

## Adaptive Mesh Refinement in Deformable Image Registration: A Posteriori Error Estimates for Primal and Mixed Formulations\*

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

**Abstract.** Deformable image registration (DIR) is a popular technique for the alignment of digital images, with highly relevant applications in medical image analysis. However, the numerical solution of DIR problems can be very challenging in computational terms, as the improvement of the DIR solution typically involves a uniform refinement of the underlying domain discretization that exponentially increases the number of degrees of freedom. In this work, we develop adaptive mesh refinement schemes particularly designed for the finite-element solution of DIR problems. We start by deriving residual-based a posteriori error estimators for the primal and mixed formulations of the DIR problem and show that they are reliable and efficient. Based on these error estimators, we implement adaptive mesh-refinement schemes into a finite-element code to register images. We assess the numerical performance of the proposed adaptive scheme on smooth synthetic images, where numerical convergence is verified. We further show that the adaptive mesh refinement scheme can deliver solutions to DIR problems with significant reductions in the number of degrees of freedom without compromising the accuracy of the solution. We also confirm that the adaptive scheme proposed for the mixed DIR formulation successfully handles volume-constrained registration problems, providing optimal convergence in analytic examples. To demonstrate the applicability of the method, we perform adaptive DIR on medical brain images and binary images and study how image noise affects the proposed refinement schemes.

**Key words.** image registration, adaptive mesh refinement, mixed formulations, a posteriori error analysis

**AMS subject classifications.** 68U10, 65N30, 65N15, 74B05

**DOI.** 10.1137/20M1364333

---

\*Received by the editors September 3, 2020; accepted for publication (in revised form) June 1, 2021; published electronically August 30, 2021.

<https://doi.org/10.1137/20M1364333>

**Funding:** The work of the authors was supported by CONICYT-Chile through grant FONDECYT Regular 1180832, project AFB170001 of the PIA program Concurso Apoyo a Centros Científicos y Tecnológicos de Excelencia con Financiamiento Basal, the Centro de Investigación Ingeniería Matemática (CI<sup>2</sup>MA), Universidad de Concepción, Monash Mathematics Research Fund S05802-3951284, and the Ministry of Science and Higher Education of the Russian Federation within the framework of state support for the creation and development of world-class research centers, “Digital biodesign and personalised healthcare,” grant 075-15-2020-926.

<sup>†</sup>Dipartimento di Matematica, Università degli Studi di Milano, Via Saldini 50, I-20133 Milan, Italy ([nicolas.barnafi@polimi.it](mailto:nicolas.barnafi@polimi.it)).

<sup>‡</sup>CI<sup>2</sup>MA and Departamento de Ingeniería Matemática, Universidad de Concepción, Casilla 160-C, Concepción, Chile ([ggatica@ci2ma.udec.cl](mailto:ggatica@ci2ma.udec.cl), [wmiranda@ci2ma.udec.cl](mailto:wmiranda@ci2ma.udec.cl)).

<sup>§</sup>Corresponding author. 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 ([dhurtado@ing.puc.cl](mailto:dhurtado@ing.puc.cl)).

<sup>¶</sup>School of Mathematics, Monash University, 9 Rainforest Walk, Melbourne 3800 VIC, Australia; Institute of Computer Science and Mathematical Modelling, Sechenov University, Moscow, Russian Federation; and Universidad Adventista de Chile, Casilla 7-D, Chillán, Chile ([ricardo.ruizbaier@monash.edu](mailto:ricardo.ruizbaier@monash.edu)).

**1. Introduction.** Deformable image registration (DIR) consists of aligning two images through a transformation that deforms one image onto the other. It arises in several applications, particularly in the medical imaging field [39]. Its mathematical formulation requires three objects: a transformation model, defined by a family of suitable mappings that warp the target image; a similarity measure, typically represented by a functional that quantifies the difference between images; and a regularizer, which renders the problem well-posed [34]. In addition to the many variants of these components, different modeling approaches exist, among which we highlight the traditional variational minimization [28, 34], optimal mass transport [25], and level-set modeling [40]. The solution of the DIR problem typically considers incorporating an auxiliary time variable. This approach can be interpreted as a semi-implicit formulation of the proximal point algorithm [38] extended to a more general class of proximal operators by using forward-backward splitting [21]. The formulation of the optical flow problem put forward by Horn and Schunk [28] leads 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 [36] in an algorithm-specific fashion and also in the optimal-control setting within a more classical Galerkin framework [33].

One active area of DIR application is the study of deformation in the lungs from the analysis of computed tomography images of the thorax [15]. In this setting, the optimal warping  $u$  that solves the DIR problem can be interpreted as a displacement field, from which deformation metrics such as the strain tensor can be computed based on  $\nabla u$  using the framework of continuum mechanics. The study of deformation using DIR has revealed that the lungs display a highly heterogeneous and anisotropic behavior [29]. Further, deformation metrics from the strain tensor have proved very insightful in understanding certain pulmonary diseases and lung injury progression [16, 37, 31]. The prediction of strain measures from DIR is not without problems, as it has been shown that estimating the strain tensor from direct differentiation of the transformation mapping yields spurious numerical errors that can distort the physical meaning of the strain tensor [30]. This problem, together with an effort of providing a rigorous analysis of the Galerkin formulation of DIR, motivated the recent development of primal and mixed continuous formulations and finite-element schemes [5]. This last work used null traction boundary conditions so as to avoid spurious stress. It relied on the mixed theory of linear elasticity problems with pure-traction conditions [23], which delivered a priori error estimates not only for the displacement solution but also for the stress and rotation fields in the mixed formulation. These analytical results provide a sound framework for the error assessment of stress and deformation estimates in DIR.

Depending on the amount of warping from the target to the reference images, the optimal warping  $u$  can typically result in localized regions with high variations. These localizations may not be accurately captured by the transformation model, which has motivated the development of adaptive refinement techniques in other areas of numerical analysis [42]. In the case of image processing, adaptive refinement finite-element techniques have been developed and analyzed for the image segmentation [10, 11], image denoising [27, 26], and optical flow [9, 8] problems. However, specific schemes developed for DIR formulations remain understudied. One exception is the work of Haber, Heldmann, and Modersitzki [24], where a finite-difference scheme was employed to solve the DIR problem, and an oct-tree strategy was used to

improve the numerical solution by adaptive refinement. Another approach is the use of ad hoc mesh-refinement techniques based on classical strategies in finite-element analysis for elasticity [46, 35]. While very useful, this approach does not directly extend to mixed formulations, and it lacks a theoretical framework that can guarantee the numerical convergence of the scheme.

In this work, we propose an a posteriori mesh-refinement scheme particularly tailored for the primal and mixed finite-element formulations of the DIR problem. We validate the proposed mesh-refinement scheme and the associated theoretical results through applications on the registration of analytic and medical images, as well as images subject to incompressibility constraints, where the performance of the methods is assessed in terms of error measures and convergence rates. We start by constructing an optimal a posteriori error estimator  $\Theta$  [42]. The estimator  $\Theta$  is of residual type, and it is decomposed into a sum of local error indicators  $\theta_T$  that give a normwise equivalent of the total error. The estimator  $\Theta$  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 for denoting higher-order terms. This estimator is designed to be effective in terms of computing cost, allowing one to rapidly identify regions with large total error that are candidates for local mesh refinement. At the same time, the use of the estimator prevents the refinement of areas where the error is small, delivering an efficient scheme for error reduction, which is in contrast to uniform refinement schemes.

**Outline.** We have organized the contents of this paper as follows. The remainder of this section introduces some standard notation and definitions of functional spaces. In section 2, we state the mathematical formulation of DIR, along with the similarity measure and regularizer considered in this work. In section 3, we state the weak problems for the primal and mixed formulations of DIR, along with their corresponding Galerkin schemes. In section 4, we develop a posteriori error indicators for the finite element formulations, to then derive the corresponding theoretical bounds yielding reliability and efficiency of each estimator under reasonable assumptions. To demonstrate the applicability of the proposed methods, in section 5 we apply the mesh-refinement scheme in the elastic registration of smooth and medical images, where we confirm the reliability and efficiency of the estimators, along with assessing their numerical performance. We end this article by discussing the results and sharing some concluding remarks in section 6.

**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  $\mathbf{n}$  the outward unit normal vector on  $\Gamma$ . 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  $\mathbf{A}$  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  $\mathbf{v} = (v_i)_{i=1,n}$ , we set the gradient and divergence operator as

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

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  $\mathbf{div}$  acting along the rows of  $\boldsymbol{\tau}$ , and define the transpose, the trace, the tensor inner product, and the deviatoric tensor, respectively, as

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

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

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

equipped with the usual norm

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

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

**2. Mathematical formulation of the deformable image registration problem.** In this section we recall from [5, section 2] the elastic DIR 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  $T \in H^1(\tilde{\Omega})$  be the target image. The DIR problem consists in finding a transformation  $\mathbf{u} : \Omega \rightarrow \mathbb{R}^n$ , also known as the displacement field, that best aligns the images  $R$  and  $T$ , which is expressed as the variational problem (cf. [34])

$$(2.1) \quad \inf_{\mathbf{u} \in \mathcal{V}} \alpha \mathcal{D}[\mathbf{u}; R, T] + \mathcal{S}[\mathbf{u}],$$

where  $\mathcal{V}$  is typically  $\mathbf{H}^1(\Omega)$ ,  $\mathcal{D} : V \rightarrow \mathbb{R}$  is the similarity measure between the images  $R$  and  $T$ ,  $\alpha > 0$  is a weighting constant, and  $\mathcal{S} : V \rightarrow \mathbb{R}$  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

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

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

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

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

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

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

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

where

$$(2.4) \quad \mathbf{f}_u(\mathbf{x}) = \{T(\mathbf{x} + \mathbf{u}(\mathbf{x})) - R(\mathbf{x})\} \nabla T(\mathbf{x} + \mathbf{u}(\mathbf{x})) \quad \forall \mathbf{x} \in \Omega \text{ a.e.}$$

We assume that there are positive constants  $L_f$  and  $M_f$  such that the nonlinear load term  $\mathbf{f}_u$  satisfies

$$(2.5) \quad |\mathbf{f}_u(\mathbf{x}) - \mathbf{f}_v(\mathbf{x})| \leq L_f |\mathbf{u}(\mathbf{x}) - \mathbf{v}(\mathbf{x})|, \quad |\mathbf{f}_u(\mathbf{x})| \leq M_f \quad \forall \mathbf{x} \in \Omega \text{ a.e.}$$

**3. Continuous and discrete weak formulations of DIR.** In this section we summarize the continuous primal and mixed variational formulations of (2.3) derived in [5, section 3] and [5, section 4], respectively, and recall the respective solvability results.

**3.1. DIR primal formulation.** The primal variational formulation for the registration problem reads, Find  $\mathbf{u} \in \mathbf{H}^1(\Omega)$  such that

$$(3.1) \quad a(\mathbf{u}, \mathbf{v}) = \alpha F_u(\mathbf{v}), \quad \mathbf{v} \in \mathbf{H}^1(\Omega),$$

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

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

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

$$F_u(\mathbf{v}) := - \int_{\Omega} \mathbf{f}_u \cdot \mathbf{v} \quad \forall \mathbf{v} \in \mathbf{H}^1(\Omega).$$

The conditions (2.5) imply the Lipschitz continuity and uniform boundedness of  $F_u$ , that is,

$$(3.3) \quad \|F_u - F_v\|_{\mathbf{H}^1(\Omega)'} \leq L_F \|\mathbf{u} - \mathbf{v}\|_{0,\Omega}, \quad \|F_u\|_{\mathbf{H}^1(\Omega)'} \leq M_F \quad \forall \mathbf{u}, \mathbf{v} \in \mathbf{H}^1(\Omega),$$

respectively. We recall the results concerning the solvability of (3.1), as developed in [5, section 3]. First, we define the following linear auxiliary problem: Given  $\mathbf{z} \in \mathbf{H}^1(\Omega)$ , find  $\mathbf{u} \in \mathbf{H}^1(\Omega)$  such that

$$(3.4) \quad a(\mathbf{u}, \mathbf{v}) = \alpha F_z(\mathbf{v}), \quad \mathbf{v} \in \mathbf{H}^1(\Omega).$$

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 [12, Equation 11.1.7])

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

which guarantees unique solvability of (3.4) since  $\mathbb{RM}(\Omega)$  is the null space of its solution operator. Defining

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

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

$$(3.6) \quad a(\mathbf{u}, \mathbf{v}) = \alpha F_z(\mathbf{v}), \quad \mathbf{v} \in H.$$

Then, we have the following result, proven in [5, Theorem 2].

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

$$\|\mathbf{u}\|_{1,\Omega} \leq \alpha C_p \|F\mathbf{z}\|_{\mathbf{H}^1(\Omega)'}$$

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

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

The following result, also proven in [5, Theorem 3], establishes the existence and uniqueness of a 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.*

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

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

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 stress tensor field  $\boldsymbol{\sigma} := \mathcal{C}\mathbf{e}(\mathbf{u})$ . Further, we note that the constitutive relation (2.2) can be inverted (cf. [6] or [22]) as

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

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

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

Introducing the spaces

$$\mathbb{H}_0(\mathbf{div}; \Omega) = \{\boldsymbol{\tau} \in \mathbb{H}(\mathbf{div}; \Omega) : \gamma_n \boldsymbol{\tau} = \mathbf{0}\}, \quad \mathbf{Q} := \mathbf{L}^2(\Omega) \times \mathbb{L}_{\text{skew}}^2(\Omega),$$

where

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

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

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

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

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

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

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

$$F_{\mathbf{u}}(\mathbf{v}, \boldsymbol{\eta}) := \int_{\Omega} \mathbf{f}_{\mathbf{u}} \cdot \mathbf{v} \quad \forall (\mathbf{v}, \boldsymbol{\eta}) \in \mathbf{Q}.$$

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

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

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

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

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

$$\begin{aligned} A((\boldsymbol{\sigma}, \boldsymbol{\chi}), (\boldsymbol{\tau}, \boldsymbol{\xi})) &:= a(\boldsymbol{\sigma}, \boldsymbol{\tau}) + \int_{\Omega} \boldsymbol{\chi} \cdot \boldsymbol{\xi} & \forall (\boldsymbol{\sigma}, \boldsymbol{\chi}), (\boldsymbol{\tau}, \boldsymbol{\xi}) \in \mathbf{H}, \\ B((\boldsymbol{\tau}, \boldsymbol{\xi}), (\mathbf{v}, \boldsymbol{\eta})) &:= b(\boldsymbol{\tau}, (\mathbf{v}, \boldsymbol{\eta})) + \int_{\Omega} \boldsymbol{\xi} \cdot \mathbf{v} & \forall ((\boldsymbol{\tau}, \boldsymbol{\xi}), (\mathbf{v}, \boldsymbol{\eta})) \in \mathbf{H} \times \mathbf{Q}. \end{aligned}$$

The following two lemmas proven in [23, Lemmas 3.3–3.4] are needed to establish the well-posedness of (3.13).

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

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

and there exists  $\hat{\alpha} > 0$  such that

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



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

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

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

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

*Proof.* See [23, Theorem 3.1]. ■

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

$$(3.15) \quad \mathbf{T}(\mathbf{u}) = \mathbf{u}.$$

And the existence of a solution to the fixed-point problem (3.15) follows directly as in [5, Theorem 12].

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

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

**3.3. The primal Galerkin finite-element scheme.** Let  $\mathbf{H}_h$  be a finite dimensional subspace of  $\mathbf{H}^1(\Omega)$  and define  $H_h := \mathbb{R}\mathbf{M}^\perp \cap \mathbf{H}_h$ . Then the primal nonlinear discrete problem is, Find  $\mathbf{u}_h \in H_h$  such that

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

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

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

and also let  $T_h : H_h \rightarrow H_h$  be the discrete operator given by  $T_h(\mathbf{z}_h) = \mathbf{u}_h$ , where  $\mathbf{u}_h$  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, proven in [5, Theorem 5].

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



**3.4. The mixed Galerkin finite-element scheme.** In this section we recall the Galerkin finite-element 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  $\Gamma$ . Let us consider finite dimensional subspaces  $H_h^\sigma$ ,  $Q_h^u$ , and  $Q_h^\rho$  of  $\mathbb{H}(\mathbf{div}; \Omega)$ ,  $L^2(\Omega)$ , and  $\mathbb{L}_{\text{skew}}^2(\Omega)$ , respectively. Then we introduce the product spaces

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

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

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

The unique solvability and stability of (3.18), being the Galerkin scheme of a linear elasticity problem with pure traction boundary conditions, has already been established in [23, Theorem 4.1]. This allows us to define the discrete operator  $\mathbf{T}_h : Q_h^u \rightarrow Q_h^u$  given by  $\mathbf{T}_h(\mathbf{z}_h) := \mathbf{u}_h$ , where  $\mathbf{u}_h$  is the unique displacement from (3.18), and then we rewrite the discrete nonlinear problem as, Find  $\mathbf{u}_h \in Q_h^u$  such that

$$(3.19) \quad \mathbf{T}_h(\mathbf{u}_h) = \mathbf{u}_h.$$

The well-posedness of problem (3.19) is now recalled from [5, Theorem 14].

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

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

**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 finite-element 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  $\mathbf{n}_e := (\mathbf{n}_1, \mathbf{n}_2)^t$  and let  $\mathbf{s}_e := (-\mathbf{n}_2, \mathbf{n}_1)^t$  be the corresponding fixed unit tangential vector along  $e$ . However, when no confusion arises, we simply write  $\mathbf{n}$  and  $\mathbf{s}$  instead of  $\mathbf{n}_e$  and  $\mathbf{s}_e$ , 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  $[\boldsymbol{\tau} \mathbf{s}]$  and  $[\boldsymbol{\tau} \mathbf{n}]$  the tangential and normal jumps of  $\boldsymbol{\tau}$  across  $e$ , that is,  $[\boldsymbol{\tau} \mathbf{s}] := (\boldsymbol{\tau}|_K - \boldsymbol{\tau}|_{K'})|_e \mathbf{s}$  and  $[\boldsymbol{\tau} \mathbf{n}] := (\boldsymbol{\tau}|_K - \boldsymbol{\tau}|_{K'})|_e \mathbf{n}$ , 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  $v$ ,  $\boldsymbol{\varphi} = (\varphi_1, \varphi_2)^t$  and  $\boldsymbol{\tau} := (\tau_{ij})_{1 \leq i, j \leq 2}$ , respectively, we let

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

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  $I_h : H^1(\Omega) \rightarrow X_h$  be the usual Clément interpolation operator (cf. [17]), 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  $I_h$ . For a proof, see [17].

**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*

$$\begin{aligned} \|v - I_h(v)\|_{0,K} &\leq c_1 h_K \|v\|_{0,\Delta(K)} \quad \forall K \in \mathcal{T}_h, \\ \|v - I_h(v)\|_{0,e} &\leq c_2 h_e^{1/2} \|v\|_{0,\Delta(e)} \quad \forall e \in \mathcal{E}_h(\Omega) \cup \mathcal{E}_h(\Gamma), \end{aligned}$$

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

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  $\psi_K$  and  $\psi_e$  be the usual triangle-bubble and edge-bubble functions [43, equations (1.5)–(1.6)], respectively, which satisfy

- (i)  $\psi_K \in P_3(K)$ ,  $\psi_K = 0$  on  $\partial K$ ,  $\text{supp}(\psi_K) \subseteq K$ , and  $0 \leq \psi_K \leq 1$  in  $K$ ,
- (ii)  $\psi_e \in P_2(K)$ ,  $\psi_e = 0$  on  $\partial K$ ,  $\text{supp}(\psi_e) \subseteq \omega_e$ , and  $0 \leq \psi_e \leq 1$  in  $\omega_e$ ,

where  $\omega_e := \cup\{K' \in \mathcal{T}_h : e \in \mathcal{E}(K')\}$ . Additional properties of  $\psi_K$  and  $\psi_e$  are collected in the following lemma (cf. [41, Lemma 1.3], [43, section 3.4], or [44, section 4]).

**Lemma 4.2.** *Given  $k \in \mathbb{N} \cup \{0\}$ , there exist positive constants  $\gamma_1, \gamma_2, \gamma_3, \gamma_4$ , and  $\gamma_5$ , 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*

$$(4.1a) \quad \gamma_1 \|q\|_{0,K}^2 \leq \left\| \psi_K^{1/2} q \right\|_{0,K}^2, \quad \|\psi_K q\|_{1,K} \leq \gamma_2 h_K^{-1} \|q\|_{0,K} \quad \forall q \in P_k(K),$$

$$(4.1b) \quad \gamma_3 \|p\|_{0,e}^2 \leq \left\| \psi_e^{1/2} p \right\|_{0,e}^2, \quad \|\psi_e p\|_{1,\omega_e} \leq \gamma_4 h_e^{-1/2} \|p\|_{0,e}, \quad \|\psi_e p\|_{0,\omega_e} \leq \gamma_5 h_e^{1/2} \|p\|_{0,e} \quad \forall p \in P_k(e).$$

**4.2. A posteriori error analysis for the primal finite-element scheme.** We now derive a reliable and efficient residual-based a posteriori error estimator for (3.16). We draw ideas from [2, 3] (see also the monograph [42]). Letting  $\mathbf{u}_h \in H_h$  be the unique solution of (3.16), we define for each  $K \in \mathcal{T}_h$  the a posteriori error indicator,

$$(4.2) \quad \Theta_K^2 := h_K^2 \|\alpha \mathbf{f}_{\mathbf{u}_h} - \text{div}(\mathcal{C}\mathbf{e}(\mathbf{u}_h))\|_{0,K}^2 + \sum_{e \in \mathcal{E}(K) \cap \mathcal{E}_h(\Omega)} h_e \|\mathcal{C}\mathbf{e}(\mathbf{u}_h) \mathbf{n}_e\|_{0,e}^2 + \sum_{e \in \mathcal{E}(K) \cap \mathcal{E}_h(\Gamma)} h_e \|\mathcal{C}\mathbf{e}(\mathbf{u}_h) \mathbf{n}_e\|_{0,e}^2,$$

where, according to (2.4),

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

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

$$(4.3) \quad C_{\text{eff}} \Theta \leq \|\mathbf{u} - \mathbf{u}_h\|_H \leq C_{\text{rel}} \Theta.$$

The reliability (upper bound in (4.3)) and the efficiency (lower bound in (4.3)) of  $\Theta$  are established separately in the following two lemmas. The corresponding proofs are postponed to Appendix A.

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

$$\|\mathbf{u} - \mathbf{u}_h\|_H \leq C_{\text{rel}} \Theta.$$

**Lemma 4.5.** *There exist constants  $\eta_1, \eta_2, \eta_3 > 0$  and  $C_{\text{eff}} > 0$ , independent of  $h$ , but depending on  $\gamma_1, \gamma_2, \gamma_3, \gamma_4$ , and  $\gamma_5$  (cf. (4.1)), such that for each  $K \in \mathcal{T}_h$  there holds*

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

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

$$C_{\text{eff}} \Theta \leq \|\mathbf{u} - \mathbf{u}_h\|_H.$$

**4.3. A posteriori error analysis for the mixed finite-element scheme.** In this section we derive a reliable and efficient residual-based a posteriori error estimator for (3.18). Throughout the rest of this section we let  $((\boldsymbol{\sigma}, \boldsymbol{\chi}), (\mathbf{u}, \boldsymbol{\rho})) \in \mathbf{H} \times \mathbf{Q}$  and  $((\boldsymbol{\sigma}_h, \boldsymbol{\chi}_h), (\mathbf{u}_h, \boldsymbol{\rho}_h)) \in \mathbf{H}_h \times \mathbf{Q}_h$  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  $K \in \mathcal{T}_h$

$$(4.4) \quad \begin{aligned} \Psi_K^2 := & \|\alpha \mathbf{f}_{\mathbf{u}_h} - \mathbf{div} \boldsymbol{\sigma}_h\|_{0,K}^2 + \|\boldsymbol{\sigma}_h - \boldsymbol{\sigma}_h^t\|_{0,K}^2 + \|\boldsymbol{\chi}_h\|_{0,K}^2 + h_K^2 \|\operatorname{curl}(\mathcal{C}^{-1} \boldsymbol{\sigma}_h + \boldsymbol{\rho}_h)\|_{0,K}^2 \\ & + h_K^2 \|\mathcal{C}^{-1} \boldsymbol{\sigma}_h + \boldsymbol{\rho}_h\|_{0,K}^2 + \sum_{e \in \mathcal{E}(K) \cap \mathcal{E}_h(\Omega)} h_e \|[(\mathcal{C}^{-1} \boldsymbol{\sigma}_h + \boldsymbol{\rho}_h) \mathbf{s}]\|_{0,e}^2 \\ & + \sum_{e \in \mathcal{E}(K) \cap \mathcal{E}_h(\Gamma)} h_e \|(\mathcal{C}^{-1} \boldsymbol{\sigma}_h + \boldsymbol{\rho}_h) \mathbf{s}\|_{0,e}^2. \end{aligned}$$

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  $C_{\text{rel}}, C_{\text{eff}} > 0$  independent of  $h$  such that*

$$(4.5) \quad C_{\text{eff}} \Psi \leq \|(\boldsymbol{\sigma}, \boldsymbol{\chi}) - (\boldsymbol{\sigma}_h, \boldsymbol{\chi}_h)\|_{\mathbf{H}} + \|(\mathbf{u}, \boldsymbol{\rho}) - (\mathbf{u}_h, \boldsymbol{\rho}_h)\|_{\mathbf{Q}} \leq C_{\text{rel}} \Psi.$$

Pertaining to the reliability of the estimator (the upper bound in (4.5)), we begin by stating a more general result due to Lemmas 3.3 and 3.4 and Theorem 3.6.

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

$$(4.6) \quad \begin{aligned} A((\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\chi}}), (\boldsymbol{\tau}, \boldsymbol{\xi})) + B((\boldsymbol{\tau}, \boldsymbol{\xi}), (\bar{\mathbf{u}}, \bar{\boldsymbol{\rho}})) &= \bar{F}((\boldsymbol{\tau}, \boldsymbol{\xi})) \quad \forall (\boldsymbol{\tau}, \boldsymbol{\xi}) \in \mathbf{H}, \\ B((\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\chi}}), (\mathbf{v}, \boldsymbol{\eta})) &= \bar{G}_u((\mathbf{v}, \boldsymbol{\eta})) \quad \forall (\mathbf{v}, \boldsymbol{\eta}) \in \mathbf{Q}. \end{aligned}$$

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

$$(4.7) \quad \|(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\chi}})\|_{\mathbf{H}} + \|(\bar{\mathbf{u}}, \bar{\boldsymbol{\rho}})\|_{\mathbf{Q}} \leq C \{\|\bar{F}\|_{\mathbf{H}'} + \|\bar{G}_u\|_{\mathbf{Q}'}\}.$$

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

$$(4.8) \quad S_h(\boldsymbol{\tau}) := a(\boldsymbol{\sigma}_h, \boldsymbol{\tau}) + b(\boldsymbol{\tau}, (\mathbf{u}_h, \boldsymbol{\rho}_h)) \quad \forall \boldsymbol{\tau} \in \mathbb{H}(\mathbf{div}; \Omega),$$

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  $B$  (cf. (3.14)). We note that  $S_h(\boldsymbol{\tau}_h) = 0$  for each  $\boldsymbol{\tau}_h \in H_h^\sigma$ .

Now, we make use of a particular problem of the form (4.6) with  $\bar{F} \in \mathbf{H}'$  and  $\bar{G}_u \in \mathbf{Q}'$  defined by

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

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

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

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

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

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

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

by the previous estimate and the continuous dependence results (4.7), we have

$$(4.9) \quad \|(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\chi}})\|_{\mathbf{H}} \leq C \left\{ \|\alpha \mathbf{f}_{u_h} - \operatorname{div} \boldsymbol{\sigma}_h\|_{0,\Omega} + \|\boldsymbol{\sigma}_h - \boldsymbol{\sigma}_h^t\|_{0,\Omega} + \|\boldsymbol{\chi}_h\|_{0,\Omega} + \alpha L_F \|\mathbf{u} - \mathbf{u}_h\|_{0,\Omega} \right\}.$$

Now, applying the triangle inequality we obtain

$$(4.10) \quad \|(\boldsymbol{\sigma}, \boldsymbol{\chi}) - (\boldsymbol{\sigma}_h, \boldsymbol{\chi}_h)\|_{\mathbf{H}} \leq \|(\boldsymbol{\sigma}, \boldsymbol{\chi}) - (\boldsymbol{\sigma}_h, \boldsymbol{\chi}_h) - (\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\chi}})\|_{\mathbf{H}} + \|(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\chi}})\|_{\mathbf{H}},$$

and hence it remains to estimate  $\|(\boldsymbol{\sigma}, \boldsymbol{\chi}) - (\boldsymbol{\sigma}_h, \boldsymbol{\chi}_h) - (\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\chi}})\|_{\mathbf{H}}$ . First observe that  $(\boldsymbol{\sigma}, \boldsymbol{\chi}) - (\boldsymbol{\sigma}_h, \boldsymbol{\chi}_h) - (\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\chi}}) \in \mathbf{V}$ ; hence applying the ellipticity of  $A$  in  $\mathbf{V}$  (cf. Lemma 3.3) and analogously to [19, Lemma 4.6], and using (4.9) and (4.10), it follows that

$$(4.11) \quad \|(\boldsymbol{\sigma}, \boldsymbol{\chi}) - (\boldsymbol{\sigma}_h, \boldsymbol{\chi}_h)\|_{\mathbf{H}} \leq C \left\{ \|S_h|_V\|_{V'} + \|\alpha \mathbf{f}_{u_h} - \operatorname{div} \boldsymbol{\sigma}_h\|_{0,\Omega} + \|\boldsymbol{\sigma}_h - \boldsymbol{\sigma}_h^t\|_{0,\Omega} + \|\boldsymbol{\chi}_h\|_{0,\Omega} + \alpha L_F \|\mathbf{u} - \mathbf{u}_h\|_{0,\Omega} \right\}.$$

To estimate  $\|S_h|_V\|_{V'}$  (cf. (4.8)) in (4.11), we use the following result, derived in [19, Lemma 4.7].

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

$$(4.12) \quad \|S_h|_V\|_{V'} \leq C \left\{ h_K^2 \|\operatorname{curl}(\mathcal{C}^{-1} \boldsymbol{\sigma}_h + \boldsymbol{\rho}_h)\|_{0,K}^2 + \sum_{e \in \mathcal{E}(K) \cap \mathcal{E}_h(\Omega)} h_e \|[(\mathcal{C}^{-1} \boldsymbol{\sigma}_h + \boldsymbol{\rho}_h) \mathbf{s}]\|_{0,e}^2 + \sum_{e \in \mathcal{E}(K) \cap \mathcal{E}_h(\Gamma)} h_e \|(\mathcal{C}^{-1} \boldsymbol{\sigma}_h + \boldsymbol{\rho}_h) \mathbf{s}\|_{0,e}^2 \right\}.$$

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*

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

where

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

*Proof.* It follows straightforwardly from (4.11) and (4.12). ■

Now we proceed to obtain the corresponding upper bound for  $\|(\mathbf{u}, \boldsymbol{\rho}) - (\mathbf{u}_h, \boldsymbol{\rho}_h)\|_{\mathcal{Q}}$ .

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

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

where  $\Psi_K^2$  is the local indicator defined in (4.4).

*Proof.* The proof follows directly from [19, Lemma 4.9] with small modifications. ■

The reliability of  $\Psi$  is a straightforward consequence of Lemmas 4.9 and 4.10, assuming  $\alpha C_m L_F < 1/2$ .

On the other hand, the efficiency analysis focuses on deriving upper bounds depending on the actual errors for the seven terms defining the local indicator  $\Psi_K^2$  (cf. (4.4)). For this, and analogously to [19, section 4.3], we begin with the first three bounds; more precisely, since  $\mathbf{div}(\boldsymbol{\sigma}) = \alpha \mathbf{f}_u$  in  $\Omega$ , we have that

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

Next, adding and subtracting  $\boldsymbol{\sigma}$ , and using that  $\boldsymbol{\sigma} = \boldsymbol{\sigma}^t$  in  $\Omega$ , leads to

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

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  $C_1, C_2, C_3, C_4 > 0$ , independent of  $h$ , such that for each  $K \in \mathcal{T}_h$  there holds*

$$\begin{aligned} h_K^2 \|\mathbf{curl}(\mathcal{C}^{-1} \boldsymbol{\sigma}_h + \boldsymbol{\rho}_h)\|_{0,K}^2 &\leq C_1 \{ \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,K}^2 + \|\boldsymbol{\rho} - \boldsymbol{\rho}_h\|_{0,K}^2 \} \\ \|\mathcal{C}^{-1} \boldsymbol{\sigma}_h + \boldsymbol{\rho}_h\|_{0,K}^2 &\leq C_2 \{ \|\mathbf{u} - \mathbf{u}_h\|_{0,K}^2 + h_K^2 \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,K}^2 + h_K^2 \|\boldsymbol{\rho} - \boldsymbol{\rho}_h\|_{0,K}^2 \} \\ h_e \|[(\mathcal{C}^{-1} \boldsymbol{\sigma}_h + \boldsymbol{\rho}_h) \mathbf{s}]\|_{0,e}^2 &\leq C_3 \sum_{K \subseteq \omega_e} \{ \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,K}^2 + \|\boldsymbol{\rho} - \boldsymbol{\rho}_h\|_{0,K}^2 \} \\ \sum_{e \in \mathcal{E}_h(\Gamma)} h_e \|(\mathcal{C}^{-1} \boldsymbol{\sigma}_h + \boldsymbol{\rho}_h) \mathbf{s}\|_{0,e}^2 &\leq C_4 \sum_{e \in \mathcal{E}_h(\Gamma)} \{ \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,K}^2 + \|\boldsymbol{\rho} - \boldsymbol{\rho}_h\|_{0,K}^2 \}, \end{aligned}$$

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

*Proof.* See [19, section 4.3]. ■

## 5. Applications and performance assessment.

**5.1. Numerical implementation.** We now turn to the implementation of some numerical tests that confirm the predicted reliability and efficiency of the a posteriori error estimators (4.2) and (4.4). 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 displacement. 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 [22]. 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 biconjugate gradient method (BiCGStab) using an incomplete LU decomposition as preconditioner.

Mesh adaptation guided by the a posteriori error estimators 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 error indicators, we proceed to tag elements for refinement using the Dörfler strategy [20], where we mark sufficiently many elements so that one establishes equidistribution 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  $\zeta$  over the initial mesh (see, for instance, [41]). For this, in each of the accuracy tests below, we consider an arbitrary constant  $\gamma_{\text{ratio}}$  such that the elements to be refined are the ones such that the local error  $\zeta_i$  is bigger than  $\gamma_{\text{ratio}} \max_i \zeta_i$  so as to generate either a roughly similar number of degrees of freedom or similar individual error magnitudes as in the case of uniform refinement. The density of the refinement process is tuned at will.

Let us also recall from [5] that the implementation of the fixed-point scheme includes an additional stabilization term associated with dynamic gradient flows, that essentially translates to having a pseudo timestep in the Euler–Lagrange equations (2.3), that then read, knowing  $\mathbf{u}^k$ , for  $k = 1, \dots$ , solve

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

Further details can be found in [5, Appendix C]. Therefore the primal and mixed Galerkin methods, as well as the a posteriori error indicators  $\Theta$  and  $\Psi$ , are modified accordingly and only affect, the residual terms associated with the momentum equation. The Picard iterations with pseudo timestepping are located inside the adaptive refinement loop, which consists of solving, estimating, marking, and refining. All numerical simulations were run on a laptop computer using a single core in an i7-4750HQ processor (2.0 GHz 6 MB L3 cache).

We stress that hypothesis  $\alpha C_m L_F < 1$  can be rarely verified in practice and should be instead used to understand that the lack of uniqueness is mainly driven by the Lipschitz constant  $L_F$ , which can be computed from the images  $R$  and  $T$ . As  $\alpha$  is a modeling choice, the actual regularization is handled by a sufficiently small pseudo timestep  $\delta t$ , which renders the problem convex. A solution to the registration problem in such a time-dependent context corresponds to the limiting steady state of the problem. The explicit treatment of the nonlinearity given by  $\mathbf{f}_{\mathbf{u}^k}$  imposes a CFL condition, which requires  $\delta t$  to be possibly smaller than what is required for convexity. In this regard, it has been shown that it suffices to consider



$\delta t \propto \alpha^{-1}$  in both primal and mixed formulations, which indeed leads to convergent numerical schemes [7].

**5.2. Example 1: Registration of smooth synthetic images.** We assess the accuracy of the primal and mixed DIR methods using a smooth synthetic image under a smooth transformation. To this end, we define the reference image  $R : [0, 1]^2 \rightarrow \mathbb{R}$  by

$$(5.1) \quad R(x_1, x_2) = \sin(2\pi x_1) \sin(2\pi x_2).$$

We further define a manufactured displacement and the corresponding stress and rotation tensor fields by

$$(5.2) \quad \mathbf{u}(x_1, x_2) = \begin{pmatrix} 0.1 \cos(2\pi x_1) \sin(2\pi x_2) + \frac{x_1^2(1-x_1)^2 x_2^2(1-x_2)^2}{2\lambda} \\ -0.1 \sin(2\pi x_1) \cos(2\pi x_2) + \frac{x_1^3(1-x_1)^3 x_2^3(1-x_2)^3}{2\lambda} \end{pmatrix},$$

$$\boldsymbol{\sigma}(x_1, x_2) = \mathcal{C}\mathbf{e}(\mathbf{u}), \quad \boldsymbol{\rho}(x_1, x_2) = \frac{1}{2}(\nabla \mathbf{u} - \nabla \mathbf{u}^t).$$

Then, we construct a synthetic target image via composition of the reference image and the inverse warping, namely  $T = R \circ (\mathbf{id} + \mathbf{u})^{-1}$ . 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 manufactured solutions satisfy the zero-traction boundary condition, and they are used to construct an additional body load (apart from  $\mathbf{f}_u$ ) that needs to be incorporated as the 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 timestep  $\delta t = 0.1\alpha^{-1}$ . The initial mesh considered elements with edge length of 0.1768. The primal adaptive scheme considered  $\gamma_{ratio} = 0.01$ , while the mixed adaptive scheme used  $\gamma_{ratio} = 0.05$ .

On sequences of uniformly or adaptive refined meshes, we solve the DIR problem with primal and mixed methods and compute (nonnormalized) errors between the approximate and exact solutions in their natural norms, that is, for the primal method  $\mathbf{e}_u = \|\mathbf{u} - \mathbf{u}_h\|_{1,\Omega}$ ; whereas for the mixed method  $\mathbf{e}_u = \|\mathbf{u} - \mathbf{u}_h\|_{0,\Omega}$  and  $\mathbf{e}_\rho = \|\boldsymbol{\rho} - \boldsymbol{\rho}_h\|_{0,\Omega}$ ,  $\mathbf{e}_\sigma = \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\text{div},\Omega}$ . We also point out that in the case of adaptive mesh refinement, the experimental rates of convergence  $\widehat{\text{rate}}$  are computed differently than in the uniform case

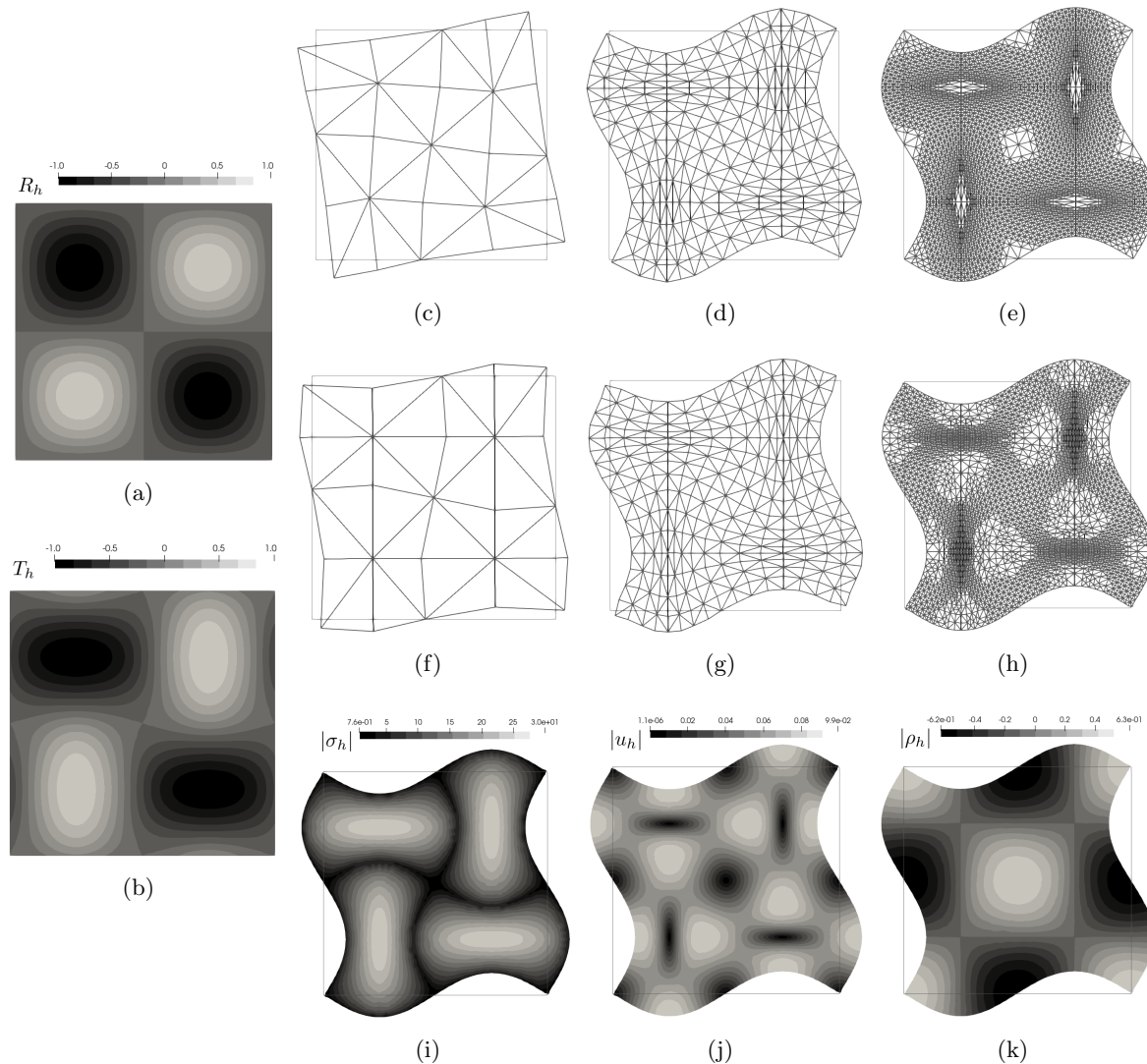
$$\widehat{\text{rate}} = \log(\mathbf{e}/\widehat{\mathbf{e}})[\log(h/\widehat{h})]^{-1}, \quad \widehat{\text{rate}} = -2 \log(\mathbf{e}/\widehat{\mathbf{e}})[\log(\text{DoF}/\widehat{\text{DoF}})]^{-1},$$

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

$$\text{eff}(\Theta) = \frac{\lambda \mathbf{e}_u}{\Theta}, \quad \text{eff}(\Psi) = \frac{\{\mathbf{e}_\sigma^2 + \mathbf{e}_u^2 + \mathbf{e}_\rho^2\}^{1/2}}{\Psi},$$

where the additional scaling (with the dilation modulus  $\lambda$ ) for the indicator  $\Theta$  is motivated by the fact that the efficiency bound arising from the proof of Lemma 4.5 is proportional to  $\lambda$  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.

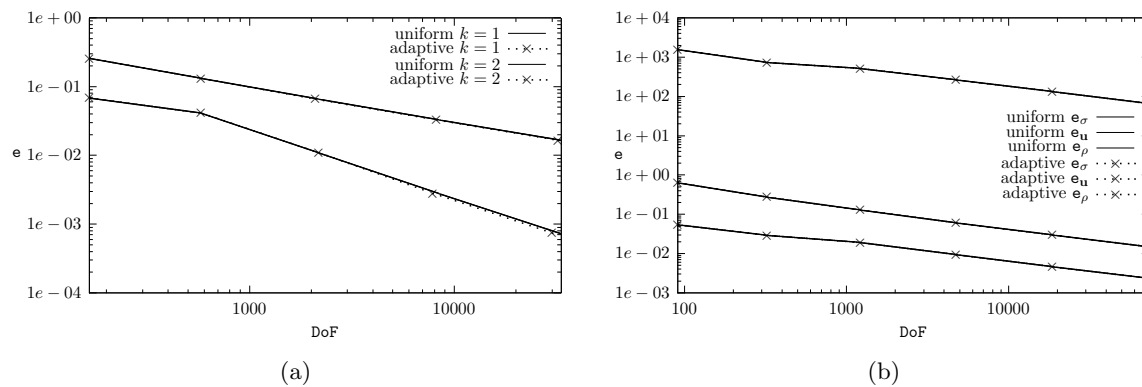
In Figure 5.1(a), (b) we show the reference image  $R_h$  and the resampled image  $T_h = T(\mathbf{x} + \mathbf{u}_h(\mathbf{x}))$ , and the panels (c)–(h) show examples of meshes adaptively refined guided by the estimators. We note that the primal method refines largely around the center of the



**Figure 5.1.** Example 1: adaptive mesh refinement in the registration of a smooth synthetic images. (a), (b) Projected fields of the reference  $R$  and composed  $T(\mathbf{x} + \mathbf{u}_h(\mathbf{x}))$  images; (c), (d), (e) evolution of the mesh adaption for the primal scheme using the error indicator  $\Theta$ ; (f), (g), (h) evolution of the mesh adaption for the mixed scheme using the error indicator  $\Psi$ ; (i), (j), (k) stress, displacement, and rotation norm fields predicted by the mixed scheme using mesh adaptivity.

domain. We also show in panels (i)–(j) 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.

The numerical convergence of the primal and mixed DIR methods is shown in Figure 5.2(a) and Figure 5.2(b), respectively. We observe that both methods do exhibit monotonic convergence. For this particular example, no major differences arise between the uniform and adaptive refinement schemes. Convergence rates for both methods are reported in Table 5.1, where we verify that optimal convergence rates are achieved with  $O(h^k)$ . No differences are observed in the number of Picard iterations required by the uniform and adaptive refinement strategies. The mixed DIR method also displays optimal convergence rates; see Table 5.2. We further note that the effectivity index values for the mixed scheme are roughly constant and



**Figure 5.2.** Example 1, smooth synthetic image registration example: Error convergence with respect to the number of degrees of freedom for both (a) primal and (b) mixed DIR formulations. Uniform refinement is shown in solid lines, while the adaptive refinement is shown in dotted lines.

**Table 5.1**

Example 1, smooth synthetic image registration example: error measures, convergence rates, and Picard iteration count for the approximate displacements  $\mathbf{u}_h$  produced with the primal method (of polynomial degrees  $k=1$  and  $k=2$ ) and tabulated according to the resolution level. (a) Uniform mesh refinement, (b) adaptive mesh refinement based on error estimator  $\Theta$ , with  $\gamma_{\text{ratio}} = 0.1$ , also displaying the rescaled effectivity index.

(a) Primal method, uniform refinement					(b) Primal method, adaptive refinement					
$k$	DoF	$h$	rate	iter	$k$	DoF	$h_{\min}$	$\widehat{\text{rate}}$	$\text{eff}(\Theta)$	iter
1	165	0.1768	1.403	15	1	165	0.1768	1.403	0.7875	15
	581	0.08839	1.073	16		565	0.08839	1.1	0.7306	16
	2181	0.04419	1.058	14		2069	0.04419	1.074	0.7056	14
	8453	0.0221	1.028	12		7525	0.0221	1.06	0.6918	11
	33285	0.01105	1.013	8		21957	0.01105	0.8978	0.6658	8
2	581	0.1768	0.789	11		581	0.1768	0.789	0.3317	11
	2181	0.08839	2.028	7		2117	0.08839	2.066	0.3353	14
	8453	0.04419	2.019	9		7237	0.04419	2.036	0.334	12
	33285	0.0221	1.947	9		7749	0.04419	3.033	0.3385	10
	132101	0.01105	1.364	9		25221	0.0221	1.596	0.3549	10

Table 5.2

Example 1, smooth synthetic image registration example: convergence rates, and Picard iteration count for the approximate Cauchy stress, displacements, and rotation  $\sigma_h, \mathbf{u}_h, \rho_h$  for the mixed formulations. (a) Uniform mesh refinement, (b) adaptive mesh refinement guided by  $\Psi$ , with  $\gamma_{ratio} = 0.05$ .

(a) Mixed method, uniform refinement						(b) Mixed method, adaptive refinement						
DoF	$h$	$rate_\sigma$	$rate_u$	$rate_\rho$	iter	DoF	$h_{min}$	$\widehat{rate}_\sigma$	$\widehat{rate}_u$	$\widehat{rate}_\rho$	$eff(\Psi)$	iter
323	0.3536	1.214	1.009	1.355	7	323	0.3536	1.214	1.009	1.355	1.004	7
1219	0.1768	0.5156	0.613	1.143	7	1219	0.1768	0.5156	0.613	1.143	0.9991	7
4739	0.08839	1.001	1.06	1.107	7	4739	0.08839	1.001	1.06	1.107	0.9959	7
18691	0.04419	1.005	1.028	1.051	10	17187	0.04419	1.034	1.059	0.9751	0.9944	10
74243	0.0221	1.004	1.01	1.019	10	64323	0.0221	1.013	1.023	0.8751	0.9922	10

Table 5.3

Example 1, volume-constrained smooth synthetic image registration: error measures, convergence rates, and Picard iteration count for the approximate displacements  $\mathbf{u}_h$  produced with the primal method. (a) Uniform mesh refinement, (b) adaptive mesh refinement based on error estimator  $\Theta$ , with  $\gamma_{ratio} = 0.1$ , also displaying the rescaled effectivity index.

(a) Primal method, uniform refinement					(b) Primal method, adaptive refinement				
DoF	$h$	rate	iter		DoF	$h_{min}$	$\widehat{rate}$	$eff(\Theta)$	iter
165	0.1768	1.37	8		165	0.1768	1.37	10330	12
581	0.08839	1.025	10		581	0.08839	1.025	9843	11
2181	0.04419	0.9901	10		2101	0.04419	1.02	9854	12
8453	0.0221	0.8515	10		8165	0.0221	0.8495	10960	12
33285	0.01105	0.5709	10		32013	0.01105	0.5718	14630	13

close to 0.43 and that the convergence rate is not substantially improved by the adaptivity in this example.

We studied the ability of the proposed DIR adaptive schemes to handle constrained problems by considering a volume-constrained transformation. To this end, we consider the same images defined in (5.1) and transformation (5.2), where we chose  $\nu = 0.49999$  to reflect the nearly incompressible behavior of the displacement field. The primal DIR method did not converge when using iterative methods, so we solved the linear system using a direct method. In contrast, the mixed DIR method did not exhibit any convergence issues using iterative methods. The convergence of the primal DIR method under uniform and adaptive refinement is shown in Table 5.3. Under both schemes, the convergence is suboptimal as the number of DoFs increases. In contrast, the mixed DIR method under both uniform and adaptive refinement shows optimal convergence rates; see Table 5.4. Figure 5.3 shows the reference and resampled target images for the mixed DIR method, along with the adaptive meshes for both formulations. Adaptive refinement in this case results in a rather uniform mesh with some small regions of larger mesh size, similarly to the trend observed in the nonconstrained case; see Figure 5.1.

**5.3. Example 2: Registration of smooth synthetic images with high gradients.** Next, we modify the closed-form displacement of Example 1 to produce higher gradients in the

Table 5.4

Example 1, volume-constrained smooth synthetic image registration example: convergence rates, and Picard iteration count for the approximate Cauchy stress, displacements, and rotation  $\sigma_h, \mathbf{u}_h, \rho_h$  for the mixed formulations. (a) Uniform mesh refinement, (b) adaptive mesh refinement guided by  $\Psi$ , with  $\gamma_{ratio} = 0.05$ .

(a) Mixed method, uniform refinement						(b) Mixed method, adaptive refinement						
DoF	$h$	$rate_\sigma$	$rate_u$	$rate_\rho$	iter	DoF	$h_{min}$	$\widehat{rate}_\sigma$	$\widehat{rate}_u$	$\widehat{rate}_\rho$	$eff(\Psi)$	iter
323	0.3536	1.215	1.012	1.249	10	323	0.3536	1.215	1.012	1.249	0.9943	6
1219	0.1768	0.5167	0.6176	1.16	10	1219	0.1768	0.5167	0.6176	1.16	0.9709	7
4739	0.08839	1.002	1.056	1.09	10	4739	0.08839	1.002	1.056	1.09	0.959	7
18691	0.04419	1.006	1.025	1.04	10	18691	0.04419	1.006	1.025	1.04	0.9313	7
74243	0.0221	1.004	1.009	1.015	10	73512	0.0221	1.01	1.016	0.9966	0.8385	7

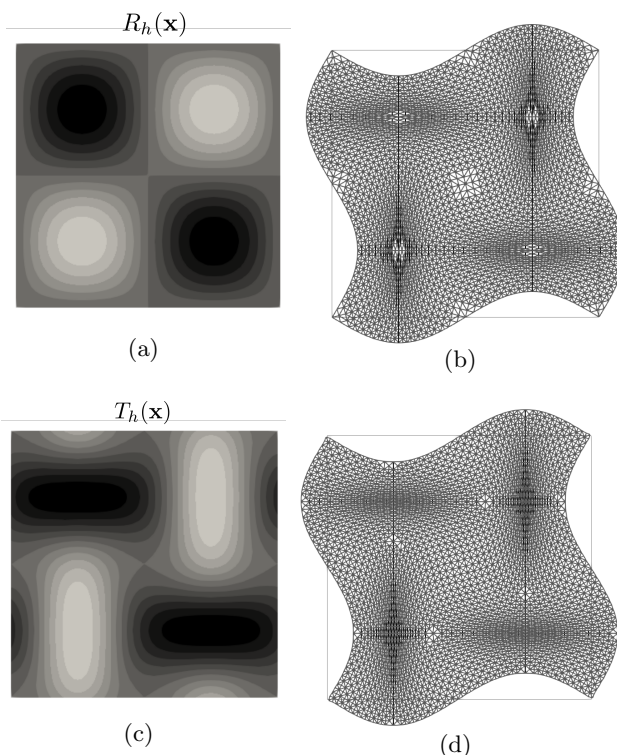


Figure 5.3. Example 1: Volume-constrained DIR. (a) Reference image, (b) mesh adaptivity in the primal method, (c) resampled target image for the mixed DIR method, and (d) mesh adaptivity in the mixed method.

reference image and initial target image. To this end, we consider the following image and displacement field expressions:

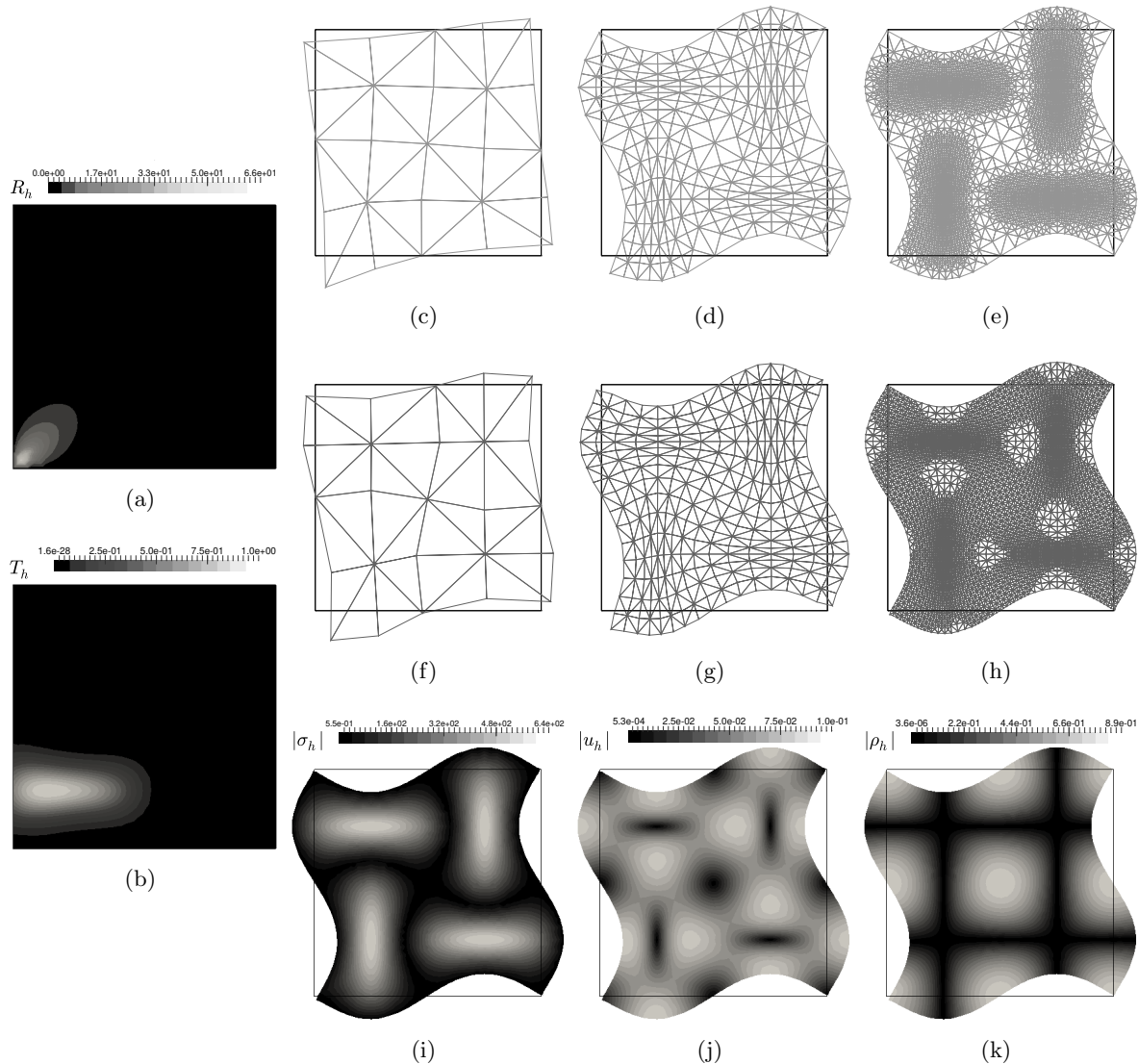
$$\mathbf{u}(x_1, x_2) = \begin{pmatrix} 0.1 \cos(\pi x_1) \sin(\pi x_2) + \frac{x_1^2(1-x_1)^2 x_2^2(1-x_2)^2}{2} \\ -0.1 \sin(\pi x_1) \cos(\pi x_2) + \frac{x_1^3(1-x_1)^3 x_2^3(1-x_2)^3}{2} \end{pmatrix}, \quad R(x_1, x_2) = \frac{x_1 x_2 (x_1 - 1)(x_2 - 1)}{(x_1 + 0.01)^4 + (x_2 + 0.01)^4},$$

$$T_0(x_1, x_2) = e^{-50[(x_1 - 0.2)^2 + (x_2 - 0.2)^2]}.$$

All other remaining model and mesh parameters are kept the same as in Example 1.



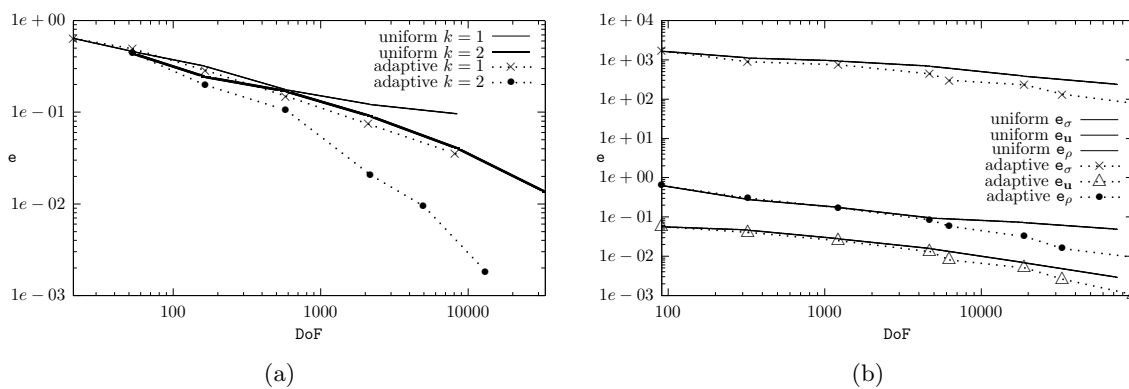
We show in Figure 5.4(a), (b) synthetic images projected onto the space of piecewise linear and continuous functions, as well as a few adapted meshes produced with the indicators (c)–(h), where one sees that the agglomeration of vertices occurs not so much due to the high gradients of the synthetic images, 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. We note that for the adaptive algorithm with  $\gamma_{\text{ratio}} = 0.01$ , the error indicator makes



**Figure 5.4.** Example 2: adaptive mesh refinement in the registration of smooth synthetic images with high gradients. (a), (b) Reference image  $R$  and composed target image  $T(\mathbf{x} + \mathbf{u}_h(\mathbf{x}))$ ; (c), (d), (e) evolution of the mesh adaption for the primal DIR method using the error indicator  $\Theta$ ; (f), (g), (h) evolution of the mesh adaption for the mixed DIR method using the error indicator  $\Psi$ ; (i), (j), (k) stress, displacement, and rotation norm fields predicted by the mixed scheme using mesh adaptivity.

the refinement to be applied uniformly for the first three iterations, after which localized meshing takes place in certain regions of the domain.

Figure 5.5(a) shows the numerical convergence of the primal DIR method under uniform and adaptive refinement. We observe monotonic convergence for all displacement field error as the number of DoFs increases. A notable improvement in convergence is observed for the particular case of the adaptive refinement scheme using second-order element interpolations. Convergence rates for the primal DIR method using uniform and adaptive refinement are reported in Table 5.5, where we observe that the case of adaptive refinement using second-order elements results in convergence rates that reach  $k = 2$ , which is notoriously higher than the convergence rate of  $k = 1.5$  reached by the primal method under uniform refinement. For the case of the mixed method, adaptive refinement always results in better convergence than



**Figure 5.5.** Example 2: Error convergence for (a) primal DIR method and (b) mixed DIR method under uniform and adaptive mesh refinement.

**Table 5.5**

Example 2. Convergence rates, and Picard iteration count for the approximate displacements  $\mathbf{u}_h$  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  $\Theta$  with  $\gamma_{\text{ratio}} = 0.01$  (b), also displaying the rescaled effectivity index).

(a) Primal method, uniform refinement

$k$	DoF	$h$	rate	iter
1	53	0.3536	0.481	4
	165	0.1768	0.526	6
	581	0.0884	0.859	19
	2181	0.0442	0.793	24
	8453	0.0221	0.620	28
2	165	0.3536	0.844	4
	581	0.1768	0.529	6
	2181	0.0884	0.900	20
	8453	0.0442	1.169	25
	33285	0.0221	1.564	29

(b) Primal method, adaptive refinement

$k$	DoF	$h_{\min}$	$\widehat{\text{rate}}$	$\text{eff}(\Theta)$	iter
1	53	0.3536	0.617	0.8314	4
	165	0.1768	1.277	0.8282	6
	581	0.0884	1.037	0.8205	11
	2105	0.0442	1.084	0.8187	15
	8177	0.0221	1.099	0.8219	18
2	165	0.3536	1.398	1.6422	4
	581	0.1768	1.489	1.6926	6
	2181	0.0884	1.891	1.6360	12
	4959	0.0442	1.786	1.6799	15
	13129	0.0221	2.097	1.6492	18



Table 5.6

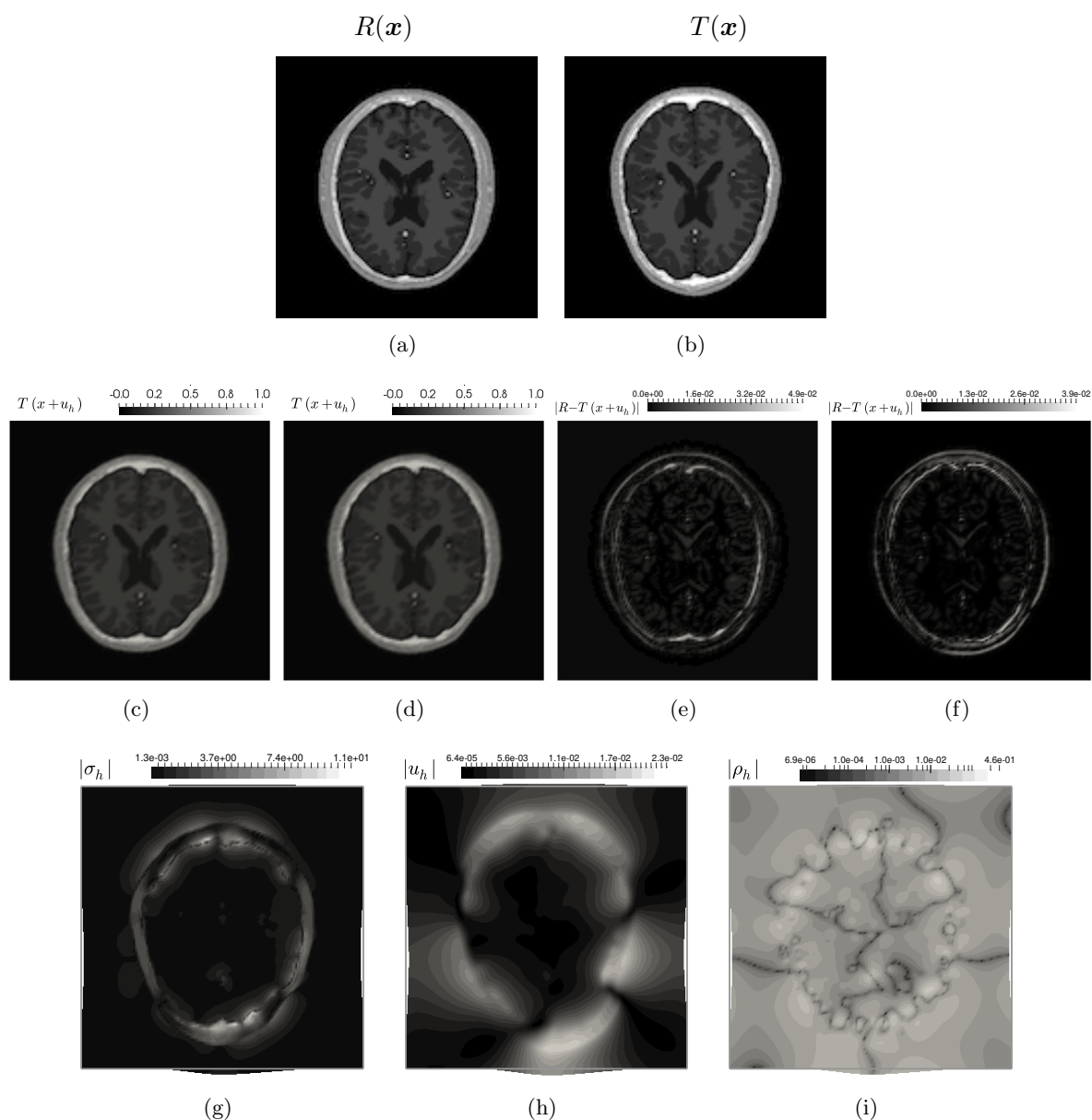
Example 2. Convergence rates, and Picard iteration count for the approximate Cauchy stress, displacements, and rotation  $\sigma_h, \mathbf{u}_h, \rho_h$  produced with the lowest-order mixed method, and tabulated according to the resolution level, under uniform (a) and adaptive mesh refinement guided by  $\Psi$  with  $\gamma_{\text{ratio}} = 0.009$  ((b), also displaying the effectivity index).

(a) Mixed method, uniform refinement						(b) Mixed method, adaptive refinement							
DoF	$h$	$\text{rate}_\sigma$	$\text{rate}_u$	$\text{rate}_\rho$	iter	DoF	$h_{\min}$	$h_{\max}$	$\widehat{\text{rate}}_\sigma$	$\widehat{\text{rate}}_u$	$\widehat{\text{rate}}_\rho$	$\text{eff}(\Psi)$	iter
323	0.3536	0.552	0.278	1.174	7	323	0.3536	0.3536	0.965	0.518	1.169	0.5333	5
1219	0.1768	0.278	0.728	0.674	13	1219	0.1768	0.1768	0.955	0.725	0.869	0.5272	8
4739	0.0884	0.443	0.846	0.882	28	4692	0.0742	0.1250	0.952	0.946	1.002	0.5188	19
18691	0.0442	0.741	1.183	0.658	45	6277	0.0264	0.1250	1.066	1.114	1.106	0.5139	21
74243	0.0221	0.598	1.235	0.606	50	18884	0.0107	0.0817	1.052	1.039	1.067	0.5205	24
						32998	0.0051	0.0730	0.986	0.958	0.975	0.5216	30
						94153	0.0020	0.0601	0.961	0.963	0.959	0.5210	30

uniform refinement for the displacement, stress and rotation fields; see Figure 5.5(b). Table 5.6 reports the convergence rates of the mixed method, where we note that the adaptive refinement always results in rates that are greater than those obtained under uniform refinement. Further, we observe that in systems with a roughly similar number of DoFs, the number of Picard iterations needed to reach the tolerance is smaller in the case of adaptive refinement.

**5.4. Example 3: Registration of brain medical images.** We now turn to the application of the adaptive primal and mixed DIR methods in the registration of medical images of human brains [18]. The reference and target images for the brain have dimensions  $258 \times 258$  and the voxel resolution corresponds to 1 mm; see top panels in Figure 5.6. 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 5.6 depict the composed images  $T \circ (id + \mathbf{u}_h)$  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,  $|R(\mathbf{x}) - T(\mathbf{x} + \mathbf{u}_h(\mathbf{x}))|$  resulting from both methods.

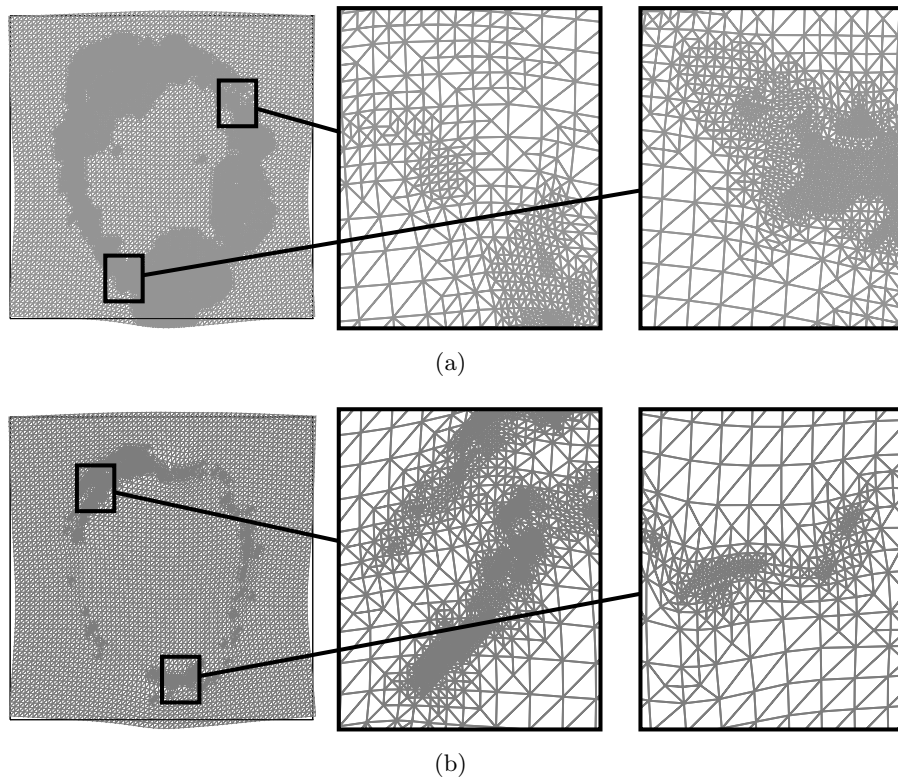
We also plot an example of a mesh obtained after four steps of adaptive refinement with the primal and mixed methods (see Figure 5.7). For illustration purposes we initiate the process from a coarse mesh of 8196 triangles (corresponding to a low resolution image of  $64 \times 64$  pixels; starting with images of higher resolution implies 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 5.6, where the Frobenius norm of the rotation tensor is plotted in log-scale for clarity).



**Figure 5.6.** Example 3. Registration of brain medical images. (a) Reference image; (b) target image; (c), (d) resampled (composed) images from solutions using primal and mixed schemes, respectively; (e), (f) similarity plots resulting from primal and mixed schemes, respectively; (g), (h), (i) stress, displacement, and rotation norm fields resulting from the mixed DIR scheme using adaptive mesh refinement.

On the other hand, the displacements are, in comparison, rather smooth and they seem not to contribute substantially to the local error indicators.

In Table 5.7 we report information about the CPU time required in each step of the overall solution algorithm. We record the wall-time during the execution of the mixed and primal DIR methods, when starting from a coarse grid (representing 8715 DoFs for the primal



**Figure 5.7.** Example 3. Adaptive mesh refinement in the registration of brain medical images. (a) Mesh after four steps of adaptive refinement using the error indicator  $\Theta$  for the primal DIR method; (b) mesh after four steps of adaptive refinement using the error indicator  $\Psi$  for the mixed DIR method.

**Table 5.7**

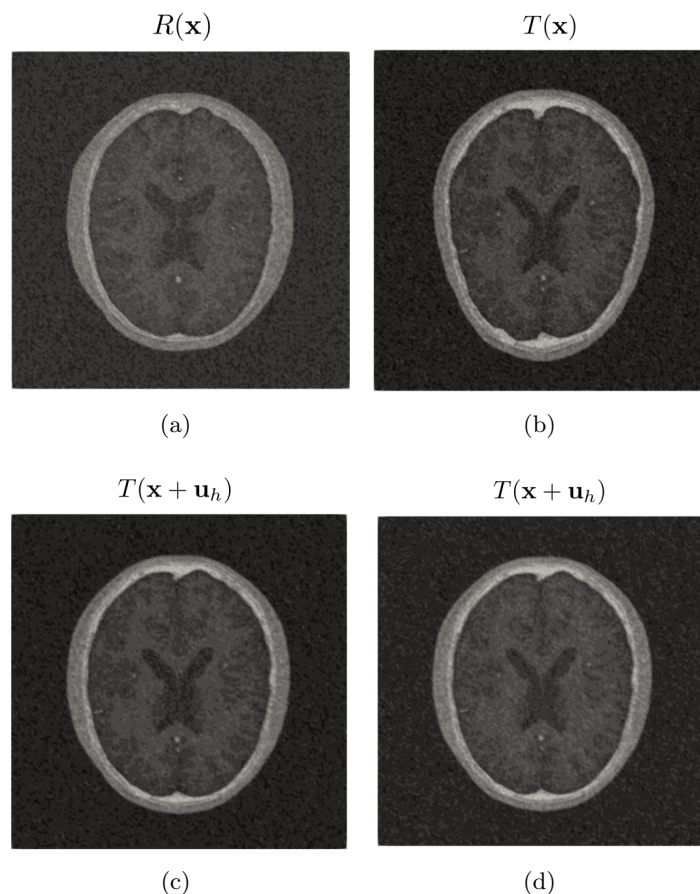
Example 3. CPU time (in [s]) of each step of the adaptive finite element method for the DIR problem, measured for the primal and mixed methods, starting from coarse meshes. The time associated with the solution of the linear systems is averaged over the number of inner Picard iterations.

	Refin. level	Matrix assembly	Solution computation	IO and residual	Evaluation of estimator	Marking and refinement
Primal method (total CPU time: 73.16)	1	0.101	0.075 (avg)	0.102	0.096	0.544
	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 (total CPU time: 997.83)	1	0.418	1.445 (avg)	0.101	0.099	0.530
	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

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.

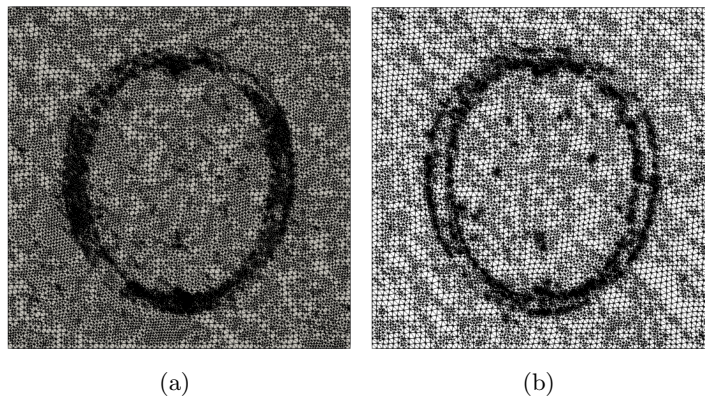
To assess the performance of the adaptive refinement DIR schemes on images under the presence of noise, we considered the case of brain images with Gaussian noise (zero mean, standard deviation of 0.1); see Figures 5.8(a) and 5.8(b) for the resulting reference and target images. The numerical solution of this problem included the use of Anderson acceleration at each time step for improved convergence [45]. Figures 5.8(c) and 5.8(d) show the resampled target images for the primal and mixed schemes, which are both similar to the reference image, confirming the ability of the proposed adaptive methods to handle Gaussian noise at moderate levels. The resulting meshes for the primal and mixed schemes are shown in Figures 5.9(a) and 5.9(b), respectively. Similar to the case of noise-free brain images, mesh refinement predominantly concentrates in the region where the skull is located. We note, however, that noise also increases the level of refinement in other regions of the brain image that do not have anatomical changes (e.g., area outside the skull). The mixed scheme displays an attenuation of this spurious refinement under noise.

**5.5. Example 4: Registration of binary images.** The last example of application addressed in this study consists of a classic benchmark in DIR which introduces two important challenges. First, reference and target images are binary-composed, i.e., they have intensity values of either



**Figure 5.8.** Example 3: Adaptive DIR of brain images with Gaussian noise; the variance is  $\sigma = 0.1$ . (a) Reference image, (b) target image, (c) resampled target image, primal scheme, and (d) resampled target image, mixed scheme.





**Figure 5.9.** Example 3: Adaptive mesh refinement in brain images under noise. (a) Primal scheme, and (b) mixed scheme.

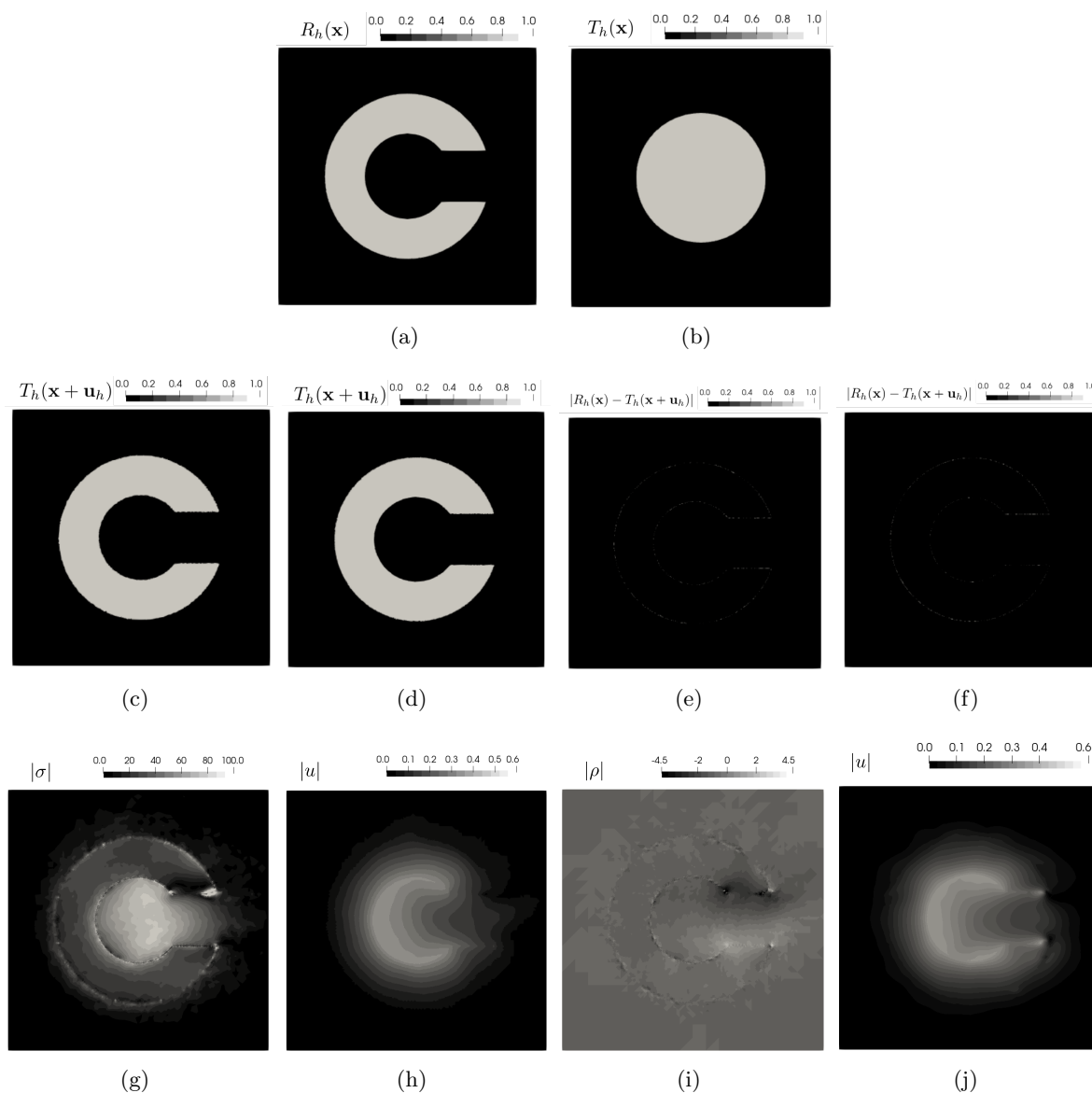
0 or 1, which creates steep numerical gradients at the binary interface of order  $1/h$ . Thus, the images do not satisfy condition (2.5). Second, the transformation between images involves large displacements and deformations, which questions the validity of the elastic strain energy from a physical viewpoint. We define the ball  $B(\mathbf{x}, r) = \{\mathbf{x} \in \mathbb{R}^2 : |\mathbf{x}| \leq r\}$  to set the images as

$$R(\mathbf{x}) = \begin{cases} 1 & \mathbf{x} \in B((0.5, 0.5), 0.32) \cap [B((0.5, 0.5), 0.16)]^c \cap [\{x_1 > 0.5\} \cap \{0.4 < x_2 < 0.6\}]^c, \\ 0 & \text{otherwise,} \end{cases}$$

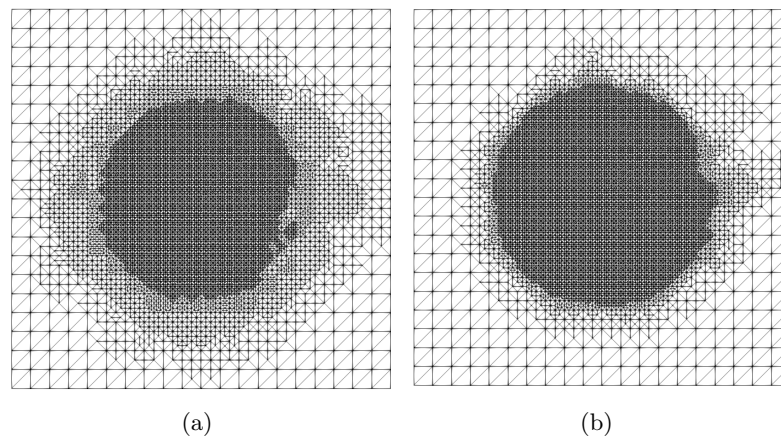
$$T(\mathbf{x}) = \begin{cases} 1 & \mathbf{x} \in B((0.5, 0.5), 0.25), \\ 0 & \text{otherwise.} \end{cases}$$

Both methods consider quadrature rules of sixth order, with an initial mesh given by a unit square with 20 elements per side, which yields a total of 800 triangular elements. In this example,  $\gamma_{ratio} = 0.075$ . We consider the parameters  $\alpha = 1000$ ,  $E = 15$ , and  $\nu = 0.3$  and set the pseudo timestep to  $\delta t = h_{\min}^2/\alpha$  for the primal case and  $\delta t = 0.01 h_{\min}^2/\alpha$ , where  $h_{\min}$  is the minimum characteristic length of the mesh. This was motivated by a possible CFL condition on the timestep arising from the explicit treatment of the nonlinearity and proved effective during numerical tests. The convergence was set through the  $\ell^\infty$  norm of the increment  $|\mathbf{u}^k - \mathbf{u}^{k-1}|_{\ell^\infty}$  with a tolerance of  $h_{\min}$ , so that iterations stop when the displacement changes by less than the smallest element. A maximum number of 100 iterations was always achieved, following previous works addressing this problem [34]. Both the primal and mixed DIR problems for this example were solved in serial with the iterative scheme BiCGStab preconditioned with an incomplete LU factorization, using the default parameters available in FEniCS. The solution of the mixed DIR problem required a considerable numerical effort to converge to a solution that met the error criterion. To overcome this difficulty, we used at each refinement level the solution of the primal formulation as an initial solution for the mixed case, and then employed five iterations of the mixed formulation only. This was already implemented in [5] to substantially improve the registration of lung images in a mixed formulation.

We report the solution with its components in Figure 5.10. In the first row we show the reference (a) and target (b) images, constructed as in [34], with the solution reported in the second row together with its absolute error  $|R(\mathbf{x}) - T(\mathbf{x} + \mathbf{u}_h(\mathbf{x}))|$  in primal (c), (e) and mixed (d), (f) form. We note that the mixed DIR performs slightly worse than the primal DIR method, which is to be expected due to the lower order of approximation used. The last



**Figure 5.10.** Example 4. Registration of binary images (O-C). (a) Reference image, (b) target image, (c), (d) resampled (composed) images from solutions using primal and mixed schemes, respectively; (e), (f) similarity images resulting from the primal and mixed methods, respectively; (g), (h), (i) stress, displacement, and rotation norm fields using the adaptive mixed DIR method; (j) displacement norm field using the adaptive primal DIR method.



**Figure 5.11.** Example 4. Registration of binary images. Mesh after three steps of adaptive refinement for (a) primal DIR problem and (b) mixed DIR problem.

**Table 5.8**

Example 4. 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.

	Refin. level	Matrix assembly	Solution computation	IO and residual	Evaluation of estimator	Marking and refinement
Primal method (total CPU time: 261.83)	1	0.018	0.036 (avg)	0.024	0.017	0.010
	2	0.031	0.076 (avg)	0.029	0.023	0.023
	3	0.293	0.181 (avg)	0.049	0.054	0.081
	4	0.293	0.627 (avg)	0.130	0.140	0.202
	5	0.751	2.381 (avg)	0.338	0.466	0.799
Mixed method (total CPU time: 326.61)	1	0.04	0.051 (avg)	0.058	0.016	0.006
	2	0.248	0.112 (avg)	0.037	0.047	0.064
	3	0.777	0.361 (avg)	0.037	0.047	0.064
	4	2.781	1.256 (avg)	0.142	0.147	0.243
	5	10.725	5.104 (avg)	0.464	0.545	1.031

row shows the magnitude of all components of the solution, in both primal (j) and mixed (g), (h), (i) formulations. Also, in Figure 5.11 we present the refined mesh after three steps, where it can be observed how the mixed scheme yields a more localized refinement even though the amount of refined elements is the same in both schemes. Finally, we provide information on the CPU time required in each step of the overall solution algorithm in Table 5.8.

**6. Discussion.** In this work, we present a novel adaptive mesh-refinement scheme for the numerical solution of primal and mixed DIR problems. Our method hinges upon the development of a posteriori error estimators for both the primal and mixed finite-element formulations that are reliable and efficient, and at the same time, they are easily computed. These estimators allow for an optimal refinement of the mesh in zones where the accuracy of the numerical approximation does not perform well. Thus, one distinctive feature of our work is the effectiveness of the mesh-adaption strategy, as they are justified on selectively reducing



the local approximation error made by the finite-element schemes employed. This contrasts with current methods of mesh adaptation employed in DIR problems, which either refine the discretization uniformly or rely on heuristic grounds to select regions that are refined.

To assess the numerical performance of the proposed method, we employ uniform and adaptive mesh refinement to solve a DIR problem based on smooth synthetic images where the displacement solution is known in advance. In Example 1, we show that both the uniform and adaptive refinement schemes display optimal convergence, and no marked differences in convergence rate arise between them. In contrast, adaptive refinement results in higher convergence rates than those delivered by uniform refinement in Example 2; see Figure 5.2. This improvement in convergence is explained by the ability of the adaptive method to refine regions where gradients in the image and in the displacement fields are considerably higher than the average gradients in the image, while keeping coarser elements in regions where the gradients are low. Such improvement is not observed in Example 1 because the magnitude of gradients was not as high nor as localized as in Example 1. It is important to remark that the numerical convergence obtained in these cases strongly relies on the smoothness of the chosen images and fabricated solution. The use of smooth and bounded images and the existence of a regular solution ensures that the Lipschitz and upper bound constants  $L_F$  and  $M_F$  are well defined and take on reasonable values. Further, the existence of a regular solution is crucial to prove the approximation properties of the finite-element scheme. The convergence properties shown in this example may not directly extend to real-life applications, where images may display high gradients, and the solution may not be regular.

We studied the ability of the proposed refinement schemes to handle constrained DIR problems. In particular, Example 1 deals with a volume-preserving problem, which is achieved by considering an elastic energy that strongly penalizes volumetric deformations. Our results show that the primal DIR method displays suboptimal convergence (see Table 5.3). This poor behavior is expected and is similar to that found in traditional elasticity problems where material incompressibility ( $\nu \rightarrow 0.5$ ) results in volumetric-locking phenomena [12]. Notably, the mixed DIR scheme adequately handles this volumetric constraint, delivering considerably better results with optimal convergence rates; see Table 5.4. These findings confirm that the mixed DIR formulation and the proposed adaptive scheme are suitable to handle volume-constrained DIR problems.

To demonstrate the applicability of the adaptive refinement method in medical images, we perform DIR on human brain images; see Figure 5.6. In this case, both the primal and mixed methods using adaptive refinement deliver acceptable results, as shown in the images of the similarity measure; see Figure 5.6(e), (f). Further, the mixed DIR method delivers the stress field, which shows that most of the deformation occurs in the vicinity of the skull. These regions are indeed the ones where most of the mesh refinement takes place, based on the error estimators  $\Theta$  and  $\Psi$ ; see Figure 5.7. Table 5.7 reports the distribution of computing time among different tasks in the solution of the DIR problem using adaptive refinement. We conclude that, for this example, the cost of evaluating the local estimator for the primal method (also including assembly of the global one) is roughly half of the time spent in the initial assembly of the left-hand side of the matrix systems, representing around 15% of the time spent in marking and refining. Also, the evaluation of the estimator in the mixed DIR method appears to be faster than in the primal DIR method, which is possibly due to fewer

applications of numerical differentiation in constructing the residual terms. Future extensions of this work could consider the use of robust preconditioners that can reduce the computing time in the solution of the linear systems [32]. We also conclude that both DIR adaptive schemes result in overall increased levels of refinement when compared to the noise-free case; see Figure 5.9. Still, both schemes are able to concentrate the mesh refinement in the same regions observed in the noise-free case, delivering acceptable resampled images despite the presence of noise; see Figure 5.8

As a final example of application, we use the adaptive refinement scheme to register O-C binary images. We remark that this challenging problem is well addressed by the proposed adaptive method, as observed in the intensity figures showing the difference between the reference and the composed target images; see Figure 5.10(e), (f). The high accuracy of the DIR solution relies on the five levels of mesh refinement performed in the region around the solid circle in the target image; see Figure 5.11. In terms of CPU time, we observe that computing efforts are largely focused on solving the systems of equations. The remaining tasks do not take more than roughly 1% and 3% of the total CPU time in the primal and mixed methods, respectively. Since the solution time directly depends on the number of DoFs, we conclude that the proposed adaptive mesh refinement scheme can significantly reduce the computational effort needed to register images in real-life applications, as it delivers attractive reductions in the number of DoFs without compromising the accuracy of the registration.

**Appendix A. Proofs of robustness in the primal case.** In this appendix we include the proofs of reliability and efficiency stated in section 4.2.

#### Proof of Lemma 4.4.

*Proof.* Let us first define

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

As a consequence of the ellipticity of  $a$  (cf. (3.2)) with ellipticity constant  $\bar{\alpha}$  (cf. [12, Corollary 11.2.22]), we obtain the following condition:

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

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

$$(A.1) \quad \bar{\alpha} \|\mathbf{u} - \mathbf{u}_h\|_H \leq \sup_{\substack{\mathbf{w} \in H \\ \mathbf{w} \neq \mathbf{0}}} \frac{\mathcal{R}_h(\mathbf{w} - \mathbf{w}_h)}{\|\mathbf{w}\|_H} \quad \forall \mathbf{w}_h \in H_h.$$

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

(A.2)

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

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

$$(A.3) \quad \|\mathbf{w} - \mathbf{w}_h\|_{0,K} \leq c_1 h_K \|\mathbf{w}\|_{1,\Delta(K)}, \quad \|\mathbf{w} - \mathbf{w}_h\|_{0,e} \leq c_2 h_e \|\mathbf{w}\|_{1,\Delta(e)}.$$

In this way, applying the Cauchy-Schwarz inequality to each term (A.2), and making use of (A.3) together with the Lipschitz continuity of  $F_{\mathbf{u}}$  (cf. (3.3)), we obtain

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

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

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

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 (A.1), we conclude that

$$\|\mathbf{u} - \mathbf{u}_h\|_H \leq C_{\text{rel}} \Theta,$$

where  $C_{\text{rel}}$  is independent of  $h$ . ■

### Proof of Lemma 4.5.

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

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

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

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

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

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

We omit further details and repeating arguments used for the remaining inequalities. ■

## 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), pp. 9–23.
- [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), pp. 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), pp. 332–346.
- [4] 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, pp. 9252–9260.
- [5] 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), pp. 2529–2567.
- [6] 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), pp. 197–222.
- [7] N. BARNAFI, G. N. GATICA, D. E. HURTADO, W. MIRANDA, AND R. RUIZ-BAIER, *New primal and dual-mixed finite element methods for stable image registration with singular regularization*, Math. Models Methods Appl. Sci., 31 (2021), pp. 979–1020, <https://doi.org/10.1142/S021820252150024X>.
- [8] Z. BELHACHMI AND F. HECHT, *An adaptive approach for the segmentation and the TV-filtering in the optic flow estimation*, J. Math. Imaging Vision, 54 (2016), pp. 358–377.
- [9] Z. BELHACHMI AND F. HECHT, *Control of the effects of regularization on variational optic flow computations*, J. Math. Imaging Vision, 40 (2011), pp. 1–19.
- [10] B. BOURDIN, *Image segmentation with a finite element method*, ESAIM Math. Model. Numer. Anal., 33 (1999), pp. 229–244.
- [11] B. BOURDIN AND A. CHAMBOLE, *Implementation of an adaptive finite-element approximation of the Mumford-Shah functional*, Numer. Math., 85 (2000), pp. 609–646.
- [12] S. C. BRENNER AND L. R. SCOTT, *The Mathematical Theory of Finite Element Methods*, 3rd ed., Texts in Appl. Math. 15. Springer, New York, 2008.

- [13] F. BREZZI AND M. FORTIN, *Mixed and Hybrid Finite Element Methods*, Springer Ser. Comput. Math. 15, Springer, New York, 1991.
- [14] C. CARSTENSEN AND G. DOLZMANN, *A posteriori error estimates for mixed FEM in elasticity*, Numer. Math., 81 (1998), pp. 187–209.
- [15] 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), pp. 2155–2163.
- [16] 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*, J. Applied Physiology, 115 (2013), pp. 730–742.
- [17] P. CLÉMENT, *Approximation by finite element functions using local regularisation*, Rev. Française Automat. Informat. Recherche Opérationnelle Sér., 9 (1975), pp. 77–84.
- [18] 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.
- [19] 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), pp. 486–504.
- [20] W. DÖRFLER, *A convergent adaptive algorithm for Poisson’s equation*, SIAM J. Numer. Anal., 33 (1994), pp. 1106–1124.
- [21] M. EBNER, M. MODAT, S. FERRARIS, S. OURSELIN, AND T. VERCAUTEREN, *Forward-backward splitting in deformable image registration: A demons approach*, in Proceedings of the IEEE 15th International Symposium on Biomedical Imaging, Washington, DC, 2018, pp. 1065–1069.
- [22] G. N. GATICA, *A Simple Introduction to the Mixed Finite Element Method. Theory and Applications*, Springer Briefs in Math., Springer, Cham, 2014.
- [23] 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), pp. 1115–1130.
- [24] E. HABER, S. HELDMANN, AND J. MODERSITZKI, *Adaptive mesh refinement for nonparametric image registration*, SIAM J. Sci. Comput., 30 (2008), 3012–3027.
- [25] S. HAKER, L. ZHU, A. TANNENBAUM, AND S. ANGENENT, *Optimal mass transport for registration and warping*, Int. J. Comput. Vis., 60 (2004), pp. 225–240.
- [26] M. HINTERMÜLLER, A. LANGER, C. N. RAUTENBERG, AND T. WU, *Adaptive regularization for image reconstruction from subsampled data*, in Imaging, Vision and Learning Based on Optimization and PDEs, Math. Vis. 326, Springer, Cham, 2018.
- [27] M. HINTERMÜLLER AND M. RINCON-CAMACHO, *An adaptive finite element method in  $L^2$ -TV-based image denoising*, Inverse Probl. Imaging, 8(3) (2014), pp. 685–711.
- [28] B. K. HORN AND B. G. SCHUNCK, *Determining Optical Flow*, Tech. report, Cambridge, MA, 1980.
- [29] 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 Modeling Mechanobiology, 16 (2017), pp. 1413–1423.
- [30] 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 Trans. on Medical Imaging, 35 (2016), pp. 580–588.
- [31] D. E. HURTADO, B. ERRANZ, F. LILLO, M. SARABIA-VALLEJOS, P. ITURRIETA, F. MORALES, K. BLAHA, T. MEDINA, F. DIAZ, AND P. CRUCES, *Progression of regional lung strain and heterogeneity in lung injury: Assessing the evolution under spontaneous breathing and mechanical ventilation*, Ann. Intensive Care, 10 (2020), 107.
- [32] M. KUCHTA, K. A. MARDAL, AND M. MORTENSEN, *On the singular Neumann problem in linear elasticity*, Numer. Linear Algebra Appl., 26 (2019).
- [33] E. LEE AND M. GUNZBURGER, *An optimal control formulation of an image registration problem*, J. Math. Imaging Vis., 36 (2010), pp. 69–80.
- [34] J. MODERSITZKI, *Numerical Methods for Image Registration*, Numer. Math. Sci. Comput., Oxford Science Publications, New York, 2004.

- [35] A. PAWAR, Y. ZHANG, Y. JIA, X. WEI, T. RABCZUK, C. L. CHAN, AND C. ANITESCU, *Adaptive FEM-based nonrigid image registration using truncated hierarchical B-splines*, *Computers and Mathematics with Applications*, 72 (2016), pp. 2028–2040.
- [36] C. PÖSCHL, J. MODERSITZKI, AND O. SCHERZER, *A variational setting for volume constrained image registration*, *Inverse Probl. Imaging*, 4 (2010), pp. 50–522.
- [37] 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), pp. 591–599.
- [38] R. T. ROCKAFELLAR, *Monotone operators and the proximal point algorithm*, *SIAM J. Control Optim.*, 14 (1976), pp. 877–898.
- [39] A. SOTIRAS, C. DAVATZIKOS, AND N. PARAGIOS, *Deformable medical image registration: A survey*, *IEEE Trans. Medical Imaging*, 32 (2013), pp. 1153–1190.
- [40] G. UNAL AND G. SLABAUGH, *Coupled PDEs for non-rigid registration and segmentation*, in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 1, 2005, pp. 168–175.
- [41] R. VERFÜRTH, *A posteriori error estimation and adaptive mesh-refinement techniques*, *J. Comput. Appl. Math.*, 50 (1994), pp. 67–83.
- [42] R. VERFÜRTH, *A posteriori error estimation techniques for finite element methods*, *Numer. Math. Sci. Comput.*, Oxford University Press, Oxford, UK, 2013.
- [43] R. VERFÜRTH, *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*, Wiley-Teubner, Stuttgart, 1996.
- [44] R. VERFÜRTH, *A review of a posteriori error estimation techniques for elasticity problems*, *Comput. Methods Appl. Mech. Engrg.*, 176 (1999), pp. 419–440.
- [45] H. F. WALKER AND P. NI, *Anderson acceleration for fixed-point iterations*, *SIAM J. Numer. Anal.*, 49 (2011), pp. 1715–1735.
- [46] J. ZHANG, J. WANG, X. WANG, AND D. FENG, *The adaptive FEM elastic model for medical image registration*, *Physics Medicine Biology*, 59 (2013), 97.
- [47] B. ZITOVÁ AND J. FLUSSER, *Image registration methods: A survey*, *Image Vision Computing*, 21 (2003), 977–1000.