d}\mathbf{A} = 1$. For $k \in \mathbb{N}$ set $\rho^k(\mathbf{B}) := k^{d^2}\rho(k\mathbf{B})$ and define the mollification of $\mathbf{S}^\star$ with respect to the last argument by

$$\mathbf{S}^k(z,\mathbf{B}) := \big(\mathbf{S}^\star * \rho^k\big)(z,\mathbf{B}) = \int_{\mathbb{R}^{d\times d}_{\mathrm{sym}}} \mathbf{S}^\star(z,\mathbf{A})\rho^k(\mathbf{B}-\mathbf{A})\,\mathrm{d}\mathbf{A}, \quad z \in Q,\ \mathbf{B} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}. \tag{2.6}$$

Lemma 3.21 in Tscherpel (2018) shows that $\mathbf{S}^k$ satisfies Assumption 2.3, see also Bulíček *et al.* (2012).

A possibly more practicable approximation based on a piecewise affine interpolant can be used in the case of a radially symmetric selection function $\mathbf{S}^\star$ under additional regularity assumptions.

Example 2.5 (Approximation by affine interpolation). Assume that $\mathrm{S}^\star\colon Q \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is a measurable function with $\mathrm{S}^\star(z,0) = 0$, for any $z \in Q$, such that $\mathbf{S}^\star\colon Q \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$, defined by

$$\mathbf{S}^\star(z,\mathbf{B}) = \begin{cases} \mathrm{S}^\star\big(z,|\mathbf{B}|\big)\,\frac{\mathbf{B}}{|\mathbf{B}|} & \text{if } \mathbf{B} \neq \mathbf{0},\\ \mathbf{0} & \text{if } \mathbf{B} = \mathbf{0},\end{cases}$$

is a measurable selection of a graph $\mathcal{A}$ satisfying Assumption 1.1. Furthermore, we assume that

(i) $\mathrm{S}^\star(z,\cdot)\colon \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is monotone for a.e. $z \in Q$.

Denote by $J^\star := \bigcup_{z\in Q} J(\mathrm{S}^\star(z,\cdot))$ the overall jump set, where $J(\mathrm{S}^\star(z,\cdot))$ is the jump set of $\mathrm{S}^\star(z,\cdot)$, which is countable by monotonicity of $\mathrm{S}^\star(z,\cdot)$ for fixed $z \in Q$, see Alberti & Ambrosio (1999, Thm. 2.2). Let us assume one of the following:

(iia) The set $J^\star$ is finite and for a.e. $z \in Q$ the function $\mathrm{S}^\star(z,\cdot)$ is locally Lipschitz continuous on each connected component of $\mathbb{R}_{\geq 0}\backslash J^\star$ (the Lipschitz constants are allowed to depend on $z \in Q$).

(iib) The set $J^\star$ is countable, without accumulation points, the jump-sizes are bounded above by a constant $H > 0$, and $\mathrm{S}^\star(z,\cdot)$ is Lipschitz continuous on each connected component of $\mathbb{R}_{\geq 0}\backslash J^\star$, with Lipschitz constants bounded uniformly in $z \in Q$ and independently of the specific component, say by $L > 0$.

Then, there exists an index set $\mathcal{I}$ ($\mathcal{I} = \{0,\dots,I\}$ for some $I \in \mathbb{N}$ in case (iia) and $\mathcal{I} = \mathbb{N}_0$ in case (iib)) and there exists a sequence $\{a_i\}_{i\in\mathcal{I}} \subset \mathbb{R}_{\geq 0}$, such that $J^\star \subset A := \cup_{i\in\mathcal{I}} a_i$. Without loss of generality, assume that $a_0 = 0$ and $a_{i-1} < a_i$ for all $i \in \mathcal{I}\backslash\{0\}$.

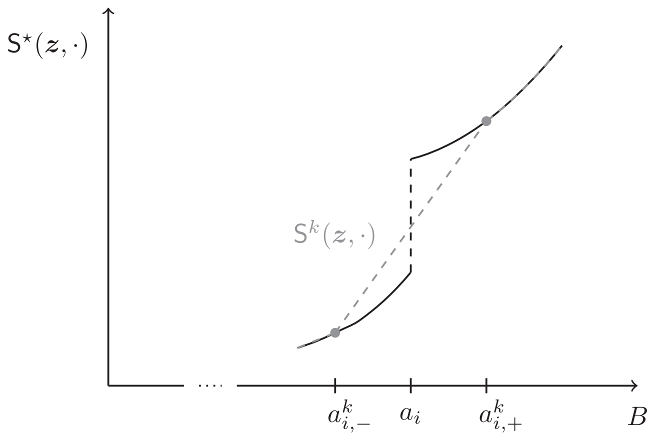

FIG. 1. Schematic representation of the construction of S^k, $k \in \mathbb{N}$.

We construct the approximation as follows. There exists a $k_0 \in \mathbb{N}$ such that $\frac{2}{k_0} < \inf_{i\in\mathcal{I}\setminus\{0\}}(a_i - a_{i-1})$ since either A is finite or A does not have any accumulation points. Let $k \in \mathbb{N}$, with $k \geq k_0$ be arbitrary but fixed. Denote, for $i \in \mathcal{I}$,

$$a^k_{i,-} := a_i - \frac{1}{k}, \quad a^k_{i,+} := a_i + \frac{1}{k}, \quad A^k_i := \left[a^k_{i,-}, a^k_{i,+}\right] \quad \text{and} \quad A^k := \bigcup_{i\in\mathcal{I}} A^k_i.$$

Let $z \in Q$ be arbitrary but fixed. First we extend $\mathsf{S}^\star(z,\cdot)$ as an odd function to $[-\frac{1}{k},\infty)$, still denoted by $\mathsf{S}^\star(z,\cdot)$. Since the point evaluations $\mathsf{S}^\star(z, a^k_{i,\pm})$, for $i \in \mathcal{I}$, are well-defined, we can define

$$\begin{aligned} \overline{\mathsf{S}}^k_i(z,B) &:= \frac{k}{2}\left(\mathsf{S}^\star\big(z,a^k_{i,-}\big)\frac{a^k_{i,+}}{a^k_{i,-}} - \mathsf{S}^\star\big(z,a^k_{i,+}\big)\right)\big(a^k_{i,-} - B\big) + \mathsf{S}^\star\big(z,a^k_{i,-}\big)\frac{B}{a^k_{i,-}}, \\ \mathsf{S}^k(z,B) &:= \begin{cases} \mathsf{S}^\star(z,B) & \text{if } B \notin A^k, \\ \overline{\mathsf{S}}^k_i(z,B) & \text{if } B \in A^k_i, \quad i \in \mathcal{I}. \end{cases} \end{aligned} \tag{2.7}$$

On A^k_i the approximation $\mathsf{S}^k(z,\cdot)$ is the affine interpolant between $\mathsf{S}^\star(z,a^k_{i,-})$ and $\mathsf{S}^\star(z,a^k_{i,+})$ and otherwise $\mathsf{S}^\star(z,\cdot)$ is unchanged, cf. Fig. 1. The resulting approximating sequence $\mathbf{S}^k(z,\cdot)$ satisfies Assumption 2.3, see Corollary 3.24 in Tscherpel (2018).

Minty's Trick. The following lemma is one of the crucial tools for the identification of the implicit constitutive law upon passage to the limit.

LEMMA 2.6 (Convergence lemma of Minty type, (Bulíček *et al.*, 2012, Lem. 2.4) and (Bulíček & Málek, 2016, Lem. 3.1)). Let $Q \ni z \mapsto \mathcal{A}(z) \subset \mathbb{R}^{d\times d}_{\rm sym} \times \mathbb{R}^{d\times d}_{\rm sym}$ satisfy (A2), (A3) in Assumption 1.1 and assume that there are sequences $\{\mathbf{S}^j\}_{j\in\mathbb{N}}$ and $\{\mathbf{D}^j\}_{j\in\mathbb{N}}$, and there is a measurable set $\widetilde{Q} \subset Q$ and a $p \in (1,\infty)$ such that

$$\begin{aligned}
\big(\mathbf{D}^j(z), \mathbf{S}^j(z)\big) &\in \mathcal{A}(z) && \text{for a.e. } z \in \widetilde{Q},\\
\mathbf{D}^j &\rightharpoonup \mathbf{D} && \text{weakly in } \mathrm{L}^p(\widetilde{Q})^{d\times d},\\
\mathbf{S}^j &\rightharpoonup \mathbf{S} && \text{weakly in } \mathrm{L}^{p'}(\widetilde{Q})^{d\times d},\\
\limsup_{j\to\infty} \langle \mathbf{S}^j, \mathbf{D}^j\rangle_{\widetilde{Q}} &\le \langle \mathbf{S}, \mathbf{D}\rangle_{\widetilde{Q}}\,.
\end{aligned}$$

Then, we have that $(\mathbf{D}(z), \mathbf{S}(z)) \in \mathcal{A}(z)$ for a.e. $z \in \widetilde{Q}$.

2.2 *Lipschitz approximation*

For small $q \in (1,\infty)$, a weak solution according to Definition 2.1 is not an admissible test function because of the presence of the convective term. The Lipschitz truncation method helps to identify the implicit relation despite the lack of admissibility. It first appeared in Acerbi & Fusco (1988), and since then the method was further developed and refined in a series of papers, see, e.g., Kinnunen & Lewis (2002), Frehse *et al.* (2003), Diening *et al.* (2008, 2010, 2013) and Breit *et al.* (2012, 2013), to mention just a few.

For a sequence of solutions to a sequence of divergence-form evolution equations a solenoidal parabolic Lipschitz truncation was developed in Breit *et al.* (2013). Note that the sets $\mathscr{B}_{l,j}$ in the following lemma satisfy $\mathscr{B}_{l,j} = \mathcal{O}_{l,j} \cap Q_0$, where $\mathcal{O}_{l,j}$ are the 'bad sets' in the construction in Breit *et al.* (2013).

LEMMA 2.7 (Parabolic solenoidal Lipschitz approximation (Breit *et al.*, 2013, Thm. 2.2, Cor. 2.4)). Let $p \in (1,\infty)$, $\sigma \in (1, \min(p,p'))$ and let $Q_0 = I_0 \times B_0 \subset \mathbb{R}\times\mathbb{R}^d$ be a parabolic cylinder, for $d = 3$, for an open interval I_0 and an open ball B_0. For $\alpha > 0$ we denote by αQ_0 the cylinder with the same center as Q_0, but scaled by α. Let $\{\boldsymbol{v}^l\}_{l\in\mathbb{N}}$ be a sequence of (weakly) divergence-free functions, which is converging to zero weakly in $\mathrm{L}^p(I_0; \mathrm{W}^{1,p}(B_0)^d)$, strongly in $\mathrm{L}^\sigma(Q_0)^d$, and is uniformly bounded in $\mathrm{L}^\infty(I_0, \mathrm{L}^\sigma(B_0)^d)$. Consider a sequence $\{\mathbf{G}_1^l\}_{l\in\mathbb{N}}$, converging to zero weakly in $\mathrm{L}^{p'}(Q_0)^{d\times d}$, and a second sequence, $\{\mathbf{G}_2^l\}_{l\in\mathbb{N}}$, converging to zero strongly in $\mathrm{L}^\sigma(Q_0)^{d\times d}$. Furthermore, denoting $\mathbf{G}^l := \mathbf{G}_1^l + \mathbf{G}_2^l$, assume that, for any $l \in \mathbb{N}$, the equation

$$\langle \partial_t \boldsymbol{v}^l, \boldsymbol{\xi}\rangle_{Q_0} = \langle \mathbf{G}^l, \nabla\boldsymbol{\xi}\rangle_{Q_0} \quad \text{for all } \boldsymbol{\xi} \in \mathrm{C}^\infty_{0,\mathrm{div}}(Q_0)^d \tag{2.8}$$

is satisfied. Then, there exists a $j_0 \in \mathbb{N}$,

- a double sequence $\{\lambda_{l,j}\}_{l,j\in\mathbb{N}} \subset \mathbb{R}$ with $\lambda_{l,j} \in \left[2^{2^j}, 2^{2^{j+1}-1}\right]$, for any $l,j \in \mathbb{N}$,
- a double sequence of open sets $\mathscr{B}_{l,j} \subset Q_0$, $l,j \in \mathbb{N}$,
- a double sequence of functions $\{\boldsymbol{v}^{l,j}\}_{l,j\in\mathbb{N}} \subset \mathrm{L}^1(Q_0)^d$ and
- a nonnegative function $\zeta \in \mathrm{C}_0^\infty(\frac{1}{6}Q_0)$ such that $\mathbb{1}_{\frac{1}{8}Q_0} \le \zeta \le \mathbb{1}_{\frac{1}{6}Q_0}$,

such that

(i) $\boldsymbol{v}^{l,j} \in \mathrm{L}^s(\frac{1}{4}I_0; \mathrm{W}^{1,s}_{0,\mathrm{div}}(\frac{1}{6}B_0)^d)$ for all $s \in [1,\infty)$ and $\mathrm{supp}(\boldsymbol{v}^{l,j}) \subset \frac{1}{6}Q_0$ for any $j \geq j_0$ and any $l \in \mathbb{N}$;

(ii) $\boldsymbol{v}^{l,j} = \boldsymbol{v}^l$ on $\frac{1}{8}Q_0 \backslash \mathscr{B}_{l,j}$, i.e., $\{\boldsymbol{v}^{l,j} \neq \boldsymbol{v}^l\} \cap \frac{1}{8}Q_0 \subset \mathscr{B}_{l,j}$ for any $j \geq j_0$ and any $l \in \mathbb{N}$;

(iii) there exists a constant $c > 0$ such that

$$\limsup_{l\to\infty} \lambda_{l,j}^p \left|\mathscr{B}_{l,j}\right| \leq c2^{-j} \quad \text{for all } j \geq j_0;$$

(iv) there exists a constant $c > 0$ such that

$$\left\|\nabla \boldsymbol{v}^{l,j}\right\|_{\mathrm{L}^\infty(\frac{1}{4}Q_0)} \leq c\lambda_{l,j} \quad \text{for all } j \geq j_0 \text{ and all } l \in \mathbb{N};$$

(v) for any fixed $j \geq j_0$ we have

$$\begin{aligned} \boldsymbol{v}^{l,j} &\to \mathbf{0} \quad \text{strongly in } \mathrm{L}^\infty(\tfrac{1}{4}Q_0)^d, \\ \nabla \boldsymbol{v}^{l,j} &\rightharpoonup \mathbf{0} \quad \text{weakly in } \mathrm{L}^s(\tfrac{1}{4}Q_0)^{d\times d} \quad \text{for all } s \in [1,\infty), \end{aligned}$$

as $l \to \infty$;

(vi) there exists a constant $c > 0$ such that

$$\limsup_{l\to\infty} \left|\langle \mathbf{G}^l, \nabla \boldsymbol{v}^{l,j}\rangle\right| \leq c2^{-j} \quad \text{for all } j \geq j_0;$$

(vii) there exists a constant $c > 0$ such that for any $\mathbf{H} \in \mathrm{L}^{p'}(\frac{1}{6}Q_0)^{d\times d}$, we have that

$$\limsup_{l\to\infty} \left|\left\langle (\mathbf{G}_1^l + \mathbf{H}), \nabla \boldsymbol{v}^l \zeta \mathbb{1}_{\mathscr{B}^c_{l,j}}\right\rangle\right| \leq c2^{-\frac{j}{p}} \quad \text{for all } j \geq j_0.$$

The lemma is stated for $d = 3$, but according to Breit *et al.* (2013, Rem. 2.1, p. 2692) the result holds for all $d \geq 2$ with minor modifications of the proof. In the convergence proof we will use the following corollary, including a lower order term in the equation.

COROLLARY 2.8 (Lower order term for parabolic solenoidal Lipschitz approximation). Let $p \in (1,\infty)$, $\sigma \in (1, \min(p,p'))$ and let $Q_0 = I_0 \times B_0 \subset \mathbb{R} \times \mathbb{R}^d$ be a parabolic cylinder, for $d \geq 2$, for an open interval I_0 and an open ball B_0. Let $\{\boldsymbol{v}^l\}_{l\in\mathbb{N}}$ be a sequence of weakly divergence-free functions, which is converging to zero weakly in $\mathrm{L}^p(I_0; \mathrm{W}^{1,p}_0(B_0)^d)$, strongly in $\mathrm{L}^\sigma(Q_0)^d$ and is uniformly bounded in $\mathrm{L}^\infty(I_0, \mathrm{L}^\sigma(B_0)^d)$. Consider a sequence, $\{\mathbf{G}_1^l\}_{l\in\mathbb{N}}$, converging to zero weakly in $\mathrm{L}^{p'}(Q_0)^{d\times d}$, a second sequence, $\{\widetilde{\mathbf{G}}_2^l\}_{l\in\mathbb{N}}$, converging to zero strongly in $\mathrm{L}^\sigma(Q_0)^{d\times d}$ and a third sequence, $\{\boldsymbol{f}^l\}_{l\in\mathbb{N}}$, converging to zero strongly in $\mathrm{L}^{\sigma}(Q_0)^d$. Furthermore, denoting $\widetilde{\mathbf{G}}^l := \mathbf{G}_1^l + \widetilde{\mathbf{G}}_2^l$, assume that, for any $l \in \mathbb{N}$, the equation

$$\langle \partial_t \boldsymbol{v}^l, \boldsymbol{\xi} \rangle_{Q_0} = \langle \widetilde{\mathbf{G}}^l, \nabla \boldsymbol{\xi} \rangle_{Q_0} + \langle \boldsymbol{f}^l, \boldsymbol{\xi} \rangle_{Q_0} \quad \text{for all } \boldsymbol{\xi} \in \mathrm{C}_{0,\mathrm{div}}^{\infty}(Q_0)^d \tag{2.9}$$

is satisfied. Then, the same statement as in Lemma 2.7 holds with $\mathbf{G}_2^l = \widetilde{\mathbf{G}}_2^l - \nabla \Delta^{-1} \boldsymbol{f}^l$, and Δ^{-1} signifying the inverse Dirichlet Laplacian.

Proof. For a.e. $t \in I_0$ we wish to find a $\boldsymbol{g}^l(t,\cdot) \in \mathrm{W}_0^{1,\sigma}(B_0)^d$ such that

$$-\left\langle \nabla \boldsymbol{g}^l, \nabla \boldsymbol{v} \right\rangle_{B_0} = \left\langle \boldsymbol{f}^l(t,\cdot), \boldsymbol{v} \right\rangle_{B_0} \quad \text{for all } \boldsymbol{v} \in \mathrm{C}_0^{\infty}(B_0)^d. \tag{2.10}$$

Standard regularity theory for Poisson's equation (see (Grisvard, 2011, Thm. 2.4.2.5) and (Gilbarg & Trudinger, 2001, Lem. 9.17)) guarantees the existence of a unique $\boldsymbol{g}^l(t,\cdot) \in \mathrm{W}^{2,\sigma}(B_0)^d \cap \mathrm{W}_0^{1,\sigma}(B_0)^d$ solving (2.10) such that

$$\|\boldsymbol{g}^l(t,\cdot)\|_{\mathrm{W}^{2,\sigma}(B_0)} \le c \|\boldsymbol{f}^l(t,\cdot)\|_{\mathrm{L}^{\sigma}(B_0)}, \tag{2.11}$$

since $\sigma \in (1,\infty)$ and ∂B_0 is smooth. Viewing $\boldsymbol{g}^l$ as a function on $Q_0 = I_0 \times B_0$ by (2.11), one has that

$$\|\boldsymbol{g}^l\|_{\mathrm{L}^{\sigma}(I_0;\mathrm{W}^{2,\sigma}(B_0))} \le c \|\boldsymbol{f}^l\|_{\mathrm{L}^{\sigma}(Q_0)} \to 0, \quad \text{as } l \to \infty, \tag{2.12}$$

by assumption. Thus, we have in particular that $\nabla \boldsymbol{g}^l \to \mathbf{0}$ strongly in $\mathrm{L}^{\sigma}(Q_0)^{d\times d}$, as $l \to \infty$, and hence $\widetilde{\mathbf{G}}_2^l - \nabla \boldsymbol{g}^l$ converges to zero strongly in $\mathrm{L}^{\sigma}(Q_0)^{d\times d}$, as $l \to \infty$. Applying (2.10) in (2.9) shows that

$$\langle \partial_t \boldsymbol{v}^l, \boldsymbol{\xi} \rangle_{Q_0} = \left\langle \mathbf{G}_1^l + \widetilde{\mathbf{G}}_2^l - \nabla \boldsymbol{g}^l, \nabla \boldsymbol{\xi} \right\rangle_{Q_0} \quad \text{for all } \boldsymbol{\xi} \in \mathrm{C}_{0,\mathrm{div}}^{\infty}(Q_0)^d,$$

and thus all assumptions of Lemma 2.7 are satisfied and the claim follows. □

2.3 *Compactness in time*

LEMMA 2.9 (Parabolic interpolation (DiBenedetto, 1993, Ch. I, Prop. 2.3)). Let $d \ge 2$, let $\Omega \subset \mathbb{R}^d$ be a Lipschitz domain, let $T \in (0,\infty)$, $Q = (0,T) \times \Omega$ and let $p > 1$. Then, there exists a constant $c > 0$ such that

$$\|v\|_{\mathrm{L}^{\frac{p(d+2)}{d}}(Q)} \le c\, \|v\|_{\mathrm{L}^p(0,T;\mathrm{W}^{1,p}(\Omega))}^{\frac{d}{d+2}} \|v\|_{\mathrm{L}^{\infty}(0,T;\mathrm{L}^2(\Omega))}^{\frac{2}{d+2}} \tag{2.13}$$

for all $v \in \mathrm{L}^p(0,T;\mathrm{W}^{1,p}(\Omega)) \cap \mathrm{L}^{\infty}(0,T;\mathrm{L}^2(\Omega))$.

LEMMA 2.10 (Simon, 1987, Thm. 3, p. 80). Let X, B be Banach spaces such that the embedding $X \hookrightarrow\hookrightarrow B$ is compact. Let $\mathcal{F} \subset \mathrm{L}^p(0,T;B)$ for some $p \in [1,\infty)$ and let

(i) $\mathcal{F}$ be bounded in $\mathrm{L}^1_{\mathrm{loc}}(0,T;X)$,

(ii) $\int_0^{T-\varepsilon} \|f(s+\varepsilon,\cdot) - f(s,\cdot)\|_B^p \,\mathrm{d}s \to 0, \quad$ as $\varepsilon \to 0$, uniformly for $f \in \mathcal{F}$.

Then, $\mathcal{F}$ is relatively compact in $\mathrm{L}^p(0,T;B)$.

3. Finite element approximation

3.1 *Finite element spaces and assumptions*

The setting here is slightly more general than the one in Diening *et al.* (2013).

ASSUMPTION 3.1 (Triangulations $\{\mathcal{T}_n\}_{n\in\mathbb{N}}$). Let us assume that $d \geq 2$ and that Ω is a bounded Lipschitz polytopal domain. Furthermore, assume that $\{\mathcal{T}_n\}_{n\in\mathbb{N}}$ is a family of simplicial partitions of $\overline{\Omega}$ (in the sense of (Ciarlet, 2002, Sec. 2.1, p. 38)) such that the following conditions hold:

(i) Each element $K \in \mathcal{T}_n$ is affine-equivalent to the closed standard reference simplex, which is given by $\widehat{K} := \mathrm{conv}\{\mathbf{0}, \boldsymbol{e}_1, \ldots, \boldsymbol{e}_d\} \subset \mathbb{R}^d$, i.e., there exists an affine invertible function $\mathbf{F}_K \colon K \to \widehat{K}$;

(ii) $\{\mathcal{T}_n\}_{n\in\mathbb{N}}$ is shape-regular, i.e., there exists a constant c_r (independent of $n \in \mathbb{N}$) such that

$$\frac{h_K}{\rho_K} \leq c_r \quad \text{for all } K \in \mathcal{T}_n \text{ and all } n \in \mathbb{N},$$

where $h_K := \mathrm{diam}(K)$ and $\rho_K := \sup\{\mathrm{diam}(B)\colon\ B \text{ is a ball contained in} K\}$.

For $n \in \mathbb{N}$ we denote by $h_n := \max\{h_K\colon\ K \in \mathcal{T}_n\}$ the spatial grid-size.

Finite element spaces. Let $\widehat{\mathbb{P}}_{\mathbb{V}} \subset \mathrm{W}^{1,\infty}(\widehat{K})^d$ and let $\widehat{\mathbb{P}}_{\mathbb{Q}} \subset \mathrm{L}^\infty(\widehat{K})$ be finite-dimensional function spaces on the reference simplex $\widehat{K}$ (with a slight abuse of notation) as in Diening *et al.* (2013). Further, let $\mathbb{V} \subset \mathrm{C}(\overline{\Omega})^d$ and let $\mathbb{Q} \subset \mathrm{L}^\infty(\Omega)$. Then we define the conforming finite element spaces $\mathbb{V}^n$ and $\mathbb{Q}^n$ with respect to $\mathcal{T}_n$ by

$$\mathbb{V}^n := \left\{\boldsymbol{V} \in \mathbb{V}\colon \quad \boldsymbol{V}|_K \circ \mathbf{F}_K^{-1} \in \widehat{\mathbb{P}}_{\mathbb{V}},\, K \in \mathcal{T}_n \text{ and } \boldsymbol{V}|_{\partial\Omega} = \mathbf{0}\right\}, \tag{3.1}$$

$$\mathbb{Q}^n := \left\{Q \in \mathbb{Q}\colon \quad Q|_K \circ \mathbf{F}_K^{-1} \in \widehat{\mathbb{P}}_{\mathbb{Q}},\, K \in \mathcal{T}_n\right\}. \tag{3.2}$$

Let us also introduce the subspace of discretely divergence-free functions of $\mathbb{V}^n$ and the subspace of zero integral mean functions of $\mathbb{Q}^n$ by

$$\mathbb{V}^n_{\mathrm{div}} := \{\boldsymbol{V} \in \mathbb{V}^n\colon\ \langle \mathrm{div}\,\boldsymbol{V}, Q\rangle_\Omega = 0 \text{ for all } Q \in \mathbb{Q}^n\}, \tag{3.3}$$

$$\mathbb{Q}^n_0 := \left\{Q \in \mathbb{Q}^n\colon\ \int_\Omega Q \,\mathrm{d}\boldsymbol{x} = 0\right\}. \tag{3.4}$$

Note that the functions in $\mathbb{V}^n_{\mathrm{div}}$ are in general not divergence-free, so in general $\mathbb{V}^n_{\mathrm{div}} \not\subset \mathrm{W}^{1,\infty}_{0,\mathrm{div}}(\Omega)^d$.

Assumption 3.2 (Approximability (Diening *et al.* (2013), Assump. 5)). Assume that for all $p \in [1, \infty)$, we have that

$$\inf_{V \in \mathbb{V}^n} \|\boldsymbol{v} - \boldsymbol{V}\|_{W^{1,p}(\Omega)} \to 0, \quad \text{as } n \to \infty \quad \text{for all } \boldsymbol{v} \in W_0^{1,p}(\Omega)^d, \tag{3.5}$$

$$\inf_{Q \in \mathbb{Q}^n} \|h - Q\|_{L^p(\Omega)} \to 0, \quad \text{as } n \to \infty \quad \text{for all } h \in L_0^p(\Omega). \tag{3.6}$$

Note that this assumption implies that $h_n \to 0$, as $n \to \infty$.

Projectors. For the convergence analysis we use certain projectors to the respective finite element spaces and we require suitable assumptions on them. Since we do not need local stability of the projector Π^n, we assume less than in Diening *et al.* (2013).

Assumption 3.3 (Projector Π^n). Assume that for each $n \in \mathbb{N}$ there exists a linear projector $\Pi^n \colon W_0^{1,1}(\Omega)^d \to \mathbb{V}^n$ such that:

(i) (preservation of the divergence in $(\mathbb{Q}^n)'$) for any $\boldsymbol{v} \in W_0^{1,1}(\Omega)^d$ one has that

$$\langle \operatorname{div} \boldsymbol{v}, Q \rangle_\Omega = \langle \operatorname{div} \Pi^n \boldsymbol{v}, Q \rangle_\Omega \qquad \text{for all } Q \in \mathbb{Q}^n;$$

(ii) ($W^{1,p}$-stability) for any $p \in (1, \infty)$ there exists a constant $c(p) > 0$ (independent of n) such that

$$\|\Pi^n \boldsymbol{v}\|_{W^{1,p}(\Omega)} \le c\, \|\boldsymbol{v}\|_{W^{1,p}(\Omega)} \quad \text{for all } \boldsymbol{v} \in W_0^{1,p}(\Omega)^d \text{ and all } n \in \mathbb{N}.$$

Assumption 3.4 (Projector $\Pi^n_{\mathbb{Q}}$). Assume that for each $n \in \mathbb{N}$ there exists a linear projector $\Pi^n_{\mathbb{Q}} \colon L^1(\Omega) \to \mathbb{Q}^n$ such that, for any $p \in (1, \infty)$, there exists a constant $c(p) > 0$ such that

$$\|\Pi^n_{\mathbb{Q}} h\|_{L^p(\Omega)} \le c\, \|h\|_{L^p(\Omega)} \quad \text{for all } h \in L^p(\Omega) \text{ and all } n \in \mathbb{N}. \tag{3.7}$$

Remark 3.5 (Properties of Π^n and $\Pi^n_{\mathbb{Q}}$).

(i) The stability in Assumption 3.3 (ii) and the approximability in (3.5) yield that

$$\|\boldsymbol{v} - \Pi^n \boldsymbol{v}\|_{W^{1,p}(\Omega)} \to 0, \quad \text{as } n \to \infty$$

for all $\boldsymbol{v} \in W_0^{1,p}(\Omega)^d$ with $p \in [1, \infty)$.

(ii) Similarly, the stability in (3.7) and the approximability in (3.6) imply that

$$\|h - \Pi^n_{\mathbb{Q}} h\|_{L^p(\Omega)} \to 0, \quad \text{as } n \to \infty$$

for all $h \in L^p(\Omega)$ with $p \in [1, \infty)$.

(iii) The existence of the Bogovskiĭ operator, see Bogovskiĭ (1979) and Diening *et al.* (2008, p. 223), implies that the continuous inf-sup condition holds for any $p \in (1, \infty)$. With this and Assumption 3.3 the corresponding discrete inf-sup condition holds uniformly in $n \in \mathbb{N}$, cf.

Fortin's Lemma for Banach spaces in Ern & Guermond (2004, Lem. 4.19). This means that the framework results in an inf-sup stable pair $(\mathbb{V}^n, \mathbb{Q}^n)$.

EXAMPLE 3.6 (Finite element spaces). The following elements satisfy Assumptions 3.2–3.4:

(i) the $\mathbb{P}_2 - \mathbb{P}_0$ element for $d = 2$, see Boffi *et al.* (2013, Sec. 8.4.3), where the projector Π^n is given and Assumption 3.3 (i) is shown; the stability in (ii) can be proved similarly as for the MINI element, see Belenki *et al.* (2012, App. A.1) and Diening *et al.* (2013, pp. 990);

(ii) the conforming Crouzeix–Raviart element, for $d = 2$, see Boffi *et al.* (2013, Ex. 8.6.1) and Crouzeix & Raviart (1973); the projector Π^n satisfying Assumption 3.3 (i) is given in Crouzeix & Raviart (1973, pp. 49) and it can be shown to satisfy Assumption 3.3 (ii), see, for example, Girault & Scott (2003, Thm. 3.3);

(iii) the Bernardi–Raugel element for $d \in \{2, 3\}$ (polynomial order $r = 1$) and $d = 3$ ($r = 2$), see Bernardi & Raugel (1985); the construction of Π^n satisfying Assumption 3.3 for $p = 2$ is contained therein and can be generalized to $p \in [1, \infty)$. See also Girault & Lions (2001) for $p \in [2, \infty)$ and $r = 1$;

(iv) the MINI element for $d \in \{2, 3\}$ ($r = 1$), see Boffi *et al.* (2013, Sec. 8.4.2, 8.7.1); the proof that Assumption 3.3 is satisfied is given in Belenki *et al.* (2012, App. A.1), see also Girault (2001, Lem. 4.5) and Diening *et al.* (2013, pp. 990);

(v) the Taylor–Hood element and its generalizations for $d \in \{2, 3\}$ and $r \geq d$, see Boffi *et al.* (2013, Sec. 8.8.2); the proof of Assumption 3.3 can be found in Girault & Scott (2003, Thm. 3.1, 3.2).

The following element satisfies Assumption 3.2–3.4 and, additionally, that $\mathbb{V}^n_{\mathrm{div}} \subset \mathrm{W}^{1,\infty}_{0,\mathrm{div}}(\Omega)^d$:

(vi) the family of Guzmán–Neilan elements for $d = 2$ ($k \geq 1$) and for $d = 3$ ($r = 1$), see Guzmán & Neilan (2014a,b). Therein Assumption 3.3 is shown for $p = 2$ when $d = 2$, and for $p \in [1, \infty)$ when $d = 3$. For stability for general $p \in [1, \infty)$ when $d = 2$ ($r = 1$), see also Diening *et al.* (2013).

L^2-Projector to $\mathbb{V}^n_{\mathrm{div}}$. Let us introduce the projector onto $\mathbb{V}^n_{\mathrm{div}}$, given by

$$\begin{aligned} &P^n_{\mathrm{div}} : \mathrm{L}^2(\Omega)^d \to \mathbb{V}^n_{\mathrm{div}}, \quad \text{and for } \boldsymbol{v} \in \mathrm{L}^2(\Omega)^d, \\ &\left\langle P^n_{\mathrm{div}}\boldsymbol{v}, \boldsymbol{V}\right\rangle_\Omega = (\boldsymbol{v}, \boldsymbol{V})_\Omega \quad \text{for all } \boldsymbol{V} \in \mathbb{V}^n_{\mathrm{div}}. \end{aligned} \tag{3.8}$$

Directly from the definition we have L^2-stability, i.e., for $\boldsymbol{v} \in \mathrm{L}^2(\Omega)^d$ we have

$$\left\| P^n_{\mathrm{div}}\boldsymbol{v} \right\|_{\mathrm{L}^2(\Omega)} \leq \|\boldsymbol{v}\|_{\mathrm{L}^2(\Omega)}. \tag{3.9}$$

By this and an approximation argument using the properties of Π^n (see Remark 3.5 (i)), we have that

$$P^n_{\mathrm{div}}\boldsymbol{w} \to \boldsymbol{w} \quad \text{strongly in } \mathrm{L}^2(\Omega)^d, \text{as } n \to \infty, \tag{3.10}$$

for any $\boldsymbol{w} \in \mathrm{L}^2_{\mathrm{div}}(\Omega)^d$.

3.2 *Convective term and its numerical approximation*

Motivated by the form of the convective term in the conservation of momentum equation, we consider the trilinear form b defined by

$$b(\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}) := -\langle \boldsymbol{u}\otimes\boldsymbol{v},\nabla\boldsymbol{w}\rangle_\Omega = \langle \boldsymbol{u}\otimes\boldsymbol{w},\nabla\boldsymbol{v}\rangle_\Omega - \langle \operatorname{div}\boldsymbol{u},\boldsymbol{v}\cdot\boldsymbol{w}\rangle_\Omega, \tag{3.11}$$

for $\boldsymbol{u},\boldsymbol{v},\boldsymbol{w} \in \mathrm{W}^{1,\infty}_0(\Omega)^d$, where the second equality follows by integration by parts. Hence, for divergence-free functions $\boldsymbol{u}$, the last term vanishes and $b(\boldsymbol{u},\cdot,\cdot)$ is skew-symmetric, i.e., $b(\boldsymbol{u},\boldsymbol{v},\boldsymbol{v})=0$ for any $\boldsymbol{u}\in\mathrm{W}^{1,\infty}_{0,\mathrm{div}}(\Omega)^d$ and any $\boldsymbol{v}\in\mathrm{W}^{1,\infty}_0(\Omega)^d$.

As in general $\mathbb{V}^n_{\mathrm{div}} \not\subset \mathrm{W}^{1,\infty}_{0,\mathrm{div}}(\Omega)^d$, the second term in (3.11) need not vanish for $\boldsymbol{u}\in\mathbb{V}^n_{\mathrm{div}}$. To preserve the skew-symmetry of the trilinear form associated with the convective term, the usual approach in the numerical analysis literature (see, e.g., (Temam, 1984)) is therefore to consider instead the skew-symmetric trilinear form

$$\widetilde{b}(\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}) := \frac{1}{2}\left(\langle \boldsymbol{u}\otimes\boldsymbol{w},\nabla\boldsymbol{v}\rangle_\Omega - \langle \boldsymbol{u}\otimes\boldsymbol{v},\nabla\boldsymbol{w}\rangle_\Omega\right) = -\langle \boldsymbol{u}\otimes\boldsymbol{v},\nabla\boldsymbol{w}\rangle_\Omega + \frac{1}{2}\langle \operatorname{div}\boldsymbol{u},\boldsymbol{v}\cdot\boldsymbol{w}\rangle_\Omega, \tag{3.12}$$

for $\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}\in\mathrm{W}^{1,\infty}_0(\Omega)^d$. Thus, we have that $\widetilde{b}(\boldsymbol{u},\boldsymbol{v},\boldsymbol{v})=0$ regardless of the solenoidality of $\boldsymbol{u}$. Note that for divergence-free functions $\boldsymbol{u}$ we have that $b(\boldsymbol{u},\cdot,\cdot)=\widetilde{b}(\boldsymbol{u},\cdot,\cdot)$.

In the equations the terms appear in the form $b(\boldsymbol{u},\boldsymbol{u},\boldsymbol{v})$ and $\widetilde{b}(\boldsymbol{u},\boldsymbol{u},\boldsymbol{v})$, for the velocity $\boldsymbol{u}$ and a test function $\boldsymbol{v}$. The natural function space for weak solutions of problem (P) is given by $\mathrm{L}^\infty(0,T;\mathrm{L}^2(\Omega)^d)\cap\mathrm{L}^q(0,T;\mathrm{W}^{1,q}_0(\Omega)^d)$, which embeds by Lemma 2.9 continuously into $\mathrm{L}^{\frac{q(d+2)}{d}}(Q)^d$. Also, provided that $q\geq\frac{2d}{d+2}$, we have that the embedding $\mathrm{L}^{\frac{q(d+2)}{d}}(Q)^d\hookrightarrow\mathrm{L}^2(Q)^d$ is continuous, which means that the expression $b(\boldsymbol{u}(t,\cdot),\boldsymbol{u}(t,\cdot),\boldsymbol{v})$ is integrable on $(0,T)$, for any $\boldsymbol{v}\in\mathrm{W}^{1,\infty}(\Omega)^d$. More specifically, with $\hat{q}:=\max\left(\left(\frac{q(d+2)}{2d}\right)',q\right)$ we have that

$$\left|\langle \boldsymbol{u}(t,\cdot)\otimes\boldsymbol{u}(t,\cdot),\nabla\boldsymbol{v}\rangle_\Omega\right| \leq c\,\|\boldsymbol{u}(t,\cdot)\|^2_{\mathrm{L}^{\frac{q(d+2)}{d}}(\Omega)}\|\boldsymbol{v}\|_{\mathrm{W}^{1,\hat{q}}(\Omega)}, \quad \text{provided that } q\geq\tfrac{2d}{d+2}. \tag{3.13}$$

On the other hand, for the modification (cf. the first term in (3.12)) of the trilinear form b associated with the convective term, one obtains

$$\left|\langle \boldsymbol{u}(t,\cdot)\otimes\boldsymbol{v},\nabla\boldsymbol{u}(t,\cdot)\rangle_\Omega\right| \leq c\,\|\boldsymbol{u}(t,\cdot)\|_{\mathrm{L}^{\frac{q(d+2)}{d}}(\Omega)}\|\boldsymbol{v}\|_{\mathrm{W}^{1,\hat{q}}(\Omega)}\|\nabla\boldsymbol{u}(t,\cdot)\|_{\mathrm{L}^q(\Omega)}, \quad \text{if } q\geq\tfrac{2(d+1)}{d+2}. \tag{3.14}$$

Evidently, the source of this more restrictive requirement on q is the modification of the trilinear form b, introduced to reinstate the skew symmetry of b, lost in the course of approximating the pointwise divergence-free solution by discretely divergence-free finite element functions. We note in passing that the restriction $q\geq\frac{2(d+1)}{d+2}$ in the unsteady case corresponds to the restriction $q\geq\frac{2d}{d+1}$ in the steady case in Diening *et al.* (2013).

Let us motivate the choice of the penalty term that we shall add to the weak form to relax the excessive restriction $q \geq \frac{2(d+1)}{d+2}$ to the natural restriction on $q > \frac{2d}{d+2}$. We note that by Hölder's inequality we have that

$$\begin{aligned}\left\|\widetilde{b}(\boldsymbol{u},\boldsymbol{u},\boldsymbol{v})\right\|_{\mathrm{L}^1((0,T))} &\leq \|\boldsymbol{u}\|^2_{\mathrm{L}^{2q'}(Q)}\|\nabla\boldsymbol{v}\|_{\mathrm{L}^q(Q)} + \|\boldsymbol{u}\|_{\mathrm{L}^{2q'}(Q)}\|\boldsymbol{v}\|_{\mathrm{L}^{2q'}(Q)}\|\nabla\boldsymbol{u}\|_{\mathrm{L}^q(Q)} \\ &\leq c\,\|\boldsymbol{u}\|^2_{\mathrm{X}(Q)}\|\boldsymbol{v}\|_{\mathrm{X}(Q)}, \end{aligned} \tag{3.15}$$

for $\boldsymbol{u},\boldsymbol{v} \in \mathrm{X}(Q) = \mathrm{L}^q(0,T;\mathrm{W}^{1,q}_0(\Omega)^d) \cap \mathrm{L}^{2q'}(Q)^d$, see (2.2), without any restrictions on the range of q, other than $q \in (1,\infty)$. This justifies the use of a regularizing term guaranteeing additional $\mathrm{L}^{2q'}$-integrability, cf. Section 4.

3.3 *Time discretization*

For the purpose of time discretization, let $l \in \mathbb{N}$ and define the time step by $\delta_l = T/l \to 0$, as $l \to \infty$. For $l \in \mathbb{N}$, we shall use the equidistant temporal grid on $[0,T]$ defined by $\{t^l_i\}_{i\in\{0,\dots,l\}}$, where $t^l_i := i\delta_l$, for $i \in \{0,\dots,l\}$. In the following we will suppress the superscript l and write t_i, $i \in \{0,\dots,l\}$.

For a Banach space X of functions, $l \in \mathbb{N}$ and a sequence $\{\varphi_i\}_{i\in\{0,\dots,l\}} \subset \mathrm{X}$, we consider the temporal difference quotient

$$d_t\varphi_i := \frac{1}{\delta_l}\left(\varphi_i - \varphi_{i-1}\right) \quad \text{for } i \in \{1,\dots,l\}. \tag{3.16}$$

Furthermore, for $l \in \mathbb{N}$ we denote by $\mathbb{P}^l_0(0,T;\mathrm{X})$ the linear space of left-continuous piecewise constant mappings from $(0,T]$ into X, with respect to the equidistant temporal grid $\{t_0,\dots,t_l\} \subset [0,T]$ and by $\mathbb{P}^l_1(0,T;\mathrm{X})$ the space of continuous, piecewise affine functions from $[0,T]$ into X, with respect to the same temporal grid. Let the piecewise constant and the piecewise affine interpolants $\overline{\varphi}$ and $\widetilde{\varphi}$ of $\{\varphi_i\}_{i\in\{0,\dots,l\}}$ be defined by

$$\overline{\varphi}(t) := \varphi_i \qquad \text{for } t \in (t_{i-1},t_i],\ i \in \{1,\dots,l\}, \tag{3.17}$$

$$\widetilde{\varphi}(t) := \varphi_i\frac{t-t_{i-1}}{\delta_l} + \varphi_{i-1}\frac{t_i - t}{\delta_l} \qquad \text{for } t \in [t_{i-1},t_i],\ i \in \{1,\dots,l\}, \tag{3.18}$$

so that $\overline{\varphi},\widetilde{\varphi},\partial_t\widetilde{\varphi} \in \mathrm{L}^\infty(0,T;\mathrm{X})$. Choosing the representative $\partial_t\widetilde{\varphi} \in \mathbb{P}^l_0(0,T;\mathrm{X})$, for $t \in (t_{i-1},t_i]$, we have $\partial_t\widetilde{\varphi}(t) = d_t\varphi_i$ and

$$\overline{\varphi}(t) - \widetilde{\varphi}(t) = (t_i - t)\partial_t\widetilde{\varphi}(t). \tag{3.19}$$

Furthermore, note that one has

$$\|\overline{\varphi}\|_{\mathrm{L}^\infty(0,T;\mathrm{X})} = \max_{i\in\{1,\dots,l\}}\left\|\varphi_i\right\|_{\mathrm{X}}, \qquad \|\overline{\varphi}\|^p_{\mathrm{L}^p(0,T;\mathrm{X})} = \delta_l\sum_{i=1}^{l}\left\|\varphi_i\right\|^p_{\mathrm{X}}, \quad \text{for } p \in [1,\infty), \tag{3.20}$$

$$\|\widetilde{\varphi}\|_{\mathrm{L}^\infty(0,T;\mathrm{X})} = \max_{i\in\{0,\dots,l\}} \|\varphi_i\|_{\mathrm{X}}, \qquad \|\widetilde{\varphi}\|^p_{\mathrm{L}^p(0,T;\mathrm{X})} \le c(p)\delta_l \sum_{i=0}^{l} \|\varphi_i\|^p_{\mathrm{X}}, \quad \text{for } p \in [1,\infty), \tag{3.21}$$

where $0 < c(p) \le 1$ by the Riesz–Thorin interpolation theorem (cf. (Bergh & Löfström, 1976, Thm. 1.1.1, p. 2)).

For a Bochner function $\psi \in \mathrm{L}^p(0,T;\mathrm{X})$, $p \in [1,\infty)$, we define the time averages with respect to the time grid $\{t_0,\dots,t_l\}$, for $l \in \mathbb{N}$, by

$$\psi_i := \fint_{t_{i-1}}^{t_i} \psi(t,\cdot)\,\mathrm{d}t \in \mathrm{X}, \quad i \in \{1,\dots,l\}. \tag{3.22}$$

For the piecewise constant interpolant $\overline{\psi}$ of the set of values $\{\psi_i\}_{i\in\{1,\dots,l\}}$, one can show that

$$\|\overline{\psi}\|_{\mathrm{L}^p(0,T;\mathrm{X})} \le \|\psi\|_{\mathrm{L}^p(0,T;\mathrm{X})} \qquad \text{for all } p \in [1,\infty], \tag{3.23}$$

$$\overline{\psi} \to \psi \qquad \text{strongly in } \mathrm{L}^p(0,T;\mathrm{X}), \quad \text{as } l \to \infty, \quad \text{for any } p \in [1,\infty). \tag{3.24}$$

The estimate (3.23) follows by Jensen's inequality, and the convergence in (3.24) is a consequence of the inequality $\|\psi - \overline{\psi}\|_{\mathrm{L}^p(0,T;\mathrm{X})} \le T^{\frac{1}{p}}\delta_l\|\psi\|_{\mathrm{C}^{0,1}([0,T];\mathrm{X})}$ for all $\psi \in \mathrm{C}^{0,1}([0,T];\mathrm{X})$ and $p \in [1,\infty]$, the density of $\mathrm{C}^{0,1}([0,T];\mathrm{X})$ in $\mathrm{L}^p(0,T;\mathrm{X})$ for $p \in [1,\infty)$, and (3.23).

To simplify the notation we denote $Q_s^t := (s,t)\times\Omega$, for $0 \le s < t \le T$, and $Q_s := Q_0^s$, for $s \in (0,T]$. Furthermore, let us introduce the notation $Q_{i-1}^i := Q_{t_{i-1}}^{t_i}$ and $Q_i := Q_{t_i}$, for $i \in \{1,\dots,l\}$.

4. Convergence proof

Motivated by the approach in Bulíček *et al.* (2012, Sec. 3.1), we consider the following levels of approximation.

$k \in \mathbb{N}$: The selection $\mathbf{S}^\star$ given in Lemma 2.2 is approximated by a family of Carathéodory functions $\{\mathbf{S}^k\}_{k\in\mathbb{N}}$, which satisfy Assumption 2.3. The approximation of the stress is then explicit and continuous in $\mathbf{D}\boldsymbol{u}$.

$l \in \mathbb{N}$: A time-stepping based on the implicit Euler method is introduced similarly as, e.g., in Temam (1984) and Carelli *et al.* (2010), see Subsection 3.3.

$n \in \mathbb{N}$: The velocity $\boldsymbol{u}$ is approximated by a Galerkin approximation in finite element spaces in the spatial variable, see Section 3.

$m \in \mathbb{N}$: The penalty/regularization term $\frac{1}{m}|\boldsymbol{u}|^{2q'-2}\boldsymbol{u}$ is added to the equation to gain admissibility of the approximate solutions in case we have $q < \frac{3d+2}{d+2}$, and to enable us to use the bound on $\widetilde{b}(\cdot,\cdot,\cdot)$ in (3.15), without imposing the restriction $q \ge \frac{2(d+1)}{d+2}$.

This results in a fully discrete approximation. The limits are taken in the order $k \to \infty$, $l, n \to \infty$, and then $m \to \infty$, and we can take the limits in $l, n \to \infty$ simultaneously. To simplify the notation we shall write

$$\boldsymbol{v}^{k,l,n,m} \underset{k}{\to} \underset{(l,n)}{\to} \underset{m}{\to} \boldsymbol{v} \quad \text{in } X, \quad \text{as } k \to \infty,\ l, n \to \infty,\ m \to \infty, \tag{4.1}$$

to denote the fact that the limits k, (l, n), m are taken successively in the order of indexing (from left to right) and the space X describes the weakest topology of the three limits. We will use the analogous notation for weak and weak* convergence. In each step one has to identify the equation and the implicit relation, which is the most challenging part. The most significant difference compared to Bulíček *et al.* (2012) lies in the passage to the limits $l, n \to \infty$ and the identification of the implicit law.

As both the external force $\boldsymbol{f}$ and the approximate stress $\mathbf{S}^k$ will be allowed to be time-dependent, and the time-dependence is not assumed to be continuous, we shall consider integral-averaged versions in the approximate problem. Recall the notation in Subsection 3.3 and for $\boldsymbol{f} \in \mathrm{L}^{q'}(0, T; \mathrm{W}^{-1,q'}(\Omega)^d)$ and $\mathbf{S}^k : Q \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$ as in Assumption 2.3, and $l \in \mathbb{N}$ let us introduce the averages with respect to the time grid $\{t_i\}_{i\in\{0,\ldots,l\}}$ defined, for $i \in \{1, \ldots, l\}$, by

$$\boldsymbol{f}_i(\boldsymbol{x}) := \fint_{t_{i-1}}^{t_i} \boldsymbol{f}(t, \boldsymbol{x})\,\mathrm{d}t, \qquad \mathbf{S}^k_i(\boldsymbol{x}, \mathbf{B}) := \fint_{t_{i-1}}^{t_i} \mathbf{S}^k(t, \boldsymbol{x}, \mathbf{B})\,\mathrm{d}t, \tag{4.2}$$

for $\boldsymbol{x} \in \Omega$ and $\mathbf{B} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$. Let the corresponding piecewise constant interpolants $\bar{\boldsymbol{f}}$ and $\bar{\mathbf{S}}^k$ be defined as in (3.17). Recall that by (3.23) and (3.24) we have that

$$\left\|\bar{\boldsymbol{f}}\right\|_{\mathrm{L}^{q'}(0,T;\mathrm{W}^{-1,q'}(\Omega))} \le \|\boldsymbol{f}\|_{\mathrm{L}^{q'}(0,T;\mathrm{W}^{-1,q'}(\Omega))} \quad \text{for all } \ l \in \mathbb{N}, \tag{4.3}$$

$$\bar{\boldsymbol{f}} \to \boldsymbol{f} \quad \text{strongly in } \mathrm{L}^{q'}(0, T; \mathrm{W}^{-1,q'}(\Omega)^d), \quad \text{as } l \to \infty. \tag{4.4}$$

For $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{V}^n$ we introduce

$$\mathfrak{L}^{k,l,n,m}_i[\boldsymbol{u}; \boldsymbol{v}] := -\,\widetilde{b}(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{v}) - \left\langle \mathbf{S}^k_i(\cdot, \mathbf{D}\boldsymbol{u}), \mathbf{D}\boldsymbol{v}\right\rangle_\Omega - \frac{1}{m}\left\langle |\boldsymbol{u}|^{2q'-2}\boldsymbol{u}, \boldsymbol{v}\right\rangle_\Omega + \langle \boldsymbol{f}_i, \boldsymbol{v}\rangle_\Omega, \tag{4.5}$$

for $k, l, n, m \in \mathbb{N}$, $i \in \{1, \ldots, l\}$ and $\widetilde{b}(\cdot, \cdot, \cdot)$ as defined in (3.12).

Approximate Problem. For $k, l, n, m \in \mathbb{N}$ find a sequence $\{\boldsymbol{U}^{k,l,n,m}_i\}_{i\in\{0,\ldots,l\}} \subset \mathbb{V}^n_{\mathrm{div}}$ such that

$$\boldsymbol{U}^{k,l,n,m}_0 = P^n_{\mathrm{div}}\boldsymbol{u}_0, \tag{4.6}$$

and for a given $\boldsymbol{U}^{k,l,n,m}_{i-1} \in \mathbb{V}^n_{\mathrm{div}}$ the approximate solution on the next time level, $\boldsymbol{U}^{k,l,n,m}_i \in \mathbb{V}^n_{\mathrm{div}}$, is defined, for $i \in \{1, \ldots, l\}$, by

$$\left\langle d_t \boldsymbol{U}^{k,l,n,m}_i, \boldsymbol{W}\right\rangle_\Omega = \mathfrak{L}^{k,l,n,m}_i\left[\boldsymbol{U}^{k,l,n,m}_i; \boldsymbol{W}\right] \qquad \text{for all } \boldsymbol{W} \in \mathbb{V}^n_{\mathrm{div}}, \tag{4.7}$$

where P^n_{div} is the L^2-projector onto $\mathbb{V}^n_{\mathrm{div}}$, defined in (3.8).

For each $i \in \{1, \ldots, l\}$ a fully implicit problem has to be solved, since the numerical solution from the previous time level only appears in the term involving $d_t \boldsymbol{U}_i^{k,l,n,m}$, as defined in (3.16).

THEOREM 4.1 (Main result). In addition to the assumptions of Definition 2.1, let $\mathbf{S}^k$ satisfy Assumption 2.3. For the finite element approximation let Assumption 3.1 on the domain and on the family of simplicial partitions be satisfied. Let $\mathbb{V}^n$ and let $\mathbb{V}^n_{\mathrm{div}}$ be as introduced in (3.1) and (3.3), respectively, and assume that Assumptions 3.2, 3.3 and 3.4 hold. Then, for all $k, l, n, m \in \mathbb{N}$ there exists a sequence $\{\boldsymbol{U}_i^{k,l,n,m}\}_{i\in\{0,\ldots,l\}} \subset \mathbb{V}^n_{\mathrm{div}}$ solving (4.6), (4.7). Moreover, if $q \in \left(\frac{2d}{d+2}, \infty\right)$, then there exists a weak solution $(\boldsymbol{u}, \mathbf{S})$ of **(P)** according to Definition 2.1 and for the piecewise constant interpolant $\overline{\boldsymbol{U}}^{k,l,n,m} \in \mathbb{P}^l_0(0, T; \mathbb{V}^n_{\mathrm{div}})$, and the continuous, piecewise affine interpolant $\widetilde{\boldsymbol{U}}^{k,l,n,m} \in \mathbb{P}^l_1(0, T; \mathbb{V}^n_{\mathrm{div}})$ of $\{\boldsymbol{U}_i^{k,l,n,m}\}_{i\in\{0,\ldots,l\}}$, and the piecewise constant interpolant $\overline{\mathbf{S}}^k(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^{k,l,n,m}) \in \mathbb{P}^l_0(0, T; \mathrm{L}^{q'}(\Omega)^{d\times d})$ of $\{\mathbf{S}_i^k(\cdot,\mathbf{D}\overline{\boldsymbol{U}}^{k,l,n,m})\}_{i\in\{1,\ldots,l\}}$ as defined in (3.17) and (3.18) (up to not relabelled subsequences), one has that

$$
\begin{aligned}
\overline{\boldsymbol{U}}^{k,l,n,m}, \widetilde{\boldsymbol{U}}^{k,l,n,m} &\underset{k}{\to} \underset{(l,n)}{\to} \underset{m}{\to} \boldsymbol{u} && \text{strongly in } \mathrm{L}^q(0, T; \mathrm{L}^2(\Omega)^d),\\
\overline{\boldsymbol{U}}^{k,l,n,m}, \widetilde{\boldsymbol{U}}^{k,l,n,m} &\underset{k}{\overset{*}{\rightharpoonup}} \underset{(l,n)}{\overset{*}{\rightharpoonup}} \underset{m}{\overset{*}{\rightharpoonup}} \boldsymbol{u} && \text{weakly in } \mathrm{L}^\infty(0, T; \mathrm{L}^2(\Omega)^d),\\
\overline{\boldsymbol{U}}^{k,l,n,m} &\underset{k}{\rightharpoonup} \underset{(l,n)}{\rightharpoonup} \underset{m}{\rightharpoonup} \boldsymbol{u} && \text{weakly in } \mathrm{L}^q\left(0, T; \mathrm{W}^{1,q}_0(\Omega)^d\right),\\
\overline{\mathbf{S}}^k(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^{k,l,n,m}),\ \mathbf{S}^k(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^{k,l,n,m}) &\underset{k}{\rightharpoonup} \underset{(l,n)}{\rightharpoonup} \underset{m}{\rightharpoonup} \mathbf{S} && \text{weakly in } \mathrm{L}^{q'}(Q)^{d\times d},
\end{aligned}
$$

as $k \to \infty$, $(l, n) \to \infty$ (combined) and $m \to \infty$, when taking the limits successively, without restrictions on the relation between the discretization parameters δ_l and h_n.

REMARK 4.2

(i) In the proof of Theorem 4.1 it is essential that the limits are taken in the indicated order.

(ii) If $\mathbf{S}^\star$ is a Carathéodory function, then the approximation level corresponding to $k \in \mathbb{N}$ can be skipped.

(iii) For Lipschitz polytopal domains, Theorem 4.1 is also a new existence result, since in Bulíček *et al.* (2012) a Navier slip boundary condition and $\partial\Omega \in \mathrm{C}^{1,1}$ are assumed.

(iv) The convergence proof is presented for discretely divergence-free velocity functions. If additionally $\mathbb{V}^n_{\mathrm{div}} \subset \mathrm{W}^{1,\infty}_{0,\mathrm{div}}(\Omega)^d$, then no modification of the convective term is required and the proof that $\boldsymbol{u}^m$ is divergence-free is also simpler.

The rest of this section consists of the proof of Theorem 4.1, which relies on Lemmas 4.3–4.5, dealing with the existence of the discrete solution, and the limit $k \to \infty$, Lemmas 4.6 and 4.7 covering the combined limit $l, n \to \infty$, and Lemmas 4.8 and 4.9 the limit $m \to \infty$. Note that Lemma 4.9 contains stronger statements regarding the weak solution than Definition 2.1.

Limit $k \to \infty$

The existence and convergence in Lemmas 4.3 and 4.4 follow by a standard approach presented, e.g., in Temam (1984), with minor modifications required to deal with the time-dependence of $\mathbf{S}^k$. Taking $k \to \infty$ we remain in the finite-dimensional setting, and hence strong convergence of the sequence of symmetric gradients follows. Consequently, the identification of the limiting equation is based on the properties of the sequence $\{\mathbf{S}^k\}_{k\in\mathbb{N}}$ according to Assumption 2.3, cf. Bulíček *et al.* (2012).

Lemma 4.3 (Existence of approximate solutions and *a priori* estimates). For each $\kappa := (k,l,n,m) \in \mathbb{N}^4$, there exists a sequence $\{\boldsymbol{U}_i^\kappa\}_{i\in\{0,\dots,l\}} \subset \mathbb{V}^n_{\mathrm{div}}$, which satisfies (4.6) and (4.7). Furthermore, there exists a constant $c > 0$ such that for all $\kappa = (k,l,n,m) \in \mathbb{N}^4$ one has that

$$\max_{j\in\{0,\dots,l\}} \left\|\boldsymbol{U}_j^\kappa\right\|^2_{\mathrm{L}^2(\Omega)} + \sum_{j=1}^{l} \left\|\boldsymbol{U}_j^\kappa - \boldsymbol{U}_{j-1}^\kappa\right\|^2_{\mathrm{L}^2(\Omega)} + \delta_l \sum_{j=1}^{l} \|\boldsymbol{U}_j^\kappa\|^q_{\mathrm{W}^{1,q}(\Omega)}$$
$$+ \sum_{j=1}^{l} \left\|\mathbf{S}^k(\cdot,\cdot,\mathbf{D}\boldsymbol{U}_j^\kappa)\right\|^{q'}_{\mathrm{L}^{q'}(Q^j_{j-1})} + \frac{\delta_l}{m}\sum_{j=1}^{l} \|\boldsymbol{U}_j^\kappa\|^{2q'}_{\mathrm{L}^{2q'}(\Omega)} \le c. \tag{4.8}$$

Proof.

Step 1: A priori *estimates*. The *a priori* estimates follow from standard arguments, see Temam (1984), in combination with the estimates for $\mathbf{S}^k$ by Assumption 2.3: testing (4.7) with $\boldsymbol{W} = \boldsymbol{U}_i^\kappa \in \mathbb{V}^n_{\mathrm{div}}$ one obtains

$$\langle d_t\boldsymbol{U}_i^\kappa, \boldsymbol{U}_i^\kappa\rangle_\Omega + \left\langle \mathbf{S}_i^k(\cdot,\mathbf{D}\boldsymbol{U}_i^\kappa), \mathbf{D}\boldsymbol{U}_i^\kappa\right\rangle_\Omega + \frac{1}{m}\|\boldsymbol{U}_i^\kappa\|^{2q'}_{\mathrm{L}^{2q'}(\Omega)} = \langle \boldsymbol{f}_i, \boldsymbol{U}_i^\kappa\rangle_\Omega, \tag{4.9}$$

since the term involving $\widetilde{b}$ vanishes by skew-symmetry. By the fact that $2a(a-b) = a^2 - b^2 + (a-b)^2$, for $a, b \in \mathbb{R}$ and by the definition of $\mathrm{d}_t\boldsymbol{U}_i^\kappa$ in (3.16), the first term in (4.9) can be rewritten as

$$\begin{aligned}\langle d_t\boldsymbol{U}_i^\kappa, \boldsymbol{U}_i^\kappa\rangle_\Omega &= \frac{1}{\delta_l}\langle \boldsymbol{U}_i^\kappa - \boldsymbol{U}_{i-1}^\kappa, \boldsymbol{U}_i^\kappa\rangle_\Omega \\ &= \frac{1}{2\delta_l}\left(\|\boldsymbol{U}_i^\kappa\|^2_{\mathrm{L}^2(\Omega)} - \|\boldsymbol{U}_{i-1}^\kappa\|^2_{\mathrm{L}^2(\Omega)} + \|\boldsymbol{U}_i^\kappa - \boldsymbol{U}_{i-1}^\kappa\|^2_{\mathrm{L}^2(\Omega)}\right).\end{aligned} \tag{4.10}$$

Using the definition of $\mathbf{S}_i^k$ in (4.2) and Assumption 2.3 (α2), one has that

$$\begin{aligned}\left\langle \mathbf{S}_i^k(\cdot,\mathbf{D}\boldsymbol{U}_i^\kappa), \mathbf{D}\boldsymbol{U}_i^\kappa\right\rangle_\Omega &\overset{(4.2)}{=} \left\langle ⨍_{t_{i-1}}^{t_i} \mathbf{S}^k(t,\cdot,\mathbf{D}\boldsymbol{U}_i^\kappa)\,\mathrm{d}t, \mathbf{D}\boldsymbol{U}_i^\kappa\right\rangle_\Omega = \frac{1}{\delta_l}\left\langle \mathbf{S}^k(\cdot,\cdot,\mathbf{D}\boldsymbol{U}_i^\kappa), \mathbf{D}\boldsymbol{U}_i^\kappa\right\rangle_{Q^i_{i-1}} \\ &\ge \frac{1}{\delta_l}\int_{Q^i_{i-1}} -|\widetilde{g}(\cdot)| + \widetilde{c}_*\left(|\mathbf{D}\boldsymbol{U}_i^\kappa|^q + \left|\mathbf{S}^k(\cdot,\cdot,\mathbf{D}\boldsymbol{U}_i^\kappa)\right|^{q'}\right)\mathrm{d}z \\ &\ge -\frac{1}{\delta_l}\|\widetilde{g}\|_{\mathrm{L}^1(Q^i_{i-1})} + c\left\|\boldsymbol{U}_i^\kappa\right\|^q_{\mathrm{W}^{1,q}(\Omega)} + \frac{\widetilde{c}_*}{\delta_l}\left\|\mathbf{S}^k(\cdot,\cdot,\mathbf{D}\boldsymbol{U}_i^\kappa)\right\|^{q'}_{\mathrm{L}^{q'}(Q^i_{i-1})},\end{aligned} \tag{4.11}$$

where the last inequality follows by Korn's and Poincaré's inequality. On the term on the right-hand side of (4.9), by duality of norms and by Young's inequality with $\varepsilon > 0$, we obtain that

$$\begin{aligned}\left\langle \boldsymbol{f}_i, \boldsymbol{U}_i^\kappa \right\rangle_\Omega &\le \left\|\boldsymbol{f}_i\right\|_{\mathrm{W}^{-1,q'}(\Omega)} \left\|\boldsymbol{U}_i^\kappa\right\|_{\mathrm{W}^{1,q}(\Omega)} \le c(\varepsilon) \left\|\boldsymbol{f}_i\right\|_{\mathrm{W}^{-1,q'}(\Omega)}^{q'} + \varepsilon \left\|\boldsymbol{U}_i^\kappa\right\|_{\mathrm{W}^{1,q}(\Omega)}^{q} \\ &\le \frac{c(\varepsilon)}{\delta_l} \|\boldsymbol{f}\|_{\mathrm{L}^{q'}(t_{i-1},t_i;\mathrm{W}^{-1,q'}(\Omega))}^{q'} + \varepsilon \left\|\boldsymbol{U}_i^\kappa\right\|_{\mathrm{W}^{1,q}(\Omega)}^{q},\end{aligned} \tag{4.12}$$

where the last inequality follows by (4.3). Applying the estimates (4.10)–(4.12) in (4.9), after rearranging, choosing $\varepsilon > 0$ small enough and multiplying by δ_l, we arrive at

$$\begin{aligned}&\left\|\boldsymbol{U}_i^\kappa\right\|_{\mathrm{L}^2(\Omega)}^2 - \left\|\boldsymbol{U}_{i-1}^\kappa\right\|_{\mathrm{L}^2(\Omega)}^2 + \left\|\boldsymbol{U}_i^\kappa - \boldsymbol{U}_{i-1}^\kappa\right\|_{\mathrm{L}^2(\Omega)}^2 \\ &\quad + \delta_l \left\|\boldsymbol{U}_i^\kappa\right\|_{\mathrm{W}^{1,q}(\Omega)}^{q} + \left\|\mathbf{S}^k\left(\cdot,\cdot,\mathbf{D}\boldsymbol{U}_i^\kappa\right)\right\|_{\mathrm{L}^{q'}(Q_{i-1}^i)}^{q'} + \frac{\delta_l}{m} \left\|\boldsymbol{U}_i^\kappa\right\|_{\mathrm{L}^{2q'}(\Omega)}^{2q'} \\ &\quad \le c\left(\|\boldsymbol{f}\|_{\mathrm{L}^{q'}(t_{i-1},t_i;\mathrm{W}^{-1,q'}(\Omega))}^{q'} + \|\widetilde{g}\|_{\mathrm{L}^1(Q_{i-1}^i)}\right).\end{aligned} \tag{4.13}$$

For arbitrary $j \in \{1, \ldots, l\}$, summing over $i \in \{1, \ldots, j\}$ yields

$$\begin{aligned}&\left\|\boldsymbol{U}_j^\kappa\right\|_{\mathrm{L}^2(\Omega)}^2 - \left\|\boldsymbol{U}_0^\kappa\right\|_{\mathrm{L}^2(\Omega)}^2 + \sum_{i=1}^{j} \left\|\boldsymbol{U}_i^\kappa - \boldsymbol{U}_{i-1}^\kappa\right\|_{\mathrm{L}^2(\Omega)}^2 \\ &\quad + \delta_l \sum_{i=1}^{j} \left\|\boldsymbol{U}_i^\kappa\right\|_{\mathrm{W}^{1,q}(\Omega)}^{q} + \sum_{i=1}^{j} \left\|\mathbf{S}^k(\cdot,\cdot,\mathbf{D}\boldsymbol{U}_i^\kappa)\right\|_{\mathrm{L}^{q'}(Q_{i-1}^i)}^{q'} + \frac{\delta_l}{m} \sum_{i=1}^{j} \left\|\boldsymbol{U}_i^\kappa\right\|_{\mathrm{L}^{2q'}(\Omega)}^{2q'} \\ &\quad \le c\left(\|\boldsymbol{f}\|_{\mathrm{L}^{q'}(0,T;\mathrm{W}^{-1,q'}(\Omega))}^{q'} + \|\widetilde{g}\|_{\mathrm{L}^1(Q)}\right),\end{aligned} \tag{4.14}$$

because of cancellation in the first term. Applying the estimate

$$\left\|\boldsymbol{U}_0^\kappa\right\|_{\mathrm{L}^2(\Omega)}^2 \overset{(4.6)}{=} \left\|P_{\mathrm{div}}^n \boldsymbol{u}_0\right\|_{\mathrm{L}^2(\Omega)}^2 \overset{(3.9)}{\le} \left\|\boldsymbol{u}_0\right\|_{\mathrm{L}^2(\Omega)}^2, \tag{4.15}$$

taking the supremum over all $j \in \{1, \ldots, l\}$ in (4.14), and using again (4.15) finishes the proof of (4.8).

Step 2: *Existence of* $\{\boldsymbol{U}_i^\kappa\}_{i\in\{0,\ldots,l\}}$. Let $\kappa \in \mathbb{N}^4$ be fixed. Since $\boldsymbol{U}_0^\kappa = P_{\mathrm{div}}^n \boldsymbol{u}_0$ by (4.6), one only has to show that for a given $\boldsymbol{U}_{i-1}^\kappa \in \mathbb{V}_{\mathrm{div}}^n$, there exists a $\boldsymbol{U}_i^\kappa \in \mathbb{V}_{\mathrm{div}}^n$ such that (4.7) is satisfied. Since $\mathbf{S}_i^k(z,\cdot)$ is continuous, the existence of such a $\boldsymbol{U}_i^\kappa \in \mathbb{V}_{\mathrm{div}}^n$ follows by a standard argument from Brouwer's fixed point theorem. For details we refer to Tscherpel (2018). Uniqueness is in general not guaranteed, so we choose one such sequence for each $\kappa \in \mathbb{N}^4$. □

For $t \in (0, T]$, $\boldsymbol{u} \in \mathbb{P}_0^l(0, T; \mathbb{V}_{\mathrm{div}}^n)$ and $\boldsymbol{v} \in \mathbb{V}_{\mathrm{div}}^n$ we introduce

$$\begin{aligned}\mathfrak{L}^\kappa[\boldsymbol{u}; \boldsymbol{v}](t) := &-\widetilde{b}(\boldsymbol{u}(t,\cdot), \boldsymbol{u}(t,\cdot), \boldsymbol{v}) - \left\langle \overline{\mathbf{S}}^k(t,\cdot,\mathbf{D}\boldsymbol{u}(t,\cdot)), \mathbf{D}\boldsymbol{v}\right\rangle_\Omega \\ &- \frac{1}{m}\left\langle |\boldsymbol{u}(t,\cdot)|^{2q'-2}\boldsymbol{u}(t,\cdot), \boldsymbol{v}\right\rangle_\Omega + \left\langle \overline{\boldsymbol{f}}(t,\cdot), \boldsymbol{v}\right\rangle_\Omega,\end{aligned} \tag{4.16}$$

for $\kappa = (k, l, n, m) \in \mathbb{N}^4$ and $\widetilde{b}(\cdot,\cdot,\cdot)$ as defined in (3.12). Recall that $\overline{\boldsymbol{f}} \in \mathbb{P}_0^l(0, T; \mathrm{W}^{-1,q'}(\Omega)^d)$ is the piecewise constant interpolant of $\{\boldsymbol{f}_i\}_{i\in\{1,\dots,l\}}$, as defined in (3.17) in Subsection 3.3 and, similarly, $\overline{\mathbf{S}}^k(t,\cdot,\cdot) = \mathbf{S}_i^k(\cdot,\cdot)$, for $t \in (t_{i-1}, t_i]$, which is piecewise constant with respect to the variable $t \in (0, T]$.

LEMMA 4.4 (Equation for $t \in (0, T]$ and Convergence $k \to \infty$). The functions $\overline{\boldsymbol{U}}^\kappa \in \mathbb{P}_0^l(0, T; \mathbb{V}_{\mathrm{div}}^n)$ and $\widetilde{\boldsymbol{U}}^\kappa \in \mathbb{P}_1^l(0, T; \mathbb{V}_{\mathrm{div}}^n)$ defined as piecewise constant and piecewise affine interpolants of $\{\boldsymbol{U}_i^\kappa\}_{i\in\{0,\dots,l\}}$ satisfy

$$\left\langle \partial_t \widetilde{\boldsymbol{U}}^\kappa(t,\cdot), \boldsymbol{W}\right\rangle_\Omega = \mathfrak{L}^\kappa[\overline{\boldsymbol{U}}^\kappa; \boldsymbol{W}](t) \quad \text{for all } \boldsymbol{W} \in \mathbb{V}_{\mathrm{div}}^n, \text{for all } t \in (0, T], \tag{4.17}$$

$$\widetilde{\boldsymbol{U}}^\kappa(0,\cdot) = P_{\mathrm{div}}^n \boldsymbol{u}_0 \qquad \text{in } \Omega, \tag{4.18}$$

for any $\kappa = (k, l, n, m) \in \mathbb{N}^4$. For each $\lambda := (l, n, m) \in \mathbb{N}^3$, i.e., $\kappa = (k, \lambda)$, there exists a sequence $\{\boldsymbol{U}_i^\lambda\}_{i\in\{0,\dots,l\}} \subset \mathbb{V}_{\mathrm{div}}^n$ and subsequences such that the interpolants $\overline{\boldsymbol{U}}^\lambda \in \mathbb{P}_0^l(0, T; \mathbb{V}_{\mathrm{div}}^n)$, and $\widetilde{\boldsymbol{U}}^\lambda \in \mathbb{P}_1^l(0, T; \mathbb{V}_{\mathrm{div}}^n)$ of $\{\boldsymbol{U}_i^\lambda\}_{i\in\{0,\dots,l\}}$, as defined in (3.17) and (3.18), satisfy

$$\sup_{t\in(0,T]} \left\| \overline{\boldsymbol{U}}^{k,\lambda}(t,\cdot) - \overline{\boldsymbol{U}}^\lambda(t,\cdot)\right\|_{\mathrm{W}^{1,\infty}(\Omega)} \to 0, \tag{4.19}$$

$$\sup_{t\in[0,T]} \left\| \widetilde{\boldsymbol{U}}^{k,\lambda}(t,\cdot) - \widetilde{\boldsymbol{U}}^\lambda(t,\cdot)\right\|_{\mathrm{W}^{1,\infty}(\Omega)} + \sup_{t\in(0,T]} \left\| \partial_t\widetilde{\boldsymbol{U}}^{k,\lambda}(t,\cdot) - \partial_t\widetilde{\boldsymbol{U}}^\lambda(t,\cdot)\right\|_{\mathrm{W}^{1,\infty}(\Omega)} \to 0, \tag{4.20}$$

as $k \to \infty$. Furthermore, for each $\lambda \in \mathbb{N}^3$ there exist $\mathbf{S}^\lambda \in \mathrm{L}^{q'}(Q)^{d\times d}$ and $\overline{\mathbf{S}}^\lambda \in \mathbb{P}_0^l(0, T; \mathrm{L}^{q'}(\Omega)^{d\times d})$ and subsequences such that

$$\mathbf{S}^k(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^{k,\lambda}) \rightharpoonup \mathbf{S}^\lambda \quad \text{weakly in } \mathrm{L}^{q'}(Q)^{d\times d}, \tag{4.21}$$

$$\overline{\mathbf{S}}^k(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^{k,\lambda}) \rightharpoonup \overline{\mathbf{S}}^\lambda \quad \text{weakly in } \mathrm{L}^{q'}(Q)^{d\times d}, \tag{4.22}$$

as $k \to \infty$, where, up to a representative, we have

$$\overline{\mathbf{S}}^\lambda(t,\cdot) = \mathbf{S}_i^\lambda(\cdot) := \fint_{t_{i-1}}^{t_i} \mathbf{S}^\lambda(t,\cdot)\,\mathrm{d}t \quad \text{for all } t \in (t_{i-1}, t_i] \text{ and all } i \in \{1,\dots,l\}. \tag{4.23}$$

Proof.

Step 1: *Identification of the equation*. We have that $\widetilde{\boldsymbol{U}}^{\kappa}(0,\cdot) = \boldsymbol{U}_0^{\kappa} = P^n_{\mathrm{div}}\boldsymbol{u}_0$ by definition of $\widetilde{\boldsymbol{U}}^{\kappa}$ and by (4.6), which shows (4.18). The equation (4.17) follows from (4.7) and the fact that for $t \in (t_{i-1}, t_i]$ we have that

$$\overline{\boldsymbol{U}}^{\kappa}(t,\cdot) = \boldsymbol{U}_i^{\kappa}, \quad \partial_t \widetilde{\boldsymbol{U}}^{\kappa}(t,\cdot) = d_t \boldsymbol{U}_i^{\kappa}, \quad \bar{\boldsymbol{f}}(t,\cdot) = \boldsymbol{f}_i \text{ and } \overline{\mathbf{S}}^k(t,\cdot,\cdot) = \mathbf{S}_i^k(\cdot,\cdot), \qquad i \in \{1,\dots,l\}.$$

Step 2: *Estimates*. Let $\lambda = (l,n,m) \in \mathbb{N}^3$ be arbitrary, but fixed. From the *a priori* estimate (4.8) it follows directly that

$$\left\| \mathbf{S}^k\big(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^{k,\lambda}\big) \right\|^{q'}_{\mathrm{L}^{q'}(Q)} \le c \quad \text{for all } k \in \mathbb{N}. \tag{4.24}$$

By the definition of $\mathbf{S}_i^k$ in (4.2), we have that

$$\begin{aligned}
\left\| \overline{\mathbf{S}}^k(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^{k,\lambda}) \right\|^{q'}_{\mathrm{L}^{q'}(Q)} &= \sum_{i=1}^{l} \left\| \mathbf{S}_i^k\big(\cdot,\mathbf{D}\boldsymbol{U}_i^{k,\lambda}\big) \right\|^{q'}_{\mathrm{L}^{q'}(Q^i_{i-1})} = \sum_{i=1}^{l} \delta_l \left\| \mathbf{S}_i^k\big(\cdot,\mathbf{D}\boldsymbol{U}_i^{k,\lambda}\big) \right\|^{q'}_{\mathrm{L}^{q'}(\Omega)} \\
&\overset{(3.23)}{\le} \sum_{i=1}^{l} \left\| \mathbf{S}^k\big(\cdot,\cdot,\mathbf{D}\boldsymbol{U}_i^{k,\lambda}\big) \right\|^{q'}_{\mathrm{L}^{q'}(Q^i_{i-1})} = \left\| \mathbf{S}^k(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^{k,\lambda}) \right\|^{q'}_{\mathrm{L}^{q'}(Q)} \overset{(4.24)}{\le} c
\end{aligned} \tag{4.25}$$

for all $k \in \mathbb{N}$. This also shows that

$$\left\| \mathbf{S}_i^k\big(\cdot,\mathbf{D}\boldsymbol{U}_i^{k,\lambda}\big) \right\|^{q'}_{\mathrm{L}^{q'}(\Omega)} \le \frac{c}{\delta_l} \le c(l) \quad \text{for any } k \in \mathbb{N} \text{ and any } i \in \{1,\dots,l\}. \tag{4.26}$$

Step 3: *Convergence as $k \to \infty$*. By the *a priori* estimate (4.8), the sequences $\{\overline{\boldsymbol{U}}^{k,\lambda}\}_{k\in\mathbb{N}} \subset \mathbb{P}^l_0(0,T;\mathbb{V}^n_{\mathrm{div}}) \subset \mathrm{L}^\infty(0,T;\mathrm{W}^{1,\infty}_0(\Omega)^d)$ and $\{\widetilde{\boldsymbol{U}}^{k,\lambda}\}_{k\in\mathbb{N}} \subset \mathbb{P}^l_1(0,T;\mathbb{V}^n_{\mathrm{div}}) \subset \mathrm{W}^{1,\infty}(0,T;\mathrm{W}^{1,\infty}_0(\Omega)^d)$ are bounded in the space $\mathrm{L}^\infty(0,T;\mathrm{L}^2(\Omega)^d)$. Since $\mathbb{P}^l_0(0,T;\mathbb{V}^n_{\mathrm{div}})$ and $\mathbb{P}^l_1(0,T;\mathbb{V}^n_{\mathrm{div}})$ are finite-dimensional function spaces, all norms on them are equivalent. Furthermore, any bounded sequence has a subsequence converging strongly in the respective norm. Hence, there exist $\overline{\boldsymbol{U}}^{\lambda} \in \mathbb{P}^l_0(0,T;\mathbb{V}^n_{\mathrm{div}})$, $\widetilde{\boldsymbol{U}}^{\lambda} \in \mathbb{P}^l_1(0,T;\mathbb{V}^n_{\mathrm{div}})$ and subsequences such that (4.19) and (4.20) are satisfied. Further, since the convergence is pointwise in time and $\widetilde{\boldsymbol{U}}^{k,\lambda}(t_i,\cdot) = \overline{\boldsymbol{U}}^{k,\lambda}(t_i,\cdot) = \boldsymbol{U}_i^{k,\lambda}$ for all $i \in \{1,\dots,l\}$ and $\widetilde{\boldsymbol{U}}^{k,\lambda}(0,\cdot) = P^n_{\mathrm{div}}\boldsymbol{u}_0$, it follows that $\widetilde{\boldsymbol{U}}^{\lambda}(t_i,\cdot) = \overline{\boldsymbol{U}}^{\lambda}(t_i,\cdot) =: \boldsymbol{U}_i^{\lambda}$ for all $i \in \{1,\dots,l\}$ and $\widetilde{\boldsymbol{U}}^{\lambda}(0,\cdot) = P^n_{\mathrm{div}}\boldsymbol{u}_0 =: \boldsymbol{U}_0^{\lambda}$. Consequently, $\overline{\boldsymbol{U}}^{\lambda}$ and $\widetilde{\boldsymbol{U}}^{\lambda}$ are the respective interpolants of $\{\boldsymbol{U}_i^{\lambda}\}_{i\in\{0,\dots,l\}} \subset \mathbb{V}^n_{\mathrm{div}}$. By the Banach–Alaoglu theorem, (4.24)–(4.26) imply that there exist $\mathbf{S}^{\lambda}, \overline{\mathbf{S}}^{\lambda} \in \mathrm{L}^{q'}(Q)^{d\times d}$ and $\mathbf{S}_i^{\lambda} \in \mathrm{L}^{q'}(\Omega)^{d\times d}$ for $i \in \{1,\dots,l\}$, and subsequences such that

$$\mathbf{S}^k(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^{k,\lambda}) \rightharpoonup \mathbf{S}^{\lambda} \quad \text{weakly in } \mathrm{L}^{q'}(Q)^{d\times d}, \tag{4.27}$$

$$\overline{\mathbf{S}}^k(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^{k,\lambda}) \rightharpoonup \overline{\mathbf{S}}^\lambda \quad \text{weakly in } \mathrm{L}^{q'}(Q)^{d\times d}, \tag{4.28}$$

$$\mathbf{S}_i^k\big(\cdot,\mathbf{D}\boldsymbol{U}_i^{k,\lambda}\big) \rightharpoonup \mathbf{S}_i^\lambda \quad \text{weakly in } \mathrm{L}^{q'}(\Omega)^{d\times d}, \quad \text{for } i\in\{1,\dots,l\}, \tag{4.29}$$

as $k\to\infty$. It remains to show the identification of $\mathbf{S}^\lambda$, $\overline{\mathbf{S}}^\lambda$ and $\{\mathbf{S}_i^\lambda\}_{i\in\{1,\dots,l\}}$. Let $i\in\{1,\dots,l\}$ be arbitrary, but fixed. First let $\varphi\in \mathrm{C}_0^\infty((t_{i-1},t_i))$ and $\mathbf{v}\in\mathrm{C}_0^\infty(\Omega)^{d\times d}$. On the one hand, by (4.28) we have

$$\left\langle \overline{\mathbf{S}}^k(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^{k,\lambda}),\varphi\mathbf{v}\right\rangle_{Q_{i-1}^i} \to \left\langle \overline{\mathbf{S}}^\lambda,\varphi\mathbf{v}\right\rangle_{Q_{i-1}^i}, \quad \text{as } k\to\infty. \tag{4.30}$$

On the other hand, by the definition of $\overline{\mathbf{S}}^k(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^{k,\lambda})$ as piecewise constant interpolant of the sequence $\{\mathbf{S}_i^k(\cdot,\mathbf{D}\boldsymbol{U}_i^{k,\lambda})\}_{i\in\{1,\dots,l\}}$ and by (4.29), we have

$$\begin{aligned}\left\langle \overline{\mathbf{S}}^k(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^{k,\lambda}),\varphi\mathbf{v}\right\rangle_{Q_{i-1}^i} &= \left\langle\left\langle \mathbf{S}_i^k\big(\cdot,\mathbf{D}\boldsymbol{U}_i^{k,\lambda}\big),\mathbf{v}\right\rangle_\Omega \varphi\right\rangle_{(t_{i-1},t_i)} = \langle 1,\varphi\rangle_{(t_{i-1},t_i)}\left\langle \mathbf{S}_i^k\big(\cdot,\mathbf{D}\boldsymbol{U}_i^{k,\lambda}\big),\mathbf{v}\right\rangle_\Omega\\ &\to \langle 1,\varphi\rangle_{(t_{i-1},t_i)}\left\langle \mathbf{S}_i^\lambda,\mathbf{v}\right\rangle_\Omega = \left\langle \mathbf{S}_i^\lambda,\mathbf{v}\varphi\right\rangle_{Q_{i-1}^i}, \quad \text{as } k\to\infty.\end{aligned} \tag{4.31}$$

Now, (4.30) and (4.31) imply, by the uniqueness of the limit, that $\overline{\mathbf{S}}^\lambda(t,\boldsymbol{x}) = \mathbf{S}_i^\lambda(\boldsymbol{x})$ for a.e. $(t,\boldsymbol{x})\in Q_{i-1}^i$, i.e., $\overline{\mathbf{S}}^\lambda$ is piecewise constant in t and we can choose the representative in $\mathbb{P}_0^l(0,T;\mathrm{L}^{q'}(\Omega)^{d\times d})$. Again, for $\mathbf{v}\in\mathrm{C}_0^\infty(\Omega)^{d\times d}$ we have by (4.29) that

$$\left\langle \mathbf{S}_i^k\big(\cdot,\mathbf{D}\boldsymbol{U}_i^{k,\lambda}\big),\mathbf{v}\right\rangle_\Omega \to \left\langle \mathbf{S}_i^\lambda,\mathbf{v}\right\rangle_\Omega, \qquad \text{as } k\to\infty. \tag{4.32}$$

On the other hand, by the definition of $\mathbf{S}_i^k(\cdot,\mathbf{D}\boldsymbol{U}_i^{k,\lambda})$ in (4.2) and by (4.27), we obtain that

$$\begin{aligned}\left\langle \mathbf{S}_i^k\big(\cdot,\mathbf{D}\boldsymbol{U}_i^{k,\lambda}\big),\mathbf{v}\right\rangle_\Omega &= \left\langle {\textstyle\fint_{t_{i-1}}^{t_i}} \mathbf{S}^k\big(t,\cdot,\mathbf{D}\boldsymbol{U}_i^{k,\lambda}\big)\,\mathrm{d}t,\mathbf{v}\right\rangle_\Omega = \frac{1}{\delta_l}\left\langle \mathbf{S}^k\big(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^{k,\lambda}\big),\mathbb{1}_{(t_{i-1},t_i)}\mathbf{v}\right\rangle_Q\\ &\to \frac{1}{\delta_l}\left\langle \mathbf{S}^\lambda,\mathbb{1}_{(t_{i-1},t_i)}\mathbf{v}\right\rangle_Q = \left\langle {\textstyle\fint_{t_{i-1}}^{t_i}} \mathbf{S}^\lambda(t,\cdot)\,\mathrm{d}t,\mathbf{v}\right\rangle_\Omega, \qquad \text{as } k\to\infty,\end{aligned} \tag{4.33}$$

so by the uniqueness of limits, we conclude from (4.32) and (4.33), that $\mathbf{S}_i^\lambda(\boldsymbol{x}) = \fint_{t_{i-1}}^{t_i}\mathbf{S}^\lambda(t,\boldsymbol{x})\,\mathrm{d}t$ for a.e. $\boldsymbol{x}\in\Omega$, which completes the proof. □

For $\lambda=(l,n,m)\in\mathbb{N}^3$, $t\in(0,T]$, $\boldsymbol{u}\in\mathbb{P}_0^l(0,T;\mathbb{V}^n)$ and $\boldsymbol{v}\in\mathbb{V}^n$, let us introduce

$$\begin{aligned}\mathfrak{L}^\lambda[\boldsymbol{u};\boldsymbol{v}](t) := &-\widetilde{b}(\boldsymbol{u}(t,\cdot),\boldsymbol{u}(t,\cdot),\boldsymbol{v}) - \left\langle \overline{\mathbf{S}}^\lambda(t,\cdot),\mathbf{D}\boldsymbol{v}\right\rangle_\Omega\\ &-\frac{1}{m}\left\langle |\boldsymbol{u}(t,\cdot)|^{2q'-2}\boldsymbol{u}(t,\cdot),\boldsymbol{v}\right\rangle_\Omega + \left\langle \overline{\boldsymbol{f}}(t,\cdot),\boldsymbol{v}\right\rangle_\Omega,\end{aligned} \tag{4.34}$$

where $\overline{\mathbf{S}}^\lambda\in\mathbb{P}_0^l(0,T;\mathrm{L}^{q'}(\Omega)^{d\times d})$ is given in Lemma 4.4.

LEMMA 4.5 (Identification of the PDE as $k \to \infty$). The functions $\overline{\boldsymbol{U}}^{\lambda} \in \mathbb{P}_0^l(0,T;\mathbb{V}^n_{\mathrm{div}})$, $\widetilde{\boldsymbol{U}}^{\lambda} \in \mathbb{P}_1^l(0,T;\mathbb{V}^n_{\mathrm{div}})$ and $\mathbf{S}^{\lambda} \in \mathrm{L}^{q'}(Q)^{d\times d}$ given in Lemma 4.4 satisfy

$$\left\langle \partial_t \widetilde{\boldsymbol{U}}^{\lambda}(t,\cdot), \boldsymbol{W}\right\rangle_{\Omega} = \mathfrak{L}^{\lambda}[\overline{\boldsymbol{U}}^{\lambda}; \boldsymbol{W}](t) \qquad \text{for all } \boldsymbol{W} \in \mathbb{V}^n_{\mathrm{div}}, \text{ for all } t \in (0,T], \tag{4.35}$$

$$\widetilde{\boldsymbol{U}}^{\lambda}(0,\cdot) = P^n_{\mathrm{div}}\boldsymbol{u}_0(\cdot) \qquad \text{in } \Omega, \tag{4.36}$$

$$(\mathbf{D}\overline{\boldsymbol{U}}^{\lambda}(z), \mathbf{S}^{\lambda}(z)) \in \mathcal{A}(z) \qquad \text{for a.e. } z \in Q, \tag{4.37}$$

for all $\lambda = (l,n,m) \in \mathbb{N}^3$, where $\mathfrak{L}^{\lambda}[\cdot;\cdot](\cdot)$ is defined by (4.34), using $\overline{\mathbf{S}}^{\lambda} \in \mathbb{P}_0^l(0,T;\mathrm{L}^{q'}(\Omega)^{d\times d})$ given by (4.23) in Lemma 4.4.

Proof. Let $\lambda = (l,n,m) \in \mathbb{N}^3$ be arbitrary but fixed. The fact that the initial condition (4.36) is satisfied follows directly by (4.18) and (4.20).

Step 1: *Identification of the limiting equation*. Let $\boldsymbol{W} \in \mathbb{V}^n_{\mathrm{div}}$ be arbitrary but fixed. With the convergence of $\partial_t \widetilde{\boldsymbol{U}}^{k,\lambda}$ in (4.20), it follows that

$$\left\langle \partial_t \widetilde{\boldsymbol{U}}^{k,\lambda}(t,\cdot), \boldsymbol{W}\right\rangle_{\Omega} \to \left\langle \partial_t \widetilde{\boldsymbol{U}}^{\lambda}(t,\cdot), \boldsymbol{W}\right\rangle_{\Omega}, \quad \text{as } k \to \infty, \tag{4.38}$$

for all $t \in (0,T]$. Further, by the strong convergence (4.19) and (4.20), it is straightforward to show that

$$\left\langle \mathfrak{L}^{k,\lambda}[\overline{\boldsymbol{U}}^{k,\lambda}, \boldsymbol{W}](t), \right\rangle_{(0,T)} \to \left\langle \mathfrak{L}^{\lambda}[\overline{\boldsymbol{U}}^{\lambda}, \boldsymbol{W}](t), \right\rangle_{(0,T)}, \quad \text{as } k \to \infty, \tag{4.39}$$

for all $t \in (0,T]$. In particular, the strong convergence in (4.19) allows us to take the limit in the numerical convective term without any restriction. Finally, (4.38) and (4.39) applied in (4.17) imply that (4.35) holds for all $t \in (0,T]$.

Step 2: *Identification of the implicit relation*. The proof of the implicit relation (4.37) relies on the strong convergence of $\{\mathbf{D}\overline{\boldsymbol{U}}^{k,\lambda}\}_{k\in\mathbb{N}}$ and the properties of $\mathbf{S}^k$ stated in Assumption 2.3. By the property (α3) in Assumption 2.3 on $\mathbf{S}^k$ and the boundedness of $\{\mathbf{D}\overline{\boldsymbol{U}}^{k,\lambda}\}_{k\in\mathbb{N}}$ in $\mathrm{L}^{\infty}(Q)^{d\times d}$ resulting from (4.19), we have

$$0 \le \liminf_{k\to\infty} \left\langle \mathbf{S}^k(\cdot, \mathbf{D}\overline{\boldsymbol{U}}^{k,\lambda}) - \mathbf{S}^{\star}(\cdot,\mathbf{B}), (\mathbf{D}\overline{\boldsymbol{U}}^{k,\lambda} - \mathbf{B})\varphi\right\rangle_Q \tag{4.40}$$

for all $\varphi \in \mathrm{C}_0^{\infty}(Q)$ such that $\varphi \ge 0$ and for all matrices $\mathbf{B} \in U$, for the dense set $U \subset \mathbb{R}^{d\times d}_{\mathrm{sym}}$ given in the assumption. Then, by the strong convergence of $\mathbf{D}\overline{\boldsymbol{U}}^{k,\lambda}$ in (4.19) and the weak convergence of $\mathbf{S}^k(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^{k,\lambda})$ in (4.21), we obtain

$$0 \le \liminf_{k\to\infty} \left\langle \mathbf{S}^k(\cdot, \mathbf{D}\overline{\boldsymbol{U}}^{k,\lambda}) - \mathbf{S}^{\star}(\cdot,\mathbf{B}), (\mathbf{D}\overline{\boldsymbol{U}}^{k,\lambda} - \mathbf{B})\varphi\right\rangle_Q = \left\langle \mathbf{S}^{\lambda} - \mathbf{S}^{\star}(\cdot,\mathbf{B}), (\mathbf{D}\overline{\boldsymbol{U}}^{\lambda} - \mathbf{B})\varphi\right\rangle_Q \tag{4.41}$$

for all $\varphi \in C_0^\infty(Q)$ such that $\varphi \geq 0$ and for all matrices $\mathbf{B} \in U$. By Lemma 2.2 (a4) this allows us to conclude that

$$\left(\mathbf{D}\overline{\boldsymbol{U}}^\lambda(z), \mathbf{S}^\lambda(z)\right) \in \mathcal{A}(z) \quad \text{for a.e. } z \in Q,$$

so (4.37) is shown. □

Limit $l, n \to \infty$

We are taking the limits $l, n \to \infty$ simultaneously without imposing any condition on δ_l and h_n. The condition $q > \frac{2d}{d+2}$ is required to gain compactness. Two additional difficulties, compared to Bulíček *et al.* (2012), arise from the discretization. The first is that in order to prove a uniform bound on the sequence of approximations to the time derivative, one would require the stability of the L^2-projector onto $\mathbb{V}^n_{\mathrm{div}}$ in Sobolev norms, which would impose stronger requirements on the finite element partition of Ω. To avoid this, instead of the Aubin–Lions lemma we shall employ an alternative compactness result due to Simon (cf. Lemma 2.10), which requires convergence properties of time-increments. The second difficulty is that, in the identification of the implicit relation, we have to deal with the discrepancy between $\overline{\mathbf{S}}^\lambda$ and $\mathbf{S}^\lambda$, since $\overline{\mathbf{S}}^\lambda$ appears in the equation (4.35) and $\mathbf{S}^\lambda$ satisfies the implicit relation in (4.37).

LEMMA 4.6 (Convergence as $l, n \to \infty$). Let the functions $\overline{\boldsymbol{U}}^{l,n,m} \in \mathbb{P}^l_0(0,T;\mathbb{V}^n_{\mathrm{div}})$, $\widetilde{\boldsymbol{U}}^{l,n,m} \in \mathbb{P}^l_1(0,T;\mathbb{V}^n_{\mathrm{div}})$, $\mathbf{S}^{l,n,m} \in L^{q'}(Q)^{d\times d}$ and $\overline{\mathbf{S}}^{l,n,m} \in \mathbb{P}^l_0(0,T;L^{q'}(\Omega)^{d\times d})$ satisfy (4.35)–(4.37), for any $l,n,m \in \mathbb{N}$, by Lemma 4.5. Further, let $\eta := \max\left(2q', \frac{q(d+2)}{d}\right) > 2$. For any $0 \leq s_0 < s \leq T$ and all $\lambda = (l,n,m) \in \mathbb{N}^3$, one has that

$$\frac{1}{2}\left\|\widetilde{\boldsymbol{U}}^\lambda(s,\cdot)\right\|^2_{L^2(\Omega)} + \left\langle \overline{\mathbf{S}}^\lambda, \mathbf{D}\overline{\boldsymbol{U}}^\lambda \right\rangle_{Q^s_{s_0}} + \frac{1}{m}\left\|\overline{\boldsymbol{U}}^\lambda\right\|^{2q'}_{L^{2q'}(Q^s_{s_0})} \leq \langle \bar{\boldsymbol{f}}, \overline{\boldsymbol{U}}^\lambda\rangle_{Q^s_{s_0}} + \frac{1}{2}\left\|\widetilde{\boldsymbol{U}}^\lambda(s_0,\cdot)\right\|^2_{L^2(\Omega)}. \tag{4.42}$$

Furthermore, for each $m \in \mathbb{N}$ there exists a $\boldsymbol{u}^m \in L^\infty(0,T;L^2_{\mathrm{div}}(\Omega)^d) \cap X_{\mathrm{div}}(Q)$, $\mathbf{S}^m \in L^{q'}(Q)^{d\times d}$ and subsequences such that, as $l, n \to \infty$,

$$\widetilde{\boldsymbol{U}}^{l,n,m} \to \boldsymbol{u}^m \quad \text{strongly in } L^p(0,T;L^2(\Omega)^d) \text{ for all } p \in [1,\infty), \tag{4.43}$$

$$\widetilde{\boldsymbol{U}}^{l,n,m}(s,\cdot) \to \boldsymbol{u}^m(s,\cdot) \quad \text{strongly in } L^2(\Omega)^d \text{ for a.e. } s \in (0,T), \tag{4.44}$$

$$\widetilde{\boldsymbol{U}}^{l,n,m}(0,\cdot) \to \boldsymbol{u}_0 \quad \text{strongly in } L^2(\Omega)^d, \tag{4.45}$$

$$\widetilde{\boldsymbol{U}}^{l,n,m}, \overline{\boldsymbol{U}}^{l,n,m} \overset{*}{\rightharpoonup} \boldsymbol{u}^m \quad \text{weakly* in } L^\infty(0,T;L^2(\Omega)^d), \tag{4.46}$$

$$\overline{\boldsymbol{U}}^{l,n,m} \to \boldsymbol{u}^m \quad \text{strongly in } L^p(0,T;L^2(\Omega)^d) \cap L^r(Q)^d \tag{4.47}$$

$$\text{for all } p \in [1,\infty) \text{ and all } r \in [1,\eta),$$

$$\overline{\boldsymbol{U}}^{l,n,m}(s,\cdot) \to \boldsymbol{u}^m(s,\cdot) \quad \text{strongly in } L^2(\Omega)^d \text{ for a.e. } s \in (0,T), \tag{4.48}$$

$$\overline{\boldsymbol{U}}^{l,n,m} \rightharpoonup \boldsymbol{u}^m \quad \text{weakly in } L^q\left(0,T;W^{1,q}_0(\Omega)^d\right) \cap L^\eta(Q)^d, \tag{4.49}$$

$$\big|\overline{\boldsymbol{U}}^{l,n,m}\big|^{2q'-2}\overline{\boldsymbol{U}}^{l,n,m} \rightharpoonup \big|\boldsymbol{u}^m\big|^{2q'-2}\boldsymbol{u}^m \quad \text{weakly in } \mathrm{L}^{(2q')'}(Q)^d, \tag{4.50}$$

$$\overline{\mathbf{S}}^{l,n,m} \rightharpoonup \mathbf{S}^m \quad \text{weakly in } \mathrm{L}^{q'}(Q)^{d\times d}, \tag{4.51}$$

$$\mathbf{S}^{l,n,m} \rightharpoonup \mathbf{S}^m \quad \text{weakly in } \mathrm{L}^{q'}(Q)^{d\times d}. \tag{4.52}$$

Proof.

Step 1: *Energy inequality*. Let $\lambda = (l,n,m) \in \mathbb{N}^3$, $i \in \{1,\dots,l\}$ and let $t \in (t_{i-1},t_i]$. In (4.35) we test with $\boldsymbol{W} = \overline{\boldsymbol{U}}^\lambda(t,\cdot) \in \mathbb{V}^n_{\mathrm{div}}$. For the first term adding and subtracting $\widetilde{\boldsymbol{U}}^\lambda(t,\cdot)$ with (3.19) we obtain

$$\begin{aligned}\left\langle \partial_t\widetilde{\boldsymbol{U}}^\lambda(t,\cdot), \overline{\boldsymbol{U}}^\lambda(t,\cdot)\right\rangle_\Omega &= \left\langle \partial_t\widetilde{\boldsymbol{U}}^\lambda(t,\cdot), \widetilde{\boldsymbol{U}}^\lambda(t,\cdot)\right\rangle_\Omega + \left\langle \partial_t\widetilde{\boldsymbol{U}}^\lambda(t,\cdot), \overline{\boldsymbol{U}}^\lambda(t,\cdot) - \widetilde{\boldsymbol{U}}^\lambda(t,\cdot)\right\rangle_\Omega \\ &= \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left\|\widetilde{\boldsymbol{U}}^\lambda(t,\cdot)\right\|^2_{\mathrm{L}^2(\Omega)} + (t_i - t)\left\|\partial_t\widetilde{\boldsymbol{U}}^\lambda(t,\cdot)\right\|^2_{\mathrm{L}^2(\Omega)} \\ &\geq \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left\|\widetilde{\boldsymbol{U}}^\lambda(t,\cdot)\right\|^2_{\mathrm{L}^2(\Omega)},\end{aligned} \tag{4.53}$$

since $t \le t_i$. By the continuity of $\widetilde{\boldsymbol{U}}^\lambda$, upon integration over (s_0, s), for $0 \le s_0 < s \le T$, this yields

$$\int_{s_0}^{s} \left\langle \partial_t\widetilde{\boldsymbol{U}}^\lambda(t,\cdot), \overline{\boldsymbol{U}}^\lambda(t,\cdot)\right\rangle_\Omega \geq \frac{1}{2}\left\|\widetilde{\boldsymbol{U}}^\lambda(s,\cdot)\right\|^2_{\mathrm{L}^2(\Omega)} - \frac{1}{2}\left\|\widetilde{\boldsymbol{U}}^\lambda(s_0,\cdot)\right\|^2_{\mathrm{L}^2(\Omega)}. \tag{4.54}$$

The other terms follow immediately and (4.42) is proved.

Step 2: *Estimates*. By the weak convergence in (4.21) and (4.22), the estimates (4.24) and (4.25) uniformly in $\lambda = (l,n,m) \in \mathbb{N}^3$, and the lower semicontinuity of the norms with respect to weak convergence, we obtain

$$\|\overline{\mathbf{S}}^\lambda\|^{q'}_{\mathrm{L}^{q'}(Q)} + \|\mathbf{S}^\lambda\|^{q'}_{\mathrm{L}^{q'}(Q)} \le c \quad \text{for all } \lambda \in \mathbb{N}^3. \tag{4.55}$$

By (4.20) we have in particular that $\boldsymbol{U}_i^{k,\lambda} \to \boldsymbol{U}_i^\lambda$ strongly in $\mathrm{W}^{1,\infty}_0(\Omega)^d$, as $k \to \infty$ for any $i \in \{1,\dots,l\}$. By the lower semicontinuity of the norm function with respect to the convergence in the respective norm we deduce from (4.8), which is uniform in $\lambda = (l,n,m) \in \mathbb{N}^3$, that

$$\max_{j\in\{0,\dots,l\}} \|\boldsymbol{U}_j^\lambda\|^2_{\mathrm{L}^2(\Omega)} + \sum_{j=1}^{l} \|\boldsymbol{U}_j^\lambda - \boldsymbol{U}_{j-1}^\lambda\|^2_{\mathrm{L}^2(\Omega)} + \delta_l \sum_{j=1}^{l} \|\boldsymbol{U}_j^\lambda\|^q_{\mathrm{W}^{1,q}(\Omega)} + \frac{\delta_l}{m}\sum_{j=1}^{l} \|\boldsymbol{U}_j^\lambda\|^{2q'}_{\mathrm{L}^{2q'}(\Omega)} \le c, \tag{4.56}$$

for all $\lambda = (l, n, m) \in \mathbb{N}^3$. By the definition of the piecewise constant interpolant according to (3.17), it follows from the discrete estimate that

$$\left\|\overline{\boldsymbol{U}}^{\lambda}\right\|_{\mathrm{L}^{\infty}(0,T;\mathrm{L}^2(\Omega))} + \left\|\overline{\boldsymbol{U}}^{\lambda}\right\|^q_{\mathrm{L}^q(0,T;\mathrm{W}^{1,q}(\Omega))} + \frac{1}{m}\left\|\overline{\boldsymbol{U}}^{\lambda}\right\|^{2q'}_{\mathrm{L}^{2q'}(Q)} \overset{(4.56)}{\leq} c, \tag{4.57}$$

for all $\lambda = (l, n, m) \in \mathbb{N}^3$. With this and the parabolic interpolation from Lemma 2.9, we have that

$$\left\|\overline{\boldsymbol{U}}^{\lambda}\right\|_{\mathrm{L}^{\frac{q(d+2)}{d}}(Q)} \leq c \quad \text{for all } \lambda = (l, n, m) \in \mathbb{N}^3. \tag{4.58}$$

For the estimates of the continuous, piecewise affine interpolant $\widetilde{\boldsymbol{U}}^{\lambda}$ according to (3.18), one also has to estimate the corresponding norms of $\boldsymbol{U}_0^{\lambda}$. In the $\mathrm{L}^2(\Omega)^d$ norm this is given by (4.56), and hence we have that

$$\left\|\widetilde{\boldsymbol{U}}^{\lambda}\right\|_{\mathrm{L}^{\infty}(0,T;\mathrm{L}^2(\Omega))} \overset{(4.56)}{\leq} c \qquad \text{for all } \lambda \in \mathbb{N}^3. \tag{4.59}$$

Since for smaller function spaces on Ω the corresponding estimate is not available, for the compactness argument we consider, instead, $\widehat{\boldsymbol{U}}^{\lambda} \in \mathrm{C}([0,T]; \mathbb{V}^n_{\mathrm{div}})$ defined by

$$\widehat{\boldsymbol{U}}^{\lambda}(t,\cdot) := \begin{cases} \widetilde{\boldsymbol{U}}^{\lambda}(t,\cdot) & \text{if } t \in (\delta_l, T], \\ \overline{\boldsymbol{U}}^{\lambda}(t,\cdot) = \boldsymbol{U}_1^{\lambda}(\cdot) & \text{if } t \in [0, \delta_l]. \end{cases} \tag{4.60}$$

This function is constant on $[0, \delta_l]$ and satisfies for $r \in [1, \infty)$ and a normed space X the bound

$$\left\|\widehat{\boldsymbol{U}}^{\lambda}\right\|^r_{\mathrm{L}^r(0,T;\mathrm{X})} \leq c\delta_l \sum_{i=1}^{l} \left\|\boldsymbol{U}_i^{\lambda}\right\|^r_{\mathrm{L}^r(0,T;\mathrm{X})} = c\left\|\overline{\boldsymbol{U}}^{\lambda}\right\|^r_{\mathrm{L}^r(0,T;\mathrm{X})}, \tag{4.61}$$

and an analogous estimate holds for $r = \infty$. Then, by (4.57) it follows that

$$\left\|\widehat{\boldsymbol{U}}^{\lambda}\right\|_{\mathrm{L}^{\infty}(0,T;\mathrm{L}^2(\Omega))} + \left\|\widehat{\boldsymbol{U}}^{\lambda}\right\|^q_{\mathrm{L}^q(0,T;\mathrm{W}^{1,q}(\Omega))} + \frac{1}{m}\left\|\widehat{\boldsymbol{U}}^{\lambda}\right\|^{2q'}_{\mathrm{L}^{2q'}(Q)} \leq c \quad \text{for all } \lambda = (l, n, m) \in \mathbb{N}^3. \tag{4.62}$$

By the fact that $\partial_t \widetilde{\boldsymbol{U}}^{\lambda}(t,\cdot) = d_t \boldsymbol{U}_i^{\lambda}$, for $t \in (t_{i-1}, t_i]$, $i \in \{1, \ldots, l\}$, and by (4.56) we obtain

$$\delta_l \left\|\partial_t \widetilde{\boldsymbol{U}}^{\lambda}\right\|^2_{\mathrm{L}^2(Q)} = \delta_l \sum_{i=1}^{l} \left\|\frac{1}{\delta_l}(\boldsymbol{U}_i^{\lambda} - \boldsymbol{U}_{i-1}^{\lambda})\right\|^2_{\mathrm{L}^2(Q^i_{i-1})} = \sum_{i=1}^{l} \left\|\boldsymbol{U}_i^{\lambda} - \boldsymbol{U}_{i-1}^{\lambda}\right\|^2_{\mathrm{L}^2(\Omega)} \overset{(4.56)}{\leq} c \tag{4.63}$$

for all $\lambda \in \mathbb{N}^3$.

Finally, we also estimate $\mathfrak{L}^{\lambda}[\boldsymbol{u};\boldsymbol{v}](t)$, as defined in (4.34): by (3.15), duality of norms and Hölder's and Poincaré's inequality, we obtain

$$\begin{aligned}
\int_a^b \mathfrak{L}^{\lambda}[\boldsymbol{u};\boldsymbol{v}](t)\,\mathrm{d}t &= \left\langle \widetilde{b}(\boldsymbol{u}(t,\cdot),\boldsymbol{u}(t,\cdot),\boldsymbol{v})\right\rangle_{(a,b)} - \left\langle \overline{\mathbf{S}}^{\lambda}(t,\cdot),\mathbf{D}\boldsymbol{v}\right\rangle_{Q_a^b} \\
&\quad - \frac{1}{m}\left\langle |\boldsymbol{u}(t,\cdot)|^{2q'-2}\boldsymbol{u}(t,\cdot),\boldsymbol{v}\right\rangle_{Q_a^b} + \left\langle \overline{\boldsymbol{f}}(t,\cdot),\boldsymbol{v}\right\rangle_{Q_a^b} \\
&\leq \|\boldsymbol{u}\|^2_{\mathrm{L}^{2q'}(Q_a^b)}\|\nabla\boldsymbol{v}\|_{\mathrm{L}^{q}(Q_a^b)} + \|\boldsymbol{u}\|_{\mathrm{L}^{2q'}(Q_a^b)}\|\nabla\boldsymbol{u}\|_{\mathrm{L}^{q}(Q_a^b)}\|\boldsymbol{v}\|_{\mathrm{L}^{2q'}(Q_a^b)} \\
&\quad + \left\|\overline{\mathbf{S}}^{\lambda}\right\|_{\mathrm{L}^{q'}(Q_a^b)}\|\mathbf{D}\boldsymbol{v}\|_{\mathrm{L}^{q}(Q_a^b)} + \frac{1}{m}\|\boldsymbol{u}\|^{2q'-1}_{\mathrm{L}^{2q'}(Q_a^b)}\|\boldsymbol{v}\|_{\mathrm{L}^{2q'}(Q_a^b)} \\
&\quad + \left\|\overline{\boldsymbol{f}}\right\|_{\mathrm{L}^{q'}(a,b;\mathrm{W}^{-1,q'}(\Omega))}\|\boldsymbol{v}\|_{\mathrm{L}^{q}(a,b;\mathrm{W}^{1,q}(\Omega))} \\
&\leq c\left(1+\|\boldsymbol{u}\|^2_{\mathrm{L}^{2q'}(Q_a^b)}\right)\|\nabla\boldsymbol{v}\|_{\mathrm{L}^{q}(Q_a^b)} \\
&\quad + c\left(\|\boldsymbol{u}\|_{\mathrm{L}^{2q'}(Q_a^b)}\|\nabla\boldsymbol{u}\|_{\mathrm{L}^{q}(Q_a^b)} + \frac{1}{m}\|\boldsymbol{u}\|^{2q'-1}_{\mathrm{L}^{2q'}(Q_a^b)}\right)\|\boldsymbol{v}\|_{\mathrm{L}^{2q'}(Q_a^b)},
\end{aligned} \tag{4.64}$$

for $0 \leq a < b \leq T$, for any $\lambda = (l,n,m) \in \mathbb{N}^3$, where we have used the estimate (4.55) on $\overline{\mathbf{S}}^{\lambda}$ and (4.3) on $\overline{\boldsymbol{f}}$. With the estimates on $\overline{\boldsymbol{U}}^{\lambda}$ in (4.57) this yields

$$\begin{aligned}
\int_a^b \mathfrak{L}^{\lambda}[\overline{\boldsymbol{U}}^{\lambda};\boldsymbol{v}](t)\,\mathrm{d}t &\overset{(4.64)}{\leq} c\left(1+\left\|\overline{\boldsymbol{U}}^{\lambda}\right\|^2_{\mathrm{L}^{2q'}(Q_a^b)}\right)\|\nabla\boldsymbol{v}\|_{\mathrm{L}^{q}(Q_a^b)} \\
&\quad + c\left(\left\|\overline{\boldsymbol{U}}^{\lambda}\right\|_{\mathrm{L}^{2q'}(Q_a^b)}\left\|\nabla\overline{\boldsymbol{U}}^{\lambda}\right\|_{\mathrm{L}^{q}(Q_a^b)} + \frac{1}{m}\left\|\overline{\boldsymbol{U}}^{\lambda}\right\|^{2q'-1}_{\mathrm{L}^{2q'}(Q_a^b)}\right)\|\boldsymbol{v}\|_{\mathrm{L}^{2q'}(Q_a^b)} \\
&\overset{(4.57)}{\leq} c(m)\left(\|\nabla\boldsymbol{v}\|_{\mathrm{L}^{q}(Q_a^b)} + \|\boldsymbol{v}\|_{\mathrm{L}^{2q'}(Q_a^b)}\right),
\end{aligned} \tag{4.65}$$

for $0 \leq a < b \leq T$ and any $\lambda = (l,n,m) \in \mathbb{N}^3$.

Step 3: *Convergence of the time increments* (cf. Carelli *et al.*, (2010, pp. 174). Instead of applying the Aubin–Lions lemma, as in Bulíček *et al.* (2012), here we apply the compactness result due to Simon, stated in Lemma 2.10. This means that we do not need uniform bounds on the time derivatives, but only convergence properties for time increments, which avoids the use of stability results in Sobolev norms for the L^2-projector onto $\mathbb{V}^n_{\mathrm{div}}$. We wish to apply Lemma 2.10 to the sequence $\{\widehat{\boldsymbol{U}}^{l,n,m}\}_{l,n\in\mathbb{N}}$, for fixed $m \in \mathbb{N}$, with $X = \mathrm{W}^{1,q}(\Omega)^d$, $B = \mathrm{L}^2(\Omega)^d$ and $p = 2$. Let us show that

$$\int_0^{T-\varepsilon}\left\|\widehat{\boldsymbol{U}}^{\lambda}(s+\varepsilon,\cdot)-\widehat{\boldsymbol{U}}^{\lambda}(s,\cdot)\right\|^2_{\mathrm{L}^2(\Omega)}\,\mathrm{d}s \to 0, \quad \text{as } \varepsilon \to 0, \quad \text{uniformly for } l,n \in \mathbb{N}. \tag{4.66}$$

Consider the term $\langle \widehat{\boldsymbol{U}}^{\lambda}(s+\varepsilon,\cdot) - \widehat{\boldsymbol{U}}^{\lambda}(s,\cdot), \boldsymbol{W}\rangle_{\Omega}$, for $\boldsymbol{W} \in \mathbb{V}^n_{\mathrm{div}}$, $s \in (0,T)$ and $\varepsilon > 0$ such that $s+\varepsilon < T$. If $s+\varepsilon \leq \delta_l$, then we have $\widehat{\boldsymbol{U}}^{\lambda}(s+\varepsilon) = \widehat{\boldsymbol{U}}^{\lambda}(s) = \boldsymbol{U}^{\lambda}_1$, so the term vanishes. Now let $s+\varepsilon > \delta_l$. By the definition of $\widehat{\boldsymbol{U}}^{\lambda}$ in (4.60) we have that $\widehat{\boldsymbol{U}}^{\lambda}(s,\cdot) = \widehat{\boldsymbol{U}}^{\lambda}(\max(s,\delta_l),\cdot)$. By the continuity of $\widehat{\boldsymbol{U}}^{\lambda}$ and since $\partial_t \widehat{\boldsymbol{U}}^{\lambda}$ is integrable, we obtain

$$\left\langle \widehat{\boldsymbol{U}}^{\lambda}(s+\varepsilon,\cdot) - \widehat{\boldsymbol{U}}^{\lambda}(s,\cdot), \boldsymbol{W}\right\rangle_{\Omega} = \int_{\max(s,\delta_l)}^{s+\varepsilon} \left\langle \partial_t \widehat{\boldsymbol{U}}^{\lambda}(t,\cdot), \boldsymbol{W}\right\rangle_{\Omega} \,\mathrm{d}t = \int_{\max(s,\delta_l)}^{s+\varepsilon} \left\langle \partial_t \widetilde{\boldsymbol{U}}^{\lambda}(t,\cdot), \boldsymbol{W}\right\rangle_{\Omega} \,\mathrm{d}t, \quad (4.67)$$

where in the last line we have used that $\widehat{\boldsymbol{U}}^{\lambda}(t,\cdot)$ and $\widetilde{\boldsymbol{U}}^{\lambda}(t,\cdot)$ coincide on $(\max(s,\delta_l), s+\varepsilon) \subset [\delta_l, T]$. Applying the equation (4.35) for a.e. $t \in (\max(s,\delta_l), s+\varepsilon)$, integrating and applying the bounds in (4.65) yields

$$\begin{aligned} \int_{\max(s,\delta_l)}^{s+\varepsilon} \left\langle \partial_t \widetilde{\boldsymbol{U}}^{\lambda}(t,\cdot), \boldsymbol{W}\right\rangle_{\Omega} \,\mathrm{d}t &= \int_{\max(s,\delta_l)}^{s+\varepsilon} \mathfrak{L}^{\lambda}[\overline{\boldsymbol{U}}^{\lambda}; \boldsymbol{W}](t)\,\mathrm{d}t \\ &\overset{(4.65)}{\leq} c(m)\left(\|\nabla \boldsymbol{W}\|_{\mathrm{L}^q\left(Q^{s+\varepsilon}_{\max(s,\delta_l)}\right)} + \|\boldsymbol{W}\|_{\mathrm{L}^{2q'}\left(Q^{s+\varepsilon}_{\max(s,\delta_l)}\right)}\right) \\ &= c(m)\left(\varepsilon^{\frac{1}{q}} + \varepsilon^{\frac{1}{2q'}}\right)\|\boldsymbol{W}\|_{\mathrm{X}(\Omega)}, \end{aligned} \quad (4.68)$$

since $\boldsymbol{W}$ is constant in time and the length of the time interval is bounded by ε.

For all $s \in (0,T)$ and $\varepsilon > 0$ such that $s+\varepsilon < T$, we have that $\widehat{\boldsymbol{U}}^{\lambda}(s+\varepsilon,\cdot), \widehat{\boldsymbol{U}}^{\lambda}(s,\cdot) \in \mathbb{V}^n_{\mathrm{div}}$; applying (4.67) and (4.68) with $\boldsymbol{W} = \widehat{\boldsymbol{U}}^{\lambda}(s+\varepsilon,\cdot) - \widehat{\boldsymbol{U}}^{\lambda}(s,\cdot)$, which is piecewise constant in time, shows that

$$\left\|\widehat{\boldsymbol{U}}^{\lambda}(s+\varepsilon,\cdot) - \widehat{\boldsymbol{U}}^{\lambda}(s,\cdot)\right\|^2_{\mathrm{L}^2(\Omega)} \leq c(m)\left(\varepsilon^{\frac{1}{q}} + \varepsilon^{\frac{1}{2q'}}\right)\left\|\widehat{\boldsymbol{U}}^{\lambda}(s+\varepsilon,\cdot) - \widehat{\boldsymbol{U}}^{\lambda}(s,\cdot)\right\|_{\mathrm{X}(\Omega)}. \quad (4.69)$$

Integrating over $(0, T-\varepsilon)$, using the triangle inequality, Hölder's inequality and the estimate in (4.62) yields

$$\begin{aligned} &\int_0^{T-\varepsilon} \left\|\widehat{\boldsymbol{U}}^{\lambda}(s+\varepsilon,\cdot) - \widehat{\boldsymbol{U}}^{\lambda}(s,\cdot)\right\|^2_{\mathrm{L}^2(\Omega)} \mathrm{d}s \\ &\overset{(4.69)}{\leq} c(m)\left(\varepsilon^{\frac{1}{q}} + \varepsilon^{\frac{1}{2q'}}\right)\int_0^{T-\varepsilon}\left(\left\|\widehat{\boldsymbol{U}}^{\lambda}(s+\varepsilon,\cdot)\right\|_{\mathrm{X}(\Omega)} + \left\|\widehat{\boldsymbol{U}}^{\lambda}(s,\cdot)\right\|_{\mathrm{X}(\Omega)}\right)\mathrm{d}s \\ &\leq c(m)\left(\varepsilon^{\frac{1}{q}} + \varepsilon^{\frac{1}{2q'}}\right)\left\|\widehat{\boldsymbol{U}}^{\lambda}\right\|_{\mathrm{X}(Q)} \overset{(4.62)}{\leq} c(m)(\varepsilon^{\frac{1}{q}} + \varepsilon^{\frac{1}{2q'}}) \to 0, \end{aligned} \quad (4.70)$$

as $\varepsilon \to 0$ uniformly in $l, n \in \mathbb{N}$, where $\lambda = (l,n,m) \in \mathbb{N}^3$. This proves (4.66).

Step 4: *Convergence as $l, n \to \infty$.* Recall that we have $\lambda = (l,n,m) \in \mathbb{N}^3$ and let $m \in \mathbb{N}$ be fixed. By estimate (4.62) we have that $\{\widehat{\boldsymbol{U}}^{l,n,m}\}_{l,n\in\mathbb{N}}$ is bounded in particular in $\mathrm{L}^2(Q)^d$ and $\mathrm{L}^1(0,T;\mathrm{W}^{1,q}(\Omega)^d)$. By the condition that $q > \frac{2d}{d+2}$, the embedding $\mathrm{W}^{1,q}(\Omega) \hookrightarrow\hookrightarrow \mathrm{L}^2(\Omega)$ is compact and with (4.66) all the assumptions in Lemma 2.10 are satisfied for $X = \mathrm{W}^{1,q}(\Omega)^d$, $B = \mathrm{L}^2(\Omega)^d$ and $p = 2$. Hence, there exists $\boldsymbol{u}^m \in \mathrm{L}^2(Q)^d$ and a subsequence such that

$$\widehat{\boldsymbol{U}}^{l,n,m} \to \boldsymbol{u}^m \quad \text{strongly in } \mathrm{L}^2(Q)^d, \quad \text{as } l, n \to \infty. \tag{4.71}$$

By the definition of $\widehat{\boldsymbol{U}}^{l,n,m}$ in (4.60) and the property (3.19) of the interpolants defined in (3.17) and (3.18), we have that

$$\begin{aligned} \left\|\widehat{\boldsymbol{U}}^{l,n,m} - \widetilde{\boldsymbol{U}}^{l,n,m}\right\|^2_{\mathrm{L}^2(Q)} &= \left\|\overline{\boldsymbol{U}}^{l,n,m} - \widetilde{\boldsymbol{U}}^{l,n,m}\right\|^2_{\mathrm{L}^2(0,\delta_l;\mathrm{L}^2(\Omega))} \\ &\overset{(3.19)}{\leq} \left\|(\delta_l - t)\partial_t \widetilde{\boldsymbol{U}}^{l,n,m}\right\|^2_{\mathrm{L}^2(0,\delta_l;\mathrm{L}^2(\Omega))} \leq \delta_l^2 \left\|\partial_t \widetilde{\boldsymbol{U}}^{l,n,m}\right\|^2_{\mathrm{L}^2(0,\delta_l;\mathrm{L}^2(\Omega))} \\ &\overset{(4.63)}{\leq} c\delta_l \to 0, \quad \text{as } l \to \infty. \end{aligned} \tag{4.72}$$

With (4.71) it follows that $\widetilde{\boldsymbol{U}}^{l,n,m} \to \boldsymbol{u}^m$ strongly in $\mathrm{L}^2(Q)^d$, as $l, n \to \infty$. By the boundedness in $\mathrm{L}^\infty(0, T; \mathrm{L}^2(\Omega)^d)$ in (4.59) and interpolation, this implies that

$$\widetilde{\boldsymbol{U}}^{l,n,m} \to \boldsymbol{u}^m \quad \text{strongly in } \mathrm{L}^p(0, T; \mathrm{L}^2(\Omega)^d), \quad \text{as } l, n \to \infty, \tag{4.73}$$

for any $p \in [1, \infty)$. Similarly, by (3.19) we have that

$$\begin{aligned} \left\|\overline{\boldsymbol{U}}^{l,n,m} - \widetilde{\boldsymbol{U}}^{l,n,m}\right\|^2_{\mathrm{L}^2(Q)} &\overset{(3.19)}{=} \sum_{i=1}^{l} \left\|(t_i - t)\partial_t \widetilde{\boldsymbol{U}}^{l,n,m}\right\|^2_{\mathrm{L}^2(Q^i_{i-1})} \\ &\leq \delta_l^2 \left\|\partial_t \widetilde{\boldsymbol{U}}^{l,n,m}\right\|^2_{\mathrm{L}^2(Q)} \overset{(4.63)}{\leq} c\delta_l \to 0, \quad \text{as } l \to \infty. \end{aligned} \tag{4.74}$$

Consequently, with (4.73) it follows that $\overline{\boldsymbol{U}}^{l,n,m} \to \boldsymbol{u}^m$ strongly in $\mathrm{L}^2(Q)^d$, as $l, n \to \infty$. In particular, $t \mapsto \|\overline{\boldsymbol{U}}^{l,n,m}(t, \cdot) - \boldsymbol{u}^m(t, \cdot)\|_{\mathrm{L}^2(\Omega)}$ converges to zero strongly in $\mathrm{L}^2(0, T)$, as $l, n \to \infty$. Thus, there exists a subsequence such that $t \mapsto \|\overline{\boldsymbol{U}}^{l,n,m}(t, \cdot) - \boldsymbol{u}^m(t, \cdot)\|_{\mathrm{L}^2(\Omega)}$ converges to zero a.e. in $(0, T)$, as $l, n \to \infty$, which implies (4.48). Analogously, (4.45) follows from the strong convergence of $\widetilde{\boldsymbol{U}}^{l,n,m}$ in (4.73).

The uniform bounds in $\mathrm{L}^\infty(0, T; \mathrm{L}^2(\Omega)^d)$ and $\mathrm{L}^\eta(Q)^d$, with $\eta = \max\left(2q', \frac{q(d+2)}{d}\right)$, by (4.57) and (4.58), and the strong convergence in $\mathrm{L}^2(Q)^d$, yield by interpolation, that

$$\overline{\boldsymbol{U}}^{l,n,m} \to \boldsymbol{u}^m \quad \text{strongly in } \mathrm{L}^p(0, T; \mathrm{L}^2(\Omega)^d) \cap \mathrm{L}^r(Q)^d, \quad \text{as } l, n \to \infty, \tag{4.75}$$

for any $p \in [1, \infty)$ and any $r \in [1, \eta)$. By the uniform bounds in (4.57), (4.58) and (4.59), and the Banach–Alaoglu theorem, up to subsequences, we have that

$$\widetilde{\boldsymbol{U}}^{l,n,m}, \overline{\boldsymbol{U}}^{l,n,m} \overset{*}{\rightharpoonup} \boldsymbol{u}^m \quad \text{weakly* in } \mathrm{L}^\infty(0, T; \mathrm{L}^2(\Omega)^d), \tag{4.76}$$

$$\overline{\boldsymbol{U}}^{l,n,m} \rightharpoonup \boldsymbol{u}^m \quad \text{weakly in } \mathrm{L}^q(0, T; \mathrm{W}^{1,q}_0(\Omega)^d) \cap \mathrm{L}^\eta(Q)^d, \tag{4.77}$$

as $l,n \to \infty$, and the identification of the limiting functions follows by the strong convergence in (4.75).

The argument that $\boldsymbol{u}^m$ is divergence-free follows as in Diening *et al.* (2013, p. 1001): let $h \in L^{q'}(\Omega)$ and note that by the Assumption 3.4 on the projector $\Pi^n_{\mathbb{Q}}$, we have that $\Pi^n_{\mathbb{Q}} h \to h$ in particular in $L^{q'}(\Omega)$, as $n \to \infty$, compare Remark 3.5 (ii). Also, let $\varphi \in C^\infty_0(0,T)$. By (4.77) we have that $\operatorname{div} \overline{\boldsymbol{U}}^{l,n,m} \rightharpoonup \operatorname{div} \boldsymbol{u}^m$ weakly in $L^q(Q)$, and hence

$$\left\langle \operatorname{div} \overline{\boldsymbol{U}}^{l,n,m}, \varphi \Pi^n_{\mathbb{Q}} h \right\rangle_Q \to \left\langle \operatorname{div} \boldsymbol{u}^m, \varphi h \right\rangle_Q, \quad \text{as } l,n \to \infty. \tag{4.78}$$

Since $\overline{\boldsymbol{U}}^{l,n,m} \in \mathbb{P}^l_0(0,T;\mathbb{V}^n_{\mathrm{div}})$, the left-hand side vanishes for all $l,n \in \mathbb{N}$, and hence we have $\langle \operatorname{div} \boldsymbol{u}^m, h\varphi\rangle_Q = 0$ for all $h \in L^{q'}(\Omega)$ and all $\varphi \in C^\infty(0,T)$, so by density $\boldsymbol{u}^m$ is (weakly) divergence-free.

By (4.57) with $(2q')'(2q'-1) = 2q'$ it follows that $\{|\overline{\boldsymbol{U}}^{l,n,m}|^{2q'-2}\,\overline{\boldsymbol{U}}^{l,n,m}\}_{l,n\in\mathbb{N}}$ is bounded in $L^{(2q')'}(Q)^d$ and thus, by the Banach–Alaoglu theorem there exists a subsequence and $\psi^m \in L^{(2q')'}(Q)^d$ such that

$$\left|\overline{\boldsymbol{U}}^{l,n,m}\right|^{2q'-2}\overline{\boldsymbol{U}}^{l,n,m} \rightharpoonup \psi^m \quad \text{weakly in } L^{(2q')'}(Q)^d, \quad \text{as } l,n \to \infty. \tag{4.79}$$

By the strong convergence in (4.75), there exists a subsequence, which converges a.e. in Q, and hence we can identify $\psi^m = |\boldsymbol{u}^m|^{2q'-2}\boldsymbol{u}^m$, which shows (4.50).

Because $\widetilde{\boldsymbol{U}}^{l,n,m}(0,\cdot) = P^n_{\mathrm{div}}\boldsymbol{u}_0$ by (4.36), with (3.10) it follows that

$$\widetilde{\boldsymbol{U}}^{l,n,m}(0,\cdot) = P^n_{\mathrm{div}}\boldsymbol{u}_0 \to \boldsymbol{u}_0 \quad \text{strongly in } L^2(\Omega)^d, \quad \text{as } n \to \infty, \tag{4.80}$$

so (4.45) is proven.

The uniform estimates in (4.55) and the Banach–Alaoglu theorem imply that there exist $\overline{\mathbf{S}}^m$, $\mathbf{S}^m \in L^{q'}(Q)^{d\times d}$ such that

$$\overline{\mathbf{S}}^{l,n,m} \rightharpoonup \overline{\mathbf{S}}^m \quad \text{weakly in } L^{q'}(Q)^{d\times d}, \tag{4.81}$$

$$\mathbf{S}^{l,n,m} \rightharpoonup \mathbf{S}^m \quad \text{weakly in } L^{q'}(Q)^{d\times d}, \tag{4.82}$$

as $l,n \to \infty$. It remains to show that $\overline{\mathbf{S}}^m = \mathbf{S}^m$; to this end, let $\mathbf{B} \in C^\infty_0(Q)^{d\times d}$ be arbitrary but fixed. On the one hand, the weak convergence in (4.81) shows that

$$\left\langle \overline{\mathbf{S}}^{l,n,m}, \mathbf{B} \right\rangle_Q \to \left\langle \overline{\mathbf{S}}^m, \mathbf{B} \right\rangle_Q, \quad \text{as } l,n \to \infty. \tag{4.83}$$

On the other hand, by the relation between $\overline{\mathbf{S}}^{l,n,m}$ and $\mathbf{S}^{l,n,m}$ according to (4.23), one can show that $\langle \overline{\mathbf{S}}^{l,n,m}, \mathbf{B}\rangle_Q = \langle \mathbf{S}^{l,n,m}, \overline{\mathbf{B}}\rangle_Q$. By (3.24) we have that $\overline{\mathbf{B}} \to \mathbf{B}$ strongly in $L^q(Q)^{d\times d}$, as $l \to \infty$, so that with the convergence in (4.82) it follows that

$$\left\langle \overline{\mathbf{S}}^{l,n,m}, \mathbf{B} \right\rangle_Q = \left\langle \mathbf{S}^{l,n,m}, \overline{\mathbf{B}} \right\rangle_Q \to \langle \mathbf{S}^m, \mathbf{B} \rangle_Q, \quad \text{as } l, n \to \infty. \tag{4.84}$$

By (4.83) and (4.84), the uniqueness of limits implies that $\overline{\mathbf{S}}^m = \mathbf{S}^m$ a.e. in Q. □

For $m \in \mathbb{N}$, $t \in (0, T)$, $\boldsymbol{u} \in \mathrm{L}^{2q'}(Q)^d$ and $\boldsymbol{v} \in \mathrm{X}(\Omega)$ let us introduce

$$\begin{aligned} \mathfrak{L}^m[\boldsymbol{u}; \boldsymbol{v}](t) := &- b(\boldsymbol{u}(t,\cdot), \boldsymbol{u}(t,\cdot), \boldsymbol{v}) - \langle \mathbf{S}^m(t,\cdot), \mathbf{D}\boldsymbol{v} \rangle_\Omega \\ &- \frac{1}{m} \left\langle |\boldsymbol{u}(t,\cdot)|^{2q'-2} \boldsymbol{u}(t,\cdot), \boldsymbol{v} \right\rangle_\Omega + \langle \boldsymbol{f}(t,\cdot), \boldsymbol{v} \rangle_\Omega, \end{aligned} \tag{4.85}$$

where $\mathbf{S}^m$ is given by Lemma 4.6 and $b(\cdot,\cdot,\cdot)$ is defined in (3.11).

LEMMA 4.7 (Identification of the PDE as $l, n \to \infty$). The functions $\boldsymbol{u}^m \in \mathrm{L}^\infty(0, T; \mathrm{L}^2_{\mathrm{div}}(\Omega)^d) \cap \mathrm{X}_{\mathrm{div}}(Q)$ given in Lemma 4.6 satisfy that $\partial_t \boldsymbol{u}^m \in \mathrm{L}^\tau(0, T; (\mathrm{X}_{\mathrm{div}}(\Omega))')$, with $\tau := \min(q', (2q')') > 1$, and $\mathrm{X}_{\mathrm{div}}(\Omega)$ defined in (2.1). (Up to a representative) we have that $\boldsymbol{u}^m \in \mathrm{C}_w([0, T], \mathrm{L}^2_{\mathrm{div}}(\Omega)^d)$ for all $m \in \mathbb{N}$. Furthermore, for each $m \in \mathbb{N}$ the functions $\boldsymbol{u}^m$ and $\mathbf{S}^m \in \mathrm{L}^{q'}(Q)^{d\times d}$ from Lemma 4.6 satisfy

$$\langle \partial_t \boldsymbol{u}^m(t,\cdot), \boldsymbol{w} \rangle_\Omega = \mathfrak{L}^m[\boldsymbol{u}^m; \boldsymbol{w}](t) \quad \text{for all } \boldsymbol{w} \in \mathrm{C}^\infty_{0,\mathrm{div}}(\Omega)^d \text{ for a.e. } t \in (0, T), \tag{4.86}$$

$$(\mathbf{D}\boldsymbol{u}^m(z), \mathbf{S}^m(z)) \in \mathcal{A}(z) \quad \text{for a.e. } z \in Q, \tag{4.87}$$

$$\operatorname{ess\,lim}_{t\to 0_+} \|\boldsymbol{u}^m(t,\cdot) - \boldsymbol{u}_0\|_{\mathrm{L}^2(\Omega)} = 0. \tag{4.88}$$

Proof. Let $m \in \mathbb{N}$ be arbitrary but fixed.

Step 1: *Identification of the limiting equation*. For $\lambda = (l, n, m) \in \mathbb{N}^3$ multiplying (4.35) by a function $\varphi \in \mathrm{C}^\infty_0((-T, T))$ and integrating over $(0, T)$ yields

$$\left\langle \partial_t \widetilde{\boldsymbol{U}}^\lambda, \boldsymbol{W}\varphi \right\rangle_Q \overset{(4.35)}{=} \left\langle \mathfrak{L}^\lambda[\overline{\boldsymbol{U}}^\lambda, \boldsymbol{W}], \varphi \right\rangle_{(0,T)} \quad \text{for any } \boldsymbol{W} \in \mathbb{V}^n_{\mathrm{div}}. \tag{4.89}$$

Then, by integration by parts and the fact that $\widetilde{\boldsymbol{U}}^\lambda \in \mathrm{C}([0, T]; \mathrm{L}^2(\Omega)^d)$, it follows that

$$-\left\langle \widetilde{\boldsymbol{U}}^\lambda, \boldsymbol{W}\partial_t\varphi \right\rangle_Q = \left\langle \widetilde{\boldsymbol{U}}^\lambda(0,\cdot), \varphi(0)\boldsymbol{W} \right\rangle_\Omega + \left\langle \mathfrak{L}^\lambda[\overline{\boldsymbol{U}}^\lambda, \boldsymbol{W}], \varphi \right\rangle_{(0,T)} \tag{4.90}$$

for all $\boldsymbol{W} \in \mathbb{V}^n_{\mathrm{div}}$ and all $\varphi \in \mathrm{C}^\infty_0((-T, T))$ and $\lambda = (l, n, m) \in \mathbb{N}^3$.

Now let $\boldsymbol{w} \in \mathrm{C}^\infty_{0,\mathrm{div}}(\Omega)^d$ and $\varphi \in \mathrm{C}^\infty_0((-T, T))$ be arbitrary. Recall that by Remark 3.5 for $\boldsymbol{w} \in \mathrm{C}^\infty_{0,\mathrm{div}}(\Omega)^d$, we have that

$$\mathbb{V}^n_{\mathrm{div}} \ni \Pi^n \boldsymbol{w} \to \boldsymbol{w} \quad \text{strongly in } \mathrm{W}^{1,s}_0(\Omega)^d, \quad \text{as } n \to \infty, \quad \text{for any } s \in [1, \infty). \tag{4.91}$$

In order to deduce the limiting equation for $\boldsymbol{u}^m$ we consider (4.90) term by term, as $l,n \to \infty$: let $s \in [1,\infty)$ be large enough that the embedding $\mathrm{W}^{1,s}(\Omega)^d \hookrightarrow \mathrm{L}^2(\Omega)^d$ is continuous. By the strong convergence of $\widetilde{\boldsymbol{U}}^{l,n,m}$ in $\mathrm{L}^p(0,T;\mathrm{L}^2(\Omega)^d)$, for $p \in [1,\infty)$ by (4.43), with (4.91) we obtain that

$$-\left\langle \widetilde{\boldsymbol{U}}^{l,n,m}, \Pi^n(\boldsymbol{w})\partial_t\varphi \right\rangle_Q \to -\langle \boldsymbol{u}^m, \boldsymbol{w}\partial_t\varphi\rangle_Q, \quad \text{as } l,n \to \infty. \tag{4.92}$$

Similarly, the strong convergence of $\widetilde{\boldsymbol{U}}^{l,n,m}(0,\cdot) \to \boldsymbol{u}_0$ in $\mathrm{L}^2(\Omega)^d$ in (4.45) yields that

$$\left\langle \widetilde{\boldsymbol{U}}^{l,n,m}(0,\cdot), \varphi(0)\Pi^n\boldsymbol{w} \right\rangle_\Omega \to \langle \boldsymbol{u}_0, \varphi(0)\boldsymbol{w}\rangle_\Omega, \quad \text{as } l,n \to \infty. \tag{4.93}$$

By the fact that $\overline{\boldsymbol{U}}^{l,n,m} \to \boldsymbol{u}^m$ strongly in $\mathrm{L}^r(Q)^d$ for all $r \in [1,\eta)$, as $l,n \to \infty$, by (4.47), it follows that $\overline{\boldsymbol{U}}^{l,n,m} \otimes \overline{\boldsymbol{U}}^{l,n,m} \to \boldsymbol{u}^m \otimes \boldsymbol{u}^m$ strongly in $\mathrm{L}^p(Q)^{d\times d}$ for all $p \in \left[1, \frac{\eta}{2}\right)$. Such a $p > 1$ exists, since $\eta = \max\left(2q', \frac{q(d+2)}{d}\right) > 2$. With (4.91) applied for $s = p' < \infty$, we obtain that $\varphi\nabla\Pi^n\boldsymbol{w} \to \varphi\nabla\boldsymbol{w}$ strongly in $\mathrm{L}^{p'}(Q)^{d\times d}$. Together these imply that

$$\left\langle \overline{\boldsymbol{U}}^{l,n,m} \otimes \overline{\boldsymbol{U}}^{l,n,m}, \varphi\nabla\Pi^n\boldsymbol{w} \right\rangle_Q \to \langle \boldsymbol{u}^m \otimes \boldsymbol{u}^m, \varphi\nabla\boldsymbol{w}\rangle_Q, \quad \text{as } l,n \to \infty. \tag{4.94}$$

For the modification of the convective term, note first that by (4.49) we have weak convergence of $\nabla\overline{\boldsymbol{U}}^{l,n,m} \rightharpoonup \nabla\boldsymbol{u}^m$ in $\mathrm{L}^q(Q)^{d\times d}$. By (4.47) we have in particular that $\overline{\boldsymbol{U}}^{l,n,m} \to \boldsymbol{u}^m$ strongly in $\mathrm{L}^{q'}(Q)^d$, as $l,n \to \infty$ since $q' < 2q' \le \eta$. For $s > d$, the embedding $\mathrm{W}^{1,s}(\Omega) \hookrightarrow \mathrm{L}^\infty(\Omega)$ is continuous, and hence we have $\varphi\Pi^n\boldsymbol{w} \to \varphi\boldsymbol{w}$ strongly in $\mathrm{L}^\infty(Q)^d$. Together, this yields that

$$\left\langle \overline{\boldsymbol{U}}^{l,n,m} \otimes \varphi\Pi^n\boldsymbol{w}, \nabla\overline{\boldsymbol{U}}^{l,n,m} \right\rangle_Q \to \langle \boldsymbol{u}^m \otimes \varphi\boldsymbol{w}, \nabla\boldsymbol{u}^m\rangle_Q, \quad \text{as } l,n \to \infty. \tag{4.95}$$

By (4.91) we have that $\varphi\mathbf{D}\Pi^n\boldsymbol{w} \to \varphi\mathbf{D}\boldsymbol{w}$ strongly in $\mathrm{L}^q(Q)^{d\times d}$ and by (4.51) that $\overline{\mathbf{S}}^{l,n,m} \rightharpoonup \mathbf{S}^m$ weakly in $\mathrm{L}^{q'}(Q)^{d\times d}$. Thus, it follows that

$$\left\langle \overline{\mathbf{S}}^{l,n,m}, \varphi\mathbf{D}\Pi^n\boldsymbol{w} \right\rangle_Q \to \langle \mathbf{S}^m, \varphi\mathbf{D}\boldsymbol{w}\rangle_Q, \quad \text{as } l,n \to \infty. \tag{4.96}$$

Since $|\overline{\boldsymbol{U}}^{l,n,m}|^{2q'-2}\overline{\boldsymbol{U}}^{l,n,m} \rightharpoonup |\boldsymbol{u}^m|^{2q'-2}\boldsymbol{u}^m$ weakly in $\mathrm{L}^{(2q')'}(Q)^d$ by (4.50) and $\varphi\Pi^n\boldsymbol{w} \to \varphi\boldsymbol{w}$, in particular in $\mathrm{L}^{2q'}(Q)^d$, we obtain

$$\frac{1}{m}\left\langle \left|\overline{\boldsymbol{U}}^{l,n,m}\right|^{2q'-2}\overline{\boldsymbol{U}}^{l,n,m}, \varphi\Pi^n\boldsymbol{w} \right\rangle_Q \to \frac{1}{m}\left\langle \left|\boldsymbol{u}^m\right|^{2q'-2}\boldsymbol{u}^m, \varphi\boldsymbol{w} \right\rangle_Q, \quad \text{as } l,n \to \infty. \tag{4.97}$$

Finally, with the strong convergence $\overline{\boldsymbol{f}} \to \boldsymbol{f}$ in $\mathrm{L}^{q'}(0,T;\mathrm{W}^{-1,q'}(\Omega)^d)$ by (4.4) and with (4.91), we have that

$$\left\langle \overline{\boldsymbol{f}}, \varphi\Pi^n\boldsymbol{w} \right\rangle_Q \to \langle \boldsymbol{f}, \varphi\boldsymbol{w}\rangle_Q, \quad \text{as } l,n \to \infty. \tag{4.98}$$

By the fact that $\boldsymbol{u}^m$ is divergence-free, it follows that $\widetilde{b}(\boldsymbol{u}^m, \boldsymbol{u}^m, \varphi\boldsymbol{w}) = b(\boldsymbol{u}^m, \boldsymbol{u}^m, \varphi\boldsymbol{w})$. So with $\mathfrak{L}^{l,n,m}$ and $\mathfrak{L}^m$ as defined in (4.34) and (4.85), respectively, the convergence results (4.94)–(4.98) yield that

$$\left\langle \mathfrak{L}^{l,n,m}[\overline{\boldsymbol{U}}^{l,n,m}, \Pi^n \boldsymbol{w}], \varphi \right\rangle_{(0,T)} \to \left\langle \mathfrak{L}^m[\boldsymbol{u}^m, \boldsymbol{w}], \varphi \right\rangle_{(0,T)}, \quad \text{as } l, n \to \infty. \tag{4.99}$$

Now, from (4.90), using (4.92), (4.93) and (4.99), as $l, n \to \infty$, we have that

$$-\left\langle \boldsymbol{u}^m, \boldsymbol{w}\partial_t \varphi \right\rangle_Q = \left\langle \boldsymbol{u}_0, \varphi(0)\boldsymbol{w} \right\rangle_\Omega + \left\langle \mathfrak{L}^m[\boldsymbol{u}^m, \boldsymbol{w}], \varphi \right\rangle_{(0,T)} \tag{4.100}$$

for all $\boldsymbol{w} \in \mathrm{C}^\infty_{0,\mathrm{div}}(\Omega)^d$ and all $\varphi \in \mathrm{C}^\infty_0((-T, T))$.

Step 2: *Bound on the time-derivative*. The distributional derivative of $\boldsymbol{u}^m$ satisfies, by definition and using (4.100), that

$$\left\langle \partial_t \boldsymbol{u}^m, \boldsymbol{w}\varphi \right\rangle_Q = -\left\langle \boldsymbol{u}^m, \boldsymbol{w}\partial_t \varphi \right\rangle_Q \overset{(4.100)}{=} \left\langle \mathfrak{L}^m[\boldsymbol{u}^m, \boldsymbol{w}], \varphi \right\rangle_{(0,T)} \tag{4.101}$$

for all $\boldsymbol{w} \in \mathrm{C}^\infty_{0,\mathrm{div}}(\Omega)^d$ and all $\varphi \in \mathrm{C}^\infty_0((0, T))$, since $\operatorname{supp}\varphi \subset (0, T)$. Using this equation we wish to show that $\partial_t \boldsymbol{u}^m \in \mathrm{L}^\tau(0, T; (\mathrm{X}_{\mathrm{div}}(\Omega))')$ (not uniformly in $m \in \mathbb{N}$), for $\tau := \min(q', (2q')') > 1$ and $\mathrm{X}_{\mathrm{div}}(\Omega)$ as in (2.1). For $\mathfrak{L}^m$ as defined in (4.85), using the fact that $\boldsymbol{u}^m \in \mathrm{L}^{2q'}(Q)^d$ and $\mathbf{S}^m \in \mathrm{L}^{q'}(Q)^{d\times d}$, similarly as in (4.64) we can estimate

$$\begin{aligned}
\left|\left\langle \mathfrak{L}^m[\boldsymbol{u}^m, \boldsymbol{w}], \varphi \right\rangle_{(0,T)}\right| &\le \left\|\boldsymbol{u}^m\right\|^2_{\mathrm{L}^{2q'}(Q)} \|\varphi\nabla \boldsymbol{w}\|_{\mathrm{L}^q(Q)} + \left\|\mathbf{S}^m\right\|_{\mathrm{L}^{q'}(Q)} \left\|\varphi \mathbf{D}\boldsymbol{w}\right\|_{\mathrm{L}^q(Q)} \\
&\quad + \frac{1}{m}\left\|\boldsymbol{u}^m\right\|^{2q'-1}_{\mathrm{L}^{2q'}(Q)} \|\varphi \boldsymbol{w}\|_{\mathrm{L}^{2q'}(Q)} + \|\boldsymbol{f}\|_{\mathrm{L}^{q'}(0,T;\mathrm{W}^{-1,q'}(\Omega))} \|\varphi \boldsymbol{w}\|_{\mathrm{L}^q(0,T;\mathrm{W}^{1,q}(\Omega))} \\
&\le c(m)\left(\|\varphi\|_{\mathrm{L}^q(0,T)} + \|\varphi\|_{\mathrm{L}^{2q'}(0,T)}\right)\left(\|\boldsymbol{w}\|_{\mathrm{W}^{1,q}(\Omega)} + \|\boldsymbol{w}\|_{\mathrm{L}^{2q'}(\Omega)}\right) \\
&\le c(m)\,\|\varphi\|_{\mathrm{L}^{\tau'}(0,T)} \|\boldsymbol{w}\|_{\mathrm{X}(\Omega)}
\end{aligned} \tag{4.102}$$

for all $\varphi \in \mathrm{C}^\infty_0((0, T))$ and all $\boldsymbol{w} \in \mathrm{C}^\infty_{0,\mathrm{div}}(\Omega)^d$, since $\tau' = \max(2q', q)$. By the density of the respective test function spaces, $\langle \mathfrak{L}^m[\boldsymbol{u}^m, \cdot], \cdot\rangle_{(0,T)}$ represents a bounded linear functional on $\mathrm{L}^{\tau'}(0, T; \mathrm{X}_{\mathrm{div}}(\Omega))$, and thus we have that $\partial_t \boldsymbol{u}^m \in \mathrm{L}^\tau(0, T; (\mathrm{X}_{\mathrm{div}}(\Omega))')$ by (4.101) and by reflexivity of the function space. Consequently, $\left\langle \partial_t \boldsymbol{u}^m, \boldsymbol{w}\right\rangle_\Omega$ is integrable for $\boldsymbol{w} \in \mathrm{C}^\infty_{0,\mathrm{div}}(\Omega)^d$, and thus we can rephrase (4.101) by the fundamental lemma of calculus of variations in the pointwise sense in time, so (4.86) is proved.

Step 3: *Identification of the initial condition*. Since the arguments are standard, let us only give an outline and refer to Tscherpel (2018) for additional details. Using the function spaces which $\boldsymbol{u}^m$ and $\partial_t \boldsymbol{u}^m$ are contained in, one can show that $\boldsymbol{u}^m \in \mathrm{C}_w([0, T]; \mathrm{L}^2_{\mathrm{div}}(\Omega)^d)$, see Temam (1984, Lem. 1.1, 1.4, Ch. III, § 1). Then, by use of the equation and integration by parts, one can identify $\boldsymbol{u}^m(0, \cdot) = \boldsymbol{u}_0 \in \mathrm{L}^2_{\mathrm{div}}(\Omega)^d$. Finally, by the strong convergence in (4.44) and (4.45), and also applying the energy inequality (4.42), the proof of (4.88) can be concluded.

Step 4: *Energy identity*. Let us recall that $\boldsymbol{u}^m \in \mathrm{X}_{\mathrm{div}}(Q) \hookrightarrow \mathrm{L}^{\min(q,2q')}(0,T;\mathrm{X}_{\mathrm{div}}(\Omega)^d)$ and that $\partial_t \boldsymbol{u}^m \in \mathrm{L}^{\tau}(0,T;(\mathrm{X}_{\mathrm{div}}(\Omega))')$, where $\tau = \min(q',(2q')')$, and equation (4.86) is satisfied. Because of the lack of integrability in time, an approximation argument by means of mollification in time can be applied to show the energy identity

$$\frac{1}{2}\left\|\boldsymbol{u}^m(s,\cdot)\right\|^2_{\mathrm{L}^2(\Omega)} + \left\langle \mathbf{S}^m, \mathbf{D}\boldsymbol{u}^m\right\rangle_{Q_s} + \frac{1}{m}\left\|\boldsymbol{u}^m\right\|^{2q'}_{\mathrm{L}^{2q'}(Q_s)} = \left\langle \boldsymbol{f}, \boldsymbol{u}^m\right\rangle_{Q_s} + \frac{1}{2}\left\|\boldsymbol{u}_0\right\|^2_{\mathrm{L}^2(\Omega)} \tag{4.103}$$

for a.e. $s \in (0,T)$. The proof follows by a standard procedure and we therefore omit the details, see, e.g., Lions (1969, Ch. 2.5). Let us note, however, that the identity can be obtained only for a.e. $s \in (0,T)$, since the limit can be taken only for Lebesgue points of the function $t \mapsto \|\boldsymbol{u}^m(t,\cdot)\|^2_{\mathrm{L}^2(\Omega)}$. Also, the attainment of the initial datum in the sense of (4.88) is used.

***Step 5**: Identification of the implicit relation*. Recall that we have by the assertion (4.37) that the inclusion $(\mathbf{D}\overline{\boldsymbol{U}}^{l,n,m}(z), \mathbf{S}^{l,n,m}(z)) \in \mathcal{A}(z)$ holds for a.e. $z \in Q$. Furthermore, by (4.49) we have that $\mathbf{D}\overline{\boldsymbol{U}}^{l,n,m} \rightharpoonup \mathbf{D}\boldsymbol{u}^m$ weakly in $\mathrm{L}^q(Q)^{d\times d}$ and by (4.52) that $\mathbf{S}^{l,n,m} \rightharpoonup \mathbf{S}^m$ weakly in $\mathrm{L}^{q'}(Q)^{d\times d}$, as $l, n \to \infty$. By Lemma 2.6 it suffices to show that

$$\limsup_{l,n\to\infty} \left\langle \mathbf{S}^{l,n,m}, \mathbf{D}\overline{\boldsymbol{U}}^{l,n,m}\right\rangle_{Q_s} \le \left\langle \mathbf{S}^m, \mathbf{D}\boldsymbol{u}^m\right\rangle_{Q_s}, \tag{4.104}$$

in order to obtain $(\mathbf{D}\boldsymbol{u}^m(z), \mathbf{S}^m(z)) \in \mathcal{A}(z)$ for a.e. $z \in Q_s$. Then we can exhaust Q by letting $s \to T$. We can only show (4.104) for a.e. $s \in (0,T)$ since the energy identity (4.103) is available only for a.e. $s \in (0,T)$, and some of the arguments used to show (4.104) are only available for a.e. $s \in (0,T)$.

Let us add and subtract the term $\langle \overline{\mathbf{S}}^{l,n,m}, \mathbf{D}\overline{\boldsymbol{U}}^{l,n,m}\rangle_{Q_s}$ to obtain

$$\left\langle \mathbf{S}^{l,n,m}, \mathbf{D}\overline{\boldsymbol{U}}^{l,n,m}\right\rangle_{Q_s} = \left\langle \overline{\mathbf{S}}^{l,n,m}, \mathbf{D}\overline{\boldsymbol{U}}^{l,n,m}\right\rangle_{Q_s} + \left\langle \mathbf{S}^{l,n,m} - \overline{\mathbf{S}}^{l,n,m}, \mathbf{D}\overline{\boldsymbol{U}}^{l,n,m}\right\rangle_{Q_s} =: \mathrm{I} + \mathrm{II}, \tag{4.105}$$

where the first term appears in the equation (4.35) for the approximate solutions and the second term has to be shown to vanish. The energy inequality (4.42) yields that

$$\begin{aligned}\mathrm{I} = \left\langle \overline{\mathbf{S}}^{l,n,m}, \mathbf{D}\overline{\boldsymbol{U}}^{l,n,m}\right\rangle_{Q_s} &\overset{(4.42)}{\le} -\frac{1}{2}\left\|\widetilde{\boldsymbol{U}}^{l,n,m}(s,\cdot)\right\|^2_{\mathrm{L}^2(\Omega)} + \frac{1}{2}\left\|\widetilde{\boldsymbol{U}}^{l,n,m}(0,\cdot)\right\|^2_{\mathrm{L}^2(\Omega)} \\ &\quad + \left\langle \overline{\boldsymbol{f}}, \overline{\boldsymbol{U}}^{l,n,m}\right\rangle_{Q_s} - \frac{1}{m}\left\|\overline{\boldsymbol{U}}^{l,n,m}\right\|^{2q'}_{\mathrm{L}^{2q'}(Q_s)}.\end{aligned} \tag{4.106}$$

For the second term in (4.105), for $l \in \mathbb{N}$ let $j \in \{1,\ldots,l\}$ be such that $s \in (t_{j-1}, t_j]$, i.e., j depends on s and on l. By the relation (4.23) we have that

$$\left\langle \mathbf{S}^{l,n,m} - \overline{\mathbf{S}}^{l,n,m}, \mathbf{D}\overline{\boldsymbol{U}}^{l,n,m}\right\rangle_{Q^i_{i-1}} = \left\langle \int_{t_{i-1}}^{t_i} \mathbf{S}^{l,n,m}(t,\cdot)\,\mathrm{d}t - \delta_l \overline{\mathbf{S}}^{l,n,m}, \mathbf{D}\overline{\boldsymbol{U}}^{l,n,m}\right\rangle_{\Omega} = 0 \tag{4.107}$$

for any $i \in \{1, \dots, l\}$. So for II we obtain that

$$\begin{aligned} \mathrm{II} &= \left\langle \mathbf{S}^{l,n,m} - \overline{\mathbf{S}}^{l,n,m}, \mathbf{D}\overline{\boldsymbol{U}}^{l,n,m} \right\rangle_{Q_{t_j}} - \left\langle \mathbf{S}^{l,n,m} - \overline{\mathbf{S}}^{l,n,m}, \mathbf{D}\overline{\boldsymbol{U}}^{l,n,m} \right\rangle_{Q_s^{t_j}} \\ &\overset{(4.107)}{=} 0 - \left\langle \mathbf{S}^{l,n,m}, \mathbf{D}\overline{\boldsymbol{U}}^{l,n,m} \right\rangle_{Q_s^{t_j}} + \left\langle \overline{\mathbf{S}}^{l,n,m}, \mathbf{D}\overline{\boldsymbol{U}}^{l,n,m} \right\rangle_{Q_s^{t_j}} \leq \left\langle \overline{\mathbf{S}}^{l,n,m}, \mathbf{D}\overline{\boldsymbol{U}}^{l,n,m} \right\rangle_{Q_s^{t_j}}, \end{aligned} \tag{4.108}$$

where the inequality follows since $\mathcal{A}(\cdot)$ is monotone a.e. in Q, $(\mathbf{0}, \mathbf{0}) \in \mathcal{A}(\cdot)$ a.e. in Q, and by the fact that $(\mathbf{D}\overline{\boldsymbol{U}}^{l,n,m}, \mathbf{S}^{l,n,m}) \in \mathcal{A}(\cdot)$ a.e. in Q by (4.37). For the remaining term we use again (4.42) on (s, t_j), noting that the term involving $\frac{1}{m}$ is nonnegative, which yields

$$\left\langle \overline{\mathbf{S}}^{l,n,m}, \mathbf{D}\overline{\boldsymbol{U}}^{l,n,m} \right\rangle_{Q_s^{t_j}} \leq \left\langle \overline{\boldsymbol{f}}, \overline{\boldsymbol{U}}^{l,n,m} \right\rangle_{Q_s^{t_j}} - \frac{1}{2} \left\| \widetilde{\boldsymbol{U}}^{l,n,m}(t_j, \cdot) \right\|^2_{\mathrm{L}^2(\Omega)} + \frac{1}{2} \left\| \widetilde{\boldsymbol{U}}^{l,n,m}(s, \cdot) \right\|^2_{\mathrm{L}^2(\Omega)}. \tag{4.109}$$

By the duality of norms, the estimate (4.57) and by (4.3), we obtain

$$\begin{aligned} \left\langle \overline{\boldsymbol{f}}, \overline{\boldsymbol{U}}^{l,n,m} \right\rangle_{Q_s^{t_j}} &\leq \left\| \overline{\boldsymbol{f}} \right\|_{\mathrm{L}^{q'}(s,t_j;\mathrm{W}^{-1,q'}(\Omega))} \left\| \overline{\boldsymbol{U}}^{l,n,m} \right\|_{\mathrm{L}^q(0,T;\mathrm{W}^{1,q}(\Omega))} \\ &\leq c(m) \, \|\boldsymbol{f}\|_{\mathrm{L}^{q'}(t_{j-1},t_j;\mathrm{W}^{-1,q'}(\Omega))} \leq c(m) \, \|\boldsymbol{f}\|_{\mathrm{L}^{q'}(s-\delta_l,s+\delta_l;\mathrm{W}^{-1,q'}(\Omega))}. \end{aligned} \tag{4.110}$$

Furthermore, we have $\widetilde{\boldsymbol{U}}^{l,n,m}(t_j, \cdot) = \overline{\boldsymbol{U}}^{l,n,m}(t_j, \cdot) = \overline{\boldsymbol{U}}^{l,n,m}(s, \cdot)$, since $s \in (t_{j-1}, t_j]$, and hence

$$\begin{aligned} \mathrm{II} &\overset{(4.108),(4.109)}{\leq} \left\langle \overline{\boldsymbol{f}}, \overline{\boldsymbol{U}}^{l,n,m} \right\rangle_{Q_s^{t_j}} - \frac{1}{2} \left\| \widetilde{\boldsymbol{U}}^{l,n,m}(t_j, \cdot) \right\|^2_{\mathrm{L}^2(\Omega)} + \frac{1}{2} \left\| \widetilde{\boldsymbol{U}}^{l,n,m}(s, \cdot) \right\|^2_{\mathrm{L}^2(\Omega)} \\ &\overset{(4.110)}{\leq} c(m) \, \|\boldsymbol{f}\|_{\mathrm{L}^{q'}(s-\delta_l,s+\delta_l;\mathrm{W}^{-1,q'}(\Omega))} - \frac{1}{2} \left\| \overline{\boldsymbol{U}}^{l,n,m}(s, \cdot) \right\|^2_{\mathrm{L}^2(\Omega)} + \frac{1}{2} \left\| \widetilde{\boldsymbol{U}}^{l,n,m}(s, \cdot) \right\|^2_{\mathrm{L}^2(\Omega)}. \end{aligned} \tag{4.111}$$

Now applying $\limsup_{l,n\to\infty}$ to (I+II) with (4.106) and (4.111), noting that the term involving $\widetilde{\boldsymbol{U}}^{l,n,m}(s, \cdot)$ drops out, we obtain

$$\begin{aligned} \limsup_{l,n\to\infty} (\mathrm{I} + \mathrm{II}) &\leq -\frac{1}{2} \lim_{l,n\to\infty} \left\| \overline{\boldsymbol{U}}^{l,n,m}(s, \cdot) \right\|^2_{\mathrm{L}^2(\Omega)} + \frac{1}{2} \lim_{l,n\to\infty} \left\| \widetilde{\boldsymbol{U}}^{l,n,m}(0, \cdot) \right\|^2_{\mathrm{L}^2(\Omega)} \\ &\quad - \frac{1}{m} \liminf_{l,n\to\infty} \left\| \overline{\boldsymbol{U}}^{l,n,m} \right\|^{2q'}_{\mathrm{L}^{2q'}(Q_s)} + \lim_{l,n\to\infty} \left\langle \overline{\boldsymbol{f}}, \overline{\boldsymbol{U}}^{l,n,m} \right\rangle_{Q_s} \\ &\quad + c(m) \lim_{l\to\infty} \|\boldsymbol{f}\|_{\mathrm{L}^{q'}(s-\delta_l,s+\delta_l;\mathrm{W}^{-1,q'}(\Omega))} \\ &\leq -\frac{1}{2} \|\boldsymbol{u}^m(s, \cdot)\|^2_{\mathrm{L}^2(\Omega)} + \frac{1}{2} \|\boldsymbol{u}_0\|^2_{\mathrm{L}^2(\Omega)} - \frac{1}{m} \|\boldsymbol{u}^m\|^{2q'}_{\mathrm{L}^{2q'}(Q_s)} + \langle \boldsymbol{f}, \boldsymbol{u}^m \rangle_{Q_s}, \end{aligned} \tag{4.112}$$

where the last inequality is based on the following arguments. By (4.48) we have the convergence $\overline{\boldsymbol{U}}^{l,n,m}(s, \cdot) \to \boldsymbol{u}^m(s, \cdot)$ strongly in $\mathrm{L}^2(\Omega)^d$, as $l, n \to \infty$, for a.e. $s \in (0, T)$. The second term converges to $\frac{1}{2}\|\boldsymbol{u}_0\|^2_{\mathrm{L}^2(\Omega)}$, since by (4.45) we have that $\widetilde{\boldsymbol{U}}^{l,n,m}(0, \cdot) \to \boldsymbol{u}_0$ strongly in $\mathrm{L}^2(\Omega)^d$. For the third term we use weak lower semicontinuity with respect to the weak convergence in $\mathrm{L}^{2q'}(Q_s)^d$ and (4.49). For the fourth term we have convergence, since $\overline{\boldsymbol{U}}^{l,n,m} \rightharpoonup \boldsymbol{u}^m$ weakly in $\mathrm{L}^q(0,T;\mathrm{W}^{1,q}(\Omega)^d)$ by (4.49) and $\overline{\boldsymbol{f}} \to \boldsymbol{f}$ strongly in $\mathrm{L}^{q'}(0,T;\mathrm{W}^{-1,q'}(\Omega)^d)$ by (4.4), as $l,n \to \infty$. The last term vanishes by the absolute continuity of the integral, as $l \to \infty$. Finally, returning to (4.105), applying $\limsup_{l,n\to\infty}$ and the energy identity (4.103) for a.e. $s \in (0,T)$, yields

$$\begin{aligned}
\limsup_{l,n\to\infty}\left\langle \mathbf{S}^{l,n,m}, \mathbf{D}\overline{\boldsymbol{U}}^{l,n,m}\right\rangle_{Q_s} &\overset{(4.105)}{\leq} \limsup_{l,n\to\infty}\,(\mathrm{I}+\mathrm{II})\\
&\overset{(4.112)}{\leq} -\frac{1}{2}\left\|\boldsymbol{u}^m(s,\cdot)\right\|^2_{\mathrm{L}^2(\Omega)} + \frac{1}{2}\left\|\boldsymbol{u}_0\right\|^2_{\mathrm{L}^2(\Omega)} - \frac{1}{m}\left\|\boldsymbol{u}^m\right\|^{2q'}_{\mathrm{L}^{2q'}(Q_s)} + \left\langle \boldsymbol{f}, \boldsymbol{u}^m\right\rangle_{Q_s}\\
&\overset{(4.103)}{=} \left\langle \mathbf{S}^m, \mathbf{D}\boldsymbol{u}^m\right\rangle_{Q_s}
\end{aligned} \tag{4.113}$$

for a.e. $s \in (0,T)$. This proves the claim in (4.104) and completes the proof. □

Limit $m \to \infty$

In this step we lose the admissibility of the solution as a test function, and we have to use Lipschitz truncation to identify the implicit relation. The availability of the solenoidal Lipschitz truncation allows to simplify the arguments in Bulíček *et al.* (2012), since no pressure has to be reconstructed.

For $q \in \left(\frac{2d}{d+2}, \infty\right)$ let us denote

$$\hat{q} := \max\left(\left(\frac{q(d+2)}{2d}\right)', q\right) = \max\left(\frac{q(d+2)}{q(d+2)-2d}, q\right) < \infty, \quad \nu := \max\left(\hat{q}, 2q'\right) < \infty, \tag{4.114}$$

and note that $\hat{q} = q$, if $q \geq \frac{3d+2}{d+2}$.

LEMMA 4.8 (Convergence $m \to \infty$). For $m \in \mathbb{N}$ let $\boldsymbol{u}^m \in \mathrm{L}^\infty(0,T;\mathrm{L}^2_{\mathrm{div}}(\Omega)^d) \cap \mathrm{X}_{\mathrm{div}}(Q)$ be such that $\partial_t \boldsymbol{u}^m \in \mathrm{L}^\tau(0,T;(\mathrm{X}_{\mathrm{div}}(\Omega))')$ and let $\mathbf{S}^m \in \mathrm{L}^{q'}(Q)^{d\times d}$ be a solution to (4.86)–(4.88). Further, let $\hat{q}, \nu \in (1,\infty)$ be defined as in (4.114). Then, there exists a constant $c > 0$ such that, for all $m \in \mathbb{N}$, we have that

$$\begin{aligned}
&\left\|\boldsymbol{u}^m\right\|^2_{\mathrm{L}^\infty(0,T;\mathrm{L}^2(\Omega))} + \left\|\boldsymbol{u}^m\right\|^q_{\mathrm{L}^q(0,T;\mathrm{W}^{1,q}(\Omega))} + \left\|\mathbf{S}^m\right\|^{q'}_{\mathrm{L}^{q'}(Q)}\\
&\qquad + \frac{1}{m}\left\|\boldsymbol{u}^m\right\|^{2q'}_{\mathrm{L}^{2q'}(Q)} + \left\|\boldsymbol{u}^m\right\|_{\mathrm{L}^{\frac{q(d+2)}{d}}(Q)} \leq c.
\end{aligned} \tag{4.115}$$

Furthermore, there exists a function $\boldsymbol{u} \in \mathrm{L}^\infty(0,T;\mathrm{L}^2_{\mathrm{div}}(\Omega)^d) \cap \mathrm{L}^q(0,T;\mathrm{W}^{1,q}_{0,\mathrm{div}}(\Omega)^d)$ such that we have that $\partial_t \boldsymbol{u} \in \mathrm{L}^{\nu'}(0,T;(\mathrm{W}^{1,\hat{q}}_{0,\mathrm{div}}(\Omega)^d)')$, an $\mathbf{S} \in \mathrm{L}^{q'}(Q)^{d\times d}$ and subsequences such that, as $m \to \infty$,

$$\boldsymbol{u}^m \to \boldsymbol{u} \qquad \text{strongly in } \mathrm{L}^q(0,T;\mathrm{L}^2_{\mathrm{div}}(\Omega)^d) \cap \mathrm{L}^r(Q)^d,\ \forall r \in \left[1, \tfrac{q(d+2)}{d}\right), \tag{4.116}$$

$$\boldsymbol{u}^m(s,\cdot) \to \boldsymbol{u}(s,\cdot) \qquad \text{strongly in } \mathrm{L}^2_{\mathrm{div}}(\Omega)^d \text{ for a.e. } s \in (0,T), \tag{4.117}$$

$$\boldsymbol{u}^m \rightharpoonup \boldsymbol{u} \qquad \text{weakly in } \mathrm{L}^q(0,T;\mathrm{W}^{1,q}_{0,\mathrm{div}}(\Omega)^d) \cap \mathrm{L}^{\frac{q(d+2)}{d}}(Q)^d, \tag{4.118}$$

$$\boldsymbol{u}^m \overset{*}{\rightharpoonup} \boldsymbol{u} \qquad \text{weakly* in } \mathrm{L}^\infty(0,T;\mathrm{L}^2_{\mathrm{div}}(\Omega)^d), \tag{4.119}$$

$$\partial_t \boldsymbol{u}^m \rightharpoonup \partial_t \boldsymbol{u} \qquad \text{weakly in } \mathrm{L}^{\nu'}(0,T;(\mathrm{W}^{1,\hat{q}}_{0,\mathrm{div}}(\Omega)^d)'), \tag{4.120}$$

$$\mathbf{S}^m \rightharpoonup \mathbf{S} \qquad \text{weakly in } \mathrm{L}^{q'}(Q)^{d\times d}, \tag{4.121}$$

$$\frac{1}{m}\left|\boldsymbol{u}^m\right|^{2q'-2}\boldsymbol{u}^m \to \mathbf{0} \qquad \text{strongly in } \mathrm{L}^{(2q')'}(Q)^d. \tag{4.122}$$

Proof.

Step 1: *Estimates*. Recall that the bounds on $\{\overline{\boldsymbol{U}}^{l,n,m}\}_{l,n,m}$ in $\mathrm{L}^\infty(0,T;\mathrm{L}^2(\Omega)^d) \cap \mathrm{L}^q(0,T;\mathrm{W}^{1,q}_0(\Omega)^d)$, on $\{m^{-\frac{1}{2q'}}\overline{\boldsymbol{U}}^{l,n,m}$ in $\mathrm{L}^{2q'}(Q)^d$ by (4.57) and on $\{\mathbf{S}^{l,n,m}\}_{l,n,m}$ in $\mathrm{L}^{q'}(Q)^{d\times d}$ by (4.55) are uniform in $m \in \mathbb{N}$. Hence, by the weak* and weak convergence in (4.46), (4.49) and (4.52), and the weak(*) lower semicontinuity of the norm, the estimate (4.115) follows.

In order to derive a uniform bound on the time derivative, let us estimate $\mathfrak{L}^m[\boldsymbol{u}^m;\boldsymbol{v}]$. Since no uniform bounds on $\|\boldsymbol{u}^m\|_{\mathrm{L}^{2q'}(Q)}$ are available at this point, we use the bound (3.13) on the convective term, with $\hat{q}$ as defined in (4.114), to deduce that

$$\left|\langle \boldsymbol{u}(t,\cdot)\otimes\boldsymbol{u}(t,\cdot),\nabla\boldsymbol{v}\rangle_\Omega\right| \le c\left\|\boldsymbol{u}(t,\cdot)\right\|^2_{\mathrm{L}^{\frac{q(d+2)}{d}}(\Omega)}\|\nabla\boldsymbol{v}\|_{\mathrm{L}^{\hat{q}}(\Omega)},$$

which holds, since $q \ge \frac{2d}{d+2}$. Note that the embedding $\mathrm{W}^{1,\hat{q}}(\Omega) \hookrightarrow \mathrm{W}^{1,q}(\Omega) \cap \mathrm{L}^{2q'}(\Omega)$ is continuous for $\hat{q}$ as in (4.114). Also, we have for ν as defined in (4.114) that the embedding $\mathrm{L}^\nu(\Omega) \hookrightarrow \mathrm{L}^{\hat{q}}(\Omega) \cap \mathrm{L}^q(\Omega) \cap \mathrm{L}^{2q'}(\Omega)$ is continuous. With this, similarly as in (4.102) applying the uniform estimates in (4.115), one has that

$$\begin{aligned}
\left|\left\langle \mathfrak{L}^m[\boldsymbol{u}^m,\boldsymbol{w}],\varphi\right\rangle_{(0,T)}\right| &\le \left\|\boldsymbol{u}^m\right\|^2_{\mathrm{L}^{\frac{q(d+2)}{d}}(Q)}\|\varphi\nabla\boldsymbol{w}\|_{\mathrm{L}^{\hat{q}}(Q)} + \left\|\mathbf{S}^m\right\|_{\mathrm{L}^{q'}(Q)}\left\|\varphi\mathbf{D}\boldsymbol{w}\right\|_{\mathrm{L}^q(Q)} \\
&\quad + \frac{1}{m}\left\|\boldsymbol{u}^m\right\|^{2q'-1}_{\mathrm{L}^{2q'}(Q)}\|\varphi\boldsymbol{w}\|_{\mathrm{L}^{2q'}(Q)} + \|\boldsymbol{f}\|_{\mathrm{L}^{q'}(0,T;\mathrm{W}^{-1,q'}(\Omega))}\|\varphi\boldsymbol{w}\|_{\mathrm{L}^q(0,T;\mathrm{W}^{1,q}(\Omega))} \\
&\le c\,\|\varphi\|_{\mathrm{L}^\nu(0,T)}\|\boldsymbol{w}\|_{\mathrm{W}^{1,\hat{q}}(\Omega)}
\end{aligned} \tag{4.123}$$

for all $\varphi \in \mathrm{C}^\infty_0((0,T))$, all $\boldsymbol{w} \in \mathrm{C}^\infty_{0,\mathrm{div}}(\Omega)^d$ and all $m \in \mathbb{N}$. With (4.86) and using the fact that $\nu < \infty$, and hence the space $\mathrm{L}^\nu(0,T;\mathrm{W}^{1,\hat{q}}_{0,\mathrm{div}}(\Omega)^d)$ is reflexive, this shows that $\{\partial_t\boldsymbol{u}^m\}_{m\in\mathbb{N}}$ is bounded in $\mathrm{L}^{\nu'}(0,T;(\mathrm{W}^{1,\hat{q}}_{0,\mathrm{div}}(\Omega)^d)')$.

Step 2: *Convergence as $m \to \infty$*. Since $q > \frac{2d}{d+2}$, the embedding $\mathrm{W}^{1,q}_{0,\mathrm{div}}(\Omega)^d \hookrightarrow\hookrightarrow \mathrm{L}^2_{\mathrm{div}}(\Omega)^d$ is compact. Because $\hat{q} \ge q > \frac{2d}{d+2}$, the embedding $\mathrm{W}^{1,\hat{q}}_{0,\mathrm{div}}(\Omega) \hookrightarrow \mathrm{L}^2_{\mathrm{div}}(\Omega)^d$ is in particular continuous and dense, which implies that $(\mathrm{L}^2_{\mathrm{div}}(\Omega)^d) \hookrightarrow (\mathrm{W}^{1,\hat{q}}_{0,\mathrm{div}}(\Omega))'$. Combined with the embedding

$\mathrm{L}^2_{\mathrm{div}}(\Omega)^d \hookrightarrow (\mathrm{L}^2_{\mathrm{div}}(\Omega)^d)'$, this yields that the embedding $\mathrm{L}^2_{\mathrm{div}}(\Omega)^d \hookrightarrow (\mathrm{W}^{1,\hat{q}}_{0,\mathrm{div}}(\Omega))'$ is continuous. Hence, the Aubin–Lions compactness lemma implies that the embedding

$$\left\{\boldsymbol{v} \in \mathrm{L}^q(0,T;\mathrm{W}^{1,q}_{0,\mathrm{div}}(\Omega)^d) : \ \partial_t \boldsymbol{v} \in \mathrm{L}^{\nu'}(0,T;(\mathrm{W}^{1,\hat{q}}_{0,\mathrm{div}}(\Omega))')\right\} \hookrightarrow\hookrightarrow \mathrm{L}^q\left(0,T,\mathrm{L}^2_{\mathrm{div}}(\Omega)^d\right)$$

is compact, see, for example, Roubíček (2013, Lem. 7.7). The fact that by (4.115) the sequence $\{\boldsymbol{u}^m\}_{m\in\mathbb{N}}$ is bounded in $\mathrm{L}^q(0,T;\mathrm{W}^{1,q}_{0,\mathrm{div}}(\Omega)^d)$ and that $\{\partial_t\boldsymbol{u}^m\}_{m\in\mathbb{N}}$ is bounded in $\mathrm{L}^{\nu'}(0,T;(\mathrm{W}^{1,\hat{q}}_{0,\mathrm{div}}(\Omega)^d)')$, then ensures the existence of a subsequence such that

$$\boldsymbol{u}^m \to \boldsymbol{u} \quad \text{strongly in } \mathrm{L}^q(0,T;\mathrm{L}^2_{\mathrm{div}}(\Omega)^d), \quad \text{as } m\to\infty. \tag{4.124}$$

By the estimates in (4.115), the uniform bound on $\{\partial_t\boldsymbol{u}^m\}_{m\in\mathbb{N}}$ in $\mathrm{L}^{\nu'}(0,T;(\mathrm{W}^{1,\hat{q}}_{0,\mathrm{div}}(\Omega)^d)')$ and the Banach–Alaoglu theorem, there exists a subsequence such that (4.118)–(4.121) holds, where the limits can be identified with the help of (4.124).

The strong convergence in $\mathrm{L}^r(Q)^d$ for all $r \in \left[1,\frac{q(d+2)}{d}\right)$ asserted in (4.116) follows from the strong convergence in $\mathrm{L}^1(Q)^d$ by (4.124), and the boundedness in $\mathrm{L}^{\frac{q(d+2)}{d}}(Q)^d$ by (4.115) by means of interpolation. The convergence (4.117) is deduced analogously to the proof of (4.44) by the arguments following (4.74). By the estimate (4.115) we have with $(2q')' = \frac{2q'}{2q'-1} > 1$ that

$$\left\|\frac{1}{m}\left|\boldsymbol{u}^m\right|^{2q'-2}\boldsymbol{u}^m\right\|^{(2q')'}_{\mathrm{L}^{(2q')'}(Q)} \le m^{-(2q')'}\int_Q \left|\boldsymbol{u}^m\right|^{2q'}\mathrm{d}z = m^{1-(2q')'}\frac{1}{m}\left\|\boldsymbol{u}^m\right\|^{2q'}_{\mathrm{L}^{2q'}(Q)} \overset{(4.115)}{\le} cm^{1-(2q')'} \to 0,$$

as $m\to\infty$, so (4.122) follows. □

For $t\in(0,T)$, $\boldsymbol{u}\in\mathrm{L}^{\frac{q(d+2)}{d}}(Q)^d$ and $\boldsymbol{v}\in\mathrm{W}^{1,\hat{q}}_0(\Omega)^d$ with $\hat{q}$ defined in (4.114), let us introduce

$$\mathfrak{L}[\boldsymbol{u};\boldsymbol{v}](t) := -\,b(\boldsymbol{u}(t,\cdot),\boldsymbol{u}(t,\cdot),\boldsymbol{v}) - \left\langle \mathbf{S}(t,\cdot),\mathbf{D}\boldsymbol{v}\right\rangle_\Omega + \langle \boldsymbol{f}(t,\cdot),\boldsymbol{v}\rangle_\Omega, \tag{4.125}$$

where $\mathbf{S}\in\mathrm{L}^{q'}(Q)^{d\times d}$ is the limiting function introduced in Lemma 4.8.

LEMMA 4.9 (Identification of the PDE as $m\to\infty$). The function $\boldsymbol{u}\in\mathrm{L}^\infty(0,T;\mathrm{L}^2_{\mathrm{div}}(\Omega)^d)\cap\mathrm{L}^q(0,T;\mathrm{W}^{1,q}_{0,\mathrm{div}}(\Omega)^d)$ from Lemma 4.8 satisfies that $\partial_t\boldsymbol{u}\in\mathrm{L}^{\hat{q}'}(0,T;(\mathrm{W}^{1,\hat{q}}_{0,\mathrm{div}}(\Omega)^d)')$, with $\hat{q}$ defined in (4.114). (Up to a representative) we have that $\boldsymbol{u}\in C_w([0,T],\mathrm{L}^2_{\mathrm{div}}(\Omega)^d)$. Furthermore, the functions $\boldsymbol{u}$ and $\mathbf{S}\in\mathrm{L}^{q'}(Q)^{d\times d}$ from Lemma 4.8 satisfy

$$\left\langle\partial_t\boldsymbol{u}(t,\cdot),\boldsymbol{w}\right\rangle_\Omega = \mathfrak{L}[\boldsymbol{u};\boldsymbol{w}](t) \qquad \text{for all } \boldsymbol{w}\in C^\infty_{0,\mathrm{div}}(\Omega)^d, \text{for a.e. } t\in(0,T), \tag{4.126}$$

$$(\mathbf{D}\boldsymbol{u}(z),\mathbf{S}(z))\in\mathcal{A}(z) \qquad \text{for a.e. } z\in Q, \tag{4.127}$$

$$\mathit{ess\,lim}_{t\to 0_+}\left\|\boldsymbol{u}(t,\cdot)-\boldsymbol{u}_0\right\|_{\mathrm{L}^2(\Omega)} = 0, \tag{4.128}$$

i.e., $(\boldsymbol{u},\mathbf{S})$ is in particular a weak solution according to Definition 2.1.

Proof.

Step 1: *Identification of the limiting equation.* Let $\boldsymbol{w} \in \mathrm{C}^{\infty}_{0,\mathrm{div}}(\Omega)^d$, let $\varphi \in \mathrm{C}^{\infty}_0((0,T))$ and let us consider each of the terms in (4.86) and (4.126). By the weak convergence in (4.120) and (4.121), we have that

$$\langle \partial_t \boldsymbol{u}^m, \varphi \boldsymbol{w} \rangle_Q \to \langle \partial_t \boldsymbol{u}, \varphi \boldsymbol{w} \rangle_Q, \tag{4.129}$$

$$\langle \mathbf{S}^m, \varphi \mathbf{D}\boldsymbol{w} \rangle_Q \to \langle \mathbf{S}, \varphi \mathbf{D}\boldsymbol{w} \rangle_Q, \tag{4.130}$$

as $m \to \infty$. Since by (4.116) we have that $\boldsymbol{u}^m \to \boldsymbol{u}$ in $\mathrm{L}^r(Q)^{d\times d}$ for all $r \in \left[1, \frac{q(d+2)}{d}\right)$, it follows that $\boldsymbol{u}^m \otimes \boldsymbol{u}^m \to \boldsymbol{u} \otimes \boldsymbol{u}$ in $\mathrm{L}^r(Q)^{d\times d}$ for all $r \in \left[1, \frac{q(d+2)}{2d}\right)$. Since $q > \frac{2d}{d+2}$, this set is nonempty and the convergence holds in particular in $\mathrm{L}^1(Q)^{d\times d}$; hence, we have that

$$\langle \boldsymbol{u}^m \otimes \boldsymbol{u}^m, \varphi \nabla \boldsymbol{w} \rangle_Q \to \langle \boldsymbol{u} \otimes \boldsymbol{u}, \varphi \nabla \boldsymbol{w} \rangle_Q, \quad \text{as } m \to \infty. \tag{4.131}$$

Taking the results in (4.129)–(4.131) and (4.122) shows that (4.86) implies (4.126).

Step 2: *Identification of the initial condition.* With similar arguments as in the proof of Lemma 4.7, Step 3, it follows that $\boldsymbol{u} \in \mathrm{C}_w([0,T]; \mathrm{L}^2_{\mathrm{div}}(\Omega)^d)$, that $\boldsymbol{u}_0 = \boldsymbol{u}(0,\cdot) \in \mathrm{L}^2_{\mathrm{div}}(\Omega)^d$ and that the initial datum is attained in the sense of (4.128).

Step 3: *Higher integrability of the time derivative.* As in Step 2 in the proof of Lemma 4.6, we can improve the integrability of $\partial_t \boldsymbol{u}$ using the fact that (4.126) is satisfied. This yields that we have that $\partial_t \boldsymbol{u} \in \mathrm{L}^{\hat{q}'}(0,T; (\mathrm{W}^{1,\hat{q}}_{0,\mathrm{div}}(\Omega)^d)')$, for $\hat{q}$ as defined in (4.114).

Step 4: *Identification of the implicit relation (cf. Bulíček et al., (2012) and Breit et al., (2013, Sec. 3)).* Recall that $\mathbf{D}\boldsymbol{u}^m \rightharpoonup \mathbf{D}\boldsymbol{u}$ weakly in $\mathrm{L}^q(Q)^{d\times d}$ by (4.118), that $\mathbf{S}^m \rightharpoonup \mathbf{S}$ weakly in $\mathrm{L}^{q'}(Q)^{d\times d}$ by (4.121) and that we have that $(\mathbf{D}\boldsymbol{u}^m(z), \mathbf{S}^m(z)) \in \mathcal{A}(z)$ for a.e. $z \in Q$ by (4.87). Hence, by Lemma 2.6, it suffices to show that

$$\limsup_{m\to\infty} \langle \mathbf{S}^m, \mathbf{D}\boldsymbol{u}^m \rangle_{\widetilde{Q}} \le \langle \mathbf{S}, \mathbf{D}\boldsymbol{u} \rangle_{\widetilde{Q}}, \tag{4.132}$$

for a set $\widetilde{Q} \subset Q$, to identify the implicit relation $(\mathbf{D}\boldsymbol{u}, \mathbf{S}) \in \mathcal{A}(\cdot)$ a.e. on $\widetilde{Q}$.

Since there is no energy identity available for $\boldsymbol{u}$, in order to identify the implicit relation, one has to truncate the elements of the approximating sequence of velocity fields suitably so as to be able to use them as test functions. In contrast with Bulíček *et al.* (2012), we will not use a parabolic Lipschitz truncation after locally reconstructing the approximations to the pressure, but work with the solenoidal Lipschitz truncation introduced subsequently in Breit *et al.* (2013) and stated in Lemma 2.7, as the argument is then more direct.

We wish to truncate $\boldsymbol{v}^m := \boldsymbol{u}^m - \boldsymbol{u}$, which satisfies, for all $\boldsymbol{\xi} \in \mathrm{C}^{\infty}_{0,\mathrm{div}}(Q)^d$, that

$$\langle \partial_t \boldsymbol{v}^m, \boldsymbol{\xi} \rangle_Q = \langle \boldsymbol{u}^m \otimes \boldsymbol{u}^m - \boldsymbol{u} \otimes \boldsymbol{u}, \nabla \boldsymbol{\xi} \rangle_Q - \langle \mathbf{S}^m - \mathbf{S}, \mathbf{D}\boldsymbol{\xi} \rangle_Q - \frac{1}{m} \left\langle |\boldsymbol{u}^m|^{2q'-2} \boldsymbol{u}^m, \boldsymbol{\xi} \right\rangle_Q, \tag{4.133}$$

by (4.86) and (4.126) and by the density of $\mathrm{C}_0^\infty(0,T) \times \mathrm{C}_{0,\mathrm{div}}^\infty(\Omega)^d$ in $\mathrm{C}_{0,\mathrm{div}}^\infty(Q)^d$. Due to the (lower order) regularizing term we aim to apply Corollary 2.8 instead of Lemma 2.7 with $p = q \in (1,\infty)$ and σ such that

$$1 < \sigma < \min\left(2, q, q', \frac{q(d+2)}{2d}, (2q')'\right) = \min\left(q', \frac{q(d+2)}{2d}, (2q')'\right). \tag{4.134}$$

Such a σ exists, since we have by assumption that $q > \frac{2d}{d+2}$. Let $Q_0 = I_0 \times B_0 \subset\subset Q$ be a parabolic cylinder. First note that $\boldsymbol{u}$ and $\boldsymbol{u}^m$ are (weakly) divergence-free, and so is $\boldsymbol{v}^m$, and $\boldsymbol{v}^m \rightharpoonup \boldsymbol{0}$ weakly in $\mathrm{L}^q(I_0;\mathrm{W}^{1,q}(B_0)^d)$, as $m \to \infty$ by (4.118). Since $\boldsymbol{u}^m \to \boldsymbol{u}$ strongly in $\mathrm{L}^p(Q)$ for $p \in \left[1, \frac{q(d+2)}{d}\right)$ by (4.116) and $\sigma < \frac{q(d+2)}{d}$, we have that $\boldsymbol{v}^m \to \boldsymbol{0}$ strongly in $\mathrm{L}^\sigma(Q_0)^d$, as $m \to \infty$. Furthermore, since $\{\boldsymbol{u}^m\}_{m\in\mathbb{N}}$ is bounded in $\mathrm{L}^\infty(0,T;\mathrm{L}^2(\Omega)^d)$ by (4.115), we have with $\sigma < 2$ that $\{\boldsymbol{v}^m\}_{m\in\mathbb{N}}$ is bounded in $\mathrm{L}^\infty(0,T;\mathrm{L}^\sigma(\Omega)^d)$. Now we set

$$\mathbf{G}_1^m := \mathbf{S} - \mathbf{S}^m, \quad \widetilde{\mathbf{G}}_2^m := \boldsymbol{u}^m \otimes \boldsymbol{u}^m - \boldsymbol{u} \otimes \boldsymbol{u} \quad \text{and} \quad \boldsymbol{f}^m := -\frac{1}{m}\left|\boldsymbol{u}^m\right|^{2q'-2}\boldsymbol{u}^m.$$

Note that $\mathbf{G}_1^m \rightharpoonup \mathbf{0}$ weakly in $\mathrm{L}^{q'}(Q_0)^{d\times d}$ by (4.121). By (4.116) we have that $\boldsymbol{u}^m \to \boldsymbol{u}$ in $\mathrm{L}^r(Q)^d$ for all $r \in \left[1, \frac{q(d+2)}{d}\right)$, and thus, $\boldsymbol{u}^m \otimes \boldsymbol{u}^m \to \boldsymbol{u} \otimes \boldsymbol{u}$ in $\mathrm{L}^r(Q)^{d\times d}$ for all $r \in \left[1, \frac{q(d+2)}{2d}\right)$. This holds in particular for $r = \sigma < \frac{q(d+2)}{2d}$. Furthermore, by (4.122) we have that $\boldsymbol{f}^m \to \boldsymbol{0}$ strongly in $\mathrm{L}^{(2q')'}(Q)^d$, as $m \to \infty$, and hence also strongly in $\mathrm{L}^\sigma(Q_0)^d$. This means that all the assumptions of Corollary 2.8 are satisfied, and hence the statement of Lemma 2.7 is available with $\mathbf{G}_2^m := \widetilde{\mathbf{G}}_2^m - \nabla\Delta^{-1}\boldsymbol{f}^m$.

With the aid of the parabolic solenoidal Lipschitz truncation, we show that

$$\lim_{m\to\infty} \int_{\frac{1}{8}Q_0} \left[\left(\mathbf{S}^m - \mathbf{S}^\star(\cdot,\mathbf{D}\boldsymbol{u})\right) : \left(\mathbf{D}\boldsymbol{u}^m - \mathbf{D}\boldsymbol{u}\right)\right]^{\frac{1}{2}} \,\mathrm{d}z = 0, \tag{4.135}$$

where the exponent $\frac{1}{2}$ is used to control the size of the set, where $\boldsymbol{v}^m = \boldsymbol{u}^m - \boldsymbol{u}$ and its truncation do not coincide. By the monotonicity of $\mathcal{A}$ and the fact that $(\mathbf{D}\boldsymbol{u}, \mathbf{S}^\star(\cdot,\mathbf{D}\boldsymbol{u})) \in \mathcal{A}(\cdot)$ and $(\mathbf{D}\boldsymbol{u}^m, \mathbf{S}^m) \in \mathcal{A}(\cdot)$ a.e. in Q by (4.87), it follows that the $\liminf_{m\to\infty}$ of the above is nonnegative. To show the other direction, denote $H^m := \left(\mathbf{S}^m - \mathbf{S}^\star(\cdot,\mathbf{D}\boldsymbol{u})\right) : \left(\mathbf{D}\boldsymbol{u}^m - \mathbf{D}\boldsymbol{u}\right) \geq 0$, and let $j \geq j_0$, $\mathscr{B}_{m,j} \subset Q_0$ and $\boldsymbol{v}^{m,j}$ be given by Lemma 2.7 applied on Q_0, and by (ii) we have that $\boldsymbol{v}^m = \boldsymbol{v}^{m,j}$ on $\frac{1}{8}Q_0 \backslash \mathscr{B}_{m,j}$. Dividing the domain into $\frac{1}{8}Q_0 \cap \mathscr{B}_{m,j}$ and $\frac{1}{8}Q_0 \backslash \mathscr{B}_{m,j}$, by Hölder's inequality, we obtain

$$\begin{aligned}
\int_{\frac{1}{8}Q_0} (H^m)^{\frac{1}{2}} \,\mathrm{d}z &= \int_{\frac{1}{8}Q_0\cap\mathscr{B}_{m,j}} (H^m)^{\frac{1}{2}} \,\mathrm{d}z + \int_{\frac{1}{8}Q_0\backslash\mathscr{B}_{m,j}} (H^m)^{\frac{1}{2}} \,\mathrm{d}z \\
&\leq \left|\tfrac{1}{8}Q_0 \cap \mathscr{B}_{m,j}\right|^{\frac{1}{2}} \left(\int_{\frac{1}{8}Q_0\cap\mathscr{B}_{m,j}} H^m \,\mathrm{d}z\right)^{\frac{1}{2}} + \left|\tfrac{1}{8}Q_0 \backslash \mathscr{B}_{m,j}\right|^{\frac{1}{2}} \left(\int_{\frac{1}{8}Q_0\backslash\mathscr{B}_{m,j}} H^m \,\mathrm{d}z\right)^{\frac{1}{2}} \\
&\leq \left|\mathscr{B}_{m,j}\right|^{\frac{1}{2}} \left(\int_Q H^m \,\mathrm{d}z\right)^{\frac{1}{2}} + |Q|^{\frac{1}{2}} \left(\int_{\frac{1}{8}Q_0\backslash\mathscr{B}_{m,j}} H^m \,\mathrm{d}z\right)^{\frac{1}{2}},
\end{aligned} \tag{4.136}$$

where we have used the nonnegativity of H^m in the first term. Since H^m is bounded in $\mathrm{L}^1(Q)$ by the *a priori* estimate in (4.115), one has that

$$\int_{\frac{1}{8}Q_0} (H^m)^{\frac{1}{2}}\,\mathrm{d}z \le c\left|\mathscr{B}_{m,j}\right|^{\frac{1}{2}} + c\left(\int_{\frac{1}{8}Q_0\setminus\mathscr{B}_{m,j}} H^m\,\mathrm{d}z\right)^{\frac{1}{2}}. \tag{4.137}$$

By Lemma 2.7 (iii) we have that

$$\limsup_{m\to\infty}\left|\mathscr{B}_{m,j}\right|^{\frac{1}{2}} \le \limsup_{m\to\infty}\left(\lambda_{m,j}^q\left|\mathscr{B}_{m,j}\right|\right)^{\frac{1}{2}} \le c2^{-\frac{j}{2}}. \tag{4.138}$$

Let $\zeta \in \mathrm{C}_0^\infty(\frac{1}{6}B_0)$ be the nonnegative function given by Lemma 2.7 such that $\zeta|_{\frac{1}{8}B_0} \equiv 1$. In the second term in (4.137), we can use the nonnegativity of H^m, the definition of H^m and $\boldsymbol{v}^m$ and, finally, the definition of $\mathbf{G}_1^m$ in order to find that $\mathbf{S}^m = \mathbf{S} - \mathbf{G}_1^m$, and we obtain

$$\begin{aligned}\int_{\frac{1}{8}Q_0\setminus\mathscr{B}_{m,j}} H^m\,\mathrm{d}z &= \int_{\frac{1}{8}Q_0\setminus\mathscr{B}_{m,j}} H^m\,\zeta\,\mathrm{d}z = \int_{\frac{1}{8}Q_0} H^m\,\zeta\,\mathbb{1}_{\mathscr{B}_{m,j}^c}\,\mathrm{d}z\\ &\le \int H^m\,\zeta\,\mathbb{1}_{\mathscr{B}_{m,j}^c}\,\mathrm{d}z = \int \left(\mathbf{S}^m - \mathbf{S}^\star(\cdot,\mathbf{D}\boldsymbol{u})\right) : \mathbf{D}\boldsymbol{v}^m\,\zeta\,\mathbb{1}_{\mathscr{B}_{m,j}^c}\,\mathrm{d}z\\ &= -\int \left(\mathbf{G}_1^m - \mathbf{S} + \mathbf{S}^\star(\cdot,\mathbf{D}\boldsymbol{u})\right) : \nabla\boldsymbol{v}^m\,\zeta\,\mathbb{1}_{\mathscr{B}_{m,j}^c}\,\mathrm{d}z.\end{aligned} \tag{4.139}$$

Since $\mathbf{S} - \mathbf{S}^\star(\cdot,\mathbf{D}\boldsymbol{u}) \in \mathrm{L}^{q'}(Q)^{d\times d}$, we can use Lemma 2.7 (vii). Applying $\limsup_{m\to\infty}$ we find that

$$\limsup_{m\to\infty}\int_{\frac{1}{8}Q_0\setminus\mathscr{B}_{m,j}} H^m\,\mathrm{d}z \overset{(4.139)}{\le} \limsup_{m\to\infty}\left|\int \left(\mathbf{G}_1^m - \mathbf{S} + \mathbf{S}^\star(\cdot,\mathbf{D}\boldsymbol{u})\right) : \nabla\boldsymbol{v}^m\,\zeta\,\mathbb{1}_{\mathscr{B}_{m,j}^c}\,\mathrm{d}z\right| \le c2^{-\frac{j}{q}}. \tag{4.140}$$

Using (4.138) and (4.140) in (4.137) yields

$$\begin{aligned}\limsup_{m\to\infty}\int_{\frac{1}{8}Q_0}\left[\left(\mathbf{S}^m - \mathbf{S}^\star(\cdot,\mathbf{D}\boldsymbol{u})\right) : \left(\mathbf{D}\boldsymbol{u}^m - \mathbf{D}\boldsymbol{u}\right)\right]^{\frac{1}{2}}\,\mathrm{d}z &= \limsup_{m\to\infty}\int_{\frac{1}{8}Q_0}(H^m)^{\frac{1}{2}}\,\mathrm{d}z\\ &\overset{(4.137)}{\le} c\limsup_{m\to\infty}\left|\mathscr{B}_{m,j}\right|^{\frac{1}{2}} + c\limsup_{m\to\infty}\left(\int_{\frac{1}{8}Q_0\setminus\mathscr{B}_{m,j}} H^m\,\mathrm{d}z\right)^{\frac{1}{2}}\\ &\overset{(4.138),(4.140)}{\le} c(2^{-\frac{j}{2}} + 2^{-\frac{j}{2q}}).\end{aligned} \tag{4.141}$$

Then taking $j \to \infty$ gives the claim and (4.135) is proved.

This means that $(H^m)^{\frac{1}{2}} \to 0$ strongly in $\mathrm{L}^1(\frac{1}{8}Q_0)$, as $m \to \infty$. However, to show (4.132) we need L^1-convergence of H^m at least on suitable subdomains. The L^1-convergence implies that $(H^m)^{\frac{1}{2}} \to 0$ a.e. in $\frac{1}{8}Q_0$, and hence $H^m \to 0$ a.e. in $\frac{1}{8}Q_0$. Egorov's theorem implies that there exists a nonincreasing sequence of measurable subsets $E_i \subset \frac{1}{8}Q_0$, $i \in \mathbb{N}$, with $|E_i| \to 0$ as $i \to \infty$, such that $H^m \to 0$ uniformly on $\frac{1}{8}Q_0 \backslash E_i$, as $m \to \infty$, for any fixed $i \in \mathbb{N}$. In particular, we have that $H^m \to 0$ in $\mathrm{L}^1(\frac{1}{8}Q_0 \backslash E_i)$, as $m \to \infty$, for any fixed $i \in \mathbb{N}$, i.e.,

$$\langle \mathbf{S}^m - \mathbf{S}^\star(\cdot, \mathbf{D}\boldsymbol{u}), \mathbf{D}\boldsymbol{u}^m - \mathbf{D}\boldsymbol{u} \rangle_{\frac{1}{8}Q_0 \backslash E_i} \to 0, \quad \text{as } m \to \infty, \tag{4.142}$$

for any fixed $i \in \mathbb{N}$. With the weak convergence of $\mathbf{S}^m \rightharpoonup \mathbf{S}$ in $\mathrm{L}^{q'}(Q)^{d\times d}$ by (4.121) and the weak convergence of $\mathbf{D}\boldsymbol{u}^m \rightharpoonup \mathbf{D}\boldsymbol{u}$ in $\mathrm{L}^q(Q)^{d\times d}$ following from (4.119), we thus deduce that

$$\lim_{m\to\infty} \langle \mathbf{S}^m, \mathbf{D}\boldsymbol{u}^m \rangle_{\frac{1}{8}Q_0 \backslash E_i} = \langle \mathbf{S}, \mathbf{D}\boldsymbol{u} \rangle_{\frac{1}{8}Q_0 \backslash E_i} \quad \text{for all } i \in \mathbb{N}.$$

This shows (4.132) for $\widetilde{Q} = \frac{1}{8}Q_0 \backslash E_i$, and thus we find that $(\mathbf{D}\boldsymbol{u}(z), \mathbf{S}(z)) \in \mathcal{A}(z)$ for a.e. $z \in \frac{1}{8}Q_0 \backslash E_i$. Since $|E_i| \to 0$, as $i \to \infty$, we have that $(\mathbf{D}\boldsymbol{u}(z), \mathbf{S}(z)) \in \mathcal{A}(z)$ for a.e. $z \in \frac{1}{8}Q_0$.

Finally, let us consider a cover of Q consisting of (open) parabolic cylinders $Q^j = I^j \times B^j$, $j \in J$, for an index set J such that $Q = \bigcup_{j\in J} \frac{1}{8}Q^j$. This can be, for example, chosen as a Whitney-type cover, cf. Diening *et al.* (2010). Then we can identify the implicit relation a.e. on $\frac{1}{8}Q^j$ for all $j \in J$ by the above, and thus have that $(\mathbf{D}\boldsymbol{u}(z), \mathbf{S}(z)) \in \mathcal{A}(z)$ for a.e. $z \in Q$, which proves (4.127). □

REMARK 4.10 As an alternative to the fully implicit approximate problem in (4.6), (4.7) with (4.5), one can consider the semi-implicit scheme by replacing $\widetilde{b}(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{v})$ by $\widetilde{b}(\boldsymbol{U}^\kappa_{i-1}, \boldsymbol{u}, \boldsymbol{v})$, for $\kappa = (k, l, n, m) \in \mathbb{N}^4$. Since this represents a linearization of the problem, the approximate solutions exist and are unique. To show uniform estimates for the shifted interpolant, one has to estimate $\boldsymbol{U}^n_0$, which is the value taken on $[0, \delta_l]$. This can be done by assuming one of the following:

(1) Assume that $\{\mathcal{T}_n\}_{n\in\mathbb{N}}$ is quasiuniform and, if $q < 2$ and $\mathbb{V}^n_{\mathrm{div}}$, consists of discretely divergence-free finite element functions, assume additionally that there exist constants $c > 0$ and $\varepsilon > 0$ such that

$$\delta_l \leq c h_n^{\frac{d}{2}(\frac{2-q}{q-1})+\varepsilon} \quad \text{for all } l, n \in \mathbb{N}.$$

(2) Or assume that $\boldsymbol{u}_0 \in \mathrm{W}^{1,q}_{0,\mathrm{div}}(\Omega)^d$ and replace (4.6) by $\boldsymbol{U}^n_0 = \Pi^n \boldsymbol{u}_0$.

Funding

This work was supported by the UK Engineering and Physical Sciences Research Council [EP/L015811/1].

REFERENCES

ACERBI, E. and FUSCO, E. N. (1988) An approximation lemma for $W^{1,p}$ functions. *Material Instabilities in Continuum Mechanics (Edinburgh, 1985–1986). Oxford Science Publications.* New York: Oxford University Press, pp. 1–5.

ALBERTI, G. and AMBROSIO, L. (1999) A geometrical approach to monotone functions in $\mathbb{R}^n$. *Math. Z.*, **230**, 259–316.

Belenki, L., Berselli, L. C., Diening, L. & Růžička, M. (2012) On the finite element approximation of p-Stokes systems. *SIAM J. Numer. Anal.*, **50**, 373–397.

Bercovier, M. & Engelman, M. (1980) A finite-element method for incompressible non-Newtonian flows. *J. Comput. Phys.*, **36**, 313–326.

Bergh, J. & Löfström, J. (1976) *Interpolation Spaces: An Introduction*. *Grundlehren der Mathematischen Wissenschaften, vol. 223.* Berlin–New York: Springer.

Bernardi, C. & Raugel, G. (1985) Analysis of some finite elements for the Stokes problem. *Math. Comp.*, **44**, 71–79.

Berselli, L. C., Diening, L. & Růžička, M. (2010) Existence of strong solutions for incompressible fluids with shear dependent viscosities. *J. Math. Fluid Mech.*, **12**, 101–132.

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

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

Bogovskiĭ, M. E. (1979) Solution of the first boundary value problem for an equation of continuity of an incompressible medium. *Dokl. Akad. Nauk SSSR*, **248**, 1037–1040.

Breit, D., Diening, L. & Fuchs, M. (2012) Solenoidal Lipschitz truncation and applications in fluid mechanics. *J. Differential Equations*, **253**, 1910–1942.

Breit, D., Diening, L. & Schwarzacher, S. (2013) Solenoidal Lipschitz truncation for parabolic PDE's. *Math. Models Methods Appl. Sci.*, **23**, 2671–2700.

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

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

Bulíček, M. & Málek, J. (2016) On unsteady internal flows of Bingham fluids subject to threshold slip on the impermeable boundary. *Recent Developments of Mathematical Fluid Mechanics. Advances in Mathematical Fluid Mechanics.* Basel: Birkhäuser/Springer, pp. 135–156.

Carelli, E., Haehnle, J. & Prohl, A. (2010) Convergence analysis for incompressible generalized Newtonian fluid flows with nonstandard anisotropic growth conditions. *SIAM J. Numer. Anal.*, **48**, 164–190.

Ciarlet, P. (2002) *The Finite Element Method for Elliptic Problems*. *Society for Industrial and Applied Mathematics*. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM).

Crouzeix, M. & Raviart, P.-A. (1973) Conforming and nonconforming finite element methods for solving the stationary Stokes equations I. *Rev. Française Automat. Informat. Recherche Opérationnelle Sér. Rouge*, **7**, 33–75.

Dean, E. J., Glowinski, R. & Guidoboni, G. (2007) On the numerical simulation of Bingham visco-plastic flow: Old and new results. *J. Non-Newton. Fluid Mech.*, **142**, 36–62.

DiBenedetto, E. (1993) *Degenerate Parabolic Equations*. New York: Springer.

Diening, L., Kreuzer, C. & Süli, E. (2013) Finite element approximation of steady flows of incompressible fluids with implicit power-law-like rheology. *SIAM J. Numer. Anal.*, **51**, 984–1015.

Diening, L., Málek, J. & Steinhauer, M. (2008) On Lipschitz truncations of Sobolev functions (with variable exponent) and their selected applications. *ESAIM Control Optim. Calc. Var.*, **14**, 211–232.

Diening, L., Růžička, M. & Wolf, J. (2010) Existence of weak solutions for unsteady motions of generalized Newtonian fluids. *Ann. Scuola Norm. Sup. Pisa Cl. Sci.*, **9**, 1–46.

Duvaut, G. & Lions, J.-L. (1976) *Inequalities in Mechanics and Physics*. Translated from the French by C. W. John. *Grundlehren der Mathematischen Wissenschaften*, vol. 219. Berlin–New York: Springer.

Eberlein, H. & Růžička, M. (2012) Existence of weak solutions for unsteady motions of Herschel–Bulkley fluids. *J. Math. Fluid Mech.*, **14**, 485–500.

Ern, A. & Guermond, J.-L. (2004) *Theory and Practice of Finite Elements*, 1st edn. Applied Mathematical Sciences, vol. 159. New York: Springer.

FARIA, C. O. & KARAM-FILHO, J. (2013) A regularized-stabilized mixed finite element formulation for viscoplasticity of Bingham type. *Comput. Math. Appl.*, **66**, 975–995.

FREHSE, J., MÁLEK, J. & STEINHAUER, M. (1997) An existence result for fluids with shear dependent viscosity—steady flows. *Proceedings of the Second World Congress of Nonlinear Analysts, Part 5 (Athens, 1996)*, vol. 30, pp. 3041–3049. Oxford: Elsevier Ltd.

FREHSE, J., MÁLEK, J. & STEINHAUER, M. (2003) On analysis of steady flows of fluids with shear-dependent viscosity based on the Lipschitz truncation method. *SIAM J. Math. Anal.*, **34**, 1064–1083.

FUCHS, M. & SEREGIN, G. (2000) *Variational Methods for Problems From Plasticity Theory and for Generalized Newtonian Fluids*. Lecture Notes in Mathematics, vol. 1749. Berlin: Springer.

GILBARG, D. & TRUDINGER, N. S. (2001). *Elliptic Partial Differential Equations of Second Order*. Classics in Mathematics. Springer, 2001. Reprint, originally in Grundlehren der mathematischen Wissenschaften, vol. 224.

GIRAULT, V. & LIONS, J.-L. (2001) Two-grid finite-element schemes for the steady Navier–Stokes problem in polyhedra. *Port. Math. (N.S.)*, **58**, 25–57.

GIRAULT, V. & SCOTT, L. R. (2003) A quasi-local interpolation operator preserving the discrete divergence. *Calcolo*, **40**, 1–19.

GRISVARD, P. (2011) *Elliptic Problems in Nonsmooth Domains*. Classics in Applied Mathematics, vol. 69. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM).

GUZMÁN, J. & NEILAN, M. (2014a) Conforming and divergence-free Stokes elements in three dimensions. *IMA J. Numer. Anal.*, **34**, 1489–1508.

GUZMÁN, J. & NEILAN, M. (2014b) Conforming and divergence-free Stokes elements on general triangular meshes. *Math. Comp.*, **83**, 15–36.

HRON, J., MÁLEK, J., STEBEL, J. & TOUŠKA, K. (2017) A novel view on computations of steady flows of Bingham fluids using implicit constitutive relations. Project MORE Preprint, MORE/2017/08.

KINNUNEN, J. & LEWIS, J. L. (2002) Very weak solutions of parabolic systems of p-Laplacian type. *Ark. Mat.*, **40**, 105–132.

KREUZER, C. & SÜLI, E. (2016) Adaptive finite element approximation of steady flows of incompressible fluids with implicit power-law-like rheology. *ESAIM Math. Model. Numer. Anal.*, **50**, 1333–1369.

LADYŽENSKAJA, O. A. (1969) *The Mathematical Theory of Viscous Incompressible Flow*, 2nd edn. (English), revised and enlarged. Translated from the Russian by Richard A. Silverman and John Chu. *Mathematics and its Applications*, vol. 2. New York–London–Paris: Gordon and Breach, Science Publishers.

LIONS, J.-L. (1969) *Quelques Méthodes De Résolution Des Problèmes Aux Limites Non Linéaires*. Paris: Dunod.

MAHMOOD, R., KOUSAR, N., YAQUB, M. & JABEEN, K. (2017) Numerical simulations of the square lid driven cavity flow of Bingham fluids using nonconforming finite elements coupled with a direct solver. *Adv. Math. Phys.*, **2017**, 10 pp., Article ID 5210708.

MÁLEK, J., RŮŽIČKA, M. & SHELUKHIN, V. V. (2005) Herschel–Bulkley fluids: existence and regularity of steady flows. *Math. Models Methods Appl. Sci.*, **15**, 1845–1861.

MORENO, E., LARESE, A. & CERVERA, M. (2016) Modelling of Bingham and Herschel–Bulkley flows with mixed P1/P1 finite elements stabilized with orthogonal subgrid scale. *J. Non-Newton. Fluid Mech.*, **228**, 1–16.

RAJAGOPAL, K. R. (2003) On implicit constitutive theories. *Appl. Math.*, **48**, 279–319.

RAJAGOPAL, K. R. (2008) On implicit constitutive theories for fluids. *J. Fluid Mech., 550RS.*, **3**, 243–249.

ROUBÍČEK, T. (2013) *Nonlinear Partial Differential Equations with Applications*, 2nd ed. International Series of Numerical Mathematics, vol. 153. Basel: Birkhäuser/Springer Basel AG.

RŮŽIČKA, M. (1997) A note on steady flow of fluids with shear dependent viscosity. *Proceedings of the Second World Congress of Nonlinear Analysts, Part 5 (Athens, 1996). Nonlinear Anal.*, vol. 30, pp. 3029–3039.

SERËGIN, G. A. (1994) On a dynamical system generated by two-dimensional equations of motion of Bingham fluid. *J. Math. Sci.*, **70**, 1806–1816.

SIMON, J. (1987) Compact sets in the space $L^p\,(0,T;B)$. *Ann. Mat. Pura Appl. (4)*, **146**, 65–96.

TEMAM, R. (1984) *Navier–Stokes Equations: Theory & Numerical Analysis*. *Studies in Mathematics and Its Applications*. Amsterdam: North-Holland Publishing Co.

TSCHERPEL, T. (2018) Finite element approximation for the unsteady flow of implicitly constituted incompressible fluids. *D. Phil. Thesis*, University of Oxford. https://ora.ox.ac.uk/objects/uuid:01b4901c-9705-4087-80c1-4d656d160aed.

WOLF, J. (2007) Existence of weak solutions to the equations of non-stationary motion of non-Newtonian fluids with shear rate dependent viscosity. *J. Math. Fluid Mech.*, **9**, 104–138.

ZHANG, J. (2010) An augmented Lagrangian approach to Bingham fluid flows in a lid-driven square cavity with piecewise linear equal-order finite elements. *Comput. Methods Appl. Mech. Engrg.*, **199**, 3051–3057.