The MHM Method for Elasticity on Polytopal Meshes

Antônio Tadeu A. Gomes, Weslley S. Pereira, Frédéric Valentin

THE MHM METHOD FOR ELASTICITY ON POLYTOPAL MESHES

ANTÔNIO TADEU A. GOMES*, WESLLEY S. PEREIRA†, AND FRÉDÉRIC VALENTIN‡

Abstract. The multiscale hybrid-mixed (MHM) method consists of a multi-level strategy to approximate the solution of boundary value problems with heterogeneous coefficients. In this context, we propose a new family of finite elements for the linear elasticity equation defined on coarse polytopal partitions of the domain. The finite elements rely on face degrees of freedom associated with multiscale bases obtained from local Neumann problems with polynomial interpolations on faces. We establish sufficient conditions on the fine-scale interpolations such that the MHM method is well-posed. Also, discrete traction stays in local equilibrium with external forces. We show by means of a multi-level analysis that the MHM method achieves optimal convergence under local regularity conditions without refining the coarse partition. The upshot is that the Poincaré and Korn's inequalities do not degenerate, and then convergence arises on general meshes. We employ two- and three-dimensional numerical tests to assess theoretical results and to verify the robustness of the method through a multi-layer media case. Also, we address computational aspects of the underlying parallel algorithm associated with different configurations of the MHM method; our aim is to find the best compromise between execution time and memory allocation to achieve a given error threshold.

1. Introduction. The modeling of elastic phenomena often involves linear differential equations with heterogeneous coefficients defined on intricate domains. For instance, in geological modeling, data include complex geological layers with structures over a wide range of length scales and, sometimes, the surface topography. In this case, details matter; the numerical solution may strongly differ if the mesh does not account for the multiple scales properly. In standard finite element methods, this constraint demands very fine meshes, which enlarges the computational cost. Furthermore, numerical methods ideally should preserve essential physical properties, such as symmetry and conformity of the stress field, and local equilibrium of the stress field with external forces, which require high-order polynomial interpolations [6] and, then, additional computational resources. The common approach has been to relax symmetry, local equilibrium, or conformity. These options have been actively pursued, with works dating from the eighties [5, 43] to the present day [22, 7].

Finite element methods on polytopal partitions for elastic models have undertaken significant development in the last years [16, 39, 45, 37]. They bring flexibility to deal with complex geometries avoiding unnecessary spread in the mesh refinement. When associated with domain decomposition approaches, polytopal finite elements become an attractive and compelling option to decrease computational costs.

In this context, the multiscale finite element methods, started with the seminal work [8], can be interpreted as members of the domain decomposition approach; examples include [29, 30, 4, 1, 32]. These methods propose upscaling numerical strategies to solve problems using coarse partitions. They use localized multiscale basis functions to recover structures of the solution lost by the unresolved fine scales. From a computational viewpoint, these methods fit well in massively parallel computer systems

*Department of Computational and Mathematical Methods, National Laboratory for Scientific Computing - LNCC, Av. Getúlio Vargas, 333, 25651-070 Petrópolis - RJ, Brazil (atagomes@lncc.br)

†Department of Computational and Mathematical Methods, National Laboratory for Scientific Computing - LNCC, Av. Getúlio Vargas, 333, 25651-070 Petrópolis - RJ, Brazil and Federal University of Juiz de Fora, R. José Lourenćo Kelmer, s/n, 36036-900 Juiz de Fora - MG, Brazil (wesleyp@lncc.br,weslley.pereira@ice.ufjf.br)

‡Department of Computational and Mathematical Methods, National Laboratory for Scientific Computing - LNCC, Av. Getúlio Vargas, 333, 25651-070 Petrópolis - RJ, Brazil and Nachos Project-Team, Inria Sophia Antipolis - Méditerranée, 2004 Route des Lucioles, 06902 Valbonne, France (valentin@lncc.br, frederic.valentin@inria.fr)

because they allow for the multiscale basis functions to be computed in a decoupled fashion. Nevertheless, the global part of these methods remains a computational bottleneck–especially in realistic three-dimensional problems—because of its coupled form. One of the strategies to decrease the number of degrees of freedom in the global part of the multiscale methods is to combine the flexibility of polytopal meshes with non-conforming, skeleton-based methods. These methods employ degrees of freedom on the faces of the mesh skeleton at the global level and can cope with jumping coefficients without aligning faces at the local level. Also, discontinuous interpolations on faces play an essential role in maintaining accuracy in such a crossing-face interface scenario.

In this work, we propose a new family of polytopal finite elements for the Multiscale Hybrid-Mixed (MHM for short) method applied to the two- and three-dimensional linear elasticity model. It extends the construction and the analysis of the MHM method proposed in [28] by following the recent strategy given in [9] for the two-dimensional Poisson equation. In addition to the extension of [28] to more general meshes, the contributions of the present work are the following:

- it allows discontinuous polynomial interpolations on faces to account for crossing-face interfaces. This feature is also new for simplicial elements in the context of the method proposed in [28];
- it accounts for face partitioning and fine-scale meshes in the analysis of well-posedness and convergence of the method. Such a multi-level analysis is absent in [28]. Here, we extend the analysis of [9] to the linear elasticity model in two- and three-dimensions thanks to a new Fortin operator;
- it establishes optimal error estimates in natural norms under local regularity assumption, which are robust regarding the Poincaré and Korn constants. The upshot is that convergence does not degenerate on general meshes. It corresponds to a new aspect of the MHM method for the elasticity model not presented in [28], even for simplicial meshes;
- it proves optimal convergence in the L^2 norms for non-convex polytopal elements under a mild condition on the shape of elements;
- it validates the MHM method numerically for two- and three-dimensional elasticity problems. Such validations are also new for the simplicial meshes in [28]. Numerical results also verify the accuracy of the solutions in a high-contrast problem with complex topography using non-aligned meshes;
- it highlights the computational performance of the underlying MHM algorithm on a parallel computing system, which complements [21].

The paper's outline is as follows: In Section 2, we present the linear elasticity model and a hybrid version of it. Then, we introduce the framework to address the decomposition of the exact solution on a general partition of the domain. We dedicate Section 3 to the MHM method and prove its well-posedness in Section 4. We present error estimates in Section 5. We dedicate Section 6 to the numerical validation and the assessment of computational performance. We present some concluding remarks in Section 7.

2. The model and its hybrid version.

2.1. The elasticity model. Let $\Omega \subset \mathbb{R}^d$, $d \in \{2, 3\}$, be an open and bounded domain with polygonal Lipschitz boundary $\partial\Omega$. Consider the elasticity problem of finding a displacement $\boldsymbol{u} : \Omega \to \mathbb{R}^d$ satisfying

$$\begin{cases} -\nabla\cdot\left(\mathbf{C}\,\underline{\boldsymbol{\varepsilon}}(\boldsymbol{u})\right) \,=\, \boldsymbol{f} & \text{in } \Omega\,, \\ \qquad\quad \boldsymbol{u} \,=\, \boldsymbol{g} & \text{on } \partial\Omega\,, \end{cases} \tag{2.1}$$

where $\boldsymbol{g} \in \boldsymbol{H}^{1/2}(\partial\Omega)$ and $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega)$ are given functions. The fourth-order stiffness tensor $\mathbf{C} = \mathbf{C}(\boldsymbol{x})$ may embed multiple geometrical scales, and satisfies the following symmetries $\mathsf{C}_{ijkl} = \mathsf{C}_{ijlk} = \mathsf{C}_{klij}$. We assume $\mathbf{C}$ is bounded and uniformly positive-definite, i.e., there exist positive constants c_{min} and c_{max} such that

$$c_{min}|\boldsymbol{\varepsilon}|^2 \leqslant (\mathbf{C}(\boldsymbol{x})\,\boldsymbol{\varepsilon}) : \boldsymbol{\varepsilon} \quad \text{and} \quad (\mathbf{C}(\boldsymbol{x})\,\boldsymbol{\tau}) : \boldsymbol{\varepsilon} \leqslant c_{max}|\boldsymbol{\tau}||\boldsymbol{\varepsilon}| \quad \text{for all } \boldsymbol{\tau},\, \boldsymbol{\varepsilon} \in \mathbb{R}^{d\times d}_{\text{sym}}\,, \tag{2.2}$$

for almost every $\boldsymbol{x} \in \Omega$. The space $\mathbb{R}^{d\times d}_{\text{sym}}$ stands for the set of symmetric $d\times d$ matrices, $d \in \mathbb{N}$, and $\underline{\boldsymbol{\varepsilon}}(\boldsymbol{u})$ for the infinitesimal strain tensor, i.e.,

$$\underline{\boldsymbol{\varepsilon}}(\boldsymbol{u}) := \frac{\nabla\boldsymbol{u} + (\nabla\boldsymbol{u})^T}{2} \quad \text{where} \quad \underline{\boldsymbol{\varepsilon}}(\boldsymbol{u})_{ij} := \frac{1}{2}\left(\partial_{x_j} u_i + \partial_{x_i} u_j\right)\,.$$

Above, and throughout, the indexes run from $1,\ldots,d$, even when not explicitly mentioned; the symbol "$\nabla\cdot$" denotes the row-wise divergence operator, and $\partial_{x_i} u_j$ stands for the i^{th} partial derivative of the j^{th} entry of $\boldsymbol{u}$.

We denote by $\boldsymbol{v}\cdot\boldsymbol{w}$ and $\boldsymbol{\tau} : \boldsymbol{\varepsilon}$ the usual inner products in spaces $\mathbb{R}^d$ and $\mathbb{R}^{d\times d}$, i.e.,

$$\boldsymbol{v}\cdot\boldsymbol{w} := \sum_{i=1}^{d} v_i\, w_i \quad \text{and} \quad \boldsymbol{\tau} : \boldsymbol{\varepsilon} := \sum_{i,j=1}^{d} \tau_{ij}\,\boldsymbol{\varepsilon}_{ij}\,,$$

and by $|\boldsymbol{v}|$ and $|\tau|$ the norms each of such inner products induce. Owing to such notation, the classical weak form of (2.1) is

$$\begin{aligned} \int_\Omega \mathbf{C}\,\underline{\boldsymbol{\varepsilon}}(\boldsymbol{u}) \,:\, \underline{\boldsymbol{\varepsilon}}(\boldsymbol{w})\,d\boldsymbol{x} \,&=\, \int_\Omega \boldsymbol{f}\cdot\boldsymbol{v}\,d\boldsymbol{x} \quad \text{for all } \boldsymbol{v} \in \boldsymbol{H}^1_0(\Omega)\,, \\ \boldsymbol{u} &= \boldsymbol{g} \quad \text{on } \partial\Omega\,. \end{aligned} \tag{2.3}$$

Under the previous assumptions, there exists a unique function $\boldsymbol{u} \in \boldsymbol{H}^1(\Omega)$ satisfying (2.3) (c.f. [20, Proposition 2.10]). As a consequence of (2.3), the stress tensor $\boldsymbol{\sigma} := \mathbf{C}\,\underline{\boldsymbol{\varepsilon}}(\boldsymbol{u})$ belongs to $H(\mathbf{div};\Omega;\mathbb{S})$, where $H(\mathbf{div};\Omega;\mathbb{S})$ is the space of symmetric tensors in $H(\mathbf{div};\Omega)$. Above, and hereafter, we adopt typical notation for Sobolev and Lebesgue spaces (see [20] for instance).

2.2. Broken spaces and some important inequalities. Let $\mathscr{P}$ be a collection of closed and bounded d-polytopes K, such that $\overline{\Omega} = \cup_{K\in\mathscr{P}} K$. Define

$\mathcal{H} := \max_{K\in\mathscr{P}} h_K$, where $h_D := \sup_{x,y\in D}|x-y|$ is the diameter of an arbitrary set $D \subset \mathbb{R}^n$, $n \in \mathbb{N}$. The radius of the largest inscribed ball in D reads ρ_D, and the shape regularity of D is denoted by $\sigma_D := \frac{h_D}{\rho_D}$. We also introduce $\partial\mathscr{P}$ as the collection of boundaries ∂K, and $\mathcal{E}$ as the set of the faces in $\mathscr{P}$, that is

$$\partial\mathscr{P} := \{\partial K \,:\, K \in \mathscr{P}\}\,,$$
$$\mathcal{E} := \{E = K\cap K' \text{ or } K\cap\partial\Omega \,:\, K, K' \in \mathscr{P}, \text{ and } E \text{ is a } (d-1)\text{-polytope }\}\,.$$

The set of internal faces is $\mathcal{E}_0$, i.e., $\mathcal{E}_0 := \{E \in \mathcal{E} \,:\, E \not\subset \partial\Omega\}$. Associated to the partition $\mathscr{P}$, we define the function spaces

$$\mathbf{V} := \boldsymbol{H}^1(\mathscr{P}) \quad \text{with norm} \quad \|\boldsymbol{v}\|^2_{\mathbf{V}} := \sum_{K\in\mathscr{P}} \left(h_\Omega^{-2}\|\boldsymbol{v}\|^2_{0,K} + \|\underline{\boldsymbol{\varepsilon}}(\boldsymbol{v})\|^2_{0,K}\right),$$

$$\boldsymbol{\Lambda} := \{\boldsymbol{\tau}\,\boldsymbol{n}^K|_{\partial K} \,:\, \boldsymbol{\tau} \in H(\mathbf{div};\Omega;\mathbb{S}) \quad \text{for all } K\in\mathscr{P}\} \quad \text{with norm}$$

$$\|\boldsymbol{\mu}\|_{\boldsymbol{\Lambda}} := \inf_{\substack{\boldsymbol{\tau}\in H(\mathbf{div};\Omega;\mathbb{S})\\ \boldsymbol{\tau}\,\boldsymbol{n}^K=\boldsymbol{\mu} \text{ on } \partial K,\, K\in\mathscr{P}}} \|\boldsymbol{\tau}\|_{\mathrm{div}}\,, \tag{2.4}$$

$$\text{where} \quad \|\boldsymbol{\tau}\|_{\mathrm{div}} := \left(\sum_{K\in\mathscr{P}} \left(\|\boldsymbol{\tau}\|^2_{0,K} + h_\Omega^2\|\nabla\cdot\boldsymbol{\tau}\|^2_{0,K}\right)\right)^{\frac{1}{2}}.$$

Above and hereafter, $(\cdot,\cdot)_D$ stands for the $L^2(D)$ inner product and $\|\cdot\|_{0,D}$ denotes its respective norm (we do not make a distinction between vector-valued and scalar-valued functions). The usual norm and semi-norm in $H^m(D)$, $m\in\mathbb{N}$, is denoted by $\|\cdot\|_{m,D}$ and $|\cdot|_{m,D}$, respectively.

Also, we shall use intensively the local space of rigid body modes

$$\mathbf{V}_{\mathrm{rm}}(K) := \left\{\boldsymbol{v}^{\mathrm{rm}} \in \boldsymbol{H}^1(K) \,:\, \underline{\boldsymbol{\varepsilon}}(\boldsymbol{v}^{\mathrm{rm}}) = \mathbf{0}\right\}, \tag{2.5}$$

and its $\boldsymbol{L}^2(\Omega)$-orthogonal complement

$$\widetilde{\mathbf{V}}(K) := \left\{\boldsymbol{v} \in \boldsymbol{H}^1(K) \,:\, \int_K \boldsymbol{v}\cdot\boldsymbol{v}^{\mathrm{rm}}\,d\boldsymbol{x} = 0,\ \forall \boldsymbol{v}^{\mathrm{rm}} \in \mathbf{V}_{\mathrm{rm}}(K)\right\}. \tag{2.6}$$

Naturally, their global space counterparts are given and denoted by

$$\mathbf{V}_{\mathrm{rm}} := \{\boldsymbol{v}\in\mathbf{V} \,:\, \boldsymbol{v}\,|_K \in \mathbf{V}_{\mathrm{rm}}(K),\ K\in\mathscr{P}\}\,,$$
$$\widetilde{\mathbf{V}} := \left\{\boldsymbol{v}\in\mathbf{V} \,:\, \boldsymbol{v}\,|_K \in \widetilde{\mathbf{V}}(K),\ K\in\mathscr{P}\right\}. \tag{2.7}$$

We denote by $(\cdot,\cdot)_{\mathbf{V}}$ the inner-product in $\mathbf{V}$ given by

$$(\boldsymbol{v},\boldsymbol{w})_{\mathbf{V}} := h_\Omega^{-2}(\boldsymbol{v},\boldsymbol{w})_{\mathscr{P}} + (\underline{\boldsymbol{\varepsilon}}(\boldsymbol{v}),\underline{\boldsymbol{\varepsilon}}(\boldsymbol{w}))_{\mathscr{P}} \quad \text{for all } \boldsymbol{v},\, \boldsymbol{w} \in \mathbf{V}\,,$$

where

$$(\boldsymbol{u},\boldsymbol{v})_{\mathscr{P}} := \sum_{K\in\mathscr{P}} \int_K \boldsymbol{u}\cdot\boldsymbol{v}\,d\boldsymbol{x} \quad \text{and} \quad \|\boldsymbol{v}\|_{0,\mathscr{P}}^2 := \sum_{K\in\mathscr{P}} \|\boldsymbol{v}\|_{0,K}^2\,.$$

Also, we define the following broken product

$$\langle\boldsymbol{\mu},\boldsymbol{v}\rangle_{\partial\mathscr{P}} := \sum_{K\in\mathscr{P}} \langle\boldsymbol{\mu},\boldsymbol{v}\rangle_{\partial K}\,,$$

for every $\boldsymbol{u}$, $\boldsymbol{v}\in\mathbf{V}$ and $\boldsymbol{\mu}\in\boldsymbol{\Lambda}$, where $\langle\cdot,\cdot\rangle_{\partial D}$ means the duality pairing between $\boldsymbol{H}^{-\frac{1}{2}}(\partial D)$ and $\boldsymbol{H}^{\frac{1}{2}}(\partial D)$ defined through

$$\langle\boldsymbol{\mu},\boldsymbol{v}\rangle_{\partial D} := \int_D (\nabla\cdot\boldsymbol{\tau})\cdot\boldsymbol{v}\,d\boldsymbol{x} + \int_D \boldsymbol{\tau}:\nabla\boldsymbol{v}\,d\boldsymbol{x} \quad \text{for all } \boldsymbol{v}\in\boldsymbol{H}^1(D)\,, \tag{2.8}$$

where $\boldsymbol{\tau}\in H(\mathbf{div};\Omega;\mathbb{S})$ is such that $\boldsymbol{\tau}\boldsymbol{n}^D = \boldsymbol{\mu}$ on ∂D. The space $\mathbf{V}$ is Hilbert with the inner product $(\cdot,\cdot)_{\mathbf{V}}$ as a consequence of the Korn's inequality (see [20, Theorem 3.78]). Space $(\boldsymbol{\Lambda},\|\cdot\|_{\boldsymbol{\Lambda}})$ is also Hilbert and the following equality holds (see Lemma A.3)

$$\|\boldsymbol{\mu}\|_{\boldsymbol{\Lambda}} = \sup_{\boldsymbol{v}\in\mathbf{V}} \frac{\langle\boldsymbol{\mu},\boldsymbol{v}\rangle_{\partial\mathscr{P}}}{\|\boldsymbol{v}\|_{\mathbf{V}}} \quad \text{for all } \boldsymbol{\mu}\in\boldsymbol{\Lambda}\,. \tag{2.9}$$

We lighten the notation and understand the supremum to be taken over sets excluding the zero function even though this is not specifically indicated.

Next, we address the local Korn's, Poincaré and trace inequalities on $K\in\mathscr{P}$ in the context of polytopal partitions. As for the Korn's inequality on the space $\widetilde{\mathbf{V}}$, we prove in Lemma A.1 that the following local Korn's inequality holds: There exists a positive constant $C_{korn,K}$, independent of h_K, such that

$$\|\nabla\tilde{\boldsymbol{v}}\|_{0,K} \leqslant C_{korn,K}\|\boldsymbol{\varepsilon}(\tilde{\boldsymbol{v}})\|_{0,K} \quad \text{for all } \tilde{\boldsymbol{v}}\in\widetilde{\mathbf{V}}(K)\,. \tag{2.10}$$

The global counterpart of the local Korn's inequality on the space $\mathbf{V}$ is presented in Lemma A.2 inspired by [13]. The result highlights the existence of a positive constant C, depending only on Ω and on the shape of the elements of $K\in\mathscr{P}$, such that

(2.11)
$$|\boldsymbol{v}|^2_{1,\mathscr{P}} \leqslant C \left(\|\underline{\boldsymbol{\varepsilon}}(\boldsymbol{v})\|^2_{0,\mathscr{P}} + \sum_{K\in\mathscr{P}} \left\| \Pi_{RM}\boldsymbol{v} - \frac{1}{|K|}\int_K \boldsymbol{v}\,\mathrm{d}\boldsymbol{x} \right\|^2_{0,K} + \sum_{E\in\mathcal{E}_0} \frac{1}{h_E}\|[\![\boldsymbol{v}]\!]\|^2_{0,E} \right),$$

for all $\boldsymbol{v} \in \mathbf{V}$. Above and hereafter, the jump function is $[\![\boldsymbol{v}]\!]\,|_E := \boldsymbol{v}\,|_K - \boldsymbol{v}\,|_{K'}$ for $E \in \mathcal{E}_0$, where $K, K' \in \mathscr{P}$ share the face E, and $[\![\boldsymbol{v}]\!]\,|_E := \boldsymbol{v}\,|_E$ for $E \in \partial\Omega$. The mapping Π_{RM} stands for the $\boldsymbol{L}^2(\Omega)$ projection onto $\mathbf{V}_{\mathrm{rm}}$, i.e., given $\boldsymbol{v} \in \mathbf{V}$, the function $\Pi_{RM}\boldsymbol{v}\,|_K$ satisfies

$$(\Pi_{RM}\boldsymbol{v}, \boldsymbol{v}^{\mathrm{rm}})_K = (\boldsymbol{v}, \boldsymbol{v}^{\mathrm{rm}})_K \quad \text{for all } \boldsymbol{v}^{\mathrm{rm}} \in \mathbf{V}_{\mathrm{rm}}(K)\,, \tag{2.12}$$

which immediately leads to the following estimates

(2.13)
$$\|\Pi_{RM}\boldsymbol{v}\|_{\mathbf{V}} \leqslant \|\boldsymbol{v}\|_{\mathbf{V}}, \quad \|\boldsymbol{v} - \Pi_{RM}\boldsymbol{v}\|_{\mathbf{V}} \leqslant \|\boldsymbol{v}\|_{\mathbf{V}} \quad \text{and} \quad \|\boldsymbol{v} - \Pi_{RM}\boldsymbol{v}\|_{0,\Omega} \leqslant \|\nabla\boldsymbol{v}\|_{1,\mathscr{P}}\,.$$

Regarding the Poincaré and trace inequalities, we recall some known results making sure to explicit their constant dependency in terms of the domain K. From [38, Lemma 1.49], there is a positive constant $C_{tr,K}$ such that the following trace inequalities hold

$$\|\boldsymbol{v}\|^2_{0,\partial K} \leqslant C_{tr,K}\,\left(d\,h_K^{-1}\|\boldsymbol{v}\|_{0,K} + 2\,\|\nabla\boldsymbol{v}\|_{0,K}\right)\,\|\boldsymbol{v}\|_{0,K}\,, \tag{2.14}$$

for all $\boldsymbol{v} \in \boldsymbol{H}^1(K)$. As for the Poincaré inequality (see [20]), it is well-known that there exists a positive constant $C_{P,K}$, depending only on the shape of K, such that

$$\|\tilde{\boldsymbol{v}}\|_{0,K} \leqslant C_{P,K}\,h_K\,\|\nabla\tilde{\boldsymbol{v}}\|_{0,K} \quad \text{for all } \tilde{\boldsymbol{v}} \in \widetilde{\mathbf{V}}(K)\,, \tag{2.15}$$

Combining (2.15) and (2.10), we obtain

$$\|\tilde{\boldsymbol{v}}\|_{0,K} \leqslant C_{korn,K}C_{P,K}\,h_K\|\boldsymbol{\varepsilon}(\tilde{\boldsymbol{v}})\|_{0,K} \quad \text{for all } \tilde{\boldsymbol{v}} \in \widetilde{\mathbf{V}}(K)\,. \tag{2.16}$$

and from (2.14) and (2.15), we get

$$\|\tilde{\boldsymbol{v}}\|^2_{0,\partial K} \leqslant C_{tr,K}\,C_{P,K}\,\left(d\,C_{P,K} + 2\right)\,h_K\,\|\nabla\tilde{\boldsymbol{v}}\|^2_{0,K} \quad \text{for all } \tilde{\boldsymbol{v}} \in \widetilde{\mathbf{V}}(K)\,. \tag{2.17}$$

Above and hereafter, we shall make systematically use of the following constants

$$C_{korn} := \max_{K\in\mathscr{P}} C_{korn,K}\,, \quad C_{tr} := \max_{K\in\mathscr{P}} C_{tr,K}\,, \quad \text{and} \quad C_P := \max_{K\in\mathscr{P}} C_{P,K} \tag{2.18}$$

where $C_{P,K}$, $C_{korn,K}$ and $C_{tr,K}$ denote the local Poincaré, Korn's and the trace constants, respectively. From now on, we use C to represent a generic constant independent of the mesh size.

Remark 2.1. We outline below some known aspects of the constants $C_{tr,K}$ and $C_{P,K}$ in terms of the shape of $K \in \mathscr{P}$.

- let $\mathcal{T}^K$ be a simplicial conforming partition of K, and define

$$\bar{\sigma}_K := \max_{\tau\in\mathcal{T}^K} \frac{h_\tau}{\rho_\tau}\,, \qquad \bar{\varrho}_K := \frac{\min_{\tau\in\mathcal{T}^K} h_\tau}{h_K}\,. \tag{2.19}$$

  Then, adapting the proof of [38, Lemma 1.49] to be used on ∂K, we conclude that the constant in (2.14) corresponds to $C_{tr,K} = (d+1)\,\bar{\sigma}_K$ if K is a simplex, and $C_{tr,K} = (d+1)\,\bar{\sigma}_K\,\bar{\varrho}_K^{-1}$ otherwise;

- it is well-known (see [34]) that the constant $C_{P,K}$ is equal to π^{-1} for convex domains no matter the aspect ratio of K; for star-shaped polytopal elements K, the constant $C_{P,K}$ in (2.15) depends only on d and the shape of K [44]. Recently, [36, Lemma 3.4] refined such a result and proved that $C_{P,K}$ depends only on d and ρ, where ρ is the radius of star-shaped domain embedding K;

- if one assumes some kind of control on the aspect ratio on K, notably $h_K^2 \leq C|K|$, as well as the existence of a natural number N such that, for every $K \in \mathscr{P}$, there exists a collection of points $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N \in K$ and a collection $M_1, \ldots, M_N$ of closed (possibly overlapping) subsets of K such that M_i is star-shaped with respect to $\boldsymbol{x}_i$ and $K = \cup_{i=1}^N M_i$, then one obtains $C_{P,K} = C\,(1+N)$, where C is independent of K (c.f. [46]).

□

2.3. Characterizing the exact solution. We extend the approach proposed in [28] for simplexes and quadrilateral elements, and propose a decomposition of the exact solution in terms of local problems defined on polytopal elements. The first step is to observe that the exact solution $\boldsymbol{u}$ decomposes locally in the following direct sum

$$\boldsymbol{u} \,=\, \boldsymbol{u}^{\mathrm{rm}} + \tilde{\boldsymbol{u}} \quad \text{with} \quad \boldsymbol{u}^{\mathrm{rm}} \in \mathbf{V}_{\mathrm{rm}}(K) \text{ and } \tilde{\boldsymbol{u}} \in \widetilde{\mathbf{V}}(K)\,. \tag{2.20}$$

Such a characterization is a direct consequence of $\mathbf{V}$ written in the following direct sum

$$\mathbf{V} = \widetilde{\mathbf{V}} \oplus \mathbf{V}_{\mathrm{rm}}\,,$$

where the spaces are given in (2.7). Then, we assume that $\tilde{\boldsymbol{u}} \in \widetilde{\mathbf{V}}$ can be parameterized in terms of a function $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}$ and the datum $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega)$ by

$$\tilde{\boldsymbol{u}} = T\boldsymbol{\lambda} + \hat{T}\boldsymbol{f}\,, \tag{2.21}$$

where the mappings $T \in \mathcal{L}(\boldsymbol{\Lambda}, \widetilde{\mathbf{V}})$ and $\hat{T} \in \mathcal{L}(\boldsymbol{L}^2(\Omega), \widetilde{\mathbf{V}})$ are defined locally as follows

- for all $\boldsymbol{\mu} \in \boldsymbol{\Lambda}$ and $K \in \mathscr{P}$, $T\boldsymbol{\mu}\,|_K \in \widetilde{\mathbf{V}}(K)$ is the unique solution of

$$\int_K \mathbf{C}\underline{\boldsymbol{\varepsilon}}(T\,\boldsymbol{\mu}) \,:\, \underline{\boldsymbol{\varepsilon}}(\tilde{\boldsymbol{v}})\,d\boldsymbol{x} = \langle \boldsymbol{\mu}, \tilde{\boldsymbol{v}} \rangle_{\partial K} \quad \text{for all } \tilde{\boldsymbol{v}} \in \widetilde{\mathbf{V}}(K)\,; \tag{2.22}$$

- for all $\boldsymbol{q} \in \boldsymbol{L}^2(\Omega)$ and $K \in \mathscr{P}$, $\hat{T}\boldsymbol{q}\,|_K \in \widetilde{\mathbf{V}}(K)$ is the unique solution of

$$\int_K \mathbf{C}\underline{\boldsymbol{\varepsilon}}(\hat{T}\boldsymbol{q}) \,:\, \underline{\boldsymbol{\varepsilon}}(\tilde{\boldsymbol{v}})\,d\boldsymbol{x} = \int_K \boldsymbol{q}\cdot\tilde{\boldsymbol{v}}\,d\boldsymbol{x} \quad \text{for all } \tilde{\boldsymbol{v}} \in \widetilde{\mathbf{V}}(K)\,. \tag{2.23}$$

The following result ensures T and $\hat{T}$ are well-defined continuous mappings and then $\tilde{\boldsymbol{u}}$ is well-defined.

Lemma 2.2. *Given $\boldsymbol{q} \in \boldsymbol{L}^2(\Omega)$ and $\boldsymbol{\mu} \in \boldsymbol{\Lambda}$ there is a unique function $\tilde{\boldsymbol{w}} := T\boldsymbol{\mu} + \hat{T}\boldsymbol{q} \in \widetilde{\mathbf{V}}(K)$ satisfying*

$$\int_K \mathbf{C}\underline{\boldsymbol{\varepsilon}}(\tilde{\boldsymbol{w}}) \,:\, \underline{\boldsymbol{\varepsilon}}(\tilde{\boldsymbol{v}})\,d\boldsymbol{x} = \int_K \boldsymbol{q}\cdot\tilde{\boldsymbol{v}}\,d\boldsymbol{x} + \langle \boldsymbol{\mu}, \tilde{\boldsymbol{v}} \rangle_{\partial K} \quad \textit{for all } \tilde{\boldsymbol{v}} \in \widetilde{\mathbf{V}}(K)\,, \tag{2.24}$$

for all $K \in \mathscr{P}$. Moreover, it holds

$$\begin{aligned} \|T\boldsymbol{\mu}\|_{\mathbf{V}} &\leqslant \frac{1}{c_{min}}\,\left(1 + (C_{korn}\,C_P)^2\right)\,\|\boldsymbol{\mu}\|_{\boldsymbol{\Lambda}} \quad \textit{for all } \boldsymbol{\mu} \in \boldsymbol{\Lambda}\,, \\ \|\hat{T}\boldsymbol{q}\|_{\mathbf{V}} &\leqslant \frac{\mathcal{H}}{c_{min}}\,\left(C_{korn}\,C_P\,\sqrt{1 + (C_{korn}\,C_P)^2}\right)\,\|\boldsymbol{q}\|_{0,\Omega} \quad \textit{for all } \boldsymbol{q} \in \boldsymbol{L}^2(\Omega)\,. \end{aligned} \tag{2.25}$$

Proof. First, notice that the bilinear form on the left-hand side and the linear forms on the right-hand side of (2.24) are continuous with respect to the norms $\|\cdot\|_{1,K}$, $\|\cdot\|_{0,K}$, and $\|\cdot\|_{-1/2,\partial K}$, respectively. In addition, the bilinear form in right-hand side of (2.24) is also coercive in the $\widetilde{\mathbf{V}}(K)$ space, with a coercive constant independent of K. Indeed, from the local Korn's and Poincaré inequalities in (2.10) and (2.15), respectively, for all $\tilde{\boldsymbol{v}} \in \widetilde{\mathbf{V}}(K)$ it holds

$$\begin{aligned}\int_K \mathbf{C}\underline{\boldsymbol{\varepsilon}}(\tilde{\boldsymbol{v}}) : \underline{\boldsymbol{\varepsilon}}(\tilde{\boldsymbol{v}})\,\mathrm{d}\boldsymbol{x} &\geqslant c_{min}\|\underline{\boldsymbol{\varepsilon}}(\tilde{\boldsymbol{v}})\|^2_{0,K}\\ &\geqslant \frac{c_{min}}{C^2_{korn,K}}\|\nabla\tilde{\boldsymbol{v}}\|^2_{0,K}\\ &\geqslant \frac{c_{min}}{2\,C^2_{korn,K}}\Big(\|\nabla\tilde{\boldsymbol{v}}\|^2_{0,K}+\frac{1}{\mathcal{H}^2_K\,C^2_{P,K}}\|\tilde{\boldsymbol{v}}\|^2_{0,K}\Big)\\ &\geqslant \frac{c_{min}}{2\,C^2_{korn}}\min\{1,(h_\Omega\,C_P)^{-2}\}\|\tilde{\boldsymbol{v}}\|^2_{1,K}\,.\end{aligned}$$

Hence, well-posedness follows from the Lax-Milgram lemma (e.g., [20, Lemma 2.2]). Using (2.9) and (2.22), and $\mathcal{H}\leqslant h_\Omega$ it holds

$$\begin{aligned}c_{min}\,\|T\boldsymbol{\mu}\|^2_{\mathbf{V}} &\leqslant \left(1+\frac{(C_{korn}\,C_P)^2\,\mathcal{H}^2}{h^2_\Omega}\right)\,\langle\boldsymbol{\mu},T\boldsymbol{\mu}\rangle_{\mathscr{P}}\\ &\leqslant \left(1+(C^2_{korn}\,C_P)^2\right)\,\|\boldsymbol{\mu}\|_{\boldsymbol{\Lambda}}\,\|T\boldsymbol{\mu}\|_{\mathbf{V}}\,,\end{aligned}$$

and the estimate (2.25) follows. Next, from (2.23), the Cauchy-Schwarz inequality, and Poincaré and Korn's inequalities, we arrive at

$$\begin{aligned}c_{min}\,\|\underline{\boldsymbol{\varepsilon}}(\hat{T}\boldsymbol{q})\|^2_{0,\mathscr{P}} &\leqslant (\boldsymbol{q}\,,\hat{T}\boldsymbol{q})_\Omega\\ &\leqslant \|\boldsymbol{q}\|_{0,\Omega}\,\|\hat{T}\boldsymbol{q}\|_{0,\Omega}\\ &\leqslant C_{korn}\,C_P\,\mathcal{H}\,\|\boldsymbol{q}\|_{0,\Omega}\,\|\underline{\boldsymbol{\varepsilon}}(\hat{T}\boldsymbol{q})\|_{0,\mathscr{P}}\,.\end{aligned}\tag{2.26}$$

Then, using the (2.16) and (2.26), and the definition of the $\mathbf{V}$-norm, we obtain

$$\begin{aligned}\|\hat{T}\boldsymbol{q}\|^2_{\mathbf{V}} &= \frac{1}{h^2_\Omega}\|\hat{T}\boldsymbol{q}\|^2_{0,\Omega}+\|\underline{\boldsymbol{\varepsilon}}(\hat{T}\boldsymbol{q})\|^2_{0,\mathscr{P}}\\ &\leqslant \frac{1}{c^2_{min}}\left(\frac{\left((C_{korn}\,C_P)^2\,\mathcal{H}^2\right)^2}{h^2_\Omega}+(C_{korn}\,C_P\,\mathcal{H})^2\right)\|\boldsymbol{q}\|^2_{0,\Omega}\\ &\leqslant \frac{1}{c^2_{min}}\left((C_{korn}\,C_P)^2\,\mathcal{H}^2\,\left(1+\frac{(C_{korn}\,C_P)^2\,\mathcal{H}^2}{h^2_\Omega}\right)\right)\|\boldsymbol{q}\|^2_{0,\Omega}\,.\end{aligned}$$ ∎

Next, it remains to set $\boldsymbol{u}^{\mathrm{rm}}$ and $\boldsymbol{\lambda}$ such that (2.20) represents the exact solution using the definition (2.21). To this end, we define $a(\cdot,\cdot)$ and $b(\cdot,\cdot)$ as the following continuous bilinear forms given by

$$\begin{aligned}a:\boldsymbol{\Lambda}\times\boldsymbol{\Lambda}\to\mathbb{R} \qquad &\text{with} \qquad a(\boldsymbol{\lambda},\boldsymbol{\mu}):=\langle\boldsymbol{\mu},T\boldsymbol{\lambda}\rangle_{\partial\mathscr{P}}\,,\\ b:\boldsymbol{\Lambda}\times\mathbf{V}_{\mathrm{rm}}\to\mathbb{R} \qquad &\text{with} \qquad b(\boldsymbol{\mu},\boldsymbol{v}^{\mathrm{rm}}):=\langle\boldsymbol{\mu},\boldsymbol{v}^{\mathrm{rm}}\rangle_{\partial\mathscr{P}}\,,\end{aligned}\tag{2.27}$$

and search the pair $(\boldsymbol{\lambda},\boldsymbol{u}^{\mathrm{rm}})\in\boldsymbol{\Lambda}\times\mathbf{V}_{\mathrm{rm}}$ as the solution of the following global problem

$$\begin{cases} a(\boldsymbol{\lambda},\boldsymbol{\mu}) + b(\boldsymbol{\mu},\boldsymbol{u}^{\mathrm{rm}}) = -\langle \boldsymbol{\mu}, \hat{T}\boldsymbol{f}\rangle_{\partial\mathscr{P}} + \langle \boldsymbol{\mu}, \boldsymbol{g}\rangle_{\partial\Omega} & \text{for all } \boldsymbol{\mu}\in\boldsymbol{\Lambda}\,, \\ b(\boldsymbol{\lambda},\boldsymbol{v}^{\mathrm{rm}}) = -(\boldsymbol{f},\boldsymbol{v}^{\mathrm{rm}})_{\mathscr{P}} & \text{for all } \boldsymbol{v}^{\mathrm{rm}}\in\mathbf{V}_{\mathrm{rm}}\,. \end{cases} \tag{2.28}$$

The next result ensures the existence and uniqueness of a solution for problem (2.28) such that the function $\boldsymbol{u}$ given through (2.20) and (2.21) recovers the exact solution of the original problem. Our proof extends the proof of [28, Theorem 4.2] for simplicial meshes to polytopal partitions. For the sake of completeness, we provide all the details.

Theorem 2.3. *There exist positive constants α and β, independent of the partition $\mathscr{P}$, such that*

$$a(\boldsymbol{\mu},\boldsymbol{\mu}) \geq \alpha\,\|\boldsymbol{\mu}\|^2_{\boldsymbol{\Lambda}} \quad \textit{for all } \boldsymbol{\mu}\in\mathscr{N}\,, \tag{2.29}$$

$$\sup_{\boldsymbol{\mu}\in\boldsymbol{\Lambda}} \frac{b(\boldsymbol{\mu},\boldsymbol{v}^{\mathrm{rm}})}{\|\boldsymbol{\mu}\|_{\boldsymbol{\Lambda}}} \geq \beta\,\|\boldsymbol{v}^{\mathrm{rm}}\|_{\mathbf{V}} \quad \textit{for all } \boldsymbol{v}^{\mathrm{rm}}\in\mathbf{V}_{\mathrm{rm}}\,, \tag{2.30}$$

where

$$\mathscr{N} := \{\boldsymbol{\mu}\in\boldsymbol{\Lambda} \,:\, \langle \boldsymbol{\mu},\boldsymbol{v}^{\mathrm{rm}}\rangle_{\partial\mathscr{P}} = 0 \quad \textit{for all } \boldsymbol{v}^{\mathrm{rm}}\in\mathbf{V}_{\mathrm{rm}}\}\,.$$

Hence, (2.28) has a unique solution $(\boldsymbol{\lambda},\boldsymbol{u}^{\mathrm{rm}})\in\boldsymbol{\Lambda}\times\mathbf{V}_{\mathrm{rm}}$. Moreover, the following characterization of the exact solution $\boldsymbol{u}$ holds

$$\boldsymbol{u} \,=\, \boldsymbol{u}^{\mathrm{rm}} + T\boldsymbol{\lambda} + \hat{T}\boldsymbol{f}\,, \tag{2.31}$$

where $\boldsymbol{\lambda}$ corresponds to

$$\boldsymbol{\lambda} \,=\, (\mathbf{C}\,\underline{\boldsymbol{\varepsilon}}(\boldsymbol{u}))\,\boldsymbol{n}^K \quad \textit{on } \partial K\,, \quad \textit{for all } K\in\mathscr{P}\,.$$

Proof. Let $\boldsymbol{\mu}\in\mathscr{N}$ and $\boldsymbol{\sigma}^{\boldsymbol{\mu}} := \mathbf{C}\,\underline{\boldsymbol{\varepsilon}}(T\boldsymbol{\mu})$. Observe from (2.22) that $\boldsymbol{\sigma}^{\boldsymbol{\mu}}\,\boldsymbol{n}^K|_{\partial K} = \boldsymbol{\mu}\in\boldsymbol{\Lambda}$ for all $K\in\mathscr{P}$, and thus, the stress tensor $\boldsymbol{\sigma}^{\boldsymbol{\mu}}\in H(\mathbf{div};\Omega;\mathbb{S})$. Also, it follows from (2.22) that $\nabla\cdot\boldsymbol{\sigma}^{\boldsymbol{\mu}} = \mathbf{0}$. As a result, we get

$$a(\boldsymbol{\mu},\boldsymbol{\mu}) = \sum_{K\in\mathscr{P}}\int_K \boldsymbol{\sigma}^{\boldsymbol{\mu}} : \mathbf{C}^{-1}\boldsymbol{\sigma}^{\boldsymbol{\mu}}\,\mathrm{d}\boldsymbol{x} \geqslant \frac{1}{c_{max}}\sum_{K\in\mathscr{P}}\|\boldsymbol{\sigma}^{\boldsymbol{\mu}}\|^2_{0,K} = \frac{\|\boldsymbol{\sigma}^{\boldsymbol{\mu}}\|^2_{\mathrm{div}}}{c_{max}} \geqslant \frac{1}{c_{max}}\,\|\boldsymbol{\mu}\|^2_{\boldsymbol{\Lambda}}\,,$$

and then (2.29) holds with $\alpha = c_{max}^{-1}$. As for (2.30), first take $\boldsymbol{v}^{\mathrm{rm}}\in\mathbf{V}_{\mathrm{rm}}$ and notice that there is a symmetric field $\boldsymbol{\sigma}^*\in H^1(\Omega)^{d\times d}$ such that

$$\nabla\cdot\boldsymbol{\sigma}^* = \boldsymbol{v}^{\mathrm{rm}} \quad \text{in } \Omega \qquad \text{and} \qquad \|\boldsymbol{\sigma}^*\|_{1,\Omega} \leqslant C\,\|\boldsymbol{v}^{\mathrm{rm}}\|_{0,\Omega}\,, \tag{2.32}$$

where C is a positive constant depending only on Ω. The result (2.32) follows from

a classical argument. First, we extend $\boldsymbol{v}^{\mathrm{rm}}$ to a larger smoothly bounded domain $B_\Omega \supset \Omega$ such that $\boldsymbol{v}^{\mathrm{rm}} = \boldsymbol{0}$ in $B_\Omega \backslash \Omega$, and define $\boldsymbol{u}^* \in \boldsymbol{H}^1_0(B_\Omega)$ as the unique solution satisfying $\nabla\cdot(\underline{\boldsymbol{\varepsilon}}(\boldsymbol{u}^*)) = \boldsymbol{v}^{\mathrm{rm}}$ in Ω. From a standard regularity argument for elliptic problems, e.g., [20, §3.1.3], it holds $\boldsymbol{u}^* \in \boldsymbol{H}^2(B_\Omega)$ and $\|\boldsymbol{u}^*\|_{2,B_\Omega} \leqslant C\,\|\boldsymbol{v}^{\mathrm{rm}}\|_{0,B_\Omega}$ with c depending only on B_Ω. Thereby, we consider the symmetric field $\boldsymbol{\sigma}^* := \underline{\boldsymbol{\varepsilon}}(\boldsymbol{u}^*) \in H^1(B_\Omega)^{d\times d}$ observing that $\nabla\cdot\boldsymbol{\sigma}^* = \boldsymbol{v}^{\mathrm{rm}}$ in B_Ω. Thereby, the inequality in (2.32) follows as

$$\|\boldsymbol{\sigma}^*\|_{1,\Omega} \leq \|\boldsymbol{\sigma}^*\|_{1,B_\Omega} \leqslant C\,\|\boldsymbol{v}^{\mathrm{rm}}\|_{0,B_\Omega} = C\,\|\boldsymbol{v}^{\mathrm{rm}}\|_{0,\Omega}\,.$$

Next, we set $\boldsymbol{\mu}^* := \boldsymbol{\sigma}^*\boldsymbol{n}^K \in \boldsymbol{\Lambda}$ for all $K \in \mathscr{P}$, and arrive at

$$\|\boldsymbol{\mu}^*\|_{\boldsymbol{\Lambda}}\,\|\boldsymbol{v}^{\mathrm{rm}}\|_{\mathbf{V}} \leq \|\boldsymbol{\sigma}^*\|_{\mathrm{div}}\,\|\boldsymbol{v}^{\mathrm{rm}}\|_{\mathbf{V}} \leqslant C\,\|\boldsymbol{v}^{\mathrm{rm}}\|^2_{0,\Omega} = C\,(\nabla\cdot\boldsymbol{\sigma}^*, \boldsymbol{v}^{\mathrm{rm}})_\Omega = C\,b(\boldsymbol{\mu}^*, \boldsymbol{v}^{\mathrm{rm}})\,.$$

and then, it holds

$$\|\boldsymbol{v}^{\mathrm{rm}}\|_{\mathbf{V}} \leqslant C\,\frac{b(\boldsymbol{\mu}^*, \boldsymbol{v}^{\mathrm{rm}})}{\|\boldsymbol{\mu}^*\|_{\boldsymbol{\Lambda}}}\,, \tag{2.33}$$

which proves (2.30). The well-posedness of (2.28) follows from the standard saddle-point theory (c.f. [20]). Notice that the function in (2.31) satisfies the original problem (2.3) using the global-local problem (2.22)-(2.23) and (2.28), and the characterization of the exact solution follows from uniqueness. The traction is represented by $\boldsymbol{\lambda}$ as a consequence of (2.31) and the definition of the local problem (2.22). ∎

The MHM method relies on the discretization of the global-local formulation (2.22)-(2.23) and (2.28) instead of the standard formulation (2.3). This is addressed next.

3. The MHM method. The MHM method uses a multi-level discretization starting from the first-level partition $\mathscr{P}$. Notably, each face $E \subset \mathcal{E}$ and polytopal element $K \in \mathscr{P}$ may carry its own family of partitions in a way that each member is *a priori* independent of each other. Nevertheless, in view of the numerical analysis of the method, a compatibility condition between the different partition levels is mandatory which yields a mild assumption in the form of the polytope element K with an underlying minimum angle condition. To be more specific, one needs some further notations first.

Without loss of generality, we shall use hereafter the terminology employed for three-dimensional domains. Let $\{\mathcal{E}_H\}_{H>0}$ be a family of partitions of $\mathcal{E}$, and set $H := \max_{F\in\mathcal{E}_H} h_F$. We call $\mathcal{E}_H^K$ the collection of faces $F \in \mathcal{E}_H$ such that $F \subset \partial K$.

Assumption (A): For each $K \in \mathscr{P}$, there exists $\{\Xi_H^K\}_{H>0}$, a shape-regular family of conforming simplicial partitions of K matching with $\{\mathcal{E}_H\}_{H>0}$, i.e., such that, for each $F \in \mathcal{E}_H^K$, there exists an element $\tau_F \in \Xi_H^K$ with $\partial\tau_F \cap \partial K = F$.

Owing to Assumption (A) and for every $H > 0$, we also define:

(i) the minimal partition Ξ_H given by

$$\Xi_H := \cup_{K\in\mathscr{P}}\Xi_H^K\,.$$

Its shape-regularity constant is $\sigma_\Xi := \max_{K\in\mathscr{P}}\max_{\tau\in\Xi_H^K}\sigma_\tau$;

(ii) a shape-regular family of simplicial partitions $\{\mathcal{T}_h^K\}_{h>0}$ on K from successive refinements of Ξ_H^K, and

$$\mathcal{T}_h := \cup_{K\in\mathscr{P}}\mathcal{T}_h^K \quad\text{and}\quad \sigma_{\mathcal{T}} := \max_{K\in\mathscr{P}}\max_{\tau\in\mathcal{T}_h^K}\sigma_\tau\,. \tag{3.1}$$

See Figure 3.1 for an illustration of the partitions $\mathscr{P}$, $\mathcal{E}_H$ and $\mathcal{T}_h$.

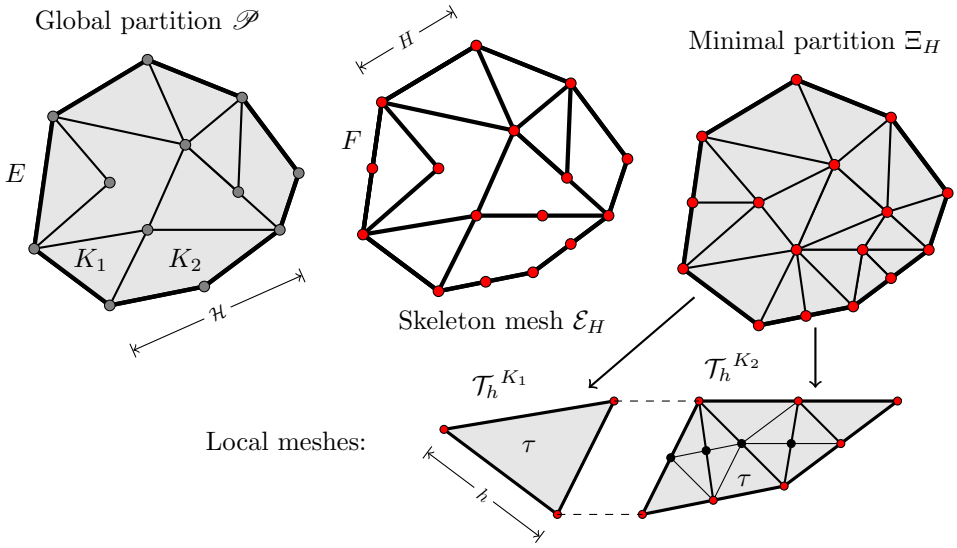

FIGURE 3.1. *A two-dimensional polygonal domain with the overlaid meshes $\mathscr{P}$, Ξ_H and $\mathcal{T}_h$. The red dots stands for the degrees of freedom associated with the faces and the black dots with the local meshes.*

To this end, we define discrete spaces using the face partition $\mathcal{E}_H$ and the second-level mesh $\mathcal{T}_h$, as follows

$$\begin{aligned}
&\boldsymbol{\Lambda}_H(K) := \{\boldsymbol{\mu}_H \in \boldsymbol{L}^2(\partial K)\,:\,\boldsymbol{\mu}_H|_F \in \mathbb{P}_\ell(F)^d\,,\,\forall\,F\in\mathcal{E}_H^K\},\quad \ell\geqslant 1\,,\\
&\mathbf{V}_h(K) := \{\boldsymbol{v}_h \in \mathbf{C}^0(K) : \boldsymbol{v}_h|_\tau \in \mathbb{P}_k(\tau)^d\,,\forall\,\tau\in\mathcal{T}_h^K\},\quad k\geqslant \ell+d\,,\\
&\widetilde{\mathbf{V}}_h(K) := \mathbf{V}_h(K)\cap\widetilde{\mathbf{V}}(K)\,,
\end{aligned} \tag{3.2}$$

where k, $\ell\in\mathbb{N}^+$ and $\mathbb{P}_s(D)$ is the space of polynomials of degree less or equal to s on D. The corresponding global spaces are

$$\begin{aligned}
&\mathbf{V}_h := \{\boldsymbol{v}_h\in\mathbf{V}\,:\,\boldsymbol{v}_h|_K\in\mathbf{V}_h(K),\,K\in\mathscr{P}\}\,,\qquad \widetilde{\mathbf{V}}_h := \mathbf{V}_h\cap\widetilde{\mathbf{V}}\,,\\
&\boldsymbol{\Lambda}_H := \{\boldsymbol{\mu}_H\in\boldsymbol{\Lambda}\,:\,\boldsymbol{\mu}_H|_{\partial K}\in\boldsymbol{\Lambda}_H(K),\,K\in\mathscr{P}\}\,.
\end{aligned} \tag{3.3}$$

The MHM method consists of finding $(\boldsymbol{\lambda}_H, \boldsymbol{u}_h^{\mathrm{rm}}) \in \boldsymbol{\Lambda}_H \times \mathbf{V}_{\mathrm{rm}}$ such that

$$\begin{aligned} a_h(\boldsymbol{\lambda}_H, \boldsymbol{\mu}_H) + b(\boldsymbol{\mu}_H, \boldsymbol{u}_h^{\mathrm{rm}}) &= -\langle \boldsymbol{\mu}_H, \hat{T}_h \boldsymbol{f} \rangle_{\partial\mathscr{P}} + \langle \boldsymbol{\mu}_H, \boldsymbol{g} \rangle_{\partial\Omega} \quad \text{for all } \boldsymbol{\mu}_H \in \boldsymbol{\Lambda}_H\,, \\ b(\boldsymbol{\lambda}_H, \boldsymbol{v}^{\mathrm{rm}}) &= -(\boldsymbol{f}, \boldsymbol{v}^{\mathrm{rm}})_{\mathscr{P}} \quad \text{for all } \boldsymbol{v}^{\mathrm{rm}} \in \mathbf{V}_{\mathrm{rm}}\,, \end{aligned} \tag{3.4}$$

where $b(\cdot,\cdot)$ is given in (2.27), and the discrete bilinear form $a_h(\cdot,\cdot)$ defined by

$$a_h : \boldsymbol{\Lambda} \times \boldsymbol{\Lambda} \to \mathbb{R} \quad \text{where} \quad a_h(\boldsymbol{\lambda}, \boldsymbol{\mu}) := \langle \boldsymbol{\mu}, T_h \boldsymbol{\lambda} \rangle_{\partial\mathscr{P}}\,. \tag{3.5}$$

The mappings $T_h \in \mathcal{L}(\boldsymbol{\Lambda}, \widetilde{\mathbf{V}}_h)$ and $\hat{T}_h \in \mathcal{L}(\boldsymbol{L}^2(\Omega), \widetilde{\mathbf{V}}_h)$ approximate T and $\hat{T}$ given in (2.22)-(2.23). In this work, these approximations correspond to the Galerkin method on the standard continuous polynomial interpolation spaces, i.e., we set

- for all $\boldsymbol{\mu} \in \boldsymbol{\Lambda}$ and $K \in \mathscr{P}$, $T_h\boldsymbol{\mu}\,|_K \in \widetilde{\mathbf{V}}_h(K)$ is the unique solution of

$$\int_K \mathbf{C}\underline{\boldsymbol{\varepsilon}}(T_h\boldsymbol{\mu}) : \underline{\boldsymbol{\varepsilon}}(\tilde{\boldsymbol{v}}_h)\, d\boldsymbol{x} = \langle \boldsymbol{\mu}, \tilde{\boldsymbol{v}}_h \rangle_{\partial K} \quad \text{for all } \tilde{\boldsymbol{v}}_h \in \widetilde{\mathbf{V}}_h(K)\,; \tag{3.6}$$

- for all $\boldsymbol{q} \in \boldsymbol{L}^2(\Omega)$ and $K \in \mathscr{P}$, $\hat{T}_h\boldsymbol{q}\,|_K \in \widetilde{\mathbf{V}}_h(K)$ is the unique solution of

$$\int_K \mathbf{C}\underline{\boldsymbol{\varepsilon}}(\hat{T}_h\boldsymbol{q}) : \underline{\boldsymbol{\varepsilon}}(\tilde{\boldsymbol{v}}_h)\, d\boldsymbol{x} = \int_K \boldsymbol{q} \cdot \tilde{\boldsymbol{v}}_h\, d\boldsymbol{x} \quad \text{for all } \tilde{\boldsymbol{v}}_h \in \widetilde{\mathbf{V}}_h(K)\,. \tag{3.7}$$

It is essential at this point to bring some comments on the MHM method (3.4) regarding its flexibility and properties, before heading to its numerical analysis.

Remark 3.1.

- Notice that, the proof of Lemma 2.2 still holds if we replace (2.22)–(2.23) by (3.6)–(3.7) and, therefore, we conclude that the mappings T_h and $\hat{T}_h$ are well-defined and bounded as follows

(3.8)

$$\begin{aligned} \|T_h\boldsymbol{\mu}\|_{\mathbf{V}} &\leqslant \frac{1}{c_{min}} \left(1 + (C_{korn}\, C_P)^2\right) \|\boldsymbol{\mu}\|_{\boldsymbol{\Lambda}} \quad \text{for all } \boldsymbol{\mu} \in \boldsymbol{\Lambda}\,, \\ \|\hat{T}_h\boldsymbol{q}\|_{\mathbf{V}} &\leqslant \frac{\mathcal{H}}{c_{min}} \left(C_{korn}\, C_P \sqrt{1 + (C_{korn}\, C_P)^2}\right) \|\boldsymbol{q}\|_{0,\Omega} \quad \text{for all } \boldsymbol{q} \in \boldsymbol{L}^2(\Omega)\,. \end{aligned}$$

- The proposed method (3.4) preserves the local equilibrium in the sense that, for all $K \in \mathscr{P}$,

$$\int_{\partial K} \boldsymbol{\lambda}_H \cdot \boldsymbol{v}^{\mathrm{rm}}\, \mathrm{d}\mathbf{s} + \int_K \boldsymbol{f} \cdot \boldsymbol{v}^{\mathrm{rm}}\, \mathrm{d}\boldsymbol{x} = 0 \quad \text{for all } \boldsymbol{v}^{\mathrm{rm}} \in \mathbf{V}_{\mathrm{rm}}\,,$$

where $\boldsymbol{\lambda}_H$ is the discrete traction field on ∂K. In the case of multi-query purposes (i.e., $\boldsymbol{f}$ changes multiple times), the method (3.4) can be rewritten without contribution $\hat{T}_h\boldsymbol{f}$, by using the following equivalence

$$\langle \boldsymbol{\mu}, \hat{T}_h\boldsymbol{f}\rangle_{\partial K} = \int_K \mathbf{C}\,\underline{\boldsymbol{\varepsilon}}(T_h\boldsymbol{\mu}) : \underline{\boldsymbol{\varepsilon}}(\hat{T}_h\boldsymbol{f})\,\mathrm{d}\boldsymbol{x} = \int_K T_h\boldsymbol{\mu}\cdot\boldsymbol{f}\,\mathrm{d}\boldsymbol{x}\,. \tag{3.9}$$

Also, observe that $\hat{T}_h\boldsymbol{f} = 0$ if $\boldsymbol{f} \in \mathbf{V}_{\mathrm{rm}}$, and then, we can disregard the term $\hat{T}_h\boldsymbol{f}$ without undermining convergence if $\boldsymbol{f}$ is a low degree polynomial function.

- The mappings T_h and $\hat{T}_h$ can be defined quite generally and encompass other methods. Such a choice depends on which unknown one wants to approximate accurately and impacts the robustness of the method in terms of physical coefficients. For instance, to ensure that the method yields a numerical stress tensor $\boldsymbol{\sigma}_{Hh} \in H(\mathbf{div};\Omega;\mathbb{S})$, we can constrain $\boldsymbol{\Lambda}_H$ and redefine T_h and $\hat{T}_h$ using a stable mixed finite element method. This alternative addressed in [18] and named MHM-Hdiv method relaxes symmetry and produces a discrete stress tensor in $H(\mathbf{div};\Omega;\mathbb{S})$ with optimal convergence. In such a case, we mitigate the increased complexity of the local problems with the outcome to release an accurate stress tensor $\boldsymbol{\sigma}_{Hh}$ in $H(\mathbf{div};\Omega;\mathbb{S})$. Although both methodologies work on general polygonal meshes and can achieve convergence without refining the global partition, the construction and analysis of the current method and the MHM-Hdiv are fundamentally different. Besides, we can improve the robustness of the method to deal with quasi-incompressible materials adopting stabilized finite element methods to define T_h and $\hat{T}_h$. This option, proposed in [35], corresponds to an alternative to using a mixed finite element method as a second-level solver on this regard.

- The one-level MHM method corresponds to replace T_h and $\hat{T}_h$ by T and $\hat{T}$ in (3.4) (see [28] for details). Although, the one-level method may not be effective since no closed formula is available for local problem solutions in general, such a perspective leads to interesting insights about the relationship between the MHM method and more standard methods (see [26, 25] for the Poisson equation case). Also, it works as a first step and a guideline towards more effective two-level methods since the one-level approach yields symmetric stress variables $\boldsymbol{\sigma}_H := \mathbf{C}\,\underline{\boldsymbol{\varepsilon}}(\boldsymbol{u}^{\mathrm{rm}} + T\boldsymbol{\lambda}_H + \hat{T}\boldsymbol{f})$ in $H(\mathbf{div};\Omega;\mathbb{S})$ that satisfy the equilibrium equation $\nabla\cdot\boldsymbol{\sigma}_H + \boldsymbol{f} = \mathbf{0}$ exactly.

□

Summary: The exact displacement $\boldsymbol{u}$ is approximated by the $\boldsymbol{H}^1(\Omega)$ non-conforming discrete solution $\boldsymbol{u}_{Hh}$ given by

$$\boldsymbol{u}_{Hh} := \boldsymbol{u}_h^{\mathrm{rm}} + T_h\boldsymbol{\lambda}_H + \hat{T}_h\boldsymbol{f}\,, \tag{3.10}$$

where $(\boldsymbol{u}_h^{\mathrm{rm}}, \boldsymbol{\lambda}_H)$ solves the global system (3.4) and T_h and $\hat{T}_h$ are computed from the local problems (3.6)-(3.7). We recover a symmetric approximation of the exact stress field $\boldsymbol{\sigma}$ by the simple post-processing $\boldsymbol{\sigma}_{Hh} := \mathbf{C}\underline{\boldsymbol{\varepsilon}}(\boldsymbol{u}_{Hh})$.

4. Well-posedness. This section proves the well-posedness of the MHM method (3.4). To this end, we extend the classical work [40] to more general polytopal elements including face partitioning. The main ingredient is the construction of a new Fortin operator under the condition $k \geq \ell + d$, where ℓ and k are the degree of polynomial functions in $\mathbf{\Lambda}_H$ and $\mathbf{V}_h$, respectively.

4.1. A Fortin operator. The Fortin operator inherits the two-level characteristic of the method. Thereby, we first define the second-level form of the Fortin operator. Roughly, the demonstration combines the strategies proposed in [10] and [23].

Lemma 4.1. *Let $\tau \in \mathcal{T}_h$ and $F \subset \partial\tau$, and assume $\ell \geqslant 1$. For all $\boldsymbol{v} \in \boldsymbol{H}^1(\tau)$, there exists a unique $\rho_F^\tau(\boldsymbol{v}) \in \mathbb{P}_{\ell+d}^F(\tau)^d$, with*

$$\mathbb{P}_{\ell+d}^F(\tau)^d := \{\mathbf{p} \in \mathbb{P}_{\ell+d}(\tau)^d : \mathbf{p}|_{\partial\tau\setminus F} = 0\},$$

such that

$$\begin{aligned} \int_\tau (\rho_F^\tau \boldsymbol{v} - \boldsymbol{v}) \cdot \boldsymbol{q}\, d\boldsymbol{x} = 0, \qquad & \text{for all } \boldsymbol{q} \in \mathbb{P}_{\ell-1}(\tau)^d, \\ \int_F (\rho_F^\tau \boldsymbol{v} - \boldsymbol{v}) \cdot \boldsymbol{\mu}\, d\mathbf{s} = 0, \qquad & \text{for all } \boldsymbol{\mu} \in \mathbb{P}_\ell(F)^d, \end{aligned} \tag{4.1}$$

and

$$\|\rho_F^\tau \boldsymbol{v}\|_{0,\tau} + h_\tau |\rho_F^\tau \boldsymbol{v}|_{1,\tau} \leq C\, (\|\boldsymbol{v}\|_{0,\tau} + h_\tau |\boldsymbol{v}|_{1,\tau}), \tag{4.2}$$

where C is a positive constant depending only on d, ℓ and σ_τ.

Proof. We first follow closely the proof of [23, Lemma 3.1] adapted to the second-level mesh. Define $\mathring{\mathbb{P}}_{\ell+d}(\tau)$ as the space of functions in $\mathbb{P}_{\ell+d}(\tau)$ that vanish on $\partial\tau$ and, analogously, define $\mathring{\mathbb{P}}_{\ell+d}(F)$ as the space of functions in $\mathbb{P}_{\ell+d}(F)$ that are zero on ∂F. Also, let b_τ and b_F be the product of all barycentric coordinates that do not vanish everywhere on τ and F, respectively. The following equivalences hold

$$\mathring{\mathbb{P}}_{\ell+d}(\tau) = b_\tau\, \mathbb{P}_{\ell-1}(\tau) \quad \text{and} \quad \mathring{\mathbb{P}}_{\ell+d}(F) = b_F\, \mathbb{P}_\ell(F), \tag{4.3}$$

and, consequently $\dim(\mathring{\mathbb{P}}_{\ell+d}(\tau)) = \dim(\mathbb{P}_{\ell-1}(\tau))$ and $\dim(\mathring{\mathbb{P}}_{\ell+d}(F)) = \dim(\mathbb{P}_\ell(F))$. Then, we arrive at

$$\dim(\mathbb{P}_{\ell+d}^F(\tau)^d) = (\dim(\mathring{\mathbb{P}}_{\ell+d}(\tau)))^d + (\dim(\mathring{\mathbb{P}}_{\ell+d}(F)))^d, \tag{4.4}$$

and it follows from (4.4) that the system (4.1) is squared. To verify the invertible of the system, it suffices to show that $\boldsymbol{v} = \mathbf{0}$ leads to $\rho_F^\tau \boldsymbol{v} = \mathbf{0}$. For this, observe that $\rho_F^\tau \boldsymbol{v}|_F \in \mathring{\mathbb{P}}_{\ell+d}(F)^d$, and then $\rho_F^\tau \boldsymbol{v}|_F = b_F \boldsymbol{\mu}_\ell$ for some $\boldsymbol{\mu}_\ell \in \mathbb{P}_\ell(F)^d$ from (4.3). Then, using the first equation in (4.1) we find $\rho_F^\tau \boldsymbol{v}$ vanishes on F, so $\rho_F^\tau \boldsymbol{v} = b_\tau \boldsymbol{q}_{\ell-1}$ for

some $\boldsymbol{q}_{\ell-1} \in \mathbb{P}_{\ell-1}(\tau)^d$. Then, the second equation in (4.1) implies that $\rho_F^\tau \boldsymbol{v} = \mathbf{0}$ on τ.

Now, we adapt [10, Lemma 3.1] to prove (4.2). Let $\{\boldsymbol{\varphi}_i\}_{i=1}^{d_1} \subset \mathbb{P}^F_{\ell+d}(\tau)^d$, $\{\boldsymbol{q}_i\}_{i=1}^{d_2} \subset \mathbb{P}_{\ell-1}(\tau)^d$ and $\{\boldsymbol{\mu}_i\}_{i=1}^{d_3} \subset \mathbb{P}_\ell(F)^d$ be basis, where d_1, d_2, $d_3 \in \mathbb{N}^+$ and $d_1 = d_2 + d_3$. Since τ is a simplex, we define an affine transformation $\mathcal{F}_\tau : \hat{\tau} \to \tau$, where $\hat{\tau} \in \mathbb{R}^d$ is a reference element. It is well-known that (see [12, §2.1.3] and [20, lemma 1.100])

$$\int_\tau \boldsymbol{\varphi}_j \cdot \boldsymbol{q}_i \, \mathrm{d}\boldsymbol{x} = \frac{\mathrm{meas}(\tau)}{\mathrm{meas}(\hat{\tau})} \int_{\hat{\tau}} \hat{\boldsymbol{\varphi}}_j \cdot \hat{\boldsymbol{q}}_i \, d\hat{\boldsymbol{x}}\,, \qquad 1 \leqslant i \leqslant d_2\,, \quad 1 \leqslant j \leqslant d_1\,,$$
$$\int_F \boldsymbol{\varphi}_j \cdot \boldsymbol{\mu}_i \, \mathrm{d}\mathbf{s} = |(\nabla(\mathcal{F}_\tau^{-1}))^t \, \hat{\boldsymbol{n}}_F| \, \frac{\mathrm{meas}(\tau)}{\mathrm{meas}(\hat{\tau})} \int_{\hat{F}} \hat{\boldsymbol{\varphi}}_j \cdot \hat{\boldsymbol{\mu}}_i \, \mathrm{d}\hat{\mathbf{s}}\,, \qquad 1 \leqslant i \leqslant d_3\,, \quad 1 \leqslant j \leqslant d_1\,,$$

where $\hat{\boldsymbol{n}}_F$ represents the outward normal vector on $\hat{\tau}$ restricted to the face $\hat{F}$. Next, define the matrices $\hat{A} := (\hat{a}_{ij})_{i,j} \in \mathbb{R}^{d_2 \times d_1}$ and $\hat{B} := (\hat{b}_{ij})_{i,j} \in \mathbb{R}^{d_3 \times d_1}$, where

$$\hat{a}_{ij} := \frac{1}{\mathrm{meas}(\hat{\tau})} \int_{\hat{\tau}} \hat{\boldsymbol{\varphi}}_j \cdot \hat{\boldsymbol{q}}_i \, d\hat{\boldsymbol{x}}\,,$$
$$\hat{b}_{ij} := \frac{1}{\mathrm{meas}(\hat{\tau})} \int_{\hat{F}} \hat{\boldsymbol{\varphi}}_j \cdot \hat{\boldsymbol{\mu}}_i \, \mathrm{d}\mathbf{s}\,,$$

and set

$$\mathbb{A} := \mathrm{meas}(\tau) \begin{pmatrix} I_{d_2 \times d_2} & \mathbf{0} \\ \mathbf{0} & |(\nabla(\mathcal{F}_\tau^{-1}))^t \, \hat{\boldsymbol{n}}_F| \, I_{d_3 \times d_3} \end{pmatrix} \begin{pmatrix} \hat{A} \\ \hat{B} \end{pmatrix} \quad \in \mathbb{R}^{d_1 \times d_1}\,,$$

where $I_{d_2 \times d_2}$ (resp. $I_{d_3 \times d_3}$) is the identity matrix in $\mathbb{R}^{d_2 \times d_2}$ (resp. $\mathbb{R}^{d_3 \times d_3}$). Given $\boldsymbol{v} \in \boldsymbol{H}^1(\tau)$, we set $\mathbf{f} := (f_i)_i \in \mathbb{R}^{d_2}$ and $\mathbf{g} := (g_i)_i \in \mathbb{R}^{d_3}$ as follows

$$\begin{aligned} f_i &:= \int_\tau \boldsymbol{v} \cdot \boldsymbol{q}_i \, d\boldsymbol{x}\,, \quad 1 \leqslant i \leqslant d_2\,, \\ g_i &:= \int_F \boldsymbol{v} \cdot \boldsymbol{\mu}_i \, \mathrm{d}\mathbf{s}\,, \quad 1 \leqslant i \leqslant d_3\,, \end{aligned} \tag{4.5}$$

and look for $\alpha = (\alpha_i)_i \in \mathbb{R}^{d_1}$, the solution of the following algebraic system

$$\mathbb{A}\,\alpha \;=\; \begin{pmatrix} \mathbf{f} \\ \mathbf{g} \end{pmatrix}.$$

Owing to the previous definitions, and using that $\mathbb{A}$ is invertible, we estimate the norm of the coordinate vector α as follows

$$\begin{aligned} \sqrt{\sum_{i=1}^{d_1} \alpha_i^2} = \left| \mathbb{A}^{-1} \begin{pmatrix} \mathbf{f} \\ \mathbf{g} \end{pmatrix} \right| &\leqslant \mathrm{meas}(\tau)^{-1} \left| \begin{pmatrix} \hat{A} \\ \hat{B} \end{pmatrix}^{-1} \right|_{max} \cdot \left| \begin{pmatrix} \mathbf{f} \\ |(\nabla(\mathcal{F}_\tau^{-1}))^t \, \hat{\boldsymbol{n}}_F|^{-1} \mathbf{g} \end{pmatrix} \right| \\ &\leqslant \frac{C}{\mathrm{meas}(\tau)} \max_{\substack{1 \leqslant i \leqslant d_2 \\ 1 \leqslant j \leqslant d_3}} \left\{ |f_i|,\ \frac{|g_j|}{|(\nabla(\mathcal{F}_\tau^{-1}))^t \, \hat{\boldsymbol{n}}_F|} \right\}, \end{aligned} \tag{4.6}$$

where $|\cdot|_{max}$ stands for the norm of maximum in the codomain of $\mathbb{A}$, and C denotes a constant depending only on d and ℓ. Now, we estimate the right-hand side of (4.6). For that, we select the nodal Lagrangian basis $\{\boldsymbol{q}_i\}_i$ and $\{\boldsymbol{\mu}_i\}_i$ such that $|\boldsymbol{q}_i(\cdot)| \leqslant 1$ and $|\boldsymbol{\mu}_i(\cdot)| \leqslant 1$. Then, using Cauchy-Schwarz inequality on (4.5) and the nodal Lagrangian basis we obtain

$$|f_i| \leqslant \|\boldsymbol{v}\|_{0,\tau}\|\boldsymbol{q}_i\|_{0,\tau} = \|\boldsymbol{v}\|_{0,\tau}\sqrt{\int_\tau |\boldsymbol{q}_i(\boldsymbol{x})|^2 \,\mathrm{d}\boldsymbol{x}} \leqslant \mathrm{meas}(\tau)^{\frac{1}{2}}\|\boldsymbol{v}\|_{0,\tau}\,.$$

To bound the other term, we follow the same path to get

$$\begin{aligned} |g_j| &= |(\nabla(\mathcal{F}_\tau^{-1}))^t\,\hat{\boldsymbol{n}}_F|\,\frac{\mathrm{meas}(\tau)}{\mathrm{meas}(\hat{\tau})}\left|\int_{\hat{F}} \hat{\boldsymbol{v}}\cdot\hat{\boldsymbol{\mu}}_j \,\mathrm{d}\hat{\mathbf{s}}\right| \\ &\leqslant |(\nabla(\mathcal{F}_\tau^{-1}))^t\,\hat{\boldsymbol{n}}_F|\,\frac{\mathrm{meas}(\tau)}{\mathrm{meas}(\hat{\tau})}\,\mathrm{meas}(\hat{F})^{\frac{1}{2}}\,\|\hat{\boldsymbol{v}}\|_{0,\hat{F}} \quad (|\hat{\boldsymbol{\mu}}_j(\cdot)| \leqslant 1)\,. \end{aligned} \tag{4.7}$$

Now, using the trace inequality (2.14) and scaling arguments, we obtain

$$\begin{aligned} \|\hat{\boldsymbol{v}}\|_{0,\hat{F}} &\leqslant \sigma_{\hat{\tau}}^{\frac{1}{2}}\,\left(d\,h_{\hat{\tau}}^{-1}\,\|\hat{\boldsymbol{v}}\|_{0,\hat{\tau}} + 2\,\|\nabla\hat{\boldsymbol{v}}\|_{0,\hat{\tau}}\right)^{\frac{1}{2}}\,\|\hat{\boldsymbol{v}}\|_{0,\hat{\tau}}^{\frac{1}{2}} \\ &\leqslant \frac{\mathrm{meas}(\hat{\tau})^{\frac{1}{2}}}{\mathrm{meas}(\tau)^{\frac{1}{2}}}\,\sigma_{\hat{\tau}}^{\frac{1}{2}}\,\left(\frac{d}{h_{\hat{\tau}}}\,\|\boldsymbol{v}\|_{0,\tau} + 2\,\frac{h_\tau}{\rho_{\hat{\tau}}}\,\|\nabla\boldsymbol{v}\|_{0,\tau}\right)^{\frac{1}{2}}\,\|\boldsymbol{v}\|_{0,\tau}^{\frac{1}{2}} \\ &\leqslant \frac{\mathrm{meas}(\hat{\tau})^{\frac{1}{2}}}{\mathrm{meas}(\tau)^{\frac{1}{2}}}\,\left(\frac{d}{\rho_{\hat{\tau}}}\,\|\boldsymbol{v}\|_{0,\tau} + 2\,\sigma_{\hat{\tau}}\,\frac{h_\tau}{\rho_{\hat{\tau}}}\,\|\nabla\boldsymbol{v}\|_{0,\tau}\right)^{\frac{1}{2}}\,\|\boldsymbol{v}\|_{0,\tau}^{\frac{1}{2}} \\ &\leqslant \frac{\mathrm{meas}(\hat{\tau})^{\frac{1}{2}}}{\mathrm{meas}(\tau)^{\frac{1}{2}}}\,\left(\frac{1}{2}\left(\frac{d}{\rho_{\hat{\tau}}}+1\right)\,\|\boldsymbol{v}\|_{0,\tau} + \sigma_{\hat{\tau}}\,\frac{h_\tau}{\rho_{\hat{\tau}}}\,\|\nabla\boldsymbol{v}\|_{0,\tau}\right)\,, \end{aligned} \tag{4.8}$$

and from (4.7) and (4.8), we arrive at

$$\frac{|g_j|}{|(\nabla(\mathcal{F}_\tau^{-1}))^t\,\hat{\boldsymbol{n}}_F|} \leqslant C\,\mathrm{meas}(\tau)^{\frac{1}{2}}\,\left(\|\boldsymbol{v}\|_{0,\tau} + h_\tau\,\|\nabla\boldsymbol{v}\|_{0,\tau}\right)\,,$$

where C depends only on d. Therefore, gathering the previous estimates, we conclude

$$\mathrm{meas}(\tau)\sum_{i=1}^{d_1}\alpha_i^2 \leqslant C\left(\|\boldsymbol{v}\|_{0,\tau}^2 + h_\tau^2|\boldsymbol{v}|_{1,\tau}^2\right)\,.$$

Now, using Cauchy-Schwarz inequality and a nodal Lagrangian basis for $\{\boldsymbol{\varphi}_i\}_i$ such that $|\boldsymbol{\varphi}_i(\cdot)| \leqslant 1$, it holds

$$\begin{aligned} \|\rho_F^\tau\,\boldsymbol{v}\|_{0,\tau}^2 = \int_\tau\left|\sum_{i=1}^{d_1}\alpha_i\boldsymbol{\varphi}_i\right|^2 &\leqslant \int_\tau\left(\sum_{i=1}^{d_1}\alpha_i^2\right)\left(\sum_{i=1}^{d_1}|\boldsymbol{\varphi}_i|^2\right) \leqslant C\,\mathrm{meas}(\tau)\left(\sum_{i=1}^{d_1}\alpha_i^2\right) \\ &\leqslant C\left(\|\boldsymbol{v}\|_{0,\tau}^2 + h_\tau^2|\boldsymbol{v}|_{1,\tau}^2\right)\,. \end{aligned} \tag{4.9}$$

To conclude the result, we use the following inverse inequality [20, lemma 1.138]

$$h_\tau \, |\rho_F^\tau \, \boldsymbol{v}|_{1,\tau} \leqslant C \|\rho_F^\tau \, \boldsymbol{v}\|_{0,\tau} \, , \tag{4.10}$$

and thus (4.2) follows from (4.9) and (4.10). ∎

Let $\mathscr{C}_h : \boldsymbol{H}^1(\mathscr{P}) \to \mathbf{V}_h$ be the Clément interpolation operator defined locally. In other words, for every $\boldsymbol{v} \in \mathbf{V}$ we define $\mathscr{C}_h(\boldsymbol{v})|_K := \mathscr{C}_h^K(\boldsymbol{v})$ where $\mathscr{C}_h^K : \boldsymbol{H}^1(K) \to \mathbf{V}_h(K)$ is the usual Clément interpolation operator. It is well-known that the operator $\mathscr{C}_h^K$ satisfies the following two properties (see [11]):

(*i*) there exists $C > 0$ such that

$$\|\mathscr{C}_h^K(\boldsymbol{v})\|_{1,K} \leqslant C \, \|\boldsymbol{v}\|_{1,K} \, , \tag{4.11}$$

(*ii*) for m and s satisfying $0 \leqslant m \leqslant s$, with $s = 0$, 1, there is $C > 0$ such that

$$\|\boldsymbol{v} - \mathscr{C}_h^K(\boldsymbol{v})\|_{m,\tau} \leq C \, h_\tau^{s-m} \, \|\boldsymbol{v}\|_{s,\omega_\tau^K} \, , \tag{4.12}$$

for all $\boldsymbol{v} \in \boldsymbol{H}^s(\omega_\tau^K)$ and all $\tau \in \Xi_H^K$, where $\omega_\tau^K := \{\tau' \in \Xi_H^K \, : \, \tau \cap \tau' \neq \emptyset\}$. The constants C depend only on k, d, and σ_Ξ.

We are ready to present the global Fortin operator we use in this work.

Lemma 4.2. *Consider that Assumption (A) holds, and $k-d \geq \ell \geq 1$. There exists a mapping $\Pi_h : \mathbf{V} \to \mathbf{V}_h$ such that, for all $\boldsymbol{v} \in \mathbf{V}$, it satisfies*

$$\begin{aligned} &\int_F \Pi_h(\boldsymbol{v}) \cdot \boldsymbol{\mu}_H \, d\mathbf{s} = \int_F \boldsymbol{v} \cdot \boldsymbol{\mu}_H \, d\mathbf{s} \quad \textit{for all } \boldsymbol{\mu}_H \in \boldsymbol{\Lambda}_H \quad \textit{and } \; F \in \mathcal{E}_H \, , \\ &\|\Pi_h(\boldsymbol{v})\|_{1,\mathscr{P}} \leq C \, \|\boldsymbol{v}\|_{1,\mathscr{P}} \, , \end{aligned} \tag{4.13}$$

where C is a positive constant depending only on ℓ, k, d and σ_Ξ.

Proof. Let $K \in \mathscr{P}$ and $F \subset \partial K \cap \mathcal{E}_H$. From Assumption (A), there exists $\tau_F \in \Xi_H^K$ such that $\tau_F \cap \partial K = F$. Then, we define $\tilde{\rho}_F^K : \boldsymbol{H}^1(K) \to \mathbf{V}_h(K)$ as follow

$$\tilde{\rho}_F^K(\boldsymbol{v}) := \begin{cases} \rho_F^{\tau_F}(\boldsymbol{v}) & \text{on } \tau_F \, , \\ \mathbf{0} & \text{otherwise} \, , \end{cases}$$

where $\rho_F^{\tau_F}(\boldsymbol{v})$ is given in Lemma 4.1. Notice that $\tilde{\rho}_F^K(\boldsymbol{v}) \in \mathbf{V}_h(K)$ since $\mathcal{T}_h^K$ is obtained from refinements of Ξ_H^K. Now, let Π_h be the operator defined by its local counterpart

$$\Pi_h(\boldsymbol{v}) \, |_K := \mathscr{C}_h^K(\boldsymbol{v}) + \sum_{F \in \mathcal{E}_H \cap \partial K} \tilde{\rho}_F^K \Big(\boldsymbol{v} - \mathscr{C}_h^K(\boldsymbol{v}) \Big) \, .$$

From Lemma 4.1, we get for all $\boldsymbol{\mu}_H \in \boldsymbol{\Lambda}_H$ that

$$\begin{aligned}\int_F \Pi_h(\boldsymbol{v})\cdot\boldsymbol{\mu}_H \,\mathrm{d}\mathbf{s} &= \int_F \mathscr{C}_h^K(\boldsymbol{v})\cdot\boldsymbol{\mu}_H \,\mathrm{d}\mathbf{s} + \int_F \tilde{\rho}_F^K\Big(\boldsymbol{v}-\mathscr{C}_h^K(\boldsymbol{v})\Big)\cdot\boldsymbol{\mu}_H \,\mathrm{d}\mathbf{s}\\ &= \int_F \mathscr{C}_h^K(\boldsymbol{v})\cdot\boldsymbol{\mu}_H \,\mathrm{d}\mathbf{s} + \int_F \rho_F^{\tau_F}\Big(\boldsymbol{v}-\mathscr{C}_h^K(\boldsymbol{v})\Big)\cdot\boldsymbol{\mu}_H \,\mathrm{d}\mathbf{s}\\ &= \int_F \mathscr{C}_h^K(\boldsymbol{v})\cdot\boldsymbol{\mu}_H \,\mathrm{d}\mathbf{s} + \int_F \Big(\boldsymbol{v}-\mathscr{C}_h^K(\boldsymbol{v})\Big)\cdot\boldsymbol{\mu}_H \,\mathrm{d}\mathbf{s}\\ &= \int_F \boldsymbol{v}\cdot\boldsymbol{\mu}_H \,\mathrm{d}\mathbf{s}\,,\end{aligned}$$

which leads to the first condition in (4.13). Next, using (4.2), (4.11), and (4.12) we get

$$\begin{aligned}\|\Pi_h(\boldsymbol{v})\|_{1,K}^2 &\le \Big(\|\mathscr{C}_h^K(\boldsymbol{v})\|_{1,K}^2 + \sum_{F\in\mathcal{E}_H\cap\partial K}\|\tilde{\rho}_F^K(\boldsymbol{v}-\mathscr{C}_h^K(\boldsymbol{v}))\|_{1,K}^2\Big)\\ &\le \Big(C\|\boldsymbol{v}\|_{1,K}^2 + \sum_{F\in\mathcal{E}_H\cap\partial K}\|\rho_F^{\tau_F}(\boldsymbol{v}-\mathscr{C}_h^K(\boldsymbol{v}))\|_{1,\tau_F}^2\Big)\\ &\le C\,\Big(\|\boldsymbol{v}\|_{1,K}^2 + \sum_{F\in\mathcal{E}_H\cap\partial K}\Big(\|\rho_F^{\tau_F}(\boldsymbol{v}-\mathscr{C}_h^K(\boldsymbol{v}))\|_{0,\tau_F}^2\\ &\qquad + |\rho_F^{\tau_F}(\boldsymbol{v}-\mathscr{C}_h^K(\boldsymbol{v}))|_{1,\tau_F}^2\Big)\\ &\le C\,\Big(\|\boldsymbol{v}\|_{1,K}^2 + \sum_{F\in\mathcal{E}_H\cap\partial K}\Big(\|\boldsymbol{v}-\mathscr{C}_h^K(\boldsymbol{v})\|_{0,\tau_F}^2 + h_{\tau_F}^2|\boldsymbol{v}-\mathscr{C}_h^K(\boldsymbol{v})|_{1,\tau_F}^2\\ &\qquad + \frac{1}{h_{\tau_F}^2}\|\boldsymbol{v}-\mathscr{C}_h^K(\boldsymbol{v})\|_{0,\tau_F}^2 + |\boldsymbol{v}-\mathscr{C}_h^K(\boldsymbol{v})|_{1,\tau_F}^2\Big)\Big)\\ &\le C\,\Big(\|\boldsymbol{v}\|_{1,K}^2 + \sum_{F\in\mathcal{E}_H\cap\partial K}\|\boldsymbol{v}\|_{1,\omega_{\tau_F}^K}^2\Big)\,. \qquad (4.14)\end{aligned}$$

From the Assumption (A), there exists a constant $C>0$, depending only on k, d and σ_Ξ, such that

$$\sum_{F\in\mathcal{E}_H\cap\partial K}\|\boldsymbol{v}\|_{1,\omega_{\tau_F}^K}^2 \leqslant (d+1)\sum_{\tau\in\Xi_H^K}\|\boldsymbol{v}\|_{1,\omega_\tau^K}^2 \leqslant C\,\|\boldsymbol{v}\|_{1,K}^2 \quad \text{for all } \boldsymbol{v}\in\boldsymbol{H}^1(K)\,. \qquad (4.15)$$

The second condition in (4.13) follows from summing (4.14) over $K\in\mathscr{P}$ and using (4.15). ∎

Remark 4.3. The constraint $k-d\ge\ell\ge 1$ in Lemma 4.2 guarantees the existence of a Fortin mapping (4.13), which leads to the well-posedness of (3.4) in two and three-dimensional problems on polytopal meshes (Theorem 4.4). As for the three-dimensional simplicial partitions with one-element local meshes, we recognize the compatibility condition proposed in [23]. For the two-dimensional case, other possibilities exist to ensure well-posedness of (3.4), such as

- simplicial meshes with one-element local meshes, and even polynomial degree ℓ under the constraint $k = \ell + 1$ [40];
- quadrilateral meshes with one-element local meshes under the constraint $k \geq \ell + 2$ [28];
- polytopal meshes with "enough" refined local meshes under the conditions $k = \ell + 1$ or $k = \ell$. The proof follows from Lemma 4 and Lemma 5 in [9] applied to each component of the vector function $\boldsymbol{\mu}_H \in \boldsymbol{\Lambda}_H$.

□

4.2. Existence and uniqueness. Before addressing the main well-posedness result, we introduce two additional discrete spaces as follows.

$$\begin{aligned} \boldsymbol{\Lambda}_{\mathrm{rm}} &:= \{\boldsymbol{\mu} \in \boldsymbol{\Lambda} \,:\, \boldsymbol{\mu}|_{\partial K} \in \boldsymbol{\Lambda}_{\mathrm{rm}}(K),\ \forall K \in \mathscr{P}\}\,, \quad \text{where} \\ \boldsymbol{\Lambda}_{\mathrm{rm}}(K) &:= \left\{\boldsymbol{\mu}_{\mathrm{rm}} \in \boldsymbol{L}^2(\partial K) \,:\, \boldsymbol{\mu}_{\mathrm{rm}}|_E = \boldsymbol{v}^{\mathrm{rm}}|_E,\ \boldsymbol{v}^{\mathrm{rm}} \in \mathbf{V}_{\mathrm{rm}}(K),\ E \in \mathcal{E} \cap \partial K\right\}, \end{aligned} \tag{4.16}$$

and $\boldsymbol{\Lambda}_{\mathrm{rm}} \subset \boldsymbol{\Lambda}_H$ as $\ell \geqslant 1$. The well-posedness of the MHM method (3.4) is proved in the following theorem, which generalizes [28, Theorem 6.2].

Theorem 4.4. *Suppose Assumption (A) holds, and $k \geq \ell + d$ and $\ell \geqslant 1$. Then, there exist positive constants α_0 and β_0, independent of any mesh diameters, such that*

$$a_h(\boldsymbol{\mu}_H, \boldsymbol{\mu}_H) \geq \alpha_0 \, \|\boldsymbol{\mu}_H\|^2_{\boldsymbol{\Lambda}} \quad \textit{for all } \boldsymbol{\mu}_H \in \mathscr{N}_H\,, \tag{4.17}$$

$$\sup_{\boldsymbol{\mu}_H \in \boldsymbol{\Lambda}_H} \frac{b(\boldsymbol{\mu}_H, \boldsymbol{v}^{\mathrm{rm}})}{\|\boldsymbol{\mu}_H\|_{\boldsymbol{\Lambda}}} \geq \beta_0 \, \|\boldsymbol{v}^{\mathrm{rm}}\|_{\mathbf{V}} \quad \textit{for all } \boldsymbol{v}^{\mathrm{rm}} \in \mathbf{V}_{\mathrm{rm}}\,, \tag{4.18}$$

where $\mathscr{N}_H$ is the discrete kernel of $b(\cdot,\cdot)$, i.e.,

$$\mathscr{N}_H := \{\boldsymbol{\mu}_H \in \boldsymbol{\Lambda}_H : b(\boldsymbol{\mu}_H, \boldsymbol{v}^{\mathrm{rm}}) = 0 \quad \textit{for all } \boldsymbol{v}^{\mathrm{rm}} \in \mathbf{V}_{\mathrm{rm}}\}\,, \tag{4.19}$$

and then (3.4) is well-posed.

Proof. Let $\boldsymbol{\mu}_H \in \mathscr{N}_H$. From local problem (2.22), it holds

$$\int_K \mathbf{C}\,\underline{\boldsymbol{\varepsilon}}(T_h\boldsymbol{\mu}_H) : \underline{\boldsymbol{\varepsilon}}(\boldsymbol{v}_h)\,\mathrm{d}\boldsymbol{x} = \langle \boldsymbol{\mu}_H, \boldsymbol{v}_h\rangle_{\partial K} \quad \text{for all } \boldsymbol{v}_h \in \mathbf{V}_h\,. \tag{4.20}$$

Next, given $\boldsymbol{v} \in \mathbf{V}$, set $\tilde{\boldsymbol{v}} = (I - \Pi_{RM})\boldsymbol{v} \in \widetilde{\mathbf{V}}$, where Π_{RM} is the operator defined in (2.12). Hence, there is a positive constant C, depending on h_Ω, c_{max}, c_{min}, C_{korn}, ℓ,

k, d and σ_Ξ but independent of $\mathcal{H}$, H and h, such that

$$
\begin{aligned}
\|\boldsymbol{\mu}_H\|_{\boldsymbol{\Lambda}} &= \sup_{\boldsymbol{v}\in\mathbf{V}} \frac{\langle \boldsymbol{\mu}_H, \boldsymbol{v}\rangle_{\partial\mathscr{P}}}{\|\boldsymbol{v}\|_{\mathbf{V}}} && \text{(using (2.9))}\\
&\le \sup_{\tilde{\boldsymbol{v}}\in\tilde{\mathbf{V}}} \frac{\langle \boldsymbol{\mu}_H, \tilde{\boldsymbol{v}}\rangle_{\partial\mathscr{P}}}{\|\tilde{\boldsymbol{v}}\|_{\mathbf{V}}} && \text{(using } \boldsymbol{\mu}_H \in \mathscr{N}_H \text{ and (2.13))}\\
&\le C \sup_{\tilde{\boldsymbol{v}}\in\tilde{\mathbf{V}}} \frac{\langle \boldsymbol{\mu}_H, \tilde{\boldsymbol{v}}\rangle_{\partial\mathscr{P}}}{\|\tilde{\boldsymbol{v}}\|_{1,\mathscr{P}}} && \text{(using (2.10))}\\
&\le C \sup_{\tilde{\boldsymbol{v}}\in\tilde{\mathbf{V}}} \frac{\langle \boldsymbol{\mu}_H, \Pi_h(\tilde{\boldsymbol{v}})\rangle_{\partial\mathscr{P}}}{\|\Pi_h(\tilde{\boldsymbol{v}})\|_{1,\mathscr{P}}} && \text{(using (4.13))}\\
&\le C \sup_{\tilde{\boldsymbol{v}}\in\tilde{\mathbf{V}}} \frac{\langle \boldsymbol{\mu}_H, \Pi_h(\tilde{\boldsymbol{v}})\rangle_{\partial\mathscr{P}}}{\|\Pi_h(\tilde{\boldsymbol{v}})\|_{\mathbf{V}}}\\
&\le C \sup_{\boldsymbol{v}_h\in\mathbf{V}_h} \frac{\langle \boldsymbol{\mu}_H, \boldsymbol{v}_h\rangle_{\partial\mathscr{P}}}{\|\boldsymbol{v}_h\|_{\mathbf{V}}}\\
&= C \sup_{\boldsymbol{v}_h\in\mathbf{V}_h} \frac{\sum_{K\in\mathscr{P}} \int_K \mathbf{C}\,\underline{\boldsymbol{\varepsilon}}(T_h\boldsymbol{\mu}_H) : \underline{\boldsymbol{\varepsilon}}(\boldsymbol{v}_h)\,\mathrm{d}\boldsymbol{x}}{\|\boldsymbol{v}_h\|_{\mathbf{V}}} && \text{(using (4.20))}\\
&\le C\, c_{max} \Big(\sum_{K\in\mathscr{P}} \|\underline{\boldsymbol{\varepsilon}}(T_h\boldsymbol{\mu}_H)\|_{0,K}^2 \Big)^{1/2}.
\end{aligned}
$$

Hence, using the definition of $a_h(\cdot,\cdot)$ in (3.5), we arrive at

$$
\begin{aligned}
a_h(\boldsymbol{\mu}_H,\boldsymbol{\mu}_H) = \langle \boldsymbol{\mu}_H, T_h\boldsymbol{\mu}_H\rangle_{\partial\mathscr{P}} &= \sum_{K\in\mathscr{P}} \int_K \mathbf{C}\,\underline{\boldsymbol{\varepsilon}}(T_h\boldsymbol{\mu}_H) : \underline{\boldsymbol{\varepsilon}}(T_h\boldsymbol{\mu}_H)\,\mathrm{d}\boldsymbol{x}\\
&\ge c_{min} \sum_{K\in\mathscr{P}} \|\underline{\boldsymbol{\varepsilon}}(T_h\boldsymbol{\mu}_H)\|_{0,K}^2 \ge C\,\frac{c_{min}}{c_{max}^2}\,\|\boldsymbol{\mu}_H\|_{\boldsymbol{\Lambda}}^2,
\end{aligned}
$$

which proves (4.17). As for (4.18), we use a slightly different version of the proof proposed in [28] in the case of polytope partitions. We give the details next for completeness. Define the local interpolation $\mathcal{I}_K$ on functions in $\boldsymbol{L}^2(\partial K)$ with value in $\boldsymbol{\Lambda}_{\mathrm{rm}}(K)$ such that, for each $E\in\mathcal{E}$, it holds

$$
\int_E \mathcal{I}_K\boldsymbol{\mu}\cdot\boldsymbol{v}^{\mathrm{rm}}\,\mathrm{d}\mathbf{s} = \int_E \boldsymbol{\mu}\cdot\boldsymbol{v}^{\mathrm{rm}}\,\mathrm{d}\mathbf{s} \quad \text{for all } \boldsymbol{v}^{\mathrm{rm}}\in\mathbf{V}_{\mathrm{rm}}(K). \tag{4.21}
$$

Observe that Equation (4.21) is well-defined in $\boldsymbol{\Lambda}_{\mathrm{rm}}(K)$ and the following local stability result holds

$$
\|\mathcal{I}_K\boldsymbol{\mu}\|_{0,\partial K} \le \|\boldsymbol{\mu}\|_{0,\partial K} \quad \text{for all } \boldsymbol{\mu}\in\boldsymbol{L}^2(\partial K). \tag{4.22}
$$

The global interpolation $\mathcal{I}$ acts on $\boldsymbol{\Lambda}\cap\boldsymbol{L}^2(\partial\mathscr{P})$ and has values in $\boldsymbol{\Lambda}_{\mathrm{rm}}$, and is fully defined assuming $\mathcal{I}|_K := \mathcal{I}_K$. Next, using (2.13) and (4.21), we obtain, for all $\boldsymbol{\mu}\in\boldsymbol{\Lambda}$ and $\boldsymbol{v}\in\mathbf{V}$,

$$
\langle \mathcal{I}\boldsymbol{\mu}, \Pi_{RM}\boldsymbol{v}\rangle_{\partial\mathscr{P}} = \langle \boldsymbol{\mu}, \Pi_{RM}\boldsymbol{v}\rangle_{\partial\mathscr{P}} \leqslant \|\boldsymbol{\mu}\|_{\boldsymbol{\Lambda}}\,\|\Pi_{RM}\boldsymbol{v}\|_{\mathbf{V}} \leqslant \|\boldsymbol{\mu}\|_{\boldsymbol{\Lambda}}\,\|\boldsymbol{v}\|_{\mathbf{V}}, \tag{4.23}
$$

and

$$\begin{aligned}
\langle \mathcal{I}\boldsymbol{\mu}, \boldsymbol{v} - \Pi_{RM}\boldsymbol{v}\rangle_{\partial\mathscr{P}} &\leq \sum_{K\in\mathscr{P}} \|\mathcal{I}_K\boldsymbol{\mu}\|_{0,\partial K}\|\boldsymbol{v} - \Pi_{RM}\,\boldsymbol{v}\|_{0,\partial K} && \text{(Cauchy-Schwarz ineq.)}\\
&\leq C \sum_{K\in\mathscr{P}} h_K^{1/2}\,\|\boldsymbol{\mu}\|_{0,\partial K}\|\nabla\boldsymbol{v}\|_{0,K} && \text{(using (4.22) and (2.17))}\\
&\leq C\,C_{korn} \sum_{K\in\mathscr{P}} \|\boldsymbol{\mu}\|_{0,\partial K}\,h_K^{1/2}\|\underline{\boldsymbol{\varepsilon}}(\boldsymbol{v})\|_{0,K} && \text{(using (2.10))}\\
&\leq C\Big[\sum_{K\in\mathscr{P}} \|\boldsymbol{\mu}\|_{0,\partial K}^2\,h_K\Big]^{1/2}\|\boldsymbol{v}\|_{\mathbf{V}}\,, && \text{(Cauchy-Schwarz ineq.)} \qquad (4.24)
\end{aligned}$$

where C is a positive constant depending on d and the constants in (2.18).

Now, given $\boldsymbol{v}^{\mathrm{rm}} \in \mathbf{V}_{\mathrm{rm}}$ let $\boldsymbol{\sigma}^* \in H^1(\Omega)^{d\times d}$ be the symmetric tensor satisfying (2.32), define $\boldsymbol{\mu}^*|_{\partial K} := \boldsymbol{\sigma}^*\,\boldsymbol{n}^K\,|_{\partial K}$ for all $K \in \mathscr{P}$. Observe that $\boldsymbol{\mu}^* \in \boldsymbol{\Lambda} \cap \boldsymbol{L}^2(\partial\mathscr{P})$, and then, from (2.8), (2.9), and (2.32) it holds

$$\|\boldsymbol{v}^{\mathrm{rm}}\|_{0,\Omega} = \frac{(\boldsymbol{v}^{\mathrm{rm}},\boldsymbol{v}^{\mathrm{rm}})_\Omega}{\|\boldsymbol{v}^{\mathrm{rm}}\|_{0,\Omega}} = h_\Omega\,\frac{(\nabla\cdot\boldsymbol{\sigma}^*,\boldsymbol{v}^{\mathrm{rm}})_\Omega}{\|\boldsymbol{v}^{\mathrm{rm}}\|_{\mathbf{V}}} = h_\Omega\,\frac{\langle\boldsymbol{\mu}^*,\boldsymbol{v}^{\mathrm{rm}}\rangle_{\partial\mathscr{P}}}{\|\boldsymbol{v}^{\mathrm{rm}}\|_{\mathbf{V}}} \leqslant h_\Omega\,\|\boldsymbol{\mu}^*\|_{\boldsymbol{\Lambda}}\,. \qquad (4.25)$$

Using (2.14) and (2.32), and (4.25) there exists a positive constant C, depending only on Ω and on the parameters (2.19) associated with the element $K \in \mathscr{P}$, such that

$$\begin{aligned}
\Big[\sum_{K\in\mathscr{P}} \|\boldsymbol{\mu}^*\|_{0,\partial K}^2\,h_K\Big]^{1/2} &\leqslant \Big[\sum_{K\in\mathscr{P}} C\,\Big(\big(d\,h_K^{-1}+2\big)\,\|\boldsymbol{\sigma}^*\|_{0,K}^2 + \|\nabla\boldsymbol{\sigma}^*\|_{0,K}^2\Big)\,h_K\Big]^{1/2}\\
&\leqslant C\,\|\boldsymbol{\sigma}^*\|_{1,\Omega} \leqslant C\,\|\boldsymbol{v}^{\mathrm{rm}}\|_{0,\Omega} \leqslant C\,\|\boldsymbol{\mu}^*\|_{\boldsymbol{\Lambda}}\,. \qquad (4.26)
\end{aligned}$$

Next, we replace (4.26) in (4.24) and add the resulting estimate to (4.23), to get

$$\|\mathcal{I}\boldsymbol{\mu}^*\|_{\boldsymbol{\Lambda}} \leqslant C\,\|\boldsymbol{\mu}^*\|_{\boldsymbol{\Lambda}}\,. \qquad (4.27)$$

Finally, using (2.33), (4.21), and (4.27), it holds

$$\|\boldsymbol{v}^{\mathrm{rm}}\|_{\mathbf{V}} \leqslant C\,\frac{b(\boldsymbol{\mu}^*,\boldsymbol{v}^{\mathrm{rm}})}{\|\boldsymbol{\mu}^*\|_{\boldsymbol{\Lambda}}} \leqslant C\,\frac{b(\mathcal{I}(\boldsymbol{\mu}^*),\boldsymbol{v}^{\mathrm{rm}})}{\|\mathcal{I}(\boldsymbol{\mu}^*)\|_{\boldsymbol{\Lambda}}} \leqslant C\sup_{\boldsymbol{\mu}_{\mathrm{rm}}\in\boldsymbol{\Lambda}_{\mathrm{rm}}}\frac{b(\boldsymbol{\mu}_{\mathrm{rm}},\boldsymbol{v}^{\mathrm{rm}})}{\|\boldsymbol{\mu}_{\mathrm{rm}}\|_{\boldsymbol{\Lambda}}}\,,$$

where we used that $\boldsymbol{\Lambda}_{\mathrm{rm}} \subset \boldsymbol{\Lambda}_H$, and (4.18) follows. The well-posedness of (3.4) arises from the saddle-point theory. ∎

Remark 4.5. The results in Theorem 4.4 still hold if one replaces $\boldsymbol{\Lambda}_H$ by $\boldsymbol{\Lambda}_{\mathrm{rm}}$ and assumes $k \geq d+1$.

5. Error Estimates. This section presents a priori error estimates for the MHM method (3.4). The properties of the method yields a multi-level error analysis; therefore, we split interpolation results in global one-level and local second-level interpolation estimates.

5.1. Interpolation estimates. To begin, the next lemma presents the approximation properties of the global space $\boldsymbol{\Lambda}_H$ in terms of the characteristic size H. It corresponds to the extension of [9] to the elasticity case.

Lemma 5.1. *Suppose Assumption (A) holds and let* $\boldsymbol{w} \in \boldsymbol{H}^{m+2}(\mathscr{P})$, *for* $0 \leqslant m \leqslant \ell$ *and* $\ell \geqslant 1$, *such that* $\mathbf{C}\,\underline{\boldsymbol{\varepsilon}}(\boldsymbol{w}) \in H^{m+1}(\mathscr{P})^{d\times d}$. *Defining* $\boldsymbol{\mu} \in \boldsymbol{\Lambda}$ *such that* $\boldsymbol{\mu}\,|_{\partial K} := \mathbf{C}\,\underline{\boldsymbol{\varepsilon}}(\boldsymbol{w})\boldsymbol{n}^K\,|_{\partial K}$ *for each* $K \in \mathscr{P}$, *there exists a positive constant* C, *depending only on* d, σ_Ξ, ℓ *and* C_{korn}, *such that*

$$\inf_{\boldsymbol{\mu}_\ell \in \boldsymbol{\Lambda}_H} \|\boldsymbol{\mu} - \boldsymbol{\mu}_\ell\|_{\boldsymbol{\Lambda}} \leq C\, H^{m+1}\, |\mathbf{C}\,\underline{\boldsymbol{\varepsilon}}(\boldsymbol{w})|_{m+1,\mathscr{P}}\,, \tag{5.1}$$

where $\boldsymbol{\Lambda}_H$ *is given in* (3.3). *Moreover, it holds*

$$\inf_{\boldsymbol{\mu}_{\rm rm} \in \boldsymbol{\Lambda}_{\rm rm}} \|\boldsymbol{\mu} - \boldsymbol{\mu}_{\rm rm}\|_{\boldsymbol{\Lambda}} \leq C\, H\, |\mathbf{C}\,\underline{\boldsymbol{\varepsilon}}(\boldsymbol{w})|_{1,\mathscr{P}}\,. \tag{5.2}$$

Proof. Given $K \in \mathscr{P}$ and $E \subset \partial K \in \partial\mathscr{P}$, define $\boldsymbol{\chi}_E^K := \mathbf{C}\,\underline{\boldsymbol{\varepsilon}}(\boldsymbol{w})\boldsymbol{n}_E^K \in \boldsymbol{H}^{m+1}(K)$ where $\boldsymbol{n}_E^K|_K$ must be understood as the trivial extension of the unit normal $\boldsymbol{n}_E^K$ to the constant vector function in K. Observe that $\boldsymbol{\chi}_E^K|_E = \boldsymbol{\mu} \in \boldsymbol{L}^2(E)$ owing to the regularity of the function $\boldsymbol{w}$. Let $\Pi_F^\ell : \boldsymbol{L}^2(F) \to \mathbb{P}_\ell(F)^d$ be the $\boldsymbol{L}^2(F)$-orthogonal projection on $F \subset E$, i.e.,

$$\int_F \Pi_F^\ell \boldsymbol{\mu} \cdot \boldsymbol{v}_\ell \,\mathrm{d}\mathbf{s} := \int_F \boldsymbol{\mu} \cdot \boldsymbol{v}_\ell \,\mathrm{d}\mathbf{s} \quad \text{for all } \boldsymbol{v}_\ell \in \mathbb{P}_\ell(F)^d \text{ and } \boldsymbol{\mu} \in \boldsymbol{L}^2(F)\,, \tag{5.3}$$

and set $\boldsymbol{\mu}_\ell \in \boldsymbol{\Lambda}_H$ such that $\boldsymbol{\mu}_\ell|_F := \Pi_F^\ell \boldsymbol{\mu}$ for all $F \subset \partial K$ and all $K \in \mathscr{P}$. Then, for each partition Ξ_H^K, and using (2.8) and [15, Lemma 3], there exists a positive constant C, depending only on d and ℓ, such that, for all $\boldsymbol{v} \in \mathbf{V}(K)$, it holds

$$\langle \boldsymbol{\mu} - \boldsymbol{\mu}_\ell, \boldsymbol{v}\rangle_{\partial K} = \sum_{F\subset\partial K} \int_F (\boldsymbol{\mu} - \Pi_F^\ell \boldsymbol{\mu}) \cdot \boldsymbol{v} \,\mathrm{d}\mathbf{s} \leq \sum_{F\subset\partial K} C\, \sigma_{\tau_F}\, h_{\tau_F}^{m+1}\, |\boldsymbol{v}|_{1,\tau_F}\, |\boldsymbol{\chi}_E^K|_{m+1,\tau_F}\,,$$

where $\tau_F \in \Xi_H^K$ is the element having the face F, and σ_{τ_F} is the shape regularity of τ_F. Then,

$$\begin{aligned}
\langle \boldsymbol{\mu} - \boldsymbol{\mu}_\ell, \boldsymbol{v}\rangle_{\partial K} &\leq C\, \sigma_\Xi^{m+2}\, H^{m+1} \sum_{F\subset\partial K} |\boldsymbol{v}|_{1,\tau_F}\, |\boldsymbol{\chi}_E^K|_{m+1,\tau_F} && \text{(Assumption (A))}\\
&\leq C\, \sigma_\Xi^{m+2}\, H^{m+1} \left(\sum_{F\subset\partial K} |\boldsymbol{\chi}_E^K|_{m+1,\tau_F}^2\right)^{1/2} \left(\sum_{F\subset\partial K} |\boldsymbol{v}|_{1,\tau_F}^2\right)^{1/2}\\
&\leq C\, \sigma_\Xi^{m+2}\, (d+1)\, H^{m+1}\, |\mathbf{C}\,\underline{\boldsymbol{\varepsilon}}(\boldsymbol{w})|_{m+1,K}\, |\boldsymbol{v}|_{1,K}\,.
\end{aligned}$$

Summing up over $\mathscr{P}$, it holds

$$\langle \boldsymbol{\mu} - \boldsymbol{\mu}_\ell, \boldsymbol{v} \rangle_{\partial \mathscr{P}} \leq C\, H^{m+1}\, |\mathbf{C}\, \underline{\boldsymbol{\varepsilon}}(\boldsymbol{w})|_{m+1,\mathscr{P}}\, |\boldsymbol{v}|_{1,\mathscr{P}} \quad \text{for all } \boldsymbol{v} \in \mathbf{V}\,, \tag{5.4}$$

where C depends only on d, ℓ and σ_Ξ. Using the definition of $\Pi^\ell_F(\cdot)$, for every $F \in \mathcal{E}_H$, we get

$$b(\boldsymbol{\mu} - \boldsymbol{\mu}_\ell, \boldsymbol{v}^{\mathrm{rm}}) = 0 \quad \text{for all } \boldsymbol{v}^{\mathrm{rm}} \in \mathbf{V}_{\mathrm{rm}}\,. \tag{5.5}$$

As such, from (2.9) and (2.13) and Korn's inequality (2.10), we obtain

$$\begin{aligned}
\|\boldsymbol{\mu} - \boldsymbol{\mu}_\ell\|_{\boldsymbol{\Lambda}} &= \sup_{\boldsymbol{v} \in \mathbf{V}} \frac{b(\boldsymbol{\mu} - \boldsymbol{\mu}_\ell,\, \boldsymbol{v})}{\|\boldsymbol{v}\|_{\mathbf{V}}} && \text{(using (2.9))} \\
&\leqslant \sup_{\boldsymbol{v} \in \mathbf{V}} \frac{b(\boldsymbol{\mu} - \boldsymbol{\mu}_\ell,\, \boldsymbol{v} - \Pi_{RM}\boldsymbol{v})}{\|\boldsymbol{v} - \Pi_{RM}\boldsymbol{v}\|_{\mathbf{V}}} && \text{(using (2.13) and (5.5))} \\
&= \sup_{\tilde{\boldsymbol{v}} \in \widetilde{\mathbf{V}}} \frac{b(\boldsymbol{\mu} - \boldsymbol{\mu}_\ell,\, \tilde{\boldsymbol{v}})}{\|\tilde{\boldsymbol{v}}\|_{\mathbf{V}}} \\
&\leqslant \sup_{\tilde{\boldsymbol{v}} \in \widetilde{\mathbf{V}}} \frac{b(\boldsymbol{\mu} - \boldsymbol{\mu}_\ell,\, \tilde{\boldsymbol{v}})}{\|\underline{\boldsymbol{\varepsilon}}(\tilde{\boldsymbol{v}})\|_{0,\mathscr{P}}} && \text{(using (2.4))} \\
&\leqslant C_{korn} \sup_{\tilde{\boldsymbol{v}} \in \widetilde{\mathbf{V}}} \frac{b(\boldsymbol{\mu} - \boldsymbol{\mu}_\ell,\, \tilde{\boldsymbol{v}})}{|\tilde{\boldsymbol{v}}|_{1,\mathscr{P}}} && \text{(Equation (2.10))} \\
&\leqslant C \sup_{\boldsymbol{v} \in \mathbf{V}} \frac{b(\boldsymbol{\mu} - \boldsymbol{\mu}_\ell,\, \boldsymbol{v})}{|\boldsymbol{v}|_{1,\mathscr{P}}}\,,
\end{aligned}$$

which together with (5.4) proves (5.1). To prove (5.2), select $\boldsymbol{\mu}_{\mathrm{rm}} \in \boldsymbol{\Lambda}_{\mathrm{rm}}$ from

$$\int_E (\boldsymbol{\mu} - \boldsymbol{\mu}_{\mathrm{rm}}) \cdot \boldsymbol{v}^{\mathrm{rm}}\, \mathrm{d}\mathbf{s} = 0 \quad \text{for all } \boldsymbol{v}^{\mathrm{rm}} \in \mathbf{V}_{\mathrm{rm}}(K)\,, \tag{5.6}$$

for all $K \in \mathscr{P}$. Observe that (5.6) is well-posed from the definition of $\boldsymbol{\Lambda}_{\mathrm{rm}}(K)$ in (4.16). Particularly, it holds from (5.6) that

$$\langle \boldsymbol{\mu} - \boldsymbol{\mu}_{\mathrm{rm}}, \boldsymbol{v} \rangle_{\partial K} \leq C\, H\, |\mathbf{C}\, \underline{\boldsymbol{\varepsilon}}(\boldsymbol{w})|_{1,K}\, |\boldsymbol{v}|_{1,K} \quad \text{for all } \boldsymbol{v} \in \mathbf{V}(K)\,,$$

since $\mathbb{P}_0(K)^d \subset \mathbf{V}_{\mathrm{rm}}(K)$, where the positive constant C, depending only on d using again [15, Lemma 3]. The result (5.2) arises summing up over $\mathscr{P}$, using (5.6), and following closely the prove for the case $\ell \geq 1$. ∎

Next, we present the approximation properties of the local operators T_h and $\hat{T}_h$. Let $\kappa := \frac{c_{max}}{c_{min}}$ be the measure of the contrast between elastic properties in Ω. The following lemma brings second-level error estimates under local regularity assumptions.

Lemma 5.2. *Suppose Assumption (A) holds and that* $T\boldsymbol{\mu}+\hat{T}\boldsymbol{q}\in\boldsymbol{H}^{s+1}(\mathscr{P})$, *with* $\boldsymbol{\mu}\in\boldsymbol{\Lambda}$, $\boldsymbol{q}\in\boldsymbol{L}^2(\Omega)$ *and* $1\leqslant s\leqslant k$. *Then, there exist a positive constant* c, *depending only on* k, d *and* $\sigma_{\mathcal{T}}$ *given in* (3.1), *such that*

$$
\begin{aligned}
&\|(T-T_h)\boldsymbol{\mu}+(\hat{T}-\hat{T}_h)\boldsymbol{q}\|_{\mathbf{V}}\leqslant c\,\kappa\,\sqrt{1+(C_{korn}\,C_P)^2}\,h^s\,|T\boldsymbol{\mu}+\hat{T}\boldsymbol{q}|_{s+1,\mathscr{P}}\\
&\|(T-T_h)\boldsymbol{\mu}+(\hat{T}-\hat{T}_h)\boldsymbol{q}\|_{0,\Omega}\leqslant c\,\kappa\,C_{korn,K}\,C_{P,K}\,\mathcal{H}\,h^s\,|T\boldsymbol{\mu}+\hat{T}\boldsymbol{q}|_{s+1,\mathscr{P}}\,,\\
&|(T-T_h)\boldsymbol{\mu}+(\hat{T}-\hat{T}_h)\boldsymbol{q}|_{1,\mathscr{P}}\leqslant c\,\kappa\,C_{korn,K}\,h^s\,|T\boldsymbol{\mu}+\hat{T}\boldsymbol{q}|_{s+1,\mathscr{P}}\,.
\end{aligned}
\tag{5.7}
$$

Proof. Let $\boldsymbol{\mu}\in\boldsymbol{\Lambda}$ and $\boldsymbol{q}\in\boldsymbol{L}^2(\Omega)$. Subtracting (2.22) and (3.6) from (2.23) and (3.7), respectively, the following orthogonalities hold

$$
\int_K \mathbf{C}\,\underline{\boldsymbol{\varepsilon}}((T_h-T)\boldsymbol{\mu}):\underline{\boldsymbol{\varepsilon}}(\boldsymbol{v}_h)\,\mathrm{d}\boldsymbol{x}=\int_K \mathbf{C}\,\underline{\boldsymbol{\varepsilon}}((\hat{T}_h-\hat{T})\boldsymbol{q}):\underline{\boldsymbol{\varepsilon}}(\boldsymbol{v}_h)\,\mathrm{d}\boldsymbol{x}=0
\tag{5.8}
$$

for all $\boldsymbol{v}_h\in\mathbf{V}_h(K)$ and all $K\in\mathscr{P}$, where we used $\underline{\boldsymbol{\varepsilon}}(\boldsymbol{v}^{\mathrm{rm}})=0$ for all $\boldsymbol{v}^{\mathrm{rm}}\in\mathbf{V}_{\mathrm{rm}}$. Using standard arguments, the following estimate holds

$$
\|\underline{\boldsymbol{\varepsilon}}((T-T_h)\boldsymbol{\mu}+(\hat{T}-\hat{T}_h)\boldsymbol{q})\|_{0,K}\leqslant\kappa\inf_{\boldsymbol{v}_h\in\mathbf{V}_h}\|\underline{\boldsymbol{\varepsilon}}((T\boldsymbol{\mu}+\hat{T}\boldsymbol{q})-\boldsymbol{v}_h)\|_{0,K}\,.
\tag{5.9}
$$

Moreover, using (2.16), and (5.9) and (2.10), we get

$$
\begin{aligned}
&\|(T-T_h)\boldsymbol{\mu}+(\hat{T}-\hat{T}_h)\boldsymbol{q}\|_{0,K}\leqslant C_{korn,K}\,C_{P,K}\,\mathcal{H}\,\kappa\,|T\boldsymbol{\mu}+\hat{T}\boldsymbol{q}-\boldsymbol{v}_h|_{1,K}\,,\\
&|(T-T_h)\boldsymbol{\mu}+(\hat{T}-\hat{T}_h)\boldsymbol{q}|_{1,K}\leqslant C_{korn,K}\,\kappa\,|T\boldsymbol{\mu}+\hat{T}\boldsymbol{q}-\boldsymbol{v}_h|_{1,K}\,.
\end{aligned}
\tag{5.10}
$$

and then, from the definition of the $\mathbf{V}$-norm (5.9) and (5.10), we get

$$
|(T-T_h)\boldsymbol{\mu}+(\hat{T}-\hat{T}_h)\boldsymbol{q}\|_{\mathbf{V}}\leqslant\sqrt{1+(C_{korn}\,C_P)^2}\,\kappa\,|T\boldsymbol{\mu}+\hat{T}\boldsymbol{q}-\boldsymbol{v}_h|_{1,\mathscr{P}}\,,
\tag{5.11}
$$

Next, we chose $\boldsymbol{v}_h|_K$ in (5.9)-(5.11) such that there exists a positive constant c, depending on k, d and $\sigma_{\mathcal{T}}$, that satisfies (see for instance the Scott-Zhang interpolation [41, Theorem 4.1])

$$
|T\boldsymbol{\mu}+\hat{T}\boldsymbol{q}-\boldsymbol{v}_h|_{1,K}\leq c\,h^s\,|T\boldsymbol{\mu}+\hat{T}\boldsymbol{q}|_{s+1,K}\,,
$$

and the results follow. ∎

Assuming additional local regularity for the solution of (2.23), we can improve the error estimates for $\|\boldsymbol{u}_{Hh}-\boldsymbol{u}\|_{0,\Omega}$ using the classical Aubin-Nitsche duality argument applied at the local level. Notably, we shall consider the following additional assumption:

Assumption (B): Let $\hat{T}\boldsymbol{q}\in\boldsymbol{H}^2(\mathscr{P})$ where $\boldsymbol{q}\in\boldsymbol{L}^2(\Omega)$.

There exists a positive constant C_B, independent of $\mathcal{H}$, such that

$$|\hat{T}\boldsymbol{q}|_{2,K} \leqslant C_B \, \|\boldsymbol{q}\|_{0,K} \quad \text{for all } K \in \mathscr{P}\,. \tag{5.12}$$

The literature dedicated to the regularity of elasticity problems is extensive. For instance, [14, Lemmas 2.2 and 2.3] presents classical results for the isotropic elasticity case and, for anisotropic cases, we found some results in [33]. See also [3] for an extensive study on the subject, which includes some sufficient conditions to reach the regularity given in Assumption (B).

Lemma 5.3. *Suppose that Assumption (A) and (B) hold, and let* $\boldsymbol{\mu} \in \boldsymbol{\Lambda}$ *and* $\boldsymbol{q} \in \boldsymbol{L}^2(\Omega)$. *Then, there exists a positive constant* c, *depending only on* k, d, $\sigma_{\mathcal{T}}$, C_B *and* c_{max}, *such that*

$$\|(T - T_h)\boldsymbol{\mu} + (\hat{T} - \hat{T}_h)\boldsymbol{q}\|_{0,\Omega} \leqslant c\, h \, \|\underline{\boldsymbol{\varepsilon}}((T - T_h)\boldsymbol{\mu} + (\hat{T} - \hat{T}_h)\boldsymbol{q})\|_{0,\Omega}\,. \tag{5.13}$$

Proof. The standard Aubin-Nitsche duality argument (see [20, Lemma 2.31]) used in the context of the present work consists of observing that the function $\mathbf{e} = (T - T_h)\boldsymbol{\mu} + (\hat{T} - \hat{T}_h)\boldsymbol{q}$ satisfies

$$\|\mathbf{e}\|_{0,K} = \sup_{\boldsymbol{g} \in \boldsymbol{L}^2(K)} \frac{\int_K \boldsymbol{g} \cdot \mathbf{e} \, \mathrm{d}\boldsymbol{x}}{\|\boldsymbol{g}\|_{0,K}} = \sup_{\boldsymbol{g} \in \boldsymbol{L}^2(K)} \frac{\int_K \mathbf{C}\underline{\boldsymbol{\varepsilon}}(\mathbf{e}) : \underline{\boldsymbol{\varepsilon}}(\hat{T}(\boldsymbol{g})) \, \mathrm{d}\boldsymbol{x}}{\|\boldsymbol{g}\|_{0,K}}\,, \tag{5.14}$$

for a given $\boldsymbol{\mu} \in \boldsymbol{\Lambda}$ and $\boldsymbol{q} \in \boldsymbol{L}^2(K)$, where we used (2.23). Then, using (5.8) and the properties of the elasticity tensor, we obtain

$$\begin{aligned}
\int_K \mathbf{C}\underline{\boldsymbol{\varepsilon}}(\mathbf{e}) : \underline{\boldsymbol{\varepsilon}}(\hat{T}(\boldsymbol{g})) \, \mathrm{d}\boldsymbol{x} &= \inf_{\boldsymbol{v}_h \in \mathbf{V}_h} \int_K \mathbf{C}\underline{\boldsymbol{\varepsilon}}(\mathbf{e}) : \underline{\boldsymbol{\varepsilon}}(\hat{T}(\boldsymbol{g}) - \boldsymbol{v}_h) \, \mathrm{d}\boldsymbol{x} \\
&\leqslant c_{max} \, \|\underline{\boldsymbol{\varepsilon}}(\mathbf{e})\|_{0,K} \inf_{\boldsymbol{v}_h \in \mathbf{V}_h} \|\underline{\boldsymbol{\varepsilon}}(\hat{T}(\boldsymbol{g}) - \boldsymbol{v}_h)\|_{0,K} \\
&\leqslant c_{max} \, \|\underline{\boldsymbol{\varepsilon}}(\mathbf{e})\|_{0,K} \inf_{\boldsymbol{v}_h \in \mathbf{V}_h} |\hat{T}(\boldsymbol{g}) - \boldsymbol{v}_h|_{1,K}
\end{aligned}$$

Choosing $\boldsymbol{v}_h \in \mathbf{V}_h$ with approximation properties as in the proof of Lemma 5.2, we arrive at

$$\begin{aligned}
\int_K \mathbf{C}\underline{\boldsymbol{\varepsilon}}(\mathbf{e}) : \underline{\boldsymbol{\varepsilon}}(\hat{T}(\boldsymbol{g})) \, \mathrm{d}\boldsymbol{x} &\leqslant c\, c_{max} \, h \, \|\underline{\boldsymbol{\varepsilon}}(\mathbf{e})\|_{0,K} |\hat{T}(\boldsymbol{g})|_{2,K} \\
&\leqslant c\, C_B \, c_{max} \, h \, \|\underline{\boldsymbol{\varepsilon}}(\mathbf{e})\|_{0,K} \|\boldsymbol{g}\|_{0,K}\,,
\end{aligned}$$

where we used (5.12). The result follows by replacing the last expression in (5.14), and summing up over $K \in \mathscr{P}$. ∎

Now, we address error estimates for the MHM method on different norms.

5.2. Estimates for $\|\boldsymbol{u}-\boldsymbol{u}_{Hh}\|_{\mathbf{V}}$ and $\|\boldsymbol{\lambda}-\boldsymbol{\lambda}_H\|_{\boldsymbol{\Lambda}}$. In this section, we propose error estimates for displacement and traction in the $\mathbf{V}$- and $\boldsymbol{\Lambda}$-norms, respectively. They highlight the impact of first- and second-level discretizations on convergence under local regularity assumptions. Moreover, we prove that, under Assumption (A), the first-level characteristic size $\mathcal{H}$ does not impact convergence rates. So far, we have explicitly tracked constants dependency in terms of on the mesh regularities $\sigma_{\mathcal{T}}$, σ_{Ξ}, on the dimension d, and on the constants C_{korn}, C_P and C_{tr}, which ultimately depends on the shape of the elements in $\mathscr{P}$ (see Remark 2.1). Hereafter, we decided to hide such dependencies in general positive constants C for the sake of clarity in the statements of next theorems, which will continue to be independent of any mesh parameter (but may depend on the polynomial degrees k and ℓ).

Theorem 5.4. *Suppose Assumption (A) holds, and that $k \geq \ell+d$ and $\ell \geqslant 1$. Let $\boldsymbol{u}$ be the solution of (2.3), and let $1 \leqslant s \leqslant k$ and $1 \leqslant m \leqslant \min\{s, \ell+1\}$ be such that $\boldsymbol{u} \in \boldsymbol{H}^{s+1}(\mathscr{P})$ and $\mathbf{C}\,\underline{\boldsymbol{\varepsilon}}(\boldsymbol{u}) \in H^m(\mathscr{P})^{d\times d}$. Then, there exists C such that*

$$\|\boldsymbol{u}^{\mathrm{rm}}-\boldsymbol{u}_h^{\mathrm{rm}}\|_{\mathbf{V}} + \|\boldsymbol{\lambda}-\boldsymbol{\lambda}_H\|_{\boldsymbol{\Lambda}} \leq C\,\left(h^s\,|\boldsymbol{u}|_{s+1,\mathscr{P}} + H^m\,|\mathbf{C}\,\underline{\boldsymbol{\varepsilon}}(\boldsymbol{u})|_{m,\mathscr{P}}\right)\,. \tag{5.15}$$

In addition, the approximation $\boldsymbol{u}_{Hh} := \boldsymbol{u}_h^{\mathrm{rm}} + T_h\boldsymbol{\lambda}_H + \hat{T}_h\boldsymbol{f}$ satisfies

$$\|\boldsymbol{u}-\boldsymbol{u}_{Hh}\|_{\mathbf{V}} \leq C\,\left(h^s\,|\boldsymbol{u}|_{s+1,\mathscr{P}} + H^m\,|\mathbf{C}\,\underline{\boldsymbol{\varepsilon}}(\boldsymbol{u})|_{m,\mathscr{P}}\right)\,. \tag{5.16}$$

Proof. Let $\boldsymbol{\Lambda}_H^\star \subset \boldsymbol{\Lambda}_H$ be defined by

$$\boldsymbol{\Lambda}_H^\star := \left\{\boldsymbol{\mu}_H \in \boldsymbol{\Lambda}_H \,:\, \int_{\partial K} \boldsymbol{\mu}_H\cdot\boldsymbol{v}^{\mathrm{rm}}\,\mathrm{d}\mathbf{s} = -\int_K \boldsymbol{f}\cdot\boldsymbol{v}^{\mathrm{rm}}\,d\boldsymbol{x} \quad \text{for all } K\in\mathscr{P}\right\}\,.$$

Observe that $\boldsymbol{\lambda}_\ell := \Pi_F^\ell\boldsymbol{\lambda} \in \boldsymbol{\Lambda}_H^\star$ since

$$\int_{\partial K}\boldsymbol{\lambda}_\ell\cdot\boldsymbol{v}^{\mathrm{rm}}\,\mathrm{d}\mathbf{s} = \sum_{F\subset\partial K}\int_F \Pi_F^\ell\boldsymbol{\lambda}\cdot\boldsymbol{v}^{\mathrm{rm}}\,\mathrm{d}\mathbf{s} = \sum_{F\subset\partial K}\int_F \boldsymbol{\lambda}\cdot\boldsymbol{v}^{\mathrm{rm}}\,\mathrm{d}\mathbf{s} = -\int_K \boldsymbol{f}\cdot\boldsymbol{v}^{\mathrm{rm}}\,d\boldsymbol{x}\,,$$

for all $\boldsymbol{v}^{\mathrm{rm}} \in \mathbf{V}_{\mathrm{rm}}$, where we used $\boldsymbol{\lambda}\in\boldsymbol{L}^2(\partial K)$ and the second equation in (2.28). Now let $\boldsymbol{\mu}_H \in \boldsymbol{\Lambda}_H^*$ be arbitrary, and remark that $\boldsymbol{\lambda}_H-\boldsymbol{\mu}_H \in \mathscr{N}_H$ given in (4.19). As such, we get from (4.17) that there exists a positive constant C such that

$$\begin{aligned}
\alpha_0\,\|\boldsymbol{\lambda}_H-\boldsymbol{\mu}_H\|_{\boldsymbol{\Lambda}}^2 &\leq a_h(\boldsymbol{\lambda}_H-\boldsymbol{\mu}_H,\boldsymbol{\lambda}_H-\boldsymbol{\mu}_H)\\
&= a_h(\boldsymbol{\lambda}_H,\boldsymbol{\lambda}_H-\boldsymbol{\mu}_H) - a(\boldsymbol{\lambda},\boldsymbol{\lambda}_H-\boldsymbol{\mu}_H) + a_h(\boldsymbol{\lambda}-\boldsymbol{\mu}_H,\boldsymbol{\lambda}_H-\boldsymbol{\mu}_H)\\
&\quad + a(\boldsymbol{\lambda},\boldsymbol{\lambda}_H-\boldsymbol{\mu}_H) - a_h(\boldsymbol{\lambda},\boldsymbol{\lambda}_H-\boldsymbol{\mu}_H)\\
&= -\langle\boldsymbol{\lambda}_H-\boldsymbol{\mu}_H,\hat{T}_h\boldsymbol{f}\rangle_{\partial\mathscr{P}} + \langle\boldsymbol{\lambda}_H-\boldsymbol{\mu}_H,\hat{T}\boldsymbol{f}\rangle_{\partial\mathscr{P}}\\
&\quad + \langle\boldsymbol{\lambda}_H-\boldsymbol{\mu}_H,T_h(\boldsymbol{\lambda}-\boldsymbol{\mu}_H)\rangle_{\partial\mathscr{P}}\\
&\quad + \langle\boldsymbol{\lambda}_H-\boldsymbol{\mu}_H,T\boldsymbol{\lambda}\rangle_{\partial\mathscr{P}} - \langle\boldsymbol{\lambda}_H-\boldsymbol{\mu}_H,T_h\boldsymbol{\lambda}\rangle_{\partial\mathscr{P}}\\
&\leq \left(C\,\|\boldsymbol{\lambda}-\boldsymbol{\mu}_H\|_{\boldsymbol{\Lambda}} + \|(T-T_h)\boldsymbol{\lambda} + (\hat{T}-\hat{T}_h)\boldsymbol{f}\|_{\mathbf{V}}\right)\|\boldsymbol{\lambda}_H-\boldsymbol{\mu}_H\|_{\boldsymbol{\Lambda}}\,,
\end{aligned} \tag{5.17}$$

where we used Theorem 4.4, (2.27)–(2.28), (3.4)–(3.5) and the stability of T_h in (3.8).

Next, using the fact that $\boldsymbol{u} = \boldsymbol{u}^{\mathrm{rm}} + T\boldsymbol{\lambda} + \hat{T}\boldsymbol{f} \in H^{m+1}(K)$, for every $K \in \mathscr{P}$, it holds from Lemma 5.2 that

$$\|T\boldsymbol{\lambda} + \hat{T}\boldsymbol{f} - \big(\hat{T}_h\boldsymbol{f} + T_h\boldsymbol{\lambda}\big)\|_{\mathbf{V}} \leq c\, h^s\, |\boldsymbol{u}|_{s+1,\mathscr{P}}\,, \tag{5.18}$$

and, therefore, from (5.17) and (5.18) it follows

$$\|\boldsymbol{\lambda}_H - \boldsymbol{\mu}_H\|_{\boldsymbol{\Lambda}} \leq c\, h^s\, |\boldsymbol{u}|_{s+1,\mathscr{P}} + C\, \|\boldsymbol{\lambda} - \boldsymbol{\mu}_H\|_{\boldsymbol{\Lambda}}\,.$$

By setting $\boldsymbol{\mu}_H = \boldsymbol{\lambda}_\ell \in \boldsymbol{\Lambda}_H^\star$ and, from Lemma 5.1 and (2.8), we get

$$\|\boldsymbol{\lambda}_H - \boldsymbol{\mu}_H\|_{\boldsymbol{\Lambda}} \leq c\, h^s\, |\boldsymbol{u}|_{s+1,\mathscr{P}} + C\, H^m\, |\mathbf{C}\,\underline{\boldsymbol{\varepsilon}}(\boldsymbol{u})|_{m,\mathscr{P}}\,,$$

and, thus, using triangle inequality it results

$$\|\boldsymbol{\lambda} - \boldsymbol{\lambda}_H\|_{\boldsymbol{\Lambda}} \leq c\, h^s\, |\boldsymbol{u}|_{s+1,\mathscr{P}} + C\, H^m\, |\mathbf{C}\,\underline{\boldsymbol{\varepsilon}}(\boldsymbol{u})|_{m,\mathscr{P}}\,. \tag{5.19}$$

From Theorem 4.4, there exists $\boldsymbol{\xi}_H \in \boldsymbol{\Lambda}_H$, with $\|\boldsymbol{\xi}_H\|_{\boldsymbol{\Lambda}} = 1$, such that

$$\begin{aligned}
\beta_0\, \|\boldsymbol{u}_h^{\mathrm{rm}} - \boldsymbol{u}^{\mathrm{rm}}\|_{\mathbf{V}} &\leq b(\boldsymbol{\xi}_H, \boldsymbol{u}_h^{\mathrm{rm}} - \boldsymbol{u}^{\mathrm{rm}}) \\
&= -a_h(\boldsymbol{\lambda}_H, \boldsymbol{\xi}_H) + a(\boldsymbol{\lambda}, \boldsymbol{\xi}_H) + \langle \boldsymbol{\xi}_H, (\hat{T} - \hat{T}_h)\boldsymbol{f}\rangle_{\partial\mathscr{P}} \\
&= -a_h(\boldsymbol{\lambda}_H - \boldsymbol{\lambda}, \boldsymbol{\xi}_H) + a(\boldsymbol{\lambda}, \boldsymbol{\xi}_H) - a_h(\boldsymbol{\lambda}, \boldsymbol{\xi}_H) + \langle \boldsymbol{\xi}_H, (\hat{T} - \hat{T}_h)\boldsymbol{f}\rangle_{\partial\mathscr{P}} \\
&\leq C\, \|\boldsymbol{\lambda}_H - \boldsymbol{\lambda}\|_{\boldsymbol{\Lambda}} + \|(T_h - T)\boldsymbol{\lambda} + (\hat{T}_h - \hat{T})\boldsymbol{f}\|_{\mathbf{V}}\,,
\end{aligned}$$

where we used again (2.27)–(2.28) and (3.4)–(3.5), and the uniform continuity of T_h from (3.8). From (5.19) and (5.18), the following estimate holds

$$\|\boldsymbol{u}_h^{\mathrm{rm}} - \boldsymbol{u}^{\mathrm{rm}}\|_{\mathbf{V}} \leq c\, h^s\, |\boldsymbol{u}|_{s+1,\mathscr{P}} + C\, H^m\, |\mathbf{C}\,\underline{\boldsymbol{\varepsilon}}(\boldsymbol{u})|_{m,\mathscr{P}}\,, \tag{5.20}$$

and summing up (5.19) and (5.20) estimate (5.15) follows. To prove (5.16), we observe that

$$\begin{aligned}
\|\boldsymbol{u} - \boldsymbol{u}_{Hh}\|_{\mathbf{V}} &\leq \|\boldsymbol{u}_h^{\mathrm{rm}} - \boldsymbol{u}^{\mathrm{rm}}\|_{\mathbf{V}} + \|T\boldsymbol{\lambda} + \hat{T}\boldsymbol{f} - (T_h\boldsymbol{\lambda}_H + \hat{T}_h\boldsymbol{f})\|_{\mathbf{V}} \\
&\leq \|\boldsymbol{u}_h^{\mathrm{rm}} - \boldsymbol{u}^{\mathrm{rm}}\|_{\mathbf{V}} + \|T_h(\boldsymbol{\lambda} - \boldsymbol{\lambda}_H)\|_{\mathbf{V}} + \|T\boldsymbol{\lambda} + \hat{T}\boldsymbol{f} - (T_h\boldsymbol{\lambda} + \hat{T}_h\boldsymbol{f})\|_{\mathbf{V}}\,,
\end{aligned}$$

and, applying the uniform continuity of T_h from (3.8), (5.15) and (5.18), estimate (5.16) follows. ∎

Remark 5.5. As for the minimal space $\mathbf{V}_{\rm rm}\times\mathbf{\Lambda}_{\rm rm}$, we can follow the same strategy of proof in Theorem 5.4, and under Assumption (A), $k\geq d+1$, and $\boldsymbol{u}\in\boldsymbol{H}^2(\mathscr{P})$ and $\mathbf{C}\underline{\boldsymbol{\varepsilon}}(\boldsymbol{u})\in H^1(\mathscr{P})^{d\times d}$, to arrive at

$$\begin{aligned}\|\boldsymbol{u}^{\rm rm}-\boldsymbol{u}_h^{\rm rm}\|_{\mathbf{V}}+\|\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\rm rm}\|_{\mathbf{\Lambda}}&\leq C\,\left(h\,|\boldsymbol{u}|_{2,\mathscr{P}}+H\,|\mathbf{C}\,\underline{\boldsymbol{\varepsilon}}(\boldsymbol{u})|_{1,\mathscr{P}}\right)\,,\\ \|\boldsymbol{u}-\boldsymbol{u}_{Hh}\|_{\mathbf{V}}&\leq C\,\left(h\,|\boldsymbol{u}|_{2,\mathscr{P}}+H\,|\mathbf{C}\,\underline{\boldsymbol{\varepsilon}}(\boldsymbol{u})|_{1,\mathscr{P}}\right)\,.\end{aligned}\tag{5.21}$$

□

5.3. Estimates for $\|\boldsymbol{u}-\boldsymbol{u}_{Hh}\|_{1,\mathscr{P}}$ and $\|\boldsymbol{u}-\boldsymbol{u}_{Hh}\|_{0,\Omega}$. In this section, we propose an error estimate in the $\boldsymbol{H}^1(\mathscr{P})$-norm providing an estimation to the anti-symmetric part of the gradient of the error with clean constants. Also, we establish optimal error estimates in the $\boldsymbol{L}^2(\Omega)$-norm under smoothing properties.

As for the convergence in the $\boldsymbol{H}^1(\mathscr{P})$-norm, we first observe that, given $K\in\mathscr{P}$, the number of faces $E\subset\partial K$ does not increase indefinitely as the diameter of K stays fixed, under a reasonable assumption on the trace constant $C_{tr,K}$. Precisely, it holds

$$h_K\leq\frac{C_{tr,K}}{2\,(d+1)}\,h_E\quad\text{for all } K\in\mathscr{P}\text{ and } E\subset\partial K\,,\tag{5.22}$$

since from Remark 2.1, we get

$$h_K=\frac{\min_{\tau\in\mathcal{T}^K}h_\tau}{\bar{\varrho}_K}=\frac{\min_{\tau\in\mathcal{T}^K}\frac{h_\tau}{\rho_\tau}\,\rho_\tau}{\bar{\varrho}_K}\leqslant\frac{\bar{\sigma}_K}{\bar{\varrho}_K}\,\min_{\tau\in\mathcal{T}^K}\rho_\tau\leqslant\frac{\bar{\sigma}_K}{\bar{\varrho}_K}\,\frac{h_E}{2}=\frac{C_{tr,K}}{2\,(d+1)}\,h_E\,.$$

Theorem 5.6. *Assume the hypothesis of Theorem 5.4 and (5.22) hold. Then, there exists C such that*

$$\|\boldsymbol{u}-\boldsymbol{u}_{Hh}\|_{1,\mathscr{P}}\leq C\,\left(h^s\,|\boldsymbol{u}|_{s+1,\mathscr{P}}+H^m\,|\mathbf{C}\,\underline{\boldsymbol{\varepsilon}}(\boldsymbol{u})|_{m,\mathscr{P}}\right)\,.\tag{5.23}$$

Proof. The estimate $\|\boldsymbol{u}-\boldsymbol{u}_{Hh}\|_{0,\mathscr{P}}$ is given in (5.16). We are left with establishing an estimate for $\|\nabla(\boldsymbol{u}-\boldsymbol{u}_{Hh})\|_{0,\mathscr{P}}$. Using (2.18) and (5.16), and the Korn's inequality (2.10), it holds

$$\begin{aligned}\|\nabla(T\boldsymbol{\lambda}-T_h\boldsymbol{\lambda}_H+\hat{T}\boldsymbol{f}-\hat{T}_h\boldsymbol{f})\|_{0,\mathscr{P}}&\leqslant C_{korn}\,\|\underline{\boldsymbol{\varepsilon}}(T\boldsymbol{\lambda}-T_h\boldsymbol{\lambda}_H+\hat{T}\boldsymbol{f}-\hat{T}_h\boldsymbol{f})\|_{0,\mathscr{P}}\\ &\leqslant c\,h^s\,|\boldsymbol{u}|_{s+1,\mathscr{P}}+C\,H^m\,|\mathbf{C}\,\underline{\boldsymbol{\varepsilon}}(\boldsymbol{u})|_{m,\mathscr{P}}\,.\end{aligned}\tag{5.24}$$

It remains to estimate $\boldsymbol{u}^{\mathrm{rm}} - \boldsymbol{u}_h^{\mathrm{rm}}$ in the $\boldsymbol{H}^1(\mathscr{P})$-semi-norm. To this end, we use the Korn's inequality (2.11) on $\boldsymbol{u}^{\mathrm{rm}} - \boldsymbol{u}_h^{\mathrm{rm}} \in \mathbf{V}_{\mathrm{rm}}$, which leads to

$$\|\nabla(\boldsymbol{u}^{\mathrm{rm}} - \boldsymbol{u}_h^{\mathrm{rm}})\|_{0,\mathscr{P}}^2 \leqslant C \left(\|\boldsymbol{u}^{\mathrm{rm}} - \boldsymbol{u}_h^{\mathrm{rm}}\|_{0,\Omega}^2 + \sum_{E \in \mathcal{E}_0} \frac{1}{h_E} \|[\![\boldsymbol{u}^{\mathrm{rm}} - \boldsymbol{u}_h^{\mathrm{rm}}]\!]\|_{0,E}^2 \right). \tag{5.25}$$

Let $K, K' \in \mathscr{P}$ be two elements sharing the face $E \in \mathcal{E}_0$. Set $\boldsymbol{\mu}_H := [\![\boldsymbol{u}^{\mathrm{rm}} - \boldsymbol{u}_h^{\mathrm{rm}}]\!]\,|_E \in \boldsymbol{\Lambda}_{\mathrm{rm}}$ and zero on the other faces. Next, testing (3.5) and (2.27) with $\boldsymbol{\mu}_H$ and subtracting them, we obtain

$$\begin{aligned}
\|[\![\boldsymbol{u}^{\mathrm{rm}} - \boldsymbol{u}_h^{\mathrm{rm}}]\!]\|_{0,E}^2 &= (\boldsymbol{\mu}_H, [\![\boldsymbol{u}^{\mathrm{rm}} - \boldsymbol{u}_h^{\mathrm{rm}}]\!])_E \\
&= |(\boldsymbol{\mu}_H, \boldsymbol{u}^{\mathrm{rm}} - \boldsymbol{u}_h^{\mathrm{rm}})_{E\cap K} - (\boldsymbol{\mu}_H, \boldsymbol{u}^{\mathrm{rm}} - \boldsymbol{u}_h^{\mathrm{rm}})_{E\cap K'}| \\
&= |\langle \boldsymbol{\mu}_H, \boldsymbol{u}^{\mathrm{rm}} - \boldsymbol{u}_h^{\mathrm{rm}} \rangle_{\partial\mathscr{P}}| \\
&= |\langle \boldsymbol{\mu}_H, T\boldsymbol{\lambda} - T_h\boldsymbol{\lambda}_H + \hat{T}\boldsymbol{f} - \hat{T}_h\boldsymbol{f} \rangle_{\partial\mathscr{P}}| \\
&= |(\boldsymbol{\mu}_H, T\boldsymbol{\lambda} - T_h\boldsymbol{\lambda}_H + \hat{T}\boldsymbol{f} - \hat{T}_h\boldsymbol{f})_{E\cap K} \\
&\quad - (\boldsymbol{\mu}_H, T\boldsymbol{\lambda} - T_h\boldsymbol{\lambda}_H + \hat{T}\boldsymbol{f} - \hat{T}_h\boldsymbol{f})_{E\cap K'}| \\
&\leqslant \|[\![\boldsymbol{u}^{\mathrm{rm}} - \boldsymbol{u}_h^{\mathrm{rm}}]\!]\|_{0,E} (\|T\boldsymbol{\lambda} - T_h\boldsymbol{\lambda}_H + \hat{T}\boldsymbol{f} - \hat{T}_h\boldsymbol{f}\|_{0,E\cap K} \\
&\quad + \|T\boldsymbol{\lambda} - T_h\boldsymbol{\lambda}_H + \hat{T}\boldsymbol{f} - \hat{T}_h\boldsymbol{f}\|_{0,E\cap K'}) .
\end{aligned}$$

Next, we apply (2.14) and (2.15) on the right-hand side of the above inequality, and use (5.22) to obtain

$$\begin{aligned}
\|[\![\boldsymbol{u}^{\mathrm{rm}} - \boldsymbol{u}_h^{\mathrm{rm}}]\!]\|_{0,E} &\leqslant (C_{tr}\, C_P\, (C_P\, d + 2))^{\frac{1}{2}}\, (h_K^{\frac{1}{2}} |T\boldsymbol{\lambda} - T_h\boldsymbol{\lambda}_H + \hat{T}\boldsymbol{f} - \hat{T}_h\boldsymbol{f}|_{1,K} \\
&\quad + h_{K'}^{\frac{1}{2}} |T\boldsymbol{\lambda} - T_h\boldsymbol{\lambda}_H + \hat{T}\boldsymbol{f} - \hat{T}_h\boldsymbol{f}|_{1,K'}) \\
&\leqslant C_{tr} \left(\frac{C_P}{d+1} \left(C_P\, \frac{d}{2} + 1 \right) \right)^{\frac{1}{2}} h_E^{\frac{1}{2}} (|T\boldsymbol{\lambda} - T_h\boldsymbol{\lambda}_H + \hat{T}\boldsymbol{f} - \hat{T}_h\boldsymbol{f}|_{1,K} \\
&\quad + |T\boldsymbol{\lambda} - T_h\boldsymbol{\lambda}_H + \hat{T}\boldsymbol{f} - \hat{T}_h\boldsymbol{f}|_{1,K'}) ,
\end{aligned}$$

with constants defined in (2.18). Now, dividing both sides by $h_E^{\frac{1}{2}}$ and taking the sum of the squares, for all $E \in \mathcal{E}_0$, we get

$$\left(\sum_{E \in \mathcal{E}_0} \frac{1}{h_E} \|[\boldsymbol{u}^{\mathrm{rm}} - \boldsymbol{u}_h^{\mathrm{rm}}]\|_{0,E}^2 \right)^{\frac{1}{2}} \leq C\, \|\nabla(T\boldsymbol{\lambda} - T_h\boldsymbol{\lambda}_H + \hat{T}\boldsymbol{f} - \hat{T}_h\boldsymbol{f})\|_{0,\mathscr{P}} . \tag{5.26}$$

Hence, from the triangle inequality, (5.25) and (5.26), we arrive at

$$\begin{aligned}
&\|\boldsymbol{u} - \boldsymbol{u}_{Hh}\|_{1,\mathscr{P}}^2 \\
&\quad \leqslant \|\boldsymbol{u} - \boldsymbol{u}_{Hh}\|_{0,\Omega}^2 + \|\nabla(\boldsymbol{u}^{\mathrm{rm}} - \boldsymbol{u}_h^{\mathrm{rm}})\|_{0,\mathscr{P}}^2 + \|\nabla(T\boldsymbol{\lambda} - T_h\boldsymbol{\lambda}_H + \hat{T}\boldsymbol{f} - \hat{T}_h\boldsymbol{f})\|_{0,\mathscr{P}}^2 \\
&\quad \leqslant \|\boldsymbol{u} - \boldsymbol{u}_{Hh}\|_{0,\Omega}^2 + C \left(\|\boldsymbol{u}^{\mathrm{rm}} - \boldsymbol{u}_h^{\mathrm{rm}}\|_{0,\mathscr{P}}^2 + \|\nabla(T\boldsymbol{\lambda} - T_h\boldsymbol{\lambda}_H + \hat{T}\boldsymbol{f} - \hat{T}_h\boldsymbol{f})\|_{0,\mathscr{P}}^2 \right),
\end{aligned}$$

and the result follows from (5.16) and (5.24). ∎

Next, we assume smoothing properties (see [20, §3.1.3] for instance) hold. Notably, we assume the following:

Assumption (C): Let $\boldsymbol{w}_{\boldsymbol{f}} \in \boldsymbol{H}^1_0(\Omega)$ be the solution of (2.3) with $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega)$ and $\boldsymbol{g} = \boldsymbol{0}$.

Then $\boldsymbol{w}_{\boldsymbol{f}} \in \boldsymbol{H}^2(\mathscr{P})$, $\mathbf{C}\boldsymbol{w}_{\boldsymbol{f}} \in H^1(\mathscr{P})^{d\times d}$, and there exists a positive constant C_A, independent of $\mathcal{H}$, such that

$$|\boldsymbol{w}_{\boldsymbol{f}}|_{2,\mathscr{P}} + |\mathbf{C}\underline{\boldsymbol{\varepsilon}}(\boldsymbol{w}_{\boldsymbol{f}})|_{1,\mathscr{P}} \leqslant C_A \, \|\boldsymbol{f}\|_{0,\Omega} \quad \text{for all } \boldsymbol{f} \in \boldsymbol{L}^2(\Omega)\,.$$

The next result improves the estimate in the $\boldsymbol{L}^2(\Omega)$-norm under Assumptions (B) and (C).

Theorem 5.7. *Assume the hypothesis of Theorem 5.4 and Assumption (B) and (C) hold. Then, there exists C such that*

$$\|\boldsymbol{u} - \boldsymbol{u}_{Hh}\|_{0,\Omega} \leq C\,(h+H)\,\left(h^k\,|\boldsymbol{u}|_{k+1,\mathscr{P}} + H^{\ell+1}\,|\mathbf{C}\,\underline{\boldsymbol{\varepsilon}}(\boldsymbol{u})|_{\ell+1,\mathscr{P}}\right)\,. \tag{5.27}$$

Proof. First, we observe that (3.9) also holds if one replaces the operator T_h by T, i.e., we have

$$\langle \boldsymbol{\mu}, \hat{T}\boldsymbol{f}\rangle_{\partial K} = \int_K \mathbf{C}\,\underline{\boldsymbol{\varepsilon}}(T\boldsymbol{\mu}) : \underline{\boldsymbol{\varepsilon}}(\hat{T}\boldsymbol{f})\,\mathrm{d}\boldsymbol{x} = \int_K T\boldsymbol{\mu}\cdot\boldsymbol{f}\,\mathrm{d}\boldsymbol{x}\,. \tag{5.28}$$

Then, define $\mathbf{e} = \boldsymbol{u} - \boldsymbol{u}_{Hh}$, and let $\boldsymbol{w}$ be the solution of the following elasticity problem

$$\nabla\cdot\,\mathbf{C}\,\underline{\boldsymbol{\varepsilon}}(\boldsymbol{w}) = \mathbf{e} \quad \text{in } \Omega\,, \quad \text{and} \quad \boldsymbol{w} = \boldsymbol{0} \quad \text{on } \partial\Omega\,. \tag{5.29}$$

Using (3.9), we can write $\boldsymbol{w} = \boldsymbol{w}^{\mathrm{rm}} + T\gamma + \hat{T}(-\mathbf{e})$, where $(\gamma, \boldsymbol{w}^{\mathrm{rm}}) \in \boldsymbol{\Lambda}\times\mathbf{V}_{\mathrm{rm}}$ satisfy

$$\begin{aligned} a(\gamma,\boldsymbol{\mu}) + b(\boldsymbol{\mu}, \boldsymbol{w}^{\mathrm{rm}}) &= \int_\Omega T\boldsymbol{\mu}\cdot\mathbf{e}\,\mathrm{d}\boldsymbol{x} \quad \text{for all } \boldsymbol{\mu}\in\boldsymbol{\Lambda}\,,\\ b(\gamma, \boldsymbol{v}^{\mathrm{rm}}) &= \int_\Omega \mathbf{e}\cdot\boldsymbol{v}^{\mathrm{rm}}\,\mathrm{d}\boldsymbol{x} \quad \text{for all } \boldsymbol{v}^{\mathrm{rm}}\in\mathbf{V}_{\mathrm{rm}}\,. \end{aligned} \tag{5.30}$$

Problem (5.30) is well-posed by Theorem 2.3, and from Assumption (C), function $\boldsymbol{w}$ belongs to $\boldsymbol{H}^1_0(\Omega)\cap\boldsymbol{H}^2(\mathscr{P})$ and

$$|\boldsymbol{w}|_{2,\mathscr{P}} + |\mathbf{C}\underline{\boldsymbol{\varepsilon}}(\boldsymbol{w})|_{1,\mathscr{P}} \leq C_A \, \|\mathbf{e}\|_{0,\Omega} \, . \tag{5.31}$$

Let $(\gamma_{\rm rm}, \boldsymbol{w}_h^{\rm rm}) \in \boldsymbol{\Lambda}_{\rm rm} \times \mathbf{V}_{\rm rm}$ be the following discrete solution of (5.30)

$$\begin{aligned} a_h(\gamma_{\rm rm}, \boldsymbol{\mu}_H) + b(\boldsymbol{\mu}_H, \boldsymbol{w}_h^{\rm rm}) &= \int_\Omega T_h \boldsymbol{\mu}_H \cdot \mathbf{e} \, {\rm d}\boldsymbol{x} \quad \text{for all } \boldsymbol{\mu}_H \in \boldsymbol{\Lambda}_{\rm rm} \, , \\ b(\gamma_{\rm rm}, \boldsymbol{v}^{\rm rm}) &= \int_\Omega \mathbf{e} \cdot \boldsymbol{v}^{\rm rm} \, {\rm d}\boldsymbol{x} \quad \text{for all } \boldsymbol{v}^{\rm rm} \in \mathbf{V}_{\rm rm} \, . \end{aligned} \tag{5.32}$$

Problem (5.32) is well-posed from (3.9) and Theorem 4.4, and $\boldsymbol{w}_h := \boldsymbol{w}_h^{\rm rm} + T_h\gamma + \hat{T}_h(-\mathbf{e})$. From (5.21), the following error estimate holds

$$\begin{aligned} \|\gamma - \gamma_{\rm rm}\|_{\boldsymbol{\Lambda}} + \|\boldsymbol{w}^{\rm rm} - \boldsymbol{w}_h^{\rm rm}\|_{\mathbf{V}} &\leq C \, \left(h \, |\boldsymbol{w}|_{2,\mathscr{P}} + H \, |\mathbf{C}\underline{\boldsymbol{\varepsilon}}(\boldsymbol{w})|_{1,\mathscr{P}} \right) \\ &\leq C \, (h + H) \, \|\mathbf{e}\|_{0,\Omega} \, , \end{aligned} \tag{5.33}$$

where we used (5.31). Then, from (5.30), (5.32) and (2.27), we arrive at

$$\begin{aligned} \|\mathbf{e}\|_{0,\Omega}^2 &= (\boldsymbol{u} - \boldsymbol{u}_{Hh}, \mathbf{e})_{\mathscr{P}} \\ &= (\boldsymbol{u}^{\rm rm} - \boldsymbol{u}_h^{\rm rm}, \mathbf{e})_{\mathscr{P}} + (T\boldsymbol{\lambda} - T_h\boldsymbol{\lambda}_H + \hat{T}\boldsymbol{f} - \hat{T}_h\boldsymbol{f}, \mathbf{e})_{\mathscr{P}} \\ &= (\boldsymbol{u}^{\rm rm} - \boldsymbol{u}_h^{\rm rm}, \mathbf{e})_{\mathscr{P}} + (T(\boldsymbol{\lambda} - \boldsymbol{\lambda}_H), \mathbf{e})_{\mathscr{P}} \\ &\quad + (T\boldsymbol{\lambda}_H - T_h\boldsymbol{\lambda}_H + \hat{T}\boldsymbol{f} - \hat{T}_h\boldsymbol{f}, \mathbf{e})_{\mathscr{P}} \\ &= b(\gamma_{\rm rm}, \boldsymbol{u}^{\rm rm} - \boldsymbol{u}_h^{\rm rm}) + a(\gamma, \boldsymbol{\lambda} - \boldsymbol{\lambda}_H) + b(\boldsymbol{\lambda} - \boldsymbol{\lambda}_H, \boldsymbol{w}^{\rm rm}) \\ &\quad + (T\boldsymbol{\lambda}_H - T_h\boldsymbol{\lambda}_H + \hat{T}\boldsymbol{f} - \hat{T}_h\boldsymbol{f}, \mathbf{e})_{\mathscr{P}} \\ &= \underbrace{\langle \gamma_{\rm rm}, \boldsymbol{u}^{\rm rm} - \boldsymbol{u}_h^{\rm rm} \rangle_{\partial\mathscr{P}} + \langle \boldsymbol{\lambda} - \boldsymbol{\lambda}_H, T\gamma \rangle_{\partial\mathscr{P}}}_{(I)} \\ &\quad + \underbrace{\|(T - T_h)\boldsymbol{\lambda}_H + (\hat{T} - \hat{T}_h)\boldsymbol{f}\|_{0,\Omega} \|\mathbf{e}\|_{0,\Omega}}_{(II)} \end{aligned} \tag{5.34}$$

where we used $b(\boldsymbol{\lambda} - \boldsymbol{\lambda}_H, \boldsymbol{w}^{\rm rm}) = 0$. Let us estimate (I). Observe that

$$\begin{aligned} \langle \gamma_{\rm rm}, \boldsymbol{u}^{\rm rm} - \boldsymbol{u}_h^{\rm rm} \rangle_{\partial\mathscr{P}} + \langle \boldsymbol{\lambda} - \boldsymbol{\lambda}_H, T\gamma \rangle_{\partial\mathscr{P}} & \\ = -\langle \gamma_{\rm rm}, T\boldsymbol{\lambda} \rangle_{\partial\mathscr{P}} + \langle \gamma_{\rm rm}, T_h\boldsymbol{\lambda}_H \rangle_{\partial\mathscr{P}} & \\ - \langle \gamma_{\rm rm}, \hat{T}\boldsymbol{f} - \hat{T}_h\boldsymbol{f} \rangle_{\partial\mathscr{P}} + \langle \gamma, T(\boldsymbol{\lambda} - \boldsymbol{\lambda}_H) \rangle_{\partial\mathscr{P}} & \\ = \langle \gamma - \gamma_{\rm rm}, T(\boldsymbol{\lambda} - \boldsymbol{\lambda}_H) \rangle_{\partial\mathscr{P}} + \langle \gamma_{\rm rm}, T_h\boldsymbol{\lambda}_H \rangle_{\partial\mathscr{P}} & \\ - \langle \gamma_{\rm rm}, T\boldsymbol{\lambda}_H \rangle_{\partial\mathscr{P}} - \langle \gamma_{\rm rm}, \hat{T}\boldsymbol{f} - \hat{T}_h\boldsymbol{f} \rangle_{\partial\mathscr{P}} & \\ \leq C \, \underbrace{\|\gamma - \gamma_{\rm rm}\|_{\boldsymbol{\Lambda}} \|\boldsymbol{\lambda} - \boldsymbol{\lambda}_H\|_{\boldsymbol{\Lambda}}}_{(III)} & \\ + \underbrace{|\langle \gamma_{\rm rm}, (T - T_h)\boldsymbol{\lambda}_H + (\hat{T} - \hat{T}_h)\boldsymbol{f} \rangle_{\partial\mathscr{P}}|}_{(IV)} \, . & \end{aligned}$$

The term (III) above is bounded as follows

$$\begin{aligned}\|\gamma-\gamma_{\rm rm}\|_{\boldsymbol{\Lambda}}\|\boldsymbol{\lambda}-\boldsymbol{\lambda}_H\|_{\boldsymbol{\Lambda}} &\leq C\,(h+H)\,\|\mathbf{e}\|_{0,\Omega}\|\boldsymbol{\lambda}-\boldsymbol{\lambda}_H\|_{\boldsymbol{\Lambda}}\\ &\leq C\,(h+H)\,\left(h^k\,|\boldsymbol{u}|_{k+1,\mathscr{P}}+H^{\ell+1}\,|\mathbf{C}\,\underline{\boldsymbol{\varepsilon}}(\boldsymbol{u})|_{\ell+1,\mathscr{P}}\right)\,\|\mathbf{e}\|_{0,\Omega}\,,\end{aligned}\tag{5.35}$$

where we used (5.15) and (5.33). For term (IV), it holds

$$\begin{aligned}&|\langle\gamma_{\rm rm},(T-T_h)\boldsymbol{\lambda}_H+(\hat{T}-\hat{T}_h)\boldsymbol{f}\rangle_{\partial\mathscr{P}}|\\ &\quad=|\sum_{K\in\mathscr{P}}\int_K\mathbf{C}\underline{\boldsymbol{\varepsilon}}(T\gamma_{\rm rm}):\underline{\boldsymbol{\varepsilon}}((T-T_h)\boldsymbol{\lambda}_H+(\hat{T}-\hat{T}_h)\boldsymbol{f})\,\mathrm{d}\boldsymbol{x}| && \text{(def. of } T)\\ &\quad\leqslant|\sum_{K\in\mathscr{P}}\int_K\mathbf{C}\underline{\boldsymbol{\varepsilon}}(T_h\gamma_{\rm rm}):\underline{\boldsymbol{\varepsilon}}((T-T_h)\boldsymbol{\lambda}_H+(\hat{T}-\hat{T}_h)\boldsymbol{f})\,\mathrm{d}\boldsymbol{x}|\\ &\qquad+|\sum_{K\in\mathscr{P}}\int_K\mathbf{C}\underline{\boldsymbol{\varepsilon}}((T-T_h)\gamma_{\rm rm}):\underline{\boldsymbol{\varepsilon}}((T-T_h)\boldsymbol{\lambda}_H+(\hat{T}-\hat{T}_h)\boldsymbol{f})\,\mathrm{d}\boldsymbol{x}| && (\pm T_h\gamma_{\rm rm})\\ &\quad=|\sum_{K\in\mathscr{P}}\int_K\mathbf{C}\underline{\boldsymbol{\varepsilon}}((T-T_h)\gamma_{\rm rm}):\underline{\boldsymbol{\varepsilon}}((T-T_h)\boldsymbol{\lambda}_H+(\hat{T}-\hat{T}_h)\boldsymbol{f})\,\mathrm{d}\boldsymbol{x}| && ((5.8))\\ &\quad\leqslant c_{max}\,\|\underline{\boldsymbol{\varepsilon}}((T-T_h)\boldsymbol{\lambda}_H+(\hat{T}-\hat{T}_h)\boldsymbol{f})\|_{0,\mathscr{P}}\|\underline{\boldsymbol{\varepsilon}}((T-T_h)\gamma_{\rm rm}))\|_{0,\mathscr{P}} && \text{(using (2.2))}\end{aligned}\tag{5.36}$$

$$\begin{aligned}&\quad\leqslant c_{max}\left(\|\underline{\boldsymbol{\varepsilon}}(\boldsymbol{u}-\boldsymbol{u}_{Hh})\|_{0,\mathscr{P}}+\|\underline{\boldsymbol{\varepsilon}}(T(\boldsymbol{\lambda}_H-\boldsymbol{\lambda}))\|_{0,\mathscr{P}}\right)\\ &\qquad\qquad\|\underline{\boldsymbol{\varepsilon}}((T-T_h)\gamma_{\rm rm}))\|_{0,\mathscr{P}} && (\pm T\boldsymbol{\lambda})\\ &\quad\leqslant C\left(\|\underline{\boldsymbol{\varepsilon}}(\mathbf{e})\|_{0,\mathscr{P}}+\|\boldsymbol{\lambda}_H-\boldsymbol{\lambda}\|_{\boldsymbol{\Lambda}}\right)\|\underline{\boldsymbol{\varepsilon}}((T-T_h)\gamma_{\rm rm}))\|_{0,\mathscr{P}} && \text{(using (2.25))}\end{aligned}\tag{5.37}$$

$$\leqslant C\left(\|\mathbf{e}\|_{\mathbf{V}}+\|\boldsymbol{\lambda}_H-\boldsymbol{\lambda}\|_{\boldsymbol{\Lambda}}\right)\|\underline{\boldsymbol{\varepsilon}}((T-T_h)\gamma_{\rm rm}))\|_{0,\mathscr{P}}\,.$$

The last term in (5.37) can be bounded as follows

$$\begin{aligned}&\|\underline{\boldsymbol{\varepsilon}}((T-T_h)\gamma_{\rm rm})\|_{0,\mathscr{P}}\\ &\qquad\leqslant\|\underline{\boldsymbol{\varepsilon}}(T(\gamma_{\rm rm}-\gamma)\|_{0,\mathscr{P}}+\|\underline{\boldsymbol{\varepsilon}}(T\gamma-T_h\gamma_{\rm rm})\|_{0,\mathscr{P}} && (\pm T\gamma)\\ &\qquad\leqslant\|\gamma_{\rm rm}-\gamma\|_{\boldsymbol{\Lambda}}+\|\underline{\boldsymbol{\varepsilon}}(T\gamma-T_h\gamma_{\rm rm})\|_{0,\mathscr{P}} && \text{(using (2.25))}\\ &\qquad\leqslant\|\gamma_{\rm rm}-\gamma\|_{\boldsymbol{\Lambda}}+\|\underline{\boldsymbol{\varepsilon}}((\hat{T}-\hat{T}_h)\mathbf{e})\|_{0,\mathscr{P}}\\ &\qquad\quad+\|\underline{\boldsymbol{\varepsilon}}(T\gamma-T_h\gamma_{\rm rm}-(\hat{T}-\hat{T}_h)\mathbf{e})\|_{0,\mathscr{P}} && (\pm(\hat{T}-\hat{T}_h)\mathbf{e})\\ &\qquad\leqslant\|\gamma_{\rm rm}-\gamma\|_{\boldsymbol{\Lambda}}+\|\underline{\boldsymbol{\varepsilon}}(\boldsymbol{w}-\boldsymbol{w}_h)\|_{0,\mathscr{P}}\\ &\qquad\quad+\|\underline{\boldsymbol{\varepsilon}}((\hat{T}-\hat{T}_h)\mathbf{e})\|_{0,\mathscr{P}}\\ &\qquad\leqslant(\|\gamma_{\rm rm}-\gamma\|_{\boldsymbol{\Lambda}}+\|\boldsymbol{w}-\boldsymbol{w}_h\|_{\mathbf{V}})+\|\underline{\boldsymbol{\varepsilon}}((\hat{T}-\hat{T}_h)\mathbf{e})\|_{0,\mathscr{P}}\,.\end{aligned}\tag{5.38}$$

From (5.37) and (5.38), and using Theorem 5.4 and (5.21) and (5.31), it holds

$$
\begin{aligned}
(5.39)\qquad & |\langle \gamma_{\mathrm{rm}}, (T-T_h)\boldsymbol{\lambda}_H + (\hat{T}-\hat{T}_h)\boldsymbol{f}\rangle_{\partial\mathscr{P}}| \\
& \leqslant C\,\left(h^k\,|\boldsymbol{u}|_{k+1,\mathscr{P}} + H^{\ell+1}\,|\mathbf{C}\,\underline{\boldsymbol{\varepsilon}}(\boldsymbol{u})|_{\ell+1,\mathscr{P}}\right) \\
& \quad \left(\left(h\,|\boldsymbol{w}|_{2,\mathscr{P}} + H\,|\mathbf{C}\,\underline{\boldsymbol{\varepsilon}}(\boldsymbol{w})|_{1,\mathscr{P}}\right) + \|\underline{\boldsymbol{\varepsilon}}((\hat{T}-\hat{T}_h)\mathbf{e})\|_{0,\mathscr{P}}\right) \\
& \leqslant C\,\left(h^k\,|\boldsymbol{u}|_{k+1,\mathscr{P}} + H^{\ell+1}\,|\mathbf{C}\,\underline{\boldsymbol{\varepsilon}}(\boldsymbol{u})|_{\ell+1,\mathscr{P}}\right) \\
(5.40)\qquad & \quad \left((h+H)\,\|\mathbf{e}\|_{0,\Omega} + \|\underline{\boldsymbol{\varepsilon}}((\hat{T}-\hat{T}_h)\mathbf{e})\|_{0,\mathscr{P}}\right)\,.
\end{aligned}
$$

Using Lemma 5.2 and Assumption (B), one estimates the last part of (5.40) as follows

$$
(5.41)\qquad \|\underline{\boldsymbol{\varepsilon}}((\hat{T}-\hat{T}_h)\mathbf{e})\|_{0,\mathscr{P}} \leqslant C\,h\,\|\hat{T}\mathbf{e}\|_{2,\mathscr{P}} \leqslant C\,C_B\,h\,\|\mathbf{e}\|_{0,\Omega}\,,
$$

and the bound for the term (IV) follows combining (5.40) and (5.41). Item (II) in (5.34) is bounded using (5.13), (5.7), the stability of the mappings T_h and T in (3.8) and (2.25), and Theorem 5.4 as follows

$$
\begin{aligned}
(5.42)\qquad & \|(T-T_h)\boldsymbol{\lambda}_H + (\hat{T}-\hat{T}_h)\boldsymbol{f}\|_{0,\Omega} \\
& \leqslant \|(T-T_h)(\boldsymbol{\lambda}_H-\boldsymbol{\lambda})\|_{0,\Omega} + \|(T-T_h)\boldsymbol{\lambda} + (\hat{T}-\hat{T}_h)\boldsymbol{f}\|_{0,\Omega} \\
& \leqslant C\,h\,\left(\|(T-T_h)(\boldsymbol{\lambda}_H-\boldsymbol{\lambda})\|_{\mathbf{V}} + \|(T-T_h)\boldsymbol{\lambda} + (\hat{T}-\hat{T}_h)\boldsymbol{f}\|_{\mathbf{V}}\right) \\
& \leqslant C\,h\,\left(\|\boldsymbol{\lambda}_H-\boldsymbol{\lambda}\|_{\boldsymbol{\Lambda}} + h^k\,|\boldsymbol{u}|_{k+1,\mathscr{P}}\right) \\
(5.43)\qquad & \leqslant C\,h\,\left(h^k\,|\boldsymbol{u}|_{k+1,\mathscr{P}} + H^{\ell+1}\,|\mathbf{C}\,\underline{\boldsymbol{\varepsilon}}(\boldsymbol{u})|_{\ell+1,\mathscr{P}}\right)\,.
\end{aligned}
$$

The result follows by replacing (5.35), (5.40), and (5.41) in (I) and (5.43) in (II). ∎

In the case Assumption (B) does not hold, we can still improve the convergence rate in the $\boldsymbol{L}^2(\Omega)$-norm. Indeed, observe that the proof employs Assumption (B) in (5.41) only and, therefore, we can revisit it using Korn's inequality (2.16). This result is presented next.

Corollary 5.8. *Assume the hypothesis of Theorem 5.4 and Assumption (C) hold. Then, there exists C such that*

$$
(5.44)\qquad \|\boldsymbol{u}-\boldsymbol{u}_{Hh}\|_{0,\Omega} \leq C\,\mathcal{H}\,\left(h^k\,|\boldsymbol{u}|_{k+1,\mathscr{P}} + H^{\ell+1}\,|\mathbf{C}\,\underline{\boldsymbol{\varepsilon}}(\boldsymbol{u})|_{\ell+1,\mathscr{P}}\right)\,.
$$

Proof. From the stability of $\hat{T}$ (2.25) and $\hat{T}_h$ in (3.8) to get

$$
\|\underline{\boldsymbol{\varepsilon}}((\hat{T}-\hat{T}_h)\mathbf{e})\|_{0,\mathscr{P}} \leqslant \|\underline{\boldsymbol{\varepsilon}}(\hat{T}\mathbf{e})\|_{0,\mathscr{P}} + \|\underline{\boldsymbol{\varepsilon}}(\hat{T}_h\mathbf{e})\|_{0,\mathscr{P}} \leqslant C\,\mathcal{H}\,\|\mathbf{e}\|_{0,\Omega}\,.
$$

Hence, item (II) can be bounded as follows

$$
\begin{aligned}
&(T\boldsymbol{\lambda}_H - T_h\boldsymbol{\lambda}_H + \hat{T}\boldsymbol{f} - \hat{T}_h\boldsymbol{f}, \mathbf{e})_{\mathscr{P}} \\
&\quad \leq \|T\boldsymbol{\lambda}_H - T_h\boldsymbol{\lambda}_H + \hat{T}\boldsymbol{f} - \hat{T}_h\boldsymbol{f}\|_{0,\Omega}\|\mathbf{e}\|_{0,\Omega} && \text{(Cauchy-Schwarz)} \\
&\quad \leq C\,\mathcal{H}\,\|\underline{\boldsymbol{\varepsilon}}(T\boldsymbol{\lambda}_H - T_h\boldsymbol{\lambda}_H + \hat{T}\boldsymbol{f} - \hat{T}_h\boldsymbol{f})\|_{0,\mathscr{P}}\|\mathbf{e}\|_{0,\Omega} && \text{(using (2.16))} \\
&\quad \leq C\,\mathcal{H}\,\Big(\|\underline{\boldsymbol{\varepsilon}}(T\boldsymbol{\lambda} - T_h\boldsymbol{\lambda} + \hat{T}\boldsymbol{f} - \hat{T}_h\boldsymbol{f})\|_{0,\mathscr{P}} \\
&\qquad + \|\underline{\boldsymbol{\varepsilon}}((T - T_h)(\boldsymbol{\lambda} - \boldsymbol{\lambda}_H))\|_{0,\mathscr{P}}\Big)\|\mathbf{e}\|_{0,\Omega} && (\pm(T - T_h)\boldsymbol{\lambda}) \\
&\quad \leq C\,\mathcal{H}\,\Big(\|\boldsymbol{u} - \boldsymbol{u}_{Hh}\|_{\mathbf{V}} + \|\boldsymbol{\lambda} - \boldsymbol{\lambda}_H\|_{\boldsymbol{\Lambda}}\Big)\|\mathbf{e}\|_{0,\Omega} && \text{(using (2.25) and (3.8))} \\
&\quad \leq C\,\mathcal{H}\,\left(h^k\,|\boldsymbol{u}|_{k+1,\mathscr{P}} + H^{\ell+1}\,|\mathbf{C}\,\underline{\boldsymbol{\varepsilon}}(\boldsymbol{u})|_{\ell+1,\mathscr{P}}\right)\,\|\mathbf{e}\|_{0,\Omega} && \text{(using Theorem 5.4)}
\end{aligned}
$$

Finally, the estimate (5.44) results from the approach used to prove (5.27) in Theorem 5.7. ∎

5.4. $\mathcal{H}$- versus H-refinement. Error estimates from the previous sections indicate that, in addition to the standard strategy to refine the first-level mesh ($\mathcal{H} \to 0$), convergence also arises from enhancing approximation spaces on faces only under mild conditions on the partition $\mathscr{P}$. Such a convergence can be achieved in two different forms (or in a combination of both):

(i) Refining the face partitions (H-refinement), where the first-level meshes satisfy Assumption (A);

(ii) Increasing polynomial degrees on faces (ℓ-refinement), where the polynomial degrees in $\mathbf{V}_h$ and $\boldsymbol{\Lambda}_H$ must respect the condition $k \geq \ell + d$.

In both scenarios, the first-level partition $\mathscr{P}$ stays fixed, and still, convergence occurs under local regularity assumptions. Here, we focus on convergence associated with item (i) since the dependence of the constants in error estimates in terms of k and ℓ stays an open problem. Thereby, we compare theoretically two types of convergence in this section (and numerically in the next section), namely,

- $H \to 0$, the H-refinement. The global partition $\mathscr{P}$ stays fixed;
- $\mathcal{H} \to 0$, the $\mathcal{H}$-refinement. This is the standard mesh-based strategy.

First, it is worth observing that the constants in Theorems 5.4 and 5.6 and (5.21) do not degenerate in the H-convergence case since $\sigma_{\mathcal{T}}$, C_{korn}, C_P and C_{tr} stays unchanged as convergence evolves. As a result, the MHM method provides optimal convergence in the norms $\|\cdot\|_{\mathbf{V}}$, $\|\cdot\|_{1,\mathscr{P}}$ and $\|\cdot\|_{\boldsymbol{\Lambda}}$ on general coarse partition $\mathscr{P}$, which only needs to fullfil Assumption (A). On the opposite, this attractive property may not occur in the standard mesh-based approach due to the constants dependency

in Theorems 5.4 and 5.6 and (5.21) in terms of $\sigma_{\mathcal{T}}$, C_{korn}, C_P and C_{tr} . To ensures convergence in the $\mathcal{H}$-refinement scenario, we replace Assumption (A) by the following stronger assumption:

Assumption (D): Let $\{\mathscr{P}_{\mathcal{H}}\}_{\mathcal{H}>0}$ be a family of global partitions of Ω.

For each $\mathscr{P}_{\mathcal{H}} \in \{\mathscr{P}_{\mathcal{H}}\}_{\mathcal{H}>0}$, there exists Ξ_H such that the family of partitions $\{\Xi_H\}_{H>0}$ satisfies Assumption (A), and such that C_{korn}, C_P, C_{tr}, $\sigma_{\mathcal{T}}$ and σ_{Ξ} stay bounded when $\mathcal{H} \to 0$.

Let $\mathscr{P}$ be a partition of Ω, and $\{\mathscr{P}_{\mathcal{H}}\}_{\mathcal{H}>0}$ be a sequence of global partitions of $\mathscr{P}$ that satisfies Assumption (D). The next result focus on the mesh-based convergence ($\mathcal{H} \to 0$) and is a straightforward consequence of Theorem 5.4 and Theorem 5.6, and Corollary 5.8 under Assumptions (C) and (D).

Corollary 5.9. *Assume $\boldsymbol{u} \in \boldsymbol{H}^{k+1}(\mathscr{P})$ solves (2.3), with $\mathbf{C}\,\underline{\boldsymbol{\varepsilon}}(\boldsymbol{u}) \in H^{\ell+1}(\mathscr{P})^{d\times d}$, where $k \geq \ell + d$ and $\ell \geqslant 1$. There exists C, such that*

$$\|\boldsymbol{\lambda} - \boldsymbol{\lambda}_{Hh}\|_{\boldsymbol{\Lambda}} + \|\boldsymbol{u} - \boldsymbol{u}_{Hh}\|_{1,\mathscr{P}_{\mathcal{H}}} \leq C\,\mathcal{H}^{\ell+1}\,\left(|\boldsymbol{u}|_{k+1,\mathscr{P}} + |\mathbf{C}\,\underline{\boldsymbol{\varepsilon}}(\boldsymbol{u})|_{\ell+1,\mathscr{P}}\right)\,. \tag{5.45}$$

Moreover, if $\mathscr{P}$ satisfies Assumption (C), then it holds

$$\|\boldsymbol{u} - \boldsymbol{u}_{Hh}\|_{0,\Omega} \leq C\,\mathcal{H}^{\ell+2}\,\left(|\boldsymbol{u}|_{k+1,\mathscr{P}} + |\mathbf{C}\,\underline{\boldsymbol{\varepsilon}}(\boldsymbol{u})|_{\ell+1,\mathscr{P}}\right)\,. \tag{5.46}$$

Remark 5.10. Observe that Assumption (D) is not too restrictive. Indeed, we can control constants $\sigma_{\mathcal{T}}$, σ_{Ξ} and C_{tr} assuming the regularity of the family of meshes $\{\mathscr{P}_{\mathcal{H}}\}_{\mathcal{H}>0}$. As for the constant C_P, it corresponds to the maximum value among the local Poincaré constants $C_{P,K}$ in (2.15), and then, constant C_P stays bounded for a quite general family of partitions as discussed in Remark 2.1. The Korn's constant C^K_{korn} is independent of the diameter of K on Lipschitz domains K as proved in Lemma A.1. However, a rigorous proof on the C^K_{korn} dependency on the shape of K stays an open question for the space $\widetilde{\mathbf{V}}(K)$. That said, some insights on the dependency of C^K_{korn} can be found in [19] when the domain is star-shaped. Its validity to functions belonging to the space $\widetilde{\mathbf{V}}(K)$ stays an open problem.

Remark 5.11. Convergence still holds under lower regularity of $\boldsymbol{u}$ in the usual norms. Indeed, assuming $\boldsymbol{u} \in \boldsymbol{H}^{m+1}(\mathscr{P})$ and $1 \leqslant m < k$, error estimates in Theorem 5.4, Theorem 5.6, and Theorem 5.7 an Corollary 5.8 remain valid if one replaces k by m and ℓ by $\min\{m-1, \ell\}$, respectively, in these estimates.

Summary: We gather in Tables 5.1 and 5.2 the error estimates obtained throughout the previous sections under the different assumptions.

TABLE 5.1
Error estimates in the $\boldsymbol{H}^1(\mathscr{P})$-norm for the displacement ($\boldsymbol{u}$), in the $\boldsymbol{\Lambda}$-norm for traction ($\boldsymbol{\lambda}$), and in $\boldsymbol{L}^2(\Omega)$-norm for stress ($\boldsymbol{\sigma}$).

Error	Assumption	
	A	D
$\|\boldsymbol{u}-\boldsymbol{u}_{Hh}\|_{1,\mathscr{P}}$	$h^k+H^{\ell+1}$	$\mathcal{H}^{\ell+1}$
$\|\boldsymbol{\lambda}-\boldsymbol{\lambda}_{Hh}\|_{\boldsymbol{\Lambda}}$	$h^k+H^{\ell+1}$	$\mathcal{H}^{\ell+1}$
$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_{Hh}\|_{0,\Omega}$	$h^k+H^{\ell+1}$	$\mathcal{H}^{\ell+1}$

TABLE 5.2
Error estimates in the $\boldsymbol{L}^2(\Omega)$-norm for the displacement ($\boldsymbol{u}$).

Error	Assumption				
	A	A and C	A, B and C	D	C and D
$\|\boldsymbol{u}-\boldsymbol{u}_{Hh}\|_{0,\Omega}$	$h^k+H^{\ell+1}$	$\mathcal{H}\,(h^k+H^{\ell+1})$	$h^{k+1}+H^{\ell+2}$	$\mathcal{H}^{\ell+1}$	$\mathcal{H}^{\ell+2}$

6. Numerical Results. This section assesses the MHM method numerically. The first part verifies the theoretical results using two- and three-dimensional test cases with analytical solutions. In the process, we perform a comparison between the two mesh refinement strategies, i.e. the $\mathcal{H}$- and H-refinements, from the perspectives of accuracy and computational cost. A second test considers an elastic model with heterogeneous coefficients for which we compare the MHM and the Galerkin methods with respect to a reference solution obtained on a finer mesh. All test cases assume isotropic elastic media, i.e., the stiffness tensor $\mathbf{C}$ is characterized by the shear modulus μ and the Poisson's ratio ν as follows:

$$\mathbf{C}_{ijkl} := 2\,\mu\,\delta_{ik}\,\delta_{jl}+\frac{2\,\mu\,\nu}{1-2\,\nu}\delta_{ij}\,\delta_{kl} \quad \text{or} \quad \mathbf{C}\,\underline{\boldsymbol{\varepsilon}}(\boldsymbol{u}) = 2\,\mu\,\underline{\boldsymbol{\varepsilon}}(\boldsymbol{u})+\frac{2\,\mu\,\nu}{1-2\,\nu}(\nabla\cdot\boldsymbol{u})I_{d\times d}\,, \tag{6.1}$$

where δ_{ij} denotes the Kronecker delta and $I_{d\times d}$ is the identity tensor in $\mathbb{R}^{d\times d}$.

Before heading to numerics, we present the underlying algorithm and some of the main implementation aspects of the MHM method.

6.1. Algorithm and implementation aspects. Let $\{\boldsymbol{\psi}_i\}_{i=1}^{m_{\boldsymbol{\Lambda}}}$ and $\{\boldsymbol{\phi}_i\}_{i=1}^{m_0}$ be bases for $\boldsymbol{\Lambda}_H$ and $\mathbf{V}_{\mathrm{rm}}$, respectively, where $m_{\boldsymbol{\Lambda}} = \dim\boldsymbol{\Lambda}_H$ and $m_0 = \dim\mathbf{V}_{\mathrm{rm}}$. Also, we define $A\in\mathbb{R}^{m_{\boldsymbol{\Lambda}}\times m_{\boldsymbol{\Lambda}}}$, $B\in\mathbb{R}^{m_{\boldsymbol{\Lambda}}\times m_0}$, $f\in\mathbb{R}^{m_{\boldsymbol{\Lambda}}}$, and $f_0\in\mathbb{R}^{m_0}$ with entries

$$\begin{aligned} A_{ij} &= \sum_{K\in\mathscr{P}}\int_{\partial K}\boldsymbol{\psi}_i T_h(\boldsymbol{\psi}_j)\,d\boldsymbol{x}\,, \qquad B_{ij} = \sum_{K\in\mathscr{P}}\int_{\partial K}\boldsymbol{\psi}_i\boldsymbol{\phi}_j\,d\boldsymbol{x}\,,\\ f_i &= \sum_{K\in\mathscr{P}}\int_{\partial K}\boldsymbol{\psi}_i\widehat{T}_h(\boldsymbol{f})\,d\boldsymbol{x}\,, \qquad \text{and} \qquad f_{0,i} = \int_\Omega \boldsymbol{f}\cdot\boldsymbol{\phi}_i\,d\boldsymbol{x}\,. \end{aligned} \tag{6.2}$$

The MHM method (3.4) naturally yields the two-level algorithm presented in Algorithm 6.1.

Algorithm 6.1 Upscaling algorithm for the MHM method in (3.4).

1: **for all** $K \in \mathscr{P}$ **do**
2: Compute $T_h(\boldsymbol{\psi}_i)\,|_K$ solving (3.6), for all $i = 1, \cdots, m_{\boldsymbol{\Lambda}}$;
3: Compute $\hat{T}_h(\boldsymbol{f})\,|_K$ solving (3.7);
4: Compute $\int_{\partial K} \boldsymbol{\psi}_i T_h(\boldsymbol{\psi}_j)\, d\boldsymbol{x}$, $\int_{\partial K} \boldsymbol{\psi}_i \boldsymbol{\phi}_r\, d\boldsymbol{x}$, $\int_{\partial K} \boldsymbol{\psi}_i \widehat{T}_h(\boldsymbol{f})\, d\boldsymbol{x}$, and $\int_K \boldsymbol{f} \cdot \boldsymbol{\phi}_r\, d\boldsymbol{x}$, for all $i, j = 1, \cdots, m_{\boldsymbol{\Lambda}}$ and $r = 1, \cdots, m_0$;
5: Assemble in A, B, f, and f_0;
6: Solve the following global system for $(\lambda, u^{rm}) \in \mathbb{R}^{m_{\boldsymbol{\Lambda}} \times m_0}$

$$\begin{pmatrix} A & B \\ B^t & 0 \end{pmatrix} \begin{pmatrix} \lambda \\ u^{rm} \end{pmatrix} = \begin{pmatrix} f \\ f_0 \end{pmatrix}; \tag{6.3}$$

7: Compute $\boldsymbol{u}_{Hh}$, on all $K \in \mathscr{P}$, as follows

$$\boldsymbol{u}_{Hh} := \sum_{i=1}^{m_0} u_i^{rm} \boldsymbol{\phi}_i + \sum_{i=1}^{m_{\boldsymbol{\Lambda}}} \lambda_i T_h(\boldsymbol{\psi}_i) + \widehat{T}_h(\boldsymbol{f})\,,$$

using the solutions $\lambda = (\lambda_i)_{i=1}^{m_{\boldsymbol{\Lambda}}}$ and $u^{rm} = (u_i^{rm})_{i=1}^{m_0}$.

At that point, some comments are necessary:

- Terms $T_h(\boldsymbol{\psi}_j)$ correspond to the multiscale basis, which along with $\hat{T}_h(\boldsymbol{f})$, respond for upscaling the information from the local level to the global one.
- The "embarrassingly" parallel stage (Steps 2-4) computes the local integrals through (6.2), which are assembled "on the fly" into the global linear system to obtain the coordinates of the solutions;
- The equivalence (3.9) is particularly useful when one deals with multi-query problems wherein $\boldsymbol{f}$ changes multiple times. In this case, Steps 2 and 3, as well as part of Steps 4 and 5, can be computed at an offline stage.

We obtained all the results presented in the remainder of this section with the help of an in-house simulation software, which is called MSL (MHM Set of Libraries). MSL uses the C++11 standard, and its software architecture accommodates OpenMP and MPI technologies to allow for both shared-memory and distributed-memory hardware architectures. MSL is also pluggable to professional software components such as:

- ParMETIS [31], for the partitioning of $\mathscr{P}$ across different processors;
- MUMPS [2], for the parallel solution of the global linear system;
- Eigen [24], for the solution of the local linear systems and other linear algebra computations; and
- TetGen [42], for the generation of the local meshes.

6.2. Two- and three-dimensional analytical cases. Let $\Omega := [0,1]^d$ be an isotropic elastic domain with Lamé coefficients $\mu = 1$ and $\lambda = 1$, and homogeneous Dirichlet boundary condition on $\partial\Omega$. Let $\boldsymbol{f}$ be such that the two-dimensional solution

$\boldsymbol{u} = \{u_1, u_2\}$ of (2.1) is defined by

$$u_1(x,y) = \sin(2\pi x)\sin(2\pi y),$$
$$u_2(x,y) = 2\sin(2\pi x)\sin(2\pi y)\,,$$

and the three-dimensional solution $\boldsymbol{u} = \{u_1, u_2, u_3\}$ of (2.1) is defined by

$$u_1(x,y,z) = \sin(2\pi x)\sin(2\pi y)\sin(2\pi z),$$
$$u_2(x,y,z) = 2\sin(2\pi x)\sin(2\pi y)\sin(2\pi z),$$
$$u_3(x,y,z) = 3\sin(2\pi x)\sin(2\pi y)\sin(2\pi z).$$

Using these problems, we first perform numerical tests to verify the theoretical convergence rates of the MHM method.

6.2.1. Verifying theoretical results. The convergence history is measured in the $\boldsymbol{L}^2(\Omega)$- and $\boldsymbol{H}^1(\mathscr{P})$-norms for displacement, and in the following $H(\mathbf{div};\mathcal{T}_h;\mathbb{S})$-norm for the post-processed stress tensor

$$\|\boldsymbol{\sigma}\|_{\mathrm{div}} := \sqrt{\sum_{K\in\mathscr{P}}\sum_{\tau\in\mathcal{T}_h{}^K} \|\boldsymbol{\sigma}\|^2_{0,\tau} + h^2_\Omega \,\|\nabla\cdot\,\boldsymbol{\sigma}\|^2_{0,\tau}} \quad \text{for all } \boldsymbol{\sigma}\in H(\mathbf{div};\mathcal{T}_h;\mathbb{S})\,. \tag{6.4}$$

We restrict ourselves to sequences of meshes satisfying Assumptions (A) and (D), and $h \leqslant H \leqslant \mathcal{H}$. The model solutions above naturally satisfy Assumption (C). In Figure 6.1, we show the convergence curves using the $\mathcal{H}$-refinement for the three dimensional problem. We adopt $k = \ell + 3$, with $\ell = 1$, 2, a family $\{\mathscr{P}_\mathcal{H}\}_{\mathcal{H}>0}$ of nested structured tetrahedral meshes, one-element face partitions $\mathcal{E}_H^K$, and one-element local meshes $\mathcal{T}_h^K$, i.e., $H = \mathcal{H} = h$.

We observe a convergence of $O(H^{\ell+2})$ in the $\boldsymbol{L}^2(\Omega)$-norm and $O(H^{\ell+1})$ using the $\boldsymbol{H}^1(\mathscr{P})$-norm as predicted by the theory in Section 5.4, and $O(H^\ell)$ convergence for the stress $\boldsymbol{\sigma}_{Hh}$ in the $H(\mathbf{div};\mathcal{T}_h;\mathbb{S})$-norm. The latter is not covered by the present theory.

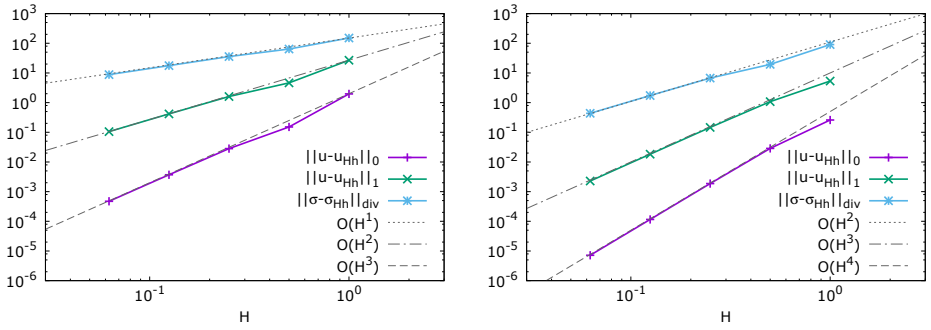

FIGURE 6.1. *Three-dimensional case with a structured simplicial mesh family.* Left: $\ell = 1$. Right: $\ell = 2$.

In Figure 6.3 (left), we show the $\mathcal{H}$-convergence curves for a family of L-shaped meshes,

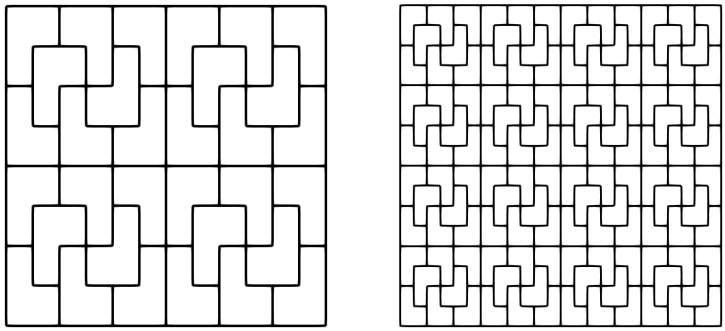

FIGURE 6.2. *Two meshes from the L-shaped mesh family.*

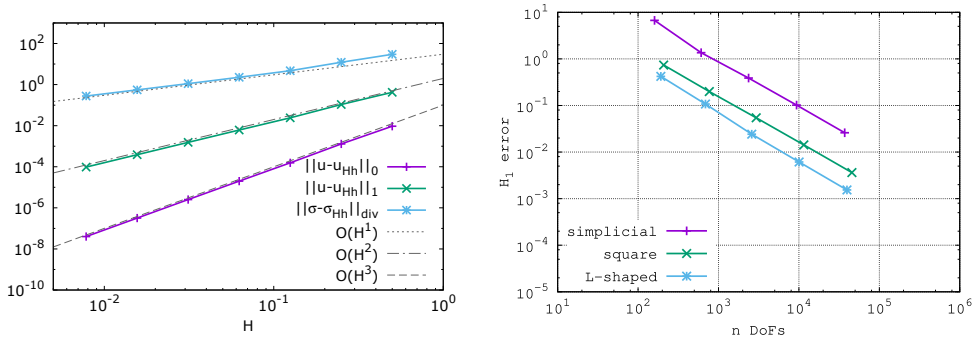

FIGURE 6.3. *Convergence curves for the two-dimensional case using $\ell = 1$.* Left: *Convergence using the L-shaped mesh family.* Right: *Error in terms of the number of global degrees of freedom (DoFs) for three types of mesh families: (structured) simplicial, square and L-shaped.*

using $H = \mathcal{H}$, $\ell = 1$ and $k = 3$. This family of structured non-convex meshes follow the pattern shown in Figure 6.2. Here we recover all the convergence rates from theory, and also the convergence of $O(H^\ell)$ for the stress $\boldsymbol{\sigma}_{Hh}$ in the $H(\mathbf{div}; \mathcal{T}_h; \mathbb{S})$-norm. Notice that the error in the $\boldsymbol{L}^2(\Omega)$-norm decays at $O(H^{\ell+2})$ even though the elements $K \in \mathscr{P}$ are not convex as anticipated in the theory.

Still in Figure 6.3 (right), we compare the error from three families of structured meshes: simplicial, square and L-shaped, using $\ell = 1$ and $k = 3$. Note that by increasing the number of sides in the mesh elements, the MHM method becomes more precise for the same number of global degrees of freedom (DoFs).

Next, we investigate the convergence in terms of H while keeping $\mathcal{H}$ fixed. Figure 6.4 depicts that the MHM method is super-convergent in this context. Indeed, additional convergence rates of order $H^{1/d}$ appear, which is not predicted by the theory. Figure 6.5 compares the $\mathcal{H}$- and H-refinements. Interesting, the result points that the H-refinement needs fewer degrees of freedom to reach the same level of error compared with $\mathcal{H}$-refinements. On the other hand, the underlying smaller global system associated with the H-refinement approach has a matrix with a broader band, and the local problems become somewhat more numerous. Thereby, the best deal turns out to be to offset H- and $\mathcal{H}$-refinements. Such a claim is verified in Subsection 6.2.2.

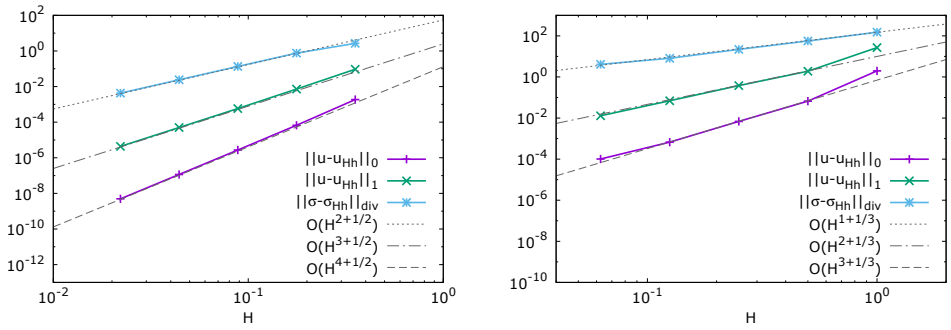

FIGURE 6.4. *Errors using the H-refinement.* Left: *global mesh composed by 4 (equal) squares in the two-dimensional case using $\ell = 2$.* Right: *global mesh composed by 24 (equal) tetrahedra in the three-dimensional case using $\ell = 1$.*

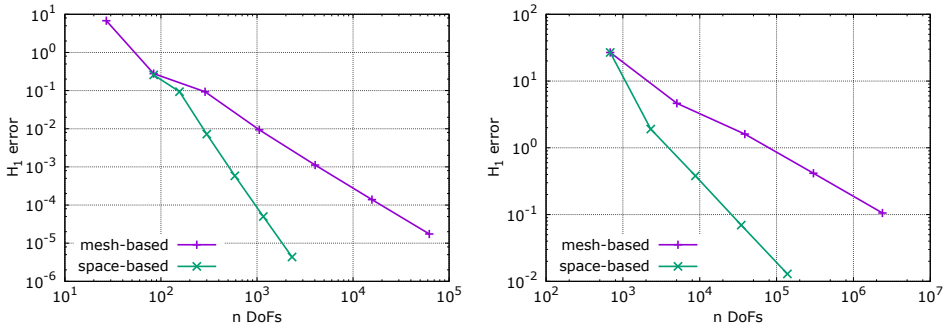

FIGURE 6.5. *A comparison between the mesh-based ($\mathcal{H}$) and space-based (H) refinements.* Left: *two-dimensional case.* Right: *three-dimensional case.*

6.2.2. Computational performance of H- and $\mathcal{H}$-refinements. In this section, we do a performance evaluation of the MHM algorithm. Our focus is on measuring the impact of the coupled global problem and the independent local problems on execution time and memory consumption.

The experiments presented herein were conducted on a cluster of computing nodes. Each node in this cluster has 2 CPU sockets Intel Xeon E5-2695v2, each one with 12 cores operating at 2.4 GHz, and 64 GB of RAM. An Infiniband FDR network interconnects the nodes. To take into account problems of reasonable computational complexity, all the configurations employed in the remainder of this section were conducted in the three-dimensional analytical case.

First, we set upper bounds on the order of approximation in the $\boldsymbol{L}^2(\Omega)$-norm ($O(10^{-4})$) and in the $\boldsymbol{H}^1(\mathscr{P})$-norm ($O(10^{-1})$). We then considered three configurations aiming at these orders of approximation, but accounting for different levels of H- and $\mathcal{H}$-refinements. We set $k = 4$, $\ell = 1$, and $h = H$ in all three configurations. We ran these configurations using 2 computing nodes, with 1 MPI process per node socket, summing up 4 MPI processes in a simulation. Table 6.1 describes these

configurations.

TABLE 6.1
Simulation configurations to compare H-refinement and $\mathcal{H}$-refinement: #DoFs = $m_{\mathbf{\Lambda}} + m_0$; #NNZ = total number of non-zero entries of matrices A, B and B^t in (6.3).

Configuration	$\mathcal{H}$	#Elements per face	#DoFs	#NNZ
1	1/16	1	2,386,944	154,440,295
2	1/3	16	206,064	197,531,136
3	1	64	34,704	120,434,688

Table 6.2 summarizes the performance results related with the solution of the global problem. Looking at these results alone suggests that the $\mathcal{H}$-refinement (represented by configuration 1) is by far the most unsatisfactory for this specific scenario. Nevertheless, a comparison between configurations 2 and 3, again taking the global problem alone, is inconclusive.

TABLE 6.2
Simulation results related to the global problem: F+S time = time (in seconds) for factorization and forward/backward substitution; RAM = maximum amount of memory (in MBytes) of the most demanding MPI process.

Configuration	F+S time	RAM	$\|\boldsymbol{u} - \boldsymbol{u}_{Hh}\|_{0,\Omega}$	$\|\boldsymbol{u} - \boldsymbol{u}_{Hh}\|_{1,\mathscr{P}}$
1	206.53	8,840.54	6.68×10^{-4}	1.49×10^{-1}
2	60.16	3,131.23	4.45×10^{-4}	6.85×10^{-2}
3	52.87	760.11	9.51×10^{-4}	9.82×10^{-2}

Table 6.3 summarizes the performance results related to the solution of the local problems. Irrespective of the fact that we set $k = 4$, $\ell = 1$, and $h = H$ in all three configurations, not only the number (which is directly related to $\mathcal{H}$) but also the size of the local problems vary in each configuration. In our implementation (see Algorithm 6.1), the set of local problems associated to an element $K \in \mathscr{P}$ is solved independently of the other elements in $\mathscr{P}$.

In all three configurations, the simulation software partitions the set of local problems across all available cores such that each core schedules a single element $K \in \mathscr{P}$ to be solved at a time. Besides, the overall procedure to solve each of the local problems comprises local mesh generation and numerical integrations, since $\mathscr{P}$ and the refinement levels of $\mathcal{E}_H$ are determined at runtime. For configurations 1 and 2, we use a single thread for the local problems in an element $K \in \mathscr{P}$ without thread synchronizations or context switches to minimize potential sources of parallel overhead. For configuration 3, we use two threads for local problems in K, since the number of cores is two times the amount of local problems in this configuration; thus, some thread synchronization is needed. This myriad of possible scheduling configurations renders a system that, in contrast to the global problem, is difficult to be easily instrumented in terms of the solution of the linear systems associated with each local problem. We therefore chose to present in Table 6.3 the aggregated performance related to the complete configuration and solution of all local problems computed by the most demanding computing resource.

TABLE 6.3

Simulation results related to the local problems: F+S time = aggregated time (in seconds) for mesh generation, numerical integration, and linear system resolution (factorization+fwd/bwd substitution) of all local problems at the most demanding core; RAM = maximum amount of memory (in MBytes) of the most demanding MPI process.

Configuration	F+S time	RAM
1	840.00	6,270.06
2	60.00	6,760.00
3	660.00	24,674.14

The results in Table 6.2 and Table 6.3 together suggest that configuration 2 provides the best balance between overall execution time and memory consumption. This balance is particularly important if the local problems are to be computed *online* with the global problem. We conclude that the best choice to accommodate precision, computational cost, and memory allocation stems from a right balance between H-refinement and $\mathcal{H}$-refinement.

6.3. A high-contrast multilayered problem. We assess the MHM method's performance in a high-contrast scenario using the geophysical domain proposed in the HPC4E Seismic Test Suite 1.0 [17]. The original domain is a three-dimensional box split into 16 horizontal layers with constant properties, as shown in Figure 6.6. The interfaces between layers are defined by 6,001 × 6,001 matrices that store depth coordinates on a fine-scale regular grid. The top interface determines the terrain topography, and the bottom layer goes 4,480.78 m deep.

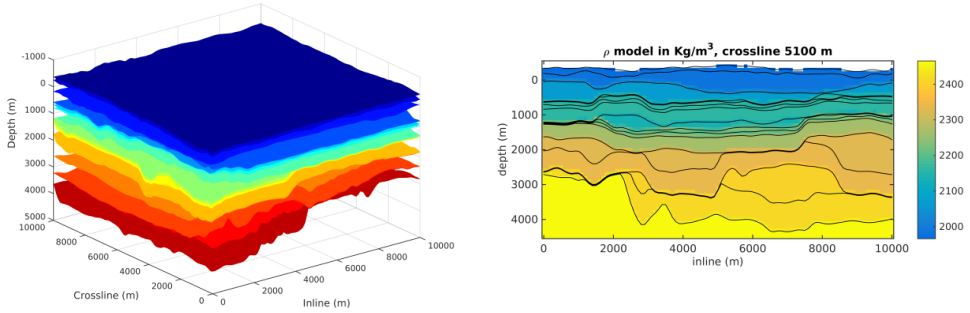

FIGURE 6.6. *The HPC4e Test Suite (adapted from [17]).* Left: *physical domain divided in layers.* Right: *a vertical cut showing the density ρ.*

We consider a two-dimensional isotropic elasticity problem defined on the slice $y = 0$ of the domain mentioned above. Fig. 6.7 shows the interfaces between layers together with two non-aligned meshes. We prescribe null displacement at the bottom boundary, and the remaining parts of the boundary are free-surface, i.e., with null traction. Layers 4 and 12 (from top to bottom) contain saturated clay (Young's modulus: 15 MPa, Poisson's ratio: 0.49, and mass density: 1,760 kg/m^3). The remaining

layers are filled with materials provided in [17]. The source function is the piecewise constant weight, using gravity acceleration of 9.8 m/s^2. The z direction represents depth, so it increases from top to bottom.

First, we address the performance of the MHM method on coarse meshes. We adopt the 128 elements mesh shown in Figure 6.7 (left) to compute the approximate solution with the MHM method. Most $K \in \mathscr{P}$ are quadrilaterals, except those in the layer touching the top boundary, which are polygons with 378 sides. Such a choice of discretization accounts for the topography precisely while it verifies the robustness of the developed MHM method to deal with non-convex general polygons. Each face of $\mathcal{E}$ uses a two-element mesh, and we choose $\ell = 3$ (with $k = \ell + 2 = 5$). We prescribed the homogeneous Dirichlet boundary condition (null displacement) as an essential condition in $\mathbf{V}_h$ following the idea proposed in [27]. Also, notice that the homogeneous Neumann boundary condition (free-surface) to the MHM formulation can be defined as an essential condition in the space $\mathbf{\Lambda}_H$. Notably, we look for the traction unknown $\boldsymbol{\lambda}$ in

$$\mathbf{\Lambda}_{0,N} := \{\boldsymbol{\tau}\, \boldsymbol{n}^K|_{\partial K} \,:\, \boldsymbol{\tau} \in H_{0,N}(\mathbf{div}; \Omega; \mathbb{S}) \quad \text{for all } K \in \mathscr{P}\}\,, \tag{6.5}$$

instead of $\mathbf{\Lambda}$, where $\partial\Omega_N \subset \partial\Omega$ denotes the free surface boundary. This setup results in an MHM formulation with only 4,048 global DoFs, while the local problems drive multiscale basis functions.

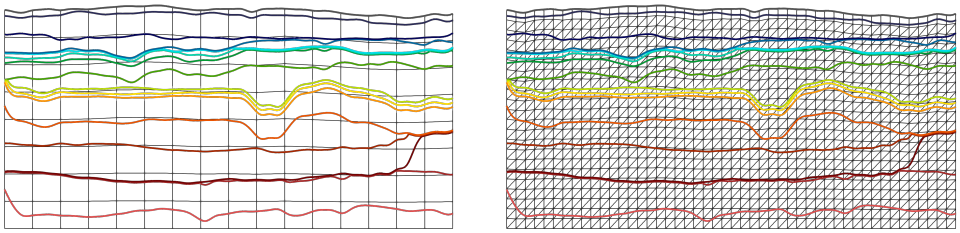

FIGURE 6.7. *Non-aligned global meshes used by the MHM method. Colorful lines represent interfaces between physical layers.*

Figure 6.8 shows the comparison between the solution from the MHM method with 4,048 global DoFs, and a reference solution from the standard Galerkin method (stdFEM) with $[\mathbb{P}_3(K)]^2$ elements on a simplicial fine mesh (7,717,979 elements) aligned with all layers, using a total of 69,491,849 DoFs. We observe that the MHM method recovers precisely the displacement components u_x and u_z of the references ones. To illustrate the importance of adopting multiscale basis functions on top of coarse meshes, we show in Figure 6.9 the numerical solution using the stdFEM with $[\mathbb{P}_2(K)]^2$ interpolation on a coarse simplicial mesh (90,258 DoFs). We see that the solution degrades considerably.

The second experiment compares the MHM method with the stdFEM on simplicial meshes. To establish a fair comparison, we use the mesh illustrated in Figure 6.7 (right) since it contains simplicial elements almost everywhere. We choose here $\ell = 1$ (with $k = \ell + 2 = 3$) and two-element meshes on edges. The resulting global MHM

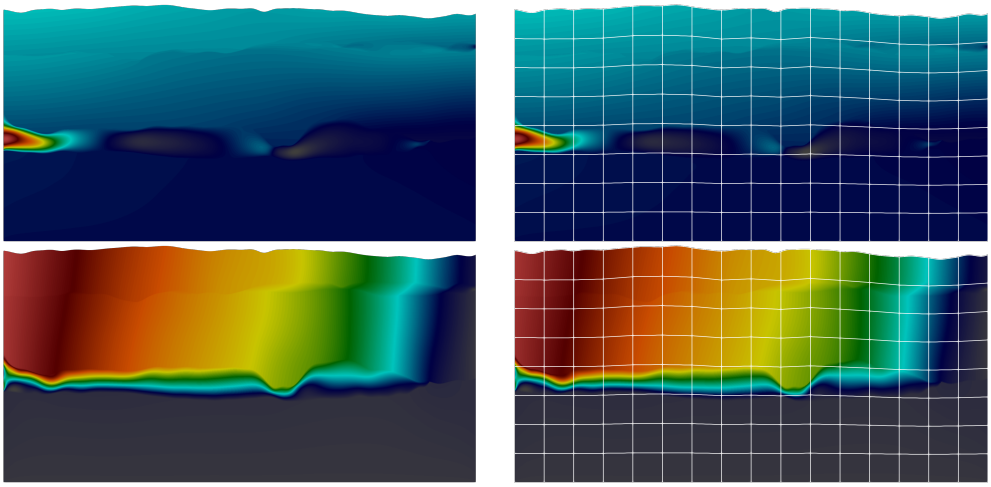

FIGURE 6.8. *Displacement component u_x (resp. u_z) obtained with the MHM method (*top-right*) (resp.* bottom-right*) and the reference solution (*top-left*) (resp.* bottom-left*) with the overlaid mesh.*

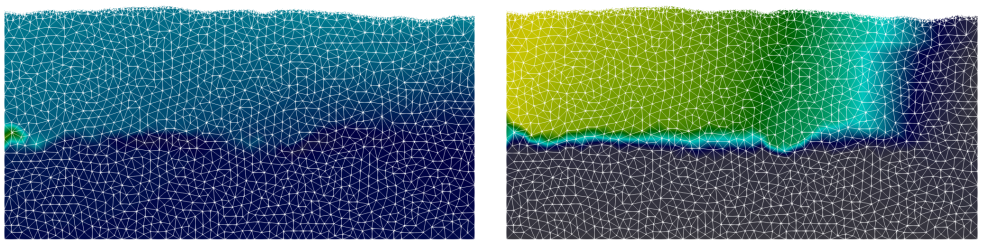

FIGURE 6.9. *Displacement components of the solution using the standard finite element method stdFEM with 90,258 DoFs with the overlaid mesh.* Left: *x-component.* Right: *z-component.*

formulation has 30,352 global DoFs. To compare methods with the same order of convergence, we employ the stdFEM with $\mathbb{P}_2(K)$ interpolation. Figure 6.10 compares the profiles of the MHM and two stdFEM solutions at $x = 0$ using two-level of refinements, namely,

- `stdFEM (coarse)`: uses mesh given Figure 6.9 with 90,258 DoFs;
- `stdFEM (fine)`: uses a non-aligned mesh with 2,912,070 DoFs (724,420 elements).

We depict the results in Figure 6.10, wherein we observe that the MHM solution coincides with the one from `stdFEM (fine)` using 95 times less DoFs.

Some clarifications are essential at that point. The significant improvements in precision as a result of the MHM method on coarse meshes comes with the price to solve 2,064 independent local elasticity problems to generate the basis functions. In the present problem, the largest local problem amounts to 10,172 DoFs. That means that if one used 2,064 processing units in parallel, the total execution time would, in theory, be proportional to the assemble+solution of two linear systems, namely the local system with 10,172 DoFs, followed by the global one with 30,352 DoFs. Such trivial parallelism for solving local problems is one of the main reasons for the adoption of

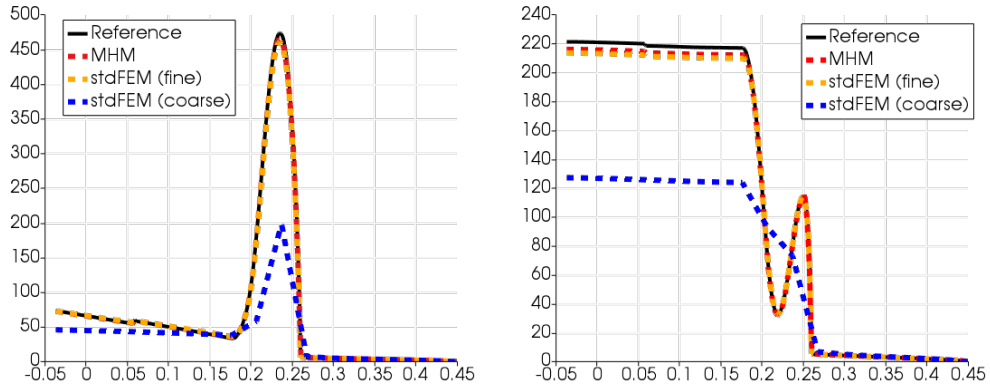

FIGURE 6.10. *Profiles of the displacement solution at $x = 0$. We compare the MHM method with the stdFEM using two levels of mesh refinement for the stdFEM, and with the reference solution. The x- (left) and z-component (right) of the displacement.*

the MHM method in larger problems, for which the lack of memory prohibits the use of very fine meshes.

7. Conclusions. We presented a new family of polytopal finite elements for the MHM method applied to the linear elasticity equations with discontinuous interpolations on faces, and refined local meshes. These properties allow the method to cope with elasticity models with heterogeneous coefficients in complex geometries using coarse meshes with non-aligned faces and interfaces with jumping coefficients. We established sufficient conditions for the well-posedness of the MHM method. We also highlighted the impact of the second-level approximation on convergence and showed that the MHM method is super-convergent under local regularity assumptions with no additional post-processing. Notably, we proved that the method achieves convergence without deteriorating Poincaré and Korn's constants under a mild condition on face partitions. Moreover, we showed that the error estimates for displacement in the $\boldsymbol{L}^2(\Omega)$-norm could be improved even on non-convex element meshes.

We also verified the theoretical finding of the MHM method for elasticity for the first time numerically. Two- and three-dimensional numerical experiments recovered theoretical results, and we noticed unexpected additional convergence rates of order $H^{\frac{1}{d}}$ in the H-refinement. Such a numerical finding also appeared in the MHM method for the Poisson equation [10] and is currently under theoretical investigation. The numerical tests also inferred that the H-refinement requires fewer degrees of freedom than the $\mathcal{H}$-refinement to reach the same error threshold. On the other hand, the computational performance tests suggested that the H- and $\mathcal{H}$-refinements should be combined to achieve better precision in shorter overall execution time and memory consumption.

Finally, we assessed the MHM method on a two-dimensional synthetic multilayered problem with complex topography. We found that current strategy needed about one hundred times fewer degrees of freedom to achieve a specified error threshold compared with the standard continuous finite element method. Such a huge gain of the MHM method hides the costs for obtaining the basis functions. Fortunately,

such computations are local and completely independent of one another, and may be performed in an offline stage.

Acknowledgements. All authors were partially supported by the MCTI/RNP-Brazil under the HPC4E Project. The second author was partially supported by the CNPq/Brazil No. 140764/2015-1, by the company Bull Ltda and the SCAC from the French Embassy in Brazil under the CIFRE program. The third author was partially supported by CNPq/Brazil No. 301576/2013-0. The authors acknowledge the National Laboratory for Scientific Computing (LNCC/MCTI, Brazil) for providing HPC resources of the SDumont supercomputer, which have contributed to the research results reported within this paper. URL: http://sdumont.lncc.br

References.

[1] A. Abdulle, E. Weinan, B. Engquist, and E. Vanden-Eijnden. The heterogeneous multiscale method. *Acta Numerica*, 21:1–87, apr 2012. doi: 10.1017/s0962492912000025. Cited on page 1.

[2] P. Amestoy, I. S. Duff, J. Koster, and J.-Y. L'Excellent. A fully asynchronous multifrontal solver using distributed dynamic scheduling. *SIAM Journal on Matrix Analysis and Applications*, 23(1):15–41, 2001. Cited on page 38.

[3] T. Apel, A.-M. Sändig, and J. R. Whiteman. Graded mesh refinement and error estimates for finite element solutions of elliptic boundary value problems in non-smooth domains. *Mathematical Methods in the Applied Sciences*, 19(1): 63–85, jan 1996. doi: 10.1002/(sici)1099-1476(19960110)19:1⟨63::aid-mma764⟩3.0.co;2-s. Cited on page 26.

[4] T. Arbogast. Analysis of a two-scale, locally conservative subgrid upscaling for elliptic problems. *SIAM Journal on Numerical Analysis*, 42(2):576–598, jan 2004. doi: 10.1137/s0036142902406636. Cited on page 1.

[5] D. N. Arnold, F. Brezzi, and J. Douglas. Peers: a new mixed finite element for plane elasticity. *Japan J. Appl. Math*, 1(2):347–367, 1984. Cited on page 1.

[6] D. N. Arnold, G. Awanou, and R. Winther. Finite elements for symmetric tensors in three dimensions. *Mathematics of Computation*, 77(263):1229–1251, sep 2008. doi: 10.1090/s0025-5718-08-02071-1. Cited on page 1.

[7] D. N. Arnold, G. Awanou, and R. Winther. Nonconforming tetrahedral mixed finite elements for elasticity. *Math. Models Methods Appl. Sci.*, 23:783–796, 2014. Cited on page 1.

[8] I. Babuška and J. E. Osborn. Generalized finite element methods: Their performance and their relation to mixed methods. *SIAM Journal on Numerical Analysis*, 20(3):510–536, jun 1983. doi: 10.1137/0720034. Cited on page 1.

[9] G. Barrenechea, F. Jaillet, D. Paredes, and F. Valentin. The multiscale hybrid mixed method in general polygonal meshes. *Numer. Math.*, https://doi.org/10.1007/s00211-020-01103-5, 2020. Cited on pages 2, 20, and 23.

[10] G. R. Barrenechea, F. Jaillet, D. Paredes, and F. Valentin. The Multiscale Hybrid Mixed Method in General Polygonal Meshes. hal-02054681, Mar. 2019. URL https://hal.inria.fr/hal-02054681. Cited on pages 15, 16, and 46.

[11] C. Bernardi. Optimal finite-element interpolation on curved domains. *SIAM Journal on Numerical Analysis*, 26(5):1212–1240, oct 1989. doi: 10.1137/0726068. Cited on page 18.

[12] D. Boffi, F. Brezzi, and M. Fortin. *Mixed Finite Element Methods and Applica-*

tions. Springer Berlin Heidelberg, 2013. doi: 10.1007/978-3-642-36519-5. Cited on page 16.

[13] S. C. Brenner. Korn's inequalities for piecewise H^1 vector fields. *Mathematics of Computation*, 73(247):1067–1088, sep 2003. doi: 10.1090/s0025-5718-03-01579-5. Cited on pages 5 and 51.

[14] S. C. Brenner and L.-Y. Sung. Linear finite element methods for planar linear elasticity. *Mathematics of Computation*, 59(200):321–321, 1992. doi: 10.1090/s0025-5718-1992-1140646-2. Cited on page 26.

[15] M. Crouzeix and P.-A. Raviart. Conforming and nonconforming finite element methods for solving the stationary stokes equations i. *Revue française d'automatique informatique recherche opérationnelle. Mathématique*, 7(R3):33–75, 1973. doi: 10.1051/m2an/197307r300331. Cited on pages 23 and 24.

[16] L. B. da Veiga, F. Brezzi, and L. D. Marini. Virtual elements for linear elasticity problems. *SIAM Journal on Numerical Analysis*, 51(2):794–812, jan 2013. doi: 10.1137/120874746. Cited on page 1.

[17] J. de la Puente. HPC4E Seismic Test Suite. https://hpc4e.eu/downloads/datasets-and-software, 2016. Copyright Josep de la Puente (Barcelona Supercomputing Center) 2016. Licenced under the Creative-Commons Attribution-ShareAlike 4.0 International License. Cited on pages 43 and 44.

[18] P. Devloo, A. M. Farias, S. Gomes, W. Pereira, J. Santos, and F. Valentin. New $H(div)$-conforming multiscale hybrid-mixed methods for the elasticity problem on polygonal meshes. https://hal.archives-ouvertes.fr/hal-02415020, 2019. Cited on page 14.

[19] R. G. Durán. An elementary proof of the continuity from $L^2_0(\Omega)$ to $H^1_0(\Omega)^n$ of bogovskii's right inverse of the divergence. http://arxiv.org/abs/1103.3718v1, Mar. 2011. Cited on page 36.

[20] A. Ern and J.-L. Guermond. *Theory and Practice of Finite Elements*, volume 159 of *Applied Mathematical Sciences*. Springer-Verlag New York, New York, 1 edition, 2004. ISBN 978-1-4757-4355-5. doi: 10.1007/978-1-4757-4355-5. Cited on pages 3, 5, 6, 9, 11, 16, 18, 26, 31, and 50.

[21] A. T. A. Gomes, D. Paredes, W. Pereira, R. P. Souto, and F. Valentin. Performance analysis of the MHM simulator in a petascale machine. In *Proceedings of the XXXVIII Iberian Latin American Congress on Computational Methods in Engineering*. ABMEC Brazilian Association of Computational Methods in Engineering, 2017. doi: 10.20906/CPS/CILAMCE2017-0381. Cited on page 2.

[22] J. Gopalakrishnan and J. Guzmán. Symmetric nonconforming mixed finite elements for linear elasticity. *SIAM J. Numer. Anal.*, 49(4):1504–1520, 2011. Cited on page 1.

[23] J. Gopalakrishnan and W. Qiu. An analysis of the practical DPG method. *Mathematics of Computation*, 83(286):537–552, may 2013. doi: 10.1090/s0025-5718-2013-02721-4. Cited on pages 15 and 19.

[24] G. Guennebaud, B. Jacob, et al. Eigen v3. http://eigen.tuxfamily.org, 2010. Cited on page 38.

[25] C. Harder and F. Valentin. Foundations of the MHM method. In *Lecture Notes in Computational Science and Engineering*, pages 401–433. Springer International Publishing, 2016. doi: 10.1007/978-3-319-41640-3_13. Cited on page 14.

[26] C. Harder, D. Paredes, and F. Valentin. A family of multiscale hybrid-mixed finite element methods for the darcy equation with rough coefficients. *Journal of Computational Physics*, 245:107–130, jul 2013. doi: 10.1016/j.jcp.2013.03.019. Cited on page 14.

[27] C. Harder, D. Paredes, and F. Valentin. On a multiscale hybrid-mixed method for advective-reactive dominated problems with heterogeneous coefficients. *Multiscale Modeling & Simulation*, 13(2):491–518, jan 2015. doi: 10.1137/130938499. Cited on page 44.

[28] C. Harder, A. L. Madureira, and F. Valentin. A hybrid-mixed method for elasticity. *ESAIM: Mathematical Modelling and Numerical Analysis*, 50(2):311–336, feb 2016. doi: 10.1051/m2an/2015046. Cited on pages 2, 7, 10, 14, 20, 21, and 51.

[29] T. Y. Hou and X.-H. Wu. A multiscale finite element method for elliptic problems in composite materials and porous media. *Journal of Computational Physics*, 134 (1):169–189, jun 1997. doi: 10.1006/jcph.1997.5682. Cited on page 1.

[30] T. J. Hughes, G. R. Feijóo, L. Mazzei, and J.-B. Quincy. The variational multiscale method—a paradigm for computational mechanics. *Computer Methods in Applied Mechanics and Engineering*, 166(1-2):3–24, nov 1998. doi: 10.1016/s0045-7825(98)00079-6. Cited on page 1.

[31] G. Karypis and V. Kumar. A fast and high quality multilevel scheme for partitioning irregular graphs. *SIAM J. Sci. Comput.*, 20(1):359–392, Dec. 1998. ISSN 1064-8275. Cited on page 38.

[32] A. Målqvist and D. Peterseim. Localization of elliptic multiscale problems. *Mathematics of Computation*, 83(290):2583–2603, jun 2014. doi: 10.1090/S0025-5718-2014-02868-8. Cited on page 1.

[33] A. L. Mazzucato and V. Nistor. Well-posedness and regularity for the elasticity equation with mixed boundary conditions on polyhedral domains and domains with cracks. *Archive for Rational Mechanics and Analysis*, 195(1):25–73, oct 2008. doi: 10.1007/s00205-008-0180-y. Cited on page 26.

[34] L. E. Payne and H. F. Weinberger. An optimal poincaré inequality for convex domains. *Archive for Rational Mechanics and Analysis*, 5(1):286–292, jan 1960. doi: 10.1007/bf00252910. Cited on page 7.

[35] W. Pereira and F. Valentin. A locking-free MHM method for elasticity. In *Proceeding Series of the Brazilian Society of Computational and Applied Mathematics*, volume 5. SBMAC, apr 2017. doi: 10.5540/03.2017.005.01.0336. Cited on page 14.

[36] D. A. D. Pietro and J. Droniou. A hybrid high-order method for leray–lions elliptic equations on general meshes. *Mathematics of Computation*, 86(307):2159–2191, dec 2016. doi: 10.1090/mcom/3180. Cited on page 7.

[37] D. A. D. Pietro and J. Droniou. *The Hybrid High-Order Method for Polytopal Meshes*. Springer International Publishing, 2020. doi: 10.1007/978-3-030-37203-3. Cited on page 1.

[38] D. A. D. Pietro and A. Ern. *Mathematical Aspects of Discontinuous Galerkin Methods*. Springer Berlin Heidelberg, 2012. doi: 10.1007/978-3-642-22980-0. Cited on pages 6 and 7.

[39] D. A. D. Pietro and A. Ern. A hybrid high-order locking-free method for linear elasticity on general meshes. *Computer Methods in Applied Mechanics and Engineering*, 283:1–21, jan 2015. doi: 10.1016/j.cma.2014.09.009. Cited on page 1.

[40] P. A. Raviart and J. M. Thomas. A mixed finite element method for 2-nd order elliptic problems. In *Lecture Notes in Mathematics*, pages 292–315. Springer Berlin Heidelberg, 1977. doi: 10.1007/bfb0064470. Cited on pages 15 and 20.

[41] L. R. Scott and S. Zhang. Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Mathematics of Computation*, 54(190):483–483, may 1990. doi: 10.1090/s0025-5718-1990-1011446-7. Cited on page 25.

[42] H. Si. Tetgen, a delaunay-based quality tetrahedral mesh generator. *ACM Trans. Math. Softw.*, 41(2), Feb. 2015. ISSN 0098-3500. doi: 10.1145/2629697. URL https://doi.org/10.1145/2629697. Cited on page 38.

[43] R. Stenberg. On the construction of optimal mixed finite element methods for the linear elasticity problem. *Numer. Math.*, 48(4):447–462, 1986. Cited on page 1.

[44] A. Veeser and R. Verfurth. Poincare constants for finite element stars. *IMA Journal of Numerical Analysis*, 32(1):30–47, may 2011. doi: 10.1093/imanum/drr011. Cited on page 7.

[45] B. Zhang, J. Zhao, Y. Yang, and S. Chen. The nonconforming virtual element method for elasticity problems. *Journal of Computational Physics*, 378(2):394–410, feb 2019. ISSN 00219991. doi: 10.1016/j.jcp.2018.11.004. URL http://link.springer.com/10.1007/s10444-020-09743-9https://linkinghub.elsevier.com/retrieve/pii/S0021999118307228. Cited on page 1.

[46] W. Zheng and H. Qi. On friedrichs–poincaré-type inequalities. *Journal of Mathematical Analysis and Applications*, 304(2):542–551, apr 2005. doi: 10.1016/j.jmaa.2004.09.066. Cited on page 7.

Appendix A. Auxiliary results.

Lemma A.1. *Let K be a bounded domain in $\mathbb{R}^d$ ($d = 2, 3$) with Lipschitz boundary. There exists a positive constant $C_{korn,K}$, independent of $\mathcal{H}_K$, such that*

$$\|\nabla\tilde{\boldsymbol{v}}\|_{0,K} \leqslant C_{korn,K}\|\boldsymbol{\varepsilon}(\tilde{\boldsymbol{v}})\|_{0,K} \quad \textit{for all } \tilde{\boldsymbol{v}} \in \widetilde{\mathbf{V}}(K)\,.$$

Proof. First, we use the Petree-Tartar's lemma (see [20, Lemma A.38]) with

$$X := \widetilde{\mathbf{V}}(K) \quad \text{and} \quad Y := [L^2(K)]^{d\times d} \quad \text{and} \quad Z = \boldsymbol{L}^2(K) \quad \text{and} \quad A := \boldsymbol{\varepsilon}\,.$$

Notice $(X, \|\cdot\|_{1,K})$ is closed and the mapping $A : X \to Y$ is injective. Next, it is well-known that $H^1(K) \overset{c}{\hookrightarrow} L^2(K)$ (see Rellich-Kondrachov [20, Theorem B.46] for $d \geqslant 2$), and therefore, the mapping $X \subset \boldsymbol{H}^1(K) \overset{c}{\hookrightarrow} Z$ defines a compact operator $T : X \to Z$. Also, from the Korn's inequality in $\boldsymbol{H}^1(K)$ (see [20, Theorem 3.78]) it holds

$$\|\tilde{\boldsymbol{v}}\|_{1,K} \leqslant \|A\,\tilde{\boldsymbol{v}}\|_{0,K} + \|T\tilde{\boldsymbol{v}}\|_{0,K} \quad \text{for all } \tilde{\boldsymbol{v}} \in \widetilde{\mathbf{V}}(K)\,.$$

We fullfil, then, the requirements of the Petree-Tartar's lemma which leads to the existence of a positive constant $C_{korn,K}$ such that

$$\|\nabla\tilde{\boldsymbol{v}}\|_{0,K} \leqslant \|\tilde{\boldsymbol{v}}\|_{1,K} \leqslant C_{korn,K}\,\|A\,\tilde{\boldsymbol{v}}\|_{0,K} = C_{korn,K}\,\|\underline{\boldsymbol{\varepsilon}}(\tilde{\boldsymbol{v}})\|_{0,K} \quad \text{for all } \tilde{\boldsymbol{v}} \in \widetilde{\mathbf{V}}(K)\,.$$

The constant $C_{korn,K}$ is independent of $\mathcal{H}_K$ by a scaling argument, and the result follows. ∎

We present a Korn's inequality over functions in $\mathbf{V}$ with a constant independent of mesh parameters.

Lemma A.2. *Let $\mathscr{P}$ be a polytopal partition. There exists a positive constant C, depending only on Ω and the shape-regularity of $\mathscr{P}$, such that*

$$|\boldsymbol{v}|^2_{1,\mathscr{P}} \leqslant C \left(\|\underline{\boldsymbol{\varepsilon}}(\boldsymbol{v})\|^2_{0,\mathscr{P}} + \sum_{K\in\mathscr{P}} \left\| \Pi_{RM}\boldsymbol{v} - \frac{1}{|K|}\int_K \boldsymbol{v}\, d\boldsymbol{x} \right\|^2_{0,K} + \sum_{E\in\mathcal{E}_0} \frac{1}{h_E} \|[\![\boldsymbol{v}]\!]\|^2_{0,E} \right),$$

for all $\boldsymbol{v}\in\mathbf{V}$.

Proof. The proof uses the technique proposed in [13]. Define the operator $\Phi : \mathbf{V} \to \mathbb{R}$ as follows

$$\Phi(\boldsymbol{v}) = \sqrt{\sum_{K\in\mathscr{P}} \left\| \Pi_{RM}\boldsymbol{v} - \frac{1}{|K|}\int_K \boldsymbol{v}\,\mathrm{d}\boldsymbol{x} \right\|^2_{0,K}}.$$

Notice that it is a semi-norm and satisfies

$$\begin{aligned}
&\Phi((I-\Pi_{RM})\boldsymbol{v}) = 0\,,\\
&\Phi(\boldsymbol{v}) \leqslant \|\boldsymbol{v}\|_{0,\Omega}\,,\\
&\Phi(\boldsymbol{v}) = 0 \quad \text{and} \quad \boldsymbol{v}\in\mathbf{V}_{\mathrm{rm}}(\Omega) \qquad \Leftrightarrow \qquad \boldsymbol{v} \text{ is a constant vector}\,,
\end{aligned}$$

and, thus, the operator $\Phi(\cdot)$ fulfils the requirements of [13, Lemma 2.2] and [13, Theorem 3.1]. As such, using [13, Theorem 4.2] ($d=2$) or [13, Theorem 5.2] ($d=3$), there exists a constant C, depending only on the shape-regularity of $\mathscr{P}$, such that

$$\|\nabla\boldsymbol{v}\|^2_{0,\mathscr{P}} \leqslant C \left(\|\boldsymbol{\varepsilon}(\boldsymbol{v})\|^2_{0,\mathscr{P}} + \Phi(\boldsymbol{v})^2 + \sum_{E\in\mathcal{E}_0} \frac{1}{h_E} \|[\![\boldsymbol{v}]\!]\|^2_{0,E} \right) \quad \text{for all } \boldsymbol{v}\in\mathbf{V}\,,$$

and the result follows. ∎

The next result slightly improves the equivalence result proved in [28, Lemma A.3].

Lemma A.3. *Given $\boldsymbol{\mu}\in\boldsymbol{\Lambda}$, the following result holds*

$$\|\boldsymbol{\mu}\|_{\boldsymbol{\Lambda}} = \sup_{\boldsymbol{v}\in\mathbf{V}} \frac{\langle \boldsymbol{\mu}, \boldsymbol{v}\rangle_{\partial\mathscr{P}}}{\|\boldsymbol{v}\|_{\mathbf{V}}}\,.$$

Proof. From [28, Lemma A.3], we have

$$\|\boldsymbol{\mu}\|_{\boldsymbol{\Lambda}} \geqslant \sup_{\boldsymbol{v}\in\mathbf{V}} \frac{\langle \boldsymbol{\mu}, \boldsymbol{v}\rangle_{\partial\mathscr{P}}}{\|\boldsymbol{v}\|_{\mathbf{V}}} \quad \text{for all } \boldsymbol{\mu}\in\boldsymbol{\Lambda}\,.$$

Next, given $\boldsymbol{\mu} \in \boldsymbol{\Lambda}$ from the Riesz representation theorem there exists $\boldsymbol{u}^{\boldsymbol{\mu}} \in \mathbf{V}$ such that

$$(\boldsymbol{u}^{\boldsymbol{\mu}}, \boldsymbol{v})_{\mathbf{V}} = \langle \boldsymbol{\mu}, \boldsymbol{v} \rangle_{\partial \mathscr{P}} \qquad \text{for all } \ \boldsymbol{v} \in \mathbf{V} \,.$$

Notice the above problem corresponds to the strong formulation (in a distributional sense)

$$\begin{cases} h_\Omega^2 \nabla \cdot (\underline{\boldsymbol{\varepsilon}}(\boldsymbol{u}^{\boldsymbol{\mu}})) = \boldsymbol{u}^{\boldsymbol{\mu}} & \text{in } K \,, \\ (\underline{\boldsymbol{\varepsilon}}(\boldsymbol{u}^{\boldsymbol{\mu}}))\boldsymbol{n}^K = \boldsymbol{\mu} & \text{on } \partial K \,, \end{cases}$$

in each $K \in \mathscr{P}$. Therefore the tensor function $\underline{\boldsymbol{\varepsilon}}(\boldsymbol{u}^{\boldsymbol{\mu}}) \in H(\mathbf{div}; \Omega; \mathbb{S})$ and is satisfies

$$\begin{aligned} \|\underline{\boldsymbol{\varepsilon}}(\boldsymbol{u}^{\boldsymbol{\mu}})\|_{\mathrm{div}}^2 &= \|\underline{\boldsymbol{\varepsilon}}(\boldsymbol{u}^{\boldsymbol{\mu}})\|_{0,\Omega}^2 + h_\Omega^2 \|\nabla \cdot (\underline{\boldsymbol{\varepsilon}}(\boldsymbol{u}^{\boldsymbol{\mu}}))\|_{0,\Omega}^2 \\ &= \|\underline{\boldsymbol{\varepsilon}}(\boldsymbol{u}^{\boldsymbol{\mu}})\|_{0,\Omega}^2 + h_\Omega^{-2} \|\boldsymbol{u}^{\boldsymbol{\mu}}\|_{0,\Omega}^2 = \|\boldsymbol{u}^{\boldsymbol{\mu}}\|_{\mathbf{V}}^2 \,. \end{aligned}$$

Also, given $\boldsymbol{\mu} \in \boldsymbol{\Lambda}$, it holds

$$\|\boldsymbol{\mu}\|_{\boldsymbol{\Lambda}} \, \|\boldsymbol{u}^{\boldsymbol{\mu}}\|_{\mathbf{V}} \leqslant \|\underline{\boldsymbol{\varepsilon}}(\boldsymbol{u}^{\boldsymbol{\mu}})\|_{\mathrm{div}} \, \|\boldsymbol{u}^{\boldsymbol{\mu}}\|_{\mathbf{V}} = \|\boldsymbol{u}^{\boldsymbol{\mu}}\|_{\mathbf{V}}^2 = \langle \boldsymbol{\mu}, \boldsymbol{u}^{\boldsymbol{\mu}} \rangle_{\partial \mathscr{P}} \,,$$

which implies that

$$\|\boldsymbol{\mu}\|_{\boldsymbol{\Lambda}} \leqslant \sup_{\boldsymbol{v} \in \mathbf{V}} \frac{\langle \boldsymbol{\mu}, \boldsymbol{v} \rangle_{\partial \mathscr{P}}}{\|\boldsymbol{v}\|_{\mathbf{V}}} \,,$$

and the result follows. ∎