Finite element approximation of a strain-limiting elastic model

Andrea Bonito, Vivette Girault, Endre Süli

IMA Journal of Numerical Analysis (2020) **40**, 29–86
doi:10.1093/imanum/dry065
Advance Access publication on 29 October 2018

Finite element approximation of a strain-limiting elastic model

Andrea Bonito*
Department of Mathematics, Texas A&M University, College Station, TX 77843-3368, USA
*Corresponding author: bonito@math.tamu.edu

Vivette Girault
Sorbonne Université, Université Paris-Diderot SPC, CNRS, Laboratoire Jacques-Louis Lions, LJLL, F-75005 Paris
girault@ann.jussieu.fr

and

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

[Received on 10 May 2018; revised on 10 May 2018]

We construct a finite element approximation of a strain-limiting elastic model on a bounded open domain in $\mathbb{R}^d$, $d \in \{2, 3\}$. The sequence of finite element approximations is shown to exhibit strong convergence to the unique weak solution of the model. A rate of convergence for the sequence of finite element approximations is shown provided that the material parameters featuring in the model are Lipschitz continuous and that the exact solution possesses additional regularity. A rate of convergence for the sequence of finite element approximations is shown provided that the material parameters featuring in the model are Lipschitz continuous and that the exact solution possesses additional regularity. An iterative algorithm is constructed for the solution of the system of nonlinear algebraic equations that arises from the finite element approximation. An appealing feature of the iterative algorithm is that it decouples the monotone and linear elastic parts of the nonlinearity in the model. In particular, our choice of piecewise constant approximation for the stress tensor (and continuous piecewise linear approximation for the displacement) allows us to compute the monotone part of the nonlinearity by solving an algebraic system with $d(d+1)/2$ unknowns independently on each element in the subdivision of the computational domain. The theoretical results are illustrated by numerical experiments.

Keywords: strain-limiting elastic model; finite element method; convergence; decoupled iterative method.

1. Introduction and statement of the problem

Until recently, the term *elasticity* referred to *Cauchy elasticity*, and within such a theory, strain-limiting models are not possible. Motivated by the work of Rajagopal discussed in Málek *et al.* (1996), see also Rajagopal (2003, 2007), the objective of this paper is to design, analyze and implement numerical approximations of models that fall outside the realm of classical Cauchy elasticity. These models are implicit and nonlinear, and are referred to as strain limiting because they permit the linearized strain

to remain bounded even when the stress is very large, a property that cannot be guaranteed within the framework of standard elastic or nonlinear elastic models.

On a bounded domain $\Omega \subset \mathbb{R}^d$, $d \in \{2, 3\}$, and for a given external force $\mathbf{f} : \Omega \to \mathbb{R}^d$ we consider the nonlinear elastic model

$$-\operatorname{div}(\mathbf{T}) = \mathbf{f} \qquad \text{in } \Omega, \tag{1.1}$$

where the symmetric stress tensor $\mathbf{T}$ is related to the strain tensor $\varepsilon(\mathbf{u}) := \frac{1}{2}(\nabla\mathbf{u} + (\nabla\mathbf{u})^{\mathrm{T}})$, for a given displacement vector $\mathbf{u}$, via a nonlinear constitutive relation of the form

$$\varepsilon(\mathbf{u}) = \lambda(\operatorname{Tr}(\mathbf{T}))\operatorname{Tr}(\mathbf{T})\mathbf{I} + \mu(|\mathbf{T}^{\mathbf{d}}|)\mathbf{T}^{\mathbf{d}} \qquad \text{in } \Omega. \tag{1.2}$$

Here $\lambda \in \mathcal{C}^0(\mathbb{R})$ and $\mu \in \mathcal{C}^0([0, +\infty))$ are given functions and $\mathbf{T}^{\mathbf{d}}$ denotes the deviatoric part of the tensor $\mathbf{T}$, defined by

$$\mathbf{T}^{\mathbf{d}} := \mathbf{T} - \frac{1}{d}\operatorname{Tr}(\mathbf{T})\mathbf{I}.$$

Additional assumptions on λ and μ are required (see (A1–A4) below), which guarantee that, in particular, the right-hand side of (1.2) is a monotone operator applied to $\mathbf{T}$. This strain-limiting model is used to describe, for example, the behavior of brittle materials in the vicinity of fracture tips, or in the neighborhood of concentrated loads, where there is concentration of stress even though the magnitude of the strain tensor is limited. The model itself is derived and analyzed in the work of Bulíček *et al.* (2014); some of the ideas introduced in Bulíček *et al.* (2014) will also be used in the numerical analysis developed in the sequel. Of course, there are several strain-limiting models: the reader will find other models in Bulíček *et al.* (2014) and the references quoted therein. This being the first effort though to construct and rigorously analyze a numerical algorithm for a strain-limiting elastic model, we shall confine ourselves to the model (1.1), (1.2).

The analysis of the model (1.1), (1.2) is far from trivial because the operator involved, although monotone, lacks coercivity. Bulíček *et al.* (2014) show the existence of a weak solution to the problem by first regularizing (1.2) with the addition of an appropriate coercive term,

$$\frac{\operatorname{Tr}(\mathbf{T})\mathbf{I}}{n|\operatorname{Tr}(\mathbf{T})|^{1-\frac{1}{n}}} + \frac{\mathbf{T}^{\mathbf{d}}}{n|\mathbf{T}^{\mathbf{d}}|^{1-\frac{1}{n}}},$$

see (3.2), eventually providing a control of $\mathbf{T}$ in $L_{1+\frac{1}{n}}(\Omega)^{d\times d}$. It is then shown in Bulíček *et al.* (2014) that, as $n \to \infty$, the limit of the sequence of solutions to the regularized system satisfies the original problem. This nonlinear regularization is necessary in order to be able to cope with possibly rough data $\mathbf{f}$. However, for smoother data, the simpler *linear* regularization $\frac{1}{n}\mathbf{T}$ has been used in Bulíček *et al.* (2014) to recover additional regularity of the solution; see (6.3).

The same framework is used here in the discrete case. More precisely, the regularized problems (3.2) and (6.3) are discretized by means of a simple finite element scheme: for instance, on simplices, by discontinuous piecewise $\mathcal{P}_0$ elements for each component of the stress tensor $\mathbf{T}$, and globally continuous, piecewise $\mathcal{P}_1$ elements for each component of the displacement vector $\mathbf{u}$; see (5.3) and (6.10). It is worth noting here that for quadrilateral subdivisions of the domain Ω, the corresponding $(\mathcal{Q}_0, \mathcal{Q}_1)$ stress/displacement pair of finite element spaces is (inf-sup) unstable, and discontinuous polynomials of degree 1 in each direction should be selected for the stress approximations instead of $\mathcal{Q}_0$ elements so as to restore (inf-sup) stability; see Sections 5.3 and 5.4. Convergence to the exact solution is established by

first passing to the limit as the mesh size tends to zero, for a fixed value of the regularization parameter n, and then we let n tend to infinity. For rough data, the delicate part in the approximation of (3.2) is the derivation of a suitable rate of convergence for the approximation error. The difficulty stems from the lack of a meaningful error bound in a standard Lebesgue norm. Our analysis therefore relies on modular forms and associated Orlicz norms (see Theorem 5.5 and the subsequent discussion). For smoother data, the $\frac{1}{n}\mathbf{T}$ regularization mentioned above can be used, and the numerical analysis of (6.10) is then somewhat simpler because estimates for the stress, for the regularized problem at least, are naturally obtained in $L_2(\Omega)^{d\times d}$ (see Theorem 6.2) instead of $L_1(\Omega)^{d\times d}$ (or $L_{1+\frac{1}{n}}(\Omega)^{d\times d}$).

The proposed finite element discretizations (5.3) and (6.10) yield nonlinear systems with constraints. Since the nonlinear operator is the sum of a monotone and a coercive operator, we take advantage of the algorithm developed in Lions & Mercier (1979) to decouple these two parts: the unconstrained monotone system is solved first, followed by solving a constrained coercive system. As the stress tensor is potentially discontinuous, its simplest possible discretization is, as was suggested above, by means of a piecewise constant approximation on simplices; thus, the associated nonlinearity can be resolved element by element. We establish convergence of this splitting algorithm when applied to (6.10) (see Theorem 7.1). When applied to (5.3), the rigorous proof of convergence of the splitting algorithm is an open problem, although our numerical experiments at least appear to indicate that the splitting algorithm may well be convergent in this case as well.

1.1 *Setting of the problem*

We consider the system (1.1), (1.2) and describe the assumptions required on λ and μ. In addition to $\lambda \in \mathcal{C}^0(\mathbb{R})$ and $\mu \in \mathcal{C}^0([0,+\infty))$, we assume that $s \in \mathbb{R} \mapsto \lambda(s)s \in \mathcal{C}^1(\mathbb{R})$. Complementing these regularity hypotheses we assume that λ and μ satisfy, for some positive constants C_1, C_2, κ and α, the following inequalities:

$$\frac{C_1 s^2}{\kappa + |s|} \le \lambda(s)s^2 \le C_2|s| \qquad \forall\, s \in \mathbb{R}, \tag{A1}$$

$$\frac{C_1 s^2}{\kappa + s} \le \mu(s)s^2 \le C_2 s \qquad \forall\, s \in \mathbb{R}_{\ge 0}, \tag{A2}$$

$$0 \le \frac{\mathrm{d}}{\mathrm{d}s}(\lambda(s)s) \qquad \forall\, s \in \mathbb{R}, \tag{A3}$$

$$\frac{C_1}{(\kappa + s)^{\alpha+1}} \le \frac{\mathrm{d}}{\mathrm{d}s}(\mu(s)s) \qquad \forall\, s \in \mathbb{R}_{>0}. \tag{A4}$$

We note that, using the continuity of λ, the first inequality in assumption (A1) implies that $\lambda(s) > 0$ when $s \neq 0$ and $\lambda(s) \ge 0$ for $s \in \mathbb{R}$. In addition, now using the second inequality in (A1), we have

$$|\lambda(s)s| = \lambda(s)|s| \le C_2 \qquad \forall\, s \in \mathbb{R}. \tag{1.3}$$

The same argument applied to the function μ gives $\mu(s) > 0$ when $s > 0$, and

$$\mu(s) \ge 0, \quad \mu(s)\, s \le C_2 \qquad \forall\, s \in \mathbb{R}_{\ge 0}. \tag{1.4}$$

In particular, these assumptions guarantee that the system will exhibit finite strain only (see Theorem 2.1 below). At this point we also recall a result from Málek *et al.* (1996) (see also Bulíček *et al.* (2014), Lemma 4.1), which will play a crucial role in the subsequent analysis.

Under assumptions (A1–A4) stated above, there exists a positive constant C such that the following inequalities hold for all $\mathbf{R}_1, \mathbf{R}_2 \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$ (the set of all $d \times d$ symmetric matrices with real-valued entries):

$$\big(\mu(|\mathbf{R}_1|)\mathbf{R}_1 - \mu(|\mathbf{R}_2|)\mathbf{R}_2\big) : (\mathbf{R}_1 - \mathbf{R}_2) \geq C\frac{|\mathbf{R}_1 - \mathbf{R}_2|^2}{(\kappa + |\mathbf{R}_1| + |\mathbf{R}_2|)^{1+\alpha}}, \tag{1.5}$$

$$\big(\mu(|\mathbf{R}_1|)\mathbf{R}_1 - \mu(|\mathbf{R}_2|)\mathbf{R}_2\big) : (\mathbf{R}_1 - \mathbf{R}_2) \geq C\left|(\kappa + |\mathbf{R}_1|)^{\frac{1-\alpha}{2}} - (\kappa + |\mathbf{R}_2|)^{\frac{1-\alpha}{2}}\right|^2, \tag{1.6}$$

$$\big(\lambda(\mathrm{Tr}(\mathbf{R}_1))\mathrm{Tr}(\mathbf{R}_1) - \lambda(\mathrm{Tr}(\mathbf{R}_2))\mathrm{Tr}(\mathbf{R}_2)\big)\,(\mathrm{Tr}(\mathbf{R}_1) - \mathrm{Tr}(\mathbf{R}_2)) \geq 0. \tag{1.7}$$

If, in addition,

$$0 < \frac{\mathrm{d}}{\mathrm{d}s}(\lambda(s)s) \qquad \forall\, s \in \mathbb{R}, \tag{A3′}$$

then, for all $\mathbf{R}_1, \mathbf{R}_2 \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$ such that $\mathrm{Tr}(\mathbf{R}_1) \neq \mathrm{Tr}(\mathbf{R}_2)$, we have

$$\big(\lambda(\mathrm{Tr}(\mathbf{R}_1))\mathrm{Tr}(\mathbf{R}_1) - \lambda(\mathrm{Tr}(\mathbf{R}_2))\mathrm{Tr}(\mathbf{R}_2)\big)\,(\mathrm{Tr}(\mathbf{R}_1) - \mathrm{Tr}(\mathbf{R}_2)) > 0. \tag{1.8}$$

The system (1.1), (1.2) is supplemented with the boundary conditions

$$\mathbf{u} = \mathbf{g} \quad \text{on } \partial_D\Omega \qquad \text{and} \qquad \mathbf{T}\boldsymbol{\nu} = \boldsymbol{\ell} \quad \text{on } \partial_N\Omega,$$

where the boundary of Ω is decomposed into two parts, $\partial_D\Omega$ and $\partial_N\Omega$, with $\partial_D\Omega \cap \partial_N\Omega = \emptyset$ and $\overline{\partial_D\Omega \cup \partial_N\Omega} = \partial\Omega$, $\boldsymbol{\nu}$ is the outward-pointing unit normal to $\partial\Omega$, $\mathbf{g} : \partial\Omega \to \mathbb{R}^d$ is a given displacement on $\partial_D\Omega$ and $\boldsymbol{\ell} : \partial\Omega \to \mathbb{R}^d$ is a given traction force on $\partial_N\Omega$.

1.2 *Notation*

We shall suppose for the rest of this section that Ω is a bounded simply connected John domain; see, for instance, Acosta *et al.* (2006) or Jiang & Kauranen (2017). Henceforth, $L_p(\Omega)$ and $W^{k,p}(\Omega)$ will denote the standard Lebesgue and Sobolev spaces, and the corresponding spaces of d-component vector-valued functions and symmetric $(d \times d)$-component tensor-valued functions will be denoted, respectively by $L_p(\Omega)^d$, $L_p(\Omega)^{d\times d}_{\mathrm{sym}}$ and $W^{k,p}(\Omega)^d$, $W^{k,p}(\Omega)^{d\times d}_{\mathrm{sym}}$. In order to characterize displacements that vanish on the boundary, $\partial\Omega$, of Ω, we consider for $p \in [1, \infty)$ the Sobolev space $W^{1,p}_0(\Omega)$, defined as the closure of the linear space $\mathcal{C}^\infty_0(\Omega)$, consisting of infinitely many times continuously differentiable functions with compact support in Ω, in the norm of the space $W^{1,p}(\Omega)$:

$$W^{1,p}_0(\Omega) = \overline{\mathcal{C}^\infty_0(\Omega)}^{\|\cdot\|_{1,p}}.$$

We recall the Poincaré and Korn inequalities, which for each $p \in (1,\infty)$ assert the existence of positive constants $\mathcal{S}_p$ and $\mathcal{K}_p$ such that, respectively (cf. Jiang & Kauranen, 2017, Theorem 1.5),

$$\|v\|_{L_p(\Omega)} \le \mathcal{S}_p \, \|\nabla v\|_{L_p(\Omega)} \qquad \forall\, v \in W_0^{1,p}(\Omega), \tag{1.9}$$

$$\|\nabla \mathbf{v}\|_{L_p(\Omega)} \le \mathcal{K}_p \, \|\varepsilon(\mathbf{v})\|_{L_p(\Omega)} \qquad \forall\, \mathbf{v} \in W_0^{1,p}(\Omega)^d. \tag{1.10}$$

By combining inequalities (1.10) and (1.9) we obtain the inequality

$$\|\mathbf{v}\|_{L_p(\Omega)} \le C_{K,p} \|\varepsilon(\mathbf{v})\|_{L_p(\Omega)} \qquad \forall\, \mathbf{v} \in W_0^{1,p}(\Omega)^d, \tag{1.11}$$

with $C_{K,p} = \mathcal{S}_p \, \mathcal{K}_p > 0$.

For any two symmetric $d \times d$ tensors $\mathbf{S} = (S_{ij})$ and $\mathbf{T} = (T_{ij})$ we shall use a colon to denote their contraction product,

$$\mathbf{S} : \mathbf{T} = \sum_{i=1}^{d} \sum_{j=1}^{d} S_{ij} T_{ij},$$

so that the Frobenius norm of $\mathbf{S}$ reads

$$|\mathbf{S}|^2 = \mathbf{S} : \mathbf{S} = \mathrm{Tr}(\mathbf{S}^2).$$

It is then easy to show

$$|\mathbf{S}|^2 = |\mathbf{S^d}|^2 + \frac{1}{d}|\mathrm{Tr}(\mathbf{S})|^2 \le |\mathbf{S^d}|^2 + |\mathrm{Tr}(\mathbf{S})|^2 \qquad \forall\, \mathbf{S} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}, \tag{1.12}$$

which implies

$$|\mathbf{S}| \le |\mathbf{S^d}| + |\mathrm{Tr}(\mathbf{S})| \qquad \forall\, \mathbf{S} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}. \tag{1.13}$$

Conversely,

$$|\mathrm{Tr}(\mathbf{S})| + |\mathbf{S^d}| \le \sqrt{2d}\, |\mathbf{S}| \qquad \forall\, \mathbf{S} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}, \tag{1.14}$$

since, by elementary inequalities and by noting the equality stated in (1.12),

$$|\mathrm{Tr}(\mathbf{S})| + |\mathbf{S^d}| \le \sqrt{d}\left(\frac{1}{\sqrt{d}}|\mathrm{Tr}(\mathbf{S})| + |\mathbf{S^d}|\right) \le \sqrt{2d}\left(\frac{1}{d}|\mathrm{Tr}(\mathbf{S})|^2 + |\mathbf{S^d}|^2\right)^{\frac{1}{2}} = \sqrt{2d}\,|\mathbf{S}|.$$

Moreover, for any non-negative real numbers a and b, and for any $p \ge 1$ and $\theta \in (0,1]$, we have

$$a^p + b^p \ge 2^{1-p}(a+b)^p \ge 2^{1-p}(a^2+b^2)^{\frac{p}{2}} \ge 2^{1-p}(\theta a^2 + b^2)^{\frac{p}{2}}.$$

Thus, by taking $a = |\mathrm{Tr}(\mathbf{S})|$, $b = |\mathbf{S^d}|$ and $\theta = 1$, we have by (1.12) that, for any $p \ge 1$,

$$2^{1-p}|\mathbf{S}|^p \le |\mathrm{Tr}(\mathbf{S})|^p + |\mathbf{S^d}|^p. \tag{1.15}$$

The remainder of this article is organized as follows. The problem is set into variational form in Section 2 and the associated existence and uniqueness results are recalled. Sections 3 and 4 are devoted to the analysis of the sequence of regularized problems (3.2) that will be discretized by finite elements in Section 5; this includes *a priori* estimates, convergence and identification of the limit. The simpler analysis of (6.3) is sketched in Section 6. In Section 7 we present an iterative algorithm that dissociates the computation of the nonlinearity from the elastic constraint, and we prove its convergence when applied to (6.10). In Section 8 we report numerical experiments aimed at assessing the performance of the iterative algorithm and the discretization scheme.

2. Weak formulation

We begin by recalling Bulíček *et al.* (2014, Theorem 4.3), which guarantees the existence and uniqueness of a solution to problem (1.1), (1.2) in the case when $\partial_D\Omega = \partial\Omega$ and $\mathbf{g} = \mathbf{0}$.

When the Neumann part of the boundary $\partial_N\Omega$ is nonempty, the structure of the solution is potentially much more complicated. It was shown in Beck *et al.* (2017) that, in general, the solution in that case belongs to the space of Radon measures, but if the problem is equipped with a so-called asymptotic radial structure, then the solution can in fact be understood as a standard weak solution, with one proviso: the attainment of the boundary value is penalized by a measure supported on $\partial_N\Omega$. For simplicity, in this initial effort to construct a provably convergent numerical algorithm for the problem under consideration, we shall therefore suppose henceforth that $\partial_D\Omega = \partial\Omega$ (i.e., $\partial_N\Omega = \emptyset$) and that the Dirichlet boundary datum is $\mathbf{g} = \mathbf{0}$ on $\partial\Omega$.

THEOREM 2.1 (Bulíček *et al.*, 2014, Theorem 4.3). Assume that $\partial_N\Omega = \emptyset$ and that λ, μ satisfy (A1–A4) with $0 \le \alpha < 1/d$; then the following statements hold:

(a) Assume that $\mathbf{f} = -\mathrm{div}(\mathbf{F})$ for $\mathbf{F} \in W^{\beta,1}(\Omega)^{d\times d}_{\mathrm{sym}}$ with $\beta \in (\alpha d, 1)$. Then there exists a pair $(\mathbf{T}, \mathbf{u})$ such that

$$\begin{aligned} \mathbf{T} &\in L_1(\Omega)^{d\times d}_{\mathrm{sym}}, \\ \mathbf{u} &\in W^{1,p}_0(\Omega)^d \qquad \forall p \in [1,\infty), \\ \varepsilon(\mathbf{u}) &\in L_\infty(\Omega)^{d\times d}_{\mathrm{sym}} \end{aligned}$$

is a weak solution in the sense that it satisfies

$$\int_\Omega \mathbf{T} : \varepsilon(\mathbf{w})\,\mathrm{d}\mathbf{x} = \int_\Omega \mathbf{F} : \varepsilon(\mathbf{w})\,\mathrm{d}\mathbf{x} \qquad \forall \mathbf{w} \in \mathcal{D}(\Omega)^d, \tag{2.1}$$

where $\mathcal{D}(\Omega)^d := C^\infty_0(\Omega)^d$, and the nonlinear relationship between the strain $\varepsilon(\mathbf{u})$ and the stress $\mathbf{T}$ stated in (1.2) holds almost everywhere in Ω.

(b) Moreover, if Ω has a continuous boundary, then equality (2.1) holds for all $\mathbf{w} \in W^{1,1}_0(\Omega)^d$ such that $\varepsilon(\mathbf{w}) \in L_\infty(\Omega)^{d\times d}_{\mathrm{sym}}$.

(c) In addition, $\mathbf{u}$ is unique and if λ satisfies assumption (A3′), then $\mathbf{T}$ is also unique.

(d) Furthermore, if $\mathbf{F}$ belongs to $W^{2,2}(\Omega)^{d\times d}_{\mathrm{sym}}$ then $\mathbf{T} \in W^{1,q}_{\mathrm{loc}}(\Omega)^{d\times d}_{\mathrm{sym}}$ with

$$q \begin{cases} := 2 - \frac{2(d-2)(1+\alpha)}{(d-2)(1+\alpha)+d(1-\alpha)} & \text{for } d \geq 3, \\ \in [1,2), & \text{arbitrary for } d = 2. \end{cases}$$

REMARK 2.2 We note in connection with part (d) of the above theorem that when $d = 3$ then

$$q = 2 - \frac{1+\alpha}{2-\alpha}$$

is a monotonic decreasing function of α. Thus, as $0 \leq \alpha < \frac{1}{3}$, we have $\frac{6}{5} < q \leq \frac{3}{2}$.

3. Analysis of a regularized problem

The proof of existence of weak solutions to the problem is based on constructing a sequence of solutions to a regularized problem, where the original stress–strain relationship (1.2) is modified to become

$$\varepsilon(\mathbf{u}) = \lambda(\mathrm{Tr}(\mathbf{T}))\mathrm{Tr}(\mathbf{T})\mathbf{I} + \mu(|\mathbf{T}^{\mathbf{d}}|)\mathbf{T}^{\mathbf{d}} + \frac{\mathrm{Tr}(\mathbf{T})\mathbf{I}}{n|\mathrm{Tr}(\mathbf{T})|^{1-\frac{1}{n}}} + \frac{\mathbf{T}^{\mathbf{d}}}{n|\mathbf{T}^{\mathbf{d}}|^{1-\frac{1}{n}}}; \tag{3.1}$$

here $n \in \mathbb{N}$ (where $\mathbb{N}$ denotes the set of all positive integers) is a regularization parameter, which we shall ultimately send to the limit $n \to \infty$.

Following this idea, we study in this work the finite element approximation of this regularized problem, stated in the following variational form: find $(\mathbf{T}_n, \mathbf{u}_n) \in \mathbb{M}_n \times \mathbb{X}_n$ satisfying

$$\begin{aligned} a_n(\mathbf{T}_n, \mathbf{S}) + c(\mathbf{T}_n; \mathbf{T}_n, \mathbf{S}) - b(\mathbf{S}, \mathbf{u}_n) &= 0 && \forall\, \mathbf{S} \in \mathbb{M}_n, \\ b(\mathbf{T}_n, \mathbf{v}) &= \int_\Omega \mathbf{F} : \varepsilon(\mathbf{v})\, \mathrm{d}\mathbf{x} && \forall\, \mathbf{v} \in \mathbb{X}_n, \end{aligned} \tag{3.2}$$

where

$$\begin{aligned} a_n(\mathbf{T}, \mathbf{S}) &:= \frac{1}{n}\int_\Omega \left(\frac{\mathrm{Tr}(\mathbf{T})\mathbf{I}}{|\mathrm{Tr}(\mathbf{T})|^{1-\frac{1}{n}}} + \frac{\mathbf{T}^{\mathbf{d}}}{|\mathbf{T}^{\mathbf{d}}|^{1-\frac{1}{n}}} \right) : \mathbf{S}\, \mathrm{d}\mathbf{x}, \\ c(\mathbf{T}; \mathbf{R}, \mathbf{S}) &:= \int_\Omega \left(\lambda(\mathrm{Tr}(\mathbf{T}))\mathrm{Tr}(\mathbf{R})\mathbf{I} + \mu(|\mathbf{T}^{\mathbf{d}}|)\mathbf{R}^{\mathbf{d}} \right) : \mathbf{S}\, \mathrm{d}\mathbf{x}, \\ b(\mathbf{S}, \mathbf{v}) &:= \int_\Omega \mathbf{S} : \varepsilon(\mathbf{v})\, \mathrm{d}\mathbf{x} \end{aligned}$$

and

$$\mathbb{M}_n := L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\mathrm{sym}}, \qquad \mathbb{X}_n := W^{1,n+1}_0(\Omega)^d, \qquad n \in \mathbb{N}.$$

Motivated by the form of the expression appearing on the right-hand side of the relationship (3.1), we define the mapping $\mathcal{A}_n : L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\mathrm{sym}} \to L_{n+1}(\Omega)^{d\times d}_{\mathrm{sym}}$ by

$$\mathcal{A}_n(\mathbf{S}) := \lambda(\mathrm{Tr}(\mathbf{S}))\mathrm{Tr}(\mathbf{S})\mathbf{I} + \mu(|\mathbf{S^d}|)\mathbf{S^d} + \frac{\mathrm{Tr}(\mathbf{S})\mathbf{I}}{n|\mathrm{Tr}(\mathbf{S})|^{1-\frac{1}{n}}} + \frac{\mathbf{S^d}}{n|\mathbf{S^d}|^{1-\frac{1}{n}}}. \tag{3.3}$$

It follows from inequalities (1.3) and (1.4) that $\mathcal{A}_n$ does indeed take its values in $L_{n+1}(\Omega)^{d\times d}_{\mathrm{sym}}$, since the first two terms belong to $L_\infty(\Omega)^{d\times d}_{\mathrm{sym}}$ for all $\mathbf{S} \in L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\mathrm{sym}}$, while the third and fourth terms belong to $L_{n+1}(\Omega)^{d\times d}_{\mathrm{sym}}$ for all $\mathbf{S} \in L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\mathrm{sym}}$, $n \in \mathbb{N}$. Moreover, the mapping $\mathcal{A}_n : L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\mathrm{sym}} \to L_{n+1}(\Omega)^{d\times d}_{\mathrm{sym}}$ is bounded, continuous and coercive for all $n \in \mathbb{N}$, as is asserted in the following lemma.

LEMMA 3.1 (Boundedness, continuity and coercivity of $\mathcal{A}_n$). Let $\lambda \in C^0(\mathbb{R})$ and $\mu \in C^0([0,+\infty))$ and suppose that hypotheses (A1) and (A2) are valid. Then the following assertions hold:

(i) For any $n \in \mathbb{N}$, the mapping $\mathcal{A}_n : L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\mathrm{sym}} \to L_{n+1}(\Omega)^{d\times d}_{\mathrm{sym}}$ is bounded; i.e., every bounded set in $L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\mathrm{sym}}$ is mapped by $\mathcal{A}_n$ into a bounded set in $L_{n+1}(\Omega)^{d\times d}_{\mathrm{sym}}$.

(ii) For any $n \in \mathbb{N}$, the mapping $\mathcal{A}_n : L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\mathrm{sym}} \to L_{n+1}(\Omega)^{d\times d}_{\mathrm{sym}}$ is continuous, i.e., for any sequence $(\mathbf{S}_k)_{k>0} \subset L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\mathrm{sym}}$, which strongly converges in the norm of $L_{1+\frac{1}{n}}(\Omega)^{d\times d}$ to some $\mathbf{S} \in L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\mathrm{sym}}$, we have

$$\mathcal{A}_n(\mathbf{S}_k) \to \mathcal{A}_n(\mathbf{S}) \qquad \text{strongly in } L_{n+1}(\Omega)^{d\times d}_{\mathrm{sym}}.$$

(iii) For any $n \in \mathbb{N}$, the mapping $\mathcal{A}_n$ is coercive, i.e.,

$$\frac{\int_\Omega \mathcal{A}_n(\mathbf{S}) : \mathbf{S}\,\mathrm{d}\mathbf{x}}{\|\mathbf{S}\|_{L_{1+\frac{1}{n}}(\Omega)}} \to \infty \qquad \text{as } \|\mathbf{S}\|_{L_{1+\frac{1}{n}}(\Omega)} \to \infty.$$

Proof. (i) It suffices to prove that any bounded ball in $L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\mathrm{sym}}$, centered at the origin, is mapped by $\mathcal{A}_n$ into a bounded set in $L_{n+1}(\Omega)^{d\times d}_{\mathrm{sym}}$. Consider, to this end, the bounded ball

$$B_R := \left\{\mathbf{S} \in L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\mathrm{sym}} : \|\mathbf{S}\|_{L_{1+\frac{1}{n}}(\Omega)} \le R\right\} \quad \text{with } R > 0.$$

For every $\mathbf{S} \in B_R$ we have

$$\begin{aligned}\|\mathcal{A}_n(\mathbf{S})\|_{L_{n+1}(\Omega)} &\le C_2 d^{\frac{1}{2}}|\Omega|^{\frac{1}{n+1}} + C_2|\Omega|^{\frac{1}{n+1}} + \frac{1}{n}d^{\frac{1}{2}}\|\mathrm{Tr}(\mathbf{S})\|^{\frac{1}{n}}_{L_{1+\frac{1}{n}}(\Omega)} + \frac{1}{n}\|\mathbf{S^d}\|^{\frac{1}{n}}_{L_{1+\frac{1}{n}}(\Omega)} \\ &\le C_2 d^{\frac{1}{2}}|\Omega|^{\frac{1}{n+1}} + C_2|\Omega|^{\frac{1}{n+1}} + \frac{1}{n}d^{\frac{1}{2}(1+\frac{1}{n})}\|\mathbf{S}\|^{\frac{1}{n}}_{L_{1+\frac{1}{n}}(\Omega)} + \frac{1}{n}\|\mathbf{S}\|^{\frac{1}{n}}_{L_{1+\frac{1}{n}}(\Omega)},\end{aligned}$$

where in the transition to the second inequality we have made use of the facts that, by the identity (1.12), we have $|\mathrm{Tr}(\mathbf{S})| \le d^{\frac{1}{2}}|\mathbf{S}|$ and $|\mathbf{S}^{\mathbf{d}}|^2 = |\mathbf{S}|^2 - \frac{1}{d}(\mathrm{Tr}(\mathbf{S}))^2$, whereby $|\mathbf{S}^{\mathbf{d}}| \le |\mathbf{S}|$. Hence,

$$\|\mathcal{A}_n(\mathbf{S})\|_{L_{n+1}(\Omega)} \le C_2 d^{\frac{1}{2}}|\Omega|^{\frac{1}{n+1}} + C_2|\Omega|^{\frac{1}{n+1}} + \frac{1}{n}d^{\frac{1}{2}\left(1+\frac{1}{n}\right)}R^{\frac{1}{n}} + \frac{1}{n}R^{\frac{1}{n}} =: R_*,$$

which implies that $\mathcal{A}_n(B_R)$ is contained in a bounded ball in $L_{n+1}(\Omega)^{d\times d}_{\mathrm{sym}}$, centered at the origin, of radius R_*. Thus, $\mathcal{A}_n : L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\mathrm{sym}} \to L_{n+1}(\Omega)^{d\times d}_{\mathrm{sym}}$ is a bounded mapping.

(ii) Suppose that $\mathbf{S}_k \to \mathbf{S}$ strongly in $L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\mathrm{sym}}$. We begin by showing that

$$\lambda(\mathrm{Tr}(\mathbf{S}_k))\mathrm{Tr}(\mathbf{S}_k)\mathbf{I} \to \lambda(\mathrm{Tr}(\mathbf{S}))\mathrm{Tr}(\mathbf{S})\mathbf{I} \quad \text{strongly in } L_{n+1}(\Omega)^{d\times d}_{\mathrm{sym}}.$$

By defining $\varphi_k := \lambda(\mathrm{Tr}(\mathbf{S}_k))\mathrm{Tr}(\mathbf{S}_k)$ and using (1.3) we get

$$|\varphi_k| \le C_2 \qquad \text{a.e. in } \Omega.$$

Now, the strong convergence of $\{\mathbf{S}_k\}_{k>0}$ in $L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\mathrm{sym}}$ implies that there exists a subsequence (not indicated) such that $\mathbf{S}_k \to \mathbf{S}$ a.e. on Ω. Thanks to the assumed continuity of λ, it then follows that

$$\varphi_k \to \varphi := \lambda(\mathrm{Tr}(\mathbf{S}))\mathrm{Tr}(\mathbf{S})$$

a.e. in Ω and $|\varphi| \le C_2$ a.e. in Ω. By Lebesgue's dominated convergence theorem we therefore have that $\varphi_k \to \varphi$ strongly in $L_1(\Omega)$. When combined with the boundedness of φ_k, the strong convergence $\varphi_k \to \varphi$ in $L_1(\Omega)$ implies that $\varphi_k \to \varphi$ strongly in $L_p(\Omega)$ for all $p \in [1,\infty)$. Therefore, taking $p = n+1$, the first term of $\mathcal{A}_n(\mathbf{S}_k)$ strongly converges in $L_{n+1}(\Omega)^{d\times d}_{\mathrm{sym}}$ to the first term in $\mathcal{A}_n(\mathbf{S})$. The same is true of the second term.

To handle the third term we note that since, for any $a \in (0,1]$,

$$\left|\frac{x}{|x|}|x|^a - \frac{y}{|y|}|y|^a\right| \le 2^{1-a}|x-y|^a \qquad \forall\, x,y \in \mathbb{R}\setminus\{0\},$$

it follows with $a = \frac{1}{n}$, $n \in \mathbb{N}$ that

$$\left|\frac{\mathrm{Tr}(\mathbf{S}_k)}{|\mathrm{Tr}(\mathbf{S}_k)|^{1-\frac{1}{n}}} - \frac{\mathrm{Tr}(\mathbf{S})}{|\mathrm{Tr}(\mathbf{S})|^{1-\frac{1}{n}}}\right| \le 2^{1-\frac{1}{n}}|\mathrm{Tr}(\mathbf{S}_k) - \mathrm{Tr}(\mathbf{S})|^{\frac{1}{n}} \le 2^{1-\frac{1}{n}}d^{\frac{1}{2n}}\,|\mathbf{S}_k - \mathbf{S}|^{\frac{1}{n}}, \tag{3.4}$$

whereby the assumed strong convergence $\mathbf{S}_k \to \mathbf{S}$ in $L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\mathrm{sym}}$ implies

$$\frac{\mathrm{Tr}(\mathbf{S}_k)}{|\mathrm{Tr}(\mathbf{S}_k)|^{1-\frac{1}{n}}}\mathbf{I} \to \frac{\mathrm{Tr}(\mathbf{S})}{|\mathrm{Tr}(\mathbf{S})|^{1-\frac{1}{n}}}\mathbf{I}$$

in $L_{n+1}(\Omega)^{d\times d}_{\mathrm{sym}}$, $n \in \mathbb{N}$. By an identical argument the fourth term strongly converges in $L_{n+1}(\Omega)^{d\times d}_{\mathrm{sym}}$, $n \in \mathbb{N}$.

(iii) Note that, by assumptions (A1) and (A2),

$$\int_\Omega \mathcal{A}_n(\mathbf{S}) : \mathbf{S}\,\mathrm{d}\mathbf{x} \geq \frac{1}{n}\int_\Omega \left(|\mathrm{Tr}(\mathbf{S})|^{1+\frac{1}{n}} + |\mathbf{S^d}|^{1+\frac{1}{n}}\right)\mathrm{d}\mathbf{x}.$$

By taking $p = 1 + \frac{1}{n}$ with $n \in \mathbb{N}$ in (1.15) and using (1.15), we then have

$$\int_\Omega \mathcal{A}_n(\mathbf{S}) : \mathbf{S}\,\mathrm{d}\mathbf{x} \geq \frac{2^{-\frac{1}{n}}}{n}\int_\Omega |\mathbf{S}|^{\frac{n+1}{n}}\,\mathrm{d}\mathbf{x} = \frac{2^{-\frac{1}{n}}}{n}\|\mathbf{S}\|^{1+\frac{1}{n}}_{L_{1+\frac{1}{n}}(\Omega)}.$$

As the exponent $1 + \frac{1}{n}$ appearing on the right-hand side of the last equality is strictly greater than 1 for all $n \in \mathbb{N}$, the coercivity of the mapping $\mathcal{A}_n : L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\mathrm{sym}} \to L_{n+1}(\Omega)^{d\times d}_{\mathrm{sym}}$ directly follows. □

REMARK 3.2 One can simplify the proof of the continuity of $\mathcal{A}_n$ asserted in Lemma 3.1(ii) by assuming that $s \mapsto \lambda(s)s$ and $s \mapsto \mu(s)s$ are globally Hölder-continuous functions over their respective domains of definition. The latter assumption will be required in Theorem 5.5 to deduce rates of convergence for the finite element approximation of the regularized problem; prior to that, we do not assume the global Hölder continuity of $s \mapsto \lambda(s)s$ and $s \mapsto \mu(s)s$.

LEMMA 3.3 (Monotonicity of $\mathcal{A}_n$). Assume that $\lambda \in C^0(\mathbb{R})$ and $\mu \in C^0([0,+\infty))$, and that hypotheses (A1–A4) are satisfied. Then, for any $n \in \mathbb{N}$, the mapping $\mathcal{A}_n : L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\mathrm{sym}} \to L_{n+1}(\Omega)^{d\times d}_{\mathrm{sym}}$ is monotone, i.e.,

$$\int_\Omega (\mathcal{A}_n(\mathbf{S}_1) - \mathcal{A}_n(\mathbf{S}_2)) : (\mathbf{S}_1 - \mathbf{S}_2)\,\mathrm{d}\mathbf{x} \geq 0 \tag{3.5}$$

for any pair of functions $\mathbf{S}_1, \mathbf{S}_2 \in L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\mathrm{sym}}$. Furthermore, monotonicity is strict, in the sense that equality holds if and only if $\mathbf{S}_1 = \mathbf{S}_2$ a.e. on Ω.

Proof. To prove the monotonicity of $\mathcal{A}_n$, note first that for any pair of matrices $\mathbf{S}, \mathbf{R} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$ one has

$$\left(\mathbf{S}|\mathbf{S}|^{\frac{1}{n}-1} - \mathbf{R}|\mathbf{R}|^{\frac{1}{n}-1}\right) : (\mathbf{S} - \mathbf{R}) \geq \frac{1}{n}|\mathbf{S} - \mathbf{R}|^2 \int_0^1 |\mathbf{R} + \theta(\mathbf{S} - \mathbf{R})|^{\frac{1}{n}-1}\,\mathrm{d}\theta \geq 0, \tag{3.6}$$

and since $n \geq 1$, the expression on the right-hand side is equal to 0 if and only if $\mathbf{S} = \mathbf{R}$. Similarly, for any $s, r \in \mathbb{R}$,

$$\left(s|s|^{\frac{1}{n}-1} - r|r|^{\frac{1}{n}-1}\right)(s - r) = \frac{1}{n}|s - r|^2 \int_0^1 |r + \theta(s - r)|^{\frac{1}{n}-1}\,\mathrm{d}\theta \geq 0, \tag{3.7}$$

and the expression on the right-hand side is equal to 0 if and only if $s = r$. Hence, and by noting inequalities (1.5) and (1.7), we have

$$\begin{aligned}\int_\Omega (\mathcal{A}_n(\mathbf{S}_1) - \mathcal{A}_n(\mathbf{S}_2)) : (\mathbf{S}_1 - \mathbf{S}_2)\,\mathrm{d}\mathbf{x} \\ \geq \frac{1}{n^2}|\mathrm{Tr}(\mathbf{S}_1 - \mathbf{S}_2)|^2 \int_0^1 \left|\mathrm{Tr}(\mathbf{S}_2 + \theta(\mathbf{S}_1 - \mathbf{S}_2))\right|^{\frac{1}{n}-1}\mathrm{d}\theta \\ + \frac{1}{n^2}|\mathbf{S}_1^{\mathbf{d}} - \mathbf{S}_2^{\mathbf{d}}|^2 \int_0^1 \left|\mathbf{S}_2^{\mathbf{d}} + \theta\left(\mathbf{S}_1^{\mathbf{d}} - \mathbf{S}_2^{\mathbf{d}}\right)\right|^{\frac{1}{n}-1}\mathrm{d}\theta.\end{aligned} \tag{3.8}$$

The expression on the right-hand side of this inequality is non-negative and it is equal to 0 if, and only if, $\mathrm{Tr}(\mathbf{S}_1) = \mathrm{Tr}(\mathbf{S}_2)$ a.e. on Ω and $\mathbf{S}_1^{\mathbf{d}} = \mathbf{S}_2^{\mathbf{d}}$ a.e. on Ω, that is, when $\mathbf{S}_1 = \mathbf{S}_2$ a.e. on Ω. □

4. *A priori* estimates for the regularized problem

Our aim in this section is to derive *a priori* estimates for the regularized problem (3.2). Clearly, problem (3.2) can be interpreted as a constrained system with (strictly) monotone nonlinearity. The constraint is the second equation in problem (3.2); it is linear and nonhomogeneous, and can be, as is usual in mixed variational problems, transformed into a homogenous constraint via an inf-sup property, which we state in the next lemma.

LEMMA 4.1 (Inf-sup property). The following inequality holds for all $n \in \mathbb{N}$:

$$\inf_{\mathbf{v}\in\mathbb{X}_n} \sup_{\mathbf{S}\in\mathbb{M}_n} \frac{b(\mathbf{S},\mathbf{v})}{\|\mathbf{S}\|_{L_{1+\frac{1}{n}}(\Omega)}\|\varepsilon(\mathbf{v})\|_{L_{n+1}(\Omega)}} \geq 1. \tag{4.1}$$

Proof. Given $\mathbf{v} \in \mathbb{X}_n = W_0^{1,n+1}(\Omega)^d$, it suffices to note that $\mathbf{R} = \varepsilon(\mathbf{v})|\varepsilon(\mathbf{v})|^{n-1} \in L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\mathrm{sym}}$ and that we have

$$b(\mathbf{R},\mathbf{v}) = \|\varepsilon(\mathbf{v})\|^{n+1}_{L_{n+1}(\Omega)} = \|\varepsilon(\mathbf{v})\|_{L_{n+1}(\Omega)}\|\varepsilon(\mathbf{v})\|^{n}_{L_{n+1}(\Omega)} = \|\varepsilon(\mathbf{v})\|_{L_{n+1}(\Omega)}\|\mathbf{R}\|_{L_{1+\frac{1}{n}}(\Omega)},$$

whence,

$$\sup_{\mathbf{S}\in\mathbb{M}_n} \frac{b(\mathbf{S},\mathbf{v})}{\|\mathbf{S}\|_{L_{1+\frac{1}{n}}(\Omega)}} \geq \|\varepsilon(\mathbf{v})\|_{L_{n+1}(\Omega)},$$

and the stated inf-sup property follows. □

We shall assume henceforth that, as in Theorem 2.1, $\mathbf{f} = -\mathrm{div}(\mathbf{F})$, with $\mathbf{F} \in W^{\beta,1}(\Omega)^{d\times d}_{\mathrm{sym}}$ and $\beta \in (\alpha d, 1)$ (recall that, by hypothesis, $0 < \alpha < \frac{1}{d}$); hence, by Sobolev embedding $\mathbf{F} \in L_{1+\frac{\beta}{d-\beta}}(\Omega)^{d\times d}_{\mathrm{sym}}$ whereby also $\mathbf{F} \in L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\mathrm{sym}} = \mathbb{M}_n$ for all $n \geq \frac{d}{\beta} - 1$ (consequently, $\mathbf{f} \in W^{-1,1+\frac{1}{n}}(\Omega)^d = (W_0^{1,n+1}(\Omega)^d)' = (\mathbb{X}_n)'$ for all $n \geq \frac{d}{\beta} - 1$), and we define

$$\mathbf{T}_n^{\mathbf{f}} := \mathbf{F}.$$

Clearly, the subscript $_n$ in the expression on the left-hand side of this equality is redundant, as $\mathbf{T}_n^{\mathbf{f}}$ is equal to $\mathbf{F}$ for all $n \geq \frac{d}{\beta} - 1$. We shall, however, continue to carry this redundant subscript in order to emphasize the fact that the problem, as a whole, is dependent on n. Should it be desired that $\mathbf{F} \in L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\mathrm{sym}}$ for all $n \in \mathbb{N}$, one can, instead, adopt the slightly stronger assumption that $\mathbf{F} \in W^{\beta,1}(\Omega)^{d\times d}_{\mathrm{sym}} \cap L_2(\Omega)^{d\times d}_{\mathrm{sym}}$.

The use of the function $\mathbf{T}_n^{\mathbf{f}}$ will allow us to lift the constraint imposed by the second equation in problem (3.2) by converting it into a homogeneous equation; we can then replace the first equation in (3.2) by one that is considered on a linear subspace $\mathbb{V}_n$ of $\mathbb{M}_n$, defined below, which we choose to be the kernel of the mapping $\mathrm{div} : \mathbb{M}_n \to (\mathbb{X}_n)'$.

Trivially,

$$\int_\Omega \mathbf{T}_n^{\mathbf{f}} : \varepsilon(\mathbf{v})\,\mathrm{d}\mathbf{x} = \int_\Omega \mathbf{F} : \varepsilon(\mathbf{v})\,\mathrm{d}\mathbf{x} \qquad \forall\, \mathbf{v} \in \mathbb{X}_n. \tag{4.2}$$

We define

$$\mathbb{V}_n := \left\{\mathbf{S} \in \mathbb{M}_n \ :\ b(\mathbf{S},\mathbf{v}) = 0 \quad \forall\, \mathbf{v} \in \mathbb{X}_n\right\} = \left\{\mathbf{S} \in \mathbb{M}_n \ :\ \mathrm{div}(\mathbf{S}) = \mathbf{0} \in (\mathbb{X}_n)'\right\} = \mathrm{Ker}(\mathrm{div}). \tag{4.3}$$

As $\mathbb{X}_n$ is a reflexive Banach space, transposition yields that the transpose $(-\mathrm{div})' : (\mathbb{X}_n)'' = \mathbb{X}_n \to (\mathbb{M}_n)'$ of the linear operator $-\mathrm{div} : \mathbb{M}_n \to (\mathbb{X}_n)'$ is $(-\mathrm{div})' = \varepsilon(\cdot)$. The annihilator $\mathbb{V}_n^\perp$ of $\mathbb{V}_n$ is, by definition,

$$\mathbb{V}_n^\perp := \{\ell \in (\mathbb{M}_n)' : \ell(\mathbf{S}) = 0 \quad \forall\, \mathbf{S} \in \mathbb{V}_n\}.$$

By the Riesz representation theorem the dual space $(\mathbb{M}_n)'$ of $L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\mathrm{sym}}$ is isometrically isomorphic to $L_{n+1}(\Omega)^{d\times d}_{\mathrm{sym}}$. Furthermore, since $-\mathrm{div} : \mathbb{M}_n \to (\mathbb{X}_n)'$ is a bounded linear operator, it is also a closed linear operator. Hence, by Banach's closed range theorem,

$$\begin{aligned}\mathbb{V}_n^\perp &:= \left\{\mathbf{R} \in L_{n+1}(\Omega)^{d\times d}_{\mathrm{sym}} \ :\ \int_\Omega \mathbf{R} : \mathbf{S}\,\mathrm{d}\mathbf{x} = 0 \quad \forall\, \mathbf{S} \in \mathbb{V}_n\right\} \\ &= [\mathrm{Ker}(\mathrm{div})]^\perp = [\mathrm{Ker}(-\mathrm{div})]^\perp = \mathrm{Range}((-\mathrm{div})') = \mathrm{Range}(\varepsilon(\cdot)).\end{aligned}$$

Furthermore, once again by the closed range theorem,

$$\mathrm{Range}(\mathrm{div}) = \mathrm{Range}(-\mathrm{div}) = [\mathrm{Ker}((-\mathrm{div})')]^\perp = [\mathrm{Ker}(\varepsilon(\cdot))]^\perp = [\{0\}]^\perp = (\mathbb{X}_n)',$$

where the penultimate equality follows from inequality (1.11).

Thanks to the definition of $\mathbf{T}_n^{\mathbf{f}}$,

$$\|\mathbf{T}_n^{\mathbf{f}}\|_{L_{1+\frac{1}{n}}(\Omega)} = \|\mathbf{F}\|_{L_{1+\frac{1}{n}}(\Omega)}. \tag{4.4}$$

Using $\mathbf{T}_n^{\mathbf{f}}$, we can eliminate the constraint $(3.2)_2$ by setting

$$\mathbf{T}_n^0 := \mathbf{T}_n - \mathbf{T}_n^{\mathbf{f}} \in \mathbb{V}_n$$

and consider the following problem: find $\mathbf{T}_n^0 \in \mathbb{V}_n$ such that

$$a_n\big(\mathbf{T}_n^0 + \mathbf{T}_n^{\mathbf{f}}, \mathbf{S}\big) + c_n\big(\mathbf{T}_n^0 + \mathbf{T}_n^{\mathbf{f}}; \mathbf{T}_n^0 + \mathbf{T}_n^{\mathbf{f}}, \mathbf{S}\big) = 0 \qquad \forall\, \mathbf{S} \in \mathbb{V}_n. \tag{4.5}$$

From here, by using Lemmas 3.1 and 3.3, we easily deduce that the mapping

$$\mathbf{S} \in \mathbb{V}_n \mapsto \mathcal{A}_n\big(\mathbf{S} + \mathbf{T}_n^{\mathbf{f}}\big) \in L_{n+1}(\Omega)^{d\times d}_{\mathrm{sym}} = (\mathbb{M}_n)' \subset (\mathbb{V}_n)'$$

is bounded, continuous (and therefore hemicontinuous), coercive and monotone; in addition, $\mathbb{V}_n$ is a separable reflexive Banach space, as it is a closed linear subspace of the separable and reflexive Banach space $\mathbb{M}_n = L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\mathrm{sym}}$. Therefore, by the Browder–Minty theorem (cf., for instance, Lions, 1969; Showalter, 1997), problem (4.5), and hence also problem (3.2), has a solution $\mathbf{T}_n = \mathbf{T}_n^0 + \mathbf{T}_n^{\mathbf{f}} \in \mathbb{M}_n$, and since by Lemma 3.3 the operator $\mathcal{A}_n$ is strictly monotone, the solution is unique.

With $\mathbf{T}_n \in \mathbb{M}_n$ thus uniquely fixed we seek $\mathbf{u}_n \in \mathbb{X}_n$ such that

$$b(\mathbf{S}, \mathbf{u}_n) = a_n(\mathbf{T}_n, \mathbf{S}) + c(\mathbf{T}_n; \mathbf{T}_n, \mathbf{S}) \qquad \forall\, \mathbf{S} \in \mathbb{M}_n.$$

Consider the linear functional $\ell_n \in (\mathbb{M}_n)'$ defined by

$$\ell_n(\mathbf{S}) := a_n(\mathbf{T}_n, \mathbf{S}) + c(\mathbf{T}_n; \mathbf{T}_n, \mathbf{S}), \qquad \mathbf{S} \in \mathbb{M}_n.$$

Hence, thanks to equation (4.5), we have $\ell_n(\mathbf{S}) = 0$ for all $\mathbf{S} \in \mathbb{V}_n$; consequently, $\ell_n \in \mathbb{V}_n^{\perp}$. Thus, we are seeking $\mathbf{u}_n \in \mathbb{X}_n$ such that

$$b(\mathbf{S}, \mathbf{u}_n) = \ell_n(\mathbf{S}) \qquad \forall\, \mathbf{S} \in \mathbb{M}_n. \tag{4.6}$$

As $\ell_n \in \mathbb{V}_n^{\perp} = [\mathrm{Ker}(\mathrm{div})]^{\perp} = \mathrm{Range}(\varepsilon(\cdot))$, there exists a $\mathbf{u}_n \in \mathbb{X}_n$ such that $\varepsilon(\mathbf{u}_n) = \ell_n$; that is, $\mathbf{u}_n \in \mathbb{X}_n$ solves problem (4.6). The inf-sup property (4.1), together with inequality (1.11), then implies that such a $\mathbf{u}_n \in \mathbb{X}_n$ is unique. Thus, we have shown the existence of a unique solution pair $(\mathbf{T}_n, \mathbf{u}_n) \in \mathbb{M}_n \times \mathbb{X}_n$ to the regularized problem (3.2).

Next we shall prove the following *a priori* bounds on $\varepsilon(\mathbf{u}_n)$ and $\mathbf{T}_n$.

LEMMA 4.2 (*A priori* estimates). Suppose that $\mathbf{F} \in L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\mathrm{sym}}$, and that λ and μ satisfy properties (A1) and (A2). We then have

$$\|\varepsilon(\mathbf{u}_n)\|_{L_{n+1}(\Omega)} \le \frac{1}{n}\, d\sqrt{2}\left[16d^2\|\mathbf{F}\|^{1+\frac{1}{n}}_{L_{1+\frac{1}{n}}(\Omega)} + 2(n+1)C_2\sqrt{2d}\,|\Omega|^{\frac{1}{n+1}}\|\mathbf{F}\|_{L_{1+\frac{1}{n}}(\Omega)} + 4(n+1)C_1\kappa|\Omega|\right]^{\frac{1}{n+1}} + C_2\sqrt{2d}\,|\Omega|^{\frac{1}{n+1}}.$$

Moreover,

$$\frac{1}{n+1}\|\mathbf{T}_n\|^{1+\frac{1}{n}}_{L_{1+\frac{1}{n}}(\Omega)} + C_1\|\mathbf{T}_n\|_{L_1(\Omega)} \le \frac{16d^2}{n+1}\|\mathbf{F}\|^{1+\frac{1}{n}}_{L_{1+\frac{1}{n}}(\Omega)} + 2C_2\sqrt{2d}\,|\Omega|^{\frac{1}{n+1}}\|\mathbf{F}\|_{L_{1+\frac{1}{n}}(\Omega)} + 4C_1\kappa|\Omega|.$$

Proof. We start by testing problem (3.2) with $\mathbf{v} = \mathbf{u}_n$ and $\mathbf{S} = \mathbf{T}_n$ to get

$$\begin{aligned} a_n(\mathbf{T}_n, \mathbf{T}_n) + c(\mathbf{T}_n; \mathbf{T}_n, \mathbf{T}_n) - b(\mathbf{T}_n, \mathbf{u}_n) = 0, \\ b(\mathbf{T}_n, \mathbf{u}_n) = \int_\Omega \mathbf{F} : \varepsilon(\mathbf{u}_n)\, \mathrm{d}\mathbf{x}, \end{aligned} \tag{4.7}$$

whence, by substituting equation $(4.7)_2$ into equation $(4.7)_1$, we have

$$\begin{aligned} &\frac{1}{n}\int_\Omega \left(\frac{|\mathrm{Tr}(\mathbf{T}_n)|^2}{|\mathrm{Tr}(\mathbf{T}_n)|^{1-\frac{1}{n}}} + \frac{|\mathbf{T}_n^{\mathbf{d}}|^2}{|\mathbf{T}_n^{\mathbf{d}}|^{1-\frac{1}{n}}} \right) \mathrm{d}\mathbf{x} + \int_\Omega \left(\lambda(\mathrm{Tr}(\mathbf{T}_n))|\mathrm{Tr}(\mathbf{T}_n)|^2 + \mu(|\mathbf{T}_n^{\mathbf{d}}|)|\mathbf{T}_n^{\mathbf{d}}|^2 \right) \mathrm{d}\mathbf{x} \\ &\quad = \int_\Omega \mathbf{F} : \varepsilon(\mathbf{u}_n)\, \mathrm{d}\mathbf{x}, \end{aligned}$$

where we have used that

$$\mathbf{T}^{\mathbf{d}} : \mathbf{T} = \mathbf{T}^{\mathbf{d}} : \left(\mathbf{T}^{\mathbf{d}} + \frac{1}{d}\mathrm{Tr}(\mathbf{T})\mathbf{I} \right) = |\mathbf{T}^{\mathbf{d}}|^2 + \frac{1}{d}\mathrm{Tr}(\mathbf{T}) \underbrace{\mathrm{Tr}(\mathbf{T}^{\mathbf{d}})}_{=0} = |\mathbf{T}^{\mathbf{d}}|^2.$$

Hence, Hölder's inequality yields

$$\begin{aligned} &\frac{1}{n}\int_\Omega \left(|\mathrm{Tr}(\mathbf{T}_n)|^{1+\frac{1}{n}} + |\mathbf{T}_n^{\mathbf{d}}|^{1+\frac{1}{n}} \right) \mathrm{d}\mathbf{x} + \int_\Omega \left(\lambda(\mathrm{Tr}(\mathbf{T}_n))|\mathrm{Tr}(\mathbf{T}_n)|^2 + \mu(|\mathbf{T}_n^{\mathbf{d}}|)|\mathbf{T}_n^{\mathbf{d}}|^2 \right) \mathrm{d}\mathbf{x} \\ &\quad \le \|\mathbf{F}\|_{L_{1+\frac{1}{n}}(\Omega)} \|\varepsilon(\mathbf{u}_n)\|_{L_{n+1}(\Omega)}. \end{aligned}$$

For the λ and μ terms on the left-hand side of this inequality we note that for $s > 0$ one has

$$\frac{s^2}{\kappa + s} = s - \frac{\kappa}{1 + \frac{\kappa}{s}}.$$

This, together with properties (A1) and (A2), leads to

$$\begin{aligned} &\frac{1}{n}\int_\Omega \left(|\mathrm{Tr}(\mathbf{T}_n)|^{1+\frac{1}{n}} + |\mathbf{T}_n^{\mathbf{d}}|^{1+\frac{1}{n}} \right) \mathrm{d}\mathbf{x} + C_1 \int_\Omega \left(|\mathrm{Tr}(\mathbf{T}_n)| + |\mathbf{T}_n^{\mathbf{d}}| \right) \mathrm{d}\mathbf{x} \\ &\quad \le \|\mathbf{F}\|_{L_{1+\frac{1}{n}}(\Omega)} \|\varepsilon(\mathbf{u}_n)\|_{L_{n+1}(\Omega)} + C_1\kappa \int_\Omega \left(\frac{1}{1 + \frac{\kappa}{|\mathrm{Tr}(\mathbf{T}_n)|}} + \frac{1}{1 + \frac{\kappa}{|\mathbf{T}_n^{\mathbf{d}}|}} \right) \mathrm{d}\mathbf{x} \\ &\quad \le \|\mathbf{F}\|_{L_{1+\frac{1}{n}}(\Omega)} \|\varepsilon(\mathbf{u}_n)\|_{L_{n+1}(\Omega)} + 2C_1\kappa|\Omega|, \end{aligned}$$

since $\sup_{s>0} \frac{1}{1+\frac{\kappa}{s}} = 1$.

Moreover, it follows from inequality (1.15) with $p = 1 + \frac{1}{n}$ that

$$|\mathrm{Tr}(\mathbf{T}_n)|^{1+\frac{1}{n}} + |\mathbf{T}_n^{\mathbf{d}}|^{1+\frac{1}{n}} \ge 2^{-\frac{1}{n}} |\mathbf{T}_n|^{1+\frac{1}{n}},$$

and therefore (1.13) yields

$$\frac{2^{-\frac{1}{n}}}{n}\|\mathbf{T}_n\|^{1+\frac{1}{n}}_{L_{1+\frac{1}{n}}(\Omega)} + C_1\|\mathbf{T}_n\|_{L_1(\Omega)} \leq \|\mathbf{F}\|_{L_{1+\frac{1}{n}}(\Omega)}\|\varepsilon(\mathbf{u}_n)\|_{L_{n+1}(\Omega)} + 2C_1\kappa|\Omega|. \tag{4.8}$$

We now derive a bound on $\|\varepsilon(\mathbf{u}_n)\|_{L_{n+1}(\Omega)}$, using the inf-sup property (4.1). We begin by noting that

$$\|\varepsilon(\mathbf{u}_n)\|_{L_{n+1}(\Omega)} \leq \sup_{\mathbf{S}\in L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\mathrm{sym}}} \frac{b(\mathbf{S},\mathbf{u}_n)}{\|\mathbf{S}\|_{L_{1+\frac{1}{n}}(\Omega)}} = \sup_{\mathbf{S}\in L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\mathrm{sym}}} \frac{a_n(\mathbf{T}_n,\mathbf{S}) + c(\mathbf{T}_n;\mathbf{T}_n,\mathbf{S})}{\|\mathbf{S}\|_{L_{1+\frac{1}{n}}(\Omega)}}.$$

We invoke Hölder's inequality, the equality $|\mathbf{I}| = \sqrt{d}$, the elementary inequality $a+b \leq 2^{1-\frac{1}{n}}(a^n+b^n)^{\frac{1}{n}}$, where $a, b \geq 0$ and $n \in \mathbb{N}$ with $a = \sqrt{d}\,|\mathrm{Tr}(\mathbf{T}_n)|^{\frac{1}{n}}$, $b = |\mathbf{T}^{\mathbf{d}}_n|^{\frac{1}{n}}$ and note that $\frac{1}{\sqrt{d}}|\mathrm{Tr}(\mathbf{T}_n)|+|\mathbf{T}^{\mathbf{d}}_n| \leq \sqrt{2}\,|\mathbf{T}_n|$, to deduce that

$$a_n(\mathbf{T}_n,\mathbf{S}) \leq \frac{2\sqrt{d}}{n}\left(\frac{d}{2}\right)^{\frac{1}{2n}}\|\mathbf{T}_n\|^{\frac{1}{n}}_{L_{1+\frac{1}{n}}(\Omega)}\|\mathbf{S}\|_{L_{1+\frac{1}{n}}(\Omega)}.$$

Further, by noting properties (A1) and (A2) again, together with inequality (1.14), we have

$$c(\mathbf{T}_n;\mathbf{T}_n,\mathbf{S}) \leq C_2\sqrt{2d}\,|\Omega|^{\frac{1}{n+1}}\|\mathbf{S}\|_{L_{1+\frac{1}{n}}(\Omega)}, \tag{4.9}$$

where we have bounded $\sqrt{d+1}$ by $\sqrt{2d}$ for the sake of simplifying the constants appearing in the subsequent calculations. Hence,

$$\|\varepsilon(\mathbf{u}_n)\|_{L_{n+1}(\Omega)} \leq \frac{2\sqrt{d}}{n}\left(\frac{d}{2}\right)^{\frac{1}{2n}}\|\mathbf{T}_n\|^{\frac{1}{n}}_{L_{1+\frac{1}{n}}(\Omega)} + C_2\sqrt{2d}\,|\Omega|^{\frac{1}{n+1}}. \tag{4.10}$$

By substituting inequality (4.10) into inequality (4.8), we obtain

$$\frac{2^{-\frac{1}{n}}}{n}\|\mathbf{T}_n\|^{1+\frac{1}{n}}_{L_{1+\frac{1}{n}}(\Omega)} + C_1\|\mathbf{T}_n\|_{L_1(\Omega)} \leq \|\mathbf{F}\|_{L_{1+\frac{1}{n}}(\Omega)}\left(\frac{2\sqrt{d}}{n}\left(\frac{d}{2}\right)^{\frac{1}{2n}}\|\mathbf{T}_n\|^{\frac{1}{n}}_{L_{1+\frac{1}{n}}(\Omega)} + C_2\sqrt{2d}\,|\Omega|^{\frac{1}{n+1}}\right) + 2C_1\kappa|\Omega|;$$

thus, by applying Young's inequality,

$$ab \leq \varepsilon\frac{a^p}{p} + \varepsilon^{-\frac{1}{p-1}}\frac{b^q}{q} \qquad \text{for } a, b \geq 0, \varepsilon > 0, p > 1 \text{ and } \frac{1}{p} + \frac{1}{q} = 1,$$

to the first term on the right-hand side with $p = n+1$, $\varepsilon = \frac{1}{n}\,2^{-\frac{1}{n}}$,

$$a = \|\mathbf{T}_n\|^{\frac{1}{n}}_{L_{1+\frac{1}{n}}(\Omega)} \qquad \text{and} \qquad b = \frac{2\sqrt{d}}{n}\left(\frac{d}{2}\right)^{\frac{1}{2n}}\|\mathbf{F}\|_{L_{1+\frac{1}{n}}(\Omega)}$$

in order to absorb the factor $\|\mathbf{T}_n\|_{L_{1+\frac{1}{n}}(\Omega)}^{\frac{1}{n}}$ into the left-hand side, we deduce that

$$\frac{2^{-\frac{1}{n}}}{n+1}\|\mathbf{T}_n\|_{L_{1+\frac{1}{n}}(\Omega)}^{1+\frac{1}{n}} + C_1\|\mathbf{T}_n\|_{L_1(\Omega)} \leq \frac{2^{1+\frac{1}{2n}+\frac{1}{2n^2}}(\sqrt{d})^{(1+\frac{1}{n})^2}}{n+1}\|\mathbf{F}\|_{L_{1+\frac{1}{n}}(\Omega)}^{1+\frac{1}{n}} + C_2\sqrt{2d}\,|\Omega|^{\frac{1}{n+1}}\|\mathbf{F}\|_{L_{1+\frac{1}{n}}(\Omega)} + 2C_1\kappa|\Omega|.$$

Hence, after multiplying by $2^{\frac{1}{n}}$ and noting that $1 \leq 2^{\frac{1}{n}} \leq 2$ and $1 + \frac{3}{2n} + \frac{1}{2n^2} \leq \left(1+\frac{1}{n}\right)^2$, we obtain

$$\frac{1}{n+1}\|\mathbf{T}_n\|_{L_{1+\frac{1}{n}}(\Omega)}^{1+\frac{1}{n}} + C_1\|\mathbf{T}_n\|_{L_1(\Omega)} \leq \frac{(2\sqrt{d})^{(1+\frac{1}{n})^2}}{n+1}\|\mathbf{F}\|_{L_{1+\frac{1}{n}}(\Omega)}^{1+\frac{1}{n}} + 2C_2\sqrt{2d}\,|\Omega|^{\frac{1}{n+1}}\|\mathbf{F}\|_{L_{1+\frac{1}{n}}(\Omega)} + 4C_1\kappa|\Omega|.$$

Bounding $1+\frac{1}{n}$ by 2 in the exponent of $2\sqrt{d}$ in the first term on the right-hand side then yields the second inequality in the statement of the lemma.

Omitting the second term from the left-hand side of that inequality and multiplying by $n+1$ then yields

$$\|\mathbf{T}_n\|_{L_{1+\frac{1}{n}}(\Omega)}^{1+\frac{1}{n}} \leq 16d^2\|\mathbf{F}\|_{L_{1+\frac{1}{n}}(\Omega)}^{1+\frac{1}{n}} + 2(n+1)C_2\sqrt{2d}\,|\Omega|^{\frac{1}{n+1}}\|\mathbf{F}\|_{L_{1+\frac{1}{n}}(\Omega)} + 4(n+1)C_1\kappa|\Omega|.$$

Therefore, by inequality (4.10), we have

$$\|\varepsilon(\mathbf{u}_n)\|_{L_{n+1}(\Omega)} \leq \frac{2\sqrt{d}}{n}\left(\frac{d}{2}\right)^{\frac{1}{2n}}\left[16d^2\|\mathbf{F}\|_{L_{1+\frac{1}{n}}(\Omega)}^{1+\frac{1}{n}} + 2(n+1)C_2\sqrt{2d}\,|\Omega|^{\frac{1}{n+1}}\|\mathbf{F}\|_{L_{1+\frac{1}{n}}(\Omega)} + 4(n+1)C_1\kappa|\Omega|\right]^{\frac{1}{n+1}} + C_2\sqrt{2d}\,|\Omega|^{\frac{1}{n+1}}.$$

Bounding the exponent $\frac{1}{2n}$ by $\frac{1}{2}$ in the prefactor on the right-hand side yields the first bound in the lemma. □

Lemma 4.2 implies in particular that

$$\limsup_{n\to\infty}\|\varepsilon(\mathbf{u}_n)\|_{L_{n+1}(\Omega)} \leq C_2\sqrt{2d}$$

and

$$\limsup_{n\to\infty}\|\mathbf{T}_n\|_{L_1(\Omega)} \leq \frac{2C_2}{C_1}\sqrt{2d}\,\|\mathbf{F}\|_{L_1(\Omega)} + 4\kappa|\Omega|.$$

These bounds are consistent with the properties of the strain-limiting model under consideration, expressed by Theorem 2.1(a), which asserts that the strain tensor is contained in $L_\infty(\Omega)^{d\times d}_{\rm sym}$, even though the stress tensor is, in general, an element of $L_1(\Omega)^{d\times d}_{\rm sym}$ only.

In connection with this we recall from Bulíček *et al.* (2014, Section 4) that the sequence of (unique) weak solution pairs $((\mathbf{T}_n,\mathbf{u}_n))_{n\in\mathbb{N}}$ to the regularized problem (3.2) converges to a weak solution pair $(\mathbf{T},\mathbf{u})$ of problem (1.1), (1.2), supplemented by a homogeneous Dirichlet boundary condition on $\partial\Omega$ (which is also unique if the condition (A3′) holds), in the sense that, as $n\to\infty$,

$$\mathbf{T}_n \to \mathbf{T} \text{ strongly in } L_q(\Omega_0)^{d\times d}_{\rm sym} \text{ for any } q\in\left[1,1+\tfrac{1}{2}\tfrac{\beta-\alpha d}{d-\beta}\right),\ \beta\in(\alpha d,1),\ 0\le\alpha<\tfrac{1}{d},\ \Omega_0\subset\subset\Omega; \tag{4.11}$$

furthermore,

$$\begin{aligned}
\mathbf{u}_n &\rightharpoonup \mathbf{u} && \text{weakly in } W^{1,2d}_0(\Omega)^d,\\
\mathbf{u}_n &\to \mathbf{u} && \text{strongly in } \mathcal{C}(\overline{\Omega})^d,\\
\frac{(\mathbf{T}_n)^{\mathbf{d}}}{n|(\mathbf{T}_n)^{\mathbf{d}}|^{1-\frac{1}{n}}} &\to \mathbf{0} && \text{strongly in } L_1(\Omega)^{d\times d}_{\rm sym},\\
\frac{{\rm Tr}(\mathbf{T}_n)}{n|{\rm Tr}(\mathbf{T}_n)|^{1-\frac{1}{n}}} &\to 0 && \text{strongly in } L_1(\Omega).
\end{aligned} \tag{4.12}$$

In particular,

$$\varepsilon(\mathbf{u}_n)\rightharpoonup\varepsilon(\mathbf{u}) \qquad \text{weakly in } L_{2d}(\Omega)^{d\times d}_{\rm sym}. \tag{4.13}$$

We note though that the weak convergence result (4.13) can be strengthened to

$$\varepsilon(\mathbf{u}_n)\to\varepsilon(\mathbf{u}) \qquad \text{strongly in } L_p(\Omega_0)^{d\times d}_{\rm sym} \quad \forall\,\Omega_0\subset\subset\Omega, \quad \forall\, p\in[1,\infty) \tag{4.14}$$

and consequently to

$$\varepsilon(\mathbf{u}_n)\to\varepsilon(\mathbf{u}) \qquad \text{strongly in } L_{2d}(\Omega)^{d\times d}_{\rm sym}. \tag{4.15}$$

To show this we fix any $\Omega_0\subset\subset\Omega$ and note that by subtracting the constitutive relation (1.2) from its regularized counterpart (3.1) we have

$$\varepsilon(\mathbf{u}_n-\mathbf{u}) = \mathcal{A}(\mathbf{T}_n)-\mathcal{A}(\mathbf{T}) + \frac{{\rm Tr}(\mathbf{T}_n)\mathbf{I}}{n|{\rm Tr}(\mathbf{T}_n)|^{1-\frac{1}{n}}} + \frac{\mathbf{T}_n^{\mathbf{d}}}{n|\mathbf{T}_n^{\mathbf{d}}|^{1-\frac{1}{n}}},$$

where $\mathcal{A}: L_1(\Omega_0)^{d\times d}_{\rm sym}\to L_\infty(\Omega_0)^{d\times d}_{\rm sym}$ is given by

$$\mathcal{A}(\mathbf{S}) := \lambda({\rm Tr}(\mathbf{S})){\rm Tr}(\mathbf{S})\mathbf{I} + \mu(|\mathbf{S}^{\mathbf{d}}|)\mathbf{S}^{\mathbf{d}}. \tag{4.16}$$

A similar argument to the one in the proof of Lemma 3.1 yields that the mapping $\mathcal{A}: L_1(\Omega_0)^{d\times d}_{\rm sym}\to L_p(\Omega_0)^{d\times d}_{\rm sym}$ is well defined and continuous for all $p\in[1,\infty)$, whence, because $\mathbf{T}_n$ converges strongly

to $\mathbf{T}$ in $L_1(\Omega_0)^{d\times d}_{\text{sym}}$, it follows that so does $\mathcal{A}(\mathbf{T}_n)$ to $\mathcal{A}(\mathbf{T})$ in $L_p(\Omega_0)^{d\times d}_{\text{sym}}$ for all $p \in [1,\infty)$. For the first regularization term, Hölder's inequality implies

$$\frac{1}{n}\||\mathrm{Tr}(\mathbf{T}_n)|^{\frac{1}{n}}\|_{L_p(\Omega_0)} \le \frac{1}{n}\|\mathrm{Tr}(\mathbf{T}_n)\|^{\frac{1}{n}}_{L_1(\Omega_0)}|\Omega_0|^{\frac{1}{p}(1-\frac{p}{n})} \to 0 \qquad \text{as } n\to\infty,$$

and similarly for the second regularization term, containing $\mathbf{T}^{\mathbf{d}}_n$. The convergence result (4.14) then follows by collecting the above results. To show (4.15) we consider a nested sequence of Ω_0 that exhausts Ω. By (4.14) there exists a subsequence (still indexed by n) such that $\varepsilon(\mathbf{u}_n) \to \varepsilon(\mathbf{u})$ almost everywhere on Ω. Hence, in view of (4.13), (4.15) follows by Vitali's theorem.

Motivated by these convergence results our objective is to construct a sequence of finite element approximations $((\mathbf{T}_{n,h},\mathbf{u}_{n,h}))_{h\in(0,1]}$ to the solution pair $(\mathbf{T}_n,\mathbf{u}_n)$ of the regularized problem, for a fixed value of n, and then pass to the limit $h\to 0_+$ with the discretization parameter $h\in(0,1]$, followed by passage to the limit $n\to\infty$ with the regularization parameter $n\in\mathbb{N}$—instead of approximating the solution pair $(\mathbf{T},\mathbf{u})$ directly by a finite element method. Our reasons for proceeding in this way will be made clear at the start of Section 5.2.

5. Finite element approximation

For the sake of simplicity we shall suppose from now on that Ω is a polygon when $d=2$ or a Lipschitz polyhedron when $d=3$.

We consider a sequence of shape-regular simplicial subdivisions $(\mathcal{T}_h)_{h\in(0,1]}$ of $\overline{\Omega}$; by this we mean that there exists a positive real number η, independent of the mesh size h, such that all closed simplices K in the subdivision $\mathcal{T}_h$ satisfy the inequality

$$\frac{h_K}{\varrho_K} \le \eta, \tag{5.1}$$

where h_K is the diameter of K and ϱ_K is the diameter of the largest ball inscribed in K; see for instance Ciarlet (1991). The extension to quadrilateral and hexahedral meshes is discussed in Sections 5.3 and 5.4.

Let $\mathcal{P}^r_h$ be the space of piecewise (subordinate to $\mathcal{T}_h$) polynomials of degree at most r. We consider the conforming finite element spaces

$$\mathbb{M}_{n,h} := \left(\mathcal{P}^0_h\right)^{d\times d}_{\text{sym}} \subset \mathbb{M}_n, \qquad \mathbb{X}_{n,h} := \left(\mathcal{P}^1_h\right)^d \cap \mathbb{X}_n \subset \mathbb{X}_n \tag{5.2}$$

for the approximation of $\mathbf{T}_n$ and $\mathbf{u}_n$, respectively. We note in passing that in the set-theoretical sense $\mathbb{M}_{n,h}$ and $\mathbb{X}_{n,h}$ are independent of n; we shall, however, continue to label them with the double subscript $_{n,h}$ instead of just $_h$ in order to emphasize that they are being thought of as finite-dimensional normed linear subspaces of $\mathbb{M}_n$ and $\mathbb{X}_n$, respectively, throughout the paper.

As the exact solution is not expected to be very smooth, we have restricted ourselves to considering a first-order finite element approximation. There are, of course, other choices of first-order spaces than the one we shall be focusing on, but for the sake of brevity we shall not dwell on those here in detail; for extensions and alternative choices of spaces we refer the reader again to Sections 5.3 and 5.4.

5.1 *Discrete scheme*

The discrete counterpart of problem (3.2), based on $\mathbb{X}_{n,h}$ and $\mathbb{M}_{n,h}$, is then defined as follows: find $(\mathbf{T}_{n,h}, \mathbf{u}_{n,h}) \in \mathbb{M}_{n,h} \times \mathbb{X}_{n,h}$ such that

$$\begin{aligned} &a_n(\mathbf{T}_{n,h}, \mathbf{S}_h) + c(\mathbf{T}_{n,h}; \mathbf{T}_{n,h}, \mathbf{S}_h) - b(\mathbf{S}_h, \mathbf{u}_{n,h}) = 0 && \forall\, \mathbf{S}_h \in \mathbb{M}_{n,h}, \\ &b(\mathbf{T}_{n,h}, \mathbf{v}_h) = \int_\Omega \mathbf{F} : \varepsilon(\mathbf{v}_h)\, \mathrm{d}\mathbf{x} && \forall\, \mathbf{v}_h \in \mathbb{X}_{n,h}. \end{aligned} \tag{5.3}$$

We start by proving the discrete version of the inf-sup property (4.1).

LEMMA 5.1 (Discrete inf-sup property). For each $n \in \mathbb{N}$ we have

$$\inf_{\mathbf{v}_h \in \mathbb{X}_{n,h}} \sup_{\mathbf{S}_h \in \mathbb{M}_{n,h}} \frac{b(\mathbf{S}_h, \mathbf{v}_h)}{\|\mathbf{S}_h\|_{L_{1+\frac{1}{n}}(\Omega)} \|\varepsilon(\mathbf{v}_h)\|_{L_{n+1}(\Omega)}} \geq 1. \tag{5.4}$$

Proof. The argument is based on mimicking the proof of Fortin's lemma. Indeed, the assertion directly follows from the continuous inf-sup property (4.1) upon noting that for all $\mathbf{v}_h \in \mathbb{X}_{n,h} \subset \mathbb{X}_n$, $\varepsilon(\mathbf{v}_h)$ belongs to $\mathbb{M}_{n,h}$. Thus, for all $\mathbf{v}_h \in \mathbb{X}_{n,h}$, all $\mathbf{S} \in \mathbb{M}_n$ and all $K \in \mathcal{T}_h$, one has

$$\int_K \varepsilon(\mathbf{v}_h) : \Pi_h \mathbf{S}\, \mathrm{d}\mathbf{x} = \int_K \varepsilon(\mathbf{v}_h) : \mathbf{S}\, \mathrm{d}\mathbf{x},$$

where $\Pi_h \mathbf{S} \in \mathbb{M}_{n,h}$ is defined componentwise by

$$\Pi_h f|_K := \frac{1}{|K|} \int_K f\, \mathrm{d}\mathbf{x} \qquad \forall\, K \in \mathcal{T}_h,\ \forall f \in L_1(K), \tag{5.5}$$

and the projector $\Pi_h : \mathbb{M}_n \to \mathbb{M}_{n,h}$ is stable in (the norm of) $\mathbb{M}_n$ because

$$\|\Pi_h f\|_{L_{1+\frac{1}{n}}(K)} \leq \|f\|_{L_{1+\frac{1}{n}}(K)} \qquad \forall\, K \in \mathcal{T}_h,\ \forall f \in L_{1+\frac{1}{n}}(K),$$

whereby

$$\|\Pi_h \mathbf{S}\|_{L_{1+\frac{1}{n}}(\Omega)} \leq \|\mathbf{S}\|_{L_{1+\frac{1}{n}}(\Omega)} \qquad \forall\, \mathbf{S} \in L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\mathrm{sym}}.$$

That completes the proof. □

The above proof suggests eliminating the constraint by defining

$$\mathbf{T}^{\mathbf{f}}_{n,h} = \Pi_h \mathbf{T}^{\mathbf{f}}_n = \Pi_h \mathbf{F}; \tag{5.6}$$

we recall that $\mathbf{T}^{\mathbf{f}}_n \in \mathbb{M}_n$ satisfies equality (4.2). Then, by setting

$$\mathbb{V}_{n,h} := \left\{ \mathbf{S}_h \in \mathbb{M}_{n,h} \ :\ b(\mathbf{S}_h, \mathbf{v}_h) = 0 \quad \forall\, \mathbf{v}_h \in \mathbb{X}_{n,h} \right\} \tag{5.7}$$

and

$$\mathbb{V}_{n,h}^{\perp} := \left\{ \mathbf{S}_h \in \mathbb{M}_{n,h} \ : \int_\Omega \mathbf{S}_h : \mathbf{R}_h \, \mathrm{d}\mathbf{x} = 0 \quad \forall \, \mathbf{R}_h \in \mathbb{V}_{n,h} \right\},$$

we deduce from equalities (4.2), (5.6), and by noting that $\varepsilon(\mathbf{v}_h) \in \mathbb{M}_{n,h}$ for all $\mathbf{v}_h \in \mathbb{X}_{n,h}$, that $\mathbf{T}_{n,h}^{\mathbf{f}} \in \mathbb{M}_{n,h}$ satisfies

$$\int_\Omega \varepsilon(\mathbf{v}_h) : \mathbf{T}_{n,h}^{\mathbf{f}} \, \mathrm{d}\mathbf{x} = \int_\Omega \mathbf{F} : \varepsilon(\mathbf{v}_h) \, \mathrm{d}\mathbf{x} \qquad \forall \, \mathbf{v}_h \in \mathbb{X}_{n,h}; \tag{5.8}$$

furthermore, equality (5.6) implies

$$\left\| \mathbf{T}_{n,h}^{\mathbf{f}} \right\|_{L_{1+\frac{1}{n}}(\Omega)} \leq \|\mathbf{F}\|_{L_{1+\frac{1}{n}}(\Omega)}. \tag{5.9}$$

We observe further that, as $h \to 0_+$,

$$\mathbf{T}_{n,h}^{\mathbf{f}} \to \mathbf{T}_n^{\mathbf{f}} \qquad \text{strongly in } \mathbb{M}_n. \tag{5.10}$$

Given $\mathbf{T}_{n,h}^{\mathbf{f}} \in \mathbb{M}_{n,h}$ defined by equality (5.6) we shall seek $\mathbf{T}_{n,h}^{0} := \mathbf{T}_{n,h} - \mathbf{T}_{n,h}^{\mathbf{f}} \in \mathbb{V}_{n,h}$ that solves

$$a_n\big(\mathbf{T}_{n,h}^{0} + \mathbf{T}_{n,h}^{\mathbf{f}}, \mathbf{S}_h\big) + c\big(\mathbf{T}_{n,h}^{0} + \mathbf{T}_{n,h}^{\mathbf{f}}; \mathbf{T}_{n,h}^{0} + \mathbf{T}_{n,h}^{\mathbf{f}}, \mathbf{S}_h\big) = 0 \qquad \forall \, \mathbf{S}_h \in \mathbb{V}_{n,h}. \tag{5.11}$$

The existence of a unique such $\mathbf{T}_{n,h}^{0} \in \mathbb{V}_{n,h}$, and therefore of a unique $\mathbf{T}_{n,h} = \mathbf{T}_{n,h}^{0} + \mathbf{T}_{n,h}^{\mathbf{f}} \in \mathbb{M}_{n,h}$ and a unique $\mathbf{u}_{n,h} \in \mathbb{X}_{n,h}$ satisfying equation $(5.3)_1$ for all $\mathbf{S}_h \in \mathbb{M}_{n,h}$, can be shown by proceeding as in the case of the continuous problem discussed in Section 4, but with the continuous inf-sup property stated in Lemma 4.1, that was used there, now replaced by the discrete inf-sup property stated in Lemma 5.1. Indeed, let $\mathcal{A}_{n,h} : \mathbb{M}_{n,h} \to (\mathbb{M}_{n,h})'$ be defined by the projection of $\mathcal{A}_n$ onto $\mathbb{M}_{n,h}$,

$$\int_\Omega \mathcal{A}_{n,h}(\mathbf{S}_h) : \mathbf{R}_h \, \mathrm{d}\mathbf{x} = a_n(\mathbf{S}_h, \mathbf{R}_h) + c(\mathbf{S}_h; \mathbf{S}_h, \mathbf{R}_h) \qquad \forall \, \mathbf{R}_h \in \mathbb{M}_{n,h}.$$

In the present case where the tensors of $\mathbb{M}_{n,h}$ are piecewise constant functions, $\mathcal{A}_{n,h}(\mathbf{S}_h)$ coincides with $\mathcal{A}_n(\mathbf{S}_h)$, but this equality is not necessary. It is easy to check that $\mathcal{A}_{n,h}$ has the same boundedness, continuity, coercivity and monotonicity properties (all uniform in h) as $\mathcal{A}_n$, as stated in Lemmas 3.1 and 3.3. The same is true of the mapping

$$\mathbf{S}_h \in \mathbb{V}_{n,h} \mapsto \mathcal{A}_{n,h}\big(\mathbf{S}_h + \mathbf{T}_{n,h}^{\mathbf{f}}\big).$$

Therefore, another application of the Browder–Minty theorem gives existence and uniqueness of $\mathbf{T}_{n,h}^{0}$ solving (5.11) and we set $\mathbf{T}_{n,h} = \mathbf{T}_{n,h}^{0} + \mathbf{T}_{n,h}^{\mathbf{f}}$. The discrete inf-sup property (5.4) then guarantees the existence of a corresponding $\mathbf{u}_{n,h} \in \mathbb{X}_{n,h}$ such that $(\mathbf{T}_{n,h}, \mathbf{u}_{n,h})$ solves system (5.3). This is summarized in the following lemma.

LEMMA 5.2 (Existence and uniqueness of the discrete solution). Assume that λ and μ satisfy hypotheses (A1–A4). Then, system (5.3) has exactly one solution $(\mathbf{T}_{n,h}, \mathbf{u}_{n,h}) \in \mathbb{M}_{n,h} \times \mathbb{X}_{n,h}$.

5.2 *Convergence of the sequence of discrete solutions*

Without regularization (i.e., with $\frac{1}{n}$ formally set equal to zero in problem (5.3), resulting in the absence of the form $a_n(\cdot,\cdot)$ from the left-hand side of $(5.3)_1$), the proof of convergence of the sequence of solutions generated by the resulting numerical method to $(\mathbf{T},\mathbf{u})$ is an open problem. The source of the technical difficulties is that, as $n\to\infty$, the only uniform (w.r.t. $n\in\mathbb{N}$) bound on $\mathbf{T}_{n,h}$, with $h\in(0,1]$ fixed, that is directly available to us is in the $L_1(\Omega)^{d\times d}$ norm; a uniform bound in the $L_1(\Omega)^{d\times d}$ norm guarantees only biting weak convergence, via Chacon's biting lemma for example, and this is insufficient to deduce even convergence of a subsequence in the weak topology of $L_1(\Omega)^{d\times d}$. The proof of existence of a solution to the continuous problem in reference Bulíček *et al.* (2014) succeeds because the $L_1(\Omega)^{d\times d}$ norm bound on $\mathbf{T}_n$ in the sequence of solution pairs $(\mathbf{T}_n,\mathbf{u}_n)$ to the regularized problem is supplemented by fractional derivative estimates. Unfortunately, the extension of those fractional derivative estimates to the finite element discretization considered here is problematic. For this reason, we freeze the parameter $n\in\mathbb{N}$ and we now discuss convergence, without rates, of the sequence of solution pairs $(\mathbf{T}_{n,h},\mathbf{u}_{n,h})$ of the discrete scheme to the solution $(\mathbf{T}_n,\mathbf{u}_n)$ of the regularized problem as $h\to 0_+$. Having done so, we will invoke the converge results stated at the end of Section 4 to pass to the limit $n\to\infty$ to deduce that $\lim_{n\to\infty}\lim_{h\to 0_+}(\mathbf{T}_{n,h},\mathbf{u}_{n,h})=(\mathbf{T},\mathbf{u})$ in the strong topology of $L_1(\Omega_0)^{d\times d}_{\rm sym}\times\mathcal{C}(\overline{\Omega})^d$, for any $\Omega_0\subset\subset\Omega$.

We begin by establishing the weak convergence of the sequence $(\mathbf{T}_{n,h})_{h\in(0,1]}\subset\mathbb{M}_n$, with $n\in\mathbb{N}$ fixed.

LEMMA 5.3 (Weak convergence of $\mathbf{T}_{n,h}$). Assume that $\mathbf{F}\in L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\rm sym}$ and that the functions λ and μ satisfy hypotheses (A1–A4). Let $(\mathbf{T}_n,\mathbf{u}_n)\in\mathbb{M}_n\times\mathbb{X}_n$ be the unique solution of the regularized problem (3.2). Then, as $h\to 0_+$,

$$\mathbf{T}_{n,h}\rightharpoonup\mathbf{T}_n\qquad\text{weakly in }\mathbb{M}_n=L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\rm sym}.$$

Proof. In this proof C denotes a generic positive constant that is independent of n and h. We use again the lift $\mathbf{T}^{\mathbf{f}}_{n,h}=\Pi_h\mathbf{T}^{\mathbf{f}}_n$ of the data satisfying equality (5.8) and set $\mathbf{T}^0_{n,h}=\mathbf{T}_{n,h}-\mathbf{T}^{\mathbf{f}}_{n,h}\in\mathbb{V}_{n,h}$, which satisfies equation (5.11). The *a priori* estimates provided by Lemma 4.2 guarantee that

$$\begin{aligned}\|\varepsilon(\mathbf{u}_{n,h})\|_{L_{n+1}(\Omega)}\le\frac{d\sqrt{2}}{n}\left[16d^2\|\mathbf{F}\|^{1+\frac{1}{n}}_{L_{1+\frac{1}{n}}(\Omega)}+2(n+1)C_2\sqrt{2d}\,|\Omega|^{\frac{1}{n+1}}\|\mathbf{F}\|_{L_{1+\frac{1}{n}}(\Omega)}+4(n+1)C_1\kappa|\Omega|\right]^{\frac{1}{n+1}}\\ +C_2\sqrt{2d}\,|\Omega|^{\frac{1}{n+1}}\end{aligned}\tag{5.12}$$

and

$$\begin{aligned}\frac{1}{n+1}\|\mathbf{T}_{n,h}\|^{1+\frac{1}{n}}_{L_{1+\frac{1}{n}}(\Omega)}+C_1\|\mathbf{T}_{n,h}\|_{L_1(\Omega)}\le\frac{16d^2}{n+1}\|\mathbf{F}\|^{1+\frac{1}{n}}_{L_{1+\frac{1}{n}}(\Omega)}\\ +2C_2\sqrt{2d}\,|\Omega|^{\frac{1}{n+1}}\|\mathbf{F}\|_{L_{1+\frac{1}{n}}(\Omega)}+4C_1\kappa|\Omega|.\end{aligned}\tag{5.13}$$

Hence, in particular,

$$\frac{1}{n+1}\|\mathbf{T}_{n,h}\|^{1+\frac{1}{n}}_{L_{1+\frac{1}{n}}(\Omega)}+\|\mathbf{T}_{n,h}\|_{L_1(\Omega)}+\|\varepsilon(\mathbf{u}_{n,h})\|_{L_{n+1}(\Omega)}\le c_n\qquad\forall\,n\in\mathbb{N},\quad\forall\,h\in(0,1],$$

where c_n is a positive constant, depending on $\|\mathbf{F}\|_{L_{1+\frac{1}{n}}(\Omega)}$, C_1, C_2, d, κ and $|\Omega|$ only. Thanks to the stability inequality (5.9) satisfied by the lift $\mathbf{T}^{\mathbf{f}}_{n,h}$ we then deduce that

$$\frac{1}{n+1}\|\mathbf{T}^0_{n,h}\|^{1+\frac{1}{n}}_{L_{1+\frac{1}{n}}(\Omega)} + \|\mathbf{T}^0_{n,h}\|_{L_1(\Omega)} \leq c_n \qquad \forall\, n \in \mathbb{N}, \quad \forall\, h \in (0,1].$$

Therefore, for each fixed $n \in \mathbb{N}$ there exists a subsequence with respect to h (and still indexed by h) and $\overline{\mathbf{T}}^0_n \in L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\mathrm{sym}}$ such that, as $h \to 0_+$,

$$\mathbf{T}^0_{n,h} \rightharpoonup \overline{\mathbf{T}}^0_n \qquad \text{weakly in } L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\mathrm{sym}}. \tag{5.14}$$

We note that $\overline{\mathbf{T}}^0_n \in \mathbb{V}_n$, in fact. Indeed, for any $\mathbf{v}_n \in \mathbb{X}_n$ there exists a sequence $(\mathbf{v}_{n,h})_{h\in(0,1]}$, with $\mathbf{v}_{n,h} \in \mathbb{X}_{n,h}$, such that $\varepsilon(\mathbf{v}_{n,h}) \to \varepsilon(\mathbf{v}_n)$ strongly in $L_{n+1}(\Omega)^{d\times d}_{\mathrm{sym}}$. As

$$b\big(\mathbf{T}^0_{n,h}, \mathbf{v}_{n,h}\big) = b(\mathbf{T}_{n,h}, \mathbf{v}_{n,h}) - b\big(\mathbf{T}^{\mathbf{f}}_{n,h}, \mathbf{v}_{n,h}\big) = (\mathbf{f}, \mathbf{v}_{n,h}) - (\mathbf{f}, \mathbf{v}_{n,h}) = 0,$$

passage to the limit $h \to 0_+$, using the weak convergence (5.14) and the strong convergence $\varepsilon(\mathbf{v}_{n,h}) \to \varepsilon(\mathbf{v}_n)$ in $L_{n+1}(\Omega)^{d\times d}_{\mathrm{sym}}$, implies that $b(\overline{\mathbf{T}}^0_n, \mathbf{v}_n) = 0$ for all $\mathbf{v}_n \in \mathbb{X}_n$. Hence, $\overline{\mathbf{T}}^0_n \in \mathbb{V}_n$ thanks to the definition of the linear space $\mathbb{V}_n$.

We now show, using Minty's method, that $\overline{\mathbf{T}}^0_n$ satisfies equation (4.5). To this end, we recall the notation (3.3) for $\mathcal{A}_n$ and first prove that, for $\mathbf{S}_{n,h} \in \mathbb{V}_{n,h}$,

$$\int_\Omega \mathcal{A}_n\big(\mathbf{S}_{n,h} + \mathbf{T}^{\mathbf{f}}_{n,h}\big) : \big(\mathbf{T}^0_{n,h} - \mathbf{S}_{n,h}\big)\,\mathrm{d}\mathbf{x} \leq 0. \tag{5.15}$$

We begin the proof of inequality (5.15) by invoking the monotonicity result (3.5) to deduce that, for $\mathbf{S}_{n,h} \in \mathbb{V}_{n,h}$,

$$\begin{aligned}\int_\Omega \Big(\mathcal{A}_n\big(\mathbf{T}^0_{n,h} + \mathbf{T}^{\mathbf{f}}_{n,h}\big) - \mathcal{A}_n\big(\mathbf{S}_{n,h} + \mathbf{T}^{\mathbf{f}}_{n,h}\big)\Big) : \big(\mathbf{T}^0_{n,h} - \mathbf{S}_{n,h}\big)\,\mathrm{d}\mathbf{x}\\ = \int_\Omega \Big(\mathcal{A}_n\big(\mathbf{T}^0_{n,h} + \mathbf{T}^{\mathbf{f}}_{n,h}\big) - \mathcal{A}_n\big(\mathbf{S}_{n,h} + \mathbf{T}^{\mathbf{f}}_{n,h}\big)\Big) : \big(\mathbf{T}^0_{n,h} + \mathbf{T}^{\mathbf{f}}_{n,h} - \mathbf{T}^{\mathbf{f}}_{n,h} - \mathbf{S}_{n,h}\big)\,\mathrm{d}\mathbf{x} \geq 0.\end{aligned}$$

Moreover, as $\mathbf{T}^0_{n,h}$ and $\mathbf{S}_{n,h}$ both belong to $\mathbb{V}_{n,h}$, we use relation (5.11) satisfied by $\mathbf{T}^0_{n,h}$ to deduce that

$$\int_\Omega \mathcal{A}_n\big(\mathbf{T}^0_{n,h} + \mathbf{T}^{\mathbf{f}}_{n,h}\big) : \big(\mathbf{T}^0_{n,h} - \mathbf{S}_{n,h}\big)\,\mathrm{d}\mathbf{x} = 0,$$

and thus we obtain inequality (5.15).

We can now use inequality (5.15) to show that $\overline{\mathbf{T}}^0_n$ solves problem (4.5). To see this, we consider $\Pi_h\mathbf{S}_n$ for a given $\mathbf{S}_n \in \mathbb{V}_n$. As $h \to 0_+$, the weak convergence (5.14), the strong convergence $\Pi_h\mathbf{S}_n \to$

$\mathbf{S}_n$ (by density) and $\mathbf{T}^{\mathbf{f}}_{n,h} = \Pi_h \mathbf{T}^{\mathbf{f}}_n \to \mathbf{T}^{\mathbf{f}}_n$ (see (5.10)) in $\mathbb{M}_n$ guarantee that

$$\mathbf{T}^0_{n,h} - \Pi_h \mathbf{S}_n \rightharpoonup \overline{\mathbf{T}}^0_n - \mathbf{S}_n \qquad \text{weakly in } \mathbb{M}_n$$

and

$$\Pi_h \mathbf{S}_n + \mathbf{T}^{\mathbf{f}}_{n,h} \to \mathbf{S}_n + \mathbf{T}^{\mathbf{f}}_n \qquad \text{strongly in } \mathbb{M}_n.$$

Hence, inequality (5.15) and the continuity of $\mathcal{A}_n$ (cf. Lemma 3.1(ii)) lead to

$$\int_\Omega \mathcal{A}_n\big(\mathbf{S}_n + \mathbf{T}^{\mathbf{f}}_n\big) : \big(\overline{\mathbf{T}}^0_n - \mathbf{S}_n\big)\, \mathrm{d}\mathbf{x} \le 0 \qquad \forall\, \mathbf{S}_n \in \mathbb{V}_n.$$

Choosing $\mathbf{S}_n = \overline{\mathbf{T}}^0_n - t\mathbf{W}_n$ for $t > 0$ and some $\mathbf{W}_n \in \mathbb{V}_n$ we get

$$\int_\Omega \mathcal{A}_n\big(\overline{\mathbf{T}}^0_n + \mathbf{T}^{\mathbf{f}}_n - t\mathbf{W}_n\big) : \mathbf{W}_n\, \mathrm{d}\mathbf{x} \le 0 \qquad \forall\, \mathbf{W}_n \in \mathbb{V}_n.$$

Thanks to the continuity (and therefore hemicontinuity) of $\mathcal{A}_n$ (cf. again Lemma 3.1(ii)), we can pass to the limit $t \to 0_+$ to deduce that

$$\int_\Omega \mathcal{A}_n\big(\overline{\mathbf{T}}^0_n + \mathbf{T}^{\mathbf{f}}_n\big) : \mathbf{W}_n\, \mathrm{d}\mathbf{x} \le 0 \qquad \forall\, \mathbf{W}_n \in \mathbb{V}_n,$$

and consequently, since $\mathbb{V}_n$ is a linear space, after replacing $\mathbf{W}_n$ by $-\mathbf{W}_n$ in the inequality above and then combining the two inequalities,

$$\int_\Omega \mathcal{A}_n\big(\overline{\mathbf{T}}^0_n + \mathbf{T}^{\mathbf{f}}_n\big) : \mathbf{W}_n\, \mathrm{d}\mathbf{x} = 0 \qquad \forall\, \mathbf{W}_n \in \mathbb{V}_n,$$

which shows that $\overline{\mathbf{T}}^0_n = \mathbf{T}^0_n = \mathbf{T}_n - \mathbf{T}^{\mathbf{f}}_n$ satisfies equation (4.5), and thus $\mathbf{T}_{n,h} \rightharpoonup \mathbf{T}_n$ in $\mathbb{M}_n$ as $h \to 0_+$. □

LEMMA 5.4 (Strong convergence). Assume that $\mathbf{F} \in L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\mathrm{sym}}$, that the functions λ and μ satisfy assumptions (A1–A4) and let $(\mathbf{T}_n, \mathbf{u}_n)$ denote the unique solution to the regularized problem (3.2), with $n \in \mathbb{N}$. Then, for each fixed $n \in \mathbb{N}$, as $h \to 0_+$,

$$\mathbf{T}_{n,h} \to \mathbf{T}_n \qquad \text{strongly in } L_p(\Omega)^{d\times d}_{\mathrm{sym}} \quad \forall\, p \in \big[1, 1+\tfrac{1}{n}\big), \tag{5.16}$$

$$\varepsilon(\mathbf{u}_{n,h}) \rightharpoonup \varepsilon(\mathbf{u}_n) \qquad \text{weakly in } L_p(\Omega)^{d\times d}_{\mathrm{sym}} \quad \forall\, p \in [1, n+1]. \tag{5.17}$$

When $n = 1$, the strong convergence result (5.16) holds for all $p \in [1, 2]$. In addition, for each $n \in \mathbb{N}$,

$$\mathbf{u}_{n,h} \to \mathbf{u}_n \qquad \begin{cases} \text{strongly in } L_p(\Omega)^d & \forall\, p \in \big[1, \frac{d(n+1)}{d-(n+1)}\big) \text{ when } 1 \le n \le d-1, \\ \text{strongly in } C^{0,\alpha}(\overline{\Omega})^d & \forall\, \alpha \in \big(0, 1 - \frac{d}{n+1}\big) \text{ when } d < n+1, \end{cases} \tag{5.18}$$

and for each $n \in \mathbb{N}$, $n \geq 2$,

$$\varepsilon(\mathbf{u}_{n,h}) \to \varepsilon(\mathbf{u}_n) \qquad \text{strongly in } L_n(\Omega)^{d\times d}_{\text{sym}}. \tag{5.19}$$

Furthermore, if λ satisfies (A3′), we have that, for any $\Omega_0 \subset\subset \Omega$,

$$\lim_{n\to\infty}\lim_{h\to 0_+} \|\mathbf{T}_{n,h} - \mathbf{T}\|_{L_1(\Omega_0)} = 0 \qquad \text{and} \qquad \lim_{n\to\infty}\lim_{h\to 0_+} \|\mathbf{u}_{n,h} - \mathbf{u}\|_{C(\overline{\Omega})} = 0$$

and

$$\lim_{n\to\infty}\lim_{h\to 0_+} \|\varepsilon(\mathbf{u}_{n,h}) - \varepsilon(\mathbf{u})\|_{L_p(\Omega_0)} = 0 \qquad \forall\, \Omega_0 \subset\subset \Omega, \ \forall p \in [1,\infty),$$

where $(\mathbf{T}, \mathbf{u})$ denotes the unique solution of the original (nonregularized) continuous problem (1.1), (1.2) subject to a homogeneous Dirichlet boundary condition on $\partial\Omega$.

Proof. In this proof, again, C denotes a generic positive constant, independent of h and n. Also, we use again the notation $\mathbf{T}_n = \mathbf{T}^0_n + \mathbf{T}^{\mathbf{f}}_n$ and $\mathbf{T}_{n,h} = \mathbf{T}^0_{n,h} + \mathbf{T}^{\mathbf{f}}_{n,h}$, where $\mathbf{T}^{\mathbf{f}}_n = \mathbf{F}$ satisfies equality (4.2) and $\mathbf{T}^{\mathbf{f}}_{n,h} = \Pi_h \mathbf{T}^{\mathbf{f}}_n$ satisfies equality (5.8).

To establish control on $\mathbf{T}^0_{n,h} - \mathbf{T}^0_n$ we write

$$\mathbf{T}^0_{n,h} - \mathbf{T}^0_n = \left(\mathbf{T}^0_{n,h} - \Pi_h \mathbf{T}^0_n\right) + \left(\Pi_h \mathbf{T}^0_n - \mathbf{T}^0_n\right). \tag{5.20}$$

Since $\Pi_h \mathbf{T}^0_n \to \mathbf{T}^0_n$ strongly in $L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\text{sym}}$ for all $n \in \mathbb{N}$, it suffices to focus on the discrepancy $\mathbf{T}^0_{n,h} - \Pi_h \mathbf{T}^0_n$.

Thanks to inequality (3.6), for any pair of matrices $\mathbf{S}, \mathbf{R} \in \mathbb{R}^{d\times d}_{\text{sym}}$, one has

$$\begin{aligned}\left(\mathbf{S}|\mathbf{S}|^{\frac{1}{n}-1} - \mathbf{R}|\mathbf{R}|^{\frac{1}{n}-1}\right) : (\mathbf{S}-\mathbf{R}) &\geq \frac{1}{n}|\mathbf{S}-\mathbf{R}|^2 \int_0^1 |\mathbf{R} + \theta(\mathbf{S}-\mathbf{R})|^{\frac{1}{n}-1}\, \mathrm{d}\theta \\ &\geq \frac{1}{n}\frac{|\mathbf{S}-\mathbf{R}|^2}{(|\mathbf{R}| + |\mathbf{S}-\mathbf{R}|)^{1-\frac{1}{n}}}.\end{aligned}$$

Analogously, for any pair of real numbers $s, r \in \mathbb{R}$,

$$\left(s|s|^{\frac{1}{n}-1} - r|r|^{\frac{1}{n}-1}\right)(s-r) \geq \frac{1}{n}\frac{|s-r|^2}{(|r| + |s-r|)^{1-\frac{1}{n}}}.$$

Hence, and by invoking inequalities (1.5) and (1.7) (guaranteed by assumptions (A1–A4)) we have

$$\begin{aligned}\int_\Omega \Big(\mathcal{A}_n\big(\mathbf{T}^0_{n,h}+\mathbf{T}^{\mathbf{f}}_{n,h}\big)-\mathcal{A}_n\big(\Pi_h\mathbf{T}^0_n+\mathbf{T}^{\mathbf{f}}_{n,h}\big)\Big):\big(\mathbf{T}^0_{n,h}-\Pi_h\mathbf{T}^0_n\big)\,\mathrm{d}\mathbf{x}\\ \geq \frac{1}{n^2}\int_\Omega \frac{\big|\mathrm{Tr}\big(\mathbf{T}^0_{n,h}+\mathbf{T}^{\mathbf{f}}_{n,h}\big)-\mathrm{Tr}\big(\Pi_h\mathbf{T}^0_n+\mathbf{T}^{\mathbf{f}}_{n,h}\big)\big|^2}{\big(\big|\mathrm{Tr}\big(\Pi_h\mathbf{T}^0_n+\mathbf{T}^{\mathbf{f}}_{n,h}\big)\big|+\big|\mathrm{Tr}\big(\mathbf{T}^0_{n,h}+\mathbf{T}^{\mathbf{f}}_{n,h}\big)-\mathrm{Tr}\big(\Pi_h\mathbf{T}^0_n+\mathbf{T}^{\mathbf{f}}_{n,h}\big)\big|\big)^{1-\frac{1}{n}}}\,\mathrm{d}\mathbf{x}\\ +\frac{1}{n^2}\int_\Omega \frac{\big|\big(\mathbf{T}^0_{n,h}+\mathbf{T}^{\mathbf{f}}_{n,h}\big)^{\mathbf{d}}-\big(\Pi_h\mathbf{T}^0_n+\mathbf{T}^{\mathbf{f}}_{n,h}\big)^{\mathbf{d}}\big|^2}{\big(\big|\big(\Pi_h\mathbf{T}^0_n+\mathbf{T}^{\mathbf{f}}_{n,h}\big)^{\mathbf{d}}\big|+\big|\big(\mathbf{T}^0_{n,h}+\mathbf{T}^{\mathbf{f}}_{n,h}\big)^{\mathbf{d}}-\big(\Pi_h\mathbf{T}^0_n+\mathbf{T}^{\mathbf{f}}_{n,h}\big)^{\mathbf{d}}\big|\big)^{1-\frac{1}{n}}}\,\mathrm{d}\mathbf{x}.\end{aligned}\tag{5.21}$$

Moreover, because $\mathbf{T}^0_n=\overline{\mathbf{T}}^0_n\in\mathbb{V}_n$ (cf. the last sentence in the proof of Lemma 5.3) we have

$$\int_\Omega \varepsilon(\mathbf{v}_h):\Pi_h\mathbf{T}^0_n\,\mathrm{d}\mathbf{x}=\int_\Omega \varepsilon(\mathbf{v}_h):\mathbf{T}^0_n\,\mathrm{d}\mathbf{x}=0\qquad\forall\,\mathbf{v}_h\in\mathbb{X}_{n,h},$$

and so $\Pi_h\mathbf{T}^0_n\in\mathbb{V}_{n,h}$. As a consequence, $\mathbf{T}^0_{n,h}-\Pi_h\mathbf{T}^0_n\in\mathbb{V}_{n,h}$ and there holds

$$\int_\Omega \mathcal{A}_n\big(\mathbf{T}^0_{n,h}+\mathbf{T}^{\mathbf{f}}_{n,h}\big):\big(\mathbf{T}^0_{n,h}-\Pi_h\mathbf{T}^0_n\big)\,\mathrm{d}\mathbf{x}=0.$$

Using this in inequality (5.21) and noting that $\Pi_h\mathbf{T}^0_n+\mathbf{T}^{\mathbf{f}}_{n,h}=\Pi_h(\mathbf{T}^0_n+\mathbf{T}^{\mathbf{f}}_n)=\Pi_h\mathbf{T}_n$ we obtain

$$\begin{aligned}-\int_\Omega \mathcal{A}_n\big(\Pi_h\mathbf{T}^0_n+\mathbf{T}^{\mathbf{f}}_{n,h}\big):\big(\mathbf{T}^0_{n,h}-\Pi_h\mathbf{T}^0_n\big)\,\mathrm{d}\mathbf{x}\\ \geq \frac{1}{n^2}\int_\Omega \frac{\big|\mathrm{Tr}\big(\mathbf{T}^0_{n,h}-\Pi_h\mathbf{T}^0_n\big)\big|^2}{\big(|\mathrm{Tr}(\Pi_h\mathbf{T}_n)|+\big|\mathrm{Tr}\big(\mathbf{T}^0_{n,h}-\Pi_h\mathbf{T}^0_n\big)\big|\big)^{1-\frac{1}{n}}}\,\mathrm{d}\mathbf{x}\\ +\frac{1}{n^2}\int_\Omega \frac{\big|\big(\mathbf{T}^0_{n,h}-\Pi_h\mathbf{T}^0_n\big)^{\mathbf{d}}\big|^2}{\big(|(\Pi_h\mathbf{T}_n)^{\mathbf{d}}|+\big|\big(\mathbf{T}^0_{n,h}-\Pi_h\mathbf{T}^0_n\big)^{\mathbf{d}}\big|\big)^{1-\frac{1}{n}}}\,\mathrm{d}\mathbf{x}\geq 0.\end{aligned}\tag{5.22}$$

On the one hand, $\mathbf{T}^0_{n,h}-\Pi_h\mathbf{T}^0_n$ weakly converges to $\mathbf{T}^0_n-\mathbf{T}^0_n=\mathbf{0}$ in $\mathbb{M}_n$ as $h\to 0_+$ (cf. Lemma 5.3). On the other hand, $\Pi_h\mathbf{T}^0_n+\mathbf{T}^{\mathbf{f}}_{n,h}$ strongly converges to $\mathbf{T}^0_n+\mathbf{T}^{\mathbf{f}}_n$ in $\mathbb{M}_n=L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\mathrm{sym}}$ as $h\to 0_+$. Therefore, the continuity of the mapping $\mathcal{A}_n:L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\mathrm{sym}}\to L_{n+1}(\Omega)^{d\times d}_{\mathrm{sym}}$ (cf. Lemma 3.1(ii)), which implies that $\mathcal{A}_n(\Pi_h\mathbf{T}^0_n+\mathbf{T}^{\mathbf{f}}_{n,h})$ strongly converges to $\mathcal{A}_n(\mathbf{T}^0_n+\mathbf{T}^{\mathbf{f}}_n)$ in $L_{n+1}(\Omega)^{d\times d}_{\mathrm{sym}}$ as $h\to 0_+$, yields

$$-\int_\Omega \mathcal{A}_n\big(\Pi_h\mathbf{T}^0_n+\mathbf{T}^{\mathbf{f}}_{n,h}\big):\big(\mathbf{T}^0_{n,h}-\Pi_h\mathbf{T}^0\big)\,\mathrm{d}\mathbf{x}\to 0\qquad\text{as } h\to 0_+,$$

whence, returning to inequality (5.22),

$$0 \leq \frac{1}{n^2}\int_\Omega \frac{\left|\mathrm{Tr}\left(\mathbf{T}^0_{n,h} - \Pi_h\mathbf{T}^0_n\right)\right|^2}{\left(|\mathrm{Tr}(\Pi_h\mathbf{T}_n)| + \left|\mathrm{Tr}\left(\mathbf{T}^0_{n,h} - \Pi_h\mathbf{T}^0_n\right)\right|\right)^{1-\frac{1}{n}}}\,\mathrm{d}\mathbf{x}$$
$$+ \frac{1}{n^2}\int_\Omega \frac{\left|\left(\mathbf{T}^0_{n,h} - \Pi_h\mathbf{T}^0_n\right)^{\mathbf{d}}\right|^2}{\left(|(\Pi_h\mathbf{T}_n)^{\mathbf{d}}| + \left|\left(\mathbf{T}^0_{n,h} - \Pi_h\mathbf{T}^0_n\right)^{\mathbf{d}}\right|\right)^{1-\frac{1}{n}}}\,\mathrm{d}\mathbf{x} \to 0 \qquad \text{as } h \to 0_+.$$

Consequently, for each $n \in \mathbb{N}$,

$$\lim_{h\to 0_+}\int_\Omega \frac{\left|\mathrm{Tr}\left(\mathbf{T}^0_{n,h} - \Pi_h\mathbf{T}^0_n\right)\right|^2}{\left(|\mathrm{Tr}(\Pi_h\mathbf{T}_n)| + \left|\mathrm{Tr}\left(\mathbf{T}^0_{n,h} - \Pi_h\mathbf{T}^0_n\right)\right|\right)^{1-\frac{1}{n}}}\,\mathrm{d}\mathbf{x} = 0,$$
$$\lim_{h\to 0_+}\int_\Omega \frac{\left|\left(\mathbf{T}^0_{n,h} - \Pi_h\mathbf{T}^0_n\right)^{\mathbf{d}}\right|^2}{\left(|(\Pi_h\mathbf{T}_n)^{\mathbf{d}}| + \left|\left(\mathbf{T}^0_{n,h} - \Pi_h\mathbf{T}^0_n\right)^{\mathbf{d}}\right|\right)^{1-\frac{1}{n}}}\,\mathrm{d}\mathbf{x} = 0.$$

In the special case when $n = 1$ we directly deduce from these, equality (1.12) and the strong convergence of $\Pi_h\mathbf{T}^0_n$ to $\mathbf{T}^0_n$ in $L_2(\Omega)^{d\times d}_{\mathrm{sym}}$, that $\mathbf{T}^0_{n,h} \to \mathbf{T}^0_n$ strongly in $L_2(\Omega)^{d\times d}_{\mathrm{sym}}$, as $h \to 0_+$. Since $\mathbf{T}^{\mathbf{f}}_{n,h} = \Pi_h\mathbf{T}^{\mathbf{f}}_n \to \mathbf{T}^{\mathbf{f}}_n$ strongly in $L_2(\Omega)^{d\times d}_{\mathrm{sym}}$ as $h \to 0_+$, it follows that, for $n = 1$, $\mathbf{T}_{n,h} \to \mathbf{T}_n$ strongly in $L_2(\Omega)^{d\times d}_{\mathrm{sym}}$, and therefore also strongly in $L_p(\Omega)^{d\times d}_{\mathrm{sym}}$ for all $p \in [1,2]$, as $h \to 0_+$. That completes the proof of the assertion of the lemma concerning $(\mathbf{T}_{n,h})_{h\in(0,1]}$ for $n = 1$.

Let us now consider the case when $n > 1$. Let $\mathcal{M}(f)$ denote the Hardy–Littlewood maximal function of $f \in L_1(\Omega)$, with f extended by zero outside Ω to the whole of $\mathbb{R}^d$, and let $B_r(\mathbf{x})$ denote the d-dimensional ball of radius r centered at $\mathbf{x} \in \mathbb{R}^d$. Clearly,

$$|\Pi_h\mathbf{T}_n(\mathbf{x})| \leq \frac{1}{|K|}\int_K |\mathbf{T}_n(\mathbf{y})|\,\mathrm{d}\mathbf{y} \leq \frac{|B_{h_K}(\mathbf{x})|}{|K|}\left(\frac{1}{|B_{h_K}(\mathbf{x})|}\int_{B_{h_K}(\mathbf{x})} |\mathbf{T}_n(\mathbf{y})|\,\mathrm{d}\mathbf{y}\right) \leq c(\eta)\,\mathcal{M}(|\mathbf{T}_n|)(\mathbf{x})$$

for all $\mathbf{x} \in K$ and all $K \in \mathcal{T}_h$, where $h_K = \mathrm{diam}(K)$ and $c(\eta)$ is a positive constant that depends only on the shape-regularity parameter η of the family $(\mathcal{T}_h)_{h\in(0,1]}$ of simplicial subdivisions of the domain $\overline{\Omega}$ (see (5.1)). Thus,

$$|\Pi_h\mathbf{T}_n(\mathbf{x})| \leq c(\eta)\,\mathcal{M}(|\mathbf{T}_n|)(\mathbf{x}) \qquad \forall\,\mathbf{x} \in \Omega. \tag{5.23}$$

Since the Hardy–Littlewood maximal function is of weak-type $(L_1, L_{1,\infty})$ (with $L_{1,\infty}$ signifying a Lorentz space) with norm at most 3^d (cf. Theorem 2.1.6 and inequality (2.1.3) in Grafakos (2014)), we have

$$|\{\mathbf{x} \in \Omega \,:\, \mathcal{M}(|\mathbf{T}_n|)(\mathbf{x}) > t\}| \leq \frac{3^d}{t}\,\|\mathbf{T}_n\|_{L_1(\Omega)} \qquad \forall\, t > 0.$$

For $k \in \mathbb{N}$ we define

$$\Omega_k := \{\mathbf{x} \in \Omega \,:\, \mathcal{M}(|\mathbf{T}_n|)(\mathbf{x}) \leq k\}.$$

Hence,

$$\Omega_1 \subset \Omega_2 \subset \cdots \subset \Omega \quad \text{and} \quad |\Omega \setminus \Omega_k| \leq \frac{3^d}{k} \|\mathbf{T}_n\|_{L_1(\Omega)} \qquad \forall\, k \in \mathbb{N}; \tag{5.24}$$

in particular,

$$\lim_{k\to\infty} |\Omega \setminus \Omega_k| = 0. \tag{5.25}$$

By recalling (1.12), (5.23) and the definition of the set Ω_k, we have

$$|\mathrm{Tr}(\Pi_h \mathbf{T}_n)(\mathbf{x})| \leq d^{\frac{1}{2}} |\Pi_h \mathbf{T}_n(\mathbf{x})| \leq d^{\frac{1}{2}}\, c(\eta)\, \mathcal{M}(|\mathbf{T}_n|)(\mathbf{x}) \leq d^{\frac{1}{2}}\, c(\eta)\, k \qquad \forall\, \mathbf{x} \in \Omega_k,\ \ \forall\, k \in \mathbb{N},$$
$$|(\Pi_h \mathbf{T}_n)^{\mathbf{d}}(\mathbf{x})| \leq |\Pi_h \mathbf{T}_n(\mathbf{x})| \leq c(\eta)\, \mathcal{M}(|\mathbf{T}_n|)(\mathbf{x}) \leq c(\eta)\, k \qquad \forall\, \mathbf{x} \in \Omega_k,\ \ \forall\, k \in \mathbb{N}.$$

Thus, we deduce that, for each $k \in \mathbb{N}$,

$$\lim_{h\to 0_+} \int_{\Omega_k} \frac{\left|\mathrm{Tr}\left(\mathbf{T}^0_{n,h} - \Pi_h \mathbf{T}^0_n\right)\right|^2}{\left(\sqrt{d} c(\eta)\, k + \left|\mathrm{Tr}\left(\mathbf{T}^0_{n,h} - \Pi_h \mathbf{T}^0_n\right)\right|\right)^{1-\frac{1}{n}}}\, \mathrm{d}\mathbf{x} = 0,$$
$$\lim_{h\to 0_+} \int_{\Omega_k} \frac{\left|\left(\mathbf{T}^0_{n,h} - \Pi_h \mathbf{T}^0_n\right)^{\mathbf{d}}\right|^2}{\left(c(\eta)\, k + \left|\left(\mathbf{T}^0_{n,h} - \Pi_h \mathbf{T}^0_n\right)^{\mathbf{d}}\right|\right)^{1-\frac{1}{n}}}\, \mathrm{d}\mathbf{x} = 0,$$

as $h \to 0_+$; hence, for each k there exists a null sequence $(h^{(k)}) \subset (0, 1]$, with $(h^{(k+1)}) \subset (h^{(k)})$ for all $k \in \mathbb{N}$, such that

$$\frac{\left|\mathrm{Tr}\left(\mathbf{T}^0_{n,h^{(k)}} - \Pi_{h^{(k)}} \mathbf{T}^0_n\right)\right|^2}{\left(\sqrt{d} c(\eta)\, k + \left|\mathrm{Tr}\left(\mathbf{T}^0_{n,h^{(k)}} - \Pi_{h^{(k)}} \mathbf{T}^0_n\right)\right|\right)^{1-\frac{1}{n}}} \to 0 \quad \text{and} \quad \frac{\left|\left(\mathbf{T}^0_{n,h^{(k)}} - \Pi_{h^{(k)}} \mathbf{T}^0_n\right)^{\mathbf{d}}\right|^2}{\left(c(\eta)\, k + \left|\left(\mathbf{T}^0_{n,h^{(k)}} - \Pi_{h^{(k)}} \mathbf{T}^0_n\right)^{\mathbf{d}}\right|\right)^{1-\frac{1}{n}}} \to 0$$

a.e. on Ω_k as $h^{(k)} \to 0_+$. Since

$$\frac{a^2}{(k+a)^{1-\frac{1}{n}}} \geq 2^{\frac{1}{n}-1} \min\left(\frac{a^2}{k^{1-\frac{1}{n}}}, a^{1+\frac{1}{n}}\right) \quad \forall\, a \geq 0, \quad \forall\, k \in \mathbb{N}, \tag{5.26}$$

it follows that

$$\left|\mathrm{Tr}\left(\mathbf{T}^0_{n,h^{(k)}} - \Pi_{h^{(k)}} \mathbf{T}^0_n\right)\right| \to 0 \quad \text{and} \quad \left|\left(\mathbf{T}^0_{n,h^{(k)}} - \Pi_{h^{(k)}} \mathbf{T}^0_n\right)^{\mathbf{d}}\right| \to 0 \qquad \text{a.e. on } \Omega_k,\ \forall\, k\ \in \mathbb{N}$$

as $h^{(k)} \to 0_+$. We then deduce from inequality (1.13) that

$$\mathbf{T}^0_{n,h^{(k)}} - \Pi_{h^{(k)}} \mathbf{T}^0_n \to 0 \qquad \text{a.e. on } \Omega_k\ \forall\, k \in \mathbb{N} \text{ as } h^{(k)} \to 0_+.$$

Hence,

$$\mathbf{T}^0_{n,h^{(k)}} \to \mathbf{T}^0_n \qquad \text{a.e. on } \Omega_k \ \forall\, k \in \mathbb{N} \text{ as } h^{(k)} \to 0_+.$$

By Cantor's diagonal argument we can then extract a 'diagonal' null sequence $(h^{(\infty)})$ such that

$$\mathbf{T}^0_{n,h^{(\infty)}} \to \mathbf{T}^0_n \qquad \text{a.e. on } \Omega_k \ \forall\, k \in \mathbb{N} \text{ as } h^{(\infty)} \to 0_+.$$

Since the sets Ω_k are nested (cf. (5.24)) and they exhaust the whole of Ω (cf. (5.25)), it follows that

$$\mathbf{T}^0_{n,h^{(\infty)}} \to \mathbf{T}^0_n \qquad \text{a.e. on } \Omega \text{ as } h^{(\infty)} \to 0_+.$$

For the sake of the simplicity of our notation we shall henceforth suppress the superscript $^{(\infty)}$ and will simply write

$$\mathbf{T}^0_{n,h} \to \mathbf{T}^0_n \qquad \text{a.e. on } \Omega \text{ as } h \to 0_+.$$

As $\mathbf{T}^{\mathbf{f}}_{n.h} = \Pi_h \mathbf{T}^{\mathbf{f}}_n \to \mathbf{T}^{\mathbf{f}}_n$ strongly in $L_{1+\frac{1}{n}}(\Omega)^{n\times n}_{\text{sym}}$, and therefore (for a subsequence, not indicated) a.e. in Ω, it follows that

$$\mathbf{T}_{n,h} = \mathbf{T}^0_{n,h} + \mathbf{T}^{\mathbf{f}}_{n.h} \to \mathbf{T}^0_n + \mathbf{T}^{\mathbf{f}}_n = \mathbf{T}_n \qquad \text{a.e. on } \Omega \text{ as } h \to 0_+.$$

As, by Lemma 5.3, $\mathbf{T}_{n,h} \rightharpoonup \mathbf{T}_n$ weakly in $L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\text{sym}}$, it follows that the sequence $(\mathbf{T}_{n,h})_{h\in(0,1]}$ is equi-integrable in $L_1(\Omega)^{d\times d}_{\text{sym}}$, and therefore by Vitali's theorem,

$$\mathbf{T}_{n,h} \to \mathbf{T}_n \qquad \text{strongly in } L_1(\Omega)^{d\times d}_{\text{sym}} \text{ as } h \to 0_+, \tag{5.27}$$

whereby, because of the weak convergence $\mathbf{T}_{n,h} \rightharpoonup \mathbf{T}_n$ in $L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\text{sym}}$, it follows that

$$\mathbf{T}_{n,h} \to \mathbf{T}_n \qquad \text{strongly in } L_p(\Omega)^{d\times d}_{\text{sym}} \quad \forall\, p \in \left[1, 1+\frac{1}{n}\right) \text{ as } \ h \to 0_+,$$

where the limiting function $\mathbf{T}_n$ is the first component of the *unique* solution $(\mathbf{T}_n, \mathbf{u}_n)$ of the regularized problem. That completes the proof of the strong convergence result (5.16) for $n > 1$. For $n = 1$, (5.16) has already been shown in the first part of this proof for all $p \in [1, 2]$; hence (5.16) holds for all $n \in \mathbb{N}$.

To prove the strong convergence of the sequence $(\mathbf{u}_{n,h})_{h\in(0,1]} \subset \mathbb{X}_{n,h}$ to $\mathbf{u}_n \in \mathbb{X}_n$ we note that inequality (5.12) implies that the sequence $(\varepsilon(\mathbf{u}_{n,h}))_{h\in(0,1]}$ is bounded in $L_{n+1}(\Omega)^{d\times d}_{\text{sym}}$. Hence, there exists a subsequence (not indicated) and $\overline{\varepsilon(\mathbf{u}_{n,h})} \in L_{n+1}(\Omega)^{d\times d}_{\text{sym}}$ such that, as $h \to 0_+$,

$$\varepsilon(\mathbf{u}_{n,h}) \rightharpoonup \overline{\varepsilon(\mathbf{u}_{n,h})} \qquad \text{weakly in } L_{n+1}(\Omega)^{d\times d}_{\text{sym}}. \tag{5.28}$$

Here, and henceforth, for any weakly (respectively, strongly) convergent sequence of the form $(a_{n,h})_{h\in(0,1]}$ in a function space, with $n \geq 1$ held fixed, $\overline{a_{n,h}}$ will denote the weak (respectively, strong) limit of the sequence as $h \to 0_+$, in instances where the limit of the sequence is yet to be identified.

This will imply assertion (5.17) once we have shown that $\overline{\varepsilon(\mathbf{u}_{n,h})} = \varepsilon(\mathbf{u}_n)$, which we shall do now. For $1 \leq n \leq d-1$, Korn's inequality (1.10) and Poincaré's inequality (1.9) together imply that $(\mathbf{u}_{n,h})_{h\in(0,1]}$ is bounded in $W_0^{1,n+1}(\Omega)^d$, and by Kondrashov's compact embedding theorem the sequence therefore possesses a strongly convergent subsequence (not indicated), with limit $\overline{\mathbf{u}_{n,h}} \in L_p(\Omega)^d$, such that

$$\mathbf{u}_{n,h} \to \overline{\mathbf{u}_{n,h}} \qquad \text{strongly in } L_p(\Omega)^d \text{ as } h \to 0_+ \quad \forall\, p \in \left[1, \frac{(n+1)d}{d-(n+1)}\right).$$

This will imply the first line of assertion (5.18) once we have shown that $\overline{\mathbf{u}_{n,h}} = \mathbf{u}_n$, which we shall do below. In any case, by the uniqueness of the weak limit it then follows that $\overline{\varepsilon(\mathbf{u}_{n,h})} = \varepsilon(\overline{\mathbf{u}_{n,h}})$, and therefore

$$\varepsilon(\mathbf{u}_{n,h}) \rightharpoonup \varepsilon(\overline{\mathbf{u}_{n,h}}) \qquad \text{weakly in } L_{n+1}(\Omega)^{d\times d}_{\text{sym}} \text{ as } h \to 0_+.$$

For $n > d-1$, by an analogous argument,

$$\mathbf{u}_{n,h} \to \overline{\mathbf{u}_{n,h}} \qquad \text{strongly in } C^{0,\alpha}(\overline{\Omega})^d \text{ as } h \to 0_+ \ \forall\, \alpha \in \left(0, 1 - \frac{d}{n+1}\right).$$

This will imply the second line of assertion (5.18) provided we show that $\overline{\mathbf{u}_{n,h}} = \mathbf{u}_n$, the second component of the unique solution $(\mathbf{T}_n, \mathbf{u}_n)$ of the regularized problem. We shall do so by passing to the limit in equation $(5.3)_1$.

To this end, take any $\mathbf{S} \in L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\text{sym}}$ and let $\mathbf{S}_h := \Pi_h \mathbf{S}$ in equation $(5.3)_1$, resulting in

$$a_n(\mathbf{T}_{n,h}, \Pi_h \mathbf{S}) + c(\mathbf{T}_{n,h}; \mathbf{T}_{n,h}, \Pi_h \mathbf{S}) - b(\Pi_h \mathbf{S}, \mathbf{u}_{n,h}) = 0. \tag{5.29}$$

As

$$\Pi_h \mathbf{S} \to \mathbf{S} \qquad \text{strongly in } L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\text{sym}} \tag{5.30}$$

it follows from the weak convergence (5.28) that, for each $n \in \mathbb{N}$,

$$\lim_{h\to 0_+} b(\Pi_h \mathbf{S}, \mathbf{u}_{n,h}) = b(\mathbf{S}, \overline{\mathbf{u}_{n,h}}) \qquad \forall\, \mathbf{S} \in L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\text{sym}}. \tag{5.31}$$

We shall now pass to the limit $h \to 0_+$ in the first two terms on the left-hand side of equation (5.29).

Thanks to the strong convergence result $\mathbf{T}_{n,h} \to \mathbf{T}_n$ in $L_1(\Omega)^{d\times d}_{\text{sym}}$, which follows from assertion (5.16) for all $n \in \mathbb{N}$, as $h \to 0_+$, an identical argument to the one in the proof of Lemma 3.1(ii) implies that, as $h \to 0_+$,

$$\lambda(\text{Tr}(\mathbf{T}_{n,h}))\text{Tr}(\mathbf{T}_{n,h})\mathbf{I} \to \lambda(\text{Tr}(\mathbf{T}_n))\text{Tr}(\mathbf{T}_n)\mathbf{I} \quad \text{strongly in } L_{n+1}(\Omega)^{d\times d}_{\text{sym}}$$

and

$$\mu\big(|\mathbf{T}^{\mathbf{d}}_{n,h}|\big)\mathbf{T}^{\mathbf{d}}_{n,h} \to \mu\big(|\mathbf{T}^{\mathbf{d}}_{n}|\big)\mathbf{T}^{\mathbf{d}}_{n} \quad \text{strongly in } L_{n+1}(\Omega)^{d\times d}_{\text{sym}}.$$

Together with the strong convergence (5.30) these then imply that, for each $n \in \mathbb{N}$,

$$\lim_{h\to 0_+} c(\mathbf{T}_{n,h}; \mathbf{T}_{n,h}, \Pi_h \mathbf{S}) = c(\mathbf{T}_n; \mathbf{T}_n, \mathbf{S}) \qquad \forall\, \mathbf{S} \in L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\mathrm{sym}}. \tag{5.32}$$

Finally, we consider the first term on the left-hand side of equation (5.29). By inequality (3.4),

$$\left| \frac{\mathrm{Tr}(\mathbf{T}_{n,h})}{|\mathrm{Tr}(\mathbf{T}_{n,h})|^{1-\frac{1}{n}}} - \frac{\mathrm{Tr}(\mathbf{T}_n)}{|\mathrm{Tr}(\mathbf{T}_n)|^{1-\frac{1}{n}}} \right| \leq 2^{1-\frac{1}{n}} d^{\frac{1}{2n}} \, |\mathbf{T}_{n,h} - \mathbf{T}_n|^{\frac{1}{n}}.$$

Thus, because of the strong convergence (5.16), we have that, as $h \to 0_+$,

$$\frac{\mathrm{Tr}(\mathbf{T}_{n,h})}{|\mathrm{Tr}(\mathbf{T}_{n,h})|^{1-\frac{1}{n}}}\mathbf{I} \to \frac{\mathrm{Tr}(\mathbf{T}_n)}{|\mathrm{Tr}(\mathbf{T}_n)|^{1-\frac{1}{n}}}\mathbf{I} \qquad \text{strongly in } L_p(\Omega)^{d\times d}_{\mathrm{sym}} \quad \forall\, p \in [1, n+1).$$

Furthermore, by the uniform bound (5.13), for each fixed $n \in \mathbb{N}$,

$$\left(\frac{\mathrm{Tr}(\mathbf{T}_{n,h})}{|\mathrm{Tr}(\mathbf{T}_{n,h})|^{1-\frac{1}{n}}}\mathbf{I} \right)_{h\in(0,1]}$$

is a bounded sequence in $L_{n+1}(\Omega)^{d\times d}_{\mathrm{sym}}$, which therefore has a weakly convergent subsequence (not indicated), whose (weak) limit in $L_{n+1}(\Omega)^{d\times d}_{\mathrm{sym}}$, by the uniqueness of the weak limit, coincides with

$$\frac{\mathrm{Tr}(\mathbf{T}_n)}{|\mathrm{Tr}(\mathbf{T}_n)|^{1-\frac{1}{n}}}\mathbf{I}.$$

Hence, as $h \to 0_+$,

$$\frac{\mathrm{Tr}(\mathbf{T}_{n,h})}{|\mathrm{Tr}(\mathbf{T}_{n,h})|^{1-\frac{1}{n}}}\mathbf{I} \rightharpoonup \frac{\mathrm{Tr}(\mathbf{T}_n)}{|\mathrm{Tr}(\mathbf{T}_n)|^{1-\frac{1}{n}}}\mathbf{I} \qquad \text{weakly in } L_{n+1}(\Omega)^{d\times d}_{\mathrm{sym}}.$$

By an identical argument,

$$\frac{\mathbf{T}^{\mathbf{d}}_{n,h}}{|\mathbf{T}^{\mathbf{d}}_{n,h}|^{1-\frac{1}{n}}} \rightharpoonup \frac{\mathbf{T}^{\mathbf{d}}_{n}}{|\mathbf{T}^{\mathbf{d}}_{n}|^{1-\frac{1}{n}}} \qquad \text{weakly in } L_{n+1}(\Omega)^{d\times d}_{\mathrm{sym}}.$$

By combining these two weak convergence results with the strong convergence result (5.30) we deduce that, for each $n \in \mathbb{N}$,

$$\lim_{h\to 0_+} a_n(\mathbf{T}_{n,h}, \Pi_h \mathbf{S}) = a_n(\mathbf{T}_n, \mathbf{S}) \qquad \forall\, \mathbf{S} \in L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\mathrm{sym}}. \tag{5.33}$$

Using the convergence results (5.31), (5.32) and (5.33) we can now pass to the limit $h \to 0_+$ in equation (5.29) to deduce that

$$a_n(\mathbf{T}_n, \mathbf{S}) + c(\mathbf{T}_n; \mathbf{T}_n, \mathbf{S}) - b(\mathbf{S}, \overline{\mathbf{u}_{n,h}}) = 0 \qquad \forall\, \mathbf{S} \in L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\rm sym}. \tag{5.34}$$

By subtracting equation (5.34) from equation (3.2) we deduce that

$$b(\mathbf{S}, \overline{\mathbf{u}_{n,h}} - \mathbf{u}_n) = 0 \qquad \forall\, \mathbf{S} \in L_{1+\frac{1}{n}}(\Omega)^{d\times d}_{\rm sym}.$$

Hence,

$$\varepsilon(\overline{\mathbf{u}_{n,h}} - \mathbf{u}_n) = \mathbf{0} \qquad \text{in } L_{n+1}(\Omega)^{d\times d}_{\rm sym}.$$

Thus, by noting inequality (1.11) we deduce that

$$\overline{\mathbf{u}_{n,h}} - \mathbf{u}_n = \mathbf{0} \qquad \text{in } W^{1,n+1}_0(\Omega)^d.$$

In other words, $\overline{\mathbf{u}_{n,h}} = \mathbf{u}_n \in W^{1,n+1}_0(\Omega)^d$, as has been asserted above.

The strong convergence (5.19) in $L_n(\Omega)^d$ for $n \geq 2$ follows by an argument that we have already used, so we only sketch the proof. For any $\mathbf{S}_h \in \mathbb{M}_{n,h}$, the constitutive relations in (3.2) and (5.3) imply

$$\int_\Omega (\varepsilon(\mathbf{u}_{n,h}) - \varepsilon(\mathbf{u}_n)) : \mathbf{S}_h \,\mathrm{d}\mathbf{x} = \int_\Omega (\mathcal{A}_n(\mathbf{T}_{n,h}) - \mathcal{A}_n(\mathbf{T}_n)) : \mathbf{S}_h \,\mathrm{d}\mathbf{x} \leq \|\mathcal{A}_n(\mathbf{T}_{n,h}) - \mathcal{A}_n(\mathbf{T}_n))\|_{L_n(\Omega)} \|\mathbf{S}_h\|_{L_{\frac{n}{n-1}}(\Omega)}.$$

Now, using an argument similar to the one leading to (3.4), we find that

$$\begin{aligned} &\frac{1}{n}\int_\Omega \left| \frac{\mathrm{Tr}(\mathbf{T}_{n,h})}{|\mathrm{Tr}(\mathbf{T}_{n,h})|^{1-\frac{1}{n}}} - \frac{\mathrm{Tr}(\mathbf{T}_n)}{|\mathrm{Tr}(\mathbf{T}_n)|^{1-\frac{1}{n}}} \right|^n \mathrm{d}\mathbf{x} + \frac{1}{n}\int_\Omega \left| \frac{\mathbf{T}^{\mathbf{d}}_{n,h}}{|\mathbf{T}^{\mathbf{d}}_{n,h}|^{1-\frac{1}{n}}} - \frac{\mathbf{T}^{\mathbf{d}}_{n}}{|\mathbf{T}^{\mathbf{d}}_{n}|^{1-\frac{1}{n}}} \right|^n \mathrm{d}\mathbf{x} \\ &\qquad \leq C\|\mathbf{T}_{n,h} - \mathbf{T}_n\|_{L_1(\Omega)} \to 0 \qquad \text{as } h \to 0_+, \end{aligned} \tag{5.35}$$

for a constant C depending only on d and n. For the monotone part, $\mathcal{A}$, in $\mathcal{A}_n$ (cf. (4.16)), we invoke a similar argument to the one used in Lemma 3.1 to deduce that

$$\|\mathcal{A}(\mathbf{T}_{n,h}) - \mathcal{A}(\mathbf{T}_n)\|_{L_n(\Omega)} \to 0 \qquad \text{as } h \to 0_+.$$

Hence, in conjunction with (5.35), we arrive at

$$\frac{1}{\|\mathbf{S}_h\|_{L_{\frac{n}{n-1}}(\Omega)}} \int_\Omega (\varepsilon(\mathbf{u}_{n,h}) - \varepsilon(\mathbf{u}_n)) : \mathbf{S}_h \,\mathrm{d}\mathbf{x} \to 0 \qquad \text{as } h \to 0_+.$$

Using the decomposition $\varepsilon(\mathbf{u}_{n,h}) - \varepsilon(\mathbf{u}_n) = \varepsilon(\mathbf{u}_{n,h}) - \Pi_h(\varepsilon(\mathbf{u}_n)) + \Pi_h(\varepsilon(\mathbf{u}_n)) - \varepsilon(\mathbf{u}_n)$ we write

$$\frac{1}{\|\mathbf{S}_h\|_{L_{\frac{n}{n-1}}(\Omega)}} \int_\Omega (\varepsilon(\mathbf{u}_{n,h}) - \Pi_h(\varepsilon(\mathbf{u}_n))) : \mathbf{S}_h \, \mathrm{d}\mathbf{x}$$
$$= \frac{1}{\|\mathbf{S}_h\|_{L_{\frac{n}{n-1}}(\Omega)}} \int_\Omega (\varepsilon(\mathbf{u}_{n,h}) - \varepsilon(\mathbf{u}_n)) : \mathbf{S}_h \, \mathrm{d}\mathbf{x} - \frac{1}{\|\mathbf{S}_h\|_{L_{\frac{n}{n-1}}(\Omega)}} \int_\Omega (\Pi_h(\varepsilon(\mathbf{u}_n)) - \varepsilon(\mathbf{u}_n)) : \mathbf{S}_h \, \mathrm{d}\mathbf{x}.$$

Choosing $\mathbf{S}_h = (\varepsilon(\mathbf{u}_{n,h}) - \Pi_h(\varepsilon(\mathbf{u}_n)))|\varepsilon(\mathbf{u}_{n,h}) - \Pi_h(\varepsilon(\mathbf{u}_n))|^{n-2} \in \mathbb{M}_{n,h}$ yields

$$\|\varepsilon(\mathbf{u}_{n,h}) - \Pi_h(\varepsilon(\mathbf{u}_n))\|_{L_n(\Omega)} \to 0 \quad \text{as } h \to 0_+.$$

It remains to employ a density argument to deduce the strong convergence result (5.19).

The final claim in the statement of the lemma follows from the strong convergence results (4.11), $(4.12)_2$, (4.14), (5.16) and $(5.18)_2$, which together imply that, for any $\Omega_0 \subset\subset \Omega$,

$$\lim_{n\to\infty} \lim_{h\to 0_+} \|\mathbf{T}_{n,h} - \mathbf{T}\|_{L_1(\Omega_0)} = 0,$$
$$\lim_{n\to\infty} \lim_{h\to 0_+} \|\mathbf{u}_{n,h} - \mathbf{u}\|_{C(\overline{\Omega})} = 0$$

and

$$\lim_{n\to\infty} \lim_{h\to 0_+} \|\varepsilon(\mathbf{u}_{n,h} - \mathbf{u})\|_{L_p(\Omega_0)} = 0 \qquad \forall\, p \in [1,\infty).$$

The assertions concerning the uniqueness of $\mathbf{u}$ and $\mathbf{T}$ follow from Theorem 2.1(c). □

Hypotheses (A3′) and (A4) adopted in the statement of Lemma 5.4 guarantee that the derivatives of the functions $s \in \mathbb{R} \mapsto \lambda(s)s$ and $s \in \mathbb{R}_{\geq 0} \mapsto \mu(s)s$ are bounded below by 0 on $\mathbb{R}$ and $\mathbb{R}_{>0}$, respectively. These two functions are, in fact, Lipschitz continuous on any compact subinterval of $\mathbb{R}$ and $\mathbb{R}_{\geq 0}$, respectively. If they are assumed to be globally Hölder continuous on $\mathbb{R}$ and $\mathbb{R}_{\geq 0}$, respectively, with Hölder exponent $\beta \in (0,1]$, then an error inequality holds, for all $n \in \mathbb{N}$, in the limit of $h \to 0_+$, as we shall now show.

THEOREM 5.5 In addition to the assumptions of Lemma 5.4, let us also suppose that the functions $s \in \mathbb{R} \mapsto \lambda(s)s \in \mathbb{R}$ and $\mathbf{S} \in \mathbb{R}^{d\times d}_{\mathrm{sym}} \mapsto \mu(|\mathbf{S}|)\mathbf{S} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$ are Hölder continuous with exponent $\beta \in (0,1]$, i.e., there exists a positive constant Λ such that

$$|\lambda(r)r - \lambda(s)s| \leq \Lambda |r-s|^\beta \quad \forall\, r,s \in \mathbb{R}, \quad |\mu(|\mathbf{R}|)\mathbf{R} - \mu(|\mathbf{S}|)\mathbf{S}| \leq \Lambda |\mathbf{R}-\mathbf{S}|^\beta \quad \forall\, \mathbf{R},\mathbf{S} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}. \tag{5.36}$$

Then, assuming that $\mathbf{T}_n \in L_\infty(\Omega)^{d\times d}_{\mathrm{sym}}$ for $n \geq 2$, the following error bound holds:

$$\int_\Omega \Phi_n(|\mathbf{T}_{n,h} - \Pi_h \mathbf{T}_n|) \, \mathrm{d}\mathbf{x}$$
$$\leq C(d,\Lambda,\beta,n,\mathrm{K}_n,\mathrm{K}) \left(\inf_{\mathbf{v}_h \in \mathbb{X}_{n,h}} \int_\Omega \Phi_n^*(|\varepsilon(\mathbf{v}_h - \mathbf{u}_n)|) \, \mathrm{d}\mathbf{x} + \int_\Omega \Phi_n^*(|\mathbf{T}_n - \Pi_h \mathbf{T}_n|^{\min(\beta,\frac{1}{n})}) \, \mathrm{d}\mathbf{x} \right). \tag{5.37}$$

Here

$$\Phi_n(s) := \frac{s^2}{(1+s)^{1-\frac{1}{n}}}, \qquad s \in [0,\infty), \quad n \in \mathbb{N},$$

Φ_n^*, defined by $\Phi_n^*(s) := \sup_{t\geq 0}(st - \Phi_n(t))$ for $s \in [0,\infty)$, is the convex conjugate of the function Φ_n, $\mathrm{K}_n := \max(1, \|\mathbf{T}_n\|_{L_\infty(\Omega)})$ and $\mathrm{K} = \mathrm{K}(n)$ is a positive constant that will be specified in the proof. When $n = 1$ inequality (5.37) holds without the additional assumption that $\mathbf{T}_n \in L_\infty(\Omega)^{d\times d}_{\mathrm{sym}}$.

Proof. We proceed similarly to the proof of Lemma 5.4. From relations (3.2) and (5.3) we have for all $\mathbf{S}_h \in \mathbb{M}_{n,h}$ that

$$\int_\Omega (\mathcal{A}_n(\mathbf{T}_{n,h}) - \mathcal{A}_n(\Pi_h\mathbf{T}_n)) : \mathbf{S}_h \,\mathrm{d}\mathbf{x} = \int_\Omega \varepsilon(\mathbf{u}_{n,h} - \mathbf{u}_n) : \mathbf{S}_h \,\mathrm{d}\mathbf{x} + \int_\Omega (\mathcal{A}_n(\mathbf{T}_n) - \mathcal{A}_n(\Pi_h\mathbf{T}_n)) : \mathbf{S}_h \,\mathrm{d}\mathbf{x}.$$

The choice $\mathbf{S}_h = \mathbf{T}_{n,h} - \Pi_h\mathbf{T}_n \in \mathbb{V}_{n,h}$ guarantees that

$$\int_\Omega \varepsilon(\mathbf{v}_h) : \mathbf{S}_h \,\mathrm{d}\mathbf{x} = 0 \qquad \forall\, \mathbf{v}_h \in \mathbb{X}_{n,h}.$$

Thus, by defining, for any $\mathbf{v}_h$ in $\mathbb{X}_{n,h}$,

$$\mathbf{U}_{n,h} := \varepsilon(\mathbf{v}_h - \mathbf{u}_n) + \big(\mathcal{A}_n(\mathbf{T}_n) - \mathcal{A}_n(\Pi_h\mathbf{T}_n)\big),$$

and proceeding similarly to the proof of inequality (5.21), we have

$$\begin{aligned} \frac{1}{n^2}\int_\Omega \frac{|\mathrm{Tr}(\mathbf{T}_{n,h}) - \mathrm{Tr}(\Pi_h\mathbf{T}_n)|^2}{(|\mathrm{Tr}(\Pi_h\mathbf{T}_n)| + |\mathrm{Tr}(\mathbf{T}_{n,h}) - \mathrm{Tr}(\Pi_h\mathbf{T}_n)|)^{1-\frac{1}{n}}}\,\mathrm{d}\mathbf{x} \\ + \frac{1}{n^2}\int_\Omega \frac{\big|\mathbf{T}^{\mathbf{d}}_{n,h} - \Pi_h\mathbf{T}^{\mathbf{d}}_n\big|^2}{\big(\big|\Pi_h\mathbf{T}^{\mathbf{d}}_n\big| + \big|\mathbf{T}^{\mathbf{d}}_{n,h} - \Pi_h\mathbf{T}^{\mathbf{d}}_n\big|\big)^{1-\frac{1}{n}}}\,\mathrm{d}\mathbf{x} \\ \leq \int_\Omega \mathbf{U}_{n,h} : (\mathbf{T}_{n,h} - \Pi_h\mathbf{T}_n)\,\mathrm{d}\mathbf{x}. \end{aligned} \tag{5.38}$$

Thanks to equality (1.12),

$$\big|\Pi_h\mathbf{T}^{\mathbf{d}}_n(\mathbf{x})\big|^2 + \frac{1}{d}|\mathrm{Tr}(\Pi_h\mathbf{T}_n)(\mathbf{x})|^2 = |\Pi_h\mathbf{T}_n(\mathbf{x})|^2 \leq \|\Pi_h\mathbf{T}_n\|^2_{L_\infty(\Omega)} \leq \|\mathbf{T}_n\|^2_{L_\infty(\Omega)} \qquad \text{for a.e. } \mathbf{x} \in \Omega.$$

Thus, by denoting $\mathrm{K}_n := \max(1, \|\mathbf{T}_n\|_{L_\infty(\Omega)})$, it follows that

$$\|\Pi_h\mathbf{T}^{\mathbf{d}}_n\|_{L_\infty(\Omega)} \leq \mathrm{K}_n \qquad \text{and} \qquad \|\mathrm{Tr}(\Pi_h\mathbf{T}_n)\|_{L_\infty(\Omega)} \leq d^{\frac{1}{2}}\mathrm{K}_n.$$

Hence we have from inequality (5.38) that

$$\frac{1}{n^2}\int_\Omega \frac{|\mathrm{Tr}(\mathbf{T}_{n,h}) - \mathrm{Tr}(\Pi_h \mathbf{T}_n)|^2}{(d^{\frac{1}{2}}\mathrm{K}_n + |\mathrm{Tr}(\mathbf{T}_{n,h}) - \mathrm{Tr}(\Pi_h \mathbf{T}_n)|)^{1-\frac{1}{n}}}\,\mathrm{d}\mathbf{x} + \frac{1}{n^2}\int_\Omega \frac{\left|\mathbf{T}^{\mathbf{d}}_{n,h} - \Pi_h \mathbf{T}^{\mathbf{d}}_n\right|^2}{\left(\mathrm{K}_n + \left|\mathbf{T}^{\mathbf{d}}_{n,h} - \Pi_h \mathbf{T}^{\mathbf{d}}_n\right|\right)^{1-\frac{1}{n}}}\,\mathrm{d}\mathbf{x}$$
$$\leq \int_\Omega \mathbf{U}_{n,h} : (\mathbf{T}_{n,h} - \Pi_h \mathbf{T}_n)\,\mathrm{d}\mathbf{x}.$$

Because $\mathrm{K}_n \geq 1$ and by noting the decomposition $\mathbf{T} = \frac{1}{d}\mathrm{Tr}(\mathbf{T})\mathbf{I} + \mathbf{T}^{\mathbf{d}}$, the above inequality implies

$$\begin{aligned}
&\frac{1}{n^2}\left(d^{\frac{1}{2}}\mathrm{K}_n\right)^{\frac{1}{n}-1}\int_\Omega \frac{|\mathrm{Tr}(\mathbf{T}_{n,h}) - \mathrm{Tr}(\Pi_h \mathbf{T}_n)|^2}{(1 + |\mathrm{Tr}(\mathbf{T}_{n,h}) - \mathrm{Tr}(\Pi_h \mathbf{T}_n)|)^{1-\frac{1}{n}}}\,\mathrm{d}\mathbf{x} \\
&\qquad + \frac{1}{n^2}(\mathrm{K}_n)^{\frac{1}{n}-1}\int_\Omega \frac{\left|\mathbf{T}^{\mathbf{d}}_{n,h} - \Pi_h \mathbf{T}^{\mathbf{d}}_n\right|^2}{\left(1 + \left|\mathbf{T}^{\mathbf{d}}_{n,h} - \Pi_h \mathbf{T}^{\mathbf{d}}_n\right|\right)^{1-\frac{1}{n}}}\,\mathrm{d}\mathbf{x} \\
&\leq \frac{1}{d}\int_\Omega \mathrm{Tr}(\mathbf{U}_{n,h})\,\mathrm{Tr}(\mathbf{T}_{n,h} - \Pi_h \mathbf{T}_n)\,\mathrm{d}\mathbf{x} + \int_\Omega \mathbf{U}^{\mathbf{d}}_{n,h} : (\mathbf{T}_{n,h} - \Pi_h \mathbf{T}_n)^{\mathbf{d}}\,\mathrm{d}\mathbf{x} \\
&\leq \frac{1}{d}\int_\Omega |\mathrm{Tr}(\mathbf{U}_{n,h})|\,|\mathrm{Tr}(\mathbf{T}_{n,h}) - \mathrm{Tr}(\Pi_h \mathbf{T}_n)|\,\mathrm{d}\mathbf{x} + \int_\Omega \left|\mathbf{U}^{\mathbf{d}}_{n,h}\right|\left|\mathbf{T}^{\mathbf{d}}_{n,h} - \Pi_h \mathbf{T}^{\mathbf{d}}_n\right|\mathrm{d}\mathbf{x}.
\end{aligned} \tag{5.39}$$

Let us consider the function $\Phi_n : \mathbb{R} \to \mathbb{R}_{\geq 0}$ defined by

$$\Phi_n(s) := \frac{s^2}{(1 + |s|)^{1-\frac{1}{n}}}, \qquad n \in \mathbb{N}. \tag{5.40}$$

The values of s of interest to us below will always be in the range $[0, \infty)$, and therefore the absolute value sign appearing in the denominator of $\Phi_n(s)$ can be ignored for such s.

Clearly, $\Phi_n(0) = 0$, Φ_n is even, continuous, strictly monotonic increasing for $s \geq 0$ and convex, with

$$\Phi_n(s) \asymp s^2 \quad \text{as } s \to 0_+ \qquad \text{and} \qquad \Phi_n(s) \asymp s^{1+\frac{1}{n}} \quad \text{as } s \to +\infty. \tag{5.41}$$

Here $A \asymp B$ means that there exist positive constants c and $\tilde{c}$ independent of A and B such that $cB \leq A \leq \tilde{c}B$. Following Rao & Ren (2002), a function $\Phi : \mathbb{R} \to \mathbb{R}_{\geq 0}$ is called a nice Young function (*N-function*), if (i) Φ is even and convex; (ii) $\Phi(s) = 0$ if and only if $s = 0$ and (iii) $\lim_{s\to 0} \Phi(s)/s = 0$ and $\lim_{s\to+\infty} \Phi(s)/s = +\infty$.

Hence, Φ_n is an N-function. Simple calculations show that

$$\Phi_n(2s) \leq 4\Phi_n(s) \quad \forall\, s \in [0, \infty) \quad \text{and} \quad \frac{1}{2c}\Phi_n(cs) \geq \Phi_n(s) \quad \forall\, s \in [0, \infty), \quad \forall\, c \geq 2^n; \tag{5.42}$$

therefore, Φ_n satisfies the Δ_2 and ∇_2 conditions on $[0, \infty)$ (cf. Rao & Ren, 2002, Definition 1 on p. 2). Now, let Φ_n^* denote the convex conjugate of the function Φ_n. Then (Φ_n, Φ_n^*) is a pair of complementary

N-functions and, by Rao & Ren (2002, Theorem 2 on p. 3), Φ_n^* also satisfies the Δ_2 and ∇_2 conditions on $[0,\infty)$; i.e., there exists a constant $\mathrm{K} = \mathrm{K}(n) > 2$ such that

$$\Phi_n^*(2s) \le \mathrm{K}\Phi_n^*(s) \qquad \forall\, s \in [0,\infty), \tag{5.43}$$

and there exists a constant $c = c(n) > 1$ such that

$$\frac{1}{2c}\Phi_n^*(cs) \ge \Phi_n^*(s) \qquad \forall\, s \in [0,\infty).$$

More precisely, by inequality (5.26),

$$2^{\frac{1}{n}-1}\min\left(s^2, s^{1+\frac{1}{n}}\right) \le \Phi_n(s) \le \min\left(s^2, s(1+s)^{\frac{1}{n}}\right) \qquad \forall\, s \in [0,\infty).$$

By recalling that $\Phi_n^*(s) := \sup_{t\ge 0}(st - \Phi_n(t))$, we get from (5.41) that

$$\Phi_n^*(s) \asymp \frac{1}{4}s^2 \quad \text{as } s \to 0_+ \qquad \text{and} \qquad \Phi_n^*(s) \asymp \frac{s^{n+1}}{n+1}\left(\frac{n}{n+1}\right)^n \quad \text{as } s \to +\infty. \tag{5.44}$$

Therefore, there exist positive constants $c_{1,n}$ and $c_{2,n}$, with $c_{1,n} \le c_{2,n}$, such that

$$0 \le \Phi_n^*(s) \le c_{1,n}s^2 \qquad \forall\, s \in [0,1]$$

and

$$c_{1,n} \le \Phi_n^*(s) \le c_{2,n}s^{n+1} \qquad \forall\, s \in [1,\infty).$$

Reverting to (5.39), by the Fenchel–Young inequality, for any real number $\delta > 0$,

$$\begin{aligned}
&\frac{1}{n^2}\left(d^{\frac{1}{2}}\mathrm{K}_n\right)^{\frac{1}{n}-1}\int_\Omega \Phi_n(|\mathrm{Tr}(\mathbf{T}_{n,h}) - \mathrm{Tr}(\Pi_h\mathbf{T}_n)|)\,\mathrm{d}\mathbf{x} + \frac{1}{n^2}(\mathrm{K}_n)^{\frac{1}{n}-1}\int_\Omega \Phi_n\left(\left|\mathbf{T}^{\mathbf{d}}_{n,h} - \Pi_h\mathbf{T}^{\mathbf{d}}_n\right|\right)\mathrm{d}\mathbf{x}\\
&\quad\le \frac{1}{d}\int_\Omega |\mathrm{Tr}(\mathbf{U}_{n,h})|\,|\mathrm{Tr}(\mathbf{T}_{n,h}) - \mathrm{Tr}(\Pi_h\mathbf{T}_n)|\,\mathrm{d}\mathbf{x} + \int_\Omega \left|\mathbf{U}^{\mathbf{d}}_{n,h}\right|\left|\mathbf{T}^{\mathbf{d}}_{n,h} - \Pi_h\mathbf{T}^{\mathbf{d}}_n\right|\mathrm{d}\mathbf{x}\\
&\quad= \frac{1}{d\delta}\int_\Omega |\mathrm{Tr}(\mathbf{U}_{n,h})|\,\delta\,|\mathrm{Tr}(\mathbf{T}_{n,h}) - \mathrm{Tr}(\Pi_h\mathbf{T}_n)|\,\mathrm{d}\mathbf{x} + \frac{1}{\delta}\int_\Omega \left|\mathbf{U}^{\mathbf{d}}_{n,h}\right|\delta\left|\mathbf{T}^{\mathbf{d}}_{n,h} - \Pi_h\mathbf{T}^{\mathbf{d}}_n\right|\mathrm{d}\mathbf{x}\\
&\quad\le \frac{1}{d\delta}\int_\Omega \Phi_n(\delta|\mathrm{Tr}(\mathbf{T}_{n,h}) - \mathrm{Tr}(\Pi_h\mathbf{T}_n)|)\,\mathrm{d}\mathbf{x} + \frac{1}{d\delta}\int_\Omega \Phi_n^*(|\mathrm{Tr}(\mathbf{U}_{n,h})|)\,\mathrm{d}\mathbf{x}\\
&\qquad+ \frac{1}{\delta}\int_\Omega \Phi_n\left(\delta\left|\mathbf{T}^{\mathbf{d}}_{n,h} - \Pi_h\mathbf{T}^{\mathbf{d}}_n\right|\right)\mathrm{d}\mathbf{x} + \frac{1}{\delta}\int_\Omega \Phi_n^*\left(\left|\mathbf{U}^{\mathbf{d}}_{n,h}\right|\right)\mathrm{d}\mathbf{x}.
\end{aligned}$$

Clearly, for any $a \in \mathbb{R}_{\ge 0}$ and $\delta \in (0,1]$, we have

$$\Phi_n(\delta a) = \frac{\delta^2 a^2}{(1+\delta a)^{1-\frac{1}{n}}} \le \delta^{1+\frac{1}{n}}\Phi_n(a).$$

Hence,

$$\begin{aligned}
&\frac{1}{n^2}\big(d^{\frac{1}{2}}\mathrm{K}_n\big)^{\frac{1}{n}-1}\int_\Omega \Phi_n(|\mathrm{Tr}(\mathbf{T}_{n,h})-\mathrm{Tr}(\Pi_h\mathbf{T}_n)|)\,\mathrm{d}\mathbf{x}+\frac{1}{n^2}(\mathrm{K}_n)^{\frac{1}{n}-1}\int_\Omega \Phi_n\big(\big|\mathbf{T}^{\mathbf{d}}_{n,h}-\Pi_h\mathbf{T}^{\mathbf{d}}_n\big|\big)\,\mathrm{d}\mathbf{x}\\
&\quad\le \frac{\delta^{\frac{1}{n}}}{d}\int_\Omega \Phi_n(|\mathrm{Tr}(\mathbf{T}_{n,h})-\mathrm{Tr}(\Pi_h\mathbf{T}_n)|)\,\mathrm{d}\mathbf{x}+\frac{1}{d\delta}\int_\Omega \Phi_n^*(|\mathrm{Tr}(\mathbf{U}_{n,h})|)\,\mathrm{d}\mathbf{x}\\
&\qquad+\delta^{\frac{1}{n}}\int_\Omega \Phi_n\big(\big|\mathbf{T}^{\mathbf{d}}_{n,h}-\Pi_h\mathbf{T}^{\mathbf{d}}_n\big|\big)\,\mathrm{d}\mathbf{x}+\frac{1}{\delta}\int_\Omega \Phi_n^*\big(\big|\mathbf{U}^{\mathbf{d}}_{n,h}\big|\big)\,\mathrm{d}\mathbf{x}.
\end{aligned}$$

Let $\delta_1,\delta_2>0$ be such that

$$\frac{d}{2n^2}\big(d^{\frac{1}{2}}\mathrm{K}_n\big)^{\frac{1}{n}-1}=\delta_1^{\frac{1}{n}}\quad\text{and}\quad\frac{1}{2n^2}(\mathrm{K}_n)^{\frac{1}{n}-1}=\delta_2^{\frac{1}{n}}.$$

Thus, with $\delta:=\min(1,\delta_1,\delta_2)$, we have

$$\begin{aligned}
&\int_\Omega \Phi_n(|\mathrm{Tr}(\mathbf{T}_{n,h})-\mathrm{Tr}(\Pi_h\mathbf{T}_n)|)\,\mathrm{d}\mathbf{x}+\int_\Omega \Phi_n\big(\big|\mathbf{T}^{\mathbf{d}}_{n,h}-\Pi_h\mathbf{T}^{\mathbf{d}}_n\big|\big)\,\mathrm{d}\mathbf{x}\\
&\qquad\le C(d,n,\mathrm{K}_n)\left(\int_\Omega \Phi_n^*(|\mathrm{Tr}(\mathbf{U}_{n,h})|)\,\mathrm{d}\mathbf{x}+\int_\Omega \Phi_n^*\big(\big|\mathbf{U}^{\mathbf{d}}_{n,h}\big|\big)\,\mathrm{d}\mathbf{x}\right).
\end{aligned}\tag{5.45}$$

Now assumptions (5.36) and (3.4) yield

$$\begin{aligned}
|\mathbf{U}_{n,h}|&\le|\varepsilon(\mathbf{v}_h-\mathbf{u}_n)|+|\mathcal{A}_n(\mathbf{T}_n)-\mathcal{A}_n(\Pi_h\mathbf{T}_n)|\\
&\le|\varepsilon(\mathbf{v}_h-\mathbf{u}_n)|+C\left(|\mathbf{T}_n-\Pi_h\mathbf{T}_n|^{\frac{1}{n}}+|\mathbf{T}_n-\Pi_h\mathbf{T}_n|^{\beta}\right).
\end{aligned}$$

As Φ_n^* is an N-function, it is strictly monotonic increasing (cf. Rao & Ren, 2002, the top of p. 2) and convex, and therefore by (5.43),

$$\begin{aligned}
\Phi_n^*(|\mathbf{U}_{n,h}|)&\le\Phi_n^*(|\varepsilon(\mathbf{v}_h-\mathbf{u}_n)|+|\mathcal{A}_n(\mathbf{T}_n)-\mathcal{A}_n(\Pi_h\mathbf{T}_n)|)\\
&\le\frac{1}{2}\left(\Phi_n^*(2|\varepsilon(\mathbf{v}_h-\mathbf{u}_n)|)+\Phi_n^*(2|\mathcal{A}_n(\mathbf{T}_n)-\mathcal{A}_n(\Pi_h\mathbf{T}_n)|)\right)\\
&\le\frac{\mathrm{K}}{2}\left(\Phi_n^*(|\varepsilon(\mathbf{v}_h-\mathbf{u}_n)|)+\Phi_n^*(|\mathcal{A}_n(\mathbf{T}_n)-\mathcal{A}_n(\Pi_h\mathbf{T}_n)|)\right).
\end{aligned}\tag{5.46}$$

In order to proceed we need to bound the right-hand side of the last inequality and that involves comparing

$$\mathcal{A}_n(\mathbf{T}_n):=\lambda(\mathrm{Tr}(\mathbf{T}_n))\mathrm{Tr}(\mathbf{T}_n)\mathbf{I}+\mu\big(\big|\mathbf{T}^{\mathbf{d}}_n\big|\big)\mathbf{T}^{\mathbf{d}}_n+\frac{\mathrm{Tr}(\mathbf{T}_n)\mathbf{I}}{n|\mathrm{Tr}(\mathbf{T}_n)|^{1-\frac{1}{n}}}+\frac{\mathbf{T}^{\mathbf{d}}_n}{n\big|\mathbf{T}^{\mathbf{d}}_n\big|^{1-\frac{1}{n}}}$$

with

$$\mathcal{A}_n(\Pi_h\mathbf{T}_n) := \lambda(\mathrm{Tr}(\Pi_h\mathbf{T}_n))\mathrm{Tr}(\Pi_h\mathbf{T}_n)\mathbf{I} + \mu\big(\big|\Pi_h\mathbf{T}_n^{\mathbf{d}}\big|\big)\Pi_h\mathbf{T}_n^{\mathbf{d}} + \frac{\mathrm{Tr}(\Pi_h\mathbf{T}_n)\mathbf{I}}{n|\mathrm{Tr}(\Pi_h\mathbf{T}_n)|^{1-\frac{1}{n}}} + \frac{\Pi_h\mathbf{T}_n^{\mathbf{d}}}{n\big|\Pi_h\mathbf{T}_n^{\mathbf{d}}\big|^{1-\frac{1}{n}}}.$$

We have from inequalities (5.36) and (3.4) that

$$\begin{aligned}
|\mathcal{A}_n(\mathbf{T}_n) - \mathcal{A}_n(\Pi_h\mathbf{T}_n)| &\le d^{\frac{1}{2}}\Lambda|\mathrm{Tr}(\mathbf{T}_n) - \Pi_h\mathrm{Tr}(\mathbf{T}_n)|^{\beta} + \Lambda\big|\mathbf{T}_n^{\mathbf{d}} - \Pi_h\mathbf{T}_n^{\mathbf{d}}\big|^{\beta} \\
&\quad + \frac{d^{\frac{1}{2n}}}{n}\,2^{1-\frac{1}{n}}\,|\mathrm{Tr}(\mathbf{T}_n) - \mathrm{Tr}(\Pi_h\mathbf{T}_n)|^{\frac{1}{n}} \\
&\quad + C(d,n)\big|\mathbf{T}_n^{\mathbf{d}} - \Pi_h\mathbf{T}_n^{\mathbf{d}}\big|^{\frac{1}{n}} \\
&\le C(d,\Lambda,\beta)\,|\mathbf{T}_n - \Pi_h\mathbf{T}_n|^{\beta} + C(d,n)\,|\mathbf{T}_n - \Pi_h\mathbf{T}_n|^{\frac{1}{n}} \\
&\le C(d,\Lambda,\beta,n,\mathrm{K}_n)|\mathbf{T}_n - \Pi_h\mathbf{T}_n|^{\min(\beta,\frac{1}{n})}.
\end{aligned}$$

By inequality (5.43), $\Phi_n^*(2^\ell s) \le \mathrm{K}^\ell\Phi_n^*(s)$ for all $s \in [0,\infty)$ and all $\ell \ge 1$. Hence, with

$$\ell := [\log_2 C(d,\Lambda,\beta,n,\mathrm{K}_n)] + 1$$

we have that $C(d,\Lambda,\beta,n,\mathrm{K}_n) \le 2^\ell$, whereby

$$\begin{aligned}
\Phi_n^*(|\mathcal{A}_n(\mathbf{T}_n) - \mathcal{A}_n(\Pi_h\mathbf{T}_n)|) &\le \Phi_n^*(C(d,\Lambda,\beta,n,\mathrm{K}_n)|\mathbf{T}_n - \Pi_h\mathbf{T}_n|^{\min(\beta,\frac{1}{n})}) \\
&\le \Phi_n^*(2^\ell|\mathbf{T}_n - \Pi_h\mathbf{T}_n|^{\min(\beta,\frac{1}{n})}) \\
&\le \mathrm{K}^\ell\Phi_n^*(|\mathbf{T}_n - \Pi_h\mathbf{T}_n|^{\min(\beta,\frac{1}{n})}).
\end{aligned}$$

By substituting this into inequality (5.46) we deduce that

$$\begin{aligned}
\Phi_n^*(|\mathbf{U}_{n,h}|) &\le \Phi_n^*(|\varepsilon(\mathbf{v}_h - \mathbf{u}_n)| + |\mathcal{A}_n(\mathbf{T}_n) - \mathcal{A}_n(\Pi_h\mathbf{T}_n)|) \\
&\le \frac{1}{2}\mathrm{K}\Phi_n^*(|\varepsilon(\mathbf{v}_h - \mathbf{u}_n)|) + \frac{1}{2}\mathrm{K}^{\ell+1}\Phi_n^*(|\mathbf{T}_n - \Pi_h\mathbf{T}_n|^{\min(\beta,\frac{1}{n})}).
\end{aligned}$$

We then substitute this into inequality (5.45) and note, once again, the monotonicity of Φ_n^*, which gives

$$\begin{aligned}
&\int_\Omega \Phi_n(|\mathrm{Tr}(\mathbf{T}_{n,h}) - \mathrm{Tr}(\Pi_h\mathbf{T}_n)|)\,\mathrm{d}\mathbf{x} + \int_\Omega \Phi_n\big(\big|\mathbf{T}_{n,h}^{\mathbf{d}} - \Pi_h\mathbf{T}_n^{\mathbf{d}}\big|\big)\,\mathrm{d}\mathbf{x} \\
&\quad \le C(d,\Lambda,\beta,n,\mathrm{K}_n,\mathrm{K})\left(\int_\Omega \Phi_n^*(|\varepsilon(\mathbf{v}_h - \mathbf{u}_n)|)\,\mathrm{d}\mathbf{x} + \int_\Omega \Phi_n^*(|\mathbf{T}_n - \Pi_h\mathbf{T}_n|^{\min(\beta,\frac{1}{n})})\,\mathrm{d}\mathbf{x}\right).
\end{aligned}$$

For any pair of numbers $a, b \in \mathbb{R}_{\geq 0}$, by (5.42) and convexity, we have $\Phi_n(a+b) \leq 2\Phi_n(a) + 2\Phi_n(b)$; hence, by inequality (1.13),

$$\int_\Omega \Phi_n(|\mathbf{T}_{n,h} - \Pi_h \mathbf{T}_n|)\,\mathrm{d}\mathbf{x} \leq C(d, \Lambda, \beta, n, \mathrm{K}_n, \mathrm{K}) \left(\int_\Omega \Phi_n^*(|\varepsilon(\mathbf{v}_h - \mathbf{u}_n)|)\,\mathrm{d}\mathbf{x} + \int_\Omega \Phi_n^*(|\mathbf{T}_n - \Pi_h \mathbf{T}_n|^{\min(\beta, \frac{1}{n})})\,\mathrm{d}\mathbf{x} \right).$$

As this inequality holds for all $\mathbf{v}_h \in \mathbb{X}_{n,h}$, the bound (5.37) directly follows. □

The error bound (5.37) can be restated in the following equivalent form. Given an N-function Ψ, let

$$\tilde{L}_\Psi(\Omega) := \left\{ \mathbf{S} : \Omega \to \mathbb{R}^{d\times d}_{\mathrm{sym}} \quad \text{measurable, such that} \quad \rho_\Psi(\mathbf{S}) := \int_\Omega \Psi(|\mathbf{S}(\mathbf{x})|)\,\mathrm{d}\mathbf{x} < \infty \right\};$$

the function $\rho_\Psi(\cdot)$ is called a *modular*. In terms of the modulars ρ_{Φ_n} and $\rho_{\Phi_n^*}$ the error bound (5.37) takes the form

$$\rho_{\Phi_n}(|\mathbf{T}_{n,h} - \Pi_h \mathbf{T}_n|) \leq C(d, \Lambda, \beta, n, \mathrm{K}_n, \mathrm{K}) \left(\inf_{\mathbf{v}_h \in \mathbb{X}_{n,h}} \rho_{\Phi_n^*}(|\varepsilon(\mathbf{v}_h - \mathbf{u}_n)|) + \rho_{\Phi_n^*}(|\mathbf{T}_n - \Pi_h \mathbf{T}_n|^{\min(\beta, \frac{1}{n})}) \right). \tag{5.47}$$

Here, as before,

$$\Phi_n(s) := \frac{s^2}{(1+s)^{1-\frac{1}{n}}}, \qquad s \in [0, \infty), \quad n \in \mathbb{N},$$

and Φ_n^* is the convex conjugate of Φ_n.

Under the above assumptions, convergence rates can be derived by strengthening the regularity hypothesis $\mathbf{T}_n \in L_\infty(\Omega)^{d\times d}_{\mathrm{sym}}$ from Theorem 5.5. Thus, for example, suppose that

$$\mathbf{T}_n \in W^{r,q}(\Omega)^{d\times d}_{\mathrm{sym}} \quad \text{with } 1 \geq r > \frac{d}{q} \quad \text{and} \quad \mathbf{u}_n \in W^{1+t,p}(\Omega)^d \cap W_0^{1,n+1}(\Omega)^d \quad \text{with } 1 \geq t > \frac{d}{p},$$

and $q, p \in (1, \infty]$, which ensure, by Morrey's embedding theorem, that

$$\mathbf{T}_n \in W^{r,q}(\Omega)^{d\times d}_{\mathrm{sym}} \hookrightarrow \mathcal{C}^{0,\gamma}(\overline{\Omega})^{d\times d}_{\mathrm{sym}} \hookrightarrow W^{\gamma,\infty}(\Omega)^{d\times d}_{\mathrm{sym}} \qquad \text{with } \gamma := r - \frac{d}{q}$$

and

$$\varepsilon(\mathbf{u}_n) \in W^{t,p}(\Omega)^{d\times d}_{\mathrm{sym}} \hookrightarrow \mathcal{C}^{0,\zeta}(\overline{\Omega})^{d\times d}_{\mathrm{sym}} \hookrightarrow W^{\zeta,\infty}(\Omega)^{d\times d}_{\mathrm{sym}} \qquad \text{with } \zeta := t - \frac{d}{p}.$$

With these stronger regularity hypotheses we then have

$$|\mathbf{T}_n(\mathbf{x}) - \Pi_h \mathbf{T}_n(\mathbf{x})| \leq C h_K^\gamma \|\mathbf{T}_n\|_{W^{\gamma,\infty}(K)} \leq C h^\gamma \|\mathbf{T}_n\|_{W^{\gamma,\infty}(\Omega)} \qquad \forall\, \mathbf{x} \in K, \quad \forall\, K \in \mathcal{T}_h.$$

Thus, thanks to the fact that Φ_n^* is monotonic increasing, and by the first asymptotic property in (5.44),

$$\rho_{\Phi_n^*}\left(|\mathbf{T}_n - \Pi_h\mathbf{T}_n|^{\min(\beta,\frac{1}{n})}\right) \leq \rho_{\Phi_n^*}\left(Ch^{\gamma\min(\beta,\frac{1}{n})}\|\mathbf{T}_n\|^{\min(\beta,\frac{1}{n})}_{W^{\gamma,\infty}(\Omega)}\right) \asymp Ch^{2\gamma\min(\beta,\frac{1}{n})}\|\mathbf{T}_n\|^{2\min(\beta,\frac{1}{n})}_{W^{\gamma,\infty}(\Omega)}$$

as $h \to 0_+$. Analogously,

$$\inf_{\mathbf{v}_h\in\mathbb{X}_{n,h}} \rho_{\Phi_n^*}(|\varepsilon(\mathbf{v}_h) - \varepsilon(\mathbf{u}_n)|) \leq \rho_{\Phi_n^*}(Ch^\zeta\|\varepsilon(\mathbf{u}_n)\|_{W^{\zeta,\infty}(\Omega)}) \asymp Ch^{2\zeta}\|\varepsilon(\mathbf{u}_n)\|^2_{W^{\zeta,\infty}(\Omega)} \qquad \text{as } h \to 0_+.$$

By substituting these bounds into the error inequality (5.47) we deduce that

$$\rho_{\Phi_n}(|\mathbf{T}_{n,h} - \Pi_h\mathbf{T}_n|) \leq C\left(h^{2\zeta}\|\varepsilon(\mathbf{u}_n)\|^2_{W^{\zeta,\infty}(\Omega)} + h^{2\gamma\min(\beta,\frac{1}{n})}\|\mathbf{T}_n\|^{2\min(\beta,\frac{1}{n})}_{W^{\gamma,\infty}(\Omega)}\right)$$

as $h \to 0_+$. In particular, if $\beta = \frac{1}{n}$ and $\zeta = \frac{\gamma}{n}$,

$$\rho_{\Phi_n}(|\mathbf{T}_{n,h} - \Pi_h\mathbf{T}_n|) \leq Ch^{2\frac{\gamma}{n}}\left(\|\varepsilon(\mathbf{u}_n)\|^2_{W^{\frac{\gamma}{n},\infty}(\Omega)} + \|\mathbf{T}_n\|^{\frac{2}{n}}_{W^{\gamma,\infty}(\Omega)}\right) \tag{5.48}$$

as $h \to 0_+$, where $\gamma \in (0,1]$ and $n \in \mathbb{N}$. The error bound (5.48) on $\rho_{\Phi_n}(|\mathbf{T}_{n,h} - \Pi_h\mathbf{T}_n|)$ can be used to derive bounds on norms of the error $|\mathbf{T}_{n,h} - \Pi_h\mathbf{T}_n|$. For example, in the special case when $n = 1$, we have $\Phi_n(s) = s^2$, and therefore

$$\|\mathbf{T}_{n,h} - \Pi_h\mathbf{T}_n\|_{L_2(\Omega)} \leq Ch^\gamma\left(\|\varepsilon(\mathbf{u}_n)\|_{W^{\gamma,\infty}(\Omega)} + \|\mathbf{T}_n\|_{W^{\gamma,\infty}(\Omega)}\right)$$

as $h \to 0_+$, where $\gamma \in (0,1]$. In this special case, the regularity requirements on $\mathbf{u}$ and $\mathbf{T}$ can, in fact, be relaxed to $\mathbf{u}_n \in W^{1+\gamma,2}(\Omega)^{d\times d}_{\text{sym}} \cap W^{1,2}_0(\Omega)^{d\times d}_{\text{sym}}$ and $\mathbf{T}_n \in W^{\gamma,2}(\Omega)^{d\times d}_{\text{sym}}$, $\gamma \in (0,1]$.

More generally, for $n \in \mathbb{N}$, we divide inequality (5.48) by $|\Omega|$, recall the definition of the modular $\rho_{\Phi_n}(\cdot)$ and apply Jensen's inequality on the left-hand side to deduce that

$$\Phi_n\left(\fint_\Omega |\mathbf{T}_{n,h} - \Pi_h\mathbf{T}_n|\,\mathrm{d}\mathbf{x}\right) \leq Ch^{2\frac{\gamma}{n}}\left(\|\varepsilon(\mathbf{u}_n)\|^2_{W^{\frac{\gamma}{n},\infty}(\Omega)} + \|\mathbf{T}_n\|^{\frac{2}{n}}_{W^{\gamma,\infty}(\Omega)}\right)$$

as $h \to 0_+$, where $\gamma \in (0,1]$. Because Φ_n^{-1}, the inverse function of Φ_n (which is uniquely defined on $[0,\infty)$ thanks to the fact that Φ_n is strictly monotonic increasing on $[0,\infty)$), is monotonic increasing, we have

$$\fint_\Omega |\mathbf{T}_{n,h} - \Pi_h\mathbf{T}_n|\,\mathrm{d}\mathbf{x} \leq \Phi_n^{-1}\left(Ch^{2\frac{\gamma}{n}}\left(\|\varepsilon(\mathbf{u}_n)\|^2_{W^{\frac{\gamma}{n},\infty}(\Omega)} + \|\mathbf{T}_n\|^{\frac{2}{n}}_{W^{\gamma,\infty}(\Omega)}\right)\right)$$

as $h \to 0_+$, where $\gamma \in (0,1]$ and $n \in \mathbb{N}$. Since $\Phi_n(s) \asymp s^2$ as $s \to 0_+$, it follows that $\Phi_n^{-1}(s) \asymp s^{\frac{1}{2}}$ as $s \to 0_+$, and therefore

$$\|\mathbf{T}_{n,h} - \Pi_h \mathbf{T}_n\|_{L_1(\Omega)} \leq C h^{\frac{\gamma}{n}} \left(\|\varepsilon(\mathbf{u}_n)\|_{W^{\frac{\gamma}{n},\infty}(\Omega)} + \|\mathbf{T}_n\|^{\frac{1}{n}}_{W^{\gamma,\infty}(\Omega)} \right) \tag{5.49}$$

as $h \to 0_+$, where $\gamma \in (0,1]$, $n \in \mathbb{N}$ and $C = C(d, \Lambda, n, \mathrm{K}_n, \mathrm{K}, \gamma, |\Omega|)$.

5.3 *Other elements that fit the theory*

We shall comment here on some alternative choices of finite element spaces to which our analysis applies. Let $\mathcal{Q}_h^r$ denote the finite element space on quadrilateral or hexahedral meshes for $d = 2$ or $d = 3$, respectively, consisting of (possibly discontinuous) mapped piecewise d-variate functions that are polynomials of degree r in each variable over each element in the subdivision. We consider the conforming finite element spaces

$$\mathbb{M}_{n,h} := \left(\mathcal{Q}_h^r\right)^{d\times d}_{\mathrm{sym}} \subset \mathbb{M}_n, \qquad \mathbb{X}_{n,h} := \left(\mathcal{Q}_h^r\right)^d \cap \mathbb{X}_n \subset \mathbb{X}_n \tag{5.50}$$

for the approximation of $\mathbf{T}_n$ and $\mathbf{u}_n$, respectively. Clearly, $\varepsilon(\mathbb{X}_{n,h}) \subset \mathbb{M}_{n,h}$ and the $L_2(\Omega)^{d\times d}$ orthogonal projector $\Pi_h : \mathbb{M}_n \mapsto \mathbb{M}_{n,h}$ is stable in the $L_p(\Omega)^{d\times d}$ norm for all $p \in [1,\infty]$.[1] Then Lemma 5.3 can be shown to hold by an identical argument; if in addition it is assumed that $\mathbf{T}_n \in L_\infty(\Omega)^{d\times d}_{\mathrm{sym}}$, then Lemma 5.4 and Theorem 5.5 also hold. We note that our proof of Lemma 5.4 in the special case of

$$\mathbb{M}_{n,h} = \left(\mathcal{P}_h^0\right)^{d\times d}_{\mathrm{sym}} \subset \mathbb{M}_n \qquad \text{and} \qquad \mathbb{X}_{n,h} = \left(\mathcal{P}_h^1\right)^d \cap \mathbb{X}_n \subset \mathbb{X}_n \tag{5.51}$$

did not require the additional assumption $\mathbf{T}_n \in L_\infty(\Omega)^{d\times d}_{\mathrm{sym}}$, thanks to the connection between the explicit formula for the projection onto piecewise constant functions and the Hardy–Littlewood maximal function.

5.4 *A simple quadrilateral/hexahedral element to which the theory does not apply*

The simplest extension to quadrilaterals or hexahedra of the spaces defined in (5.2) is of course

$$\mathbb{M}_{n,h} := \left(\mathcal{Q}_h^0\right)^{d\times d}_{\mathrm{sym}} \subset \mathbb{M}_n, \qquad \mathbb{X}_{n,h} := \left(\mathcal{Q}_h^1\right)^d \cap \mathbb{X}_n \subset \mathbb{X}_n \tag{5.52}$$

for the approximation of $\mathbf{T}_n$ and $\mathbf{u}_n$, respectively. Everything done previously applies to this pair of elements, *except the uniform discrete inf-sup condition*. Indeed, the proof of Lemma 5.1 does not carry over to this case because $\varepsilon(\mathbb{X}_{n,h})$ is not contained in $\mathbb{M}_{n,h}$.

---

[1] This stability result is a consequence of the stability in the $L_p(-1,1)$ norm of the $L_2(-1,1)$ orthogonal projection onto the space of univariate polynomials of degree r on the interval $(-1,1)$, for all $p \in [1,\infty]$, with a stability constant $C_{r,p} = C \cdot r^{\frac{1}{2}\left|1-\frac{2}{p}\right|}$; for $p = \infty$ see Gronwall (1913, eq. (29) on p. 230); for $p = 2$, $C_{r,2} = 1$ for all $r \geq 1$; for $p \in (2,\infty)$ the form of $C_{r,p}$ follows by function space interpolation and for $p \in [1,2)$ it follows from the result for $p = (2,\infty]$ by duality.

Let us look more closely at the greatest lower bound in (5.4), say β_h. First, for any given $\mathbf{v}_h$, the choice in each element K (which generalizes (5.5)),

$$\mathbf{T}_h = \frac{1}{|K|^n}\left(\int_K \varepsilon(\mathbf{v}_h)\,\mathrm{d}\mathbf{x}\right)\left|\int_K \varepsilon(\mathbf{v}_h)\,\mathrm{d}\mathbf{x}\right|^{n-1}, \tag{5.53}$$

shows that $\beta_h \geq 0$. The next lemma shows that on a structured mesh (i.e., a mesh with a Cartesian numbering) $\beta_h \neq 0$. To avoid excessive technicalities, it is stated for quadrilaterals, but it extends to structured hexahedral meshes.

PROPOSITION 5.6 Let $\mathcal{T}_h$ be a structured quadrilateral mesh. Then the greatest lower bound β_h in (5.4) is strictly positive.

Proof. We argue by contradiction. Suppose that $\beta_h = 0$. Then there is a displacement $\mathbf{v}_h$ in $\mathbb{X}_{n,h}$ such that

$$\sup_{\mathbf{S}_h \in \mathbb{M}_{n,h}} b(\mathbf{S}_h, \mathbf{v}_h) = 0.$$

In particular, $b(\mathbf{T}_h, \mathbf{v}_h) = 0$ for $\mathbf{T}_h$ defined by (5.53). This implies

$$\left|\int_K \varepsilon(\mathbf{v}_h)\,\mathrm{d}\mathbf{x}\right| = 0 \qquad \forall K \in \mathcal{T}_h. \tag{5.54}$$

Let us examine the consequences of (5.54) on specific elements K of the mesh. Let $\hat{K} = [0,1]^2$ be the reference square with vertices $\hat{\mathbf{a}}_1 = (0,0)$, $\hat{\mathbf{a}}_2 = (1,0)$, $\hat{\mathbf{a}}_3 = (1,1)$ and $\hat{\mathbf{a}}_4 = (0,1)$. Let $\mathbf{a}_i, 1 \leq i \leq 4$ denote the vertices of K and $\mathcal{F}_K$ the bilinear mapping from $\hat{K}$ onto K that maps $\hat{\mathbf{a}}_i$ to $\mathbf{a}_i$, $1 \leq i \leq 4$. Since the mesh is assumed to be nondegenerate, $\mathcal{F}_K$ is invertible and the functions of $\mathcal{Q}_h^1$ are the images by $\mathcal{F}_K^{-1}$ of the functions of $\hat{Q}^1$ defined on $\hat{K}$. Their derivatives are transformed as follows:

$$\frac{\partial v}{\partial x_1} \circ \mathcal{F}_K = \frac{1}{\mathcal{J}_K}\Big(\frac{\partial \hat{v}}{\partial \hat{x}_1}\big(a_2^4 - a_2^1 + \hat{x}_1\big(a_2^3 - a_2^2 - a_2^4 + a_2^1\big)\big) - \frac{\partial \hat{v}}{\partial \hat{x}_2}\big(a_2^2 - a_2^1 + \hat{x}_2\big(a_2^3 - a_2^2 - a_2^4 + a_2^1\big)\big)\Big),$$

$$\frac{\partial v}{\partial x_2} \circ \mathcal{F}_K = \frac{1}{\mathcal{J}_K}\Big(\frac{\partial \hat{v}}{\partial \hat{x}_2}\big(a_1^2 - a_1^1 + \hat{x}_2\big(a_1^3 - a_1^2 - a_1^4 + a_1^1\big)\big) - \frac{\partial \hat{v}}{\partial \hat{x}_1}\big(a_1^4 - a_1^1 + \hat{x}_1\big(a_1^3 - a_1^2 - a_1^4 + a_1^1\big)\big)\Big),$$

the subscript indicating the coordinate, and $\mathcal{J}_K$ the Jacobian of $\mathcal{F}_K$.

Now, let us start with a corner element; since the mesh is structured, all corner elements have at least two sides and three vertices on the boundary, say $\mathbf{a}_1$, $\mathbf{a}_2$ and $\mathbf{a}_4$. As $\mathbf{v}_h$ vanishes on $\partial\Omega$, this means that $\mathbf{v}_h(\mathbf{a}_1) = \mathbf{v}_h(\mathbf{a}_2) = \mathbf{v}_h(\mathbf{a}_4) = \mathbf{0}$ and thus

$$\begin{aligned}\left(\int_K \varepsilon(\mathbf{v}_h)\,\mathrm{d}\mathbf{x}\right) : \left(\int_K \varepsilon(\mathbf{v}_h)\,\mathrm{d}\mathbf{x}\right) &= \frac{1}{4}\Big[\big(\hat{v}_1^3\big)^2\big(a_2^4 - a_2^2\big)^2 + \big(\hat{v}_2^3\big)^2\big(a_1^4 - a_1^2\big)^2\Big] \\ &\quad + \frac{1}{4}\left[\frac{1}{4}\big(\hat{v}_2^3\big(a_2^4 - a_2^2\big) + \hat{v}_1^3\big(a_1^4 - a_1^2\big)\big)^2\right] = 0.\end{aligned}$$

As $|\mathbf{a}_4 - \mathbf{a}_2| > 0$, we easily derive from this expression that $\mathbf{v}_h(\mathbf{a}_3) = \mathbf{0}$, and hence $\mathbf{v}_h$ vanishes on K. This implies that $\mathbf{v}_h$ also vanishes at its neighbours adjacent to the boundary, and by progressing

element by element along the boundary, we have $\mathbf{v}_h = \mathbf{0}$ on all boundary elements. From here, the same argument gives $\mathbf{v}_h = \mathbf{0}$ on all elements of $\mathcal{T}_h$. □

The positivity of β_h implies that (5.4) holds with a positive constant for each h, but does not guarantee that the positive constant is uniformly bounded away from zero as h tends to zero. Let us give an example when β_h tends to zero, inspired by the checkerboard modes of the Stokes problem; see Girault & Raviart (1986). The idea is to construct a displacement $\mathbf{v}_h$ such that the integral average of $\varepsilon(\mathbf{v}_h)$ vanishes on a large number of elements, while $\varepsilon(\mathbf{v}_h)$ is nonzero there. Consider a square domain $\Omega = (0,1)^2$ divided into $(N+1)^2$ equal squares K_{ij}, $0 \le i,j \le N$, with mesh size $h = \frac{1}{N+1}$. Take $\mathbf{v}_h = \mathbf{0}$ on $\partial\Omega$ and define each component $\mathbf{v}_h$ by

$$v_h(\mathbf{x}_{ij}) = \begin{cases} 1 & \text{if } i+j \text{ is odd} \\ -1 & \text{if } i+j \text{ is even} \end{cases} \qquad \text{for } 1 \le i,j \le N.$$

It is easy to check that, in all interior elements K,

$$\int_K \varepsilon(\mathbf{v}_h)\,\mathrm{d}\mathbf{x} = \mathbf{0},$$

and in each boundary element K,

$$0 < c_1 h \le \left| \int_K \varepsilon(\mathbf{v}_h)\,\mathrm{d}\mathbf{x} \right| \le C_1 h,$$

where here and below all constants are independent of K and h. Let $\mathcal{T}_h^b$ denote the union of the boundary elements. Since the choice of $\mathbf{T}_h$ in all interior elements does not affect the value of $b(\mathbf{T}_h, \mathbf{v}_h)$, let us choose $\mathbf{T}_h = \mathbf{0}$ in these elements; this will minimize its norm there. On the boundary elements K, we choose $\mathbf{T}_h$ by (5.53); this gives

$$b(\mathbf{T}_h, \mathbf{v}_h) = \sum_{K \in \mathcal{T}_h^b} \frac{1}{|K|^n} \left| \int_K \varepsilon(\mathbf{v}_h)\,\mathrm{d}\mathbf{x} \right|^{n+1}$$

and

$$\|\mathbf{T}_h\|_{L_{1+\frac{1}{n}}(\Omega)} = \left(\sum_{K \in \mathcal{T}_h^b} \frac{1}{|K|^n} \left| \int_K \varepsilon(\mathbf{v}_h)\,\mathrm{d}\mathbf{x} \right|^{n+1} \right)^{\frac{n}{n+1}},$$

so that

$$\frac{b(\mathbf{T}_h, \mathbf{v}_h)}{\|\mathbf{T}_h\|_{L_{1+\frac{1}{n}}(\Omega)}} \le C_2 h^{-\frac{n}{n+1}}.$$

On the other hand, $\varepsilon(\mathbf{v}_h)$ does not vanish in the interior elements, and we have

$$\|\varepsilon(\mathbf{v}_h)\|_{L_{1+n}(\Omega)} \ge C_3 h^{-1}.$$

Hence, with this choice of $\mathbf{T}_h$,

$$\inf_{\mathbf{v}_h \in \mathbb{X}_{n,h}} \frac{b(\mathbf{T}_h, \mathbf{v}_h)}{\|\mathbf{T}_h\|_{L_{1+\frac{1}{n}}(\Omega)} \|\varepsilon(\mathbf{v}_h)\|_{L_{n+1}(\Omega)}} \leq C_4 h^{\frac{1}{n+1}}. \tag{5.55}$$

Of course, we have not proved that this choice of $\mathbf{T}_h$ realizes the supremum in (5.55). But since the number of interior elements, which do not contribute to the numerator of (5.55) but do contribute to the norm of $\mathbf{v}_h$, is much larger than the number of boundary elements (more precisely, this ratio is of the order of h^{-1}), no value of $\mathbf{T}_h$ can balance this ratio.

6. The case of smoother data

The regularization (3.1) is a particular case of

$$\varepsilon(\mathbf{u}) = \lambda(\mathrm{Tr}(\mathbf{T}))\mathrm{Tr}(\mathbf{T})\mathbf{I} + \mu(|\mathbf{T}^{\mathbf{d}}|)\mathbf{T}^{\mathbf{d}} + \frac{\mathrm{Tr}(\mathbf{T})\mathbf{I}}{n|\mathrm{Tr}(\mathbf{T})|^{1-\frac{1}{t}}} + \frac{\mathbf{T}^{\mathbf{d}}}{n|\mathbf{T}^{\mathbf{d}}|^{1-\frac{1}{t}}}, \tag{6.1}$$

$n \in \mathbb{N}$, $t \in \mathbb{R}_{>0}$, with $t = n$ in (3.1). When the data are smoother, as in Theorem 2.1 (d), the following simpler regularization is used in reference Bulíček *et al.* (2014):

$$\varepsilon(\mathbf{u}) = \lambda(\mathrm{Tr}(\mathbf{T}))\mathrm{Tr}(\mathbf{T})\mathbf{I} + \mu(|\mathbf{T}^{\mathbf{d}}|)\mathbf{T}^{\mathbf{d}} + \frac{1}{n}\mathbf{T}, \tag{6.2}$$

which corresponds to $t = 1$ (up to the factor $\frac{1}{d}$ multiplying $\mathbf{T}^{\mathbf{d}}$). The analysis developed in the previous sections applies to (1.1), (6.2) but is in fact much simpler. Indeed, let $(\mathbf{T}_{n,1}, \mathbf{u}_{n,1})$ denote a solution to (1.1), (6.2), i.e., $(\mathbf{T}_{n,1}, \mathbf{u}_{n,1}) \in \mathbb{M}_{n,1} \times \mathbb{X}_{n,1}$ satisfies

$$\begin{aligned} a_{n,1}(\mathbf{T}_{n,1}, \mathbf{S}) + c(\mathbf{T}_{n,1}; \mathbf{T}_{n,1}, \mathbf{S}) - b(\mathbf{S}, \mathbf{u}_{n,1}) &= 0 && \forall\, \mathbf{S} \in \mathbb{M}_{n,1}, \\ b(\mathbf{T}_{n,1}, \mathbf{v}) &= \int_\Omega \mathbf{f} \cdot \mathbf{v}\, \mathrm{d}\mathbf{x} && \forall\, \mathbf{v} \in \mathbb{X}_{n,1}, \end{aligned} \tag{6.3}$$

where

$$a_{n,1}(\mathbf{T}, \mathbf{S}) := \frac{1}{n} \int_\Omega \mathbf{T} : \mathbf{S}\, \mathrm{d}\mathbf{x}$$

and

$$\mathbb{M}_{n,1} := L_2(\Omega)^{d\times d}_{\mathrm{sym}}, \qquad \mathbb{X}_{n,1} := H^1_0(\Omega)^d.$$

The function $\mathbf{F}$ is used in deriving more regularity of the solution, but as far as the numerical scheme is concerned, we can simply proceed with the original data $\mathbf{f}$. Let us briefly sketch the analysis of (6.3). We define the mapping $\mathcal{A}_{n,1} : L_2(\Omega)^{d\times d}_{\mathrm{sym}} \to L_2(\Omega)^{d\times d}_{\mathrm{sym}}$ by

$$\mathcal{A}_{n,1}(\mathbf{S}) := \lambda(\mathrm{Tr}(\mathbf{S}))\mathrm{Tr}(\mathbf{S})\mathbf{I} + \mu(|\mathbf{S}^{\mathbf{d}}|)\mathbf{S}^{\mathbf{d}} + \frac{1}{n}\mathbf{S}, \tag{6.4}$$

and we easily prove as in Lemma 3.1 that $\mathcal{A}_{n,1}$ is bounded, continuous and coercive for all $n \in \mathbb{N}$. The inf-sup condition is satisfied, as in Lemma 4.1,

$$\inf_{\mathbf{v}\in\mathbb{X}_{n,1}} \sup_{\mathbf{S}\in\mathbb{M}_{n,1}} \frac{b(\mathbf{S},\mathbf{v})}{\|\mathbf{S}\|_{L_2(\Omega)}\|\varepsilon(\mathbf{v})\|_{L_2(\Omega)}} \geq 1. \tag{6.5}$$

The lifting $\mathbf{T}^{\mathbf{f}}_{n,1}$ is defined by the analogue of (4.2),

$$\int_\Omega \mathbf{T}^{\mathbf{f}}_{n,1} : \varepsilon(\mathbf{v})\,\mathrm{d}\mathbf{x} = \int_\Omega \mathbf{f}\cdot\mathbf{v}\,\mathrm{d}\mathbf{x} \qquad \forall\,\mathbf{v}\in\mathbb{X}_{n,1}, \tag{6.6}$$

and is bounded by

$$\|\mathbf{T}^{\mathbf{f}}_{n,1}\|_{L_2(\Omega)} \leq C_K\|\mathbf{f}\|_{L_2(\Omega)}, \tag{6.7}$$

where C_K is the constant of (1.11) with $p = 2$. In this context, the *a priori* estimates of Lemma 4.2 simplify to

$$\|\varepsilon(\mathbf{u}_{n,1})\|^2_{L_2(\Omega)} \leq \frac{4}{n^2}C_K^2\|\mathbf{f}\|^2_{L_2(\Omega)} + \frac{8}{n}C_1\kappa|\Omega| + 8C_2^2 d|\Omega|, \tag{6.8}$$

$$\frac{1}{n}\|\mathbf{T}_{n,1}\|^2_{L_2(\Omega)} + C_1\|\mathbf{T}_{n,1}\|_{L_1(\Omega)} \leq 2C_1\kappa|\Omega| + C_K\|\mathbf{f}\|_{L_2(\Omega)}\left(\frac{4}{n^2}C_K^2\|\mathbf{f}\|^2_{L_2(\Omega)} + \frac{8}{n}C_1\kappa|\Omega| + 8C_2^2 d|\Omega|\right)^{\frac{1}{2}}. \tag{6.9}$$

Thus, up to a subsequence, $\mathbf{u}_{n,1}$ converges weakly in $W^{1,2}_0(\Omega)^d$, and thanks to the results in Bulíček *et al.* (2014) (see also Theorem 2.1 (d) and Remark 2.2), the additional regularity $\mathbf{F} \in W^{2,2}(\Omega)^{d\times d}_{\mathrm{sym}}$ enables one to prove in particular that $\mathbf{T}_{n,1}$ is bounded in $W^{1,q}(\Omega_0)^{d\times d}_{\mathrm{sym}}$ for any $\Omega_0 \subset\subset \Omega$, with $q \in [1,2)$ when $d = 2$ and $q \in [1, \frac{3}{2}]$ when $d = 3$, and therefore, up to a subsequence, weakly converges to $\mathbf{T}$ in $W^{1,q}(\Omega_0)^{d\times d}_{\mathrm{sym}}$ for any $\Omega_0 \subset\subset \Omega$ for $q \in [1,2)$ when $d = 2$ and $q \in [1, \frac{3}{2}]$ when $d = 3$. Hence, by the Rellich–Kondrashov theorem, up to a subsequence, $\mathbf{T}_{n,1}$ tends to $\mathbf{T}$ strongly in $L_p(\Omega_0)^{d\times d}_{\mathrm{sym}}$ on any $\Omega_0 \subset\subset \Omega$ for all $p \in [1,\infty)$ when $d = 2$ and all $p \in [1, \frac{3}{2})$ when $d = 3$.

6.1 *Finite element discretization*

With the spaces $\mathbb{M}_{n,h}$ and $\mathbb{X}_{n,h}$ defined in (5.2) or (5.50), system (6.3) is discretized by the following: find $(\mathbf{T}_{n,1,h}, \mathbf{u}_{n,1,h})$ in $\mathbb{M}_{n,h} \times \mathbb{X}_{n,h}$ such that

$$\begin{aligned} a_{n,1}(\mathbf{T}_{n,1,h},\mathbf{S}_h) + c(\mathbf{T}_{n,1,h};\mathbf{T}_{n,1,h},\mathbf{S}_h) - b(\mathbf{S}_h,\mathbf{u}_{n,1,h}) &= 0 && \forall\,\mathbf{S}_h\in\mathbb{M}_{n,h},\\ b(\mathbf{T}_{n,1,h},\mathbf{v}_h) &= \int_\Omega \mathbf{f}\cdot\mathbf{v}_h\,\mathrm{d}\mathbf{x} && \forall\,\mathbf{v}_h\in\mathbb{X}_{n,h}. \end{aligned} \tag{6.10}$$

As previously, the constraint in the second part of (6.10) is lifted by means of the projection operator Π_h defined in (5.5), $\mathbf{T}^f_{n,1,h}$ is defined by (5.6),

$$\mathbf{T}^{\mathbf{f}}_{n,1,h} = \Pi_h\mathbf{T}^{\mathbf{f}}_{n,1}$$

and

$$\mathbf{T}^0_{n,1,h} := \mathbf{T}_{n,1,h} - \mathbf{T}^{\mathbf{f}}_{n,1,h}.$$

Existence and uniqueness of the discrete solution $(\mathbf{T}_{n,1,h}, \mathbf{u}_{n,1,h})$ is derived as in Lemma 5.2. Again, the *a priori* bounds (6.8) and (6.9) hold for $\mathbf{u}_{n,1,h}$ and $\mathbf{T}_{n,1,h}$. In fact, even without regularization, i.e., without the form $a_{n,1}(\cdot,\cdot)$, existence by Brouwer's fixed point follows and, if moreover (A3′) holds, uniqueness follows by a finite-dimensional argument. But we shall not pursue the no-regularization option, because as stated at the beginning of Section 5.2, we are then unable to show convergence.

The arguments of Lemma 5.3 under analogous assumptions show that, as $h \to 0_+$, for each n,

$$\mathbf{T}_{n,1,h} \rightharpoonup \mathbf{T}_{n,1} \qquad \text{weakly in } L_2(\Omega)^{d\times d}_{\mathrm{sym}}.$$

Let us sketch the proof of the strong convergence, which is much simpler than that of Lemma 5.4.

LEMMA 6.1 (Strong convergence). Assume that $\mathbf{f} \in L_2(\Omega)^d$, that the functions λ and μ satisfy assumptions (A1–A4), and let $(\mathbf{T}_{n,1}, \mathbf{u}_{n,1})$ denote the unique solution to the regularized problem (6.3), with $n \in \mathbb{N}$. Then for each fixed $n \in \mathbb{N}$, as $h \to 0_+$,

$$\mathbf{T}_{n,1,h} \to \mathbf{T}_{n,1} \quad \text{strongly in } \mathbb{M}_{n,1} = L_2(\Omega)^{d\times d}_{\mathrm{sym}} \quad \text{and} \quad \mathbf{u}_{n,1,h} \to \mathbf{u}_{n,1} \quad \text{strongly in } \mathbb{X}_{n,1} = W^{1,2}_0(\Omega)^d.$$

Proof. We retain the notation and the setting of the proof of Lemma 5.4. The discrepancy $\mathbf{T}^0_{n,1,h} - \Pi_h \mathbf{T}^0_{n,1}$ satisfies

$$\begin{aligned} &\frac{1}{n}\big\|\mathbf{T}^0_{n,1,h} - \Pi_h \mathbf{T}^0_{n,1}\big\|^2_{L_2(\Omega)} + C\int_\Omega \frac{\big|\mathbf{T}^0_{n,1,h} - \Pi_h \mathbf{T}^0_{n,1}\big|^2}{\big(\kappa + \big|\mathbf{T}^0_{n,1,h}\big| + \big|\Pi_h \mathbf{T}^0_{n,1}\big|\big)^{1+\alpha}}\,\mathrm{d}\mathbf{x} \\ &\qquad \le \int_\Omega \Big(\mathcal{A}_{n,1}(\mathbf{T}_{n,1,h}) - \mathcal{A}_{n,1}\big(\Pi_h \mathbf{T}^0_{n,1} + \mathbf{T}^f_{n,1,h}\big)\Big) : \big(\mathbf{T}^0_{n,1,h} - \Pi_h \mathbf{T}^0_{n,1}\big)\,\mathrm{d}\mathbf{x}, \end{aligned} \tag{6.11}$$

where C is the constant in (1.5). As $\mathbf{T}^0_{n,1,h} - \Pi_h \mathbf{T}^0_{n,1} \in \mathbb{V}_{n,h}$, (6.11) reduces to

$$\begin{aligned} &\frac{1}{n}\big\|\mathbf{T}^0_{n,1,h} - \Pi_h \mathbf{T}^0_{n,1}\big\|^2_{L_2(\Omega)} + C\int_\Omega \frac{\big|\mathbf{T}^0_{n,1,h} - \Pi_h \mathbf{T}^0_{n,1}\big|^2}{\big(\kappa + \big|\mathbf{T}^0_{n,1,h}\big| + \big|\Pi_h \mathbf{T}^0_{n,1}\big|\big)^{1+\alpha}}\,\mathrm{d}\mathbf{x} \\ &\qquad \le -\int_\Omega \mathcal{A}_{n,1}\big(\Pi_h \mathbf{T}^0_{n,1} + \mathbf{T}^f_{n,1,h}\big) : \big(\mathbf{T}^0_{n,1,h} - \Pi_h \mathbf{T}^0_{n,1}\big)\,\mathrm{d}\mathbf{x}. \end{aligned} \tag{6.12}$$

Then the weak convergence of $\mathbf{T}^0_{n,1,h} - \Pi_h \mathbf{T}^0_{n,1}$ to zero, the strong convergence of $\Pi_h \mathbf{T}^0_{n,1} + \mathbf{T}^f_{n,1,h}$ both in $\mathbb{M}_{n,1}$ as $h \to 0_+$ and the continuity of the mapping $\mathcal{A}_{n,1} : \mathbb{M}_{n,1} \to \mathbb{M}_{n,1}$ yield

$$-\int_\Omega \mathcal{A}_{n,1}\big(\Pi_h \mathbf{T}^0_{n,1} + \mathbf{T}^f_{n,1,h}\big) : \big(\mathbf{T}^0_{n,1,h} - \Pi_h \mathbf{T}^0_{n,1}\big)\,\mathrm{d}\mathbf{x} \to 0 \qquad \text{as } h \to 0_+,$$

whence, returning to (6.12),

$$\frac{1}{n}\left\|\mathbf{T}^0_{n,1,h} - \Pi_h \mathbf{T}^0_{n,1}\right\|^2_{L_2(\Omega)} \to 0 \qquad \text{as } h \to 0_+,$$

and the asserted strong convergence of $\mathbf{T}_{n,1,h}$ to $\mathbf{T}_{n,1}$ in $\mathbb{M}_{n,1} = L_2(\Omega)^{d\times d}_{\text{sym}}$, as $h \to 0_+$, follows for any $n \geq 1$.

For the strong convergence of $\mathbf{u}_{n,1,h}$, we use again the discrete inf-sup property (5.4) to define $\mathbf{R}_h \in \mathbb{V}^{\perp}_{n,h}$ satisfying

$$\int_\Omega \varepsilon\left(\mathbf{u}_{n,1,h} - \Pi^{\text{sz}}_h \mathbf{u}_{n,1}\right) : \varepsilon(\mathbf{v}_h)\,\mathrm{d}\mathbf{x} = \int_\Omega \mathbf{R}_h : \varepsilon(\mathbf{v}_h)\,\mathrm{d}\mathbf{x} \qquad \forall\, \mathbf{v}_h \in \mathbb{X}_{n,h},$$

where Π^{sz}_h is the Scott–Zhang projector onto $\mathbb{X}_{n,h}$; see Scott & Zhang (1990). In particular, we have

$$\|\mathbf{R}_h\|_{L_2(\Omega)} \leq \|\varepsilon(\mathbf{u}_{n,1,h} - \Pi^{\text{sz}}_h \mathbf{u}_{n,1})\|_{L_2(\Omega)}. \tag{6.13}$$

For $\mathbf{v}_h = \mathbf{u}_{n,1,h} - \Pi^{\text{sz}}_h \mathbf{u}_{n,1}$ we then get

$$\begin{aligned}\left\|\varepsilon\left(\mathbf{u}_{n,1,h} - \Pi^{\text{sz}}_h \mathbf{u}_{n,1}\right)\right\|^2_{L_2(\Omega)} &= \int_\Omega \mathbf{R}_h : \varepsilon(\mathbf{u}_{n,1,h} - \mathbf{u}_{n,1})\,\mathrm{d}\mathbf{x} + \int_\Omega \mathbf{R}_h : \varepsilon\left(\mathbf{u}_{n,1} - \Pi^{\text{sz}}_h \mathbf{u}_{n,1}\right)\mathrm{d}\mathbf{x}\\ &= \int_\Omega (\mathcal{A}_{n,1}(\mathbf{T}_{n,1,h}) - \mathcal{A}_{n,1}(\mathbf{T}_{n,1})) : \mathbf{R}_h\,\mathrm{d}\mathbf{x} + \int_\Omega \mathbf{R}_h : \varepsilon\left(\mathbf{u}_{n,1} - \Pi^{\text{sz}}_h \mathbf{u}_{n,1}\right)\mathrm{d}\mathbf{x},\end{aligned}$$

where we have also used relations (6.3) and (6.10) to obtain the second equality. We now argue that both terms on the right-hand side of the above equality vanish as $h \to 0_+$. To see this, it suffices to recall the uniform bound (6.13) on $\mathbf{R}_h$; hence, the strong convergence results $\Pi^{\text{sz}}_h \mathbf{u}_{n,1} \to \mathbf{u}_{n,1}$ in $\mathbb{X}_{n,1}$ and $\mathbf{T}_{n,1,h} \to \mathbf{T}_{n,1}$ in $\mathbb{M}_{n,1}$, as $h \to 0_+$, together with the continuity of $\mathcal{A}_{n,1}$ guaranteed by Lemma 3.1, imply the stated claim. Thanks to Korn's inequality (1.10),

$$\left\|\nabla\left(\mathbf{u}_{n,1,h} - \Pi^{\text{sz}}_h \mathbf{u}_{n,1}\right)\right\|_{L_2(\Omega)} \leq \mathcal{K}\left\|\varepsilon\left(\mathbf{u}_{n,1,h} - \Pi^{\text{sz}}_h \mathbf{u}_{n,1}\right)\right\|_{L_2(\Omega)} \to 0 \qquad \text{as } h \to 0_+,$$

and therefore $\mathbf{u}_{n,1,h} \to \mathbf{u}_{n,1}$ in $\mathbb{X}_{n,1}$. □

Thus, when λ satisfies (A3′), we have again, for any $\Omega_0 \subset\subset \Omega$,

$$\begin{aligned}\lim_{n\to\infty}\lim_{h\to 0_+} \|\mathbf{T}_{n,1,h} - \mathbf{T}\|_{L_1(\Omega_0)} &= 0,\\ \lim_{n\to\infty}\lim_{h\to 0_+} \|\mathbf{u}_{n,1,h} - \mathbf{u}\|_{C(\overline{\Omega})} &= 0,\\ \lim_{n\to\infty}\lim_{h\to 0_+} \|\varepsilon(\mathbf{u}_{n,1,h}) - \varepsilon(\mathbf{u})\|_{L_2(\Omega_0)} &= 0.\end{aligned}$$

As in the preceding section, an error inequality can be established when the functions $\lambda(s)s$ and $\mu(s)s$ are Lipschitz continuous, but again the situation is much simpler.

THEOREM 6.2 In addition to the assumptions of Lemma 6.1, suppose that the real-valued functions $s \in \mathbb{R} \mapsto \lambda(s)s$ and $s \in \mathbb{R}_{\geq 0} \mapsto \mu(s)s$ are Lipschitz continuous, i.e., that there exists a positive constant

Λ such that

$$|\lambda(s)s - \lambda(r)r| \le \Lambda|r-s| \qquad \forall\, r,s \in \mathbb{R}, \qquad |\mu(s)s - \mu(r)r| \le \Lambda|r-s| \qquad \forall\, r,s \in \mathbb{R}_{\ge 0}. \quad (6.14)$$

Then, the following error inequality holds:

$$\frac{1}{n}\|\mathbf{T}_{n,1,h} - \mathbf{T}_{n,1}\|_{L_2(\Omega)} \le \inf_{\mathbf{v}_h \in \mathbb{X}_h} \|\varepsilon(\mathbf{v}_h - \mathbf{u}_{n,1})\|_{L_2(\Omega)} + 2\left(\frac{1}{n} + \Lambda\right)\|\mathbf{T}_{n,1} - \Pi_h(\mathbf{T}_{n,1})\|_{L_2(\Omega)}. \quad (6.15)$$

Proof. As in the proof of Theorem 5.5, from relations (6.3) and (6.10) we infer that on the one hand,

$$\int_\Omega \varepsilon(\mathbf{v}_h) : \mathbf{S}_h \,\mathrm{d}\mathbf{x} = 0,$$

and on the other hand, for any $\mathbf{v}_h$ in $\mathbb{X}_{n,h}$,

$$\begin{aligned}\frac{1}{n}\|\mathbf{T}_{n,1,h} - \Pi_h\mathbf{T}_{n,1}\|^2_{L_2(\Omega)} + C\int_\Omega \frac{|\mathbf{T}_{n,1,h} - \Pi_h\mathbf{T}_{n,1}|^2}{(\kappa + |\mathbf{T}_{n,1,h}| + |\Pi_h\mathbf{T}_{n,1}|)^{1+\alpha}}\,\mathrm{d}\mathbf{x}\\ \le \left(\|\varepsilon(\mathbf{v}_h - \mathbf{u}_{n,1})\|_{L_2(\Omega)} + \frac{1}{n}\|\mathbf{T}_{n,1} - \Pi_h\mathbf{T}_{n,1}\|_{L_2(\Omega)}\right)\|\mathbf{T}_{n,1,h} - \Pi_h\mathbf{T}_{n,1}\|_{L_2(\Omega)}\\ + \int_\Omega \big(\mathcal{A}_{n,1}(\mathbf{T}_{n,1}) - \mathcal{A}_{n,1}(\Pi_h\mathbf{T}_{n,1})\big) : (\mathbf{T}_{n,1,h} - \Pi_h\mathbf{T}_{n,1})\,\mathrm{d}\mathbf{x},\end{aligned} \quad (6.16)$$

where C is the constant in (1.5). The Lipschitz property (6.14) implies

$$|\mathcal{A}_{n,1}(\mathbf{T}_{n,1}) - \mathcal{A}_{n,1}(\Pi_h\mathbf{T}_{n,1})| \le \frac{1}{n}|\mathbf{T}_{n,1} - \Pi_h\mathbf{T}_{n,1}| + 2\Lambda|\Pi_h\mathbf{T}_{n,1} - \mathbf{T}_{n,1}|,$$

so that

$$\frac{1}{n}\|\mathbf{T}_{n,1,h} - \Pi_h\mathbf{T}_{n,1}\|_{L_2(\Omega)} \le \left(\|\varepsilon(\mathbf{v}_h - \mathbf{u}_{n,1})\|_{L_2(\Omega)} + 2\left(\frac{1}{n} + \Lambda\right)\|\mathbf{T}_{n,1} - \Pi_h\mathbf{T}_{n,1}\|_{L_2(\Omega)}\right), \quad (6.17)$$

which yields (6.15). □

Under the above assumptions, convergence rates can be derived provided that $\mathbf{T}_{n,1} \in W^{1,q}(\Omega)^{d\times d}_{\mathrm{sym}}$ with $q > \frac{2d}{2+d}$ (ensuring that $W^{1,q}(\Omega)^{d\times d} \hookrightarrow L_2(\Omega)^{d\times d}$) and $\mathbf{u}_{n,1} \in W^{1+t,2}(\Omega)^d$, $t > 0$ (ensuring that $W^{1,1+t}(\Omega)^d \hookrightarrow W^{1,2}(\Omega)^d$). Rates of convergence for $\|\nabla(\mathbf{u}_{n,1} - \mathbf{u}_{n,1,h})\|_{L_2(\Omega)}$ are obtained using the inf-sup properties and interpolation theory again.

7. Decoupled iterative algorithm

The convergent iterative algorithm proposed in this section for the solution of the discrete problem (6.10) is designed to dissociate the computation of the nonlinearity from that of the elastic constraint. We have also applied it numerically to (5.3) in Section 8 but proving its convergence in that case is still an open problem.

The algorithm, which belongs to the class of alternating direction methods, proceeds in two steps. In both steps an artificial divided difference, analogous to a discrete time derivative, is added to enhance the stability of the algorithm. The first half-step involves the monotone nonlinearity while, in the case of (6.10), the second half-step solves for the elastic part from a system of linear algebraic equations whose matrix is the mass-matrix (Gram matrix) generated by the basis functions of the finite element space $\mathbb{X}_{n,h}$. In the case (5.3), this second system is nonlinear. But in both cases, our choice of the finite element space $\mathbb{M}_{n,h}$, consisting of piecewise constant approximations for the stress tensor $\mathbf{T}_{n,1}$ or $\mathbf{T}_n$, allows us to deal with the monotone nonlinearity involved in the first half-step in an efficient way by solving an algebraic system with $d(d+1)/2$ unknowns independently on each element K in the subdivision $\mathcal{T}_h$ of the computational domain Ω. Let us describe the algorithm applied to (6.1).

The initialization consists of finding $(\mathbf{T}_h^{(0)}, \mathbf{u}_h^{(0)}) \in \mathbb{M}_{n,h} \times \mathbb{X}_{n,h}$ satisfying

$$\int_\Omega \varepsilon(\mathbf{v}_h) : \mathbf{T}_h^{(0)}\, d\mathbf{x} = \int_\Omega \mathbf{f}\cdot\mathbf{v}_h\, d\mathbf{x}, \qquad \forall\, \mathbf{v}_h \in \mathbb{X}_{n,h}$$
$$\int_\Omega \mathbf{T}_h^{(0)} : \mathbf{S}_h\, d\mathbf{x} = \int_\Omega \varepsilon\left(\mathbf{u}_h^{(0)}\right) : \mathbf{S}_h\, d\mathbf{x} \qquad \forall\, \mathbf{S}_h \in \mathbb{M}_{n,h}.$$

Let $\tau > 0$. Given $(\mathbf{T}_h^{(k)}, \mathbf{u}_h^{(k)})$ in $\mathbb{M}_{n,h} \times \mathbb{X}_{n,h}$ for a nonnegative integer k, the algorithm proceeds in the following two steps.

Step 1. Find $\mathbf{T}_h^{(k+\frac{1}{2})}$ in $\mathbb{M}_{n,h}$ such that, for all $\mathbf{S}_h \in \mathbb{M}_{n,h}$,

$$\begin{aligned}
&\frac{1}{\tau}\int_\Omega \left(\mathbf{T}_h^{(k+\frac{1}{2})} - \mathbf{T}_h^{(k)}\right) : \mathbf{S}_h\, d\mathbf{x} \\
&\quad + \int_\Omega \left(\lambda\left(\mathrm{Tr}\left(\mathbf{T}_h^{(k+\frac{1}{2})}\right)\right)\mathrm{Tr}\left(\mathbf{T}_h^{(k+\frac{1}{2})}\right)\mathrm{Tr}(\mathbf{S}_h) + \mu\left(\left|\left(\mathbf{T}_h^{(k+\frac{1}{2})}\right)^{\mathbf{d}}\right|\right)\left(\mathbf{T}_h^{(k+\frac{1}{2})}\right)^{\mathbf{d}} : \mathbf{S}_h\right) d\mathbf{x} \\
&\quad = \int_\Omega \varepsilon(\mathbf{u}_h^{(k)}) : \mathbf{S}_h\, d\mathbf{x} - \int_\Omega \left(\frac{\mathrm{Tr}\left(\mathbf{T}_h^{(k)}\right)\mathbf{I}}{n\left|\mathrm{Tr}\left(\mathbf{T}_h^{(k)}\right)\right|^{1-\frac{1}{t}}} + \frac{\left(\mathbf{T}_h^{(k)}\right)^{\mathbf{d}}}{n\left|\left(\mathbf{T}_h^{(k)}\right)^{\mathbf{d}}\right|^{1-\frac{1}{t}}}\right) : \mathbf{S}_h\, d\mathbf{x}.
\end{aligned}$$

As has already been mentioned, because $\mathbf{T}_h^{(k+\frac{1}{2})}$ is piecewise constant, the above system reduces to decoupled algebraic systems of $d(d+1)/2$ unknowns each, in every element in the subdivision of the computational domain.

Step 2. Find $\mathbf{T}_h^{(k+1)} \in \mathbb{M}_{n,h}$ and $\mathbf{u}_h^{(k+1)} \in \mathbb{X}_{n,h}$ such that

$$\begin{aligned}
&\frac{1}{\tau}\int_\Omega \left(\mathbf{T}_h^{(k+1)} - \mathbf{T}_h^{(k+\frac{1}{2})}\right) : \mathbf{S}_h\, d\mathbf{x} + \int_\Omega \left(\frac{\mathrm{Tr}\left(\mathbf{T}_h^{(k+1)}\right)\mathbf{I}}{n|\mathrm{Tr}\left(\mathbf{T}_h^{(k+1)}\right)|^{1-\frac{1}{t}}} + \frac{\left(\mathbf{T}_h^{(k+1)}\right)^{\mathbf{d}}}{n|\left(\mathbf{T}_h^{(k+1)}\right)^{\mathbf{d}}|^{1-\frac{1}{t}}}\right) : \mathbf{S}_h\, d\mathbf{x} \\
&\quad = \int_\Omega \varepsilon\left(\mathbf{u}_h^{(k+1)}\right) : \mathbf{S}_h\, d\mathbf{x} - \int_\Omega \left(\lambda\left(\mathrm{Tr}\left(\mathbf{T}_h^{(k+\frac{1}{2})}\right)\right)\mathrm{Tr}\left(\mathbf{T}_h^{(k+\frac{1}{2})}\right)\mathrm{Tr}(\mathbf{S}_h)\right. \\
&\qquad \left. + \mu\left(\left|\left(\mathbf{T}_h^{(k+\frac{1}{2})}\right)^{\mathbf{d}}\right|\right)\left(\mathbf{T}_h^{(k+\frac{1}{2})}\right)^{\mathbf{d}} : \mathbf{S}_h\right) d\mathbf{x}
\end{aligned}$$

and such that, for all $\mathbf{v}_h \in \mathbb{X}_{n,h}$,

$$\int_\Omega \varepsilon(\mathbf{v}_h) : \mathbf{T}_h^{(k+1)} \, \mathrm{d}\mathbf{x} = \int_\Omega \mathbf{f} \cdot \mathbf{v}_h \, \mathrm{d}\mathbf{x}.$$

When $t = 1$, the initialization is unchanged and the two steps simplify as follows:

Step 1. Find $\mathbf{T}_h^{(k+\frac{1}{2})}$ in $\mathbb{M}_{n,h}$ such that, for all $\mathbf{S}_h \in \mathbb{M}_{n,h}$,

$$\begin{aligned}
&\frac{1}{\tau} \int_\Omega \left(\mathbf{T}_h^{(k+\frac{1}{2})} - \mathbf{T}_h^{(k)} \right) : \mathbf{S}_h \, \mathrm{d}\mathbf{x} \\
&\quad + \int_\Omega \left(\lambda \left(\mathrm{Tr} \left(\mathbf{T}_h^{(k+\frac{1}{2})} \right) \right) \mathrm{Tr} \left(\mathbf{T}_h^{(k+\frac{1}{2})} \right) \mathrm{Tr}(\mathbf{S}_h) + \mu \left(\left| \left(\mathbf{T}_h^{(k+\frac{1}{2})} \right)^{\mathbf{d}} \right| \right) \left(\mathbf{T}_h^{(k+\frac{1}{2})} \right)^{\mathbf{d}} : \mathbf{S}_h \right) \mathrm{d}\mathbf{x} \\
&\quad = \int_\Omega \varepsilon \left(\mathbf{u}_h^{(k)} \right) : \mathbf{S}_h \, \mathrm{d}\mathbf{x} - \frac{1}{n} \int_\Omega \mathbf{T}_h^{(k)} : \mathbf{S}_h \, \mathrm{d}\mathbf{x}.
\end{aligned}$$

Step 2. Find $\mathbf{T}_h^{(k+1)} \in \mathbb{M}_{n,h}$ and $\mathbf{u}_h^{(k+1)} \in \mathbb{X}_{n,h}$ such that

$$\begin{aligned}
&\frac{1}{\tau} \int_\Omega \left(\mathbf{T}_h^{(k+1)} - \mathbf{T}_h^{(k+\frac{1}{2})} \right) : \mathbf{S}_h \, \mathrm{d}\mathbf{x} + \frac{1}{n} \int_\Omega \mathbf{T}_h^{(k+1)} : \mathbf{S}_h \, \mathrm{d}\mathbf{x} \\
&\quad = \int_\Omega \varepsilon \left(\mathbf{u}_h^{(k+1)} \right) : \mathbf{S}_h \, \mathrm{d}\mathbf{x} - \int_\Omega \left(\lambda \left(\mathrm{Tr} \left(\mathbf{T}_h^{(k+\frac{1}{2})} \right) \right) \mathrm{Tr} \left(\mathbf{T}_h^{(k+\frac{1}{2})} \right) \mathrm{Tr}(\mathbf{S}_h) \right. \\
&\qquad \left. + \mu \left(\left| \left(\mathbf{T}_h^{(k+\frac{1}{2})} \right)^{\mathbf{d}} \right| \right) \left(\mathbf{T}_h^{(k+\frac{1}{2})} \right)^{\mathbf{d}} : \mathbf{S}_h \right) \mathrm{d}\mathbf{x}
\end{aligned}$$

and such that, for all $\mathbf{v}_h \in \mathbb{X}_{n,h}$,

$$\int_\Omega \varepsilon(\mathbf{v}_h) : \mathbf{T}_h^{(k+1)} \, \mathrm{d}\mathbf{x} = \int_\Omega \mathbf{f} \cdot \mathbf{v}_h \, \mathrm{d}\mathbf{x}.$$

Following the general theory of Lions & Mercier (1979) we now prove that the iterative algorithm for $t = 1$ converges to the solution of the decoupled system.

THEOREM 7.1 (Convergence of the iterative decoupled algorithm). Assume that λ and μ satisfy (A1–A4) and that $n \geq 1$. Let $\mathbf{T}_{n,1,h} \in \mathbb{M}_{n,h}$ be the first component of the solution of (6.10) and let $\mathbf{T}_h^{(k)} \in \mathbb{M}_{n,h}$, $k = 1, 2, \ldots$ be successive iterates computed by the iterative algorithm, with $\tau > 0$. We then have

$$\lim_{k \to \infty} \left\| \mathbf{T}_h^{(k)} - \mathbf{T}_{n,1,h} \right\|_{L_2(\Omega)} = 0.$$

Proof. The nonlinear part of the system is represented by the operator $\mathfrak{A}_h : \mathbb{M}_{n,h} \to \mathbb{M}_{n,h}$ defined by $\mathfrak{A}_h \mathbf{S}_h = \mathbf{A}_h$, where for all $\mathbf{R}_h \in \mathbb{M}_{n,h}$,

$$\int_\Omega \mathbf{A}_h : \mathbf{R}_h \, \mathrm{d}\mathbf{x} = \int_\Omega \Big(\lambda(\mathrm{Tr}(\mathbf{S}_h))\mathrm{Tr}(\mathbf{S}_h)\mathrm{Tr}(\mathbf{R}_h) + \mu(|(\mathbf{S}_h^{\mathbf{d}}|)\mathbf{S}_h^{\mathbf{d}} : \mathbf{R}_h \Big) \, \mathrm{d}\mathbf{x},$$

and the linear part, excluding the artificial time derivative, is represented by the function

$$\mathbf{B}_h^{(k)} := \frac{1}{n}\mathbf{T}_h^{(k)} - \varepsilon\big(\mathbf{u}_h^{(k)}\big).$$

With this notation, the first step of the iterative algorithm reads

$$(I + \tau\mathfrak{A}_h)\mathbf{T}_h^{\left(k+\frac{1}{2}\right)} = \mathbf{T}_h^{(k)} - \tau\mathbf{B}_h^{(k)}$$

or, equivalently,

$$\mathbf{T}_h^{\left(k+\frac{1}{2}\right)} = (I + \tau\mathfrak{A}_h)^{-1}\big(\mathbf{T}_h^{(k)} - \tau\mathbf{B}_h^{(k)}\big).$$

It is convenient to introduce the two auxiliary tensors

$$\boldsymbol{\Lambda}_h^{(k)} := \mathbf{T}_h^{(k)} + \tau\left(\frac{1}{n}\mathbf{T}_h^{(k)} - \varepsilon\big(\mathbf{u}_h^{(k)}\big)\right) = \mathbf{T}_h^{(k)} + \tau\mathbf{B}_h^{(k)} \tag{7.1}$$

and

$$\boldsymbol{\Theta}_h^{(k)} := 2\mathbf{T}_h^{(k)} - \boldsymbol{\Lambda}_h^{(k)},$$

whereby

$$\mathbf{T}_h^{(k)} = \frac{1}{2}\big(\boldsymbol{\Theta}_h^{(k)} + \boldsymbol{\Lambda}_h^{(k)}\big).$$

We shall see that the convergence of $\mathbf{T}_h^{(k)}$ will result from that of $\boldsymbol{\Lambda}_h^{(k)}$ and $\boldsymbol{\Theta}_h^{(k)}$. With these tensors the second step of the iterative algorithm reads

$$\boldsymbol{\Lambda}_h^{(k+1)} = (I - \tau\mathfrak{A}_h)\mathbf{T}_h^{\left(k+\frac{1}{2}\right)} = (I - \tau\mathfrak{A}_h)(I + \tau\mathfrak{A}_h)^{-1}\big(\mathbf{T}_h^{(k)} - \tau\mathbf{B}_h^{(k)}\big).$$

Notice that, from (7.1), $\mathbf{B}_h^{(k)} = \frac{1}{2\tau}(\boldsymbol{\Lambda}_h^{(k)} - \boldsymbol{\Theta}_h^{(k)})$, and we define $\mathbf{C}_h^{(k)} := \frac{1}{2\tau}(\boldsymbol{\Theta}_h^{(k)} - \boldsymbol{\Lambda}_h^{(k+1)})$. In addition, we note for later that

$$(I + \tau\mathfrak{A}_h)^{-1}\boldsymbol{\Theta}_h^{(k)} = \frac{1}{2}\big(\boldsymbol{\Lambda}_h^{(k+1)} + \boldsymbol{\Theta}_h^{(k)}\big),$$

which implies

$$\mathfrak{A}_h\left(\frac{\boldsymbol{\Lambda}_h^{(k+1)} + \boldsymbol{\Theta}_h^{(k)}}{2}\right) = \frac{1}{2\tau}\big(\boldsymbol{\Theta}_h^{(k)} - \boldsymbol{\Lambda}_h^{(k+1)}\big) = \mathbf{C}_h^{(k)}.$$

We also define the analogous quantities

$$\mathbf{B}_h := \frac{1}{n}\mathbf{T}_{n,1,h} - \varepsilon(\mathbf{u}_{n,1,h}), \qquad \boldsymbol{\Lambda}_h := \mathbf{T}_{n,1,h} + \tau\mathbf{B}_h, \qquad \mathbf{C}_h := \mathfrak{A}_h\mathbf{T}_{n,1,h}, \qquad \boldsymbol{\Theta}_h = \mathbf{T}_{n,1,h} + \tau\mathbf{C}_h.$$

With this notation, the first relation in (5.3) reads

$$\mathbf{C}_h + \mathbf{B}_h = \mathfrak{A}_h\mathbf{T}_{n,1,h} + \frac{1}{n}\mathbf{T}_{n,1,h} - \varepsilon(\mathbf{u}_{n,1,h}) = \mathbf{0},$$

and so

$$\boldsymbol{\Lambda}_h + \boldsymbol{\Theta}_h = 2\mathbf{T}_{n,1,h} + \tau(\mathbf{B}_h + \mathbf{C}_h) = 2\mathbf{T}_{n,1,h},$$

which in turn implies

$$\mathbf{C}_h = \frac{1}{\tau}(\boldsymbol{\Theta}_h - \mathbf{T}_{n,1,h}) = \frac{1}{2\tau}(\boldsymbol{\Theta}_h - \boldsymbol{\Lambda}_h).$$

Similarly, for $\mathbf{B}_h$ we have the decomposition

$$\mathbf{B}_h = \frac{1}{2\tau}(\boldsymbol{\Lambda}_h - \boldsymbol{\Theta}_h).$$

We can now express the discrepancy between $\mathbf{T}_{n,1,h}$ and $\mathbf{T}_h^{(k)}$ as

$$\begin{aligned}\frac{1}{n}\big\|\mathbf{T}_h^{(k)} - \mathbf{T}_{n,1,h}\big\|_{L_2(\Omega)}^2 &= \frac{1}{n}\int_\Omega \big(\mathbf{T}_h^{(k)} - \mathbf{T}_{n,1,h}\big) : \big(\mathbf{T}_h^{(k)} - \mathbf{T}_{n,1,h}\big)\,\mathrm{d}\mathbf{x}\\ &= \int_\Omega \big(\mathbf{B}_h^{(k)} - \mathbf{B}_h\big) : \big(\mathbf{T}_h^{(k)} - \mathbf{T}_{n,1,h}\big)\,\mathrm{d}\mathbf{x} + \int_\Omega \varepsilon\big(\mathbf{u}_h^{(k)} - \mathbf{u}_{n,1,h}\big) : \big(\mathbf{T}_h^{(k)} - \mathbf{T}_{n,1,h}\big)\,\mathrm{d}\mathbf{x}.\end{aligned}$$

Because, for all $\mathbf{v}_h \in \mathbb{X}_{n,h}$,

$$\int_\Omega \mathbf{T}_h^{(k)} : \varepsilon(\mathbf{v}_h)\,\mathrm{d}\mathbf{x} = \int_\Omega \mathbf{f}\cdot\mathbf{v}_h = \int_\Omega \mathbf{T}_{n,1,h} : \varepsilon(\mathbf{v}_h)\,\mathrm{d}\mathbf{x},$$

we deduce that $\mathbf{T}_h^{(k)} - \mathbf{T}_{n,1,h} \in \mathbb{V}_{n,h}$, and therefore

$$\frac{1}{n}\big\|\mathbf{T}_h^{(k)} - \mathbf{T}_{n,1,h}\big\|_{L_2(\Omega)}^2 = \int_\Omega \big(\mathbf{B}_h^{(k)} - \mathbf{B}_h\big) : \big(\mathbf{T}_h^{(k)} - \mathbf{T}_{n,1,h}\big)\,\mathrm{d}\mathbf{x}.$$

The relations

$$\mathbf{B}_h^{(k)} - \mathbf{B}_h = \frac{1}{2\tau}\Big(\boldsymbol{\Lambda}_h^{(k)} - \boldsymbol{\Lambda}_h - \big(\boldsymbol{\Theta}_h^{(k)} - \boldsymbol{\Theta}_h\big)\Big)$$

and

$$\mathbf{T}_h^{(k)} - \mathbf{T}_{n,1,h} = \frac{1}{2}\Big(\boldsymbol{\Lambda}_h^{(k)} - \boldsymbol{\Lambda}_h + \big(\boldsymbol{\Theta}_h^{(k)} - \boldsymbol{\Theta}_h\big)\Big)$$

further lead to

$$\frac{1}{n}\left\|\mathbf{T}_h^{(k)}-\mathbf{T}_{n,1,h}\right\|_{L_2(\Omega)}^2=\frac{1}{4\tau}\left(\|\boldsymbol{\Lambda}_h^{(k)}-\boldsymbol{\Lambda}_h\|_{L_2(\Omega)}^2-\|\boldsymbol{\Theta}_h^{(k)}-\boldsymbol{\Theta}_h\|_{L_2(\Omega)}^2\right). \tag{7.2}$$

This, of course, implies

$$\|\boldsymbol{\Theta}_h^{(k)}-\boldsymbol{\Theta}_h\|_{L_2(\Omega)}\le\|\boldsymbol{\Lambda}_h^{(k)}-\boldsymbol{\Lambda}_h\|_{L_2(\Omega)}. \tag{7.3}$$

In addition, we have

$$\begin{aligned}&\int_\Omega\left(\mathbf{C}_h^{(k)}-\mathbf{C}_h\right):\left(\frac{\boldsymbol{\Lambda}_h^{(k+1)}+\boldsymbol{\Theta}_h^{(k)}}{2}-\mathbf{T}_{n,1,h}\right)\mathrm{d}\mathbf{x}\\&\quad=\int_\Omega\left(\mathfrak{A}_h\left(\frac{\boldsymbol{\Lambda}_h^{(k+1)}+\boldsymbol{\Theta}_h^{(k)}}{2}\right)-\mathfrak{A}_h\mathbf{T}_{n,1,h}\right):\left(\frac{\boldsymbol{\Lambda}_h^{(k+1)}+\boldsymbol{\Theta}_h^{(k)}}{2}-\mathbf{T}_{n,1,h}\right)\mathrm{d}\mathbf{x}\\&\quad\ge 0,\end{aligned} \tag{7.4}$$

thanks to the monotonicity property of $\mathfrak{A}_h$ due to (1.5) and (1.7). On the other hand, we compute

$$\int_\Omega\left(\mathbf{C}_h^{(k)}-\mathbf{C}_h\right):\left(\frac{\boldsymbol{\Lambda}_h^{(k+1)}+\boldsymbol{\Theta}_h^{(k)}}{2}-\mathbf{T}_{n,1,h}\right)\mathrm{d}\mathbf{x}=\frac{1}{4\tau}\left(\left\|\boldsymbol{\Theta}_h^{(k)}-\boldsymbol{\Theta}_h\right\|_{L_2(\Omega)}^2-\left\|\boldsymbol{\Lambda}_h^{(k+1)}-\boldsymbol{\Lambda}_h\right\|_{L_2(\Omega)}^2\right). \tag{7.5}$$

Hence, we find that

$$\frac{1}{4\tau}\left(\left\|\boldsymbol{\Theta}_h^{(k)}-\boldsymbol{\Theta}_h\right\|_{L_2(\Omega)}^2-\left\|\boldsymbol{\Lambda}_h^{(k+1)}-\boldsymbol{\Lambda}_h\right\|_{L_2(\Omega)}^2\right)\ge 0, \tag{7.6}$$

and therefore, in view of (7.3),

$$\left\|\boldsymbol{\Lambda}_h^{(k+1)}-\boldsymbol{\Lambda}_h\right\|_{L_2(\Omega)}\le\left\|\boldsymbol{\Theta}_h^{(k)}-\boldsymbol{\Theta}_h\right\|_{L_2(\Omega)}\le\left\|\boldsymbol{\Lambda}_h^{(k)}-\boldsymbol{\Lambda}_h\right\|_{L_2(\Omega)}. \tag{7.7}$$

This guarantees that the sequence $\|\boldsymbol{\Lambda}_h^{(k)}-\boldsymbol{\Lambda}_h\|_{L_2(\Omega)}$ of nonnegative real numbers is monotonic nonincreasing, and so converging. In particular, we have

$$\lim_{k\to\infty}\left(\left\|\boldsymbol{\Lambda}_h^{(k)}-\boldsymbol{\Lambda}_h\right\|_{L_2(\Omega)}-\left\|\boldsymbol{\Lambda}_h^{(k+1)}-\boldsymbol{\Lambda}_h\right\|_{L_2(\Omega)}\right)=0.$$

Consequently, (7.7) yields that

$$\lim_{k\to\infty}\left\|\boldsymbol{\Theta}_h^{(k)}-\boldsymbol{\Theta}_h\right\|_{L_2(\Omega)}=0.$$

From (7.6), this also means that

$$\lim_{k\to\infty}\left\|\boldsymbol{\Lambda}_h^{(k)}-\boldsymbol{\Lambda}_h\right\|_{L_2(\Omega)}=0.$$

With these two limits, (7.2) implies

$$\lim_{k\to\infty} \frac{1}{n} \left\| \mathbf{T}_h^{(k)} - \mathbf{T}_{n,1,h} \right\|_{L_2(\Omega)} = 0.$$

That completes the proof. □

REMARK 7.2 (Post-processing). Since $\mathbf{T}_h^{(k+\frac{1}{2})}$ within the iterative algorithm does not satisfy the constraint, it seems difficult to prove its convergence to $\mathbf{T}_{n,1,h}$, and as a consequence the convergence of $\mathbf{u}_h^{(k)}$ to $\mathbf{u}_{n,1,h}$, as $k \to \infty$. Instead, given $\mathbf{T}_h^{(k)}$, one can define $\widetilde{\mathbf{u}}_h^{(k)} \in \mathbb{X}_{n,h}$ as the solution to the elasticity problem

$$\int_\Omega \varepsilon\big(\widetilde{\mathbf{u}}_h^{(k)}\big) : \varepsilon(\mathbf{v}_h)\,\mathrm{d}\mathbf{x} = \frac{1}{n}\int_\Omega \mathbf{T}_h^{(k)} : \varepsilon(\mathbf{v}_h)\,\mathrm{d}\mathbf{x} + \int_\Omega \mathfrak{A}_h\big(\mathbf{T}_h^{(k)}\big) : \varepsilon(\mathbf{v}_h)\,\mathrm{d}\mathbf{x} \qquad \forall\, \mathbf{v}_h \in \mathbb{X}_{n,h}.$$

The convergence of $\widetilde{\mathbf{u}}_h^{(k)}$ towards $\mathbf{u}_{n,1,h}$ follows from the convergence of $\mathbf{T}_h^{(k)}$ towards $\mathbf{T}_{n,1,h}$, as $k \to \infty$.

8. Numerical experiments

We now illustrate the performance of the decoupled algorithm in several situations. We start with a setting where the exact solution is accessible in order to demonstrate the asymptotic behavior of the algorithm and to determine adequate values for the numerical parameters to be used in other situations. We then challenge our algorithm in the two-dimensional case of a crack.

The numerical results presented below are obtained using the *deal.ii* library (Arndt *et al.*, 2017). The subdivisions of Ω consist of quadrilaterals/hexahedra. Unless stated otherwise, the stress tensor $\mathbf{T}$ is approximated using piecewise constant polynomials while the displacement $\mathbf{u}$ is approximated by piecewise polynomials of degree 1 in each coordinate direction; see Section 5.3.

8.1 *Details of the implementation*

For a given tolerance parameter TOL>0, the decoupled iterative algorithm described in Section 7 is terminated once the relative tolerance on the increment

$$\frac{\left\| \mathbf{T}_h^{(k+1)} - \mathbf{T}_h^{(k)} \right\|_{L_p(\Omega)} + \left\| \nabla\big(\mathbf{u}_h^{(k+1)} - \mathbf{u}_h^{(k)}\big) \right\|_{L_2(\Omega)}}{\left\| \mathbf{T}_h^{(k)} \right\|_{L_p(\Omega)} + \left\| \nabla \mathbf{u}_h^{(k)} \right\|_{L_2(\Omega)}} \le \mathrm{TOL} \tag{8.1}$$

is satisfied, where $p = 2$ when $t = 1$ and $p = 1$ otherwise.

Each step of the decoupled algorithm requires subiterations (only in Step 1 when $t = 1$), which are terminated once the relative tolerance on the increments is smaller than TOL/5.

TABLE 1 *Asymptotic behavior of $e_{\mathbf{u}}$ and $e_{\mathbf{T}}$ for $n = t = 1$ and $n = 2$ with $t = 1$ or $t = 2$. The method exhibits convergence of order 1 in all cases. This is in accordance with Theorem 5.5 when $t = 1$ but better than predicted for $t > 1$.*

h	$n = 1$		$n = 2$			
	$t = 1$		$t = 1$		$t = 2$	
	$e_{\mathbf{u}}$	$e_{\mathbf{T}}$	$e_{\mathbf{u}}$	$e_{\mathbf{T}}$	$e_{\mathbf{u}}$	$e_{\mathbf{T}}$
2^{-2}	0.14438	0.03946	0.14436	0.05453	0.14434	0.05182
2^{-3}	0.07217	0.01973	0.07217	0.02725	0.07217	0.02486
2^{-4}	0.03609	0.00986	0.03609	0.01363	0.03609	0.01224
2^{-5}	0.01804	0.00493	0.01804	0.00681	0.01804	0.00625
2^{-6}	0.00902	0.00247	0.00902	0.00341	0.00902	0.00327
2^{-7}	0.00451	0.00124	0.00451	0.00171	0.00451	0.00177

8.2 *Validation on smooth solutions*

We illustrate the performance of the decoupled algorithm introduced in Section 7 on the discretization of the regularized system

$$\begin{aligned} a_n(\mathbf{T},\mathbf{S}) + c(\mathbf{T};\mathbf{T},\mathbf{S}) - b(\mathbf{S},\mathbf{u}) &= \int_\Omega \mathbf{G} : \mathbf{S}\,\mathrm{d}\mathbf{x} \qquad \forall \mathbf{S} \in \mathbb{M}, \\ b(\mathbf{T},\mathbf{v}) &= \int_\Omega \mathbf{f}\cdot\mathbf{v}\,\mathrm{d}\mathbf{x} \qquad \forall \mathbf{v} \in \mathbb{X}. \end{aligned} \tag{8.2}$$

The presence of the given tensor $\mathbf{G} : \Omega \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$ on the right-hand side of the first equation allows us to exhibit an exact solution in closed form; compare with (3.2). In fact, we let $\lambda(s) = \mu(s) = (1+s^2)^{-\frac{1}{2}}$, $\Omega = (0,1)^2$ and, given $n \geq 1$, we define $\mathbf{f}$ and $\mathbf{G}$ so that

$$\mathbf{u}(x,y) = \begin{pmatrix} y(1-y) \\ 0 \end{pmatrix}, \qquad \mathbf{T}(x,y) = \begin{pmatrix} e^x & 0 \\ 0 & \cos y \end{pmatrix} \tag{8.3}$$

solves (8.2).

Regarding the numerical parameters, we fix the pseudo-time increment parameter $\tau = 0.01$ and perform simulations for several values of the regularization parameter n and for $t = 1$ (linear regularization) and $t = n$. The computational domain Ω is subdivided by using a sequence of uniform partitions consisting of squares of diameter $h = 2^{-i}$, $i = 0,\ldots,7$. The target tolerance for the iterative algorithm is set to $\mathrm{TOL} = 10^{-5}$.

Convergence as $h \to 0$. We provide in Table 1 the corresponding errors $e_{\mathbf{u}} := \|\nabla(\mathbf{u}_n - \mathbf{u}_{n,h})\|_{L_2(\Omega)}$ and $e_{\mathbf{T}} := \|\mathbf{T}_n - \mathbf{T}_{n,h}\|_{L_p(\Omega)}$. Theorem 5.5 predicts a rate of convergence of $\mathcal{O}(h^{\frac{1}{t}})$ for both quantities which seems to be pessimistic (in this model problem with a smooth solution) since convergence of order $\mathcal{O}(h)$ is observed for $t = 1$ and $t = n$. In fact, we also ran tests with other values of $t > 1$ and observed the same order $\mathcal{O}(h)$.

TABLE 2 *Convergence of the decoupled algorithm when $n \to \infty$ for a fixed spatial resolution ($h = 2^{-7}$) using linear ($t = 1$) and nonlinear ($t = n$) regularization. In the nonlinear regularization case, the error in the stress is always measured in $L_1(\Omega)$ (instead of $L_2(\Omega)$ when $t = 1$). The two algorithms yield similar results.*

	$t = 1$				$t = n$		
	$n = 1.0$	$n = 500$	$n = 1000$		$n = 1.0$	$n = 500$	$n = 1000$
$e_{\mathbf{u}}$	0.80168	0.00927	0.00617	$e_{\mathbf{u}}$	0.80167	0.00519	0.00470
$e_{\mathbf{T}}$	1.53397	0.06777	0.03583	$e_{\mathbf{T}}$	2.18173	0.04052	0.02234

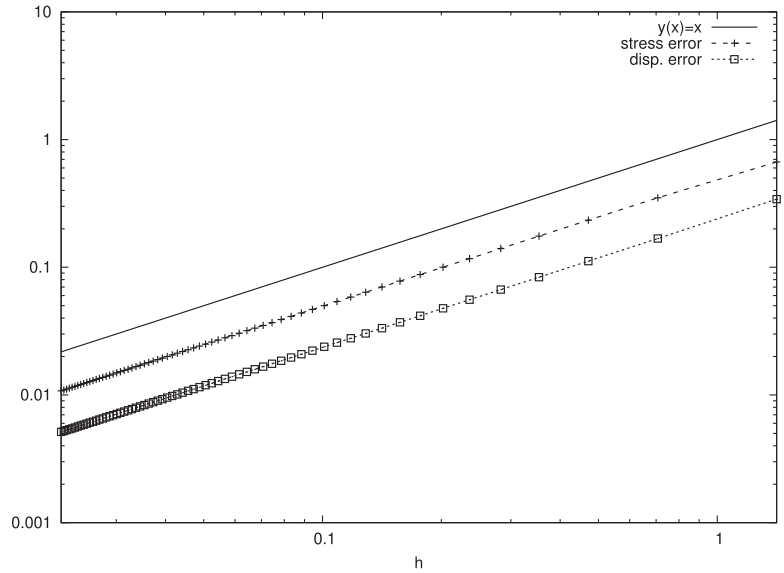

FIG. 1. Decay of $\|\mathbf{u} - \mathbf{u}_h\|_{L_2(\Omega)}$ and $\|\mathbf{T} - \mathbf{T}_h\|_{L_1(\Omega)}$ as a function of the mesh size h using the unstable pair in (5.52). Both quantities decay linearly.

Convergence as $n \to \infty$. We now turn our attention to the convergence of the algorithm when $n \to \infty$ for a fixed subdivision corresponding to $h = 2^{-7}$. Again we consider two cases: $t = 1$ (linear regularization) and $t = n$. The data $\mathbf{f}$ and $\mathbf{G}$ are modified so that $(\mathbf{u}, \mathbf{T})$ given by (8.3) solves (8.2) without regularization, i.e., without the bilinear form $a_n(\cdot, \cdot)$. The results are reported in Table 2; they indicate that in this smooth setting, $e_{\mathbf{u}} + e_{\mathbf{T}} \to 0$ as $n \to \infty$.

8.3 *Inf-sup condition*

We conclude the section containing our numerical experiments with an observation on the inf-sup condition when using quadrilaterals. We consider the discretization of the linear problem, for which

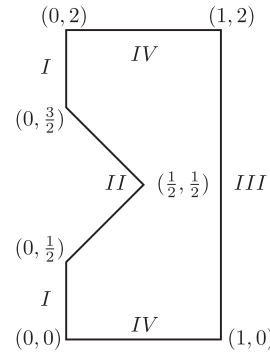

FIG. 2. Crack problem. A horizontal compressive force $\mathbf{Tn} = (f, 0)^{\mathrm{T}}$ for $f > 0$ is applied on side *III*, while no force (i.e., $\mathbf{Tn} = \mathbf{0}$) is imposed on the side marked by *I* and *II*. The top and bottom sides are fixed, i.e., $\mathbf{u} = \mathbf{0}$.

FIG. 3. Crack problem. The deformed domain for different force magnitudes $f = 0.25, 0.5, 0.75, 1$ (from left to right) pulling the right face of the computational domain. The gray scale describes the magnitude of the displacement $|\mathbf{u}|$, where white corresponds to 0 and black to 0.92.

the solution $(\mathbf{u}, \mathbf{T}) \in \mathbb{X} \times \mathbb{M}$ is defined as the one satisfying

$$\int_\Omega \mathbf{T} : \mathbf{S} - \int_\Omega \varepsilon(\mathbf{u}) : \mathbf{S} + \int_\Omega \varepsilon(\mathbf{v}) : \mathbf{T} = \int_\Omega \mathbf{f} \cdot \mathbf{v} \qquad \forall\, (\mathbf{v}, \mathbf{S}) \in \mathbb{X} \times \mathbb{M}.$$

In view of the discussion in Section 5.3, any pair of discrete spaces satisfying $\varepsilon(\mathbb{X}_{n,h}) \subset \mathbb{M}_{n,h}$, such as in (5.50) or in (5.51), yields an inf-sup stable scheme. In contrast, unstable modes (that violate the discrete inf-sup condition with an h-independent positive inf-sup constant) can be proved to exist when using the pair in (5.52). However, for the exact (smooth) solution

$$\mathbf{u}(x, y) = \begin{pmatrix} x\, e^{y} \\ \sin x \end{pmatrix}, \qquad \mathbf{T}(x, y) = \varepsilon(\mathbf{u}(x, y))$$

on a square domain $\Omega = (0, 1)^2$, the finite element approximations using this unstable pair showed no signs of instability in our numerical experiments. In fact, a linear rate of convergence for $\|\nabla(\mathbf{u} - \mathbf{u}_h)\|_{L_2(\Omega)}$ and $\|\mathbf{T} - \mathbf{T}_h\|_{L_1(\Omega)}$ was observed in the limit of $h \to 0$; see Fig. 1.

It is worth mentioning that, when using $(\mathcal{Q}_h^1)_{\mathrm{sym}}^{d\times d}$ instead of $(\mathcal{Q}_h^0)_{\mathrm{sym}}^{d\times d}$ for $\mathbf{T}_h$, the approximation of $\mathbf{u}_h$ remains exactly the same while the approximation of $\mathbf{T}_h$ is more accurate on any given subdivision, but it still exhibits only first-order convergence as $h \to 0$. The intriguing fact that, for the exact solution

TABLE 3 *Evolutions of* $\|\nabla \mathbf{u}_h\|_{L_\infty(\Omega)}$ *and* $\|\mathbf{T}_h\|_{L_\infty(\Omega)}$ *as functions of the force magnitude* f *pulling the right face of the domain. The influence of increasing the magnitude of the force is severe on the stress while relatively moderate on the strain. This is in accordance with the properties of the strain-limiting model considered.*

	$f = 0.25$	$f = 0.5$	$f = 0.75$	$f = 1$	$f = 1.25$	$f = 1.5$
$\|\nabla \mathbf{u}_{n,1,h}\|_{L_\infty(\Omega)}$	1.0656	2.2510	3.5032	5.2703	7.0492	8.8003
$\|\mathbf{T}_{n,1,h}\|_{L_\infty(\Omega)}$	0.92231	5.3090	18.17	46.5215	95.3902	166.335

$(\mathbf{u}, \mathbf{T})$ considered above, the scheme exhibits the optimal rate of convergence dictated by interpolation theory, even though an inf-sup unstable finite element pair is being used, will be the subject of future work.

8.4 *Crack problem*

We consider the 'crack problem' described in Fig. 2. A horizontal force of magnitude f is applied to the right face of the domain (*III*), while the left faces (*I* and *II*) are free to deform (i.e., no external force is being applied there). The top and bottom (*IV*) are fixed with $\mathbf{u} = \mathbf{0}$.

We set $\lambda(s) = \mu(s) = (1 + s^2)^{-\frac{1}{2}}$. In view of the performance observed in Section 8.2 we set the numerical parameters to $\tau = 2$, $n = 100$ and $t = 1$. The domain is partitioned into 16384 quadrilaterals of minimal diameter $h = 0.011$. The stress is approximated in $(\mathcal{Q}_h^0)^{d\times d}_{\text{sym}}$ and the displacement in $(\mathcal{Q}_h^1)^d \cap \mathbb{X}_n$. In Fig. 3 we provide the deformed domain predicted by the algorithm for different values of f. We also report in Table 3 the evolution of $\|\nabla \mathbf{u}_h\|_{L_\infty(\Omega)}$ and $\|\mathbf{T}_h\|_{L_\infty(\Omega)}$ as the magnitude of the force increases. The influence of the latter is severe on $\|\mathbf{T}_h\|_{L_\infty(\Omega)}$ while relatively moderate on $\|\varepsilon(\mathbf{u}_h)\|_{L_\infty(\Omega)} \leq \|\nabla \mathbf{u}_h\|_{L_\infty(\Omega)}$. This is in accordance with the properties of the strain-limiting model considered.

Funding

National Science Foundation (DMS-1817691 to A.B).

REFERENCES

ACOSTA, G., DURÁN, R. G. & MUSCHIETTI, M. A. (2006) Solutions of the divergence operator on John domains. *Adv. Math.*, **206**, 373–401.

ARNDT, D., BANGERTH, W., DAVYDOV, D., HEISTER, T., HELTAI, L., KRONBICHLER, M., MAIER, M., PELTERET, J.-P., TURCKSIN, B. & WELLS, D. (2017) The deal. II library, version 8.5. *J. Numer. Math.*, **25**, 137–146.

BECK, L., BULÍČEK, M., MÁLEK, J. & SÜLI, E. (2017) On the existence of integrable solutions to nonlinear elliptic systems and variational problems with linear growth. *Arch. Ration. Mech. Anal.*, **225**, 717–769.

BULÍČEK, M., MÁLEK, J., RAJAGOPAL, K. & SÜLI, E. (2014) On elastic solids with limiting small strain: modelling and analysis. *EMS Surv. Math. Sci.*, **1**, 283–332.

CIARLET, P. G. (1991) Basic error estimates for elliptic problems. *Handbook of Numerical Analysis*, vol. II, Amsterdam: North-Holland, pp. 17–351.

GIRAULT, V. & RAVIART, P.-A. (1986) *Finite Element Methods for Navier-Stokes Equations: Theory and Algorithms.* Springer Series in Computational Mathematics, vol. 5. Berlin: Springer.

GRAFAKOS, L. (2014) *Classical Fourier Analysis.* Graduate Texts in Mathematics, vol. 249, 3rd edn. New York: Springer.

GRONWALL, T. H. (1913) Über die Laplacesche Reihe. *Math. Ann.*, **74**, 213–270.

JIANG, R. & KAURANEN, A. (2017) Korn's inequality and John domains. *Calc. Var. Partial Differ. Equ.*, **56**, Art. 109, 18.

LIONS, J.-L. (1969) *Quelques Méthodes de Résolution des Problèmes aux Limites Non Linéaires*. Paris, France: Dunod.

LIONS, P.-L. & MERCIER, B. (1979) Splitting algorithms for the sum of two nonlinear operators. *SIAM J. Numer. Anal.*, **16**, 964–979.

MÁLEK, J., NEČAS, J., ROKYTA M. & RUŽIČKA, M. (1996) *Weak and Measure-Valued Solutions to Evolutionary PDEs.* Applied Mathematics and Mathematical Computation, vol. 13. London: Chapman & Hall.

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

RAJAGOPAL, K. R. (2007) The elasticity of elasticity. *Z. Angew. Math. Phys.*, **58**, 309–317.

RAO, M. M. & REN, Z. D. (2002) *Applications of Orlicz Spaces.* Monographs and Textbooks in Pure and Applied Mathematics, vol. 250. New York: Marcel Dekker.

SCOTT, L. R. & ZHANG, S. (1990) Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Math. Comput.*, **54**, 483–493.

SHOWALTER, R. E. (1997) *Monotone Operators in Banach Spaces and Nonlinear Partial Differential Equations.* Math. Surveys and Monographs, vol. 49. Providence, RI: AMS.