Finite element approximation for the unsteady flow of implicitly constituted incompressible fluids

Tabea Tscherpel

The Queen's College
University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy in Mathematics

2018

Declaration of Authorship

This thesis is entirely my own work except where otherwise indicated. I have clearly signalled the presence of quoted or paraphrased material and referenced all sources. It has not been submitted, either wholly or substantially, for another Honour School or degree of this University, or for a degree at any other institution.

Abstract

The purpose of this work is to provide qualitative convergence results of a sequence of numerical approximate solutions for the flow of incompressible fluids subject to an implicit constitutive relation. The constitutive law between shear stress tensor and shear rate tensor is encoded by a $(t, \boldsymbol{x})$-dependent maximal monotone graph with q-growth, $q \in (1, \infty)$. The resulting framework includes both generalised Newtonian fluids and discontinuous constitutive relations. Existence of weak solutions was established in [BGMS09, BGMS12] for steady and unsteady flows, respectively, for $q > \frac{2d}{d+2}$.

Due to the generality of the framework, higher regularity results and error estimates are out of reach. Thus we aim for convergence of a sequence of approximate solutions to a weak solution based on compactness results. For the finite element approximation, the assumptions result in a pair of (conforming) inf-sup stable finite element spaces for the velocity and the pressure.

We formulate the hypotheses on the graph so that the weak formulation is consistent with the thermodynamic framework. A number of graph approximations, all of which satisfy sufficient conditions to take the graph approximation limit before and separately from the discretisation limit are investigated. This includes the generalised Yosida graph approximation, which satisfies a Minty type convergence lemma, allowing us to take the graph approximation limit and the discretisation limit simultaneously.

Thanks to this, in the steady case we can simplify and generalise the proof in [DKS13a]: by means of a discrete Lipschitz approximation convergence is established for $q > \frac{2d}{d+1}$ when using discretely divergence-free finite element functions, and for $q > \frac{2d}{d+2}$ when using exactly divergence-free ones. In addition, we take a regularisation approach, allowing us to show convergence for the whole range $q > \frac{2d}{d+2}$ by separating the discretisation and the regularisation limit and using a non-discrete Lipschitz truncation.

For the unsteady problem no results in the context of implicit relations are available and the ones focused on explicit relations assume a restricted range of q, see [Car07, CHP10]. We set up a fully discrete and fully implicit approximate scheme based on finite element approximation in space and an implicit Euler stepping in time. For the unregularised problem the lack of a discrete truncation restricts the proof to the admissible range $q \geq \frac{3d+2}{d+2}$. Under additional assumptions on the finite element setting all steps apart from the identification of the constitutive relation are performed for $q > \frac{2d}{d+2}$. For the regularised problem, as in the steady case one can avoid the extra restrictions arising from the discretisation and show convergence for the full range of existence when taking the regularisation limit separately and after the discretisation limit. Additional difficulties arise from the non-continuous time-dependence of the graph and the discretisation in space requires careful compactness techniques. This result was established in collaboration with Endre Süli, cf. [ST18].

Acknowledgements

First of all I would like to express my deep thankfulness to my advisor Professor Endre Süli for his guidance and support during these past years. It has been a real pleasure to learn from him and guided by him. I very much appreciate his time, effort and enthusiasm with which he approaches both mathematics and people.

I would also like to thank Professors Lars Diening and Yves Capdeboscq for their willingness to act as examiners for this thesis.

My special thanks also goes to Professors Lars Diening, Christian Kreuzer and Sebastian Noelle for their kind respective invitations and the opportunity to present and discuss mathematics in a very pleasant atmosphere.

For helpful discussions and memorable times in and around Prague I would like to thank Professors Josef Málek and Miroslav Bulíček.

Furthermore, I would like to thank the OxPDE group for the great environment to learn and do research. In particular I would like to thank Professor Melanie Rupflin for numerous helpful and enjoyable conversations and her outstanding hospitability; also for helpful remarks on my work I am thankful to her, Dr. Jan Burczak and Professor Jan Kristensen.

I enjoyed stimulating discussions on fluids and more with Jan Blechta, Alexei Gazca, Dr. Giacomo Canevari and Professor Ridgway Scott - thank you for this.

I am thankful to my CDT colleagues for many shared discussions, frustration, joy and cups of coffee; to Matthew Schrecker, Ellya Kawecki, Miles Caddick and my dear office mates, Luca Alasio and Francesco Della Porta: Thank you for making our office a great place to be and work, for many discussions about mathematics and other, and your patience to introduce me to the Italian language and food. Molte grazie!

I also would like to thank the EPSRC and the German Academic Scholarship Foundation for their contribution towards funding my studies in Oxford and before.

Thanks to my friends from Oxford, to Miles, Avner, Liz, Andy, Almut and my three angels Sonja, Sandra and Rachel, for sharing a lot of wonderful times.

A big thank you to my friends from home and abroad, to Lena, Alexey, Jeroen, Gerald and Dorina, who have walked important steps with me and have a part in me being in this place.

I am deeply indebted to my family for their unconditional support, for their confidence in my capabilities, and for letting me find my way. A special thanks goes to my grandfather who would have done anything for his grandchildren to have access to the education that he could not have.

My endless gratitude goes to Franz for his encouragement and backing. You are the song that makes the difference between walking and dancing. Thank you for wonderful times in Oxford and for *more than this*.

Contents

List of Figures

List of Notation

$:$	Frobenius scalar product in $\mathbb{R}^{d\times d}$ 21
$\oplus$	direct sum between function spaces 173
$\otimes$	tensor product between vectors or product σ-algebra 22, 56
$\fint_\omega$	average integral on ω 22
$\mathbb{1}_\omega$	characteristic function of a set ω 22
$\mathcal{A}$	maximal monotone graph, subset of $\mathbb{R}^{d\times d}_{\mathrm{sym}} \times \mathbb{R}^{d\times d}_{\mathrm{sym}}$ 5, 57
$\widetilde{\mathcal{A}}$	maximal monotone graph, subset of $\mathbb{R}^{d\times d}_{\mathrm{sym},0} \times \mathbb{R}^{d\times d}_{\mathrm{sym},0}$ 58
$\mathcal{A}^k$	approximate maximal monotone graph, subset of $\mathbb{R}^{d\times d}_{\mathrm{sym}} \times \mathbb{R}^{d\times d}_{\mathrm{sym}}$ 9
$\mathfrak{A}$	maximal monotone subset of $\mathrm{L}^q(M)^{d\times d} \times \mathrm{L}^{q'}(M)^{d\times d}$ 59
$b(\cdot,\cdot,\cdot)$	trilinear form representing the convective term 42, 88, 115
$\widetilde{b}(\cdot,\cdot,\cdot)$	trilinear form representing the numerical convective term 43, 88, 115
$\mathfrak{B}$	Bogovskiĭ operator 165
$\mathfrak{B}^n$	discrete Bogovskiĭ operator 168
$\mathcal{B}(\mathbb{R}^s)$	σ-algebra of all Borel subsets of $\mathbb{R}^s$ 56
conv	convex hull 22
$\mathrm{C}(\overline{\Omega})$	space of continuous functions on $\overline{\Omega}$ 22
$\mathrm{C}^\infty_0(\Omega)$	space of smooth, compactly supported functions on Ω 22
$\mathrm{C}^\infty_{0,\mathrm{div}}(\Omega)^d$	subspace of divergence-free functions in $\mathrm{C}^\infty_0(\Omega)$ 22
$\mathrm{C}([0,T];\mathrm{X})$	space of continuous functions on $[0,T]$ with values in X 22
$\mathrm{C}^{0,1}([0,T];\mathrm{X})$	space of Lipschitz continuous functions on $[0,T]$ with values in X 22
$\mathrm{C}_w([0,T];\mathrm{X})$	space of weakly continuous functions on $[0,T]$ with values in X 23
dom	domain of a (set-valued) function 54
$\mathbf{A}_\delta$	deviatoric part of a matrix $\mathbf{A}$ 2, 22, 60
d	spacial dimension, $d \in \{2,3\}$ 1, 7, 22, 114

$\mathscr{d}_n$	dimension of $\mathbb{Q}^n$ 34
d_n	dimension of $\mathbb{V}^n_{\text{div}}$ 93
δ_l	time step, for $l \in \mathbb{N}$ 46
∂_t	short notation for partial time derivative 23
d_t	backward temporal difference quotient 46
∂f	subdifferential of a proper, lower semi-continuous convex function f 56
$\mathbf{D}\boldsymbol{v}$	symmetric gradient of the function $\boldsymbol{v} : \mathbb{R}^d \to \mathbb{R}^d$ 2
e	specific internal energy 1
$\boldsymbol{e}_i$	ith unit vector in $\mathbb{R}^d$, $i \in \{1, \ldots, d\}$ 22
$\boldsymbol{\mathcal{E}}$	matrix valued function of q'/q-growth 60
$\boldsymbol{\mathcal{E}}^k$	family of matrix valued function of q'/q-growth, $k \in \mathbb{N}$ 77, 108
$\boldsymbol{f}$	given external force 1, 7, 114
$\{\boldsymbol{f}_i\}_{i\in\{1,\ldots,l\}}$	time discrete external force 117
$\overline{\boldsymbol{f}}$	piecewise constant approximation of external force 117
F_K	affine function mapping K to $\widehat{K}$ 33
$\{F_i\}_{i\in\{1,\ldots,d+1\}}$	faces of an element $K \in \mathcal{T}_n$ 173
$\mathbf{G}$	continuous tensor-valued function representing the implicit constitutive relation 4
γ	exponent in $(-1, \infty)$ depending on q 60, 77, 108
Γ	graph of a (set-valued) function 54
h_K	diameter of element $K \in \mathcal{T}_n$ 33
h_n	grid size, $n \in \mathbb{N}$ 33
$\mathbf{I}$	identity matrix in $\mathbb{R}^{d\times d}$ 2, 22
$\overline{\varphi}$	piecewise constant interpolant in $\mathbb{P}^l_0(0, T; \mathrm{X})$ of $\{\varphi_i\}_{i\in\{0,\ldots,l\}} \subset \mathrm{X}$ 46
$\widetilde{\varphi}$	continuous piecewise affine interpolant in $\mathbb{P}^l_1(0, T; \mathrm{X})$ of $\{\varphi_i\}_{i\in\{0,\ldots,l\}} \subset \mathrm{X}$ 46
$k \in \mathbb{N}$	index denoting the sequence of graph approximations 89, 116
K	closed element in simplicial partition $\mathcal{T}_n$, $n \in \mathbb{N}$ 33
$\widehat{K}$	closed standard reference simplex in $\mathbb{R}^d$ 33
$l \in \mathbb{N}$	index denoting the time-stepping 116
$\mathcal{L}(\omega)$	σ-algebra of all Lebesgue measurable subsets of $\omega \subset \mathbb{R}^n$ 57
$\mathrm{L}^p(\omega)$	Lebesgue space on $\omega \subset \mathbb{R}^n$, $p \in [1, \infty]$, $n \in \mathbb{N}$ 22
$\mathrm{L}^1_{\text{loc}}(\omega)$	space of locally $\mathrm{L}^1(\omega)$ functions 22

$\widehat{\mathbb{P}}_{\mathbb{V}}$ local space for the velocity space $\mathbb{V}^n$ on $\widehat{K}$ 34
$\widehat{\mathbb{P}}_{\mathbb{Q}}$ local space for the pressure space $\mathbb{Q}^n$ on $\widehat{K}$ 34
$\mathbb{P}_0^l(0,T;\mathrm{X})$ function space of left-continuous piecewise constant mappings from $(0,T]$ into X with respect to temporal grid $\{t_0,\dots,t_l\}$, $l \in \mathbb{N}$ 46
$\mathbb{P}_1^l(0,T;\mathrm{X})$ function space of continuous, piecewise affine mappings from $[0,T]$ into X with respect to temporal grid $\{t_0,\dots,t_l\}$, $l \in \mathbb{N}$ 46
Π^n divergence-preserving projector to $\mathbb{V}^n$ 36
P_r^n local projection operator mapping to $\mathbb{V}_r^n$ 36
P_{div}^n L^2-projector to $\mathbb{V}_{\mathrm{div}}^n$, $n \in \mathbb{N}$ 47
$\Pi_{\mathbb{Q}}^n$ projector to $\mathbb{Q}^n$ 35
$\psi_{\mathcal{G}}$ 1-Lipschitz function associated with a maximal monotone set $\mathcal{G} \subset \mathbb{R}^s \times \mathbb{R}^s$ 55
ψ Carathéodory contraction characterising $\mathcal{A}$ 59

q growth and coercivity exponent 5, 57
$\hat{q}$ exponent 45, 115
$\overline{q}$ exponent 88
$\widetilde{q}$ exponent 43, 88
q_d critical exponent depending on $d \in \{2,3\}$ 116
$\widetilde{q}_d$ critical exponent depending on $d \in \{2,3\}$ 116
Q parabolic cylinder $(0,T) \times \Omega$ 7, 22, 114
Q_s^t parabolic cylinder $(s,t) \times \Omega$ 47
Q_s parabolic cylinder $Q_0^s = (0,s) \times \Omega$ 47
Q_i parabolic cylinder $Q_{t_i} = (0,t_i) \times \Omega$ 47
Q_{i-1}^i parabolic cylinder $Q_{t_{i-1}}^{t_i} = (t_{i-1},t_i) \times \Omega$ 47
$\mathbb{Q}^n$ finite element space for the pressure, $n \in \mathbb{N}$ 34
$\mathbb{Q}_0^n$ subspace of $\mathbb{Q}^n$ consisting of functions with zero mean integral 35

ρ density 1
ρ_K interior ball radius of element K 33
$\mathbb{R}_{\mathrm{sym}}^{d\times d}$ subset of symmetric matrices in $\mathbb{R}^{d\times d}$ 3, 21, 57
$\mathbb{R}_{\mathrm{sym},0}^{d\times d}$ subset of symmetric, trace-free matrices in $\mathbb{R}^{d\times d}$ 3, 21, 57
$\mathbb{R}_{\mathrm{sym,tr}}^{d\times d}$ orthogonal complement of $\mathbb{R}_{\mathrm{sym},0}^{d\times d}$ in $\mathbb{R}_{\mathrm{sym}}^{d\times d}$ 60

$\mathbf{S}$ extra (deviatoric) stress tensor 2, 7, 114
$\mathcal{S}$ single-valued stress tensor function with continuous dependence on the shear rate 3
$\mathcal{S}$ possibly set-valued stress tensor function 9, 58
$\mathcal{S}^\star$ selection function corresponding to $\mathcal{A}$ 9, 63
$\mathcal{S}^k$ approximate Carathéodory stress tensor function 9, 64
$\mathcal{S}_i^k$ time discrete approximate stress tensor function 117

CHAPTER 1

Introduction

Since the introduction of the Navier–Stokes equations for incompressible fluids in the 19th century to describe the flow of Newtonian fluids, various modifications and generalisations have been proposed. Implicit constitutive theory provides a very general framework to formulate and investigate such generalisations. In this context we aim to investigate the convergence of numerical approximation schemes for the flow of implicitly constituted fluids. To cover a class of fluids as large as possible, very weak assumptions are imposed. Due to the lack of structure one cannot expect quantitative results in the sense of error estimates, but merely qualitative results. More specifically, we aim to show convergence up to subsequences of approximate solutions to a weak solution, the existence of which was shown in [BGMS09, BGMS12].

1.1. Framework of Implicitly Constituted Incompressible Fluids

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 generalised Newtonian fluids it is usually a power-law-like nonlinear, but still explicit, functional relation. Implicit constitutive theory was introduced to describe a wide range of non-Newtonian rheology admitting also implicit and discontinuous constitutive laws, see [Raj03, Raj06].

In what follows, we give a formal derivation of the equations describing the motion of a fluid in the framework of continuum mechanics and thermodynamics.

1.1.1. Derivation of the Governing Equations

The motion of a fluid is described by a collection of *balance laws*, supplemented by so-called *constitutive relations*, which depend on the material properties.

The balance laws of mass, linear momentum, angular momentum and (specific) internal energy read as follows:

$$\partial_t \rho + \operatorname{div}(\rho \boldsymbol{u}) = 0, \tag{1.1a}$$

$$\rho \partial_t \boldsymbol{u} + \rho \boldsymbol{u} \cdot \nabla \boldsymbol{u} = \operatorname{div} \mathbf{T} + \rho \boldsymbol{f}, \tag{1.1b}$$

$$\mathbf{T} = \mathbf{T}^\top, \tag{1.1c}$$

$$\rho \partial_t e + \rho \boldsymbol{u} \cdot \nabla e = \mathbf{T} : \mathbf{D}\boldsymbol{u} - \operatorname{div} \boldsymbol{j}, \tag{1.1d}$$

in a domain in $\mathbb{R}^d$ for $d \in \{2,3\}$, where ρ is the density, $\boldsymbol{u}$ is the velocity field, $\mathbf{T}$ is the Cauchy stress tensor, $\boldsymbol{f}$ is a given external force, e is the specific internal energy, $\boldsymbol{j}$ is the heat flux and $\mathbf{D}\boldsymbol{u} := \frac{1}{2}(\nabla\boldsymbol{u} + (\nabla\boldsymbol{u})^\top)$ is the symmetric gradient of the velocity field; the latter represents the rate of strain of the fluid flow. For the derivation of these balance laws in continuum mechanics and continuum thermodynamics we refer to [Sch78, Ch. 4], [MR05, Ch. 2] and also to [MP16, Sec. 2]. The *entropy inequality*, which ensures that the second law of thermodynamics is satisfied, can be stated as

$$\mathbf{T} : \mathbf{D}\boldsymbol{u} + p\,\mathrm{div}\,\boldsymbol{u} - \frac{1}{\theta}\boldsymbol{j}\cdot\nabla\theta \geq 0, \tag{1.1e}$$

where $\theta > 0$ is the thermodynamic temperature and p is the thermodynamic pressure, see [MP16, Sec. 2.3.3].

We assume that the fluid is *incompressible*, i.e., the material derivative $\partial_t\rho + \boldsymbol{u}\cdot\nabla\rho$ vanishes. Additionally we assume that the fluid is *homogeneous*, meaning that the density is spatially constant $\rho(t,\boldsymbol{x}) = \rho(t)$. In conjunction with (1.1a) this implies that $0 = \mathrm{div}\,\boldsymbol{u} = \mathrm{tr}(\mathbf{D}\boldsymbol{u})$, where tr denotes the trace of a matrix. Hence $\mathbf{D}\boldsymbol{u}$ is in fact the shear rate tensor. Furthermore, we assume that the temperature $\theta = \theta_0 > 0$ is constant, so the motion is *isothermal*. Assuming a constitutive relation for the specific internal energy of the form $e = e(\rho,\theta)$ the homogeneity assumption and the isothermality imply that e is constant. Then, by the balance law of internal energy, the heat flux $\boldsymbol{j}$ is determined by the Cauchy stress tensor $\mathbf{T}$ and the symmetric velocity gradient $\mathbf{D}\boldsymbol{u}$, which is why we will not consider the energy balance equation any further. For simplicity of presentation we choose $\rho_0 = \theta_0 = 1$. Then, the balance laws (1.1a)–(1.1c) reduce to

$$\mathrm{div}\,\boldsymbol{u} = 0, \tag{1.2a}$$
$$\partial_t\boldsymbol{u} + \mathrm{div}(\boldsymbol{u}\otimes\boldsymbol{u}) = \mathrm{div}\,\mathbf{T} + \boldsymbol{f}, \tag{1.2b}$$
$$\mathbf{T} = \mathbf{T}^\top, \tag{1.2c}$$

and the entropy inequality (1.1e) simplifies to

$$\mathbf{T} : \mathbf{D}\boldsymbol{u} \geq 0. \tag{1.2d}$$

For similar derivations and more details we refer to [Lio13, Ch. 1] and [MP16, Sec. 4.2.3].

Since $\mathbf{D}\boldsymbol{u}$ is trace-free, we infer that

$$\mathbf{T} : \mathbf{D}\boldsymbol{u} = \big(\mathbf{T}_\delta + \tfrac{1}{d}\,\mathrm{tr}(\mathbf{T})\mathbf{I}\big) : (\mathbf{D}\boldsymbol{u})_\delta = \mathbf{T}_\delta : \mathbf{D}\boldsymbol{u},$$

where $\mathbf{I}$ is the identity matrix in $\mathbb{R}^{d\times d}$ and $\mathbf{A}_\delta := \mathbf{A} - \frac{1}{d}\,\mathrm{tr}(\mathbf{A})\mathbf{I}$ is the deviatoric part of a matrix $\mathbf{A} \in \mathbb{R}^{d\times d}$. Now, in the entropy inequality one can replace $\mathbf{T}$ by its deviatoric part, the so-called extra stress tensor (or shear stress tensor) $\mathbf{S} := \mathbf{T}_\delta$. The scalar $\frac{1}{d}\,\mathrm{tr}(\mathbf{T})$ is called mean field stress and its negative $\pi = -\frac{1}{d}\,\mathrm{tr}(\mathbf{T})$ is the mean field pressure, hence $\mathbf{T}$ can be decomposed into

$$\mathbf{T} = \mathbf{S} - \pi\mathbf{I}, \tag{1.3}$$

see also [Sch78, p. 126] and [Lio13, p. 2]. Now (1.2a)–(1.2d) can be stated as

$$\mathrm{div}\,\boldsymbol{u} = 0, \tag{1.4a}$$
$$\partial_t\boldsymbol{u} + \mathrm{div}(\boldsymbol{u}\otimes\boldsymbol{u}) = \mathrm{div}\,\mathbf{S} - \nabla\pi + \boldsymbol{f}, \tag{1.4b}$$
$$\mathbf{S} = \mathbf{S}^\top, \quad \mathrm{tr}(\mathbf{S}) = 0, \tag{1.4c}$$
$$\mathbf{S} : \mathbf{D}\boldsymbol{u} \geq 0. \tag{1.4d}$$

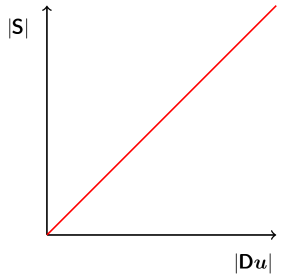

Fig. 1.1: $q = 2$, Navier Stokes (explicit law)

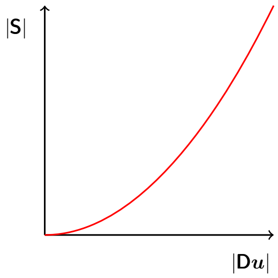

Fig. 1.2: $q > 2$, power-law fluid (explicit law)

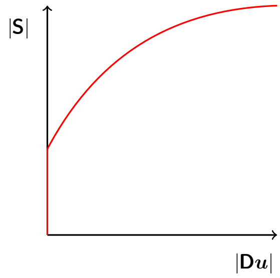

Fig. 1.3: Herschel–Bulkley fluid for $q < 2$

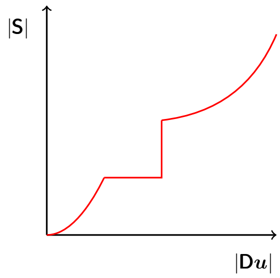

Fig. 1.4: (fictitious) fully implicit relation

Note that (1.4a) and (1.4c) are constraints on $\boldsymbol{u}$ and $\mathbf{S}$, respectively. The pressure π in (1.4b) has the role of a Lagrange multiplier with respect to the incompressibility condition and can be determined once $\boldsymbol{u}$ and $\mathbf{S}$ are known.

In order to close the system one has to impose a relation between the unknown extra stress tensor $\mathbf{S}$ and the other physical quantities in a way that (1.4d) is satisfied. Such a relation is called a *constitutive relation* and can be derived from material properties in two ways: Either from the microstructure of the fluid, some examples of which can be found in [Sch78, Ch. 10, 11], or from the phenomenological behaviour of the fluid, which is the perspective taken, e.g., in [MP16]. Therein the authors derive constitutive relations starting from the second law of thermodynamics and assumptions on the specific internal energy and the specific entropy.

1.1.2. Implicit Constitutive Laws for the Stress Tensor

In the classical literature on thermodynamics it is assumed that the extra stress tensor $\mathbf{S}$ is a function of $\nabla \boldsymbol{u}$ and, by frame indifference and isotropy, it is deduced that $\mathbf{S} = \mathbf{S}(\mathbf{D}\boldsymbol{u})$, see [Sch78, Ch. 9.4] and [Tru91, Ch. 10.2] for details. In this case, the explicit expression for the extra stress $\mathbf{S}$ can be inserted into the balance law of linear momentum (1.4a). In the following, in situations where we consider the relation between the shear rate and the stress tensor we will write $\mathbf{D}$ instead of $\mathbf{D}\boldsymbol{u}$, and denote by $\mathcal{S}$ a function such that $\mathbf{S} = \mathcal{S}(\mathbf{D})$, provided that the dependence is continuous.

Examples for models with explicit constitutive relations include the Navier–Stokes equation and various forms of q-fluid models for $q \in (1, \infty)$: For the Navier–Stokes equation the viscosity $\nu_0 \in (0, \infty)$ is constant and the constitutive law is given by the linear relation

$$\mathcal{S}(\mathbf{D}) = 2\nu_0 \mathbf{D}, \tag{1.5a}$$

for $\mathbf{D} \in \mathbb{R}^{d\times d}_{\mathrm{sym},0}$, where $\mathbb{R}^{d\times d}_{\mathrm{sym}}$ is the space of symmetric $d \times d$ matrices and $\mathbb{R}^{d\times d}_{\mathrm{sym},0}$ is the subspace of zero trace matrices. The relation of the absolute values of $\mathbf{S}$ and $\mathbf{D}$ is shown in Fig. 1.1. For q-fluid models the constitutive relation is given by

$$\mathcal{S}(\mathbf{D}) = 2\nu_0 \left(\delta + |\mathbf{D}|^{q-2}\right) \mathbf{D}, \tag{1.5b}$$

for $q \in (1, \infty)$ and $\delta \in [0, \infty)$ constant. For $\delta = 0$ this is the Ostwald–de Waele power law model, and for $\delta \in (0, \infty)$ its (non-degenerate) generalisation. The nonlinear relations (for $q \neq 2$) feature a viscosity $\nu > 0$ depending on the shear rate; more precisely there exists a function $\nu\colon \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ such that

$$\mathcal{S}(\mathbf{D}) = \nu(|\mathbf{D}|)\mathbf{D},$$

for $\mathbf{D} \in \mathbb{R}^{d\times d}_{\mathrm{sym},0}$. For more examples of explicit constitutive relations and the original references we refer to [MP16, Sec. 4.5] and the references contained therein.

There are fluids for which the extra stress tensor cannot be described by a (continuous) function of the shear rate. For the class of Herschel–Bulkley fluids the shear rate is non-zero only once the Euclidean norm of the extra stress tensor exceeds a given constant yield stress $\tau_* > 0$. These fluids belong to the class of *activated fluids* and for $q \in (1, \infty)$ the constitutive law is given by

$$\begin{cases} |\mathbf{S}| \leq \tau_* & \Leftrightarrow \quad \mathbf{D} = \mathbf{0}, \\ |\mathbf{S}| > \tau_* & \Leftrightarrow \quad \mathbf{S} = \tau_* \frac{\mathbf{D}}{|\mathbf{D}|} + |\mathbf{D}|^{q-2}\mathbf{D} \quad \Leftrightarrow \quad \mathbf{D} = (|\mathbf{S}| - \tau_*)^{\frac{1}{q-1}} \frac{\mathbf{S}}{|\mathbf{S}|}, \end{cases} \tag{1.6}$$

for $\mathbf{D}, \mathbf{S} \in \mathbb{R}^{d\times d}_{\mathrm{sym},0}$, see Fig. 1.3, [BGMS12, Sec. 1.1], [MP16, Sec. 4.5] and the references therein. For $q = 2$, (1.6) reduces to the constitutive law for Bingham fluids, see [DL76, Ch. VI.1.2]. Evidently $\mathbf{D}$ can be expressed as an explicit continuous function in $\mathbf{S}$, but to describe $\mathbf{S}$ as function of $\mathbf{D}$ one has to rely on discontinuous functions or set-valued mappings. However, the constitutive relation can equivalently be stated as

$$|\mathbf{D}|^{q-2}\,(\tau_* + (|\mathbf{S}| - \tau_*)^+)\mathbf{D} - (|\mathbf{S}| - \tau_*)^+\mathbf{S} = \mathbf{0}, \tag{1.7}$$

for $\mathbf{D}, \mathbf{S} \in \mathbb{R}^{d\times d}_{\mathrm{sym},0}$, where $(s)^+ := \max(s, 0)$ denotes the positive part of $s \in \mathbb{R}$. Note that the left-hand side of the expression is continuous in $\mathbf{S}$ and $\mathbf{D}$. This suggests considering the more general class of constitutive relations as in [BGMS12], which can be written as

$$\mathbf{G}(\mathbf{D}, \mathbf{S}) = \mathbf{0}, \tag{1.8}$$

with a continuous tensor-valued function $\mathbf{G}$. Since this includes constitutive laws, for which $\mathbf{S}$ and $\mathbf{D}$ are related in a fully implicit manner, for an example see Fig. 1.4, we refer to such a relation as *implicit constitutive relation*.

Implicit Constitutive Theory

Examples of implicit constitutive relations have been known for a long time. In 1845 Stokes ([Sto45]) remarked that the viscosity is independent of the pressure only under suitable conditions. Indeed, if the viscosity of an incompressible fluid depends on the pressure $\pi = -\frac{1}{d}\operatorname{tr}(\mathbf{T})$, then the constitutive relation between the Cauchy stress tensor $\mathbf{T}$ and the shear rate $\mathbf{D}$, given by

$$\mathbf{T} = -\pi\mathbf{I} + \nu(\pi, \mathbf{D})\mathbf{D} = \tfrac{1}{d}\operatorname{tr}(\mathbf{T})\mathbf{I} + \nu\left(\tfrac{1}{d}\operatorname{tr}(\mathbf{T}), \mathbf{D}\right)\mathbf{D} \tag{1.9}$$

is implicit, see [Raj03, Sec. 4.1] and [Raj06]. Another situation, in which implicit relations arise naturally, is the flow of viscoelastic fluids. These constitutive relations, involving both the stress tensor and certain time-derivatives thereof, were first described by Maxwell in 1867 ([Max67]), see [Raj03, Sec. 4.3]. Furthermore, if the shear rate $\mathbf{D}$ non-monotonically dependends on $\mathbf{T}_\delta$ for certain fluids, it is not possible to write $\mathbf{T}_\delta$ as function of $\mathbf{D}$, see [RS16].

A general framework of *implicit constitutive theory* for fluids was developed in [Raj03, MR05, Raj06, RS08, MP16] to mention just a view. In the situation of (1.2a)–(1.2d), an implicit relation of the form

$$\mathbf{f}(\mathbf{T},\mathbf{D},\rho,\theta)=\mathbf{0} \tag{1.10}$$

can be considered. Even more generally, the implicit constitutive law may include material derivatives of $\mathbf{D}$ and $\mathbf{T}$, such that the framework also includes rate-type models. Physical principles, such as frame-indifference and isotropy (as well as the entropy inequality), impose restrictions on the constitutive relations, see for example [Raj06]. One advantage of considering such general relations is that the framework includes a large class of fluids models with non-Newtonian behaviour. Furthermore, it provides a unified approach to the derivation of previously proposed constitutive laws, and gives rise to new models. From a thermodynamical point of view another strength of implicit constitutive theory lies in the fact that it allows one to derive thermodynamically consistent constitutive laws starting from the entropy inequality rather than imposing it afterwards, for details see [MP16].

1.1.3. Assumptions on the Implicit Constitutive Relation

Returning to the flow of a homogeneous, incompressible fluid in an isothermal situation and assuming that the constitutive relation is independent of the pressure π, the relation (1.10) reduces to the relation (1.8), which is the object of this work.

Following [BGMS09, p. 110] and [BGMS12, Sec. 1.2] we assume that the relation given by $\mathbf{G}$ in (1.8) can be identified with a maximal monotone set $\mathcal{A}\subset\mathbb{R}^{d\times d}_{\mathrm{sym}}\times\mathbb{R}^{d\times d}_{\mathrm{sym}}$ as

$$\mathbf{G}(\mathbf{D},\mathbf{S})=\mathbf{0}\quad\Leftrightarrow\quad(\mathbf{D},\mathbf{S})\in\mathcal{A}, \tag{1.11}$$

in a pointwise sense. Since $\mathcal{A}$ can be viewed as the graph of a maximal monotone set-valued function, we refer to $\mathcal{A}$ as a *maximal monotone graph*. Furthermore, let us sketch the assumptions on $\mathcal{A}$, which we will impose:

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

(ii) $\mathcal{A}$ is a maximal monotone set, i.e., for all $(\mathbf{D}_1,\mathbf{S}_1),(\mathbf{D}_2,\mathbf{S}_2)\in\mathcal{A}$,

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

and $\mathcal{A}$ is maximal in the class of monotone sets with respect to inclusion;

(iii) There exists a constant $c_*>0$ and a $q\in(1,\infty)$ such that

$$\mathbf{D}:\mathbf{S}\geq c_*(|\mathbf{D}|^q+|\mathbf{S}|^{q'})\quad\text{for all }\ (\mathbf{D},\mathbf{S})\in\mathcal{A},$$

where q' is the Hölder conjugate of q.

We postpone the precise assumptions on $\mathcal{A}$ used in this work to Assumption 3.11 in Chapter 3. In particular, accounting for possible extensions where the constitutive relation depends on more quantities like the temperature θ ([MŽ18]), the pressure π ([FMR05, MÓ7, Bv15]) or a concentration, we will consider the case where $\mathbf{G}$ and hence also $\mathcal{A}$ depends on points in time and space $(t,\boldsymbol{x})$.

If we consider $\widetilde{\mathcal{A}}\subset\mathbb{R}^{d\times d}_{\mathrm{sym},0}\times\mathbb{R}^{d\times d}_{\mathrm{sym},0}$ with the same properties, the trace constraint on $\mathbf{S}$ in (1.4c) is automatically satisfied and considering trace-free matrices suffices when working with divergence-free functions. Here however, we will adapt the assumptions on $\mathcal{A}\subset\mathbb{R}^{d\times d}_{\mathrm{sym}}\times\mathbb{R}^{d\times d}_{\mathrm{sym}}$ in a way that the constraints on $\mathbf{S}$ are enforced and hence the framework is consistent with the thermodynamic derivation. By the monotonicity of $\mathcal{A}$ and the fact that $(\mathbf{0},\mathbf{0})$, one has that the entropy condition (1.4d) and thus the second law of thermodynamics is satisfied. However, the assumptions on $\mathcal{A}$ do not include any frame-indifference properties, which means that those have to be checked individually for specific models of interest.

The resulting framework includes explicit relations as in (1.5b): Newtonian fluids, where $q = 2$ as well as generalised Newtonian fluids with shear-thinning behaviour ($q \in (1, 2)$), as for example paint, certain hydrogels, blood and ketchup, and shear-thickening fluids ($q > 2$), such as a suspension of starch and water, see also Fig. 1.2. Further, relations, where the stress is a set-valued or discontinuous function of the symmetric gradient, but the symmetric gradient can be expressed as continuous function of the stress tensor, are included. This is the case for Bingham or Herschel–Bulkley fluids, such as drilling mud and certain toothpaste. Fully implicit constitutive relations, for which neither the stress not the symmetric gradient can be expressed as continuous function of the other (see Fig. 1.4), are included as well, even though at present we are not aware of any examples.

Models including viscoelasticity (or in fact any rate-dependence of the constitutive relation) and pressure-dependent constitutive relations are not covered by the assumptions. Furthermore, there are fluids, which do not satisfy the polynomial growth and coercivity condition with respect to some $q \in (1, \infty)$, like the Prandtl-Eyring fluid, see [BDF12].

Thanks to the generality of the framework the existence results of weak solutions proved in [BGMS09, BGMS12] for the steady and the unsteady case, respectively, generalise a number of previous existence results both on explicit constitutive relations and relations for which the stress tensor is discontinuous in the shear rate ([DL76, Ser91, FS00, MRS05, ER12] and [GMS07]). Furthermore, since the implicit constitutive relation is imposed in a pointwise sense almost everywhere in the domain, the notion of solution is stronger than notions based on weak inequalities, where the stress tensor is not a part of the weak solution, see [DL76, Ser91, FS00].

A Note on Boundary Conditions

The boundary condition on the boundary $\partial\Omega$ of the domain $\Omega \subset \mathbb{R}^d$ occupied by the fluid depends on the properties of the materials involved. Thus one can view the boundary condition as a constitutive relation, see [MP16]. The impermeability of the boundary can be phrased as

$$\boldsymbol{u} \cdot \boldsymbol{\nu} = 0 \quad \text{on } \partial\Omega, \tag{1.12a}$$

where $\boldsymbol{\nu}(\boldsymbol{x})$ denotes the outward unit normal vector in a point $\boldsymbol{x} \in \partial\Omega$. Denoting by $\boldsymbol{w}_{\boldsymbol{\tau}} := \boldsymbol{w} - (\boldsymbol{w} \cdot \boldsymbol{\nu})\boldsymbol{\nu}$ one specific tangential component of a vector $\boldsymbol{w} \in \mathbb{R}^d$ at $\boldsymbol{x} \in \partial\Omega$ we impose the following condition on the tangential velocity

$$\boldsymbol{u}_{\boldsymbol{\tau}} = -\kappa_* \left(\mathbf{S}\boldsymbol{\nu}\right)_{\boldsymbol{\tau}}, \tag{1.12b}$$

for $\kappa_* \geq 0$. In case $\kappa_* = 0$ is constant this reduces to the so-called *no-slip* (homogeneous Dirichlet) condition. For constant $\kappa_* > 0$ the resulting condition is called *Navier's slip* boundary condition. More generally, κ_* could be related to $|(\mathbf{S}\boldsymbol{\nu})_{\boldsymbol{\tau}}|$ and $|\mathbf{D}|$ in an explicit or implicit way, which includes also the stick-slip condition investigated for example in [BM16]. This fits in the framework of implicit constitutive relations but will not be subject to investigation in this work, instead we will focus on the cases when κ_* is constant.

From a thermodynamic point of view the no-slip boundary condition has for a long time been the prevailing one and only relatively recently have alternatives been examined, see [Den04]. Mathematically a Navier's slip boundary condition or even periodic boundary conditions are preferable since they allow for better regularity, see Subsection 1.2.6 below.

1.2. Existence of Weak Solutions

Let us state the full problem introduced in (1.4a)-(1.4c) and (1.11) subject to homogeneous Dirichlet boundary conditions ((1.12a) and (1.12b) with $\kappa_* = 0$).

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 $T \in (0,\infty)$. We assume that $\boldsymbol{f}\colon Q \to \mathbb{R}^d$ is a given external force, that $\boldsymbol{u}_0\colon \Omega \to \mathbb{R}^d$ is an initial velocity field and that $\mathcal{A} \subset \mathbb{R}^{d\times d}_{\mathrm{sym}} \times \mathbb{R}^{d\times d}_{\mathrm{sym}}$ is a maximal monotone graph satisfying the assumptions outlined in Section 1.1.3.

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) extra stress tensor field $\mathbf{S}\colon Q \to \mathbb{R}^{d\times d}_{\mathrm{sym},0}$ satisfying

$$\begin{aligned} \partial_t \boldsymbol{u} + \operatorname{div}(\boldsymbol{u} \otimes \boldsymbol{u}) - \operatorname{div} \mathbf{S} &= -\nabla \pi + \boldsymbol{f} && \text{in } Q, \\ \operatorname{div} \boldsymbol{u} &= 0 && \text{in } Q, \end{aligned} \tag{1.13}$$

subject to the initial condition and no-slip boundary condition

$$\boldsymbol{u}(0,\cdot) = \boldsymbol{u}_0(\cdot) \qquad \text{in } \Omega, \tag{1.14}$$

$$\boldsymbol{u} = \mathbf{0} \qquad \text{on } (0,T) \times \partial\Omega, \tag{1.15}$$

and satisfying the constitutive relation

$$(\mathbf{D}\boldsymbol{u}(\cdot), \mathbf{S}(\cdot)) \in \mathcal{A}(\cdot) \quad \text{pointwise a.e. in } Q. \tag{1.16}$$

In the following we refer to the unsteady problem consisting of (1.13)–(1.16) as **(PU)**. The corresponding steady problem, such that $\partial_t \boldsymbol{u} = \mathbf{0}$, will be denoted by **(PS)**. The precise notions of weak solutions we aim for are introduced in Chapter 2 and the precise assumptions on $\mathcal{A}$ will be postponed to Chapter 3.

1.2.1. Galerkin Approximation

To prove existence of weak solutions of the above problem or of special cases thereof, one of the common approaches is a Galerkin approximation in Ω. In addition to the Galerkin approximation in a suitable function space, further approximations may be included in order to guarantee existence of an approximate solution or in order to simplify the argument. Due to the properties of $\mathcal{A}$ and Korn's and Poincarés inequality the usual a priori estimates yield uniform bounds on the velocity in $\mathrm{L}^\infty(0,T;\mathrm{L}^2(\Omega)^d) \cap \mathrm{L}^q(0,T;\mathrm{W}^{1,q}_0(\Omega)^d)$ (in the steady case in $\mathrm{W}^{1,q}_0(\Omega)^d$) and on the stress tensor in $\mathrm{L}^{q'}(Q)^{d\times d}$ (in the steady case in $\mathrm{L}^{q'}(\Omega)^{d\times d}$). Then, by compactness arguments weak (and strong) convergence of subsequences can be obtained. Finally we have to identify the limiting equation and the constitutive relation, the latter of which is the key part of the proof.

1.2.2. Nonlinearities and Challenges

The main challenges arise from the two nonlinearities in the problem: the convective term and the constitutive relation.

The Convective Term and Admissibility

Writing M for Q or Ω, respectively, the weak form of the convective term is given by

$$\int_M (\boldsymbol{u} \otimes \boldsymbol{u}) : \nabla \boldsymbol{w} \,\mathrm{d}\boldsymbol{z}$$

for a suitable test function $\boldsymbol{w}$. This term is well-posed in the steady case if $q \geq \frac{2d}{d+2}$, because then we have that $\boldsymbol{u} \in \mathrm{W}_0^{1,q}(\Omega)^d \hookrightarrow \mathrm{L}^2(\Omega)^d$, and in the unsteady case without extra condition on q, since $\boldsymbol{u} \in \mathrm{L}^\infty(0,T;\mathrm{L}^2(\Omega)^d)$.

However, to identify the term in the limiting equation when taking the Galerkin limit, strong convergence of the sequence of velocities is required. In both the steady and the unsteady case for compactness we require that $q > \frac{2d}{d+2}$. For small $q \in (1,\infty)$ the term causes lack of admissibility. More precisely, if $q < \frac{3d+2}{d+2}$ in the unsteady case (and if $q < \frac{3d}{d+2}$ in the steady case), then $\boldsymbol{u}$ is in general not an admissible test function in the weak form of the convective term.

Identification of the Constitutive Relation for Explicit Laws

The lack of admissibility caused by the convective term is problematic, because the straightforward way to identify the constitutive relation includes testing the equation with the velocity function $\boldsymbol{u}$. Let us collect the options available for the identification of the constitutive relation for explicit laws used in the literature (see [FM03]):

- In the admissible case an energy identity is available and by use of monotone operator theory (Minty's trick, see the following subsection) the constitutive relation can be identified, see [Lad69, Lio69].
- If higher regularity of the velocity can be established this may restore admissibility. Thus, the constitutive relation can be identified also for q in part of the range in which admissibility is not guaranteed a priori.
- If $\boldsymbol{u}$ is not an admissible test function, one can recover admissibility by truncation. Instead of testing with the velocity function, we test with a truncation, which is in a smaller function space and hence admissible. Such a truncation can be constructed in a way that it coincides with the original function on a large part of the domain. There are the following two type of truncation:
  - the L^∞-truncation, covering the range $q > \frac{2(d+1)}{d+2}$ in the unsteady case, ([Wol07]), and $q > \frac{2d}{d+1}$ in the steady case ([FMS97, Růž97]).
  - and the Lipschitz approximation suitable for the whole range $q > \frac{2d}{d+2}$ in both the steady and the unsteady case, see for example [FMS03, DMS08, DRW10]. In Subsection 1.2.5 we will provide more detail.

For implicitly constituted fluid the situation is a bit more difficult: On the one hand the higher regularity approach is not available due to the lack of structure of $\mathcal{A}$. And on the other hand, when applying the Lipschitz truncation technique additional work has to be done to identify the implicit constitutive relation, using various L^1 weak compactness results and a generalised Minty type convergence lemma, see [BGMS12].

1.2.3. Monotonicity and Minty Type Convergence Lemma

Whenever monotonicity (but not necessarily strict monotonicity) of the constitutive relation is available, Minty's trick and its generalisations are a powerful tool to identify the constitutive relation. Denoting by $\{\mathbf{D}\boldsymbol{u}^N\}_{N\in\mathbb{N}}$ and $\{\mathbf{S}^N\}_{N\in\mathbb{N}}$ the sequences of approximate shear rates and approximate stress tensors, we assume that $(\mathbf{D}\boldsymbol{u}^N, \mathbf{S}^N)$ are related by the constitutive relation, for each $N \in \mathbb{N}$. Furthermore, from the a priori estimates weak convergence (up to subsequences) can be obtained in the sense that $\mathbf{D}\boldsymbol{u}^N \rightharpoonup \mathbf{D}\boldsymbol{u}$ weakly in $\mathrm{L}^q(M)^{d\times d}$ and $\mathbf{S}^N \rightharpoonup \overline{\mathbf{S}}$ weakly in $\mathrm{L}^{q'}(M)^{d\times d}$, as $N \to \infty$. Now we want to show that the limiting functions $(\mathbf{D}\boldsymbol{u}, \overline{\mathbf{S}})$ are related by the constitutive relation:

(i) If the constitutive relation is explicit such that $\mathbf{S}^N = \mathcal{S}(\mathbf{D}\boldsymbol{u}^N)$ for a continuous monotone function $\mathcal{S}$, then one can apply the *standard Minty trick* (e.g., [Rou13, Lem. 2.13]):

If one can prove that

$$\limsup_{N\to\infty} \int_M \mathbf{S}^N : \mathbf{D}\boldsymbol{u}^N \,\mathrm{d}\boldsymbol{z} \le \int_M \overline{\mathbf{S}} : \mathbf{D}\boldsymbol{u} \,\mathrm{d}\boldsymbol{z}, \tag{1.17}$$

then for all $\mathbf{A} \in \mathrm{L}^q(M)^{d\times d}$ with values in $\mathbb{R}^{d\times d}_{\mathrm{sym}}$ by monotonicity it follows that

$$0 \le \limsup_{N\to\infty} \int_M \left(\mathcal{S}(\mathbf{D}\boldsymbol{u}^N) - \mathcal{S}(\mathbf{A})\right) : \left(\mathbf{D}\boldsymbol{u}^N - \mathbf{A}\right) \mathrm{d}\boldsymbol{z} \le \int_M \left(\overline{\mathbf{S}} - \mathcal{S}(\mathbf{A})\right) : \left(\mathbf{D}\boldsymbol{u} - \mathbf{A}\right) \mathrm{d}\boldsymbol{z}.$$

Choosing $\mathbf{A} = \mathbf{D}\boldsymbol{u} \pm \varepsilon\mathbf{B}$, for $\varepsilon > 0$, and $\mathbf{B} \in \mathrm{L}^q(M)^{d\times d}$ with values in $\mathbb{R}^{d\times d}_{\mathrm{sym}}$, dividing by ε and letting $\varepsilon \to 0$ yields by continuity of $\mathcal{S}$ that $\overline{\mathbf{S}} = \mathcal{S}(\mathbf{D}\boldsymbol{u})$, which shows that the constitutive relation is satisfied. Whenever $\boldsymbol{u}$ is an admissible test function in the weak formulation of the problem, (1.17) can be verified using energy identities for $\boldsymbol{u}^N$ and $\boldsymbol{u}$. Otherwise, (1.17) or a local version thereof can be shown by testing the approximate equation with a truncation of the sequence $\{\boldsymbol{u}^N - \boldsymbol{u}\}_{N\in\mathbb{N}}$.

(ii) For implicit constitutive relations a generalisation of the above is available which does not rely on continuity, first proved in [BGMS12] and reproved in [BM16]. This *Minty type convergence lemma* states, that if the implicit relation is satisfied pointwise almost everywhere in the domain, i.e., $(\mathbf{D}\boldsymbol{u}^N(\cdot), \mathbf{S}^N(\cdot)) \in \mathcal{A}(\cdot)$ a.e. in M for the weakly converging sequences and (1.17) is satisfied, then it follows that $(\mathbf{D}\boldsymbol{u}(\cdot), \mathbf{S}(\cdot)) \in \mathcal{A}(\cdot)$ a.e. in M. In fact using the pointwise properties of $\mathcal{A}$ the authors prove a local version for subdomains $\widetilde{M} \subset M$. This makes the lemma applicable also when using truncations, see Lemma 3.16.

1.2.4. Graph Approximation for Implicitly Constituted Fluids

For explicit constitutive relations the fact that $\mathcal{S}$ is continuous suffices to show existence of solutions to the Galerkin approximate problem. In case $\mathcal{S}$ depends also on points in the domain the corresponding assumption would be that $\mathcal{S}\colon M \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$ is a Carathéodory function.

For implicitly constituted fluids $\mathcal{A}$ has to be approximated by a sequence of graphs $\mathcal{A}^k$ such that $(\mathbf{D}, \mathbf{S}) \in \mathcal{A}^k(\cdot)$ if and only if $\mathbf{S} = \mathcal{S}^k(\cdot, \mathbf{D})$ for a (single-valued) Carathéodory function $\mathcal{S}^k\colon M \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$. There are various ways to achieve this, one of which is smoothing a possibly discontinuous selection function $\mathcal{S}^\star\colon M \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$ of the possibly multivalued function $\mathcal{S}\colon M \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \rightrightarrows \mathbb{R}^{d\times d}_{\mathrm{sym}}$ such that $(\mathbf{D}, \mathbf{S}) \in \mathcal{A}(\cdot)$ if and only if $\mathbf{S} \in \mathcal{S}(\cdot, \mathbf{D})$, for $\mathbf{D}, \mathbf{S} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$. This approach is chosen in the existence proofs in [BGMS09, BGMS12] by means of mollification of $\mathcal{S}^\star$ in the second argument. In the context of systems of divergence-form equations formulated with a maximal monotone operator in [FMT04] further graph approximations have been investigated, see also Chapter 3. The minimum requirement on such a graph approximation is that one can take the limit $k \to \infty$ before taking the Galerkin limit $N \to \infty$. Then, the Galerkin approximate solutions satisfy the implicit relation encoded by $\mathcal{A}$, and hence one can use a Minty type convergence lemma for the identification of the implicit relation in the limit $N \to \infty$.

However, classes of graph approximations, for which one can show a Minty type convergence lemma itself, are even more attractive. More precisely, instead of $(\mathbf{D}\boldsymbol{u}^N, \mathbf{S}^N) \in \mathcal{A}(\cdot)$ a.e. in the domain one assumes that $(\mathbf{D}\boldsymbol{u}^{k,N}, \mathbf{S}^{k,N}) \in \mathcal{A}(\cdot)^k$ a.e. in the domain. If this allows us to show that $(\mathbf{D}\boldsymbol{u}, \overline{\mathbf{S}}) \in \mathcal{A}(\cdot)$ a.e. in the domain, then one can take the limit in the graph approximation $k \to \infty$ simultaneous with the Galerkin limit $N \to \infty$, see Lemma 3.31.

1.2.5. Lipschitz Approximation

As mentioned before the Lipschitz truncation technique is a method to recover admissibility in the convective term for small $q \in (1, \infty)$ and to verify condition (1.17) on suitable subdomains. It can be used to approximate a Sobolev function by a sequence of Lipschitz functions, so that each member of the sequence is an admissible test function in the weak formulation of the convective term. The construction is based on a truncation at the level of the Hardy–Littlewood maximal operator of the gradients resulting in a sequence of functions with increasing Lipschitz constant such that the set on which they differ from the original function has small (decreasing) measure.

The Lipschitz truncation method first appeared in [AF88] and since then was further developed and refined in a series of papers, see, e.g., [KL02, FMS03, DMS08, DRW10, BDF12, BDS13, DKS13a] to mention just a few. In particular there are versions available, which preserve zero-traces, various forms of parabolic truncations have been introduced for which an equation is truncated rather than a function. Furthermore, there are truncations which preserve solenoidality of the given Sobolev function both in the steady and the unsteady case, which is especially useful if the pressure behaves badly because solenoidality of the sequence of truncations means that one can work with the pressure-free formulation. Finally, for the steady case also a discrete Lipschitz approximation is available. For each of these truncations a sequential version can be obtained, such that the sequence of truncations has improved convergence properties for a fixed truncation level.

Survey on the Existence Result in the Steady Case

Let us give a brief summary of the existence proof for the steady problem **(PS)**, in [BGMS09]. The authors show existence of weak solutions consisting of velocity, stress tensor and pressure to the steady problem subject to homogeneous Dirichlet boundary conditions in the natural Sobolev and Lebesgue spaces for $q > \frac{2d}{d+2}$. Additionally to the assumptions on $\mathcal{A}$ mentioned above it is assumed that $\mathcal{A}$ is strictly monotone in a generalised sense. They set up an approximate problem which includes a graph approximation based on mollification of a selection function, a Galerkin approximation in the space $\mathrm{W}^{1,q}(\Omega)^d \cap \mathrm{L}^{2q'}(\Omega)^d$ for the velocity and a regularisation term giving additional $\mathrm{L}^{2q'}$-integrability. First they take the Galerkin limit and then simultaneously the graph approximation limit and the limit corresponding to the regularisation. In order to identify the implicit constitutive relation they apply a divergence corrected steady Lipschitz approximation. Further, since the (generalised) Minty type convergence lemma for implicit constitutive relations was not available at the time, they employ Young measures combined with weak compactness results in L^1 to identify the constitutive law.

1.2.6. Pressure and Boundary Conditions

In the steady case the reconstruction of the pressure as an integrable function can be done once the velocity and the stress tensor functions are known by means of the de Rham theorem and Nečas negative norm theorem.

In the unsteady case the interaction between the pressure π and the time derivative $\partial_t \boldsymbol{u}$ in combination with the homogeneous Dirichlet boundary condition on a Lipschitz domain complicate things. Considering the pressure-free equation one can show that $\partial_t \boldsymbol{u} \in (\mathrm{L}^s(0, T; \mathrm{W}^{1,s}_{0,\mathrm{div}}(\Omega)^d))'$ for some $s \geq q$. Because $\partial_t \boldsymbol{u}$ is not a distribution in space on the pressure one obtains initially only that $\pi \in \mathrm{W}^{-1,\infty}(0, T; \mathrm{L}^2(\Omega))$, i.e., it is a distribution in time, compare [Tem84, Ch. III, § 3, pp. 307, Rem. 3.8] for the Navier–Stokes equations. On smoother domains the regularity of the pressure can be improved and hence locally one obtains $\pi \in L^1_{loc}(Q)$ or even better, see [Wol07, DRW10].

For $\partial\Omega \in C^{1,1}$ and alternative boundary conditions the pressure can be shown to be globally integrable, which in turn improves the regularity of $\partial_t \boldsymbol{u}$. This is the case for Navier's slip boundary condition, see [FM03, BMR09, BGMS12] and more recently this was shown for stick-slip boundary conditions, see [BM16].

For a numerical approximation by means of finite element functions one cannot handle $\partial\Omega \in C^{1,1}$ without extra difficulty, so none of the above approaches to obtain globally integrable pressure is available. If $\pi \in L^1_{loc}(Q)$ one has two options to identify the implicit constitutive relation: either decomposing the pressure locally and working with a local parabolic Lipschitz truncation, as done in [DRW10, ER12] and [BGMS12]; or applying the parabolic solenoidal Lipschitz approximation presented in [BDS13], with which one can stay in the pressure-free formulation.

Survey on the Existence Result in the Unsteady Case

Since it is the starting point of the convergence analysis let us give an outline of the existence proof in [BGMS12] for weak solutions to the unsteady problem.

The assumptions on $\mathcal{A}$ are even weaker, formulated for an N-function ψ and giving rise to an Orlicz–Sobolev space setting (with Δ_2- and ∇_2-condition) rather than to a Sobolev space setting. Then, existence of weak solutions (consisting of velocity, pressure and stress tensor) of the unsteady problem subject to Navier's slip boundary condition on the boundary $\partial\Omega \in C^{1,1}$ is proved for $q > \frac{2d}{d+2}$. For the proof a three level approximation is introduced: a graph approximation based on mollification of a selection function $\boldsymbol{\mathcal{S}}^\star$, a Galerkin approximation in spaces of solenoidal eigenfunctions of a higher order eigenvalue problem and an $L^{2q'}$-regularisation term restoring admissibility of the approximate solution. First the graph approximation limit is taken and since the dimension stays finite, strong convergence is obtained and the implicit constitutive relation can be identified by the properties of the graph approximation. Thanks to the regularising the approximate solutions are still admissible after the Galerkin limit. Hence an energy identity is available and the identification can be shown by the aforementioned Minty type convergence lemma. Before taking the final limit corresponding to the regularisation, the sequence of pressure functions is reconstructed in $L^{1+\varepsilon}(Q)$ for some $\varepsilon > 0$, which is achieved by solving certain suitable Poisson problems subject to a Neumann boundary condition. Finally, after taking the limit a Lipschitz truncation method tailored to the Orlicz setting is applied, which - supplemented by some L^1 weak compactness results - allows one to verify the assumptions of the Minty type convergence lemma to show the implicit constitutive relation.

1.2.7. Limitations: Regularity and Uniqueness in the Unsteady Case

For weak solutions to the problem with implicit constitutive law subject to homogeneous Dirichlet boundary conditions on a Lipschitz domain we cannot hope for more regularity than the one given by the a priori estimates because $\mathcal{A}$ does not have any structure. Also uniqueness of the velocity in the unsteady problem cannot be hoped for unless one considers the Stokes problem neglecting the convective term or one assumes a suitable strong monotonicity property of $\mathcal{A}$ and further restriction of q, see [Lio69, Ch. II.5.3] and [MNRR96, Thm. 4.29] for the corresponding proofs for explicit laws. Even if uniqueness of $\boldsymbol{u}$ should be given, uniqueness of the stress tensor is not clear unless $\mathcal{A}$ is strongly monotone.

For comparison let us mention some of the available results for explicit constitutive laws of the form (1.5b) in the *unsteady case*: Uniqueness for the non-degenerate version of (1.5b), i.e., $\delta \in (0, \infty)$, subject to homogeneous Dirichlet boundary conditions was shown in [Lad67, Lad69, Lio69] for $q \geq \frac{d+2}{2}$ and later extended to $q \geq \frac{11}{5}$ if $d = 3$ for sufficiently smooth initial data, see [BEKP10, BKP18]. For smaller q uniqueness is not known, not even for the special case of Navier-Stokes equations ($q = 2$) if $d = 3$.

If $\delta > 0$ and one considers the problem subject to homogeneous Dirichlet boundary conditions on sufficiently smooth domains, then global in time strong solutions exist for $q > \frac{11}{5}$ if $d = 3$ ([BadVKR11]), and for $q \geq 2$ if $d = 2$ ([MNR01]). Furthermore, local in time strong solutions exist for $q > 1$ in both cases $d \in \{2, 3\}$ ([BP07]). In the case of space periodic boundary conditions the available results are slightly better. For $d = 2$ global in time solutions exist for the larger range $q > 1$, see [MNRR96] for global and [DR05] for local in time existence of strong solutions. In the degenerate case, i.e., $\delta = 0$, there are few regularity results available, see [BJ14] for the Dirichlet problem and [BDR10] for the periodic problem.

d	$\frac{2d}{d+2}$	$\frac{2d}{d+1}$	$\frac{2(d+1)}{d+2}$	$\frac{3d}{d+2}$	q_d	$\widetilde{q}_d$	$\frac{3d+2}{d+2}$
2	1	$\frac{4}{3}$	$\frac{3}{2}$		$\frac{1}{2}(1+\sqrt{5}) \approx 1.62$	$1 + \frac{1}{\sqrt{2}} \approx 1.71$	2
3	$\frac{6}{5}$	$\frac{3}{2}$	$\frac{8}{5}$	$\frac{9}{5}$	$\frac{1}{5}\left(3+\sqrt{39}\right) \approx 1.85$	$\frac{1}{10}\left(11+\sqrt{61}\right) \approx 1.88$	$\frac{11}{5}$

Fig. 1.5: Values of exponents relevant for the steady and unsteady problem

1.3. Numerical Approximation

As we have seen the existence proofs of weak solutions in [BGMS09, BGMS12] are constructive, however, we cannot lay our hands on the finite-dimensional spaces used for the Galerkin approximation to approximate weak solutions computationally. For this reason we aim to set up an approximate problem using finite element function spaces in Ω and a time-stepping in $(0, T)$. For a discretisation by means of finite element approximation polytopal domains are the straightforward choice, because this is compatible with triangulation. We choose homogeneous Dirichlet boundary conditions, since integrability of the pressure is out of reach anyway.

If higher regularity is available, then error estimates provide convergence of a numerical scheme with a certain rate. As outlined such regularity results exist for explicit constitutive relations under suitable conditions on the degeneracy, the boundary conditions and the range of q at least for short times. The first result concerning error estimates is [PR01b] and subsequently, based on improved regularity results, a series of contributions on error estimates for strong solutions was published, improving the range of q and the convergence rates: [DPR02, DPR06, BDR09, BDR15]. The most recent result holds for $d = 3$, $q \in \left(\frac{3}{2}, 2\right]$ and shows optimal convergence rates in both space and time, based on the regularity result in [BDR10] in the periodic case. This was achieved using the natural quasi-norms suggested by the constitutive relation, implicit Euler steps for the discretisation in time, and semi-implicit approximation of the convective term. Note also that in the case of homogeneous Dirichlet boundary conditions at present only the Stokes problem can be handled, see [ER18].

1.3.1. Convergence to Weak Solutions for Implicit Laws

Since higher regularity results are not available in the more general framework of implicitly constituted fluids, convergence by means of error estimates is out of reach. The most one can hope for is a convergence result of qualitative nature, i.e., convergence up to subsequences to a weak solution based on compactness arguments. How to choose the subsequences or at which rate the convergence happens cannot be answered in this context.

1.3.1.1 Finite Element Approximation

The standard inf-sup stable pairs of mixed finite element spaces $(\mathbb{V}^n, \mathbb{Q}^n)_{n\in\mathbb{N}}$ for the velocity and the pressure give rise to spaces $\mathbb{V}^n_{\mathrm{div}}$ of *discretely divergence-free* finite element functions. Discretely divergence-free means that the divergence is zero in the dual of the pressure space $\mathbb{Q}^n$, but in general not in a pointwise sense. Skew-symmetry of the weak form of the convective term, however, is in general only given for (exactly) divergence-free functions though. Hence, the usual approach (see [Tem84]) consists in introducing the numerical convective term

$$\frac{1}{2}\left(\int_M (\boldsymbol{u}\otimes\boldsymbol{w}) : \nabla\boldsymbol{u}\,\mathrm{d}\boldsymbol{z} - \int_M (\boldsymbol{u}\otimes\boldsymbol{u}) : \nabla\boldsymbol{w}\,\mathrm{d}\boldsymbol{z}\right) \tag{1.18}$$

for suitable test functions $\boldsymbol{w}$. Note that the first term is added to maintain skew-symmetry. This extra term causes an additional restriction of q, since one can only use the information that $\boldsymbol{w}$ is a bounded function rather than a Lipschitz function. Consequently, for the convergence proof one requires that $q > \frac{2d}{d+1}$ in the steady case and that $q > \frac{2(d+1)}{d+2}$ in the unsteady case.

This modification of the convective term and the resulting extra restriction of q are not needed when working with the less standard finite element spaces, for which $\mathbb{V}^n_{\mathrm{div}}$ consists of exactly (pointwise) divergence-free functions. To ensure this conformity with respect to the divergence, those finite element spaces tend to have higher local dimension than the ones for which $\mathbb{V}^n_{\mathrm{div}}$ is only discretely divergence-free. Because of their higher complexity and more recent introduction exactly divergence-free finite element spaces are less well investigated both in theory and for practical computations. Aside from the restrictions on q using exactly divergence-free spaces has the advantage that they possess better pressure-robustness properties than the standard spaces, see [JLM+17, LR18] for details.

Survey on the Convergence Proof in the Steady Case

To the best of our knowledge the only convergence results available for implicit constituted fluids are those contained in [DKS13a, KS16] for the steady case.

In [DKS13a] the authors prove the convergence (up to subsequences) 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 the larger range $q > \frac{2d}{d+2}$ for exactly divergence-free finite element functions for the velocity field. The convercence proof builds on the existence results in [BGMS09] and uses Young measures and a similar set of weak compactness results. Since no regularising term is used admissibility is lost in the discretisation limit if $q < \frac{3d}{d+2}$. Thus, in order to identify the implicit relation a discrete divergence-corrected Lipschitz approximation is developed and applied.

In [KS16] an a posteriori analysis is performed for implicitly constituted fluid flow models using discretely divergence-free finite element functions, for which the finite element spaces are nested. Furthermore the authors assume that the graph approximation has slightly better properties than the mollification based approximation possesses and illustrate the construction for some examples. The use of Young measures can be avoided using a characterisation for a couple of matrices to be in $\mathcal{A}(\cdot)$, contained in [BGMS12].

1.3.1.2 Time Discretisation

To obtain a fully discrete approximation the time derivative $\partial_t\boldsymbol{u}$ in the equation is replaced by a difference quotient d_t with respect to a time grid. Then the approximate problem involves solving a steady problem on each time level. Because of its stability properties, we choose a backward Euler method.

In the context of error extimates for explicit constitutive relations it has proven useful to consider a semi-implicit approximation of the numerical convective, see [DPR06, BDR09, BDR15], because it has better stability properties than the fully implicit approximate problem. When proving convergence up to subsequences for implicit constitutive laws the advantage of a semi-implicit approximation lies in the fact that the approximate solutions are unique. In the convergence proof not much changes when assuming slightly more either on the initial data or on the finite element setting.

Additional difficulties arise from the application of the Aubin–Lions or the Simon compactness lemma. The Aubin–Lions lemma requires uniform estimates on the sequence of time derivatives, which means that if applied in the discretisation limit one requires stability properties of an L^2-projection P^n_{div} mapping to $\mathbb{V}^n_{\mathrm{div}}$. More precisely, if the linear operator $P^n_{\mathrm{div}}\colon \mathrm{Y} \to \mathrm{X}$ is continuous for a suitably chosen Sobolev space X and a Banach space Y, one can use the approximate equation to obtain

$$\int_\Omega \partial_t \boldsymbol{u}^n(t,\cdot)\cdot \boldsymbol{w}\,\mathrm{d}\boldsymbol{x} = \int_\Omega \partial_t \boldsymbol{u}^n(t,\cdot)\cdot P^n_{\mathrm{div}}\boldsymbol{w}\,\mathrm{d}\boldsymbol{x} \overset{(\mathrm{eq})}{=} \int_\Omega \boldsymbol{F}^n(t,\cdot)\cdot P^n_{\mathrm{div}}\boldsymbol{w}\,\mathrm{d}\boldsymbol{x}$$
$$\le \|\boldsymbol{F}^n(t,\cdot)\|_{\mathrm{X}'}\,\|P^n_{\mathrm{div}}\boldsymbol{w}\|_{\mathrm{X}} \le c\,\|\boldsymbol{F}^n(t,\cdot)\|_{\mathrm{X}'}\,\|\boldsymbol{w}\|_{\mathrm{Y}}\,,$$

where by $\boldsymbol{F}^n$ we denote the right-hand side of the approximate equation. Already for a L^2-projection mapping to continuous, piecewise polynomial spaces the proof of $\mathrm{W}^{1,p}$-stability requires additional restrictions on the finite element setting, see [DDW74, CT87, EJ95, Bom06]. Initially quasi-uniformity of the mesh was assumed and then stepwise relaxed to a graded mesh condition. Here $\mathbb{V}^n_{\mathrm{div}}$ is not a space of continuous piecewise polynomial functions due to the divergence condition, which is why (as in [Car07]) we cannot get around the assumption of quasi-uniformity of the mesh and the assumption that the local space of $\mathbb{V}^n$ contains polynomials of a certain order.

In this respect applying the Simon compactness lemma, which relies on uniform convergence of certain time increments instead, is favourable. But for the estimates one requires a certain amount of admissibility, restricting the range of q.

Additional difficulties arise from the time-dependence of $\mathcal{A}$ compared to [BDR15, CHP10]. To set up a fully discrete approximate scheme one has to introduce time-averaged versions of the (time-dependent) graph approximations, similar to the approach for the external force, see [Tem84]. This gap between the stress tensor and its time-averaged version has to be bridged when identifying the implicit constitutive relation.

Survey on Convergence Results in the Unsteady Case

As far as we are aware there is no convergence result available for the unsteady flow of implicitly constituted fluids. Even those contributions that are focussed on explicit laws assume additional restrictions on q.

In [Car07] a fully discrete, fully implicit approximation scheme is introduced and convergence (up to subsequences) to a weak solution is shown for explicit laws, provided that $q \ge \frac{3d+2}{d+2}$. For compactness the author uses the Aubin–Lions compactness lemma and stability of the projector P^n_{div} is obtained under the above mentioned conditions.

In [CHP10] the authors show convergence up to subsequences for a regularised explicit law satisfying a $q(\boldsymbol{x})$-growth condition, with $\boldsymbol{x} \in \Omega$. They set up a fully discrete, fully implicit approximation scheme with regularisation in the convective term and show convergence (up to a subsequence) to a weak solution of the regularised problem, for continuous $q(\boldsymbol{x}) > \frac{2d}{d+2}$. Further, they show convergence to weak solutions of the non-regularised problem by means of an L^∞-truncation, if $q > \frac{2(d+1)}{d+2}$ is constant. Thanks to the regularisation present in the discretisation limit there is still full admissibility and hence the Simon compactness lemma can be applied.

1.3.2. Discrete Truncation and Regularisation

As we have seen before, the modified convective term used for discretely divergence-free finite element functions imposes extra restrictions on the range of q, see Subsection 1.3.1.1.

In Subsection 1.3.1.2 we have discussed that the compactness results in the unsteady case require either a further restriction on q or additional assumptions on the finite element setting.

The more severe problem is, however, that in the unsteady case no *discrete Lipschitz truncation* is available and its construction can be expected to be difficult. Some of the difficulties occurring are related to the parabolic scaling likely to carry over from the truncation to restrictions on the discretisation parameters. Furthermore, the divergence constraint and the lack of regularity of the time-derivative pose a severe difficulty, especially for discretely divergence-free finite element functions. Without a discrete truncation, in the unsteady case the implicit constitutive relation can be identified only if $q \geq \frac{3d+2}{d+2}$.

For this reason in part of this work we will employ a regularisation and take the limit after the discretisation limit, similarly as in the existence proof in [BGMS12]. This separates the discretisation limit from the admissibility problem; in the final limit corresponding to the regularisation term all discrete reminiscence has disappeared such that a *continuous truncation* can be used to identify the constitutive relation. Consequently, all the additional restrictions on q are irrelevant and the whole range of existence $q > \frac{2d}{d+2}$ can be dealt with.

1.3.3. Numerical Simulations

For computations and simulations one has to choose a specific constitutive law. There are a number of contributions devoted to numerical simulations available for the special cases of Bingham and Herschel–Bulkley fluids, see [MKYJ17, MLC16, FKF13, Zha10, BE80, DGG07] for an incomplete list. Let us highlight the numerical experiments for Bingham fluids in [HMST17] comparing a range of mixed finite element approximations, motivated by implicit constitutive theory. Here we aim to stay in the general setting, for which the lack of rigorous numerical analysis results in the literature is evident.

1.4. Aim and Outline

The objective of this work is to establish a convergence result for implicitly constituted fluids for the full range $q > \frac{2d}{d+2}$, for which existence of weak solutions is estabilished in [BGMS09, BGMS12]. More precisely, the aim is to show weak convergence (up to subsequences) of the sequence of approximate solutions to a weak solution of the problem. The main challenges arise from the implicit constitutive relation and the small exponent q in the coercivity and boundedness assumption on the implicit law in conjunction with the presence of the convective term. In the unsteady case no such result is available for the general framework of implicit constitutive relations and this is the gap we aim to fill. As a byproduct in the steady case we will be able to extend the range of q and slightly generalise the assumptions in [DKS13a].

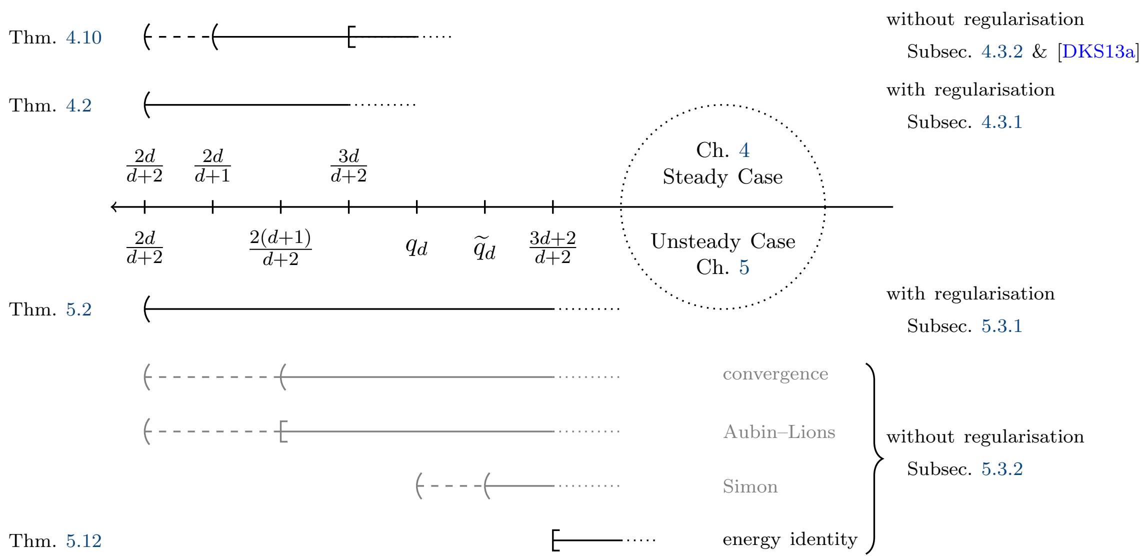

Fig. 1.6: Overview of the range of q, for which the convergence results hold; the dashed line represents the special case of exactly divergence-free finite element functions.

1.4.1. Results and Summary

We adapt the framework for implicitly constituted fluids in a thermodynamically consistent way and so that it enables us to use non-divergence-free functions for the approximation. This can be done adding a trace condition to the assumptions on $\mathcal{A} \subset \mathbb{R}^{d\times d}_{\text{sym}} \times \mathbb{R}^{d\times d}_{\text{sym}}$. Starting from this we investigate various kinds of graph approximations. In particular we verify a set of assumptions sufficient to take the graph approximation limit before the discretisation limit and identify the implicit constitutive relation. Further, we find that the generalised Yosida graph approximation additionally satisfies a Minty type convergence lemma which allows us to identify the implicit constitutive relation when taking the graph approximation simultaneously with the discretisation limit.

Let us give a summary of the convergence results we obtain for the steady and the unsteady problem **(PS)** and **(PU)**, respectively, subject to homogeneous Dirichlet boundary conditions on a bounded polytopal domain, see Fig. 1.6:

In the *steady case* we revisit the *non-regularised case*, see Theorem 4.10, where the discretisation and the graph approximation limit can be taken together thanks to the generalised Yosida graph approximation and the Minty type convergence result. This allows us to cover the same range of q as in [DKS13a], namely for $q > \frac{2d}{d+1}$ when using discretely divergence-free finite element functions and $q > \frac{2d}{d+2}$ when using exactly divergence-free finite element spaces, but under slightly weaker assumptions and with a simpler proof. In particular we do not assume (generalised) strict monotonicity, but only monotonicity and the proof does not require any Young measure tools. For $q \geq \frac{3d}{d+2}$ no truncation is required and (global) $\mathrm{W}^{1,p}$-stability of a divergence-preserving projection Π^n suffices to take the limit. For $q < \frac{3d}{d+2}$ we rely on the discrete truncation developed in [DKS13a], which requires local $\mathrm{W}^{1,1}$-stability of the projection operator Π^n.

Additionally, in Theorem 4.2 we consider the *regularised case for the steady problem*, where we first take the graph approximation limit, then the discretisation limit and finally the regularisation limit, separately and successively. This can be done for a larger class of graph approximations, it only requires a continuous truncation and thus (global) $\mathrm{W}^{1,p}$-stability of Π^n suffices. The regularisation term allows us to obtain convergence for $q > \frac{2d}{d+2}$ also for discretely divergence-free finite element functions.

The main contribution consists in the convergence proof in the *unsteady case*, for which we set up a fully discrete (pressure-free) approximate problem using an implicit Euler time-stepping and treating the (numerical) convective term fully implicitly.

Theorem 5.2 states convergence of approximate solutions for the *regularised case* and the whole range $q > \frac{2d}{d+2}$, where as in the steady case the limits are taken successively and the assumptions on the finite element setting are the same as before. Here we can take the limit in both finite element and space discretisation simultaneously without restriction of the discretisation parameters. Thanks to the regularisation present in the discretisation limit we have sufficient admissibility to apply the Simon compactness lemma, which does not require further assumptions on the finite element setting. Furthermore, there is an energy identity available for the regularised problem for almost all times which allows us to identify the implicit constitutive relation by means of a Minty type convergence result. In the last limit a (continuous) solenoidal parabolic Lipschitz approximation is applied. Let us remark that this is also an existence proof for polytopal domains and homogeneous Dirichlet boundary conditions.

Finally, in Theorem 5.12 we obtain convergence for the non-regularised case, provided that $q \geq \frac{3d+2}{d+2}$, i.e., for the admissible range. As in the steady case we employ the Yosida approximation in order to take the limits simultaneously. For convergence, the additional restriction $q > \frac{2(d+1)}{d+2}$ arises from the numerical convective term when using discretely divergence-free finite element functions. If $q > \widetilde{q}_d$ (and $q > q_d$ if the space $\mathbb{V}^n_{\mathrm{div}}$ consists of exactly divergence-free finite element functions) there is sufficient admissibility available that the Simon lemma provides compactness, see Fig. 1.5 for the values of $\widetilde{q}_d$ and q_d. Otherwise the Aubin–Lions lemma is applied, which requires additional assumptions on the finite element setting, namely quasi-uniformity of the mesh and in some cases that the local velocity space contains all quadratic functions. The bottleneck for the range of q is the identification of the implicit constitutive relation, which works if $q \geq \frac{3d+2}{d+2}$ since no discrete truncation is available and hence an energy identity is required to apply the Minty type convergence lemma.

Under additional assumptions on the initial data or assuming quasi-uniformity of the mesh and a restrictions on the discretisation parameters one can set up the corresponding semi-implicit schemes and show convergence of weak solutions, see Remark. 5.16.

1.4.2. Outline

Let us give a brief overview of the structure of this work.

Chapter 2: We introduce the preliminary results for both the analysis and the numerical approximation of the problem. These include results on time-dependent function spaces, the relevant Lipschitz approximation and weak compactness results. Furthermore, we introduce the finite element setup in a rather general manner as well as the setup for the discretisation in time. Finally we show a stability result for the L^2-projection mapping to the space of (discretely) divergence-free finite element functions.

Chapter 3: We formulate the precise assumptions for the graph encoding the implicit constitutive relation and investigate the properties of a range of graph approximations.

Chapter 4: This chapter contains the convergence proofs in the steady case and is intended to showcase some of the ideas used in the following chapter. After introducing the approximation levels, the regularised case is contained in Section 4.3.1 and the non-regularised case is revisited in Section 4.3.2.

Chapter 5: This chapter contains the main results regarding convergence in the unsteady case. The regularised case is treated in Section 5.3.1 and is based on the preprint [ST18]. Furthermore, in Section 5.3.2 we consider the unregularised case.

Appendix A: For reasons of readability we have moved various technical results concerning the pressure for the steady problem to the Appendix. Also we reproduce the proofs of the Bogovskiĭ corrections of the continuous and discrete Lipschitz approximation in the steady case. Furthermore, we investigate a range of examples of finite element spaces and collect local approximation properties of divergence-preserving projection operators Π^n sufficient to satisfy the assumptions in the convergence results.

1.4.3. Discussion and Outlook

Compared to the previous contributions [BGMS09, BGMS12, DKS13a, KS16] by adapting the assumptions on $\mathcal{A}$ we obtain notions of weak solutions which are automatically consistent with the thermodynamic framework for fluids. Also the Yosida approximation has proven to be preferable over the graph approximation based on mollification as previously used, because it allows us to take the graph approximation limit simultaneously with the discretisation limits without resorting to Young measures and without assuming (generalised) strict monotonicity of $\mathcal{A}$. As in [CHP10] we have seen that if there is sufficient admissibility, using the Simon compactness lemma in the discretisation limit requires less restrictive assumptions on the finite element setting than using the Aubin–Lions lemma. Using a regularisation term as was done in the existence proofs in [BGMS09, BGMS12] allows us to recover convergence for the full range of existence $q > \frac{2d}{d+2}$ and to avoid the additional restriction to $q > \frac{2d}{d+1}$ in the steady case and $q > \frac{2(d+1)}{d+2}$ in the unsteady case when using discretely divergence-free finite element functions. Furthermore, it separates the discretisation limit from the limit in which admissibility is lost, which means that a continuous Lipschitz approximation method suffices. This is especially important for the unsteady case, for which no discrete truncation is available. Let us remark that the regularisation term can be viewed as numerical stabilisation.

In the steady case our results extend the one in [DKS13a] to $q \in \left(\frac{2d}{d+2}, \frac{2d}{d+1}\right]$ when using discretely divergence-free finite element functions. Furthermore, the assumptions on $\mathcal{A}$ are weaker and the convergence proof is simpler. Overall one has the following options: if $q \geq \frac{3d}{d+2}$, neither regularisation nor truncation is required. If $q \in \left(\frac{2d}{d+1}, \frac{3d}{d+2}\right)$ one can use both discretely and exactly divergence-free finite element functions using a discrete truncation. If $q \in \left(\frac{2d}{d+2}, \frac{2d}{d+1}\right)$ one has to choose between using the less investigated exactly divergence-free finite element functions or using a regularisation term.

The result for the unsteady case is new in the context of implicitly constituted fluids. The regularised version is a generalisation of the result in [CHP10] both in the range of q from $q > \frac{2(d+1)}{d+2}$ to $q > \frac{2d}{d+2}$ as well as from explicit laws to implicit laws. The unregularised version is a generalisation of the result in [Car07] from explicit laws to implicit laws. In this case, clearly a discrete truncation is missing, which would allow us to obtain a result corresponding to the steady one. For this reason we currently do not enter the range of q, for which there is a difference between discretely and exactly divergence-free finite element spaces. Hence, if $q \geq \frac{3d+2}{d+2}$, one does not require a regularisation, whereas for the range $q \in \left(\frac{2d}{d+2}, \frac{3d+2}{d+2}\right)$ a regularisation approach is currently the only way to obtain convergence.

In general, whenever one wants more than a qualitative convergence result one should pass to a special class of constitutive relations providing more structure and the same is true for numerical computations.

The results presented could be improved by a discrete truncation in the unsteady case, even though it is likely to impose further restrictions. As a first step one could construct a truncation in time only for exactly divergence-free finite element spaces and $q > q_d$. For a fully discrete Lipschitz approximation it may be beneficial to assume periodic boundary conditions such that the pressure can be reconstructed globally and the time derivative is a distribution in space, in order to avoid problems concerning solenoidality. As soon as, by truncation, one actually enters the range of q for which only the Aubin–Lions compactness lemma is available, it might be worth weakening the assumptions for the stability result of the L^2-projection mapping to $\mathbb{V}^n_{\text{div}}$. With respect to numerical simulations it would be interesting to see how the different types of graph approximation perform and what influence the regularisation term has, for example for the prototypical example of Herschel–Bulkley fluids.

Extending the model one could also include additional dependencies (and equations) such as the temperature into the equations, see [MŽ18] for an existence result in this direction. Furthermore, it may be interesting to consider implicit boundary conditions of stick-slip type for convergence of numerical approximations, see [BM16]. Note however, that one has to deal with $C^{1,1}$ boundaries in order to obtain an integrable pressure.

Using a Minty type convergence lemma the methods used in this work crucially depend on monotonicity. If one drops the assumption of monotonicity, the problem becomes much harder, and to date very few results are available in this direction, see for example [CR04] for nonlinear wave equations. Recently in [JMPT18] some numerical experiments have been performed on a subclass of non-monotone (implicit) relations (and existence of approximate solutions was shown), but existence of weak solutions is not available to date. Overall, non-monotone constitutive relations represent a wide field of open problems.

CHAPTER 2

Preliminaries

In the present chapter we want to introduce and collect the notions, assumptions and results related to the analysis and the numerical approximation of the problems presented in Chapter 1. In Chapter 4 and 5 these results will be applied in the convergence proofs.

Section 2.1 includes the analysis framework and results. More specifically, here we introduce the notions of weak solutions, we collect results in time-dependent function spaces, on the Lipschitz truncation method and on weak compactness in L^1.

In the second half of the chapter we present the setting for the numerical approximation in Section 2.2. This includes the finite element setting, estimates on the numerical convective term, the time discretisation. Furthermore, we recall and present discrete tools including the discrete Lipschitz approximation method in the steady case and the L^2-projection mapping to the space of discretely divergence-free finite element functions.

All results related to the graph $\mathcal{A}$ and its approximation are contained in Chapter 3. Since the steady problem is not our main focus we refer to the appendix for results related to the pressure in the steady case. Furthermore, some details on examples of finite element spaces and constructions of certain projection operators will be postponed to the appendix. Note that except of in cases, where the arguments are hidden in a larger proof, we will not present the proofs of the results available in the literature.

2.1. Analysis

Here we want to recap the definitions and the results for the analysis of fluid flows. First we set the notation in particular for the function spaces involved. Then, in Subsection 2.1.1 we give the definition of a weak solution in the steady case and the unsteady case. Subsection 2.1.2 contains results on time-dependent functions concerning interpolation, compactness and continuity. In Subsection 2.1.3 we collect the Lipschitz approximation results used for the identification of the constitutive relation. This includes the divergence corrected version in the steady case and the solenoidal version in the unsteady case. Finally, in Subsection 2.1.4 we recall and explain the weak compactness results in L^1 used to identify the implicit constitutive relation.

Notation and Function Spaces

For the rest of the thesis we adopt the following notation, see also the list of notation.

By $\mathbb{N}_0$ we mean $\mathbb{N}\cup\{0\}$. Recall that by $\mathbb{R}^{d\times d}_{\mathrm{sym}}$ we denote the set of all real-valued symmetric $d\times d$-matrices, and that $\mathbb{R}^{d\times d}_{\mathrm{sym},0}$ is the subspace of trace-free matrices in $\mathbb{R}^{d\times d}_{\mathrm{sym}}$. We use the notation $:$ for the Frobenius scalar product in $\mathbb{R}^{d\times d}$ and with $|\cdot|$ the induced Euclidean norm on $\mathbb{R}^{d\times d}$. For a matrix $\mathbf{A} \in \mathbb{R}^{d\times d}_{\text{sym}}$ we denote by $\mathbf{A}_\delta = \mathbf{A} - \frac{1}{d}\operatorname{tr}(\mathbf{A})\mathbf{I} \in \mathbb{R}^{d\times d}_{\text{sym},0}$ its deviatoric part, where tr denotes the trace of a matrix and $\mathbf{I} \in \mathbb{R}^{d\times d}_{\text{sym}}$ is the identity matrix. By $\boldsymbol{e}_i\mathbb{R}^d$ we denote the ith unit vector, for $i \in \{1,\dots,d\}$ and by conv the convex hull. In the following we will denote vectors by bold letters and matrices by bold serifless letters.

Furthermore, the tensor product is denoted by $\otimes$. For vectors $\boldsymbol{a},\boldsymbol{b} \in \mathbb{R}^d$ this means that $(\boldsymbol{a}\otimes\boldsymbol{b})_{i,j} := a_i b_j$, for $i,j \in \{1,\dots,d\}$. Note that in the context of measurable spaces we will also denote the product σ-algebra on the product space by $\otimes$.

For a Lebesgue measurable set $\omega \subset \mathbb{R}^d$ we denote by $|\omega|$ the d-dimensional Lebesgue measure of ω. By $\mathbb{1}_\omega$ we denote the characteristic function of the set ω, $\overline{\omega}$ is the closure and $\operatorname{int}(\omega)$ the interior of ω.

In the following, $c, C > 0$ will denote generic constants, which can change from line to line depending only on the given data unless specified otherwise.

For spaces of vector-valued and tensor-valued functions we use superscripts d and $d\times d$, respectively (except for in norms).

For an open set $\omega \subset \mathbb{R}^d$ 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 and denote by $|\cdot|_{\mathrm{W}^{s,p}(\Omega)}$ the corresponding semi-norm. By $\mathrm{L}^p_0(\omega)$ we denote the set of functions in $\mathrm{L}^p(\omega)$ with zero mean integral. The space $\mathrm{L}^1_{\text{loc}}(\omega)$ denotes the space of functions, which are defined a.e. in ω and contained in $\mathrm{L}^1(O)$ for all open $O \subset\subset \omega$. Furthermore, if $\omega \subset \mathbb{R}^d$ is measurable and $0 < |\omega| < \infty$, then we denote $\fint_\omega f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} := \frac{1}{|\omega|}\int_\omega f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$.

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 $f\cdot g \in \mathrm{L}^1(\omega)$.

For the following let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain, with $d \in \{2,3\}$, and let $C(\overline{\Omega})$ be the set of all continuous real-valued functions on $\overline{\Omega}$. Denote by $\mathrm{C}^\infty_0(\Omega)$ the set of all smooth and compactly supported functions on Ω and by $\mathrm{C}^\infty_{0,\text{div}}(\Omega)^d$ the set of all functions in $\mathrm{C}^\infty_0(\Omega)^d$ with vanishing divergence. 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 $\mathrm{W}^{1,\infty}_0(\Omega) := \mathrm{W}^{1,1}_0(\Omega) \cap \mathrm{W}^{1,\infty}(\Omega)$, which coincides with the subspace of functions in $\mathrm{W}^{1,p}_0(\Omega)$ with zero boundary trace, since $\partial\Omega$ is Lipschitz. For $p \in [1,\infty]$ we let the Hölder exponent $p' \in [1,\infty]$ be defined by $\frac{1}{p} + \frac{1}{p'} = 1$, if $p \in (1,\infty)$ and with slight misuse of notation we set $1' = \infty$ and $\infty' = 1$. 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 $\mathrm{W}^{1,p}_0(\Omega)$. For $p \in (1,d)$ let $p^* = \frac{dq}{d-q}$ be the critical Sobolev exponent, such that the embedding $\mathrm{W}^{1,p}(\Omega) \hookrightarrow \mathrm{L}^{p^*}(\Omega)$ is continuous.

Further, we define the spaces of divergence-free functions: The spaces $\mathrm{L}^2_{\text{div}}(\Omega)^d$ and $\mathrm{W}^{1,p}_{0,\text{div}}(\Omega)^d$, for $p \in [1,\infty)$, are the closures of $\mathrm{C}^\infty_{0,\text{div}}(\Omega)^d$ with respect to the norms $\|\cdot\|_{\mathrm{L}^2(\Omega)}$ and $\|\cdot\|_{\mathrm{W}^{1,p}(\Omega)}$, respectively, and $\mathrm{W}^{1,\infty}_{0,\text{div}}(\Omega)^d := \mathrm{W}^{1,1}_{0,\text{div}}(\Omega)^d \cap \mathrm{W}^{1,\infty}(\Omega)^d$.

Let $T \in (0,\infty)$ and $Q = (0,T)\times\Omega$ the parabolic cylinder. Analogously as in the steady case we define $\mathrm{C}^\infty_0(Q)$ and $\mathrm{C}^\infty_{0,\text{div}}(Q)^d$. With $\mathrm{C}([0,T];\mathrm{X})$ we denote the space 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 $[0,T]$ taking values in a Banach space X. Furthermore, we define the space of weakly continuous functions with values in X by

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

We use the notation $\operatorname{ess\,lim}_{t\to 0_+} f(t)$ 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 $t \to 0_+$. We denote by $L^p(0,T;X)$ the standard Bochner space of p-integrable X-valued functions. If X is separable and reflexive and $p \in (1,\infty)$, then $L^p(0,T;X)$ is separable and reflexive with $(L^p(0,T;X))' = L^{p'}(0,T;X')$ (see[Cen92]).

For the (distributional) partial derivatives with respect to time we use the shorthand notation $\partial_t f := \frac{\partial f}{\partial t} \in \mathcal{D}'(0,T;X)$ for $f \in L^1(0,T;X)$ and X a (reflexive) Banach space.

For simplicity we will refer to the sequential version of the Banach–Alaoglu theorem as Banach–Alaoglu theorem. Also note that in general we do not relabel subsequences, and we will mention this explicitly only in theorems.

2.1.1. Weak Solutions

For the following, let $q \in (1,\infty)$ be given and assume that $\mathcal{A}(\cdot) \subset \mathbb{R}^{d\times d}_{\text{sym}} \times \mathbb{R}^{d\times d}_{\text{sym}}$ is a maximal monotone graph as introduced in Chapter 1 in Subsection 1.1.3 with respect to q a.e. in Ω in the steady case and a.e. in Q in the unsteady case. The more detailed assumptions on $\mathcal{A}(\cdot)$ will be given in in Chapter 3, Assumption 3.11. The weak solution we aim for satisfy the implicit constitutive relation encoded by $\mathcal{A}$ pointwise a.e. in the domain.

Let us present the notion of weak (distributional) solutions to the steady problem **(PS)** (see Section 1.2) for which existence was proved in [BGMS09].

Definition 2.1 (Weak Solution to the Steady Problem)**.**
*For a given $\boldsymbol{f} \in W^{-1,q'}(\Omega)^d$ we call $(\boldsymbol{u}, \mathbf{S}, \pi,)$ a weak solution to problem **(PS)**, if*

$$\boldsymbol{u} \in W^{1,q}_{0,\text{div}}(\Omega)^d, \quad \mathbf{S} \in L^{q'}(\Omega)^{d\times d}, \quad \pi \in L^1_0(\Omega),$$

and

$$-\langle \boldsymbol{u}\otimes\boldsymbol{u}, \mathbf{D}\boldsymbol{v}\rangle_\Omega + \langle \mathbf{S}, \mathbf{D}\boldsymbol{v}\rangle_\Omega - \langle \operatorname{div}\boldsymbol{v}, \pi\rangle_\Omega = \langle \boldsymbol{f}, \boldsymbol{v}\rangle_\Omega \quad \text{for all } \boldsymbol{v} \in C_0^\infty(\Omega)^d, \tag{2.1}$$

$$(\mathbf{D}\boldsymbol{u}(\boldsymbol{x}), \mathbf{S}(\boldsymbol{x})) \in \mathcal{A}(\boldsymbol{x}) \quad \text{for a.e. } \boldsymbol{x} \in \Omega. \tag{2.2}$$

Recall, that the convective term is well-defined, provided that $q \geq \frac{2d}{d+2}$, since this ensures that the embedding $W^{1,q}(\Omega) \hookrightarrow L^2(\Omega)$ is continuous. Note that the pressure can be reconstructed in a smaller Lebesgue space with exponent depending on q and d, once $\boldsymbol{u}$ and $\mathbf{S}$ are known, see also Definition 2.1.

For the unsteady problem **(PU)** (see Subsection 1.2, let us formulate the notion of a distributional solution subject to homogeneous Dirichlet boundary conditions. In contrast, the existence result in [BGMS12] deals with Navier's slip boundary condition and it is assumed that $\partial\Omega \in C^{1,1}$.

Definition 2.2 (Weak Solution to the Unsteady Problem)**.**
*For a given $\boldsymbol{u}_0 \in L^2_{\text{div}}(\Omega)^d$ and $\boldsymbol{f} \in L^{q'}(0,T;W^{-1,q'}(\Omega)^d)$ we call $(\boldsymbol{u}, \mathbf{S})$ a weak solution to problem **(PU)**, if*

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

and

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

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

Note that the convective term is well-defined since $\boldsymbol{u} \in \mathrm{L}^\infty(0,T;\mathrm{L}^2(\Omega)^d)$. A stronger definition will be given in Definition 5.1. As we have discussed in Chapter 1, Subsection 1.2.6 the pressure can be obtained as locally integrable function only due to the boundary condition and the boundary regularity. Hence we choose the pressure-free formulation. In fact, existence of weak solution could be shown for $\boldsymbol{f} \in \mathrm{L}^{q'}(0,T;\mathrm{W}^{1,q}_{0,\mathrm{div}}(\Omega)^d)')$, since the pressure is neglected anyway. But then approximation in spaces of discretely divergence-free finite element functions is not possible.

2.1.2. Results on Time-dependent Function Spaces

Here we want to present some refined interpolation inequalities, recall the Simon and the Aubin–Lions–Simon compactness lemma and collect some standard results on continuity in time of Bochner functions.

Interpolation Inequalities

We present some interpolation results, which will be applied to the velocity function $\boldsymbol{u} \in \mathrm{L}^\infty(0,T;\mathrm{L}^2(\Omega)^d) \cap \mathrm{L}^q(0,T;\mathrm{W}^{1,q}_0(\Omega)^d)$ in order to obtain estimates for the (numerical) convective term for the unsteady problem. The proof is based on the following multiplicative embedding inequality.

Lemma 2.3 (Gagliardo–Nirenberg Embedding Inequality, [DiB93, Thm. I.2.1]). *For given $p, q \in [1,\infty)$ there exists a constant $c = c(d,p,q) > 0$ such that*

$$\|v\|_{\mathrm{L}^s(\Omega)} \le c\,\|\nabla v\|^\gamma_{\mathrm{L}^q(\Omega)}\,\|v\|^{1-\gamma}_{\mathrm{L}^p(\Omega)} \tag{2.5}$$

for all $v \in \mathrm{W}^{1,q}_0(\Omega)\cap\mathrm{L}^s(\Omega)$, whenever $s\in[1,\infty)$ and $\gamma\in[0,1]$ satisfy

$$\gamma = \frac{\frac{1}{p}-\frac{1}{s}}{\frac{1}{d}-\frac{1}{q}+\frac{1}{p}}. \tag{2.6}$$

As stated in [DiB93, Thm. 2.1], for $d \ge 2$ the admissible range for $p,q,s\in[1,\infty)$ and $\gamma\in[0,1]$ satisfying (2.6) is given by:

$$\text{if } q\in[1,d): \quad \gamma\in[0,1], \quad \text{and} \quad s\in\begin{cases}\left[p,\frac{dq}{d-q}\right] & \text{if } p\in\left[1,\frac{dq}{d-q}\right],\\ \left[\frac{dq}{d-q},p\right] & \text{if } p\in\left[\frac{dq}{d-q},\infty\right);\end{cases} \tag{2.7a}$$

$$\text{if } q\in[d,\infty): \quad s\in[p,\infty), \quad \text{and} \quad \gamma\in\left[0,\frac{dq}{dq+p(q-d)}\right). \tag{2.7b}$$

Slightly more general than in [DiB93, Ch. I.3] but with analogous arguments we can prove the following interpolation result.

Lemma 2.4 (Parabolic Interpolation)**.**
Let $p, q, s \in [1, \infty)$ and $\gamma \in [0,1]$ such that (2.6) is satisfied and assume that $s \in [p, \infty)$. Then, there exists a constant $c > 0$ such that

$$\|v\|_{\mathrm{L}^r(0,T;\mathrm{L}^s(\Omega))} \leq c \|\nabla v\|^{\gamma}_{\mathrm{L}^q(Q)} \|v\|^{1-\gamma}_{\mathrm{L}^\infty(0,T;\mathrm{L}^p(\Omega))}, \tag{2.8}$$

for all $v \in \mathrm{L}^q(0,T;\mathrm{W}^{1,q}_0(\Omega)) \cap \mathrm{L}^\infty(0,T;\mathrm{L}^p(\Omega))$, with

$$r := \frac{s\,(q(p+d) - dp)}{(s-p)d} \in (1, \infty]. \tag{2.9}$$

Proof. By the assumption on p, q, s and γ we can apply the Gagliardo–Nirenberg inequality, in Lemma 2.3 and obtain that

$$\|v(t,\cdot)\|_{\mathrm{L}^s(\Omega)} \leq c \|\nabla v(t,\cdot)\|^{\gamma}_{\mathrm{L}^q(\Omega)} \|v(t,\cdot)\|^{1-\gamma}_{\mathrm{L}^p(\Omega)}, \tag{2.10}$$

for a.e $t \in (0,T)$, with

$$\gamma = \frac{(s-p)dq}{s\,(q(p+d) - dp)}.$$

If $s > p$, then we have that $0 < \gamma \leq 1 < q$, and hence for r as defined in (2.9) we obtain that $r = \frac{q}{\gamma} \in (1, \infty)$. Raising the inequality (2.10) to the power r, integrating in $(0,T)$, pulling out the L^∞-norm of the second factor of the integrand and taking the rth root shows the claim.

If $s = p$, then the claim follows trivially with $r = \infty$ and $\gamma = 0$ using Hölder's inequality in Ω. □

The special case $r = s = \frac{q(d+p)}{d}$ is proven in [DiB93, Prop. 3.1] and the case $p = q$ is shown in [DiB93, Prop. 3.3].

Corollary 2.5.
If $q \geq \frac{2d}{d+2}$, then there exists a constant $c(q) > 0$ such that

$$\|v\|_{\mathrm{L}^{\frac{q(d+2)}{d}}(Q)} \leq c \|\nabla v\|^{\gamma}_{\mathrm{L}^q(Q)} \|v\|^{1-\gamma}_{\mathrm{L}^\infty(0,T;\mathrm{L}^2(\Omega))}, \tag{2.11}$$

for all $v \in \mathrm{L}^q(0,T;\mathrm{W}^{1,q}_0(\Omega)) \cap \mathrm{L}^\infty(0,T;\mathrm{L}^2(\Omega))$, with $\gamma = \frac{d}{d+2}$.

Proof. We apply Lemma 2.4 with $p = 2$ and with $r = s = \frac{q(d+2)}{d}$, noting that one has admissibility by (2.7), if $q \geq \frac{2d}{d+2}$. Indeed, this is obvious if $q \in [d, \infty)$. For $q \in [1, d)$ one can see that $s = \frac{q(d+2)}{d} \in [2, \frac{dq}{d-q}]$ if and only if $q \geq \frac{2d}{d+2}$. □

This means that for a function $\boldsymbol{u}$ as in Definition 2.2 we have that $\boldsymbol{u} \otimes \boldsymbol{u} \in \mathrm{L}^{\frac{q(d+2)}{2d}}(Q)^{d \times d} \hookrightarrow \mathrm{L}^1(Q)^{d\times d}$, provided that $q \geq \frac{2d}{d+2}$. Also one can see that $\boldsymbol{u}$ is an admissible test function in the convective term in (2.3), if $q \geq \frac{3d+2}{d+2}$.

Corollary 2.6.
If $q \geq \frac{3d}{d+2} > 1$, then there exists a constant $c(q) > 0$ such that

$$\|v\|_{\mathrm{L}^\nu(0,T;\mathrm{L}^{2q'}(\Omega))} \leq c \|\nabla v\|^{\gamma}_{\mathrm{L}^q(Q)} \|v\|^{1-\gamma}_{\mathrm{L}^\infty(0,T;\mathrm{L}^2(\Omega))}, \tag{2.12}$$

for all $v \in \mathrm{L}^q(0,T;\mathrm{W}_0^{1,q}(\Omega)) \cap \mathrm{L}^\infty(0,T;\mathrm{L}^2(\Omega))$, *with* $\gamma = \frac{1}{q}$ *and*

$$\nu := q\left(\frac{q(d+2)}{d} - 2\right) \in (1,\infty). \tag{2.13}$$

Proof. We apply Lemma 2.4 with $q > 1$, $p = 2$ and $s = 2q' > 2$ noting that ν corresponds to $r > 1$ in this case. If $q \in [d,\infty)$ admissibility in (2.7) is clear. If $q \in (1,d)$, then we have that $2 \leq \frac{dq}{d-q}$, which is equivalent to $q \geq \frac{2d}{d+2}$ and $s = 2q' \in [2, \frac{dq}{d-q}]$, since $q \geq \frac{3d}{d+2}$. □

Note that:

$$\nu > \begin{cases} 2, \\ q', \end{cases} \quad \text{provided that} \quad \begin{cases} q^2 - \frac{2d}{d+2}q - \frac{2d}{d+2} > 0, \\ q^2 - \frac{3d+2}{d+2}q + \frac{d}{d+2} > 0. \end{cases}$$

Computing the positive roots one can see that

$$\nu > 2 \quad \text{if } q > q_d := \frac{1}{d+2}\left(d + \sqrt{3d^2+4d}\right), \tag{2.14}$$

$$\nu > q' \quad \text{if } q > \widetilde{q}_d := \frac{1}{2(d+2)}\left(3d + 2 + \sqrt{5d^2+4d+4}\right). \tag{2.15}$$

With the values

$$q_2 = \frac{1}{2}(1+\sqrt{5}) \approx 1.62, \qquad \widetilde{q}_2 = 1 + \frac{\sqrt{2}}{2} \approx 1.71, \tag{2.16}$$

$$q_3 = \frac{1}{5}\left(3+\sqrt{39}\right) \approx 1.85, \qquad \widetilde{q}_3 = \frac{1}{10}\left(11+\sqrt{61}\right) \approx 1.88, \tag{2.17}$$

it is evident that

$$\frac{3d}{d+2} < q_d < \widetilde{q}_d < \frac{3d+2}{d+2}, \qquad \text{for } d \in \{2,3\}. \tag{2.18}$$

Similarly as before, for $\boldsymbol{u}$ as in Definition 2.1 this means that $\boldsymbol{u} \otimes \boldsymbol{u} \in \mathrm{L}^{\nu/2}(0,T;\mathrm{L}^{q'}(\Omega)^{d\times d})$, which means that $\boldsymbol{u}(t,\cdot)$ is an admissible test function in Ω, if $\nu \geq 2$. The condition that $\nu \geq q'$ will be relevant for the numerical modification of the convective term, as we will see below.

Compactness

To obtain compactness we have two options: to apply the Aubin–Lions–Simon lemma for which uniform estimates on the sequence of distributional time derivatives are required; or to use the Simon lemma, which is based on uniform convergence of time increments. We will see that depending on the situation one or the other might be preferable.

There are various versions of the Aubin–Lions or Aubin–Lions–Simon lemma available, see for example[Mou16]. The following version is convenient, because we do not have to specify the exponent in time, of the Bochner space, in which the sequence of time derivatives is bounded in.

Lemma 2.7 (Aubin–Lions–Simon, [Sim87, Cor. 6, p. 87])**.**
Let $\mathrm{X}, \mathrm{B}, \mathrm{Z}$ *be Banach spaces such that the embedding* $\mathrm{X} \hookrightarrow\hookrightarrow \mathrm{B}$ *is compact and the embedding* $\mathrm{B} \hookrightarrow \mathrm{Z}$ *is continuous. Further, let* $p \in (1,\infty]$ *be given. Let*

(i) $\mathcal{F}$ *be bounded in* $\mathrm{L}^p(0,T;\mathrm{B}) \cap \mathrm{L}^1_{\mathrm{loc}}(0,T;\mathrm{X})$ *and let*

(ii) the set of distributional derivatives $\{\partial_t f : f \in \mathcal{F}\}$ *be bounded in* $\mathrm{L}^1_{\mathrm{loc}}(0,T;\mathrm{Z})$.

Then, $\mathcal{F}$ is relatively compact in $\mathrm{L}^s(0,T;\mathrm{B})$*, for any* $s \in [1,p)$.

Lemma 2.8 (Simon, [Sim87, Thm. 3, p. 80])**.**
Let X, B *be Banach spaces such that the embedding* $\mathrm{X} \hookrightarrow\hookrightarrow \mathrm{B}$ *is compact. Let* $\mathcal{F} \subset \mathrm{L}^p(0,T;\mathrm{B})$ *for some* $p \in [1,\infty)$ *and let*
(i) $\mathcal{F}$ *be bounded in* $\mathrm{L}^1_{\mathrm{loc}}(0,T;\mathrm{X})$,
(ii) $\int_0^{T-\varepsilon} \|f(s+\varepsilon,\cdot) - f(s,\cdot)\|_{\mathrm{B}}^p \,\mathrm{d}s \to 0, \quad$ *as* $\varepsilon \to 0, \quad$ *uniformly for* $f \in \mathcal{F}$.
Then, $\mathcal{F}$ *is relatively compact in* $\mathrm{L}^p(0,T;\mathrm{B})$.

We will apply these compactness results for $\mathrm{X} = \mathrm{W}^{1,q}(\Omega)^d$ and $\mathrm{B} = \mathrm{L}^2(\Omega)^d$ and for the assumption of compactness of the embedding $\mathrm{X} \hookrightarrow\hookrightarrow \mathrm{B}$ we require that $q > \frac{2d}{d+2}$.

Continuity in Time

Considering purely the functions spaces a Bochner function and its (distributional) time derivative are contained in, one can deduce certain continuity properties.

Lemma 2.9 (Continuity I)**.**
Let Z *be a reflexive Banach space. Then, any function* $v \in \mathrm{L}^1(0,T;Z')$ *with distributional derivative* $\partial_t v \in \mathrm{L}^1(0,T;Z')$ *is contained in* $\mathrm{C}_w([0,T];Z')$.

Proof. The proof follows from [Tem84, Lem. 1.1, Ch. III, § 1]. □

Lemma 2.10 (Continuity II, [Tem84, Lem. 1.4, Ch. III, § 1])**.**
Let B *and* Z *be Banach spaces and let* B *be reflexive such that the embedding* $\mathrm{B} \hookrightarrow \mathrm{Z}$ *is continuous. Then one has that*

$$\mathrm{L}^\infty(0,T;\mathrm{B}) \cap \mathrm{C}_w([0,T];\mathrm{Z}) \subset \mathrm{C}_w([0,T];\mathrm{B}).$$

Lemma 2.11 (Continuity III, [Ser15, Thm. 2.2])**.**
Let X *be a reflexive Banach space and let* B *be a Hilbert space, such that the embedding* $\mathrm{X} \hookrightarrow \mathrm{B}$ *is continuous and dense. Furthermore for* $p \in (1,\infty)$*, let* $v \in \mathrm{L}^p(0,T;\mathrm{X}) \cap \mathrm{L}^2(0,T;\mathrm{B})$ *and the distributional time derivative* $\partial_t v \in \mathrm{L}^{p'}(0,T;\mathrm{X}')$*. Then (up to a representative) one has that* $v \in \mathrm{C}([0,T];\mathrm{B})$.

We will apply these results to show that the initial conditions are attained in a way which is helpful for the identification of the implicit constitutive relation.

2.1.3. Lipschitz Approximation

For small $q \in (1,\infty)$ weak solutions of the steady and the unsteady problem according to the Definitions 2.1 and 2.2 are not an admissible test function due to the presence of the convective term. As mentioned in Chapter 1 the Lipschitz truncation method can be used to restore the admissibility which is the key to identify the implicit constitutive relation. The construction is based on a truncation at the level of the Hardy–Littlewood maximal operator of the gradients resulting in approximations that differ from the original function only on a set of small measure. In the interest of space we do not go into more depth here.

A sequential version can be applied to sequences which converge to zero weakly in a Sobolev space and the truncated (double) sequence will have better convergence properties for fixed truncation level.

In the context of fluid equations the pressure might be existing as integrable function only locally and the approximating sequence converges weakly only. Hence, it is convenient, if the sequence of truncations is divergence-free, provided that the original functions are divergence-free, since this allows us to use the pressure-free equations. Here we will see a Bogovskiĭ corrected version of the Lipschitz truncation in the steady case (see [DMS08]), where the Bogovskiĭ operator is the right-inverse of the divergence operator, see Lemma A.2. This has motivated a corresponding approach in the discrete situation, developed in [DKS13a], see Section 2.2.1.2. Futhermore, in the unsteady case we will recall a solenoidal parabolic Lipschitz approximation, developed in [BDS13].

2.1.3.1 The Divergence Corrected Lipschitz Approximation in the Steady Case

Let us first state a sequential version for Sobolev functions with zero boundary trace.

Lemma 2.12 (Lipschitz Approximation in the Steady Case, [DKS13a, Cor. 25])**.**
Let $p \in (1,\infty)$ and let $\{\boldsymbol{v}^l\}_{l\in\mathbb{N}} \subset \mathrm{W}_0^{1,p}(\Omega)^d$ be a sequence converging to zero weakly in $\mathrm{W}_0^{1,p}(\Omega)^d$, as $l \to \infty$.
Then, there exist:

- *a double sequence $\{\lambda_{l,j}\}_{l,j\in\mathbb{N}} \subset \mathbb{R}$ with $\lambda_{l,j} \in [2^{2^j}, 2^{2^{j+1}-1}]$, for any $l,j \in \mathbb{N}$,*
- *a double sequence of open sets $\mathcal{B}_{l,j} \subset \Omega$, $l,j \in \mathbb{N}$, and*
- *a double sequence of functions $\{\boldsymbol{v}^{l,j}\}_{l,j\in\mathbb{N}} \subset \mathrm{W}_0^{1,\infty}(\Omega)^d$,*

such that

(i) $\boldsymbol{v}^{l,j} = \boldsymbol{v}^l$ on $\Omega\backslash\mathcal{B}_{l,j}$, i.e., $\{\boldsymbol{v}^{l,j} \neq \boldsymbol{v}^l\} \subset \mathcal{B}_{l,j}$ for all $l,j \in \mathbb{N}$;

(ii) there exists a constant $c = c(p) > 0$ such that

$$\left\|\lambda_{l,j}\mathbb{1}_{\mathcal{B}_{l,j}}\right\|_{\mathrm{L}^p(\Omega)} \le c2^{-\frac{j}{p}} \quad \text{for all } l,j \in \mathbb{N};$$

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

$$\left\|\nabla \boldsymbol{v}^{l,j}\right\|_{\mathrm{L}^\infty(\Omega)} \le c\lambda_{l,j} \quad \text{for all } l,j \in \mathbb{N};$$

(iv) for any fixed $j \in \mathbb{N}$ we have

$$\begin{aligned} \boldsymbol{v}^{l,j} &\to \mathbf{0} \quad &&\text{strongly in } \mathrm{L}^\infty(\Omega)^d, \\ \nabla\boldsymbol{v}^{l,j} &\overset{*}{\rightharpoonup} \mathbf{0} \quad &&\text{weakly* in } \mathrm{L}^\infty(\Omega)^{d\times d}, \end{aligned}$$

as $l \to \infty$.

Here we have denoted $\mathcal{B}_{l,j} := \mathcal{O}^{l,j} \cap \Omega$, where $\mathcal{O}^{l,j}$ are the respective (global) super-level sets of the maximal function, referred to as "bad" sets in the respective references.

For divergence-free functions $\boldsymbol{v}^l$ the truncations $\boldsymbol{v}^{l,j}$ are in general only divergence-free on the set $\Omega\backslash\mathcal{B}_{l,j}$. For this reason a modification is required to obtain a divergence-free sequence of function $\boldsymbol{w}^{l,j}$ with similar convergence properties, which allows to use the pressure-free equation. With the Bogovskiĭ operator $\mathfrak{B}$ introduced in Section A.1 one considers

$$\boldsymbol{w}^{l,j} := \boldsymbol{v}^{l,j} - \mathfrak{B}(\operatorname{div}\boldsymbol{v}^{l,j}), \tag{2.19}$$

which is divergence-free by the properties of the Bogovskiĭ operator, see Lemma A.2. Further, if $\boldsymbol{v}^l$ is divergence-free, then $\boldsymbol{w}^{l,j}$ can be shown to have similarly good convergence properties as $\boldsymbol{v}^{l,j}$. The following lemma can be extracted from the proof of [DMS08, Thm. 3.1].

Lemma 2.13 (Divergence Correction, [DMS08]).
Let $p \in (1,\infty)$ *and let* $\{\boldsymbol{v}^l\}_{l\in\mathbb{N}} \subset \mathrm{W}^{1,p}_{0,\mathrm{div}}(\Omega)^d$ *be a sequence converging to zero weakly in* $\mathrm{W}^{1,p}_0(\Omega)^d$, *as* $l \to \infty$. *Furthermore, let* $\{\boldsymbol{v}^{l,j}\}_{l,j\in\mathbb{N}} \subset \mathrm{W}^{1,\infty}_0(\Omega)^d$ *be the sequence given by Lemma 2.12.*
Then, there exists a double sequence $\{\boldsymbol{w}^{l,j}\}_{l,j\in\mathbb{N}}$ *such that*

(i) $\boldsymbol{w}^{l,j} \in \mathrm{W}^{1,s}_{0,\mathrm{div}}(\Omega)^d$, *for all* $l,j \in \mathbb{N}$ *and all* $s \in [1,\infty)$;

(ii) there exists a constant $c = c(p) > 0$ *such that*

$$\left\| \boldsymbol{v}^{l,j} - \boldsymbol{w}^{l,j} \right\|_{\mathrm{W}^{1,p}(\Omega)} \le c2^{-\frac{j}{p}} \quad \text{for all } l,j \in \mathbb{N};$$

(iii) for any fixed $j \in \mathbb{N}$ *(up to subsequences) we have that*

$$\begin{aligned} \boldsymbol{w}^{l,j} &\to \mathbf{0} && \text{strongly in } \mathrm{L}^s(\Omega)^d, \\ \boldsymbol{w}^{l,j} &\rightharpoonup \mathbf{0} && \text{weakly in } \mathrm{W}^{1,s}_0(\Omega)^d, \end{aligned}$$

as $l \to \infty$, *for any* $s \in (1,\infty)$.

Proof. The proof can be found in [DMS08, Thm. 3.1], see pp. 223–224 therein, and is recalled in the Appendix A in Section A.1 for the reader's convenience. □

Note that the Bogovskiĭ operator is not a local operator; the correction changes the function on the whole of Ω, which means that $\boldsymbol{w}^{l,j}$ does not coincide with $\boldsymbol{v}^l$ on a large set anymore. However, due to the stability of the operator $\mathfrak{B}$ and because of the fact that $\boldsymbol{v}^{l,j}$ is divergence-free on a large set, the difference $\boldsymbol{v}^{l,j} - \boldsymbol{w}^{l,j}$ can be controlled.

Alternatively there are solenoidal truncations available, for which the truncation itself is divergence-free and coincides with the original function on a large set. This was achieved in [BDF12] using a local Bogovkiĭ correction and in [BDS13] by means of a construction based on the fact that $\operatorname{div}(\operatorname{curl} \boldsymbol{v})$ vanishes. The latter was actually developed for the unsteady situation and will be stated in the following subsection.

2.1.3.2 Parabolic Solenoidal Lipschitz Approximation

Truncations for evolutionary problems differ from the truncations we have seen for steady problems in that they involve an equation and the truncation also takes place on the level of the right-hand side of the equation, see for example [DRW10, BDS13, BGMS12, DSSV17]. Furthermore, they are local in nature because then the boundary conditions do not have to be met.

As we have seen in Section 1.2.6 for unsteady problems subject to homogeneous Dirichlet boundary conditions we can only expect locally integrable pressure. Furthermore, a Bogovskiĭ correction is not available in the unsteady case, since the distributional time derivative and $\mathfrak{B}$ do not behave well for the regularity available. Thus, one has two options: Either the pressure is reconstructed locally and decomposed, see [DRW10]. Or a solenoidal Lipschitz truncation is applied, which ensures solenoidality by other means than a Bogovskiĭ correction. Such a Lipschitz approximation was developed in [BDS13] and relies on the fact that $\operatorname{div}\operatorname{curl}$ vanishes. The corresponding result for the steady case is contained in [BDS13] as well. Note that the sets $\mathcal{B}_{l,j}$ in the following Lemma satisfy $\mathcal{B}_{l,j} = \mathcal{O}_{l,j} \cap Q_0$, where $\mathcal{O}_{l,j}$ are the "bad sets" in the construction in [BDS13].

Lemma 2.14 (Parabolic Solenoidal Lipschitz Approximation, [BDS13, 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

$$\left\langle \partial_t \boldsymbol{v}^l, \boldsymbol{\xi} \right\rangle_{Q_0} = \left\langle \mathbf{G}^l, \nabla\boldsymbol{\xi} \right\rangle_{Q_0} \qquad \text{for all } \boldsymbol{\xi}\in \mathrm{C}^\infty_{0,\mathrm{div}}(Q_0)^d \tag{2.20}$$

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 $\mathcal{B}_{l,j}\subset Q_0$, $l,j\in\mathbb{N}$,*
- *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* $\operatorname{supp}(\boldsymbol{v}^{l,j})\subset\frac{1}{6}Q_0$, *for any* $j\ge j_0$ *and any* $l\in\mathbb{N}$;

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

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

$$\limsup_{l\to\infty}\lambda_{l,j}^p\,|\mathcal{B}_{l,j}|\le c2^{-j} \qquad \text{for all } j\ge j_0;$$

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

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

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

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

as $l\to\infty$;

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

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

(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

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

The lemma is stated for $d=3$, but according to [BDS13, Rem. 2.1, p. 2692] the result holds for all $d\ge 2$, with minor modifications of the proof.

For the convergence proof we want to apply the solenoidal Lipschitz approximation for a regularised problem, for which the regularisation term is not in divergence form. Thus let us give the following slight modification including a lower order term, which can be proved by applying an inverse Laplacian.

Corollary 2.15 (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 $\mathrm{L}^p(I_0;\mathrm{W}^{1,p}(B_0)^d)$, 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 $\mathrm{L}^\sigma(Q_0)^{d\times d}$ and a third sequence $\{\boldsymbol{f}^l\}_{l\in\mathbb{N}}$ converging to zero strongly in $\mathrm{L}^\sigma(Q_0)^d$. Furthermore, denoting $\mathbf{G}^l := \mathbf{G}_1^l + \mathbf{G}_2^l$, assume that, for any $l \in \mathbb{N}$, the equation*

$$\left\langle \partial_t \boldsymbol{v}^l, \boldsymbol{\xi} \right\rangle_{Q_0} = \left\langle \mathbf{G}^l, \nabla \boldsymbol{\xi} \right\rangle_{Q_0} + \left\langle \boldsymbol{f}^l, \boldsymbol{\xi} \right\rangle_{Q_0} \quad \text{for all } \boldsymbol{\xi} \in \mathrm{C}^\infty_{0,\mathrm{div}}(Q_0)^d \tag{2.21}$$

is satisfied. Then, the same statement as in Lemma 2.14 holds.

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

$$\left\langle \nabla \boldsymbol{g}^l, \nabla \boldsymbol{v} \right\rangle_{B_0} = \left\langle \boldsymbol{f}^l(t,\cdot), \boldsymbol{v} \right\rangle_{B_0} \qquad \text{for all } \boldsymbol{v} \in \mathrm{C}^\infty_0(B_0)^d. \tag{2.22}$$

Standard regularity theory for Poisson's equation (see [Gri11, Thm. 2.4.2.5] and [GT01, Lem. 9.17] guarantees the existence of a unique $\boldsymbol{g}^l(t,\cdot) \in \mathrm{W}^{2,\sigma}(B_0)^d \cap \mathrm{W}^{1,\sigma}_0(B_0)^d$ solving (2.22) such that

$$\left\| \boldsymbol{g}^l(t,\cdot) \right\|_{\mathrm{W}^{2,\sigma}(B_0)} \leq c \left\| \boldsymbol{f}^l(t,\cdot) \right\|_{\mathrm{L}^\sigma(B_0)}, \tag{2.23}$$

since $\sigma \in (1,\infty)$ and ∂B_0 is smooth. Viewing $\boldsymbol{g}^l$ as a function defined on $Q_0 = I_0 \times B_0$ by (2.23) one has that

$$\left\| \boldsymbol{g}^l \right\|_{\mathrm{L}^\sigma(I_0;\mathrm{W}^{2,\sigma}(B_0))} \leq c \left\| \boldsymbol{f}^l \right\|_{\mathrm{L}^\sigma(Q_0)} \to 0, \quad \text{as } l \to \infty, \tag{2.24}$$

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 $\mathbf{G}_2^l + \nabla \boldsymbol{g}^l$ converges to zero strongly in $\mathrm{L}^\sigma(Q_0)^{d\times d}$, as $l \to \infty$. Applying (2.22) in (2.21) shows that

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

and thus all assumptions of Lemma 2.14 are satisfied and the claim follows. ☐

2.1.4. Convergence Results in L^1

The results collected in this section close the gap between applying a Lipschitz truncation and a Minty type convergence lemma for implicit constitutive relations, mentioned in Chapter 1, Subsection 1.2.3, see Lemma 3.16.

Let M be Ω or Q, respectively. Further, let $\{\boldsymbol{u}^l\}_{l\in\mathbb{N}}$ is the sequence of approximate velocity functions and $\{\mathbf{S}^l\}_{l\in\mathbb{N}}$ the sequence of approximate stress tensor functions, such that $\mathbf{D}\boldsymbol{u}^l \rightharpoonup \mathbf{D}\boldsymbol{u}$ weakly in $\mathrm{L}^q(M)^{d\times d}$ and $\mathbf{S}^l \rightharpoonup \overline{\mathbf{S}}$ weakly in $\mathrm{L}^{q'}(M)^{d\times d}$.

Application of the Lipschitz approximation typically shows that

$$(H^l)^{1/2} := \Big((\mathbf{S}^l - \mathcal{S}^\star(\mathbf{D}\boldsymbol{u})) : (\mathbf{D}\boldsymbol{u}^l - \mathbf{D}\boldsymbol{u}) \Big)^{1/2} \to 0 \quad \text{strongly in } \mathrm{L}^1(\widetilde{M}), \tag{2.25}$$

as $l \to \infty$, where $\widetilde{M}$ is a subdomain of M and $\mathcal{S}^\star$ is a selection function representing the constitutive relation. Here the root stems from using Hölder's inequality which is applied to control the size of the bad set.

If the constitutive relation is explicit (i.e., $\mathbf{S}^l = \mathcal{S}(\mathbf{D}\boldsymbol{u}^l)$ and $\mathcal{S}^\star(\mathbf{D}\boldsymbol{u}) = \mathcal{S}(\mathbf{D}\boldsymbol{u})$ for a continuous function $\mathcal{S}$) and strictly monotone, this suffices to identify the implicit constitutive relation. Indeed, the fact that $(H^l)^{1/2} \to 0$ strongly in $\mathrm{L}^1(\widetilde{M})$ implies that (up to a subsequence) $(H^l)^{1/2} \to 0$ a.e. in $\widetilde{M}$ and thus also $H^l \to 0$ a.e. in $\widetilde{M}$. By strict monotonicity and continuity of $\mathcal{S}$ one can show that this implies that $\mathbf{D}\boldsymbol{u}^l \to \mathbf{D}\boldsymbol{u}$ a.e. in $\widetilde{M}$, see [DMM98, Lem. 6]. This suffices to identify the constitutive relation.

Here, however, we consider implicit constitutive relations which are not assumed to be strictly monotone. To obtain weak convergence in L^1 one has to pay the price of passing passing to a sequence of subdomains with increasing measure. This approach was introduced in the existence proof in [BGMS12] and will be stated here since it will be used several times in the following both in the steady and the unsteady case.

Lemma 2.16 (Chacon's Biting Lemma, [BM89])**.**
Let $\omega \subset \mathbb{R}^n$ be a an open and bounded set and let $\{H^l\}_{l\in\mathbb{N}}$ be a sequence, which is bounded in $\mathrm{L}^1(\omega)$. Then, there exists an $H \in \mathrm{L}^1(\omega)$ and a non-increasing sequence of measurable sets $E_i \subset \omega$, $i \in \mathbb{N}$, such that $|E_i| \to 0$, as $i \to \infty$, and a (non-relabelled) subsequence, such that

$$H^l \rightharpoonup H \qquad \textit{weakly in } \mathrm{L}^1(\omega\backslash E_i), \quad \textit{as } l \to \infty, \tag{2.26}$$

for any fixed $i \in \mathbb{N}$.

If in addition one has convergence of a root to 0 strongly in $\mathrm{L}^1(\omega)$, this allows one to obtain strong convergence and to identify the limit with 0.

Lemma 2.17 (L^1-Convergence, [BGMS12])**.**
Let $\omega \subset \mathbb{R}^n$ be an open and bounded set and let $\{H^l\}_{l\in\mathbb{N}}$ be a sequence, which is bounded in $\mathrm{L}^1(\omega)$ such that

$$(H^l)^{1/2} \to 0 \qquad \textit{strongly in } \mathrm{L}^1(\omega), \quad \textit{as } l \to \infty.$$

Then, there exists a non-increasing sequence of measurable sets $E_i \subset \omega$, $i \in \mathbb{N}$, such that $|E_i| \to 0$, as $i \to \infty$ and (up to a subsequence) one has that

$$H^l \to 0 \qquad \textit{strongly in } \mathrm{L}^1(\omega\backslash E_i), \quad \textit{as } l \to \infty,$$

for any fixed $i \in \mathbb{N}$.

Proof. The proof can be found in [BGMS12, p. 2788] and is reproduced for the reader's convenience: The strong convergence of $(H^l)^{1/2}$ in $\mathrm{L}^1(\omega)$ implies that there exists a subsequence such that $(H^l)^{1/2} \to 0$ a.e. in ω and hence also $H^l \to 0$ a.e. in ω. Furthermore, since $\{H^l\}_{l\in\mathbb{N}}$ is bounded in $\mathrm{L}^1(\omega)$ Chacon's biting lemma stated in Lemma 2.16 shows the existence of a further subsequence, a function $H \in \mathrm{L}^1(\omega)$ and a non-increasing sequence of measurable subsets $E_i \subset \omega$, $i \in \mathbb{N}$, such that $|E_i| \to 0$, as $i \to \infty$, such that

$$H^l \rightharpoonup H \quad \text{ in } \mathrm{L}^1(\omega\backslash E_i), \qquad \text{as } l \to \infty, \qquad \text{for each fixed } i \in \mathbb{N}.$$

Now let $i \in \mathbb{N}$ be arbitrary but fixed. By the Dunford–Pettis compactness criterion (see for example [ET76, Ch. 8, Thm. 1.3]) it follows that the sequence $\{H^l\}_{l\in\mathbb{N}}$ is equi-integrable on $\omega\backslash E_i$. By the a.e. convergence of H^l to zero in particular in $\omega\backslash E_i$ the Vitali convergence theorem implies that $H^l \to 0$ in $\mathrm{L}^1(\omega\backslash E_i)$, as $l \to \infty$, for any fixed $i \in \mathbb{N}$. □

Note that the localisation effect by Chacon's biting lemma requires us to use a localised Minty type convergence lemma, even when using a Lipschitz approximation arguments on the whole domain (in the steady case).

2.2. Numerical Approximation

Since the Galerkin function spaces in the existence proof are not available for computational purposes we want to approximate the problem by a numerical approximation scheme. More specifically, we aim for a finite element approximation in Ω and in the unsteady case a time-stepping in $(0, T)$. Here we want to introduce the setting and assumptions used for the convergence proofs.

Subsection 2.2.1 contains the finite element setting and results relevant in the steady case. Then, in Subsection 2.2.2 we will focus on the convective term, its numerical modification and estimates of both. Finally, in Subsection 2.2.3 we will describe the time discretisation and discrete results relevant for the unsteady case.

2.2.1. Approximation in Space

Here we will first describe the finite element setting. This includes the assumptions the convergence proofs will be based on and examples of finite element spaces satisfying the assumptions.

Furthermore we will state the discrete divergence corrected Lipschitz approximation result obtained in [DKS13a].

2.2.1.1 Finite Element Spaces and Assumptions

The setting for the discretisation in space is similar to the one used in [DKS13a] and the assumptions imposed result in a pair of inf-sup stable conforming finite element spaces. Note that for the convergence proof we will assume subsets of the set of assumptions presented, depending on the case.

The following assumption concerning regularity of the triangulation will be used in the rest of this chapter without explicitly mentioning it.

Assumption 2.18 (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 [Cia02, 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 $\widehat{K} := \operatorname{conv}\{\mathbf{0}, \boldsymbol{e}_1, \ldots, \boldsymbol{e}_d\} \subset \mathbb{R}^d$, i.e., there exists an affine invertible function $\mathbf{F}_K : K \to \widehat{K}$;*

(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*

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

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

For $n \in \mathbb{N}$ we denote the grid size $h_n := \max\{h_K : K \in \mathcal{T}_n\}$. For $K \in \mathcal{T}_n$, $n \in \mathbb{N}$ we denote the *neighbouring patch* of elements of K by

$$\omega_n(K) := \operatorname{int}\left(\bigcup\{K' \in \mathcal{T}_n \colon K \cap K' \neq \emptyset\}\right). \tag{2.27}$$

By the regularity of $\{\mathcal{T}_n\}_{n\in\mathbb{N}}$ the number of (distinct) elements $K' \in \mathcal{T}_n$ such that $K' \subset \overline{\omega_n(K)}$ is bounded, uniformly in $K \in \mathcal{T}_n$ and in $n \in \mathbb{N}$, i.e., the overlap of the neighbourhoods is bounded. Furthermore, there exists a constant $c_R > 0$ such that

$$\frac{1}{c_R} h'_K \leq h_K \leq c_R h'_K \quad \text{for all } K' \subset \overline{\omega_n(K)}, \tag{2.28}$$

for all $K \in \mathcal{T}_n$ and all $n \in \mathbb{N}$. Let us inductively define the neighbouring patch of K of level $\ell \in \mathbb{N}$. Let $\omega_n^1(K) := \omega_n(K)$, and then define

$$\omega_n^\ell(K) := \bigcup \{\omega_n(K') \colon\ K' \in \mathcal{T}_n \ \text{ and } \ K' \subset \overline{\omega_n^{\ell-1}(K)}\} \quad \text{for } \ell \geq 2. \tag{2.29}$$

For fixed $\ell \in \mathbb{N}$ the overlap of $\omega_n^\ell(K)$ is bounded by a constant, depending on ℓ, but independent of $K \in \mathcal{T}_n$ and $n \in \mathbb{N}$ and an estimate analogous to (2.28) holds with constants depending on ℓ.

For part of the convergence proof in the unsteady case we additionally assume quasiuniformity of the mesh so that inverse estimates are applicable, see Lemma 5.14.

Assumption 2.19 (Quasiuniform Triangulations $\{\mathcal{T}_n\}_{n\in\mathbb{N}}$)**.**
Assume that the family $\{\mathcal{T}_n\}_{n\in\mathbb{N}}$ as in Assumption 2.18 is quasiuniform*, i.e., there exists a constant $c_q > 0$ (independent of $n \in \mathbb{N}$) such that*

$$h_K \geq c_q h_n \quad \textit{for all } K \in \mathcal{T}_n \textit{ and all } n \in \mathbb{N}.$$

Finite Element Spaces

Let us introduce the pairs of conforming finite element spaces for velocity and pressure.

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 [DKS13a]. Denoting by $\widehat{\mathcal{P}}_r$ the space of all polynomial functions (in d variables) on the standard reference simplex of order $\leq r \in \mathbb{N}_0$, we assume that $\widehat{\mathcal{P}}_1^d \subset \widehat{\mathbb{P}}_{\mathbb{V}}$. 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 the partition $\mathcal{T}_n$ by

$$\mathbb{V}^n := \{\boldsymbol{V} \in \mathbb{V} \colon\ \boldsymbol{V}|_K \circ \mathbf{F}_K^{-1} \in \widehat{\mathbb{P}}_{\mathbb{V}},\ \text{ for all } K \in \mathcal{T}_n \text{ and } \boldsymbol{V}|_{\partial\Omega} = \mathbf{0}\}, \tag{2.30}$$

$$\mathbb{Q}^n := \{Q \in \mathbb{Q} \colon\ Q|_K \circ \mathbf{F}_K^{-1} \in \widehat{\mathbb{P}}_{\mathbb{Q}},\ \text{ for all } K \in \mathcal{T}_n\}. \tag{2.31}$$

Further, for the choice $\widehat{\mathbb{P}}_{\mathbb{V}} = \widehat{\mathcal{P}}_1^d$, with $r \in \mathbb{N}$ we denote $\mathbb{V}_r^n := \mathbb{V}^n$, which is the space of vector-valued continuous, piecewise polynomial functions of degree $\leq r$ in d variables.

The following assumption regarding locally supported basis is not a strong restriction, since it is satisfied by all relevant finite element spaces, also for $\mathbb{V}^n$. For spaces with locally supported bases the stiffness matrix, arising when solving the approximate problem computationally, is sparse.

Assumption 2.20 (Locally Supported Basis of $\mathbb{Q}^n$)**.**
Let us assume, that $\mathbb{V}^n$ and $\mathbb{Q}^n$ have locally supported bases. In the case of $\mathbb{Q}^n$ this means that for $n \in \mathbb{N}$ there exists $d_n \in \mathbb{N}$ and $\{Q_i^n\}_{i\in\{1,\ldots,d_n\}}$ such that

$$\mathbb{Q}^n = \operatorname{span}\{Q_i^n \colon\ i \in \{1,\ldots,d_n\}\},$$

and for any basis function Q_i^n, $i \in \{1,\ldots,d_n\}$ holds: If there exists an element $K \in \mathcal{T}_n$

such that $Q_i^n|_K \neq 0$, *then this implies that* $\operatorname{supp} Q_i^n \subset \omega_n(K)$, *with the neighbourhood* $\omega_n(K)$ *defined as in* (2.27)*.*

This assumption is used for the proof of the local divergence corrected Lipschitz approximation developed in [DKS13a].

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

$$\mathbb{V}^n_{\mathrm{div}} := \{\boldsymbol{V} \in \mathbb{V}^n \colon \ \langle \operatorname{div} \boldsymbol{V}, Q\rangle_\Omega = 0 \ \text{ for all } Q \in \mathbb{Q}^n\}, \tag{2.32}$$

$$\mathbb{Q}^n_0 := \{Q \in \mathbb{Q}^n : \int_\Omega Q \,\mathrm{d}\boldsymbol{x} = 0\}. \tag{2.33}$$

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

Projectors

For the convergence analysis we require certain projectors to the respective finite element spaces and suitable assumptions on them. Given a Banach space X and a subspace Y we call $P\colon \mathrm{X} \to \mathrm{Y}$ a *projection operator* (or projector), if

$$P(y) = y, \qquad \text{for all } y \in \mathrm{Y}. \tag{2.34}$$

For the convergence proof we assume the existence of projection operators mapping to the respective finite element functions.

Assumption 2.21 (Approximability in $\mathbb{Q}^n$ and projector $\Pi^n_{\mathbb{Q}}$, [DKS13a, Asm. 5, 7])**.** *Assume that for all* $p \in [1, \infty)$*, we have that*

$$\inf_{Q \in \mathbb{Q}^n} \|h - Q\|_{\mathrm{L}^p(\Omega)} \to 0, \qquad \text{as } n \to \infty \qquad \text{for all } h \in \mathrm{L}^p(\Omega). \tag{2.35}$$

Further, assume that for each $n \in \mathbb{N}$ *there exists a linear projection operator* $\Pi^n_{\mathbb{Q}} \colon \mathrm{L}^1(\Omega) \to \mathbb{Q}^n$ *such that, for any* $p \in (1, \infty)$*, there exists a constant* $c(p) > 0$ *such that*

$$\left\|\Pi^n_{\mathbb{Q}} h\right\|_{\mathrm{L}^p(\Omega)} \leq c \,\|h\|_{\mathrm{L}^p(\Omega)} \qquad \text{for all } h \in \mathrm{L}^p(\Omega) \text{ and all } n \in \mathbb{N}. \tag{2.36}$$

As mentioned in [DKS13b] the approximation property in (2.35) implies that $h_n \to 0$, as $n \to \infty$. The stability in (2.36) and the approximability in (2.35) imply that

$$\left\|h - \Pi^n_{\mathbb{Q}} h\right\|_{\mathrm{L}^p(\Omega)} \to 0, \qquad \text{as } n \to \infty \tag{2.37}$$

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

Assumption 2.21 is satisfied, if $\widehat{\mathcal{P}}_0 \subset \widehat{\mathbb{P}}_{\mathbb{Q}}$ and $\mathbb{Q} = \mathrm{L}^\infty(\Omega)$, i.e., the pressure space contains discontinuous functions. In this case the projector $\Pi^n_{\mathbb{Q}}$ can be chosen a local L^2-projection (taking the average integral over an element as value) or more generally a Clément type interpolation operator, see [Clé75]. Also the approximation property is satisfied for continuous pressure spaces, i.e., $\mathbb{Q} = \mathrm{C}(\overline{\Omega})$, if $\widehat{\mathcal{P}}_1 \subset \widehat{\mathbb{P}}_{\mathbb{Q}}$. Then $\Pi^n_{\mathbb{Q}}$ can be chosen as a Clément interpolation operator, see [EG04, Sec. 1.6.1].

The following lemma regarding a local projection operator mapping to spaces of continuous piecewise polynomial functions is satisfied, e.g., by the Scott–Zhang interpolation operator ([SZ90]) or a variant of the Clément interpolation operator ([EG04, Rem. 1.129]).

Lemma 2.22 (Projector P_r^n)**.**
Let Assumption 2.18 on the domain Ω and on the family of triangulations $\{\mathcal{T}_n\}_{n\in\mathbb{N}}$ be satisfied and let $\mathbb{V}^n$ as defined in (2.30) with $\widehat{\mathcal{P}}_1^d \subset \widehat{\mathbb{P}}_{\mathbb{V}} \subset \mathrm{W}^{1,\infty}(\widehat{K})^d$. Let $r \in \mathbb{N}$ be maximal such that $\widehat{\mathcal{P}}_r^d \subset \widehat{\mathbb{P}}_{\mathbb{V}}$.

Then there exists a linear projection operator $P_r^n \colon \mathrm{W}_0^{1,1}(\Omega)^d \to \mathbb{V}_r^n \subset \mathbb{V}^n$ such that for any $p \in [1,\infty)$, any $\mu \in \{0,1\}$ and for any $s \in \{1,\dots,r+1\}$ there exists a constant $c > 0$ (independent of K and n) such that

$$|\boldsymbol{v} - P_r^n \boldsymbol{v}|_{\mathrm{W}^{\mu,p}(K)} \le c h_K^{s-\mu} |\boldsymbol{v}|_{\mathrm{W}^{s,p}(\omega_n(K))}, \tag{2.38}$$

for all $K \in \mathcal{T}_n$, all $n \in \mathbb{N}$ and all $\boldsymbol{v} \in \mathrm{W}_0^{1,p}(\Omega)^d \cap \mathrm{W}^{s,p}(\Omega)^d$.

The approximation property of P_r^n implies that $\mathbb{V}^n$ satisfies a (global) approximation property corresponding to (2.35) in $\mathrm{W}_0^{1,p}(\Omega)^d$. To deal with the (discrete) solenoidality similar as in [BBDR12, DKS13a] we assume that there exists a divergence-preserving projector with suitable stability properties which will vary depending on the situations of the convergence proof.

Assumption 2.23 (Projector Π^n)**.**
Assume that for each $n \in \mathbb{N}$ there exists a linear projection operator $\Pi^n \colon \mathrm{W}_0^{1,1}(\Omega)^d \to \mathbb{V}^n$ such that:

(i) (preservation of the divergence in $(\mathbb{Q}^n)'$) for any $\boldsymbol{v} \in \mathrm{W}_0^{1,1}(\Omega)^d$ one has that

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

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

$$\|\Pi^n \boldsymbol{v}\|_{\mathrm{W}^{1,p}(\Omega)} \le c\, \|\boldsymbol{v}\|_{\mathrm{W}^{1,p}(\Omega)} \qquad \textit{for all } \boldsymbol{v} \in \mathrm{W}_0^{1,p}(\Omega)^d \textit{ and all } n \in \mathbb{N}.$$

In some cases we additionally assume one of the following stronger stability properties:

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

$$\|\Pi^n \boldsymbol{v}\|_{\mathrm{L}^p(\Omega)} \le c \left(\|\boldsymbol{v}\|_{\mathrm{L}^p(\Omega)} + h_n \|\nabla \boldsymbol{v}\|_{\mathrm{L}^p(\Omega)} \right)$$

for all $\boldsymbol{v} \in \mathrm{W}_0^{1,p}(\Omega)^d$ and all $n \in \mathbb{N}$.

(iib) (local $\mathrm{W}^{1,1}$-stability with level $\ell \in \mathbb{N}$) there exists an $\ell \in \mathbb{N}$ and there exists a constant $c_1 > 0$ (both independent of K and n) such that

$$\fint_K |\Pi^n \boldsymbol{v}| \,\mathrm{d}\boldsymbol{x} \le c_1 \fint_{\omega_n^\ell(K)} (|\boldsymbol{v}| + h_K |\nabla \boldsymbol{v}|)\,\mathrm{d}\boldsymbol{x},$$

for all $\boldsymbol{v} \in \mathrm{W}_0^{1,1}(\Omega)^d$, all $K \in \mathcal{T}_n$ and all $n \in \mathbb{N}$.

Projection operators Π_p^n satisfying Assumption 2.23 (i) and (ii) (for a fixed $p \in (1,\infty)$ and they need not be linear or idempotent) are sometimes referred to as *Fortin operators*, compare [For77, EG04, BBF13, EG16]. For fixed $p \in (1,\infty)$ the Fortin lemma states that, given the continuous inf-sup condition in Corollary A.3, the existence of a Fortin operator Π_p^n implies the discrete inf-sup condition, see [EG04, Lem. 4.19], [BBF13, Prop. 5.4.2], and in the same way the discrete inf-sup condition can be shown to hold here, see Lemma A.4.

Remark 2.24 (Properties of Π^n)**.**

(i) The global stability in Assumption 2.23 (ii) and approximation properties of P_1^n according to Lemma 2.22 with $s = 2$ and an approximation argument yield that

$$\|\boldsymbol{v} - \Pi^n \boldsymbol{v}\|_{W^{1,p}(\Omega)} \to 0, \qquad \text{as } n \to \infty, \tag{2.39}$$

for all $\boldsymbol{v} \in W_0^{1,p}(\Omega)^d$ with $p \in [1, \infty)$. This property is required for all convergence proofs in Chapters 4 and 5.

(ii) The additional stability property in Assumption 2.23 (iia) is used to derive finer approximation properties, which include rates, see Lemma 2.25 below. The latter is needed to establish stability properties of a L^2-projection applied for a part of the convergence proof for the unregularised problem in the unsteady case, see Chaper 5, Lemma 5.14.

(iii) The strongest assumption of local $W^{1,1}$-stability in Assumption 2.23 (iib) with $\ell = 1$ is required for the discrete Lipschitz approximation in Lemma 2.29 used in the (unregularised) steady case, see Chapter 4, Section 4.3.2 in order to control the size of the bad set, compare [DKS13a]. Let us show that the local $W^{1,1}$-stability with $\ell \in \mathbb{N}_0$ indeed implies both (ii) and (iia). With convexity arguments as in [DKS13a] (iib) implies that for any $p \in [1, \infty)$ there exists a constant $c > 0$ such that

$$\fint_K |\Pi^n \boldsymbol{v}|^p \, d\boldsymbol{x} \le c \fint_{\omega_n^\ell(K)} |\boldsymbol{v}|^p + h_K^p |\nabla \boldsymbol{v}|^p \, d\boldsymbol{x}, \tag{2.40}$$

for any $K \in \mathcal{T}_n$, all $n \in \mathbb{N}$ and all $\boldsymbol{v} \in W_0^{1,p}(\Omega)^d$. Summing over $K \in \mathcal{T}_n$ and using the fact that the number of elements in $\omega_n^\ell(K)$ with constant only depending on the fixed $\ell \in \mathbb{N}$ is bounded, uniformly in $K \in \mathcal{T}_n$ and in $n \in \mathbb{N}$ we obtain (iia).

To show that (iib) implies (ii) we have to bound the gradients as well. We add and subtract $\nabla \boldsymbol{v}$ and $\nabla P_r^n \boldsymbol{v} \in \mathbb{V}^n \subset W_0^{1,1}(\Omega)^d$ and use the fact that Π^n is a projector, apply a local inverse estimate, (2.40) and the approximation property of P_r^n according to Lemma 2.22 with $s = 1$, which shows that

$$\begin{aligned} \|\nabla \Pi^n \boldsymbol{v}\|_{L^p(K)} &\le \|\nabla \Pi^n (\boldsymbol{v} - P_r^n \boldsymbol{v})\|_{L^p(K)} + \|\nabla (P_r^n \boldsymbol{v} - \boldsymbol{v})\|_{L^p(K)} \\ &\le c h_K^{-1} \|\Pi^n (\boldsymbol{v} - P_r^n \boldsymbol{v})\|_{L^p(K)} + \|\nabla (P_r^n \boldsymbol{v} - \boldsymbol{v})\|_{L^p(K)} \\ &\le c h_K^{-1} \|\boldsymbol{v} - P_r^n \boldsymbol{v}\|_{L^p(\omega_n^\ell(K))} + c \|\nabla (\boldsymbol{v} - P_r^n \boldsymbol{v})\|_{L^p(\omega_n^\ell(K))} \\ &\overset{(2.38)}{\le} c \|\nabla \boldsymbol{v}\|_{L^p(\omega_n^{\ell+1}(K))}, \end{aligned} \tag{2.41}$$

noting that by the regularity of $\{\mathcal{T}_n\}_{n \in \mathbb{N}}$ and estimate (2.28), h'_K and h_K are comparable (with constant only depending on ℓ for any $K' \subset \overline{\omega_n^{\ell+1}(K)}$). Then summing taking pth powers, summing over $K \in \mathcal{T}_n$, using the bounded overlap of $\omega_n^{\ell+1}(K)$ and taking the pth root implies together with (iia) that (ii) holds.

To show stability properties of an L^2-projection mapping to $\mathbb{V}_{\mathrm{div}}^n$ we will require global approximation properties of Π^n of a certain order. With the following lemma this reduces to assuming that $\mathbb{V}^n$ contains continuous, piecewise polynomials of a certain degree.

Lemma 2.25 (Global Approximation Properties of Π^n)**.**

Let Assumption 2.18 on the domain Ω and the family of triangulations $\{\mathcal{T}_n\}_{n \in \mathbb{N}}$ be satisfied and let $\mathbb{V}^n$ be as defined in (2.30). Let $r \in \mathbb{N}$ be maximal such that $\widehat{\mathcal{P}}_r^d \subset \widehat{\mathbb{P}}_{\mathbb{V}}$. Furthermore, assume that $\Pi^n \colon W_0^{1,1}(\Omega)^d \to \mathbb{V}^n$ satisfies Assumption 2.23 (i)–(iia). Then, for any $p \in [1, \infty)$, any $\mu \in \{0, 1\}$ and for any $s \in \{1, \dots, r+1\}$ there exists a constant $c > 0$ (independent *of n) such that*

$$|\boldsymbol{v} - \Pi^n \boldsymbol{v}|_{\mathrm{W}^{\mu,p}(\Omega)} \leq c h_n^{s-\mu} \, |\boldsymbol{v}|_{\mathrm{W}^{s,p}(\Omega)} \tag{2.42}$$

for all $n \in \mathbb{N}$ and all $\boldsymbol{v} \in \mathrm{W}_0^{1,p}(\Omega)^d \cap \mathrm{W}^{s,p}(\Omega)^d$.

Proof. Adding and subtracting $P_r^n \boldsymbol{v} = \Pi^n P_r^n \boldsymbol{v}$ and applying the triangle inequality yields

$$|\boldsymbol{v} - \Pi^n \boldsymbol{v}|_{\mathrm{W}^{\mu,p}(\Omega)} \leq |\boldsymbol{v} - P_r^n \boldsymbol{v}|_{\mathrm{W}^{\mu,p}(\Omega)} + |\Pi^n(\boldsymbol{v} - P_r^n \boldsymbol{v})|_{\mathrm{W}^{\mu,p}(\Omega)} \, . \tag{2.43}$$

On the second term, if $\mu = 0$ we apply the stability in Assumption 2.23 (iia) to obtain

$$\|\Pi^n(\boldsymbol{v} - P_r^n \boldsymbol{v})\|_{\mathrm{L}^p(\Omega)} \leq \|\boldsymbol{v} - P_r^n \boldsymbol{v}\|_{\mathrm{L}^p(\Omega)} + h_n \, \|\nabla(\boldsymbol{v} - P_r^n \boldsymbol{v})\|_{\mathrm{L}^p(\Omega)} \, . \tag{2.44}$$

And if $\mu = 1$, applying the stability in Assumption 2.23 (ii) shows that

$$\|\nabla \Pi^n(\boldsymbol{v} - P_r^n \boldsymbol{v})\|_{\mathrm{L}^p(\Omega)} \leq \|\boldsymbol{v} - P_r^n \boldsymbol{v}\|_{\mathrm{L}^p(\Omega)} + \|\nabla(\boldsymbol{v} - P_r^n \boldsymbol{v})\|_{\mathrm{L}^p(\Omega)} \, . \tag{2.45}$$

Applying both (2.44) and (2.45) in (2.43) and using Lemma 2.22 shows that

$$\begin{aligned} |\boldsymbol{v} - \Pi^n \boldsymbol{v}|_{\mathrm{W}^{\mu,p}(\Omega)} &\leq |\boldsymbol{v} - P_r^n \boldsymbol{v}|_{\mathrm{W}^{\mu,p}(\Omega)} + \|\boldsymbol{v} - P_r^n \boldsymbol{v}\|_{\mathrm{L}^p(\Omega)} + h_n^{1-\mu} \, \|\nabla(\boldsymbol{v} - P_r^n \boldsymbol{v})\|_{\mathrm{L}^p(\Omega)} \\ &\leq c h_n^{s-\mu} \, |\boldsymbol{v}|_{\mathrm{W}^{s,p}(\Omega)} + c h_n^s \, |\boldsymbol{v}|_{\mathrm{W}^{s,p}(\Omega)} \\ &\leq c h_n^{s-\mu} \, |\boldsymbol{v}|_{\mathrm{W}^{s,p}(\Omega)} \, , \end{aligned}$$

for all $s \in \{1, \ldots, r+1\}$, since we have without loss of generality that $h_n \in (0,1)$. This finishes the proof. □

In the application we will require this approximation property for $s \in \{2, 3\}$, depending on q and d, i.e., we require that $r \in \{1, 2\}$ which is a rather moderate assumption.

In the same manner, assuming that Π^n satisfies the local $\mathrm{W}^{1,1}$-stability with fixed level $\ell \in \mathbb{N}$ in Assumption 2.23 (iib) one can show the corresponding local approximations property of level $\ell + 1$. With this we mean that the estimate has K on the left-hand side and $\omega_n^\ell(K)$ on the right-hand side of the inequality corresponding to (2.42).

The following two approaches to constructing a projection operator Π^n satisfying Assumption 2.23 can be found in the literature: If the pressure space is of low enough order, then one can construct Π^n as (local) divergence correction of a (local) interpolation, see, e.g., [BF13]. For higher order finite element spaces, [GS03, Thm. 2.1] reduces the construction to the verification of a local inf-sup condition and the construction of a local projection, which preserves the divergence in the dual of piecewise constant functions in stead of $\mathbb{Q}^n$.

Let us now collect finite element spaces for which Π^n satisfies the local $\mathrm{W}^{1,1}$-stability in Assumption 2.23 (iib) for some $\ell \in \mathbb{N}$, see also [BBDR12]. For more details let us refer to Appendix A.2.2.

Example 2.26 (Finite Element Spaces)**.**
Let us collect pairs of finite element spaces $(\mathbb{V}^n, \mathbb{Q}^n)$, for which a projection $\Pi^n \colon \mathrm{W}_0^{1,1}(\Omega)^d \to \mathbb{V}^n$ exists, that satisfies Assumption 2.23 (i), i.e., it is divergence-preserving, and it is locally $\mathrm{W}^{1,1}$-stable with level $\ell \in \mathbb{N}$ as formulated in (iib).

(i) the $\mathbb{P}_2$–$\mathbb{P}_0$ element for $d = 2$, $r = 2$, see [BBF13, Sec. 8.4.3] and the projection operator Π^n satisfying Assumption 2.23 (i) is constructed in the proof of Prop. 8.4.3 therein. The proof of a local approximation property implying (iib) with $\ell = 1$ is contained in Appendix A.2.2.1.

(ii) *the* first order Bernardi–Raugel element *($r = 1$) for $d \in \{2, 3\}$ introduced in [BR85, Sec. II]. For $d = 2$ this is sometimes referred to as reduced $\mathbb{P}_2 - \mathbb{P}_0$ element or as 2D SMALL element, see also [BBF13, Rem. 8.4.2] and [GR86, Ch. II.2.1]. A local divergence-preserving projector Π^n is introduced in [BR85] and a local approximation result is proved for $p = 2$. In [GL01] this is extended to $p \in [2, \infty)$. The proof of the local approximation property for all $p \in [1, \infty)$ is contained in Appendix A.2.2.1, which in turn implies the local $\mathrm{W}^{1,1}$-stability with level $\ell = 1$.*

(iii) *the* conforming Crouzeix–Raviart element*, for $d = 2$ and $r = 2$, introduced in [CR73] and presented in [BBF13, Ex. 8.6.1] and its generalisations for $r \geq 2$, see [CR73, Man82], [BBF13, Prop. 8.6.2] and [GR86, Ch. II.2.2]. For $r = 2$ in [GS03, Thm. 3.3] it is shown that a divergence-preserving projector Π^n exists, that satisfies a quasi-local approximation property and it is stated that the proof extends to general $r \geq 2$. Investigating the proof of [GS03, Thm 2.1] carefully one can see that the local approximation property implies (iib) with $\ell = 2$, see Appendix A.2.2.2.*

(iv) *the* second order Bernardi–Raugel element *($r = 2$) for $d = 3$ as introduced in [BR85, Sec. III], see also [BBF13, Ex. 8.7.2] and [GR86, Ch. II.2.3]. This element can be seen as a 3D generalisation of the standard Crouzeix–Raviart element. The projection Π^n is introduced in (III.3) in [BR85] and a local approximation property is shown for $p = 2$, which implies local $\mathrm{W}^{1,2}$-stability. The proof of the local approximation property for all $p \in [1, \infty)$ implying the local $\mathrm{W}^{1,1}$-stability with level $\ell = 1$ is contained in Appendix A.2.2.1.*

(v) *the* MINI element *for $d \in \{2, 3\}$ ($r = 1$), as introduced in [ABF84] (for $d = 2$), see also [GR86, CH. II.4.1] [BBF13, Sec. 8.4.2, 8.7.1]. For $d = 3$ the projection is constructed in (4.17) in [GL01] and a local approximation property is proved for $p \in [2, \infty)$. In [BBDR12, A.1] local $\mathrm{W}^{1,1}$-stability with level $\ell = 1$ is shown, see also Appendix A.2.2.1.*

(vi) *the* Taylor–Hood element *due to [TH73] (for $d = r = 2$), see also [GR86, Ch. II.4.2], and its generalisations for $r \geq 2$, $d \in \{2, 3\}$, see [BBF13, Sec. 8.8.2]. It is proven in [GS03, Sec. 3.1, 3.2] that there exists a divergence-preserving projection Π^n satisfying a quasi-local approximation property, if $r \geq d$ and if each element $K \in \mathcal{T}_n$, $n \in \mathbb{N}$ has at least one interior vertex. The quasi-local approximation property formulated for macro-elements implies the local $\mathrm{W}^{1,1}$-stability of level $\ell = 4$, see Appendix A.2.2.2.*

To summarise, let us note that all the above examples satisfy the global stability properties in Assumption 2.23 (ii) and (iia) and hence are suitable for the regularised problems or the unregularised cases for the admissible range, see Sections 4.3.1, 5.3.1 and 5.3.2.

The construction of the discrete Lipschitz approximation in [DKS13a] (see Lemma 2.29 below), is based on the assumption of local $\mathrm{W}^{1,1}$-stability in Assumption 2.23 (iib) with $\ell = 1$. This is satisfied for all examples except of the family of Taylor–Hood elements and hence they are suited for the unregularised problem in the steady case for $q \in \left(\frac{2d}{d+1}, \frac{3d}{d+2}\right)$, when applying a discrete Lipschitz approximation, see Section 4.3.2. However, the construction of the discrete Lipschitz approximation can be adapted, such that the weaker local $\mathrm{W}^{1,1}$-stability with some fixed level $\ell \in \mathbb{N}$ suffices and then all examples are suited for the convergence proof in the non-admissible case for the steady problem.

Note that Assumptions 2.20 and 2.21 are satisfied for all the examples of space $\mathbb{Q}^n$.

Exactly Divergence-free Finite Element Functions

For some families of mixed finite element space $(\mathbb{V}^n, \mathbb{Q}^n)$ the functions in $\mathbb{V}^n_{\mathrm{div}}$, the space of of discretely divergence-free finite element functions, are already exactly divergence-free. This fact simplifies some arguments in the analysis, since it means that the velocity space is conforming with respect to the incompressibility condition.

Assumption 2.27 (Exact Solenoidality of $\mathbb{V}^n_{\mathrm{div}}$)**.**
For $(\mathbb{V}^n, \mathbb{Q}^n)$, with $n \in \mathbb{N}$, as defined in (2.30), (2.31) and $\mathbb{Q}^n_0$ as defined in 2.33 assume that

$$\operatorname{div} \mathbb{V}^n \subset \mathbb{Q}^n_0 \quad \textit{for all } n \in \mathbb{N}. \tag{2.46}$$

The condition (2.46) is sufficient to show that the functions in $\mathbb{V}^n_{\mathrm{div}}$ as defined in (2.32) are exactly (pointwise) divergence-free, i.e.,

$$\mathbb{V}^n_{\mathrm{div}} \subset \mathrm{W}^{1,\infty}_{0,\mathrm{div}}(\Omega)^d \quad \text{for all } n \in \mathbb{N}. \tag{2.47}$$

Examples of finite element spaces satisfying Assumption 2.27 available in the literature are based on local spaces $\widehat{\mathbb{P}}_{\mathbb{V}}$ of comparably high dimension to account for the solenoidality condition. This makes them less attractive from a computational point of view. Also their complexity makes it more difficult to analyse them theoretically.

However, they do not have some of the shortcomings observed for discretely divergence-free finite element functions. For example exactly divergence-free finite element spaces have better conservation properties and pressure robustness properties, see for details [JLM+17, LR18]. For this reason, recently some effort has been made to investigate such finite element spaces.

Example 2.28 (Exactly Divergence-free Finite Element Spaces)**.**
The following families of finite element spaces satisfy Assumption 2.27 and there exists a projection Π^n satisfying Assumption 2.23 (i), (iib) with some $\ell \in \mathbb{N}$:

(vii) the family of Guzmán–Neilan elements *for $d = 3$ ($r = 1$), and $d = 2$ ($r \geq 1$), as introduced in [GN14a, GN14b]. The construction of the velocity space $\mathbb{V}^n$ is based on enriching piecewise polynomials by polynomial and rational bubble functions. The additional degrees of freedom allow one to verify Assumption 2.27, see [GN14b, GN14a, Sec. 4] but they also increase the dimension of the local space considerably. The dimension of the local space for $d = 3$ is 60 (and there is a reduced version with dimension 16). For $d = 2$ in the lowest order case ($r = 1$) the local dimension is 7 (and 4 for a reduced version). The projection operator Π^n is constructed explicitly. For $d = 2$ [GN14b] contains a local approximation property for $p = 2$, which implies local $\mathrm{W}^{1,2}$-stability with $\ell = 1$. The proof in [GN14b] can be extended to $p \in [1, \infty)$, which shows the $\mathrm{W}^{1,1}$-stability in (iib) with $\ell = 1$; see also [DKS13a, Ex. 12] for a direct proof. For $d = 3$ local approximation properties for $p \in [1, \infty)$ are proved in [GN14a], which implies the local $\mathrm{W}^{1,1}$-stability with $\ell = 1$.*

(viii) the Scott–Vogelius element for $d = 2$ and $r \geq 4$, introduced in [SV85], see also [BBF13, Sec. 8.8.1] and [GS17], under suitable conditions on the singular vertices of the triangulation $\{\mathcal{T}_n\}_{n\in\mathbb{N}}$. The fact that Assumption 2.27 is satisfied is proved in [SV85] without condition on the mesh, see also [GS17, Lem. 2]. In [GS17] the authors show inf-sup stability without relying on mesh restrictions assumed in the earlier proof in [SV85]. Based on the work in [GS17] we construct a projection operator Π^n satisfying Assumption 2.23 (i). Furthermore, we can verify the local $\mathrm{W}^{1,1}$-stability as stated in (iib) with $\ell \geq 2$, provided that the singular vertices do not cluster in neighbourhoods larger than $\ell - 2$; for details see Appendix A.2.2.2.

Note that as before both examples also satisfy Assumption 2.20 and 2.21, since the respective pressure spaces consist of discontinuous, piecewise polynomial functions.

Here, their main benefit is, that the range of q, for which the convergence can be shown, is the whole range of existence $q > \frac{2d}{d+2}$, for exactly divergence-free finite element functions when using a discrete truncation in the steady case.

2.2.1.2 Discrete Divergence Corrected Lipschitz Approximation

The discrete Lipschitz approximation for steady problems was introduced in [DKS13a] and is constructed as composition of the continuous Lipschitz truncation of a function in $\mathbb{V}^n$, see Lemma 2.12 with the projection operator Π^n satisfying Assumption 2.23. It is used in case admissibility is lost in the discretisation limit. Since Π^n is continuous, the convergence and stability properties carry over. The assumptions of locality of a basis of $\mathbb{Q}^n$ and the local $\mathrm{W}^{1,1}$-stability of Π^n are required in order to control that the size of the bad set does not increase too much by projecting to $\mathbb{V}^n_{\mathrm{div}}$. Note that at this stage for the following lemma one could use any projection satisfying the local $\mathrm{W}^{1,1}$-stability. Only for the discrete Bogovskiĭ correction we will require that Π^n is divergence-preserving.

Lemma 2.29 (Discrete Lipschitz Approximation, see [DKS13a, Cor. 17])**.**
Let Assumption 2.18 on Ω and the family of simplicial partitions be satisfied. Let $\mathbb{V}^n$ be as defined in (2.30) and assume that Assumption 2.23 (iib) with $\ell = 1$ holds. Let $p \in (1,\infty)$ and let $\{\boldsymbol{V}^n\}_{n\in\mathbb{N}}$ be a sequence such that $\boldsymbol{V}^n \in \mathbb{V}^n$ for each $n \in \mathbb{N}$, converging to zero weakly in $\mathrm{W}^{1,p}_0(\Omega)^d$, as $n \to \infty$.
Then, there exist

- *a double sequence $\{\lambda_{n,j}\}_{n,j\in\mathbb{N}} \subset \mathbb{R}$ with $\lambda_{n,j} \in [2^{2^j}, 2^{2^{j+1}-1}]$ for any $n, j \in \mathbb{N}$;*
- *a double sequence of open sets $\mathcal{B}_{n,j} \subset \Omega$, $n, j \in \mathbb{N}$ of the form*
$$\mathcal{B}_{n,j} = \mathrm{int}\left(\bigcup\{K\colon\ K \in \mathcal{T}_{n,j}\}\right),$$
where $\mathcal{T}_{n,j} \subset \mathcal{T}_n$, $n, j \in \mathbb{N}$ is a subset of elements of the triangulation $\mathcal{T}_n$;
- *a double sequence of functions $\{\boldsymbol{V}^{n,j}\}_{n,j\in\mathbb{N}} \subset \mathrm{W}^{1,\infty}_0(\Omega)^d$ such that $\boldsymbol{V}^{n,j} \in \mathbb{V}^n$ for any $n, j \in \mathbb{N}$,*

such that

(i) $\boldsymbol{V}^{n,j} = \boldsymbol{V}^n$ *on* $\Omega\backslash\mathcal{B}_{n,j}$, *i.e.,* $\{\boldsymbol{V}^{n,j} \neq \boldsymbol{V}^n\} \subset \mathcal{B}_{n,j}$ *for all* $n, j \in \mathbb{N}$;

(ii) there exists a constant $c = c(p) > 0$ such that
$$\left\|\lambda_{n,j}\mathbb{1}_{\mathcal{B}_{n,j}}\right\|_{\mathrm{L}^p(\Omega)} \leq c2^{-\frac{j}{p}}, \qquad \textit{for all } n, j \in \mathbb{N};$$

(iii) there exists a constant $c > 0$ such that
$$\left\|\nabla\boldsymbol{V}^{n,j}\right\|_{\mathrm{L}^\infty(\Omega)} \leq c\lambda_{n,j} \qquad \textit{for all } n, j \in \mathbb{N};$$

(iv) for any fixed $j \in \mathbb{N}$ we have
$$\begin{aligned} \boldsymbol{V}^{n,j} &\to \mathbf{0} && \textit{strongly in } \mathrm{L}^\infty(\Omega)^d, \\ \nabla\boldsymbol{V}^{n,j} &\overset{*}{\rightharpoonup} \mathbf{0} && \textit{weakly* in } \mathrm{L}^\infty(\Omega)^{d\times d}, \end{aligned}$$
as $n \to \infty$.

Note that the analogous construction works can be done also for general (fixed) $\ell \in \mathbb{N}$ in Assumption 2.23 (iib).

By use of the projection operator Π^n it is possible to construct a discrete Bogovskiĭ operator $\mathfrak{B}^n\colon \operatorname{div}\mathbb{V}^n \to \mathbb{V}^n$, see Lemma A.6, which was proved in [DKS13a]. Then, similarly as in the continuous situation in Lemma 2.13 a (discrete) divergence-correction of the discrete truncation $\{\boldsymbol{V}^{n,j}\}_{n,j\in\mathbb{N}}$ is obtained by

$$\boldsymbol{W}^{n,j} := \boldsymbol{V}^{n,j} - \mathfrak{B}^n(\operatorname{div}\boldsymbol{V}^{n,j}). \tag{2.48}$$

The so defined functions are discretely divergence-free. If the original sequence $\{\boldsymbol{V}^n\}_{n\in\mathbb{N}}$ is discretely divergence-free the lemma corresponding to Lemma 2.13 can be shown under the additional condition that Q^n has a locally supported basis.

Since the proof is contained in the proof of the main result in [DKS13a] we reproduce it in Subsection A.2.1.

Lemma 2.30 (Discrete Divergence Correction, [DKS13a, pp. 1006–1007])**.**
Additionally to the assumptions in Lemma 2.29 let $\mathbb{V}^n_{\mathrm{div}}$ and $\mathbb{Q}^n$ be as defined in (2.32) and (2.31), respectively, and let Assumptions 2.20, 2.21 and 2.23 (i) be satisfied.
Let $p \in (1,\infty)$ and let $\{\boldsymbol{V}^n\}_{n\in\mathbb{N}}$ be a sequence such that $\boldsymbol{V}^n \in \mathbb{V}^n_{\mathrm{div}}$ for all $n \in \mathbb{N}$, converging to zero weakly in $\mathrm{W}^{1,p}_0(\Omega)^d$, as $n \to \infty$. Furthermore, let $\{\boldsymbol{V}^{n,j}\}_{n,j\in\mathbb{N}}$ be the sequence of discrete Lipschitz truncations given by Lemma 2.29.
Then, there exists a double sequence $\{\boldsymbol{W}^{n,j}\}_{n,j\in\mathbb{N}}$ such that

(i) $\boldsymbol{W}^{n,j} \in \mathbb{V}^n_{\mathrm{div}}$ *for all* $n, j \in \mathbb{N}$*;*

(ii) there exists a constant $c = c(p) > 0$ *such that*

$$\left\|\boldsymbol{V}^{n,j} - \boldsymbol{W}^{n,j}\right\|_{\mathrm{W}^{1,p}(\Omega)} \le c2^{-\frac{j}{p}} \qquad \textit{for all } n, j \in \mathbb{N};$$

(iii) for any fixed $j \in \mathbb{N}$ *we have (up to subsequences)*

$$\begin{aligned} \boldsymbol{W}^{n,j} &\to \mathbf{0} \qquad \textit{strongly in } \mathrm{L}^s(\Omega)^d, \\ \boldsymbol{W}^{n,j} &\rightharpoonup \mathbf{0} \qquad \textit{weakly in } \mathrm{W}^{1,s}_0(\Omega)^d, \end{aligned}$$

as $n \to \infty$*, for any* $s \in (1,\infty)$*.*

2.2.2. Convective Term and its Numerical Approximation

We have mentioned in Chapter 1 that the convective term represents one of the difficulties in the analysis of the problem; it causes problems concerning admissibility of the weak solution as a test function for small $q \in (1,\infty)$. When considering a numerical approximation this problem is further aggravated by a modification of the convective term required to preserve some of the properties of the convective term.

Here we want to collect all estimates on the convective term and its numerical modification. For this we introduce the ranges of q both in the steady and the unsteady case, for which the estimates hold, and set the notation for Chapters 4 and 5.

Motivated by the form of the convective term in the conservation of momentum equation and the term corresponding to it in the weak formulation, we consider the trilinear form b defined by

$$b(\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}) := -\langle \boldsymbol{u}\otimes\boldsymbol{v}, \nabla\boldsymbol{w}\rangle_\Omega = \langle \boldsymbol{u}\otimes\boldsymbol{w}, \nabla\boldsymbol{v}\rangle_\Omega - \langle \operatorname{div}\boldsymbol{u}, \boldsymbol{v}\cdot\boldsymbol{w}\rangle_\Omega\,, \tag{2.49}$$

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 $\boldsymbol{u} \in \mathrm{W}^{1,\infty}_{0,\mathrm{div}}(\Omega)^d$ and $\boldsymbol{v} \in \mathrm{W}^{1,\infty}_0(\Omega)^d$.

As in general $\mathbb{V}^n_{\mathrm{div}} \not\subset \mathrm{W}^{1,\infty}_{0,\mathrm{div}}(\Omega)^d$, the second term in (2.49) need not vanish. 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., [Tem84]) is therefore to consider instead the skew-symmetric trilinear form

$$\begin{aligned}\widetilde{b}(\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}) &:= \tfrac{1}{2}\left(\langle \boldsymbol{u}\otimes\boldsymbol{w},\nabla\boldsymbol{v}\rangle_\Omega - \langle \boldsymbol{u}\otimes\boldsymbol{v},\nabla\boldsymbol{w}\rangle_\Omega\right)\\ &= -\langle \boldsymbol{u}\otimes\boldsymbol{v},\nabla\boldsymbol{w}\rangle_\Omega + \tfrac{1}{2}\langle \operatorname{div}\boldsymbol{u},\boldsymbol{v}\cdot\boldsymbol{w}\rangle_\Omega\,,\end{aligned} \tag{2.50}$$

for $\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}\in \mathrm{W}_0^{1,\infty}(\Omega)^d$. Thus we have that $\widetilde{b}(\boldsymbol{u},\boldsymbol{v},\boldsymbol{v})=0$ regardless of the solenoidality of $\boldsymbol{u}$. Note that we have $b(\boldsymbol{u},\cdot,\cdot)=\widetilde{b}(\boldsymbol{u},\cdot,\cdot)$ for divergence-free functions $\boldsymbol{u}$, and in particular for functions in $\mathbb{V}^n_{\operatorname{div}}$, if Assumption 2.27 is satisfied. Note that the bounds below allow us to extend $b(\cdot,\cdot,\cdot)$ and $\widetilde{b}(\cdot,\cdot,\cdot)$ to the respective Sobolev spaces by an approximation argument.

The Steady Case

In the steady situation we will aim for $\boldsymbol{u}\in \mathrm{W}_0^{1,q}(\Omega)^d$ and we have the Sobolev embedding $\mathrm{W}^{1,q}(\Omega)\hookrightarrow \mathrm{L}^{q^*}(\Omega)$, provided that $q\in[1,d)$ and otherwise $\mathrm{W}^{1,q}(\Omega)\hookrightarrow \mathrm{L}^{p}(\Omega)$ for all $p\in[1,\infty)$. Hence, when testing with $\boldsymbol{v}\in \mathrm{W}_0^{1,q}(\Omega)^d$ by Hölder's inequality and Sobolev embedding we find that

$$|\langle \boldsymbol{u}\otimes\boldsymbol{u},\nabla\boldsymbol{v}\rangle_\Omega| \le \|\boldsymbol{u}\|^2_{\mathrm{L}^{2q'}(\Omega)}\|\nabla\boldsymbol{v}\|_{\mathrm{L}^{q}(\Omega)} \le c\,\|\boldsymbol{u}\|^2_{\mathrm{W}^{1,q}(\Omega)}\|\boldsymbol{v}\|_{\mathrm{W}^{1,q}(\Omega)}\,, \tag{2.51}$$

provided that $2q'\ge q^*$ if $q\in[1,d)$, which is equivalent to $q\in\left[\frac{3d}{d+2},d\right)$, and without restriction if $q\in[d,\infty)$. For $q<\frac{3d}{d+2}$ the function space for the test function $\boldsymbol{v}$ is smaller, and $\boldsymbol{u}$ is not an admissible test function anymore. Again by Hölder's inequality one has

$$|\langle \boldsymbol{u}\otimes\boldsymbol{u},\nabla\boldsymbol{v}\rangle_\Omega| \le \|\boldsymbol{u}\|^2_{\mathrm{L}^{q^*}(\Omega)}\|\nabla\boldsymbol{v}\|_{\mathrm{L}^{(q^*/2)'}(\Omega)} \le c\,\|\boldsymbol{u}\|^2_{\mathrm{W}^{1,q}(\Omega)}\|\boldsymbol{v}\|_{\mathrm{W}^{1,(q^*/2)'}(\Omega)}\,, \tag{2.52}$$

for $\boldsymbol{v}\in \mathrm{W}^{1,(q^*/2)'}(\Omega)^d$, provided that $\frac{q^*}{2}\ge 1$, which is equivalent to $q\ge\frac{2d}{d+2}$. Let us define

$$q^\times := \begin{cases} q^*=\frac{dq}{d-1} & \text{if } q\in[1,d),\\ \infty & \text{otherwise,}\end{cases} \tag{2.53}$$

and $\widetilde{q}$ by which coincides with the critical Sobolev embedding exponent for $q<d$, and

$$\widetilde{q} := \min\left(\frac{q^\times}{2},q'\right) = \begin{cases} \frac{q^*}{2} & \text{if } q\in\left[\frac{2d}{d+2},\frac{3d}{d+2}\right),\\ q' & \text{if } q\in\left[\frac{3d}{d+2},\infty\right),\end{cases} \tag{2.54}$$

and note that $\widetilde{q}>1$, provided that $q>\frac{2d}{d+2}$. Now taking the two cases (2.51) and (2.52) together, we obtain

$$|\langle \boldsymbol{u}\otimes\boldsymbol{u},\nabla\boldsymbol{v}\rangle_\Omega| \le c\,\|\boldsymbol{u}\|^2_{\mathrm{W}^{1,q}(\Omega)}\|\boldsymbol{v}\|_{\mathrm{W}^{1,\widetilde{q}'}(\Omega)}\,, \qquad \text{provided that } q\ge\tfrac{2d}{d+2}. \tag{2.55}$$

Similarly, for the first term in (2.50) by Hölder's inequality and Sobolev embedding, one has

$$|\langle \boldsymbol{u}\otimes\boldsymbol{v},\nabla\boldsymbol{u}\rangle_\Omega| \le \|\boldsymbol{u}\|_{\mathrm{L}^{2q'}(\Omega)}\|\boldsymbol{v}\|_{\mathrm{L}^{2q'}(\Omega)}\|\boldsymbol{u}\|_{\mathrm{W}^{1,q}(\Omega)} \le c\,\|\boldsymbol{u}\|^2_{\mathrm{W}^{1,q}(\Omega)}\|\boldsymbol{v}\|_{\mathrm{W}^{1,q}(\Omega)}\,, \tag{2.56}$$

for $\boldsymbol{u},\boldsymbol{v}\in \mathrm{W}_0^{1,q}(\Omega)^d$, provided that $2q'\ge q^*$ if $q\in[1,d)$, which is equivalent to $q\in\left[\frac{3d}{d+2},d\right)$, and without restriction if $q \in [d, \infty)$. For smaller q, note that there exists an $s \in (1, \infty]$, such that $\frac{1}{q} + \frac{1}{q^*} + \frac{1}{s} = 1$, provided that $q \geq \frac{2d}{d+1}$, and hence

$$|\langle \boldsymbol{u} \otimes \boldsymbol{v}, \nabla \boldsymbol{u} \rangle_\Omega| \leq \|\boldsymbol{u}\|_{\mathrm{L}^{q^*}(\Omega)} \|\boldsymbol{v}\|_{\mathrm{L}^s(\Omega)} \|\boldsymbol{u}\|_{\mathrm{W}^{1,q}(\Omega)} \leq c \|\boldsymbol{u}\|^2_{\mathrm{W}^{1,q}(\Omega)} \|\boldsymbol{v}\|_{\mathrm{W}^{1,\widetilde{q}'}(\Omega)}, \tag{2.57}$$

and the last inequality follows from the embedding $\mathrm{W}^{1,\widetilde{q}'}(\Omega) \hookrightarrow \mathrm{L}^s(\Omega)$. Again taking the two cases (2.56) and (2.57) together one obtains

$$|\langle \boldsymbol{u} \otimes \boldsymbol{v}, \nabla \boldsymbol{u} \rangle_\Omega| \leq c \|\boldsymbol{u}\|^2_{\mathrm{W}^{1,q}(\Omega)} \|\boldsymbol{v}\|_{\mathrm{W}^{1,\widetilde{q}'}(\Omega)}, \quad \text{provided that } q \geq \tfrac{2d}{d+1}. \tag{2.58}$$

This gives an additional restriction on the parameter range of q when working with the numerical convective term.

To avoid this, note that by (2.51) and (2.56) we have that

$$\left|\widetilde{b}(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{v})\right| \leq \|\boldsymbol{u}\|^2_{\mathrm{L}^{2q'}(\Omega)} \|\boldsymbol{v}\|_{\mathrm{W}^{1,q}(\Omega)} + \|\boldsymbol{u}\|_{\mathrm{L}^{2q'}(\Omega)} \|\boldsymbol{v}\|_{\mathrm{L}^{2q'}(\Omega)} \|\boldsymbol{u}\|_{\mathrm{W}^{1,q}(\Omega)}, \tag{2.59}$$

for $\boldsymbol{u}, \boldsymbol{v} \in \mathrm{W}^{1,q}_0(\Omega)^d \cap \mathrm{L}^{2q'}(\Omega)^d$ without any restriction on q. This motivates the use of a regularising term giving additional $\mathrm{L}^{2q'}$-integrability as long as the numerical modification of the convective term is present. For the regularisation let us introduce the notation

$$\mathrm{X}(\Omega) := \mathrm{W}^{1,q}_0(\Omega)^d \cap \mathrm{L}^{2q'}(\Omega)^d, \qquad \text{with } \|\cdot\|_{\mathrm{X}(\Omega)} := \|\cdot\|_{\mathrm{W}^{1,q}(\Omega)} + \|\cdot\|_{\mathrm{L}^{2q'}(\Omega)} \tag{2.60}$$

and denote by $\mathrm{X}_{\mathrm{div}}(\Omega)$ the subspace of solenoidal functions in $\mathrm{X}(\Omega)$.

The Unsteady Case

For the unsteady problem the space $\mathrm{L}^\infty(0, T; \mathrm{L}^2(\Omega)^d) \cap \mathrm{L}^q(0, T; \mathrm{W}^{1,q}_0(\Omega)^d)$ is the natural function space for a weak solution, and by Corollary 2.5 this space embeds continuously into $\mathrm{L}^{\frac{q(d+2)}{d}}(Q)^d$, if $q \geq \frac{2d}{d+2}$ Furthermore, by Corollary 2.6, if $q \geq \frac{3d}{d+2}$, this space embeds continuously into the space $\mathrm{L}^\nu(0, T; \mathrm{L}^{2q'}(\Omega)^d)$, with ν as defined in (2.13).

By Hölder's inequality we obtain for $\boldsymbol{v} \in \mathrm{L}^q(0, T; \mathrm{W}^{1,q}_0(\Omega)^d)$ that

$$\|\langle \boldsymbol{u} \otimes \boldsymbol{u}, \nabla \boldsymbol{v} \rangle_\Omega\|_{\mathrm{L}^1((0,T))} \leq \|\boldsymbol{u}\|^2_{\mathrm{L}^{2q'}(Q)} \|\nabla \boldsymbol{v}\|_{\mathrm{L}^q(Q)} \leq c \|\boldsymbol{u}\|^2_{\mathrm{L}^{\frac{q(d+2)}{d}}(Q)} \|\nabla \boldsymbol{v}\|_{\mathrm{L}^q(Q)}. \tag{2.61}$$

provided that $2q' \geq \frac{q(d+2)}{d}$, which is equivalent to $q \geq \frac{3d+2}{d+2}$.

If q is smaller, then the test function space has to be chosen smaller: If $q \geq \frac{3d}{d+2}$ the above mentioned interpolation result and Hölder's inequality yield that

$$\|\langle \boldsymbol{u} \otimes \boldsymbol{u}, \nabla \boldsymbol{v} \rangle_\Omega\|_{\mathrm{L}^1((0,T))} \leq \|\boldsymbol{u}\|^2_{\mathrm{L}^\nu(0,T;\mathrm{L}^{2q'}(\Omega))} \|\nabla \boldsymbol{v}\|_{\mathrm{L}^{(\nu/2)'}(0,T;\mathrm{L}^q(\Omega))}, \tag{2.62}$$

for $\boldsymbol{v} \in \mathrm{L}^{(\nu/2)'}(0, T; \mathrm{W}^{1,q}_0(\Omega)^d)$ provided that $\nu \geq 2$, which is equivalent to $q \geq q_d$, as defined in (2.14).

If $q \geq \frac{2d}{d+2}$, we have by interpolation and by Hölder's inequality that

$$\|\langle \boldsymbol{u} \otimes \boldsymbol{u}, \nabla \boldsymbol{v} \rangle_\Omega\|_{\mathrm{L}^1((0,T))} \leq \|\boldsymbol{u}\|^2_{\mathrm{L}^{\frac{q(d+2)}{d}}(Q)} \|\nabla \boldsymbol{v}\|_{\mathrm{L}^{\left(\frac{q(d+2)}{2d}\right)'}(Q)}, \tag{2.63}$$

for all $\boldsymbol{v} \in \mathrm{L}^\infty(0, T; \mathrm{W}^{1,\infty}(\Omega)^d)$, since under this condition we have that $\frac{q(d+2)}{2d} \geq 1$. Let us denote

$$\hat{q} := \max\left(\left(\frac{q(d+2)}{2d}\right)', q\right) = \max\left(\frac{q(d+2)}{q(d+2)-2d}, q\right), \tag{2.64}$$

and note that $\hat{q} < \infty$ for any $q > \frac{2d}{d+2}$. Now taking the two cases (2.61) and (2.63) together yields that

$$\|\langle \boldsymbol{u}\otimes\boldsymbol{u}, \nabla\boldsymbol{v}\rangle_\Omega\|_{\mathrm{L}^1((0,T))} \le \|\boldsymbol{u}\|^2_{\mathrm{L}^{\frac{q(d+2)}{d}}(Q)} \|\nabla\boldsymbol{v}\|_{\mathrm{L}^{\hat{q}}(Q)}, \tag{2.65}$$

for all $\boldsymbol{v} \in \mathrm{L}^{\hat{q}}(0,T;\mathrm{W}^{1,\hat{q}}_0(\Omega)^d)$, provided that $q \ge \frac{2d}{d+2}$, and the restriction on q is the same one as in the steady case.

For the modification of the convective term, i.e., the first term in (2.50), we have as in (2.61) that

$$\begin{aligned}\|\langle \boldsymbol{u}\otimes\boldsymbol{v}, \nabla\boldsymbol{u}\rangle_\Omega\|_{\mathrm{L}^1((0,T))} &\le \|\boldsymbol{u}\|_{\mathrm{L}^{2q'}(Q)} \|\boldsymbol{v}\|_{\mathrm{L}^{2q'}(Q)} \|\nabla\boldsymbol{u}\|_{\mathrm{L}^{q}(Q)}\\ &\le c\,\|\boldsymbol{u}\|_{\mathrm{L}^{\frac{q(d+2)}{d}}(Q)} \|\boldsymbol{v}\|_{\mathrm{L}^{\frac{q(d+2)}{d}}(Q)} \|\nabla\boldsymbol{u}\|_{\mathrm{L}^{q}(Q)},\end{aligned} \tag{2.66}$$

provided that $q \ge \frac{3d+2}{d+2}$.

For smaller q, if $q \ge \frac{3d}{d+2}$ there exists an $r \in (1,\infty]$ such that $\frac{1}{\nu}+\frac{1}{q}+\frac{1}{r} = 1$, provided that $\nu \ge q'$, which is equivalent to $q \ge \widetilde{q}_d$, defined in (2.15). Hence, we have that

$$\|\langle \boldsymbol{u}\otimes\boldsymbol{v}, \nabla\boldsymbol{u}\rangle_\Omega\|_{\mathrm{L}^1((0,T))} \le \|\boldsymbol{u}\|_{\mathrm{L}^{\nu}(0,T;\mathrm{L}^{2q'}(\Omega))} \|\boldsymbol{v}\|_{\mathrm{L}^{r}(0,T;\mathrm{L}^{2q'}(\Omega))} \|\nabla\boldsymbol{u}\|_{\mathrm{L}^{q}(Q)}, \tag{2.67}$$

provided that $q \ge \widetilde{q}_d > \frac{3d}{d+2}$.

Furthermore, there exists an $s \in (1,\infty]$, such that $\frac{1}{q} + \frac{d}{q(d+2)} + \frac{1}{s} = 1$, provided that $q \ge \frac{2(d+1)}{d+2}$. Using this with Hölder's inequality we have that

$$\|\langle \boldsymbol{u}\otimes\boldsymbol{v}, \nabla\boldsymbol{u}\rangle_\Omega\|_{\mathrm{L}^1((0,T))} \le \|\boldsymbol{u}\|_{\mathrm{L}^{\frac{q(d+2)}{d}}(Q)} \|\boldsymbol{v}\|_{\mathrm{L}^{s}(Q)} \|\nabla\boldsymbol{u}\|_{\mathrm{L}^{q}(Q)}. \tag{2.68}$$

We note in passing that this restriction is slightly stronger than the corresponding restriction $q \ge \frac{2d}{d+1}$ in the steady case.

As in the steady case, incorporating a regularisation enables us to relax the restriction $q \ge \frac{2(d+1)}{d+2}$. By (2.61) and (2.66) it follows that

$$\left\|\widetilde{b}(\boldsymbol{u},\boldsymbol{u},\boldsymbol{v})\right\|_{\mathrm{L}^1((0,T))} \le \|\boldsymbol{u}\|^2_{\mathrm{L}^{2q'}(Q)} \|\nabla\boldsymbol{v}\|_{\mathrm{L}^{q}(Q)} + \|\boldsymbol{u}\|_{\mathrm{L}^{2q'}(Q)} \|\boldsymbol{v}\|_{\mathrm{L}^{2q'}(Q)} \|\nabla\boldsymbol{u}\|_{\mathrm{L}^{q}(Q)} \tag{2.69}$$

for $\boldsymbol{u},\boldsymbol{v} \in \mathrm{L}^q(0,T;\mathrm{W}^{1,q}_0(\Omega)^d) \cap \mathrm{L}^{2q'}(Q)^d$ without any restriction on $q \in (1,\infty)$. For the regularisation let us introduce

$$\mathrm{X}(Q) := \mathrm{L}^q(0,T;\mathrm{W}^{1,q}_0(\Omega)^d) \cap \mathrm{L}^{2q'}(Q)^d, \tag{2.70}$$

with norm $\|\cdot\|_{\mathrm{X}(Q)} := \|\cdot\|_{\mathrm{L}^q(0,T;\mathrm{W}^{1,q}(\Omega))} + \|\cdot\|_{\mathrm{L}^{2q'}(Q)}$ and denote by $\mathrm{X}_{\mathrm{div}}(Q)$ the subspace of solenoidal functions in $\mathrm{X}(Q)$.

Evidently, in both the steady and the unsteady case the additional restriction on q arises from the modification of the trilinear form b. Recall that this modification was introduced in order to reinstate the skew symmetry of b, which is lost when approximating the pointwise divergence-free solution by discretely divergence-free finite element functions.

2.2.3. Approximation in Time

To obtain a fully discrete approximate problem we will use a time-stepping based on backward finite differences in time in Chapter 5.

Here we want to introduce the notation and some basic results on piecewise constant and continuous, piecewise affine interpolants in time. Furthermore, we investigate the stability properties of a L^2-projection mapping to $\mathbb{V}^n_{\mathrm{div}}$, which can be used to prove uniform bounds on the sequence of approximate time derivatives in the discretisation limit, see Chapter 1, Subsection 1.3.1.2.

2.2.3.1 Time Discretization

For the purpose of time discretisation, 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 $\{t_i^l\}_{i\in\{0,\dots,l\}}$, where $t_i^l := i\delta_l$, 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

$$\mathrm{d}_t\varphi_i := \frac{1}{\delta_l}\left(\varphi_i - \varphi_{i-1}\right), \quad i \in \{1,\dots,l\}. \tag{2.71}$$

Furthermore, for $l \in \mathbb{N}$ we denote by $\mathbb{P}_0^l(0,T;\mathrm{X})$ 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 $\mathbb{P}_1^l(0,T;\mathrm{X})$ 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 $\overline{\varphi} \in \mathbb{P}_0^l(0,T;\mathrm{X})$ and $\widetilde{\varphi} \in \mathbb{P}_1^l(0,T;\mathrm{X})$ of $\{\varphi_i\}_{i\in\{0,\dots,l\}}$ be defined by

$$\overline{\varphi}(t) := \varphi_i, \qquad \text{for } t \in (t_{i-1},t_i],\ i \in \{1,\dots,l\}, \tag{2.72}$$

$$\widetilde{\varphi}(t) := \varphi_i\frac{t-t_{i-1}}{\delta_l} + \varphi_{i-1}\frac{t_i - t}{\delta_l}, \qquad \text{for } t \in [t_{i-1},t_i],\ i \in \{1,\dots,l\}, \tag{2.73}$$

so that $\overline{\varphi}, \widetilde{\varphi}, \partial_t\widetilde{\varphi} \in \mathrm{L}^\infty(0,T;\mathrm{X})$. Choosing the representative $\partial_t\widetilde{\varphi} \in \mathbb{P}_0^l(0,T;\mathrm{X})$, for $t \in (t_{i-1},t_i]$ we have $\partial_t\widetilde{\varphi}(t) = \mathrm{d}_t\varphi_i$ and

$$\overline{\varphi}(t) - \widetilde{\varphi}(t) = (t_i - t)\partial_t\widetilde{\varphi}(t). \tag{2.74}$$

Furthermore, note that one has

$$\|\overline{\varphi}\|_{\mathrm{L}^\infty(0,T;\mathrm{X})} = \max_{i\in\{1,\dots,l\}}\|\varphi_i\|_{\mathrm{X}}, \qquad \|\overline{\varphi}\|^p_{\mathrm{L}^p(0,T;\mathrm{X})} = \delta_l\sum_{i=1}^{l}\|\varphi_i\|^p_{\mathrm{X}}, \quad \text{for } p \in [1,\infty), \tag{2.75}$$

$$\|\widetilde{\varphi}\|_{\mathrm{L}^\infty(0,T;\mathrm{X})} = \max_{i\in\{0,\dots,l\}}\|\varphi_i\|_{\mathrm{X}}, \qquad \|\widetilde{\varphi}\|^p_{\mathrm{L}^p(0,T;\mathrm{X})} \le c(p)\delta_l\sum_{i=0}^{l}\|\varphi_i\|^p_{\mathrm{X}}, \quad \text{for } p \in [1,\infty), \tag{2.76}$$

where $0 < c(p) \le 1$ by the Riesz–Thorin interpolation theorem (cf. [BL76, 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 $\{t_0, \ldots, t_l\}$, for $l \in \mathbb{N}$, by

$$\psi_i := \fint_{t_{i-1}}^{t_i} \psi(t,\cdot)\,\mathrm{d}t \in \mathrm{X}, \quad i \in \{1,\ldots,l\}. \tag{2.77}$$

Considering the piecewise constant interpolant $\overline{\psi}$ of the set of values $\{\psi_i\}_{i\in\{1,\ldots,l\}}$, with ψ_i defined by (2.77), it follows by Jensen's inequality that

$$\|\overline{\psi}\|_{\mathrm{L}^p(0,T;\mathrm{X})} \le \|\psi\|_{\mathrm{L}^p(0,T;\mathrm{X})} \qquad \text{for all } p \in [1,\infty], \tag{2.78}$$

and, for any $p \in [1,\infty)$,

$$\overline{\psi} \to \psi \quad \text{ strongly in } \mathrm{L}^p(0,T;\mathrm{X}), \quad \text{ as } l \to \infty, \tag{2.79}$$

thanks to 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 $p \in [1,\infty)$, which, together with (2.78) implies (2.79) for $p \in [1,\infty)$.

To simplify the notation we will 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 $i \in \{1,\ldots,l\}$.

2.2.3.2 L^2-Projector to $\mathbb{V}^n_{\mathrm{div}}$

The L^2-projection mapping to $\mathbb{V}^n_{\mathrm{div}}$ can be used in the unsteady problem in two ways: It is useful to project the initial data $\boldsymbol{u}_0 \in \mathrm{L}^2_{\mathrm{div}}(\Omega)^d$ to $\mathbb{V}^n_{\mathrm{div}}$. And it can be used to obtain uniform estimates on the approximate time derivatives in the discretisation limit to apply the Aubin–Lions compactness lemma. However, as outlined in Subsection 1.3.1.2, such estimates require stability properties of Π^n, which we investigate here.

Let us introduce the projector onto $\mathbb{V}^n_{\mathrm{div}}$, given by

$$\begin{aligned} P^n_{\mathrm{div}}\colon \mathrm{L}^2(\Omega)^d \to \mathbb{V}^n_{\mathrm{div}}, \quad & \text{and for } \boldsymbol{v} \in \mathrm{L}^2(\Omega)^d, \\ \langle P^n_{\mathrm{div}}\boldsymbol{v}, \boldsymbol{V}\rangle_\Omega = (\boldsymbol{v}, \boldsymbol{V})_\Omega \quad & \text{for all } \boldsymbol{V} \in \mathbb{V}^n_{\mathrm{div}}. \end{aligned} \tag{2.80}$$

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

$$\|P^n_{\mathrm{div}}\boldsymbol{v}\|_{\mathrm{L}^2(\Omega)} \le \|\boldsymbol{v}\|_{\mathrm{L}^2(\Omega)}\,. \tag{2.81}$$

By the approximation property of Π^n in (2.39) one can show that

$$P^n_{\mathrm{div}}\boldsymbol{w} \to \boldsymbol{w} \quad \text{ strongly in } \mathrm{L}^2(\Omega)^d, \text{ as } n \to \infty, \tag{2.82}$$

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

Since $\mathbb{V}^n_{\mathrm{div}}$ is not polynomial and in general we do not have much information about it, we cannot expect to obtain stability properties as in the case of L^2-projections mapping to spaces of continuous, piecewise polynomial functions. For such finite element spaces in [DDW74, CT87, EJ95, Bom06] the authors show $\mathrm{W}^{1,p}$-stability of the L^2-projection under weaker and weaker mesh grading conditions. See also [BY14, GHS16] for $\mathrm{W}^{1,2}$-stability and even weaker mesh conditions.

Here we derive the stability properties of P^n_{div} from the approximation properties of Π^n according to Lemma 2.25. For this we require quasiuniformity of the triangulation so that inverse estimates are available, which we recall now.

Inverse Estimates

In order to deduce stability of the L^2-projection P^n_{div} we require global inverse estimates. If $p, q \in [0,1]$ and $s, \mu \in \mathbb{N}$ such that the embedding $\mathrm{W}^{s,q}(\Omega)^d \hookrightarrow \mathrm{W}^{\mu,p}(\Omega)^d$ is continuous, which is equivalent to $s - \frac{d}{q} \geq \mu - \frac{d}{p}$, then on the reference simplex one has that

$$\|\widehat{\boldsymbol{v}}\|_{\mathrm{W}^{\mu,p}(\widehat{K})} \leq c\, \|\widehat{\boldsymbol{v}}\|_{\mathrm{W}^{s,q}(\widehat{K})} \qquad \text{for all } \widehat{\boldsymbol{v}} \in \mathrm{W}^{s,q}(\widehat{K})^d,$$

If $s - \frac{d}{q} \leq \mu - \frac{d}{p}$ and so the embedding does not hold, this inequality is in general wrong. But if we consider functions in a finite-dimensional subspace $\widehat{\mathbb{P}} \subset \mathrm{W}^{\mu,p}(\widehat{K})^d \cap \mathrm{W}^{s,q}(\widehat{K})^d$, then we can obtain the estimate

$$\left\|\widehat{\boldsymbol{V}}\right\|_{\mathrm{W}^{\mu,p}(\widehat{K})} \leq c \left\|\widehat{\boldsymbol{V}}\right\|_{\mathrm{W}^{s,q}(\widehat{K})} \qquad \text{for all } \widehat{\boldsymbol{V}} \in \widehat{\mathbb{P}},$$

with constant c depending on the dimension of $\widehat{\mathbb{P}}$, the reason being that all norms on finite-dimensional spaces are equivalent. With careful scaling one can deduce that

$$\|\boldsymbol{V}\|_{\mathrm{W}^{s,q}(K)} \leq c h_K^{s-\mu+d\left(\frac{1}{p}-\frac{1}{q}\right)} \|\boldsymbol{V}\|_{\mathrm{W}^{\mu,p}(K)} \qquad \text{for all } \boldsymbol{V} \circ \mathsf{F}_K^{-1} \in \widehat{\mathbb{P}}, \tag{2.83}$$

and the exponent is negative, see [BS08, Lem. (4.5.3)]. Since this is the interesting case, in order to deduce a global version, quasiuniformity of the triangulation as in Assumption 2.19 is assumed to bound $h_K^\alpha \leq c h_n^\alpha$, for $\alpha < 0$.

Lemma 2.31 (Global Inverse Estimates, [BS08, Thm. (4.5.11)])**.**
Let $\mathbb{V}^n \subset \mathrm{W}^{1,\infty}_0(\Omega)^d$ be as defined in (2.30), assume that $\{\mathcal{T}_n\}_{n\in\mathbb{N}}$ is quasiuniform as defined in Assumption 2.19 and let $p, q \in [1,\infty]$ and $s, \mu \in \mathbb{N}_0$ be given such that $0 \leq s \leq \mu$ and $\widehat{\mathbb{P}}_{\mathbb{V}} \subset \mathrm{W}^{\mu,p}(\widehat{K})^d \cap \mathrm{W}^{s,q}(\widehat{K})^d$. Then there exists a constant $c > 0$ (independent of n) such that

$$\|\boldsymbol{V}\|_{\mathrm{W}^{\mu,p}(\Omega)} \leq c h_n^{s-\mu+d\min\left(0,\frac{1}{p}-\frac{1}{q}\right)} \|\boldsymbol{V}\|_{\mathrm{W}^{s,q}(\Omega)}, \tag{2.84}$$

for all $\boldsymbol{V} \in \mathbb{V}^n$ and all $n \in \mathbb{N}$.

When summing over $K \in \mathcal{T}_n$ to show the global inverse estimate, the minimum arises since the embedding $\ell_q \hookrightarrow \ell_p$ is continuous only for $1 \leq q \leq p < \infty$. Otherwise merely Hölder's inequality is at ones disposal, which results in the loss of a factor $h_n^{\frac{1}{p}-\frac{1}{q}} < 1$.
Now we are in the position to show the stability results on P^n_{div}. Note that the corresponding (approximation) result was obtained in [Car07] for the case $p \geq \frac{3d+2}{d+2} \geq 2$ using approximation properties of a divergence-preserving projection operator corresponding to Lemma 2.25 with $p = 2$.

Lemma 2.32 (Stability of P^n_{div})**.**
Let $d \in \{2,3\}$ and $p \in [1,\infty]$ be given and let $\beta = \beta(d,p) \in \{1,2,3\}$ be minimal, such that the embedding $\mathrm{W}^{\beta,2}(\Omega) \hookrightarrow \mathrm{W}^{1,p}(\Omega)$ is continuous. Then define the space

$$\mathrm{Y}_p \coloneqq \mathrm{W}^{1,p}_{0,\mathrm{div}}(\Omega)^d \cap \mathrm{W}^{\beta,2}(\Omega)^d, \qquad \text{with } \|\cdot\|_{\mathrm{Y}_p} = \|\cdot\|_{\mathrm{W}^{1,p}(\Omega)} + \|\cdot\|_{\mathrm{W}^{\beta,2}(\Omega)}. \tag{2.85}$$

Let $\{\mathcal{T}_n\}_{n\in\mathbb{N}}$ satisfy Assumptions 2.18 and 2.19 and let $(\mathbb{V}^n, \mathbb{Q}^n)$ be defined as in (2.30), (2.31). Further, assume that there exists $r \in \mathbb{N}$ such that $\widehat{\mathcal{P}}_r^d \subset \widehat{\mathbb{P}}_{\mathbb{V}}$ and such that $r + 1 \geq \beta$. Assume that there exists a projection operator Π^n satisfying Assumption 2.23 (i)–(iia).

Then, there exists a constant $c > 0$ such that

$$\|P^n_{\mathrm{div}}\boldsymbol{v}\|_{\mathrm{W}^{1,p}(\Omega)} \leq c\,\|\boldsymbol{v}\|_{\mathrm{Y}_p}\,, \qquad \textit{for any } \boldsymbol{v} \in \mathrm{Y}_p \textit{ and any } n \in \mathbb{N}. \tag{2.86}$$

Proof. Let $\boldsymbol{v} \in \mathrm{Y}_p \subset \mathrm{L}^2(\Omega)^d$. Because P^n_{div} maps to $\mathbb{V}^n_{\mathrm{div}} \subset \mathrm{W}^{1,\infty}_0(\Omega)^d \subset \mathrm{W}^{1,p}_0(\Omega)^d$ by Poincaré's inequality it is enough to show that

$$\|\nabla(P^n_{\mathrm{div}}\boldsymbol{v})\|_{\mathrm{L}^p(\Omega)} \leq c\,\|\boldsymbol{v}\|_{\mathrm{Y}_p}\,.$$

Since $\boldsymbol{v}$ is divergence-free by the definition of Y_p and Π^n is divergence-preserving in the sense of Assumption 2.23 (i) we have that $\Pi^n\boldsymbol{v} \in \mathbb{V}^n_{\mathrm{div}}$. Consequently, since P^n_{div} is a linear projection operator we add and subtract $\Pi^n\boldsymbol{v} = P^n_{\mathrm{div}}(\Pi^n\boldsymbol{v})$ and $\boldsymbol{v}$ to obtain that

$$\|\nabla(P^n_{\mathrm{div}}\boldsymbol{v})\|_{\mathrm{L}^p(\Omega)} \leq \|\nabla P^n_{\mathrm{div}}(\boldsymbol{v} - \Pi^n\boldsymbol{v})\|_{\mathrm{L}^p(\Omega)} + \|\nabla(\Pi^n\boldsymbol{v} - \boldsymbol{v})\|_{\mathrm{L}^p(\Omega)} + \|\nabla\boldsymbol{v}\|_{\mathrm{L}^p(\Omega)}\,. \tag{2.87}$$

The difficulty lies in estimating the first term, since the third one is fine already and the second one can be estimated by means of the approximability property of Π^n in Lemma 2.25 with $\mu = s = 1 < r + 1$ for $r \in \mathbb{N}$, as

$$\|\nabla(\Pi^n\boldsymbol{v} - \boldsymbol{v})\|_{\mathrm{L}^p(\Omega)} \leq c\,|\boldsymbol{v}|_{\mathrm{W}^{1,p}(\Omega)}\,. \tag{2.88}$$

For the first term in (2.87) we want to use the stability of P^n_{div} in $\mathrm{L}^2(\Omega)$, see (2.81). Since $P^n_{\mathrm{div}}(\boldsymbol{v} - \Pi^n\boldsymbol{v}) \in \mathbb{V}^n$ we can use the global inverse estimate in Lemma 2.31, with $\mu = 1$, $s = 0$, $p = p$ and $q = 2$ from which it follows that

$$\begin{aligned}\|\nabla P^n_{\mathrm{div}}(\boldsymbol{v} - \Pi^n\boldsymbol{v})\|_{\mathrm{L}^p(\Omega)} &\leq c h_n^{-1+d\min\left(0,\frac{1}{p}-\frac{1}{2}\right)}\|P^n_{\mathrm{div}}(\boldsymbol{v} - \Pi^n\boldsymbol{v})\|_{\mathrm{L}^2(\Omega)} \\ &\overset{(2.81)}{\leq} c h_n^{-1+d\min\left(0,\frac{1}{p}-\frac{1}{2}\right)}\|\boldsymbol{v} - \Pi^n\boldsymbol{v}\|_{\mathrm{L}^2(\Omega)}\,,\end{aligned} \tag{2.89}$$

assuming quasiuniformity of the triangulations and we have also used the stability of P^n_{div}. Finally we apply the global approximation property in Lemma 2.25 with $\mu = 0$, $p = 2$, which shows that

$$\begin{aligned}\|\nabla P^n_{\mathrm{div}}(\boldsymbol{v} - \Pi^n\boldsymbol{v})\|_{\mathrm{L}^p(\Omega)} &\overset{(2.89)}{\leq} c h_n^{-1+d\min\left(0,\frac{1}{p}-\frac{1}{2}\right)}\|\boldsymbol{v} - \Pi^n\boldsymbol{v}\|_{\mathrm{L}^2(\Omega)} \\ &\leq c h_n^{s-1+d\min\left(0,\frac{1}{p}-\frac{1}{2}\right)}\|\boldsymbol{v}\|_{\mathrm{W}^{s,2}(\Omega)}\,,\end{aligned} \tag{2.90}$$

for all $s \in \{1, \ldots, r+1\}$. The only thing left to do is to choose $s = \beta$ (assuming that $r + 1 \geq \beta$) such that the exponent is non-negative. Let us distinguish the cases $p \leq 2$ and $p > 2$

If $p \leq 2$ we have that $\min\left(0, \frac{1}{p} - \frac{1}{2}\right) = 0$ and hence $s = \beta = 1 \leq r + 1$ suffices and does not impose any extra restriction on r. Note that in this case we have $\mathrm{W}^{1,2}(\Omega) \hookrightarrow \mathrm{W}^{1,p}(\Omega)$ which shows the claim.

If $p > 2$ we have that $\min\left(0, \frac{1}{p} - \frac{1}{2}\right) = \frac{1}{p} - \frac{1}{2}$ and we thus require that $s = \beta$ are such that

$$s - 1 + d\left(\frac{1}{p} - \frac{1}{2}\right) \geq 0. \tag{2.91}$$

This is the case, if $s = \beta \geq 1$ and the embedding $\mathrm{W}^{\beta,2}(\Omega) \hookrightarrow \mathrm{W}^{1,p}(\Omega)$ is continuous. In this case we require that $\beta = s \leq r + 1$, and impose this restriction on $\mathbb{V}^n$. $\square$

Let us specify the values of β depending on d and p and also the corresponding assumptions on r (depending on β). For $p > 2$ we can determine β and see that $\beta = 2$, provided that $p \leq 2_*$, which is the case for

$$p \in \begin{cases} [1, \infty) & \text{if } d = 2, \\ [1, 6] & \text{if } d = 3. \end{cases}$$

However, since $r \geq 1$ and $r + 1 \geq 2 = \beta$ this does not require any extra assumption on $\widehat{\mathbb{P}}_{\mathbb{V}}$ and hence on $\mathbb{V}^n$. Since we do not require $p = \infty$ for the application the only situation, for which we need need $\beta = 3$ is if $d = 3$ and $p > 6$. In this case we assume that $r \geq 2$, which restricts the choice of finite element spaces slightly.

If $d = 3$ the lemma is applied for $p > 6$, then the restriction on r excludes the first order finite element spaces introduced in Examples 2.26 and 2.28. More specifically this affects the Bernardi–Raugel element for $r = 1$ and $d = 3$, the MINI element for $d = 3$ and the family Guzmán–Neilan elements of exactly divergence-free functions, for $d = 3$ ($r = 1$).

CHAPTER 3

Graph Representation of the Constitutive Law

3.1. Introduction

The implicit constitutive relation, which determines the fluid behaviour, is at the core of the problem we want to consider. In Section 1.1 we have introduced the general thermodynamic framework for fluids flows. Here we will be concerned with the mathematical formulation of implicit constitutive relations. Furthermore we aim for approximations thereof, which allow us to show convergence of numerical approximation schemes in the subsequent chapters.

3.1.1. Mathematical Description of Constitutive Relations

In Chapter 1 we have introduced constitutive relations, which relate the deviatoric part of the symmetric stress-tensor $\mathbf{S}$ and the symmetric velocity gradient $\mathbf{D}\boldsymbol{u}$. Such a relation is imposed additionally to the balance laws in order to obtain a closed system of equations.

Explicit Constitutive Laws

Explicit constitutive relations are relations for which there exists a (single-valued) continuous function $\mathcal{S}$ such that $\mathbf{S} = \mathcal{S}(\mathbf{D}\boldsymbol{u})$, see (1.5b) for examples. If the constitutive relation depends on the point $\boldsymbol{z}$ in the domain M, it is assumed that $\mathbf{S} = \mathcal{S}(\boldsymbol{z}, \mathbf{D}\boldsymbol{u})$ for a Carathéodory function $\mathcal{S}$. This means that $\mathcal{S}$ is measurable in the first argument for each fixed value of the second argument and continuous in the second one for almost every value of the first argument. Both assumptions on $\mathcal{S}$ are sufficient to prove existence of a weak solution to an approximate problem, e.g., by a Galerkin approximation, see Section 1.2. For $q \in (1, \infty)$ the a priori estimates allow one to obtain weakly converging subsequences $\{\mathbf{D}\boldsymbol{u}^N\}_{N\in\mathbb{N}}$ and $\{\mathbf{S}^N\}_{N\in\mathbb{N}}$ with $\mathbf{S}^N := \mathcal{S}(\cdot, \mathbf{D}\boldsymbol{u}^N)$ such that

$$\mathbf{D}\boldsymbol{u}^N \rightharpoonup \mathbf{D}\boldsymbol{u} \quad \text{weakly in } \mathrm{L}^q(M)^{d\times d} \quad \text{and} \quad \mathbf{S}^N \rightharpoonup \overline{\mathbf{S}} \quad \text{weakly in } \mathrm{L}^{q'}(M)^{d\times d},$$

as $N \to \infty$, and one has to identify the constitutive relation $\overline{\mathbf{S}} = \mathcal{S}(\cdot, \mathbf{D}\boldsymbol{u})$. In case the function $\mathcal{S}$ is strictly monotone in the second argument, this identification can be done by means of Vitali's convergence theorem, see, e.g., [DMM98] for divergence-form nonlinear elliptic systems and [FMS03, DMS08] for fluid equations. In Subsection 2.1.4 we have seen the argument for the case without $\boldsymbol{z}$-dependence, but including a truncation.

If $\mathcal{S}$ is merely monotone in the second argument, Minty's trick allows one to identify the constitutive relation, provided that

$$\limsup_{N\to\infty} \left\langle \boldsymbol{\mathcal{S}}(\cdot, \mathbf{D}\boldsymbol{u}^N), \mathbf{D}\boldsymbol{u}^N \right\rangle_M \leq \left\langle \overline{\mathbf{S}}, \mathbf{D}\boldsymbol{u} \right\rangle_M . \tag{3.1}$$

The proof relies on the monotonicity and continuity of $\boldsymbol{\mathcal{S}}$ in the second argument and was presented in Subsection 1.2.3 in the case where $\boldsymbol{\mathcal{S}}$ does not depend on $\boldsymbol{z} \in M$. This result dates back to [Min63] and since then was applied in various contributions. For example in [Lad69] and [Lio69, Ch. 2.5] existence of weak solutions for fluid equations with q-growth was shown in the admissible range. By means of truncation (3.1) can be verified also in situations, when admissibility of weak solutions as test function is lost. If the respective truncation requires localisation, then a local version of (3.1) can be shown, see [Wol07, Lem. A.2] for a localised Minty type convergence result.

Implicit Constitutive Laws

In Section 1.1 we have introduced implicit constitutive relations of the form $\mathbf{G}(\boldsymbol{z}, \mathbf{D}\boldsymbol{u}, \mathbf{S}) = \mathbf{0}$, where $\mathbf{G}$ is a continuous tensor-valued function, which may depend on $\boldsymbol{z} \in M$. Furthermore, the implicit relation was assumed to be encoded by a maximal monotone graph $\mathcal{A} \subset \mathbb{R}^{d\times d}_{\mathrm{sym}} \times \mathbb{R}^{d\times d}_{\mathrm{sym}}$ in the pointwise sense

$$\mathbf{G}(\boldsymbol{z}, \mathbf{D}\boldsymbol{u}(\boldsymbol{z}), \mathbf{S}(\boldsymbol{z})) = \mathbf{0} \quad \Leftrightarrow \quad (\mathbf{D}(\boldsymbol{z}), \mathbf{S}(\boldsymbol{z})) \in \mathcal{A}(\boldsymbol{z}), \qquad \text{for } \boldsymbol{z} \in M. \tag{3.2}$$

Note that the maximal monotone set $\mathcal{A}(\boldsymbol{z})$ can be identified with the graph of a set-valued maximal monotone function $\mathcal{S}(\boldsymbol{z}, \cdot)$.

The existence proof of solutions to an approximate problem requires an approximation of $\mathcal{S}$ by Carathéodory functions, which can be obtained in various ways. Some examples in the literature, which we will consider in this chapter, are the following: Starting from a possibly discontinuous selection $\boldsymbol{\mathcal{S}}^\star$ of the set-valued function $\mathcal{S}$ one can introduce a sequence of Carathéodory functions by smoothing $\boldsymbol{\mathcal{S}}^\star$ by means of a mollification. This is the path taken in, e.g., [GZG07, BGMS09, BGMS12, DKS13a]. Alternatively, under additional assumptions on the selection one can approximate $\boldsymbol{\mathcal{S}}^\star$ by affine interpolation in the neighbourhood of jumps. Another class of approximations can be obtained by approximating the characterising Carathéodory contraction, which can be identified with any maximal monotone graph. This approach can be found in [FMT04, Sec. 3]. Finally, a generalised Yosida approximation, which is well-known and investigated in the theory of monotone operators, provides a way to approximate $\mathcal{A}$, see for example [FMT04, Sec. 4]. More details and references will be given in the course of the present chapter.

The main challenge of any convergence proof of approximate solutions to a weak solution of the respective problem is the identification of the implicit relation, for which the lack of continuity of the constitutive relation has to be overcome. For (generalised) strictly monotone constitutive relations arguments based on Young measures and Vitali's theorem have been used to identify the implicit relation satisfied by the limiting functions, see for example [MRS05, GMS07, GZG07, BGMS09, DKS13a]. Subsequently, for the larger class of only monotone constitutive relations, in [BGMS12] a convergence lemma of Minty type was proved (and reproved in [BM16]), which simplifies the argument considerably: it allows one to conclude that $(\mathbf{D}\boldsymbol{u}, \overline{\mathbf{S}}) \in \mathcal{A}(\cdot)$ a.e. in the domain M, provided that there are approximating sequences such that $(\mathbf{D}\boldsymbol{u}^N, \mathbf{S}^N) \in \mathcal{A}(\cdot)$ a.e. in M, $\mathbf{D}\boldsymbol{u}^N \rightharpoonup \mathbf{D}\boldsymbol{u}$ weakly in $\mathrm{L}^q(M)^{d\times d}$, $\mathbf{S}^N \rightharpoonup \overline{\mathbf{S}}$

weakly in $\mathrm{L}^{q'}(M)^{d\times d}$, as $N\to\infty$, and that

$$\limsup_{N\to\infty} \langle \mathbf{S}^N, \mathbf{D}\boldsymbol{u}^N\rangle_M \le \left\langle \overline{\mathbf{S}}, \mathbf{D}\boldsymbol{u}\right\rangle_M . \tag{3.3}$$

which is the property corresponding to (3.1). In fact, this was shown for subdomains $\widetilde{M}\subset M$, so that it can be applied also if admissibility of the weak solution is lost, see Subsection 2.1.4. Note however, that this convergence result can only be applied, if the approximate sequences satisfy the implicit constitutive relation, and in general not, if they satisfy an approximate constitutive relation.

For specific choices of approximation, the hope is to show such a convergence result assuming that $(\mathbf{D}\boldsymbol{u}^N, \mathbf{S}^N)\in\mathcal{A}^N$, where $\mathcal{A}^N$ is an approximation of $\mathcal{A}$ and the graph of a Carathéodory function, i.e., existence of the approximate problem can be shown by standard arguments. Some results in this direction can be found in [GZG07, Thm. 5.1] for graph approximations based on mollification of the selection function $\mathcal{S}^\star$ under the assumption of strict monotonicity, and in the arguments contained in [FMT04, pp. 41] for generalised Yosida approximations for problems without loss of admissibility.

3.1.2. Outline

First we introduce some of the relevant definitions and results concerning maximal monotonicity and measurability in Section 3.2. Then, in Section 3.3, we state and justify the exact assumptions we pose on the graph $\mathcal{A}$ encoding the implicit constitutive law. Furthermore we give some examples of constitutive relations, which fit into this framework. In Section 3.4 we consider approximations of $\mathcal{A}$, which allow us to show existence of the approximate problem. In Assumption 3.18 we will introduce sufficient conditions on graph approximations, which allow us to take the limit corresponding to the graph approximation first and separately from any other limits. All variants of graph approximation investigated in the course of that section satisfy those assumptions. In addition, some of the graph approximations will satisfy a convergence result of Minty type as outlined before. This makes them more powerful, since it allows us to take the graph approximation limit together with the discretisation limit in the subsequent chapters. Finally, in Section 3.5 we summarise and discuss the results of this chapter and compare the classes of graph approximation.

3.2. Preliminaries on Maximal Monotonicity and Measurability

After introducing the notation for the rest of this chapter we state some basic definitions regarding monotonicity and set-valued functions. Furthermore, we provide the reader with some of the properties and characterisations of maximal monotone functions, required for the formulation of the assumptions on $\mathcal{A}$ and the investigation of examples of graph approximation. We will only introduce the results to as much generality as required for the presentation, which means in particular maximal monotonicity of functions mapping from $\mathbb{R}^s$ to $\mathbb{R}^s$ suffices, and we point the reader to [AA99] as one of the main references; for details on the general theory of maximal monotone operators see for example [Phe97]. Since we consider $\boldsymbol{z}$-dependent graphs $\mathcal{A}$, measurability plays a role and we will recall some of the results in [CPDMD90] for later use.

Since we will use it repeatedly, let us record that for any $p\in(0,\infty)$ there exists a constant $c(p)>0$ such that

$$(a+b)^p \le c(a^p+b^p), \qquad \text{for all } a,b\in\mathbb{R}_{\ge 0}. \tag{3.4}$$

This can be shown with $c(p) = 1$ for $p \in (0,1]$, and with $c(p) = 2^{p-1}$ for $p \in (1,\infty)$.
First let us introduce the notion of a set-valued function.

Definition 3.1 (Set-valued Functions)**.**
Let X, Y be sets. For a set-valued function $g\colon X \rightrightarrows Y$*, i.e.,* $X \ni x \mapsto g(x) \subset Y$ *we denote by* $\Gamma(g) := \{(x,y) \in X \times Y\colon\ y \in g(x)\}$ *its* graph *and by* $\operatorname{dom} g := \{x \in X\colon\ g(x) \neq \emptyset\}$ *its* domain*.*
The (set-valued) inverse function $g^{-1}\colon Y \rightrightarrows X$ *is defined by* $g^{-1}(y) := \{x \in X\colon\ y \in \boldsymbol{g}(x)\}$ *and the* preimage *of a set* $A \subset Y$ *is defined as usual as* $g^{-1}(A) := \{x \in X\colon\ y \in g(x) \text{ for some } y \in A\}$*.*
If for all $x \in \operatorname{dom}(g)$ *there is a unique* $y \in Y$ *such that* $g(x) = \{y\}$*, then we call* g single-valued*.*

3.2.1. Maximal Monotonicity

Before introducing the notions of maximal monotone set-valued functions, let us state the definitions of (maximal) monotonicity and the inverse of a set.

Definition 3.2 (Maximal Monotone Set)**.**
A subset $\mathcal{G} \subset \mathbb{R}^s \times \mathbb{R}^s$*,* $s \in \mathbb{N}$*, is called* monotone*, if*

$$(\boldsymbol{y}_1 - \boldsymbol{y}_2) : (\boldsymbol{x}_1 - \boldsymbol{x}_2) \geq 0 \qquad \text{for all } (\boldsymbol{x}_1, \boldsymbol{y}_1), (\boldsymbol{x}_2, \boldsymbol{y}_2) \in \mathcal{G}.$$

A set $\mathcal{G} \subset \mathbb{R}^s \times \mathbb{R}^s$ *is called* maximal monotone*, if it is monotone and it is maximal with respect to inclusion in the set of monotone sets, which it is contained in. I.e., if for* $(\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{R}^s \times \mathbb{R}^s$ *the fact that*

$$(\boldsymbol{y} - \overline{\boldsymbol{y}}) : (\boldsymbol{x} - \overline{\boldsymbol{x}}) \geq 0 \qquad \text{for all } (\overline{\boldsymbol{x}}, \overline{\boldsymbol{y}}) \in \mathcal{G},$$

implies that $(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{G}$*.*
The inverse set *of* $\mathcal{G} \subset \mathbb{R}^s \times \mathbb{R}^s$ *is defined as*

$$\mathcal{G}^{-1} := \{(\boldsymbol{y}, \boldsymbol{x}) \in \mathbb{R}^s \times \mathbb{R}^s\colon\ (\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{G}\}.$$

Clearly, by definition $\mathcal{G} \subset \mathbb{R}^s \times \mathbb{R}^s$ is (maximal) monotone if and only if $\mathcal{G}^{-1} \subset \mathbb{R}^s \times \mathbb{R}^s$ is (maximal) monotone. Now let us give the definition of a (maximal) monotone function and state some of its properties, see, e.g., [Phe97].

Definition 3.3 (Maximal Monotone (Set-valued) Functions)**.**
A set-valued function $\mathfrak{g}\colon \mathbb{R}^s \rightrightarrows \mathbb{R}^s$ *is* (maximal) monotone*, if its graph* $\Gamma(\mathfrak{g})$ *is a (maximal) monotone set.*
If for any $(\boldsymbol{x}_1, \boldsymbol{y}_1), (\boldsymbol{x}_2, \boldsymbol{y}_2) \in \Gamma(\mathfrak{g})$ *such that* $\boldsymbol{x}_1 \neq \boldsymbol{x}_2$ *we have that* $(\boldsymbol{y}_1 - \boldsymbol{y}_2) : (\boldsymbol{x}_1 - \boldsymbol{x}_2) > 0$*, then* $\mathfrak{g}$ *is called* strictly monotone*.*

Since the graphs of $\mathfrak{g}$ and $\mathfrak{g}^{-1}$ are related via $(\Gamma(\mathfrak{g}))^{-1} = \Gamma(\mathfrak{g}^{-1})$, we have that $\mathfrak{g}$ is (maximal) monotone if and only if $\mathfrak{g}^{-1}$ is (maximal) monotone. Also it is straightforward to see that $\mathfrak{g}$ is strictly monotone if and only if $\mathfrak{g}^{-1}$ is single-valued.

Lemma 3.4 (Maximal Monotone Functions, [AA99, Prop.1.2(1), Cor.1.3(3),(4), Cor.1.4])**.**
Let $\mathfrak{g}\colon \mathbb{R}^s \rightrightarrows \mathbb{R}^s$ *be a set-valued monotone function.*

(i) If $\mathfrak{g}$ *is maximal monotone, then* $\Gamma(\mathfrak{g})$ *is closed and* $\mathfrak{g}(\boldsymbol{x})$ *is closed (possibly empty) and convex for any* $\boldsymbol{x} \in \mathbb{R}^s$*.*

(ii) *If* $\mathfrak{g}$ *is continuous and single-valued and* $\operatorname{dom}\mathfrak{g} = \mathbb{R}^s$*, then* $\mathfrak{g}$ *is maximal monotone.*
(iii) *If* $\mathfrak{g}$ *is maximal monotone and single-valued at a point* $\boldsymbol{x} \in \mathbb{R}^s$*, then* $\boldsymbol{x} \in \operatorname{int}\operatorname{dom}\mathfrak{g}$ *and* $\mathfrak{g}$ *is continuous at* $\boldsymbol{x}$*.*
(iv) *If* $\mathfrak{g}$ *is maximal monotone, then for any set* $B \subset \mathbb{R}^s$*, which is relatively compact in* $\operatorname{int}(\operatorname{dom}(\mathfrak{g}))$*, the image* $\mathfrak{g}(B)$ *is bounded.*

The following theorem due to Rockafellar and originally formulated for general maximal monotone operators, gives a sufficient condition under which the sum of maximal monotone functions is maximal monotone.

Lemma 3.5 (Sum of Maximal Monotone Functions, [Roc70, Thm. 1])**.**
Let $\mathfrak{g}_1, \mathfrak{g}_2 \colon \mathbb{R}^s \rightrightarrows \mathbb{R}^s$ *be maximal monotone set-valued functions. If we have that* $\operatorname{dom}\mathfrak{g}_1 \cap \operatorname{int}(\operatorname{dom}(\mathfrak{g}_2)) \neq \emptyset$*, then the sum* $\mathfrak{g}_1 + \mathfrak{g}_2 \colon \mathbb{R}^s \rightrightarrows \mathbb{R}^s$*, defined by*

$$(\mathfrak{g}_1 + \mathfrak{g}_2)(\boldsymbol{x}) := \{\boldsymbol{y}_1 + \boldsymbol{y}_2 \in \mathbb{R}^s \colon\ \boldsymbol{y}_1 \in \mathfrak{g}_1(\boldsymbol{x}),\ \boldsymbol{y}_2 \in \mathfrak{g}_2(\boldsymbol{x})\},$$

is a maximal monotone set-valued function.

Characterisation of Maximal Monotone Functions

Let us now present the connection between graphs of monotone (set-valued) functions and graphs of 1-Lipschitz functions, which goes back to [Min62], see [AA99] for a well-presented survey. There exists a one-to-one correspondence between the two concepts, which in the 1D case can be visualised by a rotation by $\frac{\pi}{4}$ around the origin. For $s \in \mathbb{N}$ let us denote by $\phi \colon \mathbb{R}^s \to \mathbb{R}^s$ the so-called *Cayley transformation*, see [AA99, p. 265], defined by

$$(\boldsymbol{x}, \boldsymbol{y}) \mapsto \frac{1}{\sqrt{2}}(\boldsymbol{x} + \boldsymbol{y}, -\boldsymbol{x} + \boldsymbol{y}) \quad \text{for } \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^s. \tag{3.5}$$

This is a linear isometry and its inverse ϕ^{-1} is defined by $(\boldsymbol{x}, \boldsymbol{y}) \mapsto \frac{1}{\sqrt{2}}(\boldsymbol{x} - \boldsymbol{y}, \boldsymbol{x} + \boldsymbol{y})$.

For a set $\mathcal{G} \subset \mathbb{R}^s \times \mathbb{R}^s$ let us introduce the set-valued function $\psi_{\mathcal{G}} \colon \mathbb{R}^s \rightrightarrows \mathbb{R}^s$, defined by

$$\phi(\mathcal{G}) = \Gamma(\psi_{\mathcal{G}}), \tag{3.6}$$

and vice versa for any set-valued function $\psi \colon \mathbb{R}^s \rightrightarrows \mathbb{R}^s$, let us denote by $\mathcal{G}_\psi \subset \mathbb{R}^s \times \mathbb{R}^s$ the set, such that

$$\mathcal{G}_\psi = \phi^{-1}(\Gamma(\psi)). \tag{3.7}$$

By the properties of the Cayley transformation ϕ it follows that $\mathcal{G}_{\psi_{\mathcal{G}}} = \mathcal{G}$ and $\psi_{\mathcal{G}_\psi} = \psi$, i.e., the correspondence is one-to-one, see [FMT04, Sec. 2]. Furthermore, we have the following relationships between $\mathcal{G}$ and ψ.

Lemma 3.6 (Characterising Contraction)**.**
Let a set $\mathcal{G} \subset \mathbb{R}^s \times \mathbb{R}^s$ *and a set-valued function* $\psi \colon \mathbb{R}^s \rightrightarrows \mathbb{R}^s$ *be related by* (3.6) *(or equivalently* (3.7)*). Furthermore, let the set-valued function* $\mathfrak{g} \colon \mathbb{R}^s \rightrightarrows \mathbb{R}^s$ *be such that* $\Gamma(\mathfrak{g}) = \mathcal{G}$*. Then the following hold:*
(i) $\mathfrak{g}$ *is monotone if and only if* ψ *is* 1*-Lipschitz on* $\operatorname{dom}(\psi) \subset \mathbb{R}^s$*;*
(ii) *If* (i) *holds,* $\mathfrak{g}$ *is maximal if and only if* $\operatorname{dom}(\psi) = \mathbb{R}^s$*;*
(iii) *If* ψ *is* λ*-Lipschitz for some* $\lambda \in (0, 1)$*, then* $\mathfrak{g}$ *is strictly monotone and single-valued.*

Proof. The claims (i) and (ii) are proved in [FMT04, Lem. 2.1] and also in [AA99, Prop. 1.1].

For (iii) let $(\boldsymbol{a}_1,\boldsymbol{b}_1),(\boldsymbol{a}_2,\boldsymbol{b}_2)\in\Gamma(\mathfrak{g})$, then there exist $\boldsymbol{x}_1,\boldsymbol{x}_2\in\mathbb{R}^s$, such that

$$\phi^{-1}(\boldsymbol{x}_i,\psi(\boldsymbol{x}_i))=\frac{1}{\sqrt{2}}(\boldsymbol{x}_i-\psi(\boldsymbol{x}_i),\boldsymbol{x}_i+\psi(\boldsymbol{x}_i))=(\boldsymbol{a}_i,\boldsymbol{b}_i),\quad i\in\{1,2\}.$$

By the fact that ψ is λ-Lipschitz with $\lambda\in(0,1)$ we have that

$$\begin{aligned}2(\boldsymbol{a}_1-\boldsymbol{a}_2)&:(\boldsymbol{b}_1-\boldsymbol{b}_2)\\&=(\boldsymbol{x}_1-\psi(\boldsymbol{x}_1)-(\boldsymbol{x}_2-\psi(\boldsymbol{x}_2))):(\boldsymbol{x}_1+\psi(\boldsymbol{x}_1)-(\boldsymbol{x}_2+\psi(\boldsymbol{x}_2)))\\&=\left(|\boldsymbol{x}_1-\boldsymbol{x}_2|^2-|\psi(\boldsymbol{x}_1)-\psi(\boldsymbol{x}_2)|^2\right)\geq(1-\lambda^2)\,|\boldsymbol{x}_1-\boldsymbol{x}_2|^2\geq 0,\end{aligned}\tag{3.8}$$

since $\lambda^2<1$. Now if $\boldsymbol{a}_1=\boldsymbol{a}_2$, then the left-hand side vanishes, so it follows that $\boldsymbol{x}_1=\boldsymbol{x}_2$, and hence $\psi(\boldsymbol{x}_1)=\psi(\boldsymbol{x}_2)$, which together imply that $\boldsymbol{b}_1=\boldsymbol{b}_2$. Thus $\mathfrak{g}$ is single-valued. By the analogous argument we obtain that $\mathfrak{g}^{-1}$ is single-valued, which implies that $\mathfrak{g}$ is strictly monotone. □

Subdifferentials of Convex Functions

One particular class of maximal monotone functions is given by the subdifferentials of lower semi-continuous, convex, proper functions, which was shown by Rockafellar.

Definition 3.7 (Subdifferential, [Phe97, Def. 2.7])**.**
A function $f\colon\mathbb{R}^s\to\mathbb{R}\cup\{\infty\}$ is called proper, *if there exists a point $\boldsymbol{x}\in\mathbb{R}^s$ such that $f(\boldsymbol{x})<\infty$. Furthermore, a function $f\colon\mathbb{R}^s\to\mathbb{R}\cup\{\infty\}$ is called* lower semicontinuous, *if $\{\boldsymbol{x}\in\mathbb{R}^s\colon\ f(\boldsymbol{x})\leq\lambda\}$ is closed in $\mathbb{R}^s$ for every $\lambda\in\mathbb{R}$.*
For a proper, lower semi-continuous, convex function $f\colon\mathbb{R}^s\to\mathbb{R}\cup\{\infty\}$ the subdifferential mapping *∂f at a point $\boldsymbol{x}\in\mathbb{R}^s$ such that $f(\boldsymbol{x})<\infty$ is defined by*

$$\partial f(\boldsymbol{x}):=\{\boldsymbol{x}'\in\mathbb{R}^s\colon\ \langle\boldsymbol{x}',\boldsymbol{y}\rangle\leq f(\boldsymbol{x}+\boldsymbol{y})-f(\boldsymbol{x})\ \text{ for all }\boldsymbol{y}\in\mathbb{R}^s\},$$

and $\partial f(\boldsymbol{x})=\emptyset$ otherwise.

Lemma 3.8 (Subdifferentials and Maximal Monotonicity, [Phe97, Thm. 2.15])**.**
For a proper, lower semi-continuous, convex function f, the subdifferential ∂f is a maximal monotone (set-valued) function.

3.2.2. Measurability

Due to the $\boldsymbol{z}$-dependence ($\boldsymbol{z}\in M$) of the implicit constitutive relation let us also state some results on measurability.

We denote by $\mathcal{B}(\mathbb{R}^s)$ the σ-algebra of all Borel sets in $\mathbb{R}^s$. For two measurable spaces (Z_1,Σ_1) and (Z_2,Σ_2) we say that a set-valued function $g\colon Z_1\rightrightarrows Z_2$ is $\Sigma_1-\Sigma_2$ measurable, provided that $g^{-1}(A)\in\Sigma_1$ for any $A\in\Sigma_2$. By $(Z_1\times Z_2,\Sigma_1\otimes\Sigma_2)$ we denote the product measure space with product σ-algebra $\Sigma_1\otimes\Sigma_2$. Recall that completeness of a measure space is the property that subsets of zero sets are zero sets, and a σ-finite measure space is a measure space for which the set is a countable union of measurable subsets with finite measure.

Lemma 3.9 (Measurability, [CPDMD90, Thm. 1.3])**.**
Let (Z,Σ,μ) be a measurable space such that μ is a complete σ-finite measure defined on Σ and let $n,m\in\mathbb{N}$. Let $\mathsf{F}\colon\mathrm{Z}\rightrightarrows\mathbb{R}^n\times\mathbb{R}^m$ be a set-valued function with non-empty closed

values. Let $\boldsymbol{H}\colon \mathrm{Z}\times\mathbb{R}^n \rightrightarrows \mathbb{R}^m$ *be defined by* $\boldsymbol{H}(z,\boldsymbol{A}) = \{\boldsymbol{B}\in\mathbb{R}^m\colon (\boldsymbol{A},\boldsymbol{B})\in \mathsf{F}(z)\}$, *for* $(z,\boldsymbol{A})\in \mathrm{Z}\times\mathbb{R}^n$. *Then the following are equivalent:*
(i) F *is* $\Sigma - (\mathcal{B}(\mathbb{R}^n)\otimes\mathcal{B}(\mathbb{R}^m))$ *measurable;*
(ii) $\boldsymbol{H}$ *is* $(\Sigma\otimes\mathcal{B}(\mathbb{R}^n)) - \mathcal{B}(\mathbb{R}^m)$ *measurable.*

Now let $\omega\subset\mathbb{R}^n$ be an open set and we denote by $\mathcal{L}(\omega)$ the σ-algebra of all Lebesgue measurable subsets of ω.

Corollary 3.10.
Let $\mathsf{F}\colon \omega \rightrightarrows \mathbb{R}^s\times\mathbb{R}^s$ *be a set-valued function, such that* $(\mathbf{0},\mathbf{0})\in\mathsf{F}(\boldsymbol{z})$ *and* $\mathsf{F}(\boldsymbol{z})$ *is maximal monotone for a.e.* $\boldsymbol{z}\in\omega$. *Let the set-valued function* $\boldsymbol{H}\colon\omega\times\mathbb{R}^s\rightrightarrows\mathbb{R}^s$ *be defined by* $\boldsymbol{H}(\boldsymbol{z},\boldsymbol{x}) := \{\boldsymbol{y}\in\mathbb{R}^s\colon (\boldsymbol{x},\boldsymbol{y})\in\mathsf{F}(\boldsymbol{z})\}$. *Then the following are equivalent:*
(i) F *is* $\mathcal{L}(\omega) - (\mathcal{B}(\mathbb{R}^s)\otimes\mathcal{B}(\mathbb{R}^s))$ *measurable;*
(ii) $\boldsymbol{H}$ *is* $(\mathcal{L}(\omega)\otimes\mathcal{B}(\mathbb{R}^s)) - \mathcal{B}(\mathbb{R}^s)$ *measurable.*

Proof. Since $(\mathbf{0},\mathbf{0})\in\mathsf{F}(\boldsymbol{z})$ for a.e. $\boldsymbol{z}\in\omega$ the values of F are non-empty. Whenever $\mathsf{F}(\boldsymbol{z})$ is maximal monotone, the function $\mathfrak{f}(\boldsymbol{z})\colon\mathbb{R}^{d\times d}_{\mathrm{sym}}\rightrightarrows\mathbb{R}^{d\times d}_{\mathrm{sym}}$ such that $\Gamma(\mathfrak{f}(\boldsymbol{z})) = \mathsf{F}(\boldsymbol{z})$ is maximal monotone. By Lemma 3.4 (i) this implies that $\Gamma(\mathfrak{f}(\boldsymbol{z})) = \mathsf{F}(\boldsymbol{z})$ is closed for a.e. $\boldsymbol{z}\in\omega$. The Lebesgue measure $\mathcal{L}$ is a complete σ-finite measure defined on the Lebesgue measurable subsets of ω. Thus, Lemma 3.9 applies and the claim follows. □

3.3. Assumptions on $\mathcal{A}$

For the rest of this chapter let $M\subset\mathbb{R}^n$, for $n\geq 2$, be an open bounded set. Furthermore, let $d\geq 2$ and recall that $\mathbb{R}^{d\times d}_{\mathrm{sym}}$ is the space of symmetric $(d\times d)$ matrices and $\mathbb{R}^{d\times d}_{\mathrm{sym},0}$ is the subspace of trace-free matrices.

In Subsection 1.1.3 we have assumed that the implicit constitutive relation can be identified with a mapping $\boldsymbol{z}\mapsto\mathcal{A}(\boldsymbol{z})$ via

$$\mathsf{G}(\boldsymbol{z},\mathsf{D},\mathsf{S}) = \mathbf{0} \Leftrightarrow (\mathsf{D},\mathsf{S})\in\mathcal{A}(\boldsymbol{z}). \tag{3.9}$$

Here $\mathcal{A}(\boldsymbol{z})$ is the graph of a maximal monotone set-valued function for a.e. $\boldsymbol{z}\in M$, which is why we refer to it as maximal monotone graph. This will be applied with

$$M = \begin{cases} \Omega & \text{in the steady case,}\\ Q & \text{in the unsteady case,}\end{cases}$$

in the Chapters 4 and 5, respectively.

Let us formulate the specific assumptions we pose on the mapping encoding the implicit constitutive relation, which are similar to the ones used in [BGMS09, BGMS12, DKS13a]:

Assumption 3.11 (Properties of $\mathcal{A}$)**.**
We assume that the set-valued mapping $\mathcal{A}\colon M\rightrightarrows\mathbb{R}^{d\times d}_{\mathrm{sym}}\times\mathbb{R}^{d\times d}_{\mathrm{sym}}$, *such that* $\boldsymbol{z}\mapsto\mathcal{A}(\boldsymbol{z})\subset\mathbb{R}^{d\times d}_{\mathrm{sym}}\times\mathbb{R}^{d\times d}_{\mathrm{sym}}$ *satisfies the following conditions for a.e.* $\boldsymbol{z}\in M$:
(A1) $(\mathbf{0},\mathbf{0})\in\mathcal{A}(\boldsymbol{z})$;
(A2) $\mathcal{A}(\boldsymbol{z})$ *is a maximal monotone set;*
(A3) There exists a constant $c_*>0$ *(independent of* $\boldsymbol{z}$*), a nonnegative function* $g\in\mathrm{L}^1(M)$ *and* $q\in(1,\infty)$ *such that*

$$\mathsf{D}:\mathsf{S}\geq -g(\boldsymbol{z}) + c_*(|\mathsf{D}|^q + |\mathsf{S}|^{q'}) \quad \textit{for all } (\mathsf{D},\mathsf{S})\in\mathcal{A}(\boldsymbol{z}),$$

where q' is the Hölder conjugate of q;

(A4) *The set-valued map $\boldsymbol{z} \mapsto \mathcal{A}(\boldsymbol{z})$ is $\mathcal{L}(M) - (\mathcal{B}(\mathbb{R}^{d\times d}_{\text{sym}}) \otimes \mathcal{B}(\mathbb{R}^{d\times d}_{\text{sym}}))$ measurable;*

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

$$\operatorname{tr}(\mathbf{D}) = 0 \Leftrightarrow \operatorname{tr}(\mathbf{S}) = 0.$$

Remark 3.12 (Properties of $\mathcal{A}$)**.**

(i) In [BGMS12] the authors phrase the coercivity and boundedness condition (A3) in the more general context of Orlicz–Sobolev spaces. For simplicity we will restrict ourselves to the Sobolev setting.

(ii) Given that we want to use the implicit constitutive relation for fluid equations, *we aim for a notion of weak solution with trace-free stress tensor to ensure consistence with the thermodynamic framework introduced in Section 1.1. For this reason, we introduce condition (A5) additionally to the assumptions in [BGMS09, BGMS12, DKS13a, KS16]. This implies that the stress tensor, which is part of the weak solution, is trace-free without imposing this condition separately. More specifically, if $\boldsymbol{u}$ is a (pointwise) divergence-free function, then we have that $\operatorname{tr}(\mathbf{D}\boldsymbol{u}) = \operatorname{div}(\boldsymbol{u}) = 0$. Thus, if $(\mathbf{D}\boldsymbol{u}(\boldsymbol{z}),\mathbf{S}(\boldsymbol{z})) \in \mathcal{A}(\boldsymbol{z})$ for a.e. $\boldsymbol{z} \in M$, this implies with (A5) that $\mathbf{S}$ is (pointwise) trace-free.*
Alternatively, one could consider a graph $\widetilde{\mathcal{A}} \subset \mathbb{R}^{d\times d}_{\text{sym},0} \times \mathbb{R}^{d\times d}_{\text{sym},0}$, where $\mathbb{R}^{d\times d}_{\text{sym},0}$ denotes the subspace of trace-free matrices in $\mathbb{R}^{d\times d}_{\text{sym}}$, then condition (A5) is trivially satisfied. However, dealing with such a graph is only possible if one deals with (exactly) divergence-free velocity functions, which is the case in the existence proof in [BGMS09, BGMS12, BM16] and if one chooses exactly divergence-free finite element functions for the numerical approximation, see Assumption 2.27. When using discretely divergence-free finite element functions this is not the case and for the numerical approximation $\widetilde{\mathcal{A}}$ has to be extended to $\mathbb{R}^{d\times d}_{\text{sym}} \times \mathbb{R}^{d\times d}_{\text{sym}}$.

(iii) Given (A1) and (A2), by Corollary 3.10 the measurability of the set-valued mapping $\mathcal{A}\colon M \rightrightarrows \mathbb{R}^{d\times d}_{\text{sym}} \times \mathbb{R}^{d\times d}_{\text{sym}}$ is equivalent to the fact that the set-valued mapping $\mathcal{S}\colon M \times \mathbb{R}^{d\times d}_{\text{sym}} \rightrightarrows \mathbb{R}^{d\times d}_{\text{sym}}$ defined by

$$\mathcal{S}(\boldsymbol{z},\mathbf{D}) = \{\mathbf{S} \in \mathbb{R}^{d\times d}_{\text{sym}} \colon (\mathbf{D},\mathbf{S}) \in \mathcal{A}(\boldsymbol{z})\}, \quad \boldsymbol{z} \in M,\, \mathbf{D} \in \mathbb{R}^{d\times d}_{\text{sym}}, \tag{3.10}$$

is $(\mathcal{L}(M) \otimes \mathcal{B}(\mathbb{R}^{d\times d}_{\text{sym}})) - \mathcal{B}(\mathbb{R}^{d\times d}_{\text{sym}})$ measurable. By definition (see [CPDMD90, (1.2)]) this means that for any closed $C \subset \mathbb{R}^{d\times d}_{\text{sym}}$ the set

$$\{(\boldsymbol{z},\mathbf{D}) \in M \times \mathbb{R}^{d\times d}_{\text{sym}} \colon \; \textit{there exists an } \mathbf{S} \in C \textit{ such that } (\mathbf{D},\mathbf{S}) \in \mathcal{A}(\boldsymbol{z})\}$$

is measurable with respect to the product σ-algebra $\mathcal{L}(M) \otimes \mathcal{B}(\mathbb{R}^{d\times d}_{\text{sym}})$. Note that using also (A3), [CPDMD90, Thm. 1.1 (i)] provides a set of equivalent characterisations for measurability of $\mathcal{S}$ and it shows the existence of a measurable selection, see Lemma 3.17.

(iv) We require $\big(\mathcal{L}(M) \otimes \mathcal{B}(\mathbb{R}^{d\times d}_{\text{sym}})\big) - \mathcal{B}(\mathbb{R}^{d\times d}_{\text{sym}})$ measurability of $\mathcal{S}$, so that the composition $\boldsymbol{z} \mapsto \mathcal{S}(\boldsymbol{z},\mathbf{D}\boldsymbol{u}(\boldsymbol{z}))$ is Lebesgue measurable (i.e. $\mathcal{L}(M) - \mathcal{B}(\mathbb{R}^{d\times d}_{\text{sym}})$ measurable) for Lebesgue measurable $\mathbf{D}\boldsymbol{u}$.

Remark 3.13 (Relaxed Assumptions on $\mathcal{A}$)**.**

(i) As pointed out in [FMT04, p. 28], condition (A1) can be relaxed to: there exists a function $\mathbf{S}_0 \in \mathrm{L}^{q'}(M)^{d\times d}$ with values in $\mathbb{R}^{d\times d}_{\text{sym},0}$, such that $(\mathbf{0},\mathbf{S}_0(\boldsymbol{z})) \in \mathcal{A}(\boldsymbol{z})$ for a.e. $\boldsymbol{z} \in M$; or equivalently, there exists a function $\mathbf{D}_0 \in \mathrm{L}^{q}(M)^{d\times d}$ with values in $\mathbb{R}^{d\times d}_{\text{sym},0}$

such that $(\mathbf{D}_0(\boldsymbol{z}), \mathbf{0}) \in \mathcal{A}(\boldsymbol{z})$ *for a.e.* $\boldsymbol{z} \in M$*. Then one can deduce Assumption 3.11 by shifting the graph* $\mathcal{A}$*.*

(ii) *In some situations one can consider a more general case, when* $\mathfrak{A} \subset \mathrm{L}^q(M)^{d\times d} \times \mathrm{L}^{q'}(M)^{d\times d}$*, for which the pointwise properties (A1)–(A3) are relaxed to the corresponding properties "integrated" in* M *and (A4) is dropped. These are the objects considered in [FMT04, Sec. 4].*

In the spirit of the one-to-one correspondence of maximal monotone functions and 1-Lipschitz functions, see Lemma 3.6, equivalently to assuming Assumption 3.11 (A1)–(A4) on $\mathcal{A}$ one could also assume the existence of a Carathéodory contraction with the corresponding properties. The latter is the approach taken in [FMT04, p. 28] and has the advantage that measurability is less intricate to deal with. The equivalence was shown in Rem. 2.2 therein. The following lemma is based on this equivalence, see also [BGMS09, Sec. 2.1].

Lemma 3.14 (Characterisation of $\mathcal{A}$ via Carathéodory Contraction ψ)**.**
Let a set-valued function $\mathcal{A}\colon M \rightrightarrows \mathbb{R}^{d\times d}_{\mathrm{sym}} \times \mathbb{R}^{d\times d}_{\mathrm{sym}}$ *and a set-valued function* $\psi\colon M \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \rightrightarrows \mathbb{R}^{d\times d}_{\mathrm{sym}}$ *be given, such that*

$$\mathcal{A}(\boldsymbol{z}) = \phi^{-1}(\Gamma(\psi(\boldsymbol{z},\cdot))) \qquad \text{for a.e. } \boldsymbol{z} \in M, \tag{3.11}$$

where ϕ *is the Cayley transformation, defined in* (3.5)*.*
Then, $\mathcal{A}$ *satisfies Assumption 3.11 (A1)–(A3) for a.e.* $\boldsymbol{z} \in M$ *if and only if* $\psi\colon M \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$ *is a (single-valued) function, satisfying the following conditions for a.e.* $\boldsymbol{z} \in M$*:*

(p1) $\psi(\boldsymbol{z}, \mathbf{0}) = \mathbf{0}$*;*

(p2) $\psi(\boldsymbol{z},\cdot)$ *is* 1*-Lipschitz and* $\mathrm{dom}(\psi(\boldsymbol{z},\cdot)) = \mathbb{R}^{d\times d}_{\mathrm{sym}}$*, i.e.,*

$$|\psi(\boldsymbol{z},\mathbf{B}_1) - \psi(\boldsymbol{z},\mathbf{B}_2)| \le |\mathbf{B}_1 - \mathbf{B}_2| \qquad \text{for all } \mathbf{B}_1, \mathbf{B}_2 \in \mathbb{R}^{d\times d}_{\mathrm{sym}};$$

(p3) *There exists a constant* $c_* > 0$ *(independent of* $\boldsymbol{z}$*), a nonnegative function* $g \in \mathrm{L}^1(M)$ *and* $q \in (1,\infty)$ *such that for any* $\mathbf{B} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$ *and a.e.* $\boldsymbol{z} \in M$ *one has that*

$$\mathbf{D}(\boldsymbol{z}) : \mathbf{S}(\boldsymbol{z}) \ge -g(\boldsymbol{z}) + c_*(|\mathbf{D}(\boldsymbol{z})|^q + |\mathbf{S}(\boldsymbol{z})|^{q'}),$$

where $(\mathbf{D}(\boldsymbol{z}), \mathbf{S}(\boldsymbol{z})) := \phi^{-1}(\mathbf{B}, \psi(\boldsymbol{z},\mathbf{B})) \in \mathcal{A}(\boldsymbol{z})$*;*

Note that c_**,* g *and* q *are the same in both coercivity estimates (A3) and (p3).*

Furthermore, if $\mathcal{A}$ *satisfies (A1)–(A3) for a.e.* $\boldsymbol{z} \in M$*, then:*
$\mathcal{A}$ *satisfies (A4) if and only if*

(p4) ψ *is a Carathéodory function, i.e.,* ψ *is* $\big(\mathcal{L}(M) \otimes \mathcal{B}(\mathbb{R}^{d\times d}_{\mathrm{sym}})\big) - \mathcal{B}(\mathbb{R}^{d\times d}_{\mathrm{sym}})$ *measurable.*

Proof. First note that we have by Lemma 3.6 that $\mathcal{A}(\boldsymbol{z})$ is maximal monotone if and only if $\mathrm{dom}\,\psi(\boldsymbol{z},\cdot) = \mathbb{R}^{d\times d}_{\mathrm{sym}}$ and $\psi(\boldsymbol{z},\cdot)$ is 1-Lipschitz (which in particular implies that ψ is a single-valued function). This shows that (A2) is equivalent to (p2).

Note that by the definition of ϕ in (3.5) we have that $(\mathbf{0},\mathbf{0}) \in \mathcal{A}(\boldsymbol{z}) = \phi^{-1}(\Gamma(\psi(\boldsymbol{z},\cdot)))$ is equivalent to $(\mathbf{0},\mathbf{0}) = \phi(\mathbf{0},\mathbf{0}) \in \Gamma(\psi(\boldsymbol{z},\cdot))$. So (p1) directly implies (A1). On the other hand (A1), together with the fact that ψ is single-valued by (A2), implies also that (p1) holds. The coercivity and boundedness conditions (A3) and (p3) are equivalent by definition.

Finally, if $\mathcal{A}$ satisfies (A1)–(A3) for a.e. $\boldsymbol{z} \in M$, then the arguments in [FMT04, Rem. 2.2] based on [CPDMD90, Thm. 1.3] and attributed to G. Dal Maso show the equivalence of (p4) and (A4). □

Is can easily be checked that explicit constitutive laws of power-law type as in (1.5b) satisfy Assumption 3.11.

However, for Herschel–Bulkley fluids the constitutive relation

$$\widetilde{\mathbf{G}}(\mathbf{D},\mathbf{S}) = |\mathbf{D}|^{q-2}\,(\tau_* + (|\mathbf{S}| - \tau_*)^+)\mathbf{D} - (|\mathbf{S}| - \tau_*)^+\mathbf{S} = \mathbf{0}, \tag{3.12}$$

is only defined for trace-free matrices $\mathbf{D},\mathbf{S} \in \mathbb{R}^{d\times d}_{\mathrm{sym},0}$, see (1.6) and (1.7). In [BGMS12, Lem. 1.1] it is proven that the graph $\widetilde{\mathcal{A}} \subset \mathbb{R}^{d\times d}_{\mathrm{sym},0} \times \mathbb{R}^{d\times d}_{\mathrm{sym},0}$ identified with $\widetilde{\mathbf{G}}$ as in (3.9) satisfies Assumption 3.11, if we replace $\mathbb{R}^{d\times d}_{\mathrm{sym}}$ by $\mathbb{R}^{d\times d}_{\mathrm{sym},0}$ in the statement. To obtain $\mathcal{A}$ satisfying Assumption 3.11, we want to find an extension of $\widetilde{\mathcal{A}}$ to $\mathbb{R}^{d\times d}_{\mathrm{sym}} \times \mathbb{R}^{d\times d}_{\mathrm{sym}}$. For a matrix $\mathbf{B} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$, we denote the trace part of $\mathbf{B}$ by $\mathsf{T}(\mathbf{B}) := \frac{1}{d}\,\mathrm{tr}(\mathbf{B})\mathbf{I}$, i.e., we have $\mathrm{tr}(\mathbf{B}) = \mathrm{tr}(\mathsf{T}(\mathbf{B}))$ and the deviatoric part of $\mathbf{B}$ by $\mathbf{B}_\delta := \mathbf{B} - \mathsf{T}(\mathbf{B}) \in \mathbb{R}^{d\times d}_{\mathrm{sym},0}$. Then it follows that

$$\mathsf{T}(\mathbf{B}) : \mathbf{B}_\delta = 0, \tag{3.13}$$

so the subspace $\mathbb{R}^{d\times d}_{\mathrm{sym,tr}} := \{\mathsf{T}(\mathbf{B})\colon\ \mathbf{B} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}\}$ is the orthogonal complement of $\mathbb{R}^{d\times d}_{\mathrm{sym},0}$ in $\mathbb{R}^{d\times d}_{\mathrm{sym}}$.

Lemma 3.15 ($\mathcal{A}$ for Herschel–Bulkley fluids)**.**
Let $\boldsymbol{\mathcal{E}}\colon \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$ *be defined by*

$$\boldsymbol{\mathcal{E}}(\mathbf{B}) = \begin{cases} |\mathbf{B}|^{\gamma}\,\mathbf{B} & \text{if } \mathbf{B} \neq \mathbf{0},\\ \mathbf{0}, & \text{if } \mathbf{B} = \mathbf{0}, \end{cases} \qquad \text{for} \qquad \gamma = \gamma(q) = \frac{2-q}{q-1} = \frac{1}{q-1} - 1 \in (-1,\infty), \tag{3.14}$$

and consider

$$\mathbf{G}(\mathbf{D},\mathbf{S}) := (\tau_* + (|\mathbf{S}_\delta| - \tau_*)^+)\,(\mathbf{D} - \boldsymbol{\mathcal{E}}(\mathsf{T}(\mathbf{S}))) - \left((|\mathbf{S}_\delta| - \tau_*)^+\right)^{\frac{1}{q-1}}\mathbf{S}_\delta, \tag{3.15}$$

for $\mathbf{D},\mathbf{S} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$*. Then, the set* $\mathcal{A} \subset \mathbb{R}^{d\times d}_{\mathrm{sym}} \times \mathbb{R}^{d\times d}_{\mathrm{sym}}$ *identified with* $\mathbf{G}$ *via* (3.9) *satisfies Assumption 3.11. Furthermore,* $\mathcal{A} \cap \mathbb{R}^{d\times d}_{\mathrm{sym},0} \times \mathbb{R}^{d\times d}_{\mathrm{sym},0} = \widetilde{\mathcal{A}}$*, where* $\widetilde{\mathcal{A}}$ *is the graph identified with* $\widetilde{\mathbf{G}}$ *in* (3.12)*.*

Proof. Clearly we have that $\boldsymbol{\mathcal{E}}(\mathbf{T}(\mathbf{B})) = \mathbf{0}$, for any $\mathbf{B} \in \mathbb{R}^{d\times d}_{\mathrm{sym},0}$, so it follows that

$$(\tau_* + (|\mathbf{S}| - \tau_*)^+)\mathbf{D} - \left((|\mathbf{S}| - \tau_*)^+\right)^{\frac{1}{q-1}}\mathbf{S} = \mathbf{0},$$

for all $(\mathbf{D},\mathbf{S}) \in \widetilde{\mathcal{A}} \subset \mathbb{R}^{d\times d}_{\mathrm{sym},0} \times \mathbb{R}^{d\times d}_{\mathrm{sym},0}$. Considering the two regimes separately one can see that this is equivalent to $\widetilde{\mathbf{G}}(\mathbf{D},\mathbf{S}) = \mathbf{0}$, for $\mathbf{D},\mathbf{S} \in \mathbb{R}^{d\times d}_{\mathrm{sym},0}$, with $\widetilde{\mathbf{G}}$ as in (3.12). Hence $\mathcal{A}$ is an extension of $\widetilde{\mathcal{A}}$ and $(\mathbf{0},\mathbf{0}) \in \mathcal{A}$, so (A1) is satisfied. By the monotonicity of the function $s \mapsto s^{\gamma+1}$ for $s \geq 0$, since $\gamma + 1 > 0$ we have that

$$\begin{aligned}(\boldsymbol{\mathcal{E}}(\mathbf{B}_1) - \boldsymbol{\mathcal{E}}(\mathbf{B}_2)) : (\mathbf{B}_1 - \mathbf{B}_2) &= (|\mathbf{B}_1|^{\gamma}\,\mathbf{B}_1 - |\mathbf{B}_2|^{\gamma}\,\mathbf{B}_2) : (\mathbf{B}_1 - \mathbf{B}_2)\\ &\geq \left(|\mathbf{B}_1|^{\gamma+1} - |\mathbf{B}_2|^{\gamma+1}\right)(|\mathbf{B}_1| - |\mathbf{B}_2|) \geq 0,\end{aligned} \tag{3.16}$$

and thus $\boldsymbol{\mathcal{E}}$ is monotone. Furthermore $\boldsymbol{\mathcal{E}}$ is continuous and we have that

$$\boldsymbol{\mathcal{E}}(\mathbf{B}) : \mathbf{B} = |\mathbf{B}|^{\gamma+2} = |\mathbf{B}|^{q'} = |\boldsymbol{\mathcal{E}}(\mathbf{B})|^{q} = \frac{1}{2}\left(|\boldsymbol{\mathcal{E}}(\mathbf{B})|^{q} + |\mathbf{B}|^{q'}\right). \tag{3.17}$$

Distinguishing cases from (3.15) we obtain that

$$(\mathbf{D},\mathbf{S}) \in \mathcal{A} \;\Leftrightarrow\; \mathbf{G}(\mathbf{D},\mathbf{S}) = \mathbf{0} \;\Leftrightarrow\; \begin{cases} |\mathbf{S}_\delta| \le \tau_* \;\Leftrightarrow\; \mathbf{D} = \mathfrak{E}(\mathsf{T}(\mathbf{S})) \\ |\mathbf{S}_\delta| > \tau_* \;\Leftrightarrow\; \mathbf{D} = (|\mathbf{S}_\delta| - \tau_*)^{\frac{1}{q-1}} \frac{\mathbf{S}_\delta}{|\mathbf{S}_\delta|} + \mathfrak{E}(\mathsf{T}(\mathbf{S})). \end{cases} \tag{3.18}$$

This implies that $\mathbf{T}(\mathbf{D}) = \mathfrak{E}(\mathbf{T}(\mathbf{S}))$, for $(\mathbf{D},\mathbf{S}) \in \mathcal{A}$ and hence $\mathbf{T}(\mathbf{D}) = \mathbf{0}$ if and only if $\mathbf{T}(\mathbf{S}) = \mathbf{0}$, which shows that Assumption 3.11 (A5) is satisfied.

Let $(\mathbf{D},\mathbf{S}), (\overline{\mathbf{B}},\overline{\mathbf{S}}) \in \mathcal{A}$. Since $\mathbf{T}(\mathbf{D}) = \mathfrak{E}(\mathbf{T}(\mathbf{S}))$ for all $(\mathbf{D},\mathbf{S}) \in \mathcal{A}$, by the definition of $\mathcal{A}$ and $\widetilde{\mathcal{A}}$ it follows that $(\mathbf{D}_\delta,\mathbf{S}_\delta), (\overline{\mathbf{D}}_\delta,\overline{\mathbf{S}}_\delta) \in \widetilde{\mathcal{A}}$. With the orthogonality in (3.13) this implies that

$$\begin{aligned} \left(\mathbf{D} - \overline{\mathbf{D}}\right) : \left(\mathbf{S} - \overline{\mathbf{S}}\right) &= \left(\mathbf{D}_\delta - \overline{\mathbf{D}}_\delta\right) : \left(\mathbf{S}_\delta - \overline{\mathbf{S}}_\delta\right) + \left(\mathbf{T}(\mathbf{D}) - \mathbf{T}(\overline{\mathbf{D}})\right) : \left(\mathbf{T}(\mathbf{S}) - \mathbf{T}(\overline{\mathbf{S}})\right) \\ &\ge 0 + \left(\mathfrak{E}(\mathbf{T}(\mathbf{S})) - \mathfrak{E}(\mathbf{T}(\overline{\mathbf{S}}))\right) : \left(\mathbf{T}(\mathbf{S}) - \mathbf{T}(\overline{\mathbf{S}})\right) \ge 0, \end{aligned} \tag{3.19}$$

where the first term on the right-hand side in the first line is nonnegative due to the monotonicity of $\widetilde{\mathcal{A}}$ according to [BGMS12, Lem. 1.1] and the second term in the second line is nonnegative due to the monotony of $\mathfrak{E}$ in (3.16). Hence, $\mathcal{A}$ is monotone. From (3.18) one can see that the function $\mathfrak{D}\colon \mathbb{R}^{d\times d}_{\mathrm{sym}} \rightrightarrows \mathbb{R}^{d\times d}_{\mathrm{sym}}$, such that $\mathbf{S} \mapsto \{\mathbf{D}\colon\ (\mathbf{D},\mathbf{S}) \in \mathcal{A}\}$, i.e., $\Gamma(\mathfrak{D}) = \mathcal{A}^{-1}$, is single-valued, continuous and $\mathrm{dom}(\mathfrak{D}) = \mathbb{R}^{d\times d}_{\mathrm{sym}}$. Then, by Lemma 3.4 (ii) it follows that $\mathfrak{D}$ is maximal monotone, and hence $\mathcal{A}^{-1}$ is maximal monotone, and so is $\mathcal{A}$, which shows Assumption 3.11 (A2). Since $\mathfrak{D}$ is continuous, its graph is closed, and hence $\mathcal{A}^{-1} = \Gamma(\mathfrak{D})$ is $\mathcal{B}(\mathbb{R}^{d\times d}_{\mathrm{sym}}) \otimes \mathcal{B}(\mathbb{R}^{d\times d}_{\mathrm{sym}})$ measurable. Thus $\mathcal{A}$, which does not depend on $\boldsymbol{z} \in M$, satisfies Assumption 3.11 (A4).

Finally for $(\mathbf{D},\mathbf{S}) \in \mathcal{A}$, again by the fact that $(\mathbf{D}_\delta,\mathbf{S}_\delta) \in \widetilde{\mathcal{A}}$ and $\mathbf{T}(\mathbf{D}) = \mathfrak{E}(\mathbf{T}(\mathbf{S}))$, it follows with the orthonormality in (3.13) that

$$\begin{aligned} \mathbf{D} : \mathbf{S} = \mathbf{D}_\delta : \mathbf{S}_\delta + \mathbf{T}(\mathbf{D}) : \mathbf{T}(\mathbf{S}) &\ge -\widetilde{c} + \widetilde{c}_* \left(|\mathbf{D}_\delta|^q + |\mathbf{S}_\delta|^{q'}\right) + \frac{1}{2}\left(|\mathfrak{E}(\mathbf{T}(\mathbf{S}))|^q + |\mathbf{T}(\mathbf{S})|^{q'}\right) \\ &= -\widetilde{c} + \widetilde{c}_* \left(|\mathbf{D}_\delta|^q + |\mathbf{S}_\delta|^{q'}\right) + \frac{1}{2}\left(|\mathbf{T}(\mathbf{D})|^q + |\mathbf{T}(\mathbf{S})|^{q'}\right), \end{aligned} \tag{3.20}$$

where we have used that $\widetilde{\mathcal{A}}$ satisfies the coercivity estimate corresponding to (A3) with $g = c > 0$ constant, by [BGMS12, Lem. 1.1] and (3.17). Since $(a+b)^p \le c(p)(a^p + b^p)$ for all $a, b \in \mathbb{R}_{\ge 0}$ and $p \in (0,\infty)$ it follows with $|\mathbf{D}|^2 = |\mathbf{D}_\delta|^2 + |\mathbf{T}(\mathbf{D})|^2$ that

$$|\mathbf{D}|^q = \left(|\mathbf{D}_\delta|^2 + |\mathbf{T}(\mathbf{D})|^2\right)^{\frac{q}{2}} \le c(q)\left(|\mathbf{D}_\delta|^q + |\mathbf{T}(\mathbf{D})|^q\right), \tag{3.21}$$

and the corresponding estimate holds for $|\mathbf{S}|^{q'}$. Applying this in (3.20) and choosing the constants suitably yields that $\mathcal{A}$ satisfies Assumption 3.11 (A3), which finishes the proof. □

This example shows that the following results on Bingham or Herschel–Bulkley type fluids are based on assumptions stronger than the ones in Assumption 3.11: [DL76, Ch. VI.3], [Ser91, MRS05, GMS07, ER12]. However, there are constitutive relations not included in the framework, such as the relation for the Prandtl–Eyring fluid, see [FS00, Ch. 4] and also [BDF12].

The following lemma is one of the crucial tools for the identification of the implicit constitutive law upon passage to the limit and was first proved in [BGMS12]; for a simpler proof see also [BM16, Lem. 3.1].

Lemma 3.16 (Localised Convergence Lemma of Minty type, [BGMS12, Lem. 2.4]).
Let $\mathcal{A}\colon M \rightrightarrows \mathbb{R}^{d\times d}_{\mathrm{sym}} \times \mathbb{R}^{d\times d}_{\mathrm{sym}}$, such that $\mathcal{A}(\boldsymbol{z})$ is a maximal monotone set for a.e. $\boldsymbol{z} \in M$, i.e., it satisfies Assumption 3.11 (A2). 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 $\widetilde{M} \subset M$ and a $p \in (1,\infty)$ such that

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

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

The proof of the convergence lemma formulated for subdomains $\widetilde{M} \subset M$ is based on the point-wise properties of $\mathcal{A}$. Recall that such a local version is required if Chacon's Biting lemma is applied. For implicitly constituted relations this is the case when a truncation is used, see Subsection 2.1.4.

Note that a non-local version of this lemma (i.e., with $\widetilde{M} = M$) can be proven for the more general graphs $\mathfrak{A} \subset \mathrm{L}^q(M)^{d\times d} \times \mathrm{L}^{q'}(M)^{d\times d}$, mentioned in Remark 3.13 (ii).

3.4. Approximation of the Implicit Constitutive Relation

For the existence proof of approximate solutions one typically applies a fixed point theorem or in the semidiscrete time dependent case one has the standard ODE theory at one's disposal. To be able to apply either of them, $\mathcal{A}(\cdot)$ has to be approximated by a sequence of graphs $\{\mathcal{A}^k(\cdot)\}_{k\in\mathbb{N}}$, such that $\mathcal{A}^k(\boldsymbol{z})$ is the graph of $\boldsymbol{\mathcal{S}}^k(\boldsymbol{z},\cdot)$ for a Carathéodory function $\boldsymbol{\mathcal{S}}^k$.

In this section we shall consider graph approximations and their properties: On the one hand we derive a general set of assumptions, see Assumption 3.18, which includes a large class of graphs. The assumptions we pose are the weakest we can think of, in order to take the graph approximation limit before and separately from the discretisation limit in the convergence proof. On the other hand we want to review various examples of graph approximation, most of which can be found in the literature. All of them satisfy the general Assumption 3.18 mentioned above, however, when investigating their properties we hope for a Minty type convergence lemma as outlined in Section 3.1, which can be applied for sequences $\{\mathbf{D}^k\}_{k\in\mathbb{N}}$, $\{\mathbf{S}^k\}_{k\in\mathbb{N}}$ such that $(\mathbf{D}^k(\boldsymbol{z}), \mathbf{S}^k(\boldsymbol{z})) \in \mathcal{A}^k(\boldsymbol{z})$ for all $k \in \mathbb{N}$ and a.e. $\boldsymbol{z} \in M$. It requires some knowledge about the "distance" between the graph $\mathcal{A}$ and the sequence of approximations $\mathcal{A}^k$, if one wants to deduce such a convergence result from the version in Lemma 3.16, see Lemma 3.27 and 3.31.

In the following we will first consider a selection of the graph $\mathcal{A}$ and introduce the general set of assumptions on a sequence of Carathéodory functions $\boldsymbol{\mathcal{S}}^k$ representing the graph approximation. Then we will introduce and investigate three types of approximation, which satisfy those assumptions. In Subsection 3.4.1 we consider approximations which arise through smoothing from a measurable selection. In Subsection 3.4.2 we deal with an approximation, which is constructed starting from the characterising Carathéodory contraction ψ representing $\mathcal{A}$. Unfortunately we can show the coercivity property only for $q = 2$, but this approximation allows us to show a convergence lemma of Minty type. Finally in Subsection 3.4.3 a so-called generalised Yosida graph approximation is investigated, for which we can show a convergence lemma of Minty type for any $q \in (1,\infty)$.

The following result about the existence of a measurable selection follows from Assumption 3.11 on $\mathcal{A}(\cdot)$ and the properties of the function $\mathcal{S}$, such that $\Gamma(\mathcal{S}(\boldsymbol{z},\cdot)) = \mathcal{A}(\boldsymbol{z})$.

Lemma 3.17 (Measurable Selection, [BGMS12, Rem. 1.1, Lem. 2.2])**.**
Let $q \in (1,\infty)$ and let the map $\boldsymbol{z} \mapsto \mathcal{A}(\boldsymbol{z}) \subset \mathbb{R}^{d\times d}_{\mathrm{sym}} \times \mathbb{R}^{d\times d}_{\mathrm{sym}}$, for $\boldsymbol{z} \in M$, satisfy Assumption 3.11 with respect to q. Then, there exists a selection $\mathbf{\mathcal{S}}^\star \colon M \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$, i.e.,

$$(\mathbf{B}, \mathbf{\mathcal{S}}^\star(\boldsymbol{z},\mathbf{B})) \in \mathcal{A}(\boldsymbol{z}) \qquad \text{for all } \mathbf{B} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}, \text{ for a.e. } \boldsymbol{z} \in M, \tag{3.22}$$

which is $\big(\mathcal{L}(M) \otimes \mathcal{B}(\mathbb{R}^{d\times d}_{\mathrm{sym}})\big) - \mathcal{B}(\mathbb{R}^{d\times d}_{\mathrm{sym}})$ measurable. Furthermore, for a.e. $\boldsymbol{z} \in M$, one has the following:

(S1) $\operatorname{dom} \mathbf{\mathcal{S}}^\star(\boldsymbol{z},\cdot) = \mathbb{R}^{d\times d}_{\mathrm{sym}}$*;*

(S2) $\mathbf{\mathcal{S}}^\star(\boldsymbol{z},\cdot)$ *is monotone;*

(S3) For the constant $c_* > 0$ and the nonnegative function $g \in \mathrm{L}^1(M)$ from Assumption 3.11 we have that

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

for any $\mathbf{B} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$;

(S4) 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)* $(\mathbf{S} - \mathbf{\mathcal{S}}^\star(\boldsymbol{z},\mathbf{B})) : (\mathbf{D} - \mathbf{B}) \geq 0 \quad$ *for all* $\mathbf{B} \in U$*;*
- *(ii)* $(\mathbf{D},\mathbf{S}) \in \mathcal{A}(\boldsymbol{z})$*;*

(S5) $\mathbf{\mathcal{S}}^\star$ is locally bounded, in the sense that for a given $r > 0$ there exists a constant $c = c(r)$ such that

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

Proof. The proof of (S1)–(S4) follows along the lines of [BGMS12], see Rem. 1.1 and Lem. 2.2, therein. Let us give a bit more detail. Recall that the function $\mathcal{S} \colon M \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \rightrightarrows \mathbb{R}^{d\times d}_{\mathrm{sym}}$, defined in (3.10) is $(\mathcal{L}(M) \otimes \mathcal{B}(\mathbb{R}^{d\times d}_{\mathrm{sym}})) - \mathcal{B}(\mathbb{R}^{d\times d}_{\mathrm{sym}})$ measurable. Let us consider the extension $\overline{\mathcal{S}} \colon \mathbb{R}^n \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$, defined by

$$\overline{\mathcal{S}}(\boldsymbol{z},\mathbf{B}) \coloneqq \mathbb{1}_M(\boldsymbol{z})\mathcal{S}(\boldsymbol{z},\mathbf{B}) + \mathbb{1}_{\mathbb{R}^n\setminus M}(\boldsymbol{z})\, |\mathbf{B}|^{q-2}\mathbf{B} \qquad \text{for } \boldsymbol{z} \in \mathbb{R}^n,\ \mathbf{B} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}.$$

Since the function $\mathbf{B} \mapsto |\mathbf{B}|^{q-2}\mathbf{B}$ is continuous, single-valued and monotone and the graph $\overline{\mathcal{A}}(\boldsymbol{z}) \coloneqq \Gamma(\overline{\mathcal{S}}(\boldsymbol{z},\cdot))$ satisfies the properties corresponding to (A1)–(A5), it follows that $\overline{\mathcal{S}}(\boldsymbol{z})$ satisfies the same properties for a.e. $\boldsymbol{z} \in \mathbb{R}^n$ as $\mathcal{S}$ does for a.e. $\boldsymbol{z} \in M$. By the maximal monotonicity of $\overline{\mathcal{A}}(\boldsymbol{z})$ also $\overline{\mathcal{S}}(\boldsymbol{z},\cdot)$ is maximal monotone, and hence the set of values is closed by Lemma 3.4 (i). Furthermore, by the maximal monotonicity and the bounds in (A3) one can show that $\operatorname{dom}(\overline{\mathcal{S}}(\boldsymbol{z},\cdot)) = \mathbb{R}^{d\times d}_{\mathrm{sym}}$ for a.e. $\boldsymbol{z} \in \mathbb{R}^n$ and consequently the set of values is non-empty. Then, with a selection lemma, see e.g. [CPDMD90, Thm. 1.1] or [AF09, Thm. 8.1.3], there exists a measurable selection $\overline{\mathbf{\mathcal{S}}}^\star \colon \mathbb{R}^n \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$ and denote by $\mathbf{\mathcal{S}}^\star \colon M \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$ its restriction to $\boldsymbol{z} \in M$, which satisfies (3.22). Then by the definition of the selection $\mathbf{\mathcal{S}}^\star$ and the fact that $\operatorname{dom}(\mathcal{S}(\boldsymbol{z},\cdot)) = \mathbb{R}^{d\times d}_{\mathrm{sym}}$ for a.e. $\boldsymbol{z} \in M$, (S1) follows. The properties (S2) and (S3) are inherited from the properties (A2) and (A3). The property (S4) follows as in [BGMS12, Lem. 2.2] and uses the fact that any monotone graph can be extended to a maximal monotone graph.

In order to show (S5) we shall use the property (iv) in Lemma 3.4: Let us consider the diffeomorphism $h \colon \mathbb{R}^n \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}^{(n+d)\times(n+d)}_{\mathrm{sym}}$ defined by

$$(\boldsymbol{z},\mathbf{D}) \mapsto \begin{pmatrix} \operatorname{diag}(\boldsymbol{z}) & \mathbf{0} \\ \mathbf{0} & \mathbf{D} \end{pmatrix}.$$

Then we define the set-valued function $\widetilde{\mathcal{S}}\colon h(\mathbb{R}^n \times \mathbb{R}^{d\times d}_{\mathrm{sym}}) \rightrightarrows \mathbb{R}^{(n+d)\times(n+d)}_{\mathrm{sym}}$ by

$$\widetilde{\mathcal{S}}(h(\boldsymbol{z},\mathbf{D})) = \left\{ \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{S} \end{pmatrix} : \mathbf{S} \in \overline{\mathcal{S}}(\boldsymbol{z},\mathbf{D}) \right\}, \qquad \text{for } \boldsymbol{z} \in \mathbb{R}^n,\, \mathbf{D} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}.$$

This function is maximal monotone on $h(\mathbb{R}^n \times \mathbb{R}^{d\times d}_{\mathrm{sym}})$ by the maximal monotonicity of $\overline{\mathcal{S}}(\boldsymbol{z},\cdot)$.

Now for arbitrary but fixed $r > 0$ let $B_r(\mathbf{0}) \subset \mathbb{R}^{d\times d}_{\mathrm{sym}}$. Since $\mathrm{dom}(\overline{\mathcal{S}}) = \mathbb{R}^n \times \mathbb{R}^{d\times d}_{\mathrm{sym}}$ it follows that $\mathrm{dom}(\widetilde{\mathcal{S}}) = h(\mathbb{R}^n \times \mathbb{R}^{d\times d}_{\mathrm{sym}})$. As h is a diffeomorphism, it follows that $h(M \times B_r(\mathbf{0}))$ is relatively compact in $h(\mathbb{R}^n \times \mathbb{R}^{d\times d}_{\mathrm{sym}}) = \operatorname{int}\operatorname{dom}\widetilde{\mathcal{S}}$. Then, Lemma 3.4 (iv) shows that $\widetilde{\mathcal{S}}(h(M \times B_r(\mathbf{0})))$ is bounded (by a constant depending on $r > 0$). This implies that $\overline{\mathcal{S}}(M \times B_r(\mathbf{0}))) = \mathcal{S}(M \times B_r(\mathbf{0})))$ is bounded and so is $\mathbf{\mathcal{S}}^\star(M \times B_r(\mathbf{0})))$. This finishes the proof of (S5).

□

To show existence of approximate solutions we require the approximate stress tensor function to be a Carathéodory function with the same properties as $\mathcal{S}$, but uniformly in the approximation parameter. The following assumptions are distilled from the existence proof in [BGMS12], in which an approximation based on mollification of the selection is used.

Assumption 3.18 (Properties of $\mathbf{\mathcal{S}}^k$, $k \in \mathbb{N}$)**.**
Assume that there is a sequence $\{\mathbf{\mathcal{S}}^k\}_{k\in\mathbb{N}}$ *of Carathéodory functions* $\mathbf{\mathcal{S}}^k\colon M \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$ *such that:*

(σ1) $\mathbf{\mathcal{S}}^k(\boldsymbol{z},\cdot)$ *is monotone;*

(σ2) There exists a constant $\widetilde{c}_* > 0$ *and a nonnegative function* $\widetilde{g} \in \mathrm{L}^1(M)$ *such that, for all* $k \in \mathbb{N}$, *for any* $\mathbf{B} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$ *and for a.e.* $\boldsymbol{z} \in M$, *one has that*

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

(σ3) 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* $\mathrm{L}^\infty(M)^{d\times d}$ *and with values in* $\mathbb{R}^{d\times d}_{\mathrm{sym}}$, *for any* $\mathbf{B} \in U$ *and all* $\varphi \in C_0^\infty(M)$ *such that* $\varphi \geq 0$, *we have*

$$\liminf_{k\to\infty} \int_M \left(\mathbf{\mathcal{S}}^k(\cdot,\mathbf{D}^k) - \mathbf{\mathcal{S}}^\star(\cdot,\mathbf{B}) \right) : \left(\mathbf{D}^k - \mathbf{B} \right) \varphi \,\mathrm{d}\boldsymbol{z} \geq 0.$$

This assumption is satisfied by all the approximations we consider in the remainder of the chapter.

Remark 3.19 (Identification of the Implicit Constitutive Relation as $k \to \infty$)**.**
The condition (σ3) allows us to take the limit $k \to \infty$ *before the discretisation limit in the convergence proof. Indeed, staying in the finite-dimensional situation we obtain strong convergence of* $\mathbf{D}\boldsymbol{u}^k \to \mathbf{D}\boldsymbol{u}$ *in* $\mathrm{L}^\infty(M)^{d\times d}$ *and from the a priori bounds we also have weak convergence of* $\mathbf{\mathcal{S}}^k(\cdot,\mathbf{D}\boldsymbol{u}^k) \rightharpoonup \mathbf{S}$ *(not indicating the discretisation parameter), as* $k \to \infty$. *This means that taking the limit in the expression in (σ3) yields that*

$$\int_M (\mathbf{S} - \mathbf{\mathcal{S}}^\star(\cdot,\mathbf{B})) : (\mathbf{D}\boldsymbol{u} - \mathbf{B})\, \varphi \,\mathrm{d}\boldsymbol{z} \geq 0,$$

for all $\mathbf{B} \in U$ *and all nonnegative* $\varphi \in \mathrm{C}_0^\infty(M)$. *By Lemma 3.17 (S4) this implies that* $(\mathbf{D}\boldsymbol{u},\mathbf{S}) \in \mathcal{A}(\cdot)$ *a.e. in* M, *and in the following limits the Lemma 3.16 can be applied to identify the implicit constitutive relation.*

3.4.1. Approximation of a Measurable Selection

Since the function $\mathcal{S}\colon M \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \rightrightarrows \mathbb{R}^{d\times d}_{\mathrm{sym}}$ is possibly a set-valued mapping, the measurable selection $\mathbf{S}^\star\colon M \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \rightrightarrows \mathbb{R}^{d\times d}_{\mathrm{sym}}$ is possibly discontinuous. To obtain a sequence of Carathéodory functions satisfying Assumption 3.18 one has to smooth out the jumps. Here we will consider two ways of doing so: The first one is a mollification of $\mathbf{S}^\star$ in the second argument, which has good properties but is unlikely to be convenient from a computational point of view. The second one is based on piecewise affine interpolation, which one can hope to be easier to implement, however, additional restrictions on $\mathcal{A}$ are required, such as radial symmetry, in order to show Assumption 3.18.

Mollification

In the existence proofs in [BGMS09, BGMS12] and also in [DKS13a] the sequence $\mathbf{S}^k$ is chosen as the convolution of the selection $\mathbf{S}^\star$ in the second argument with a mollification kernel.

Example 3.20 (Approximation by Mollification)**.**
Let $\rho \in C^\infty_0(\mathbb{R}^{d\times d}_{\mathrm{sym}})$ 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 $\rho^k(\mathbf{A}) := k^{d^2}\rho(k\mathbf{A})$ and define the mollification of $\mathbf{S}^\star$ with respect to the last argument by

$$\mathbf{S}^k(\boldsymbol{z},\mathbf{B}) := (\mathbf{S}^\star * \rho^k)(\boldsymbol{z},\mathbf{B}) = \int_{\mathbb{R}^{d\times d}_{\mathrm{sym}}} \mathbf{S}^\star(\boldsymbol{z},\mathbf{A})\rho^k(\mathbf{B}-\mathbf{A})\,\mathrm{d}\mathbf{A}, \tag{3.23}$$

for $\boldsymbol{z} \in M$ and $\mathbf{B} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$.

Lemma 3.21 (Properties of $\mathbf{S}^k$, $k \in \mathbb{N}$)**.**
For each $k \in \mathbb{N}$ the function $\mathbf{S}^k\colon M \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$ defined in (3.23) is measurable with respect to its first argument and smooth with respect to the second argument. Furthermore, the sequence $\{\mathbf{S}^k\}_{k\in\mathbb{N}}$ satisfies the Assumption 3.18 with $\widetilde{g} = g$, $\widetilde{c}_* = c_*$ and $q \in (1,\infty)$, as in Lemma 3.17.

Proof. Smoothness and measurability follow by the definition of the convolution and Fubini's theorem. To show the properties $(\sigma 1)$ and $(\sigma 2)$ is straightforward from the definition of $\mathbf{S}^k$ in (3.23). The proof of property $(\sigma 3)$ can be extracted from [BGMS12, Sec. 3.2] and we will reproduce it for the sake of completeness: let $\boldsymbol{z} \in M$ be arbitrary but fixed and not in any of the zero-sets for which the properties of $\mathbf{S}^\star$ do not hold.

Let $\{\mathbf{D}^k\}_{k\in\mathbb{N}} \subset \mathrm{L}^\infty(M)^{d\times d}$ be bounded and have values in $\mathbb{R}^{d\times d}_{\mathrm{sym}}$ and let $\mathbf{B} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$ be arbitrary. The monotonicity of $\mathbf{S}^\star(\boldsymbol{z})$ gives that

$$(\mathbf{S}^\star(\boldsymbol{z},\mathbf{A}) - \mathbf{S}^\star(\boldsymbol{z},\mathbf{B})) : (\mathbf{A}-\mathbf{B}) \geq 0 \quad \text{for all } \mathbf{A} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}, \quad \text{for a.e. } \boldsymbol{z} \in M. \tag{3.24}$$

Then, using the definition of $\mathbf{S}^k$ in (3.23), the fact that ρ^k integrates to 1 on $\mathbb{R}^{d\times d}_{\mathrm{sym}}$, the pointwise estimate (3.24) and the nonnegativity of ρ^k, we obtain that

$$\begin{aligned}
&\Big(\mathbf{S}^k(\boldsymbol{z},\mathbf{D}^k(\boldsymbol{z})) - \mathbf{S}^\star(\boldsymbol{z},\mathbf{B})\Big) : \Big(\mathbf{D}^k(\boldsymbol{z}) - \mathbf{B}\Big)\\
&\overset{(3.23)}{=} \int_{\mathbb{R}^{d\times d}_{\mathrm{sym}}} (\mathbf{S}^\star(\boldsymbol{z},\mathbf{A}) - \mathbf{S}^\star(\boldsymbol{z},\mathbf{B})) : (\mathbf{D}^k(\boldsymbol{z}) - \mathbf{B})\rho^k(\mathbf{D}^k(\boldsymbol{z}) - \mathbf{A})\,\mathrm{d}\mathbf{A}
\end{aligned}$$

$$\begin{aligned}&\overset{(3.24)}{\geq} \int_{\mathbb{R}^{d\times d}_{\mathrm{sym}}} (\mathcal{S}^\star(\boldsymbol{z},\mathbf{A}) - \mathcal{S}^\star(\boldsymbol{z},\mathbf{B})) : (\mathbf{D}^k(\boldsymbol{z}) - \mathbf{A})\rho^k(\mathbf{D}^k(\boldsymbol{z}) - \mathbf{A})\,\mathrm{d}\mathbf{A} \\ &\geq - \int_{\mathbb{R}^{d\times d}_{\mathrm{sym}}} |\mathcal{S}^\star(\boldsymbol{z},\mathbf{A}) - \mathcal{S}^\star(\boldsymbol{z},\mathbf{B})| \left|\mathbf{D}^k(\boldsymbol{z}) - \mathbf{A}\right| \rho^k(\mathbf{D}^k(\boldsymbol{z}) - \mathbf{A})\,\mathrm{d}\mathbf{A} \\ &\geq -\frac{1}{k}\int_{B_{1/k}(\mathbf{D}^k(\boldsymbol{z}))} |\mathcal{S}^\star(\boldsymbol{z},\mathbf{A}) - \mathcal{S}^\star(\boldsymbol{z},\mathbf{B})|\,\rho^k(\mathbf{D}^k(\boldsymbol{z}) - \mathbf{A})\,\mathrm{d}\mathbf{A} =: (\star),\end{aligned} \tag{3.25}$$

where in the last step we have used that by the definition of ρ^k we have

$$\operatorname{supp}(\rho^k(\mathbf{D}^k(\boldsymbol{z}) - \mathbf{A})) \subset B_{1/k}(\mathbf{D}^k(\boldsymbol{z})),$$

and hence $\left|\mathbf{D}^k(\boldsymbol{z}) - \mathbf{A}\right| \leq \frac{1}{k}$. Since the integral is supported on $B_{1/k}(\mathbf{D}^k(\boldsymbol{z}))$, and since $\mathbf{D}^k$ is uniformly bounded in $\mathrm{L}^\infty(\Omega)^{d\times d}$, the matrices $\mathbf{A}$ over which we integrate are bounded by

$$|\mathbf{A}| \leq \tfrac{1}{k} + \left\|\mathbf{D}^k\right\|_{\mathrm{L}^\infty(M)} \leq c, \tag{3.26}$$

where the constant c is independent of $\boldsymbol{z} \in \Omega$ and of $k \in \mathbb{N}$.

Then, since $\mathcal{S}^\star(\boldsymbol{z},\cdot)$ is locally bounded by Lemma 3.17 (S5), $B_{1/k}(\mathbf{D}^k(\boldsymbol{z}))$ is bounded and $\mathbf{B} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$ is fixed, there exists a constant $c > 0$, independent of k and $\boldsymbol{z}$, such that

$$\begin{aligned}|\mathcal{S}^\star(\boldsymbol{z},\mathbf{A})| &\leq c && \text{for all } \mathbf{A} \in B_{1/k}(\mathbf{D}^k(\boldsymbol{z})),\\ |\mathcal{S}^\star(\boldsymbol{z},\mathbf{B})| &\leq c(\mathbf{B}).\end{aligned}$$

This implies that independently of $\boldsymbol{z} \in M$ we have that

$$|\mathcal{S}^\star(\boldsymbol{z},\mathbf{A}) - \mathcal{S}^\star(\boldsymbol{z},\mathbf{B})| \leq |\mathcal{S}^\star(\boldsymbol{z},\mathbf{A})| + |\mathcal{S}^\star(\boldsymbol{z},\mathbf{B})| \leq c(\mathbf{B})$$

for all $\mathbf{A} \in B_{1/k}(\mathbf{D}^k(\boldsymbol{z}))$. With this we can estimate $(\star)$ further by

$$\begin{aligned}(\star) &= -\frac{1}{k}\int_{B_{1/k}(\mathbf{D}^k(\boldsymbol{z}))} |\mathcal{S}^\star(\boldsymbol{z},\mathbf{A}) - \mathcal{S}^\star(\boldsymbol{z},\mathbf{B})|\,\rho^k(\mathbf{D}^k(\boldsymbol{z}) - \mathbf{A})\,\mathrm{d}\mathbf{A}\\ &\geq -\frac{1}{k}c(\mathbf{B})\int_{B_{1/k}(\mathbf{D}^k(\boldsymbol{z}))} \rho^k(\mathbf{D}^k(\boldsymbol{z}) - \mathbf{A})\,\mathrm{d}\mathbf{A} = -\frac{c(\mathbf{B})}{k},\end{aligned} \tag{3.27}$$

where we have again used that ρ^k integrates to 1. Now we take (3.25) and (3.27) together, multiply by $\varphi \in C_0^\infty(M)$ such that $\varphi \geq 0$, and then integrate over M. Recalling that the constant is independent of $\boldsymbol{z} \in M$ this shows that

$$\int_M (\mathcal{S}^k(\cdot,\mathbf{D}^k) - \mathcal{S}^\star(\cdot,\mathbf{B})) : (\mathbf{D}^k - \mathbf{B})\varphi\,\mathrm{d}\boldsymbol{z} \geq -\frac{c(\mathbf{B})}{k}\,\|\varphi\|_{\mathrm{L}^1(M)}\,.$$

Then, applying $\liminf_{k\to\infty}$ yields

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

which proves the claim. □

The approximation based on mollification of the selection $\mathcal{S}^\star$ has the advantage that it is relatively simple and most properties of $\mathcal{S}^\star$ carry over to $\mathcal{S}^k$. Unfortunately one should expect it to be computationally expensive to implement and there is no Minty type convergence result available.

Piecewise Affine Interpolation

Motivated by the computationally not very appealing mollification-based approximation, let us turn turn our attention to an approximation based on piecewise affine interpolation.

Unfortunately, for vector-valued functions monotonicity is in general not preserved under piecewise affine interpolation. This has to do with the cross terms arising from the monotonicity relation and can be seen from the equivalent problem of preserving the Lipschitz constant under piecewise affine interpolation. By the fact that the Cayley transformation ϕ is a linear isometry any piecewise affine interpolant of a monotone function coincides, after applying ϕ to it, with a piecewise affine interpolation of the characterising 1-Lipschitz function, given by Lemma 3.6 (i).

Even though it is not a counter example, the following example highlights that piecewise affine interpolation need not preserve 1-Lipschitz continuity: Let $f\colon [0,1] \to \mathbb{R}$ be defined by $f(x,y) = \frac{1}{\sqrt{5}}(x-1)(y+1)$; one can show that $|\nabla f|^2 = \frac{1}{5}\left((y+1)^2 + (x-1)^2\right) \leq 1$. The piecewise affine interpolant with respect to the points $(0,0)$, $(0,\delta)$ and $(\delta,0)$, for $\delta \in (0,1)$, is given by $g(x,y) = \frac{1}{\sqrt{5}}(x-(2-\delta)y-1)$. Then, one can compute that $|\nabla g|^2 = \frac{1}{5}(1+(2+\delta)^2) > 1$, for all $\delta > 0$.

Sufficient conditions for a vector-valued function to have a sequence of 1-Lipschitz interpolations is, that the original function is Lipschitz continuous with Lipschitz constant < 1 and has a bounded Hessian matrix. This, however, implies by Lemma 3.6 (iii) that the corresponding monotone function is already strictly monotone and single-valued. Hence such an interpolation is not of much interest for the approximation of implicit constitutive laws.

Another option is to assume radial symmetry of the constitutive relation. This is for example the case for any explicit relation of the form 1.5b and for Herschel–Bulkley fluids when considering $\widetilde{\mathcal{A}} \subset \mathbb{R}^{d\times d}_{\mathrm{sym},0} \times \mathbb{R}^{d\times d}_{\mathrm{sym},0}$, see (1.6). In this situation monotonicity reduces to monotonicity of a scalar function, which is preserved by interpolants. To show Assumption 3.18 one would additionally want to assume some additional regularity of the selection function away from the jumps.

Example 3.22 (Approximation by Affine Interpolation)**.**
Assume that $\mathsf{S}^\star\colon M \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ *is a measurable function with* $\mathsf{S}^\star(\boldsymbol{z}, 0) = 0$*, for any* $\boldsymbol{z} \in M$*, such that* $\boldsymbol{\mathcal{S}}^\star\colon M \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$*, defined by*

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

is a measurable selection of a graph $\mathcal{A}$ *satisfying Assumption 3.11. Furthermore, we assume that*

(i) $\mathsf{S}^\star(\boldsymbol{z}, \cdot)\colon \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ *is monotone for a.e.* $\boldsymbol{z} \in M$*;*

Denoting by $J^\star \coloneqq \bigcup_{\boldsymbol{z}\in M} J(\mathsf{S}^\star(\boldsymbol{z}, \cdot))$ *the overall jump set, where* $J(\mathsf{S}^\star(\boldsymbol{z}, \cdot))$ *is the jump set of* $\mathsf{S}^\star(\boldsymbol{z}, \cdot)$*, which is countable by monotonicity of* $\mathsf{S}^\star(\boldsymbol{z}, \cdot)$ *for fixed* $\boldsymbol{z} \in M$*, see [AA99, Thm. 2.2], let us assume one of the following:*

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

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

Then there exists an index set $\mathcal{I}$, which is $\mathcal{I} = \mathbb{N}_0$ in case (iib) and there exists an $I \in \mathbb{N}$ such that $\mathcal{I} = \{0, \dots, I\}$ in case (iia), 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 $i \in \mathcal{I}\backslash\{0\}$.

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

$$a_{i,-}^k := a_i - \frac{1}{k}, \quad a_{i,+}^k := a_i + \frac{1}{k}, \quad A_i^k := [a_{i,-}^k, a_{i,+}^k] \quad \text{and} \quad A^k := \bigcup_{i\in\mathcal{I}} A_i^k.$$

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

$$\begin{aligned} \overline{\mathsf{S}}_i^k(\boldsymbol{z}, B) &:= \frac{k}{2}\left(\mathsf{S}^\star(\boldsymbol{z}, a_{i,-}^k)\frac{a_{i,+}^k}{a_{i,-}^k} - \mathsf{S}^\star(\boldsymbol{z}, a_{i,+}^k)\right)(a_{i,-}^k - B) + \mathsf{S}^\star(\boldsymbol{z}, a_{i,-}^k)\frac{B}{a_{i,-}^k}, \\ \mathsf{S}^k(\boldsymbol{z}, B) &:= \begin{cases} \mathsf{S}^\star(\boldsymbol{z}, B) & \text{if } B \notin A^k, \\ \overline{\mathsf{S}}_i^k(\boldsymbol{z}, B) & \text{if } B \in A_i^k, \quad i \in \mathcal{I}. \end{cases} \end{aligned} \tag{3.28}$$

On A_i^k the approximation $\mathsf{S}^k(\boldsymbol{z}, \cdot)$ is the affine interpolant between $\mathsf{S}^\star(\boldsymbol{z}, a_{i,-}^k)$ and $\mathsf{S}^\star(\boldsymbol{z}, a_{i,+}^k)$ and otherwise $\mathsf{S}^\star(\boldsymbol{z}, \cdot)$ is unchanged.

Note that assumption (iib) cannot hold for $q > 2$.

Lemma 3.23 (Properties of S^k, $k \in \mathbb{N}$)**.**
Each function $\mathsf{S}^k\colon M \times \mathbb{R} \to \mathbb{R}$ of the sequence $\{\mathsf{S}^k\}_{k\geq k_0}$ as defined in (3.28) is a Carathéodory function such that:

(σ1') $\mathsf{S}^k(\boldsymbol{z}, \cdot)$ is monotone for a.e. $\boldsymbol{z} \in M$ and all $k \geq k_0$;

(σ2') There exists a constant $\widetilde{c}_* > 0$ and a nonnegative function $\widetilde{g} \in \mathrm{L}^1(M)$ and $\widetilde{k}_0 \in \mathbb{N}$ such that

$$\mathsf{S}^k(\boldsymbol{z}, B)B \geq -\widetilde{g}(\boldsymbol{z}) + \widetilde{c}_*\left(|B|^q + \left|\mathsf{S}^k(\boldsymbol{z}, B)\right|^{q'}\right),$$

for all $B \in \mathbb{R}_{\geq 0}$ and for a.e. $\boldsymbol{z} \in M$ and all $k \geq \widetilde{k}_0$;

(σ3') for any sequence $\{D^k\}_{k\in\mathbb{N}}$ bounded in $\mathrm{L}^\infty(M)$, for any $B \in \mathbb{R}_{\geq 0}$ and all $\varphi \in C_0^\infty(M)$ such that $\varphi \geq 0$, we have

$$\liminf_{k\to\infty} \int_M \left(\mathsf{S}^k(\cdot, D^k) - \mathsf{S}^\star(\cdot, B)\right) : \left(D^k - B\right)\varphi \, \mathrm{d}\boldsymbol{z} \geq 0.$$

Proof. Since $\mathsf{S}^\star(\cdot, \cdot)$ is measurable in the first argument for every value of the second argument by Lemma 3.17, so is $\mathsf{S}^k(\cdot, \cdot)$, and by construction $\mathsf{S}^k(\cdot, \cdot)$ is continuous in the second argument for almost every value of the second argument. Hence, $\mathsf{S}^k(\cdot, \cdot)$ is a Carathéodory function.

(σ1') By Lemma 3.17 (S2), $\mathsf{S}^\star(\boldsymbol{z}, \cdot)$ is monotone for a.e. $\boldsymbol{z} \in M$. Then, by construction $\mathsf{S}^k(\boldsymbol{z}, \cdot)$ is piecewise monotone for a.e. $\boldsymbol{z} \in M$, and consequently monotone for a.e. $\boldsymbol{z} \in M$.

(σ2') Let $\boldsymbol{z} \in M$ be arbitrary but fixed such that all properties hold for $\mathsf{S}^\star(\boldsymbol{z}, \cdot)$, excluding zero sets where necessary. Also let $B \in \mathbb{R}_{\geq 0}$ and $k \geq k_0$ be arbitrary but fixed.
If $B \notin A^k$, then we have that $\mathsf{S}^k(\boldsymbol{z}, B) = \mathsf{S}^\star(\boldsymbol{z}, B)$ and hence the claim holds by the

estimate for $\boldsymbol{\mathcal{S}}^\star(\boldsymbol{z},\cdot)$ in Lemma 3.17 (S3), which is equivalent to the one for $\mathsf{S}^\star(\boldsymbol{z},\cdot)$, with the same g and c_*.
If $B \in A^k$, using Young's inequality with ε in the first term and applying the estimate corresponding to (S3) in Lemma 3.17 in the second term, and choosing $\varepsilon > 0$ small enough, we obtain

$$\begin{aligned}\mathsf{S}^k(\boldsymbol{z},B)B &= \Big(\mathsf{S}^k(\boldsymbol{z},B) - \mathsf{S}^\star(\boldsymbol{z},B)\Big)\, B + \mathsf{S}^\star(\boldsymbol{z},B)B \\ &\geq -\Big|\mathsf{S}^k(\boldsymbol{z},B) - \mathsf{S}^\star(\boldsymbol{z},B)\Big|\, |B| - g(\boldsymbol{z}) + c_*\left(|B|^q + |\mathsf{S}^\star(\boldsymbol{z},B)|^{q'}\right) \\ &\geq -c(\varepsilon,q)\Big|\mathsf{S}^k(\boldsymbol{z},B) - \mathsf{S}^\star(\boldsymbol{z},B)\Big|^{q'} - g(\boldsymbol{z}) + c(c_*,\varepsilon)\left(|B|^q + |\mathsf{S}^\star(\boldsymbol{z},B)|^{q'}\right).\end{aligned} \tag{3.29}$$

Also by the triangle inequality and the estimate (3.4) it follows that

$$\begin{aligned}\Big|\mathsf{S}^k(\boldsymbol{z},B)\Big|^{q'} &\leq \left(\Big|\mathsf{S}^k(\boldsymbol{z},B) - \mathsf{S}^\star(\boldsymbol{z},B)\Big| + |\mathsf{S}^\star(\boldsymbol{z},B)|\right)^{q'} \\ &\leq c(q)\left(\Big|\mathsf{S}^k(\boldsymbol{z},B) - \mathsf{S}^\star(\boldsymbol{z},B)\Big|^{q'} + |\mathsf{S}^\star(\boldsymbol{z},B)|^{q'}\right),\end{aligned} \tag{3.30}$$

which, applied in (3.29) yields

$$\mathsf{S}^k(\boldsymbol{z},B)B \geq -c\Big|\mathsf{S}^k(\boldsymbol{z},B) - \mathsf{S}^\star(\boldsymbol{z},B)\Big|^{q'} - g(\boldsymbol{z}) + c\left(|B|^q + \Big|\mathsf{S}^k(\boldsymbol{z},B)\Big|^{q'}\right), \tag{3.31}$$

for some constant $c > 0$, depending on q and c_*. Once the term $\big|\mathsf{S}^k(\boldsymbol{z},B) - \mathsf{S}^\star(\boldsymbol{z},B)\big|$ is estimated, the claim follows. Let us distinguish, whether (iia) or (iia) are satisfied:
If (iia) is satisfied, i.e., $J^\star$ is finite, we have that A^k is bounded and since $B \in A^k$ it follows that $B \leq a^k_{I,+}$. By Lemma 3.17 (S5) $\mathsf{S}^\star$ is locally bounded, and there exists a constant $C > 0$ such that $|\mathsf{S}^\star(\boldsymbol{z},\beta)| \leq C$ for any $\boldsymbol{z} \in M$ and any $\beta \in [0, a_I + 1]$, which applies in particular for B. Using this and the monotonicity of $\mathsf{S}^k(\boldsymbol{z},\cdot)$ we find that

$$\Big|\mathsf{S}^k(\boldsymbol{z},B)\Big| \leq \Big|\mathsf{S}^k(\boldsymbol{z},a^k_{I,+})\Big| = \Big|\mathsf{S}^\star(\boldsymbol{z},a^k_{I,+})\Big| \leq C. \tag{3.32}$$

Thus, we obtain

$$\Big|(\mathsf{S}^k(\boldsymbol{z},B) - \mathsf{S}^\star(\boldsymbol{z},B))\Big| \leq \Big|\mathsf{S}^k(\boldsymbol{z},B)\Big| + |\mathsf{S}^\star(\boldsymbol{z},B)| \leq 2C, \tag{3.33}$$

for $C > 0$ independent of $k \geq k_0$, $\boldsymbol{z} \in M$ and $B \in A^k$, which applied in (3.31) shows the claim in (σ2').
Now if (iib) is satisfied, i.e. $J^\star$ is countable and $\mathcal{I} = \mathbb{N}_0$ let $i \in \mathcal{I}$, such that $B \in A^k_i$ and choose $\widetilde{k}_0 \in \mathbb{N}$ large enough that $\widetilde{k}_0 > 2L$, with L the uniform bound on the ($\boldsymbol{z}$ and component dependent) Lipschitz constants. Let us consider the case when $B \in [a^k_{i,-}, a_i]$, which implies that $\mathsf{S}^k(\boldsymbol{z},B) \geq \mathsf{S}^\star(\boldsymbol{z},B)$: Indeed, by monotonicity and Lipschitz continuity for $\mathsf{S}^\star(\boldsymbol{z},\cdot)$ and by construction of $\mathsf{S}^k(\boldsymbol{z},\cdot)$ it follows that

$$\begin{aligned}&\mathsf{S}^\star(\boldsymbol{z},B) - \mathsf{S}^\star(\boldsymbol{z},a^k_{i,-}) \leq \|\nabla \mathsf{S}^\star(\boldsymbol{z},\cdot)\|_{\mathrm{L}^\infty((a^k_{i,-},a_i))}\,(B - a^k_{i,-}) \leq L(B - a^k_{i,-}), \\ &\mathsf{S}^k(\boldsymbol{z},B) - \mathsf{S}^k(\boldsymbol{z},a^k_{i,-}) = \tfrac{k}{2}(B - a^k_{i,-}),\end{aligned}$$

which implies that $\mathsf{S}^\star(\boldsymbol{z},B) \leq \mathsf{S}^k(\boldsymbol{z},B)$, since $\mathsf{S}^\star(\boldsymbol{z},a^k_{i,-}) = \mathsf{S}^k(\boldsymbol{z},a^k_{i,-})$. Let us denote by $\mathsf{S}_-(\boldsymbol{z},a_i) := \lim_{\delta\to 0}\mathsf{S}^\star(\boldsymbol{z},a_i - \delta)$ and by $\mathsf{S}_+(\boldsymbol{z},a_i) := \lim_{\delta\to 0}\mathsf{S}^\star(\boldsymbol{z},a_i + \delta)$ the values at a_i of the left- and right-continuous representatives of $\mathsf{S}^\star(\boldsymbol{z},\cdot)$, respectively, and recall that by (iib) we have that $|\mathsf{S}_+(\boldsymbol{z},a_i)-\mathsf{S}_-(\boldsymbol{z},a_i)|\le H$, independently of $\boldsymbol{z}\in M$ and of $i\in\mathcal{I}$. Then, using the fact that $\mathsf{S}^\star(\boldsymbol{z},B)\le\mathsf{S}^k(\boldsymbol{z},B)$, monotonicity of $\mathsf{S}^\star(\boldsymbol{z},\cdot)$, $\mathsf{S}^k(\boldsymbol{z},\cdot)$ and the fact that $\mathsf{S}^\star(\boldsymbol{z},a^k_{i,\pm})=\mathsf{S}^k(\boldsymbol{z},a^k_{i,\pm})$ we obtain

$$\begin{aligned}0&\le \mathsf{S}^k(\boldsymbol{z},B)-\mathsf{S}^\star(\boldsymbol{z},B)\le \mathsf{S}^k(\boldsymbol{z},a^k_{i,+})-\mathsf{S}^\star(\boldsymbol{z},a^k_{i,-})\\&\le \Big|\mathsf{S}^\star(\boldsymbol{z},a^k_{i,+})-\mathsf{S}_+(\boldsymbol{z},a_i)\Big|+H+\Big|\mathsf{S}_-(\boldsymbol{z},a^k_i)-\mathsf{S}^\star(\boldsymbol{z},a_{i,-})\Big|\\&\le \tfrac{1}{k}\Big(\|\nabla\mathsf{S}^\star(\boldsymbol{z},\cdot)\|_{\mathrm{L}^\infty((a^k_{i,-},a_i))}+\|\nabla\mathsf{S}^\star(\boldsymbol{z},\cdot)\|_{\mathrm{L}^\infty((a_i,a^k_{i,+}))}\Big)+H\le\tfrac{2L}{k}+H,\end{aligned}\tag{3.34}$$

and hence $\big|\mathsf{S}^k(\boldsymbol{z},B)-\mathsf{S}^\star(\boldsymbol{z},B)\big|$ is bounded, uniformly in $\boldsymbol{z}\in M$ and independently of $i\in\mathcal{I}$. The case, if $B\in[a_i,a^k_{i,+}]$ follows analogously, and then applying the estimate in (3.31) shows that (σ2') holds.

(σ3') Let $\{D^k\}_{k\in\mathbb{N}}$ be bounded in $\mathrm{L}^\infty(M)$, let $B\in\mathbb{R}_{\ge0}$ and let $\boldsymbol{z}\in M$ be arbitrary but fixed, possibly excluding zero sets. We wish to show that there exists a constant $c>0$ independent of $\boldsymbol{z}\in M$ such that

$$(\star):=\Big(\mathsf{S}^k(\boldsymbol{z},D^k(\boldsymbol{z}))-\mathsf{S}^\star(\boldsymbol{z},B)\Big):\Big(D^k(\boldsymbol{z})-B\Big)\ge-\tfrac{c}{k}\quad\text{for all }k\ge k_0.\tag{3.35}$$

Now let also $k\ge k_0$ be arbitrary but fixed, and distinguish the following three cases:
If $B\notin A^k$, then we have that $\mathsf{S}^k(\boldsymbol{z},B)=\mathsf{S}^\star(\boldsymbol{z},B)$ and by monotonicity of $\mathsf{S}^k(\boldsymbol{z},\cdot)$, it follows that $(\star)\ge0$.
Similarly, if $D^k(\boldsymbol{z})\notin A^k$, then we have $\mathsf{S}^k(\boldsymbol{z},D^k(\boldsymbol{z}))=\mathsf{S}^\star(\boldsymbol{z},D^k(\boldsymbol{z}))$ and by monotonicity of $\mathsf{S}^\star(\boldsymbol{z},\cdot)$ it again follows that $(\star)\ge0$.
Finally, if $B,D^k(\boldsymbol{z})\in A^k$, let $i,j\in\mathcal{I}$ be such that $B\in A^k_i$ and $D^k(\boldsymbol{z})\in A^k_j$.
If $i\ne j$, then we can use of the fact that $\mathsf{S}^k(\boldsymbol{z},\cdot)$ and $\mathsf{S}^\star(\boldsymbol{z},\cdot)$ agree at $a^k_{i,\pm}$, $a^k_{j,\pm}$ and are monotone. Assume that $i<j$, i.e., $B\le D^k(\boldsymbol{z})$; then we have that

$$\mathsf{S}^\star(\boldsymbol{z},B)\le\mathsf{S}^\star(\boldsymbol{z},a^k_{i,+})=\mathsf{S}^k(\boldsymbol{z},a^k_{i,+})\le\mathsf{S}^k(\boldsymbol{z},a^k_{j,-})\le\mathsf{S}^k(\boldsymbol{z},D^k(\boldsymbol{z})),$$

which implies that

$$\Big(\mathsf{S}^k(\boldsymbol{z},D^k(\boldsymbol{z}))-\mathsf{S}^\star(\boldsymbol{z},B)\Big):\Big(D^k(\boldsymbol{z})-B\Big)\ge0.$$

The case $i>j$ follows analogously.
If $i=j$, then we make use of the fact that $\big|D^k(\boldsymbol{z})-B\big|\le\frac{2}{k}$ and that $\mathsf{S}^\star$ is locally bounded by Lemma 3.17 (S5): Since $\{D^k\}_{k\in\mathbb{N}}$ is bounded in $\mathrm{L}^\infty(M)$ and $\big|D^k(\boldsymbol{z})-B\big|\le\frac{2}{k}$, there exists a constant $\widetilde{c}>0$ independent of $k\in\mathbb{N}$ and $\boldsymbol{z}\in M$ such that $\big|D^k(\boldsymbol{z})\big|+|B|\le\widetilde{c}$. Then, by the local boundedness of $\mathsf{S}^\star$ and the arguments in (3.32) it follows that

$$\Big|\mathsf{S}^k(\boldsymbol{z},D^k(\boldsymbol{z}))-\mathsf{S}^\star(\boldsymbol{z},B)\Big|\le\Big|\mathsf{S}^k(\boldsymbol{z},D^k(\boldsymbol{z}))\Big|+|\mathsf{S}^\star(\boldsymbol{z},B)|\le c(\widetilde{c}),$$

which yields with $\big|D^k(\boldsymbol{z})-B\big|\le\frac{2}{k}$ that

$$(\star)\ge-\Big(\Big|\mathsf{S}^k(\boldsymbol{z},D^k(\boldsymbol{z}))\Big|+|\mathsf{S}^\star(\boldsymbol{z},B)|\Big)\Big|D^k(\boldsymbol{z})-B\Big|\ge-\tfrac{c}{k},$$

where the constant $c>0$ is independent of $\boldsymbol{z}\in M$ and $k\in\mathbb{N}$.
Altogether, these imply (3.35). Multiplying by $\varphi\in C^\infty_0(M)$ such that $\varphi\ge0$ and

integrating over M yields

$$\int_M \Big(\mathsf{S}^k(\cdot, D^k) - \mathsf{S}^\star(\cdot, B)\Big) : \Big(D^k - B\Big)\, \varphi \,\mathrm{d}\boldsymbol{z} \geq -\frac{c}{k}\, \|\varphi\|_{\mathrm{L}^1(M)}\,,$$

for all $k \geq k_0$. Then applying $\liminf_{k\to\infty}$ yields the assertion.

□

Note that estimate (3.34) suggests, that we can weaken the assumptions on the Lipschitz constants of $\mathsf{S}^\star$, if we move away from constant width of the intervals $[a^k_{i,+}, a^k_{i,-}]$: As long as $\|\nabla \mathsf{S}^\star(\boldsymbol{z},\cdot)\|_{\mathrm{L}^\infty((a_i,a^k_{i,+}))} < c(k,i)$, for a.e. $\boldsymbol{z} \in M$, and $\Big|a_i - a^k_{i,+}\Big|\, \|\nabla \mathsf{S}^\star(\boldsymbol{z},\cdot)\|_{\mathrm{L}^\infty((a_i,a^k_{i,+}))}$ is bounded uniformly in $k \in \mathbb{N}$ and in $i \in \mathcal{I}$, and the corresponding holds for $a^k_{i,-}$, then the proof still works.

Corollary 3.24.
The family of functions $\boldsymbol{\mathcal{S}}^k\colon M \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$, $k \geq k_0$, defined by

$$\boldsymbol{\mathcal{S}}^k(\boldsymbol{z}, \mathbf{B}) \coloneqq \mathsf{S}^k(\boldsymbol{z}, |\mathbf{B}|)\frac{\mathbf{B}}{|\mathbf{B}|}, \qquad \text{for } \boldsymbol{z} \in M,\ \mathbf{B} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}\backslash\{\mathbf{0}\},$$

and $\boldsymbol{\mathcal{S}}^k(\boldsymbol{z},\mathbf{0}) \coloneqq \mathbf{0}$, where S^k is defined in (3.28), satisfies Assumption 3.18.

Proof. The fact that $\boldsymbol{\mathcal{S}}^k\colon M \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$ is a Carathéodory function follows directly from the fact that $\mathsf{S}^k\colon M \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is a Carathéodory function. It is straightforward to show that monotonicity of S^k implies monotonicity of $\boldsymbol{\mathcal{S}}^k$ (and by the same argument as for (σ3') implies that $\boldsymbol{\mathcal{S}}^k$ satisfies (σ3) in Assumption 3.18) and that the growth and coercivity bounds are equivalent for S^k and $\boldsymbol{\mathcal{S}}^k$. □

We have seen that in case one assumes radial symmetry and suitable additional regularity of the selection function $\boldsymbol{\mathcal{S}}^\star$ away from the jumps, approximations based on piecewise affine interpolation in the neighbourhood of the jumps satisfy the general Assumption 3.18. This approximation is likely to be easier to implement than the mollification-based one, but similarly as for the latter, there is no Minty type convergence result available.

3.4.2. Approximation of the Characterising Carathéodory Contraction

The starting point for the approximation investigated in this subsection is the equivalence between maximal monotone graphs and 1-Lipschitz functions through application of the Cayley transformation ϕ. Recall also that Lipschitz functions with Lipschitz constant smaller than 1 correspond to graphs of strictly monotone single-valued functions and this is in fact what we will aim for for the purpose of approximation.

This approximation was used in [FMT04, Sec. 3] for $q = 2$ in order to show existence to divergence-form equations with $\boldsymbol{z}$-dependent graph. Some of the properties in Lemma 3.26 can be found therein but for the sake of completeness we will reprove it here.

In the following we will first present the construction of the approximation, which can be done for any $q \in (1,\infty)$. Then we will show that the resulting sequence of approximations is strictly monotone and satisfies Assumption 3.18, however only for $q = 2$. And finally we will see in Lemma 3.27 that a convergence result of Minty type is satisfied.

Recall the equivalent characterisation of $\mathcal{A}$ by means of the Carathéodory contraction $\psi\colon M \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$: Assuming that $\mathcal{A}$ satisfies Assumption 3.11, there exists a (single-valued) ψ, which satisfies (p1)–(p4) in Lemma 3.14 and is related to $\mathcal{A}$ via

$$\mathcal{A}(\boldsymbol{z}) = \phi^{-1}(\Gamma(\psi(\boldsymbol{z},\cdot))) \qquad \text{for a.e. } \boldsymbol{z} \in M, \tag{3.36}$$

where ϕ is the Cayley transformation defined in (3.5).

Example 3.25 (Graph Approximation based on ψ)**.**
Let $\mathcal{A}$ satisfy Assumption 3.11 with $q = 2$ and let $\psi\colon M \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$ be the Carathéodory contraction from Lemma 3.14. Then, for $k \in \mathbb{N}$, let us define $\psi^k\colon M \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$, by

$$\psi^k(\boldsymbol{z}, \mathbf{B}) := \tfrac{k-1}{k}\psi(\boldsymbol{z}, \mathbf{B}) \qquad \text{for } \boldsymbol{z} \in M,\, \mathbf{B} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}. \tag{3.37}$$

Then let $\mathcal{A}^k\colon M \rightrightarrows \mathbb{R}^{d\times d}_{\mathrm{sym}} \times \mathbb{R}^{d\times d}_{\mathrm{sym}}$ be the set-valued mapping such that

$$\mathcal{A}^k(\boldsymbol{z}) = \phi^{-1}(\Gamma(\psi^k(\boldsymbol{z},\cdot))), \quad \boldsymbol{z} \in M, \tag{3.38}$$

and denote by $\mathcal{S}^k\colon M \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \rightrightarrows \mathbb{R}^{d\times d}_{\mathrm{sym}}$ the (set-valued) function defined by

$$\mathcal{S}^k(\boldsymbol{z}, \mathbf{D}) = \{\mathbf{S} \in \mathbb{R}^{d\times d}_{\mathrm{sym}} \colon (\mathbf{D}, \mathbf{S}) \in \mathcal{A}^k\}, \qquad \text{for } \boldsymbol{z} \in M,\, \mathbf{D} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}, \tag{3.39}$$

i.e., we have that $\Gamma(\mathcal{S}^k(\boldsymbol{x},\cdot)) = \mathcal{A}^k(\boldsymbol{z})$, for $\boldsymbol{z} \in M$.

Lemma 3.26 (Properties of $\mathcal{S}^k$, $k \in \mathbb{N}$, for $q = 2$)**.**
There exists a $k_0 \in \mathbb{N}$ such that the family of functions $\{\mathcal{S}^k\}_{k\in\mathbb{N}}$ defined in (3.39) satisfy Assumption 3.18 for $q = 2$ and $k \geq k_0$. Furthermore, for each $k \in \mathbb{N}$ and a.e. $\boldsymbol{z} \in M$ the function $\mathcal{S}^k(\boldsymbol{z},\cdot)$ is strictly monotone.

Proof. First note that $\psi^k(\boldsymbol{z}, \mathbf{0}) = \frac{k-1}{k}\psi(\boldsymbol{z}, \mathbf{0}) = \mathbf{0}$, for a.e. $\boldsymbol{z} \in M$ and for all $k \in \mathbb{N}$, so the property corresponding to (p1) holds.

By the definition of ψ^k in (3.37) and the fact that $\psi(\boldsymbol{z},\cdot)$ is 1-Lipschitz for a.e. $\boldsymbol{z} \in M$, it is obvious that $\psi^k(\boldsymbol{z},\cdot)$ is $\frac{k-1}{k}$-Lipschitz and $\frac{k-1}{k} \in (0,1)$ for $k \geq 2$. Hence, it is in particular 1-Lipschitz for all $k \in \mathbb{N}$ so the property corresponding to (p2) is satisfied.

Let $k \in \mathbb{N}$ and $\boldsymbol{z} \in M$ be arbitrary but fixed (possibly avoiding zero sets) and let $(\mathbf{D}, \mathbf{S}) \in \mathcal{A}^k(\boldsymbol{z}) = \phi^{-1}(\Gamma(\psi^k(\boldsymbol{z},\cdot)))$ be arbitrary. Then, by the definition of ϕ in (3.5) there exists an $m^k = m^k(\boldsymbol{z})$ such that

$$\sqrt{2}(\mathbf{D}, \mathbf{S}) = \left(m^k - \psi^k(\boldsymbol{z}, m^k), m^k + \psi^k(\boldsymbol{z}, m^k)\right).$$

Note also that $\frac{1}{\sqrt{2}}\left(m^k - \psi(\boldsymbol{z}, m^k), m^k + \psi(\boldsymbol{z}, m^k)\right) =: (\overline{\mathbf{D}}, \overline{\mathbf{S}}) \in \mathcal{A}(\boldsymbol{z})$, by (3.36). Thus, by Assumption 3.11 (A3), with $q = 2$, we have that

$$\overline{\mathbf{D}} : \overline{\mathbf{S}} \geq -g(\boldsymbol{z}) + c_*\left(\left|\overline{\mathbf{D}}\right|^2 + \left|\overline{\mathbf{S}}\right|^2\right). \tag{3.40}$$

Denoting, for $k \geq 2$,

$$\varepsilon^k(\cdot) := \psi(\cdot, m^k) - \psi^k(\cdot, m^k) = \frac{1}{k}\psi(\cdot, m^k) = \frac{1}{k-1}\psi^k(\cdot, m^k), \tag{3.41}$$

where we have used the definition of ψ^k in (3.37). We have that

$$\sqrt{2}(\mathbf{D} - \overline{\mathbf{D}}) = \varepsilon^k = -\sqrt{2}(\mathbf{S} - \overline{\mathbf{S}}). \tag{3.42}$$

Hence, it follows that

$$|\mathbf{D}|^2 + |\mathbf{S}|^2 \le c\left(\left|\overline{\mathbf{D}}\right|^2 + \left|\overline{\mathbf{S}}\right|^2 + \left|\varepsilon^k\right|^2\right). \tag{3.43}$$

Using the definitions of $(\mathbf{D},\mathbf{S})$, $(\overline{\mathbf{D}},\overline{\mathbf{S}})$ and ε^k and then applying the estimates (3.40) and (3.43) and finally that $\psi^k(\cdot,m^k) = \frac{k-1}{k}\psi(\cdot,m^k)$ we obtain that

$$\begin{aligned}
\mathbf{D} : \mathbf{S} &\overset{(3.42)}{=} \left(\overline{\mathbf{D}} + \tfrac{1}{\sqrt{2}}\varepsilon^k\right) : \left(\overline{\mathbf{S}} - \tfrac{1}{\sqrt{2}}\varepsilon^k\right) = \overline{\mathbf{D}} : \overline{\mathbf{S}} - \tfrac{1}{2}\left|\varepsilon^k\right|^2 + \varepsilon^k : \psi(\cdot,m^k) \\
&\overset{(3.40)}{\ge} -g(\boldsymbol{z}) + c_*\left(\left|\overline{\mathbf{D}}\right|^2 + \left|\overline{\mathbf{S}}\right|^2\right) - \tfrac{1}{2}\left|\varepsilon^k\right|^2 - \left|\varepsilon^k\right|\left|\psi(\cdot,m^k)\right| \\
&\overset{(3.37),(3.43)}{\ge} -g(\boldsymbol{z}) + c\left(|\mathbf{D}|^2 + |\mathbf{S}|^2\right) - \left(c\left|\varepsilon^k\right|^2 + \left|\varepsilon^k\right|\left|\psi^k(\cdot,m^k)\right|\right) \\
&\overset{(3.41)}{\ge} -g(\boldsymbol{z}) + c\left(|\mathbf{D}|^2 + |\mathbf{S}|^2\right) - c\frac{k}{(k-1)^2}\left|\psi^k(\cdot,m^k)\right|^2,
\end{aligned} \tag{3.44}$$

with constants depending on c_*, but not on $\boldsymbol{z} \in M$ or on $k \in \mathbb{N}$. Note that we have used that for example one has $\frac{k}{k-1} \le 2$, for $k \ge 2$, and hence $\left|\psi(\cdot,m^k)\right| \le c\left|\psi^k(\cdot,m^k)\right|$. With the fact that $\psi^k(\cdot,m^k) = \sqrt{2}(\mathbf{S}-\mathbf{D})$, and hence

$$\left|\psi^k(\cdot,m^k)\right|^2 = 2\,|\mathbf{S}|^2 + 2\,|\mathbf{D}|^2 - \mathbf{D} : \mathbf{S},$$

one can choose $k_0 \in \mathbb{N}$ large enough such that the resulting terms can be absorbed into the left-hand side and the term $(|\mathbf{D}|^2 + |\mathbf{S}|^2)$ on the right-hand side of (3.44) for all $k \ge k_0$. This finishes the proof of the property corresponding to (p3), with $q = 2$ and L^1-function and constant independent of $k \ge k_0$.

By the definition of $\mathcal{A}^k$ in (3.38) and Lemma 3.14 it follows that $\mathcal{A}^k$ satisfies the properties corresponding to (A1)–(A3), uniformly in $k \ge k_0$. Thus, $\mathcal{S}^k(\boldsymbol{z},\cdot)$ is monotone for a.e. $\boldsymbol{z} \in M$, so it satisfies Assumption 3.18 (σ1), and it satisfies the coercivity and boundedness property in (σ2).

Since ψ^k is a Carathéodory function for each $k \ge 2$ and $\mathcal{A}^k$ satisfies the corresponding properties to (A1)–(A3), the second part of Lemma 3.14 implies that each $\mathcal{A}^k$ satisfies (A4) and by the arguments in Remark 3.12 the (set-valued) function $\mathcal{S}^k$ is $\left(\mathcal{L}(M) \otimes \mathcal{B}(\mathbb{R}^{d\times d}_{\mathrm{sym}})\right) - \mathcal{B}(\mathbb{R}^{d\times d}_{\mathrm{sym}})$ measurable. Since $\psi^k(\boldsymbol{z},\cdot)$ is $\frac{k-1}{k}$-Lipschitz for a.e. $\boldsymbol{z} \in M$, this implies by Lemma 3.6 (iii) that $\mathcal{S}^k(\boldsymbol{z},\cdot)$ is strictly monotone and single-valued for all $k \ge 2$ and a.e. $\boldsymbol{z} \in M$. Together with the measurability this implies that each $\mathcal{S}^k$ is a Carathéodory function and strictly monotone. For a direct proof we refer to [FMT04, Lem. 3.1].

To show that also (σ3) in Assumption 3.18 is satisfied, let $\{\mathbf{D}^k\}_{k\in\mathbb{N}}$ with values in $\mathbb{R}^{d\times d}_{\mathrm{sym}}$ be bounded in $\mathrm{L}^\infty(M)^{d\times d}$ and in particular in $\mathrm{L}^2(M)^{d\times d}$. Then by ($\sigma$2) it follows that $\{\mathcal{S}^k(\cdot,\mathbf{D}^k)\}_{k\in\mathbb{N}}$ is bounded in $\mathrm{L}^2(\Omega)^{d\times d}$ as well. As before, there exists a sequence $\{m^k\}_{k\in\mathbb{N}} \subset \mathrm{L}^2(M)^{d\times d}$ with values in $\mathbb{R}^{d\times d}_{\mathrm{sym}}$ such that

$$\sqrt{2}\left(\mathbf{D}^k, \mathcal{S}^k(\cdot,\mathbf{D}^k)\right) = (m^k - \psi^k(\cdot,m^k), m^k + \psi^k(\cdot,m^k)), \tag{3.45}$$

and also $\{\psi^k(\cdot,m^k)\}_{k\in\mathbb{N}} \subset \mathrm{L}^2(M)^{d\times d}$ has values in $\mathbb{R}^{d\times d}_{\mathrm{sym}}$. Now let a constant matrix $\mathbf{B} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$ be arbitrary, but fixed. By the local boundedness of $\mathcal{S}^\star$ in Lemma 3.17 (S5) it follows that $\mathcal{S}^\star(\cdot,\mathbf{B}) \in \mathrm{L}^\infty(M)^{d\times d}$. Now by (3.36) there exists an $m \in \mathrm{L}^\infty(M)^{d\times d}$ depending on $\mathbf{B}$, such that $\psi(\cdot,m) \in \mathrm{L}^\infty(M)^{d\times d}$ and

$$\sqrt{2}(\mathbf{B}, \mathcal{S}^\star(\cdot,\mathbf{B})) = (m - \psi(\cdot,m), m + \psi(\cdot,m)). \tag{3.46}$$

For a.e. $\boldsymbol{z} \in M$ we have that $\psi^k(\boldsymbol{z},\cdot)$ is $\frac{k-1}{k}$-Lipschitz and by the definition of ψ^k in (3.37) this shows that pointwise for a.e. $\boldsymbol{z} \in M$ (suppressing the dependence on $\boldsymbol{z}$ in the following estimates) we have that

$$\begin{aligned}\left|\psi^k(\cdot,m^k) - \psi(\cdot,m)\right| &\le \left|\psi^k(\cdot,m^k) - \psi^k(\cdot,m)\right| + \left|\psi^k(\cdot,m) - \psi(\cdot,m)\right| \\ &\le \tfrac{k-1}{k}\left|m^k - m\right| + \tfrac{1}{k}\left|\psi(\cdot,m)\right|.\end{aligned} \tag{3.47}$$

Then, using Young's inequality with ε, we obtain

$$\begin{aligned}\left|\psi^k(\cdot,m^k) - \psi(\cdot,m)\right|^2 &\overset{(3.47)}{\le} \tfrac{(k-1)^2}{k^2}\left|m^k - m\right|^2 + \tfrac{1}{k^2}\left|\psi(\cdot,m)\right|^2 + 2\tfrac{k-1}{k^2}\left|m^k - m\right|\left|\psi(\cdot,m)\right| \\ &= \tfrac{(k-1)^2}{k^2}\left|m^k - m\right|^2 + \tfrac{1}{k^2}\left|\psi(\cdot,m)\right|^2 \\ &\qquad + 2\left(\tfrac{\sqrt{2k-1}}{k}\left|m^k - m\right|\right)\left(\tfrac{k-1}{k\sqrt{2k-1}}\left|\psi(\cdot,m)\right|\right) \\ &\le \left(\tfrac{(k-1)^2}{k^2} + \tfrac{2k-1}{k^2}\right)\left|m^k - m\right|^2 + \left(\tfrac{1}{k^2} + \tfrac{(k-1)^2}{k^2(2k-1)}\right)\left|\psi(\cdot,m)\right|^2 \\ &= \left|m^k - m\right|^2 + \tfrac{1}{2k-1}\left|\psi(\cdot,m)\right|^2.\end{aligned} \tag{3.48}$$

Using the identities (3.45), (3.46) and estimate (3.48) we obtain that

$$\begin{aligned}\left(\mathcal{S}^k(\cdot,\mathbf{D}^k) - \mathcal{S}^\star(\cdot,\mathbf{B})\right) : \left(\mathbf{D}^k - \mathbf{B}\right) &= \left(m^k - m + \psi^k(\cdot,m^k) - \psi(\cdot,m)\right) \\ &\qquad : \left(m^k - m - (\psi^k(\cdot,m^k) - \psi(\cdot,m))\right) \\ &= \left|m^k - m\right|^2 - \left|\psi^k(\cdot,m^k) - \psi(\cdot,m)\right|^2 \\ &\overset{(3.48)}{\ge} -\tfrac{1}{2k-1}\left|\psi(\cdot,m)\right|^2.\end{aligned} \tag{3.49}$$

Now for any $\varphi \in \mathrm{C}_0^\infty(M)$ such that $\varphi \ge 0$ this implies that

$$\begin{aligned}\int_M \left(\mathcal{S}^k(\cdot,\mathbf{D}^k) - \mathcal{S}^\star(\cdot,\mathbf{B})\right) : \left(\mathbf{D}^k - \mathbf{B}\right)\varphi(\cdot)\,\mathrm{d}\boldsymbol{z} &\overset{(3.49)}{\ge} -\tfrac{1}{2k-1}\int_M \left|\psi(\cdot,m)\right|^2\varphi(\cdot)\,\mathrm{d}\boldsymbol{z} \\ &\ge -\tfrac{1}{2k-1}\left\|\varphi\right\|_{\mathrm{L}^\infty(M)}\left\|\psi(\cdot,m)\right\|^2_{\mathrm{L}^2(M)}.\end{aligned} \tag{3.50}$$

Then, applying $\limsup_{k\to\infty}$ shows that

$$\limsup_{k\to\infty}\int_M \left(\mathcal{S}^k(\cdot,\mathbf{D}^k) - \mathcal{S}^\star(\cdot,\mathbf{B})\right) : \left(\mathbf{D}^k - \mathbf{B}\right)\varphi(\cdot)\,\mathrm{d}\boldsymbol{z} \ge 0, \tag{3.51}$$

since we have in particular that $\psi(\cdot,m) \in \mathrm{L}^2(M)^{d\times d}$ and it depends only on $\mathbf{B}$. Thus (σ3) in Assumption 3.18 is satisfied. □

Note that the coercivity and boundedness property in (σ2) is the one, which works for $q = 2$ only. Since the transformation ϕ has linear combinations of the arguments as values, this is no surprise.

The following lemma gives a convergence result of Minty type and makes use of the fact that we know how to measure the "difference" between $\mathcal{A}$ and $\mathcal{A}^k$ and hence we can apply Lemma 3.16. Even though it holds for general $p \in (1,\infty)$, we will be able to use it only for $p = 2$ unless weak convergence in the respective spaces is known by other means than the coercivity estimate in Lemma 3.26. Let us mention that if $p = 2$, then the statement of the lemma also follows from the arguments in [FMT04, Sec. 4, pp. 41–42].

Lemma 3.27 (Convergence Lemma for ψ-based graph approximation $\mathcal{A}^k$)**.**
Let $\mathcal{A}\colon M \rightrightarrows \mathbb{R}^{d\times d}_{\mathrm{sym}} \times \mathbb{R}^{d\times d}_{\mathrm{sym}}$ be such that $\mathcal{A}(\boldsymbol{z})$ is a maximal monotone set for a.e. $\boldsymbol{z} \in M$, i.e., it satisfies Assumption 3.11 (A2). Let $\{\mathcal{A}^k\}_{k\in\mathbb{N}}$ be a family of approximations of $\mathcal{A}$ given in Example 3.25. Assume that there are sequences $\{\mathbf{S}^k\}_{k\in\mathbb{N}}$ and $\{\mathbf{D}^k\}_{k\in\mathbb{N}}$ and there is a measurable set $\widetilde{M} \subset M$ such that

$$
\begin{aligned}
(\mathbf{D}^k(\boldsymbol{z}),\mathbf{S}^k(\boldsymbol{z})) \in \mathcal{A}^k(\boldsymbol{z}) \qquad & \textit{for a.e. } \boldsymbol{z}\in\widetilde{M} \textit{ and all } k\in\mathbb{N},\\
\mathbf{D}^k \rightharpoonup \mathbf{D} \qquad & \textit{weakly in } \mathrm{L}^p(\widetilde{M})^{d\times d}, \quad \textit{as } k\to\infty,\\
\mathbf{S}^k \rightharpoonup \mathbf{S} \qquad & \textit{weakly in } \mathrm{L}^{p'}(\widetilde{M})^{d\times d}, \quad \textit{as } k\to\infty,\\
\limsup_{k\to\infty}\left\langle \mathbf{S}^k,\mathbf{D}^k\right\rangle_{\widetilde{M}} \le \langle \mathbf{S},\mathbf{D}\rangle_{\widetilde{M}}. &
\end{aligned}
$$

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

Proof. As $(\mathbf{D}^k(\boldsymbol{z}),\mathbf{S}^k(\boldsymbol{z})) \in \mathcal{A}^k(\boldsymbol{z}) = \phi^{-1}(\Gamma(\psi^k(\boldsymbol{z},\cdot)))$ for a.e. $\boldsymbol{z}\in\widetilde{M}$, with the Cayley transformation ϕ in (3.5), for each $k\in\mathbb{N}$ there exists an $m^k \in \mathrm{L}^s(\widetilde{M})^{d\times d}$, for $s=\min(p,p')$, such that pointwise a.e. in $\widetilde{M}$ we have that

$$\sqrt{2}\left(\mathbf{D}^k,\mathbf{S}^k\right) = (m^k - \psi^k(\cdot,m^k), m^k + \psi^k(\cdot,m^k)),$$

where ψ^k is defined by (3.37). Since $\{\mathbf{D}^k\}_{k\in\mathbb{N}}$ and $\{\mathbf{S}^k\}_{k\in\mathbb{N}}$ are bounded in $\mathrm{L}^s(\widetilde{M})^{d\times d}$, for $s=\min(p,p')$, also $\{m^k\}_{k\in\mathbb{N}}$ and $\{\psi^k(\cdot,m^k)\}_{k\in\mathbb{N}}$ are bounded in $\mathrm{L}^s(\widetilde{M})^{d\times d}$.

Now let us introduce the sequence of functions $\{\overline{\mathbf{D}}^k\}_{k\in\mathbb{N}}, \{\overline{\mathbf{S}}^k\}_{k\in\mathbb{N}}$ on $\widetilde{M}$ and mapping to $\mathbb{R}^{d\times d}_{\mathrm{sym}}$, pointwise given by

$$\sqrt{2}(\overline{\mathbf{D}}^k,\overline{\mathbf{S}}^k) = (m^k - \psi(\cdot,m^k), m^k + \psi(\cdot,m^k)),$$

i.e., we have that $(\overline{\mathbf{D}}^k(\boldsymbol{z}),\overline{\mathbf{S}}^k(\boldsymbol{z})) \in \mathcal{A}(\boldsymbol{z})$ for a.e. $\boldsymbol{z}\in\widetilde{M}$. We aim to apply Lemma 3.16 to this sequence in order to show the claim. By the definition of ψ^k in (3.37) we obtain

$$
\begin{aligned}
\overline{\mathbf{D}}^k &= \mathbf{D}^k + \tfrac{1}{\sqrt{2}}\left(\psi^k(\cdot,m^k) - \psi(\cdot,m^k)\right) = \mathbf{D}^k - \tfrac{1}{\sqrt{2}(k-1)}\psi^k(\cdot,m^k),\\
\overline{\mathbf{S}}^k &= \mathbf{S}^k - \tfrac{1}{\sqrt{2}}\left(\psi^k(\cdot,m^k) - \psi(\cdot,m^k)\right) = \mathbf{S}^k + \tfrac{1}{\sqrt{2}(k-1)}\psi^k(\cdot,m^k).
\end{aligned}
\tag{3.52}
$$

By the convergence $\mathbf{D}^k \rightharpoonup \mathbf{D}$ and $\mathbf{S}^k \rightharpoonup \mathbf{S}$ weakly in $\mathrm{L}^s(\widetilde{M})^{d\times d}$, as $k\to\infty$, and by the fact that $\{\psi^k(\cdot,m^k)\}_{k\in\mathbb{N}}$ is bounded in $\mathrm{L}^s(\widetilde{M})^{d\times d}$, it follows that $\overline{\mathbf{D}}^k \rightharpoonup \mathbf{D}$ and $\overline{\mathbf{S}}^k \rightharpoonup \mathbf{S}$ weakly in $\mathrm{L}^s(\widetilde{M})^{d\times d}$, as $k\to\infty$. But by assumption we have that $\mathbf{D}\in\mathrm{L}^p(\widetilde{M})^{d\times d}$ and $\mathbf{S}\in\mathrm{L}^{p'}(\widetilde{M})^{d\times d}$, which means that we can improve the convergence to $\overline{\mathbf{D}}^k \rightharpoonup \mathbf{D}$ weakly in $\mathrm{L}^p(\widetilde{M})^{d\times d}$ and $\overline{\mathbf{S}}^k \rightharpoonup \mathbf{S}$ weakly in $\mathrm{L}^{p'}(\widetilde{M})^{d\times d}$, as $k\to\infty$.

Using again (3.52), the definition of $\psi^k(\cdot,m^k)$ and nonnegativity of the terms involving ψ^k

we have that

$$\begin{aligned}\left\langle \overline{\mathbf{D}}^k, \overline{\mathbf{S}}^k \right\rangle_{\widetilde{M}} &= \left\langle \mathbf{D}^k - \tfrac{1}{\sqrt{2}(k-1)}\psi^k(\cdot, m^k), \mathbf{S}^k + \tfrac{1}{\sqrt{2}(k-1)}\psi^k(\cdot, m^k) \right\rangle_{\widetilde{M}} \\ &= \left\langle \mathbf{D}^k, \mathbf{S}^k \right\rangle_{\widetilde{M}} - \tfrac{1}{2(k-1)^2} \left\| \psi^k(\cdot, m^k) \right\|^2_{\mathrm{L}^2(\widetilde{M})} - \tfrac{1}{(k-1)} \left\| \psi^k(\cdot, m^k) \right\|^2_{\mathrm{L}^2(\widetilde{M})} \\ &\leq \left\langle \mathbf{D}^k, \mathbf{S}^k \right\rangle_{\widetilde{M}} + 0.\end{aligned} \tag{3.53}$$

Then applying $\limsup_{k\to\infty}$ and using the assumption on $\langle \mathbf{S}^k, \mathbf{D}^k \rangle_{\widetilde{M}}$ yields that

$$\limsup_{k\to\infty} \left\langle \overline{\mathbf{D}}^k, \overline{\mathbf{S}}^k \right\rangle_{\widetilde{M}} \overset{(3.53)}{\leq} \limsup_{k\to\infty} \left\langle \mathbf{D}^k, \mathbf{S}^k \right\rangle_{\widetilde{M}} \leq \langle \mathbf{D}, \mathbf{S} \rangle_{\widetilde{M}}.$$

Now the sequences $\{\overline{\mathbf{D}}^k\}_{k\in\mathbb{N}}$, $\{\overline{\mathbf{S}}^k\}_{k\in\mathbb{N}}$ satisfy all the assumptions of Lemma 3.16 and hence we can conclude that $(\mathbf{D}(\boldsymbol{z}), \mathbf{S}(\boldsymbol{z})) \in \mathcal{A}(\boldsymbol{z})$ for a.e. $\boldsymbol{z} \in \widetilde{M}$. □

The graph approximation based on the characterising Carathéodory contraction satisfies both Assumption 3.18 under the restriction that $q = 2$, and a convergence lemma of Minty type. Additionally, the functions $\mathcal{S}^k$ are strictly monotone in the second argument, which is for example one of the assumptions for the adaptive approximation in [KS16].

3.4.3. Approximation of the Graph: Generalised Yosida Approximation

The so-called Yosida approximation yields a sequence of Lipschitz monotone operators for a maximal monotone operator on Hilbert spaces. Let us present the idea if the Hilbert space is $\mathbb{R}^s$, for $s \in \mathbb{N}$: For a maximal monotone (set-valued) function $\mathfrak{g}\colon \mathbb{R}^s \to \mathbb{R}^s$ the family of *Yosida approximations* $\mathfrak{g}^k\colon \mathbb{R}^s \rightrightarrows \mathbb{R}^s$, for $k \in \mathbb{N}$, given by

$$\left(\boldsymbol{x} + \tfrac{1}{k}\boldsymbol{y}, \boldsymbol{y}\right) \in \Gamma(\mathfrak{g}^k) \;\Leftrightarrow\; (\boldsymbol{x}, \boldsymbol{y}) \in \Gamma(\mathfrak{g}), \tag{3.54}$$

is single-valued and $\frac{1}{k}$-Lipschitz. Equivalently this can be written as $\mathfrak{g}^k = \left(\frac{1}{k}\mathbf{I} + (\mathfrak{g})^{-1}\right)^{-1}$, where $\mathbf{I}$ is the identity. For details let us refer to [AA99, Sec. 6] and note that we adopt the choice of names for the approximations from this reference.

The so-called *Moreau–Yosida approximation* gives the counter part of regularising proper lower semi-continuous convex function $f\colon \mathbb{R}^s \to \mathbb{R} \cup \{\infty\}$ by means of

$$f_k(\boldsymbol{x}) \coloneqq \inf_{\boldsymbol{y}\in\mathbb{R}^s} \left(f(\boldsymbol{y}) + \tfrac{k}{2} \left|\boldsymbol{x} - \boldsymbol{y}\right|^2 \right),$$

see [AA99, Def. 7.12]. For more details let us refer to [AA99, Sec. 7] and the references contained within. The connection between the two both is given by the fact that the sub-differential ∂f of a lower semicontinuous, proper convex function is maximal monotone by Lemma 3.8 and one has

$$\partial f_k = (\partial f)^k, \tag{3.55}$$

see [AA99, Prop. 7.13].

In [PR01a] the authors present a generalisation to Banach spaces with a regularising term corresponding to some potential instead of $\frac{1}{k}\left|\cdot\right|^2$. This generalisation proves to be fruitful to preserve coercivity and boundedness in an L^q-setting for $q \neq 2$.

Here we want to investigate a generalised Yosida approximation, which uses a "potential" regularisation as in [PR01a], but acts on functions with $\boldsymbol{x}$-dependence on the space $\mathbb{R}^s$, for $s \in \mathbb{N}$. This is a special case of [PR01a] in the sense that the space is not a general Banach space, but $\mathbb{R}^s$, and is an extension due to the $\boldsymbol{z}$-dependence. Even though the authors do not refer to it as (generalised) Yosida approximation, in [FMT04, Sec. 4] it is used to approximate maximal monotone graphs $\mathfrak{A} \subset \mathrm{L}^p(M)^s \times \mathrm{L}^{p'}(M)^s$. The class of graphs in question is more general than the class of $\boldsymbol{z}$-dependent graphs $\mathcal{A}$ satisfying Assumption 3.11, see also Remark 3.13. Thus our contribution is a special case of theirs and some of the properties in Lemma 3.29 can be directly deduced from [FMT04] by applying the arguments pointwise. However, for the reader's convenience we include the full proof. Note also that such an approximation was used in [BM16] for Bingham fluids, i.e., for the special case when $q = 2$.

In the following, we will first introduce the construction of the generalised Yosida approximation and derive some properties of the approximate graphs $\mathcal{A}^k$. Those allow us to show that the corresponding functions $\mathcal{S}^k$ satisfy the general Assumption 3.18 and a convergence result of Minty type for any $q \in (1, \infty)$, see Lemma 3.31.

The construction of the generalised Yosida approximation depends on $q \in (1, \infty)$, which appears in the coercivity and boundedness estimate for $\mathcal{A}$ in Assumption 3.11 (A3).

Example 3.28 (Generalised Yosida Approximation for $q \in (1, \infty)$).
For given $q \in (1, \infty)$ let

$$\gamma = \gamma(q) := \frac{2-q}{q-1} = \frac{1}{q-1} - 1 \in (-1, \infty), \tag{3.56}$$

and note that $\gamma > 0$ if and only if $q < 2$. Let $\mathcal{A}\colon M \rightrightarrows \mathbb{R}^{d\times d}_{\mathrm{sym}} \times \mathbb{R}^{d\times d}_{\mathrm{sym}}$ be a set-valued mapping satisfying Assumption 3.11 with q in (A3). For fixed $\boldsymbol{z} \in M$ and $k \in \mathbb{N}$ let us define the generalised Yosida approximation *of $\mathcal{A}(\boldsymbol{z})$ by*

$$\mathcal{A}^k(\boldsymbol{z}) := \left\{ \left(\mathbf{D} + \mathcal{E}^k(\mathbf{S}), \mathbf{S}\right) \in \mathbb{R}^{d\times d}_{\mathrm{sym}} \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \colon\ (\mathbf{D}, \mathbf{S}) \in \mathcal{A}(\boldsymbol{z}) \right\}, \tag{3.57}$$

where $\mathcal{E}^k\colon \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$ is defined by

$$\mathcal{E}^k(\mathbf{S}) = \begin{cases} \mathbf{0} & \text{if } \mathbf{S} = \mathbf{0}, \\ \frac{1}{k}\,|\mathbf{S}|^{\gamma}\,\mathbf{S}, & \text{if } \mathbf{S} \neq \mathbf{0}. \end{cases} \tag{3.58}$$

Furthermore let the set-valued function $\mathcal{S}^k\colon M \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \rightrightarrows \mathbb{R}^{d\times d}_{\mathrm{sym}}$ be defined by

$$\mathcal{S}^k(\boldsymbol{z}, \mathbf{D}) = \{\mathbf{S} \in \mathbb{R}^{d\times d}_{\mathrm{sym}} \colon\ (\mathbf{D}, \mathbf{S}) \in \mathcal{A}^k(\boldsymbol{z})\}, \tag{3.59}$$

for $\boldsymbol{z} \in M$ and $\mathbf{D} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$.

Denoting by $\mathfrak{s}[\boldsymbol{z}]\colon \mathbb{R}^{d\times d}_{\mathrm{sym}} \rightrightarrows \mathbb{R}^{d\times d}_{\mathrm{sym}}$ the set-valued function such that $\Gamma(\mathfrak{s}[\boldsymbol{z}]) = \mathcal{A}(\boldsymbol{z})$ we can write the function $\mathfrak{s}^k[\boldsymbol{z}]\colon \mathbb{R}^{d\times d}_{\mathrm{sym}} \rightrightarrows \mathbb{R}^{d\times d}_{\mathrm{sym}}$ such that $\Gamma(\mathfrak{s}^k[\boldsymbol{z}]) = \mathcal{A}^k(\boldsymbol{z})$ as

$$\mathfrak{s}^k[\boldsymbol{z}] = \left(\mathcal{E}^k + (\mathfrak{s}[\boldsymbol{z}])^{-1}\right)^{-1}, \tag{3.60}$$

and note that $\mathcal{S}(\boldsymbol{z}, \cdot) = \mathfrak{s}[\boldsymbol{z}]$ and $\mathcal{S}^k(\boldsymbol{z}, \cdot) = \mathfrak{s}^k[\boldsymbol{z}]$. This shows that to compute $\mathcal{S}^k$, one has to do two inversions, which may require some effort, see Example 3.32 below.

Note that for $q = 2$ we have that $\gamma = 0$, and hence $\mathcal{E}^k$ is linear and the approximation corresponds to (3.54). Furthermore, from the properties of the function $\mathcal{E} = k\mathcal{E}^k$ used in Lemma 3.15 it follows directly, that for each $k \in \mathbb{N}$ the function $\mathcal{E}^k$ is continuous, strictly monotone and $\operatorname{dom}\mathcal{E}^k = \mathbb{R}^{d\times d}_{\text{sym}}$, so it is in particular maximal monotone. Also it is invertible with inverse $(\mathcal{E}^k)^{-1}(\mathbf{B}) = k^{q-2}\,|\mathbf{B}|^{q-2}\,\mathbf{B}$, if $\mathbf{B} \neq \mathbf{0}$ and $(\mathcal{E}^k)^{-1}(\mathbf{0}) = \mathbf{0}$.

Before showing that the family of functions $\{\mathcal{S}^k\}_{k\in\mathbb{N}}$ defined in (3.59) satisfies Assumption 3.18 with respect to q, let us collect some properties of $\mathcal{A}^k$. Note that coercivity and monotonicity can also be shown by the arguments in the proof of [FMT04, Thm. 4.4].

Lemma 3.29 (Properties of the Generalised Yosida Approximation $\mathcal{A}^k$)**.**
For $q \in (1,\infty)$ let $\mathcal{A}\colon M \times \mathbb{R}^{d\times d}_{\text{sym}} \rightrightarrows \mathbb{R}^{d\times d}_{\text{sym}}$ satisfy Assumption 3.11 with respect to q. For $k \in \mathbb{N}$ let $\mathcal{A}^k\colon M \rightrightarrows \mathbb{R}^{d\times d}_{\text{sym}} \times \mathbb{R}^{d\times d}_{\text{sym}}$ be the set-valued mapping defined in (3.57). Then, for a.e. $\boldsymbol{z} \in M$ the following conditions are satisfied:

(α1) $(\mathbf{0},\mathbf{0}) \in \mathcal{A}^k(\boldsymbol{z})$;

(α2) $\mathcal{A}^k(\boldsymbol{z})$ is maximal monotone and it is the graph of a single-valued, continuous function.

(α3) There exists a constant $\widetilde{c}_*$, a nonnegative function $\widetilde{g} \in \mathrm{L}^1(M)$ and a $k_0 \in \mathbb{N}$, such that

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

for any $(\mathbf{D},\mathbf{S}) \in \mathcal{A}^k(\boldsymbol{z})$, for any $k \geq k_0$ and for a.e. $\boldsymbol{z} \in M$;

(α4) The set-valued map $\boldsymbol{z} \mapsto \mathcal{A}^k(\boldsymbol{z})$ is $\mathcal{L}(M) - \big(\mathcal{B}(\mathbb{R}^{d\times d}_{\text{sym}}) \otimes \mathcal{B}(\mathbb{R}^{d\times d}_{\text{sym}})\big)$ measurable.

Proof. Let $\boldsymbol{z} \in M$ be arbitrary (possibly excluding zero sets) but fixed.

(α1) Since $\mathcal{E}^k(\mathbf{0}) = \mathbf{0}$ and $(\mathbf{0},\mathbf{0}) \in \mathcal{A}(\boldsymbol{z})$ by Assumption 3.11, by the definition of $\mathcal{A}^k(\boldsymbol{z})$ in (3.57) it follows that $(\mathbf{0},\mathbf{0}) \in \mathcal{A}^k(\boldsymbol{z})$.

(α2) Since $\mathcal{A}(\boldsymbol{z})$ is maximal monotone by Assumption 3.11 (A2), so is $\mathcal{A}(\boldsymbol{z})^{-1}$, as well as $\mathcal{E}^k$. By the fact that $\operatorname{dom}(\mathcal{E}^k) = \mathbb{R}^{d\times d}_{\text{sym}}$ and by $\mathbf{0} \in \operatorname{dom}((\mathcal{A}(\boldsymbol{z}))^{-1})$ by Assumption 3.11 (A1) it follows that $\operatorname{int}\operatorname{dom}(\mathcal{E}^k) \cap \operatorname{dom}((\mathcal{A}(\boldsymbol{z}))^{-1}) \neq \emptyset$. By the definition of $\mathcal{A}^k$ in (3.57) we have that $(\mathcal{A}^k(\boldsymbol{z}))^{-1} = (\mathcal{A}(\boldsymbol{z}))^{-1} + \mathcal{E}^k$, when each of the graphs is viewed as a set-valued function. Thus Lemma 3.5 shows that $(\mathcal{A}^k(\boldsymbol{z}))^{-1}$ is maximal monotone, and consequently also $\mathcal{A}^k(\boldsymbol{z})$ is maximal monotone.
Also since $(\mathcal{A}^k(\boldsymbol{z}))^{-1}$ is a sum of a (set-valued) monotone and a (single-valued) strictly monotone function, it is the graph of a strictly monotone function. Then, $\mathcal{A}^k(\boldsymbol{z})$ is the graph of a single-valued function.
Because $\mathcal{A}^k(\boldsymbol{z})$ is the graph of a single-valued and maximal monotone function, by Lemma 3.4 (iii) the function, $\mathcal{A}^k(\boldsymbol{z})$ is the graph of, is continuous.

(α3) Let $(\mathbf{D},\mathbf{S}) \in \mathcal{A}^k(\boldsymbol{z})$. Then, by the definition in (3.57) we have that $(\mathbf{D} - \mathcal{E}^k(\mathbf{S}),\mathbf{S}) \in \mathcal{A}(\boldsymbol{z})$. If $\mathbf{S} \neq \mathbf{0}$, by Assumption 3.11 (A3) we obtain

$$\begin{aligned}\mathbf{D} : \mathbf{S} - \tfrac{1}{k}\,|\mathbf{S}|^{q'} = \mathbf{D} : \mathbf{S} - \tfrac{1}{k}\,|\mathbf{S}|^{\gamma+2} &= \left(\mathbf{D} - \mathcal{E}^k(\mathbf{S})\right) : \mathbf{S} \\ &\geq -g(\boldsymbol{z}) + c_* \left(\left|\mathbf{D} - \mathcal{E}^k(\mathbf{S})\right|^q + |\mathbf{S}|^{q'}\right),\end{aligned} \tag{3.61}$$

where $g \in \mathrm{L}^1(M)$ is the nonnegative function and $c_* > 0$ the constant given in Assumption 3.11, both of which are independent of $k \in \mathbb{N}$, $\boldsymbol{z} \in M$ and $(\mathbf{D},\mathbf{S})$. By the triangle inequality, applying the inequality (3.4) and using the definition of $\mathcal{E}^k$ in (3.58) and of $\gamma(q)$, we have that

$$|\mathbf{D}|^q \leq c(q)\left(\left|\mathbf{D} - \mathcal{E}^k(\mathbf{S})\right|^q + \left|\mathcal{E}^k(\mathbf{S})\right|^q\right) = c(q)\left(\left|\mathbf{D} - \mathcal{E}^k(\mathbf{S})\right|^q + \tfrac{1}{k^q}\,|\mathbf{S}|^{q'}\right). \tag{3.62}$$

Rearranging and applying this in (3.61) yields

$$\begin{aligned}\mathbf{D} : \mathbf{S} &\overset{(3.61)}{\geq} -g(\boldsymbol{z}) + c_* \left(\left| \mathbf{D} - \mathcal{E}^k(\mathbf{S}) \right|^q + |\mathbf{S}|^{q'} \right) + \tfrac{1}{k} |\mathbf{S}|^{q'} \\ &\overset{(3.62)}{\geq} -g(\boldsymbol{z}) + c \left(|\mathbf{D}|^q + \left(1 - \tfrac{1}{k^q} + \tfrac{1}{k}\right) |\mathbf{S}|^{q'} \right),\end{aligned} \tag{3.63}$$

with constant $c > 0$ depending on c_* and q. Now one can see that choosing $k_0 \in \mathbb{N}$ large enough, there exists a constant $\widetilde{c}_* > 0$, independent of $k \geq k_0$ and $\boldsymbol{z} \in M$, such that (α3) is satisfied with $\widetilde{g} = g$.

If $\mathbf{S} = \mathbf{0}$, then $\mathbf{D} = \mathbf{0}$, since $\mathcal{A}^k(\boldsymbol{z})$ is the graph of a single-valued function by (α2) and $(\mathbf{0}, \mathbf{0}) \in \mathcal{A}^k(\boldsymbol{z})$ by (A1). Hence, the estimates holds trivially.

(α4) The measurability of $\boldsymbol{z} \mapsto \mathcal{A}^k(\boldsymbol{z})$ follows using Corollary 3.10 several times: First we want to apply it to the set-valued function $\boldsymbol{z} \mapsto (\mathcal{A}(\boldsymbol{z}))^{-1}$. By the definition of the inverse of a set, we have that $(\mathbf{0}, \mathbf{0}) \in (\mathcal{A}(\boldsymbol{z}))^{-1}$ and that $(\mathcal{A}(\boldsymbol{z}))^{-1}$ is maximal monotone for a.e. $\boldsymbol{z} \in M$. The measurability of $\boldsymbol{z} \mapsto \mathcal{A}(\boldsymbol{z})$ is equivalent to measurability of $\boldsymbol{z} \mapsto (\mathcal{A}(\boldsymbol{z}))^{-1}$, so with $\mathcal{L}(M) - \mathcal{B}(\mathbb{R}^{d\times d}_{\text{sym}}) \otimes \mathcal{B}(\mathbb{R}^{d\times d}_{\text{sym}})$ measurability of the latter, Corollary 3.10 implies that the set-valued function $\mathfrak{D}\colon M \times \mathbb{R}^{d\times d}_{\text{sym}} \rightrightarrows \mathbb{R}^{d\times d}_{\text{sym}}$, defined by

$$\mathfrak{D}(\boldsymbol{z}, \mathbf{S}) = \{\mathbf{B} \in \mathbb{R}^{d\times d}_{\text{sym}} \colon (\mathbf{S}, \mathbf{B}) \in (\mathcal{A}(\boldsymbol{z}))^{-1}\}, \qquad \text{for} \qquad (\boldsymbol{z}, \mathbf{S}) \in M \times \mathbb{R}^{d\times d}_{\text{sym}},$$

is $\big(\mathcal{L}(M) \otimes \mathcal{B}(\mathbb{R}^{d\times d}_{\text{sym}})\big) - \mathcal{B}(\mathbb{R}^{d\times d}_{\text{sym}})$ measurable. Since $\mathcal{E}^k\colon M \times \mathbb{R}^{d\times d}_{\text{sym}} \to \mathbb{R}^{d\times d}_{\text{sym}}$ defined in (3.58) is constant in $\boldsymbol{z}$, and continuous in $\mathbf{S}$ in particular $\big(\mathcal{L}(M) \otimes \mathcal{B}(\mathbb{R}^{d\times d}_{\text{sym}})\big) - \mathcal{B}(\mathbb{R}^{d\times d}_{\text{sym}})$ measurable. Consequently, the sum $\mathfrak{D}^k = \mathfrak{D} + \mathcal{E}^k\colon M \times \mathbb{R}^{d\times d}_{\text{sym}} \to \mathbb{R}^{d\times d}_{\text{sym}}$, mapping $(\boldsymbol{z}, \mathbf{S}) \mapsto \mathfrak{D}(z, \mathbf{S}) + \mathcal{E}^k(\mathbf{S})$ is $\big(\mathcal{L}(M) \otimes \mathcal{B}(\mathbb{R}^{d\times d}_{\text{sym}})\big) - \mathcal{B}(\mathbb{R}^{d\times d}_{\text{sym}})$ measurable.

Now we apply Corollary 3.10 to the function $\boldsymbol{z} \mapsto (\mathcal{A}^k(\boldsymbol{z}))^{-1}$, which again has the properties that $(\mathbf{0}, \mathbf{0}) \in (\mathcal{A}^k(\boldsymbol{z}))^{-1}$ and $(\mathcal{A}^k(\boldsymbol{z}))^{-1}$ is maximal monotone for a.a. $\boldsymbol{z} \in M$, since the same holds for $\mathcal{A}^k(\boldsymbol{z})$ by (α1) and (α2). Noting that $\mathfrak{D}^k\colon M \times \mathbb{R}^{d\times d}_{\text{sym}} \to \mathbb{R}^{d\times d}_{\text{sym}}$ satisfies that $\mathfrak{D}^k(\boldsymbol{z}, \mathbf{S}) = \{\mathbf{B} \in \mathbb{R}^{d\times d}_{\text{sym}} \colon (\mathbf{S}, \mathbf{B}) \in (\mathcal{A}^k(\boldsymbol{z}))^{-1}\}$, Corollary 3.10 shows that $\boldsymbol{z} \mapsto (\mathcal{A}^k(\boldsymbol{z}))^{-1}$ is $\mathcal{L}(M) - \big(\mathcal{B}(\mathbb{R}^{d\times d}_{\text{sym}}) \otimes \mathcal{B}(\mathbb{R}^{d\times d}_{\text{sym}})\big)$ measurable. Then we also have that $\boldsymbol{z} \mapsto \mathcal{A}^k(\boldsymbol{z})$ is $\mathcal{L}(M) - \big(\mathcal{B}(\mathbb{R}^{d\times d}_{\text{sym}}) \otimes \mathcal{B}(\mathbb{R}^{d\times d}_{\text{sym}})\big)$ measurable.

□

Lemma 3.30 (Properties of $\mathcal{S}^k$, $q \in (1, \infty)$).
For each $k \in \mathbb{N}$ the function $\mathcal{S}^k\colon M \times \mathbb{R}^{d\times d}_{\text{sym}} \to \mathbb{R}^{d\times d}_{\text{sym}}$, as defined in (3.59) is a (single-valued) Carathéodory function, and there is a $\widetilde{g} \in \mathrm{L}^1(M)$, a constant $\widetilde{c}_* > 0$ and a $k_0 \in \mathbb{N}$, such that Assumption 3.18 is satisfied for all $k \geq k_0$.

Furthermore, for any function $\mathbf{B}\colon M \to \mathbb{R}^{d\times d}_{\text{sym}}$ in $\mathrm{L}^\infty(M)^{d\times d}$, one has

$$\Big(\mathbf{B}(\boldsymbol{z}) - \mathcal{E}^k(\mathcal{S}^k(\boldsymbol{z}, \mathbf{B}(\boldsymbol{z}))), \mathcal{S}^k(\boldsymbol{z}, \mathbf{B}(\boldsymbol{z}))\Big) \in \mathcal{A}(\boldsymbol{z}), \tag{3.64}$$

for a.a. $\boldsymbol{z} \in M$ and any $k \in \mathbb{N}$, with $\mathcal{E}^k$ as defined in (3.58).

Proof. The properties of $\mathcal{S}^k$ rely on the properties of $\mathcal{A}^k$ according to Lemma 3.29: The fact that $\mathcal{A}^k$ is the graph of a single-valued, continuous function by (α2), shows that $\mathcal{S}^k$ is single-valued and continuous in the second argument. Since $\boldsymbol{z} \mapsto \mathcal{A}^k(\boldsymbol{z})$ is also $\big(\mathcal{L}(M) \otimes \mathcal{B}(\mathbb{R}^{d\times d}_{\text{sym}})\big) - \mathcal{B}(\mathbb{R}^{d\times d}_{\text{sym}})$ measurable by (α4), it follows that $\mathcal{S}^k$ is a Carathéodory function. The monotonicity of $\mathcal{S}^k$ in the sense of Assumption 3.18 (σ1) follows from monotonicity of $\mathcal{A}^k(\boldsymbol{z})$ in (α2) and the estimate in 3.18 (σ1) follows from (α3). The relation (3.64) follows directly from the definition of $\mathcal{A}^k$ in (3.57).

The only thing left to show is Assumption 3.18 (σ3). For this let $\{\mathbf{D}^k\}_{k\in\mathbb{N}}$ be a sequence, bounded in $\mathrm{L}^\infty(M)^{d\times d}$ with values in $\mathbb{R}^{d\times d}_{\mathrm{sym}}$. Note first, that by the local boundedness of the selection $\mathcal{S}^\star\colon M\times\mathbb{R}^{d\times d}_{\mathrm{sym}}\to\mathbb{R}^{d\times d}_{\mathrm{sym}}$ given in Lemma 3.17, there exists a constant $C>0$, such that

$$\left|\mathcal{S}^\star(\boldsymbol{z},\mathbf{D}^k(\boldsymbol{z}))\right|\le C,\qquad \text{for all } k\in\mathbb{N} \text{ and a.e. } \boldsymbol{z}\in M. \tag{3.65}$$

Furthermore, by the coercivity estimate for $\mathcal{S}^\star$ in Lemma 3.17 (S3) and for $\mathcal{S}^k$ in Assumption 3.18 (σ2), which we have proven to hold for $\mathcal{S}^k$, $k\ge k_0$, one can show that there exists a nonnegative function $\widehat{g}\in\mathrm{L}^1(M)$ and a constant $\widehat{c}_*>0$, such that

$$\left|\mathcal{S}^k(\boldsymbol{z},\mathbf{B})\right|^{q'}\le\widehat{g}(\boldsymbol{z})+\widehat{c}_*\left|\mathcal{S}^\star(\boldsymbol{z},\mathbf{B})\right|^{q'}, \tag{3.66}$$

for all $k\ge k_0$, all $\mathbf{B}\in\mathbb{R}^{d\times d}_{\mathrm{sym}}$ and a.e. $\boldsymbol{z}\in M$. Indeed, using Young's inequality with $\varepsilon>0$ in the estimates

$$\begin{aligned}\mathcal{S}^k(\boldsymbol{z},\mathbf{B}):\mathbf{B}&\ge-g(\boldsymbol{z})+\widetilde{c}_*(|\mathbf{B}|^q+\left|\mathcal{S}^k(\boldsymbol{z},\mathbf{B})\right|^{q'}),\\ \mathcal{S}^\star(\boldsymbol{z},\mathbf{B}):\mathbf{B}&\ge-g(\boldsymbol{z})+c_*(|\mathbf{B}|^q+|\mathcal{S}^\star(\boldsymbol{z},\mathbf{B})|^{q'}),\end{aligned}$$

which hold for all $\mathbf{B}\in\mathbb{R}^{d\times d}_{\mathrm{sym}}$ and a.e. $\boldsymbol{z}\in M$, allow us to show that

$$\begin{aligned}\left|\mathcal{S}^k(\boldsymbol{z},\mathbf{B})\right|^{q'}&\le c\left(|\mathbf{B}|^q+g(\boldsymbol{z})\right),\\ |\mathbf{B}|^q&\le c\left(|\mathcal{S}^\star(\boldsymbol{z},\mathbf{B})|^{q'}+g(\boldsymbol{z})\right),\end{aligned}$$

and these together imply (3.66). Then, with (3.65) and (3.66) we have that

$$\begin{aligned}\left|\mathcal{S}^k(\boldsymbol{z},\mathbf{D}^k(\boldsymbol{z}))\right|&\le 1+\left|\mathcal{S}^k(\boldsymbol{z},\mathbf{D}^k(\boldsymbol{z}))\right|\le\left(1+\left|\mathcal{S}^k(\boldsymbol{z},\mathbf{D}^k(\boldsymbol{z}))\right|\right)^{q'}\\ &\le c\left(1+\left|\mathcal{S}^k(\boldsymbol{z},\mathbf{D}^k(\boldsymbol{z}))\right|^{q'}\right)\overset{(3.66)}{\le}c\left(1+\widehat{g}(\boldsymbol{z})+\left|\mathcal{S}^\star(\boldsymbol{z},\mathbf{D}^k(\boldsymbol{z}))\right|^{q'}\right)\\ &\overset{(3.65)}{\le}c\left(1+\widehat{g}(\boldsymbol{z})\right),\end{aligned} \tag{3.67}$$

uniformly in $k\in\mathbb{N}$ such that $k\ge k_0$, and for a.e. $\boldsymbol{z}\in M$.

Now let $\boldsymbol{z}\in M$ be arbitrary, but for now fixed. For arbitrary $\mathbf{B}\in\mathbb{R}^{d\times d}_{\mathrm{sym}}$, we have that $(\mathbf{B},\mathcal{S}^\star(\boldsymbol{z},\mathbf{B}))\in\mathcal{A}(\boldsymbol{z})$, and thus $(\mathbf{B}+\mathcal{E}^k(\mathcal{S}^\star(\boldsymbol{z},\mathbf{B}),\mathcal{S}^\star(\boldsymbol{z},\mathbf{B}))\in\mathcal{A}^k(\boldsymbol{z})$, by the definition of $\mathcal{A}^k(\boldsymbol{z})$ in (3.57). By the monotonicity of $\mathcal{A}^k(\boldsymbol{z})$ according to Lemma 3.29 (α2), it follows that

$$\left(\mathcal{S}^k(\boldsymbol{z},\mathbf{D}^k(\boldsymbol{z}))-\mathcal{S}^\star(\boldsymbol{z},\mathbf{B})\right):\left(\mathbf{D}^k(\boldsymbol{z})-\mathbf{B}-\mathcal{E}^k(\mathcal{S}^\star(\boldsymbol{z},\mathbf{B}))\right)\ge 0. \tag{3.68}$$

Using this and the definition of $\mathcal{E}^k$ in (3.58), for $\mathbf{B}\in\mathbb{R}^{d\times d}_{\mathrm{sym}}$ and for $\boldsymbol{z}\in M$ such that $\mathcal{S}^\star(\boldsymbol{z},\mathbf{B})\ne\mathbf{0}$, we have that

$$\begin{aligned}&\left(\mathcal{S}^k(\boldsymbol{z},\mathbf{D}^k(\boldsymbol{z}))-\mathcal{S}^\star(\boldsymbol{z},\mathbf{B})\right):\left(\mathbf{D}^k(\boldsymbol{z})-\mathbf{B}\right)\\ &\overset{(3.68)}{\ge}\left(\mathcal{S}^k(\boldsymbol{z},\mathbf{D}^k(\boldsymbol{z}))-\mathcal{S}^\star(\boldsymbol{z},\mathbf{B})\right):\mathcal{E}^k(\mathcal{S}^\star(\boldsymbol{z},\mathbf{B}))\end{aligned}$$

$$
\begin{aligned}
&\geq \tfrac{1}{k}\left(|\mathcal{S}^\star(\boldsymbol{z},\mathbf{B})|^\gamma\, \mathcal{S}^k(\boldsymbol{z},\mathbf{D}^k(\boldsymbol{z})) : \mathcal{S}^\star(\boldsymbol{z},\mathbf{B}) - |\mathcal{S}^\star(\boldsymbol{z},\mathbf{B})|^{\gamma+2}\right)\\
&\geq -\tfrac{1}{k}\left(|\mathcal{S}^\star(\boldsymbol{z},\mathbf{B})|^{\gamma+1}\left|\mathcal{S}^k(\boldsymbol{z},\mathbf{D}^k(\boldsymbol{z}))\right| + |\mathcal{S}^\star(\boldsymbol{z},\mathbf{B})|^{\gamma+2}\right) \qquad (3.69)\\
&\overset{(3.67)}{\geq} -\tfrac{c}{k}\,|\mathcal{S}^\star(\boldsymbol{z},\mathbf{B})|^{\gamma+1}\,(1+\widehat{g}(\boldsymbol{z}) + |\mathcal{S}^\star(\boldsymbol{z},\mathbf{B})|)\\
&\geq -\tfrac{c(\mathbf{B})}{k}(1+\widehat{g}(\boldsymbol{z})),
\end{aligned}
$$

where we have used the fact that $\mathcal{S}^\star$ is locally bounded, by Lemma 3.17 (S5), and that $\gamma+1>0$. If $\mathbf{B}\in\mathbb{R}^{d\times d}_{\mathrm{sym}}$, and $\boldsymbol{z}\in M$ such that $\mathcal{S}^\star(\boldsymbol{z},\mathbf{B})=\mathbf{0}$, we have $\mathcal{E}^k(\mathcal{S}^\star(\boldsymbol{z},\mathbf{B}))=\mathbf{0}$ and the expression on the left-hand side of (3.69) is nonnegative.

For any nonnegative $\varphi\in C_0^\infty(M)$, multiplying the expression in (3.69) with φ and integrating over M yields

$$
\begin{aligned}
\int_M \left(\mathcal{S}^k(\cdot,\mathbf{D}^k) - \mathcal{S}^\star(\cdot,\mathbf{B})\right) : \left(\mathbf{D}^k - \mathbf{B}\right)\varphi \,\mathrm{d}\boldsymbol{z} &\overset{(3.69)}{\geq} -\frac{c(\mathbf{B})}{k}\int_M (1+\widehat{g})\varphi\,\mathrm{d}\boldsymbol{z}\\
&\geq -\frac{1}{k}c(\mathbf{B},\varphi,\|\widehat{g}\|_{\mathrm{L}^1(M)}),
\end{aligned}
$$

for all $k\in\mathbb{N}$ for $k\geq k_0$. Applying $\liminf_{k\to\infty}$ shows that

$$
\liminf_{k\to\infty}\int_M \left(\mathcal{S}^k(\cdot,\mathbf{D}^k) - \mathcal{S}^\star(\cdot,\mathbf{B})\right) : \left(\mathbf{D}^k - \mathbf{B}\right)\varphi\,\mathrm{d}\boldsymbol{z} \geq 0, \qquad (3.70)
$$

for any $\mathbf{B}\in\mathbb{R}^{d\times d}_{\mathrm{sym}}$ and any $\varphi\in C_0^\infty(M)$, such that $\varphi\geq 0$, which proves (σ3) in Assumption 3.18. ☐

Let us now show a convergence lemma of Minty type for general $q\in(1,\infty)$, which relies on the relation (3.64).

Lemma 3.31 (Localised Convergence Lemma for Generalised Yosida Approximation). *Let the set-valued mapping $\mathcal{A}\colon M \rightrightarrows \mathbb{R}^{d\times d}_{\mathrm{sym}}\times\mathbb{R}^{d\times d}_{\mathrm{sym}}$ be such that $\mathcal{A}(\boldsymbol{z})$ is a maximal monotone set for a.e. $\boldsymbol{z}\in M$, i.e., it satisfies Assumption 3.11 (A2). For $q\in(1,\infty)$ let $\mathcal{A}^k(\cdot)_{k\in\mathbb{N}}$ be a family of generalised Yosida approximations of $\mathcal{A}(\cdot)$ (depending on q) given by (3.57). Assume that there are sequences $\{\mathbf{S}^k\}_{k\in\mathbb{N}}$ and $\{\mathbf{D}^k\}_{k\in\mathbb{N}}$ and there is a measurable set $\widetilde{M}\subset M$ such that*

$$
\begin{aligned}
(\mathbf{D}^k(\boldsymbol{z}),\mathbf{S}^k(\boldsymbol{z}))&\in\mathcal{A}^k(\boldsymbol{z}) && \text{for a.e. } \boldsymbol{z}\in\widetilde{M} \text{ and all } k\in\mathbb{N},\\
\mathbf{D}^k &\rightharpoonup \mathbf{D} && \text{weakly in } \mathrm{L}^q(\widetilde{M})^{d\times d}, \text{ as } k\to\infty,\\
\mathbf{S}^k &\rightharpoonup \mathbf{S} && \text{weakly in } \mathrm{L}^{q'}(\widetilde{M})^{d\times d}, \text{ as } k\to\infty,\\
\limsup_{k\to\infty}\left\langle \mathbf{S}^k,\mathbf{D}^k\right\rangle_{\widetilde{M}} &\leq \langle\mathbf{S},\mathbf{D}\rangle_{\widetilde{M}}.
\end{aligned}
$$

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

Proof. Recall the definition of $\mathcal{A}^k$ in (3.57):

$$
\mathcal{A}^k(\boldsymbol{z}) := \left\{\left(\mathbf{D}+\mathcal{E}^k(\mathbf{S}),\mathbf{S}\right)\in\mathbb{R}^{d\times d}_{\mathrm{sym}}\times\mathbb{R}^{d\times d}_{\mathrm{sym}}\colon\ (\mathbf{D},\mathbf{S})\in\mathcal{A}(\boldsymbol{z})\right\},
$$

where $\mathcal{E}^k(\mathbf{S}) = \frac{1}{k}|\mathbf{S}|^{\gamma(q)}\mathbf{S}$ if $\mathbf{S}\in\mathbb{R}^{d\times d}_{\mathrm{sym}}\backslash\{\mathbf{0}\}$, for $\gamma(q)=\frac{1}{q-1}-1$, and $\mathcal{E}^k(\mathbf{0})=\mathbf{0}$. The fact that $(\mathbf{D}^k(\boldsymbol{z}),\mathbf{S}^k(\boldsymbol{z}))\in\mathcal{A}^k(\boldsymbol{z})$, for a.e. $\boldsymbol{z}\in M$, implies that $(\mathbf{D}^k(\boldsymbol{z})-\mathcal{E}^k(\mathbf{S}^k(\boldsymbol{z})),\mathbf{S}^k(\boldsymbol{z}))\in\mathcal{A}(\boldsymbol{z})$ for a.a. $\boldsymbol{z} \in M$ and all $k \in \mathbb{N}$. By the definition of $\mathcal{E}^k$ we have that

$$\left\|\mathcal{E}^k(\mathbf{S}^k)\right\|^q_{\mathrm{L}^q(\widetilde{M})} = \frac{1}{k^q}\int \left|\mathbf{S}^k\right|^{(\gamma+1)q} = \frac{1}{k^q}\left\|\mathbf{S}^k\right\|^{q'}_{\mathrm{L}^{q'}(\widetilde{M})} \le \frac{c}{k^q} \to 0,$$

as $k \to \infty$, since $\{\mathbf{S}^k\}_{k\in\mathbb{N}}$ is bounded in $\mathrm{L}^{q'}(\widetilde{M})^{d\times d}$. Thus, it follows with the weak convergence $\mathbf{D}^k \rightharpoonup \mathbf{D}$ in $\mathrm{L}^{q'}(\widetilde{M})^{d\times d}$, that

$$\mathbf{D}^k - \mathcal{E}^k(\mathbf{S}^k(\boldsymbol{z})) \rightharpoonup \mathbf{D}, \qquad \text{weakly in } \mathrm{L}^{q'}(\widetilde{M})^{d\times d}, \qquad \text{as } k \to \infty.$$

Furthermore, by the assumptions and since $\{\mathbf{S}^k\}_{k\in\mathbb{N}}$ is bounded in $\mathrm{L}^{q'}(\widetilde{M})^{d\times d}$, we obtain

$$\begin{aligned}
&\limsup_{k\to\infty}\left\langle \mathbf{D}^k - \mathcal{E}^k(\mathbf{S}^k(\boldsymbol{z})), \mathbf{S}^k\right\rangle_{\widetilde{M}} \\
&\quad \le \limsup_{k\to\infty}\left\langle \mathbf{D}^k, \mathbf{S}^k\right\rangle_{\widetilde{M}} - \liminf_{k\to\infty}\left\langle \mathcal{E}^k(\mathbf{S}^k(\boldsymbol{z})), \mathbf{S}^k\right\rangle_{\widetilde{M}} \\
&\quad \le \langle \mathbf{D}, \mathbf{S}\rangle_{\widetilde{M}} - \liminf_{k\to\infty}\frac{1}{k}\left\|\mathbf{S}^k\right\|^{q'}_{\mathrm{L}^{q'}(\widetilde{M})} = \langle \mathbf{D}, \mathbf{S}\rangle_{\widetilde{M}}.
\end{aligned}$$

Now, we can apply Lemma 3.16 to the sequences $\{\mathbf{D}^k - \mathcal{E}^k(\mathbf{S}^k(\boldsymbol{z}))\}_{k\in\mathbb{N}}$ and $\{\mathbf{S}^k\}_{k\in\mathbb{N}}$, which yields that $(\mathbf{D}(\boldsymbol{z}), \mathbf{S}(\boldsymbol{z})) \in \mathcal{A}(\boldsymbol{z})$ for a.e. $\boldsymbol{z} \in \widetilde{M}$. □

Since the generalised Yosida approximation is the most promising graph approximation let us show how to compute $\mathcal{S}^k$ explicitly for $\mathcal{A}$ encoding the implicit constitutive relation for Herschel–Bulkley fluids, see Lemma 3.15. We will find that we cannot perform the inversion explicitly for most choices of $q \in (1, \infty)$, so computational inversion by Newton's method or other means come to the fore.

Example 3.32 (Generalised Yosida Approximation for Herschel–Bulkley Fluids)**.**

(a) Let us consider first the graph $\widetilde{\mathcal{A}} \subset \mathbb{R}^{d\times d}_{\mathrm{sym},0} \times \mathbb{R}^{d\times d}_{\mathrm{sym},0}$, with which the implicit constitutive relation can be identified when working with divergence-free functions only. Recall from (3.12) that for a given $\tau_* > 0$ the constitutive relation is determined by

$$(\mathbf{D}, \mathbf{S}) \in \widetilde{\mathcal{A}} \Leftrightarrow \begin{cases} \mathbf{D} = \mathbf{0} \Leftrightarrow |\mathbf{S}| \le \tau_*, \\ \mathbf{D} \ne \mathbf{0} \Leftrightarrow |\mathbf{S}| > \tau_* \Leftrightarrow \mathbf{D} = (|\mathbf{S}| - \tau_*)^{\frac{1}{q-1}}\frac{\mathbf{S}}{|\mathbf{S}|}, \end{cases} \tag{3.71}$$

As before in (3.57) we introduce

$$\widetilde{\mathcal{A}}^k := \left\{\left(\mathbf{D} + \mathcal{E}^k(\mathbf{S}), \mathbf{S}\right) \in \mathbb{R}^{d\times d}_{\mathrm{sym},0} \times \mathbb{R}^{d\times d}_{\mathrm{sym},0} \colon (\mathbf{D}, \mathbf{S}) \in \widetilde{\mathcal{A}}(\boldsymbol{z})\right\}, \tag{3.72}$$

which is essentially a restriction of $\mathcal{A}^k$ to $\mathbb{R}^{d\times d}_{\mathrm{sym},0} \times \mathbb{R}^{d\times d}_{\mathrm{sym},0}$.
We aim for a characterisation of $\widetilde{\mathcal{S}}^k \colon \mathbb{R}^{d\times d}_{\mathrm{sym},0} \to \mathbb{R}^{d\times d}_{\mathrm{sym},0}$ such that $\Gamma(\widetilde{\mathcal{S}}^k) = \widetilde{\mathcal{A}}^k$. Recall that $(\mathbf{D}, \mathbf{S}) \in \widetilde{\mathcal{A}}^k$ if and only if $(\mathbf{D} - \mathcal{E}^k(\mathbf{S}), \mathbf{S}) \in \widetilde{\mathcal{A}}$. By (3.71) this means that if $|\mathbf{S}| \le \tau_*$, then $\mathbf{D} - \mathcal{E}^k(\mathbf{S}) = \mathbf{0}$, which is equivalent to $\mathbf{S} = (\mathcal{E}^k)^{-1}(\mathbf{D})$, if $\mathbf{D} \in \mathbb{R}^{d\times d}_{\mathrm{sym},0}$.
On the other hand, if $|\mathbf{S}| > \tau_*$, i.e., in particular $\mathbf{S} \ne 0$, then we have that

$$\mathbf{D} - \mathcal{E}^k(\mathbf{S}) = (|\mathbf{S}| - \tau_*)^{\frac{1}{q-1}}\frac{\mathbf{S}}{|\mathbf{S}|},$$

which by the definition of $\mathcal{E}^k$ *in* (3.58) *and with* $\gamma+1=\frac{1}{q-1}$, *is equivalent to*

$$\mathsf{D}=\left((|\mathsf{S}|-\tau_*)^{\frac{1}{q-1}}+\frac{1}{k}|\mathsf{S}|^{\frac{1}{q-1}}\right)\frac{\mathsf{S}}{|\mathsf{S}|}=:\mathsf{F}_q^k(\mathsf{S}). \tag{3.73}$$

To invert F_q^k *it is enough to consider the absolute values, since* S *and* D *are (scalar) multiples of each other. Hence, one has to solve*

$$|\mathsf{D}|=(|\mathsf{S}|-\tau_*)^{\frac{1}{q-1}}+\frac{1}{k}|\mathsf{S}|^{\frac{1}{q-1}}=:f_q^k(|\mathsf{S}|), \tag{3.74}$$

for $|\mathsf{S}|$, *i.e., to invert* f_q^k. *In case* $\frac{1}{q-1}\in\{\frac{1}{2},1,2,3,4\}$, *which is equivalent to* $q\in\{\frac{5}{4},\frac{4}{3},\frac{3}{2},2,3\}$, *this is possible explicitly, since it comes down to solving a polynomial equation of degree* ≤ 4. *For other* $q\in(1,\infty)$ *one has to use computational methods to invert the function.*
Overall, the function $\widetilde{\mathsf{S}}^k:\mathbb{R}^{d\times d}_{\mathrm{sym},0}\to\mathbb{R}^{d\times d}_{\mathrm{sym},0}$ *is then given by,*

$$\widetilde{\mathsf{S}}^k(\mathsf{D}):=\begin{cases}(\mathcal{E}^k)^{-1}(\mathsf{D}) & \text{if } |\mathsf{D}|\le\frac{1}{k}\tau_*^{q'/q},\\ (\mathsf{F}_q^k)^{-1}(\mathsf{D}) & \text{else,}\end{cases} \tag{3.75}$$

and recall that $(\mathcal{E}^k)^{-1}(\mathsf{D})=k^{q-2}|\mathsf{D}|^{q-2}\mathsf{D}$, *if* $\mathsf{D}\in\mathbb{R}^{d\times d}_{\mathrm{sym},0}\backslash\{\mathbf{0}\}$ *and* $(\mathcal{E}^k)^{-1}(\mathbf{0})=\mathbf{0}$.

(b) Now let us consider the graph $\mathcal{A}\subset\mathbb{R}^{d\times d}_{\mathrm{sym}}\times\mathbb{R}^{d\times d}_{\mathrm{sym}}$, *with which the implicit constitutive relation can be identified in the general case, defined by*

$$(\mathsf{D},\mathsf{S})\in\mathcal{A}\;\Leftrightarrow\begin{cases}|\mathsf{S}_\delta|\le\tau_*\;\Leftrightarrow\;\mathsf{D}=\mathcal{E}(\mathsf{T}(\mathsf{S})),\\ |\mathsf{S}_\delta|>\tau_*\;\Leftrightarrow\;\mathsf{D}=(|\mathsf{S}_\delta|-\tau_*)^{\frac{1}{q-1}}\frac{\mathsf{S}_\delta}{|\mathsf{S}_\delta|}+\mathcal{E}(\mathsf{T}(\mathsf{S})).\end{cases} \tag{3.76}$$

where $\tau_*>0$ *is given and* $\mathcal{E}$ *is defined in* (3.14), *see* (3.18). *In this case, instead of regularising with* $\mathcal{E}^k(\mathsf{S})$, *as in* (3.57), *it is more convenient to regularise with*

$$\overline{\mathcal{E}}^k(\mathsf{S}):=\mathcal{E}^k(\mathsf{S}_\delta)+\mathcal{E}^k(\mathsf{T}(\mathsf{S}))\qquad\text{for } \mathsf{S}=\mathsf{S}_\delta+\mathsf{T}(\mathsf{S})\in\mathbb{R}^{d\times d}_{\mathrm{sym}}, \tag{3.77}$$

which splits the regularisation of the trace part and of the deviatoric part. The approximate graph

$$\overline{\mathcal{A}}^k:=\left\{\left(\mathsf{D}+\overline{\mathcal{E}}^k(\mathsf{S}),\mathsf{S}\right)\in\mathbb{R}^{d\times d}_{\mathrm{sym}}\times\mathbb{R}^{d\times d}_{\mathrm{sym}}:\;(\mathsf{D},\mathsf{S})\in\mathcal{A}(\boldsymbol{z})\right\}, \tag{3.78}$$

has the same properties as $\mathcal{A}^k$, *defined in* (3.57), *since the spaces* $\mathbb{R}^{d\times d}_{\mathrm{sym},0}$ *and* $\mathbb{R}^{d\times d}_{\mathrm{sym,tr}}$ *are orthogonal, see* (3.13).
We want to find $\overline{\mathsf{S}}^k:\mathbb{R}^{d\times d}_{\mathrm{sym}}\to\mathbb{R}^{d\times d}_{\mathrm{sym}}$, *corresponding to* $\overline{\mathcal{A}}^k$. *By* (3.78) *we have that* $(\mathsf{D},\mathsf{S})\in\overline{\mathcal{A}}^k$ *if and only if* $(\mathsf{D}-\overline{\mathcal{E}}^k(\mathsf{S}),\mathsf{S})\in\mathcal{A}$. *By* (3.76) *this means, if* $|\mathsf{S}_\delta|\le\tau_*$, *that*

$$\mathcal{E}(\mathsf{T}(\mathsf{S}))\overset{(3.76)}{=}\mathsf{D}-\overline{\mathcal{E}}^k(\mathsf{S})=\mathsf{D}_\delta+\mathsf{T}(\mathsf{D})-\mathcal{E}^k(\mathsf{S}_\delta)-\mathcal{E}^k(\mathsf{T}(\mathsf{S})),$$

which is equivalent to

$$\begin{cases}\mathsf{D}_\delta &=\mathcal{E}^k(\mathsf{S}_\delta),\\ \mathsf{T}(\mathsf{D}) &=\mathcal{E}(\mathsf{T}(\mathsf{S}))+\mathcal{E}^k(\mathsf{T}(\mathsf{S}))=(k+1)\mathcal{E}^k(\mathsf{T}(\mathsf{S})),\end{cases} \tag{3.79}$$

since $\mathcal{E}$ *and* $\mathcal{E}^k$ *map* $\mathbb{R}^{d\times d}_{\mathrm{sym},0}$ *to* $\mathbb{R}^{d\times d}_{\mathrm{sym},0}$ *and* $\mathbb{R}^{d\times d}_{\mathrm{sym,tr}}$ *to* $\mathbb{R}^{d\times d}_{\mathrm{sym,tr}}$. *Note also that by the definition* $\mathcal{E} = k\mathcal{E}^k$. *Since* $\mathcal{E}^k$ *is invertible, this means that* $\mathbf{S} = \mathbf{S}_\delta + \mathbf{T}(\mathbf{S}) = (\mathcal{E}^k)^{-1}(\mathbf{D}_\delta) + (\mathcal{E}^k)^{-1}(\frac{1}{k+1}\mathbf{T}(\mathbf{D}))$.
If on the other hand $|\mathbf{S}_\delta| > \tau_*$, *then* (3.76) *shows that*

$$(|\mathbf{S}_\delta| - \tau_*)^{\frac{1}{q-1}}\frac{\mathbf{S}_\delta}{|\mathbf{S}_\delta|} + \mathcal{E}(\mathsf{T}(\mathbf{S})) \overset{(3.76)}{=} \mathbf{D} - \overline{\mathcal{E}}^k(\mathbf{S}) = \mathbf{D}_\delta + \mathbf{T}(\mathbf{D}) - \mathcal{E}^k(\mathbf{S}_\delta) - \mathcal{E}^k(\mathbf{T}(\mathbf{S})).$$

Similarly as before, this is equivalent to

$$\begin{cases}\mathbf{D}_\delta = \mathcal{E}^k(\mathbf{S}_\delta) + (|\mathbf{S}_\delta| - \tau_*)^{\frac{1}{q-1}}\frac{\mathbf{S}_\delta}{|\mathbf{S}_\delta|},\\ \mathbf{T}(\mathbf{D}) = \mathcal{E}(\mathbf{T}(\mathbf{S})) + \mathcal{E}^k(\mathbf{T}(\mathbf{S})) = (k+1)\mathcal{E}^k(\mathbf{T}(\mathbf{S})).\end{cases}$$

The first equation corresponds to (3.73) *and the second equation corresponds to the second equation in* (3.79). *So then we have that* $\mathbf{S} = \mathbf{S}_\delta + \mathbf{T}(\mathbf{S}) = (\mathbf{F}^k_q)^{-1}(\mathbf{D}_\delta) + (\mathcal{E}^k)^{-1}(\frac{1}{k+1}\mathbf{T}(\mathbf{D}))$. *This means that the function* $\overline{\mathcal{S}}^k : \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$ *is given by*

$$\overline{\mathcal{S}}^k(\mathbf{D}) := \begin{cases}(\mathcal{E}^k)^{-1}(\mathbf{D}_\delta) + (\mathcal{E}^k)^{-1}(\frac{1}{k+1}\mathbf{T}(\mathbf{D})) & \text{if } |\mathbf{D}_\delta| \le \frac{1}{k}\tau_*^{q'/q},\\ (\mathbf{F}^k_q)^{-1}(\mathbf{D}_\delta) + (\mathcal{E}^k)^{-1}(\frac{1}{k+1}\mathbf{T}(\mathbf{D})) & \text{otherwise.}\end{cases} \tag{3.80}$$

Recall that the inversion of $\mathcal{E}^k$ *is explicitly known, but for* $\boldsymbol{F}^k_q$ *an explicit inverse is guaranteed only for special choices of* q, *see (a). Otherwise* $(\mathbf{F}^k_q)^{-1}$ *has to be approximated computationally.*

Even though the calculation of $\mathcal{S}^k$ requires computational inversion in most cases, the approximation is very useful in the sense that a Minty type convergence lemma is available for all $q \in (1,\infty)$.

3.5. Summary and Discussion

In this chapter we have introduced and discussed the detailed assumptions on the graph $\mathcal{A}(\boldsymbol{z})$, $\boldsymbol{z} \in M$, in Assumption 3.11. In particular we have seen that adding a condition on the trace yields constitutive relations, which are automatically consistent with the thermodynamic framework.

Then, we have introduced general assumptions on graph approximations in Assumption 3.18, which includes a large class of approximations and in particular all the examples presented here. These assumptions allow us to take the limit in the graph approximation first and separately from the discretisation limit in the convergence proof.

Finally we have presented a range of examples of graph approximation and have investigated their properties. This includes approximations of the selection function $\mathcal{S}^\star$ in Subsection 3.4.1, based on mollification and on piecewise affine interpolation, the approximation of the characterising Carathéodory function ψ in Subsection 3.4.2 and the generalised Yosida approximation in 3.4.3. For both the selection-based approximations and the Yosida approximation we were able to show Assumption 3.18 for all $q \in (1,\infty)$, whereas for the ψ-based approximation we succeeded only for $q = 2$. Furthermore, we have proved convergence lemmas of Minty type for the ψ-based approximation and for the generalised Yosida approximation.

In the following Chapters 4 and 5 we will consider two situations for the convergence of a sequence of numerical approximations to the steady and the unsteady problem, respectively: The first is an approximation, for which the graph approximation satisfies Assumption 3.18

and the limit $k \to \infty$ is taken first and separately. And the second one uses the generalised Yosida approximation and Lemma 3.31, which allows us to take the limit $k \to \infty$ together with the discretisation limit. The same could be done using the ψ-based approximation if $q = 2$, but we will omit the details for this.

CHAPTER 4

Steady Case Revisited

In this chapter we want to revisit and improve the convergence results for the steady problem, which were obtained in [DKS13a, KS16]. The main ingredient of the existing convergence proof is a discrete Lipschitz truncation method. Convergence of a sequence of finite element approximations to a weak solution was shown for $q > \frac{2d}{d+1}$, if discretely divergence-free finite element spaces are used, and for the whole range of existence $q > \frac{2d}{d+2}$, if exactly divergence-free finite element spaces are used.

Here we want to showcase some approaches to improving the previous result, which turn out to be helpful also in the unsteady situation in Chapter 5. This includes on the one hand a regularisation procedure allowing us to cover the whole range of existence $q > \frac{2d}{d+2}$ also for discretely divergence-free finite element spaces, which in addition avoids a discrete Lipschitz approximation. On the other hand we utilise a generalised Yosida graph approximation presented in Section 3.4 to simplify the result in [DKS13a] and achieve a slight extension with respect to the assumptions on the graph $\mathcal{A}$.

First, in Section 4.1 we recall the problem and the notion of weak solution we are aiming for. Furthermore, we recall and summarise the relevant notation and the estimates on the convective term used in this chapter. Then, in Section 4.2 we introduce the levels of approximation and give a more detailed outline of the subsequent convergence proofs in Section 4.3: in Subsection 4.3.1 we present the approach using a regularising term and in Subsection 4.3.2 the Yosida approximation is employed for the unregularised problem.

4.1. The Steady Problem and Notation

Let us recall the steady problem and introduce the notion of weak solution we are interested in, cf. Section 2.1. Furthermore, we want to recall the estimates on the convective term and its numerical approximation from Subsection 2.2.2, which are relevant for this chapter.

Statement of the Problem

Let $\Omega \subset \mathbb{R}^d$, with $d \geq 2$, be a bounded Lipschitz domain and let $\boldsymbol{f}\colon \Omega \to \mathbb{R}^d$ be a given external force. We seek a velocity field $\boldsymbol{u}\colon \overline{\Omega} \to \mathbb{R}^d$, a pressure $\pi\colon \Omega \to \mathbb{R}$, and a trace-free stress tensor field $\mathbf{S}\colon \Omega \to \mathbb{R}^{d\times d}_{\mathrm{sym},0}$ satisfying the balance law of linear momentum and the incompressibility condition:

$$\begin{aligned} \operatorname{div}(\boldsymbol{u}\otimes\boldsymbol{u}) - \operatorname{div}\mathbf{S} &= -\nabla\pi + \boldsymbol{f} && \text{on } \Omega,\\ \operatorname{div}\boldsymbol{u} &= 0 && \text{on } \Omega, \end{aligned} \tag{4.1}$$

subject to the homogeneous Dirichlet boundary condition:

$$\boldsymbol{u} = \mathbf{0} \qquad \text{on } \partial\Omega. \tag{4.2}$$

The *constitutive law* is encoded by means of a maximal monotone graph $\mathcal{A}(\boldsymbol{x}) \subset \mathbb{R}^{d\times d}_{\text{sym}} \times \mathbb{R}^{d\times d}_{\text{sym}}$, for $\boldsymbol{x} \in \Omega$, via

$$(\mathbf{D}\boldsymbol{u}(\boldsymbol{x}), \mathbf{S}(\boldsymbol{x})) \in \mathcal{A}(\boldsymbol{x}), \tag{4.3}$$

where $\mathcal{A}(\cdot)$ satisfies Assumption 3.11. In the following we will refer to the problem consisting of (4.1)–(4.3) as **(PS)**.

For $q \in (1,\infty)$ and q^* its critical Sobolev exponent let us denote

$$\widetilde{q} = \begin{cases} \frac{q^*}{2} & \text{if } q \in \left[\frac{2d}{d+2}, \frac{3d}{d+2}\right), \\ q' & \text{if } q \in \left[\frac{3d}{d+2}, \infty\right), \end{cases} \tag{4.4}$$

$$\overline{q} := \min\left(\frac{2qd}{q(d-2)+d}, q'\right), \tag{4.5}$$

where $\widetilde{q}$ was mentioned in Section 2.2.2 and we have that $\widetilde{q} > 1$ if and only if $q > \frac{2d}{d+2}$.

Definition 4.1 (Weak Solution to Problem **(PS)**)**.**
*Let $\Omega \subset \mathbb{R}^d$, with $d \in \{2,3\}$, be a bounded Lipschitz domain. Furthermore, assume that $q \in (1,\infty)$ is given and let $\mathcal{A}(\cdot) \subset \mathbb{R}^{d\times d}_{\text{sym}} \times \mathbb{R}^{d\times d}_{\text{sym}}$ be a maximal monotone graph satisfying Assumption 3.11 with respect to $q \in (1,\infty)$ for $M = \Omega$. For a given $\boldsymbol{f} \in \mathrm{W}^{-1,q'}(\Omega)^d$ we call $(\boldsymbol{u}, \mathbf{S}, \pi,)$ a weak solution to problem **(PS)**, if*

$$\boldsymbol{u} \in \mathrm{W}^{1,q}_{0,\text{div}}(\Omega)^d, \quad \mathbf{S} \in \mathrm{L}^{q'}(\Omega)^{d\times d}, \quad \pi \in \mathrm{L}^{\widetilde{q}}_0(\Omega),$$

and

$$-\langle \boldsymbol{u} \otimes \boldsymbol{u}, \mathbf{D}\boldsymbol{v}\rangle_\Omega + \langle \mathbf{S}, \mathbf{D}\boldsymbol{v}\rangle_\Omega - \langle \operatorname{div} \boldsymbol{v}, \pi\rangle_\Omega = \langle \boldsymbol{f}, \boldsymbol{v}\rangle_\Omega \quad \textit{for all } \boldsymbol{v} \in \mathrm{C}^\infty_0(\Omega)^d, \tag{4.6}$$

$$(\mathbf{D}\boldsymbol{u}(\boldsymbol{x}), \mathbf{S}(\boldsymbol{x})) \in \mathcal{A}(\boldsymbol{x}) \quad \textit{for a.e. } \boldsymbol{x} \in \Omega. \tag{4.7}$$

In Section 2.2.2 for the weak form of the convective term and its numerical modification we adopted the notation

$$b(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) := -\langle \boldsymbol{u} \otimes \boldsymbol{v}, \nabla \boldsymbol{w}\rangle_\Omega, \tag{4.8}$$

$$\widetilde{b}(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) := \tfrac{1}{2}\left(\langle \boldsymbol{u} \otimes \boldsymbol{w}, \nabla \boldsymbol{v}\rangle_\Omega - \langle \boldsymbol{u} \otimes \boldsymbol{v}, \nabla \boldsymbol{w}\rangle_\Omega\right), \tag{4.9}$$

for $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \mathrm{W}^{1,\infty}_0(\Omega)^d$. Recall also the notation

$$\mathrm{X}(\Omega) := \mathrm{W}^{1,q}_0(\Omega)^d \cap \mathrm{L}^{2q'}(\Omega)^d, \qquad \text{with } \|\cdot\|_{\mathrm{X}(\Omega)} := \|\cdot\|_{\mathrm{W}^{1,q}(\Omega)} + \|\cdot\|_{\mathrm{L}^{2q'}(\Omega)}, \tag{4.10}$$

and that $\mathrm{X}_{\text{div}}(\Omega)$ denotes the subspace of solenoidal functions in $\mathrm{X}(\Omega)$. Note that we have the following embedding

$$\mathrm{W}^{1,\widetilde{q}'}(\Omega) \hookrightarrow \mathrm{W}^{1,\overline{q}'}(\Omega) \hookrightarrow \mathrm{W}^{1,q}(\Omega) \cap \mathrm{L}^{2q'}(\Omega), \tag{4.11}$$

for $\widetilde{q}$ and $\overline{q}$ as defined in (4.4) and (4.5), respectively. In this chapter we use the following estimates:

$$|\langle \boldsymbol{u} \otimes \boldsymbol{u}, \nabla \boldsymbol{v}\rangle_\Omega| \leq c\, \|\boldsymbol{u}\|^2_{\mathrm{W}^{1,q}(\Omega)} \|\boldsymbol{v}\|_{\mathrm{W}^{1,\widetilde{q}'}(\Omega)}\,, \quad \text{provided that } q \geq \tfrac{2d}{d+2}, \tag{4.12}$$

$$|\langle \boldsymbol{u} \otimes \boldsymbol{v}, \nabla \boldsymbol{u}\rangle_\Omega| \leq c\, \|\boldsymbol{u}\|^2_{\mathrm{W}^{1,q}(\Omega)} \|\boldsymbol{v}\|_{\mathrm{W}^{1,\widetilde{q}'}(\Omega)}\,, \quad \text{provided that } q \geq \tfrac{2d}{d+1}, \tag{4.13}$$

$$\begin{aligned} \left|\widetilde{b}(\boldsymbol{u},\boldsymbol{u},\boldsymbol{v})\right| &\leq \|\boldsymbol{u}\|^2_{\mathrm{L}^{2q'}(\Omega)} \|\boldsymbol{v}\|_{\mathrm{W}^{1,q}(\Omega)} + \|\boldsymbol{u}\|_{\mathrm{L}^{2q'}(\Omega)} \|\boldsymbol{v}\|_{\mathrm{L}^{2q'}(\Omega)} \|\boldsymbol{u}\|_{\mathrm{W}^{1,q}(\Omega)} \\ &\leq \|\boldsymbol{u}\|^2_{\mathrm{X}(\Omega)} \|\boldsymbol{v}\|_{\mathrm{X}(\Omega)}\,, \end{aligned} \tag{4.14}$$

for $\boldsymbol{u}, \boldsymbol{v}$ in the respective Sobolev spaces, see (2.55), (2.58) and (2.59) for the derivation thereof.

4.2. Approximation Levels and Outline

Similarly as in [BGMS12, Sec. 3.1] let us introduce the following levels of approximation:

- $k \in \mathbb{N}$: The implicit constitutive relation is approximated using a family of Carathéodory functions $\{\boldsymbol{\mathcal{S}}^k\}_{k\in\mathbb{N}}$, which satisfy Assumption 3.18. Then the approximation of the stress is continuous and explicit in $\mathbf{D}\boldsymbol{u}$.
- $n \in \mathbb{N}$: The velocity and the pressure $(\boldsymbol{u}, \pi)$ are approximated by a Galerkin approximation based on a pair of finite element spaces $(\mathbb{V}^n, \mathbb{Q}^n)$, see Section 2.2.1. Recall that if the velocity space $\mathbb{V}^n_{\mathrm{div}}$ is discretely divergence-free rather than exactly divergence-free, then the convective term is replaced by the numerical convective term $\widetilde{b}$ defined in (4.9) in order to ensure skew-symmetry.
- $m \in \mathbb{N}$: The regularising term $\frac{1}{m}\, |\boldsymbol{u}|^{2q'-2}\, \boldsymbol{u}$ is added to the equation to gain admissibility of the approximate solutions, if $q < \frac{3d}{d+2}$, and to enable us to use the bound on $\widetilde{b}(\cdot,\cdot,\cdot)$ in (4.14), which do not require a restriction on q.

Note that in case the selection $\boldsymbol{\mathcal{S}}^\star$ given in Lemma 3.17 is already a Carathéodory function, the approximation in $k \in \mathbb{N}$ can be skipped. This is in particular the case for the explicit constitutive relations presented in Section 1.1.2. As discussed in Chapter 3 there are two approaches to the graph approximation: One option is to consider a family of $\{\boldsymbol{\mathcal{S}}^k\}_{k\in\mathbb{N}}$ satisfying Assumption 3.18. They arise for example as sequence of approximations of a measurable selection $\boldsymbol{\mathcal{S}}^\star$ of $\mathcal{A}$, given in Lemma 3.17. If this is all the information available, then the limit in k is taken before and separately from the other limits, using the property (σ3) in Assumption 3.18. Another option is to consider the family of approximate graphs $\{\mathcal{A}^k\}_{k\in\mathbb{N}}$ satisfying a Minty type convergence result. For example the generalised Yosida approximation given in Example 3.28 is a suitable choice. The Minty type convergence lemma, see Lemma 3.31, allows us to take the limit in k simultaneously with the discretisation limit.

Without the regularising term due to the modification of the numerical convective term the additional restriction $q > \frac{2d}{d+1}$ is required when using discretely divergence-free finite element functions. Furthermore, the discrete Lipschitz approximation is employed to show the implicit constitutive relation. On the other hand, including the regularising term and taking the regularising limit last allows us to cover the full range $q > \frac{2d}{d+2}$ also when using discretely divergence-free finite element functions. The additional restriction of q can be avoided by first taking the limit $n \to \infty$, where the modification term vanishes, and then taking $m \to \infty$. Furthermore, since the loss of admissibility appears only in the limit $m \to \infty$ no discrete Lipschitz approximation is required, but a continuous one is sufficient. Note also, that for the admissible range $q \geq \frac{3d}{d+2}$ neither a truncation nor a regularising term are required.

The convergence proof for a sequence of finite element approximations of the steady problem in [DKS13a] is based on the existence proof in [BGMS09]. It uses Chacon's biting lemma and the fundamental lemma of Young measures, and to show strong convergence of the symmetric gradients a generalised strict monotonicity property of $\mathcal{A}$ is assumed. Since no regularisation is used a discrete Lipschitz approximation is introduced in order to show the implicit relation and the restrictions $q > \frac{2d}{d+2}$ for exactly divergence-free finite element spaces and $q > \frac{2d}{d+1}$ for discretely divergence-free finite element spaces arise from the convective term and its numerical modification. The approximation of the graph is based on a mollification of a selection function, as in Example 3.20 and the limits corresponding to the graph approximation and the discretisation are taken simultaneously.

In [KS16] convergence (up to subsequences) of a sequence of adaptive finite element approximation to a weak solution is shown for $q > \frac{2d}{d+1}$ and under more restrictive assumptions on the graph approximation, see [KS16, Sec. 7].

<table>
<tr><td colspan="2" rowspan="2">Options
steady problem</td><td colspan="2">regularisation</td></tr>
<tr><td>with</td><td>without</td></tr>
<tr><td rowspan="2">graph
approximation</td><td>general</td><td>k, n, m
Subsection 4.3.1</td><td>k, n</td></tr>
<tr><td>Yosida</td><td>$(k, n), m$</td><td>(k, n)
Subsection 4.3.2</td></tr>
</table>

Fig. 4.1: Options for the approximation levels for the steady problem

Altogether there are four options to take the limits, two of which we want to present, see Figure 4.1:

In Subsection 4.3.1 we will consider the three approximate levels $(k, n, m) \in \mathbb{N}^3$ and take them successively in the order $k \to \infty$, $n \to \infty$ and finally $m \to \infty$. The convergence result in Theorem 4.2 extends the result in [DKS13a] in the sense that the convergence holds for the full range of $q > \frac{2d}{d+2}$ also for discretely divergence-free finite element functions, $\mathcal{A}$ need not be strictly monotone anymore, and more general graph approximations are allowed.

In Subsection 4.3.2 we deal with the two levels of approximation $(k, n) \in \mathbb{N}$ and we will take the two limits together, reproving the result in [DKS13a] in Theorem 4.10. However, using the generalised Yosida approximation instead of an approximation based on mollification as presented in Example 3.20 and used in [DKS13a] simplifies the proof and allows for slightly more general assumptions in the sense that $\mathcal{A}$ is not required to be (generalised) strictly monotone.

4.3. Convergence Results

Let us now show convergence (up to subsequences) of the sequence of solutions to the respective approximate problems to weak solutions of the steady problem. Existence follows from a fixed point argument and is straightforward since $\mathcal{S}^k$ is a Carathéodory function for each $k \in \mathbb{N}$. Then, for each limit taken, a priori estimates allow us to extract weakly converging subsequences, compactness yields strong convergence and the limiting equations can be identified under suitable conditions on the exponent q. The key step in the proofs is the identification of the implicit relation.

4.3.1. With Regularisation

Here we include all three levels of approximation $k, n, m \in \mathbb{N}$ and take the limit successively. Splitting the limits $k \to \infty$ and $n \to \infty$ allows for general graph approximations satisfying Assumption 3.18. Taking the limit $k \to \infty$ first means that afterwards the approximate solutions satisfy the implicit relation formulated with the original graph $\mathcal{A}$ and in the limits $n \to \infty$, $m \to \infty$ the Minty Lemma 3.16 is used. Due to the presence of the regularising term the approximate solutions are admissible in the first two limits and then (4.14) can be used. Note that this is only required for $q < \frac{3d}{d+2}$ but we present the proof for $q \in \left(\frac{2d}{d+2}, \infty\right)$ to avoid distinguishing cases. Only in the limit $m \to \infty$ we loose admissibility, but then the modification of the convective term has vanished already, so that estimate (4.12) can be applied. Since then we are in the continuous situation, the continuous Lipschitz approximation in Lemma 2.12 and the corresponding divergence correction in Lemma 2.13 are applied to show the conditions for the Minty Lemma.
For $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{V}^n$ let us introduce let us introduce

$$L^{k,n,m}[\boldsymbol{u}; \boldsymbol{v}] := \widetilde{b}(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{v}) + \left\langle \boldsymbol{\mathcal{S}}^k(\cdot, \mathbf{D}\boldsymbol{u}), \mathbf{D}\boldsymbol{v} \right\rangle_\Omega + \frac{1}{m} \left\langle |\boldsymbol{u}|^{2q'-2} \boldsymbol{u}, \boldsymbol{v} \right\rangle_\Omega - \langle \boldsymbol{f}, \boldsymbol{v} \rangle_\Omega \,, \tag{4.15}$$

for $k, n, m \in \mathbb{N}$, where $\boldsymbol{\mathcal{S}}^k$ satisfies Assumption 3.18 and $\widetilde{b}(\cdot,\cdot,\cdot)$ is introduced in (2.50) and recalled in (4.9).

The mixed formulation for the approximate problem can be formulated as follows.

Approximate Problem:

For $k, n, m \in \mathbb{N}$ find $(\boldsymbol{U}^{k,n,m}, \pi^{k,n,m}) \in \mathbb{V}^n \times \mathbb{Q}_0^n$ such that

$$L^{k,n,m}[\boldsymbol{U}^{k,n,m}; \boldsymbol{V}] = \left\langle \operatorname{div} \boldsymbol{V}, \pi^{k,n,m} \right\rangle_\Omega \qquad \text{for all } \boldsymbol{V} \in \mathbb{V}^n, \tag{4.16}$$

$$\left\langle \operatorname{div} \boldsymbol{U}^{k,n,m}, Q \right\rangle_\Omega = 0 \qquad \text{for all } Q \in \mathbb{Q}^n. \tag{4.17}$$

Using the definition of $\mathbb{V}^n_{\operatorname{div}}$ and (4.17) for $\boldsymbol{U}^{k,n,m}$ as well as testing with $\boldsymbol{W} \in \mathbb{V}^n_{\operatorname{div}}$ leads to the decoupled problem.

Decoupled Approximate Problem:

For $k, n, m \in \mathbb{N}$ find $\boldsymbol{U}^{k,n,m} \in \mathbb{V}^n_{\operatorname{div}}$ such that

$$L^{k,n,m}[\boldsymbol{U}^{k,n,m}; \boldsymbol{W}] = 0 \qquad \text{for all } \boldsymbol{W} \in \mathbb{V}^n_{\operatorname{div}}. \tag{4.18}$$

For $k, n, m \in \mathbb{N}$ and given $\boldsymbol{U}^{k,n,m} \in \mathbb{V}^n_{\operatorname{div}}$ find $\pi^{k,n,m} \in \mathbb{Q}_0^n$ such that

$$L^{k,n,m}[\boldsymbol{U}^{k,n,m}; \boldsymbol{V}] = \left\langle \operatorname{div} \boldsymbol{V}, \pi^{k,n,m} \right\rangle_\Omega \qquad \text{for all } \boldsymbol{V} \in \mathbb{V}^n. \tag{4.19}$$

To simplify the notation we shall write

$$\boldsymbol{v}^{k,n,m} \underset{k}{\to}\underset{n}{\to}\underset{m}{\to} \boldsymbol{v} \quad \text{strongly in X,} \quad \text{as } k \to \infty,\ n \to \infty,\ m \to \infty, \tag{4.20}$$

for the fact that the limits k, 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 corresponding notation for weak and weak* convergence.

Theorem 4.2 (Convergence in the Steady Case with Regularisation).
In addition to the assumptions of Definition 4.1, let $\mathcal{S}^k$ satisfy Assumption 3.18 with $M = \Omega$. For the finite element approximation let Assumption 2.18 on the domain and on the family of simplicial partitions be satisfied. Let $\mathbb{V}^n$, $\mathbb{V}^n_{\mathrm{div}}$ and $\mathbb{Q}^n_0$ be as introduced in (2.30), (2.32) and (2.33), respectively, and assume that Assumption 2.21 as well as Assumption 2.23 (i), (ii) are satisfied.
*Then, for all $k, n, m \in \mathbb{N}$ there exists a $\boldsymbol{U}^{k,n,m} \in \mathbb{V}^n_{\mathrm{div}}$ and a $\pi^{k,n,m} \in \mathbb{Q}^n_0$ satisfying (4.18), (4.19). Moreover, if $q \in \left(\frac{2d}{d+2}, \infty\right)$, then there exists a weak solution $(\boldsymbol{u}, \mathbf{S}, \pi)$ of **(PS)** according to Definition 4.1 such that with $\widetilde{q}$ as defined in (4.4) (up to non-relabelled subsequences) one has that*

$$\begin{aligned}
\boldsymbol{U}^{k,n,m} &\underset{k}{\to}\underset{n}{\to}\underset{m}{\to} \boldsymbol{u} && \text{strongly in } \mathrm{L}^p(\Omega)^d, \ \text{for all } p \in [1, q^*),\\
\boldsymbol{U}^{k,n,m} &\underset{k}{\rightharpoonup}\underset{n}{\rightharpoonup}\underset{m}{\rightharpoonup} \boldsymbol{u} && \text{weakly in } \mathrm{W}^{1,q}_0(\Omega)^d,\\
\mathcal{S}^k(\cdot, \mathbf{D}\boldsymbol{U}^{k,n,m}) &\underset{k}{\rightharpoonup}\underset{n}{\rightharpoonup}\underset{m}{\rightharpoonup} \mathbf{S} && \text{weakly in } \mathrm{L}^{q'}(\Omega)^{d\times d},\\
\pi^{k,n,m} &\underset{k}{\rightharpoonup}\underset{n}{\rightharpoonup}\underset{m}{\rightharpoonup} \pi && \text{weakly in } \mathrm{L}^{\widetilde{q}}(\Omega),
\end{aligned}$$

as $k \to \infty$, $n \to \infty$ and $m \to \infty$, when taking the limits successively.

The proof of this theorem relies on the Lemmas 4.3–4.8, stated and proved below, where the most challenging part is the identification of the implicit relation after each limit.

Limit $k \to \infty$

The existence proof follows from a standard fixed point argument. Taking $k \to \infty$ we stay in the finite-dimensional setting and hence all bounds are straightforward and strong convergence of the symmetric gradients is obtained. Thus, the identification of the implicit relation relies on the properties of the sequence $\{\mathcal{S}^k\}_{k\in\mathbb{N}}$ according to Assumption 3.18, cf. [BGMS12].

Lemma 4.3 (Existence of Approximate Solutions and Convergence $k \to \infty$).
For each $\kappa := (k, n, m) \in \mathbb{N}^3$ there exists a pair $(\boldsymbol{U}^\kappa, \pi^\kappa) \in \mathbb{V}^n_{\mathrm{div}} \times \mathbb{Q}^n_0$, which satisfies (4.18), (4.19). Furthermore, there exists a constant $c > 0$ such that we have that

$$\|\boldsymbol{U}^\kappa\|^q_{\mathrm{W}^{1,q}(\Omega)} + \left\|\mathcal{S}^k(\cdot, \mathbf{D}\boldsymbol{U}^\kappa)\right\|^{q'}_{\mathrm{L}^{q'}(\Omega)} + \frac{1}{m}\|\boldsymbol{U}^\kappa\|^{2q'}_{\mathrm{L}^{2q'}(\Omega)} + c(m)\|\pi^\kappa\|_{\mathrm{L}^{\overline{q}}(\Omega)} \le c, \tag{4.21}$$

for all $\kappa = (k, n, m) \in \mathbb{N}^3$, where $\overline{q}$ is defined in (4.5).
Also, for each $\nu := (n, m) \in \mathbb{N}^2$ there exists a $\boldsymbol{U}^\nu \in \mathbb{V}^n_{\mathrm{div}}$, a $\pi^\nu \in \mathbb{Q}^n_0$, an $\mathbf{S}^\nu \in \mathrm{L}^\infty(\Omega)^{d\times d}$ and subsequences such that

$$\boldsymbol{U}^{k,\nu} \to \boldsymbol{U}^\nu \qquad \text{strongly in } \mathrm{C}(\overline{\Omega})^d, \tag{4.22}$$
$$\mathbf{D}\boldsymbol{U}^{k,\nu} \to \mathbf{D}\boldsymbol{U}^\nu \qquad \text{strongly in } \mathrm{L}^\infty(\Omega)^{d\times d}, \tag{4.23}$$
$$\nabla\boldsymbol{U}^{k,\nu} \to \nabla\boldsymbol{U}^\nu \qquad \text{strongly in } \mathrm{L}^\infty(\Omega)^{d\times d}, \tag{4.24}$$
$$\mathcal{S}^k(\cdot, \mathbf{D}\boldsymbol{U}^{k,\nu}) \rightharpoonup \mathbf{S}^\nu \qquad \text{weakly in } \mathrm{L}^{q'}(\Omega)^{d\times d}, \tag{4.25}$$
$$\pi^{k,\nu} \to \pi^\nu \qquad \text{strongly in } \mathrm{L}^\infty(\Omega), \tag{4.26}$$

as $k \to \infty$.

Proof.

Step 1: A priori estimates
Testing (4.18) with $\boldsymbol{W} = \boldsymbol{U}^\kappa \in \mathbb{V}^n_{\mathrm{div}}$ for $\kappa = (k,n,m) \in \mathbb{N}^3$, we obtain

$$\left\langle \mathcal{S}^k(\cdot, \mathbf{D}\boldsymbol{U}^\kappa), \mathbf{D}\boldsymbol{U}^\kappa \right\rangle_\Omega + \frac{1}{m}\left\|\boldsymbol{U}^\kappa\right\|^{2q'}_{\mathrm{L}^{2q'}(\Omega)} = \langle \boldsymbol{f}, \boldsymbol{U}^\kappa\rangle_\Omega\,, \tag{4.27}$$

since the term involving $\widetilde{b}$ vanishes due to the skew-symmetry of $\widetilde{b}(\boldsymbol{U}^\kappa,\cdot,\cdot)$, see (2.50). By the properties of $\mathcal{S}^k$ according to Assumption 3.18 ($\sigma2$) and Korn's and Poincaré's inequalities we obtain

$$\left\langle \mathcal{S}^k(\cdot, \mathbf{D}\boldsymbol{U}^\kappa), \mathbf{D}\boldsymbol{U}^\kappa \right\rangle_\Omega \geq -\left\|\widetilde{g}\right\|_{\mathrm{L}^1(\Omega)} + \widetilde{c}_*\left(c\left\|\boldsymbol{U}^\kappa\right\|^q_{\mathrm{W}^{1,q}(\Omega)} + \left\|\mathcal{S}^k(\cdot,\mathbf{D}\boldsymbol{U}^\kappa)\right\|^{q'}_{\mathrm{L}^{q'}(\Omega)}\right). \tag{4.28}$$

On the right-hand side of (4.27) we use duality and Young's inequality with $\varepsilon > 0$ to obtain

$$\langle \boldsymbol{f}, \boldsymbol{U}^\kappa\rangle_\Omega \leq \left\|\boldsymbol{f}\right\|_{\mathrm{W}^{-1,q'}(\Omega)}\left\|\boldsymbol{U}^\kappa\right\|_{\mathrm{W}^{1,q}(\Omega)} \leq c(\varepsilon)\left\|\boldsymbol{f}\right\|^{q'}_{\mathrm{W}^{-1,q'}(\Omega)} + \varepsilon\left\|\boldsymbol{U}^\kappa\right\|^q_{\mathrm{W}^{1,q}(\Omega)}\,. \tag{4.29}$$

Inserting both (4.28) and (4.29) into (4.27) and choosing $\varepsilon > 0$ small enough yields

$$\left\|\boldsymbol{U}^\kappa\right\|^q_{\mathrm{W}^{1,q}(\Omega)} + \left\|\mathcal{S}^k(\cdot,\mathbf{D}\boldsymbol{U}^\kappa)\right\|^{q'}_{\mathrm{L}^{q'}(\Omega)} + \frac{1}{m}\left\|\boldsymbol{U}^\kappa\right\|^{2q'}_{\mathrm{L}^{2q'}(\Omega)} \leq c, \tag{4.30}$$

for any $\kappa = (k,n,m) \in \mathbb{N}^3$.

Let $\{\boldsymbol{W}_j\}_{j\in\{1,\dots,d_n\}}$ with $d_n \in \mathbb{N}$ be a basis of $\mathbb{V}^n_{\mathrm{div}}$, which is L^2-orthonormal (e.g., by the Gram–Schmidt method). We can write $\boldsymbol{U}^\kappa = \sum_{j=1}^{d_n}\alpha^\kappa_j\boldsymbol{W}_j \in \mathbb{V}^n_{\mathrm{div}}$ for $\boldsymbol{\alpha}^\kappa \in \mathbb{R}^{d_n}$. By the L^2-orthonormality we have that

$$\begin{aligned}\left\|\boldsymbol{U}^\kappa\right\|^2_{\mathrm{L}^2(\Omega)} = \left\|\sum_{j=1}^{d_n}\alpha^\kappa_j\boldsymbol{W}_j\right\|^2_{\mathrm{L}^2(\Omega)} &= \sum_{j=1}^{d_n}(\alpha^\kappa_j)^2\left\|\boldsymbol{W}_j\right\|^2_{\mathrm{L}^2(\Omega)} \\ &\geq \min_{j\in\{1,\dots,d_n\}}\left\|\boldsymbol{W}_j\right\|^2_{\mathrm{L}^2(\Omega)}\sum_{j=1}^{d_n}(\alpha^\kappa_j)^2 = c(n)\sum_{j=1}^{d_n}(\alpha^\kappa_j)^2.\end{aligned} \tag{4.31}$$

This implies with the equivalence of norms on finite-dimensional spaces that

$$|\boldsymbol{\alpha}^\kappa|^2 = \sum_{j=1}^{d_n}(\alpha^\kappa_j)^2 \overset{(4.31)}{\leq} c(n)\left\|\boldsymbol{U}^\kappa\right\|^2_{\mathrm{L}^2(\Omega)} \leq c(n)\left\|\boldsymbol{U}^\kappa\right\|^2_{\mathrm{W}^{1,q}(\Omega)} \overset{(4.30)}{\leq} c(n), \tag{4.32}$$

for all $\kappa = (k,n,m) \in \mathbb{N}^3$, where in the last step we have used the a priori estimate (4.30). Hence $\{\boldsymbol{\alpha}^{k,\nu}\}_{k\in\mathbb{N}}$ is bounded in $\mathbb{R}^{d_n}$ for each $\nu = (n,m) \in \mathbb{N}^2$.

In order to estimate $L^\kappa[\cdot;\cdot]$ introduced in (4.15) with $\kappa = (k,n,m) \in \mathbb{N}^3$ let $\boldsymbol{u},\boldsymbol{v} \in \mathrm{X}(\Omega)$ be arbitrary. By the estimate (4.14) on $\widetilde{b}(\cdot,\cdot,\cdot)$, by Hölder's inequality and duality of norms one has

$$\begin{aligned}|L^\kappa[\boldsymbol{u};\boldsymbol{v}]| \leq{}& \left(\left\|\boldsymbol{u}\right\|^2_{\mathrm{L}^{2q'}(\Omega)} + \left\|\mathcal{S}^k(\cdot,\mathbf{D}\boldsymbol{u})\right\|_{\mathrm{L}^{q'}(\Omega)} + \left\|\boldsymbol{f}\right\|_{\mathrm{W}^{-1,q'}(\Omega)}\right)\left\|\boldsymbol{v}\right\|_{\mathrm{W}^{1,q}(\Omega)} \\ &+ \left(\left\|\boldsymbol{u}\right\|_{\mathrm{L}^{2q'}(\Omega)}\left\|\boldsymbol{u}\right\|_{\mathrm{W}^{1,q}(\Omega)} + \frac{1}{m}\left\|\boldsymbol{u}\right\|^{2q'-1}_{\mathrm{L}^{2q'}(\Omega)}\right)\left\|\boldsymbol{v}\right\|_{\mathrm{L}^{2q'}(\Omega)}\,.\end{aligned} \tag{4.33}$$

Considering $\boldsymbol{u} = \boldsymbol{U}^\kappa$ and using the a priori estimate (4.30) in (4.33) yields

$$|L^\kappa[\boldsymbol{U}^\kappa; \boldsymbol{v}]| \leq c(m) \left(\|\boldsymbol{v}\|_{\mathrm{W}^{1,q}(\Omega)} + \|\boldsymbol{v}\|_{\mathrm{L}^{2q'}(\Omega)} \right) \leq c(m) \|\boldsymbol{v}\|_{\mathrm{W}^{1,\bar{q}'}(\Omega)}, \tag{4.34}$$

for all $\kappa = (k, n, m) \in \mathbb{N}^3$, where we have used the embedding $\mathrm{W}^{1,\bar{q}'}(\Omega) \hookrightarrow \mathrm{W}^{1,q}(\Omega) \cap \mathrm{L}^{2q'}(\Omega)$, see (4.11). By the discrete inf-sup condition in Lemma A.4 and the equation (4.19) it follows that

$$\|\pi^\kappa\|_{\mathrm{L}^{\bar{q}}(\Omega)} \leq c \|L^\kappa[\boldsymbol{U}^\kappa; \cdot]\|_{\mathrm{W}^{-1,\bar{q}}(\Omega)} \overset{(4.34)}{\leq} c(m), \tag{4.35}$$

for all $\kappa = (k, n, m) \in \mathbb{N}^3$, which together with (4.30) implies (4.21).

Step 2: Existence of $(\boldsymbol{U}^\kappa, \pi^\kappa)$, $\kappa = (k, n, m) \in \mathbb{N}^3$
To establish the existence of the approximate velocities we use a consequence of Brouwer's fixed point theorem, since we are in a finite-dimensional setting. First, for any fixed $\kappa = (k, n, m) \in \mathbb{N}^3$ we are seeking $\alpha^\kappa \in \mathbb{R}^{d_n}$ such that $\boldsymbol{U}^\kappa = \sum_{i=1}^{d_n} \alpha_i^\kappa \boldsymbol{W}_i$ solves (4.18), where $\{\boldsymbol{W}_i\}_{i \in \{1,\dots,d_n\}}$ is a L^2-orthonormal basis of $\mathbb{V}^n_{\mathrm{div}}$. With

$$\boldsymbol{F} \colon \mathbb{R}^{d_n} \to \mathbb{R}^{d_n}, \quad \boldsymbol{F}(\boldsymbol{\alpha}) = \left(L^\kappa \left(\sum_{i=1}^{d_n} \alpha_i \boldsymbol{W}_i; \boldsymbol{W}_j \right) \right)_{j=1,\dots,d_n},$$

this means that we want to solve $\boldsymbol{F}(\boldsymbol{\alpha}^\kappa) = \boldsymbol{0}$. Since the term including $\widetilde{b}$ is quadratic in $\boldsymbol{U}^\kappa$, $\boldsymbol{\mathcal{S}}^k$ is continuous in its last argument by Assumption 3.18 and the regularising term is continuous in $\boldsymbol{U}^\kappa$, and $\boldsymbol{U}^\kappa$ is linear in $\boldsymbol{\alpha}^\kappa$, we have that $\boldsymbol{F}$ is continuous in $\boldsymbol{\alpha}^\kappa$. Furthermore, setting $\boldsymbol{U} = \sum_{i=1}^{d_n} \alpha_i \boldsymbol{W}_i \in \mathbb{V}^n_{\mathrm{div}}$ we find that

$$\boldsymbol{F}(\boldsymbol{\alpha}) \cdot \boldsymbol{\alpha} = \widetilde{b}(\boldsymbol{U}, \boldsymbol{U}, \boldsymbol{U}) + \left\langle \boldsymbol{\mathcal{S}}^k(\cdot, \mathbf{D}\boldsymbol{U}), \mathbf{D}\boldsymbol{U} \right\rangle_\Omega + \frac{1}{m} \|\boldsymbol{U}\|^{2q'}_{\mathrm{L}^{2q'}(\Omega)} - \langle \boldsymbol{f}, \boldsymbol{U} \rangle_\Omega .$$

The skew-symmetry of $\widetilde{b}$ implies that the first term on the right-hand side vanishes and the other terms we can treat similarly to (4.28) and (4.29) to obtain

$$\boldsymbol{F}(\boldsymbol{\alpha}) \cdot \boldsymbol{\alpha} \geq -\widetilde{c} + \frac{1}{2} \|\mathbf{D}\boldsymbol{U}\|^q_{\mathrm{L}^q(\Omega)},$$

where the constant $\widetilde{c} > 0$ can be assumed to be ≥ 1. With the estimate in (4.31), the equivalence of norms on finite-dimensional function spaces and Poincaré's and Korn's inequality we have that

$$|\boldsymbol{\alpha}| \overset{(4.31)}{\leq} c(n) \|\boldsymbol{U}\|_{\mathrm{L}^2(\Omega)} \leq c(n) \|\boldsymbol{U}\|_{\mathrm{W}^{1,q}(\Omega)} \leq c(n) \|\mathbf{D}\boldsymbol{U}\|_{\mathrm{L}^q(\Omega)} .$$

Hence, for all $\boldsymbol{\alpha} \in \mathbb{R}^{d_n}$ such that $|\boldsymbol{\alpha}| = R := \frac{2\widetilde{c}}{c(n)}$ we have that

$$\boldsymbol{F}(\boldsymbol{\alpha}) \cdot \boldsymbol{\alpha} > 0.$$

This means that $\boldsymbol{F}$ is everywhere outward normal, and thus, as a consequence of Brouwer's fixed point theorem, has a zero in $B_R(\boldsymbol{0})$, compare [GD03, § 5.7, (G.7), p. 104]. So the existence of $\boldsymbol{U}^{k,n,m}$ is proven.

For a given $\boldsymbol{U}^{k,n,m}$ solving (4.18), by the estimate (4.34) on $L^{k,n,m}[\boldsymbol{U}^{k,n,m}; \cdot]$ the existence of a unique (discrete) pressure $\pi^{k,n,m}$ solving (4.19) follows from Corollary A.5.

Step 3: Convergence as $k \to \infty$
Let $\nu := (n,m) \in \mathbb{N}^2$ be fixed. By (4.32) $\{\boldsymbol{\alpha}^{k,\nu}\}_{k\in\mathbb{N}}$ is bounded in $\mathbb{R}^{d_n}$. So by the Bolzano–Weierstraß theorem there exists a (not relabelled) subsequence such that

$$\boldsymbol{\alpha}^{k,\nu} \to \boldsymbol{\alpha}^{\nu} \quad \text{in } \mathbb{R}^{d_n}, \quad \text{as } k \to \infty. \tag{4.36}$$

Setting $\boldsymbol{U}^{\nu} := \sum_{j=1}^{d_n} \alpha_j^{\nu} \boldsymbol{W}_j \in \mathbb{V}^n_{\mathrm{div}}$, we have that

$$\begin{aligned} \boldsymbol{U}^{k,\nu} &\to \boldsymbol{U}^{\nu} && \text{strongly in } C(\overline{\Omega})^d, && (4.37)\\ \mathbf{D}\boldsymbol{U}^{k,\nu} &\to \mathbf{D}\boldsymbol{U}^{\nu} && \text{strongly in } \mathrm{L}^{\infty}(\Omega)^{d\times d}, && (4.38)\\ \nabla\boldsymbol{U}^{k,\nu} &\to \nabla\boldsymbol{U}^{\nu} && \text{strongly in } \mathrm{L}^{\infty}(\Omega)^{d\times d}, && (4.39) \end{aligned}$$

as $k \to \infty$.

By the a priori estimate (4.21), uniformly in k, the (sequential) Banach–Alaoglu theorem[1] allows us to extract a weakly converging (non-relabelled) subsequence such that

$$\begin{aligned} \mathcal{S}^k(\cdot, \mathbf{D}\boldsymbol{U}^{k,\nu}) &\rightharpoonup \mathbf{S}^{\nu} && \text{weakly in } \mathrm{L}^{q'}(\Omega)^{d\times d}, && (4.40)\\ \pi^{k,\nu} &\rightharpoonup \pi^{\nu} && \text{weakly in } \mathrm{L}^{\overline{q}}(\Omega), && (4.41) \end{aligned}$$

as $k \to \infty$. For each $n \in \mathbb{N}$ the space $\mathbb{Q}^n_0 \subset \mathrm{L}^{\overline{q}}_0(\Omega)$ is finite-dimensional, so the convergence is strong in $\mathrm{L}^{\overline{q}}(\Omega)$ and by the equivalence of norms on finite-dimensional spaces (4.26) follows. Since $\mathbb{Q}^n_0$ is closed, we have that $\pi^{\nu} \in \mathbb{Q}^n_0$.

□

Note that $\boldsymbol{U}^{k,n,m}$ solving (4.18) is in general not unique. But for any $k, n, m \in \mathbb{N}$ we select one function $\boldsymbol{U}^{k,n,m}$ and the corresponding pressure $\pi^{k,n,m}$.

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

$$L^{\nu}[\boldsymbol{u};\boldsymbol{v}] := \widetilde{b}(\boldsymbol{u},\boldsymbol{u},\boldsymbol{v}) + \langle \mathbf{S}^{\nu}, \mathbf{D}\boldsymbol{v}\rangle_{\Omega} + \frac{1}{m}\left\langle |\boldsymbol{u}|^{2q'-2}\boldsymbol{u}, \boldsymbol{v}\right\rangle_{\Omega} - \langle \boldsymbol{f}, \boldsymbol{v}\rangle_{\Omega}, \tag{4.42}$$

for any $\nu = (n,m) \in \mathbb{N}^2$, where $\mathbf{S}^{\nu}$ is given in Lemma 4.3.

Lemma 4.4 (Identification of the PDE as $k \to \infty$)**.**
Let $\nu := (n,m) \in \mathbb{N}^2$ be arbitrary but fixed. The functions $\boldsymbol{U}^{\nu} \in \mathbb{V}^n_{\mathrm{div}}$, $\pi^{\nu} \in \mathbb{Q}^n_0$ and $\mathbf{S}^{\nu} \in \mathrm{L}^{q'}(\Omega)^{d\times d}$ from Lemma 4.3 satisfy

$$\begin{aligned} L^{\nu}[\boldsymbol{U}^{\nu};\boldsymbol{V}] &= \langle \operatorname{div}\boldsymbol{V}, \pi^{\nu}\rangle_{\Omega} && \text{for all } \boldsymbol{V} \in \mathbb{V}^n, && (4.43)\\ (\mathbf{D}\boldsymbol{U}^{\nu}(\boldsymbol{x}), \mathbf{S}^{\nu}(\boldsymbol{x})) &\in \mathcal{A}(\boldsymbol{x}) && \text{for a.e. } \boldsymbol{x} \in \Omega, && (4.44) \end{aligned}$$

for all $\nu = (n,m) \in \mathbb{N}^2$.

Proof. Let $\nu = (n,m) \in \mathbb{N}^2$ be arbitrary but fixed.

Step 1: Identification of the limiting equations
For $\boldsymbol{V} \in \mathbb{V}^n$ let us consider $L^{k,\nu}[\boldsymbol{U}^{k,\nu};\boldsymbol{V}]$ and $L^{\nu}[\boldsymbol{U}^{\nu};\boldsymbol{V}]$, defined in (4.15) and (4.42), respectively, term by term: The convergence of $\boldsymbol{U}^{k,n,m}$ in $\mathrm{C}(\overline{\Omega})^d$ due to (4.22) implies that $\boldsymbol{U}^{k,\nu} \otimes \boldsymbol{U}^{k,\nu} \to \boldsymbol{U}^{\nu} \otimes \boldsymbol{U}^{\nu}$ strongly in $\mathrm{L}^1(\Omega)^{d\times d}$, as $k \to \infty$. Since $\nabla\boldsymbol{V} \in \mathrm{L}^{\infty}(\Omega)^{d\times d}$ we obtain

$$\left\langle \boldsymbol{U}^{k,\nu} \otimes \boldsymbol{U}^{k,\nu}, \nabla\boldsymbol{V}\right\rangle_{\Omega} \to \langle \boldsymbol{U}^{\nu} \otimes \boldsymbol{U}^{\nu}, \nabla\boldsymbol{V}\rangle_{\Omega}, \quad \text{as } k \to \infty. \tag{4.45}$$

[1] We refer to the sequential version and omit the “sequential” from now on.

By (4.24) we have that $\nabla \boldsymbol{U}^{k,\nu} \to \nabla \boldsymbol{U}^{\nu}$ in $\mathrm{L}^\infty(\Omega)^{d\times d}$ and (4.22) implies that $\boldsymbol{U}^{k,\nu} \otimes \boldsymbol{V} \to \boldsymbol{U}^{\nu} \otimes \boldsymbol{V}$ in $\mathrm{L}^\infty(\Omega)^{d\times d}$, hence we have that

$$\left\langle \boldsymbol{U}^{k,\nu} \otimes \boldsymbol{V}, \nabla \boldsymbol{U}^{k,\nu} \right\rangle_\Omega \to \langle \boldsymbol{U}^{\nu} \otimes \boldsymbol{V}, \nabla \boldsymbol{U}^{\nu} \rangle_\Omega, \quad \text{as } k \to \infty. \tag{4.46}$$

So taking (4.45) and (4.46) together, yields

$$\widetilde{b}(\boldsymbol{U}^{k,\nu}, \boldsymbol{U}^{k,\nu}, \boldsymbol{V}) \to \widetilde{b}(\boldsymbol{U}^{\nu}, \boldsymbol{U}^{\nu}, \boldsymbol{V}), \quad \text{as } k \to \infty. \tag{4.47}$$

Noting that $\mathbf{D}\boldsymbol{V} \in \mathrm{L}^\infty(\Omega)^{d\times d}$, using the weak convergence of $\mathcal{S}^k(\cdot, \mathbf{D}\boldsymbol{U}^{k,\nu})$ in $\mathrm{L}^{q'}(\Omega)^{d\times d}$ due to (4.25) shows that

$$\left\langle \mathcal{S}^k(\cdot, \mathbf{D}\boldsymbol{U}^{k,\nu}), \mathbf{D}\boldsymbol{V} \right\rangle_\Omega \to \langle \mathbf{S}^{\nu}, \mathbf{D}\boldsymbol{V} \rangle_\Omega, \quad \text{as } k \to \infty. \tag{4.48}$$

Furthermore, the strong convergence of $\boldsymbol{U}^{k,\nu}$ in (4.22) implies a.e. convergence of a subsequence of $\boldsymbol{U}^{k,\nu}$ and consequently a.e. convergence of $\frac{1}{m}\left|\boldsymbol{U}^{k,\nu}\right|^{2q'-2}\boldsymbol{U}^{k,\nu}$. Thus, we find that

$$\frac{1}{m}\left\langle \left|\boldsymbol{U}^{k,\nu}\right|^{2q'-2}\boldsymbol{U}^{k,\nu}, \boldsymbol{V} \right\rangle_\Omega \to \frac{1}{m}\left\langle |\boldsymbol{U}^{\nu}|^{2q'-2}\boldsymbol{U}^{\nu}, \boldsymbol{V} \right\rangle_\Omega, \quad \text{as } k \to \infty. \tag{4.49}$$

Taking (4.47)–(4.49) yields that

$$L^{k,\nu}(\boldsymbol{U}^{k,\nu}; \boldsymbol{V}) \to L^{\nu}(\boldsymbol{U}^{\nu}; \boldsymbol{V}), \quad \text{as } k \to \infty, \tag{4.50}$$

for any $\boldsymbol{V} \in \mathbb{V}^n$ and all $\nu = (n, m) \in \mathbb{N}^2$. By the strong convergence of $\pi^{k,\nu}$ in $\mathrm{L}^\infty(\Omega)$ by (4.26) and the fact that $\operatorname{div} \boldsymbol{V} \in \mathrm{L}^\infty(\Omega)$ it also follows that

$$\left\langle \operatorname{div} \boldsymbol{V}, \pi^{k,\nu} \right\rangle_\Omega \to \langle \operatorname{div} \boldsymbol{V}, \pi^{\nu} \rangle_\Omega, \quad \text{as } k \to \infty. \tag{4.51}$$

Now the convergence results (4.50) and (4.51) applied in (4.19) yield (4.43).

Step 2: Identification of the implicit relation

It remains to show the implicit relation in (4.44), which relies on the strong convergence of $\mathbf{D}\boldsymbol{U}^{k,\nu} \to \mathbf{D}\boldsymbol{U}^{\nu}$ and the properties of $\mathcal{S}^k$ stated in Assumption 3.18.

Since the sequence $\{\mathbf{D}\boldsymbol{U}^{k,\nu}\}_{k\in\mathbb{N}}$ is bounded in $\mathrm{L}^\infty(\Omega)^{d\times d}$ by (4.23), it follows by the Assumption 3.18 (σ3) that for all $\varphi \in \mathrm{C}_0^\infty(\Omega)$ such that $\varphi \geq 0$ and for all matrices $\mathbf{B} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$ we have that

$$0 \leq \liminf_{k\to\infty} \left\langle \mathcal{S}^k(\cdot, \mathbf{D}\boldsymbol{U}^{k,\nu}) - \mathcal{S}^\star(\cdot, \mathbf{B}), (\mathbf{D}\boldsymbol{U}^{k,\nu} - \mathbf{B})\varphi \right\rangle_\Omega. \tag{4.52}$$

By the strong convergence of $\mathbf{D}\boldsymbol{U}^{k,\nu}$ in (4.23) and the weak convergence of $\mathcal{S}^k(\cdot, \mathbf{D}\boldsymbol{U}^{k,\nu})$ in (4.25) we have that

$$\left\langle \mathcal{S}^k(\cdot, \mathbf{D}\boldsymbol{U}^{k,\nu}) - \mathcal{S}^\star(\cdot, \mathbf{B}), (\mathbf{D}\boldsymbol{U}^{k,\nu} - \mathbf{B})\varphi \right\rangle_\Omega \to \langle \mathbf{S}^{\nu} - \mathcal{S}^\star(\cdot, \mathbf{B}), (\mathbf{D}\boldsymbol{U}^{\nu} - \mathbf{B})\varphi \rangle_\Omega, \tag{4.53}$$

as $k \to \infty$. With (4.52) this implies that for all $\varphi \in \mathrm{C}_0^\infty(\Omega)$ such that $\varphi \geq 0$ and for all matrices $\mathbf{B} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$ we have that

$$0 \leq \langle \mathbf{S}^{\nu} - \mathcal{S}^\star(\cdot, \mathbf{B}), (\mathbf{D}\boldsymbol{U}^{\nu} - \mathbf{B})\varphi \rangle_\Omega. \tag{4.54}$$

By Lemma 3.17 (S4) this allows us to conclude that

$$(\mathbf{D}\boldsymbol{U}^{\nu}(\boldsymbol{x}), \mathbf{S}^{\nu}(\boldsymbol{x})) \in \mathcal{A}(\boldsymbol{x}) \quad \text{for a.e. } \boldsymbol{x} \in \Omega,$$

which finishes the proof. □

Limit $n \to \infty$

To gain compactness the restriction $q > \frac{2d}{d+2}$ is required. Also we need the assumptions on the finite element setting, phrased in Theorem 4.10. The implicit relation can be identified by means of the Minty type convergence result. Due to the presence of the regularising term the approximate solutions are still admissible and hence energy identities can be used to verify the assumptions of the convergence lemma.

Lemma 4.5 (Convergence $n \to \infty$).
Let $\nu := (n, m) \in \mathbb{N}^2$ and let $\boldsymbol{U}^{\nu} \in \mathbb{V}^n_{\mathrm{div}}$, $\pi^{\nu} \in \mathbb{Q}^n_0$, and $\mathbf{S}^{\nu} \in \mathrm{L}^{\infty}(\Omega)^{d\times d}$ be solutions to (4.43), (4.44). Then, there exists a constant $c > 0$ such that we have that

$$\|\boldsymbol{U}^{\nu}\|^{q}_{\mathrm{W}^{1,q}(\Omega)} + \|\mathbf{S}^{\nu}\|^{q'}_{\mathrm{L}^{q'}(\Omega)} + \frac{1}{m}\|\boldsymbol{U}^{\nu}\|^{2q'}_{\mathrm{L}^{2q'}(\Omega)} + c(m)\,\|\pi^{\nu}\|_{\mathrm{L}^{\overline{q}}(\Omega)} \le c, \tag{4.55}$$

for all $\nu = (n, m) \in \mathbb{N}^2$.
Furthermore, with $\overline{q}$ as defined in (4.5), for each $m \in \mathbb{N}$ there exists a $\boldsymbol{u}^m \in \mathrm{X}_{\mathrm{div}}(\Omega)$, a $\pi^m \in \mathrm{L}^{\overline{q}}_0(\Omega)$, an $\mathbf{S}^m \in \mathrm{L}^{q'}(Q)^{d\times d}$ and subsequences such that

$$\begin{aligned}
\boldsymbol{U}^{n,m} &\to \boldsymbol{u}^m && \textit{strongly in } \mathrm{L}^p(\Omega)^d, \\
& && \quad \textit{for all } p \in [1, \max(2q', q^*)), && (4.56)\\
\boldsymbol{U}^{n,m} &\rightharpoonup \boldsymbol{u}^m && \textit{weakly in } \mathrm{W}^{1,q}_0(\Omega)^d \cap \mathrm{L}^{2q'}(\Omega)^d, && (4.57)\\
\mathbf{S}^{n,m} &\rightharpoonup \mathbf{S}^m && \textit{weakly in } \mathrm{L}^{q'}(\Omega)^{d\times d}, && (4.58)\\
|\boldsymbol{U}^{n,m}|^{2q'-2}\boldsymbol{U}^{n,m} &\rightharpoonup |\boldsymbol{u}^m|^{2q'-2}\boldsymbol{u}^m && \textit{weakly in } \mathrm{L}^{(2q')'}(\Omega)^d, && (4.59)\\
\pi^{n,m} &\rightharpoonup \pi^m && \textit{weakly in } \mathrm{L}^{\overline{q}}(\Omega), && (4.60)
\end{aligned}$$

as $n \to \infty$.

Proof.
Step 1: Estimates
Testing (4.43) with $\boldsymbol{V} = \boldsymbol{U}^{\nu} \in \mathbb{V}^n_{\mathrm{div}}$ considering the skew-symmetry of $\widetilde{b}(\cdot,\cdot,\cdot)$ and noting that the pressure term vanishes, one has that

$$\langle \mathbf{S}^{\nu}, \mathbf{D}\boldsymbol{U}^{\nu}\rangle_{\Omega} + \frac{1}{m}\|\boldsymbol{U}^{\nu}\|^{2q'}_{\mathrm{L}^{2q'}(\Omega)} = \langle \boldsymbol{f}, \boldsymbol{U}^{\nu}\rangle_{\Omega} \quad \text{for any } \nu = (n, m) \in \mathbb{N}^2. \tag{4.61}$$

By the same estimates as performed in (4.28), (4.29), where the bounds on $\langle \mathbf{S}^{\nu}, \mathbf{D}\boldsymbol{U}^{\nu}\rangle_{\Omega}$ now follow by Assumption 3.11 (A3) with the fact that $(\mathbf{D}\boldsymbol{U}^{\nu}, \mathbf{S}^{\nu}) \in \mathcal{A}(\cdot)$ a.e. in Ω by (4.44), it follows that

$$\|\boldsymbol{U}^{\nu}\|^{q}_{\mathrm{W}^{1,q}(\Omega)} + \|\mathbf{S}^{\nu}\|^{q'}_{\mathrm{L}^{q'}(\Omega)} + \frac{1}{m}\|\boldsymbol{U}^{\nu}\|^{2q'}_{\mathrm{L}^{2q'}(\Omega)} \le c, \tag{4.62}$$

for any $\nu = (n,m) \in \mathbb{N}^2$. Exactly as in (4.33) and (4.34) we obtain

$$|\mathrm{L}^\nu[\boldsymbol{U}^\nu;\boldsymbol{v}]| \le c(m)\left(\|\boldsymbol{v}\|_{\mathrm{W}^{1,q}(\Omega)} + \|\boldsymbol{v}\|_{\mathrm{L}^{2q'}(\Omega)}\right) \le c(m)\,\|\boldsymbol{v}\|_{\mathrm{W}^{1,\overline{q}'}(\Omega)}\,, \tag{4.63}$$

for all $\nu = (n,m) \in \mathbb{N}^2$. As in (4.35) by the discrete inf-sup condition in Lemma A.4 and the equation (4.43), we obtain

$$\|\pi^\nu\|_{\mathrm{L}^{\overline{q}}(\Omega)} \le c\,\|\mathrm{L}^\nu[\boldsymbol{U}^{n,m};\cdot]\|_{\mathrm{W}^{-1,\overline{q}}(\Omega)} \overset{(4.63)}{\le} c(m), \tag{4.64}$$

for all $\nu = (n,m) \in \mathbb{N}^2$, which together with (4.62) implies (4.55).

Step 2: Convergence as $n \to \infty$
By the uniform estimates from Step 1, the convergence results in (4.57)–(4.60) follow using the Banach–Alaoglu theorem. For (4.56) compactness of the embedding $\mathrm{W}^{1,q}(\Omega) \hookrightarrow\hookrightarrow \mathrm{L}^p(\Omega)$, for all $p \in [1, q^*)$, and interpolation between $\mathrm{L}^p(\Omega)^d$ and $\mathrm{L}^{2q'}(\Omega)^d$ is used.

The proof that $\boldsymbol{u}^m$ is divergence-free follows as in [DKS13a, p. 1001]: Let $h \in \mathrm{L}^{q'}(\Omega)$ and note that by the Assumption 2.21 it follows that $\Pi^n_{\mathbb{Q}} h \to h$ in particular in $\mathrm{L}^{q'}(\Omega)$, see (2.37). By (4.57) one has that $\operatorname{div}\boldsymbol{U}^{n,m} \rightharpoonup \operatorname{div}\boldsymbol{u}^m$ weakly in $\mathrm{L}^q(\Omega)$, and hence

$$\left\langle \operatorname{div}\boldsymbol{U}^{n,m}, \Pi^n_{\mathbb{Q}} h\right\rangle_\Omega \to \langle \operatorname{div}\boldsymbol{u}^m, h\rangle_\Omega\,, \qquad \text{as } n \to \infty. \tag{4.65}$$

Since $\boldsymbol{U}^{n,m} \in \mathbb{V}^n_{\mathrm{div}}$ the left-hand side vanishes for all $n \in \mathbb{N}$, and hence it follows that $\langle \operatorname{div}\boldsymbol{u}^m, h\rangle_\Omega = 0$ for all $h \in \mathrm{L}^{q'}(\Omega)$, so $\boldsymbol{u}^m$ is (weakly) divergence-free.

For each $n \in \mathbb{N}$ we have that $\pi^{n,m} \in \mathbb{Q}^n_0$ and $1 \in \mathrm{L}^{\overline{q}'}(\Omega)$, since Ω is bounded. With the weak convergence of $\pi^{n,m}$ in $\mathrm{L}^{\overline{q}}(\Omega)$ according to (4.60) we have that $0 = \langle 1, \pi^{n,m}\rangle_\Omega \to \langle 1, \pi^m\rangle_\Omega$ as $n \to \infty$, and hence $\langle 1, \pi^m\rangle_\Omega = 0$, so $\pi^m \in \mathrm{L}^{\overline{q}}_0(\Omega)$. □

For $\boldsymbol{u} \in \mathrm{L}^{2q'}(\Omega)^d$ and $\boldsymbol{v} \in \mathrm{X}(\Omega)$, let us introduce

$$L^m[\boldsymbol{u};\boldsymbol{v}] \coloneqq b(\boldsymbol{u},\boldsymbol{u},\boldsymbol{v}) + \langle \mathbf{S}^m, \mathbf{D}\boldsymbol{v}\rangle_\Omega + \frac{1}{m}\left\langle |\boldsymbol{u}|^{2q'-2}\boldsymbol{u}, \boldsymbol{v}\right\rangle_\Omega - \langle \boldsymbol{f}, \boldsymbol{v}\rangle_\Omega\,, \tag{4.66}$$

for $m \in \mathbb{N}$ and $\mathbf{S}^m$ given by Lemma 4.5 and $b(\cdot,\cdot,\cdot)$ as defined in (4.8).

Lemma 4.6 (Identification of the PDE as $n \to \infty$)**.**
The limiting functions $\boldsymbol{u}^m \in \mathrm{X}_{\mathrm{div}}(\Omega)$, $\pi^m \in \mathrm{L}^{\overline{q}}_0(\Omega)$ and $\mathbf{S}^m \in \mathrm{L}^{q'}(\Omega)^{d\times d}$ from Lemma 4.5 satisfy

$$L^m[\boldsymbol{u}^m;\boldsymbol{v}] = \langle \operatorname{div}\boldsymbol{v}, \pi^m\rangle_\Omega \qquad \text{for all } \boldsymbol{v} \in \mathrm{C}^\infty_0(\Omega)^d, \tag{4.67}$$

$$(\mathbf{D}\boldsymbol{u}^m(\boldsymbol{x}), \mathbf{S}^m(\boldsymbol{x})) \in \mathcal{A}(\boldsymbol{x}) \qquad \text{for a.e. } \boldsymbol{x} \in \Omega, \tag{4.68}$$

for all $m \in \mathbb{N}$.

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

Step 1: Identification of the limiting equation
Let $\boldsymbol{v} \in \mathrm{C}^\infty_0(\Omega)^d$ and recall that by Remark 2.24 (i), for any $s \in [1,\infty)$, we have the convergence

$$\Pi^n\boldsymbol{v} \to \boldsymbol{v} \qquad \text{strongly in } \mathrm{W}^{1,s}_0(\Omega)^d, \qquad \text{as } n \to \infty. \tag{4.69}$$

One can show with the convergence results in (4.56)–(4.59) that

$$L^{n,m}[\boldsymbol{U}^{n,m};\Pi^n\boldsymbol{v}] \to L^m[\boldsymbol{u}^m;\boldsymbol{v}], \qquad \text{as } n\to\infty, \tag{4.70}$$

considering the expressions term by term. We will demonstrate the argument only for the convective term, since this term contains the main difficulty:

Since $q^*>2$, whenever $q>\frac{2d}{d+2}$, by (4.56) there exists a $p>2$ such that $\boldsymbol{U}^{n,m}\to\boldsymbol{u}^m$ in $\mathrm{L}^p(\Omega)^d$ as $n\to\infty$, and hence $\boldsymbol{U}^{n,m}\otimes\boldsymbol{U}^{n,m}\to\boldsymbol{u}^m\otimes\boldsymbol{u}^m$ in $\mathrm{L}^{p/2}(\Omega)^{d\times d}$, as $n\to\infty$. Since $\frac{p}{2}>1$, (4.69) holds for $s=\left(\frac{p}{2}\right)'\in(1,\infty)$, so we have that $\nabla\Pi^n\boldsymbol{v}\to\nabla\boldsymbol{v}$ in $\mathrm{L}^{(p/2)'}(\Omega)^{d\times d}$ and hence

$$\langle\boldsymbol{U}^{n,m}\otimes\boldsymbol{U}^{n,m},\nabla\Pi^n\boldsymbol{v}\rangle_\Omega\to\langle\boldsymbol{u}^m\otimes\boldsymbol{u}^m,\nabla\boldsymbol{v}\rangle_\Omega\,, \qquad \text{as } n\to\infty. \tag{4.71}$$

Furthermore, we have that $\nabla\boldsymbol{U}^{n,m}\rightharpoonup\nabla\boldsymbol{u}^m$ converges weakly in $\mathrm{L}^q(\Omega)^{d\times d}$ by (4.57), and again by (4.56) there exists a $p>q'$ such that $\boldsymbol{U}^{n,m}\to\boldsymbol{u}^m$ in $\mathrm{L}^p(\Omega)^d$. Then, there exists an $s\in(1,\infty)$ such that $\frac{1}{q}+\frac{1}{p}+\frac{1}{s}=1$, and with (4.69) we have that $\Pi^n\boldsymbol{v}\to\boldsymbol{v}$ in $\mathrm{L}^s(\Omega)^d$. Together these yield

$$\langle\boldsymbol{U}^{n,m}\otimes\Pi^n\boldsymbol{v},\nabla\boldsymbol{U}^{n,m}\rangle_\Omega\to\langle\boldsymbol{u}^m\otimes\boldsymbol{v},\nabla\boldsymbol{u}^m\rangle_\Omega\,, \qquad \text{as } n\to\infty. \tag{4.72}$$

Taking (4.71) and (4.72) together shows that

$$\widetilde{b}(\boldsymbol{U}^{n,m},\boldsymbol{U}^{n,m},\Pi^n\boldsymbol{v})\to\widetilde{b}(\boldsymbol{u}^m,\boldsymbol{u}^m,\boldsymbol{v})=b(\boldsymbol{u}^m,\boldsymbol{u}^m,\boldsymbol{v}), \qquad \text{as } n\to\infty, \tag{4.73}$$

and the equality holds, because $\boldsymbol{u}^m$ is divergence-free, see (4.9).

Similarly, the weak convergence of $\pi^{n,m}\rightharpoonup\pi^m$ in $\mathrm{L}^{\overline{q}}(\Omega)$ by (4.60), implies with $\overline{q}>1$ and (4.69) that

$$\langle\operatorname{div}\Pi^n\boldsymbol{v},\pi^{n,m}\rangle_\Omega\to\langle\operatorname{div}\boldsymbol{v},\pi^m\rangle_\Omega\,, \qquad \text{as } n\to\infty. \tag{4.74}$$

Now (4.70) and (4.74) allow us to deduce (4.67) from (4.43). By density we can test with $\boldsymbol{v}\in\mathrm{X}(\Omega)$ in (4.67). In particular $\boldsymbol{u}^m\in\mathrm{X}_{\mathrm{div}}(\Omega)$ is an admissible test function and testing with it yields the energy identity

$$\langle\mathbf{S}^m,\mathbf{D}\boldsymbol{u}^m\rangle_\Omega+\frac{1}{m}\left\|\boldsymbol{u}^m\right\|^{2q'}_{\mathrm{L}^{2q'}(\Omega)}=\langle\boldsymbol{f},\boldsymbol{u}^m\rangle_\Omega\,. \tag{4.75}$$

Here $b(\boldsymbol{u}^m,\cdot,\cdot)$ and the pressure term vanish since $\boldsymbol{u}^m$ is divergence-free.

Step 2: Identification of the implicit relation (cf. [BGMS12])
Since we have weak convergence of $\mathbf{D}\boldsymbol{U}^{n,m}$ in (4.57) and of $\mathbf{S}^{n,m}$ in (4.58) as well as the implicit relation (4.44) for $(\mathbf{D}\boldsymbol{U}^{n,m},\mathbf{S}^{n,m})$, by the Minty Lemma 3.16 it is enough to show that

$$\limsup_{n\to\infty}\,\langle\mathbf{S}^{n,m},\mathbf{D}\boldsymbol{U}^{n,m}\rangle_\Omega\le\langle\mathbf{S}^m,\mathbf{D}\boldsymbol{u}^m\rangle_\Omega\,, \tag{4.76}$$

in order to conclude that (4.68) holds.

To show (4.76) we proceed similarly as in [BGMS12, p. 2784] using the energy identities for $\boldsymbol{U}^{n,m}$ in (4.61) and for $\boldsymbol{u}^m$ in (4.75), which is available since $\boldsymbol{u}^m$ is still admissible. Taking both energy identities into account and using the fact that $\boldsymbol{U}^{n,m}\rightharpoonup\boldsymbol{u}^m$ weakly in $\mathrm{W}^{1,q}(\Omega)^d$

and in $\mathrm{L}^{2q'}(\Omega)^d$ by (4.58) and the lower semicontinuity of the norm we obtain

$$\begin{aligned}\limsup_{n\to\infty}\langle \mathbf{S}^{n,m},\mathbf{D}\boldsymbol{U}^{n,m}\rangle_\Omega &\overset{(4.61)}{=} \limsup_{n\to\infty}\left(\langle \boldsymbol{f},\boldsymbol{U}^{n,m}\rangle_\Omega - \frac{1}{m}\|\boldsymbol{U}^{n,m}\|^{2q'}_{\mathrm{L}^{2q'}(\Omega)}\right)\\ &\le \limsup_{n\to\infty}\langle \boldsymbol{f},\boldsymbol{U}^{n,m}\rangle_\Omega - \liminf_{n\to\infty}\frac{1}{m}\|\boldsymbol{U}^{n,m}\|^{2q'}_{\mathrm{L}^{2q'}(\Omega)}\\ &\le \langle \boldsymbol{f},\boldsymbol{u}^m\rangle_\Omega - \frac{1}{m}\|\boldsymbol{u}^m\|^{2q'}_{\mathrm{L}^{2q'}(\Omega)} \overset{(4.75)}{=} \langle \mathbf{S}^m,\mathbf{D}\boldsymbol{u}^m\rangle_\Omega\,.\end{aligned} \tag{4.77}$$

This proves (4.76) and hence (4.68) is shown. □

Limit $m\to\infty$

Since the limit in n is already taken, the modification of the convective term has vanished. Here the limiting function will not be an admissible test function in the equation anymore. In order to identify the constitutive law an approximating sequence of admissible functions is considered, which is where the Lipschitz truncation method is used. Then we can again use a localised version of the Minty lemma to prove the implicit relation.

Lemma 4.7 (Convergence $m\to\infty$)**.**
Let $\boldsymbol{u}^m\in \mathrm{X}_{\mathrm{div}}(\Omega)$, $\pi^m\in \mathrm{L}^{\widetilde{q}}_0(\Omega)$ and $\mathbf{S}^m\in \mathrm{L}^{q'}(\Omega)^{d\times d}$ be solutions to (4.67), (4.68), for $m\in\mathbb{N}$. Further, let $\widetilde{q}$ be defined by (4.4).
Then, there exists a constant $c>0$ such that we have that

$$\|\boldsymbol{u}^m\|^q_{\mathrm{W}^{1,q}(\Omega)} + \|\mathbf{S}^m\|^{q'}_{\mathrm{L}^{q'}(\Omega)} + \frac{1}{m}\|\boldsymbol{u}^m\|^{2q'}_{\mathrm{L}^{2q'}(\Omega)} + \|\pi^m\|_{\mathrm{L}^{\widetilde{q}}(\Omega)} \le c, \tag{4.78}$$

for all $m\in\mathbb{N}$. Furthermore, there exists a $\boldsymbol{u}\in \mathrm{W}^{1,q}_{0,\mathrm{div}}(\Omega)^d$, a $\pi\in \mathrm{L}^{\widetilde{q}}_0(\Omega)^d$, an $\mathbf{S}\in \mathrm{L}^{q'}(\Omega)^{d\times d}$ and subsequences such that

$$\boldsymbol{u}^m \to \boldsymbol{u} \quad \text{strongly in } \mathrm{L}^p(\Omega)^d, \quad \text{for all } p\in[1,q^*)\,, \tag{4.79}$$

$$\boldsymbol{u}^m \rightharpoonup \boldsymbol{u} \quad \text{weakly in } \mathrm{W}^{1,q}_{0,\mathrm{div}}(\Omega)^d), \tag{4.80}$$

$$\mathbf{S}^m \rightharpoonup \mathbf{S} \quad \text{weakly in } \mathrm{L}^{q'}(\Omega)^{d\times d}, \tag{4.81}$$

$$\pi^m \rightharpoonup \pi \quad \text{weakly in } \mathrm{L}^{\widetilde{q}}_0(\Omega), \tag{4.82}$$

$$\frac{1}{m}|\boldsymbol{u}^m|^{2q'-2}\boldsymbol{u}^m \rightharpoonup \mathbf{0} \quad \text{weakly in } \mathrm{L}^{(2q')'}(\Omega)^d, \tag{4.83}$$

as $m\to\infty$.

Proof.

Step 1: Estimates
The first three terms in (4.78) can be estimated as in (4.62) starting from the energy identity (4.75).

For the estimates on L^m as defined in (4.66) note that we do not have uniform (in m) bounds in the $\mathrm{L}^{2q'}(\Omega)$-norm and that the numerical modification of the convective term is not present anymore. Hence, instead of estimate (4.14) we apply estimate (4.12), which holds provided that $q\ge\frac{2d}{d+2}$. The other terms are estimated as before in (4.33) and (4.63). Using the estimates on the first three terms in (4.78) and the embedding $\mathrm{W}^{1,\widetilde{q}'}(\Omega)\hookrightarrow \mathrm{W}^{1,q}(\Omega)\cap$

$L^{2q'}(\Omega)$ in (4.11) we obtain

$$\begin{aligned} |L^m[\boldsymbol{u}^m;\boldsymbol{v}]| &\overset{(4.12)}{\leq} \|\boldsymbol{u}^m\|^2_{W^{1,q}(\Omega)} \|\boldsymbol{v}\|_{W^{1,\widetilde{q}'}(\Omega)} + \left(\|\mathbf{S}^m\|_{L^{q'}(\Omega)} + \|\boldsymbol{f}\|_{W^{-1,q'}(\Omega)}\right) \|\boldsymbol{v}\|_{W^{1,q}(\Omega)} \\ &\quad + \frac{1}{m} \|\boldsymbol{u}^m\|^{2q'-1}_{L^{2q'}(\Omega)} \|\boldsymbol{v}\|_{L^{2q'}(\Omega)} \\ &\leq c\left(\|\boldsymbol{v}\|_{W^{1,\widetilde{q}'}(\Omega)} + \|\boldsymbol{v}\|_{W^{1,q}(\Omega)} + \|\boldsymbol{v}\|_{L^{2q'}(\Omega)}\right) \leq c\,\|\boldsymbol{v}\|_{W^{1,\widetilde{q}'}(\Omega)}\,, \end{aligned} \tag{4.84}$$

for all $m \in \mathbb{N}$. With the continuous inf-sup condition in Lemma A.4 and the equation (4.67) we have that

$$\|\pi^m\|_{L^{\widetilde{q}}(\Omega)} \leq c\,\|L^m[\boldsymbol{u}^m;\cdot]\|_{W^{-1,\widetilde{q}}(\Omega)} \overset{(4.84)}{\leq} c, \tag{4.85}$$

for all $m \in \mathbb{N}$. This finishes the proof of (4.78).

Step 2: Convergence as $m \to \infty$
The uniform estimate (4.78) allows us to use the Banach–Alaoglu theorem to extract subsequences such that (4.80)–(4.83) hold. By compactness also (4.79) follows and the limit in (4.83) can be identified with $\mathbf{0}$.

Furthermore, since the respective function spaces are closed with respect to the associated weak convergences, we have that $\boldsymbol{u} \in W^{1,q}_{0,\mathrm{div}}(\Omega)^d$ and $\pi \in L^{\widetilde{q}}_0(\Omega)$. □

For $\boldsymbol{u} \in W^{1,q}_0(\Omega)^d$ and $\boldsymbol{v} \in W^{1,\widetilde{q}'}_0(\Omega)^d$, with $\widetilde{q}$ defined in (4.4), let us introduce

$$L[\boldsymbol{u};\boldsymbol{v}] := b(\boldsymbol{u},\boldsymbol{u},\boldsymbol{v}) + \langle \mathbf{S}, \mathbf{D}\boldsymbol{v}\rangle_\Omega - \langle \boldsymbol{f}, \boldsymbol{v}\rangle_\Omega\,, \tag{4.86}$$

where $\mathbf{S}$ is the limiting function introduced in Lemma 4.7.

Lemma 4.8 (Identification of the PDE as $m \to \infty$)**.**
The limiting functions $\boldsymbol{u} \in W^{1,q}_{0,\mathrm{div}}(\Omega)^d$, $\mathbf{S} \in L^{q}(\Omega)^{d\times d}$ *and* $\pi \in L^{\widetilde{q}}_0(\Omega)$ *from Lemma 4.7 satisfy*

$$L[\boldsymbol{u};\boldsymbol{v}] = \langle \operatorname{div} \boldsymbol{v}, \pi\rangle_\Omega \qquad \text{for all } \boldsymbol{v} \in C^\infty_0(\Omega)^d, \tag{4.87}$$

$$(\mathbf{D}\boldsymbol{u}(\boldsymbol{x}), \mathbf{S}(\boldsymbol{x})) \in \mathcal{A}(\boldsymbol{x}) \qquad \text{for a.e. } \boldsymbol{x} \in \Omega, \tag{4.88}$$

i.e., $(\boldsymbol{u}, \mathbf{S}, \pi)$ *is a weak solution according to Definition 4.1.*

Proof.
Step 1: Identification of the limiting equation
With the convergence results in (4.79), (4.81)–(4.83) it can be shown that (4.87) follows from (4.67). The most difficult term is the convective term and thanks to (4.79) and the condition $q > \frac{2d}{d+2}$, which is equivalent to $q^* > 2$ we know that $\boldsymbol{u}^m \otimes \boldsymbol{u}^m \to \boldsymbol{u} \otimes \boldsymbol{u}$ strongly in $L^1(Q)^{d\times d}$, which allows us to take the limit in the convective term.

Note that the test functions in (4.87) can be extended to $\boldsymbol{v} \in W^{1,\widetilde{q}'}_0(\Omega)^d$, but for $q < \frac{3d}{d+2}$ we have that $\widetilde{q}' > q$, and the function $\boldsymbol{u} \in W^{1,q}_{0,\mathrm{div}}(\Omega)^d$ is not an admissible test function.

Step 2: Identification of the implicit relation (compare [BGMS09, BGMS12])
Note that $\mathbf{D}\boldsymbol{u}^m \rightharpoonup \mathbf{D}\boldsymbol{u}$ weakly in $L^q(\Omega)^{d\times d}$ by (4.80) and $\mathbf{S}^m \rightharpoonup \mathbf{S}$ weakly in $L^{q'}(\Omega)^{d\times d}$ by (4.81) and $(\mathbf{D}\boldsymbol{u}^m, \mathbf{S}^m)$ satisfy the implicit relation by (4.68) a.e. in Ω. By Lemma 3.16, one has to show that

$$\limsup_{m\to\infty} \langle \mathbf{S}^m, \mathbf{D}\boldsymbol{u}^m\rangle_{\widetilde{\Omega}} \leq \langle \mathbf{S}, \mathbf{D}\boldsymbol{u}\rangle_{\widetilde{\Omega}}\,, \tag{4.89}$$

for a set $\widetilde{\Omega} \subset \Omega$, to conclude that $(\mathbf{D}\boldsymbol{u}, \mathbf{S})$ satisfies the implicit relation a.e. on $\widetilde{\Omega}$.

Because $\boldsymbol{u}$ is not admissible in the equation (4.87) for $q < \frac{3d}{d+2}$ there is no energy identity available to be applied on the right-hand side of (4.89). Thus, the function $\boldsymbol{u}$ has to be truncated in a suitable manner to identify the implicit relation. As in [BGMS09, DMS08] we use a divergence-corrected version of the Lipschitz approximation, which is summarised in Lemma 2.13. This has the advantage that the pressure term in the equation vanishes.

We aim to truncate the sequence $\boldsymbol{v}^m := \boldsymbol{u}^m - \boldsymbol{u}$, to obtain a sequence which we can use as test functions in the equation (4.67). Recall that $\boldsymbol{v}^m \rightharpoonup \mathbf{0}$ weakly in $\mathrm{W}^{1,q}_{0,\mathrm{div}}(\Omega)^d$ by (4.81) and each $\boldsymbol{v}^m$ is weakly divergence-free. We will apply Lemma 2.12 and Lemma 2.13 to show that

$$\lim_{m\to\infty} \int_\Omega \left[(\mathbf{S}^m - \mathcal{S}^\star(\cdot, \mathbf{D}\boldsymbol{u})) : (\mathbf{D}\boldsymbol{u}^m - \mathbf{D}\boldsymbol{u})\right]^{1/2} \,\mathrm{d}\boldsymbol{x} = 0. \tag{4.90}$$

and the exponent $1/2$ is used to control the size of the "bad" set, which is where $\boldsymbol{v}^m = \boldsymbol{u}^m - \boldsymbol{u}$ and its truncation do not coincide.

The monotonicity of $\mathcal{A}$ and the fact that $(\mathbf{D}\boldsymbol{u}, \mathcal{S}^\star(\cdot, \mathbf{D}\boldsymbol{u})) \in \mathcal{A}(\cdot)$ by the definition of the selection $\mathcal{S}^\star$ and $(\mathbf{D}\boldsymbol{u}^m, \mathbf{S}^m) \in \mathcal{A}(\cdot)$ a.e. in Ω by (4.68) imply that the $\liminf_{m\to\infty}$ is non-negative. To show the other inequality, we denote $H^m := (\mathbf{S}^m - \mathcal{S}^\star(\cdot, \mathbf{D}\boldsymbol{u})) : (\mathbf{D}\boldsymbol{u}^m - \mathbf{D}\boldsymbol{u}) \geq 0$ and let $\mathcal{B}_{m,j} \subset \Omega$ and $\{\boldsymbol{v}^{m,j}\}_{m,j\in\mathbb{N}}$ be given by Lemma 2.12 and let $\{\boldsymbol{w}^{m,j}\}_{m,j\in\mathbb{N}}$ be given by Lemma 2.13.

Splitting the domain and applying Hölder's inequality gives

$$\begin{aligned}\int_\Omega (H^m)^{1/2} \,\mathrm{d}\boldsymbol{x} &\leq |\mathcal{B}_{m,j}|^{\frac12} \left(\int_{\mathcal{B}_{m,j}} H^m \,\mathrm{d}\boldsymbol{x}\right)^{1/2} + |\Omega\backslash\mathcal{B}_{m,j}|^{\frac12} \left(\int_{\Omega\backslash\mathcal{B}_{m,j}} H^m \,\mathrm{d}\boldsymbol{x}\right)^{1/2} \\ &\leq |\mathcal{B}_{m,j}|^{\frac12} \left(\int_\Omega H^m \,\mathrm{d}\boldsymbol{x}\right)^{1/2} + |\Omega|^{\frac12} \left(\int_{\Omega\backslash\mathcal{B}_{m,j}} H^m \,\mathrm{d}\boldsymbol{x}\right)^{1/2},\end{aligned} \tag{4.91}$$

where we have used the non-negativity of H^m in the first term. Because H^m is bounded in $\mathrm{L}^1(\Omega)$ by the estimate (4.78), we obtain

$$\int_\Omega (H^m)^{\frac12} \,\mathrm{d}\boldsymbol{x} \leq c\,|\mathcal{B}_{m,j}|^{\frac12} + c \left(\int_{\Omega\backslash\mathcal{B}_{m,j}} H^m \,\mathrm{d}\boldsymbol{x}\right)^{\frac12} \qquad \text{for all } m, j \in \mathbb{N}. \tag{4.92}$$

By Hölder's inequality, Lemma 2.12 (ii) and since $\lambda_{m,j} \geq 1$, we find for the first term that

$$|\mathcal{B}_{m,j}| = \left\|\mathbb{1}_{\mathcal{B}_{m,j}}\right\|_{\mathrm{L}^1(\Omega)} \leq c \left\|\mathbb{1}_{\mathcal{B}_{m,j}}\right\|_{\mathrm{L}^q(\Omega)} \leq \frac{c}{\lambda_{m,j}} \left\|\lambda_{m,j}\mathbb{1}_{\mathcal{B}_{m,j}}\right\|_{\mathrm{L}^q(\Omega)} \overset{(ii)}{\leq} c2^{-\frac{j}{q}}, \tag{4.93}$$

for all $m \in \mathbb{N}$. In the second term in (4.92) we use Lemma 2.12 (i), which states that on $\Omega\backslash\mathcal{B}_{m,j}$ the approximation $\boldsymbol{v}^{m,j}$ and $\boldsymbol{v}^m$ agree. Splitting the domain of integration and adding and subtracting the term $\langle \mathbf{S}^m - \mathcal{S}^\star(\cdot, \mathbf{D}\boldsymbol{u}), \mathbf{D}\boldsymbol{w}^{m,j}\rangle_\Omega$ yields

$$\begin{aligned}\int_{\Omega\backslash\mathcal{B}_{m,j}} H^m \,\mathrm{d}\boldsymbol{x} &= \langle \mathbf{S}^m - \mathcal{S}^\star(\cdot, \mathbf{D}\boldsymbol{u}), \mathbf{D}\boldsymbol{v}^m\rangle_{\Omega\backslash\mathcal{B}_{m,j}} \overset{(i)}{=} \langle \mathbf{S}^m - \mathcal{S}^\star(\cdot, \mathbf{D}\boldsymbol{u}), \mathbf{D}\boldsymbol{v}^{m,j}\rangle_{\Omega\backslash\mathcal{B}_{m,j}} \\ &= \langle \mathbf{S}^m - \mathcal{S}^\star(\cdot, \mathbf{D}\boldsymbol{u}), \mathbf{D}\boldsymbol{v}^{m,j}\rangle_\Omega - \langle \mathbf{S}^m - \mathcal{S}^\star(\cdot, \mathbf{D}\boldsymbol{u}), \mathbf{D}\boldsymbol{v}^{m,j}\mathbb{1}_{\mathcal{B}_{m,j}}\rangle_\Omega \\ &= \langle \mathbf{S}^m, \mathbf{D}\boldsymbol{w}^{m,j}\rangle_\Omega - \langle \mathcal{S}^\star(\cdot, \mathbf{D}\boldsymbol{u}), \mathbf{D}\boldsymbol{w}^{m,j}\rangle_\Omega \\ &\quad + \langle \mathbf{S}^m - \mathcal{S}^\star(\cdot, \mathbf{D}\boldsymbol{u}), \mathbf{D}\boldsymbol{v}^{m,j} - \mathbf{D}\boldsymbol{w}^{m,j}\rangle_\Omega\end{aligned} \tag{4.94}$$

$$
\begin{aligned}
&- \left\langle \mathbf{S}^m - \mathcal{S}^\star(\cdot, \mathbf{D}\boldsymbol{u}), \mathbf{D}\boldsymbol{v}^{m,j} \mathbb{1}_{\mathcal{B}_{m,j}} \right\rangle_\Omega \\
=:&\ \mathrm{I} - \mathrm{II} + \mathrm{III} - \mathrm{IV}.
\end{aligned}
$$

Since $\mathbf{D}\boldsymbol{u} \in \mathrm{L}^q(\Omega)^{d\times d}$, by Lemma 3.17 (S3) it follows that $\mathcal{S}^\star(\cdot, \mathbf{D}\boldsymbol{u}) \in \mathrm{L}^{q'}(\Omega)^{d\times d}$. The sequence $\mathbf{S}^m$ is bounded in $\mathrm{L}^{q'}(\Omega)^{d\times d}$ uniformly in m by (4.78), and so is $\mathbf{S}^m - \mathcal{S}^\star(\cdot, \mathbf{D}\boldsymbol{u})$. Thus in the terms III and IV, by Hölder's inequality, the uniform bounds, and applying by Lemma 2.13 (ii) in the first term and Lemma 2.12 (iii) and (ii) in the second term, we obtain that

$$
\begin{aligned}
|\mathrm{III}| + |\mathrm{IV}| &\le \left\| \mathbf{S}^m - \mathcal{S}^\star(\cdot, \mathbf{D}\boldsymbol{u}) \right\|_{\mathrm{L}^{q'}(\Omega)} \left(\left\| \mathbf{D}\boldsymbol{v}^{m,j} - \mathbf{D}\boldsymbol{w}^{m,j} \right\|_{\mathrm{L}^q(\Omega)} + \left\| \mathbf{D}\boldsymbol{v}^{m,j} \mathbb{1}_{\mathcal{B}_{m,j}} \right\|_{\mathrm{L}^q(\Omega)} \right) \\
&\le c \left(\left\| \mathbf{D}\boldsymbol{v}^{m,j} - \mathbf{D}\boldsymbol{w}^{m,j} \right\|_{\mathrm{L}^q(\Omega)} + \left\| \mathbf{D}\boldsymbol{v}^{m,j} \right\|_{\mathrm{L}^\infty(\Omega)} \left\| \mathbb{1}_{\mathcal{B}_{m,j}} \right\|_{\mathrm{L}^q(\Omega)} \right) \\
&\le c \left(2^{-\frac{j}{q}} + \lambda_{m,j} \left\| \mathbb{1}_{\mathcal{B}_{m,j}} \right\|_{\mathrm{L}^q(\Omega)} \right) \le c 2^{-\frac{j}{q}},
\end{aligned}
\tag{4.95}
$$

for all $m, j \in \mathbb{N}$.

Since $\mathcal{S}^\star(\cdot, \mathbf{D}\boldsymbol{u}) \in \mathrm{L}^{q'}(\Omega)^{d\times d}$ and since for any fixed $j \in \mathbb{N}$ we have that $\mathbf{D}\boldsymbol{w}^{m,j}$ converges to zero weakly in $\mathrm{L}^q(\Omega)^{d\times d}$, as $m \to \infty$, by Lemma 2.13 (iii), we obtain for the second term in (4.94) that

$$
\lim_{m\to\infty} \mathrm{II} = \lim_{m\to\infty} \left\langle \mathcal{S}^\star(\cdot, \mathbf{D}\boldsymbol{u}), \mathbf{D}\boldsymbol{w}^{m,j} \right\rangle_\Omega = 0 \qquad \text{for any fixed } j \in \mathbb{N}. \tag{4.96}
$$

Since $\boldsymbol{w}^{m,j} \in \mathrm{W}^{1,s}_{0,\mathrm{div}}(\Omega)^d$ for any $s \in (1, \infty)$ by Lemma 2.13 (i), in the first term in (4.94) we can use the equation (4.67) to find that

$$
\begin{aligned}
\mathrm{I} &= \left\langle \mathbf{S}^m, \mathbf{D}\boldsymbol{w}^{m,j} \right\rangle_\Omega \\
&\overset{(4.67)}{=} -b(\boldsymbol{u}^m, \boldsymbol{u}^m, \boldsymbol{w}^{m,j}) - \frac{1}{m} \left\langle \left|\boldsymbol{u}^m\right|^{2q'-2} \boldsymbol{u}^m, \boldsymbol{w}^{m,j} \right\rangle_\Omega + \left\langle \boldsymbol{f}, \boldsymbol{w}^{m,j} \right\rangle_\Omega,
\end{aligned}
\tag{4.97}
$$

where the pressure term vanishes thanks to the solenoidality of $\boldsymbol{w}^{m,j}$. Let $j \in \mathbb{N}$ be fixed and consider the limit $m \to \infty$. By Lemma 2.13 (iii) $\boldsymbol{w}^{m,j}$ converges to zero weakly in $\mathrm{W}^{1,q}_0(\Omega)^d$, so that the last term vanishes since $\boldsymbol{f} \in \mathrm{W}^{-1,q'}(\Omega)^d$. In the second term we use that $\boldsymbol{w}^{m,j}$ converges to zero strongly in $\mathrm{L}^{2q'}(\Omega)^d$ by (iii), and that $\frac{1}{m} \left|\boldsymbol{u}^m\right|^{2q'-1} \boldsymbol{u}^m$ converges to zero weakly in $\mathrm{L}^{(2q')'}(\Omega)^d$ by (4.83); so that the second term vanishes, as $m \to \infty$. Recall, that (4.79) implies that $\boldsymbol{u}^m \otimes \boldsymbol{u}^m \to \boldsymbol{u} \otimes \boldsymbol{u}$ in $\mathrm{L}^p(\Omega)^{d\times d}$ for all $p \in [1, q^*/2)$, as $m \to \infty$. Since $q > \frac{2d}{d+2}$ is equivalent to $q^* > 2$, there exists a $p > 1$ for which the convergence holds. By Lemma 2.13 (iii) we have that $\nabla \boldsymbol{w}^{m,j} \rightharpoonup \mathbf{0}$ weakly in $\mathrm{L}^{p'}(\Omega)^{d\times d}$, as $m \to \infty$ and the first term vanishes, as $m \to \infty$. Altogether we have that

$$
\lim_{m\to\infty} \mathrm{I} = \lim_{m\to\infty} \left\langle \mathbf{S}^m, \mathbf{D}\boldsymbol{w}^{m,j} \right\rangle_\Omega = 0 \qquad \text{for any fixed } j \in \mathbb{N}. \tag{4.98}
$$

Applying the results (4.95)–(4.98) in (4.94), then applying $\limsup_{m\to\infty}$, yields that

$$
\begin{aligned}
\limsup_{m\to\infty} \int_{\Omega\setminus\mathcal{B}_{m,j}} H^m \,\mathrm{d}\boldsymbol{x} &\overset{(4.94)}{\le} \lim_{m\to\infty} |\mathrm{I}| + \lim_{m\to\infty} |\mathrm{II}| + \limsup_{m\to\infty} \left(|\mathrm{III}| + |\mathrm{IV}|\right) \\
&\le 0 + c 2^{-\frac{j}{q}} \qquad \text{for any fixed } j \in \mathbb{N}.
\end{aligned}
\tag{4.99}
$$

Finally using this and (4.93) in (4.92) implies that

$$\limsup_{m\to\infty}\int_\Omega (H^m)^{\frac{1}{2}}\,\mathrm{d}\boldsymbol{x} \overset{(4.92)}{\leq} c\limsup_{m\to\infty}|\mathcal{B}_{m,j}|^{\frac{1}{2}} + c\limsup_{m\to\infty}\left(\int_{\Omega\setminus\mathcal{B}_{m,j}} H^m\,\mathrm{d}\boldsymbol{x}\right)^{\frac{1}{2}} \leq c2^{-\frac{j}{2q}} \quad (4.100)$$

for any $j\in\mathbb{N}$, where we have used the fact that $\limsup_{m\to\infty}(a^m)^{1/2} \leq (\limsup_{m\to\infty} a^m)^{1/2}$, for any sequence $\{a^m\}_{m\in\mathbb{N}}\subset\mathbb{R}_{\geq 0}$. Since $j\in\mathbb{N}$ is arbitrary, we can apply $\limsup_{j\to\infty}$ to both sides of (4.100), and the claim in (4.90) is shown. This means that $(H^m)^{\frac{1}{2}}\to 0$ strongly in $\mathrm{L}^1(\Omega)$, as $m\to\infty$. However, to show (4.89) we need L^1-convergence of H^m at least on suitable subdomains. This can be achieved by use of Chacon's biting lemma, which was applied to prove Lemma 2.17. Recall that $\{H^m\}_{m\in\mathbb{N}}$ is bounded in $\mathrm{L}^1(\Omega)$ by (4.78) and hence Lemma 2.17 shows that there is a non-increasing sequence of measurable subsets $E_i\subset\Omega$, $i\in\mathbb{N}$, such that $|E_i|\to 0$, as $i\to\infty$, and

$$\int_{\Omega\setminus E_i} H^m(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \langle \mathbf{S}^m - \mathcal{S}^\star(\cdot,\mathbf{D}\boldsymbol{u}), \mathbf{D}\boldsymbol{u}^m - \mathbf{D}\boldsymbol{u}\rangle_{\Omega\setminus E_i} \to 0, \qquad \text{as } m\to\infty, \quad (4.101)$$

for any fixed $i\in\mathbb{N}$. With the weak convergence of $\mathbf{S}^m \rightharpoonup \mathbf{S}$ in $\mathrm{L}^{q'}(\Omega)^{d\times d}$ by (4.81) and the weak convergence of $\mathbf{D}\boldsymbol{u}^m \rightharpoonup \mathbf{D}\boldsymbol{u}$ in $\mathrm{L}^q(\Omega)^{d\times d}$ by (4.80) it follows that

$$\lim_{m\to\infty}\langle \mathbf{S}^m, \mathbf{D}\boldsymbol{u}^m\rangle_{\Omega\setminus E_i} = \langle \mathbf{S}, \mathbf{D}\boldsymbol{u}\rangle_{\Omega\setminus E_i} \qquad \text{for all } i\in\mathbb{N}.$$

This shows (4.89) for $\widetilde{\Omega} = \Omega\setminus E_i$, and by the above and the Minty type Lemma 3.16 we find that $(\mathbf{D}\boldsymbol{u}(\boldsymbol{x}), \mathbf{S}(\boldsymbol{x}))\in\mathcal{A}(\boldsymbol{x})$ for a.e. $\boldsymbol{x}\in\Omega\setminus E_i$. Since $|E_i|\to 0$ as $i\to\infty$, we conclude that $(\mathbf{D}\boldsymbol{u}(\boldsymbol{x}), \mathbf{S}(\boldsymbol{x}))\in\mathcal{A}(\boldsymbol{x})$ for a.e. $\boldsymbol{x}\in\Omega$. □

Remark 4.9 (Simplifications and Alternatives)**.**

(i) In case one additionally assumes that the space $\mathbb{V}^n_{\mathrm{div}}$ is exactly divergence-free, i.e., Assumption 2.27 holds, some arguments simplify: In this case, the convective term does not need any modification and the proof that $\boldsymbol{u}^m$ is divergence-free in the proof of Lemma 4.5 is trivial, since the space $\mathrm{W}^{1,q}_{0,\mathrm{div}}(\Omega)^d$ is closed.

(ii) If one works with the generalised Yosida graph approximation $\mathcal{A}^k$ as introduced in Example 3.28, then the limits $k\to\infty$ and $n\to\infty$ can be taken simultaneously without any restriction on the relation between $\frac{1}{k}$ and h_n, see Section 4.3.2.

4.3.2. Without Regularisation (Revisited)

Here we consider the approximation levels $k,n\in\mathbb{N}$. Restricting our attention to graph approximations, which satisfy the Minty type convergence result stated in Lemma 3.31, the limits in k and n can be taken simultaneously. This is the case for the (generalised) Yosida graph approximation in Example 3.28. Since no regularisation is used, the relevant estimates of the convective term and its numerical modification are the ones in (4.12) and (4.13), which is where an additional restriction on q arises from when using discretely divergence-free finite element spaces. For $q\leq\frac{3d}{d+2}$ admissibility is lost and the discrete Lipschitz approximation in Lemma 2.29 and its divergence correction in Lemma 2.30 is applied to show the implicit constitutive relation.

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

$$L^{k,n}[\boldsymbol{u}; \boldsymbol{v}] := \widetilde{b}(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{v}) + \left\langle \mathcal{S}^k(\cdot, \mathbf{D}\boldsymbol{u}), \mathbf{D}\boldsymbol{v} \right\rangle_\Omega - \langle \boldsymbol{f}, \boldsymbol{v} \rangle_\Omega, \tag{4.102}$$

for $k, n \in \mathbb{N}$, where $\mathcal{S}^k$ is defined in Example 3.28 and satisfies Lemma 3.30, and $\widetilde{b}(\cdot,\cdot,\cdot)$ is defined in (4.9). As before we arrive at the following approximate problem, which can be decoupled.

Approximate Problem:

For $k, n \in \mathbb{N}$ find $\boldsymbol{U}^{k,n} \in \mathbb{V}^n_{\mathrm{div}}$ and $\pi^{k,n} \in \mathbb{Q}^n_0$ such that

$$L^{k,n}[\boldsymbol{U}^{k,n}; \boldsymbol{V}] = \left\langle \operatorname{div} \boldsymbol{V}, \pi^{k,n} \right\rangle_\Omega \qquad \text{for all } \boldsymbol{V} \in \mathbb{V}^n. \tag{4.103}$$

Theorem 4.10 (Convergence in the Steady Case without Regularisation)**.**
In addition to the assumptions of Definition 4.1 let $\{\mathcal{S}^k\}_{k \geq k_0}$ be the sequence of Carathéodory functions, defined in (3.59), corresponding to the generalised Yosida graph approximation defined in Example 3.28. For the finite element approximation let Assumption 2.18 on the domain and on the family of simplicial partitions be satisfied. Let $\mathbb{V}^n$, $\mathbb{V}^n_{\mathrm{div}}$ and $\mathbb{Q}^n_0$ be as introduced in (2.30), (2.32) and (2.33), respectively, and assume that Assumption 2.20 as well as Assumption 2.21 and 2.23 (i), (iib) are satisfied.
*Then, for all $k, n \in \mathbb{N}$ such that $k \geq k_0$ there exists a $\boldsymbol{U}^{k,n} \in \mathbb{V}^n_{\mathrm{div}}$ and a $\pi^{k,n} \in \mathbb{Q}^n_0$ satisfying (4.103). Moreover, if $q \in \left(\frac{2d}{d+1}, \infty\right)$ (or if $q \in \left(\frac{2d}{d+2}, \infty\right)$, in case the spaces $\mathbb{V}^n_{\mathrm{div}}$ consist of exactly divergence-free functions, i.e., Assumption 2.27 is satisfied), then there exists a weak solution $(\boldsymbol{u}, \mathbf{S}, \pi)$ of **(PS)** according to Definition 4.1 such that with $\widetilde{q}$ as defined in (4.4) (up to non-relabelled subsequences) one has that*

$$\begin{aligned}
\boldsymbol{U}^{k,n} &\to \boldsymbol{u} && \text{strongly in } \mathrm{L}^p(\Omega)^d, \ \text{for all } p \in [1, q^*), \\
\boldsymbol{U}^{k,n} &\rightharpoonup \boldsymbol{u} && \text{weakly in } \mathrm{W}^{1,q}_0(\Omega)^d, \\
\mathcal{S}^k(\cdot, \mathbf{D}\boldsymbol{U}^{k,n}) &\rightharpoonup \mathbf{S} && \text{weakly in } \mathrm{L}^{q'}(\Omega)^{d \times d}, \\
\pi^{k,n} &\rightharpoonup \pi && \text{weakly in } \mathrm{L}^{\widetilde{q}}(\Omega),
\end{aligned}$$

as $k, n \to \infty$ (combined), without restrictions on the relation between $\frac{1}{k}$ and the discretisation parameter h_n.

The proof of this theorem relies on the Lemmas 4.11 and 4.12, stated and proved below. Note that Assumption 2.20 and Assumption 2.23 (iib) is used only for the discrete Lipschitz approximation, if $q < \frac{3d}{d+2}$, otherwise Assumption 2.23 (ii) suffices. We assume that $k \in \mathbb{N}$ such that $k \geq k_0$ without repeating this restriction.

Lemma 4.11 (Existence of Approximate Solutions and Convergence $k, n \to \infty$)**.**
For each $\kappa := (k, n) \in \mathbb{N}^2$ there exists a pair $(\boldsymbol{U}^\kappa, \pi^\kappa) \in \mathbb{V}^n_{\mathrm{div}} \times \mathbb{Q}^n_0$ which satisfies (4.103). Furthermore, if $q \geq \frac{2d}{d+1}$ (and if $q \geq \frac{2d}{d+2}$ when the spaces $\mathbb{V}^n_{\mathrm{div}}$ are exactly divergence-free), there exists a constant $c > 0$ such that we have that

$$\|\boldsymbol{U}^\kappa\|^q_{\mathrm{W}^{1,q}(\Omega)} + \left\|\mathcal{S}^k(\cdot, \mathbf{D}\boldsymbol{U}^\kappa)\right\|^{q'}_{\mathrm{L}^{q'}(\Omega)} + c\, \|\pi^\kappa\|_{\mathrm{L}^{\widetilde{q}}(\Omega)} \leq c, \tag{4.104}$$

for all $\kappa = (k, n) \in \mathbb{N}^2$, where $\widetilde{q}$ is defined in (4.4). Also, there exists a $\boldsymbol{u} \in \mathrm{W}^{1,q}_{0,\mathrm{div}}(\Omega)^d$, a

$\pi \in \mathrm{L}_0^{\widetilde{q}}(\Omega)$, *an* $\mathbf{S} \in \mathrm{L}^{q'}(\Omega)^{d\times d}$ *and subsequences such that*

$$\boldsymbol{U}^{k,n} \to \boldsymbol{u} \quad \textit{strongly in } \mathrm{L}^p(\Omega)^d, \quad \textit{for all } p \in [1, q^*), \tag{4.105}$$

$$\boldsymbol{U}^{k,n} \rightharpoonup \boldsymbol{u} \quad \textit{weakly in } \mathrm{W}^{1,q}(\Omega)^d, \tag{4.106}$$

$$\mathcal{S}^k(\cdot, \mathbf{D}\boldsymbol{U}^{k,n}) \rightharpoonup \mathbf{S} \quad \textit{weakly in } \mathrm{L}^{q'}(\Omega)^{d\times d}, \tag{4.107}$$

$$\pi^{k,n} \rightharpoonup \pi \quad \textit{weakly in } \mathrm{L}^{\widetilde{q}}(\Omega), \tag{4.108}$$

as $k, n \to \infty$ *(combined).*

Proof.

Step 1: A priori estimates
The first two terms in estimate (4.104) can be estimated as in the proof of Lemma 4.3, Step 1, which yields

$$\|\boldsymbol{U}^\kappa\|_{\mathrm{W}^{1,q}(\Omega)}^q + \left\|\mathcal{S}^k(\cdot, \mathbf{D}\boldsymbol{U}^\kappa)\right\|_{\mathrm{L}^{q'}(\Omega)}^{q'} \le c \quad \text{for all } \kappa = (k, n) \in \mathbb{N}^2. \tag{4.109}$$

In order to estimate L^κ for $\kappa = (k, n) \in \mathbb{N}^2$ we use the estimates (4.12) on the convective term and if the space $\mathbb{V}_{\mathrm{div}}^n$ is only discretely divergence-free also estimate (4.13) on the modification of the numerical convective term since no regularisation is present. Recall that the first estimate holds provided that $q \ge \frac{2d}{d+2}$ and the latter, provided that $q \ge \frac{2d}{d+1}$, which shows that the additional restriction on q is only required in the case of discretely divergence-free finite element functions. With these estimates, applying also duality of norms, Hölder's inequality, estimate (4.109) and the embedding $\mathrm{W}^{1,\widetilde{q}'}(\Omega) \hookrightarrow \mathrm{W}^{1,q}(\Omega)$ (due to $\widetilde{q}' \ge q$) we obtain that

$$\begin{aligned} |L^\kappa[\boldsymbol{U}^\kappa; \boldsymbol{v}]| &\le \|\boldsymbol{U}^\kappa\|_{\mathrm{W}^{1,q}(\Omega)}^2 \|\boldsymbol{v}\|_{\mathrm{W}^{1,\widetilde{q}'}(\Omega)} \\ &\quad + \left(\left\|\mathcal{S}^k(\cdot, \mathbf{D}\boldsymbol{U}^\kappa)\right\|_{\mathrm{L}^{q'}(\Omega)} + \|\boldsymbol{f}\|_{\mathrm{W}^{-1,q'}(\Omega)}\right) \|\boldsymbol{v}\|_{\mathrm{W}^{1,q}(\Omega)} \\ &\overset{(4.109)}{\le} c\left(\|\boldsymbol{v}\|_{\mathrm{W}^{1,\widetilde{q}'}(\Omega)} + \|\boldsymbol{v}\|_{\mathrm{W}^{1,q}(\Omega)}\right) \le c\,\|\boldsymbol{v}\|_{\mathrm{W}^{1,\widetilde{q}'}(\Omega)}, \end{aligned} \tag{4.110}$$

for all $\kappa = (k, n) \in \mathbb{N}^2$.

With the discrete inf-sup condition in Lemma A.4, applying equation (4.103) and the estimate (4.110) it follows that

$$\|\pi^\kappa\|_{\mathrm{L}^{\widetilde{q}}(\Omega)} \le c\,\|L^\kappa[\boldsymbol{U}^\kappa; \cdot]\|_{\mathrm{W}^{-1,\widetilde{q}}(\Omega)} \overset{(4.110)}{\le} c, \tag{4.111}$$

for all $\kappa = (k, n) \in \mathbb{N}^2$. This finishes the proof of (4.104).

Step 2: Existence of $(\boldsymbol{U}^\kappa, \pi^\kappa)$, $\kappa = (k, n) \in \mathbb{N}^2$
The existence of $\boldsymbol{U}^\kappa \in \mathbb{V}_{\mathrm{div}}^n$ solving the pressure-free version of (4.103) is a consequence of Brouwer's fixed point theorem, see Step 2 in the proof of Lemma 4.3. Given $\boldsymbol{U}^\kappa \in \mathbb{V}_{\mathrm{div}}^n$ the existence of a unique $\pi^\kappa \in \mathbb{Q}_0^n$ solving (4.103) is a consequence of the discrete inf-sup condition in Lemma A.4, see Corollary A.5.

Step 3: Convergence as $k \to \infty$
By the uniform estimates in (4.104) we can apply the Banach–Alaoglu theorem to obtain subsequences such that (4.106)–(4.108) hold. Then, by compactness also (4.105) follows.

The proof that $\boldsymbol{u}$ is divergence-free and that π has vanishing mean integral proceeds as in the proof of Lemma 4.5, Step 2. □

For $\boldsymbol{u} \in \mathrm{W}_0^{1,q}(\Omega)^d$ and $\boldsymbol{v} \in \mathrm{W}_0^{1,\widetilde{q}'}(\Omega)^d$, with $\widetilde{q}$ defined in (4.4), let us introduce

$$L[\boldsymbol{u};\boldsymbol{v}] := b(\boldsymbol{u},\boldsymbol{u},\boldsymbol{v}) + \langle \mathbf{S}, \mathbf{D}\boldsymbol{v}\rangle_\Omega - \langle \boldsymbol{f}, \boldsymbol{v}\rangle_\Omega \,, \tag{4.112}$$

where $\mathbf{S}$ is the limiting function introduced in Lemma 4.11 and b is as defined in (4.8).

Lemma 4.12 (Identification of the PDE as $k,n\to\infty$)**.**
Provided that $q > \frac{2d}{d+1}$ (and provided that $q > \frac{2d}{d+2}$ in case $\mathbb{V}^n_{\mathrm{div}}$ is exactly divergence-free) the limiting functions $\boldsymbol{u} \in \mathrm{W}^{1,q}_{0,\mathrm{div}}(\Omega)^d$, $\mathbf{S} \in \mathrm{L}^q(\Omega)^{d\times d}$ and $\pi \in \mathrm{L}_0^{\widetilde{q}}(\Omega)$ from Lemma 4.11 satisfy

$$L[\boldsymbol{u};\boldsymbol{v}] = \langle \operatorname{div}\boldsymbol{v}, \pi\rangle_\Omega \qquad \text{for all } \boldsymbol{v} \in \mathrm{C}_0^\infty(\Omega)^d, \tag{4.113}$$

$$(\mathbf{D}\boldsymbol{u}(\boldsymbol{x}), \mathbf{S}(\boldsymbol{x})) \in \mathcal{A}(\boldsymbol{x}) \qquad \text{for a.e. } \boldsymbol{x} \in \Omega, \tag{4.114}$$

i.e., $(\boldsymbol{u}, \mathbf{S}, \pi)$ is a weak solution according to Definition 4.1.

Proof.
Step 1: Identification of the limiting equation
Let $\boldsymbol{v} \in \mathrm{C}_0^\infty(\Omega)^d$ and recall Remark 2.24 (i) which states that for any $s \in [1,\infty)$ we have that

$$\Pi^n \boldsymbol{v} \to \boldsymbol{v} \qquad \text{strongly in } \mathrm{W}_0^{1,s}(\Omega)^d, \qquad \text{as } n\to\infty. \tag{4.115}$$

One can show that

$$L^{k,n}[\boldsymbol{U}^{k,n}; \Pi^n \boldsymbol{v}] \to L[\boldsymbol{u};\boldsymbol{v}], \qquad \text{as } k,n\to\infty, \tag{4.116}$$

using the convergence results in (4.105)–(4.108) considering $L^{k,n}$ and L, defined in (4.102) and (4.112), respectively, term by term. We will only present the argument for the numerical convective term, since this is where the restrictions on q arise from.

Since by (4.105) we have that $\boldsymbol{U}^{k,n} \to \boldsymbol{u}$ strongly in $\mathrm{L}^p(\Omega)^d$ for any $p \in [1, q^*)$, it follows that $\boldsymbol{U}^{k,n} \otimes \boldsymbol{U}^{k,n} \to \boldsymbol{u}\otimes\boldsymbol{u}$ strongly in $\mathrm{L}^p(\Omega)^{d\times d}$ for any $p \in [1, q^*/2)$, as $k,n\to\infty$. Since $q > \frac{2d}{d+2}$ is equivalent to $q^* > 2$ there exists such $p > 1$ and (4.115) holds in particular for $s = p'$. Then we have that

$$\left\langle \boldsymbol{U}^{k,n} \otimes \boldsymbol{U}^{k,n}, \nabla \Pi^n \boldsymbol{v}\right\rangle_\Omega \to \langle \boldsymbol{u}\otimes\boldsymbol{u}, \nabla\boldsymbol{v}\rangle_\Omega \,, \qquad \text{as } k,n\to\infty, \tag{4.117}$$

provided that $q > \frac{2d}{d+2}$. In the case of exactly divergence-free finite element functions this is the only restriction on q, since the modification of the numerical convective term vanishes, see (2.50).

For the extra term in the numerical convective term note that by (4.106) we have that $\nabla \boldsymbol{U}^{k,n} \rightharpoonup \nabla\boldsymbol{u}$ converges weakly in $\mathrm{L}^q(\Omega)^{d\times d}$, as $k,n\to\infty$. Again by (4.105) the convergence $\boldsymbol{U}^{k,n} \to \boldsymbol{u}$ strongly in $\mathrm{L}^p(\Omega)^d$ holds for all $p \in [1, q^*)$. If $q' < q^*$, which is equivalent to $q > \frac{2d}{d+1}$, then there exists an $s \in (1,\infty)$ such that $\frac{1}{q} + \frac{1}{p} + \frac{1}{s} = 1$ and by (4.115) we have that $\Pi^n\boldsymbol{v} \to \boldsymbol{v}$ strongly in $\mathrm{L}^s(\Omega)^d$, as $n\to\infty$. Thus, it follows that

$$\left\langle \boldsymbol{U}^{k,n} \otimes \Pi^n\boldsymbol{v}, \nabla \boldsymbol{U}^{k,n}\right\rangle_\Omega \to \langle \boldsymbol{u}\otimes\boldsymbol{v}, \nabla\boldsymbol{u}\rangle_\Omega \,, \qquad \text{as } k,n\to\infty, \tag{4.118}$$

provided that $q > \frac{2d}{d+1}$, which represents an additional restriction on q when working with discretely divergence-free finite element functions. Taking (4.116) and (4.117) together, we obtain

$$\left.\begin{aligned} b(\boldsymbol{U}^{k,n},\boldsymbol{U}^{k,n},\Pi^n\boldsymbol{v}) &\to b(\boldsymbol{u},\boldsymbol{u},\boldsymbol{v}) \\ \widetilde{b}(\boldsymbol{U}^{k,n},\boldsymbol{U}^{k,n},\Pi^n\boldsymbol{v}) &\to \widetilde{b}(\boldsymbol{u},\boldsymbol{u},\boldsymbol{v}) = b(\boldsymbol{u},\boldsymbol{u},\boldsymbol{v}) \end{aligned}\right\} \quad \text{as } n\to\infty, \quad \text{if } \begin{cases} q > \frac{2d}{d+2}, \\ q > \frac{2d}{d+1}, \end{cases} \tag{4.119}$$

where the equality in the last line follows from the the solenoidality of $\boldsymbol{u}$.

Also the weak convergence of $\pi^{k,n} \rightharpoonup \pi$ in $L^{\widetilde{q}}(\Omega)$ by (4.108), implies with $\widetilde{q} > 1$ and (4.115) that

$$\left\langle \operatorname{div} \Pi^n \boldsymbol{v}, \pi^{k,n} \right\rangle_\Omega \to \langle \operatorname{div} \boldsymbol{v}, \pi \rangle_\Omega, \qquad \text{as } k,n\to\infty. \tag{4.120}$$

Then, (4.116) and (4.120) applied in (4.103) show that (4.113) holds.

Step 2: Identification of the implicit relation

As before we have that $\mathbf{D}\boldsymbol{U}^{k,n} \rightharpoonup \mathbf{D}\boldsymbol{u}$ weakly in $L^q(\Omega)^{d\times d}$ by (4.106) and $\mathcal{S}^k(\cdot,\mathbf{D}\boldsymbol{U}^{k,n}) \rightharpoonup \mathbf{S}$ weakly in $L^{q'}(\Omega)^{d\times d}$ by (4.107). Furthermore, by the definition of $\mathcal{S}^k$ in (3.59) we have that $(\mathbf{D}\boldsymbol{U}^{k,n}, \mathcal{S}^k(\cdot,\mathbf{D}\boldsymbol{U}^{k,n})) \in \mathcal{A}^k(\cdot)$ a.e. in Ω. Hence, we aim to apply Lemma 3.31 instead of the usual Minty Lemma 3.16 in order to identify the implicit relation.

First let us choose a diagonal sequence such that $n\to\infty \Leftrightarrow k\to\infty$, and set $k = k_n$. Furthermore, we denote

$$\boldsymbol{U}^n := \boldsymbol{U}^{k_n,n}, \quad \mathcal{S}^n := \mathcal{S}^{k_n} \quad \text{and} \quad \mathcal{A}^n := \mathcal{A}^{k_n}.$$

Note that we do not impose any relation between h_n and k_n. Once we have that

$$\limsup_{n\to\infty} \langle \mathcal{S}^n(\cdot,\mathbf{D}\boldsymbol{U}^n), \mathbf{D}\boldsymbol{U}^n\rangle_{\widetilde{\Omega}} \leq \langle \mathbf{S}, \mathbf{D}\boldsymbol{u}\rangle_{\widetilde{\Omega}}, \tag{4.121}$$

for a set $\widetilde{\Omega} \subset \Omega$, then by Lemma 3.31 it follows that $(\mathbf{D}\boldsymbol{u}(\boldsymbol{x}), \mathbf{S}(\boldsymbol{x})) \in \mathcal{A}(\boldsymbol{x})$ for a.e. $\boldsymbol{x} \in \widetilde{\Omega}$. As before for $q < \frac{3d}{d+2}$ the function $\boldsymbol{u}$ is not an admissible test function in (4.113) and hence a truncation is required. Here we need even more: In the equation (4.103) one can test only with functions in $\mathbb{V}^n$ and therefore a discrete Lipschitz approximation shall be used, which was introduced in [DKS13a] and stated in Lemma 2.29. Similarly to the continuous case a divergence-corrected version was developed in [DKS13a], see Lemma 2.30, which allows us to consider the pressure-free version of (4.103).

Since we are working with the graph approximation in Example 3.28 the arguments need some modification compared to the ones in the previous section, see Lemma 4.8, Step 2: First we aim to show that

$$\lim_{n\to\infty} \int_\Omega \left[(\mathcal{S}^n(\cdot,\mathbf{D}\boldsymbol{U}^n) - \mathcal{S}^\star(\cdot,\mathbf{D}\boldsymbol{u})) : (\mathbf{D}\boldsymbol{U}^n - \mathbf{D}\boldsymbol{u} + \mathcal{E}^n(\mathcal{S}^\star(\cdot,\mathbf{D}\boldsymbol{u})))\right]^{1/2} \,\mathrm{d}\boldsymbol{x} = 0, \tag{4.122}$$

and recall that the function $\mathcal{E}^n \colon \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$ is defined by

$$\mathcal{E}^n(\mathbf{S}) = \begin{cases} \mathbf{0} & \text{if } \mathbf{S} = \mathbf{0}, \\ \frac{1}{k_n} |\mathbf{S}|^\gamma \mathbf{S} & \text{if } \mathbf{S} \neq \mathbf{0}. \end{cases} \tag{4.123}$$

with $\gamma = \frac{1}{q-1} - 1$, compare Example 3.28. Since $\mathbf{D}\boldsymbol{u} \in L^q(\Omega)^{d\times d}$, by Lemma 3.17 (S3) it follows that $\mathcal{S}^\star(\cdot,\mathbf{D}\boldsymbol{u}) \in L^{q'}(\Omega)^{d\times d}$ and then we have that

$$\|\mathcal{E}^n(\mathcal{S}^\star(\cdot,\mathbf{D}\boldsymbol{u}))\|^q_{L^q(\Omega)} = \frac{1}{k_n}\int_\Omega |\mathcal{S}^\star(\cdot,\mathbf{D}\boldsymbol{u})|^{q(\gamma+1)} \,\mathrm{d}\boldsymbol{x} = \frac{1}{k_n} \|\mathcal{S}^\star(\cdot,\mathbf{D}\boldsymbol{u})\|^{q'}_{L^{q'}(\Omega)} \to 0, \tag{4.124}$$

as $n \to \infty$.

Note that $(\mathbf{D}\boldsymbol{u}, \boldsymbol{\mathcal{S}}^\star(\cdot, \mathbf{D}\boldsymbol{u})) \in \mathcal{A}(\cdot)$ a.e. in Ω implies by the definition of $\mathcal{A}^n := \mathcal{A}^{k_n}$ in (3.57) that $(\mathbf{D}\boldsymbol{u} - \boldsymbol{\mathcal{E}}^n(\boldsymbol{\mathcal{S}}^\star(\cdot, \mathbf{D}\boldsymbol{u})), \boldsymbol{\mathcal{S}}^\star(\cdot, \mathbf{D}\boldsymbol{u})) \in \mathcal{A}^n(\cdot)$ a.e. in Ω. Since also $(\mathbf{D}\boldsymbol{U}^n, \boldsymbol{\mathcal{S}}^n(\cdot, \mathbf{D}\boldsymbol{U}^n)) \in \mathcal{A}^n(\cdot)$ a.e. in Ω, the monotonicity of $\mathcal{A}^n$ according to Lemma 3.30 shows that

$$H^n(\cdot) := (\boldsymbol{\mathcal{S}}^n(\cdot, \mathbf{D}\boldsymbol{U}^n) - \boldsymbol{\mathcal{S}}^\star(\cdot, \mathbf{D}\boldsymbol{u})) : (\mathbf{D}\boldsymbol{U}^n - \mathbf{D}\boldsymbol{u} + \boldsymbol{\mathcal{E}}^n(\boldsymbol{\mathcal{S}}^\star(\cdot, \mathbf{D}\boldsymbol{u}))) \geq 0, \tag{4.125}$$

a.e. in Ω. Thus, we have that

$$\liminf_{n\to\infty} \int_\Omega (H^n)^{1/2} \, \mathrm{d}\boldsymbol{x} \geq 0, \tag{4.126}$$

which shows one inequality of (4.122). Note also that by the estimates in (4.104) and (4.124) we have that $\{H^n\}_{n\in\mathbb{N}}$ is bounded in $\mathrm{L}^1(\Omega)$.

For the other inequality we want to use the discrete Lipschitz truncation in Lemma 2.29 and the divergence corrected version in Lemma 2.30: We aim to truncate the sequence

$$\boldsymbol{V}^n := \Pi^n(\boldsymbol{U}^n - \boldsymbol{u}) = \boldsymbol{U}^n - \Pi^n(\boldsymbol{u}), \tag{4.127}$$

where the equality follows since Π^n is assumed to be a projection by Assumption 2.23. Note that $\Pi^n(\boldsymbol{u}) \in \mathbb{V}^n_{\mathrm{div}}$, because $\boldsymbol{u}$ is divergence-free and Π^n preserves the divergence in the dual of $\mathbb{Q}^n$ by Assumption 2.23 (i), and hence $\boldsymbol{V}^n \in \mathbb{V}^n_{\mathrm{div}}$ for any $n \in \mathbb{N}$. Furthermore, by Remark 2.24 (i) we have that $\Pi^n\boldsymbol{u} \to \boldsymbol{u}$ strongly in $\mathrm{W}^{1,q}(\Omega)^d$, as $n \to \infty$. Consequently, with the convergence of $\boldsymbol{U}^n \rightharpoonup \boldsymbol{u}$ weakly in $\mathrm{W}^{1,q}(\Omega)^d$, as $n \to \infty$ by (4.106), it follows that $\boldsymbol{V}^n \rightharpoonup \boldsymbol{0}$ weakly in $\mathrm{W}^{1,q}(\Omega)^d$, as $n \to \infty$. Now let $\mathcal{B}_{n,j} \subset \Omega$ and $\{\boldsymbol{V}^{n,j}\}_{n,j\in\mathbb{N}}$ be given by Lemma 2.29 and let $\{\boldsymbol{W}^{n,j}\}_{n,j\in\mathbb{N}}$ be given by Lemma 2.30.

As in (4.91), (4.92) the non-negativity of H^n and the uniform boundedness of H^n in $\mathrm{L}^1(\Omega)$ one can show that

$$\int_\Omega (H^n)^{\frac{1}{2}} \, \mathrm{d}\boldsymbol{x} \leq c\,|\mathcal{B}_{n,j}|^{\frac{1}{2}} + c\left(\int_{\Omega\setminus\mathcal{B}_{n,j}} H^n \, \mathrm{d}\boldsymbol{x}\right)^{\frac{1}{2}} \qquad \text{for all } n, j \in \mathbb{N}. \tag{4.128}$$

As before, by Hölder's inequality, Lemma 2.29 (ii) and since $\lambda_{n,j} \geq 1$, we find for the first term that

$$|\mathcal{B}_{n,j}| \leq c \left\|\mathbb{1}_{\mathcal{B}_{n,j}}\right\|_{\mathrm{L}^q(\Omega)} \leq \frac{c}{\lambda_{n,j}} \left\|\lambda_{n,j}\mathbb{1}_{\mathcal{B}_{n,j}}\right\|_{\mathrm{L}^q(\Omega)} \leq c2^{-\frac{j}{q}}, \tag{4.129}$$

for all $n \in \mathbb{N}$. For the second term in (4.128) we first split

$$\begin{aligned} \mathbf{D}\boldsymbol{U}^n - \mathbf{D}\boldsymbol{u} + \boldsymbol{\mathcal{E}}^n(\boldsymbol{\mathcal{S}}^\star(\cdot, \mathbf{D}\boldsymbol{u})) &= \mathbf{D}\boldsymbol{U}^n - \mathbf{D}\Pi^n\boldsymbol{u} + \mathbf{D}\Pi^n\boldsymbol{u} - \mathbf{D}\boldsymbol{u} + \boldsymbol{\mathcal{E}}^n(\boldsymbol{\mathcal{S}}^\star(\cdot, \mathbf{D}\boldsymbol{u})) \\ &\overset{(4.127)}{=} \mathbf{D}\boldsymbol{V}^n + \mathbf{D}(\Pi^n\boldsymbol{u} - \boldsymbol{u}) + \boldsymbol{\mathcal{E}}^n(\boldsymbol{\mathcal{S}}^\star(\cdot, \mathbf{D}\boldsymbol{u})), \end{aligned} \tag{4.130}$$

where we have used the definition of $\boldsymbol{V}^n$. Compared to Section 4.3.1 we have the extra terms $\mathbf{D}(\Pi^n\boldsymbol{u} - \boldsymbol{u})$ arising from the discretisation and $\boldsymbol{\mathcal{E}}^n(\boldsymbol{\mathcal{S}}^\star(\cdot, \mathbf{D}\boldsymbol{u}))$ stemming from the graph approximation in Example 3.28. Due to Remark 2.24 (i) and (4.124) both terms converge to zero strongly in $\mathrm{L}^q(\Omega)^{d\times d}$, as $n \to \infty$, and hence it follows that

$$\lim_{n\to\infty} \int_{\Omega\setminus\mathcal{B}_{n,j}} (\boldsymbol{\mathcal{S}}^n(\cdot, \mathbf{D}\boldsymbol{U}^n) - \boldsymbol{\mathcal{S}}^\star(\cdot, \mathbf{D}\boldsymbol{u})) : (\mathbf{D}(\Pi^n\boldsymbol{u} - \boldsymbol{u}) + \boldsymbol{\mathcal{E}}^n(\boldsymbol{\mathcal{S}}^\star(\cdot, \mathbf{D}\boldsymbol{u}))) \, \mathrm{d}\boldsymbol{x} = 0, \tag{4.131}$$

since $\{\boldsymbol{\mathcal{S}}^n(\cdot, \mathbf{D}\boldsymbol{U}^n)\}_{k\in\mathbb{N}}$ is bounded in $\mathrm{L}^{q'}(\Omega)^{d\times d}$ by (4.104) and $\boldsymbol{\mathcal{S}}^\star(\cdot, \mathbf{D}\boldsymbol{u}) \in \mathrm{L}^{q'}(\Omega)^{d\times d}$.

The remaining part of the second term in (4.128) can be dealt with as in the proof of Lemma 4.8 Step 2, since the properties of the discrete truncation in Lemmas 2.29 and 2.30 correspond to the properties of the continuous truncation: As in (4.94) one can find that

$$\begin{aligned}\int_{\Omega\setminus\mathcal{B}_{n,j}} (\mathcal{S}^n(\cdot,\mathbf{D}\boldsymbol{U}^n) - \mathcal{S}^\star(\cdot,\mathbf{D}\boldsymbol{u})) : \mathbf{D}\boldsymbol{V}^n \,\mathrm{d}\boldsymbol{x} &= \left\langle \mathcal{S}^n(\cdot,\mathbf{D}\boldsymbol{U}^n), \mathbf{D}\boldsymbol{W}^{n,j}\right\rangle_\Omega \\ &- \left\langle \mathcal{S}^\star(\cdot,\mathbf{D}\boldsymbol{u}), \mathbf{D}\boldsymbol{W}^{n,j}\right\rangle_\Omega + \left\langle \mathcal{S}^n(\cdot,\mathbf{D}\boldsymbol{U}^n) - \mathcal{S}^\star(\cdot,\mathbf{D}\boldsymbol{u}), \mathbf{D}\boldsymbol{V}^{n,j} - \mathbf{D}\boldsymbol{W}^{n,j}\right\rangle_\Omega \\ &- \left\langle \mathcal{S}^n(\cdot,\mathbf{D}\boldsymbol{U}^n) - \mathcal{S}^\star(\cdot,\mathbf{D}\boldsymbol{u}), \mathbf{D}\boldsymbol{V}^{n,j}\mathbb{1}_{\mathcal{B}_{n,j}}\right\rangle_\Omega =: \mathrm{I} - \mathrm{II} + \mathrm{III} - \mathrm{IV}.\end{aligned} \tag{4.132}$$

The terms II, III, IV are treated as in (4.95) and (4.96), respectively:

$$|\mathrm{III}| + |\mathrm{IV}| \le c2^{-\frac{j}{q}} \qquad \text{for any } n, j \in \mathbb{N}, \tag{4.133}$$

$$\lim_{n\to\infty} \mathrm{II} = 0 \qquad \text{for any fixed } j \in \mathbb{N}, \tag{4.134}$$

using Lemmas 2.29 and 2.30. As before, since $\boldsymbol{W}^{n,j} \in \mathbb{V}^n_{\mathrm{div}}$, in term I we can use the pressure-free version of the approximate equation (4.103), which yields that

$$\mathrm{I} = \left\langle \mathcal{S}^n(\cdot,\mathbf{D}\boldsymbol{U}^n), \mathbf{D}\boldsymbol{W}^{n,j}\right\rangle_\Omega \overset{(4.113)}{=} -\widetilde{b}(\boldsymbol{U}^n, \boldsymbol{U}^n, \boldsymbol{W}^{n,j}) + \left\langle \boldsymbol{f}, \boldsymbol{W}^{n,j}\right\rangle_\Omega. \tag{4.135}$$

Let $j \in \mathbb{N}$ be fixed and consider the limit $n \to \infty$. By Lemma 2.30 (iii) $\boldsymbol{W}^{n,j}$ converges to zero weakly in $\mathrm{W}^{1,q}_0(\Omega)^d$, so that the last term vanishes as $\boldsymbol{f} \in \mathrm{W}^{-1,q'}(\Omega)^d$. For the convective term recall that $\boldsymbol{U}^n \to \boldsymbol{u}$ strongly in $\mathrm{L}^p(\Omega)^d$ for any $p \in [1, q^*)$ by (4.105), and hence $\boldsymbol{U}^n \otimes \boldsymbol{U}^n \to \boldsymbol{u} \otimes \boldsymbol{u}$ strongly in $\mathrm{L}^p(\Omega)^{d\times d}$ for any $p \in [1, q^*/2)$. Since $q^* > 2$ if and only if $q > \frac{2d}{d+2}$, there exists such a $p > 1$. Lemma 2.30 (iii) shows that we have weak convergence $\nabla \boldsymbol{W}^{n,j} \rightharpoonup \mathbf{0}$ weakly in $\mathrm{L}^{p'}(\Omega)^{d\times d}$, thus we obtain that

$$\left\langle \boldsymbol{U}^n \otimes \boldsymbol{U}^n, \nabla \boldsymbol{W}^{n,j}\right\rangle_\Omega \to 0, \qquad \text{as } n \to \infty, \qquad \text{for any fixed } j \in \mathbb{N}.$$

In case $\mathbb{V}^n_{\mathrm{div}}$ is only discretely divergence-free and thus the modification of the convective term is present, also the condition $q > \frac{2d}{d+1}$ is needed: The strong convergence $\boldsymbol{U}^n \to \boldsymbol{u}$ strongly in $\mathrm{L}^p(\Omega)^d$ for all $p \in [1, q^*)$ by (4.105), and the weak convergence $\nabla \boldsymbol{U}^n \rightharpoonup \nabla \boldsymbol{u}$ in $\mathrm{L}^q(\Omega)^{d\times d}$ imply, that there is an $s \in (1,\infty)$ such that $\frac{1}{p} + \frac{1}{q} + \frac{1}{s} = 1$, provided that $q' < q^*$, which is equivalent to $q > \frac{2d}{d+1}$. Then, by Lemma 2.30 (iii) we have that $\boldsymbol{W}^{n,j} \to \mathbf{0}$ strongly in $\mathrm{L}^s(\Omega)^d$ and hence

$$\left\langle \boldsymbol{U}^n \otimes \boldsymbol{W}^{n,j}, \nabla \boldsymbol{U}^n\right\rangle_\Omega \to 0, \qquad \text{as } n \to \infty, \qquad \text{for any fixed } j \in \mathbb{N},$$

under the additional condition $q > \frac{2d}{d+1}$. Altogether it follows that

$$\lim_{n\to\infty} \mathrm{I} = \lim_{n\to\infty} \left\langle \mathcal{S}^n(\cdot,\mathbf{D}\boldsymbol{U}^n), \mathbf{D}\boldsymbol{W}^{n,j}\right\rangle_\Omega = 0 \qquad \text{for any fixed } j \in \mathbb{N}. \tag{4.136}$$

Applying the splitting in (4.130), then both (4.132) and (4.131) and finally the results on the terms I–IV in (4.136), (4.133) and (4.134) yield that

$$\begin{aligned}&\limsup_{n\to\infty} \int_{\Omega\setminus\mathcal{B}_{n,j}} H^n \,\mathrm{d}\boldsymbol{x} \\ &\overset{(4.130)}{\le} \limsup_{n\to\infty} \int_{\Omega\setminus\mathcal{B}_{n,j}} (\mathcal{S}^n(\cdot,\mathbf{D}\boldsymbol{U}^n) - \mathcal{S}^\star(\cdot,\mathbf{D}\boldsymbol{u})) : \mathbf{D}\boldsymbol{V}^n \,\mathrm{d}\boldsymbol{x}\end{aligned}$$

$$
\begin{aligned}
&+ \lim_{n\to\infty} \int_{\Omega\setminus\mathcal{B}_{n,j}} (\mathcal{S}^n(\cdot,\mathbf{D}\boldsymbol{U}^n) - \mathcal{S}^\star(\cdot,\mathbf{D}\boldsymbol{u})) : (\mathbf{D}(\Pi^n\boldsymbol{u} - \boldsymbol{u}))\,\mathrm{d}\boldsymbol{x} \\
&+ \lim_{n\to\infty} \int_{\Omega\setminus\mathcal{B}_{n,j}} (\mathcal{S}^n(\cdot,\mathbf{D}\boldsymbol{U}^n) - \mathcal{S}^\star(\cdot,\mathbf{D}\boldsymbol{u})) : (\mathcal{E}^n(\mathcal{S}^\star(\cdot,\mathbf{D}\boldsymbol{u})))\,\mathrm{d}\boldsymbol{x} \qquad (4.137)\\
&\overset{(4.132),(4.131)}{\leq} \lim_{n\to\infty}(\mathrm{I} - \mathrm{II}) + \limsup_{n\to\infty}(|\mathrm{III}| + |\mathrm{IV}|) + 0 \\
&\overset{(4.133)-(4.136)}{\leq} c2^{-\frac{j}{q}} \qquad \text{for any fixed } j \in \mathbb{N}.
\end{aligned}
$$

Now we are in the same situation as in (4.99) and the rest of the proof of (4.122) follows as in (4.100) and the arguments thereafter. Again with Lemma 2.17 it follows from (4.122) and the boundedness of $\{H^n\}_{n\in\mathbb{N}}$ in $\mathrm{L}^1(\Omega)$ that there exists a non-increasing sequence of measurable subsets $E_i \subset \Omega$, $i \in \mathbb{N}$, such that $|E_i| \to 0$, as $i \to \infty$ and

$$
\begin{aligned}
\int_{\Omega\setminus E_i} H^n(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} &= \Big\langle \mathcal{S}^k(\cdot,\mathbf{D}\boldsymbol{U}^n) - \mathcal{S}^\star(\cdot,\mathbf{D}\boldsymbol{u}), \mathbf{D}\boldsymbol{U}^n - \mathbf{D}\boldsymbol{u} + \mathcal{E}^n(\mathcal{S}^\star(\cdot,\mathbf{D}\boldsymbol{u}))\Big\rangle_{\Omega\setminus E_i} \qquad (4.138)\\
&\to 0 \quad \text{as } n \to \infty,
\end{aligned}
$$

for any fixed $i \in \mathbb{N}$. Since $\mathcal{S}^n(\cdot,\mathbf{D}\boldsymbol{U}^n) \rightharpoonup \mathbf{S}$ weakly in $\mathrm{L}^{q'}(\Omega)^{d\times d}$ by (4.107), $\mathbf{D}\boldsymbol{U}^n \rightharpoonup \mathbf{D}\boldsymbol{u}$ weakly in $\mathrm{L}^q(\Omega)^{d\times d}$ by (4.106) and $\mathcal{E}^n(\mathcal{S}^\star(\cdot,\mathbf{D}\boldsymbol{u})) \to \mathbf{0}$ strongly in $\mathrm{L}^q(\Omega)^{d\times d}$ by (4.124), as $n \to \infty$ for any fixed $i \in \mathbb{N}$, it follows that

$$
\langle \mathcal{S}^n(\cdot,\mathbf{D}\boldsymbol{U}^n), \mathbf{D}\boldsymbol{U}^n\rangle_{\Omega\setminus E_i} \to \langle \mathbf{S}, \mathbf{D}\boldsymbol{u}\rangle_{\Omega\setminus E_i}, \qquad \text{as } n \to \infty, \qquad (4.139)
$$

for any fixed $i \in \mathbb{N}$. This shows (4.121) for $\widetilde{\Omega} = \Omega\setminus E_i$, and by the above we find that $(\mathbf{D}\boldsymbol{u}(\boldsymbol{x}), \mathbf{S}(\boldsymbol{x})) \in \mathcal{A}(\boldsymbol{x})$ for a.e. $\boldsymbol{x} \in \Omega\setminus E_i$. Since $|E_i| \to 0$ as $i \to \infty$, we conclude that $(\mathbf{D}\boldsymbol{u}(\boldsymbol{x}), \mathbf{S}(\boldsymbol{x})) \in \mathcal{A}(\boldsymbol{x})$ for a.e. $\boldsymbol{x} \in \Omega$. □

In order to allow for more general graph approximations satisfying Assumption 3.18 the limits $k \to \infty$ and $n \to \infty$ can be taken successively, compare Section 4.3.1.

d	$\frac{2d}{d+2}$	$\frac{2d}{d+1}$	$\frac{3d}{d+2}$
2	1	$\frac{4}{3} \approx 1.33$	$\frac{3}{2} = 1.5$
3	$\frac{6}{5} = 1.2$	$\frac{3}{2} = 1.5$	$\frac{9}{5} = 1.8$

Fig. 4.2: Values of exponents relevant for the steady problem

4.4. Discussion

In Section 4.3.1 we have seen that the use of a regularisation allows us to show the convergence of the sequence of approximate solutions for the full range of $q > \frac{2d}{d+2}$, also when using discretely divergence-free finite element functions, see Theorem 4.2. Furthermore, it allows us to avoid employing a discrete Lipschitz approximation and hence the additional assumptions on the finite element setting are not needed. The presence of the regularising term can be seen as disadvantage, however, one can view the term as a numerical stabilisation.

Compared to the situation in [DKS13a] the graph and its approximation assumed for Theorem 4.2 is more general, since we do not require $\mathcal{A}$ to be generalised strictly monotone

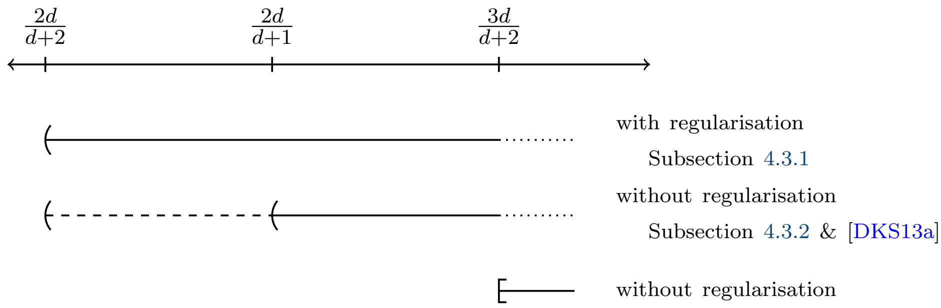

Fig. 4.3: Overview of range of q for which the convergence results in the steady case hold; the dashed line represents the special case, if $\mathbb{V}^n_{\text{div}}$ satisfies Assumption 2.27.

and any graph approximation satisfying Assumption 3.18 works for the argument as long as the limits in k and n are taken successively.

In Subsection 4.3.2 the generalised Yosida graph approximation allowed us to show the same result as in [DKS13a], see Theorem 4.2, but in a simpler manner, avoiding the use of Young measures and relaxing the assumption on the graph $\mathcal{A}$. Note that the use of the discrete Lipschitz approximation compared to the use of the continuous Lipschitz approximation requires the stronger assumption of local $\mathrm{W}^{1,1}$-stability of the projection Π^n in Assumption 2.23 (iib), compared to global $\mathrm{W}^{1,p}$-stability in Assumption 2.23 (ii), and in addition the existence of a locally supported basis of $\mathbb{Q}^n$ is assumed, see Assumption 2.20.

Let us give an overview of the options one has for the convergence results, see Figure 4.3:

- If $q \geq \frac{3d}{d+2}$, the solution $\boldsymbol{u}$ is an admissible test function and neither a regularisation nor a truncation is required, hence also the additional assumptions for the discrete Lipschitz approximation are not needed. Note that in this case we could relax the assumptions on $\mathcal{A}$ and instead work with $\mathfrak{A} \subset \mathrm{L}^q(\Omega)^{d\times d} \times \mathrm{L}^{q'}(\Omega)^{d\times d}$, satisfying the corresponding properties globally, see Remark 3.13 (ii). The proof still works since for such $\mathfrak{A}$ a global version of the Minty type convergence lemma in Lemma 3.16 can be proved.
- If $q \in \left(\frac{2d}{d+1}, \frac{3d}{d+2}\right)$, one can either use a regularisation. Or if no regularisation is used then the discrete Lipschitz approximation is applied under the additional assumptions. It does not make a difference, whether $\mathbb{V}^n_{\text{div}}$ is exactly or discretely divergence-free.
- If $q \in \left(\frac{2d}{d+2}, \frac{2d}{d+1}\right]$ and discretely divergence-free finite element spaces are used, then a regularisation term allows us to show the convergence. If exactly divergence-free finite element spaces are used, then the regularisation is not needed. However, the examples for exactly divergence-free finite element spaces in Example 2.26 are less standard for computations. Also they are likely to be computationally more expensive, since the order is rather high compared to some of the more classically used mixed finite element spaces, for which $\mathbb{V}^n_{\text{div}}$ is merely discretely divergence-free.

Independent of q one has the choice between a general graph approximation satisfying Assumption 3.18 in Chapter 3, which allows one to prove convergence when the limits $k \to \infty$, $n \to \infty$ are taken successively. Or the (generalised) Yosida graph approximation can be employed, which allows one to take the limits in k and n together.

For the unsteady case no discrete Lipschitz approximation is available, which is why the regularising approach in Subsection 4.3.1 is the most promising one to cover the non-admissible range of exponents in the unsteady case; this will be the content of Chapter 5, Subsection 5.3.1. On the other hand Section 5.3.2 will cover the unregularised unsteady problem, for which we can prove convergence only for the admissible range of exponents.

CHAPTER 5

Unsteady Case

In the present chapter we aim to show convergence results for the unsteady problem, in the same spirit as the results for the steady problem in Chapter 4. For this purpose we introduce a fully discrete approximate problem, based on a spatial mixed finite element approximation and a backward Euler discretisation with respect to the temporal variable. Then, we prove convergence up to subsequences under suitable conditions on the finite element setting and the range of q. For the unsteady flow of implicitly constituted fluids such a result is unknown to the best of our knowledge, and even contributions focused on explicit constitutive laws assume additional restrictions on q compared to the full range of existence $q > \frac{2d}{d+2}$, see for example [CHP10] for $q > \frac{2(d+1)}{d+2}$ and [Car07] for $q \geq \frac{3d+2}{d+2}$.

The main challenge consists in the lack of a discrete truncation, a steady version of which was developed in [DKS13a] and applied in Subsection 4.3.2. Since in the unsteady case such a truncation is not available we will present a regularisation approach, similar to Subsection 4.3.2, for which a discrete truncation is not needed, but a continuous one suffices. For the case without regularisation we restrict q to the admissible range, for which a truncation is not required.

Even if a discrete truncation would be at hand, in the case of discretely divergence-free finite element functions one would face additional restrictions on q. Hence, the regularisation approach has some justification independent of the availability of a truncation, because it allows to cover the whole range $q \in \left(\frac{2d}{d+2}, \infty\right)$ for discretely divergence-free finite element spaces when taking the regularisation limit last. Another tool, which we will apply again, is the generalised Yosida graph approximation, which was presented and investigated in Subsection 3.4.3. Since it satisfies a Minty type convergence lemma, we can take the graph approximation limit simultaneously with the discretisation limit.

Additional problems arising from time-dependence and the time-stepping are the following: The stress tensor might be explicitly dependent on the time variable, which requires averaging in order to formulate the equation on each time level. To use the properties of the implicit constitutive relation and also for the identification thereof after taking the limit, one has to to relate the stress tensor function and its averaged, piecewise constant in time version.

For the compactness arguments we will use Simon's lemma or the Aubin–Lions–Simon lemma, see Subsection 2.1.2, depending on the situation. The first has the advantage that it does not rely on uniform estimates on the time derivative, which in the discretisation limit would require the use of a L^2-projection to the space $\mathbb{V}^n_{\mathrm{div}}$ and continuity properties thereof, see Subsection 1.3.1.2. Those properties are investigated in Subsection 2.2.3.2, and the proof requires additional assumptions on the finite element setting. However, the use of Simon's lemma requires a certain amount of admissibility, and therefore stronger restrictions on q, than the Aubin–Lions lemma. We will present these differences in more detail in the course of this chapter.

In the regularised case the identification of the implicit relation relies on a parabolic solenoidal Lipschitz approximation, see Subsection 2.1.3.2, the construction of which is more intricate than the one of the (continuous) divergence-corrected Lipschitz approximation in the steady case. However, the application in proving the validity of the implicit constitutive relation is similar in the steady case and in the unsteady case. In the case without regularisation we restrict our attention to the range of q for which an energy identity is available and thus we can apply the Minty type convergence lemma without a discrete truncation.

The present chapter is structured as follows: In Section 5.1 we will recall the problem in the unsteady case, introduce the notion of weak solution we are aiming for and summarise the notation and estimates on the convective term relevant for this chapter. Then, in Section 5.2, we introduce the layers of approximation, we discuss the choices of setting up the approximate problem and give a more detailed outline of the rest of this chapter. Section 5.3 contains the statement of the convergence results and its proofs. In Subsection 5.3.1 we consider the regularised case, which is based on the preprint [ST18] in collaboration with Endre Süli and covers the whole range $q \in \left(\frac{2d}{d+2}, \infty\right)$. Then, in Subsection 5.3.2 we investigate the case without regularisation, for which the convergence result can be shown for $q \in \left[\frac{3d+2}{d+2}, \infty\right)$.

5.1. The Unsteady Problem and Notation

Let us recall the unsteady problem and introduce the notion of weak solution we are considering. We also recall some of the notation as well as estimates on the (numerical) convective term from Chapter 2 applied in this chapter.

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 $\mathbf{S} \colon Q \to \mathbb{R}^{d\times d}_{\mathrm{sym},0}$ satisfying the balance law of linear momentum and the incompressibility condition:

$$\begin{aligned} \partial_t \boldsymbol{u} + \operatorname{div}(\boldsymbol{u} \otimes \boldsymbol{u}) - \operatorname{div} \mathbf{S} &= -\nabla \pi + \boldsymbol{f} && \text{on } Q, \\ \operatorname{div} \boldsymbol{u} &= 0 && \text{on } Q, \end{aligned} \tag{5.1}$$

subject to the initial condition and homogeneous Dirichlet boundary condition:

$$\boldsymbol{u}(0,\cdot) = \boldsymbol{u}_0(\cdot) \qquad \text{in } \Omega, \tag{5.2}$$

$$\boldsymbol{u} = \mathbf{0} \qquad \text{on } (0,T) \times \partial\Omega. \tag{5.3}$$

The *constitutive relation* is phrased via a maximal monotone set $\mathcal{A}(\boldsymbol{z}) \subset \mathbb{R}^{d\times d}_{\mathrm{sym}} \times \mathbb{R}^{d\times d}_{\mathrm{sym}}$, for $\boldsymbol{z} \in Q$, as

$$(\mathbf{D}\boldsymbol{u}(\boldsymbol{z}), \mathbf{S}(\boldsymbol{z})) \in \mathcal{A}(\boldsymbol{z}), \tag{5.4}$$

with $\mathcal{A}(\cdot)$ satisfying Assumption 3.11 with respect to a given $q \in (1,\infty)$. For the rest of this chapter we refer to the problem consisting of (5.1)–(5.4) as **(PU)**.

For $q \in (1,\infty)$ recall the notation

$$\hat{q} := \max\left(\left(\frac{q(d+2)}{2d}\right)', q\right) = \max\left(\frac{q(d+2)}{q(d+2)-2d}, q\right), \tag{5.5}$$

which was introduced in (2.64). Note that $\hat{q} < \infty$ for any $q > \frac{2d}{d+2}$ and that $\hat{q} = q$, if $q \geq \frac{3d+2}{d+2}$.

Let us now specify the notion of weak solution we are interested in, cf. Definition 2.2 in Chapter 2.

Definition 5.1 (Weak Solution to Problem **(PU)**).
*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 3.11 with respect to q for $M = Q$. 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 **(PU)**, if*

$$\begin{aligned}
\boldsymbol{u} &\in \mathrm{L}^q(0,T;\mathrm{W}^{1,q}_{0,\mathrm{div}}(\Omega)^d) \cap \mathrm{C}_w([0,T];\mathrm{L}^2_{\mathrm{div}}(\Omega)^d), \ \ \textit{s.t.}\\
\partial_t \boldsymbol{u} &\in \mathrm{L}^{\hat{q}'}(0,T;(\mathrm{W}^{1,\hat{q}}_{0,\mathrm{div}}(\Omega)^d)'),\\
\mathbf{S} &\in \mathrm{L}^{q'}(Q)^{d\times d},
\end{aligned}$$

and

$$\langle \partial_t \boldsymbol{u}, \boldsymbol{w}\rangle_\Omega - \langle \boldsymbol{u}\otimes\boldsymbol{u}, \mathbf{D}\boldsymbol{w}\rangle_\Omega + \langle \mathbf{S}, \mathbf{D}\boldsymbol{w}\rangle_\Omega = \langle \boldsymbol{f}, \boldsymbol{w}\rangle_\Omega \quad \textit{for all } \boldsymbol{w} \in \mathrm{C}^\infty_{0,\mathrm{div}}(\Omega)^d, \tag{5.6}$$

$$\textit{for a.e. } t \in (0,T),$$

$$(\mathbf{D}\boldsymbol{u}(\boldsymbol{z}), \mathbf{S}(\boldsymbol{z})) \in \mathcal{A}(\boldsymbol{z}) \quad \textit{for a.e. } \boldsymbol{z} \in Q, \tag{5.7}$$

$$\operatorname*{ess\,lim}_{t\to 0_+} \|\boldsymbol{u}(t,\cdot) - \boldsymbol{u}_0(\cdot)\|_{\mathrm{L}^2(\Omega)} = 0. \tag{5.8}$$

We have chosen 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, cf. Subsection 1.2.6.

Recall from Subsection 2.2.2 the notation for the convective term and its numerical modification:

$$b(\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}) := -\langle \boldsymbol{u}\otimes\boldsymbol{v}, \nabla\boldsymbol{w}\rangle_\Omega\,, \tag{5.9}$$

$$\widetilde{b}(\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}) := \tfrac{1}{2}\left(\langle \boldsymbol{u}\otimes\boldsymbol{w}, \nabla\boldsymbol{v}\rangle_\Omega - \langle \boldsymbol{u}\otimes\boldsymbol{v}, \nabla\boldsymbol{w}\rangle_\Omega\right), \tag{5.10}$$

for $\boldsymbol{u},\boldsymbol{v},\boldsymbol{w} \in \mathrm{W}^{1,\infty}_0(\Omega)^d$, see (2.49) and (2.50), respectively. Furthermore, recall the notation of the function spaces and their associated norms:

$$\mathrm{X}(\Omega) := \mathrm{W}^{1,q}_0(\Omega)^d \cap \mathrm{L}^{2q'}(\Omega)^d, \qquad \|\cdot\|_{\mathrm{X}(\Omega)} := \|\cdot\|_{\mathrm{W}^{1,q}(\Omega)} + \|\cdot\|_{\mathrm{L}^{2q'}(\Omega)}\,, \tag{5.11}$$

$$\mathrm{X}(Q) := \mathrm{L}^q(0,T;\mathrm{W}^{1,q}_0(\Omega)^d) \cap \mathrm{L}^{2q'}(Q)^d, \quad \|\cdot\|_{\mathrm{X}(Q)} := \|\cdot\|_{\mathrm{L}^q(0,T;\mathrm{W}^{1,q}(\Omega))} + \|\cdot\|_{\mathrm{L}^{2q'}(Q)}\,, \tag{5.12}$$

and their solenoidal subspaces are denoted by $\mathrm{X}_{\mathrm{div}}(\Omega)$ and $\mathrm{X}_{\mathrm{div}}(Q)$, respectively.

The following exponents are relevant for the range of q in this chapter:

$$\frac{2d}{d+2} < \frac{2(d+1)}{d+2} < q_d < \widetilde{q}_d < \frac{3d+2}{d+2}, \tag{5.13}$$

with $q_d = \frac{d+\sqrt{3d^2+4d}}{d+2}$ and $\widetilde{q}_d = \frac{3d+2+\sqrt{5d^2+4d+4}}{2(d+2)}$, see (2.14) and (2.15), respectively.

Also we will use the following estimates on the convective term and its numerical modification from Subsection 2.2.2, for $\boldsymbol{u}, \boldsymbol{v}$ in the respective function spaces.

For $\nu := q\left(\frac{q(d+2)}{d} - 2\right)$ we have that

$$\|\langle \boldsymbol{u} \otimes \boldsymbol{u}, \nabla \boldsymbol{v}\rangle_\Omega\|_{\mathrm{L}^1((0,T))} \le c\,\|\boldsymbol{u}\|^2_{\mathrm{L}^\nu(0,T;\mathrm{L}^{2q'}(\Omega))}\,\|\nabla \boldsymbol{v}\|_{\mathrm{L}^{(\nu/2)'}(0,T;\mathrm{L}^q(\Omega))}\,, \tag{5.14}$$

provided that $q \ge q_d$, see (2.62), and

$$\|\langle \boldsymbol{u} \otimes \boldsymbol{v}, \nabla \boldsymbol{u}\rangle_\Omega\|_{\mathrm{L}^1((0,T))} \le c\,\|\boldsymbol{u}\|_{\mathrm{L}^\nu(0,T;\mathrm{L}^{2q'}(\Omega))}\,\|\boldsymbol{v}\|_{\mathrm{L}^r(0,T;\mathrm{L}^{2q'}(\Omega))}\,\|\nabla \boldsymbol{u}\|_{\mathrm{L}^q(Q)} \tag{5.15}$$

for $r \in (1,\infty]$, provided that $q \ge \widetilde{q}_d$, see (2.67). Furthermore we have that $r < \infty$, provided that $q > \widetilde{q}_d$.

With $\hat{q}$ as defined in (5.5) by (2.65) we have that

$$\|\langle \boldsymbol{u} \otimes \boldsymbol{u}, \nabla \boldsymbol{v}\rangle_\Omega\|_{\mathrm{L}^1((0,T))} \le \|\boldsymbol{u}\|^2_{\mathrm{L}^{\frac{q(d+2)}{d}}(Q)}\,\|\nabla \boldsymbol{v}\|_{\mathrm{L}^{\hat{q}}(Q)}\,, \tag{5.16}$$

provided that $q \ge \frac{2d}{d+2}$. Also by (2.68) one has that

$$\|\langle \boldsymbol{u} \otimes \boldsymbol{v}, \nabla \boldsymbol{u}\rangle_\Omega\|_{\mathrm{L}^1((0,T))} \le \|\boldsymbol{u}\|_{\mathrm{L}^{\frac{q(d+2)}{d}}(Q)}\,\|\boldsymbol{v}\|_{\mathrm{L}^s(Q)}\,\|\nabla \boldsymbol{u}\|_{\mathrm{L}^q(Q)} \tag{5.17}$$

for $s \in (1,\infty]$, provided that $q \ge \frac{2(d+1)}{d+2}$, and $s < \infty$, if $q > \frac{2(d+1)}{d+2}$. For the regularised case let us also recall that by (2.69) we have that

$$\begin{aligned}\left\|\widetilde{b}(\boldsymbol{u},\boldsymbol{u},\boldsymbol{v})\right\|_{\mathrm{L}^1((0,T))} &\le \|\boldsymbol{u}\|^2_{\mathrm{L}^{2q'}(Q)}\,\|\nabla \boldsymbol{v}\|_{\mathrm{L}^q(Q)} + \|\boldsymbol{u}\|_{\mathrm{L}^{2q'}(Q)}\,\|\boldsymbol{v}\|_{\mathrm{L}^{2q'}(Q)}\,\|\nabla \boldsymbol{u}\|_{\mathrm{L}^q(Q)} \\ &\le c\,\|\boldsymbol{u}\|^2_{\mathrm{X}(Q)}\,\|\boldsymbol{v}\|_{\mathrm{X}(Q)} \qquad \text{for any } q \in (1,\infty).\end{aligned} \tag{5.18}$$

In the following we will repeatedly use the continuous embedding

$$\mathrm{L}^2_{\mathrm{div}}(\Omega)^d \hookrightarrow (\mathrm{L}^2_{\mathrm{div}}(\Omega)^d)', \tag{5.19}$$

which follows by the Riesz representation theorem.

5.2. Approximation Levels and Outline

In addition to the approximation levels in Chapter 4 we introduce a time-stepping and arrive at the following four levels of approximation:

- $k \in \mathbb{N}$: By means of a family of Carathéodory functions $\{\boldsymbol{\mathcal{S}}^k\}_{k\in\mathbb{N}}$ satisfying Assumption 3.18 in order to approximate the implicit constitutive relation, we arrive at a formulation for which the stress tensor is 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 [CHP10, Tem84], see Subsection 2.2.3.1.
- $n \in \mathbb{N}$: The velocity $\boldsymbol{u}$ is approximated by a Galerkin approximation from finite element spaces $\mathbb{V}^n_{\mathrm{div}}$ in the spatial variable, see Section 2.2. If the space $\mathbb{V}^n_{\mathrm{div}}$ consists of functions that are discretely divergence-free (rather than exactly divergence-free), then we replace the convective term b by the numerical convective term $\widetilde{b}$ defined in (5.10) to maintain skew-symmetry.
- $m \in \mathbb{N}$: As before the regularising term $\frac{1}{m}|\boldsymbol{u}|^{2q'-2}\boldsymbol{u}$ is added to the equation to gain ad-

missibility of the approximate solutions, if $q < \frac{3d+2}{d+2}$. This enables us to use the estimate on $\widetilde{b}(\cdot,\cdot,\cdot)$ in (5.18), without imposing a restriction on q.

For the graph approximation we have two options, as mentioned in more detail in Section 4.2 and discussed in Chapter 3: One is to choose a family $\{\boldsymbol{\mathcal{S}}^k\}_{k\in\mathbb{N}}$ satisfying Assumption 3.18, which is the case for all approximations presented in Chapter 3. This allows for a large class of approximations. The assumptions are formulated in a way that the limit $k \to \infty$ can be taken before the discretisation limit. The other option is to choose the special case of the generalised Yosida graph approximation $\{\mathcal{A}^k\}_{k\in\mathbb{N}}$ presented in Example 3.28. The corresponding family $\{\boldsymbol{\mathcal{S}}^k\}_{k\in\mathbb{N}}$ satisfies a Minty type convergence lemma, see Lemma 3.31, and thus the limit $k \to \infty$ can be taken simultaneously with the discretisation limit.

The time-dependence and time discretisation are the origin for the extra difficulty compared to Chapter 4. For the notion of the backward difference quotient d_t with respect to an equidistant time grid replacing the time derivative in the equation, we refer to Subsection 2.2.3.1. We consider the piecewise constant and continuous, piecewise affine interpolant of the sequence of approximate velocity functions at each time level.

Since both the external force $\boldsymbol{f} \in \mathrm{L}^{q'}(0,T;\mathrm{W}^{-1,q'}(\Omega)^d)$ and the approximate stress $\boldsymbol{\mathcal{S}}^k : Q \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$ are time-dependent, and the time-dependence is not assumed to be continuous, we shall consider an integral-averaged version in the approximate problem. For $l \in \mathbb{N}$ let us introduce the averages with respect to the time grid $\{t_i\}_{i\in\{0,\dots,l\}}$ defined by

$$\boldsymbol{f}_i(\boldsymbol{x}) := \fint_{t_{i-1}}^{t_i} \boldsymbol{f}(t,\boldsymbol{x})\,\mathrm{d}t, \qquad \boldsymbol{\mathcal{S}}_i^k(\boldsymbol{x},\mathbf{B}) := \fint_{t_{i-1}}^{t_i} \boldsymbol{\mathcal{S}}^k(t,\boldsymbol{x},\mathbf{B})\,\mathrm{d}t, \tag{5.20}$$

for $i \in \{1,\dots,l\}$, for $\boldsymbol{x} \in \Omega$ and $\mathbf{B} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$. Furthermore let the corresponding piecewise constant interpolants $\overline{\boldsymbol{f}}$ and $\overline{\boldsymbol{\mathcal{S}}}^k$ be defined as in Subsection 2.2.3.1 (2.72). Recall that by (2.78) and (2.79) we have that

$$\|\overline{\boldsymbol{f}}\|_{\mathrm{L}^{q'}(0,T;\mathrm{W}^{-1,q'}(\Omega))} \le \|\boldsymbol{f}\|_{\mathrm{L}^{q'}(0,T;\mathrm{W}^{-1,q'}(\Omega))} \qquad \text{for all } l \in \mathbb{N}, \tag{5.21}$$

$$\overline{\boldsymbol{f}} \to \boldsymbol{f} \qquad \text{strongly in } \mathrm{L}^{q'}(0,T;\mathrm{W}^{-1,q'}(\Omega)^d), \quad \text{as } l \to \infty. \tag{5.22}$$

Note that we can use the pointwise properties of the (approximate) constitutive relation in an almost everywhere sense only when considering a set of non-zero $(d+1)$-dimensional Lebesgue measure, which entails some additional technicalities.

Furthermore, time-dependence plays a crucial role in the compactness arguments and it requires that $q > \frac{2d}{d+2}$. On the one hand one can apply Simon's compactness lemma, see Lemma 2.8, which relies on uniform convergence of time increments. This assumption can be verified for q large enough that the available estimates ensure admissibility in space and L^p-integrability in time for some $p > 1$. On the other hand the Aubin–Lions compactness lemma, see Lemma 2.7 does not require admissibility in space, but uniform estimates for the sequence of time derivatives, and hence it can be applied for a larger range of q. However, in the discretisation limit such uniform estimates rely on stability properties of the L^2-projection P^n_{div} mapping to $\mathbb{V}^n_{\mathrm{div}}$ defined in Subsection 2.2.3.2. Under additional assumptions on the finite element setting those are proved in Subsection 2.2.3.2.

As in the steady case adding a regularisation term and taking the corresponding limit after the discretisation limit allows us to show convergence for the whole range of existence of weak solutions $q \in \left(\frac{2d}{d+2}, \infty\right)$.

Even if a (discrete) truncation were available the presence of the modification of the numerical convective term would require the restriction $q > \frac{2(d+1)}{d+2}$, see Subsection 2.2.2. This corresponds to the restriction $q > \frac{2d}{d+1}$ in the steady case in [DKS13a] and Subsection 4.3.2. Note that the restriction in the unsteady case is slightly stronger. As before, taking the regularisation limit last circumvents this additional restriction, since at that stage the modification of the convective term is not present anymore.

For q so small that for compactness only the Aubin–Lions lemma is available in the discretisation limit one would have to accept the additional restrictions on the finite element setting mentioned above. By regularising this can be avoided since in the discretisation limit the approximate solutions are still admissible and Simon's compactness lemma can be applied.

Finally, at present there is no discrete parabolic solenoidal Lipschitz approximation available, which means that without regularisation we can identify the implicit relation only for the restricted range of q, for which an energy identity is available.

Overall, by introducing the regularisation and taking the corresponding limit last, one separates the discretisation limit from the admissibility problem, which means that the respective restrictions on q are avoided.

The only convergence results available in the literature in the unsteady case concern explicit constitutive relations. There is a range of results on convergence based on error estimates and proving the required regularity for strong solutions under suitable conditions, see for example [PR01b, DPR02, DPR06, BDR09, BDR15]. Since such regularity is out of reach for the more general framework of implicitly constituted fluids we are interested in the approximation of weak solutions. Let us mention one result for explicit constitutive relations, which is closest to what we are aiming for: In [CHP10] the authors show convergence (up to subsequence) for a fully time-discretised approximation scheme including a regularisation and taking the regularisation limit last, for $q > \frac{2(d+1)}{d+2}$. This restriction arises because of the use of an L^∞-truncation.

Since some of the more standard techniques applied in this chapter stem from the theory of Navier–Stokes equations and their numerical approximation, we include a selection of references: [Tem84, GR86, HR82, Ran99].

Options unsteady problem		regularisation with	regularisation without
graph approximation	general	$k,(l,n),m$ Subsection 5.3.1	k,l,n
graph approximation	Yosida	$(k,l,n),m$	(k,l,n) Subsection 5.3.2

Fig. 5.1: Options for the approximation levels for the unsteady problem

As in the steady case there are four options to take the limits, two of which we want to present, see Figure 5.1.

In Subsection 5.3.1 we will consider the approximation levels $(k,l,n,m) \in \mathbb{N}$ and take them successively, first $k \to \infty$, then simultaneously $l,n \to \infty$ (without restrictions on the discretisation parameters), and finally $m \to \infty$. The convergence result in Theorem 5.2 holds for any $q \in \left(\frac{2d}{d+2}, \infty\right)$ independent of whether $\mathbb{V}^n_{\mathrm{div}}$ is exactly divergence-free or not and is inspired by the existence result in [BGMS12]. It corresponds to the steady result proved in Subsection 4.3.1. Compared to the convergence result in [CHP10] for explicit constitutive relations the result in Theorem 5.2 is a twofold extension: from explicit constitutive relations with suitable smoothness conditions to implicit constitutive relations, and from the range $q > \frac{2(d+1)}{d+2}$ to $q > \frac{2d}{d+2}$.

In Subsection 5.3.2 we consider the levels $(k, l, n) \in \mathbb{N}^3$ and take all three limits together without restriction on the respective parameters. As in the steady situation in Subsection 4.3.2 we utilise the generalised Yosida graph approximation for this. The bottleneck for the restrictions on q is the identification of the implicit constitutive relation, for which we assume $q \in \left[\frac{3d+2}{d+2}, \infty\right)$. However, the compactness and convergence results used hold for a larger range of q. We will present the detailed assumptions to show where the restrictions arise from and what one can expect from the (potential future) use of a discrete truncation. The convergence result in Theorem 5.12 generalises the result in [Car07] from explicit to implicit constitutive relations and generalises the finite element setting.

In the following we will be brief in those arguments, which are similar to the corresponding ones in the steady case in Chapter 4.

5.3. Convergence Results

In this section we present the results on the convergence of a sequence of solutions to the respective approximate problem (up to subsequences) to a weak solution of the unsteady problem. The existence on each time level follows by a fixed point argument as in the steady case. Because of the time-stepping we have to consider piecewise constant and continuous, piecewise affine interpolants of the sequence of approximate velocity functions. For each limiting process we derive uniform estimates and use compactness arguments to obtain convergence of subsequences. Then both the equation and the initial conditions have to be identified. The final step of identifying the constitutive relation is the most challenging part.

5.3.1. With Regularisation

Here we introduce all four levels of approximation referred to by $k, l, n, m \in \mathbb{N}$ and take the limits successively, where the limits in l and n are taken simultaneously. Including the regularising term is strictly speaking only required, if $q < \frac{3d+2}{d+2}$, but we present the proof for $p \in \left(\frac{2d}{d+2}, \infty\right)$ to reduce case distinctions. Taking the limit $k \to \infty$ first and separately allows for general sequences of graph approximations satisfying Assumption 3.18. After having taken the limit in k the sequence of approximate solutions satisfies the implicit constitutive relation encoded by $\mathcal{A}$ and thus the Minty type convergence lemma can be applied in the following limits. Thanks to the regularisation in the limit $l, n \to \infty$ we still have enough admissibility to apply Simon's compactness lemma and to avoid a truncation to identify the implicit relation. Only in the last limit $m \to \infty$, when all problems stemming from the discretisation are gone, is admissibility lost, and thus we apply the Aubin–Lions compactness lemma, and a parabolic solenoidal Lipschitz approximation to identify the implicit relation.

The most significant difference compared to the existence proof in [BGMS12] lies in the passage to the limits $l, n \to \infty$ and the identification of the implicit law for $m \to \infty$ since at the time a parabolic solenoidal truncation was not available.

First we will formulate the approximate problem, then state the convergence theorem and the rest of the subsection consists of the proof thereof.

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

$$\mathfrak{L}_i^{k,l,n,m}[\boldsymbol{u};\boldsymbol{v}] := -\,\widetilde{b}(\boldsymbol{u},\boldsymbol{u},\boldsymbol{v}) - \left\langle \boldsymbol{\mathcal{S}}_i^k(\cdot,\mathbf{D}\boldsymbol{u}),\mathbf{D}\boldsymbol{v}\right\rangle_\Omega - \frac{1}{m}\left\langle |\boldsymbol{u}|^{2q'-2}\,\boldsymbol{u},\boldsymbol{v}\right\rangle_\Omega + \langle \boldsymbol{f}_i,\boldsymbol{v}\rangle_\Omega\,, \quad (5.23)$$

for $k,l,n,m \in \mathbb{N}$ and $i \in \{1,\ldots,l\}$, where $\boldsymbol{\mathcal{S}}_i^k$ and $\boldsymbol{f}_i \in \mathrm{W}^{-1,q'}(\Omega)^d$ as defined in (5.20) and $\widetilde{b}(\cdot,\cdot,\cdot)$ extended to the the finite element space, as recalled in (5.10).

The pressure-free formulation of the approximate problem can be stated as follows.

Approximate Problem:

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

$$\boldsymbol{U}_0^{k,l,n,m} = P^n_{\mathrm{div}}\boldsymbol{u}_0, \quad (5.24)$$

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

$$\left\langle \mathrm{d}_t\boldsymbol{U}_i^{k,l,n,m},\boldsymbol{W}\right\rangle_\Omega = \mathfrak{L}_i^{k,l,n,m}[\boldsymbol{U}_i^{k,l,n,m};\boldsymbol{W}] \qquad \text{for all } \boldsymbol{W} \in \mathbb{V}^n_{\mathrm{div}}, \quad (5.25)$$

where P^n_{div} is the L^2-projector onto $\mathbb{V}^n_{\mathrm{div}}$, defined in (2.80) and d_t is the backward temporal difference quotient in (2.71).

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 $\mathrm{d}_t\boldsymbol{U}_i^{k,l,n,m}$.

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 $l,n \to \infty$ simultaneously. Note that the order of the limits is crucial for the convergence proof. To simplify the notation we shall write

$$\boldsymbol{v}^{k,l,n,m} \underset{k}{\to}\underset{(l,n)}{\to}\underset{m}{\to} \boldsymbol{v} \quad \text{in X}, \quad \text{as } k\to\infty,\ l,n\to\infty,\ m\to\infty, \quad (5.26)$$

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 analogous notation for weak and weak* convergence.

Let us now state the main convergence result in this subsection.

Theorem 5.2 (Convergence in the Unsteady Case with Regularisation)**.**
In addition to the assumptions of Definition 5.1 let $\{\boldsymbol{\mathcal{S}}^k\}_{k\in\mathbb{N}}$ satisfy Assumption 3.18 for $M = Q$. For the finite element approximation let Assumption 2.18 on the domain and on the family of simplicial partitions be satisfied. Let $\mathbb{V}^n$, $\mathbb{V}^n_{\mathrm{div}}$ and $\mathbb{Q}^n$ be as introduced in (2.30), (2.32) and (2.31), respectively, and assume that Assumption 2.21 as well as Assumptions 2.23 (i), (ii) are satisfied.
*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 (5.24), (5.25). Moreover, if $q \in \left(\frac{2d}{d+2},\infty\right)$, then there exists a weak solution $(\boldsymbol{u},\mathbf{S})$ of **(PU)** according to Definition 5.1 and for the piecewise constant interpolant $\overline{\boldsymbol{U}}^{k,l,n,m}$ and the continuous, piecewise affine interpolant $\widetilde{\boldsymbol{U}}^{k,l,n,m}$ of $\{\boldsymbol{U}_i^{k,l,n,m}\}_{i\in\{0,\ldots,l\}}$ and the piecewise constant interpolant $\overline{\boldsymbol{\mathcal{S}}}^k(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^{k,l,n})$ of $\{\boldsymbol{\mathcal{S}}_i^k(\cdot,\mathbf{D}\overline{\boldsymbol{U}}^{k,l,n})\}_{i\in\{1,\ldots,l\}}$ as defined in (2.72) and (2.73), (up to not*

relabelled subsequences) one has that

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

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 discretisation parameters δ_l and h_n.

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, exactly as in the steady case in Chapter 4, no modification of the convective term is required and the proof that $\boldsymbol{u}^m$ is divergence-free is simpler.

The rest of this section consists of the proof of Theorem 5.2, which relies on Lemmas 5.3–5.5 dealing with the existence of the discrete solution, and the limit $k\to\infty$, Lemmas 5.6 and 5.7 covering the combined limit $l,n\to\infty$, and Lemmas 5.9 and 5.10 the limit $m\to\infty$.

Limit $k\to\infty$

As in the steady case the proof of existence of a sequence of approximate solutions on each time level is based on a fixed point argument. The a priori estimates and the convergence in Lemmas 5.3 and 5.4 are slightly more complicated than in the steady case, but follow by a standard approach presented, e.g., in [Tem84], with minor modifications required to deal with the time-dependence of $\boldsymbol{\mathcal{S}}^k$. In particular, one has to identify the limits of the piecewise constant version with the piecewise constant version of the limit of the approximate stress function. Taking $k\to\infty$ we remain in the finite-dimensional setting, and hence strong convergence of the sequence of symmetric gradients follows from the bounds on the sequence of the coefficients appearing in the expansions in terms of the basis functions of the finite element spaces. Consequently, as in the steady case the identification of the limiting equation is based on the properties of the sequence $\{\boldsymbol{\mathcal{S}}^k\}_{k\in\mathbb{N}}$ according to Assumption 3.18, see also [BGMS12].

Lemma 5.3 (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 (5.24), (5.25). Furthermore, there exists a constant $c>0$ such that we have that

$$\begin{aligned}
\max_{j\in\{1,\dots,l\}}\|\boldsymbol{U}_j^\kappa\|^2_{\mathrm{L}^2(\Omega)}+\sum_{j=1}^l\|\boldsymbol{U}_j^\kappa-\boldsymbol{U}_{j-1}^\kappa\|^2_{\mathrm{L}^2(\Omega)}+\delta_l\sum_{j=1}^l\|\boldsymbol{U}_j^\kappa\|^q_{\mathrm{W}^{1,q}(\Omega)}\\
+\sum_{j=1}^l\left\|\boldsymbol{\mathcal{S}}^k(\cdot,\cdot,\mathbf{D}\boldsymbol{U}_j^\kappa)\right\|^{q'}_{\mathrm{L}^{q'}(Q^j_{j-1})}+\frac{\delta_l}{m}\sum_{j=1}^l\|\boldsymbol{U}_j^\kappa\|^{2q'}_{\mathrm{L}^{2q'}(\Omega)}\le c,
\end{aligned}\tag{5.27}$$

for all $\kappa=(k,l,n,m)\in\mathbb{N}^4$.

Proof.

Step 1: A priori estimates
The a priori estimates follow from standard arguments for time-discrete approximation, see

[Tem84], in combination with the estimates available for $\mathbf{S}^k$ by Assumption 3.18: testing (5.25) with $\boldsymbol{W} = \boldsymbol{U}_i^\kappa \in \mathbb{V}_{\mathrm{div}}^n$ for $\kappa = (k,l,n,m) \in \mathbb{N}^4$ and $i \in \{1,\dots,l\}$ one obtains

$$\langle \mathrm{d}_t \boldsymbol{U}_i^\kappa, \boldsymbol{U}_i^\kappa\rangle_\Omega + \left\langle \mathbf{S}_i^k(\cdot, \mathbf{D}\boldsymbol{U}_i^\kappa), \mathbf{D}\boldsymbol{U}_i^\kappa \right\rangle_\Omega + \frac{1}{m}\|\boldsymbol{U}_i^\kappa\|_{\mathrm{L}^{2q'}(\Omega)}^{2q'} = \langle \boldsymbol{f}_i, \boldsymbol{U}_i^\kappa\rangle_\Omega \,, \tag{5.28}$$

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 (2.71), the first term in (5.28) can be rewritten as

$$\begin{aligned}\langle \mathrm{d}_t \boldsymbol{U}_i^\kappa, \boldsymbol{U}_i^\kappa\rangle_\Omega &= \frac{1}{\delta_l}\langle \boldsymbol{U}_i^\kappa - \boldsymbol{U}_{i-1}^\kappa, \boldsymbol{U}_i^\kappa\rangle_\Omega \\ &= \frac{1}{2\delta_l}\left(\|\boldsymbol{U}_i^\kappa\|_{\mathrm{L}^2(\Omega)}^2 - \|\boldsymbol{U}_{i-1}^\kappa\|_{\mathrm{L}^2(\Omega)}^2 + \|\boldsymbol{U}_i^\kappa - \boldsymbol{U}_{i-1}^\kappa\|_{\mathrm{L}^2(\Omega)}^2\right).\end{aligned} \tag{5.29}$$

Using the definition of $\mathbf{S}_i^k$ in (5.20) and Assumption 3.18 ($\sigma 2$) one has that

$$\begin{aligned}\left\langle \mathbf{S}_i^k(\cdot, \mathbf{D}\boldsymbol{U}_i^\kappa), \mathbf{D}\boldsymbol{U}_i^\kappa \right\rangle_\Omega &\overset{(5.20)}{=} \left\langle \fint_{t_{i-1}}^{t_i} \mathbf{S}^k(t,\cdot,\mathbf{D}\boldsymbol{U}_i^\kappa)\,\mathrm{d}t, \mathbf{D}\boldsymbol{U}_i^\kappa \right\rangle_\Omega \\ &= \frac{1}{\delta_l}\left\langle \mathbf{S}^k(\cdot,\cdot,\mathbf{D}\boldsymbol{U}_i^\kappa), \mathbf{D}\boldsymbol{U}_i^\kappa \right\rangle_{Q_{i-1}^i} \\ &\geq \frac{1}{\delta_l}\int_{Q_{i-1}^i} -\widetilde{g}(\cdot) + \widetilde{c}_*\left(|\mathbf{D}\boldsymbol{U}_i^\kappa|^q + \left|\mathbf{S}^k(\cdot,\cdot,\mathbf{D}\boldsymbol{U}_i^\kappa)\right|^{q'}\right)\mathrm{d}\boldsymbol{z} \\ &\geq -\frac{1}{\delta_l}\|\widetilde{g}\|_{\mathrm{L}^1(Q_{i-1}^i)} + c\,\|\boldsymbol{U}_i^\kappa\|_{\mathrm{W}^{1,q}(\Omega)}^q + \frac{\widetilde{c}_*}{\delta_l}\left\|\mathbf{S}^k(\cdot,\cdot,\mathbf{D}\boldsymbol{U}_i^\kappa)\right\|_{\mathrm{L}^{q'}(Q_{i-1}^i)}^{q'},\end{aligned} \tag{5.30}$$

where the last inequality follows by Korn's and Poincaré's inequalities. For the term on the right-hand side of (5.28) by duality of norms and by Young's inequality with $\varepsilon > 0$ we obtain that

$$\begin{aligned}\langle \boldsymbol{f}_i, \boldsymbol{U}_i^\kappa\rangle_\Omega &\leq \|\boldsymbol{f}_i\|_{\mathrm{W}^{-1,q'}(\Omega)}\|\boldsymbol{U}_i^\kappa\|_{\mathrm{W}^{1,q}(\Omega)} \leq c(\varepsilon)\,\|\boldsymbol{f}_i\|_{\mathrm{W}^{-1,q'}(\Omega)}^{q'} + \varepsilon\,\|\boldsymbol{U}_i^\kappa\|_{\mathrm{W}^{1,q}(\Omega)}^q \\ &\leq \frac{c(\varepsilon)}{\delta_l}\|\boldsymbol{f}\|_{\mathrm{L}^{q'}(t_{i-1},t_i;\mathrm{W}^{-1,q'}(\Omega))}^{q'} + \varepsilon\,\|\boldsymbol{U}_i^\kappa\|_{\mathrm{W}^{1,q}(\Omega)}^q\,,\end{aligned} \tag{5.31}$$

where the last inequality follows by (5.21). Applying the estimates (5.29)–(5.31) in (5.28), after rearranging, choosing $\varepsilon > 0$ small enough and multiplying by δ_l, we arrive at

$$\begin{aligned}&\|\boldsymbol{U}_i^\kappa\|_{\mathrm{L}^2(\Omega)}^2 - \|\boldsymbol{U}_{i-1}^\kappa\|_{\mathrm{L}^2(\Omega)}^2 + \|\boldsymbol{U}_i^\kappa - \boldsymbol{U}_{i-1}^\kappa\|_{\mathrm{L}^2(\Omega)}^2 \\ &\quad + \delta_l\,\|\boldsymbol{U}_i^\kappa\|_{\mathrm{W}^{1,q}(\Omega)}^q + \left\|\mathbf{S}^k(\cdot,\cdot,\mathbf{D}\boldsymbol{U}_i^\kappa)\right\|_{\mathrm{L}^{q'}(Q_{i-1}^i)}^{q'} + \frac{\delta_l}{m}\|\boldsymbol{U}_i^\kappa\|_{\mathrm{L}^{2q'}(\Omega)}^{2q'} \\ &\quad \leq c\left(\|\boldsymbol{f}\|_{\mathrm{L}^{q'}(t_{i-1},t_i;\mathrm{W}^{-1,q'}(\Omega))}^{q'} + \|\widetilde{g}\|_{\mathrm{L}^1(Q_{i-1}^i)}\right).\end{aligned} \tag{5.32}$$

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

$$\begin{aligned}&\|\boldsymbol{U}_j^\kappa\|_{\mathrm{L}^2(\Omega)}^2 - \|\boldsymbol{U}_0^\kappa\|_{\mathrm{L}^2(\Omega)}^2 + \sum_{i=1}^j \|\boldsymbol{U}_i^\kappa - \boldsymbol{U}_{i-1}^\kappa\|_{\mathrm{L}^2(\Omega)}^2 + \delta_l \sum_{i=1}^j \|\boldsymbol{U}_i^\kappa\|_{\mathrm{W}^{1,q}(\Omega)}^q \\ &\qquad + \sum_{i=1}^j \left\|\mathbf{S}^k(\cdot,\cdot,\mathbf{D}\boldsymbol{U}_i^\kappa)\right\|_{\mathrm{L}^{q'}(Q_{i-1}^i)}^{q'} + \frac{\delta_l}{m}\sum_{i=1}^j \|\boldsymbol{U}_i^\kappa\|_{\mathrm{L}^{2q'}(\Omega)}^{2q'}\end{aligned} \tag{5.33}$$

$$\leq c(\|\boldsymbol{f}\|^{q'}_{\mathrm{L}^{q'}(0,T;\mathrm{W}^{-1,q'}(\Omega))} + \|\widetilde{g}\|_{\mathrm{L}^1(Q)})$$

because of cancellation in the first term. Applying the estimate

$$\|\boldsymbol{U}^\kappa_0\|^2_{\mathrm{L}^2(\Omega)} \overset{(5.24)}{=} \|P^n_{\mathrm{div}}\boldsymbol{u}_0\|^2_{\mathrm{L}^2(\Omega)} \overset{(2.81)}{\leq} \|\boldsymbol{u}_0\|^2_{\mathrm{L}^2(\Omega)}\,, \tag{5.34}$$

and taking the supremum over all $j \in \{1,\dots,l\}$ in (5.33) finishes the proof of (5.27).

Step 2: Existence of $\{\boldsymbol{U}^\kappa_i\}_{i\in\{0,\dots,l\}} \subset \mathbb{V}^n_{\mathrm{div}}$
Let $\kappa \in \mathbb{N}^4$ be fixed. Since $\boldsymbol{U}^\kappa_0 = P^n_{\mathrm{div}}\boldsymbol{u}_0$ by (5.24), one only has to show that for a given $\boldsymbol{U}^\kappa_{i-1} \in \mathbb{V}^n_{\mathrm{div}}$, there exists a $\boldsymbol{U}^\kappa_i \in \mathbb{V}^n_{\mathrm{div}}$ such that (5.25) is satisfied. Since $\boldsymbol{\mathcal{S}}^k_i(\boldsymbol{z},\cdot)$ is continuous, the existence of such a $\boldsymbol{U}^\kappa_i \in \mathbb{V}^n_{\mathrm{div}}$ follows by a standard argument from Brouwer's fixed point theorem, similarly as in Chapter 4 in Step 2 in the proof of Lemma 4.3. ☐

For $\kappa = (k,l,n,m) \in \mathbb{N}^4$ let the sequence of coefficients $\{\boldsymbol{\alpha}^\kappa_i\}_{i\in\{0,\dots,l\}} \subset \mathbb{R}^{d_n}$ be such that $\boldsymbol{U}^\kappa_i = \sum_{j=1}^{d_n}(\alpha^\kappa_i)_j \boldsymbol{W}_j$, where $\{\boldsymbol{W}_1,\dots,\boldsymbol{W}_{d_n}\}$ is a basis of $\mathbb{V}^n_{\mathrm{div}}$. Uniqueness is in general not guaranteed, so we choose one such sequence for each $\kappa \in \mathbb{N}^4$. Let $\overline{\boldsymbol{\alpha}}^\kappa \in \mathbb{P}^l_0(0,T;\mathbb{R}^{d_n}) \subset \mathrm{L}^\infty(0,T)^{d_n}$ and $\widetilde{\boldsymbol{\alpha}}^\kappa \in \mathbb{P}^l_1(0,T;\mathbb{R}^{d_n}) \subset \mathrm{W}^{1,\infty}(0,T)^{d_n}$ be the piecewise constant and piecewise affine interpolants as in (2.72) and (2.73). We denote

$$\begin{aligned}\overline{\boldsymbol{U}}^\kappa(t,\boldsymbol{x}) &:= \sum_{j=1}^{d_n}\overline{\alpha}^\kappa_j(t)\boldsymbol{W}_j(\boldsymbol{x}) \in \mathbb{P}^l_0(0,T;\mathbb{V}^n_{\mathrm{div}}),\\ \widetilde{\boldsymbol{U}}^\kappa(t,\boldsymbol{x}) &:= \sum_{j=1}^{d_n}\widetilde{\alpha}^\kappa_j(t)\boldsymbol{W}_j(\boldsymbol{x}) \in \mathbb{P}^l_1(0,T;\mathbb{V}^n_{\mathrm{div}}).\end{aligned} \tag{5.35}$$

These functions coincide with the respective interpolants of $\{\boldsymbol{U}^\kappa_i\}_{i\in\{0,\dots,l\}}$, defined in (2.72) and (2.73). Furthermore, for $t \in (0,T]$, $\boldsymbol{u} \in \mathbb{P}^l_0(0,T;\mathbb{V}^n)$ and $\boldsymbol{v} \in \mathbb{V}^n$ we introduce

$$\begin{aligned}\mathfrak{L}^\kappa[\boldsymbol{u};\boldsymbol{v}](t) := &-\widetilde{b}(\boldsymbol{u}(t,\cdot),\boldsymbol{u}(t,\cdot),\boldsymbol{v}) - \left\langle \overline{\boldsymbol{\mathcal{S}}}^k(t,\cdot,\mathbf{D}\boldsymbol{u}(t,\cdot)),\mathbf{D}\boldsymbol{v}\right\rangle_\Omega\\ &-\frac{1}{m}\left\langle |\boldsymbol{u}(t,\cdot)|^{2q'-2}\boldsymbol{u}(t,\cdot),\boldsymbol{v}\right\rangle_\Omega + \langle \overline{\boldsymbol{f}}(t,\cdot),\boldsymbol{v}\rangle_\Omega\end{aligned} \tag{5.36}$$

for $\kappa = (k,l,n,m) \in \mathbb{N}^4$. Recall that $\overline{\boldsymbol{f}} \in \mathbb{P}^l_0(0,T;\mathrm{W}^{-1,q'}(\Omega)^d)$ is the piecewise constant interpolant of $\{\boldsymbol{f}_i\}_{i\in\{1,\dots,l\}}$, as in (5.20), see (2.72) in Subsection 2.2.3.1. Similarly, $\overline{\boldsymbol{\mathcal{S}}}^k(t,\cdot,\cdot) = \boldsymbol{\mathcal{S}}^k_i(\cdot,\cdot)$, for $t \in (t_{i-1},t_i]$, so it is piecewise constant with respect to the variable $t \in (0,T]$.

Lemma 5.4 (Equation for $t \in (0,T]$ and Convergence $k \to \infty$)**.**
The functions $\overline{\boldsymbol{U}}^\kappa$ *and* $\widetilde{\boldsymbol{U}}^\kappa$ *defined as in* (5.35) *satisfy*

$$\left\langle \partial_t\widetilde{\boldsymbol{U}}^\kappa(t,\cdot),\boldsymbol{W}\right\rangle_\Omega = \mathfrak{L}^\kappa[\overline{\boldsymbol{U}}^\kappa;\boldsymbol{W}](t) \qquad \textit{for all } \boldsymbol{W} \in \mathbb{V}^n_{\mathrm{div}},\ \textit{for all } t \in (0,T], \tag{5.37}$$

$$\widetilde{\boldsymbol{U}}^\kappa(0,\cdot) = P^n_{\mathrm{div}}\boldsymbol{u}_0 \qquad \textit{in } \Omega, \tag{5.38}$$

for any $\kappa = (k,l,n,m) \in \mathbb{N}^4$*. For each* $\lambda := (l,n,m) \in \mathbb{N}^3$*, there exists a sequence* $\{\boldsymbol{U}^\lambda_i\}_{i\in\{0,\dots,l\}} \subset \mathbb{V}^n_{\mathrm{div}}$*, and subsequences such that the piecewise constant and continuous piecewise affine interpolants* $\overline{\boldsymbol{U}}^\lambda \in \mathbb{P}^l_0(0,T;\mathbb{V}^n_{\mathrm{div}})$ *and* $\widetilde{\boldsymbol{U}}^\lambda \in \mathbb{P}^l_1(0,T;\mathbb{V}^n_{\mathrm{div}})$ *of* $\{\boldsymbol{U}^\lambda_i\}_{i\in\{0,\dots,l\}}$*,* *as defined in* (2.72) *and* (2.73), *satisfy*

$$\widetilde{\boldsymbol{U}}^{k,\lambda} \to \widetilde{\boldsymbol{U}}^{\lambda} \qquad \textit{strongly in } \mathrm{C}([0,T];\mathrm{C}(\overline{\Omega})^d), \tag{5.39}$$

$$\partial_t\widetilde{\boldsymbol{U}}^{k,\lambda} \to \partial_t\widetilde{\boldsymbol{U}}^{\lambda} \qquad \textit{strongly in } \mathrm{L}^\infty(0,T;\mathrm{C}(\overline{\Omega})^d), \tag{5.40}$$

$$\overline{\boldsymbol{U}}^{k,\lambda} \to \overline{\boldsymbol{U}}^{\lambda} \qquad \textit{strongly in } \mathrm{L}^\infty(0,T;\mathrm{C}(\overline{\Omega})^d), \tag{5.41}$$

$$\mathbf{D}\overline{\boldsymbol{U}}^{k,\lambda} \to \mathbf{D}\overline{\boldsymbol{U}}^{\lambda} \qquad \textit{strongly in } \mathrm{L}^\infty(0,T;\mathrm{L}^\infty(\Omega)^{d\times d}), \tag{5.42}$$

$$\nabla\overline{\boldsymbol{U}}^{k,\lambda} \to \nabla\overline{\boldsymbol{U}}^{\lambda} \qquad \textit{strongly in } \mathrm{L}^\infty(0,T;\mathrm{L}^\infty(\Omega)^{d\times d}), \tag{5.43}$$

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

$$\mathcal{S}^k(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^{k,\lambda}) \rightharpoonup \mathbf{S}^\lambda \qquad \textit{weakly in } \mathrm{L}^{q'}(Q)^{d\times d}, \tag{5.44}$$

$$\overline{\mathcal{S}}^k(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^{k,\lambda}) \rightharpoonup \overline{\mathbf{S}}^\lambda \qquad \textit{weakly in } \mathrm{L}^{q'}(Q)^{d\times d}, \tag{5.45}$$

as $k \to \infty$, *and (up to a representative) we have that*

$$\overline{\mathbf{S}}^\lambda(t,\cdot) = \mathbf{S}_i^\lambda(\cdot) := \fint_{t_{i-1}}^{t_i} \mathbf{S}^\lambda(t,\cdot)\,\mathrm{d}t \qquad \textit{for all } t \in (t_{i-1},t_i],\ i \in \{1,\dots,l\}. \tag{5.46}$$

Proof.

Step 1: Identification of the equation
We have that $\widetilde{\boldsymbol{U}}^\kappa(0,\cdot) = \boldsymbol{U}_0^\kappa = P_{\mathrm{div}}^n \boldsymbol{u}_0$ by definition of $\widetilde{\boldsymbol{U}}^\kappa$ and by (5.24), which shows (5.38). The equation (5.37) follows from (5.25) and the fact that for $t \in (t_{i-1},t_i]$ we have that

$$\overline{\boldsymbol{U}}^\kappa(t,\cdot) = \boldsymbol{U}_i^\kappa, \quad \partial_t\widetilde{\boldsymbol{U}}^\kappa(t,\cdot) = \mathrm{d}_t\boldsymbol{U}_i^\kappa, \quad \overline{\boldsymbol{f}}(t,\cdot) = \boldsymbol{f}_i \quad \text{and} \quad \overline{\mathcal{S}}^k(t,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^\kappa) = \mathcal{S}_i^k(\cdot,\mathbf{D}\boldsymbol{U}_i^\kappa),$$

for any $i \in \{1,\dots,l\}$.

Step 2: Estimates
Let $\lambda = (l,n,m) \in \mathbb{N}^3$ be arbitrary, but fixed. When taking $k \to \infty$ we stay in the finite-dimensional setting, hence it suffices to focus on estimates for the coefficient functions $\overline{\boldsymbol{\alpha}}^\kappa = \overline{\boldsymbol{\alpha}}^{k,\lambda}$ and $\widetilde{\boldsymbol{\alpha}}^\kappa = \widetilde{\boldsymbol{\alpha}}^{k,\lambda}$, uniformly in $k \in \mathbb{N}$. By use of an L^2-orthonormal basis of $\mathbb{V}_{\mathrm{div}}^n$, as in (4.31), one can show that

$$|\boldsymbol{\alpha}|^2 \le c(n)\,\|\boldsymbol{U}\|_{\mathrm{L}^2(\Omega)}^2 \qquad \text{for } \boldsymbol{U} = \sum_j^{d_n} \alpha_j \boldsymbol{W}_j. \tag{5.47}$$

Using the definition of $\overline{\boldsymbol{U}}^{k,\lambda}$ in (5.35) with this estimate, the fact that the interpolants are piecewise constant and also the a priori estimate in (5.27), we obtain

$$\left|\overline{\boldsymbol{\alpha}}^{k,\lambda}(t)\right|^2 \overset{(5.47)}{\le} c(n)\left\|\overline{\boldsymbol{U}}^{k,\lambda}(t,\cdot)\right\|_{\mathrm{L}^2(\Omega)}^2 \overset{(2.75)}{\le} c(n)\max_{i\in\{1,\dots,l\}}\left\|\boldsymbol{U}_i^{k,\lambda}\right\|_{\mathrm{L}^2(\Omega)}^2 \overset{(5.27)}{\le} c(n)$$

for any $t \in (0,T)$ uniformly in $k \in \mathbb{N}$. Using the following estimate, based on the L^2-stability of P_{div}^n,

$$\left\|\boldsymbol{U}_0^{k,\lambda}\right\|_{\mathrm{L}^2(\Omega)} \overset{(5.38)}{=} \|P_{\mathrm{div}}^n\boldsymbol{u}_0\|_{\mathrm{L}^2(\Omega)} \overset{(2.81)}{\le} \|\boldsymbol{u}_0\|_{\mathrm{L}^2(\Omega)}\,, \tag{5.48}$$

we also have that

$$\left|\widetilde{\boldsymbol{\alpha}}^{k,\lambda}(t)\right|^2 \overset{(5.47)}{\leq} c(n)\left\|\widetilde{\boldsymbol{U}}^{k,\lambda}(t,\cdot)\right\|_{\mathrm{L}^2(\Omega)}^2 \overset{(2.76)}{\leq} c(n)\max_{i\in\{0,\dots,l\}}\left\|\boldsymbol{U}_i^{k,\lambda}\right\|_{\mathrm{L}^2(\Omega)}^2 \overset{(5.27),(5.48)}{\leq} c(n)$$

for any $t \in (0,T)$ uniformly in $k \in \mathbb{N}$. Thus, we have that

$$\left\|\overline{\boldsymbol{\alpha}}^{k,\lambda}\right\|_{\mathrm{L}^\infty(0,T)} + \left\|\widetilde{\boldsymbol{\alpha}}^{k,\lambda}\right\|_{\mathrm{L}^\infty(0,T)} \leq c \quad \text{for all } k \in \mathbb{N}. \tag{5.49}$$

Since the space of continuous piecewise affine functions $\mathbb{P}_1^l(0,T;\mathbb{R}^{d_n}) \subset \mathrm{W}^{1,\infty}(0,T)^{d_n}$ with respect to the time grid $\{t_0,\dots,t_l\} \subset [0,T]$ is finite-dimensional, all norms on it are equivalent, with the norm-equivalence constants depending on the (here fixed) dimension. This shows that we also have that

$$\left\|\widetilde{\boldsymbol{\alpha}}^{k,\lambda}\right\|_{\mathrm{W}^{1,\infty}(0,T)} \leq c(l,n) \quad \text{for all } k \in \mathbb{N}. \tag{5.50}$$

It is a direct consequence of the a priori estimate (5.27) that

$$\left\|\boldsymbol{\mathcal{S}}^k\left(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^{k,\lambda}\right)\right\|_{\mathrm{L}^{q'}(Q)}^{q'} \leq c \quad \text{for all } k \in \mathbb{N}, \tag{5.51}$$

and then, by the definition of $\boldsymbol{\mathcal{S}}_i^k$ in (5.20), it follows that

$$\begin{aligned}
\left\|\overline{\boldsymbol{\mathcal{S}}}^k(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^{k,\lambda})\right\|_{\mathrm{L}^{q'}(Q)}^{q'} &= \sum_{i=1}^{l}\left\|\boldsymbol{\mathcal{S}}_i^k(\cdot,\mathbf{D}\boldsymbol{U}_i^{k,\lambda})\right\|_{\mathrm{L}^{q'}(Q_{i-1}^i)}^{q'} = \sum_{i=1}^{l}\delta_l\left\|\boldsymbol{\mathcal{S}}_i^k(\cdot,\mathbf{D}\boldsymbol{U}_i^{k,\lambda})\right\|_{\mathrm{L}^{q'}(\Omega)}^{q'} \\
&\overset{(2.78)}{\leq} \sum_{i=1}^{l}\left\|\boldsymbol{\mathcal{S}}^k(\cdot,\cdot,\mathbf{D}\boldsymbol{U}_i^{k,\lambda})\right\|_{\mathrm{L}^{q'}(Q_{i-1}^i)}^{q'} \\
&= \left\|\boldsymbol{\mathcal{S}}^k(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^{k,\lambda})\right\|_{\mathrm{L}^{q'}(Q)}^{q'} \overset{(5.51)}{\leq} c
\end{aligned} \tag{5.52}$$

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

$$\left\|\boldsymbol{\mathcal{S}}_i^k(\cdot,\mathbf{D}\boldsymbol{U}_i^{k,\lambda})\right\|_{\mathrm{L}^{q'}(\Omega)}^{q'} \leq \frac{c}{\delta_l} \leq c(l) \quad \text{for any } k \in \mathbb{N},\ i \in \{1,\dots,l\}. \tag{5.53}$$

Step 3: Convergence as $k \to \infty$

Let again $\lambda \in \mathbb{N}^3$ be arbitrary but fixed. Since $\{\overline{\boldsymbol{\alpha}}^{k,\lambda}\}_{k\in\mathbb{N}} \subset \mathbb{P}_0^l(0,T;\mathbb{R}^{d_n})$ and the space $\mathbb{P}_0^l(0,T;\mathbb{R}^{d_n})$ is finite-dimensional, (5.49) implies strong convergence of a subsequence, i.e., there exists an $\overline{\boldsymbol{\alpha}}^\lambda \in \mathbb{P}_0^l(0,T;\mathbb{R}^{d_n})$ such that

$$\overline{\boldsymbol{\alpha}}^{k,\lambda} \to \overline{\boldsymbol{\alpha}}^\lambda \quad \text{strongly in } \mathrm{L}^\infty(0,T)^{d_n}, \quad \text{as } k \to \infty. \tag{5.54}$$

Similarly, we obtain from (5.50) that there exists a subsequence and an $\widetilde{\boldsymbol{\alpha}}^\lambda \in \mathbb{P}_1^l(0,T;\mathbb{R}^{d_n})$ such that

$$\widetilde{\boldsymbol{\alpha}}^{k,\lambda} \to \widetilde{\boldsymbol{\alpha}}^\lambda \quad \text{strongly in } \mathrm{W}^{1,\infty}(0,T)^{d_n}, \quad \text{as } k \to \infty. \tag{5.55}$$

Note that the convergence holds pointwise everywhere in $(0,T]$, and hence,

$$\overline{\boldsymbol{\alpha}}^{\lambda}(t_i) \leftarrow \overline{\boldsymbol{\alpha}}^{k,\lambda}(t_i) = \boldsymbol{\alpha}_i^{k,\lambda} = \widetilde{\boldsymbol{\alpha}}^{k,\lambda}(t_i) \to \widetilde{\boldsymbol{\alpha}}^{\lambda}(t_i), \qquad \text{as } k \to \infty,$$

for any $i \in \{1,\dots,l\}$, so the limits coincide and we can set $\boldsymbol{\alpha}_i^{\lambda} = \widetilde{\boldsymbol{\alpha}}^{\lambda}(t_i) = \overline{\boldsymbol{\alpha}}^{\lambda}(t_i)$, for $i \in \{1,\dots,l\}$. Since we also have that $\widetilde{\boldsymbol{\alpha}}^{k,\lambda}(0,\cdot) \to \widetilde{\boldsymbol{\alpha}}^{\lambda}(0,\cdot)$, we can set $\boldsymbol{\alpha}_0^{\lambda} = \widetilde{\boldsymbol{\alpha}}^{\lambda}(0,\cdot)$. Then $\overline{\boldsymbol{\alpha}}^{\lambda} \in \mathbb{P}_0^l(0,T;\mathbb{R}^{d_n})$ and $\widetilde{\boldsymbol{\alpha}}^{\lambda} \in \mathbb{P}_1^l(0,T;\mathbb{R}^{d_n})$ are the piecewise constant and continuous piecewise affine interpolants of $\{\boldsymbol{\alpha}_i^{\lambda}\}_{i\in\{0,\dots,l\}}$, respectively. Let us set

$$\overline{\boldsymbol{U}}^{\lambda}(t,\boldsymbol{x}) \coloneqq \sum_{j=1}^{d_n} \overline{\alpha}_j^{\lambda}(t)\boldsymbol{W}_j(\boldsymbol{x}), \qquad \widetilde{\boldsymbol{U}}^{\lambda}(t,\boldsymbol{x}) \coloneqq \sum_{j=1}^{d_n} \widetilde{\alpha}_j^{\lambda}(t)\boldsymbol{W}_j(\boldsymbol{x}), \tag{5.56}$$

and let $\boldsymbol{U}_i^{\lambda} = \sum_{j=1}^{d_n} (\alpha_i^{\lambda})_j \boldsymbol{W}_j \in \mathbb{V}_{\mathrm{div}}^n$ for $i \in \{0,\dots,l\}$. Note that by the above considerations concerning the coefficients we have that $\overline{\boldsymbol{U}}^{\lambda}$ and $\widetilde{\boldsymbol{U}}^{\lambda}$ coincide with the respective interpolants of $\{\boldsymbol{U}_i^{\lambda}\}_{i\in\{0,\dots,l\}}$. By the convergence in (5.54), (5.55) one obtains for the so-defined functions the convergence results (5.41)–(5.40), as $k\to\infty$. By the Banach–Alaoglu theorem, (5.51)–(5.53) 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

$$\begin{aligned}
\mathcal{S}^k(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^{k,\lambda}) &\rightharpoonup \mathbf{S}^{\lambda} && \text{weakly in } \mathrm{L}^{q'}(Q)^{d\times d}, && (5.57)\\
\overline{\mathcal{S}}^k(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^{k,\lambda}) &\rightharpoonup \overline{\mathbf{S}}^{\lambda} && \text{weakly in } \mathrm{L}^{q'}(Q)^{d\times d}, && (5.58)\\
\mathcal{S}_i^k(\cdot,\mathbf{D}\boldsymbol{U}_i^{k,\lambda}) &\rightharpoonup \mathbf{S}_i^{\lambda} && \text{weakly in } \mathrm{L}^{q'}(\Omega)^{d\times d}, \quad \text{for } i \in \{1,\dots,l\}, && (5.59)
\end{aligned}$$

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 (5.58) we have that

$$\left\langle \overline{\mathcal{S}}^k(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^{k,\lambda}), \varphi\mathbf{v} \right\rangle_{Q_{i-1}^i} \to \left\langle \overline{\mathbf{S}}^{\lambda}, \varphi\mathbf{v} \right\rangle_{Q_{i-1}^i}, \qquad \text{as } k\to\infty. \tag{5.60}$$

On the other hand by the definition of $\overline{\mathcal{S}}^k(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^{k,\lambda})$ as piecewise constant interpolant of the sequence $\{\mathcal{S}_i^k(\cdot,\mathbf{D}\boldsymbol{U}_i^{k,\lambda})\}_{i\in\{1,\dots,l\}}$ and by (5.59) we have that

$$\begin{aligned}
\left\langle \overline{\mathcal{S}}^k(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^{k,\lambda}), \varphi\mathbf{v} \right\rangle_{Q_{i-1}^i} &= \left\langle \left\langle \mathcal{S}_i^k(\cdot,\mathbf{D}\boldsymbol{U}_i^{k,\lambda}), \mathbf{v} \right\rangle_{\Omega}, \varphi \right\rangle_{(t_{i-1},t_i)} \\
&= \langle 1,\varphi\rangle_{(t_{i-1},t_i)} \left\langle \mathcal{S}_i^k(\cdot,\mathbf{D}\boldsymbol{U}_i^{k,\lambda}), \mathbf{v} \right\rangle_{\Omega} \\
&\to \langle 1,\varphi\rangle_{(t_{i-1},t_i)} \left\langle \mathbf{S}_i^{\lambda}, \mathbf{v} \right\rangle_{\Omega} = \left\langle \mathbf{S}_i^{\lambda}, \mathbf{v}\varphi \right\rangle_{Q_{i-1}^i},
\end{aligned} \tag{5.61}$$

as $k\to\infty$. Now, (5.60) and (5.61) 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 (5.59) that

$$\left\langle \mathcal{S}_i^k(\cdot,\mathbf{D}\boldsymbol{U}_i^{k,\lambda}), \mathbf{v} \right\rangle_{\Omega} \to \left\langle \mathbf{S}_i^{\lambda}, \mathbf{v} \right\rangle_{\Omega}, \qquad \text{as } k\to\infty. \tag{5.62}$$

On the other hand by the definition of $\mathcal{S}_i^k(\cdot, \mathbf{D}\boldsymbol{U}_i^{k,\lambda})$ in (5.20) and by (5.57) we obtain that

$$\begin{aligned}
\left\langle \mathcal{S}_i^k(\cdot, \mathbf{D}\boldsymbol{U}_i^{k,\lambda}), \mathbf{v} \right\rangle_\Omega &= \left\langle \fint_{t_{i-1}}^{t_i} \mathcal{S}^k(t, \cdot, \mathbf{D}\boldsymbol{U}_i^{k,\lambda})\, \mathrm{d}t, \mathbf{v} \right\rangle_\Omega \\
&= \frac{1}{\delta_l} \left\langle \mathcal{S}^k(\cdot, \cdot, \mathbf{D}\overline{\boldsymbol{U}}^{k,\lambda}), \mathbb{1}_{(t_{i-1},t_i)}\mathbf{v} \right\rangle_Q \\
&\to \frac{1}{\delta_l} \left\langle \mathbf{S}^\lambda, \mathbb{1}_{(t_{i-1},t_i)}\mathbf{v} \right\rangle_Q = \left\langle \fint_{t_{i-1}}^{t_i} \mathbf{S}^\lambda(t, \cdot)\, \mathrm{d}t, \mathbf{v} \right\rangle_\Omega ,
\end{aligned} \tag{5.63}$$

as $k \to \infty$, so by the uniqueness of limits, we conclude from (5.62), (5.63), 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

$$\begin{aligned}
\mathfrak{L}^\lambda[\boldsymbol{u}; \boldsymbol{v}](t) := &- \widetilde{b}(\boldsymbol{u}(t, \cdot), \boldsymbol{u}(t, \cdot), \boldsymbol{v}) - \left\langle \overline{\mathbf{S}}^\lambda(t, \cdot), \mathbf{D}\boldsymbol{v} \right\rangle_\Omega \\
&- \frac{1}{m} \left\langle |\boldsymbol{u}(t, \cdot)|^{2q'-2}\, \boldsymbol{u}(t, \cdot), \boldsymbol{v} \right\rangle_\Omega + \langle \overline{\boldsymbol{f}}(t, \cdot), \boldsymbol{v} \rangle_\Omega ,
\end{aligned} \tag{5.64}$$

where $\overline{\mathbf{S}}^\lambda \in \mathbb{P}_0^l(0, T; \mathrm{L}^{q'}(\Omega)^{d\times d})$ is given in Lemma 5.4.

Furthermore, for $\lambda = (l, n, m) \in \mathbb{N}^3$, $i \in \{1, \ldots, l\}$ and $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{V}^n$ let us denote

$$\mathfrak{L}_i^\lambda[\boldsymbol{u}; \boldsymbol{v}] := -\widetilde{b}(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{v}) - \left\langle \mathbf{S}_i^\lambda, \mathbf{D}\boldsymbol{v} \right\rangle_\Omega - \frac{1}{m} \left\langle |\boldsymbol{u}|^{2q'-2}\, \boldsymbol{u}, \boldsymbol{v} \right\rangle_\Omega + \langle \boldsymbol{f}_i, \boldsymbol{v} \rangle_\Omega , \tag{5.65}$$

where $\mathbf{S}_i^\lambda \in \mathrm{L}^{q'}(\Omega)^{d\times d}$ is given in Lemma 5.4 (5.46).

Lemma 5.5 (Identification of the PDE as $k \to \infty$)**.**
The functions $\overline{\boldsymbol{U}}^\lambda \in \mathbb{P}_0^l(0, T; \mathbb{V}_{\mathrm{div}}^n)$, $\widetilde{\boldsymbol{U}}^\lambda \in \mathbb{P}_1^l(0, T; \mathbb{V}_{\mathrm{div}}^n)$ *and* $\mathbf{S}^\lambda \in \mathrm{L}^{q'}(Q)^{d\times d}$ *given in Lemma 5.4 satisfy*

$$\left\langle \partial_t \widetilde{\boldsymbol{U}}^\lambda(t, \cdot), \boldsymbol{W} \right\rangle_\Omega = \mathfrak{L}^\lambda[\overline{\boldsymbol{U}}^\lambda; \boldsymbol{W}](t) \qquad \text{for all } \boldsymbol{W} \in \mathbb{V}_{\mathrm{div}}^n, \text{ for all } t \in (0, T], \tag{5.66}$$

$$\widetilde{\boldsymbol{U}}^\lambda(0, \cdot) = P_{\mathrm{div}}^n \boldsymbol{u}_0(\cdot) \qquad \text{in } \Omega, \tag{5.67}$$

$$(\mathbf{D}\overline{\boldsymbol{U}}^\lambda(\boldsymbol{z}), \mathbf{S}^\lambda(\boldsymbol{z})) \in \mathcal{A}(\boldsymbol{z}) \qquad \text{for a.e. } \boldsymbol{z} \in Q, \tag{5.68}$$

for all $\lambda = (l, n, m) \in \mathbb{N}^3$, *where* $\mathfrak{L}^\lambda[\cdot; \cdot](\cdot)$ *is defined by (5.64), using* $\overline{\mathbf{S}}^\lambda \in \mathbb{P}_0^l(0, T; \mathrm{L}^{q'}(\Omega)^{d\times d})$ *given by (5.46) in Lemma 5.4. Furthermore, the sequence* $\{\boldsymbol{U}_i^\lambda\}_{i\in\{0,\ldots,l\}} \subset \mathbb{V}_{\mathrm{div}}^n$ *given in Lemma 5.4 satisfies*

$$\boldsymbol{U}_0^\lambda = P_{\mathrm{div}}^n \boldsymbol{u}_0, \tag{5.69}$$

$$\left\langle \mathrm{d}_t \boldsymbol{U}_i^\lambda, \boldsymbol{W} \right\rangle_\Omega = \mathfrak{L}_i^\lambda[\boldsymbol{U}_i^\lambda; \boldsymbol{W}] \qquad \text{for all } \boldsymbol{W} \in \mathbb{V}_{\mathrm{div}}^n, \quad \text{for all } i \in \{1, \ldots, l\}. \tag{5.70}$$

Proof. Let $\lambda = (l, n, m) \in \mathbb{N}^3$ be arbitrary but fixed.

Step 1: Identification of the initial condition
By (5.38) the family of continuous functions $\widetilde{\boldsymbol{U}}^{k,\lambda}$ satisfies the initial condition $\widetilde{\boldsymbol{U}}^{k,\lambda}(0, \cdot) = P_{\mathrm{div}}^n \boldsymbol{u}_0(\cdot)$ in Ω for all $k \in \mathbb{N}$. Thus, the sequence $\{\widetilde{\boldsymbol{U}}^{k,\lambda}(0, \cdot)\}_{k\in\mathbb{N}}$ is constant and the strong convergence in (5.39) implies that $P_{\mathrm{div}}^n \boldsymbol{u}_0(\cdot) = \widetilde{\boldsymbol{U}}^{k,\lambda}(0, \cdot) = \widetilde{\boldsymbol{U}}^\lambda(0, \cdot) = \boldsymbol{U}_0^\lambda$ in Ω, so (5.67) and

(5.69) are satisfied.

Step 2: Identification of the limiting equation
Let $\boldsymbol{W} \in \mathbb{V}^n_{\mathrm{div}}$ and let $\varphi \in C_0^\infty((0,T))$ be arbitrary but fixed. With the convergence of $\partial_t \widetilde{\boldsymbol{U}}^{k,\lambda}$ in (5.40) it follows that

$$\left\langle \partial_t \widetilde{\boldsymbol{U}}^{k,\lambda}, \varphi \boldsymbol{W} \right\rangle_Q \to \left\langle \partial_t \widetilde{\boldsymbol{U}}^{\lambda}, \varphi \boldsymbol{W} \right\rangle_Q, \qquad \text{as } k \to \infty. \tag{5.71}$$

Further, similarly as in Step 1 in the proof of Lemma 4.4 by the strong convergence in (5.41), (5.43) and the weak convergence in (5.45) it is straightforward to show that

$$\left\langle \mathfrak{L}^{k,\lambda}[\overline{\boldsymbol{U}}^{k,\lambda}; \boldsymbol{W}](\cdot), \varphi \right\rangle_{(0,T)} \to \left\langle \mathfrak{L}^{\lambda}[\overline{\boldsymbol{U}}^{\lambda}; \boldsymbol{W}](\cdot), \varphi \right\rangle_{(0,T)}, \qquad \text{as } k \to \infty. \tag{5.72}$$

In particular, the strong convergence in (5.41) and in (5.43) allows us to take the limit in the numerical convective term without any restriction. Finally, (5.71) and (5.72) applied in (5.37) imply that (5.66) holds for a.e. $t \in (0,T)$. Recall that $\overline{\mathbf{S}}^{\lambda}(t,\cdot) = \mathbf{S}_i^{\lambda}$, for any $t \in (t_{i-1}, t_i]$ and $i \in \{1,\dots,l\}$ by (5.46). Since now the terms in (5.66) are constant on each interval $(t_{i-1}, t_i]$, for $i \in \{1,\dots,l\}$, the equation also holds for all $t \in (0,T]$. Also, (5.70) follows from (5.66) since $\partial_t \widetilde{\boldsymbol{U}}^{\lambda}(t,\cdot) = \mathrm{d}_t \boldsymbol{U}_i^{\lambda}$, $\overline{\boldsymbol{U}}^{\lambda}(t,\cdot) = \boldsymbol{U}_i^{\lambda}$ and the corresponding holds for $\mathbf{S}_i^{\lambda}$ and $\boldsymbol{f}_i$, for any $t \in (t_{i-1}, t_i]$ and $i \in \{1,\dots,l\}$.

Step 3: Identification of the implicit relation
The proof of the implicit relation (5.68) relies on the strong convergence of $\{\mathbf{D}\overline{\boldsymbol{U}}^{k,\lambda}\}_{k\in\mathbb{N}}$ by (5.42), the weak convergence of $\{\boldsymbol{\mathcal{S}}^k(\cdot, \mathbf{D}\overline{\boldsymbol{U}}^{k,\lambda})\}_{k\in\mathbb{N}}$ by (5.44) and the property of $\boldsymbol{\mathcal{S}}^k$ stated in Assumption 3.18 (σ3) and follows as in Step 2 in the proof of Lemma 4.4; see also Remark 3.19.

□

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 and the condition $q > \frac{2d}{d+2}$ is required to gain compactness.

Two additional difficulties, compared to the existence proof in [BGMS12], arise from the discretisation. 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 has been proven in Subsection 2.2.3.2 but under stronger assumptions on the finite element setting. To avoid this, instead of the Aubin–Lions lemma we shall employ Simon's compactness lemma, see Lemma 2.8, which requires convergence properties of time-increments and can be applied since we still have admissibility.

The second difficulty is that, due to the time-stepping we have to identify the limits of the respective interpolants with each other both for the approximate velocity function and for the approximate stress tensor function. Furthermore, 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 (5.66) and $\mathbf{S}^{\lambda}$ satisfies the implicit relation in (5.68). Note that despite the regularisation and the fact that we still have full admissibility of the approximate solutions in the convective term, in the time derivative some integrability in time is missing. Hence, we obtain an energy identity only for almost all times and – in contrast with the steady case – we require a localised version of the Minty type convergence lemma to identify the implicit relation already at this stage.

For the following let us denote

$$\eta \coloneqq \max\left(2q', \frac{q(d+2)}{d}\right) > 2, \tag{5.73}$$

since $q > \frac{2d}{d+2}$.

Lemma 5.6 (Convergence $l, n \to \infty$)**.**
Let $\lambda \coloneqq (l,n,m) \in \mathbb{N}^3$ and let $\overline{\boldsymbol{U}}^\lambda \in \mathbb{P}_0^l(0,T;\mathbb{V}_{\mathrm{div}}^n)$, $\widetilde{\boldsymbol{U}}^\lambda \in \mathbb{P}_1^l(0,T;\mathbb{V}_{\mathrm{div}}^n)$, $\mathbf{S}^\lambda \in \mathrm{L}^{q'}(Q)^{d\times d}$ and $\overline{\mathbf{S}}^\lambda \in \mathbb{P}_0^l(0,T;\mathrm{L}^{q'}(\Omega)^{d\times d})$ be solutions to (5.66)–(5.68).
Then, for any $0 \le s_0 < s \le T$ and all $\lambda = (l,n,m) \in \mathbb{N}^3$ one has that

$$\begin{aligned}\frac{1}{2}\left\|\widetilde{\boldsymbol{U}}^\lambda(s,\cdot)\right\|_{\mathrm{L}^2(\Omega)}^2 + \left\langle \overline{\mathbf{S}}^\lambda, \mathbf{D}\overline{\boldsymbol{U}}^\lambda\right\rangle_{Q_{s_0}^s} + \frac{1}{m}\left\|\overline{\boldsymbol{U}}^\lambda\right\|_{\mathrm{L}^{2q'}(Q_{s_0}^s)}^{2q'} \\ \le \left\langle \overline{\boldsymbol{f}}, \overline{\boldsymbol{U}}^\lambda\right\rangle_{Q_{s_0}^s} + \frac{1}{2}\left\|\widetilde{\boldsymbol{U}}^\lambda(s_0,\cdot)\right\|_{\mathrm{L}^2(\Omega)}^2.\end{aligned} \tag{5.74}$$

Further, with η as defined in (5.73), for each $m \in \mathbb{N}$ there exists a $\boldsymbol{u}^m \in \mathrm{L}^\infty(0,T;\mathrm{L}^2_{\mathrm{div}}(\Omega)^d) \cap \mathrm{X}_{\mathrm{div}}(Q)$, a function $\mathbf{S}^m \in \mathrm{L}^{q'}(Q)^{d\times d}$ and subsequences such that

$$\widetilde{\boldsymbol{U}}^{l,n,m} \to \boldsymbol{u}^m \quad \text{strongly in } \mathrm{L}^p(0,T;\mathrm{L}^2(\Omega)^d) \text{ for all } p\in[1,\infty), \tag{5.75}$$

$$\widetilde{\boldsymbol{U}}^{l,n,m}(s,\cdot) \to \boldsymbol{u}^m(s,\cdot) \quad \text{strongly in } \mathrm{L}^2(\Omega)^d \text{ for a.e. } s\in(0,T), \tag{5.76}$$

$$\widetilde{\boldsymbol{U}}^{l,n,m}(0,\cdot) \to \boldsymbol{u}_0 \quad \text{strongly in } \mathrm{L}^2(\Omega)^d, \tag{5.77}$$

$$\widetilde{\boldsymbol{U}}^{l,n,m}, \overline{\boldsymbol{U}}^{l,n,m} \overset{*}{\rightharpoonup} \boldsymbol{u}^m \quad \text{weakly* in } \mathrm{L}^\infty(0,T;\mathrm{L}^2(\Omega)^d), \tag{5.78}$$

$$\overline{\boldsymbol{U}}^{l,n,m} \to \boldsymbol{u}^m \quad \text{strongly in } \mathrm{L}^p(0,T;\mathrm{L}^2(\Omega)^d) \cap \mathrm{L}^r(Q)^d \text{ for all } p\in[1,\infty) \text{ and all } r\in[1,\eta), \tag{5.79}$$

$$\overline{\boldsymbol{U}}^{l,n,m}(s,\cdot) \to \boldsymbol{u}^m(s,\cdot) \quad \text{strongly in } \mathrm{L}^2(\Omega)^d \text{ for a.e. } s\in(0,T), \tag{5.80}$$

$$\overline{\boldsymbol{U}}^{l,n,m} \rightharpoonup \boldsymbol{u}^m \quad \text{weakly in } \mathrm{L}^q(0,T;\mathrm{W}_0^{1,q}(\Omega)^d) \cap \mathrm{L}^\eta(Q)^d, \tag{5.81}$$

$$\left|\overline{\boldsymbol{U}}^{l,n,m}\right|^{2q'-2}\overline{\boldsymbol{U}}^{l,n,m} \rightharpoonup |\boldsymbol{u}^m|^{2q'-2}\boldsymbol{u}^m \quad \text{weakly in } \mathrm{L}^{(2q')'}(Q)^d, \tag{5.82}$$

$$\overline{\mathbf{S}}^{l,n,m} \rightharpoonup \mathbf{S}^m \quad \text{weakly in } \mathrm{L}^{q'}(Q)^{d\times d}, \tag{5.83}$$

$$\mathbf{S}^{l,n,m} \rightharpoonup \mathbf{S}^m \quad \text{weakly in } \mathrm{L}^{q'}(Q)^{d\times d}, \tag{5.84}$$

as $l, n \to \infty$.

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 (5.66) we test with $\boldsymbol{W} = \overline{\boldsymbol{U}}^\lambda(t,\cdot) \in \mathbb{V}_{\mathrm{div}}^n$. For the first term adding and subtracting $\widetilde{\boldsymbol{U}}^\lambda(t,\cdot)$ one obtains with (2.74) that

$$\begin{aligned}\left\langle \partial_t\widetilde{\boldsymbol{U}}^\lambda(t,\cdot), \overline{\boldsymbol{U}}^\lambda(t,\cdot)\right\rangle_\Omega &= \left\langle \partial_t\widetilde{\boldsymbol{U}}^\lambda(t,\cdot), \widetilde{\boldsymbol{U}}^\lambda(t,\cdot)\right\rangle_\Omega + \left\langle \partial_t\widetilde{\boldsymbol{U}}^\lambda(t,\cdot), \overline{\boldsymbol{U}}^\lambda(t,\cdot) - \widetilde{\boldsymbol{U}}^\lambda(t,\cdot)\right\rangle_\Omega \\ &= \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left\|\widetilde{\boldsymbol{U}}^\lambda(t,\cdot)\right\|_{\mathrm{L}^2(\Omega)}^2 + (t_i - t)\left\|\partial_t\widetilde{\boldsymbol{U}}^\lambda(t,\cdot)\right\|_{\mathrm{L}^2(\Omega)}^2\end{aligned} \tag{5.85}$$

$$\geq \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left\|\widetilde{\boldsymbol{U}}^{\lambda}(t,\cdot)\right\|_{\mathrm{L}^2(\Omega)}^2,$$

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

$$\int_{s_0}^{s}\left\langle \partial_t \widetilde{\boldsymbol{U}}^{\lambda}(t,\cdot), \overline{\boldsymbol{U}}^{\lambda}(t,\cdot)\right\rangle_{\Omega} \geq \frac{1}{2}\left\|\widetilde{\boldsymbol{U}}^{\lambda}(s,\cdot)\right\|_{\mathrm{L}^2(\Omega)}^2 - \frac{1}{2}\left\|\widetilde{\boldsymbol{U}}^{\lambda}(s_0,\cdot)\right\|_{\mathrm{L}^2(\Omega)}^2. \tag{5.86}$$

The other terms follow immediately and (5.74) is proved.

Step 2: Estimates on the discrete level

Similarly as in the proof of Lemma 5.3, testing (5.70) with $\boldsymbol{W} = \boldsymbol{U}_i^{\lambda} \in \mathbb{V}_{\mathrm{div}}^n$ for $i \in \{1, \ldots, l\}$ yields

$$\left\langle \mathrm{d}_t \boldsymbol{U}_i^{\lambda}, \boldsymbol{U}_i^{\lambda}\right\rangle_{\Omega} + \left\langle \mathbf{S}_i^{\lambda}, \mathbf{D}\boldsymbol{U}_i^{\lambda}\right\rangle_{\Omega} + \frac{1}{m}\left\|\boldsymbol{U}_i^{\lambda}\right\|_{\mathrm{L}^{2q'}(\Omega)}^{2q'} = \left\langle \boldsymbol{f}_i, \boldsymbol{U}_i^{\lambda}\right\rangle_{\Omega}. \tag{5.87}$$

The first term on the left-hand side is bounded as in (5.29) and the term on the right-hand side is bounded as in (5.31). The only difference arises in bounding the term involving the stress tensor, see (5.30): we use that $(\mathbf{D}\overline{\boldsymbol{U}}^{\lambda}(\boldsymbol{z}), \mathbf{S}^{\lambda}(\boldsymbol{z})) \in \mathcal{A}(\boldsymbol{z})$ for a.e. $\boldsymbol{z} \in Q$ by (5.68) and Assumption 3.11 (A3) to obtain

$$\begin{aligned}
\left\langle \mathbf{S}_i^{\lambda}, \mathbf{D}\boldsymbol{U}_i^{\lambda}\right\rangle_{\Omega} &\overset{(5.46)}{=} \left\langle \fint_{t_{i-1}}^{t_i} \mathbf{S}^{\lambda}(t,\cdot)\,\mathrm{d}t, \mathbf{D}\boldsymbol{U}_i^{\lambda}\right\rangle_{\Omega} = \frac{1}{\delta_l}\left\langle \mathbf{S}^{\lambda}, \mathbf{D}\boldsymbol{U}_i^{\lambda}\right\rangle_{Q_{i-1}^i} \\
&\geq \frac{1}{\delta_l}\int_{Q_{i-1}^i}\left(-g(\cdot) + c_*\left(\left|\mathbf{D}\boldsymbol{U}_i^{\lambda}\right|^q + \left|\mathbf{S}^{\lambda}\right|^{q'}\right)\right)\mathrm{d}\boldsymbol{z} \\
&\geq -\frac{1}{\delta_l}\|g\|_{\mathrm{L}^1(Q_{i-1}^i)} + c\left\|\boldsymbol{U}_i^{\lambda}\right\|_{\mathrm{W}^{1,q}(\Omega)}^q + \frac{c_*}{\delta_l}\left\|\mathbf{S}^{\lambda}\right\|_{\mathrm{L}^{q'}(Q_{i-1}^i)}^{q'},
\end{aligned} \tag{5.88}$$

where again Poincaré's and Korn's inequalities were used. Following the same procedure as in (5.32)–(5.34) we arrive at

$$\begin{aligned}
\max_{j\in\{1,\ldots,l\}}\left\|\boldsymbol{U}_j^{\lambda}\right\|_{\mathrm{L}^2(\Omega)}^2 + \sum_{j=1}^{l}\left\|\boldsymbol{U}_j^{\lambda} - \boldsymbol{U}_{j-1}^{\lambda}\right\|_{\mathrm{L}^2(\Omega)}^2 + \delta_l\sum_{j=1}^{l}\left\|\boldsymbol{U}_j^{\lambda}\right\|_{\mathrm{W}^{1,q}(\Omega)}^q \\
+ \left\|\mathbf{S}^{\lambda}\right\|_{\mathrm{L}^{q'}(Q)}^{q'} + \frac{\delta_l}{m}\sum_{j=1}^{l}\left\|\boldsymbol{U}_j^{\lambda}\right\|_{\mathrm{L}^{2q'}(\Omega)}^{2q'} \leq c
\end{aligned} \tag{5.89}$$

for all $\lambda = (l, n, m) \in \mathbb{N}^3$. It follows by the relation between $\overline{\mathbf{S}}^{\lambda}$ and $\mathbf{S}^{\lambda}$ in (5.46), by (2.78) and the estimate (5.89) that

$$\left\|\overline{\mathbf{S}}^{\lambda}\right\|_{\mathrm{L}^{q'}(Q)}^{q'} \leq \left\|\mathbf{S}^{\lambda}\right\|_{\mathrm{L}^{q'}(Q)}^{q'} \leq c \quad \text{for all } \lambda \in \mathbb{N}^3. \tag{5.90}$$

Step 3: Estimates on the continuous level

By the definition of the piecewise constant interpolant according to (2.72) it follows from the discrete estimates that

$$\begin{aligned}
&\left\|\overline{\boldsymbol{U}}^{\lambda}\right\|_{\mathrm{L}^{\infty}(0,T;\mathrm{L}^2(\Omega))} + \left\|\overline{\boldsymbol{U}}^{\lambda}\right\|_{\mathrm{L}^{2q'}(Q)}^{2q'} + \left\|\overline{\boldsymbol{U}}^{\lambda}\right\|_{\mathrm{L}^{q}(0,T;\mathrm{W}^{1,q}(\Omega))}^{q} \\
&\overset{(2.75)}{=} \max_{j\in\{1,\dots,l\}} \left\|\boldsymbol{U}_j^{\lambda}\right\|_{\mathrm{L}^2(\Omega)} + \delta_l \sum_{j=1}^{l} \left\|\boldsymbol{U}_j^{\lambda}\right\|_{\mathrm{L}^{2q'}(\Omega)}^{2q'} + \delta_l \sum_{j=1}^{l} \left\|\boldsymbol{U}_j^{\lambda}\right\|_{\mathrm{W}^{1,q}(\Omega)}^{q} \\
&\overset{(5.89)}{\leq} c(m),
\end{aligned} \tag{5.91}$$

for all $\lambda = (l,n,m) \in \mathbb{N}^3$. With this and the parabolic interpolation from Corollary 2.5 we also have that

$$\left\|\overline{\boldsymbol{U}}^{\lambda}\right\|_{\mathrm{L}^{\frac{q(d+2)}{d}}(Q)} \leq c(m) \quad \text{ for all } \lambda = (l,n,m) \in \mathbb{N}^3. \tag{5.92}$$

For the estimates of the continuous, piecewise affine interpolant $\widetilde{\boldsymbol{U}}^{\lambda}$ according to (2.73) one also has to control the corresponding norms of $\boldsymbol{U}_0^{\lambda}$, which is why we can only show bounds of $\{\widetilde{\boldsymbol{U}}^{\lambda}\}_{\lambda\in\mathbb{N}^3}$ in $\mathrm{L}^{\infty}(0,T;\mathrm{L}^2(\Omega)^d)$ unless one wants to assume more regularity on the initial data; as in the argument following (5.48) we obtain by (5.69), with stability of the L^2-projector in (2.81), and with the discrete estimate in (5.89) that

$$\left\|\widetilde{\boldsymbol{U}}^{\lambda}\right\|_{\mathrm{L}^{\infty}(0,T;\mathrm{L}^2(\Omega))} \overset{(2.76)}{=} \max_{j\in\{0,\dots,l\}} \left\|\boldsymbol{U}_j^{\lambda}\right\|_{\mathrm{L}^2(\Omega)} \overset{(5.69),(5.89)}{\leq} c \quad \text{ for all } \lambda \in \mathbb{N}^3. \tag{5.93}$$

For the compactness argument instead of $\widetilde{\boldsymbol{U}}^{\lambda}$ we consider $\widehat{\boldsymbol{U}}^{\lambda} \in \mathrm{C}([0,T];\mathbb{V}^n_{\mathrm{div}})$ defined by

$$\widehat{\boldsymbol{U}}^{\lambda}(t,\cdot) := \begin{cases} \widetilde{\boldsymbol{U}}^{\lambda}(t,\cdot) & \text{if } t \in (\delta_l, T], \\ \overline{\boldsymbol{U}}^{\lambda}(t,\cdot) = \boldsymbol{U}_1^{\lambda}(\cdot) & \text{if } t \in [0,\delta_l], \end{cases} \tag{5.94}$$

which is constant on $[0,\delta_l]$. By the definition of $\widehat{\boldsymbol{U}}^{\lambda}$ and the relations (2.75), (2.76) , for $r \in [1,\infty)$ and a normed space X, we have that

$$\begin{aligned}
\left\|\widehat{\boldsymbol{U}}^{\lambda}\right\|_{\mathrm{L}^r(0,T;\mathrm{X})}^{r} &= \left\|\overline{\boldsymbol{U}}^{\lambda}\right\|_{\mathrm{L}^r(0,\delta_l;\mathrm{X})}^{r} + \left\|\widetilde{\boldsymbol{U}}^{\lambda}\right\|_{\mathrm{L}^r(\delta_l,T;\mathrm{X})}^{r} \leq \delta_l \left\|\boldsymbol{U}_1^{\lambda}\right\|_{\mathrm{X}}^{r} + \delta_l \sum_{i=1}^{l} \left\|\boldsymbol{U}_i^{\lambda}\right\|_{\mathrm{X}}^{r} \\
&\leq c\delta_l \sum_{i=1}^{l} \left\|\boldsymbol{U}_i^{\lambda}\right\|_{\mathrm{X}}^{r} = c \left\|\overline{\boldsymbol{U}}^{\lambda}\right\|_{\mathrm{L}^r(0,T;\mathrm{X})}^{r} \quad \text{ for any } \lambda \in \mathbb{N}^3,
\end{aligned} \tag{5.95}$$

and an analogous estimate holds for $r = \infty$ with the obvious modifications. Hence, by the estimate (5.91) we have that

$$\left\|\widehat{\boldsymbol{U}}^{\lambda}\right\|_{\mathrm{L}^{\infty}(0,T;\mathrm{L}^2(\Omega))} + \left\|\widehat{\boldsymbol{U}}^{\lambda}\right\|_{\mathrm{L}^{2q'}(Q)}^{2q'} + \left\|\widehat{\boldsymbol{U}}^{\lambda}\right\|_{\mathrm{L}^{q}(0,T;\mathrm{W}^{1,q}(\Omega))}^{q} \leq c(m) \tag{5.96}$$

for all $\lambda = (l,n,m) \in \mathbb{N}^3$.

By the fact that $\partial_t \widetilde{\boldsymbol{U}}^{\lambda}(t,\cdot) = \mathrm{d}_t \boldsymbol{U}_i^{\lambda}$, for $t \in (t_{i-1}, t_i]$, $i \in \{1,\dots,l\}$, and by the discrete estimate (5.89) we obtain

$$\delta_l \left\| \partial_t \widetilde{\boldsymbol{U}}^{\lambda} \right\|^2_{\mathrm{L}^2(Q)} = \delta_l \sum_{i=1}^{l} \left\| \frac{1}{\delta_l} (\boldsymbol{U}_i^{\lambda} - \boldsymbol{U}_{i-1}^{\lambda}) \right\|^2_{\mathrm{L}^2(Q_{i-1}^i)} = \sum_{i=1}^{l} \left\| \boldsymbol{U}_i^{\lambda} - \boldsymbol{U}_{i-1}^{\lambda} \right\|^2_{\mathrm{L}^2(\Omega)} \overset{(5.89)}{\leq} c \tag{5.97}$$

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

Finally, we also estimate $\mathfrak{L}^{\lambda}[\boldsymbol{u}; \boldsymbol{v}](t)$, as defined in (5.64): by the bound (5.18) on the numerical convective term, duality of norms, and Hölder's and Poincaré's inequalities we obtain

$$\begin{aligned}
\int_a^b \mathfrak{L}^{\lambda}[\boldsymbol{u}; \boldsymbol{v}](t)\,\mathrm{d}t &= \left\langle \widetilde{b}(\boldsymbol{u}(t,\cdot), \boldsymbol{u}(t,\cdot), \boldsymbol{v}) \right\rangle_{(a,b)} - \left\langle \overline{\mathbf{S}}^{\lambda}(t,\cdot), \mathbf{D}\boldsymbol{v} \right\rangle_{Q_a^b} \\
&\quad - \frac{1}{m} \left\langle |\boldsymbol{u}(t,\cdot)|^{2q'-2}\, \boldsymbol{u}(t,\cdot), \boldsymbol{v} \right\rangle_{Q_a^b} + \left\langle \overline{\boldsymbol{f}}(t,\cdot), \boldsymbol{v} \right\rangle_{Q_a^b} \\
&\overset{(5.18)}{\leq} \|\boldsymbol{u}\|^2_{\mathrm{L}^{2q'}(Q_a^b)} \|\nabla \boldsymbol{v}\|_{\mathrm{L}^q(Q_a^b)} + \|\boldsymbol{u}\|_{\mathrm{L}^{2q'}(Q_a^b)} \|\nabla \boldsymbol{u}\|_{\mathrm{L}^q(Q_a^b)} \|\boldsymbol{v}\|_{\mathrm{L}^{2q'}(Q_a^b)} \\
&\quad + \left\| \overline{\mathbf{S}}^{\lambda} \right\|_{\mathrm{L}^{q'}(Q_a^b)} \|\mathbf{D}\boldsymbol{v}\|_{\mathrm{L}^q(Q_a^b)} + \frac{1}{m} \|\boldsymbol{u}\|^{2q'-1}_{\mathrm{L}^{2q'}(Q_a^b)} \|\boldsymbol{v}\|_{\mathrm{L}^{2q'}(Q_a^b)} \\
&\quad + \left\| \overline{\boldsymbol{f}} \right\|_{\mathrm{L}^{q'}(a,b;\mathrm{W}^{-1,q'}(\Omega))} \|\boldsymbol{v}\|_{\mathrm{L}^q(a,b;\mathrm{W}^{1,q}(\Omega))} \\
&\leq c \left(1 + \|\boldsymbol{u}\|^2_{\mathrm{L}^{2q'}(Q_a^b)} \right) \|\nabla \boldsymbol{v}\|_{\mathrm{L}^q(Q_a^b)} \\
&\quad + c \left(\|\boldsymbol{u}\|_{\mathrm{L}^{2q'}(Q_a^b)} \|\nabla \boldsymbol{u}\|_{\mathrm{L}^q(Q_a^b)} + \frac{1}{m} \|\boldsymbol{u}\|^{2q'-1}_{\mathrm{L}^{2q'}(Q_a^b)} \right) \|\boldsymbol{v}\|_{\mathrm{L}^{2q'}(Q_a^b)},
\end{aligned} \tag{5.98}$$

for $0 \leq a < b \leq T$, for any $\lambda = (l,n,m) \in \mathbb{N}^3$, where we have used the uniform estimate (5.90) on $\overline{\mathbf{S}}^{\lambda}$ and (5.21) on $\overline{\boldsymbol{f}}$. With the estimates on $\overline{\boldsymbol{U}}^{\lambda}$ in (5.91) this yields

$$\begin{aligned}
\int_a^b \mathfrak{L}^{\lambda}[\overline{\boldsymbol{U}}^{\lambda}; \boldsymbol{v}](t)\,\mathrm{d}t &\overset{(5.98)}{\leq} c \left(\left\| \overline{\boldsymbol{U}}^{\lambda} \right\|_{\mathrm{L}^{2q'}(Q_a^b)} \left\| \nabla \overline{\boldsymbol{U}}^{\lambda} \right\|_{\mathrm{L}^q(Q_a^b)} + \frac{1}{m} \left\| \overline{\boldsymbol{U}}^{\lambda} \right\|^{2q'-1}_{\mathrm{L}^{2q'}(Q_a^b)} \right) \|\boldsymbol{v}\|_{\mathrm{L}^{2q'}(Q_a^b)} \\
&\quad + c \left(1 + \left\| \overline{\boldsymbol{U}}^{\lambda} \right\|^2_{\mathrm{L}^{2q'}(Q_a^b)} \right) \|\nabla \boldsymbol{v}\|_{\mathrm{L}^q(Q_a^b)} \\
&\overset{(5.91)}{\leq} c(m) \left(\|\nabla \boldsymbol{v}\|_{\mathrm{L}^q(Q_a^b)} + \|\boldsymbol{v}\|_{\mathrm{L}^{2q'}(Q_a^b)} \right)
\end{aligned} \tag{5.99}$$

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

Step 4: Convergence of the time increments (compare [CHP10, pp. 174])
Instead of applying the Aubin–Lions lemma, as in [BGMS12], here we apply the compactness result due to Simon, stated in Lemma 2.8. We wish to apply it to the sequence $\{\widehat{\boldsymbol{U}}^{l,n,m}\}_{l,n \in \mathbb{N}}$, for fixed $m \in \mathbb{N}$, with $\mathrm{X} = \mathrm{W}^{1,q}(\Omega)^d$, $\mathrm{B} = \mathrm{L}^2(\Omega)^d$ and $p = 2$. Let us show that

$$\int_0^{T-\varepsilon} \left\| \widehat{\boldsymbol{U}}^{\lambda}(s+\varepsilon,\cdot) - \widehat{\boldsymbol{U}}^{\lambda}(s,\cdot) \right\|^2_{\mathrm{L}^2(\Omega)} \mathrm{d}s \to 0, \quad \text{as } \varepsilon \to 0, \quad \text{uniformly for } l,n \in \mathbb{N}. \tag{5.100}$$

Consider the term $\left\langle \widehat{\boldsymbol{U}}^{\lambda}(s+\varepsilon,\cdot) - \widehat{\boldsymbol{U}}^{\lambda}(s,\cdot), \boldsymbol{W} \right\rangle_{\Omega}$, 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 $\widehat{\boldsymbol{U}}^{\lambda}(s+\varepsilon) = \widehat{\boldsymbol{U}}^{\lambda}(s) = \boldsymbol{U}_1^{\lambda}$, so the term vanishes. Now let $s+\varepsilon > \delta_l$. By the definition of $\widehat{\boldsymbol{U}}^{\lambda}$ in (5.94) 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

$$\begin{aligned}\left\langle \widehat{\boldsymbol{U}}^{\lambda}(s+\varepsilon,\cdot) - \widehat{\boldsymbol{U}}^{\lambda}(s,\cdot), \boldsymbol{W} \right\rangle_{\Omega} &= \int_{\max(s,\delta_l)}^{s+\varepsilon} \left\langle \partial_t \widehat{\boldsymbol{U}}^{\lambda}(t,\cdot), \boldsymbol{W} \right\rangle_{\Omega} \mathrm{d}t \\ &= \int_{\max(s,\delta_l)}^{s+\varepsilon} \left\langle \partial_t \widetilde{\boldsymbol{U}}^{\lambda}(t,\cdot), \boldsymbol{W} \right\rangle_{\Omega} \mathrm{d}t,\end{aligned} \tag{5.101}$$

where in the last line we have used that $\widehat{\boldsymbol{U}}^{\lambda}(t,\cdot)$ and $\widetilde{\boldsymbol{U}}^{\lambda}(t,\cdot)$ coincide on $(\max(s,\delta_l), s+\varepsilon) \subset [\delta_l, T]$ by (5.94). Applying the equation (5.66) for a.e. $t \in (\max(s,\delta_l), s+\varepsilon)$, integrating and applying the bounds in (5.99) yields

$$\begin{aligned}\int_{\max(s,\delta_l)}^{s+\varepsilon} \left\langle \partial_t \widetilde{\boldsymbol{U}}^{\lambda}(t,\cdot), \boldsymbol{W} \right\rangle_{\Omega} \mathrm{d}t &\overset{(5.66)}{=} \int_{\max(s,\delta_l)}^{s+\varepsilon} \mathfrak{L}^{\lambda}[\overline{\boldsymbol{U}}^{\lambda}; \boldsymbol{W}](t)\, \mathrm{d}t \\ &\overset{(5.99)}{\leq} c(m) \left(\|\nabla \boldsymbol{W}\|_{\mathrm{L}^q\left(Q^{s+\varepsilon}_{\max(s,\delta_l)}\right)} + \|\boldsymbol{W}\|_{\mathrm{L}^{2q'}\left(Q^{s+\varepsilon}_{\max(s,\delta_l)}\right)} \right) \\ &= c(m) \left(\varepsilon^{1/q} + \varepsilon^{1/2q'} \right) \|\boldsymbol{W}\|_{\mathrm{X}(\Omega)},\end{aligned} \tag{5.102}$$

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 $\widehat{\boldsymbol{U}}^{\lambda}(s+\varepsilon,\cdot)$, $\widehat{\boldsymbol{U}}^{\lambda}(s,\cdot) \in \mathbb{V}^n_{\mathrm{div}}$; so, applying (5.101) and (5.102) with $\boldsymbol{W} = \widehat{\boldsymbol{U}}^{\lambda}(s+\varepsilon,\cdot) - \widehat{\boldsymbol{U}}^{\lambda}(s,\cdot)$, which is piecewise constant in time, shows that

$$\left\| \widehat{\boldsymbol{U}}^{\lambda}(s+\varepsilon,\cdot) - \widehat{\boldsymbol{U}}^{\lambda}(s,\cdot) \right\|^2_{\mathrm{L}^2(\Omega)} \leq c(m) \left(\varepsilon^{1/q} + \varepsilon^{1/2q'} \right) \left\| \widehat{\boldsymbol{U}}^{\lambda}(s+\varepsilon,\cdot) - \widehat{\boldsymbol{U}}^{\lambda}(s,\cdot) \right\|_{\mathrm{X}(\Omega)}. \tag{5.103}$$

Integrating over $(0, T-\varepsilon)$, using the triangle inequality, Hölder's inequality and the estimate in (5.96) yields

$$\begin{aligned}&\int_0^{T-\varepsilon} \left\| \widehat{\boldsymbol{U}}^{\lambda}(s+\varepsilon,\cdot) - \widehat{\boldsymbol{U}}^{\lambda}(s,\cdot) \right\|^2_{\mathrm{L}^2(\Omega)} \mathrm{d}s \\ &\overset{(5.103)}{\leq} c(m) \left(\varepsilon^{1/q} + \varepsilon^{1/2q'} \right) \int_0^{T-\varepsilon} \left(\left\| \widehat{\boldsymbol{U}}^{\lambda}(s+\varepsilon,\cdot) \right\|_{\mathrm{X}(\Omega)} + \left\| \widehat{\boldsymbol{U}}^{\lambda}(s,\cdot) \right\|_{\mathrm{X}(\Omega)} \right) \mathrm{d}s \\ &\leq c(m) \left(\varepsilon^{1/q} + \varepsilon^{1/2q'} \right) \left\| \widehat{\boldsymbol{U}}^{\lambda} \right\|_{\mathrm{X}(Q)} \overset{(5.96)}{\leq} c(m)(\varepsilon^{1/q} + \varepsilon^{1/2q'}) \to 0,\end{aligned} \tag{5.104}$$

as $\varepsilon \to 0$ uniformly in $l, n \in \mathbb{N}$, where $\lambda = (l,n,m) \in \mathbb{N}^3$. This proves (5.100) .

Step 5: 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 the estimates (5.96) 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 (5.100) all the assumptions in Lemma 2.8 are satisfied for $\mathrm{X} = \mathrm{W}^{1,q}(\Omega)^d$, $\mathrm{B} = \mathrm{L}^2(\Omega)^d$ and $p = 2$. Hence, there exists a $\boldsymbol{u}^m \in \mathrm{L}^2(Q)^d$ and a subsequence such that

$$\widehat{\boldsymbol{U}}^{l,n,m} \to \boldsymbol{u}^m \quad \text{strongly in } \mathrm{L}^2(Q)^d, \quad \text{as } l, n \to \infty. \tag{5.105}$$

By the definition of $\widehat{\boldsymbol{U}}^{l,n,m}$ in (5.94) and the property (2.74) of the interpolants defined in

(2.72) and (2.73) we have that

$$\begin{aligned}\left\|\widehat{\boldsymbol{U}}^{l,n,m}-\widetilde{\boldsymbol{U}}^{l,n,m}\right\|_{\mathrm{L}^2(Q)}^2&=\left\|\overline{\boldsymbol{U}}^{l,n,m}-\widetilde{\boldsymbol{U}}^{l,n,m}\right\|_{\mathrm{L}^2(0,\delta_l;\mathrm{L}^2(\Omega))}^2\\&\overset{(2.74)}{\leq}\left\|(\delta_l-t)\partial_t\widetilde{\boldsymbol{U}}^{l,n,m}\right\|_{\mathrm{L}^2(0,\delta_l;\mathrm{L}^2(\Omega))}^2\leq\delta_l^2\left\|\partial_t\widetilde{\boldsymbol{U}}^{l,n,m}\right\|_{\mathrm{L}^2(0,\delta_l;\mathrm{L}^2(\Omega))}^2\\&\overset{(5.97)}{\leq}c\delta_l\to0,\qquad\text{as }l\to\infty.\end{aligned}\tag{5.106}$$

With (5.105) 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 (5.93) and interpolation, this implies that

$$\widetilde{\boldsymbol{U}}^{l,n,m}\to\boldsymbol{u}^m\quad\text{strongly in }\mathrm{L}^p(0,T;\mathrm{L}^2(\Omega)^d),\quad\text{as }l,n\to\infty,\tag{5.107}$$

for any $p\in[1,\infty)$. Similarly, by (2.74) we have that

$$\begin{aligned}\left\|\overline{\boldsymbol{U}}^{l,n,m}-\widetilde{\boldsymbol{U}}^{l,n,m}\right\|_{\mathrm{L}^2(Q)}^2&\overset{(2.74)}{=}\sum_{i=1}^{l}\left\|(t_i-t)\partial_t\widetilde{\boldsymbol{U}}^{l,n,m}\right\|_{\mathrm{L}^2(Q_{i-1}^i)}^2\\&\leq\delta_l^2\left\|\partial_t\widetilde{\boldsymbol{U}}^{l,n,m}\right\|_{\mathrm{L}^2(Q)}^2\overset{(5.97)}{\leq}c\delta_l\to0,\qquad\text{as }l\to\infty.\end{aligned}\tag{5.108}$$

Consequently, with (5.107) 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 (5.80). Analogously, (5.77) follows from the strong convergence of $\widetilde{\boldsymbol{U}}^{l,n,m}$ in (5.107).

The uniform bounds in $\mathrm{L}^\infty(0,T;\mathrm{L}^2(\Omega)^d)$ and $\mathrm{L}^\eta(Q)^d$, with $\eta=\max(2q',\frac{q(d+2)}{d})$, by (5.91) and (5.92), and the strong convergence in $\mathrm{L}^2(Q)^d$, yield by interpolation, that

$$\overline{\boldsymbol{U}}^{l,n,m}\to\boldsymbol{u}^m\quad\text{strongly in }\mathrm{L}^p(0,T;\mathrm{L}^2(\Omega)^d)\cap\mathrm{L}^r(Q)^d,\ \text{as }l,n\to\infty,\tag{5.109}$$

for any $p\in[1,\infty)$ and any $r\in[1,\eta)$. By the uniform bounds in (5.91) and (5.92) and the Banach–Alaoglu theorem, up to subsequences, we have that

$$\overline{\boldsymbol{U}}^{l,n,m}\overset{*}{\rightharpoonup}\boldsymbol{u}^m\quad\text{weakly* in }\mathrm{L}^\infty(0,T;\mathrm{L}^2(\Omega)^d),\tag{5.110}$$

$$\overline{\boldsymbol{U}}^{l,n,m}\rightharpoonup\boldsymbol{u}^m\quad\text{weakly in }\mathrm{L}^q(0,T;\mathrm{W}_0^{1,q}(\Omega)^d)\cap\mathrm{L}^\eta(Q)^d,\tag{5.111}$$

as $l,n\to\infty$, where the identification of the limiting functions with $\boldsymbol{u}^m$ follows by the strong convergence in (5.109).

The argument that $\boldsymbol{u}^m$ is divergence-free follow analogously to [DKS13a, p. 1001], see also Step 2 in the proof of Lemma 4.5: Let $h\in\mathrm{L}^{q'}(\Omega)$ and note that by the Assumption 2.21 on the projector $\Pi_{\mathbb{Q}}^n$ we have that $\Pi_{\mathbb{Q}}^n h\to h$ in particular in $\mathrm{L}^{q'}(\Omega)$, as $n\to\infty$, see (2.37). Also, let $\varphi\in\mathrm{C}_0^\infty(0,T)$. By (5.111) we have that $\operatorname{div}\overline{\boldsymbol{U}}^{l,n,m}\rightharpoonup\operatorname{div}\boldsymbol{u}^m$ weakly in $\mathrm{L}^q(Q)$, and hence

$$\left\langle\operatorname{div}\overline{\boldsymbol{U}}^{l,n,m},\varphi\Pi_{\mathbb{Q}}^n h\right\rangle_Q\to\langle\operatorname{div}\boldsymbol{u}^m,\varphi h\rangle_Q,\quad\text{as }l,n\to\infty.\tag{5.112}$$

Since $\overline{\boldsymbol{U}}^{l,n,m}\in\mathbb{P}_0^l(0,T;\mathbb{V}_{\operatorname{div}}^n)$ 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 $h \in \mathrm{L}^{q'}(\Omega)$ and all $\varphi \in \mathrm{C}^\infty(0,T)$, so by density $\boldsymbol{u}^m$ is (weakly) divergence-free.

By (5.91) 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 $\mathrm{L}^{(2q')'}(Q)^d$ and thus, by the Banach–Alaoglu theorem there exists a subsequence and $\psi^m \in \mathrm{L}^{(2q')'}(Q)^d$ such that

$$\left|\overline{\boldsymbol{U}}^{l,n,m}\right|^{2q'-2} \overline{\boldsymbol{U}}^{l,n,m} \rightharpoonup \psi^m \quad \text{weakly in } \mathrm{L}^{(2q')'}(Q)^d, \quad \text{as } l,n \to \infty. \tag{5.113}$$

By the strong convergence in (5.109), 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 (5.82).

Because $\widetilde{\boldsymbol{U}}^{l,n,m}(0,\cdot) = P^n_{\operatorname{div}}\boldsymbol{u}_0$ by (5.67), with (2.82) it follows that

$$\widetilde{\boldsymbol{U}}^{l,n,m}(0,\cdot) = P^n_{\operatorname{div}}\boldsymbol{u}_0 \to \boldsymbol{u}_0 \quad \text{strongly in } \mathrm{L}^2(\Omega)^d, \quad \text{as } n \to \infty, \tag{5.114}$$

so (5.77) is proven.

The uniform estimates in (5.90) and the Banach–Alaoglu theorem imply that there exist $\overline{\mathbf{S}}^m, \mathbf{S}^m \in \mathrm{L}^{q'}(Q)^{d\times d}$ and subsequences such that

$$\overline{\mathbf{S}}^{l,n,m} \rightharpoonup \overline{\mathbf{S}}^m \quad \text{weakly in } \mathrm{L}^{q'}(Q)^{d\times d}, \tag{5.115}$$
$$\mathbf{S}^{l,n,m} \rightharpoonup \mathbf{S}^m \quad \text{weakly in } \mathrm{L}^{q'}(Q)^{d\times d}, \tag{5.116}$$

as $l,n \to \infty$. It remains to show that $\overline{\mathbf{S}}^m = \mathbf{S}^m$; to this end, let $\mathbf{B} \in \mathrm{C}^\infty_0(Q)^{d\times d}$ be arbitrary but fixed. On the one hand the weak convergence in (5.115) shows that

$$\left\langle \overline{\mathbf{S}}^{l,n,m}, \mathbf{B} \right\rangle_Q \to \left\langle \overline{\mathbf{S}}^m, \mathbf{B} \right\rangle_Q, \quad \text{as } l,n \to \infty. \tag{5.117}$$

On the other hand, by the relation between $\overline{\mathbf{S}}^{l,n,m}$ and $\mathbf{S}^{l,n,m}$ according to (5.46), one can show that $\left\langle \overline{\mathbf{S}}^{l,n,m}, \mathbf{B} \right\rangle_Q = \left\langle \mathbf{S}^{l,n,m}, \overline{\mathbf{B}} \right\rangle_Q$. By (2.79) we have that $\overline{\mathbf{B}} \to \mathbf{B}$ strongly in $\mathrm{L}^q(Q)^{d\times d}$, as $l \to \infty$, so that with the convergence in (5.116) it follows that

$$\left\langle \overline{\mathbf{S}}^{l,n,m}, \mathbf{B} \right\rangle_Q = \left\langle \mathbf{S}^{l,n,m}, \overline{\mathbf{B}} \right\rangle_Q \to \left\langle \mathbf{S}^m, \mathbf{B} \right\rangle_Q, \quad \text{as } l,n \to \infty. \tag{5.118}$$

By (5.117) and (5.118), 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

$$\begin{aligned} \mathfrak{L}^m[\boldsymbol{u};\boldsymbol{v}](t) :=& -b(\boldsymbol{u}(t,\cdot),\boldsymbol{u}(t,\cdot),\boldsymbol{v}) - \langle \mathbf{S}^m(t,\cdot), \mathbf{D}\boldsymbol{v}\rangle_\Omega \\ & - \frac{1}{m}\left\langle |\boldsymbol{u}(t,\cdot)|^{2q'-2}\,\boldsymbol{u}(t,\cdot), \boldsymbol{v}\right\rangle_\Omega + \langle \boldsymbol{f}(t,\cdot), \boldsymbol{v}\rangle_\Omega, \end{aligned} \tag{5.119}$$

where $\mathbf{S}^m$ is given by Lemma 5.6 and $b(\cdot,\cdot,\cdot)$ is defined in (5.9). Let us denote

$$\tau := \min\left(q', (2q')'\right) > 1. \tag{5.120}$$

Lemma 5.7 (Identification of the PDE as $l,n \to \infty$)**.**
The limiting functions $\boldsymbol{u}^m \in \mathrm{L}^\infty(0,T;\mathrm{L}^2_{\operatorname{div}}(\Omega)^d) \cap \mathrm{X}_{\operatorname{div}}(Q)$ given in Lemma 5.6 satisfy that $\partial_t \boldsymbol{u}^m \in \mathrm{L}^\tau(0,T;(\mathrm{X}_{\operatorname{div}}(\Omega))')$, with τ defined in (5.120), and $\mathrm{X}_{\operatorname{div}}(\Omega)$ defined in (5.11). (Up to a representative) we have that $\boldsymbol{u}^m \in \mathrm{C}_w([0,T],\mathrm{L}^2_{\operatorname{div}}(\Omega)^d)$ 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 5.6 satisfy

$$\langle \partial_t \boldsymbol{u}^m(t,\cdot), \boldsymbol{w}\rangle_\Omega = \mathfrak{L}^m[\boldsymbol{u}^m; \boldsymbol{w}](t) \quad \text{for all } \boldsymbol{w} \in \mathrm{C}^\infty_{0,\mathrm{div}}(\Omega)^d \text{ for a.e. } t \in (0,T), \tag{5.121}$$

$$(\mathbf{D}\boldsymbol{u}^m(\boldsymbol{z}), \mathbf{S}^m(\boldsymbol{z})) \in \mathcal{A}(\boldsymbol{z}) \quad \text{for a.e. } \boldsymbol{z} \in Q, \tag{5.122}$$

$$\operatorname*{ess\,lim}_{t\to 0_+} \|\boldsymbol{u}^m(t,\cdot) - \boldsymbol{u}_0\|_{\mathrm{L}^2(\Omega)} = 0. \tag{5.123}$$

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 (5.66) by $\varphi \in \mathrm{C}^\infty_0((-T,T))$ and integrating over $(0,T)$ yields

$$\left\langle \partial_t \widetilde{\boldsymbol{U}}^\lambda, \boldsymbol{W}\varphi \right\rangle_Q \overset{(5.66)}{=} \left\langle \mathfrak{L}^\lambda[\overline{\boldsymbol{U}}^\lambda; \boldsymbol{W}], \varphi \right\rangle_{(0,T)} \quad \text{for any } \boldsymbol{W} \in \mathbb{V}^n_{\mathrm{div}}. \tag{5.124}$$

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

$$-\left\langle \widetilde{\boldsymbol{U}}^\lambda, \boldsymbol{W}\partial_t\varphi \right\rangle_Q = \left\langle \widetilde{\boldsymbol{U}}^\lambda(0,\cdot), \varphi(0)\boldsymbol{W} \right\rangle_\Omega + \left\langle \mathfrak{L}^\lambda[\overline{\boldsymbol{U}}^\lambda; \boldsymbol{W}], \varphi \right\rangle_{(0,T)} \tag{5.125}$$

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 2.24 (i) for $\boldsymbol{w} \in \mathrm{C}^\infty_{0,\mathrm{div}}(\Omega)^d$ we have that

$$\mathbb{V}^n_{\mathrm{div}} \ni \Pi^n \boldsymbol{w} \to \boldsymbol{w} \quad \text{strongly in } \mathrm{W}^{1,s}_0(\Omega)^d, \quad \text{as } n\to\infty, \quad \text{for any } s \in [1,\infty). \tag{5.126}$$

To deduce the limiting equation for $\boldsymbol{u}^m$ we consider (5.125) 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 (5.75), with (5.126) we obtain that

$$-\left\langle \widetilde{\boldsymbol{U}}^{l,n,m}, \Pi^n(\boldsymbol{w})\partial_t\varphi \right\rangle_Q \to -\left\langle \boldsymbol{u}^m, \boldsymbol{w}\partial_t\varphi\right\rangle_Q, \quad \text{as } l,n\to\infty. \tag{5.127}$$

Similarly, the strong convergence of $\widetilde{\boldsymbol{U}}^{l,n,m}(0,\cdot) \to \boldsymbol{u}_0$ in $\mathrm{L}^2(\Omega)^d$ in (5.77) yields that

$$\left\langle \widetilde{\boldsymbol{U}}^{l,n,m}(0,\cdot), \varphi(0)\Pi^n\boldsymbol{w} \right\rangle_\Omega \to \left\langle \boldsymbol{u}_0, \varphi(0)\boldsymbol{w}\right\rangle_\Omega, \quad \text{as } l,n\to\infty. \tag{5.128}$$

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 (5.79), 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 $p \in [1,\frac{\eta}{2})$. Such a $p>1$ exists, since $\eta = \max\left(2q', \frac{q(d+2)}{d}\right) > 2$. With (5.126) 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

$$\left\langle \overline{\boldsymbol{U}}^{l,n,m} \otimes \overline{\boldsymbol{U}}^{l,n,m}, \varphi\nabla\Pi^n\boldsymbol{w} \right\rangle_Q \to \left\langle \boldsymbol{u}^m \otimes \boldsymbol{u}^m, \varphi\nabla\boldsymbol{w}\right\rangle_Q, \quad \text{as } l,n\to\infty. \tag{5.129}$$

For the modification of the convective term note first that 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 (5.81). By (5.79) 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 $\mathrm{W}^{1,s}(\Omega) \hookrightarrow$

$L^\infty(\Omega)$ is continuous, and hence we have $\varphi\Pi^n\boldsymbol{w} \to \varphi\boldsymbol{w}$ strongly in $L^\infty(Q)^d$. Together, these yield that

$$\left\langle \overline{\boldsymbol{U}}^{l,n,m} \otimes \varphi\Pi^n\boldsymbol{w}, \nabla\overline{\boldsymbol{U}}^{l,n,m} \right\rangle_Q \to \langle \boldsymbol{u}^m \otimes \varphi\boldsymbol{w}, \nabla\boldsymbol{u}^m\rangle_Q\,, \quad \text{as } l,n\to\infty. \tag{5.130}$$

By (5.126) we have in particular that $\varphi\mathbf{D}\Pi^n\boldsymbol{w} \to \varphi\mathbf{D}\boldsymbol{w}$ strongly in $L^q(Q)^{d\times d}$ and by (5.83) that $\overline{\mathbf{S}}^{l,n,m} \rightharpoonup \mathbf{S}^m$ weakly in $L^{q'}(Q)^{d\times d}$. Thus, it follows that

$$\left\langle \overline{\mathbf{S}}^{l,n,m}, \varphi\mathbf{D}\Pi^n\boldsymbol{w} \right\rangle_Q \to \langle \mathbf{S}^m, \varphi\mathbf{D}\boldsymbol{w}\rangle_Q\,, \quad \text{as } l,n\to\infty. \tag{5.131}$$

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 $L^{(2q')'}(Q)^d$ by (5.82) and $\varphi\Pi^n\boldsymbol{w} \to \varphi\boldsymbol{w}$ in particular in $L^{2q'}(Q)^d$, we obtain

$$\frac{1}{m}\left\langle \left|\overline{\boldsymbol{U}}^{l,n,m}\right|^{2q'-2}\overline{\boldsymbol{U}}^{l,n,m}, \varphi\Pi^n\boldsymbol{w} \right\rangle_Q \to \frac{1}{m}\left\langle |\boldsymbol{u}^m|^{2q'-2}\,\boldsymbol{u}^m, \varphi\boldsymbol{w}\right\rangle_Q\,, \quad \text{as } l,n\to\infty. \tag{5.132}$$

Finally, with the strong convergence $\overline{\boldsymbol{f}} \to \boldsymbol{f}$ in $L^{q'}(0,T;W^{-1,q'}(\Omega)^d)$ by (5.22) and with (5.126) we have that

$$\left\langle \overline{\boldsymbol{f}}, \varphi\Pi^n\boldsymbol{w} \right\rangle_Q \to \langle \boldsymbol{f}, \varphi\boldsymbol{w}\rangle_Q\,, \quad \text{as } l,n\to\infty. \tag{5.133}$$

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 (5.64) and (5.119), respectively, the convergence results (5.129)–(5.133) yield that

$$\left\langle \mathfrak{L}^{l,n,m}[\overline{\boldsymbol{U}}^{l,n,m}; \Pi^n\boldsymbol{w}], \varphi \right\rangle_{(0,T)} \to \langle \mathfrak{L}^m[\boldsymbol{u}^m; \boldsymbol{w}], \varphi\rangle_{(0,T)}\,, \quad \text{as } l,n\to\infty. \tag{5.134}$$

Now, from (5.125), using (5.127), (5.128) and (5.134), as $l,n\to\infty$, we have that

$$-\langle \boldsymbol{u}^m, \boldsymbol{w}\partial_t\varphi\rangle_Q = \langle \boldsymbol{u}_0, \varphi(0)\boldsymbol{w}\rangle_\Omega + \langle \mathfrak{L}^m[\boldsymbol{u}^m; \boldsymbol{w}], \varphi\rangle_{(0,T)} \tag{5.135}$$

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

Step 2: Regularity of the time-derivative
The distributional derivative of $\boldsymbol{u}^m$ satisfies, by definition and using (5.135), that

$$\langle \partial_t\boldsymbol{u}^m, \boldsymbol{w}\varphi\rangle_Q = -\langle \boldsymbol{u}^m, \boldsymbol{w}\partial_t\varphi\rangle_Q \overset{(5.135)}{=} \langle \mathfrak{L}^m[\boldsymbol{u}^m\boldsymbol{w}], \varphi\rangle_{(0,T)} \tag{5.136}$$

for all $\boldsymbol{w} \in C^\infty_{0,\mathrm{div}}(\Omega)^d$ and all $\varphi \in C^\infty_0((0,T))$, since $\operatorname{supp}\varphi \subset (0,T)$. Using this equation we wish to show that $\partial_t\boldsymbol{u}^m \in L^\tau(0,T;(X_{\mathrm{div}}(\Omega))')$ (not uniformly in $m\in\mathbb{N}$), for τ as in (5.120) and $X_{\mathrm{div}}(\Omega)$ as in (5.11). For $\mathfrak{L}^m$ as defined in (5.119), using the fact that $\boldsymbol{u}^m \in L^{2q'}(Q)^d$ and $\mathbf{S}^m \in L^{q'}(Q)^{d\times d}$, similarly as in (5.98) with (5.18) we can estimate

$$\begin{aligned}
\left|\langle \mathfrak{L}^m[\boldsymbol{u}^m, \boldsymbol{w}], \varphi\rangle_{(0,T)}\right| &\le \|\boldsymbol{u}^m\|^2_{L^{2q'}(Q)}\,\|\varphi\nabla\boldsymbol{w}\|_{L^q(Q)} + \|\mathbf{S}^m\|_{L^{q'}(Q)}\,\|\varphi\mathbf{D}\boldsymbol{w}\|_{L^q(Q)}\\
&\quad + \tfrac{1}{m}\,\|\boldsymbol{u}^m\|^{2q'-1}_{L^{2q'}(Q)}\,\|\varphi\boldsymbol{w}\|_{L^{2q'}(Q)} \qquad (5.137)\\
&\quad + \|\boldsymbol{f}\|_{L^{q'}(0,T;W^{-1,q'}(\Omega))}\,\|\varphi\boldsymbol{w}\|_{L^q(0,T;W^{1,q}(\Omega))} \qquad (5.138)\\
&\le c(m)\left(\|\varphi\|_{L^q(0,T)} + \|\varphi\|_{L^{2q'}(0,T)}\right)\left(\|\boldsymbol{w}\|_{W^{1,q}(\Omega)} + \|\boldsymbol{w}\|_{L^{2q'}(\Omega)}\right)
\end{aligned}$$

$$\leq c(m)\,\|\varphi\|_{\mathrm{L}^{\tau'}(0,T)}\,\|\boldsymbol{w}\|_{\mathrm{X}(\Omega)}\,,$$

for all $\varphi \in \mathrm{C}_0^\infty((0,T))$ 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 (5.136) and by reflexivity of the function space. Consequently, $\langle \partial_t \boldsymbol{u}^m, \boldsymbol{w}\rangle_\Omega$ is integrable for $\boldsymbol{w} \in \mathrm{C}^\infty_{0,\mathrm{div}}(\Omega)^d$, and thus, we can rephrase (5.136) by the fundamental lemma of calculus of variations in the pointwise sense in time, so (5.121) is proved.

Step 3: Identification of the initial condition
Let us first show that $\boldsymbol{u}^m \in \mathrm{C}_w([0,T];\mathrm{L}^2_{\mathrm{div}}(\Omega)^d)$. The embedding $\mathrm{X}_{\mathrm{div}}(\Omega) \hookrightarrow \mathrm{L}^2_{\mathrm{div}}(\Omega)^d$ is dense and hence we have $(\mathrm{L}^2_{\mathrm{div}}(\Omega)^d)' \hookrightarrow (\mathrm{X}_{\mathrm{div}}(\Omega))'$. Together with the embedding $\mathrm{L}^2_{\mathrm{div}}(\Omega)^d \hookrightarrow (\mathrm{L}^2_{\mathrm{div}}(\Omega)^d)'$ by (5.19) this shows that the embedding $\mathrm{L}^2_{\mathrm{div}}(\Omega)^d \hookrightarrow (\mathrm{X}_{\mathrm{div}}(\Omega))'$ is continuous.

Consequently, we have that $\boldsymbol{u}^m \in \mathrm{L}^\infty(0,T;\mathrm{L}^2_{\mathrm{div}}(\Omega)^d) \hookrightarrow \mathrm{L}^1(0,T;(\mathrm{X}_{\mathrm{div}}(\Omega))')$. With this and the fact that in particular $\partial_t \boldsymbol{u}^m \in \mathrm{L}^1(0,T;(\mathrm{X}_{\mathrm{div}}(\Omega))')$, Lemma 2.9 implies that $\boldsymbol{u}^m \in \mathrm{C}_w([0,T];(\mathrm{X}_{\mathrm{div}}(\Omega))')$. Furthermore, by this and the fact that $\boldsymbol{u}^m \in \mathrm{L}^\infty(0,T;\mathrm{L}^2_{\mathrm{div}}(\Omega)^d)$, again with $\mathrm{L}^2_{\mathrm{div}}(\Omega)^d \hookrightarrow (\mathrm{X}_{\mathrm{div}}(\Omega))'$, Lemma 2.10 shows that $\boldsymbol{u}^m \in \mathrm{C}_w([0,T];\mathrm{L}^2_{\mathrm{div}}(\Omega)^d)$.

Next, we shall show that $\boldsymbol{u}^m(0,\cdot) = \boldsymbol{u}_0 \in \mathrm{L}^2_{\mathrm{div}}(\Omega)^d$. Let $\varphi \in \mathrm{C}_0^\infty((-T,T))$ be such that $\varphi(0) = 1$. Multiplying (5.121) by φ and integrating over $(0,T)$, yields that

$$\langle \partial_t \boldsymbol{u}^m, \boldsymbol{w}\varphi\rangle_Q = \langle \mathfrak{L}^m[\boldsymbol{u}^m;\boldsymbol{w}],\varphi\rangle_{(0,T)} \tag{5.139}$$

for all $\boldsymbol{w} \in \mathrm{C}^\infty_{0,\mathrm{div}}(\Omega)^d$. On the other hand, by integration by parts, the fundamental theorem of calculus and the fact that $\boldsymbol{u}^m \in \mathrm{C}_w([0,T];\mathrm{L}^2_{\mathrm{div}}(\Omega)^d)$ and also applying (5.135), we have

$$\begin{aligned}\langle \partial_t \boldsymbol{u}^m, \boldsymbol{w}\varphi\rangle_Q &= \langle \partial_t(\boldsymbol{u}^m\varphi), \boldsymbol{w}\rangle_Q - \langle \boldsymbol{u}^m, \boldsymbol{w}\partial_t\varphi\rangle_Q \\ &\overset{(5.135)}{=} -\langle \boldsymbol{u}^m(0,\cdot)\varphi(0), \boldsymbol{w}\rangle_\Omega + \langle \boldsymbol{u}_0, \varphi(0)\boldsymbol{w}\rangle_\Omega + \langle \mathfrak{L}^m[\boldsymbol{u}^m;\boldsymbol{w}],\varphi\rangle_{(0,T)}\end{aligned} \tag{5.140}$$

for all $\boldsymbol{w} \in \mathrm{C}^\infty_{0,\mathrm{div}}(\Omega)^d$. Comparing with (5.139) and noting that $\varphi(0) = 1$, we obtain

$$\langle \boldsymbol{u}^m(0,\cdot), \boldsymbol{w}\rangle_\Omega = \langle \boldsymbol{u}_0, \boldsymbol{w}\rangle_\Omega \qquad \text{for all } \boldsymbol{w} \in \mathrm{C}^\infty_{0,\mathrm{div}}(\Omega)^d. \tag{5.141}$$

Since $\boldsymbol{u}_0, \boldsymbol{u}^m(0,\cdot) \in \mathrm{L}^2_{\mathrm{div}}(\Omega)^d$ are divergence-free, this suffices to conclude that $\boldsymbol{u}_0 = \boldsymbol{u}^m(0,\cdot)$.

By (5.76) we have strong convergence $\widetilde{\boldsymbol{U}}^{l,n,m}(s,\cdot) \to \boldsymbol{u}^m(s,\cdot)$ in $\mathrm{L}^2(\Omega)^d$, as $l,n \to \infty$, for a.e. $s \in (0,T)$. Furthermore, by (5.77) we have $\widetilde{\boldsymbol{U}}^{l,n,m}(0,\cdot) \to \boldsymbol{u}_0$ strongly in $\mathrm{L}^2(\Omega)^d$, as $l,n \to \infty$, and consequently for a.e. $s \in (0,T)$ we obtain that

$$\|\boldsymbol{u}^m(s,\cdot) - \boldsymbol{u}_0\|^2_{\mathrm{L}^2(\Omega)} = \lim_{l,n\to\infty} \left\|\widetilde{\boldsymbol{U}}^{l,n,m}(s,\cdot) - \widetilde{\boldsymbol{U}}^{l,n,m}(0,\cdot)\right\|^2_{\mathrm{L}^2(\Omega)}. \tag{5.142}$$

By (5.74), for all $s \in (0,T)$ and $\lambda = (l,n,m) \in \mathbb{N}^3$ we have that

$$\left\|\widetilde{\boldsymbol{U}}^\lambda(s,\cdot)\right\|^2 - \left\|\widetilde{\boldsymbol{U}}^\lambda(0,\cdot)\right\|^2 \leq 2\left\langle \overline{\boldsymbol{f}}, \overline{\boldsymbol{U}}^\lambda\right\rangle_{Q_s} \qquad \text{for } \lambda = (l,n,m) \in \mathbb{N}^3, \tag{5.143}$$

since the other terms can be shown to be nonnegative. Indeed, the nonnegativity of the term $\langle \overline{\mathbf{S}}^\lambda, \mathbf{D}\overline{\boldsymbol{U}}^\lambda\rangle_{Q_s}$ can be shown as follows: By the relation (5.46) we have that

$$\left\langle \overline{\mathbf{S}}^\lambda, \mathbf{D}\overline{\boldsymbol{U}}^\lambda\right\rangle_{Q^i_{i-1}} = \delta_l \left\langle \fint_{t_{i-1}}^{t_i} \mathbf{S}^\lambda(t,\cdot)\,\mathrm{d}t, \mathbf{D}\boldsymbol{U}^\lambda_i\right\rangle_\Omega = \left\langle \mathbf{S}^\lambda, \mathbf{D}\overline{\boldsymbol{U}}^\lambda\right\rangle_{Q^i_{i-1}}, \tag{5.144}$$

for any $i \in \{1,\dots,l\}$. For $j \in \{1,\dots,l\}$ such that $s \in (t_{j-1},t_j]$, we can split the term and then apply (5.144) to the first term and again (5.46) to the second term to obtain

$$\begin{aligned}\left\langle \overline{\mathbf{S}}^\lambda, \mathbf{D}\overline{\boldsymbol{U}}^\lambda \right\rangle_{Q_s} &= \left\langle \overline{\mathbf{S}}^\lambda, \mathbf{D}\overline{\boldsymbol{U}}^\lambda \right\rangle_{Q_{j-1}} + \left\langle \overline{\mathbf{S}}^\lambda, \mathbf{D}\overline{\boldsymbol{U}}^\lambda \right\rangle_{Q^s_{t_{j-1}}} \\ &\overset{(5.144)}{=} \left\langle \mathbf{S}^\lambda, \mathbf{D}\overline{\boldsymbol{U}}^\lambda \right\rangle_{Q_{j-1}} + \frac{(s-t_{j-1})}{\delta_l} \left\langle \mathbf{S}^\lambda, \mathbf{D}\overline{\boldsymbol{U}}^\lambda \right\rangle_{Q^j_{j-1}} .\end{aligned} \tag{5.145}$$

By the fact that $(\mathbf{D}\overline{\boldsymbol{U}}^\lambda(\boldsymbol{z}), \mathbf{S}^\lambda(\boldsymbol{z})) \in \mathcal{A}(\boldsymbol{z})$ for a.e. $\boldsymbol{z} \in Q$ by (5.68), and that $\mathcal{A}(\cdot)$ is monotone and $(\mathbf{0},\mathbf{0}) \in \mathcal{A}(\cdot)$ a.e. in Q by Assumption 3.11 (A1) and (A2), it follows that all terms in (5.145) are nonnegative, which shows that $\left\langle \overline{\mathbf{S}}^\lambda, \mathbf{D}\overline{\boldsymbol{U}}^\lambda \right\rangle_{Q_s} \geq 0$.

Expanding the norm on the right-hand side in (5.142), adding and subtracting twice the term $\|\widetilde{\boldsymbol{U}}^{l,n,m}(0,\cdot)\|^2_{\mathrm{L}^2(\Omega)}$ and applying (5.143), for $\lambda = (l,n,m)$, we obtain that

$$\begin{aligned}\left\|\widetilde{\boldsymbol{U}}^\lambda(s,\cdot) - \widetilde{\boldsymbol{U}}^\lambda(0,\cdot)\right\|^2_{\mathrm{L}^2(\Omega)} &= \left\|\widetilde{\boldsymbol{U}}^\lambda(s,\cdot)\right\|^2_{\mathrm{L}^2(\Omega)} - 2\left\langle \widetilde{\boldsymbol{U}}^\lambda(s,\cdot), \widetilde{\boldsymbol{U}}^\lambda(0,\cdot)\right\rangle_\Omega + \left\|\widetilde{\boldsymbol{U}}^\lambda(0,\cdot)\right\|^2_{\mathrm{L}^2(\Omega)} \\ &= \left\|\widetilde{\boldsymbol{U}}^\lambda(s,\cdot)\right\|^2_{\mathrm{L}^2(\Omega)} - \left\|\widetilde{\boldsymbol{U}}^\lambda(0,\cdot)\right\|^2_{\mathrm{L}^2(\Omega)} + 2\left\langle \widetilde{\boldsymbol{U}}^\lambda(0,\cdot) - \widetilde{\boldsymbol{U}}^\lambda(s,\cdot), \widetilde{\boldsymbol{U}}^\lambda(0,\cdot)\right\rangle_\Omega \\ &\overset{(5.143)}{\leq} \left\langle \overline{\boldsymbol{f}}, \overline{\boldsymbol{U}}^\lambda \right\rangle_{Q_s} + 2\left\langle \widetilde{\boldsymbol{U}}^\lambda(0,\cdot) - \widetilde{\boldsymbol{U}}^\lambda(s,\cdot), \widetilde{\boldsymbol{U}}^\lambda(0,\cdot)\right\rangle_\Omega .\end{aligned} \tag{5.146}$$

Then, applying $\limsup_{l,n\to\infty}$ gives that, for a.e. $s \in (0,T)$,

$$\begin{aligned}\limsup_{l,n\to\infty} \left\|\widetilde{\boldsymbol{U}}^{l,n,m}(s,\cdot) - \widetilde{\boldsymbol{U}}^{l,n,m}(0,\cdot)\right\| &\overset{(5.146)}{\leq} \limsup_{l,n\to\infty} \left\langle \overline{\boldsymbol{f}}, \overline{\boldsymbol{U}}^{l,n,m} \right\rangle_{Q_s} \\ &\quad + 2\limsup_{l,n\to\infty} \left\langle \widetilde{\boldsymbol{U}}^{l,n,m}(0,\cdot) - \widetilde{\boldsymbol{U}}^{l,n,m}(s,\cdot), \widetilde{\boldsymbol{U}}^{l,n,m}(0,\cdot)\right\rangle_\Omega \\ &= \langle \boldsymbol{f}, \boldsymbol{u}^m\rangle_{Q_s} + 2\langle \boldsymbol{u}_0 - \boldsymbol{u}^m(s,\cdot), \boldsymbol{u}_0\rangle_\Omega \\ &= \langle \boldsymbol{f}, \boldsymbol{u}^m\rangle_{Q_s} + 2\langle \boldsymbol{u}^m(0,\cdot) - \boldsymbol{u}^m(s,\cdot), \boldsymbol{u}_0\rangle_\Omega ,\end{aligned} \tag{5.147}$$

since we have the convergence $\overline{\boldsymbol{U}}^{l,n,m} \rightharpoonup \boldsymbol{u}^m$ weakly in $\mathrm{L}^q(0,T;\mathrm{W}^{1,q}_0(\Omega)^d)$ by (5.81), $\overline{\boldsymbol{f}} \to \boldsymbol{f}$ strongly in $\mathrm{L}^{q'}(0,T;\mathrm{W}^{-1,q'}(\Omega)^d)$ by (5.22), $\widetilde{\boldsymbol{U}}^{l,n,m}(0,\cdot) \to \boldsymbol{u}_0$ strongly in $\mathrm{L}^2(\Omega)^d$ by (5.77) and $\widetilde{\boldsymbol{U}}^{l,n,m}(s,\cdot) \to \boldsymbol{u}^m(s,\cdot)$ strongly in $\mathrm{L}^2(\Omega)^d$ for a.e. $s \in (0,T)$ by (5.76). In the final line we have used that $\boldsymbol{u}^m(0,\cdot) = \boldsymbol{u}_0$. Let us denote by $N^m \subset [0,T]$ the zero subset of times for which (5.147) does not hold. Applying (5.147) in (5.142) and taking $\liminf_{s\to 0_+}$ omitting the zero set N^m we have that

$$\begin{aligned}0 &\leq \liminf_{(0,T)\setminus N^m \ni s\to 0} \|\boldsymbol{u}^m(s,\cdot) - \boldsymbol{u}_0\|^2_{\mathrm{L}^2(\Omega)} \\ &\overset{(5.142)}{\leq} \liminf_{(0,T)\setminus N^m \ni s\to 0} \limsup_{l,n\to\infty} \left\|\widetilde{\boldsymbol{U}}^{l,n,m}(s,\cdot) - \widetilde{\boldsymbol{U}}^{l,n,m}(0,\cdot)\right\|^2_{\mathrm{L}^2(\Omega)} \\ &\overset{(5.147)}{\leq} \liminf_{(0,T)\setminus N^m \ni s\to 0} \langle \boldsymbol{f}, \boldsymbol{u}^m\rangle_{Q_s} + 2\langle \boldsymbol{u}^m(0,\cdot) - \boldsymbol{u}^m(s,\cdot), \boldsymbol{u}_0\rangle_\Omega = 0,\end{aligned} \tag{5.148}$$

where the last equality follows from the absolute continuity of the integral and from the fact that $\boldsymbol{u}^m \in \mathrm{C}_w([0,T];\mathrm{L}^2_{\mathrm{div}}(\Omega)^d)$, $\boldsymbol{u}_0 \in \mathrm{L}^2_{\mathrm{div}}(\Omega)^d$ and $\mathrm{L}^2_{\mathrm{div}}(\Omega)^d \hookrightarrow (\mathrm{L}^2_{\mathrm{div}}(\Omega)^d)'$. This shows (5.123).

Step 4: Energy identity
Recall that $\boldsymbol{u}^m \in \mathrm{X}_{\mathrm{div}}(Q) \hookrightarrow \mathrm{L}^{\min(q,2q')}(0,T;\mathrm{X}_{\mathrm{div}}(\Omega))$ and $\partial_t \boldsymbol{u}^m \in \mathrm{L}^{\tau}(0,T;(\mathrm{X}_{\mathrm{div}}(\Omega))')$, with $\tau = \min(q',(2q')')$, and equation (5.121) is satisfied. Because of the lack of integrability in time, but still full admissibility in the convective term, an approximation procedure by means of mollification in time can be applied to prove an energy identity for almost all times, which is similar to [Lio69, Ch. 2.5] and [BGMS12, Sec. 3.3].

For this let $0 < s_0 < s < T$ be arbitrary (possibly omitting a zero set) but fixed. Choose $\delta \in (0, \frac{1}{3}\min(s_0, s-s_0, T-s)$ and define a piecewise affine function $\psi_\delta \colon \mathbb{R} \to [0,T]$, such that $\psi_\delta|_{(s_0+\delta,s-\delta)} = 1$, and both $\psi_\delta|_{(s_0-\delta,s_0+\delta)}$ and $\psi_\delta|_{(s-\delta,s+\delta)}$ are affine and $\operatorname{supp}\psi_\delta = [s_0-\delta, s+\delta] \subset (0,T)$. Then, for $\varepsilon > 0$ let $\{\rho_\varepsilon\}_{\varepsilon>0}$ be a family of standard mollifiers in time, i.e., for each $\varepsilon > 0$ let $\rho_\varepsilon \colon \mathbb{R} \to \mathbb{R}_{\geq 0}$ be compactly supported in $[-\varepsilon,\varepsilon]$, smooth, symmetric and such that it integrates to 1. For a Banach space X and a function $f \in \mathrm{L}^1(0,T;\mathrm{X})$ extended to $\mathbb{R}$ by 0 we define the mollification in time in the usual manner

$$(\rho_\varepsilon * f)(t,\boldsymbol{x}) \coloneqq \int_{\mathbb{R}} \rho_\varepsilon(t-s) f(s,\boldsymbol{x})\,\mathrm{d}s.$$

Extending $\boldsymbol{u}^m \in \mathrm{L}^\infty(0,T;\mathrm{L}^2_{\mathrm{div}}(\Omega)^d) \cap \mathrm{L}^q(0,T;\mathrm{W}^{1,q}_{0,\mathrm{div}}(\Omega)^d)$ by $\mathbf{0}$ to $\mathbb{R}$ in time without changing the notation, for $\delta > 0$ as specified above and for $\varepsilon \in (0,\delta)$ let us introduce

$$\boldsymbol{u}^m_{\varepsilon,\delta} \coloneqq \psi_\delta(\rho_\varepsilon * \rho_\varepsilon * (\psi_\delta \boldsymbol{u}^m)),$$

where the convolution is understood in the componentwise sense.

Since we have that $\boldsymbol{u}^m \in \mathrm{L}^{\min(q,2q')}(0,T;\mathrm{X}_{\mathrm{div}}(\Omega))$ it follows by construction that $\boldsymbol{u}^m_{\varepsilon,\delta} \in \mathrm{C}([0,T];\mathrm{X}_{\mathrm{div}}(\Omega))$. Testing with $\boldsymbol{u}^m_{\varepsilon,\delta}(t,\cdot) \in \mathrm{X}_{\mathrm{div}}(\Omega)$ in (5.121), which is admissible for any $t \in [0,T]$, and integrating in $(0,T)$ we obtain

$$\begin{aligned}\left\langle \partial_t \boldsymbol{u}^m, \boldsymbol{u}^m_{\varepsilon,\delta}\right\rangle_Q + \left\langle \boldsymbol{u}^m \otimes \boldsymbol{u}^m, \nabla \boldsymbol{u}^m_{\varepsilon,\delta}\right\rangle_Q + \left\langle \mathbf{S}^m, \mathbf{D}\boldsymbol{u}^m_{\varepsilon,\delta}\right\rangle_Q &\\ + \frac{1}{m}\left\langle |\boldsymbol{u}^m|^{2q'-2}\,\boldsymbol{u}^m, \boldsymbol{u}^m_{\varepsilon,\delta}\right\rangle_Q = \left\langle \boldsymbol{f}, \boldsymbol{u}^m_{\varepsilon,\delta}\right\rangle_Q,&\end{aligned} \tag{5.149}$$

and since $\boldsymbol{u}^m \otimes \boldsymbol{u}^m \in \mathrm{L}^{q'}(Q)^{d\times d}$ all the terms are well-defined.

Let us examine term by term: For the first term on the left-hand side, by the fact that $\operatorname{supp}\psi_\delta \subset\subset (0,T)$, by integration by parts and the product rule, we find that

$$\begin{aligned}\left\langle \partial_t \boldsymbol{u}^m, \boldsymbol{u}^m_{\varepsilon,\delta}\right\rangle_Q &= \langle \partial_t \boldsymbol{u}^m, \psi_\delta\left[\rho_\varepsilon * \rho_\varepsilon * (\psi_\delta \boldsymbol{u}^m)\right]\rangle_{\mathbb{R}\times\Omega}\\ &= -\left\langle \boldsymbol{u}^m, \partial_t\left(\psi_\delta\left[\rho_\varepsilon * \rho_\varepsilon * (\psi_\delta \boldsymbol{u}^m)\right]\right)\right\rangle_{\mathbb{R}\times\Omega}\\ &= -\left\langle \boldsymbol{u}^m \psi_\delta', \rho_\varepsilon * \rho_\varepsilon * (\psi_\delta \boldsymbol{u}^m)\right\rangle_{\mathbb{R}\times\Omega}\\ &\quad - \langle \psi_\delta \boldsymbol{u}^m, \partial_t\left(\rho_\varepsilon * \rho_\varepsilon * (\psi_\delta \boldsymbol{u}^m)\right)\rangle_{\mathbb{R}\times\Omega}.\end{aligned} \tag{5.150}$$

For the second term of (5.150) we find, again with the compact support and by the definition of the convolution, that

$$\begin{aligned}-\langle \psi_\delta \boldsymbol{u}^m, \partial_t\left(\rho_\varepsilon * \rho_\varepsilon * (\psi_\delta \boldsymbol{u}^m)\right)\rangle_{\mathbb{R}\times\Omega} &= \langle \partial_t(\psi_\delta \boldsymbol{u}^m), \rho_\varepsilon * \rho_\varepsilon * (\psi_\delta \boldsymbol{u}^m)\rangle_{\mathbb{R}\times\Omega}\\ &= \langle \rho_\varepsilon * \partial_t(\psi_\delta \boldsymbol{u}^m), \rho_\varepsilon * (\psi_\delta \boldsymbol{u}^m)\rangle_{\mathbb{R}\times\Omega}\\ &= \langle \partial_t(\rho_\varepsilon * (\psi_\delta \boldsymbol{u}^m)), \rho_\varepsilon * (\psi_\delta \boldsymbol{u}^m)\rangle_{\mathbb{R}\times\Omega}\\ &= \frac{1}{2}\int_{\mathbb{R}} \frac{\mathrm{d}}{\mathrm{d}t}\left\|(\rho_\varepsilon * (\psi_\delta \boldsymbol{u}^m))(t,\cdot)\right\|^2_{\mathrm{L}^2(\Omega)}\,\mathrm{d}t = 0,\end{aligned} \tag{5.151}$$

since the function $\rho_\varepsilon * (\psi_\delta \boldsymbol{u}^m) \in \mathrm{C}([0,T]; \mathrm{L}^2_{\operatorname{div}}(\Omega)^d)$ has compact (in time) support. Using this in (5.150) and by the fact that $\psi_\delta \in \mathrm{W}^{1,\infty}(\mathbb{R})$ and $\boldsymbol{u}^m \in \mathrm{L}^\infty(0,T; \mathrm{L}^2_{\operatorname{div}}(\Omega)^d)$, we obtain that

$$
\begin{aligned}
\left\langle \partial_t \boldsymbol{u}^m, \boldsymbol{u}^m_{\varepsilon,\delta} \right\rangle_Q &\overset{(5.150),(5.151)}{=} - \left\langle \boldsymbol{u}^m \psi'_\delta, \rho_\varepsilon * \rho_\varepsilon * (\psi_\delta \boldsymbol{u}^m) \right\rangle_{\mathbb{R}\times\Omega} + 0 \\
&\to - \int_{\mathbb{R}} \|\boldsymbol{u}^m(t,\cdot)\|^2_{\mathrm{L}^2(\Omega)} \psi_\delta(t) \psi'_\delta(t) \,\mathrm{d}t, \qquad \text{as } \varepsilon \to 0,
\end{aligned}
\tag{5.152}
$$

since in particular $\boldsymbol{u}^m \psi'_\delta \in \mathrm{L}^\infty(0,T; \mathrm{L}^2_{\operatorname{div}}(\Omega)^d)$ and $\rho_\varepsilon * \rho_\varepsilon * (\psi_\delta \boldsymbol{u}^m) \to \psi_\delta \boldsymbol{u}^m$ strongly in $\mathrm{L}^\infty(0,T; \mathrm{L}^2_{\operatorname{div}}(\Omega)^d)$, as $\varepsilon \to 0$. For the second term in (5.149), we have by the fact that $\boldsymbol{u}^m \in \mathrm{L}^q(0,T; \mathrm{W}^{1,q}(\Omega)^d)$ that $\nabla \boldsymbol{u}^m_{\varepsilon,\delta} = \psi_\delta(\rho_\varepsilon * \rho_\varepsilon * (\psi_\delta \nabla \boldsymbol{u}^m)) \to \psi_\delta^2 \nabla \boldsymbol{u}^m$ strongly in $\mathrm{L}^q(Q)^{d\times d}$, as $\varepsilon \to 0$, and hence

$$
\begin{aligned}
\left\langle \boldsymbol{u}^m \otimes \boldsymbol{u}^m, \nabla \boldsymbol{u}^m_{\varepsilon,\delta} \right\rangle_Q &\to \left\langle \boldsymbol{u}^m \otimes \boldsymbol{u}^m, \psi_\delta^2 \nabla \boldsymbol{u}^m \right\rangle_Q \\
&= \int_0^T \psi_\delta(t)^2 \left\langle \boldsymbol{u}^m(t,\cdot) \otimes \boldsymbol{u}^m(t,\cdot), \nabla \boldsymbol{u}^m(t,\cdot) \right\rangle_\Omega \,\mathrm{d}t = 0,
\end{aligned}
\tag{5.153}
$$

since we still have admissibility in the convective term as $\boldsymbol{u}^m \otimes \boldsymbol{u}^m \in \mathrm{L}^{q'}(Q)^{d\times d}$, and it vanishes for a.e. $t \in (0,T)$, since $b(\boldsymbol{v},\boldsymbol{v},\boldsymbol{v}) = 0$ for divergence-free functions $\boldsymbol{v}$.

For the remaining terms in (5.149) note that $\boldsymbol{u}^m \in \mathrm{L}^q(0,T; \mathrm{W}^{1,q}_{0,\operatorname{div}}(\Omega)^d) \cap \mathrm{L}^{2q'}(Q)^d$, and hence we have that $\boldsymbol{u}^m_{\varepsilon,\delta} \to \psi_\delta^2 \boldsymbol{u}^m$ strongly in $\mathrm{L}^q(0,T; \mathrm{W}^{1,q}_{0,\operatorname{div}}(\Omega)^d) \cap \mathrm{L}^{2q'}(Q)^d$. Applying this, together with (5.152) and (5.153), taking the limit $\varepsilon \to 0$ in (5.149) yields

$$
\begin{aligned}
&- \int_0^T \|\boldsymbol{u}^m(t,\cdot)\|^2_{\mathrm{L}^2(\Omega)} \psi_\delta(t) \psi'_\delta(t) \,\mathrm{d}t + \left\langle \mathbf{S}^m, \mathbf{D}\boldsymbol{u}^m \psi_\delta^2 \right\rangle_Q \\
&\qquad + \frac{1}{m} \int_Q |\boldsymbol{u}^m|^{2q'} \psi_\delta^2 \,\mathrm{d}\boldsymbol{z} = \left\langle \boldsymbol{f}, \boldsymbol{u}^m \psi_\delta^2 \right\rangle_Q .
\end{aligned}
\tag{5.154}
$$

Now taking the limit $\delta \to 0$, it follows that

$$
\begin{aligned}
&\frac{1}{2} \|\boldsymbol{u}^m(s,\cdot)\|^2_{\mathrm{L}^2(\Omega)} - \tfrac{1}{2} \|\boldsymbol{u}^m(s_0,\cdot)\|^2_{\mathrm{L}^2(\Omega)} + \langle \mathbf{S}^m, \mathbf{D}\boldsymbol{u}^m \rangle_{Q^s_{s_0}} \\
&\qquad + \frac{1}{m} \|\boldsymbol{u}^m\|^{2q'}_{\mathrm{L}^{2q'}(Q^s_{s_0})} = \langle \boldsymbol{f}, \boldsymbol{u}^m \rangle_{Q^s_{s_0}} ,
\end{aligned}
\tag{5.155}
$$

if $s, s_0 \in (0,T)$ are chosen as Lebesgue points of the function $t \mapsto \|\boldsymbol{u}^m(t,\cdot)\|^2_{\mathrm{L}^2(\Omega)}$, see, e.g. [Lio69, p. 214]. Hence (5.155) holds for a.e. $s, s_0 \in (0,T)$. Since the initial condition is attained in the sense of (5.123) this implies that

$$
\frac{1}{2} \|\boldsymbol{u}^m(s,\cdot)\|^2_{\mathrm{L}^2(\Omega)} + \langle \mathbf{S}^m, \mathbf{D}\boldsymbol{u}^m \rangle_{Q_s} + \frac{1}{m} \|\boldsymbol{u}^m\|^{2q'}_{\mathrm{L}^{2q'}(Q_s)} = \langle \boldsymbol{f}, \boldsymbol{u}^m \rangle_{Q_s} + \frac{1}{2} \|\boldsymbol{u}_0\|^2_{\mathrm{L}^2(\Omega)} , \tag{5.156}
$$

for a.e. $s \in (0,T)$.

Step 5: Identification of the implicit relation

Recall that we have by the assertion (5.68) that the inclusion $(\mathbf{D}\overline{\boldsymbol{U}}^{l,n,m}(\boldsymbol{z}), \mathbf{S}^{l,n,m}(\boldsymbol{z})) \in \mathcal{A}(\boldsymbol{z})$ holds for a.e. $\boldsymbol{z} \in Q$. Furthermore, by (5.81) 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 (5.84) 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 3.16

it suffices to show that

$$\limsup_{l,n\to\infty} \left\langle \mathbf{S}^{l,n,m}, \mathbf{D}\overline{\boldsymbol{U}}^{l,n,m} \right\rangle_{Q_s} \le \langle \mathbf{S}^m, \mathbf{D}\boldsymbol{u}^m \rangle_{Q_s} , \tag{5.157}$$

in order to obtain $(\mathbf{D}\boldsymbol{u}^m(\boldsymbol{z}), \mathbf{S}^m(\boldsymbol{z})) \in \mathcal{A}(\boldsymbol{z})$ for a.e. $\boldsymbol{z} \in Q_s$. Then we can exhaust Q by letting $s \to T$. We can only show (5.157) for a.e. $s \in (0,T)$ since the energy identity (5.156) is available only for a.e. $s \in (0,T)$ and some of the arguments used to show (5.157) are only available for a.e. $s \in (0,T)$. For this reason we require a localised version of the Minty type convergence lemma.

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

$$\begin{aligned} \left\langle \mathbf{S}^{l,n,m}, \mathbf{D}\overline{\boldsymbol{U}}^{l,n,m} \right\rangle_{Q_s} &= \left\langle \overline{\mathbf{S}}^{l,n,m}, \mathbf{D}\overline{\boldsymbol{U}}^{l,n,m} \right\rangle_{Q_s} + \left\langle \mathbf{S}^{l,n,m} - \overline{\mathbf{S}}^{l,n,m}, \mathbf{D}\overline{\boldsymbol{U}}^{l,n,m} \right\rangle_{Q_s} \\ &=: \mathrm{I} + \mathrm{II}, \end{aligned} \tag{5.158}$$

where the first term appears in the equation (5.66) for the approximate solutions and the second term has to be shown to vanish. The energy inequality (5.74) yields that

$$\begin{aligned} \mathrm{I} = \left\langle \overline{\mathbf{S}}^{l,n,m}, \mathbf{D}\overline{\boldsymbol{U}}^{l,n,m} \right\rangle_{Q_s} \overset{(5.74)}{\le} & -\frac{1}{2} \left\| \widetilde{\boldsymbol{U}}^{l,n,m}(s,\cdot) \right\|_{\mathrm{L}^2(\Omega)}^2 + \frac{1}{2} \left\| \widetilde{\boldsymbol{U}}^{l,n,m}(0,\cdot) \right\|_{\mathrm{L}^2(\Omega)}^2 \\ & + \left\langle \overline{\boldsymbol{f}}, \overline{\boldsymbol{U}}^{l,n,m} \right\rangle_{Q_s} - \frac{1}{m} \left\| \overline{\boldsymbol{U}}^{l,n,m} \right\|_{\mathrm{L}^{2q'}(Q_s)}^{2q'} . \end{aligned} \tag{5.159}$$

For the second term in (5.158), for $l \in \mathbb{N}$ let $j \in \{1, \dots, l\}$ be such that $s \in (t_{j-1}, t_j]$, i.e., j depends on s and on l. As before in (5.144) by the relation (5.46) we have that

$$\left\langle \mathbf{S}^{l,n,m} - \overline{\mathbf{S}}^{l,n,m}, \mathbf{D}\overline{\boldsymbol{U}}^{l,n,m} \right\rangle_{Q_{i-1}^i} = 0 \tag{5.160}$$

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

$$\begin{aligned} \mathrm{II} &= \left\langle \mathbf{S}^{l,n,m} - \overline{\mathbf{S}}^{l,n,m}, \mathbf{D}\overline{\boldsymbol{U}}^{l,n,m} \right\rangle_{Q_{t_j}} - \left\langle \mathbf{S}^{l,n,m} - \overline{\mathbf{S}}^{l,n,m}, \mathbf{D}\overline{\boldsymbol{U}}^{l,n,m} \right\rangle_{Q_s^{t_j}} \\ &\overset{(5.160)}{=} 0 - \left\langle \mathbf{S}^{l,n,m}, \mathbf{D}\overline{\boldsymbol{U}}^{l,n,m} \right\rangle_{Q_s^{t_j}} + \left\langle \overline{\mathbf{S}}^{l,n,m}, \mathbf{D}\overline{\boldsymbol{U}}^{l,n,m} \right\rangle_{Q_s^{t_j}} \le \left\langle \overline{\mathbf{S}}^{l,n,m}, \mathbf{D}\overline{\boldsymbol{U}}^{l,n,m} \right\rangle_{Q_s^{t_j}} , \end{aligned} \tag{5.161}$$

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 (5.68). For the remaining term we use again (5.74) on (s, t_j), noting that the term involving $\frac{1}{m}$ is nonnegative, which yields

$$\begin{aligned} \mathrm{I} \le \left\langle \overline{\mathbf{S}}^{l,n,m}, \mathbf{D}\overline{\boldsymbol{U}}^{l,n,m} \right\rangle_{Q_s^{t_j}} &\le \left\langle \overline{\boldsymbol{f}}, \overline{\boldsymbol{U}}^{l,n,m} \right\rangle_{Q_s^{t_j}} \\ & \quad - \frac{1}{2} \left\| \widetilde{\boldsymbol{U}}^{l,n,m}(t_j,\cdot) \right\|_{\mathrm{L}^2(\Omega)}^2 + \frac{1}{2} \left\| \widetilde{\boldsymbol{U}}^{l,n,m}(s,\cdot) \right\|_{\mathrm{L}^2(\Omega)}^2 . \end{aligned} \tag{5.162}$$

By the duality of norms, the estimate (5.91) (the dependence on m arises from the regularisation term only) and by (5.21) we obtain

$$\begin{aligned} \left\langle \overline{\boldsymbol{f}}, \overline{\boldsymbol{U}}^{l,n,m} \right\rangle_{Q_s^{t_j}} &\le \left\| \overline{\boldsymbol{f}} \right\|_{\mathrm{L}^{q'}(s,t_j;\mathrm{W}^{-1,q'}(\Omega))} \left\| \overline{\boldsymbol{U}}^{l,n,m} \right\|_{\mathrm{L}^q(0,T;\mathrm{W}^{1,q}(\Omega))} \\ &\le c \left\| \boldsymbol{f} \right\|_{\mathrm{L}^{q'}(t_{j-1},t_j;\mathrm{W}^{-1,q'}(\Omega))} \le c \left\| \boldsymbol{f} \right\|_{\mathrm{L}^{q'}(s-\delta_l,s+\delta_l;\mathrm{W}^{-1,q'}(\Omega))} . \end{aligned} \tag{5.163}$$

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

$$\begin{aligned}
\mathrm{II} &\overset{(5.161),(5.162)}{\leq} \left\langle \overline{\boldsymbol{f}}, \overline{\boldsymbol{U}}^{l,n,m} \right\rangle_{Q_s^{t_j}} - \frac{1}{2}\left\| \widetilde{\boldsymbol{U}}^{l,n,m}(t_j,\cdot)\right\|^2_{\mathrm{L}^2(\Omega)} + \frac{1}{2}\left\| \widetilde{\boldsymbol{U}}^{l,n,m}(s,\cdot)\right\|^2_{\mathrm{L}^2(\Omega)} \\
&\overset{(5.163)}{\leq} c(m)\left\| \boldsymbol{f}\right\|_{\mathrm{L}^{q'}(s-\delta_l,s+\delta_l;\mathrm{W}^{-1,q'}(\Omega))} \\
&\qquad - \frac{1}{2}\left\| \overline{\boldsymbol{U}}^{l,n,m}(s,\cdot)\right\|^2_{\mathrm{L}^2(\Omega)} + \frac{1}{2}\left\| \widetilde{\boldsymbol{U}}^{l,n,m}(s,\cdot)\right\|^2_{\mathrm{L}^2(\Omega)} .
\end{aligned} \tag{5.164}$$

Note that in order to take the limit $l, n \to \infty$, we had to get rid of any dependence on t_j. Applying $\limsup_{l,n\to\infty}$ to $(\mathrm{I}+\mathrm{II})$ with (5.159) and (5.164), noting that the term involving $\widetilde{\boldsymbol{U}}^{l,n,m}(s,\cdot)$ drops out, we obtain

$$\begin{aligned}
\limsup_{l,n\to\infty} (\mathrm{I}+\mathrm{II}) &\leq -\frac{1}{2}\lim_{l,n\to\infty}\left\| \overline{\boldsymbol{U}}^{l,n,m}(s,\cdot)\right\|^2_{\mathrm{L}^2(\Omega)} + \frac{1}{2}\lim_{l,n\to\infty}\left\| \widetilde{\boldsymbol{U}}^{l,n,m}(0,\cdot)\right\|^2_{\mathrm{L}^2(\Omega)} \\
&\quad - \frac{1}{m}\liminf_{l,n\to\infty}\left\| \overline{\boldsymbol{U}}^{l,n,m}\right\|^{2q'}_{\mathrm{L}^{2q'}(Q_s)} + \lim_{l,n\to\infty}\left\langle \overline{\boldsymbol{f}}, \overline{\boldsymbol{U}}^{l,n,m}\right\rangle_{Q_s} \\
&\quad + \lim_{l\to\infty}\left\|\boldsymbol{f}\right\|_{\mathrm{L}^{q'}(s-\delta_l,s+\delta_l;\mathrm{W}^{-1,q'}(\Omega))} \\
&\leq -\frac{1}{2}\left\|\boldsymbol{u}^m(s,\cdot)\right\|^2_{\mathrm{L}^2(\Omega)} + \frac{1}{2}\left\|\boldsymbol{u}_0\right\|^2_{\mathrm{L}^2(\Omega)} - \frac{1}{m}\left\|\boldsymbol{u}^m\right\|^{2q'}_{\mathrm{L}^{2q'}(Q_s)} + \left\langle \boldsymbol{f}, \boldsymbol{u}^m\right\rangle_{Q_s},
\end{aligned} \tag{5.165}$$

where the last inequality is based on the following arguments. By (5.80) we have that $\overline{\boldsymbol{U}}^{l,n,m}(s,\cdot) \to \boldsymbol{u}^m(s,\cdot)$ 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 (5.77) 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 (5.81). For the forth 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 (5.81) and $\overline{\boldsymbol{f}} \to \boldsymbol{f}$ strongly in $\mathrm{L}^{q'}(0,T;\mathrm{W}^{-1,q'}(\Omega)^d)$ by (5.22), as $l,n\to\infty$. The last term vanishes by the absolute continuity of the integral, as $l\to\infty$. Finally returning to (5.158), applying $\limsup_{l,n\to\infty}$ and the energy identity (5.156) for a.e. $s\in(0,T)$, yields

$$\begin{aligned}
\limsup_{l,n\to\infty}\left\langle \mathbf{S}^{l,n,m}, \mathbf{D}\overline{\boldsymbol{U}}^{l,n,m}\right\rangle_{Q_s} &\overset{(5.158)}{\leq} \limsup_{l,n\to\infty}(\mathrm{I}+\mathrm{II}) \\
&\overset{(5.165)}{\leq} -\frac{1}{2}\left\|\boldsymbol{u}^m(s,\cdot)\right\|^2_{\mathrm{L}^2(\Omega)} + \frac{1}{2}\left\|\boldsymbol{u}_0\right\|^2_{\mathrm{L}^2(\Omega)} - \frac{1}{m}\left\|\boldsymbol{u}^m\right\|^{2q'}_{\mathrm{L}^{2q'}(Q_s)} + \left\langle \boldsymbol{f}, \boldsymbol{u}^m\right\rangle_{Q_s} \\
&\overset{(5.156)}{=} \left\langle \mathbf{S}^m, \mathbf{D}\boldsymbol{u}^m\right\rangle_{Q_s}
\end{aligned} \tag{5.166}$$

for a.e. $s\in(0,T)$. This proves the claim in (5.157) and thus completes the proof. □

Remark 5.8.

(i) *Note that in the existence proof in [BGMS12, Sec. 3.3] the compactness argument is based on the Aubin–Lions lemma, which shows that $\partial_t\boldsymbol{u}^m \in \mathrm{L}^1(0,T,(\mathcal{V}_{\mathrm{div}})')$, where $\mathcal{V}_{\mathrm{div}}$ is a higher order divergence-free function space used for the Galerkin approximation, i.e., $\boldsymbol{u}^{n,m} \in \mathrm{C}([0,T];\mathcal{V}_{\mathrm{div}})$. To show an energy identity as in (5.156) the authors consider the test functions $\boldsymbol{u}^{n,m}_{\varepsilon,\delta}$ instead of $\boldsymbol{u}^m_{\varepsilon,\delta}$, since the term $\left\langle \partial_t\boldsymbol{u}^m, \boldsymbol{u}^{n,m}_{\varepsilon,\delta}\right\rangle$ is well-defined in their situation. Alternatively they could have performed a finer estimate on*

$\partial_t \boldsymbol{u}^m$ and used the arguments above. For the numerical analysis with finite element functions this approach is not attractive, since in general the spaces $\mathbb{V}^n_{\mathrm{div}}$ are not exactly divergence-free. For this reason one cannot test with a mollified version of $\overline{\boldsymbol{U}}^{l,n,m} \in \mathbb{P}^l_0(0,T;\mathbb{V}^n_{\mathrm{div}})$ in (5.121).

(ii) The implicit constitutive relation can also be identified, if the initial condition is not verified. In Step 5, one would consider the parabolic cylinders $Q^s_{s_0}$ instead of $Q_s = Q^s_0$ and then let $s \to T$ and $s_0 \to 0$.

Limit $m \to \infty$

In this step admissibility of the solution as test function is lost and hence we apply the compactness lemma due to Aubin–Lions rather than the one by Simon. Since the numerical convective term is not present anymore and the equation (5.121) is continuous, this does not require any restriction on q. The loss of admissibility of the solution as test function means that we have to use Lipschitz truncation to identify the implicit relation. The availability of the solenoidal Lipschitz truncation allows to simplify the arguments in [BGMS12], since no pressure has to be reconstructed. The proof of the identification is similar to the one in the steady case, but is done locally.

Let us denote

$$\mu := \min\left(\frac{q(d+2)}{2d}, (2q')', q'\right) = \min\left(\hat{q}', \tau\right), \tag{5.167}$$

where τ defined in (5.120). Note that since $q > \frac{2d}{d+2}$, we have that $\mu > 1$.

Lemma 5.9 (Convergence $m \to \infty$).
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 $\mathbf{S}^m \in \mathrm{L}^{q'}(Q)^{d\times d}$ satisfy (5.121)–(5.123), for $m \in \mathbb{N}$. Further, let $\mu > 1$ and $\hat{q}$ be defined in (5.167) and (5.5), respectively.
Then, there exists a constant $c > 0$ such that we have that

$$\begin{aligned} &\|\boldsymbol{u}^m\|^2_{\mathrm{L}^\infty(0,T;\mathrm{L}^2(\Omega))} + \|\boldsymbol{u}^m\|^q_{\mathrm{L}^q(0,T;\mathrm{W}^{1,q}(\Omega))} + \|\mathbf{S}^m\|^{q'}_{\mathrm{L}^{q'}(Q)} \\ &\qquad + \frac{1}{m}\|\boldsymbol{u}^m\|^{2q'}_{\mathrm{L}^{2q'}(Q)} + \|\boldsymbol{u}^m\|_{\mathrm{L}^{\frac{q(d+2)}{d}}(Q)} \le c \end{aligned} \tag{5.168}$$

for all $m \in \mathbb{N}$.
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 $\partial_t \boldsymbol{u} \in \mathrm{L}^\mu(0,T;(\mathrm{W}^{1,\hat{q}}_{0,\mathrm{div}}(\Omega)^d)')$, an $\mathbf{S} \in \mathrm{L}^{q'}(Q)^{d\times d}$ and subsequences such that

$$\boldsymbol{u}^m \to \boldsymbol{u} \quad \text{strongly in } \mathrm{L}^p(0,T;\mathrm{L}^2_{\mathrm{div}}(\Omega)^d) \cap \mathrm{L}^r(Q)^d, \text{ for all } p \in [1,\infty) \text{ and all } r \in [1, \tfrac{q(d+2)}{d}), \tag{5.169}$$

$$\boldsymbol{u}^m(s,\cdot) \to \boldsymbol{u}(s,\cdot) \quad \text{strongly in } \mathrm{L}^2_{\mathrm{div}}(\Omega)^d \text{ for a.e. } s \in (0,T), \tag{5.170}$$

$$\boldsymbol{u}^m \rightharpoonup \boldsymbol{u} \quad \text{weakly in } \mathrm{L}^q(0,T;\mathrm{W}^{1,q}_{0,\mathrm{div}}(\Omega)^d) \cap \mathrm{L}^{\frac{q(d+2)}{d}}(Q)^d, \tag{5.171}$$

$$\boldsymbol{u}^m \overset{*}{\rightharpoonup} \boldsymbol{u} \quad \text{weakly* in } \mathrm{L}^\infty(0,T;\mathrm{L}^2_{\mathrm{div}}(\Omega)^d), \tag{5.172}$$

$$\partial_t \boldsymbol{u}^m \rightharpoonup \partial_t \boldsymbol{u} \quad \text{weakly in } \mathrm{L}^\mu(0,T;(\mathrm{W}^{1,\hat{q}}_{0,\mathrm{div}}(\Omega)^d)'), \tag{5.173}$$

$$\mathbf{S}^m \rightharpoonup \mathbf{S} \quad \text{weakly in } \mathrm{L}^{q'}(Q)^{d\times d}, \tag{5.174}$$

$$\frac{1}{m}|\boldsymbol{u}^m|^{2q'-2}\boldsymbol{u}^m \to \mathbf{0} \quad \text{strongly in } \mathrm{L}^p(Q)^{d\times d}, \text{ for all } p \in [1,(2q')'), \tag{5.175}$$

as $m \to \infty$*.*

Proof.

Step 1: Estimates

Recall that by (5.156) we have the following energy identity for a.e. $t \in (0,T)$:

$$\frac{1}{2}\|\boldsymbol{u}^m(t,\cdot)\|^2_{\mathrm{L}^2(\Omega)} + \langle \mathbf{S}^m, \mathbf{D}\boldsymbol{u}^m\rangle_{Q_t} + \frac{1}{m}\|\boldsymbol{u}^m\|^{2q'}_{\mathrm{L}^{2q'}(Q_t)} = \langle \boldsymbol{f}, \boldsymbol{u}^m\rangle_{Q_t} + \frac{1}{2}\|\boldsymbol{u}_0\|^2_{\mathrm{L}^2(\Omega)}. \tag{5.176}$$

By the fact that $(\mathbf{D}\boldsymbol{u}^m, \mathbf{S}^m) \in \mathcal{A}(\cdot)$ a.e. in Q by (5.122), we can use Assumption 3.11 (A3) to show that

$$\begin{aligned}\langle \mathbf{S}^m, \mathbf{D}\boldsymbol{u}^m\rangle_{Q_t} &\geq -\|g\|_{\mathrm{L}^1(Q_t)} + c_*\left(\|\mathbf{D}\boldsymbol{u}^m\|^q_{\mathrm{L}^q(Q_t)} + \|\mathbf{S}^m\|^{q'}_{\mathrm{L}^{q'}(Q_t)}\right)\\ &\geq -\|g\|_{\mathrm{L}^1(Q_t)} + c\|\boldsymbol{u}^m\|^q_{\mathrm{L}^q(0,t;\mathrm{W}^{1,q}(Q))} + c_*\|\mathbf{S}^m\|^{q'}_{\mathrm{L}^{q'}(Q_t)},\end{aligned} \tag{5.177}$$

where we have used Poincaré's and Korn's inequalities in the last line. Similarly as before, we use duality of norms and Young's inequality with $\varepsilon > 0$ to bound

$$\langle \boldsymbol{f}, \boldsymbol{u}^m\rangle_{Q_t} \leq c(\varepsilon)\|\boldsymbol{f}\|^{q'}_{\mathrm{L}^{q'}(0,T;\mathrm{W}^{-1,q'}(\Omega))} + \varepsilon\|\boldsymbol{u}^m\|^q_{\mathrm{L}^q(0,t;\mathrm{W}^{1,q}(\Omega))}. \tag{5.178}$$

Applying (5.177) and (5.178) in (5.156), rearranging and choosing $\varepsilon > 0$ small enough yields

$$\|\boldsymbol{u}^m(t,\cdot)\|^2_{\mathrm{L}^2(\Omega)} + \|\boldsymbol{u}^m\|^q_{\mathrm{L}^q(0,t;\mathrm{W}^{1,q}(\Omega))} + \|\mathbf{S}^m\|^{q'}_{\mathrm{L}^{q'}(Q_t)} + \frac{1}{m}\|\boldsymbol{u}^m\|^{2q'}_{\mathrm{L}^{2q'}(Q_t)} \leq c \tag{5.179}$$

for a.e. $t \in (0,T)$ and all $m \in \mathbb{N}$. Taking the essential supremum over $t \in (0,T)$ and also applying the parabolic interpolation from Corollary 2.5 shows (5.168).

Step 2: Bound on the time derivative

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 estimate (5.16) on the convective term, which holds since $q \geq \frac{2d}{d+2}$. 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 (5.5). Also we have that $\mu' = \max\left(2q', q, \left(\frac{q(d+2)}{2d}\right)'\right) = \max(2q', \hat{q})$ for μ as in (5.167), i.e., the embedding $\mathrm{L}^{\mu'}(\Omega) \hookrightarrow \mathrm{L}^{\hat{q}}(\Omega) \cap \mathrm{L}^q(\Omega) \cap \mathrm{L}^{2q'}(\Omega)$ is continuous. With this, similarly as in (5.138) applying the uniform estimates in (5.168) one has that

$$\begin{aligned}\left|\langle \mathfrak{L}^m[\boldsymbol{u}^m;\boldsymbol{w}], \varphi\rangle_{(0,T)}\right| &\overset{(5.16)}{\leq} \|\boldsymbol{u}^m\|^2_{\mathrm{L}^{\frac{q(d+2)}{d}}(Q)}\|\varphi\nabla\boldsymbol{w}\|_{\mathrm{L}^{\hat{q}}(Q)} + \|\mathbf{S}^m\|_{\mathrm{L}^{q'}(Q)}\|\varphi\mathbf{D}\boldsymbol{w}\|_{\mathrm{L}^q(Q)}\\ &\quad + \frac{1}{m}\|\boldsymbol{u}^m\|^{2q'-1}_{\mathrm{L}^{2q'}(Q)}\|\varphi\boldsymbol{w}\|_{\mathrm{L}^{2q'}(Q)}\\ &\quad + \|\boldsymbol{f}\|_{\mathrm{L}^{q'}(0,T;\mathrm{W}^{-1,q'}(\Omega))}\|\varphi\boldsymbol{w}\|_{\mathrm{L}^q(0,T;\mathrm{W}^{1,q}(\Omega))}\\ &\leq c\|\varphi\|_{\mathrm{L}^{\mu'}(0,T)}\|\boldsymbol{w}\|_{\mathrm{W}^{1,\hat{q}}(\Omega)}\end{aligned} \tag{5.180}$$

for all $\varphi \in \mathrm{C}^\infty_0((0,T))$ and all $\boldsymbol{w} \in \mathrm{C}^\infty_{0,\mathrm{div}}(\Omega)^d$, and all $m \in \mathbb{N}$. With (5.121) and using the fact that $\mu > 1$ and that the space $\mathrm{L}^\mu(0,T;(\mathrm{W}^{1,\hat{q}}_{0,\mathrm{div}}(\Omega)^d)')$ is reflexive, this shows that $\{\partial_t\boldsymbol{u}^m\}_{m\in\mathbb{N}}$ is bounded in $\mathrm{L}^\mu(0,T;(\mathrm{W}^{1,\hat{q}}_{0,\mathrm{div}}(\Omega)^d)')$.

Step 3: Convergence as $m \to \infty$

Due to the restriction $q > \frac{2d}{d+2}$, the embedding $\mathrm{W}^{1,q}_{0,\mathrm{div}}(\Omega)^d \hookrightarrow\hookrightarrow \mathrm{L}^2_{\mathrm{div}}(\Omega)^d$ is compact. Because $\hat{q} \geq q > \frac{2d}{d+2}$, 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 in (5.19), 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, we can apply the Aubin–Lions compactness lemma (see Lemma 2.7) to obtain from the estimate in (5.168) and the fact that $\{\partial_t \boldsymbol{u}^m\}_{m\in\mathbb{N}}$ is bounded in $\mathrm{L}^\mu(0,T;(\mathrm{W}^{1,\hat{q}}_{0,\mathrm{div}}(\Omega)^d)')$ that there exists a subsequence such that

$$\boldsymbol{u}^m \to \boldsymbol{u} \quad \text{strongly in } \mathrm{L}^q(0,T;\mathrm{L}^2_{\mathrm{div}}(\Omega)^d), \quad \text{as } m \to \infty. \tag{5.181}$$

By the estimates in (5.168), the uniform bound on $\{\partial_t \boldsymbol{u}^m\}_{m\in\mathbb{N}}$ in $\mathrm{L}^\mu(0,T;(\mathrm{W}^{1,\hat{q}}_{0,\mathrm{div}}(\Omega))')$ and the Banach–Alaoglu theorem, there exists a subsequence such that (5.171)–(5.174) hold, where the limits can be identified with the help of (5.181).

The strong convergence in $\mathrm{L}^p(0,T;\mathrm{L}^2_{\mathrm{div}}(\Omega)^d)$ for all $p \in [1,\infty)$ and in $\mathrm{L}^r(Q)^d$ for all $r \in [1, \frac{q(d+2)}{d})$ asserted in (5.169) follows from the strong convergence in $\mathrm{L}^1(Q)^d$ by (5.181), and the boundedness in $\mathrm{L}^\infty(0,T;\mathrm{L}^2_{\mathrm{div}}(\Omega)^d)$ and in $\mathrm{L}^{\frac{q(d+2)}{d}}(Q)^d$ by (5.168) by means of interpolation. The convergence (5.170) is deduced analogously to the proof of (5.76) by the arguments following (5.108). With Hölder's inequality and the estimate in (5.168) we find that

$$\left\| \frac{1}{m} |\boldsymbol{u}^m|^{2q'-2} \boldsymbol{u}^m \right\|_{\mathrm{L}^1(Q)} \leq \frac{c}{m} \|\boldsymbol{u}^m\|^{2q'-1}_{\mathrm{L}^{2q'}(Q)} \overset{(5.168)}{\leq} c m^{-1/2q'} \to 0, \quad \text{as } m \to \infty, \tag{5.182}$$

so strong convergence to $\mathbf{0}$ of the regularisation term in $\mathrm{L}^1(Q)^d$ is proved. One can show uniform boundedness of $\frac{1}{m}|\boldsymbol{u}^m|^{2q'-2}\boldsymbol{u}^m$ in $\mathrm{L}^{(2q')'}(Q)^d$; interpolation between $\mathrm{L}^1(Q)^d$ and $\mathrm{L}^{(2q')'}(Q)^d$ then gives strong convergence to $\mathbf{0}$ in $\mathrm{L}^p(Q)^d$, for any $p \in [1,(2q')')$. Hence (5.175) 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 (5.5), let us introduce

$$\mathfrak{L}[\boldsymbol{u};\boldsymbol{v}](t) := -\, b(\boldsymbol{u}(t,\cdot),\boldsymbol{u}(t,\cdot),\boldsymbol{v}) - \langle \mathbf{S}(t,\cdot), \mathbf{D}\boldsymbol{v}\rangle_\Omega + \langle \boldsymbol{f}(t,\cdot),\boldsymbol{v}\rangle_\Omega\,, \tag{5.183}$$

where $\mathbf{S} \in \mathrm{L}^{q'}(Q)^{d\times d}$ is the limiting function introduced in Lemma 5.9.

Lemma 5.10 (Identification of the PDE as $m \to \infty$).
The limiting 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 5.9 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 (5.5). (Up to a representative) we have that $\boldsymbol{u} \in \mathrm{C}_w([0,T],\mathrm{L}^2_{\mathrm{div}}(\Omega)^d)$. Furthermore, the functions $\boldsymbol{u}$ and $\mathbf{S} \in \mathrm{L}^{q'}(Q)^{d\times d}$ from Lemma 5.9 satisfy

$$\langle \partial_t \boldsymbol{u}(t,\cdot), \boldsymbol{w}\rangle_\Omega = \mathfrak{L}[\boldsymbol{u};\boldsymbol{w}](t) \quad \textit{for all } \boldsymbol{w} \in \mathrm{C}^\infty_{0,\mathrm{div}}(\Omega)^d,\ \textit{for a.e. } t \in (0,T), \tag{5.184}$$

$$(\mathbf{D}\boldsymbol{u}(\boldsymbol{z}),\mathbf{S}(\boldsymbol{z})) \in \mathcal{A}(\boldsymbol{z}) \quad \textit{for a.e. } \boldsymbol{z} \in Q, \tag{5.185}$$

$$\operatorname*{ess\,lim}_{t\to 0_+} \|\boldsymbol{u}(t,\cdot) - \boldsymbol{u}_0\|_{\mathrm{L}^2(\Omega)} = 0, \tag{5.186}$$

i.e., $(\boldsymbol{u},\mathbf{S})$ is a weak solution according to Definition 5.1.

Proof.

Step 1: Identification of the limiting equation
Let $\boldsymbol{w} \in \mathrm{C}^\infty_{0,\mathrm{div}}(\Omega)^d$ and $\varphi \in \mathrm{C}^\infty_0((0,T))$ and let us consider each of the terms in (5.121) and

(5.184). By the weak convergence in (5.173) and (5.174) we have that

$$\langle \partial_t \boldsymbol{u}^m, \varphi \boldsymbol{w} \rangle_Q \to \langle \partial_t \boldsymbol{u}, \varphi \boldsymbol{w} \rangle_Q \,, \tag{5.187}$$

$$\langle \mathbf{S}^m, \varphi \mathbf{D}\boldsymbol{w} \rangle_Q \to \langle \mathbf{S}, \varphi \mathbf{D}\boldsymbol{w} \rangle_Q \,, \tag{5.188}$$

as $m \to \infty$. Since by (5.169) we have that $\boldsymbol{u}^m \to \boldsymbol{u}$ in $\mathrm{L}^r(Q)^d$ for all $r \in \left[1, \frac{q(d+2)}{d}\right)$ 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

$$\langle \boldsymbol{u}^m \otimes \boldsymbol{u}^m, \varphi \nabla \boldsymbol{w} \rangle_Q \to \langle \boldsymbol{u} \otimes \boldsymbol{u}, \varphi \nabla \boldsymbol{w} \rangle_Q \,, \qquad \text{as } m \to \infty. \tag{5.189}$$

Taking the results in (5.187)–(5.189) and (5.175) shows that (5.121) implies (5.184).

Step 2: Identification of the initial condition
Similarly as in the proof of Lemma 5.7, 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 (5.186) using (5.170), (5.171), the fact that $\boldsymbol{u}^m(0,\cdot) = \boldsymbol{u}_0$ shown in Step 3 in the proof of Lemma 5.7 and the identity (5.151).

Step 3: Higher integrability of the time derivative
As in Step 2 in the proof of Lemma 5.6 we can improve the integrability of $\partial_t \boldsymbol{u}$ using the fact that (5.184) is satisfied. Using the estimate (5.16) on the convective term, this yields 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 (5.5).

Step 4: Identification of the implicit relation (compare [BGMS12] and [BDS13, Sec. 3])
Recall that $\mathbf{D}\boldsymbol{u}^m \rightharpoonup \mathbf{D}\boldsymbol{u}$ weakly in $\mathrm{L}^q(Q)^{d\times d}$ by (5.171), that $\mathbf{S}^m \rightharpoonup \mathbf{S}$ weakly in $\mathrm{L}^{q'}(Q)^{d\times d}$ by (5.174) and that we have that $(\mathbf{D}\boldsymbol{u}^m(\boldsymbol{z}), \mathbf{S}^m(\boldsymbol{z})) \in \mathcal{A}(\boldsymbol{z})$ for a.e. $\boldsymbol{z} \in Q$ by (5.122). Hence, by Lemma 3.16, it suffices to show that

$$\limsup_{m\to\infty} \langle \mathbf{S}^m, \mathbf{D}\boldsymbol{u}^m \rangle_{\widetilde{Q}} \le \langle \mathbf{S}, \mathbf{D}\boldsymbol{u} \rangle_{\widetilde{Q}} \,, \tag{5.190}$$

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 $\boldsymbol{u}$ if $q < \frac{3d+2}{d+2}$, 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 [BGMS12] we will not use a parabolic Lipschitz truncation after locally reconstructing the approximations to the pressure, but we apply the solenoidal Lipschitz truncation introduced subsequently in [BDS13] and stated in Lemma 2.14, since this simplifies the argument.

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

$$\langle \partial_t \boldsymbol{v}^m, \boldsymbol{\xi} \rangle_Q = \langle \boldsymbol{u}^m \otimes \boldsymbol{u}^m - \boldsymbol{u} \otimes \boldsymbol{u}, \nabla \boldsymbol{\xi} \rangle_Q - \langle \mathbf{S}^m - \mathbf{S}, \mathbf{D}\boldsymbol{\xi} \rangle_Q - \frac{1}{m} \left\langle |\boldsymbol{u}^m|^{2q'-2} \boldsymbol{u}^m, \boldsymbol{\xi} \right\rangle_Q , \tag{5.191}$$

by (5.121) and (5.184) and by the density of $\mathrm{C}^\infty_0(0,T) \times \mathrm{C}^\infty_{0,\mathrm{div}}(\Omega)^d$ in $\mathrm{C}^\infty_{0,\mathrm{div}}(Q)^d$.

Due to the (lower order) regularising term we aim to apply Corollary 2.15 instead of Lemma 2.14 with $p = q \in (1, \infty)$ and σ such that

$$1 < \sigma < \min\left(2, q, q', \frac{q(d+2)}{2d}, (2q')'\right) = \min\left(q', \frac{q(d+2)}{2d}, (2q')'\right). \tag{5.192}$$

Such a σ exists, since we have by assumption that $q > \frac{2d}{d+2}$. First note that $\boldsymbol{u}$ and $\boldsymbol{u}^m$ are (weakly) divergence-free, and so is $\boldsymbol{v}^m$, and $\boldsymbol{v}^m \rightharpoonup \mathbf{0}$ weakly in $\mathrm{L}^q(I_0;\mathrm{W}^{1,q}(B_0)^d)$, as $m\to\infty$, by (5.171). Since $\boldsymbol{u}^m \to \boldsymbol{u}$ strongly in $\mathrm{L}^p(Q)$ for $p\in[1,\frac{q(d+2)}{d})$ by (5.169) and $\sigma<\frac{q(d+2)}{d}$ we have that $\boldsymbol{v}^m\to\mathbf{0}$ 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 (5.168) 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

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

Note that $\mathbf{G}_1^m \rightharpoonup \mathbf{0}$ weakly in $\mathrm{L}^{q'}(Q_0)^{d\times d}$ by (5.174). By (5.169) we have that $\boldsymbol{u}^m\to\boldsymbol{u}$ in $\mathrm{L}^r(Q_0)^d$ 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 $\mathrm{L}^r(Q_0)^{d\times d}$ 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 (5.175) we have that $\boldsymbol{g}^m\to\mathbf{0}$ strongly in $\mathrm{L}^p(Q_0)^d$, as $m\to\infty$, for any $p\in[1,(2q')')$ and hence also for $p=\sigma<(2q')'$. This means that all the assumptions of Corollary 2.15 are satisfied and hence the statement of Lemma 2.14 applies.

Analogously to the steady case in Step 2 in the proof of Lemma 4.8, with the aid of the parabolic solenoidal Lipschitz truncation we show that

$$\lim_{m\to\infty}\int_{\frac18 Q_0}\left[(\mathbf{S}^m-\mathcal{S}^\star(\cdot,\mathbf{D}\boldsymbol{u})):(\mathbf{D}\boldsymbol{u}^m-\mathbf{D}\boldsymbol{u})\right]^{1/2}\,\mathrm{d}\boldsymbol{z}=0, \tag{5.193}$$

where the exponent $1/2$ 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 $(\mathbf{D}\boldsymbol{u},\mathcal{S}^\star(\cdot,\mathbf{D}\boldsymbol{u}))\in\mathcal{A}(\cdot)$ and $(\mathbf{D}\boldsymbol{u}^m,\mathbf{S}^m)\in\mathcal{A}(\cdot)$ a.e. in Q by (5.122), it follows that the $\liminf_{m\to\infty}$ of the above is nonnegative. To show the other direction, denote $H^m:=(\mathbf{S}^m-\mathcal{S}^\star(\cdot,\mathbf{D}\boldsymbol{u})):(\mathbf{D}\boldsymbol{u}^m-\mathbf{D}\boldsymbol{u})\geq 0$, and let $j\geq j_0$, $\mathcal{B}_{m,j}\subset Q_0$ and $\boldsymbol{v}^{m,j}$ be given by Lemma 2.14 applied on Q_0, and by (ii) therein we have that $\boldsymbol{v}^m=\boldsymbol{v}^{m,j}$ on $\frac18 Q_0\backslash\mathcal{B}_{m,j}$. Dividing the domain into $\frac18 Q_0\cap\mathcal{B}_{m,j}$ and $\frac18 Q_0\backslash\mathcal{B}_{m,j}$ and applying Hölder's inequality as in the proof of Lemma 4.8, (4.91), (4.92) we obtain

$$\int_{\frac18 Q_0}(H^m)^{\frac12}\,\mathrm{d}\boldsymbol{z}\leq c\,|\mathcal{B}_{m,j}|^{\frac12}+c\left(\int_{\frac18 Q_0\backslash\mathcal{B}_{m,j}}H^m\,\mathrm{d}\boldsymbol{z}\right)^{\frac12}\qquad\text{for all } m,j\in\mathbb{N}, \tag{5.194}$$

using also the nonnegativity of H^m and the fact that $\{H^m\}_{m\in\mathbb{N}}$ is bounded in $\mathrm{L}^1(Q)$ by the a priori estimate in (5.168). By Lemma 2.14 (iii) we have that

$$\limsup_{m\to\infty}|\mathcal{B}_{m,j}|^{\frac12}\leq\limsup_{m\to\infty}(\lambda_{m,j}^q\,|\mathcal{B}_{m,j}|)^{\frac12}\leq c2^{-\frac{j}{2}}. \tag{5.195}$$

Let $\zeta\in\mathrm{C}_0^\infty(\frac16 B_0)$ be the nonnegative function given by Lemma 2.14 such that $\zeta|_{\frac18 B_0}\equiv 1$. In the second term in (5.194) 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 have

$$\begin{aligned}\int_{\frac18 Q_0\backslash\mathcal{B}_{m,j}}H^m\,\mathrm{d}\boldsymbol{z}&=\int_{\frac18 Q_0\backslash\mathcal{B}_{m,j}}H^m\,\zeta\,\mathrm{d}\boldsymbol{z}=\int_{\frac18 Q_0}H^m\,\zeta\,\mathbb{1}_{\mathcal{B}_{m,j}^c}\,\mathrm{d}\boldsymbol{z}\\&\leq\int H^m\,\zeta\,\mathbb{1}_{\mathcal{B}_{m,j}^c}\,\mathrm{d}\boldsymbol{z}=\int(\mathbf{S}^m-\mathcal{S}^\star(\cdot,\mathbf{D}\boldsymbol{u})):\mathbf{D}\boldsymbol{v}^m\,\zeta\,\mathbb{1}_{\mathcal{B}_{m,j}^c}\,\mathrm{d}\boldsymbol{z}\\&=-\int(\mathbf{G}_1^m-\mathbf{S}+\mathcal{S}^\star(\cdot,\mathbf{D}\boldsymbol{u})):\nabla\boldsymbol{v}^m\,\zeta\,\mathbb{1}_{\mathcal{B}_{m,j}^c}\,\mathrm{d}\boldsymbol{z}.\end{aligned} \tag{5.196}$$

Since $\mathbf{S} - \mathcal{S}^\star(\cdot, \mathbf{D}\boldsymbol{u}) \in \mathrm{L}^{q'}(Q)^{d\times d}$, we are in the position to use Lemma 2.14 (vii). Applying $\limsup_{m\to\infty}$ we find that

$$\begin{aligned}\limsup_{m\to\infty} \int_{\frac{1}{8}Q_0\setminus \mathcal{B}_{m,j}} H^m \,\mathrm{d}\boldsymbol{z} &\overset{(5.196)}{\leq} \limsup_{m\to\infty} \left| \int [\mathbf{G}_1^m - \mathbf{S} + \mathcal{S}^\star(\cdot, \mathbf{D}\boldsymbol{u})] : \nabla \boldsymbol{v}^m \,\zeta\, \mathbb{1}_{\mathcal{B}_{m,j}^c} \,\mathrm{d}\boldsymbol{z} \right| \\ &\leq c 2^{-j/q}.\end{aligned} \tag{5.197}$$

Using (5.195) and (5.197) in (5.194) yields

$$\begin{aligned}\limsup_{m\to\infty} \int_{\frac{1}{8}Q_0} [(\mathbf{S}^m - \mathcal{S}^\star(\cdot, \mathbf{D}\boldsymbol{u})) : (\mathbf{D}\boldsymbol{u}^m - \mathbf{D}\boldsymbol{u})]^{1/2} \,\mathrm{d}\boldsymbol{z} &= \limsup_{m\to\infty} \int_{\frac{1}{8}Q_0} (H^m)^{1/2} \,\mathrm{d}\boldsymbol{z} \\ &\overset{(5.194)}{\leq} c \limsup_{m\to\infty} |\mathcal{B}_{m,j}|^{\frac{1}{2}} + c \limsup_{m\to\infty} \left(\int_{\frac{1}{8}Q_0\setminus\mathcal{B}_{m,j}} H^m \,\mathrm{d}\boldsymbol{z} \right)^{\frac{1}{2}} \\ &\overset{(5.195),(5.197)}{\leq} c(2^{-\frac{j}{2}} + 2^{-\frac{j}{2q}}).\end{aligned} \tag{5.198}$$

Then taking $j \to \infty$ gives the claim and (5.193) 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 (5.190) we need L^1-convergence of H^m at least on suitable subdomains, which follows by Lemma 2.17, as in the steady situation in Lemma 4.8: Since $\{H^m\}_{m\in\mathbb{N}}$ is bounded in $\mathrm{L}^1(\frac{1}{8}Q_0)$ Lemma 2.17 shows 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

$$\int_{\frac{1}{8}Q_0\setminus E_i} H^m(\boldsymbol{z}) \,\mathrm{d}\boldsymbol{z} = \langle \mathbf{S}^m - \mathcal{S}^\star(\cdot, \mathbf{D}\boldsymbol{u}), \mathbf{D}\boldsymbol{u}^m - \mathbf{D}\boldsymbol{u} \rangle_{\frac{1}{8}Q_0\setminus E_i} \to 0, \quad \text{as } m \to \infty, \tag{5.199}$$

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 (5.174) 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 (5.172) we thus deduce that

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

This shows (5.190) for $\widetilde{Q} = \frac{1}{8}Q_0\setminus E_i$, and thus we find that $(\mathbf{D}\boldsymbol{u}(\boldsymbol{z}), \mathbf{S}(\boldsymbol{z})) \in \mathcal{A}(\boldsymbol{z})$ for a.e. $\boldsymbol{z} \in \frac{1}{8}Q_0\setminus E_i$. Since $|E_i| \to 0$, as $i \to \infty$, we have that $(\mathbf{D}\boldsymbol{u}(\boldsymbol{z}), \mathbf{S}(\boldsymbol{z})) \in \mathcal{A}(\boldsymbol{z})$ for a.e. $\boldsymbol{z} \in \frac{1}{8}Q_0$.

Finally let us consider a cover of Q consisting of (open) 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, compare, e.g., [DRW10]. 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 $(\mathbf{D}\boldsymbol{u}(\boldsymbol{z}), \mathbf{S}(\boldsymbol{z})) \in \mathcal{A}(\boldsymbol{z})$ for a.e. $\boldsymbol{z} \in Q$, which proves (5.185). □

Remark 5.11.

(i) For Lipschitz polytopal domains Theorem 5.2 is also a new existence result, since in [BGMS12] a Navier slip boundary condition and $\partial\Omega \in \mathrm{C}^{1,1}$ are assumed. This was done in order to reduce technicalities related to the pressure because at the time no solenoidal parabolic Lipschitz approximation was available.

(ii) As in the steady case one could show the corresponding result using the generalised Yosida graph approximation taking the limits $k, l, n \to \infty$ simultaneously.

(iii) In the regularised case the assumptions on q and the finite element setting do not differ, depending on whether $\mathbb{V}^n_{\mathrm{div}}$ is assumed to be exactly divergence-free or not. As in the steady case there are merely some simplifications in the argument if $\mathbb{V}^n_{\mathrm{div}}$ consists of exactly divergence-free functions.

5.3.2. Without Regularisation

Here we consider the approximation levels $k, l, n \in \mathbb{N}$, i.e., there is no regularisation term, and we take all three limits simultaneously. As in the steady case this can be achieved by use of the generalised Yosida graph approximation introduced and investigated in Subsection 3.4.3, since a Minty type convergence lemma is available, see Lemma 3.31.

Because of the unavailability of a discrete truncation the strongest restriction on q arises from the identification of the implicit constitutive relation, for which we assume that $q \in \left[\frac{3d+2}{d+2}, \infty\right)$. The restriction on q required for the compactness argument and the convergence in the equation are weaker. We still present the detailed assumptions to show what one can expect from a discrete truncation and what sort of difficulties the regularisation approach presented in the previous subsection circumvents.

For the compactness argument we can use either Simon's compactness lemma, provided that $q > \widetilde{q}_d$ (or $q > q_d$ if the spaces $\mathbb{V}^n_{\mathrm{div}}$ are exactly divergence-free). The restriction on q arises from the convective term and its numerical modification. Or one can use the Aubin–Lions compactness lemma for a larger range of q, more specifically for $q \geq \frac{2(d+1)}{d+2}$ (or $q > \frac{2d}{d+2}$ if the spaces $\mathbb{V}^n_{\mathrm{div}}$ are exactly divergence-free). However, this requires more restrictive assumptions on the finite element setting since stability of the L^2-projection P^n_{div} mapping to $\mathbb{V}^n_{\mathrm{div}}$ is required to obtain uniform estimates on the sequence of time derivatives.

Due to the convective term and its numerical modification one can show that the limiting equation is satisfies for $q > \frac{2(d+1)}{d+2}$ (or $q > \frac{2d}{d+2}$ if the spaces $\mathbb{V}^n_{\mathrm{div}}$ are exactly divergence-free), which corresponds to the situation in the steady case.

If $q \geq \frac{3d+2}{d+2}$ one has full admissibility of the solution as test function and hence an energy identity can be shown to hold for all times, which allows us to use a Minty type convergence lemma to identify the implicit constitutive relation of the limiting functions. For smaller q, the only hope to identify it is by means of a discrete truncation, which is, other than in the steady case, not available at present.

As before, we will first state the approximate problem, then state the convergence result in Theorem 5.2 and then provide the proof of it in the remaining part of the subsection.

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

$$\mathfrak{L}^{k,l,n}_i[\boldsymbol{u}; \boldsymbol{v}] := -\widetilde{b}(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{v}) - \left\langle \boldsymbol{\mathcal{S}}^k_i(\cdot, \mathbf{D}\boldsymbol{u}), \mathbf{D}\boldsymbol{v} \right\rangle_\Omega + \langle \boldsymbol{f}_i, \boldsymbol{v} \rangle_\Omega \,, \tag{5.200}$$

for $k, l, n \in \mathbb{N}$ and $i \in \{1, \ldots, l\}$, where $\boldsymbol{\mathcal{S}}^k_i$ and $\boldsymbol{f}_i$ as defined in (5.20) and $\widetilde{b}(\cdot, \cdot, \cdot)$ as recalled in (5.10).

Analogously to the previous subsection the pressure-free formulation of the approximate problem can be stated as follows.

Approximate Problem:

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

$$\boldsymbol{U}^{k,l,n}_0 = P^n_{\mathrm{div}} \boldsymbol{u}_0, \tag{5.201}$$

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

$$\left\langle \mathrm{d}_t \boldsymbol{U}^{k,l,n}_i, \boldsymbol{W} \right\rangle_\Omega = \mathfrak{L}^{k,l,n}_i[\boldsymbol{U}^{k,l,n}_i; \boldsymbol{W}] \qquad \text{for all } \boldsymbol{W} \in \mathbb{V}^n_{\mathrm{div}}, \tag{5.202}$$

where P^n_{div} is the L^2-projector onto $\mathbb{V}^n_{\mathrm{div}}$, defined in (2.80) and d_t is the backward temporal difference quotient as defined in (2.71).

As before this results in a fully implicit approximate problem.

Theorem 5.12 (Convergence in the Unsteady Case Without Regularisation)**.** *In addition to the assumptions of Definition 5.1 let $\{\mathcal{S}^k\}_{k\geq k_0}$ be the sequence of Carathéodory functions, defined in (3.59), which corresponds to the generalised Yosida graph approximation presented in Example 3.28. For the finite element approximation let Assumption 2.18 on the domain and on the family of simplicial partitions be satisfied. Let $\mathbb{V}^n$, $\mathbb{V}^n_{\mathrm{div}}$ and $\mathbb{Q}^n$ be as introduced in (2.30), (2.32)and (2.31) respectively, and assume that Assumptions 2.21 and 2.23 (i), (ii) are satisfied.*
Then, for all $k,l,n \in \mathbb{N}$ such that $k \geq k_0$, there exists a sequence $\{\boldsymbol{U}^{k,l,n}_i\}_{i\in\{0,\dots,l\}} \subset \mathbb{V}^n_{\mathrm{div}}$ solving (5.201), (5.202).
*If $q \in \left[\frac{3d+2}{d+2}, \infty\right)$, then there exists a weak solution $(\boldsymbol{u}, \mathbf{S})$ of **(PU)** according to Definition 5.1 and for the piecewise constant interpolant $\overline{\boldsymbol{U}}^{k,l,n}$ and the continuous, piecewise affine interpolant $\widetilde{\boldsymbol{U}}^{k,l,n}$ of $\{\boldsymbol{U}^{k,l,n}_i\}_{i\in\{0,\dots,l\}}$ and the piecewise constant interpolant $\overline{\mathcal{S}}^k(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^{k,l,n})$ of $\{\mathcal{S}^k_i(\cdot,\mathbf{D}\overline{\boldsymbol{U}}^{k,l,n})\}_{i\in\{1,\dots,l\}}$ as defined in (2.72) and (2.73), (up to not relabelled subsequences) one has that*

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

as $k,l,n \to \infty$ (combined), without restrictions on the relation between the discretisation parameters δ_l, h_n and $\frac{1}{k}$.

The rest of this section will deal with the proof of Theorem 5.12. Lemma 5.13 states the existence and a priori estimates, which follows as in Subsection 5.3.1. Lemma 5.14 covers the compactness arguments and convergence and finally in Lemma 5.15 it is proved that the pair of limiting functions $(\boldsymbol{u}, \mathbf{S})$ is a weak solution of **(PU)**, under the condition that $q \geq \frac{3d+2}{d+2}$. Note that we present the compactness and convergence results for a larger range of q under suitable additional assumptions in order to show where the restrictions arise from.

Recall, that if $\mathbb{V}^n_{\mathrm{div}} \subset \mathrm{W}^{1,\infty}_{0,\mathrm{div}}(\Omega)^d$, i.e., Assumption 2.27 is satisfied, then we have that $\widetilde{b}(\boldsymbol{W},\cdot,\cdot) = b(\boldsymbol{W},\cdot,\cdot)$, for $\boldsymbol{W} \in \mathbb{V}^n_{\mathrm{div}}$. This affects some of the following estimates and convergence results by the restrictions on the range of q, for which they hold.

Note that we will assume that $k \in \mathbb{N}$ such that $k \geq k_0$ without repeating this restriction.

For $t \in (0,T]$, $\boldsymbol{u} \in \mathbb{P}^l_0(0,T;\mathbb{V}^n)$ and $\boldsymbol{v} \in \mathbb{V}^n$ we introduce

$$
\mathfrak{L}^\kappa[\boldsymbol{u};\boldsymbol{v}](t) := -\widetilde{b}(\boldsymbol{u}(t,\cdot),\boldsymbol{u}(t,\cdot),\boldsymbol{v}) - \left\langle \overline{\mathcal{S}}^k(t,\cdot,\mathbf{D}\boldsymbol{u}(t,\cdot)),\mathbf{D}\boldsymbol{v}\right\rangle_\Omega + \left\langle \overline{\boldsymbol{f}}(t,\cdot),\boldsymbol{v}\right\rangle_\Omega, \tag{5.203}
$$

for $\kappa = (k,l,n) \in \mathbb{N}^3$. Recall that $\overline{\boldsymbol{f}} \in \mathbb{P}^l_0(0,T;\mathrm{W}^{-1,q'}(\Omega)^d)$ is the piecewise constant interpolant of $\{\boldsymbol{f}_i\}_{i\in\{1,\dots,l\}}$, see (2.72) in Subsection 2.2.3.1, and similarly, $\overline{\mathcal{S}}^k(t,\cdot,\cdot) = \mathcal{S}^k_i(\cdot,\cdot)$, for $t \in (t_{i-1},t_i]$, which is piecewise constant with respect to the variable $t \in (0,T]$.

Lemma 5.13 (Existence of Approximate Solutions and Estimates).
For each $\kappa := (k,l,n) \in \mathbb{N}^3$*, there exists a sequence* $\{\boldsymbol{U}_i^\kappa\}_{i\in\{0,\ldots,l\}} \subset \mathbb{V}^n_{\mathrm{div}}$*, which satisfies* (5.201)*,* (5.202)*. Furthermore, there exists a constant* $c > 0$ *such that for all* $\kappa = (k,l,n) \in \mathbb{N}^3$ *one has that*

$$\begin{aligned}\max_{j\in\{1,\ldots,l\}} \|\boldsymbol{U}_j^\kappa\|^2_{\mathrm{L}^2(\Omega)} &+ \sum_{j=1}^{l} \|\boldsymbol{U}_j^\kappa - \boldsymbol{U}_{j-1}^\kappa\|^2_{\mathrm{L}^2(\Omega)} \\ &+ \delta_l \sum_{j=1}^{l} \|\boldsymbol{U}_j^\kappa\|^q_{\mathrm{W}^{1,q}(\Omega)} + \sum_{j=1}^{l} \left\|\boldsymbol{\mathcal{S}}^k(\cdot,\cdot,\mathbf{D}\boldsymbol{U}_j^\kappa)\right\|^{q'}_{\mathrm{L}^{q'}(Q^j_{j-1})} \le c.\end{aligned} \tag{5.204}$$

For $\kappa := (k,l,n) \in \mathbb{N}^3$ *let the functions* $\overline{\boldsymbol{U}}^\kappa \in \mathbb{P}^l_0(0,T;\mathbb{V}^n_{\mathrm{div}})$ *and* $\widetilde{\boldsymbol{U}}^\kappa \in \mathbb{P}^l_1(0,T;\mathbb{V}^n_{\mathrm{div}})$ *be the piecewise constant and piecewise affine interpolants, of* $\{\boldsymbol{U}_i^\kappa\}_{i\in\{1,\ldots,l\}}$*, as defined in* (2.72) *and* (2.73)*, respectively. Then,* $\overline{\boldsymbol{U}}^\kappa$ *and* $\widetilde{\boldsymbol{U}}^\kappa$ *satisfy*

$$\left\langle \partial_t \widetilde{\boldsymbol{U}}^\kappa(t,\cdot), \boldsymbol{W}\right\rangle_\Omega = \mathfrak{L}^\kappa[\overline{\boldsymbol{U}}^\kappa; \boldsymbol{W}](t) \qquad \textit{for all } \boldsymbol{W} \in \mathbb{V}^n_{\mathrm{div}},\ \textit{for all } t \in (0,T], \tag{5.205}$$

$$\widetilde{\boldsymbol{U}}^\kappa(0,\cdot) = P^n_{\mathrm{div}}\boldsymbol{u}_0 \qquad \textit{in } \Omega, \tag{5.206}$$

for any $\kappa = (k,l,n) \in \mathbb{N}^3$*. For any* $0 \le s_0 < s \le T$ *and all* $\kappa = (k,l,n) \in \mathbb{N}^3$ *one has that*

$$\frac{1}{2}\left\|\widetilde{\boldsymbol{U}}^\kappa(s,\cdot)\right\|^2_{\mathrm{L}^2(\Omega)} + \left\langle \overline{\boldsymbol{\mathcal{S}}}^k(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^\kappa), \mathbf{D}\overline{\boldsymbol{U}}^\kappa\right\rangle_{Q^s_{s_0}} \le \langle \overline{\boldsymbol{f}}, \overline{\boldsymbol{U}}^\kappa\rangle_{Q^s_{s_0}} + \frac{1}{2}\left\|\widetilde{\boldsymbol{U}}^\kappa(s_0,\cdot)\right\|^2_{\mathrm{L}^2(\Omega)}. \tag{5.207}$$

Proof. Step 1: A priori estimates
The a priori estimates follow as in Step 1 in the proof of Lemma 5.3 noting that the regularising term was not used in any of the estimates and noting that the family $\{\boldsymbol{\mathcal{S}}^k\}_{k\in\mathbb{N}}$ satisfies Assumption 3.18 (σ2) by Lemma 3.30.

Step 2: Existence of $\{\boldsymbol{U}_i^\kappa\}_{i\in\{0,\ldots,l\}}$
Let $\kappa \in \mathbb{N}^3$ be fixed. Since $\boldsymbol{U}_0^\kappa = P^n_{\mathrm{div}}\boldsymbol{u}_0$ by (5.201), one again only has to show that for a given $\boldsymbol{U}_{i-1}^\kappa \in \mathbb{V}^n_{\mathrm{div}}$, there exists a $\boldsymbol{U}_i^\kappa \in \mathbb{V}^n_{\mathrm{div}}$ such that (5.202) is satisfied. Since the sequence $\{\boldsymbol{\mathcal{S}}^k\}_{k\ge k_0}$ satisfies Assumption 3.18 and the regularising term is not used, this follows as in Step 2 in the proof of Lemma 5.3 by means of a standard fixed point argument.

For $\kappa = (k,l,n) \in \mathbb{N}^3$ let $\{\boldsymbol{U}_i^\kappa\}_{i\in\{0,\ldots,l\}}$ be the sequence of solutions to (5.201), (5.202). Uniqueness is in general not guaranteed, so we choose one such sequence for each κ.

The fact that (5.205), (5.206) are satisfied by the piecewise constant interpolant $\overline{\boldsymbol{U}}^\kappa \in \mathbb{P}^l_0(0,T,\mathbb{V}^n_{\mathrm{div}})$ and the continuous, piecewise affine interpolant $\widetilde{\boldsymbol{U}}^\kappa \in \mathbb{P}^l_1(0,T,\mathbb{V}^n_{\mathrm{div}})$ of the sequence $\{\boldsymbol{U}_i^\kappa\}_{i\in\{0,\ldots,l\}}$ follows as in Step 1 in the proof of Lemma 5.4 and the energy inequality (5.207) can be derived with the same arguments as in Steps 1 in the proof of Lemma 5.6. □

For $q \in \left[\frac{3d}{d+2}, \infty\right)$ let us recall

$$\nu := q\left(\frac{q(d+2)}{d} - 2\right), \tag{5.208}$$

see (2.13). Recall also that $\nu > 2$, provided that $q > q_d$, and $\nu > q'$, provided that $q > \widetilde{q}_d$, with q_d and $\widetilde{q}_d$ as in (2.14) and (2.15), respectively. Recall that $q_d < \widetilde{q}_d < \frac{3d+2}{d+2}$, see (5.13). Since for the identification of the implicit constitutive relation we will assume that $q \ge \frac{3d+2}{d+2}$, distinguishing between these cases does affect Theorem 5.12. However, it helps to understand where the restrictions on q arise from.

Lemma 5.14 (Convergence $k,l,n\to\infty$)**.**
For $\kappa:=(k,l,n)\in\mathbb{N}^3$ let the functions $\overline{\boldsymbol{U}}^\kappa\in\mathbb{P}_0^l(0,T;\mathbb{V}_{\mathrm{div}}^n)$ and $\widetilde{\boldsymbol{U}}^\kappa\in\mathbb{P}_1^l(0,T;\mathbb{V}_{\mathrm{div}}^n)$ be as in Lemma 5.13, and in addition to the assumptions in Theorem 5.12 assume one of the following:

(I) Assume that $q\in(\widetilde{q}_d,\infty)$ (and $q\in(q_d,\infty)$ if the elements of the spaces $\mathbb{V}_{\mathrm{div}}^n$ are exactly divergence-free, i.e., Assumption 2.27 is satisfied).

(II) Assume that $q\in\left[\frac{2(d+1)}{d+2},\infty\right)$ (and $q\in\left(\frac{2d}{d+2},\infty\right)$ if the spaces $\mathbb{V}_{\mathrm{div}}^n$ satisfy Assumption 2.27). In addition to before, let Assumption 2.19 on the family of simplicial partitions and Assumption 2.23 (iia) be satisfied. If $d=3$ and $q\in\left(\frac{6}{5},\left(\frac{6}{5}\right)^2\right)$, assume that $\mathbb{V}_2^n\subset\mathbb{V}^n$, i.e., $\widehat{\mathcal{P}}_2^2\subset\widehat{\mathbb{P}}_{\mathbb{V}}$.

Then, there exists a $\boldsymbol{u}\in\mathrm{L}^\infty(0,T;\mathrm{L}_{\mathrm{div}}^2(\Omega)^d)\cap\mathrm{L}^q(0,T;\mathrm{W}_{0,\mathrm{div}}^{1,q}(\Omega)^d)$, an $\mathbf{S}\in\mathrm{L}^{q'}(Q)^{d\times d}$ and subsequences such that

$$\widetilde{\boldsymbol{U}}^{k,l,n}\to\boldsymbol{u}\qquad\text{strongly in } \mathrm{L}^p(0,T;\mathrm{L}^2(\Omega)^d)\text{ for all } p\in[1,\infty),\tag{5.209}$$

$$\widetilde{\boldsymbol{U}}^{k,l,n}(s,\cdot)\to\boldsymbol{u}(s,\cdot)\qquad\text{strongly in } \mathrm{L}^2(\Omega)^d\text{ for a.e. } s\in(0,T),\tag{5.210}$$

$$\widetilde{\boldsymbol{U}}^{k,l,n}(0,\cdot)\to\boldsymbol{u}_0\qquad\text{strongly in } \mathrm{L}^2(\Omega)^d,\tag{5.211}$$

$$\widetilde{\boldsymbol{U}}^{k,l,n},\overline{\boldsymbol{U}}^{k,l,n}\overset{*}{\rightharpoonup}\boldsymbol{u}\qquad\text{weakly* in } \mathrm{L}^\infty(0,T;\mathrm{L}^2(\Omega)^d),\tag{5.212}$$

$$\overline{\boldsymbol{U}}^{k,l,n}\to\boldsymbol{u}\qquad\text{strongly in } \mathrm{L}^p(0,T;\mathrm{L}^2(\Omega)^d)\cap\mathrm{L}^r(Q)^d\tag{5.213}$$

$$\text{for all } p\in[1,\infty)\text{ and all } r\in[1,\tfrac{q(d+2)}{d}),$$

$$\overline{\boldsymbol{U}}^{k,l,n}(s,\cdot)\to\boldsymbol{u}(s,\cdot)\qquad\text{strongly in } \mathrm{L}^2(\Omega)^d\text{ for a.e. } s\in(0,T),\tag{5.214}$$

$$\overline{\boldsymbol{U}}^{k,l,n}\rightharpoonup\boldsymbol{u}\qquad\text{weakly in } \mathrm{L}^q(0,T;\mathrm{W}_0^{1,q}(\Omega)^d)\cap\mathrm{L}^{\frac{q(d+2)}{d}}(Q)^d,\tag{5.215}$$

$$\overline{\mathcal{S}}^k(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^{k,l,n})\rightharpoonup\mathbf{S}\qquad\text{weakly in } \mathrm{L}^{q'}(Q)^{d\times d},\tag{5.216}$$

$$\mathcal{S}^k(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^{k,l,n})\rightharpoonup\mathbf{S}\qquad\text{weakly in } \mathrm{L}^{q'}(Q)^{d\times d},\tag{5.217}$$

as $k,l,n\to\infty$.

Proof. Step 1: General estimates
Since the arguments are the same as in Step 3 in the proof of Lemma 5.6 let us only outline the proofs of the following estimates. By the definition of the piecewise constant interpolant, the discrete estimates in (5.204) and the parabolic interpolation in Corollary 2.5 one can show that

$$\|\overline{\boldsymbol{U}}^\kappa\|_{\mathrm{L}^\infty(0,T;\mathrm{L}^2(\Omega))}+\|\overline{\boldsymbol{U}}^\kappa\|_{\mathrm{L}^q(0,T;\mathrm{W}^{1,q}(\Omega))}^q+\|\overline{\boldsymbol{U}}^\kappa\|_{\mathrm{L}^{\frac{q(d+2)}{d}}(Q)}\le c\quad\text{for all }\kappa\in\mathbb{N}^3,\tag{5.218}$$

since $q\ge\frac{2d}{d+2}$. Furthermore, by the definition of the continuous, piecewise affine interpolant $\widetilde{\boldsymbol{U}}^\kappa$, equation 5.206, the L^2-stability of P_{div}^n and the discrete estimate (5.204) one has that

$$\left\|\widetilde{\boldsymbol{U}}^\kappa\right\|_{\mathrm{L}^\infty(0,T;\mathrm{L}^2(\Omega))}\le c\qquad\text{for all }\kappa\in\mathbb{N}^3.\tag{5.219}$$

By the fact that $\partial_t\widetilde{\boldsymbol{U}}^\kappa(t,\cdot)=\mathrm{d}_t\boldsymbol{U}_i^\kappa$, for $t\in(t_{i-1},t_i]$, $i\in\{1,\dots,l\}$, and by (5.204) one also obtains that

$$\delta_l\left\|\partial_t\widetilde{\boldsymbol{U}}^\kappa\right\|_{\mathrm{L}^2(Q)}^2\le c\quad\text{for all }\kappa\in\mathbb{N}^3.\tag{5.220}$$

By the definition of $\overline{\mathbf{S}}^k(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^\kappa)$, the discrete estimate (5.204), the fact that $\mathbf{S}^k$ satisfies Assumption 3.18 ($\sigma2$) and (5.218) it follows that

$$\left\|\overline{\mathbf{S}}^k(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^\kappa)\right\|_{\mathrm{L}^{q'}(Q)}^{q'} + \left\|\mathbf{S}^k(\cdot,\mathbf{D}\overline{\boldsymbol{U}}^\kappa)\right\|_{\mathrm{L}^{q'}(Q)}^{q'} \le c \quad \text{for all } \kappa\in\mathbb{N}^3. \tag{5.221}$$

As before, for the compactness argument we consider $\widehat{\boldsymbol{U}}^\kappa \in \mathrm{C}([0,T];\mathbb{V}^n_{\mathrm{div}})$ defined by

$$\widehat{\boldsymbol{U}}^\kappa(t,\cdot) := \begin{cases} \widetilde{\boldsymbol{U}}^\kappa(t,\cdot) & \text{if } t\in(\delta_l,T],\\ \overline{\boldsymbol{U}}^\kappa(t,\cdot) = \boldsymbol{U}^\kappa_1(\cdot) & \text{if } t\in[0,\delta_l], \end{cases} \tag{5.222}$$

which is constant on $[0,\delta_l]$. As in (5.95) for any $r\in[1,\infty]$ and a normed space X we have

$$\left\|\widehat{\boldsymbol{U}}^\kappa\right\|_{\mathrm{L}^r(0,T;\mathrm{X})} \le c\left\|\overline{\boldsymbol{U}}^\kappa\right\|_{\mathrm{L}^r(0,T;\mathrm{X})} \quad \text{for any } \kappa\in\mathbb{N}^3. \tag{5.223}$$

With this and the estimate (5.218) we obtain that

$$\left\|\widehat{\boldsymbol{U}}^\kappa\right\|_{\mathrm{L}^\infty(0,T;\mathrm{L}^2(\Omega))} + \left\|\widehat{\boldsymbol{U}}^\kappa\right\|_{\mathrm{L}^q(0,T;\mathrm{W}^{1,q}(\Omega))} + \left\|\widehat{\boldsymbol{U}}^\kappa\right\|_{\mathrm{L}^{\frac{q(d+2)}{d}}(Q)} \le c \quad \text{for all } \kappa\in\mathbb{N}^3. \tag{5.224}$$

Step 2: Estimates and compactness in case (I)
Here we apply Simon's compactness lemma, and the interpolation result required for this is the reason for the stronger restriction of q in (I). By the assumptions on q we have in particular that $q\ge\frac{3d}{d+2}$, see (2.18). Hence, by the parabolic interpolation in Corollary 2.6 and the estimate (5.218) we obtain

$$\left\|\overline{\boldsymbol{U}}^\kappa\right\|_{\mathrm{L}^\nu(0,T;\mathrm{L}^{2q'}(\Omega))} \le c\left\|\nabla\overline{\boldsymbol{U}}^\kappa\right\|_{\mathrm{L}^q(Q)}^{1/q}\left\|\overline{\boldsymbol{U}}^\kappa\right\|_{\mathrm{L}^\infty(0,T;\mathrm{L}^2(\Omega))}^{1/q'} \le c \tag{5.225}$$

for any $\kappa\in\mathbb{N}^3$, with ν as in (5.208). With property (5.223) on $\widehat{\boldsymbol{U}}^\kappa$ this implies that

$$\left\|\widehat{\boldsymbol{U}}^\kappa\right\|_{\mathrm{L}^\nu(0,T;\mathrm{L}^{2q'}(\Omega))} \le c \quad \text{for all } \kappa\in\mathbb{N}^3. \tag{5.226}$$

For $\boldsymbol{v}\in\mathrm{L}^q(0,T;\mathrm{W}^{1,q}_0(\Omega)^d)\cap\mathrm{L}^{\frac{q(d+2)}{d}}(Q)^d$ we also estimate $\mathfrak{L}^\kappa[\overline{\boldsymbol{U}}^\kappa;\boldsymbol{v}](t)$, as defined in (5.203), which differs from the estimate in the regularised case: in case $\mathbb{V}^n_{\mathrm{div}}$ is exactly divergence-free the convective term $b(\overline{\boldsymbol{U}}^\kappa,\cdot,\cdot)$ and its numerical modification $\widetilde{b}(\overline{\boldsymbol{U}}^\kappa,\cdot,\cdot)$ coincide, so we only require estimate (5.14), which holds provided that $q\ge q_d$. For the following let us denote

$$\check{q} := \max\left(\left(\tfrac{\nu}{2}\right)',q\right). \tag{5.227}$$

If $\mathbb{V}^n_{\mathrm{div}}$ is discretely divergence-free, we also apply estimate (5.15), with $r\in(1,\infty]$ such that $\frac{1}{\nu}+\frac{1}{r}+\frac{1}{q}=1$, which is available, provided that $q\ge\widetilde{q}_d$. With these estimates, duality of norms, Hölder's and Poincaré's inequalities and the estimates in (5.218), (5.221) and (5.21) we obtain

$$\begin{aligned}\int_a^b \mathfrak{L}^\kappa[\overline{\boldsymbol{U}}^\kappa;\boldsymbol{v}](t)\,\mathrm{d}t &= \left\langle\widetilde{b}(\overline{\boldsymbol{U}}^\kappa,\overline{\boldsymbol{U}}^\kappa,\boldsymbol{v})\right\rangle_{(a,b)} - \left\langle\overline{\mathbf{S}}^k(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^\kappa),\mathbf{D}\boldsymbol{v}\right\rangle_{Q_a^b} + \left\langle\overline{\boldsymbol{f}},\boldsymbol{v}\right\rangle_{Q_a^b}\\ &\le c\left\|\overline{\boldsymbol{U}}^\kappa\right\|_{\mathrm{L}^\nu(a,b;\mathrm{L}^{2q'}(\Omega))}^2\left\|\nabla\boldsymbol{v}\right\|_{\mathrm{L}^{(\nu/2)'}(a,b;\mathrm{L}^q(\Omega))}\\ &\quad + c\left\|\overline{\boldsymbol{U}}^\kappa\right\|_{\mathrm{L}^\nu(a,b;\mathrm{L}^{2q'}(\Omega))}\left\|\nabla\overline{\boldsymbol{U}}^\kappa\right\|_{\mathrm{L}^q(Q_a^b)}\left\|\boldsymbol{v}\right\|_{\mathrm{L}^r(a,b;\mathrm{L}^{2q'}(\Omega))}\end{aligned} \tag{5.228}$$

$$\begin{aligned}&+\left\|\overline{\mathbf{S}}^k(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^\kappa)\right\|_{\mathrm{L}^{q'}(Q_a^b)}\|\mathbf{D}\boldsymbol{v}\|_{\mathrm{L}^q(Q_a^b)}\\&+\left\|\overline{\boldsymbol{f}}\right\|_{\mathrm{L}^{q'}(a,b;\mathrm{W}^{-1,q'}(\Omega))}\|\boldsymbol{v}\|_{\mathrm{L}^q(a,b;\mathrm{W}^{1,q}(\Omega))}\\&\le c\left(\|\nabla\boldsymbol{v}\|_{\mathrm{L}^{\check{q}}(a,b;\mathrm{L}^q(\Omega))}+\|\boldsymbol{v}\|_{\mathrm{L}^r(a,b;\mathrm{L}^{2q'}(\Omega))}\right),\end{aligned} \tag{5.229}$$

for $0 \le a < b \le T$, for any $\kappa \in \mathbb{N}^3$, where we have used that $\check{q} \ge q$. Let us note that the second term on the right-hand side stems from the modification of the numerical convective term, which is required if $\mathbb{V}^n_{\mathrm{div}}$ is only discretely divergence-free. In case $\mathbb{V}^n_{\mathrm{div}}$ is exactly divergence-free we have that $q > q_d$, which implies that $\check{q} < \infty$, and in case $\mathbb{V}^n_{\mathrm{div}}$ is only discretely divergence-free, we additionally assume that $q > \widetilde{q}_d$, which means that $r < \infty$, see (5.14) and (5.15).

We wish to apply the compactness result due to Simon stated in Lemma 2.8 and used before in the regularised case in Lemma 5.6, to the sequence $\{\widehat{\boldsymbol{U}}^{k,l,n}\}_{k,l,n\in\mathbb{N}}$, with $\mathrm{X} = \mathrm{W}^{1,q}(\Omega)^d$, $\mathrm{B} = \mathrm{L}^2(\Omega)^d$ and $p = 2$. As in Step 4 in the proof of Lemma 5.6 let us show that

$$\int_0^{T-\varepsilon}\left\|\widehat{\boldsymbol{U}}^\kappa(s+\varepsilon,\cdot)-\widehat{\boldsymbol{U}}^\kappa(s,\cdot)\right\|^2_{\mathrm{L}^2(\Omega)}\mathrm{d}s \to 0, \quad \text{as } \varepsilon\to 0, \text{ uniformly for } \kappa\in\mathbb{N}^3. \tag{5.230}$$

The proof proceeds analogously as before, where we use the estimate (5.229) instead of (5.99): for $\boldsymbol{W} \in \mathbb{V}^n_{\mathrm{div}}$, $s \in (0,T)$ and $\varepsilon > 0$ such that $s+\varepsilon < T$ one can show that, by the definition of $\widehat{\boldsymbol{U}}^\kappa$, the equation (5.205) for a.e. $t \in (0,T)$ and applying (5.229) one has that

$$\begin{aligned}\left\langle\widehat{\boldsymbol{U}}^\kappa(s+\varepsilon,\cdot)-\widehat{\boldsymbol{U}}^\kappa(s,\cdot),\boldsymbol{W}\right\rangle_\Omega &= \int_{\max(s,\delta_l)}^{s+\varepsilon}\left\langle\partial_t\widetilde{\boldsymbol{U}}^\kappa(t,\cdot),\boldsymbol{W}\right\rangle_\Omega \mathrm{d}t,\\&=\int_{\max(s,\delta_l)}^{s+\varepsilon}\mathfrak{L}^\kappa[\overline{\boldsymbol{U}}^\kappa;\boldsymbol{W}](t)\,\mathrm{d}t\\&\overset{(5.229)}{\le} c\left(\|\nabla\boldsymbol{W}\|_{\mathrm{L}^{\check{q}}(s,s+\varepsilon;\mathrm{L}^q(\Omega))}+\|\boldsymbol{W}\|_{\mathrm{L}^r(s,s+\varepsilon;\mathrm{L}^{2q'}(\Omega))}\right)\\&= c\left(\varepsilon^{1/\check{q}}\|\boldsymbol{W}\|_{\mathrm{W}^{1,q}(\Omega)}+\varepsilon^{1/r}\|\boldsymbol{W}\|_{\mathrm{L}^{2q'}(\Omega)}\right),\end{aligned} \tag{5.231}$$

since $\boldsymbol{W}$ is constant in time and the length of the time interval is bounded by ε. Choosing $\boldsymbol{W} = \widehat{\boldsymbol{U}}^\kappa(s+\varepsilon,\cdot) - \widehat{\boldsymbol{U}}^\kappa(s,\cdot) \in \mathbb{V}^n_{\mathrm{div}}$, which is piecewise constant in time, yields

$$\begin{aligned}\left\|\widehat{\boldsymbol{U}}^\kappa(s+\varepsilon,\cdot)-\widehat{\boldsymbol{U}}^\kappa(s,\cdot)\right\|^2_{\mathrm{L}^2(\Omega)} \le c\Big(&\varepsilon^{1/\check{q}}\left\|\widehat{\boldsymbol{U}}^\kappa(s+\varepsilon,\cdot)-\widehat{\boldsymbol{U}}^\kappa(s,\cdot)\right\|_{\mathrm{W}^{1,q}(\Omega)}\\&+\varepsilon^{1/r}\left\|\widehat{\boldsymbol{U}}^\kappa(s+\varepsilon,\cdot)-\widehat{\boldsymbol{U}}^\kappa(s,\cdot)\right\|_{\mathrm{L}^{2q'}(\Omega)}\Big).\end{aligned} \tag{5.232}$$

Integrating in $(0, T-\varepsilon)$, using the triangle inequality, Hölder's inequality and the estimates (5.224) and (5.226) show as in (5.104) that

$$\begin{aligned}&\int_0^{T-\varepsilon}\left\|\widehat{\boldsymbol{U}}^\kappa(s+\varepsilon,\cdot)-\widehat{\boldsymbol{U}}^\kappa(s,\cdot)\right\|^2_{\mathrm{L}^2(\Omega)}\mathrm{d}s\\&\overset{(5.232)}{\le}\left(\varepsilon^{1/\check{q}}+\varepsilon^{1/r}\right)\left(\left\|\widehat{\boldsymbol{U}}^\kappa\right\|_{\mathrm{L}^1(0,T;\mathrm{W}^{1,q}(\Omega))}+\left\|\widehat{\boldsymbol{U}}^\kappa\right\|_{\mathrm{L}^1(0,T;\mathrm{L}^{2q'}(\Omega))}\right)\\&\overset{(5.224),(5.226)}{\le} c\left(\varepsilon^{1/\check{q}}+\varepsilon^{1/r}\right),\end{aligned} \tag{5.233}$$

which vanishes, as $\varepsilon \to 0$ uniformly in $\kappa \in \mathbb{N}^3$, since $\check{q}, r \in (1,\infty)$. This proves (5.230).

Let us choose subsequences such that $k \to \infty \Leftrightarrow l \to \infty \Leftrightarrow n \to \infty$ and set $k = k_n$ and $l = l_n$. Note that we do not require any relation between the corresponding parameters $\frac{1}{k}$, h_n and δ_l. Furthermore, for ease of readability we denote, for $n \in \mathbb{N}$,

$$\mathfrak{S}^n := \mathfrak{S}^{k_n},\ \overline{\mathfrak{S}}^n := \overline{\mathfrak{S}}^{k_n},\ \mathcal{A}^n := \mathcal{A}^{k_n}, \text{ and } F^n := F^{k_n,l_n,n}, \tag{5.234}$$

for $F \in \{\overline{\boldsymbol{U}}, \widetilde{\boldsymbol{U}}, \widehat{U}\}$.

By (5.224) we have that $\{\widehat{\boldsymbol{U}}^n\}_{n\in\mathbb{N}}$ is bounded in $\mathrm{L}^2(Q)^d$ and $\mathrm{L}^1(0,T;\mathrm{W}^{1,q}(\Omega)^d)$. Since we have in particular that $q > \frac{2d}{d+2}$, the embedding $\mathrm{W}^{1,q}(\Omega) \hookrightarrow\hookrightarrow \mathrm{L}^2(\Omega)$ is compact and with (5.230) all the assumptions in Lemma 2.8 are satisfied for $\mathrm{X} = \mathrm{W}^{1,q}(\Omega)^d$, $\mathrm{B} = \mathrm{L}^2(\Omega)^d$ and $p = 2$. Hence, there exists a $\boldsymbol{u} \in \mathrm{L}^2(Q)^d$ and a subsequence such that

$$\widehat{\boldsymbol{U}}^n \to \boldsymbol{u} \quad \text{strongly in } \mathrm{L}^2(Q)^d, \quad \text{as } n \to \infty. \tag{5.235}$$

Step 3: Estimates and compactness in case (II)
For $q \le \widetilde{q}_d$ (or $q \le q_d$) the estimates above are not strong enough to apply Simon's Lemma. Instead, for the larger range of $q > \frac{2(d+1)}{d+2}$ (or $q > \frac{2d}{d+2}$) we aim to apply the Aubin–Lions compactness lemma, see Lemma 2.7. For this purpose we have to show uniform bounds on the time derivatives $\partial_t \widehat{\boldsymbol{U}}^\kappa$ in suitable negative Sobolev spaces. This can be achieved using the equation (5.205) and a stability result on the L^2-projection P^n_{div} to $\mathbb{V}^n_{\mathrm{div}}$, see Lemma 2.32, which is the reason for the additional restrictions on the finite element setting in (II).

For any $\boldsymbol{w} \in \mathrm{C}^\infty_{0,\mathrm{div}}(\Omega)^d$ we have that $P^n_{\mathrm{div}}(\boldsymbol{w}) \in \mathbb{V}^n_{\mathrm{div}}$, since P^n_{div} defined in (2.80) maps to $\mathbb{V}^n_{\mathrm{div}}$, so it is an admissible test function in (5.205). Since $\widetilde{\boldsymbol{U}}^\kappa \in \mathbb{P}^l_1(0,T;\mathbb{V}^n_{\mathrm{div}})$ we have that $\partial_t \widetilde{\boldsymbol{U}}^\kappa(t,\cdot) \in \mathbb{V}^n_{\mathrm{div}}$ for all $t \in (0,T)$, and hence with the properties of the L^2-projection P^n_{div} in (2.80) and with equation (5.205) we have that

$$\left\langle \partial_t \widetilde{\boldsymbol{U}}^\kappa(t,\cdot), \boldsymbol{w} \right\rangle_\Omega = \left\langle \partial_t \widetilde{\boldsymbol{U}}^\kappa(t,\cdot), P^n_{\mathrm{div}}(\boldsymbol{w}) \right\rangle_\Omega = \mathfrak{L}^\kappa[\overline{\boldsymbol{U}}^\kappa; P^n_{\mathrm{div}}(\boldsymbol{w})](t). \tag{5.236}$$

We have to estimate the right-hand side in order to obtain uniform bounds on $\{\partial_t \widetilde{\boldsymbol{U}}^\kappa\}_{\kappa\in\mathbb{N}^3}$.

First let us estimate $\mathfrak{L}^\kappa[\boldsymbol{U};\boldsymbol{V}](t)$, for $\boldsymbol{U} \in \mathrm{L}^{\frac{q(d+2)}{d}}(Q)^d$, $\boldsymbol{V} \in \mathrm{W}^{1,\hat{q}}(\Omega)^d$, with $\hat{q}$ as in (5.5), pointwise in time. The pointwise version of the estimate (5.16) on the convective term reads

$$|\langle \boldsymbol{u}(t,\cdot) \otimes \boldsymbol{u}(t,\cdot), \nabla \boldsymbol{v} \rangle_\Omega| \le \|\boldsymbol{u}(t,\cdot)\|^2_{\mathrm{L}^{\frac{q(d+2)}{d}}(\Omega)} \|\nabla \boldsymbol{v}\|_{\mathrm{L}^{\hat{q}}(\Omega)}, \tag{5.237}$$

and holds, if $q \ge \frac{2d}{d+2}$. If $\mathbb{V}^n_{\mathrm{div}}$ is only discretely divergence-free, the modification of the convective term is present and the pointwise in time version of (5.17) can be stated as

$$\begin{aligned} |\langle \boldsymbol{u}(t,\cdot) \otimes \boldsymbol{v}, \nabla \boldsymbol{u}(t,\cdot) \rangle_\Omega| &\le \|\boldsymbol{u}(t,\cdot)\|_{\mathrm{L}^{\frac{q(d+2)}{d}}(\Omega)} \|\nabla \boldsymbol{u}(t,\cdot)\|_{\mathrm{L}^q(\Omega)} \|\boldsymbol{v}\|_{\mathrm{L}^s(\Omega)} \\ &\le c\, \|\boldsymbol{u}(t,\cdot)\|_{\mathrm{L}^{\frac{q(d+2)}{d}}(\Omega)} \|\nabla \boldsymbol{u}(t,\cdot)\|_{\mathrm{L}^q(\Omega)} \|\boldsymbol{v}\|_{\mathrm{W}^{1,\hat{q}}(\Omega)}, \end{aligned} \tag{5.238}$$

if $q \ge \frac{2(d+1)}{d+2}$. Here we have also used that $\mathrm{W}^{1,\hat{q}}(\Omega) \hookrightarrow \mathrm{L}^s(\Omega)$, with $\hat{q}$ as in (5.5) and $s \in (1,\infty]$ such that $\frac{d}{q(d+2)} + \frac{1}{q} + \frac{1}{s} = 1$. Since $q \ge \frac{2(d+1)}{d+2}$ (and $q \ge \frac{2d}{d+2}$ if the spaces $\mathbb{V}^n_{\mathrm{div}}$ are exactly divergence-free), we use (5.237) on the convective term and (5.238) on the modification of the numerical convective term, as well as duality of norms and Hölder's inequality to obtain

$$
\begin{aligned}
|\mathfrak{L}^\kappa[\boldsymbol{U};\boldsymbol{V}](t)| &\leq c\,|\langle \boldsymbol{U}(t,\cdot)\otimes \boldsymbol{U}(t,\cdot), \nabla \boldsymbol{V}\rangle_\Omega| + c\,|\langle \boldsymbol{U}(t,\cdot)\otimes \boldsymbol{V}, \nabla \boldsymbol{U}(t,\cdot)\rangle_\Omega| \\
&\quad + \left|\left\langle \overline{\boldsymbol{\mathfrak{S}}}^k(t,\cdot,\mathbf{D}\boldsymbol{U}(t,\cdot)), \mathbf{D}\boldsymbol{V}\right\rangle_\Omega\right| + |\langle \overline{\boldsymbol{f}}, \boldsymbol{V}\rangle_\Omega| \\
&\leq c\Bigg(\|\boldsymbol{U}(t,\cdot)\|^2_{\mathrm{L}^{\frac{q(d+2)}{d}}(\Omega)} + \|\boldsymbol{U}(t,\cdot)\|_{\mathrm{L}^{\frac{q(d+2)}{d}}(\Omega)} \|\nabla \boldsymbol{U}(t,\cdot)\|_{\mathrm{L}^q(\Omega)} \\
&\quad + \left\|\overline{\boldsymbol{\mathfrak{S}}}^k(t,\cdot,\mathbf{D}\boldsymbol{U}(t,\cdot))\right\|_{\mathrm{L}^{q'}(\Omega)} + \|\overline{\boldsymbol{f}}(t,\cdot)\|_{\mathrm{W}^{-1,q'}(\Omega)}\Bigg) \|\boldsymbol{V}\|_{\mathrm{W}^{1,\hat{q}}(\Omega)} \\
&=: \mathfrak{I}^\kappa[\boldsymbol{U}](t)\,\|\boldsymbol{V}\|_{\mathrm{W}^{1,\hat{q}}(\Omega)}\,,
\end{aligned}
\tag{5.239}
$$

where we have also used that $\hat{q} \geq q$.

Let us investigate the integrability of $\mathfrak{I}^\kappa[\overline{\boldsymbol{U}}^\kappa]$. The last two terms of $\mathfrak{I}^\kappa[\overline{\boldsymbol{U}}^\kappa]$ have $\mathrm{L}^{q'}((0,T))$ integrability, the first term has $\mathrm{L}^{\frac{q(d+2)}{2d}}((0,T))$ integrability, and we have that $q \geq \frac{2d}{d+2}$. For the second term, which stems from the modification of the convective term and is present only if $\mathbb{V}^n_{\mathrm{div}}$ is discretely divergence-free, by Hölder's inequality with $\frac{d+2}{2(d+1)} + \frac{d}{2(d+1)} = 1$ we obtain

$$
\begin{aligned}
&\int_0^T \left(\|\boldsymbol{U}(t,\cdot)\|_{\mathrm{L}^{\frac{q(d+2)}{d}}(\Omega)} \|\nabla \boldsymbol{U}(t,\cdot)\|_{\mathrm{L}^q(\Omega)} \right)^{\frac{q(d+2)}{2(d+1)}} \mathrm{d}t \\
&\qquad\qquad \leq \left(\|\boldsymbol{U}\|_{\mathrm{L}^{\frac{q(d+2)}{d}}(Q)} \|\nabla \boldsymbol{U}\|_{\mathrm{L}^q(Q)} \right)^{\frac{q(d+2)}{2(d+1)}},
\end{aligned}
\tag{5.240}
$$

so we have $\mathrm{L}^{\frac{q(d+2)}{2(d+1)}}((0,T))$ integrability and $\frac{q(d+2)}{2(d+1)} \geq 1$, since $q \geq \frac{2(d+1)}{d+2}$ in that case. Setting

$$
\widetilde{\tau} := \min\left(\frac{q(d+2)}{2(d+1)}, \frac{q(d+2)}{2d}, q' \right) = \min\left(\frac{q(d+2)}{2(d+1)}, q' \right) \geq 1, \tag{5.241}
$$

provided that $q \geq \frac{2(d+1)}{d+2}$, it follows from these considerations, using also (5.218), (5.221) and (5.21) that

$$
\begin{aligned}
\left\|\mathfrak{I}^\kappa[\overline{\boldsymbol{U}}^\kappa]\right\|_{\mathrm{L}^{\widetilde{\tau}}(0,T)} &\overset{(5.240)}{\leq} c\Bigg(\left\|\overline{\boldsymbol{U}}^\kappa\right\|^2_{\mathrm{L}^{\frac{q(d+2)}{d}}(Q)} + \left\|\overline{\boldsymbol{U}}^\kappa\right\|_{\mathrm{L}^{\frac{q(d+2)}{d}}(Q)} \left\|\nabla \overline{\boldsymbol{U}}^\kappa\right\|_{\mathrm{L}^q(Q)} \\
&\qquad + \left\|\overline{\boldsymbol{\mathfrak{S}}}^k(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^\kappa)\right\|_{\mathrm{L}^{q'}(Q)} + \|\overline{\boldsymbol{f}}\|_{\mathrm{L}^{q'}(0,T;\mathrm{W}^{-1,q'}(\Omega))}\Bigg) \\
&\leq c \qquad \text{for all } \kappa \in \mathbb{N}^3,
\end{aligned}
\tag{5.242}
$$

in case the spaces $\mathbb{V}^n_{\mathrm{div}}$ are discretely divergence-free. With the same arguments, if the spaces $\mathbb{V}^n_{\mathrm{div}}$ are exactly divergence-free one has that

$$
\begin{aligned}
\|\mathfrak{I}^\kappa[\boldsymbol{U}]\|_{\mathrm{L}^{\hat{q}'}(0,T)} &\leq c\left(\left\|\overline{\boldsymbol{U}}^\kappa\right\|^2_{\mathrm{L}^{\frac{q(d+2)}{d}}(Q)} + \left\|\overline{\boldsymbol{\mathfrak{S}}}^k(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^\kappa)\right\|_{\mathrm{L}^{q'}(Q)} + \|\overline{\boldsymbol{f}}\|_{\mathrm{L}^{q'}(0,T;\mathrm{W}^{-1,q'}(\Omega))} \right) \\
&\leq c \qquad \text{for all } \kappa \in \mathbb{N}^3,
\end{aligned}
\tag{5.243}
$$

since $\hat{q}' = \min\left(\frac{q(d+2)}{2d}, q'\right) \geq 1$, if $q \geq \frac{2d}{d+2}$.

Now we use Lemma 2.32 concerning the stability of the L^2-projection P^n_{div} as defined in (2.80) in order to bound $\|P^n_{\mathrm{div}}\boldsymbol{w}\|_{\mathrm{W}^{1,\hat{q}}(\Omega)}$. For this we require that the family of simplicial partitions is quasiuniform as formulated in Assumption 2.19 and the fact that Assumption 2.23 (iia) is assumed. Note that for $d = 3$ we have that $q < \left(\frac{6}{5}\right)^2$ if and only if $\hat{q} > 6$. The assumption that in this case the maximal $r \in \mathbb{N}$ such that $\mathbb{V}_r^n \subset \mathbb{V}^n$ is ≥ 2 implies that the minimal β, such that $\mathrm{W}^{\beta,2}(\Omega) \hookrightarrow \mathrm{W}^{1,\hat{q}}(\Omega)$ satisfies $\beta \leq r + 1$. Hence, the assumptions of Lemma 2.32 are satisfied for $p = \hat{q}$. Let us denote $\mathrm{Y} := \mathrm{W}^{1,\hat{q}}_{0,\mathrm{div}}(\Omega)^d \cap \mathrm{W}^{\beta,2}(\Omega)^d \hookrightarrow \mathrm{L}^2(\Omega)^d$, which is reflexive as a closed subspace of a reflexive Banach space. Then, Lemma 2.32 shows that

$$\|P^n_{\mathrm{div}} \boldsymbol{w}\|_{\mathrm{W}^{1,\hat{q}}(\Omega)} \leq c\, \|\boldsymbol{w}\|_{\mathrm{Y}} \qquad \text{for all } n \in \mathbb{N}, \tag{5.244}$$

independently of $\boldsymbol{w} \in \mathrm{C}^\infty_{0,\mathrm{div}}(\Omega)^d \subset \mathrm{Y}$. Now we use (5.236), (5.239) and (5.244) to obtain

$$\begin{aligned} \left|\left\langle \partial_t \widetilde{\boldsymbol{U}}^\kappa(t,\cdot), \boldsymbol{w}\right\rangle_\Omega\right| &= \left|\mathfrak{L}^\kappa[\overline{\boldsymbol{U}}^\kappa; P^n_{\mathrm{div}}(\boldsymbol{w})](t)\right| \leq \mathfrak{I}^\kappa[\overline{\boldsymbol{U}}^\kappa](t)\, \|P^n_{\mathrm{div}}(\boldsymbol{w})\|_{\mathrm{W}^{1,\hat{q}}(\Omega)} \\ &\leq c\mathfrak{I}^\kappa[\overline{\boldsymbol{U}}^\kappa](t)\, \|\boldsymbol{w}\|_{\mathrm{Y}} \end{aligned} \tag{5.245}$$

for all $\kappa = (k,l,n) \in \mathbb{N}^3$. By the definition of the dual norm and the integrability of $t \mapsto \mathfrak{I}^\kappa[\overline{\boldsymbol{U}}^\kappa](t)$ according to (5.242) or (5.243), this shows that

$$\left\|\partial_t \widetilde{\boldsymbol{U}}^\kappa\right\|_{\mathrm{L}^1(0,T;\mathrm{Y}')} \leq c \qquad \text{for all } \kappa \in \mathbb{N}^3, \tag{5.246}$$

since $\widetilde{\tau} \geq 1$, if $q \geq \frac{2(d+1)}{d+2}$, and $\widehat{q}' \geq 1$, provided that $q \geq \frac{2d}{d+2}$ (if $\mathbb{V}^n_{\mathrm{div}}$ is exactly divergence-free). By the definition of $\widehat{\boldsymbol{U}}^\kappa$ in (5.222), it follows that

$$\left\|\partial_t \widehat{\boldsymbol{U}}^\kappa\right\|_{\mathrm{L}^1(0,T;\mathrm{Y}')} = \left\|\partial_t \widetilde{\boldsymbol{U}}^\kappa\right\|_{\mathrm{L}^1(\delta_l,T;\mathrm{Y}')} \leq \left\|\partial_t \widetilde{\boldsymbol{U}}^\kappa\right\|_{\mathrm{L}^1(0,T;\mathrm{Y}')} \leq c \qquad \text{for all } \kappa \in \mathbb{N}^3, \tag{5.247}$$

if $q \geq \frac{2(d+1)}{d+2}$ (and if $q \geq \frac{2d}{d+2}$ in case the elements of the spaces $\mathbb{V}^n_{\mathrm{div}}$ are exactly divergence-free).

As in Step 2 let us choose subsequences $(k_n, l_n, n) \in \mathbb{N}$, such that $k_n \to \infty$ and $l_n \to \infty$ as $n \to \infty$ and use the notation in (5.234).

Now we want to apply the Aubin–Lions–Simon compactness lemma, stated in Lemma 2.7, to $\{\widehat{\boldsymbol{U}}^n\}_{n\in\mathbb{N}}$, which is by (5.218) bounded in $\mathrm{L}^\infty(0,T;\mathrm{L}^2(\Omega)^d) \cap \mathrm{L}^q(0,T;\mathrm{W}^{1,q}_0(\Omega)^d)$ and by (5.247) the sequence $\{\partial_t \widehat{\boldsymbol{U}}^n\}_{n\in\mathbb{N}}$ is bounded in $\mathrm{L}^1(0,T;\mathrm{Y}')$. Since $q > \frac{2d}{d+2}$ we have that the embedding $\mathrm{W}^{1,q}(\Omega)^d \hookrightarrow\hookrightarrow \mathrm{L}^2(\Omega)^d$ is compact. Let us show that the embedding $\mathrm{L}^2(\Omega)^d \hookrightarrow \mathrm{Y}'$ is continuous. First note that the embedding $\mathrm{W}^{\beta,2}(\Omega)^d \hookrightarrow \mathrm{L}^2(\Omega)^d$ is dense, and thus it follows that

$$\mathrm{L}^2(\Omega)^d \hookrightarrow (\mathrm{W}^{\beta,2}(\Omega)^d)'. \tag{5.248}$$

Furthermore, we have that $(\mathrm{W}^{\beta,2}(\Omega)^d)' \hookrightarrow (\mathrm{W}^{\beta,2}(\Omega)^d)' + (\mathrm{W}^{1,\hat{q}}_{0,\mathrm{div}}(\Omega)^d)'$, by the observation that for normed spaces X, Y one has that

$$\|g\|_{X'+Z'} = \inf\{\|g_1\|_{X'} + \|g_2\|_{Z'} \,:\, g = g_1 + g_2\} \leq \|g\|_{X'}\,, \tag{5.249}$$

for any $g \in X'$. Taking both embeddings together, this yields for $\mathrm{Y} = \mathrm{W}^{\beta,2}(\Omega)^d \cap \mathrm{W}^{1,\hat{q}}_{0,\mathrm{div}}(\Omega)^d$ that

$$\mathrm{L}^2(\Omega)^d \hookrightarrow (\mathrm{W}^{\beta,2}(\Omega)^d)' \hookrightarrow (\mathrm{W}^{\beta,2}(\Omega)^d)' + (\mathrm{W}^{1,\hat{q}}_{0,\mathrm{div}}(\Omega)^d)' = \mathrm{Y}'. \tag{5.250}$$

Thus, we are in the position to apply Lemma 2.7 with $\mathrm{X} = \mathrm{W}^{1,q}(\Omega)^d$, $\mathrm{B} = \mathrm{L}^2(\Omega)^d$, $\mathrm{Z} = \mathrm{Y}'$

and $p = \infty$. Hence, there exists a (non-relabelled) subsequence and a $\boldsymbol{u} \in \mathrm{L}^2(Q)^d$ such that in particular

$$\widehat{\boldsymbol{U}}^n \to \boldsymbol{u} \quad \text{strongly in } \mathrm{L}^2(Q)^d, \quad \text{as } n \to \infty. \tag{5.251}$$

and we are in the same situation as in the end of Step 2.

Step 4: Convergence as $k, l, n \to \infty$
Starting from (5.235) and (5.251), respectively, the arguments for the convergence follow along the lines of Step 5 in the proof of Lemma 5.6, and hence we will be brief.

By the definition of $\widehat{\boldsymbol{U}}^n$ in (5.222), and the property (2.74) of the interpolants and estimate (5.220) as before one can show that $\widetilde{\boldsymbol{U}}^n - \widehat{\boldsymbol{U}}^n \to \mathbf{0}$ and $\overline{\boldsymbol{U}}^n - \widetilde{\boldsymbol{U}}^n \to \mathbf{0}$ strongly in $\mathrm{L}^2(Q)^d$, as $n \to \infty$, which implies with (5.235) and (5.251), respectively, that $\widetilde{\boldsymbol{U}}^n \to \boldsymbol{u}$ strongly in $\mathrm{L}^2(Q)^d$, and also $\overline{\boldsymbol{U}}^n \to \boldsymbol{u}$ strongly in $\mathrm{L}^2(Q)^d$, as $n \to \infty$. In particular, $t \mapsto \|\widetilde{\boldsymbol{U}}^n(t,\cdot) - \boldsymbol{u}(t,\cdot)\|_{\mathrm{L}^2(\Omega)}$ converges to zero strongly in $\mathrm{L}^2(0,T)$, as $n \to \infty$. Thus, there exists a subsequence such that $t \mapsto \|\widetilde{\boldsymbol{U}}^n(t,\cdot) - \boldsymbol{u}(t,\cdot)\|_{\mathrm{L}^2(\Omega)}$ converges to zero a.e. in $(0,T)$, as $n \to \infty$, which implies (5.210). The same argument holds for $\overline{\boldsymbol{U}}^n$, and so (5.214) is shown. By the strong convergence in $\mathrm{L}^2(Q)^d$ and boundedness of $\{\widetilde{\boldsymbol{U}}^n\}_{n\in\mathbb{N}}$ in $\mathrm{L}^\infty(0,T;\mathrm{L}^2(\Omega)^d)$ by (5.219) and boundedness of $\{\overline{\boldsymbol{U}}^n\}_{n\in\mathbb{N}}$ in $\mathrm{L}^\infty(0,T;\mathrm{L}^2(\Omega)^d) \cap \mathrm{L}^{\frac{q(d+2)}{d}}(Q)^d$ by (5.218) implies by interpolation that

$$\widetilde{\boldsymbol{U}}^n \to \boldsymbol{u} \quad \text{strongly in } \mathrm{L}^p(0,T;\mathrm{L}^2(\Omega)^d), \tag{5.252}$$

$$\overline{\boldsymbol{U}}^n \to \boldsymbol{u} \quad \text{strongly in } \mathrm{L}^p(0,T;\mathrm{L}^2(\Omega)^d) \cap \mathrm{L}^r(Q)^d, \tag{5.253}$$

as $n \to \infty$, for any $p \in [1,\infty)$ and any $r \in [1, \frac{q(d+2)}{d})$, which shows (5.209), (5.213).

By the uniform bounds in (5.218) and (5.219) and the Banach–Alaoglu theorem, up to subsequences, we have that

$$\widetilde{\boldsymbol{U}}^n, \overline{\boldsymbol{U}}^n \overset{*}{\rightharpoonup} \boldsymbol{u} \quad \text{weakly* in } \mathrm{L}^\infty(0,T;\mathrm{L}^2(\Omega)^d), \tag{5.254}$$

$$\overline{\boldsymbol{U}}^n \rightharpoonup \boldsymbol{u} \quad \text{weakly in } \mathrm{L}^q(0,T;\mathrm{W}^{1,q}_0(\Omega)^d) \cap \mathrm{L}^{\frac{q(d+2)}{d}}(Q)^d, \tag{5.255}$$

as $n \to \infty$, which shows (5.212), (5.215). The argument that $\boldsymbol{u}$ is divergence-free follow as in Step 5 in the proof of Lemma 5.6. The fact that $\widetilde{\boldsymbol{U}}^n(0,\cdot) = P^n_{\mathrm{div}}\boldsymbol{u}_0 \to \boldsymbol{u}_0$ strongly in $\mathrm{L}^2(\Omega)^d$ as $n \to \infty$ follows from (5.205) and (2.82), so (5.211) is shown.

Finally the uniform estimate in (5.221) and the Banach–Alaoglu theorem imply that there exist $\mathbf{S}, \overline{\mathbf{S}} \in \mathrm{L}^{q'}(Q)^{d\times d}$ such that (up to subsequences) one has that

$$\overline{\mathcal{S}}^n(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^n) \rightharpoonup \mathbf{S} \quad \text{weakly in } \mathrm{L}^{q'}(Q)^{d\times d}, \tag{5.256}$$

$$\mathcal{S}^n(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^n) \rightharpoonup \overline{\mathbf{S}} \quad \text{weakly in } \mathrm{L}^{q'}(Q)^{d\times d}, \tag{5.257}$$

as $n \to \infty$. Then the identification $\mathbf{S} = \overline{\mathbf{S}}$ follows as in Step 5 in the proof of Lemma 5.6. □

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 $\hat{q}$ as defined in (5.5) let us introduce

$$\mathfrak{L}[\boldsymbol{u};\boldsymbol{v}](t) := -b(\boldsymbol{u}(t,\cdot),\boldsymbol{u}(t,\cdot),\boldsymbol{v}) - \langle \mathbf{S}(t,\cdot), \mathbf{D}\boldsymbol{v}\rangle_\Omega + \langle \boldsymbol{f}(t,\cdot),\boldsymbol{v}\rangle_\Omega\,, \tag{5.258}$$

where $\mathbf{S} \in \mathrm{L}^{q'}(Q)^{d\times d}$ is given in Lemma 5.14 and $b(\cdot,\cdot,\cdot)$ is recalled in (5.9).

Lemma 5.15 (Identification of the PDE as $k, l, n \to \infty$)**.**
Let $q > \frac{2(d+1)}{d+2}$ (and $q > \frac{2d}{d+2}$ if the elements of the spaces $\mathbb{V}^n_{\mathrm{div}}$ are exactly divergence-free). The limiting function $\boldsymbol{u} \in \mathrm{L}^\infty(0,T;\mathrm{L}^2_{\mathrm{div}}(\Omega)^d) \cap \mathrm{L}^q(0,T;\mathrm{W}^{1,q}_{0,\mathrm{div}}(\Omega)^d)$ given in Lemma 5.14 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 (5.5). (Up to a representative) we have that $\boldsymbol{u} \in \mathrm{C}_w([0,T],\mathrm{L}^2_{\mathrm{div}}(\Omega)^d)$. Furthermore, the functions $\boldsymbol{u}$ and $\mathbf{S} \in \mathrm{L}^{q'}(Q)^{d\times d}$ from Lemma 5.14 satisfy

$$\langle \partial_t \boldsymbol{u}(t,\cdot), \boldsymbol{w}\rangle_\Omega = \mathfrak{L}[\boldsymbol{u};\boldsymbol{w}](t) \quad \text{for all } \boldsymbol{w} \in \mathrm{C}^\infty_{0,\mathrm{div}}(\Omega)^d, \text{ for a.e. } t \in (0,T), \tag{5.259}$$

$$\operatorname*{ess\,lim}_{t\to 0_+} \|\boldsymbol{u}(t,\cdot) - \boldsymbol{u}_0\|_{\mathrm{L}^2(\Omega)} = 0. \tag{5.260}$$

If additionally $q \geq \frac{3d+2}{d+2}$, then $(\boldsymbol{u}, \mathbf{S})$ satisfy also

$$(\mathbf{D}\boldsymbol{u}(\boldsymbol{z}), \mathbf{S}(\boldsymbol{z})) \in \mathcal{A}(\boldsymbol{z}) \qquad \text{for a.e. } \boldsymbol{z} \in Q, \tag{5.261}$$

i.e., $(\boldsymbol{u}, \mathbf{S})$ is a weak solution according to Definition 5.1.

Proof.

Step 1: Identification of the limiting equation
Let us choose the (arbitrary) subsequences introduced in Step 2 in the proof of Lemma 5.14 (for the notation we replace the index $\kappa = (k_n, l_n, n) \in \mathbb{N}^3$ by $n \in \mathbb{N}$). Starting from (5.205), by integration by parts and due the fact that $\widetilde{\boldsymbol{U}}^n \in \mathrm{C}([0,T];\mathrm{L}^2(\Omega)^d)$, we arrive at

$$-\left\langle \widetilde{\boldsymbol{U}}^n, \boldsymbol{W}\partial_t\varphi\right\rangle_Q = \left\langle \widetilde{\boldsymbol{U}}^n(0,\cdot), \varphi(0)\boldsymbol{W}\right\rangle_\Omega + \left\langle \mathfrak{L}^n[\overline{\boldsymbol{U}}^\kappa, \boldsymbol{W}], \varphi\right\rangle_{(0,T)}, \tag{5.262}$$

for all $\boldsymbol{W} \in \mathbb{V}^n_{\mathrm{div}}$ and all $\varphi \in \mathrm{C}^\infty_0((-T,T))$ and $n \in \mathbb{N}$.
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 2.24 (i) for $\boldsymbol{w} \in \mathrm{C}^\infty_{0,\mathrm{div}}(\Omega)^d$ we have that

$$\mathbb{V}^n_{\mathrm{div}} \ni \Pi^n\boldsymbol{w} \to \boldsymbol{w} \quad \text{strongly in } \mathrm{W}^{1,s}_0(\Omega)^d, \quad \text{as } n\to\infty, \quad \text{for any } s\in[1,\infty). \tag{5.263}$$

To deduce the limiting equation for $\boldsymbol{u}$ the equation (5.262) is considered term by term, $n \to \infty$, as was done in Step 1 in the proof of Lemma 5.7 using the convergence results (5.209), (5.211), (5.213), (5.215) and (5.216). The only term, which is different here is the numerical convective term, so let us focus on this term.
By the fact that $\overline{\boldsymbol{U}}^n \to \boldsymbol{u}$ strongly in $\mathrm{L}^r(Q)^d$ for all $r \in [1, \frac{q(d+2)}{d})$, as $n \to \infty$, by (5.213), it follows that $\overline{\boldsymbol{U}}^n \otimes \overline{\boldsymbol{U}}^n \to \boldsymbol{u}\otimes\boldsymbol{u}$ strongly in $\mathrm{L}^p(Q)^{d\times d}$ for all $p \in [1, \frac{q(d+2)}{2d})$. Such a $p > 1$ exists, since we have that $q > \frac{2d}{d+2}$. With (5.263) 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

$$\left\langle \overline{\boldsymbol{U}}^n \otimes \overline{\boldsymbol{U}}^n, \varphi\nabla\Pi^n\boldsymbol{w}\right\rangle_Q \to \langle \boldsymbol{u}\otimes\boldsymbol{u}, \varphi\nabla\boldsymbol{w}\rangle_Q, \quad \text{as } n\to\infty. \tag{5.264}$$

For the modification of the convective term, which appears only if $\mathbb{V}^n_{\mathrm{div}}$ is discretely divergence-free, note first that we have weak convergence of $\nabla\overline{\boldsymbol{U}}^n \rightharpoonup \nabla\boldsymbol{u}$ in $\mathrm{L}^q(Q)^{d\times d}$ by (5.215). By (5.213) we have in particular that $\overline{\boldsymbol{U}}^n \to \boldsymbol{u}$ strongly in $\mathrm{L}^{q'}(Q)^d$, as $n\to\infty$ since $q' < \frac{q(d+2)}{d}$, provided that $q > \frac{2(d+1)}{d+2}$. 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 $\mathrm{L}^\infty(Q)^d$, as $n\to\infty$. Together, this yields that

$$\left\langle \overline{\boldsymbol{U}}^n \otimes \varphi \Pi^n \boldsymbol{w}, \nabla \overline{\boldsymbol{U}}^n \right\rangle_Q \to \langle \boldsymbol{u} \otimes \varphi \boldsymbol{w}, \nabla \boldsymbol{u} \rangle_Q \,, \quad \text{as } n \to \infty. \tag{5.265}$$

As before, since $\boldsymbol{u}$ is divergence-free, it follows that $\widetilde{b}(\boldsymbol{u}, \boldsymbol{u}, \varphi \boldsymbol{w}) = b(\boldsymbol{u}, \boldsymbol{u}, \varphi \boldsymbol{w})$. All of the other terms can be treated as in Step 1 in the proof of Lemma 5.7, and from (5.262), taking $n \to \infty$, one can conclude that

$$-\langle \boldsymbol{u}, \boldsymbol{w} \partial_t \varphi \rangle_Q = \langle \boldsymbol{u}_0, \varphi(0) \boldsymbol{w} \rangle_\Omega + \langle \mathfrak{L}[\boldsymbol{u}, \boldsymbol{w}], \varphi \rangle_{(0,T)} \tag{5.266}$$

for all $\boldsymbol{w} \in \mathrm{C}^\infty_{0,\mathrm{div}}(\Omega)^d$ and all $\varphi \in \mathrm{C}^\infty_0((-T,T))$, with $\mathfrak{L}$ as defined in (5.183).

The distributional derivative of $\boldsymbol{u}$ satisfies, by definition and using (5.266), that

$$\langle \partial_t \boldsymbol{u}, \boldsymbol{w} \varphi \rangle_Q = -\langle \boldsymbol{u}, \boldsymbol{w} \partial_t \varphi \rangle_Q \overset{(5.266)}{=} \langle \mathfrak{L}[\boldsymbol{u}, \boldsymbol{w}], \varphi \rangle_{(0,T)} \tag{5.267}$$

for all $\boldsymbol{w} \in \mathrm{C}^\infty_{0,\mathrm{div}}(\Omega)^d$ and all $\varphi \in \mathrm{C}^\infty_0((0,T))$, since $\operatorname{supp} \varphi \subset (0,T)$. This is the same equation as in (5.184) in Lemma 5.10, where we have seen 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 (5.5). Consequently, $\langle \partial_t \boldsymbol{u}, \boldsymbol{w} \rangle_\Omega$ is integrable for $\boldsymbol{w} \in \mathrm{C}^\infty_{0,\mathrm{div}}(\Omega)^d$, and thus, we can rephrase (5.267) by the fundamental lemma of calculus of variations in the pointwise sense in time, so (5.259) is proved.

Step 2: Identification of the initial condition
The fact that $\boldsymbol{u} \in \mathrm{C}_w([0,T], \mathrm{L}^2_{\mathrm{div}}(\Omega)^d)$, that $\boldsymbol{u}(0,\cdot) = \boldsymbol{u}_0$ and (5.260) follow as in Step 2 in the proof of Lemma 5.10 using (5.210), (5.211) and (5.215).

Step 3: Energy identity
If $q \geq \frac{3d+2}{d+2}$, then we have that $\hat{q} = q$, and thus $\partial_t \boldsymbol{u} \in \mathrm{L}^{q'}(0,T; (\mathrm{W}^{1,q}_{0,\mathrm{div}}(\Omega)^d)')$ and in particular $\boldsymbol{u} \in \mathrm{L}^q(0,T; \mathrm{W}^{1,q}_{0,\mathrm{div}}(\Omega)^d) \cap \mathrm{L}^2(0,T; \mathrm{L}^2_{\mathrm{div}}(\Omega)^d)$. For $q \geq \frac{2d}{d+2}$, the embedding $\mathrm{W}^{1,q}_{0,\mathrm{div}}(\Omega)^d \hookrightarrow \mathrm{L}^2_{\mathrm{div}}(\Omega)^d$ is continuous and dense. Hence, by Lemma 2.11 (with $q > 1$) we have that $\boldsymbol{u} \in \mathrm{C}([0,T]; \mathrm{L}^2_{\mathrm{div}}(\Omega)^d)$ up to a representative.

The function $\boldsymbol{u} \in \mathrm{L}^q(0,T; \mathrm{W}^{1,q}_{0,\mathrm{div}}(\Omega)^d)$ is an admissible test function in (5.259) and by the fact that $\boldsymbol{u} \in \mathrm{C}([0,T]; \mathrm{L}^2_{\mathrm{div}}(\Omega)^d)$ we obtain that

$$\langle \mathbf{S}, \mathbf{D}\boldsymbol{u} \rangle_Q = \langle \boldsymbol{f}, \boldsymbol{u} \rangle_Q + \frac{1}{2} \|\boldsymbol{u}_0\|^2_{\mathrm{L}^2(\Omega)} - \frac{1}{2} \|\boldsymbol{u}(T,\cdot)\|^2_{\mathrm{L}^2(\Omega)} \,, \tag{5.268}$$

because the convective term vanishes and $\boldsymbol{u}(0,\cdot) = \boldsymbol{u}_0$.

Step 4: Identification of the implicit relation
Because $q \geq \frac{3d+2}{d+2}$ we can use a Minty type convergence result and the available energy identities to show that the implicit constitutive relation (5.261) is satisfied. Recall that by the definition of $\boldsymbol{\mathcal{S}}^k$ in (3.59), with the notation for the subsequences as in (5.234), we have that

$$(\mathbf{D}\overline{\boldsymbol{U}}^n(\boldsymbol{z}), \boldsymbol{\mathcal{S}}^n(\boldsymbol{z}, \mathbf{D}\overline{\boldsymbol{U}}^n(\boldsymbol{z})) \in \mathcal{A}^n(\boldsymbol{z}), \tag{5.269}$$

for any $n \in \mathbb{N}$ and a.e. $\boldsymbol{z} \in Q$. By (5.215) we have that $\mathbf{D}\overline{\boldsymbol{U}}^n \rightharpoonup \mathbf{D}\boldsymbol{u}$ weakly in $\mathrm{L}^q(Q)^{d\times d}$, and by (5.217) that $\boldsymbol{\mathcal{S}}^n(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^n) \rightharpoonup \mathbf{S}$ weakly in $\mathrm{L}^{q'}(Q)^{d\times d}$, as $n \to \infty$.

The only thing that is left to show in order to apply Lemma 3.31 for the generalised Yosida approximation, is that

$$\limsup_{n\to\infty} \langle \boldsymbol{\mathcal{S}}^n(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^n),\mathbf{D}\overline{\boldsymbol{U}}^n\rangle_Q \le \langle \mathbf{S},\mathbf{D}\boldsymbol{u}\rangle_Q\,. \tag{5.270}$$

Then all the assumptions are satisfied and we can conclude that $(\mathbf{D}\boldsymbol{u}(\boldsymbol{z}),\mathbf{S}(\boldsymbol{z}))\in\mathcal{A}(\boldsymbol{z})$ for a.e. $\boldsymbol{z}\in Q$, which finishes the proof.

Recall that the energy inequality (5.207) for $\overline{\boldsymbol{U}}^n$ is available for all $s_0,s\in[0,T]$, so we have in particular that

$$\left\langle \overline{\boldsymbol{\mathcal{S}}}^n(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^n),\mathbf{D}\overline{\boldsymbol{U}}^n\right\rangle_Q \le \langle \overline{\boldsymbol{f}},\overline{\boldsymbol{U}}^n\rangle_Q + \frac12\left\|\widetilde{\boldsymbol{U}}^n(0,\cdot)\right\|^2_{\mathrm{L}^2(\Omega)} - \frac12\left\|\widetilde{\boldsymbol{U}}^n(T,\cdot)\right\|^2_{\mathrm{L}^2(\Omega)}\,. \tag{5.271}$$

As in (5.144) one can show that

$$\langle \boldsymbol{\mathcal{S}}^n(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^n),\mathbf{D}\overline{\boldsymbol{U}}^n\rangle_Q = \left\langle \overline{\boldsymbol{\mathcal{S}}}^n(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^n),\mathbf{D}\overline{\boldsymbol{U}}^n\right\rangle_Q\,, \tag{5.272}$$

since $\mathbf{D}\overline{\boldsymbol{U}}^n\in\mathbb{P}_0^{l_n}(0,T;\mathrm{L}^q(\Omega)^{d\times d})$.

Using this and then the energy inequality (5.271) for $\overline{\boldsymbol{U}}^n$ and the identity (5.268) for $\boldsymbol{u}$, it follows that

$$\begin{aligned}\limsup_{n\to\infty} \langle \boldsymbol{\mathcal{S}}^n(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^n),\mathbf{D}\overline{\boldsymbol{U}}^n\rangle_Q &\overset{(5.272)}{=} \limsup_{n\to\infty}\left\langle \overline{\boldsymbol{\mathcal{S}}}^n(\cdot,\cdot,\mathbf{D}\overline{\boldsymbol{U}}^n),\mathbf{D}\overline{\boldsymbol{U}}^n\right\rangle_Q\\ &\overset{(5.271)}{\le} \lim_{n\to\infty}\langle \overline{\boldsymbol{f}},\overline{\boldsymbol{U}}^n\rangle_Q + \lim_{n\to\infty}\frac12\left\|\widetilde{\boldsymbol{U}}^n(0,\cdot)\right\|^2_{\mathrm{L}^2(\Omega)}\\ &\qquad - \lim_{n\to\infty}\frac12\left\|\widetilde{\boldsymbol{U}}^n(T,\cdot)\right\|^2_{\mathrm{L}^2(\Omega)}\\ &\overset{(\#)}{=} \langle \boldsymbol{f},\boldsymbol{u}\rangle_Q + \frac12\|\boldsymbol{u}_0\|^2_{\mathrm{L}^2(\Omega)} - \frac12\|\boldsymbol{u}(T,\cdot)\|^2_{\mathrm{L}^2(\Omega)}\\ &\overset{(5.268)}{=} \langle \mathbf{S},\mathbf{D}\boldsymbol{u}\rangle_Q\,,\end{aligned} \tag{5.273}$$

where the equality denoted by $(\#)$ relies on the following arguments: the convergence in the first term follows from the strong convergence $\overline{\boldsymbol{f}}\to\boldsymbol{f}$ in $\mathrm{L}^{q'}(0,T;\mathrm{W}^{-1,q'}(\Omega)^d)$ in (5.22) and the weak convergence of $\overline{\boldsymbol{U}}^n$ in $\mathrm{L}^q(0,T;\mathrm{W}^{1,q}_0(\Omega)^d)$ in (5.215). The convergence of the second term follows from (5.211) and the convergence of the last term follows from (5.210), which holds for all $t\in[0,T]$, since $\boldsymbol{u}\in\mathrm{C}([0,T];\mathrm{L}^2_{\mathrm{div}}(\Omega)^d)$. This shows (5.270) and hence finishes the proof.

□

Remark 5.16 (Possible Extensions)**.**

(i) *For $q\ge\frac{3d+2}{d+2}$, the assumption (I) in Lemma 5.14 is automatically satisfied and we do not enter the range, where only the Aubin–Lions lemma can be used. However, if a discrete truncation would be at hand, thanks to the convergence lemma and the first part of Lemma 5.15 one would only have to identify the implicit constitutive relation.*

(ii) *An analogous result can be shown for the Stokes problem for $q\in\left(\frac{2d}{d+2},\infty\right)$ since all the restrictions arose from the convective term and its numerical modification.*

(iii) *The approximate problem in (5.201), (5.202), can also be formulated semi-implicitly by replacing $\widetilde{b}(\boldsymbol{u},\boldsymbol{u},\boldsymbol{v})$ by $\widetilde{b}(\boldsymbol{U}^\kappa_{i-1},\boldsymbol{u},\boldsymbol{v})$, for $\kappa=(k,l,n)\in\mathbb{N}^3$. Since this represents a linearisation of the problem, the approximate solutions exist and are unique. Then one works with the piecewise constant interpolant of $\{\boldsymbol{U}^\kappa_i\}_{i\in\{0,\dots,l-1\}}$. To show the uniform*

estimates corresponding to the ones for $\overline{\boldsymbol{U}}^{\kappa}$ *one has to bound* $\delta_l \|\boldsymbol{U}_0^n\|_{\mathrm{W}^{1,q}(\Omega)}^q$ *uniformly in* $l, n \in \mathbb{N}$*. This can be done by assuming one of the following:*

(1) Assume that $\{\mathcal{T}_n\}_{n\in\mathbb{N}}$ *is quasiuniform, i.e., Assumption 2.19 is satisfied and*

$$\delta_l \leq c h_n^{\max(q, q\frac{d+2}{2}-d)} \tag{5.274}$$

for a constant independent of $l, n \in \mathbb{N}$*;*

(2) Or assume that $\boldsymbol{u}_0 \in \mathrm{W}^{1,q}_{0,\mathrm{div}}(\Omega)^d$ *and replace (5.201) by* $\boldsymbol{U}_0^n = \Pi^n \boldsymbol{u}_0$*.*

Then, the rest of the convergence proof works. Obviously, the same can be done for the regularised problem.

5.4. Discussion

Let us first recall the relevant values for the exponent range for q:

$$\frac{2d}{d+2} < \frac{2(d+1)}{d+2} < q_d < \widetilde{q}_d < 2 \leq \frac{3d+2}{d+2}, \quad d \in \{2,3\}, \tag{5.275}$$

and refer to Figure 5.2 for the values.

In Subsection 5.3.1 we have shown convergence of the interpolants of solutions to a fully time-discretised approximate problem including a regularising term, when taking the regularising limit after the discretisation limit. This allowed us to show weak convergence (up to subsequences) to a weak solution for the whole range of existence $q \in \left(\frac{2d}{d+2}, \infty\right)$ when taking the limits successively. The use of a regularisation allowed us to apply Simon's compactness lemma in the discretisation limit instead of the Aubin–Lions compactness lemma. This means that additional assumptions, which are used in order to prove stability of the L^2-projection onto $\mathbb{V}^n_{\mathrm{div}}$, need not be imposed. Furthermore, by regularising we do not obtain any additional restrictions on the parameter range of q, stemming from the numerical modification of the convective term in case the elements of $\mathbb{V}^n_{\mathrm{div}}$ are discretely divergence-free only. Separating the discretisation limit and the limit in which we lose admissibility allows us to use a continuous parabolic solenoidal Lipschitz approximation to identify the implicit constitutive relation. This is useful, since other than in the steady case a discrete Lipschitz truncation is not available at present.

In Subsection 5.3.2 we have investigated the approximate problem without regularisation and thanks to the generalised Yosida approximation we were able to take the limits $k, l, n \to \infty$ simultaneously. The strongest restriction on q arises from the identification of the implicit constitutive relation, where, due to the lack of a discrete truncation, we assume that q is in the admissible range $q \in \left[\frac{3d+2}{d+2}, \infty\right)$. All arguments except the identification of the constitutive law can in fact be shown for a larger range of q. We have presented the details of those arguments to show where the restrictions on q come from, and what one can at most

d	$\frac{2d}{d+2}$	$\frac{2(d+1)}{d+2}$	q_d	$\widetilde{q}_d$	$\frac{3d+2}{d+2}$
2	1	$\frac{3}{2} = 1.5$	$\frac{1}{2}(1+\sqrt{5}) \approx 1.62$	$1 + \frac{1}{\sqrt{2}} \approx 1.71$	2
3	$\frac{6}{5} = 1.2$	$\frac{8}{5} = 1.6$	$\frac{1}{5}\left(3+\sqrt{39}\right) \approx 1.85$	$\frac{1}{10}\left(11+\sqrt{61}\right) \approx 1.88$	$\frac{11}{5} = 2.2$

Fig. 5.2: Values of exponents relevant for the unsteady problem

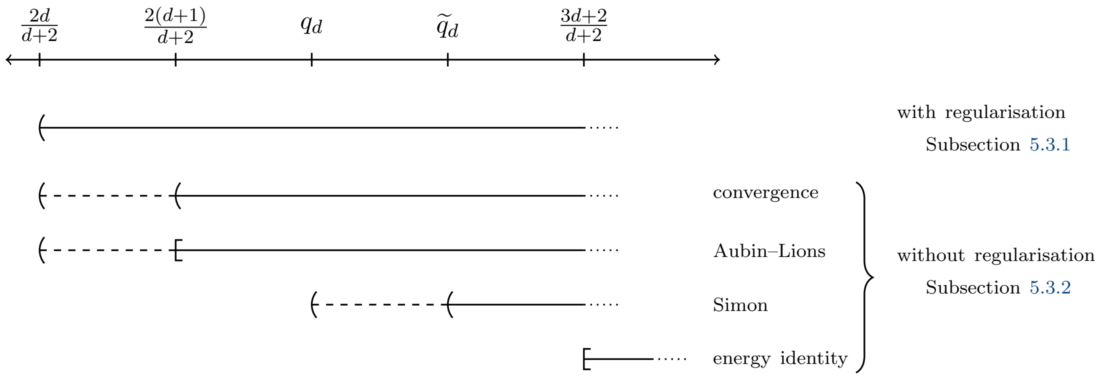

Fig. 5.3: Overview of the range of q for which the convergence results in the unsteady case hold; the dashed line represents the special case, if $\mathbb{V}^n_{\text{div}}$ satisfies Assumption 2.27.

expect from the use of a discrete truncation. The modification of the convective term is the reason that we have to restrict the range of q to $q > \frac{2(d+1)}{d+2}$, in case $\mathbb{V}^n_{\text{div}}$ is only discretely divergence-free, in order to identify the limiting equation. If the spaces $\mathbb{V}^n_{\text{div}}$ consist of exactly divergence-free functions one only requires that $q > \frac{2d}{d+1}$. The compactness can be obtained using the Simon lemma for the parameter range $q > \widetilde{q}_d$ (or $q > q_d$) and for the larger range $q \geq \frac{2(d+1)}{d+2}$ (or $q > \frac{2d}{d+2}$) by the Aubin–Lions lemma. For the application of the latter additional assumptions on the finite element setting were required, in order to show the stability properties of the L^2-projection, mentioned before.

Recall that under slightly stronger assumptions for a semi-implicit approximation the corresponding convergence result can be shown.

Let us give an overview of the options we have for the convergence results, see also Figure 5.3:

- If $q \in \left[\frac{3d+2}{d+2}, \infty\right)$, then the solution $\boldsymbol{u}$ is admissible as a test function and neither a regularisation nor a truncation is required in order to show convergence to a weak solution.
- If $q \in \left(\frac{2d}{d+2}, \frac{3d+2}{d+2}\right)$, independently of $\mathbb{V}^n_{\text{div}}$, only the regularisation approach is available. Since no discrete truncation is available we do not enter the range for which it makes a difference whether the spaces $\mathbb{V}^n_{\text{div}}$ are discretely or exactly divergence-free.

As in the steady case one can choose between taking the limit $k \to \infty$ for general graph approximation satisfying Assumption 3.18 or the more specific choice of the (generalised) Yosida approximation, which allows to take the limit in k together with the other limits.

Comparing the exponent ranges for the respective arguments for the steady and the unsteady case, even if a discrete parabolic truncation were available, one can see that the ranges in the unsteady case are smaller. However, since there is no discrete truncation available yet, there is still some work to be done before the corresponding result can be shown to hold in the unsteady case for the unregularised problem. More specifically, for $q \in \left(\frac{2(d+1)}{d+2}, \frac{3d+2}{d+2}\right)$ (or $q \in \left(\frac{2d}{d+2}, \frac{3d+2}{d+2}\right)$ if $\mathbb{V}^n_{\text{div}}$ is exactly divergence-free) the identification of the implicit constitutive relation is still an open problem.

APPENDIX A

Results Related to Pressure and Solenoidality

Here we collect results related to the pressure reconstruction, the solenoidality of the velocity function. These concern both the continuous case and the discrete case and are relevant for the steady problem in Chapter 4. Furthermore, we investigate the divergence-preserving projection operator Π^n satisfying Assumption 2.23 for a range of mixed finite element spaces.

In Section A.1 we will collect results for the continuous case. This includes the Bogovskiĭ operator, the inf-sup condition and a divergence-corrected Lipschitz approximation in the steady case. Section A.2 includes the discrete counterparts of those results and the investigation of divergence-preserving projection operators.

First let us recall the following well-known abstract characterisation.

Lemma A.1 ([EG04, Lem. A.40, and Thm. A.34]).
Let X, Y be real Banach spaces and let $A\colon X \to Y$ a linear bounded operator and A' its adjoint operator. Then the following are equivalent:

(i) $A\colon X \to Y$ is surjective;

(ii) $A'\colon Y' \to X'$ is injective and $\operatorname{Im}(A') = (\ker(A))^\perp$;

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

$$\|A'y'\|_{X'} \geq c\,\|y'\|_{Y'} \qquad \text{for all } y' \in Y';$$

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

$$\inf_{y' \in Y' \setminus \{0\}} \sup_{x \in X \setminus \{0\}} \frac{\langle A'y', x\rangle_{X',X}}{\|y'\|_{Y'}\,\|x\|_X} \geq c. \tag{A.1}$$

A.1. The Continuous Case

By means of singular integrals one can construct a right inverse of the divergence operator, often referred to as Bogovskiĭ operator. The following result regarding such an operator dates back to [Bog79] and can also be found in [Gal11, Thm. III.3.1]. For more general versions we refer to [ADM06] and [DRS10, Thm. 5.2].

Lemma A.2 (Bogovskiĭ Operator, [Gal11, Thm. III.3.1]).
There exists a linear map $\mathfrak{B}$, which for all $p \in (1,\infty)$ maps $\mathrm{L}^p_0(\Omega) \to \mathrm{W}^{1,p}_0(\Omega)^d$ and satisfies

$$\operatorname{div} \mathfrak{B}h = h \qquad \text{and} \qquad \|\mathfrak{B}h\|_{\mathrm{W}^{1,p}(\Omega)} \leq c(p)\,\|h\|_{\mathrm{L}^p(\Omega)}, \tag{A.2}$$

for all $h \in \mathrm{L}_0^p(\Omega)$*, where the constant depends only on* p*,* d *and on* Ω*.*

The existence of the Bogovskiĭ operator implies that the continuous inf-sup condition holds for any $p \in (1, \infty)$:

Corollary A.3 (Continuous Inf-sup Condition, [DKS13a, (2.1)])**.** *For any* $p \in (1, \infty)$ *and* p' *its Hölder conjugate, there exists a constant* $c_i = c_i(p) > 0$ *such that*

$$\sup_{\boldsymbol{v} \in \mathrm{W}_0^{1,p}(\Omega)^d \setminus \{\boldsymbol{0}\}} \frac{\langle \operatorname{div} \boldsymbol{v}, r \rangle_\Omega}{\|\boldsymbol{v}\|_{\mathrm{W}^{1,p}(\Omega)}} \geq c_i \|r\|_{\mathrm{L}^{p'}(\Omega)} \qquad \text{for all } r \in \mathrm{L}_0^{p'}(\Omega). \tag{A.3}$$

Proof. This follows directly from Lemma A.1: For $p \in (1, \infty)$ fixed choose $X = \mathrm{W}_0^{1,p}(\Omega)^d$, $Y = \mathrm{L}_0^p(\Omega)$ and $A = \operatorname{div}$. Lemma A.2 shows that (i) is satisfied. Hence (iv) holds and the claim follows with $(\mathrm{L}_0^p(\Omega))' \cong \mathrm{L}_0^{p'}(\Omega)$ and constant depending on p. □

Also, thanks to Lemma A.1 the pressure can be reconstructed for the continuous steady problem, but this is not needed for our purposes.

Using the Bogovskiĭ operator one can correct the divergence of the Lipschitz truncation of a divergence-free function in Ω. The following proof is contained in the proof of Thm. 3.1 in [DMS08] and will be reproduced for the reader's convenience.

Proof of Lemma 2.13: Divergence Corrected Lipschitz Approximation

We set $\boldsymbol{w}^{l,j} := \boldsymbol{v}^{l,j} - \mathfrak{B}(\operatorname{div}(\boldsymbol{v}^{l,j}))$, for any $l, j \in \mathbb{N}$, where $\mathfrak{B}$ is the Bogovskiĭ operator from Lemma A.2. Since $\boldsymbol{v}^{l,j} \in \mathrm{W}_0^{1,s}(\Omega)^d$ and because $\mathfrak{B} \colon \mathrm{L}_0^s(\Omega) \to \mathrm{W}_0^{1,s}(\Omega)^d$ for any $s \in (1, \infty)$ we have that $\boldsymbol{w}^{l,j} \in \mathrm{W}_0^{1,s}(\Omega)^d$ for any $s \in (1, \infty)$. Due to (A.2) it follows that $\operatorname{div}(\mathfrak{B}(\operatorname{div} \boldsymbol{v}^{l,j})) = \operatorname{div} \boldsymbol{v}^{l,j}$, so $\boldsymbol{w}^{l,j}$ is divergence-free.
By assumption $\boldsymbol{v}^l$ is divergence-free for each $l \in \mathbb{N}$ and we have that $\boldsymbol{v}^{l,j} = \boldsymbol{v}^l$ on $\Omega \backslash \mathcal{B}_{l,j}$. Hence $\boldsymbol{v}^{l,j}$ is divergence-free on $\Omega \backslash \mathcal{B}_{l,j}$, so that $\operatorname{div} \boldsymbol{v}^{l,j} = \mathbb{1}_{\mathcal{B}_{l,j}} \operatorname{div} \boldsymbol{v}^{l,j} \in \mathrm{L}_0^p(\Omega)$. Thus, it follows that

$$\boldsymbol{w}^{l,j} - \boldsymbol{v}^{l,j} = -\mathfrak{B}(\operatorname{div}(\boldsymbol{v}^{l,j})) = -\mathfrak{B}(\mathbb{1}_{\mathcal{B}_{l,j}} \operatorname{div}(\boldsymbol{v}^{l,j})),$$

and consequently, by the continuity of $\mathfrak{B}$ in (A.2) and the properties of the Lipschitz approximation in Lemma 2.12 (ii), (iii)

$$\begin{aligned}
\left\|\boldsymbol{w}^{l,j} - \boldsymbol{v}^{l,j}\right\|_{\mathrm{W}^{1,p}(\Omega)} &= \left\|\mathfrak{B}(\mathbb{1}_{\mathcal{B}_{l,j}} \operatorname{div}(\boldsymbol{v}^{l,j}))\right\|_{\mathrm{W}^{1,p}(\Omega)} \leq c \left\|\mathbb{1}_{\mathcal{B}_{l,j}} \operatorname{div} \boldsymbol{v}^{l,j}\right\|_{\mathrm{L}^p(\Omega)} \\
&\leq c \left\|\mathbb{1}_{\mathcal{B}_{l,j}}\right\|_{\mathrm{L}^p(\Omega)} \left\|\nabla \boldsymbol{v}^{l,j}\right\|_{\mathrm{L}^\infty(\Omega)} \leq c 2^{-\frac{j}{p}},
\end{aligned}$$

which proves (ii).

Because $\mathfrak{B}$ is continuous with respect to the strong convergences in the respective spaces, it is also weakly-weakly continuous, compare [AB06, Thm. 6.17]. By Lemma 2.12 (iv) we have that $\operatorname{div} \boldsymbol{v}^{l,j} \rightharpoonup 0$ weakly in $\mathrm{L}_0^s(\Omega)$ as $l \to \infty$ for foxed $j \in \mathbb{N}$, for any $s \in (1, \infty)$. With (A.2) this implies that for any fixed $j \in \mathbb{N}$ we have that

$$\mathfrak{B}(\operatorname{div} \boldsymbol{v}^{l,j}) \rightharpoonup \boldsymbol{0} \quad \text{weakly in } \mathrm{W}_0^{1,s}(\Omega)^d, \quad \text{as } l \to \infty.$$

This holds in particular for $s > d$, so by the compact embedding there exists a subsequence such that we have

$$\mathfrak{B}(\operatorname{div} \boldsymbol{v}^{l,j}) \to \mathbf{0} \quad \text{in } \mathrm{L}^s(\Omega)^d, \quad \text{as } l \to \infty,$$

for any $s \in [1,\infty)$ and fixed $j \in \mathbb{N}$. Together with the convergence of $\boldsymbol{v}^{l,j}$ in Lemma 2.12 (iv) this implies (iii). □

A.2. The Discrete Situation

Let us now summarise the discrete versions of the above and also state the pressure reconstruction in the discrete (steady) case. First, recall the finite element setting from Subsection 2.2.1.

The proof of the following lemma is a consequence of the continuous inf-sup condition in Corollary A.3 and Assumption 2.23 on Π^n.

Lemma A.4 (Discrete Inf-sup Condition, [DKS13a, Prop. 8])**.**
Let us assume that Assumption 2.18 is satisfied by Ω and the family of simplicial partitions $\{\mathcal{T}_n\}_{n\in\mathbb{N}}$. Furthermore, let $\mathbb{V}^n$, $\mathbb{Q}^n$ and $\mathbb{Q}^n_0$ as defined in (2.30), (2.31) and (2.33) and let Assumption 2.23 (i), (ii) hold.
Then, for any $p \in (1,\infty)$ there exists a constant $c_I = c_I(p)$ (independent of n) such that

$$\sup_{\boldsymbol{V}\in\mathbb{V}^n\setminus\{\mathbf{0}\}} \frac{\langle \operatorname{div} \boldsymbol{V}, Q\rangle_\Omega}{\|\boldsymbol{V}\|_{\mathrm{W}^{1,p}(\Omega)}} \geq c_I \|Q\|_{\mathrm{L}^{p'}(\Omega)}, \quad \textit{for all } Q \in \mathbb{Q}^n_0 \textit{ and all } n \in \mathbb{N}. \tag{A.4}$$

Proof. The the proof is contained in [EG04, Lem. 4.19], where a more general statement was proved for constants possibly depending on $n \in \mathbb{N}$: For any $Q \in \mathbb{Q}^n_0 \subset \mathrm{L}^{p'}_0(\Omega)$ we apply the continuous inf-sup condition in Corollary A.3, the divergence preserving property and the $\mathrm{W}^{1,p}(\Omega)$-stability of Π^n according to Assumption 2.23 (i), (ii) to obtain

$$\begin{aligned} c_i \|Q\|_{\mathrm{L}^{p'}(\Omega)} &\leq \sup_{\boldsymbol{v}\in\mathrm{W}^{1,p}_0(\Omega)^d\setminus\{\mathbf{0}\}} \frac{\langle \operatorname{div} \boldsymbol{v}, Q\rangle_\Omega}{\|\boldsymbol{v}\|_{\mathrm{W}^{1,p}(\Omega)}} = \sup_{\boldsymbol{v}\in\mathrm{W}^{1,p}_0(\Omega)^d\setminus\{\mathbf{0}\}} \frac{\langle \operatorname{div}(\Pi^n\boldsymbol{v}), Q\rangle_\Omega}{\|\boldsymbol{v}\|_{\mathrm{W}^{1,p}(\Omega)}} \\ &\leq c \sup_{\boldsymbol{v}\in\mathrm{W}^{1,p}_0(\Omega)^d:\, \|\Pi^n\boldsymbol{v}\|>0} \frac{\langle \operatorname{div}(\Pi^n\boldsymbol{v}), Q\rangle_\Omega}{\|\Pi^n\boldsymbol{v}\|_{\mathrm{W}^{1,p}(\Omega)}} \leq c \sup_{\boldsymbol{V}\in\mathbb{V}^n\setminus\{\mathbf{0}\}} \frac{\langle \operatorname{div}(\boldsymbol{V}), Q\rangle_\Omega}{\|\boldsymbol{V}\|_{\mathrm{W}^{1,p}(\Omega)}}, \end{aligned}$$

where in the last step we have used that $\Pi^n(\mathrm{W}^{1,p}_0(\Omega)^d) \subset \mathbb{V}^n$. Then we have that $c_I = \frac{c_i}{c}$, which is independent of n. □

Again with Lemma A.1 one can prove the following result, for which various proofs are known in the literature.

Corollary A.5 (Discrete Pressure Reconstruction)**.**
Under the assumptions of Lemma A.4 let $n \in \mathbb{N}$, $p \in (1,\infty)$ and $f \in \mathrm{W}^{-1,p'}(\Omega)^d$ be given, such that

$$\langle \boldsymbol{f}, \boldsymbol{W}\rangle_\Omega = 0 \quad \textit{for all } \boldsymbol{W} \in \mathbb{V}^n_{\operatorname{div}}.$$

Then there exists a unique $\pi^n \in \mathbb{Q}^n_0$, such that

$$\langle \boldsymbol{f}, \boldsymbol{V}\rangle_\Omega = \langle \operatorname{div} \boldsymbol{V}, \pi\rangle_\Omega \quad \textit{for all } \boldsymbol{V} \in \mathbb{V}^n,$$

and there exists a constant $c(p) > 0$ (independent of n) such that

$$\|\pi^n\|_{L^{p'}(\Omega)} \le c\,\|\boldsymbol{f}\|_{W^{-1,p'}(\Omega)} \qquad \text{for all } n \in \mathbb{N}.$$

Proof. This can be shown by Lemma A.1: Choosing $X = \mathbb{V}^n$ with norm $\|\cdot\|_{W^{1,p}(\Omega)}$, $A = \operatorname{div}$ and $Y = (\mathbb{Q}_0^n)'$ with $\|\cdot\|_{L^p(\Omega)}$ (which is reflexive). Then $Y' = \mathbb{Q}_0^n$ and by Lemma A.4 the condition (iv) is satisfied. Then $\boldsymbol{f} \in (\ker(A))^\perp$ and existence follows from (ii) and the boundedness follows from (iii). The linearity and the boundedness also imply uniqueness. □

A.2.1. Discrete Bogovskiĭ Operator and Corrected Lipschitz Truncation

The existence of a discrete Bogovskiĭ operator can be shown by Lemma A.1, but this operator would depend on $p \in (1,\infty)$ and linearity would not be guaranteed. For this reason it is better to construct it directly using the projection operator Π^n according to Assumption 2.23 and the Bogovskiĭ operator $\mathfrak{B}$.

Corollary A.6 (Discrete Bogovskiĭ Operator, [DKS13b, Cor. 10])**.**
Under the assumptions of Lemma A.4 let additionally Assumption 2.21 hold. Then, for each $n \in \mathbb{N}$, there exists a linear operator $\mathfrak{B}^n \colon \operatorname{div}\mathbb{V}^n \to \mathbb{V}^n$ and for any $p \in (1,\infty)$, there exists a constant $c = c(p) > 0$ (independent of $n \in \mathbb{N}$), such that

$$\operatorname{div}(\mathfrak{B}^n H) = H \quad \text{in } (\mathbb{Q}^n)' \quad \text{and} \quad \|\mathfrak{B}^n H\|_{W^{1,p}(\Omega)} \le c \sup_{Q \in \mathbb{Q}^n \setminus \{\mathbf{0}\}} \frac{\langle H, Q\rangle_\Omega}{\|Q\|_{L^{p'}(\Omega)}}, \tag{A.5}$$

for all $H \in \operatorname{div}\mathbb{V}^n$ and all $n \in \mathbb{N}$.
Furthermore, if also Assumption 2.23 (iia) holds and if $\{\boldsymbol{V}^n\}_{n\in\mathbb{N}}$ is a sequence such that $\boldsymbol{V}^n \in \mathbb{V}^n$ for each $n \in \mathbb{N}$, and $\boldsymbol{V}^n \rightharpoonup \mathbf{0}$ weakly in $W^{1,p}(\Omega)^d$, as $n \to \infty$, for some $p \in (1,\infty)$, then we have that

$$\mathfrak{B}^n \operatorname{div} \boldsymbol{V}^n \rightharpoonup \mathbf{0} \qquad \text{weakly in } W_0^{1,p}(\Omega)^d, \tag{A.6}$$

as $n \to \infty$, where $\mathfrak{B}$ is the Bogovskiĭ operator in Lemma A.2.

The following proof is contained in the proof of [DKS13a, Lem. 20] and presented for the reader's convenience.

Proof of Lemma 2.30: Discrete Divergence Corrected Lipschitz Truncation

Proof. Set $\boldsymbol{W}^{n,j} := \boldsymbol{V}^{n,j} - \mathfrak{B}^n(\operatorname{div}(\boldsymbol{V}^{n,j}))$, for any $n, j \in \mathbb{N}$, where $\mathfrak{B}^n$ is the discrete Bogovskiĭ operator introduced in Corollary A.6. Since $\boldsymbol{V}^{n,j} \in \mathbb{V}^n$ and $\mathfrak{B}^n \colon \operatorname{div}\mathbb{V}^n \to \mathbb{V}^n$ we have that $\boldsymbol{W}^{n,j} \in \mathbb{V}^n$ for any $n, j \in \mathbb{N}$. Furthermore, since $\operatorname{div}\mathfrak{B}^n(\operatorname{div}\boldsymbol{V}^{n,j}) = \operatorname{div}\boldsymbol{V}^{n,j}$ in $(\mathbb{Q}^n)'$ by (A.5) testing with $Q \in \mathbb{Q}^n$ yields that $\boldsymbol{W}^{n,j} \in \mathbb{V}^n_{\operatorname{div}}$ for any $n, j \in \mathbb{N}$, so (i) is proven.

Using the definition of $\boldsymbol{W}^{n,j}$ and the properties of $\mathfrak{B}^n$ in Corollary A.6 we obtain that

$$\big\|\boldsymbol{W}^{n,j} - \boldsymbol{V}^{n,j}\big\|_{W^{1,p}(\Omega)} = \big\|\mathfrak{B}^n(\operatorname{div}\boldsymbol{V}^{n,j})\big\|_{W^{1,p}(\Omega)} \le c \sup_{Q \in \mathbb{Q}^n \setminus \{\mathbf{0}\}} \frac{\langle \operatorname{div}\boldsymbol{V}^{n,j}, Q\rangle_\Omega}{\|Q\|_{L^{p'}(\Omega)}}. \tag{A.7}$$

Next we use the fact that $\boldsymbol{V}^n \in \mathbb{V}^n_{\operatorname{div}}$ and we expand $Q \in \mathbb{Q}^n$ in the locally supported basis given by Assumption 2.20, i.e., writing $Q = \sum_{i=1}^{d_n} \beta_i Q_i^n$ for $\boldsymbol{\beta} \in \mathbb{R}^{d_n}$. By Lemma 2.29 (i) we have that $\boldsymbol{V}^{n,j} = \boldsymbol{V}^n$ on $\Omega \setminus \mathcal{B}_{n,j}$ and hence, those terms that correspond to $i \in \{1,\dots,d_n\}$

such that $\operatorname{supp} Q_i^n \subset \Omega \backslash \mathcal{B}_{n,j}$ are zero. Thus, we obtain

$$\begin{aligned}
\left\langle \operatorname{div} \boldsymbol{V}^{n,j}, Q \right\rangle_\Omega &= \left\langle \operatorname{div} \boldsymbol{V}^{n,j} - \operatorname{div} \boldsymbol{V}^n, Q \right\rangle_\Omega \\
&= \sum_{i=1}^{d_n} \left\langle \operatorname{div} \boldsymbol{V}^{n,j} - \operatorname{div} \boldsymbol{V}^n, \beta_i Q_i^n \right\rangle_\Omega \\
&= \sum_{\substack{i \in \{1,\dots,d_n\}: \\ \operatorname{supp} Q_i^n \cap \mathcal{B}_{n,j} \neq \emptyset}} \left\langle \operatorname{div} \boldsymbol{V}^{n,j} - \operatorname{div} \boldsymbol{V}^n, \beta_i Q_i^n \right\rangle_\Omega \\
&= \sum_{\substack{i \in \{1,\dots,d_n\}: \\ \operatorname{supp} Q_i^n \cap \mathcal{B}_{n,j} \neq \emptyset}} \left\langle \operatorname{div} \boldsymbol{V}^{n,j}, \beta_i Q_i^n \right\rangle_\Omega \\
&= \left\langle \operatorname{div} \boldsymbol{V}^{n,j}, \sum_{\substack{i \in \{1,\dots,d_n\}: \\ \operatorname{supp} Q_i^n \cap \mathcal{B}_{n,j} \neq \emptyset}} \beta_i Q_i^n \right\rangle_\Omega ,
\end{aligned} \tag{A.8}$$

where we have used again that $\boldsymbol{V}^n \in \mathbb{V}^n_{\operatorname{div}}$. Recall that by Lemma 2.29 for any $n, j \in \mathbb{N}$ there is a subset $\mathcal{T}_{n,j} \subset \mathcal{T}_n$ such that

$$\mathcal{B}_{n,j} = \operatorname{int}\left(\bigcup \{K : \; K \in \mathcal{T}_{n,j}\} \right).$$

Hence, whenever $\operatorname{supp} Q_i^n \cap \mathcal{B}_{n,j} \neq \emptyset$, then there exists $K \in \mathcal{T}_{n,j} \subset \mathcal{T}_n$, such that $K \subset \overline{\mathcal{B}_{n,j}}$ and $\operatorname{supp} Q_i^n \cap K \neq \emptyset$, because $\mathcal{B}_{n,j}$ is open. Therefore it follows that the integrand is supported on the set

$$\bigcup \{\operatorname{supp} Q_i^n : \; \operatorname{supp} Q_i^n \cap \mathcal{B}_{n,j} \neq \emptyset\} \subset \bigcup \{\omega_n(K) : \; K \in \mathcal{T}_n, K \subset \overline{\mathcal{B}_{n,j}}\} =: \omega_n(\mathcal{B}_{n,j}),$$

because the basis of $\mathbb{Q}^n$ is assumed to be locally supported by Assumption 2.20. Due to the regularity of the triangulation we can estimate

$$|\omega_n(\mathcal{B}_{n,j})| \leq c\, |\mathcal{B}_{n,j}| ,$$

where the constant is independent of n and $j \in \mathbb{N}$. With Lemma 2.29 (ii) this yields

$$\left\| \mathbb{1}_{\omega_n(\mathcal{B}_{n,j})} \right\|_{L^p(\Omega)} = |\omega_n(\mathcal{B}_{n,j})|^{1/p} \leq c\, |\mathcal{B}_{n,j}|^{1/p} \leq \frac{c}{\lambda_{n,j}} 2^{-\frac{j}{p}}. \tag{A.9}$$

Also note that for $Q = \sum_{i=1}^{d_n} \beta_i Q_i^n$ the norms

$$Q \mapsto \|Q\|_{L^{p'}(\Omega)} \quad \text{and} \quad Q \mapsto \left(\sum_{i=1}^{d_n} |\beta_i|^{p'} \, \|Q_i^n\|^{p'}_{L^{p'}(\Omega)} \right)^{1/p'}$$

are equivalent and the constants only depend on the shape-regularity and the dimension of $\widehat{\mathbb{P}}_{\mathbb{Q}}$, but are independent of $n \in \mathbb{N}$. Then, in (A.8) we can further estimate

$$\left\langle \operatorname{div} \boldsymbol{V}^{n,j}, \sum_{\substack{i \in \{1,\dots,d_n\}: \\ \operatorname{supp} Q_i^n \cap \mathcal{B}_{n,j} \neq \emptyset}} \beta_i Q_i^n \right\rangle$$

$$\leq \left\| \mathbb{1}_{\omega_n(\mathcal{B}_{n,j})} \right\|_{\mathrm{L}^p(\Omega)} \left\| \operatorname{div} \boldsymbol{V}^{n,j} \right\|_{\mathrm{L}^\infty(\Omega)} \left\| \sum_{\substack{i\in\{1,\ldots,d_n\}:\\ \operatorname{supp} Q_i^n \cap \mathcal{B}_{n,j}\neq\emptyset}} \beta_i Q_i^n \right\|_{\mathrm{L}^{p'}(\Omega)} \tag{A.10}$$
$$\overset{(A.9)}{\leq} \frac{c}{\lambda_{n,j}} 2^{-\frac{j}{p}} \left\| \nabla \boldsymbol{V}^{n,j} \right\|_{\mathrm{L}^\infty(\Omega)} \|Q\|_{\mathrm{L}^{p'}(\Omega)} \leq c 2^{-\frac{j}{p}} \|Q\|_{\mathrm{L}^{p'}(\Omega)},$$

where in the last step we have used the property (iii) of the discrete Lipschitz approximation in Lemma 2.29. Applying (A.10) in (A.8) yields

$$\left\langle \operatorname{div} \boldsymbol{V}^{n,j}, Q \right\rangle_\Omega \overset{(A.8)}{=} \left\langle \operatorname{div} \boldsymbol{V}^{n,j}, \sum_{\substack{i\in\{1,\ldots,d_n\}:\\ \operatorname{supp} Q_i^n \cap \mathcal{B}_{n,j}\neq\emptyset}} \beta_i Q_i^n \right\rangle_\Omega \overset{(A.10)}{\leq} c 2^{-\frac{j}{p}} \|Q\|_{\mathrm{L}^{p'}(\Omega)}, \tag{A.11}$$

and using this in (A.7) shows that

$$\left\| \boldsymbol{W}^{n,j} - \boldsymbol{V}^{n,j} \right\|_{\mathrm{W}^{1,p}(\Omega)} \overset{(A.7)}{\leq} c \sup_{Q\in\mathbb{Q}^n\setminus\{0\}} \frac{\left\langle \operatorname{div} \boldsymbol{V}^{n,j}, Q \right\rangle_\Omega}{\|Q\|_{\mathrm{L}^{p'}(\Omega)}} \overset{(A.11)}{\leq} c 2^{-\frac{j}{p}},$$

which proves (ii).

Finally for (iii) let $j \in \mathbb{N}$ be fixed. Note that Lemma 2.29 (iv) implies in particular that $\boldsymbol{V}^{n,j} \rightharpoonup \mathbf{0}$ weakly in $\mathrm{W}_0^{1,s}(\Omega)^d$, as $n \to \infty$, for any $s \in (1,\infty)$. Then, by the property (A.6) of the discrete Bogovskiĭ operator in Corollary A.6, we have that $\mathfrak{B}^n(\operatorname{div} \boldsymbol{V}^{n,j}) \rightharpoonup \mathbf{0}$ weakly in $\mathrm{W}_0^{1,s}(\Omega)^d$, as $n \to \infty$, for any $s \in (1,\infty)$. By the definition of $\boldsymbol{W}^{n,j}$ it follows that

$$\boldsymbol{W}^{n,j} \rightharpoonup \mathbf{0} \quad \text{weakly in } \mathrm{W}_0^{1,s}(\Omega)^d, \text{ as } n \to \infty,$$

for any $s \in (1,\infty)$. In particular this holds for $s > d$ and thus, with the compact embedding, we have (up to a subsequence) that

$$\boldsymbol{W}^{n,j} \to \mathbf{0} \quad \text{strongly in } \mathrm{L}^s(\Omega)^d, \text{ as } n \to \infty,$$

for any $s \in [1,\infty)$. □

A.2.2. Examples of Finite Element Spaces and Projectors Π^n

In this section we want to provide the details on some of the pairs of finite element spaces, for which we stated in Examples 2.26 and 2.28 that they satisfy the assumptions in Section 2.2.1. More specifically, we aim to present the construction and investigate the properties of a divergence-preserving projection Π^n as defined in Assumption 2.23. Some of this can be found in the literature and is usually used to prove a discrete inf-sup condition, see for example [BBF13, Sec. 8.4.1].

Let Assumption 2.18 on the domain $\Omega \subset \mathbb{R}^d$ and on the family of partitions $\{\mathcal{T}_n\}_{n\in\mathbb{N}}$ be satisfied. Further, let $(\mathbb{V}^n, \mathbb{Q}^n)$, for $n \in \mathbb{N}$, be as introduced in (2.30) and (2.31), respectively, and let $r \in \mathbb{N}$ be maximal, such that $\mathbb{V}_r^n \subset \mathbb{V}^n$.

The following properties are sufficient for Assumption 2.23 to be satisfied: There exists a projection operator $\Pi^n \colon \mathrm{W}_0^{1,1}(\Omega)^n \to \mathbb{V}^n$ such that

(i) (preservation of the divergence in $(\mathbb{Q}^n)'$) for any $\boldsymbol{v} \in \mathrm{W}_0^{1,1}(\Omega)^d$ 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) (local approximation property with level $\ell \in \mathbb{N}$) there exists a constant $\ell \in \mathbb{N}$ such that for any $p \in [1, \infty)$, any $\mu \in \{0, 1\}$ and for any $s \in \{1, \ldots, r+1\}$ there exists a constant $c > 0$ (independent of K and n) such that

$$|\boldsymbol{v} - \Pi^n \boldsymbol{v}|_{\mathrm{W}^{\mu,p}(K)} \leq c h_K^{s-\mu} |\boldsymbol{v}|_{\mathrm{W}^{s,p}(\omega_n^\ell(K))}$$

for all $K \in \mathcal{T}_n$, all $n \in \mathbb{N}$ and all $\boldsymbol{v} \in \mathrm{W}_0^{1,p}(\Omega)^d \cap \mathrm{W}^{s,p}(\Omega)^d$.

Note that (ii) with $\ell \in \mathbb{N}$ implies Assumption 2.23 (iib) with level ℓ.

In the following we want to present some examples of finite element spaces and the construction of Π^n satisfying the assumptions (i)–(ii).

First we will summarise a general approach for the construction, which can be found in [BBF13, Sec. 8.4.1]. Then we will first focus on lower order spaces for which the construction is comparatively simple, see Subsection A.2.2.1. In Subsection A.2.2.2 we will present some higher order spaces.

General Construction

In a number of cases Π^n can be constructed by first using a local projector to $\mathbb{V}^n$ and then correcting the discrete divergence by means of a local correction mapping. For the local projector we choose $\Pi_r^n \colon \mathrm{W}_0^{1,1}(\Omega)^d \to \mathbb{V}_r^n \subset \mathbb{V}^n$, given by Lemma 2.25. With a correction mapping $\Pi_c^n \colon \mathrm{W}_0^{1,1}(\Omega)^d \to \mathbb{V}^n$ we can then define $\Pi^n \colon \mathrm{W}_0^{1,1}(\Omega)^d \to \mathbb{V}^n$ by

$$\Pi^n(\boldsymbol{v}) = \Pi_r^n(\boldsymbol{v}) + \Pi_c^n(\boldsymbol{v} - \Pi_r^n(\boldsymbol{v})). \tag{A.12}$$

Note that if $\Pi_c^n(\mathbf{0}) = \mathbf{0}$, then Π^n is a projection operator, since Π_r^n is a projection operator on $\mathbb{V}^n$. The local correction mapping Π_c^n is chosen such that

$$\langle \operatorname{div} \Pi_c^n \boldsymbol{v}, Q \rangle_\Omega = \langle \operatorname{div} \boldsymbol{v}, Q \rangle_\Omega \qquad \text{for all } Q \in \mathbb{Q}^n, \tag{A.13}$$

for $n \in \mathbb{N}$ and $\boldsymbol{v} \in \mathrm{W}_0^{1,1}(\Omega)^d$. Then, Π^n as defined in (A.12) satisfies (i).

According to Lemma 2.25 $P_r^n \colon \mathrm{W}_0^{1,1}(\Omega)^d \to \mathbb{V}_r^n$ satisfies the approximation property (ii) with $\ell = 1$. For the correction mapping Π_c^n we assume the following local $\mathrm{W}^{1,1}$-stability with level $\ell - 1$, $\ell \in \mathbb{N}$: There exists a constant $\widetilde{c}_1 > 0$ (independent of K and n) such that

$$⨍_K |\Pi_c^n \boldsymbol{v}| \, \mathrm{d}\boldsymbol{x} \leq \widetilde{c}_1 ⨍_{\omega_n^{\ell-1}(K)} (|\boldsymbol{v}| + h_K |\nabla \boldsymbol{v}|) \, \mathrm{d}\boldsymbol{x}, \tag{A.14}$$

for all $\boldsymbol{v} \in \mathrm{W}_0^{1,1}(\Omega)^d$, all $K \in \mathcal{T}_n$ and all $n \in \mathbb{N}$. Note that this implies automatically that $\Pi_c^n(\mathbf{0}) = \mathbf{0}$. In Remark 2.24 (iii) we have seen that (A.14) implies the corresponding estimate in $\mathrm{L}^p(K)$ for $p \in [1, \infty]$. This in turn allows us to verify the local approximation property (ii) with level $\ell \in \mathbb{N}$. Indeed, by the construction of Π^n in (A.12), an inverse estimate, then applying the stability in (A.14) with $p \in [1, \infty)$ and finally using the local approximation property of P_r^n and regularity of $\{\mathcal{T}_n\}_{n \in \mathbb{N}}$ (cf.(2.28)) we obtain that

$$\begin{aligned}
|\boldsymbol{v} - \Pi^n \boldsymbol{v}|_{\mathrm{W}^{\mu,p}(K)} &\leq |\boldsymbol{v} - \Pi_r^n \boldsymbol{v}|_{\mathrm{W}^{\mu,p}(K)} + |\Pi_c^n(\boldsymbol{v} - \Pi_r^n \boldsymbol{v})|_{\mathrm{W}^{\mu,p}(K)} \\
&\leq c \, |\boldsymbol{v} - \Pi_r^n \boldsymbol{v}|_{\mathrm{W}^{\mu,p}(K)} + c h_K^{-\mu} \, \|\Pi_c^n(\boldsymbol{v} - \Pi_r^n \boldsymbol{v})\|_{\mathrm{L}^p(K)} \\
&\leq c h_K^{-\mu} \, \|\boldsymbol{v} - \Pi_r^n \boldsymbol{v}\|_{\mathrm{L}^p(\omega_n^{\ell-1}(K))} + c h_K^{1-\mu} \, |\boldsymbol{v} - \Pi_r^n \boldsymbol{v}|_{\mathrm{W}^{1,p}(\omega_n^{\ell-1}(K))} \\
&\leq c h_K^{s-\mu} \, |\boldsymbol{v}|_{\mathrm{W}^{s,p}(\omega_n^\ell(K))} \, ,
\end{aligned}$$

for $\mu \in \{0, 1\}$, $s \in \{1, \ldots, r+1\}$ and $p \in [1, \infty)$. These observations can be summarised as follows.

Lemma A.7 (Reduction to Π_c^n)**.**
Let $(\mathbb{V}^n, \mathbb{Q}^n)$, for $n \in \mathbb{N}$ be a sequence of finite element spaces as defined in (2.30), (2.31) and assume that $r \in \mathbb{N}$ is maximal such that $\widehat{\mathcal{P}}_r^d \subset \mathbb{P}_{\mathbb{V}}$. Furthermore, let $\ell \in \mathbb{N}$ be given and independent of n. Assume that there exists a mapping $\Pi_c^n \colon \mathrm{W}_0^{1,1}(\Omega)^d \to \mathbb{V}^n$ such that (A.13) is satisfied and additionally we have that the local $\mathrm{W}^{1,1}$-stability in (A.14) is satisfied with level $\ell - 1$.
Then, Π^n as defined in (A.12) is a projection operator satisfying (i) and (ii) with level ℓ and in particular Assumption 2.23.

A.2.2.1 Low Order Examples

For many pairs of finite element spaces with low order a correction mapping Π_c^n is simple to construct. This is based on the fact that for $\mathbb{Q}^n$ consisting of piecewise polynomial function of low enough order, the condition (A.13) can be simplified. If $\mathbb{Q}^n$ consists of piecewise constant functions, then the condition

$$\int_K \operatorname{div} \Pi_c^n \boldsymbol{v} \, \mathrm{d}\boldsymbol{x} = \int_K \operatorname{div} \boldsymbol{v} \, \mathrm{d}\boldsymbol{x} \qquad \text{for all } K \in \mathcal{T}_n, \tag{A.15}$$

for $n \in \mathbb{N}$ and $\boldsymbol{v} \in \mathrm{W}_0^{1,1}(\Omega)^d$, implies (A.13). If the pressure space is continuous, i.e., $\mathbb{Q}^n \subset \mathrm{C}(\overline{\Omega})$, the condition (A.13) is by integration by parts equivalent to

$$\langle \Pi_c^n \boldsymbol{v}, \nabla Q \rangle_\Omega = \langle \boldsymbol{v}, \nabla Q \rangle_\Omega \qquad \text{for all } Q \in \mathbb{Q}^n,$$

for any $\boldsymbol{v} \in \mathrm{W}_0^{1,1}(\Omega)^d$. If $\mathbb{Q}^n$ consists of continuous, piecewise affine functions a sufficient condition for this to hold is that

$$\int_K \Pi_c^n \boldsymbol{v} \, \mathrm{d}\boldsymbol{x} = \int_K \boldsymbol{v} \, \mathrm{d}\boldsymbol{x} \qquad \text{for all } K \in \mathcal{T}_n, \tag{A.16}$$

for any $\boldsymbol{v} \in \mathrm{W}_0^{1,1}(\Omega)^d$, since the gradients of functions in $\mathbb{Q}^n$ are piecewise constant. The construction of Π_c^n such that (A.15) or (A.16) are satisfied, has to be tailored to the specific choice of $\mathbb{V}^n$. We will see that for low order pressure spaces this can be done such that the local stability property (A.14) is satisfied with level 0, which corresponds to $\ell = 1$.
By Lemma A.7 one can show the following lemma.

Lemma A.8 (Reduction to Π_c^n for Low Order $\mathbb{Q}^n$)**.**
Let $(\mathbb{V}^n, \mathbb{Q}^n)$, for $n \in \mathbb{N}$ be a sequence of finite element spaces as defined in (2.30), (2.31) and assume that $r \in \mathbb{N}$ is maximal such that $\widehat{\mathcal{P}}_r^d \subset \mathbb{P}_{\mathbb{V}}$. Assume that there exists a mapping $\Pi_c^n \colon \mathrm{W}_0^{1,1}(\Omega)^d \to \mathbb{V}^n$ such that the local $\mathrm{W}^{1,1}$-stability in (A.14) is satisfied with $\ell = 1$ (this corresponds to level 0) and additionally we have one of the following:

- *(piecewise constant pressure) (A.15) holds, if $\mathbb{Q} = \mathrm{L}^\infty(\Omega)$ and $\widehat{\mathbb{P}}_{\mathbb{Q}} = \widehat{\mathcal{P}}_0$;*
- *(continuous piecewise affine pressure) (A.16) holds, if $\mathbb{Q} = \mathrm{C}(\overline{\Omega})$ and $\widehat{\mathbb{P}}_{\mathbb{Q}} = \widehat{\mathcal{P}}_1$.*

Then, Π^n as constructed in (A.12) is a projection operator satisfying (i) and (ii) with level $\ell = 1$ and in particular Assumption 2.23 with $\ell = 1$.

In the rest of this subsection we will review some examples of finite element spaces for which one can verify the assumptions on Π_c^n.
Whenever the normal vector of an element K enters the definition of the velocity space $\mathbb{V}^n$ in (2.30) one has to replace the pullback by $\mathbf{F}_K^{-1}$, by the Piola transformation (see [BBF13, Ch. 2.1]). Here, instead we directly specify the local spaces $\mathbb{P}_{\mathbb{V}}(K)$ on K and then de-

fine

$$\mathbb{V}^n := \{\boldsymbol{V} \in \mathbb{V}\colon\ \boldsymbol{V}|_K \in \mathbb{P}_{\mathbb{V}}(K) \ \text{ for all } K \in \mathcal{T}_n \ \text{ and } \boldsymbol{V}|_{\partial\Omega} = \boldsymbol{0}\}. \tag{A.17}$$

Note that those two variants are equivalent but (A.17) involves less notation. Let us also denote by $\mathcal{P}_r(K)$ the space of all polynomials on K with degree $\le r$.

For a fixed element $K \in \mathcal{T}_n$, $n \in \mathbb{N}$, let $\{\boldsymbol{x}_i\}_{i\in\{1,\dots,d+1\}}$ be the set of vertices of K and let $\{F_i\}_{i\in\{1,\dots,d+1\}}$ be the set of (relatively closed) faces of K (with indexing such that $\boldsymbol{x}_i$ is not contained in the face F_i). Further, let $\boldsymbol{\nu}_i \in \mathbb{R}^d$ be the outward unit normal of the face F_i for each $i \in \{1,\dots,d+1\}$, and let $\{\lambda_i\}_{i\in\{1,\dots,d+1\}}$ be the set of (scalar) barycentric coordinates, satisfying $\lambda_i(\boldsymbol{x}_j) = \delta_{i,j}$, $i,j \in \{1,\dots,d+1\}$. Note that this implies that $\lambda_i|_{F_i} \equiv 0$ for all $i \in \{1,\dots,d+1\}$. In the following we use the indices modulo $d+1$, if $i > d+1$.

For later reference let us introduce the bubble functions and local bubble spaces

$$b := \prod_{j=1}^{d+1} \lambda_j, \qquad B(K) := \operatorname{span}\{b\} \subset \mathcal{P}_{d+1}(K), \tag{A.18}$$

$$b_i := \prod_{j=1, j\neq i}^{d+1} \lambda_j, \qquad B_F(K) := \operatorname{span}\{b_i\colon\ i \in \{1,\dots,d+1\}\} \subset \mathcal{P}_d(K), \tag{A.19}$$

$$\boldsymbol{b}_i := \boldsymbol{\nu}_i p_i, \qquad \boldsymbol{B}_\nu(K) = \operatorname{span}\{\boldsymbol{b}_i\colon\ i \in \{1,\dots,d+1\}\} \subset \mathcal{P}_d(K)^d, \tag{A.20}$$

We denote by $X \oplus Y$ the direct sum between two function spaces X, Y.

The $\mathbb{P}_2 - \mathbb{P}_0$ Element

The $\mathbb{P}_2 - \mathbb{P}_0$ element for $d = 2$ arises from the choice $\mathbb{P}_{\mathbb{V}}(K) = \mathcal{P}_2(K)^2$ $(r = 2)$ and $\mathbb{V} = \mathrm{C}(\overline{\Omega})^2$ for the velocity space, defined in (A.17) and $\widehat{\mathbb{P}}_{\mathbb{Q}} = \widehat{\mathcal{P}}_0$ and $\mathbb{Q} = \mathrm{L}^\infty(\Omega)$ for the pressure space, defined in (2.31) resulting in discontinuous piecewise constants pressure functions and $\mathbb{V}^n = \mathbb{V}_2^n$, see [BBF13, Sec. 8.4.3]. As in the construction of Π^n in the proof of Prop. 8.4.3 in [BBF13] we consider the mapping $\Pi_c^n\colon \mathrm{W}_0^{1,1}(\Omega)^2 \to \mathbb{V}_2^n$, defined by

$$\Pi_c^n \boldsymbol{v}|_K(\boldsymbol{x}_i) = \boldsymbol{0}, \tag{A.21}$$

$$\int_{F_i} \Pi_c^n \boldsymbol{v} \,\mathrm{d}\sigma = \int_{F_i} \boldsymbol{v} \,\mathrm{d}\sigma, \tag{A.22}$$

for all $i \in \{1,2,3\}$. The so defined interpolation is unique and continuous. Taking the scalar product of (A.22) with the (constant) normal vector $\boldsymbol{\nu}_i$, and summing over $i \in \{1,2,3\}$ shows by Gauß theorem that (A.15) is satisfied. By construction for fixed $K \in \mathcal{T}_n$ we can see that $\Pi_c^n|_K$ maps into $B_F(K)^2$, hence there exists $\{\boldsymbol{\alpha}_i\}_{i\in\{1,2,3\}} \subset \mathbb{R}^2$ such that

$$\Pi_c^n(\boldsymbol{v})|_K = \sum_{j=1}^{3} \boldsymbol{\alpha}_j b_j = \sum_{j=1}^{3} \boldsymbol{\alpha}_j \lambda_{j+1}\lambda_{j+2}.$$

Since $\lambda_k|_{F_k} \equiv 0$, $k \in \{1,2,3\}$ we obtain that

$$\boldsymbol{\alpha}_i \int_{F_i} \lambda_{i+1}\lambda_{i+2} \,\mathrm{d}\sigma = \int_{F_i} \Pi_c^n \boldsymbol{v} \,\mathrm{d}\sigma = \int_{F_i} \boldsymbol{v} \,\mathrm{d}\sigma, \quad i \in \{1,2,3\}.$$

If $|\boldsymbol{\alpha}_i| > 0$, since $\lambda_{i+1}\lambda_{i+2} \geq 0$ we have that

$$\begin{aligned}\|\Pi_c^n \boldsymbol{v}\|_{\mathrm{L}^1(F_i)} &= |\boldsymbol{\alpha}_i| \int_{F_i} \lambda_{i+1}\lambda_{i+2}\,\mathrm{d}\sigma = \frac{\boldsymbol{\alpha}_i}{|\boldsymbol{\alpha}_i|}\cdot\boldsymbol{\alpha}_i \int_{F_i} \lambda_{i+1}\lambda_{i+2}\,\mathrm{d}\sigma \\ &\leq \left|\int_{F_i} \boldsymbol{v}\,\mathrm{d}\sigma\right| \leq \|\boldsymbol{v}\|_{\mathrm{L}^1(F_i)}\,,\end{aligned} \tag{A.23}$$

and if $|\boldsymbol{\alpha}_i| = 0$, the same holds trivially. By scaling arguments, summing (A.23) over $i \in \{1,2,3\}$ and applying a trace inequality yields that

$$\|\Pi_c^n \boldsymbol{v}\|_{\mathrm{L}^1(K)} \leq c h_K \|\Pi_c^n \boldsymbol{v}\|_{\mathrm{L}^1(\partial K)} \overset{(\mathrm{A.23})}{\leq} c h_K \|\boldsymbol{v}\|_{\mathrm{L}^1(\partial K)} \leq c\left(\|\boldsymbol{v}\|_{\mathrm{L}^1(K)} + h_K \|\nabla \boldsymbol{v}\|_{\mathrm{L}^1(K)}\right), \tag{A.24}$$

which shows that (A.14) is satisfied with $\ell = 1$ and hence Lemma A.8 applies.

The Bernardi–Raugel Element ($r = 1$)

The Bernardi–Raugel Element for $r = 1$ and $d \in \{2,3\}$ can be defined as follows, see [BR85, Sec. II]: For the definition of the pressure space in (2.31) we choose $\widehat{\mathbb{P}}_{\mathbb{Q}} = \widehat{\mathcal{P}}_0$ and $\mathbb{Q} = \mathrm{L}^\infty(\Omega)$ resulting in a space of discontinuous pressure functions. For the velocity space $\mathbb{V}^n$ we use the definition (A.17) with $\mathbb{V} = \mathrm{C}(\overline{\Omega})^d$ and $\mathbb{P}_{\mathbb{V}}(K) = \mathcal{P}_1(K)^d \oplus \boldsymbol{B}_\nu(K)$, with $\boldsymbol{B}_\nu(K)$ defined in (A.20).

Again we want to use Lemma A.8 to verify the properties of Π^n. Note that (i), (ii) are proved in [BR85] and [GL01] for part of the range for $p \in [1,\infty)$. The correction mapping $\Pi_c^n \colon \mathrm{W}_0^{1,1}(\Omega)^d \to \mathbb{V}^n$ is defined in [BR85, GL01] by

$$\Pi_c^n \boldsymbol{v}|_K(\boldsymbol{x}_i) = \boldsymbol{0}, \tag{A.25}$$

$$\int_{F_i} \Pi_c^n \boldsymbol{v}\cdot\boldsymbol{\nu}_i\,\mathrm{d}\sigma = \int_{F_i} \boldsymbol{v}\cdot\boldsymbol{\nu}_i\,\mathrm{d}\sigma, \tag{A.26}$$

for all $i \in \{1,\dots,d+1\}$. Considering the degrees of freedom one can see that $\Pi_c^n|_K$ maps to $\boldsymbol{B}_\nu(K)$, it is uniquely defined and the global function $\Pi_c^n\boldsymbol{v}$ is continuous. As before, summing over $i \in \{1,\dots,d+1\}$ in (A.26) and applying the Gauß theorem shows that (A.15) is satisfied.

Now let $\{\alpha_i\}_{i\in\{1,\dots,d+1\}} \subset \mathbb{R}$ such that $\Pi_c^n(\boldsymbol{v})|_K = \sum_{l=1}^{d+1} \alpha_l \boldsymbol{b}_l = \sum_{l=1}^{d+1} \alpha_l \boldsymbol{\nu}_l \prod_{j=1,j\neq l}^{d+1} \lambda_j$. As before, noting that $\lambda_i|_{F_i} = 0$, and hence $\prod_{j=1,j\neq l}^{d+1} \lambda_j|_{F_i} = 0$ for all $l \neq i$, it follows from (A.26) with $\boldsymbol{\nu}_i\cdot\boldsymbol{\nu}_i = 1$ that

$$\int_{F_i} \boldsymbol{v}\cdot\boldsymbol{\nu}_i\,\mathrm{d}\sigma = \int_{F_i} \Pi_c^n\boldsymbol{v}\cdot\boldsymbol{\nu}_i\,\mathrm{d}\sigma = \alpha_i \int_{F_i} \prod_{j=1,j\neq i}^{d+1} \lambda_j \boldsymbol{\nu}_i\cdot\boldsymbol{\nu}_i\,\mathrm{d}\sigma = \alpha_i \int_{F_i} \prod_{j=1,j\neq i}^{d+1} \lambda_j\,\mathrm{d}\sigma.$$

Since the integrand on the right-hand side is non-negative, if $\alpha_i \neq 0$ this implies that

$$\begin{aligned}\|\Pi_c^n \boldsymbol{v}\|_{\mathrm{L}^1(F_i)} &= |\alpha_i| \int_{F_i} \prod_{j=1,j\neq i}^{d+1} \lambda_j\,\mathrm{d}\sigma = \frac{|\alpha_i|}{\alpha_i}\alpha_i \int_{F_i} \prod_{j=1,j\neq i}^{d+1} \lambda_j\,\mathrm{d}\sigma \\ &\leq \left|\int_{F_i} \boldsymbol{v}\cdot\boldsymbol{\nu}_i\,\mathrm{d}\sigma\right| \leq \|\boldsymbol{v}\|_{\mathrm{L}^1(F_i)}\,,\end{aligned} \tag{A.27}$$

and if $\alpha_i = 0$, the same holds trivially. Now the same arguments as in (A.24) yield that

(A.14) with $\ell = 1$ is satisfied. Hence Lemma A.8 applies.

The Second Order Bernardi–Raugel Element

The Bernardi–Raugel Element for $d = 3$ and $r = 2$ is an extension of the previous example and was introduced in [BR85, Sec. III]: For the pressure space we choose $\widehat{\mathbb{P}}_{\mathbb{Q}} = \widehat{\mathcal{P}}_1$ and $\mathbb{Q} = \mathrm{L}^\infty(\Omega)$, resulting in discontinuous pressure spaces $\mathbb{Q}^n$. Further, let $\mathbb{V}^n$ be defined by (A.17) choosing $\mathbb{V} = \mathrm{C}(\overline{\Omega})^d$ and

$$\mathbb{P}_{\mathbb{V}}(K) = \mathcal{P}_2(K)^3 \oplus \boldsymbol{B}_\nu(K) \oplus B(K)^3, \tag{A.28}$$

with the face bubble space $\boldsymbol{B}_\nu(K) \subset \mathcal{P}_3(K)^3$ defined in (A.20) and the element bubble space $B(K) \subset \mathcal{P}_4(K)^3$ defined in (A.18).

Let us show that Π^n satisfying (i) and (ii) exists by verifying the conditions in Lemma A.7. The correction mapping $\Pi^n_c \colon \mathrm{W}^{1,1}_0(\Omega)^3 \to \mathbb{V}^n$ was introduced in [BR85, Lem. III.1] and is defined by

$$\begin{aligned} \Pi^n_c \boldsymbol{v}|_K(\boldsymbol{x}_i) &= \boldsymbol{0} && \text{for all } i \in \{1,\dots,4\}, && \text{(A.29)}\\ \Pi^n_c \boldsymbol{v}|_K(\boldsymbol{x}_{ij}) &= \boldsymbol{0} && \text{for all } i,j \in \{1,\dots,4\},\ i<j, && \text{(A.30)}\\ \int_{F_i} \Pi^n_c \boldsymbol{v}\cdot\boldsymbol{\nu}_i \,\mathrm{d}\sigma &= \int_{F_i} \boldsymbol{v}\cdot\boldsymbol{\nu}_i\,\mathrm{d}\sigma && \text{for all } i \in \{1,\dots,4\}, && \text{(A.31)}\\ \int_K x_l \operatorname{div}(\boldsymbol{v} - \Pi^n_c \boldsymbol{v})\,\mathrm{d}\boldsymbol{x} &= 0 && \text{for all } l \in \{1,2,3\}, && \text{(A.32)} \end{aligned}$$

with $\boldsymbol{x}_{ij} = \frac{1}{2}(\boldsymbol{x}_i + \boldsymbol{x}_j)$ the midpoints of the edges $\boldsymbol{x}_i, bx_j$, for $i,j \in \{1,\dots,4\}$, and noting that $\{1, x_1, x_2, x_3\}$ is a basis of $\mathbb{Q}^n|_K$. Again by considering the degrees of freedom we obtain that $\Pi^n_c|_K$ maps to $\boldsymbol{B}_\nu(K) \oplus B(K)^3$ and is uniquely defined. Further, the global function Π^n_c is continuous. By (A.31), (A.32) it follows that Π^n_c satisfies (A.13), since $\mathbb{Q}^n$ consists of discontinuous, piecewise affine functions. Without loss of generality we assume that $\boldsymbol{0} \in K$, otherwise we shift the basis functions $\{x_1, x_2, x_3\}$ by constants $\{c_1, c_2, c_3\} \subset \mathbb{R}$ such that $(c_1, c_2, c_3)^\top \in K$ and the argument follows analogously.

We have seen before, in (A.27) and below, that the part of $\Pi^n_c(\boldsymbol{v})|_K$ mapping to $\boldsymbol{B}_\nu(K)$ determined by (A.31) satisfies an estimate of the form (A.14) with $\ell = 1$. Thus, we only have to consider the part in $B(K)^3$, determined by (A.32). For this let $\boldsymbol{\alpha} \in \mathbb{R}^3$ such that

$$\begin{aligned} \int_K x_l \operatorname{div} \boldsymbol{v}\,\mathrm{d}\boldsymbol{x} &= \int_K x_l \operatorname{div}(\boldsymbol{\alpha} b)\,\mathrm{d}\boldsymbol{x} = -\int_K \nabla x_l \cdot \boldsymbol{\alpha} b\,\mathrm{d}\boldsymbol{x}\\ &= -\int_K \boldsymbol{e}_l \cdot \boldsymbol{\alpha} b\,\mathrm{d}\boldsymbol{x} = -\alpha_l \int_K b\,\mathrm{d}\boldsymbol{x}, \end{aligned} \tag{A.33}$$

for all $l \in \{1,2,3\}$, and the boundary term vanishes because $b|_{\partial K} = 0$. If $\alpha_l \neq 0$, then we can estimate

$$\begin{aligned} |\alpha_l| \int_K b\,\mathrm{d}\boldsymbol{x} &= \frac{|\alpha_l|}{\alpha_l}\alpha_l \int_K b\,\mathrm{d}\boldsymbol{x} \overset{\text{(A.33)}}{=} -\frac{|\alpha_l|}{\alpha_l}\int_K x_l \operatorname{div}\boldsymbol{v}\,\mathrm{d}\boldsymbol{x}\\ &\le \left| \int_K x_l \operatorname{div}\boldsymbol{v}\,\mathrm{d}\boldsymbol{x} \right| \le \|x_l\|_{\mathrm{L}^\infty(K)} \|\operatorname{div}\boldsymbol{v}\|_{\mathrm{L}^1(K)} \le h_K \|\nabla \boldsymbol{v}\|_{\mathrm{L}^1(K)}, \end{aligned} \tag{A.34}$$

where the last step follows from the fact that $\boldsymbol{0} \in K$ and $|x_l| \le \operatorname{diam}(K) = h_K$ for any $\boldsymbol{x} \in K$, $l \in \{1,2,3\}$. If $\alpha_l = 0$ the corresponding estimate follows trivially. Finally the non-negativity of b and the estimate (A.34) yield that

$$\|\boldsymbol{\alpha} b\|_{\mathrm{L}^1(K)} = |\boldsymbol{\alpha}| \int_K b \,\mathrm{d}\boldsymbol{x} \le \sum_{l=1}^{3} |\alpha_l| \int_K b \,\mathrm{d}\boldsymbol{x} \le c h_K \|\nabla \boldsymbol{v}\|_{\mathrm{L}^1(K)}, \tag{A.35}$$

which finishes the proof that Π_c^n, defined in (A.29)–(A.32) satisfies (A.14) with $\ell = 1$. Hence Lemma A.7 applies with $\ell = 1$.

The MINI Element

The MINI element for $d \in \{2, 3\}$ and $r = 1$, see [BBF13, Sec. 8.4.2, 8.7.1], [GL01, p. 50], arises from the choices $\mathbb{Q} = C(\overline{\Omega})$ and $\widehat{\mathbb{P}}_{\mathbb{Q}} = \widehat{\mathcal{P}}_1$, i.e., the pressure space consists of continuous, piecewise affine functions. For the velocity space we let $\mathbb{V}^n$ as defined in (A.17) with $\mathbb{V} = C(\overline{\Omega})^d$ and the local space $\mathbb{P}_{\mathbb{V}}(K) = \mathcal{P}_1(K)^d \oplus B(K)^d$, with $B(K) \subset \mathcal{P}_{d+1}(K)$ as defined in (A.18). The correction mapping $\Pi_c^n \colon \mathrm{W}_0^{1,1}(\Omega)^d \to \mathbb{V}^n$ is defined by

$$\begin{aligned} \Pi_c^n \boldsymbol{v}|_K(\boldsymbol{x}_i) &= \mathbf{0} \qquad \text{for all } i \in \{1, \dots, d+1\}, \\ \int_K \Pi_c^n \boldsymbol{v} \,\mathrm{d}\boldsymbol{x} &= \int_K \boldsymbol{v} \,\mathrm{d}\boldsymbol{x}. \end{aligned}$$

Looking at the degrees of freedom it is obvious that $\Pi_c^n(\boldsymbol{v})|_K \in B(K)^d$ and the projection is uniquely defined. Furthermore, since the bubbles vanish on the boundary of the element, $\Pi_c^n \boldsymbol{v}$ is continuous and satisfies the boundary conditions.

Obviously Π_c^n satisfies (A.16). What is left to show in order to apply Lemma A.8 is the stability in (A.14) with $\ell = 1$. This follow by the proof of local $\mathrm{W}^{1,1}$-stability of Π^n as shown in [BBDR12, App. A.1]. Let us give the direct argument for the sake of completeness: For $\boldsymbol{\alpha} \in \mathbb{R}^d$ such that $\Pi_c^n \boldsymbol{v} = \boldsymbol{\alpha} b$ we have by definition of Π_c^n and by non-negativity of b that

$$\|\Pi_c^n \boldsymbol{v}\|_{\mathrm{L}^1(\Omega)} = |\boldsymbol{\alpha}| \int_K b \,\mathrm{d}\boldsymbol{x} = \frac{1}{|\boldsymbol{\alpha}|} \boldsymbol{\alpha} \cdot \boldsymbol{\alpha} \int_K b \,\mathrm{d}\boldsymbol{x} = \frac{1}{|\boldsymbol{\alpha}|} \boldsymbol{\alpha} \cdot \int_K \boldsymbol{v} \,\mathrm{d}\boldsymbol{x} \le \|\boldsymbol{v}\|_{\mathrm{L}^1(K)}, \tag{A.36}$$

provided that $\boldsymbol{\alpha} \neq \mathbf{0}$. Otherwise the estimate is obvious and (A.14) with $\ell = 1$ is satisfied. Now Lemma A.7 applies.

A.2.2.2 Higher Order Examples

For higher order pressure spaces the construction of a correction mapping Π_c^n satisfying (A.13) is not that easy.

In [GS03, Thm. 2.1] the authors reduce the construction of a divergence preserving projector Π^n with a quasi-local approximation property to the following: the verification of a quasi-local discrete inf-sup condition; and the construction of a projector $\overline{\Pi}^n$, which preserves the divergence in the dual of piecewise constant functions and satisfies the local approximation property in (ii) with level $\ell = 1$. By a *quasi-local* approximation property we mean the following variant of (ii), for families of macro-elements $\{\mathcal{O}_i\}_{i \in \mathcal{I}^n}$ and (larger macro-elements) $\{\widetilde{\Delta}_i\}_{i \in \mathcal{I}^n}$:

(ii') (quasi-local approximation property) for any $p \in [1, \infty)$, any $\mu \in \{0, 1\}$ and for any $s \in \{1, \dots, r+1\}$ there exists a constant $c > 0$ (independent of i and n) such that

$$|\boldsymbol{v} - \Pi^n \boldsymbol{v}|_{\mathrm{W}^{\mu,p}(\mathcal{O}_i)} \le c h_i^{s-\mu} |\boldsymbol{v}|_{\mathrm{W}^{s,p}(\widetilde{\Delta}_i)}$$

for all $i \in \mathcal{I}^n$, all $n \in \mathbb{N}$ and all $\boldsymbol{v} \in \mathrm{W}_0^{1,p}(\Omega)^d \cap \mathrm{W}^{s,p}(\Omega)^d$. Note that $h_i = \max_{K \subset \mathcal{O}_i} h_K$

for $i \in \mathcal{I}^n$ and that $\operatorname{diam}(\widetilde{\Delta}_i) \le c h_i$, for some constant $c > 0$ independent of $i \in \mathcal{I}^n$ and of $n \in \mathbb{N}$.

Note that the quasi-local approximation property implies the local $\mathrm{W}^{1,1}$-stability in (ii) for some level $\ell \in \mathbb{N}$.

The projector $\overline{\Pi}^n$ can then be constructed similarly as Π^n above in (A.12), with Π^n_c satisfying (A.15), since the divergence has to be preserved in the dual of piecewise constant functions only. This projector is constructed in [GS03] for finite element spaces such that $\mathbb{V}^n_r \subset \mathbb{V}^n$ with $r \ge d$, for $d \in \{2,3\}$. Under this condition the face bubble space $B(K)$ defined in (A.19) is contained in the local space $\mathbb{P}_{\mathbb{V}}(K)$ used for the definition of $\mathbb{V}^n$ in (A.17). Hence, for $d = 2$, the projection operator $\overline{\Pi}^n$ is defined as Π^n for the $\mathbb{P}_2 - \mathbb{P}_0$ element, and the analogous is possible for $d = 3$. However, for $d = 3$ and $r = 2$, the number of degrees of freedom is too small to construct a suitable correction mapping Π^n_c.

Let us in the following give a rough sketch of the quasi-local approximation results obtained for the conforming Crouzeix–Raviart element and the Taylor–Hood element by the results in [GS03], before moving on to the Scott–Vogelius element. For the latter we can construct a correction projection Π^n_c from the proof of inf-sup condition in [GS17] under suitable mesh conditions.

The Conforming Crouzeix–Raviart Element ($r \ge 2$)

For $d = 2$ the lowest order conforming Crouzeix–Raviart element (with $r = 2$) arises from the choice $\mathbb{Q} = \mathrm{L}^\infty(\Omega)$ and $\widehat{\mathbb{P}}_{\mathbb{Q}} = \widehat{\mathcal{P}}_1$, i.e., the pressure functions are discontinuous. Furthermore, for the velocity space $\mathbb{V}^n$ as defined in (A.17) we choose $\mathbb{V} = \mathrm{C}(\overline{\Omega})^2$ and $\mathbb{P}_{\mathbb{V}}(K) = \mathcal{P}_2(K)^2 \oplus B(K)^2 \subset \mathcal{P}_3(K)^2$. With b as in (A.18) the higher order local space of element bubbles on K is defined by

$$B_{r+1}(K) \coloneqq b\mathcal{P}_{r-2}(K) \subset \mathcal{P}_{r+1}(K), \tag{A.37}$$

understood as the elementwise product. Then the generalisations for $d = 2$ and $r \ge 2$ can be defined by $\widehat{\mathbb{P}}_{\mathbb{Q}} = \widehat{\mathcal{P}}_{r-1}$ and $\mathbb{P}_{\mathbb{V}}(K) = \mathcal{P}_r(K)^2 \oplus B_{r+1}(K)^2$, see [BBF13, Ex. 8.6.1, Prop. 8.6.2] and [GR86, Ch. II.2.2].

In [GS03, Thm. 3.3] for $r = 2$ it is proved that there exists a projection Π^n satisfying (i) and (ii') and it is mentioned that the proof extends to the generalisations for $r \ge 2$. When exploring the proof one can see that the local inf-sup condition assumed in [GS03, Thm. 2.1] and used in the proof of [GS03, Thm. 3.3] holds for macro-elements $\mathcal{O}_i = K$, see [GR86, Ch. II, Thm. 2.2]. In this case the macro-element $\widetilde{\Delta}_i$ on the right-hand side of the quasi-local estimate in (ii') coincides with $\omega_n^2(K)$. This means that also the local approximation property (ii) with $\ell = 2$ is satisfied.

The Taylor–Hood Element ($r \ge d$)

The family of (generalised) Taylor–Hood elements for $d \in \{2,3\}$ and $r \ge 2$ is given by $(\mathbb{V}^n, \mathbb{Q}^n)$, for $n \in \mathbb{N}$, as defined in (2.30) and (2.31) with $\mathbb{V} = \mathrm{C}(\overline{\Omega})^d$, $\widehat{\mathbb{P}}_{\mathbb{V}} = \widehat{\mathcal{P}}_r^d$ and $\mathbb{Q} = \mathrm{C}(\overline{\Omega})$, $\widehat{\mathbb{P}}_{\mathbb{V}} = \widehat{\mathcal{P}}_{r-1}$, so the pressure space consists of continuous functions, see [BBF13, Sec. 8.8.2] and [GR86, Ch. II.4.2].

The existence of a divergence-preserving projector Π^n satisfying (i) and the quasi-local approximation property (ii') is proved in [GS03, Sec. 3.1, 3.2], provided that if $r \ge d$ and if each element $K \in \mathcal{T}_n$, $n \in \mathbb{N}$ has at least one interior vertex. The results in [GS03, Thm. 3.1, 3.2] use a quasi-local inf-sup condition, formulated for macro-elements $\mathcal{O}_i$ of the form

$$\omega_n(\boldsymbol{x}) = \operatorname{int}\left(\bigcup\{K \in \mathcal{T}_n \colon\ \boldsymbol{x} \in K\}\right), \tag{A.38}$$

for a vertex $\boldsymbol{x}$ in the simplicial partition $\mathcal{T}_n$, $n \in \mathbb{N}$ and proved in [GS03, Lem. 3.1] stated for general $r \geq 2$ and $d \in \{2,3\}$, see also [GR86, Ch. II, Thm. 4.2] and [BBF13, Sec. 8.8.2].

Thus, (ii') is satisfied with $\mathcal{O}_i = \omega_n(\boldsymbol{x})$ and one can see from the proof of [GS03, Thm. 2.1] that $\widetilde{\Delta}_i$ is contained in a three-layer neighbourhood of $\mathcal{O}_i$. Hence, (ii) is satisfied with $\ell = 4$.

The Scott–Vogelius Element for $r \geq 4$

Let $d = 2$ and $r \geq 1$ be given. Then $\mathbb{V}^n$ is defined by (2.30) choosing $\mathbb{V} = \mathrm{C}(\overline{\Omega})^2$ and $\widehat{\mathbb{P}}_{\mathbb{V}} = \widehat{\mathcal{P}}_r^2$, i.e., $\mathbb{V}^n = \mathbb{V}_2^n$. Let us first define the space $\widetilde{\mathbb{Q}}^n$ as in (2.31), with $\mathbb{Q} = \mathrm{L}^\infty(\Omega)$ and $\widehat{\mathbb{P}}_{\mathbb{Q}} = \widehat{\mathcal{P}}_{r-1}$ and the corresponding space $\widetilde{\mathbb{Q}}_0^n$ as defined in (2.33). To ensure that the space $\mathbb{V}^n_{\mathrm{div}}$ of discretely divergence-free finite element functions as in (2.32) consists of exactly divergence-free functions in the sense of Assumption 2.27, the pressure spaces $\mathbb{Q}_0^n$ have to be chosen as a suitable subspaces of $\widetilde{\mathbb{Q}}_0^n$.

First let us introduce the following notation: By $S^n = \{\boldsymbol{x}_i\}_{i \in \mathcal{I}^n}$ we denote the set of vertices in $\mathcal{T}_n$, $n \in \mathbb{N}$. For a fixed $\boldsymbol{x} \in S^n$ let $\{K_1, \dots, K_N\}$ (with $N = N(\boldsymbol{x}) \leq c$ for a constant independent of $\boldsymbol{x}$ and of n) be the set of (distinct) elements such that $K_j \subset \overline{\omega_n(\boldsymbol{x})}$ and let $\{\theta_1, \dots, \theta_N\}$ be the set of angles such that θ_j is the interior angle of K_j at $\boldsymbol{x}$. Furthermore we choose the numbering such that K_j and K_{j+1} share an edge for $j \in \{1, \dots, N-1\}$. If $\boldsymbol{x} \in \partial\Omega$ this means that K_1 and K_N share an edge with the boundary, and otherwise we assume that also K_1 and K_N share an edge and we formally set the index $N + 1 \equiv 1$. Then, the vertex $\boldsymbol{x}$ is called *singular*, if

$$\begin{aligned} \max_{j \in \{1,\dots,N-1\}} |\theta_j + \theta_{j+1} - \pi| &= 0 \text{ if } \boldsymbol{x} \in \partial\Omega, \\ \max_{j \in \{1,\dots,N\}} |\theta_j + \theta_{j+1} - \pi| &= 0 \text{ otherwise.} \end{aligned}$$

We denote by $Z^n \subset S^n$ the set of singular vertices in S^n. As mentioned in [GS17] an interior vertex $\boldsymbol{x}$ can only be singular if $N = 4$ and the edges of $\{K_1, \dots, K_4\}$ are contained in the union of two straight lines through $\boldsymbol{x}$. Note that this implies that no two interior singular vertices can lie on one edge. For more details and figures of singular vertives, see [GS17].

Then, the pressure space is defined by

$$\mathbb{Q}_0^n := \{Q \in \widetilde{\mathbb{Q}}^n\colon\ A_{\boldsymbol{x}}^n(Q) = 0\ \forall \boldsymbol{x} \in Z^n,\ \fint_\Omega Q \,\mathrm{d}\boldsymbol{x} = 0\} \subset \widetilde{\mathbb{Q}}_0^n, \tag{A.39}$$

with $A_{\boldsymbol{x}}^n\colon \widetilde{\mathbb{Q}}_0^n \to \mathbb{R}$ defined by

$$A_{\boldsymbol{x}}^n(Q) = \sum_{j=1}^N (-1)^{N-j} Q|_{K_j}(\boldsymbol{x}), \qquad \text{for } \boldsymbol{x} \in Z^n,\ Q \in \widetilde{\mathbb{Q}}_0^n,\ n \in \mathbb{N}.$$

This family of elements was introduced in [SV85] and inf-sup stability for $p = 2$ was shown under the condition that the triangulation is quasiuniform, as formulated in Assumption 2.19 and that $r \geq 4$. Recently, in [GS17] the assumption of quasiuniformity was removed. The condition involving $A_{\boldsymbol{x}}^n$ is required in the proof that $\operatorname{div} \mathbb{V}^n \subset \mathbb{Q}_0^n$, see [GS17, Lem. 1]. Hence, the pair of spaces satisfies Assumption 2.27.

For our purposes we require also a space $\mathbb{Q}^n$, such that $\mathbb{Q}_0^n$ is the subspace of zero mean value functions in $\mathbb{Q}^n$. However, the two conditions in the definition in (A.39) are in general not independent of each other. Indeed, note that $A_{\boldsymbol{x}}^n(1) = 0$ for all interior singular vertices

$\boldsymbol{x} \in Z^n \backslash \partial\Omega$, but for singular vertices on the boundary this may not be the case.

Assumption A.9 (Singular Boundary Vertices)**.** *Assume that all singular vertices on the boundary $\boldsymbol{x} \in Z^n \cap \partial\Omega$ have an even number of elements in their neighbourhood $\omega_n(\boldsymbol{x})$.*

By the definition of $A^n_{\boldsymbol{x}}$ this assumption is sufficient for the fact that

$$1 \in \mathbb{Q}^n := \{Q \in \widetilde{\mathbb{Q}}^n \colon\ A^n_{\boldsymbol{x}}(Q) = 0\,\forall \boldsymbol{x} \in Z^n\}, \tag{A.40}$$

and then the above defined space $\mathbb{Q}^n_0$ is the subspace of $\mathbb{Q}^n$ with zero mean value.

We want to construct a projection operator $\Pi^n \colon \mathrm{W}^{1,1}_0(\Omega)^d \to \mathbb{V}^n$ satisfying (i) and (ii) with some level $\ell \in \mathbb{N}$. Note that in general $\mathbb{Q}^n$ does not contain all the piecewise polynomial functions of degree $\leq r-1$, i.e., $\widetilde{\mathbb{Q}}^n$, because of the extra condition for singular vertices. Hence, the reduction result in [GS03, Thm. 2.1] does not apply. Consequently, we aim to construct Π^n as in (A.12), and thus want to apply Lemma A.7. With this we only have to construct the mapping $\Pi^n_c \colon \mathrm{W}^{1,1}_0(\Omega)^d \to \mathbb{V}^n$ satisfying (A.13) and local $\mathrm{W}^{1,1}$-stability of a certain level, see (A.14). This involves two steps:

(a) Given $\boldsymbol{v} \in \mathrm{W}^{1,1}_0(\Omega)^d$ we want to find $R \in \mathbb{Q}^n_0$ such that

$$\langle R, Q\rangle_\Omega = \langle \operatorname{div} \boldsymbol{v}, Q\rangle_\Omega \qquad \text{for all } Q \in \mathbb{Q}^n_0, \tag{A.41}$$

and

$$\|R\|_{\mathrm{L}^1(K)} \leq c\,\|\nabla \boldsymbol{v}\|_{\mathrm{L}^1(\omega^\beta_n(K))}\,, \tag{A.42}$$

for some $\beta \in \mathbb{N}$ and a constant $c > 0$, both of which are independent of $\boldsymbol{v} \in \mathrm{W}^{1,1}_0(\Omega)^d$, of $K \in \mathcal{T}_n$ and of $n \in \mathbb{N}$.

(b) Then, given $R \in \mathbb{Q}^n_0$ we want to find $\boldsymbol{V}$ such that

$$\langle \operatorname{div} \boldsymbol{V}, Q\rangle_\Omega = \langle R, Q\rangle_\Omega \qquad \text{for all } Q \in \mathbb{Q}^n_0, \tag{A.43}$$

and

$$\|\boldsymbol{V}\|_{\mathrm{L}^1(K)} \leq c h_K\,\|R\|_{\mathrm{L}^1(\omega^1_n(K))}\,, \tag{A.44}$$

with constant independent of $R \in \mathbb{Q}^n_0$, of $K \in \mathcal{T}_n$ and of $n \in \mathbb{N}$.

Together, setting $\Pi^n_c(\boldsymbol{v}) := \boldsymbol{V}$ this implies that

$$\langle \operatorname{div} \boldsymbol{v}, Q\rangle_\Omega \overset{\text{(A.41)}}{=} \langle R, Q\rangle_\Omega \overset{\text{(A.43)}}{=} \langle \operatorname{div} \boldsymbol{V}, Q\rangle_\Omega = \langle \operatorname{div} \Pi^n_c(\boldsymbol{v}), Q\rangle_\Omega \qquad \text{for all } Q \in \mathbb{Q}^n_0. \tag{A.45}$$

Since $1 \in \mathbb{Q}^n$, this shows (A.13). Furthermore, we have that

$$\|\boldsymbol{V}\|_{\mathrm{L}^1(K)} \overset{\text{(A.44)}}{\leq} c h_K\,\|R\|_{\mathrm{L}^1(\omega^1_n(K))} \overset{\text{(A.42)}}{\leq} c h_K\,\|\nabla \boldsymbol{v}\|_{\mathrm{L}^1(\omega^{\beta+1}_n(K))}\,, \tag{A.46}$$

again with constants independent of $\boldsymbol{v} \in \mathrm{W}^{1,1}_0(\Omega)^d$, of $K \in \mathcal{T}_n$ and of $n \in \mathbb{N}$. This verifies (A.14) with $\beta+1 = \ell-1$. Then Lemma A.7 applies and shows that (i) and (ii) with $\ell = \beta+2$ is satisfied.

The size of β in condition (a) depends on how large the union of elements in a triangulation is, which are "connected" by the condition on the singular vertices in (A.39).

For this, let us define by $\mathcal{M}^n$ a partition of Ω into closed (connected) sets, such that each

$\mathcal{M} \in \mathcal{M}^n$ is the union of some elements in $\{\mathcal{T}_n\}_{n\in\mathbb{N}}$ and each $\mathcal{M}$ is minimal such that

$$\mathbb{1}_{\mathcal{M}} Q \in \mathbb{Q}^n \quad \text{for all } Q \in \mathbb{Q}^n. \tag{A.47}$$

Note that such a partition $\mathcal{M}^n$ can be obtained by stepwise dividing Ω along edges which do not contain singular vertices (by Zorn's lemma this terminates). The size of the *singular neighbourhoods* $\mathcal{M} \in \mathcal{M}^n$ is a measure for how much the singular vertices "connect" the elements in $\mathcal{T}_n$.

Assumption A.10 (Boundedness of Singular Neighbourhoods)**.**
Assume that one of the following holds:

- $Z^n = \emptyset$*, then set* $\beta = 0$*;*
- *there exists* $\beta \in \mathbb{N}$ *such that: For all* $n \in \mathbb{N}$ *and all* $\mathcal{M} \in \mathcal{M}^n$ *there exists an element* $K \in \mathcal{T}_n$ *such that*

$$K \subset \mathcal{M} \subset \overline{\omega_n^\beta(K)}. \tag{A.48}$$

If there exists a (relatively closed) edge e in $\mathcal{T}_n$ and $\boldsymbol{x} \neq \boldsymbol{y}$ such that $\boldsymbol{x}, \boldsymbol{y} \in Z^n \cap e$, then there exists an $\mathcal{M} \in \mathcal{M}^n$ such that $\omega_n(\boldsymbol{x}) \cup \omega_n(\boldsymbol{y}) \subset \mathcal{M}$, and hence $\beta \geq 2$. But note that by the characterisation of interior singularities mentioned above this can only happen if one of the edges lies on the boundary $\partial\Omega$. Thus, if there are no singular vertices on the boundary, then we have that $\beta \leq 1$.

Now we are in the position to formulate the assumptions under which one can show that Π^n exists and satisfies the respective properties.

Lemma A.11 (Π^n for Scott–Vogelius element, $r \geq 4$)**.**
Let $\mathbb{V}^n$ *and* $\mathbb{Q}^n$ *as defined above, with* $r \geq 4$ *and assuming that Assumption A.9 is satisfied. Furthermore, assume that there exists a constant* $\beta \in \mathbb{N}_0$ *with respect to which Assumption A.10 is satisfied.*
Then, there exists a projection operator $\Pi^n \colon \mathrm{W}_0^{1,1}(\Omega)^d \to \mathbb{V}^n$ *such that the properties (i) and (ii) with* $\ell = \beta + 2$ *are satisfied.*

Proof.
Let us start to show (b) and then prove (a) based on Assumption A.10.

(b) The mapping $R \mapsto \boldsymbol{V}$ can be "read off" from the proof of inf-sup stability for $p = 2$ provided that $r \geq 4$ in [GS17], and the corresponding local stability estimates have to be shown for $p = 1$. Let $R \in \mathbb{Q}_0^n$ be arbitrary but fixed. Following the proof of [GS17, Thm. 1] there is a $\boldsymbol{V} = \boldsymbol{V}_1 + \boldsymbol{V}_2 + \boldsymbol{V}_3 \in \mathbb{V}^n$ such that

$$\operatorname{div} \boldsymbol{V} = R \quad \text{in } \Omega, \tag{A.49}$$

which implies (A.43). To show the local $\mathrm{W}^{1,1}$-stability in (A.44) we have to investigate the functions $\boldsymbol{V}_1, \boldsymbol{V}_2$ and $\boldsymbol{V}_3$ in the construction in the proof of [GS17, Thm. 1] closer. More specifically we have to obtain local estimates for $p = 1$, wherever [GS17] contains global estimates for $p = 2$.
By [GS17, Prop. 1], for given $R \in \mathbb{Q}_0^n$ there exists a function $\boldsymbol{V}_1 \in \mathbb{V}_2^n$ such that

$$\int_K \operatorname{div} \boldsymbol{V}_1 \,\mathrm{d}\boldsymbol{x} = \int_K R \,\mathrm{d}\boldsymbol{x} \quad \text{for all } K \in \mathcal{T}_n.$$

This is based on the correction mapping for the Bernardi–Raugel element, and indeed one can see that $\boldsymbol{V}_1$ arises from (A.25), (A.26) by replacing $\int_{F_i} \boldsymbol{v} \cdot \boldsymbol{\nu}_i \,\mathrm{d}\sigma$ for example by $\frac{1}{3}\int_K R\,\mathrm{d}\boldsymbol{x}$. Then the stability proof above shows that

$$\|\boldsymbol{V}_1\|_{\mathrm{L}^1(\partial K)} \le c\,\|R\|_{\mathrm{L}^1(K)}\,,$$

which yields by scaling and norm equivalence (see (A.24)) that

$$\|\boldsymbol{V}_1\|_{\mathrm{L}^1(K)} \le ch_K\,\|\boldsymbol{V}_1\|_{\mathrm{L}^1(\partial K)} \le ch_K\,\|R\|_{\mathrm{L}^1(K)}\,. \tag{A.50}$$

As in [GS17] we set $R_1 = R - \operatorname{div}\boldsymbol{V}_1 \in \mathbb{Q}_0^n$ and apply first Lemma 6 and then Lemma 7 therein to obtain $\boldsymbol{V}_2$: Lemma 6 shows that for each $\boldsymbol{x}\in S^n$ there exists $\boldsymbol{V}_{\boldsymbol{x}} \in \mathbb{V}_4^n$ with specific properties such that in particular $\operatorname{supp}\boldsymbol{V}_{\boldsymbol{x}} \subset \omega_n(\boldsymbol{x})$ and

$$\|\nabla\boldsymbol{V}_{\boldsymbol{x}}\|_{\mathrm{L}^2(\omega_n(\boldsymbol{x}))} \le c\,\|R_1\|_{\mathrm{L}^2(\omega_n(\boldsymbol{x}))}\,.$$

Note that by the regularity of the triangulation according to Assumption 2.18 we have that $\operatorname{diam}(\omega_n(\boldsymbol{x}))$ is comparable to h_K for any $K\subset\omega_n(\boldsymbol{x})$. By Poincaré's inequality and a scaling argument this implies that

$$\|\boldsymbol{V}_{\boldsymbol{x}}\|_{\mathrm{L}^1(\omega_n(\boldsymbol{x}))} \le ch_K\,\|\nabla\boldsymbol{V}_{\boldsymbol{x}}\|_{\mathrm{L}^1(\omega_n(\boldsymbol{x}))} \le ch_K\,\|R_1\|_{\mathrm{L}^1(\omega_n(\boldsymbol{x}))}\,,$$

with constant $c>0$ is independent of $\boldsymbol{z}\in S^n$, of $n\in\mathbb{N}$ and of R_1. In Lemma 7 in [GS17] the authors set $\boldsymbol{V}_2 = \sum_{\boldsymbol{x}\in S^n}\boldsymbol{V}_{\boldsymbol{x}} \in \mathbb{V}_4^n$, and one has that $\boldsymbol{V}_2|_K = \sum_{i=1}^3 \boldsymbol{V}_{\boldsymbol{x}_i}$, where $\{\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3\}$ are the vertices of the element $K\in\mathcal{T}_n$. Consequently, one has that

$$\begin{aligned}\|\boldsymbol{V}_2\|_{\mathrm{L}^1(K)} &\le \sum_{i=1}^3 \|\boldsymbol{V}_{\boldsymbol{x}_i}\|_{\mathrm{L}^1(K)} \le \sum_{i=1}^3 \|\boldsymbol{V}_{\boldsymbol{x}_i}\|_{\mathrm{L}^1(\omega_n(\boldsymbol{x}_i))}\\ &\le ch_K \sum_{i=1}^3 \|R_1\|_{\mathrm{L}^1(\omega_n(\boldsymbol{x}_i))} = ch_K\,\|R_1\|_{\mathrm{L}^1(\omega_n(K))}\,.\end{aligned} \tag{A.51}$$

Finally, setting $R_2 = R_1 - \operatorname{div}\boldsymbol{V}_2 \in \mathbb{Q}_0^n$ and applying Lemma 3 in [GS17] shows the existence of $\boldsymbol{V}_3\in\mathbb{V}_r^n$ such that $\operatorname{div}\boldsymbol{V}_3 = R_2$ in Ω and by [GS17, Prop. 2] one has the local estimate

$$\|\boldsymbol{V}_3\|_{\mathrm{W}^{1,2}(K)} \le c\,\|R_2\|_{\mathrm{L}^2(K)}\,,$$

again with constant independent of K, R_2 and n. Adapting the scaling and using again the finite-dimensionality of the spaces we arrive at

$$\|\boldsymbol{V}_3\|_{\mathrm{L}^1(K)} \le ch_K\,\|R_2\|_{\mathrm{L}^1(K)}\,, \tag{A.52}$$

for all $K\in\mathcal{T}_n$. Now we have all estimates in place and using the inverse estimates

$h_K \|\nabla \boldsymbol{V}_i\|_{L^1(K)} \leq c \|\boldsymbol{V}_i\|_{L^1(K)}$ we obtain

$$\begin{aligned}
\|\boldsymbol{V}\|_{L^1(K)} &\overset{(A.52)}{\leq} \|\boldsymbol{V}_1\|_{L^1(K)} + \|\boldsymbol{V}_2\|_{L^1(K)} + \|\boldsymbol{V}_3\|_{L^1(K)} \\
&\leq \|\boldsymbol{V}_1\|_{L^1(K)} + \|\boldsymbol{V}_2\|_{L^1(K)} + ch_K \|R_2\|_{L^1(K)} \\
&\leq \|\boldsymbol{V}_1\|_{L^1(K)} + \|\boldsymbol{V}_2\|_{L^1(K)} + ch_K \|\nabla \boldsymbol{V}_2\|_{L^1(K)} + ch_K \|R_1\|_{L^1(K)} \\
&\overset{(A.51)}{\leq} \|\boldsymbol{V}_1\|_{L^1(K)} + ch_K \|R_1\|_{L^1(\omega_n(K))} \\
&\leq \|\boldsymbol{V}_1\|_{L^1(K)} + ch_K \|\nabla \boldsymbol{V}_1\|_{L^1(\omega_n(K))} + ch_K \|R\|_{L^1(\omega_n(K))} \\
&\overset{(A.50)}{\leq} ch_K \|R\|_{L^1(\omega_n(K))} ,
\end{aligned}$$

with constant independent of $K \in \mathcal{T}_n$, $R \in \mathbb{Q}_0^n$ and $n \in \mathbb{N}$. This shows (A.44).

(a) The property (A.41) suggests that R is the L^2-projection of $\operatorname{div} \boldsymbol{v}$ to $\mathbb{Q}_0^n$. Hence, for $n \in \mathbb{N}$ let us introduce the L^2-projection $P^n \colon L^2(\Omega) \to \mathbb{Q}^n$, defined by

$$\langle P^n f, Q \rangle_\Omega = \langle f, Q \rangle_\Omega \qquad \text{for all } Q \in \mathbb{Q}^n \tag{A.53}$$

for $f \in L^2(\Omega)$. Since $1 \in \mathbb{Q}^n$ by Assumption A.9, it follows that

$$\int_\Omega P^n f \,\mathrm{d}\boldsymbol{x} = \int_\Omega f \,\mathrm{d}\boldsymbol{x} \qquad \text{for all } f \in L^2(\Omega). \tag{A.54}$$

and hence $P^n(L_0^2(\Omega)) \subset \mathbb{Q}_0^n$.

Let us now show local L^p stability of P^n. By definition of the partition of singular neighbourhoods $\mathcal{M}^n$ we have that the function $\mathbb{1}_{\mathcal{M}} Q \in \mathbb{Q}^n$, for any $\mathcal{M} \in \mathcal{M}^n$ and for each $Q \in \mathbb{Q}^n$. Testing in (A.53) with $\mathbb{1}_{\mathcal{M}} P^n f$ yields for each $K \subset \mathcal{M}$ that

$$\begin{aligned}
\|P^n f\|_{L^2(\mathcal{M})}^2 &= \langle P^n f, P^n f \rangle_{\mathcal{M}} = \langle P^n f, \mathbb{1}_{\mathcal{M}} P^n f \rangle_\Omega \\
&\overset{(A.53)}{=} \langle f, \mathbb{1}_{\mathcal{M}} P^n f \rangle_\Omega \leq \|f\|_{L^2(\mathcal{M})} \|P^n f\|_{L^2(\mathcal{M})} ,
\end{aligned}$$

which implies that

$$\|P^n f\|_{L^2(\mathcal{M})} \leq \|f\|_{L^2(\mathcal{M})} .$$

By Assumption A.10 it follows that

$$\|P^n f\|_{L^2(K)} \leq \|f\|_{L^2(\omega_n^\beta(K))} , \tag{A.55}$$

for all $f \in L^2(\Omega)$, all $K \in \mathcal{T}_n$ and all $n \in \mathbb{N}$. Furthermore, by scaling with $d = 2$, Hölder's inequality and regularity of the triangulation we have

$$\begin{aligned}
\|P^n f\|_{L^\infty(K)} &\leq ch_K^{-d/2} \|P^n f\|_{L^2(K)} \overset{(A.55)}{\leq} ch_K^{-d/2} \|f\|_{L^2(\omega_n^\beta(K))} \\
&\leq ch_K^{-d/2} \left|\omega_n^\beta(K)\right|^{1/2} \|f\|_{L^\infty(\omega_n^\beta(K))} \leq c \|f\|_{L^\infty(\omega_n^\beta(K))} ,
\end{aligned}$$

for all $f \in L^2(\Omega)$, all $K \in \mathcal{T}_n$ and all $n \in \mathbb{N}$. Then, by interpolation one has local L^p stability for all $p \in [2, \infty]$ with $c = c(p)$ and a duality argument allows one to show it also for $p \in [1, \infty]$. Let $\boldsymbol{v} \in W_0^{1,1}(\Omega)$ be arbitrary but fixed, and hence $\operatorname{div} \boldsymbol{w} \in L_0^1(\Omega)$.

Setting $R \coloneqq P^n(\operatorname{div} \boldsymbol{w})$ we have that (A.41) is satisfied and that

$$\|R\|_{\mathrm{L}^1(K)} = \|P^n(\operatorname{div} \boldsymbol{w})\|_{\mathrm{L}^1(K)} \leq c\,\|\nabla \boldsymbol{v}\|_{\mathrm{L}^1(\omega_n^\beta(K))}\,, \tag{A.56}$$

with constant independent of $K \in \mathcal{T}_n$ and $n \in \mathbb{N}$.

This shows that (a) and (b) are satisfied and hence by the above arguments we obtain a mapping $\Pi_c^n \colon \mathrm{W}_0^{1,1}(\Omega)^d \to \mathbb{V}^n$ satisfying (A.13) and A.14 with $\ell - 1 = \beta + 1$. Then Lemma A.7 applies with $\ell = \beta + 2$, which finishes the proof.

□

References

[AA99] G. Alberti and L. Ambrosio. A geometrical approach to monotone functions in $\mathbb{R}^n$. *Mathematische Zeitschrift*, 230(2):259–316, 1999.

[AB06] C. D. Aliprantis and K. C. Border. *Infinite dimensional analysis*. Springer, Berlin, 3rd edition, 2006. A hitchhiker's guide.

[ABF84] D. N. Arnold, F. Brezzi, and M. Fortin. A stable finite element for the Stokes equations. *Calcolo*, 21(4):337–344 (1985), 1984.

[ADM06] G. Acosta, R. G. Durán, and M. A. Muschietti. Solutions of the divergence operator on John domains. *Adv. Math.*, 206(2):373–401, 2006.

[AF88] E. Acerbi and N. Fusco. An approximation lemma for $W^{1,p}$ functions. In *Material instabilities in continuum mechanics (Edinburgh, 1985–1986)*, Oxford Sci. Publ., pages 1–5. Oxford Univ. Press, New York, 1988.

[AF09] J. P. Aubin and H. Frankowska. *Set-valued Analysis.* Modern Birkhäuser Classics. Birkhäuser, 2009. reprint.

[BadVKR11] H. Beirão da Veiga, P. Kaplický, and M. Růžička. Boundary regularity of shear thickening flowss. *J. Math. Fluid Mech.*, 13(3):387–404, 2011.

[BBDR12] L. Belenki, L. C. Berselli, L. Diening, and M. Růžička. On the finite element approximation of p-Stokes systems. *SIAM Journal on Numerical Analysis*, 50(2):373–397, 2012.

[BBF13] D. Boffi, F. Brezzi, and M. Fortin. *Mixed Finite Element Methods and Applications*, volume 44 of *Springer Series in Computational Mathematics.* Springer, 2013.

[BDF12] D. Breit, L. Diening, and M. Fuchs. Solenoidal Lipschitz truncation and applications in fluid mechanics. *Journal of Differential Equations*, 253(6):1910–1942, 2012.

[BDR09] L. C. Berselli, L. Diening, and M. Růžička. Optimal error estimates for a semi-implicit Euler scheme for incompressible fluids with shear dependent viscosities. *SIAM J. Numer. Anal.*, 47(3):2177–2202, 2009.

[BDR10] L. C. Berselli, L. Diening, and M. Růžička. Existence of strong solutions for incompressible fluids with shear dependent viscosities. *J. Math. Fluid Mech.*, 12(1):101–132, 2010.

[BDR15] L. C. Berselli, L. Diening, and M. Růžička. Optimal error estimate for semi-implicit space-time discretization for the equations describing incompressible generalized Newtonian fluids. *IMA J. Numer. Anal.*, 35(2):680–697, 2015.

[BDS13] D. Breit, L. Diening, and S. Schwarzacher. Solenoidal Lipschitz truncation for parabolic PDE's. *Mathematical Models and Methods in Applied Sciences*, 23(14):2671–2700, 2013.

[BE80] M. Bercovier and M. Engelman. A finite-element method for incompressible non-Newtonian flows. *Journal of Computational Physics*, 36(3):313–326, 1980.

[BEKP10] M. Bulíček, F. Ettwein, P. Kaplický, and D. Pražák. On uniqueness and time regularity of flows of power-law like non-Newtonian fluids. *Math. Methods Appl. Sci.*, 33(16):1995–2010, 2010.

[BF13] F. Boyer and P. Fabrie. *Mathematical tools for the study of the incompressible Navier-Stokes equations and related models*, volume 183 of *Applied Mathematical Sciences*. Springer, New York, 2013.

[BGMS09] M. Bulíček, P. Gwiazda, J. Málek, and A. Świerczewska-Gwiazda. On steady flows of incompressible fluids with implicit power-law-like rheology. *Advances in Calculus of Variations*, 2:109–136, February 2009.

[BGMS12] M. Bulíček, P. Gwiazda, J. Málek, and A. Świerczewska-Gwiazda. On unsteady flows of implicitly constituted incompressible fluids. *SIAM Journal on Mathematical Analysis*, 44:2756–2801, 2012.

[BJ14] H.-O. Bae and B. J. Jin. Regularity of non-Newtonian fluids. *Journal of Mathematical Fluid Mechanics*, 16(2):225–241, 6 2014.

[BKP18] M. Bulíček, P. Kaplický, and D. Pražák. Time regularity of flows of non-Newtonian fluids with critical power-law growth. *ArXiv e-prints*, February 2018.

[BL76] J. Bergh and J. Löfström. *Interpolation spaces. An introduction.* Springer-Verlag, Berlin-New York, 1976. Grundlehren der Mathematischen Wissenschaften, No. 223.

[BM89] J. M. Ball and F. Murat. Remarks on Chacon's Biting Lemma. *Proceedings of the American Mathematical Society*, 107(3):655–663, 1989.

[BM16] M. Bulíček and J. Málek. On unsteady internal flows of Bingham fluids subject to threshold slip on the impermeable boundary. In *Recent developments of mathematical fluid mechanics*, Adv. Math. Fluid Mech., pages 135–156. Birkhäuser/Springer, Basel, 2016.

[BMR09] M. Bulíček, J. Málek, and K. R. Rajagopal. Mathematical analysis of unsteady flows of fluids with pressure, shear-rate, and temperature dependent material moduli that slip at solid boundaries. *SIAM J. Math. Anal.*, 41(2):665–707, 2009.

[Bog79] M. E. Bogovskiĭ. Solution of the first boundary value problem for an equation of continuity of an incompressible medium. *Dokl. Akad. Nauk SSSR*, 248(5):1037–1040, 1979.

[Bom06] M. Boman. Estimates for the L_2-projection onto continuous finite element spaces in a weighted L_p-norm. *BIT Numerical Mathematics*, 46(2):249–260, 2006.

[BP07] D. Bothe and J. Prüss. L_P-theory for a class of non-Newtonian fluids. *SIAM J. Math. Anal.*, 39(2):379–421, 2007.

[BR85] C. Bernardi and G. Raugel. Analysis of some finite elements for the Stokes problem. *Math. Comp.*, 44(169):71–79, 1985.

[BS08] S. Brenner and S. Scott. *The Mathematical Theory of Finite Element Methods*, volume 15 of *Texts in Applied Mathematics*. Springer, 3rd edition, 2008.

[Bv15] M. Bulíček and J. Žabenský. Large data existence theory for unsteady flows of fluids with pressure- and shear-dependent viscosities. *Nonlinear Anal.*, 127:94–127, 2015.

[BY14] R. E. Bank and H. Yserentant. On the H^1-stability of the L_2-projection onto finite element spaces. *Numerische Mathematik*, 126(2):361–381, 2014.

[Car07] E. Carelli. Finite element discretization of motions of incompressible generalized Newtonian flows. online, Dec 2007.

[Cen92] B. Cengiz. On the duals of Lebesgue-Bochner L^p spaces. *Proc. Amer. Math. Soc.*, 114(4):923–926, 1992.

[CHP10] E. Carelli, J. Haehnle, and A. Prohl. Convergence analysis for incompressible generalized Newtonian fluid flows with nonstandard anisotropic growth conditions. *SIAM Journal on Numerical Analysis*, 48(1):164–190, 2010.

[Cia02] P. Ciarlet. *The Finite Element Method for Elliptic Problems*. Society for Industrial and Applied Mathematics, 2002.

[Clé75] P. Clément. Approximation by finite element functions using local regularization. *Revue française d'automatique, informatique, recherche opérationnelle. Analyse numérique*, 9(R-2):77–84, 1975.

[CPDMD90] V. Chiadò Piat, G. Dal Maso, and A. Defranceschi. G-convergence of monotone operators. *Annales de l'I.H.P. Analyse non linéaire*, 7(3):123–160, 1990.

[CR73] M. Crouzeix and P.-A. Raviart. Conforming and nonconforming finite element methods for solving the stationary Stokes equations I. *Rev. Française Automat. Informat. Recherche Opérationnelle Sér. Rouge*, 7(R-3):33–75, 1973.

[CR04] C. Carstensen and M. O. Rieger. Young-measure approximations for elastodynamics with non-monotone stress-strain relations. *ESAIM: M2AN*, 38(3):397–418, 2004.

[CT87] M. Crouzeix and V. Thomée. The stability in L_p and W_p^1 of the L_2-projection onto finite element function spaces. *Math. Comp.*, 48(178):521–532, 1987.

[DDW74] J. jr. Douglas, T. Dupont, and L. Wahlbin. The stability in L_q of the L_2-projection into finite element function spaces. *Numerische Mathematik*, 23(3):193–198, 1974.

[Den04] M. M. Denn. Fifty years of non-Newtonian fluid dynamics. *AIChE Journal*, 50(10):2335–2345, 2004.

[DGG07] E. J. Dean, R. Glowinski, and G. Guidoboni. On the numerical simulation of Bingham visco-plastic flow: Old and new results. *Journal of Non-Newtonian Fluid Mechanics*, 142(1):36–62, 2007. Viscoplastic fluids: From theory to application.

[DiB93] E. DiBenedetto. *Degenerate Parabolic Equations*. Springer Verlag, 1993.

[DKS13a] L. Diening, C. Kreuzer, and E. Süli. Finite element approximation of steady flows of incompressible fluids with implicit power-law-like rheology. *SIAM J. Numer. Anal.*, 51(2):984–1015, 2013.

[DKS13b] L. Diening, C. Kreuzer, and E. Süli. Finite element approximation of steady flows of incompressible fluids with implicit power-law-like rheology. *ArXiv e-prints*, 2013. version 3 (28 Oct 2013).

[DL76] G. Duvaut and J.-L. Lions. *Inequalities in mechanics and physics*. Springer-Verlag, Berlin-New York, 1976. Translated from the French by C. W. John, Grundlehren der Mathematischen Wissenschaften, 219.

[DMM98] G. Dal Maso and F. Murat. Almost everywhere convergence of gradients of solutions to nonlinear elliptic systems. *Nonlinear Anal.*, 31(3-4):405–412, 1998.

[DMS08] L. Diening, J. Málek, and M. Steinhauer. On Lipschitz truncations of Sobolev functions (with variable exponent) and their selected applications. *ESAIM: Control, Optimisation and Calculus of Variations*, 14(2):211–232, 3 2008.

[DPR02] L. Diening, A. Prohl, and M. Ružička. On time-discretizations for generalized Newtonian fluids. In *Nonlinear problems in mathematical physics and related topics*, volume 2 of *Int. Math. Ser. (N. Y.)*, pages 89–118. Kluwer/Plenum, New York, 2002.

[DPR06] L. Diening, A. Prohl, and M. Růžička. Semi-implicit Euler scheme for generalized Newtonian fluids. *SIAM J. Numer. Anal.*, 44(3):1172–1190, 2006.

[DR05] L. Diening and M. Růžička. Strong solutions for generalized Newtonian fluids. *J. Math. Fluid Mech.*, 7(3):413–450, 2005.

[DRS10] L. Diening, M. Růžička, and K. Schumacher. A decomposition technique for John domains. *Ann. Acad. Sci. Fenn. Math.*, 35:87–114, 2010.

[DRW10] L. Diening, M. Růžička, and J. Wolf. Existence of weak solutions for unsteady motions of generalized Newtonian fluids. *Ann. Scuola Norm. Sup. Pisa Cl. Sci.*, 9(1):1–46, 2010.

[DSSV17] L. Diening, S. Schwarzacher, B. Stroffolini, and A. Verde. Parabolic Lipschitz truncation and caloric approximation. *Calc. Var. Partial Differential Equations*, 56(4):Art. 120, 27, 2017.

[EG04] A. Ern and J.-L. Guermond. *Theory and Practice of Finite Elements*, volume 159 of *Applied Mathematical Sciences*. Springer-Verlag New York, 1st edition, 2004.

[EG16] A. Ern and J.-L. Guermond. A converse to Fortin's lemma in Banach spaces. *C. R. Math. Acad. Sci. Paris*, 354(11):1092–1095, 2016.

[EJ95] K. Eriksson and C. Johnson. Adaptive finite element methods for parabolic problems II: Optimal error estimates in $L_\infty L_2$ and $L_\infty L_\infty$. *SIAM Journal on Numerical Analysis*, 32(3):706–740, 1995.

[ER12] H. Eberlein and M. Růžička. Existence of weak solutions for unsteady motions of Herschel-Bulkley fluids. *J. Math. Fluid Mech.*, 14(3):485–500, 2012.

[ER18] S. Eckstein and M. Růžička. On the full space-time discretization of the generalized Stokes equations: The Dirichlet case. *SIAM J. Numer. Anal.*, 56(4):2234–2261, 2018.

[ET76] I. Ekeland and R. Temam. *Convex analysis and variational problems*. North-Holland Publishing Co., Amsterdam-Oxford; American Elsevier Publishing Co., Inc., New York, 1976. Translated from the French, Studies in Mathematics and its Applications, Vol. 1.

[FKF13] C. O. Faria and J. Karam-Filho. A regularized-stabilized mixed finite element formulation for viscoplasticity of Bingham type. *Comput. Math. Appl.*, 66(6):975–995, 2013.

[FM03] J. Frehse and J. Málek. Problems due to the no-slip boundary in incompressible fluid dynamics. In S. Hildebrandt and H. Karcher, editors, *Geometric analysis and nonlinear partial differential equations*, pages 559–571. Springer Berlin Heidelberg, Berlin, Heidelberg, 2003.

[FMR05] M. Franta, J. Málek, and K. R. Rajagopal. On steady flows of fluids with pressure- and shear-dependent viscosities. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 461(2055):651–670, 2005.

[FMS97] J. Frehse, J. Málek, and M. Steinhauer. An existence result for fluids with shear dependent viscosity—steady flows. In *Proceedings of the Second World Congress of Nonlinear Analysts, Part 5 (Athens, 1996)*, volume 30, pages 3041–3049, 1997.

[FMS03] J. Frehse, J. Málek, and M. Steinhauer. On analysis of steady flows of fluids with shear-dependent viscosity based on the Lipschitz truncation method. *SIAM Journal on Mathematical Analysis*, 34(5):1064–1083, 2003.

[FMT04] G. Francfort, F. Murat, and L. Tartar. Monotone operators in divergence form with x-dependent multivalued graphs. *Bollettino dell'Unione Matematica Italiana*, 7-B(1):23–59, 2 2004.

[For77] M. Fortin. An analysis of the convergence of mixed finite element methods. *RAIRO Anal. Numér.*, 11(4):341–354, iii, 1977.

[FS00] M. Fuchs and G. Seregin. *Variational methods for problems from plasticity theory and for generalized Newtonian fluids*, volume 1749 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2000.

[Gal11] G. P. Galdi. *An Introduction to the Mathematical Theory of the Navier-Stokes Equations - Steady-State Problems*. Springer Monographs in Mathematics. Springer Verlag New York, 2nd edition, 2011.

[GD03] A. Granas and J. Dugundji. *Fixed point theory*. Springer Monographs in Mathematics. Springer-Verlag, New York, 2003.

[GHS16] F. D. Gaspoz, C.-J. Heine, and K. G. Siebert. Optimal grading of the newest vertex bisection and H^1-stability of the L_2-projection. *IMA J. Numer. Anal.*, 36(3):1217–1241, 2016.

[GL01] V. Girault and J.-L. Lions. Two-grid finite-element schemes for the steady Navier-Stokes problem in polyhedra. *Port. Math. (N.S.)*, 58(1):25–57, 2001.

[GMS07] P. Gwiazda, J. Málek, and A. Świerczewska. On flows of an incompressible fluid with a discontinuous power-law-like rheology. *Computers & Mathematics with Applications*, 53:531–546, 2007. Recent Advances in the Mathematical Analysis of Non-Linear Phenomena.

[GN14a] J. Guzmán and M. Neilan. Conforming and divergence-free Stokes elements in three dimensions. *IMA Journal of Numerical Analysis*, 34(4):1489–1508, 2014.

[GN14b] J. Guzmán and M. Neilan. Conforming and divergence-free Stokes elements on general triangular meshes. *Math. Comp.*, 83:15–36, 2014.

[GR86] V. Girault and P.-A. Raviart. *Finite Element Methods for Navier-Stokes Equations*, volume 5 of *Springer Series in Computational Mathematics*. Springer Berlin Heidelberg, 1986.

[Gri11] P. Grisvard. *Elliptic Problems in Nonsmooth Domains*, volume 69 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics, 2011.

[GS03] V. Girault and L. R. Scott. A quasi-local interpolation operator preserving the discrete divergence. *Calcolo*, 40(1):1–19, 2003.

[GS17] J. Guzmán and R. Scott. The Scott-Vogelius finite elements revisited. *ArXiv e-prints*, April 2017.

[GT01] D. Gilbarg and N. S. Trudinger. *Elliptic Partial Differential Equations of Second Order*. Classics in Mathematics. Springer, 2001. reprint, originally in Grundlehren der mathematischen Wissenschaften, Vol. 224.

[GZG07] P. Gwiazda and A. Zatorska-Goldstein. On elliptic and parabolic systems with x-dependent multivalued graphs. *Math. Methods Appl. Sci.*, 30(2):213–236, 2007.

[HMST17] J. Hron, J. Málek, J. Stebel, and K. Touška. A novel view on computations of steady flows of Bingham fluids using implicit constitutive relations. *Project MORE Preprint*, 2017.

[HR82] J. G. Heywood and R. Rannacher. Finite element approximation of the non-stationary Navier-Stokes problem. I. Regularity of solutions and second-order error estimates for spatial discretization. *SIAM J. Numer. Anal.*, 19(2):275–311, 1982.

[JLM+17] V. John, A. Linke, C. Merdon, M. Neilan, and L. G. Rebholz. On the divergence constraint in mixed finite element methods for incompressible flows. *SIAM Rev.*, 59(3):492–544, 2017.

[JMPT18] A. Janečka, J. Málek, V. Průša, and G. Tierra. Numerical scheme for simulation of transient flows of non-Newtonian fluids characterised by a non-monotone relation between the symmetric part of the velocity gradient and the Cauchy stress tensor. *ArXiv e-prints*, September 2018.

[KL02] J. Kinnunen and J. L. Lewis. Very weak solutions of parabolic systems of p-Laplacian type. *Ark. Mat.*, 40:105–132, 2002.

[KS16] C. Kreuzer and E. Süli. Adaptive finite element approximation of steady flows of incompressible fluids with implicit power-law-like rheology. *ESAIM Math. Model. Numer. Anal.*, 50(5):1333–1369, 2016.

[Lad67] O. A. Ladyženskaja. New equations for the description of the motions of viscous incompressible fluids, and global solvability for their boundary value problems. *Trudy Mat. Inst. Steklov.*, 102:85–104, 1967.

[Lad69] O. A. Ladyženskaja. *The mathematical theory of viscous incompressible flow.* Second English edition, revised and enlarged. Translated from the Russian by Richard A. Silverman and John Chu. Mathematics and its Applications, Vol. 2. Gordon and Breach, Science Publishers, New York-London-Paris, 1969.

[Lio69] J.-L. Lions. *Quelques Méthodes De Résolution Des Problèmes Aux Limites Non Linéaires.* Dunod, Paris, 1969.

[Lio13] P. L. Lions. *Mathematical topics in fluid mechanics*, volume 1 of *Oxford lecture series in mathematics and its applications ; 3, 10.* Oxford University Press, 2013.

[LR18] A. Linke and L. G. Rebholz. Pressure-induced locking in mixed methods for time-dependent (Navier-)Stokes equations. *ArXiv e-prints*, August 2018.

[M0́7] J. Málek. Mathematical properties of the flows of incompressible fluids with pressure and shear rate dependent viscosities, 2007. DSc. thesis.

[Man82] L. Mansfield. Finite element subspaces with optimal rates of convergence for the stationary Stokes problem. *ESAIM: Mathematical Modelling and Numerical Analysis - Modélisation Mathématique et Analyse Numérique*, 16(1):49–66, 1982.

[Max67] J. C. Maxwell. On the dynamical theory of gases. *Philosophical Transactions of the Royal Society of London*, 157:49–88, 1867.

[Min62] G. J. Minty. Monotone (nonlinear) operators in Hilbert space. *Duke Math. J.*, 29(3):341–346, Sept 1962.

[Min63] G. J. Minty. On a monotonicity method for the solution of non-linear equations in Banach spaces. *Proc. Nat. Acad. Sci. U.S.A.*, 50:1038–1041, 1963.

[MKYJ17] R. Mahmood, N. Kousar, M. Yaqub, and K. Jabeen. 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. Article ID 5210708.

[MLC16] E. Moreno, A. Larese, and M. Cervera. 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, 2016.

[MNR01] J. Málek, J. Nečas, and M. Růžička. On weak solutions to a class of non-Newtonian incompressible fluids in bounded three-dimensional domains: the case $p \geq 2$. *Adv. Differential Equations*, 6(3):257–302, 2001.

[MNRR96] J. Málek, J. Nečas, M. Rokyta, and M. Růžička. *Weak and measure-valued solutions to evolutionary PDEs*, volume 13 of *Applied Mathematics and Mathematical Computation.* Chapman & Hall, London, 1996.

[Mou16] A. Moussa. Some variants of the classical Aubin-Lions lemma. *J. Evol. Equ.*, 16(1):65–93, 2016.

[MP16] J. Málek and V. Průša. Derivation of equations for continuum mechanics and thermodynamics of fluids. In Y. Giga and A. Novotny, editors, *Handbook of Mathematical Analysis in Mechanics of Viscous Fluids*, pages 1–70. Springer International Publishing, Cham, 2016.

[MR05] J. Málek and K. R. Rajagopal. Mathematical issues concerning the Navier-Stokes equations and some of its generalizations. In *Evolutionary equations. Vol. II*, Handb. Differ. Equ., pages 371–459. Elsevier/North-Holland, Amsterdam, 2005.

[MRS05] J. Málek, M. Růžička, and V. V. Shelukhin. Herschel-Bulkley fluids: existence and regularity of steady flows. *Math. Models Methods Appl. Sci.*, 15(12):1845–1861, 2005.

[MŽ18] E. Maringová and J. Žabenský. On a Navier-Stokes-Fourier-like system capturing transitions between viscous and inviscid fluid regimes and between no-slip and perfect-slip boundary conditions. *Nonlinear Anal. Real World Appl.*, 41:152–178, 2018.

[Phe97] R. R. Phelps. Lectures on maximal monotone operators. *Extracta Mathematicae*, 12(3):193–230, 1997.

[PR01a] J.-P. Penot and R. Ratsimahalo. On the Yosida approximation of operators. *Proc. Roy. Soc. Edinburgh Sect. A*, 131(4):945–966, 2001.

[PR01b] A. Prohl and M. Ružička. On fully implicit space-time discretization for motions of incompressible fluids with shear-dependent viscosities: the case $p \leq 2$. *SIAM J. Numer. Anal.*, 39(1):214–249 (electronic), 2001.

[Raj03] K. R. Rajagopal. On implicit constitutive theories. *Applications of Mathematics*, 48(4):279–319, 2003.

[Raj06] K. R. Rajagopal. On implicit constitutive theories for fluids. *Journal of Fluid Mechanics*, 550RS.2008:243–249, 3 2006.

[Ran99] R. Rannacher. Finite element methods for the incompressible Navier-Stokes equations. online, Aug 1999.

[Roc70] R. T. Rockafellar. On the maximality of sums of nonlinear monotone operators. *Transactions of the American Mathematical Society*, 149(1):75–88, 1970.

[Rou13] T. Roubíček. *Nonlinear partial differential equations with applications*, volume 153 of *International Series of Numerical Mathematics*. Birkhäuser/Springer Basel AG, Basel, 2nd edition, 2013.

[RS08] K. R. Rajagopal and A. R. Srinivasa. On the thermodynamics of fluids defined by implicit constitutive relations. *Zeitschrift für angewandte Mathematik und Physik*, 59(4):715–729, 2008.

[RS16] K. R. Rajagopal and G. Saccomandi. A novel approach to the description of constitutive relations. *Frontiers in Materials*, 3:36, 2016.

[Růž97] M. Růžička. A note on steady flow of fluids with shear dependent viscosity. *Nonlinear Analysis: Theory, Methods & Applications*, 30(5):3029–3039, 1997. Proceedings of the Second World Congress of Nonlinear Analysts.

[Sch78] W. R. Schowalter. *Mechanics of non-Newtonian fluids*. Pergamon Press Oxford, England ; New York, 1st edition, 1978.

[Ser91] G. A. Serëgin. On a dynamical system generated by two-dimensional equations of motion of Bingham fluid. *Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI)*, 188(Kraev. Zadachi Mat. Fiz. i Smezh. Voprosy Teor. Funktsiĭ. 22):128–142, 187, 1991.

[Ser15] G. Seregin. *Lecture Notes on Regularity Theory for the Navier-Stokes Equations*. World Scientific, 2015.

[Sim87] J. Simon. Compact sets in the space $L^p(0,T;B)$. *Ann. Mat. Pura Appl. (4)*, 146:65–96, 1987.

[ST18] E. Süli and T. Tscherpel. Fully discrete finite element approximation of unsteady flows of implicitly constituted incompressible fluids. *ArXiv e-prints*, April 2018. (accepted at IMA J. Numer. Anal.).

[Sto45] G. G. Stokes. On the theories of the internal friction of fluids in motion and of the equilibrium and motion of elastic solids. *Trans. Cambridge Philos. Soc.*, 8:287–319, 1845.

[SV85] L. R. Scott and M. Vogelius. Norm estimates for a maximal right inverse of the divergence operator in spaces of piecewise polynomials. *RAIRO Modél. Math. Anal. Numér.*, 19(1):111–143, 1985.

[SZ90] L. R. Scott and S. Zhang. Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Mathematics of Computation*, 54(190):483–493, April 1990.

[Tem84] R. Temam. *Navier-Stokes Equations: Theory & Numerical Analysis.* Studies in Mathematics and Its Applications. North-Holland, 1984.

[TH73] C. Taylor and P. Hood. A numerical solution of the Navier-Stokes equations using the finite element technique. *Internat. J. Comput. & Fluids*, 1:73–100, 1973.

[Tru91] C. A. Truesdell. *A first course in rational continuum mechanics. Vol. 1*, volume 71 of *Pure and Applied Mathematics.* Academic Press, Inc., Boston, MA, 2nd edition, 1991. General concepts.

[Wol07] J. Wolf. 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(1):104–138, 2007.

[Zha10] J. Zhang. 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(45-48):3051–3057, 2010.

Remember the LORD, your God,
for it is he who provides you with strength.

Deut. 8:18