A posteriori error estimates for primal and mixed finite element approximations of the deformable image registration problem*

Nicolas Barnafi†, Gabriel N. Gatica‡, Daniel E. Hurtado§,
Willian Miranda‡, Ricardo Ruiz-Baier¶

Abstract

In this paper we consider primal and mixed variational formulations that have been recently proposed for the deformable image registration (DIR) problem, and derive reliable and efficient residual-based a posteriori error estimators suitable for adaptive mesh-refinement methods. Our theoretical results, being based on the a posteriori error analysis for the linear elasticity problem with Neumann boundary conditions, make use of the standard tools for that purpose. In particular, these include global inf-sup conditions, Helmholtz decompositions, and the approximation properties of the Raviart-Thomas and Clément interpolants for proving reliability. Localization techniques using bubble functions and inverse inequalities are employed to prove the corresponding efficiency estimates. The adaptive mesh-refinement schemes for the primal and mixed DIR formulations are implemented and tested using synthetic images as well as brain images, and the corresponding numerical results confirm the theoretical properties of the estimators.

Key words: image registration, linear elasticity with weakly imposed symmetry, primal and mixed formulations, finite element methods, a posteriori error analysis.

Mathematics subject classifications (2000): 65N30, 65N15, 68U10, 74B05.

1 Introduction

Deformable image registration (DIR) is the process of aligning a given set of images by means of a transformation that warps one or more of these images. It arises in several applications, particularly in the medical imaging field [30]. Its formulation requires three inputs: a transformation model (composed by a family of mappings that warp the target images), a function that measures the differences between images known as similarity measure and a regularizer that renders the problem

*This work was supported by CONICYT-Chile through the grants FONDECYT Regular 1180832, FONDEF Concurso IDeA en dos etapas ID17I1030, and the project AFB170001 of the PIA program: Concurso Apoyo a Centros Cientificos y Tecnológicos de Excelencia con Financiamiento Basal; and by Centro de Investigación Ingeniería Matemática (CI^{2}MA), Universidad de Concepción.

†MOX - Modellistica e Calcolo Scientifico, Dipartimento di Matematica "F. Brioschi", Politecnico di Milano, via Bonardi 9, 20133 Milano, Italy, email: nicolas.barnafi@polimi.it.

‡CI^{2}MA and Departamento de Ingeniería Matemática, Universidad de Concepción, Casilla 160-C, Concepción, Chile, email: {ggatica,wmiranda}@ci2ma.udec.cl.

§Department of Structural and Geotechnical Engineering, School of Engineering, and Institute for Biological and Medical Engineering, Schools of Engineering, Medicine and Biological Sciences, Pontificia Universidad Católica de Chile, Vicuña Mackenna 4860, Santiago, Chile, email: dhurtado@ing.puc.cl.

¶Mathematical Institute, University of Oxford, A. Wiles Building, Woodstock Road, Oxford OX2 6GG, United Kingdom, email: ruizbaier@maths.ox.ac.uk.

well-posed. In addition to the many variants of these components, different modeling approaches exist, between which we highlight: traditional variational minimization [22, 27], optimal mass transport [21] and level-set modeling [31]. The solution strategy in general considers the incorporation of an auxiliary time variable, which can be seen as a semi-implicit formulation of the proximal point algorithm [29] recently extended to a more general class of proximal operators by using Forward-backward splitting [18]. Also, machine learning techniques have been developed for the solution of this problem, which do not depend on the existence of ground truth solutions and support large deformations [6]. This last work proved competitive against the well established software ANTs [5].

Our work has been mainly motivated by the study of lung regional deformation computed from tomography images of the thorax [12]. The optimal warping u can be interpreted as a displacement field, from which the gradient ∇u can be calculated to obtain the strain tensor and thus give local descriptions in a Continuum Mechanics framework. The study of deformation from one side has revealed the lungs to present a highly heterogeneous and anisotropic behaviour [23], thus providing new deformation-based markers to understand lung disease and injury progression [4, 13]. The proposal of the optical flow formulation by Horn & Schunk [22] lead to a more rigorous mathematical analysis of the DIR problem continuous formulation, which is in contrast with the lack of rigorous numerical analysis of the discrete counterpart, recently developed in the variational formulation [28] in an algorithm-specific fashion and also in the optimal-control setting within a more classical Galerkin framework [26]. The estimation of strain by direct differentiation has been shown to be higly inaccurate [24], which together with the lack of a general discrete analysis motivated the development of mixed finite element schemes for DIR in [7]. This last work used null traction boundary conditions so as to avoid spurious stress, and relied heavily on the mixed theory of linear elasticity problems with pure-traction conditions [20].

The mixed finite element method (MFEM) is a well-established technique which allows to incorporate unknowns of physical interest, such as stress and rotation, and also delivers locking-free schemes in the context of incompressible elasticity (see, e.g. [10, 19]). It also introduces additional difficulties: (i) the new variables increase the dimension of the numerical scheme, making its computational solution more expensive, (ii) the mixed formulation possesses a saddle-point structure, which results in linear systems of equations that are harder to solve numerically and (iii) only discrete spaces that satisfy the inf-sup condition grant a stable scheme, which restricts the choices for low-order approximations and also demands more attention in the formulation of the finite element spaces. For a mixed formulation of DIR with elastic regularization and a target image with Lipschitz gradient, it has been shown that classical existence of solutions is independent of the regularization parameter in the primal case, both primal and mixed give existence and uniqueness for a sufficiently small regularization and PEERS elements, as well as BDM-$\mathbb{P}_0$ for stress-displacement, are inf-sup stable [7]. Furthermore, the drawback described by (iii) is overcome in [7] by using an augmented mixed variational formulation whose discrete analysis does not require the verification of any inf-sup condition, and hence arbitrary finite element subspaces can be employed.

DIR is a highly local problem, which involves many coarse-to-fine type strategies, such as iterated gaussian convolution, image subsampling and total variation smoothing, to name a few [36]. In the context of finite elements, this motivates the use of an a posteriori error estimator [33], which is a global quantity Θ usually decomposed into a sum of local error indicators θ_T that give a norm-wise equivalent of the error. The estimator Θ is said to be reliable (resp. efficient) if there exists $C_{\text{rel}} > 0$ (resp. $C_{\text{eff}} > 0$) independent of the mesh sizes such that

$$C_{\text{eff}}\,\Theta + h.o.t. \leq \|\text{error}\| \leq C_{\text{rel}}\,\Theta + h.o.t.,$$

where $h.o.t.$ is a generic expression denoting one or several terms of higher order. This estimator is

constructed so as to be easily computable, and thus gives a way to choose regions where the error is bigger, so as to perform local refinement. This prevents the refinement of areas where the error is not significant, and so gives an efficient way of error reduction, in contrast to uniform refinement. A posteriori estimates for MFEM in elasticity with weakly imposed symmetry were proposed in [11], and later extended to a residual based estimator in the pure traction case in [16]. Thus, the aim of this work is to propose an a posteriori estimator for the primal registration formulation, as well as extend the current estimators developed for linear elasticity to the mixed formulation developed in [7]. We validate our theoretical results with numerical experiments: one simple analytical case, another one including steep gradients and real data.

Outline

We have organized the contents of this paper as follows. The remainder of this section introduces some standard notations and definition of functional spaces. In Section 2 we recall from [7, Section 2] the model problem. Next, in Section 3 we define the associated continuous primal and mixed variational formulations, and we introduce the corresponding Galerkin schemes. In Section 4 we introduce a posteriori error indicators for both discrete formulations, and assuming small data and certain regularity assumptions, we derive the corresponding theoretical bounds yielding reliability and efficiency of each estimator. Finally, in Section 5 we present numerical examples confirming the reliability and efficiency of the estimators, and illustrating the performance of the associated adaptive algorithms.

Preliminaries

Let us denote $\Omega \subseteq \mathbb{R}^n$, $n \in \{2,3\}$ a given bounded domain with Lipschitz boundary $\Gamma := \partial\Omega$, and denote by $\boldsymbol{n}$ the outward unit normal vector on Γ. Standard notation will be adopted for Lebesgue spaces $L^p(\Omega)$ and Sobolev spaces $H^s(\Omega)$ with norm $\|\cdot\|_{s,\Omega}$ and seminorm $|\cdot|_{s,\Omega}$. Given a scalar space A, we let $\boldsymbol{A}$ and $\mathbb{A}$ be its vectorial and tensor versions, respectively, and $\|\cdot\|$, with no subscripts, will stand for the natural norm of either an element or an operator in any product functional space.

As usual, for any vector field $\boldsymbol{v} = (v_i)_{i=1,n}$, we set the gradient and divergence operator as

$$\nabla \boldsymbol{v} := \left(\frac{\partial v_i}{\partial x_j}\right)_{i,j=1,n} \quad \text{and} \quad \operatorname{div} \boldsymbol{v} := \sum_{j=1}^{n} \frac{\partial v_j}{\partial x_j}.$$

Furthermore, for any tensor fields $\boldsymbol{\tau} = (\tau_{ij})_{i,j=1,n}$ and $\boldsymbol{\zeta} = (\zeta_{ij})_{i,j=1,n}$, we let $\mathbf{div}\,\boldsymbol{\tau}$ be the divergence operator div acting along the rows of $\boldsymbol{\tau}$, and define the transpose, the trace, the tensor inner product, and the deviatoric tensor, respectively, as

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

where $\mathbb{I}$ stands for the identy tensor in $\mathbb{R}^{n\times n}$. Then we recall that the space

$$\mathbb{H}(\mathbf{div};\Omega) := \left\{\boldsymbol{\tau} \in \mathbb{L}^2(\Omega) : \quad \mathbf{div}\,\boldsymbol{\tau} \in \boldsymbol{L}^2(\Omega)\right\},$$

equipped with the usual norm

$$\|\boldsymbol{\tau}\|^2_{\mathbf{div},\Omega} := \|\boldsymbol{\tau}\|^2_{0,\Omega} + \|\mathbf{div}\,\boldsymbol{\tau}\|^2_{0,\Omega},$$

is a Hilbert space. Finally, we employ $\mathbf{0}$ to denote a generic null vector.

2 The model problem

In this section we recall from [7, Section 2] the elastic deformable image registration model. Let $n \in \{2, 3\}$ be the dimension of the images we are interested in analyzing, and let $\Omega \subseteq \mathbb{R}^n$ be a compact domain with Lipschitz boundary $\Gamma := \partial\Omega$. Let $R \in H^1(\Omega)$ be the reference image and $T \in H^1(\widetilde{\Omega})$ be the target image. The DIR problem consists in finding a transformation $\boldsymbol{u} : \Omega \to \mathbb{R}$, also known as the displacement field, that best aligns the images R and T, which is expressed as the variational problem (cf. [27])

$$\inf_{\boldsymbol{u}\in\mathcal{V}} \alpha\mathcal{D}[\boldsymbol{u}; R, T] + \mathcal{S}[\boldsymbol{u}], \tag{2.1}$$

where $\mathcal{V}$ is typically $\boldsymbol{H}^1(\Omega)$, $\mathcal{D} : V \to \mathbb{R}$ is the similarity measure between the images R and T, $\alpha > 0$ is a weighting constant, and $\mathcal{S} : V \to \mathbb{R}$ is a regularization term rendering the problem well-posed. A common choice for the similarity measure is the sum of squares difference, i.e, the L^2 error that takes the form

$$\mathcal{D}[\boldsymbol{u}; R, T] := \frac{1}{2}\int_\Omega (T(\boldsymbol{x} + \boldsymbol{u}(\boldsymbol{x})) - R(\boldsymbol{x}))^2.$$

For the case of elastic DIR, the regularizing term is commonly taken to be the elastic deformation energy, defined by

$$S[\boldsymbol{u}] := \frac{1}{2}\int_\Omega \mathcal{C}\mathbf{e}(\boldsymbol{u}) : \mathbf{e}(\boldsymbol{u}),$$

where

$$\mathbf{e}(\boldsymbol{u}) = \frac{1}{2}\{\nabla\boldsymbol{u} + (\nabla\boldsymbol{u})^{\mathrm{t}}\}$$

is the infinitesimal strain tensor, i.e., the symmetric component of the displacement field gradient, and $\mathcal{C}$ is the elasticity tensor for isotropic solids, that is

$$\mathcal{C}\boldsymbol{\tau} = \lambda\mathrm{tr}(\boldsymbol{\tau})\mathbb{I} + 2\mu\boldsymbol{\tau} \quad \forall\boldsymbol{\tau} \in \mathbb{L}^2(\Omega). \tag{2.2}$$

Assuming that (2.1) has at least one solution with sufficient regularity, the associated Euler-Lagrange equations deliver the following strong problem: Find $\boldsymbol{u}$ such that

$$\begin{aligned} \mathbf{div}(\mathcal{C}\mathbf{e}(\boldsymbol{u})) &= \alpha\boldsymbol{f}_{\boldsymbol{u}} && \text{in} \quad \Omega, \\ \mathcal{C}\mathbf{e}(\boldsymbol{u})\boldsymbol{n} &= \mathbf{0} && \text{on} \quad \partial\Omega, \end{aligned} \tag{2.3}$$

where

$$\boldsymbol{f}_{\boldsymbol{u}}(\boldsymbol{x}) = \big\{T(\boldsymbol{x} + \boldsymbol{u}(\boldsymbol{x})) - R(\boldsymbol{x})\big\}\nabla T(\boldsymbol{x} + \boldsymbol{u}(\boldsymbol{x})) \quad \forall\boldsymbol{x} \in \Omega \ a.e. \tag{2.4}$$

We assume the following conditions on the nonlinear load term $\boldsymbol{f}_{\boldsymbol{u}}$:

$$\begin{aligned} |\boldsymbol{f}_{\boldsymbol{u}}(\boldsymbol{x}) - \boldsymbol{f}_{\boldsymbol{v}}(\boldsymbol{x})| \le L_f|\boldsymbol{u}(\boldsymbol{x}) - \boldsymbol{v}(\boldsymbol{x})| \quad &\forall\,\boldsymbol{x} \in \Omega \ a.e, \\ |\boldsymbol{f}_{\boldsymbol{u}}(\boldsymbol{x})| \le M_f \qquad\qquad\qquad\quad &\forall\,\boldsymbol{x} \in \Omega \ a.e, \end{aligned} \tag{2.5}$$

where L_f and M_f are positive constants.

3 The continuous variational formulations

In this section we introduce the continuous primal and mixed variational formulations of (2.3) derived in [7, Section 3] and [7, Section 4], respectively, and recall the respective solvability results.

3.1 Primal DIR formulation

The primal variational formulation for the registration problem reads: Find $\boldsymbol{u} \in \boldsymbol{H}^1(\Omega)$ such that

$$a(\boldsymbol{u}, \boldsymbol{v}) = \alpha F_{\boldsymbol{u}}(\boldsymbol{v}), \quad \boldsymbol{v} \in \boldsymbol{H}^1(\Omega), \tag{3.1}$$

where $a : \boldsymbol{H}^1(\Omega) \times \boldsymbol{H}^1(\Omega) \to \mathbb{R}$ is the bilinear form defined by

$$a(\boldsymbol{u}, \boldsymbol{v}) := \int_\Omega \mathcal{C}\mathbf{e}(\boldsymbol{u}) : \mathbf{e}(\boldsymbol{v}) \qquad \forall\, \boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{H}^1(\Omega), \tag{3.2}$$

and for every $\boldsymbol{u} \in \boldsymbol{H}^1(\Omega)$, $F_{\boldsymbol{u}} : \boldsymbol{H}^1(\Omega) \to \mathbb{R}$ is the linear functional given by

$$F_{\boldsymbol{u}}(\boldsymbol{v}) := -\int_\Omega \boldsymbol{f}_{\boldsymbol{u}} \cdot \boldsymbol{v} \qquad \forall\, \boldsymbol{v} \in \boldsymbol{H}^1(\Omega).$$

By imposing the conditions (2.5), we can deduce the Lipschitz continuity and uniform boundedness properties for the functional $F_{\boldsymbol{u}}$, that is

$$\|F_{\boldsymbol{u}} - F_{\boldsymbol{v}}\|_{\boldsymbol{H}^1(\Omega)'} \leq L_F \|\boldsymbol{u} - \boldsymbol{v}\|_{0,\Omega} \qquad \forall\, \boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{H}^1(\Omega), \tag{3.3}$$

and

$$\|F_{\boldsymbol{u}}\|_{\boldsymbol{H}^1(\Omega)'} \leq M_F \qquad \forall\, \boldsymbol{u} \in \boldsymbol{H}^1(\Omega),$$

respectively. We recall the results concerning the solvability of (3.1), as developed in [7, Section 3]. First, we define the following partial problem: Given $\boldsymbol{z} \in \boldsymbol{H}^1(\Omega)$, find $\boldsymbol{u} \in \boldsymbol{H}^1(\Omega)$ such that

$$a(\boldsymbol{u}, \boldsymbol{v}) = \alpha F_{\boldsymbol{z}}(\boldsymbol{v}), \quad \boldsymbol{v} \in \boldsymbol{H}^1(\Omega). \tag{3.4}$$

Since this problem does not have unisolvency, we modify it by imposing weak orthogonality to the rigid motions space, denoted by $\mathbb{RM}(\Omega)$ and defined as (see [9, Eq. 11.1.7])

$$\mathbb{RM}(\Omega) := \left\{ \boldsymbol{v} \in \boldsymbol{H}^1(\Omega) : \quad \mathbf{e}(\boldsymbol{v}) = 0 \right\}, \tag{3.5}$$

which guarantees unique solvability of problem (3.4) since $\mathbb{RM}(\Omega)$ is its null space. In fact, defining

$$H := \mathbb{RM}(\Omega)^\perp = \left\{ \boldsymbol{v} \in \boldsymbol{H}^1(\Omega) : \quad \int_\Omega \boldsymbol{v} = \mathbf{0}, \quad \int_\Omega \operatorname{rot} \boldsymbol{v} = 0 \right\},$$

where $\operatorname{rot} \boldsymbol{v} = -\partial v_1/\partial x_2 + \partial v_2/\partial x_1$, for $\boldsymbol{v} = (v_1, v_2)^{\mathrm{t}}$, we consider the following restricted problem: Given $\boldsymbol{z} \in H$, find $\boldsymbol{u} \in H$ such that

$$a(\boldsymbol{u}, \boldsymbol{v}) = \alpha F_{\boldsymbol{z}}(\boldsymbol{v}), \quad \boldsymbol{v} \in H. \tag{3.6}$$

Then, we have the following result:

Theorem 3.1 *Given $\boldsymbol{z} \in H$, the problem (3.6) has a unique solution $\boldsymbol{u} \in H$, and there exists a constant $C_p > 0$ such that*

$$\|\boldsymbol{u}\|_{1,\Omega} \leq \alpha C_p \|F_{\boldsymbol{z}}\|_{\boldsymbol{H}^1(\Omega)'}.$$

Proof. See [7, Theorem 2]. □

We now define the operator $\widehat{\mathbf{T}} : H \to H$ given by $\widehat{\mathbf{T}}(\boldsymbol{z}) = \boldsymbol{u}$, where $\boldsymbol{u}$ is the unique solution to problem (3.6) and thus rewrite (3.1) as the fixed-point equation: Find $\boldsymbol{u} \in H$ such that

$$\widehat{\mathbf{T}}(\boldsymbol{u}) = \boldsymbol{u}. \tag{3.7}$$

The following result establishes the existence of solution to the fixed-point equation (3.7).

Theorem 3.2 *Under data assumptions (2.5), the operator $\widehat{\mathbf{T}}$ has at least one fixed point. Moreover, if $\alpha C_p L_F < 1$, the fixed point is unique.*

Proof. See [7, Theorem 3]. □

3.2 Continuous mixed DIR formulation

In what follows, we introduce a mixed variational formulation of (2.3). We begin by defining an auxiliary field given by the skew symmetric component of the displacement field gradient as

$$\boldsymbol{\rho} := \frac{1}{2}(\nabla \boldsymbol{u} - \nabla \boldsymbol{u}^{\mathrm{t}}).$$

We note that from a continuum mechanics perspective, $\boldsymbol{\rho}$ corresponds to the rotation tensor, which accounts for displacement gradients that do not induce deformation energy. We further define the auxiliary field $\boldsymbol{\sigma} := \mathcal{C}\mathbf{e}(\boldsymbol{u})$. Further, we note that the constitutive relation (2.2) can be inverted (cf. [8] or [19]) as

$$\mathcal{C}^{-1}\boldsymbol{\sigma} = \frac{1}{2\mu}\boldsymbol{\sigma} - \frac{\lambda}{2\mu(2\mu + n\lambda)}\mathrm{tr}(\boldsymbol{\sigma})\mathbb{I}.$$

Then, the strong form of the mixed registration BVP problem of (2.3) becomes: Find $\boldsymbol{u}$, $\boldsymbol{\sigma}$ and $\boldsymbol{\rho}$ such that

$$\begin{aligned} \mathcal{C}^{-1}\boldsymbol{\sigma} &= \nabla \boldsymbol{u} - \boldsymbol{\rho} && \text{in} \quad \Omega, \\ \mathbf{div}(\boldsymbol{\sigma}) &= \alpha \boldsymbol{f_u} && \text{in} \quad \Omega, \\ \boldsymbol{\sigma} &= \boldsymbol{\sigma}^{\mathrm{t}} && \text{in} \quad \Omega, \\ \boldsymbol{\sigma}\boldsymbol{n} &= \mathbf{0} && \text{on} \quad \partial\Omega. \end{aligned} \tag{3.8}$$

Introducing the spaces

$$\mathbb{H}_0(\mathbf{div};\Omega) = \left\{\boldsymbol{\tau} \in \mathbb{H}(\mathbf{div};\Omega) : \ \gamma_{\boldsymbol{n}}\boldsymbol{\tau} = 0\right\},$$

and

$$\boldsymbol{Q} := \boldsymbol{L}^2(\Omega) \times \mathbb{L}^2_{\mathrm{skew}}(\Omega),$$

where

$$\mathbb{L}^2_{\mathrm{skew}}(\Omega) := \{\boldsymbol{\eta} \in \mathbb{L}^2(\Omega) : \ \boldsymbol{\eta}^{\mathrm{t}} = -\boldsymbol{\eta}\},$$

and using a standard integration by parts procedure, the weak formulation of the mixed DIR problem (3.8) reads: Find $(\boldsymbol{\sigma}, (\boldsymbol{u}, \boldsymbol{\rho})) \in \mathbb{H}_0(\mathbf{div};\Omega) \times \boldsymbol{Q}$ such that

$$\begin{aligned} a(\boldsymbol{\sigma}, \boldsymbol{\tau}) + b(\boldsymbol{\tau}, (\boldsymbol{u}, \boldsymbol{\rho})) &= 0 && \forall\, \boldsymbol{\tau} \in \mathbb{H}_0(\mathbf{div};\Omega), \\ b(\boldsymbol{\sigma}, (\boldsymbol{v}, \boldsymbol{\eta})) &= \alpha F_{\boldsymbol{u}}(\boldsymbol{v}, \boldsymbol{\eta}) && \forall\, (\boldsymbol{v}, \boldsymbol{\eta}) \in \boldsymbol{Q}, \end{aligned} \tag{3.9}$$

where $a : \mathbb{H}_0(\mathbf{div};\Omega) \times \mathbb{H}_0(\mathbf{div};\Omega) \to \mathbb{R}$ and $b : \mathbb{H}_0(\mathbf{div};\Omega) \times \boldsymbol{Q} \to \mathbb{R}$ are the bilinear forms defined by

$$a(\boldsymbol{\sigma}, \boldsymbol{\tau}) := \int_\Omega \mathcal{C}^{-1}\boldsymbol{\sigma} : \boldsymbol{\tau} \quad \forall\, \boldsymbol{\sigma}, \boldsymbol{\tau} \in \mathbb{H}_0(\mathbf{div};\Omega), \tag{3.10}$$

and

$$b(\boldsymbol{\tau}, (\boldsymbol{v}, \boldsymbol{\eta})) := \int_\Omega \boldsymbol{v} \cdot \mathbf{div}\boldsymbol{\tau} + \int_\Omega \boldsymbol{\eta} : \boldsymbol{\tau} \quad \forall\, \boldsymbol{\tau} \in \mathbb{H}_0(\mathbf{div};\Omega),\ \forall (\boldsymbol{v}, \boldsymbol{\eta}) \in \boldsymbol{Q}. \tag{3.11}$$

In turn, given $\boldsymbol{u} \in \boldsymbol{L}^2(\Omega)$, $F_{\boldsymbol{u}} : \boldsymbol{Q} \to \mathbb{R}$ is the linear functional defined by

$$F_{\boldsymbol{u}}(\boldsymbol{v}, \boldsymbol{\eta}) := \int_\Omega \boldsymbol{f_u} \cdot \boldsymbol{v} \qquad \forall (\boldsymbol{v}, \boldsymbol{\eta}) \in \boldsymbol{Q}\,.$$

In order to have unisolvency of (3.9), we define the following partial problem: Given $\boldsymbol{z} \in \boldsymbol{L}^2(\Omega)$, find $(\boldsymbol{\sigma}, (\boldsymbol{u}, \boldsymbol{\rho})) \in \mathbb{H}_0(\mathbf{div};\Omega) \times \boldsymbol{Q}$ such that

$$\begin{aligned} a(\boldsymbol{\sigma}, \boldsymbol{\tau}) + b(\boldsymbol{\tau}, (\boldsymbol{u}, \boldsymbol{\rho})) &= 0 && \forall\, \boldsymbol{\tau} \in \mathbb{H}_0(\mathbf{div};\Omega), \\ b(\boldsymbol{\sigma}, (\boldsymbol{v}, \boldsymbol{\eta})) &= \alpha F_{\boldsymbol{z}}(\boldsymbol{v}, \boldsymbol{\eta}) && \forall\, (\boldsymbol{v}, \boldsymbol{\eta}) \in \boldsymbol{Q}, \end{aligned} \tag{3.12}$$

which corresponds to a mixed formulation of the linear elasticity problem with Neumann boundary conditions. Since this problem does not yield unique solvability, we impose weak orthogonality to the rigid motions space $\mathbb{RM}(\Omega)$ (c.f. (3.5)). In this way, defining $\boldsymbol{H} := \mathbb{H}_0(\mathbf{div};\Omega) \times \mathbb{RM}(\Omega)$, we arrive at the following equivalent mixed variational formulation of (3.12): Given $\boldsymbol{z} \in \boldsymbol{L}^2(\Omega)$, find $((\boldsymbol{\sigma},\boldsymbol{\chi}),(\boldsymbol{u},\boldsymbol{\rho})) \in \boldsymbol{H} \times \boldsymbol{Q}$ such that

$$\begin{array}{rcll} A((\boldsymbol{\sigma},\boldsymbol{\chi}),(\boldsymbol{\tau},\boldsymbol{\xi})) + B((\boldsymbol{\tau},\boldsymbol{\xi}),(\boldsymbol{u},\boldsymbol{\rho})) & = & 0 & \forall\,(\boldsymbol{\tau},\boldsymbol{\xi}) \in \boldsymbol{H}, \\ B((\boldsymbol{\sigma},\boldsymbol{\chi}),(\boldsymbol{v},\boldsymbol{\eta})) & = & \alpha F_{\boldsymbol{z}}(\boldsymbol{v},\boldsymbol{\eta}) & \forall\,(\boldsymbol{v},\boldsymbol{\eta}) \in \boldsymbol{Q}, \end{array} \tag{3.13}$$

where $A : \boldsymbol{H} \times \boldsymbol{H} \to \mathbb{R}$ and $B : \boldsymbol{H} \times \boldsymbol{Q} \to \mathbb{R}$ are the bilinear forms given by

$$A((\boldsymbol{\sigma},\boldsymbol{\chi}),(\boldsymbol{\tau},\boldsymbol{\xi})) := a(\boldsymbol{\sigma},\boldsymbol{\tau}) + \int_\Omega \boldsymbol{\chi}\cdot\boldsymbol{\xi} \qquad \forall\,(\boldsymbol{\sigma},\boldsymbol{\chi}),(\boldsymbol{\tau},\boldsymbol{\xi}) \in \boldsymbol{H},$$

$$B((\boldsymbol{\tau},\boldsymbol{\xi}),(\boldsymbol{v},\boldsymbol{\eta})) := b(\boldsymbol{\tau},(\boldsymbol{v},\boldsymbol{\eta})) + \int_\Omega \boldsymbol{\xi}\cdot\boldsymbol{v} \qquad \forall\,((\boldsymbol{\tau},\boldsymbol{\xi}),(\boldsymbol{v},\boldsymbol{\eta})) \in \boldsymbol{H} \times \boldsymbol{Q}.$$

The following two lemmas are needed to establish the well-posedness of (3.13).

Lemma 3.3 *Let* $\boldsymbol{V} := \{(\boldsymbol{\tau},\boldsymbol{\xi}) \in \boldsymbol{H} : B((\boldsymbol{\tau},\boldsymbol{\xi}),(\boldsymbol{v},\boldsymbol{\eta})) = 0,\ \forall (\boldsymbol{v},\boldsymbol{\eta}) \in \boldsymbol{Q}\}$. *Then, there holds* $\boldsymbol{V} = V \times \{\mathbf{0}\}$, *with*

$$V := \{\boldsymbol{\tau} \in \mathbb{H}(\mathbf{div};\Omega) : \mathbf{div}\,\boldsymbol{\tau} = \mathbf{0} \ \textit{and}\ \boldsymbol{\tau} = \boldsymbol{\tau}^{\mathrm{t}} \ \text{in}\ \Omega\}, \tag{3.14}$$

and there exists $\widehat{\alpha} > 0$, *such that*

$$\widehat{\alpha}\|(\boldsymbol{\tau},\boldsymbol{\xi})\|^2_{\boldsymbol{H}} \leq A((\boldsymbol{\tau},\boldsymbol{\xi}),(\boldsymbol{\tau},\boldsymbol{\xi})) \quad \forall\,(\boldsymbol{\tau},\boldsymbol{\xi}) \in \boldsymbol{V}.$$

Proof. See [20, Lemma 3.3]. □

Lemma 3.4 *There exists* $\widehat{\beta} > 0$, *such that*

$$\widehat{\beta}\|(\boldsymbol{v},\boldsymbol{\eta})\|_{\boldsymbol{Q}} \leq \sup_{\substack{(\boldsymbol{\tau},\boldsymbol{\xi})\in\boldsymbol{H}\\(\boldsymbol{\tau},\boldsymbol{\xi})\neq\mathbf{0}}} \frac{|B((\boldsymbol{\tau},\boldsymbol{\xi}),(\boldsymbol{v},\boldsymbol{\eta}))|}{\|(\boldsymbol{\tau},\boldsymbol{\xi})\|_{\boldsymbol{H}}} \quad \forall\,(\boldsymbol{v},\boldsymbol{\eta}) \in \boldsymbol{Q}.$$

Proof. See [20, Lemma 3.4]. □

The well-posedness of the variational formulation (3.13) is stated as follows.

Theorem 3.5 *There exists a unique solution* $((\boldsymbol{\sigma},\boldsymbol{\chi}),(\boldsymbol{u},\boldsymbol{\rho})) \in \boldsymbol{H} \times \boldsymbol{Q}$ *of (3.13). In addition,* $\boldsymbol{\chi} = \mathbf{0}$ *and there exist* $C_m > 0$, *such that*

$$\|((\boldsymbol{\sigma},\boldsymbol{\chi}),(\boldsymbol{u},\boldsymbol{\rho}))\|_{\boldsymbol{H}\times\boldsymbol{Q}} \leq \alpha C_m \|F_{\boldsymbol{z}}\|_{\boldsymbol{Q}'}.$$

Proof. See [20, Theorem 3.1]. □

The above allows us to define a fixed-point operator. Let $\boldsymbol{T} : \boldsymbol{L}^2(\Omega) \to \boldsymbol{L}^2(\Omega)$ given by $\boldsymbol{T}(\boldsymbol{z}) := \boldsymbol{u}$ $\forall \boldsymbol{z} \in \boldsymbol{L}^2(\Omega)$, where $\boldsymbol{u}$ is the displacement component of the unique solution of problem (3.13), and so the mixed formulation (3.9) can be restated as: Find $\boldsymbol{u} \in \boldsymbol{L}^2(\Omega)$ such that

$$\boldsymbol{T}(\boldsymbol{u}) = \boldsymbol{u}\,. \tag{3.15}$$

The following result establishes the existence of solution to the fixed-point problem (3.15):

Theorem 3.6 *Under data conditions (2.5) and assuming $\alpha C_m L_F < 1$, there is a unique fixed point for problem (3.15). With this, the mixed formulation (3.9) has a unique solution $(\boldsymbol{\sigma}, (\boldsymbol{u}, \boldsymbol{\rho})) \in \mathbb{H}_0(\mathbf{div}; \Omega) \times \boldsymbol{Q}$. Moreover, there holds*

$$\|(\boldsymbol{\sigma}, (\boldsymbol{u}, \boldsymbol{\rho}))\|_{\mathbb{H}_0(\mathbf{div};\Omega)\times \boldsymbol{Q}} \leq \alpha C_m M_F.$$

Proof. See [7, Theorem 12]. □

3.3 The primal Galerkin scheme

In the following we formulate a Galerkin scheme to the primal DIR formulation (3.1). Let $\boldsymbol{H}_h$ be a finite dimensional subspace of $\boldsymbol{H}^1(\Omega)$ and define $H_h := \mathbb{RM}^{\perp} \cap \boldsymbol{H}_h$. Then we formulate the nonlinear discrete problem as follows: Find $\boldsymbol{u}_h \in H_h$ such that

$$a(\boldsymbol{u}_h, \boldsymbol{v}_h) = \alpha F_{\boldsymbol{u}_h}(\boldsymbol{v}_h), \quad \boldsymbol{v}_h \in H_h. \tag{3.16}$$

Analogously to the continuous case, we consider the discrete partial problem: Given $\boldsymbol{z}_h \in H_h$, find $\boldsymbol{u}_h \in H_h$ such that

$$a(\boldsymbol{u}_h, \boldsymbol{v}_h) = \alpha F_{\boldsymbol{z}_h}(\boldsymbol{v}_h), \quad \boldsymbol{v}_h \in H_h, \tag{3.17}$$

and also let $T_h : H_h \to H_h$ be the discrete operator given by $T_h(\boldsymbol{z}_h) = \boldsymbol{u}_h$, where $\boldsymbol{u}_h$ is the solution to problem (3.17). Considering the same data assumptions as in the continuous case, as well as the continuity and bound obtained before, we arrive at the following result.

Theorem 3.7 *Assume that data assumptions (2.5) hold. Then, the operator T_h has at least one fixed point. Moreover, if $\alpha C_p L_F < 1$, the fixed point is unique.*

Proof. See [7, Theorem 5]. □

3.4 The mixed Galerkin scheme

In this section we recall the Galerkin scheme for (3.9). First, let $\{\mathcal{T}_h\}_{h>0}$ be a regular family of triangulations of the polygonal region $\bar{\Omega}$ by triangles K of diameter h_K with global mesh size $h := \max\{h_K : K \in \mathcal{T}_h\}$, such that they are quasi-uniform around Γ. Let us consider finite dimensional subspaces $H_h^{\boldsymbol{\sigma}}$, $Q_h^{\boldsymbol{u}}$, and $Q_h^{\boldsymbol{\rho}}$ of $\mathbb{H}(\mathbf{div}; \Omega)$, $\boldsymbol{L}^2(\Omega)$, and $\mathbb{L}^2_{\mathrm{skew}}(\Omega)$, respectively. Then we introduce the product spaces

$$\boldsymbol{H}_h := (H_h^{\boldsymbol{\sigma}} \cap \mathbb{H}_0(\mathbf{div}; \Omega)\,) \times \mathbb{RM}, \qquad \boldsymbol{Q}_h := Q_h^{\boldsymbol{u}} \times Q_h^{\boldsymbol{\rho}},$$

and define the discrete version of (3.13): Given $\boldsymbol{z}_h \in Q_h^{\boldsymbol{u}}$, find $((\boldsymbol{\sigma}_h, \boldsymbol{\chi}_h), (\boldsymbol{u}_h, \boldsymbol{\rho}_h)) \in \boldsymbol{H}_h \times \boldsymbol{Q}_h$ such that

$$\begin{aligned} A((\boldsymbol{\sigma}_h, \boldsymbol{\chi}_h), (\boldsymbol{\tau}_h, \boldsymbol{\xi}_h)) + B((\boldsymbol{\tau}_h, \boldsymbol{\xi}_h), (\boldsymbol{u}_h, \boldsymbol{\rho}_h)) &= 0 && \forall\, (\boldsymbol{\tau}_h, \boldsymbol{\xi}_h) \in \boldsymbol{H}_h, \\ B((\boldsymbol{\sigma}_h, \boldsymbol{\chi}_h), (\boldsymbol{v}_h, \boldsymbol{\eta}_h)) &= \alpha F_{\boldsymbol{z}_h}(\boldsymbol{v}_h, \boldsymbol{\eta}_h) && \forall\, (\boldsymbol{v}_h, \boldsymbol{\eta}_h) \in \boldsymbol{Q}_h. \end{aligned} \tag{3.18}$$

The unique solvability and stability of (3.18), being the Galerkin scheme of a linear elasticity problem with Neumann boundary conditions, has already been established in [20, Theorem 4.1]. This allows us to define the discrete operator $\boldsymbol{T}_h : Q_h^{\boldsymbol{u}} \to Q_h^{\boldsymbol{u}}$ given by $\boldsymbol{T}_h(\boldsymbol{z}_h) := \boldsymbol{u}_h$, where $\boldsymbol{u}_h$ is the displacement component of the unique solution of problem (3.18), and then we rewrite the discrete nonlinear problem as: Find $\boldsymbol{u}_h \in Q_h^{\boldsymbol{u}}$ such that

$$\boldsymbol{T}_h(\boldsymbol{u}_h) = \boldsymbol{u}_h. \tag{3.19}$$

Now we establish the well-posedness of problem (3.19).

Theorem 3.8 *Assuming (2.5) and $\alpha C_m L_F < 1$, the problem (3.19) has a unique solution $\boldsymbol{u}_h \in Q_h^{\boldsymbol{u}}$, which yields $((\boldsymbol{\sigma}_h, \boldsymbol{\chi}_h), (\boldsymbol{u}_h, \boldsymbol{\rho}_h)) \in \boldsymbol{H}_h \times \boldsymbol{Q}_h$ the unique solution of (3.18) with $\boldsymbol{z}_h = \boldsymbol{u}_h$, which satifies*

$$\|((\boldsymbol{\sigma}_h, \boldsymbol{\chi}_h), (\boldsymbol{u}_h, \boldsymbol{\rho}_h))\|_{\boldsymbol{H}\times\boldsymbol{Q}} \leq \alpha C_m M_F.$$

Proof. See [7, Theorem 14]. □

4 Residual-based a posteriori error estimators

In this section we derive a reliable and efficient residual-based a posteriori error estimator for each one of the Galerkin schemes (3.16) and (3.18).

4.1 Preliminaries

We first let $\mathcal{E}_h$ be the set of all edges of the triangulation $\mathcal{T}_h$, and given $K \in \mathcal{T}_h$, we let $\mathcal{E}(K)$ be the set of its edges. Then we decompose $\mathcal{E}_h$ as $\mathcal{E}_h = \mathcal{E}_h(\Omega) \cup \mathcal{E}_h(\Gamma)$, where $\mathcal{E}_h(\Omega) := \{e \in \mathcal{E}_h : e \subseteq \Omega\}$ and $\mathcal{E}_h(\Gamma) := \{e \in \mathcal{E}_h : e \subseteq \Gamma\}$. Further, h_e stands for the length of a given edge e. Also, for each edge $e \in \mathcal{E}_h$ we fix a unit normal vector $\boldsymbol{n}_e := (\boldsymbol{n}_1, \boldsymbol{n}_2)^{\mathrm{t}}$ and let $\boldsymbol{s}_e := (-\boldsymbol{n}_2, \boldsymbol{n}_1)^{\mathrm{t}}$ be the corresponding fixed unit tangential vector along e. However, when no confusion arises, we simple write $\boldsymbol{n}$ and $\boldsymbol{s}$ instead of $\boldsymbol{n}_e$ and $\boldsymbol{s}_e$, respectively. Now, let $\boldsymbol{\tau} \in \mathbb{L}^2(\Omega)$ such that $\boldsymbol{\tau}|_K \in \mathbb{C}(K)$ on each $K \in \mathcal{T}_h$. Then, given $e \in \mathcal{E}_h(\Omega)$, we denote by $[\boldsymbol{\tau}\ \boldsymbol{s}]$ and $[\boldsymbol{\tau}\ \boldsymbol{n}]$ the tangential and normal jumps of $\boldsymbol{\tau}$ across e, that is, $[\boldsymbol{\tau}\ \boldsymbol{s}] := (\boldsymbol{\tau}|_K - \boldsymbol{\tau}|_{K'})|_e \boldsymbol{s}$ and $[\boldsymbol{\tau}\ \boldsymbol{n}] := (\boldsymbol{\tau}|_K - \boldsymbol{\tau}|_{K'})|_e \boldsymbol{n}$, respectively, where K and K' are the triangles of $\mathcal{T}_h$ having e as a common edge. Additionally, given scalar, vector and tensor valued fields v, $\boldsymbol{\varphi} = (\varphi_1, \varphi_2)^{\mathrm{t}}$ and $\boldsymbol{\tau} := (\tau_{ij})_{1\leq i,j\leq 2}$, respectively, we let

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

Next, we collect a few preliminary definitions and results that we need in what follows. Given an integer $k \leq 0$ and $S \subseteq \mathbb{R}^2$, we let $P_k(S)$ be the space of polynomials of degree $\leq k$. Then, we let $I_h : H^1(\Omega) \to X_h$ be the usual Clément interpolation operator (cf. [14]), where

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

The following lemma establishes the local approximation properties of I_h.

Lemma 4.1 *There exist constants $c_1, c_2 > 0$, independent of h, such that for all $v \in H^1(\Omega)$ there holds*

$$\|v - I_h(v)\|_{0,K} \leq c_1 h_K \|v\|_{0,\Delta(K)} \quad \forall\, K \in \mathcal{T}_h$$

$$\|v - I_h(v)\|_{0,e} \leq c_2 h_e^{1/2} \|v\|_{0,\Delta(e)} \quad \forall\, e \in \mathcal{E}_h(\Omega) \cup \mathcal{E}_h(\Gamma),$$

where

$$\Delta(K) := \cup\{K' \in \mathcal{T} : \ K' \cap K \neq \emptyset\} \quad \textit{and} \quad \Delta(e) := \cup\{K' \in \mathcal{T} : \ K' \cap e \neq \emptyset\}.$$

Proof. See [14]. □

The main techniques involved below in the proof of efficiency include the localization technique based on element-bubble and edge-bubble functions. Given $K \in \mathcal{T}_h$ and $e \in \mathcal{E}(K)$, we let ψ_K and ψ_e be the usual triangle-bubble and edge-bubble functions [34, eqs. (1.5)-(1.6)], respectively, which satisfy:

(i) $\psi_K \in P_3(K)$, $\psi_K = 0$ on ∂K, $\mathrm{supp}(\psi_K) \subseteq K$, and $0 \leq \psi_K \leq 1$ in K,

(ii) $\psi_e \in P_2(K)$, $\psi_e = 0$ on ∂K, $\mathrm{supp}(\psi_e) \subseteq \omega_e$, and $0 \leq \psi_e \leq 1$ in ω_e,

where $\omega_e := \cup\{K' \in \mathcal{T}_h : e \in \mathcal{E}(K')\}$. Additional properties of ψ_K and ψ_e are collected in the following lemma (c.f. [32, Lemma 1.3], [34, Section 3.4] or [35, Section 4]).

Lemma 4.2 *Given $k \in \mathbb{N} \cup \{0\}$, there exist positive constants γ_1, γ_2, γ_3, γ_4 and γ_5, depending only on k and the shape regularity of the triangulations, such that for each $K \in \mathcal{T}_h$ and $e \in \mathcal{E}(K)$, there hold*

$$\begin{aligned}
\gamma_1 \|q\|^2_{0,K} &\leq \left\|\psi_K^{1/2} q\right\|^2_{0,K} && \forall\, q \in P_k(K),\\
\|\psi_K q\|_{1,K} &\leq \gamma_2 h_K^{-1} \|q\|_{0,K} && \forall\, q \in P_k(K),\\
\gamma_3 \|p\|^2_{0,e} &\leq \left\|\psi_e^{1/2} p\right\|^2_{0,e} && \forall\, p \in P_k(e),\\
\|\psi_e p\|_{1,\omega_e} &\leq \gamma_4 h_e^{-1/2} \|p\|_{0,e} && \forall\, p \in P_k(e),\\
\|\psi_e p\|_{0,\omega_e} &\leq \gamma_5 h_e^{1/2} \|p\|_{0,e} && \forall\, p \in P_k(e).
\end{aligned} \tag{4.1}$$

4.2 A posteriori error analysis for the primal scheme

In this section we derive a reliable and efficient residual-based a posteriori error estimator for the Galerkin scheme (3.16). We basically follow the approach from [2, 3] (see also the monograph [33]). Letting $\boldsymbol{u}_h \in H_h$ be the unique solution of (3.16), we define for each $K \in \mathcal{T}_h$ the a posteriori error indicator:

$$\begin{aligned}
\Theta_K^2 := h_K^2 \left\|\alpha \boldsymbol{f}_{\boldsymbol{u}_h} - \mathbf{div}(\mathcal{C}\mathbf{e}(\boldsymbol{u}_h))\right\|^2_{0,K} &+ \sum_{e \in \mathcal{E}(K) \cap \mathcal{E}_h(\Omega)} h_e \|[\mathcal{C}\mathbf{e}(\boldsymbol{u}_h)\boldsymbol{n}_e]\|^2_{0,e}\\
&+ \sum_{e \in \mathcal{E}(K) \cap \mathcal{E}_h(\Gamma)} h_e \|\mathcal{C}\mathbf{e}(\boldsymbol{u}_h)\boldsymbol{n}_e\|^2_{0,e},
\end{aligned} \tag{4.2}$$

where, according to (2.4),

$$\boldsymbol{f}_{\boldsymbol{u}_h}|_K(\boldsymbol{x}) := \big\{T(\boldsymbol{x} + \boldsymbol{u}_h(\boldsymbol{x})) - R(\boldsymbol{x})\big\} \nabla T(\boldsymbol{x} + \boldsymbol{u}_h(\boldsymbol{x})) \quad \forall\, \boldsymbol{x} \in K,$$

and introduce the global a posteriori error estimator

$$\Theta := \left\{ \sum_{K \in \mathcal{T}_h} \Theta_K^2 \right\}^{1/2}.$$

The following theorem constitutes the main result of this section.

Theorem 4.3 *Let $\boldsymbol{u} \in H$ and $\boldsymbol{u}_h \in H_h$ be the solutions of the continuous and discrete formulations (3.1) and (3.16), respectively and assume that $\alpha C_p L_F < 1/2$. Then, there exist constants h, C_{rel}, $C_{\mathrm{eff}} > 0$, independent of h, such that for $h \leq h_0$ there holds*

$$C_{\mathrm{eff}} \Theta \leq \|\boldsymbol{u} - \boldsymbol{u}_h\|_H \leq C_{\mathrm{rel}} \Theta. \tag{4.3}$$

The reliability of the global a posteriori error estimator (upper bound in (4.3)) and the corresponding efficiency (lower bound in (4.3)) are established in Sections 4.2.1 and 4.2.2, respectively.

4.2.1 Reliability

The upper bound for (4.3) is established as follows.

Lemma 4.4 *Assume that $\alpha C_p L_F < 1/2$. Then, there exist h_0, $C_{\rm rel} > 0$, independent of h, such that for $h \leq h_0$ there holds*

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_H \leq C_{\rm rel}\, \Theta.$$

Proof. Let us first define

$$\mathcal{R}_h(\boldsymbol{w} - \boldsymbol{w}_h) := \alpha F_{\boldsymbol{u}}(\boldsymbol{w} - \boldsymbol{w}_h) - a(\boldsymbol{u}_h, \boldsymbol{w} - \boldsymbol{w}_h) \quad \forall \boldsymbol{w}_h \in H_h.$$

As a consequence of the ellipticity of a (c.f (3.2)) with ellipticity constant $\bar{\alpha}$ (c.f. [9, Corollary 11.2.22]), we obtain the following condition

$$\bar{\alpha}\|\boldsymbol{v}\|_{1,\Omega} \leq \sup_{\substack{\boldsymbol{w}\in H\\ \boldsymbol{w}\neq \boldsymbol{0}}} \frac{a(\boldsymbol{v},\boldsymbol{w})}{\|\boldsymbol{w}\|_H} \quad \forall \boldsymbol{v} \in H.$$

In particular, for $\boldsymbol{v} = \boldsymbol{u} - \boldsymbol{u}_h \in H$, we notice from (3.1) and (3.16) that $a(\boldsymbol{u} - \boldsymbol{u}_h, \boldsymbol{w}_h) = 0 \ \forall \boldsymbol{w}_h \in H_h$, and hence we obtain $a(\boldsymbol{u} - \boldsymbol{u}_h, \boldsymbol{w}) = a(\boldsymbol{u} - \boldsymbol{u}_h, \boldsymbol{w} - \boldsymbol{w}_h) = \mathcal{R}_h(\boldsymbol{w} - \boldsymbol{w}_h)$, which yields

$$\bar{\alpha}\|\boldsymbol{u} - \boldsymbol{u}_h\|_H \leq \sup_{\substack{\boldsymbol{w}\in H\\ \boldsymbol{w}\neq \boldsymbol{0}}} \frac{\mathcal{R}_h(\boldsymbol{w} - \boldsymbol{w}_h)}{\|\boldsymbol{w}\|_H} \quad \forall \boldsymbol{w}_h \in H_h. \tag{4.4}$$

From the definition of $\mathcal{R}_h(\boldsymbol{w} - \boldsymbol{w}_h)$, integrating by parts on each $K \in \mathcal{T}_h$, and adding and subtracting a suitable term, we can write

$$\begin{aligned}
\mathcal{R}_h(\boldsymbol{w} - \boldsymbol{w}_h) =& \ \alpha F_{\boldsymbol{u}_h}(\boldsymbol{w} - \boldsymbol{w}_h) + \alpha F_{\boldsymbol{u}}(\boldsymbol{w} - \boldsymbol{w}_h) - a(\boldsymbol{u}_h, \boldsymbol{w} - \boldsymbol{w}_h) - \alpha F_{\boldsymbol{u}_h}(\boldsymbol{w} - \boldsymbol{w}_h),\\
=& \ \alpha \left\{F_{\boldsymbol{u}}(\boldsymbol{w} - \boldsymbol{w}_h) - F_{\boldsymbol{u}_h}(\boldsymbol{w} - \boldsymbol{w}_h)\right\} - \alpha \int_\Omega \boldsymbol{f}_{\boldsymbol{u}_h} \cdot (\boldsymbol{w} - \boldsymbol{w}_h)\\
& - \sum_{K\in\mathcal{T}_h} \int_K \mathcal{C}\mathbf{e}(\boldsymbol{u}_h) : \mathbf{e}(\boldsymbol{w} - \boldsymbol{w}_h),\\
=& \ \alpha\{(F_{\boldsymbol{u}} - F_{\boldsymbol{u}_h})(\boldsymbol{w} - \boldsymbol{w}_h)\} - \alpha \int_\Omega \boldsymbol{f}_{\boldsymbol{u}_h} \cdot (\boldsymbol{w} - \boldsymbol{w}_h)\\
& - \sum_{K\in\mathcal{T}_h} \left\{ - \int_K \mathbf{div}(\mathcal{C}\mathbf{e}(\boldsymbol{u}_h)) \cdot (\boldsymbol{w} - \boldsymbol{w}_h) + \int_{\partial K} (\mathcal{C}\mathbf{e}(\boldsymbol{u}_h)\boldsymbol{n}_e) \cdot (\boldsymbol{w} - \boldsymbol{w}_h) \right\},\\
=& \ \alpha\{(F_{\boldsymbol{u}} - F_{\boldsymbol{u}_h})(\boldsymbol{w} - \boldsymbol{w}_h)\} + \sum_{K\in\mathcal{T}_h} \int_K (\mathbf{div}(\mathcal{C}\mathbf{e}(\boldsymbol{u}_h)) - \alpha \boldsymbol{f}_{\boldsymbol{u}_h}) \cdot (\boldsymbol{w} - \boldsymbol{w}_h)\\
& - \sum_{e\in\mathcal{E}_h(\Omega)} \int_e [(\mathcal{C}\mathbf{e}(\boldsymbol{u}_h)\boldsymbol{n}_e)] \cdot (\boldsymbol{w} - \boldsymbol{w}_h) - \sum_{e\in\mathcal{E}_h(\Gamma)} \int_e (\mathcal{C}\mathbf{e}(\boldsymbol{u}_h)\boldsymbol{n}_e) \cdot (\boldsymbol{w} - \boldsymbol{w}_h).
\end{aligned} \tag{4.5}$$

Then, choosing $\boldsymbol{w}_h$ as the Clément interpolant of $\boldsymbol{w}$, that is $\boldsymbol{w}_h := I_h(\boldsymbol{w})$, the approximation properties of I_h (cf. Lemma 4.1) yield

$$\begin{aligned}
\|\boldsymbol{w} - \boldsymbol{w}_h\|_{0,K} \leq& c_1 h_K \|\boldsymbol{w}\|_{1,\Delta(K)},\\
\|\boldsymbol{w} - \boldsymbol{w}_h\|_{0,e} \leq& c_2 h_e \|\boldsymbol{w}\|_{1,\Delta(e)}.
\end{aligned} \tag{4.6}$$

In this way, applying the Cauchy-Schwarz inequality to each term (4.5), and making use of (4.6) together with the Lipschitz continuity of $F_{\boldsymbol{u}}$ (cf. (3.3)), we obtain

$$\begin{aligned}\mathcal{R}_h(\boldsymbol{w}-\boldsymbol{w}_h) \;\leq\; & \alpha c_1 L_F h_K \|\boldsymbol{u}-\boldsymbol{u}_h\|_H \|\boldsymbol{w}\|_{1,\Delta(K)} \\ & + \widehat{C}\left\{\sum_{K\in\mathcal{T}_h}\Theta_K^2\right\}^{1/2}\left\{\sum_{K\in\mathcal{T}_h}\|\boldsymbol{w}\|^2_{1,\Delta(K)} + \sum_{e\in\mathcal{E}_h(\Omega)}\|\boldsymbol{w}\|^2_{1,\Delta(e)}\right\}^{1/2},\end{aligned}$$

where $\widehat{C}$ is a constant depending on c_1 and c_2 and Θ_K^2 defined by (4.2). Additionally using the fact that the number of triangles in $\Delta(K)$ and $\Delta(e)$ are bounded, we have

$$\sum_{K\in\mathcal{T}_h}\|\boldsymbol{w}\|^2_{1,\Delta(K)} \leq C_1\|\boldsymbol{w}\|^2_{1,\Omega} \qquad \text{and} \qquad \sum_{e\in\mathcal{E}_h(\Omega)}\|\boldsymbol{w}\|^2_{1,\Delta(e)} \leq C_2\|\boldsymbol{w}\|^2_{1,\Omega}$$

where C_1, C_2 are positive constant, and using that $\alpha C_p L_F \leq 1/2$, it follows that $h_0 := 1/(2c_1\alpha L_F)$, finally substituting in (4.4), we conclude that

$$\|\boldsymbol{u}-\boldsymbol{u}_h\|_H \leq C_{\rm rel}\,\Theta,$$

where $C_{\rm rel}$ is independent of h. □

4.2.2 Efficiency

Now we focus on establish the lower bound in (4.3). We begin with the following lemma whose proof is a slight modification of [35, Section 6].

Lemma 4.5 *There exist constants $\eta_1,\eta_2,\eta_3>0$, independent of h, but depending on $\gamma_1,\gamma_2,\gamma_3,\gamma_4$ and γ_5 (c.f. (4.1)), such that for each $K\in\mathcal{T}_h$ there holds*

$$\begin{aligned} h_K\left\|\alpha\boldsymbol{f}_{\boldsymbol{u}_h} - \mathbf{div}(\mathcal{C}\mathbf{e}(\boldsymbol{u}_h))\right\|_{0,K} &\leq \eta_1\|\boldsymbol{u}-\boldsymbol{u}_h\|_{0,K}, \\ h_e^{1/2}\|[\mathcal{C}\mathbf{e}(\boldsymbol{u}_h)\cdot\boldsymbol{n}_e]\|_{0,e} &\leq \eta_2\left\{\|\boldsymbol{u}-\boldsymbol{u}_h\|_{0,\omega_e} + \sum_{K\in\omega_e} h_K\|\boldsymbol{u}-\boldsymbol{u}_h\|_{0,K}\right\}, \\ h_e^{1/2}\|\mathcal{C}\mathbf{e}(\boldsymbol{u}_h)\cdot\boldsymbol{n}_e\|_{0,e} &\leq \eta_3\|\boldsymbol{u}-\boldsymbol{u}_h\|_{0,K}, \end{aligned}$$

where $\omega_e := \cup\{K'\in\mathcal{T}_h:\ e\in\mathcal{E}(K')\}$.

Proof. Using the first inequality in (4.1), and let $R_K(\boldsymbol{u}_h) := \alpha\boldsymbol{f}_{\boldsymbol{u}_h} - \mathbf{div}(\mathcal{C}\mathbf{e}(\boldsymbol{u}_h))$ we have

$$\begin{aligned}\|R_K(\boldsymbol{u}_h)\|^2_{0,K} &\leq \gamma_1^{-1}\left\|\psi_K^{1/2}R_K(\boldsymbol{u}_h)\right\|^2_{0,K}, \\ &= \gamma_1^{-1}\int_K \psi_K R_K(\boldsymbol{u}_h)\{\alpha\boldsymbol{f}_{\boldsymbol{u}_h} - \mathbf{div}(\mathcal{C}\mathbf{e}(\boldsymbol{u}_h))\}, \\ &= \gamma_1^{-1}\int_K \alpha\psi_K R_K(\boldsymbol{u}_h)\{\boldsymbol{f}_{\boldsymbol{u}_h} - \boldsymbol{f}_{\boldsymbol{u}}\} - \gamma_1^{-1}\int_K \psi_K R_K(\boldsymbol{u}_h)\{\mathbf{div}(\mathcal{C}\mathbf{e}(\boldsymbol{u}_h) - \mathcal{C}\mathbf{e}(\boldsymbol{u}))\}, \\ &= \gamma_1^{-1}\int_K \alpha\psi_K R_K(\boldsymbol{u}_h)\{\boldsymbol{f}_{\boldsymbol{u}_h} - \boldsymbol{f}_{\boldsymbol{u}}\} + \gamma_1^{-1}\int_K (\mathcal{C}\mathbf{e}(\boldsymbol{u}_h) - \mathcal{C}\mathbf{e}(\boldsymbol{u}))\cdot\nabla(\psi_K R_K(\boldsymbol{u}_h)), \\ &\leq \alpha\gamma_1^{-1}\|R_K(\boldsymbol{u}_h)\|_{0,K}\|\boldsymbol{f}_{\boldsymbol{u}_h} - \boldsymbol{f}_{\boldsymbol{u}}\|_{0,K} + \gamma_1^{-1}\gamma_2 h_K^{-1}\|\mathcal{C}\mathbf{e}(\boldsymbol{u}_h) - \mathcal{C}\mathbf{e}(\boldsymbol{u})\|_{0,K}\|R_K(\boldsymbol{u}_h)\|_{0,K},\end{aligned}$$

where, for the last inequality we used the second inequality in (4.1), also know as inverse inequality. Next, we have

$$h_K\|R_K(\boldsymbol{u}_h)\|_{0,K} \leq \alpha h_K\gamma_1^{-1}\|\boldsymbol{f}_{\boldsymbol{u}_h} - \boldsymbol{f}_{\boldsymbol{u}}\|_{0,K} + \gamma_1^{-1}\gamma_2\|\mathcal{C}\mathbf{e}(\boldsymbol{u}_h) - \mathcal{C}\mathbf{e}(\boldsymbol{u})\|_{0,K},$$

now, using (2.5) and grouping terms, we conclude with $\eta_1 > 0$ independent of h, that

$$h_K\left\|\alpha\boldsymbol{f}_{\boldsymbol{u}_h} - \mathbf{div}(\mathcal{C}\mathbf{e}(\boldsymbol{u}_h))\right\|_{0,K} \leq \eta_1\|\boldsymbol{u} - \boldsymbol{u}_h\|_{0,K},$$

We omit further details, for the remaining inequalities. □

4.3 A posteriori error analysis for the mixed scheme

In this section we derive a reliable and efficient residual-based a posteriori error estimator for the Galerkin scheme (3.18). Throughout the rest of this section we let $((\boldsymbol{\sigma},\boldsymbol{\chi}),(\boldsymbol{u},\boldsymbol{\rho})) \in \boldsymbol{H}\times\boldsymbol{Q}$ and $((\boldsymbol{\sigma}_h,\boldsymbol{\chi}_h),(\boldsymbol{u}_h,\boldsymbol{\rho}_h)) \in \boldsymbol{H}_h\times\boldsymbol{Q}_h$ be the solutions of the continuous and discrete formulations (3.13) and (3.18), respectively. We introduce the global a posteriori error estimator

$$\Psi := \left\{\sum_{K\in\mathcal{T}_h}\Psi_K^2\right\}^{1/2},$$

where we define for each $K\in\mathcal{T}_h$

$$\begin{aligned}\Psi_K^2 := &\left\|\alpha\boldsymbol{f}_{\boldsymbol{u}_h} - \mathbf{div}\boldsymbol{\sigma}_h\right\|_{0,K}^2 + \|\boldsymbol{\sigma}_h - \boldsymbol{\sigma}_h^{\mathrm{t}}\|_{0,K}^2 + \|\boldsymbol{\chi}_h\|_{0,K}^2\\ &+ h_K^2\|\mathrm{curl}(\mathcal{C}^{-1}\boldsymbol{\sigma}_h + \boldsymbol{\rho}_h)\|_{0,K}^2 + h_K^2\|\mathcal{C}^{-1}\boldsymbol{\sigma}_h + \boldsymbol{\rho}_h\|_{0,K}^2\\ &+ \sum_{e\in\mathcal{E}(K)\cap\mathcal{E}_h(\Omega)} h_e\|[(\mathcal{C}^{-1}\boldsymbol{\sigma}_h + \boldsymbol{\rho}_h)\mathbf{s}]\|_{0,e}^2 + \sum_{e\in\mathcal{E}(K)\cap\mathcal{E}_h(\Gamma)} h_e\|(\mathcal{C}^{-1}\boldsymbol{\sigma}_h + \boldsymbol{\rho}_h)\mathbf{s}\|_{0,e}^2.\end{aligned} \tag{4.7}$$

The following theorem constitutes the main result of this section.

Theorem 4.6 *Assume that $\alpha C_m L_F < 1/2$. Then, there exist constants $C_{\mathrm{rel}}, C_{\mathrm{eff}} > 0$, independent of h, such that*

$$C_{\mathrm{eff}}\Psi \leq \|(\boldsymbol{\sigma},\boldsymbol{\chi}) - (\boldsymbol{\sigma}_h,\boldsymbol{\chi}_h)\|_{\boldsymbol{H}} + \|(\boldsymbol{u},\boldsymbol{\rho}) - (\boldsymbol{u}_h,\boldsymbol{\rho}_h)\|_{\boldsymbol{Q}} \leq C_{\mathrm{rel}}\Psi. \tag{4.8}$$

The reliability of the global a posteriori error estimator (upper bound in (4.8)) and the corresponding efficiency (lower bound in (4.8)) are established in Sections 4.3.1 and 4.3.2, respectively.

4.3.1 Reliability

We begin by establishing a more general result due to Lemmas 3.3, 3.4 and Theorem 3.6, and that we will use to establish the upper bound in (4.8). This result we establish in the following theorem.

Theorem 4.7 *Given $\bar{F}\in\boldsymbol{H}'$ and $\bar{G}_{\boldsymbol{u}}\in\boldsymbol{Q}'$, there exists a unique $((\bar{\boldsymbol{\sigma}},\bar{\boldsymbol{\chi}}),(\bar{\boldsymbol{u}},\bar{\boldsymbol{\rho}})) \in \boldsymbol{H}\times\boldsymbol{Q}$ such that*

$$\begin{aligned}A((\bar{\boldsymbol{\sigma}},\bar{\boldsymbol{\chi}}),(\boldsymbol{\tau},\boldsymbol{\xi})) + B((\boldsymbol{\tau},\boldsymbol{\xi}),(\bar{\boldsymbol{u}},\bar{\boldsymbol{\rho}})) &= \bar{F}((\boldsymbol{\tau},\boldsymbol{\xi})) \qquad \forall(\boldsymbol{\tau},\boldsymbol{\xi})\in\boldsymbol{H},\\ B((\bar{\boldsymbol{\sigma}},\bar{\boldsymbol{\chi}}),(\boldsymbol{v},\boldsymbol{\eta})) &= \bar{G}_{\boldsymbol{u}}((\boldsymbol{v},\boldsymbol{\eta})) \qquad \forall(\boldsymbol{v},\boldsymbol{\eta})\in\boldsymbol{Q}.\end{aligned} \tag{4.9}$$

In addition, there exists $C>0$, depending only on $\widehat{\alpha}$, $\widehat{\beta}$, $\|a\|$, and $\|b\|$, such that

$$\|(\bar{\boldsymbol{\sigma}},\bar{\boldsymbol{\chi}})\|_{\boldsymbol{H}} + \|(\bar{\boldsymbol{u}},\bar{\boldsymbol{\rho}})\|_{\boldsymbol{Q}} \leq C\{\|\bar{F}\|_{\boldsymbol{H}'} + \|\bar{G}_{\boldsymbol{u}}\|_{\boldsymbol{Q}'}\}. \tag{4.10}$$

To establish an upper bound for $\|(\boldsymbol{\sigma},\boldsymbol{\chi})-(\boldsymbol{\sigma}_h,\boldsymbol{\chi}_h)\|_{\boldsymbol{H}}$, we consider the functional $S_h:\mathbb{H}(\mathbf{div};\Omega)\to\mathbb{R}$ defined by

$$S_h(\boldsymbol{\tau}):=a(\boldsymbol{\sigma}_h,\boldsymbol{\tau})+b(\boldsymbol{\tau},(\boldsymbol{u}_h,\boldsymbol{\rho}_h)) \qquad \forall\,\boldsymbol{\tau}\in\mathbb{H}(\mathbf{div};\Omega), \tag{4.11}$$

where a and b are the bilinear forms defined in (3.10) and (3.11), respectively, and let $S_h|_V$ be the restriction of S to V, the first component of the kernel $\mathbf{V}$ of B (c.f. (3.14)) We note that $S_h(\boldsymbol{\tau}_h)=0$ for each $\boldsymbol{\tau}_h\in H_h^{\boldsymbol{\sigma}}$.

Now, we make use of a particular problem of the form (4.9) with $\bar{F}\in\boldsymbol{H}'$ and $\bar{G}_{\boldsymbol{u}}\in\boldsymbol{Q}'$ defined by

$$\bar{F}\left((\boldsymbol{\tau},\boldsymbol{\xi})\right):=0\ \ \forall\,(\boldsymbol{\tau},\boldsymbol{\xi})\in\boldsymbol{H} \quad\text{and}\quad \bar{G}_{\boldsymbol{u}}((\boldsymbol{v},\boldsymbol{\eta})):=B((\boldsymbol{\sigma},\boldsymbol{\chi})-(\boldsymbol{\sigma}_h,\boldsymbol{\chi}_h),(\boldsymbol{v},\boldsymbol{\eta}))\ \ \forall\,(\boldsymbol{v},\boldsymbol{\eta})\in\boldsymbol{Q},$$

and let $((\bar{\boldsymbol{\sigma}},\bar{\boldsymbol{\chi}}),(\bar{\boldsymbol{u}},\bar{\boldsymbol{\rho}}))\in\boldsymbol{H}\times\boldsymbol{Q}$ be the unique solution of this particular problem. We note that

$$\bar{G}_{\boldsymbol{u}}((\boldsymbol{v},\boldsymbol{\eta}))=\int_\Omega(\alpha\boldsymbol{f}_{\boldsymbol{u}}-\mathbf{div}\boldsymbol{\sigma}_h)\cdot\boldsymbol{v}-\int_\Omega\boldsymbol{\chi}_h\cdot\boldsymbol{v}-\int_\Omega\boldsymbol{\sigma}_h:\boldsymbol{\eta},$$

this conforming the definition of B and the second equation of (3.13). Adding and subtracting a suitable term we can rewrite the above equation as:

$$\bar{G}_{\boldsymbol{u}}((\boldsymbol{v},\boldsymbol{\eta}))=\int_\Omega\left(\alpha\boldsymbol{f}_{\boldsymbol{u}_h}-\mathbf{div}\boldsymbol{\sigma}_h\right)\cdot\boldsymbol{v}-\int_\Omega\boldsymbol{\chi}_h\cdot\boldsymbol{v}-\int_\Omega\boldsymbol{\sigma}_h:\boldsymbol{\eta}+\alpha\int_\Omega(\boldsymbol{f}_{\boldsymbol{u}}-\boldsymbol{f}_{\boldsymbol{u}_h})\cdot\boldsymbol{v}.$$

Applying Cauchy Schwarz inequality and noting that $\boldsymbol{\sigma}_h:\boldsymbol{\eta}=\frac{1}{2}(\boldsymbol{\sigma}_h-\boldsymbol{\sigma}_h^{\mathrm{t}}):\boldsymbol{\eta}$, together with the condition (2.5), we can establish

$$\|\bar{G}_{\boldsymbol{u}}\|_{\boldsymbol{Q}'}\leq C\left\{\left\|\alpha\boldsymbol{f}_{\boldsymbol{u}_h}-\mathbf{div}\boldsymbol{\sigma}_h\right\|_{0,\Omega}+\|\boldsymbol{\sigma}_h-\boldsymbol{\sigma}_h^{\mathrm{t}}\|_{0,\Omega}+\|\boldsymbol{\chi}_h\|_{0,\Omega}+\alpha L_F\|\boldsymbol{u}-\boldsymbol{u}_h\|_{0,\Omega}\right\},$$

by the previous estimate and the continuous dependence results (4.10), we have

$$\|(\bar{\boldsymbol{\sigma}},\bar{\boldsymbol{\chi}})\|_{\boldsymbol{H}}\leq C\left\{\left\|\alpha\boldsymbol{f}_{\boldsymbol{u}_h}-\mathbf{div}\boldsymbol{\sigma}_h\right\|_{0,\Omega}+\|\boldsymbol{\sigma}_h-\boldsymbol{\sigma}_h^{\mathrm{t}}\|_{0,\Omega}+\|\boldsymbol{\chi}_h\|_{0,\Omega}+\alpha L_F\|\boldsymbol{u}-\boldsymbol{u}_h\|_{0,\Omega}\right\}. \tag{4.12}$$

Now, applying the triangle inequality we obtain

$$\|(\boldsymbol{\sigma},\boldsymbol{\chi})-(\boldsymbol{\sigma}_h,\boldsymbol{\chi}_h)\|_{\boldsymbol{H}}\leq\|(\boldsymbol{\sigma},\boldsymbol{\chi})-(\boldsymbol{\sigma}_h,\boldsymbol{\chi}_h)-(\bar{\boldsymbol{\sigma}},\bar{\boldsymbol{\chi}})\|_{\boldsymbol{H}}+\|(\bar{\boldsymbol{\sigma}},\bar{\boldsymbol{\chi}})\|_{\boldsymbol{H}}, \tag{4.13}$$

and hence, it remains to estimate $\|(\boldsymbol{\sigma},\boldsymbol{\chi})-(\boldsymbol{\sigma}_h,\boldsymbol{\chi}_h)-(\bar{\boldsymbol{\sigma}},\bar{\boldsymbol{\chi}})\|_{\boldsymbol{H}}$. First observe that $(\boldsymbol{\sigma},\boldsymbol{\chi})-(\boldsymbol{\sigma}_h,\boldsymbol{\chi}_h)-(\bar{\boldsymbol{\sigma}},\bar{\boldsymbol{\chi}})\in\boldsymbol{V}$, hence applying the ellipticity of A in $\boldsymbol{V}$ (cf. Lemma 3.3) and analogously to [16, Lemma 4.6], we obtain an estimate for this term that replacing together with (4.12) in (4.13), allows us to establish that

$$\begin{aligned}\|(\boldsymbol{\sigma},\boldsymbol{\chi})-(\boldsymbol{\sigma}_h,\boldsymbol{\chi}_h)\|_{\boldsymbol{H}}\leq C\Big\{&\|S_h|_V\|_{V'}+\left\|\alpha\boldsymbol{f}_{\boldsymbol{u}_h}-\mathbf{div}\boldsymbol{\sigma}_h\right\|_{0,\Omega}\\&+\|\boldsymbol{\sigma}_h-\boldsymbol{\sigma}_h^{\mathrm{t}}\|_{0,\Omega}+\|\boldsymbol{\chi}_h\|_{0,\Omega}+\alpha L_F\|\boldsymbol{u}-\boldsymbol{u}_h\|_{0,\Omega}\Big\}.\end{aligned} \tag{4.14}$$

To estimate $\|S_h|_V\|_{V'}$, (cf. (4.11)) in (4.14), we have the following result

Lemma 4.8 *There exists $C>0$, such that*

$$\begin{aligned}\|S_h|_V\|_{V'}\ \leq C\Bigg\{&h_K^2\|\mathbf{curl}(\mathcal{C}^{-1}\boldsymbol{\sigma}_h+\boldsymbol{\rho}_h)\|_{0,K}^2+\sum_{e\in\mathcal{E}(K)\cap\mathcal{E}_h(\Omega)}h_e\|[(\mathcal{C}^{-1}\boldsymbol{\sigma}_h+\boldsymbol{\rho}_h)\boldsymbol{s}]\|_{0,e}^2\\&+\sum_{e\in\mathcal{E}(K)\cap\mathcal{E}_h(\Gamma)}h_e\|(\mathcal{C}^{-1}\boldsymbol{\sigma}_h+\boldsymbol{\rho}_h)\boldsymbol{s}\|_{0,e}^2\Bigg\}.\end{aligned} \tag{4.15}$$

Proof. See [16, Lemma 4.7] for details. □

From the above the following lemma is configured.

Lemma 4.9 *Assume that $\alpha C_m L_F < 1/2$. Then, there exists $C > 0$ such that*

$$\|(\boldsymbol{\sigma}, \boldsymbol{\chi}) - (\boldsymbol{\sigma}_h, \boldsymbol{\chi}_h)\|_{\boldsymbol{H}} \leq C \left\{ \sum_{K \in \mathcal{T}_h} \widetilde{\Psi}_K^2 \right\}^{1/2},$$

where

$$\begin{aligned} \widetilde{\Psi}_K^2 \; := \; & h_K^2 \|\mathrm{curl}(\mathcal{C}^{-1}\boldsymbol{\sigma}_h + \boldsymbol{\rho}_h)\|_{0,K}^2 + \sum_{e \in \mathcal{E}(K) \cap \mathcal{E}_h(\Omega)} h_e \|[(\mathcal{C}^{-1}\boldsymbol{\sigma}_h + \boldsymbol{\rho}_h)\boldsymbol{s}]\|_{0,e}^2 \\ & + \sum_{e \in \mathcal{E}(K) \cap \mathcal{E}_h(\Gamma)} h_e \|(\mathcal{C}^{-1}\boldsymbol{\sigma}_h + \boldsymbol{\rho}_h)\boldsymbol{s}\|_{0,e}^2 + \|\alpha \boldsymbol{f}_{\boldsymbol{u}} - \mathbf{div}\, \boldsymbol{\sigma}_h\|_{0,\Omega} \\ & + \|\boldsymbol{\sigma}_h - \boldsymbol{\sigma}_h^{\mathrm{t}}\|_{0,\Omega} + \|\boldsymbol{\chi}_h\|_{0,\Omega} + \alpha L_F \|\boldsymbol{u} - \boldsymbol{u}_h\|_{0,\Omega}. \end{aligned}$$

Proof. It follows straightforwardly from (4.14) and (4.15). □

Now we proceed to obtain the corresponding upper bound for $\|(\boldsymbol{u}, \boldsymbol{\rho}) - (\boldsymbol{u}_h, \boldsymbol{\rho}_h)\|_{\boldsymbol{Q}}$. More precisely, we have the following result.

Lemma 4.10 *Assume that $\alpha C_m L_F < 1/2$. Then, there exists $C > 0$ such that*

$$\|(\boldsymbol{u}, \boldsymbol{\rho}) - (\boldsymbol{u}_h, \boldsymbol{\rho}_h)\|_{\boldsymbol{Q}} \leq C \left\{ \sum_{K \in \mathcal{T}_h} \Psi_K^2 \right\}^{1/2},$$

where Ψ_K^2 is the local indicator defined in (4.7).

Proof. The proof follows directly from [16, Lemma 4.9] with small modifications. □

The reliability of Ψ, is a straightforward consequence of the Lemmas 4.9 and 4.10, under the assumption that $\alpha C_m L_F < 1/2$.

4.3.2 Efficiency

In this section, we focus on the efficiency of our a posteriori error estimator Ψ and provide upper bounds depending on the actual errors for the seven terms defining the local indicator Ψ_K^2 (c.f. (4.7)). For this, analogously to [16, Section 4.3] we begin with the first three ones appearing there, more precisely, since $\mathbf{div}(\boldsymbol{\sigma}) = \alpha \boldsymbol{f}_{\boldsymbol{u}}$ in Ω, we have that

$$\|\alpha \boldsymbol{f}_{\boldsymbol{u}} - \mathbf{div}\boldsymbol{\sigma}_h\|_{0,K}^2 \leq \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\mathbf{div},K}^2.$$

Next, adding and subtracting $\boldsymbol{\sigma}$, and we use that $\boldsymbol{\sigma} = \boldsymbol{\sigma}^{\mathrm{t}}$ in Ω, we see that

$$\|\boldsymbol{\sigma}_h - \boldsymbol{\sigma}_h^{\mathrm{t}}\|_{0,K}^2 \leq 4\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,K}^2.$$

Finally, since $\boldsymbol{\chi} = \mathbf{0}$, we obtain

$$\|\boldsymbol{\chi}_h\|_{0,K}^2 = \|\boldsymbol{\chi} - \boldsymbol{\chi}_h\|_{0,K}^2.$$

The upper bounds for the terms involving only the tensor $\mathcal{C}^{-1}\boldsymbol{\sigma}_h + \boldsymbol{\rho}_h$, are established in the following result.

Lemma 4.11 *There exist constants $C_1, C_2, C_3, C_4 > 0$, independents of h, such that for each $K \in \mathcal{T}_h$ there holds*

$$h_K^2 \|\mathrm{curl}(\mathcal{C}^{-1}\boldsymbol{\sigma}_h + \boldsymbol{\rho}_h)\|_{0,K}^2 \leq C_1 \left\{ \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,K}^2 + \|\boldsymbol{\rho} - \boldsymbol{\rho}_h\|_{0,K}^2 \right\}$$

$$\|\mathcal{C}^{-1}\boldsymbol{\sigma}_h + \boldsymbol{\rho}_h\|_{0,K}^2 \leq C_2 \left\{ \|\boldsymbol{u} - \boldsymbol{u}_h\|_{0,K}^2 + h_K^2 \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,K}^2 + h_K^2 \|\boldsymbol{\rho} - \boldsymbol{\rho}_h\|_{0,K}^2 \right\}$$

$$h_e \|[(\mathcal{C}^{-1}\boldsymbol{\sigma}_h + \boldsymbol{\rho}_h)\boldsymbol{s}]\|_{0,e}^2 \leq C_3 \sum_{K \subseteq \omega_e} \left\{ \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,K}^2 + \|\boldsymbol{\rho} - \boldsymbol{\rho}_h\|_{0,K}^2 \right\}$$

$$\sum_{e \in \mathcal{E}_h(\Gamma)} h_e \|(\mathcal{C}^{-1}\boldsymbol{\sigma}_h + \boldsymbol{\rho}_h)\boldsymbol{s}\|_{0,e}^2 \leq C_4 \sum_{e \in \mathcal{E}_h(\Gamma)} \left\{ \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,K}^2 + \|\boldsymbol{\rho} - \boldsymbol{\rho}_h\|_{0,K}^2 \right\}.$$

where $\omega_e := \cup \{K' \in \mathcal{T}_h : \ e \in \mathcal{E}(K')\}$.

Proof. See [16, Section 4.3]. □

5 Numerical results

5.1 Preliminaries and implementation details

We now turn to the implementation of some numerical tests that confirm the predicted reliability and efficiency of the proposed a posteriori error estimators (4.2) and (4.7). The DIR problem is in all cases restricted to images mapped to the unit square $\Omega = (0,1)^2$, and uniform triangular partitions are employed for all initial meshes. The discretization of the primal problem is done with continuous piecewise linear and continuous piecewise quadratic approximations for the displacement vector field. For the case of the mixed formulation, we consider the lowest-order family of Brezzi-Douglas-Marini elements for the rows of the Cauchy stress tensor, and piecewise constant approximations of the entries of the displacement vector and the rotation tensor. The Picard method is used to linearize the problem and we set a fixed tolerance of 1e-5 on the energy norm of the difference between two consecutive solutions. Unless otherwise specified, all linear solves related to the fixed-point iteration (in both primal and mixed formulations) are carried out with the stabilized bi-conjugated gradient method (BiCGStab) using an incomplete LU decomposition as preconditioner.

Mesh adaptation guided by the a posteriori error estimators proposed for the primal and mixed methods is carried out by a classical conforming partitioning. No coarsening is applied (mainly due to the capabilities of the current version of the finite element library we use herein [1]). After computing locally the a posteriori error indicators, we proceed to tag elements for refinement using the Dörfler strategy [17], where we mark sufficiently many elements so that one establishes equi-distribution of the error indicator mass, and then the diameter of each triangle in the new adapted mesh (contained in a generic element K on the initial grid) is set proportional to the diameter of the initial element times the ratio $\bar{\zeta}_h/\zeta_K$, where $\bar{\zeta}_h$ is the mean value of a generic error estimator ζ over the initial mesh (see for instance, [32]). In each of the accuracy tests below (Examples 1 and 2), these ratios are multiplied by a constant γ_{ratio} that is arbitrarily chosen so as to generate either a roughly similar number of degrees of freedom, or similar individual error magnitudes than in the case of uniform refinement. However the density of the refinement process is tuned at will.

Let us also recall from [7] that the implementation of the fixed-point scheme includes an additional stabilization term associated with dynamic gradient flows, that essentially translates in having a pseudo-time step in the Euler-Lagrange equations (2.3), that then read: knowing $\boldsymbol{u}^k$, for $k = 1, \ldots,$

(a) Primal method, uniform refinement

k	DoF	h	$\mathtt{e}_{\boldsymbol{u}}$	rate	iter
1	21	0.7071	2.050e-01	–	3
	53	0.3536	1.389e-01	0.561	4
	165	0.1768	7.283e-02	0.931	4
	581	0.0884	3.360e-02	1.116	4
	2181	0.0442	1.587e-02	1.082	4
	8453	0.0221	7.774e-03	1.030	4
2	53	0.7071	7.094e-02	–	3
	165	0.3536	2.261e-02	1.649	4
	581	0.1768	5.872e-03	1.945	4
	2181	0.0884	1.443e-03	2.025	4
	8453	0.0442	3.530e-04	2.031	4
	33285	0.0221	8.577e-05	2.041	4

(b) Primal method, adaptive refinement

k	DoF	$h_{\min}$	$\mathtt{e}_{\boldsymbol{u}}$	$\widehat{\mathtt{rate}}$	$\mathtt{eff}(\Theta)$	iter
1	21	0.7071	2.050e-01	–	0.6797	3
	53	0.3536	1.219e-01	1.123	0.6984	4
	165	0.1768	6.440e-02	1.123	0.7083	4
	557	0.0884	3.252e-02	1.123	0.6964	4
	2101	0.0442	1.629e-02	1.041	0.6962	4
	8149	0.0221	8.150e-03	1.022	0.6950	4
2	53	0.7071	7.094e-02	–	0.2559	3
	165	0.3536	2.061e-02	2.177	0.3599	4
	581	0.1768	5.756e-03	2.026	0.3675	4
	2181	0.0884	1.446e-03	2.088	0.3602	4
	8453	0.0442	3.620e-04	2.045	0.3546	4
	32933	0.0221	8.914e-05	2.061	0.3457	4

Table 5.1: Example 1A. Errors, convergence rates, and Picard iteration count for the approximate displacements $\boldsymbol{u}_h$ produced with the primal method (of degrees k=1 and k=2); and tabulated according to the resolution level, under uniform (a) and adaptive mesh refinement guided by Θ, with $\gamma_{\text{ratio}} = 0.1$ (b, also displaying the rescaled effectivity index).

solve

$$\frac{\boldsymbol{u}^{k+1}}{\delta t} - \mathbf{div}(\mathcal{C}\mathbf{e}(\boldsymbol{u}^{k+1})) = \frac{\boldsymbol{u}^{k}}{\delta t} - \alpha \boldsymbol{f}_{\boldsymbol{u}^k}.$$

Further details can be found in [7, Appendix C]. Therefore the primal and mixed weak formulations, the corresponding Galerkin methods, as well as the a posteriori error indicators Θ and Ψ are modified accordingly, and only affecting the residual terms associated with the momentum equation. The Picard iterations with pseudo time-stepping are located inside the adaptive refinement loop which consists in solving, estimating, marking and refining.

(a) Mixed method, uniform refinement

DoF	h	$\mathtt{e}_{\boldsymbol{\sigma}}$	rate	$\mathtt{e}_{\boldsymbol{u}}$	rate	$\mathtt{e}_{\boldsymbol{\rho}}$	rate	iter
91	0.7071	304.821	–	3.521e-02	–	1.337e-01	–	6
323	0.3536	150.763	1.016	1.845e-02	0.932	6.126e-02	1.125	7
1219	0.1768	70.2653	1.101	9.260e-03	0.994	2.961e-02	1.049	8
4739	0.0884	33.9119	1.051	4.629e-03	1.000	1.463e-02	1.017	8
18691	0.0442	16.7754	1.015	2.314e-03	1.000	7.284e-03	1.006	8
74243	0.0221	8.36553	1.004	1.157e-03	1.000	3.637e-03	1.002	10

(b) Mixed method, adaptive refinement

DoF	$h_{\min}$	$\mathtt{e}_{\boldsymbol{\sigma}}$	$\widehat{\mathtt{rate}}$	$\mathtt{e}_{\boldsymbol{u}}$	$\widehat{\mathtt{rate}}$	$\mathtt{e}_{\boldsymbol{\rho}}$	$\widehat{\mathtt{rate}}$	eff(Ψ)	iter
91	0.7071	304.821	–	3.521e-02	–	1.337e-01	–	0.4427	6
323	0.3536	138.151	1.249	1.842e-02	1.023	6.212e-02	1.210	0.4266	8
1219	0.1768	70.8962	1.005	9.252e-03	1.037	2.964e-02	1.114	0.4264	8
4128	0.0884	36.1260	1.105	4.896e-03	1.044	1.873e-02	0.752	0.4288	8
5902	0.0442	32.3187	0.623	4.410e-03	0.839	1.400e-02	1.628	0.4372	8
15846	0.0313	18.3526	1.146	2.497e-03	1.152	1.010e-02	0.661	0.4368	8
19534	0.0156	16.8697	0.805	2.316e-03	0.727	8.120e-03	2.085	0.4355	8
41440	0.0110	12.8837	0.716	1.764e-03	0.724	6.499e-03	0.592	0.4316	9

Table 5.2: Example 1B. Errors, convergence rates, and Picard iteration count for the approximate Cauchy stress, displacements, and rotation $\boldsymbol{\sigma}_h, \boldsymbol{u}_h, \boldsymbol{\rho}_h$ produced with our mixed method; and tabulated according to the resolution level, under uniform (a) and adaptive mesh refinement guided by Ψ, with $\gamma_{\text{ratio}} = 0.275$ (sub-table b, also displaying the effectivity index).

5.2 Example 1: errors with respect to smooth solutions

First we assess the accuracy of the primal and mixed schemes using the following closed form exact solutions to (2.3) defined on the unit square

$$\boldsymbol{u}(x_1,x_2) = \begin{pmatrix} 0.1\cos(\pi x_1)\sin(\pi x_2) + \frac{x_1^2(1-x_1)^2x_2^2(1-x_2)^2}{2\lambda} \\ -0.1\sin(\pi x_1)\cos(\pi x_2) + \frac{x_1^3(1-x_1)^3x_2^3(1-x_2)^3}{2\lambda} \end{pmatrix},$$
$$\boldsymbol{\sigma}(x_1,x_2) = \mathcal{C}\mathbf{e}(\boldsymbol{u}), \quad \boldsymbol{\rho}(x_1,x_2) = \frac{1}{2}(\nabla\boldsymbol{u} - \nabla\boldsymbol{u}^{\mathrm{t}}),$$

for a synthetic smooth reference image defined by $R(x_1,x_2) = \sin(2\pi x_1)\sin(2\pi x_2)$, and constructing a synthetic target image simply via the composition of the reference and the inverse warping, $T = R\circ(\boldsymbol{id}+\boldsymbol{u})^{-1}$. An initial target in the fixed-point scheme is a perturbation of the reference image, that is $T_0(x_1,x_2) = \sin(2\pi x_1)\sin(2\pi[x_2+0.01])$. These exact solutions satisfy the zero-traction boundary condition, and they are used to construct an additional body load (apart from $\boldsymbol{f}_{\boldsymbol{u}}$) that needs to be incorporated as right-hand side in the discrete problems, as well as in the residual term associated with the momentum conservation equation in the definition of the error indicators. The model parameters employed in this test are Young modulus $E = 1000$, Poisson ratio $\nu = 0.4$ (used to obtain the Lamé constants of the solid, $\lambda = \frac{E\nu}{(1+\nu)(1-2\nu)}$ and $\mu = \frac{E}{2+2\nu}$), a weight constant $\alpha = 100$, and pseudo time-step $\delta t = \alpha^{-2}$.

On sequences of uniformly or adaptively refined meshes, we solve the DIR problem with primal and mixed methods and compute (non-normalized) errors between the approximate and exact solutions in

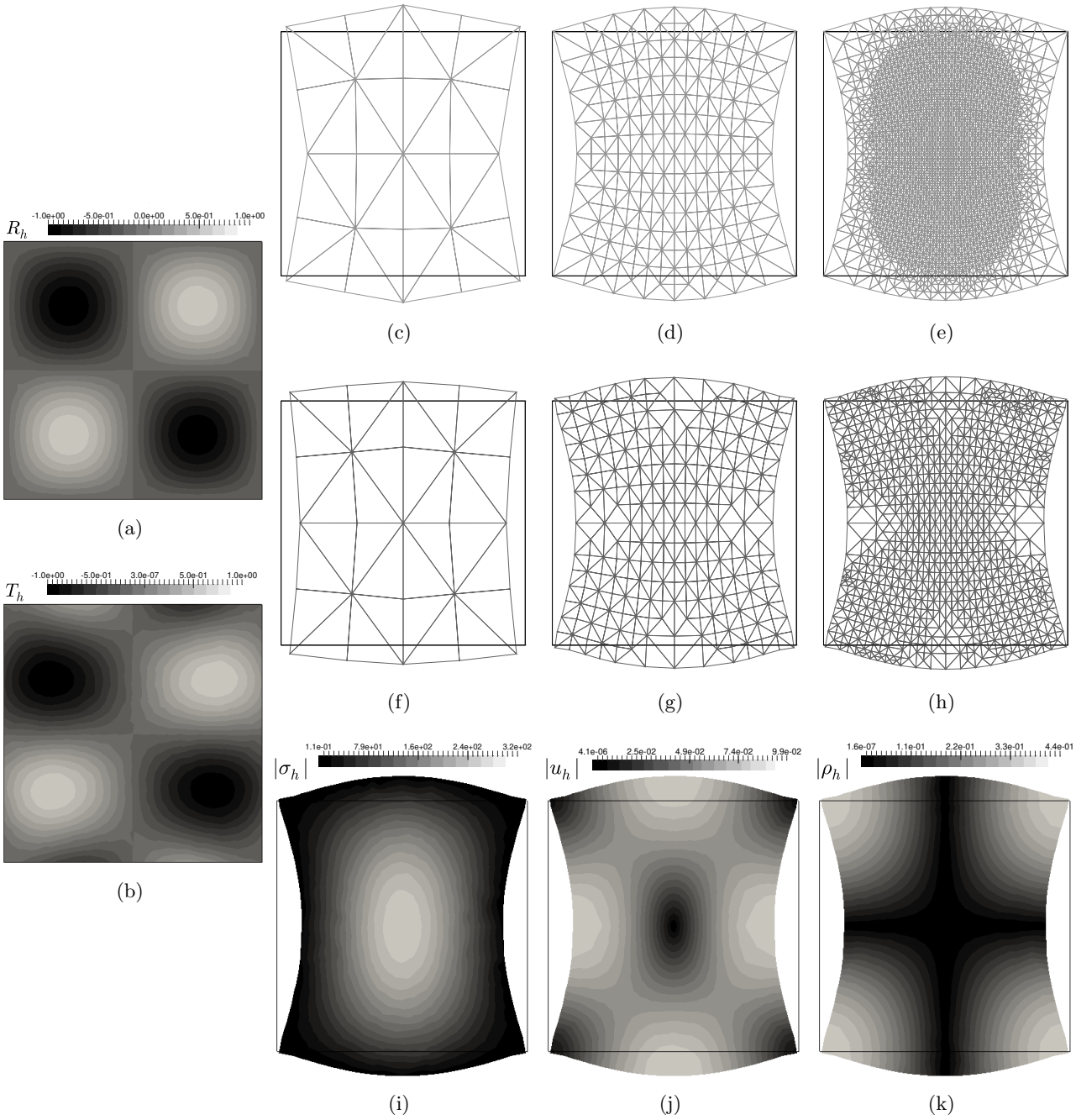


Figure 5.1: Example 1. Projected fields of the reference R and composed $T(\boldsymbol{x}+\boldsymbol{u}_h(\boldsymbol{x}))$ images (a,b); sample meshes refined according to the indicator Θ for the primal method (c,d,e), according to Ψ after solving the problem with the mixed method (f,g,h); and approximate solutions computed with the mixed method (i,j,k).

their natural norms, that is, for the primal method $\mathtt{e}_{\boldsymbol{u}} = \|\boldsymbol{u}-\boldsymbol{u}_h\|_{1,\Omega}$; whereas for the mixed method $\mathtt{e}_{\boldsymbol{u}} = \|\boldsymbol{u}-\boldsymbol{u}_h\|_{0,\Omega}$ and $\mathtt{e}_{\boldsymbol{\rho}} = \|\boldsymbol{\rho}-\boldsymbol{\rho}_h\|_{0,\Omega}$, $\mathtt{e}_{\boldsymbol{\sigma}} = \|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{\mathbf{div},\Omega}$. We also point out that in the case of adaptive mesh refinement, the experimental rates of convergence $\widehat{\mathtt{rate}}$ are computed differently than

in the uniform case

$$\texttt{rate} = \log(\texttt{e}/\widehat{\texttt{e}})[\log(h/\widehat{h})]^{-1}, \qquad \widehat{\texttt{rate}} = -2\log(\texttt{e}/\widehat{\texttt{e}})[\log(\texttt{DoF}/\widehat{\texttt{DoF}})]^{-1},$$

where $\texttt{e}$ and $\widehat{\texttt{e}}$ denote errors produced on two consecutive meshes. These grids have respective mesh sizes h and h' (needed to compute the experimental order of convergence $\texttt{rate}$), or they are associated with $\texttt{DoF}$ and $\widehat{\texttt{DoF}}$ degrees of freedom, respectively (in when computing $\widehat{\texttt{rate}}$). In addition, the effectivity index associated with the global estimators proposed for the primal and mixed discretizations is computed as

$$\texttt{eff}(\Theta) = \frac{\lambda \texttt{e}_{\boldsymbol{u}}}{\Theta}, \qquad \texttt{eff}(\Psi) = \frac{\left\{\texttt{e}_{\boldsymbol{\sigma}}^2 + \texttt{e}_{\boldsymbol{u}}^2 + \texttt{e}_{\boldsymbol{\rho}}^2\right\}^{1/2}}{\Psi},$$

where the additional scaling (in this case with the dilation modulus λ) for the indicator Θ is motivated by the fact that the efficiency bound arising from the proof of Lemma 4.5 is proportional to λ due to the definition of the Hooke tensor $\mathcal{C}$. Such an explicit scaling is however not required for the a posteriori estimation in the mixed method.

Detailed information on the convergence of the primal scheme under uniform and adaptive mesh refinement is collected in Table 5.1. Apart from errors and convergence rates that indicate optimal convergence with $O(h^k)$, we also show the number of Picard iterations required to reach the desired tolerance, where we remark that very little differences are observed in the two cases. Similarly, the convergence history of the lowest-order mixed method is given in Table 5.2. The effectivity index for the adaptive primal scheme has a roughly constant value of 0.69 for piecewise linear displacements and roughly 0.36 for the second-order method. The values for the mixed scheme are also maintained roughly constant to 0.43.

In Figure 5.1(a,b) we portray the projected reference image R_h and the composed image $T_h = T(\boldsymbol{x} + \boldsymbol{u}_h(\boldsymbol{x}))$ on the undeformed domain, and the panels (c-h) show examples of adaptively refined meshes resulting from the estimators for the primal and mixed approximations. For instance, the primal method refines largely near the domain center. We also display in panels (i,j,k) the approximate solutions (the Frobenius norm of stress, displacement magnitude, and Frobenius norm of the rotation matrix) generated with the mixed method at the final refinement level.

5.3 Example 2: convergence to solutions with higher gradients

Next we slightly modify the closed-form displacement and impose a higher gradient on the reference image and initial target image (and we remark that typical benchmark tests in this context would be L-shaped or other non-convex geometries, but they lack a strong motivation in view of the application of the present methods in solving DIR problems). The model parameters are kept as in Example 1, however the transformation used to construct the target image is now

$$\boldsymbol{u}(x_1, x_2) = \begin{pmatrix} 0.1\cos(\pi x_1)\sin(\pi x_2) + \frac{x_1^2(1-x_1)^2 x_2^2(1-x_2)^2}{2} \\ -0.1\sin(\pi x_1)\cos(\pi x_2) + \frac{x_1^3(1-x_1)^3 x_2^3(1-x_2)^3}{2} \end{pmatrix},$$

and the synthetic reference image and initial target image are, respectively

$$R(x_1, x_2) = \frac{x_1 x_2 (x_1 - 1)(x_2 - 1)}{(x_1 + 0.01)^4 + (x_2 + 0.01)^4}, \qquad T_0(x_1, x_2) = e^{-50[(x_1-0.2)^2 + (x_2-0.2)^2]}.$$

We perform again a series of convergence tests whose results are reported in Table 5.3(a) and Table 5.4(a). An consequence of these higher gradients is that optimal convergence for all the fields is

(a) Primal method, uniform refinement

k	DoF	h	$\mathtt{e}_{\boldsymbol{u}}$	rate	iter
1	21	0.7071	6.390e-01	–	3
	53	0.3536	4.578e-01	0.481	4
	165	0.1768	3.180e-01	0.526	6
	581	0.0884	1.752e-01	0.859	19
	2181	0.0442	1.212e-01	0.793	24
	8453	0.0221	9.567e-02	0.620	28
2	53	0.7071	4.364e-01	–	3
	165	0.3536	2.429e-01	0.844	4
	581	0.1768	1.684e-01	0.529	6
	2181	0.0884	9.021e-02	0.900	20
	8453	0.0442	4.013e-02	1.169	25
	33285	0.0221	1.357e-02	1.564	29

(b) Primal method, adaptive refinement

k	DoF	$h_{\min}$	$\mathtt{e}_{\boldsymbol{u}}$	$\widehat{\mathtt{rate}}$	eff(Θ)	iter
1	21	0.7071	6.390e-01	–	0.4912	3
	53	0.3536	4.907e-01	0.617	0.8314	4
	165	0.1768	2.860e-01	1.277	0.8282	6
	581	0.0884	1.489e-01	1.037	0.8205	11
	2105	0.0442	7.413e-02	1.084	0.8187	15
	8177	0.0221	3.516e-02	1.099	0.8219	18
2	53	0.7071	4.364e-01	–	0.9015	3
	165	0.3536	1.973e-01	1.398	1.6422	4
	581	0.1768	1.050e-01	1.489	1.6926	6
	2181	0.0884	2.046e-02	1.891	1.6360	12
	4959	0.0442	9.383e-03	1.786	1.6799	15
	13129	0.0221	1.819e-03	2.097	1.6492	18

Table 5.3: Example 2A. Errors, convergence rates, and Picard iteration count for the approximate displacements $\boldsymbol{u}_h$ produced with the first and second-order primal method; and tabulated according to the resolution level, under uniform (a) and adaptive mesh refinement guided by Θ, with $\gamma_{\text{ratio}} = 0.01$ (b, also displaying the rescaled effectivity index).

no longer evidenced under uniform refinement. In contrast, when adaptive mesh refinement guided by Θ and Ψ is applied, optimal convergence (in this case, linear) is restored. This is particularly visible for the stress and rotation tensor approximation with the mixed method. Another clear indication of the hindered convergence is the number of fixed-point steps needed to reach the tolerance. Irrespective of the method (primal or mixed), the last columns in all sub-tables show that the adaptive algorithm essentially halves the iteration count. As in Example 1, the effectivity indexes are again close to constant values, with the ones generated by the primal scheme being quite small.

We show in Figure 5.2(a,b) the synthetic images projected onto the space of piecewise linear and continuous functions, as well as a few adapted meshes produced using the two indicators (c-h), where one sees that the agglomeration of vertices occurs not so much due to the high gradients of the

(a) Mixed method, uniform refinement

DoF	h	$\mathtt{e}_{\boldsymbol{\sigma}}$	rate	$\mathtt{e}_{\boldsymbol{u}}$	rate	$\mathtt{e}_{\boldsymbol{\rho}}$	rate	iter
91	0.7071	1648.52	–	5.626e-02	–	6.323e-01	–	5
323	0.3536	1124.02	0.552	4.638e-02	0.278	2.801e-01	1.174	7
1219	0.1768	926.769	0.278	2.799e-02	0.728	1.755e-01	0.674	13
4739	0.0884	681.592	0.443	1.557e-02	0.846	9.520e-02	0.882	28
18691	0.0442	381.548	0.741	6.856e-03	1.183	7.267e-02	0.658	45
74243	0.0221	234.824	0.598	2.913e-03	1.235	4.850e-02	0.606	50

(b) Mixed method, adaptive refinement

DoF	$h_{\min}$	$\mathtt{e}_{\boldsymbol{\sigma}}$	$\widehat{\mathtt{rate}}$	$\mathtt{e}_{\boldsymbol{u}}$	$\widehat{\mathtt{rate}}$	$\mathtt{e}_{\boldsymbol{\rho}}$	$\widehat{\mathtt{rate}}$	eff(Ψ)	iter
91	0.7071	1648.54	–	5.626e-02	–	6.326e-01	–	0.7323	3
323	0.3536	894.299	0.964	4.053e-02	0.517	3.017e-01	1.169	0.7352	5
1219	0.1768	747.041	0.860	2.504e-02	0.725	1.695e-01	0.868	0.7425	8
4739	0.0884	546.386	0.933	1.348e-02	0.912	8.456e-02	1.024	0.7413	15
18680	0.0442	322.782	0.967	6.662e-03	1.027	4.131e-02	1.044	0.7447	20
64706	0.0221	128.559	1.148	2.937e-03	1.118	2.170e-02	1.036	0.7329	22

Table 5.4: Example 2B. Errors, convergence rates, and Picard iteration count for the approximate Cauchy stress, displacements, and rotation $\boldsymbol{\sigma}_h, \boldsymbol{u}_h, \boldsymbol{\rho}_h$ produced with our mixed method; and tabulated according to the resolution level, under uniform (a) and adaptive mesh refinement guided by Ψ, with $\gamma_{\text{ratio}} = 0.001$ (b, also displaying the effectivity index).

synthetic images, but mainly because of the features in the solutions to the elasticity problem. Panels (i,j,k) have snapshots of approximate solutions generated with the mixed method after five steps of adaptive refinement, and plotted on the deformed domain.

5.4 Example 3: deformable registration for brain images

We now turn to the application of the adaptive primal and mixed methods in the solution of a DIR problem involving medical images of human brains [15]. The reference and target images for the brain have dimensions 258×258 and the voxel resolution corresponds to 1 mm (see top panels in Figure 5.3). Interpolation from the initial images into finite element fields is done through B-splines applied to the pixel images. We proceed to solve the DIR problem using both primal and mixed adaptive schemes, starting from structured meshes with 32768 triangular elements. The elasticity parameters are set to $E = 15$, $\nu = 0.3$, the weight constant is $\alpha = 50$, and the pseudo timestep is $\delta t = 0.01/\alpha$. The tolerance for the Picard scheme is increased to 1e-04, and for the mixed method the refinement density proportion is ruled by the constant $\gamma_{\text{ratio}} = 0.1$. The primal method requires an average (over the number of mesh refinement steps, here assigned to 4) of 19 Picard steps to reach convergence, which is slightly larger for the mixed method (22 iterations). The first two plots on the middle row of Figure 5.3 depict the converged composed images $T \circ (\boldsymbol{id} + \boldsymbol{u}_h)$ generated with the primal and mixed methods, where we can notice very similar patterns in both cases. The two other figures on the right show the similarity between reference and warped images, $|R(\boldsymbol{x}) - T(\boldsymbol{x} + \boldsymbol{u}_h(\boldsymbol{x}))|$ resulting from both methods.

We also plot an example of a mesh obtained after four steps of adaptive refinement with the primal and mixed methods (see Figure 5.4). For illustration purposes we initiate the process from a coarse mesh of 8196 triangles (corresponding to a low resolution image of 64×64 pixels. Starting with images

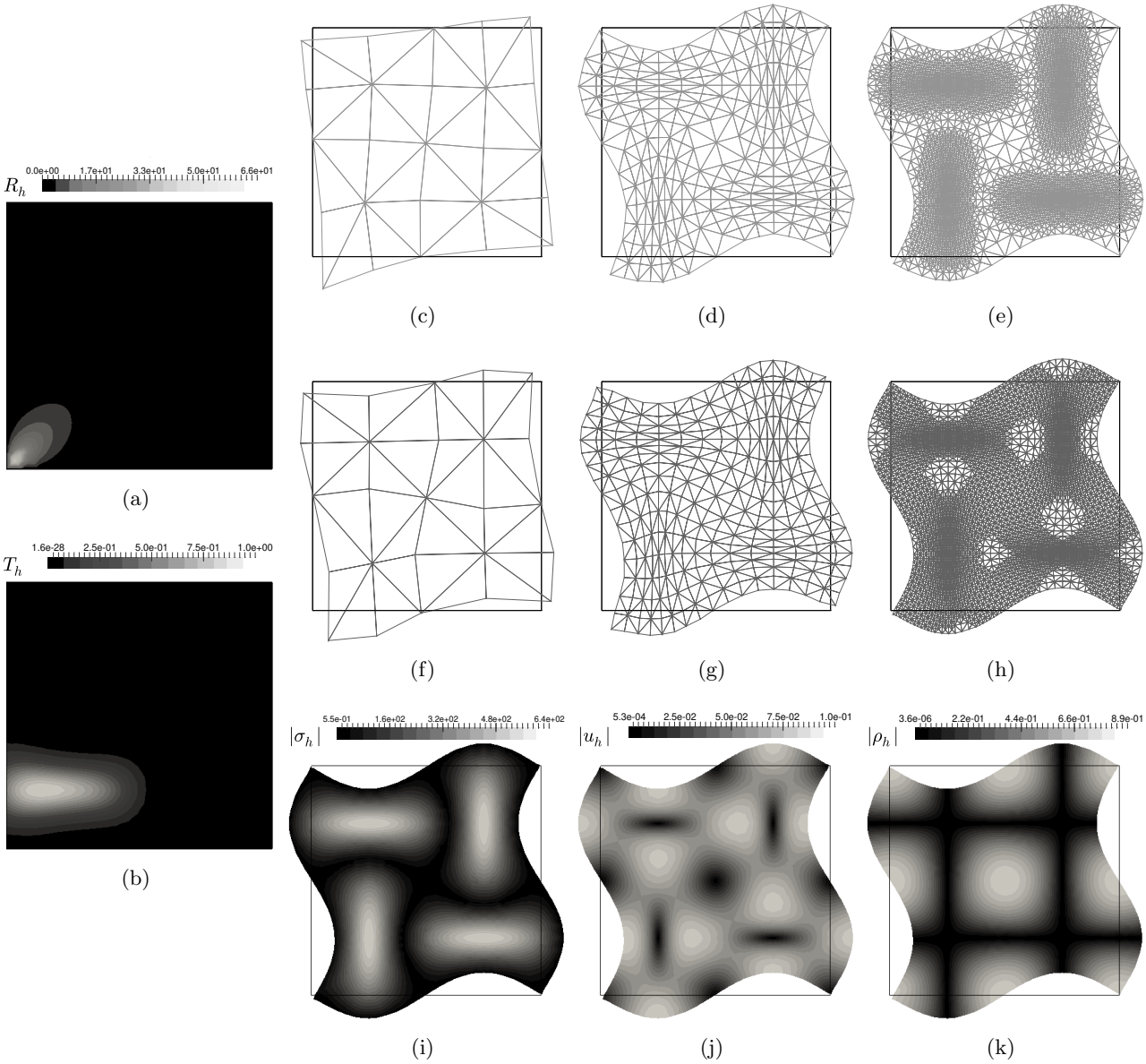


Figure 5.2: Example 2. Projected fields of the reference R and composed $T(\boldsymbol{x} + \boldsymbol{u}_h(\boldsymbol{x}))$ images (a,b); sample meshes refined according to the indicator Θ for the primal method (c,d,e), according to Ψ after solving the problem with the mixed method (f,g,h); and approximate solutions computed with the mixed method (i,j,k).

of higher resolution imply that the meshes obtained after adaptive refinement are too dense to be easily visualized). The figures exemplify the concentration of refinement near the skull, which is consistently the zone with highest gradients in the reference and target images, as well as in stress and rotations (as inferred from panels (g,h,i) in Figure 5.3, where the Frobenius norm of the rotation tensor is plotted in log-scale for clarity). On the other hand, the displacements are, in comparison, rather smooth and they seem not to contribute substantially to the local error indicators.

Finally, for these tests we also provide information on the CPU time required in each step of the

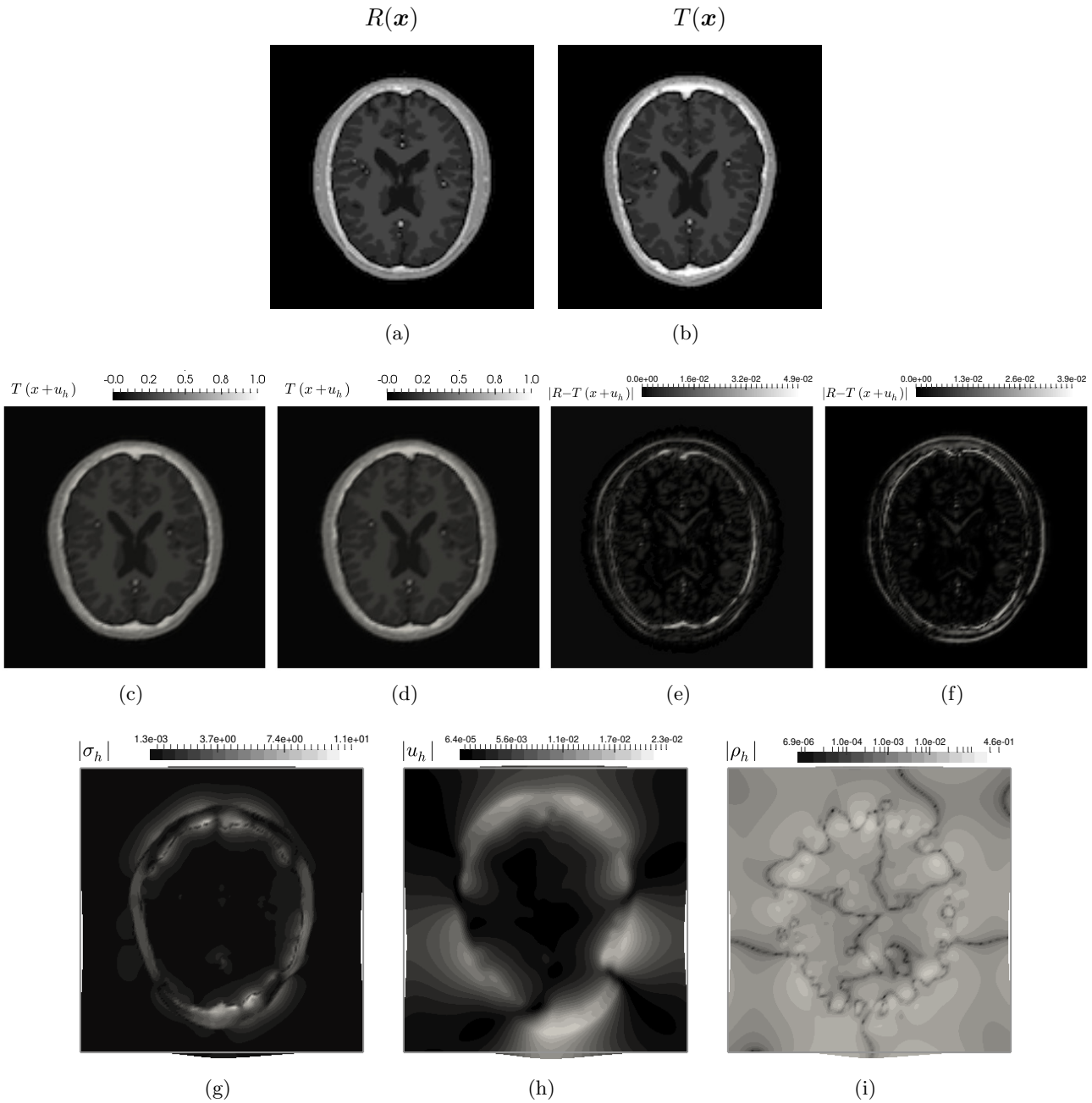
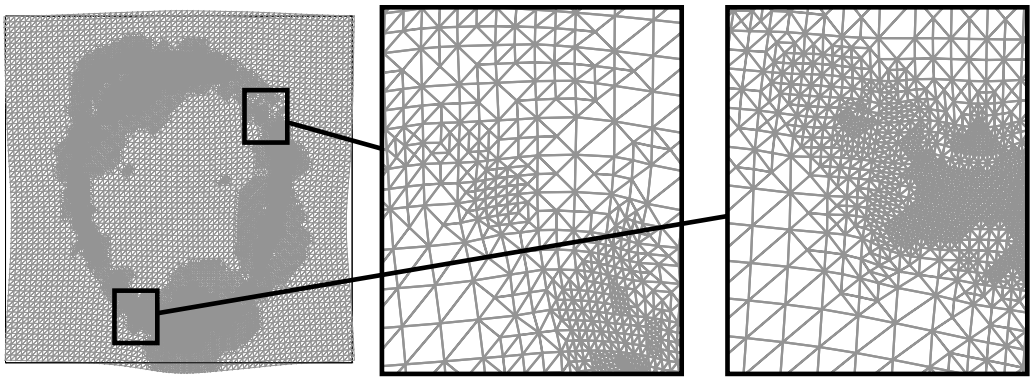
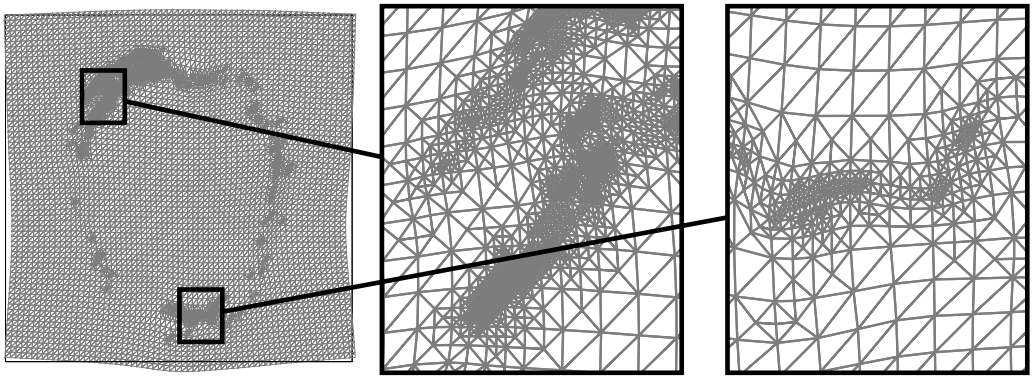


Figure 5.3: Example 3. Reference (a) and target (b) images for medical image registration; composed images for primal and mixed (c,d), and similarity generated with primal and mixed methods (e,f); and solutions to the mixed DIR problem obtained with the adaptive mixed scheme (g,h,i).

overall solution algorithm. We record the wall-time for the mixed and primal methods, when starting from a coarse grid (representing 8715 DoFs for the primal method and 76573 DoFs for the mixed scheme) and in both cases applying five iterations of adaptive mesh refinement. An average of 17 fixed-point iterations are needed for the primal approximations and 25 for the mixed scheme. We collect the information in Table 5.5. For instance, the results indicate that the cost of evaluating the local estimator for the primal method (also including assembly of the global one) is roughly half



(a)



(b)

Figure 5.4: Example 3. Coarse mesh generated by four steps of adaptive refinement using the a posteriori error indicators Θ (a) and Ψ (b), and sampling of two 7x zooms near the skull, all portrayed on the deformed domain.

the time spent in the initial assembly of the left-hand side of the matrix systems, and representing around 15% of the time spent in marking and refining. Also, evaluating the estimator for the mixed method appears to be faster than the one for the primal scheme, possibly due to fewer applications of numerical differentiation in constructing the residual terms. For the primal method, it should be possible to construct robust preconditioners directly following the ideas in [25], however for the case of mixed schemes the adaptation would be much more delicate.

Acknowledgement. The authors thank the fruitful discussions with Bryan Gomez-Vargas (Concepción) and Kent-Andre Mardal (Oslo), regarding the implementation of finite element schemes for elasticity with weakly imposed symmetry and pure traction conditions with Lagrange multipliers and rigid motions.

References

[1] M.S Alnæs, J. Blechta, J. Hake, A. Johansson, B. Kehlet, A. Logg, C. Richardson, J. Ring, M.E. Rognes, and G.N. Wells, *The FEniCS project version 1.5.* Arch. Numer. Softw. 3 (2015), no. 100, 9–23.

	refin. level	matrix assembly	solution computation	IO and residual	evaluation of estimator	marking and refinement
Primal method	1	0.101	0.075 (avg)	0.102	0.096	0.544
(total CPU time: 73.16)	2	0.099	0.163 (avg)	0.110	0.130	0.757
	3	0.162	0.312 (avg)	0.192	0.235	1.284
	4	0.489	1.127 (avg)	0.481	0.704	3.351
	5	0.853	2.093 (avg)	0.758	0.812	5.246
Mixed method	1	0.418	1.445 (avg)	0.101	0.099	0.530
(total CPU time: 997.83)	2	0.443	2.373 (avg)	0.109	0.141	0.668
	3	0.578	4.746 (avg)	0.135	0.154	0.719
	4	0.704	8.390 (avg)	0.204	0.237	1.298
	5	0.921	22.45 (avg)	0.439	0.304	2.616

Table 5.5: Example 3. CPU time (in [s]) of each step of the adaptive finite element method for the DIR problem, measured for the primal and mixed methods, starting from coarse meshes. The time associated with the solution of the linear systems is averaged over the number of inner Picard iterations.

[2] M. Alvarez, G.N. Gatica, and R. Ruiz-Baier, *A posteriori error analysis for a viscous flow-transport problem*. ESAIM Math. Model. Numer. Anal. 50 (2016), no. 6, 1789–1816.

[3] M. Alvarez, G.N. Gatica, and R. Ruiz-Baier, *A posteriori error estimation for an augmented mixed-primal method applied to sedimentation-consolidation systems*. J. Comput. Phys. 367 (2018), 332–346, 2018.

[4] J. Retamal, D.E. Hurtado, N. Villarroel, A. Bruhn, G. Bugedo M.B.P. Amato, E.L.V Costa, G. Hedenstierna, A. Larsson and J.B. Borges *Does Regional Lung Strain Correlate With Regional Inflammation in Acute Respiratory Distress Syndrome During Nonprotective Ventilation? An Experimental Porcine Study*. Critical Care Medicine 46 (2018), no. 6, pp. 591–599.

[5] B.B. Avants, N.J. Tustison, G. Song, P.A. Cook, A. Klein and J.C. Gee, *A reproducible evaluation of ants similarity metric performance in brain image registration*. Neuroimage 54 (2011), no. 3, pp. 2033–2044.

[6] G. Balakrishnan, A. Zhao, M.R. Sabuncu, J. Guttag and A.V. Dalca, *An unsupervised learning model for deformable medical image registration*. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2018), 9252–9260.

[7] N. Barnafi, G.N. Gatica, and D.E. Hurtado, *Primal and mixed finite element methods for deformable image registration problems*. SIAM J. Imaging Sci. 11 (2018), no. 4, pp. 2529–2567.

[8] M.A. Barrientos, G.N. Gatica, and E.P. Stephan, *A mixed finite element method for nonlinear elasticity: two-fold saddle point approach and a-posteriori error estimate*. Numer. Math. 91 (2002), no. 2, 197–222.

[9] S.C. Brenner and L.R. Scott, *The Mathematical Theory of Finite Element Method*. Third edition. Texts in Applied Mathematics, 15. Springer, New York, 2008.

[10] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*. Vol. 15 of Springer Series in Computational Mathematics, Springer-Verlag, New York, 1991.

[11] C. Carstensen and G. Dolzmann, *A posteriori error estimates for mixed FEM in elasticity.* Numer. Math., 81 (1998), no. 2, 187–209.

[12] G. E. Christensen, J. H. Song, W. Lu, I. El Naqa, and D. a. Low, *Tracking lung tissue motion and expansion/compression with inverse consistent image registration and spirometry.* Medical Physics, 34 (2007), 2155–2163.

[13] S. Choi, E. A. Hoffman, S. E. Wenzel, M. H. Tawhai, Y. Yin, M. Castro, and C. L. Lin, *Registration-based assessment of regional lung function via volumetric CT images of normal subjects vs. severe asthmatics.* Journal of Applied Physiology, 115 (2013), 730–742.

[14] P. Clément, *Approximation by finite element functions using local regularisation.* Rev. Française Automat. Informat. Recherche Opérationnelle Sér. 9 (1975), no. R-2, 77–84.

[15] D. L. Collins, A. P. Zijdenbos, V. Kollokian, J. G. Sled, N. J. Kabani, C.J. Holmes, and A.C. Evans, *Design and construction of a realistic digital brain phantom.* IEEE Trans. Med. Imaging, 17 (1998), pp. 463–468.

[16] C. Domínguez, G.N. Gatica, and A. Márquez, *A residual-based a posteriori error estimator for the plane linear elasticity problem with pure traction boundary conditions.* J. Comput. Appl. Math. 292 (2016), 486–504.

[17] W. Dörfler, *A convergent adaptive algorithm for Poisson's equation.* SIAM J. Numer. Anal., 33 (1994), no. 3, 1106–1124.

[18] M. Ebner, M. Modat, S. Ferraris, S. Ourselin and T. Vercauteren, *Forward-backward splitting in deformable image registration: A demons approach.* 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), Washington, DC, 2018, pp. 1065-1069.

[19] G.N. Gatica, *A Simple Introduction to the Mixed Finite Element Method. Theory and Applications* Springer Briefs in Mathematics. Springer, Cham, 2014.

[20] G.N. Gatica, A. Márquez, and S. Meddahi, *A new dual-mixed finite element method for the plane linear elasticity problem with pure traction boundary conditions.* Comput. Methods Appl. Mech. Engrg. 197 (2008), no. 9-12, 1115–1130.

[21] S. Haker, L. Zhu, A. Tannenbaum, and S. Angenent, *Optimal mass transport for registration and warping.* International Journal of Computer Vision, 60 (2004), pp. 225240.

[22] B. K. Horn and B. G. Schunck, *Determining optical flow.* Tech. rep., Cambridge, MA, USA, 1980.

[23] D. E. Hurtado, N. Villarroel, C. Andrade, J. Retamal, G. Bugedo, and A. R. Bruhn, *Spatial patterns and frequency distributions of regional deformation in the healthy human lung.* Biomechanics and Modeling in Mechanobiology, 16 (2017), pp. 1413–1423.

[24] D. E. Hurtado, N. Villarroel, J. Retamal, G. Bugedo, and A. Bruhn, *Improving the accuracy of registration-based biomechanical analysis: A finite element approach to lung regional strain quantification.* IEEE Transactions on Medical Imaging, 35 (2016), pp. 580–588.

[25] M. Kuchta, K.A. Mardal, and M. Mortensen, *On the singular Neumann problem in linear elasticity.* Numer. Linear Alg. Appl. S1-23 (2018).

[26] E. Lee and M. Gunzburger, *An optimal control formulation of an image registration problem.* Journal of Mathematical Imaging and Vision, 36 (2010), pp. 69–80.

[27] J. Modersitzki, *Numerical Methods for Image Registration.* Numerical Mathematics and Scientific Computation, Oxford Science Publications, New York, 2004.

[28] C. Pschl, J. Modersitzki, and O. Scherzer, *A variational setting for volume constrained image registration.* Inverse Problems and Imaging, 4 (2010), pp. 50–522.

[29] R. T. Rockafellar, *Monotone operators and the proximal point algorithm.* SIAM Journal on Control and Optimization, 14 (1976), pp. 877–898.

[30] A. Sotiras, C. Davatzikos, and N. Paragios, *Deformable medical image registration: A survey.* IEEE Transactions on Medical Imaging 32 (2013), no. 7, 1153–1190.

[31] G. Unal and G. Slabaugh, *Coupled pdes for non-rigid registration and segmentation.* IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR05), vol. 1, June 2005, pp. 168–175.

[32] R. Verfürth, *A posteriori error estimation and adaptive mesh-refinement techniques.* J. Comput. Appl. Math. 50 (1994), no. 1-3, 67–83.

[33] R. Verfürth, *A Posteriori Error Estimation Techniques for Finite Element Methods.* Numerical Mathematics and Scientific Computation. Oxford University Press, Oxford, 2013.

[34] R. Verfürth, *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques.* Wiley-Teubner, Stuttgart, 1996.

[35] R. Verfürth, *A review of a posteriori error estimation techniques for elasticity problems.* Comput. Methods Appl. Mech. Engrg. 176 (1999), no. 1-4, 419–440.

[36] B. Zitov, J. Flusser, *Image registration methods: a survey.* Image and Vision Computing 21 (2003), no. 11, 977-1000.