A priori and a posteriori error analyses of augmented twofold saddle point formulations for nonlinear elasticity problems*

Gabriel N. Gatica[†] Antonio Márquez[‡] Walter Rudolph[§]

Abstract

In this paper we introduce and analyze new augmented mixed finite element methods for a class of nonlinear elasticity problems arising in hyperelasticity. The starting mixed method is based on the incorporation of the strain tensor as an auxiliary unknown, which, together with the usual stress-displacement-rotation approach employed in linear elasticity, yields a nonlinear twofold saddle point operator equation as the resulting weak formulation. We first extend known results on the well-posedness of the associated Galerkin scheme with PEERS of order $k = 0$ to the case $k \geq 1$. Then the augmented schemes are obtained by adding consistent Galerkin-type terms arising first from the constitutive equation, and then from the equilibrium equation and the relations defining the rotation in terms of the displacement and the strain tensor as independent unknown, all of them multiplied by suitably chosen stabilization parameters. We apply classical results on the solvability analysis of nonlinear saddle point and strongly monotone operator equations to prove that the corresponding continuous and discrete augmented schemes are well-posed. In particular, we show that the well-posedness of a partially augmented Galerkin scheme is ensured by any finite element subspace for the strain tensor together with the PEERS space of order $k \geq 0$ for the remaining unknowns, whereas any finite element subspace of the whole continuous space will do in the case of a fully augmented scheme. Then, we derive reliable and efficient residual-based a posteriori error estimators for all the schemes. Finally, we provide several numerical results illustrating the good performance of the mixed finite element methods, confirming the theoretical properties of the estimators, and showing the behaviour of the associated adaptive algorithms.

Key words: twofold saddle point formulation, augmented approach, mixed finite element method, a posteriori error estimator

Mathematics Subject Classifications (2010): 65J15, 65N12, 65N15, 65N30, 65N50, 74B20

1 Introduction

The nonlinear twofold saddle point operator equations, also called dual-dual variational formulations, arised about a decade ago from the necessity of applying dual-mixed methods to a class of nonlinear boundary value problems appearing in continuum mechanics, particularly in potential theory, heat conduction, elasticity, and fluid mechanics. At that time, the usual procedure for treating nonlinear elliptic equations in divergence form was based on the inversion, thanks to the implicit function

[*]This research was partially supported by BASAL project CMM, Universidad de Chile, and by Centro de Investigación en Ingeniería Matemática (CI^{2}MA), Universidad de Concepción.

[†]CI^{2}MA and Departamento de Ingeniería Matemática, Universidad de Concepción, Casilla 160-C, Concepción, Chile, email: ggatica@ing-mat.udec.cl

[‡]Departamento de Construcción e Ingeniería de Fabricación, Universidad de Oviedo, Oviedo, España, email: amarquez@uniovi.es

[§]Departamento de Ingeniería Matemática, Universidad de Concepción, Casilla 160-C, Concepción, Chile, email: wrudolph@udec.cl

theorem, of the constitutive equations involved. In heat conduction, for instance, the gradient of the temperature was expressed, when possible, as a function of the temperature and the flux variable. Then, in order to deal with the case of constitutive equations that are not explicitly invertible, an alternative approach was proposed first in [45] and [46], in connection with the coupling of mixed finite element and boundary element methods for solving nonlinear transmission problems. This new methodology, which has been later on extended to several other nonlinear boundary value problems, is based on the introduction of auxiliary unknowns such as the gradient of the temperature (in heat conduction) or the strain tensor (in elasticity and fluid mechanics), which yields twofold saddle point operator equations as the resulting weak formulations (see, e.g. [1], [4], [5], [28], [31], [32], [33], [34], [35], [36], [37], [38], [39], [43], and [44]). Actually, the idea of introducing further incognitas to deal with the nonlinearities of a boundary value problem had been employed before in [24], [14], and [15], in which the associated Galerkin schemes were named expanded mixed finite element methods. However, it is important to remark that the twofold saddle point structure has only been obtained and analyzed in the works mentioned above. Now, particularly significant for the current paper are the results from [5], where a detailed analysis of the continuous and discrete dual-mixed formulations of a two-dimensional nonlinear boundary value problem arising in hyperelasticity was developed. In particular, it is shown there that stable mixed finite elements for linear elasticity, such as PEERS of order 0, also lead to well-posed Galerkin schemes for that nonlinear problem. The corresponding extensions to nonlinear incompressible elasticity and quasi-Newtonian Stokes flows were short after provided in [44], [32] and [31]. Further applications to diverse transmission problems are given in [4], [34], [38], and [43]. The results in [32] were also extended in [19] and [49] to a setting in reflexive Banach spaces, thus allowing other nonlinear models such as the Carreau law for viscoplastic flows. More recently, a velocity-pseudostress formulation for the same quasi-Newtonian Stokes flows considered in [32] and [19] is analyzed in [42]. In this case, in addition to introducing the gradient of the velocity of the fluid as an auxiliary unknown, the pressure is eliminated using the incompressibility condition, and similarly as in [32] the resulting variational formulation still shows a twofold saddle point structure.

In turn, the abstract framework that is needed for the solvability analysis of the continuous and discrete nonlinear twofold saddle point formulations, which constitutes a natural extension of the classical Babuška-Brezzi theory, was derived in [28] and [39]. It is quite clear from these works that, as for the case of linear saddle point problems (see, e.g. [10]), the hardest aspect of the associated numerical analysis refer to the choice of suitable finite element subspaces satisfying the discrete inf-sup conditions involved. Moreover, while it is usually possible to establish specific well-posed Galerkin schemes for each one of the problems studied so far, it is also true that not any polynomial degree can be employed for the local approximations of the unknowns and that additional necessary conditions among the global subspaces need to be satisfied.

In the case of linear problems, the restrictions and conditions mentioned in the previous paragraph have been somehow overcomed through the application of several stabilization procedures developed during the last two decades, which have allowed more flexibility in the choice of the corresponding finite element subspaces. In particular, the augmented variational formulations, also known as Galerkin least-squares methods, and which go back to [22] and [23], have already been extended in different directions. Some applications to elasticity problems can be found in [26] and [12], and a non-symmetric variant was considered in [18] for the Stokes problem. In addition, stabilized mixed finite element methods for related problems, including Darcy and incompressible flows, can be seen in [3], [9], [25], and [52]. For an abstract framework concerning the stabilization of general mixed finite element methods, we refer to [11]. Furthermore, a new stabilized mixed finite element method for plane linear elasticity was introduced and analyzed in [29]. The approach there is based on the incorporation of suitable Galerkin least-squares terms arising from the constitutive and equilibrium equations, and from the relation defining the rotation in terms of the displacement. It is shown that the resulting continuous and discrete augmented formulations are well posed, and that the latter becomes locking-free for both Dirichlet and mixed boundary conditions. Moreover, in the case of pure Dirichlet conditions, the augmented formulation becomes strongly coercive, and hence arbitrary finite element subspaces can be employed in the associated Galerkin scheme, which constitutes one of its main advantages with respect to other methods. In particular, Raviart-Thomas spaces of lowest order for the stress tensor, continuous piecewise linear elements for the displacement, and piecewise constants for the rotation can be used. The corresponding extensions to non-homogeneous Dirichlet boundary conditions and to three-dimensional elasticity were provided in [30] and [40], respectively. In addition, residual based a posteriori error analyses yielding reliable and efficient estimators for the augmented method from [29], are presented in [7] and [6]. Furthermore, augmented mixed finite element methods for pseudostress-based formulations of the stationary Stokes equations, which further extends the results from [29], [30], and [40], are introduced and analyzed in [21]. The corresponding augmented mixed finite element schemes for the velocity-pressure-stress formulation of the Stokes problem, in which the vorticity is introduced as the Lagrange multiplier taking care of the weak symmetry of the stress, are studied in [20]. The results in [21] and [20] also include the derivation of reliable and efficient residual-based a posteriori error estimators.

Motivated by the discussion from the preceding paragraphs, and in order to achieve also more flexibility in the choice of the finite element subspaces for the dual-mixed formulations of nonlinear boundary value problems, we now aim to extend the applicability of the augmented dual-mixed formulations developed in [29], [30], [40], [21], and [20] to the class of nonlinear twofold saddle point operator equations described above. We are interested in the a priori and a posteriori error analyses of the resulting augmented schemes, and as a model we consider the problem in hyperelasticity studied in [5]. Up to the authors' knowledge, the closest contribution in the proposed direction is given by a partially augmented approach introduced in [42] for the velocity-pseudostress formulation of quasi-Newtonian Stokes flows. Indeed, the redundant incorporation of the constitutive law relating the stress and the strain tensors transforms the original twofold saddle point structure of the nonlinear problem in [42] into a single saddle point operator equation, which certainly simplifies the requirements for well-posedness of the associated Galerkin scheme. The adaptation of this idea to our model problem from [5] constitutes the starting point of the augmented formulations that are introduced and analyzed in the present paper. As a by product of our preliminary analysis, we extend the results from [5] and show that the Galerkin scheme becomes well-posed for any PEERS space of order $k \geq 1$ (not only for $k = 0$ as proved in [5]).

The rest of our work is organized as follows. In Section 2 we introduce the model problem, derive the associated nonlinear operator equation, which, as shown originally in [5], has a twofold saddle point structure, and then discuss the solvability and stability of the continuous and discrete formulations. Next, in Section 3 we propose and analyze a partially augmented approach for our twofold saddle point problem. Classical results on nonlinear functional analysis are applied to prove the well-posedness of the resulting continuous and discrete formulations. In particular, a discrete inf-sup condition for one of the forms involved is no longer required, and hence a larger class of finite element subspaces can be employed to define the Galerkin schemes. Several examples in this direction are specified. The idea and analysis from Section 3 are extended in Section 4 through the introduction of a fully augmented approach. In this case, no discrete inf-sup conditions are needed at all, and therefore the associated discrete scheme becomes well-posed for any finite element subspace. Then, in Section 5 we derive reliable and efficient residual-based a posteriori error estimators for the three Galerkin schemes defined in the previous sections. Finally, several numerical results illustrating the performance of the methods, confirming the reliability and efficiency of the a posteriori estimators, and showing the good behavior of the associated adaptive algorithms, are reported in Section 6.

We end this section with several notations to be used below. In what follows, $\mathrm{R}^{2\times 2}$ is the space of square matrices of orden 2 with real entries, $\mathbb{I} := (\delta_{ij})$ is the identity matrix of $\mathrm{R}^{2\times 2}$, and given $\boldsymbol{\tau} := (\tau_{ij})$, $\boldsymbol{\zeta} := (\zeta_{ij})$ in $\mathrm{R}^{2\times 2}$, we write as usual

$$\boldsymbol{\tau}^{\mathtt{t}} := (\tau_{ji}), \quad \mathrm{tr}(\boldsymbol{\tau}) := \sum_{i=1}^{2} \tau_{ii}, \quad \boldsymbol{\tau}^{\mathtt{d}} := \boldsymbol{\tau} - \frac{1}{2}\,\mathrm{tr}(\boldsymbol{\tau})\,\mathbb{I}, \quad \text{and} \quad \boldsymbol{\tau} : \boldsymbol{\zeta} := \sum_{i,j=1}^{2} \tau_{ij}\,\zeta_{ij},$$

which corresponds, respectively, to the transpose, the trace, and the deviator of a tensor $\boldsymbol{\tau}$, and to the tensorial product between $\boldsymbol{\tau}$ and $\boldsymbol{\zeta}$. In turn, in what follows we utilize standard simplified terminology for Sobolev spaces and norms. In particular, if $\mathcal{O}$ is a domain, $\mathcal{S}$ is a closed Lipschitz curve, and $r \in \mathrm{R}$, we define

$$\mathbf{H}^r(\mathcal{O}) := [H^r(\mathcal{O})]^2, \quad \mathbb{H}^r(\mathcal{O}) := [H^r(\mathcal{O})]^{2\times 2}, \quad \text{and} \quad \mathbf{H}^r(\mathcal{S}) := [H^r(\mathcal{S})]^2.$$

However, when $r = 0$ we usually write $\mathbf{L}^2(\mathcal{O})$, $\mathbb{L}^2(\mathcal{O})$, and $\mathbf{L}^2(\mathcal{S})$ instead of $\mathbf{H}^0(\mathcal{O})$, $\mathbb{H}^0(\mathcal{O})$, and $\mathbf{H}^0(\mathcal{S})$, respectively. The corresponding norms are denoted by $\|\cdot\|_{r,\mathcal{O}}$ (for $H^r(\mathcal{O})$, $\mathbf{H}^r(\mathcal{O})$, and $\mathbb{H}^r(\mathcal{O})$) and $\|\cdot\|_{r,\mathcal{S}}$ (for $H^r(\mathcal{S})$ and $\mathbf{H}^r(\mathcal{S})$). In general, given any Hilbert space H, we use $\mathbf{H}$ and $\mathbb{H}$ to denote $[H]^2$ and $[H]^{2\times 2}$, respectively. In addition, we use $\langle\cdot,\cdot\rangle_{\mathcal{S}}$ to denote the usual duality pairings between $H^{-1/2}(\mathcal{S})$ and $H^{1/2}(\mathcal{S})$, and between $\mathbf{H}^{-1/2}(\mathcal{S})$ and $\mathbf{H}^{1/2}(\mathcal{S})$. Furthermore, with div denoting the usual divergence operator, the Hilbert space

$$\mathbf{H}(\mathrm{div};\mathcal{O}) := \left\{ \mathbf{w} \in \mathbf{L}^2(\mathcal{O}) : \quad \mathrm{div}\,\mathbf{w} \in L^2(\mathcal{O}) \right\},$$

is standard in the realm of mixed problems (see [10], [47]). The space of matrix valued functions whose rows belong to $\mathbf{H}(\mathrm{div};\mathcal{O})$ will be denoted $\mathbb{H}(\mathbf{div};\mathcal{O})$, where $\mathbf{div}$ stands for the action of div along each row of a tensor. The Hilbert norms of $\mathbf{H}(\mathrm{div};\mathcal{O})$ and $\mathbb{H}(\mathbf{div};\mathcal{O})$ are denoted by $\|\cdot\|_{\mathrm{div};\mathcal{O}}$ and $\|\cdot\|_{\mathbf{div};\mathcal{O}}$, respectively. Note that if $\boldsymbol{\tau} \in \mathbb{H}(\mathbf{div};\mathcal{O})$, then $\mathbf{div}\,\boldsymbol{\tau} \in \mathbf{L}^2(\mathcal{O})$. Finally, we employ $\mathbf{0}$ to denote a generic null vector (including the null functional and operator), and use C and c, with or without subscripts, bars, tildes or hats, to denote generic constants independent of the discretization parameters, which may take different values at different places.

2 The twofold saddle point approach

2.1 The model problem

In order to describe the nonlinear problem studied in [5], we now let Ω be a bounded and simply connected polygonal domain in R^2 with Lipschitz-continuous boundary Γ. Our goal is to determine the displacement $\mathbf{u}$ and stress $\boldsymbol{\sigma}$ of a hyperelastic material occupying the region Ω, which is subject to a volume force and has a known displacement on Γ. As a description of the hyperelasticity we assume the validity of the Hencky-Mises stress-strain relation as discussed in [53] (see also [59]). In other words, given $\mathbf{f} \in \mathbf{L}^2(\Omega)$ and $\mathbf{g} \in \mathbf{H}^{1/2}(\Gamma)$, the nonlinear boundary value problem reads as follows: Find a tensor field $\boldsymbol{\sigma}$ and a vector field $\mathbf{u}$ such that

$$\begin{aligned} \boldsymbol{\sigma} &= \widetilde{\lambda}(\|\mathbf{e}(\mathbf{u})^{\mathtt{d}}\|)\,(\mathrm{div}\,\mathbf{u})\,\mathbb{I} + \widetilde{\mu}(\|\mathbf{e}(\mathbf{u})^{\mathtt{d}}\|)\,\mathbf{e}(\mathbf{u}) \quad \text{in} \quad \Omega, \\ \mathbf{div}\,\boldsymbol{\sigma} &= -\mathbf{f} \quad \text{in} \quad \Omega, \qquad \mathbf{u} = \mathbf{g} \quad \text{on} \quad \Gamma, \end{aligned} \tag{2.1}$$

where $\widetilde{\lambda}, \widetilde{\mu} : \mathrm{R}^+ \to \mathrm{R}$ are the nonlinear Lamé functions, $\nabla\mathbf{u}$ is the tensor gradient of $\mathbf{u}$, $\mathbf{e}(\mathbf{u}) := \frac{1}{2}\left(\nabla\mathbf{u} + (\nabla\mathbf{u})^{\mathtt{t}}\right)$ is the strain tensor of small deformations, $\|\cdot\|$ is the euclidean norm in $\mathrm{R}^{2\times 2}$, and

$\boldsymbol{\nu}$ stands for the unit outward normal to Γ. In addition, from now on we suppose that $\widetilde{\lambda},\, \widetilde{\mu} \in C^1(\mathrm{R}^+)$ and that there exist $\kappa, \mu_0, \mu_1, \mu_2 > 0$ such that for all $\rho \geq 0$,

$$\widetilde{\lambda}(\rho) \,=\, \kappa - \frac{1}{2}\widetilde{\mu}(\rho)\,, \qquad \mu_0 \,\leq\, \widetilde{\mu}(\rho) \,<\, 2\kappa\,, \qquad \mu_1 \,\leq\, \widetilde{\mu}(\rho) \,+\, \rho\,\widetilde{\mu}'(\rho) \,\leq\, \mu_2\,. \tag{2.2}$$

2.2 The continuous variational formulation

We now recall from [5] the dual-mixed variational formulation of (2.1). For this purpose we set

$$\lambda(\mathbf{r}) \,:=\, \widetilde{\lambda}(\|\mathbf{r}^{\mathrm{d}}\|) \quad \text{and} \quad \mu(\mathbf{r}) \,:=\, \widetilde{\mu}(\|\mathbf{r}^{\mathrm{d}}\|) \qquad \forall\, \mathbf{r} \in \mathbb{L}^2(\Omega)\,,$$

so that, defining the new unknown $\mathbf{t} \,:=\, \mathbf{e}(\mathbf{u}) \,\in\, \mathbb{L}^2(\Omega)$, problem (2.1) adopts the equivalent form

$$\begin{gathered} \mathbf{t} \,=\, \mathbf{e}(\mathbf{u}) \quad \text{in} \quad \Omega\,, \quad \boldsymbol{\sigma} \,=\, \lambda(\mathbf{t})\,\mathrm{tr}(\mathbf{t})\,\mathbb{I} \,+\, \mu(\mathbf{t})\,\mathbf{t} \quad \text{in} \quad \Omega\,, \\ \mathbf{div}\,\boldsymbol{\sigma} \,=\, -\,\mathbf{f} \quad \text{in} \quad \Omega\,, \qquad \mathbf{u} \,=\, \mathbf{g} \quad \text{on} \quad \Gamma\,. \end{gathered} \tag{2.3}$$

Next, we introduce the subspace of $\mathbb{L}^2(\Omega)$ given by

$$\mathbb{L}^2_{\mathtt{skew}}(\Omega) \,:=\, \Big\{ \boldsymbol{\eta} \in \mathbb{L}^2(\Omega) : \quad \boldsymbol{\eta} \,+\, \boldsymbol{\eta}^{\mathrm{t}} \,=\, \mathbf{0} \Big\}\,.$$

Then, rewriting the identity $\mathbf{t} = \mathbf{e}(\mathbf{u})$ as

$$\mathbf{t} \,=\, \nabla\mathbf{u} \,-\, \boldsymbol{\gamma}\,, \tag{2.4}$$

where

$$\boldsymbol{\gamma} \,:=\, \frac{1}{2}\left(\nabla\mathbf{u} - (\nabla\mathbf{u})^{\mathrm{t}}\right) \tag{2.5}$$

is an auxiliary unknown (named rotation) living in $\mathbb{L}^2_{\mathtt{skew}}(\Omega)$, and following the usual integration by parts procedure (see [5] for details), we arrive at the problem: Find $(\mathbf{t}, \boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\gamma}) \,\in\, \mathbb{L}^2(\Omega) \times \mathbb{H}(\mathbf{div};\Omega) \times \mathbf{L}^2(\Omega) \times \mathbb{L}^2_{\mathtt{skew}}(\Omega)$ such that

$$\begin{aligned} \int_\Omega \Big\{ \lambda(\mathbf{t})\,\mathrm{tr}(\mathbf{t})\,\mathrm{tr}(\mathbf{s}) \,+\, \mu(\mathbf{t})\,\mathbf{t} : \mathbf{s} \Big\} \,-\int_\Omega \boldsymbol{\sigma} : \mathbf{s} \qquad\qquad\qquad &= \quad 0\,, \\ -\int_\Omega \mathbf{t} : \boldsymbol{\tau} \qquad\qquad -\int_\Omega \mathbf{u}\cdot\mathbf{div}\,\boldsymbol{\tau} \,-\int_\Omega \boldsymbol{\gamma} : \boldsymbol{\tau} &= \,-\,\langle\,\boldsymbol{\tau}\,\boldsymbol{\nu}\,,\,\mathbf{g}\,\rangle_\Gamma\,, \\ -\int_\Omega \mathbf{v}\cdot\mathbf{div}\,\boldsymbol{\sigma} \,-\int_\Omega \boldsymbol{\eta} : \boldsymbol{\sigma} &= \quad \int_\Omega \mathbf{f}\cdot\mathbf{v}\,, \end{aligned} \tag{2.6}$$

for all $(\mathbf{s}, \boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\eta}) \in \mathbb{L}^2(\Omega) \times \mathbb{H}(\mathbf{div};\Omega) \times \mathbf{L}^2(\Omega) \times \mathbb{L}^2_{\mathtt{skew}}(\Omega)$.

Next, we notice that (2.6) has the typical twofold saddle point structure (see, e.g. [28], [39]). In fact, let us define the Hilbert spaces $X_1 \,:=\, \mathbb{L}^2(\Omega)$, $M_1 \,:=\, \mathbb{H}(\mathbf{div};\Omega)$, and $M \,:=\, \mathbf{L}^2(\Omega)\times\mathbb{L}^2_{\mathtt{skew}}(\Omega)$, provided with the usual norms and product norms, respectively, and the nonlinear operator $\mathbb{A}_1 : X_1 \to X_1'$, the bounded linear operators $\mathbb{B}_1 : X_1 \to M_1'$ and $\mathbb{B} : M_1 \to M'$, and the bounded linear functionals $\mathbb{H} \in X_1'$, $\mathbb{G} \in M_1'$ and $\mathbb{F} \in M'$, given for each $\mathbf{r}, \mathbf{s} \in X_1$, $\boldsymbol{\zeta}, \boldsymbol{\tau} \in M_1$ and $(\mathbf{v}, \boldsymbol{\eta}) \in M$ as

$$\begin{aligned} [\mathbb{A}_1(\mathbf{r}), \mathbf{s}] &:= \int_\Omega \Big\{ \lambda(\mathbf{r})\,\mathrm{tr}(\mathbf{r})\,\mathrm{tr}(\mathbf{s}) \,+\, \mu(\mathbf{r})\,\mathbf{r} : \mathbf{s} \Big\}\,, \\ [\mathbb{B}_1(\mathbf{r}), \boldsymbol{\tau}] &:= -\int_\Omega \mathbf{r} : \boldsymbol{\tau}\,, \\ [\mathbb{B}(\boldsymbol{\zeta}), (\mathbf{v}, \boldsymbol{\eta})] &:= -\int_\Omega \mathbf{v}\cdot\mathbf{div}\,\boldsymbol{\zeta} - \int_\Omega \boldsymbol{\zeta} : \boldsymbol{\eta}\,, \end{aligned} \tag{2.7}$$

and

$$[\mathbb{H}, \mathbf{s}] := 0, \quad [\mathbb{G}, \boldsymbol{\tau}] \;:=\; -\langle \boldsymbol{\tau}\boldsymbol{\nu}\,,\,\mathbf{g}\rangle_\Gamma\,, \quad \text{and} \quad [\mathbb{F}, (\mathbf{v},\boldsymbol{\eta})] := \int_\Omega \mathbf{f}\cdot\mathbf{v}\,,$$

where the brackets $[\cdot,\cdot]$ denote the duality pairings induced by the corresponding operators and functionals.

Then, it is easy to see that the variational formulation (2.6) can be rewritten as: Find $(\mathbf{t},\boldsymbol{\sigma},(\mathbf{u},\gamma)) \in \mathbf{X} := X_1 \times M_1 \times M$ such that

$$\begin{array}{llllll} [\mathbb{A}_1(\mathbf{t}),\mathbf{s}] & + \; [\mathbb{B}_1(\mathbf{s}),\boldsymbol{\sigma}] & & = & [\mathbb{H},\mathbf{s}] & \forall\,\mathbf{s}\in X_1\,, \\ [\mathbb{B}_1(\mathbf{t}),\boldsymbol{\tau}] & & + \; [\mathbb{B}(\boldsymbol{\tau}),(\mathbf{u},\gamma)] & = & [\mathbb{G},\boldsymbol{\tau}] & \forall\,\boldsymbol{\tau}\in M_1\,, \\ & [\mathbb{B}(\boldsymbol{\sigma}),(\mathbf{v},\boldsymbol{\eta})] & & = & [\mathbb{F},(\mathbf{v},\boldsymbol{\eta})] & \forall\,(\mathbf{v},\boldsymbol{\eta})\in M\,. \end{array} \tag{2.8}$$

The abstract theory for this kind of twofold saddle point operator equation, including the analysis of the associated discrete formulation (whose definition is pretty straightforward from (2.8)), is already well known (see [28], [39]), and their main results are recalled in the following section.

2.3 Abstract theory for twofold saddle point operator equations

Let X_1, M_1, and M be Hilbert spaces, and consider a nonlinear operator $\mathbf{A}_1 : X_1 \to X_1'$, and linear bounded operators $\mathbf{B}_1 : X_1 \to M_1'$ and $\mathbf{B} : M_1 \to M'$, with transposes $\mathbf{B}_1' : M_1 \to X_1'$ and $\mathbf{B}' : M \to M_1'$, respectively. Then, given $(\mathbf{H},\mathbf{G},\mathbf{F}) \in X_1' \times M_1' \times M'$, we are interested in the following nonlinear variational problem (written as a matrix operator equation): Find $(\mathbf{t},\boldsymbol{\sigma},u) \in X_1 \times M_1 \times M$ such that

$$\begin{pmatrix} \mathbf{A}_1 & \mathbf{B}_1' & \mathbf{O} \\ \mathbf{B}_1 & \mathbf{O} & \mathbf{B}' \\ \mathbf{O} & \mathbf{B} & \mathbf{O} \end{pmatrix} \begin{pmatrix} \mathbf{t} \\ \boldsymbol{\sigma} \\ u \end{pmatrix} = \begin{pmatrix} \mathbf{H} \\ \mathbf{G} \\ \mathbf{F} \end{pmatrix}. \tag{2.9}$$

We have the following theorem.

THEOREM 2.1 *Let* $V := \mathrm{Ker}\,(\mathbf{B})$, *define* $V_1 := \{\mathbf{s}\in X_1 \colon [\mathbf{B}_1(\mathbf{s}),\boldsymbol{\tau}] = 0 \;\forall\,\boldsymbol{\tau}\in V\}$, *and let* $\Pi_1 : X_1' \to V_1'$ *be the operator defined by* $\Pi_1(\mathbf{H}) = \mathbf{H}|_{V_1}$ *for all* $\mathbf{H}\in X_1'$. *Assume that*

i) *the nonlinear operator* $\mathbf{A}_1 : X_1 \to X_1'$ *is Lipschitz continuous with a Lipschitz constant* $\gamma > 0$, *and for any* $\tilde{\mathbf{t}} \in X_1$, *the nonlinear operator* $\Pi_1\mathbf{A}_1(\cdot + \tilde{\mathbf{t}}) : V_1 \to V_1'$ *is strongly monotone with a monotonicity constant* $\alpha > 0$ *independent of* $\tilde{\mathbf{t}}$.

ii) *there exists* $\beta > 0$ *such that for all* $v \in M$

$$\sup_{\boldsymbol{\tau}\in M_1\setminus\{\mathbf{0}\}} \frac{[\mathbf{B}(\boldsymbol{\tau}),v]}{||\boldsymbol{\tau}||_{M_1}} \;\geq\; \beta\,||v||_M\,; \tag{2.10}$$

iii) *there exists* $\beta_1 > 0$ *such that for all* $\boldsymbol{\tau}\in V$

$$\sup_{\mathbf{s}\in X_1\setminus\{\mathbf{0}\}} \frac{[\mathbf{B}_1(\mathbf{s}),\boldsymbol{\tau}]}{||\mathbf{s}||_{X_1}} \;\geq\; \beta_1\,||\boldsymbol{\tau}||_{M_1}\,; \tag{2.11}$$

Then, for each $(\mathbf{H},\mathbf{G},\mathbf{F}) \in X_1' \times M_1' \times M'$ *there exists a unique* $(\mathbf{t},\boldsymbol{\sigma},u) \in X_1 \times M_1 \times M$ *solution of* (2.9). *Moreover, there exists* $C > 0$, *independent of the solution, such that*

$$\|(\mathbf{t},\boldsymbol{\sigma},u)\|_{X_1\times M_1\times M} \;\leq\; C\Big\{\,\|\mathbf{H}\| + \|\mathbf{G}\| + \|\mathbf{F}\| + \|\mathbf{A}_1(\mathbf{0})\|\,\Big\}\,. \tag{2.12}$$

Proof. See [28, Theorem 2.4] (see also [39, Theorem 2.1] or [43, Theorem 4.1]). □

Now, let $X_{1,h}$, $M_{1,h}$ and M_h be finite dimensional subspaces of X_1, M_1 and M, respectively. Then the Galerkin scheme associated with (2.9) reads as follows: Find $(\mathbf{t}_h, \boldsymbol{\sigma}_h, u_h) \in X_{1,h} \times M_{1,h} \times M_h$ such that

$$\begin{aligned}
[\mathbf{A}_1(\mathbf{t}_h), \mathbf{s}_h] \; + \; [\mathbf{B}_1(\mathbf{s}_h), \boldsymbol{\sigma}_h] \qquad\qquad &= [\mathbf{H}, \mathbf{s}_h] \quad \forall\, \mathbf{s}_h \in X_{1,h}\,, \\
[\mathbf{B}_1(\mathbf{t}_h), \boldsymbol{\tau}_h] \qquad\qquad + \; [\mathbf{B}(\boldsymbol{\tau}_h), u_h] &= [\mathbf{G}, \boldsymbol{\tau}_h] \quad \forall\, \boldsymbol{\tau}_h \in M_{1,h}\,, \\
[\mathbf{B}(\boldsymbol{\sigma}_h), v_h] \qquad\qquad &= [\mathbf{F}, v_h] \quad \forall\, v_h \in M_h\,.
\end{aligned} \tag{2.13}$$

The discrete analogue of Theorem 2.1 is established next.

THEOREM 2.2 *Let $V_h := \{\boldsymbol{\tau}_h \in M_{1,h} : \; [\mathbf{B}(\boldsymbol{\tau}_h), v_h] = 0 \;\; \forall\, v_h \in M_h\}$, define the space $V_{1,h} := \{\mathbf{s}_h \in X_{1,h} : \; [\mathbf{B}_1(\mathbf{s}_h), \boldsymbol{\tau}_h] = 0 \;\; \forall\, \boldsymbol{\tau}_h \in V_h\}$ and let $\Pi_{1,h} : X'_{1,h} \to V'_{1,h}$ be the operator defined by $\Pi_{1,h}(\mathbf{H}_h) = \mathbf{H}_h|_{V_{1,h}}$ for all $\mathbf{H}_h \in X'_{1,h}$. Further, let $\mathbf{A}_{1,h} := p'_h \mathbf{A}_1 : X_1 \to X'_{1,h}$ where $p_h : X_{1,h} \to X_1$ is the canonical injection with adjoint $p'_h : X'_1 \to X'_{1,h}$. Assume that*

i) *the nonlinear operator $\mathbf{A}_{1,h} : X_1 \to X'_{1,h}$ is Lipschitz-continuous with a Lipschitz constant $\gamma_h > 0$, and for any $\tilde{\mathbf{t}} \in X_{1,h}$, the nonlinear operator $\Pi_{1,h}\mathbf{A}_{1,h}(\cdot + \tilde{\mathbf{t}}) : V_{1,h} \to V'_{1,h}$ is strongly monotone with a monotonicity constant $\alpha_h > 0$ independent of $\tilde{\mathbf{t}}$.*

ii) *there exists $\beta_h > 0$ such that for all $v_h \in M_h$*

$$\sup_{\boldsymbol{\tau}_h \in M_{1,h}\backslash\{\mathbf{0}\}} \frac{[\mathbf{B}(\boldsymbol{\tau}_h), v_h]}{||\boldsymbol{\tau}_h||_{M_1}} \;\geq\; \beta_h\, ||v_h||_M; \tag{2.14}$$

iii) *there exists $\beta_{1,h} > 0$ such that for all $\boldsymbol{\tau}_h \in V_h$*

$$\sup_{\mathbf{s}_h \in X_{1,h}\backslash\{\mathbf{0}\}} \frac{[\mathbf{B}_1(\mathbf{s}_h), \boldsymbol{\tau}_h]}{||\mathbf{s}_h||_{X_1}} \;\geq\; \beta_{1,h}\, ||\boldsymbol{\tau}_h||_{M_1}; \tag{2.15}$$

Then, for each $(\mathbf{H}, \mathbf{G}, \mathbf{F}) \in X'_1 \times M'_1 \times M'$ there exists a unique $(\mathbf{t}_h, \boldsymbol{\sigma}_h, u_h) \in X_{1,h} \times M_{1,h} \times M_h$ solution of (2.13). Moreover, there exists $C_h > 0$, independent of the solution, but depending on h, such that

$$\|(\mathbf{t}_h, \boldsymbol{\sigma}_h, u_h)\|_{X_1 \times M_1 \times M} \;\leq\; C_h \left\{ \|\mathbf{H}_h\| + \|\mathbf{G}_h\| + \|\mathbf{F}_h\| + \|\mathbf{A}_{1,h}(\mathbf{0})\| \right\},$$

where $\mathbf{H}_h := \mathbf{H}|_{X_{1,h}}$, $\mathbf{G}_h := \mathbf{G}|_{M_{1,h}}$, and $\mathbf{F}_h := \mathbf{F}|_{M_h}$.

Proof. See [28, Theorem 3.2] (see also [39, Theorem 3.1] or [43, Theorem 4.2]). □

Finally, concerning the error analysis, we have the following result.

THEOREM 2.3 *Assume that the hypotheses of Theorems 2.1 and 2.2 are satisfied, and let $(\mathbf{t}, \boldsymbol{\sigma}, u) \in X_1 \times M_1 \times M$ and $(\mathbf{t}_h, \boldsymbol{\sigma}_h, u_h) \in X_{1,h} \times M_{1,h} \times M_h$ be the unique solutions of (2.9) and (2.13), respectively. In addition, suppose that there exist positive constants $\tilde{\gamma}$, $\tilde{\alpha}$, $\tilde{\beta}$, and $\tilde{\beta}_1$ such that $\gamma_h \leq \tilde{\gamma}$, $\alpha_h \geq \tilde{\alpha}$, $\beta_h \geq \tilde{\beta}$, and $\beta_{1,h} \geq \tilde{\beta}_1$ for all h. Then, there exists $C > 0$, independent of h, such that the following Céa error estimate holds:*

$$\|(\mathbf{t}, \boldsymbol{\sigma}, u) - (\mathbf{t}_h, \boldsymbol{\sigma}_h, u_h)\| \;\leq\; C \inf_{\substack{(\mathbf{s}_h, \boldsymbol{\tau}_h, v_h) \\ \in X_{1,h} \times M_{1,h} \times M_h}} \|(\mathbf{t}, \boldsymbol{\sigma}, u) - (\mathbf{s}_h, \boldsymbol{\tau}_h, v_h)\|\,.$$

Proof. See [28, Section 4] (see also [39, Theorem 3.3]). □

2.4 Analysis of the continuous formulation

The well-posedness of the continuous formulation (2.6) (equivalently (2.8)) was already established in [5, Theorem 4.5]. However, for the sake of completeness, and for later use throughout the paper, in what follows we collect the intermediate results yielding that main theorem. In addition, we remark in advance that some parts of the analysis will employ alternative arguments to those given in [5]. We begin with the Gâteaux differentiability of the nonlinear operator $\mathbb{A}_1$ (cf. (2.7)).

LEMMA 2.1 *The nonlinear operator* $\mathbb{A}_1 : X_1 \to X_1'$ *is Gâteaux differentiable in* X_1*, and the family of Gâteaux derivatives* $\{\mathcal{D}\mathbb{A}_1(\mathbf{x})\}_{\mathbf{x}\in X_1}$ *is both uniformly bounded and uniformly elliptic on* $X_1 \times X_1$*. More precisely, there exist positive constants* γ_1, α_1*, depending only on* κ, μ_0, μ_1*, and* μ_2 *(cf.* (2.2)*), such that for all* $\mathbf{x}$, $\mathbf{r}$, $\mathbf{s} \in X_1$*, there hold*

$$| \mathcal{D}\mathbb{A}_1(\mathbf{x})(\mathbf{r},\mathbf{s}) | \,\le\, \gamma_1 \, \|\mathbf{r}\|_{X_1} \, \|\mathbf{s}\|_{X_1} \tag{2.16}$$

and

$$\mathcal{D}\mathbb{A}_1(\mathbf{x})(\mathbf{r},\mathbf{r}) \,\ge\, \alpha_1 \, \|\mathbf{r}\|^2_{X_1} \,. \tag{2.17}$$

Proof. Simple computations and the C^1-regularity of $\widetilde{\mu}$ and $\widetilde{\lambda}$ yield for $\mathbf{x}$, $\mathbf{r}$, $\mathbf{s} \in X_1$, $\mathbf{x}^{\mathrm{d}} \neq \mathbf{0}$,

$$\begin{aligned} \lim_{\epsilon\to 0} \frac{\big[\mathbb{A}_1(\mathbf{x}+\epsilon\,\mathbf{r}) - \mathbb{A}_1(\mathbf{x}),\mathbf{s}\big]}{\epsilon} \,&=\, \int_\Omega \widetilde{\lambda}'(\|\mathbf{x}^{\mathrm{d}}\|) \frac{(\mathbf{x}^{\mathrm{d}} : \mathbf{r}^{\mathrm{d}})}{\|\mathbf{x}^{\mathrm{d}}\|} \operatorname{tr}(\mathbf{x}) \, \operatorname{tr}(\mathbf{s}) \\ +\, \int_\Omega \widetilde{\lambda}(\|\mathbf{x}^{\mathrm{d}}\|) \operatorname{tr}(\mathbf{r}) \, \operatorname{tr}(\mathbf{s}) \,&+\, \int_\Omega \widetilde{\mu}'(\|\mathbf{x}^{\mathrm{d}}\|) \frac{(\mathbf{x}^{\mathrm{d}} : \mathbf{r}^{\mathrm{d}})}{\|\mathbf{x}^{\mathrm{d}}\|} \, \mathbf{x} : \mathbf{s} \,+\, \int_\Omega \widetilde{\mu}(\|\mathbf{x}^{\mathrm{d}}\|) \, \mathbf{r} : \mathbf{s} \,, \end{aligned} \tag{2.18}$$

whereas for $\mathbf{x}^{\mathrm{d}} \,=\, \mathbf{0}$ there holds

$$\begin{aligned} \lim_{\epsilon\to 0} \frac{\big[\mathbb{A}_1(\mathbf{x}+\epsilon\,\mathbf{r}) - \mathbb{A}_1(\mathbf{x}),\mathbf{s}\big]}{\epsilon} \,&=\, \int_\Omega \widetilde{\lambda}'(0) \, \|\mathbf{r}^{\mathrm{d}}\| \operatorname{tr}(\mathbf{x}) \, \operatorname{tr}(\mathbf{s}) \\ +\, \int_\Omega \widetilde{\lambda}(0) \operatorname{tr}(\mathbf{r}) \, \operatorname{tr}(\mathbf{s}) \,&+\, \int_\Omega \widetilde{\mu}'(0) \, \|\mathbf{r}^{\mathrm{d}}\| \, \mathbf{x} : \mathbf{s} \,+\, \int_\Omega \widetilde{\mu}(0) \, \mathbf{r} : \mathbf{s} \,. \end{aligned} \tag{2.19}$$

The above identities show that $\mathbb{A}_1$ is Gâteaux differentiable at $\mathbf{x}$. Moreover, $\mathcal{D}\mathbb{A}_1(\mathbf{x})$ is the bounded linear operator from X_1 into X_1' that can be identified with the bilinear form $\mathcal{D}\mathbb{A}_1(\mathbf{x}) : X_1 \times X_1 \to \mathrm{R}$ defined by

$$\mathcal{D}\mathbb{A}_1(\mathbf{x})(\mathbf{r},\mathbf{s}) \,:=\, \lim_{\epsilon\to 0} \frac{\big[\mathbb{A}_1(\mathbf{x}+\epsilon\,\mathbf{r}) - \mathbb{A}_1(\mathbf{x}),\mathbf{s}\big]}{\epsilon} \qquad \forall\, \mathbf{r},\,\mathbf{s} \,\in\, X_1 \,. \tag{2.20}$$

Hence, the derivation of (2.16) and (2.17) follows from (2.20), using (2.18), (2.19), and the assumptions on $\widetilde{\lambda}$ and $\widetilde{\mu}$ (cf. (2.2)). We omit further details and refer to [27, Lemma 5.1]. In particular, from the analysis there we find that

$$\gamma_1 \,=\, \max\Big\{\mu_2 + 2\,\kappa, 6\,\kappa\Big\} \qquad \text{and} \qquad \alpha_1 \,=\, \min\Big\{\mu_0, \mu_1, 2\,\kappa\Big\} \,. \tag{2.21}$$

Alternatively, one may look at the corresponding analysis within the proof of [5, Lemma 4.1]. □

The Lipschitz-continuity and strong monotonicity of $\mathbb{A}_1$, which is a straightforward consequence of Lemma 2.1, is established next.

LEMMA 2.2 *Let* γ_1 *and* α_1 *be the constants from Lemma* 2.1 *(cf.* (2.21)*). Then, for each* $\mathbf{t}$, $\mathbf{r} \in X_1$ *there hold*

$$\|\mathbb{A}_1(\mathbf{t}) - \mathbb{A}_1(\mathbf{r})\|_{X_1'} \,\le\, \gamma_1 \, \|\mathbf{t} - \mathbf{r}\|_{X_1} \,, \tag{2.22}$$

and

$$[\mathbb{A}_1(\mathbf{t}) - \mathbb{A}_1(\mathbf{r}), \mathbf{t} - \mathbf{r}] \,\ge\, \alpha_1 \, \|\mathbf{t} - \mathbf{r}\|^2_{X_1} \,. \tag{2.23}$$

Proof. Given $\mathbf{t}, \mathbf{r} \in X_1$, a direct application of the mean value theorem yields the existence of a convex combination of $\mathbf{t}$ and $\mathbf{r}$, say $\tilde{\mathbf{r}} \in X_1$, such that

$$[\mathbb{A}_1(\mathbf{t}) - \mathbb{A}_1(\mathbf{r}), \mathbf{s}] = \mathcal{D}\mathbf{A}_1(\tilde{\mathbf{r}})(\mathbf{t} - \mathbf{r}, \mathbf{s}) \quad \forall \mathbf{s} \in X_1 . \tag{2.24}$$

Hence, (2.22) and (2.23) follow easily from (2.24) and the estimates (2.16) and (2.17). □

It is quite clear from Lemma 2.2 that the hypothesis i) of Theorem 2.1 is satisfied by the operator $\mathbb{A}_1$, and hence by our twofold saddle point variational formulation (2.8). In particular, for the strong monotonicity property, let $\mathbb{V} := \operatorname{Ker}(\mathbb{B})$, $\mathbb{V}_1 := \{\mathbf{s} \in X_1 : \quad [\mathbb{B}_1(\mathbf{s}), \boldsymbol{\tau}] = 0 \; \forall \boldsymbol{\tau} \in \mathbb{V}\}$, and $\Pi_1 : X_1' \to \mathbb{V}_1'$ be the operator defined by $\Pi_1(\mathbb{H}) = \mathbb{H}|_{\mathbb{V}_1}$ for all $\mathbb{H} \in X_1'$. Hence, given $\tilde{\mathbf{t}} \in X_1$ and $\mathbf{r}, \mathbf{s} \in \mathbb{V}_1$, we find that

$$\begin{aligned} [\Pi_1 \mathbb{A}_1(\mathbf{r} + \tilde{\mathbf{t}}) - \Pi_1 \mathbb{A}_1(\mathbf{s} + \tilde{\mathbf{t}}), \mathbf{r} - \mathbf{s}] &= [\mathbb{A}_1(\mathbf{r} + \tilde{\mathbf{t}}) - \mathbb{A}_1(\mathbf{s} + \tilde{\mathbf{t}}), \mathbf{r} - \mathbf{s}] \\ = [\mathbb{A}_1(\mathbf{r} + \tilde{\mathbf{t}}) - \mathbb{A}_1(\mathbf{s} + \tilde{\mathbf{t}}), (\mathbf{r} + \tilde{\mathbf{t}}) - (\mathbf{s} + \tilde{\mathbf{t}})] &\geq \alpha_1 \|\mathbf{r} - \mathbf{s}\|_{X_1}^2 , \end{aligned} \tag{2.25}$$

which shows that for each $\tilde{\mathbf{t}} \in X_1$, $\Pi_1 \mathbb{A}_1(\cdot + \tilde{\mathbf{t}}) : \mathbb{V}_1 \to \mathbb{V}_1'$ is strongly monotone with constant $\alpha = \alpha_1$.

On the other hand, the continuous inf-sup condition (2.10) for $\mathbb{B}$, which is equivalent to the surjectivity of this operator, is already a classical requirement in the analysis of the dual-mixed variational formulation for linear elasticity. However, just a few places seem to provide a proof of this inequality (see, e.g. [5, Lemma 4.3]), and therefore a simple version of it is given now.

LEMMA 2.3 *The operator $\mathbb{B} : M_1 \to M'$ is surjective.*

Proof. We first observe from (2.7) that $\mathbb{B}(\boldsymbol{\tau}) = \mathcal{R}^{-1}\big(- \operatorname{\mathbf{div}} \boldsymbol{\tau}, -\frac{1}{2}(\boldsymbol{\tau} - \boldsymbol{\tau}^{\mathrm{t}})\big)$ for each $\boldsymbol{\tau} \in M_1$, where $\mathcal{R} : M' \to M$ is the corresponding Riesz operator. Then, given $\mathbb{F} \in M'$, we let $(\mathbf{v}, \boldsymbol{\eta}) := \mathcal{R}(\mathbb{F}) \in M$ and consider the boundary value problem

$$\operatorname{\mathbf{div}} \mathbf{e}(\mathbf{z}) = \mathbf{v} - \operatorname{\mathbf{div}} \boldsymbol{\eta} \quad \text{in} \quad \Omega , \qquad \mathbf{z} = \mathbf{0} \quad \text{in} \quad \Gamma ,$$

whose weak formulation reduces to: Find $\mathbf{z} \in \mathbf{H}_0^1(\Omega)$ such that

$$\int_\Omega \mathbf{e}(\mathbf{z}) : \mathbf{e}(\mathbf{w}) = - \int_\Omega \mathbf{v} \cdot \mathbf{w} - \int_\Omega \boldsymbol{\eta} : \nabla \mathbf{w} \qquad \forall \mathbf{w} \in \mathbf{H}_0^1(\Omega) .$$

It follows, thanks to Korn's inequality and Lax-Milgram's Lemma, that the above problem has a unique solution $\mathbf{z}$ depending continuously on the data $\mathbf{v}$ and $\boldsymbol{\eta}$. In this way, defining $\widehat{\boldsymbol{\tau}} := \mathbf{e}(\mathbf{z}) + \boldsymbol{\eta}$, we easily see that $\widehat{\boldsymbol{\tau}} \in \mathbb{H}(\operatorname{\mathbf{div}}; \Omega)$, $\big(\operatorname{\mathbf{div}} \widehat{\boldsymbol{\tau}}, \frac{1}{2}(\widehat{\boldsymbol{\tau}} - \widehat{\boldsymbol{\tau}}^{\mathrm{t}})\big) = (\mathbf{v}, \boldsymbol{\eta})$, and therefore $\mathbb{B}(-\widehat{\boldsymbol{\tau}}) = \mathbb{F}$, which confirms the surjectivity of $\mathbb{B}$. □

In turn, the continuous inf-sup condition (2.11) for $\mathbb{B}_1$ requires first to identify $\mathbb{V} := \operatorname{Ker}(\mathbb{B})$, which, acording to the definition of $\mathbb{B}$ (cf. (2.7)), is given by

$$\mathbb{V} = \Big\{ \boldsymbol{\tau} \in \mathbb{H}(\operatorname{\mathbf{div}}; \Omega) : \quad \operatorname{\mathbf{div}} \boldsymbol{\tau} = \mathbf{0} \quad \text{and} \quad \boldsymbol{\tau} = \boldsymbol{\tau}^{\mathrm{t}} \quad \text{in} \quad \Omega \Big\} .$$

Then, using that $\mathbb{V} \subseteq \mathbb{X}_1 := \mathbb{L}^2(\Omega)$ and that the tensors of $\mathbb{V}$ are divergence-free, it follows that

$$\sup_{\substack{\mathbf{s} \in \mathbb{X}_1 \\ \mathbf{s} \neq \mathbf{0}}} \frac{[\mathbb{B}_1(\mathbf{s}), \boldsymbol{\tau}]}{\|\mathbf{s}\|_{0,\Omega}} \geq \frac{[\mathbb{B}_1(-\boldsymbol{\tau}), \boldsymbol{\tau}]}{\|\boldsymbol{\tau}\|_{0,\Omega}} = \|\boldsymbol{\tau}\|_{0,\Omega} = \|\boldsymbol{\tau}\|_{\operatorname{\mathbf{div}};\Omega} \quad \forall \boldsymbol{\tau} \in \mathbb{V} , \tag{2.26}$$

which shows that $\mathbb{B}_1$ satisfies (2.11) with a constant $\beta_1 = 1$.

Alternatively, we see from the definition of $\mathbb{B}_1$ (cf. (2.7)) that the above condition is equivalent to the surjectivity of the operator

$$\Pi\,\mathbb{B}_1 \,:=\, -\,i'\,\mathcal{R}_1^{-1} : X_1 \to \mathbb{V}'\,,$$

where $\Pi : M_1' \to \mathbb{V}'$ is defined by $\Pi(\mathbb{G}) = \mathbb{G}|_{\mathbb{V}}$ for all $\mathbb{G} \in M_1'$, i' is the adjoint of the canonical injection $i : \mathbb{V} \to X_1$, and $\mathcal{R}_1 : X_1' \to X_1$ is the Riesz operator. Thus, given $\mathbb{F} \in \mathbb{V}'$, the fact that $\|\cdot\|_{\mathbf{div};\Omega}$ and $\|\cdot\|_{0,\Omega}$ coincide in $\mathbb{V}$ says that actually $\mathbb{F}$ is bounded with respect to the $\mathbb{L}^2(\Omega)$-norm. Hence, we let $\mathcal{F} \in X_1'$ be any extension (by Riesz or Hahn-Banach) of $\mathbb{F}$, define $\mathbf{r} \,:=\, -\mathcal{R}_1(\mathcal{F}) \in X_1$, and observe that $-\,i'\,\mathcal{R}_1^{-1}(\mathbf{r}) \,=\, \mathbb{F}$, which shows the surjectivity of $\Pi\,\mathbb{B}_1$. Note again that the inclusion $\mathbb{V} \subseteq X_1$ and the divergence-free property of the elements in $\mathbb{V}$ are crucial here.

Having proved the above results, the well-posedness of (2.8) can be established next.

THEOREM 2.4 *There exists a unique* $(\mathbf{t},\boldsymbol{\sigma},(\mathbf{u},\boldsymbol{\gamma})) \in \mathbf{X} := X_1 \times M_1 \times M$ *solution of problem* (2.8). *Moreover, there exists* $C > 0$, *independent of the solution and the data, such that*

$$\|(\mathbf{t},\boldsymbol{\sigma},(\mathbf{u},\boldsymbol{\gamma}))\|_{\mathbf{X}} \,\le\, C\Big\{\, \|\mathbf{f}\|_{0,\Omega} \,+\, \|\mathbf{g}\|_{1/2,\Gamma}\,\Big\}\,.$$

Proof. It follows from a straightforward application of Theorem 2.1, taking also into account that $\mathbb{A}_1(\mathbf{0})$ becomes the null functional $\mathbf{0}$. □

2.5 The discrete formulation and its analysis

We now let $X_{1,h}$, $M_{1,h}$ and $M_h \,:=\, M_h^{\mathbf{u}} \times M_h^{\boldsymbol{\gamma}}$ be arbitrary finite dimensional subspaces of X_1, M_1 and $M \,:=\, \mathbf{L}^2(\Omega) \times \mathbb{L}^2_{\mathtt{skew}}(\Omega)$, respectively. Then the Galerkin scheme associated with (2.8) reads as follows: Find $(\mathbf{t}_h,\boldsymbol{\sigma}_h,(\mathbf{u}_h,\boldsymbol{\gamma}_h)) \in X_{1,h} \times M_{1,h} \times M_h$ such that

$$\begin{array}{llll} [\mathbb{A}_1(\mathbf{t}_h),\mathbf{s}_h] \;+\; [\mathbb{B}_1(\mathbf{s}_h),\boldsymbol{\sigma}_h] & & = [\mathbb{H},\mathbf{s}_h] & \forall\,\mathbf{s}_h \in X_{1,h}\,,\\ [\mathbb{B}_1(\mathbf{t}_h),\boldsymbol{\tau}_h] & +\; [\mathbb{B}(\boldsymbol{\tau}_h),(\mathbf{u}_h,\boldsymbol{\gamma}_h)] & = [\mathbb{G},\boldsymbol{\tau}_h] & \forall\,\boldsymbol{\tau}_h \in M_{1,h}\,,\\ \qquad [\mathbb{B}(\boldsymbol{\sigma}_h),(\mathbf{v}_h,\boldsymbol{\eta}_h)] & & = [\mathbb{F},(\mathbf{v}_h,\boldsymbol{\eta}_h)] & \forall\,(\mathbf{v}_h,\boldsymbol{\eta}_h) \in M_h\,. \end{array} \tag{2.27}$$

In order to define specific finite element subspaces $X_{1,h}$, $M_{1,h}$ and M_h satisfying the hypotheses of the abstract Theorem 2.2, our strategy is to follow/adapt as much as possible the continuous analysis from Section 2.4. To this respect, we first observe, thanks to Lemma 2.2, that the strong monotonicity and Lipschitz-continuity properties concerning the operator $\mathbb{A}_{1,h} := p_h'\,\mathbb{A}_1 : X_1 \to X_{1,h}$ (p_h being the canonical injection from $X_{1,h}$ to X_1), which constitute the assumption i) of Theorem 2.2, are satisfied for any finite dimensional subspace $X_{1,h}$, and with the same constants γ_1 and α_1 from Lemma 2.1 (cf. (2.21)). In particular, it is easy to see that the simple computation given in (2.25) also works for $\Pi_{1,h}\,\mathbb{A}_{1,h}$ instead of $\Pi_1\,\mathbb{A}_1$, where $\Pi_{1,h}$ is defined as in Theorem 2.2.

Now, given a pair of finite element subspaces $M_{1,h}$ and M_h satisfying the discrete inf-sup condition (2.14) for $\mathbb{B}$ uniformly (which means that there exists $\tilde{\beta} > 0$ such that $\beta_h \,\ge\, \tilde{\beta}$ for all $h > 0$), we let $\mathbb{V}_h$ be the discrete kernel of $\mathbb{B}$, that is

$$\mathbb{V}_h \,:=\, \Big\{\,\boldsymbol{\tau}_h \in M_{1,h} : \quad [\mathbb{B}(\boldsymbol{\tau}_h),(\mathbf{v}_h,\boldsymbol{\eta}_h)] = 0 \qquad \forall\,(\mathbf{v}_h,\boldsymbol{\eta}_h) \,\in\, M_h\,\Big\}\,.$$

which, according to the definition of $\mathbb{B}$ (cf. (2.7)), reduces to

$$\mathbb{V}_h := \Big\{\boldsymbol{\tau}_h \in M_{1,h} : \quad \int_\Omega \mathbf{v}_h\cdot\mathbf{div}\,\boldsymbol{\tau}_h = 0 \quad \forall\,\mathbf{v}_h \in M_h^{\mathbf{u}} \quad\text{and}\quad \int_\Omega \boldsymbol{\tau}_h : \boldsymbol{\eta}_h = 0 \quad \forall\,\boldsymbol{\eta}_h \in M_h^{\boldsymbol{\gamma}}\Big\}\,. \tag{2.28}$$

Next, in order to be able to apply analogous arguments to those employed in the previous section (cf. (2.26)) to conclude the discrete inf-sup condition (2.15) for $\mathbb{B}_1$ uniformly, we just need to assume that

$$\mathbb{V}_h \subseteq X_{1,h} \qquad \text{and} \qquad \mathbf{div}\,\boldsymbol{\tau}_h = \mathbf{0} \quad \forall\, \boldsymbol{\tau}_h \in \mathbb{V}_h\,. \tag{2.29}$$

In particular, note from the first identity defining $\mathbb{V}_h$ (cf. (2.28)) that a sufficient condition for the second requirement in (2.29) is that $\mathbf{div}\big(M_{1,h}\big) \subseteq M_h^{\mathbf{u}}$.

The above analysis and the abstract Theorem 2.2 induce the following general result.

THEOREM 2.5 *Let $X_{1,h}$, $M_{1,h}$ and $M_h := M_h^{\mathbf{u}} \times M_h^{\gamma}$ be finite dimensional subspaces of X_1, M_1 and $M := \mathbf{L}^2(\Omega) \times \mathbb{L}^2_{\mathtt{skew}}(\Omega)$, respectively, and let $\mathbb{V}_h$ be the associated discrete kernel of $\mathbb{B}$ (as defined by (2.28)). Assume that:*

(H.1) *$M_{1,h}$ and M_h satisfy the discrete inf-sup condition (2.14) for $\mathbb{B}$ uniformly.*

(H.2) *$\mathbb{V}_h$ is contained in $X_{1,h}$.*

(H.3) $\mathbf{div}\,\boldsymbol{\tau}_h = \mathbf{0} \quad \forall\, \boldsymbol{\tau}_h \in \mathbb{V}_h$, *or in particular*

$\widetilde{\text{(H.3)}}$ $\mathbf{div}\big(M_{1,h}\big) \subseteq M_h^{\mathbf{u}}$.

Then, there exists a unique $(\mathbf{t}_h, \boldsymbol{\sigma}_h, (\mathbf{u}_h, \gamma_h)) \in X_{1,h} \times M_{1,h} \times M_h$ solution of (2.27). Moreover, there exist C, $\tilde{C} > 0$, independent of h, such that

$$\|(\mathbf{t}_h, \boldsymbol{\sigma}_h, (\mathbf{u}_h, \gamma_h))\|_{\mathbf{X}} \,\leq\, C \Big\{ \|\mathbb{H}_h\| + \|\mathbb{G}_h\| + \|\mathbb{F}_h\| \Big\}$$

and

$$\|(\mathbf{t}, \boldsymbol{\sigma}, (\mathbf{u}, \gamma)) - (\mathbf{t}_h, \boldsymbol{\sigma}_h, (\mathbf{u}_h, \gamma_h))\|_{\mathbf{X}} \,\leq\, \tilde{C} \inf_{\substack{(\mathbf{s}_h, \boldsymbol{\tau}_h, (\mathbf{v}_h, \boldsymbol{\eta}_h)) \\ \in X_{1,h} \times M_{1,h} \times M_h}} \|(\mathbf{t}, \boldsymbol{\sigma}, (\mathbf{u}, \gamma)) - (\mathbf{s}_h, \boldsymbol{\tau}_h, (\mathbf{v}_h, \boldsymbol{\eta}_h))\|\,,$$

where $\mathbb{H}_h := \mathbb{H}|_{X_{1,h}}$, $\mathbb{G}_h := \mathbb{G}|_{M_{1,h}}$, and $\mathbb{F}_h := \mathbb{F}|_{M_h}$.

2.6 Specific finite element subspaces

In order to provide concrete examples of finite element subspaces satisfying the assumptions of Theorem 2.5, we assume from now on that Γ is a polygonal curve and let $\{\mathcal{T}_h\}_{h>0}$ be a regular family of triangulations of Ω, made up of triangles T of diameter h_T, such that $h := \max\{h_T : T \in \mathcal{T}_h\}$ and $\overline{\Omega} := \bigcup\{T : T \in \mathcal{T}_h\}$. Given an integer $\ell \geq 0$ and a subset S of $\mathbb{R}^2$, we denote by $P_\ell(S)$ and $\tilde{P}_\ell(S)$ the spaces of polynomials defined on S of total degree at most ℓ and equal ℓ, respectively. Then, for each $T \in \mathcal{T}_h$ and for each integer $k \geq 0$ we define the local Raviart-Thomas space of order k (see, e.g. [10], [55])

$$\mathbf{RT}_k(T) \,:=\, \mathbf{P}_k(T) \oplus \tilde{P}_k(T)\,\mathbf{x},$$

where $\mathbf{x} := (x_1, x_2)$ is a generic vector of R^2. Recall here that, according to the notation described in Section 1, $\mathbf{P}_k(T)$ stands for $[P_k(T)]^2$. In addition, we let b_T be the triangle-bubble function defined as the unique polynomial in $P_3(T)$ vanishing on ∂T with $\int_T b_T = 1$. Then, for each $T \in \mathcal{T}_h$ and for each integer $k \geq 0$ we define the local bubble space of order k

$$\mathbf{B}_k(T) \,:=\, \mathbf{curl}^{\mathtt{t}}\big(b_T\, P_k(T)\big)\,,$$

where, given a scalar field v, $\mathbf{curl^t}\, v$ is the vector field

$$\mathbf{curl^t}\, v \;:=\; \left(\frac{\partial v}{\partial x_2} \quad -\frac{\partial v}{\partial x_1}\right). \tag{2.30}$$

Next, given an integer $k \geq 0$, we introduce the finite element subspaces

$$M_{1,h} := \Big\{ \boldsymbol{\tau}_h \in \mathbb{H}(\mathbf{div};\Omega): \quad \boldsymbol{\tau}_h|_T \in [\mathbf{RT}_k(T) \oplus \mathbf{B}_k(T)]^2 \quad \forall\, T \in \mathcal{T}_h \Big\}, \tag{2.31}$$

$$M_h^{\mathbf{u}} \;:=\; \Big\{ \mathbf{v}_h \in \mathbf{L}^2(\Omega): \quad \mathbf{v}_h|_T \;\in\; \mathbf{P}_k(T) \quad \forall\, T \in \mathcal{T}_h \Big\}, \tag{2.32}$$

$$M_h^{\boldsymbol{\gamma}} \;:=\; \Big\{ \boldsymbol{\eta}_h \in \mathbb{L}^2_{\mathtt{skew}}(\Omega) \cap \mathbb{C}(\Omega): \quad \boldsymbol{\eta}_h|_T \;\in\; \mathbb{P}_{k+1}(T) \quad \forall\, T \in \mathcal{T}_h \Big\}, \tag{2.33}$$

and

$$M_h \;:=\; M_h^{\mathbf{u}} \,\times\, M_h^{\boldsymbol{\gamma}}\,.$$

The resulting product space $M_{1,h} \times M_h^{\mathbf{u}} \times M_h^{\boldsymbol{\gamma}}$ (with $k = 0$) corresponds to the classical $PEERS$-space introduced originally in [2] for the linear elasticity problem. Moreover, it was shown in [2, Lemma 4.4] that these particular spaces $M_{1,h}$ and M_h satisfy the discrete inf-sup condition (2.14) uniformly, thus providing one of the first stable Galerkin schemes for the mixed variational formulation of the elasticity problem with weak symmetry. In turn, the general case $k \geq 0$ corresponds to the $PEERS$-space of order k introduced in [50], which is denoted $PEERS_k \;:=\; M_{1,h} \times M_h^{\mathbf{u}} \times M_h^{\boldsymbol{\gamma}}$. More precisely, it is shown in [50, Theorem 4.5 and Section 5], as a simple corollary of the corresponding stability result for the BDMS element, that $PEERS_k$ also satisfies the discrete inf-sup condition (2.14) uniformly.

Hence, knowing that the spaces defined by (2.31), (2.32), and (2.33) satisfy the hypothesis (H.1) of Theorem 2.5 for each integer $k \geq 0$, we now aim to prove that, defining a suitable subspace $X_{1,h}$, they all verify (H.2) and (H.3) as well. In fact, it is quite straightforward to see that in this case there holds $\mathbf{div}\big(M_{1,h}\big) \subseteq M_h^{\mathbf{u}}$, which certainly implies that $\mathbf{div}\,\boldsymbol{\tau}_h \;=\; \mathbf{0} \quad \forall\, \boldsymbol{\tau}_h \;\in\; \mathbb{V}_h$, thus satisfying (H.3), where $\mathbb{V}_h$ is the discrete kernel of $\mathbb{B}$ defined according to (2.28), (2.31), (2.32), and (2.33). Moreover, given $\boldsymbol{\tau}_h \in \mathbb{V}_h \subseteq M_{1,h}$, $T \in \mathcal{T}_h$, and $i \in \{1, 2\}$, there exist $\mathbf{q} \in \mathbf{P}_k(T)$, $\tilde{q} \in \tilde{P}_k(T)$, and $\mathbf{b} \in \mathbf{B}_k(T)$, such that, denoting by $\boldsymbol{\tau}_{h,i}$ the i-th row of $\boldsymbol{\tau}_h$,

$$\boldsymbol{\tau}_{h,i} \;=\; \mathbf{q} \,+\, \tilde{q}\,\mathbf{x} \,+\, \mathbf{b} \quad \text{in} \quad T\,.$$

It follows, performing simple algebraic computations, that

$$0 \;=\; \operatorname{div}\boldsymbol{\tau}_{h,i} \;=\; \operatorname{div}\mathbf{q} \,+\, (k+2)\,\tilde{q} \quad \text{in} \quad T\,,$$

which yields $\tilde{q} \;=\; 0$ since $\operatorname{div}\mathbf{q} \;=\; 0$ for $k = 0$, and for $k \geq 1$ there also holds $\tilde{q} \;=\; 0$ since otherwise $\tilde{q} \;=\; -\dfrac{1}{(k+2)}\operatorname{div}\mathbf{q} \;\in\; P_{k-1}(T)$, which contradicts the fact that $\tilde{q} \in \tilde{P}_k(T)$. In this way, we actually have that $\boldsymbol{\tau}_{h,i} \;=\; \mathbf{q} \,+\, \mathbf{b} \quad \text{in} \quad T$, from which we conclude that

$$\boldsymbol{\tau}_h|_T \;\in\; [\mathbf{P}_k(T) \,\oplus\, \mathbf{B}_k(T)]^2 \qquad \forall\, T \,\in\, \mathcal{T}_h\,.$$

Therefore, in order to accomplish (H.2), the above suggests to simply define for each $k \geq 0$

$$X_{1,h} := \Big\{ \mathbf{s}_h \in \mathbb{L}^2(\Omega): \quad \mathbf{s}_h|_T \in [\mathbf{P}_k(T) \oplus \mathbf{B}_k(T)]^2 \quad \forall\, T \in \mathcal{T}_h \Big\}. \tag{2.34}$$

We have thus demonstrated the following theorem, which extends to the case $k \geq 1$ the corresponding result provided in [5, Theorem 5.1].

THEOREM 2.6 *Given an integer $k \geq 0$, we let $X_{1,h}$, $M_{1,h}$ and $M_h := M_h^{\mathbf{u}} \times M_h^{\gamma}$ be the finite element subspaces defined by* (2.34)*,* (2.31)*,* (2.32)*, and* (2.33)*, respectively. Then, there exists a unique $(\mathbf{t}_h, \boldsymbol{\sigma}_h, (\mathbf{u}_h, \boldsymbol{\gamma}_h)) \in X_{1,h} \times M_{1,h} \times M_h$ solution of* (2.27)*. Moreover, there exist $C, \tilde{C} > 0$, independent of h, such that*

$$\|(\mathbf{t}_h, \boldsymbol{\sigma}_h, (\mathbf{u}_h, \boldsymbol{\gamma}_h))\|_{\mathbf{X}} \leq C \left\{ \|\mathbf{f}\|_{0,\Omega} + \|\mathbf{g}\|_{1/2,\Gamma} \right\}$$

and

$$\|(\mathbf{t}, \boldsymbol{\sigma}, (\mathbf{u}, \boldsymbol{\gamma})) - (\mathbf{t}_h, \boldsymbol{\sigma}_h, (\mathbf{u}_h, \boldsymbol{\gamma}_h))\|_{\mathbf{X}} \leq \tilde{C} \inf_{\substack{(\mathbf{s}_h, \boldsymbol{\tau}_h, (\mathbf{v}_h, \boldsymbol{\eta}_h)) \\ \in X_{1,h} \times M_{1,h} \times M_h}} \|(\mathbf{t}, \boldsymbol{\sigma}, (\mathbf{u}, \boldsymbol{\gamma})) - (\mathbf{s}_h, \boldsymbol{\tau}_h, (\mathbf{v}_h, \boldsymbol{\eta}_h))\| . \quad (2.35)$$

Proof. It is a straightforward application of Theorem 2.5 and the fact that the discrete functionals are bounded by the data as indicated here. □

Furthermore, in order to establish the rate of convergence of the Galerkin solution provided by Theorem 2.6, we need the approximation properties of the finite element subspaces involved. For this purpose, we first define the global Raviart-Thomas, bubble, and piecewise polynomial spaces, all of order $k \geq 0$, as

$$\mathbb{RT}_k(\mathcal{T}_h) := \left\{ \boldsymbol{\tau}_h \in \mathbb{H}(\mathbf{div}; \Omega) : \quad \boldsymbol{\tau}_h|_T \in [\mathbf{RT}_k(T)]^2 \quad \forall T \in \mathcal{T}_h \right\},$$

$$\mathbb{B}_k(\mathcal{T}_h) := \left\{ \boldsymbol{\tau}_h \in \mathbb{H}(\mathbf{div}; \Omega) : \quad \boldsymbol{\tau}_h|_T \in [\mathbf{B}_k(T)]^2 \quad \forall T \in \mathcal{T}_h \right\},$$

and

$$P_k(\mathcal{T}_h) := \left\{ v_h \in L^2(\Omega) : \quad v_h|_T \in P_k(T) \quad \forall T \in \mathcal{T}_h \right\}.$$

Note, in particular, that the finite element subspaces $X_{1,h}$ (cf. (2.34)), $M_{1,h}$ (cf. (2.31)), and $M_h^{\mathbf{u}}$ (cf. (2.32)) can also be defined, for each $k \geq 0$, as

$$X_{1,h} = [P_k(\mathcal{T}_h)]^{2\times 2} \oplus \mathbb{B}_k(\mathcal{T}_h), \quad M_{1,h} := \mathbb{RT}_k(\mathcal{T}_h) \oplus \mathbb{B}_k(\mathcal{T}_h), \quad \text{and} \quad M_h^{\mathbf{u}} := [P_k(\mathcal{T}_h)]^2 .$$

Now, we let $\mathcal{E}_h^k : \mathbb{H}^1(\Omega) \longrightarrow \mathbb{RT}_k(\mathcal{T}_h)$ be the usual equilibrium interpolation operator (see, e.g. [55], [10]), which, given $\boldsymbol{\tau} \in \mathbb{H}^1(\Omega)$, is characterized by the following identities:

$$\int_e \mathcal{E}_h^k(\boldsymbol{\tau})\boldsymbol{\nu} \cdot \boldsymbol{\psi} = \int_e \boldsymbol{\tau}\boldsymbol{\nu} \cdot \boldsymbol{\psi} \qquad \forall \text{ edge } e \in \mathcal{T}_h , \quad \forall\, \boldsymbol{\psi} \in \mathbf{P}_k(e), \quad \text{when} \quad k \geq 0 , \quad (2.36)$$

and

$$\int_T \mathcal{E}_h^k(\boldsymbol{\tau}) : \boldsymbol{\psi} = \int_T \boldsymbol{\tau} : \boldsymbol{\psi} \qquad \forall\, T \in \mathcal{T}_h , \quad \forall\, \boldsymbol{\psi} \in \mathbb{P}_{k-1}(T), \quad \text{when} \quad k \geq 1 . \quad (2.37)$$

It is easy to show, using (2.36) and (2.37), that

$$\mathbf{div}(\mathcal{E}_h^k(\boldsymbol{\tau})) = \mathcal{P}_h^k(\mathbf{div}(\boldsymbol{\tau})), \quad (2.38)$$

where $\mathcal{P}_h^k$ is the orthogonal projector from $\mathbf{L}^2(\Omega)$ into $[P_k(\mathcal{T}_h)]^2$. Note that $\mathcal{P}_h^k$ can also be identified with $(\mathbf{P}_h^k, \mathbf{P}_h^k)$, where $\mathbf{P}_h^k$ is the orthogonal projector from $L^2(\Omega)$ into $P_k(\mathcal{T}_h)$. It is well known (see, e.g. [16]) that for each $v \in H^m(\Omega)$, with $0 \leq m \leq k+1$, there holds

$$\|v - \mathbf{P}_h^k(v)\|_{0,T} \leq C\, h_T^m\, |v|_{m,T} \qquad \forall T \in \mathcal{T}_h . \quad (2.39)$$

In addition, the operator $\mathcal{E}_h^k$ satisfies the following approximation properties (see, e.g. [10], [55]):

$$\|\boldsymbol{\tau} - \mathcal{E}_h^k(\boldsymbol{\tau})\|_{0,T} \leq C\, h_T^m\, |\boldsymbol{\tau}|_{m,T} \qquad \forall T \in \mathcal{T}_h , \quad (2.40)$$

for each $\boldsymbol{\tau} \in \mathbb{H}^m(\Omega)$, with $1 \leq m \leq k+1$,

$$\|\mathbf{div}(\boldsymbol{\tau} - \mathcal{E}_h^k(\boldsymbol{\tau}))\|_{0,T} \leq C\, h_T^m\, |\mathbf{div}(\boldsymbol{\tau})|_{m,T} \qquad \forall\, T \in \mathcal{T}_h\,, \tag{2.41}$$

for each $\boldsymbol{\tau} \in \mathbb{H}^1(\Omega)$ such that $\mathbf{div}(\boldsymbol{\tau}) \in \mathbf{H}^m(\Omega)$, with $0 \leq m \leq k+1$, and

$$\|\boldsymbol{\tau}\,\boldsymbol{\nu} - \mathcal{E}_h^k(\boldsymbol{\tau})\,\boldsymbol{\nu}\|_{0,e} \leq C\, h_e^{1/2}\, \|\boldsymbol{\tau}\|_{1,T_e} \qquad \forall \text{ edge } e \in \mathcal{T}_h\,, \tag{2.42}$$

for each $\boldsymbol{\tau} \in \mathbb{H}^1(\Omega)$, where $T_e \in \mathcal{T}_h$ contains e on its boundary. In particular, note that (2.41) follows easily from (2.38) and (2.39). Moreover, it turns out (see, e.g. Theorem 3.16 in [48]) that $\mathcal{E}_h^k$ can also be defined as a bounded linear operator from the larger space $\mathbb{H}^\delta(\Omega) \cap \mathbb{H}(\mathbf{div};\Omega)$ into $\mathbb{RT}_k(\mathcal{T}_h)$ for all $\delta \in (0,1]$. Furthermore, it is easy to show, using the well-known Bramble-Hilbert Lemma and the boundedness of the local interpolation operators on the reference element $\widehat{T}$ (see, e.g. [48, equation (3.39)]), that in this case there holds the following interpolation error estimate

$$\|\boldsymbol{\tau} - \mathcal{E}_h^k(\boldsymbol{\tau})\|_{0,T} \leq C\, h_T^\delta \Big\{ \|\boldsymbol{\tau}\|_{\delta,T} + \|\mathbf{div}(\boldsymbol{\tau})\|_{0,T} \Big\} \qquad \forall\, T \in \mathcal{T}_h\,. \tag{2.43}$$

Then, as a consequence of (2.39), (2.40), (2.41), (2.42), (2.43), and the usual interpolation estimates, we find that the finite element subspaces $X_{1,h}$, $M_{1,h}$, and $M_h^{\mathbf{u}}$ given by (2.34), (2.31), and (2.32) for $k \geq 0$, satisfy the following approximation properties:

$(\mathrm{AP}_{1,h}^{\mathbf{t}})$ *For each* $\delta \in [0,k+1]$ *and for each* $\mathbf{s} \in \mathbb{H}^\delta(\Omega)$ *there exists* $\mathbf{s}_h \in X_{1,h}$ *such that*

$$\|\mathbf{s} - \mathbf{s}_h\|_{0,\Omega} \leq C\, h^\delta\, \|\mathbf{s}\|_{\delta,\Omega}\,.$$

$(\mathrm{AP}_{1,h}^{\boldsymbol{\sigma}})$ *For each* $\delta \in (0,k+1]$ *and for each* $\boldsymbol{\tau} \in \mathbb{H}^\delta(\Omega) \cap \mathbb{H}(\mathbf{div};\Omega)$ *with* $\mathbf{div}\,\boldsymbol{\tau} \in \mathbf{H}^\delta(\Omega)$ *there exists* $\boldsymbol{\tau}_h \in M_{1,h}$ *such that*

$$\|\boldsymbol{\tau} - \boldsymbol{\tau}_h\|_{\mathrm{div},\Omega} \leq C\, h^\delta \Big\{ \|\boldsymbol{\tau}\|_{\delta,\Omega} + \|\mathbf{div}\,\boldsymbol{\tau}\|_{\delta,\Omega} \Big\}\,.$$

$(\mathrm{AP}_h^{\mathbf{u}})$ *For each* $\delta \in [0,k+1]$ *and for each* $\mathbf{v} \in \mathbf{H}^\delta(\Omega)$ *there exists* $\mathbf{v}_h \in M_h^{\mathbf{u}}$ *such that*

$$\|\mathbf{v} - \mathbf{v}_h\|_{0,\Omega} \leq C\, h^\delta\, \|\mathbf{v}\|_{\delta,\Omega}\,.$$

In turn, the approximation property of $M_h^{\boldsymbol{\gamma}}$ is given as follows (cf. [10]):

$(\mathrm{AP}_h^{\boldsymbol{\gamma}})$ *For each* $\delta \in [0,k+1]$ *and for each* $\boldsymbol{\eta} \in \mathbb{H}^\delta(\Omega) \cap \mathbb{L}^2_{\mathtt{skew}}(\Omega)$ *there exists* $\boldsymbol{\eta}_h \in M_h^{\boldsymbol{\gamma}}$ *such that*

$$\|\boldsymbol{\eta} - \boldsymbol{\eta}_h\|_{0,\Omega} \leq C\, h^\delta\, \|\boldsymbol{\eta}\|_{\delta,\Omega}\,.$$

The following theorem establishes the corresponding rate of convergence of the Galerkin scheme (2.27).

Theorem 2.7 *Given an integer* $k \geq 0$, *we let* $X_{1,h}$, $M_{1,h}$ *and* $M_h := M_h^{\mathbf{u}} \times M_h^{\boldsymbol{\gamma}}$ *be the finite element subspaces defined by* (2.34), (2.31), (2.32), *and* (2.33), *respectively. Let* $(\mathbf{t},\boldsymbol{\sigma},(\mathbf{u},\boldsymbol{\gamma})) \in \mathbf{X} := X_1 \times M_1 \times M$ *and* $(\mathbf{t}_h,\boldsymbol{\sigma}_h,(\mathbf{u}_h,\boldsymbol{\gamma}_h)) \in X_{1,h} \times M_{1,h} \times M_h$ *be the unique solutions of the continuous and discrete formulations* (2.8) *and* (2.27), *respectively. Assume that* $\mathbf{t} \in \mathbb{H}^\delta(\Omega)$, $\boldsymbol{\sigma} \in \mathbb{H}^\delta(\Omega)$, $\mathbf{div}\,\boldsymbol{\sigma} \in \mathbf{H}^\delta(\Omega)$, $\mathbf{u} \in \mathbf{H}^\delta(\Omega)$ *and* $\boldsymbol{\gamma} \in \mathbb{H}^\delta(\Omega)$, *for some* $\delta \in (0,k+1]$. *Then there exists* $C > 0$, *independent of* h, *such that*

$$\begin{aligned} &\|(\mathbf{t},\boldsymbol{\sigma},(\mathbf{u},\boldsymbol{\gamma})) - (\mathbf{t}_h,\boldsymbol{\sigma}_h,(\mathbf{u}_h,\boldsymbol{\gamma}_h))\|_{\mathbf{X}} \\ &\qquad \leq C\, h^\delta \Big\{ \|\mathbf{t}\|_{\delta,\Omega} + \|\boldsymbol{\sigma}\|_{\delta,\Omega} + \|\mathbf{div}\,\boldsymbol{\sigma}\|_{\delta,\Omega} + \|\mathbf{u}\|_{\delta,\Omega} + \|\boldsymbol{\gamma}\|_{\delta,\Omega} \Big\}\,. \end{aligned} \tag{2.44}$$

Proof. It follows from the Céa estimate (2.35) and the above approximation properties. □

3 The augmented variational formulation

In this section we propose an augmented formulation for (2.8) and a corresponding discrete scheme whose main advantage is the elimination of the assumption (H.2) in Theorem 2.5, which means that the discrete inf-sup condition for $\mathbb{B}_1$ is no longer required. More precisely, we show that a suitable enrichment of (2.8) yields an associated Galerkin scheme whose well-posedness is guaranteed by any finite dimensional subspace $X_{1,h}$ of X_1 and any pair $(M_{1,h}, M_h)$ satisfying (H.1) and (H.3) only. In particular, the eventual need of approximating $\mathbf{t}$ by either continuous or discontinuous piecewise polynomial tensors of any degree can be satisfied with this approach.

3.1 The continuous augmented formulation

As mentioned above, we now enrich the formulation (2.8) with the further introduction of the constitutive law relating $\boldsymbol{\sigma}$ and $\mathbf{t}$ (written as in the second equation of (2.3)) multiplied by a stabilization parameter. More precisely, given $\kappa_0 > 0$, to be chosen later, we add

$$\kappa_0 \int_\Omega \Big(\boldsymbol{\sigma} - \Big\{\boldsymbol{\lambda}(\mathbf{t})\,\mathrm{tr}(\mathbf{t})\,\mathbb{I} \,+\, \boldsymbol{\mu}(\mathbf{t})\,\mathbf{t}\Big\}\Big) : \boldsymbol{\tau} \;=\; 0 \qquad \forall\, \boldsymbol{\tau} \,\in\, \mathbb{H}(\mathbf{div};\Omega)$$

to the first equation of (2.8), and subtract the second equation of (2.8) to the resulting expression. In addition, we keep the third equation as it is, but multiplied by -1. In this way, denoting the product space $X := X_1 \times M_1$, we arrive at the following augmented formulation (written as a single saddle point system): Find $((\mathbf{t},\boldsymbol{\sigma}),(\mathbf{u},\boldsymbol{\gamma})) \in X \times M$ such that

$$\begin{aligned} [\mathcal{A}(\mathbf{t},\boldsymbol{\sigma}),(\mathbf{s},\boldsymbol{\tau})] \,+\, [\mathcal{B}(\mathbf{s},\boldsymbol{\tau}),(\mathbf{u},\boldsymbol{\gamma})] \;&=\; [\mathcal{F},(\mathbf{s},\boldsymbol{\tau})] \quad \forall\,(\mathbf{s},\boldsymbol{\tau}) \in X\,,\\ [\mathcal{B}(\mathbf{t},\boldsymbol{\sigma}),(\mathbf{v},\boldsymbol{\eta})] \;&=\; [\mathcal{G},(\mathbf{v},\boldsymbol{\eta})] \quad \forall\,(\mathbf{v},\boldsymbol{\eta}) \in M\,, \end{aligned} \tag{3.1}$$

where the nonlinear operator $\mathcal{A} : X \to X'$, the linear operator $\mathcal{B} : X \to M'$, and the functionals $\mathcal{F} \in X'$ and $\mathcal{G} \in M'$, are defined by:

$$[\mathcal{A}(\mathbf{t},\boldsymbol{\sigma}),(\mathbf{s},\boldsymbol{\tau})] \,:=\, [\mathbb{A}_1(\mathbf{t}),\mathbf{s}] + [\mathbb{B}_1(\mathbf{s}),\boldsymbol{\sigma}] - [\mathbb{B}_1(\mathbf{t}),\boldsymbol{\tau}] + \kappa_0 \int_\Omega \Big(\boldsymbol{\sigma} - \Big\{\boldsymbol{\lambda}(\mathbf{t})\,\mathrm{tr}(\mathbf{t})\,\mathbb{I} + \boldsymbol{\mu}(\mathbf{t})\,\mathbf{t}\Big\}\Big) : \boldsymbol{\tau}\,, \tag{3.2}$$

$$[\mathcal{B}(\mathbf{s},\boldsymbol{\tau}),(\mathbf{v},\boldsymbol{\eta})] \,:=\, -\,[\mathbb{B}(\boldsymbol{\tau}),(\mathbf{v},\boldsymbol{\eta})] \,=\, \int_\Omega \mathbf{v}\cdot\mathbf{div}\,\boldsymbol{\tau} + \int_\Omega \boldsymbol{\tau} \,:\, \boldsymbol{\eta}\,, \tag{3.3}$$

$$[\mathcal{F},(\mathbf{s},\boldsymbol{\tau})] \,:=\, [\mathbb{H},\mathbf{s}] \,-\, [\mathbb{G},\boldsymbol{\tau}] \,=\, \langle \boldsymbol{\tau}\,\boldsymbol{\nu},\mathbf{g}\rangle_\Gamma\,, \tag{3.4}$$

and

$$[\mathcal{G},\mathbf{v}] \,:=\, -\,[\mathbb{F},\mathbf{v}] \,=\, -\int_\Omega \mathbf{f}\cdot\mathbf{v}\,. \tag{3.5}$$

Our next goal is to show the unique solvability of the variational formulation (3.1), whence (2.8) and (3.1) share the same unique solution. We first recall from [57] the following abstract theorem.

Theorem 3.1 *Let X, M be Hilbert spaces and let $\mathcal{A} : X \to X'$ and $\mathcal{B} : X \to M'$ be nonlinear and linear operators, respectively. Let $V := \mathrm{Ker}\,(\mathcal{B}) = \{\, x \in X : \;\; [\mathcal{B}(x), q] = 0 \;\; \forall\, q \in M\,\}$. Assume that $\mathcal{A}$ is Lipschitz-continuous on X and that for all $\tilde{z} \in X$, $\mathcal{A}(\tilde{z} + \cdot)$ is uniformly strongly monotone on V, that is, there exist constants γ, $\alpha > 0$ such that*

$$||\mathcal{A}(x) \,-\, \mathcal{A}(y)||_{X'} \,\le\, \gamma\,\|x - y\|_X \quad \forall\, x,\, y \in X\,,$$

and

$$[\mathcal{A}(\tilde{z}+x)-\mathcal{A}(\tilde{z}+y), x-y] \;\geq\; \alpha\,\|x-y\|_X^2$$

for all $\tilde{z}\in X$ *and for all* $x,\, y\in V$. *In addition, assume that there exists* $\beta>0$ *such that for all* $q\in M$

$$\sup_{x\in X\backslash\{\mathbf{0}\}} \frac{[\mathcal{B}(x), q]}{\|x\|_X} \;\geq\; \beta\,\|q\|_M\,.$$

Then, given $(\mathcal{F},\mathcal{G})\in X'\times M'$, *there exists a unique* $(x,p)\in X\times M$ *such that*

$$\begin{aligned}
[\mathcal{A}(x), y] \,+\, [\mathcal{B}(y), p] &\;=\; [\mathcal{F}, y] \quad \forall\, y\in X\,,\\
[\mathcal{B}(x), q] &\;=\; [\mathcal{G}, q] \quad \forall\, q\in M\,.
\end{aligned}$$

Further, the following estimates hold

$$\|x\|_X \;\leq\; \frac{1}{\alpha}\,\|\mathcal{F}\| + \frac{1}{\beta}\left(1+\frac{\gamma}{\alpha}\right)\,\|\mathcal{G}\|\,, \tag{3.6}$$

$$\|p\|_M \;\leq\; \frac{1}{\beta}\,\left(1+\frac{\gamma}{\alpha}\right)\,\left(\|\mathcal{F}\| \,+\, \frac{\gamma}{\beta}\|\mathcal{G}\|\right)\,. \tag{3.7}$$

Proof. It is a particular case of Proposition 2.3 in [57]. □

The discrete analogue of Theorem 3.1 and the corresponding Céa estimate are provided in [57, Proposition 2.6, Theorem 2.1]. We omit details here.

In order to apply Theorem 3.1 to the augmented formulation (3.1), we need several preliminary results establishing the required properties for our nonlinear operator $\mathcal{A}$ (cf. (3.2)). We begin with the following lemma.

Lemma 3.1 *Let* $\mathcal{A}$ *be the nonlinear operator defined by* (3.2). *Then, there exists a constant* $\gamma>0$ *such that*

$$||\mathcal{A}(\mathbf{t},\boldsymbol{\sigma}) \,-\, \mathcal{A}(\mathbf{s},\boldsymbol{\tau})||_{X'} \;\leq\; \gamma\,\|(\mathbf{t},\boldsymbol{\sigma})-(\mathbf{s},\boldsymbol{\tau})\|_X \qquad \forall\,(\mathbf{t},\boldsymbol{\sigma}),\,(\mathbf{s},\boldsymbol{\tau})\,\in\, X\,. \tag{3.8}$$

Proof. Given $(\mathbf{t},\boldsymbol{\sigma})$, $(\mathbf{s},\boldsymbol{\tau})$, $(\mathbf{r},\boldsymbol{\zeta})\,\in\, X$, we obtain, according to (3.2) and the definition of $\mathbb{A}_1$ (cf. (2.7)), that

$$\begin{aligned}
[\mathcal{A}(\mathbf{t},\boldsymbol{\sigma})-\mathcal{A}(\mathbf{s},\boldsymbol{\tau}), (\mathbf{r},\boldsymbol{\zeta})] \;=\; [\mathbb{A}_1(\mathbf{t})-\mathbb{A}_1(\mathbf{s}), \mathbf{r}] \,+\, [\mathbb{B}_1(\mathbf{r}), \boldsymbol{\sigma}-\boldsymbol{\tau}] \,-\, [\mathbb{B}_1(\mathbf{t}-\mathbf{s}), \boldsymbol{\zeta}]\\
+\;\; \kappa_0 \int_\Omega (\boldsymbol{\sigma}-\boldsymbol{\tau}) : \boldsymbol{\zeta} \;\;-\;\; \kappa_0\,[\mathbb{A}_1(\mathbf{t}) \,-\, \mathbb{A}_1(\mathbf{s}), \boldsymbol{\zeta}]\,,
\end{aligned} \tag{3.9}$$

which, employing Cauchy-Schwarz's inequality, yields

$$\begin{aligned}
\big|\,[\mathcal{A}(\mathbf{t},\boldsymbol{\sigma})-\mathcal{A}(\mathbf{s},\boldsymbol{\tau}), (\mathbf{r},\boldsymbol{\zeta})]\,\big| \;\leq\; ||\mathbb{A}_1(\mathbf{t})-\mathbb{A}_1(\mathbf{s})||_{X_1'}\,||\mathbf{r}||_{X_1} \,+\, ||\mathbb{B}_1(\mathbf{r})||_{M_1'}\,||\boldsymbol{\sigma}-\boldsymbol{\tau}||_{M_1}\\
+\;\; ||\mathbb{B}_1(\mathbf{t}-\mathbf{s})||_{M_1'}\,||\boldsymbol{\zeta}||_{M_1} \,+\, \kappa_0\,||(\boldsymbol{\sigma}-\boldsymbol{\tau})||_{M_1}\,||\boldsymbol{\zeta}||_{M_1} \,+\, \kappa_0\,||\mathbb{A}_1(\mathbf{t})-\mathbb{A}_1(\mathbf{s})||_{X_1'}\,||\boldsymbol{\zeta}||_{M_1}\,.
\end{aligned}$$

Hence, applying the Lipschitz-continuity of $\mathbb{A}_1$ (cf. (2.22) in Lemma 2.2) and the boundedness of $\mathbb{B}_1$, we conclude from the above inequality that $\mathcal{A}$ is Lipschitz continuous on X with a constant γ depending on γ_1, $\|\mathbb{B}_1\|$, and κ_0. □

The following estimate is applied later on to show that $\mathcal{A}$ satisfies the strong monotonicity property.

Lemma 3.2 *Let $\mathcal{A}$ be the operator defined by (3.2) and assume that the parameter κ_0 lies in $\left(0, \frac{2\,\alpha_1}{\gamma_1^2}\right)$, where γ_1 and α_1 are the positive constants from Lemma 2.1 (cf. (2.21)). Then, there exists a constant $\alpha > 0$ such that*

$$[\mathcal{A}((\mathbf{r},\boldsymbol{\zeta}) + (\mathbf{t},\boldsymbol{\sigma})) \,-\, \mathcal{A}((\mathbf{r},\boldsymbol{\zeta}) + (\mathbf{s},\boldsymbol{\tau})), (\mathbf{t},\boldsymbol{\sigma}) \,-\, (\mathbf{s},\boldsymbol{\tau})] \,\geq\, \alpha \Big\{ \|\mathbf{t} \,-\, \mathbf{s}\|_{X_1}^2 \,+\, \|\boldsymbol{\sigma} \,-\, \boldsymbol{\tau}\|_{0,\Omega}^2 \Big\} \tag{3.10}$$

for all $(\mathbf{r},\boldsymbol{\zeta})\,,\ (\mathbf{t},\boldsymbol{\sigma})\,,\ (\mathbf{s},\boldsymbol{\tau}) \,\in\, X$.

Proof. Given $(\mathbf{r},\boldsymbol{\zeta})\,,\ (\mathbf{t},\boldsymbol{\sigma})\,,\ (\mathbf{s},\boldsymbol{\tau}) \,\in\, X$, we find, using the identity (3.9) and noting that the terms involving $\mathbb{B}_1$ cancell out, that

$$\begin{aligned}
[\mathcal{A}((\mathbf{r},\boldsymbol{\zeta}) + (\mathbf{t},\boldsymbol{\sigma})) - \mathcal{A}((\mathbf{r},\boldsymbol{\zeta}) + (\mathbf{s},\boldsymbol{\tau})), (\mathbf{t},\boldsymbol{\sigma}) - (\mathbf{s},\boldsymbol{\tau})] \;&=\; [\mathbb{A}_1(\mathbf{r} + \mathbf{t}) \,-\, \mathbb{A}_1(\mathbf{r} + \mathbf{s}), \mathbf{t} - \mathbf{s}] \\
&+\;\; \kappa_0\, \|\boldsymbol{\sigma} - \boldsymbol{\tau}\|_{0,\Omega}^2 \,-\, \kappa_0\, [\mathbb{A}_1(\mathbf{r} + \mathbf{t}) - \mathbb{A}_1(\mathbf{r} + \mathbf{s}), \boldsymbol{\sigma} - \boldsymbol{\tau}]\,.
\end{aligned}$$

Then, using that $[\mathbb{A}_1(\mathbf{r}+\mathbf{t}) - \mathbb{A}_1(\mathbf{r}+\mathbf{s}), \mathbf{t}-\mathbf{s}] \,=\, [\mathbb{A}_1(\mathbf{r}+\mathbf{t}) - \mathbb{A}_1(\mathbf{r}+\mathbf{s}), (\mathbf{r}+\mathbf{t}) - (\mathbf{r}+\mathbf{s})]$, and applying the strong monotonicity and Lipschitz-continuity of $\mathbb{A}_1$ (cf. Lemma 2.2), we deduce from the above equation that

$$\begin{aligned}
&[\mathcal{A}((\mathbf{r},\boldsymbol{\zeta}) + (\mathbf{t},\boldsymbol{\sigma})) - \mathcal{A}((\mathbf{r},\boldsymbol{\zeta}) + (\mathbf{s},\boldsymbol{\tau})), (\mathbf{t},\boldsymbol{\sigma}) - (\mathbf{s},\boldsymbol{\tau})] \\
&\geq\; \alpha_1\, \|\mathbf{t} - \mathbf{s}\|_{X_1}^2 \,+\, \kappa_0\, \|\boldsymbol{\sigma} - \boldsymbol{\tau}\|_{0,\Omega}^2 \,-\, \kappa_0\, \gamma_1\, \|\mathbf{t} - \mathbf{s}\|_{X_1}\, \|\boldsymbol{\sigma} - \boldsymbol{\tau}\|_{0,\Omega} \\
&\geq\; \alpha_1\, \|\mathbf{t} - \mathbf{s}\|_{X_1}^2 \,+\, \kappa_0\, \|\boldsymbol{\sigma} - \boldsymbol{\tau}\|_{0,\Omega}^2 \,-\, \kappa_0\, \gamma_1 \left\{ \frac{\|\mathbf{t} - \mathbf{s}\|_{X_1}^2}{2\,\delta} \,+\, \frac{\delta}{2}\, \|\boldsymbol{\sigma} - \boldsymbol{\tau}\|_{0,\Omega}^2 \right\} \\
&=\; \left(\alpha_1 - \frac{\kappa_0\, \gamma_1}{2\,\delta}\right) \|\mathbf{t} - \mathbf{s}\|_{X_1}^2 \,+\, \kappa_0 \left(1 \,-\, \frac{\gamma_1\, \delta}{2}\right) \|\boldsymbol{\sigma} - \boldsymbol{\tau}\|_{0,\Omega}^2 \qquad \forall\, \delta > 0\,.
\end{aligned}$$

It follows that the constants multiplying the norms above become positive if $\delta \,\in\, \left(0, \frac{2}{\gamma_1}\right)$ and $\kappa_0 \,\in\, \left(0, \frac{2\,\alpha_1\,\delta}{\gamma_1}\right)$. In particular, for $\delta \,=\, \frac{1}{\gamma_1}$ we require $\kappa_0 \,\in\, \left(0, \frac{2\,\alpha_1}{\gamma_1^2}\right)$, whence we find that

$$\begin{aligned}
&[\mathcal{A}((\mathbf{r},\boldsymbol{\zeta}) + (\mathbf{t},\boldsymbol{\sigma})) - \mathcal{A}((\mathbf{r},\boldsymbol{\zeta}) + (\mathbf{s},\boldsymbol{\tau})), (\mathbf{t},\boldsymbol{\sigma}) - (\mathbf{s},\boldsymbol{\tau})] \\
&\geq\;\; \left(\alpha_1 \,-\, \frac{\kappa_0\, \gamma_1^2}{2}\right) \|\mathbf{t} - \mathbf{s}\|_{X_1}^2 \,+\, \frac{\kappa_0}{2}\, \|\boldsymbol{\sigma} - \boldsymbol{\tau}\|_{0,\Omega}^2\,.
\end{aligned}$$

Finally, this inequality implies the required estimate with $\alpha \,:=\, \min\left\{\alpha_1 \,-\, \frac{\kappa_0\,\gamma_1^2}{2}, \frac{\kappa_0}{2}\right\}$. □

It is quite straightforward from Lemma 3.2 that, defining

$$\widetilde{\mathcal{V}} \,:=\, X_1 \,\times\, \Big\{ \boldsymbol{\tau} \,\in\, M_1 : \quad \mathbf{div}\,\boldsymbol{\tau} \,=\, \mathbf{0} \quad \text{in} \quad \Omega \Big\}\,, \tag{3.11}$$

and assuming again that the parameter κ_0 lies in $\left(0, \frac{2\,\alpha_1}{\gamma_1^2}\right)$, there holds with the same constant α,

$$[\mathcal{A}((\mathbf{r},\boldsymbol{\zeta}) + (\mathbf{t},\boldsymbol{\sigma})) - \mathcal{A}((\mathbf{r},\boldsymbol{\zeta}) + (\mathbf{s},\boldsymbol{\tau})), (\mathbf{t},\boldsymbol{\sigma}) - (\mathbf{s},\boldsymbol{\tau})] \,\geq\, \alpha\, \|(\mathbf{t},\boldsymbol{\sigma}) - (\mathbf{s},\boldsymbol{\tau})\|_X^2 \tag{3.12}$$

for all $(\mathbf{r}, \boldsymbol{\zeta}) \in X$, and for all $(\mathbf{t}, \boldsymbol{\sigma})$, $(\mathbf{s}, \boldsymbol{\tau}) \in \widetilde{\mathcal{V}}$. In particular, noting from the definition of $\mathcal{B}$ (cf. (3.3)) that its kernel $\mathcal{V}$ reduces to

$$\mathcal{V} \,=\, X_1 \times \left\{ \boldsymbol{\tau} \in M_1 : \quad \mathbf{div}\,\boldsymbol{\tau} \,=\, \mathbf{0} \quad \text{and} \quad \boldsymbol{\tau} \,=\, \boldsymbol{\tau}^{\mathbf{t}} \quad \text{in} \quad \Omega \right\},$$

which is certainly contained in $\widetilde{V}$, we find that

$$[\mathcal{A}((\mathbf{r},\boldsymbol{\zeta}) + (\mathbf{t},\boldsymbol{\sigma})) - \mathcal{A}((\mathbf{r},\boldsymbol{\zeta}) + (\mathbf{s},\boldsymbol{\tau})), (\mathbf{t},\boldsymbol{\sigma}) - (\mathbf{s},\boldsymbol{\tau})] \,\geq\, \alpha \, \|(\mathbf{t},\boldsymbol{\sigma}) - (\mathbf{s},\boldsymbol{\tau})\|_X^2$$

for all $(\mathbf{r}, \boldsymbol{\zeta}) \in X$, and for all $(\mathbf{t}, \boldsymbol{\sigma})$, $(\mathbf{s}, \boldsymbol{\tau}) \in \mathcal{V}$.

On the other hand, it is clear from the definition of our linear operator $\mathcal{B}$ (cf. (3.3)) that

$$\sup_{(\mathbf{s},\boldsymbol{\tau}) \in X\backslash\{\mathbf{0}\}} \frac{[\mathcal{B}(\mathbf{s},\boldsymbol{\tau}), (\mathbf{v},\boldsymbol{\eta})]}{\|(\mathbf{s},\boldsymbol{\tau})\|_X} \,=\, \sup_{\boldsymbol{\tau} \in M_1\backslash\{0\}} \frac{[\mathbb{B}(\boldsymbol{\tau}), (\mathbf{v},\boldsymbol{\eta})]}{\|\boldsymbol{\tau}\|_{M_1}} \qquad \forall\,(\mathbf{v},\boldsymbol{\eta}) \,\in\, M\,, \tag{3.13}$$

which implies that the continuous inf-sup conditions for $\mathcal{B}$ and $\mathbb{B}$, the latter already proved in Lemma 2.3, coincide.

Hence, we are ready to establish the well-posedness of our augmented formulation (3.1).

Theorem 3.2 *Assume that the parameter* κ_0 *lies in* $\left(0, \dfrac{2\,\alpha_1}{\gamma_1^2}\right)$, *where* γ_1 *and* α_1 *are the positive constants from Lemma* 2.1. *Then, there exists a unique* $((\mathbf{t},\boldsymbol{\sigma}),(\mathbf{u},\boldsymbol{\gamma})) \,\in\, X \times M$ *solution of* (3.1). *Moreover, there exists* $C > 0$ *such that*

$$\|((\mathbf{t},\boldsymbol{\sigma}),(\mathbf{u},\boldsymbol{\gamma}))\|_{X\times M} \,\leq\, C \left\{ \|\mathbf{f}\|_{0,\Omega} \,+\, \|\mathbf{g}\|_{1/2,\Gamma} \right\}.$$

Proof. By virtue of the previous discussion and the fact that the functionals $\mathcal{F}$ (cf. (3.4)) and $\mathcal{G}$ (cf. (3.5)) are bounded by the data (as indicated here), the proof follows from a straightforward application of Theorem 3.1. □

3.2 The discrete augmented formulation

We now come to the analysis of the Galerkin scheme associated with the augmented formulation (3.1). To this end, we now let $X_{1,h}$, $M_{1,h}$, and $M_h \,:=\, M_h^{\mathbf{u}} \times M_h^{\boldsymbol{\gamma}}$ be finite dimensional subspaces of X_1, M_1, and M, respectively, and define $X_h \,:=\, X_{1,h} \times M_{1,h}$. Then, we are interested in the following discrete scheme: Find $((\mathbf{t}_h,\boldsymbol{\sigma}_h),(\mathbf{u}_h,\boldsymbol{\gamma}_h)) \,\in\, X_h \times M_h$ such that

$$\begin{aligned} [\mathcal{A}(\mathbf{t}_h,\boldsymbol{\sigma}_h),(\mathbf{s},\boldsymbol{\tau})] \,+\, [\mathcal{B}(\mathbf{s},\boldsymbol{\tau}),(\mathbf{u}_h,\boldsymbol{\gamma}_h)] \,&=\, [\mathcal{F},(\mathbf{s},\boldsymbol{\tau})] \quad \forall\,(\mathbf{s},\boldsymbol{\tau}) \in X_h\,, \\ [\mathcal{B}(\mathbf{t}_h,\boldsymbol{\sigma}_h),(\mathbf{v},\boldsymbol{\eta})] \,&=\, [\mathcal{G},(\mathbf{v},\boldsymbol{\eta})] \quad \forall\,(\mathbf{v},\boldsymbol{\eta}) \in M_h\,. \end{aligned} \tag{3.14}$$

In order to analyze the solvability of (3.14), we first notice from (3.3) that the discrete kernel of $\mathcal{B}$, that is $\mathcal{V}_h \,:=\, \left\{ (\mathbf{s}_h,\boldsymbol{\tau}_h) \in X_h : \quad [\mathcal{B}(\mathbf{s}_h,\boldsymbol{\tau}_h),(\mathbf{v}_h,\boldsymbol{\eta}_h)] \,=\, 0 \quad \forall\,(\mathbf{v}_h,\boldsymbol{\eta}_h) \in M_h \right\}$, reduces to

$$\mathcal{V}_h \,=\, X_{1,h} \times \left\{ \boldsymbol{\tau}_h \,\in\, M_{1,h} : \quad \int_\Omega \mathbf{v}_h \cdot \mathbf{div}\,\boldsymbol{\tau}_h \,=\, 0 \quad \text{and} \quad \int_\Omega \boldsymbol{\tau}_h \,:\, \boldsymbol{\eta}_h \,=\, 0 \quad \forall\,(\mathbf{v}_h,\boldsymbol{\eta}_h) \in M_h \right\}.$$

In addition, as in (3.13), we realize that

$$\sup_{(\mathbf{s}_h,\boldsymbol{\tau}_h) \in X_h\backslash\{\mathbf{0}\}} \frac{[\mathcal{B}(\mathbf{s}_h,\boldsymbol{\tau}_h), (\mathbf{v}_h,\boldsymbol{\eta}_h)]}{\|(\mathbf{s}_h,\boldsymbol{\tau}_h)\|_X} \,=\, \sup_{\boldsymbol{\tau}_h \in M_{1,h}\backslash\{0\}} \frac{[\mathbb{B}(\boldsymbol{\tau}_h), (\mathbf{v}_h,\boldsymbol{\eta}_h)]}{\|\boldsymbol{\tau}_h\|_{M_1}} \qquad \forall\,(\mathbf{v}_h,\boldsymbol{\eta}_h) \,\in\, M_h\,,$$

which implies that the discrete inf-sup conditions for $\mathcal{B}$ and $\mathbb{B}$ also coincide.

Hence, we are in a position to establish the following result.

THEOREM 3.3 *Besides the hypotheses of Theorem* 3.2*, assume that* $X_{1,h}$ *is any finite element subspace of* X_1*, that* $\mathbf{div}\,\boldsymbol{\tau}_h = \mathbf{0} \quad \forall\,(\mathbf{s}_h, \boldsymbol{\tau}_h) \in \mathcal{V}_h$ *(equivalently,* $\mathcal{V}_h \subseteq \widetilde{\mathcal{V}}$*), and that* $\mathbb{B}$ *satisfies the discrete inf-sup condition on* $M_{1,h} \times M_h$*, that is there exists* $\tilde{\beta} > 0$*, independent of* h*, such that*

$$\sup_{\boldsymbol{\tau}_h \in M_{1,h}\setminus\{0\}} \frac{[\mathbb{B}(\boldsymbol{\tau}_h), (\mathbf{v}_h, \boldsymbol{\eta}_h)]}{\|\boldsymbol{\tau}_h\|_{M_1}} \,\geq\, \tilde{\beta}\,\|(\mathbf{v}_h, \boldsymbol{\eta}_h)\|_M \qquad \forall\,(\mathbf{v}_h, \boldsymbol{\eta}_h) \,\in\, M_h\,.$$

Then there exists a unique $((\mathbf{t}_h, \boldsymbol{\sigma}_h), (\mathbf{u}_h, \boldsymbol{\gamma}_h)) \in X_h \times M_h$ *solution of* (3.14)*. Moreover, there exist* C_1, $C_2 > 0$*, independent of* h*, such that*

$$\|((\mathbf{t}_h, \boldsymbol{\sigma}_h), (\mathbf{u}_h, \boldsymbol{\gamma}_h))\|_{X\times M} \,\leq\, C_1 \left\{ \|\mathbf{f}\|_{0,\Omega} \,+\, \|\mathbf{g}\|_{1/2,\Gamma} \right\}, \tag{3.15}$$

and

$$\begin{aligned} &\|((\mathbf{t}, \boldsymbol{\sigma}), (\mathbf{u}, \boldsymbol{\gamma})) - ((\mathbf{t}_h, \boldsymbol{\sigma}_h), ((\mathbf{u}_h, \boldsymbol{\gamma}_h))\|_{X\times M} \\ &\leq\quad C_2 \left\{ \inf_{(\mathbf{s}_h, \boldsymbol{\tau}_h)\in X_h} \|(\mathbf{t}, \boldsymbol{\sigma}) - (\mathbf{s}_h, \boldsymbol{\tau}_h)\|_X \,+\, \inf_{(\mathbf{v}_h, \boldsymbol{\eta}_h)\in M_h} \|(\mathbf{u}, \boldsymbol{\gamma}) - (\mathbf{v}_h, \boldsymbol{\eta}_h)\|_M \right\}. \end{aligned} \tag{3.16}$$

Proof. It is clear that the Lipschitz-continuity of $\mathcal{A}$ (cf. Lemma 3.1) is also valid on $X_h \times X_h'$, which means that, with the same constant γ from Lemma 3.1, there holds

$$||\mathcal{A}(\mathbf{t}_h, \boldsymbol{\sigma}_h) \,-\, \mathcal{A}(\mathbf{s}_h, \boldsymbol{\tau}_h)||_{X_h'} \,\leq\, \gamma\,\|(\mathbf{t}_h, \boldsymbol{\sigma}_h) - (\mathbf{s}_h, \boldsymbol{\tau}_h)\|_X \qquad \forall\,(\mathbf{t}_h, \boldsymbol{\sigma}_h),\, (\mathbf{s}_h, \boldsymbol{\tau}_h) \,\in\, X_h\,.$$

Furthermore, since $\mathcal{V}_h \subseteq \widetilde{\mathcal{V}}$ (cf. (3.11)), the strong monotonicity of $\mathcal{A}$ provided by (3.12) also holds for all $(\mathbf{r}_h, \boldsymbol{\zeta}_h) \in X_h$, and for all $(\mathbf{t}_h, \boldsymbol{\sigma}_h)$, $(\mathbf{s}_h, \boldsymbol{\tau}_h) \in \mathcal{V}_h$. Therefore, the unique solvability of (3.14) and the estimate (3.15) are again consequence of Theorem 3.1 (see also the discrete analogue given by [57, Proposition 2.6]). Finally, the Céa estimate (3.16) constitutes a particular application of the general result given by [57, Theorem 2.1]. We omit further details. □

It is important to notice that, on the contrary to the condition (H.2) in Theorem 2.5, the well-posedness of the present discrete augmented scheme (3.14) does not require any additional restriction on $X_{1,h}$, but being only a finite dimensional subspace of X_1. Furthermore, as established by the hypothesis $\widetilde{\text{(H.3)}}$ in Theorem 2.5, we recall that a sufficient condition for $\mathcal{V}_h \subseteq \widetilde{\mathcal{V}}$ to hold is that $\mathbf{div}(M_{1,h}) \subseteq M_h^{\mathbf{u}}$. Finally, we remark that, though the unique solutions of the discrete schemes (2.13) and (3.14) are denoted in the same way, they do not necessarily coincide.

3.3 Specific finite element subspaces

We now provide several examples of subspaces verifying the hypotheses of Theorem 3.3. First of all, it is quite clear from the analysis in Section 2.6 that, given an integer $k \geq 0$, the subspaces $M_{1,h}$ and $M_h := M_h^{\mathbf{u}} \times M_h^{\boldsymbol{\gamma}}$ defined by (2.31), (2.32), and (2.33), and the resulting discrete kernel $\mathcal{V}_h$ of $\mathcal{B}$ (irrespective of the chosen subspace $X_{1,h}$), satisfy the corresponding assumptions in Theorem 3.3. Consequently, and since any finite element subspace $X_{1,h}$ of X_1 will yield a well-posed discrete augmented scheme (3.14), we can establish the following result.

THEOREM 3.4 *Besides the hypotheses of Theorem* 3.2*, assume that* $X_{1,h}$ *is any finite element subspace of* X_1*, and that given an integer* $k \geq 0$*,* $M_{1,h}$ *and* $M_h := M_h^{\mathbf{u}} \times M_h^{\boldsymbol{\gamma}}$ *are defined by* (2.31)*,* (2.32)*, and* (2.33)*, respectively. Then there exists a unique* $((\mathbf{t}_h, \boldsymbol{\sigma}_h), (\mathbf{u}_h, \boldsymbol{\gamma}_h)) \in X_h \times M_h$ *solution of* (3.14)*. Moreover, there exist* C_1, $C_2 > 0$*, independent of* h*, such that*

$$\|((\mathbf{t}_h, \boldsymbol{\sigma}_h), (\mathbf{u}_h, \boldsymbol{\gamma}_h))\|_{X\times M} \,\leq\, C_1 \left\{ \|\mathbf{f}\|_{0,\Omega} \,+\, \|\mathbf{g}\|_{1/2,\Gamma} \right\},$$

and

$$\begin{aligned} &\|((\mathbf{t},\boldsymbol{\sigma}),(\mathbf{u},\boldsymbol{\gamma})) - ((\mathbf{t}_h,\boldsymbol{\sigma}_h),(\mathbf{u}_h,\boldsymbol{\gamma}_h))\|_{X\times M} \\ &\quad\leq\quad C_2\left\{\inf_{(\mathbf{s}_h,\boldsymbol{\tau}_h)\in X_h}\|(\mathbf{t},\boldsymbol{\sigma}) - (\mathbf{s}_h,\boldsymbol{\tau}_h)\|_X \;+\; \inf_{(\mathbf{v}_h,\boldsymbol{\eta}_h)\in M_h}\|(\mathbf{u},\boldsymbol{\gamma}) - (\mathbf{v}_h,\boldsymbol{\eta}_h)\|_M\right\}. \end{aligned} \tag{3.17}$$

Proof. It is a direct consequence of the previous analysis and Theorem 3.3. □

Next, for the rate of convergence of (3.14) we proceed similarly as we did for Theorem 2.7, using now the Céa estimate (3.16) (or (3.17)), and the approximation properties of the subspaces involved. In particular, if the discrete augmented scheme (3.14) is defined with the subspaces from Theorem 2.7, we obtain exactly the same estimate (2.44) provided there. Moreover, this result also holds if we take the bubble functions away in (2.34) and consider the simpler subspace

$$\widetilde{X}_{1,h} := \left\{\mathbf{s}_h \in \mathbb{L}^2(\Omega): \quad \mathbf{s}_h|_T \in \mathbb{P}_k(T) \quad \forall\, T\in\mathcal{T}_h\right\}, \tag{3.18}$$

which, failing to satisfy (H.2) in Theorem 2.5, is certainly not suitable for the non-augmented discrete scheme (2.27). Note, however, that the approximation property ($\mathrm{AP}^{\mathbf{t}}_{1,h}$) of $X_{1,h}$ (cf. (2.34)), which was introduced in Section 2.6, actually corresponds to the approximation property of $\widetilde{X}_{1,h}$ (cf. (3.18)). The results described in this paragraph are summarized as follows.

THEOREM 3.5 *Given an integer $k \geq 0$, we take $X_{1,h}$ (cf. (2.34)) or $\widetilde{X}_{1,h}$ (cf. (3.18)) as the finite element subspace of X_1, and let $M_{1,h}$ and $M_h := M_h^{\mathbf{u}} \times M_h^{\gamma}$ be the finite element subspaces defined by* (2.31)*,* (2.32)*, and* (2.33)*, respectively. Let $((\mathbf{t},\boldsymbol{\sigma}),(\mathbf{u},\boldsymbol{\gamma})) \in X\times M$ and $((\mathbf{t}_h,\boldsymbol{\sigma}_h),(\mathbf{u}_h,\boldsymbol{\gamma}_h)) \in X_h\times M_h$ be the unique solutions of the continuous and resulting discrete formulations* (3.1) *and* (3.14)*, respectively. Assume that $\mathbf{t}\in\mathbb{H}^\delta(\Omega)$, $\boldsymbol{\sigma}\in\mathbb{H}^\delta(\Omega)$, $\mathbf{div}\,\boldsymbol{\sigma}\in\mathbf{H}^\delta(\Omega)$, $\mathbf{u}\in\mathbf{H}^\delta(\Omega)$ and $\boldsymbol{\gamma}\in\mathbb{H}^\delta(\Omega)$, for some $\delta\in(0,k+1]$. Then there exists $C>0$, independent of h, such that*

$$\begin{aligned} &\|((\mathbf{t},\boldsymbol{\sigma}),(\mathbf{u},\boldsymbol{\gamma})) - ((\mathbf{t}_h,\boldsymbol{\sigma}_h),(\mathbf{u}_h,\boldsymbol{\gamma}_h))\|_{X\times M} \\ &\qquad\leq\quad C\,h^\delta\left\{\|\mathbf{t}\|_{\delta,\Omega} + \|\boldsymbol{\sigma}\|_{\delta,\Omega} + \|\mathbf{div}\,\boldsymbol{\sigma}\|_{\delta,\Omega} + \|\mathbf{u}\|_{\delta,\Omega} + \|\boldsymbol{\gamma}\|_{\delta,\Omega}\right\}. \end{aligned}$$

On the other hand, we could keep $M_{1,h}$ and M_h as given by (2.31), (2.32), and (2.33), but use a lower polynomial degree for approximating $\mathbf{t}$ in the case $k\geq 1$. For example, instead of (3.18), we could consider:

$$\widehat{X}_{1,h} := \left\{\mathbf{s}_h \in \mathbb{L}^2(\Omega): \quad \mathbf{s}_h|_T \in \mathbb{P}_{k-1}(T) \quad \forall\, T\in\mathcal{T}_h\right\}, \tag{3.19}$$

which clearly does not satisfy (H.2) in Theorem 2.5 either. It follows, applying (2.39), that the approximation property of $\widehat{X}_{1,h}$ (cf. (3.19)) becomes as ($\mathrm{AP}^{\mathbf{t}}_{1,h}$), but with regularity range $[0,k]$ instead of $[0,k+1]$. Hence, thanks to the approximation properties of $M_{1,h}$ and M_h (cf. ($\mathrm{AP}^{\boldsymbol{\sigma}}_{1,h}$) and ($\mathrm{AP}^{\mathbf{u}}_{h}$) in Section 2.6), we also obtain in this case the same rate of convergence provided by Theorem 3.5, but limited to $\delta\in(0,k]$.

Another possibility is to approximate $\mathbf{t}$ by continuous piecewise polynomial tensors. For instance, given $k\geq 0$, we could keep again $M_{1,h}$ and M_h as given by (2.31), (2.32), and (2.33), and consider now:

$$\underline{X}_{1,h} := \left\{\mathbf{s}_h \in \mathbb{C}(\Omega): \quad \mathbf{s}_h|_T \in \mathbb{P}_{k+1}(T) \quad \forall\, T\in\mathcal{T}_h\right\}, \tag{3.20}$$

which, due to the continuity requirement, does not verify (H.2) in Theorem 2.5 either. In this case, assuming a convex domain Ω, one can show (cf. [54, eq. (3.5.15) and Remark 6.2.1]) that $\underline{X}_{1,h}$ (cf. (3.20)) satisfies the approximation property

($\underline{\mathrm{AP}}^{\mathbf{t}}_{1,h}$) *For each* $\delta \in [0, k+1]$ *and for each* $\mathbf{s} \in \mathbb{H}^{\delta}(\Omega) \cap X_1$ *there exists* $\mathbf{s}_h \in \underline{X}_{1,h}$ *such that*

$$\|\mathbf{s} - \mathbf{s}_h\|_{0,\Omega} \,\leq\, C\, h^{\delta}\, \|\mathbf{s}\|_{\delta,\Omega}\,.$$

Hence, the rate of convergence of the resulting augmented scheme is again the same provided by Theorem 3.5.

4 The fully augmented variational formulation

In this section we propose a fully augmented formulation for (2.8) and a corresponding discrete scheme whose main advantage is the elimination of the remaining assumptions on the finite element subspaces (cf. Theorem 3.3), which means that the discrete inf-sup condition for $\mathbb{B}$ is no longer needed. In other words, we show that a further enrichment of (3.1) yields an associated Galerkin scheme whose well-posedness is guaranteed by any finite element subspace of the resulting global space.

4.1 The continuous fully augmented formulation

In what follows we proceed as in [29] and enrich the variational formulation (3.1) with additional terms arising from the equilibrium equation and from the relations defining $\mathbf{t}$ and the rotation $\boldsymbol{\gamma}$ as functions of the displacement $\mathbf{u}$. In addition, in order to deal with the non-homogeneous Dirichlet boundary condition on Γ, we apply the idea from [30] (see also [40]) and introduce a consistent boundary term. More precisely, we first substract the second from the first equation of (3.1) and then add the redundant equations:

$$\begin{aligned}
&\kappa_1 \int_{\Omega} \big(\mathbf{div}\,\boldsymbol{\sigma} \,+\, \mathbf{f}\big)\cdot \mathbf{div}\,\boldsymbol{\tau} \,=\, 0\,,\\
&\kappa_2 \int_{\Omega} \big(\mathbf{e}(\mathbf{u}) \,-\, \mathbf{t}\big) : \mathbf{e}(\mathbf{v}) \,=\, 0\,,\\
&\kappa_3 \int_{\Omega} \Big\{\boldsymbol{\gamma} \,-\, \frac{1}{2}(\nabla\mathbf{u} \,-\, (\nabla\mathbf{u})^{\mathtt{t}})\Big\} : \boldsymbol{\eta} \,=\, 0\,,
\end{aligned}$$

and

$$\kappa_4 \int_{\Gamma} \mathbf{u}\cdot\mathbf{v} \,=\, \kappa_4 \int_{\Gamma} \mathbf{g}\cdot\mathbf{v}\,,$$

for all $(\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\eta}) \in \mathbb{H}(\mathbf{div};\Omega)\times\mathbf{H}^1(\Omega)\times\mathbb{L}^2_{\mathtt{skew}}(\Omega)$, where $(\kappa_1, \kappa_2, \kappa_3, \kappa_4)$ is a vector of positive constants, also named stabilization parameters, to be suitably chosen later on. It is important to observe here that the above terms require now the displacement $\mathbf{u}$ to live in $\mathbf{H}^1(\Omega)$ (instead of $\mathbf{u} \in \mathbf{L}^2(\Omega)$ as in (2.8) and (3.1)).

In this way, we now look at the following fully augmented variational formulation: Find $(\mathbf{t}, \boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\gamma}) \in \mathbb{X} := \mathbb{L}^2(\Omega)\times\mathbb{H}(\mathbf{div};\Omega)\times\mathbf{H}^1(\Omega)\times\mathbb{L}^2_{\mathtt{skew}}(\Omega)$ such that

$$[\mathbb{A}(\mathbf{t}, \boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\gamma}), (\mathbf{s}, \boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\eta})] \,=\, [\mathbb{F}, (\mathbf{s}, \boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\eta})] \qquad \forall\, (\mathbf{s}, \boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\eta}) \in \mathbb{X}\,, \tag{4.1}$$

where the nonlinear operator $\mathbb{A} : \mathbb{X} \to \mathbb{X}'$ and the functional $\mathbb{F} \in \mathbb{X}'$ are defined by

$$\begin{aligned}
{[\mathbb{A}(\mathbf{t}, \boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\gamma}), (\mathbf{s}, \boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\eta})]} \,:=\, & [\mathcal{A}(\mathbf{t}, \boldsymbol{\sigma}), (\mathbf{s}, \boldsymbol{\tau})] \,+\, [\mathcal{B}(\mathbf{s}, \boldsymbol{\tau}), (\mathbf{u}, \boldsymbol{\gamma})] \,-\, [\mathcal{B}(\mathbf{t}, \boldsymbol{\sigma}), (\mathbf{v}, \boldsymbol{\eta})]\\
& +\; \kappa_1 \int_{\Omega} \mathbf{div}\,\boldsymbol{\sigma}\cdot\mathbf{div}\,\boldsymbol{\tau} \,+\, \kappa_2 \int_{\Omega} \big(\mathbf{e}(\mathbf{u}) \,-\, \mathbf{t}\big) : \mathbf{e}(\mathbf{v})\\
& +\; \kappa_3 \int_{\Omega} \Big\{\boldsymbol{\gamma} \,-\, \frac{1}{2}(\nabla\mathbf{u} \,-\, (\nabla\mathbf{u})^{\mathtt{t}})\Big\} : \boldsymbol{\eta} \,+\, \kappa_4 \int_{\Gamma} \mathbf{u}\cdot\mathbf{v}\,,
\end{aligned} \tag{4.2}$$

and

$$[\mathbb{F},(\mathbf{s},\boldsymbol{\tau},\mathbf{v},\boldsymbol{\eta})] \,:=\, [\mathcal{F},(\mathbf{s},\boldsymbol{\tau})] \,-\, [\mathcal{G},(\mathbf{v},\boldsymbol{\eta})] \,-\, \kappa_1 \int_\Omega \mathbf{f}\cdot\mathbf{div}\,\boldsymbol{\tau} \,+\, \kappa_4 \int_\Gamma \mathbf{g}\cdot\mathbf{v}\,.$$

Our next goal is to show the unique solvability of the variational formulation (4.1), whence (3.1) and (4.1) share the same unique solution. We first recall from [53] the following abstract theorem.

Theorem 4.1 *Let X be a Hilbert space and let $A : X \to X'$ be a nonlinear operator. Assume that A is Lipschitz-continuous and strongly monotone on X, that is, there exist constants $\tilde{\gamma}$, $\tilde{\alpha} > 0$ such that*

$$||A(x) \,-\, A(y)||_{X'} \,\le\, \tilde{\gamma}\,\|x-y\|_X \qquad \forall\, x,\, y \in X\,,$$

and

$$[A(x)-A(y),x-y] \,\ge\, \tilde{\alpha}\,\|x-y\|_X^2 \qquad \forall\, x,\, y \in X\,.$$

Then, given $F \in X'$, there exists a unique $x \in X$ such that

$$[A(x),y] \;=\; [F,y] \qquad \forall\, x \in X\,.$$

Further, the following estimate holds

$$\|x\|_X \;\le\; \frac{1}{\tilde{\alpha}}\,\|F\|_{X'}\,. \tag{4.3}$$

Proof. It is a particular case of [53, Theorem 3.3.23]. □

In order to apply Theorem 4.1 to the fully augmented formulation (4.1), we need to prove first the required properties for our nonlinear operator $\mathbb{A}$ (cf. (4.2)). We begin with the Lipschitz-continuity.

Lemma 4.1 *Let $\mathbb{A}$ be the nonlinear operator defined by (4.2). Then, there exists a constant $\tilde{\gamma} > 0$, depending on γ (cf. (3.8)), $\|\mathcal{B}\|$, and the parameters κ_i, $i \in \{1,...,4\}$, such that*

$$||\mathbb{A}(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma}) \,-\, \mathbb{A}(\mathbf{s},\boldsymbol{\tau},\mathbf{v},\boldsymbol{\eta})||_{\mathbb{X}'} \;\le\; \tilde{\gamma}\,\|(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma}) - (\mathbf{s},\boldsymbol{\tau},\mathbf{v},\boldsymbol{\eta})\|_{\mathbb{X}}$$

for all $(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma})\,,\,(\mathbf{s},\boldsymbol{\tau},\mathbf{v},\boldsymbol{\eta}) \,\in\, \mathbb{X}$.

Proof. It basically follows from the Lipschitz-continuity of the nonlinear operator $\mathcal{A}$ (cf. Lemma 3.1) and the boundedness of the remaining terms (all bilinear) defining $\mathbb{A}$, together with applications of the Cauchy-Schwarz inequality and the trace theorem in $\mathbf{H}^1(\Omega)$. We omit further details. □

In turn, the strong monotonicity of $\mathbb{A}$ makes use of a slight extension of the second Korn inequality, which establishes the existence of a constant $c_1 > 0$ such that

$$\|\mathbf{e}(\mathbf{v})\|_{0,\Omega}^2 \,+\, \|\mathbf{v}\|_{0,\Gamma}^2 \,\ge\, c_1\,\|\mathbf{v}\|_{1,\Omega}^2 \qquad \forall\,\mathbf{v} \,\in\, \mathbf{H}^1(\Omega)\,. \tag{4.4}$$

The proof of (4.4) follows from a direct application of the Peetre-Tartar Lemma (see, e.g. [47, Theorem 2.1, Chapter I]). Alternatively, (4.4) is a particular case of [30, Lemma 3.1], whose proof employs analogue arguments to those given in the proof of [8, Theorem 9.2.16].

Lemma 4.2 *Let $\mathbb{A}$ be the nonlinear operator defined by (4.2), and let the parameter $\kappa_0 \in \left(0, \dfrac{2\,\alpha_1}{\gamma_1^2}\right)$, where γ_1 and α_1 are the positive constants from Lemma 2.1 (cf. (2.21)). In addition, assume that the parameters κ_1, κ_2, κ_3, and κ_4 are chosen such that $0 < \kappa_1$, $\;0 < \kappa_2 < 2\,\alpha$, $\;0 < \kappa_3 < \alpha_3$, and*

$0 < \kappa_4$, *where* α *is the constant from* (3.10) (*cf.* *Lemma* 3.2), *that is* $\alpha := \min\left\{\alpha_1 - \dfrac{\kappa_0\,\gamma_1^2}{2}, \dfrac{\kappa_0}{2}\right\}$, *and* $\alpha_3 := c_1 \min\{\kappa_2, 2\,\kappa_4\}$. *Then, there exists a constant* $\tilde{\alpha} > 0$, *depending on* α, c_1, κ_1, κ_2, κ_3, *and* κ_4, *such that*

$$[\mathbb{A}(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma}) - \mathbb{A}(\mathbf{s},\boldsymbol{\tau},\mathbf{v},\boldsymbol{\eta}), (\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma}) - (\mathbf{s},\boldsymbol{\tau},\mathbf{v},\boldsymbol{\eta})] \,\geq\, \tilde{\alpha}\,\|(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma}) - (\mathbf{s},\boldsymbol{\tau},\mathbf{v},\boldsymbol{\eta})\|_{\mathbb{X}}^2$$

for all $(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma})$, $(\mathbf{s},\boldsymbol{\tau},\mathbf{v},\boldsymbol{\eta}) \in \mathbb{X}$.

Proof. Given $(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma})$, $(\mathbf{s},\boldsymbol{\tau},\mathbf{v},\boldsymbol{\eta}) \in \mathbb{X}$, we observe, according to the definition of $\mathbb{A}$ (cf. (4.2)) and the fact that the terms involving $\mathcal{B}$ cancell out, that

$$\begin{aligned}
&[\mathbb{A}(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma}) - \mathbb{A}(\mathbf{s},\boldsymbol{\tau},\mathbf{v},\boldsymbol{\eta}), (\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma}) - (\mathbf{s},\boldsymbol{\tau},\mathbf{v},\boldsymbol{\eta})] \,=\, [\mathcal{A}(\mathbf{t},\boldsymbol{\sigma}) - \mathcal{A}(\mathbf{s},\boldsymbol{\tau}), (\mathbf{t},\boldsymbol{\sigma}) - (\mathbf{s},\boldsymbol{\tau})] \\
&\quad + \kappa_1\,\|\mathbf{div}(\boldsymbol{\sigma}-\boldsymbol{\tau})\|_{0,\Omega}^2 + \kappa_2 \int_\Omega \Big(\mathbf{e}(\mathbf{u}-\mathbf{v}) - (\mathbf{t}-\mathbf{s})\Big) : \mathbf{e}(\mathbf{u}-\mathbf{v}) \\
&\quad + \kappa_3 \int_\Omega \Big\{(\boldsymbol{\gamma}-\boldsymbol{\eta}) - \frac{1}{2}\Big(\nabla(\mathbf{u}-\mathbf{v}) - \big(\nabla(\mathbf{u}-\mathbf{v})\big)^{\mathtt{t}}\Big)\Big\} : (\boldsymbol{\gamma}-\boldsymbol{\eta}) \,+\, \kappa_4\,\|\mathbf{u}-\mathbf{v}\|_{0,\Gamma}^2\,,
\end{aligned}$$

which, applying (3.10) (cf. Lemma 3.2), the Cauchy-Schwarz inequality, and the basic estimate $a\,b \leq \frac{1}{2}\,(a^2+b^2)$, yields

$$\begin{aligned}
&[\mathbb{A}(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma}) - \mathbb{A}(\mathbf{s},\boldsymbol{\tau},\mathbf{v},\boldsymbol{\eta}), (\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma}) - (\mathbf{s},\boldsymbol{\tau},\mathbf{v},\boldsymbol{\eta})] \,\geq\, \alpha\Big\{\,\|\mathbf{t}-\mathbf{s}\|_{0,\Omega}^2 + \|\boldsymbol{\sigma}-\boldsymbol{\tau}\|_{0,\Omega}^2\Big\} \\
&\quad + \kappa_1\,\|\mathbf{div}(\boldsymbol{\sigma}-\boldsymbol{\tau})\|_{0,\Omega}^2 + \frac{\kappa_2}{2}\,\|\mathbf{e}(\mathbf{u}-\mathbf{v})\|_{0,\Omega}^2 - \frac{\kappa_2}{2}\,\|\mathbf{t}-\mathbf{s}\|_{0,\Omega}^2 + \frac{\kappa_3}{2}\,\|\boldsymbol{\gamma}-\boldsymbol{\eta}\|_{0,\Omega}^2 \\
&\quad - \frac{\kappa_3}{2}\,\Big\|\frac{1}{2}\Big(\nabla(\mathbf{u}-\mathbf{v}) - \big(\nabla(\mathbf{u}-\mathbf{v})\big)^{\mathtt{t}}\Big)\Big\|_{0,\Omega}^2 + \kappa_4\,\|\mathbf{u}-\mathbf{v}\|_{0,\Gamma}^2\,.
\end{aligned}$$

Then, noting that

$$\Big\|\frac{1}{2}\Big(\nabla(\mathbf{u}-\mathbf{v}) - \big(\nabla(\mathbf{u}-\mathbf{v})\big)^{\mathtt{t}}\Big)\Big\|_{0,\Omega}^2 \,=\, |\mathbf{u}-\mathbf{v}|_{1,\Omega}^2 - \|\mathbf{e}(\mathbf{u}-\mathbf{v})\|_{0,\Omega}^2\,,$$

and employing the Korn inequality (4.4), we find that

$$\begin{aligned}
&[\mathbb{A}(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma}) - \mathbb{A}(\mathbf{s},\boldsymbol{\tau},\mathbf{v},\boldsymbol{\eta}), (\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma}) - (\mathbf{s},\boldsymbol{\tau},\mathbf{v},\boldsymbol{\eta})] \,\geq\, \Big(\alpha - \frac{\kappa_2}{2}\Big)\,\|\mathbf{t}-\mathbf{s}\|_{0,\Omega}^2 + \alpha_2\,\|\boldsymbol{\sigma}-\boldsymbol{\tau}\|_{\mathbf{div};\Omega}^2 \\
&\quad + \frac{(\kappa_2+\kappa_3)}{2}\,\|\mathbf{e}(\mathbf{u}-\mathbf{v})\|_{0,\Omega}^2 + \kappa_4\,\|\mathbf{u}-\mathbf{v}\|_{0,\Gamma}^2 - \frac{\kappa_3}{2}\,|\mathbf{u}-\mathbf{v}|_{1,\Omega}^2 + \frac{\kappa_3}{2}\,\|\boldsymbol{\gamma}-\boldsymbol{\eta}\|_{0,\Omega}^2 \\
&\geq\, \Big(\alpha - \frac{\kappa_2}{2}\Big)\,\|\mathbf{t}-\mathbf{s}\|_{0,\Omega}^2 + \alpha_2\,\|\boldsymbol{\sigma}-\boldsymbol{\tau}\|_{\mathbf{div};\Omega}^2 + \frac{(\alpha_3-\kappa_3)}{2}\,\|\mathbf{u}-\mathbf{v}\|_{1,\Omega}^2 + \frac{\kappa_3}{2}\,\|\boldsymbol{\gamma}-\boldsymbol{\eta}\|_{0,\Omega}^2
\end{aligned} \tag{4.5}$$

where $\alpha_2 := \min\{\alpha, \kappa_1\}$, thus finishing the proof with $\tilde{\alpha} := \min\Big\{\Big(\alpha - \frac{\kappa_2}{2}\Big), \alpha_2, \frac{(\alpha_3-\kappa_3)}{2}, \frac{\kappa_3}{2}\Big\}$.

□

The well-posedness of the fully augmented formulation (4.1) can be established as follows.

Theorem 4.2 *Assume that the parameters* κ_0, κ_1, κ_2, κ_3, *and* κ_4 *are chosen as indicated in Lemma* 4.2. *Then, there exists a unique* $(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma}) \in \mathbb{X}$ *solution of* (4.1). *Moreover, there exists* $C > 0$, *depending on* $\tilde{\alpha}$ (*cf. Lemma* 4.2), *such that*

$$\|(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma})\|_{\mathbb{X}} \,\leq\, C\Big\{\,\|\mathbf{f}\|_{0,\Omega} + \|\mathbf{g}\|_{1/2,\Gamma}\Big\}\,.$$

Proof. Thanks to Lemmas 4.1 and Lemma 4.2, the proof is a direct application of Theorem 4.1. □

It is important to remark here that all the parameters but κ_3, which depends on the unknown constant c_1 (cf. (4.4)), can be chosen explicitly. In particular, and adopting as criterion the choice of the average value of each feasible range, we take $\kappa_0 = \dfrac{\alpha_1}{\gamma_1^2}$ and $\kappa_2 = \alpha$, which yields

$$\kappa_2 = \frac{\alpha_1}{2} \min\left\{1, \frac{1}{\gamma_1^2}\right\}.$$

Next, in order to maximize the values of the minima involved in the definition of α_2 and α_3, thus maximizing these constants, we choose $\kappa_1 = \alpha$ and $2\,\kappa_4 = \kappa_2$, which gives

$$\kappa_1 = \frac{\alpha_1}{2} \min\left\{1, \frac{1}{\gamma_1^2}\right\} \quad \text{and} \quad \kappa_4 = \frac{\alpha_1}{4} \min\left\{1, \frac{1}{\gamma_1^2}\right\}.$$

Then, the theoretical feasible range of κ_3 becomes the interval $\big(0, c_1\,\kappa_2\big)$, whose average value is

$$\kappa_3 = c_1 \frac{\alpha_1}{4} \min\left\{1, \frac{1}{\gamma_1^2}\right\}.$$

The numerical results shown below in Section 6, which simply assume $c_1 = 1$ in the above expression, illustrate that not knowing this constant does not really affect, at least for the examples considered there, the well-posedness of the resulting discrete fully augmented scheme.

On the other hand, we remark that when $\mathbf{g} = \mathbf{0}$, that is in the case of homogeneous Dirichlet boundary conditions, there is no need to introduce the boundary term on Γ, and hence no parameter κ_4 appears in the fully augmented formulation. In fact, the corresponding product space is then $\mathbb{X}_0 := \mathbb{L}^2(\Omega) \times \mathbb{H}(\mathbf{div};\Omega) \times \mathbf{H}_0^1(\Omega) \times \mathbb{L}^2_{\mathtt{skew}}(\Omega)$, and according to the first Korn inequality, which says that

$$\|\mathbf{e}(\mathbf{v})\|_{0,\Omega}^2 \geq \frac{1}{2}\,|\mathbf{v}|_{1,\Omega}^2 \qquad \forall\, \mathbf{v} \in \mathbf{H}_0^1(\Omega),$$

the estimate (4.5) now becomes

$$\begin{aligned}[\mathbb{A}(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma}) - \mathbb{A}(\mathbf{s},\boldsymbol{\tau},\mathbf{v},\boldsymbol{\eta}),(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma}) - (\mathbf{s},\boldsymbol{\tau},\mathbf{v},\boldsymbol{\eta})] &\geq \Big(\alpha - \frac{\kappa_2}{2}\Big)\,\|\mathbf{t} - \mathbf{s}\|_{0,\Omega}^2 \\ + \; \alpha_2\,\|\boldsymbol{\sigma} - \boldsymbol{\tau}\|_{\mathbf{div};\Omega}^2 + \frac{(\kappa_2 - \kappa_3)}{4}\,|\mathbf{u} - \mathbf{v}|_{1,\Omega}^2 &+ \frac{\kappa_3}{2}\,\|\boldsymbol{\gamma} - \boldsymbol{\eta}\|_{0,\Omega}^2 .\end{aligned}$$

In this way, the strong monotonicity of $\mathbb{A}$ is guaranteed by any explicit choice of the parameters satisfying $0 < \kappa_2 < 2\,\alpha$ and $0 < \kappa_3 < \kappa_2$ (besides the already mentioned choices for κ_0 and κ_1).

4.2 The discrete fully augmented formulation

We now consider the Galerkin scheme associated with the fully augmented formulation (4.1). For this purpose, we now let $X_{1,h}$, $M_{1,h}$, $M_h^{\mathbf{u}}$, and $M_h^{\boldsymbol{\gamma}}$ be finite dimensional subspaces of $\mathbb{L}^2(\Omega)$, $\mathbb{H}(\mathbf{div};\Omega)$, $\mathbf{H}^1(\Omega)$, and $\mathbb{L}^2_{\mathtt{skew}}(\Omega)$, respectively, and define $\mathbb{X}_h := X_{1,h} \times M_{1,h} \times M_h^{\mathbf{u}} \times M_h^{\boldsymbol{\gamma}}$. Then, we are interested in the following discrete scheme: Find $(\mathbf{t}_h,\boldsymbol{\sigma}_h,\mathbf{u}_h,\boldsymbol{\gamma}_h) \in \mathbb{X}_h$ such that

$$[\mathbb{A}(\mathbf{t}_h,\boldsymbol{\sigma}_h,\mathbf{u}_h,\boldsymbol{\gamma}_h),(\mathbf{s},\boldsymbol{\tau},\mathbf{v},\boldsymbol{\eta})] = [\mathbb{F},(\mathbf{s},\boldsymbol{\tau},\mathbf{v},\boldsymbol{\eta})] \qquad \forall\,(\mathbf{s},\boldsymbol{\tau},\mathbf{v},\boldsymbol{\eta}) \in \mathbb{X}_h. \tag{4.6}$$

The following theorem establishes the well-posedness and convergence properties of (4.6).

THEOREM 4.3 *Assume that the parameters κ_0, κ_1, κ_2, κ_3, and κ_4 are chosen as indicated in Lemma 4.2. In addition, let $X_{1,h}$, $M_{1,h}$, $M_h^{\mathbf{u}}$, and M_h^{γ} be arbitrary finite dimensional subspaces of $\mathbb{L}^2(\Omega)$, $\mathbb{H}(\mathbf{div};\Omega)$, $\mathbf{H}^1(\Omega)$, and $\mathbb{L}^2_{\mathtt{skew}}(\Omega)$, respectively. Then there exists a unique $(\mathbf{t}_h, \boldsymbol{\sigma}_h, \mathbf{u}_h, \boldsymbol{\gamma}_h) \in \mathbb{X}_h$ solution of* (4.6). *Moreover, there exist C_1, $C_2 > 0$, independent of h, such that*

$$\|(\mathbf{t}_h, \boldsymbol{\sigma}_h, \mathbf{u}_h, \boldsymbol{\gamma}_h)\|_{\mathbb{X}} \;\le\; C_1 \Big\{ \|\mathbf{f}\|_{0,\Omega} \,+\, \|\mathbf{g}\|_{1/2,\Gamma} \Big\}, \tag{4.7}$$

and

$$\begin{aligned} &\|(\mathbf{t}, \boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\gamma}) - (\mathbf{t}_h, \boldsymbol{\sigma}_h, \mathbf{u}_h, \boldsymbol{\gamma}_h)\|_{\mathbb{X}} \\ &\le \quad C_2 \inf_{(\mathbf{s}_h, \boldsymbol{\tau}_h, \mathbf{v}_h, \boldsymbol{\eta}_h) \in \mathbb{X}_h} \|(\mathbf{t}, \boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\gamma}) - (\mathbf{s}_h, \boldsymbol{\tau}_h, \mathbf{v}_h, \boldsymbol{\eta}_h)\|_{\mathbb{X}}\,. \end{aligned} \tag{4.8}$$

Proof. It is clear that the Lipschitz-continuity and strong monotonicity of $\mathbb{A}$ (cf. Lemmas 4.1 and 4.2) are certainly valid on $\mathbb{X}_h \times \mathbb{X}'_h$, with the same constants $\tilde{\gamma}$ and $\tilde{\alpha}$, respectively. Therefore, the unique solvability of (4.6) and the estimate (4.7) are again consequence of Theorem 4.1. In turn, the Céa estimate (4.8) follows from standard arguments, similarly as for linear problems, and using obviously the above mentioned properties of $\mathbb{A}$. We omit further details.

□

Next, we consider the canonical finite element subspaces $X_{1,h}$, $M_{1,h}$, $M_h^{\mathbf{u}}$, and M_h^{γ} of $\mathbb{L}^2(\Omega)$, $\mathbb{H}(\mathbf{div};\Omega)$, $\mathbf{H}^1(\Omega)$, and $\mathbb{L}^2_{\mathtt{skew}}(\Omega)$, respectively. More precisely, given an integer $k \ge 0$, we now define

$$X_{1,h} \,:=\, \Big\{ \mathbf{s}_h \in \mathbb{L}^2(\Omega) : \quad \mathbf{s}_h|_T \in \mathbb{P}_k(T) \quad \forall\, T \in \mathcal{T}_h \Big\}, \tag{4.9}$$

$$M_{1,h} \,:=\, \Big\{ \boldsymbol{\tau}_h \in \mathbb{H}(\mathbf{div};\Omega) : \quad \boldsymbol{\tau}_h|_T \in [\mathbf{RT}_k(T)]^2 \quad \forall\, T \in \mathcal{T}_h \Big\}, \tag{4.10}$$

$$M_h^{\mathbf{u}} \,:=\, \Big\{ \mathbf{v}_h \in \mathbf{C}(\bar{\Omega}) : \quad \mathbf{v}_h|_T \,\in\, \mathbf{P}_{k+1}(T) \quad \forall\, T \in \mathcal{T}_h \Big\}, \tag{4.11}$$

and

$$M_h^{\gamma} \,:=\, \Big\{ \boldsymbol{\eta}_h \in \mathbb{L}^2_{\mathtt{skew}}(\Omega) : \quad \boldsymbol{\eta}_h|_T \,\in\, \mathbb{P}_k(T) \quad \forall\, T \in \mathcal{T}_h \Big\}. \tag{4.12}$$

It is easy to see that the approximation properties of $X_{1,h}$ (cf. (4.9)) and $M_{1,h}$ (cf. (4.10)) are given by ($\mathrm{AP}^{\mathbf{t}}_{1,h}$) and ($\mathrm{AP}^{\boldsymbol{\sigma}}_{1,h}$) in Section 2.6. Note, in particular, that the present $X_{1,h}$ coincides with the subspace $\widetilde{X}_{1,h}$ (cf. (3.18)), whose approximation property was already identified in Section 3.3. In turn, the approximation property of M_h^{γ} (cf. (4.12)) is basically the same as that of $X_{1,h}$ (except for the skew-symmetry), while the one of $M_h^{\mathbf{u}}$ (cf. (4.11)), which is the classical Lagrange finite element subspace of order $k+1$, reduces to the following (see [16]):

($\mathrm{AP}^{\mathbf{u}}_h$) *For each $\delta \in (0, k+1]$ and for each $\mathbf{v} \in \mathbf{H}^{1+\delta}(\Omega)$ there exists $\mathbf{v}_h \in M_h^{\mathbf{u}}$ such that*

$$\|\mathbf{v} - \mathbf{v}_h\|_{1,\Omega} \;\le\; C\, h^{\delta}\, \|\mathbf{v}\|_{1+\delta,\Omega}\,.$$

The following theorem provides the corresponding rate of convergence of (4.6).

THEOREM 4.4 *Assume that the parameters κ_0, κ_1, κ_2, κ_3, and κ_4 are chosen as indicated in Lemma 4.2. In addition, given an integer $k \ge 0$, we let $X_{1,h}$, $M_{1,h}$, $M_h^{\mathbf{u}}$, and M_h^{γ} be the finite element subspaces defined by* (4.9), (4.10), (4.11), *and* (4.12), *respectively. Let $(\mathbf{t}, \boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\gamma}) \in \mathbb{X}$ and $(\mathbf{t}_h, \boldsymbol{\sigma}_h, \mathbf{u}_h, \boldsymbol{\gamma}_h) \in \mathbb{X}_h$ be the unique solutions of the continuous and discrete formulations* (4.1) *and* (4.6), *respectively, and* *suppose that* $\mathbf{t} \in \mathbb{H}^\delta(\Omega)$, $\boldsymbol{\sigma} \in \mathbb{H}^\delta(\Omega)$, $\mathbf{div}\,\boldsymbol{\sigma} \in \mathbf{H}^\delta(\Omega)$, $\mathbf{u} \in \mathbf{H}^{1+\delta}(\Omega)$ *and* $\boldsymbol{\gamma} \in \mathbb{H}^\delta(\Omega)$, *for some* $\delta \in (0, k+1]$. *Then there exists* $C > 0$, *independent of* h, *such that*

$$\|(\mathbf{t}, \boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\gamma}) - (\mathbf{t}_h, \boldsymbol{\sigma}_h, \mathbf{u}_h, \boldsymbol{\gamma}_h)\|_{\mathbb{X}}$$

$$\leq \quad C\, h^\delta \Big\{ \|\mathbf{t}\|_{\delta,\Omega} + \|\boldsymbol{\sigma}\|_{\delta,\Omega} + \|\mathbf{div}\,\boldsymbol{\sigma}\|_{\delta,\Omega} + \|\mathbf{u}\|_{1+\delta,\Omega} + \|\boldsymbol{\gamma}\|_{\delta,\Omega} \Big\}.$$

Proof. It follows from the Céa estimate (4.8) and the above indicated approximation properties. □

5 A posteriori error analysis

In this section we derive reliable and efficient residual-based a posteriori error estimators for the Galerkin schemes (2.27), (3.14) and (4.6).

5.1 Preliminaries and main results

We begin by introducing several notations. We let $\mathcal{E}_h$ be the set of all edges of the triangulation $\mathcal{T}_h$, and given $T \in \mathcal{T}_h$, we let $\mathcal{E}(T)$ be the set of its edges. Then we write $\mathcal{E}_h = \mathcal{E}_h(\Omega) \cup \mathcal{E}_h(\Gamma)$, where $\mathcal{E}_h(\Omega) := \{e \in \mathcal{E}_h : e \subseteq \Omega\}$ and $\mathcal{E}_h(\Gamma) := \{e \in \mathcal{E}_h : e \subseteq \Gamma\}$. In what follows, h_e stands for the length of the edge e. Also, for each edge $e \in \mathcal{E}_h$ we fix a unit normal vector $\boldsymbol{\nu}_e := (\nu_1, \nu_2)^{\mathtt{t}}$, and let $s_e := (-\nu_2, \nu_1)^{\mathtt{t}}$ be the corresponding fixed unit tangential vector along e. Then, given $e \in \mathcal{E}_h(\Omega)$ and $\boldsymbol{\tau} \in \mathbb{L}^2(\Omega)$ such that $\boldsymbol{\tau}|_T \in \mathbb{C}(T)$ on each $T \in \mathcal{T}_h$, we let $[\boldsymbol{\tau}\, s_e]$ be the corresponding jump across e, that is $[\boldsymbol{\tau}\, s_e] := (\boldsymbol{\tau}|_T - \boldsymbol{\tau}|_{T'})|_e\, s_e$, where T and T' are the triangles of $\mathcal{T}_h$ having e as a common edge. Abusing notation, when $e \in \mathcal{E}_h(\Gamma)$, we also write $[\boldsymbol{\tau}\, s_e] := \boldsymbol{\tau}|_e\, s_e$. Similar definitions hold for the tangential jumps of scalar fields $v \in L^2(\Omega)$ such that $v|_T \in C(T)$ on each $T \in \mathcal{T}_h$. From now on, when no confusion arises, we simply write s and $\boldsymbol{\nu}$ instead of s_e and $\boldsymbol{\nu}_e$, respectively. Finally, given scalar, vector and tensor valued fields v, $\boldsymbol{\varphi} := (\varphi_1, \varphi_2)$ and $\boldsymbol{\tau} := (\tau_{ij})$, respectively, we recall that $\mathbf{curl}^{\mathtt{t}}\, v$ is defined in (2.30), and now let

$$\mathbf{curl}\, v := \begin{pmatrix} \dfrac{\partial v}{\partial x_2} \\ -\dfrac{\partial v}{\partial x_1} \end{pmatrix}, \quad \underline{\mathbf{curl}}(\boldsymbol{\varphi}) := \begin{pmatrix} \mathbf{curl}^{\mathtt{t}}\, \varphi_1 \\ \mathbf{curl}^{\mathtt{t}}\, \varphi_2 \end{pmatrix} \quad \text{and} \quad \mathrm{curl}(\boldsymbol{\tau}) := \begin{pmatrix} \frac{\partial \tau_{12}}{\partial x_1} - \frac{\partial \tau_{11}}{\partial x_2} \\ \frac{\partial \tau_{22}}{\partial x_1} - \frac{\partial \tau_{21}}{\partial x_2} \end{pmatrix}.$$

Then, letting $(\mathbf{t}, \boldsymbol{\sigma}, (\mathbf{u}, \boldsymbol{\gamma})) \in X_1 \times M_1 \times M$ and $(\mathbf{t}_h, \boldsymbol{\sigma}_h, (\mathbf{u}_h, \boldsymbol{\gamma}_h)) \in X_{1,h} \times M_{1,h} \times M_h$ be the unique solutions of the continuous and discrete formulations (2.8) and (2.27), respectively, we define for each $T \in \mathcal{T}_h$ a local error indicator θ_T as follows:

$$\begin{aligned} \theta_T^2 &:= \big\|\boldsymbol{\sigma}_h - \big\{\boldsymbol{\lambda}(\mathbf{t}_h)\,\mathrm{tr}(\mathbf{t}_h)\,\mathbb{I} + \boldsymbol{\mu}(\mathbf{t}_h)\,\mathbf{t}_h\big\}\big\|_{0,T}^2 + \|\mathbf{f} + \mathbf{div}\,\boldsymbol{\sigma}_h\|_{0,T}^2 + \big\|\boldsymbol{\sigma}_h - \boldsymbol{\sigma}_h^{\mathtt{t}}\big\|_{0,T}^2 \\ &+ h_T^2\, \big\|\mathrm{curl}\{\mathbf{t}_h + \boldsymbol{\gamma}_h\}\big\|_{0,T}^2 + h_T^2\, \|\nabla \mathbf{u}_h - (\mathbf{t}_h + \boldsymbol{\gamma}_h)\|_{0,T}^2 + \sum_{e \in \mathcal{E}(T) \cap \mathcal{E}_h(\Omega)} h_e\, \|[(\mathbf{t}_h + \boldsymbol{\gamma}_h)\, s]\|_{0,e}^2 \\ &+ \sum_{e \in \mathcal{E}(T) \cap \mathcal{E}_h(\Gamma)} h_e \left\{ \left\|\frac{d\mathbf{g}}{ds} - (\mathbf{t}_h + \boldsymbol{\gamma}_h)\, s\right\|_{0,e}^2 + \|\mathbf{g} - \mathbf{u}_h\|_{0,e}^2 \right\}. \end{aligned} \tag{5.1}$$

Note that the above requires that $\left.\dfrac{d\mathbf{g}}{ds}\right|_e \in \mathbf{L}^2(e)$ for each $e \in \mathcal{E}_h(\Gamma)$. This is fixed below by assuming that $\mathbf{g} \in \mathbf{H}^1(\Gamma)$.

Similarly, letting $((\mathbf{t},\boldsymbol{\sigma}),(\mathbf{u},\boldsymbol{\gamma})) \in X \times M$ and $((\mathbf{t}_h,\boldsymbol{\sigma}_h),(\mathbf{u}_h,\boldsymbol{\gamma}_h)) \in X_h \times M_h$ be the unique solutions of the continuous and discrete formulations (3.1) and (3.14), respectively, we define for each $T \in \mathcal{T}_h$ a local error indicator $\widetilde{\theta}_T$ as follows:

$$\begin{aligned}\widetilde{\theta}_T^2 &:= \theta_T^2 + h_T^2 \left\|\operatorname{curl}\big(\boldsymbol{\sigma}_h - \{\boldsymbol{\lambda}(\mathbf{t}_h)\operatorname{tr}(\mathbf{t}_h)\mathbb{I} + \boldsymbol{\mu}(\mathbf{t}_h)\mathbf{t}_h\}\big)\right\|_{0,T}^2 \\ &+ \sum_{e\in\mathcal{E}(T)\cap\mathcal{E}_h(\Omega)} h_e \left\|\big[\big(\boldsymbol{\sigma}_h - \{\boldsymbol{\lambda}(\mathbf{t}_h)\operatorname{tr}(\mathbf{t}_h)\mathbb{I} + \boldsymbol{\mu}(\mathbf{t}_h)\mathbf{t}_h\}\big)s\big]\right\|_{0,e}^2 .\end{aligned} \tag{5.2}$$

In turn, letting $(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma}) \in \mathbb{X}$ and $(\mathbf{t}_h,\boldsymbol{\sigma}_h,\mathbf{u}_h,\boldsymbol{\gamma}_h) \in \mathbb{X}_h$ be the unique solutions of the continuous and discrete formulations (4.1) and (4.6), respectively, we define for each $T \in \mathcal{T}_h$ a local error indicator $\widehat{\theta}_T$ as follows:

$$\begin{aligned}\widehat{\theta}_T^2 &:= \left\|\boldsymbol{\sigma}_h - \{\boldsymbol{\lambda}(\mathbf{t}_h)\operatorname{tr}(\mathbf{t}_h)\mathbb{I} + \boldsymbol{\mu}(\mathbf{t}_h)\mathbf{t}_h\}\right\|_{0,T}^2 + \|\mathbf{f} + \mathbf{div}\,\boldsymbol{\sigma}_h\|_{0,T}^2 + \left\|\boldsymbol{\sigma}_h - \boldsymbol{\sigma}_h^{\mathtt{t}}\right\|_{0,T}^2 \\ &+ h_T^2 \left\|\operatorname{curl}\{\mathbf{t}_h + \boldsymbol{\gamma}_h\}\right\|_{0,T}^2 + h_T^2 \|\nabla\mathbf{u}_h - (\mathbf{t}_h + \boldsymbol{\gamma}_h)\|_{0,T}^2 + \sum_{e\in\mathcal{E}(T)\cap\mathcal{E}_h(\Omega)} h_e \|[(\mathbf{t}_h + \boldsymbol{\gamma}_h)s]\|_{0,e}^2 \\ &+ \sum_{e\in\mathcal{E}(T)\cap\mathcal{E}_h(\Gamma)} \left\{ h_e \left\|\frac{d\mathbf{g}}{ds} - (\mathbf{t}_h + \boldsymbol{\gamma}_h)s\right\|_{0,e}^2 + \|\mathbf{g} - \mathbf{u}_h\|_{0,e}^2 \right\} \\ &+ h_T^2 \left\|\operatorname{curl}\big(\boldsymbol{\sigma}_h - \{\boldsymbol{\lambda}(\mathbf{t}_h)\operatorname{tr}(\mathbf{t}_h)\mathbb{I} + \boldsymbol{\mu}(\mathbf{t}_h)\mathbf{t}_h\}\big)\right\|_{0,T}^2 + \left\|\boldsymbol{\gamma}_h - \frac{1}{2}(\nabla\mathbf{u}_h - (\nabla\mathbf{u}_h)^{\mathtt{t}})\right\|_{0,T}^2 \\ &+ \|\mathbf{e}(\mathbf{u}_h) - \mathbf{t}_h\|_{0,T}^2 + \sum_{e\in\mathcal{E}(T)\cap\mathcal{E}_h(\Omega)} h_e \left\|\big[\big(\boldsymbol{\sigma}_h - \{\boldsymbol{\lambda}(\mathbf{t}_h)\operatorname{tr}(\mathbf{t}_h)\mathbb{I} + \boldsymbol{\mu}(\mathbf{t}_h)\mathbf{t}_h\}\big)s\big]\right\|_{0,e}^2 .\end{aligned} \tag{5.3}$$

Equivalently,

$$\begin{aligned}\widehat{\theta}_T^2 &= \widetilde{\theta}_T^2 + \sum_{e\in\mathcal{E}(T)\cap\mathcal{E}_h(\Gamma)} (1 - h_e)\|\mathbf{g} - \mathbf{u}_h\|_{0,e}^2 \\ &+ \left\|\boldsymbol{\gamma}_h - \frac{1}{2}(\nabla\mathbf{u}_h - (\nabla\mathbf{u}_h)^{\mathtt{t}})\right\|_{0,T}^2 + \|\mathbf{e}(\mathbf{u}_h) - \mathbf{t}_h\|_{0,T}^2 .\end{aligned}$$

The residual character of each term on the right hand sides of (5.1), (5.2) and (5.3) is quite clear. As usual the expressions

$$\boldsymbol{\theta} := \left\{\sum_{T\in\mathcal{T}_h} \theta_T^2\right\}^{1/2}, \quad \widetilde{\boldsymbol{\theta}} := \left\{\sum_{T\in\mathcal{T}_h} \widetilde{\theta}_T^2\right\}^{1/2} \quad \text{and} \quad \widehat{\boldsymbol{\theta}} := \left\{\sum_{T\in\mathcal{T}_h} \widehat{\theta}_T^2\right\}^{1/2}$$

are employed as the respective global residual error estimators.

The following theorems constitute the main results of this section.

THEOREM 5.1 *Let* $(\mathbf{t},\boldsymbol{\sigma},(\mathbf{u},\boldsymbol{\gamma})) \in \mathbf{X} := X_1 \times M_1 \times M$ *and* $(\mathbf{t}_h,\boldsymbol{\sigma}_h,(\mathbf{u}_h,\boldsymbol{\gamma}_h)) \in X_{1,h} \times M_{1,h} \times M_h$ *be the unique solutions of the continuous and discrete formulations* (2.8) *and* (2.27)*, respectively, and assume that* $\mathbf{g} \in \mathbf{H}^1(\Gamma)$*. Then there exist positive constants* $C_{\mathtt{eff}}$ *and* $C_{\mathtt{rel}}$*, independent of* h*, such that*

$$C_{\mathtt{eff}}\,\boldsymbol{\theta} + \text{h.o.t.} \le \|(\mathbf{t},\boldsymbol{\sigma},(\mathbf{u},\boldsymbol{\gamma})) - (\mathbf{t}_h,\boldsymbol{\sigma}_h,(\mathbf{u}_h,\boldsymbol{\gamma}_h))\|_{\mathbf{X}} \le C_{\mathtt{rel}}\,\boldsymbol{\theta}, \tag{5.4}$$

where h.o.t. *stands for one or several terms of higher order.*

THEOREM 5.2 *Let* $((\mathbf{t},\boldsymbol{\sigma}),(\mathbf{u},\gamma)) \in X \times M$ *and* $((\mathbf{t}_h,\boldsymbol{\sigma}_h),(\mathbf{u}_h,\gamma_h)) \in X_h \times M_h$ *be the unique solutions of the continuous and discrete augmented formulations* (3.1) *and* (3.14), *respectively, and assume that* $\mathbf{g} \in \mathbf{H}^1(\Gamma)$. *Then there exist positive constants* $\widetilde{C}_{\tt eff}$ *and* $\widetilde{C}_{\tt rel}$, *independent of* h, *such that*

$$\widetilde{C}_{\tt eff}\,\widetilde{\boldsymbol{\theta}} + \text{h.o.t.} \;\le\; \|((\mathbf{t},\boldsymbol{\sigma}),(\mathbf{u},\gamma)) - ((\mathbf{t}_h,\boldsymbol{\sigma}_h),(\mathbf{u}_h,\gamma_h))\|_{X\times M} \;\le\; \widetilde{C}_{\tt rel}\,\widetilde{\boldsymbol{\theta}}\,, \tag{5.5}$$

where h.o.t. *stands for one or several terms of higher order.*

THEOREM 5.3 *Let* $(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\gamma) \in \mathbb{X}$ *and* $(\mathbf{t}_h,\boldsymbol{\sigma}_h,\mathbf{u}_h,\gamma_h) \in \mathbb{X}_h$ *be the unique solutions of the continuous and discrete augmented formulations* (4.1) *and* (4.6), *respectively, and assume that* $\mathbf{g} \in \mathbf{H}^1(\Gamma)$. *Then there exist positive constants* $\widehat{C}_{\tt eff}$ *and* $\widehat{C}_{\tt rel}$, *independent of* h, *such that*

$$\widehat{C}_{\tt eff}\,\widehat{\boldsymbol{\theta}} + \text{h.o.t.} \;\le\; \|(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\gamma) - (\mathbf{t}_h,\boldsymbol{\sigma}_h,\mathbf{u}_h,\gamma_h)\|_{\mathbb{X}} \;\le\; \widehat{C}_{\tt rel}\,\widehat{\boldsymbol{\theta}}\,, \tag{5.6}$$

where h.o.t. *stands for one or several terms of higher order.*

We remark in advance that for the proofs of these theorems we follow very closely the approaches introduced in [42] and [27]. The efficiency of the global error estimators (lower bounds in (5.4), (5.5) and (5.6)) is proved below in Section 5.5, whereas the corresponding reliability (upper bounds in (5.4), (5.5) and (5.6)) is derived next in Sections 5.2, 5.3 and 5.4. However, since the reliability and efficiency of $\boldsymbol{\theta}$ (cf. Theorem 5.1) were already proved in [27] within a general framework for nonlinear twofold saddle point formulations, in the corresponding sections below we just provide, for sake of completeness, the main aspects of the associated analysis.

5.2 Reliability of the a posteriori error estimator $\boldsymbol{\theta}$

We begin by recalling from the analysis in Section 2.4 that the Gâteaux derivatives $\big\{\mathcal{D}\mathbb{A}_1(\tilde{\mathbf{r}})\big\}_{\tilde{\mathbf{r}}\in X_1}$ constitute a family of uniformly bounded and uniformly elliptic bilinear forms on $X_1 \times X_1$ (cf. (2.16) and (2.17) in Lemma 2.1), and that the operators $\mathbb{B}$ and $\mathbb{B}_1$ satisfy the corresponding continuous inf-sup conditions (cf. Lemma 2.3 and the discussion right after it). Hence, as a consequence of the continuous dependence result provided by the linear version of Theorem 2.1 (cf. (2.12) with $\mathbf{A}_1$ linear), we conclude that the linear operator $\mathcal{L}$ obtained by adding the three equations of the left hand side of (2.8), after replacing $\mathbb{A}_1$ by the Gâteaux derivative $\mathcal{D}\mathbb{A}_1(\tilde{\mathbf{r}})$ at any $\tilde{\mathbf{r}} \in X_1$, satisfies a global inf-sup condition. More precisely, there exists a constant $\tilde{C} > 0$ such that

$$\tilde{C}\,\|(\mathbf{r},\boldsymbol{\zeta},(\mathbf{w},\boldsymbol{\xi}))\|_{\mathbf{X}} \;\le\; \sup_{(\mathbf{s},\boldsymbol{\tau},(\mathbf{v},\boldsymbol{\eta}))\in\mathbf{X}\setminus\{\mathbf{0}\}} \frac{[\mathcal{L}(\mathbf{s},\boldsymbol{\tau},(\mathbf{v},\boldsymbol{\eta})),(\mathbf{r},\boldsymbol{\zeta},(\mathbf{w},\boldsymbol{\xi}))]}{\|(\mathbf{s},\boldsymbol{\tau},(\mathbf{v},\boldsymbol{\eta}))\|_{\mathbf{X}}} \tag{5.7}$$

for all $(\tilde{\mathbf{r}},(\mathbf{r},\boldsymbol{\zeta},(\mathbf{w},\boldsymbol{\xi}))) \in X_1 \times \mathbf{X}$, where

$$\begin{aligned}[\mathcal{L}(\mathbf{s},\boldsymbol{\tau},(\mathbf{v},\boldsymbol{\eta})),(\mathbf{r},\boldsymbol{\zeta},(\mathbf{w},\boldsymbol{\xi}))] \;:=\; &\mathcal{D}\mathbb{A}_1(\tilde{\mathbf{r}})(\mathbf{r},\mathbf{s}) \;+\; [\mathbb{B}_1(\mathbf{s}),\boldsymbol{\zeta}] \;+\; [\mathbb{B}_1(\mathbf{r}),\boldsymbol{\tau}]\\ &+\; [\mathbb{B}(\boldsymbol{\tau}),(\mathbf{w},\boldsymbol{\xi})] \;+\; [\mathbb{B}(\boldsymbol{\zeta}),(\mathbf{v},\boldsymbol{\eta})]\,.\end{aligned} \tag{5.8}$$

We now have the following preliminary result.

LEMMA 5.1 *Let* $(\mathbf{t},\boldsymbol{\sigma},(\mathbf{u},\gamma)) \in \mathbf{X} := X_1 \times M_1 \times M$ *and* $(\mathbf{t}_h,\boldsymbol{\sigma}_h,(\mathbf{u}_h,\gamma_h)) \in X_{1,h} \times M_{1,h} \times M_h$ *be the unique solutions of the continuous and discrete formulations* (2.8) *and* (2.27), *respectively. Then there exists* $C > 0$, *independent of* h, *such that*

$$\begin{aligned}C\,\|(\mathbf{t},\boldsymbol{\sigma},(\mathbf{u},\gamma)) - (\mathbf{t}_h,\boldsymbol{\sigma}_h,(\mathbf{u}_h,\gamma_h))\|_{\mathbf{X}} \;\le\; &\big\|\boldsymbol{\sigma}_h - \big\{\boldsymbol{\lambda}(\mathbf{t}_h)\,\mathrm{tr}(\mathbf{t}_h)\,\mathbb{I} + \boldsymbol{\mu}(\mathbf{t}_h)\,\mathbf{t}_h\big\}\big\|_{0,\Omega}\\ +\;\; \|R\|_{M_1'} \;+\; \|\mathbf{f} + \mathbf{div}\,\boldsymbol{\sigma}_h\|_{0,\Omega} \;+\; &\big\|\boldsymbol{\sigma}_h \;-\; \boldsymbol{\sigma}_h^{\tt t}\big\|_{0,\Omega}\,,\end{aligned} \tag{5.9}$$

where

$$R(\boldsymbol{\tau}) \;:=\; -\langle \boldsymbol{\tau}\,\boldsymbol{\nu}, \mathbf{g}\rangle_\Gamma \;+\; \int_\Omega (\mathbf{t}_h \,+\, \boldsymbol{\gamma}_h) : \boldsymbol{\tau} \;+\; \int_\Omega \mathbf{u}_h \cdot \mathbf{div}\,\boldsymbol{\tau} \qquad \forall\, \boldsymbol{\tau} \,\in\, M_1\,. \tag{5.10}$$

In addition, there holds

$$R(\boldsymbol{\tau}_h) \;=\; 0 \qquad \forall\, \boldsymbol{\tau}_h \,\in\, M_{1,h}\,.$$

Proof. We first proceed as in the proof of Lemma 2.2 (cf. (2.24)) and observe, thanks to the mean value theorem, that there exists a convex combination of $\mathbf{t}$ and $\mathbf{t}_h$, say $\tilde{\mathbf{r}}_h \,\in\, X_1$, such that

$$\mathcal{D}\mathbb{A}_1(\tilde{\mathbf{r}}_h)(\mathbf{t} - \mathbf{t}_h, \mathbf{s}) \;=\; [\mathbb{A}_1(\mathbf{t}), \mathbf{s}] \,-\, [\mathbb{A}_1(\mathbf{t}_h), \mathbf{s}] \qquad \forall\, \mathbf{s} \,\in\, X_1\,. \tag{5.11}$$

Then, applying (5.7)-(5.8) to the error $(\mathbf{r}, \boldsymbol{\zeta}, (\mathbf{w}, \xi)) \,:=\, (\mathbf{t}, \boldsymbol{\sigma}, (\mathbf{u}, \gamma)) - (\mathbf{t}_h, \boldsymbol{\sigma}_h, (\mathbf{u}_h, \gamma_h))$, and making use of the identity (5.11), the equations forming (2.8), and the definitions of the operators $\mathbb{A}_1$, $\mathbb{B}_1$, and $\mathbb{B}$ (cf. (2.7)), we find that

$$\begin{aligned} \tilde{C}\,\|(\mathbf{t}, \boldsymbol{\sigma}, (\mathbf{u}, \gamma)) - (\mathbf{t}_h, \boldsymbol{\sigma}_h, (\mathbf{u}_h, \gamma_h))\|_{\mathbf{X}} \,&\leq\, \sup_{(\mathbf{s},\boldsymbol{\tau},(\mathbf{v},\boldsymbol{\eta}))\,\in\,\mathbf{X}\backslash\{\mathbf{0}\}} \left\{ \frac{Q(\mathbf{s}) \,+\, R(\boldsymbol{\tau}) \,+\, S(\mathbf{v}, \boldsymbol{\eta})}{\|(\mathbf{s}, \boldsymbol{\tau}, (\mathbf{v}, \boldsymbol{\eta}))\|_{\mathbf{X}}} \right\} \\ &\leq\, \|Q\|_{X_1'} \,+\, \|R\|_{M_1'} \,+\, \|S\|_{M'}\,, \end{aligned} \tag{5.12}$$

where R is defined by (5.10), and $Q \in X_1'$ and $S \in M'$ are given by

$$\begin{aligned} Q(\mathbf{s}) \;&=\; \int_\Omega \big(\boldsymbol{\sigma}_h - \big\{\boldsymbol{\lambda}(\mathbf{t}_h)\,\mathrm{tr}(\mathbf{t}_h)\,\mathbb{I} \,+\, \boldsymbol{\mu}(\mathbf{t}_h)\,\mathbf{t}_h\big\}\big) : \mathbf{s} \qquad \forall\, \mathbf{s} \in X_1\,, \\ S(\mathbf{v}, \boldsymbol{\eta}) \;&=\; \int_\Omega \big\{\mathbf{f} \,+\, \mathbf{div}\,\boldsymbol{\sigma}_h\big\} \cdot \mathbf{v} \,+\, \int_\Omega \boldsymbol{\sigma}_h : \boldsymbol{\eta} \qquad\qquad \forall\,(\mathbf{v}, \boldsymbol{\eta}) \in M\,. \end{aligned}$$

It follows, using Cauchy-Schwarz's inequality and the fact that $\displaystyle\int_\Omega \boldsymbol{\sigma}_h : \boldsymbol{\eta} = \frac{1}{2}\int_\Omega \big(\boldsymbol{\sigma}_h - \boldsymbol{\sigma}_h^{\mathtt{t}}\big) : \boldsymbol{\eta}$, that

$$\|Q'\|_{M_1'} \,\leq\, \big\|\boldsymbol{\sigma}_h - \big\{\boldsymbol{\lambda}(\mathbf{t}_h)\,\mathrm{tr}(\mathbf{t}_h)\,\mathbb{I} \,+\, \boldsymbol{\mu}(\mathbf{t}_h)\,\mathbf{t}_h\big\}\big\|_{0,\Omega} \tag{5.13}$$

and

$$\|S'\|_{M'} \,\leq\, \|\mathbf{f} + \mathbf{div}\,\boldsymbol{\sigma}_h\|_{0,\Omega} \,+\, \|\boldsymbol{\sigma}_h \,-\, \boldsymbol{\sigma}_h^{\mathtt{t}}\|_{0,\Omega}\,. \tag{5.14}$$

In this way, (5.9) is a direct consequence of (5.12), (5.13) and (5.14). In turn, it is easy to see from the second equation of (2.27) that $R(\boldsymbol{\tau}_h) \,=\, 0 \quad \forall\, \boldsymbol{\tau}_h \in M_{1,h}$, which completes the proof.

□

It remains to bound $\|R\|_{M_1'}$ in (5.9), for which we proceed as in [41, Section 4.1] (see also [42, Section 4.2]) and use that $R(\boldsymbol{\tau}) = R(\boldsymbol{\tau} - \boldsymbol{\tau}_h)$ for each $\boldsymbol{\tau}_h \in M_{1,h}$. Hence, in order to define a suitable $\boldsymbol{\tau}_h \,\in\, M_{1,h}$ for the computation of $R(\boldsymbol{\tau} - \boldsymbol{\tau}_h)$ according to (5.10), we now let $I_h : H^1(\Omega) \longrightarrow X_h$ be the Clément interpolation operator (cf. [17]), where

$$X_h \,:=\, \big\{v_h \in C(\bar{\Omega}) : \quad v_h|_T \in P_1 \quad \forall\, T \in \mathcal{T}_h\big\}\,.$$

The following lemma establishes the local approximation properties of I_h.

LEMMA 5.2 *There exist constants $C_1, C_2 > 0$, independent of h, such that for all $v \in H^1(\Omega)$ there hold*

$$\|v - I_h(v)\|_{0,T} \,\leq\, C_1\,h_T\,\|v\|_{1,\Delta(T)} \quad \forall\, T \in \mathcal{T}_h,$$

and

$$\|v - I_h(v)\|_{0,e} \,\leq\, C_2\,h_e^{1/2}\,\|v\|_{1,\Delta(e)} \quad \forall\, e \in \mathcal{E}_h\,,$$

where $\Delta(T)$ and $\Delta(e)$ are the union of all elements intersecting with T and e, respectively.

Proof. See [17].

□

Next, given $\boldsymbol{\tau} \in M_1$, we consider the Helmholtz decomposition

$$\boldsymbol{\tau} \,=\, \underline{\mathbf{curl}}(\boldsymbol{\varphi}) \,+\, \nabla\mathbf{z}\,, \tag{5.15}$$

where $\boldsymbol{\varphi} \,:=\, (\varphi_1, \varphi_2)^{\mathrm{t}} \in \mathbf{H}^1(\Omega)$, with $\displaystyle\int_\Omega \varphi_1 \,=\, \int_\Omega \varphi_2 \,=\, 0$, $\mathbf{z} \in \mathbf{H}^2(\Omega)$, and

$$\|\boldsymbol{\varphi}\|_{1,\Omega} \,+\, \|\mathbf{z}\|_{2,\Omega} \,\leq\, C\,\|\boldsymbol{\tau}\|_{\mathbf{div};\Omega}\,. \tag{5.16}$$

Then, we set $\boldsymbol{\varphi}_h \,:=\, (I_h(\varphi_1), I_h(\varphi_2))^{\mathrm{t}}$ and define the discrete Helmholtz decomposition

$$\overline{\boldsymbol{\tau}}_h \,:=\, \underline{\mathbf{curl}}(\boldsymbol{\varphi}_h) \,+\, \mathcal{E}_h^k(\nabla\mathbf{z})\,, \tag{5.17}$$

where $\mathcal{E}_h^k : \mathbb{H}^1(\Omega) \longrightarrow \mathbb{RT}_k(\mathcal{T}_h)$ is the Raviart-Thomas interpolation operator (cf. (2.36), (2.37)). In this way, replacing $\boldsymbol{\tau}$ (cf. (5.15)) and $\overline{\boldsymbol{\tau}}_h$ (cf. (5.17)) into the expression $R(\boldsymbol{\tau}) = R(\boldsymbol{\tau} - \overline{\boldsymbol{\tau}}_h)$, observing that $\mathbf{div}(\nabla\mathbf{z}) \,=\, \mathbf{div}\,\boldsymbol{\tau}$, and noting, according to (2.38) and the definition of $\mathcal{P}_h^k$, that

$$\int_\Omega \mathbf{u}_h \cdot \mathbf{div}(\nabla\mathbf{z} - \mathcal{E}_h^k(\nabla\mathbf{z})) \,=\, \int_\Omega \mathbf{u}_h \cdot \big(\mathbf{div}\,\boldsymbol{\tau} \,-\, \mathcal{P}_h^k(\mathbf{div}\,\boldsymbol{\tau})\big) \,=\, 0\,,$$

we find that $R(\boldsymbol{\tau})$ can be decomposed as $R(\boldsymbol{\tau}) \,=\, R_1(\boldsymbol{\varphi}) \,+\, R_2(\mathbf{z})$, where

$$R_1(\boldsymbol{\varphi}) \,:=\, -\,\langle \underline{\mathbf{curl}}(\boldsymbol{\varphi} - \boldsymbol{\varphi}_h)\,\boldsymbol{\nu}, \mathbf{g}\rangle_\Gamma \,+\, \int_\Omega (\mathbf{t}_h \,+\, \boldsymbol{\gamma}_h) : \underline{\mathbf{curl}}(\boldsymbol{\varphi} - \boldsymbol{\varphi}_h)$$

and

$$R_2(\mathbf{z}) \,:=\, -\,\langle (\nabla\mathbf{z} - \mathcal{E}_h^k(\nabla\mathbf{z}))\,\boldsymbol{\nu}, \mathbf{g}\rangle_\Gamma \,+\, \int_\Omega (\mathbf{t}_h \,+\, \boldsymbol{\gamma}_h) : (\nabla\mathbf{z} - \mathcal{E}_h^k(\nabla\mathbf{z}))\,.$$

The following two lemmas provide upper bounds for $|R_1(\varphi)|$ and $|R_2(\mathbf{z})|$.

Lemma 5.3 *Assume that* $\mathbf{g} \in [H^1(\Gamma_D)]^2$. *Then there exists* $C > 0$, *independent of* h, *such that*

$$|R_1(\varphi)| \,\leq\, C \left\{ \sum_{T\in\mathcal{T}_h} \theta_{1,T}^2 \right\}^{1/2} \|\boldsymbol{\tau}\|_{\mathbf{div};\Omega}\,,$$

where

$$\theta_{1,T}^2 \,:=\, h_T^2\, \big\|\mathrm{curl}\{\mathbf{t}_h \,+\, \boldsymbol{\gamma}_h\}\big\|_{0,T}^2 \,+\, \sum_{e\in\mathcal{E}(T)\cap\mathcal{E}_h(\Omega)} h_e\, \|[(\mathbf{t}_h \,+\, \boldsymbol{\gamma}_h)\,s]\|_{0,e}^2$$
$$+\, \sum_{e\in\mathcal{E}(T)\cap\mathcal{E}_h(\Gamma)} h_e \left\| \frac{d\mathbf{g}}{ds} - (\mathbf{t}_h \,+\, \boldsymbol{\gamma}_h)\, s \right\|_{0,e}^2 .$$

Proof. It follows analogously to the proof of [41, Lemma 4.3] by employing $\mathbf{t}_h + \boldsymbol{\gamma}_h$ instead of just $\mathbf{t}_h$. The main tools employed are integration by parts, the Cauchy-Schwarz inequality, the approximation properties of the Clément interpolant (cf. Lemma 5.2), the fact that the number of triangles in $\Delta(T)$ and $\Delta(e)$ are bounded, and the estimate (5.16). We omit further details here.

□

Lemma 5.4 *There exists $C > 0$, independent of h, such that*

$$|R_2(\mathbf{z})| \,\le\, C \left\{ \sum_{T\in\mathcal{T}_h} \theta_{2,T}^2 \right\}^{1/2} \|\boldsymbol{\tau}\|_{\mathbf{div};\Omega}\,,$$

where

$$\theta_{2,T}^2 \,=\, h_T^2\, \|\nabla \mathbf{u}_h \,-\, (\mathbf{t}_h \,+\, \boldsymbol{\gamma}_h)\|_{0,T}^2 \,+ \sum_{e\in\mathcal{E}(T)\,\cap\,\mathcal{E}_h(\Gamma)} h_e\, \|\mathbf{g} - \mathbf{u}_h\|_{0,e}^2\,.$$

Proof. It follows analogously to the proof of [41, Lemma 4.4] by employing again $\mathbf{t}_h + \boldsymbol{\gamma}_h$ instead of just $\mathbf{t}_h$. In this case the main tools are given by the identities (2.36) and (2.37) characterizing $\mathcal{E}_h^k$, the Cauchy-Schwarz inequality, the approximation properties (2.42) and (2.40) (with $m = 1$), and the estimate (5.16). Further details are omitted here.

□

Finally, it follows from the decomposition of R and Lemmas 5.3 and 5.4 that

$$|\, R(\boldsymbol{\tau})\,| \,=\, |R(\boldsymbol{\tau} - \overline{\boldsymbol{\tau}}_h)| \,\le\, \left\{ \sum_{T\in\mathcal{T}_h} \left(\theta_{1,T}^2 + \theta_{2,T}^2\right) \right\}^{1/2} \|\boldsymbol{\tau}\|_{M_1} \qquad \forall\, \boldsymbol{\tau} \,\in\, M_1\,, \tag{5.18}$$

which, together with the estimate (5.9) (cf. Lemma 5.1), yields the reliability of $\boldsymbol{\theta}$.

5.3 Reliability of the a posteriori error estimator $\widetilde{\boldsymbol{\theta}}$

We now consider the augmented formulation (3.1) and let $\mathcal{M}$ be the linear operator obtained by adding the two equations of its left hand side, after replacing $\mathbb{A}_1$ within $\mathcal{A}$ (see (3.2)) by the Gâteaux derivative $\mathcal{D}\mathbb{A}_1(\tilde{\mathbf{r}})$ at any $\tilde{\mathbf{r}} \in X_1$, that is

$$\begin{aligned}
[\mathcal{M}((\mathbf{s},\boldsymbol{\tau}),(\mathbf{v},\boldsymbol{\eta})),((\mathbf{r},\boldsymbol{\zeta}),(\mathbf{w},\boldsymbol{\xi}))] \,&:=\, \mathcal{D}\mathbb{A}_1(\tilde{\mathbf{r}})(\mathbf{r},\mathbf{s} \,-\, \kappa_0\,\boldsymbol{\tau}) \,+\, [\mathbb{B}_1(\mathbf{s}),\boldsymbol{\zeta}] \,-\, [\mathbb{B}_1(\mathbf{r}),\boldsymbol{\tau}]\\
&+\kappa_0 \int_\Omega \boldsymbol{\zeta} : \boldsymbol{\tau} \,+\, [\mathcal{B}(\mathbf{s},\boldsymbol{\tau}),(\mathbf{w},\boldsymbol{\xi})] \,+\, [\mathcal{B}(\mathbf{r},\boldsymbol{\zeta}),(\mathbf{v},\boldsymbol{\eta})]
\end{aligned} \tag{5.19}$$

for all $((\mathbf{s},\boldsymbol{\tau}),(\mathbf{v},\boldsymbol{\eta}))$, $((\mathbf{r},\boldsymbol{\zeta}),(\mathbf{w},\boldsymbol{\xi})) \in X\times M$. Note that we have used here that the nonlinear operator $\mathcal{A}$ (cf. (3.2)) can be rewritten as

$$[\mathcal{A}(\mathbf{r},\boldsymbol{\zeta}),(\mathbf{s},\boldsymbol{\tau})] \,:=\, [\mathbb{A}_1(\mathbf{r}),\mathbf{s} \,-\, \kappa_0\,\boldsymbol{\tau}] \,+\, [\mathbb{B}_1(\mathbf{s}),\boldsymbol{\zeta}] \,-\, [\mathbb{B}_1(\mathbf{r}),\boldsymbol{\tau}] \,+\, \kappa_0 \int_\Omega \boldsymbol{\zeta} : \boldsymbol{\tau}\,. \tag{5.20}$$

Then, applying the continuous dependence result provided by the linear version of Theorem 3.1 (cf. (3.6)-(3.7) with $\mathcal{A}$ linear), which is actually the usual estimate provided by the Babuška-Brezzi theory (see, e.g. [47, Theorem 4.1, Chapter I]), and having in mind again the uniform estimates (2.16) and (2.17), we deduce that $\mathcal{M}$ satisfies a global inf-sup condition uniformly with respect to $\tilde{\mathbf{r}} \in X_1$, that is there exists a constant $\tilde{C} > 0$ such that

$$\tilde{C}\, \|((\mathbf{r},\boldsymbol{\zeta}),(\mathbf{w},\boldsymbol{\xi}))\|_{X\times M} \,\le \sup_{((\mathbf{s},\boldsymbol{\tau}),(\mathbf{v},\boldsymbol{\eta}))\in X\times M\setminus\{\mathbf{0}\}} \frac{[\mathcal{M}((\mathbf{s},\boldsymbol{\tau}),(\mathbf{v},\boldsymbol{\eta})),((\mathbf{r},\boldsymbol{\zeta}),(\mathbf{w},\boldsymbol{\xi}))]}{\|((\mathbf{s},\boldsymbol{\tau}),(\mathbf{v},\boldsymbol{\eta}))\|_{X\times M}} \tag{5.21}$$

for all $(\tilde{\mathbf{r}},((\mathbf{r},\boldsymbol{\zeta}),(\mathbf{w},\boldsymbol{\xi}))) \,\in\, X_1 \times (X \times M)$.

The analogue of Lemma 5.1 is established as follows.

LEMMA 5.5 *Let* $((\mathbf{t},\boldsymbol{\sigma}),(\mathbf{u},\boldsymbol{\gamma})) \in X \times M$ *and* $((\mathbf{t}_h,\boldsymbol{\sigma}_h),(\mathbf{u}_h,\boldsymbol{\gamma}_h)) \in X_h \times M_h$ *be the unique solutions of the continuous and discrete augmented formulations* (3.1) *and* (3.14)*, respectively. Then there exists* $C > 0$*, independent of* h*, such that*

$$\begin{aligned} C\,\|((\mathbf{t},\boldsymbol{\sigma}),(\mathbf{u},\boldsymbol{\gamma})) - ((\mathbf{t}_h,\boldsymbol{\sigma}_h),(\mathbf{u}_h,\boldsymbol{\gamma}_h))\|_{X\times M} \;&\le\; \big\|\boldsymbol{\sigma}_h - \big\{\boldsymbol{\lambda}(\mathbf{t}_h)\,\mathrm{tr}(\mathbf{t}_h)\,\mathbb{I} \,+\, \boldsymbol{\mu}(\mathbf{t}_h)\,\mathbf{t}_h\big\}\big\|_{0,\Omega} \\ +\quad \|\,R + \widetilde{R}\,\|_{M_1'} \;+\; \|\mathbf{f} + \mathbf{div}\,\boldsymbol{\sigma}_h\|_{0,\Omega} &\;+\; \big\|\boldsymbol{\sigma}_h \;-\; \boldsymbol{\sigma}_h^{\mathtt{t}}\big\|_{0,\Omega}\,, \end{aligned} \tag{5.22}$$

where R *is defined by* (5.10) *and*

$$\widetilde{R}(\boldsymbol{\tau}) \;:=\; \kappa_0 \int_\Omega \big(\boldsymbol{\sigma}_h \;-\; \big\{\boldsymbol{\lambda}(\mathbf{t}_h)\,\mathrm{tr}(\mathbf{t}_h)\,\mathbb{I} \,+\, \boldsymbol{\mu}(\mathbf{t}_h)\,\mathbf{t}_h\big\}\big) : \boldsymbol{\tau} \qquad \forall\,\boldsymbol{\tau} \,\in\, M_1\,. \tag{5.23}$$

In addition, there holds

$$R(\boldsymbol{\tau}_h) \,+\, \widetilde{R}(\boldsymbol{\tau}_h) \;=\; 0 \qquad \forall\,\boldsymbol{\tau}_h \,\in\, M_{1,h}\,.$$

Proof. We proceed analogously to the proof of Lemma 5.1, though we omit several similar details. Indeed, according to (5.19) and (5.21), we easily deduce that

$$\begin{aligned} &\tilde{C}\,\|((\mathbf{t},\boldsymbol{\sigma}),(\mathbf{u},\boldsymbol{\gamma})) - ((\mathbf{t}_h,\boldsymbol{\sigma}_h),(\mathbf{u}_h,\boldsymbol{\gamma}_h))\|_{X\times M} \\ &\qquad\le\; \sup_{((\mathbf{s},\boldsymbol{\tau}),(\mathbf{v},\boldsymbol{\eta}))\in X\times M\setminus\{\mathbf{0}\}} \left\{\frac{Q(\mathbf{s}) \,+\, R(\boldsymbol{\tau}) \,+\, \widetilde{R}(\boldsymbol{\tau}) \,+\, S(\mathbf{v},\boldsymbol{\eta})}{\|((\mathbf{s},\boldsymbol{\tau}),(\mathbf{v},\boldsymbol{\eta}))\|_{X\times M}}\right\} \\ &\qquad\le\; \|Q\|_{X_1'} \,+\, \|R + \widetilde{R}\|_{M_1'} \,+\, \|S\|_{M'}\,, \end{aligned} \tag{5.24}$$

where R and $\widetilde{R}$ are defined by (5.10) and (5.23), and $Q \in X_1'$ and $S \in M'$ are given as in the proof of Lemma 5.1, that is

$$\begin{aligned} Q(\mathbf{s}) \;&=\; \int_\Omega \big(\boldsymbol{\sigma}_h - \big\{\boldsymbol{\lambda}(\mathbf{t}_h)\,\mathrm{tr}(\mathbf{t}_h)\,\mathbb{I} \,+\, \boldsymbol{\mu}(\mathbf{t}_h)\,\mathbf{t}_h\big\}\big) : \mathbf{s} \qquad \forall\,\mathbf{s} \in X_1\,, \\ S(\mathbf{v},\boldsymbol{\eta}) \;&=\; \int_\Omega \big\{\mathbf{f} \,+\, \mathbf{div}\,\boldsymbol{\sigma}_h\big\}\cdot\mathbf{v} \,+\, \int_\Omega \boldsymbol{\sigma}_h : \boldsymbol{\eta} \qquad \forall\,(\mathbf{v},\boldsymbol{\eta}) \in M\,. \end{aligned}$$

The rest of the derivation of (5.22) is pretty straightforward from (5.24) and the above expressions for Q and S. Finally, taking $\mathbf{s} = \mathbf{0}$ and $\boldsymbol{\tau} = \boldsymbol{\tau}_h \in M_{1,h}$ in the first equation of (3.14), we deduce that $R + \widetilde{R}$ vanishes in $M_{1,h}$, which completes the proof.

□

We now aim to bound $\|R + \widetilde{R}\|_{M_1'}$ in (5.22) by proceeding similarly as we did before for $\|R\|_{M_1'}$. Indeed, we first note that $R(\boldsymbol{\tau}) + \widetilde{R}(\boldsymbol{\tau}) \;=\; R(\boldsymbol{\tau} - \boldsymbol{\tau}_h) + \widetilde{R}(\boldsymbol{\tau} - \boldsymbol{\tau}_h)$ for each $\boldsymbol{\tau}_h \in M_{1,h}$, and then employ again the Helmholtz decompositions (5.15) and (5.17) for rewriting $\boldsymbol{\tau}$ and introducing the particular tensor $\overline{\boldsymbol{\tau}}_h \in M_{1,h}$, respectively. In this way, since $R(\boldsymbol{\tau} - \overline{\boldsymbol{\tau}}_h)$ is already bounded by (5.18), it only remains to estimate the extra-term given by $\widetilde{R}(\boldsymbol{\tau} - \overline{\boldsymbol{\tau}}_h)$, which becomes

$$\begin{aligned} \widetilde{R}(\boldsymbol{\tau} - \overline{\boldsymbol{\tau}}_h) \;:=\; \kappa_0 &\int_\Omega \big(\boldsymbol{\sigma}_h - \big\{\boldsymbol{\lambda}(\mathbf{t}_h)\,\mathrm{tr}(\mathbf{t}_h)\,\mathbb{I} \,+\, \boldsymbol{\mu}(\mathbf{t}_h)\,\mathbf{t}_h\big\}\big) : \underline{\mathbf{curl}}(\boldsymbol{\varphi} - \boldsymbol{\varphi}_h) \\ +\,\kappa_0 &\int_\Omega \big(\boldsymbol{\sigma}_h - \big\{\boldsymbol{\lambda}(\mathbf{t}_h)\,\mathrm{tr}(\mathbf{t}_h)\,\mathbb{I} \,+\, \boldsymbol{\mu}(\mathbf{t}_h)\,\mathbf{t}_h\big\}\big) : (\nabla\mathbf{z} \,-\, \mathcal{E}_h^k(\nabla\mathbf{z}))\,. \end{aligned}$$

Moreover, by applying the same techniques employed to prove Lemmas 5.3 and 5.4 (see also [41, Lemmas 4.3 and 4.4] for further details), we arrive at the following estimate for $\widetilde{R}(\boldsymbol{\tau} - \overline{\boldsymbol{\tau}}_h)$.

Lemma 5.6 *There exists $C > 0$, independent of h, such that*

$$|\widetilde{R}(\boldsymbol{\tau} - \overline{\boldsymbol{\tau}}_h)| \,\le\, C \left\{ \sum_{T\in\mathcal{T}_h} \widetilde{\theta}_{1,T}^2 \right\}^{1/2} \|\boldsymbol{\tau}\|_{\mathbf{div};\Omega} \qquad \forall\, \boldsymbol{\tau} \in M_1\,, \tag{5.25}$$

where

$$\begin{aligned}\widetilde{\theta}_{1,T}^2 \,&=\, h_T^2\, \big\|\mathrm{curl}\big(\boldsymbol{\sigma}_h - \big\{\boldsymbol{\lambda}(\mathbf{t}_h)\,\mathrm{tr}(\mathbf{t}_h)\,\mathbb{I} \,+\, \boldsymbol{\mu}(\mathbf{t}_h)\,\mathbf{t}_h\big\}\big)\big\|_{0,T}^2 \\ &+\, \sum_{e\in\mathcal{E}(T)} h_e\, \big\|\big[\big(\boldsymbol{\sigma}_h - \big\{\boldsymbol{\lambda}(\mathbf{t}_h)\,\mathrm{tr}(\mathbf{t}_h)\,\mathbb{I} \,+\, \boldsymbol{\mu}(\mathbf{t}_h)\,\mathbf{t}_h\big\}\big)\, s\big]\big\|_{0,e}^2 \\ &+\, h_T^2\, \big\|\boldsymbol{\sigma}_h - \big\{\boldsymbol{\lambda}(\mathbf{t}_h)\,\mathrm{tr}(\mathbf{t}_h)\,\mathbb{I} \,+\, \boldsymbol{\mu}(\mathbf{t}_h)\,\mathbf{t}_h\big\}\big\|_{0,T}^2\,.\end{aligned}$$

As a consequence of (5.18) and (5.25) we deduce that

$$\begin{aligned}|R(\boldsymbol{\tau}) \,+\, \widetilde{R}(\boldsymbol{\tau})| &= |R(\boldsymbol{\tau} - \overline{\boldsymbol{\tau}}_h) \,+\, \widetilde{R}(\boldsymbol{\tau} - \overline{\boldsymbol{\tau}}_h)| \\ &\le\, \left\{ \sum_{T\in\mathcal{T}_h} \Big(\theta_{1,T}^2 \,+\, \theta_{2,T}^2 \,+\, \widetilde{\theta}_{1,T}^2\Big) \right\}^{1/2} \|\boldsymbol{\tau}\|_{M_1} \quad \forall\, \boldsymbol{\tau} \in M_1\,,\end{aligned}$$

which, together with (5.22), and noting that the third term in the definition of $\widetilde{\theta}_{1,T}^2$ is certainly dominated by $\big\|\boldsymbol{\sigma}_h - \big\{\boldsymbol{\lambda}(\mathbf{t}_h)\,\mathrm{tr}(\mathbf{t}_h)\,\mathbb{I} \,+\, \boldsymbol{\mu}(\mathbf{t}_h)\,\mathbf{t}_h\big\}\big\|_{0,T}^2$, yields the reliability of the a posteriori error estimator $\widetilde{\boldsymbol{\theta}}$.

5.4 Reliability of the a posteriori error estimator $\widehat{\boldsymbol{\theta}}$

Following the same reasoning of the previous sections, we now consider the fully augmented formulation (4.1) and let $\mathcal{N}$ be the linear operator obtained by replacing $\mathbb{A}_1$ within $\mathbb{A}$ (see (4.2) and (5.20)) by the Gâteaux derivative $\mathcal{D}\mathbb{A}_1(\tilde{\mathbf{r}})$ at any $\tilde{\mathbf{r}} \in X_1$, that is

$$\begin{aligned}[\mathcal{N}(\mathbf{s},\boldsymbol{\tau},\mathbf{v},\boldsymbol{\eta}),(\mathbf{r},\boldsymbol{\zeta},\mathbf{w},\boldsymbol{\xi})] \,&:=\, \mathcal{D}\mathbb{A}_1(\tilde{\mathbf{r}})(\mathbf{r},\mathbf{s} \,-\, \kappa_0\,\boldsymbol{\tau}) \,+\, [\mathbb{B}_1(\mathbf{s}),\boldsymbol{\zeta}] \,-\, [\mathbb{B}_1(\mathbf{r}),\boldsymbol{\tau}] \\ +\,\kappa_0 \int_\Omega \boldsymbol{\zeta} : \boldsymbol{\tau} \,+\, [\mathcal{B}(\mathbf{s},\boldsymbol{\tau}),(\mathbf{w},\boldsymbol{\xi})] \,&-\, [\mathcal{B}(\mathbf{r},\boldsymbol{\zeta}),(\mathbf{v},\boldsymbol{\eta})] \,+\, \kappa_1 \int_\Omega \mathbf{div}\,\boldsymbol{\zeta}\cdot\mathbf{div}\,\boldsymbol{\tau} \\ +\,\kappa_2 \int_\Omega (\mathbf{e}(\mathbf{w}) \,-\, \mathbf{r}) : \mathbf{e}(\mathbf{v}) \,+\, \kappa_3 \int_\Omega \Big\{\boldsymbol{\xi} \,&-\, \frac{1}{2}(\nabla\mathbf{w} \,-\, (\nabla\mathbf{w})^{\mathtt{t}})\Big\} : \boldsymbol{\eta} \,+\, \kappa_4 \int_{\Gamma_D} \mathbf{w}\cdot\mathbf{v}\end{aligned} \tag{5.26}$$

for all $(\mathbf{s},\boldsymbol{\tau},\mathbf{v},\boldsymbol{\eta})$, $(\mathbf{r},\boldsymbol{\zeta},\mathbf{w},\boldsymbol{\xi}) \in \mathbb{X}$. Then, applying the continuous dependence result provided by the linear version of Theorem 4.1 (cf. (4.3) with $\mathbb{A}$ linear), we deduce that $\mathcal{N}$ satisfies a global inf-sup condition uniformly with respect to $\tilde{\mathbf{r}} \in X_1$, which means that there exists a constant $\tilde{C} > 0$ such that

$$\tilde{C}\,\|(\mathbf{r},\boldsymbol{\zeta},\mathbf{w},\boldsymbol{\xi})\|_{\mathbb{X}} \,\le\, \sup_{(\mathbf{s},\boldsymbol{\tau},\mathbf{v},\boldsymbol{\eta})\in\mathbb{X}\setminus\{\mathbf{0}\}} \frac{[\mathcal{N}(\mathbf{s},\boldsymbol{\tau},\mathbf{v},\boldsymbol{\eta}),(\mathbf{r},\boldsymbol{\zeta},\mathbf{w},\boldsymbol{\xi})]}{\|(\mathbf{s},\boldsymbol{\tau},\mathbf{v},\boldsymbol{\eta})\|_{\mathbb{X}}} \tag{5.27}$$

for all $(\tilde{\mathbf{r}},(\mathbf{r},\boldsymbol{\zeta},\mathbf{w},\boldsymbol{\xi})) \,\in\, X_1 \times \mathbb{X}$.

The analogue of Lemma 5.5 is established as follows.

LEMMA 5.7 *Let* $(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma}) \in \mathbb{X}$ *and* $(\mathbf{t}_h,\boldsymbol{\sigma}_h,\mathbf{u}_h,\boldsymbol{\gamma}_h) \in \mathbb{X}_h$ *be the unique solutions of the continuous and discrete fully augmented formulations* (4.1) *and* (4.6)*, respectively. Then there exists* $C > 0$*, independent of* h*, such that*

$$\begin{aligned} C\,\|(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma}) - (\mathbf{t}_h,\boldsymbol{\sigma}_h,\mathbf{u}_h,\boldsymbol{\gamma}_h)\|_{\mathbb{X}} \;&\le\; \big\|\boldsymbol{\sigma}_h - \big\{\boldsymbol{\lambda}(\mathbf{t}_h)\,\mathrm{tr}(\mathbf{t}_h)\mathbb{I} \,+\, \boldsymbol{\mu}(\mathbf{t}_h)\,\mathbf{t}_h\big\}\big\|_{0,\Omega} \\ +\;\; \|R \,+\, \widetilde{R} \,+\, \widehat{R}\|_{M_1'} \,+\, &\|\mathbf{f} + \mathbf{div}\,\boldsymbol{\sigma}_h\|_{0,\Omega} \,+\, \big\|\boldsymbol{\sigma}_h \,-\, \boldsymbol{\sigma}_h^{\mathtt{t}}\big\|_{0,\Omega} \\ +\;\; \|\mathbf{e}(\mathbf{u}_h) \,-\, \mathbf{t}_h\|_{0,\Omega} \,+\, &\Big\|\boldsymbol{\gamma}_h \,-\, \frac{1}{2}(\nabla\mathbf{u}_h \,-\, (\nabla\mathbf{u}_h)^{\mathtt{t}})\Big\|_{0,\Omega} \,+\, \|\mathbf{u}_h - \mathbf{g}\|_{0,\Gamma} \end{aligned} \tag{5.28}$$

where R *and* $\widetilde{R}$ *are defined by* (5.10) *and* (5.23)*, respectively, and*

$$\widehat{R}(\boldsymbol{\tau}) \;:=\; \kappa_1 \int_\Omega \big(\mathbf{div}\,\boldsymbol{\sigma}_h \,+\, \mathbf{f}\big)\cdot\mathbf{div}\,\boldsymbol{\tau} \qquad \forall\,\boldsymbol{\tau} \in M_1\,. \tag{5.29}$$

In addition, there holds

$$R(\boldsymbol{\tau}_h) \,+\, \widetilde{R}(\boldsymbol{\tau}_h) \,+\, \widehat{R}(\boldsymbol{\tau}_h) \;=\; 0 \qquad \forall\,\boldsymbol{\tau}_h \in M_{1,h}\,.$$

Proof. We proceed analogously to the proofs of Lemmas 5.1 and 5.5. Indeed, applying now (5.27) and (5.26), we deduce that

$$\begin{aligned} &\tilde{C}\,\|(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma}) - (\mathbf{t}_h,\boldsymbol{\sigma}_h,\mathbf{u}_h,\boldsymbol{\gamma}_h)\|_{\mathbb{X}} \\ &\quad\le\; \sup_{(\mathbf{s},\boldsymbol{\tau},\mathbf{v},\boldsymbol{\eta})\in\mathbb{X}\setminus\{\mathbf{0}\}} \left\{\frac{Q(\mathbf{s}) \,+\, R(\boldsymbol{\tau}) \,+\, \widetilde{R}(\boldsymbol{\tau}) \,+\, \widehat{R}(\boldsymbol{\tau}) \,+\, \widehat{S}(\mathbf{v},\boldsymbol{\eta})}{\|(\mathbf{s},\boldsymbol{\tau},\mathbf{v},\boldsymbol{\eta})\|_{\mathbb{X}}}\right\} \\ &\quad\le\; \|Q\|_{X_1'} \,+\, \|R + \widetilde{R} + \widehat{R}\|_{M_1'} \,+\, \|\widehat{S}\|_{M'}\,, \end{aligned} \tag{5.30}$$

where R, $\widetilde{R}$, and $\widehat{R}$ are defined by (5.10), (5.23), and (5.29), respectively, $Q \in X_1'$ is given as in the proof of Lemma 5.5, that is

$$Q(\mathbf{s}) \;=\; \int_\Omega \big(\boldsymbol{\sigma}_h - \big\{\boldsymbol{\lambda}(\mathbf{t}_h)\,\mathrm{tr}(\mathbf{t}_h)\,\mathbb{I} \,+\, \boldsymbol{\mu}(\mathbf{t}_h)\,\mathbf{t}_h\big\}\big) : \mathbf{s} \qquad \forall\,\mathbf{s} \in X_1\,,$$

and

$$\begin{aligned} \widehat{S}(\mathbf{v},\boldsymbol{\eta}) \;&=\; \int_\Omega \big\{\mathbf{f} \,+\, \mathbf{div}\,\boldsymbol{\sigma}_h\big\}\cdot\mathbf{v} \,+\, \int_\Omega \boldsymbol{\sigma}_h : \boldsymbol{\eta} \,+\, \kappa_2\int_\Omega (\mathbf{e}(\mathbf{u}_h) \,-\, \mathbf{t}_h) : \mathbf{e}(\mathbf{v}) \\ +\;\; \kappa_3 \int_\Omega &\Big\{\boldsymbol{\gamma}_h \,-\, \frac{1}{2}(\nabla\mathbf{u}_h \,-\, (\nabla\mathbf{u}_h)^{\mathtt{t}})\Big\} : \boldsymbol{\eta} \,+\, \kappa_4 \int_\Gamma (\mathbf{u}_h \,-\, \mathbf{g})\cdot\mathbf{v} \qquad \forall\,(\mathbf{v},\boldsymbol{\eta}) \in M \end{aligned}$$

The rest of the derivation of (5.28) follows from (5.30) and the application of the Cauchy-Schwarz inequality to the above expressions for Q and $\widehat{S}$. In particular, the fact that the test functions $\mathbf{v}$ belong now to $\mathbf{H}^1(\Omega)$, and the corresponding trace theorem, imply that

$$\left|\int_\Gamma (\mathbf{u}_h \,-\, \mathbf{g})\cdot\mathbf{v}\right| \;\le\; \|\mathbf{u}_h \,-\, \mathbf{g}\|_{0,\Gamma}\,\|\mathbf{v}\|_{0,\Gamma} \;\le\; c\,\|\mathbf{u}_h \,-\, \mathbf{g}\|_{0,\Gamma}\,\|\mathbf{v}\|_{1,\Omega}\,.$$

Finally, it is straightforward to see that, taking $\mathbf{s} = \mathbf{0}$, $\mathbf{v} = \mathbf{0}$, $\boldsymbol{\eta} = \mathbf{0}$ and $\boldsymbol{\tau} = \boldsymbol{\tau}_h \in M_{1,h}$ in (4.6), we find that $R + \widetilde{R} + \widehat{R}$ vanishes in $M_{1,h}$, which ends the proof.

□

It remains to bound $\|R+\widetilde{R}+\widehat{R}\|_{M_1'}$ in (5.28), for which we proceed as we did before for $\|R\|_{M_1'}$ and $\|R+\widetilde{R}\|_{M_1'}$. In other words, we now use that $R(\boldsymbol{\tau})+\widetilde{R}(\boldsymbol{\tau})+\widehat{R}(\boldsymbol{\tau}) \,=\, R(\boldsymbol{\tau}-\boldsymbol{\tau}_h)+\widetilde{R}(\boldsymbol{\tau}-\boldsymbol{\tau}_h)+\widehat{R}(\boldsymbol{\tau}-\boldsymbol{\tau}_h)$ for each $\boldsymbol{\tau}_h \in M_{1,h}$, and then employ once again the Helmholtz decompositions (5.15) and (5.17). In this way, since $R(\boldsymbol{\tau}-\overline{\boldsymbol{\tau}}_h)$ and $\widetilde{R}(\boldsymbol{\tau}-\overline{\boldsymbol{\tau}}_h)$ are already bounded by (5.18) and (5.25), we just need to estimate the extra-term given by $\widehat{R}(\boldsymbol{\tau}-\overline{\boldsymbol{\tau}}_h)$, which is done as follows.

LEMMA 5.8 *There exists $C>0$, independent of h, such that*

$$|\widehat{R}(\boldsymbol{\tau}-\overline{\boldsymbol{\tau}}_h)| \,\le\, C \left\{ \sum_{T\in\mathcal{T}_h} \widehat{\theta}_{1,T}^2 \right\}^{1/2} \|\boldsymbol{\tau}\|_{\mathbf{div};\Omega}\,, \tag{5.31}$$

where

$$\widehat{\theta}_{1,T}^2 \,:=\, h_T^2\, \|\mathbf{div}\,\boldsymbol{\sigma}_h \,+\, \mathbf{f}\|_{0,T}^2\,.$$

Proof. It suffices to observe, having in mind (5.15) and (5.17), that

$$\widehat{R}(\boldsymbol{\tau}-\overline{\boldsymbol{\tau}}_h) \,=\, \kappa_1 \int_\Omega (\mathbf{div}\,\boldsymbol{\sigma}_h \,+\, \mathbf{f})\cdot \mathbf{div}\,(\nabla \mathbf{z} - \mathcal{E}_h^k(\nabla\mathbf{z}))\,,$$

and then apply the Cauchy-Schwarz inequality, (2.38), and (2.39).

□

As a consequence of (5.18), (5.25), and (5.31) we deduce that

$$\begin{aligned} |\,R(\boldsymbol{\tau}) \,+\, \widetilde{R}(\boldsymbol{\tau}) \,+\, \widehat{R}(\boldsymbol{\tau})\,| \,&=\, |\,R(\boldsymbol{\tau}-\overline{\boldsymbol{\tau}}_h) \,+\, \widetilde{R}(\boldsymbol{\tau}-\overline{\boldsymbol{\tau}}_h) \,+\, \widehat{R}(\boldsymbol{\tau}-\overline{\boldsymbol{\tau}}_h)\,| \\ &\le\, \left\{ \sum_{T\in\mathcal{T}_h} \left(\theta_{1,T}^2 \,+\, \theta_{2,T}^2 \,+\, \widetilde{\theta}_{1,T}^2 \,+\, \widehat{\theta}_{1,T}^2 \right) \right\}^{1/2} \|\boldsymbol{\tau}\|_{M_1} \qquad \forall\, \boldsymbol{\tau}\in M_1\,, \end{aligned}$$

which, replaced back into (5.28) for estimating $\|R \,+\, \widetilde{R} \,+\, \widehat{R}|_{M_1'}$, and noting that $h_T^2\, \|\mathbf{div}\,\boldsymbol{\sigma}_h \,+\, \mathbf{f}\|_{0,T}^2$ and $h_e\, \|\mathbf{g}-\mathbf{u}_h\|_{0,e}^2$ are dominated by $\|\mathbf{div}\,\boldsymbol{\sigma}_h \,+\, \mathbf{f}\|_{0,T}^2$ and $\|\mathbf{g}-\mathbf{u}_h\|_{0,e}^2$, respectively, yields the reliability of the a posteriori error estimator $\widehat{\boldsymbol{\theta}}$.

At this point we find it important to remark that the derivation of $\widehat{\boldsymbol{\theta}}$ does not take into account that actually $\mathbf{u}_h$ also belongs to $\mathbf{H}^1(\Omega)$. To this respect, we show next that this fact allows to simplify the upper bound of $\|R \,+\, \widetilde{R} \,+\, \widehat{R}\|_{M_1'}$ (cf. (5.28)), which yields a simpler reliable and efficient a posteriori error estimator. However, unless the Dirichlet datum is homogeneous, this alternative estimator does not become localizable, which makes it unsuitable for adaptive computations. More precisely, integrating by parts the third term in the definition of R (cf. (5.10)), we find that

$$R(\boldsymbol{\tau}) \,=\, \int_\Omega \big(\mathbf{t}_h + \boldsymbol{\gamma}_h \,-\, \nabla\mathbf{u}_h\big) : \boldsymbol{\tau} \,+\, \langle \boldsymbol{\tau}\,\boldsymbol{\nu}, \mathbf{u}_h - \mathbf{g}\rangle_\Gamma \qquad \forall\, \boldsymbol{\tau}\in M_1\,,$$

which gives

$$\|R\|_{M_1'} \,\le\, C \Big\{ \|\mathbf{t}_h + \boldsymbol{\gamma}_h \,-\, \nabla\mathbf{u}_h\|_{0,\Omega} \,+\, \|\mathbf{u}_h - \mathbf{g}\|_{1/2,\Gamma} \Big\}\,.$$

In turn, simple applications of the Cauchy-Schwarz inequality in (5.23) and (5.29) imply, respectively,

$$\|\widetilde{R}\|_{M_1'} \,\le\, \kappa_0\, \big\|\boldsymbol{\sigma}_h - \big\{\boldsymbol{\lambda}(\mathbf{t}_h)\,\mathrm{tr}(\mathbf{t}_h)\mathbb{I} \,+\, \boldsymbol{\mu}(\mathbf{t}_h)\,\mathbf{t}_h\big\}\big\|_{0,\Omega} \quad \text{and} \quad \|\widehat{R}\|_{M_1'} \,\le\, \kappa_1\, \|\mathbf{f} + \mathbf{div}\,\boldsymbol{\sigma}_h\|_{0,\Omega}\,.$$

In this way, employing the above estimates to bound $\|R + \widetilde{R} + \widehat{R}\|_{M_1'}$ in (5.28), we arrive at

$$\begin{aligned} C\,\|(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma}) - (\mathbf{t}_h,\boldsymbol{\sigma}_h,\mathbf{u}_h,\boldsymbol{\gamma}_h)\|_{\mathbb{X}} \;\le\; & \big\|\boldsymbol{\sigma}_h - \big\{\boldsymbol{\lambda}(\mathbf{t}_h)\,\mathrm{tr}(\mathbf{t}_h)\mathbb{I} \,+\, \boldsymbol{\mu}(\mathbf{t}_h)\,\mathbf{t}_h\big\}\big\|_{0,\Omega} \\ + \quad \|\mathbf{t}_h + \boldsymbol{\gamma}_h \,-\, \nabla\mathbf{u}_h\|_{0,\Omega} \,+\, & \|\mathbf{f} + \mathbf{div}\,\boldsymbol{\sigma}_h\|_{0,\Omega} \,+\, \big\|\boldsymbol{\sigma}_h \,-\, \boldsymbol{\sigma}_h^{\mathtt{t}}\big\|_{0,\Omega} \\ + \quad \|\mathbf{e}(\mathbf{u}_h) \,-\, \mathbf{t}_h\|_{0,\Omega} \,+\, \Big\|\boldsymbol{\gamma}_h \,-\, \frac{1}{2}(\nabla\mathbf{u}_h \,-\, (\nabla\mathbf{u}_h)^{\mathtt{t}})\Big\|_{0,\Omega} & \,+\, \|\mathbf{u}_h - \mathbf{g}\|_{1/2,\Gamma}\,, \end{aligned} \tag{5.32}$$

from which it is clear that the last term on the right hand side is not localizable. Certainly, one could use interpolation results to handle $\|\mathbf{u}_h - \mathbf{g}\|_{1/2,\Gamma}$ in terms of local terms, but then it is easy to see that the resulting estimator does not become efficient. Nevertheless, if the Dirichlet datum $\mathbf{g}$ vanishes on Γ, the last term in (5.32) dissapears, and the above reduces to

$$\|(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma}) - (\mathbf{t}_h,\boldsymbol{\sigma}_h,\mathbf{u}_h,\boldsymbol{\gamma}_h)\|_{\mathbb{X}} \;\le\; C\,\overline{\boldsymbol{\theta}} \;:=\; C\,\left\{\sum_{T\in\mathcal{T}_h} \overline{\theta}_T^2\right\}^{1/2}\,,$$

where

$$\begin{aligned} \overline{\theta}_T^2 \;:=\; & \big\|\boldsymbol{\sigma}_h - \big\{\boldsymbol{\lambda}(\mathbf{t}_h)\,\mathrm{tr}(\mathbf{t}_h)\mathbb{I} \,+\, \boldsymbol{\mu}(\mathbf{t}_h)\,\mathbf{t}_h\big\}\big\|_{0,T}^2 \\ & + \quad \|\mathbf{t}_h + \boldsymbol{\gamma}_h \,-\, \nabla\mathbf{u}_h\|_{0,T}^2 \,+\, \|\mathbf{f} + \mathbf{div}\,\boldsymbol{\sigma}_h\|_{0,T}^2 \,+\, \big\|\boldsymbol{\sigma}_h \,-\, \boldsymbol{\sigma}_h^{\mathtt{t}}\big\|_{0,T}^2 \\ & + \quad \|\mathbf{e}(\mathbf{u}_h) \,-\, \mathbf{t}_h\|_{0,T}^2 \,+\, \Big\|\boldsymbol{\gamma}_h \,-\, \frac{1}{2}(\nabla\mathbf{u}_h \,-\, (\nabla\mathbf{u}_h)^{\mathtt{t}})\Big\|_{0,T}^2 . \end{aligned} \tag{5.33}$$

The efficiency of $\overline{\boldsymbol{\theta}}$, which is quite straightforward, is briefly mentioned at the end of Section 5.5.

5.5 Efficiency of the a posteriori error estimators $\boldsymbol{\theta}$, $\widetilde{\boldsymbol{\theta}}$ and $\widehat{\boldsymbol{\theta}}$

In this section we establish the efficiency of our main a posteriori error estimators $\boldsymbol{\theta}$, $\widetilde{\boldsymbol{\theta}}$ and $\widehat{\boldsymbol{\theta}}$ (lower bounds in (5.4), (5.5) and (5.6), respectively). In other words, we provide suitable upper bounds for the eight terms defining the local error indicator θ_T^2 (cf. (5.1)), for the remaining two terms completing the definition of the local error indicator $\widetilde{\theta}_T^2$ (cf. (5.2)) and for the remaining three terms completing the local error indicator $\widehat{\theta}_T^2$ (cf. (5.3)). For this purpose, we first notice that the converses of the derivations of (2.8), (3.1) and (4.1) from (2.1) hold true. Indeed, it is not difficult to prove, applying integration by parts backwardly and using appropriate test functions, that the unique solution $(\mathbf{t},\boldsymbol{\sigma},(\mathbf{u},\boldsymbol{\gamma})) \in X_1 \times M_1 \times M$ of (2.8) (which is easily shown to coincide with that of (3.1) and (4.1)) solves the original problem (2.1).

We begin with three simple estimates. Since $\mathbf{f} \,=\, -\,\mathbf{div}\,\boldsymbol{\sigma}$ in Ω, it is clear that

$$\|\mathbf{f} \,+\, \mathbf{div}\,\boldsymbol{\sigma}_h\|_{0,T} \;=\; \big\|\mathbf{div}\,(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\big\|_{0,T}\,. \tag{5.34}$$

In addition, using that $\boldsymbol{\sigma} \,=\, \boldsymbol{\lambda}(\mathbf{t})\,\mathrm{tr}(\mathbf{t})\mathbb{I} \,+\, \boldsymbol{\mu}(\mathbf{t})\,\mathbf{t}$ in Ω and applying the Lipschitz-continuity of $\mathbb{A}_1$ (cf. Lemma 2.2), but restricted to the triangle $T \in \mathcal{T}_h$ instead of Ω, we deduce that

$$\big\|\boldsymbol{\sigma}_h \,-\, \big\{\boldsymbol{\lambda}(\mathbf{t}_h)\,\mathrm{tr}(\mathbf{t}_h)\,\mathbb{I} \,+\, \boldsymbol{\mu}(\mathbf{t}_h)\,\mathbf{t}_h\big\}\big\|_{0,T} \le c\Big\{\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,T} \,+\, \|\mathbf{t} - \mathbf{t}_h\|_{0,T}\Big\}. \tag{5.35}$$

Furthermore, using the symmetry of $\boldsymbol{\sigma}$, we easily find that

$$\big\|\boldsymbol{\sigma}_h \,-\, \boldsymbol{\sigma}_h^{\mathtt{t}}\big\|_{0,T} \;\le\; 2\,\|\boldsymbol{\sigma} \,-\, \boldsymbol{\sigma}\|_{0,T}\,. \tag{5.36}$$

Next, in order to bound the terms involving the mesh parameters h_T and h_e, we make use of the general results and estimates available in the analysis of related linear problems (see, e.g. [41, Section 4.2]). The techniques applied there are based on triangle-bubble and edge-bubble functions, extension operators, and discrete trace and inverse inequalities. For further details on these tools we refer particularly to [41, Lemmas 4.7 and 4.8, and eq. (4.34)].

We have the following efficiency estimates.

Lemma 5.9 *There exist C_1, $C_2 > 0$, independent of h, such that*

$$h_T^2 \, \left\|\operatorname{curl}\{\mathbf{t}_h + \boldsymbol{\gamma}_h\}\right\|_{0,T}^2 \,\le\, C_1 \left\{\|\mathbf{t} - \mathbf{t}_h\|_{0,T}^2 + \|\boldsymbol{\gamma} - \boldsymbol{\gamma}_h\|_{0,T}^2\right\} \qquad \forall\, T \in \mathcal{T}_h\,,$$

and

$$h_e \, \| \, [(\mathbf{t}_h + \boldsymbol{\gamma}_h)\, s] \, \|_{0,e}^2 \,\le\, C_2 \left\{\| \, \mathbf{t} - \mathbf{t}_h \, \|_{0,\omega_e}^2 + \| \, \boldsymbol{\gamma} - \boldsymbol{\gamma}_h \, \|_{0,\omega_e}^2\right\} \qquad \forall\, e \in \mathcal{E}_h(\Omega)\,,$$

where $\omega_e := \cup\big\{T \in \mathcal{T}_h : \quad e \in \mathcal{E}(T)\big\}$.

Proof. It suffices to apply the general results stated in [41, Lemmas 4.9 and 4.10] to $\boldsymbol{\rho}_h = \mathbf{t}_h + \boldsymbol{\gamma}_h$ and $\boldsymbol{\rho} = \mathbf{t} + \boldsymbol{\gamma} = \nabla\mathbf{u}$, noting that $\operatorname{curl}(\boldsymbol{\rho}) = \operatorname{curl}(\nabla\mathbf{u}) = \mathbf{0}$ in Ω (cf. (2.4) and (2.5)).

□

Lemma 5.10 *There exists $C_3 > 0$, independent of h, such that*

$$h_T^2 \, \|\nabla\mathbf{u}_h - (\mathbf{t}_h + \boldsymbol{\gamma}_h)\|_{0,T}^2 \,\le\, C_3 \left\{\| \, \mathbf{u} - \mathbf{u}_h \, \|_{0,T}^2 + h_T^2 \, \| \, \mathbf{t} - \mathbf{t}_h \, \|_{0,T}^2 + h_T^2 \, \| \, \boldsymbol{\gamma} - \boldsymbol{\gamma}_h \, \|_{0,T}^2\right\} \tag{5.37}$$

for all $T \in \mathcal{T}_h$.

Proof. It follows from the proof of [41, Lemma 4.13], which itself is a slight modification of the proof of [13, Lemma 6.3], by replacing the tensor utilized there by $\nabla\mathbf{u}_h - (\mathbf{t}_h + \boldsymbol{\gamma}_h)$, and recalling that $\nabla\mathbf{u} = \mathbf{t} + \boldsymbol{\gamma}$.

□

Lemma 5.11 *Assume that* $\mathbf{g}$ *is piecewise polynomial. Then there exists $C_4 > 0$, independent of h, such that*

$$h_e \left\| \frac{d\mathbf{g}}{ds} - (\mathbf{t}_h + \boldsymbol{\gamma}_h) s \right\|_{0,e}^2 \,\le\, C_4 \left\{\|\mathbf{t} - \mathbf{t}_h\|_{0,T}^2 + \|\boldsymbol{\gamma} - \boldsymbol{\gamma}_h\|_{0,T}^2\right\} \qquad \forall\, e \in \mathcal{E}_h(\Gamma)\,, \tag{5.38}$$

where T is the triangle of $\mathcal{T}_h$ having e as an edge.

Proof. It suffices to modify the proof of [41, Lemma 4.15], by using $\dfrac{d\mathbf{g}}{d\mathbf{s}} - (\mathbf{t}_h + \boldsymbol{\gamma}_h)\,\mathbf{s}$ instead of $\dfrac{d\mathbf{g}}{d\mathbf{s}} - \dfrac{1}{2\mu}\boldsymbol{\sigma}_h^{\mathtt{t}}\,\mathbf{s}$, and noting in the present case that $\dfrac{d\mathbf{g}}{d\mathbf{s}} = (\nabla\mathbf{u})\,\mathbf{s} = (\mathbf{t} + \boldsymbol{\gamma})\,\mathbf{s}$ on Γ.

□

Lemma 5.12 *There exists $C_5 > 0$, independent of h, such that*

$$h_e \, \|\mathbf{g} - \mathbf{u}_h\|_{0,e}^2 \,\le\, C_5 \left\{\|\mathbf{u} - \mathbf{u}_h\|_{0,T}^2 + h_T^2 \, \|\mathbf{t} - \mathbf{t}_h\|_{0,T}^2 + h_T^2 \, \|\boldsymbol{\gamma} - \boldsymbol{\gamma}_h\|_{0,T}^2\right\} \qquad \forall\, e \in \mathcal{E}_h(\Gamma)\,,$$

where T is the triangle of $\mathcal{T}_h$ having e as an edge.

Proof. Similarly to the previous lemmas, it follows as in the proof of [41, Lemma 4.14] by utilizing the tensor $\nabla \mathbf{u}_h - (\mathbf{t}_h + \boldsymbol{\gamma}_h)$, and then using that $\nabla \mathbf{u} = \mathbf{t} + \boldsymbol{\gamma}$ in Ω. At the end, the above estimate (5.37) for $h_T^2 \, \|\nabla \mathbf{u}_h - (\mathbf{t}_h + \boldsymbol{\gamma}_h)\|_{0,T}^2$ is also employed.

□

We remark here that if $\mathbf{g}$ were not piecewise polynomial but sufficiently smooth, then higher order terms given by the errors arising from suitable polynomial approximations would appear in (5.38). This explains the eventual expression h.o.t. in (5.4). In this way, the efficiency of $\boldsymbol{\theta}$ follows straightforwardly from estimates (5.34), (5.35) and (5.36), together with Lemmas 5.9 throughout 5.12, after summing up over $T \in \mathcal{T}_h$ and using that the number of triangles on each domain ω_e is bounded by two.

Next, for the efficiency of $\widetilde{\boldsymbol{\theta}}$ it only remains to provide upper bounds for the two terms completing the definition of the local error indicator $\widetilde{\theta}_T^2$ (cf. (5.2)), which is established in the following lemma.

Lemma 5.13 *There exist C_6, $C_7 > 0$, independent of h, such that*

$$h_T^2 \, \left\|\operatorname{curl}\left(\boldsymbol{\sigma}_h - \left\{\boldsymbol{\lambda}(\mathbf{t}_h)\operatorname{tr}(\mathbf{t}_h)\,\mathbb{I} + \boldsymbol{\mu}(\mathbf{t}_h)\,\mathbf{t}_h\right\}\right)\right\|_{0,T}^2 \,\leq\, C_6 \left\{\|\mathbf{t} - \mathbf{t}_h\|_{0,T}^2 \,+\, \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,T}^2\right\}$$

for all $T \in \mathcal{T}_h$, and

$$h_e \left\|\left[\left(\boldsymbol{\sigma}_h - \left\{\boldsymbol{\lambda}(\mathbf{t}_h)\operatorname{tr}(\mathbf{t}_h)\,\mathbb{I} + \boldsymbol{\mu}(\mathbf{t}_h)\,\mathbf{t}_h\right\}\right) s\right]\right\|_{0,e}^2 \,\leq\, C_7 \left\{\|\mathbf{t} - \mathbf{t}_h\|_{0,\omega_e}^2 \,+\, \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,\omega_e}^2\right\}$$

for all $e \in \mathcal{E}_h(\Omega)$.

Proof. As in the proof of Lemma 5.9, it suffices now to apply the general results stated in [41, Lemmas 4.9 and 4.10] to $\boldsymbol{\rho}_h = \boldsymbol{\sigma}_h - \left\{\boldsymbol{\lambda}(\mathbf{t}_h)\operatorname{tr}(\mathbf{t}_h)\,\mathbb{I} + \boldsymbol{\mu}(\mathbf{t}_h)\,\mathbf{t}_h\right\}$ and $\boldsymbol{\rho} = \boldsymbol{\sigma} - \left\{\boldsymbol{\lambda}(\mathbf{t})\operatorname{tr}(\mathbf{t})\,\mathbb{I} + \boldsymbol{\mu}(\mathbf{t})\,\mathbf{t}\right\} = 0$ in Ω, and then use the Lipschitz-continuity of $\mathbb{A}_1$ (cf. (2.22) in Lemma 2.2) restricted to T and ω_e.

□

Now, for the efficiency of $\widehat{\boldsymbol{\theta}}$ it remains to provide upper bounds for the three terms completing the definition of the local error indicator $\widehat{\theta}_T^2$ (cf. (5.3)), which is done in what follows. In fact, using that $\mathbf{t} = \mathbf{e}(\mathbf{u})$ and that $\boldsymbol{\gamma} = \frac{1}{2}(\nabla \mathbf{u} - (\nabla \mathbf{u})^{\mathtt{t}})$ in Ω, we easily deduce that

$$\|\mathbf{e}(\mathbf{u}_h) - \mathbf{t}_h\|_{0,T} \leq C\left\{\|\mathbf{u} - \mathbf{u}_h\|_{1,T} + \|\mathbf{t} - \mathbf{t}_h\|_{0,T}\right\} \tag{5.39}$$

and

$$\left\|\boldsymbol{\gamma}_h - \frac{1}{2}(\nabla \mathbf{u}_h - (\nabla \mathbf{u}_h)^{\mathtt{t}})\right\|_{0,T} \leq C\left\{\|\mathbf{u} - \mathbf{u}_h\|_{1,T} + \|\boldsymbol{\gamma} - \boldsymbol{\gamma}_h\|_{0,T}\right\}. \tag{5.40}$$

In addition, employing that $\mathbf{u} = \mathbf{g}$ on Γ and applying the trace theorem, we find that

$$\sum_{e \in \mathcal{E}_h(\Gamma)} \|\mathbf{g} - \mathbf{u}_h\|_{0,e}^2 \,=\, \|\mathbf{u} - \mathbf{u}_h\|_{0,\Gamma}^2 \,\leq\, c\,\|\mathbf{u} - \mathbf{u}_h\|_{1,\Omega}^2 \,.$$

Finally, in order to complete the efficiency estimate for $\overline{\boldsymbol{\theta}}$, we just need to bound the second term defining the local error indicator $\overline{\theta}_T^2$ (cf. (5.33)), which, using again that $\mathbf{t} + \boldsymbol{\gamma} = \nabla \mathbf{u}$, yields

$$\|\mathbf{t}_h + \boldsymbol{\gamma}_h - \nabla \mathbf{u}_h\|_{0,T} \,\leq\, \|\mathbf{t} - \mathbf{t}_h\|_{0,T} + \|\boldsymbol{\gamma} - \boldsymbol{\gamma}_h\|_{0,T} + \|\mathbf{u} - \mathbf{u}_h\|_{1,T}\,. \tag{5.41}$$

Therefore, the required lower bound for $\overline{\boldsymbol{\theta}}$ is a straightforward consequence of (5.34), (5.35), (5.36), (5.39), (5.40), and (5.41).

6 Numerical results

In this section we present numerical examples illustrating the performance of the Galerkin schemes (2.27), (3.14), and (4.6), confirming the reliability and efficiency of the a posteriori error estimators derived in Section 5, and showing the behaviour of the associated adaptive algorithms. The specific finite element subspaces $X_{1,h}$, $M_{1,h}$, $M_h^{\mathbf{u}}$, and M_h^{γ} that are employed for the respective computational implementations are indicated below in Table 6.1. In each case we consider $k = 0$. In addition, all the nonlinear algebraic systems arising from the Galerkin schemes are solved by the Newton method with a tolerance of 1E-06 and taking as initial iteration the solution of the associated linear problems with $\widetilde{\mu}$, and hence $\widetilde{\lambda}$, constant.

Table 6.1: Finite element subspaces employed

Galerkin scheme	$X_{1,h}$	$M_{1,h}$	$M_h^{\mathbf{u}}$	M_h^{γ}
(2.27)	(2.34)	(2.31)	(2.32)	(2.33)
(3.14)	(3.18)	(2.31)	(2.32)	(2.33)
(4.6)	(4.9)	(4.10)	(4.11)	(4.12)

In what follows, N stands for the total number of degrees of freedom (unknowns) of each Galerkin scheme, which can be proved to behave asymptotically as the number of elements of each triangulation, multiplied by the factors 13.5, 11.5, and 9, for (2.27), (3.14), and (4.6), respectively. Also, the individual and total errors are given by

$$\mathsf{e}(\mathbf{t}) := \|\mathbf{t} - \mathbf{t}_h\|_{0,\Omega}\,, \quad \mathsf{e}(\boldsymbol{\sigma}) := \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\mathbf{div};\Omega}\,, \quad \mathsf{e}_0(\mathbf{u}) := \|\mathbf{u} - \mathbf{u}_h\|_{0,\Omega}\,,$$

$$\mathsf{e}_1(\mathbf{u}) := \|\mathbf{u} - \mathbf{u}_h\|_{1,\Omega}\,, \qquad \mathsf{e}(\gamma) := \|\gamma - \gamma_h\|_{0,\Omega}\,,$$

$$\mathsf{e}(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\gamma) := \Big\{ [\mathsf{e}(\mathbf{t})]^2 + [\mathsf{e}(\boldsymbol{\sigma})]^2 + [\mathsf{e}_0(\mathbf{u})]^2 + [\mathsf{e}(\gamma)]^2 \Big\}^{1/2},$$

$$\widetilde{\mathsf{e}}(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\gamma) := \Big\{ [\mathsf{e}(\mathbf{t})]^2 + [\mathsf{e}(\boldsymbol{\sigma})]^2 + [\mathsf{e}_0(\mathbf{u})]^2 + [\mathsf{e}(\gamma)]^2 \Big\}^{1/2},$$

$$\widehat{\mathsf{e}}(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\gamma) := \Big\{ [\mathsf{e}(\mathbf{t})]^2 + [\mathsf{e}(\boldsymbol{\sigma})]^2 + [\mathsf{e}_1(\mathbf{u})]^2 + [\mathsf{e}(\gamma)]^2 \Big\}^{1/2},$$

and

$$\overline{\mathsf{e}}(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\gamma) := \Big\{ [\mathsf{e}(\mathbf{t})]^2 + [\mathsf{e}(\boldsymbol{\sigma})]^2 + [\mathsf{e}_1(\mathbf{u})]^2 + [\mathsf{e}(\gamma)]^2 \Big\}^{1/2},$$

whereas the effectivity indexes are defined by

$$\mathtt{ef}(\boldsymbol{\theta}) := \mathsf{e}(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\gamma)/\boldsymbol{\theta}\,, \qquad \mathtt{ef}(\widetilde{\boldsymbol{\theta}}) := \widetilde{\mathsf{e}}(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\gamma)/\widetilde{\boldsymbol{\theta}}\,,$$

$$\mathtt{ef}(\widehat{\boldsymbol{\theta}}) := \widehat{\mathsf{e}}(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\gamma)/\widehat{\boldsymbol{\theta}}\,, \quad \text{and} \quad \mathtt{ef}(\overline{\boldsymbol{\theta}}) := \overline{\mathsf{e}}(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\gamma)/\overline{\boldsymbol{\theta}}\,.$$

In addition, we introduce the experimental rates of convergence

$$\mathtt{r}(\mathbf{t}) := \frac{\log(\mathsf{e}(\mathbf{t})/\mathsf{e}'(\mathbf{t}))}{\log(h/h')}\,, \quad \mathtt{r}(\boldsymbol{\sigma}) := \frac{\log(\mathsf{e}(\boldsymbol{\sigma})/\mathsf{e}'(\boldsymbol{\sigma}))}{\log(h/h')}\,, \quad \mathtt{r}_0(\mathbf{u}) := \frac{\log(\mathsf{e}_0(\mathbf{u})/\mathsf{e}_0'(\mathbf{u}))}{\log(h/h')}\,,$$

$$\mathtt{r}_1(\mathbf{u}) \,:=\, \frac{\log(\mathsf{e}_1(\mathbf{u})/\mathsf{e}_1'(\mathbf{u}))}{\log(h/h')}\,, \qquad \mathtt{r}(\boldsymbol{\gamma}) \,:=\, \frac{\log(\mathsf{e}(\boldsymbol{\gamma})/\mathsf{e}'(\boldsymbol{\gamma}))}{\log(h/h')}\,,$$

$$\mathtt{r}(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma}) \,:=\, \frac{\log(\mathsf{e}(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma})/\mathsf{e}'(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma}))}{\log(h/h')}\,,$$

and analogously for $\widetilde{\mathtt{r}}(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma})$, $\widehat{\mathtt{r}}(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma})$, and $\overline{\mathtt{r}}(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma})$, where e and e' denote the corresponding errors at two consecutive triangulations with mesh sizes h and h', respectively. However, when the adaptive algorithm is applied (see details below), the expression $\log(h/h')$ appearing in the computation of the above rates is replaced by $-\frac{1}{2}\log(N/N')$, where N and N' denote the corresponding degrees of freedom of each triangulation.

The examples to be considered in this section are described next. Example 1 is employed to illustrate the performance of the discrete schemes and to confirm the reliability and efficiency of the a posteriori error estimators when a sequence of quasi-uniform meshes is considered. Then, Examples 2 and 3 are utilized to show the behavior of the associated adaptive algorithms, which apply the following procedure from [58] for each $\boldsymbol{\chi} \in \left\{\boldsymbol{\theta}, \widetilde{\boldsymbol{\theta}}, \widehat{\boldsymbol{\theta}}, \overline{\boldsymbol{\theta}}\right\}$ with local indicators χ_T, $T \in \mathcal{T}_h$:

1) Start with a coarse mesh $\mathcal{T}_h$.

2) Solve the discrete problem for the actual mesh $\mathcal{T}_h$.

3) Compute χ_T for each triangle $T \in \mathcal{T}_h$.

4) Evaluate stopping criterion and decide to finish or go to next step.

5) Use *blue-green* procedure to refine each $T' \in \mathcal{T}_h$ whose indicator $\chi_{T'}$ satisfies

$$\chi_{T'} \,\geq\, \frac{1}{2}\max\left\{\chi_T : \quad T \in \mathcal{T}_h\right\}$$

6) Define resulting mesh as actual mesh $\mathcal{T}_h$ and go to step 2.

In all the examples we consider the Lamé functions $\widetilde{\lambda}$, $\widetilde{\mu} : \mathrm{R}^+ \to \mathrm{R}$ defined by

$$\widetilde{\lambda}(\rho) \,:=\, \kappa - \frac{1}{2}\,\widetilde{\mu}(\rho) \quad \text{and} \quad \widetilde{\mu}(\rho) \,:=\, \beta_0 \,+\, \beta_1\,(1+\rho^2)^{(\beta-2)/2} \qquad \forall\,\rho \,\in\, \mathrm{R}^+\,,$$

with $\kappa = \beta_0 = \beta_1 = 1/4$, and $\beta = 3/2$, which are easily shown to verify the assumptions (2.2) with $\mu_0 = \mu_1 = 1/4$ and $\mu_2 = 5/8$. This function $\widetilde{\mu}$ corresponds to the Carreau law for viscoplastic materials (see, e.g. [51], [56]). Now, according to (2.21), we obtain $\alpha_1 = 1/4$ and $\gamma_1 = 3/2$, which, as indicated in Section 4.1, yields the following stabilization parameters for the partial and fully augmented formulations:

$$\begin{gathered}\kappa_0 = \frac{\alpha_1}{\gamma_1^2} \,=\, 1/9\,, \quad \kappa_1 = \frac{\alpha_1}{2}\,\min\left\{1, \frac{1}{\gamma_1^2}\right\} \,=\, 1/18\,, \quad \kappa_2 \,=\, \frac{\alpha_1}{2}\,\min\left\{1, \frac{1}{\gamma_1^2}\right\} \,=\, 1/18\,,\\ \kappa_3 \,=\, \frac{\alpha_1}{4}\,\min\left\{1, \frac{1}{\gamma_1^2}\right\} \,=\, 1/36\,, \quad \text{and} \quad \kappa_4 \,=\, \frac{\alpha_1}{4}\,\min\left\{1, \frac{1}{\gamma_1^2}\right\} \,=\, 1/36\,.\end{gathered} \tag{6.1}$$

In Example 1 we set $\Omega =]0,1[^2$ and choose the data $\mathbf{f}$ and $\mathbf{g}$ so that the exact solution is given by

$$\mathbf{u}(\mathbf{x}) \,:=\, \begin{pmatrix} \sin x_1 \cos x_2 \exp(x_1 x_2) \\ \cos x_1 \sin x_2 \exp(-x_1 x_2) \end{pmatrix} \qquad \forall\,\mathbf{x} := (x_1, x_2)^{\mathtt{t}} \,\in\, \Omega\,.$$

In turn, in Example 2 we consider the T-shaped domain $\Omega =]-1,1[^2 \setminus \big([-1,-0.25]\times[-1,0.5] \cup [0.25,1]\times[-1,0.5]\big)$, and choose the data $\mathbf{f}$ and $\mathbf{g}$ so that the exact solution is given by

$$\mathbf{u}(\mathbf{x}) := \left(\big\|\mathbf{x} - (-0.25, 0.5)\big\|^{4/3} \sin\left(\frac{2\theta_1 + \pi}{3}\right), \big\|\mathbf{x} - (0.25, 0.5)\big\|^{5/3} \sin\left(\frac{2\theta_2}{3}\right)\right)^{\mathtt{t}}$$

for all $\mathbf{x} := (x_1, x_2)^{\mathtt{t}} \in \Omega$, with

$$\theta_1 := \mathrm{Arctan}\left(\frac{x_2 - 0.50}{x_1 + 0.25}\right) \quad \text{and} \quad \theta_2 := \mathrm{Arctan}\left(\frac{x_2 - 0.50}{x_1 - 0.25}\right).$$

Note that the partial derivatives of this solution are singular at the points $(-0.25, 0.5)$ and $(0.25, 0.5)$, which are the middle corners of the T.

Finally, in Example 3, we consider the L-shaped domain $\Omega :=]-1,1[^2 \setminus [0,1]^2$ and choose the data $\mathbf{f}$ and $\mathbf{g}$ so that the exact solution is given by

$$\mathbf{u}(\mathbf{x}) := \sin(\pi\, x_1)\, \sin(\pi\, x_2)\, r^{-2/3} \sin\left(\frac{2\,\theta - \pi}{3}\right) (1,1)^{\mathtt{t}} \qquad \forall\, \mathbf{x} := (x_1, x_2)^{\mathtt{t}} \in \Omega\,,$$

where (r, θ) stands for the usual polar coordinates, that is $r := \|\mathbf{x}\|$ and $\theta := \mathrm{Arctan}\left(\frac{x_2}{x_1}\right)$. Note that the trace of $\mathbf{u}$ vanishes on $\Gamma := \partial\Omega$ and that its partial derivatives are singular at the origin, which is the inner corner of the L.

In Tables 6.2, 6.3, and 6.4 we summarize the convergence history of the finite element schemes (2.27), (3.14), and (4.6) as applied to Example 1 for sequences of quasi-uniform triangulations of the domains. The number of Newton iterations required, for the tolerance given, ranges between 3 and 6 for all the nonlinear systems involved. We observe in these tables, looking at the corresponding experimental rates of convergence, that the $O(h)$ predicted by Theorems 2.7, 3.5, and 4.4 (with $\delta = 1$ in the three cases) is attained by all the unknowns. In addition, we also highlight, according to the last column of each one of the tables, that the effectivity indexes $\mathtt{ef}(\boldsymbol{\theta})$, $\mathtt{ef}(\widetilde{\boldsymbol{\theta}})$, and $\mathtt{ef}(\widehat{\boldsymbol{\theta}})$ remain all bounded (they lie in neighborhoods of 0.34, 0.19, and 0.17, respectively), which illustrates, in this case of a regular solution, the reliability and efficiency of the three a posteriori error estimators $\boldsymbol{\theta}$, $\widetilde{\boldsymbol{\theta}}$, and $\widehat{\boldsymbol{\theta}}$. On the other hand, in Figure 6.1, which for sake of completeness includes additional inputs that are not listed in the corresponding tables, we display the total errors $\mathtt{e}(\mathbf{t}, \boldsymbol{\sigma}, \mathbf{u}, \gamma)$, $\widetilde{\mathtt{e}}(\mathbf{t}, \boldsymbol{\sigma}, \mathbf{u}, \gamma)$, and $\widehat{\mathtt{e}}(\mathbf{t}, \boldsymbol{\sigma}, \mathbf{u}, \gamma)$ vs. the degrees of freedom N. It is interesting to notice there that, though the three schemes yield the same rate of convergence (which was already confirmed by the tables), the augmented one requires less degrees of freedom than the other two to achieve a given accuracy. This fact is particularly important when comparing the non-augmented and augmented approaches since both measure their respective errors with exactly the same norm, and hence, this example would suggest to better employ the latter one instead of the former. Actually, this observation could have been announced in advance since, on the contrary to the finite element subspace $X_{1,h}$ (cf. (2.34)) employed in (2.27), the corresponding finite element subspace $\widetilde{X}_{1,h}$ (cf. (3.18)) utilized in the augmented scheme (3.14) does not include the bubble functions, which certainly yields a less amount of degrees of freedom.

Next, in Tables 6.5 up to 6.12, we provide the convergence history of the quasi-uniform and adaptive schemes (2.27), (3.14), and (4.6) as applied to Examples 2 and 3. More precisely, Example 2 is utilized to illustrate the behavior of the three methods, while Example 3, which considers homogeneous Dirichlet boundary conditions, is employed only to show the performance of the fully-augmented approach (4.6) with the a posteriori error estimator $\overline{\boldsymbol{\theta}}$. The number of Newton iterations required now ranges between 3 and 8, and between 11 and 18, respectively. We notice, as expected, that the errors of the adaptive methods decrease faster than those obtained by the quasi-uniform ones. This fact is better illustrated in Figures 6.2, 6.3, 6.4, and 6.7, where we display the total errors vs. the degrees of freedom N for the corresponding refinements. In addition, in Figures 6.5 and 6.6 we summarize the results of Example 2 by displaying the total errors vs. N for the quasi-uniform and adaptive refinements of the three schemes. It is interesting to observe there that, at least for this example, both augmented approaches perform much better than the non-augmented one. Note that all these figures include additional data that are not shown in the corresponding tables. Furthermore, we see from the last column of the tables that the effectivity indexes remain again bounded from above and below, which confirms the reliability and efficiency of $\boldsymbol{\theta}$, $\widetilde{\boldsymbol{\theta}}$, $\widehat{\boldsymbol{\theta}}$, and $\overline{\boldsymbol{\theta}}$, in these cases of non-smooth solutions, as well. Some intermediate meshes obtained with the associated adaptive algorithms are displayed in Figures 6.8 and 6.10 for the augmented and fully-augmented schemes, respectively. It is important to observe here that the adapted meshes concentrate the refinements around the points $(-0.25, 0.5)$ and $(0.25, 0.5)$ in Example 2, and around the origin in Example 3, which confirms that the methods are able to recognize the singularity regions of the solutions. On the other hand, in Figures 6.9 and 6.11 we consider fixed meshes (according to the values of N indicated there) and display the total errors vs. κ_0 and κ_3, respectively, for the quasi-uniform augmented and fully-augmented approaches as applied to Example 2. The other parameters needed are taken from (6.1). The corresponding non-augmented schemes yield $N = 252868$ and $N = 312172$ with total errors $\mathtt{e}(\mathbf{t}, \boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\gamma})$ given by $3.188E-01$ and $3.003E-01$, respectively. It is quite clear from these figures that for each one of the parameters κ_0 and κ_3 there is a sufficiently large range yielding stable Galerkin schemes in the sense that the corresponding errors remain bounded. This fact, which was theoretically known in advance for κ_0 (cf. Theorems 3.3 and 4.3), certainly confirms the robustness of the augmented and fully-augmented methods with respect to these stabilization parameters. This remark is specially significant for κ_3, which, as explained in Section 4.1, can only be determined heuristically. Note in particular that the parameters $\kappa_0 = 1/9$ and $\kappa_3 = 1/36$ employed in our computations lie precisely in the ranges identified by Figures 6.9 and 6.11. Finally, in order to illustrate the accurateness of the finite element schemes and their associated adaptive algorithms, in Figures 6.12, 6.13, and 6.14, we display some components of the approximate (left) and exact (right) solutions for Examples 2 and 3.

We conclude this paper by emphasizing that we have provided enough support to consider the augmented and fully-aumented mixed finite element schemes (3.14) and (4.6), together with its associated adaptive algorithms, as valid and competitive alternatives to solve the present class of nonlinear elasticity problems.

N	h	$\mathtt{e}(\mathbf{t})$	$\mathtt{r}(\mathbf{t})$	$\mathtt{e}(\boldsymbol{\sigma})$	$\mathtt{r}(\boldsymbol{\sigma})$	$\mathtt{e}_0(\mathbf{u})$	$\mathtt{r}_0(\mathbf{u})$	$\mathtt{e}(\boldsymbol{\gamma})$	$\mathtt{r}(\boldsymbol{\gamma})$	$\mathtt{ef}(\boldsymbol{\theta})$
7009	1/16	3.808E−02	–	7.034E−02	–	2.003E−02	–	1.472E−02	–	0.3418
10921	1/20	3.047E−02	1.000	5.628E−02	1.000	1.602E−02	1.000	1.106E−02	1.294	0.3412
15697	1/24	2.539E−02	1.001	4.690E−02	1.000	1.335E−02	1.000	8.693E−03	1.327	0.3409
21337	1/28	2.176E−02	1.002	4.020E−02	1.000	1.145E−02	1.000	7.065E−03	1.351	0.3407
27841	1/32	1.903E−02	1.002	3.517E−02	1.000	1.001E−02	1.000	5.887E−03	1.370	0.3407
35209	1/36	1.691E−02	1.002	3.126E−02	1.000	8.902E−03	1.000	5.003E−03	1.385	0.3406
62497	1/48	1.268E−02	1.002	2.344E−02	1.000	6.676E−03	1.000	3.342E−03	1.407	0.3407
110977	1/64	9.502E−03	1.002	1.758E−02	1.000	5.007E−03	1.000	2.217E−03	1.432	0.3408
173281	1/80	7.598E−03	1.002	1.406E−02	1.000	4.006E−03	1.000	1.607E−03	1.443	0.3409
249409	1/96	6.330E−03	1.002	1.172E−02	1.000	3.338E−03	1.000	1.233E−03	1.453	0.3409
443137	1/128	4.746E−03	1.000	8.790E−03	0.999	2.504E−03	1.000	8.165E−04	1.403	0.3411
692161	1/160	3.796E−03	1.000	7.032E−03	1.000	2.003E−03	1.000	5.909E−04	1.448	0.3412
996481	1/192	3.164E−03	1.000	5.861E−03	1.000	1.669E−03	1.000	4.545E−04	1.440	0.3413

Table 6.2: EXAMPLE 1, quasi–uniform non–augmented scheme (2.27)

N	h	$\mathtt{e}(\mathbf{t})$	$\mathtt{r}(\mathbf{t})$	$\mathtt{e}(\boldsymbol{\sigma})$	$\mathtt{r}(\boldsymbol{\sigma})$	$\mathtt{e}_0(\mathbf{u})$	$\mathtt{r}_0(\mathbf{u})$	$\mathtt{e}(\boldsymbol{\gamma})$	$\mathtt{r}(\boldsymbol{\gamma})$	$\mathtt{ef}(\widetilde{\boldsymbol{\theta}})$
5985	1/16	4.047E−02	–	3.547E−02	–	2.003E−02	–	5.355E−03	–	0.1990
7561	1/18	3.597E−02	1.000	3.169E−02	0.957	1.780E−02	1.000	4.499E−03	1.478	0.1985
11265	1/22	2.942E−02	1.001	2.612E−02	0.966	1.457E−02	1.000	3.343E−03	1.481	0.1978
15705	1/26	2.489E−02	1.001	2.221E−02	0.971	1.233E−02	1.000	2.612E−03	1.481	0.1973
20881	1/30	2.157E−02	1.001	1.932E−02	0.975	1.068E−02	1.000	2.113E−03	1.481	0.1970
26793	1/34	1.903E−02	1.001	1.709E−02	0.979	9.425E−03	1.000	1.756E−03	1.482	0.1967
37041	1/40	1.617E−02	1.001	1.458E−02	0.981	8.012E−03	1.000	1.380E−03	1.482	0.1964
72465	1/56	1.154E−02	1.001	1.046E−02	0.986	5.723E−03	1.000	8.381E−04	1.482	0.1960
147681	1/80	8.078E−03	1.001	7.352E−03	0.990	4.006E−03	1.000	4.942E−04	1.481	0.1957
289185	1/112	5.773E−03	0.995	5.268E−03	0.989	2.861E−03	1.000	3.098E−04	1.280	0.1956
477793	1/144	4.492E−03	0.998	4.104E−03	0.994	2.226E−03	1.000	2.271E−04	1.188	0.1955
589761	1/160	4.044E−03	0.998	3.696E−03	0.995	2.003E−03	1.000	2.025E−04	1.089	0.1955
849025	1/192	3.370E−03	0.999	3.078E−03	1.003	1.669E−03	1.000	1.804E−04	0.635	0.1955
1155393	1/224	2.891E−03	0.996	2.640E−03	0.996	1.431E−03	1.000	1.594E−04	0.803	0.1955

Table 6.3: EXAMPLE 1, quasi–uniform augmented scheme (3.14)

N	h	$\mathtt{e}(\mathbf{t})$	$\mathtt{r}(\mathbf{t})$	$\mathtt{e}(\boldsymbol{\sigma})$	$\mathtt{r}(\boldsymbol{\sigma})$	$\mathtt{e}_1(\mathbf{u})$	$\mathtt{r}_1(\mathbf{u})$	$\mathtt{e}(\boldsymbol{\gamma})$	$\mathtt{r}(\boldsymbol{\gamma})$	$\mathtt{ef}(\widehat{\boldsymbol{\theta}})$
4738	1/16	4.223E−02	–	2.633E−02	–	5.051E−02	–	1.154E−01	–	0.1755
8890	1/22	3.030E−02	1.040	1.895E−02	1.030	3.613E−02	1.049	8.808E−02	0.872	0.1726
12378	1/26	2.549E−02	1.034	1.596E−02	1.025	3.035E−02	1.042	7.588E−02	0.900	0.1715
16442	1/30	2.199E−02	1.030	1.379E−02	1.021	2.616E−02	1.037	6.657E−02	0.919	0.1708
21082	1/34	1.934E−02	1.026	1.214E−02	1.018	2.298E−02	1.033	5.926E−02	0.933	0.1702
29122	1/40	1.638E−02	1.022	1.030E−02	1.015	1.945E−02	1.028	5.084E−02	0.945	0.1696
56898	1/56	1.163E−02	1.015	7.326E−03	1.010	1.379E−02	1.019	3.680E−02	0.966	0.1687
74242	1/64	1.016E−02	1.012	6.404E−03	1.008	1.204E−02	1.016	3.232E−02	0.973	0.1685
166658	1/96	6.750E−03	1.007	4.260E−03	1.004	7.989E−03	1.010	2.171E−02	0.985	0.1679
295938	1/128	5.061E−03	1.003	3.193E−03	1.002	5.986E−03	1.004	1.633E−02	0.991	0.1677
462082	1/160	4.048E−03	1.000	2.554E−03	1.001	4.787E−03	1.001	1.308E−02	0.994	0.1675
665090	1/192	3.372E−03	1.001	2.128E−03	1.001	3.989E−03	1.000	1.091E−02	0.995	0.1675
904962	1/224	2.892E−03	0.997	1.824E−03	1.000	3.421E−03	0.997	9.359E−03	0.996	0.1675
1181698	1/256	2.532E−03	0.995	1.596E−03	0.999	2.995E−03	0.995	8.192E−03	0.997	0.1674

Table 6.4: EXAMPLE 1, quasi–uniform fully-augmented scheme (4.6)

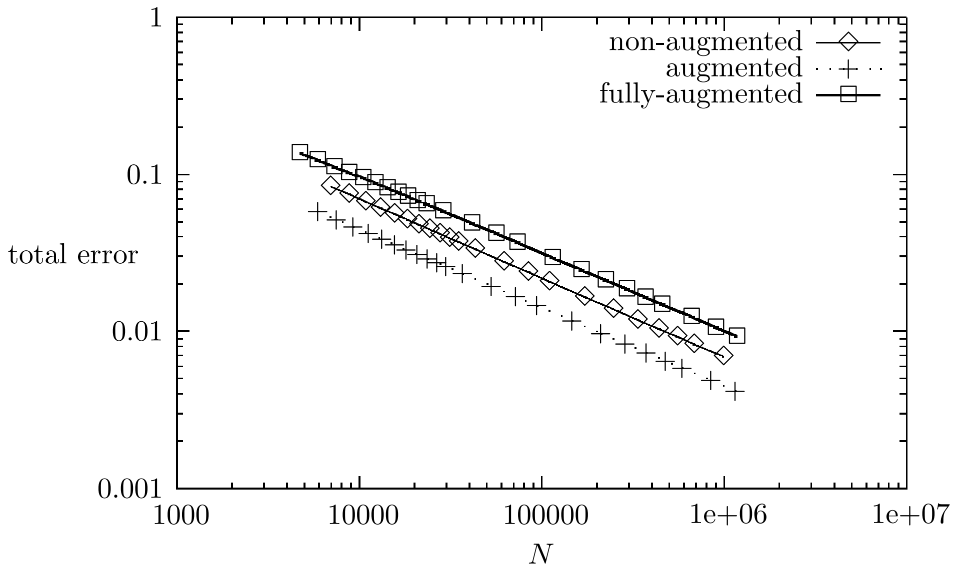

Figure 6.1: EXAMPLE 1, total error vs. N for the quasi-uniform schemes

N	h	$\mathtt{e}(\mathbf{t})$	$\mathtt{e}(\boldsymbol{\sigma})$	$\mathtt{e}_0(\mathbf{u})$	$\mathtt{e}(\gamma)$	$\mathtt{e}(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma})$	$\mathtt{r}(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma})$	$\mathtt{ef}(\boldsymbol{\theta})$
184	1/1	3.315E−01	1.168E−00	2.361E−01	4.371E−01	1.311E−00	–	0.4497
745	1/3	1.584E−01	8.296E−01	9.472E−02	1.251E−01	8.590E−01	0.613	0.6926
3817	1/7	7.368E−02	6.473E−01	4.206E−02	4.717E−02	6.545E−01	0.274	0.8589
6256	1/9	5.511E−02	5.812E−01	3.230E−02	2.601E−02	5.853E−01	0.175	0.9000
9211	1/11	4.484E−02	5.181E−01	2.685E−02	2.164E−02	5.211E−01	0.405	0.9125
13297	1/13	3.803E−02	5.056E−01	2.245E−02	1.792E−02	5.078E−01	0.433	0.9324
15115	1/14	3.559E−02	5.058E−01	-2.101E−02	1.466E−02	5.077E−01	0.002	0.9416
25828	1/18	2.763E−02	4.486E−01	1.616E−02	1.243E−02	4.499E−01	0.866	0.9540
38188	1/22	2.298E−02	4.151E−01	1.328E−02	9.829E−03	4.160E−01	0.621	0.9626
67732	1/29	1.699E−02	3.717E−01	9.862E−03	5.512E−03	3.723E−01	0.485	0.9746
142486	1/42	1.197E−02	3.507E−01	6.824E−03	3.858E−03	3.510E−01	0.108	0.9858
252868	1/56	8.936E−03	3.186E−01	5.112E−03	2.911E−03	3.188E−01	0.117	0.9903
321172	1/63	7.982E−03	3.002E−01	4.535E−03	2.301E−03	3.003E−01	0.507	0.9914
519349	1/80	6.257E−03	2.758E−01	3.562E−03	1.778E−03	2.759E−01	0.428	0.9937
660085	1/90	5.596E−03	2.584E−01	3.156E−03	1.632E−03	2.585E−01	0.552	0.9943
813916	1/100	4.978E−03	2.391E−01	2.846E−03	1.364E−03	2.391E−01	0.739	0.9947

Table 6.5: EXAMPLE 2, quasi–uniform non-augmented scheme (2.27)

N	h	$\mathtt{e}(\mathbf{t})$	$\mathtt{e}(\boldsymbol{\sigma})$	$\mathtt{e}_0(\mathbf{u})$	$\mathtt{e}(\gamma)$	$\mathtt{e}(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma})$	$\mathtt{r}(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma})$	$\mathtt{ef}(\boldsymbol{\theta})$
184	1.000	3.315E−01	1.168E−00	2.361E−01	4.371E−01	1.311E−00	–	0.4497
1288	0.451	1.190E−01	7.572E−01	9.565E−02	1.088E−01	7.801E−01	0.481	0.6742
3340	0.375	8.009E−02	5.796E−01	6.607E−02	7.456E−02	5.935E−01	0.683	0.7048
5881	0.354	6.565E−02	4.499E−01	5.269E−02	5.932E−02	4.615E−01	0.852	0.6963
10315	0.250	5.654E−02	3.534E−01	4.681E−02	4.376E−02	3.636E−01	1.007	0.6732
16111	0.188	4.523E−02	2.925E−01	3.259E−02	3.274E−02	2.996E−01	0.975	0.7048
18856	0.188	3.898E−02	2.648E−01	3.068E−02	2.898E−02	2.709E−01	1.277	0.7071
25870	0.125	3.277E−02	2.206E−01	2.627E−02	2.219E−02	2.257E−01	1.085	0.7075
56914	0.094	2.219E−02	1.520E−01	1.683E−02	1.515E−02	1.553E−01	0.981	0.7226
83722	0.094	1.772E−02	1.246E−01	1.450E−02	1.204E−02	1.272E−01	1.091	0.7203
124744	0.063	1.457E−02	1.021E−01	1.204E−02	8.918E−03	1.043E−01	0.968	0.7247
202879	0.047	1.140E−02	8.102E−02	8.635E−03	7.565E−03	8.262E−02	1.045	0.7362
484858	0.031	7.374E−03	5.211E−02	6.112E−03	4.357E−03	5.316E−02	0.999	0.7258
811063	0.023	5.713E−03	4.126E−02	4.351E−03	3.431E−03	4.202E−02	0.987	0.7432
1002865	0.023	5.031E−03	3.670E−02	4.060E−03	2.975E−03	3.738E−02	1.102	0.7382

Table 6.6: EXAMPLE 2, adaptive non-augmented scheme (2.27)

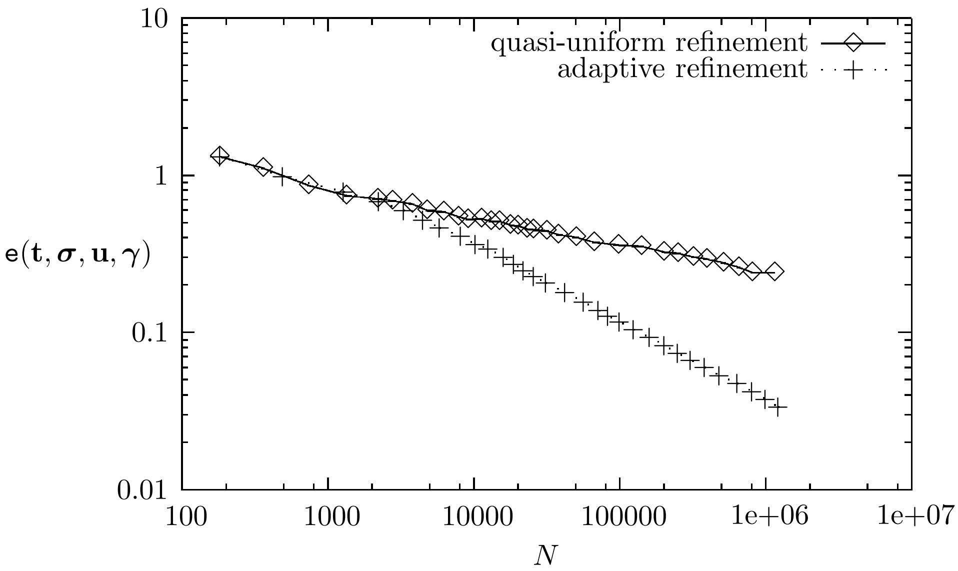

Figure 6.2: EXAMPLE 2, $\mathtt{e}(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma})$ vs. N for non-augmented scheme (2.27)

N	h	$\mathtt{e}(\mathbf{t})$	$\mathtt{e}(\boldsymbol{\sigma})$	$\mathtt{e}_0(\mathbf{u})$	$\mathtt{e}(\boldsymbol{\gamma})$	$\widetilde{\mathtt{e}}(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma})$	$\widetilde{\mathtt{r}}(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma})$	$\mathtt{ef}(\widetilde{\boldsymbol{\theta}})$
160	1/1	2.476E−01	4.578E−01	2.291E−01	9.596E−02	5.767E−01	–	0.2685
641	1/3	1.307E−01	3.193E−01	9.468E−02	4.036E−02	3.600E−01	0.721	0.3645
1894	1/5	7.750E−02	2.555E−01	5.534E−02	2.062E−02	2.734E−01	0.407	0.4225
5346	1/9	4.686E−02	2.105E−01	3.231E−02	1.098E−02	2.183E−01	0.325	0.4918
7867	1/11	3.801E−02	1.867E−01	2.686E−02	7.872E−03	1.925E−01	0.486	0.5052
11351	1/13	3.214E−02	1.804E−01	2.245E−02	6.398E−03	1.847E−01	0.519	0.5436
19863	1/17	2.487E−02	1.623E−01	1.688E−02	4.792E−03	1.651E−01	0.745	0.5949
27187	1/20	2.134E−02	1.571E−01	1.453E−02	3.909E−03	1.592E−01	0.201	0.6436
43166	1/25	1.687E−02	1.414E−01	1.151E−02	2.745E−03	1.429E−01	0.349	0.6849
84032	1/35	1.223E−02	1.260E−01	8.181E−03	1.830E−03	1.268E−01	0.231	0.7512
173459	1/50	8.548E−03	1.136E−01	5.698E−03	1.181E−03	1.141E−01	0.478	0.8230
215506	1/56	7.637E−03	1.116E−01	5.112E−03	1.046E−03	1.120E−01	0.161	0.8497
442551	1/80	5.353E−03	9.618E−02	3.562E−03	6.419E−04	9.640E−02	0.455	0.8907
562455	1/90	4.785E−03	9.118E−02	3.156E−03	5.895E−04	9.137E−02	0.455	0.9007
693514	1/100	4.262E−03	8.424E−02	2.846E−03	4.910E−04	8.440E−02	0.752	0.9054

Table 6.7: EXAMPLE 2, quasi–uniform augmented scheme (3.14)

N	h	$\mathtt{e}(\mathbf{t})$	$\mathtt{e}(\boldsymbol{\sigma})$	$\mathtt{e}_0(\mathbf{u})$	$\mathtt{e}(\boldsymbol{\gamma})$	$\widetilde{\mathtt{e}}(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma})$	$\widetilde{\mathtt{r}}(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma})$	$\mathtt{ef}(\widetilde{\boldsymbol{\theta}})$
160	1.000	2.476E−01	4.578E−01	2.291E−01	9.596E−02	5.767E−01	–	0.2685
1569	0.500	8.140E−02	2.640E−01	6.894E−02	2.496E−02	2.858E−01	0.544	0.3910
6105	0.250	4.152E−02	1.845E−01	3.585E−02	9.050E−03	1.927E−01	0.595	0.3973
14792	0.188	2.875E−02	1.380E−01	2.583E−02	5.233E−03	1.434E−01	0.683	0.4147
23877	0.125	2.273E−02	1.063E−01	2.010E−02	3.875E−03	1.106E−01	1.239	0.3984
31951	0.125	1.971E−02	9.002E−02	1.802E−02	3.139E−03	9.395E−02	1.683	0.3925
36300	0.125	1.846E−02	8.232E−02	1.698E−02	2.770E−03	8.610E−02	1.368	0.3850
44567	0.094	1.657E−02	7.462E−02	1.520E−02	2.371E−03	7.797E−02	0.966	0.3851
55891	0.088	1.482E−02	6.750E−02	1.359E−02	2.081E−03	7.047E−02	0.894	0.3866
69790	0.063	1.329E−02	6.113E−02	1.227E−02	1.870E−03	6.378E−02	0.898	0.3890
88006	0.063	1.181E−02	5.432E−02	1.103E−02	1.555E−03	5.669E−02	1.015	0.3870
114690	0.063	1.023E−02	4.805E−02	9.387E−03	1.298E−03	5.003E−02	0.944	0.3895
142100	0.063	9.170E−03	4.349E−02	8.309E−03	1.113E−03	4.523E−02	0.941	0.3914
174543	0.047	8.284E−03	3.888E−02	7.528E−03	9.659E−04	4.047E−02	1.083	0.3860
223591	0.044	7.395E−03	3.487E−02	6.867E−03	1.121E−03	3.632E−02	0.873	0.3884

Table 6.8: EXAMPLE 2, adaptive augmented scheme (3.14)

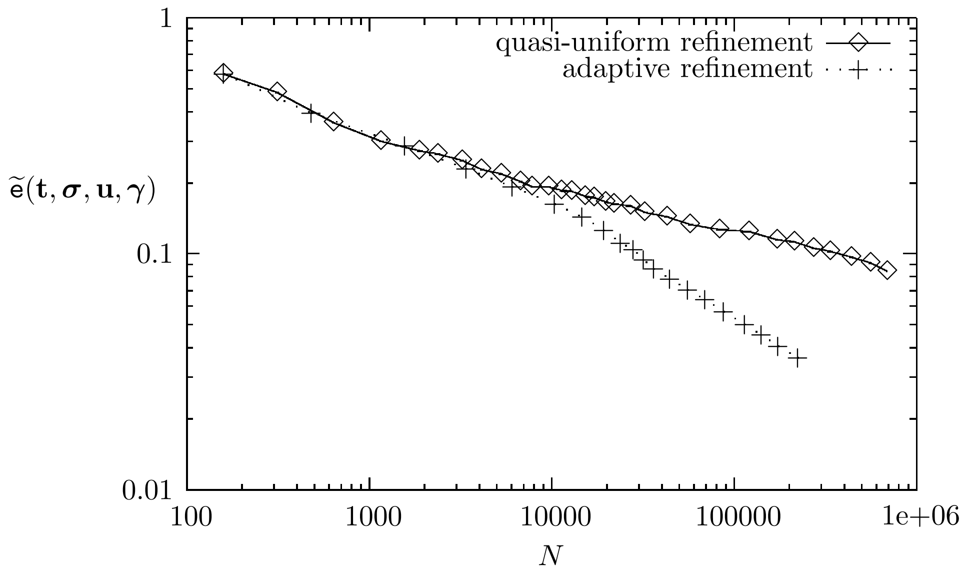

Figure 6.3: EXAMPLE 2, $\widetilde{\mathtt{e}}(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma})$ vs. N for augmented scheme (3.14)

N	h	$\mathtt{e}(\mathbf{t})$	$\mathtt{e}(\boldsymbol{\sigma})$	$\mathtt{e}_1(\mathbf{u})$	$\mathtt{e}(\boldsymbol{\gamma})$	$\widehat{\mathtt{e}}(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma})$	$\widehat{\mathtt{r}}(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma})$	$\mathtt{ef}(\widehat{\boldsymbol{\theta}})$
138	1/1	2.389E−01	4.817E−01	7.515E−01	9.610E−02	9.290E−01	–	0.7159
526	1/3	1.325E−01	3.210E−01	2.506E−01	1.289E−01	4.472E−01	0.968	0.5157
1519	1/5	7.912E−02	2.535E−01	1.539E−01	1.001E−01	3.228E−01	0.695	0.4972
6234	1/11	3.916E−02	1.832E−01	6.733E−02	7.062E−02	2.112E−01	0.583	0.4606
12094	1/15	2.899E−02	1.688E−01	5.364E−02	5.144E−02	1.867E−01	0.735	0.5308
15661	1/17	2.548E−02	1.607E−01	4.966E−02	4.272E−02	1.754E−01	0.925	0.5794
21411	1/20	2.192E−02	1.557E−01	4.383E−02	3.874E−02	1.678E−01	0.265	0.5984
33951	1/25	1.732E−02	1.404E−01	3.133E−02	3.137E−02	1.483E−01	0.393	0.6307
66000	1/35	1.261E−02	1.254E−01	2.083E−02	2.367E−02	1.299E−01	0.260	0.6845
136083	1/50	8.807E−03	1.133E−01	1.435E−02	1.646E−02	1.157E−01	0.494	0.7689
214624	1/63	7.034E−03	1.048E−01	1.049E−02	1.355E−02	1.064E−01	0.541	0.8016
346875	1/80	5.530E−03	9.606E−02	8.087E−03	1.073E−02	9.715E−02	0.469	0.8391
440779	1/90	4.943E−03	9.112E−02	7.438E−03	9.560E−03	9.206E−02	0.457	0.8538
543413	1/100	4.401E−03	8.420E−02	6.289E−03	8.664E−03	8.499E−02	0.758	0.8582
779875	1/120	3.721E−03	8.440E−02	5.497E−03	7.126E−03	8.496E−02	0.002	0.8971

Table 6.9: Example 2, quasi–uniform fully-augmented scheme (4.6)

N	h	$\mathtt{e}(\mathbf{t})$	$\mathtt{e}(\boldsymbol{\sigma})$	$\mathtt{e}_1(\mathbf{u})$	$\mathtt{e}(\boldsymbol{\gamma})$	$\widehat{\mathtt{e}}(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma})$	$\widehat{\mathtt{r}}(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma})$	$\mathtt{ef}(\widehat{\boldsymbol{\theta}})$
138	1.000	2.389E−01	4.817E−01	7.515E−01	9.610E−02	9.290E−01	–	0.7159
976	0.707	9.092E−02	2.702E−01	1.721E−01	1.088E−01	3.503E−01	0.739	0.4652
3796	0.500	5.491E−02	1.825E−01	7.484E−02	6.519E−02	2.149E−01	0.812	0.4612
7538	0.250	4.004E−02	1.414E−01	4.926E−02	4.594E−02	1.617E−01	0.797	0.4850
12319	0.250	3.503E−02	1.130E−01	4.231E−02	3.729E−02	1.311E−01	0.740	0.4827
19871	0.177	2.551E−02	9.059E−02	3.037E−02	2.994E−02	1.033E−01	0.985	0.4802
30635	0.125	2.039E−02	7.395E−02	2.404E−02	2.357E−02	8.377E−02	0.952	0.4933
46093	0.125	1.714E−02	6.098E−02	2.042E−02	1.911E−02	6.924E−02	0.957	0.4978
73803	0.125	1.323E−02	4.800E−02	1.551E−02	1.572E−02	5.447E−02	0.994	0.4814
120383	0.094	1.028E−02	3.849E−02	1.191E−02	1.211E−02	4.331E−02	1.065	0.4925
185623	0.063	8.290E−03	3.112E−02	9.650E−03	9.901E−03	3.504E−02	0.983	0.4889
301403	0.063	6.490E−03	2.442E−02	7.533E−03	7.847E−03	2.751E−02	0.881	0.4850
486087	0.044	5.134E−03	1.960E−02	5.950E−03	6.096E−03	2.198E−02	0.987	0.4950
612070	0.044	4.663E−03	1.794E−02	5.406E−03	5.497E−03	2.008E−02	0.785	0.4999

Table 6.10: Example 2, adaptive fully-augmented scheme (4.6)

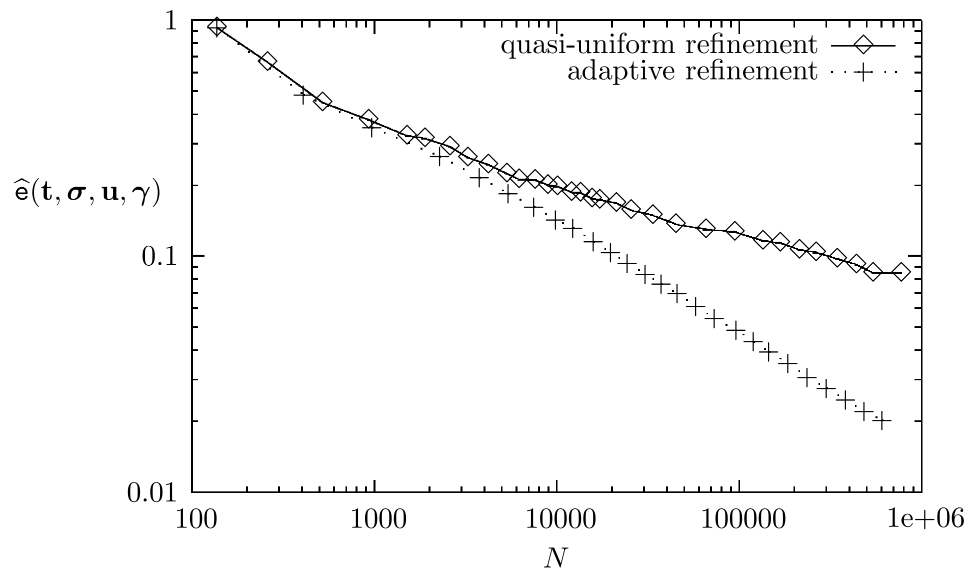

Figure 6.4: Example 2, $\widehat{\mathtt{e}}(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma})$ vs. N for fully-augmented scheme (4.6)

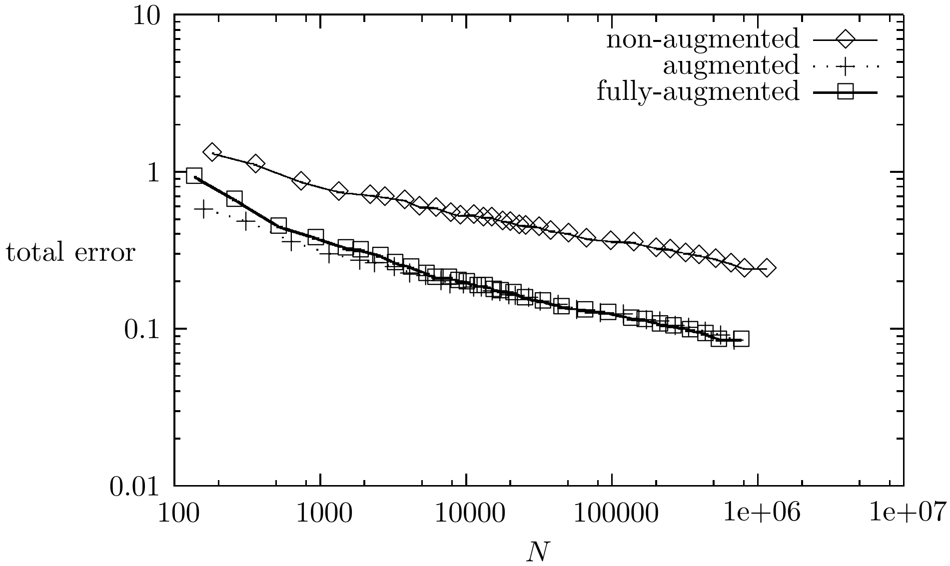

Figure 6.5: EXAMPLE 2, total error vs. N for the quasi-uniform refinements

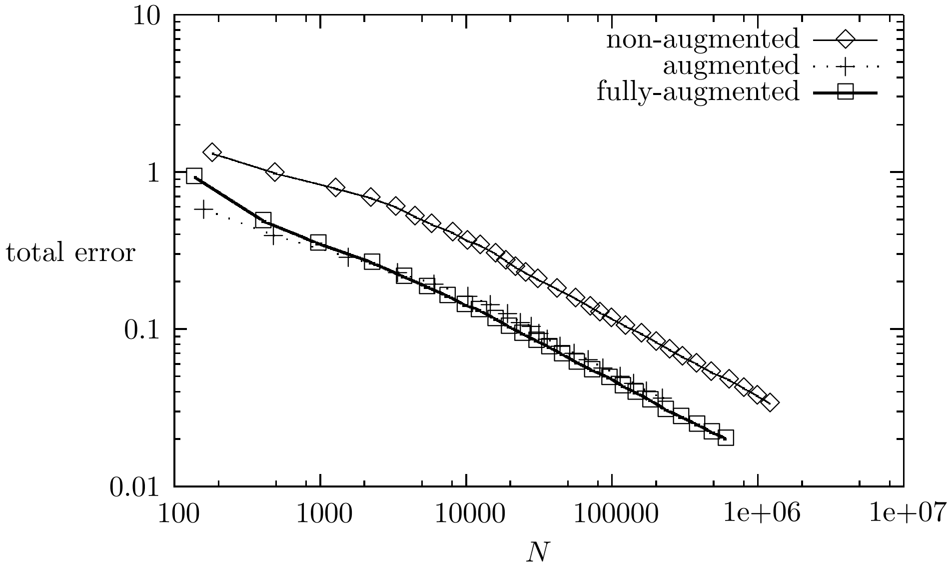

Figure 6.6: EXAMPLE 2, total error vs. N for the adaptive refinements

N	h	$\mathtt{e}(\mathbf{t})$	$\mathtt{e}(\boldsymbol{\sigma})$	$\mathtt{e}_1(\mathbf{u})$	$\mathtt{e}(\boldsymbol{\gamma})$	$\overline{\mathtt{e}}(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma})$	$\overline{\mathtt{r}}(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma})$	$\mathtt{ef}(\overline{\boldsymbol{\theta}})$
72	1/1	4.382E−00	8.486E−00	1.048E+01	7.985E−01	1.420E+01	–	0.8621
824	1/3	2.058E−00	6.466E−00	4.903E−00	3.122E−00	8.935E−00	1.203	0.7309
2206	1/5	1.321E−00	5.356E−00	3.093E−00	2.202E−00	6.696E−00	0.043	0.7667
4614	1/7	9.479E−01	4.609E−00	2.050E−00	1.815E−00	5.445E−00	0.337	0.7872
7382	1/9	7.771E−01	3.871E−00	1.258E−00	1.560E−00	4.428E−00	0.656	0.7971
11194	1/11	6.347E−01	3.940E−00	1.175E−00	1.407E−00	4.392E−00	-0.725	0.8183
15888	1/13	5.392E−01	3.616E−00	9.299E−01	1.191E−00	3.956E−00	-0.532	0.8395
26662	1/17	4.172E−01	3.297E−00	6.863E−01	9.214E−01	3.517E−00	0.253	0.8731
58920	1/25	2.802E−01	2.993E−00	4.572E−01	5.925E−01	3.098E−00	0.078	0.9263
113638	1/35	2.057E−01	2.834E−00	3.457E−01	4.663E−01	2.900E−0	-0.259	0.9462
229726	1/50	1.471E−01	2.368E−00	2.119E−01	3.478E−01	2.407E−00	1.282	0.9584
370046	1/63	1.150E−01	2.159E−00	1.522E−01	2.671E−01	2.184E−00	0.258	0.9703
457062	1/70	1.025E−01	2.110E−00	1.358E−01	2.584E−01	2.132E−00	0.226	0.9710
596632	1/80	8.987E−02	2.146E−00	1.301E−01	2.082E−01	2.162E−00	-0.104	0.9808
749216	1/90	8.001E−02	2.037E−00	1.102E−01	1.889E−01	2.050E−00	0.453	0.9826

Table 6.11: Example 3, quasi–uniform fully-augmented scheme (4.6)

N	h	$\mathtt{e}(\mathbf{t})$	$\mathtt{e}(\boldsymbol{\sigma})$	$\mathtt{e}_1(\mathbf{u})$	$\mathtt{e}(\boldsymbol{\gamma})$	$\overline{\mathtt{e}}(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma})$	$\overline{\mathtt{r}}(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma})$	$\mathtt{ef}(\overline{\boldsymbol{\theta}})$
72	1.000	4.382E−00	8.486E−00	1.048E+01	7.985E−01	1.420E+01	–	0.8621
828	0.500	2.085E−00	7.518E−00	5.603E−00	2.675E−00	9.971E−00	0.884	0.7717
2519	0.500	1.328E−00	5.213E−00	2.092E−00	1.539E−00	5.974E−00	1.014	0.8418
7419	0.354	8.816E−01	3.233E−00	1.143E−00	1.222E−00	3.746E−00	0.989	0.8128
9694	0.250	7.603E−01	2.827E−00	9.475E−01	9.652E−01	3.225E−00	1.041	0.8394
12043	0.250	6.977E−01	2.440E−00	8.382E−01	9.335E−01	2.831E−00	1.228	0.8165
16637	0.250	6.173E−01	2.006E−00	7.332E−01	8.056E−01	2.364E−00	1.018	0.8107
24816	0.250	4.949E−01	1.625E−00	5.711E−01	6.824E−01	1.917E−00	1.074	0.8093
38405	0.125	3.943E−01	1.315E−00	4.469E−01	5.334E−01	1.539E−00	0.903	0.8129
61030	0.125	3.147E−01	1.033E−00	3.475E−01	4.375E−01	1.216E−00	1.001	0.8049
94214	0.125	2.511E−01	8.360E−01	2.803E−01	3.495E−01	9.812E−01	0.999	0.8086
154303	0.088	1.990E−01	6.632E−01	2.210E−01	2.754E−01	7.772E−01	0.959	0.8099
195155	0.063	1.757E−01	5.871E−01	1.930E−01	2.459E−01	6.880E−01	1.039	0.8068
247607	0.063	1.563E−01	5.214E−01	1.708E−01	2.176E−01	6.106E−01	1.002	0.8089
305215	0.063	1.393E−01	4.701E−01	1.526E−01	1.982E−01	5.504E−01	0.993	0.8072

Table 6.12: Example 3, adaptive fully-augmented scheme (4.6)

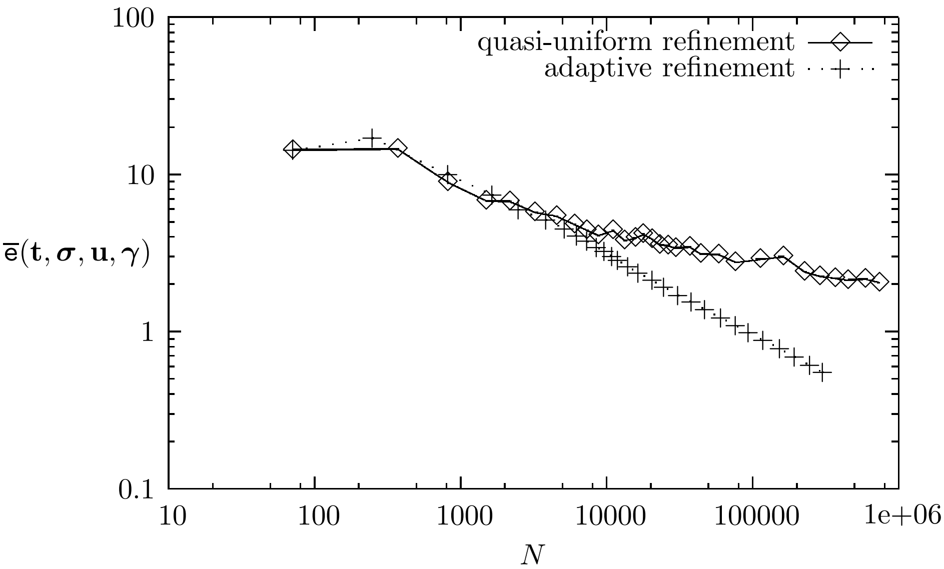

Figure 6.7: Example 3, $\overline{\mathtt{e}}(\mathbf{t},\boldsymbol{\sigma},\mathbf{u},\boldsymbol{\gamma})$ vs. N for fully-augmented scheme (4.6)

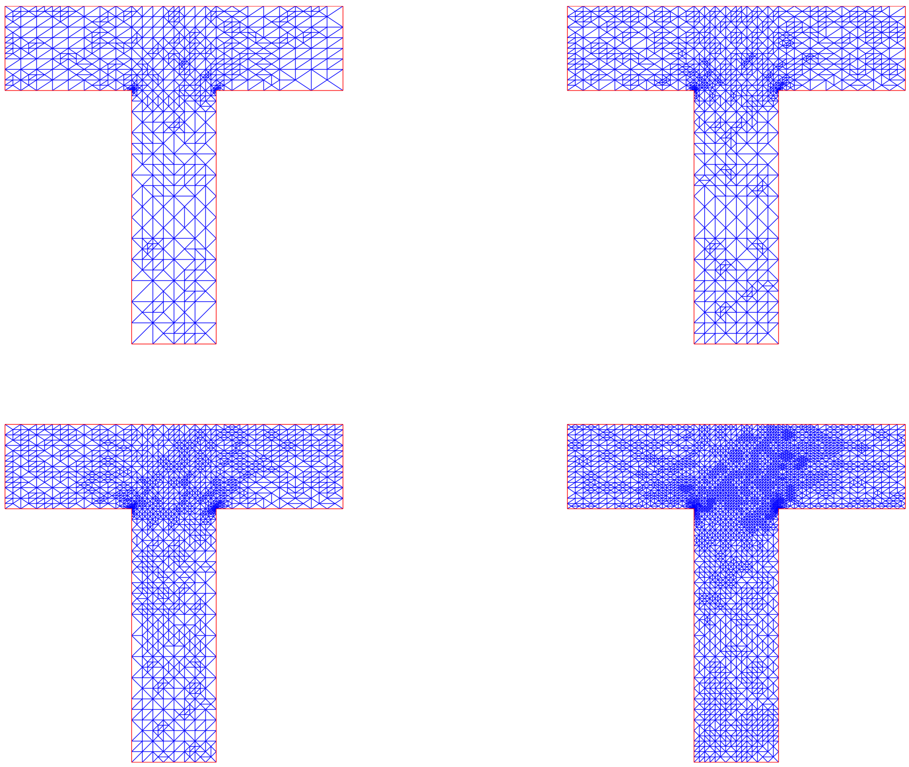

Figure 6.8: EXAMPLE 2 (augmented), adapted meshes for $N \in \{14792, 23877, 36300, 69790\}$

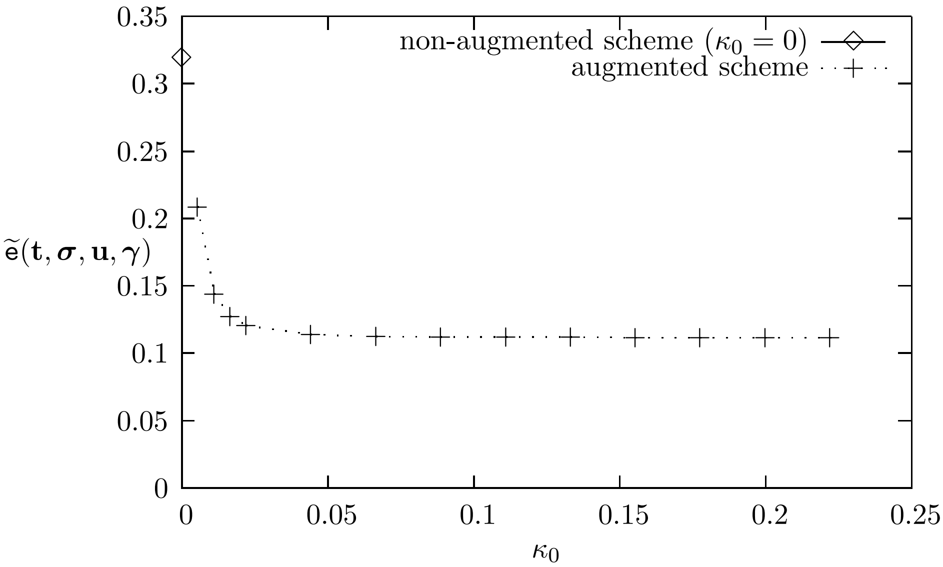

Figure 6.9: EXAMPLE 2, $\widetilde{\mathsf{e}}(\mathbf{t}, \boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\gamma})$ vs. κ_0 for $N = 215506$

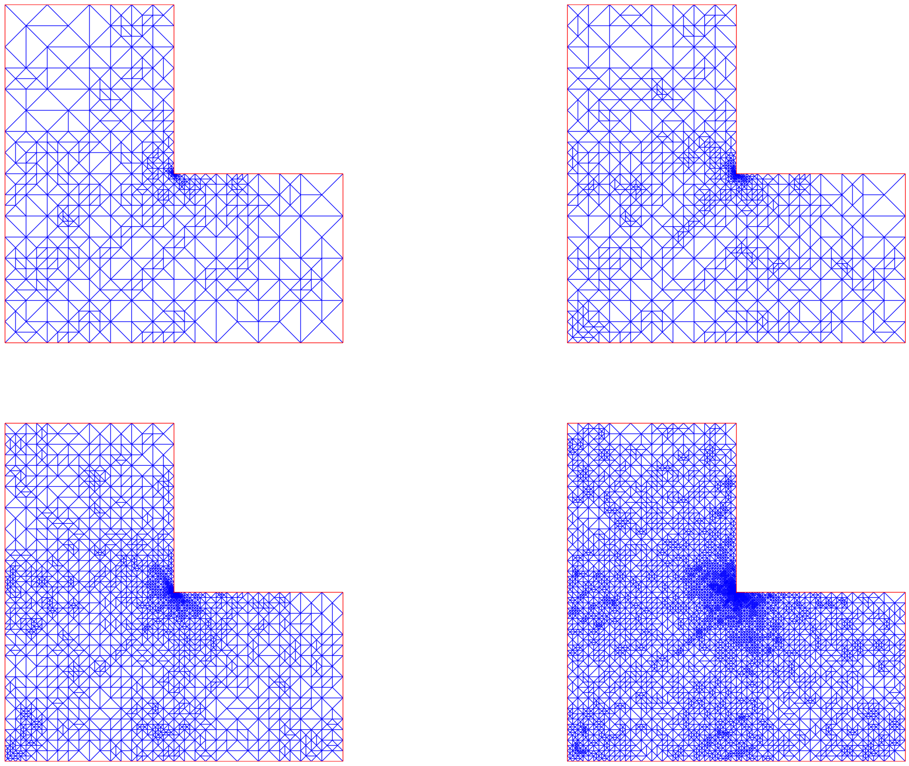

Figure 6.10: EXAMPLE 3 (fully-augmented), adapted meshes for $N \in \{9694, 16637, 38405, 94214\}$

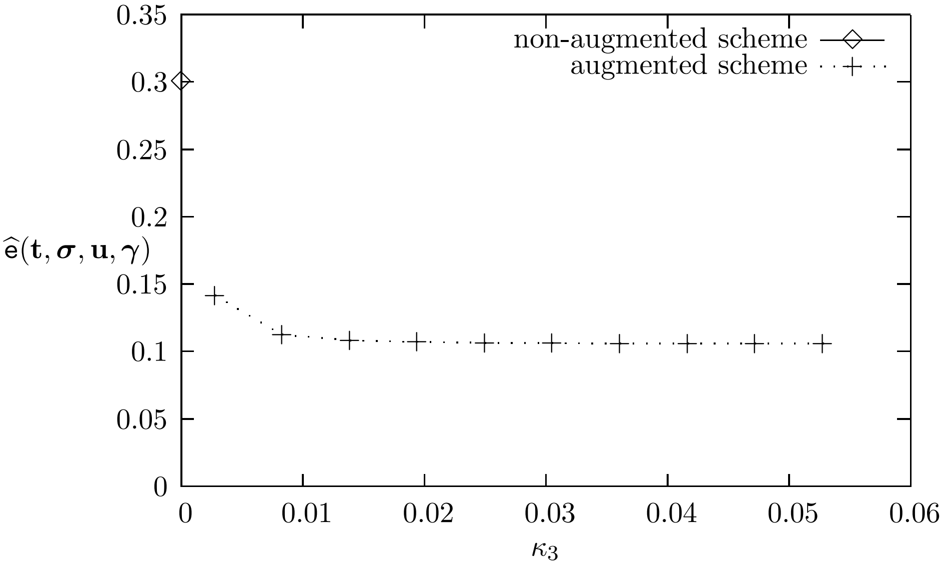

Figure 6.11: EXAMPLE 2, $\widehat{\mathsf{e}}(\mathbf{t}, \boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\gamma})$ vs. κ_3 for $N = 214624$

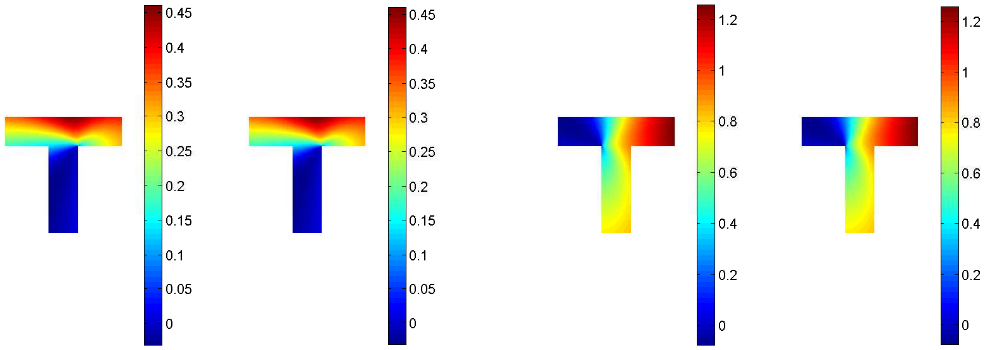

Figure 6.12: EXAMPLE 2, σ_{22} and t_{11} ($N = 174543$) for adaptive augmented scheme

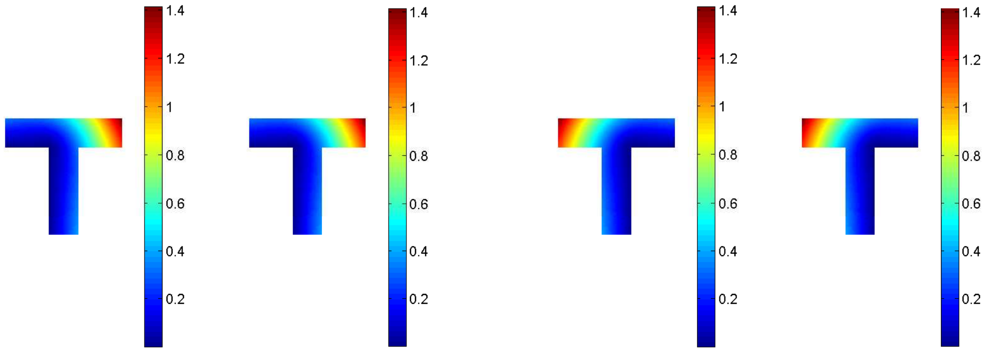

Figure 6.13: EXAMPLE 2, u_1 and u_2 ($N = 185623$) for adaptive fully-augmented scheme

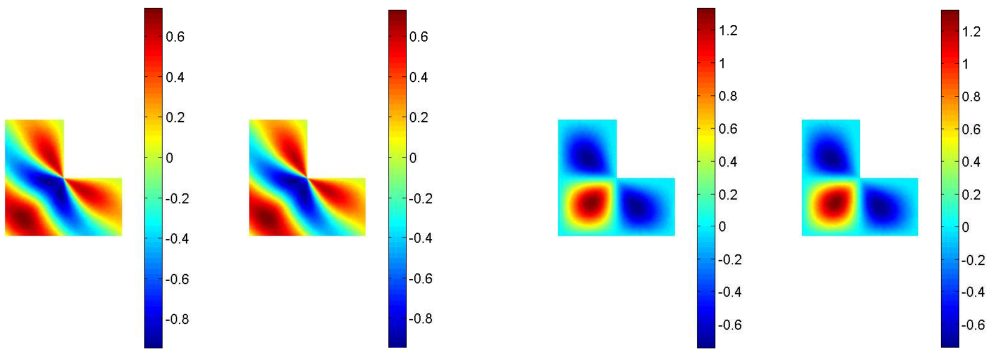

Figure 6.14: EXAMPLE 3, σ_{12} and u_2 ($N = 195155$) for adaptive fully-augmented scheme

References

[1] R. Araya, T.P. Barrios, G.N. Gatica, and N. Heuer, *A-posteriori error estimates for a mixed-FEM formulation of a non-linear elliptic problem.* Computer Methods in Applied Mechanics and Engineering, vol. 191, 21-22, pp. 2317-2336, (2002).

[2] D.N. Arnold, F. Brezzi, and J. Douglas, *PEERS: A new mixed finite element method for plane elasticity.* Japan Journal of Applied Mathematics, vol. 1, 2, pp. 347-367, (1984).

[3] D.N. Arnold, F. Brezzi, and M. Fortin, *A stable finite element method for the Stokes equations.* Calcolo, vol. 21, pp. 337-344, (1984).

[4] M.A. Barrientos, G.N. Gatica, N. Heuer, *An a-posteriori error estimate for a linear-nonlinear transmission problem in plane elastostatics.* Calcolo, vol. 39, 2, pp. 61-86, (2002).

[5] M. Barrientos, G.N. Gatica, and E.P. Stephan, *A mixed finite element method for nonlinear elasticity: two-fold saddle approach and a-posteriori error estimate.* Numerische Mathematik, vol. 91, 2, pp. 197-222, (2002).

[6] T.P. Barrios, E. Behrens, and M. González, *A posteriori error analysis of an augmented mixed formulation in linear elasticity with mixed and Dirichlet boundary conditions.* Computer Methods in Applied Mechanics and Engineering, vol. 200, 1-4, pp. 101-113, (2011).

[7] T.P. Barrios, G.N. Gatica, M. González, and N. Heuer, *A residual based a posteriori error estimator for an augmented mixed finite element method in linear elasticity.* ESAIM: Mathematical Modelling and Numerical Analysis, vol. 40, 5, pp. 843-869, (2006).

[8] S.C. Brenner and L.R. Scott, The Mathematical Theory of Finite Element Methods, Springer-Verlag New York, 1994.

[9] F. Brezzi and J. Douglas, *Stabilized mixed methods for the Stokes problem.* Numerische Mathematik, vol. 53, 1&2, pp. 225-235, (1988).

[10] F. Brezzi and M. Fortin, Mixed and Hybrid Finite Element Methods. Springer Verlag, 1991.

[11] F. Brezzi and M. Fortin, *A minimal stabilisation procedure for mixed finite element methods.* Numerische Mathematik, vol. 89, 3, pp. 457-491, (2001).

[12] F. Brezzi, M. Fortin, and L.D. Marini, *Mixed finite element methods with continuous stresses.* Mathematical Models and Methods in Applied Sciences, vol. 3, 2, pp. 275-287, (1993).

[13] C. Carstensen, *A posteriori error estimate for the mixed finite element method.* Mathematics of Computation, vol. 66, 218, pp. 465-476, (1997).

[14] Z. Chen, *Expanded mixed finite element methods for linear second-order elliptic problems.* ESAIM: Mathematical Modelling and Numerical Analysis, vol. 32, pp. 479-499, (1998).

[15] Z. Chen, *Expanded mixed finite element methods for quasilinear second-order elliptic problems.* ESAIM: Mathematical Modelling and Numerical Analysis, vol. 32, pp. 501-520, (1998).

[16] P.G. Ciarlet, The Finite Element Method for Elliptic Problems. North-Holland, 1978.

[17] P. Clément, *Approximation by finite element functions using local regularisation.* RAIRO Modélisation Mathématique et Analyse Numérique, vol. 9, pp. 77-84, (1975).

[18] J. Douglas and J. Wan, *An absolutely stabilized finite element method for the Stokes problem.* Mathematics of Computation, vol. 52, 186, pp. 495-508, (1989).

[19] V.J. Ervin, J.S. Howell, and I. Stanculescu, *A dual-mixed approximation method for a three-field model of a nonlinear generalized Stokes problem.* Computer Methods in Applied Mechanics and Engineering, vol. 197, 33-40, pp. 2886-2900, (2008).

[20] L. Figueroa, G.N. Gatica, and N. Heuer, *A priori and a posteriori error analysis of an augmented mixed finite element method for incompressible fluid flows.* Computer Methods in Applied Mechanics and Engineering, vol. 198, 2, pp. 280-291, (2008).

[21] L. Figueroa, G.N. Gatica, and A. Márquez, *Augmented mixed finite element methods for the stationary Stokes Equations.* SIAM Journal on Scientific Computing, vol. 31, 2, pp. 1082-1119, (2008).

[22] L.P. Franca, New Mixed Finite Element Methods. Ph.D. Thesis, Stanford University, 1987.

[23] L.P. Franca and T.J.R. Hughes, *Two classes of finite element methods.* Computer Methods in Applied Mechanics and Engineering, vol. 69, pp. 89-129, (1988).

[24] L. Franca and A. Loula, *A new mixed finite element method for the Timoshenko beam problem.* Modélisation Mathématique et Analyse Numérique, vol. 25, pp. 561-578, (1991).

[25] L.P. Franca and A. Russo, *Unlocking with residual-free bubbles.* Computer Methods in Applied Mechanics and Engineering, vol. 142, 3&4, pp. 361-364, (1997).

[26] L.P. Franca and R. Stenberg, *Error analysis of Galerkin least squares methods for the elasticity equations.* SIAM Journal on Numerical Analysis, vol. 28, 6, pp. 1680-1697, (1991).

[27] A.I. Garralda-Guillem, G.N. Gatica, A. Márquez, and M. Ruiz Galán, *A posteriori error analysis of twofold saddle point variational formulations for nonlinear boundary value problems.* Preprint 2012-07, Centro de Investigación en Ingeniería Matemática, Universidad de Concepción, Chile, (2012).

[28] G.N. Gatica, *Solvability and Galerkin approximations of a class of nonlinear operator equations.* Zeitschrif für Analysis und ihre Anwendungen, vol. 21, 3, pp. 761-781, (2002).

[29] G.N. Gatica, *Analysis of a new augmented mixed finite element method for linear elasticity allowing* $\mathbb{RT}_0-\mathbb{P}_1-\mathbb{P}_0$ *approximations.* ESAIM: Mathematical Modelling and Numerical Analysis, vol. 40, 1, pp. 1-28, (2006).

[30] G.N. Gatica, *An augmented mixed finite element method for linear elasticity with non-homogeneous Dirichlet conditions.* Electronic Transactions on Numerical Analysis, vol. 26, pp. 421-438, (2007).

[31] G.N. Gatica, L.F. Gatica, and E.P. Stephan, *A dual-mixed finite element method for nonlinear incompressible elasticity with mixed boundary conditions.* Computer Methods in Applied Mechanics and Engineering, vol. 196, 35-36, pp. 3348-3369, (2007).

[32] G.N. Gatica, M. González, and S. Meddahi, *A low-order mixed finite element method for a class of quasi-Newtonian Stokes flows. Part I: a-priori error analysis.* Computer Methods in Applied Mechanics and Engineering, vol. 193, 9-11, pp. 881-892, (2004).

[33] G.N. Gatica, M. González, and S. Meddahi, *A low-order mixed finite element method for a class of quasi-Newtonian Stokes flows. Part II: a-posteriori error analysis.* Computer Methods in Applied Mechanics and Engineering, vol. 193, 9-11, pp. 893-911, (2004).

[34] G.N. Gatica and N. Heuer, *A dual-dual formulation for the coupling of mixed-FEM and BEM in hyperelasticity.* SIAM Journal on Numerical Analysis, vol. 38, 2, pp. 380-400, (2000).

[35] G.N. Gatica and N. Heuer, *Minimum residual iteration for a dual-dual mixed formulation of exterior transmission problems.* Numerical Linear Algebra with Applications, vol. 8, 3, pp. 147-164, (2001).

[36] G.N. Gatica and N. Heuer, *An expanded mixed finite element approach via a dual-dual formulation and the minimum residual method.* Journal of Computational and Applied Mathematics, vol. 132, 2, pp. 371-385, (2001).

[37] G.N. Gatica and N. Heuer, *Conjugate gradient method for dual-dual mixed formulations.* Mathematics of Computation, vol. 71, 240, pp. 1455-1472, (2002).

[38] G.N. Gatica and N. Heuer, *A preconditioned MINRES method for the coupling of MIXED-FEM and BEM for some nonlinear problems.* SIAM Journal on Scientific Computing, vol. 24, 2, pp. 572-596, (2002).

[39] G.N. Gatica, N. Heuer, and S. Meddahi, *On the numerical analysis of nonlinear twofold saddle point problems.* IMA Journal of Numerical Analysis, vol. 23, 2, pp. 301-330, (2003).

[40] G.N. Gatica, A. Márquez, and S. Meddahi *An augmented mixed finite element method for 3D linear elasticity problems.* Journal of Computational and Applied Mathematics, vol. 231, 2, pp. 526-540, (2009).

[41] G.N. Gatica, A. Márquez, A. and M.A. Sánchez, *Analysis of a velocity-pressure-pseudo-stress formulation for the stationary Stokes equations.* Computer Methods in Applied Mechanics and Engineering, vol. 199, 17-20, pp. 1064-1079, (2010).

[42] G.N. Gatica, A. Márquez, and M. Sánchez, *A priori and a posteriori error analyses of a velocity-pseudostress formulation for a class of quasi-Newtonian Stokes flows.* Computer Methods in Applied Mechanics and Engineering, vol. 200, 17-20, pp. 1619-1636, (2011).

[43] G.N. Gatica and S. Meddahi, *A dual-dual mixed formulation for nonlinear exterior transmission problems.* Mathematics of Computation, vol. 70, 236, pp. 1461-1480, (2001).

[44] G.N. Gatica and E.P. Stephan, *A mixed-FEM formulation for nonlinear incompressible elasticity in the plane.* Numerical Methods for Partial Differential Equations, vol. 18, 1, pp. 105-128, (2002).

[45] G.N. Gatica and W.L. Wendland, *Coupling of mixed finite elements and boundary elements for linear and nonlinear elliptic problems.* Applicable Analysis, vol. 63, pp. 39-75, (1996).

[46] G.N. Gatica and W.L. Wendland, *Coupling of mixed finite elements and boundary elements for a hyperelastic interface problem.* SIAM Journal on Numerical Analysis, vol. 34, 6, pp. 2335-2356, (1997).

[47] V. Girault and P.-A. Raviart, Finite Element Methods for Navier-Stokes Equations. Theory and Algorithms. Springer Series in Computational Mathematics, vol. 5, Springer-Verlag, 1986.

[48] R. Hiptmair, *Finite elements in computational electromagnetism.* Acta Numerica, vol. 11, pp. 237-339, (2002).

[49] J.S. Howell, *Dual-mixed finite element approximation of Stokes and nonlinear Stokes problems using trace-free velocity gradients.* Journal of Computational and Applied Mathematics, vol. 231, 2, pp. 780-792, (2009).

[50] M. Lonsing and R. Verfürth, *On the stability of BDMS and PEERS elements.* Numerische Mathematik, vol. 99, 1, pp. 131-140, (2004).

[51] A.F.D. Loula and J.N.C. Guerreiro, *Finite element analysis of nonlinear creeping flows.* Computer Methods in Applied Mechanics and Engineering, vol. 79, 1, pp. 87-109, (1990).

[52] A. Masud and T.J.R. Hughes, *A stabilized mixed finite element method for Darcy flow.* Computer Methods in Applied Mechanics and Engineering, vol. 191, 39&40, pp. 4341-4370, (2002).

[53] J. Nečas, Introduction to Theory of Nonlinear Elliptic Equations. University of Prague, Czechoslovakia, 1983.

[54] A. Quarteroni and A. Valli, Numerical Approximation of Partial Differential Equations. Springer Series in Computational Mathematics, vol. 23. Springer-Verlag Berlin Heidelberg, 1994.

[55] J.E. Roberts and J.M. Thomas, Mixed and Hybrid Methods. In: Handbook of Numerical Analysis, edited by P.G. Ciarlet and J.L. Lions, vol. II, Finite Element Methods (Part 1), 1991, North-Holland, Amsterdam.

[56] D. Sandri, *Sur l'approximation numérique des écoulements quasi-Newtoniens dont la viscosité suit la loi puissance ou la loi de Carreau.* Mathematical Modelling and Numerical Analysis, vol. 27, 2, pp. 131-155, (1993).

[57] B. Scheurer, *Existence et approximation de point selles pour certain problemes non linéaires.* R.A.I.R.O. Analyse Numérique, vol. 11, 4, pp. 369-400, (1977).

[58] R. Verfürth, A Review of A-Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques. John Wiley and Teubner Series. Advances in Numerical Mathematics 1996.

[59] E. Zeidler, *Nonlinear Functional Analysis and its Applications IV.* New York:Springer, 1988.