ESAIM: M2AN 50 (2016) 1333–1369
DOI: 10.1051/m2an/2015085

ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

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

Christian Kreuzer[1] and Endre Süli[2]

Abstract. We develop the *a posteriori* error analysis of finite element approximations to implicit power-law-like models for viscous incompressible fluids in d space dimensions, $d \in \{2, 3\}$. 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 $\frac{2d}{d+1} < r < \infty$. We establish upper and lower bounds on the finite element residual, as well as the local stability of the error bound. We then consider an adaptive finite element approximation of the problem, and, under suitable assumptions, we show the weak convergence of the adaptive algorithm to a weak solution of the boundary-value problem. The argument is based on a variety of weak compactness techniques, including Chacon's biting lemma and a finite element counterpart of the Acerbi–Fusco Lipschitz truncation of Sobolev functions, introduced by [L. Diening, C. Kreuzer and E. Süli, *SIAM J. Numer. Anal.* **51** (2013) 984–1015].

Mathematics Subject Classification. 65N30, 65N12, 76A05, 35Q35.

Received March 20, 2015. Revised September 16, 2015. Accepted October 30, 2015.

1. Introduction

Typical physical models of 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 constitutive hypothesis then leads to the Navier–Stokes system and its nonlinear generalizations, such as fluids with shear-rate-dependent viscosity including power-law fluids with constant or variable power-law index. 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 for the accurate description of 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, referred to as implicit constitutive theory, which was proposed recently in a series of papers by Rajagopal and collaborators; see, for example, [35–37]. The underlying principle of implicit constitutive theory in the context of viscous flows is the following: instead of demanding that the Cauchy stress

Keywords and phrases. Adaptive finite element methods, implicit constitutive models, power-law fluids, *a posteriori* analysis, convergence.

[1] Fakultät für Mathematik, Ruhr-Universität Bochum, Universitätsstrasse 150, 44801 Bochum, Germany.
christan.kreuzer@rub.de

[2] Mathematical Institute, University of Oxford, 24-29 St Giles', Oxford OX1 3LB, UK. Endre.Suli@maths.ox.ac.uk

is an explicit function of the symmetric part of the velocity gradient, one may allow an implicit relationship between these quantities. This then leads to a general theory, which admits fluid flow models with implicit and possibly discontinuous power-law-like rheology; see, [32, 33]. Very recently a rigorous mathematical existence theory was developed for these models by Bulíček *et al.* in [14], for $r > \frac{2d}{d+2}$; for the range $\frac{2d}{d+2} < r \leqslant \frac{3d}{d+2}$ the Acerbi–Fusco Lipschitz truncation [1] was used in order to prove the existence of a weak solution. In [22], using a variety of weak compactness techniques, we showed that a subsequence of the sequence of finite element solutions converges weakly to a weak solution of the problem as the finite element discretization parameter h tends to 0. A key new technical tool in the analysis presented in [22] was a finite element counterpart of the Acerbi–Fusco Lipschitz truncation of Sobolev functions. However, in the case of velocity approximations that are not exactly divergence-free the convergence theory developed there was restricted to the range $\frac{2d}{d+1} < r < \infty$.

The focus of the present paper is on the adaptive finite element approximation of implicitly constituted power-law-like models for viscous incompressible fluids. As in [22], the implicit constitutive relation between the stress and the symmetric part of the velocity gradient is approximated by an explicit (smooth) constitutive law. The resulting steady non-Newtonian flow problem is then discretized by a mixed finite element method. Guided by an *a posteriori* error analysis, we propose a numerical method with competing adaptive strategies for the mesh refinement and the approximation of the implicit constitutive law, and we present a rigorous convergence proof generalizing the ideas in [34] and [38]. More precisely, we show that a subsequence of the adaptively generated sequence of discrete approximations converges, in the weak topology of the ambient function space, to a weak solution of the model when $\frac{2d}{d+1} < r < \infty$. In contrast with [22], stimulated by ideas from [16] we shall be able to avoid resorting to the theory of Young measures. We emphasize that even in the case when the weak solution is unique we have only weak convergence of a subsequence; in this case, however, such a subsequence can be identified with the aid of the *a posteriori* bounds derived herein; *cf.* Remark 5.9.

The paper is structured as follows. In Section 2 we shall formulate the problem under consideration and will introduce some known mathematical results. In Section 3 we define the finite element approximation of the problem and present related technical properties and tools, such as the discrete Lipschitz truncation from [22]. Section 4 is concerned with the *a posteriori* error analysis for both the error in the approximation of the graph and the finite element approximation. The adaptive algorithm together with our main result are stated in Section 5; for the sake of clarity of the presentation certain technical parts of the proof are deferred to Section 6. We conclude the paper by discussing concrete graph approximations for certain problems of practical relevance. While the emphasis in this paper is on the mathematical analysis of adaptive finite element algorithms for implicitly constituted fluid flow models, the ideas developed herein may be of more general interest in the convergence analysis of adaptive finite element methods for other nonlinear problems in continuum mechanics with possibly nonunique weak solutions.

2. Implicitly constituted power-law-like fluids

In this section we introduce the variational model of steady flow, in a bounded open Lipschitz domain $\Omega \subset \mathbb{R}^d$, $d \in \{2, 3\}$, with polyhedral boundary $\partial\Omega$, of an incompressible fluid with an implicit constitutive law given by a maximal monotone x-dependent r-graph. We then recall the existence result from [14] together with some known results and mathematical tools from the literature.

2.1. The variational formulation

Before stating the weak formulation of the problem we need to introduce basic notations and recall some well-known properties of Lebesgue and Sobolev function spaces.

For a measurable subset $\omega \subset \mathbb{R}^d$, we denote the classical spaces of Lebesgue and vector-valued Sobolev functions by $(L^s(\omega) := L^s(\omega;\mathbb{R}), \|\cdot\|_{s;\omega})$ and $(W^{1,s}(\omega)^d := W^{1,s}(\omega;\mathbb{R}^d), \|\cdot\|_{1,s;\omega})$, $s \in [1, \infty]$, respectively. Henceforth ω will be assumed to have Lipschitz continuous boundary. We denote the space of functions in $W^{1,s}(\omega)^d$ with zero trace by $W^{1,s}_0(\omega)^d$ and let $W^{1,s}_{0,\,\mathrm{div}}(\omega)^d := \{\boldsymbol{v} \in W^{1,s}_0(\omega)^d\colon \operatorname{div} \boldsymbol{v} = 0\}$. 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 $\frac{1}{s} + \frac{1}{s'} = 1$ we have that $L^{s'}(\Omega)$

and $L_0^{s'}(\Omega)$ are the dual spaces of $L^s(\Omega)$ and $L_0^s(\Omega)$, respectively. We have, for such s and s', that $W_0^{1,s}(\Omega)^d$ is the closure of $\mathcal{D}(\Omega)^d := C_0^\infty(\omega)^d$ and its dual 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 $\|\cdot\|_{s,\Omega}$.

For $r \in (1,\infty)$, we define $r' \in (1,\infty)$ by $\frac{1}{r} + \frac{1}{r'} = 1$, and set

$$\tilde{r} := \begin{cases} \frac{1}{2}\frac{dr}{d-r} & \text{if } r \leqslant \frac{3d}{d+2} \\ r' & \text{otherwise.} \end{cases} \tag{2.1}$$

With such r, r' and $\tilde{r}$, we shall consider the following boundary-value problem.

Problem. For $\boldsymbol{f} \in L^{r'}(\Omega)^d$ find $(\boldsymbol{u}, p, \boldsymbol{S}) \in W_0^{1,r}(\Omega)^d \times L_0^{\tilde{r}}(\Omega) \times L^{r'}(\Omega; \mathbb{R}^{d\times d}_{\mathrm{sym}})$ such that

$$\begin{aligned} \operatorname{div}(\boldsymbol{u}\otimes\boldsymbol{u} + p\mathbf{1} - \boldsymbol{S}) &= \boldsymbol{f} && \text{in } \mathcal{D}'(\Omega)^d, \\ \operatorname{div}\boldsymbol{u} &= 0 && \text{in } \mathcal{D}'(\Omega), \\ (\mathrm{D}\boldsymbol{u}(x), \boldsymbol{S}(x)) &\in \mathcal{A}(x) && \text{for almost every } x \in \Omega. \end{aligned} \tag{2.2}$$

Here, $\mathrm{D}\boldsymbol{u} := \frac{1}{2}(\nabla\boldsymbol{u} + (\nabla\boldsymbol{u})^{\mathrm{T}}) \in \mathbb{R}^{d\times d}_{\mathrm{sym}} := \{\boldsymbol{\delta} \in \mathbb{R}^{d\times d} : \boldsymbol{\delta} = \boldsymbol{\delta}^{\mathrm{T}}\}$ signifies the symmetric part of the gradient of $\boldsymbol{u}$. As is indicated in our choice of the solution space for the velocity $\boldsymbol{u}$ in the statement of the above boundary-value problem, we shall suppose a homogenous Dirichlet boundary condition for $\boldsymbol{u}$. The integrability of the pressure p is inherited from the convective term and therefore the definition (2.1) of $\tilde{r}$ is a consequence of the embedding $W_0^{1,r}(\Omega) \hookrightarrow L^{2\tilde{r}}(\Omega)$. The implicit constitutive law, which relates the shear rate to the shear stress, is given by an inhomogeneous maximal monotone r-graph $\mathcal{A} : x \mapsto \mathcal{A}(x) \subset \mathbb{R}^{d\times d}_{\mathrm{sym}} \times \mathbb{R}^{d\times d}_{\mathrm{sym}}$. In particular, we assume that for almost every $x \in \Omega$ the following properties hold:

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

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

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

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

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

(A5) The set-valued mapping $\mathcal{A} : \Omega \to \mathbb{R}^{d\times d}_{\mathrm{sym}} \times \mathbb{R}^{d\times d}_{\mathrm{sym}}$ is measurable, *i.e.*, for any closed sets $\mathcal{C}_1, \mathcal{C}_2 \subset \mathbb{R}^{d\times d}_{\mathrm{sym}}$, we have that

$$\{x \in \Omega : \mathcal{A}(x) \cap (\mathcal{C}_1 \times \mathcal{C}_2) \neq \varnothing\}$$

is a Lebesgue measurable subset of Ω.

The following existence result was originally proved by Bulíček *et al.* in [14] assuming additionally that if $\boldsymbol{\delta}_1 \neq \boldsymbol{\delta}_2$ and $\boldsymbol{\sigma}_1 \neq \boldsymbol{\sigma}_2$, then the inequality in (A2) is strict. In fact, based on a generalization of the fundamental theorem on Young measures, this condition was required in order to prove that the implicit constitutive law is satisfied. For the unsteady case, they presented a new technique in [16] avoiding the additional condition. This technique can also be applied to steady problems (2.2); compare with [15].

Proposition 2.1. *For $r > \frac{2d}{d+2}$ there exists a (not necessarily unique) weak solution to problem* (2.2).

Remark 2.2. Two remarks concerning the definition of x-dependent maximal monotone graph $\mathcal{A}$ are now in order:

- Let $\mathcal{D} \subset \mathbb{R}^{d\times d}_{\text{sym}} \times \mathbb{R}^{d\times d}_{\text{sym}}$ be a closed set; then, for a.e. $x \in \Omega$ we have that the set $\mathcal{A}(x) \cap \mathcal{D}$ is closed. To see this, we assume w.l.o.g. that $\mathcal{A}(x) \cap \mathcal{D} \neq \emptyset$ and let $\{(\boldsymbol{\delta}_k, \boldsymbol{\sigma}_k)\}_{k\in\mathbb{N}} \subset \mathcal{A}(x) \cap \mathcal{D}$, such that $\boldsymbol{\delta}_k \to \boldsymbol{\delta} \in \mathbb{R}^{d\times d}_{\text{sym}}$ and $\boldsymbol{\sigma}_k \to \boldsymbol{\sigma} \in \mathbb{R}^{d\times d}_{\text{sym}}$ as $n \to \infty$. Let $(\bar{\boldsymbol{\delta}}, \bar{\boldsymbol{\sigma}}) \in \mathcal{A}(x)$ be arbitrary. We then have that
$$0 \leqslant (\bar{\boldsymbol{\sigma}} - \boldsymbol{\sigma}_k) : (\bar{\boldsymbol{\delta}} - \boldsymbol{\delta}_k) \to (\bar{\boldsymbol{\sigma}} - \boldsymbol{\sigma}) : (\bar{\boldsymbol{\delta}} - \boldsymbol{\delta})$$
as $k \to \infty$. This proves that $(\boldsymbol{\delta}, \boldsymbol{\sigma}) \in \mathcal{A}(x) \cap \mathcal{D}$ thanks to condition (A3) and the closedness of $\mathcal{D}$. Taking $\mathcal{D} := \{\boldsymbol{\sigma}\} \times \mathbb{R}^{d\times d}_{\text{sym}}$, we then deduce that the set
$$\{\boldsymbol{\sigma} \in \mathbb{R}^{d\times d}_{\text{sym}} : (\boldsymbol{\delta}, \boldsymbol{\sigma}) \in \mathcal{A}(x)\}$$
is closed. This is condition (A5)(i) of [14].
- According to ([4], Thm. 8.1.4) Property (A5) is equivalent to the fact that the graph of the set-valued map $\mathcal{A}(x)$ belongs to the product σ-algebra $\mathfrak{L}(\Omega) \otimes \mathfrak{B}(\mathbb{R}^{d\times d}_{\text{sym}}) \otimes \mathfrak{B}(\mathbb{R}^{d\times d}_{\text{sym}})$. Here $\mathfrak{L}(\Omega)$ denotes the Lebesgue measurable subsets of Ω and $\mathfrak{B}(\mathbb{R}^{d\times d}_{\text{sym}})$ the Borel subsets of $\mathbb{R}^{d\times d}_{\text{sym}}$. With the same argument it follows that (A5) is equivalent to the fact that, for any closed $\mathcal{C} \subset \mathbb{R}^{d\times d}_{\text{sym}}$, the sets
$$\big\{(x, \boldsymbol{\sigma}) \in \Omega \times \mathbb{R}^{d\times d}_{\text{sym}} : \text{there exists } \boldsymbol{\delta} \in \mathcal{C}, \text{ such that } (\boldsymbol{\delta}, \boldsymbol{\sigma}) \in \mathcal{A}(x)\big\},$$
$$\big\{(x, \boldsymbol{\delta}) \in \Omega \times \mathbb{R}^{d\times d}_{\text{sym}} : \text{there exists } \boldsymbol{\sigma} \in \mathcal{C}, \text{ such that } (\boldsymbol{\delta}, \boldsymbol{\sigma}) \in \mathcal{A}(x)\big\}$$
are measurable relative to $\mathfrak{L}(\Omega) \otimes \mathfrak{B}(\mathbb{R}^{d\times d}_{\text{sym}})$. These equivalent conditions imply that there exist measurable functions (so-called selections) $\boldsymbol{S}^\star, \boldsymbol{D}^\star : \Omega \times \mathbb{R}^{d\times d}_{\text{sym}} \to \mathbb{R}^{d\times d}_{\text{sym}}$ such that $\big(\boldsymbol{\delta}, \boldsymbol{S}^\star(x, \boldsymbol{\delta})\big), \big(\boldsymbol{D}^\star(x, \boldsymbol{\sigma}), \boldsymbol{\sigma}\big) \in \mathcal{A}(x)$ for a.e. $x \in \Omega$ and all $\boldsymbol{\delta}, \boldsymbol{\sigma} \in \mathbb{R}^{d\times d}_{\text{sym}}$; compare also with ([16], Rem. 1.1).

2.2. Analytical framework

We shall briefly recall some results that are crucial for the existence theory for problem (2.2).

Inf-sup condition. The inf-sup condition has a central role in the analysis of the Stokes problem. It states that, for $s, s' \in (1, \infty)$ with $\frac{1}{s} + \frac{1}{s'} = 1$, there exists an $\alpha_s > 0$ such that
$$\sup_{0\neq \boldsymbol{v} \in W^{1,s}_0(\Omega)^d} \frac{\int_\Omega q \operatorname{div} \boldsymbol{v} \,\mathrm{d}x}{\|\boldsymbol{v}\|_{1,s}} \geqslant \alpha_s \|q\|_{s'} \qquad \text{for all } q \in L^{s'}_0(\Omega). \tag{2.3}$$
This is the consequence of the existence of the *Bogovskiĭ* operator $\mathfrak{B} : L^s_0(\Omega) \to W^{1,s}_0(\Omega)^d$, with
$$\operatorname{div} \mathfrak{B}h = h \qquad \text{and} \qquad \alpha_s \|\mathfrak{B}h\|_{1,s} \leqslant \|h\|_s$$
for all $s \in (1, \infty)$; compare *e.g.* with [9, 20]. It follows from ([12], Sect. II, Prop. 1.2) that condition (2.3) is equivalent to the isomorphism
$$L^{s'}_0(\Omega) \cong \left\{\boldsymbol{v}' \in W^{-1,s'}(\Omega)^d : \langle \boldsymbol{v}', \boldsymbol{w}\rangle = 0 \text{ for all } \boldsymbol{w} \in W^{1,s}_{0,\operatorname{div}}(\Omega)^d\right\}. \tag{2.4}$$

Korn's inequality. According to (2.2) the maximal monotone graph defined in (A1)–(A5) provides control only of the symmetric part of the velocity gradient. Korn's inequality states that this already suffices to control the norm of a Sobolev function; *i.e.*, for $s \in (1, \infty)$, there exists a $\gamma_s > 0$ such that
$$\gamma_s \|\boldsymbol{v}\|_{1,s} \leqslant \|\mathrm{D}\boldsymbol{v}\|_s \qquad \text{for all } \boldsymbol{v} \in W^{1,s}_0(\Omega)^d; \tag{2.5}$$
compare e. g. with [20].

We conclude this subsection with Chacon's biting lemma and a corollary of it that is relevant for our purposes; compare *e.g.* with [13] and ([27], Lem. 7.3).

Lemma 2.3 (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 $j \in \mathbb{N}$.*

In other words, there exists a $v \in L^1(\Omega)$, such that for a subsequence (not relabelled) of $\{v_n\}_{n\in\mathbb{N}}$, $v_n \rightharpoonup v$ weakly in $L^1(\Omega\backslash E_j)$ as $n \to \infty$ for all $j \in \mathbb{N}$. We denote this by writing

$$v_n \overset{b}{\rightharpoonup} v \quad \textit{in } L^1(\Omega)$$

and call v the biting limit of $\{v_n\}_{n\in\mathbb{N}}$.

Lemma 2.4. *Let $\{v_n\}_{n\in\mathbb{N}} \subset L^1(\Omega)$ be a sequence of nonnegative functions such that $v_n \overset{b}{\rightharpoonup} v$ for some $v \in L^1(\Omega)$. Then,*

$$\lim_{n\to\infty} \int_\Omega (v_n - v)\,\mathrm{d}x = 0 \quad \textit{implies that} \quad v_n \rightharpoonup v \quad \textit{weakly in } L^1(\Omega) \textit{ as } n \to \infty.$$

3. Finite element approximation

This section is concerned with approximating problem (2.2) by the finite element method. To this end, we first approximate (2.2) by an explicitly constituted problem. We then introduce a general finite element framework for inf-sup stable Stokes elements. This, together with some representative examples of velocity-pressure pairs of finite element spaces, is the subject of Section 3.3. The finite element approximation of (2.2) is stated in Section 3.4.

3.1. 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, it can be approximated by a regular single-valued monotone tensor field based on a regularized *measurable selection* $\boldsymbol{S}^\star$ with the following properties; compare with [14, 16] and Remark 2.2.

Lemma 3.1 ([16], Lem. 2.2)**.** *Let $\boldsymbol{S}^\star : \Omega\times\mathbb{R}^{d\times d}_{sym} \to \times\mathbb{R}^{d\times d}_{sym}$ be a measurable selection of the x-dependent maximal monotone r-graph $\mathcal{A}$ with the properties* (A1)–(A5)*. Then, for $\boldsymbol{\delta}, \boldsymbol{\sigma} \in \mathbb{R}^{d\times d}_{sym}$, the following two statements are equivalent for almost all $x \in \Omega$:*

- $\big(\boldsymbol{\sigma} - \boldsymbol{S}^\star(x, \boldsymbol{D})\big) : (\boldsymbol{\delta} - \boldsymbol{D}) \geqslant 0 \qquad$ *for all $\boldsymbol{D} \in \mathbb{R}^{d\times d}_{sym}$;*
- $(\boldsymbol{\delta}, \boldsymbol{\sigma}) \in \mathcal{A}(x)$.

In [14, 26, 27] the selection $\boldsymbol{S}^\star$ is used to approximate the maximal monotone graph $\mathcal{A}$ by a single-valued monotone mapping $\boldsymbol{S}^n : \Omega \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$ based on a mollification technique. In order to allow for different practical implementations of such an approximation, we shall formulate its required properties and demonstrate in Section 7 how such graph-approximations can be constructed for some typical problems of practical interest within the class of problems under consideration.

Assumption 3.2. For $n \in \mathbb{N}$, there exists a mapping $\boldsymbol{S}^n : \Omega \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$, such that

- $\boldsymbol{S}^n(\cdot, \boldsymbol{\delta}) : \Omega \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$ is measurable for all $\boldsymbol{\delta} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$;
- $\boldsymbol{S}^n(x, \cdot) : \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$ is continuous for almost every $x \in \Omega$;
- $\boldsymbol{S}^n$ is strictly monotone; *i.e.*, for almost every $x \in \Omega$ we have

$$(\boldsymbol{S}^n(x, \boldsymbol{\delta}_1) - \boldsymbol{S}^n(x, \boldsymbol{\delta}_2)) : (\boldsymbol{\delta}_1 - \boldsymbol{\delta}_2) > 0 \quad \text{for all} \quad \boldsymbol{\delta}_1 \neq \boldsymbol{\delta}_2 \in \mathbb{R}^{d\times d}_{\mathrm{sym}};$$

- 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

$$|\boldsymbol{S}^n(x,\boldsymbol{\delta})| \leqslant \tilde{c}_1 |\boldsymbol{\delta}|^{r-1} + \tilde{k}(x) \quad \text{and} \quad \boldsymbol{S}^n(x,\boldsymbol{\delta}) : \boldsymbol{\delta} \geqslant \tilde{c}_2 |\boldsymbol{\delta}|^r - \tilde{m}(x)$$

for all $\boldsymbol{\delta} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$ and almost every $x \in \Omega$.

We emphasize that in contrast with [14, 16] we assume that $\boldsymbol{S}^n$ is strictly monotone. Of course we have to assume additionally that the graph of $\boldsymbol{S}^n$ converges to $\mathcal{A}$ is some sense. This will be specified in Assumption 5.6 with the aid of the *a posteriori* graph approximation indicator formulated in Section 4.1.

Having at hand a mapping $\boldsymbol{S}^n : \Omega \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$ as in Assumption 3.2, we aim to approximate the solution of (2.2) by solving the following explicitly constituted nonlinear boundary-value problem: For $\boldsymbol{f} \in L^{r'}(\Omega)^d$ find $(\boldsymbol{u}, p, \boldsymbol{S}) \in W^{1,r}_0(\Omega)^d \times L^{\tilde{r}}_0(\Omega) \times L^{r'}(\Omega, \mathbb{R}^{d\times d}_{\mathrm{sym}})$ such that

$$\begin{aligned} \operatorname{div}(\boldsymbol{u}\otimes\boldsymbol{u} + p\mathbf{1} - \boldsymbol{S}) &= \boldsymbol{f} && \text{in } \mathcal{D}'(\Omega)^d, \\ \operatorname{div}\boldsymbol{u} &= 0 && \text{in } \mathcal{D}'(\Omega), \\ \boldsymbol{S}(x) &= \boldsymbol{S}^n(x, \mathrm{D}\boldsymbol{u}(x)) && \text{for almost every } x \in \Omega. \end{aligned} \tag{3.1}$$

3.2. Domain partition and refinement framework

In this section we provide the framework for adaptive grid refinement. For the sake of simplicity, we restrict our presentation to conforming simplicial meshes and refinement by bisection. To be more precise, let $\mathcal{G}_0$ be a regular conforming partition of Ω into closed simplexes, the so-called macro mesh. Each simplex in the partition is referred to as an element. We assume that there exists a refinement routine `REFINE` with the following properties.

- The refinement routine has two input arguments: a regular conforming partition $\mathcal{G}$ and a subset $\mathcal{M} \subset \mathcal{G}$ of marked elements. The output is a refined regular conforming triangulation of Ω, where all elements in $\mathcal{M}$ have been bisected at least once. The input grid can be $\mathcal{G}_0$ or the output of a previous application of `REFINE`.
- *Shape-regularity:* We call $\mathcal{G}'$ a refinement of $\mathcal{G}$ (briefly $\mathcal{G}' \geqslant \mathcal{G}$), when it can be produced from $\mathcal{G}$ by a finite number of applications of `REFINE`. The set

$$\mathbb{G} := \{\mathcal{G} : \mathcal{G} \text{ is a refinement of } \mathcal{G}_0\}$$

is shape-regular, *i.e.*, for any element $E \in \mathcal{G}$ with $\mathcal{G} \in \mathbb{G}$, the ratio of its diameter to the diameter of the largest inscribed ball is bounded uniformly with respect to all partitions $\mathcal{G}$ with $\mathcal{G} \in \mathbb{G}$.

For the proof of existence of such a procedure, we refer to [5, 29, 40] or the monograph [39] and the references therein.

For every element $E \in \mathcal{G}$, $\mathcal{G} \in \mathbb{G}$, there exists an invertible affine mapping

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

where $\hat{E}$ is the standard reference d-simplex. The neighbourhood of an element $E \in \mathcal{G}$, with $\mathcal{G} \in \mathbb{G}$, is denoted by

$$\mathcal{N}^{\mathcal{G}}(E) := \{E' \in \mathcal{G} : E' \cap E \neq \varnothing\}.$$

Let $\omega \subset \Omega$ and define $\mathcal{U}^{\mathcal{G}}(\omega) := \bigcup\{E \in \mathcal{G} \mid E \cap \omega \neq \varnothing\}$. For subsets $\mathcal{M} \subset \mathcal{G}$, let

$$\Omega(\mathcal{M}) := \bigcup\{E \mid E \in \mathcal{M}\} \subset \Omega \quad \text{and} \quad \mathcal{U}^{\mathcal{G}}(\mathcal{M}) := \mathcal{U}^{\mathcal{G}}(\Omega(\mathcal{M})) \subset \Omega,$$

i.e., we have $\Omega = \Omega(\mathcal{G})$. Thanks to the shape-regularity of $\mathbb{G}$, we have that $\#\mathcal{N}^{\mathcal{G}}(E) \leqslant C$ and $|\mathcal{U}^{\mathcal{G}}(E)| = |\Omega(\mathcal{N}^{\mathcal{G}}(E))| \leqslant C\,|E|$ with a constant $C > 0$ independent of $\mathcal{G} \in \mathbb{G}$. For $\mathcal{G} \in \mathbb{G}$, we define the mesh-size function

$$\Omega \ni x \mapsto h_{\mathcal{G}}(x) := \left|\mathcal{U}^{\mathcal{G}}(\{x\})\right|^{1/d}.$$

For $x \in \operatorname{interior}(E)$, this coincides with the usual definition $h_{\mathcal{G}}(x) = |E|^{1/d} =: h_E$. The mesh-size function is monotonically decreasing under refinement.

We call the $(d-1)$-dimensional sub-simplexes of any simplex $E \in \mathcal{G}$, whose interiors lie inside Ω, the sides of $\mathcal{G}$ and denote the set of all of them by $\mathcal{S}(\mathcal{G})$. For $S \in \mathcal{S}(\mathcal{G})$, we define $h_S := |S|^{1/(d-1)}$ and observe for $x \in S$ that $ch_S \leqslant h_{\mathcal{G}}(x) \leqslant Ch_S$, with constants $C, c > 0$ depending solely on the shape-regularity of $\mathcal{G}$.

3.3. Finite element spaces

Denote by $\mathbb{P}_m$ the space of polynomials of degree at most $m \in \mathbb{N}$. For a given grid $\mathcal{G} \in \mathbb{G}$ and certain subspaces $\mathbb{Q} \subseteq L^\infty(\Omega)$ and $\mathbb{V} \subseteq W_0^{1,\infty}(\Omega)^d$ the finite element spaces are given by

$$\mathbb{V}(\mathcal{G}) := \left\{\boldsymbol{V} \in \mathbb{V} : \ \boldsymbol{V}|_E \circ \boldsymbol{F}_E^{-1} \in \hat{\mathbb{P}}_{\mathbb{V}},\ E \in \mathcal{G} \text{ and } \boldsymbol{V}|_{\partial\Omega} = 0\right\}, \tag{3.2a}$$

$$\mathbb{Q}(\mathcal{G}) := \left\{Q \in \mathbb{Q} : \ Q|_E \circ \boldsymbol{F}_E^{-1} \in \hat{\mathbb{P}}_{\mathbb{Q}},\ E \in \mathcal{G}\right\}, \tag{3.2b}$$

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 such that $\mathbb{P}_1^2 \subseteq \hat{\mathbb{P}}_{\mathbb{V}} \subseteq \mathbb{P}_\ell^d$ and $\mathbb{P}_0 \subseteq \hat{\mathbb{P}}_{\mathbb{Q}} \subseteq \mathbb{P}_\jmath$ for some $\ell \geqslant \jmath \in \mathbb{N}$. For convenience, we introduce the space of piecewise polynomials of degree at most $m \in \mathbb{N}$ over $\mathcal{G}$ by

$$\mathbb{P}_m(\mathcal{G}) := \{R : \bar{\Omega} \to \mathbb{R} \,:\, R|_E \in \mathbb{P}_m,\ E \in \mathcal{G}\}.$$

Note that $\mathbb{Q}(\mathcal{G}) \subset L^\infty(\Omega) \cap \mathbb{P}_\jmath(\mathcal{G})$ and since $\mathbb{V}(\mathcal{G}) \subset C_0(\bar{\Omega})^d \cap \mathbb{P}_\ell(\mathcal{G})^d$ it follows that $\mathbb{V}(\mathcal{G}) \subset W_0^{1,\infty}(\Omega)^d$. Additionally, we assume that the finite element spaces are nested, *i.e.*, if $\mathcal{G}_\star$ is a refinement of $\mathcal{G}$, then

$$\mathbb{V}(\mathcal{G}) \subset \mathbb{V}(\mathcal{G}_\star) \qquad \text{and} \qquad \mathbb{Q}(\mathcal{G}) \subset \mathbb{Q}(\mathcal{G}_\star). \tag{3.3}$$

Each of the above spaces is supposed to have a finite and locally supported basis; e. g. for the discrete velocity space this means that for $\mathcal{G} \in \mathbb{G}$ there exists an $N_{\mathcal{G}} \in \mathbb{N}$ such that

$$\mathbb{V}(\mathcal{G}) = \operatorname{span}\{\boldsymbol{V}_1^{\mathcal{G}}, \ldots, \boldsymbol{V}_{N_{\mathcal{G}}}^{\mathcal{G}}\}$$

and for each basis function $\boldsymbol{V}_i^{\mathcal{G}}$, $i = 1, \ldots, N_{\mathcal{G}}$, we have that if there exists an $E \in \mathcal{G}$ with $\boldsymbol{V}_i^{\mathcal{G}} \not\equiv 0$ on E, then $\operatorname{supp}\boldsymbol{V}_\jmath^{\mathcal{G}} \subset \mathcal{U}^{\mathcal{G}}(E)$. We introduce the subspace $\mathbb{V}_0(\mathcal{G})$ of discretely divergence-free functions by

$$\mathbb{V}_0(\mathcal{G}) := \left\{\boldsymbol{V} \in \mathbb{V}(\mathcal{G})\colon \int_\Omega Q \operatorname{div} \boldsymbol{V} \,\mathrm{d}x = 0 \text{ for all } Q \in \mathbb{Q}(\mathcal{G})\right\}$$

and we define

$$\mathbb{Q}_0(\mathcal{G}) := \left\{Q \in \mathbb{Q}(\mathcal{G}) : \int_\Omega Q \,\mathrm{d}x = 0\right\}.$$

It will be assumed throughout the paper that all pairs of velocity-pressure finite element spaces considered possess the following properties.

Assumption 3.3 (Projector $\mathfrak{I}_{\mathrm{div}}^{\mathcal{G}}$)**.** We assume that for each $\mathcal{G} \in \mathbb{G}$ there exists a linear projection operator $\mathfrak{I}_{\mathrm{div}}^{\mathcal{G}} : W_0^{1,1}(\Omega)^d \to \mathbb{V}(\mathcal{G})$ such that, for all $s \in (1, \infty)$,

- $\mathfrak{I}^{\mathcal{G}}_{\operatorname{div}}$ preserves divergence in $\mathbb{Q}(\mathcal{G})^*$; *i.e.*, for $\boldsymbol{v} \in W^{1,s}_0(\Omega)^d$ we have

$$\int_\Omega Q \operatorname{div} \boldsymbol{v} \,\mathrm{d}x = \int_\Omega Q \operatorname{div} \mathfrak{I}^{\mathcal{G}}_{\operatorname{div}} \boldsymbol{v} \,\mathrm{d}x \qquad \text{for all } Q \in \mathbb{Q}(\mathcal{G}).$$

- $\mathfrak{I}^{\mathcal{G}}_{\operatorname{div}}$ is locally defined; *i.e.*, for any other partition $\mathcal{G}_\star \in \mathbb{G}$ we have

$$\mathfrak{I}^{\mathcal{G}_\star}_{\operatorname{div}} \boldsymbol{v}|_{\mathcal{U}^{\mathcal{G}_\star}(E)} = \mathfrak{I}^{\mathcal{G}}_{\operatorname{div}} \boldsymbol{v}|_{\mathcal{U}^{\mathcal{G}}(E)} \tag{3.4}$$

  for all $\boldsymbol{v} \in W^{1,s}_0(\Omega)^d$ and all $E \in \mathcal{G}$ with $\mathcal{N}^{\mathcal{G}}(E) \subset \mathcal{G}_\star$.
- $\mathfrak{I}^{\mathcal{G}}_{\operatorname{div}}$ is locally $W^{1,1}$-stable; *i.e.*, there exists a $c_1 > 0$, independent of $\mathcal{G}$, such that

$$\int_E \left|\mathfrak{I}^{\mathcal{G}}_{\operatorname{div}} \boldsymbol{v}\right| + h_{\mathcal{G}} \left|\nabla \mathfrak{I}^{\mathcal{G}}_{\operatorname{div}} \boldsymbol{v}\right| \,\mathrm{d}x \leqslant c_1 \int_{\mathcal{U}^{\mathcal{G}}(E)} |\boldsymbol{v}| + h_{\mathcal{G}} \,|\nabla \boldsymbol{v}| \,\mathrm{d}x \tag{3.5}$$

  for all $\boldsymbol{v} \in W^{1,s}_0(\Omega)^d$ and all $E \in \mathcal{G}$.

As in [6, 17, 22], the local $W^{1,1}$-stability property (3.5) implies global $W^{1,s}$-stability, *i.e.*, for each $s \in [1,\infty]$, there exists a $c_s > 0$, such that

$$\left\|\mathfrak{I}^{\mathcal{G}}_{\operatorname{div}} \boldsymbol{v}\right\|_{1,s} \leqslant c_s \,\|\boldsymbol{v}\|_{1,s} \qquad \text{for all } \boldsymbol{v} \in W^{1,s}_0(\Omega)^d. \tag{3.6}$$

Moreover, since $\mathbb{V}(\mathcal{G})$ contains piecewise affine functions, we have the following interpolation error bound. For each $s \in [1,\infty]$ there exists a $c_s > 0$ such that

$$\int_E \left|\boldsymbol{v} - \mathfrak{I}^{\mathcal{G}}_{\operatorname{div}} \boldsymbol{v}\right|^s + h^s_{\mathcal{G}} \left|\nabla \boldsymbol{v} - \nabla \mathfrak{I}^{\mathcal{G}}_{\operatorname{div}} \boldsymbol{v}\right|^s \,\mathrm{d}x \leqslant c_s h_E^{s(1+\delta)} \,|\boldsymbol{v}|^s_{W^{1+\delta,s}(\mathcal{U}^{\mathcal{G}}(E))} \tag{3.7}$$

for all $E \in \mathcal{G}$ and $\boldsymbol{v} \in W^{1+\delta,s}(\Omega)^d \cap W^{1,s}_0(\Omega)^d$, $\delta \in \{0,1\}$.

As a consequence, we deduce the following result for weak limits in nested spaces. Before stating the result, we adopt the following notational convention: we shall write $A \lesssim B$ to denote $A \leqslant C \cdot B$ with a constant $C > 0$ that is independent of the discretization parameter h.

Proposition 3.4. *Let $\{\boldsymbol{v}_k\}_{k\in\mathbb{N}} \subset W^{1,s}_0(\Omega)^d$, $s \in (1,\infty)$, be such that $\boldsymbol{v}_k \rightharpoonup \mathbf{0}$ weakly in $W^{1,s}_0(\Omega)^d$ as $k \to \infty$ and let $\{\mathcal{G}_k\}_{k\in\mathbb{N}} \subset \mathbb{G}$ be a sequence of nested partitions of Ω,* i.e., *$\mathcal{G}_k \leqslant \mathcal{G}_{k+1}$ for all $k \in \mathbb{N}$. Then,*

$$\mathfrak{I}^{\mathcal{G}_k}_{\operatorname{div}} \boldsymbol{v}_k \rightharpoonup \mathbf{0} \quad \textit{weakly in } W^{1,s}_0(\Omega)^d \textit{ as } k \to \infty.$$

Proof. Thanks to the uniform boundedness (3.6) of the sequence of linear operators $\{\mathfrak{I}^{\mathcal{G}_k}_{\operatorname{div}} : W^{1,s}_0(\Omega)^d \to \mathbb{V}(\mathcal{G}_k) \subset W^{1,s}_0(\Omega)^d\}_{k\in\mathbb{N}}$, we have that there exists a not relabelled weakly converging subsequence of $\{\mathfrak{I}^{\mathcal{G}_k}_{\operatorname{div}} \boldsymbol{v}_k\}_{k\in\mathbb{N}}$ in $W^{1,s}_0(\Omega)^d$. By the compact embedding $W^{1,s}_0(\Omega)^d \hookrightarrow\hookrightarrow L^s(\Omega)^d$ the sequence $\{\mathfrak{I}^{\mathcal{G}_k}_{\operatorname{div}} \boldsymbol{v}_k\}_{k\in\mathbb{N}}$ converges strongly in $L^s(\Omega)^d$. Thanks to the uniqueness of the strong limit, it suffices to identify the limit of $\{\mathfrak{I}^{\mathcal{G}_k}_{\operatorname{div}} \boldsymbol{v}_k\}_{k\in\mathbb{N}}$ in $L^s(\Omega)^d$. To this end, we introduce the sets

$$\mathcal{G}^+_k := \bigcap_{j \geqslant k} \mathcal{G}_j \qquad \text{and} \qquad \mathring{\mathcal{G}}^+_k := \{E \in \mathcal{G}^+_k : \mathcal{N}^{\mathcal{G}_k}(E) \subset \mathcal{G}^+_k\},$$

i.e., $\mathcal{N}^{\mathcal{G}_j}(E) = \mathcal{N}^{\mathcal{G}_k}(E)$ for all $j \geqslant k$ and $E \in \mathring{\mathcal{G}}^+_k$. For $j \geqslant k$, we consider the decomposition

$$\mathfrak{I}^{\mathcal{G}_j}_{\operatorname{div}} \boldsymbol{v}_j = (\mathfrak{I}^{\mathcal{G}_j}_{\operatorname{div}} \boldsymbol{v}_j)\chi_{\Omega(\mathring{\mathcal{G}}^+_k)} + (\mathfrak{I}^{\mathcal{G}_j}_{\operatorname{div}} \boldsymbol{v}_j)\chi_{\Omega(\mathcal{G}_k \setminus \mathring{\mathcal{G}}^+_k)}.$$

For the latter term, we have according to (3.7) that

$$\left\|(\boldsymbol{v}_j - \mathfrak{I}_{\mathrm{div}}^{\mathcal{G}_j} \boldsymbol{v}_j)\chi_{\Omega(\mathcal{G}_k\setminus\mathring{\mathcal{G}}_k^+)}\right\|_s \lesssim \left\|h_{\mathcal{G}_j}\chi_{\mathcal{U}(\mathcal{G}_k\setminus\mathring{\mathcal{G}}_k^+)}\right\|_\infty \|\nabla \boldsymbol{v}_j\|_s \leqslant \left\|h_{\mathcal{G}_k}\chi_{\mathcal{U}(\mathcal{G}_k\setminus\mathring{\mathcal{G}}_k^+)}\right\|_\infty \|\nabla \boldsymbol{v}_j\|_s .$$

Here we have used the monotonicity of the mesh-size under refinement in the last step. It follows from ([34], Cor. 4.1 and (4.15)) that $\left\|h_{\mathcal{G}_k}\chi_{\Omega(\mathcal{G}_k\setminus\mathring{\mathcal{G}}_k^+)}\right\|_{L^\infty(\Omega)} \to 0$ as $k \to \infty$. Thanks to the shape-regularity of $\mathbb{G}$, this readily implies that

$$\lim_{k\to\infty} \left\|h_{\mathcal{G}_k}\chi_{\mathcal{U}(\mathcal{G}_k\setminus\mathring{\mathcal{G}}_k^+)}\right\|_{L^\infty(\Omega)} = 0. \tag{3.8}$$

By the compact embedding $W_0^{1,s}(\Omega)^d \hookrightarrow\hookrightarrow L^s(\Omega)^d$ we have that $\boldsymbol{v}_j \to \mathbf{0}$ strongly in $L^s(\Omega)^d$ as $j \to \infty$. Combining these observations, we deduce that for any $\epsilon > 0$ there exists a $K_\epsilon > 0$ such that

$$\left\|(\mathfrak{I}_{\mathrm{div}}^{\mathcal{G}_j} \boldsymbol{v}_j)\chi_{\Omega(\mathcal{G}_k\setminus\mathring{\mathcal{G}}_k^+)}\right\|_s \leqslant \epsilon \qquad \text{for all } j \geqslant k \geqslant K_\epsilon. \tag{3.9}$$

We next investigate the term $(\mathfrak{I}_{\mathrm{div}}^{\mathcal{G}_j} \boldsymbol{v}_j)\chi_{\Omega(\mathring{\mathcal{G}}_k^+)}$. Thanks to the definition of $\mathring{\mathcal{G}}_k^+$ and (3.4) we have

$$(\mathfrak{I}_{\mathrm{div}}^{\mathcal{G}_j} \boldsymbol{v}_j)|_{\Omega(\mathring{\mathcal{G}}_k^+)} = (\mathfrak{I}_{\mathrm{div}}^{\mathcal{G}_k} \boldsymbol{v}_j)|_{\Omega(\mathring{\mathcal{G}}_k^+)} \quad \text{for all } j \geqslant k.$$

Since a linear operator between two normed linear spaces is norm-continuous if and only if it is weakly continuous (*cf.* Theorem 6.17 in [3], for example,) we deduce that, for fixed $k \in \mathbb{N}$, we have

$$(\mathfrak{I}_{\mathrm{div}}^{\mathcal{G}_j} \boldsymbol{v}_j)|_{\Omega(\mathring{\mathcal{G}}_k^+)} \rightharpoonup \mathbf{0} \qquad \text{weakly in } W^{1,s}(\Omega(\mathring{\mathcal{G}}_k^+))^d \text{ as } j \to \infty.$$

By the compact embedding $W_0^{1,s}(\Omega)^d \hookrightarrow\hookrightarrow L^s(\Omega)^d$ this implies that

$$(\mathfrak{I}_{\mathrm{div}}^{\mathcal{G}_j} \boldsymbol{v}_j)\chi_{\Omega(\mathring{\mathcal{G}}_k^+)} \to \mathbf{0} \qquad \text{strongly in } L^s(\Omega)^d \text{ as } j \to \infty.$$

Together with (3.9) we have, for all $j \geqslant k \geqslant K_\epsilon$, that

$$\left\|\mathfrak{I}_{\mathrm{div}}^{\mathcal{G}_j} \boldsymbol{v}_j\right\|_s \leqslant \left\|(\mathfrak{I}_{\mathrm{div}}^{j} \boldsymbol{v}_j)\chi_{\Omega(\mathcal{G}_k\setminus\mathring{\mathcal{G}}_k^+)}\right\|_s + \left\|(\mathfrak{I}_{\mathrm{div}}^{\mathcal{G}_j} \boldsymbol{v}_j)\chi_{\Omega(\mathring{\mathcal{G}}_k^+)}\right\|_s \leqslant \epsilon + \left\|(\mathfrak{I}_{\mathrm{div}}^{\mathcal{G}_j} \boldsymbol{v}_j)\chi_{\Omega(\mathring{\mathcal{G}}_k^+)}\right\|_s \to \epsilon \qquad \text{as } j \to \infty.$$

Since $\epsilon > 0$ was arbitrary, this proves the assertion. □

Next, we shall introduce a quasi-interpolation operator, which will be important for the treatment of the, generally non-polynomial, stress approximation.

Assumption 3.5. We assume that for each $\mathcal{G} \in \mathbb{G}$ there exists a linear projection operator $\Pi_{\mathcal{G}} : L^1(\Omega; \mathbb{R}^{d\times d}_{\mathrm{sym}}) \to \mathbb{P}_{\ell-1}(\mathcal{G}; \mathbb{R}^{d\times d}_{\mathrm{sym}})$, such that $\Pi_{\mathcal{G}}$ is locally L^1 stable, *i.e.*, there exists a $c > 0$, depending on $\mathcal{G}_0$, such that

$$\int_E |\Pi_{\mathcal{G}} \boldsymbol{S}| \,\mathrm{d}x \leqslant c \int_{\mathcal{U}^{\mathcal{G}}(E)} |\boldsymbol{S}| \,\mathrm{d}x \qquad \text{for all } \boldsymbol{S} \in L^1(\Omega; \mathbb{R}^{d\times d}_{\mathrm{sym}}).$$

This implies that

$$\|\Pi_{\mathcal{G}} \boldsymbol{S}\|_s \leqslant c_s \|\boldsymbol{S}\|_s \qquad \text{for all } \boldsymbol{S} \in L^s(\Omega; \mathbb{R}^{d\times d}_{\mathrm{sym}}), \tag{3.10}$$

with a constant c_s depending on $\mathcal{G}_0$ and s; compare also with (3.6).

Assumption 3.6 (Projector $\mathfrak{I}_{\mathbb{Q}}^{\mathcal{G}}$)**.** We assume that for each $\mathcal{G} \in \mathbb{G}$ there exists a linear projection operator $\mathfrak{I}_{\mathbb{Q}}^{\mathcal{G}} : L^1(\Omega) \to \mathbb{Q}(\mathcal{G})$ such that, for all $s' \in (1, \infty)$, $\mathfrak{I}_{\mathbb{Q}}^{\mathcal{G}}$ is locally L^1 stable, *i.e.*, there exists a $c > 0$, independent of $\mathcal{G}$, such that

$$\int_E \left|\mathfrak{I}_{\mathbb{Q}}^{\mathcal{G}} q\right| \,\mathrm{d}x \leqslant c \int_{\mathcal{U}^{\mathcal{G}}(E)} |q| \,\mathrm{d}x \qquad \text{for all } q \in L^1(\Omega) \text{ and all } E \in \mathcal{G}.$$

We may argue similarly as for $\mathfrak{I}_{\mathrm{div}}^{\mathcal{G}}$ to deduce that

$$\left\|\mathfrak{I}_{\mathbb{Q}}^{\mathcal{G}} q\right\|_s \leqslant c_s \|q\|_s \quad \text{and} \quad \int_E \left|q - \mathfrak{I}_{\mathbb{Q}}^{\mathcal{G}} q\right|^s \,\mathrm{d}x \leqslant c_s h_E^{\delta s} |q|^s_{W^{\delta,s}(\mathcal{U}^{\mathcal{G}}(E))} \tag{3.11}$$

for all $E \in \mathcal{G}$ and $q \in W^{\delta,s}(\Omega)$, $\delta \in \{0, 1\}$.

As a consequence of (2.3) and Assumption 3.3 (compare also with (3.6)) the following discrete counterpart of (2.3) holds; see [6].

Proposition 3.7 (Inf-sup stability)**.** *For all $s, s' \in (1, \infty)$ with $\frac{1}{s} + \frac{1}{s'} = 1$, there exists a $\beta_s > 0$, independent of $\mathcal{G} \in \mathbb{G}$, such that*

$$\sup_{0 \neq \boldsymbol{V} \in \mathbb{V}(\mathcal{G})} \frac{\int_\Omega Q \operatorname{div} \boldsymbol{V} \,\mathrm{d}x}{\|\boldsymbol{V}\|_{1,s}} \geqslant \beta_s \|Q\|_{s'} \qquad \textit{for all} \quad Q \in \mathbb{Q}_0(\mathcal{G}).$$

Thanks to the above considerations, there exists a discrete Bogovskiĭ operator, which has the following properties; compare also with ([22], Cor. 9).

Corollary 3.8 (Discrete Bogovskiĭ operator)**.** *The linear operator $\mathfrak{B}^{\mathcal{G}} := \mathfrak{I}_{\mathrm{div}}^{\mathcal{G}} \circ \mathfrak{B} : \operatorname{div} \mathbb{V}(\mathcal{G}) \to \mathbb{V}(\mathcal{G})$ satisfies*

$$\operatorname{div}(\mathfrak{B}^{\mathcal{G}} H) = H \qquad \textit{and} \qquad \beta_s \left\|\mathfrak{B}^{\mathcal{G}} H\right\|_{1,s} \leqslant \sup_{Q \in \mathbb{Q}(\mathcal{G})} \frac{\int_\Omega HQ \,\mathrm{d}x}{\|Q\|_{s'}}$$

for all $H \in \operatorname{div} \mathbb{V}(\mathcal{G})$ and $s \in (1, \infty)$, with a positive constant β_s, independent of $\mathcal{G} \in \mathbb{G}$.

Moreover, let $\{\mathcal{G}_k\}_{k \in \mathbb{N}} \subset \mathbb{G}$ be a sequence of nested partitions of Ω, i.e., *$\mathcal{G}_{k+1} \geqslant \mathcal{G}_k$ for all $k \in \mathbb{N}$, and let $\boldsymbol{V}_k \in \mathbb{V}(\mathcal{G}_k)$ be such that $\boldsymbol{V}_k \rightharpoonup \mathbf{0}$ weakly in $W_0^{1,s}(\Omega)^d$ as $k \to \infty$. We then have that*

$$\mathfrak{B}^{\mathcal{G}_k} \operatorname{div} \boldsymbol{V}_k \rightharpoonup \mathbf{0} \quad \textit{weakly in } W_0^{1,s}(\Omega)^d \textit{ as } k \to \infty.$$

Proof. The claim follows as in ([21], Cor. 10) after replacing ([21], Prop. 7) in the proof by Proposition 3.4 here. □

Upon integration by parts, it follows that

$$-\int_\Omega (\boldsymbol{v} \otimes \boldsymbol{w}) : \nabla \boldsymbol{h} \,\mathrm{d}x = \int_\Omega (\boldsymbol{v} \otimes \boldsymbol{h}) : \nabla \boldsymbol{w} + (\operatorname{div} \boldsymbol{v})(\boldsymbol{w} \cdot \boldsymbol{h}) \,\mathrm{d}x \tag{3.12}$$

for all $\boldsymbol{v}, \boldsymbol{w}, \boldsymbol{h} \in \mathcal{D}(\Omega)^d$. The last term vanishes provided that $\operatorname{div} \boldsymbol{v} \equiv 0$, *i.e.*, the convection term is skew-symmetric with respect to the second and the third argument, which implies that

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

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

$$\int_\Omega Q \operatorname{div} \boldsymbol{V} \,\mathrm{d}x = 0 \qquad \text{for all } Q \in \mathbb{Q}(\mathcal{G}), \tag{3.13}$$

i.e., if $\boldsymbol{V}$ is discretely divergence-free. As in [41], we wish to ensure that the discrete counterpart of the convection term inherits this skew-symmetry of the convection term. To this end, we observe from (3.12) that

$$-\int_{\Omega}(\boldsymbol{v}\otimes\boldsymbol{w}):\nabla\boldsymbol{h}\,\mathrm{d}x=\frac{1}{2}\int_{\Omega}(\boldsymbol{v}\otimes\boldsymbol{h}):\nabla\boldsymbol{w}-(\boldsymbol{v}\otimes\boldsymbol{w}):\nabla\boldsymbol{h}\,\mathrm{d}x=:\mathcal{B}[\boldsymbol{v},\,\boldsymbol{w},\,\boldsymbol{h}] \tag{3.14}$$

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

$$\mathcal{B}[\boldsymbol{v},\,\boldsymbol{v},\,\boldsymbol{v}]=0\qquad\text{for all }\boldsymbol{v}\in W^{1,\infty}(\Omega)^d. \tag{3.15}$$

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 (2.1). 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

$$W^{1,r}(\Omega)^d\hookrightarrow L^{2\tilde{r}}(\Omega)^d \tag{3.16}$$

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

$$\int_{\Omega}(\boldsymbol{v}\otimes\boldsymbol{w}):\nabla\boldsymbol{h}\,\mathrm{d}x\leqslant c\,\|\boldsymbol{v}\|_{1,r}\,\|\boldsymbol{w}\|_{1,r}\,\|\boldsymbol{h}\|_{1,\tilde{r}'} \tag{3.17}$$

for all $\boldsymbol{v},\boldsymbol{w},\boldsymbol{h}\in W^{1,\infty}(\Omega)^d$; see [14]. Because of the extension (3.14) 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 an $s\in(1,\infty)$ 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 (3.16) and $W^{1,\tilde{r}'}_0(\Omega)^d\hookrightarrow L^s(\Omega)^d$. Consequently, together with (3.17) we thus obtain

$$\mathcal{B}[\boldsymbol{v},\,\boldsymbol{w},\,\boldsymbol{h}]\leqslant c\,\|\boldsymbol{v}\|_{1,r}\,\|\boldsymbol{w}\|_{1,r}\,\|\boldsymbol{h}\|_{1,\tilde{r}'}\,. \tag{3.18}$$

In view of (3.14), for $\boldsymbol{v}=(v_1,\ldots,v_d)^{\mathrm{T}}\in W^{1,r}_0(\Omega)^d$, the convective term can be reformulated as

$$\int_{\Omega}\mathrm{B}[\boldsymbol{v},\boldsymbol{v}]\cdot\boldsymbol{w}\,\mathrm{d}x=\mathcal{B}[\boldsymbol{v},\,\boldsymbol{v},\,\boldsymbol{w}],\qquad\boldsymbol{w}\in W^{1,\tilde{r}}_0(\Omega)^d, \tag{3.19}$$

where $\mathrm{B}[\boldsymbol{v},\boldsymbol{v}]\in L^{\tilde{r}}(\Omega)^d$ is defined by $(\mathrm{B}[\boldsymbol{v},\boldsymbol{v}])_j=\frac{1}{2}\sum_{i=1}^d v_i\frac{\partial v_i}{\partial x_j}+\frac{\partial}{\partial x_i}(v_iv_j)$ for $j=1,\ldots,d$. In particular, for $\boldsymbol{v}=\boldsymbol{V}\in\mathbb{V}(\mathcal{G})$, we have that $\mathrm{B}[\boldsymbol{V},\boldsymbol{V}]\in\mathbb{P}_{2\ell-1}(\mathcal{G})^d$.

Example 3.9. The following velocity-pressure pairs of finite elements satisfy Assumptions 3.3 and 3.6 for $d=2,3$ (see, *e.g.*, [6, 24, 25]):

- The lowest order Taylor–Hood element;
- Spaces of continuous piecewise quadratic elements for the velocity and piecewise constants for the pressure (see *e.g.* [12], Sect. VI Ex. 3.6).

We note that the MINI element and the conforming Crouzeix–Raviart Stokes element do not satisfy the nestedness hypothesis stated in (3.3).

Remark 3.10. The boundedness of the trilinear form $\mathcal{B}[\cdot,\cdot,\cdot]$ stated in (3.18) requires that $r > \frac{2d}{d+1}$. In [21] and [22] the set of admissible values of r was the same range, $r \in (\frac{2d}{d+2}, \infty)$, as in the existence theorem for the continuous problem in [14]; however, for $r \in (\frac{2d}{d+2}, \frac{2d}{d+1}]$ the finite element space for the velocity was assumed in [21] and [22] to consist of pointwise divergence-free functions, whose construction is more complicated. For simplicity, we shall therefore confine ourselves here to the limited range of $r > \frac{2d}{d+1}$ so as to be able to admit standard discretely divergence-free (*cf.* (3.13)) finite element velocity spaces.

3.4. The Galerkin approximation

We are now ready to state the discrete problem. Let $\{\mathbb{V}(\mathcal{G}), \mathbb{Q}(\mathcal{G})\}_{\mathcal{G}\in\mathbb{G}}$ be the finite element spaces of Section 3.3.

For $n \in \mathbb{N}$ and $\mathcal{G} \in \mathbb{G}$ we call a triple of functions $\big(\boldsymbol{U}_{\mathcal{G}}^n, P_{\mathcal{G}}^n\big) \in \mathbb{V}(\mathcal{G}) \times \mathbb{Q}_0(\mathcal{G})$ a Galerkin approximation of (3.1) if it satisfies

$$\begin{aligned}\int_\Omega \boldsymbol{S}^n(\cdot, \mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n) : \mathrm{D}\boldsymbol{V} + \mathrm{B}[\boldsymbol{U}_{\mathcal{G}}^n, \boldsymbol{U}_{\mathcal{G}}^n] \cdot \boldsymbol{V} - P_{\mathcal{G}}^n \operatorname{div} \boldsymbol{V} \,\mathrm{d}x &= \int_\Omega \boldsymbol{f} \cdot \boldsymbol{V} \,\mathrm{d}x, \\ \int_\Omega Q \operatorname{div} \boldsymbol{U}_{\mathcal{G}}^n \,\mathrm{d}x &= 0,\end{aligned} \tag{3.20}$$

for all $\boldsymbol{V} \in \mathbb{V}(\mathcal{G})$ and $Q \in \mathbb{Q}(\mathcal{G})$.

Restricting the test-functions to $\mathbb{V}_0(\mathcal{G})$ the discrete problem (3.20) reduces to finding $\boldsymbol{U}_{\mathcal{G}}^n \in \mathbb{V}_0(\mathcal{G})$ such that

$$\int_\Omega \boldsymbol{S}^n(\cdot, \mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n) : \mathrm{D}\boldsymbol{V} \,\mathrm{d}x + \mathcal{B}[\boldsymbol{U}_{\mathcal{G}}^n, \boldsymbol{U}_{\mathcal{G}}^n, \boldsymbol{V}] = \int_\Omega \boldsymbol{f} \cdot \boldsymbol{V} \,\mathrm{d}x \tag{3.21}$$

for all $\boldsymbol{V} \in \mathbb{V}_0(\mathcal{G})$. Thanks to (3.15), it follows from Assumption 3.2 and Korn's inequality (2.5) that the nonlinear operator defined on $\mathbb{V}_0(\mathcal{G})$ by the left-hand side of (3.21) is coercive and continuous on $\mathbb{V}_0(\mathcal{G})$. Since the dimension of $\mathbb{V}(\mathcal{G})$ is finite, Brouwer's fixed point theorem ensures the existence of a solution to (3.21). The existence of a solution triple to (3.20) then follows by the discrete inf-sup stability, Proposition 3.7. Of course, because of the weak assumptions in the definition of the maximal monotone r-graph, (3.20) does not define the Galerkin approximation $(\boldsymbol{U}_{\mathcal{G}}^n, P_{\mathcal{G}}^n)$ uniquely. However, supposing the axiom of choice, for each $n \in \mathbb{N}$, $\mathcal{G} \in \mathbb{G}$, we may choose an arbitrary one among possibly infinitely many solution triples and thus obtain

$$\big\{\big(\boldsymbol{U}_{\mathcal{G}}^n, P_{\mathcal{G}}^n, \boldsymbol{S}^n(\cdot, \mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n)\big)\big\}_{n\in\mathbb{N}, \mathcal{G}\in\mathbb{G}}. \tag{3.22}$$

From (3.20) we see that $\boldsymbol{U}_{\mathcal{G}}^n$ is discretely divergence-free and thus, thanks to (3.21) and (3.15), we have that

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

The coercivity of $\boldsymbol{S}^n$ (Assump. 3.2) and Korn's inequality (2.5) imply that the sequence $\{\boldsymbol{U}_{\mathcal{G}}^n\}_{n\in\mathbb{N}}$ is bounded in the norm of $W_0^{1,r}(\Omega)^d$, independently of $\mathcal{G} \in \mathbb{G}$ and $n \in \mathbb{N}$. This in turn implies, again by Assumption 3.2, the uniform boundedness of $\boldsymbol{S}^n(\cdot, \mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n)$ in $L^{r'}(\Omega; \mathbb{R}^{d\times d}_{\mathrm{sym}})$. In other words, there exists a constant $c_{\boldsymbol{f}} > 0$ depending on the data $\boldsymbol{f}$, such that

$$\big\|\boldsymbol{U}_{\mathcal{G}}^n\big\|_{1,r} + \big\|\boldsymbol{S}^n(\cdot, \mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n)\big\|_{r'} \leqslant c_{\boldsymbol{f}}, \qquad \text{for all } \mathcal{G} \in \mathbb{G} \text{ and } n \in \mathbb{N}. \tag{3.23}$$

For the sake of simplicity of the presentation, if there is no risk of confusion, we will denote in what follows $\boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n) = \boldsymbol{S}^n(\cdot, \mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n)$.

Remark 3.11. An alternative formulation of (3.21) is as follows: find a triple $\big(\boldsymbol{U}_{\mathcal{G}}^n,\, P_{\mathcal{G}}^n,\, \boldsymbol{S}_{\mathcal{G}}^n\big) \in \mathbb{V}(\mathcal{G}) \times \mathbb{Q}_0(\mathcal{G}) \times \mathbb{P}_{\ell-1}(\mathcal{G}; \mathbb{R}^{d\times d}_{\text{sym}})$ such that

$$\int_\Omega \boldsymbol{S}_{\mathcal{G}}^n : \mathrm{D}\boldsymbol{V} + \mathrm{B}[\boldsymbol{U}_{\mathcal{G}}^n, \boldsymbol{U}_{\mathcal{G}}^n] \cdot \boldsymbol{V} - P_{\mathcal{G}}^n \operatorname{div} \boldsymbol{V} \,\mathrm{d}x = \int_\Omega \boldsymbol{f} \cdot \boldsymbol{V} \,\mathrm{d}x,$$

$$\int_\Omega Q \operatorname{div} \boldsymbol{U}_{\mathcal{G}}^n \,\mathrm{d}x = 0,$$

$$\int_\Omega \boldsymbol{S}_{\mathcal{G}}^n : \boldsymbol{D} \,\mathrm{d}x = \int_\Omega \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n) : \boldsymbol{D} \,\mathrm{d}x,$$

for all $\boldsymbol{V} \in \mathbb{V}(\mathcal{G})$, $Q \in \mathbb{Q}(\mathcal{G})$, and $\boldsymbol{D} \in \mathbb{P}_{\ell-1}(\mathcal{G}; \mathbb{R}^{d\times d}_{\text{sym}})$. Here $\mathbb{P}_{\ell-1}(\mathcal{G}; \mathbb{R}^{d\times d}_{\text{sym}})$ denotes the space of all piecewise polynomials of degree $\leqslant \ell - 1$ on $\mathcal{G}$ with values in $\mathbb{R}^{d\times d}_{\text{sym}}$. In particular, if we define $\varPi_{\mathcal{G}} : L^1(\Omega; \mathbb{R}^{d\times d}_{\text{sym}}) \to \mathbb{P}_{\ell-1}(\mathcal{G}; \mathbb{R}^{d\times d}_{\text{sym}})$ by

$$\int_\Omega \varPi_{\mathcal{G}} \boldsymbol{S} : \boldsymbol{D} \,\mathrm{d}x = \int_\Omega \boldsymbol{S} : \boldsymbol{D} \,\mathrm{d}x, \quad \text{for all } \boldsymbol{D} \in \mathbb{P}_{\ell-1}(\mathcal{G}; \mathbb{R}^{d\times d}_{\text{sym}}),$$

then Assumption 3.5 can be easily verified and $\boldsymbol{S}_{\mathcal{G}}^n$ may take the role of $\varPi_{\mathcal{G}} \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n)$ in the subsequent analysis.

3.5. Discrete Lipschitz truncation

In this section we shall recall a discrete counterpart of Lipschitz truncation, which acts on finite element spaces. This *discrete Lipschitz truncation* is a composition of a continuous Lipschitz truncation with a projection onto the finite element space. The continuous Lipschitz truncation used here is based on results from [10, 11, 19], which provides finer estimates than the original Lipschitz truncation technique proposed by Acerbi and Fusco in [1]; for details consider [22].

We summarize the properties of the discrete Lipschitz truncation in the following result. Similar results for Sobolev functions can be found in [19] and [10].

Proposition 3.12. *Let $1 < s < \infty$ and let $\{\boldsymbol{E}_k\}_{k\in\mathbb{N}}$ be a sequence such that for all $k \in \mathbb{N}$ we have $\boldsymbol{E}_k \in \mathbb{V}(\mathcal{G}_k)$ for some $\mathcal{G}_k \in \mathbb{G}$. In addition, assume that $\{\boldsymbol{E}_k\}_{k\in\mathbb{N}} \subset W_0^{1,s}(\Omega)^d$ converges to zero weakly in $W_0^{1,s}(\Omega)^d$, as $k \to \infty$.*

Then, there exists a sequence $\{\lambda_{k,j}\}_{k,j\in\mathbb{N}} \subset \mathbb{R}$ with $2^{2^j} \leqslant \lambda_{k,j} \leqslant 2^{2^{j+1}-1}$ and Lipschitz truncated functions $\boldsymbol{E}_{k,j} = \boldsymbol{E}_{k,\lambda_{k,j}}$, $k, j \in \mathbb{N}$, with the following properties:

(a) $\boldsymbol{E}_{k,j} \in \mathbb{V}(\mathcal{G}_k)$;
(b) $\|\boldsymbol{E}_{k,j}\|_{1,s} \leqslant c\, \|\boldsymbol{E}_k\|_{1,s}$ *for* $1 < s \leqslant \infty$;
(c) $\|\nabla \boldsymbol{E}_{k,j}\|_\infty \leqslant c\, \lambda_{k,j}$;
(d) $\boldsymbol{E}_{k,j} \to 0$ *in* $L^\infty(\Omega)^d$ *as* $k \to \infty$;
(e) $\nabla \boldsymbol{E}_{k,j} \rightharpoonup^* 0$ *in* $L^\infty(\Omega)^{d\times d}$ *as* $k \to \infty$;
(f) *For all* $k, j \in \mathbb{N}$ *we have* $\big\|\lambda_{k,j}\, \chi_{\{\boldsymbol{E}_k \neq \boldsymbol{E}_{k,j}\}}\big\|_s \leqslant c\, 2^{-\frac{j}{s}}\, \|\nabla \boldsymbol{E}_k\|_s$.

The constants c appearing in the inequalities (b), (c) and (f) depend on d, Ω, $\hat{\mathbb{P}}_{\mathbb{V}}$ and the shape-regularity of $\{\mathcal{G}_k\}_{k\in\mathbb{N}}$. The constants in (b) and (f) also depend on s.

Proof. The proof is exactly the same as that of ([21], Thm. 17 and Cor. 18 replacing [21], Prop. 7) by Proposition 3.4. □

4. Error Analysis

4.1. Graph approximation error

In order to quantify the error committed in the approximation of the graph $\mathcal{A}(x)$, $x \in \Omega$, we introduce the following indicator. For $\boldsymbol{D} \in L^r(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}})$, $\boldsymbol{S} \in L^{r'}(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}})$, we define

$$\mathcal{E}_{\mathcal{A}}(\boldsymbol{D},\boldsymbol{S}) := \int_\Omega \inf_{(\boldsymbol{\delta},\boldsymbol{\sigma})\in\mathcal{A}(x)} |\boldsymbol{D}-\boldsymbol{\delta}|^r + |\boldsymbol{S}-\boldsymbol{\sigma}|^{r'} \,\mathrm{d}x. \tag{4.1}$$

The following result shows that this indicator is well-defined.

Proposition 4.1. *Let $\boldsymbol{D} \in L^r(\Omega;\mathbb{R}^{d\times d}_{sym})$ and $\boldsymbol{S} \in L^{r'}(\Omega;\mathbb{R}^{d\times d}_{sym})$; then, the mapping*

$$x \mapsto \inf_{(\boldsymbol{\delta},\boldsymbol{\sigma})\in\mathcal{A}(x)} |\boldsymbol{D}(x)-\boldsymbol{\delta}|^r + |\boldsymbol{S}(x)-\boldsymbol{\sigma}|^{r'}$$

is integrable. Moreover, there exist $\tilde{\boldsymbol{D}} \in L^r(\Omega;\mathbb{R}^{d\times d}_{sym})$ and $\tilde{\boldsymbol{S}} \in L^{r'}(\Omega;\mathbb{R}^{d\times d}_{sym})$ such that $\big(\tilde{\boldsymbol{D}}(x),\tilde{\boldsymbol{S}}(x)\big) \in \mathcal{A}(x)$ for a.e. $x \in \Omega$ and

$$\mathcal{E}_{\mathcal{A}}(\boldsymbol{D},\boldsymbol{S}) = \int_\Omega \left|\boldsymbol{D}-\tilde{\boldsymbol{D}}\right|^r + \left|\boldsymbol{S}-\tilde{\boldsymbol{S}}\right|^{r'} \,\mathrm{d}x.$$

Proof. The first claim is an immediate consequence of the second one. The second assertion follows from ([4], Thm. 8.2.11) by observing that the mapping

$$\Omega \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \ni \big(x;(\boldsymbol{\delta},\boldsymbol{\sigma})\big) \mapsto |\boldsymbol{D}(x)-\boldsymbol{\delta}|^r + |\boldsymbol{S}(x)-\boldsymbol{\sigma}|^{r'}$$

is Carathéodory, *i.e.*, $x \mapsto |\boldsymbol{D}(x)-\boldsymbol{\delta}|^r + |\boldsymbol{S}(x)-\boldsymbol{\sigma}|^{r'}$ is measurable for all $\boldsymbol{\delta},\boldsymbol{\sigma} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$ and

$$(\boldsymbol{\delta},\boldsymbol{\sigma}) \mapsto |\boldsymbol{D}(x)-\boldsymbol{\delta}|^r + |\boldsymbol{S}(x)-\boldsymbol{\sigma}|^{r'}$$

is continuous for a.e. $x \in \Omega$. □

4.2. *A posteriori* finite element error estimates

In this section we shall prove bounds on the residual

$$\mathcal{R}(\boldsymbol{U}^n_{\mathcal{G}}, P^n_{\mathcal{G}}, \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}^n_{\mathcal{G}})) = (\mathcal{R}^{\mathsf{pde}}(\boldsymbol{U}^n_{\mathcal{G}}, P^n_{\mathcal{G}}, \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}^n_{\mathcal{G}})), \mathcal{R}^{\mathsf{ic}}(\boldsymbol{U}^n_{\mathcal{G}})) \in W^{-1,\tilde{r}}(\Omega)^d \times L^r_0(\Omega)$$

of (3.1). In particular, for $(\boldsymbol{v},q,\boldsymbol{T}) \in W^{1,r}_0(\Omega)^d \times L^{\tilde{r}}_0(\Omega) \times L^{r'}(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}})$ we have

$$\begin{aligned}\big\langle \mathcal{R}(\boldsymbol{v},q,\boldsymbol{T}),\,(\boldsymbol{w},o)\big\rangle &:= \big\langle \mathcal{R}^{\mathsf{pde}}(\boldsymbol{v},q,\boldsymbol{T}),\,\boldsymbol{w}\big\rangle + \big\langle \mathcal{R}^{\mathsf{ic}}(\boldsymbol{v}),\,o\big\rangle \\ &:= \int_\Omega \boldsymbol{T} : \mathrm{D}\boldsymbol{w} + \mathrm{B}[\boldsymbol{v},\boldsymbol{v}]\cdot\boldsymbol{w} - q\,\mathrm{div}\,\boldsymbol{w} - \boldsymbol{f}\cdot\boldsymbol{w}\,\mathrm{d}x - \int_\Omega o\,\mathrm{div}\,\boldsymbol{v}\,\mathrm{d}x,\end{aligned} \tag{4.2}$$

where $(\boldsymbol{w},o) \in W^{1,\tilde{r}'}_0(\Omega)^d \times L^{r'}(\Omega)/\mathbb{R}$. Although for the sake of simplicity we restrict ourselves here to residual-based estimates, we note that in principle other *a posteriori* techniques, such as hierarchical estimates, flux-equilibration or estimates based on local problems, can be used as well; compare with [34, 38]. For $n \in \mathbb{N}$ and $\mathcal{G} \in \mathbb{G}$ let $\big(\boldsymbol{U}^n_{\mathcal{G}},\, P^n_{\mathcal{G}}\big) \in \mathbb{V}(\mathcal{G}) \times \mathbb{Q}_0(\mathcal{G})$ be the Galerkin approximation defined in (3.20). We begin with some preliminary observations.

The first part of the residual in (4.2), $\mathcal{R}^{\mathsf{pde}}(\boldsymbol{v},q,\boldsymbol{T}) \in W^{-1,\tilde{r}}(\Omega)^d$, provides information about how well the functions $\boldsymbol{v},q,\boldsymbol{T}$ satisfy the first equation in (3.1). For the second part, we have $\mathcal{R}^{\mathsf{ic}}(\boldsymbol{v}) \in (L^r_0(\Omega))^*$. We note that the space $(L_0^r(\Omega))^*$ is isometrically isomorphic to $L^{r'}(\Omega)/\mathbb{R}$, which is, in turn, isomorphic to $L_0^{r'}(\Omega)$ since $r \in (1,\infty)$. The term $\langle \mathcal{R}^{\mathsf{ic}}(\boldsymbol{v}), o\rangle$ provides information about the compressibility of $\boldsymbol{v}$.

We emphasize that $\mathcal{R}(\boldsymbol{v}, q, \boldsymbol{T}) = 0$ if and only if $\mathcal{R}^{\mathsf{pde}}(\boldsymbol{v}, q, \boldsymbol{T}) = 0$ and $\mathcal{R}^{\mathsf{ic}}(\boldsymbol{v}) = 0$, but that a vanishing residual itself does not guarantee that $\big(\mathrm{D}\boldsymbol{v}(x), \boldsymbol{T}(x)\big) \in \mathcal{A}(x)$ for almost every $x \in \Omega$. For this, additionally $\mathcal{E}_{\mathcal{A}}(\mathrm{D}\boldsymbol{v}, \boldsymbol{T}) = 0$ is needed.

For the rest of the paper let t and $\tilde{t}$ be such that

$$\frac{2d}{d+1} < t < r \quad \text{and} \quad \tilde{t} := \frac{1}{2}\frac{dt}{d-t}, \quad \text{if } r \leqslant \frac{3d}{d+2}, \tag{4.3a}$$

$$t = r \quad \text{and} \quad \tilde{t} = t' = \tilde{r} = r', \quad \text{otherwise.} \tag{4.3b}$$

Note that (4.3a) implies that if $r \leqslant \frac{3d}{d+2}$, then $t < r$ and $\tilde{t} < \tilde{r}$.

Lemma 4.2. *The triple $(\boldsymbol{u}, p, \boldsymbol{S}) \in W_0^{1,r}(\Omega)^d \times L_0^{\tilde{r}}(\Omega) \times L^{r'}(\Omega; \mathbb{R}^{d\times d}_{sym})$ is a solution of (2.2) if and only if*

$$\mathcal{R}(\boldsymbol{u}, p, \boldsymbol{S}) = 0 \quad \textit{in } W^{-1,\tilde{t}}(\Omega)^d \times L_0^t(\Omega) \qquad \textit{and} \qquad \mathcal{E}_{\mathcal{A}}(\boldsymbol{u}, \boldsymbol{S}) = 0.$$

Proof. Thanks to the fact, that $W_0^{1,\tilde{t}'}(\Omega)^d \times L^{t'}(\Omega)/\mathbb{R}$ is dense in $W_0^{1,\tilde{r}'}(\Omega)^d \times L^{r'}(\Omega)/\mathbb{R}$ we have that $\mathcal{R}(\boldsymbol{u}, p, \boldsymbol{S}) = 0$ in $W^{-1,\tilde{t}}(\Omega)^d \times L_0^t(\Omega)$ is equivalent to $\mathcal{R}(\boldsymbol{u}, p, \boldsymbol{S}) = 0$ in $W^{-1,\tilde{r}}(\Omega)^d \times L_0^r(\Omega)$. This is, in turn, equivalent to the fact that the triple $(\boldsymbol{u}, p, \boldsymbol{S})$ satisfies the system of partial differential equations (2.2).

On the other hand we have that $(\mathrm{D}\boldsymbol{u}(x), \boldsymbol{S}(x)) \in \mathcal{A}(x)$ for almost every $x \in \Omega$ if and only if $\mathcal{E}_{\mathcal{A}}(\mathrm{D}\boldsymbol{u}, \boldsymbol{S}) = 0$, and that completes the proof. □

Note that Lemma 4.2 does not provide a quantitative relation between the error and the residual. Even for simple r-Laplacian type problems, such a relation requires complicated techniques and problem-adapted error notions (*e.g.* a suitable quasi-norm); *cf.* [7, 18, 31]. However, because of the possible nonuniqueness of solutions to (2.2), such a relation cannot be guaranteed in our situation. We shall therefore restrict the a posteriori analysis to bounding the residual of the problem instead of bounding the error.

Recalling the quasi-interpolation $\Pi_{\mathcal{G}}$ from Assumption 3.5 as well as the representation of the discrete convective term in (3.19), we define the local indicators on $E \in \mathcal{G}$ as follows:

$$\begin{aligned}\mathcal{E}^{\mathsf{pde}}_{\mathcal{G}}\left(\boldsymbol{U}^n_{\mathcal{G}}, P^n_{\mathcal{G}}, \boldsymbol{S}^n\left(\mathrm{D}\boldsymbol{U}^n_{\mathcal{G}}\right); E\right) &:= \left\|h_{\mathcal{G}}\left(-\operatorname{div}\Pi_{\mathcal{G}}\boldsymbol{S}^n\left(\mathrm{D}\boldsymbol{U}^n_{\mathcal{G}}\right) + \mathrm{B}[\boldsymbol{U}^n_{\mathcal{G}}, \boldsymbol{U}^n_{\mathcal{G}}] + \nabla P^n_{\mathcal{G}} - \boldsymbol{f}\right)\right\|^{\tilde{t}}_{\tilde{t},E} \\ &\quad + \left\|h^{1/\tilde{t}}_{\mathcal{G}}\left[\!\left[\Pi_{\mathcal{G}}\boldsymbol{S}^n\left(\mathrm{D}\boldsymbol{U}^n_{\mathcal{G}}\right) - P^n_{\mathcal{G}}\,\mathrm{id}\right]\!\right]\right\|^{\tilde{t}}_{\tilde{t},\partial E} + \left\|\boldsymbol{S}^n\left(\mathrm{D}\boldsymbol{U}^n_{\mathcal{G}}\right) - \Pi_{\mathcal{G}}\boldsymbol{S}^n\left(\mathrm{D}\boldsymbol{U}^n_{\mathcal{G}}\right)\right\|^{\tilde{t}}_{\tilde{t},E},\end{aligned} \tag{4.4a}$$

$$\mathcal{E}^{\mathsf{ic}}_{\mathcal{G}}\left(\boldsymbol{U}^n_{\mathcal{G}}; E\right) := \left\|\operatorname{div}\boldsymbol{U}^n_{\mathcal{G}}\right\|^t_{t,E}, \tag{4.4b}$$

and

$$\mathcal{E}_{\mathcal{G}}\left(\boldsymbol{U}^n_{\mathcal{G}}, P^n_{\mathcal{G}}, \boldsymbol{S}^n\left(\mathrm{D}\boldsymbol{U}^n_{\mathcal{G}}\right); E\right) := \mathcal{E}^{\mathsf{pde}}_{\mathcal{G}}\left(\boldsymbol{U}^n_{\mathcal{G}}, P^n_{\mathcal{G}}, \boldsymbol{S}^n\left(\mathrm{D}\boldsymbol{U}^n_{\mathcal{G}}\right); E\right) + \mathcal{E}^{\mathsf{ic}}_{\mathcal{G}}\left(\boldsymbol{U}^n_{\mathcal{G}}; E\right). \tag{4.4c}$$

Here, for $S \in \mathcal{S}(\mathcal{G})$, $[\![\cdot]\!]|_S$ denotes the normal jump across S and $[\![\cdot]\!]|_{\partial\Omega} := 0$. Moreover, we define the error bounds to be the sums of the local indicators, *i.e.*, for $\mathcal{M} \subset \mathcal{G}$, we have

$$\mathcal{E}^{\mathsf{pde}}_{\mathcal{G}}\big(\boldsymbol{U}^n_{\mathcal{G}}, P^n_{\mathcal{G}}, \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}^n_{\mathcal{G}}); \mathcal{M}\big) := \sum_{E\in\mathcal{M}} \mathcal{E}^{\mathsf{pde}}_{\mathcal{G}}\big(\boldsymbol{U}^n_{\mathcal{G}}, P^n_{\mathcal{G}}, \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}^n_{\mathcal{G}}); E\big),$$

$$\mathcal{E}^{\mathsf{ic}}_{\mathcal{G}}\big(\boldsymbol{U}^n_{\mathcal{G}}; \mathcal{M}\big) := \left\|\operatorname{div}\boldsymbol{U}^n_{\mathcal{G}}\right\|^t_{t;\Omega(\mathcal{M})}$$

and

$$\begin{aligned}\mathcal{E}_{\mathcal{G}}\big(\boldsymbol{U}^n_{\mathcal{G}}, P^n_{\mathcal{G}}, \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}^n_{\mathcal{G}})\big) &:= \mathcal{E}^{\mathsf{pde}}_{\mathcal{G}}\big(\boldsymbol{U}^n_{\mathcal{G}}, P^n_{\mathcal{G}}, \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}^n_{\mathcal{G}})\big) + \mathcal{E}^{\mathsf{ic}}_{\mathcal{G}}\big(\boldsymbol{U}^n_{\mathcal{G}}\big) \\ &:= \mathcal{E}^{\mathsf{pde}}_{\mathcal{G}}\big(\boldsymbol{U}^n_{\mathcal{G}}, P^n_{\mathcal{G}}, \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}^n_{\mathcal{G}}); \mathcal{G}\big) + \mathcal{E}^{\mathsf{ic}}_{\mathcal{G}}\big(\boldsymbol{U}^n_{\mathcal{G}}; \mathcal{G}\big).\end{aligned}$$

Theorem 4.3 (Upper bound on the residual)**.** *Let $n \in \mathbb{N}$ and $\mathcal{G} \in \mathbb{G}$, and denote by $\big(\boldsymbol{U}^n,\, P^n_{\mathcal{G}}\big) \in \mathbb{V}(\mathcal{G}) \times \mathbb{Q}_0(\mathcal{G})$ a Galerkin approximation of* (3.20)*. We then have the following bounds:*

$$\big\|\mathcal{R}^{pde}\big(\boldsymbol{U}^n_{\mathcal{G}}, P^n_{\mathcal{G}}, \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}^n_{\mathcal{G}})\big)\big\|_{W^{-1,\tilde{t}}(\Omega)} \leqslant C_1\, \mathcal{E}^{pde}_{\mathcal{G}}\big(\boldsymbol{U}^n_{\mathcal{G}}, P^n_{\mathcal{G}}, \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}^n_{\mathcal{G}})\big)^{1/\tilde{t}}, \tag{4.5a}$$

$$\sup_{o \in L^{t'}(\Omega)/\mathbb{R}} \Big\langle \mathcal{R}^{ic}\big(\boldsymbol{U}^n_{\mathcal{G}}\big),\, \frac{o}{\inf_{c\in\mathbb{R}} \|o-c\|_{t'}} \Big\rangle = \mathcal{E}^{ic}_{\mathcal{G}}\big(\boldsymbol{U}^n_{\mathcal{G}}\big)^{1/t}. \tag{4.5b}$$

The constant $C_1 > 0$ depends only on the shape-regularity of $\mathcal{G}$, $\tilde{t}$, and on the dimension d.

Proof. The assertions are proved using standard techniques; compare *e.g.* with [2, 42]. For the reader's convenience we sketch the arguments. For arbitrary $(\boldsymbol{v}, q) \in W^{1,\tilde{t}'}_0(\Omega)^d \times L^{t'}(\Omega)/\mathbb{R}$ with $\|\boldsymbol{v}\|_{1,\tilde{t}'} = \|p\|_{t'} = 1$ we deduce from (3.20) that

$$\begin{aligned}\big\langle \mathcal{R}^{\mathsf{pde}}\big(\boldsymbol{U}^n_{\mathcal{G}}, P^n_{\mathcal{G}}, \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}^n_{\mathcal{G}})\big),\, \boldsymbol{v}\big\rangle &= \int_\Omega \Pi_{\mathcal{G}}\boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}^n_{\mathcal{G}}) : \mathrm{D}(\boldsymbol{v} - \mathfrak{I}^{\mathcal{G}}_{\mathrm{div}}\boldsymbol{v}) + \mathrm{B}[\boldsymbol{U}^n_{\mathcal{G}}, \boldsymbol{U}^n_{\mathcal{G}}]\cdot(\boldsymbol{v} - \mathfrak{I}^{\mathcal{G}}_{\mathrm{div}}\boldsymbol{v}) - \boldsymbol{f}\cdot(\boldsymbol{v} - \mathfrak{I}^{\mathcal{G}}_{\mathrm{div}}\boldsymbol{v})\,\mathrm{d}x \\ &\quad - \int_\Omega P^n_{\mathcal{G}}\,\mathrm{div}\,(\boldsymbol{v} - \mathfrak{I}^{\mathcal{G}}_{\mathrm{div}}\boldsymbol{v})\,\mathrm{d}x + \int_\Omega (\boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}^n_{\mathcal{G}}) - \Pi_{\mathcal{G}}\boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}^n_{\mathcal{G}})) : \mathrm{D}(\boldsymbol{v} - \mathfrak{I}^{\mathcal{G}}_{\mathrm{div}}\boldsymbol{v})\,\mathrm{d}x.\end{aligned}$$

Thanks to (3.19), local integration by parts and using Hölder's inequality, we obtain

$$\begin{aligned}\big\langle \mathcal{R}^{\mathsf{pde}}\big(\boldsymbol{U}^n_{\mathcal{G}}, P^n_{\mathcal{G}}, \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}^n_{\mathcal{G}})\big),\, \boldsymbol{v}\big\rangle &\leqslant \sum_{E\in\mathcal{G}} \Big\{ \big\|-\mathrm{div}\,\Pi_{\mathcal{G}}\boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}^n_{\mathcal{G}}) + \mathrm{B}[\boldsymbol{U}^n_{\mathcal{G}}, \boldsymbol{U}^n_{\mathcal{G}}] + \nabla P^n_{\mathcal{G}} - \boldsymbol{f}\big\|_{\tilde{t},E}\, \big\|\boldsymbol{v} - \mathfrak{I}^{\mathcal{G}}_{\mathrm{div}}\boldsymbol{v}\big\|_{\tilde{t}',E} \\ &\qquad + \frac{1}{2}\, \big\|[\![\Pi_{\mathcal{G}}\boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}^n_{\mathcal{G}}) - P^n_{\mathcal{G}}\,\mathrm{id}]\!]\big\|_{\tilde{t},\partial E}\, \big\|\boldsymbol{v} - \mathfrak{I}^{\mathcal{G}}_{\mathrm{div}}\boldsymbol{v}\big\|_{\tilde{t}',\partial E} \\ &\qquad + \big\|\boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}^n_{\mathcal{G}}) - \Pi_{\mathcal{G}}\boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}^n_{\mathcal{G}})\big\|_{\tilde{t},E}\, \|\mathrm{D}\boldsymbol{v}\|_{\tilde{t}',E} \Big\} \\ &\leqslant C \Bigg(\sum_{E\in\mathcal{G}} \Big\{ \big\|h_{\mathcal{G}}\big(-\,\mathrm{div}\,\Pi_{\mathcal{G}}\boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}^n_{\mathcal{G}}) + \mathrm{B}[\boldsymbol{U}^n_{\mathcal{G}}, \boldsymbol{U}^n_{\mathcal{G}}] + \nabla P^n_{\mathcal{G}} - \boldsymbol{f}\big)\big\|^{\tilde{t}}_{\tilde{t},E} \\ &\qquad + \Big\|h^{1/\tilde{t}}_{\mathcal{G}}\, [\![\Pi_{\mathcal{G}}\boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}^n_{\mathcal{G}}) - P^n_{\mathcal{G}}\,\mathrm{id}]\!]\Big\|^{\tilde{t}}_{\tilde{t},\partial E} \\ &\qquad + \big\|\boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}^n_{\mathcal{G}}) - \Pi_{\mathcal{G}}\boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}^n_{\mathcal{G}})\big\|^{\tilde{t}}_{\tilde{t},E} \Big\} \Bigg)^{1/\tilde{t}} \|\boldsymbol{v}\|_{1,\tilde{t}'}\,.\end{aligned}$$

Here, in the last inequality, we used the stability of $\mathfrak{I}^{\mathcal{G}}_{\mathrm{div}}$ (see (3.5)), a scaled trace theorem, and the interpolation estimate for $\mathfrak{I}^{\mathcal{G}}_{\mathrm{div}}$ in (3.7), as well as the finite overlapping of patches and a scaled trace theorem.

To prove the bound (4.5b), we first deduce from $\int_\Omega 1\,\mathrm{div}\,\boldsymbol{U}^n_{\mathcal{G}}\,\mathrm{d}x = 0$ and Hölder's inequality that for all $c \in \mathbb{R}$, we have

$$\int_\Omega o\,\mathrm{div}\,\boldsymbol{U}^n_{\mathcal{G}}\,\mathrm{d}x = \int_\Omega (o-c)\,\mathrm{div}\,\boldsymbol{U}^n_{\mathcal{G}}\,\mathrm{d}x \leqslant \big\|\,\mathrm{div}\,\boldsymbol{U}^n_{\mathcal{G}}\big\|_t\, \|o-c\|_{t'}\,.$$

Taking the infimum over all $c \in \mathbb{R}$ and then the supremum over all $o \in L^{r'}(\Omega)$ proves '$\leqslant$' in (4.5b). In order to prove '$\geqslant$', we observe that

$$\mathcal{E}^{\mathsf{ic}}_{\mathcal{G}}\big(\boldsymbol{U}^n_{\mathcal{G}}\big) = \int_\Omega \mathrm{div}\,\boldsymbol{U}^n_{\mathcal{G}}\,|\,\mathrm{div}\,\boldsymbol{U}^n_{\mathcal{G}}|^{t-2}\,\mathrm{div}\,\boldsymbol{U}^n_{\mathcal{G}}\,\mathrm{d}x \leqslant \sup_{o\in L^{t'}(\Omega)/\mathbb{R}} \Big\langle \mathcal{R}^{\mathsf{ic}}\big(\boldsymbol{U}^n_{\mathcal{G}}\big),\, \frac{o}{\inf_{c\in\mathbb{R}}\|o-c\|_{t'}} \Big\rangle \big\||\,\mathrm{div}\,\boldsymbol{U}^n_{\mathcal{G}}|^{t-2}\,\mathrm{div}\,\boldsymbol{U}^n_{\mathcal{G}}\big\|_{t'}\,.$$

Together with the definition of $\mathcal{E}_{\mathcal{G}}^{\mathsf{ic}}(\boldsymbol{U}_{\mathcal{G}}^n)$ and noting that

$$\big\||\operatorname{div}\boldsymbol{U}_{\mathcal{G}}^n|^{t-2}\operatorname{div}\boldsymbol{U}_{\mathcal{G}}^n\big\|_{t'} = \big\|\operatorname{div}\boldsymbol{U}_{\mathcal{G}}^n\big\|_t^{t-1} = \mathcal{E}_{\mathcal{G}}^{\mathsf{ic}}\big(\boldsymbol{U}_{\mathcal{G}}^n\big)^{1-\frac{1}{t}} = \mathcal{E}_{\mathcal{G}}^{\mathsf{ic}}\big(\boldsymbol{U}_{\mathcal{G}}^n\big)^{\frac{1}{t'}},$$

this yields (4.5b). □

Corollary 4.4. *Under the conditions of Theorem 4.3, we have*

$$\left\langle \mathcal{R}^{\mathsf{pde}}\big(\boldsymbol{U}_{\mathcal{G}}^n, P_{\mathcal{G}}^n, \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n)\big),\, \boldsymbol{v}\right\rangle \leqslant C_1 \sum_{E\in\mathcal{G}} \mathcal{E}_{\mathcal{G}}^{\mathsf{pde}}\big(\boldsymbol{U}_{\mathcal{G}}^n, P_{\mathcal{G}}^n, \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n); E\big)^{1/\tilde{t}} \,\|\nabla \boldsymbol{v}\|_{\tilde{t}',\mathcal{U}^{\mathcal{G}}(E)}$$

and

$$\left\langle \mathcal{R}^{\mathsf{ic}}\big(\boldsymbol{U}_{\mathcal{G}}^n\big),\, q\right\rangle \leqslant \sum_{E\in\mathcal{G}} \mathcal{E}_{\mathcal{G}}^{\mathsf{ic}}\big(\boldsymbol{U}_{\mathcal{G}}^n; E\big)^{1/t} \,\|q\|_{t',\mathcal{U}^{\mathcal{G}}(E)}$$

for all $\boldsymbol{v} \in W_0^{1,\tilde{t}'}(\Omega)^d$ *and* $q \in L^{t'}(\Omega)$.

Theorem 4.5 (Lower bound on the residual)**.** *Under the conditions of Theorem 4.3, we have*

$$c_2\, \mathcal{E}_{\mathcal{G}}^{\mathsf{pde}}\big(\boldsymbol{U}_{\mathcal{G}}^n, P_{\mathcal{G}}^n, \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n)\big)^{1/\tilde{t}} \leqslant \big\|\mathcal{R}^{\mathsf{pde}}\big(\boldsymbol{U}_{\mathcal{G}}^n, P_{\mathcal{G}}^n, \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n)\big)\big\|_{W^{-1,\tilde{t}}(\Omega)} + \operatorname{osc}_{\mathcal{G}}\big(\boldsymbol{U}_{\mathcal{G}}^n, \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n)\big)^{1/\tilde{t}}. \tag{4.6}$$

The constant $c_2 > 0$ *depends solely on the shape-regularity of* $\mathcal{G}$, $\tilde{t}$, *and on the dimension* d. *The oscillation term is defined by*

$$\begin{aligned}\operatorname{osc}_{\mathcal{G}}\big(\boldsymbol{U}_{\mathcal{G}}^n, \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n)\big) &:= \sum_{E\in\mathcal{G}} \operatorname{osc}\big(\boldsymbol{U}_{\mathcal{G}}^n, \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n), E\big)\\ &:= \sum_{E\in\mathcal{G}} \min_{\boldsymbol{f}_E\in\mathbb{P}_{2\ell-1}^d} \big\|h_{\mathcal{G}}\big(\boldsymbol{f}-\boldsymbol{f}_E\big)\big\|_{\tilde{t},E}^{\tilde{t}} + \big\|\boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n) - \Pi_{\mathcal{G}}\boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n)\big\|_{\tilde{t},E}^{\tilde{t}}.\end{aligned}$$

Proof. Let $E \in \mathcal{G}$ and let $S \in \mathcal{S}(\mathcal{G})$, *i.e.*, there exist $E_1, E_2 \in \mathcal{G}$, $E_1 \neq E_2$, such that $S = E_1 \cap E_2$. Let $\boldsymbol{f}_E \in \mathbb{P}_{2\ell-1}^d$ be arbitrary; for convenience we use the notation

$$R_E := -\operatorname{div}\Pi_{\mathcal{G}}\boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n) + \mathrm{B}[\boldsymbol{U}_{\mathcal{G}}^n, \boldsymbol{U}_{\mathcal{G}}^n] + \nabla P_{\mathcal{G}}^n - \boldsymbol{f}_E \in \mathbb{P}_{2\ell-1}^d,$$

and

$$J_S := \big[\!\big[\Pi_{\mathcal{G}}\boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n) - P_{\mathcal{G}}^n\,\mathrm{id}\big]\!\big]\,|_S \in \mathbb{P}_m^{d\times d}, \quad \text{where } m = \max\{\ell-1, \jmath\}.$$

It is well-known that there exist local bubble functions $b_E, b_S \in W_0^{1,\infty}(\Omega)$, such that

$$0 \leqslant b_E, b_S \leqslant 1, \qquad \operatorname{supp} b_E = E \qquad \text{and} \qquad \operatorname{supp} b_S = \omega_S := E_1 \cup E_2. \tag{4.7a}$$

Moreover, we have that there exist $\rho_E \in \mathbb{P}_{2\ell-1}^d$ and $\rho_S \in \mathbb{P}_m^{d\times d}$, with $\|\rho_E\|_{\tilde{t}',E} = 1 = \|\rho_S\|_{\tilde{t}',S}$, such that

$$\begin{aligned}&\|R_E\|_{\tilde{t},E} \leqslant C\int_E R_E b_E\rho_E\,\mathrm{d}x, \qquad \|\nabla(b_E\rho_E)\|_{\tilde{t}',E} \leqslant C\left\|h_{\mathcal{G}}^{-1}\rho_E\right\|_{\tilde{t}',E},\\ &\|J_S\|_{\tilde{t},S} \leqslant C\int_S J_S b_S\rho_S\,\mathrm{d}x, \qquad \|\nabla(b_S\rho_S)\|_{\tilde{t}',\omega_S} \leqslant C\left\|h_{\mathcal{G}}^{-1/\tilde{t}'}\rho_S\right\|_{\tilde{t}',S},\\ &\qquad\text{and} \qquad \|b_S\rho_S\|_{\tilde{t}',\omega_S} \leqslant C\left\|h_{\mathcal{G}}^{1/\tilde{t}'}\rho_S\right\|_{\tilde{t}',S};\end{aligned} \tag{4.7b}$$

compare, for example, with ([43], Chap. 3.6). Here the constants only depend on r, the polynomial degree of R_E, respectively J_S, on the shape-regularity of $\mathcal{G}$, and on the dimension d. Hence, for the element residual, we deduce that

$$\|R_E\|_{\tilde{t},E} \leqslant C\left\{\left\langle \mathcal{R}^{\mathsf{pde}}(\boldsymbol{U}_{\mathcal{G}}^n, P_{\mathcal{G}}^n, \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n)),\, \mathscr{b}_E\rho_E\right\rangle + \left\langle \varPi_{\mathcal{G}}\boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n) - \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n),\, \boldsymbol{D}(\mathscr{b}_E\rho_E)\right\rangle + \|\boldsymbol{f} - \boldsymbol{f}_E\|_{\tilde{t},E}\right\},$$

where we have used Hölder's inequality and that $0 \leqslant \mathscr{b}_E \leqslant 1$. Together with a triangle inequality and (4.7b), this implies that

$$\begin{aligned}\left\|h_{\mathcal{G}}\big(-\operatorname{div}\varPi_{\mathcal{G}}\boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n) + \mathrm{B}[\boldsymbol{U}_{\mathcal{G}}^n, \boldsymbol{U}_{\mathcal{G}}^n] + \nabla P_{\mathcal{G}}^n - f\big)\right\|_{\tilde{t},E} &\leqslant C\left\{\left\langle \mathcal{R}^{\mathsf{pde}}(\boldsymbol{U}_{\mathcal{G}}^n, P_{\mathcal{G}}^n, \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n)),\, h_{\mathcal{G}}\mathscr{b}_E\rho_E\right\rangle\right.\\ &\quad\left. + \left\|\varPi_{\mathcal{G}}\boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n) - \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n)\right\|_{\tilde{t},E} + \|h_{\mathcal{G}}(\boldsymbol{f} - \boldsymbol{f}_E)\|_{\tilde{t},E}\right\}.\end{aligned} \tag{4.8}$$

For the jump residual, we deduce from (4.7a) and integration by parts, that

$$\begin{aligned}\|J_S\|_{\tilde{t},S} &\leqslant C\int_S \left[\!\left[\varPi_{\mathcal{G}}\boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n) - P_{\mathcal{G}}^n\,\mathrm{id}\right]\!\right] \mathscr{b}_S\rho_S\,\mathrm{d}s\\ &= C\Big\{\left\langle \mathcal{R}^{\mathsf{pde}}(\boldsymbol{U}_{\mathcal{G}}^n, P_{\mathcal{G}}^n, \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n)),\, \mathscr{b}_S\rho_S\right\rangle\\ &\quad + \sum_{i=1,2}\int_{E_i}\big(\operatorname{div}\varPi_{\mathcal{G}}\boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n) - \mathrm{B}[\boldsymbol{U}_{\mathcal{G}}^n, \boldsymbol{U}_{\mathcal{G}}^n] - \nabla P_{\mathcal{G}}^n + \boldsymbol{f}\big)\mathscr{b}_S\rho_S\,\mathrm{d}x\Big\}.\end{aligned}$$

Therefore, we obtain, with (4.7b), Hölder's inequality and (4.8), that

$$\begin{aligned}\left\|h_{\mathcal{G}}^{1/\tilde{t}}\left[\!\left[\boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n) - P_{\mathcal{G}}^n\,\mathrm{id}\right]\!\right]\right\|_{\tilde{t},S} &\leqslant C\Big\{\left\langle \mathcal{R}^{\mathsf{pde}}(\boldsymbol{U}_{\mathcal{G}}^n, P_{\mathcal{G}}^n, \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n)),\, h_S^{1/\tilde{t}}\mathscr{b}_S\rho_S\right\rangle\\ &\quad + \sum_{i=1,2}\Big[\left\langle \mathcal{R}^{\mathsf{pde}}(\boldsymbol{U}_{\mathcal{G}}^n, P_{\mathcal{G}}^n, \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n)),\, h_{\mathcal{G}}\mathscr{b}_{E_i}\rho_{E_i}\right\rangle\\ &\quad + \left\|\varPi_{\mathcal{G}}\boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n) - \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n)\right\|_{\tilde{t},E_i} + \left\|h_{\mathcal{G}}(\boldsymbol{f} - \boldsymbol{f}_{E_i})\right\|_{\tilde{t},E_i}\Big]\Big\}.\end{aligned} \tag{4.9}$$

We define the constants $\alpha_E := \left\|h_{\mathcal{G}}\big(-\operatorname{div}\varPi_{\mathcal{G}}\boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n) + \mathrm{B}[\boldsymbol{U}_{\mathcal{G}}^n, \boldsymbol{U}_{\mathcal{G}}^n] + \nabla P_{\mathcal{G}}^n - \boldsymbol{f}\big)\right\|_{\tilde{t},E}^{\tilde{t}-1}$, $E \in \mathcal{G}$, and $\beta_S := \left\|h_{\mathcal{G}}^{1/\tilde{t}}\left[\!\left[\varPi_{\mathcal{G}}\boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n) - P_{\mathcal{G}}^n\,\mathrm{id}\right]\!\right]\right\|_{\tilde{t},S}^{\tilde{t}-1}$, $S \in \mathcal{S}(\mathcal{G})$. Then, combining (4.8) and (4.9) and summing over all $E \in \mathcal{G}$, $S \in \mathcal{S}(\mathcal{G})$, yields

$$\begin{aligned}\mathcal{E}_{\mathcal{G}}^{\mathsf{pde}}\big(\boldsymbol{U}_{\mathcal{G}}^n, P_{\mathcal{G}}^n, \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n)\big) &= \sum_{E\in\mathcal{G}}\alpha_E\left\|h_{\mathcal{G}}\big(-\operatorname{div}\varPi_{\mathcal{G}}\boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n) + \mathrm{B}[\boldsymbol{U}_{\mathcal{G}}^n, \boldsymbol{U}_{\mathcal{G}}^n] + \nabla P_{\mathcal{G}}^n - \boldsymbol{f}\big)\right\|_{\tilde{t},E}\\ &\quad + \sum_{s\in\mathcal{S}(\mathcal{G})}\beta_S\left\|h_{\mathcal{G}}^{1/\tilde{t}}\left[\!\left[\boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n) - P_{\mathcal{G}}^n\,\mathrm{id}\right]\!\right]\right\|_{\tilde{t},S}\\ &\leqslant C\left\{\left\langle \mathcal{R}^{\mathsf{pde}}(\boldsymbol{U}_{\mathcal{G}}^n, P_{\mathcal{G}}^n, \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n)),\, \sum_{E\in\mathcal{G}}\left(\alpha_E + \sum_{S\subset\partial E\cap\Omega}\beta_S\right)h_{\mathcal{G}}\mathscr{b}_E\rho_E\right\rangle\right.\\ &\quad + \left\langle \mathcal{R}^{\mathsf{pde}}(\boldsymbol{U}_{\mathcal{G}}^n, P_{\mathcal{G}}^n, \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n)),\, \sum_{S\in\mathcal{S}(\mathcal{G})}\beta_S h_S^{1/\tilde{t}}\mathscr{b}_S\rho_S\right\rangle\\ &\quad\left. + \sum_{E\in\mathcal{G}}\left(\alpha_E + \sum_{S\subset\partial E\cap\Omega}\beta_S\right)\operatorname{osc}\big(\boldsymbol{U}_{\mathcal{G}}^n, \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n), E\big)^{1/\tilde{t}}\right\}.\end{aligned}$$

Here we have used in the last step that $\boldsymbol{f}_E \in \mathbb{P}_{2\ell-1}^d$, with $E \in \mathcal{G}$, are arbitrary.

Thanks to the fact that the supp$\mathscr{b}_E$, $E \in \mathcal{G}$, are mutually disjoint up to a null-set, together with (4.7b), we have that

$$\Big\| \sum_{E\in\mathcal{G}} \Big(\alpha_E + \sum_{S\subset\partial E\cap\Omega} \beta_S \Big) \nabla(h_{\mathcal{G}} \mathscr{b}_E \rho_E) \Big\|_{\tilde{t}'}^{\tilde{t}'} = \sum_E \Big(\alpha_E + \sum_{S\subset\partial E\cap\Omega} \beta_S \Big)^{\tilde{t}'} \int_E \big| h_{\mathcal{G}} \nabla(\mathscr{b}_E \rho_E) \big|^{\tilde{t}'} \,\mathrm{d}x$$
$$\leqslant C \left(\sum_E \alpha_E^{\tilde{t}'} + \sum_{S\in\mathcal{S}(\mathcal{G})} \beta_S^{\tilde{t}'} \right) \leqslant C\, \mathcal{E}_{\mathcal{G}}^{\mathsf{pde}}\big(\boldsymbol{U}_{\mathcal{G}}^n, P_{\mathcal{G}}^n, \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n)\big)^{1/\tilde{t}'},$$

where we have used that each element E has at most $(d+1)$ sides $S \in \mathcal{S}(\mathcal{G})$ with $S \subset \partial E$. The constants C depend only on the shape-regularity of $\mathcal{G}$. Analogously, we deduce from the fact that only finitely many of the supp$\mathscr{b}_S$, $S \in \mathcal{S}(\mathcal{G})$, overlap, that

$$\Big\| \nabla \Big(\sum_{S\in\mathcal{S}(\mathcal{G})} \beta_S h_{\mathcal{G}}^{1/\tilde{t}} \mathscr{b}_S \rho_S \Big) \Big\|_{\tilde{t}'}^{\tilde{t}'} \leqslant C \sum_{S\in\mathcal{S}(\mathcal{G})} \beta^{\tilde{t}'} \int_{\omega_S} \big| h_{\mathcal{G}}^{1/\tilde{t}} \nabla \mathscr{b}_S \rho_S \big|^{\tilde{t}'} \leqslant C \sum_{S\in\mathcal{S}(\mathcal{G})} \beta^{\tilde{t}'} \leqslant C \mathcal{E}_{\mathcal{G}}^{\mathsf{pde}}\big(\boldsymbol{U}_{\mathcal{G}}^n, P_{\mathcal{G}}^n, \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n)\big)^{1/\tilde{t}'}.$$

Combining Hölder's inequality with similar arguments yields for the last term

$$\sum_{E\in\mathcal{G}} \Big(\alpha_E + \sum_{S\subset\partial E\cap\Omega} \beta_S \Big) \operatorname{osc}\big(\boldsymbol{U}_{\mathcal{G}}^n, \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n); E\big)^{1/\tilde{t}} \leqslant C \left(\sum_{E\in\mathcal{G}} \alpha_E^{\tilde{t}'} + \sum_{S\in\mathcal{S}(\mathcal{G})} \beta_S^{\tilde{t}'} \right)^{1/\tilde{t}'} \operatorname{osc}(\boldsymbol{U}_{\mathcal{G}}^n, \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n))^{1/\tilde{t}}$$
$$\leqslant C\, \mathcal{E}_{\mathcal{G}}^{\mathsf{pde}}\big(\boldsymbol{U}_{\mathcal{G}}^n, P_{\mathcal{G}}^n, \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n)\big)^{1/\tilde{t}'} \operatorname{osc}(\boldsymbol{U}_{\mathcal{G}}^n, \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n))^{1/\tilde{t}}.$$

Altogether, we have thus proved that

$$\mathcal{E}_{\mathcal{G}}^{\mathsf{pde}}\big(\boldsymbol{U}_{\mathcal{G}}^n, P_{\mathcal{G}}^n, \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n)\big) \leqslant C \Big\{ \big\| \mathcal{R}^{\mathsf{pde}}(\boldsymbol{U}_{\mathcal{G}}^n, P_{\mathcal{G}}^n, \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n)) \big\|_{W^{-1,\tilde{t}'}(\Omega)} \mathcal{E}_{\mathcal{G}}^{\mathsf{pde}}\big(\boldsymbol{U}_{\mathcal{G}}^n, P_{\mathcal{G}}^n, \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n)\big)^{1/\tilde{t}'}$$
$$+ \operatorname{osc}(\boldsymbol{U}_{\mathcal{G}}^n, \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n))\, \mathcal{E}_{\mathcal{G}}^{\mathsf{pde}}\big(\boldsymbol{U}_{\mathcal{G}}^n, P_{\mathcal{G}}^n, \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n)\big)^{1/\tilde{t}'} \Big\}.$$

This is the desired bound. ∎

The following result states the local stability of the error bound and is referred to as *local lower bound* in the context of linear elliptic problems.

Corollary 4.6 (Local stability)**.** *Suppose the conditions of Theorem 4.3 and let $\mathcal{M} \subset \mathcal{G}$; then, there exists a constant C, depending solely on the shape-regularity of $\mathcal{G}$, $\tilde{t}$, d and Ω, such that*

$$\mathcal{E}_{\mathcal{G}}^{\mathsf{pde}}\big(\boldsymbol{U}_{\mathcal{G}}^n, P_{\mathcal{G}}^n, \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n); \mathcal{M}\big)^{1/\tilde{t}} \leqslant C \left(\big\| \mathcal{R}^{\mathsf{pde}}(\boldsymbol{U}_{\mathcal{G}}^n, P_{\mathcal{G}}^n, \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n)) \big\|_{W^{-1,\tilde{t}}(\mathcal{U}^{\mathcal{G}}(\mathcal{M}))} + \operatorname{osc}(\boldsymbol{U}_{\mathcal{G}}^n, \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n); \mathcal{M})^{1/\tilde{t}} \right)$$
$$\leqslant C \left(\|\boldsymbol{U}_{\mathcal{G}}^n\|_{1,t;\mathcal{U}^{\mathcal{G}}(\mathcal{M})} + \|\boldsymbol{U}_{\mathcal{G}}^n\|_{1,t;\mathcal{U}^{\mathcal{G}}(\mathcal{M})}^2 + \|P_{\mathcal{G}}^n\|_{\tilde{t};\mathcal{U}^{\mathcal{G}}(\mathcal{M})} + \|\boldsymbol{f}\|_{\tilde{t},\mathcal{U}^{\mathcal{G}}(\mathcal{M})} + \big\|\tilde{k}\big\|_{r,\mathcal{U}^{\mathcal{G}}(\mathcal{M})} \right).$$

Proof. The first bound follows as in the proof of Theorem 4.5. In order to prove the second bound, let $\boldsymbol{v} \in W_0^{1,\tilde{t}'}(\mathcal{U}^{\mathcal{G}}(\mathcal{M}))^d$. We then have with Hölder's inequality and (3.18), with t and $\tilde{t}$ instead of r and $\tilde{r}$, that

$$\big\langle \mathcal{R}^{\mathsf{pde}}(\boldsymbol{U}_{\mathcal{G}}^n, P_{\mathcal{G}}^n, \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n)), \boldsymbol{v} \big\rangle = \int_\Omega \Pi_{\mathcal{G}} \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n) : \mathrm{D}\boldsymbol{v} + \mathrm{B}[\boldsymbol{U}_{\mathcal{G}}^n, \boldsymbol{U}_{\mathcal{G}}^n] \cdot \boldsymbol{v} - P_{\mathcal{G}}^n \operatorname{div} \boldsymbol{v} - \boldsymbol{f} \cdot \boldsymbol{v} \,\mathrm{d}x$$
$$\leqslant C \Big(\|\Pi_{\mathcal{G}} \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n)\|_{\tilde{t};\mathcal{U}^{\mathcal{G}}(\mathcal{M})} + \|\boldsymbol{U}_{\mathcal{G}}^n\|_{1,t;\mathcal{U}^{\mathcal{G}}(\mathcal{M})}^2 + \|P_{\mathcal{G}}^n\|_{\tilde{t};\mathcal{U}^{\mathcal{G}}(\mathcal{M})} + \|\boldsymbol{f}\|_{\tilde{t};\mathcal{U}^{\mathcal{G}}(\mathcal{M})} \Big)$$
$$\times \|\boldsymbol{v}\|_{1,\tilde{t}';\mathcal{U}^{\mathcal{G}}(\mathcal{M})}.$$

Note that in the case of (4.3a), we have

$$\tilde{t} = \frac{1}{2}\frac{\mathrm{d}t}{d-t} < \frac{1}{2}\frac{\mathrm{d}t}{d-r} = \tilde{r}\frac{t}{r} \leqslant r'\frac{t}{r} = \frac{t}{r-1} =: s'.$$

Hence, Hölder's inequality, the stability of $\varPi_{\mathcal{G}}$ (Assump. 3.5) and Assumption 3.2 yield

$$\begin{aligned}\left\|\varPi_{\mathcal{G}}\boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}^n_{\mathcal{G}})\right\|_{\tilde{t};\mathcal{U}^{\mathcal{G}}(\mathcal{M})} &\leqslant \left|\mathcal{U}^{\mathcal{G}}(\mathcal{M})\right|^{\frac{s'-\tilde{t}}{s'\tilde{t}}}\left\|\varPi_{\mathcal{G}}\boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}^n_{\mathcal{G}})\right\|_{s';\mathcal{U}^{\mathcal{G}}(\mathcal{M})} \leqslant |\varOmega|^{\frac{s'-\tilde{t}}{s'\tilde{t}}}\left\|\varPi_{\mathcal{G}}\boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}^n_{\mathcal{G}})\right\|_{s';\mathcal{U}^{\mathcal{G}}(\mathcal{M})}\\ &\leqslant C\left(\left\|\mathrm{D}\boldsymbol{U}^n_{\mathcal{G}}\right\|_{t;\mathcal{U}^{\mathcal{G}}(\mathcal{M})} + \left\|\tilde{k}\right\|_{r';\mathcal{U}^{\mathcal{G}}(\mathcal{M})}\right).\end{aligned}$$

The oscillation term can be bounded above similarly, and the assertions follows. □

Remark 4.7. Corollary 4.6 states the stability properties of the estimator, which are required in order to apply the convergence theory in [34, 38]; compare with ([38], Eq. (2.10b)), for example. The stability of the estimator is also of importance for the efficiency of the estimator. If Corollary 4.6 fails to hold, it may happen that the *a posteriori* error estimator is unbounded even though the sequence of discrete solutions is convergent; in particular, $\operatorname{div}\boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}^n_{\mathcal{G}})$ need not belong to $L^{r'}(E)$ when $1 < r < 2$. This problem already appears in the *a posteriori* analysis of quadratic finite element approximations of the r-Laplacian, or the r-Stokes problem (*cf.* [8]) for $1 < r < 2$. In order to avoid this, we use $\varPi_{\mathcal{G}}\boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}^n_{\mathcal{G}})$ instead of $\boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}^n_{\mathcal{G}})$ in the element residual (4.4a). This is compensated by the term $\left\|\boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}^n_{\mathcal{G}}) - \varPi_{\mathcal{G}}\boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}^n_{\mathcal{G}})\right\|_{\tilde{t}}^{\tilde{t}}$ in the *a posteriori* bounds (4.5a) and (4.6); *cf.* the Appendix in [30] for further details.

5. Convergent adaptive finite elements

This section is concerned with the proof of convergence of an adaptive finite element algorithm for the implicit constitutive model under consideration.

5.1. The adaptive finite element method (AFEM)

In this section, we shall introduce an adaptive finite element method for (2.2).

Algorithm 5.1 (AFEM)**.**

Let $k = 0$, $n_0 = 1$, and let $\mathcal{G}_0$ be a given partition of $\varOmega$.

loop
  let $\boldsymbol{S}_k = \boldsymbol{S}^{n_k}$.
  $(\boldsymbol{U}_k, P_k, \boldsymbol{S}_k(\cdot, \mathrm{D}\boldsymbol{U}_k)) = \texttt{SOLVE}(n_k, \mathcal{G}_k)$
  compute $\left\{\mathcal{E}_{\mathcal{G}_k}\left(\boldsymbol{U}_k, P_k, \boldsymbol{S}_k(\cdot, \mathrm{D}\boldsymbol{U}_k); E\right)\right\}_{E\in\mathcal{G}_k}$, and $\mathcal{E}_{\mathcal{A}}(\mathrm{D}\boldsymbol{U}_k, \boldsymbol{S}_k(\cdot, \mathrm{D}\boldsymbol{U}_k))$
  if $\mathcal{E}_{\mathcal{G}_k}\left(\boldsymbol{U}_k, P_k, \boldsymbol{S}_k(\cdot, \mathrm{D}\boldsymbol{U}_k)\right) \geqslant \mathcal{E}_{\mathcal{A}}(\mathrm{D}\boldsymbol{U}_k, \boldsymbol{S}_k(\cdot, \mathrm{D}\boldsymbol{U}_k))$ **then**
    $\mathcal{M}_k = \texttt{MARK}\left(\left\{\mathcal{E}_{\mathcal{G}_k}\left(\boldsymbol{U}_k, P_k, \boldsymbol{S}_k(\cdot, \mathrm{D}\boldsymbol{U}_k); E\right)\right\}_{E\in\mathcal{G}_k}, \mathcal{G}_k\right)$
    $\mathcal{G}_{k+1} = \texttt{REFINE}\left(\mathcal{G}_k, \mathcal{M}_k\right)$ % mesh-refinement
    $n_{k+1} = n_k$
  else
    $n_{k+1} = n_k + 1$ % graph-refinement
  end if
  $k = k + 1$
end loop

The details of the subroutines used in the process are listed below:
The routine `SOLVE`. We assume that for arbitrary $n \in \mathbb{N}$, $\mathcal{G} \in \mathbb{G}$, the routine $\texttt{SOLVE}(n,\mathcal{G}) = (\boldsymbol{U}_{\mathcal{G}}^n, P_n, \boldsymbol{S}^n(\cdot, \mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n))$ computes an exact solution $(\boldsymbol{U}_{\mathcal{G}}^n, P_n) \in \mathbb{V}(\mathcal{G}) \times \mathbb{Q}(\mathcal{G})$ of (3.20).
The routine `MARK`. For a fixed function $g : \mathbb{R}_0^+ \to \mathbb{R}_0^+$, which is continuous at 0 with $g(0) = 0$, we assume that the set $\mathcal{M} = \texttt{MARK}\big(\{\mathcal{E}_{\mathcal{G}}\big(\boldsymbol{U}_{\mathcal{G}}^n, P_{\mathcal{G}}^n, \boldsymbol{S}^n(\mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n); E\big)\}_{E\in\mathcal{G}}, \mathcal{G}\big)$ satisfies

$$\max\big\{\mathcal{E}_{\mathcal{G}}\big(\boldsymbol{U}_{\mathcal{G}}^n, P_{\mathcal{G}}^n, \boldsymbol{S}^n(\cdot, \mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n); E\big) : E \in \mathcal{G}\backslash\mathcal{M}\big\} \leqslant g\big(\max\big\{\mathcal{E}_{\mathcal{G}}\big(\boldsymbol{U}_{\mathcal{G}}^n, P_{\mathcal{G}}^n, \boldsymbol{S}^n(\cdot, \mathrm{D}\boldsymbol{U}_{\mathcal{G}}^n); E\big) : E \in \mathcal{M}\big\}\big). \tag{5.1}$$

Hence the marking criterion guarantees that all indicators in $\mathcal{G}$ are controlled by the maximal indicator in $\mathcal{M}$. Note that this criterion covers most commonly used marking strategies with $g(s) = s$; *cf.* [34, 38].

For the definition of **the routine** `REFINE` see Section 3.2.

For the sake of simplicity of the presentation, in the following, we will suppress the dependence on x in our notation and write $\boldsymbol{S}_k(\mathrm{D}\boldsymbol{U}_k) = \boldsymbol{S}_k(\cdot, \mathrm{D}\boldsymbol{U}_k)$ if there is no risk of confusion.

5.2. Convergence of the AFEM

Let $\{\mathcal{G}_k\}_{k\in\mathbb{N}} \subset \mathbb{G}$ be the sequence of meshes produced by AFEM. For $s \in (1,\infty]$, we define

$$\mathbb{V}_\infty^s := \overline{\bigcup_{k\geqslant 0} \mathbb{V}(\mathcal{G}_k)}^{\|\cdot\|_{1,s}} \subset W_0^{1,s}(\Omega)^d \quad \text{and} \quad \mathbb{Q}_\infty^s := \overline{\bigcup_{k\geqslant 0} \mathbb{Q}(\mathcal{G}_k)}^{\|\cdot\|_{s}} \subset L_0^s(\Omega). \tag{5.2}$$

Lemma 5.2. *Let* $\big\{\big(\boldsymbol{U}_k, P_k, \boldsymbol{S}_k(\mathrm{D}\boldsymbol{U}_k)\big)\big\}_{k\in\mathbb{N}} \subset W_0^r(\Omega)^d \times L_0^{\tilde{r}}(\Omega) \times L^{r'}(\Omega; \mathbb{R}^{d\times d}_{sym})$ *be the sequence produced by AFEM; then, at least for a not relabelled subsequence, we have*

$$\begin{aligned} \boldsymbol{U}_k &\rightharpoonup \boldsymbol{u}_\infty && \textit{weakly in } W_0^{1,r}(\Omega)^d,\\ P_k &\rightharpoonup p_\infty && \textit{weakly in } L_0^{\tilde{r}}(\Omega),\\ \boldsymbol{S}_k(\mathrm{D}\boldsymbol{U}_k) &\rightharpoonup \boldsymbol{S}_\infty && \textit{weakly in } L^{r'}(\Omega; \mathbb{R}^{d\times d}_{sym}), \end{aligned}$$

for some $(\boldsymbol{u}_\infty, p_\infty, \boldsymbol{S}_\infty) \in \mathbb{V}_\infty^r \times \mathbb{Q}_\infty^{\tilde{r}} \times L_0^{r'}(\Omega)$. *Moreover, we have that*

$$\mathcal{R}(\boldsymbol{U}_k, P_k, \boldsymbol{S}_k(\mathrm{D}\boldsymbol{U}_k)) \rightharpoonup^* \mathcal{R}(\boldsymbol{u}_\infty, p_\infty, \boldsymbol{S}_\infty) \qquad \textit{weakly* in } W^{-1,\tilde{r}}(\Omega)^d$$

and

$$\big\langle \mathcal{R}(\boldsymbol{u}_\infty, p_\infty, \boldsymbol{S}_\infty), (\boldsymbol{v}, q)\big\rangle = 0 \qquad \textit{for all } q \in \mathbb{Q}_\infty^{r'}, \boldsymbol{v} \in \mathbb{V}_\infty^{\tilde{r}'}.$$

Proof. The proof is postponed to Section 6.1. ☐

Corollary 5.3. *Let* $\big\{\big(\boldsymbol{U}_k, P_k, \boldsymbol{S}_k(\mathrm{D}\boldsymbol{U}_k)\big)\big\}_{k\in\mathbb{N}} \subset W_0^r(\Omega)^d \times L_0^{\tilde{r}}(\Omega) \times L^{r'}(\Omega; \mathbb{R}^{d\times d}_{sym})$ *be a not relabelled subsequence with weak limit* $(\boldsymbol{u}_\infty, p_\infty, \boldsymbol{S}_\infty) \in \mathbb{V}_\infty^r \times \mathbb{Q}_\infty^{\tilde{r}} \times L^{r'}(\Omega; \mathbb{R}^{d\times d}_{sym})$ *as in Lemma* 5.2. *Then,*

$$\mathcal{E}_{\mathcal{G}_k}(\boldsymbol{U}_k, P_k, \boldsymbol{S}_k(\mathrm{D}\boldsymbol{U}_k)) \to 0, \quad \textit{as } k \to \infty;$$

implies that

$$\mathcal{R}(\boldsymbol{u}_\infty, p_\infty, \boldsymbol{S}_\infty) = 0 \in W^{-1,\tilde{r}}(\Omega)^d.$$

Proof. The upper bound, Theorem 4.3, together with $\mathcal{E}_{\mathcal{G}_k}(\boldsymbol{U}_k, P_k, \boldsymbol{S}_k(\mathrm{D}\boldsymbol{U}_k)) \to 0$ as $k \to \infty$, implies that

$$\mathcal{R}(\boldsymbol{U}_k, P_k, \boldsymbol{S}_k(\mathrm{D}\boldsymbol{U}_k)) \to 0 \quad \text{strongly in } W^{-1,\tilde{t}}(\Omega)^d.$$

Thus the assertion follows from Lemma 5.2 and the uniqueness of the limit. ☐

Lemma 5.4. *Let* $\{(\boldsymbol{U}_k, P_k, \boldsymbol{S}_k(\mathrm{D}\boldsymbol{U}_k))\}_{k\in\mathbb{N}} \subset W_0^r(\Omega)^d \times L_0^{\tilde{r}}(\Omega) \times L^{r'}(\Omega; \mathbb{R}^{d\times d}_{sym})$ *be a not relabelled subsequence with weak limit* $(\boldsymbol{u}_\infty, p_\infty, \boldsymbol{S}_\infty) \in \mathbb{V}^r_\infty \times \mathbb{Q}^{\tilde{r}}_\infty \times L^r(\Omega; \mathbb{R}^{d\times d}_{sym})$ *as in Lemma* 5.2. *Assume that*

$$\mathcal{E}_{\mathcal{A}}(\mathrm{D}\boldsymbol{U}_k, \boldsymbol{S}_k(\mathrm{D}\boldsymbol{U}_k)) \to 0 \quad \text{as } k \to \infty;$$

then,

$$(\mathrm{D}\boldsymbol{u}_\infty(x), \boldsymbol{S}_\infty(x)) \in \mathcal{A}(x) \quad \text{for almost every } x \in \Omega.$$

Proof. The proof of this lemma is postponed to Section 6.2 below. ☐

Lemma 5.5. *Assume that the sequence* $\{n_k\}_{k\in\mathbb{N}}$ *satisfies* $n_k \to N < \infty$ *as* $k \to \infty$. *We then have that*

$$\mathcal{E}_{\mathcal{G}_k}(\boldsymbol{U}_k, P_k, \boldsymbol{S}_k(\mathrm{D}\boldsymbol{U}_k)) \to 0 \qquad \text{as } k \to \infty.$$

Proof. The proof of this lemma is postponed to Section 6.3 below. ☐

We further assume that the graph approximation is uniform with respect to the graph approximation indicator.

Assumption 5.6. For every $\epsilon > 0$, there exists an $N = N(\epsilon) \in \mathbb{N}$, such that

$$\mathcal{E}_{\mathcal{A}}(\mathrm{D}\boldsymbol{v}, \boldsymbol{S}^n(\cdot, \mathrm{D}\boldsymbol{v})) < \epsilon \qquad \text{for all } \boldsymbol{v} \in W_0^{1,r}(\Omega)^d \text{ and } n > N.$$

We note that this and Assumption 3.2 are the only strong assumptions among the ones we have made; Assumption 5.6 is, however, only used in the proof of the next theorem, and is not required for any of the preceding results.

Theorem 5.7. *Suppose that Assumption* 5.6 *holds and let* $\{(\boldsymbol{U}_k, P_k, \boldsymbol{S}_k(\mathrm{D}\boldsymbol{U}_k))\}$ *be the sequence of function triples produced by the AFEM. We then have that*

$$\mathcal{E}_{\mathcal{A}}(\mathrm{D}\boldsymbol{U}_k, \boldsymbol{S}_k(\mathrm{D}\boldsymbol{U}_k)) \to 0 \qquad \text{as } k \to \infty$$

and, for a not relabelled subsequence, we have that

$$\mathcal{E}_{\mathcal{G}_k}(\boldsymbol{U}_k, P_k, \boldsymbol{S}_k(\mathrm{D}\boldsymbol{U}_k)) \to 0 \qquad \text{as } k \to \infty.$$

Proof. We argue by contradiction. First assume that there exists an $\epsilon > 0$ such that, for some subsequence, we have that

$$\mathcal{E}_{\mathcal{A}}(\mathrm{D}\boldsymbol{U}_{k_\ell}, \boldsymbol{S}_{k_\ell}(\mathrm{D}\boldsymbol{U}_{k_\ell})) > \epsilon \qquad \text{for all } \ell \in \mathbb{N}.$$

Consequently, by Assumption 5.6, we have that $n_{k_\ell} = N$, for some $\ell_0, N \in \mathbb{N}$, and all $\ell \geqslant \ell_0$. Moreover, thanks to Lemma 5.2, there exists a not relabelled subsequence $\{(\boldsymbol{U}_{k_\ell}, P_{k_\ell}, \boldsymbol{S}_{k_\ell}(\mathrm{D}\boldsymbol{U}_{k_\ell}))\}_{\ell\in\mathbb{N}}$ that converges weakly in $W_0^r(\Omega)^d \times L_0^{\tilde{r}}(\Omega) \times L^{r'}(\Omega; \mathbb{R}^{d\times d}_{\mathrm{sym}})$. Combining these facts, we deduce with Lemma 5.5 that

$$\mathcal{E}_{\mathcal{G}_{k_\ell}}(\boldsymbol{U}_{k_\ell}, P_{k_\ell}, \boldsymbol{S}_{k_\ell}(\mathrm{D}\boldsymbol{U}_{k_\ell})) \to 0.$$

In particular, there exists an $\ell > \ell_0$, such that $\mathcal{E}_{\mathcal{G}_{k_\ell}}(\boldsymbol{U}_{k_\ell}, P_{k_\ell}, \boldsymbol{S}_{k_\ell}(\mathrm{D}\boldsymbol{U}_{k_\ell})) < \epsilon$. Therefore, by line 10 of AFEM we have that $n_{k_\ell+1} = N + 1$, a contradiction. Consequently, we have (for the full sequence) that

$$\mathcal{E}_{\mathcal{A}}(\mathrm{D}\boldsymbol{U}_k, \boldsymbol{S}_k(\mathrm{D}\boldsymbol{U}_k)) \to 0 \qquad \text{as } k \to \infty.$$

This proves the first claim.

Assume now that there exists an $\epsilon > 0$ such that we have that

$$\mathcal{E}_{\mathcal{G}_k}(\boldsymbol{U}_k, P_k, \boldsymbol{S}_k(\mathrm{D}\boldsymbol{U}_k)) > \epsilon \qquad \text{for all } k \in \mathbb{N}. \tag{5.3}$$

By the above considerations, there exists a $k_0 \in \mathbb{N}$ such that $\mathcal{E}_{\mathcal{A}}(\mathrm{D}\boldsymbol{U}_k, \boldsymbol{S}_k(\mathrm{D}\boldsymbol{U}_k)) < \epsilon$ for all $k \geqslant k_0$. Therefore, according to line 5 of AFEM, we have that $n_k = n_{k_0}$ for all $k \geqslant k_0$. Consequently, Lemma 5.5 contradicts (5.3).

Combining the two cases proves the assertion. □

Corollary 5.8. *Let $\{(\boldsymbol{U}_k, P_k, \boldsymbol{S}_k(\mathrm{D}\boldsymbol{U}_k))\}$ be the sequence of function triples produced by the AFEM. Then, there exists a not relabelled subsequence with weak limit $(\boldsymbol{u}_\infty, p_\infty, \boldsymbol{S}_\infty) \in W_0^{1,r}(\Omega)^d \times L_0^{\tilde{r}}(\Omega) \times L^{r'}(\Omega; \mathbb{R}^{d\times d}_{sym})$ such that*

$$\mathcal{E}_{\mathcal{A}}(\mathrm{D}\boldsymbol{U}_k, \boldsymbol{S}_k(\mathrm{D}\boldsymbol{U}_k)) \to 0 \qquad \text{and} \qquad \mathcal{E}_{\mathcal{G}_k}(\boldsymbol{U}_k, P_k, \boldsymbol{S}_k(\mathrm{D}\boldsymbol{U}_k)) \to 0,$$

as $k \to \infty$ and $(\boldsymbol{u}_\infty, p_\infty, \boldsymbol{S}_\infty)$ solves (2.2).

Proof. The claim follows from Theorem 5.7, Lemma 5.4, Corollary 5.3, and Lemma 4.2. □

Remark 5.9. We emphasize that even in the case when the exact solution of (2.2) is unique, we do not have that the statement of Corollary 5.8 is true for the full sequence. This is due to the fact that the finite element error estimator is not necessarily decreasing with respect to the refinement of the graph approximation. However, when the exact solution is unique, it is easy to select a converging subsequence with the help of the estimators; one can choose, for example, a subsequence, such that $\mathcal{E}_{\mathcal{G}_k}(\boldsymbol{U}_{k_\ell}, P_{k_\ell}, \boldsymbol{S}_{k_\ell}(\mathrm{D}\boldsymbol{U}_{k_\ell}))$ is monotonic decreasing in ℓ.

6. The Proofs of the Auxiliary Results

6.1. Proof of Lemma 5.2

We recall (3.23) and observe that the spaces $W_0^{1,r}(\Omega)^d$ and $L^{r'}(\Omega; \mathbb{R}^{d\times d}_{\mathrm{sym}})$, $r \in (1, \infty)$, are reflexive. Therefore, there exist $\boldsymbol{u}_\infty \in \mathbb{V}^r_\infty$ and $\boldsymbol{S}_\infty \in L^{r'}(\Omega; \mathbb{R}^{d\times d}_{\mathrm{sym}})$ such that for a not relabelled subsequence we have

$$\boldsymbol{U}_k \rightharpoonup \boldsymbol{u}_\infty \qquad \text{weakly in} \quad W_0^{1,r}(\Omega)^d \tag{6.1}$$

and

$$\boldsymbol{S}_k(\mathrm{D}\boldsymbol{U}_k) \rightharpoonup \boldsymbol{S}_\infty \qquad \text{weakly in} \quad L^{r'}(\Omega; \mathbb{R}^{d\times d}_{\mathrm{sym}}), \tag{6.2}$$

as $k \to \infty$. The function $\boldsymbol{u}_\infty$ is discretely divergence-free with respect to $\mathbb{Q}^{r'}_\infty$, *i.e.*,

$$\int_\Omega q \operatorname{div} \boldsymbol{u}_\infty \,\mathrm{d}x = \lim_{k\to\infty} \int_\Omega (\mathfrak{I}^{\mathcal{G}_k}_{\mathbb{Q}} q) \operatorname{div} \boldsymbol{U}_k \,\mathrm{d}x = 0 \qquad \text{for all } q \in \mathbb{Q}^{r'}_\infty.$$

This follows from (3.11) as in the proof of ([22], Lem. 19), replacing ([22], (3.5)) with the density of the union of the discrete pressure spaces in $\mathbb{Q}^{r'}_\infty$.

Moreover, using compact embeddings of Sobolev spaces, we have that

$$\boldsymbol{U}_k \to \boldsymbol{u}_\infty \qquad \text{strongly in} \quad L^s(\Omega)^d \quad \text{for all} \quad \begin{cases} s \in \left(1, \frac{rd}{d-r}\right), & \text{if } r < d, \\ s \in (1, \infty), & \text{otherwise.} \end{cases} \tag{6.3}$$

This implies, for arbitrary $\boldsymbol{v} \in W_0^{1,\infty}(\Omega)^d$, that

$$\mathcal{B}[\boldsymbol{U}_k, \boldsymbol{U}_k, \boldsymbol{v}] \to \mathcal{B}[\boldsymbol{u}_\infty, \boldsymbol{u}_\infty, \boldsymbol{v}],$$

or equivalently,

$$\mathrm{B}[\boldsymbol{U}_k,\boldsymbol{U}_k] \to \mathrm{B}[\boldsymbol{u}_\infty,\boldsymbol{u}_\infty] \quad \text{weakly in } W^{-1,1}(\Omega)^d$$

as $k \to \infty$; compare also with ([22], Lem. 19).

We now prove convergence of the pressure. Thanks to (3.18), we have

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

for all $\boldsymbol{V} \in \mathbb{V}(\mathcal{G}_k)$. By (3.23) and the discrete inf-sup condition stated in Proposition 3.7, it follows that the sequence $\{P_k\}_{k\in\mathbb{N}}$ is bounded in the reflexive Banach space $L_0^{\tilde{r}}(\Omega)$. Hence, there exists a $p_\infty \in \mathbb{Q}_\infty^{\tilde{r}} \subset L_0^{\tilde{r}}(\Omega)$ such that, for a (not relabelled) subsequence,

$$P_k \rightharpoonup p_\infty \qquad \text{weakly in } L_0^{\tilde{r}}(\Omega).$$

On the other hand we deduce for an arbitrary $\boldsymbol{v} \in \mathbb{V}_\infty^\infty \subset W_0^{1,\infty}(\Omega)^d$ that

$$\begin{aligned}\int_\Omega p_\infty \operatorname{div}\boldsymbol{v}\,\mathrm{d}x \leftarrow \int_\Omega P_k \operatorname{div}\boldsymbol{v}\,\mathrm{d}x &= \int_\Omega P_k \operatorname{div}\mathfrak{I}_{\operatorname{div}}^{\mathcal{G}_k}\boldsymbol{v}\,\mathrm{d}x \\ &= \int_\Omega \boldsymbol{S}_k(\mathrm{D}\boldsymbol{U}_k) : \mathrm{D}\mathfrak{I}_{\operatorname{div}}^{\mathcal{G}_k}\boldsymbol{v} - \boldsymbol{f}\cdot\mathfrak{I}_{\operatorname{div}}^{\mathcal{G}_k}\boldsymbol{v} + \mathrm{B}[\boldsymbol{U}_k,\boldsymbol{U}_k]\cdot\mathfrak{I}_{\operatorname{div}}^{\mathcal{G}_k}\boldsymbol{v}\,\mathrm{d}x \\ &\to \int_\Omega \boldsymbol{S}_\infty : \mathrm{D}\boldsymbol{v}\,\mathrm{d}x + \mathrm{B}[\boldsymbol{u}_\infty,\boldsymbol{u}_\infty]\cdot\boldsymbol{v} - \boldsymbol{f}\cdot\boldsymbol{v}\,\mathrm{d}x\end{aligned}$$

as $k \to \infty$, where we have used (6.2), the properties of $\mathfrak{I}_{\operatorname{div}}^{\mathcal{G}_k}$ together with the density of the union of the discrete velocity spaces in $\mathbb{V}_\infty^\infty$ and the boundedness of the sequence $\{P_k\}_{k\in\mathbb{N}}$ in $L_0^{\tilde{r}}(\Omega)$. The assertion for all $\boldsymbol{v} \in \mathbb{V}_\infty^{\tilde{r}'}$ then follows from the density of $\mathbb{V}_\infty^\infty$ in $\mathbb{V}_\infty^{\tilde{r}'}$. □

6.2. Proof of Lemma 5.4

According to Proposition 4.1, for $k \in \mathbb{N}$ there exist $\tilde{\boldsymbol{D}}_k \in L^r(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}})$ and $\tilde{\boldsymbol{S}}_k \subset L^{r'}(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}})$, such that $(\tilde{\boldsymbol{D}}_k(x),\tilde{\boldsymbol{S}}_k(x)) \in \mathcal{A}(x)$ for a.e. $x \in \Omega$ and

$$\left\|\mathrm{D}\boldsymbol{U}_k - \tilde{\boldsymbol{D}}_k\right\|_r^r + \left\|\boldsymbol{S}_k(\mathrm{D}\boldsymbol{U}_k) - \tilde{\boldsymbol{S}}_k\right\|_{r'}^{r'} = \mathcal{E}_{\mathcal{A}}(\mathrm{D}\boldsymbol{U}_k,\boldsymbol{S}_k(\mathrm{D}\boldsymbol{U}_k)) \to 0 \quad \text{as } k \to \infty. \tag{6.4}$$

Thanks to (3.23), the sequences $\{\tilde{\boldsymbol{D}}_k\}_{k\in\mathbb{N}}$ and $\{\tilde{\boldsymbol{S}}_k\}_{k\in\mathbb{N}}$ are bounded in $L^r(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}})$ and $L^{r'}(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}})$ respectively. Since both spaces are reflexive, together with the uniqueness of the limit, we obtain that

$$\tilde{\boldsymbol{D}}_k \rightharpoonup \boldsymbol{D}\boldsymbol{u}_\infty \qquad \text{weakly in } L^r(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}}), \tag{6.5a}$$

$$\tilde{\boldsymbol{S}}_k \rightharpoonup \boldsymbol{S}_\infty \qquad \text{weakly in } L^{r'}(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}}) \tag{6.5b}$$

as $k \to \infty$.

Let $\boldsymbol{S}^\star : \Omega \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$ be a measurable selection with $(\boldsymbol{\delta},\boldsymbol{S}^\star(x,\boldsymbol{\delta})) \in \mathcal{A}(x)$ for a.e. $x \in \Omega$ and thus $(\mathrm{D}\boldsymbol{u}_\infty(x),\boldsymbol{S}^\star(x,\mathrm{D}\boldsymbol{u}_\infty(x))) \in \mathcal{A}(x)$ for a.e. $x \in \Omega$; compare with Remark 2.2. Consequently, for every bounded sequence $\{\phi_k\}_{k\in\mathbb{N}} \in L^\infty(\Omega)$ of nonnegative functions, we have (recall (A3)) that

$$\begin{aligned} 0 &\leqslant \limsup_{k\to\infty} \int_\Omega \left| \left(\tilde{\boldsymbol{S}}_k - \boldsymbol{S}^\star(\mathrm{D}\boldsymbol{u}_\infty)\right) : (\tilde{\boldsymbol{D}}_k - \boldsymbol{D}\boldsymbol{u}_\infty) \right| \phi_k \,\mathrm{d}x \\ &= \limsup_{k\to\infty} \int_\Omega \left(\tilde{\boldsymbol{S}}_k - \boldsymbol{S}^\star(\mathrm{D}\boldsymbol{u}_\infty)\right) : (\tilde{\boldsymbol{D}}_k - \boldsymbol{D}\boldsymbol{u}_\infty)\, \phi_k \,\mathrm{d}x \\ &= \limsup_{k\to\infty} \int_\Omega \underbrace{(\boldsymbol{S}_k(\mathrm{D}\boldsymbol{U}_k) - \boldsymbol{S}^\star(\mathrm{D}\boldsymbol{u}_\infty)) : (\mathrm{D}\boldsymbol{U}_k - \boldsymbol{D}\boldsymbol{u}_\infty)}_{=:a_k(x)} \phi_k \,\mathrm{d}x, \end{aligned} \tag{6.6}$$

where we have used (6.4) in the last step. We assume for the moment that $a_k \to 0$ in measure and therefore

$$a_k \to 0 \quad \text{a.e. in } \Omega \tag{6.7}$$

for at least a subsequence of a_k. Since a_k is bounded in $L^1(\Omega)$, we obtain with the biting lemma (Lem. 2.3) and Vitali's theorem, that there exists a nonincreasing sequence of measurable subsets $E_j \subset \Omega$ with $|E_j| \to 0$ as $j \to \infty$, such that for all $j \in \mathbb{N}$, we have

$$a_k \to 0 \qquad \text{strongly in } L^1(\Omega\backslash E_j) \quad \text{as } k \to \infty.$$

This, together with (6.6) and (6.5), implies for all nonnegative $\phi \in L^\infty(\Omega\backslash E_j) \subset L^\infty(\Omega)$ (extend ϕ by zero on E_j) and each fixed $j \in \mathbb{N}$, that

$$\lim_{k\to\infty} \int_{\Omega\backslash E_j} \tilde{\boldsymbol{S}}_k : \tilde{\boldsymbol{D}}_k \phi \,\mathrm{d}x = \int_{\Omega\backslash E_j} \boldsymbol{S}_\infty : \boldsymbol{D}\boldsymbol{u}_\infty \phi \,\mathrm{d}x.$$

Consequently, since the graph is monotone and $(\tilde{\boldsymbol{D}}_k(x), \tilde{\boldsymbol{S}}_k(x)) \in \mathcal{A}(x)$ for a.e. $x \in \Omega$, we observe for arbitrary $\boldsymbol{\delta} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$ and all nonnegative $\phi \in L^\infty(\Omega\backslash E_j)$, that

$$0 \leqslant \lim_{k\to\infty} \int_{\Omega\backslash E_j} \left(\tilde{\boldsymbol{S}}_k - \boldsymbol{S}^\star(\cdot, \boldsymbol{\delta})\right) : (\tilde{\boldsymbol{D}}_k - \boldsymbol{\delta})\phi \,\mathrm{d}x = \int_{\Omega\backslash E_j} \left(\boldsymbol{S}_\infty - \boldsymbol{S}^\star(\cdot, \boldsymbol{\delta})\right) : (\boldsymbol{D}\boldsymbol{u}_\infty - \boldsymbol{\delta})\phi \,\mathrm{d}x.$$

Since ϕ was arbitrary, we have that

$$\left(\boldsymbol{S}_\infty - \boldsymbol{S}^\star(\cdot, \boldsymbol{\delta})\right) : (\boldsymbol{D}\boldsymbol{u}_\infty - \boldsymbol{\delta}) \geqslant 0 \qquad \text{for all } \boldsymbol{\delta} \in \mathbb{R}^{d\times d}_{\mathrm{sym}} \text{ and a.e. } x \in \Omega\backslash E_j.$$

According to Lemma 3.1, this implies that

$$(\mathrm{D}\boldsymbol{u}_\infty(x), \boldsymbol{S}_\infty(x)) \in \mathcal{A}(x) \qquad \text{for almost every } x \in \Omega\backslash E_j.$$

The assertion then follows from $|E_j| \to 0$ as $j \to \infty$.

It remains to verify that $a_k \to 0$ in measure as $k \to \infty$. We divide the proof into four steps.

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

$$\boldsymbol{E}_k := \mathfrak{I}^{\mathcal{G}_k}_{\mathrm{div}}(\boldsymbol{U}_k - \boldsymbol{u}_\infty) = \boldsymbol{U}_k - \mathfrak{I}^{\mathcal{G}_k}_{\mathrm{div}} \boldsymbol{u}_\infty \in \mathbb{V}(\mathcal{G}_k)$$

and let $\{\boldsymbol{E}_{k,j}\}_{k,j\in\mathbb{N}}$ be the sequence of Lipschitz-truncated finite element functions according to Proposition 3.12. Recall from Lemma 5.2 that $\boldsymbol{E}_k \rightharpoonup 0$ weakly in $W^{1,r}_0(\Omega)^d$, *i.e.*, we are exactly in the situation of Proposition 3.12. Although $\boldsymbol{E}_k \in \mathbb{V}_0(\mathcal{G}_k)$, *i.e.*, $\boldsymbol{E}_k$ is discretely divergence-free, this does not necessarily imply that $\boldsymbol{E}_{k,j} \in \mathbb{V}_0(\mathcal{G}_k)$

and thus we need to modify $\boldsymbol{E}_{k,j}$ in order to be able to use it as a test function in (3.21). With the discrete Bogovskiĭ operator $\mathfrak{B}^k := \mathfrak{B}^{\mathcal{G}_k}$ from Corollary 3.8, we define

$$\boldsymbol{\Psi}_{k,j} := \mathfrak{B}^k(\operatorname{div} \boldsymbol{E}_{k,j}) \in \mathbb{V}(\mathcal{G}_k). \tag{6.8a}$$

The 'corrected' function

$$\boldsymbol{\Phi}_{k,j} := \boldsymbol{E}_{k,j} - \boldsymbol{\Psi}_{k,j} \in \mathbb{V}_0(\mathcal{G}_k) \tag{6.8b}$$

is then discretely divergence-free. We need to control the correction in a norm. To this end we recall from Section 3.3 that $\mathbb{Q}(\mathcal{G}_k) = \operatorname{span}\{Q_1^k, \dots, Q_{N_k}^k\}$ for a certain locally supported basis. Then, thanks to properties of the discrete Bogovskiĭ operator and Corollary 3.8, we have that

$$\begin{aligned}
\beta_r \|\boldsymbol{\Psi}_{k,j}\|_{1,r} &\leqslant \sup_{Q\in\mathbb{Q}(\mathcal{G}_k)} \frac{\int_\Omega Q \operatorname{div} \boldsymbol{E}_{k,j}\,\mathrm{d}x}{\|Q\|_{r'}} = \sup_{Q\in\mathbb{Q}(\mathcal{G}_k)} \frac{\int_\Omega Q \operatorname{div} \boldsymbol{E}_{k,j} - \operatorname{div} \boldsymbol{E}_k\,\mathrm{d}x}{\|Q\|_{r'}} \\
&= \sup_{Q=\sum_{i=1}^{N_k}\rho_i Q_i^k} \Bigg(\sum_{\operatorname{supp} Q_i^k \subset \{\boldsymbol{E}_{k,j}=\boldsymbol{E}_k\}} \frac{\int_\Omega \rho_i Q_i^k \operatorname{div}(\boldsymbol{E}^{k,j} - \boldsymbol{E}_k)\,\mathrm{d}x}{\|Q\|_{r'}} \\
&\quad + \sum_{\operatorname{supp} Q_i^k \cap \{\boldsymbol{E}_{k,j}\neq\boldsymbol{E}_k\}\neq\emptyset} \frac{\int_\Omega \rho_i Q_i^k \operatorname{div}(\boldsymbol{E}_{k,j} - \boldsymbol{E}_k)\,\mathrm{d}x}{\|Q\|_{r'}} \Bigg) \\
&= \sup_{Q=\sum_{i=1}^{N_k}\rho_i Q_i^n} \Bigg(\sum_{\operatorname{supp} Q_i^k \cap \{\boldsymbol{E}_{k,j}\neq\boldsymbol{E}_k\}\neq\emptyset} \frac{\int_\Omega \rho_i Q_i^k \operatorname{div}(\boldsymbol{E}_{k,j} - \boldsymbol{E}_k)\,\mathrm{d}x}{\|Q\|_{r'}} \Bigg) \\
&= \sup_{Q=\sum_{i=1}^{N_k}\rho_i Q_i^k} \Bigg(\sum_{\operatorname{supp} Q_i^k \cap \{\boldsymbol{E}_{k,j}\neq\boldsymbol{E}_k\}\neq\emptyset} \frac{\int_\Omega \rho_i Q_i^k \operatorname{div} \boldsymbol{E}_{k,j}\,\mathrm{d}x}{\|Q\|_{r'}} \Bigg) \\
&\leqslant \left\| \operatorname{div} \boldsymbol{E}_{k,j} \chi_{\Omega^k_{\{\boldsymbol{E}_{k,j}\neq\boldsymbol{E}_k\}}} \right\|_r \sup_{Q=\sum_{i=1}^{N_k}\rho_i Q_i^k} \frac{\left\| \sum_{\operatorname{supp} Q_i^k \cap \{\boldsymbol{E}_{k,j}\neq\boldsymbol{E}_k\}\neq\emptyset} \rho_i Q_i^k \right\|_{r'}}{\|Q\|_{r'}} \\
&\leqslant c \left\| \operatorname{div} \boldsymbol{E}_{k,j}\, \chi_{\Omega^k_{\{\boldsymbol{E}_{k,j}\neq\boldsymbol{E}_k\}}} \right\|_r \leqslant c \left\| \nabla \boldsymbol{E}_{k,j}\, \chi_{\Omega^k_{\{\boldsymbol{E}_{k,j}\neq\boldsymbol{E}_k\}}} \right\|_r,
\end{aligned}$$

where $\chi_{\Omega^k_{\{\boldsymbol{E}_{k,j}\neq\boldsymbol{E}_k\}}}$ is the characteristic function of the set

$$\Omega^k_{\{\boldsymbol{E}_{k,j}\neq\boldsymbol{E}_k\}} := \bigcup \left\{ \Omega^k_E \mid E \in \mathcal{G}_k \text{ such that } E \subset \overline{\{\boldsymbol{E}_{k,j} \neq \boldsymbol{E}_k\}} \right\}.$$

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 (3.2b). In particular, we see by means of standard scaling arguments that for $Q = \sum_{i=1}^{N_k} \rho_i Q_i^k$ the norms

$$Q \mapsto \left(\sum_{i=1}^{N_k} |\rho_i|^{r'} \|Q_i^k\|_{r'}^{r'} \right)^{1/r'} \qquad \text{and} \qquad Q \mapsto \|Q\|_{r'}$$

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

Observe that $|\Omega^k_E| \leqslant c\,|E|$ for all $E \in \mathcal{G}_k$, $k \in \mathbb{N}$, with a shape-dependent constant $c > 0$; hence, $\left|\Omega^k_{\{\boldsymbol{E}_{k,j}\neq\boldsymbol{E}_k\}}\right| \leqslant c\,|\{\boldsymbol{E}_{k,j} \neq \boldsymbol{E}_k\}|$, and it follows from Proposition 3.12f that

$$\beta_r \|\boldsymbol{\Psi}_{k,j}\|_{1,r} \leqslant c \left\| \lambda_{k,j} \chi_{\Omega^k_{\{\boldsymbol{E}_{k,j}\neq\boldsymbol{E}_k\}}} \right\|_r \leqslant c\, 2^{-j/r} \|\nabla \boldsymbol{E}_k\|_r. \tag{6.9}$$

Moreover, we have from Proposition 3.12 and the continuity properties of $\mathfrak{B}^n$ (see Cor. 3.8) that

$$\boldsymbol{\Phi}_{k,j}, \boldsymbol{\Psi}_{k,j} \rightharpoonup 0 \qquad \text{weakly in } W^{1,s}_0(\Omega)^d \quad \text{for all } s \in [1,\infty), \tag{6.10a}$$

$$\boldsymbol{\Phi}_{k,j}, \boldsymbol{\Psi}_{k,j} \to 0 \qquad \text{strongly in } L^s(\Omega)^d \quad \text{for all } s \in [1,\infty), \tag{6.10b}$$

as $k \to \infty$.

Step 2: We shall prove (recall the last line of (6.6) for the definition of a_k) that

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

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

$$\limsup_{k\to\infty} \int_{\{\boldsymbol{E}^n = \boldsymbol{E}^{n,j}\}} |a_k| \,\mathrm{d}x \leqslant \limsup_{k\to\infty} \int_{\{\boldsymbol{E}^n = \boldsymbol{E}^{n,j}\}} a_k \,\mathrm{d}x + 2 \limsup_{k\to\infty} \int_{\{\boldsymbol{E}^n = \boldsymbol{E}^{n,j}\}} a_k^- \,\mathrm{d}x. \tag{6.11}$$

By choosing $\phi_k := \chi_{\mathrm{supp}(a_k^-)} \in L^\infty(\Omega)$ in (6.6), we observe that the latter term is zero. In order to bound the first term, we recall (6.8) and observe that

$$\begin{aligned}
\int_{\{\boldsymbol{E}_k = \boldsymbol{E}_{k,j}\}} a_k \,\mathrm{d}x &= \int_{\{\boldsymbol{E}_k = \boldsymbol{E}_{k,j}\}} (\boldsymbol{S}_k - \boldsymbol{S}^\star(\cdot, \mathrm{D}\boldsymbol{u}_\infty)) : (\mathrm{D}\mathfrak{I}^{\mathcal{G}_k}_{\mathrm{div}} \boldsymbol{u}_\infty - \mathrm{D}\boldsymbol{u}_\infty) \,\mathrm{d}x \\
&\quad + \int_\Omega \boldsymbol{S}_k : \mathrm{D}\boldsymbol{\Phi}_{k,j} \,\mathrm{d}x + \int_\Omega \boldsymbol{S}_k : \mathrm{D}\boldsymbol{\Psi}_{k,j} \,\mathrm{d}x - \int_\Omega \boldsymbol{S}^\star(\cdot, \mathrm{D}\boldsymbol{u}_\infty) : \mathrm{D}\boldsymbol{E}_{k,j} \,\mathrm{d}x \\
&\quad + \int_{\{\boldsymbol{E}_k \neq \boldsymbol{E}_{k,j}\}} \big(\boldsymbol{S}^\star(\cdot, \mathrm{D}\boldsymbol{u}_\infty) - \boldsymbol{S}_k\big) : \mathrm{D}\boldsymbol{E}_{k,j} \,\mathrm{d}x \\
&= \mathrm{I}_{k,j} + \mathrm{II}_{k,j} + \mathrm{III}_{k,j} + \mathrm{IV}_{k,j} + \mathrm{V}_{k,j}.
\end{aligned}$$

Thanks to (5.2) and (3.23) we have that

$$\begin{aligned}
|\mathrm{I}_{k,j}| &\leqslant \int_{\{\boldsymbol{E}_k = \boldsymbol{E}_{k,j}\}} |\boldsymbol{S}_k(\cdot, \mathrm{D}\boldsymbol{U}_k) - \boldsymbol{S}^\star(\cdot, \mathrm{D}\boldsymbol{u}_\infty)| \left|\mathrm{D}\mathfrak{I}^{\mathcal{G}_k}_{\mathrm{div}} \boldsymbol{u}_\infty - \mathrm{D}\boldsymbol{u}_\infty\right| \,\mathrm{d}x \\
&\leqslant \|\boldsymbol{S}_k(\cdot, \mathrm{D}\boldsymbol{U}_k) - \boldsymbol{S}^\star(\cdot, \mathrm{D}\boldsymbol{u}_\infty)\|_{r'} \left\|\mathrm{D}\mathfrak{I}^{\mathcal{G}_k}_{\mathrm{div}} \boldsymbol{u}_\infty - \mathrm{D}\boldsymbol{u}_\infty\right\|_r \to 0
\end{aligned}$$

as $k \to \infty$. In order to estimate $\mathrm{II}_{k,j}$ we recall that $\boldsymbol{\Phi}_{k,j} \in \mathbb{V}_0(\mathcal{G}_k)$ is discretely divergence-free, and we can therefore use it as a test function in (3.21) to deduce that

$$\mathrm{II}_{k,j} = -\mathcal{B}[\boldsymbol{U}_k, \boldsymbol{U}_k, \boldsymbol{\Phi}_{k,j}] + \int_\Omega \boldsymbol{f} \cdot \boldsymbol{\Phi}_{k,j} \,\mathrm{d}x \to 0 \qquad \text{as } k \to \infty.$$

Indeed, the second term vanishes thanks to (6.10a). The first term vanishes thanks to (6.3) and the weak convergence (6.10a) of $\boldsymbol{\Phi}_{k,j}$. The term $\mathrm{III}_{k,j}$ can be bounded by means of (6.9); in particular,

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

where we have used (3.23). Proposition 3.12 implies that

$$\lim_{k\to\infty} \mathrm{IV}_{k,j} = 0.$$

Finally, by (3.23) and Proposition 3.12, we have that

$$\limsup_{k\to\infty}\ |\mathrm{V}_{k,j}| \leqslant \limsup_{k\to\infty}\big(\|\boldsymbol{S}^\star(\cdot,\mathrm{D}\boldsymbol{u}_\infty)\|_{r'}+\|\boldsymbol{S}_k(\cdot,\mathrm{D}\boldsymbol{U}_k)\|_{r'}\big)\left\|\mathrm{D}\boldsymbol{E}_{k,j}\chi_{\{\boldsymbol{E}_k\neq\boldsymbol{E}_{k,j}\}}\right\|_r \leqslant c\,2^{-j/r}.$$

In view of (6.11), this completes Step 2.

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

$$\lim_{k\to\infty}\int_\Omega |a_k|^\vartheta\ \mathrm{d}x = 0, \tag{6.12}$$

which then implies the assertion $a_k\to 0$ in measure as $k\to\infty$.

Using Hölder's inequality, we easily obtain that

$$\int_\Omega |a_k|^\vartheta\ \mathrm{d}x = \int_{\{\boldsymbol{E}_k=\boldsymbol{E}_{k,j}\}}|a_k|^\vartheta\ \mathrm{d}x + \int_{\{\boldsymbol{E}_k\neq\boldsymbol{E}_{k,j}\}}|a_k|^\vartheta\ \mathrm{d}x \leqslant |\Omega|^{1-\vartheta}\left(\int_{\{\boldsymbol{E}_k=\boldsymbol{E}_{k,j}\}}|a_k|\ \mathrm{d}x\right)^\vartheta + \left(\int_\Omega |a_k|\ \mathrm{d}x\right)^\vartheta |\{\boldsymbol{E}_k\neq\boldsymbol{E}_{k,j}\}|^{1-\vartheta}.$$

Thanks to (3.23), we have that $(\int_\Omega |a_k|\ \mathrm{d}x)^\vartheta$ is bounded uniformly in k and by Proposition 3.12 we have that

$$|\{\boldsymbol{E}_k\neq\boldsymbol{E}_{k,j}\}| \leqslant c\,\frac{\|\boldsymbol{E}_k\|_{1,r}^r}{\lambda_{k,j}^r} \leqslant \frac{c}{2^{2^j r}},$$

where we have used that $\{\boldsymbol{E}_k\}_{k\in\mathbb{N}}$ is bounded in $W_0^{1,r}(\Omega)^d$ according to (3.23) and Assumption 3.3. Consequently, from Step 2 we deduce that

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

The left-hand side is independent of j and we can thus pass to the limit $j\to\infty$. This proves (6.12). □

6.3. Proof of Lemma 5.5

Since $n_k\to N$ as $k\to\infty$, we may, w.l.o.g., assume that $n_k=N$ for all $k\in\mathbb{N}$.

Step 1: We shall first prove that, in this case, we have that the (sub)sequences in Lemma 5.2 do actually converge strongly, *i.e.*,

$$\begin{aligned}
\boldsymbol{U}_k &\to \boldsymbol{u}_\infty && \text{in } W_0^{1,t}(\Omega)^d,\\
P_k &\to p_\infty && \text{in } L_0^{\tilde{t}}(\Omega),\\
\boldsymbol{S}_k(\mathrm{D}\boldsymbol{U}_k) = \boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k) &\to \boldsymbol{S}_\infty = \boldsymbol{S}^N(\mathrm{D}\boldsymbol{u}_\infty) && \text{in } L^{\tilde{t}}(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}}).
\end{aligned} \tag{6.13}$$

To this end, we investigate

$$a_k := \big(\boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k)-\boldsymbol{S}^N(\mathrm{D}\boldsymbol{u}_\infty)\big):\mathrm{D}(\boldsymbol{U}_k-\boldsymbol{u}_\infty)\geqslant 0$$

(compare with Assump. 3.2) distinguishing two cases: $r\leqslant\frac{3d}{d+2}$ and $r>\frac{3d}{d+2}$.

If $r\leqslant\frac{3d}{d+2}$, then we can deduce, as in the proof of Lemma 5.4 in Section 6.2, that

$$0\leqslant\int_\Omega |a_k|^\vartheta\,\mathrm{d}x = \int_\Omega\Big(\big(\boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k)-\boldsymbol{S}^N(\mathrm{D}\boldsymbol{u}_\infty)\big):\mathrm{D}(\boldsymbol{U}_k-\boldsymbol{u}_\infty)\Big)^\vartheta\,\mathrm{d}x\to 0, \tag{6.14}$$

where we have used that $a_k^- = 0$ almost everywhere in Ω. Thus, recalling that $\boldsymbol{S}^N$ is strictly monotone, we obtain that

$$\mathrm{D}\boldsymbol{U}_k \to \mathrm{D}\boldsymbol{u}_\infty \qquad \text{and} \qquad \boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k) \to \boldsymbol{S}^N(\mathrm{D}\boldsymbol{u}_\infty) \qquad \text{a.e. in } \Omega, \tag{6.15}$$

at least for a subsequence of $k \to \infty$. Since $1 \leqslant t < r$ and $1 \leqslant \tilde{t} < r'$ (compare with (4.3a)), we obtain with Lemma 5.2 and Vitali's theorem that

$$\boldsymbol{U}_k \to \boldsymbol{u}_\infty \quad \text{in } W_0^{1,t}(\Omega), \quad \text{and} \quad \boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k) \to \boldsymbol{S}^N(\mathrm{D}\boldsymbol{u}_\infty) = \boldsymbol{S}_\infty \quad \text{in } L^{\tilde{t}}(\Omega; \mathbb{R}^{d\times d}_{\text{sym}}).$$

Indeed, using Hölder's inequality, we obtain with (3.23) that

$$\|\mathrm{D}\boldsymbol{U}_k - \mathrm{D}\boldsymbol{u}_\infty\|_{t,\omega} \leqslant |\omega|^{\frac{r-t}{rt}} \|\mathrm{D}\boldsymbol{U}_k - \mathrm{D}\boldsymbol{u}_\infty\|_{r,\omega}$$

for all measurable $\omega \subset \Omega$; *i.e.*, $\{|\,\mathrm{D}\boldsymbol{U}_k - \mathrm{D}\boldsymbol{u}_\infty|^t\}_{k\in\mathbb{N}}$ is uniformly integrable and the claim follows from Vitali's theorem and (6.15). The convergence of the stress sequence follows analogously and the claim $\boldsymbol{S}^N(\mathrm{D}\boldsymbol{u}_\infty) = \boldsymbol{S}_\infty$ is a consequence of the uniqueness of the limit.

If $r > \frac{3d}{d+2}$, then $t = r$ and $\tilde{t} = \tilde{r} = r'$; compare with (4.3b). We deduce from Lemma 5.2 and (3.20) that

$$\begin{aligned} 0 &\leqslant \limsup_{k\to\infty} \int_\Omega \big(\boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k) - \boldsymbol{S}^N(\mathrm{D}\boldsymbol{u}_\infty)\big) : \mathrm{D}(\boldsymbol{U}_k - \boldsymbol{u}_\infty)\,\mathrm{d}x \\ &= \limsup_{k\to\infty} \int_\Omega \boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k) : \mathrm{D}\boldsymbol{U}_k - \boldsymbol{S}_\infty : \mathrm{D}\boldsymbol{u}_\infty\,\mathrm{d}x \\ &= \limsup_{k\to\infty} \int_\Omega \boldsymbol{f}\cdot\boldsymbol{U}_k - \boldsymbol{S}_\infty : \mathrm{D}\boldsymbol{u}_\infty\,\mathrm{d}x = \int_\Omega \boldsymbol{f}\cdot\boldsymbol{u}_\infty - \boldsymbol{S}_\infty : \mathrm{D}\boldsymbol{u}_\infty\,\mathrm{d}x = 0. \end{aligned}$$

As before, thanks to the strict monotonicity of $\boldsymbol{S}^N$, we have that

$$\boldsymbol{U}_k \to \boldsymbol{u}_\infty \qquad \text{and} \qquad \boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k) \to \boldsymbol{S}^N(\mathrm{D}\boldsymbol{u}_\infty) \qquad \text{a.e. in } \Omega \tag{6.16}$$

at least for a subsequence of $k \to \infty$. Moreover,

$$\lim_{k\to\infty} \int_\Omega \boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k) : \mathrm{D}\boldsymbol{U}_k\,\mathrm{d}x = \int_\Omega \boldsymbol{S}^N(\mathrm{D}\boldsymbol{u}_\infty) : \mathrm{D}\boldsymbol{u}_\infty\,\mathrm{d}x,$$

and thanks to (3.23), we have that $\boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k) : \mathrm{D}\boldsymbol{U}_k$ is bounded in $L^1(\Omega)$, hence $\boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k) : \mathrm{D}\boldsymbol{U}_k \xrightarrow{b} \boldsymbol{S}^N(\mathrm{D}\boldsymbol{u}_\infty) : \mathrm{D}\boldsymbol{u}_\infty$; compare with Lemma 2.3. Recalling Assumption 3.2 we have $0 \leqslant \tilde{m} + \boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k) : \mathrm{D}\boldsymbol{U}_k$ almost everywhere in Ω. Combining these properties, it follows from Lemma 2.4 that $\tilde{m} + \boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k) : \mathrm{D}\boldsymbol{U}_k \rightharpoonup \tilde{m} + \boldsymbol{S}^N(\mathrm{D}\boldsymbol{u}_\infty) : \mathrm{D}\boldsymbol{u}_\infty$ in $L^1(\Omega)$ and thus $\boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k) : \mathrm{D}\boldsymbol{U}_k \rightharpoonup \boldsymbol{S}^N(\mathrm{D}\boldsymbol{u}_\infty) : \mathrm{D}\boldsymbol{u}_\infty$ in $L^1(\Omega)$. Consequently, by the Dunford–Pettis theorem, $\{\boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k) : \mathrm{D}\boldsymbol{U}_k\}_{k\in\mathbb{N}}$ is uniformly integrable. Thanks to the coercivity of $\boldsymbol{S}^N$, we have that $\{|\,\mathrm{D}\boldsymbol{U}_k|^r\}_{k\in\mathbb{N}}$ and $\{|\boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k)|^{r'}\}_{k\in\mathbb{N}}$ are uniformly integrable and hence we deduce from (6.16), with Vitali's theorem, that

$$\boldsymbol{U}_k \to \boldsymbol{u}_\infty \quad \text{in } W_0^{1,r}(\Omega)^d \quad \text{and} \quad \boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k) \to \boldsymbol{S}_\infty \quad \text{in } L^{r'}(\Omega; \mathbb{R}^{d\times d}_{\text{sym}}).$$

It remains to prove the strong convergence of the pressure (sub)sequence $\{P_k\}_{k\in\mathbb{N}}$ in $L_0^{\tilde{t}}(\Omega)$. Thanks to (2.3), for $k \in \mathbb{N}$ there exists a $\boldsymbol{v}_k \in W_0^{1,\tilde{t}'}(\Omega)^d$ with $\|\boldsymbol{v}_k\|_{\tilde{r}} = 1$, such that

$$\alpha_{\tilde{r}}\,\|p_\infty - P_k\|_{\tilde{t}} \leqslant \int_\Omega (p_\infty - P_k)\,\mathrm{div}\,\boldsymbol{v}_k\,\mathrm{d}x.$$

Since $\{\boldsymbol{v}_k\}_{k\in\mathbb{N}}$ is bounded in $W_0^{1,\tilde{t}'}(\Omega)^d$, there exists a not relabelled weakly converging subsequence with weak limit $\boldsymbol{v} \in W_0^{1,\tilde{t}'}(\Omega)^d$. Therefore, we deduce using the properties of $\mathfrak{I}_{\mathrm{div}}^{\mathcal{G}_k}$ (see Assump. 3.3) that

$$\alpha_{\tilde{r}} \|p_\infty - P_k\|_{\tilde{t}} \leqslant \int_\Omega (p_\infty - P_k) \operatorname{div} \boldsymbol{v}_k \,\mathrm{d}x = \int_\Omega (p_\infty - P_k) \operatorname{div} \mathfrak{I}_{\mathrm{div}}^{\mathcal{G}_k} \boldsymbol{v}_k \,\mathrm{d}x + \int_\Omega p_\infty \operatorname{div} (\boldsymbol{v}_k - \mathfrak{I}_{\mathrm{div}}^{\mathcal{G}_k} \boldsymbol{v}_k) \,\mathrm{d}x.$$

The second term vanishes thanks to Proposition 3.4. For the first term, we have, thanks to Lemma 5.2 and (3.20), that

$$\int_\Omega (p_\infty - P_k) \operatorname{div} \mathfrak{I}_{\mathrm{div}}^{\mathcal{G}_k} \boldsymbol{v}_k \,\mathrm{d}x = \int_\Omega \big(\boldsymbol{S}_\infty - \boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k)\big) : \mathfrak{I}_{\mathrm{div}}^{\mathcal{G}_k} \boldsymbol{v}_k + \big(\mathrm{B}[\boldsymbol{u}_\infty, \boldsymbol{u}_\infty] - \mathrm{B}[\boldsymbol{U}_k, \boldsymbol{U}_k]\big) \cdot \mathfrak{I}_{\mathrm{div}}^{\mathcal{G}_k} \boldsymbol{v}_k \,\mathrm{d}x \to 0,$$

as $k \to \infty$. Here we have used in the last step the strong convergence of $\{\boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k)\}_{k\in\mathbb{N}}$ and $\{\boldsymbol{U}_k\}_{k\in\mathbb{N}}$ in $L^{\tilde{t}}(\Omega; \mathbb{R}^{d\times d}_{\mathrm{sym}})$ respectively $W_0^{1,t}(\Omega)^d$, as well as that the latter result implies that $\mathrm{B}[\boldsymbol{U}_k, \boldsymbol{U}_k] \to \mathrm{B}[\boldsymbol{u}_\infty, \boldsymbol{u}_\infty]$ strongly in $L^{\tilde{t}'}(\Omega)^d$. This completes the proof of (6.13).

Step 2: As a consequence of (6.13) we shall prove that

$$\mathcal{R}\big(\boldsymbol{U}_k, P_k, \boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k)\big) \to \mathcal{R}\big(\boldsymbol{u}_\infty, p_\infty, \boldsymbol{S}^N(\mathrm{D}\boldsymbol{u}_\infty)\big) \quad \text{strongly in} \quad W^{-1,\tilde{t}}(\Omega)^d. \tag{6.17}$$

To this end we observe, for $\boldsymbol{v} \in W_0^{1,\tilde{t}'}(\Omega)^d$, that

$$\begin{aligned}
&\Big\langle \mathcal{R}^{\mathsf{pde}}\big(\boldsymbol{U}_k, P_k, \boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k)\big) - \mathcal{R}^{\mathsf{pde}}\big(\boldsymbol{u}_\infty, p_\infty, \boldsymbol{S}_\infty\big), \boldsymbol{v} \Big\rangle \\
&\qquad = \int_\Omega \big(\boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k) - \boldsymbol{S}_\infty\big) : \mathrm{D}\boldsymbol{v} + \big(\mathrm{B}[\boldsymbol{U}_k, \boldsymbol{U}_k] - \mathrm{B}[\boldsymbol{u}_\infty, \boldsymbol{u}_\infty]\big) \cdot \boldsymbol{v} \,\mathrm{d}x \\
&\qquad\quad + \int_\Omega (P_k - p_\infty) \operatorname{div} \boldsymbol{v} \,\mathrm{d}x \\
&\qquad \leqslant \Big\{ \Big\| \boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k) - \boldsymbol{S}_\infty \Big\|_{\tilde{t}} + \|\mathrm{B}[\boldsymbol{U}_k, \boldsymbol{U}_k] - \mathrm{B}[\boldsymbol{u}_\infty, \boldsymbol{u}_\infty]\|_{\tilde{t}} + \|P_k - p_\infty\|_{\tilde{t}} \Big\} \|\boldsymbol{v}\|_{1,\tilde{t}'}.
\end{aligned}$$

Hence, thanks to (2.1), (6.13), and the fact, that $\mathrm{B}[\boldsymbol{U}_k, \boldsymbol{U}_k] \to \mathrm{B}[\boldsymbol{u}_\infty, \boldsymbol{u}_\infty]$ strongly in $L^{\tilde{t}'}(\Omega)$ (see Step 1), this proves the assertion for $\mathcal{R}^{\mathsf{pde}}$. The assertion for $\mathcal{R}^{\mathsf{ic}}$ is an immediate consequence of (6.13).

Step 3: We use the techniques of [38] to prove that

$$\mathcal{R}\big(\boldsymbol{u}_\infty, p_\infty, \boldsymbol{S}^N(\mathrm{D}\boldsymbol{u}_\infty)\big) = 0. \tag{6.18}$$

To this end, we first need to recall some results from [38].

For each $x \in \Omega$, the mesh-size $h_{\mathcal{G}_k}(x)$ is monotonically decreasing and bounded from below by zero; hence, there exists an $h_\infty \in L^\infty(\Omega)$, such that

$$\lim_{k\to\infty} h_{\mathcal{G}_k} = h_\infty \quad \text{in } L^\infty(\Omega); \tag{6.19}$$

compare *e.g.* with ([38], Lem. 3.2). We next split the domain Ω according to

$$\mathcal{G}_k^+ := \bigcap_{i\geqslant k} \mathcal{G}_i = \{E \in \mathcal{G}_k : E \in \mathcal{G}_i \text{ for all } i \geqslant k\} \quad \text{and} \quad \mathcal{G}_k^0 := \mathcal{G}_k \backslash \mathcal{G}_k^+,$$

i.e., setting $\Omega_k^+ := \Omega(\mathcal{G}_k^+)$ and $\Omega_k^0 := \Omega(\mathcal{G}_k^0)$, we have $\bar{\Omega} = \Omega_k^+ \cup \Omega_k^0$. It is proved in ([38], Cor. 3.3) (compare also (3.8)) that, in the limit, the mesh-size function $h_{\mathcal{G}}$ vanishes on Ω_k^0, *i.e.*,

$$\lim_{k\to\infty} \Big\| h_{\mathcal{G}_k} \chi_{\Omega_k^0} \Big\|_\infty = 0 = \lim_{k\to\infty} \big\| h_{\mathcal{G}_k} \chi_{\Omega_k^\star} \big\|_\infty. \tag{6.20}$$

Here $\Omega_k^\star := \mathcal{U}^{\mathcal{G}}(\Omega_k^0)$ and we have used the local quasi-uniformity of meshes for the latter limit. Since $\mathcal{G}_i^+ \subset \mathcal{G}_k^+ \subset \mathcal{G}_k$ for any $k \geqslant i$, we have

$$\Omega_i^0 = \Omega(\mathcal{G}_i^0) = \Omega(\mathcal{G}_k \backslash \mathcal{G}_i^+).$$

Now, fix $\boldsymbol{v} \in W^{2,\tilde{t}'}(\Omega)^d \cap W_0^{1,\tilde{t}'}(\Omega)^d$ and $q \in W^{1,t'}(\Omega)$, with $\|\boldsymbol{v}\|_{2,\tilde{t}'} = 1 = \|q\|_{1,t'}$. We shall prove that

$$\Big\langle \mathcal{R}\big(\boldsymbol{U}_k, P_k, \boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k)\big), (\boldsymbol{v}, q)\Big\rangle = \Big\langle \mathcal{R}^{\mathsf{pde}}\big(\boldsymbol{U}_k, P_k, \boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k)\big), \boldsymbol{v} - \mathfrak{I}_{\mathrm{div}}^k \boldsymbol{v}\Big\rangle + \big\langle \mathcal{R}^{\mathsf{ic}}(\boldsymbol{U}_k), q - \mathfrak{I}_{\mathbb{Q}}^k q\big\rangle$$

vanishes as $k \to \infty$. Here we use the abbreviations $\mathfrak{I}_{\mathrm{div}}^k := \mathfrak{I}_{\mathrm{div}}^{\mathcal{G}_k}$ and $\mathfrak{I}_{\mathbb{Q}}^k := \mathfrak{I}_{\mathbb{Q}}^{\mathcal{G}_k}$. Then, (6.18) follows from (6.17) and the density of $W^{2,\tilde{t}}(\Omega)^d \cap W_0^{1,\tilde{t}}(\Omega)^d$ in $W_0^{1,\tilde{t}}(\Omega)^d$ and of $W^{1,t'}(\Omega)$ in $L^{t'}(\Omega)$. We shall estimate the two terms on the right-hand side separately. For the first one, we have with Corollary 4.4 that

$$\begin{aligned}
\Big\langle \mathcal{R}^{\mathsf{pde}}\big(\boldsymbol{U}_k, P_k, \boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k)\big), \boldsymbol{v} - \mathfrak{I}_{\mathrm{div}}^k \boldsymbol{v}\Big\rangle &\lesssim \sum_{E \in \mathcal{G}_k} \mathcal{E}_{\mathcal{G}_k}^{\mathsf{pde}}\big(\boldsymbol{U}_k, P_k, \boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k); E\big)^{1/\tilde{t}} \|\nabla \boldsymbol{v} - \nabla \mathfrak{I}_{\mathrm{div}}^k \boldsymbol{v}\|_{\tilde{t}', \mathcal{U}^{\mathcal{G}_k}(E)} \\
&\lesssim \sum_{E \in \mathcal{G}_k \backslash \mathcal{G}_i^+} \mathcal{E}_{\mathcal{G}_k}^{\mathsf{pde}}\big(\boldsymbol{U}_k, P_k, \boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k); E\big)^{1/\tilde{t}} \|\nabla \boldsymbol{v} - \nabla \mathfrak{I}_{\mathrm{div}}^k \boldsymbol{v}\|_{\tilde{t}', \mathcal{U}^{\mathcal{G}_k}(E)} \\
&\quad + \sum_{E \in \mathcal{G}_i^+} \mathcal{E}_{\mathcal{G}_k}^{\mathsf{pde}}\big(\boldsymbol{U}_k, P_k, \boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k); E\big)^{1/\tilde{t}} \|\nabla \boldsymbol{v} - \nabla \mathfrak{I}_{\mathrm{div}}^k \boldsymbol{v}\|_{\tilde{t}', \mathcal{U}^{\mathcal{G}_k}(E)} \\
&\lesssim \mathcal{E}_{\mathcal{G}_k}^{\mathsf{pde}}\big(\boldsymbol{U}_k, P_k, \boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k); \mathcal{G}_k \backslash \mathcal{G}_i^+\big)^{1/\tilde{t}} \|\nabla \boldsymbol{v} - \nabla \mathfrak{I}_{\mathrm{div}}^k \boldsymbol{v}\|_{\tilde{t}', \Omega_i^\star} \\
&\quad + \mathcal{E}_{\mathcal{G}_k}^{\mathsf{pde}}\big(\boldsymbol{U}_k, P_k, \boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k); \mathcal{G}_i^+\big)^{1/\tilde{t}} \|\nabla \boldsymbol{v} - \nabla \mathfrak{I}_{\mathrm{div}}^k \boldsymbol{v}\|_{\tilde{t}', \mathcal{U}^{\mathcal{G}}(\mathcal{G}_i^+)},
\end{aligned}$$

where we have used Hölder's inequality and the finite overlapping of the $\mathcal{U}^{\mathcal{G}_k}(E)$, $E \in \mathcal{G}_k$. In view of Lemma 5.2 and Corollary 4.6 we obtain that

$$\mathcal{E}_{\mathcal{G}_k}^{\mathsf{pde}}\big(\boldsymbol{U}_k, P_k, \boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k); \mathcal{G}_k \backslash \mathcal{G}_i^+\big) \leqslant \mathcal{E}_{\mathcal{G}_k}^{\mathsf{pde}}\big(\boldsymbol{U}_k, P_k, \boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k)\big) \lesssim 1.$$

Recalling (3.7), we thus obtain from the monotonicity of the mesh-size function that

$$\Big\langle \mathcal{R}^{\mathsf{pde}}\big(\boldsymbol{U}_k, P_k, \boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k)\big), \boldsymbol{v} - \mathfrak{I}_{\mathrm{div}}^k \boldsymbol{v}\Big\rangle \lesssim \|h_{\mathcal{G}_i} \chi_{\Omega_i^\star}\|_\infty + \mathcal{E}_{\mathcal{G}_k}^{\mathsf{pde}}\big(\boldsymbol{U}_k, P_k, \boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k); \mathcal{G}_i^+\big)^{1/\tilde{t}}.$$

A similar argument shows that

$$\big\langle \mathcal{R}^{\mathsf{ic}}(\boldsymbol{U}_k), q - \mathfrak{I}_{\mathbb{Q}}^k q\big\rangle \lesssim \|h_{\mathcal{G}_i}\|_{\infty, \Omega_i^\star} + \mathcal{E}_{\mathcal{G}_k}^{\mathsf{ic}}\big(\boldsymbol{U}_k; \mathcal{G}_i^+\big)^{1/t'}.$$

Thanks to (6.20), for $\epsilon > 0$ there exists an $i \in \mathbb{N}$ such that

$$\Big\langle \mathcal{R}\big(\boldsymbol{U}_k, P_k, \boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k)\big), (\boldsymbol{v}, q)\Big\rangle \lesssim \epsilon + \mathcal{E}_{\mathcal{G}_k}\big(\boldsymbol{U}_k, P_k, \boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k); \mathcal{G}_i^+\big),$$

and it therefore remains to prove that

$$\mathcal{E}_{\mathcal{G}_k}\big(\boldsymbol{U}_k, P_k, \boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k); \mathcal{G}_i^+\big) \to 0 \quad \text{as } k \to \infty \tag{6.21}$$

in order to deduce (6.18). To this end, let

$$\mathcal{E}_k := \mathcal{E}_{\mathcal{G}_k}\big(\boldsymbol{U}_k, P_k, \boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k); E_k\big) := \max\Big\{\mathcal{E}_{\mathcal{G}_k}\big(\boldsymbol{U}_k, P_k, \boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k); E\big) : E \in \mathcal{M}_k\Big\}.$$

Then, by the stability estimate, Corollary 4.6, and (4.4b) we have that

$$\begin{aligned}\mathcal{E}_k &\lesssim \|\boldsymbol{U}_k\|_{1,t;\mathcal{U}^{\mathcal{G}}(E_k)}^{\tilde{t}} + \|\boldsymbol{U}_k\|_{1,t;\mathcal{U}^{\mathcal{G}}(E_k)}^{2\tilde{t}} + \|P_k\|_{\tilde{t};\mathcal{U}^{\mathcal{G}}(E_k)}^{\tilde{t}} + \|\boldsymbol{f}\|_{\tilde{t},\mathcal{U}^{\mathcal{G}}(E_k)}^{\tilde{t}} + \left\|\tilde{k}\right\|_{r,\mathcal{U}^{\mathcal{G}}(E_k)}^{\tilde{t}} + \|\operatorname{div}\boldsymbol{U}_k\|_{t;\mathcal{U}^{\mathcal{G}}(E_k)}^{t}\\ &\lesssim \|\boldsymbol{U}_k-\boldsymbol{u}_\infty\|_{1,t;\mathcal{U}^{\mathcal{G}}(E_k)}^{\tilde{t}} + \|\boldsymbol{U}_k-\boldsymbol{u}_\infty\|_{1,t;\mathcal{U}^{\mathcal{G}}(E_k)}^{2\tilde{t}} + \|P_k-p_\infty\|_{\tilde{t};\mathcal{U}^{\mathcal{G}}(E_k)}^{\tilde{t}} + \|\operatorname{div}\boldsymbol{U}_k\|_{t;\mathcal{U}^{\mathcal{G}}(E_k)}^{t}\\ &\quad + \|\boldsymbol{u}_\infty\|_{1,t;\mathcal{U}^{\mathcal{G}}(E_k)}^{\tilde{t}} + \|\boldsymbol{u}_\infty\|_{1,t;\mathcal{U}^{\mathcal{G}}(E_k)}^{2\tilde{t}} + \|p_\infty\|_{\tilde{t};\mathcal{U}^{\mathcal{G}}(E_k)}^{\tilde{t}} + \|\boldsymbol{f}\|_{\tilde{t},\mathcal{U}^{\mathcal{G}}(E_k)}^{\tilde{t}} + \left\|\tilde{k}\right\|_{t,\mathcal{U}^{\mathcal{G}}(E_k)}^{\tilde{t}}.\end{aligned}$$

The first line of this bound vanishes thanks to (6.13). Since $E_k \in \Omega_k^0$, we have that $|E_k|^{1/d} \lesssim \|h_{\mathcal{G}_k}\|_{\infty;\Omega_k^0}$ and the remaining terms therefore vanish thanks to (6.20) and the observation that $E_k \in \Omega_k^0$. Therefore, we deduce with (5.1) that

$$\mathcal{E}_{\mathcal{G}_k}\big(\boldsymbol{U}_k, P_k, \boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k); \mathcal{G}_i^+\big) \leqslant \#\mathcal{G}_i^+ \max\left\{\mathcal{E}_{\mathcal{G}_k}\big(\boldsymbol{U}_k, P_k, \boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k); E\big) : E \in \mathcal{G}_i^+\right\} \leqslant \#\mathcal{G}_i^+ g(\mathcal{E}_k) \to 0$$

as $k \to \infty$, where we have used the continuity of g at zero and that $\mathcal{G}_i^+ \subset \mathcal{G}_k^+ \subset \mathcal{G}_k \backslash \mathcal{M}_k$. Combining these observations proves (6.18).

Step 4: In this step, we shall prove that

$$\mathcal{E}_{\mathcal{G}_k}\big(\boldsymbol{U}_k, P_k, \boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k)\big) \to 0 \qquad \text{as } k \to \infty.$$

To this end, we observe from (6.21) that it suffices to prove that

$$\mathcal{E}_{\mathcal{G}_k}\big(\boldsymbol{U}_k, P_k, \boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k); \mathcal{G}_k \backslash \mathcal{G}_i^+\big) \to 0 \qquad \text{as } k \to \infty$$

for some fixed $i \geqslant 0$. In view of Corollary 4.6, (6.17) and (6.18), it thus suffices to show that

$$\operatorname{osc}(\boldsymbol{U}_k, \boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k); \mathcal{G}_k \backslash \mathcal{G}_i^+) \to 0 \qquad \text{as } k \to \infty.$$

This is a consequence of the properties of the oscillation, (6.20), (6.13), and Assumption 3.5, noting that

$$\begin{aligned}\left\|\boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k) - \varPi_{\mathcal{G}_k}\boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k)\right\|_{\tilde{t};\Omega_i^0} &\leqslant \left\|\varPi_{\mathcal{G}_k}\boldsymbol{S}_\infty - \varPi_{\mathcal{G}_k}\boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k)\right\|_{\tilde{t};\Omega_i^0}\\ &\quad + \|\varPi_{\mathcal{G}_k}\boldsymbol{S}_\infty - \boldsymbol{S}_\infty\|_{\tilde{t};\Omega_i^0} + \left\|\boldsymbol{S}_\infty - \boldsymbol{S}^N(\mathrm{D}\boldsymbol{U}_k)\right\|_{\tilde{t};\Omega_i^0}.\end{aligned}$$

Observing that this readily implies that the estimator vanishes on the whole sequence completes the proof. ◻

7. Graph approximation

In this section we shall discuss the approximation of certain typical maximal monotone graphs satisfying Assumption 5.6. Admittedly, for particular problems the approximations suggested here might not always represent the best possible choices, and in the context of discrete nonlinear solvers, such as Newton's method, properties of the smoothness of the approximation may become important as well. We believe however that the following examples provide a reasonable guideline for constructing graph approximations with properties that are required in applications.

7.1. Discontinuous stresses

Typical examples of discontinuous dependence of the stress on the shear rate are Bingham or Herschel–Bulkley fluids. In this case, the fluid behaves like a rigid body when the shear stress is below a certain critical

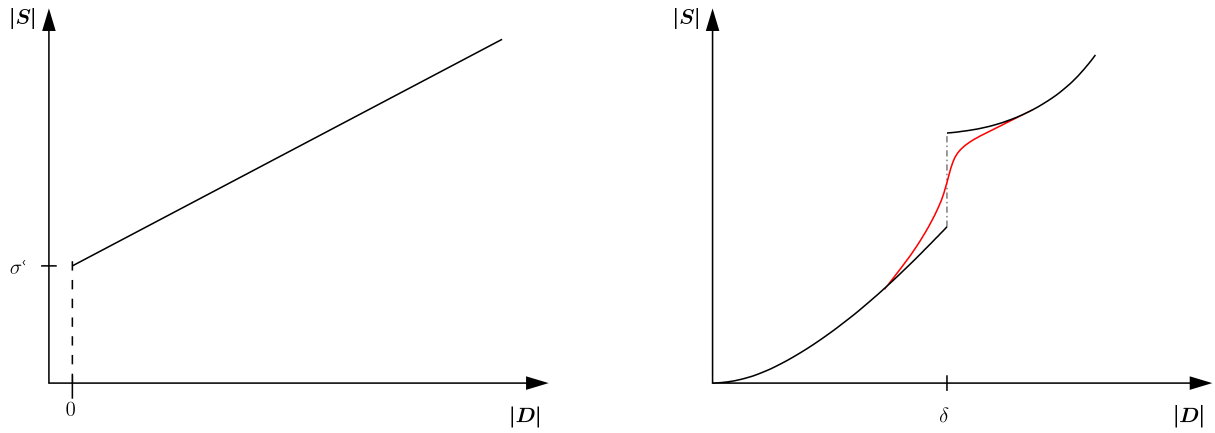

FIGURE 1. Bingham fluid (*left*) and schematic approximation of a more complex law (*right*).

value and like a Navier–Stokes fluid, respectively power-law fluid, otherwise; compare with Figure 1. To be more precise, for some yield stress $\sigma \geqslant 0$, we have

$$\begin{aligned} |\boldsymbol{S}| \leqslant \sigma \quad &\Leftrightarrow \quad \boldsymbol{D} = \boldsymbol{0}, \\ |\boldsymbol{S}| > \sigma \quad &\Leftrightarrow \quad \boldsymbol{S} = \sigma \frac{\boldsymbol{D}}{|\boldsymbol{D}|} + 2\nu(|\boldsymbol{D}|^2)\boldsymbol{D}, \end{aligned} \tag{7.1}$$

where $\nu > 0$ denotes the viscosity $\nu > 0$; see [23]. A selection of the corresponding maximal monotone graph is given, for example, by

$$\boldsymbol{S}^\star(\boldsymbol{D}) := S^\star(|\boldsymbol{D}|)\frac{\boldsymbol{D}}{|\boldsymbol{D}|}, \qquad \text{with } S^\star(D) := \begin{cases} 0, & \text{if } D = 0 \\ \sigma + 2\nu(D)D, & \text{otherwise.} \end{cases} \tag{7.2a}$$

For the sake of simplicity of presentation, we restrict ourselves in the following to $\nu > 0$ being a constant. However, we emphasize, that the approximation techniques presented below can be generalized to more complex relations such as, for example,

$$S^\star(D) = \begin{cases} S_1^\star(D), & \text{if } D < \delta \\ S_2^\star(D), & \text{otherwise,} \end{cases} \qquad S_i^\star(D) = c_i(\kappa_i^2 + D^2)^{\frac{q_i-2}{2}} D, \tag{7.2b}$$

for $D \geqslant 0$. Here $\delta \geqslant 0$ and $c_1, c_2, \kappa_1, \kappa_2 \geqslant 0$, $q_1 > 1$, $q_2 = r$, such that $S_1(\delta) \leqslant S_2(\delta)$.

We denote the maximal monotone graph containing $\{(D, S^\star(D)) \colon t \geqslant 0\}$ by $\mathfrak{a}$ and observe that $(\boldsymbol{\delta}, \boldsymbol{\sigma}) \in \mathcal{A}$ if and only if $(|\boldsymbol{\delta}|, |\boldsymbol{\sigma}|) \in \mathfrak{a}$. Therefore, the approximation of the monotone graph reduces to approximating the univariate function $S^\star$ by some smooth $S^n : \mathbb{R}_0^+ \to \mathbb{R}_0^+$. The explicit smooth approximation of $\boldsymbol{S}^\star$ is then obtained by setting

$$\boldsymbol{S}^n(\boldsymbol{D}) := S^n(|\boldsymbol{D}|)\frac{\boldsymbol{D}}{|\boldsymbol{D}|} \qquad \text{for all } \boldsymbol{D} \in \mathbb{R}^{d\times d}_{\text{sym}}. \tag{7.3}$$

A simple approximation. A simple approach to approximating $S^\star$ in (7.2a) is to use the following smooth explicit law (*cf.* [28]):

$$S^\tau(D) := \left(2\nu + \frac{\sigma}{D_\tau}\right)D, \qquad \text{where } D_\tau := \sqrt{D^2 + \tau^2}.$$

First assume that $\sigma \geqslant S^\tau(D)$; then, $(0, S^\tau(D)) \in \mathfrak{a}$ according to (7.1). If $D \leqslant \tau$, then

$$|S^\tau(D) - S^\tau(D)|^2 + |D - 0|^2 = |D|^2 \leqslant \tau^2.$$

Otherwise, we have

$$\sigma \geqslant \left(2\nu + \frac{\sigma}{D_\tau}\right) D \quad \Leftrightarrow \quad \sigma\tau^2 \geqslant 2\nu D\, D_\tau(D + D_\tau) \geqslant 4\nu D^3,$$

and hence

$$|S^\tau(D) - S^\tau(D)|^2 + |D - 0|^2 = |D|^2 \leqslant \left(\frac{\sigma}{4\nu}\right)^{1/3} \tau^{2/3}.$$

Assume now that $0 \leqslant \sigma < S^\tau(D)$; then, $D \leqslant \tau$ implies

$$2\nu D + \sigma \frac{D}{D_\tau} = S^\tau(D) > \sigma \quad \Leftrightarrow \quad \frac{2\nu}{\sigma} D > \frac{\tau^2}{D_\tau(D + D_\tau)} \geqslant \frac{1}{4}.$$

In other words this case can occur only for 'large' $\tau \geqslant \frac{\sigma}{8\nu}$. If $D \geqslant \tau$ then we have similarly

$$\sigma\tau^2 < 2\nu D\, D_\tau(D + D_\tau) \leqslant 8\nu D^3 \quad \Leftrightarrow \quad D > \tau^{2/3}\left(\frac{\sigma}{8\nu}\right)^{1/3}.$$

Therefore, we obtain

$$|S^\star(D) - S^\tau(D)| = \sigma \frac{D_\tau - D}{D_\tau} = \frac{\sigma\tau^2}{D_\tau(D + D_\tau)} \leqslant \frac{\sigma\tau^2}{D^2} < 4\nu^{2/3}\sigma^{1/3}\tau^{2/3}.$$

Combining the above cases shows the validity of Assumption 5.6 with $\tau = \frac{1}{n}$, for example. The verification of Assumption 3.2 is left to the reader.

Approximation by mollification. We can extend $S^\star$ to an odd function on the whole real axis by setting

$$S^\star(D) := -S^\star(-D) \quad \text{for } D < 0.$$

Then, for $n \in \mathbb{N}$, we define an approximation of $S^\star$ by

$$S^n(t) := \int_{-\infty}^{\infty} S^\star(s)\, \eta^n(s - t)\, \mathrm{d}s,$$

with $\eta^n(t) = n\eta(nt)$; here $\eta \in C^0(\mathbb{R})$ is a nonnegative even function with support $(-1, 1)$ such that $\int_{\mathbb{R}} \eta(s)\, \mathrm{d}s = 1$. Consequently, the function $S^n \in C(\mathbb{R})$ is odd and thus $S^n(0) = 0$.

For $D \in \mathbb{R}_0^+$ we have, by the monotonicity of $S^\star$ and the definition of the function S^n, that there exists a $D^\star$ with $0 \leqslant D^\star \in (D - \frac{1}{n}, D + \frac{1}{n})$, such that $(D^\star, S^n(D)) \in \mathfrak{a}$. Therefore, we have

$$|S^n(D) - S^n(D)|^{r'} + |D - D^\star|^r \leqslant 0 + \frac{1}{n^r},$$

and consequently

$$\mathcal{E}_{\mathcal{A}}(\boldsymbol{\delta}, \boldsymbol{S}^n(\boldsymbol{\delta})) \leqslant \frac{1}{n^r} \to 0 \quad \text{as } n \to \infty$$

for all $\boldsymbol{\delta} \in \mathbb{R}^{d \times d}_{\mathrm{sym}}$. This shows that Assumption 5.6 holds. Moreover, ϕ_n satisfies Assumption 3.2; compare *e.g.* with [14, 26, 27].

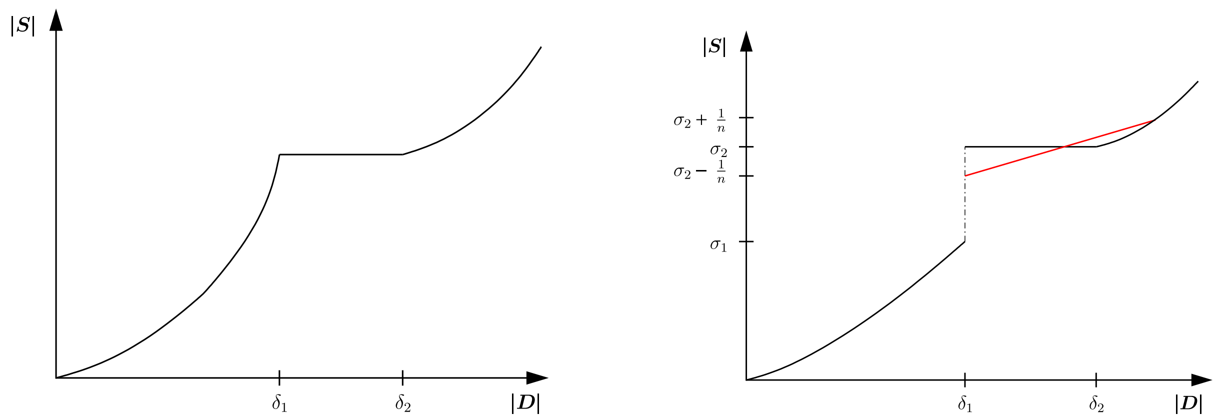

FIGURE 2. Graph with plateau (*left*) and plateau and jump (*right*).

7.2. Monotone graph with plateaus

Similarly to (7.2), we consider a maximal monotone graph with selection $\boldsymbol{S}^\star(\boldsymbol{D}) = S^\star(|\boldsymbol{D}|)\frac{\boldsymbol{D}}{|\boldsymbol{D}|}$, but now assume that $S^\star : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ is continuous with

$$S^\star(D) = \begin{cases} S_1^\star(D), & \text{if } D < \delta_1 \\ \sigma = \text{const.}, & \text{if } \delta_1 \leqslant D < \delta_2 \\ S_2^\star(D), & \text{else,} \end{cases}$$

with

$$S_i^\star(D) = c_i(\kappa_i^2 + D^2)^{\frac{q_i-2}{2}} D, \qquad i = 1, 2.$$

Here $c_1, c_2, \kappa_1, \kappa_2 \geqslant 0$, $q_1 > 1$, $q_2 = r$, such that $S_1^\star(\delta_1^\star) = S_2^\star = S_2^\star(\delta_2^\star)$; compare with Figure 2(*left*). In this case, we are basically in the same situation as in Section 7.1 with interchanged roles of S and D. Therefore, using the approximation techniques of Section 7.1, we can construct an approximation of the monotone graph where the shear rate depends explicitly on the shear stress. However, in a practical numerical method this relation typically has to be inverted, which may cause additional computational difficulties.

Another approach is to use an approximation of the form

$$\tilde{S}^n(D) := \begin{cases} S_1(D) & \text{if } S_1(D) < \sigma - \frac{1}{n} \\ S_2(D) & \text{if } S_1(D) > \sigma + \frac{1}{n} \\ S_\sigma^n(D) & \text{otherwise,} \end{cases}$$

where S_σ^n is the linear interpolant between $\sigma - \frac{1}{n}$ and $\sigma + \frac{1}{n}$ with corresponding values for D.

Combined with an approximation strategy as in Section 7.1, this procedure can also be applied to cases where jumps and plateaus are both present; compare with Figure 2(*right*).

Remark 7.1. The arguments of Sections 7.1 and 7.2 can be obviously extended to finitely many jumps/plateaus and even to cases with countably many jumps/plateaus.

References

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

[2] M. Ainsworth and J.T. Oden, *A posteriori error estimation in finite element analysis*. Wiley-Interscience, New York (2000).

[3] C.D. Aliprantis and K.C. Border, Infinite dimensional analysis, *A hitchhiker's guide*, 3rd edition. Springer, Berlin (2006).

[4] J.-P. Aubin and H. Frankowska, Set-valued analysis, *Modern Birkhäuser Classics*. Birkhäuser Boston Inc., Boston, MA (2009), Reprint of the 1990 edition.

[5] E. Bänsch, Local mesh refinement in 2 and 3 dimensions. *IMPACT Comput. Sci. Engrg.* **3** (1991) 181–191.

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

[7] L. Belenki, L. Diening and C. Kreuzer, Quasi-optimality of an adaptive finite element method for the p-Laplacian equation. *IMA J. Numer. Anal.* **32** (2012) 484–510.

[8] S. Berrone and E. Süli, Two-sided *a posteriori* error bounds for incompressible quasi-Newtonian flows. *IMA J. Numer. Anal.* **28** (2008) 382–421.

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

[10] D. Breit, L. Diening and M. Fuchs, Solenoidal Lipschitz truncation and applications in fluid mechanics. *J. Differ. Eq.* **253** (2012) 1910–1942.

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

[12] F. Brezzi and M. Fortin, Mixed and hybrid finite element methods. Vol. 15 of *Springer Series in Computational Mathematics* (1991).

[13] J.K. Brooks and R.V. Chacon, Continuity and compactness of measures. *Adv. Math.* **37** (1980) 16–26.

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

[15] M. Bulíček, P. Gwiazda, J. Málek, K.R. Rajagopal and A. Świerczewska-Gwiazda, On flows of fluids described by an implicit constitutive equation characterized by a maximal monotone graph. *Mathematical Aspects of Fluid Mechanics, London Mathematical Society Lecture Note Series* (No. 402). Cambridge University Press (2012) 23–51.

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

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

[18] L. Diening and C. Kreuzer, Convergence of an adaptive finite element method for the p-Laplacian equation. *SIAM J. Numer. Anal.* **46** (2008) 614–638.

[19] L. Diening, J. Málek, and M. Steinhauer, On Lipschitz truncations of Sobolev functions (with variable exponent) and their selected applications. *ESAIM: COCV* **14** (2008) 211–232.

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

[21] L. Diening, C. Kreuzer and E. Süli, Finite element approximation of steady flows of incompressible fluids with implicit power-law-like rheology. Prpeprint arXiv:1204.2145v3 (2012).

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

[23] G. Duvaut and J.-L. Lions, Inequalities in mechanics and physics. [Translated from the French by C.W. John. Grundlehren der Mathematischen Wissenschaften, 219.] Springer-Verlag, Berlin (1976).

[24] V. Girault and J.-L. Lions, Two-grid finite-element schemes for the steady Navier–Stokes problem in polyhedra. *Port. Math. (N.S.)* **58** (2001) 25–57.

[25] V. Girault and L.R. Scott, A quasi-local interpolation operator preserving the discrete divergence. *Calcolo* **40** (2003) 1–19.

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

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

[28] J. Hron, J. Málek, J. Stebel and K. Touška, A novel view of computations of steady flows of Bingham and Herschel–Bulkley fluids using implicit constitutive relations. *Proc. of the 26th Nordic Seminar on Computational Mechanics, Oslo, 23–25 October 2013*, edited by A. Logg, K.-A. Mardal and A. Massing. Center for Biomedical Computing, Simula Research Laboratory, Oslo (2013) 217–219.

[29] I. Kossaczký, A recursive approach to local mesh refinement in two and three dimensions. *J. Comput. Appl. Math.* **55** (1994) 275–288.

[30] C. Kreuzer and E. Süli, Adaptive finite element approximation of steady flows of incompressible fluids with implicit power-law-like rheology. Preprint arXiv:1503.05378 (2015).

[31] W. Liu and N. Yan, Quasi-norm local error estimators for p-Laplacian.. *SIAM J. Numer. Anal.* **39** (2001) 100–127.

[32] 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).

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

[34] P. Morin, K.G. Siebert, A. Veeser, A basic convergence result for conforming adaptive finite elements. *Math. Models Methods Appl.* **18** (2008) 707–737.

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

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

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

[38] K.G. Siebert, A convergence proof for adaptive finite elements without lower bound. *IMA J. Numer. Anal.* **31** (2011) 947–970.

[39] A. Schmidt and K.G. Siebert, Design of adaptive finite element software. *The finite element toolbox ALBERTA*. Vol. 42 of *Lect. Notes Comput. Sci. Engrg.* Springer (2005).

[40] R. Stevenson, The completion of locally refined simplicial partitions created by bisection. *Math. Comput.* **77** (2008) 227–241.

[41] R. Temam, Navier–Stokes equations. Theory and numerical analysis. Vol. 2 of *Studies in Mathematics and its Applications*, 3rd edition. North-Holland, Amsterdam, New York, Oxford (1984).

[42] R. Verfürth, A review of *a posteriori* error estimation and adaptive mesh-refinement techniques. *Adv. Numer. Math.* John Wiley, Chichester, UK (1996).

[43] R. Verfürth, A posteriori error estimation techniques for finite element methods. *Numerical Mathematics and Scientific Computation.* The Clarendon Press, Oxford University Press, New York (2013).