

Primal and mixed finite element methods for deformable image registration problems*

NICOLÁS BARNAFI[†] GABRIEL N. GATICA[‡] DANIEL E. HURTADO[§]

Abstract

Deformable image registration (DIR) represent a powerful computational method for image analysis, with promising applications in the diagnosis of human disease. Despite being widely used in the medical imaging community, the mathematical and numerical analysis of DIR methods remain understudied. Further, recent applications of DIR include the quantification of mechanical quantities apart from the aligning transformation, which justifies the development of novel DIR formulations where the accuracy and convergence of fields other than the aligning transformation can be studied. In this work we propose and analyze a primal, mixed and augmented formulations for the DIR problem, together with their finite-element discretization schemes for their numerical solution. The DIR variational problem is equivalent to the linear elasticity problem with a nonlinear source term that depends on the unknown field. Fixed point arguments and small data assumptions are employed to derive the well-posedness of both the continuous and discrete schemes for the usual primal and mixed variational formulations, as well as for an augmented version of the later. In particular, continuous piecewise linear elements for the displacement in the case of the primal method, and Brezzi-Douglas-Marini of order 1 (resp. Raviart-Thomas of order 0) for the stress together with piecewise constants (resp. continuous piecewise linear) for the displacement when using the mixed approach (resp. its augmented version), constitute feasible choices that guarantee the stability of the associated Galerkin systems. A-priori error estimates derived by using Strang-type Lemmas, and their associated rates of convergence depending on the corresponding approximation properties are also provided. Numerical convergence tests and DIR examples are included to demonstrate the applicability of the method.

Key words: image registration, primal formulation, mixed formulation, finite element methods

1 Introduction

Deformable image registration (DIR) concerns the problem of aligning two or more images by means of a transformation that allows for distortion (warping) of the images under analysis. Such problem arises in a number of important applications, particularly in the field of medical imaging [39]. DIR

*This work was partially supported by VRI Puente P1617 from Pontificia Universidad Católica de Chile, and CONICYT-Chile through BASAL project CMM, Universidad de Chile, and by Centro de Investigación en Ingeniería Matemática (CI²MA), Universidad de Concepción.

[†]Department of Structural and Geotechnical Engineering, School of Engineering, Pontificia Universidad Católica de Chile, Vicuña Mackenna 4860, Santiago, Chile, email: nabarnaf@uc.cl.

[‡]CI²MA and Departamento de Ingeniería Matemática, Universidad de Concepción, Casilla 160-C, Concepción, Chile, email: ggatica@ci2ma.udec.cl.

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

is commonly formulated as a variational problem, where the main ingredients of the DIR are i) the transformation model, a family of mappings that warp a target image into the reference image, ii) the similarity measure, a functional that weighs the differences between the reference image and the resampled target image, and iii) the regularizer, which renders the problem well-conditioned by adding regularity to the DIR solution.

Despite DIR is widely used in the medical imaging community, the mathematical and numerical analysis of DIR remains understudied. The DIR continuous problem has been formulated using mainly three approaches: minimization of similarity measures (with or without constraints), as an optimal mass transport problem [22, 31, 11], or as a level set segmentation-registration combined problem [42, 17]. The problem of minimizing similarity measures has been studied in [5, 43], where the direct method of calculus of variations has been used to establish existence of solutions. The optical flow formulation, an associated problem which can be seen as a sequence of registration problems in time, was proposed by Horn & Schunk in 1980 [23], and has been the subject of analysis from an optimal-control problem point of view [7, 27]. Well-posedness of optical flow schemes has been established for Dirichlet boundary conditions under reasonable assumptions [18, 41]. Besides providing existence and uniqueness of the solution, by assuming only uniform boundedness on the images, these studies show that the solution is a step-wise diffeomorphism, which is a desirable regularity property when it comes to warping images. The analysis of the numerical schemes proposed to solve similarity-minimization formulations has received less attention. A noteworthy approach is the work of Pöschl *et al.* [33], where both the continuous and discretized problems are analyzed, and a solution is found using a primal finite-element approximation that is shown to be convergent. However, the analysis is restricted to polyconvex energy densities (both for the similarity measure and regularizer) and volume-preserving transformations, and does not account for the convergence of the transformation gradients and stresses. A more traditional Galerkin approach has been introduced in [28] for optimal-control-based registration, but requires a considerable degree of regularity ($H^{2+\delta}$) of the target and reference image functions, not required by other traditional formulations. While most approaches to DIR problems are based on primal formulations, a mixed formulation of the similarity minimization problem has been proposed in the setting of fluid registration schemes [12, 35], where a sequence of incompressible Stokes problems are solved to find the optimal displacement and pressure fields. While directly solving for the pressure field, which is desirable to understand the mechanical behavior of the images being registered, limited analysis has been provided to understand the well-posedness of the continuous problem and convergence of numerical discretizations of mixed formulations of DIR problems that use elastic regularizers.

One important and recent application of DIR is the study of regional deformation of lung tissue from computed-tomography (CT) images of the thorax [14]. From a continuum-mechanics perspective, the optimal transformation \mathbf{u} obtained from the solution of the DIR problem can be considered as a displacement field, from which a strain tensor field can be computed using the gradient of the displacement field $\nabla\mathbf{u}$. The study of not only motion but also deformation at a regional level in the lung has revealed that deformation in the lung tissue can be highly heterogenous and anisotropic [3, 25], thus providing new deformation-based markers to understand lung physiology [13], showcasing the potential of DIR strain analysis as a tool in the detection and diagnosis of pulmonary disease. These advances notwithstanding, it has been recently shown that state-of-the-art strain analysis techniques based on direct differentiation of DIR solutions can be highly inaccurate as they are very sensitive to noise, discretization level and embedded anatomical boundaries [26]. While these deficiencies are ameliorated when using L^2 projection smoothing techniques [26], such approach remains largely heuristic as it lacks of a rigorous mathematical framework that can support the stability and accuracy of the resulting strain fields. This last observation motivates the development of DIR methods that can accurately

predict not only the image transformation but also its gradient by proving convergence of the associated numerical scheme.

Motivated by discussion above, the main goal of this work is to present a rigorous formulation and analysis of the registration problem, both for the continuous and discrete settings. Formulations of DIR using elastic regularizers can be interpreted as an elasticity problem with non-linear sources, where well-posedness is handled by requiring Lipschitz continuity and boundedness of the distributional gradient of the similarity-measure density. Taking this approach, the primal formulation of the DIR problem can be conveniently analyzed using Schauder's and Brower's fixed-point theorems for the continuous and discrete formulations, respectively. To directly account for the solution of not only the displacement field, but also of the associated stresses (and displacement gradients thereof), we propose and analyze mixed and augmented formulations of the DIR problem. Finite-element schemes are developed for all the formulations considered in this work, using a \mathbb{P}_1 formulation for the primal problem, a $BDM_1 - \mathbb{P}_0$ scheme for the mixed problem, and a $RT_0 - \mathbb{P}_1$ scheme for the augmented problem, all of which can be shown to be stable under the small data assumption.

In what follows, given a scalar expression A , we let \mathbf{A} and \mathbf{A} be its vectorial and tensor versions, respectively. In general, the regular font will be used for scalars, bold for vectors and bold slanted for tensors. In the same fashion, we define the usual Sobolev framework with

$$L^2(\Omega) := \{v : \Omega \rightarrow \mathbb{R} : \int v^2 < \infty\}, \quad \mathbf{L}^2(\Omega) := [L^2(\Omega)]^n, \quad \mathbf{L}^2(\Omega) := [L^2(\Omega)]^{n \times n},$$

and given $m \in \mathbb{N} \cup \{0\}$ use accordingly the spaces $H^m(\Omega)$, $\mathbf{H}^m(\Omega)$ and $\mathbf{H}^m(\Omega)$, where each v in these spaces has at least m distributional derivatives in $L^2(\Omega)$, $\mathbf{L}^2(\Omega)$ and $\mathbf{L}^2(\Omega)$ respectively. These use the inner product

$$\langle u, v \rangle_{m, \Omega} = \sum_{i=0}^m \langle D^i u, D^i v \rangle_{0, \Omega},$$

where $m = 0$ is just the $L^2(\Omega)$, meaningly $H^0 = L^2$. Typically, mixed-FEM schemes employ the space

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

which is a Hilbert space with inner product

$$\langle \boldsymbol{\tau}, \boldsymbol{\sigma} \rangle_{\text{div}; \Omega} = \int_{\Omega} \boldsymbol{\sigma} : \boldsymbol{\tau} + \int_{\Omega} \text{div } \boldsymbol{\tau} \cdot \text{div } \boldsymbol{\sigma}.$$

Hereafter, div stands for the usual divergence operator acting along the rows of a tensor. Finally, we define the trace operator $\gamma_D(\mathbf{u}) = \mathbf{u}|_{\Gamma}$, and the space

$$\mathbf{H}^{1/2}(\Gamma) := \left\{ \gamma_D(\mathbf{u}) : \mathbf{u} \in \mathbf{H}^1(\Omega) \right\},$$

where its dual is written $\mathbf{H}^{-1/2}(\Gamma) = (\mathbf{H}^{1/2}(\Gamma))'$ with the usual operator norm. In turn, we let $\gamma_{\nu} : H(\text{div}; \Omega) \rightarrow H^{-1/2}(\partial\Omega)$ be the normal trace operator on the boundary, which is defined distributionally (see [20] for more details).

2 Elastic deformable image registration (DIR) problem

Let $n \in \{2, 3\}$ be the dimension of the images we are interested in analyzing and $\Omega \subset \mathbb{R}^n$ be a compact domain with Lipschitz boundary $\Gamma := \partial\Omega$. Let $R : \Omega \rightarrow \mathbb{R}$ be the reference image and $T : \tilde{\Omega} \rightarrow \mathbb{R}$

be the target image with $\Omega \subseteq \tilde{\Omega}$, both in H^1 . The requirement that the domain of T is larger than that of R is necessary because in the definition of the registration problem we will need to evaluate T possibly outside Ω . In practice, images used in DIR problems have the same domain, and therefore we consider the underlying image $T_0 : \Omega \rightarrow \mathbb{R}$ and construct T by extending it (typically by zero) to $\tilde{\Omega}$. 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 a variational problem [30] that reads

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

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

$$\mathcal{D}[\mathbf{u}; R, T] := \frac{1}{2} \int_{\Omega} (T(\mathbf{x} + \mathbf{u}(\mathbf{x})) - R(\mathbf{x}))^2. \quad (2.2)$$

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

$$\mathcal{S}[\mathbf{u}] := \frac{1}{2} \int_{\Omega} \mathbb{C} \boldsymbol{\varepsilon}(\mathbf{u}) : \boldsymbol{\varepsilon}(\mathbf{u}), \quad (2.3)$$

where

$$\boldsymbol{\varepsilon}(\mathbf{u}) := \frac{1}{2} (\nabla \mathbf{u} + \nabla \mathbf{u}^\top) \quad (2.4)$$

is the infinitesimal strain tensor, i.e., the symmetric component of the displacement field gradient, and \mathbb{C} is the elasticity tensor, which for isotropic solids is defined by the expression

$$\mathbb{C} \boldsymbol{\tau} := \lambda \text{trace}(\boldsymbol{\tau}) \mathbf{I} + n \mu \boldsymbol{\tau}.$$

Assuming that (2.1) has at least one solution with sufficient regularity, the associated Euler-Lagrange equations deliver the following strong problem: Find $\mathbf{u} \in \mathbf{C}^2(\Omega) \cap \mathbf{C}^1(\bar{\Omega})$ such that

$$\begin{aligned} \text{div}(\mathbb{C} \boldsymbol{\varepsilon}(\mathbf{u})) &= \alpha \mathbf{f}_{\mathbf{u}} & \text{in } \Omega, \\ \mathbb{C} \boldsymbol{\varepsilon}(\mathbf{u}) \mathbf{n} &= 0 & \text{on } \partial\Omega, \end{aligned}$$

where

$$\mathbf{f}_{\mathbf{u}}(\mathbf{x}) := \{T(\mathbf{x} + \mathbf{u}(\mathbf{x})) - R(\mathbf{x})\} \nabla T(\mathbf{x} + \mathbf{u}(\mathbf{x})), \quad \forall \mathbf{x} \in \Omega \text{ a.e.} \quad (2.5)$$

3 Primal DIR formulation

Let $\mathcal{V} := \mathbf{H}^1(\Omega)$ and define the following forms by considering the registration problem defined in (2.1)

$$\begin{aligned} a(\mathbf{u}, \mathbf{v}) &:= \int \mathbb{C} \boldsymbol{\varepsilon}(\mathbf{u}) : \boldsymbol{\varepsilon}(\mathbf{v}) \quad \forall \mathbf{u}, \mathbf{v} \in \mathcal{V}, \\ F_{\mathbf{u}}(\mathbf{v}) &:= - \int \mathbf{f}_{\mathbf{u}} \cdot \mathbf{v} \quad \forall \mathbf{u}, \mathbf{v} \in \mathcal{V}, \end{aligned}$$

where $a : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}$ is bilinear and $F_{\mathbf{u}} : \mathcal{V} \rightarrow \mathbb{R}$ is linear for every \mathbf{u} . We can then write the DIR problem as

$$\min_{\mathbf{v} \in \mathcal{V}} \left\{ \alpha \mathcal{D}[\mathbf{v}] + a(\mathbf{v}, \mathbf{v}) \right\}, \quad (3.1)$$

and its Euler-Lagrange equations in weak form give the following problem: Find $\mathbf{u} \in \mathcal{V}$ such that

$$a(\mathbf{u}, \mathbf{v}) = \alpha F_{\mathbf{u}}(\mathbf{v}) \quad \forall \mathbf{v} \in \mathcal{V}. \quad (3.2)$$

We will make the following assumptions on the data:

$$\|F_{\mathbf{u}} - F_{\mathbf{v}}\|_{\mathcal{V}'} \leq L_F \|\mathbf{u} - \mathbf{v}\|_{0,\Omega} \quad \forall \mathbf{u}, \mathbf{v} \in \mathcal{V}, \quad (3.3)$$

$$\|F_{\mathbf{u}}\|_{\mathcal{V}'} \leq M_F \quad \forall \mathbf{u} \in \mathcal{V}. \quad (3.4)$$

We remark that assumptions (3.3) and (3.4) are achieved by imposing the following conditions on the nonlinear load term $\mathbf{f}_{\mathbf{u}}$:

$$\begin{aligned} |\mathbf{f}_{\mathbf{u}}(\mathbf{x}) - \mathbf{f}_{\mathbf{v}}(\mathbf{x})| &\leq L_f |\mathbf{u}(\mathbf{x}) - \mathbf{v}(\mathbf{x})| \quad \forall \mathbf{x} \in \Omega \text{ a.e.}, \\ |\mathbf{f}_{\mathbf{u}}(\mathbf{x})| &\leq M_f \quad \forall \mathbf{x} \in \Omega \text{ a.e.} \end{aligned}$$

In addition, we notice that in engineering practice, R, T are interpolations of an array of data (image) where values are defined at the nodes of a Cartesian grid. The most popular interpolation schemes used to construct R and T are cubic B-splines, which implies that $R, T \in C^2$. This in turn implies that the load term is C^1 , and therefore locally Lipschitz. This can be extended to the entire domain because Ω is compact.

In the following, we will consider the compact embedding $\mathbf{i}_c : \mathbf{H}^1(\Omega) \rightarrow \mathbf{L}^2(\Omega)$ given by Rellich-Kondrachov's theorem. We further recall Schauder's fixed point theorem (see [15] for a proof):

Theorem 3.1 (Schauder's Fixed Point Theorem). *Let W be a closed and convex subset of a Banach space \mathcal{V} and let $T : W \rightarrow W$ be a continuous mapping with the property that $\overline{T(W)}$ is compact. Then T has at least one fixed point in W .*

We define the following partial problem: Given $\mathbf{z} \in \mathcal{V}$, find $\mathbf{u} \in \mathcal{V}$ such that

$$a(\mathbf{u}, \mathbf{v}) = \alpha F_{\mathbf{z}}(\mathbf{v}) \quad \forall \mathbf{v} \in \mathcal{V}. \quad (3.5)$$

This problem does not have unisolvence, so we will modify the problem by imposing weak orthogonality to the rigid motions space, denoted by $\mathbb{RM}(\Omega)$ and defined as (see [8, Eq 11.1.7])

$$\mathbb{RM}(\Omega) := \{ \mathbf{v} \in H^1(\Omega) : \boldsymbol{\varepsilon}(\mathbf{v}) = 0 \},$$

which guarantees unisolvence of problem (3.5) since $\mathbb{RM}(\Omega)$ is precisely its null space. Defining $H = \mathbb{RM}(\Omega)^\perp$, we define the restricted problem as: Given $\mathbf{z} \in H$, find $\mathbf{u} \in H$ such that

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

An application of the Lax-Milgram lemma and Korn's inequality yields the following result:

Theorem 3.2. *Given $\mathbf{z} \in H$, problem (3.6) has a unique solution $\mathbf{u} \in H$ which satisfies the following a priori estimate for a constant $C > 0$:*

$$\|\mathbf{u}\|_{1,\Omega} \leq \alpha C \|F_{\mathbf{z}}\|_{\mathcal{V}'}. \quad (3.7)$$

Proof. See [8, Thm 11.2.30] □

We now define the operator $T : H \rightarrow H$ given by

$$T(\mathbf{z}) = \mathbf{u}, \quad (3.8)$$

where \mathbf{u} is the solution to problem (3.6) and thus rewrite problem (3.2) as: Find \mathbf{u} such that

$$T(\mathbf{u}) = \mathbf{u}, \quad (3.9)$$

which shows that the existence of solutions to the original nonlinear problem reduces to the existence of fixed points for the operator T . The following lemmas prove that the conditions required by Schauder's fixed point theorem hold.

Lemma 3.1. *Let T be an operator defined by (3.8). Then, under data assumption (3.3), there exists $L_T > 0$ such that*

$$\|T(\mathbf{z}_1) - T(\mathbf{z}_2)\|_{1,\Omega} \leq \alpha C L_F \|\mathbf{z}_1 - \mathbf{z}_2\|_{0,\Omega} \quad \forall \mathbf{z}_1, \mathbf{z}_2 \in H.$$

Proof. Given $\mathbf{z}_1, \mathbf{z}_2 \in H$, we let $\mathbf{u}_1 := T(\mathbf{z}_1)$ and $\mathbf{u}_2 := T(\mathbf{z}_2)$, that is \mathbf{u}_1 and \mathbf{u}_2 are the unique solutions to the following problems:

$$a(\mathbf{u}_1, \mathbf{v}) = \alpha F_{\mathbf{z}_1}(\mathbf{v}) \quad \forall \mathbf{v} \in H \quad \text{and} \quad a(\mathbf{u}_2, \mathbf{v}) = \alpha F_{\mathbf{z}_2}(\mathbf{v}) \quad \forall \mathbf{v} \in H.$$

Their difference gives a new problem

$$a(\mathbf{u}_1 - \mathbf{u}_2, \mathbf{v}) = \alpha (F_{\mathbf{z}_1} - F_{\mathbf{z}_2})(\mathbf{v}) \quad \forall \mathbf{v} \in H,$$

which satisfies the a priori estimate

$$\|\mathbf{u}_1 - \mathbf{u}_2\|_{1,\Omega} \leq \alpha C \|F_{\mathbf{z}_1} - F_{\mathbf{z}_2}\|_{\mathcal{V}}.$$

Using the condition over the data (3.3), we arrive to the desired result:

$$\|T(\mathbf{z}_1) - T(\mathbf{z}_2)\|_{1,\Omega} = \|\mathbf{u}_1 - \mathbf{u}_2\|_{1,\Omega} \leq \alpha C \|F_{\mathbf{z}_1} - F_{\mathbf{z}_2}\|_{\mathcal{V}} \leq \alpha C L_F \|\mathbf{z}_1 - \mathbf{z}_2\|_{0,\Omega},$$

with Lipschitz constant $L_T = \alpha C L_F$. □

Lemma 3.2. *Let T be an operator defined by (3.8). Then, under data assumption (3.4), there exists $r_0 > 0$ such that*

$$T(\overline{B}(0, r_0)) \subseteq \overline{B}(0, r_0) := \left\{ \mathbf{z} \in H : \|\mathbf{z}\|_{1,\Omega} \leq r_0 \right\}.$$

Proof. We conclude from the a priori estimate. Given $\mathbf{z} \in H$, we have from (3.7) and (3.8) that

$$\|T(\mathbf{z})\|_{1,\Omega} = \|\mathbf{u}\|_{1,\Omega} \leq \alpha C \|F_{\mathbf{z}}\|_{\mathcal{V}} \leq \alpha C M_F,$$

which shows that $T(\mathbf{z}) \in \overline{B}(0, r_0)$, with $r_0 = \alpha C M_F$, and hence in particular $T(\overline{B}(0, r_0)) \subseteq \overline{B}(0, r_0)$. □

We are now in a position to prove the existence of solution to the fixed-point problem (3.9).

Theorem 3.3. *Let T be the operator defined by (3.8). Then, under data assumptions (3.3) and (3.4), T has at least one fixed point. Moreover, if $\alpha C L_F < 1$, the fixed point is unique.*

Proof. Let $\{z_j\}_{j=1}^\infty$ be a sequence in $B(0, r_0)$ with $r_0 = \alpha CM_F$ as shown in Lemma 3.2. It follows that there exists a subsequence $\{z_j^{(1)}\}_{j \in \mathbb{N}} \subseteq \{z_j\}_{j \in \mathbb{N}}$ weakly convergent to some z in H . Using the compact embedding i_c we have that $z_j^{(1)} \xrightarrow{j} z$ in $\mathbf{L}^2(\Omega)$. In this way, using Lemma 3.1 we see that $T(z_j^{(1)}) \xrightarrow{j} T(z)$ in H , which means that $\overline{T(\overline{B}(0, r_0))}$ is compact and thus by Schauder's fixed point theorem we conclude the existence of a fixed point. Finally, if $\alpha CL_F < 1$, T is contraction, and so the conclusion is a consequence of Banach's fixed point theorem. \square

At this point we observe that, in the context of image registration, the foregoing result shows the existence of solutions to classic schemes such as diffusion and elastic registration together with SSD, cross-correlation or mutual information similarities. Also, data conditions (3.3) and (3.4) give a rule for how much regularity to ask from the images. For example, similarities involving gradients require a Lipschitz gradient, so at least H^2 is required.

3.1 Analysis of the discrete problem

In the following we formulate a Galerkin scheme to the primal DIR formulation. To this end, let $(\mathcal{V}_h)_{h>0}$ be a conforming family of discrete spaces indexed by a mesh size $h > 0$. We define $H_h := \mathbb{R}\mathbb{M}^\perp \cap \mathcal{V}_h$, and formulate the nonlinear discrete problem as follows: Find $\mathbf{u}_h \in H_h$ such that

$$a(\mathbf{u}_h, \mathbf{v}_h) = \alpha F_{\mathbf{u}_h}(\mathbf{v}_h) \quad \forall \mathbf{v}_h \in H_h. \quad (3.10)$$

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

$$a(\mathbf{u}_h, \mathbf{v}_h) = \alpha F_{\mathbf{z}_h}(\mathbf{v}_h) \quad \forall \mathbf{v}_h \in H_h, \quad (3.11)$$

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 unique solution of problem (3.11) given \mathbf{z}_h . To prove the existence of fixed points of T_h , we rely on Brouwer's fixed point theorem, which we include next for reference [15]:

Theorem 3.4 (Brouwer's Fixed Point Theorem). *Let K be a compact and convex subset of a finite-dimensional normed vector space and let $T_h : K \rightarrow K$ be a continuous mapping. Then T_h has at least one fixed point.*

Considering the same data assumptions as in the continuous case, as well as the continuity and bound obtained before, we arrive at the following result:

Theorem 3.5. *Let T_h be the discrete operator and assume data conditions (3.3) and (3.4) hold. Then, T_h has at least one fixed point. Further, if $\alpha CL_F < 1$, the fixed point it is unique.*

Proof. From the discrete analysis of Lemmas 3.1 and 3.2 we know that T_h is continuous and that there exists r_0 such that $T_h : \overline{B}(0, r_0) \rightarrow \overline{B}(0, r_0)$. Since H_h is finite-dimensional, $\overline{B}(0, r_0)$ is compact, and so Brouwer's conditions hold, from where we can conclude the existence of a fixed point. Uniqueness is established as in the continuous case with Banach's fixed point theorem. \square

Having proved the existence of solutions for the discrete problem, we can show convergence under uniqueness regime, and so we leave Cea's estimate for reference.

Theorem 3.6. *Let $\mathbf{u} \in H$ and $\mathbf{u}_h \in H_h$ be the solutions to the continuous and discrete problems (3.2), (3.10). Then, there exist $\alpha, C > 0$ such that*

$$\|\mathbf{u} - \mathbf{u}_h\|_{1,\Omega} \leq C \inf_{\mathbf{v}_h \in H_h} \|\mathbf{u} - \mathbf{v}_h\|_H. \quad (3.12)$$

Proof. Strang's first Lemma (see [40, Theorem 8.2]) implies that there exists a constant $\tilde{C} > 0$ such that

$$\|\mathbf{u} - \mathbf{u}_h\| \leq \tilde{C} \left\{ \inf_{\mathbf{v}_h \in H_h} \|\mathbf{u} - \mathbf{v}_h\|_H + \alpha \|F_{\mathbf{u}} - F_{\mathbf{u}_h}\|_{\mathcal{V}'} \right\}.$$

Using this inequality, the continuity of the compact embedding \mathbf{i}_c and data condition (3.3), we obtain

$$\|\mathbf{u} - \mathbf{u}_h\| \leq \tilde{C} \inf_{\mathbf{v}_h \in H_h} \|\mathbf{u} - \mathbf{v}_h\|_H + \alpha \tilde{C} L_F \|\mathbf{u} - \mathbf{u}_h\|_H.$$

Imposing $\alpha \tilde{C} L_F < 1$ gives the desired result for $C = \frac{\tilde{C}}{1 - \alpha \tilde{C} L_F}$. \square

Sufficiently small α allows the bound to be independent of it as shown in the following corollary.

Corollary 3.1. *Under the same assumptions of the previous Theorem, sufficiently small α allows C to depend only on \tilde{C} .*

Proof. Repeat the argument in Theorem 3.12 assuming $\alpha \tilde{C} L_F \leq \frac{1}{2}$, then $C \geq 2\tilde{C}$. \square

4 Mixed DIR formulation

In the following, we introduce a mixed variational formulation for (4.3). Following [6] and particularizing the method for the 2D case, we note that the constitutive relation can be inverted, to obtain that

$$\mathbb{C}^{-1}\boldsymbol{\sigma} := \frac{1}{2\mu}\boldsymbol{\sigma} - \frac{\lambda}{2\mu(2\mu + n\lambda)} \text{trace}(\boldsymbol{\sigma})\mathbf{I}. \quad (4.1)$$

We further define an auxiliary field given by the skew symmetric component of the displacement field gradient as

$$\boldsymbol{\rho} := \frac{1}{2}(\nabla\mathbf{u} - \nabla\mathbf{u}^t). \quad (4.2)$$

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. Then, the strong form of the mixed registration BVP problem reads: Find $\mathbf{u} \in \mathbf{C}^1(\Omega)$, $\boldsymbol{\sigma} \in \mathbf{C}^1(\Omega) \cap \mathbf{C}(\bar{\Omega})$ and $\boldsymbol{\rho} \in \mathbf{C}_{\text{skew}}^0(\Omega)$ such that

$$\begin{aligned} \mathbb{C}^{-1}\boldsymbol{\sigma} &= \nabla\mathbf{u} - \boldsymbol{\rho} \quad \text{in } \Omega, & \text{div } \boldsymbol{\sigma} &= \alpha \mathbf{f}_{\mathbf{u}} \quad \text{in } \Omega, \\ \boldsymbol{\sigma} &= \boldsymbol{\sigma}^t \quad \text{in } \Omega, & \boldsymbol{\sigma} \mathbf{n} &= \mathbf{0} \quad \text{in } \partial\Omega, \end{aligned} \quad (4.3)$$

4.1 Analysis of the continuous problem

Following the standard integration by parts procedure, the weak variational formulation of the mixed registration problem (4.3) reads: Find $(\boldsymbol{\sigma}, (\mathbf{u}, \boldsymbol{\rho})) \in \mathcal{H} \times Q$ such that

$$\begin{aligned} a(\boldsymbol{\sigma}, \boldsymbol{\tau}) + b(\boldsymbol{\tau}, (\mathbf{u}, \boldsymbol{\rho})) &= 0 & \forall \boldsymbol{\tau} \in \mathcal{H}, \\ b(\boldsymbol{\sigma}, (\mathbf{v}, \boldsymbol{\eta})) &= \alpha F_{\mathbf{u}}(\mathbf{v}, \boldsymbol{\eta}) & \forall (\mathbf{v}, \boldsymbol{\eta}) \in Q. \end{aligned} \quad (4.4)$$

where

$$\begin{aligned}\mathcal{H} &:= \mathbf{H}_0(\operatorname{div}; \Omega) = \{\boldsymbol{\tau} \in \mathbf{H}(\operatorname{div}; \Omega) : \gamma_\nu \boldsymbol{\tau} = 0\}, \\ Q &:= \mathbf{L}^2(\Omega) \times \mathbf{L}_{\text{skew}}^2(\Omega),\end{aligned}$$

and the bilinear forms $a : \mathcal{H} \times \mathcal{H} \rightarrow \mathbb{R}$ and $b : \mathcal{H} \times Q \rightarrow \mathbb{R}$ are defined as follows:

$$\begin{aligned}a(\boldsymbol{\sigma}, \boldsymbol{\tau}) &:= \int_{\Omega} \mathbb{C}^{-1} \boldsymbol{\sigma} : \boldsymbol{\tau} \quad \forall \boldsymbol{\sigma}, \boldsymbol{\tau} \in \mathcal{H}, \\ b(\boldsymbol{\tau}, (\mathbf{v}, \boldsymbol{\eta})) &:= \int_{\Omega} \mathbf{v} \cdot \operatorname{div} \boldsymbol{\tau} + \int_{\Omega} \boldsymbol{\eta} : \boldsymbol{\tau} \quad \forall \boldsymbol{\tau} \in \mathcal{H}, (\mathbf{v}, \boldsymbol{\eta}) \in Q,\end{aligned}$$

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

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

To study the solvability of (4.4), we define the partial problem: Given $\mathbf{z} \in \mathbf{L}^2(\Omega)$, find $(\boldsymbol{\sigma}, (\mathbf{u}, \boldsymbol{\rho})) \in \mathcal{H} \times Q$ such that

$$\begin{aligned}a(\boldsymbol{\sigma}, \boldsymbol{\tau}) + b(\boldsymbol{\tau}, (\mathbf{u}, \boldsymbol{\rho})) &= 0 \quad \forall \boldsymbol{\tau} \in \mathcal{H}, \\ b(\boldsymbol{\sigma}, (\mathbf{v}, \boldsymbol{\eta})) &= \alpha F_{\mathbf{z}}(\mathbf{v}, \boldsymbol{\eta}) \quad \forall (\mathbf{v}, \boldsymbol{\eta}) \in Q,\end{aligned} \tag{4.5}$$

which is a linear elasticity problem with Neumann boundary conditions. This problem does not have unisolvence, so we impose orthogonality to the rigid motions space \mathbb{RM} weakly. With this consideration, we define $H := \mathcal{H} \times \mathbb{RM}(\Omega)$, as well as the bilinear forms

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

We then consider the following equivalent mixed variational formulation: Given $\mathbf{z} \in \mathbf{L}^2(\Omega)$, find $((\boldsymbol{\sigma}, \boldsymbol{\rho}), (\mathbf{u}, \boldsymbol{\gamma})) \in H \times Q$ such that

$$\begin{aligned}A((\boldsymbol{\sigma}, \boldsymbol{\rho}), (\boldsymbol{\tau}, \boldsymbol{\xi})) + B((\boldsymbol{\tau}, \boldsymbol{\xi}), (\mathbf{u}, \boldsymbol{\gamma})) &= 0 \quad \forall (\boldsymbol{\tau}, \boldsymbol{\xi}) \in H, \\ B((\boldsymbol{\sigma}, \boldsymbol{\rho}), (\mathbf{v}, \boldsymbol{\eta})) &= \alpha F_{\mathbf{z}}((\mathbf{v}, \boldsymbol{\eta})) \quad \forall (\mathbf{v}, \boldsymbol{\eta}) \in Q.\end{aligned} \tag{4.6}$$

This formulation delivers a well-posed problem, as proved in [21, Theorem 3.1], and thus it allows for the definition of a fixed-point operator. Let $\mathbf{T} : \mathbf{L}^2(\Omega) \rightarrow \mathbf{L}^2(\Omega)$ given by

$$\mathbf{T}(\mathbf{z}) := \mathbf{u} \quad \forall \mathbf{z} \in \mathbf{L}^2(\Omega), \tag{4.7}$$

where \mathbf{u} is the displacement component of the unique solution of problem (4.6), and so the mixed formulation can be restated as: Find $\mathbf{u} \in \mathbf{L}^2(\Omega)$ such that

$$\mathbf{T}(\mathbf{u}) = \mathbf{u}. \tag{4.8}$$

Note here that if $((\boldsymbol{\tau}, \boldsymbol{\xi}), (\mathbf{v}, \boldsymbol{\eta})) = ((\mathbf{0}, \boldsymbol{\xi}), (\boldsymbol{\xi}, \nabla \boldsymbol{\xi}))$, we obtain

$$\int_{\Omega} \boldsymbol{\rho} \cdot \boldsymbol{\xi} = \int_{\Omega} \mathbf{f}_{\mathbf{z}} \cdot \boldsymbol{\xi} \quad \forall \boldsymbol{\xi} \in \mathbb{RM},$$

so the Lagrange multiplier $\boldsymbol{\rho}$ removes the compatibility requirement from the data.

To prove the existence of a solution to problem (4.8) we use Banach's fixed point theorem for a sufficiently small α and data conditions (3.4) and (3.3). The continuity of the operator \mathbf{T} is established in the following lemma.

Lemma 4.1. *Let the operator \mathbf{T} defined by (4.7), assume the data condition (3.3), and let $C > 0$ be the constant of continuous dependence on data of (4.6). Then there holds:*

$$\|\mathbf{T}(\mathbf{z}_1) - \mathbf{T}(\mathbf{z}_2)\|_{0,\Omega} \leq \alpha C L_F \|\mathbf{z}_1 - \mathbf{z}_2\|_{0,\Omega} \quad \forall \mathbf{z}_1, \mathbf{z}_2 \in \mathbf{L}^2(\Omega). \quad (4.9)$$

Proof. Set $\mathbf{u}_i = \mathbf{T}(\mathbf{z}_i), i \in \{1, 2\}$. Subtracting terms we arrive at the following system of equations:

$$\begin{aligned} A((\boldsymbol{\sigma}_1 - \boldsymbol{\sigma}_2, \boldsymbol{\rho}_1 - \boldsymbol{\rho}_2), (\boldsymbol{\tau}, \boldsymbol{\xi})) + B((\boldsymbol{\tau}, \boldsymbol{\xi}), (\mathbf{u}_1 - \mathbf{u}_2, \boldsymbol{\gamma}_1 - \boldsymbol{\gamma}_2)) &= 0 & \forall (\boldsymbol{\tau}, \boldsymbol{\xi}) \in H, \\ B((\boldsymbol{\sigma}_1 - \boldsymbol{\sigma}_2, \boldsymbol{\rho}_1 - \boldsymbol{\rho}_2), (\mathbf{v}, \boldsymbol{\eta})) &= \alpha(F_{\mathbf{z}_1} - F_{\mathbf{z}_2})(\mathbf{v}, \boldsymbol{\eta}) & \forall (\mathbf{v}, \boldsymbol{\eta}) \in Q, \end{aligned}$$

from which, using the continuous dependence of (4.6) and the Lipschitz continuity of F , we get

$$\|\mathbf{T}(\mathbf{z}_1) - \mathbf{T}(\mathbf{z}_2)\|_{0,\Omega} = \|\mathbf{u}_1 - \mathbf{u}_2\|_{0,\Omega} \leq \alpha C \|F_{\mathbf{z}_1} - F_{\mathbf{z}_2}\|_{Q'} \leq \alpha C L_F \|\mathbf{z}_1 - \mathbf{z}_2\|_{0,\Omega},$$

thus completing the proof. \square

Theorem 4.1. *Under data conditions (3.3), (3.4) and assuming $\alpha C L_F < 1$, there is a unique fixed point for problem (4.8). With this, the mixed formulation (4.4) has a unique solution and the a priori estimation*

$$\|((\boldsymbol{\sigma}, \boldsymbol{\rho}), (\mathbf{u}, \boldsymbol{\gamma}))\|_H \leq \alpha C M_F$$

holds.

Proof. From Lemma 4.1 we have that

$$\|\mathbf{T}(\mathbf{z}_1) - \mathbf{T}(\mathbf{z}_2)\|_{0,\Omega} \leq \alpha C L_F \|\mathbf{z}_1 - \mathbf{z}_2\|_{0,\Omega} \quad \forall \mathbf{z}_1, \mathbf{z}_2 \in \mathbf{L}^2(\Omega),$$

which thanks to the assumption $\alpha C L_F < 1$ makes \mathbf{T} a contraction, and thus prove the existence of the unique fixed point by virtue of Banach's fixed point theorem. For the a priori bound, we use the same estimate of the partial problem setting $\mathbf{z} = \mathbf{u}$ and condition (3.4), which gives

$$\|((\boldsymbol{\sigma}, \boldsymbol{\rho}), (\mathbf{u}, \boldsymbol{\gamma}))\|_H \leq \alpha C \|F_{\mathbf{u}}\|_{Q'} \leq \alpha C M_F,$$

which concludes the proof. \square

4.2 Analysis of the discrete problem

In this section we analyze a Galerkin scheme for problem (4.4). Let $\{\mathcal{T}_h\}_{h>0}$ be a regular family of triangularizations of $\bar{\Omega}$ of characteristic size h , and the following set of inf-sup stable discrete spaces:

$$\begin{aligned} H_h^\sigma &:= \left\{ \boldsymbol{\tau}_h \in \mathbf{H}(\text{div}; \Omega) : \boldsymbol{\tau}_{h,i}|_T \in BDM_1(T) \quad \forall T \in \mathcal{T}_h \right\}, \\ Q_h^u &:= \left\{ \mathbf{v}_h \in \mathbf{L}^2(\Omega) : \mathbf{v}_h \in [\mathbb{P}_0(T)]^n \quad \forall T \in \mathcal{T}_h \right\}, \\ Q_h^\gamma &:= \left\{ \begin{bmatrix} 0 & \psi_h \\ -\psi_h & 0 \end{bmatrix} : \psi_h \in \mathbb{P}_1(T) \quad \forall T \in \mathcal{T}_h \right\}, \end{aligned}$$

where $BDM_k = [\mathbb{P}_k]^n$ are the Brezzi-Douglas-Marini elements [9] and $\boldsymbol{\tau}_{h,i}$ is the i th row of $\boldsymbol{\tau}_h$. Then, we introduce

$$H_{0,h}^\sigma := H_h^\sigma \cap \mathbf{H}_0(\text{div}; \Omega), \quad H_h = H_{0,h}^\sigma \times \mathbb{RM}(\Omega),$$

and

$$Q_h = Q_h^u \times Q_h^\gamma.$$

Thus, we define the discrete version of (4.6) as follows: Given $z_h \in Q_h^u$, find $((\sigma_h, \rho_h), (\mathbf{u}_h, \boldsymbol{\gamma}_h)) \in H_h \times Q_h$ such that

$$\begin{aligned} A((\sigma_h, \rho_h), (\boldsymbol{\tau}_h, \boldsymbol{\xi}_h)) + B((\boldsymbol{\tau}_h, \boldsymbol{\xi}_h), (\mathbf{u}_h, \boldsymbol{\gamma}_h)) &= 0 & \forall (\boldsymbol{\tau}_h, \boldsymbol{\xi}_h) \in H_h \\ B((\sigma_h, \rho_h), (\mathbf{v}_h, \boldsymbol{\eta}_h)) &= \alpha F_{z_h}((\mathbf{v}_h, \boldsymbol{\eta}_h)) & \forall (\mathbf{v}_h, \boldsymbol{\eta}_h) \in Q_h. \end{aligned} \quad (4.10)$$

The unique solvability and stability of (4.10), being the Galerkin scheme of a linear elasticity problem with Neumann boundary conditions has already been established in [21, Theorem 4.1]. This allows us to define the discrete operator $\mathbf{T}_h : Q_h^u \rightarrow Q_h^u$, given by

$$\mathbf{T}_h(z_h) = \mathbf{u}_h,$$

and then rewrite the discretized nonlinear problem as: Find $\mathbf{u}_h \in Q_h^u$ such that

$$\mathbf{T}_h(\mathbf{u}_h) = \mathbf{u}_h. \quad (4.11)$$

Concerning the other components $\sigma_h, \boldsymbol{\gamma}_h$ and ρ_h of this problem, we will refer to $((\sigma_h, \rho_h), (\mathbf{u}_h, \boldsymbol{\gamma}_h))$ as the mixed solution for a given z_h .

Lemma 4.2. *Let \mathbf{T}_h be the discrete operator given by (4.11) and C be the constant of continuous dependence on data of (4.10). Then, given z_1, z_2 in Q_h^u , there holds*

$$\|\mathbf{T}_h(z_1) - \mathbf{T}_h(z_2)\|_{0,\Omega} \leq \alpha C L_F \|z_1 - z_2\|_{0,\Omega}.$$

Proof. Repeat argument in Lemma 4.1 to \mathbf{T}_h . □

Now we are in position to establish the well-posedness of problem (4.11), as well as to prove the Cea's best approximation property.

Theorem 4.2. *Under data assumptions (3.3), (3.4) and assuming $\alpha C L_F < 1$, the problem (4.11) has a unique solution $((\sigma_h, \rho_h), (\mathbf{u}_h, \boldsymbol{\gamma}_h)) \in H_h \times Q_h$ such that*

$$\|((\sigma_h, \rho_h), (\mathbf{u}_h, \boldsymbol{\gamma}_h))\|_{H \times Q} \leq \alpha C M_F.$$

In addition, there exists $\tilde{C} > 0$ such that Cea's estimate holds for the unique solution $((\sigma, \rho), (\mathbf{u}, \boldsymbol{\gamma}))$ to problem (4.8), i.e.,

$$\|((\sigma, \rho), (\mathbf{u}, \boldsymbol{\gamma})) - ((\sigma_h, \rho_h), (\mathbf{u}_h, \boldsymbol{\gamma}_h))\|_{H \times Q} \leq \tilde{C} \inf_{\substack{(\boldsymbol{\tau}_h, \mathbf{0}) \in H_h, \\ (\mathbf{v}_h, \boldsymbol{\eta}_h) \in Q_h}} \|((\sigma, \mathbf{0}), (\mathbf{u}, \boldsymbol{\gamma})) - ((\boldsymbol{\tau}_h, \mathbf{0}), (\mathbf{v}_h, \boldsymbol{\eta}_h))\|_{H \times Q}.$$

Proof. The first part is analogous to the continuous case, so we are only left with Cea's estimate. Let $((\sigma, \rho), (\mathbf{u}, \boldsymbol{\gamma})) \in H \times Q$ and $((\sigma_h, \rho_h), (\mathbf{u}_h, \boldsymbol{\gamma}_h)) \in H_h \times Q_h$ be the solutions arising from the continuous and discrete problems (4.8) and (4.11), respectively. Equivalently, $((\sigma, \rho), (\mathbf{u}, \boldsymbol{\gamma}))$ (resp. $((\sigma_h, \rho_h), (\mathbf{u}_h, \boldsymbol{\gamma}_h))$) solves (4.6) with $z = \mathbf{u}$ (resp. (4.10) with $z_h = \mathbf{u}_h$). In addition, let also $((\zeta_h, \varphi_h), (\mathbf{w}_h, \mathbf{s}_h)) \in H_h \times Q_h$ be the solution to the discrete version of (4.6) with $z = \mathbf{u}$, but without changing \mathbf{u} by \mathbf{u}_h , that is

$$\begin{aligned} A((\zeta_h, \varphi_h), (\boldsymbol{\tau}_h, \boldsymbol{\xi}_h)) + B((\boldsymbol{\tau}_h, \boldsymbol{\xi}_h), (\mathbf{w}_h, \mathbf{s}_h)) &= 0 & \forall (\boldsymbol{\tau}_h, \boldsymbol{\xi}_h) \in H_h \\ B((\zeta_h, \varphi_h), (\mathbf{v}_h, \boldsymbol{\eta}_h)) &= \alpha F_{\mathbf{u}}((\mathbf{v}_h, \boldsymbol{\eta}_h)) & \forall (\mathbf{v}_h, \boldsymbol{\eta}_h) \in Q_h. \end{aligned} \quad (4.12)$$

By virtue of [20, Theorem 2.6], the following estimate holds for a constant $\bar{C} > 0$, which depends only on the bilinear forms A, B :

$$\begin{aligned} & \|((\boldsymbol{\sigma}, \boldsymbol{\rho}), (\mathbf{u}, \boldsymbol{\gamma})) - ((\boldsymbol{\zeta}_h, \boldsymbol{\varphi}_h), (\mathbf{w}_h, \mathbf{s}_h))\|_{H \times Q} \\ & \leq \bar{C} \inf_{\substack{(\boldsymbol{\tau}_h, \boldsymbol{\xi}_h) \in H_h, \\ (\mathbf{v}_h, \boldsymbol{\eta}_h) \in Q_h}} \|((\boldsymbol{\sigma}, \boldsymbol{\rho}), (\mathbf{u}, \boldsymbol{\gamma})) - ((\boldsymbol{\tau}_h, \boldsymbol{\xi}_h), (\mathbf{v}_h, \boldsymbol{\eta}_h))\|_{H \times Q}. \end{aligned} \quad (4.13)$$

Subtracting systems (4.10) with $\mathbf{z}_u = \mathbf{u}_h$ and (4.12) it holds that

$$\begin{aligned} A((\boldsymbol{\sigma}_h - \boldsymbol{\zeta}_h, \boldsymbol{\rho}_h - \boldsymbol{\varphi}_h), (\boldsymbol{\tau}_h, \boldsymbol{\xi}_h)) + B((\boldsymbol{\tau}_h, \boldsymbol{\xi}_h), (\mathbf{u}_h - \mathbf{w}_h, \boldsymbol{\gamma}_h - \mathbf{s}_h)) &= 0 \\ B((\boldsymbol{\sigma}_h - \boldsymbol{\zeta}_h, \boldsymbol{\rho}_h - \boldsymbol{\varphi}_h), (\mathbf{v}_h, \boldsymbol{\eta}_h)) &= \alpha(F_{\mathbf{u}_h} - F_{\mathbf{u}})((\mathbf{v}_h, \boldsymbol{\eta}_h)), \end{aligned} \quad (4.14)$$

for all $(\boldsymbol{\tau}_h, \boldsymbol{\xi}_h) \in H_h, (\mathbf{v}_h, \boldsymbol{\eta}_h) \in Q_h$, and once again the Babuška-Brezzi theory gives the estimate

$$\|((\boldsymbol{\sigma}_h, \boldsymbol{\rho}_h), (\mathbf{u}_h, \boldsymbol{\gamma}_h)) - ((\boldsymbol{\zeta}_h, \boldsymbol{\varphi}_h), (\mathbf{w}_h, \mathbf{s}_h))\|_{H \times Q} \leq \alpha C \|F_{\mathbf{u}} - F_{\mathbf{u}_h}\|_{Q'}. \quad (4.15)$$

Finally, using (4.13) and (4.15) we obtain the following Strang-type estimate

$$\begin{aligned} & \|((\boldsymbol{\sigma}, \boldsymbol{\rho}), (\mathbf{u}, \boldsymbol{\gamma})) - ((\boldsymbol{\sigma}_h, \boldsymbol{\rho}_h), (\mathbf{u}_h, \boldsymbol{\gamma}_h))\|_{H \times Q} \\ & \leq \|((\boldsymbol{\sigma}, \boldsymbol{\rho}), (\mathbf{u}, \boldsymbol{\gamma})) - ((\boldsymbol{\zeta}_h, \boldsymbol{\varphi}_h), (\mathbf{w}_h, \mathbf{s}_h))\|_{H \times Q} + \|((\boldsymbol{\sigma}_h, \boldsymbol{\rho}_h), (\mathbf{u}_h, \boldsymbol{\gamma}_h)) - ((\boldsymbol{\zeta}_h, \boldsymbol{\varphi}_h), (\mathbf{w}_h, \mathbf{s}_h))\|_{H \times Q} \\ & \leq \bar{C} \inf_{\substack{(\boldsymbol{\tau}_h, \boldsymbol{\xi}_h) \in H_h, \\ (\mathbf{v}_h, \boldsymbol{\eta}_h) \in Q_h}} \|((\boldsymbol{\sigma}, \boldsymbol{\rho}), (\mathbf{u}, \boldsymbol{\gamma})) - ((\boldsymbol{\tau}_h, \boldsymbol{\xi}_h), (\mathbf{v}_h, \boldsymbol{\eta}_h))\|_{H \times Q} + \alpha C \|F_{\mathbf{u}} - F_{\mathbf{u}_h}\|_{Q'}. \end{aligned}$$

In this way, using the data condition (3.3), and assuming that $\alpha C L_F < 1$, we find that

$$\begin{aligned} & \|((\boldsymbol{\sigma}, \boldsymbol{\rho}), (\mathbf{u}, \boldsymbol{\gamma})) - ((\boldsymbol{\sigma}_h, \boldsymbol{\rho}_h), (\mathbf{u}_h, \boldsymbol{\gamma}_h))\|_{H \times Q} \\ & \leq \tilde{C} \inf_{\substack{(\boldsymbol{\tau}_h, \boldsymbol{\xi}_h) \in H_h, \\ (\mathbf{v}_h, \boldsymbol{\eta}_h) \in Q_h}} \|((\boldsymbol{\sigma}, \boldsymbol{\rho}), (\mathbf{u}, \boldsymbol{\gamma})) - ((\boldsymbol{\tau}_h, \boldsymbol{\xi}_h), (\mathbf{v}_h, \boldsymbol{\eta}_h))\|_{H \times Q}, \end{aligned}$$

where $\tilde{C} = \frac{\bar{C}}{1 - \alpha C L_F}$. The absence of the rigid motion variable is due to the space not being discretized, and thus we have that $d(\boldsymbol{\rho}, \mathbb{RM}(\Omega)) = 0$. \square

Notice that the proof used to establish this Strang-type estimate is analogous to the one employed in the primal case in [40, Theorem 8.2].

Corollary 4.1. *The above Theorem holds for $\tilde{C} = 2\bar{C}$ provided α is sufficiently small.*

Proof. In the preceding proof, take $\alpha C L_F \leq \frac{1}{2}$. \square

We end this section by providing the rate of convergence of the solution to (4.11). We first recall the classic approximation results from [10] and proceed as in [21]:

(AP_h^σ) For each $\boldsymbol{\tau} \in \mathbf{H}^1(\Omega) \cap \mathbf{H}_0(\text{div}; \Omega)$ with $\text{div } \boldsymbol{\tau} \in \mathbf{H}^1(\Omega)$ there exists $\boldsymbol{\tau}_h \in H_{0,h}^\sigma$ such that

$$\|\boldsymbol{\tau} - \boldsymbol{\tau}_h\|_{\text{div}; \Omega} \leq Ch \{ \|\boldsymbol{\tau}\|_{1, \Omega} + \|\text{div } \boldsymbol{\tau}\|_{1, \Omega} \}.$$

(AP_h^u) For each $\mathbf{v} \in \mathbf{H}^1(\Omega)$ there there exists $\mathbf{v}_h \in H_h^u$ such that

$$\|\mathbf{v} - \mathbf{v}_h\|_{0, \Omega} \leq Ch \|\mathbf{v}\|_{1, \Omega}.$$

($AP_h^\boldsymbol{\gamma}$) For each $\boldsymbol{\eta} \in \mathbf{H}^1(\Omega) \cap \mathbf{L}_{asym}^2(\Omega)$ there exists $\boldsymbol{\eta}_h \in H_h^\boldsymbol{\gamma}$ such that

$$\|\boldsymbol{\eta} - \boldsymbol{\eta}_h\|_{0,\Omega} \leq Ch\|\boldsymbol{\eta}\|_{1,\Omega}.$$

These allow us to establish the following theorem:

Theorem 4.3. *Under data assumptions (3.3), (3.4) and assuming $\alpha CL_F \leq \frac{1}{2}$, let $((\boldsymbol{\sigma}_h, \boldsymbol{\rho}_h), (\mathbf{u}_h, \boldsymbol{\gamma}_h)) \in H_h \times Q_h$ be the mixed solution of (4.11) and $((\boldsymbol{\sigma}, \boldsymbol{\rho}), (\mathbf{u}, \boldsymbol{\gamma})) \in H \times Q$ the solution of the continuous mixed problem (4.8). Then, there exists a constant $C > 0$, independent of h , such that whenever $\boldsymbol{\sigma} \in \mathbf{H}^1(\Omega)$, $\text{div } \boldsymbol{\sigma} \in \mathbf{H}^1(\Omega)$, $\mathbf{u} \in \mathbf{H}^1(\Omega)$ and $\boldsymbol{\gamma} \in \mathbf{H}^1(\Omega)$, there holds*

$$\|((\boldsymbol{\sigma}, \boldsymbol{\rho}), (\mathbf{u}, \boldsymbol{\gamma})) - ((\boldsymbol{\sigma}_h, \boldsymbol{\rho}_h), (\mathbf{u}_h, \boldsymbol{\gamma}_h))\|_{H \times Q} \leq Ch \{ \|\boldsymbol{\tau}\|_{1,\Omega} + \|\text{div } \boldsymbol{\tau}\|_{1,\Omega} + \|\mathbf{v}\|_{1,\Omega} + \|\boldsymbol{\eta}\|_{1,\Omega} \}.$$

Proof. It follows from applying Cea's estimate and the approximation properties ($AP_h^\boldsymbol{\sigma}$), ($AP_h^{\mathbf{u}}$) and ($AP_h^\boldsymbol{\gamma}$). \square

5 Augmented DIR formulation

This section proposes an augmented mixed variational formulation for the BVP (4.3). This scheme gives additional regularity to the displacement field, allows the use of a weaker fixed point theorem guaranteeing the existence of solutions for any positive α , and permits more flexibility for the choice of the finite element subspaces. Again, sufficiently small α grants uniqueness. It has been studied in [19] that the augmented Dirichlet and mixed-boundary condition linear elasticity problems have a unique solution, and here we prove that the same holds for the null traction problem.

5.1 Analysis of the continuous problem

Let $H := \mathbf{H}(\text{div}; \Omega) \times \mathbf{H}^1(\Omega) \times \mathbf{L}_{skew}^2$, $Q := \mathbb{RM}(\Omega)$. We further define the operators $A : H \times H \rightarrow \mathbb{R}$, $B : H \times Q \rightarrow \mathbb{R}$ and $F_z \in H'$ as follows

$$\begin{aligned} A((\boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\gamma}), (\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\eta})) &:= a(\boldsymbol{\sigma}, \boldsymbol{\tau}) + b(\boldsymbol{\tau}, (\mathbf{u}, \boldsymbol{\gamma})) - b(\boldsymbol{\sigma}, (\mathbf{v}, \boldsymbol{\eta})) \\ &\quad + \kappa_1 \int_{\Omega} (\boldsymbol{\varepsilon}(\mathbf{u})) - \mathbb{C}^{-1}\boldsymbol{\sigma} : (\boldsymbol{\varepsilon}(\mathbf{v}) + \mathbb{C}^{-1}\boldsymbol{\tau}) \\ &\quad + \kappa_2 \int_{\Omega} \text{div } \boldsymbol{\sigma} \cdot \text{div } \boldsymbol{\tau} \\ &\quad + \kappa_3 \int_{\Omega} (\boldsymbol{\gamma} - \text{skew } \nabla \mathbf{u}) : (\boldsymbol{\eta} + \text{skew } \nabla \mathbf{v}), \\ B((\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\eta}), \boldsymbol{\xi}) &:= \int_{\Omega} \mathbf{v} \cdot \boldsymbol{\xi}, \\ F_z((\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\eta})) &:= \int_{\Omega} \mathbf{f}_z \cdot (-\mathbf{v} + \kappa_2 \text{div } \boldsymbol{\tau}), \end{aligned} \tag{5.1}$$

where κ_1 , κ_2 , and κ_3 are positive parameters to be suitably chosen later on. With the definitions above, the augmented formulation reads: Given $\mathbf{z} \in \mathbf{H}^1(\Omega)$, find $((\boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\gamma}), \boldsymbol{\rho}) \in H \times Q$ such that:

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

Now we define the augmented fixed point operator $\mathbf{T} : \mathbf{H}^1(\Omega) \rightarrow \mathbf{H}^1(\Omega)$ given by $\mathbf{T}(\mathbf{z}) = \mathbf{u}$, where \mathbf{u} is the displacement component of the solution to problem (5.2), and so we write the nonlinear augmented problem as: Find $\mathbf{u} \in \mathbf{H}^1(\Omega)$ such that

$$\mathbf{T}(\mathbf{u}) = \mathbf{u}. \quad (5.2)$$

We apply the Babuška-Brezzi conditions to see that the proposed problem has a unique solution and depends continuously on the data. We will use the following definitions:

$$\boldsymbol{\tau}^d = \boldsymbol{\tau} - \frac{1}{n} \text{trace}(\boldsymbol{\tau}) \mathbf{I} \quad \text{and} \quad \boldsymbol{\tau}_0 = \boldsymbol{\tau} - \frac{1}{|\Omega|} \int \text{trace}(\boldsymbol{\tau}) \mathbf{I},$$

where the first one is known as the deviatoric tensor, and is such that it has null trace. The second one is the decomposition that arises from $\mathbf{H}(\text{div}; \Omega) = \tilde{\mathbf{H}}(\text{div}; \Omega) \oplus \mathbb{R} \mathbf{I}$, where \mathbf{I} is the identity tensor and

$$\tilde{\mathbf{H}}(\text{div}; \Omega) := \left\{ \boldsymbol{\tau} \in \mathbf{H}(\text{div}; \Omega) : \int \text{trace} \boldsymbol{\tau} = 0 \right\}.$$

We will use the following lemmas:

Lemma 5.1. *There exists $C_1 > 0$ such that*

$$\|\boldsymbol{\tau}^d\|_{0,\Omega}^2 + \|\text{div} \boldsymbol{\tau}\|_{0,\Omega}^2 \geq C_1 \|\boldsymbol{\tau}\|_{\text{div};\Omega}^2 \quad \forall \boldsymbol{\tau} \in \mathbf{H}_0(\text{div}; \Omega). \quad (5.3)$$

Proof. We start from [19, Lemma 2.1], which guarantees the existence of $c_1 > 0$ such that

$$\|\text{div} \boldsymbol{\tau}\|_{0,\Omega}^2 + \|\boldsymbol{\tau}^d\|_{0,\Omega}^2 \geq c_1 \|\boldsymbol{\tau}_0\|_{0,\Omega}^2 \quad \forall \boldsymbol{\tau} \in \mathbf{H}(\text{div}; \Omega).$$

Then, noting that $\boldsymbol{\tau}_0 = \boldsymbol{\tau}_0^d$ and $\text{div} \boldsymbol{\tau} = \text{div} \boldsymbol{\tau}_0$, we readily find that

$$\|\boldsymbol{\tau}^d\|_{0,\Omega}^2 + \|\text{div} \boldsymbol{\tau}\|_{0,\Omega}^2 \geq \frac{c_1}{2} \|\boldsymbol{\tau}_0\|_{0,\Omega}^2 + \frac{1}{2} \|\text{div} \boldsymbol{\tau}\|_{0,\Omega}^2 \geq \frac{1}{2} \min\{c_1, 1\} \|\boldsymbol{\tau}_0\|_{\text{div};\Omega}^2. \quad (5.4)$$

Now, writing explicitly $\boldsymbol{\tau} = \boldsymbol{\tau}_0 + d\mathbf{I}$ we impose the null normal trace condition from $0 = \gamma_\nu(\boldsymbol{\tau}) = \gamma_\nu(\boldsymbol{\tau}_0 + d\mathbf{I})$ and obtain

$$|d| \|\nu\|_{-1/2,\Gamma} = \|\gamma_\nu \boldsymbol{\tau}_0\|_{-1/2,\Omega} \leq \|\boldsymbol{\tau}_0\|_{\text{div};\Omega},$$

or equivalently,

$$|d| \leq \frac{1}{\|\nu\|_{-1/2,\Gamma}} \|\boldsymbol{\tau}_0\|_{\text{div};\Omega},$$

where the continuity of the normal trace was used. Then, applying this to the norm of $\boldsymbol{\tau}$ we get

$$\|\boldsymbol{\tau}\|_{\text{div};\Omega} = \|\boldsymbol{\tau}_0\|_{\text{div};\Omega} + n|d|\|\Omega\| \leq \left(1 + \frac{n|\Omega|}{\|\nu\|_{-1/2,\Gamma}}\right) \|\boldsymbol{\tau}_0\|_{\text{div};\Omega},$$

and everything together in equation (5.4) gives the desired result, with $C_1 = 2 \frac{\left(1 + \frac{n|\Omega|}{\|\nu\|_{-1/2,\Gamma}}\right)^2}{\min\{1, c_1\}}$. \square

Lemma 5.2. *There exists $\tilde{\kappa} > 0$ such that :*

$$\|\boldsymbol{\varepsilon}(\mathbf{v})\|_{0,\Omega}^2 \geq \tilde{\kappa} \|\mathbf{v}\|_{1,\Omega}^2 \quad \forall \mathbf{v} \in \mathbb{R} \mathbf{M}^\perp. \quad (5.5)$$

Proof. The proof is a combination of [8, Theorem 9.2.12] (Korn's second inequality) and the fact that $H^1 = \widehat{H}^1 \oplus \mathbb{RM}$, where

$$\widehat{H}^1(\Omega) = \left\{ v \in H^1(\Omega) : \int_{\Omega} v = 0, \int_{\Omega} \operatorname{rot} v = 0 \right\}.$$

□

Theorem 5.1. *Let $V := N(\mathbb{B})$, where \mathbb{B} is the operator induced by the bilinear form B defined in (5.2). Then, there exist κ_1, κ_2 and κ_3 for (5.1) such that*

1. *the bilinear form A is V -elliptic,*
2. *the bilinear form B satisfies the inf-sup condition,*

so that there exists a unique solution $((\boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\gamma}), \boldsymbol{\xi}) \in H \times Q$ to the problem (5.2) for a given $\mathbf{z} \in \mathbf{H}^1(\Omega)$. In addition, there exists a constant $C > 0$ such that

$$\|((\boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\gamma}), \boldsymbol{\xi})\|_{H \times Q} \leq \alpha C \|F_{\mathbf{z}}\|_{H'}.$$

Proof. It is important to mention that we are considering rigid motions endowed with the $L^2(\Omega)$ inner product, and that the main idea is to find values for $\kappa_1, \kappa_2, \kappa_3$ (cf. (5.1)) such that ellipticity holds. First we prove ellipticity in V , where the norm used will be the following:

$$\|(\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\eta})\|_V^2 := \|\boldsymbol{\tau}\|_{\operatorname{div}; \Omega}^2 + \|\mathbf{v}\|_{1, \Omega}^2 + \|\boldsymbol{\eta}\|_{0, \Omega}^2.$$

In fact, since $V := \{(\boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\gamma}) \in H : B((\boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\gamma}), \boldsymbol{\xi}) = 0 \quad \forall \boldsymbol{\xi} \in Q\}$, we find that

$$V = \{(\boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\gamma}) \in H : \int_{\Omega} \mathbf{u} \cdot \boldsymbol{\xi} = 0 \quad \forall \boldsymbol{\xi} \in Q\} = \{(\boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\gamma}) \in H : \mathbf{u} \in \mathbb{RM}^{\perp}\}.$$

In turn, for the ellipticity we are seeking, we use from [19, Theorem 3.1] that

$$\begin{aligned} A((\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\eta}), (\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\eta})) &= \int_{\Omega} \mathbb{C}^{-1} \boldsymbol{\tau} : \boldsymbol{\tau} - \kappa_1 \|\mathbb{C}^{-1} \boldsymbol{\tau}\|_{0, \Omega}^2 + (\kappa_1 + \kappa_3) \|\mathbf{e}(\mathbf{v})\|_{0, \Omega}^2 \\ &\quad + \kappa_2 \|\operatorname{div} \boldsymbol{\tau}\|_{0, \Omega}^2 + \kappa_3 \|\boldsymbol{\eta}\|_{0, \Omega}^2 - \kappa_3 \|\mathbf{v}\|_{1, \Omega}^2, \end{aligned} \quad (5.6)$$

and that

$$\int_{\Omega} \mathbb{C}^{-1} \boldsymbol{\tau} : \boldsymbol{\tau} - \kappa_1 \|\mathbb{C}^{-1} \boldsymbol{\tau}\|_{0, \Omega}^2 \geq \frac{1}{2\mu} \left(1 - \frac{\kappa_1}{2\mu}\right) \|\boldsymbol{\tau}^d\|_{0, \Omega}^2, \quad (5.7)$$

which require $0 < \kappa_1 < 2\mu$. Then, proceeding as in [19, Section 3.2], that is applying Lemma 5.1 and Korn's second inequality (cf. Lemma 5.2), it follows from (5.6) and (5.7) that

$$A((\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\eta}), (\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\eta})) \geq \alpha_1 \|\boldsymbol{\tau}\|_{\operatorname{div}; \Omega}^2 + \{\kappa_1 \tilde{\kappa} - \kappa_3 (1 - \tilde{\kappa})\} \|\mathbf{v}\|_{1, \Omega}^2 + \kappa_3 \|\boldsymbol{\eta}\|_{0, \Omega}^2, \quad (5.8)$$

where $\alpha_1 := C_1 \min \left\{ \frac{1}{2\mu} \left(1 - \frac{\kappa_1}{2\mu}\right), \kappa_2 \right\}$. In this way, the bilinear form becomes V -elliptic for the ranges of the parameters given by $0 < \kappa_1 < 2\mu$, $0 < \kappa_2$, and $\kappa_3 > 0$ if $\tilde{\kappa} \geq 1$, or $0 < \kappa_3 < \frac{\kappa_1 \tilde{\kappa}}{1 - \tilde{\kappa}}$ if $\tilde{\kappa} < 1$. Next, the inf-sup condition for B is established directly, because taking $(\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\eta}) = (0, \boldsymbol{\xi}, 0)$ gives

$$\sup_{(\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\eta}) \in H} \frac{B((\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\eta}), \boldsymbol{\xi})}{\|(\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\eta})\|_H} \geq \frac{\|\boldsymbol{\xi}\|_{0, \Omega}^2}{\|\boldsymbol{\xi}\|_{1, \Omega}} \geq c^2 \|\boldsymbol{\xi}\|_{1, \Omega},$$

where the last step comes from the fact that \mathbb{RM} is a finite dimensional space, thus the Open Mapping theorem gives the existence of a constant $c > 0$ such that $\|\boldsymbol{\xi}\|_{0, \Omega} \geq c \|\boldsymbol{\xi}\|_{1, \Omega}$. To end the proof, we note that both bilinear forms have only L^2 inner products, so that their boundedness is directly established and thus Babuška-Brezzi conditions hold, which imply continuous dependence on data. □

The existence of a solution for problem (5.2) is given by Schauder's Fixed Point Theorem.

Lemma 5.3. *Let \mathbf{T} be the augmented fixed point operator given by (5.2) and assume that data assumptions (3.3) and (3.4) hold. Then*

$$\mathbf{T}(\overline{B}(0, r_0)) \subset \overline{B}(0, r_0) := \left\{ \mathbf{v} \in \mathbf{H}^1(\Omega) : \|\mathbf{v}\|_{1,\Omega} \leq r_0 \right\},$$

where $r \leq r_0 := \alpha C M_F$ with C the constant arising from the stability estimate of the augmented problem. In addition, there holds

$$\|\mathbf{T}(z_1) - \mathbf{T}(z_2)\|_{1,\Omega} \leq \alpha \max(1, |\kappa_2|) C L_F \|z_1 - z_2\|_{0,\Omega} \quad \forall z_1, z_2 \in \mathbf{H}^1(\Omega). \quad (5.9)$$

Proof. Let $z \in \mathbf{H}^1(\Omega)$. Then, from the continuous dependence on data and (3.4) it follows:

$$\|\mathbf{T}(z)\|_{1,\Omega} = \|\mathbf{u}\|_{1,\Omega} \leq \alpha C \|F_z\|_{H'} \leq \alpha C M_F =: r_0.$$

Now, given $z_1, z_2 \in \mathbf{H}^1(\Omega)$, we subtract both augmented systems and get a new solution $u_1 - u_2$ with $F := F_{z_1} - F_{z_2}$. Applying continuity on data and (3.3) we get the desired result:

$$\begin{aligned} \|\mathbf{T}(z_1) - \mathbf{T}(z_2)\|_{1,\Omega} &\leq \alpha C \|F\|_{H'} \leq \alpha \max(1, |\kappa_2|) C \|F_{z_1} - F_{z_2}\|_{H'} \\ &\leq \alpha \max(1, |\kappa_2|) C L_F \|z_1 - z_2\|_{0,\Omega}. \end{aligned}$$

□

These results are enough to guarantee the existence of at least one solution of problem (5.2).

Theorem 5.2. *Assume that the data satisfy (3.3) and let $r_0 := \alpha C M_F$. Then the augmented problem (5.2) has at least one solution in $\overline{B}(0, r_0)$, all of which have the following a priori estimate:*

$$\|((\boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\gamma}), \boldsymbol{\xi})\|_{H \times Q} \leq \alpha C M_F.$$

Moreover, if $\alpha C L_F \max(1, |\kappa_2|) < 1$, the solution is unique.

Proof. Let $\{z_k\}_{k \in \mathbb{N}} \subset \overline{B}(0, r_0)$ be a bounded sequence given by iterated solutions of problem (5.2), then it has a subsequence $\{z_k^{(1)}\}_{k \in \mathbb{N}} \subseteq \{z_k\}_{k \in \mathbb{N}}$ weakly convergent to some z in $\mathbf{H}^1(\Omega)$. Rellich-Kondrachov's compactness theorem states that there exists a compact embedding $\mathbf{i}_c : \mathbf{H}^1(\Omega) \rightarrow \mathbf{L}^2(\Omega)$, so that $z_{k_i} \rightarrow z$ in $\mathbf{L}^2(\Omega)$. Using estimate (5.9) from Lemma 5.3 we see that also $\mathbf{T}(z_{k_i}) \rightarrow \mathbf{T}(z)$ in $\mathbf{H}^1(\Omega)$ and so there exists a fixed point due to Schauder's fixed point Theorem. The bound on α is established if \mathbf{T} is forced to be a contraction. □

5.2 Analysis of the discrete problem

For a Galerkin scheme of the augmented formulation, let $\{\mathcal{T}_h\}_{h>0}$ be a regular family of triangularizations of $\overline{\Omega}$ of size h . Now, it is crucial for the analysis to notice that the bilinear form B does not change when a discretization is made because the rigid motions space is already discrete. With this in mind, the inf-sup condition is proved trivially and so the ellipticity of A in the discrete kernel is just a consequence of the continuous case. In virtue of this, any conforming set of discrete spaces will suffice, and as such we take the following:

$$\begin{aligned} \mathbf{H}_h^\sigma &:= \left\{ \boldsymbol{\tau}_h \in \mathbf{H}(\text{div}; \Omega) : \boldsymbol{\tau}_{h,i}|_T \in \mathbb{RT}_0 \quad \forall T \in \mathcal{T}_h \right\}, \\ \mathbf{H}_h^u &:= \left\{ \mathbf{v} \in \mathbf{H}^1(\Omega) : \mathbf{v}|_T \in [\mathbb{P}_1(T)]^n \quad \forall T \in \mathcal{T}_h \right\}, \\ \mathbf{H}_h^\gamma &:= \left\{ \begin{bmatrix} 0 & \psi_h \\ -\psi_h & 0 \end{bmatrix} : \psi_h \in \mathbb{P}_0(T) \quad \forall T \in \mathcal{T}_h \right\}. \end{aligned}$$

Then, setting $H_h = (H_h^\sigma \cap \mathbf{H}_0(\text{div}; \Omega)) \times H_h^v \times H_h^\gamma$, $Q_h = \mathbb{R}\mathbb{M}$, the Galerkin scheme of the augmented partial problem reads: Given $z_h \in H_h^u$, find $((\sigma_h, \mathbf{u}_h, \boldsymbol{\gamma}_h), \boldsymbol{\xi}_h) \in H_h \times Q_h$ such that

$$\begin{aligned} A((\sigma_h, \mathbf{u}_h, \boldsymbol{\gamma}_h), (\boldsymbol{\tau}_h, \mathbf{v}_h, \boldsymbol{\eta}_h)) + B((\boldsymbol{\tau}_h, \mathbf{v}_h, \boldsymbol{\eta}_h), \boldsymbol{\xi}_h) &= \alpha F_{z_h}((\boldsymbol{\tau}_h, \mathbf{v}_h, \boldsymbol{\eta}_h)) \quad \forall (\boldsymbol{\tau}_h, \mathbf{v}_h, \boldsymbol{\eta}_h) \in H_h, \\ B((\sigma_h, \mathbf{u}_h, \boldsymbol{\gamma}_h), \boldsymbol{\rho}_h) &= 0 \quad \forall \boldsymbol{\rho} \in Q_h. \end{aligned} \quad (5.10)$$

We finally define the discrete operator $\mathbf{T}_h : H_h^u \rightarrow H_h^u$ given by

$$\mathbf{T}_h(z_h) = \mathbf{u}_h,$$

where \mathbf{u}_h is the second component of the solution in H_h from the discrete partial augmented problem (5.10), so that the nonlinear discrete problem can be stated as: Find $\mathbf{u} \in \mathbf{H}^1(\Omega)$ such that

$$\mathbf{T}_h(\mathbf{u}) = \mathbf{u}. \quad (5.11)$$

As we did with the mixed formulation, we will first establish the well posedness of the discrete partial problem, and then extend this to the nonlinear problem.

Theorem 5.3. *Given $z \in H_h^u$, problem (5.10) has a unique solution $((\sigma, \mathbf{u}, \boldsymbol{\gamma}), \boldsymbol{\xi}) \in H_h \times Q_h$ and there exists a constant C such that*

$$\|((\sigma_h, \mathbf{u}_h, \boldsymbol{\gamma}_h), \boldsymbol{\xi}_h)\|_{H \times Q} \leq \alpha C \|F_{z_h}\|_{H_h'}.$$

Proof. First we note that $N(\mathbb{B}_h) \subseteq N(\mathbb{B})$. From there, the inf-sup condition and ellipticity are a consequence of the continuous case and thus Babuška-Brezzi conditions hold. \square

Now we establish the regularity of the discrete operator. The following lemma is analogous to the continuous case, so we leave it without proof.

Lemma 5.4. *Let \mathbf{T}_h be the discrete fixed point operator associated to the discrete augmented problem (5.11) and assume that data conditions (3.3) and (3.4) hold. Then*

$$\mathbf{T}_h(\overline{B}_h(0, r_0)) \subset \overline{B}_h(0, r_0) := \{\mathbf{v} \in H_h^u : \|\mathbf{u}\|_{1, \Omega} \leq r_0\},$$

where $r_0 := \alpha C M_F$ with C the constant that arises from the a priori estimate of the partial problem. In addition, there holds

$$\|\mathbf{T}_h(z_1) - \mathbf{T}_h(z_2)\|_{1, \Omega} \leq \alpha \max(1, |\kappa_2|) C L_F \|z_1 - z_2\|_{1, \Omega} \quad \forall z_1, z_2 \in H_h^u.$$

This is enough to prove the well-posedness of the original nonlinear problem, which we show in the following theorem.

Theorem 5.4. *Assuming data conditions (3.3) and (3.4), problem (5.11) has at least one solution. All solutions verify the a priori estimate*

$$\|((\sigma_h, \mathbf{u}_h, \boldsymbol{\gamma}_h), \boldsymbol{\xi}_h)\|_{H \times Q} \leq \alpha C M_F. \quad (5.12)$$

Moreover, if $\alpha C L_F \max(1, |\kappa_2|) < 1$, the solution is unique. If under the same hypothesis $((\sigma, \mathbf{u}, \boldsymbol{\gamma}), \boldsymbol{\xi}) \in H \times Q$ is the unique solution to the continuous problem, Cea's estimate holds for a certain constant $C_4 > 0$:

$$\|((\sigma, \mathbf{u}, \boldsymbol{\gamma}), \boldsymbol{\xi}) - ((\sigma_h, \mathbf{u}_h, \boldsymbol{\gamma}_h), \boldsymbol{\xi}_h)\|_{H \times Q} \leq C_4 \inf_{(\boldsymbol{\tau}_h, \mathbf{v}_h, \boldsymbol{\eta}_h) \in H_h} \|(\sigma, \mathbf{u}, \boldsymbol{\gamma}) - (\boldsymbol{\tau}_h, \mathbf{v}_h, \boldsymbol{\eta}_h)\|_H. \quad (5.13)$$

Proof. First for the a priori estimate, we note from Lemma 5.4 that Brower's fixed point theorem holds, so that there exists a fixed point with

$$\mathbf{T}_h(\mathbf{u}_h) = \mathbf{u}_h,$$

and then the rest of the proof follows exactly as in the mixed case. We obtain the a priori estimate from the partial a priori estimate

$$\|((\boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\gamma}), \boldsymbol{\xi})\|_{H \times Q} \leq \alpha C \|F_{\mathbf{u}_h}\|_{H'_h} \leq \alpha C M_F.$$

In the case of the mixed formulation, proving Cea's estimate was done by taking partial problems and then subtracting. We only need for the same technique to work to have an approximation property for a given \mathbf{z} . Indeed, plugging \mathbf{z} in (5.10), then the solution $((\boldsymbol{\zeta}_h, \mathbf{w}_h, \mathbf{s}_h), \varphi_h) \in H_h \times Q_h$ to such problem would have the desired property for a constant $C_3 > 0$ independent of \mathbf{z} :

$$\begin{aligned} & \|((\boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\gamma}), \boldsymbol{\xi}) - ((\boldsymbol{\zeta}_h, \mathbf{w}_h, \mathbf{s}_h), \varphi_h)\|_{H \times Q} \\ & \leq C_3 \inf_{(\boldsymbol{\tau}_h, \mathbf{v}_h, \boldsymbol{\eta}_h) \in H_h} \|(\boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\gamma}) - (\boldsymbol{\tau}_h, \mathbf{v}_h, \boldsymbol{\eta}_h)\|_H. \end{aligned} \quad (5.14)$$

Using equation (5.14) and proceeding as in the proof of the mixed case, the result holds for the constant $C_4 := \frac{C_3}{1 - \alpha C \max(1, |\kappa_2|) L_F}$. \square

Corollary 5.1. *Under the assumptions from the previous theorem, a sufficiently small α grants C_4 independent of it.*

Proof. The proof is analogous to that of the primal and mixed cases. \square

We end this section with the rates of convergence for the solutions of (5.11). We first recall classic approximation results from [10] and proceed as in [19]:

(AP_h^σ) For each $\boldsymbol{\tau} \in \mathbf{H}^1(\Omega) \cap \mathbf{H}_0$ with $\operatorname{div} \boldsymbol{\tau} \in \mathbf{H}^1(\Omega)$ there exists $\boldsymbol{\tau}_h \in H_{0,h}^\sigma$ such that

$$\|\boldsymbol{\tau} - \boldsymbol{\tau}_h\|_{\operatorname{div}; \Omega} \leq Ch \{ \|\boldsymbol{\tau}\|_{1, \Omega} + \|\operatorname{div} \boldsymbol{\tau}\|_{1, \Omega} \}.$$

(AP_h^u) For each $\mathbf{v} \in \mathbf{H}^1(\Omega)$ there there exists $\mathbf{v}_h \in H_h^u$ such that

$$\|\mathbf{v} - \mathbf{v}_h\|_{0, \Omega} \leq Ch \|\mathbf{v}\|_{1, \Omega}.$$

($AP_h^\boldsymbol{\gamma}$) For each $\boldsymbol{\eta} \in \mathbf{H}^1(\Omega) \cap \mathbf{L}_{asym}^2(\Omega)$ there exists $\boldsymbol{\eta}_h \in H_h^\boldsymbol{\gamma}$ such that

$$\|\boldsymbol{\eta} - \boldsymbol{\eta}_h\|_{0, \Omega} \leq Ch \|\boldsymbol{\eta}\|_{1, \Omega}.$$

Consequently, we can prove the following result:

Theorem 5.5. *Under data assumptions (3.3) and (3.4) and assuming $\alpha C L_F \max(1, |\kappa_2|) \leq \frac{1}{2}$, let $((\boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\gamma}), \boldsymbol{\rho}) \in H \times Q$ and $((\boldsymbol{\sigma}_h, \mathbf{u}_h, \boldsymbol{\gamma}_h), \boldsymbol{\rho}_h) \in H_h \times Q_h$ be the unique solutions of the continuous problem (5.2) and the discrete one (5.11) respectively. Then, there exists a constant $C > 0$ independent of h such that whenever $\boldsymbol{\sigma} \in \mathbf{H}^1(\Omega)$, $\operatorname{div} \boldsymbol{\sigma} \in \mathbf{H}^1(\Omega)$, $\mathbf{u} \in \mathbf{H}^2(\Omega)$ and $\boldsymbol{\gamma} \in \mathbf{H}^1(\Omega)$, there holds*

$$\|((\boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\gamma}), \boldsymbol{\xi}) - ((\boldsymbol{\sigma}_h, \mathbf{u}_h, \boldsymbol{\gamma}_h), \boldsymbol{\xi}_h)\|_{H \times Q} \leq Ch \left\{ \|\boldsymbol{\tau}\|_{1, \Omega} + \|\operatorname{div} \boldsymbol{\tau}\|_{1, \Omega} + \|\mathbf{v}\|_{1, \Omega} + \|\boldsymbol{\eta}\|_{1, \Omega} \right\}. \quad (5.15)$$

Proof. The proof relies in the application of Cea's estimate together with the approximation properties (AP_h^σ), (AP_h^u) and ($AP_h^\boldsymbol{\gamma}$). \square

6 Numerical simulations

The registration problem was implemented in `python` using the library FEniCS ([2], for reference see [29]). Numerical results were visualized using Paraview [1]. The numerical solution of the discrete problems was computed using a regularization technique of the gradient-flow type, which is extensively used in the DIR community (see [36], also known as the proximal point method in the optimization community [34]). To this end, we consider an artificial time variable such that the velocity of the movement points towards the greatest-descent direction of the original functional and then discretize in time using an implicit Euler method. The resulting iterative scheme (time-discretized gradient-flow problem) reads

$$\frac{1}{\Delta t} \langle \mathbf{u} - \mathbf{u}_n, \mathbf{v} \rangle = -a(\mathbf{u}, \mathbf{v}) + f_{\mathbf{u}}(\mathbf{v}) \quad \forall \mathbf{v} \in \mathcal{V}. \quad (6.1)$$

It is easy to show that the previous problem derives from the following variational principle

$$\min_{\mathbf{u} \in \mathcal{V}} \frac{1}{2} a(\mathbf{u}, \mathbf{u}) - \mathcal{D}[\mathbf{u}] + \frac{1}{2\Delta t} \|\mathbf{u} - \mathbf{u}_n\|_{0,\Omega}^2 := \Pi_n[\mathbf{u}], \quad (6.2)$$

where Δt is a regularizing parameter chosen so that the incremental functional is strictly convex [24], guaranteeing the convergence of the iterative problem. The stop criterion is based on the L^2 error of the solution to (6.2) with respect to the previous converged solution.

In the following, the DIR methods developed in this work are tested in the registration of a synthetic image. To this end, we let $\Omega = [0, 1]^2$, and consider the reference image described by

$$R(x_0, x_1) = \sin(2\pi x_0) \sin(2\pi x_1). \quad (6.3)$$

To construct the target image, we first consider an affine displacement field of the form

$$\mathbf{u}(x) = \begin{bmatrix} 0.4 & 0 \\ 0 & -0.2 \end{bmatrix} \mathbf{x}, \quad (6.4)$$

where $\mathbf{x} = (x_0, x_1)^\dagger$. Then, the target image is found by composition, i.e., $T = R \circ (\text{id} + \mathbf{u})^{-1}$. Figure 1 shows the resulting reference and target images.

We perform numerical convergence tests for the three formulations using the images just defined. To this end, a fine mesh with 65,536 triangular elements and characteristic length $h = 0.0055$ is considered as the baseline solution. The degrees of freedom involved in each case are denoted by N and the corresponding errors are quantified as

$$\begin{aligned} e_0(\boldsymbol{\sigma}_h) &:= \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,\Omega}, & e(\boldsymbol{\sigma}_h) &:= \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\text{div};\Omega}, \\ e_0(\mathbf{u}_h) &:= \|\mathbf{u} - \mathbf{u}_h\|_{0,\Omega}, & e(\mathbf{u}_h) &:= \|\mathbf{u} - \mathbf{u}_h\|_{1,\Omega}, \end{aligned}$$

and

$$e(\boldsymbol{\gamma}_h) := \|\boldsymbol{\gamma} - \boldsymbol{\gamma}_h\|_{0,\Omega},$$

where $(\boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\gamma})$ is the baseline solution to the mixed or the augmented case when it corresponds, and $(\boldsymbol{\sigma}_h, \mathbf{u}_h, \boldsymbol{\gamma}_h)$ is the associated finite element solution with coarser mesh sizes. We further define the experimental rates of numerical convergence as

$$\begin{aligned} r_0(\boldsymbol{\sigma}_h) &:= \frac{\log(e_0(\boldsymbol{\sigma})/e'_0(\boldsymbol{\sigma}_h))}{\log(h/h')}, & r(\boldsymbol{\sigma}_h) &:= \frac{\log(e(\boldsymbol{\sigma})/e'(\boldsymbol{\sigma}_h))}{\log(h/h')}, \\ r_0(\mathbf{u}_h) &:= \frac{\log(e_0(\mathbf{u})/e'_0(\mathbf{u}_h))}{\log(h/h')}, & r(\mathbf{u}_h) &:= \frac{\log(e(\mathbf{u})/e'(\mathbf{u}_h))}{\log(h/h')}, \end{aligned}$$

and

$$r(\boldsymbol{\gamma}_h) := \frac{\log(e(\boldsymbol{\gamma}_h)/e'(\boldsymbol{\gamma}_h))}{\log(h/h')},$$

where h and h' denote two consecutive mesh sizes with corresponding errors e and e' . The experiments were performed using the primal, mixed and augmented schemes with $\alpha = 0.5$, $\Delta t = 0.1$ and $\mu = \lambda = 0.5$, and the termination criterion used was a relative error of 10^{-8} . In turn, the stabilization parameters κ_i , $i \in \{1, 2, 3\}$, of the augmented scheme were chosen according to the ranges derived in the proof of Theorem 5.1. Figures 2, 3 and 4 show the results of the numerical convergence studies measured under several error norms for the primal, mixed and augmented schemes, respectively. In all cases, the error norms monotonically decrease as the number of degrees of freedom increases (equivalently, the mesh size is reduced). The convergence rates obtained for the primal, mixed and augmented schemes are documented in Tables 1, 2, and 3, respectively. We note that, in most of the cases, the expected (linear) convergence rate is met by the numerical tests, and sometimes exceeded. The exception is the $H(\text{div})$ norm of $\boldsymbol{\sigma}$ in the mixed and augmented schemes where convergence rates are positive but well below 1, which simply says that the computation of the rates of convergence for this unknown depends strongly on how accurately the continuous solution is approximated.

Figure 5 shows the reference image, and the composed image $T(\mathbf{x} + \mathbf{u}_h(\mathbf{x}))$ for all three formulations using a structured mesh with 2048 elements, for the sake of comparison of the registration solution between methods. All three DIR methods delivered very similar solutions for the composed target image, which are qualitatively similar to the reference image. The associated stress fields are depicted in Figure (6) in terms of the Frobenius norm of the stress tensor. Stress fields are displayed using an L^2 -projection to a \mathbb{P}_1 FE space constructed on the mesh, for visualization purposes. While all three solutions qualitatively agree, the solution of the primal scheme differs from the solutions provided by the mixed and augmented schemes, which are nearly identical.

7 Discussion

In this work we have proposed and analyzed a mixed and augmented formulations for the DIR problem, along with suitable finite-element discretization for their numerical solution. To this end, we consider the variational formulation of the Euler-Lagrange equations associated to the original DIR problem, which presents a structure similar to that of a linear elasticity problem with a nonlinear load source. In this form, we leverage a large body of results in mathematical and numerical analysis to study the DIR problem, namely the Babuška-Brezzi theory for mixed formulations as well as fixed-point theorems. Two key assumptions needed to prove existence of solution, both in the continuous and discrete settings, are the Lipschitz continuity and boundedness of the distributional gradient of the similarity measure, which we argue can be easily verified in DIR applications where images are typically constructed using smooth interpolation schemes. This result, which we particularize for similarity measures of the sum-of-squares-difference's type, is readily extendable to other standard similarity measures, such as correlation and mutual information. For the case of the elastic regularization term, we assume the bijectivity of its Gâteaux gradient, possibly quotiented by its kernel, or simply its surjectivity and the availability of a space reduction that renders the operator injective. Under these assumptions, fixed point arguments allowed us to establish well-posedness of the continuous and discrete problems. A key result of our work, that can be exploited in the feasibility analysis of DIR problems, is the necessary condition for uniqueness of the DIR solution: $\alpha CL_F < 1$. This inequality establishes a novel relation between the similarity weighting constant α and the Lipschitz constant of the nonlinear source term L_F given by $\alpha \propto L_F^{-1}$. The Lipschitz constant, is interpreted as the largest

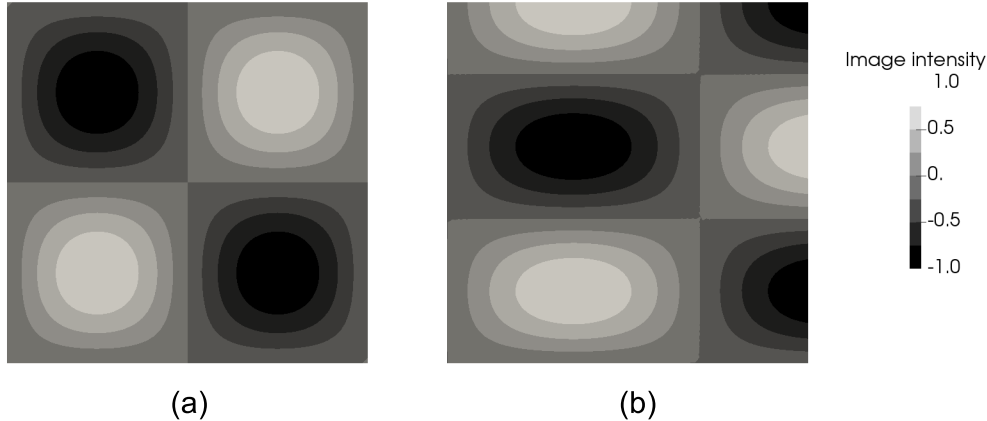


Figure 1: Synthetic images used in numerical tests: (a) Reference image, (b) Target image.

slope between two points as seen in the definition:

$$\frac{\|F(\mathbf{u}) - F(\mathbf{v})\|}{\|\mathbf{u} - \mathbf{v}\|} \leq L_F \quad \forall \mathbf{u}, \mathbf{v}.$$

In virtue of the preceding discussion, we show that DIR of images with higher gradients (i.e., high image contrast, or rapidly oscillating intensity fields due to noise) require the reducing the relevance of the similarity term for well-posedness to hold. This result provides a new insight for the success of DIR, for which a standard practice is to preprocess the reference and target images to reduce image contrast. One such approach is pyramidal gaussian convolution, where images are filtered and then subsampled to perform a sequence of chained registrations, with increasing level of detail, but more sophisticated strategies have been proposed to reduce noise [32, 4].

Another key contribution resulting from this work is the formulation and analysis of mixed and augmented numerical schemes for DIR, where convergence can be proven not only for the transformation (displacement) mapping but also for the stress (and in turn strain) and the rotation tensor. Our results show that, while the transformation mapping that results from solving the primal DIR problem can be very similar to those obtained from the mixed and augmented formulation, the stress fields can be different. Traditionally, the transformation mapping has been the fundamental field sought in DIR applications, particularly in medical applications where the goal is to align the reference and target images [39]. However, recent applications of medical image quantification have highlighted the importance of guaranteeing certain accuracy and convergence when estimating stress tensor fields [38] and rotation tensor fields [37], due to their important connection to medical conditions. Further developments to improve the accuracy of the numerical solution, that are naturally developed within the finite-element framework adopted in this work, are the introduction of *a-posteriori* mesh refinement methods, where recent results in linear elasticity for mixed formulations with Neumann boundary conditions [16] can be extended to the case of DIR problems addressed here.

References

- [1] J. AHRENS, B. GEVECI, AND C. LAW, *Paraview: An end-user tool for large data visualization*, The Visualization Handbook, 717 (2005).

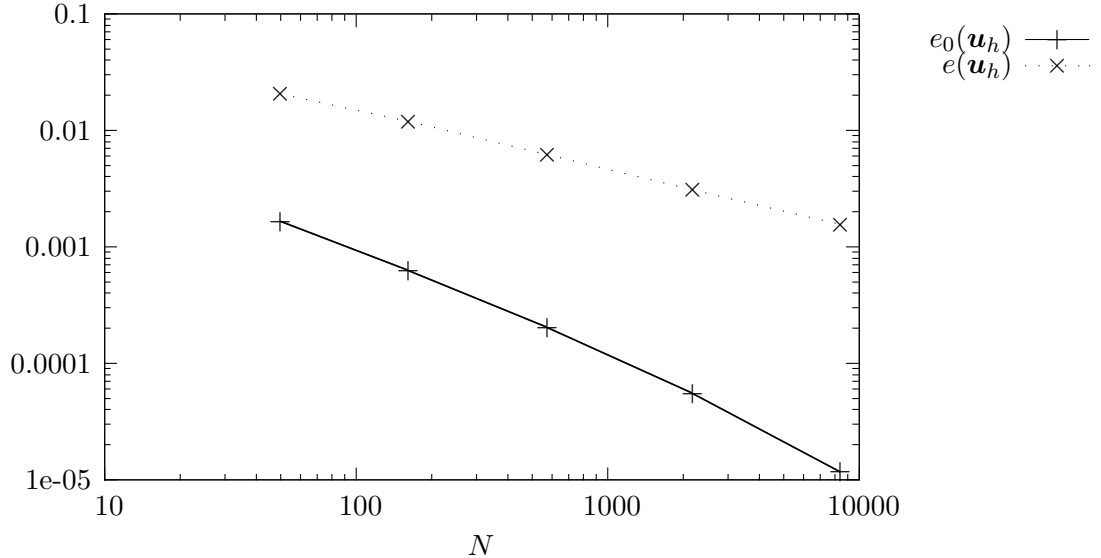


Figure 2: Errors vs. N for the primal formulation

N	$e_0(\mathbf{u}_h)$	$r_0(\mathbf{u}_h)$	$e(\mathbf{u}_h)$	$r(\mathbf{u}_h)$
50	1.66E-3	—	2.06E-2	—
162	6.23E-4	1.42	1.19E-2	0.80
578	2.02E-4	1.62	6.18E-3	0.94
2178	5.55E-5	1.86	3.08E-3	1.01
8450	1.16E-5	2.26	1.55E-3	0.99

Table 1: Errors and convergence rates for the primal scheme.

- [2] 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*, Archive of Numerical Software, 3 (2015).
- [3] R. E. AMELON, K. CAO, K. DING, G. E. CHRISTENSEN, J. M. REINHARDT, AND M. L. RAGHAVAN, *Three-dimensional characterization of regional lung deformation*, Journal of Biomechanics, 44 (2011), pp. 2489–2495.
- [4] P. ATHAVALE AND E. TADMOR, *Integro-differential equations based on (BV, L^1) image decomposition*, SIAM Journal on Imaging Sciences, 4 (2011), pp. 300–312.
- [5] G. AUBERT AND P. KORNPBST, *Mathematical Problems in Image Processing*, vol. 147 of Applied Mathematical Sciences, Springer, New York, second ed., 2006. Partial differential equations and the calculus of variations, With a foreword by Olivier Faugeras.
- [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*, Numerische Mathematik, 91 (2002), pp. 197–222.
- [7] A. BORZI, K. ITO, AND K. KUNISCH, *Optimal control formulation for determining optical flow*, SIAM Journal on Scientific Computing, 24 (2002), pp. 818–847.

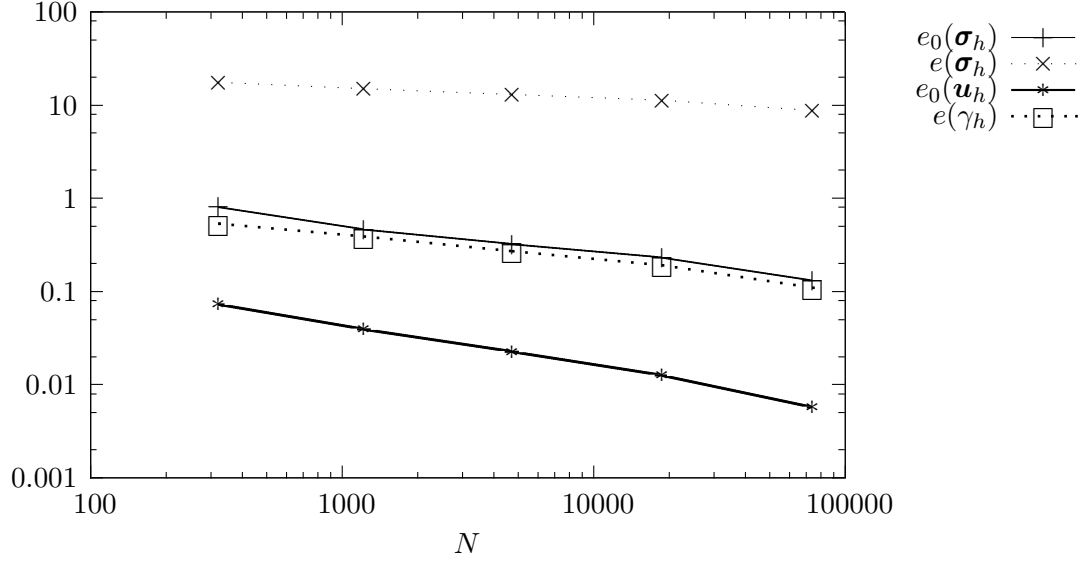


Figure 3: Errors vs. N for the mixed formulation

N	$e_0(\boldsymbol{\sigma}_h)$	$r_0(\boldsymbol{\sigma}_h)$	$e(\boldsymbol{\sigma}_h)$	$r(\boldsymbol{\sigma}_h)$	$e_0(\mathbf{u}_h)$	$r_0(\mathbf{u}_h)$	$e(\boldsymbol{\gamma}_h)$	$r(\boldsymbol{\gamma}_h)$
323	0.80	—	17.49	—	7.25E-2	—	0.53	—
1219	0.46	0.81	14.88	0.23	3.93E-2	0.88	0.39	0.45
4739	0.32	0.50	12.94	0.20	2.25E-2	0.81	0.27	0.52
18691	0.23	0.50	11.34	0.19	1.26E-2	0.83	0.19	0.51
74243	0.13	0.85	8.69	0.38	5.72E-3	1.14	0.11	0.75

Table 2: Errors and convergence rates for the mixed scheme.

- [8] S. C. BRENNER AND L. R. SCOTT, *The Mathematical Theory of Finite Element Methods*, vol. 15 of Texts in Applied Mathematics, Springer, New York, third ed., 2008.
- [9] F. BREZZI, J. DOUGLAS, JR., AND L. D. MARINI, *Two families of mixed finite elements for second order elliptic problems*, Istituto di Analisi Numerica del Consiglio Nazionale delle Ricerche [Institute of Numerical Analysis of the National Research Council], 435 (1984), p. 32.
- [10] F. BREZZI AND M. FORTIN, *Mixed and Hybrid Finite Element Methods*, vol. 15 of Springer Series in Computational Mathematics, Springer-Verlag, New York, 1991.
- [11] M. BURGER, J. MODERSITZKI, AND L. RUTHOTTO, *A hyperelastic regularization energy for image registration*, SIAM Journal on Scientific Computing, 35 (2013), pp. B132–B148.
- [12] K. CHEN AND D. A. LORENZ, *Image sequence interpolation using optimal control*, Journal of Mathematical Imaging and Vision, 41 (2011), pp. 222–238.
- [13] S. CHOI, E. A. HOFFMAN, S. E. WENZEL, M. H. TAWHAI, Y. YIN, M. CASTRO, AND C. L. LIN, *Registration-based assessment of regional lung function via volumetric CT images of normal subjects vs. severe asthmatics*, Journal of Applied Physiology, 115 (2013), pp. 730–742.

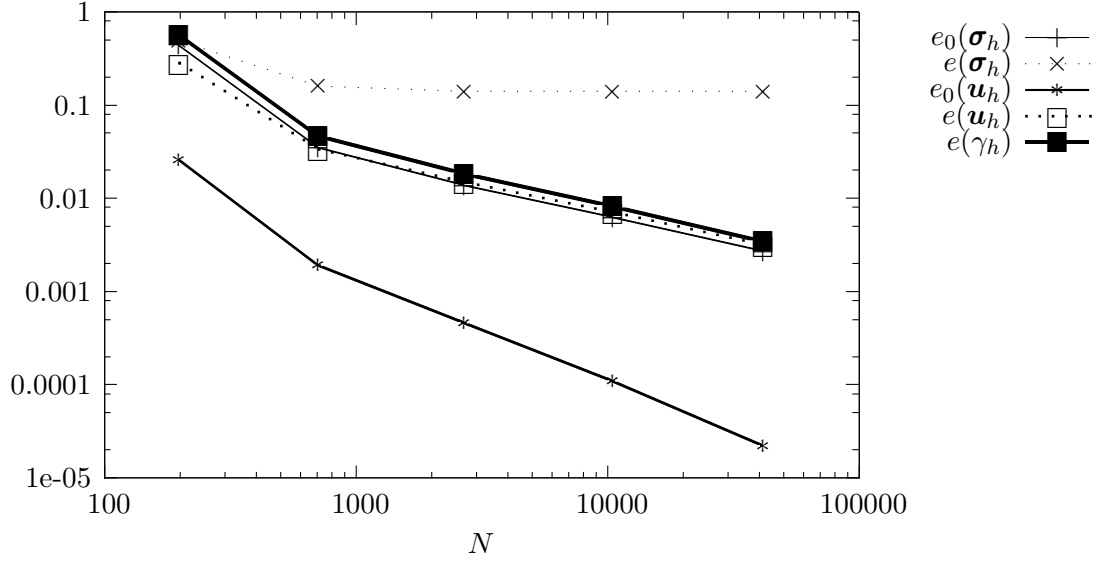


Figure 4: Errors vs. N for the augmented formulation

N	$e_0(\boldsymbol{\sigma}_h)$	$r_0(\boldsymbol{\sigma}_h)$	$e(\boldsymbol{\sigma}_h)$	$r(\boldsymbol{\sigma}_h)$	$e_0(\mathbf{u}_h)$	$r_0(\mathbf{u}_h)$	$e(\mathbf{u}_h)$	$r(\mathbf{u}_h)$	$e(\boldsymbol{\gamma}_h)$	$r(\boldsymbol{\gamma}_h)$
197	4.40E-1	—	0.49	—	2.58E-2	—	2.80E-1	—	5.70E-1	—
709	3.51E-2	3.65	0.16	1.570	1.91E-3	3.76	3.33E-2	3.08	4.66E-2	3.62
2693	1.38E-2	1.35	0.14	0.190	4.57E-4	2.06	1.48E-2	1.17	1.82E-2	1.36
10501	6.26E-3	1.14	0.14	0.061	1.09E-4	2.06	7.06E-3	1.07	8.14E-3	1.16
41477	2.71E-3	1.21	0.14	0.004	2.22E-5	2.30	3.14E-3	1.17	3.43E-3	1.25

Table 3: Errors and convergence rates for the augmented scheme.

- [14] 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.
- [15] P. G. CIARLET, *Linear and Nonlinear Functional Analysis with Applications*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 2013.
- [16] C. DOMÍNGUEZ, G. N. GATICA, AND A. MÁRQUEZ, *A residual-based a posteriori error estimator for the plane linear elasticity problem with pure traction boundary conditions*, Journal of Computational and Applied Mathematics, 292 (2016).
- [17] M. DROSKE AND W. RING, *A Mumford-Shah level-set approach for geometric image registration*, SIAM Journal on Applied Mathematics, 66 (2006), pp. 2127–2148.
- [18] P. DUPUIS, U. GRENANDER, AND M. I. MILLER, *Variational problems on flows of diffeomorphisms for image matching*, Quarterly of Applied Mathematics, 56 (1998), pp. 587–600.
- [19] G. N. GATICA, *Analysis of a new augmented mixed finite element method for linear elasticity allowing \mathbb{RT}_0 - \mathbb{P}_1 - \mathbb{P}_0 approximations*, M2AN Mathematical Modelling and Numerical Analysis, 40 (2006), pp. 1–28.

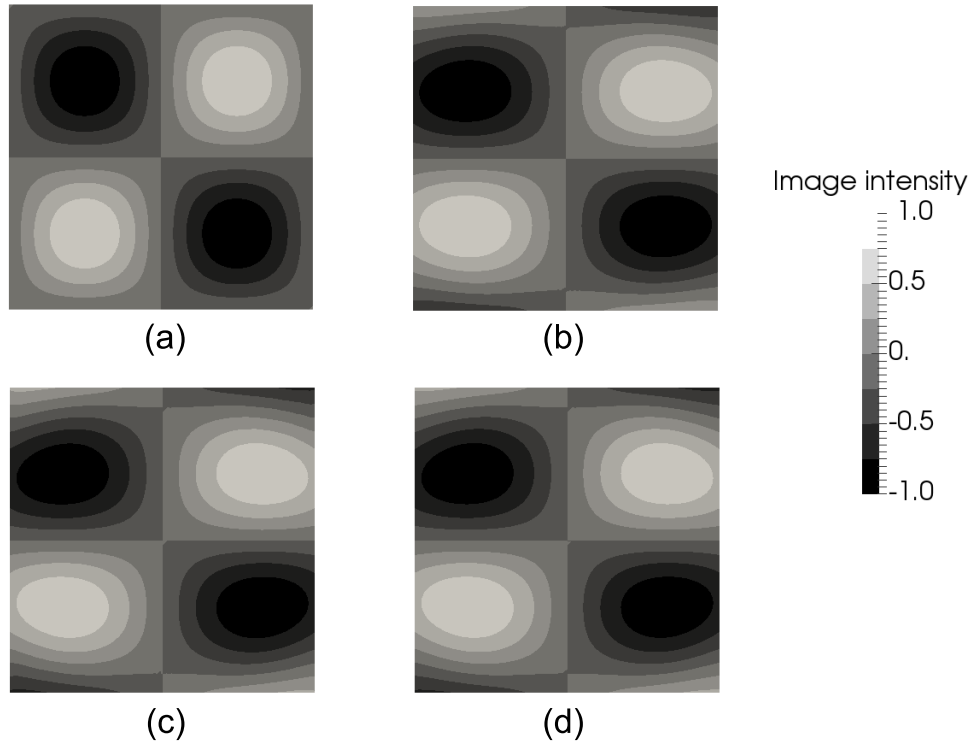


Figure 5: Registration results and comparison for $\alpha = \mu = \lambda = 0.5$. (a) Reference image R . Composed image $T(\mathbf{x} + \mathbf{u}_h(\mathbf{x}))$ using solutions of (b) the primal method, (c) the mixed method, and (d) the augmented method.

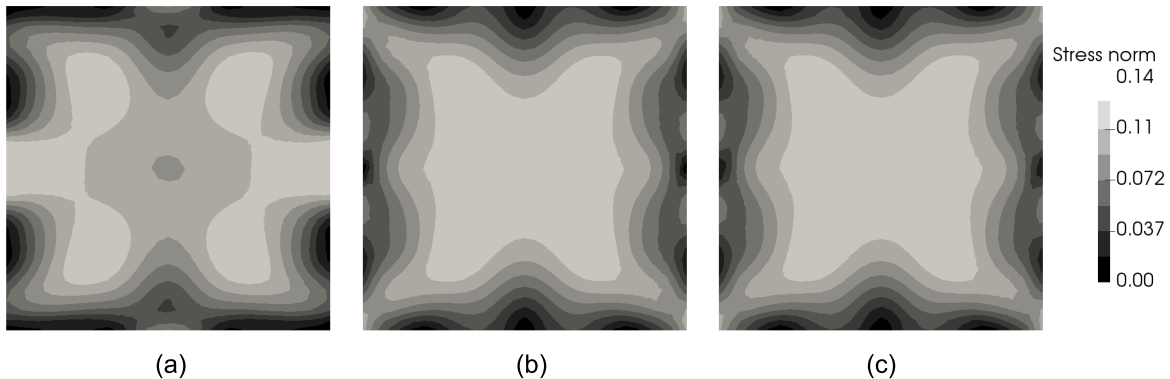


Figure 6: Frobenius norm of the stress tensor field for $\alpha = \mu = \lambda = 0.5$ and $\Delta t = 0.1$ with 2^5 elements per side: (a) primal scheme, (b) mixed scheme, and (c) augmented scheme. Stress fields are displayed using an L^2 -projection to a \mathbb{P}_1 FE space constructed on the mesh.

[20] —, *A Simple Introduction to the Mixed Finite Element Method. Theory and Applications*, (2014), pp. xii+132.

[21] 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*, *Computer Methods in Applied Mechanics and Engineering*, 197 (2008), pp. 1115–1130.

- [22] S. HAKER, L. ZHU, A. TANNENBAUM, AND S. ANGENENT, *Optimal mass transport for registration and warping*, International Journal of Computer Vision, 60 (2004), pp. 225–240.
- [23] B. K. HORN AND B. G. SCHUNCK, *Determining Optical Flow*, tech. rep., Cambridge, MA, USA, 1980.
- [24] D. E. HURTADO AND D. HENAO, *Gradient flows and variational principles for cardiac electrophysiology: toward efficient and robust numerical simulations of the electrical activity of the heart*, Computer Methods in Applied Mechanics and Engineering, 273 (2014), pp. 238–254.
- [25] D. E. HURTADO, N. VILLARROEL, C. ANDRADE, J. RETAMAL, G. BUGEDO, AND A. R. BRUHN, *Spatial patterns and frequency distributions of regional deformation in the healthy human lung.*, Biomechanics and Modeling in Mechanobiology, 16 (2017), pp. 1413–1423.
- [26] D. E. HURTADO, N. VILLARROEL, J. RETAMAL, G. BUGEDO, AND A. BRUHN, *Improving the accuracy of registration-based biomechanical analysis: A finite element approach to lung regional strain quantification*, IEEE Transactions on Medical Imaging, 35 (2016), pp. 580–588.
- [27] E. LEE AND M. GUNZBURGER, *An optimal control formulation of an image registration problem*, Journal of Mathematical Imaging and Vision, 36 (2010), pp. 69–80.
- [28] ———, *Analysis of finite element discretizations of an optimal control formulation of the image registration problem*, SIAM Journal on Numerical Analysis, 49 (2011), pp. 1321–1349.
- [29] A. LOGG, K.-A. MARDAL, G. N. WELLS, ET AL., *Automated Solution of Differential Equations by the Finite Element Method*, Springer, 2012.
- [30] J. MODERSITZKI, *Numerical Methods for Image Registration*, Numerical Mathematics and Scientific Computation, Oxford University Press, New York, 2004. Oxford Science Publications.
- [31] O. MUSEYKO, M. STIGLMAYR, K. KLAMROTH, AND G. LEUGERING, *On the application of the Monge-Kantorovich problem to image registration*, SIAM Journal on Imaging Sciences, 2 (2009), pp. 1068–1097.
- [32] D. PAQUIN, D. LEVY, AND L. XING, *Multiscale deformable registration of noisy medical images*, Mathematical Biosciences and Engineering. MBE, 5 (2008), pp. 125–144.
- [33] C. PÖSCHL, J. MODERSITZKI, AND O. SCHERZER, *A variational setting for volume constrained image registration*, Inverse Problems and Imaging, 4 (2010), pp. 505–522.
- [34] R. T. ROCKAFELLAR, *Monotone operators and the proximal point algorithm*, SIAM Journal on Control and Optimization, 14 (1976), pp. 877–898.
- [35] P. RUHNAU AND C. SCHNÖRR, *Optical Stokes flow estimation: an imaging-based control approach*, Experiments in Fluids, 42 (2007), pp. 61–78.
- [36] A. SCHMIDT-RICHBERG, *Registration Methods for Pulmonary Image Analysis: Integration of Morphological and Physiological Knowledge*, Springer Science & Business Media, 2014.
- [37] J. SOTELO, J. URBINA, I. VALVERDE, J. MURA, C. TEJOS, P. IRARRAZAVAL, M. E. ANDIA, D. E. HURTADO, AND S. URIBE, *Three-dimensional quantification of vorticity and helicity from 3D cine PC-MRI using finite-element interpolations*, Magnetic Resonance in Medicine, (2017).

- [38] J. SOTELO, J. URBINA, I. VALVERDE, C. TEJOS, P. IRARRAZAVAL, M. E. ANDIA, S. URIBE, AND D. E. HURTADO, *3D Quantification of wall shear stress and oscillatory shear index using a finite-element method in 3D CINE PC-MRI data of the thoracic aorta*, IEEE Transactions on Medical Imaging, 35 (2016), pp. 1475–1487.
- [39] A. SOTIRAS, C. DAVATZIKOS, AND N. PARAGIOS, *Deformable medical image registration: A survey*, IEEE Transactions on Medical Imaging, 32 (2013), pp. 1153–1190.
- [40] O. STEINBACH, *Numerical Approximation Methods for Elliptic Boundary Value Problems: Finite and Boundary Elements*, Springer Science & Business Media, 2007.
- [41] A. TROUVÉ, *Diffeomorphisms groups and pattern matching in image analysis*, International Journal of Computer Vision, 28 (1998), pp. 213–221.
- [42] G. UNAL AND G. SLABAUGH, *Coupled pdes for Non-Rigid Registration and Segmentation*, in 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), vol. 1, June 2005, pp. 168–175 vol. 1.
- [43] L. A. VESE AND C. LE GUYADER, *Variational Methods in Image Processing*, Chapman & Hall/CRC Mathematical and Computational Imaging Sciences, CRC Press, Boca Raton, FL, 2016.