On the numerical analysis of nonlinear two-fold saddle point problems *

Gabriel N. Gatica[1], Norbert Heuer[1] and Salim Meddahi[2]

[1] *Departamento de Ingeniería Matemática, Universidad de Concepción, Casilla 160-C, Concepción, Chile, email:* {ggatica,norbert}@ing-mat.udec.cl

[2] *Departamento de Matemática, Universidad de Oviedo, Calvo Sotelo s/n, 33007 Oviedo, España, email:* salim@orion.ciencias.uniovi.es

Abstract

We provide a general abstract theory for the solvability and Galerkin approximations of nonlinear two-fold saddle point problems. In particular, a Strang error estimate containing the consistency terms arising from the approximation of the continuous operators involved is deduced. Then we apply these results to analyse a fully discrete Galerkin scheme for a two-fold saddle point formulation of a nonlinear elliptic boundary value problem in divergence form. Some numerical results are also presented.

1. Introduction

Let X_1, M_1 and M be Hilbert spaces with duals X_1', M_1' and M', respectively. We consider linear bounded operators $\mathbf{S} : M_1 \to M_1'$, $\mathbf{B}_1 : X_1 \to M_1'$ and $\mathbf{B} : M_1 \to M'$, with corresponding adjoints $\mathbf{B}_1^* : M_1 \to X_1'$ and $\mathbf{B}^* : M \to M_1'$. Also, let $\mathbf{A}_1 : X_1 \to X_1'$ be an operator which may be linear or nonlinear. Then, given $(\mathbf{F}_1, \mathbf{G}_1, \mathbf{G}) \in X_1' \times M_1' \times M'$, we are interested in the following variational problem: *Find* $(\mathbf{t}, \boldsymbol{\sigma}, \mathbf{u}) \in X_1 \times M_1 \times M$ *such that*

$$\begin{bmatrix} \mathbf{A}_1 & \mathbf{B}_1^* & \mathbf{O} \\ \mathbf{B}_1 & -\mathbf{S} & \mathbf{B}^* \\ \mathbf{O} & \mathbf{B} & \mathbf{O} \end{bmatrix} \begin{bmatrix} \mathbf{t} \\ \boldsymbol{\sigma} \\ \mathbf{u} \end{bmatrix} = \begin{bmatrix} \mathbf{F}_1 \\ \mathbf{G}_1 \\ \mathbf{G} \end{bmatrix}, \tag{1}$$

*The work of G. N. Gatica and N. Heuer was partially supported by Fondecyt-Chile through the project No. 1980122 and the Program A on Numerical Analysis of the FONDAP in Applied Mathematics. The work of S. Meddahi was partially supported by D.G.E.S. through the project PB98-1564.

or equivalently,

$$\begin{array}{lllll} [\mathbf{A}_1(\mathbf{t}),\mathbf{s}] & +[\mathbf{B}_1^*(\boldsymbol{\sigma}),\mathbf{s}] & & = & [\mathbf{F}_1,\mathbf{s}]\,, \\ [\mathbf{B}_1(\mathbf{t}),\boldsymbol{\tau}] & -[\mathbf{S}(\boldsymbol{\sigma}),\boldsymbol{\tau}] & +[\mathbf{B}^*(\mathbf{u}),\boldsymbol{\tau}] & = & [\mathbf{G}_1,\boldsymbol{\tau}]\,, \\ & [\mathbf{B}(\boldsymbol{\sigma}),\mathbf{v}] & & = & [\mathbf{G},\mathbf{v}]\,, \end{array} \tag{2}$$

for all $(\mathbf{s},\boldsymbol{\tau},\mathbf{v}) \in X_1 \times M_1 \times M$.

Hereafter, $\mathbf{O}$ stands for a generic null operator/functional, and $[\cdot,\cdot]$ denotes the duality pairing induced by the operators and functionals used in each case. Note that $[\mathbf{B}_1^*(\boldsymbol{\sigma}),\mathbf{s}] = [\mathbf{B}_1(\mathbf{s}),\boldsymbol{\sigma}]$ and $[\mathbf{B}^*(\mathbf{u}),\boldsymbol{\tau}] = [\mathbf{B}(\boldsymbol{\tau}),\mathbf{u}]$.

Because of the structure of (1) (or (2)), these operator equations have been called *dual-dual variational formulations* and also *two-fold saddle point problems*. They have been introduced recently for solving a large class of linear and nonlinear boundary value problems via the dual-mixed finite element method (FEM) and the combined use of it with either the boundary element method (BEM) or Dirichlet-to-Neumann (DtN) mappings (see, e.g., [7–12] and [1]). We remark that an extended Babuška-Brezzi theory was introduced in [7] to deal with the particular case of (1) given by $\mathbf{S} = \mathbf{O}$. In there, sufficient conditions for the existence and uniqueness of solution to the continuous and discrete formulations, and the corresponding error estimates, were provided. Furthermore, an application of the abstract results from [7] to the combination of dual-mixed FEM with a DtN mapping (based on BEM) for a nonlinear exterior transmission problem in potential theory was described in [12].

In addition, the dual-dual approach from [12] was extended in [11] to the mixed-FEM/BEM formulation of the hyperelastic interface problem in [15]. The work in [11] is the first one providing complete continuous and discrete analyses, including the definition of explicit finite element subspaces satisfying the compatibility conditions, for the coupling of mixed-FEM and BEM, as applied to a nonlinear transmission problem. However, the resulting variational formulation did not fit into the framework of [7] since it became a matrix operator equation of the form (1) with a non-null operator $\mathbf{S}$. Thus, some modifications of the theory from [7] were given in [11] to properly analyse that interface problem. It is also important to observe that none of the above mentioned works has considered the numerical analysis of fully discrete Galerkin schemes (arising mainly from the use of numerical integration) for the dual-dual variational formulations. This is particularly necessary and unavoidable when $\mathbf{A}_1$ is nonlinear.

The purpose of this paper is to unify the main abstract results from [7] and [11] and to provide a complete and general theory for the solvability and Galerkin approximations of the nonlinear two-fold saddle point problem (1). Our present approach introduces two new distinguishable aspects. On one hand, we consider discrete operators $\mathbf{A}_{1,h}$, $\mathbf{B}_{1,h}$, $\mathbf{B}_{1,h}^*$, $\mathbf{S}_h$, $\mathbf{B}_h$ and $\mathbf{B}_h^*$, that do not necessarily coincide with the corresponding continuous ones, thus allowing the utilisation of fully discrete Galerkin schemes for the numerical analysis. On the other hand, we add some assumptions on the first-order Gâteaux derivative of $\mathbf{A}_1$ (commonly found in applications), which yield a much simpler error analysis than the ones provided in [7] and [11]. However, the results in these references are more general since no further properties of the nonlinear operator $\mathbf{A}_1$ were considered there. Anyway, with the analysis developed in the present work, the Strang and Cea error estimates are easily obtained.

The rest of the paper is organised as follows. In Section 2 we provide sufficient conditions for the existence of unique solution to the continuous two-fold saddle point formulation (1). The solvability of the associated Galerkin scheme and the corresponding error analysis are given in Section 3. We deduce there a general Strang error estimate that contains the consistency terms arising from the utilisation of fully discrete operators. Then, an application of our theory to a nonlinear elliptic problem in divergence form, appearing in the computation of magnetic fields and in steady heat conduction, is fully described in Section 4. Finally, several numerical results for the problem in Section 4 are presented in Section 5.

2. The continuous two-fold saddle point formulation

From now on, we always assume that the linear operators $\mathbf{S}$, $\mathbf{B}_1$ and $\mathbf{B}$ are bounded. In order to prove our main result for the solvability of the continuous formulation (2), we first need to consider an auxiliary problem. More precisely, given a closed subspace $\hat{M}_1$ of M_1 and $(\hat{\mathbf{F}}_1, \hat{\mathbf{G}}_1) \in X_1' \times \hat{M}_1'$, we look for $(\hat{\mathbf{t}}, \hat{\boldsymbol{\sigma}}) \in X_1 \times \hat{M}_1$ such that

$$\begin{aligned} [\mathbf{A}_1(\hat{\mathbf{t}}), \mathbf{s}] \,+\, [\mathbf{B}_1^*(\hat{\boldsymbol{\sigma}}), \mathbf{s}] &= [\hat{\mathbf{F}}_1, \mathbf{s}]\,, \\ [\mathbf{B}_1(\hat{\mathbf{t}}), \boldsymbol{\tau}] \,-\, [\mathbf{S}(\hat{\boldsymbol{\sigma}}), \boldsymbol{\tau}] &= [\hat{\mathbf{G}}_1, \boldsymbol{\tau}]\,, \end{aligned} \tag{3}$$

for all $(\mathbf{s}, \boldsymbol{\tau}) \in X_1 \times \hat{M}_1$.

Then, we have the following lemma.

Lemma 2.1: *Assume that:*

i) *the operator $\mathbf{A}_1 : X_1 \to X_1'$ is nonlinear, Lipschitz continuous (with a Lipschitz constant $\gamma > 0$) and strongly monotone (with a monotonicity constant $\alpha > 0$).*

ii) *the linear operator $\mathbf{S}$ is positive semi-definite on $\hat{M}_1$, that is*

$$[\mathbf{S}(\boldsymbol{\tau}), \boldsymbol{\tau}] \geq 0 \quad \forall\, \boldsymbol{\tau} \in \hat{M}_1\,.$$

iii) *the linear operator $\mathbf{B}_1$ satisfies an inf-sup condition on $X_1 \times \hat{M}_1$, that is, there exists $\beta_1 > 0$ such that*

$$\sup_{\substack{\mathbf{s} \in X_1 \\ \mathbf{s} \neq 0}} \frac{[\mathbf{B}_1(\mathbf{s}), \boldsymbol{\tau}]}{\|\mathbf{s}\|_{X_1}} \geq \beta_1 \, \|\boldsymbol{\tau}\|_{M_1} \quad \forall\, \boldsymbol{\tau} \in \hat{M}_1\,.$$

Then, there exists a unique solution $(\hat{\mathbf{t}}, \hat{\boldsymbol{\sigma}}) \in X_1 \times \hat{M}_1$ *of* (3)*. In addition, there exists* $C > 0$*, depending only on* γ*,* α*,* β_1 *and* $\|\mathbf{B}_1\|$*, such that*

$$\|(\hat{\mathbf{t}}, \hat{\boldsymbol{\sigma}})\| \,\leq\, C \left\{ \|\hat{\mathbf{F}}_1\| + \|\hat{\mathbf{G}}_1\| + \|\mathbf{A}_1(0)\| \right\}. \tag{4}$$

Proof: Due to a classical result in nonlinear functional analysis (see, e.g., Theorem 3.3.23 in [20]), we deduce from i) that $\mathbf{A}_1 : X_1 \to X_1'$ is an isomorphism. Then, block eliminations show that (3) is uniquely solvable if and only if the nonlinear operator $\mathbf{S} + \mathbf{B}_1 \mathbf{A}_1^{-1} \mathbf{B}_1^* : \hat{M}_1 \to \hat{M}_1'$ is invertible (see, e.g., Theorem 2.1 in [5]).

Since $\mathbf{S}$ is bounded and positive semi-definite, we only need to prove that $\mathbf{B}_1 \mathbf{A}_1^{-1} \mathbf{B}_1^*$ is strongly monotone and Lipschitz continuous. In fact, for any $\boldsymbol{\tau}_1$, $\boldsymbol{\tau}_2 \in \hat{M}_1$, define $\mathbf{s}_i := \mathbf{A}_1^{-1}(\mathbf{B}_1^*(\boldsymbol{\tau}_i)) \in X_1$, $i = 1, 2$. Then there holds

$$\begin{aligned}
&[\mathbf{B}_1 \mathbf{A}_1^{-1} \mathbf{B}_1^*(\boldsymbol{\tau}_1) - \mathbf{B}_1 \mathbf{A}_1^{-1} \mathbf{B}_1^*(\boldsymbol{\tau}_2), \boldsymbol{\tau}_1 - \boldsymbol{\tau}_2] \\
&= [\mathbf{A}_1^{-1}(\mathbf{B}_1^*(\boldsymbol{\tau}_1)) - \mathbf{A}_1^{-1}(\mathbf{B}_1^*(\boldsymbol{\tau}_2)), \mathbf{B}_1^*(\boldsymbol{\tau}_1) - \mathbf{B}_1^*(\boldsymbol{\tau}_2)] \\
&= [\mathbf{s}_1 - \mathbf{s}_2, \mathbf{A}_1(\mathbf{s}_1) - \mathbf{A}_1(\mathbf{s}_2)] \\
&\geq \alpha \, \|\mathbf{s}_1 - \mathbf{s}_2\|_{X_1}^2 \,\geq\, \frac{\alpha}{\gamma^2} \, \|\mathbf{A}_1(\mathbf{s}_1) - \mathbf{A}_1(\mathbf{s}_2)\|_{X_1'}^2 \\
&= \frac{\alpha}{\gamma^2} \, \|\mathbf{B}_1^*(\boldsymbol{\tau}_1 - \boldsymbol{\tau}_2)\|_{X_1'}^2 \,\geq\, \beta_1^2 \, \frac{\alpha}{\gamma^2} \, \|\boldsymbol{\tau}_1 - \boldsymbol{\tau}_2\|_{M_1}^2 \,,
\end{aligned} \tag{5}$$

which proves the strong monotonicity of $\mathbf{B}_1 \mathbf{A}_1^{-1} \mathbf{B}_1^*$.

The last inequality follows from the inf-sup condition for $\mathbf{B}_1$, which can equivalently be stated as $\|\mathbf{B}_1^*(\boldsymbol{\tau})\|_{X_1'} \,\geq\, \beta_1 \, \|\boldsymbol{\tau}\|_{M_1}$ for all $\boldsymbol{\tau} \in \hat{M}_1$.

Now we show that the operator $\mathbf{A}_1^{-1} : X_1' \to X_1$ is Lipschitz continuous. Let $\mathbf{s}_1$, $\mathbf{s}_2 \in X_1$ and consider $\mathbf{A}_1(\mathbf{s}_i) \in X_1'$, $i = 1, 2$. Then we have

$$\begin{aligned}
&\|\mathbf{A}_1^{-1}(\mathbf{A}_1(\mathbf{s}_1)) - \mathbf{A}_1^{-1}(\mathbf{A}_1(\mathbf{s}_2))\|_{X_1}^2 \,=\, \|\mathbf{s}_1 - \mathbf{s}_2\|_{X_1}^2 \\
&\leq\, \frac{1}{\alpha} \, [\mathbf{A}_1(\mathbf{s}_1) - \mathbf{A}_1(\mathbf{s}_2), \mathbf{s}_1 - \mathbf{s}_2] \\
&\leq\, \frac{1}{\alpha} \, \|\mathbf{A}_1(\mathbf{s}_1) - \mathbf{A}_1(\mathbf{s}_2)\|_{X_1'} \, \|\mathbf{s}_1 - \mathbf{s}_2\|_{X_1} \,,
\end{aligned}$$

whence

$$\|\mathbf{A}_1^{-1}(\mathbf{A}_1(\mathbf{s}_1)) - \mathbf{A}_1^{-1}(\mathbf{A}_1(\mathbf{s}_2))\|_{X_1} \leq\, \frac{1}{\alpha} \, \|\mathbf{A}_1(\mathbf{s}_1) - \mathbf{A}_1(\mathbf{s}_2)\|_{X_1'} \,. \tag{6}$$

Thus, the Lipschitz continuity of $\mathbf{B}_1 \mathbf{A}_1^{-1} \mathbf{B}_1^*$ follows from the linearity and boundedness of $\mathbf{B}_1$ and from the Lipschitz continuity of $\mathbf{A}_1^{-1}$.

Now, for the estimate (4) we first observe from (3) that

$$\mathbf{A}_1(\hat{\mathbf{t}}) = \hat{\mathbf{F}}_1 - \mathbf{B}_1^*(\hat{\boldsymbol{\sigma}}) \tag{7}$$

and

$$(\mathbf{S} + \mathbf{B}_1 \mathbf{A}_1^{-1} \mathbf{B}_1^*)(\hat{\boldsymbol{\sigma}}) \;=\; \mathbf{B}_1 \mathbf{A}_1^{-1} \hat{\mathbf{F}}_1 - \hat{\mathbf{G}}_1 \,. \tag{8}$$

Then, using ii), (5) and (8), we get

$$\|\hat{\boldsymbol{\sigma}}\|_{\tilde{M}_1} \;\le\; \frac{\gamma^2}{\alpha\,\beta_1^2} \left\{ \|\mathbf{B}_1\| \, \Big[\|\mathbf{A}_1^{-1}(\hat{\mathbf{F}}_1)\| + \|\mathbf{A}_1^{-1}(\mathbf{O})\| \Big] \;+\; \|\hat{\mathbf{G}}_1\| \right\} . \tag{9}$$

Next, by the strong monotonicity of $\mathbf{A}_1$ (cf. i)) we obtain that

$$\|\mathbf{A}_1^{-1}(\hat{\mathbf{F}}_1)\| \le \frac{1}{\alpha} \left\{ \|\hat{\mathbf{F}}_1\| + \|\mathbf{A}_1(0)\| \right\} \tag{10}$$

and

$$\|\mathbf{A}_1^{-1}(\mathbf{O})\| \le \frac{1}{\alpha} \, \|\mathbf{A}_1(0)\| \,. \tag{11}$$

In addition, using now (7) and the same property of $\mathbf{A}_1$, we deduce that

$$\|\hat{\mathbf{t}}\| \;\le\; \frac{1}{\alpha} \left\{ \|\hat{\mathbf{F}}_1\| + \|\mathbf{B}_1\| \, \|\hat{\boldsymbol{\sigma}}\| + \|\mathbf{A}_1(0)\| \right\} . \tag{12}$$

In this way, (9)–(12) yield (4), which completes the proof of the lemma. $\square$

We are now in a position to state sufficient conditions for the unique solvability of (1) (or (2)). To this end, we first set $\mathbf{X} := X_1 \times M_1 \times M$ and define the global operator $\mathbf{P} : \mathbf{X} \to \mathbf{X}'$ associated to the variational formulation (2), that is,

$$\begin{aligned} [\mathbf{P}(\mathbf{r}, \boldsymbol{\rho}, \mathbf{w}), (\mathbf{s}, \boldsymbol{\tau}, \mathbf{v})] &:= [\mathbf{A}_1(\mathbf{r}), \mathbf{s}] + [\mathbf{B}_1^*(\boldsymbol{\rho}), \mathbf{s}] + [\mathbf{B}_1(\mathbf{r}), \boldsymbol{\tau}] \\ &\quad - [\mathbf{S}(\boldsymbol{\rho}), \boldsymbol{\tau}] + [\mathbf{B}^*(\mathbf{w}), \boldsymbol{\tau}] + [\mathbf{B}(\boldsymbol{\rho}), \mathbf{v}] \,, \end{aligned} \tag{13}$$

for all $(\mathbf{r}, \boldsymbol{\rho}, \mathbf{w})$, $(\mathbf{s}, \boldsymbol{\tau}, \mathbf{v}) \in \mathbf{X}$.

Then we have the following main theorem.

THEOREM 2.1: *Let* $\tilde{M}_1 := \text{kernel}\,(\mathbf{B})$*, that is*

$$\tilde{M}_1 := \{\, \boldsymbol{\tau} \in M_1 : \quad [\mathbf{B}(\boldsymbol{\tau}), \mathbf{v}] = 0 \quad \forall\, \mathbf{v} \in M \} \,.$$

Assume that:

i) *the operator* $\mathbf{A}_1 : X_1 \to X_1'$ *is nonlinear, Lipschitz continuous (with a Lipschitz constant* $\gamma > 0$*) and strongly monotone (with a monotonicity constant* $\alpha > 0$*).*

ii) *the linear operator* $\mathbf{S}$ *is positive semi-definite on* $\tilde{M}_1$*, that is*

$$[\mathbf{S}(\boldsymbol{\tau}), \boldsymbol{\tau}] \ge 0 \quad \forall\, \boldsymbol{\tau} \in \tilde{M}_1 \,.$$

iii) *the linear operator* $\mathbf{B}_1$ *satisfies an inf-sup condition on* $X_1 \times \tilde{M}_1$*, that is, there exists* $\beta_1 > 0$ *such that*

$$\sup_{\substack{\mathbf{s} \in X_1 \\ \mathbf{s} \neq 0}} \frac{[\mathbf{B}_1(\mathbf{s}), \boldsymbol{\tau}]}{\|\mathbf{s}\|_{X_1}} \;\ge\; \beta_1 \, \|\boldsymbol{\tau}\|_{M_1} \quad \forall\, \boldsymbol{\tau} \in \tilde{M}_1 \,.$$

iv) *the linear operator* $\mathbf{B}$ *satisfies an inf-sup condition on* $M_1 \times M$*, that is, there exists* $\beta > 0$ *such that*

$$\sup_{\substack{\boldsymbol{\tau} \in M_1 \\ \boldsymbol{\tau} \neq 0}} \frac{[\mathbf{B}(\boldsymbol{\tau}), \mathbf{v}]}{\|\boldsymbol{\tau}\|_{M_1}} \geq \beta \|\mathbf{v}\|_M \quad \forall\, \mathbf{v} \in M\,.$$

Then, given $\mathbf{F} := (\mathbf{F}_1, \mathbf{G}_1, \mathbf{G}) \in \mathbf{X}' := X_1' \times M_1' \times M'$*, there exists a unique* $(\mathbf{t}, \boldsymbol{\sigma}, \mathbf{u}) \in \mathbf{X}$ *solution of* (2)*, that is such that*

$$[\mathbf{P}(\mathbf{t}, \boldsymbol{\sigma}, \mathbf{u}), (\mathbf{s}, \boldsymbol{\tau}, \mathbf{v})] \;=\; [\mathbf{F}, (\mathbf{s}, \boldsymbol{\tau}, \mathbf{v})] \quad \forall\, (\mathbf{s}, \boldsymbol{\tau}, \mathbf{v}) \in \mathbf{X}\,. \tag{14}$$

In addition, there exists $\tilde{C} > 0$*, depending only on* γ*,* α*,* β_1*,* $\|\mathbf{B}_1\|$*,* β *and* $\|\mathbf{S}\|$*, such that*

$$\|(\mathbf{t}, \boldsymbol{\sigma}, \mathbf{u})\|_{\mathbf{X}} \;\leq\; \tilde{C}\, \left\{ \|\mathbf{F}\|_{\mathbf{X}'} \,+\, \|\mathbf{A}_1(0)\|_{X_1'} \right\}\,. \tag{15}$$

Proof: We adapt the analysis from [16] to the present situation. First, according to the inf-sup condition for $\mathbf{B}$ (cf. iv)) and Lemma 4.1 in Chapter I of [16], we know that $\mathbf{B} : \tilde{M}_1^{\perp} \to M'$ and $\mathbf{B}^* : M \to \tilde{M}_1^{o}$ are isomorphisms with

$$\|\mathbf{B}^{-1}\|\,, \quad \|(\mathbf{B}^*)^{-1}\| \;\leq\; \frac{1}{\beta}\,. \tag{16}$$

Here, as usual, $\tilde{M}_1^{\perp}$ denotes the set of elements in M_1 that are orthogonal to $\tilde{M}_1$, and $\tilde{M}_1^{o}$ stands for the set of functionals in M_1' that vanish on the elements of $\tilde{M}_1$.

Now, let $\boldsymbol{\sigma}_0 := \mathbf{B}^{-1}(\mathbf{G}) \in \tilde{M}_1^{\perp}$. We remark that, due to (16),

$$\|\boldsymbol{\sigma}_0\|_{M_1} \;\leq\; \frac{1}{\beta}\, \|\mathbf{G}\|_{M'}\,. \tag{17}$$

With this, we consider the functionals $\tilde{\mathbf{F}}_1 := \mathbf{F}_1 - \mathbf{B}_1^*(\boldsymbol{\sigma}_0)$ and $\tilde{\mathbf{G}}_1 := \mathbf{G}_1 + \mathbf{S}(\boldsymbol{\sigma}_0)$. Then, by applying Lemma 2.1 we conclude that there exists a unique $(\tilde{\mathbf{t}}, \tilde{\boldsymbol{\sigma}}) \in X_1 \times \tilde{M}_1$ such that

$$\begin{aligned} [\mathbf{A}_1(\tilde{\mathbf{t}}), \mathbf{s}] + [\mathbf{B}_1^*(\tilde{\boldsymbol{\sigma}}), \mathbf{s}] \;&=\; [\tilde{\mathbf{F}}_1, \mathbf{s}]\,, \\ [\mathbf{B}_1(\tilde{\mathbf{t}}), \boldsymbol{\tau}] - [\mathbf{S}(\tilde{\boldsymbol{\sigma}}), \boldsymbol{\tau}] \;&=\; [\tilde{\mathbf{G}}_1, \boldsymbol{\tau}]\,, \end{aligned} \tag{18}$$

for all $(\mathbf{s}, \boldsymbol{\tau}) \in X_1 \times \tilde{M}_1$.

Since $(\tilde{\mathbf{G}}_1 - \mathbf{B}_1(\tilde{\mathbf{t}}) + \mathbf{S}(\tilde{\boldsymbol{\sigma}})) \in \tilde{M}_1^{o}$, we can take $\tilde{\mathbf{u}} := (\mathbf{B}^*)^{-1}\,(\tilde{\mathbf{G}}_1 - \mathbf{B}_1(\tilde{\mathbf{t}}) + \mathbf{S}(\tilde{\boldsymbol{\sigma}})) \in M$. Thus, it is clear that

$$[\mathbf{B}^*(\tilde{\mathbf{u}}), \boldsymbol{\tau}] \;=\; [\tilde{\mathbf{G}}_1 - \mathbf{B}_1(\tilde{\mathbf{t}}) + \mathbf{S}(\tilde{\boldsymbol{\sigma}}), \boldsymbol{\tau}] \quad \forall\, \boldsymbol{\tau} \in M_1\,, \tag{19}$$

and, due to (16),

$$\|\tilde{\mathbf{u}}\|_M \;\leq\; \frac{1}{\beta}\, \|\tilde{\mathbf{G}}_1 - \mathbf{B}_1(\tilde{\mathbf{t}}) + \mathbf{S}(\tilde{\boldsymbol{\sigma}})\|_{M_1'}\,. \tag{20}$$

In this way, noting also that $\mathbf{B}(\tilde{\boldsymbol{\sigma}} + \boldsymbol{\sigma}_0) = \mathbf{B}(\boldsymbol{\sigma}_0) = \mathbf{G}$, we conclude that $(\tilde{\mathbf{t}}, \tilde{\boldsymbol{\sigma}} + \boldsymbol{\sigma}_0, \tilde{\mathbf{u}}) \in \mathbf{X}$ is a solution of (2), which proves the existence.

For the uniqueness, let $(\mathbf{t}, \boldsymbol{\sigma}, \mathbf{u}) \in \mathbf{X}$ be another solution of (2). Since $[\mathbf{B}^*(\mathbf{u}), \boldsymbol{\tau}] = [\mathbf{B}(\boldsymbol{\tau}), \mathbf{u}] = 0$ for all $\boldsymbol{\tau} \in \tilde{M}_1$, it follows from (2) that $(\mathbf{t}, \boldsymbol{\sigma} - \boldsymbol{\sigma}_0) \in X_1 \times \tilde{M}_1$ is also a solution of (18), and hence $(\mathbf{t}, \boldsymbol{\sigma} - \boldsymbol{\sigma}_0) = (\tilde{\mathbf{t}}, \tilde{\boldsymbol{\sigma}})$. This relation and (19) show that $(\mathbf{t}, \boldsymbol{\sigma}, \mathbf{u}) = (\tilde{\mathbf{t}}, \tilde{\boldsymbol{\sigma}} + \boldsymbol{\sigma}_0, \tilde{\mathbf{u}})$.

On the other hand, for the estimate (15) we first observe from (18) and Lemma 2.1 (cf. (4)) that there exists $C_1 > 0$, depending only on γ, α, β_1 and $\|\mathbf{B}_1\|$, such that

$$\begin{aligned} \|\mathbf{t}\|_{X_1} + \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_0\|_{M_1} \,\le\, C_1 \Big\{ & \|\mathbf{F}_1 - \mathbf{B}_1^*(\boldsymbol{\sigma}_0)\|_{X_1'} \\ & + \|\mathbf{G}_1 + \mathbf{S}(\boldsymbol{\sigma}_0)\|_{M_1'} + \|\mathbf{A}_1(0)\|_{X_1'} \Big\}. \end{aligned} \tag{21}$$

Also, from (20) we can write

$$\|\mathbf{u}\|_M \,\le\, \frac{1}{\beta} \|\mathbf{G}_1 - \mathbf{B}_1(\mathbf{t}) + \mathbf{S}(\boldsymbol{\sigma})\|_{M_1'}\,. \tag{22}$$

Therefore, (17), (21) and (22) yield (15), thus finishing the proof. □

Now, if $\mathbf{A}_1$ becomes linear, the above theorem reduces to the following.

THEOREM 2.2: *Let* $\tilde{M}_1 := \operatorname{kernel}(\mathbf{B})$. *Assume that:*

a) *the operator* $\mathbf{A}_1 : X_1 \to X_1'$ *is linear, bounded and* X_1*-elliptic, that is, there exist* γ, $\alpha > 0$ *such that*

$$\|\mathbf{A}_1(\mathbf{s})\|_{X_1'} \,\le\, \gamma\, \|\mathbf{s}\|_{X_1} \quad \textit{and} \quad [\mathbf{A}_1(\mathbf{s}), \mathbf{s}] \,\ge\, \alpha\, \|\mathbf{s}\|_{X_1}\ \forall\, \mathbf{s} \in X_1\,.$$

b) *the conditions* ii), iii) *and* iv) *from Theorem* 2.1 *are satisfied.*

Then, given $\mathbf{F} := (\mathbf{F}_1, \mathbf{G}_1, \mathbf{G}) \in \mathbf{X}' := X_1' \times M_1' \times M'$, *there exists a unique* $(\mathbf{t}, \boldsymbol{\sigma}, \mathbf{u}) \in \mathbf{X}$ *such that*

$$[\mathbf{P}(\mathbf{t}, \boldsymbol{\sigma}, \mathbf{u}), (\mathbf{s}, \boldsymbol{\tau}, \mathbf{v})] \,=\, [\mathbf{F}, (\mathbf{s}, \boldsymbol{\tau}, \mathbf{v})] \quad \forall\, (\mathbf{s}, \boldsymbol{\tau}, \mathbf{v}) \in \mathbf{X}\,.$$

In addition, there exists $\tilde{C} > 0$, *depending only on* γ, α, β_1, $\|\mathbf{B}_1\|$, β *and* $\|\mathbf{S}\|$, *such that*

$$\|(\mathbf{t}, \boldsymbol{\sigma}, \mathbf{u})\|_{\mathbf{X}} \,\le\, \tilde{C}\, \|\mathbf{F}\|_{\mathbf{X}'}\,. \tag{23}$$

Before ending this section, we remark that (23) is equivalent to the following global inf-sup condition:

$$\|(\mathbf{r}, \boldsymbol{\rho}, \mathbf{w})\|_{\mathbf{X}} \,\le\, \tilde{C} \sup_{\substack{(\mathbf{s}, \boldsymbol{\tau}, \mathbf{v}) \in \mathbf{X} \\ (\mathbf{s}, \boldsymbol{\tau}, \mathbf{v}) \neq 0}} \frac{[\mathbf{P}(\mathbf{r}, \boldsymbol{\rho}, \mathbf{w}), (\mathbf{s}, \boldsymbol{\tau}, \mathbf{v})]}{\|(\mathbf{s}, \boldsymbol{\tau}, \mathbf{v})\|_{\mathbf{X}}} \tag{24}$$

for all $(\mathbf{r}, \boldsymbol{\rho}, \mathbf{w}) \in \mathbf{X}$. The discrete analogues of this estimate will be used for the error analysis later on.

3. The discrete two-fold saddle point formulations

We are now interested in the Galerkin approximations of (2). To this end, we let $X_{1,h}$, $M_{1,h}$ and M_h be finite dimensional subspaces of X_1, M_1 and M, respectively. Here, the index h is taken in a numerable family $\mathtt{I} := \{h_j\}_{j\in\mathbb{N}}$ with $h_j \geq h_{j+1}$ for all $j \in \mathbb{N}$.

3.1. The Galerkin scheme

Let $\mathbf{X}_h := X_{1,h} \times M_{1,h} \times M_h$. Then the usual Galerkin scheme associated with (2) reads: *Find* $(\mathbf{t}_h, \boldsymbol{\sigma}_h, \mathbf{u}_h) \in \mathbf{X}_h$ *such that*

$$
\begin{array}{llll}
[\mathbf{A}_1(\mathbf{t}_h), \mathbf{s}_h] & +[\mathbf{B}_1^*(\boldsymbol{\sigma}_h), \mathbf{s}_h] & & = \quad [\mathbf{F}_1, \mathbf{s}_h]\,, \\
[\mathbf{B}_1(\mathbf{t}_h), \boldsymbol{\tau}_h] & -[\mathbf{S}(\boldsymbol{\sigma}_h), \boldsymbol{\tau}_h] & +[\mathbf{B}^*(\mathbf{u}_h), \boldsymbol{\tau}_h] & = \quad [\mathbf{G}_1, \boldsymbol{\tau}_h]\,, \\
 & [\mathbf{B}(\boldsymbol{\sigma}_h), \mathbf{v}_h] & & = \quad [\mathbf{G}, \mathbf{v}_h]\,,
\end{array}
\tag{25}
$$

for all $(\mathbf{s}_h, \boldsymbol{\tau}_h, \mathbf{v}_h) \in \mathbf{X}_h$, or equivalently, such that

$$
[\mathbf{P}(\mathbf{t}_h, \boldsymbol{\sigma}_h, \mathbf{u}_h), (\mathbf{s}_h, \boldsymbol{\tau}_h, \mathbf{v}_h)] \;=\; [\mathbf{F}, (\mathbf{s}_h, \boldsymbol{\tau}_h, \mathbf{v}_h)] \tag{26}
$$

for all $(\mathbf{s}_h, \boldsymbol{\tau}_h, \mathbf{v}_h) \in \mathbf{X}_h$.

For shortness of the presentation, we omit the discrete version of Lemma 2.1, whose statement and proof are straightforward.

Since $\mathbf{X}_h \subseteq \mathbf{X}$ throughout this subsection we let $i_h : X_{1,h} \to X_1$, $j_h : M_{1,h} \to M_1$, $k_h : M_h \to M$ and $\mathbf{i}_h : \mathbf{X}_h \to \mathbf{X}$ be the canonical imbeddings with adjoints $i_h^* : X_1' \to X_{1,h}'$, $j_h^* : M_1' \to M_{1,h}'$, $k_h^* : M' \to M_h'$ and $\mathbf{i}_h^* : \mathbf{X}' \to \mathbf{X}_h'$, respectively.

The discrete analogue of Theorem 2.1 is established next.

THEOREM 3.1: *Let* $\tilde{M}_{1,h} := \{\boldsymbol{\tau}_h \in M_{1,h} : \; [\mathbf{B}(\boldsymbol{\tau}_h), \mathbf{v}_h] = 0 \;\forall\, \mathbf{v}_h \in M_h\}$. *Assume that:*

i) *the operator* $\mathbf{A}_1 : X_1 \to X_1'$ *is nonlinear, Lipschitz continuous (with a Lipschitz constant* $\gamma > 0$*) and strongly monotone (with a monotonicity constant* $\alpha > 0$*).*

ii) *the linear operator* $\mathbf{S}$ *is positive semi-definite on* $\tilde{M}_{1,h}$, *that is*

$$
[\mathbf{S}(\boldsymbol{\tau}_h), \boldsymbol{\tau}_h] \geq 0 \quad \forall\, \boldsymbol{\tau}_h \in \tilde{M}_{1,h}\,.
$$

iii) *the linear operator* $\mathbf{B}_1$ *satisfies a uniform inf-sup condition on* $X_{1,h} \times \tilde{M}_{1,h}$, *that is, there exists* $\beta_1^* > 0$, *independent of* $h \in \mathtt{I}$, *such that*

$$
\sup_{\substack{\mathbf{s}_h \in X_{1,h} \\ \mathbf{s}_h \neq 0}} \frac{[\mathbf{B}_1(\mathbf{s}_h), \boldsymbol{\tau}_h]}{\|\mathbf{s}_h\|_{X_1}} \;\geq\; \beta_1^* \, \|\boldsymbol{\tau}_h\|_{M_1} \quad \forall\, \boldsymbol{\tau}_h \in \tilde{M}_{1,h}\,.
$$

iv) *the linear operator* $\mathbf{B}$ *satisfies a uniform inf-sup condition on* $M_{1,h} \times M_h$, *that is, there exists* $\beta^* > 0$, *independent of* $h \in \mathtt{I}$, *such that*

$$\sup_{\substack{\boldsymbol{\tau}_h \in M_{1,h} \\ \boldsymbol{\tau}_h \neq 0}} \frac{[\mathbf{B}(\boldsymbol{\tau}_h), \mathbf{v}_h]}{\|\boldsymbol{\tau}_h\|_{M_1}} \geq \beta^* \|\mathbf{v}_h\|_M \quad \forall\, \mathbf{v}_h \in M_h\,.$$

Then, there exists a unique $(\mathbf{t}_h, \boldsymbol{\sigma}_h, \mathbf{u}_h) \in \mathbf{X}_h$ *solution of* (25). *In addition, there exists* $\tilde{C} > 0$, *depending only on* γ, α, β_1^*, $\|\mathbf{B}_1\|$, β^* *and* $\|\mathbf{S}\|$, *such that*

$$\|(\mathbf{t}_h, \boldsymbol{\sigma}_h, \mathbf{u}_h)\|_{\mathbf{X}} \leq \tilde{C} \left\{ \|\mathbf{i}_h^* \mathbf{F}\|_{\mathbf{X}_h'} + \|i_h^* \mathbf{A}_1(0)\|_{X_{1,h}'} \right\}. \tag{27}$$

Proof: According to the discrete inf-sup condition for $\mathbf{B}$ (cf. iv)) and Lemma 4.1 in Chapter I of [16], we know that $\mathbf{B}_h := k_h^* \mathbf{B} j_h : \tilde{M}_{1,h}^{\perp} \cap M_{1,h} \to M_h'$ and $\mathbf{B}_h^* := j_h^* \mathbf{B}^* k_h : M_h \to \tilde{M}_{1,h}^{o} \cap M_{1,h}'$ are isomorphisms with

$$\|\mathbf{B}_h^{-1}\|\,,\ \|(\mathbf{B}_h^*)^{-1}\| \leq \frac{1}{\beta^*}\,. \tag{28}$$

Then, we take $\boldsymbol{\sigma}_{0,h} := \mathbf{B}_h^{-1}(k_h^* \mathbf{G}) \in \tilde{M}_{1,h}^{\perp} \cap M_{1,h}$, which satisfies

$$\|\boldsymbol{\sigma}_{0,h}\|_{M_1} \leq \frac{1}{\beta^*} \|k_h^* \mathbf{G}\|_{M_h'}\,,$$

and define the functionals $\tilde{\mathbf{F}}_1 := \mathbf{F}_1 - \mathbf{B}_1^*(\boldsymbol{\sigma}_{0,h})$, $\tilde{\mathbf{G}}_1 := \mathbf{G}_1 + \mathbf{S}(\boldsymbol{\sigma}_{0,h})$. Thus, by applying the discrete version of Lemma 2.1, we deduce that there exists a unique $(\tilde{\mathbf{t}}_h, \tilde{\boldsymbol{\sigma}}_h) \in X_{1,h} \times \tilde{M}_{1,h}$ such that

$$\begin{aligned} [\mathbf{A}_1(\tilde{\mathbf{t}}_h), \mathbf{s}_h] + [\mathbf{B}_1^*(\tilde{\boldsymbol{\sigma}}_h), \mathbf{s}_h] &= [\tilde{\mathbf{F}}_1, \mathbf{s}_h]\,, \\ [\mathbf{B}_1(\tilde{\mathbf{t}}_h), \boldsymbol{\tau}_h] - [\mathbf{S}(\tilde{\boldsymbol{\sigma}}_h), \boldsymbol{\tau}_h] &= [\tilde{\mathbf{G}}_1, \boldsymbol{\tau}_h]\,, \end{aligned}$$

for all $(\mathbf{s}_h, \boldsymbol{\tau}_h) \in X_{1,h} \times \tilde{M}_{1,h}$. The rest of the proof is as in Theorem 2.1. We omit further details. □

The linear version of the above theorem, which corresponds to the discrete analogue of Theorem 2.2, is as follows.

THEOREM 3.2: *Let* $\tilde{M}_{1,h}$ *as in Theorem* 3.1. *Assume that:*

a) *the operator* $\mathbf{A}_1 : X_1 \to X_1'$ *is linear, bounded and* X_1*-elliptic, that is, there exist* γ, $\alpha > 0$ *such that*

$$\|\mathbf{A}_1(\mathbf{s})\|_{X_1'} \leq \gamma \|\mathbf{s}\|_{X_1} \quad \textit{and} \quad [\mathbf{A}_1(\mathbf{s}), \mathbf{s}] \geq \alpha \|\mathbf{s}\|_{X_1} \ \forall\, \mathbf{s} \in X_1\,.$$

b) *the conditions* ii), iii) *and* iv) *from Theorem* 3.1 *are satisfied.*

Then there exists a unique $(\mathbf{t}_h, \boldsymbol{\sigma}_h, \mathbf{u}_h) \in \mathbf{X}_h$ *solution of* (25). *In addition, there exists* $\tilde{C} > 0$, *depending only on* γ, α, β_1^*, $\|\mathbf{B}_1\|$, β^* *and* $\|\mathbf{S}\|$, *such that*

$$\|(\mathbf{t}_h, \boldsymbol{\sigma}_h, \mathbf{u}_h)\|_{\mathbf{X}} \,\leq\, \tilde{C}\,\|\mathbf{i}_h^*\,\mathbf{F}\|_{\mathbf{X}_h'}\,. \tag{29}$$

It is important to remark that the estimate (29) is equivalent to the following global discrete inf-sup condition:

$$\|(\mathbf{r}_h, \boldsymbol{\rho}_h, \mathbf{w}_h)\|_{\mathbf{X}} \,\leq\, \tilde{C} \sup_{\substack{(\mathbf{s}_h, \boldsymbol{\tau}_h, \mathbf{v}_h) \in \mathbf{X}_h \\ (\mathbf{s}_h, \boldsymbol{\tau}_h, \mathbf{v}_h) \neq 0}} \frac{[\mathbf{P}(\mathbf{r}_h, \boldsymbol{\rho}_h, \mathbf{w}_h), (\mathbf{s}_h, \boldsymbol{\tau}_h, \mathbf{v}_h)]}{\|(\mathbf{s}_h, \boldsymbol{\tau}_h, \mathbf{v}_h)\|_{\mathbf{X}}} \tag{30}$$

for all $(\mathbf{r}_h, \boldsymbol{\rho}_h, \mathbf{w}_h) \in \mathbf{X}_h$.

3.1.1. Error analysis

We now turn our attention to the error analysis for the Galerkin scheme (25). In what follows, given any Hilbert spaces U and V, $\mathcal{L}(U, V)$ denotes the space of linear bounded operators from U into V.

We first need to add further assumptions on the nonlinear operator $\mathbf{A}_1$. In fact, we assume that the following conditions hold:

(H.1) *The operator* $\mathbf{A}_1 : X_1 \to X_1'$ *has a hemi-continuous first-order Gâteaux derivative* $D\mathbf{A}_1 : X_1 \to \mathcal{L}(X_1, X_1')$. *Here, given* $\mathbf{s}_1 \in X_1$, *the Gâteaux derivative* $D\mathbf{A}_1(\mathbf{s}_1) \in \mathcal{L}(X_1, X_1')$ *can be interpreted as a mapping acting from* $X_1 \times X_1$ *into* $\mathbb{R}$, *which is defined by:*

$$D\mathbf{A}_1(\mathbf{s}_1)(\mathbf{r}, \mathbf{s}) := \lim_{\epsilon \to 0} \frac{[\mathbf{A}_1(\mathbf{s}_1 + \epsilon\,\mathbf{r}), \mathbf{s}] - [\mathbf{A}_1(\mathbf{s}_1), \mathbf{s}]}{\epsilon} \tag{31}$$

for all $\mathbf{r}$, $\mathbf{s} \in X_1$. *Also, we recall that the hemi-continuity of* $D\mathbf{A}_1$ *means that for any* $\mathbf{r}$, $\mathbf{s} \in X_1$, *the mapping* $\mathbb{R} \ni \mu \to D\mathbf{A}_1(\mathbf{s} + \mu\,\mathbf{r})(\mathbf{r}, \cdot) \in \mathbf{X}_1'$ *is continuous.*

(H.2) *For any* $\mathbf{s}_1 \in X_1$, $D\mathbf{A}_1(\mathbf{s}_1) : X_1 \times X_1 \to \mathbb{R}$ *is a bounded and* X_1*-elliptic bilinear form, with boundedness and ellipticity constants given by* $\tilde{\gamma}(\mathbf{s}_1)$ *and* $\tilde{\alpha}(\mathbf{s}_1)$, *respectively. In addition, there exist* $\tilde{\gamma}$, $\tilde{\alpha} > 0$ *such that*

$$\tilde{\gamma}(\mathbf{s}_1) \,\leq\, \tilde{\gamma} \quad \textit{and} \quad \tilde{\alpha}(\mathbf{s}_1) \,\geq\, \tilde{\alpha} \quad \forall\, \mathbf{s}_1 \in X_1\,. \tag{32}$$

Then, we have the following main result.

THEOREM 3.3: *Let* $\vec{\mathbf{t}} := (\mathbf{t}, \boldsymbol{\sigma}, \mathbf{u}) \in \mathbf{X}$ *and* $\vec{\mathbf{t}}_h := (\mathbf{t}_h, \boldsymbol{\sigma}_h, \mathbf{u}_h) \in \mathbf{X}_h$ *be the unique solutions of* (2) *and* (25), *respectively. Then there exists* $\hat{C} > 0$, *independent of* $h \in \mathtt{I}$, *such that*

$$\|\vec{\mathbf{t}} - \vec{\mathbf{t}}_h\|_{\mathbf{X}} \,\leq\, \hat{C} \inf_{\vec{\mathbf{r}}_h := (\mathbf{r}_h, \boldsymbol{\rho}_h, \mathbf{w}_h) \in \mathbf{X}_h} \|\vec{\mathbf{t}} - \vec{\mathbf{r}}_h\|_{\mathbf{X}}\,. \tag{33}$$

Proof: Given $\vec{\mathbf{r}}_h := (\mathbf{r}_h, \boldsymbol{\rho}_h, \mathbf{w}_h)$, $\vec{\mathbf{s}}_h := (\mathbf{s}_h, \boldsymbol{\tau}_h, \mathbf{v}_h) \in \mathbf{X}_h$, we can write

$$
\begin{aligned}
[\mathbf{P}(\vec{\mathbf{t}}_h), \vec{\mathbf{s}}_h] - [\mathbf{P}(\vec{\mathbf{r}}_h), \vec{\mathbf{s}}_h] &= \int_0^1 \frac{d}{d\mu} \Big\{ [\mathbf{P}(\mu \vec{\mathbf{t}}_h + (1-\mu)\vec{\mathbf{r}}_h), \vec{\mathbf{s}}_h] \Big\} \, d\mu \\
&= \int_0^1 D\mathbf{P}(\mu \vec{\mathbf{t}}_h + (1-\mu)\vec{\mathbf{r}}_h)(\vec{\mathbf{t}}_h - \vec{\mathbf{r}}_h, \vec{\mathbf{s}}_h) \, d\mu \,, \qquad (34)
\end{aligned}
$$

where $D\mathbf{P} : \mathbf{X} \to \mathcal{L}(\mathbf{X}, \mathbf{X}')$ is the first-order Gâteaux derivative of the operator $\mathbf{P} : \mathbf{X} \to \mathbf{X}'$. More precisely, for any $\vec{\mathbf{s}}_1 := (\mathbf{s}_1, \boldsymbol{\tau}_1, \mathbf{v}_1)$, $\vec{\mathbf{r}} := (\mathbf{r}, \boldsymbol{\rho}, \mathbf{w})$, $\vec{\mathbf{s}} := (\mathbf{s}, \boldsymbol{\tau}, \mathbf{v}) \in \mathbf{X}$, $D\mathbf{P}(\vec{\mathbf{s}}_1)(\vec{\mathbf{r}}, \vec{\mathbf{s}})$ is obtained from (13) by replacing $[\mathbf{A}_1(\mathbf{r}), \mathbf{s}]$ by $D\mathbf{A}_1(\mathbf{s}_1)(\mathbf{r}, \mathbf{s})$, that is:

$$
\begin{aligned}
D\mathbf{P}(\vec{\mathbf{s}}_1)(\vec{\mathbf{r}}, \vec{\mathbf{s}}) := {} & D\mathbf{A}_1(\mathbf{s}_1)(\mathbf{r}, \mathbf{s}) + [\mathbf{B}_1^*(\boldsymbol{\rho}), \mathbf{s}] + [\mathbf{B}_1(\mathbf{r}), \boldsymbol{\tau}] \\
& - [\mathbf{S}(\boldsymbol{\rho}), \boldsymbol{\tau}] + [\mathbf{B}^*(\mathbf{w}), \boldsymbol{\tau}] + [\mathbf{B}(\boldsymbol{\rho}), \mathbf{v}] \,.
\end{aligned} \qquad (35)
$$

Thus, for any $\vec{\mathbf{s}}_1 \in \mathbf{X}$, (35) induces the definition of an operator in $\mathcal{L}(\mathbf{X}, \mathbf{X}')$, which, according to **(H.2)**, satisfies the hypotheses of Theorem 3.2 with constants independent of h and of $\vec{\mathbf{s}}_1$. Consequently, there exists $\tilde{C} > 0$, depending only on $\tilde{\gamma}$, $\tilde{\alpha}$, β_1^*, $\|\mathbf{B}_1\|$, β^* and $\|\mathbf{S}\|$, such that (30) holds, and hence, in particular

$$
\|\vec{\mathbf{t}}_h - \vec{\mathbf{r}}_h\|_{\mathbf{X}} \,\le\, \tilde{C} \sup_{\substack{\vec{\mathbf{s}}_h \in \mathbf{X}_h \\ \vec{\mathbf{s}}_h \neq 0}} \frac{D\mathbf{P}(\vec{\mathbf{s}}_1)(\vec{\mathbf{t}}_h - \vec{\mathbf{r}}_h, \vec{\mathbf{s}}_h)}{\|\vec{\mathbf{s}}_h\|_X} \qquad (36)
$$

for all $\vec{\mathbf{s}}_1 \in \mathbf{X}$.

On the other hand, the hemi-continuity of $D\mathbf{A}_1$ (cf. **(H.1)**) implies the same property for $D\mathbf{P}$ and hence, there exists $\mu_0 \in (0,1)$ such that (34) becomes

$$
\begin{aligned}
& [\mathbf{P}(\vec{\mathbf{t}}_h), \vec{\mathbf{s}}_h] - [\mathbf{P}(\vec{\mathbf{r}}_h), \vec{\mathbf{s}}_h] \\
& = D\mathbf{P}(\mu_0 \vec{\mathbf{t}}_h + (1-\mu_0)\vec{\mathbf{r}}_h)(\vec{\mathbf{t}}_h - \vec{\mathbf{r}}_h, \vec{\mathbf{s}}_h) \,.
\end{aligned} \qquad (37)
$$

It follows from (36) (with $\vec{\mathbf{s}}_1 := \mu_0 \vec{\mathbf{t}}_h + (1-\mu_0)\vec{\mathbf{r}}_h$) and (37) that

$$
\|\vec{\mathbf{t}}_h - \vec{\mathbf{r}}_h\|_{\mathbf{X}} \,\le\, \tilde{C} \sup_{\substack{\vec{\mathbf{s}}_h \in \mathbf{X}_h \\ \vec{\mathbf{s}}_h \neq 0}} \frac{[\mathbf{P}(\vec{\mathbf{t}}_h), \vec{\mathbf{s}}_h] - [\mathbf{P}(\vec{\mathbf{r}}_h), \vec{\mathbf{s}}_h]}{\|\vec{\mathbf{s}}_h\|_X} \,. \qquad (38)
$$

Next, since $[\mathbf{P}(\vec{\mathbf{t}}_h), \vec{\mathbf{s}}_h] = [\mathbf{P}(\vec{\mathbf{t}}), \vec{\mathbf{s}}_h]$ for all $\vec{\mathbf{s}}_h \in \mathbf{X}_h$, and since $\mathbf{P}$ is Lipschitz continuous, with a constant $\bar{\gamma}$ depending only on γ, $\|\mathbf{B}_1\|$, $\|\mathbf{S}\|$ and $\|\mathbf{B}\|$, we conclude from (38) that

$$
\|\vec{\mathbf{t}}_h - \vec{\mathbf{r}}_h\|_{\mathbf{X}} \,\le\, \bar{\gamma}\, \tilde{C} \, \|\vec{\mathbf{t}} - \vec{\mathbf{r}}_h\|_{\mathbf{X}} \quad \forall\, \vec{\mathbf{r}}_h \in \mathbf{X}_h \,. \qquad (39)
$$

Finally, triangle's inequality and (39) yield the Cea estimate (33). $\square$

3.2. The fully discrete Galerkin scheme

In order to allow the utilisation and corresponding numerical analysis of fully discrete Galerkin schemes for (2), we first let $X_{1,h}$, $M_{1,h}$ and M_h be finite dimensional spaces approximating, in some sense, the spaces X_1, M_1 and M, respectively. This means, in other words, that $\mathbf{X}_h := X_{1,h} \times M_{1,h} \times M_h$ is not necessarily assumed to be a subspace of $\mathbf{X} := X_1 \times M_1 \times M$.

Then, we introduce discrete operators $\mathbf{A}_{1,h} : X_{1,h} \to X'_{1,h}$, $\mathbf{B}_{1,h} : X_{1,h} \to M'_{1,h}$, $\mathbf{B}_h : M_{1,h} \to M'_h$, $\mathbf{S}_h : M_{1,h} \to M'_{1,h}$, and the corresponding adjoints $\mathbf{B}^*_{1,h} : M_{1,h} \to X'_{1,h}$ and $\mathbf{B}^*_h : M_h \to M'_{1,h}$. In addition, we consider discrete functionals $\mathbf{F}_{1,h} \in X'_{1,h}$, $\mathbf{G}_{1,h} \in M'_{1,h}$ and $\mathbf{G}_h \in M'_h$.

Thus, our fully discrete Galerkin scheme associated with (2) reads: *Find* $(\mathbf{t}_h, \boldsymbol{\sigma}_h, \mathbf{u}_h) \in \mathbf{X}_h$ *such that*

$$\begin{array}{llll} [\mathbf{A}_{1,h}(\mathbf{t}_h), \mathbf{s}_h] & +[\mathbf{B}^*_{1,h}(\boldsymbol{\sigma}_h), \mathbf{s}_h] & & = \; [\mathbf{F}_{1,h}, \mathbf{s}_h]\,, \\[2ex] [\mathbf{B}_{1,h}(\mathbf{t}_h), \boldsymbol{\tau}_h] & -[\mathbf{S}_h(\boldsymbol{\sigma}_h), \boldsymbol{\tau}_h] & +[\mathbf{B}^*_h(\mathbf{u}_h), \boldsymbol{\tau}_h] & = \; [\mathbf{G}_{1,h}, \boldsymbol{\tau}_h]\,, \\[2ex] & [\mathbf{B}_h(\boldsymbol{\sigma}_h), \mathbf{v}_h] & & = \; [\mathbf{G}_h, \mathbf{v}_h]\,, \end{array} \tag{40}$$

for all $(\mathbf{s}_h, \boldsymbol{\tau}_h, \mathbf{v}_h) \in \mathbf{X}_h$. Equivalently, we can write

$$[\mathbf{P}_h(\mathbf{t}_h, \boldsymbol{\sigma}_h, \mathbf{u}_h), (\mathbf{s}_h, \boldsymbol{\tau}_h, \mathbf{v}_h)] \;=\; [\mathbf{F}_h, (\mathbf{s}_h, \boldsymbol{\tau}_h, \mathbf{v}_h)] \tag{41}$$

for all $(\mathbf{s}_h, \boldsymbol{\tau}_h, \mathbf{v}_h) \in \mathbf{X}_h$, where $\mathbf{F}_h := (\mathbf{F}_{1,h}, \mathbf{G}_{1,h}, \mathbf{G}_h) \in \mathbf{X}'_h := X'_{1,h} \times M'_{1,h} \times M'_h$, and $\mathbf{P}_h : \mathbf{X}_h \to \mathbf{X}'_h$ is given by

$$\begin{array}{c} [\mathbf{P}_h(\mathbf{r}_h, \boldsymbol{\rho}_h, \mathbf{w}_h), (\mathbf{s}_h, \boldsymbol{\tau}_h, \mathbf{v}_h)] = [\mathbf{A}_{1,h}(\mathbf{r}_h), \mathbf{s}_h] + [\mathbf{B}^*_{1,h}(\boldsymbol{\rho}_h), \mathbf{s}_h] \\[2ex] +[\mathbf{B}_{1,h}(\mathbf{r}_h), \boldsymbol{\tau}_h] - [\mathbf{S}_h(\boldsymbol{\rho}_h), \boldsymbol{\tau}_h] + [\mathbf{B}^*_h(\mathbf{w}_h), \boldsymbol{\tau}_h] + [\mathbf{B}_h(\boldsymbol{\rho}_h), \mathbf{v}_h]\,, \end{array} \tag{42}$$

for all $(\mathbf{r}_h, \boldsymbol{\rho}_h, \mathbf{w}_h)$, $(\mathbf{s}_h, \boldsymbol{\tau}_h, \mathbf{v}_h) \in \mathbf{X}_h$. At this point we observe, as for the continuous case, that

$$[\mathbf{B}^*_{1,h}(\boldsymbol{\rho}_h), \mathbf{s}_h] = [\mathbf{B}_{1,h}(\mathbf{s}_h), \boldsymbol{\rho}_h] \quad \text{and} \quad [\mathbf{B}^*_h(\mathbf{w}_h), \boldsymbol{\tau}_h] = [\mathbf{B}_h(\boldsymbol{\tau}_h), \mathbf{w}_h]\,.$$

Then, the analogue of Theorem 3.1 is stated as follows.

THEOREM 3.4: *Let* $\bar{M}_{1,h} := \{\boldsymbol{\tau}_h \in M_{1,h} : \; [\mathbf{B}_h(\boldsymbol{\tau}_h), \mathbf{v}_h] = 0 \;\forall\, \mathbf{v}_h \in M_h\}$. *Assume that:*

i) *the operator* $\mathbf{A}_{1,h} : X_{1,h} \to X'_{1,h}$ *is nonlinear, Lipschitz continuous (with a Lipschitz constant* $\gamma > 0$*, independent of* $h \in \mathtt{I}$*) and strongly monotone (with a monotonicity constant* $\alpha > 0$*, independent of* $h \in \mathtt{I}$*).*

ii) *the linear operator* $\mathbf{S}_h$ *is positive semi-definite on* $\bar{M}_{1,h}$*, that is*

$$[\mathbf{S}_h(\boldsymbol{\tau}_h), \boldsymbol{\tau}_h] \geq 0 \quad \forall\, \boldsymbol{\tau}_h \in \bar{M}_{1,h}\,.$$

iii) *the linear operator* $\mathbf{B}_{1,h}$ *satisfies a uniform inf-sup condition on* $X_{1,h} \times \bar{M}_{1,h}$, *that is, there exists* $\beta_1^* > 0$, *independent of* $h \in \mathtt{I}$, *such that*

$$\sup_{\substack{\mathbf{s}_h \in X_{1,h} \\ \mathbf{s}_h \neq 0}} \frac{[\mathbf{B}_{1,h}(\mathbf{s}_h), \boldsymbol{\tau}_h]}{\|\mathbf{s}_h\|_{X_{1,h}}} \geq \beta_1^* \|\boldsymbol{\tau}_h\|_{M_{1,h}} \quad \forall\, \boldsymbol{\tau}_h \in \bar{M}_{1,h}\,.$$

iv) *the linear operator* $\mathbf{B}_h$ *satisfies a uniform inf-sup condition on* $M_{1,h} \times M_h$, *that is, there exists* $\beta^* > 0$, *independent of* $h \in \mathtt{I}$, *such that*

$$\sup_{\substack{\boldsymbol{\tau}_h \in M_{1,h} \\ \boldsymbol{\tau}_h \neq 0}} \frac{[\mathbf{B}_h(\boldsymbol{\tau}_h), \mathbf{v}_h]}{\|\boldsymbol{\tau}_h\|_{M_{1,h}}} \geq \beta^* \|\mathbf{v}_h\|_{M_h} \quad \forall\, \mathbf{v}_h \in M_h\,.$$

Then, there exists a unique $(\mathbf{t}_h, \boldsymbol{\sigma}_h, \mathbf{u}_h) \in \mathbf{X}_h$ *solution of* (40). *In addition, there exists* $\tilde{C} > 0$, *depending only on* γ, α, β_1^*, $\|\mathbf{B}_{1,h}\|$, β^* *and* $\|\mathbf{S}_h\|$, *such that*

$$\|(\mathbf{t}_h, \boldsymbol{\sigma}_h, \mathbf{u}_h)\|_{\mathbf{X}_h} \leq \tilde{C} \left\{ \|\mathbf{F}_h\|_{\mathbf{X}_h'} + \|\mathbf{A}_{1,h}(0)\|_{X_{1,h}'} \right\}. \tag{43}$$

Proof: It is similar to the proof of Theorem 3.1, and hence we omit details. □

The linear version of the previous theorem is as follows.

THEOREM 3.5: *Let* $\bar{M}_{1,h} := \{\boldsymbol{\tau}_h \in M_{1,h} : [\mathbf{B}_h(\boldsymbol{\tau}_h), \mathbf{v}_h] = 0 \;\forall\, \mathbf{v}_h \in M_h\}$. *Assume that:*

a) *the operator* $\mathbf{A}_{1,h} : X_{1,h} \to X_{1,h}'$ *is linear, and uniformly bounded and* $X_{1,h}$-*elliptic, that is, there exist* γ, $\alpha > 0$, *independent of* $h \in \mathtt{I}$, *such that*

$$\|\mathbf{A}_{1,h}(\mathbf{s}_h)\|_{X_{1,h}'} \leq \gamma \|\mathbf{s}_h\|_{X_{1,h}} \quad \text{and} \quad [\mathbf{A}_{1,h}(\mathbf{s}_h), \mathbf{s}_h] \geq \alpha \|\mathbf{s}_h\|_{X_{1,h}}$$

for all $\mathbf{s}_h \in X_{1,h}$.

b) *the conditions* ii), iii) *and* iv) *from Theorem* 3.4 *are satisfied.*

Then, there exists a unique $(\mathbf{t}_h, \boldsymbol{\sigma}_h, \mathbf{u}_h) \in \mathbf{X}_h$ *solution of* (40). *In addition, there exists* $\tilde{C} > 0$, *depending only on* γ, α, β_1^*, $\|\mathbf{B}_{1,h}\|$, β^* *and* $\|\mathbf{S}_h\|$, *such that*

$$\|(\mathbf{t}_h, \boldsymbol{\sigma}_h, \mathbf{u}_h)\|_{\mathbf{X}_h} \leq \tilde{C}\, \|\mathbf{F}_h\|_{\mathbf{X}_h'}\,. \tag{44}$$

It is important to remark that the estimate (44) is equivalent to the following global discrete inf-sup condition:

$$\|(\mathbf{r}_h, \boldsymbol{\rho}_h, \mathbf{w}_h)\|_{\mathbf{X}_h} \leq \tilde{C} \sup_{\substack{(\mathbf{s}_h, \boldsymbol{\tau}_h, \mathbf{v}_h) \in \mathbf{X}_h \\ (\mathbf{s}_h, \boldsymbol{\tau}_h, \mathbf{v}_h) \neq 0}} \frac{[\mathbf{P}_h(\mathbf{r}_h, \boldsymbol{\rho}_h, \mathbf{w}_h), (\mathbf{s}_h, \boldsymbol{\tau}_h, \mathbf{v}_h)]}{\|(\mathbf{s}_h, \boldsymbol{\tau}_h, \mathbf{v}_h)\|_{\mathbf{X}_h}} \tag{45}$$

for all $(\mathbf{r}_h, \boldsymbol{\rho}_h, \mathbf{w}_h) \in \mathbf{X}_h$.

3.2.1. Error analysis

Similarly as in Section 3.1.1, we now add further assumptions on the nonlinear operator $\mathbf{A}_{1,h}$. In fact, we assume that the following conditions hold:

(H.3) *The operator* $\mathbf{A}_{1,h} : X_{1,h} \to X'_{1,h}$ *has a hemi-continuous first-order Gâteaux derivative* $D\mathbf{A}_{1,h} : X_{1,h} \to \mathcal{L}(X_{1,h}, X'_{1,h})$.

(H.4) *For any* $\mathbf{s}_{1,h} \in X_{1,h}$, $D\mathbf{A}_{1,h}(\mathbf{s}_{1,h}) : X_{1,h} \times X_{1,h} \to \mathbb{R}$ *is a bounded and* $X_{1,h}$*-elliptic bilinear form, with boundedness and ellipticity constants given by* $\tilde{\gamma}(\mathbf{s}_{1,h})$ *and* $\tilde{\alpha}(\mathbf{s}_{1,h})$*, respectively. In addition, there exist* $\tilde{\gamma}$, $\tilde{\alpha} > 0$*, independent of* $h \in \mathtt{I}$*, such that*

$$\tilde{\gamma}(\mathbf{s}_{1,h}) \le \tilde{\gamma} \quad \text{and} \quad \tilde{\alpha}(\mathbf{s}_{1,h}) \ge \tilde{\alpha} \quad \forall\, \mathbf{s}_{1,h} \in X_{1,h}\,. \tag{46}$$

Next, since the spaces $\mathbf{X}_h$ are not necessarily included in $\mathbf{X}$, we also assume that:

(H.5) *For each* $h \in \mathtt{I}$ *there exist a nonlinear operator* $\mathcal{P}_h : \mathbf{X}_h \cup \mathbf{X} \to \mathbf{X}'_h$ *and a norm* $\|\cdot\|_h : \mathbf{X}_h \cup \mathbf{X} \to \mathbb{R}^+$ *such that*

$$\begin{aligned} \|\vec{\mathbf{r}}_h\|_h &= \|\vec{\mathbf{r}}_h\|_{\mathbf{X}_h} \quad \forall\, \vec{\mathbf{r}}_h \in \mathbf{X}_h\,, \quad \text{and} \\ \|\mathcal{P}_h(\vec{\mathbf{r}}) - \mathcal{P}_h(\vec{\mathbf{s}})\|_{\mathbf{X}'_h} &\le \bar{\gamma}\, \|\vec{\mathbf{r}} - \vec{\mathbf{s}}\|_h \quad \forall\, \vec{\mathbf{r}},\, \vec{\mathbf{s}} \in \mathbf{X}_h \cup \mathbf{X}\,, \end{aligned} \tag{47}$$

where $\bar{\gamma} > 0$ *is a constant, independent of* $h \in \mathtt{I}$.

Then, we have the following main result.

THEOREM 3.6: *Let* $\vec{\mathbf{t}} := (\mathbf{t}, \boldsymbol{\sigma}, \mathbf{u}) \in \mathbf{X}$ *and* $\vec{\mathbf{t}}_h := (\mathbf{t}_h, \boldsymbol{\sigma}_h, \mathbf{u}_h) \in \mathbf{X}_h$ *be the unique solutions of* (2) *and* (40)*, respectively. Then there exists* $\hat{C} > 0$*, independent of* $h \in \mathtt{I}$*, such that the following Strang-type error estimate holds:*

$$\begin{aligned} \|\vec{\mathbf{t}} - \vec{\mathbf{t}}_h\|_h \le \hat{C} \Big\{ &\|\mathbf{F}_h - \mathcal{P}_h(\vec{\mathbf{t}})\|_{\mathbf{X}'_h} \\ &+ \inf_{\vec{\mathbf{r}}_h \in \mathbf{X}_h} \Big\{ \|\vec{\mathbf{t}} - \vec{\mathbf{r}}_h\|_h + \|\mathcal{P}_h(\vec{\mathbf{r}}_h) - \mathbf{P}_h(\vec{\mathbf{r}}_h)\|_{\mathbf{X}'_h} \Big\} \Big\}\,. \end{aligned} \tag{48}$$

Proof: We have first by triangle's inequality that

$$\|\vec{\mathbf{t}} - \vec{\mathbf{t}}_h\|_h \le \|\vec{\mathbf{t}} - \vec{\mathbf{r}}_h\|_h + \|\vec{\mathbf{t}}_h - \vec{\mathbf{r}}_h\|_h \quad \forall\, \vec{\mathbf{r}}_h \in \mathbf{X}_h\,. \tag{49}$$

Next, proceeding analogously as in the proof of Theorem 3.3, using now **(H.3)**, **(H.4)** and Theorem 3.5, one can prove that there exists $\tilde{C} > 0$, depending only on $\tilde{\gamma}$, $\tilde{\alpha}$, β_1^*, $\|\mathbf{B}_{1,h}\|$, β^* and $\|\mathbf{S}_h\|$, such that

$$\|\vec{\mathbf{t}}_h - \vec{\mathbf{r}}_h\|_h \le \tilde{C} \sup_{\substack{\vec{\mathbf{s}}_h \in \mathbf{X}_h \\ \vec{\mathbf{s}}_h \neq 0}} \frac{[\mathbf{P}_h(\vec{\mathbf{t}}_h), \vec{\mathbf{s}}_h] - [\mathbf{P}_h(\vec{\mathbf{r}}_h), \vec{\mathbf{s}}_h]}{\|\vec{\mathbf{s}}_h\|_h} \tag{50}$$

for all $\vec{\mathbf{r}}_h \in \mathbf{X}_h$.

Then, since $[\mathbf{P}_h(\vec{\mathbf{t}}_h), \vec{\mathbf{s}}_h] = [\mathbf{F}_h, \vec{\mathbf{s}}_h]$ for all $\vec{\mathbf{s}}_h \in \mathbf{X}_h$, and after adding and subtracting appropriate terms, we can write

$$[\mathbf{P}_h(\vec{\mathbf{t}}_h), \vec{\mathbf{s}}_h] - [\mathbf{P}_h(\vec{\mathbf{r}}_h), \vec{\mathbf{s}}_h] = [\mathbf{F}_h, \vec{\mathbf{s}}_h] - [\mathcal{P}_h(\vec{\mathbf{t}}), \vec{\mathbf{s}}_h]$$

$$+ [\mathcal{P}_h(\vec{\mathbf{t}}), \vec{\mathbf{s}}_h] - [\mathcal{P}_h(\vec{\mathbf{r}}_h), \vec{\mathbf{s}}_h] \, + \, [\mathcal{P}_h(\vec{\mathbf{r}}_h), \vec{\mathbf{s}}_h] - [\mathbf{P}_h(\vec{\mathbf{r}}_h), \vec{\mathbf{s}}_h] \,,$$

which, according to the Lipschitz continuity of $\mathcal{P}_h$ (cf. **(H.5)**), leads to

$$\begin{aligned} |[\mathbf{P}_h(\vec{\mathbf{t}}_h), \vec{\mathbf{s}}_h] - [\mathbf{P}_h(\vec{\mathbf{r}}_h), \vec{\mathbf{s}}_h]| \, &\leq \, |[\mathbf{F}_h, \vec{\mathbf{s}}_h] - [\mathcal{P}_h(\vec{\mathbf{t}}), \vec{\mathbf{s}}_h]| \\ &+ \bar{\gamma} \, \|\vec{\mathbf{t}} - \vec{\mathbf{r}}_h\|_h \, \|\vec{\mathbf{s}}_h\|_h + |[\mathcal{P}_h(\vec{\mathbf{r}}_h), \vec{\mathbf{s}}_h] - [\mathbf{P}_h(\vec{\mathbf{r}}_h), \vec{\mathbf{s}}_h]| \end{aligned} \tag{51}$$

for all $\vec{\mathbf{r}}_h \in \mathbf{X}_h$.

In this way, (49), (50) and (51) yield (48), which finishes the proof of the theorem. □

As a particular case of Theorem 3.6 we have the following result.

Theorem 3.7: *Assume that* $\mathbf{X}_h \subseteq \mathbf{X}$ *and let* $\vec{\mathbf{t}} := (\mathbf{t}, \boldsymbol{\sigma}, \mathbf{u}) \in \mathbf{X}$ *and* $\vec{\mathbf{t}}_h := (\mathbf{t}_h, \boldsymbol{\sigma}_h, \mathbf{u}_h) \in \mathbf{X}_h$ *be the unique solutions of* (2) *and* (40)*, respectively. In addition, let* $\mathbf{i}_h^* : \mathbf{X}' \to \mathbf{X}_h'$ *be the adjoint of the canonical imbedding* $\mathbf{i}_h : \mathbf{X}_h \to \mathbf{X}$*. Then there exists* $\hat{C} > 0$*, independent of* $h \in \mathtt{I}$*, such that the following Strang-type error estimate holds:*

$$\begin{aligned} \|\vec{\mathbf{t}} - \vec{\mathbf{t}}_h\|_X \, &\leq \, \hat{C} \Big\{ \, \|\mathbf{F}_h - \mathbf{i}_h^* \, \mathbf{F}\|_{\mathbf{X}_h'} \\ &+ \inf_{\vec{\mathbf{r}}_h \in \mathbf{X}_h} \Big\{ \|\vec{\mathbf{t}} - \vec{\mathbf{r}}_h\|_X + \|\mathbf{i}_h^* \, \mathbf{P}(\vec{\mathbf{r}}_h) - \mathbf{P}_h(\vec{\mathbf{r}}_h)\|_{\mathbf{X}_h'} \Big\} \Big\} . \end{aligned} \tag{52}$$

Proof: In order to guarantee **(H.5)**, it suffices to take $\| \cdot \|_h := \| \cdot \|_{\mathbf{X}}$ and $\mathcal{P}_h := \mathbf{i}_h^* \, \mathbf{P}$. It follows that $\mathcal{P}_h(\vec{\mathbf{t}}) = \mathbf{i}_h^* \, \mathbf{P}(\vec{\mathbf{t}}) = \mathbf{i}_h^* \, \mathbf{F}$, and hence (48) yields (52). □

We end this section by remarking that, if $\mathbf{F}_h = \mathbf{i}_h^* \, \mathbf{F}$ and $\mathbf{P}_h = \mathbf{i}_h^* \, \mathbf{P}$, then one recovers from (52) the usual Cea estimate (33) of Theorem 3.3.

4. An application to a nonlinear elliptic problem

Let Ω be a simply connected and bounded domain in $\mathbb{R}^2$ with Lipschitz continuous boundary $\Gamma := \partial\Omega$. Then, given $f \in L^2(\Omega)$, $g \in H^{1/2}(\Gamma)$ and a function

$\boldsymbol{\kappa} : \Omega \times \mathbb{R}^+ \to \mathbb{R}$, we consider the nonlinear boundary value problem with non-homogeneous Dirichlet data: *Find* $u \in H^1(\Omega)$ *such that*

$$\begin{aligned} -\operatorname{div}\left(\boldsymbol{\kappa}(\cdot, |\nabla u|)\, \nabla u\right) \quad &= f \quad \text{in} \quad \Omega\,, \\ u \quad &= g \quad \text{on} \quad \Gamma\,, \end{aligned} \tag{53}$$

where $|\cdot|$ stands for the Euclidean norm in $\mathbb{R}^2$ and div denotes the divergence operator.

Here, we assume that $\boldsymbol{\kappa} \in C^1(\Omega \times \mathbb{R}^+)$ and that there exist constants κ_0, $\kappa_1 > 0$, such that for all $(x, s) \in \Omega \times \mathbb{R}^+$:

$$\begin{gathered} \kappa_0 \,\le\, \boldsymbol{\kappa}(x, s) \,\le\, \kappa_1\,, \\ \kappa_0 \,\le\, \boldsymbol{\kappa}(x, s) + s\, \tfrac{\partial}{\partial s}\, \boldsymbol{\kappa}(x, s) \,\le\, \kappa_1\,, \quad \text{and} \\ |\nabla_x \boldsymbol{\kappa}(x, s)| \,\le\, \kappa_1\,. \end{gathered} \tag{54}$$

This kind of nonlinear problem appears in the computation of the magnetic field of electromagnetic devices and also in steady heat conduction. For instance, for the stationary model of Maxwell's equations, and under certain symmetry conditions, a simple magnetic field problem can be reduced to a problem of the type given by (53) in which Ω becomes the cross-section of a domain in $\mathbb{R}^3$ filled up with a ferromagnetic material, such as iron. Thus, the behaviour of this metal yields the nonlinearity given by $\boldsymbol{\kappa}$. For recent results concerning the finite element analysis of nonlinear magnetic field problems, we refer to [17], [18].

4.1. The continuous two-fold saddle point formulation

We are interested in using mixed finite element methods for solving (53). Hence, introducing the further unknowns $\mathbf{t} := \nabla u$ and $\boldsymbol{\sigma} := \boldsymbol{\kappa}(\cdot, |\nabla u|)\, \nabla u$ in Ω, and proceeding as in [9], we arrive at the following dual-dual variational formulation of (53): *Find* $(\mathbf{t}, \boldsymbol{\sigma}, u) \in [L^2(\Omega)]^2 \times H(\operatorname{div}; \Omega) \times L^2(\Omega)$ *such that*

$$\begin{array}{lllcl} \displaystyle\int_\Omega \boldsymbol{\kappa}(\cdot, |\mathbf{t}|)\, \mathbf{t} \cdot \mathbf{s}\, dx & \displaystyle - \int_\Omega \boldsymbol{\sigma} \cdot \mathbf{s}\, dx & & = & 0\,, \\ \displaystyle - \int_\Omega \mathbf{t} \cdot \boldsymbol{\tau}\, dx & & \displaystyle - \int_\Omega u \operatorname{div} \boldsymbol{\tau}\, dx & = & -\langle g, \boldsymbol{\tau} \cdot \boldsymbol{\nu} \rangle\,, \\ & \displaystyle - \int_\Omega v \operatorname{div} \boldsymbol{\sigma}\, dx & & = & \displaystyle\int_\Omega f v\, dx\,, \end{array} \tag{55}$$

for all $(\mathbf{s}, \boldsymbol{\tau}, v) \in [L^2(\Omega)]^2 \times H(\operatorname{div}; \Omega) \times L^2(\Omega)$, where $\langle \cdot, \cdot \rangle$ stands for the duality pairing of $H^{1/2}(\Gamma)$ and $H^{-1/2}(\Gamma)$ with respect to the $L^2(\Gamma)$-inner product, and $\boldsymbol{\nu}$ denotes the unit outward normal to Γ.

We also recall here that $H(\operatorname{div}; \Omega)$ is the space of functions $\boldsymbol{\tau} \in [L^2(\Omega)]^2$ such that $\operatorname{div} \boldsymbol{\tau} \in L^2(\Omega)$. It is well known that $H(\operatorname{div}; \Omega)$ is a Hilbert space and that for all $\boldsymbol{\tau} \in H(\operatorname{div}; \Omega)$, $\boldsymbol{\tau} \cdot \boldsymbol{\nu} \in H^{-1/2}(\Gamma)$ with $\|\boldsymbol{\tau} \cdot \boldsymbol{\nu}\|_{H^{-1/2}(\Gamma)} \le \|\boldsymbol{\tau}\|_{H(\operatorname{div};\Omega)}$ (see [16] for a proof of these results).

Now, it is easy to see that (55) has the two-fold saddle point structure (2) with the operator $\mathbf{S} = \mathbf{O}$. In fact, let $X_1 := [L^2(\Omega)]^2$, $M_1 := H(\mathrm{div};\Omega)$, $M := L^2(\Omega)$, $\mathbf{X} := X_1 \times M_1 \times M$, and define the nonlinear operator $\mathbf{A}_1 : X_1 \to X_1'$, the bounded linear operators $\mathbf{B}_1 : X_1 \to M_1'$, $\mathbf{B} : M_1 \to M'$, and the functionals $\mathbf{F}_1 \in X_1'$, $\mathbf{G}_1 \in M_1'$, $\mathbf{G} \in M'$, as follows

$$[\mathbf{A}_1(\mathbf{r}), \mathbf{s}] := \int_\Omega \boldsymbol{\kappa}(\cdot, |\mathbf{r}|)\, \mathbf{r} \cdot \mathbf{s}\, dx\,,$$

$$[\mathbf{B}_1(\mathbf{r}), \boldsymbol{\tau}] := -\int_\Omega \mathbf{r} \cdot \boldsymbol{\tau}\, dx\,,$$

$$[\mathbf{B}(\boldsymbol{\rho}), v] := -\int_\Omega v\, \mathrm{div}\, \boldsymbol{\rho}\, dx\,,$$

$$\mathbf{F}_1 := \mathbf{O}, \quad [\mathbf{G}_1, \boldsymbol{\tau}] := -\langle g, \boldsymbol{\tau} \cdot \boldsymbol{\nu} \rangle \quad \text{and} \quad [\mathbf{G}, v] := \int_\Omega f v\, dx$$

for all $(\mathbf{r}, \boldsymbol{\rho})$, $(\mathbf{s}, \boldsymbol{\tau}) \in X_1 \times M_1$ and for all $v \in M$, where, as before, $[\cdot,\cdot]$ stands for the duality pairing induced by the corresponding operators. We remark here that the first condition in (54) ensures that $\mathbf{A}_1(\mathbf{r}) \in X_1'$ for all $\mathbf{r} \in X_1$, which confirms that $\mathbf{A}_1$ is well defined.

Therefore, the variational formulation (55) can be rewritten as: *Find* $(\mathbf{t}, \boldsymbol{\sigma}, u) \in \mathbf{X}$ *such that*

$$\begin{array}{lllll}
[\mathbf{A}_1(\mathbf{t}), \mathbf{s}] & + [\mathbf{B}_1^*(\boldsymbol{\sigma}), \mathbf{s}] & & = & [\mathbf{F}_1, \mathbf{s}]\,, \\[2ex]
[\mathbf{B}_1(\mathbf{t}), \boldsymbol{\tau}] & + & [\mathbf{B}^*(u), \boldsymbol{\tau}] & = & [\mathbf{G}_1, \boldsymbol{\tau}]\,, \\[2ex]
 & [\mathbf{B}(\boldsymbol{\sigma}), v] & & = & [\mathbf{G}, v]\,,
\end{array} \tag{56}$$

for all $(\mathbf{s}, \boldsymbol{\tau}, v) \in \mathbf{X}$.

Our main result concerning the solvability of (56) is provided now.

THEOREM 4.1: *There exists a unique* $(\mathbf{t}, \boldsymbol{\sigma}, u) \in \mathbf{X}$ *solution of the nonlinear two-fold saddle point formulation* (56). *Moreover, there exists* $C > 0$, *independent of the solution, such that*

$$\|(\mathbf{t}, \boldsymbol{\sigma}, u)\|_{\mathbf{X}} \le C \left\{ \|f\|_{L^2(\Omega)} + \|g\|_{H^{1/2}(\Gamma)} \right\}.$$

Proof: The assumption (54) allows to prove that the nonlinear operator $\mathbf{A}_1$ is strongly monotone and Lipschitz continuous (for details, see Theorem 3.8 in [14]). Also, it is well known that the operator $\mathbf{B}$ satisfies the continuous inf-sup condition on $M_1 \times M$ (see, e.g., [16], [14] or [19]).

On the other hand, it is straightforward to see that

$$\tilde{M}_1 := \mathrm{kernel}(\mathbf{B}) := \{ \boldsymbol{\tau} \in M_1 : \quad \mathrm{div}\, \boldsymbol{\tau} = 0 \quad \text{in} \quad \Omega \}.$$

Thus, since $\tilde{M}_1 \subseteq M_1 \subseteq X_1$, we deduce that

$$\sup_{\substack{\mathbf{s}\in X_1\\ \mathbf{s}\neq 0}} \frac{[\mathbf{B}_1(\mathbf{s}),\boldsymbol{\tau}]}{\|\mathbf{s}\|_{X_1}} = \sup_{\substack{\mathbf{s}\in X_1\\ \mathbf{s}\neq 0}} \frac{-\int_\Omega \mathbf{s}\cdot\boldsymbol{\tau}\,dx}{\|\mathbf{s}\|_{[L^2(\Omega)]^2}} = \|\boldsymbol{\tau}\|_{[L^2(\Omega)]^2} = \|\boldsymbol{\tau}\|_{H(\mathrm{div};\Omega)}$$

for all $\boldsymbol{\tau} \in \tilde{M}_1$, which establishes the continuous inf-sup condition for $\mathbf{B}_1$.

Finally, it follows from the definition of $\mathbf{A}_1$ that $\mathbf{A}_1(0) = \mathbf{O}$.

Therefore, these results and a direct application of Theorem 2.1 complete the proof. □

It is worth mentioning that examples of dual-dual variational formulations of the form (1), with $\mathbf{S} \neq \mathbf{O}$, are obtained when one applies the coupling of mixed-FEM and BEM to a certain class of transmission problems (see, e.g., [11] and [13]).

4.2. The fully discrete Galerkin scheme

For simplicity, in what follows we assume that the boundary Γ of Ω is a polygonal curve. In order to define our fully discrete Galerkin scheme, we first introduce specific finite element subspaces.

Thus, let $\mathcal{T}_h$ be a regular triangulation of Ω made up of triangles T of diameter h_T such that $h := \sup\limits_{T\in\mathcal{T}_h} h_T$ and $\bar{\Omega} = \cup\{\, T : \quad T \in \mathcal{T}_h \,\}$. Then, we consider the canonical triangle with vertices $\hat{P}_1 = (0,0)^T$, $\hat{P}_2 = (1,0)^T$ and $\hat{P}_3 = (0,1)^T$ as a reference triangle $\hat{T}$, and introduce the family of bijective affine mappings $\{F_T\}_{T\in\mathcal{T}_h}$, such that $F_T(\hat{T}) = T$. It is well known that $F_T(\hat{x}) = B_T\,\hat{x} + b_T$ for all $\hat{x} := (\hat{x}_1,\hat{x}_2)^T \in \hat{T}$, where B_T, a square matrix of order 2, and $b_T \in \mathbb{R}^2$, depend only on the vertices of T.

Now, given the reference triangle $\hat{T}$ and a triangle $T \in \mathcal{T}_h$, we let $\mathbb{P}_0(\hat{T})$ and $\mathbb{P}_0(T)$ be the spaces of constant functions defined on $\hat{T}$ and T, respectively. In addition, we consider the lowest order Raviart-Thomas spaces

$$\mathcal{R}T_0(T) := \{\, \boldsymbol{\tau} : \quad \boldsymbol{\tau} = |\det(B_T)|^{-1}\, B_T\,\hat{\boldsymbol{\tau}} \circ F_T^{-1}\,,\ \hat{\boldsymbol{\tau}} \in \mathcal{R}T_0(\hat{T}) \,\}\,,$$

where

$$\mathcal{R}T_0(\hat{T}) := \mathrm{span}\left\{ \begin{pmatrix} 1\\ 0\end{pmatrix}, \begin{pmatrix} 0\\ 1\end{pmatrix}, \begin{pmatrix} \hat{x}_1\\ \hat{x}_2\end{pmatrix} \right\}.$$

Then we define the following finite element subspaces:

$$X_{1,h} := \{\, \mathbf{s}_h \in [L^2(\Omega)]^2 : \quad \mathbf{s}_h|_T \in [\mathbb{P}_0(T)]^2 \quad \forall\, T \in \mathcal{T}_h \,\}\,, \tag{57}$$

$$M_{1,h} := \{\boldsymbol{\tau}_h \in H(\mathrm{div};\Omega) : \quad \boldsymbol{\tau}_h|_T \in \mathcal{R}T_0(T) \quad \forall\, T \in \mathcal{T}_h\}\,, \tag{58}$$

$$M_h := \{\, v_h \in L^2(\Omega) : \quad v_h|_T \in \mathbb{P}_0(T) \quad \forall\, T \in \mathcal{T}_h \,\}\,. \tag{59}$$

It is easy to see, in this case, that $\mathbf{X}_h := X_{1,h} \times M_{1,h} \times M_h$ is a subspace of

X. We remark that the present choice of $X_{1,h}$ is simpler than the one used in [9] and [12], and, as we will show below, it also satisfies the required compatibility conditions.

In the following we associate to each scalar function $\varphi : T \to \mathbb{R}$ the function $\hat{\varphi} := \varphi \circ F_T : \hat{T} \to \mathbb{R}$, and conversely. Then, we introduce a quadrature formula on $\hat{T}$:

$$\hat{Q}(\hat{\varphi}) := \sum_{j=1}^{N} \hat{w}_j \, \hat{\varphi}(\hat{z}_j) \simeq \int_{\hat{T}} \hat{\varphi} \, d\hat{x} \,, \tag{60}$$

where $\{\hat{z}_1, \hat{z}_2, ..., \hat{z}_N\}$ is a set of nodes on $\hat{T}$ and $\{\hat{w}_1, \hat{w}_2, ..., \hat{w}_N\}$ is the corresponding set of weights. We assume that (60) is exact for constant functions $\hat{\varphi} \in \mathbb{P}_0(\hat{T})$, which means that $\sum_{j=1}^{N} \hat{w}_j = \frac{1}{2}$. Next, we introduce the error functional $\hat{E} : C^0(\hat{T}) \to \mathbb{R}$, given by

$$\hat{E}(\hat{\varphi}) := \int_{\hat{T}} \hat{\varphi} \, d\hat{x} \;-\; \hat{Q}(\hat{\varphi}) \,, \tag{61}$$

which, because of the continuous imbedding of $W^{1,\infty}(\hat{T})$ into $C^0(\hat{T})$, satisfies

$$|\hat{E}(\hat{\varphi})| \;\leq\; \left(\frac{1}{2} + \sum_{j=1}^{N} |\hat{w}_j| \right) \|\hat{\varphi}\|_{W^{1,\infty}(\hat{T})} \quad \forall\, \hat{\varphi} \in W^{1,\infty}(\hat{T}) \,. \tag{62}$$

Hereafter, given any integer $p \geq 0$ and a bounded domain D, $|\cdot|_{W^{p,\infty}(D)}$ and $\|\cdot\|_{W^{p,\infty}(D)}$ denote the usual semi-norm and norm of the Sobolev space $W^{p,\infty}(D)$.

In this way, the Bramble-Hilbert Lemma and the fact that $\hat{E}(\hat{\varphi}) = 0$ for all $\hat{\varphi} \in \mathbb{P}_0(\hat{T})$, yield the existence of a constant $\hat{C} > 0$, depending only on $\hat{T}$, such that

$$|\hat{E}(\hat{\varphi})| \;\leq\; \hat{C}\, |\hat{\varphi}|_{W^{1,\infty}(\hat{T})} \quad \forall\, \hat{\varphi} \in W^{1,\infty}(\hat{T}) \,. \tag{63}$$

On the other hand, the quadrature formula on each $T \in \mathcal{T}_h$ is defined by

$$Q_T(\varphi) := |\det(B_T)|\, \hat{Q}(\hat{\varphi}) \;=\; 2\,|T|\, \hat{Q}(\hat{\varphi}) \simeq \int_T \varphi \, dx \tag{64}$$

for all $\varphi \in C^0(T)$, with corresponding error functional $E_T : C^0(T) \to \mathbb{R}$ satisfying

$$E_T(\varphi) := \int_T \varphi \, dx \;-\; Q_T(\varphi) \;=\; |\det(B_T)|\, \hat{E}(\hat{\varphi}) \quad \forall\, \varphi \in C^0(T) \,. \tag{65}$$

Similarly, we introduce a quadrature formula on $\tilde{e} := [0,1]$:

$$\tilde{Q}(\tilde{\varphi}) := \sum_{j=1}^{m} \tilde{w}_j \tilde{\varphi}(\tilde{z}_j) \simeq \int_{\tilde{e}} \tilde{\varphi}(\tilde{t}) \, d\tilde{t} \quad \forall\, \tilde{\varphi} \in C^0(\tilde{e}) \,, \tag{66}$$

where $\{\tilde{z}_1, \tilde{z}_2, ..., \tilde{z}_m\}$ is a set of nodes in $\tilde{e}$ and $\{\tilde{w}_1, \tilde{w}_2, ..., \tilde{w}_m\}$ is the corresponding set of weights. We assume that (66) is exact for linear functions $\tilde{\varphi} \in \mathbb{P}_1(\tilde{e})$. In this way, the error functional $\tilde{E} : C^0(\tilde{e}) \to \mathbb{R}$ is given by

$$\tilde{E}(\tilde{\varphi}) := \int_{\tilde{e}} \tilde{\varphi}(\tilde{t})\, d\tilde{t} - \tilde{Q}(\tilde{\varphi})\,,$$

which, because of the continuous imbedding of $W^{2,\infty}(\tilde{e})$ into $C^0(\tilde{e})$, satisfies

$$|\tilde{E}(\tilde{\varphi})| \,\leq\, \left(1 + \sum_{j=1}^{m} |\tilde{w}_j|\right) \|\tilde{\varphi}\|_{W^{2,\infty}(\tilde{e})} \quad \forall\, \tilde{\varphi} \in W^{2,\infty}(\tilde{e})\,.$$

Thus, the Bramble-Hilbert Lemma and the fact that $\tilde{E}(\tilde{\varphi}) = 0$ for all $\tilde{\varphi} \in \mathbb{P}_1(\tilde{e})$, implies the existence of $\tilde{C} > 0$, depending only on $\tilde{e}$, such that

$$|\tilde{E}(\tilde{\varphi})| \,\leq\, \tilde{C}\, |\tilde{\varphi}|_{W^{2,\infty}(\tilde{e})} \quad \forall\, \tilde{\varphi} \in W^{2,\infty}(\tilde{e})\,. \tag{67}$$

In addition, given an edge e of a triangle $T \in \mathcal{T}_h$, the associated quadrature formula is defined by

$$Q_e(\varphi) := |e|\, \tilde{Q}(\tilde{\varphi}) \simeq \int_e \varphi(t)\, dt\,, \tag{68}$$

with $\tilde{\varphi} := \varphi \circ F_T$, and the corresponding error functional $E_e : C^0(e) \to \mathbb{R}$ satisfies

$$E_e(\varphi) := \int_e \varphi(t)\, dt - Q_e(\varphi) = |e|\, \tilde{E}(\tilde{\varphi}) \tag{69}$$

for all $\varphi \in C^0(e)$.

In what follows, given $\mathbf{s}_h \in X_{1,h}$, $\boldsymbol{\tau}_h \in M_{1,h}$ and $v_h \in M_h$ we denote, respectively, by $\mathbf{s}_{h,T}$, $\boldsymbol{\tau}_{h,T}$ and $v_{h,T}$ the corresponding restrictions to the triangle $T \in \mathcal{T}_h$.

Then, we introduce the discrete operators $\mathbf{A}_{1,h} : X_{1,h} \to X'_{1,h}$, $\mathbf{B}_{1,h} : X_{1,h} \to M'_{1,h}$ and $\mathbf{B}_h : M_{1,h} \to M'_h$, as follows:

$$\begin{aligned} [\mathbf{A}_{1,h}(\mathbf{r}_h), \mathbf{s}_h] &:= \sum_{T\in\mathcal{T}_h} Q_T(\boldsymbol{\kappa}(\cdot, |\mathbf{r}_{h,T}|))\, \mathbf{r}_{h,T} \cdot \mathbf{s}_{h,T} \\ &= \sum_{T\in\mathcal{T}_h} 2\,|T| \left\{ \sum_{j=1}^{N} \hat{w}_j\, \boldsymbol{\kappa}(F_T(\hat{z}_j), |\mathbf{r}_{h,T}|) \right\} \mathbf{r}_{h,T} \cdot \mathbf{s}_{h,T} \end{aligned} \tag{70}$$

for all $\mathbf{r}_h$, $\mathbf{s}_h \in X_{1,h}$,

$$\begin{aligned} [\mathbf{B}_{1,h}(\mathbf{r}_h), \boldsymbol{\tau}_h] &:= -\sum_{T\in\mathcal{T}_h} Q_T(\mathbf{r}_{h,T} \cdot \boldsymbol{\tau}_{h,T}) \\ &= -\sum_{T\in\mathcal{T}_h} 2|T| \left\{ \sum_{j=1}^{N} \hat{w}_j\, \mathbf{r}_{h,T} \cdot \boldsymbol{\tau}_{h,T}(F_T(\hat{z}_j)) \right\} \end{aligned} \tag{71}$$

for all $\mathbf{r}_h \in X_{1,h}$, $\boldsymbol{\tau}_h \in M_{1,h}$, and

$$\begin{aligned} [\mathbf{B}_h(\boldsymbol{\tau}_h), v_h] &:= - \sum_{T \in \mathcal{T}_h} Q_T(v_{h,T} \operatorname{div} \boldsymbol{\tau}_{h,T}) \\ &= - \sum_{T \in \mathcal{T}_h} 2|T| \left\{ \sum_{j=1}^{N} \hat{w}_j \, v_{h,T} \operatorname{div} \boldsymbol{\tau}_{h,T} \right\} \end{aligned} \tag{72}$$

for all $\boldsymbol{\tau}_h \in M_{1,h}$, $v_h \in M_h$.

It is easy to see that $\mathbf{A}_{1,h}$, $\mathbf{B}_{1,h}$ and $\mathbf{B}_h$ are well defined in the sense that they do map, respectively, $X_{1,h}$, $X_{1,h}$ and $M_{1,h}$, into the indicated dual spaces. In particular, using (54) we observe from (70) that for all $\mathbf{r}_h$, $\mathbf{s}_h \in X_{1,h}$, it holds

$$\begin{aligned} |[\mathbf{A}_{1,h}(\mathbf{r}_h), \mathbf{s}_h]| &\le \left(2\kappa_1 \sum_{j=1}^{N} |\hat{w}_j| \right) \sum_{T \in \mathcal{T}_h} |T| \, |\mathbf{r}_{h,T}| \, |\mathbf{s}_{h,T}| \\ &= \left(2\kappa_1 \sum_{j=1}^{N} |\hat{w}_j| \right) \sum_{T \in \mathcal{T}_h} \|\mathbf{r}_{h,T}\|_{[L^2(T)]^2} \, \|\mathbf{s}_{h,T}\|_{[L^2(T)]^2} \\ &\le \left(2\kappa_1 \sum_{j=1}^{N} |\hat{w}_j| \right) \|\mathbf{r}_h\|_{[L^2(\Omega)]^2} \, \|\mathbf{s}_h\|_{[L^2(\Omega)]^2} \,, \end{aligned}$$

which proves that $\mathbf{A}_{1,h}(\mathbf{r}_h) \in X'_{1,h}$ for all $\mathbf{r}_h \in X_{1,h}$.

Also, since $v_{h,T} \operatorname{div} \boldsymbol{\tau}_{h,T}$ is constant, we observe from (72) that

$$[\mathbf{B}_h(\boldsymbol{\tau}_h), v_h] \, = \, [\mathbf{B}(\boldsymbol{\tau}_h), v_h] \quad \forall \, (\boldsymbol{\tau}_h, v_h) \in M_{1,h} \times M_h \,. \tag{73}$$

Now, assuming additionally that $g \in W^{2,\infty}(\Gamma)$ and $f \in W^{1,\infty}(\Omega)$, one can define the discrete functionals $\mathbf{G}_{1,h} \in M'_{1,h}$ and $\mathbf{G}_h \in M'_h$, as follows:

$$[\mathbf{G}_{1,h}, \boldsymbol{\tau}_h] := - \sum_{e \subseteq \Gamma} Q_e(g) \, \boldsymbol{\tau}_h \cdot \boldsymbol{\nu}|_e = \, - \sum_{e \subseteq \Gamma} |e| \left\{ \sum_{j=1}^{m} \tilde{w}_j \, \tilde{g}(\tilde{z}_j) \right\} \boldsymbol{\tau}_h \cdot \boldsymbol{\nu}|_e \tag{74}$$

for all $\boldsymbol{\tau}_h \in M_{1,h}$, and

$$[\mathbf{G}_h, v_h] := \sum_{T \in \mathcal{T}_h} Q_T(f) \, v_{h,T} \, = \, \sum_{T \in \mathcal{T}_h} |T| \left\{ \sum_{j=1}^{N} \hat{w}_j \, f(F_T(\hat{z}_j)) \right\} v_{h,T} \tag{75}$$

for all $v_h \in M_h$.

The reason for considering the above regularity assumptions on g and f will become clear from the error estimate given in Lemma 4.4, from which one also deduces the uniform boundedness of $\mathbf{G}_{1,h}$ and $\mathbf{G}_h$. Indeed, noting that

$$\|g\|_{H^{1/2}(\Gamma)} \, \le \, C_1 \, \|g\|_{W^{2,\infty}(\Gamma)} \quad \text{and} \quad \|f\|_{L^2(\Omega)} \, \le \, C_2 \, \|f\|_{W^{1,\infty}(\Omega)} \,,$$

with positive constants C_1 and C_2 depending only on Γ and Ω, respectively, we deduce, using triangle's inequality, that

$$\|\mathbf{G}_{1,h}\|_{M'_{1,h}} \leq C\,\|g\|_{W^{2,\infty}(\Gamma)} \text{ and } \|\mathbf{G}_h\|_{M'_h} \leq C\,\|f\|_{W^{1,\infty}(\Omega)}\,, \tag{76}$$

with a constant $C > 0$ independent of h.

Thus, the fully discrete Galerkin scheme associated with (56) reads: *Find* $(\mathbf{t}_h, \boldsymbol{\sigma}_h, u_h) \in \mathbf{X}_h$ *such that*

$$\begin{array}{llllll} [\mathbf{A}_{1,h}(\mathbf{t}_h), \mathbf{s}_h] & +[\mathbf{B}_{1,h}(\mathbf{s}_h), \boldsymbol{\sigma}_h] & & = & 0\,, \\ [\mathbf{B}_{1,h}(\mathbf{t}_h), \boldsymbol{\tau}_h] & + & [\mathbf{B}_h(\boldsymbol{\tau}_h), u_h] & = & [\mathbf{G}_{1,h}, \boldsymbol{\tau}_h]\,, \\ & [\mathbf{B}_h(\boldsymbol{\sigma}_h), v_h] & & = & [\mathbf{G}_h, v_h]\,, \end{array} \tag{77}$$

for all $(\mathbf{s}_h, \boldsymbol{\tau}_h, v_h) \in \mathbf{X}_h$.

4.3. Analysis of the fully discrete Galerkin scheme

We begin with the following lemma.

Lemma 4.1: *The discrete nonlinear operator $\mathbf{A}_{1,h}$ is Lipschitz continuous (with a Lipschitz constant $\gamma > 0$, independent of $h \in \mathtt{I}$) and strongly monotone (with a monotonicity constant $\alpha > 0$, independent of $h \in \mathtt{I}$).*

Proof: Let $\mathbf{r}_h$, $\tilde{\mathbf{r}}_h$, $\mathbf{s}_h \in X_{1,h}$. Then, from the definition of $\mathbf{A}_{1,h}$ (cf. (70)), we can write

$$\begin{aligned} & |\,[\mathbf{A}_{1,h}(\mathbf{r}_h) - \mathbf{A}_{1,h}(\tilde{\mathbf{r}}_h), \mathbf{s}_h]\,| \\ \leq\; & 2 \sum_{T\in\mathcal{T}_h} |T| \left(\sum_{j=1}^{N} |\hat{w}_j| \left| \boldsymbol{\kappa}(F_T(\hat{z}_j), |\mathbf{r}_{h,T}|)\,\mathbf{r}_{h,T} - \boldsymbol{\kappa}(F_T(\hat{z}_j), |\tilde{\mathbf{r}}_{h,T}|)\,\tilde{\mathbf{r}}_{h,T} \right| \right) |\mathbf{s}_{h,T}|\,, \end{aligned}$$

which, according to the assumptions on $\boldsymbol{\kappa}$ (cf. (54)), yields

$$\begin{aligned} & |\,[\mathbf{A}_{1,h}(\mathbf{r}_h) - \mathbf{A}_{1,h}(\tilde{\mathbf{r}}_h), \mathbf{s}_h]\,| \\ \leq\; & C\,\kappa_1 \left(\sum_{j=1}^{N} |\hat{w}_j| \right) \sum_{T\in\mathcal{T}_h} \|\mathbf{r}_{h,T} - \tilde{\mathbf{r}}_{h,T}\|_{[L^2(T)]^2}\, \|\mathbf{s}_{h,T}\|_{[L^2(T)]^2}\,. \end{aligned}$$

In this way, the above estimate and the Cauchy-Schwarz inequality prove the Lipschitz continuity of $\mathbf{A}_{1,h}$.

On the other hand, from (70) we also get

$$\begin{aligned} & [\mathbf{A}_{1,h}(\mathbf{r}_h) - \mathbf{A}_{1,h}(\mathbf{s}_h), \mathbf{r}_h - \mathbf{s}_h] \\ =\; & 2 \sum_{T\in\mathcal{T}_h} |T| \sum_{j=1}^{N} \hat{w}_j \left[\boldsymbol{\kappa}(F_T(\hat{z}_j), |\mathbf{r}_{h,T}|)\,\mathbf{r}_{h,T} - \boldsymbol{\kappa}(F_T(\hat{z}_j), |\mathbf{s}_{h,T}|)\,\mathbf{s}_{h,T} \right] \cdot (\mathbf{r}_{h,T} - \mathbf{s}_{h,T})\,, \end{aligned}$$

which, using again the assumptions on $\boldsymbol{\kappa}$ (cf. (54)), implies that

$$[\mathbf{A}_{1,h}(\mathbf{r}_h) - \mathbf{A}_{1,h}(\mathbf{s}_h), \mathbf{r}_h - \mathbf{s}_h]$$

$$\geq\ 2\,\kappa_0 \sum_{j=1}^{N} \hat{w}_j \sum_{T\in\mathcal{T}_h} |T|\,|\mathbf{r}_{h,T} - \mathbf{s}_{h,T}|^2 = \kappa_0\,\|\mathbf{r}_h - \mathbf{s}_h\|^2_{[L^2(\Omega)]^2}\,.$$

This shows the uniform strong monotonicity of $\mathbf{A}_{1,h}$ and completes the proof of the Lemma. □

The following lemma provides the uniform discrete inf-sup conditions for the operators $\mathbf{B}_{1,h}$ and $\mathbf{B}_h$.

LEMMA 4.2: *Let* $\bar{M}_{1,h} := \{\boldsymbol{\tau}_h \in M_{1,h} :\ [\mathbf{B}_h(\boldsymbol{\tau}_h), v_h] = 0\ \forall\, v_h \in M_h\}$. *Then, there exists* $\beta_1^* > 0$, *independent of* $h \in \mathtt{I}$, *such that*

$$\sup_{\substack{\mathbf{s}_h\in X_{1,h}\\ \mathbf{s}_h\neq 0}} \frac{[\mathbf{B}_{1,h}(\mathbf{s}_h), \boldsymbol{\tau}_h]}{\|\mathbf{s}_h\|_{X_1}} \geq\ \beta_1^*\,\|\boldsymbol{\tau}_h\|_{M_1} \quad \forall\, \boldsymbol{\tau}_h \in \bar{M}_{1,h}\,.$$

In addition, there exists $\beta^* > 0$, *independent of* $h \in \mathtt{I}$, *such that*

$$\sup_{\substack{\boldsymbol{\tau}_h\in M_{1,h}\\ \boldsymbol{\tau}_h\neq 0}} \frac{[\mathbf{B}_h(\boldsymbol{\tau}_h), v_h]}{\|\boldsymbol{\tau}_h\|_{M_1}} \geq\ \beta^*\,\|v_h\|_{M} \quad \forall\, v_h \in M_h\,.$$

Proof: We first observe from (73) that

$$\bar{M}_{1,h} := \{\boldsymbol{\tau}_h \in M_{1,h} :\ [\mathbf{B}(\boldsymbol{\tau}_h), v_h] = 0\ \forall\, v_h \in M_h\}$$

$$= \{\boldsymbol{\tau}_h \in M_{1,h} :\ \int_\Omega v_h \operatorname{div} \boldsymbol{\tau}_h\, dx = 0\ \forall\, v_h \in M_h\}\,,$$

and hence

$$\bar{M}_{1,h} = \{\boldsymbol{\tau}_h \in M_{1,h} :\ \operatorname{div} \boldsymbol{\tau}_h\ =\ 0\,\}\,. \tag{78}$$

It follows that $\boldsymbol{\tau}_{h,T} \in [\mathbb{P}_0(T)]^2$ for all $\boldsymbol{\tau}_h \in \bar{M}_{1,h}$, which, in virtue of the definition of $\mathbf{B}_{1,h}$ (cf. (71)), yields

$$[\mathbf{B}_{1,h}(\mathbf{s}_h), \boldsymbol{\tau}_h] = [\mathbf{B}_1(\mathbf{s}_h), \boldsymbol{\tau}_h] \quad \forall\,(\mathbf{s}_h, \boldsymbol{\tau}_h) \in X_{1,h} \times \bar{M}_{1,h}\,.$$

Equivalently, (78) implies that $\bar{M}_{1,h} \subseteq X_{1,h}$ and therefore, for all $\boldsymbol{\tau}_h \in \bar{M}_{1,h}$ we obtain

$$\sup_{\substack{\mathbf{s}_h\in X_{1,h}\\ \mathbf{s}_h\neq 0}} \frac{[\mathbf{B}_{1,h}(\mathbf{s}_h), \boldsymbol{\tau}_h]}{\|\mathbf{s}_h\|_{X_1}} = \sup_{\substack{\mathbf{s}_h\in X_{1,h}\\ \mathbf{s}_h\neq 0}} \frac{[\mathbf{B}_1(\mathbf{s}_h), \boldsymbol{\tau}_h]}{\|\mathbf{s}_h\|_{X_1}}$$

$$= \sup_{\substack{\mathbf{s}_h\in X_{1,h}\\ \mathbf{s}_h\neq 0}} \frac{-\int_\Omega \mathbf{s}_h \cdot \boldsymbol{\tau}_h\, dx}{\|\mathbf{s}_h\|_{X_1}} = \|\boldsymbol{\tau}_h\|_{[L^2(\Omega)]^2} = \|\boldsymbol{\tau}_h\|_{M_1}\,,$$

which proves the required property for $\mathbf{B}_{1,h}$ with $\beta_1^* = 1$.

On the other hand, because of the identity (73), the discrete inf-sup condition for $\mathbf{B}_h$ reduces to the same property for $\mathbf{B}$, which has already been proved in several places (see, e.g., Lemma 4.3 in [19] or Lemma 5.6 in [12]). We omit further details. □

As a consequence of Lemmas 4.1 and 4.2, we can state the following result.

THEOREM 4.2: *Assume that $f \in W^{1,\infty}(\Omega)$ and that $g \in W^{2,\infty}(\Gamma)$. Then, there exists a unique $(\mathbf{t}_h, \boldsymbol{\sigma}_h, u_h) \in \mathbf{X}_h$ solution of (77). Moreover, there exists $\tilde{C} > 0$, independent of $h \in \mathtt{I}$, such that*

$$\|(\mathbf{t}_h, \boldsymbol{\sigma}_h, u_h)\|_{\mathbf{X}} \leq \tilde{C} \left\{ \|g\|_{W^{2,\infty}(\Gamma)} + \|f\|_{W^{1,\infty}(\Omega)} \right\} .$$

Proof: It is a straightforward application of the abstract Theorem 3.4 and the estimates (76). □

Our next purpose is to apply the Strang-type error estimate provided by Theorem 3.7. We first remark that the operator $\mathbf{A}_{1,h}$ defined by (70) verifies the conditions **(H.3)** and **(H.4)**. Since this is a technical result, which mainly follows from the assumptions stated in (54), we omit the corresponding details.

The following lemma gives the necessary estimates for the second consistency term in the Strang-type error bound (52).

LEMMA 4.3: *There exists $C > 0$, independent of $h \in \mathtt{I}$, such that*

$$|\,[\mathbf{A}_1(\mathbf{r}_h), \mathbf{s}_h] - [\mathbf{A}_{1,h}(\mathbf{r}_h), \mathbf{s}_h]\,| \leq C\, h\, \|\mathbf{r}_h\|_{[L^2(\Omega)]^2} \|\mathbf{s}_h\|_{[L^2(\Omega)]^2} \tag{79}$$

for all $\mathbf{r}_h,\, \mathbf{s}_h \in X_{1,h}$, and

$$|\,[\mathbf{B}_1(\mathbf{s}_h), \boldsymbol{\tau}_h] - [\mathbf{B}_{1,h}(\mathbf{s}_h), \boldsymbol{\tau}_h] \leq C\, h\, \|\mathbf{s}_h\|_{[L^2(\Omega)]^2} \|\boldsymbol{\tau}_h\|_{H(\mathrm{div};\Omega)} \tag{80}$$

for all $(\mathbf{s}_h, \boldsymbol{\tau}_h) \in X_{1,h} \times M_{1,h}$.

Proof: Let $\mathbf{r}_h,\, \mathbf{s}_h \in X_{1,h}$. Then, from the definitions of $\mathbf{A}_1$ and $\mathbf{A}_{1,h}$ we can write

$$|[\mathbf{A}_1(\mathbf{r}_h), \mathbf{s}_h] - [\mathbf{A}_{1,h}(\mathbf{r}_h), \mathbf{s}_h]| \leq \sum_{T \in \mathcal{T}_h} |E_T(\boldsymbol{\kappa}(\cdot, |\mathbf{r}_{h,T}|))\, \mathbf{r}_{h,T} \cdot \mathbf{s}_{h,T}|. \tag{81}$$

Now, using (65), (63) and (54), we obtain

$$\begin{aligned} |E_T(\boldsymbol{\kappa}(\cdot, |\mathbf{r}_{h,T}|))\, \mathbf{r}_{h,T} \cdot \mathbf{s}_{h,T}\,| &\leq 2\, |T|\, |\mathbf{r}_{h,T}|\, |\mathbf{s}_{h,T}|\, \left|\hat{E}(\hat{\boldsymbol{\kappa}}(\cdot, |\mathbf{r}_{h,T}|))\right| \\ &\leq 2\, \hat{C}\, |T|\, |\mathbf{r}_{h,T}|\, |\mathbf{s}_{h,T}|\, |\hat{\boldsymbol{\kappa}}(\cdot, |\mathbf{r}_{h,T}|)|_{W^{1,\infty}(\hat{T})} \\ &\leq C\, h_T\, |T|\, |\mathbf{r}_{h,T}|\, |\mathbf{s}_{h,T}|\, |\boldsymbol{\kappa}(\cdot, |\mathbf{r}_{h,T}|)|_{W^{1,\infty}(T)} \\ &\leq C \kappa_1\, h_T\, |T|\, |\mathbf{r}_{h,T}|\, |\mathbf{s}_{h,T}| \end{aligned}$$

$$= C\,\kappa_1\, h_T\, \|\mathbf{r}_{h,T}\|_{[L^2(T)]^2}\, \|\mathbf{s}_{h,T}\|_{[L^2(T)]^2}\,. \tag{82}$$

For the last equality, we also used that $|T|^{1/2}\,|\mathbf{r}_{h,T}| = \|\mathbf{r}_{h,T}\|_{[L^2(T)]^2}$ and that $|T|^{1/2}\,|\mathbf{s}_{h,T}| = \|\mathbf{s}_{h,T}\|_{[L^2(T)]^2}$.

In this way, (81), (82) and the Cauchy-Schwarz inequality yield (79).

On the other hand, from the definitions of $\mathbf{B}_1$ and $\mathbf{B}_{1,h}$ we have

$$|\,[\mathbf{B}_1(\mathbf{s}_h), \boldsymbol{\tau}_h] \,-\, [\mathbf{B}_{1,h}(\mathbf{s}_h), \boldsymbol{\tau}_h] \,\le\, \sum_{T\in\mathcal{T}_h} |E_T(\mathbf{s}_{h,T}\cdot\boldsymbol{\tau}_{h,T})| \tag{83}$$

for all $(\mathbf{s}_h, \boldsymbol{\tau}_h) \in X_{1,h} \times M_{1,h}$. Next, using again (65) and (63), we get

$$|E_T(\mathbf{s}_{h,T}\cdot\boldsymbol{\tau}_{h,T})| \,=\, 2\,|T|\;|\hat{E}(\hat{\mathbf{s}}_{h,T}\cdot\hat{\boldsymbol{\tau}}_{h,T})|$$

$$\le\, 2\,|T|\,\hat{C}\;|\,\hat{\mathbf{s}}_{h,T}\cdot\hat{\boldsymbol{\tau}}_{h,T}\,|_{W^{1,\infty}(\hat{T})} \le\, 2\,|T|\,\hat{C}\;|\mathbf{s}_{h,T}|\;|\mathrm{div}\,\hat{\boldsymbol{\tau}}_{h,T}|$$

$$\le\, C\,h_T\,|T|\,|\mathbf{s}_{h,T}|\;|\mathrm{div}\,\boldsymbol{\tau}_{h,T}| \,\le\, C\,h_T\,\|\mathbf{s}_{h,T}\|_{[L^2(T)]^2}\;\|\boldsymbol{\tau}_{h,T}\|_{H(\mathrm{div};T)}\,. \tag{84}$$

Therefore, (83), (84) and the Cauchy-Schwarz inequality imply (80), thus completing the proof of the lemma. □

We now provide the necessary estimates for the first consistency term in (52).

LEMMA 4.4: *Assume that $f \in W^{1,\infty}(\Omega)$ and that $g \in W^{2,\infty}(\Gamma)$. Then there exists $C > 0$, independent of $h \in \mathtt{I}$, such that*

$$|\,[\mathbf{G}, v_h] - [\mathbf{G}_h, v_h]\,| \,\le\, C\,h\,|f|_{W^{1,\infty}(\Omega)}\,\|v_h\|_{L^2(\Omega)} \tag{85}$$

for all $v_h \in M_h$, and

$$|\,[\mathbf{G}_1, \boldsymbol{\tau}_h] - [\mathbf{G}_{1,h}, \boldsymbol{\tau}_h]\,| \,\le\, C\,h^{3/2}\,|g|_{W^{2,\infty}(\Gamma)}\,\|\boldsymbol{\tau}_h\|_{H(\mathrm{div};\Omega)} \tag{86}$$

for all $\boldsymbol{\tau}_h \in M_{1,h}$.

Proof: Let $v_h \in M_h$. Then it is easy to see that

$$|\,[\mathbf{G}, v_h] - [\mathbf{G}_h, v_h]\,| \,\le\, \sum_{T\in\mathcal{T}_h} |v_{h,T}|\;|E_T(f)|\,. \tag{87}$$

Now, proceeding as in the previous lemmas, we also get

$$|E_T(f)| \,=\, 2\,|T|\,|\hat{E}(\hat{f})| \,\le\, 2\,|T|\,|\hat{f}|_{W^{1,\infty}(\hat{T})} \,\le\, C\,h_T\,|T|\,|f|_{W^{1,\infty}(T)}\,.$$

Hence, this estimate, (87) and the Cauchy-Schwarz inequality yield

$$|\,[\mathbf{G}, v_h] - [\mathbf{G}_h, v_h]\,| \,\le\, C\,h\,|f|_{W^{1,\infty}(\Omega)} \sum_{T\in\mathcal{T}_h} |T|^{1/2}\,\|v_{h,T}\|_{L^2(T)}$$

$$\le\, C\,h\,|\Omega|^{1/2}\,|f|_{W^{1,\infty}(\Omega)}\,\|v_h\|_{L^2(\Omega)}\,,$$

which proves (85).

On the other hand, given $\boldsymbol{\tau}_h \in M_{1,h}$, we obtain

$$|\,[\mathbf{G}_1, \boldsymbol{\tau}_h] - [\mathbf{G}_{1,h}, \boldsymbol{\tau}_h]\,| \;\leq\; \sum_{e \subseteq \Gamma} \Big|(\boldsymbol{\tau}_h \cdot \boldsymbol{\nu})|_e\Big| \; \Big|E_e(g)\Big| \,, \tag{88}$$

which, using (69) and (67), gives

$$|\,[\mathbf{G}_1, \boldsymbol{\tau}_h] - [\mathbf{G}_{1,h}, \boldsymbol{\tau}_h]\,| \;\leq\; \sum_{e \subseteq \Gamma} \Big|(\boldsymbol{\tau}_h \cdot \boldsymbol{\nu})|_e\Big| \; |e| \; \Big|\,\tilde{E}(\tilde{g})\,\Big|$$

$$\leq\; \tilde{C} \sum_{e \subseteq \Gamma} \Big|(\boldsymbol{\tau}_h \cdot \boldsymbol{\nu})|_e\Big| \; |e| \; \Big|\tilde{g}\Big|_{W^{2,\infty}(\tilde{e})} \;\leq\; C \sum_{e \subseteq \Gamma} \Big|(\boldsymbol{\tau}_h \cdot \boldsymbol{\nu})|_e\Big| \; |e|^3 \; \Big|g\Big|_{W^{2,\infty}(e)} \,.$$

We remark here that $(\boldsymbol{\tau}_h \cdot \boldsymbol{\nu})|_e$ is constant through each edge e.

Therefore, noting that $|e| \leq h$, and using the Cauchy-Schwarz inequality and the inverse estimate for the functions in $M_{1,h}$, we conclude that

$$\begin{aligned} |\,[\mathbf{G}_1, \boldsymbol{\tau}_h] - [\mathbf{G}_{1,h}, \boldsymbol{\tau}_h]\,| \;&\leq\; C\, h^2 \, |g|_{W^{2,\infty}(\Gamma)} \sum_{e \subseteq \Gamma} |e|^{1/2} \Big|(\boldsymbol{\tau}_h \cdot \boldsymbol{\nu})|_e\Big| \; |e|^{1/2} \\ &\leq\; C\, h^2 \, |g|_{W^{2,\infty}(\Gamma)} \; |\Gamma|^{1/2} \, \|\boldsymbol{\tau}_h \cdot \boldsymbol{\nu}\|_{L^2(\Gamma)} \\ &\leq\; C\, h^{3/2} \, |g|_{W^{2,\infty}(\Gamma)} \; |\Gamma|^{1/2} \, \|\boldsymbol{\tau}_h \cdot \boldsymbol{\nu}\|_{H^{-1/2}(\Gamma)} \\ &\leq\; C\, h^{3/2} \, |g|_{W^{2,\infty}(\Gamma)} \; |\Gamma|^{1/2} \, \|\boldsymbol{\tau}_h\|_{H(\mathrm{div};\Omega)} \,, \end{aligned}$$

which proves (86). □

We are now in a position to establish the following main result.

THEOREM 4.3: *Let* $\vec{\mathbf{t}} := (\mathbf{t}, \boldsymbol{\sigma}, u) \in \mathbf{X}$ *and* $\vec{\mathbf{t}}_h := (\mathbf{t}_h, \boldsymbol{\sigma}_h, u_h) \in \mathbf{X}_h$ *be the unique solutions of* (56) *and* (77)*, respectively. Assume that* $f \in W^{1,\infty}(\Omega)$ *and that* $g \in W^{2,\infty}(\Gamma)$*. Then there exists* $\tilde{C} > 0$*, independent of* $h \in \mathtt{I}$*, such that the following error estimate holds:*

$$\|\vec{\mathbf{t}} - \vec{\mathbf{t}}_h\|_{\mathbf{X}} \;\leq\; \tilde{C}\, h \,\Big[1 + \|\vec{\mathbf{t}}\|_{\mathbf{X}}\Big] \;+\; \tilde{C} \inf_{\vec{\mathbf{r}}_h \in \mathbf{X}_h} \|\vec{\mathbf{t}} - \vec{\mathbf{r}}_h\|_{\mathbf{X}} \,.$$

Proof: It follows from Lemmas 4.3, 4.4 and the abstract Theorem 3.7. □

We end this section with the following corollary of the previous theorem.

THEOREM 4.4: *In addition to the hypotheses of Theorem* 4.3*, assume that* $\mathbf{t} \in [H^1(\Omega)]^2$, $\boldsymbol{\sigma} \in [H^1(\Omega)]^2$, $\mathrm{div}\,\boldsymbol{\sigma} \in H^1(\Omega)$ *and* $u \in H^1(\Omega)$*. Then there exists* $\bar{C} > 0$*, independent of* $h \in \mathtt{I}$*, such that the following rate of convergence holds:*

$$\begin{aligned} &\|\vec{\mathbf{t}} - \vec{\mathbf{t}}_h\|_{\mathbf{X}} \\ &\leq\; \bar{C}\, h \,\left[1 + \|\mathbf{t}\|_{[H^1(\Omega)]^2} + \|\boldsymbol{\sigma}\|_{[H^1(\Omega)]^2} + \|\mathrm{div}\,\boldsymbol{\sigma}\|_{H^1(\Omega)} + \|u\|_{H^1(\Omega)}\right] \,. \end{aligned}$$

Proof: It follows from Theorem 4.3 and classical error estimates for interpolation and projection operators in the corresponding Sobolev spaces. □

5. Numerical results

We consider two examples to demonstrate the convergence of the fully discrete Galerkin scheme (77). In both cases we choose $\Omega = (0,1) \times (0,1)$. We take the right hand side functions f and g in (53) such that $u(x_1, x_2) = x_1^2 + x_2^2$ and $\boldsymbol{\kappa}(x,s) = 2 + 1/(1+s)$ in the first example and $u(x_1,x_2) = 1/(x_1 + x_2 + 0.1)$ and $\boldsymbol{\kappa}$ as before in the second one. Obviously, both solutions are smooth and $\boldsymbol{\kappa}$ satisfies (54) with $\kappa_0 = 1$ and $\kappa_1 = 3$ (see Example 1 in [21]). In the first example the corresponding unknowns $\mathbf{t} = \nabla u$, $\boldsymbol{\sigma} = \boldsymbol{\kappa}(\cdot, |\nabla u|)\nabla u$ and u have only small norms whereas in the second example the norms are larger. Indeed, in the second example $x = (0,0)$ is the most critical point and an adaptive strategy could be worthwhile, see [1]. However, the error for uniform meshes eventually must decrease like $O(h)$ in both examples. Only the absolute value of the error can be expected to be larger in the second case or the asymptotic range of $O(h)$-error appears later.

As described in the previous section we use regular triangulations $\mathcal{T}_h$ of Ω (which are uniform for our examples) and the ansatz spaces are given by (57), (58) and (59). As quadrature rule $\hat{Q}$ on the reference triangle we simply take the 1-point formula in the centre of mass and for the quadrature on the boundary pieces we take the midpoint formula. Then constants on triangles and linears on boundary pieces are integrated exactly as required. We note that in contrast to the definition (71) we use exact integration to define the discrete operator $\mathbf{B}_{1,h}$ since a simple local stiffness matrix can be used.

In Figures 1 and 2 the approximation errors for the individual unknowns $\mathbf{t}$, $\boldsymbol{\sigma}$ and u are given for the two examples, respectively. Obviously, in all cases the errors individually decay like h, and this confirms the theoretical error bound given by Theorem 4.4. The nonlinear Galerkin systems are solved by the Newton method where we stop the iteration when the l_2-error of the residual vector is less than 10^{-7}.

The linear systems appearing in the Newton steps are solved by the (preconditioned) minimum residual method as analysed in [8] and [9]. More details concerning the solution of the nonlinear system will appear in a forthcoming work. The approximation errors shown in the figures are calculated on each triangle by the 7 knots quadrature rule from [22] (p. 314).

Acknowledgement. Part of this work was done while the third author was visiting the Departamento de Ingeniería Matemática at the Universidad de Concepción during January 2000.

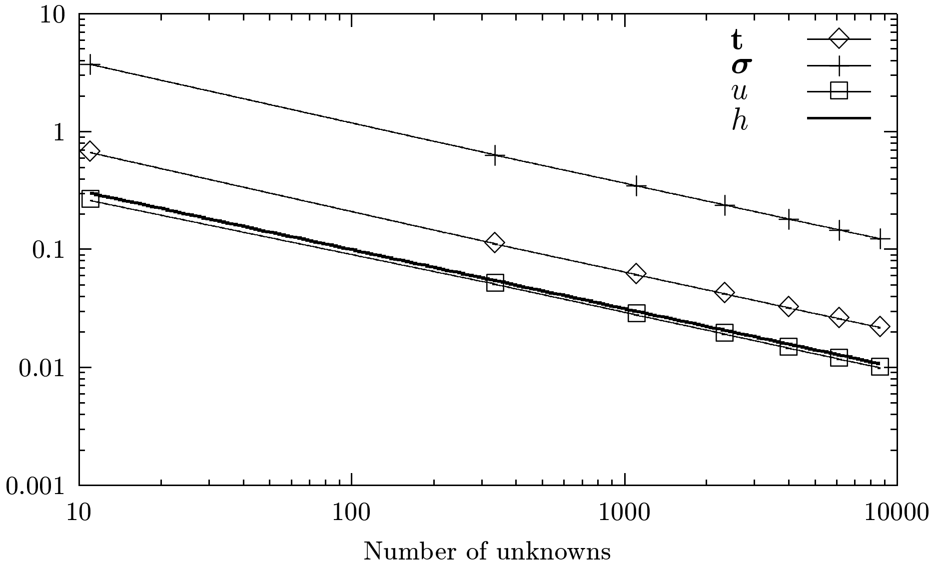

Figure 1: Approximation errors for the example $u(x_1, x_2) = x_1^2 + x_2^2$, $\boldsymbol{\kappa}(\cdot, s) = 2 + 1/(1+s)$

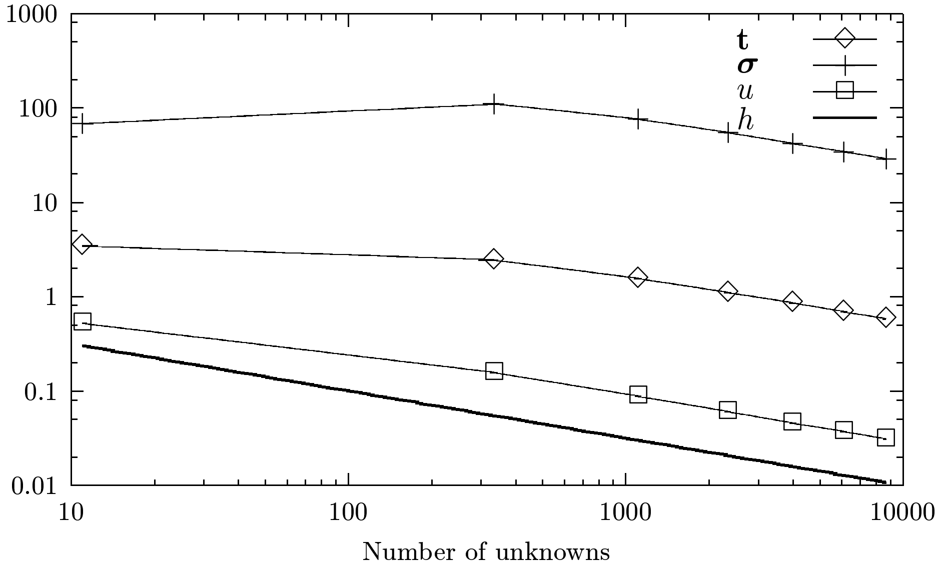

Figure 2: Approximation errors for the example $u(x_1, x_2) = 1/(x_1 + x_2 + 0.1)$, $\boldsymbol{\kappa}(\cdot, s) = 2 + 1/(1+s)$

References

[1] M A Barrientos, G N Gatica, E P Stephan, A mixed finite element method for nonlinear elasticity: two-fold saddle point approach and a-posteriori error estimate, Technical Report 99-25, Departamento de Ingeniería Matemática, Universidad de Concepción, Chile, (1999). http://www.ing-mat.udec.cl/inf-loc-dim.html

[2] F Brezzi, M Fortin, *Mixed and Hybrid Finite Element Methods*, Springer Verlag, (1991).

[3] Z Chen, Expanded mixed finite element methods for linear second-order elliptic problems, *RAIRO Math. Model. Numer. Anal.*, **32**, 479–499. (1998).

[4] Z Chen, Expanded mixed finite element methods for quasilinear second-order elliptic problems, *RAIRO Math. Model. Numer. Anal.*, **32**, 501–520. (1998).

[5] M Costabel, E P Stephan, Coupling of finite element and boundary element methods for an elasto-plastic interface problem, *SIAM J. Numer. Anal.*, **27**, 1212–1226. (1990).

[6] L Franca, A Loula, A new mixed finite element method for the Timoshenko beam problem, *RAIRO Math. Model. Numer. Anal.*, **25**, 561–578. (1991).

[7] G N Gatica, Solvability and Galerkin approximations of a class of nonlinear operator equations, Technical Report 99-03, Departamento de Ingeniería Matemática, Universidad de Concepción, Chile, (1999). http://www.ing-mat.udec.cl/inf-loc-dim.html

[8] G N Gatica, N Heuer, Minimum residual iteration for a dual-dual mixed formulation of exterior transmission problems, Technical Report 99-07, Departamento de Ingeniería Matemática, Universidad de Concepción, Chile (to appear in *Numer. Linear Algebra Appl.* **8**, (2001)), (1999). http://www.ing-mat.udec.cl/inf-loc-dim.html

[9] G N Gatica, N Heuer, An expanded mixed finite element approach via a dual-dual formulation and the minimum residual method, Technical Report 99-08, Departamento de Ingeniería Matemática, Universidad de Concepción, Chile (to appear in *J. Comput. Appl. Math.*), (1999). http://www.ing-mat.udec.cl/inf-loc-dim.html

[10] G N Gatica, N Heuer, Conjugate gradient method for dual-dual mixed formulations, Technical Report 99-16, Departamento de Ingeniería Matemática, Universidad de Concepción, Chile, (1999). http://www.ing-mat.udec.cl/inf-loc-dim.html

[11] G N Gatica, N Heuer, A dual-dual formulation for the coupling of mixed-FEM and BEM in hyperelasticity, *SIAM J. Numer. Anal.*, **38**, 380–400. (2000).

[12] G N Gatica, S Meddahi, A dual-dual mixed formulation for nonlinear exterior transmission problems, Technical Report 99-04, Departamento de Ingeniería Matemática, Universidad de Concepción, Chile (to appear in *Math. Comp.*), (1999). `http://www.ing-mat.udec.cl/inf-loc-dim.html`

[13] G N Gatica, S Meddahi, A fully discrete Galerkin scheme for a two-fold saddle point formulation of an exterior nonlinear problem, Technical Report 00-05, Departamento de Ingeniería Matemática, Universidad de Concepción, Chile, (2000). `http://www.ing-mat.udec.cl/inf-loc-dim.html`

[14] G N Gatica, W L Wendland, Coupling of mixed finite elements and boundary elements for linear and nonlinear elliptic problems, *Appl. Anal.*, **63**, 39–75. (1996).

[15] G N Gatica, W L Wendland, Coupling of mixed finite elements and boundary elements for a hyperelastic interface problem, *SIAM J. Numer. Anal.*, **34**, 2335–2356. (1997).

[16] V Girault, P-A Raviart, *Finite Element Methods for Navier-Stokes Equations. Theory and Algorithms*, Springer Verlag, (1986).

[17] B Heise, Nonlinear field calculations with multigrid-Newton methods, *Impact Comput. Sci. Engrg.*, **5**, 75–110. (1993).

[18] B Heise, Analysis of a fully discrete finite element method for a nonlinear magnetic field problem, *SIAM J. Numer. Anal.*, **31**, 745–759. (1994).

[19] S Meddahi, J Valdés, O Menéndez, P Pérez, On the coupling of boundary integral and mixed finite element methods, *J. Comput. Appl. Math.*, **69**, 113–124. (1996).

[20] J Nečas, *Introduction to the Theory of Nonlinear Elliptic Equations*, John Wiley & Sons, (1986).

[21] E P Stephan, Coupling of finite elements and boundary elements for some nonlinear interface problems, *Comput. Methods Appl. Mech. Engrg.*, **101**, 61–72. (1992).

[22] A Stroud, *Approximate Calculation of Multiple Integrals*, Prentice-Hall, Englewood Cliffs, (1971).

[23] E Zeidler, *Nonlinear Functional Analysis and its Applications IV*, Springer Verlag, New York, (1988).