ESAIM: M2AN 50 (2016) 311–336
DOI: 10.1051/m2an/2015046

ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

A HYBRID-MIXED METHOD FOR ELASTICITY *

Christopher Harder[1], Alexandre L. Madureira[2] and Frédéric Valentin[3]

Abstract. This work presents a family of stable finite element methods for two- and three-dimensional linear elasticity models. The weak form posed on the skeleton of the partition is a byproduct of the primal hybridization of the elasticity problem. The unknowns are the piecewise rigid body modes and the Lagrange multipliers used to relax the continuity of displacements. They characterize the exact displacement through a direct sum of rigid body modes and solutions to local elasticity problems with Neumann boundary conditions driven by the multipliers. The local problems define basis functions which are in a one-to-one correspondence with the basis of the subspace of Lagrange multipliers used to discretize the problem. Under the assumption that such a basis is available exactly, we prove that the underlying method is well posed, and the stress and the displacement are super-convergent in natural norms driven by (high-order) interpolating multipliers. Also, a local post-processing computation yields strongly symmetric stress which is in local equilibrium and possesses continuous traction on faces. A face-based *a posteriori* estimator is shown to be locally efficient and reliable with respect to the natural norms of the error. Next, we propose a second level of discretization to approximate the basis functions. A two-level numerical analysis establishes sufficient conditions under which the well-posedness and super-convergent properties of the one-level method is preserved.

Mathematics Subject Classification. 65N30, 65N55, 65Y05, 65N12.

Received December 24, 2014. Revised May 22, 2015.
Published online February 5, 2016.

1. Introduction

When modeling elasticity in solid mechanics, the quantity of primary interest is often the stress variable, which should be symmetric and in equilibrium with respect to internal and external forces. Ideally, finite element methods should preserve those fundamental physical properties. However, very few schemes are able to do so and still maintain simplicity in terms of the nature and the quantity of the basis functions and degrees of freedom. A few finite elements satisfying both requirements have been created by either using nested meshes to approach the stress variable [8], augmented spaces [22, 26], or adopting the same mesh for both displacement

Keywords and phrases. Elasticity equation, mixed method, hybrid method, finite element, multiscale.

* *The authors were funded by CNPq/Brazil and CAPES/Brazil.*

[1] Mathematical and Computer Sciences Department, Metropolitan State University of Denver, P.O. Box 173362, Campus Box 38, Denver, CO 80217-3362, USA. harderc@msudenver.edu

[2] National Laboratory for Scientific Computing – LNCC, and Fundação Getúlio Vargas – FGV, Brazil.
alm@lncc.br alexandre.madureira@fgv.br

[3] National Laboratory for Scientific Computing – LNCC, Av. Getúlio Vargas, 333, 25651-070 Petrópolis – RJ, Brazil.
valentin@lncc.br

and stress variables with the price of having to deal with many degrees of freedom [1,5,6]. Recently, a promising methodology closely related to mimetic methods was proposed in [11].

The common approach taken by researchers has been to relax symmetry, local equilibrium, or conformity. These options have been actively researched, each with a volume of work dating from the eighties [2, 7, 33] to the present day [9, 20]. The former idea uses Lagrange multipliers to impose weak symmetry while localizing problems so that local equilibrium is preserved. In other approaches, strong symmetry is achieved by relaxing the conformity of the approach, which leads to the loss of local equilibrium.

We present a new family of finite elements that uses a small number of degrees of freedom. The strategy was introduced for the transport equation in references [3,23,24] and is based on a hybridization scheme [28], which relaxes the continuity of the displacement on element boundaries using Lagrange multipliers. The next step is to decompose the displacement space into a direct sum between the space of local rigid body modes and its orthogonal complement. With such a decomposition we reformulate the original problem in a set of independent, element-wise elasticity problems plus a coupling global system posed on the skeleton of the partition. The unknowns are the piecewise rigid body modes and the Lagrange multipliers used to relax the continuity of displacements. The displacement and stress variables are recovered from them. The local problems are driven by the Lagrange multipliers, which impose traction boundary conditions on element boundaries. Also, the global weak form may be interpreted as the mixed formulation of the original elliptic elasticity problem with a modified right-hand side.

Under the assumption that basis functions are computed exactly, a whole family of stable finite element methods arises from the choice of interpolating space for the Lagrange multipliers. Face and element degrees of freedom define the discrete Lagrange multiplier and the rigid body modes, respectively, in association with the basis functions obtained from the local elasticity problems. The approximation of the stress tensor results from a simple post-processing of the discrete displacement, with strong symmetry being a natural consequence. This strategy leads to $H(\text{div};\Omega)$ conformity for the stresses, and can be interpreted as a H^1-non-conforming well-posed finite element method. It also preserves local equilibrium and strong symmetry, achieves error optimality for the stress and the displacement, and may easily incorporate multiscale or high-contrast aspects of the model.

Such a hybrid-mixed strategy shares some similarities (and the same goals) with other approaches as the Discontinuous Petrov−Galerkin (DPG) method [12, 21] or the Hybrid Discontinuous Galerkin (HDG) method [17, 29, 32]. However, the primal hybridization of the elasticity model selected as the starting point in this work as well as the nature of the solution decomposition leads to different global-local family of methods compared to the ones proposed in the mentioned papers, with fewer degrees of freedom and basis functions. Recently a DPG method has been also proposed for the elliptic Laplace problem [18] differing from [23] in its construction and form. When applied to multiscale or heterogeneous material models, the present method may be seen as a member of the family of multiscale finite element methods [10]. Indeed, it shares a similar structure and the same goals with the multiscale methods proposed in [4, 16], for instance. Thereby, it has been called Multiscale Hybrid-Mixed method (MHM for short). Furthermore, the local computations are completely independent of one another, thereby easily taking advantage of high-performance parallel computing environments.

The impact of second level discretization on the basis functions is also investigated. We provide a sufficient condition for the two-level methods to keep the main features of their one-level counterpart. In fact, properties such as the robustness with respect to the physical coefficients (locking) and local equilibrium are completely locally expressed in terms of the choice made to approximate the local problems with an immediate impact in the global method. In this work, we propose a general framework to analyze such two-level methods and illustrate it using the simplest and cheapest method at the second level, namely, the standard Galerkin method on classical continuous polynomial spaces applied to the solution of the elasticity model in its primal form. This natural choice turns out to be enough to preserve most of the properties of the one-level method. It is important to stress that care should be taken at the second level discretization to preserve some of the nice properties of the method, in particular $H(\text{div};\Omega)$ conformity for the stresses. However, local equilibrium holds in the sense of Remark 6.11. We leave both the study of the use of mixed methods in the second level and the important question of robustness of the proposed method for nearly incompressible materials for forthcoming works.

The main theoretical results of this work are summarized as follows: weak formulation (2.6) and its discrete version, the face-based MHM method (3.1), are proved to be well-posed in Theorem 4.2. We then present a best approximation result showing that the error only depends on the quality of the approximation on faces (Lem. 4.3). This is used to prove that the MHM method provides super-convergent numerical approximations to the displacement and stress variables in natural norms (Thm. 5.2). Furthermore, an *a posteriori* error estimator (see Eqs. (5.4)-(5.6)), established in terms of the jump of the displacement variable on the faces, is shown to control the natural norms of the displacement and stress variables (Thm. 5.5). The two-level version of the MHM method (6.4) is shown to be well-posed in Theorem 6.2 under the space compatibility condition (6.6). We also measure the impact of the second-level discretization, which is related to a consistency error. This is highlighted in Lemma 6.3. Some local spaces fulfilling the local space compatibility assumption are presented in (6.15) and analyzed in Lemmas 6.6 and 6.8. Convergence estimates for these are proved in Lemma 6.10.

The paper is outlined as follows: the remainder of this section presents the necessary steps towards hybridization. An equivalent global-local form of the hybrid formulation and its variants are left to Section 2, while statement of the method is in Section 3. Its well-posedness and best approximation properties are addressed in Section 4. Section 5 is dedicated to *a priori* and *a posteriori* error estimates. The two-level version of the method is presented and analyzed in Section 6. Conclusions follow in Section 7 and some complementary results in the Appendix.

1.1. Statement and preliminaries

In what follows, let $\Omega \subset \mathbb{R}^n$, $n \in \{2, 3\}$, be an open bounded domain with polygonal boundary $\partial\Omega$. We consider the elliptic problem to find the displacement $\boldsymbol{u} : \Omega \to \mathbb{R}^n$ such that

$$\begin{cases} -\mathbf{div}\mathcal{A}\,\mathcal{E}(\boldsymbol{u}) = \boldsymbol{f} & \text{in } \Omega, \\ \boldsymbol{u} = \boldsymbol{g} & \text{on } \partial\Omega, \end{cases} \tag{1.1}$$

where $\boldsymbol{g} \in \boldsymbol{H}^{1/2}(\partial\Omega)$ and $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega)$ are given functions with values in $\mathbb{R}^n$. As such, problem (1.1) has a unique solution $\boldsymbol{u} \in \boldsymbol{H}^1(\Omega)$, where the spaces have their usual meaning. The linearized strain tensor is given by the symmetric part of the gradient

$$\mathcal{E}(\boldsymbol{u}) := \frac{1}{2}(\nabla\boldsymbol{u} + \nabla^T\boldsymbol{u}),\ \textit{i.e.},\quad \big(\mathcal{E}(\boldsymbol{u})\big)_{ij} := \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right)\cdot$$

Above, and throughout the paper, the indices i, j run from $1, \ldots, n$, even when not explicitly mentioned. The fourth-order rigidity tensor $\mathcal{A}$ acts on the space $\mathbb{R}^{n\times n}_{\text{sym}}$ of $n \times n$ symmetric matrices. If $\tau \in \mathbb{R}^{n\times n}_{\text{sym}}$, then $\sigma := \mathcal{A}\,\tau \in \mathbb{R}^{n\times n}_{\text{sym}}$ is such that

$$\sigma_{ij} := \sum_{k,l=1}^{n} \mathcal{A}_{ijkl}\,\tau_{kl}.$$

The rigidity tensor is quite general, possibly depending on $\boldsymbol{x} \in \Omega$ and embedding multiple geometrical scales. However, it satisfies the usual symmetry properties $\mathcal{A}_{ijkl} = \mathcal{A}_{klij} = \mathcal{A}_{jikl} = \mathcal{A}_{ijlk}$, and is uniformly positive definite and bounded, *i.e.*, there exist positive constants $c_{\min}$ and $c_{\max}$ such that

$$c_{\min}^2|\tau|^2 \le \mathcal{A}\,\tau : \tau \le c_{\max}^2|\tau|^2 \quad \text{for all } \tau \in \mathbb{R}^{n\times n}_{\text{sym}}, \tag{1.2}$$

for almost every $\boldsymbol{x} \in \Omega$. Here, $\tau : \sigma := \sum_{i,j=1}^{n} \tau_{ij}\,\sigma_{ij}$ denotes the inner product between the matrices τ, σ, and $|\tau| := (\tau : \tau)^{1/2}$. Finally, for a given matrix σ, the row-wise divergence $\mathbf{div}\,\sigma$ is defined by $(\mathbf{div}\,\sigma)_i := \sum_{j=1}^{n} \partial\sigma_{ij}/\partial x_j$.

It follows from (1.1) that the stress tensor $\sigma := \mathcal{A}\,\mathcal{E}(\boldsymbol{u}) \in H(\mathrm{div};\Omega)$, again with the space taking its usual meaning. However, instead of working directly with this form of the problem, we adopt the following perspective: we seek $\boldsymbol{u}$ as the solution of the elliptic elasticity equation in a weaker, broken space which relaxes continuity, localizes computations, and allows reconstruction of a symmetric stress tensor which preserves equilibrium. Ultimately, the approach allows for the construction of $\boldsymbol{u}$ and σ using local problems. This feature is particularly attractive in the presence of heterogeneous coefficients since fine-scale contributions may be upscaled in parallel.

The first step is to partition the domain Ω with a family of regular meshes $\{\mathcal{T}_h\}_{h>0}$ into elements K, where h is a characteristic length of $\mathcal{T}_h$. The mesh can be very general, composed of heterogeneous element geometries. Without loss of generality, we shall use here the terminology usually employed for three-dimensional domains. As such, each element K has a boundary ∂K consisting of faces F, and we collect in $\mathcal{E}_h$ the faces associated with $\mathcal{T}_h$. Let $\mathcal{E}_D$ be the set of faces on $\partial\Omega$, and $\mathcal{E}_0 = \mathcal{E}_h \setminus \mathcal{E}_D$ be the set of internal faces. To each $F \in \mathcal{E}_h$ we associate a normal $\boldsymbol{n}$, taking care to ensure this is facing outward on $\partial\Omega$. For each $K \in \mathcal{T}_h$, we further denote by $\boldsymbol{n}^K$ the outward normal on ∂K, and let $\boldsymbol{n}^K_F := \boldsymbol{n}^K|_F$ for each $F \subset \partial K$. Also, the space of displacements $\mathbf{V}$ consists of

$$\mathbf{V} := \left\{ \boldsymbol{v} \in \boldsymbol{L}^2(\Omega) \,:\, \boldsymbol{v}\,|_K \in \boldsymbol{H}^1(K), \quad K \in \mathcal{T}_h \right\},$$

and the space of tractions $\boldsymbol{\Lambda}$ is formed as follows

$$\boldsymbol{\Lambda} := \left\{ \sigma\,\boldsymbol{n}^K\,|_{\partial K} \,:\, \sigma \in H(\mathrm{div};\Omega), \quad K \in \mathcal{T}_h \right\}.$$

The definition of the norms for theses spaces is postponed to Section 4. For now, we denote $(.,.)_{\mathcal{T}_h}$ and $(.,.)_{\partial\mathcal{T}_h}$ the summation of the respective inner (or dual) products, for all $K \in \mathcal{T}_h$, over the sets K and ∂K, respectively, namely,

$$(\boldsymbol{w}, \boldsymbol{v})_{\mathcal{T}_h} := \sum_{K\in\mathcal{T}_h} \int_K \boldsymbol{w}\cdot\boldsymbol{v}\,\mathrm{d}\boldsymbol{x} \quad \text{and} \quad (\boldsymbol{\mu}, \boldsymbol{v})_{\partial\mathcal{T}_h} := \sum_{K\in\mathcal{T}_h} (\boldsymbol{\mu}, \boldsymbol{v})_{\partial K},$$

where $\boldsymbol{w}$, $\boldsymbol{v} \in \mathbf{V}$ and $\boldsymbol{\mu} \in \boldsymbol{\Lambda}$, and $(\cdot,\cdot)_{\partial K}$ is the dual product involving $\boldsymbol{H}^{-1/2}(\partial K)$ and $\boldsymbol{H}^{1/2}(\partial K)$ defined as follows

$$(\boldsymbol{\mu}, \boldsymbol{v})_{\partial K} := \int_K \mathbf{div}\,\sigma\cdot\boldsymbol{v}\,\mathrm{d}\boldsymbol{x} + \int_K \sigma : \nabla\boldsymbol{v}\,\mathrm{d}\boldsymbol{x}.$$

We consider the hybrid formulation of problem (1.1): *find* $(\boldsymbol{u}, \boldsymbol{\lambda}) \in \mathbf{V}\times\boldsymbol{\Lambda}$ *such that*

$$\begin{cases} (\mathcal{A}\,\mathcal{E}(\boldsymbol{u}), \mathcal{E}(\boldsymbol{v}))_{\mathcal{T}_h} + (\boldsymbol{\lambda}, \boldsymbol{v})_{\partial\mathcal{T}_h} = (\boldsymbol{f},\, \boldsymbol{v})_{\mathcal{T}_h} & \text{for all } \boldsymbol{v} \in \mathbf{V}, \\ (\boldsymbol{\mu}, \boldsymbol{u})_{\partial\mathcal{T}_h} = (\boldsymbol{\mu},\, \boldsymbol{g})_{\partial\Omega} & \text{for all } \boldsymbol{\mu} \in \boldsymbol{\Lambda}. \end{cases} \tag{1.3}$$

In formulation (1.3), the displacement belongs *a priori* to a larger space than the solution of the original problem (1.1). However, the space of Lagrange multipliers $\boldsymbol{\Lambda}$ imposes $\boldsymbol{H}^1(\Omega)$-conformity on the solution and the boundary condition $\boldsymbol{u} = \boldsymbol{g}$. Also, problem (1.3) has a unique solution $(\boldsymbol{u}, \boldsymbol{\lambda}) \in \mathbf{V}\times\boldsymbol{\Lambda}$ where $\boldsymbol{u} \in \boldsymbol{H}^1(\Omega)$ is also the solution of problem (1.1). Such results are summarized next.

Lemma 1.1. *Assume that* $(\boldsymbol{u}, \boldsymbol{\lambda}) \in \mathbf{V}\times\boldsymbol{\Lambda}$. *Therefore,* $(\boldsymbol{u}, \boldsymbol{\lambda})$ *is the solution of* (1.3) *if and only if* $\boldsymbol{u} \in \boldsymbol{H}^1(\Omega)$ *solves* (1.1). *Furthermore, for all* $K \in \mathcal{T}_h$, *it holds*

$$\boldsymbol{\lambda} = -\mathcal{A}\,\mathcal{E}(\boldsymbol{u})\,\boldsymbol{n}^K \quad \text{on } \partial K.$$

Proof. Following closely the proof of Theorem 1 in [31] the results follow. □

2. A global-local formulation

Rather than selecting a pair of finite subspaces of $\mathbf{V}\times\boldsymbol{\Lambda}$ at this point, we rewrite (1.3) in an equivalent form which is suitable to reduce the statement to a system of locally- and globally-defined problems. Such an approach will guide the definition of stable finite subspaces. The key to the approach is the operator-driven decomposition

$$\mathbf{V}=\mathbf{V}_{\mathrm{rm}}\oplus\widetilde{\mathbf{V}}. \tag{2.1}$$

Here, $\mathbf{V}_{\mathrm{rm}}$ is the finite dimensional subspace of $\mathbf{V}$ composed of those functions $\boldsymbol{v}^{\mathrm{rm}}\in\mathbf{V}$ such that $(\mathcal{A}\,\mathcal{E}(\boldsymbol{v}^{\mathrm{rm}}),\mathcal{E}(\boldsymbol{v}))_{\mathcal{T}_h}=0$ for all $\boldsymbol{v}\in\mathbf{V}$, and $\widetilde{\mathbf{V}}$ is its L^2-orthogonal complement in $\mathbf{V}$. In fact, $\mathbf{V}_{\mathrm{rm}}$ is the space of piecewise rigid body modes, *i.e.*,

$$\mathbf{V}_{\mathrm{rm}}:=\{\boldsymbol{v}\in\mathbf{V}\,:\,\boldsymbol{v}|_K\in\mathbf{V}_{\mathrm{rm}}(K),\,K\in\mathcal{T}_h\},\qquad \mathbf{V}_{\mathrm{rm}}(K):=\{\boldsymbol{v}\in\mathbf{V}\,:\,\mathcal{E}(\boldsymbol{v})\,|_K=0\}.$$

Using decomposition (2.1), problem (1.3) is equivalent to: *find* $(\boldsymbol{u}^{\mathrm{rm}}+\tilde{\boldsymbol{u}},\boldsymbol{\lambda})\in(\mathbf{V}_{\mathrm{rm}}\oplus\tilde{\mathbf{V}})\times\boldsymbol{\Lambda}$ *such that*

$$\begin{cases}(\boldsymbol{\lambda},\boldsymbol{v}^{\mathrm{rm}})_{\partial\mathcal{T}_h}=(\boldsymbol{f},\boldsymbol{v}^{\mathrm{rm}})_{\mathcal{T}_h} & \text{for all } \boldsymbol{v}^{\mathrm{rm}}\in\mathbf{V}_{\mathrm{rm}},\\ (\boldsymbol{\mu},\boldsymbol{u}^{\mathrm{rm}}+\tilde{\boldsymbol{u}})_{\partial\mathcal{T}_h}=(\boldsymbol{\mu},\,\boldsymbol{g})_{\partial\Omega} & \text{for all } \boldsymbol{\mu}\in\boldsymbol{\Lambda},\end{cases} \tag{2.2}$$

$$(\mathcal{A}\,\mathcal{E}(\tilde{\boldsymbol{u}}),\mathcal{E}(\tilde{\boldsymbol{v}}))_{\mathcal{T}_h}+(\boldsymbol{\lambda},\tilde{\boldsymbol{v}})_{\partial\mathcal{T}_h}=(\boldsymbol{f},\tilde{\boldsymbol{v}})_{\mathcal{T}_h}\quad\text{for all } \tilde{\boldsymbol{v}}\in\widetilde{\mathbf{V}}. \tag{2.3}$$

Notice that (2.3) implies that $\tilde{\boldsymbol{u}}$ can be computed in each element from $\boldsymbol{f}$ and from $\boldsymbol{\lambda}$ once the latter is known. In fact, we find from (2.3) that $\tilde{\boldsymbol{u}}=T\,\boldsymbol{\lambda}+\hat{T}\,\boldsymbol{f}$, with $T:\,\boldsymbol{\Lambda}\to\tilde{\mathbf{V}}$ and $\hat{T}:\,\boldsymbol{L}^2(\Omega)\to\tilde{\mathbf{V}}$ being bounded linear operators (see Lems. A.1 and A.2 in the Appendix) defined, on each $K\in\mathcal{T}_h$, by

$$\big(\mathcal{A}\,\mathcal{E}(T\,\boldsymbol{\mu}),\mathcal{E}(\tilde{\boldsymbol{v}})\big)_K=-(\boldsymbol{\mu},\tilde{\boldsymbol{v}})_{\partial K}\quad\text{for all } \tilde{\boldsymbol{v}}\in\widetilde{\mathbf{V}}, \tag{2.4}$$

$$\big(\mathcal{A}\,\mathcal{E}(\hat{T}\,\boldsymbol{q}),\mathcal{E}(\tilde{\boldsymbol{v}})\big)_K=(\boldsymbol{q},\tilde{\boldsymbol{v}})_K\quad\text{for all } \tilde{\boldsymbol{v}}\in\widetilde{\mathbf{V}}, \tag{2.5}$$

for $\boldsymbol{\mu}\in\boldsymbol{\Lambda}$ and $\boldsymbol{q}\in\boldsymbol{L}^2(\Omega)$. We substitute this decomposition $\tilde{\boldsymbol{u}}=T\,\boldsymbol{\lambda}+\hat{T}\,\boldsymbol{f}$ in (2.2) to yield the problem: *find* $(\boldsymbol{u}^{\mathrm{rm}},\boldsymbol{\lambda})\in\mathbf{V}_{\mathrm{rm}}\times\boldsymbol{\Lambda}$ *such that*

$$\begin{cases}(\boldsymbol{\lambda},\boldsymbol{v}^{\mathrm{rm}})_{\partial\mathcal{T}_h}=(\boldsymbol{f},\boldsymbol{v}^{\mathrm{rm}})_{\mathcal{T}_h} & \text{for all } \boldsymbol{v}^{\mathrm{rm}}\in\mathbf{V}_{\mathrm{rm}},\\ (\boldsymbol{\mu},\boldsymbol{u}^{\mathrm{rm}}+T\,\boldsymbol{\lambda})_{\partial\mathcal{T}_h}=-(\boldsymbol{\mu},\hat{T}\,\boldsymbol{f})_{\partial\mathcal{T}_h}+(\boldsymbol{\mu},\,\boldsymbol{g})_{\partial\Omega} & \text{for all } \boldsymbol{\mu}\in\boldsymbol{\Lambda}.\end{cases} \tag{2.6}$$

As a result, the weak formulation (1.3) is analogous to the coupled system (2.4) and (2.5) (with $\boldsymbol{\mu}=\boldsymbol{\lambda}$ and $\boldsymbol{q}=\boldsymbol{f}$) and (2.6). We use the unknowns $\boldsymbol{u}^{\mathrm{rm}}$ and $\boldsymbol{\lambda}$ of the latter problem to reconstruct the displacement $\boldsymbol{u}\in\mathbf{V}$ and the stress tensor $\sigma\in H(\mathrm{div};\Omega)$ as follows:

$$\boldsymbol{u}\;=\;\boldsymbol{u}^{\mathrm{rm}}+T\,\boldsymbol{\lambda}+\hat{T}\,\boldsymbol{f},\quad \sigma\;=\;\mathcal{A}\,\mathcal{E}(T\,\boldsymbol{\lambda}+\hat{T}\,\boldsymbol{f}). \tag{2.7}$$

2.1. From hybrid to mixed

Next, we detail the procedure to translate (2.2) into a mixed formulation. To this end, we rewrite (2.4) and (2.5) in strong form. For that, suppose $\boldsymbol{\mu}\in\boldsymbol{\Lambda}$ and define the (unique) rigid body mode $\mathbf{R}_K^{\mu}\in\mathbf{V}_{\mathrm{rm}}(K)$ in each $K\in\mathcal{T}_h$ by,

$$(\mathbf{R}_K^{\mu},\,\boldsymbol{v}^{\mathrm{rm}})_K=(\boldsymbol{\mu},\,\boldsymbol{v}^{\mathrm{rm}})_{\partial K}\quad\text{for all } \boldsymbol{v}^{\mathrm{rm}}\in\mathbf{V}_{\mathrm{rm}}(K). \tag{2.8}$$

From problem (2.4), we conclude that $T\,\boldsymbol{\mu}$ is the unique solution of the local elasticity problem (in a distributional sense)

$$\begin{cases}-\mathbf{div}\mathcal{A}\,\mathcal{E}(T\,\boldsymbol{\mu})=\mathbf{R}_K^{\mu} & \text{in } K,\\ \mathcal{A}\,\mathcal{E}(T\,\boldsymbol{\mu})\,\boldsymbol{n}^K=-\boldsymbol{\mu} & \text{on } F\subset\partial K.\end{cases} \tag{2.9}$$

Following the same procedure, we may also use (2.5) to rewrite

$$\begin{cases} -\mathbf{div}\mathcal{A}\,\mathcal{E}(\hat{T}\,\boldsymbol{q}) = \boldsymbol{q} - \Pi_K(\boldsymbol{q}) & \text{in } K, \\ \mathcal{A}\,\mathcal{E}(\hat{T}\,\boldsymbol{q})\,\boldsymbol{n}^K = \mathbf{0} & \text{on } F \subset \partial K, \end{cases} \tag{2.10}$$

where $\Pi_K(.)$ is the L^2-orthogonal projection onto $\mathbf{V}_{\rm rm}(K)$.

Now, hidden in the statement of the global problem (2.6) is a mixed form of the elliptic problem (1.1). To highlight this, we use equations (2.8) and (2.9) to first establish that for each $\boldsymbol{v}^{\rm rm} \in \mathbf{V}_{\rm rm}(K)$,

$$(\boldsymbol{\mu}, \boldsymbol{v}^{\rm rm})_{\partial\mathcal{T}_h} = (\mathbf{R}_K^{\mu}, \boldsymbol{v}^{\rm rm})_{\mathcal{T}_h} = -(\mathbf{div}\,\sigma^{\mu}, \boldsymbol{v}^{\rm rm})_{\mathcal{T}_h}, \tag{2.11}$$

where we defined $\sigma^{\mu} := \mathcal{A}\,\mathcal{E}(T\,\boldsymbol{\mu})$, $\boldsymbol{\mu} \in \boldsymbol{\Lambda}$, and used $\boldsymbol{v}^{\rm rm} \in \mathbf{V}_{\rm rm}(K)$. Next, we choose (arbitrarily, and without loss of generality) to lift $\boldsymbol{\mu}$ from ∂K into K by using problem (2.4). This choice conveniently yields a form of (2.6) which is completely defined in terms of integrals on element interiors rather than their boundaries. From (2.4), (2.11) and (2.5) with the fact $T\,\boldsymbol{\lambda} + \hat{T}\,\boldsymbol{f} \in \widetilde{\mathbf{V}}$, it holds that

$$\begin{aligned} \left(\boldsymbol{\mu}, \boldsymbol{u}^{\rm rm} + T\,\boldsymbol{\lambda} + \hat{T}\,\boldsymbol{f}\right)_{\partial\mathcal{T}_h} &= -\left(\mathcal{A}\,\mathcal{E}(T\,\boldsymbol{\mu}),\, \mathcal{E}(T\boldsymbol{\lambda} + \hat{T}\,\boldsymbol{f})\right)_{\mathcal{T}_h} - (\mathbf{div}\,\sigma^{\mu}, \boldsymbol{u}^{\rm rm})_{\mathcal{T}_h} \\ &= -\left(\mathcal{A}\,\mathcal{E}(T\,\boldsymbol{\lambda}), \mathcal{E}(T\,\boldsymbol{\mu})\right)_{\mathcal{T}_h} - (\boldsymbol{f}, T\,\boldsymbol{\mu})_{\mathcal{T}_h} - (\mathbf{div}\,\sigma^{\mu}, \boldsymbol{u}^{\rm rm})_{\mathcal{T}_h}\,. \end{aligned} \tag{2.12}$$

Finally, we gather (2.11) and (2.12) and substitute them into the global problem (2.6) to propose the following equivalent weak mixed form: *find* $(\boldsymbol{\lambda}, \boldsymbol{u}^{\rm rm}) \in \boldsymbol{\Lambda} \times \mathbf{V}_{\rm rm}$ *such that*

$$\begin{cases} (\mathcal{A}^{-1}\sigma^{\lambda}, \sigma^{\mu})_{\mathcal{T}_h} + (\boldsymbol{u}^{\rm rm}, \mathbf{div}\,\sigma^{\mu})_{\mathcal{T}_h} = -(\boldsymbol{f}, T\,\boldsymbol{\mu})_{\mathcal{T}_h} - (\boldsymbol{\mu}, \boldsymbol{g})_{\partial\Omega} & \text{for all } \boldsymbol{\mu} \in \boldsymbol{\Lambda}, \\ (\mathbf{div}\,\sigma^{\lambda}, \boldsymbol{v}^{\rm rm})_{\mathcal{T}_h} = -(\boldsymbol{f}, \boldsymbol{v}^{\rm rm})_{\mathcal{T}_h} & \text{for all } \boldsymbol{v}^{\rm rm} \in \mathbf{V}_{\rm rm}, \end{cases} \tag{2.13}$$

where $\mathcal{A}^{-1}$ is the compliance tensor.

3. The Multiscale Hybrid-Mixed (MHM) method

We have not introduced any discretization up to this point, although global problem (2.6) involves the finite-dimensional space $\mathbf{V}_{\rm rm}$. Since $\boldsymbol{\lambda}$ uniquely determines $T\,\boldsymbol{\lambda}$ (see (2.4)), it is enough to pick a finite element space $\boldsymbol{\Lambda}_h$ in order to fully discretize problem (2.6). In this case, we find the discrete method: *find* $(\boldsymbol{u}_h^{\rm rm}, \boldsymbol{\lambda}_h) \in \mathbf{V}_{\rm rm} \times \boldsymbol{\Lambda}_h$ *such that*

$$\begin{cases} (\boldsymbol{\lambda}_h, \boldsymbol{v}^{\rm rm})_{\partial\mathcal{T}_h} = (\boldsymbol{f}, \boldsymbol{v}^{\rm rm})_{\mathcal{T}_h} & \text{for all } \boldsymbol{v}^{\rm rm} \in \mathbf{V}_{\rm rm}, \\ (\boldsymbol{\mu}_h, \boldsymbol{u}_h^{\rm rm} + T\,\boldsymbol{\lambda}_h)_{\partial\mathcal{T}_h} = -(\boldsymbol{\mu}_h, \hat{T}\,\boldsymbol{f})_{\partial\mathcal{T}_h} + (\boldsymbol{\mu}_h, \boldsymbol{g})_{\partial\Omega} & \text{for all } \boldsymbol{\mu}_h \in \boldsymbol{\Lambda}_h, \end{cases} \tag{3.1}$$

where T and $\hat{T}$ are as in (2.4) and (2.5), respectively. We use the unknowns of this problem to construct approximations to $\boldsymbol{u}$ and σ given in (2.7), namely,

$$\boldsymbol{u}_h := \boldsymbol{u}_h^{\rm rm} + T\,\boldsymbol{\lambda}_h + \hat{T}\,\boldsymbol{f}, \quad \sigma_h := \mathcal{A}\,\mathcal{E}(T\,\boldsymbol{\lambda}_h + \hat{T}\,\boldsymbol{f}). \tag{3.2}$$

Remark 3.1. Observe we may recast the MHM method (3.1) in the same mixed form (2.13), *i.e.*, find $(\boldsymbol{\lambda}_h, \boldsymbol{u}^{\rm rm}) \in \boldsymbol{\Lambda}_h \times \mathbf{V}_{\rm rm}$ such that

$$\begin{cases} (\mathcal{A}^{-1}\sigma^{\lambda_h}, \sigma^{\mu_h})_{\mathcal{T}_h} + (\boldsymbol{u}^{\rm rm}, \mathbf{div}\,\sigma^{\mu_h})_{\mathcal{T}_h} = -(\boldsymbol{f}, T\,\boldsymbol{\mu}_h)_{\mathcal{T}_h} - (\boldsymbol{\mu}, \boldsymbol{g})_{\partial\Omega} & \text{for all } \boldsymbol{\mu}_h \in \boldsymbol{\Lambda}_h, \\ (\mathbf{div}\,\sigma^{\lambda_h}, \boldsymbol{v}^{\rm rm})_{\mathcal{T}_h} = -(\boldsymbol{f}, \boldsymbol{v}^{\rm rm})_{\mathcal{T}_h} & \text{for all } \boldsymbol{v}^{\rm rm} \in \mathbf{V}_{\rm rm}. \end{cases} \tag{3.3}$$

Note that the second equation in (3.3) imposes the weak local equilibrium of σ^{λ_h}. Moreover, the stress σ_h in (3.2) satisfies the equilibrium equation exactly, almost everywhere, in each element K. Indeed, for each K, from (3.2) and (2.10),

$$\mathbf{div}\,\sigma_h + \boldsymbol{f} = \mathbf{div}\big(\mathcal{A}\,\mathcal{E}(T\,\boldsymbol{\lambda}_h + \hat{T}\,\boldsymbol{f})\big) + \boldsymbol{f} = -\mathbf{R}_K^{\lambda_h} + \Pi_K(\boldsymbol{f}) = \mathbf{0},$$

since $\mathbf{R}_K^{\lambda_h} = \Pi_K(\boldsymbol{f})$ using (3.1) and (2.8).

It is instructive to consider $T\,\boldsymbol{\lambda}_h$ in more detail as it plays a central role in (3.1) or (3.3). Suppose $\{\boldsymbol{\psi}_i\}_{i=1}^{\dim \boldsymbol{\Lambda}_h}$ is a basis for $\boldsymbol{\Lambda}_h$. We define the set $\{\boldsymbol{\eta}_i\}_{i=1}^{\dim \boldsymbol{\Lambda}_h} \subset \widetilde{\mathbf{V}}$ with $\boldsymbol{\eta}_i := T\,\boldsymbol{\psi}_i$, and then $\boldsymbol{\eta}_i\,|_K$ satisfies

$$\big(\mathcal{A}\,\mathcal{E}(\boldsymbol{\eta}_i), \mathcal{E}(\tilde{\boldsymbol{w}})\big)_K = -\big(\boldsymbol{\psi}_i\,\boldsymbol{n}\cdot\boldsymbol{n}^K, \tilde{\boldsymbol{w}}\big)_{\partial K} \quad \text{for all } \tilde{\boldsymbol{w}} \in \widetilde{\mathbf{V}}, \tag{3.4}$$

or equivalently,

$$\begin{cases} -\mathbf{div}\mathcal{A}\,\mathcal{E}(\boldsymbol{\eta}_i) = \mathbf{R}_K^{\psi_i} & \text{in } K, \\ \ \ \mathcal{A}\,\mathcal{E}(\boldsymbol{\eta}_i)\,\boldsymbol{n}^K = -\boldsymbol{\psi}_i\,\boldsymbol{n}\cdot\boldsymbol{n}^K & \text{on } F \subset \partial K, \end{cases} \tag{3.5}$$

where $\mathbf{R}_K^{\psi_i} \in \mathbf{V}_{\rm rm}$ satisfies $(\mathbf{R}_K^{\psi_i}, \boldsymbol{v}^{\rm rm})_K = (\boldsymbol{\psi}_i, \boldsymbol{v}^{\rm rm})_{\partial K}$, for all $\boldsymbol{v}^{\rm rm} \in \mathbf{V}_{\rm rm}$. Taking $\boldsymbol{\lambda}_h = \sum_{i=1}^{\dim \boldsymbol{\Lambda}_h} \beta_i\,\boldsymbol{\psi}_i$ in $\boldsymbol{\Lambda}_h$, the linearity of (2.4) implies we can uniquely write

$$T\,\boldsymbol{\lambda}_h = \sum_{i=1}^{\dim \boldsymbol{\Lambda}_h} \beta_i\,T\boldsymbol{\psi}_i = \sum_{i=1}^{\dim \boldsymbol{\Lambda}_h} \beta_i\,\boldsymbol{\eta}_i.$$

It then follows that

$$\boldsymbol{u}_h = \boldsymbol{u}_h^{\rm rm} + \sum_{i=1}^{\dim \boldsymbol{\Lambda}_h} \beta_i\,\boldsymbol{\eta}_i + \hat{T}\,\boldsymbol{f}. \tag{3.6}$$

In this sense, the method can be seen as a nonconforming method to find an approximation of $\boldsymbol{u}$ in a finite dimensional subspace of $\mathbf{V} \nsubseteq \boldsymbol{H}^1(\Omega)$. On the other hand, the stress tensor σ is approximated by σ_h given by:

$$\sigma_h = \sum_{i=1}^{\dim \boldsymbol{\Lambda}_h} \beta_i\,\mathcal{A}\,\mathcal{E}(\boldsymbol{\eta}_i) + \mathcal{A}\,\mathcal{E}(\hat{T}\,\boldsymbol{f}) \in H(\mathrm{div};\Omega), \tag{3.7}$$

and, then, the method is *conforming* with respect to the variable σ. Also, the post-processed stress σ_h is *strongly* symmetric.

Note that heterogeneous and/or high-contrast aspects of the media automatically impact the design of the basis functions as they are driven by the local problems (3.4) (equivalently (3.5)) and the choice of $\boldsymbol{\Lambda}_h$. Also, embedded interfaces are naturally taken care of by these local problems, which easily accommodate edge-crossing interfaces thanks to the local boundary conditions. Furthermore, the strategy allows the present methodology to address multiscale aspects of the solution should they still persist inside of each local problem (2.4)–(2.5) for $T\,\boldsymbol{\lambda}$ and $\hat{T}\,\boldsymbol{f}$. Indeed, the current framework may be used recursively on the elliptic local problem, thereby incorporating multiple scales into problem (2.6).

For practical purposes, closed formulas are not available in general for $T\boldsymbol{\psi}_i$ and $\hat{T}\boldsymbol{f}$. This prevents (3.1) or (3.3) to be solved exactly, though some cases exist for which exact solutions are known. For instance, observe that $\hat{T}\,\boldsymbol{f} = 0$ if $\boldsymbol{f} \in \mathbf{V}_{\rm rm}$. Thereby, we propose a two-level methodology such that the functions $T\,\boldsymbol{\lambda}_h$ (*e.g.* $T\,\boldsymbol{\psi}_i$) and $\hat{T}\,\boldsymbol{f}$ taking part in (3.1) (or (3.3)) are replaced by their locally approximated counterparts $T_h\,\boldsymbol{\lambda}_h$ (*e.g.* $T_h\,\boldsymbol{\psi}_i$) and $\hat{T}_h\,\boldsymbol{f}$.

It is important to note that in either case, method (3.1) (or (3.3)) consists of the *same* number of degrees of freedom, with the local approximation appearing as a pre-processing step which is easily parallelized. The two-level computations may be based on primal Galerkin methods, their mixed counterpart obtained from the recursive procedure mentioned in the previous paragraphs, or even adopting a classical mixed method [14], which preserves local $H(\mathrm{div};K)$ conformity. In a broad sense, the two-level MHM method starts by selecting T_h and $\hat{T}_h$, and looking for $(\boldsymbol{\lambda}_h, \boldsymbol{u}_h^{\rm rm})$ in $\boldsymbol{\Lambda}_h \times \mathbf{V}_{\rm rm}$ as the solution of the following problem

$$\begin{cases} (\boldsymbol{\lambda}_h, \boldsymbol{v}^{\rm rm})_{\partial\mathcal{T}_h} = (\boldsymbol{f}, \boldsymbol{v}^{\rm rm})_{\mathcal{T}_h} & \text{for all } \boldsymbol{v}^{\rm rm} \in \mathbf{V}_{\rm rm}, \\ (\boldsymbol{\mu}_h, \boldsymbol{u}^{\rm rm} + T_h\,\boldsymbol{\lambda})_{\partial\mathcal{T}_h} = -(\boldsymbol{\mu}_h, \hat{T}_h\,\boldsymbol{f})_{\partial\mathcal{T}_h} + (\boldsymbol{\mu}_h, \boldsymbol{g})_{\partial\Omega} & \text{for all } \boldsymbol{\mu} \in \boldsymbol{\Lambda}_h. \end{cases} \tag{3.8}$$

It is worth mentioning that the choice of method at local level will impact the qualitative features of the global method. In Section 6, we detail and analyze a two-level strategy which makes the MHM method effective from a practical viewpoint. Particularly, we propose a compatibility condition between finite dimensional subspaces of $\tilde{\mathbf{V}}$ and $\boldsymbol{\Lambda}_h$ such that the two-level approach (3.8) preserves the key features of method (3.1) (or (3.3)).

Remark 3.2. We can view the MHM method (3.1) as penalizing jump terms. To see this, we first observe that given $\boldsymbol{\mu}_h \in \boldsymbol{\Lambda}_h$ with property $\boldsymbol{\mu}_h|_F \in \boldsymbol{H}^{-1/2}(F)$ (this space having its usual meaning, see [34] for instance), the decomposition $(\boldsymbol{\mu}_h, \boldsymbol{u})_{\partial K} = \sum_{F\subset\partial K}(\boldsymbol{\mu}_h, \boldsymbol{u})_F$, has meaning for all $\boldsymbol{u} \in \boldsymbol{H}^1(K)$. From this, the following identity holds

$$(\boldsymbol{\mu}_h, \boldsymbol{u})_{\partial\mathcal{T}_h} = \sum_{F\in\mathcal{E}_h} (\boldsymbol{\mu}_h \otimes \boldsymbol{n}, [\![\boldsymbol{u}]\!])_F, \tag{3.9}$$

where

$$[\![\boldsymbol{v}]\!]\,|_F := \boldsymbol{v}^{K_1}|_F \otimes \boldsymbol{n}_F^{K_1} + \boldsymbol{v}^{K_2}|_F \otimes \boldsymbol{n}_F^{K_2}, \tag{3.10}$$

and $(\boldsymbol{v}\otimes\boldsymbol{n})_{ij} = \boldsymbol{v}_i\boldsymbol{n}_j$, and $\boldsymbol{v}^{K_i}$ indicate the restriction of function $\boldsymbol{v}$ to either of the two elements sharing $F \in \mathcal{E}_h$. Also, upon selecting K_1 such that $\boldsymbol{n}_F^{K_1} = \boldsymbol{n}_F$, we set

$$\boldsymbol{\mu}_h\,|_F = \boldsymbol{\mu}_h^{K_1}\,|_F \tag{3.11}$$

Using identity (3.9) in (3.1), we have that $(\boldsymbol{u}_h^{\mathrm{rm}}, \boldsymbol{\lambda}_h) \in \mathbf{V}_{\mathrm{rm}} \times \boldsymbol{\Lambda}_h$ satisfies

$$\begin{cases} \displaystyle\sum_{F\in\mathcal{E}_h} (\boldsymbol{\lambda}_h \otimes \boldsymbol{n}, [\![\boldsymbol{v}^{\mathrm{rm}}]\!])_F = (\boldsymbol{f}, \boldsymbol{v}^{\mathrm{rm}})_{\mathcal{T}_h} \quad \text{for all } \boldsymbol{v}^{\mathrm{rm}} \in \mathbf{V}_{\mathrm{rm}}, \\ \displaystyle\sum_{F\in\mathcal{E}_h} (\boldsymbol{\mu}_h \otimes \boldsymbol{n}, [\![\boldsymbol{u}_h^{\mathrm{rm}} + T\,\boldsymbol{\lambda}_h]\!])_F = -\sum_{F\in\mathcal{E}_h} (\boldsymbol{\mu}_h \otimes \boldsymbol{n}, [\![\hat{T}\,\boldsymbol{f}]\!])_F + (\boldsymbol{\mu}_h,\, \boldsymbol{g})_{\partial\Omega} \quad \text{for all } \boldsymbol{\mu}_h \in \boldsymbol{\Lambda}_h, \end{cases}$$

which emphasizes that the MHM method works as a penalization on jumps.

Summarizing the algorithm for computing an approximation to (1.3), we get:

(i) compute $\hat{T}\boldsymbol{f}$ from (2.5) and the basis $\{\boldsymbol{\eta}_i\}_{i=1}^{\dim \boldsymbol{\Lambda}_h}$ from (3.4) as a local, completely parallelizable preprocessing step;
(ii) compute the degrees of freedom of $(\boldsymbol{u}_h^{\mathrm{rm}}, \boldsymbol{\lambda}_h)$ from (3.1) (or (3.3)), noting that $T\,\boldsymbol{\lambda}_h$ expands in terms of $\{\boldsymbol{\eta}_i\}_{i=1}^{\dim \boldsymbol{\Lambda}_h}$ using the degrees of freedom for $\boldsymbol{\lambda}_h$;
(iii) bring the results together to build the approximated stress σ_h from (3.7) and the approximated displacement $\boldsymbol{u}_h$ from (3.6).

Recall that for the two-level method (3.8), it is enough to replace T and $\hat{T}$ by T_h and $\hat{T}_h$ above, respectively.

4. Well-posedness and best approximation

We preserve conformity by selecting $\boldsymbol{\Lambda}_h$ such that

$$\boldsymbol{\Lambda}_{\mathrm{rm}} \subseteq \boldsymbol{\Lambda}_h \subset \boldsymbol{\Lambda}. \tag{4.1}$$

The well-posedness of method (3.1) and its best approximation result rely on the space $\boldsymbol{\Lambda}_{\mathrm{rm}}$. Indeed, a central result of this section shows (see Thm. 4.2) that $\boldsymbol{\Lambda}_{\mathrm{rm}}$ set as

$$\boldsymbol{\Lambda}_{\mathrm{rm}} := \{\boldsymbol{\mu} \in \boldsymbol{\Lambda} \,:\, \boldsymbol{\mu}|_{\partial K} \in \boldsymbol{\Lambda}_{\mathrm{rm}}(K),\, K \in \mathcal{T}_h\}, \tag{4.2}$$

where

$$\boldsymbol{\Lambda}_{\mathrm{rm}}(K) := \{\boldsymbol{v}^{\mathrm{rm}}|_F \,:\, \boldsymbol{v}^{\mathrm{rm}} \in \mathbf{V}_{\mathrm{rm}}(K),\, F \subset \partial K\},$$

ensures that an inf-sup condition between the spaces $\boldsymbol{\Lambda}_{\rm rm}$ and $\mathbf{V}_{\rm rm}$ holds. Henceforth, C represents a positive constant independent of h which can differ in each occurrence and have a possible dependence on $\mathcal{A}$.

Let us define the bilinear forms $a : \boldsymbol{\Lambda} \times \boldsymbol{\Lambda} \to \mathbb{R}$ and $b : \boldsymbol{\Lambda} \times \mathbf{V} \to \mathbb{R}$

$$a(\boldsymbol{\lambda}, \boldsymbol{\mu}) := (\boldsymbol{\mu}, T\,\boldsymbol{\lambda})_{\partial\mathcal{T}_h}, \quad b(\boldsymbol{\mu}, \boldsymbol{v}) := (\boldsymbol{\mu}, \boldsymbol{v})_{\partial\mathcal{T}_h},$$

where we recall that the linear operator T is defined in (2.4). It is convenient to rewrite problem (2.6) by adding both equations and proposing the formulation in the following way: *find* $(\boldsymbol{\lambda}, \boldsymbol{u}^{\rm rm}) \in \boldsymbol{\Lambda} \times \mathbf{V}_{\rm rm}$ *such that*

$$\mathbf{B}(\boldsymbol{\lambda}, \boldsymbol{u}^{\rm rm}; \boldsymbol{\mu}, \boldsymbol{v}^{\rm rm}) = \mathbf{F}(\boldsymbol{\mu}, \boldsymbol{v}^{\rm rm}) \quad \text{for all } (\boldsymbol{\mu}, \boldsymbol{v}^{\rm rm}) \in \boldsymbol{\Lambda} \times \mathbf{V}_{\rm rm}, \tag{4.3}$$

where

$$\begin{aligned} \mathbf{B}(\boldsymbol{\lambda}, \boldsymbol{u}^{\rm rm}; \boldsymbol{\mu}, \boldsymbol{v}^{\rm rm}) &:= a(\boldsymbol{\lambda}, \boldsymbol{\mu}) + b(\boldsymbol{\mu}, \boldsymbol{u}^{\rm rm}) + b(\boldsymbol{\lambda}, \boldsymbol{v}^{\rm rm}), \\ \mathbf{F}(\boldsymbol{\mu}, \boldsymbol{v}^{\rm rm}) &:= (\boldsymbol{f}, \boldsymbol{v}^{\rm rm})_{\mathcal{T}_h} - (\boldsymbol{\mu}, \hat{T}\,\boldsymbol{f})_{\partial\mathcal{T}_h} + (\boldsymbol{\mu}, \boldsymbol{g})_{\partial\Omega}. \end{aligned}$$

Note that $\mathbf{B}(.,.)$ is symmetric due to (2.4). Similarly, we rewrite the MHM method (3.1) as: *find* $(\boldsymbol{\lambda}_h, \boldsymbol{u}_h^{\rm rm}) \in \boldsymbol{\Lambda}_h \times \mathbf{V}_{\rm rm}$ *such that*

$$\mathbf{B}(\boldsymbol{\lambda}_h, \boldsymbol{u}_h^{\rm rm}; \boldsymbol{\mu}_h, \boldsymbol{v}^{\rm rm}) = \mathbf{F}(\boldsymbol{\mu}_h, \boldsymbol{v}^{\rm rm}) \quad \text{for all } (\boldsymbol{\mu}_h, \boldsymbol{v}^{\rm rm}) \in \boldsymbol{\Lambda}_h \times \mathbf{V}_{\rm rm}. \tag{4.4}$$

Our analysis requires the norm on $\boldsymbol{\Lambda} \times \mathbf{V}_{\rm rm}$

$$\|(\boldsymbol{\mu}, \boldsymbol{v}^{\rm rm})\|_{\boldsymbol{\Lambda}\times\mathbf{V}} := \|\boldsymbol{\mu}\|_{\boldsymbol{\Lambda}} + \|\boldsymbol{v}^{\rm rm}\|_{\mathbf{V}}, \tag{4.5}$$

where each contribution reads (here d_Ω stands for the diameter of Ω)

$$\begin{aligned} \|\boldsymbol{\mu}\|_{\boldsymbol{\Lambda}} &:= \inf_{\substack{\sigma\in H(\mathrm{div};\Omega) \\ \sigma\,\boldsymbol{n}^K=\boldsymbol{\mu} \text{ on } \partial K,\, K\in\mathcal{T}_h}} \|\sigma\|_{\rm div}, \qquad \|\sigma\|^2_{\rm div} := \sum_{K\in\mathcal{T}_h} \Big(\|\sigma\|^2_{0,K} + d_\Omega^2 \|\mathbf{div}\,\sigma\|^2_{0,K} \Big), \\ \|\boldsymbol{v}\|^2_{\mathbf{V}} &:= \sum_{K\in\mathcal{T}_h} \Big(d_\Omega^{-2} \|\boldsymbol{v}\|^2_{0,K} + \|\mathcal{E}(\boldsymbol{v})\|^2_{0,K} \Big). \end{aligned} \tag{4.6}$$

Denote by Π the global L^2 projection onto $\mathbf{V}_{\rm rm}$ such that $\Pi\,|_K = \Pi_K$. From standard stability result of projections and (4.6), it holds for all $\boldsymbol{v} \in \mathbf{V}$

$$\|\Pi\,\boldsymbol{v}\|_{\mathbf{V}} \le \|\boldsymbol{v}\|_{\mathbf{V}}. \tag{4.7}$$

Also, from trace inequality (A.3), the property $\Pi\,\boldsymbol{v}^{\rm rm} = \boldsymbol{v}^{\rm rm}$ and Korn inequality (A.1), we get

$$\|\boldsymbol{v} - \Pi_K\,\boldsymbol{v}\|_{0,\partial K} \le C\,h_K^{1/2} \|\,\mathcal{E}(\boldsymbol{v})\|_{0,K} \quad \text{for all } \boldsymbol{v} \in \mathbf{V}. \tag{4.8}$$

Before heading to the analysis of the MHM method, we must first characterize the space $\boldsymbol{\Lambda}_{\rm rm}$ defined in (4.2) through the action of an interpolation operator. To be precise, let us define the local interpolation $\mathcal{I}_K$ on functions in $\boldsymbol{L}^2(\partial K)$ with value in $\boldsymbol{\Lambda}_{\rm rm}(K)$, such that for each $F \subset \partial K$, it holds

$$\int_F \mathcal{I}_K \boldsymbol{\mu}\,\boldsymbol{v}^{\rm rm} \mathrm{d}s = \int_F \boldsymbol{\mu}\,\boldsymbol{v}^{\rm rm} \mathrm{d}s \quad \text{for all } \boldsymbol{v}^{\rm rm} \in \mathbf{V}_{\rm rm}(K). \tag{4.9}$$

Observe that conditions (4.9) imply unisolvence in $\boldsymbol{\Lambda}_{\rm rm}(K)$ and the following local stability result holds

$$\|\mathcal{I}_K \boldsymbol{\mu}\|_{0,\partial K} \le \|\boldsymbol{\mu}\|_{0,\partial K}. \tag{4.10}$$

The global interpolation $\mathcal{I}$ acts on the trace of functions in $[H^1(\Omega)]^{n\times n}$ (with its usual meaning) with values in $\boldsymbol{\Lambda}_{\rm rm}$, and is fully defined assuming $\mathcal{I}|_K = \mathcal{I}_K$. To investigate the stability of $\mathcal{I}$ in the $\|.\|_{\boldsymbol{\Lambda}}$ norm (4.6), we first observe that from (4.10) and (4.8) we get

$$\begin{aligned} b(\mathcal{I}\boldsymbol{\mu}, \boldsymbol{v} - \Pi\,\boldsymbol{v}) &\leq \sum_{K\in\mathcal{T}_h} \|\mathcal{I}_K\boldsymbol{\mu}\|_{0,\partial K}\|\boldsymbol{v} - \Pi_K\,\boldsymbol{v}\|_{0,\partial K} \\ &\leq C \sum_{K\in\mathcal{T}_h} \|\boldsymbol{\mu}\|_{0,\partial K}\, h_K^{1/2}\|\,\mathcal{E}(\boldsymbol{v})\|_{0,K} \\ &\leq C\Big[\sum_{K\in\mathcal{T}_h}\|\boldsymbol{\mu}\|^2_{0,\partial K}\,h_K\Big]^{1/2}\|\boldsymbol{v}\|_{\mathbf{V}}. \end{aligned} \tag{4.11}$$

In what follows, we make consistent use of the following equivalence of norms (see Lem. A.3 in the Appendix)

$$\frac{\sqrt{2}}{2\,C_{\rm korn}}\|\boldsymbol{\mu}\|_{\boldsymbol{\Lambda}} \leq \sup_{\boldsymbol{v}\in\mathbf{V}} \frac{b(\boldsymbol{\mu},\boldsymbol{v})}{\|\boldsymbol{v}\|_{\mathbf{V}}} \leq \|\boldsymbol{\mu}\|_{\boldsymbol{\Lambda}}, \tag{4.12}$$

where $C_{\rm korn}$ is a positive constant independent of h. Above and hereafter we lighten the notation and understand the supremum to be taken over sets excluding the zero function.

Next, from (4.12), (4.9), the definition of the norm $\|.\|_{\boldsymbol{\Lambda}}$ in (4.6), and (4.11) the operator $\mathcal{I}$ is stable as follows

$$\begin{aligned} \|\mathcal{I}\boldsymbol{\mu}\|_{\boldsymbol{\Lambda}} \leq C \sup_{\boldsymbol{v}\in\mathbf{V}} \frac{b(\mathcal{I}\boldsymbol{\mu},\boldsymbol{v})}{\|\boldsymbol{v}\|_{\mathbf{V}}} &\leq C\left(\sup_{\boldsymbol{v}\in\mathbf{V}} \frac{b(\mathcal{I}\boldsymbol{\mu},\Pi\,\boldsymbol{v})}{\|\boldsymbol{v}\|_{\mathbf{V}}} + \sup_{\boldsymbol{v}\in\mathbf{V}} \frac{b(\mathcal{I}\boldsymbol{\mu},\boldsymbol{v}-\Pi\,\boldsymbol{v})}{\|\boldsymbol{v}\|_{\mathbf{V}}}\right) \\ &\leq C\left(\sup_{\boldsymbol{v}\in\mathbf{V}} \frac{b(\boldsymbol{\mu},\Pi\,\boldsymbol{v})}{\|\Pi\,\boldsymbol{v}\|_{\mathbf{V}}} + \left[\sum_{K\in\mathcal{T}_h}\|\boldsymbol{\mu}\|^2_{0,\partial K}\,h_K\right]^{1/2}\right) \\ &\leq C\left(\sup_{\boldsymbol{v}\in\mathbf{V}} \frac{b(\boldsymbol{\mu},\boldsymbol{v})}{\|\boldsymbol{v}\|_{\mathbf{V}}} + \Big[\sum_{K\in\mathcal{T}_h}\|\boldsymbol{\mu}\|^2_{0,\partial K}\,h_K\Big]^{1/2}\right) \\ &\leq C\left(\|\boldsymbol{\mu}\|_{\boldsymbol{\Lambda}} + \left[\sum_{K\in\mathcal{T}_h}\|\boldsymbol{\mu}\|^2_{0,\partial K}\,h_K\right]^{1/2}\right). \end{aligned} \tag{4.13}$$

We are ready to prove that the operator associated with the linear form $b(.,.)$ is surjective. Hereafter the domain Ω is assumed to be a simply connected polygon.

Lemma 4.1. *Given $\boldsymbol{v}^{\rm rm} \in \mathbf{V}_{\rm rm}$, there exists C such that*

$$C\,\|\boldsymbol{v}^{\rm rm}\|_{\mathbf{V}} \leq \sup_{\boldsymbol{\mu}_{\rm rm}\in\boldsymbol{\Lambda}_{\rm rm}} \frac{b(\boldsymbol{\mu}_{\rm rm},\boldsymbol{v}^{\rm rm})}{\|\boldsymbol{\mu}_{\rm rm}\|_{\boldsymbol{\Lambda}}}.$$

Proof. Assume $\boldsymbol{v}^{\rm rm} \in \mathbf{V}_{\rm rm}$. From the assumption on Ω, observe that there exists a symmetric matrix function $\sigma^\star \in [H^1(\Omega)]^{n\times n}$ (see [6] for instance) such that

$$\mathbf{div}\,\sigma^\star = \boldsymbol{v}^{\rm rm} \quad \text{and} \quad \|\sigma^\star\|_{1,\Omega} \leq C\,\|\boldsymbol{v}^{\rm rm}\|_{0,\Omega}. \tag{4.14}$$

Take $\boldsymbol{\mu}^\star \in \boldsymbol{\Lambda}$ such that $\boldsymbol{\mu}^\star|_{\partial K} := \sigma^\star\,\boldsymbol{n}^K|_{\partial K}$, $K\in\mathcal{T}_h$. From (4.13), and the definition of norm $\|.\|_{\boldsymbol{\Lambda}}$ in (4.6) and (4.14), and observing that from a scaling argument (*cf.* [15], p. 111) we get

$$\left[\sum_{K\in\mathcal{T}_h}\|\boldsymbol{\mu}^\star\|^2_{0,\partial K}\,h_K\right]^{1/2} \leq C\,\|\boldsymbol{\mu}^\star\|_{\boldsymbol{\Lambda}},$$

we arrive at the following result

$$\|\mathcal{I}\boldsymbol{\mu}^\star\|_{\boldsymbol{\Lambda}} \leq C\,\|\boldsymbol{\mu}^\star\|_{\boldsymbol{\Lambda}} \leq C\,\|\sigma^\star\|_{\mathrm{div}} \leq C\,\|\boldsymbol{v}^{\mathrm{rm}}\|_{\mathbf{V}}. \tag{4.15}$$

Next, from (4.14)-(4.15) and (4.9) it holds

$$d_\Omega\,\|\boldsymbol{v}^{\mathrm{rm}}\|_{\mathbf{V}} = \frac{(\mathbf{div}\,\sigma^\star, \boldsymbol{v}^{\mathrm{rm}})_{\mathcal{T}_h}}{\|\boldsymbol{v}^{\mathrm{rm}}\|_{\mathbf{V}}} = \frac{(\boldsymbol{\mu}^\star, \boldsymbol{v}^{\mathrm{rm}})_{\partial\mathcal{T}_h}}{\|\boldsymbol{v}^{\mathrm{rm}}\|_{\mathbf{V}}} = \frac{(\mathcal{I}\boldsymbol{\mu}^\star, \boldsymbol{v}^{\mathrm{rm}})_{\partial\mathcal{T}_h}}{\|\boldsymbol{v}^{\mathrm{rm}}\|_{\mathbf{V}}} \leq C\,\frac{(\mathcal{I}\boldsymbol{\mu}^\star, \boldsymbol{v}^{\mathrm{rm}})_{\partial\mathcal{T}_h}}{\|\mathcal{I}\boldsymbol{\mu}^\star\|_{\boldsymbol{\Lambda}}},$$

and the result follows taking the supremum. □

Hereafter, we will make use of the following tensor norm on $\mathcal{A}$

$$\|\mathcal{A}\| := \operatorname*{ess\,sup}_{\boldsymbol{x}\in\Omega}\max_{|\xi|=1}\,(\mathcal{A}(\boldsymbol{x})\,\xi,\xi)^{1/2}, \quad \xi\in\mathbb{R}^{n\times n}_{\mathrm{sym}}, \tag{4.16}$$

which from (1.2) satisfies $c_{\min} \leq \|\mathcal{A}\| \leq c_{\max}$. We are ready to prove the well-posedness result. Observe that the proof is unchanged for any finite dimensional space $\boldsymbol{\Lambda}_h \subset \boldsymbol{\Lambda}$ under the condition (4.1). Therefore, the following result applies to the MHM formulation (4.4) as well.

Theorem 4.2. *Suppose* $(\boldsymbol{\lambda}, \boldsymbol{u}^{\mathrm{rm}})$, $(\boldsymbol{\mu}, \boldsymbol{v}^{\mathrm{rm}}) \in \boldsymbol{\Lambda}\times\mathbf{V}_{\mathrm{rm}}$. *Then, there exists* C *such that*

$$\mathbf{B}(\boldsymbol{\lambda}, \boldsymbol{u}^{\mathrm{rm}}; \boldsymbol{\mu}, \boldsymbol{v}^{\mathrm{rm}}) \;\leq\; C\,\|(\boldsymbol{\lambda}, \boldsymbol{u}^{\mathrm{rm}})\|_{\boldsymbol{\Lambda}\times\mathbf{V}}\,\|(\boldsymbol{\mu}, \boldsymbol{v}^{\mathrm{rm}})\|_{\boldsymbol{\Lambda}\times\mathbf{V}}. \tag{4.17}$$

Moreover, there exists a positive constant β*, independent of* h*, such that*

$$\sup_{(\boldsymbol{\mu},\boldsymbol{v}^{\mathrm{rm}})\in\boldsymbol{\Lambda}\times\mathbf{V}_{\mathrm{rm}}} \frac{\mathbf{B}(\boldsymbol{\lambda}, \boldsymbol{u}^{\mathrm{rm}}; \boldsymbol{\mu}, \boldsymbol{v}^{\mathrm{rm}})}{\|(\boldsymbol{\mu}, \boldsymbol{v}^{\mathrm{rm}})\|_{\boldsymbol{\Lambda}\times\mathbf{V}}} \;\geq\; \beta\,\|(\boldsymbol{\lambda}, \boldsymbol{u}^{\mathrm{rm}})\|_{\boldsymbol{\Lambda}\times\mathbf{V}} \quad \text{for all } (\boldsymbol{\lambda}, \boldsymbol{u}^{\mathrm{rm}}) \in \boldsymbol{\Lambda}\times\mathbf{V}_{\mathrm{rm}}. \tag{4.18}$$

Also,

$$\mathbf{B}(\boldsymbol{\lambda}, \boldsymbol{u}^{\mathrm{rm}};\, \boldsymbol{\mu}, \boldsymbol{v}^{\mathrm{rm}}) = 0 \quad \text{for all } \;(\boldsymbol{\lambda}, \boldsymbol{u}^{\mathrm{rm}}) \in \boldsymbol{\Lambda}\times\mathbf{V}_{\mathrm{rm}} \quad\Longrightarrow\quad (\boldsymbol{\mu}, \boldsymbol{v}^{\mathrm{rm}}) = (\mathbf{0}, \mathbf{0}), \tag{4.19}$$

for all $(\boldsymbol{\mu}, \boldsymbol{v}^{\mathrm{rm}}) \in \boldsymbol{\Lambda}\times\mathbf{V}_{\mathrm{rm}}$. *We conclude problem* (4.3) *is well-posed.*

Proof. The proof follows closely [3]. First, we prove (4.17). Since by definition $a(\boldsymbol{\lambda}, \boldsymbol{\mu}) = b(\boldsymbol{\mu}, T\boldsymbol{\lambda})$, it follows by the equivalence result (4.12) and Lemma A.1 in the Appendix, and definition of norm (4.5) that

$$\begin{aligned}
\mathbf{B}(\boldsymbol{\lambda}, \boldsymbol{u}^{\mathrm{rm}}; \boldsymbol{\mu}, \boldsymbol{v}^{\mathrm{rm}}) &= b(\boldsymbol{\mu}, T\,\boldsymbol{\lambda} + \boldsymbol{u}^{\mathrm{rm}}) + b(\boldsymbol{\lambda}, \boldsymbol{v}^{\mathrm{rm}}) \\
&\leq \sup_{\boldsymbol{w}\in\mathbf{V}} \frac{b(\boldsymbol{\mu}, \boldsymbol{w})}{\|\boldsymbol{w}\|_{\mathbf{V}}}\|T\,\boldsymbol{\lambda} + \boldsymbol{u}^{\mathrm{rm}}\|_{\mathbf{V}} + \sup_{\boldsymbol{w}\in\mathbf{V}} \frac{b(\boldsymbol{\lambda}, \boldsymbol{w})}{\|\boldsymbol{w}\|_{\mathbf{V}}}\|\boldsymbol{v}^{\mathrm{rm}}\|_{\mathbf{V}} \\
&\leq \|\boldsymbol{\mu}\|_{\boldsymbol{\Lambda}}(\|T\,\boldsymbol{\lambda}\|_{\mathbf{V}} + \|\boldsymbol{u}^{\mathrm{rm}}\|_{\mathbf{V}}) + \|\boldsymbol{\lambda}\|_{\boldsymbol{\Lambda}}\|\boldsymbol{v}^{\mathrm{rm}}\|_{\mathbf{V}} \\
&\leq \frac{(C_{\mathrm{korn}}^2 + \pi^2)c_{\max}}{\pi^2 c_{\min}^2}\|\boldsymbol{\mu}\|_{\boldsymbol{\Lambda}}\|\boldsymbol{\lambda}\|_{\boldsymbol{\Lambda}} + \|\boldsymbol{\mu}\|_{\boldsymbol{\Lambda}}\|\boldsymbol{u}^{\mathrm{rm}}\|_{\mathbf{V}} + \|\boldsymbol{\lambda}\|_{\boldsymbol{\Lambda}}\|\boldsymbol{v}^{\mathrm{rm}}\|_{\mathbf{V}},
\end{aligned}$$

and (4.17) follows immediately.

Next, we prove a coercivity condition for $-a(.,.)$, for which we require the following nullspace

$$\mathcal{N} := \{\boldsymbol{\mu}\in\boldsymbol{\Lambda} \,:\, b(\boldsymbol{\mu}, \boldsymbol{v}^{\mathrm{rm}}) = 0,\,\forall\,\boldsymbol{v}^{\mathrm{rm}}\in\mathbf{V}_{\mathrm{rm}}\}. \tag{4.20}$$

Assume $\boldsymbol{\mu} \in \mathcal{N}$ and first note that problems (2.4) and (2.11) imply that $\mathbf{div}\, \mathcal{A}\,\mathcal{E}(T\,\boldsymbol{\mu}) = \mathbf{0}$. Using (2.4) and (1.2), it holds

$$
\begin{aligned}
-a(\boldsymbol{\mu},\boldsymbol{\mu}) &= (\mathcal{A}^{-1}\mathcal{A}\,\mathcal{E}(T\,\boldsymbol{\mu}), \mathcal{A}\,\mathcal{E}(T\,\boldsymbol{\mu}))_{\mathcal{T}_h} \\
&\geq \frac{1}{c_{\max}} \|\mathcal{A}\,\mathcal{E}(T\,\boldsymbol{\mu})\|^2_{0,\Omega} \\
&= \frac{1}{c_{\max}} \|\mathcal{A}\,\mathcal{E}(T\,\boldsymbol{\mu})\|^2_{\mathrm{div}} \\
&\geq \frac{1}{c_{\max}} \|\boldsymbol{\mu}\|^2_{\boldsymbol{\Lambda}},
\end{aligned}
$$

since $\mathcal{A}\,\mathcal{E}(T\,\boldsymbol{\mu})\boldsymbol{n}^K = \boldsymbol{\mu}$ on ∂K for all $K \in \mathcal{T}_h$. We conclude $-a(.,.)$ is coercive on the subspace $\mathcal{N}$. This result, along with the inf-sup condition for $b(.,.)$ from Lemma 4.1, are the requirements of the abstract setting in [3] (see also [19], p. 101) from which (4.18) and (4.19) hold with a positive constant β independent of h. □

We close this section by showing the MHM method produces the best approximation, where the convergence of both $\boldsymbol{\lambda}_h$ and $\boldsymbol{u}_h^{\mathrm{rm}}$ is governed by the approximation properties of $\boldsymbol{\Lambda}_h$. Indeed, observe that the accuracy of $\boldsymbol{u}_h^{\mathrm{rm}}$ approaching $\boldsymbol{u}^{\mathrm{rm}}$ depends on how well $\boldsymbol{\Lambda}_h$ approximates $\boldsymbol{\Lambda}$. In consequence, optimal convergence for $\boldsymbol{u}_h$ and σ_h given in (3.2) in the natural norms is expected to rely only on the capacity of $\boldsymbol{\lambda}$ to be optimally interpolated by $\boldsymbol{\lambda}_h$ on faces. This is established in the next lemma.

Lemma 4.3. *Let $(\boldsymbol{\lambda}, \boldsymbol{u}^{\mathrm{rm}})$ and $(\boldsymbol{\lambda}_h, \boldsymbol{u}_h^{\mathrm{rm}})$ be the solutions of (4.3) and (4.4), respectively. Under the assumptions of Theorem 4.2, it holds that*

$$
\mathbf{B}(\boldsymbol{\lambda}-\boldsymbol{\lambda}_h, \boldsymbol{u}^{\mathrm{rm}}-\boldsymbol{u}_h^{\mathrm{rm}}; \boldsymbol{\mu}_h, \boldsymbol{v}^{\mathrm{rm}}) = 0 \quad \textit{for all } \ (\boldsymbol{\mu}_h, \boldsymbol{v}^{\mathrm{rm}}) \in \boldsymbol{\Lambda}_h \times \mathbf{V}_{\mathrm{rm}}. \tag{4.21}
$$

Moreover, there exists C such that

$$
\|(\boldsymbol{\lambda}-\boldsymbol{\lambda}_h, \boldsymbol{u}^{\mathrm{rm}}-\boldsymbol{u}_h^{\mathrm{rm}})\|_{\boldsymbol{\Lambda}\times\mathbf{V}} \;\leq\; C \inf_{\boldsymbol{\mu}_h\in\boldsymbol{\Lambda}_h} \|\boldsymbol{\lambda}-\boldsymbol{\mu}_h\|_{\boldsymbol{\Lambda}}. \tag{4.22}
$$

Proof. The result (4.21) follows immediately from (4.3) and (4.4). Next, from standard arguments using Theorem 4.2 and (4.21), there is C such that

$$
\|(\boldsymbol{\lambda}-\boldsymbol{\lambda}_h, \boldsymbol{u}^{\mathrm{rm}}-\boldsymbol{u}_h^{\mathrm{rm}})\|_{\boldsymbol{\Lambda}\times\mathbf{V}} \leq C\,\|(\boldsymbol{\lambda}-\boldsymbol{\mu}_h, \boldsymbol{u}^{\mathrm{rm}}-\boldsymbol{v}^{\mathrm{rm}})\|_{\boldsymbol{\Lambda}\times\mathbf{V}},
$$

which is valid for all $(\boldsymbol{\mu}_h, \boldsymbol{v}^{\mathrm{rm}}) \in \boldsymbol{\Lambda}_h \times \mathbf{V}_{\mathrm{rm}}$. Thereby, selecting $(\boldsymbol{\mu}_h, \boldsymbol{v}^{\mathrm{rm}}) = (\boldsymbol{\mu}_h, \boldsymbol{u}^{\mathrm{rm}})$ in $\boldsymbol{\Lambda}_h \times \mathbf{V}_{\mathrm{rm}}$, the result follows by taking the infimum. □

Observe that the approximate solution fulfills the local equilibrium constraint exactly as shown in the next result. Hereafter, we shall consistently use the characterization of $\boldsymbol{u}$ and σ, and $\boldsymbol{u}_h$ and σ_h given in (2.7) and (3.2), respectively.

Corollary 4.4. *Let $(\boldsymbol{\lambda}, \boldsymbol{u}^{\mathrm{rm}})$ and $(\boldsymbol{\lambda}_h, \boldsymbol{u}_h^{\mathrm{rm}})$ be the solutions of (4.3) and (4.4), respectively. The following result holds*

$$
\mathbf{div}\, \sigma_h \;=\; \mathbf{div}\, \sigma \quad \textit{in } \Omega. \tag{4.23}
$$

Proof. See Remark 2. □

We shall make use of the assumption that problem (1.1) has smoothing properties in the sense of ([19], Def. 3.14).

Lemma 4.5. *Let* $(\boldsymbol{\lambda}, \boldsymbol{u}^{\rm rm})$ *and* $(\boldsymbol{\lambda}_h, \boldsymbol{u}_h^{\rm rm})$ *be the solutions of* (4.3) *and* (4.4)*, respectively. Under the assumptions of Theorem* 4.2*, it holds*

$$\|\sigma - \sigma_h\|_{\rm div} \leq C \inf_{\boldsymbol{\mu}_h \in \boldsymbol{\Lambda}_h} \|\boldsymbol{\lambda} - \boldsymbol{\mu}_h\|_{\boldsymbol{\Lambda}}, \tag{4.24}$$

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{0,\Omega} \leq C \inf_{\boldsymbol{\mu}_h \in \boldsymbol{\Lambda}_h} \|\boldsymbol{\lambda} - \boldsymbol{\mu}_h\|_{\boldsymbol{\Lambda}}. \tag{4.25}$$

Furthermore, if problem (1.1) *has smoothing properties, it holds*

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{0,\Omega} \leq C\, h \inf_{\boldsymbol{\mu}_h \in \boldsymbol{\Lambda}_h} \|\boldsymbol{\lambda} - \boldsymbol{\mu}_h\|_{\boldsymbol{\Lambda}}, \tag{4.26}$$

$$\|\boldsymbol{u}^{\rm rm} - \boldsymbol{u}_h^{\rm rm}\|_{0,\Omega} \leq C\, h \inf_{\boldsymbol{\mu}_h \in \boldsymbol{\Lambda}_h} \|\boldsymbol{\lambda} - \boldsymbol{\mu}_h\|_{\boldsymbol{\Lambda}}. \tag{4.27}$$

Proof. The proof closely follows the one presented in [3] for the Laplace's equation. We revisit and adapt it here for sake of clarity. First, Lemma A.1 implies $\|\mathcal{A}\, \mathcal{E}(\boldsymbol{u} - \boldsymbol{u}_h)\|_{\rm div} \leq C\, \|\boldsymbol{\lambda} - \boldsymbol{\lambda}_h\|_{\boldsymbol{\Lambda}}$ so that result (4.24) follows from (4.22) in Lemma 4.3. Again using Lemma A.1, (2.7) and (3.2), and triangle inequality, we get

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{0,\Omega} \leq \|\boldsymbol{u}^{\rm rm} - \boldsymbol{u}_h^{\rm rm}\|_{0,\Omega} + C\, \|\boldsymbol{\lambda} - \boldsymbol{\lambda}_h\|_{\boldsymbol{\Lambda}},$$

and estimate (4.25) results from Lemma 4.3.

To prove result (4.26), we employ a duality argument. Define

$$\boldsymbol{e} := \boldsymbol{u} - \boldsymbol{u}_h = \boldsymbol{u}^{\rm rm} - \boldsymbol{u}_h^{\rm rm} + T(\boldsymbol{\lambda} - \boldsymbol{\lambda}_h)$$

and suppose that $(\boldsymbol{\gamma}, \boldsymbol{w}^{\rm rm}) \in \boldsymbol{\Lambda} \times \mathbf{V}_{\rm rm}$ satisfies

$$\mathbf{B}(\boldsymbol{\mu}, \boldsymbol{v}^{\rm rm}; \boldsymbol{\gamma}, \boldsymbol{w}^{\rm rm}) = (T\, \boldsymbol{\mu} + \boldsymbol{v}^{\rm rm}, \boldsymbol{e})_{\mathcal{T}_h} \quad \text{for all } (\boldsymbol{\mu}, \boldsymbol{v}^{\rm rm}) \in \boldsymbol{\Lambda} \times \mathbf{V}_{\rm rm}. \tag{4.28}$$

The problem of finding such a $(\boldsymbol{\gamma}, \boldsymbol{w}^{\rm rm})$ is the adjoint problem of(4.3) with homogeneous Dirichlet boundary condition prescribed on $\partial\Omega$, and the right-hand side rewritten using (2.4) and (2.5). Furthermore, define $(\boldsymbol{\gamma}_{\rm rm}, \boldsymbol{w}_h^{\rm rm}) \in \boldsymbol{\Lambda}_{\rm rm} \times \mathbf{V}_{\rm rm}$ by the finite-dimensional adjoint problem

$$\mathbf{B}(\boldsymbol{\mu}_{\rm rm}, \boldsymbol{v}^{\rm rm}; \boldsymbol{\gamma}_{\rm rm}, \boldsymbol{w}_h^{\rm rm}) = (T\, \boldsymbol{\mu}_{\rm rm} + \boldsymbol{v}^{\rm rm}, \boldsymbol{e})_{\mathcal{T}_h}, \quad \text{for all } (\boldsymbol{\mu}_{\rm rm}, \boldsymbol{v}^{\rm rm}) \in \boldsymbol{\Lambda}_{\rm rm} \times \mathbf{V}_{\rm rm}. \tag{4.29}$$

Both (4.28) and (4.29) have unique solutions by Theorem 4.2 and the symmetry of the problem statements. Under the assumption that problem (1.1) has smoothing properties, we observe that the solution $\boldsymbol{w} := \boldsymbol{w}^{\rm rm} + T\,\boldsymbol{\gamma} + \hat{T}\boldsymbol{e}$ has extra regularity since $\boldsymbol{e} \in \boldsymbol{L}^2(\Omega)$, and there is a positive constant C (depending only on Ω) such that $\|\boldsymbol{w}\|_{2,\Omega} \leq \frac{C}{c_{\min}}\, \|\boldsymbol{e}\|_{0,\Omega}$. From this, Lemma 4.3, and the interpolation estimate (5.3) we find

$$\inf_{\boldsymbol{\mu}_{\rm rm} \in \boldsymbol{\Lambda}_{\rm rm}} \|\boldsymbol{\gamma} - \boldsymbol{\mu}_{\rm rm}\|_{\boldsymbol{\Lambda}} \leq C\, h\, \|\boldsymbol{w}\|_{2,\Omega},$$

and we then use (4.22) to show

$$\|(\boldsymbol{\gamma} - \boldsymbol{\gamma}_{\rm rm}, \boldsymbol{w}^{\rm rm} - \boldsymbol{w}_h^{\rm rm})\|_{\boldsymbol{\Lambda} \times \mathbf{V}} \leq C\, h \|\boldsymbol{w}\|_{2,\Omega} \leq \frac{C}{c_{\min}}\, h\, \|\boldsymbol{e}\|_{0,\Omega}.$$

Therefore, by (4.28), the consistency result of Lemma 4.3, the continuity result of Theorem 4.2, and the best approximation result of Lemma 4.3, we find

$$\begin{aligned}
\|\boldsymbol{e}\|_{0,\Omega}^2 &= (\boldsymbol{e}, \boldsymbol{e})_{\mathcal{T}_h} \\
&= (T\,(\boldsymbol{\lambda} - \boldsymbol{\lambda}_h) + (\boldsymbol{u}^{\rm rm} - \boldsymbol{u}_h^{\rm rm}), \boldsymbol{e})_{\mathcal{T}_h} \\
&= \mathbf{B}(\boldsymbol{\lambda} - \boldsymbol{\lambda}_h, \boldsymbol{u}^{\rm rm} - \boldsymbol{u}_h^{\rm rm}; \boldsymbol{\gamma}, \boldsymbol{w}^{\rm rm}) \\
&= \mathbf{B}(\boldsymbol{\lambda} - \boldsymbol{\lambda}_h, \boldsymbol{u}^{\rm rm} - \boldsymbol{u}_h^{\rm rm}; \boldsymbol{\gamma} - \boldsymbol{\gamma}_{\rm rm}, \boldsymbol{w}^{\rm rm} - \boldsymbol{w}_h^{\rm rm}) \\
&\leq C\, \|(\boldsymbol{\lambda} - \boldsymbol{\lambda}_h, \boldsymbol{u}^{\rm rm} - \boldsymbol{u}_h^{\rm rm})\|_{\boldsymbol{\Lambda} \times \mathbf{V}} \|(\boldsymbol{\gamma} - \boldsymbol{\gamma}_{\rm rm}, \boldsymbol{w}^{\rm rm} - \boldsymbol{w}_h^{\rm rm})\|_{\boldsymbol{\Lambda} \times \mathbf{V}} \\
&\leq C\, h \inf_{\boldsymbol{\mu}_h \in \boldsymbol{\Lambda}_h} \|\boldsymbol{\lambda} - \boldsymbol{\mu}_h\|_{\boldsymbol{\Lambda}} \|\boldsymbol{e}\|_{0,\Omega},
\end{aligned}$$

which establishes (4.26). As for (4.27), using the triangle inequality, the local inequality (A.2) and Lemma A.1, it holds

$$\|\boldsymbol{u}^{\mathrm{rm}} - \boldsymbol{u}_h^{\mathrm{rm}}\|_{0,\Omega} \leq \|\boldsymbol{u} - \boldsymbol{u}_h\|_{0,\Omega} + C\, h\, \|\boldsymbol{\lambda} - \boldsymbol{\lambda}_h\|_{\boldsymbol{\Lambda}},$$

and the result follows from (4.26) and Lemma 4.3. □

5. Convergence results

5.1. *A priori* estimates

Note that the result in Lemma 4.3 holds for any discrete space $\boldsymbol{\Lambda}_h$ under the assumption $\boldsymbol{\Lambda}_{\mathrm{rm}} \subset \boldsymbol{\Lambda}_h$. As such, method (4.4) achieves optimal convergence rates for any finite element subspace $\boldsymbol{\Lambda}_h$ with known best approximation properties. To illustrate, we use the polynomial space $\boldsymbol{\Lambda}_l$

$$\boldsymbol{\Lambda}_h \equiv \boldsymbol{\Lambda}_l := \left\{\boldsymbol{\mu} \in \boldsymbol{\Lambda} : \boldsymbol{\mu}|_F \in [\mathbb{P}^l(F)]^n,\, F \in \mathcal{E}_h\right\}, \tag{5.1}$$

where $\mathbb{P}^l(F)$, $l \geq 1$, stands for the space of piecewise polynomials of degree less than or equal to l on F, and $\boldsymbol{\mu}|_F$ was defined in (3.11). We closely follow the proof of a result in [30] to show that $\boldsymbol{\Lambda}_l$ has the desired approximation properties.

Lemma 5.1. *Suppose $\boldsymbol{w} \in \boldsymbol{H}^{m+1}(\Omega)$ with $1 \leq m \leq l+1$ and $l \geq 0$, and let $\boldsymbol{\mu} \in \boldsymbol{\Lambda}$ be such that $\boldsymbol{\mu} := \mathcal{E}(\boldsymbol{w})\,\boldsymbol{n}|_F$ for each $F \in \mathcal{E}_h$. There exists C such that*

$$\inf_{\boldsymbol{\mu}_l \in \boldsymbol{\Lambda}_l} \|\boldsymbol{\mu} - \boldsymbol{\mu}_l\|_{\boldsymbol{\Lambda}} \leq C\, h^m \|\boldsymbol{w}\|_{m+1,\Omega}, \tag{5.2}$$

where $\boldsymbol{\Lambda}_l$ is given in (5.1). Moreover, it holds

$$\inf_{\boldsymbol{\mu}_{\mathrm{rm}} \in \boldsymbol{\Lambda}_{\mathrm{rm}}} \|\boldsymbol{\mu} - \boldsymbol{\mu}_{\mathrm{rm}}\|_{\boldsymbol{\Lambda}} \leq C\, h\, \|\boldsymbol{w}\|_{2,\Omega}. \tag{5.3}$$

Proof. Assume $\boldsymbol{w} \in \boldsymbol{H}^{m+1}(\Omega)$, set $\boldsymbol{\mu} = \mathcal{E}(\boldsymbol{w})\,\boldsymbol{n}|_F$ for each $F \in \mathcal{E}_h$ and denote by Π^l the orthogonal projector in $\boldsymbol{L}^2(F)$ upon $[\mathbb{P}_l(F)]^n$, $l \geq 0$. Defining $\boldsymbol{\mu}_l := \Pi^l\, \boldsymbol{\mu}$, using the regularity of the mesh, and following closely the proof of Lemma 9 in [30] for each component of $\boldsymbol{w}$, and for each $K \in \mathcal{T}_h$, we get

$$(\boldsymbol{\mu} - \boldsymbol{\mu}_l, \boldsymbol{v})_{\partial K} \leq C\, h_K^m |\boldsymbol{w}|_{m+1,K}\, |\boldsymbol{v}|_{1,K} \quad \text{for all } \boldsymbol{v} \in \boldsymbol{H}^1(K).$$

Summing up over $K \in \mathcal{T}_h$ it holds

$$b(\boldsymbol{\mu} - \boldsymbol{\mu}_l, \boldsymbol{v}) = (\boldsymbol{\mu} - \boldsymbol{\mu}_l, \boldsymbol{v})_{\partial\mathcal{T}_h} \leq C\, h^m |\boldsymbol{w}|_{m+1,\Omega}\, |\boldsymbol{v}|_{1,\mathcal{T}_h} \leq C\, h^m \|\boldsymbol{w}\|_{m+1,\Omega}\, \|\boldsymbol{v}\|_{\mathbf{V}},$$

which immediately leads to

$$\sup_{\boldsymbol{v} \in \mathbf{V}} \frac{b(\boldsymbol{\mu} - \boldsymbol{\mu}_l,\, \boldsymbol{v})}{\|\boldsymbol{v}\|_{\mathbf{V}}} \leq C\, h^m \|\boldsymbol{w}\|_{m+1,\Omega}.$$

The result (5.2) follows using the equivalence of norms in Lemma A.3, and (5.3) follows using that $\boldsymbol{\Lambda}_0 \subset \boldsymbol{\Lambda}_{\mathrm{rm}}$ and (5.2). □

With the choice of $\boldsymbol{\Lambda}_l$ in (5.1), we are then ready to present rates of convergence.

Theorem 5.2. *Let* $(\boldsymbol{\lambda}, \boldsymbol{u}^{\mathrm{rm}}) \in \boldsymbol{\Lambda} \times \mathbf{V}_{\mathrm{rm}}$ *and* $(\boldsymbol{\lambda}_l, \boldsymbol{u}_h^{\mathrm{rm}}) \in \boldsymbol{\Lambda}_l \times \mathbf{V}_{\mathrm{rm}}$ *be the exact and the approximate solution of* (2.6) *and* (4.4)*, respectively, where* $\boldsymbol{\Lambda}_l$ *is given in* (5.1)*. Assuming* $\boldsymbol{u} \in \boldsymbol{H}^{m+1}(\Omega)$*, there exist* C *such that*

$$\|(\boldsymbol{\lambda} - \boldsymbol{\lambda}_l, \boldsymbol{u}^{\mathrm{rm}} - \boldsymbol{u}_h^{\mathrm{rm}})\|_{\boldsymbol{\Lambda}\times\mathbf{V}} \leq C\, h^m \,\|\boldsymbol{u}\|_{m+1,\Omega},$$
$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{0,\Omega} + \|\sigma - \sigma_h\|_{\mathrm{div}} \leq C\, h^m \,\|\boldsymbol{u}\|_{m+1,\Omega}.$$

Furthermore, if problem (1.1) *has smoothing properties, the following estimates hold*

$$\|\boldsymbol{u}^{\mathrm{rm}} - \boldsymbol{u}_h^{\mathrm{rm}}\|_{0,\Omega} + \|\boldsymbol{u} - \boldsymbol{u}_h\|_{0,\Omega} \leq C\, h^{m+1} \,\|\boldsymbol{u}\|_{m+1,\Omega},$$

where $1 \leq m \leq l+1$ *and* $l \geq 1$ *is the degree of polynomial functions in* $\boldsymbol{\Lambda}_l$*.*

Proof. The results follow using Lemmas 4.3, 4.5 and 5.1. □

Remark 5.3. Estimates in Theorem 5.2 point out that the errors in the natural norms for the displacement and the stress are super-convergent. For instance, if one adopts linear polynomial interpolation ($l = 1$) on faces to approximate the Lagrange multiplier, then Theorem 5.2 shows that $\|\nabla(\boldsymbol{u} - \boldsymbol{u}_h)\|_{0,\mathcal{T}_h} = O(h^2)$ and $\|\boldsymbol{u} - \boldsymbol{u}_h\|_{0,\Omega} = O(h^3)$.

Remark 5.4. At this point, it is interesting to count the number of local degrees of freedom necessary to approximate the variables using the space $\mathbf{V}_{\mathrm{rm}}(K) \times \boldsymbol{\Lambda}_l(K)$ in the case of a simplicial mesh, where $\boldsymbol{\Lambda}_l(K)$ stands for the space of functions in $\boldsymbol{\Lambda}_l$ restricted to K. First, $\dim \mathbf{V}_{\mathrm{rm}}(K) = \binom{n+1}{n-1}$, where we recall that $n \in \{2, 3\}$ is the dimension of Ω. Now, on a particular face, there are $n\binom{l+n-1}{n-1}$ degrees of freedom for $\boldsymbol{\Lambda}_l(K)$ if $l \geq 1$. Therefore, since there are $n+1$ faces belonging to K, $\dim \boldsymbol{\Lambda}_l(K) = n(n+1)\binom{l+n-1}{n-1}$, for $l \geq 1$. Therefore, the total number of local degrees of freedom is

$$\binom{n+1}{n-1}\left[1 + 2\binom{l+n-1}{n-1}\right], \quad l \geq 1.$$

As for the simplest element, *i.e.*, the pair of spaces $\mathbf{V}_{\mathrm{rm}}(K) \times \boldsymbol{\Lambda}_{\mathrm{rm}}(K)$, the total number of local degrees of freedom is

$$\binom{n+1}{n-1}(n+2)\,.$$

As such, there are 12 degrees of freedom total in 2D, while in the 3D case there are 30 degrees of freedom per element. Also, Theorem 5.2 holds (with $m = 1$) if one replaces $\boldsymbol{\Lambda}_l$ by $\boldsymbol{\Lambda}_{\mathrm{rm}}$ from Lemmas 4.3 and 4.5, and from (5.3).

5.2. *A posteriori* estimates

Now, we turn to *a posteriori* error estimates. With the definition of $\boldsymbol{u}_h$ given in (3.2), we propose the residual on faces as follows

$$\boldsymbol{r}_F := \begin{cases} \dfrac{1}{2}\,[\![\boldsymbol{u}_h]\!], & F \in \mathcal{E}_0 \\ \boldsymbol{g} - \boldsymbol{u}_h, & F \in \mathcal{E}_D, \end{cases} \tag{5.4}$$

where we recall $[\![.]\!]$ is given in (3.10), and we set

$$\eta_F := \frac{c_{\min}}{h_F^{1/2}}\|\boldsymbol{r}_F\|_{0,F}. \tag{5.5}$$

The estimator η is given by

$$\eta := \left[\sum_{K\in\mathcal{T}_h} \eta_K^2\right]^{1/2} \quad \text{with} \quad \eta_K^2 := \sum_{F\subset\partial K} \eta_F^2. \tag{5.6}$$

It will be also useful to adopt the following norm on $H(\mathrm{div};\Omega)$

$$\|\sigma\|_{\mathrm{div},h}^2 := \sum_{K\in\mathcal{T}_h} \Big(\|\sigma\|_{0,K}^2 + h_K^2\|\mathbf{div}\,\sigma\|_{0,K}^2\Big), \tag{5.7}$$

and the local norm

$$\|\boldsymbol{v}\|_{\mathbf{V},\omega_F}^2 := \sum_{K\in\omega_F} \Big(h_K^{-2}\|\boldsymbol{v}\|_{0,K}^2 + \|\mathcal{E}(\boldsymbol{v})\|_{0,K}^2\Big), \tag{5.8}$$

where ω_F corresponds to the set of elements sharing the face $F\in\mathcal{E}_h$. We are ready to establish the following *a posteriori* error estimate.

Theorem 5.5. *Let* $(\boldsymbol{\lambda}, \boldsymbol{u}^{\mathrm{rm}})$ *and* $(\boldsymbol{\lambda}_h, \boldsymbol{u}_h^{\mathrm{rm}})$ *be the solutions of* (4.3) *and* (4.4)*, respectively. Then, there exist positive constants* C_1 *and* C_2*, independent of* h*, such that*

$$\begin{aligned} c_{\min}\|\boldsymbol{u}-\boldsymbol{u}_h\|_{\mathbf{V}} + \|\sigma-\sigma_h\|_{\mathrm{div},h} &\le C_1\,\eta,\\ \eta_F &\le C_2\, c_{\min}\|\boldsymbol{u}-\boldsymbol{u}_h\|_{\mathbf{V},\omega_F}. \end{aligned}$$

Moreover, if we suppose smoothing properties, there exists C *such that*

$$c_{\min}\left(\frac{1}{h}\|\boldsymbol{u}-\boldsymbol{u}_h\|_{0,\Omega} + \|\,\mathcal{E}(\boldsymbol{u}-\boldsymbol{u}_h)\|_{0,\mathcal{T}_h}\right) + \|\sigma-\sigma_h\|_{\mathrm{div},h} \le C\,\eta.$$

Proof. We shall adapt the technique proposed in [3]. To first establish the lower bound, we use the definition of $\mathbf{B}(.,.)$ and follow closely ([3], Thm. 5.2) (in its vectorial version), to get

$$\begin{aligned} \mathbf{B}(\boldsymbol{\lambda}-\boldsymbol{\lambda}_h, \boldsymbol{u}^{\mathrm{rm}}-\boldsymbol{u}_h^{\mathrm{rm}}; \boldsymbol{\mu}, \boldsymbol{v}^{\mathrm{rm}}) &= \mathbf{B}(\boldsymbol{\lambda}-\boldsymbol{\lambda}_h, \boldsymbol{u}^{\mathrm{rm}}-\boldsymbol{u}_h^{\mathrm{rm}}; \boldsymbol{\mu}, \mathbf{0})\\ &= (\boldsymbol{\mu}, \boldsymbol{u}-\boldsymbol{u}_h)_{\partial\mathcal{T}_h}\\ &= -(\boldsymbol{\mu}, \boldsymbol{u}_h)_{\partial\mathcal{T}_h} + (\boldsymbol{\mu}, \boldsymbol{g})_{\partial\Omega_D}\\ &\le C\,\|\boldsymbol{\mu}\|_{\boldsymbol{\Lambda}}\,\eta. \end{aligned}$$

We therefore find from Lemma A.1 and Theorem 4.2 that

$$\begin{aligned} c_{\min}\|\boldsymbol{u}-\boldsymbol{u}_h\|_{\mathbf{V}} + \|\sigma-\sigma_h\|_{\mathrm{div},h} &\le C\,\|(\boldsymbol{\lambda}-\boldsymbol{\lambda}_h, \boldsymbol{u}^{\mathrm{rm}}-\boldsymbol{u}_h^{\mathrm{rm}})\|_{\boldsymbol{\Lambda}\times\mathbf{V}}\\ &\le C \sup_{(\boldsymbol{\mu},\boldsymbol{v}^{\mathrm{rm}})\in\boldsymbol{\Lambda}\times\mathbf{V}_{\mathrm{rm}}} \frac{\mathbf{B}(\boldsymbol{\lambda}-\boldsymbol{\lambda}_h, \boldsymbol{u}^{\mathrm{rm}}-\boldsymbol{u}_h^{\mathrm{rm}}; \boldsymbol{\mu}, \boldsymbol{v}^{\mathrm{rm}})}{\|(\boldsymbol{\mu}, \boldsymbol{v}^{\mathrm{rm}})\|_{\boldsymbol{\Lambda}\times\mathbf{V}}}\\ &\le C\,\eta. \end{aligned}$$

Now, we prove the upper bound. Let $\boldsymbol{\mu}^*\in\boldsymbol{\Lambda}$ such that $\boldsymbol{\mu}^*\,|_F = \boldsymbol{r}_F$ and $\boldsymbol{\mu}^*\,|_{F'} = \mathbf{0}$ for all $F'\neq F\in\mathcal{E}_0$, hence we get

$$\|\boldsymbol{r}_F\|_{0,F}^2 = 2\,(\boldsymbol{r}_F, [\![\boldsymbol{u}-\boldsymbol{u}_h]\!])_F \le 2\,\|\boldsymbol{r}_F\|_{0,F}\|[\![\boldsymbol{u}-\boldsymbol{u}_h]\!]\|_{0,F}.$$

Therefore, using the triangle inequality and trace inequality (A.3), and the mesh regularity it holds

$$\begin{aligned}\|\boldsymbol{r}_F\|_{0,F} &\leq 2\,\|[\![\boldsymbol{u}-\boldsymbol{u}_h]\!]\|_{0,F}\\ &\leq C\sum_{K\in\omega_F}\left[h_K^{-1}\|\boldsymbol{u}-\boldsymbol{u}_h\|_{0,K}^2+h_K\|\,\mathcal{E}(\boldsymbol{u}-\boldsymbol{u}_h)\|_{0,K}^2\right]^{1/2}\\ &\leq C\,h_F^{1/2}\,\|\boldsymbol{u}-\boldsymbol{u}_h\|_{\mathbf{V},\omega_F}.\end{aligned}$$

As for $F\in\mathcal{E}_D$, we observe that $\|\boldsymbol{r}_F\|_{0,F}\leq\|\boldsymbol{u}-\boldsymbol{u}_h\|_{0,F}$ and, then, the above estimate also holds following an analogous argument. The last result follows from Lemma 4.5 and following closely the arguments presented in ([3], Cor. 5.3) (in its vectorial version). □

6. Two-level analysis

We establish first in this section the conditions under which the two-level version of the MHM method (3.8) remains well-posed and keeps its best approximation property. There is a great deal of flexibility in the choice of the local finite dimensional spaces and in the second-level numerical method. Here, we keep the two-level approach as simple as possible, and select a conforming second-level finite dimensional space $\tilde{\mathbf{V}}_h(K)\subset\tilde{\mathbf{V}}(K)$, where $\tilde{\mathbf{V}}(K)$ stands for the functions in $\tilde{\mathbf{V}}$ restricted to $K\in\mathcal{T}_h$, and define

$$\tilde{\mathbf{V}}_h\;:=\;\oplus_{K\in\mathcal{T}_h}\tilde{\mathbf{V}}_h(K)\subset\tilde{\mathbf{V}}.$$

We include the impact of the second-level discretization in the MHM method by replacing the bilinear form $a(\cdot,\cdot)$ in (4.3) with

$$a_h(\boldsymbol{\mu}_h,\boldsymbol{\lambda}_h):=(\boldsymbol{\mu}_h,T_h\,\boldsymbol{\lambda}_h)_{\partial\mathcal{T}_h},\tag{6.1}$$

where $T_h:\boldsymbol{\Lambda}\to\tilde{\mathbf{V}}_h$ is such that $T_h\,\boldsymbol{\mu}_h\,|_K$ satisfies

$$(\mathcal{A}\,\mathcal{E}(T_h\,\boldsymbol{\mu}_h),\mathcal{E}(\tilde{\boldsymbol{v}}_h))_K=-(\boldsymbol{\mu}_h,\tilde{\boldsymbol{v}}_h)_{\partial K}\quad\text{for all }\tilde{\boldsymbol{v}}_h\in\tilde{\mathbf{V}}_h,\tag{6.2}$$

and $(\boldsymbol{\mu},\hat{T}\,\boldsymbol{f})_{\partial\mathcal{T}_h}$ is replaced by $(\boldsymbol{\mu},\hat{T}_h\,\boldsymbol{f})_{\partial\mathcal{T}_h}$, where $\hat{T}_h:\boldsymbol{L}^2(\Omega)\to\tilde{\mathbf{V}}_h$ is such that $\hat{T}_h\,\boldsymbol{q}\,|_K$ satisfies

$$(\mathcal{A}\,\mathcal{E}(\hat{T}_h\,\boldsymbol{q}),\mathcal{E}(\tilde{\boldsymbol{v}}_h))_K=(\boldsymbol{q},\tilde{\boldsymbol{v}}_h)_K\quad\text{for all }\tilde{\boldsymbol{v}}_h\in\tilde{\mathbf{V}}_h.\tag{6.3}$$

The problems above are the standard Galerkin method set over $\tilde{\mathbf{V}}_h$ restricted to each $K\in\mathcal{T}_h$. The goal is to approximate the solutions of elliptic problems (2.4)–(2.5).

The corresponding two-level MHM method reads: *find* $(\boldsymbol{\lambda}_h,\boldsymbol{u}_h^{\rm rm})\in\boldsymbol{\Lambda}_h\times\mathbf{V}_{\rm rm}$ *such that*

$$\mathbf{B}_h(\boldsymbol{\lambda}_h,\boldsymbol{u}_h^{\rm rm};\boldsymbol{\mu}_h,\boldsymbol{v}^{\rm rm})=\mathbf{F}_h(\boldsymbol{\mu}_h,\boldsymbol{v}^{\rm rm})\quad\text{for all }(\boldsymbol{\mu}_h,\boldsymbol{v}^{\rm rm})\in\boldsymbol{\Lambda}_h\times\mathbf{V}_{\rm rm},\tag{6.4}$$

where

$$\begin{aligned}\mathbf{B}_h(\boldsymbol{\lambda},\boldsymbol{u}^{\rm rm};\boldsymbol{\mu},\boldsymbol{v}^{\rm rm})&:=a_h(\boldsymbol{\lambda},\boldsymbol{\mu})+b(\boldsymbol{\mu},\boldsymbol{u}^{\rm rm})+b(\boldsymbol{\lambda},\boldsymbol{v}^{\rm rm}),\\ \mathbf{F}_h(\boldsymbol{\mu},\boldsymbol{v}^{\rm rm})&:=(\boldsymbol{f},\,\boldsymbol{v}^{\rm rm})_{\mathcal{T}_h}-(\boldsymbol{\mu},\hat{T}_h\,\boldsymbol{f})_{\partial\mathcal{T}_h}+(\boldsymbol{\mu},\,\boldsymbol{g}_D)_{\partial\Omega}.\end{aligned}$$

Observe that the invertibility of the matrix associated with $\mathbf{B}_h(\cdot,\cdot)$ comes down to invertibility of the symmetric form $a_h(.,.)$ on the nullspace

$$\mathcal{N}_h:=\mathcal{N}\cap\boldsymbol{\Lambda}_h,\tag{6.5}$$

where $\mathcal{N}$ is given in (4.20) and $\boldsymbol{\Lambda}_h$ satisfies (4.1) (we have already dealt with the other constituents in the previous analysis). Such a result will be achieved by showing the form $-a_h(.,.)$ is coercive on $\mathcal{N}_h$ with respect to the norm $\|.\|_{\boldsymbol{\Lambda}}$. Overall, this relies on the choice of the space $\tilde{\mathbf{V}}_h$. A *sufficient* condition for $\tilde{\mathbf{V}}_h$ is proposed in the next lemma.

Lemma 6.1. *Let $a_h(\cdot,\cdot)$ be given in (6.1), and assume that it holds*

$$\boldsymbol{\mu}_h \in \mathcal{N}_h, \quad (\boldsymbol{\mu}_h, \tilde{\boldsymbol{v}}_h)_{\partial K} = 0 \quad \textit{for all } \tilde{\boldsymbol{v}}_h \in \tilde{\mathbf{V}}_h \textit{ and } K \in \mathcal{T}_h \Rightarrow \boldsymbol{\mu}_h = 0. \tag{6.6}$$

Then, there exists C such that

$$-a_h(\boldsymbol{\mu}_h, \boldsymbol{\mu}_h) \geq C\, \|\boldsymbol{\mu}_h\|^2_{\boldsymbol{\Lambda}} \quad \textit{for all } \boldsymbol{\mu}_h \in \mathcal{N}_h. \tag{6.7}$$

Proof. First notice that (6.6) implies T_h is injective on $\mathcal{N}_h$. In fact, if $T_h\,\boldsymbol{\mu}_h = 0$, then from (6.2) it holds, for all $K \in \mathcal{T}_h$,

$$0 = (\mathcal{A}\,\mathcal{E}(T_h\,\boldsymbol{\mu}_h), \mathcal{E}(\tilde{\boldsymbol{v}}_h))_K = -(\boldsymbol{\mu}_h, \tilde{\boldsymbol{v}}_h)_{\partial K} \quad \text{for all } \boldsymbol{v}_h \in \tilde{\mathbf{V}}_h,$$

and from (6.6) we get $\boldsymbol{\mu}_h = 0$. As a result, $\dim \mathcal{N}_h = \dim T_h(\mathcal{N}_h)$, where $T_h(\mathcal{N}_h) \subset \tilde{\mathbf{V}}_h$ is the image of T_h restricted to functions in $\mathcal{N}_h$, and $\|T_h\,\boldsymbol{\mu}_h\|_{\mathbf{V}}$ turns out to be a norm over the space $\mathcal{N}_h$. To establish the aforementioned coercivity result, we first prove that such a result holds with respect to the norm $\|T_h\,\boldsymbol{\mu}_h\|_{\mathbf{V}}$. In fact, let $\boldsymbol{\lambda}_h = \boldsymbol{\mu}_h \in \mathcal{N}_h$ in (6.1). Using (6.2), there is C such that

$$\begin{aligned}
-a_h(\boldsymbol{\mu}_h, \boldsymbol{\mu}_h) &= -(\boldsymbol{\mu}_h, T_h\,\boldsymbol{\mu}_h)_{\partial\mathcal{T}_h} \\
&= (\mathcal{A}\,\mathcal{E}(T_h\,\boldsymbol{\mu}_h), \mathcal{E}(T_h\,\boldsymbol{\mu}_h))_{\mathcal{T}_h} \\
&\geq c_{\min} \|\,\mathcal{E}(T_h\,\boldsymbol{\mu}_h)\|^2_{0,\mathcal{T}_h} \\
&\geq C\, c_{\min} \|T_h\,\boldsymbol{\mu}_h\|^2_{\mathbf{V}},
\end{aligned} \tag{6.8}$$

where we used the Korn's inequality. Next, observe that from (6.6) there exits a C such that

$$\sup_{\tilde{\boldsymbol{v}}_h \in \tilde{\mathbf{V}}_h} \frac{(\boldsymbol{\mu}_h, \tilde{\boldsymbol{v}}_h)_{\partial\mathcal{T}_h}}{\|\tilde{\boldsymbol{v}}_h\|_{\mathbf{V}}} \geq C \sup_{\tilde{\boldsymbol{v}} \in \tilde{\mathbf{V}}} \frac{(\boldsymbol{\mu}_h, \tilde{\boldsymbol{v}})_{\partial\mathcal{T}_h}}{\|\tilde{\boldsymbol{v}}\|_{\mathbf{V}}} \quad \text{for all } \boldsymbol{\mu}_h \in \mathcal{N}_h, \tag{6.9}$$

as the left-hand side above turns out to be a norm over $\mathcal{N}_h$. The independence of C with respect to h follows from a standard scaling argument (see [15], p. 111, for instance). Inequality (6.9) is also found in ([30], Lem. 10) with minor differences but used for a different purpose. Now, from (6.9) and since $\|\tilde{\boldsymbol{v}}\|_{\mathbf{V}} \leq \|\boldsymbol{v}\|_{\mathbf{V}}$ for all $\boldsymbol{v} \in \mathbf{V}$, and recalling that $b(\cdot,\cdot) = (\cdot,\cdot)_{\partial\mathcal{T}_h}$, we get from (4.12) that

$$\sup_{\tilde{\boldsymbol{v}}_h \in \tilde{\mathbf{V}}_h} \frac{(\boldsymbol{\mu}_h, \tilde{\boldsymbol{v}}_h)_{\partial\mathcal{T}_h}}{\|\tilde{\boldsymbol{v}}_h\|_{\mathbf{V}}} \geq C \sup_{\boldsymbol{v} \in \mathbf{V}} \frac{(\boldsymbol{\mu}_h, \boldsymbol{v})_{\partial\mathcal{T}_h}}{\|\boldsymbol{v}\|_{\mathbf{V}}} \geq C\, \|\boldsymbol{\mu}_h\|_{\boldsymbol{\Lambda}} \quad \text{for all } \boldsymbol{\mu}_h \in \mathcal{N}_h.$$

The previous inequality, (6.2) and the Cauchy−Schwarz's inequality together imply,

$$\begin{aligned}
C\, \|\boldsymbol{\mu}_h\|_{\boldsymbol{\Lambda}} &\leq \sup_{\tilde{\boldsymbol{v}}_h \in \tilde{\mathbf{V}}_h} \frac{(\boldsymbol{\mu}_h, \tilde{\boldsymbol{v}}_h)_{\partial\mathcal{T}_h}}{\|\tilde{\boldsymbol{v}}_h\|_{\mathbf{V}}} \\
&= \sup_{\tilde{\boldsymbol{v}}_h \in \tilde{\mathbf{V}}_h} -\frac{(\mathcal{A}\,\mathcal{E}(T_h\,\boldsymbol{\mu}_h), \mathcal{E}(\tilde{\boldsymbol{v}}_h))_{\mathcal{T}_h}}{\|\tilde{\boldsymbol{v}}_h\|_{\mathbf{V}}} \\
&\leq C\, c_{\max} \|T_h\,\boldsymbol{\mu}_h\|_{\mathbf{V}}.
\end{aligned}$$

We conclude that, for $\boldsymbol{\mu}_h \in \mathcal{N}_h$, there exists a positive constant C such that

$$\|\boldsymbol{\mu}_h\|_{\boldsymbol{\Lambda}} \leq C\, c_{\max}\, \|T_h\,\boldsymbol{\mu}_h\|_{\mathbf{V}}, \tag{6.10}$$

and consequently, we arrive at the required result for $a_h(.,.)$, *i.e.*,

$$-a_h(\boldsymbol{\mu}_h, \boldsymbol{\mu}_h) \geq C\, \frac{c_{\min}}{c_{\max}} \|\boldsymbol{\mu}_h\|^2_{\boldsymbol{\Lambda}} \quad \text{for all } \boldsymbol{\mu}_h \in \mathcal{N}_h. \qquad \square$$

Notice that the operators T_h and $\hat{T}_h$ are bounded, *e.g.*, there are constants C such that

$$\|T_h\,\boldsymbol{\mu}\|_{\mathbf{V}} \le C\,\|\boldsymbol{\mu}\|_{\boldsymbol{\Lambda}} \quad \text{and} \quad \|\hat{T}_h\,\boldsymbol{q}\|_{\mathbf{V}} \le C\,\|\boldsymbol{q}\|_{0,\Omega} \quad \text{for all } \boldsymbol{\mu} \in \boldsymbol{\Lambda},\, \boldsymbol{q} \in \boldsymbol{L}^2(\Omega). \tag{6.11}$$

The inequality for T_h follows from (A.2), (6.2) and Lemma A.3, and for $\hat{T}_h$ from (A.2), (6.3) and Cauchy−Schwarz's inequality. Now, using (6.11) and Lemma 6.1, and following the proof of Theorem 4.2, we conclude the well-posedness of the two-level MHM method (6.4) in the next theorem.

Theorem 6.2. *Suppose* $(\boldsymbol{\lambda}_h, \boldsymbol{u}^{\rm rm})$, $(\boldsymbol{\mu}_h, \boldsymbol{v}^{\rm rm}) \in \boldsymbol{\Lambda}_h \times \mathbf{V}_{\rm rm}$. *Then, there exists* C *such that*

$$\mathbf{B}_h(\boldsymbol{\lambda}_h, \boldsymbol{u}^{\rm rm}; \boldsymbol{\mu}_h, \boldsymbol{v}^{\rm rm}) \;\le\; C\,\|(\boldsymbol{\lambda}_h, \boldsymbol{u}^{\rm rm})\|_{\boldsymbol{\Lambda}\times\mathbf{V}}\,\|(\boldsymbol{\mu}_h, \boldsymbol{v}^{\rm rm})\|_{\boldsymbol{\Lambda}\times\mathbf{V}}. \tag{6.12}$$

Moreover, assuming that (6.6) *holds, there exists a positive constant* α*, independent of* h*, such that*

$$\sup_{(\boldsymbol{\mu}_h, \boldsymbol{v}^{\rm rm}) \in \boldsymbol{\Lambda}_h \times \mathbf{V}_{\rm rm}} \frac{\mathbf{B}_h(\boldsymbol{\lambda}_h, \boldsymbol{u}^{\rm rm}; \boldsymbol{\mu}_h, \boldsymbol{v}^{\rm rm})}{\|(\boldsymbol{\mu}_h, \boldsymbol{v}^{\rm rm})\|_{\boldsymbol{\Lambda}\times\mathbf{V}}} \;\ge\; \alpha\,\|(\boldsymbol{\lambda}_h, \boldsymbol{u}^{\rm rm})\|_{\boldsymbol{\Lambda}\times\mathbf{V}}, \tag{6.13}$$

for all $(\boldsymbol{\lambda}_h, \boldsymbol{u}^{\rm rm}) \in \boldsymbol{\Lambda}_h \times \mathbf{V}_{\rm rm}$. *Hence, problem* (6.4) *is well-posed.*

Now, let us quantify the impact of the two-level approach on approximation results. To this end, observe that the two-level discretization impacts the consistency of MHM method (3.1). This can be measured through the following best approximation result, which is an incarnation of the first Strang's lemma.

Lemma 6.3. *Let* $(\boldsymbol{\lambda}, \boldsymbol{u}^{\rm rm})$ *and* $(\boldsymbol{\lambda}_h, \boldsymbol{u}_h^{\rm rm})$ *be the solutions of* (4.3) *and* (6.4)*, respectively. It holds that there exists* C *such that*

$$\|(\boldsymbol{\lambda} - \boldsymbol{\lambda}_h, \boldsymbol{u}^{\rm rm} - \boldsymbol{u}_h^{\rm rm})\|_{\boldsymbol{\Lambda}\times\mathbf{V}} \;\le\; C\,\left(\inf_{\boldsymbol{\mu}_h \in \boldsymbol{\Lambda}_h} \|\boldsymbol{\lambda} - \boldsymbol{\mu}_h\|_{\boldsymbol{\Lambda}} + \|(T - T_h)\,\boldsymbol{\lambda} + (\hat{T} - \hat{T}_h)\,\boldsymbol{f}\|_{\mathbf{V}}\right).$$

Proof. Choose arbitrary $\boldsymbol{\mu}_h \in \boldsymbol{\Lambda}_h$. By (4.12) and (4.17), for all $(\boldsymbol{\gamma}_h, \boldsymbol{w}^{\rm rm}) \in \boldsymbol{\Lambda}_h \times \mathbf{V}_{\rm rm}$, there is a constant C such that

$$\begin{aligned}
\mathbf{B}_h(\boldsymbol{\mu}_h - \boldsymbol{\lambda}_h, \boldsymbol{u}^{\rm rm} - \boldsymbol{u}_h^{\rm rm}; \boldsymbol{\gamma}_h, \boldsymbol{w}^{\rm rm}) &= \mathbf{B}_h(\boldsymbol{\mu}_h, \boldsymbol{u}^{\rm rm}; \boldsymbol{\gamma}_h, \boldsymbol{w}^{\rm rm}) - \mathbf{F}_h(\boldsymbol{\gamma}_h, \boldsymbol{w}^{\rm rm}) \\
&= \mathbf{B}_h(\boldsymbol{\mu}_h, \boldsymbol{u}^{\rm rm}; \boldsymbol{\gamma}_h, \boldsymbol{w}^{\rm rm}) + (\mathbf{F}(\boldsymbol{\gamma}_h, \boldsymbol{w}^{\rm rm}) - \mathbf{F}_h(\boldsymbol{\gamma}_h, \boldsymbol{w}^{\rm rm})) - \mathbf{F}(\boldsymbol{\gamma}_h, \boldsymbol{w}^{\rm rm}) \\
&= (\boldsymbol{\gamma}_h, (T_h - T)\,\boldsymbol{\mu}_h - (\hat{T}_h - \hat{T})\boldsymbol{f})_{\mathcal{T}_h} + \mathbf{B}(\boldsymbol{\mu}_h - \boldsymbol{\lambda}, 0; \boldsymbol{\gamma}_h, \boldsymbol{w}^{\rm rm}) \\
&\le (\|(T_h - T)\,\boldsymbol{\mu}_h + (\hat{T}_h - \hat{T})\,\boldsymbol{f}\|_{\mathbf{V}} + C\,\|\boldsymbol{\mu}_h - \boldsymbol{\lambda}\|_{\boldsymbol{\Lambda}})\|(\boldsymbol{\gamma}_h, \boldsymbol{w}^{\rm rm})\|_{\boldsymbol{\Lambda}\times\mathbf{V}} \\
&\le (\|(T_h - T)\,(\boldsymbol{\lambda} - \boldsymbol{\mu}_h)\|_{\mathbf{V}} + \|(T_h - T)\,\boldsymbol{\lambda} + (\hat{T}_h - \hat{T})\,\boldsymbol{f}\|_{\mathbf{V}} \\
&\qquad + C\,\|\boldsymbol{\mu}_h - \boldsymbol{\lambda}\|_{\boldsymbol{\Lambda}})\|(\boldsymbol{\gamma}_h, \boldsymbol{w}^{\rm rm})\|_{\boldsymbol{\Lambda}\times\mathbf{V}} \\
&\le (\|(T_h - T)\,\boldsymbol{\lambda} + (\hat{T}_h - \hat{T})\,\boldsymbol{f}\|_{\mathbf{V}} + C\,\|\boldsymbol{\mu}_h - \boldsymbol{\lambda}\|_{\boldsymbol{\Lambda}})\|(\boldsymbol{\gamma}_h, \boldsymbol{w}^{\rm rm})\|_{\boldsymbol{\Lambda}\times\mathbf{V}},
\end{aligned}$$

where we used the stability result for T and T_h from Lemma A.1 and (6.11), respectively. Therefore, from (6.13), we get

$$\|(\boldsymbol{\mu}_h - \boldsymbol{\lambda}_h, \boldsymbol{u}_h^{\rm rm} - \boldsymbol{u}^{\rm rm})\|_{\boldsymbol{\Lambda}\times\mathbf{V}} \le \frac{1}{\alpha}\left(\|(T - T_h)\,\boldsymbol{\lambda} + (\hat{T} - \hat{T}_h)\,\boldsymbol{f}\|_{\mathbf{V}} + C\,\|\boldsymbol{\lambda} - \boldsymbol{\mu}_h\|_{\boldsymbol{\Lambda}}\right),$$

and from the triangle inequality

$$\|(\boldsymbol{\lambda} - \boldsymbol{\lambda}_h, \boldsymbol{u}^{\rm rm} - \boldsymbol{u}_h^{\rm rm})\|_{\boldsymbol{\Lambda}\times\mathbf{V}} \le \|(\boldsymbol{\lambda} - \boldsymbol{\mu}_h, \mathbf{0})\|_{\boldsymbol{\Lambda}\times\mathbf{V}} + \|(\boldsymbol{\mu}_h - \boldsymbol{\lambda}_h, \boldsymbol{u}^{\rm rm} - \boldsymbol{u}_h^{\rm rm})\|_{\boldsymbol{\Lambda}\times\mathbf{V}}$$

the result follows. ∎

Hereafter, we shall use the characterization of $\boldsymbol{u}$ and σ given in (2.7), and $\boldsymbol{u}_h$, σ_h redefined by

$$\boldsymbol{u}_h := \boldsymbol{u}_h^{\rm rm} + T_h\,\boldsymbol{\lambda}_h + \hat{T}_h\,\boldsymbol{f}, \quad \sigma_h = \mathcal{A}\,\mathcal{E}(T_h\,\boldsymbol{\lambda}_h + \hat{T}_h\,\boldsymbol{f}), \tag{6.14}$$

where $(\boldsymbol{u}_h^{\rm rm}, \boldsymbol{\lambda}_h) \in \mathbf{V}_{\rm rm} \times \boldsymbol{\Lambda}_h$ is the solution of (6.4). A similar best approximation result of Lemma 6.3 is also available adopting natural norms. This is presented next.

Corollary 6.4. *Let $(\boldsymbol{\lambda}, \boldsymbol{u}^{\rm rm})$ and $(\boldsymbol{\lambda}_h, \boldsymbol{u}_h^{\rm rm})$ be the solutions of (4.3) and (6.4), respectively. There is a constant C such that*

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{\mathbf{V}} \;\le\; C \left(\inf_{\boldsymbol{\mu}_h \in \boldsymbol{\Lambda}_h} \|\boldsymbol{\lambda} - \boldsymbol{\mu}_h\|_{\boldsymbol{\Lambda}} + \|(T - T_h)\boldsymbol{\lambda} + (\hat{T} - \hat{T}_h)\boldsymbol{f}\|_{\mathbf{V}} \right).$$

Furthermore, if problem (1.1) has smoothing properties, it holds

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{0,\Omega} \;\le\; C\,h \left(\inf_{\boldsymbol{\mu}_h \in \boldsymbol{\Lambda}_h} \|\boldsymbol{\lambda} - \boldsymbol{\mu}_h\|_{\boldsymbol{\Lambda}} + \|(T - T_h)\boldsymbol{\lambda} + (\hat{T} - \hat{T}_h)\boldsymbol{f}\|_{\mathbf{V}} \right).$$

Proof. Observing that

$$\boldsymbol{e} = \boldsymbol{u} - \boldsymbol{u}_h = \boldsymbol{u}^{\rm rm} - \boldsymbol{u}_h^{\rm rm} + (T - T_h)\boldsymbol{\lambda} + T_h\,(\boldsymbol{\lambda} - \boldsymbol{\lambda}_h) + (\hat{T} - \hat{T}_h)\boldsymbol{f},$$

the proof then follows using the stability of T_h in (6.11) and Lemma 6.3 for the estimate in the $\|.\|_{\mathbf{V}}$ norm. As for the estimate in the $\boldsymbol{L}^2$ norm, we follows the same lines as the proof of Lemma 4.5. □

Remark 6.5. Another option to approximate the functions $T\,\boldsymbol{\lambda}_h$ (*e.g.* $T\,\boldsymbol{\psi}_i$) and $\hat{T}\,\boldsymbol{f}$ locally is to rewrite (2.4)−(2.5) in their mixed form and adopt a well-established mixed method to solve them. In this case, the increased complexity could be offset by a precise stress tensor σ_h in $H(\mathrm{div};\Omega)$ post-processed from the two-level primal variable (6.14) or an improvement in the robustness of the MHM method in the incompressible limiting case. This alternative deserves further investigation and will be addressed in the future.

6.1. Selecting $\tilde{\mathbf{V}}_h$

Here we propose a two-dimensional family of spaces $\tilde{\mathbf{V}}_h$ that fulfills (6.6), defined by

$$\tilde{\mathbf{V}}_h := \left\{ \boldsymbol{v}_h \in \tilde{\mathbf{V}} \;:\; \boldsymbol{v}_h\,|_K \in [\mathbb{S}^k(K)]^2,\, K \in \mathcal{T}_h \right\}, \tag{6.15}$$

where $\mathbb{S}^k(K) := \mathbb{P}^k(K)$ or $\mathbb{Q}^k(K)$, and $\mathbb{Q}^k(K)$ stands for the space of tensor polynomial function of order k at most in $K \in \mathcal{T}_h$. Note that it is based on a *single element*, and requires no further discretization of each element. It can be seen as a p-method at the second level.

Let $\boldsymbol{\Lambda}_h \equiv \boldsymbol{\Lambda}_l$ be given in (5.1). Clearly, if $k \le l$ then condition (6.6) cannot occur since $p\,k = \dim \mathbb{S}^k(K) \le p\,\dim \mathbb{P}^l(F) = p\,(l+1)$, where $p = 3$ or $p = 4$ if $\mathbb{S}^k(K) = \mathbb{P}^k(K)$ or $\mathbb{Q}^k(K)$, respectively. The following result establishes the compatibility condition between the degrees l and k such that the condition (6.6) holds.

Lemma 6.6. *Assume that $l \ge 0$. Then, if k satisfies*

$$\begin{cases} k \ge l + 1 & \text{if } l \text{ is even}, \\ k \ge l + 2 & \text{if } l \text{ is odd}, \end{cases} \tag{6.16}$$

when $\mathbb{S}^k(K) = \mathbb{P}^k(K)$ or $k \ge l + 2$ when $\mathbb{S}^k(K) = \mathbb{Q}^k(K)$, then (6.6) holds.

Proof. Assume that $k \geq l+1$ with $l \geq 0$, and take $\boldsymbol{\mu}_h \in \mathcal{N}_h$ such that, for all $\tilde{\boldsymbol{v}}_h \in \tilde{\mathbf{V}}_h$, it holds $(\boldsymbol{\mu}_h, \tilde{\boldsymbol{v}}_h)_{\partial K} = 0$ for all $K \in \mathcal{T}_h$. Observe that this condition corresponds to the assumption that each component μ_h^i of $\boldsymbol{\mu}_h \in \mathcal{N}_h$ satisfies

$$(\mu_h^i, v_h^i)_{\partial K} = 0 \quad \text{for all } K \in \mathcal{T}_h \quad \text{and } i = 1,\, 2, \tag{6.17}$$

where v_h^i is the i-component of $\boldsymbol{v}_h \in \mathbf{V}_h$. Now, from Lemmas 4 and 7 in [30] μ_h^i vanishes if and only if $v_h^i \in \mathbb{P}^k(K)$, with k satisfying (6.16), or if $v_h^i \in \mathbb{Q}^k(K)$ with $k \geq l+2$, and the result follows. □

Remark 6.7. In the case that l is odd or $\mathbb{S}^k(K) = \mathbb{Q}^k(K)$, Lemma 6.6 points out that the minimal interpolation, namely polynomial functions of degree $k = l+1$, cannot be adopted to approximate second-level solutions. In [30], Lemmas 4 and 7, it has been shown that the local space of non-trivial polynomial functions μ_h^i satisfying (6.17) is one dimensional with basis of degree l in both cases. As a result, $\tilde{\mathbf{V}}_h(K)$ enhanced with a polynomial function of degree $l+2$ (resp. $l+2$ or $l+3$ depending whether l is even or odd) in each element $K \in \mathcal{T}_h$ when $\mathbb{S}^k(K) = \mathbb{P}^k(K)$ (resp. $\mathbb{Q}^k(K)$), hereafter denoted by b_K, leads condition (6.17) to be fulfilled (see [30], Lems. 6 and 8, for details).

We may take advantage of the characterization of the non-trivial functions satisfying (6.17) given in the previous remark to decrease the computational cost involved in solving the second level problem. This is accomplished by making the minimal interpolation choice (*i.e.* $k = l+1$) available for the odd case if $\mathbb{S}^k(K) = \mathbb{P}^k(K)$ and for the case $\mathbb{S}^k(K) = \mathbb{Q}^k(K)$.

To this end, we redefine the operator T_h given in (6.2). Let us denote by $\mathcal{B}_K$ the one-dimensional local space generated through the function b_K. The desired result is presented next using that $\sum_{K\in\mathcal{T}_h} [\mathcal{B}_K]^2 \cap \tilde{\mathbf{V}}_h = \{\mathbf{0}\}$.

Lemma 6.8. *Let $k = l+1$ with $l \geq 0$. Assume either $\mathbb{S}^k(K) = \mathbb{Q}^k(K)$ or $\mathbb{S}^k(K) = \mathbb{P}^k(K)$ and l is odd. Redefine $T_h : \boldsymbol{\Lambda} \to \tilde{\mathbf{V}}_h$ replacing (6.2) by*

$$\begin{aligned}&(\mathcal{A}\,\mathcal{E}(T_h\,\boldsymbol{\lambda}_h), \mathcal{E}(\tilde{\boldsymbol{v}}_h))_K + (\mathcal{A}\,\mathcal{E}(\mathcal{P}\,T_h\,\boldsymbol{\lambda}_h), \mathcal{E}(\tilde{\boldsymbol{v}}_h))_K \\ &= -(\boldsymbol{\lambda}_h, \tilde{\boldsymbol{v}}_h)_{\partial K} - (\mathcal{A}\,\mathcal{E}(\hat{\mathcal{P}}\,\boldsymbol{\lambda}_h), \mathcal{E}(\tilde{\boldsymbol{v}}_h))_K \quad \textit{for all } \tilde{\boldsymbol{v}}_h \in \tilde{\mathbf{V}}_h,\end{aligned} \tag{6.18}$$

where $\mathcal{P} : \tilde{\mathbf{V}}_h \to \sum_{K\in\mathcal{T}_h} [\mathcal{B}_K]^2$ and $\hat{\mathcal{P}} : \boldsymbol{\Lambda}_h \to \sum_{K\in\mathcal{T}_h} [\mathcal{B}_K]^2$ are such that, given $\tilde{\boldsymbol{v}}_h \in \tilde{\mathbf{V}}_h$ and $\boldsymbol{\mu}_h \in \boldsymbol{\Lambda}_h$, they satisfy respectively,

$$\begin{aligned}(\mathcal{A}\,\mathcal{E}(\mathcal{P}\,\tilde{\boldsymbol{v}}_h), \mathcal{E}(b_K\,\boldsymbol{e}_i))_K &= -(\mathcal{A}\,\mathcal{E}(\tilde{\boldsymbol{v}}_h), \mathcal{E}(b_K\,\boldsymbol{e}_i))_K \\ (\mathcal{A}\,\mathcal{E}(\hat{\mathcal{P}}\,\boldsymbol{\mu}_h), \mathcal{E}(b_K\,\boldsymbol{e}_i))_K &= -(\boldsymbol{\mu}_h, b_K\,\boldsymbol{e}_i)_{\partial K} \quad i = 1,\, 2,\end{aligned} \tag{6.19}$$

and $b_K \in \mathcal{B}_K$ and $\boldsymbol{e}_i$ is the canonical basis in $\mathbb{R}^2$. Hence, T_h is an injective operator when restricted to $\mathcal{N}_h$.

Proof. From [30] we have that $\bar{T}_h : \mathcal{N}_h \to \tilde{\mathbf{V}}_h \oplus \sum_{K\in\mathcal{T}_h} [\mathcal{B}_K]^2$ satisfying

$$(\mathcal{A}\,\mathcal{E}(\bar{T}_h\,\boldsymbol{\lambda}_h), \mathcal{E}(\tilde{\boldsymbol{w}}_h))_K = -(\boldsymbol{\lambda}_h, \tilde{\boldsymbol{w}}_h)_{\partial K} \quad \text{for all } \tilde{\boldsymbol{w}}_h \in \tilde{\mathbf{V}}_h \oplus \sum_{K\in\mathcal{T}_h} [\mathcal{B}_K]^2 \tag{6.20}$$

is injective. Using the (unique) decomposition

$$\bar{T}_h\,\boldsymbol{\lambda}_h \;=\; \tilde{\boldsymbol{u}}_h^k + \tilde{\boldsymbol{u}}_h^b,$$

with $\tilde{\boldsymbol{u}}_h^k \in \tilde{\mathbf{V}}_h$ and $\tilde{\boldsymbol{u}}_h^b \in \sum_{K\in\mathcal{T}_h} [\mathcal{B}_K]^2$, we observe that (6.20) is completely equivalent to the following system

$$\begin{cases} (\mathcal{A}\,\mathcal{E}(\tilde{\boldsymbol{u}}_h^k), \mathcal{E}(\tilde{\boldsymbol{v}}_h))_K + (\mathcal{A}\,\mathcal{E}(\tilde{\boldsymbol{u}}_h^b), \mathcal{E}(\tilde{\boldsymbol{v}}_h))_K = -(\boldsymbol{\lambda}_h, \tilde{\boldsymbol{v}}_h)_{\partial K} & \text{for all } \tilde{\boldsymbol{v}}_h \in \tilde{\mathbf{V}}_h \\ (\mathcal{A}\,\mathcal{E}(\tilde{\boldsymbol{u}}_h^k), \mathcal{E}(b_K\,\boldsymbol{e}_i))_K + (\mathcal{A}\,\mathcal{E}(\tilde{\boldsymbol{u}}_h^b), \mathcal{E}(b_K\,\boldsymbol{e}_i))_K = -(\boldsymbol{\lambda}_h, b_K\,\boldsymbol{e}_i)_{\partial K} & \text{for } b_K \in \mathcal{B}_K. \end{cases} \tag{6.21}$$

Observe that the second equation leads to the following characterization

$$\tilde{\boldsymbol{u}}_h^b \;=\; \mathcal{P}\,\tilde{\boldsymbol{u}}_h^k + \hat{\mathcal{P}}\boldsymbol{\lambda}_h,$$

where $\mathcal{P}$ and $\hat{\mathcal{P}}$ are given in (6.19). To obtain the desired result, we substitute it into the first equation in (6.21) and define $T_h\,\boldsymbol{\lambda}_h := \tilde{\boldsymbol{u}}_h^k$. □

Remark 6.9. Clearly, assumption (6.6) also holds if one adopts piecewise polynomial spaces constructed on top of a sub-mesh in place of (6.15) (under the same constraint between the degrees l and k). Such an option becomes attractive in the case of highly heterogenous material problems since the basis functions naturally upscale the multi-scale features of the media into the numerical solution. This viewpoint makes the MHM method a member of the multi-scale finite element family [25].

6.2. Error estimates

Now, let us turn to the second-level discretization from the viewpoint of convergence. We shall demonstrate results assuming $\tilde{\mathbf{V}}_h \subset \tilde{\mathbf{V}}$ is given by (6.15) (where the order of the approximating polynomial is k). From these choices it is well known that, assuming $T\,\boldsymbol{\mu}+\hat{T}\,\boldsymbol{q} \in \boldsymbol{H}^{m+1}(\Omega)$, $1 \leq m \leq k$, the Galerkin method adopted locally to define the global operators T_h and $\hat{T}_h$ delivers the following interpolation errors

$$\begin{aligned} \|(T-T_h)\,\boldsymbol{\mu} + (\hat{T}-\hat{T}_h)\,\boldsymbol{q}\|_{\mathbf{V}} &\leq C\,h^m\,\|T\,\boldsymbol{\mu}+\hat{T}\,\boldsymbol{q}\|_{m+1},\\ \|(T-T_h)\,\boldsymbol{\mu} + (\hat{T}-\hat{T}_h)\,\boldsymbol{q}\|_{0,\Omega} &\leq C\,h^{m+1}\,\|T\,\boldsymbol{\mu}+\hat{T}\,\boldsymbol{q}\|_{m+1}. \end{aligned} \tag{6.22}$$

As such, using the estimates above, we are ready to present the convergence of the MHM method with a two-level discretization which adopts space (6.15).

Lemma 6.10. *Let* $(\boldsymbol{\lambda}, \boldsymbol{u}^{\rm rm}) \in \boldsymbol{\Lambda}\times\mathbf{V}_{\rm rm}$ *and* $(\boldsymbol{\lambda}_l, \boldsymbol{u}_h^{\rm rm}) \in \boldsymbol{\Lambda}_l\times\mathbf{V}_{\rm rm}$ *be the solutions of* (2.6) *and* (6.4)*, respectively. Assuming* $\boldsymbol{\Lambda}_l$ *and* $\tilde{\mathbf{V}}_h$ *satisfy the conditions in* (6.6) *and* $\boldsymbol{u} \in \boldsymbol{H}^{m+1}(\Omega)$*, then for* $1 \leq m \leq l+1 \leq k$ *with* $l \geq 1$*, it holds that there exist* C *such that*

$$\|(\boldsymbol{\lambda}-\boldsymbol{\lambda}_l, \boldsymbol{u}^{\rm rm}-\boldsymbol{u}_h^{\rm rm})\|_{\boldsymbol{\Lambda}\times\mathbf{V}} \;\leq\; C\,h^m\,\|\boldsymbol{u}\|_{m+1} \tag{6.23}$$

$$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{\mathbf{V}} \leq C\,h^m\,\|\boldsymbol{u}\|_{m+1}. \tag{6.24}$$

Furthermore, if problem (1.1) *has smoothing properties, there exists* C *such that*

$$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{0,\Omega} \;\leq\; C\,h^{m+1}\,\|\boldsymbol{u}\|_{m+1}. \tag{6.25}$$

Proof. From Lemma 6.3, the Cauchy−Schwarz and Poincaré inequalities we get

$$\begin{aligned} \|(\boldsymbol{\lambda}-\boldsymbol{\lambda}_l, \boldsymbol{u}^{\rm rm}-\boldsymbol{u}_h^{\rm rm})\|_{\boldsymbol{\Lambda}\times\mathbf{V}} &\leq C\left(\inf_{\boldsymbol{\mu}_l\in\boldsymbol{\Lambda}_l}\|\boldsymbol{\lambda}-\boldsymbol{\mu}_l\|_{\boldsymbol{\Lambda}} + \|(T-T_h)\boldsymbol{\lambda}+(\hat{T}-\hat{T}_h)\boldsymbol{f}\|_{\mathbf{V}}\right)\\ &\leq C\,h^m\left(\|\boldsymbol{u}\|_{m+1}+\|T\boldsymbol{\lambda}+\hat{T}\boldsymbol{f}\|_{m+1}\right)\\ &\leq C\,h^m\,\|\boldsymbol{u}\|_{m+1}, \end{aligned}$$

where we used (6.22) and $\|T\boldsymbol{\lambda}+\hat{T}\boldsymbol{f}\|_{m+1} \leq \|\boldsymbol{u}\|_{m+1}$ since $\boldsymbol{u}^{\rm rm} \in \mathbf{V}_{\rm rm}$ is $\mathbf{V}$-orthogonal to $T\boldsymbol{\lambda}+\hat{T}\boldsymbol{f}$. The last two results follow analogously from Corollary 6.4. □

Remark 6.11. Note that the proposed two-level method (6.4) preserves some of the main features of its one-level counterpart, like the well-posedness, the super-convergence of the error and local equilibrium. For the latter, for all $K \in \mathcal{T}_h$, it holds from (6.4) that the approximate two-level traction $\boldsymbol{\lambda}_h$ satisfies

$$\int_{\partial K}\boldsymbol{\lambda}_h\,\boldsymbol{v}^{\rm rm} = \int_K \boldsymbol{f}\,\boldsymbol{v}^{\rm rm} \quad \text{for all } \boldsymbol{v}^{\rm rm} \in \mathbf{V}_{\rm rm}.$$

Strict conformity in $H(\mathrm{div};\Omega)$ and optimal error estimates to the stress σ_h (post-processed from displacement (6.14)) in the $H(\mathrm{div};\Omega)$ norm should be expected only if a mixed finite element method is adopted to approximate local problems. This will be addressed in the future.

7. Conclusion

We proposed a new family of $H(\mathrm{div};\Omega)$ conforming and stable finite elements for the linear elasticity equation. The simplest member of this family has, per element, 12 degrees of freedom in 2D and 30 in 3D in total, respectively. Also interesting is that the approximate stress tensor preserves the local equilibrium property as well as the strong symmetry using a simple post-processing of the primal variable. Our analysis provided super-convergent *a priori* error estimates in natural norms and a face-based *a posteriori* estimator. For the latter, we proved that reliability and efficiency hold.

Element-wise elasticity problems with prescribed traction on faces drove basis functions. First, the numerical analysis was done using existence of unique solutions of these problems. Next, we used this "optimal" context to highlight how a second-level discretization influences well-posedness and consistency of the method. In particular, a two-level analysis showed the conditions under which the two-level MHM methods preserve super-convergent error estimates. As a result, the solutions presented high-order precision even with simple, one-element discretizations at the second level.

It is worth mentioning that the computation of completely independent local problems for the basis functions is embedded in the upscaling procedure, so their solutions may be naturally obtained within parallel computational environments. This is particularly attractive when precisely handling large elasticity problems with heterogeneous coefficients on coarse meshes.

The important question of robustness of the MHM method for the incompressible limit case was left open. We observed that such a feature is tightly attached to the choice of the second-level numerical method. For instance, basis functions obtained from mixed finite element methods are expected to yield locking-free two-level MHM methods.

Appendix A.

Throughout this work, we need some auxiliary results such as the optimal local Poincaré inequality (on convex domains) [27] and the following second Korn's inequality: for $\tilde{\boldsymbol{v}} \in \tilde{\mathbf{V}}$ it holds (see [14] for instance)

$$\|\tilde{\boldsymbol{v}}\|_{0,K} \le \frac{h_K}{\pi}\|\nabla\tilde{\boldsymbol{v}}\|_{0,K} \quad \text{and} \quad \|\nabla\tilde{\boldsymbol{v}}\|_{0,K} \le C_{\mathrm{korn}}\,\|\,\mathcal{E}(\tilde{\boldsymbol{v}})\|_{0,K}, \tag{A.1}$$

where C_{korn} is a positive constant independent of h. Combining both previous inequalities, we arrive at

$$\|\tilde{\boldsymbol{v}}\|_{0,K} \le C_{\mathrm{korn}}\,\frac{h_K}{\pi}\|\,\mathcal{E}(\tilde{\boldsymbol{v}})\|_{0,K}. \tag{A.2}$$

Also, combining the classical Korn's inequality [13] with a local trace inequality it follows that: Given $\boldsymbol{v} \in \boldsymbol{H}^1(K)$, we obtain

$$\|\boldsymbol{v}\|_{0,\partial K} \le C\left(\frac{1}{h_K}\|\boldsymbol{v}\|_{0,K}^2 + h_K\|\,\mathcal{E}(\boldsymbol{v})\|_{0,K}^2\right)^{1/2}, \tag{A.3}$$

and using (A.2) and (A.3) it holds

$$\|\tilde{\boldsymbol{v}}\|_{0,\partial K} \le C\,h_K^{1/2}\|\,\mathcal{E}(\tilde{\boldsymbol{v}})\|_{0,K}, \tag{A.4}$$

for $\tilde{\boldsymbol{v}} \in \tilde{\mathbf{V}}$. We shall make extensive use of the following value

$$\kappa := \frac{c_{\max}}{c_{\min}}\cdot \tag{A.5}$$

Next, we prove some of the auxiliary results which are used in previous sections.

Lemma A.1. *Let $\boldsymbol{\mu} \in \boldsymbol{\Lambda}$ and define $T : \boldsymbol{\Lambda} \to \tilde{\mathbf{V}}$ as in (2.4),* i.e.*, for each $K \in \mathcal{T}_h$, $T\boldsymbol{\mu} \in \tilde{\mathbf{V}}$ is the unique solution of*

$$(\mathcal{A}\,\mathcal{E}(T\,\boldsymbol{\mu}),\,\mathcal{E}(\boldsymbol{w}))_K = -\,(\boldsymbol{\mu},\,\boldsymbol{w})_{\partial K} \quad \text{for all } \boldsymbol{w} \in \tilde{\mathbf{V}}.$$

Then, T is a bounded linear operator satisfying the following bounds

$$\|\mathcal{A}\,\mathcal{E}(T\,\boldsymbol{\mu})\|_{\mathrm{div}} \le \max\left\{\frac{\kappa}{\pi}\sqrt{2C_{\mathrm{korn}}^2+\pi^2},\sqrt{2}\right\}\,\|\boldsymbol{\mu}\|_{\boldsymbol{\Lambda}}, \tag{A.6}$$

$$\|T\,\boldsymbol{\mu}\|_V \le \frac{(C_{\mathrm{korn}}^2+\pi^2)\kappa}{\pi^2 c_{\min}}\,\|\boldsymbol{\mu}\|_{\boldsymbol{\Lambda}}. \tag{A.7}$$

Proof. First, the problem has a unique solution if and only if $-\mathbf{div}\mathcal{A}\,\mathcal{E}(T\,\boldsymbol{\mu}) = \mathbf{R}^{\mu} \in \mathbf{V}_{\mathrm{rm}}$ from (2.9), where

$$(\mathbf{R}^{\mu}, \boldsymbol{v}^{\mathrm{rm}})_{\mathcal{T}_h} = \sum_{K\in\mathcal{T}_h} (\boldsymbol{\mu}, \boldsymbol{v}^{\mathrm{rm}})_{\partial K}.$$

By definition (4.6) of $\|.\|_{\mathrm{div}}$, integrating by parts and the fact $\mathbf{div}(\mathcal{A}\,\mathcal{E}(T\,\boldsymbol{\mu}))\,|_K \in \mathbf{V}_{\mathrm{rm}}(K)$ we arrive at

$$\|\mathcal{A}\,\mathcal{E}(T\,\boldsymbol{\mu})\|_{\mathrm{div}}^2 \le \sum_{K\in\mathcal{T}_h} (\boldsymbol{\mu}, -c_{\max}\,T\,\boldsymbol{\mu} + d_{\Omega}^2\,\mathbf{div}(\mathcal{A}\,\mathcal{E}(T\,\boldsymbol{\mu})))_{\partial K},$$

where we used (1.2). Therefore, since $-c_{\max}\,T\,\boldsymbol{\mu}+d_{\Omega}^2\,\mathbf{div}(\mathcal{A}\,\mathcal{E}(T\,\boldsymbol{\mu})) \in \mathbf{V}$, it follows by the local inequality (A.2) and the fact $\mathbf{div}(\mathcal{A}\,\mathcal{E}(T\,\boldsymbol{\mu}))|_K \in \mathbf{V}_{\mathrm{rm}}(K)$,

$$\begin{aligned}
\|\mathcal{A}\,\mathcal{E}(T\,\boldsymbol{\mu})\|_{\mathrm{div}}^2 &\le \sup_{\boldsymbol{v}\in\mathbf{V}} \frac{b(\boldsymbol{\mu},\boldsymbol{v})}{\|\boldsymbol{v}\|_{\mathbf{V}}}\left[\sum_{K\in\mathcal{T}_h}\left(d_{\Omega}^{-2}\|c_{\max}T\,\boldsymbol{\mu}+d_{\Omega}^2\mathbf{div}(\mathcal{A}\,\mathcal{E}(T\,\boldsymbol{\mu}))\|_{0,K}^2 + \|c_{\max}\,\mathcal{E}(T\,\boldsymbol{\mu})\|_{0,K}^2\right)\right]^{1/2}\\
&\le \sup_{\boldsymbol{v}\in\mathbf{V}} \frac{b(\boldsymbol{\mu},\boldsymbol{v})}{\|\boldsymbol{v}\|_{V}}\left[\sum_{K\in\mathcal{T}_h}\left(2\,d_{\Omega}^{-2}\,c_{\max}^2\|T\,\boldsymbol{\mu}\|_{0,K}^2 + 2\,d_{\Omega}^2\|\mathbf{div}(\mathcal{A}\,\mathcal{E}(T\,\boldsymbol{\mu}))\|_{0,K}^2 + c_{\max}^2\|\,\mathcal{E}(T\,\boldsymbol{\mu})\|_{0,K}^2\right)\right]^{1/2}\\
&\le \sup_{\boldsymbol{v}\in\mathbf{V}} \frac{b(\boldsymbol{\mu},\boldsymbol{v})}{\|\boldsymbol{v}\|_{\mathbf{V}}}\left[\sum_{K\in\mathcal{T}_h}\left(\frac{(2C_{\mathrm{korn}}^2+\pi^2)\,c_{\max}^2}{\pi^2\,c_{\min}^2}\|\mathcal{A}\,\mathcal{E}(T\,\boldsymbol{\mu})\|_{0,K}^2 + 2\,d_{\Omega}^2\|\mathbf{div}(\mathcal{A}\,\mathcal{E}(T\,\boldsymbol{\mu}))\|_{0,K}^2\right)\right]^{1/2}.
\end{aligned}$$

Then, using definition of κ in (A.5), we get

$$\|\mathcal{A}\,\mathcal{E}(T\,\boldsymbol{\mu})\|_{\mathrm{div}} \le C_{\mathrm{korn}}\,\sqrt{2}\,\kappa\,\sup_{\boldsymbol{v}\in V}\frac{b(\boldsymbol{\mu},\boldsymbol{v})}{\|\boldsymbol{v}\|_V}. \tag{A.8}$$

Now, choose arbitrary $\boldsymbol{v}\in\mathbf{V}$, and suppose that $\sigma \in H(\mathrm{div};\Omega)$ satisfies the property $\sigma\,\boldsymbol{n}^K\,|_{\partial K} = \boldsymbol{\mu}$ for $\boldsymbol{\mu}\in\boldsymbol{\Lambda}$. It follows by Green's theorem and the Cauchy−Schwarz's inequality that

$$\begin{aligned}
\sum_{K\in\mathcal{T}_h}(\boldsymbol{\mu},\boldsymbol{v})_{\partial K} &= \sum_{K\in\mathcal{T}_h}(\sigma\,\boldsymbol{n}^K,\boldsymbol{v})_{\partial K}\\
&= \sum_{K\in\mathcal{T}_h}\left[(\mathbf{div}\,\sigma,\boldsymbol{v})_K + (\sigma,\mathcal{E}(\boldsymbol{v}))_K\right]\\
&\le \sum_{K\in\mathcal{T}_h}\left[d_{\Omega}\,\|\mathbf{div}\,\sigma\|_{0,K}\,d_{\Omega}^{-1}\,\|\boldsymbol{v}\|_{0,K} + \|\sigma\|_{0,K}\|\,\mathcal{E}(\boldsymbol{v})\|_{0,K}\right]\\
&\le \|\sigma\|_{\mathrm{div}}\|\boldsymbol{v}\|_V.
\end{aligned}$$

Then, by definition of supremum, it follows that

$$\sup_{\boldsymbol{v}\in\mathbf{V}}\frac{b(\boldsymbol{\mu},\boldsymbol{v})}{\|\boldsymbol{v}\|_{\mathbf{V}}} = \sup_{\boldsymbol{v}\in\mathbf{V}}\frac{(\boldsymbol{\mu},\boldsymbol{v})_{\partial\mathcal{T}_h}}{\|\boldsymbol{v}\|_{\mathbf{V}}} \le \|\sigma\|_{\mathrm{div}}.$$

Since σ was arbitrarily taken, the above inequality and definition of infimum imply

$$\sup_{\boldsymbol{v}\in\mathbf{V}} \frac{b(\boldsymbol{\mu},\boldsymbol{v})}{\|\boldsymbol{v}\|_{\mathbf{V}}} \leq \|\boldsymbol{\mu}\|_{\boldsymbol{\Lambda}}, \tag{A.9}$$

and result (A.6) follows immediately replacing the result above in (A.8). The bound (A.7) follows from Korn's inequality (A.2). □

Lemma A.2. *Let* $\boldsymbol{q} \in \boldsymbol{L}^2(\Omega)$ *and define* $\hat{T} : \boldsymbol{L}^2(\Omega) \to \mathbf{V}$ *as in* (2.5), i.e., *for each* $K \in \mathcal{T}_h$, $\hat{T}\,\boldsymbol{q} \in \tilde{\mathbf{V}}$ *is the unique solution of*

$$(\mathcal{A}\,\mathcal{E}(\hat{T}\,\boldsymbol{q}),\,\mathcal{E}(\boldsymbol{w}))_K = (\boldsymbol{q},\boldsymbol{w})_K \quad \textit{for all } \boldsymbol{w} \in \tilde{\mathbf{V}}. \tag{A.10}$$

Then, $\hat{T}$ *is a bounded linear operator satisfying the following bounds*

$$\|\mathcal{A}\,\mathcal{E}(\hat{T})\,\boldsymbol{q}\|_{\mathrm{div}} \leq \max\{C_{\mathrm{korn}},1\}\,\sqrt{2}\,d_\Omega\,\kappa\,\|\boldsymbol{q}-\Pi\,\boldsymbol{q}\|_{0,\Omega}, \tag{A.11}$$

$$\|\hat{T}\,\boldsymbol{q}\|_V \leq \max\{C_{\mathrm{korn}},1\}\,\frac{2\,d_\Omega\,\kappa}{c_{\min}}\,\|\boldsymbol{q}-\Pi\,\boldsymbol{q}\|_{0,\Omega}. \tag{A.12}$$

Proof. First, we establish (A.11). Note that (1.2), the fact $\hat{T}\,\boldsymbol{q} \in \tilde{\mathbf{V}}$, and the Cauchy−Schwarz and the local inequality (A.2), and $h_K \leq d_\Omega$ imply

$$\begin{aligned}
\|\mathcal{A}\,\mathcal{E}(\hat{T}\,\boldsymbol{q})\|^2_{0,K} &\leq \|\mathcal{A}\|(\mathcal{A}\,\mathcal{E}(\hat{T}\,\boldsymbol{q}),\mathcal{E}(\hat{T}\,\boldsymbol{q}))_K\\
&\leq c_{\max}\,(\boldsymbol{q},\hat{T}\,\boldsymbol{q})_K\\
&= c_{\max}\,(\boldsymbol{q}-\Pi_K\,\boldsymbol{q},\hat{T}\,\boldsymbol{q})_K\\
&\leq c_{\max}\,\|\boldsymbol{q}-\Pi_K\,\boldsymbol{q}\|_{0,K}\|\hat{T}\,\boldsymbol{q}\|_{0,K}\\
&\leq C_{\mathrm{korn}}\frac{\kappa}{\pi}\,d_\Omega\,\|\boldsymbol{q}-\Pi_K\,\boldsymbol{q}\|_{0,K}\|\mathcal{A}\,\mathcal{E}(\hat{T}\,\boldsymbol{q})\|_{0,K}.
\end{aligned}$$

Therefore, from (2.10) and by definition (4.6) of $\|.\|_{\mathrm{div}}$, and observing that $1 \leq \kappa$, we get

$$\begin{aligned}
\|\mathcal{A}\,\mathcal{E}(\hat{T}\,\boldsymbol{q})\|^2_{\mathrm{div}} &= \sum_{K\in\mathcal{T}_h}\left[\|\mathcal{A}\,\mathcal{E}(\hat{T}\,\boldsymbol{q})\|^2_{0,K} + d^2_\Omega\,\|\boldsymbol{q}-\Pi_K\,\boldsymbol{q}\|^2_{0,K}\right]\\
&\leq 2\,d^2_\Omega\,\max\left\{\left(C_{\mathrm{korn}}\frac{\kappa}{\pi}\right)^2,1\right\}\,\|\boldsymbol{q}-\Pi\,\boldsymbol{q}\|^2_{0,\Omega},
\end{aligned}$$

from which the bound (A.11) follows immediately. The bound (A.12) follows using the local Poincaré's inequality (A.1) and the result (A.11). □

Lemma A.3. *Suppose* $\boldsymbol{\mu} \in \boldsymbol{\Lambda}$*. It follows that*

$$\frac{\sqrt{2}}{2\,C_{\mathrm{korn}}}\,\|\boldsymbol{\mu}\|_{\boldsymbol{\Lambda}} \leq \sup_{\boldsymbol{v}\in\mathbf{V}}\frac{b(\boldsymbol{\mu},\boldsymbol{v})}{\|\boldsymbol{v}\|_{\mathbf{V}}} \leq \|\boldsymbol{\mu}\|_{\boldsymbol{\Lambda}}.$$

Proof. Choose arbitrary $\boldsymbol{\mu} \in \boldsymbol{\Lambda}$. The left-hand side bound follows from equation (A.8) (with $\mathcal{A}$ as the identity matrix). The right-hand bound is inequality (A.9) in the proof of Lemma A.1. □

References

[1] S. Adams and B. Cockburn, A mixed finite element method for elasticity in three dimensions. *J. Sci. Comput.* **25** (2005) 515–521.

[2] M. Amara and J.M. Thomas, Equilibrium finite elements for the linear elastic problem. *Numer. Math.* **33** (1979) 367–383.

[3] R. Araya, C. Harder, D. Paredes and F. Valentin, Multiscale hybrid-mixed method. *SIAM J. Numer. Anal.* **51** (2013) 3505–3531.

[4] T. Arbogast and K. Boyd, Subgrid upscaling and mixed multiscale finite elements. *SIAM J. Numer. Anal.* **44** (2006) 1150–1171.

[5] D.N. Arnold, G. Awanou and R. Winther, Finite elements for symmetric tensors in three dimensions. *Math. Comput.* **77** (2008) 1229–1251.

[6] D.N. Arnold and R. Winther, Mixed finite elements for elasticity. *Numer. Math.* **92** (2002) 401–419.

[7] D.N. Arnold, F. Brezzi and J. Douglas, Peers: a new mixed finite element for plane elasticity. *Japan J. Appl. Math.* **1** (1984) 347–367.

[8] D.N. Arnold, J.J. Douglas and C.P. Gupta, A family of higher order mixed finite element methods for plane elasticity. *Numer. Math.* **45** (1984) 1–22.

[9] D.N. Arnold, G. Awanou and R. Winther, Nonconforming tetrahedral mixed finite elements for elasticity. *Math. Models Methods Appl. Sci.* **23** (2014) 783–796.

[10] I. Babuska and E. Osborn, Generalized finite element methods: Their performance and their relation to mixed methods. *SIAM J. Numer. Anal.* **20** (1983) 510–536.

[11] L. Beirão da Veiga, F. Brezzi and L.D. Marini, Virtual elements for linear elasticity problems. *SIAM J. Numer. Anal.* **51** (2013) 794–812.

[12] J. Bramwell, L. Demkowicz, J. Gopalakrishnan and W. Qiu, A locking-free hp dpg method for linear elasticity with symmetric stresses. *Numer. Math.* **122** (2012) 671–707.

[13] S.C. Brenner, Korn's inequalities for piecewise H^1 vector fields. *Math. Comput.* **73** (2004) 1067–1087.

[14] S.C. Brenner and L.R. Scott, The Mathematical Foundations of the Finite Element Methods. Springer (2002).

[15] F. Brezzi and M. Fortin, Mixed and hybrid finite element methods. Vol. 15 of *Springer Ser. Comput. Math.* Springer-Verlag, Berlin, New-York (1991).

[16] Z. Chen and T. Hou, A mixed multiscale finite element method for elliptic problems with oscillating coefficients. *Math. Comput.* **72** (2002) 541–576.

[17] B. Cockburn and K. Chi, Superconvergent hdg methods for linear elasticity with weakly symmetric stresses. *IMA J. Numer. Anal.* (2012) 1–24.

[18] L. Demkowicz and J. Gopalakrishnan, A primal dpg method without a first order reformulation. *Comput. Math. Appl.* **66** (2013) 1058–1064.

[19] A. Ern and J.-L. Guermond, Theory and practice of finite elements. Springer-Verlag, Berlin, New-York (2004).

[20] J. Gopalakrishnan and J. Guzmán, Symmetric nonconforming mixed finite elements for linear elasticity. SIAM J. Numer. Anal. **49** (2011) 1504–1520.

[21] J. Gopalakrishnan and W. Qiu, An analysis of the practical dpg method. *Math. Comput.* **83** (2014) 537–552.

[22] J. Guzmán and M. Neilan, Symmetric and conforming mixed finite elements for plane elasticity using rational bubble functions. *Numer. Math.* **126** (2014) 153–171.

[23] C. Harder, D. Paredes and F. Valentin, A family of multiscale hybrid-mixed finite element methods for the Darcy equation with rough coefficients. *J. Comput. Phys.* **245** (2013) 107–130.

[24] C. Harder, D. Paredes and F. Valentin, On a multiscale hybrid-mixed method for advective-reactive dominated problems with heterogenous coefficients. *SIAM Multiscale Model. Simul.* **13** (2015) 491–518.

[25] T.Y. Hou and X. Wu, A multiscale finite element method for elliptic problems in composite materials and porous media. *J. Comput. Phys.* **134** (1997) 169–189.

[26] J. Hu, A new family of efficient conforming mixed finite elements on both rectangular and cuboid meshes for linear elasticity in the symmetric formulation. Preprint arXiv:1311.4718v3 [math.NA] (2015).

[27] L.E. Payne and H.F. Weinberger, An optimal Poincaré inequality for convex domains. *Arch. Ration. Mech. Anal.* **5** (1960) 286–292.

[28] T. Pian and P. Tong, Basis of finite element methods for solid continua, *Int. J. Numer. Methods Engrg.* **1** (1969) 3–28.

[29] W. Qiu and K. Shi, An hdg method for linear elasticity with strong symmetric stresses. Preprint arXiv:1312.1407v2 [math.NA] (2014).

[30] P. Raviart and J. Thomas, A mixed finite element method for 2nd order elliptic problems, Mathematical aspect of finite element methods, No. 606 in *Lect. Notes Math.* Springer-Verlag, New York (1977) 292–315.

[31] P. Raviart and J. Thomas, Primal hybrid finite element methods for 2nd order elliptic equations. *Math. Comput.* **31** (1977) 391–413.

[32] S. Soon, B. Cockburn and H. Stolarski, A hybridizable discontinuous galerkin method for linear elasticity. *Int. J. Numer. Methods Engrg.* **80** (2009) 1058–1092.

[33] R. Stenberg, On the construction of optimal mixed finite element methods for the linear elasticity problem. *Numer. Math.* **48** (1986) 447–462.

[34] a. Toselli and O. Widlund, Domain decomposition methods-algorithms and theory. Vol. 34 of *Springer Ser. Comput. Math.* Springer-Verlag, Berlin (2005).