FINITE ELEMENT APPROXIMATION OF STEADY FLOWS OF INCOMPRESSIBLE FLUIDS WITH IMPLICIT POWER-LAW-LIKE RHEOLOGY

LARS DIENING*, CHRISTIAN KREUZER†, AND ENDRE SÜLI‡

Abstract. We develop the analysis of finite element approximations of implicit power-law-like models for viscous incompressible fluids. The Cauchy stress and the symmetric part of the velocity gradient in the class of models under consideration are related by a, possibly multi–valued, maximal monotone r-graph, with $1 < r < \infty$. Using a variety of weak compactness techniques, including Chacon's biting lemma and Young measures, we show that a subsequence of the sequence of finite element solutions converges to a weak solution of the problem as the finite element discretization parameter h tends to 0. A key new technical tool in our analysis is a finite element counterpart of the Acerbi–Fusco Lipschitz truncation of Sobolev functions.

Key words. Finite element methods, implicit constitutive models, power-law fluids, convergence, weak compactness, discrete Lipschitz truncation

AMS subject classifications. Primary 65N30, 65N12. Secondary 76A05, 35Q35

1. Introduction. Most physical models describing fluid flow rely on the assumption that the Cauchy stress is an explicit function of the symmetric part of the velocity gradient of the fluid. This assumption leads to the Navier–Stokes equations and its nonlinear generalizations, such as various electrorheological flow models; see, e. g., [Lad69, Lio69, Růž00]. It is known however that the framework of classical continuum mechanics, built upon the notions of current and reference configuration and an explicit constitutive equation for the Cauchy stress, is too narrow to enable one to model inelastic behavior of solid-like materials or viscoelastic properties of materials. Our starting point in this paper is therefore a generalization of the classical framework of continuum mechanics, called the implicit constitutive theory, which was proposed recently in a series of papers by Rajagopal; see, for example, [Raj03, Raj06]. The underlying principle of the implicit constitutive theory in the context of viscous flows is the following: instead of demanding that the Cauchy stress is an explicit function of the symmetric part of the velocity gradient, one may allow an implicit and not necessarily continuous relationship between these quantities. The resulting general theory therefore admits fluid flow models with implicit and possibly discontinuous power-law-like rheology; see, [Mál08, Mál07]. Very recently a rigorous mathematical existence theory was developed for these models by Bulíček, Gwiazda, Málek, and Świerczewska-Gwiazda in [BGMŚG09]. Motivated by the ideas in [BGMŚG09], we consider the construction of finite element approximations of implicit constitutive models for incompressible fluids and we develop the convergence theory of these numerical methods by exploiting a range of weak compactness arguments.

Let $\Omega \subset \mathbb{R}^d$, $d \in \mathbb{N}$, be a bounded open Lipschitz domain with polyhedral boundary. For $r \in (1,\infty)$ we define $r' \in (1,\infty)$ by the equality $\frac{1}{r} + \frac{1}{r'} = 1$ and we set

$$\tilde{r} := \min\{r', r^*/2\}, \qquad \text{where} \qquad r^* := \begin{cases} \frac{dr}{d-r} & \text{if } r < d, \\ \infty & \text{otherwise.} \end{cases} \tag{1.1}$$

*Mathematisches Institut der Universität München, Theresienstrasse 39, D-80333 München, Germany (lars.diening@mathematik.uni-muenchen.de).

†Fakultät für Mathematik, Ruhr-Universität Bochum, Universitätsstrasse 150, D-44801 Bochum, Germany (christan.kreuzer@rub.de).

‡Mathematical Institute, University of Oxford, 24-29 St Giles', Oxford OX1 3LB, UK (endre.suli@maths.ox.ac.uk).

We refer the reader to the first paragraph in Section 2.1 for the definitions of the various function spaces used in the paper and for a list of our notational conventions.

Problem. For $\boldsymbol{f} \in W^{-1,r'}(\Omega)^d$ find $(\boldsymbol{u}, p, \boldsymbol{S}) \in W_0^{1,r}(\Omega)^d \times L_0^{\tilde{r}}(\Omega) \times L^{r'}(\Omega)^{d\times d}$ such that

$$\begin{aligned} \operatorname{div}(\boldsymbol{u} \otimes \boldsymbol{u} + p\mathbf{1} - \boldsymbol{S}) &= \boldsymbol{f} && \text{in } \mathcal{D}'(\Omega), \\ \operatorname{div} \boldsymbol{u} &= 0 && \text{in } \mathcal{D}'(\Omega), \\ (\boldsymbol{D}\boldsymbol{u}(x), \boldsymbol{S}(x)) &\in \mathcal{A}(x) && \text{for almost every } x \in \Omega. \end{aligned} \tag{1.2}$$

The symmetric part of the gradient is defined by $\boldsymbol{D}\boldsymbol{u} := \frac{1}{2}(\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^{\mathrm{T}})$. As is implied by the choice of the solution space $W_0^{1,r}(\Omega)^d$ for the velocity, a homogenous Dirichlet boundary condition is assumed here for $\boldsymbol{u}$.

The implicit law, which relates the shear rate $\boldsymbol{D}\boldsymbol{u}$ to the shear stress $\boldsymbol{S}$, is given by an nonhomogeneous maximal monotone r-graph $\mathcal{A} : x \mapsto \mathcal{A}(x)$. In particular, we assume that the following properties hold for almost every $x \in \Omega$:

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

(A2) For all $(\boldsymbol{\delta}_1, \boldsymbol{\sigma}_1), (\boldsymbol{\delta}_2, \boldsymbol{\sigma}_2) \in \mathcal{A}(x)$,

$$(\boldsymbol{\sigma}_1 - \boldsymbol{\sigma}_2) : (\boldsymbol{\delta}_1 - \boldsymbol{\delta}_2) \geqslant 0 \qquad (\mathcal{A}(x) \text{ is a monotone graph}),$$

and if $\boldsymbol{\delta}_1 \neq \boldsymbol{\delta}_2$ and $\boldsymbol{\sigma}_1 \neq \boldsymbol{\sigma}_2$, then the inequality is strict;

(A3) If $(\boldsymbol{\delta}, \boldsymbol{\sigma}) \in \mathbb{R}^{d\times d}_{\mathrm{sym}} \times \mathbb{R}^{d\times d}_{\mathrm{sym}}$ and

$$(\bar{\boldsymbol{\sigma}} - \boldsymbol{\sigma}) : (\bar{\boldsymbol{\delta}} - \boldsymbol{\delta}) \geqslant 0 \quad \text{for all } (\bar{\boldsymbol{\delta}}, \bar{\boldsymbol{\sigma}}) \in \mathcal{A}(x),$$

then $(\boldsymbol{\delta}, \boldsymbol{\sigma}) \in \mathcal{A}(x)$ (i.e., $\mathcal{A}(x)$ is a maximal monotone graph);

(A4) There exists a constant $c > 0$ and a nonnegative $m \in L^1(\Omega)$, such that for all $(\boldsymbol{\delta}, \boldsymbol{\sigma}) \in \mathcal{A}(x)$ we have

$$\boldsymbol{\sigma} : \boldsymbol{\delta} \geqslant c(|\boldsymbol{\delta}|^r + |\boldsymbol{\sigma}|^{r'}) - m(x) \qquad (\text{i.e., } \mathcal{A}(x) \text{ is an } r\text{-graph});$$

(A5) (i) For all $\boldsymbol{\delta} \in \mathbb{R}^{d\times d}_{\mathrm{sym}} := \{\boldsymbol{\zeta} \in \mathbb{R}^{d\times d} \colon \boldsymbol{\zeta} = \boldsymbol{\zeta}^{\mathrm{T}}\}$, the set

$$\{\boldsymbol{\sigma} \in \mathbb{R}^{d\times d}_{\mathrm{sym}} : (\boldsymbol{\delta}, \boldsymbol{\sigma}) \in \mathcal{A}(x)\}$$

is closed;

(ii) For any closed $\mathcal{C} \subset \mathbb{R}^{d\times d}_{\mathrm{sym}}$, the set

$$\big\{(x, \boldsymbol{\delta}) \in \Omega \times \mathbb{R}^{d\times d}_{\mathrm{sym}} : \text{there exists } \boldsymbol{\sigma} \in \mathcal{C}, \text{ such that } (\boldsymbol{\delta}, \boldsymbol{\sigma}) \in \mathcal{A}(x)\big\}$$

is measurable relative to the smallest σ-algebra $\mathfrak{L}(\Omega) \otimes \mathfrak{B}(\mathbb{R}^{d\times d}_{\mathrm{sym}})$ of the product of the σ-algebra $\mathfrak{L}(\Omega)$ of Lebesgue measurable subsets of Ω and all Borel subsets $\mathfrak{B}(\mathbb{R}^{d\times d}_{\mathrm{sym}})$ of $\mathbb{R}^{d\times d}_{\mathrm{sym}}$.

The class of fluids described by (1.2) is very general and includes not only Newtonian (Navier–Stokes) fluids ($\boldsymbol{S} = 2\mu_* \boldsymbol{D}\boldsymbol{v}$ with μ_* being a positive constant), but also standard power-law fluid models, where $\boldsymbol{S} = 2\mu_* |\boldsymbol{D}\boldsymbol{v}|^{r-2} \boldsymbol{D}\boldsymbol{v}$, $1 < r < \infty$, and their generalizations ($\boldsymbol{S} = 2\tilde{\mu}(|\boldsymbol{D}\boldsymbol{v}|^2) \boldsymbol{D}\boldsymbol{v}$), stress power-law fluid flow models and their generalizations of the form $\boldsymbol{D}\boldsymbol{v} = \alpha(|\boldsymbol{S}|^2) \boldsymbol{S}$, fluids with the viscosity depending on the shear rate and the shear stress $\boldsymbol{S} = 2\hat{\mu}(|\boldsymbol{D}\boldsymbol{v}|^2, |\boldsymbol{S}|^2) \boldsymbol{D}\boldsymbol{v}$, as well as activated fluids, such as Bingham and Herschel–Bulkley fluids. For further details concerning the physical background of the implicit constitutive theory we refer the reader to the papers by Rajagopal and Rajagopal & Srinivasa [Raj03, Raj06, RS08], and the introductory sections of Bulíček, Gwiazda, Málek & Świerczewska-Gwiazda [BGMŚG09, BGMŚG11] and Bulíček, Málek & Süli [BMS12].

It is proved in [BGMŚG09] that under the assumption $r > \frac{2d}{d+2}$ problem (1.2) has a weak solution. The proof in [BGMŚG09] uses a sequence of approximation spaces spanned by finite subsets of a Schauder basis of an infinite-dimensional subspace of a Sobolev space, consisting of exactly divergence-free functions. Since such a Schauder basis is not explicitly available for computational purposes, here, instead, we shall approximate (1.2) from two classes of inf-sup stable pairs of finite element spaces. The first class contains velocity-pressure space-pairs that do not lead to exactly divergence-free velocity approximations. For finite element spaces of this kind our convergence result is restricted to $r > \frac{2d}{d+1}$. In the case of exactly divergence-free finite element spaces for the velocity we show that the resulting (sub)sequence of finite element approximations converges to a weak solution of the problem for any $r > \frac{2d}{d+2}$, as in [BGMŚG09].

The paper is structured as follows. In Section 2 we introduce the necessary analytical tools, including Young measures and Chacon's biting lemma. In Section 3 we define the finite element approximation of the problem with both discretely divergence-free and exactly divergence-free finite element spaces for the velocity. A key technical tool in our analysis is a new *discrete Lipschitz truncation* technique, which can be seen as the finite element counterpart of the Lipschitz truncation of Sobolev functions discovered by Acerbi and Fusco [AF88] and further refined by Diening, Málek, and Steinhauer [DMS08]; see also [DHHR11, BDF12, BDS12]. The central result of the paper is stated in Section 4, in Theorem 19, and concerns the convergence of the finite element approximations constructed in Section 3.

2. Preliminaries. In this section we recall some known results and mathematical tools from the literature. We shall first introduce basic notations and recall some well-known properties of Lebesgue and Sobolev spaces. We shall then discuss the approximation of an x-dependent r-graph by a sequence of regular single-valued tensor fields using a graph-mollification technique by Francfort, Murat and Tartar [FMT04] (which the authors of [FMT04] attribute to Dal Maso). We close the section by recalling a generalization from [Gwi05, GZG07] of the so-called fundamental theorem on Young measures; cf. [Bal89].

2.1. Analytical framework. Let $C(\bar{\Omega})^d$ be the space of d-component vector-valued continuous functions on $\bar{\Omega}$ and let $C_0(\mathbb{R}^{d\times d}_{\mathrm{sym}})$ denote the space of continuous functions with compact support in $\mathbb{R}^{d\times d}_{\mathrm{sym}}$. For a measurable subset $\omega \subset \mathbb{R}^d$, we denote the classical spaces of Lebesgue and vector-valued Sobolev functions by $(L^s(\omega), \|\cdot\|_{s;\omega})$ and $(W^{1,s}(\omega)^d, \|\cdot\|_{1,s;\omega})$, $s \in [1,\infty]$, respectively. Let $\mathcal{D}(\omega) := C_0^\infty(\omega)^d$ be the set of infinitely many times differentiable d-component vector-valued functions with compact support in ω; we denote by $\mathcal{D}'(\omega)$ the corresponding dual space, consisting of distributions on ω. For $s \in [1,\infty)$, denote by $W_0^{1,s}(\omega)^d$ the closure of $\mathcal{D}(\omega)$ in $W^{1,s}(\omega)^d$ and let $W^{1,s}_{0,\mathrm{div}}(\omega)^d := \{\boldsymbol{v} \in W_0^{1,s}(\omega)^d \colon \operatorname{div} \boldsymbol{v} \equiv 0\}$. The case $s = \infty$ has to be treated differently. We define

$$W_0^{1,\infty}(\Omega)^d := W_0^{1,1}(\Omega)^d \cap W^{1,\infty}(\Omega)^d$$

and

$$W^{1,\infty}_{0,\mathrm{div}}(\Omega)^d := W^{1,1}_{0,\mathrm{div}}(\Omega)^d \cap W^{1,\infty}(\Omega)^d.$$

Moreover, we denote the space of functions in $L^s(\omega)$ with zero integral mean by $L^s_0(\omega)$. For $s, s' \in (1,\infty)$ with $\frac{1}{s} + \frac{1}{s'} = 1$, $L^{s'}(\Omega)$ and $L_0^{s'}(\Omega)$ are the dual spaces of $L^s(\Omega)$ and $L^s_0(\Omega)$, respectively. The dual of $W_0^{1,s}(\Omega)^d$ is denoted by $W^{-1,s'}(\Omega)^d$. For $\omega = \Omega$ we omit the domain in our notation for norms; e. g., we write $\|\cdot\|_s$ instead of $\|\cdot\|_{s,\Omega}$.

Inf-sup condition. The inf-sup condition has a crucial role in the analysis of incompressible flow problems. It states that, for $s, s' \in (1, \infty)$ with $\frac{1}{s} + \frac{1}{s'} = 1$, there exists a constant $\alpha_s > 0$ such that

$$\sup_{0 \neq \boldsymbol{v} \in W_0^{1,s}(\Omega)^d} \frac{\langle \operatorname{div} \boldsymbol{v},\, q \rangle_\Omega}{\|\boldsymbol{v}\|_{1,s}} \geqslant \alpha_s \|q\|_{s'} \qquad \text{for all } q \in L_0^{s'}(\Omega). \tag{2.1a}$$

This follows from the existence of the *Bogovskiĭ* operator $\mathfrak{B} : L_0^s(\Omega) \to W_0^{1,s}(\Omega)$, with

$$\operatorname{div} \mathfrak{B} h = h \qquad \text{and} \qquad \alpha_s \|\mathfrak{B} h\|_{1,s} \leqslant \|h\|_s \tag{2.1b}$$

for all $s \in (1, \infty)$; compare e.g. with [DRS10, Bog79].

Korn's inequality. According to (1.2) the maximal monotone graph defined in (A1)–(A5) provides control over the symmetric part of the velocity gradient only. Korn's inequality implies that this suffices in order to control the norm of a Sobolev function; i. e., for $s \in (1, \infty)$, there exists a $\gamma_s > 0$ such that

$$\gamma_s \|\boldsymbol{v}\|_{1,s} \leqslant \|\boldsymbol{D}\boldsymbol{v}\|_s \qquad \text{for all } \boldsymbol{v} \in W_0^{1,s}(\Omega)^d; \tag{2.2}$$

compare, for example, with [DRS10].

2.2. Approximation of maximal monotone r-graphs. In general an x-dependent maximal monotone r-graph $\mathcal{A}$ satisfying (A1)–(A5) cannot be represented in an explicit fashion. However, based on a regularized *measurable selection*, it can be approximated by a regular single-valued monotone tensor field. Following [FMT04], there exists a mapping $\boldsymbol{S}^* : \Omega \times \mathbb{R}^{d \times d}_{\text{sym}} \to \mathbb{R}^{d \times d}_{\text{sym}}$ (a selection) such that, for all $\boldsymbol{\delta} \in \mathbb{R}^{d \times d}_{\text{sym}}$, $(\boldsymbol{\delta}, \boldsymbol{S}^*(x, \boldsymbol{\delta})) \in \mathcal{A}(x)$ for almost every $x \in \Omega$ and

(a1) $\boldsymbol{S}^*$ is measurable with respect to the product σ-algebra $\mathfrak{L}(\Omega) \otimes \mathfrak{B}(\mathbb{R}^{d \times d}_{\text{sym}})$;

(a2) For almost all $x \in \Omega$ the domain of $\boldsymbol{S}^*$ is $\mathbb{R}^{d \times d}_{\text{sym}}$;

(a3) $\boldsymbol{S}^*$ is monotone, i. e., for every $\boldsymbol{\delta}_1, \boldsymbol{\delta}_2 \in \mathbb{R}^{d \times d}_{\text{sym}}$ and almost all $x \in \Omega$,

$$\big(\boldsymbol{S}^*(x, \boldsymbol{\delta}_1) - \boldsymbol{S}^*(x, \boldsymbol{\delta}_2)\big) : (\boldsymbol{\delta}_1 - \boldsymbol{\delta}_2) \geqslant 0; \tag{2.3}$$

(a4) For almost all $x \in \Omega$ and all $\boldsymbol{\delta} \in \mathbb{R}^{d \times d}_{\text{sym}}$ the following growth and coercivity conditions hold:

$$\big|\boldsymbol{S}^*(x, \boldsymbol{\delta})\big| \leqslant c_1 \, |\boldsymbol{\delta}|^{r-1} + k(x) \quad \text{and} \quad \boldsymbol{S}^*(x, \boldsymbol{\delta}) : \boldsymbol{\delta} \geqslant c_2 \, |\boldsymbol{\delta}|^r - m(x), \tag{2.4}$$

where $c_1, c_2 > 0$, and $k \in L^{r'}(\Omega)$ and $m \in L^1(\Omega)$ are nonnegative functions.

Let $\eta \in C_0(\mathbb{R}^{d \times d}_{\text{sym}})$ be a radially symmetric nonnegative function with support in the unit ball $B_1(\mathbf{0}) \subset \mathbb{R}^{d \times d}_{\text{sym}}$ and $\int_{\mathbb{R}^{d \times d}_{\text{sym}}} \eta \, \mathrm{d}\boldsymbol{\zeta} = 1$. For $n \in \mathbb{N}$ we then set $\eta^n(\boldsymbol{\zeta}) = n^{d^2} \eta(n \boldsymbol{\zeta})$ and define

$$\boldsymbol{S}^n(x, \boldsymbol{\delta}) := (\boldsymbol{S}^* * \eta^n)(x, \boldsymbol{\delta}) = \int_{\mathbb{R}^{d \times d}_{\text{sym}}} \boldsymbol{S}^*(x, \boldsymbol{\zeta}) \, \eta^n(\boldsymbol{\delta} - \boldsymbol{\zeta}) \, \mathrm{d}\boldsymbol{\zeta} = \int_{\mathbb{R}^{d \times d}_{\text{sym}}} \boldsymbol{S}^*(x, \boldsymbol{\zeta}) \, \mathrm{d}\mu^n_{\boldsymbol{\delta}}. \tag{2.5}$$

Here, thanks to the equality $\int_{\mathbb{R}^{d \times d}_{\text{sym}}} \eta^n \, \mathrm{d}\boldsymbol{\zeta} = 1$ and the nonnegativity of η, for each $\boldsymbol{\delta} \in \mathbb{R}^{d \times d}_{\text{sym}}$, $\mathrm{d}\mu^n_{\boldsymbol{\delta}} := \eta^n(\boldsymbol{\delta} - \boldsymbol{\zeta}) \, \mathrm{d}\boldsymbol{\zeta}$ defines a probability measure that is absolutely continuous with respect to the Lebesgue measure, with density $\eta^n(\boldsymbol{\delta} - (\cdot))$.

We recall the following properties of the matrix function $\boldsymbol{S}^n$ from [GMŚ07, BGMŚG09, GZG07].

Lemma 1. *The x-dependent matrix function $\boldsymbol{S}^n$, defined in* (2.5)*, satisfies*

$$(\boldsymbol{S}^n(x, \boldsymbol{\delta}_1) - \boldsymbol{S}^n(x, \boldsymbol{\delta}_2)) : (\boldsymbol{\delta}_1 - \boldsymbol{\delta}_2) \geqslant 0 \quad \textit{for all } \boldsymbol{\delta}_1, \boldsymbol{\delta}_2 \in \mathbb{R}^{d \times d}_{sym}.$$

Moreover, there exist constants $\tilde{c}_1, \tilde{c}_2 > 0$ *and nonnegative functions* $\tilde{m} \in L^1(\Omega)$, $\tilde{k} \in L^{r'}(\Omega)$ *such that, uniformly in* $n \in \mathbb{N}$, *we have*

$$|\boldsymbol{S}^n(x,\boldsymbol{\delta})| \leqslant \tilde{c}_1 \, |\boldsymbol{\delta}|^{r-1} + \tilde{k}(x) \qquad \text{for all } \boldsymbol{\delta} \in \mathbb{R}^{d\times d}_{sym},$$
$$\boldsymbol{S}^n(x,\boldsymbol{\delta}) : \boldsymbol{\delta} \geqslant \tilde{c}_2 \, |\boldsymbol{\delta}|^{r} - \tilde{m}(x) \qquad \text{for all } \boldsymbol{\delta} \in \mathbb{R}^{d\times d}_{sym}.$$

REMARK 2. *The selection* $\boldsymbol{S}^*$ *enters in the definition of the finite element method in the form of* $\boldsymbol{S}^n$ *through the Galerkin ansatz; compare with Section 3.4 below. The natural question is then how one can gain access to such a selection. In fact, in most physical models it appears that the selection* $\boldsymbol{S}^*$ *is given and* $\mathcal{A}(x)$ *is defined as the maximal monotone graph containing the set* $\{(\boldsymbol{D}, \boldsymbol{S}^*(\boldsymbol{D}))\colon \boldsymbol{D} \in \mathbb{R}^{d\times d}_{sym}\}$; *compare with [BGMŚG09, GMŚ07] and the references therein.*

2.3. Weak convergence tools. The result in [Gwi05, GZG07] extends[Bal89] from limits of single distributed measures to limits of general probability measures. To this end we need to introduce some standard notation from the theory of Young measures. We denote by $\mathfrak{M}(\mathbb{R}^{d\times d}_{\mathrm{sym}})$ the space of bounded Radon measures. We call $\mu : \Omega \to \mathfrak{M}(\mathbb{R}^{d\times d}_{\mathrm{sym}})$, $x \mapsto \mu_x$, weak-$*$ measurable if the mapping $x \mapsto \int_{\mathbb{R}^{d\times d}_{\mathrm{sym}}} h(\boldsymbol{\zeta}) \,\mathrm{d}\mu_x(\boldsymbol{\zeta})$ is measurable for all $h \in C_0(\mathbb{R}^{d\times d}_{\mathrm{sym}})$. The associated non-negative measure is defined by $|\mu_x|(\mathcal{C}) := \mu_x^+ + \mu_x^-$, via the Jordan decomposition $\mu_x = \mu_x^+ - \mu_x^-$ into two bounded non-negative measures μ_x^+, μ_x^-. By means of the norm $\|\mu\|_{L^\infty_w(\Omega;\mathfrak{M}(\mathbb{R}^{d\times d}_{\mathrm{sym}}))} := \operatorname{ess\,sup}_{x\in\Omega} \int_{\mathbb{R}^{d\times d}_{\mathrm{sym}}} \mathrm{d}\,|\mu_x|$ the space $L^\infty_w(\Omega;\mathfrak{M}(\mathbb{R}^{d\times d}_{\mathrm{sym}}))$ of essentially bounded, weak-$*$ measurable functions turns into a Banach space with separable predual $L^1(\Omega, C_0(\mathbb{R}^{d\times d}_{\mathrm{sym}}))$. The support of a non-negative measure is defined to be the largest closed subset of $\mathbb{R}^{d\times d}_{\mathrm{sym}}$ for which every open neighborhood of every point of the set has positive measure and $\operatorname{supp}\mu_x := \operatorname{supp}\mu_x^+ \cup \operatorname{supp}\mu_x^-$.

THEOREM 3 (Young measures). *Let* Ω *be an open and bounded subset of* $\mathbb{R}^d$. *Suppose that* $\{\nu^j\}_{j\in\mathbb{N}} \subset L^\infty_w(\Omega;\mathfrak{M}(\mathbb{R}^{d\times d}_{sym}))$ *is such that* ν^j_x *is a probability measure on* $\mathbb{R}^{d\times d}_{sym}$ *for all* $j \in \mathbb{N}$ *and a.e.* $x \in \Omega$. *Assume that* ν^j *converges to* ν *in the weak-$*$ topology of* $L^\infty_w(\Omega;\mathfrak{M}(\mathbb{R}^{d\times d}_{sym}))$ *for some* $\nu \in L^\infty_w(\Omega;\mathfrak{M}(\mathbb{R}^{d\times d}_{sym}))$.

Suppose further that the sequence $\{\nu^j\}_{j\in\mathbb{N}}$ *satisfies the tightness condition*

$$\lim_{R\to\infty} \sup_{j\in\mathbb{N}} \left|\left\{x \in \Omega\colon\ \operatorname{supp}\nu^j_x \setminus B_R(\boldsymbol{0}) \neq \varnothing\right\}\right| \to 0,$$

where $B_R(\boldsymbol{0})$ *denotes the ball in* $\mathbb{R}^{d\times d}_{sym}$ *with center* $\boldsymbol{0} \in \mathbb{R}^{d\times d}_{sym}$ *and radius* $R > 0$.

Then, the following statements hold:

(i) ν_x *is a probability measure, i. e.,* $\|\nu_x\|_{\mathfrak{M}(\mathbb{R}^{d\times d}_{sym})} = \int_{\mathbb{R}^{d\times d}_{sym}} \mathrm{d}\,|\nu_x| = 1$ *a.e. in* Ω;

(ii) for every $h \in L^\infty(\Omega; C_b(\mathbb{R}^{d\times d}_{sym}))$,

$$\int_{\mathbb{R}^{d\times d}_{sym}} h(x,\boldsymbol{\zeta}) \,\mathrm{d}\nu^j_x(\boldsymbol{\zeta}) \rightharpoonup^* \int_{\mathbb{R}^{d\times d}_{sym}} h(x,\boldsymbol{\zeta}) \,\mathrm{d}\nu_x(\boldsymbol{\zeta}) \qquad \text{weak-}* \text{ in } L^\infty(\Omega);$$

(iii) for every measurable subset $\omega \subset \Omega$ *and for every Carathéodory function* h *such that*

$$\lim_{R\to\infty} \sup_{j\in\mathbb{N}} \int_\omega \int_{\left\{\boldsymbol{\zeta}\in\mathbb{R}^{d\times d}_{sym}\colon\ |h(x,\boldsymbol{\zeta})|>R\right\}} |h(x,\boldsymbol{\zeta})| \,\mathrm{d}\nu^j_x(\boldsymbol{\zeta}) \,\mathrm{d}\boldsymbol{\zeta} = 0 \tag{2.6}$$

we have that

$$\int_{\mathbb{R}^{d\times d}_{sym}} h(x,\boldsymbol{\zeta}) \,\mathrm{d}\nu^j_x(\boldsymbol{\zeta}) \rightharpoonup \int_{\mathbb{R}^{d\times d}_{sym}} h(x,\boldsymbol{\zeta}) \,\mathrm{d}\nu_x(\boldsymbol{\zeta}) \qquad \text{weakly in } L^1(\omega).$$

LEMMA 4 (Chacon's biting lemma). *Let Ω be a bounded domain in $\mathbb{R}^d$ and let $\{v^n\}_{n\in\mathbb{N}}$ be a bounded sequence in $L^1(\Omega)$. Then, there exists a nonincreasing sequence of measurable subsets $E_j \subset \Omega$ with $|E_j| \to 0$ as $j \to \infty$, such that $\{v^n\}_{n\in\mathbb{N}}$ is precompact in the weak topology of $L^1(\Omega\backslash E_j)$, for each $j \in \mathbb{N}$.*

3. Finite Element Approximation. This section is concerned with approximating problem (1.2) by a finite element method. To this end we introduce a general framework covering inf-sup stable Stokes elements, which are discretely divergence-free, as well as exactly divergence-free finite elements for the velocity. These two classes of velocity elements require different treatment of the convection term. The discussion of these, including representative examples of velocity-pressure pairs of finite element spaces from each class, is the subject of §3.2 and §3.3. The finite element approximation of (1.2) is stated in §3.4. We close with a new Lipschitz truncation method for finite element spaces, which plays a crucial role in the proof of our main result, Theorem 19.

3.1. Finite element spaces. We consider a family $\{\mathbb{V}^n, \mathbb{Q}^n\}_{n\in\mathbb{N}} \subset W_0^{1,\infty}(\Omega)^d \times L^\infty(\Omega)$ of pairs of conforming finite-dimensional subspaces of $W_0^{1,r}(\Omega)^d \times L^\infty(\Omega)$. To be more precise, let $\mathbb{G} := \{\mathcal{G}_n\}_{n\in\mathbb{N}}$ be a sequence of shape-regular partitions of $\bar{\Omega}$, i. e., a sequence of regular finite element partitions of $\bar{\Omega}$ satisfying the following structural assumptions.

- *Affine equivalence:* For every element $E \in \mathcal{G}_n$, $n \in \mathbb{N}$, there exists an invertible affine mapping

$$\boldsymbol{F}_E : E \to \hat{E},$$

  where $\hat{E}$ is the closed standard reference d-simplex or the closed standard unit cube in $\mathbb{R}^d$.
- *Shape-regularity:* For any element $E \in \mathcal{G}_n$, $n \in \mathbb{N}$, the ratio of its diameter to the diameter of the largest inscribed ball is bounded, uniformly with respect to all partitions $\mathcal{G}_n$, $n \in \mathbb{N}$.

For a given partition $\mathcal{G}_n$, $n \in \mathbb{N}$, and certain subspaces $\mathbb{V} \subset C(\bar{\Omega})^d$ and $\mathbb{Q} \subset L^\infty(\Omega)$, the finite element spaces are then given by

$$\mathbb{V}^n = \mathbb{V}(\mathcal{G}_n) := \Big\{ \boldsymbol{V} \in \mathbb{V} : \ \boldsymbol{V}_{|E} \circ \boldsymbol{F}_E^{-1} \in \hat{\mathbb{P}}_{\mathbb{V}}, \ E \in \mathcal{G}_n \text{ and } \boldsymbol{V}_{|\partial\Omega} = 0 \Big\}, \tag{3.1a}$$

$$\mathbb{Q}^n = \mathbb{Q}(\mathcal{G}_n) := \Big\{ Q \in \mathbb{Q} : \ Q_{|E} \circ \boldsymbol{F}_E^{-1} \in \hat{\mathbb{P}}_{\mathbb{Q}}, \ E \in \mathcal{G}_n \Big\}, \tag{3.1b}$$

where $\hat{\mathbb{P}}_{\mathbb{V}} \subset W^{1,\infty}(\hat{E})^d$ and $\hat{\mathbb{P}}_{\mathbb{Q}} \subset L^\infty(\hat{E})$ are finite-dimensional subspaces, with $\dim \hat{\mathbb{P}}_{\mathbb{V}} = \ell$ and $\dim \hat{\mathbb{P}}_{\mathbb{V}} = \jmath$, respectively, for some $\ell, \jmath \in \mathbb{N}$. Note that $\mathbb{Q}^n \subset L^\infty(\Omega)$ and since $\mathbb{V}^n \subset C(\bar{\Omega})^d$ it follows that $\mathbb{V}^n \subset W_0^{1,\infty}(\Omega)^d$. Each of the above spaces is assumed to have a finite and locally supported basis; e. g. for the discrete pressure space this means that for $n \in \mathbb{N}$ there exists $N_n \in \mathbb{N}$ such that

$$\mathbb{Q}^n = \operatorname{span}\{Q_1^n, \dots, Q_{N_n}^n\}$$

and for each basis function Q_i^n, $i = 1, \dots, N_n$, we have that if there exists $E \in \mathcal{G}_n$ with $Q_i^n \not\equiv 0$ on E, then

$$\operatorname{supp} Q_j^n \subset \bigcup \big\{ E' \in \mathcal{G}_n \mid E' \cap E \neq \emptyset \big\} =: \Omega_E^n \quad \text{with} \quad \Omega_E^n \leqslant c\,|E|$$

for some constant $c > 0$ depending on the shape-regularity of $\mathbb{G}$. The piecewise constant mesh size function $h_{\mathcal{G}_n} \in L^\infty(\Omega)$ is almost everywhere in Ω defined by

$$h_{\mathcal{G}_n}(x) := |E|\,, \quad \text{if } E \in \mathcal{G}_n \text{ with } x \in E.$$

We introduce the subspace $\mathbb{V}^n_{\mathrm{div}}$ of discretely divergence-free functions by

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

and we define

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

Throughout the paper we assume that all pairs of velocity-pressure finite element spaces possess the following properties.

Assumption 5 (Approximability). *For all $s \in [1,\infty)$,*

$$\inf_{\boldsymbol{V}\in\mathbb{V}^n} \|\boldsymbol{v} - \boldsymbol{V}\|_{1,s} \to 0 \qquad \text{for all } \boldsymbol{v} \in W^{1,s}_0(\Omega)^d, \text{ as } n \to \infty; \text{ and}$$
$$\inf_{Q\in\mathbb{Q}^n} \|q - Q\|_{s} \to 0 \qquad \text{for all } q \in L^s(\Omega), \text{ as } n \to \infty.$$

For this, a necessary condition is that the maximal mesh size vanishes, i.e. we have $\|h_{\mathcal{G}_n}\|_{L^\infty(\Omega)} \to 0$ as $n \to \infty$.

Assumption 6 (Projector Π^n_{div}). *For each $n \in \mathbb{N}$ there exists a linear projection operator $\Pi^n_{\mathrm{div}} : W^{1,1}_0(\Omega)^d \to \mathbb{V}^n$ such that,*

- *Π^n_{div} preserves divergence in the dual of $\mathbb{Q}^n$; i. e., for any $\boldsymbol{v} \in W^{1,1}_0(\Omega)^d$, we have*

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

- *Π^n_{div} is locally $W^{1,1}$-stable; i. e., there exists $c_1 > 0$, independent of n, such that*

$$\fint_E |\Pi^n_{\mathrm{div}}\boldsymbol{v}| + h_{\mathcal{G}_n} |\nabla \Pi^n_{\mathrm{div}}\boldsymbol{v}| \,\mathrm{d}x \leqslant c_1 \fint_{\Omega^n_E} |\boldsymbol{v}| + h_{\mathcal{G}_n} |\nabla \boldsymbol{v}| \,\mathrm{d}x \tag{3.2}$$

for all $\boldsymbol{v} \in W^{1,1}_0(\Omega)^d$ and all $E \in \mathcal{G}_n$. Here we have used the notation $\fint_B \cdot \,\mathrm{d}x := \frac{1}{|B|}\int_B \cdot \,\mathrm{d}x$ for the integral mean-value over a measurable set $B \subset \mathbb{R}^d$, $|B| \neq 0$.

It was shown in [BBDR10, DR07] that the local $W^{1,1}$-stability of Π^n_{div} implies its local and global $W^{1,s}$-stability, $s \in [1,\infty]$. In fact, by noting that the power function $t \mapsto t^s$ is convex for $s \in [1,\infty)$, we obtain for almost every $x \in E$, $E \in \mathcal{G}_n$, by the equivalence of norms on finite-dimensional spaces and standard scaling arguments, that

$$\begin{aligned}
|\Pi^n_{\mathrm{div}}\boldsymbol{v}(x)| + h_{\mathcal{G}_n} |\nabla \Pi^n_{\mathrm{div}}\boldsymbol{v}(x)| &\leqslant \|\Pi^n_{\mathrm{div}}\boldsymbol{v}(x)\|_{L^\infty(E)} + h_{\mathcal{G}_n} \|\nabla \Pi^n_{\mathrm{div}}\boldsymbol{v}\|_{L^\infty(E)} \\
&\leqslant c \fint_E |\Pi^n_{\mathrm{div}}\boldsymbol{v}(x)| + h_{\mathcal{G}_n} |\nabla \Pi^n_{\mathrm{div}}\boldsymbol{v}| \,\mathrm{d}x \\
&\leqslant c \fint_{\Omega^n_E} |\boldsymbol{v}(x)| + h_{\mathcal{G}_n} |\nabla \boldsymbol{v}| \,\mathrm{d}x \\
&\leqslant c \left(\fint_{\Omega^n_E} |\boldsymbol{v}(x)|^s + h^s_{\mathcal{G}_n} |\nabla \boldsymbol{v}|^s \,\mathrm{d}x \right)^{\frac{1}{s}},
\end{aligned}$$

where we have used Jensen's inequality in the last step; recall that $|\Omega^n_E| \leqslant c\,|E|$ with a constant depending solely on the shape-regularity of $\mathbb{G}$. Raising this inequality to the s-th power and integrating over E yields

$$\int_E |\Pi^n_{\mathrm{div}}\boldsymbol{v}|^s + h^s_{\mathcal{G}_n} |\nabla \Pi^n_{\mathrm{div}}\boldsymbol{v}|^s \,\mathrm{d}x \leqslant c \int_{\Omega^n_E} |\boldsymbol{v}|^s + h^s_{\mathcal{G}_n} |\nabla \boldsymbol{v}|^s \,\mathrm{d}x.$$

Summing over all elements $E \in \mathcal{G}_n$ and accounting for the locally finite overlap of patches yields, for any $s \in [1,\infty)$, that

$$\|\Pi^n_{\mathrm{div}}\boldsymbol{v}\|_{1,s} \leqslant c_s \|\boldsymbol{v}\|_{1,s} \qquad \text{for all } \boldsymbol{v} \in W^{1,s}_0(\Omega)^d, \tag{3.3}$$

with a constant $c_s > 0$ independent of $n \in \mathbb{N}$. Note that for $s = \infty$ the inequality (3.3) follows from an obvious modification of the argument above.

Hence, by invoking the approximation properties of the sequence of finite element spaces for the velocity, stated in Assumption 5, we obtain that

$$\|\boldsymbol{v} - \Pi^n_{\mathrm{div}} \boldsymbol{v}\|_{1,s} \to 0 \qquad \text{for all } \boldsymbol{v} \in W^{1,s}_0(\Omega)^d, \text{ as } n \to \infty \text{ and } s \in [1, \infty). \tag{3.4}$$

Moreover, we have the following result in the weak topology of $W^{1,s}_0(\Omega)^d$.

Proposition 7. *Let $\{\boldsymbol{v}_n\}_{n\in\mathbb{N}} \subset W^{1,s}_0(\Omega)^d$, $s \in (1,\infty)$, such that $\boldsymbol{v}_n \rightharpoonup \boldsymbol{v}$ weakly in $W^{1,s}_0(\Omega)^d$ as $n \to \infty$. Then*

$$\Pi^n_{\mathrm{div}} \boldsymbol{v}_n \rightharpoonup \boldsymbol{v} \quad \textit{weakly in } W^{1,s}_0(\Omega)^d \textit{ as } n \to \infty.$$

Proof. Thanks to the uniform boundedness of the sequence of linear operators $\{\Pi^n_{\mathrm{div}} : W^{1,s}_0(\Omega)^d \to \mathbb{V}^n \subset W^{1,s}_0(\Omega)^d\}_{n\in\mathbb{N}}$ (cf. (3.3)), we have that there exists a weakly converging subsequence of $\{\Pi^n_{\mathrm{div}}\boldsymbol{v}_n\}_{n\in\mathbb{N}}$ in $W^{1,s}_0(\Omega)^d$. By the uniqueness of the weak limit, it therefore suffices to identify the limit of $\{\Pi^n_{\mathrm{div}}\boldsymbol{v}_n\}_{n\in\mathbb{N}}$ in $L^s(\Omega)^d$. We deduce from the above considerations that

$$\begin{aligned}\|\boldsymbol{v} - \Pi^n_{\mathrm{div}}\boldsymbol{v}_n\|_{L^s(\Omega)} &\leqslant \|\boldsymbol{v} - \Pi^n_{\mathrm{div}}\boldsymbol{v}\|_{L^s(\Omega)} + \|\Pi^n_{\mathrm{div}}(\boldsymbol{v}_n - \boldsymbol{v})\|_{L^s(\Omega)} \\ &\leqslant \|\boldsymbol{v} - \Pi^n_{\mathrm{div}}\boldsymbol{v}\|_{L^s(\Omega)} + c\|\boldsymbol{v}_n - \boldsymbol{v}\|_{L^s(\Omega)} + c\|h_{\mathcal{G}_n}\nabla(\boldsymbol{v}_n - \boldsymbol{v})\|_{L^s(\Omega)}.\end{aligned}$$

The first term vanishes because of (3.4) and the second term vanishes since $\boldsymbol{v}_n \to \boldsymbol{v}$ strongly in $L^s(\Omega)^d$, thanks to the compact embedding $W^{1,s}_0(\Omega)^d \hookrightarrow\hookrightarrow L^s(\Omega)^d$. The last term vanishes since $\|h_{\mathcal{G}_n}\|_{L^\infty(\Omega)} \to 0$ as $n \to \infty$, by Assumption 5. ∎

Assumption 8 (Projector $\Pi^n_{\mathbb{Q}}$). *For each $n \in \mathbb{N}$ there exists a linear projection operator $\Pi^n_{\mathbb{Q}} : L^1(\Omega) \to \mathbb{Q}^n$ such that, for all $s' \in (1,\infty)$, $\Pi^n_{\mathbb{Q}}$ is stable. In other words, there exists a constant $\tilde{c}_{s'} > 0$, independent of n, such that*

$$\|\Pi^n_{\mathbb{Q}} q\|_{s'} \leqslant \tilde{c}_{s'} \|q\|_{s'} \qquad \textit{for all } q \in L^{s'}(\Omega).$$

The stability of $\Pi^n_{\mathbb{Q}}$ and the approximation properties of $\mathbb{Q}^n \subset L^{s'}(\Omega)$, stated in Assumption 5, imply that $\Pi^n_{\mathbb{Q}}$ satisfies

$$\|q - \Pi^n_{\mathbb{Q}} q\|_{s'} \to 0, \qquad \text{as } n \to \infty \quad \text{for all } q \in L^{s'}(\Omega) \text{ and } s' \in (1,\infty). \tag{3.5}$$

As a consequence of (2.1a) and Assumption 6 (compare also with (3.3)) the following discrete counterpart of (2.1a) holds; see [BBDR10].

Proposition 9 (Inf-sup stability). *For all $s, s' \in (1,\infty)$ with $\frac{1}{s} + \frac{1}{s'} = 1$, there exists a constant $\beta_s > 0$, independent of n, such that*

$$\sup_{0 \neq \boldsymbol{V} \in \mathbb{V}^n} \frac{\langle \operatorname{div} \boldsymbol{V}, Q\rangle_\Omega}{\|\boldsymbol{V}\|_{1,s}} \geqslant \beta_s \|Q\|_{s'} \qquad \textit{for all } Q \in \mathbb{Q}^n_0 \textit{ and all } n \in \mathbb{N}.$$

Thanks to the above considerations, there is a *discrete Bogovskiĭ* operator, which admits the following properties.

Corollary 10 (Discrete Bogovskiĭ operator). *Under the conditions of this section, for all $n \in \mathbb{N}$, there exists $\mathfrak{B}^n : \operatorname{div} \mathbb{V}^n \to \mathbb{V}^n$ with*

$$\operatorname{div}(\mathfrak{B}^n H) = H \qquad \textit{and} \qquad \beta_s \|\mathfrak{B}^n H\|_{1,s} \leqslant \sup_{Q \in \mathbb{Q}^n} \frac{\langle H, Q\rangle_\Omega}{\|Q\|_{s'}}$$

for all $H \in \operatorname{div} \mathbb{V}^n$. Moreover, if $\boldsymbol{V}^n \in \mathbb{V}^n$, $n \in \mathbb{N}$, such that $\boldsymbol{V}^n \rightharpoonup \boldsymbol{V}$ weakly in $W_0^{1,s}(\Omega)^d$ as $n \to \infty$, then we have that

$$\mathfrak{B}^n \operatorname{div} \boldsymbol{V}^n \rightharpoonup \mathfrak{B} \operatorname{div} \boldsymbol{V} \quad \textit{weakly in } W_0^{1,s}(\Omega)^d \textit{ as } n \to \infty.$$

It follows from Corollary 10 by Hölder's inequality that $\beta_s \|\mathfrak{B}^n H\|_{1,s} \leqslant \|H\|_s$. However, we shall need in the proof of Lemma 21 the stronger statement from Corollary 10.

Proof. Thanks to the discrete inf-sup stability (Proposition 9), we may identify $\operatorname{div} \mathbb{V}^n$ with the dual of $\mathbb{Q}^n/\mathbb{R}$. Next, we extend $H \in \operatorname{div} \mathbb{V}^n$, to $h_H = L_0^s(\Omega)$, $s \in (1, \infty)$, by means of the projection operator $\Pi_{\mathbb{Q}}^n : L^{s'}(\Omega) \to \mathbb{Q}^n$, $\frac{1}{s} + \frac{1}{s'} = 1$, from Assumption 8. In fact, $h_H \in L_0^s(\Omega)$ is uniquely defined by

$$\int_\Omega H \, \Pi_{\mathbb{Q}}^n q \, \mathrm{d}x = \int_\Omega h_H q \, \mathrm{d}x \quad \text{for all } q \in L^{s'}(\Omega).$$

Moreover, we have

$$\|h_H\|_s = \sup_{q \in L^{s'}(\Omega)} \frac{\int_\Omega h_H q \, \mathrm{d}x}{\|q\|_{s'}} = \sup_{q \in L^{s'}(\Omega)} \frac{\int_\Omega H \, \Pi_{\mathbb{Q}}^n q \, \mathrm{d}x}{\|q\|_{s'}} \leqslant \tilde{c}_{s'} \sup_{Q \in \mathbb{Q}^n} \frac{\int_\Omega H Q \, \mathrm{d}x}{\|Q\|_{s'}}.$$

We define $\mathfrak{B}^n H := \Pi_{\operatorname{div}}^n \mathfrak{B} h_H \in \mathbb{V}^n$. Thanks to the above considerations and the stability properties (3.3) and (2.1b) of $\Pi_{\operatorname{div}}^n$ and $\mathfrak{B}$ respectively, we have proved the first claim.

In order to prove the second assertion, we set $H^n := \operatorname{div} \boldsymbol{V}^n$ and conclude that $H^n \rightharpoonup H := \operatorname{div} \boldsymbol{V}$ weakly in $L_0^s(\Omega)$ as $n \to \infty$. Consequently, thanks to (3.5), we have for all $q \in L^{s'}(\Omega)$, that

$$\int_\Omega h_{H^n} q \, \mathrm{d}x = \int_\Omega H^n \Pi_{\mathbb{Q}}^n q \, \mathrm{d}x \to \int_\Omega H q \, \mathrm{d}x \quad \text{as } n \to \infty.$$

In other words, we have that $h_{H^n} \rightharpoonup H$ weakly in $L_0^s(\Omega)$ as $n \to \infty$. The Bogovskiĭ operator $\mathfrak{B} : L_0^s(\Omega) \to W_0^{1,s}(\Omega)^d$ is continuous and therefore it is also continuous with respect to the weak topologies of the respective spaces; compare e.g. with [AB2006, Theorem 6.17]. Therefore, we have $\mathfrak{B} h_{H^n} \rightharpoonup \mathfrak{B} H$ weakly in $W_0^{1,s}(\Omega)^d$ as $n \to \infty$ and the assertion follows from Proposition 7. ∎

3.2. Discretely divergence-free finite elements. As in [Tem84] we wish to ensure that the discrete counterpart of the convection term inherits the skew-symmetry of the convection term. In particular, upon integration by parts, it follows that

$$-\int_\Omega (\boldsymbol{v} \otimes \boldsymbol{w}) : \nabla \boldsymbol{h} \, \mathrm{d}x = \int_\Omega (\boldsymbol{v} \otimes \boldsymbol{h}) : \nabla \boldsymbol{w} + (\operatorname{div} \boldsymbol{v})(\boldsymbol{w} \cdot \boldsymbol{h}) \, \mathrm{d}x \tag{3.6}$$

for all $\boldsymbol{v}, \boldsymbol{w}, \boldsymbol{h} \in \mathcal{D}(\Omega)^d$. The last term vanishes provided that $\operatorname{div} \boldsymbol{v} \equiv 0$, and then

$$\int_\Omega (\boldsymbol{v} \otimes \boldsymbol{v}) : \nabla \boldsymbol{v} \, \mathrm{d}x = 0.$$

It can be easily seen that this is not generally true for finite element functions $\boldsymbol{V} \in \mathbb{V}^n$, even if

$$\langle \operatorname{div} \boldsymbol{V}, Q \rangle_\Omega = 0 \qquad \text{for all } Q \in \mathbb{Q}^n,$$

i. e., if $\boldsymbol{V}$ is discretely divergence-free. However, we observe from (3.6) that

$$-\int_\Omega (\boldsymbol{v} \otimes \boldsymbol{w}) : \nabla \boldsymbol{h} \, \mathrm{d}x = \tfrac{1}{2} \int_\Omega (\boldsymbol{v} \otimes \boldsymbol{h}) : \nabla \boldsymbol{w} - (\boldsymbol{v} \otimes \boldsymbol{w}) : \nabla \boldsymbol{h} \, \mathrm{d}x =: \mathcal{B}[\boldsymbol{v}, \boldsymbol{w}, \boldsymbol{h}] \tag{3.7}$$

for all $\boldsymbol{v}, \boldsymbol{w}, \boldsymbol{h} \in W^{1,\infty}_{0,\mathrm{div}}(\Omega)^d$. We extend this definition to $W^{1,\infty}(\Omega)^d$ in the obvious way and deduce that

$$\mathcal{B}[\boldsymbol{v},\, \boldsymbol{v},\, \boldsymbol{v}] = 0 \qquad \text{for all } \boldsymbol{v} \in W^{1,\infty}(\Omega)^d. \tag{3.8}$$

We further investigate this modified convection term for fixed $r, r' \in (1, \infty)$ with $\frac{1}{r} + \frac{1}{r'} = 1$; recall the definition of $\tilde{r}$ from (1.1). We note that $\tilde{r} > 1$ is equivalent to the condition $r > \frac{2d}{d+2}$. In this case we can define its dual $\tilde{r}' \in (1, \infty)$ by $\frac{1}{\tilde{r}} + \frac{1}{\tilde{r}'} = 1$ and we note that the Sobolev embedding

$$W^{1,r}(\Omega)^d \hookrightarrow L^{2\tilde{r}}(\Omega)^d \tag{3.9}$$

holds. This is a crucial condition in the continuous problem, which guarantees

$$\int_\Omega (\boldsymbol{v} \otimes \boldsymbol{w}) : \nabla \boldsymbol{h} \,\mathrm{d}x \leqslant c\, \|\boldsymbol{v}\|_{1,r} \|\boldsymbol{w}\|_{1,r} \|\boldsymbol{h}\|_{1,\tilde{r}'} \tag{3.10}$$

for all $\boldsymbol{v}, \boldsymbol{w}, \boldsymbol{h} \in W^{1,\infty}(\Omega)^d$; see [BGMŚG09] and Section 3.3 below. Because of the extension (3.7) of the convection term to functions that are not necessarily pointwise divergence-free, we have to adopt the following stronger condition in order to ensure that the trilinear form $\mathcal{B}[\cdot,\cdot,\cdot]$ is bounded on $W^{1,r}(\Omega)^d \times W^{1,r}(\Omega)^d \times W^{1,\tilde{r}'}(\Omega)^d$. In particular, let $r > \frac{2d}{d+1}$, in order to ensure that there exists $s \in (1, \infty)$ such that $\frac{1}{r} + \frac{1}{2\tilde{r}} + \frac{1}{s} = 1$. In other words, we have for $\boldsymbol{v}, \boldsymbol{w}, \boldsymbol{h} \in W^{1,\infty}(\Omega)^d$ that

$$\int_\Omega (\operatorname{div} \boldsymbol{v})\, (\boldsymbol{w} \cdot \boldsymbol{h}) \,\mathrm{d}x \leqslant \|\operatorname{div} \boldsymbol{v}\|_r \|\boldsymbol{w}\|_{2\tilde{r}} \|\boldsymbol{h}\|_s \leqslant c\, \|\boldsymbol{v}\|_{1,r} \|\boldsymbol{w}\|_{1,r} \|\boldsymbol{h}\|_{1,\tilde{r}'},$$

with a constant c depending on r, Ω and d. Here we have used the embeddings (3.9) and $W^{1,\tilde{r}'}_0(\Omega)^d \hookrightarrow L^s(\Omega)^d$. Consequently, together with (3.10) we thus obtain

$$\mathcal{B}[\boldsymbol{v},\, \boldsymbol{w},\, \boldsymbol{h}] \leqslant c\, \|\boldsymbol{v}\|_{1,r} \|\boldsymbol{w}\|_{1,r} \|\boldsymbol{h}\|_{1,\tilde{r}'}. \tag{3.11}$$

EXAMPLE 11. *In [BBDR10] it is shown that Assumptions 6 and 8 are satisfied by the following velocity-pressure pairs of finite elements:*

- *The conforming Crouzeix–Raviart Stokes element, i. e., continuous piecewise quadratic plus bubble velocity and discontinuous piecewise linear pressure approximations (compare e. g. with [BF91, §VI Example 3.6]);*
- *The Mini element; see, [BF91, §VI Example 3.7];*
- *The spaces of continuous piecewise quadratic elements for the velocity and piecewise constants for the pressure ([BF91, §VI Example 3.6]);*

Moreover, it is stated without proof in [BBDR10] that the lowest order Taylor–Hood element also satisfies Assumptions 6 and 8.

3.3. Exactly divergence-free finite elements. Another way of retaining the skew-symmetry of the convection term and ensuring that (3.8) holds is to use an exactly divergence-free finite element approximation of the velocity. In addition to Assumptions 6 and 8 in Section 3.1 we suppose that the following condition holds.

ASSUMPTION 12. *The finite element spaces defined in Section 3.1 satisfy*

$$\operatorname{div} \mathbb{V}^n \subset \mathbb{Q}^n_0, \qquad n \in \mathbb{N}.$$

This inclusion obviously implies that discretely divergence-free functions are automatically exactly divergence-free, i. e.,

$$\mathbb{V}^n_{\mathrm{div}} = \{\boldsymbol{V} \in \mathbb{V}^n :\ \operatorname{div} \boldsymbol{V} \equiv 0\}, \qquad n \in \mathbb{N}.$$

According to (3.6), in this case, we define

$$\mathcal{B}[\boldsymbol{v},\, \boldsymbol{w},\, \boldsymbol{h}] := -\int_\Omega (\boldsymbol{v} \otimes \boldsymbol{w}) : \nabla \boldsymbol{h} \,\mathrm{d}x \tag{3.12}$$

for all $\boldsymbol{v}, \boldsymbol{w}, \boldsymbol{h} \in W^{1,\infty}_0(\Omega)^d$ and obtain

$$\mathcal{B}[\boldsymbol{v},\, \boldsymbol{v},\, \boldsymbol{v}] = 0 \qquad \text{for all } \boldsymbol{v} \in W^{1,\infty}_{0,\mathrm{div}}(\Omega)^d. \tag{3.13}$$

Recalling (3.10), with Assumption 12, the convection term can be controlled under the weaker restriction $r > \frac{2d}{d+2}$, i. e., for $\boldsymbol{v}, \boldsymbol{w}, \boldsymbol{h} \in W^{1,\infty}(\Omega)^d$, we have that

$$\mathcal{B}[\boldsymbol{v},\, \boldsymbol{w},\, \boldsymbol{h}] \leqslant c\, \|\boldsymbol{v}\|_{1,r} \|\boldsymbol{w}\|_{1,r} \|\boldsymbol{h}\|_{1,\tilde{r}'}, \tag{3.14}$$

where, as before, $\frac{1}{\tilde{r}} + \frac{1}{\tilde{r}'} = 1$ with $\tilde{r}$ from (1.1). The constant $c > 0$ only depends on r, Ω and d.

Admittedly, finite element spaces that simultaneously satisfy Assumptions 6, 8 and 12 are not very common. Most constructions of exactly divergence-free finite element spaces in the literature are not very practical in that they require a sufficiently high polynomial degree and/or restrictions on the geometry of the mesh; see [AQ92, SV85, QZ07, Zha08]. In a very recent work [GN11] Guzmán and Neilan proposed inf-sup stable finite element pairs in two space-dimensions, which admit exactly divergence-free velocity approximations for $r = 2$. A generalization of the Guzmán–Neilan elements to three dimensions is contained in [GN12]. We shall show below that the lowest order spaces introduced in [GN11] simultaneously satisfy Assumptions 6, 8 and 12 for $r \in [1, \infty)$.

EXAMPLE 13 (Guzmán–Neilan elements [GN11]). *We consider the finite element spaces introduced by Guzmán and Neilan in [GN11, Section 3] on simplicial triangulations of a bounded open polygonal domain Ω in $\mathbb{R}^2$. In particular, we define*

$$\hat{\mathbb{P}}_{\mathbb{V}} := \mathbb{P}_1(\hat{E}) \;\oplus\; \mathrm{span}\{\mathrm{curl}(\hat{b}_i) : i = 1, 2, 3\} \;\oplus\; \mathrm{span}\{\mathrm{curl}(\hat{B}_i) : i = 1, 2, 3\}.$$

Here $\mathbb{P}_1(\hat{E})$ denotes the space of affine vector-valued functions over $\hat{E}$. Let further $\{\hat{\lambda}_i\}_{i=1,2,3}$ be the barycentric coordinates on $\hat{E}$ associated with the three vertices $\{\hat{z}_i\}_{i=1,2,3}$ of $\hat{E}$, i. e., $\hat{\lambda}_i(\hat{z}_j) = \delta_{ij}$. Then, for $i \in \{1, 2, 3\}$, we set $\hat{b}_i := \hat{\lambda}^2_{i+1}\hat{\lambda}_{i+2}$, and $\hat{B}_i$ denotes the rational bubble function

$$\hat{B}_i := \frac{\hat{\lambda}_i \hat{\lambda}^2_{i+1} \hat{\lambda}^2_{i+2}}{(\hat{\lambda}_i + \hat{\lambda}_{i+1})(\hat{\lambda}_i + \hat{\lambda}_{i+2})},$$

which can be continuously extended by zero at $\hat{z}_{i+1}$ and $\hat{z}_{i+2}$; the index i has to be understood modulo 3. Thanks to properties of the curl *operator, the local pressure space*

$$\hat{\mathbb{P}}_{\mathbb{Q}} := \mathrm{div}\, \hat{\mathbb{P}}_{\mathbb{V}}$$

is the space of constant functions over $\hat{E}$.

It is clear from [GN11] that the related pairs of spaces $\{\mathbb{V}^n, \mathbb{Q}^n\}_{n\in\mathbb{N}}$ (compare with (3.1)*) satisfy Assumption 12. For $\Pi^n_{\mathbb{Q}}$ we can use a Clément type interpolation or simply the best-approximation in $\mathbb{Q}^n$; clearly, both satisfy Assumption 8. The approximability assumption, Assumption 5, follows with the mesh-size tending to zero. It remains to verify Assumption 6. To this end we analyze the interpolation operator proposed in [GN11]. In particular, let $\Pi^n_S : W^{1,1}_0(\Omega)^2 \to \mathbb{L}^n$ be the Scott–Zhang interpolant [SZ90] into the linear Lagrange finite element space $\mathbb{L}^n := \mathbb{L}(\mathcal{G}_n)$*

over a triangulation $\mathcal{G}_n$*, belonging to a shape-regular family of triangulations* $\mathbb{G} = \{\mathcal{G}_n\}_{n\in\mathbb{N}}$ *of* Ω*. Then* $\Pi^n_{\mathrm{div}} : W^{1,1}_0(\Omega)^2 \to \mathbb{V}^n$ *is defined by*

$$\begin{aligned}(\Pi^n_{\mathrm{div}}\boldsymbol{v})(z) &:= (\Pi^n_S\boldsymbol{v})(z), && z \in \mathcal{N}^n,\\ \int_S \Pi^n_{\mathrm{div}}\boldsymbol{v}\,\mathrm{d}s &= \int_S \boldsymbol{v}\,\mathrm{d}s, && S \in \mathcal{S}^n,\end{aligned} \tag{3.15}$$

where $\mathcal{N}^n$ *and* $\mathcal{S}^n$ *denote the vertices, respectively edges, of the triangulation* $\mathcal{G}_n$*,* $n \in \mathbb{N}$*. This operator is a projector and thanks to [GN11, (3.14)] and the fact that* $\mathbb{L}^n \subset \mathbb{V}^n$ *it thus remains to prove the stability estimate* (3.2) *in Assumption 6.*

To this end we fix $n \in \mathbb{N}$*. Although the claim can be proved using the techniques in [GN11], this would necessitate the introduction of additional notation. Thus, for the sake of brevity of the presentation, we give an alternative proof. According to [GN11] the interpolation operator* Π^n_{div} *is correctly defined by* (3.15)*. Let* $\{\hat{z}_i\}_{i=1,2,3}$ *be the vertices of* $\hat{E}$ *and let* $\{\hat{S}_i\}_{i=1,2,3}$ *be its edges. Then, any function* $\hat{\boldsymbol{V}} \in \hat{\mathbb{P}}_{\mathbb{V}}$ *is uniquely defined by* $\hat{\boldsymbol{V}}(\hat{z}_i)$ *and* $\int_{\hat{S}_i}\hat{\boldsymbol{V}}\,\mathrm{d}s$*,* $i = 1, 2, 3$*. This implies that the mapping*

$$\hat{\boldsymbol{V}} \mapsto \sum_{i=1}^{3} \big|\hat{\boldsymbol{V}}(\hat{z}_i)\big| + \left|\int_{\hat{S}_i}\hat{\boldsymbol{V}}\,\mathrm{d}s\right|, \qquad \hat{\boldsymbol{V}} \in \hat{\mathbb{P}}_{\mathbb{V}},$$

is a norm on $\hat{\mathbb{P}}_{\mathbb{V}}$*. Hence, equivalence of norms on finite-dimensional spaces together with* (3.15) *yield that*

$$\begin{aligned}\int_{\hat{E}} \big|\Pi^n_{\mathrm{div}}\boldsymbol{v}\circ\boldsymbol{F}_E^{-1}\big|\,\mathrm{d}x &\leqslant c\sum_{i=1}^{3}\Big(\big|\Pi^n_{\mathrm{div}}\boldsymbol{v}\circ\boldsymbol{F}_E^{-1}(\hat{z}_i)\big| + \Big|\int_{\hat{S}_i}\Pi^n_{\mathrm{div}}\boldsymbol{v}\circ\boldsymbol{F}_E^{-1}\,\mathrm{d}s\Big|\Big)\\ &= c\sum_{i=1}^{3}\Big(\big|\Pi^n_{S}\boldsymbol{v}\circ\boldsymbol{F}_E^{-1}(\hat{z}_i)\big| + \Big|\int_{\hat{S}_i}\boldsymbol{v}\circ\boldsymbol{F}_E^{-1}\,\mathrm{d}s\Big|\Big), \quad \boldsymbol{v} \in W^{1,1}_0(\Omega)^2.\end{aligned}$$

By a scaled trace theorem and properties of the Scott–Zhang operator we arrive at

$$\int_E |\Pi^n_{\mathrm{div}}\boldsymbol{v}|\,\mathrm{d}x \leqslant c\int_{\Omega^n_E} |\boldsymbol{v}| + h_{\mathcal{G}_n}|\nabla\boldsymbol{v}|\,\mathrm{d}x, \qquad \boldsymbol{v} \in W^{1,1}_0(\Omega)^2.$$

Note that $\Pi^n_{\mathrm{div}} : W^{1,1}_0(\Omega)^2 \to \mathbb{V}^n$ *is a projector and that* $\mathbb{L}^n \subset \mathbb{V}^n$*. Thus the inequality* (3.2) *follows from a standard inverse estimate and the Bramble–Hilbert Lemma; compare also with [BBDR10, Theorem 3.5].*

3.4. The Galerkin approximation. We are now ready to state the discrete problem. Let $\{\mathbb{V}^n, \mathbb{Q}^n\}_{n\in\mathbb{N}}$ be the finite element spaces of Section 3.2 or 3.3 and let $\mathcal{B} : W^{1,\infty}_0(\Omega)^d \times W^{1,\infty}_0(\Omega)^d \times W^{1,\infty}_0(\Omega)^d \to \mathbb{R}$ be defined correspondingly.

For $n \in \mathbb{N}$ we call a triple of functions $\big(\boldsymbol{U}^n, P^n, \boldsymbol{S}^n(\boldsymbol{D}\boldsymbol{U}^n)\big) \in \mathbb{V}^n \times \mathbb{Q}^n_0 \times L^{r'}(\Omega)^{d\times d}$ a Galerkin approximation of (1.2) if it satisfies

$$\begin{aligned}\int_\Omega \boldsymbol{S}^n(\cdot, \boldsymbol{D}\boldsymbol{U}^n) : \boldsymbol{D}\boldsymbol{V}\,\mathrm{d}x + \mathcal{B}[\boldsymbol{U}^n, \boldsymbol{U}^n, \boldsymbol{V}] - \langle \operatorname{div}\boldsymbol{V}, P^n\rangle_\Omega &= \langle \boldsymbol{f}, \boldsymbol{V}\rangle\\ &\quad \text{for all } \boldsymbol{V} \in \mathbb{V}^n,\\ \int_\Omega Q\operatorname{div}\boldsymbol{U}^n\,\mathrm{d}x = 0 &\quad \text{for all } Q \in \mathbb{Q}^n.\end{aligned} \tag{3.16}$$

Restricting the test-functions to $\mathbb{V}^n_{\mathrm{div}}$ the discrete problem (3.16) reduces to finding $\boldsymbol{U}^n \in \mathbb{V}^n_{\mathrm{div}}$ such that

$$\int_\Omega \boldsymbol{S}^n(\cdot, \boldsymbol{D}\boldsymbol{U}^n) : \boldsymbol{D}\boldsymbol{V}\,\mathrm{d}x + \mathcal{B}[\boldsymbol{U}^n, \boldsymbol{U}^n, \boldsymbol{V}] = \langle \boldsymbol{f}, \boldsymbol{V}\rangle \qquad \text{for all } \boldsymbol{V} \in \mathbb{V}^n_{\mathrm{div}}. \tag{3.17}$$

Thanks to (3.8), respectively (3.13), it follows from Lemma 1 and Korn's inequality (2.2) that the nonlinear operator defined by the left-hand side of (3.17) is coercive on $\mathbb{V}^n_{\mathrm{div}}$. Since the dimension of $\mathbb{V}^n_{\mathrm{div}}$ is finite, Brouwer's fixed point theorem ensures the existence of a solution to (3.17). The existence of a solution triple to (3.16) then follows by the discrete inf-sup stability implied by Proposition 9.

Of course, because of the weak assumptions in the definition of the maximal monotone r-graph, (3.16) does not define the Galerkin approximation $\boldsymbol{U}^n$ uniquely. However for each $n \in \mathbb{N}$ we may select an arbitrary one among possibly infinitely many solution triples and thus obtain a sequence

$$\big\{\big(\boldsymbol{U}^n, P^n, \boldsymbol{S}^n(\cdot, \boldsymbol{D}\boldsymbol{U}^n)\big)\big\}_{n\in\mathbb{N}}. \tag{3.18}$$

3.5. Discrete Lipschitz truncation. In this section we shall present a modified Lipschitz truncation, which acts on finite element spaces. This *discrete Lipschitz truncation* is basically a composition of a continuous Lipschitz truncation and the projector from Assumption 6. For this reason we first introduce a new Lipschitz truncation on $W^{1,1}_0(\Omega)^d$, based on the results in [DMS08, BDF12, BDS12], which provides finer estimates than the original Lipschitz truncation technique proposed by Acerbi and Fusco in [AF88].

For $v \in L^1(\mathbb{R}^d)$ we define the Hardy–Littlewood maximal function

$$M(v)(x) := \sup_{R>0} \fint_{B_R(x)} |v| \,\mathrm{d}y. \tag{3.19}$$

For $s \in (1,\infty]$ the Hardy–Littlewood maximal operator M is continuous from $L^s(\mathbb{R}^d)$ to $L^s(\mathbb{R}^d)$, i. e., there exists a constant $c_s > 0$ such that

$$\|M(v)\|_{L^s(\mathbb{R}^d)} \leqslant c_s \|v\|_{L^s(\mathbb{R}^d)} \qquad \text{for all } v \in L^s(\mathbb{R}^d), \tag{3.20}$$

and it is of weak type $(1,1)$, i. e., there exists a constant $c_1 > 0$ such that

$$\sup_{\lambda>0} \lambda \,|\{M(v) > \lambda\}| \leqslant c_1 \|v\|_{L^1(\mathbb{R}^d)} \qquad \text{for all } v \in L^1(\mathbb{R}^d); \tag{3.21}$$

see, e. g., [G04]. For any $\boldsymbol{v} \in W^{1,1}(\mathbb{R}^d)^d$ we set $M(\boldsymbol{v}) := M(|\boldsymbol{v}|)$ and $M(\nabla \boldsymbol{v}) := M(|\nabla \boldsymbol{v}|)$.

Let $\boldsymbol{v} \in W^{1,1}_0(\Omega)^d$; we may then assume that $\boldsymbol{v} \in W^{1,1}(\mathbb{R}^d)^d$ by extending $\boldsymbol{v}$ by zero outside Ω. For fixed $\lambda > 0$ we define

$$\mathcal{U}_\lambda(\boldsymbol{v}) := \{M(\nabla \boldsymbol{v}) > \lambda\}, \tag{3.22a}$$

and

$$\mathcal{H}_\lambda(\boldsymbol{v}) := \mathbb{R}^d \backslash (\mathcal{U}_\lambda(\boldsymbol{v}) \cap \Omega) = \{M(\nabla \boldsymbol{v}) \leqslant \lambda\} \cup (\mathbb{R}^d \backslash \Omega). \tag{3.22b}$$

Since $M(\nabla \boldsymbol{v})$ is lower semi-continuous, the set $\mathcal{U}_\lambda(\boldsymbol{v})$ is open and the set $\mathcal{H}_\lambda(\boldsymbol{v})$ is closed. According to [DMS08] it follows that $\boldsymbol{v}$ restricted to $\mathcal{H}_\lambda(\boldsymbol{v})$ is Lipschitz continuous and therefore also bounded. More precisely, we have that

$$|\boldsymbol{v}(x) - \boldsymbol{v}(y)| \leqslant c\,\lambda\,|x-y| \tag{3.23}$$

for all $x, y \in \mathcal{H}_\lambda(\boldsymbol{v})$, where the constant $c > 0$ depends on Ω.

It remains to extend $\boldsymbol{v}_{|\mathcal{H}_\lambda(\boldsymbol{v})}$ to a Lipschitz continuous function on $\mathbb{R}^d$. The result in [DMS08] is based on the so-called Kirszbraun extension theorem (cf. Theorem 2.10.43 in [Fed69]) and uses an additional truncation of $\boldsymbol{v}$ with respect to $M(\boldsymbol{v})$. This can be avoided by proceeding similarly as in [BDF12, BDS12], i. e., extending $\boldsymbol{v}_{|\mathcal{H}_\lambda(\boldsymbol{v})}$ by means of a partition of unity on a Whitney covering of the open and bounded set $\mathcal{U}_\lambda(\boldsymbol{v})$. To this end, we assume w.l.o.g. that $\mathcal{U}_\lambda(\boldsymbol{v}) \neq \emptyset$; otherwise $\boldsymbol{v}$ does not need to be extended since $\mathcal{H}_\lambda(\boldsymbol{v}) = \mathbb{R}^d$. According to [G04, BDF12] there exists a decomposition of the open set $\mathcal{U}_\lambda(\boldsymbol{v})$ into a family of (dyadic) closed cubes $\{\mathcal{Q}_j\}_{j\in\mathbb{N}}$, with side lengths $\ell_j := \ell(\mathcal{Q}_j)$, $j \in \mathbb{N}$, such that

(W1) $\bigcup_{j\in\mathbb{N}} \mathcal{Q}_j = \mathcal{U}_\lambda(\boldsymbol{v})$ and the $\mathcal{Q}_j$'s have pair-wise disjoint interiors.

(W2) $8\sqrt{d}\,\ell(\mathcal{Q}_j) \leqslant \operatorname{dist}(\mathcal{Q}_j, \partial\mathcal{U}_\lambda(\boldsymbol{v})) \leqslant 32\sqrt{d}\,\ell(\mathcal{Q}_j)$.

(W3) If $\mathcal{Q}_j \cap \mathcal{Q}_k \neq \emptyset$ for some $j, k \in \mathbb{N}$, then

$$\tfrac{1}{2} \leqslant \tfrac{\ell_j}{\ell_k} \leqslant 2.$$

(W4) For a given $\mathcal{Q}_j$ there exist at most $(3^d - 1)2^d$ cubes $\mathcal{Q}_k$ with $\mathcal{Q}_j \cap \mathcal{Q}_k \neq \emptyset$.

For a fixed cube $\mathcal{Q} \in \mathbb{R}^d$ with barycenter z and any $c > 0$, we define

$$c\mathcal{Q} := \left\{ x \in \mathbb{R}^d \colon \max_{i=1,\dots,d} |x_i - z_i| \leqslant c\,\ell(\mathcal{Q}) \right\}.$$

Hence, it follows from (W2) with $\theta_d := 2 + 64\sqrt{d}$, that

$$(\theta_d \mathcal{Q}_j) \cap \mathcal{H}_\lambda(\boldsymbol{v}) \neq \emptyset \quad \text{for all } j \in \mathbb{N}. \tag{3.24}$$

Let $\mathcal{Q}_j^* := \sqrt{\frac{9}{8}}\mathcal{Q}_j$ and $\mathcal{Q}_j^{**} := \frac{9}{8}\mathcal{Q}_j$. Thanks to (W4), the enlarged cubes $\mathcal{Q}_j^{**}$, $j \in \mathbb{N}$, are locally finite, i.e., they satisfy $\sum_j \chi_{\mathcal{Q}_j^{**}} \leqslant c$ with a constant $c > 0$, which depends on d only. Thanks to the overlaps of the $\mathcal{Q}_j^*$'s, there exists a partition of unity $\{\psi_j\}_{j\in\mathbb{N}}$ subordinated to the family $\{\mathcal{Q}_j^*\}_{j\in\mathbb{N}}$ with the following properties:

- $\sum_j \psi_j = \chi_{\mathcal{U}_\lambda(\boldsymbol{v})}$ and $0 \leqslant \psi_j \leqslant 1$ for all $j \in \mathbb{N}$;
- $\chi_{\frac{7}{8}\mathcal{Q}_j} \leqslant \psi_j \leqslant \chi_{\mathcal{Q}_j^*}$, for all $j \in \mathbb{N}$;
- $\psi_j \in C_0^\infty(\mathcal{Q}_j^*)$ with $|\nabla\psi_j| \leqslant c\,\ell_j^{-1}$, for all $j \in \mathbb{N}$.

The Lipschitz truncation of $\boldsymbol{v}$ is then denoted by $\boldsymbol{v}_\lambda$ and is defined by

$$\boldsymbol{v}_\lambda := \begin{cases} \sum_{j\in\mathbb{N}} \psi_j \boldsymbol{v}_j & \text{in } \mathcal{U}_\lambda(\boldsymbol{v}), \\ \boldsymbol{v} & \text{elsewhere}, \end{cases} \tag{3.25a}$$

where

$$\boldsymbol{v}_j := \begin{cases} \fint_{\mathcal{Q}_j^{**}} \boldsymbol{v}\,\mathrm{d}x & \text{if } \mathcal{Q}_j^* \subset \Omega, \\ 0 & \text{elsewhere}. \end{cases} \tag{3.25b}$$

We emphasize that the definition of the functions $\boldsymbol{v}_j$, $j \in \mathbb{N}$, here differs from the one in [BDF12], since we need to preserve the no-slip boundary condition for the velocity field on $\partial\Omega$ under Lipschitz truncation. Combining the techniques of [DMS08] and [BDF12] we obtain the following result; for ease of readability of the main body of the paper, the proof of Theorem 14 is deferred to the Appendix.

THEOREM 14. *Let $\lambda > 0$ and $\boldsymbol{v} \in W_0^{1,1}(\Omega)^d$. Then, the Lipschitz truncation defined in* (3.25) *has the following properties: $\boldsymbol{v}_\lambda \in W_0^{1,\infty}(\Omega)^d$, and*

(a) *$\boldsymbol{v}_\lambda = \boldsymbol{v}$ on $\mathcal{H}_\lambda(\boldsymbol{v})$, i. e., $\{\boldsymbol{v} \neq \boldsymbol{v}_\lambda\} \subset \mathcal{U}_\lambda(\boldsymbol{v}) \cap \Omega = \{M(\nabla\boldsymbol{v}) > \lambda\} \cap \Omega$;*

(b) *$\|\boldsymbol{v}_\lambda\|_s \leqslant c\,\|\boldsymbol{v}\|_s$ for all $s \in [1, \infty]$, with $\boldsymbol{v} \in L^s(\Omega)^d$;*

(c) *$\|\nabla\boldsymbol{v}_\lambda\|_s \leqslant c\,\|\nabla\boldsymbol{v}\|_s$ for all $s \in [1, \infty]$, with $\boldsymbol{v} \in W_0^{1,s}(\Omega)^d$;*

(d) *$|\nabla\boldsymbol{v}_\lambda| \leqslant c\,\lambda\chi_{\mathcal{U}_\lambda(\boldsymbol{v})\cap\Omega} + |\nabla\boldsymbol{v}|\,\chi_{\mathcal{H}_\lambda(\boldsymbol{v})} \leqslant c\lambda$ almost everywhere in $\mathbb{R}^d$.*

The constants appearing in the inequalities stated in parts (b), (c) *and* (d) *depend on Ω and d. In* (b) *and* (c) *they additionally depend on s.*

We next modify the Lipschitz truncation so that for finite element functions the truncation is again a finite element function. To this end we recall the definition of the finite element space $\mathbb{V}^n = \mathbb{V}(\mathcal{G}_n)$ of Section 3.1 or 3.3.

Let $\lambda > 0$ and fix $n \in \mathbb{N}$. Since $\mathbb{V}^n \subset W_0^{1,1}(\Omega)^d$, we could apply the truncation defined in (3.25). However, since in general the Lipschitz truncation $\boldsymbol{V}_\lambda$ of $\boldsymbol{V} \in \mathbb{V}^n$ does not belong to $\mathbb{V}^n$, we shall define the discrete Lipschitz truncation by

$$\boldsymbol{V}_{n,\lambda} := \Pi^n_{\mathrm{div}} \boldsymbol{V}_\lambda \in \mathbb{V}^n. \tag{3.26}$$

According to the next lemma the interpolation operator Π^n_{div} is local, in the sense that it modifies $\boldsymbol{V}_\lambda$ in a neighborhood of $\mathcal{U}_\lambda(\boldsymbol{V})$ only.

LEMMA 15. *Let $\boldsymbol{V} \in \mathbb{V}^n$. With the notation adopted in this section, we have that*

$$\{\boldsymbol{V}_{n,\lambda} \neq \boldsymbol{V}\} \subset \Omega^n_\lambda(\boldsymbol{V}) := \text{interior}\left(\bigcup \{\Omega^n_E \mid E \in \mathcal{G}_n \text{ with } E \cap \mathcal{U}_\lambda(\boldsymbol{V}) \neq \emptyset\}\right).$$

Proof. The stated inclusion follows from (3.2) in Assumption 6. In particular, let $E \in \mathcal{G}_n$ be such that $E \subset \mathbb{R}^d \backslash \Omega^n_\lambda(\boldsymbol{V})$; then, $\Omega^n_E \subset \mathcal{H}_\lambda(\boldsymbol{V})$. Consequently, by Theorem 14(a), we have $\boldsymbol{V}_\lambda = \boldsymbol{V}$ on Ω^n_E. Hence we deduce from (3.2), our assumption that $\boldsymbol{V} \in \mathbb{V}^n$, and the fact that Π^n_{div} is a projector, that

$$\begin{aligned}\fint_E |\boldsymbol{V} - \Pi^n_{\mathrm{div}} \boldsymbol{V}_\lambda| \,\mathrm{d}x &= \fint_E |\Pi^n_{\mathrm{div}}(\boldsymbol{V} - \boldsymbol{V}_\lambda)| \,\mathrm{d}x \\ &\leqslant c \fint_{\Omega^n_E} |\boldsymbol{V} - \boldsymbol{V}_\lambda| + h_{\mathcal{G}_n} |\nabla(\boldsymbol{V} - \boldsymbol{V}_\lambda)| \,\mathrm{d}x = 0,\end{aligned}$$

i. e., $\boldsymbol{V} = \boldsymbol{V}_{n,\lambda} = \Pi^n_{\mathrm{div}} \boldsymbol{V}_\lambda$ on E. This proves the assertion. □

The set $\Omega^n_\lambda(\boldsymbol{V})$ from Lemma 15 is larger than $\mathcal{U}_\lambda(\boldsymbol{V}) \cap \Omega$. However, the next result states that we can keep the increase of the set under control. This is the key observation for the construction of the discrete Lipschitz truncation.

LEMMA 16. *For $n \in \mathbb{N}$, $\boldsymbol{V} \in \mathbb{V}^n$ and $\lambda > 0$, let $\Omega^n_\lambda(\boldsymbol{V})$ be defined as in Lemma 15. Then, there exists a $\kappa \in (0,1)$ only depending on $\hat{\mathbb{P}}_{\mathbb{V}}$ and the shape-regularity of $\mathbb{G}$, such that*

$$\mathcal{U}_\lambda(\boldsymbol{V}) \cap \Omega \subset \Omega^n_\lambda(\boldsymbol{V}) \subset \mathcal{U}_{\kappa\lambda}(\boldsymbol{V}) \cap \Omega.$$

Proof. Thanks to the definition of $\Omega^n_\lambda(\boldsymbol{V})$, the first inclusion is clear. It thus remains to show the second inclusion. In order to avoid problems at the boundary $\partial\Omega$ we extend $\boldsymbol{V}$ to $\mathbb{R}^d$ by zero outside Ω. Let $x \in \Omega^n_\lambda(\boldsymbol{V})$; then, there exists $E \in \mathcal{G}_n$, $E \cap \mathcal{U}_\lambda(\boldsymbol{V}) \neq \emptyset$ such that $x \in \Omega^n_E$. Consequently, by (3.22a) and (3.19), there exists an $x_0 \in E$ and an $R > 0$ such that

$$\fint_{B_R(x_0)} |\nabla \boldsymbol{V}| \,\mathrm{d}y > \lambda.$$

Suppose that $B_R(x_0) \subset \left(\Omega^n_E \cup (\mathbb{R}^d \backslash \Omega)\right)$; then, thanks to norm-equivalence in finite-dimensional spaces, we have, by a standard scaling argument, that

$$\lambda < \fint_{B_R(x_0)} |\nabla \boldsymbol{V}| \,\mathrm{d}y \leqslant \|\nabla \boldsymbol{V}\|_{L^\infty(\Omega^n_E)} \leqslant \tilde{c}_1 \fint_{\Omega^n_E} |\nabla \boldsymbol{V}| \,\mathrm{d}y,$$

where the constant $\tilde{c}_1$ depends solely on $\hat{\mathbb{P}}_{\mathbb{V}}$ and the shape-regularity of $\mathbb{G}$. Let $B_\rho(x)$ be the smallest ball with $\overline{B_\rho(x)} \supset \Omega^n_E$ and observe that $|B_\rho(x)| \leqslant \tilde{c}_2 |\Omega^n_E|$ with a constant $\tilde{c}_2 > 0$ depending only on the shape-regularity of $\mathbb{G}$. Consequently,

$$M(\nabla \boldsymbol{V})(x) \geqslant \fint_{B_\rho(x)} |\nabla \boldsymbol{V}| \,\mathrm{d}y \geqslant \frac{1}{\tilde{c}_2} \fint_{\Omega^n_E} |\nabla \boldsymbol{V}| \,\mathrm{d}y \geqslant \frac{1}{\tilde{c}_1 \tilde{c}_2} \lambda.$$

In other words, we have that $x \in \mathcal{U}_{(\tilde{c}_1\tilde{c}_2)^{-1}\lambda}(\boldsymbol{V}) \cap \Omega$.

We now consider the case $B_R(x_0) \not\subset \big(\Omega_E^n \cup (\mathbb{R}^d \setminus \Omega)\big)$. Since $x_0 \in E$, it follows that $\tilde{c}_3 R \geqslant \operatorname{diam}(E)$, with a constant $\tilde{c}_3 > 0$ only depending on the shape-regularity of $\mathbb{G}$. As $x \in \Omega_E^n$, there exists a constant $\tilde{c}_4 > 1$ such that $B_{\tilde{c}_4 R}(x) \supset (E \cup B_R(x_0))$. Hence,

$$M(\nabla \boldsymbol{V})(x) \geqslant \fint_{B_{\tilde{c}_4 R}(x)} |\nabla \boldsymbol{V}| \,\mathrm{d}y \geqslant \tilde{c}_4^{-d} \fint_{B_R(x_0)} |\nabla \boldsymbol{V}| \,\mathrm{d}y > \tilde{c}_4^{-d}\lambda,$$

and we deduce that $x \in \mathcal{U}_{\tilde{c}_4^{-d}\lambda}(\boldsymbol{V}) \cap \Omega$. Combining the two cases, the claim follows with $\kappa := \min\{(\tilde{c}_1\tilde{c}_2)^{-1}, \tilde{c}_4^{-d}\}$. ◻

We are now ready to state the following analogue of Theorem 14 for the discrete Lipschitz truncation (3.26).

THEOREM 17. *Let $\lambda > 0$, $n \in \mathbb{N}$ and $\boldsymbol{V} \in \mathbb{V}^n$. Then, the Lipschitz truncation defined in* (3.26) *satisfies $\boldsymbol{V}_{n,\lambda} \in \mathbb{V}^n$, and the following statements hold:*

(a) *$\boldsymbol{V}_{n,\lambda} = \boldsymbol{V}$ on $\mathbb{R}^d \setminus \Omega_\lambda^n(\boldsymbol{V})$;*

(b) *$\|\boldsymbol{V}_{n,\lambda}\|_{1,s} \leqslant c\,\|\boldsymbol{V}\|_{1,s}$ for $1 < s \leqslant \infty$;*

(c) *$|\nabla \boldsymbol{V}_{n,\lambda}| \leqslant c\,\lambda \chi_{\Omega_\lambda^n(\boldsymbol{V})} + |\nabla \boldsymbol{V}|\,\chi_{\mathbb{R}^d \setminus \Omega_\lambda^n(\boldsymbol{V})} \leqslant c\,\lambda$ almost everywhere in $\mathbb{R}^d$.*

The constants c appearing in the inequalities in parts (b) *and* (c) *depend on Ω, d, $\hat{\mathbb{P}}_{\mathbb{V}}$ and the shape-regularity of $\mathbb{G}$. In* (b) *the constant c also depends on s.*

Proof. Assertion (a) is proved in Lemma 16. Estimate (b) is a consequence of Theorem 14 and the $W^{1,q}$-stability of Π_{div}^n; compare with (3.3). The bound (c) follows from Theorem 14(d) and the $W^{1,\infty}$ stability of Π_{div}^n; see (3.3). ◻

The following corollary is an application of the discrete Lipschitz truncation to (weak) null sequences. It is similar to the results in [DMS08] and [BDF12]. Its analogue for Sobolev functions is stated in Corollary 26 in the Appendix.

COROLLARY 18. *Let $1 < s < \infty$ and let $\{\boldsymbol{E}^n\}_{n\in\mathbb{N}} \subset W_0^{1,s}(\Omega)^d$ be a sequence, which converges to zero weakly in $W_0^{1,s}(\Omega)^d$, as $n \to \infty$.*

Then, there exists a sequence $\{\lambda_{n,j}\}_{n,j\in\mathbb{N}} \subset \mathbb{R}$ with $2^{2^j} \leqslant \lambda_{n,j} \leqslant 2^{2^{j+1}-1}$ such that the Lipschitz truncations $\boldsymbol{E}^{n,j} := \boldsymbol{E}^n_{n,\lambda_{n,j}}$, $n, j \in \mathbb{N}$, have the following properties:

(a) *$\boldsymbol{E}^{n,j} \in \mathbb{V}^n$ and $\boldsymbol{E}^{n,j} = \boldsymbol{E}^n$ on $\mathbb{R}^d \setminus \Omega^n_{\lambda_{n,j}}(\boldsymbol{E}^n)$;*

(b) *$\|\nabla \boldsymbol{E}^{n,j}\|_\infty \leqslant c\,\lambda_{n,j}$;*

(c) *$\boldsymbol{E}^{n,j} \to 0$ in $L^\infty(\Omega)^d$ as $n \to \infty$;*

(d) *$\nabla \boldsymbol{E}^{n,j} \rightharpoonup^* 0$ in $L^\infty(\Omega)^{d\times d}$ as $n \to \infty$;*

(e) *For all $n, j \in \mathbb{N}$ we have $\|\lambda_{n,j}\chi_{\Omega^n_{\lambda_{n,j}}(\boldsymbol{E}^n)}\|_s \leqslant c\,2^{-\frac{j}{s}}\|\nabla \boldsymbol{E}^n\|_s$.*

The constants c appearing in the inequalities (b) *and* (e) *depend on d, Ω, $\hat{\mathbb{P}}_{\mathbb{V}}$ and the shape-regularity of $\mathbb{G}$. The constant in part* (e) *also depends on s.*

Proof. We first construct the sequence $\lambda_{n,j}$ and prove (e). Let $\kappa > 0$ be the constant in Lemma 16. Then, for $g \in L^s(\mathbb{R}^d)$, we have

$$\begin{aligned}\int_{\mathbb{R}^d} |g|^s \,\mathrm{d}x &= \int_{\mathbb{R}^d} \int_0^\infty \kappa^s s t^{s-1} \chi_{\{|g|>\kappa t\}} \,\mathrm{d}t \,\mathrm{d}x \geqslant \int_{\mathbb{R}^d} \sum_{m\in\mathbb{N}} \kappa^s 2^{ms} \chi_{\{|g|>\kappa 2^{m+1}\}} \,\mathrm{d}x \\ &\geqslant \sum_{j\in\mathbb{N}} \sum_{m=2^j}^{2^{j+1}-1} \kappa^s 2^{ms} \left|\{|g| > \kappa 2^{m+1}\}\right|.\end{aligned}$$

We apply this estimate to $g = 2\,M(\nabla \boldsymbol{E}^n)$ and use the boundedness of the maximal operator M (cf. (3.20)) to obtain

$$\kappa^s \sum_{j\in\mathbb{N}} \sum_{m=2^j}^{2^{j+1}-1} 2^{ms} \left|\{M(\nabla \boldsymbol{E}^n) > \kappa 2^m\}\right| \leqslant 2^s \left\|M(\nabla \boldsymbol{E}^n)\right\|_s^s \leqslant 2^s c_s \|\nabla \boldsymbol{E}^n\|_s^s.$$

For fixed n, j the sum over m involves 2^j summands. Consequently, there exists an integer $\lambda_{n,j} \in \{2^{2^j}, \dots, 2^{2^{j+1}-1}\}$ such that

$$\lambda_{n,j}^s \left|\{M(\nabla \boldsymbol{E}^n) > \kappa\lambda_{n,j}\}\right| \leqslant 2^{-j}\kappa^{-s}\, 2^s c_s \|\nabla \boldsymbol{E}^n\|_s^s.$$

This, together with the second inclusion in Lemma 16, proves (e). Assertions (a) and (b) are then direct consequences of Theorem 17(a) and (b). It remains to prove (c) and (d).

To prove (d), we proceed as follows. Thanks to the uniqueness of the limits, it suffices to prove that $\boldsymbol{E}^{n,j} \rightharpoonup 0$ weakly in $W_0^{1,s}(\Omega)^d$. To this end, we note that the compact embedding $W_0^{1,s}(\Omega)^d \hookrightarrow\hookrightarrow L^s(\Omega)^d$ implies that

$$\boldsymbol{E}^n \to 0 \quad \text{in } L^s(\Omega)^d \text{ as } n \to \infty.$$

Let $\{\boldsymbol{E}^n_{\lambda_{n,j}}\}_{n\in\mathbb{N}}$ be the sequence of Lipschitz-truncated functions, defined according to (3.25). Then, thanks to the boundedness of $\{\boldsymbol{E}^n\}_{n\in\mathbb{N}}$ in $W_0^{1,s}(\Omega)^d$, Theorem 14(c) and (b), we have that

$$\boldsymbol{E}^n_{\lambda_{n,j}} \rightharpoonup 0 \quad \text{weakly in } W_0^{1,s}(\Omega)^d, \text{ as } n \to \infty.$$

Thanks to the definition of the discrete Lipschitz truncation in (3.26), the desired assertion follows from Proposition 7. Moreover, using a compact embedding, this also proves (c). □

4. The main theorem. After the preceding considerations, we are now ready to state our main result. Its proof is presented in subsections §4.1–§4.4.

THEOREM 19. *Let $\{\mathbb{V}^n, \mathbb{Q}^n\}_{n\in\mathbb{N}}$ be the sequence of finite element space pairs from Section 3.2 (respectively 3.3) and let $\{(\boldsymbol{U}^n, P^n, \boldsymbol{S}^n(\cdot, \boldsymbol{D}\boldsymbol{U}^n))\}_{n\in\mathbb{N}}$ be the sequence of discrete solution triples to* (3.16) *constructed in* (3.18).

If $r > \frac{2d}{d+1}$ (respectively $r > \frac{2d}{d+2}$), then there exists a solution $(\boldsymbol{u}, p, \boldsymbol{S}) \in W_0^{1,r}(\Omega)^d \times L_0^{\tilde{r}}(\Omega) \times L^{r'}(\Omega)^{d\times d}$ of (1.2)*, such that, for a (not relabeled) subsequence, we have*

$$\begin{aligned} \boldsymbol{U}^n &\rightharpoonup \boldsymbol{u} && \textit{weakly in } W_0^{1,r}(\Omega)^d,\\ P^n &\rightharpoonup p && \textit{weakly in } L_0^{\tilde{r}}(\Omega),\\ \boldsymbol{S}^n(\cdot, \boldsymbol{D}\boldsymbol{U}^n) &\rightharpoonup \boldsymbol{S} && \textit{weakly in } L^{r'}(\Omega)^{d\times d}. \end{aligned}$$

4.1. Convergence of the finite element approximations. We begin the proof of Theorem 19 by showing the existence of a weak limit for the sequence of solution triples.

LEMMA 20. *Let $\{\mathbb{V}^n, \mathbb{Q}^n\}_{n\in\mathbb{N}}$ be the sequence of finite element space pairs from Section 3.2 (respectively 3.3) and let $\{(\boldsymbol{U}^n, P^n, \boldsymbol{S}^n(\cdot, \boldsymbol{D}\boldsymbol{U}^n))\}_{n\in\mathbb{N}}$ be the sequence of discrete solution triples to* (3.16) *constructed in* (3.18).

If $r > \frac{2d}{d+1}$ (respectively $r > \frac{2d}{d+2}$), then there exists $(\boldsymbol{u}, p, \boldsymbol{S}) \in W_0^{1,r}(\Omega)^d \times L_0^{\tilde{r}}(\Omega) \times L^{r'}(\Omega)^{d\times d}$, such that, for a (not relabeled) subsequence, we have

$$\begin{aligned} \boldsymbol{U}^n &\rightharpoonup \boldsymbol{u} && \textit{weakly in } W_0^{1,r}(\Omega)^d,\\ P^n &\rightharpoonup p && \textit{weakly in } L_0^{\tilde{r}}(\Omega),\\ \boldsymbol{S}^n(\cdot, \boldsymbol{D}\boldsymbol{U}^n) &\rightharpoonup \boldsymbol{S} && \textit{weakly in } L^{r'}(\Omega)^{d\times d}. \end{aligned}$$

Moreover, the triple $(\boldsymbol{u}, p, \boldsymbol{S})$ *satisfies*

$$\begin{aligned} \int_\Omega \boldsymbol{S} : \boldsymbol{D}\boldsymbol{v} - (\boldsymbol{u} \otimes \boldsymbol{u}) : \nabla \boldsymbol{v} \,\mathrm{d}x - \langle \operatorname{div} \boldsymbol{v},\, p\rangle_\Omega = \langle \boldsymbol{f},\, \boldsymbol{v}\rangle, &\quad \textit{for all } \boldsymbol{v} \in W_0^{1,\tilde{r}'}(\Omega)^d \\ \int_\Omega q \operatorname{div} \boldsymbol{u} \,\mathrm{d}x = 0, &\quad \textit{for all } q \in L^{r'}(\Omega). \end{aligned} \tag{4.1}$$

Proof. We divide the proof into four steps.

Step 1: From (3.16) we see that $\boldsymbol{U}^n$ is discretely divergence-free and thus, thanks to (3.17) and (3.8) (respectively (3.13)), we have that

$$\int_\Omega \boldsymbol{S}^n(\cdot, \boldsymbol{D}\boldsymbol{U}^n) : \boldsymbol{D}\boldsymbol{U}^n \,\mathrm{d}x = \langle \boldsymbol{f},\, \boldsymbol{U}^n\rangle \leqslant \|\boldsymbol{f}\|_{-1,r'} \|\boldsymbol{U}^n\|_{1,r}.$$

The coercivity of $\boldsymbol{S}^n$ (Lemma 1) and Korn's inequality (2.2) imply that the sequence $\{\boldsymbol{U}^n\}_{n\in\mathbb{N}} \subset W_0^{1,r}(\Omega)^d$ is bounded, independent of $n \in \mathbb{N}$. This in turn implies, again by Lemma 1, the boundedness of $\{\boldsymbol{S}^n(\boldsymbol{D}\boldsymbol{U}^n)\}_{n\in\mathbb{N}}$ in $L^{r'}(\Omega)^{d\times d}$. In other words, there exists a constant $c > 0$ such that

$$\|\boldsymbol{U}^n\|_{1,r} + \|\boldsymbol{S}^n(\cdot, \boldsymbol{D}\boldsymbol{U}^n)\|_{r'} \leqslant c, \qquad \text{for all } n \in \mathbb{N}. \tag{4.2}$$

As $r \in (1, \infty)$, the spaces $W_0^{1,r}(\Omega)^d$ and $L^{r'}(\Omega)^{d\times d}$ are reflexive and thus for a (not relabeled) subsequence there exist $\boldsymbol{u} \in W_0^{1,r}(\Omega)^d$ and $\boldsymbol{S} \in L^{r'}(\Omega)^{d\times d}$, such that

$$\boldsymbol{U}^n \rightharpoonup \boldsymbol{u} \qquad \text{weakly in } W_0^{1,r}(\Omega)^d \tag{4.3}$$

and

$$\boldsymbol{S}^n(\cdot, \boldsymbol{D}\boldsymbol{U}^n) \rightharpoonup \boldsymbol{S} \qquad \text{weakly in } L^{r'}(\Omega)^{d\times d}, \tag{4.4}$$

as $n \to \infty$. Moreover, using compact embeddings of Sobolev spaces, we have that

$$\boldsymbol{U}^n \to \boldsymbol{u} \qquad \text{strongly in} \quad L^s(\Omega)^d \quad \text{for all} \quad \begin{cases} s \in \left(1, \frac{rd}{d-r}\right), & \text{if } r < d, \\ s \in (1, \infty), & \text{otherwise.} \end{cases} \tag{4.5}$$

Thanks to (4.3) we have by (3.5), for arbitrary $q \in L^{r'}(\Omega)$, that

$$0 = \int_\Omega (\Pi_{\mathbb{Q}}^n q) \operatorname{div} \boldsymbol{U}^n \,\mathrm{d}x \to \int_\Omega q \operatorname{div} \boldsymbol{u} \,\mathrm{d}x, \tag{4.6}$$

i. e., the function $\boldsymbol{u} \in W_0^{1,r}(\Omega)^d$ is exactly divergence-free.

Step 2: Next, we investigate the convection term. Let $\boldsymbol{v} \in W_0^{1,\infty}(\Omega)^d$ be arbitrary and define $\boldsymbol{V}^n := \Pi_{\mathrm{div}}^n \boldsymbol{v}$. We show that

$$\mathcal{B}[\boldsymbol{U}^n,\, \boldsymbol{U}^n,\, \boldsymbol{V}^n] \to -\int_\Omega (\boldsymbol{u} \otimes \boldsymbol{u}) : \nabla \boldsymbol{v} \,\mathrm{d}x. \tag{4.7}$$

Thanks to the assumption $r > \frac{2d}{d+2}$ and (4.5), it follows that

$$\boldsymbol{U}^n \otimes \boldsymbol{U}^n \to \boldsymbol{u} \otimes \boldsymbol{u} \qquad \text{in} \quad L^s(\Omega)^{d\times d} \quad \text{for all } s \in [1, \tilde{r}),$$

with $\tilde{r} > 1$ as in (1.1). By (3.4), we have that $\boldsymbol{V}^n \to \boldsymbol{v}$ in $W_0^{1,s'}(\Omega)^d$, $s' \in (\tilde{r}', \infty)$, and hence we obtain that, as $n \to \infty$,

$$-\int_\Omega (\boldsymbol{U}^n \otimes \boldsymbol{U}^n) : \nabla \boldsymbol{V}^n \,\mathrm{d}x \to -\int_\Omega (\boldsymbol{u} \otimes \boldsymbol{u}) : \nabla \boldsymbol{v} \,\mathrm{d}x. \tag{4.8}$$

This proves (4.7) for the exactly divergence-free approximations from Section 3.3. We emphasize that we have only required so far that $r > \frac{2d}{d+2}$.

In order to prove (4.7) for the finite element spaces of Section 3.2 and thus for the modified convection term defined in (3.7), we recall from (3.6) that

$$\int_\Omega (\boldsymbol{U}^n \otimes \boldsymbol{V}^n) : \nabla \boldsymbol{U}^n \,\mathrm{d}x = -\int_\Omega (\boldsymbol{U}^n \otimes \boldsymbol{U}^n) : \nabla \boldsymbol{V}^n + (\operatorname{div} \boldsymbol{U}^n)\, \boldsymbol{U}^n \cdot \boldsymbol{V}^n \,\mathrm{d}x.$$

For the first term we have already shown convergence in (4.8). In view of the definition of $\mathcal{B}$ in (3.7) it thus remains to prove that the second term vanishes in the limit $n \to \infty$. To this end, we observe by (4.5) and Assumption 6 that

$$\boldsymbol{U}^n \cdot \boldsymbol{V}^n \to \boldsymbol{u} \cdot \boldsymbol{v} \quad \text{strongly in} \quad L^s(\Omega) \quad \text{for all} \quad \begin{cases} s \in \left(1, \frac{rd}{d-r}\right), & \text{if } r < d, \\ s \in (1, \infty), & \text{otherwise.} \end{cases}$$

Thanks to the stronger restriction $r > \frac{2d}{d+1}$ now, this last statement holds in particular for $s = r'$. Hence, together with (4.3) and (4.6), we deduce that

$$\int_\Omega (\operatorname{div} \boldsymbol{U}^n)\, \boldsymbol{U}^n \cdot \boldsymbol{V}^n \,\mathrm{d}x \to \int_\Omega (\operatorname{div} \boldsymbol{u})\, \boldsymbol{u} \cdot \boldsymbol{v} \,\mathrm{d}x = 0$$

as $n \to \infty$.

Step 3: We combine the above results. Recall that by (4.6) we have $\operatorname{div} \boldsymbol{u} \equiv 0$, which is the second equation in (4.1). For an arbitrary $\boldsymbol{v} \in W^{1,\infty}_{0,\operatorname{div}}(\Omega)^d$ let $\boldsymbol{V}^n := \Pi^n_{\operatorname{div}} \boldsymbol{v}$, $n \in \mathbb{N}$. Thanks to (3.4), we have that $\boldsymbol{V}^n \in \mathbb{V}^n_{\operatorname{div}}$ and $\boldsymbol{V}^n \to \boldsymbol{v}$ in $W^{1,s}_0(\Omega)^d$ for all $s \in (1, \infty)$. Therefore, using (4.3), (4.4) and (4.7), we obtain

$$\begin{array}{ccc} \displaystyle\int_\Omega \boldsymbol{S}^n(\cdot, \boldsymbol{D}\boldsymbol{U}^n) : \boldsymbol{D}\boldsymbol{V}^n \,\mathrm{d}x + \mathcal{B}[\boldsymbol{U}^n, \boldsymbol{U}^n, \boldsymbol{V}^n] & = & \langle \boldsymbol{f}, \boldsymbol{V}^n \rangle \\ \downarrow & & \downarrow \\ \displaystyle\int_\Omega \boldsymbol{S} : \boldsymbol{D}\boldsymbol{v} + \operatorname{div}(\boldsymbol{u} \otimes \boldsymbol{u}) \cdot \boldsymbol{v} \,\mathrm{d}x & & \langle \boldsymbol{f}, \boldsymbol{v} \rangle \end{array}$$

as $n \to \infty$. Since $\boldsymbol{S} \in L^{r'}(\Omega)^{d\times d}$, $\boldsymbol{f} \in W^{-1,r'}(\Omega)^d$ and $\boldsymbol{u} \otimes \boldsymbol{u} \in L^{\tilde{r}}(\Omega)^{d \times d}$, by a density argument, we arrive at

$$\int_\Omega \boldsymbol{S} : \boldsymbol{D}\boldsymbol{v} + \operatorname{div}(\boldsymbol{u} \otimes \boldsymbol{u}) \cdot \boldsymbol{v} \,\mathrm{d}x = \langle \boldsymbol{f}, \boldsymbol{v} \rangle$$

for all $\boldsymbol{v} \in W^{1,\tilde{r}'}_{0,\operatorname{div}}(\Omega)^d$.

Step 4: We now prove convergence of the pressure. Thanks to the restriction $r > \frac{2d}{d+1}$ we have, as in (3.11), that

$$\begin{aligned} \langle \operatorname{div} \boldsymbol{V}, P^n \rangle_\Omega &= \int_\Omega \boldsymbol{S}^n(\cdot, \boldsymbol{D}\boldsymbol{U}^n) : \boldsymbol{D}\boldsymbol{V} \,\mathrm{d}x + \mathcal{B}[\boldsymbol{U}^n, \boldsymbol{U}^n, \boldsymbol{V}] - \langle \boldsymbol{f}, \boldsymbol{V} \rangle \\ &\leqslant \|\boldsymbol{S}^n(\cdot, \boldsymbol{D}\boldsymbol{U}^n)\|_{r'} \|\boldsymbol{D}\boldsymbol{V}\|_r + c\, \|\boldsymbol{U}^n\|^2_{1,r} \|\boldsymbol{V}\|_{1,\tilde{r}'} + \|\boldsymbol{f}\|_{-1,r'} \|\boldsymbol{V}\|_{1,r} \end{aligned}$$

for all $\boldsymbol{V} \in \mathbb{V}^n$. By (4.2) and the discrete inf-sup condition stated in Proposition 9, it follows that the sequence $\{P^n\}_{n\in\mathbb{N}}$ is bounded in the reflexive Banach space $L^{\tilde{r}}_0(\Omega)$. Hence, there exists $p \in L^{\tilde{r}}_0(\Omega)$ such that, for a (not relabeled) subsequence, $P^n \rightharpoonup p$ weakly in $L^{\tilde{r}}_0(\Omega)$. On the other hand we deduce for an arbitrary $\boldsymbol{v} \in W^{1,\infty}_0(\Omega)^d$ that

$$\begin{aligned} \langle \operatorname{div} \boldsymbol{v}, P^n \rangle_\Omega &= \langle \operatorname{div} \Pi^n_{\operatorname{div}} \boldsymbol{v}, P^n \rangle_\Omega + \langle \operatorname{div}(\boldsymbol{v} - \Pi^n_{\operatorname{div}} \boldsymbol{v}), P^n \rangle_\Omega \\ &= \int_\Omega \boldsymbol{S}^n(\cdot, \boldsymbol{D}\boldsymbol{U}^n) : \boldsymbol{D}\Pi^n_{\operatorname{div}} \boldsymbol{v} \,\mathrm{d}x - \langle \boldsymbol{f}, \Pi^n_{\operatorname{div}} \boldsymbol{v} \rangle + \mathcal{B}[\boldsymbol{U}^n, \boldsymbol{U}^n, \Pi^n_{\operatorname{div}} \boldsymbol{v}] \\ &\quad + \langle \operatorname{div}(\boldsymbol{v} - \Pi^n_{\operatorname{div}} \boldsymbol{v}), P^n \rangle_\Omega \\ &\to \int_\Omega \boldsymbol{S} : \boldsymbol{D}\boldsymbol{v} + \operatorname{div}(\boldsymbol{u} \otimes \boldsymbol{u}) \cdot \boldsymbol{v} \,\mathrm{d}x - \langle \boldsymbol{f}, \boldsymbol{v} \rangle + 0 \end{aligned}$$

as $n \to \infty$, where we have used (4.4), (4.7), (3.4) and the boundedness of the sequence $\{P^n\}_{n\in\mathbb{N}}$ in $L_0^{\tilde{r}}(\Omega)$. This completes the proof of the lemma. □

For the main result, Theorem 19, it remains to prove that

$$\big(\boldsymbol{D}\boldsymbol{u}(x),\ \boldsymbol{S}(x)\big) \in \mathcal{A}(x) \tag{4.9}$$

for almost every $x \in \Omega$. The proof of this is the subject of the rest of Section 4.

4.2. Identification of the limits. In this section we shall first briefly discuss properties of the maximal monotone r-graph introduced in (A1)–(A5). Here we follow the presentation in [BGMŚG09]. Application of the fundamental theorem on Young measures (cf. Theorem 3) leads to a representation of weak limits, which is a crucial step in proving (4.9).

According to [FMT04] there exists a function $\phi : \Omega \times \mathbb{R}^{d\times d}_{\mathrm{sym}}$ such that

$$\mathcal{A}(x) = \left\{(\boldsymbol{\delta},\boldsymbol{\sigma}) \in \mathbb{R}^{d\times d}_{\mathrm{sym}} \times \mathbb{R}^{d\times d}_{\mathrm{sym}} : \boldsymbol{\sigma} - \boldsymbol{\delta} = \phi(x, \boldsymbol{\sigma} + \boldsymbol{\delta})\right\}, \tag{4.10}$$

and

(a) $\phi(x,\boldsymbol{0}) = \boldsymbol{0}$ for almost every $x \in \Omega$;

(b) $\phi(\cdot,\boldsymbol{\chi})$ is measurable for all $\boldsymbol{\chi} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$;

(c) for almost all $x \in \Omega$ the mapping $\phi(x,\cdot)$ is 1-Lipschitz continuous;

(d) the functions $\boldsymbol{s}, \boldsymbol{d} : \Omega \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$, defined as

$$\boldsymbol{s}(x,\boldsymbol{\chi}) := \tfrac{1}{2}\left(\boldsymbol{\chi} + \phi(x,\boldsymbol{\chi})\right), \quad \boldsymbol{d}(x,\boldsymbol{\chi}) := \tfrac{1}{2}\left(\boldsymbol{\chi} - \phi(x,\boldsymbol{\chi})\right) \tag{4.11}$$

satisfy, for almost every $x \in \Omega$ and all $\boldsymbol{\chi} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$, the estimate

$$\boldsymbol{s}(x,\boldsymbol{\chi}) : \boldsymbol{d}(x,\boldsymbol{\chi}) \geqslant -m(x) + c\left(|\boldsymbol{d}(x,\boldsymbol{\chi})|^r + |\boldsymbol{s}(x,\boldsymbol{\chi})|^{r'}\right).$$

We emphasize that this is in fact a characterization of maximal monotone r-graphs $\mathcal{A}$ satisfying (A1)–(A5) without the second part of (A2).

We recall the selection $\boldsymbol{S}^*$ from Section 2.2 and, as in [BGMŚG09], we define

$$b_n(x) := \int_{\mathbb{R}^{d\times d}_{\mathrm{sym}}} \big(\boldsymbol{S}^*(x,\boldsymbol{\zeta}) - \boldsymbol{S}^*(x,\boldsymbol{D}\boldsymbol{u})\big) : (\boldsymbol{\zeta} - \boldsymbol{D}\boldsymbol{u}(x))\,\mathrm{d}\mu^n_x(\boldsymbol{\zeta}), \tag{4.12}$$

where we have used the abbreviation $\mu^n_x := \mu^n_{\boldsymbol{D}\boldsymbol{U}^n(x)}$. The next result, whose proof is postponed to the next section, states that b_n vanishes in measure.

LEMMA 21. *With the definitions of this section we have that $b_n \to 0$ in measure.*

Actually, employing the above characterization of $\mathcal{A}$, the limit of the sequence $\{b_n\}_{n\in\mathbb{N}}$ can be identified in another way by using the theory of Young measures. To this end we introduce

$$\boldsymbol{G}_x(\boldsymbol{\zeta}) := \boldsymbol{S}^*(x,\boldsymbol{\zeta}) + \boldsymbol{\zeta}, \qquad x \in \Omega,\ \boldsymbol{\zeta} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}, \tag{4.13}$$

and define the push-forward measure of the measure $\mu^n_{\boldsymbol{\zeta}}$ from (2.5) by setting

$$\nu^n_{x,\boldsymbol{\zeta}}(\mathcal{C}) = \mu^n_{\boldsymbol{\zeta}}\big(\boldsymbol{G}_x^{-1}(\mathcal{C})\big) \qquad \text{for all } \mathcal{C} \in \mathfrak{B}(\mathbb{R}^{d\times d}_{\mathrm{sym}}). \tag{4.14}$$

We recall from §2.2(a1) that $\boldsymbol{S}^*$ is measurable with respect to the product σ-algebra $\mathfrak{L}(\Omega) \otimes \mathfrak{B}(\mathbb{R}^{d\times d}_{\mathrm{sym}})$, and therefore so is $\boldsymbol{G}_x$. Consequently, the measure $\nu^n_{x,\boldsymbol{\zeta}}$ is well-defined and, thanks to properties of the mollifier η^n from Section 2.2, it is a probability measure. From the definitions of the functions $\boldsymbol{s}$ and $\boldsymbol{d}$ it follows that

$\boldsymbol{S}^*(x,\boldsymbol{\zeta}) = \boldsymbol{s}(x,\boldsymbol{G}_x(\boldsymbol{\zeta}))$ and $\boldsymbol{\zeta} = \boldsymbol{d}(x,\boldsymbol{G}_x(\boldsymbol{\zeta}))$. We thus have, by simple substitution, the identities

$$\boldsymbol{S}^n(x,\boldsymbol{D}\boldsymbol{U}^n) = \int_{\mathbb{R}^{d\times d}_{\mathrm{sym}}} \boldsymbol{s}(x,\boldsymbol{\zeta})\,\mathrm{d}\nu^n_x(\boldsymbol{\zeta}), \tag{4.15a}$$

$$\boldsymbol{D}\boldsymbol{U}^n(x) = \int_{\mathbb{R}^{d\times d}_{\mathrm{sym}}} \boldsymbol{d}(x,\boldsymbol{\zeta})\,\mathrm{d}\nu^n_x(\boldsymbol{\zeta}), \tag{4.15b}$$

as well as

$$b_n(x) = \int_{\mathbb{R}^{d\times d}_{\mathrm{sym}}} \big(\boldsymbol{s}(x,\boldsymbol{\zeta}) - \boldsymbol{S}^*(x,\boldsymbol{D}\boldsymbol{u}(x))\big) : (\boldsymbol{d}(x,\boldsymbol{\zeta}) - \boldsymbol{D}\boldsymbol{u}(x))\,\mathrm{d}\nu^n_x(\boldsymbol{\zeta}), \tag{4.15c}$$

where we have used the abbreviation $\nu^n_x := \nu^n_{x,\boldsymbol{D}\boldsymbol{U}^n(x)}$.

In order to identify the limit we apply the generalized version of the classical fundamental theorem on Young measures stated in Theorem 3. Recall from Section 2.3 that $L^\infty_w(\Omega;\mathfrak{M}(\mathbb{R}^{d\times d}_{\mathrm{sym}}))$ is a separable Banach space with predual $L^1(\Omega,C_0(\mathbb{R}^{d\times d}_{\mathrm{sym}}))$. For every $n\in\mathbb{N}$ the mapping $x\mapsto\nu^n_x$ belongs to $L^\infty_w(\Omega;\mathfrak{M}(\mathbb{R}^{d\times d}_{\mathrm{sym}}))$. To see this let $g\in L^1(\Omega;C_0(\mathbb{R}^{d\times d}_{\mathrm{sym}}))$. In view of (4.14) and (2.5), a simple substitution yields

$$\int_{\mathbb{R}^{d\times d}_{\mathrm{sym}}} g(x,\boldsymbol{\zeta})\,\mathrm{d}\nu^n_x(\boldsymbol{\zeta}) = \int_{\mathbb{R}^{d\times d}_{\mathrm{sym}}} \eta^n(\boldsymbol{D}\boldsymbol{U}^n(x)-\boldsymbol{\zeta})\,g(x,\boldsymbol{G}_x(\boldsymbol{\zeta}))\,\mathrm{d}\boldsymbol{\zeta}.$$

It remains to prove that $x\mapsto\int_{\mathbb{R}^{d\times d}_{\mathrm{sym}}}\eta^n(\boldsymbol{D}\boldsymbol{U}^n(x)-\boldsymbol{\zeta})\,g(x,\boldsymbol{G}_x(\boldsymbol{\zeta}))\,\mathrm{d}\boldsymbol{\zeta}$ is measurable and integrable. It follows from (A5), the definition (4.13) of $\boldsymbol{G}_x$ and property (a1) of $\boldsymbol{S}^*$ that h is $\mathfrak{L}(\Omega)\otimes\mathfrak{B}(\mathbb{R}^{d\times d}_{\mathrm{sym}})$ measurable. Moreover, η^n, $n\geqslant 1$, are smooth functions and $g\in L^1(\Omega,C_0(\mathbb{R}^{d\times d}_{\mathrm{sym}}))$, and therefore integrability follows from Fubini's theorem.

Thanks to the properties of η^n it is clear that ν^n_x is a probability measure a.e. in Ω. Hence $\|\nu^n_x\|_{L^\infty_w(\Omega;\mathfrak{M}(\mathbb{R}^{d\times d}_{\mathrm{sym}}))} = 1$ and thus the sequence $\{\nu^n\}_{n\in\mathbb{N}}$ is bounded in $L^\infty_w(\Omega;\mathfrak{M}(\mathbb{R}^{d\times d}_{\mathrm{sym}}))$. Therefore, by the Banach–Alaoglu theorem, there exists $\nu\in L^\infty_w(\Omega;\mathfrak{M}(\mathbb{R}^{d\times d}_{\mathrm{sym}}))$ such that, for a (not relabeled) subsequence,

$$\nu^n \rightharpoonup^* \nu \qquad \text{weak-}* \text{ in } L^\infty_w(\Omega;\mathfrak{M}(\mathbb{R}^{d\times d}_{\mathrm{sym}})). \tag{4.16}$$

The next Lemma follows from the generalization of the fundamental theorem on Young measures from [Gwi05] (see Theorem 3) and Chacon's biting lemma (Lemma 4). Its proof is postponed to Section 4.4.

LEMMA 22. *With the notations of this section, ν_x is a probability measure a.e. in Ω and there exists a nonincreasing sequence of measurable subsets $E_k\subset\Omega$, with $|E_k|\to 0$, such that for all $k\in\mathbb{N}$ we have for a (not relabeled) subsequence that*

$$b_n(x) \rightharpoonup \int_{\mathbb{R}^{d\times d}_{sym}} \big(\boldsymbol{s}(x,\boldsymbol{\zeta}) - \boldsymbol{S}^*(x,\boldsymbol{\zeta})\big) : (\boldsymbol{d}(x,\boldsymbol{\zeta}) - \boldsymbol{D}\boldsymbol{u}(x))\,\mathrm{d}\nu_x(\boldsymbol{\zeta}) =: b(x) \tag{4.17}$$

weakly in $L^1(\Omega\backslash E_k)$ as $n\to\infty$. Moreover, we have that

$$\boldsymbol{S}(x) = \int_{\mathbb{R}^{d\times d}_{sym}} \boldsymbol{s}(x,\boldsymbol{\zeta})\,\mathrm{d}\nu_x(\boldsymbol{\zeta}) \qquad \textit{and} \qquad \boldsymbol{D}\boldsymbol{u}(x) = \int_{\mathbb{R}^{d\times d}_{sym}} \boldsymbol{d}(x,\boldsymbol{\zeta})\,\mathrm{d}\nu_x(\boldsymbol{\zeta}).$$

We deduce from (4.9) and Lemma 22 that, to complete the proof of Theorem 19, we need to show that

$$\left(\boldsymbol{D}\boldsymbol{u}(x),\ \int_{\mathbb{R}^{d\times d}_{\mathrm{sym}}} \boldsymbol{s}(x,\boldsymbol{\zeta})\,\mathrm{d}\nu_x(\boldsymbol{\zeta})\right) \in \mathcal{A}(x) \qquad \text{for a.e. } x\in\Omega. \tag{4.18}$$

This follows from the two preceding Lemmas exactly as in [BGMŚG09, p. 131ff]. To be more precise, the proof is based on noting that for each $\boldsymbol{\delta} \in \mathbb{R}^{d\times d}$ the set

$$\mathcal{C}^x_{\boldsymbol{\delta}} := \{\boldsymbol{\sigma} \in \mathbb{R}^{d\times d}_{\mathrm{sym}} \colon (\boldsymbol{\delta}, \boldsymbol{\sigma}) \in \mathcal{A}(x)\} \quad \text{is convex for a.e. } x \in \Omega; \tag{4.19}$$

compare with (A2) and (A3). Combining Lemmas 21 and 22 we deduce that $b = 0$ a.e. in Ω. Hence it follows from (A2) that

$$\{\boldsymbol{\zeta} \in \mathbb{R}^{d\times d}_{\mathrm{sym}} \colon \big(\boldsymbol{s}(x,\boldsymbol{\zeta}) - \boldsymbol{S}^*(x, \boldsymbol{D}\boldsymbol{u}(x))\big) : (\boldsymbol{d}(x,\boldsymbol{\zeta}) - \boldsymbol{D}\boldsymbol{u}(x)) > 0\} \not\subset \operatorname{supp} \nu_x.$$

We split $\operatorname{supp}\nu_x$ into the two sets

$$\omega_1(x) := \{\boldsymbol{\zeta} \in \operatorname{supp}\nu_x \colon \boldsymbol{s}(x,\boldsymbol{\zeta}) = \boldsymbol{S}^*(x, \boldsymbol{D}\boldsymbol{u}(x))\} \quad \text{and} \quad \omega_2(x) := \operatorname{supp}\nu_x \setminus \omega_1(x).$$

We investigate the pairing in (4.18) on the two sets $\omega_1(x)$ and $\omega_2(x)$ separately. On $\omega_2(x)$ we have by (A2) that $\boldsymbol{d}(x,\boldsymbol{\zeta}) = \boldsymbol{D}\boldsymbol{u}(x)$. Therefore, on noting that $\frac{\nu_x(\boldsymbol{\zeta})}{\nu_x(\omega_2(x))}$ is a probability measure on $\omega_2(x)$, one can show with (4.10), (4.11) and (4.19) that

$$\left(\boldsymbol{D}\boldsymbol{u}(x), \int_{\omega_2(x)} \boldsymbol{s}(x,\boldsymbol{\zeta}) \,\mathrm{d}\big(\tfrac{\nu_x(\boldsymbol{\zeta})}{\nu_x(\omega_2(x))}\big)\right) \in \mathcal{A}(x) \quad \text{for a.e. } x \in \Omega.$$

On the other hand it follows from the definition of $\omega_1(x)$ that

$$\int_{\omega_1(x)} \boldsymbol{s}(x,\boldsymbol{\zeta}) \,\mathrm{d}\nu_x(\boldsymbol{\zeta}) = \nu_x\big(\omega_1(x)\big)\boldsymbol{S}^*(x, \boldsymbol{D}\boldsymbol{u}(x)).$$

Thanks to the properties of $\boldsymbol{S}^*$, we have that $\big(\boldsymbol{D}\boldsymbol{u}(x),\, \boldsymbol{S}^*(x, \boldsymbol{D}\boldsymbol{u}(x))\big) \in \mathcal{A}(x)$; compare with Section 2.2. Using the fact that ν_x is a probability measure, we deduce that

$$\begin{aligned}\int_{\mathbb{R}^{d\times d}_{\mathrm{sym}}} \boldsymbol{s}(x,\boldsymbol{\zeta}) \,\mathrm{d}\nu_x(\boldsymbol{\zeta}) &= \int_{\omega_1(x)} \boldsymbol{s}(x,\boldsymbol{\zeta}) \,\mathrm{d}\nu_x(\boldsymbol{\zeta}) + \int_{\omega_2(x)} \boldsymbol{s}(x,\boldsymbol{\zeta}) \,\mathrm{d}\nu_x(\boldsymbol{\zeta}) \\ &= \nu_x(\omega_1(x))\boldsymbol{S}^*(x, \boldsymbol{D}\boldsymbol{u}(x)) + \nu_x(\omega_2(x)) \int_{\omega_2(x)} \boldsymbol{s}(x,\boldsymbol{\zeta}) \,\mathrm{d}\big(\tfrac{\nu_x(\boldsymbol{\zeta})}{\nu_x(\omega_2(x))}\big)\end{aligned}$$

is a convex combination of functions. Moreover, due to the above observations, for a.e. $x \in \Omega$, each of the two functions in this convex combination is an element of the set $\mathcal{C}^x_{\boldsymbol{D}\boldsymbol{u}(x)}$. Hence, by (4.19), this completes the proof of Theorem 19.

As in [BGMŚG09] we can establish from the above observations strong convergence of the symmetric velocity gradient and the stress on the subsets

$$\begin{aligned}\Omega_{\boldsymbol{D}} := \{x \in \Omega \colon \forall (\boldsymbol{\sigma}_1, \boldsymbol{\delta}_1) \in \mathcal{A}(x) \text{ with } \big(\boldsymbol{\sigma}_1 - \boldsymbol{S}^*(x, \boldsymbol{D}\boldsymbol{u}(x))\big) : (\boldsymbol{\delta}_1 - \boldsymbol{D}\boldsymbol{u}(x)) = 0 \\ \text{implies that } \boldsymbol{\delta}_1 = \boldsymbol{D}\boldsymbol{u}(x)\}, \quad \text{and} \\ \Omega_{\boldsymbol{S}} := \{x \in \Omega \colon \forall (\boldsymbol{\sigma}_1, \boldsymbol{\delta}_1) \in \mathcal{A}(x) \text{ with } \big(\boldsymbol{\sigma}_1 - \boldsymbol{S}^*(x, \boldsymbol{D}\boldsymbol{u}(x))\big) : (\boldsymbol{\delta}_1 - \boldsymbol{D}\boldsymbol{u}(x)) = 0 \\ \text{implies that } \boldsymbol{\sigma}_1 = \boldsymbol{S}^*(x, \boldsymbol{D}\boldsymbol{u}(x))\},\end{aligned}$$

respectively. Since the proof is identical to the proof of [BGMŚG09, Lemma 5.2] we omit it here and we only state the result.

COROLLARY 23. *Assume the conditions of Theorem 19 and let $r' \in (1,\infty)$ be such that $\frac{1}{r} + \frac{1}{r'} = 1$. Then, for all $1 \leqslant s < r$ and $1 \leqslant s' < r'$, we have that, as $n \to \infty$,*

$$\begin{aligned}\boldsymbol{D}\boldsymbol{U}^n &\to \boldsymbol{D}\boldsymbol{u} && \text{strongly in } L^s(\Omega_{\boldsymbol{D}})^{d\times d}, \\ \boldsymbol{S}^n &\to \boldsymbol{S}^*(\cdot, \boldsymbol{D}\boldsymbol{u}) && \text{strongly in } L^{s'}(\Omega_{\boldsymbol{S}})^{d\times d}.\end{aligned}$$

4.3. Proof of Lemma 21. The proof of this assertion is motivated by the proof of [BGMŚG09, Lemma 4.6]. However, since we are approximating problem (1.2) with finite element functions here, we need to use the discrete Lipschitz truncation from §3.5.

Let us define the auxiliary function

$$a_n(x) := \big(\boldsymbol{S}^n(x, \boldsymbol{D}\boldsymbol{U}^n(x)) - \boldsymbol{S}^*(x, \boldsymbol{D}\boldsymbol{u}(x))\big) : (\boldsymbol{D}\boldsymbol{U}^n(x) - \boldsymbol{D}\boldsymbol{u}(x)) \tag{4.20}$$

and observe that

$$\int_\Omega |a_n - b_n| \,\mathrm{d}x \to 0 \qquad \text{as} \quad n \to \infty. \tag{4.21}$$

Indeed, thanks to (2.5) and the properties of η^n, we have that

$$\begin{aligned}
\int_\Omega |a_n - b_n| \,\mathrm{d}x &= \int_\Omega \Big| \int_{\mathbb{R}^{d\times d}_{\mathrm{sym}}} \big(\boldsymbol{S}^*(x,\boldsymbol{\zeta}) - \boldsymbol{S}^*(x,\boldsymbol{D}\boldsymbol{u})\big) : (\boldsymbol{D}\boldsymbol{U}^n - \boldsymbol{D}\boldsymbol{u}) \,\mathrm{d}\mu^n_x(\boldsymbol{\zeta}) \\
&\qquad - \int_{\mathbb{R}^{d\times d}_{\mathrm{sym}}} \big(\boldsymbol{S}^*(x,\boldsymbol{\zeta}) - \boldsymbol{S}^*(x,\boldsymbol{D}\boldsymbol{u})\big) : (\boldsymbol{\zeta} - \boldsymbol{D}\boldsymbol{u}) \,\mathrm{d}\mu^n_x(\boldsymbol{\zeta}) \Big| \,\mathrm{d}x \\
&= \int_\Omega \Big| \int_{\mathbb{R}^{d\times d}_{\mathrm{sym}}} \big(\boldsymbol{S}^*(x,\boldsymbol{\zeta}) - \boldsymbol{S}^*(x,\boldsymbol{D}\boldsymbol{u})\big) : (\boldsymbol{D}\boldsymbol{U}^n - \boldsymbol{\zeta}) \,\mathrm{d}\mu^n_x(\boldsymbol{\zeta}) \Big| \,\mathrm{d}x \\
&\leqslant \int_\Omega \int_{\mathbb{R}^{d\times d}_{\mathrm{sym}}} \big|\boldsymbol{S}^*(x,\boldsymbol{\zeta}) - \boldsymbol{S}^*(x,\boldsymbol{D}\boldsymbol{u})\big| \big|\boldsymbol{D}\boldsymbol{U}^n - \boldsymbol{\zeta}\big| \,\mathrm{d}\mu^n_x(\boldsymbol{\zeta}) \,\mathrm{d}x \\
&\leqslant \frac{c}{n} \int_\Omega \sup_{|\boldsymbol{\zeta} - \boldsymbol{D}\boldsymbol{U}^n(x)| \leqslant \frac{1}{n}} \big|\boldsymbol{S}^*(x,\boldsymbol{\zeta}) - \boldsymbol{S}^*(x,\boldsymbol{D}\boldsymbol{U}^n)\big| \,\mathrm{d}x \leqslant \frac{c}{n}.
\end{aligned}$$

Consequently, in order to prove that $b_n \to 0$ in measure it suffices to prove that $a_n \to 0$ in measure. We shall establish the second claim in several steps.

Step 1: First we introduce some preliminary facts concerning discrete Lipschitz truncations. For convenience we use the notation

$$\boldsymbol{E}^n := \Pi^n_{\mathrm{div}}(\boldsymbol{U}^n - \boldsymbol{u}) = \boldsymbol{U}^n - \Pi^n_{\mathrm{div}}\boldsymbol{u} \in \mathbb{V}^n$$

and let $\{\boldsymbol{E}^{n,j}\}_{n,j\in\mathbb{N}} \subset \mathbb{V}^n$ be the sequence of Lipschitz-truncated finite element functions according to Corollary 18. Recall from Lemma 20 that $\boldsymbol{E}^n \rightharpoonup 0$ weakly in $W^{1,r}_0(\Omega)^d$, i. e., we are exactly in the situation of Corollary 18. Although $\boldsymbol{E}^n \in \mathbb{V}^n_{\mathrm{div}}$, i. e., it is discretely divergence-free, this does not necessarily imply that $\boldsymbol{E}^{n,j} \in \mathbb{V}^n_{\mathrm{div}}$ and thus we need to modify $\boldsymbol{E}^{n,j}$ in order to be able to use it as a test function in (3.17). Recalling Corollary 10 we define

$$\boldsymbol{\Psi}^{n,j} := \mathfrak{B}^n(\operatorname{div} \boldsymbol{E}^{n,j}) \in \mathbb{V}^n. \tag{4.22a}$$

The 'corrected' function

$$\boldsymbol{\Phi}^{n,j} := \boldsymbol{E}^{n,j} - \boldsymbol{\Psi}^{n,j} \in \mathbb{V}^n_{\mathrm{div}} \tag{4.22b}$$

is then discretely divergence-free. We need to control the correction in a norm. To this end we recall from Section 3.1 that $\mathbb{Q}^n = \operatorname{span}\{Q^n_1, \dots, Q^n_{N_n}\}$ for a certain locally supported basis. Then, thanks to properties of the discrete Bogovskiĭ operator and Corollary 10, we have that

$$\begin{aligned}
\beta_r \|\boldsymbol{\Psi}^{n,j}\|_{1,r} &\leqslant \sup_{Q\in\mathbb{Q}^n} \frac{\langle \operatorname{div} \boldsymbol{E}^{n,j},\, Q\rangle_\Omega}{\|Q\|_{r'}} = \sup_{Q\in\mathbb{Q}^n} \frac{\langle \operatorname{div} \boldsymbol{E}^{n,j} - \operatorname{div} \boldsymbol{E}^{n},\, Q\rangle_\Omega}{\|Q\|_{r'}} \\
&= \sup_{Q=\sum_{i=1}^{N_n}\rho_i Q_i^n} \Bigg(\sum_{\operatorname{supp} Q_i^n \subset \{\boldsymbol{E}^{n,j}=\boldsymbol{E}^n\}} \frac{\langle \operatorname{div} \boldsymbol{E}^{n,j} - \operatorname{div} \boldsymbol{E}^{n},\, \rho_i Q_i^n\rangle_\Omega}{\|Q\|_{r'}} \\
&\qquad\qquad + \sum_{\operatorname{supp} Q_i^n \cap \{\boldsymbol{E}^{n,j}\neq\boldsymbol{E}^n\}\neq\emptyset} \frac{\langle \operatorname{div} \boldsymbol{E}^{n,j} - \operatorname{div} \boldsymbol{E}^{n},\, \rho_i Q_i^n\rangle_\Omega}{\|Q\|_{r'}} \Bigg) \\
&= \sup_{Q=\sum_{i=1}^{N_n}\rho_i Q_i^n} \Bigg(\sum_{\operatorname{supp} Q_i^n \cap \{\boldsymbol{E}^{n,j}\neq\boldsymbol{E}^n\}\neq\emptyset} \frac{\langle \operatorname{div} \boldsymbol{E}^{n,j} - \operatorname{div} \boldsymbol{E}^{n},\, \rho_i Q_i^n\rangle_\Omega}{\|Q\|_{r'}} \Bigg) \\
&= \sup_{Q=\sum_{i=1}^{N_n}\rho_i Q_i^n} \Bigg(\sum_{\operatorname{supp} Q_i^n \cap \{\boldsymbol{E}^{n,j}\neq\boldsymbol{E}^n\}\neq\emptyset} \frac{\langle \operatorname{div} \boldsymbol{E}^{n,j},\, \rho_i Q_i^n\rangle_\Omega}{\|Q\|_{r'}} \Bigg) \\
&\leqslant \|\operatorname{div} \boldsymbol{E}^{n,j} \chi_{\Omega^n_{\{\boldsymbol{E}^{n,j}\neq\boldsymbol{E}^n\}}}\|_r \sup_{Q=\sum_{i=1}^{N_n}\rho_i Q_i^n} \frac{\|\sum_{\operatorname{supp} Q_i^n \cap \{\boldsymbol{E}^{n,j}\neq\boldsymbol{E}^n\}\neq\emptyset} \rho_i Q_i^n\|_{r'}}{\|Q\|_{r'}} \\
&\leqslant c\,\|\operatorname{div} \boldsymbol{E}^{n,j}\, \chi_{\Omega^n_{\{\boldsymbol{E}^{n,j}\neq\boldsymbol{E}^n\}}}\|_r \leqslant c\,\|\nabla \boldsymbol{E}^{n,j}\, \chi_{\Omega^n_{\{\boldsymbol{E}^{n,j}\neq\boldsymbol{E}^n\}}}\|_r,
\end{aligned}$$

where $\chi_{\Omega^n_{\{\boldsymbol{E}^{n,j}\neq\boldsymbol{E}^n\}}}$ is the characteristic function of the set

$$\Omega^n_{\{\boldsymbol{E}^{n,j}\neq\boldsymbol{E}^n\}} := \bigcup \Big\{ \Omega^n_E \mid E \in \mathcal{G}_n \text{ such that } E \subset \overline{\{\boldsymbol{E}^{n,j} \neq \boldsymbol{E}^n\}} \Big\}.$$

Note that in the penultimate step of the above estimate we have used norm equivalence on the reference space $\hat{\mathbb{P}}_{\mathbb{Q}}$ from (3.1b). In particular, we see by means of standard scaling arguments that for $Q = \sum_{i=1}^{N_n} \rho_i Q_i^n$ the norms

$$Q \mapsto \Big(\sum_{i=1}^{N_n} |\rho_i|^{r'} \|Q_i^n\|_{r'}^{r'} \Big)^{1/r'} \qquad \text{and} \qquad Q \mapsto \|Q\|_{r'}$$

are equivalent with constants depending on the shape-regularity of $\mathbb{G}$ and $\hat{\mathbb{P}}_{\mathbb{Q}}$. This directly implies the desired estimate.

Observe that $|\Omega^n_E| \leqslant c\,|E|$ for all $E \in \mathcal{G}_n$, $n \in \mathbb{N}$, with a shape-dependent constant $c > 0$; hence, $\Big|\Omega^n_{\{\boldsymbol{E}^{n,j}\neq\boldsymbol{E}^n\}}\Big| \leqslant c\,\big|\{\boldsymbol{E}^{n,j} \neq \boldsymbol{E}^n\}\big|$, and it follows from Theorem 17 and Corollary 18(e) that

$$\beta_r \|\boldsymbol{\Psi}^{n,j}\|_{1,r} \leqslant c\,\|\lambda_{n,j} \chi_{\Omega^n_{\{\boldsymbol{E}^{n,j}\neq\boldsymbol{E}^n\}}}\|_r \leqslant c\, 2^{-j/r} \|\nabla \boldsymbol{E}^n\|_r. \tag{4.23}$$

Moreover, we have from Corollaries 18 and 10, that

$$\boldsymbol{\Phi}^{n,j}, \boldsymbol{\Psi}^{n,j} \rightharpoonup 0 \qquad \text{weakly in } W^{1,s}_0(\Omega)^d \quad \text{for all } s \in [1,\infty), \tag{4.24a}$$

$$\boldsymbol{\Phi}^{n,j}, \boldsymbol{\Psi}^{n,j} \to 0 \qquad \text{strongly in } L^{s}(\Omega)^d \quad \text{for all } s \in [1,\infty), \tag{4.24b}$$

as $n \to \infty$.

Step 2: We claim that

$$\limsup_{n\to\infty} \int_{\{\boldsymbol{E}^n = \boldsymbol{E}^{n,j}\}} |a_n| \,\mathrm{d}x \leqslant c\, 2^{-j/r},$$

with a constant $c > 0$ independent of j. To see this we first observe that $|a_n| = a_n + 2a_n^-$ with the usual notation $a_n^-(x) = \max\{-a_n(x), 0\}$, $x \in \Omega$. Therefore, we have that

$$\limsup_{n\to\infty}\int_{\{\boldsymbol{E}^n=\boldsymbol{E}^{n,j}\}}|a_n|\,\mathrm{d}x \leqslant \limsup_{n\to\infty}\int_{\{\boldsymbol{E}^n=\boldsymbol{E}^{n,j}\}}a_n\,\mathrm{d}x + 2\limsup_{n\to\infty}\int_{\{\boldsymbol{E}^n=\boldsymbol{E}^{n,j}\}}a_n^-\,\mathrm{d}x. \tag{4.25}$$

We bound the two terms on the right-hand side separately. As a consequence of (4.21) and the fact that $b_n(x)\geqslant 0$ for a.e. $x\in\Omega$ (cf. (4.12)) it follows that

$$\int_{\{\boldsymbol{E}^n=\boldsymbol{E}^{n,j}\}}a_n^-\,\mathrm{d}x \leqslant \int_\Omega a_n^-\,\mathrm{d}x \leqslant \int_\Omega |a_n-b_n|\,\mathrm{d}x \to 0, \qquad \text{as } n\to\infty.$$

The bound on the first term on the right-hand side of (4.25) is more involved. In particular, recalling the definitions (4.20) and (4.22) we have that

$$\begin{aligned}
&\int_{\{\boldsymbol{E}^n=\boldsymbol{E}^{n,j}\}}a_n\,\mathrm{d}x\\
&\quad=\int_{\{\boldsymbol{E}^n=\boldsymbol{E}^{n,j}\}}\big(\boldsymbol{S}^n(\cdot,\boldsymbol{D}\boldsymbol{U}^n)-\boldsymbol{S}^*(\cdot,\boldsymbol{D}\boldsymbol{u})\big):(\boldsymbol{D}\Pi^n_{\mathrm{div}}\boldsymbol{u}-\boldsymbol{D}\boldsymbol{u})\,\mathrm{d}x\\
&\qquad+\int_\Omega \boldsymbol{S}^n(\cdot,\boldsymbol{D}\boldsymbol{U}^n):\boldsymbol{D}\boldsymbol{\Phi}^{n,j}\,\mathrm{d}x+\int_\Omega \boldsymbol{S}^n(\cdot,\boldsymbol{D}\boldsymbol{U}^n):\boldsymbol{D}\boldsymbol{\Psi}^{n,j}\,\mathrm{d}x\\
&\qquad-\int_\Omega \boldsymbol{S}^*(\cdot,\boldsymbol{D}\boldsymbol{u}):\boldsymbol{D}\boldsymbol{E}^{n,j}\,\mathrm{d}x\\
&\qquad+\int_{\{\boldsymbol{E}^n\neq\boldsymbol{E}^{n,j}\}}\big(\boldsymbol{S}^*(\cdot,\boldsymbol{D}\boldsymbol{u})-\boldsymbol{S}^n(\cdot,\boldsymbol{D}\boldsymbol{U}^n)\big):\boldsymbol{D}\boldsymbol{E}^{n,j}\,\mathrm{d}x\\
&\quad=\mathrm{I}^n+\mathrm{II}^{n,j}+\mathrm{III}^{n,j}+\mathrm{IV}^{n,j}+\mathrm{V}^{n,j}.
\end{aligned}$$

Thanks to (3.4) and (4.2) we have that, as $n\to\infty$,

$$\begin{aligned}
|\mathrm{I}^n| &\leqslant \int_{\{\boldsymbol{E}^n=\boldsymbol{E}^{n,j}\}}\big|\boldsymbol{S}^n(x,\boldsymbol{D}\boldsymbol{U}^n(x))-\boldsymbol{S}^*(x,\boldsymbol{D}\boldsymbol{u}(x))\big|\,|\boldsymbol{D}\Pi^n_{\mathrm{div}}\boldsymbol{u}(x)-\boldsymbol{D}\boldsymbol{u}(x)|\,\mathrm{d}x\\
&\leqslant \|\boldsymbol{S}^n(\cdot,\boldsymbol{D}\boldsymbol{U}^n(\cdot))-\boldsymbol{S}^*(\cdot,\boldsymbol{D}\boldsymbol{u}(\cdot)\|_{r'}\|\boldsymbol{D}\Pi^n_{\mathrm{div}}\boldsymbol{u}-\boldsymbol{D}\boldsymbol{u}\|_r\to 0.
\end{aligned}$$

In order to estimate $\mathrm{II}^{n,j}$ we recall that $\boldsymbol{\Phi}^{n,j}\in\mathbb{V}^n_{\mathrm{div}}$ is discretely divergence-free, and we can therefore use it as a test function in (3.17) to deduce that

$$\mathrm{II}^{n,j}=-\mathcal{B}[\boldsymbol{U}^n,\,\boldsymbol{U}^n,\,\boldsymbol{\Phi}^{n,j}]+\big\langle \boldsymbol{f},\,\boldsymbol{\Phi}^{n,j}\big\rangle_\Omega\to 0 \qquad \text{as } n\to\infty.$$

Indeed, the second term vanishes thanks to (4.24a). The first term vanishes by arguing as in (4.7) — observe that for (4.8) the weak convergence (4.24a) of $\boldsymbol{\Phi}^{n,j}$ is sufficient. The term $\mathrm{III}^{n,j}$ can be bounded by means of (4.23); in particular,

$$\limsup_{n\to\infty}\big|\mathrm{III}^{n,j}\big|\leqslant \limsup_{n\to\infty}\|\boldsymbol{S}(\cdot,\boldsymbol{D}\boldsymbol{U}^n)\|_{r'}\|\boldsymbol{D}\boldsymbol{\Psi}^{n,j}\|_r\leqslant c\,2^{-j/r},$$

where we have used (4.2). Corollary 18 implies that

$$\lim_{n\to\infty}\mathrm{IV}^{n,j}=0.$$

Finally, by (4.2) and Corollary 18, we have that

$$\begin{aligned}
\limsup_{n\to\infty}\big|\mathrm{V}^{n,j}\big| &\leqslant \limsup_{n\to\infty}\big(\|\boldsymbol{S}^*(\cdot,\boldsymbol{D}\boldsymbol{u})\|_{r'}+\|\boldsymbol{S}^n(\cdot,\boldsymbol{D}\boldsymbol{U}^n)\|_{r'}\big)\|\boldsymbol{D}\boldsymbol{E}^{n,j}\chi_{\{\boldsymbol{E}^n\neq\boldsymbol{E}^{n,j}\}}\|_r\\
&\leqslant c\,2^{-j/r}.
\end{aligned}$$

In view of (4.25) this completes Step 2.

Step 3: We prove, for any $\vartheta \in (0,1)$, that

$$\lim_{n\to\infty} \int_\Omega |a_n|^\vartheta \,\mathrm{d}x = 0.$$

Using Hölder's inequality, we easily obtain that

$$\begin{aligned}\int_\Omega |a_n|^\vartheta \,\mathrm{d}x &= \int_{\{\boldsymbol{E}^n=\boldsymbol{E}^{n,j}\}} |a_n|^\vartheta \,\mathrm{d}x + \int_{\{\boldsymbol{E}^n\neq\boldsymbol{E}^{n,j}\}} |a_n|^\vartheta \,\mathrm{d}x \\ &\leqslant |\Omega|^{1-\vartheta} \left(\int_{\{\boldsymbol{E}^n=\boldsymbol{E}^{n,j}\}} |a_n| \,\mathrm{d}x\right)^\vartheta + \left(\int_\Omega |a_n| \,\mathrm{d}x\right)^\vartheta \left|\{\boldsymbol{E}^n \neq \boldsymbol{E}^{n,j}\}\right|^{1-\vartheta}.\end{aligned}$$

Thanks to (4.2), we have that $(\int_\Omega |a_n| \,\mathrm{d}x)^\vartheta$ is bounded uniformly in n and by Corollary 18 we have that

$$\left|\{\boldsymbol{E}^n \neq \boldsymbol{E}^{n,j}\}\right| \leqslant c\, \frac{\|\boldsymbol{E}^n\|_{1,r}^r}{\lambda_{n,j}^r} \leqslant \frac{c}{2^{2^j r}},$$

where we have used that $\{\boldsymbol{E}^n\}_{n\in\mathbb{N}}$ is bounded in $W_0^{1,r}(\Omega)^d$ according to (4.2) and Assumption 6. Consequently, from Step 2 we deduce that

$$\limsup_{n\to\infty} \int_\Omega |a_n|^\vartheta \,\mathrm{d}x \leqslant c\,|\Omega|^{1-\vartheta}\, 2^{-j\vartheta/r} + \frac{c}{2^{2^j r(1-\vartheta)}}.$$

The left-hand side is independent of j and we can thus pass to the limit $j \to \infty$. This proves the assertion and actually implies that $a_n \to 0$ in measure as $n \to \infty$. According to (4.21) we have that $b_n \to 0$ in measure and thus we have completed the proof of Lemma 21.

4.4. Proof of Lemma 22. The proof of this Lemma is given in [BGMŚG09]. In order to keep the paper self-contained, we shall reproduce it here.

The assertion is an immediate consequence of the result on Young measures from [Gwi05] stated in Theorem 3. It therefore suffices to check the assumptions therein. The first assumption has already been verified in (4.16).

Step 1: We prove that the sequence $\{\nu^n\}_{n\in\mathbb{N}}$ satisfies the tightness condition. From the definition of ν_x^n (cf. (4.14)) it follows that

$$\gamma^n(x) := \max_{\boldsymbol{\zeta}\in\operatorname{supp}\nu_x^n} |\boldsymbol{\zeta}| = \max_{\boldsymbol{\zeta}\in\operatorname{supp}\mu_x^n} |\boldsymbol{G}_x(\boldsymbol{\zeta})| \leqslant \max_{\boldsymbol{\zeta}\in\operatorname{supp}\mu_x^n} \big(|\boldsymbol{\zeta}| + \left|\boldsymbol{S}^*(x,\boldsymbol{\zeta})\right|\big).$$

We deduce from the inclusion $\operatorname{supp}\mu_x^n \subset B_{1/n}(\boldsymbol{D}\boldsymbol{V}^n(x))$ that $\|\gamma^n\|_s \leqslant c$ for some constant $c > 0$ and $s = \max\{r, r'\} > 1$. Since Ω is bounded, the sequence is uniformly bounded in $L^1(\Omega)$, and for $M > 0$ we have

$$\left|\{x \in \Omega\colon \operatorname{supp}\nu_x^n \backslash B_M(0)\}\right| = \left|\{x \in \Omega\colon \gamma^n(x) > M\}\right| \leqslant \int_\Omega \frac{\gamma^n(x)}{M} \,\mathrm{d}x \leqslant \frac{c}{M}.$$

This yields the tightness of $\{\nu^n\}_{n\in\mathbb{N}}$ and it follows from part (i) of Theorem 3 that ν_x is a probability measure, i. e., $\|\nu_x\|_{\mathfrak{M}(\mathbb{R}^{d\times d}_{\mathrm{sym}})} = 1$ for a.e. $x \in \Omega$.

Step 2: We turn to proving (4.17). Recalling (4.15c), the assertion follows if there exists a nonincreasing sequence of measurable subsets $\{E_i\}_{i\in\mathbb{N}}$ with $|E_i| \to 0$ as $i \to \infty$, such that the function

$$h(x,\boldsymbol{\zeta}) := \big(\boldsymbol{s}(x,\boldsymbol{\zeta}) - \boldsymbol{S}^*(x, \boldsymbol{D}\boldsymbol{u}(x))\big) : (\boldsymbol{d}(x,\boldsymbol{\zeta}) - \boldsymbol{D}\boldsymbol{u}(x))$$

satisfies (2.6) on $A = \Omega \backslash E_i$ for each $i \in \mathbb{N}$. This can be seen as follows. From (4.2) it follows that $\|\boldsymbol{D}\boldsymbol{U}^n\|_r + \|\boldsymbol{D}\boldsymbol{u}\|_r$ is bounded uniformly in $n \in \mathbb{N}$. Consequently, the sequence $\{c_n\}_{n\in\mathbb{N}}$, defined by

$$c_n(x) := c\Big(|\boldsymbol{D}\boldsymbol{U}^n(x)|^{r-1} + |\boldsymbol{D}\boldsymbol{u}(x)|^{r-1} + \frac{1}{n^{r-1}} + k^{\frac{1}{r'}}(x)\Big) \times \Big(|\boldsymbol{D}\boldsymbol{U}^n(x)| + |\boldsymbol{D}\boldsymbol{u}(x)| + \frac{1}{n}\Big), \quad (4.26)$$

is bounded in $L^1(\Omega)$, where $k \in L^{r'}(\Omega)$ from (2.4) and $c > 0$ is a constant to be chosen later. Hence, Chacon's biting lemma (Lemma 4) implies that there exists a nonincreasing sequence of measurable subsets $\{E_i\}_{i\in\mathbb{N}}$ with $|E_i| \to 0$ as $i \to \infty$ such that $\{c_n\}_{n\in\mathbb{N}}$ is weakly precompact in $L^1(\Omega\backslash E_i)$ for each $i \in \mathbb{N}$. Fix $i \in \mathbb{N}$ and set $\omega := \Omega\backslash E_i$. Thanks to the de la Valleé-Poussin theorem (see, [Mey66]), there exists a nonnegative increasing convex function $\phi : \mathbb{R}_+ \to \mathbb{R}_+$ such that

$$\frac{\phi(t)}{t} \to \infty \quad \text{as} \quad t \to \infty \qquad \text{and} \qquad \sup_{n\in\mathbb{N}} \int_\omega \phi(|c_n|)\,\mathrm{d}x < \infty. \quad (4.27)$$

Let us also define

$$H(x, \boldsymbol{\zeta}) := \big(\boldsymbol{S}^*(x, \boldsymbol{\zeta}) - \boldsymbol{S}^*(x, \boldsymbol{D}\boldsymbol{u}(x))\big) : (\boldsymbol{\zeta} - \boldsymbol{D}\boldsymbol{u}(x)).$$

By a simple substitution in the spirit of (4.15) it follows that

$$\begin{aligned}
&\sup_{n\in\mathbb{N}} \int_\omega \int_{\{\boldsymbol{\zeta}\in\mathbb{R}^{d\times d}_{\mathrm{sym}}\,:\,|h(x,\boldsymbol{\zeta})|>R\}} h(x,\boldsymbol{\zeta})\,\mathrm{d}\nu^n_x(\boldsymbol{\zeta})\,\mathrm{d}x \\
&\quad = \sup_{n\in\mathbb{N}} \int_\omega \int_{\{\boldsymbol{\xi}\in\mathbb{R}^{d\times d}_{\mathrm{sym}}\,:\,|H(x,\boldsymbol{\xi})|>R\}} H(x,\boldsymbol{\xi})\,\mathrm{d}\mu^n_x(\boldsymbol{\xi})\,\mathrm{d}x \\
&\quad \leqslant \sup_{t\geqslant R} \frac{t}{\phi(t)} \underbrace{\sup_{n\in\mathbb{N}} \int_\omega \int_{\{\boldsymbol{\xi}\in\mathbb{R}^{d\times d}_{\mathrm{sym}}\,:\,|H(x,\boldsymbol{\xi})|>R\}} \phi\big(H(x,\boldsymbol{\xi})\big)\,\mathrm{d}\mu^n_x(\boldsymbol{\xi})\,\mathrm{d}x}_{=:J_R}.
\end{aligned}$$

Thanks to the properties (4.27) of ϕ the assertion follows once J_R has been shown to remain bounded. To this end, we observe that

$$J_R \leqslant \sup_{n\in\mathbb{N}} \int_\omega \sup_{\boldsymbol{\xi}\in B_{\frac{1}{n}}(\boldsymbol{D}\boldsymbol{u}(x))} \phi\big(H(x,\boldsymbol{\xi})\big)\,\mathrm{d}x \leqslant \sup_{n\in\mathbb{N}} \int_\omega \phi(|c_n|)\,\mathrm{d}x,$$

where we have used that we can choose the constant in (4.26) so that

$$H(x,\boldsymbol{\xi}) \leqslant c\left(|\boldsymbol{\xi}|^{r-1} + |\boldsymbol{D}\boldsymbol{u}(x)|^{r-1} + k(x)\right)(|\boldsymbol{\xi}| + |\boldsymbol{D}\boldsymbol{u}(x)|)\,.$$

The assertion then follows from (4.27).

Finally the identities for $\boldsymbol{D}\boldsymbol{u}$ and $\boldsymbol{S}$ follow similarly from the representations (4.15a) and (4.15b) and the uniqueness of the weak limits (4.3) and (4.4).

Thus we have completed the proof of Lemma 22.

5. Conclusions. We have established the convergence of finite element approximations of implicitly constituted power-law-like models for viscous incompressible fluids. A key new technical tool in our analysis was a finite element counterpart of the Acerbi–Fusco Lipschitz truncation of Sobolev functions, which was used in combination with a variety of weak compactness techniques, including Chacon's biting lemma and Young measures. An interesting direction for future research is the extension of the results obtained herein to unsteady implicitly constituted models of incompressible fluids.

Acknowledgments. The research of Christian Kreuzer was supported by the DFG research grant KR 3984/1-1. Lars Diening and Endre Süli were supported by the EPSRC Science and Innovation award to the Oxford Centre for Nonlinear PDE (EP/E035027/1).

REFERENCES

[AF88] E. Acerbi and N. Fusco, *An approximation lemma for $W^{1,p}$ functions*, Material instabilities in continuum mechanics (Edinburgh, 1985–1986), Oxford Sci. Publ., Oxford Univ. Press, New York, 1988, pp. 1–5.

[AB2006] Ch. D. Aliprantis and K. C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide, 3rd ed.*, Springer, Berlin, 2006.

[AQ92] D. N. Arnold and J. Qin, *Quadratic velocity/linear pressure Stokes elements*, Advances in Computer Methods for Partial Differential Equations–VII (R. Vichnevetsky, D. Knight, and G. Richter, eds.), IMACS, 1992, pp. 28–34.

[Bal89] J. M. Ball, *A version of the fundamental theorem for Young measures*, PDEs and continuum models of phase transitions (Nice, 1988), Lecture Notes in Phys., vol. 344, Springer, Berlin, 1989, pp. 207–215.

[BBDR10] L. Belenki, L. Berselli, L. Diening, and M. Růžička, *On the finite element approximation of p-Stokes systems*, SIAM J. Numer. Anal. (2012), **50**, pp. 373–397.

[BDF12] D. Breit, L. Diening, M. Fuchs. *Solenoidal Lipschitz truncation and applications in fluid mechanics*, J. Differential Equations **253**(6) (2012), pp. 1910–1942.

[BDS12] D. Breit, L. Diening, S. Schwarzacher. *Solenoidal Lipschitz truncation for parabolic PDE's*, M3AS (to appear).

[BF91] F. Brezzi and M. Fortin, *Mixed and hybrid finite element methods*, Springer Series in Computational Mathematics 15, 1991.

[BGMŚG09] M. Bulíček, P. Gwiazda, J. Málek, and A. Świerczewska-Gwiazda, *On steady flows of incompressible fluids with implicit power-law-like rheology*, Adv. Calc. Var. **2**(2) (2009), pp. 109–136.

[BGMŚG11] M. Bulíček, P. Gwiazda, J. Málek, and A. Świerczewska-Gwiazda. On unsteady flows of implicitly constituted incompressible fluids. SIAM J. Math. Anal. **44**(4) (2012), pp. 2756–2801.

[BMS12] M. Bulíček, J. Málek, and E. Süli. Existence of global weak solutions to implicitly constituted kinetic models of incompressible homogeneous dilute polymers. Commun. Part. Diff. Eq. (to appear), DOI: 10.1080/03605302.2012.742104.

[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) (1979), pp. 1037–1040.

[DHHR11] L. Diening, P. Harjulehto, P. Hästö, and M. Růžička, *Lebesgue and Sobolev spaces with variable exponents*, Lecture Notes in Mathematics, vol. 2017, Springer, Heidelberg, 2011.

[DMS08] L. Diening, J. Málek, and M. Steinhauer, *On Lipschitz truncations of Sobolev functions (with variable exponent) and their selected applications*, ESAIM Control Optim. Calc. Var. **14**(2) (2008), pp. 211–232.

[DR07] L. Diening and M. Růžička, *Interpolation operators in Orlicz–Sobolev spaces*, Numer. Math. **107**(1) (2007), pp. 107–129.

[DRS10] L. Diening, M. Růžička, and K. Schumacher, *A decomposition technique for John domains*, Ann. Acad. Sci. Fenn. Math. **35**(1) (2010), pp. 87–114.

[EG92] L. C. Evans and R. F. Gariepy, *Measure theory and fine properties of functions*, Studies in Advanced Mathematics, CRC Press, Boca Raton, FL, 1992.

[Fed69] H. Federer, *Geometric measure theory*, Die Grundlehren der mathematischen Wissenschaften, Band 153, Springer-Verlag New York Inc., New York, 1969.

[FMT04] G. Francfort, F. Murat, and L. Tartar, *Monotone operators in divergence form with x-dependent multivalued graphs*, Boll. Unione Mat. Ital. Sez. B Artic. Ric. Mat. (8) **7**(1) (2004), pp. 23–59.

[G04] L. Grafakos, *Classical and Modern Fourier Analysis*, Pearson Education, Inc., Upper Saddle River, NJ, (2008).

[GN11] J. Guzmán and M. Neilan, *Conforming and divergence-free Stokes elements on general triangular meshes*, Math. Comput. (to appear).

[GN12] J. Guzmán and M. Neilan, private communication, January 2012.

[GMŚ07] P. Gwiazda, J. Málek, and A. Świerczewska, *On flows of an incompressible fluid with a discontinuous power-law-like rheology*, Comput. Math. Appl. **53**(3-4) (2007), pp. 531–546.

[Gwi05] P. Gwiazda, *On measure-valued solutions to a two-dimensional gravity-driven avalanche flow model*, Math. Methods Appl. Sci. **28**(18) (2005), pp. 2201–2223.

[GZG07] P. Gwiazda and A. Zatorska-Goldstein, *On elliptic and parabolic systems with x-dependent multivalued graphs*, Math. Methods Appl. Sci. **30**(2) (2007), pp. 213–236.

[Lad69] O. A. Ladyzhenskaya, *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, 1969.

[Lio69] J.-L. Lions, *Quelques méthodes de résolution des problèmes aux limites non linéaires*, Dunod, 1969.

[Mál07] J. Málek, *Mathematical properties of the flows of incompressible fluids with pressure and shear rate dependent viscosities*, D.Sc. thesis, Academy of Sciences of the Czech Republic, Prague, 2007.

[Mál08] ———, *Mathematical properties of flows of incompressible power-law-like fluids that are described by implicit constitutive relations*, ETNA, Electron. Trans. Numer. Anal. **31** (2008), pp. 110–125.

[MZ97] Malý, J. and Ziemer, W. P., *Fine regularity of solutions of elliptic partial differential equations*, American Mathematical Society, 1997.

[Mey66] P.-A. Meyer, *Probability and potentials*, Blaisdell Publishing Co. Ginn and Co., Waltham, Mass.-Toronto, Ont.-London, 1966.

[QZ07] J. Qin and S. Zhang, *Stability and approximability of the $\mathcal{P}_1$-$\mathcal{P}_0$ element for Stokes equations*, Internat. J. Numer. Methods Fluids **54**(5) (2007), pp. 497–515.

[Raj03] K. R. Rajagopal, *On implicit constitutive theories*, Appl. Math., Prague, **48**(4) (2003), pp. 279–319.

[Raj06] ———, *On implicit constitutive theories for fluids*, J. Fluid Mech. **550** (2006), pp. 243–249.

[RS08] K. R. Rajagopal and A. R. Srinivasa. *On the thermodynamics of fluids defined by implicit constitutive relations*, Z. Angew. Math. Phys. **59**(4) (2008), pp. 715–729.

[Růž00] M. Růžička, *Electrorheological Fluids: Modeling and Mathematical Theory,* Springer, Belin, 2000.

[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) (1985), pp. 111–143.

[SZ90] L. R. Scott and S. Zhang, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Math. Comput. **54**(190) (1990), pp. 483–493.

[Tem84] R. Temam, *Navier–Stokes Equations. Theory and Numerical Analysis. 3rd (rev.) ed.*, Studies in Mathematics and its Applications, Vol. 2. North-Holland, Amsterdam-New York-Oxford, 1984.

[Zha08] S. Zhang, *On the P1 Powell–Sabin divergence-free finite element for the Stokes equations*, J. Comput. Math. **26**(3) (2008), pp. 456–470.

Appendix: Auxiliary comments on Lipschitz truncation. Although similar techniques were used in [BDF12] to prove the properties of the Lipschitz truncation, we decided to present a complete proof of Theorem 14 for the following two reasons:

- In contrast with the Lipschitz truncation in [BDF12], the Lipschitz truncation in (3.25) preserves boundary values. This requires changes to the proof that are not always obvious.
- The concept of Lipschitz truncation seems to be new to the numerical analysis community. For this reason we have aimed to keep the presentation as self-contained as possible.

Recall the notational conventions introduced in Section 3.5 prior to Theorem 14, and the definition (3.25) of the Lipschitz truncation. We start with some basic estimates.

Lemma 24. *Let $\lambda > 0$ and $\boldsymbol{v} \in W_0^{1,1}(\Omega)^d$ and let $\{\boldsymbol{v}_j\}_{j\in\mathbb{N}} \subset \mathbb{R}^d$ be defined as in* (3.25b). *We then have, for all $j \in \mathbb{N}$, that*

(a) $⨍_{\mathcal{Q}_j^{**}} \left|\frac{\boldsymbol{v}-\boldsymbol{v}_j}{\ell_j}\right| \mathrm{d}x \leqslant c ⨍_{\mathcal{Q}_j^{**}} |\nabla \boldsymbol{v}| \,\mathrm{d}x \leqslant c\, M(\nabla \boldsymbol{v})(y)$ *for all $y \in \mathcal{Q}_j^{**}$;*

(b) $⨍_{\mathcal{Q}_j^{**}} |\nabla \boldsymbol{v}| \,\mathrm{d}x \leqslant c\,\lambda$*;*

(c) *for $k \in \mathbb{N}$ with $\mathcal{Q}_j^* \cap \mathcal{Q}_k^* \neq \emptyset$, we have*

$$|\boldsymbol{v}_j - \boldsymbol{v}_k| \leqslant c \fint_{\mathcal{Q}_j^{**}} |\boldsymbol{v} - \boldsymbol{v}_j| \,\mathrm{d}x + c \fint_{\mathcal{Q}_k^{**}} |\boldsymbol{v} - \boldsymbol{v}_k| \,\mathrm{d}x;$$

(d) *for $k \in \mathbb{N}$ with $\mathcal{Q}_j^* \cap \mathcal{Q}_k^* \neq \emptyset$ we have $|\boldsymbol{v}_j - \boldsymbol{v}_k| \leqslant c\,\ell_j\,\lambda$.*

Proof. We extend $\boldsymbol{v}$ by zero outside Ω.

(a) This statement is a consequence of Poincaré's inequality and the Friedrichs inequality. Indeed, recalling (3.25b), for $\mathcal{Q}_j^* \subset \Omega$ we have by Poincaré's inequality that

$$\fint_{\mathcal{Q}_j^{**}} \left| \frac{\boldsymbol{v} - \boldsymbol{v}_j}{\ell_j} \right| \mathrm{d}x \leqslant c \fint_{\mathcal{Q}_j^{**}} |\nabla \boldsymbol{v}| \,\mathrm{d}x \leqslant c \fint_{B_{\operatorname{diam}(\mathcal{Q}_j^{**})}(y)} |\nabla \boldsymbol{v}| \,\mathrm{d}x \leqslant c\, M(\nabla \boldsymbol{v})(y)$$

for all $y \in \mathcal{Q}_j^{**}$; the constant c depends only on d.

In the case $\mathcal{Q}_j^* \not\subset \Omega$, it follows from the fact that Ω is a Lipschitz domain and $\mathcal{Q}_j^{**} = \sqrt{\frac{9}{8}}\mathcal{Q}_j^*$, that $\left|Q_j^{**} \backslash \Omega\right| \geqslant c\left|\mathcal{Q}_j^*\right|$, with a constant $c > 0$ depending on Ω. Hence $\boldsymbol{v}$ is zero on a portion of $\mathcal{Q}_j^{**}$ whose measure is bounded below by a positive constant, which depends on the Lipschitz constant of $\partial\Omega$. Consequently, we can apply Friedrichs' inequality (cf. [MZ97, Lemma 1.65]) to deduce that

$$\fint_{\mathcal{Q}_j^{**}} \left| \frac{\boldsymbol{v}}{\ell_j} \right| \mathrm{d}x \leqslant c \fint_{\mathcal{Q}_j^{**}} |\nabla \boldsymbol{v}| \,\mathrm{d}x \leqslant c\, M(\nabla \boldsymbol{v})(y) \quad \forall y \in \mathcal{Q}_j^{**}.$$

(b) It follows from (W2) that $(\theta_d \mathcal{Q}_j) \cap (\mathbb{R}^d \backslash \mathcal{U}_\lambda(\boldsymbol{v})) = (\theta_d \mathcal{Q}_j) \cap \{M(\nabla \boldsymbol{v}) \leqslant \lambda\} \neq \emptyset$; compare with (3.22). For $z \in (\theta_d \mathcal{Q}_j) \cap \{M(\nabla \boldsymbol{v}) \leqslant \lambda\}$ let $R_j := \theta_d \sqrt{d}\frac{9}{8}\ell_j = \theta_d \operatorname{diam}(\mathcal{Q}_j^{**})$; then, $\theta_d \mathcal{Q}_j^{**} \subset B_{R_j}(z)$. Consequently,

$$\fint_{\mathcal{Q}_j^{**}} |\nabla \boldsymbol{v}| \,\mathrm{d}x \leqslant c \fint_{\theta_d \mathcal{Q}_j^{**}} |\nabla \boldsymbol{v}| \,\mathrm{d}x \leqslant c \fint_{B_{R_j}(z)} |\nabla \boldsymbol{v}| \,\mathrm{d}x \leqslant c\, M(\nabla \boldsymbol{v})(z) \leqslant c\,\lambda.$$

(c) Observe that $\mathcal{Q}_j^* \cap \mathcal{Q}_k^* \neq \emptyset$ is equivalent to $\mathcal{Q}_j \cap \mathcal{Q}_k \neq \emptyset$ and hence we obtain from (W3) and $\mathcal{Q}_i^{**} = \sqrt{\frac{9}{8}}\mathcal{Q}_i^*$, $i \in \mathbb{N}$, that

$$\left|\mathcal{Q}_j^{**} \cap \mathcal{Q}_k^{**}\right| \geqslant (4\sqrt{2})^{-d} \max\{\left|\mathcal{Q}_j^*\right|, |\mathcal{Q}_k^*|\}.$$

Therefore, there exists a constant $c > 0$, depending on d, such that

$$\begin{aligned} |\boldsymbol{v}_j - \boldsymbol{v}_k| &\leqslant \fint_{\mathcal{Q}_j^{**} \cap \mathcal{Q}_j^{**}} |\boldsymbol{v} - \boldsymbol{v}_j| \,\mathrm{d}x + \fint_{\mathcal{Q}_j^{**} \cap \mathcal{Q}_j^{**}} |\boldsymbol{v} - \boldsymbol{v}_k| \,\mathrm{d}x \\ &\leqslant c \fint_{\mathcal{Q}_j^{**}} |\boldsymbol{v}_j - \boldsymbol{v}| \,\mathrm{d}x + c \fint_{\mathcal{Q}_k^{**}} |\boldsymbol{v} - \boldsymbol{v}_k| \,\mathrm{d}x. \end{aligned}$$

(d) The claim is a combination of (c), (a), (b) and (W3). ∎

The next result proves that the Lipschitz truncation is a proper Sobolev function.

LEMMA 25. *Let $\lambda > 0$, $\boldsymbol{v} \in W_0^{1,1}(\Omega)^d$ and let $\boldsymbol{v}_\lambda$ be defined as in* (3.25)*. Then, $\boldsymbol{v}_\lambda - \boldsymbol{v} = \sum_{j \in \mathbb{N}} \psi_j(\boldsymbol{v}_j - \boldsymbol{v}) \in W_0^{1,1}(\mathcal{U}_\lambda(\boldsymbol{v}) \cap \Omega)^d$.*

Proof. It follows from (3.25) and properties of the partition of unity $\{\psi_j\}_{j\in\mathbb{N}}$ that $\boldsymbol{v}_\lambda - \boldsymbol{v} = \sum_{j\in\mathbb{N}} \psi_j(\boldsymbol{v}_j - \boldsymbol{v})$ pointwise on $\mathbb{R}^d$ and $\boldsymbol{v}_\lambda - \boldsymbol{v} = 0$ in the complement of $\mathcal{U}_\lambda(\boldsymbol{v})$. Moreover, we have that $\psi_j(\boldsymbol{v}_j - \boldsymbol{v}) \in W_0^{1,1}(\mathcal{U}_\lambda(\boldsymbol{v}) \cap \Omega)^d$. Indeed, for $\mathcal{Q}_j^* \subset \Omega$ this follows from the fact that $\psi \in C_0^\infty(\mathcal{Q}_j^*)$. If on the other hand $\mathcal{Q}_j^* \not\subset \Omega$, then this follows from $\boldsymbol{v}_j = 0$ and $\boldsymbol{v} \in W_0^{1,1}(\mathcal{U}_\lambda(\boldsymbol{v}) \cap \Omega)$. We need to show that the sum converges in $W_0^{1,1}(\mathcal{U}_\lambda(\boldsymbol{v}) \cap \Omega)^d$. Since Ω is bounded, it suffices to prove that the sum of the gradients converges absolutely in $L^1(\Omega)^{d\times d}$. We have, pointwise, the equality

$$\sum_{j\in\mathbb{N}} \nabla\Big(\psi_j(\boldsymbol{v}_j - \boldsymbol{v})\Big) = \sum_{j\in\mathbb{N}} \big((\nabla\psi_j)(\boldsymbol{v}_j - \boldsymbol{v}) + \psi_j(\nabla\boldsymbol{v}_j - \nabla\boldsymbol{v})\big),$$

where we have used that both sums are just finite sums, since the family $\mathcal{Q}_j^*$ is locally finite. Every summand in the last sum belongs to $L^1(\mathcal{U}_\lambda(\boldsymbol{v}) \cap \Omega)^{d\times d}$. For a finite subset $I \subset \mathbb{N}$, we have, thanks to Lemma 24 and the locally finite overlaps of the $\mathcal{Q}_j^*$, that

$$\begin{aligned}
\Sigma_I &:= \int_{\mathcal{U}_\lambda(\boldsymbol{v})} \sum_{j\in\mathbb{N}\setminus I} |(\nabla\psi_j)(\boldsymbol{v}_j - \boldsymbol{v}) + \psi_j(\nabla\boldsymbol{v}_j - \nabla\boldsymbol{v})| \,\mathrm{d}x \\
&\leqslant c \sum_{j\in\mathbb{N}\setminus I} \int_{\mathcal{Q}_j^*} \frac{|\boldsymbol{v}_j - \boldsymbol{v}|}{\ell_j} \,\mathrm{d}x + \sum_{j\in\mathbb{N}\setminus I} \int_{\mathcal{Q}_j^*} |\nabla\boldsymbol{v}| \,\mathrm{d}x \\
&\leqslant c \sum_{j\in\mathbb{N}\setminus I} \int_{\mathcal{Q}_j^{**}} |\nabla\boldsymbol{v}| \,\mathrm{d}x \leqslant c \sum_{j\in\mathbb{N}\setminus I} \lambda\,|\mathcal{Q}_j| \leqslant c \int_{\mathcal{U}_\lambda(\boldsymbol{v})} \chi_{\cup_{j\in\mathbb{N}\setminus I}\mathcal{Q}_j^*} \lambda \,\mathrm{d}x.
\end{aligned}$$

Note that $\bigcup_{j\in\mathbb{N}\setminus I} \mathcal{Q}_j^* \subset \mathcal{U}_\lambda(\boldsymbol{v})$ and $\lambda\,|\mathcal{U}_\lambda(\boldsymbol{v})| = \lambda\,|\{M(\nabla\boldsymbol{v}) > \lambda\}| \leqslant c\,\|\nabla\boldsymbol{v}\|_{L^1(\Omega)}$ by the weak type estimate (3.21) and $\boldsymbol{v} \in W_0^{1,1}(\Omega)^d$. Thus, $\chi_{\cup_{j\in\mathbb{N}\setminus I}\mathcal{Q}_j^*}\lambda \leqslant \chi_{\mathcal{U}_\lambda(\boldsymbol{v})}\lambda \in L^1(\mathbb{R}^d)$. Therefore, it follows by $\chi_{\bigcup_{j\in\mathbb{N}\setminus I}\mathcal{Q}_j^*} \to 0$ and the Lebesgue's dominated convergence theorem that $\Sigma_I \to 0$ as $I \to \mathbb{N}$. Hence the sum $\sum_{j\in\mathbb{N}} \nabla\Big(\psi_j(\boldsymbol{v}_j - \boldsymbol{v})\Big)$ converges absolutely in $L^1(\Omega)^{d\times d}$, and the claim follows. □

Proof. [Proof of Theorem 14] We shall consider parts (a)–(d) in the statement of the theorem separately.

(a) The claim directly follows from $\boldsymbol{v}_\lambda - \boldsymbol{v} \in W_0^{1,1}(\mathcal{U}_\lambda \cap \Omega)^d$ (see Lemma 25).

(b) We begin by noting that

$$\chi_{\mathcal{U}_\lambda(\boldsymbol{v})}\,|\boldsymbol{v}_\lambda| \leqslant \sum_{j\in\mathbb{N}} \chi_{\mathcal{Q}_j^*}\,|\boldsymbol{v}_j| \leqslant \sum_{j\in\mathbb{N}} \chi_{\mathcal{Q}_j^*} \fint_{\mathcal{Q}_j^{**}} |\boldsymbol{v}| \,\mathrm{d}x.$$

By Jensen's inequality and the local finiteness of the $\mathcal{Q}_j^{**}$ we then deduce that

$$\int_\Omega \chi_{\mathcal{U}_\lambda(\boldsymbol{v})}\,|\boldsymbol{v}_\lambda|^s \,\mathrm{d}x \leqslant \sum_{j\in\mathbb{N}} \big|\mathcal{Q}_j^*\big| \bigg(\fint_{\mathcal{Q}_j^{**}} |\boldsymbol{v}| \,\mathrm{d}x\bigg)^s \leqslant \sum_{j\in\mathbb{N}} \int_{\mathcal{Q}_j^{**}} |\boldsymbol{v}|^s \,\mathrm{d}x \leqslant c \int_\Omega |\boldsymbol{v}|^s \,\mathrm{d}x,$$

for $s \in [1,\infty)$, which then proves (b) for $s \in [1,\infty)$ using also that $\boldsymbol{v}_\lambda = \boldsymbol{v}$ outside of $\mathcal{U}_\lambda(\boldsymbol{v})$. The case $s = \infty$ follows by obvious modifications of the argument.

(c) We define $I_j := \{k \in \mathbb{N} : \mathcal{Q}_j^* \cap \mathcal{Q}_k^* \neq \emptyset\}$. Then, on every $\mathcal{Q}_j^*$ we have that

$$\nabla\boldsymbol{v}_\lambda = \nabla\Big(\sum_{k\in\mathbb{N}} \psi_k\boldsymbol{v}_k\Big) = \nabla\Big(\sum_{k\in I_j} \psi_k(\boldsymbol{v}_k - \boldsymbol{v}_j)\Big) = \sum_{k\in I_j} (\nabla\psi_k)(\boldsymbol{v}_k - \boldsymbol{v}_j),$$

where we have used that $\sum_{k\in I_j} \psi_k = 1$ on $\mathcal{Q}_j^*$. By Lemma 24 we thus obtain

$$\chi_{\mathcal{U}_\lambda(\boldsymbol{v})}\,|\nabla\boldsymbol{v}_\lambda| \leqslant c \sum_{j\in\mathbb{N}} \chi_{\mathcal{Q}_j^*} \sum_{k\in I_j} \fint_{\mathcal{Q}_k^*} \frac{|\boldsymbol{v} - \boldsymbol{v}_k|}{\ell_k} \,\mathrm{d}x \leqslant c \sum_{j\in\mathbb{N}} \chi_{\mathcal{Q}_j^*} \sum_{k\in I_j} \fint_{\mathcal{Q}_k^{**}} |\nabla\boldsymbol{v}| \,\mathrm{d}x.$$

The inequality in part (c) now follows by arguing as in part (b).

(d) It follows from the final chain of inequalities in the proof of part (c) above, Lemma 24 (b) and the local finiteness of the $\mathcal{Q}_j^{**}$ that

$$\chi_{\mathcal{U}_\lambda(\boldsymbol{v})} |\nabla \boldsymbol{v}_\lambda| \leqslant c\,\lambda.$$

Since $\boldsymbol{v}_\lambda = \boldsymbol{v}$ on $\mathcal{H}_\lambda(\boldsymbol{v})$, we get the first part of the claim

$$|\nabla \boldsymbol{v}_\lambda| \leqslant c\,\lambda \chi_{\mathcal{U}_\lambda(\boldsymbol{v}) \cap \Omega} + |\nabla \boldsymbol{v}|\, \chi_{\mathcal{H}_\lambda(\boldsymbol{v})}.$$

Recall that $\mathcal{H}_\lambda(\boldsymbol{v}) = (\mathbb{R}^d \backslash \Omega) \cup \{M(\nabla \boldsymbol{v}) \leqslant \lambda\}$. Now, $\boldsymbol{v}_\lambda = 0$ on $\mathbb{R}^d \backslash \Omega$ and $|\nabla \boldsymbol{v}| \leqslant M(\nabla \boldsymbol{v})$ proves that $|\nabla \boldsymbol{v}|\, \chi_{\mathcal{H}_\lambda(\boldsymbol{v})} \leqslant \lambda$. This proves the second part of the claim. □

The following theorem is the analogue of Corollary 18 for Sobolev functions. Similar results can be found in [DMS08] and [BDF12].

Corollary 26. *Let $1 < s < \infty$ and let $\{\boldsymbol{e}^n\}_{n \in \mathbb{N}} \subset W_0^{1,s}(\Omega)^d$ be a sequence, which converges to zero weakly in $W_0^{1,s}(\Omega)^d$, as $n \to \infty$.*

Then, there exists a sequence $\{\lambda_{n,j}\}_{n,j \in \mathbb{N}} \subset \mathbb{R}$ with $2^{2^j} \leqslant \lambda_{n,j} \leqslant 2^{2^{j+1}-1}$ such that the Lipschitz truncations $\boldsymbol{e}^{n,j} := \boldsymbol{e}^n_{\lambda_{n,j}}$, $n, j \in \mathbb{N}$, have the following properties:

(a) $\boldsymbol{e}^{n,j} \in W_0^{1,\infty}(\Omega)^d$ *and* $\boldsymbol{e}^{n,j} = \boldsymbol{e}^n$ *on* $\mathcal{H}_{\lambda_{n,j}}$;

(b) $\|\nabla \boldsymbol{e}^{n,j}\|_\infty \leqslant c\lambda_{n,j}$;

(c) $\boldsymbol{e}^{n,j} \to 0$ *in* $L^\infty(\Omega)^d$ *as* $n \to \infty$;

(d) $\nabla \boldsymbol{e}^{n,j} \rightharpoonup^* 0$ *in* $L^\infty(\Omega)^{d \times d}$ *as* $n \to \infty$;

(e) *For all $n, j \in \mathbb{N}$ we have $\|\lambda_{n,j} \chi_{\mathcal{U}_\lambda(\boldsymbol{e}^n)}\|_s \leqslant c\, 2^{-\frac{j}{s}} \|\nabla \boldsymbol{e}^n\|_s$, with a constant $c > 0$ depending on s.*

Proof. The assertions follow by adopting the proof of Corollary 18. □