The three-field formulation for elasticity problems

Franco Brezzi[*1,2] and **Luisa Donatella Marini** [1,2]

[1] Dipartimento di Matematica, Università di Pavia, Via Ferrata 1, I-27100 Pavia, Italy
[2] IMATI-CNR Pavia, Via Ferrata 1, I-27100 Pavia, Italy

Received 15 November 2004, revised 30 November 2004, accepted 2 December 2004
Published online 3 December 2004

Key words Domain Decomposition, Non matching grids, Three field formulation, Stabilizations, Inf-Sup conditions
MSC (2000) 65N12, 65N22, 65N30, 65N55

The three-field decomposition method is particularly suited for decompositions with non matching grids. It corresponds to introduce an additional grid (usually uniform, or "easy") at the interface. The unknown is then represented independently in each subdomain *and* on the interface. The matching between its value in each subdomain and on the interface is provided by suitable Lagrange multipliers. Here we discuss the main features of the method for a linear three-dimensional elasticity problem, in the simplest case of two subdomains. An easy numerical test to check whether the *inf-sup* conditions (necessary for the stability) are satisfied is also presented.

1 Introduction

Sometimes, there are coincidences, in research, so striking that the only explanation seems to be: *times were ripe*. Indeed, in April 1992 two papers were presented almost simultaneously at two different Meetings (one [18] in Nice, France, from 8 to 10, and one [1] in Dallas, Texas, from 13 to 15) containing exactly the same idea. The two groups of authors did not know each other, and they belonged to two different communities (Mathematicians and Engineers). It is totally impossible that the information could have travelled from one group to another, and the only possible explanation is that *they had the same idea at the same time*. Actually both groups traced back the origin of their idea to previous works in the engineering literature: [36], and [4], respectively. Still, as both references were already relatively "old and cold" at the time, their simultaneous revival is a remarkable coincidence.

In order to see the idea, let us first see the problem: in several applications, one has to solve problems that couple several domains arriving each with its own decomposition, done by an independent team within the factory. As an example, the wing and the fuselage of an airplane are typically studied first by different groups, and the relative grids are constructed independently one from another. If you are in charge of solving a problem involving the whole plane, you have to do something about these two *non matching grids*. Similarly, in several applications, one needs a local refinement in a specific subdomain, and would like to do it independently of the decomposition of the remaining part of the domain. Here again, if one

[*] Corresponding author: e-mail: brezzi@imati.cnr.it, Phone: +39 0382 548 211, Fax: +39 0382 548 300

wants to avoid a local remeshing near the transition interface, possibly having to use rather distorted elements, the problem of *non matching grids* pops out again. There are obviously other applications where you cannot avoid non matching grids: for instance when studying sliding pieces [14], or contact problems [25], or several other types of problems (see e.g. [3], [34], [32], [39], [33] and the references therein). In a very crude and schematic way, the situation is depicted in Fig. 1.

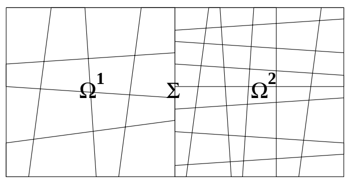

Fig. 1 Example of non matching grids

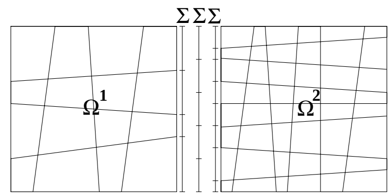

Fig. 2 The three grids on Σ

Using a linear elasticity problem as a model problem, assume that $\mathbf{u}^1$ and $\mathbf{u}^2$ represent the displacement fields in the two subdomains Ω^1 and Ω^2, respectively, and that, for the sake of simplicity, we are setting the displacements to zero all over the boundary of the whole domain Ω, union of Ω^1 and Ω^2. The total potential energy can be presented as the sum of the works $\mathcal{F}_k$ of the external forces in each subdomain Ω^k, plus the contributions $\mathcal{E}_k$ to the internal energy of each subdomain Ω^k :

$$\mathcal{E}_{tot} = \mathcal{F}_1 + \mathcal{F}_2 + \mathcal{E}_1 + \mathcal{E}_2, \tag{1.1}$$

where $\mathcal{F}_1$ and $\mathcal{E}_1$ depend on $\mathbf{u}^1$ while $\mathcal{F}_2$ and $\mathcal{E}_2$ depend on $\mathbf{u}^2$. We would like to minimize $\mathcal{E}_{tot}$ over all displacement fields that satisfy the kinematic boundary conditions on the boundary of Ω plus suitable continuity requirements on the interface Σ between Ω^1 and Ω^2.

The first possibility is to force the continuity by means of a suitable set of Lagrange multipliers. This amounts to add to $\mathcal{E}_{tot}$ an interface contribution

$$\mathcal{L} = L(\mathbf{u}, \boldsymbol{\lambda}) := \int_\Sigma \boldsymbol{\lambda} \cdot (\mathbf{u}^1 - \mathbf{u}^2)\, \mathrm{d}S \tag{1.2}$$

and to require the stationarity of the functional

$$\mathcal{F}_1 + \mathcal{F}_2 + \mathcal{E}_1 + \mathcal{E}_2 + \mathcal{L}. \tag{1.3}$$

At the equilibrium, $\boldsymbol{\lambda}$ will represent the normal component of the stress field on Σ. This is possibly the most common way of dealing with the problem, and the celebrated *mortar method* (see e.g. [8], [31], [33], [39]) is actually based on it.

The idea that we are talking about here is however different. It amounts to introduce a new mesh on the interface Σ (different, in general, from both the decompositions induced on Σ by the two given decompositions of Ω^1 and Ω^2). There we introduce a new representation, $\boldsymbol{\psi}$, of the displacement field, and we introduce *two* Lagrange multipliers (one for each subdomain);

$$\mathcal{L}_1 = \mathcal{L}_1(\mathbf{u}^1, \boldsymbol{\lambda}^1, \boldsymbol{\psi}) := \int_\Sigma \boldsymbol{\lambda}^1 \cdot (\mathbf{u}^1 - \boldsymbol{\psi})\, \mathrm{d}S \tag{1.4}$$

and

$$\mathcal{L}_2 = \mathcal{L}_2(\mathbf{u}^2, \boldsymbol{\lambda}^2, \boldsymbol{\psi}) := \int_\Sigma \boldsymbol{\lambda}^2 \cdot (\mathbf{u}^2 - \boldsymbol{\psi})\, \mathrm{d}S. \tag{1.5}$$

Then we consider the stationarity of the functional

$$\mathcal{F}_1 + \mathcal{F}_2 + \mathcal{E}_1 + \mathcal{E}_2 + \mathcal{L}_1 + \mathcal{L}_2. \tag{1.6}$$

At the equilibrium, $\boldsymbol{\lambda}^k$ will represent the (outward) normal component on Σ of the stress field in the subdomain Ω^k.

As we shall see with more details in the following sections, the stationarity of the functional (1.6) implies the equilibrium equations in each subdomain (obtained when taking the derivatives with respect to $\mathbf{u}^1$ and $\mathbf{u}^2$, respectively), but it also implies the continuity equations

$$\mathbf{u}^1 = \boldsymbol{\psi}, \qquad \mathbf{u}^2 = \boldsymbol{\psi}, \tag{1.7}$$

obtained when taking the derivatives with respect to $\boldsymbol{\lambda}^1$ and $\boldsymbol{\lambda}^2$, respectively, and the equilibrium on the interface

$$\boldsymbol{\lambda}^1 + \boldsymbol{\lambda}^2 = 0, \tag{1.8}$$

obtained when taking the derivative with respect to $\boldsymbol{\psi}$.

It is clear that, for the discretized problem, both the equilibrium equations in the subdomains and the interface conditions (1.7)-(1.8) will be imposed only *in a weak sense*.

The main features of the three-field formulation will be presented in the next section. In Section 3 we shall discuss its discretization (essentially using conforming finite elements for each field, possibly with different degrees from one field to another). The interest in using the three field formulation in the context of Domain Decomposition Methods, with the possible use of parallel computers, will be outlined in Section 4. In Section 5 we present sufficient conditions that ensure stability and optimal error bounds in an almost immediate way. As these conditions might be difficult to check in practice, in particular for non mathematicians, in Section 6 we present alternative conditions, that make the convergence proof more difficult but allow a very easy numerical test that can give reliable indications on their validity in each particular case. The test itself is presented in Section 7. Some conclusions are drawn in Section 8. Finally, in Appendices A and B we report the detailed proofs of the two theorems containing the error estimates. More precisely, Appendix A contains the proof of error estimates using the *inf-sup* conditions of Section 5, based on more difficult norms, while Appendix B contains the more complicated proof of error estimates using the *inf-sup* conditions of Section 6, based on easier norms (the ones that can be easily checked with the test of Section 7).

2 The three-field formulation

Let us consider, for the sake of simplicity, a polyhedral domain $\Omega \subset \mathbb{R}^3$, that will be the region occupied by our elastic body. For simplicity we shall only consider the case of a decomposition into *two (polyhedral) subdomains* Ω^1 and Ω^2. Most of the theory (and basically all the practice) will hold unchanged in the case of an arbitrary finite number of subdomains.

We shall try to point out, with suitable remarks, the issues that could require a special attention when dealing with more subdomains. We then assume that

$$\Omega = \Omega^1 \cup \Omega^2, \tag{2.1}$$

and we do not make an issue on whether Ω or the Ω^k's are assumed to be open or closed.

Let us see now the *data* that we are taking for our *linear elasticity model problem*.

We assume that we are given a distributed load vector field $\mathbf{f}$ in Ω, and homogeneous kinematic boundary conditions

$$\mathbf{u} = 0 \tag{2.2}$$

on the whole boundary of Ω (this is not very realistic, but it is just to simplify the formulae).

As we said, we take the model of linear elasticity for a homogeneous and isotropic material, with Hooke's law and corresponding Lamé coefficients L_λ and L_μ. Note that we cannot use the more common symbols λ and μ for Lamé coefficients, since we are going to use these symbols (or, actually, very similar ones) for the Lagrange multipliers at the interface. For simplicity, we are not going to discuss the case of *nearly incompressible materials*, and therefore we shall assume that

$$L_\lambda \simeq L_\mu. \tag{2.3}$$

Our basic unknown will be, as usual, the displacement field $\mathbf{u}$. *A priori* we are not assuming neither $\mathbf{u}$ nor any virtual displacement $\mathbf{v}$ to be continuous across the interface Σ, defined as

$$\Sigma := \partial\Omega^1 \cap \partial\Omega^2. \tag{2.4}$$

The restrictions of $\mathbf{u}$ to each subdomain Ω^k will be denoted with $\mathbf{u}^k$ ($k = 1, 2$), while the three components of the vector $\mathbf{u}$ will be denoted by (u_1, u_2, u_3).

In terms of $\mathbf{u}$ we define in each subdomain the strain tensor $\boldsymbol{\varepsilon}$ as

$$\{\boldsymbol{\varepsilon}\}_{i,j} := \Big(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\Big)/2. \tag{2.5}$$

More generally, for a given (virtual) displacement field $\mathbf{v}$ we define the corresponding (virtual) strain tensor

$$\{\boldsymbol{\varepsilon}(\mathbf{v})\}_{i,j} := \Big(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i}\Big)/2. \tag{2.6}$$

As in (2.5), we shall often simply write $\boldsymbol{\varepsilon}$ for $\boldsymbol{\varepsilon}(\mathbf{u})$. Note that we can easily use (2.5) to define $\boldsymbol{\varepsilon}(\mathbf{u})$ on the whole domain (including the interface Σ) but we do not have the right to define $\boldsymbol{\varepsilon}(\mathbf{v})$ on the interface Σ, as we have in mind virtual displacements that might be discontinuous across it. This is an important point, because, even if the *true solution* $\mathbf{u}$ will surely be continuous, we are going to look for an *approximate solution* $\mathbf{u}_h$ that is discontinuous from one subdomain to another. And we cannot really do better, as we are going to use different meshes in the two subdomains.

Finally, always in terms of $\mathbf{u}$, in each subdomain, we define the stress field as

$$\boldsymbol{\sigma} := 2L_\mu\, \boldsymbol{\varepsilon} + L_\lambda \mathbf{I}\, \mathrm{tr}(\boldsymbol{\varepsilon}), \tag{2.7}$$

where $\mathbf{I}$ is the identity matrix and $\mathrm{tr}(\boldsymbol{\varepsilon})$ is the trace of $\boldsymbol{\varepsilon}$ defined as $\varepsilon_{11}+\varepsilon_{22}+\varepsilon_{33}$. The (virtual) stress field $\boldsymbol{\sigma}(\mathbf{v})$ associated to a virtual displacement $\mathbf{v}$ will then be defined as

$$\boldsymbol{\sigma}(\mathbf{v}) := 2L_\mu\, \boldsymbol{\varepsilon}(\mathbf{v}) + L_\lambda \mathbf{I}\, \mathrm{tr}(\boldsymbol{\varepsilon}(\mathbf{v})). \tag{2.8}$$

As before, we will often write $\boldsymbol{\sigma}$ in place of $\boldsymbol{\sigma}(\mathbf{u})$. We point out again that $\boldsymbol{\sigma}(\mathbf{u})$ is clearly defined in the whole domain Ω, but the generic $\boldsymbol{\sigma}(\mathbf{v})$ is only defined separately in each subdomain Ω^k, $k = 1, 2$.

The solution $\mathbf{u}$ that we are looking for in Ω could be seen, in strong form, as the solution of

$$\mathrm{div}\boldsymbol{\sigma}(\mathbf{u}) + \mathbf{f} = 0 \text{ in } \Omega, \qquad \mathbf{u} = 0 \text{ on } \partial\Omega, \tag{2.9}$$

or alternatively as the minimizer of the total potential energy, as usual. Here however we are going to work with virtual displacements that are allowed to be discontinuous across the interface Σ, and some additional care is required.

The contribution of each subdomain to the internal energy and to the work of external forces can now be made precise as follows:

$$\mathcal{E}_k := \frac{1}{2}\int_{\Omega^k} \boldsymbol{\sigma} : \boldsymbol{\varepsilon}\,\mathrm{d}V \quad \mathcal{F}_k := -\int_{\Omega^k} \mathbf{f}\cdot\mathbf{u}\,\mathrm{d}V. \tag{2.10}$$

The "total energy" associated to a virtual displacement $\mathbf{v}$ will then be

$$\mathcal{E}_{tot}(\mathbf{v}) = \sum_{k=1}^{2}\left(\frac{1}{2}\int_{\Omega^k} \boldsymbol{\sigma}(\mathbf{v}) : \boldsymbol{\varepsilon}(\mathbf{v})\,\mathrm{d}V \;-\int_{\Omega^k} \mathbf{f}\cdot\mathbf{v}\,\mathrm{d}V\right). \tag{2.11}$$

It is now time to make precise the space where we allow the virtual displacements to vary. For this we define first the *internal energy norm* in each subdomain Ω^k ($k = 1, 2$)

$$\|\mathbf{v}\|^2_{E,k} := \int_{\Omega^k} \boldsymbol{\sigma}(\mathbf{v}) : \boldsymbol{\varepsilon}(\mathbf{v})\,\mathrm{d}V \tag{2.12}$$

(that is actually *twice the energy*), and the global (broken) internal energy norm

$$\|\mathbf{v}\|^2_{E,b} := \|\mathbf{v}\|^2_{E,1} + \|\mathbf{v}\|^2_{E,2}. \tag{2.13}$$

We define then the space $\mathbf{V}$ of virtual displacements as

$$\mathbf{V} := \{\mathbf{v} \text{ such that } \|\mathbf{v}\|_{E,b} \text{ is finite and } \mathbf{v} = 0 \text{ on } \partial\Omega\}. \tag{2.14}$$

It is worth noticing that the norm (2.13) is naturally associated with the scalar product

$$(\mathbf{u},\mathbf{v})_{E,b} := \sum_{k=1}^{2}\int_{\Omega^k} \boldsymbol{\sigma}(\mathbf{u}) : \boldsymbol{\varepsilon}(\mathbf{v})\,\mathrm{d}V \;\equiv\; \sum_{k=1}^{2}\int_{\Omega^k} \boldsymbol{\sigma}(\mathbf{v}) : \boldsymbol{\varepsilon}(\mathbf{u})\,\mathrm{d}V, \tag{2.15}$$

and that the usual Cauchy-Schwarz inequality holds

$$(\mathbf{u},\mathbf{v})_{E,b} \leq \|\mathbf{u}\|_{E,b}\,\|\mathbf{v}\|_{E,b}. \tag{2.16}$$

The idea would be to minimize

$$\mathcal{E}_{tot}(\mathbf{v}) \equiv \frac{1}{2}\|\mathbf{v}\|^2_{E,b} - \int_\Omega \mathbf{f}\cdot\mathbf{v}\,\mathrm{d}S \tag{2.17}$$

over all possible virtual displacements $\mathbf{v} \in \mathbf{V}$, but this will clearly be wrong, since we did not (yet) require any continuity at the interface Σ. The minimum, in this case, would be attained by the solution of *two independent problems*, one in each subdomain, with (homogeneous) kinematic boundary conditions on $\partial\Omega^k \cap \partial\Omega$ and no-tension boundary conditions ($\boldsymbol{\sigma}\cdot\mathbf{n} = 0$) on both sides of Σ (which is *not* what we want).

Hence we will introduce *a second representation* of the displacement field (that we denote by $\boldsymbol{\psi}$) on Σ, and we force the matching of both $\mathbf{u}^1$ and $\mathbf{u}^2$ with $\boldsymbol{\psi}$ by means of suitable *Lagrange multipliers* (one for each side of Σ). Note that each virtual displacement $\mathbf{v} \in \mathbf{V}$ is double-valued on Σ, but $\boldsymbol{\psi}$ is assumed to be single-valued. Actually, it will be better to introduce a whole space $\boldsymbol{\Phi}$ of *virtual displacements on* Σ. We first define, for $k = 1, 2$, the space $\mathbf{D}^k$ as the space of the restrictions to Σ (or *traces* on Σ) of the virtual displacements $\mathbf{v}^k$ that have finite energy in Ω^k:

$$\mathbf{D}^k := \{\boldsymbol{\delta}^k \text{ such that } \exists\, \mathbf{v}^k \in \mathbf{V}^k \text{ with } \mathbf{v}^k_{|\Sigma} = \boldsymbol{\delta}^k\}. \tag{2.18}$$

Actually, we remark that, with our definition, and in our particular case of *two subdomains*, we have $\mathbf{D}^1 \equiv \mathbf{D}^2$. In each $\mathbf{D}^k$ we introduce the (natural) norm:

$$\|\boldsymbol{\delta}^k\|_{\mathbf{D},k} := \inf_{\mathbf{v}^k_{|\Sigma} = \boldsymbol{\delta}^k} \|\mathbf{v}\|_{E,k} \tag{2.19}$$

that is, in other words, *the lowest possible energy of a virtual displacement in* Ω^k *that is equal to* $\boldsymbol{\delta}^k$ *on* Σ. Note that, even though the two spaces $\mathbf{D}^1$ and $\mathbf{D}^2$ are equal (since we have only two subdomains) *the two norms* $\|\cdot\|_{\mathbf{D},1}$ *and* $\|\cdot\|_{\mathbf{D},2}$ *will be different*, in general, unless the domain Ω is symmetric with respect to Σ. It can be proved, however, that the two norms are *equivalent*, in the sense that you can bound one of them by a constant (depending only on Ω^1 and Ω^2) times the other. We then define

$$\mathbf{D} := \{\boldsymbol{\delta} = (\boldsymbol{\delta}^1, \boldsymbol{\delta}^2) \text{ with } \boldsymbol{\delta}^1 \in \mathbf{D}^1 \text{ and } \boldsymbol{\delta}^2 \in \mathbf{D}^2\} \tag{2.20}$$

with

$$\|\boldsymbol{\delta}\|^2_{\mathbf{D}} := \sum_{k=1}^{2} \|\boldsymbol{\delta}^k\|^2_{\mathbf{D},k}. \tag{2.21}$$

As we are actually interested in virtual displacements that are *single valued* on Σ, it will be convenient to consider the space $\boldsymbol{\Phi} \subset \mathbf{D}$, defined as the space of *pairs of identical (vector valued) functions*, one for each side of Σ:

$$\boldsymbol{\Phi} := \{\boldsymbol{\varphi} = (\boldsymbol{\varphi}^1, \boldsymbol{\varphi}^2) \in \mathbf{D} \text{ such that } \boldsymbol{\varphi}^1 = \boldsymbol{\varphi}^2\}. \tag{2.22}$$

In the sequel, as $\boldsymbol{\varphi}^1$ and $\boldsymbol{\varphi}^2$ are always equal, we shall often call them both $\boldsymbol{\varphi}$. This is an *abuse of notation*, but we hope that it will not cause confusion. More generally, when speaking of a pair $\boldsymbol{\chi} = (\boldsymbol{\chi}^1, \boldsymbol{\chi}^2) = ((\chi_1^1, \chi_2^1, \chi_3^1), (\chi_1^2, \chi_2^2, \chi_3^2))$ we say that $\boldsymbol{\chi}$ is *single valued* on Σ if $\boldsymbol{\chi}^1 = \boldsymbol{\chi}^2$. Otherwise we say that $\boldsymbol{\chi}$ is *double valued*. The space $\boldsymbol{\Phi}$ will inherit the norm of $\mathbf{D}$, that is

$$\|\varphi\|_{\boldsymbol{\Phi}}^2 := \sum_{k=1}^{2} \|\varphi\|_{\mathbf{D},k}^2. \tag{2.23}$$

A nice consequence of the above definition (2.21) is the following property. For every $\mathbf{w} \in \mathbf{V}$ and for every $k = 1, 2$ we have first, in an obvious way, that

$$\|\mathbf{w}^k\|_{\mathbf{D},k} \leq \|\mathbf{w}\|_{E,k}, \tag{2.24}$$

since in (2.19) we took the infimum. As a consequence of (2.24) and (2.21) we have then

$$\|\mathbf{w}\|_{\mathbf{D}} \leq \|\mathbf{w}\|_{E,b}. \tag{2.25}$$

Together with the spaces $\mathbf{D}$ and $\boldsymbol{\Phi}$ we introduce then two spaces of Lagrange multipliers $\mathbf{M}^1$ and $\mathbf{M}^2$. Again, in our particular case (since we have only two subdomains), the two spaces will be equal to each other (but this will not be the case for more than two subdomains, see e.g. [18] or [19]). Still it is convenient to use two different names for them, as we are going to discretize them in two different ways. We set, for k=1,2,

$$\mathbf{M}^k := \{ \text{ dual space of } \mathbf{D}^k \}. \tag{2.26}$$

We consider then the global space of Lagrange multipliers as

$$\mathbf{M} := \{\boldsymbol{\mu} = (\boldsymbol{\mu}^1, \boldsymbol{\mu}^2) \text{ with } \boldsymbol{\mu}^1 \in \mathbf{M}^1 \text{ and } \boldsymbol{\mu}^2 \in \mathbf{M}^2\}. \tag{2.27}$$

The norm in $\mathbf{M}$ will therefore be defined as

$$\|\boldsymbol{\mu}\|_{\mathbf{M}} := \sup_{\boldsymbol{\delta} \in \mathbf{D}} \sum_{k=1}^{2} \frac{\int_\Sigma \boldsymbol{\mu}^k \cdot \boldsymbol{\delta}^k \,\mathrm{d}S}{\|\boldsymbol{\delta}\|_{\mathbf{D}}}. \tag{2.28}$$

The space $\mathbf{M}$ could also be defined, in a rough way, as the space of vector valued pairs $\boldsymbol{\mu} = (\boldsymbol{\mu}^1, \boldsymbol{\mu}^2)$ such that the above quantity (2.28) is finite. From the physical point of view, the elements of $\mathbf{M}$ are *tensions* (force per unit surface). When defining the norm in $\mathbf{D}$ (see (2.19) and (2.21)) we associated, in a natural way, an energy to each virtual displacement on Σ. Here, the square of the $\mathbf{M}$-norm of a tension $\boldsymbol{\mu}$ could now be interpreted as *the maximum work that $\boldsymbol{\mu}$ can make on a virtual dispacement* $\mathbf{v}$ *having a unitary associated energy*. An easy consequence of the definition of the norm (2.28) is the following Cauchy-Schwarz-like inequality, valid for any $\boldsymbol{\mu} \in \mathbf{M}$ and $\boldsymbol{\chi} \in \mathbf{D}$ (and physically obvious):

$$\sum_{k=1}^{2} \int_\Sigma \boldsymbol{\mu}^k \cdot \boldsymbol{\chi}^k \,\mathrm{d}S \leq \|\boldsymbol{\mu}\|_{\mathbf{M}} \, \|\boldsymbol{\chi}\|_{\mathbf{D}}. \tag{2.29}$$

Indeed, we easily have

$$\sum_{k=1}^{2} \frac{\int_\Sigma \boldsymbol{\mu}^k \cdot \boldsymbol{\chi}^k \,\mathrm{d}S}{\|\boldsymbol{\chi}\|_{\boldsymbol{\chi}}} \leq \sup_{\boldsymbol{\delta} \in \mathbf{D}} \sum_{k=1}^{2} \frac{\int_\Sigma \boldsymbol{\mu}^k \cdot \boldsymbol{\delta}^k \,\mathrm{d}S}{\|\boldsymbol{\delta}\|_{\mathbf{D}}} \equiv \|\boldsymbol{\mu}\|_{\mathbf{M}}. \tag{2.30}$$

Using (2.29) and (2.25) we also obtain, for every $\boldsymbol{\mu} \in \mathbf{M}$ and for every $\mathbf{v} \in \mathbf{V}$,

$$\sum_{k=1}^{2} \int_{\Sigma} \boldsymbol{\mu}^k \cdot \mathbf{v}^k \,\mathrm{d}S \leq \|\boldsymbol{\mu}\|_{\mathbf{M}} \, \|\mathbf{v}\|_{E,b}. \tag{2.31}$$

We also point out that, since the two norms $\|\cdot\|_{\mathbf{D},1}$ and $\|\cdot\|_{\mathbf{D},2}$ are equivalent, if, by chance, $\boldsymbol{\mu}$ is single valued (that is $\boldsymbol{\mu}^1 = \boldsymbol{\mu}^2$), then its norm in $\mathbf{M}$ could be bounded, up to a constant, by taking the supremum only on $\boldsymbol{\Phi}$, instead of $\mathbf{D}$. This means that there exists a constant $\gamma > 0$ such that, *for all $\boldsymbol{\mu}$ such that $\boldsymbol{\mu}^1 = \boldsymbol{\mu}^2$*:

$$\gamma \|\boldsymbol{\mu}\|_{\mathbf{M}} \leq \sup_{\boldsymbol{\varphi} \in \boldsymbol{\Phi}} \sum_{k=1}^{2} \frac{\int_{\Sigma} \boldsymbol{\mu}^k \cdot \boldsymbol{\varphi}^k \,\mathrm{d}S}{\|\boldsymbol{\varphi}\|_{\boldsymbol{\Phi}}}. \tag{2.32}$$

We realize that both norms (2.23) and (2.28) are not very friendly for a certain number of readers, and we promise that we are going to make *a very moderate use* of them. It might however help the reader to know that *all the norms we are using in this paper are "energy norms"*, in the sense that their square has the physical dimensions of an *energy*.

With all the machinery ready to use Lagrange multipliers, we consider now the functional

$$\mathcal{L}(\mathbf{v}, \boldsymbol{\mu}, \boldsymbol{\varphi}) := \frac{1}{2}\|\mathbf{v}\|^2_{E,b} - \int_{\Omega} \mathbf{f} \cdot \mathbf{v} \,\mathrm{d}S + \int_{\Sigma} (\boldsymbol{\varphi} - \mathbf{v}^1) \cdot \boldsymbol{\mu}^1 \,\mathrm{d}S + \int_{\Sigma} (\boldsymbol{\varphi} - \mathbf{v}^2) \cdot \boldsymbol{\mu}^2 \,\mathrm{d}S, \tag{2.33}$$

and *we look for the stationarity point $(\mathbf{u}, \boldsymbol{\lambda}, \boldsymbol{\psi})$ of $\mathcal{L}$ when $(\mathbf{v}, \boldsymbol{\mu}, \boldsymbol{\varphi})$ varies over $\mathbf{V} \times \mathbf{M} \times \boldsymbol{\Phi}$*. It is not difficult to show that the following result holds true.

Theorem 2.1 *Assume that $(\mathbf{u}, \boldsymbol{\lambda}, \boldsymbol{\psi})$ is the stationarity point of $\mathcal{L}$ when $(\mathbf{v}, \boldsymbol{\mu}, \boldsymbol{\varphi})$ varies over $\mathbf{V} \times \mathbf{M} \times \boldsymbol{\Phi}$. Then we have*

$$\mathrm{div}\, \boldsymbol{\sigma}(\mathbf{u}) + \mathbf{f} = 0 \textit{ in each } \Omega^k, \qquad (k = 1, 2), \tag{2.34}$$

$$\boldsymbol{\sigma}(\mathbf{u}^k) \cdot \mathbf{n}^k = \boldsymbol{\lambda}^k \textit{ on } \Sigma, \quad \mathbf{n}^k = \text{outward unit normal to } \partial\Omega^k, \quad (k = 1, 2), \tag{2.35}$$

$$\mathbf{u}^1 = \mathbf{u}^2 = \boldsymbol{\psi} \textit{ on } \Sigma, \tag{2.36}$$

$$\boldsymbol{\lambda}^1 + \boldsymbol{\lambda}^2 = 0, \qquad \textit{on } \Sigma, \tag{2.37}$$

and therefore $\mathbf{u}$ coincides with the unique solution of the linear elasticity problem (2.9) *set directly on the whole domain Ω.*

P r o o f. Take first the derivative of $\mathcal{L}(\mathbf{v}, \boldsymbol{\mu}, \boldsymbol{\varphi})$ with respect to $\mathbf{v}$ at the point $(\mathbf{u}, \boldsymbol{\lambda}, \boldsymbol{\psi})$. We have

$$\int_{\Omega^k} \boldsymbol{\sigma}(\mathbf{u}^k) : \boldsymbol{\varepsilon}(\mathbf{v}^k) \,\mathrm{d}V - \int_{\Sigma} \boldsymbol{\lambda}^k \cdot \mathbf{v}^k \,\mathrm{d}S = \int_{\Omega^k} \mathbf{f} \cdot \mathbf{v}^k \,\mathrm{d}V \tag{2.38}$$

for all $\mathbf{v}^k$ having finite energy in Ω^k, and vanishing on $\partial\Omega^k \cap \partial\Omega$, $(k = 1, 2)$. Assuming $\mathbf{f}$ *and* $\boldsymbol{\lambda}^k$ as given, and considering (2.38) as a variational equation in the unknown $\mathbf{u}^k$, we easily obtain that its (unique) solution $\mathbf{u}^k$ satisfies conditions (2.34) and (2.35), for each k. Taking now the derivative of $\mathcal{L}(\mathbf{v}, \boldsymbol{\mu}, \boldsymbol{\varphi})$ with respect to $\boldsymbol{\mu}$ at the point $(\mathbf{u}, \boldsymbol{\lambda}, \boldsymbol{\psi})$ we have

$$\int_{\Sigma} (\boldsymbol{\psi} - \mathbf{u}^k) \cdot \boldsymbol{\mu}^k \,\mathrm{d}S = 0, \tag{2.39}$$

for all $\boldsymbol{\mu}^k$ $(k = 1, 2)$, that immediately gives (2.36). Taking finally the derivative of $\mathcal{L}(\mathbf{v}, \boldsymbol{\mu}, \boldsymbol{\varphi})$ with respect to $\boldsymbol{\varphi}$ at the point $(\mathbf{u}, \boldsymbol{\lambda}, \boldsymbol{\psi})$ we have

$$\int_\Sigma \boldsymbol{\varphi} \cdot \boldsymbol{\lambda}^1 \,\mathrm{d}S + \int_\Sigma \boldsymbol{\varphi} \cdot \boldsymbol{\lambda}^2 \,\mathrm{d}S = 0, \tag{2.40}$$

for all $\boldsymbol{\varphi}$, that immediately gives (2.37).

Finally, the vector-valued function defined as $\mathbf{u}^1$ in Ω^1 and $\mathbf{u}^2$ in Ω^2 satisfies the equilibrium conditions (2.34) in each Ω^k, is continuous on Σ (thanks to (2.36)), and its corresponding stress field has its normal component continuous on Σ, thanks to the joint use of (2.35) and (2.37). Hence it is the solution of the global problem.

□

3 Discretization of the three-field formulation

We can now tackle the problem of discretizing the three-field formulation described in the previous section. The first step should be to choose finite element subspaces $\mathbf{V}_h$, $\mathbf{M}_h$, and $\boldsymbol{\Phi}_h$, of $\mathbf{V}$, $\mathbf{M}$ and $\boldsymbol{\Phi}$, respectively.

Instead of actually making a precise choice for them, we are basically going to indicate general guidelines.

Choice of $\mathbf{V}_h^k$. We assume (as it was *the name of the game* from the very beginning) that we are *given* two independent meshes $\mathcal{K}_h^1$ and $\mathcal{K}_h^2$ in Ω^1 and in Ω^2, respectively. We play the game that we cannot (or we do not want to) touch them, and we leave them unchanged. Hence the first step will be to choose, in each Ω^k, a finite element space $\mathbf{V}_h^k$ on the mesh $\mathcal{K}_h^k$. To fix the ideas, let us take conforming finite elements of degree $r = r(k)$ (that might vary from one subdomain to another), verifying the homogeneous boundary conditions on $\partial\Omega^k \cap \partial\Omega$. No continuity will be required at the interface Σ. Once the $\mathbf{V}_h^k$'s have been chosen, we set, in a natural manner,

$$\mathbf{V}_h = \{\mathbf{v} = (\mathbf{v}^1, \mathbf{v}^2) \text{ such that } \mathbf{v}^k \in \mathbf{V}_h^k \ (k = 1, 2)\}. \tag{3.1}$$

The other two fields, however, are *our* job, and we have more freedom in their construction.

Choice of $\mathbf{M}_h^k$. A general, preferred choice is to take the mesh for $\mathbf{M}_h^k$ on Σ as *the restriction to Σ of the mesh $\mathcal{K}_h^k$* (for $k = 1, 2$). Then we choose finite element spaces $\mathbf{M}_h^k$ by taking, on these meshes, piecewise polynomials of degree $\ell = \ell(k)$. An important choice to be made is whether to use continuous or discontinuous finite elements. Although the choice of continuous finite element approximations for the Lagrange multipliers $\boldsymbol{\mu}^k$ was advocated, for instance, in [1], [2], we believe that *allowing discontinuities is a healthy choice,* at least when the interface Σ is not smooth (which means, in this context, that its normal unit vector $\mathbf{n}$ has jumps that do not tend to zero with the mesh size), as it will almost always be the case when dealing with more than two subdomains. Indeed, in view of (2.35), the solution $\boldsymbol{\lambda}^k$ has to jump whenever $\mathbf{n}$ jumps (as we expect each $\boldsymbol{\sigma}(\mathbf{u}^k)$ to be smooth). And it will be impossible, for the approximate solution $\boldsymbol{\lambda}_h^k$, to produce a good approximation of $\boldsymbol{\lambda}^k$ if we force continuity on it. Once the $\mathbf{M}_h^k$'s have been chosen, we set, in a natural manner,

$$\mathbf{M}_h = \{\boldsymbol{\mu} = (\boldsymbol{\mu}^1, \boldsymbol{\mu}^2) \text{ such that } \boldsymbol{\mu}^k \in \mathbf{M}_h^k \ (k = 1, 2)\}. \tag{3.2}$$

Choice of Φ_h. Here, finally, we are totally freed of the two decompositions $\mathcal{K}_h^k$ that have been given to us, and we can choose a *new decomposition* $\mathcal{S}_h$ on Σ. Here the preferred choice is *a uniform decomposition*, as much as the geometry of Σ allows an easy choice for it. For two-dimensional problems, where Σ is therefore one-dimensional, the uniform grid is an obvious winner ([1], [18]). For three-dimensional problems, the geometry of Σ can be much more complicated, and a uniform grid might even lack any sense (what is "a uniform decomposition of an ellipse"?) The main interest in having a uniform grid, whenever possible, is to allow an easy way to find the element of $\mathcal{S}_h$ that contains a given node of $\mathcal{K}_h^k$. A second important point, that will be also mentioned in the next section in connection with Domain Decomposition Methods, is the possibility of building powerful preconditioners on Σ. For both aspects, when a uniform grid could not be chosen, a structured hierarchical grid might be a reasonable substitute. One way or another, we shall *assume* here that you have a sort of favorite grid $\mathcal{S}_h$. Then we take as $\mathbf{\Phi}_h$ the space of piecewise polynomials of degree, say, p on the grid $\mathcal{S}_h$. Here, continuity is recommended: in view of (2.36), we expect the discrete solution $\boldsymbol{\psi}_h$ to approximate $\mathbf{u}$ on Σ. And it is quite reasonable to assume that $\mathbf{u}$ is smooth there. Sometimes, in particular for two-dimensional problems (where Σ is a line) the choice of a very coarse $\mathcal{S}_h$ with a rather high p has been advocated [1], [2]. The use of *wavelets* on Σ has also been proposed and analyzed in [12]. We finally point out that, from the theoretical point of view (that is, when proving theorems) we shall actually treat $\mathbf{\Phi}_h$ as a space of pairs of identical (vector valued) functions (in agreement with the definition (2.22) of the space $\mathbf{\Phi}$), but *in practice* one shall obviously consider it as *just one space*.

We point out that, at this point, we have *three* different grids defined on the interface Σ: one coming from the (given) decomposition $\mathcal{K}_h^1$, another coming from the (given) decomposition $\mathcal{K}_h^2$, and a third (of our choice) defined on Σ independently of the other two, and used to approximate the variable $\boldsymbol{\psi}$. The situation is illustrated in Figure 2 where, for didactic purposes, the two subdomains have been drawn far apart, and the interface Σ has been triplicated in order to show the three decompositions. Once the three discrete spaces $\mathbf{V}_h$, $\mathbf{M}_h$ and $\mathbf{\Phi}_h$ have been chosen, we can write the *discrete problem* as follows.

- *Find the stationarity point* $(\mathbf{u}_h, \boldsymbol{\lambda}_h, \boldsymbol{\psi}_h)$ *of the functional* $\mathcal{L}(\mathbf{v}, \boldsymbol{\mu}, \boldsymbol{\varphi})$ *given in* (2.33) *when* $\mathbf{v}$, $\boldsymbol{\mu}$ *and* $\boldsymbol{\varphi}$ *vary over* $\mathbf{V}_h$, $\mathbf{M}_h$, *and* $\mathbf{\Phi}_h$, *respectively*.

To see the discrete problem under a better light, it will be convenient to write the corresponding equations that come out when we impose that the derivatives of $\mathcal{L}(\mathbf{v}, \boldsymbol{\mu}, \boldsymbol{\varphi})$ vanish. Taking the derivatives with respect to $\mathbf{v}$ as in (2.38) we have now

$$\int_{\Omega^k} \boldsymbol{\sigma}(\mathbf{u}_h^k) : \boldsymbol{\varepsilon}(\mathbf{v}^k)\,\mathrm{d}V - \int_{\Sigma} \boldsymbol{\lambda}_h^k \cdot \mathbf{v}^k\,\mathrm{d}S = \int_{\Omega^k} \mathbf{f} \cdot \mathbf{v}^k\,\mathrm{d}V \quad \forall \mathbf{v}^k \in \mathbf{V}_h^k \quad (k=1,2). \tag{3.3}$$

Taking now the derivative with respect to $\boldsymbol{\mu}$ as in (2.39) we have

$$\int_{\Sigma} (\boldsymbol{\psi}_h - \mathbf{u}_h^k) \cdot \boldsymbol{\mu}^k\,\mathrm{d}S = 0, \quad \forall \boldsymbol{\mu}^k \in \mathbf{M}_h^k \quad (k=1,2). \tag{3.4}$$

Taking finally the derivative with respect to $\boldsymbol{\varphi}$ as in (2.40) we have

$$\int_{\Sigma} \boldsymbol{\varphi} \cdot \boldsymbol{\lambda}_h^1\,\mathrm{d}S + \int_{\Sigma} \boldsymbol{\varphi} \cdot \boldsymbol{\lambda}_h^2\,\mathrm{d}S = 0, \quad \forall \boldsymbol{\varphi} \in \mathbf{\Phi}_h. \tag{3.5}$$

4 The three-field decomposition as a DDM

In this section we shall give some hints on the possible use of the three-field formulation in the framework of DDM (Domain Decomposition Methods), and the use of parallel computers. It is clear that the particular case presented here, using only *two* subdomains, will not, as such, be very appealing for using a parallel computer. However it will be clear that what we say will hold for a decomposition in an arbitrary finite number of subdomains, and equally clear will come out the interest of this approach for parallel computations.

The first important point is that the above equations (3.3)-(3.5) can be grouped in a different way. In particular, we might think that for any given (tentative choice) of $\boldsymbol{\psi}_h$ we solve, *independently in each subdomain*, the problems (corresponding to (3.3)-(3.4)):

$$\begin{cases} \text{find } (\mathbf{u}_h^k, \boldsymbol{\lambda}_h^k) \in \mathbf{V}_h^k \times \mathbf{M}_h^k \text{ such that} \\ \int_{\Omega^k} \boldsymbol{\sigma}(\mathbf{u}_h^k) : \boldsymbol{\varepsilon}(\mathbf{v}^k)\,\mathrm{d}V - \int_\Sigma \boldsymbol{\lambda}_h^k \cdot \mathbf{v}^k\,\mathrm{d}S = \int_{\Omega^k} \mathbf{f}\cdot\mathbf{v}^k\,\mathrm{d}V \quad \forall \mathbf{v}^k \in \mathbf{V}_h^k \\ \int_\Sigma \mathbf{u}_h^k \cdot \boldsymbol{\mu}^k\,\mathrm{d}S = \int_\Sigma \boldsymbol{\psi}_h \cdot \boldsymbol{\mu}^k\,\mathrm{d}S \quad \forall \boldsymbol{\mu}^k \in \mathbf{M}_h^k, \end{cases} \tag{4.1}$$

and that to check whether the choice of $\boldsymbol{\psi}_h$ is correct we use equation (3.5). We repeat now, in more details, the same concept from a different point of view (and with different notation). Let us denote by A^k the matrix associated with the bilinear form on $\mathbf{V}_h^k \times \mathbf{V}_h^k$

$$a^k(\mathbf{w}^k, \mathbf{v}^k) := \int_{\Omega^k} \boldsymbol{\sigma}(\mathbf{w}^k) : \boldsymbol{\varepsilon}(\mathbf{v}^k)\,\mathrm{d}V, \tag{4.2}$$

then denote by B^k the matrix associated with the bilinear form on $\mathbf{V}_h^k \times \mathbf{M}_h^k$

$$b^k(\mathbf{v}^k, \boldsymbol{\mu}^k) := -\int_\Sigma \mathbf{v}^k \cdot \boldsymbol{\mu}^k\,\mathrm{d}S, \tag{4.3}$$

and finally denote by C^k the matrix associated with the bilinear form on $\mathbf{M}_h^k \times \boldsymbol{\Phi}_h$

$$c^k(\boldsymbol{\mu}^k, \boldsymbol{\varphi}) = \int_\Sigma \boldsymbol{\varphi} \cdot \boldsymbol{\mu}^k\,\mathrm{d}S. \tag{4.4}$$

Equations (3.3)-(3.5) can now be written in matrix form (with rather obvious meaning of the notation):

$$\begin{bmatrix} A^1 & 0 & (B^1)^T & 0 & 0 \\ 0 & A^2 & 0 & (B^2)^T & 0 \\ B^1 & 0 & 0 & 0 & (C^1)^T \\ 0 & B^2 & 0 & 0 & (C^2)^T \\ 0 & 0 & C^1 & C^2 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{u}^1 \\ \mathbf{u}^2 \\ \boldsymbol{\lambda}^1 \\ \boldsymbol{\lambda}^2 \\ \boldsymbol{\psi} \end{bmatrix} = \begin{bmatrix} \mathbf{f}^1 \\ \mathbf{f}^2 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{4.5}$$

Changing the order of the unknowns and of the equations (4.5) becomes

$$\begin{bmatrix} A^1 & (B^1)^T & 0 & 0 & 0 \\ B^1 & 0 & 0 & 0 & (C^1)^T \\ 0 & 0 & A^2 & (B^2)^T & 0 \\ 0 & 0 & B^2 & 0 & (C^2)^T \\ 0 & C^1 & 0 & C^2 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{u}^1 \\ \boldsymbol{\lambda}^1 \\ \mathbf{u}^2 \\ \boldsymbol{\lambda}^2 \\ \boldsymbol{\psi} \end{bmatrix} = \begin{bmatrix} \mathbf{f}^1 \\ 0 \\ \mathbf{f}^2 \\ 0 \\ 0 \end{bmatrix} \tag{4.6}$$

Setting now

$$\mathcal{A} = \begin{bmatrix} A^1 & (B^1)^T & 0 & 0 \\ B^1 & 0 & 0 & 0 \\ 0 & 0 & A^2 & (B^2)^T \\ 0 & 0 & B^2 & 0 \end{bmatrix} \qquad \mathbf{U} = \begin{bmatrix} \mathbf{u}^1 \\ \boldsymbol{\lambda}^1 \\ \mathbf{u}^2 \\ \boldsymbol{\lambda}^2 \end{bmatrix} \qquad \mathbf{F} = \begin{bmatrix} \mathbf{f}^1 \\ 0 \\ \mathbf{f}^2 \\ 0 \end{bmatrix} \tag{4.7}$$

and

$$\mathcal{C} = \begin{bmatrix} 0 & C^1 & 0 & C^2 \end{bmatrix}, \tag{4.8}$$

the system can be written as

$$\begin{bmatrix} \mathcal{A} & \mathcal{C}^T \\ \mathcal{C} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{U} \\ \boldsymbol{\psi} \end{bmatrix} = \begin{bmatrix} \mathbf{F} \\ 0 \end{bmatrix}. \tag{4.9}$$

What is important in (4.9) is that the matrix $\mathcal{A}$ is *block diagonal*. And this will be the case even for a decomposition into an arbitrary finite number of subdomains. Hence if we eliminate $\mathbf{U}$, and write the system as

$$\mathcal{C}\mathcal{A}^{-1}\mathcal{C}^T\boldsymbol{\psi} = \mathcal{C}\mathcal{A}^{-1}\mathbf{F}, \tag{4.10}$$

and solve it by, say, preconditioned conjugate gradient, the solution of the system $\mathcal{A}\mathbf{U}^{n+1} = \mathbf{G}^n$ (to be computed at each step) can be performed working *in parallel* (each processor being dedicated to a subdomain).

It might be interesting to point out that, from the mechanical point of view, the solution of a problem of the type

$$\begin{bmatrix} A^1 & (B^1)^T \\ B^1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{u}^1 \\ \boldsymbol{\lambda}^1 \end{bmatrix} = \begin{bmatrix} \mathbf{f}^1 \\ -(C^1)^T\boldsymbol{\psi} \end{bmatrix} \tag{4.11}$$

(that is one of the problems that we have to solve, in parallel, at each step of our iterative procedure), corresponds to solving a problem in Ω^1 with distributed load $\mathbf{f}^1$ and prescribed kinematic boundary conditions $\mathbf{u}^1 = \boldsymbol{\psi}$ on Σ and $\mathbf{u}^1 = 0$ on the rest of the boundary (that is $\partial\Omega^1 \setminus \Sigma$). In particular, we solve the problem in the formulation *with Lagrange multipliers*, in the style of [5].

It is also relevant to point out that the choice of a uniform grid (or at least a nested grid) on Σ is not just a commodity, but a potential powerful instrument to construct suitable *preconditioners* for problem (4.10). We have no time here to discuss the matter, in particular since the literature on preconditioners for Domain Decomposition Methods is impressive. We just refer for instance to [11], [13], [15], [22], [26], [27], [28], [30], [31], [33], [35], [37], [39], [40].

5 Stability conditions and error estimates

We discuss here the conditions on the choices of finite element spaces that will ensure *stability* and *optimal error estimates* for the discrete problem (3.3)-(3.5). For this, we first need to choose suitable norms for all the spaces at play. As *there are many ways to skin a cat* we can do this as well in many different ways. Here, we are going to see two of them. The first choice will make use of the (difficult) norms (2.23) and (2.28), but the error bounds will come out rather easily. However, checking the *inf-sup* conditions in each particular case using these difficult norms would be much more complicated. Since, in the sequel, we are going to suggest a ”Check yourself the inf-sup conditions” strategy, we would like, instead, to use *easier norms* (no matter how difficult the proof of error estimates might become).

In this section, we deal with the “easy” error estimates with the difficult norms. To make life easier to the reader, we start from the study of the *stability* of our discrete problem. This means that we would like to prove that there exists a constant C, independent of the mesh sizes, such that: for every given $\mathbf{f}$, every possible solution $(\mathbf{u}_h, \boldsymbol{\lambda}_h, \boldsymbol{\psi}_h)$ of the discrete problem (3.3)-(3.5) (that we repeat here for convenience of the reader)

$$\int_{\Omega^k} \boldsymbol{\sigma}(\mathbf{u}_h^k) : \boldsymbol{\varepsilon}(\mathbf{v}^k)\,\mathrm{d}V - \int_{\Sigma} \boldsymbol{\lambda}_h^k \cdot \mathbf{v}^k\,\mathrm{d}S = \int_{\Omega^k} \mathbf{f} \cdot \mathbf{v}^k\,\mathrm{d}V \quad \forall \mathbf{v}^k \in \mathbf{V}_h^k \quad (k = 1, 2), \tag{5.1}$$

$$\int_{\Sigma} (\boldsymbol{\psi}_h - \mathbf{u}_h^k) \cdot \boldsymbol{\mu}^k\,\mathrm{d}S = 0 \quad \forall \boldsymbol{\mu}^k \in \mathbf{M}_h^k \quad (k = 1, 2), \tag{5.2}$$

$$\int_{\Sigma} \boldsymbol{\varphi} \cdot \boldsymbol{\lambda}_h^1\,\mathrm{d}S + \int_{\Sigma} \boldsymbol{\varphi} \cdot \boldsymbol{\lambda}_h^2\,\mathrm{d}S = 0 \quad \forall \boldsymbol{\varphi} \in \boldsymbol{\Phi}_h, \tag{5.3}$$

will satisfy the *a priori estimate*

$$\|\mathbf{u}_h\|_{E,b} + \|\boldsymbol{\lambda}_h\|_{\mathbf{M}} + \|\boldsymbol{\psi}_h\|_{\boldsymbol{\Phi}} \le C\,\|\mathbf{f}\|_*, \tag{5.4}$$

where $\|\mathbf{f}\|_*$ is the *load norm* (defined as *the dual of the energy norm*), that is

$$\|\mathbf{f}\|_* := \sup_{\mathbf{v} \in \mathbf{V}} \frac{\int_\Omega \mathbf{v} \cdot \mathbf{f}\,\mathrm{d}V}{\|\mathbf{v}\|_{E,b}}. \tag{5.5}$$

The property (5.4) will follow immediately (well, almost) from the two following assumptions on the grid.

Assumption 1 (control $\boldsymbol{\mu}$ by the $\mathbf{v}$'s). *There exists a constant $\beta_M > 0$ such that: for every $\boldsymbol{\mu} \in \mathbf{M}_h$ we can find a $\mathbf{v}_\mu \in \mathbf{V}_h$ such that*

$$\|\mathbf{v}_\mu\|_{E,b} = 1 \qquad \text{and} \qquad \|\boldsymbol{\mu}\|_{\mathbf{M}} \le \frac{1}{\beta_M} \sum_{k=1}^{2} \int_\Sigma \mathbf{v}_\mu^k \cdot \boldsymbol{\mu}^k\,\mathrm{d}S. \tag{5.6}$$

Assumption 2 (control $\boldsymbol{\varphi}$ by the $\boldsymbol{\mu}$'s). *There exists a constant $\beta_\Phi > 0$ such that: for every $\boldsymbol{\varphi} \in \boldsymbol{\Phi}_h$ we can find a $\boldsymbol{\mu}_\varphi \in \mathbf{M}_h$ such that*

$$\|\boldsymbol{\mu}_\varphi\|_{\mathbf{M}} = 1 \qquad \text{and} \qquad \|\boldsymbol{\varphi}\|_{\boldsymbol{\Phi}} \le \frac{1}{\beta_\Phi} \sum_{k=1}^{2} \int_\Sigma \boldsymbol{\mu}_\varphi^k \cdot \boldsymbol{\varphi}\,\mathrm{d}S. \tag{5.7}$$

Assumptions (5.6) and (5.7) are just *inf-sup conditions in disguise*. We point out that, using (5.6) and (2.31) we easily have $\beta_M \le 1$, while using (5.7) and (2.29) we have $\beta_\Phi \le 1$.

Let us see how the *inf-sup* conditions can give us the desired stability property (5.4).

Take first $\mathbf{v} = \mathbf{u}_h$ in (5.1), $\boldsymbol{\mu} = -\boldsymbol{\lambda}_h$ in (5.2), and $\boldsymbol{\varphi} = \boldsymbol{\psi}_h$ in (5.3). Summing the three equations, and then using the definition (5.5) of load norm we get

$$\|\mathbf{u}_h\|^2_{E,b} = \int_\Omega \mathbf{f} \cdot \mathbf{u}_h \,\mathrm{d}V \leq \|\mathbf{f}\|_* \,\|\mathbf{u}_h\|_{E,b}, \tag{5.8}$$

that gives immediately

$$\|\mathbf{u}_h\|_{E,b} \leq \|\mathbf{f}\|_*. \tag{5.9}$$

To estimate $\boldsymbol{\lambda}_h$ consider the $\mathbf{v} = \mathbf{v}_\lambda$ that we get from Assumption 1 when $\boldsymbol{\mu} = \boldsymbol{\lambda}_h$. Using (5.6), then (5.1) with $\mathbf{v} = \mathbf{v}_\lambda$, then the Cauchy-Schwarz inequality (2.16) and the definition (5.5) of the load norm, and finally again (5.6) (to see that the norm of $\mathbf{v}_\lambda$ is 1), we have

$$\begin{aligned}\beta_M \,\|\boldsymbol{\lambda}_h\|_{\mathbf{M}} \leq \sum_{k=1}^{2} \int_\Sigma \mathbf{v}_\lambda^k \cdot \boldsymbol{\lambda}_h^k \,\mathrm{d}S = \sum_{k=1}^{2} \int_{\Omega^k} \boldsymbol{\sigma}(\mathbf{u}_h^k) : \boldsymbol{\varepsilon}(\mathbf{v}_\lambda^k) \,\mathrm{d}V - \int_\Omega \mathbf{f} \cdot \mathbf{v}_\lambda \,\mathrm{d}V \\ \leq \|\mathbf{u}_h\|_{E,b} \,\|\mathbf{v}_\lambda\|_{E,b} + \|\mathbf{f}\|_* \,\|\mathbf{v}_\lambda\|_{E,b} \leq \|\mathbf{u}_h\|_{E,b} + \|\mathbf{f}\|_*.\end{aligned} \tag{5.10}$$

In a similar way we can derive the estimate for $\boldsymbol{\psi}_h$. Indeed consider the $\boldsymbol{\mu} = \boldsymbol{\mu}_\psi$ that we get from Assumption 2 when $\boldsymbol{\varphi} = \boldsymbol{\psi}_h$. Using (5.7), then (5.2), then the Cauchy-Schwarz-like inequality (2.31), and finally again (5.7) (to see that the norm of $\boldsymbol{\mu}_\psi$ is 1), we have

$$\begin{aligned}\beta_\Phi \|\boldsymbol{\psi}_h\|_{\boldsymbol{\Phi}} \leq \sum_{k=1}^{2} \int \boldsymbol{\mu}_\psi^k \cdot \boldsymbol{\psi}_h \,\mathrm{d}S = \sum_{k=1}^{2} \int \boldsymbol{\mu}_\psi^k \cdot \mathbf{u}_h^k \,\mathrm{d}S \\ \leq \|\boldsymbol{\mu}_\psi\|_{\mathbf{M}} \,\|\mathbf{u}_h\|_{E,b} = \|\mathbf{u}_h\|_{E,b}.\end{aligned} \tag{5.11}$$

Collecting (5.9), (5.10), and (5.11) we have then

$$\|\mathbf{u}_h\|_{E,b} + \|\boldsymbol{\lambda}_h\|_{\mathbf{M}} + \|\boldsymbol{\psi}_h\|_{\boldsymbol{\Phi}} \leq \big(1 + \frac{2}{\beta_M} + \frac{1}{\beta_\Phi}\big) \,\|\mathbf{f}\|_*. \tag{5.12}$$

We have therefore the following theorem.

Theorem 5.1 *Let the spaces $\mathbf{V}_h$, $\mathbf{M}_h$ and $\boldsymbol{\Phi}_h$ satisfy Assumptions 1 and 2 (namely* (5.6) *and* (5.7)*). Then for every given load* $\mathbf{f}$ *problem* (5.1)-(5.3) *has a unique solution* $(\mathbf{u}_h, \boldsymbol{\lambda}_h, \boldsymbol{\psi}_h)$ *in $\mathbf{V}_h \times \mathbf{M}_h \times \boldsymbol{\Phi}_h$. Moreover, such solution satisfies the stability bound* (5.4) *with*

$$C = 1 + \frac{2}{\beta_M} + \frac{1}{\beta_\Phi}. \tag{5.13}$$

P r o o f. The problem (5.1)-(5.3) has as many equations as unknowns. If we take $\mathbf{f} = 0$ in (5.12) we see that the homogeneous system has only the trivial solution. This implies that the determinant of the corresponding matrix is different from 0. Hence for every given right-hand side $\mathbf{f}$ the problem has a unique solution. The stability bound (5.4) (with the prescribed value for the constant C) follows again from (5.12). □

Under Assumptions 1 and 2 we can also prove *optimal a priori error estimates*, as shown, with classical arguments (see [17]), in the following theorem, whose detailed proof will be reported in Appendix A.

Theorem 5.2 *Under the same assumptions of Theorem 5.1, if* $(\mathbf{u}, \boldsymbol{\lambda}, \boldsymbol{\psi})$ *and* $(\mathbf{u}_h, \boldsymbol{\lambda}_h, \boldsymbol{\psi}_h)$ *are the solutions of the elasticity problem* (2.9) *(as given in Theorem 2.1) and of the discretized problem* (5.1)-(5.3)*, respectively, then we have*

$$\begin{aligned}\|\mathbf{u}-\mathbf{u}_h\|_{E,b} + \beta_M\|\boldsymbol{\lambda}-\boldsymbol{\lambda}_h\|_{\mathbf{M}} + \beta_\Phi\|\boldsymbol{\psi}-\boldsymbol{\psi}_h\|_{\boldsymbol{\Phi}} \\ \leq C\left(\inf_{\mathbf{v}\in\mathbf{V}_h}\|\mathbf{u}-\mathbf{v}\|_{E,b} + \inf_{\boldsymbol{\mu}\in\mathbf{M}_h}\|\boldsymbol{\lambda}-\boldsymbol{\mu}\|_{\mathbf{M}} + \inf_{\boldsymbol{\varphi}\in\boldsymbol{\Phi}_h}\|\boldsymbol{\psi}-\boldsymbol{\varphi}\|_{\boldsymbol{\Phi}}\right),\end{aligned} \quad (5.14)$$

where C depends only on the constants β_M and β_Φ appearing in (5.6) *and* (5.7)*, respectively, and is bounded by* $1/\beta_M + 1/\beta_\Phi$.

We are not going to discuss the conditions that need to be imposed on the meshes and on the type of finite elements in order to ensure that Assumptions 1 and 2 hold true. Roughly speaking, in order to ensure Assumption 2 we just have to choose a mesh $\mathcal{S}_h$ on Σ in such a way that it is *coarser* than at least one of the two meshes $\mathcal{K}^1_{h|\Sigma}$ and $\mathcal{K}^2_{h|\Sigma}$. Actually it has to be coarser, somehow, than "the union" of the two meshes (provided that we can give a meaning to the union of two meshes). One can also obviously lower the degree of the local polynomials instead of coarsening the mesh. However, the use of a high degree p (for the elements of $\boldsymbol{\Phi}_h$) together with a very coarse mesh was advocated for instance in [1], [2]. On the other hand, Assumption 1 would require, roughly speaking, that the degrees of freedom for $\mathbf{V}^k_h$ on Σ are more than the degrees of freedom for $\mathbf{M}^k_h$ (for both $k=1$ and $k=2$). This can be obtained by taking for $\mathbf{M}^k_h$ polynomials of one degree lower than that used for $\mathbf{V}^k_h$ (in other words, $\ell(k) = r(k)-1$), as proposed in [1] and [2]. Often it would be sufficient to lower the degree ℓ *only for the elements near* $\partial\Sigma$ as it is done (roughly speaking) in the *mortar method* (see [8], [9], [33], [39]). Another possibility (advocated for instance in [16], [20], and [19]) is to stabilize the problem by augmenting the spaces $\mathbf{V}^k_h$, at the boundary, by means of suitable *boundary bubbles*. Or one can add suitable *stabilizing terms* at the boundary, as in [7], [6], [10], [34].

6 More difficult error estimates using easier norms

We have seen that Assumptions 1 and 2 ensure, with a rather easy proof, optimal error bounds. Proving that these assumptions actually hold true for a given choice of the finite element spaces can be easy or difficult, according with the different particular case, *if* you are a specialist. If you are not, however, you would like to have some practical instrument to check whether a certain choice (that you are willing to use) has reasonable chances or not.

Following [17], and in the spirit of [23], we would like to present in the next section a simple way to check, in each particular case, whether your choice of finite element spaces is reliable or not. This cannot however be easily done with our original choice for the norms in $\boldsymbol{\Phi}$ and in $\mathbf{M}$ (see e.g. [29]). We introduce therefore some different norms to be used instead of the original ones. In particular, all the new norms will be defined by simple *integrals*, and, as announced, they will all be *energy norms*. For the spaces $\mathbf{M}^k_h$ we choose

$$\|\boldsymbol{\mu}^k\|^2_{\mathbf{M}^k_h} := \sum_{K\in\mathcal{K}^k_{h|\Sigma}} \int_K \frac{h_K}{L_\mu}|\boldsymbol{\mu}^k|^2\,\mathrm{d}S, \qquad (k=1,2), \quad (6.1)$$

where, for all K that is a face, on Σ, of an element belonging to $\mathcal{K}_h^k$, we denoted by h_K its diameter. From (6.1) we construct the norm for the space $\mathbf{M}_h$ in a natural way:

$$\|\boldsymbol{\mu}\|^2_{\mathbf{M}_h} := \|\boldsymbol{\mu}^1\|^2_{\mathbf{M}_h^1} + \|\boldsymbol{\mu}^2\|^2_{\mathbf{M}_h^2}. \tag{6.2}$$

It will also be handy to define norms on the space $\mathbf{D}$ of *traces on* Σ *of the functions of* $\mathbf{V}$, as defined in (2.18) and (2.20). We set

$$\|\mathbf{v}^k\|^2_{\mathbf{D}_h^k} := \sum_{K\in\mathcal{K}^k_{h|\Sigma}} \int_K \frac{L_\mu}{h_K}|\mathbf{v}^k|^2 \,\mathrm{d}S \qquad (k=1,2), \tag{6.3}$$

and in a way similar to (6.2)

$$\|\mathbf{v}\|^2_{\mathbf{D}_h} := \|\mathbf{v}^1\|^2_{\mathbf{D}_h^1} + \|\mathbf{v}^2\|^2_{\mathbf{D}_h^2}. \tag{6.4}$$

It will be convenient, in this section, to denote by $(\ ,\ )_{0,\Sigma}$ the usual L^2-inner product on Σ, that is

$$(\mathbf{v},\boldsymbol{\mu})_{0,\Sigma} := \sum_{k=1}^2 \int_\Sigma \mathbf{v}^k\cdot\boldsymbol{\mu}^k \,\mathrm{d}V. \tag{6.5}$$

We note that for the two norms $\mathbf{M}_h$ and $\mathbf{D}_h$ we have the Cauchy-Schwarz-like inequality:

$$(\mathbf{v},\boldsymbol{\mu})_{0,\Sigma} := \sum_{k=1}^2 \int_\Sigma \mathbf{v}^k\cdot\boldsymbol{\mu}^k \,\mathrm{d}V \le \|\mathbf{v}\|_{\mathbf{D}_h}\,\|\boldsymbol{\mu}\|_{\mathbf{M}_h}. \tag{6.6}$$

Both the norms (6.2) and (6.4) can be easily seen to be *energy norms*. In order to introduce norms on $\mathbf{D}$ and $\mathbf{M}$ that mimic the $L^2(\Sigma)$-norm, but are also energy norms, we define d_Ω to be a typical length of the problem, as for instance the diameter of Ω, and we set

$$\kappa := \frac{L_\mu}{d_\Omega} \tag{6.7}$$

$$\|\mathbf{v}\|^2_{0,D} := \sum_{k=1}^2 \int_\Sigma \kappa|\mathbf{v}^k|^2 \,\mathrm{d}S, \qquad \|\boldsymbol{\mu}\|^2_{0,M} := \sum_{k=1}^2 \int_\Sigma \kappa^{-1}|\boldsymbol{\mu}^k|^2 \,\mathrm{d}S. \tag{6.8}$$

Note that $(\ ,\ )_{0,D}$ (scalar product associated with $\|\cdot\|_{0,D}$) is a scalar product *for displacements*, while $(\ ,\ )_{0,M}$ (scalar product associated with $\|\cdot\|_{0,M}$) is a scalar product *for stresses*. Moreover, $(\ ,\ )_{0,\Sigma}$ *couples displacements and stresses*. In general, recalling (6.7) we have

$$(\boldsymbol{\mu},\mathbf{v})_{0,\Sigma} = (\boldsymbol{\mu},\kappa\mathbf{v})_{0,M} = (\kappa^{-1}\boldsymbol{\mu},\mathbf{v})_{0,D}. \tag{6.9}$$

We point out that here too we have the Cauchy-Schwarz inequalities

$$(\kappa\mathbf{v},\boldsymbol{\mu})_{0,M} \le \|\kappa\mathbf{v}\|_{0,M}\,\|\boldsymbol{\mu}\|_{0,M} \quad\text{and}\quad (\mathbf{v},\boldsymbol{\mu})_{0,\Sigma} \le \|\mathbf{v}\|_{0,D}\,\|\boldsymbol{\mu}\|_{0,M} \tag{6.10}$$

It will also be convenient to denote by $h_{\mathcal{S}}$ and $h_{\mathcal{K}}$ the *numbers* obtained taking the ratio between the maximum diameters of the elements in $\mathcal{S}_h$ and in $\mathcal{K}_{h|\Sigma}$, respectively, and the characteristic length d_Ω:

$$h_{\mathcal{K}} := \max_{K \in \mathcal{K}_{h|\Sigma}} \frac{h_K}{d_\Omega} \qquad h_{\mathcal{S}} := \max_{K \in \mathcal{S}_h} \frac{h_K}{d_\Omega}. \tag{6.11}$$

We are finally ready to present the modified version of Assumptions 1 and 2.

Assumption 1h (control $\boldsymbol{\mu}$ by the $\mathbf{v}$'s). *There exists a constant $\beta^*_M > 0$ such that: for every $\boldsymbol{\mu} \in \mathbf{M}_h$ we can find a $\mathbf{v}_\mu \in \mathbf{V}_h$ such that*

$$\|\mathbf{v}_\mu\|_{\mathbf{D}_h} = 1 \qquad \text{and} \qquad \|\boldsymbol{\mu}\|_{\mathbf{M}_h} \le \frac{1}{\beta^*_M}(\mathbf{v}_\mu, \boldsymbol{\mu})_{0,\Sigma}. \tag{6.12}$$

Assumption 2h (control $\boldsymbol{\varphi}$ by the $\boldsymbol{\mu}$'s). *There exists a constant $\beta^*_\Phi > 0$ such that: for every $\boldsymbol{\varphi} \in \boldsymbol{\Phi}_h$ we can find a $\boldsymbol{\mu}_\varphi \in \mathbf{M}_h$ such that*

$$\|\boldsymbol{\mu}_\varphi\|_{0,M} = 1 \qquad \text{and} \qquad \|\boldsymbol{\varphi}\|_{0,\mathbf{D}} \le \frac{1}{\beta^*_\Phi}(\boldsymbol{\mu}_\varphi, \boldsymbol{\varphi})_{0,\Sigma}. \tag{6.13}$$

We notice that using (6.12) and (6.6) we easily have $\beta^*_M \le 1$, while using (6.13) and (6.10) we have $\beta^*_\Phi \le 1$.

We have the following theorem.

Theorem 6.1 *Assume that our choices of finite element spaces for $\mathbf{V}_h$, $\mathbf{M}_h$ and $\boldsymbol{\Phi}_h$ satisfy Assumptions $1h$ and $2h$. If $(\mathbf{u}, \boldsymbol{\lambda}, \boldsymbol{\psi})$ and $(\mathbf{u}_h, \boldsymbol{\lambda}_h, \boldsymbol{\psi}_h)$ are the solutions of the elasticity problem* (2.9) *(as given in Theorem 2.1) and of the discretized problem* (5.1)-(5.3)*, respectively, then we have the following error estimate:*

$$\begin{aligned} &\|\mathbf{u} - \mathbf{u}_h\|_{E,b} + \beta^*_M \|\boldsymbol{\lambda} - \boldsymbol{\lambda}_h\|_{\mathbf{M}} \\ \le\ & C_B \Big(\inf_{\mathbf{v} \in \mathbf{V}_h} \|\mathbf{u} - \mathbf{v}\|_{E,b} + (h_{\mathcal{K}}^{1/2} + h_{\mathcal{S}}^{1/2}) \inf_{\boldsymbol{\mu} \in \mathbf{M}_h} \|\boldsymbol{\lambda} - \boldsymbol{\mu}\|_{0,M} + \inf_{\boldsymbol{\varphi} \in \boldsymbol{\Phi}_h} \|\boldsymbol{\psi} - \boldsymbol{\varphi}\|_{\boldsymbol{\Phi}} \Big), \end{aligned} \tag{6.14}$$

where C_B has the form

$$C_B = c_B \left(\frac{1}{\beta^*_M} + \frac{1}{(\beta^*_\Phi)^2} \right) \tag{6.15}$$

and c_B depends only on the shape of the elements in $\mathcal{S}_h$ and of those elements in $\mathcal{K}_h$ that have at least a vertex on Σ.

The proof of the theorem is rather technical, and could be omitted by the readers without at least a certain amount of Mathematical curiosity. We decided therefore to present it in Appendix B.

7 Check yourself your own *inf-sup*

Our aim here is to discuss a test that everybody can perform in order to check whether Assumptions $1h$ and $2h$ are satisfied or not. We shall discuss in more detail Assumption $1h$, (in particular for $k = 1$), as the arguments for Assumption $1h$ for $k = 2$, as well as (*mutatis mutandis*) for Assumption $2h$, will be identical.

As we are going to use arguments from linear algebra, it will be convenient to pass from our functional spaces to $\mathbb{R}^N$. We choose therefore a basis in each of the spaces $\mathbf{M}_h^1$, $\mathbf{M}_h^2$, $\mathbf{\Phi}_h$, as well as for the restrictions of $\mathbf{V}_h^1$ and $\mathbf{V}_h^2$ to Σ. For discussing Assumption $1h$ with $k = 1$ only the spaces $\mathbf{M}_h^1$ and the restriction of $\mathbf{V}_h^1$ to Σ will enter the game. Let therefore $\{\boldsymbol{\mu}^{(1)}, \boldsymbol{\mu}^{(2)}, \boldsymbol{\mu}^{(3)}, ..., \boldsymbol{\mu}^{(NM1)}\}$ be a basis for the former, and $\{\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \mathbf{v}^{(3)}, ..., \mathbf{v}^{(NV1)}\}$ be a basis for the latter. The numbers $NM1$ and $NV1$ are obviously the respective dimensions of these spaces. Then we construct a one-to-one mapping from $\mathbf{M}_h^1$ to $\mathbb{R}^{NM1}$ defined by

$$\underline{\mathbf{m}} \equiv (m_1, m_2, ...m_{NM1}) \leftrightarrow \sum_{i=1}^{NM1} m_i \boldsymbol{\mu}^{(i)}. \tag{7.1}$$

Similarly we construct a one-to-one mapping from "the restrictions to Σ of $\mathbf{V}_h^1$" to $\mathbb{R}^{NV1}$ defined by

$$\underline{\mathbf{q}} \equiv (q_1, q_2, ...q_{NV1}) \leftrightarrow \sum_{j=1}^{NV1} q_j \mathbf{v}^{(j)}. \tag{7.2}$$

We consider then the $NM1 \times NV1$ matrix B defined by

$$B_{i,j} := \int_\Sigma \boldsymbol{\mu}^{(i)} \cdot \mathbf{v}^{(j)} \, \mathrm{d}S \qquad (i = 1, .., NM1;\ j = 1, .., NV1), \tag{7.3}$$

together with the two matrices that define the norm $\| \cdot \|_{\mathbf{D}_h,1}$ (see (6.3)) and the norm $\| \cdot \|_{\mathbf{M}_h^1}$ (see (6.1)): the $NM1 \times NM1$ matrix R, defined by

$$R_{i,r} := \sum_{T \in \mathcal{K}^1_{h|\Sigma}} \int_\Sigma \frac{h_T}{L_\mu} \boldsymbol{\mu}^{(i)} \cdot \boldsymbol{\mu}^{(r)} \, \mathrm{d}S \qquad (i, r = 1, .., NM1), \tag{7.4}$$

and the $NV1 \times NV1$ matrix Q, defined by

$$Q_{j,s} := \sum_{T \in \mathcal{K}^1_{h|\Sigma}} \int_\Sigma \frac{L_\mu}{h_T} \mathbf{v}^{(j)} \cdot \mathbf{v}^{(s)} \, \mathrm{d}S \qquad (j, s = 1, .., NV1). \tag{7.5}$$

In terms of the vectors $\underline{\mathbf{m}}$ and $\underline{\mathbf{q}}$, Assumption $1h$ can now be written as follows. *There exist a $\beta^* > 0$ such that: for every $\underline{\mathbf{m}} \in \mathbb{R}^{NM1}$ there exists a $\underline{\mathbf{q}} \in \mathbb{R}^{NV1}$ with*

$$(\underline{\mathbf{q}}^t \, Q \, \underline{\mathbf{q}})^{1/2} = 1 \quad \text{ and } \quad \underline{\mathbf{q}}^t \, B^t \, \underline{\mathbf{m}} \, \geq \, \beta^* (\underline{\mathbf{m}}^t \, R \, \underline{\mathbf{m}})^{1/2}. \tag{7.6}$$

Our goal is to relate (7.6) with some known and computable properties of the matrices B, R, and Q. For this, consider, for every fixed $\underline{\mathbf{m}} \in \mathbb{R}^{NM1}$, the quantity

$$S(\underline{\mathbf{m}}) := \sup_{(\underline{\mathbf{q}}^t \, Q \, \underline{\mathbf{q}})^{1/2} = 1} \underline{\mathbf{q}}^t \, B^t \, \underline{\mathbf{m}} \, \equiv \, \sup_{\underline{\mathbf{q}} \neq 0} \frac{\underline{\mathbf{q}}^t \, B^t \, \underline{\mathbf{m}}}{(\underline{\mathbf{q}}^t \, Q \, \underline{\mathbf{q}})^{1/2}}, \tag{7.7}$$

where we take the supremum (which actually will be a maximum) to pick up *the best possible choice for* $\underline{\mathbf{q}}$ once $\underline{\mathbf{m}}$ has been fixed. The supremum in (7.7) is actually easily computable: with the change of variable $\underline{\mathbf{z}} := Q^{1/2}\underline{\mathbf{q}}$ (implying $\underline{\mathbf{q}}^t = \underline{\mathbf{z}}^t Q^{-1/2}$), we easily have

$$S(\underline{\mathbf{m}}) \;=\; \sup_{\underline{\mathbf{z}}\neq 0} \frac{\underline{\mathbf{z}}^t\, Q^{-1/2}\, B^t\, \underline{\mathbf{m}}}{(\underline{\mathbf{z}}^t \underline{\mathbf{z}})^{1/2}}, \tag{7.8}$$

and it is now easy to see that the supremum must be taken when $\underline{\mathbf{z}} = Q^{-1/2}\, B^t\, \underline{\mathbf{m}}$, giving

$$S(\underline{\mathbf{m}}) \;=\; (\underline{\mathbf{m}}^t\, B\, Q^{-1}\, B^t\, \underline{\mathbf{m}})^{1/2}. \tag{7.9}$$

On the other hand, it is clear that to check whether Assumption $1h$ holds you must, in terms of (7.6): *find the biggest* β *such that*

$$S(\underline{\mathbf{m}}) \;\geq\; \beta\, (\underline{\mathbf{m}}^t\, R\underline{\mathbf{m}})^{1/2}, \qquad \forall \underline{\mathbf{m}} \in \mathbb{R}^{NM1}. \tag{7.10}$$

Substituting the computed value (7.9) of $S(\underline{\mathbf{m}})$ into (7.10) and squaring both sides you have to find the biggest β such that

$$\underline{\mathbf{m}}^t\, B\, Q^{-1}\, B^t\, \underline{\mathbf{m}} \;\geq\; \beta^2 \underline{\mathbf{m}}^t\, R\underline{\mathbf{m}}, \qquad \forall \underline{\mathbf{m}} \in \mathbb{R}^{NM1}. \tag{7.11}$$

But this corresponds to find *the smallest eigenvalue of the (generalized) eigenvalue problem*

$$B\, Q^{-1}\, B^t\, \underline{\mathbf{m}} \;=\; \lambda\, R\underline{\mathbf{m}}. \tag{7.12}$$

Hence, here is the recipe: you compute the matrices B, R and Q, and you solve the generalized eigenvalue problem (7.12). If the smallest eigenvalue is equal to zero, you lost. If the smallest eigenvalue is not small (for reasonably fine meshes), you won. If it is small, try halving both meshes: if the smallest eigenvalue stays essentially constant, you won. If, say, the smallest eigenvalue is divided by four, *you definitely have a problem*. If you are not convinced, try halving the meshes again: if it is again divided by four, give up: Assumption $1h$, in the subdomain Ω^1, holds only with a constant β^*_M that goes to zero with h, and this will spoil your accuracy. As a general rule: if you perform the test on the same grid that you will use for computing your finite element solution, then *the smallest the constant you compute in the test, the more spoiled is likely to be the accuracy of your finite element solution* (compared with the best possible accuracy that your grid could provide, if you used a stable method).

A similar analysis can obviously be performed for $k = 2$ (that is, in the subdomain Ω^2) just by repeating the same procedure.

For assumption $2h$, instead, you have to choose a basis $\{\boldsymbol{\varphi}^{(1)}, ..., \boldsymbol{\varphi}^{(NP)}\}$ in $\mathbf{\Phi}_h$ (where NP denotes the dimension of $\mathbf{\Phi}_h$), and *a basis for the whole space* $\mathbf{M}_h$ (which means, taking into account both $\mathbf{M}^1_h$ *and* $\mathbf{M}^2_h$). Let then $\{\boldsymbol{\mu}^{(1)}, ..., \boldsymbol{\mu}^{(NM)}\}$ be such a basis, where the number $NM = NM1 + NM2$ is the dimension of $\mathbf{M}_h$. Now construct the matrices

$$C_{i,j} := \int_\Sigma \boldsymbol{\varphi}^{(i)} \cdot \boldsymbol{\mu}^{(j)}\, \mathrm{d}S \qquad (i = 1, .., NP;\; j = 1, .., NM), \tag{7.13}$$

$$R_{i,r} := \frac{L_\mu}{d_\Omega} \int_\Sigma \boldsymbol{\varphi}^{(i)} \cdot \boldsymbol{\varphi}^{(r)}\, \mathrm{d}S \qquad (i, r = 1, .., NP), \tag{7.14}$$

$$Q_{j,s} := \frac{d_\Omega}{L_\mu} \int_\Sigma \boldsymbol{\mu}^{(j)} \cdot \boldsymbol{\mu}^{(s)} \, \mathrm{d}S \qquad (j, s = 1, .., NM), \tag{7.15}$$

and solve (in $\mathbb{R}^{NP}$) the generalized eigenvalue problem

$$C\, Q^{-1}\, C^t\, \underline{\mathbf{p}} \;=\; \lambda\, R \underline{\mathbf{p}}. \tag{7.16}$$

The constant β^*_Φ will be the square root of the smallest eigenvalue of (7.16). And so on.

8 Conclusions

We have seen that the three-field formulation can be a viable method to deal with non-matching grids. In the context of elasticity problems it corresponds to have three different types of approximating fields: the displacements within each subdomain, the displacements on the interface between subdomains, and the normal component of the stress fields at the boundary of each subdomain (acting as a Lagrange multiplier to force the continuity between the displacements inside the subdomains and the displacements on the interface). One weak point is the necessity to have *two* types of *inf-sup* conditions satisfied. However, we propose here an alternative way to write them (using *mesh dependent norms*) that allows the use of a simple and reliable numerical test in order to check whether these *inf-sup* conditions are satisfied or not. For Engineers this might be even better than a theorem stating that *there exists a positive constant* β^**, independent of* h *such that...*, without actually knowing how small such a constant is. With the numerical test, for every mesh $\mathcal{K}_h$ you compute a $\beta^*(\mathcal{K}_h)$, and you might have difficulties in seeing whether or not there exists a positive β^*, independent of h, such that $\beta^*(\mathcal{K}_h) \geq \beta^*$ for all possible h: the best you can do is to just try a few meshes to see *the trend*. However, if you compute $\beta^*(\mathcal{K}_h)$ for the grid that you are willing to use, then at least you know *exactly* how small the constant is *on that grid*.

A Appendix: proof of Theorem 5.2

In this Appendix we report the detailed proof of Theorem 5.2, whose statement we recall for the convenience of the reader.

Theorem A.1 *Assume that our choices of finite element spaces for* $\mathbf{V}_h$*,* $\mathbf{M}_h$ *and* $\boldsymbol{\Phi}_h$ *satisfy Assumptions* $1h$ *and* $2h$*. If* $(\mathbf{u}, \boldsymbol{\lambda}, \boldsymbol{\psi})$ *and* $(\mathbf{u}_h, \boldsymbol{\lambda}_h, \boldsymbol{\psi}_h)$ *are the solutions of the elasticity problem* (2.9) *(as given in Theorem 2.1) and of the discretized problem* (5.1)-(5.3)*, respectively, then we have*

$$\begin{aligned} &\|\mathbf{u} - \mathbf{u}_h\|_{E,b} + \beta_M \|\boldsymbol{\lambda} - \boldsymbol{\lambda}_h\|_{\mathbf{M}} + \beta_\Phi \|\boldsymbol{\psi} - \boldsymbol{\psi}_h\|_{\boldsymbol{\Phi}} \\ &\qquad \leq C_A \Big(\inf_{\mathbf{v} \in \mathbf{V}_h} \|\mathbf{u} - \mathbf{v}\|_{E,b} + \inf_{\boldsymbol{\mu} \in \mathbf{M}_h} \|\boldsymbol{\lambda} - \boldsymbol{\mu}\|_{\mathbf{M}} + \inf_{\boldsymbol{\varphi} \in \boldsymbol{\Phi}_h} \|\boldsymbol{\psi} - \boldsymbol{\varphi}\|_{\boldsymbol{\Phi}} \Big), \end{aligned} \tag{A.1}$$

where C_A *depends only on the constants* β_M *and* β_Φ *appearing in* (5.6) *and* (5.7)*, respectively, and is bounded by* $1/\beta_M + 1/\beta_\Phi$*.*

Proof. The proof will be rather long and a little boring, but not difficult. In particular, as we shall see, the use of the “difficult norms” makes everything turn smoothly as a well lubricated engine.

We begin by noting that the following classical *Galerkin Orthogonality* property holds:

$$a(\mathbf{u}-\mathbf{u}_h,\mathbf{v}_h)+b(\mathbf{v}_h,\boldsymbol{\lambda}-\boldsymbol{\lambda}_h)=0 \quad \forall \mathbf{v}_h\in\mathbf{V}_h, \tag{A.2}$$

$$b(\mathbf{u}-\mathbf{u}_h,\boldsymbol{\mu}_h)+c(\boldsymbol{\mu}_h,\boldsymbol{\psi}-\boldsymbol{\psi}_h)=0 \quad \forall \boldsymbol{\mu}_h\in\mathbf{M}_h, \tag{A.3}$$

$$c(\boldsymbol{\lambda}-\boldsymbol{\lambda}_h,\boldsymbol{\varphi}_h)=0 \quad \forall \boldsymbol{\varphi}_h\in\boldsymbol{\Phi}_h, \tag{A.4}$$

with obvious meaning of the bilinear forms (see (4.2)-(4.4))

$$a=a^1+a^2, \qquad b=b^1+b^2, \qquad c=c^1+c^2. \tag{A.5}$$

Next, we define the *interpolants* $\mathbf{u}_i\in\mathbf{V}_h$, $\boldsymbol{\lambda}_i\in\mathbf{M}_h$, and $\boldsymbol{\psi}_i\in\boldsymbol{\Phi}_h$ of $\mathbf{u}$, $\boldsymbol{\lambda}$, and $\boldsymbol{\psi}$, respectively, as follows:

$$\begin{aligned}
&\|\mathbf{u}-\mathbf{u}_i\|_{E,b}\equiv\inf_{\mathbf{v}\in\mathbf{V}_h}\|\mathbf{u}-\mathbf{v}\|_{E,b},\\
&\|\boldsymbol{\lambda}-\boldsymbol{\lambda}_i\|_{\mathbf{M}}\equiv\inf_{\boldsymbol{\mu}\in\mathbf{M}_h}\|\boldsymbol{\lambda}-\boldsymbol{\mu}\|_{M},\\
&\|\boldsymbol{\psi}-\boldsymbol{\psi}_i\|_{\boldsymbol{\Phi}}\equiv\inf_{\boldsymbol{\varphi}\in\boldsymbol{\Phi}_h}\|\boldsymbol{\psi}-\boldsymbol{\varphi}\|_{\boldsymbol{\Phi}}.
\end{aligned} \tag{A.6}$$

(To be rigorous, we should actually call them *projections*, but the name *interpolants* is more evocative.) Finally, we start a lengthy add-and-subtract procedure. In all the following formulae, a-s stands for *add and subtract*:

$$\begin{aligned}
&\|\mathbf{u}-\mathbf{u}_h\|^2_{E,b}=(\text{ use (2.12)-(2.13) with (A.5) and (4.2)})\\
&=a(\mathbf{u}-\mathbf{u}_h,\mathbf{u}-\mathbf{u}_h)=(\text{a-s }\mathbf{u}_i)\\
&=\underbrace{a(\mathbf{u}-\mathbf{u}_h,\mathbf{u}-\mathbf{u}_i)}_{I}+a(\mathbf{u}-\mathbf{u}_h,\mathbf{u}_i-\mathbf{u}_h)=(\text{ use (A.2)})\\
&=I+b(\mathbf{u}_h-\mathbf{u}_i,\boldsymbol{\lambda}-\boldsymbol{\lambda}_h)=(\text{a-s }\mathbf{u})\\
&=I+\underbrace{b(\mathbf{u}-\mathbf{u}_i,\boldsymbol{\lambda}-\boldsymbol{\lambda}_h)}_{II}+b(\mathbf{u}_h-\mathbf{u},\boldsymbol{\lambda}-\boldsymbol{\lambda}_h)=(\text{ a-s }\boldsymbol{\lambda}_i)\\
&=I+II+\underbrace{b(\mathbf{u}_h-\mathbf{u},\boldsymbol{\lambda}-\boldsymbol{\lambda}_i)}_{III}+b(\mathbf{u}_h-\mathbf{u},\boldsymbol{\lambda}_i-\boldsymbol{\lambda}_h)=(\text{ use (A.3)})\\
&=I+II+III+c(\boldsymbol{\lambda}_i-\boldsymbol{\lambda}_h,\boldsymbol{\psi}-\boldsymbol{\psi}_h)=(\text{ a-s }\boldsymbol{\lambda})\\
&=I+II+III+\underbrace{c(\boldsymbol{\lambda}_i-\boldsymbol{\lambda},\boldsymbol{\psi}-\boldsymbol{\psi}_h)}_{IV}+c(\boldsymbol{\lambda}-\boldsymbol{\lambda}_h,\boldsymbol{\psi}-\boldsymbol{\psi}_h)=(\text{ a-s }\boldsymbol{\psi}_i)\\
&=I+II+III+IV+\underbrace{c(\boldsymbol{\lambda}-\boldsymbol{\lambda}_h,\boldsymbol{\psi}-\boldsymbol{\psi}_i)}_{V}+c(\boldsymbol{\lambda}-\boldsymbol{\lambda}_h,\boldsymbol{\psi}_i-\boldsymbol{\psi}_h)=(\text{ use (A.4)})\\
&=I+II+III+IV+V.
\end{aligned} \tag{A.7}$$

Our Cauchy-Schwarz inequalities (2.16), (2.31) and (2.29), together with the definitions (A.5) and (4.2)-(4.4), will provide an immediate estimate for the five terms I-V appearing in (A.7):

$$I \equiv a(\mathbf{u}-\mathbf{u}_h, \mathbf{u}-\mathbf{u}_i) \leq \|\mathbf{u}-\mathbf{u}_h\|_{E,b} \|\mathbf{u}-\mathbf{u}_i\|_{E,b}. \tag{A.8}$$

$$II \equiv b(\mathbf{u}-\mathbf{u}_i, \boldsymbol{\lambda}-\boldsymbol{\lambda}_h) \leq \|\mathbf{u}-\mathbf{u}_i\|_{E,b} \|\boldsymbol{\lambda}-\boldsymbol{\lambda}_h\|_{\mathbf{M}}. \tag{A.9}$$

$$III \equiv b(\mathbf{u}_h-\mathbf{u}, \boldsymbol{\lambda}-\boldsymbol{\lambda}_i) \leq \|\mathbf{u}-\mathbf{u}_h\|_{E,b} \|\boldsymbol{\lambda}-\boldsymbol{\lambda}_i\|_{\mathbf{M}}. \tag{A.10}$$

$$IV \equiv c(\boldsymbol{\lambda}_i-\boldsymbol{\lambda}, \boldsymbol{\psi}-\boldsymbol{\psi}_h) \leq \|\boldsymbol{\lambda}-\boldsymbol{\lambda}_i\|_{\mathbf{M}} \|\boldsymbol{\psi}-\boldsymbol{\psi}_h\|_{\mathbf{\Phi}}. \tag{A.11}$$

$$V \equiv c(\boldsymbol{\lambda}-\boldsymbol{\lambda}_h, \boldsymbol{\psi}-\boldsymbol{\psi}_i) \leq \|\boldsymbol{\lambda}-\boldsymbol{\lambda}_h\|_{\mathbf{M}} \|\boldsymbol{\psi}-\boldsymbol{\psi}_i\|_{\mathbf{\Phi}}. \tag{A.12}$$

To end the proof, we have to estimate $\boldsymbol{\lambda}-\boldsymbol{\lambda}_h$ and $\boldsymbol{\psi}-\boldsymbol{\psi}_h$ in terms of $\mathbf{u}-\mathbf{u}_h$. Let us start from $\boldsymbol{\lambda}-\boldsymbol{\lambda}_h$, and let us bound first the difference $\boldsymbol{\lambda}_i-\boldsymbol{\lambda}_h$. We use Assumption 1 with $\boldsymbol{\mu} = \boldsymbol{\lambda}_i-\boldsymbol{\lambda}_h$: we start using (5.6) and (4.3), then

$$\begin{aligned}
&\beta_M \|\boldsymbol{\lambda}_i-\boldsymbol{\lambda}_h\|_M \leq b(\mathbf{v}_\mu, \boldsymbol{\lambda}_h-\boldsymbol{\lambda}_i) = \text{ (a-s } \boldsymbol{\lambda}) \\
&= b(\mathbf{v}_\mu, \boldsymbol{\lambda}_h-\boldsymbol{\lambda}) + b(\mathbf{v}_\mu, \boldsymbol{\lambda}-\boldsymbol{\lambda}_i) = \text{ (use (A.2))} \\
&= a(\mathbf{u}-\mathbf{u}_h, \mathbf{v}_\mu) + b(\mathbf{v}_\mu, \boldsymbol{\lambda}-\boldsymbol{\lambda}_i) + \leq \text{(use (2.31) and (2.16))} \\
&\leq \|\mathbf{u}-\mathbf{u}_h\|_{E,b} \|\mathbf{v}_\mu\|_{E,b} + \|\boldsymbol{\lambda}-\boldsymbol{\lambda}_i\|_{\mathbf{M}} \|\mathbf{v}_\mu\|_{E,b} \text{ (use } \|\mathbf{v}_\mu\|_{E,b} = 1) \\
&= \|\boldsymbol{\lambda}-\boldsymbol{\lambda}_i\|_{\mathbf{M}} + \|\mathbf{u}-\mathbf{u}_h\|_{E,b}.
\end{aligned} \tag{A.13}$$

Combining (A.13) with the triangle inequality we easily get

$$\begin{aligned}
&\|\boldsymbol{\lambda}-\boldsymbol{\lambda}_h\|_{\mathbf{M}} \leq \|\boldsymbol{\lambda}-\boldsymbol{\lambda}_i\|_{\mathbf{M}} + \|\boldsymbol{\lambda}_i-\boldsymbol{\lambda}_h\|_{\mathbf{M}} \\
&\leq \|\boldsymbol{\lambda}-\boldsymbol{\lambda}_i\|_{\mathbf{M}} + \frac{1}{\beta_M}(\|\boldsymbol{\lambda}-\boldsymbol{\lambda}_i\|_{\mathbf{M}} + \|\mathbf{u}-\mathbf{u}_h\|_{E,b}) \\
&\leq \frac{\beta_M+1}{\beta_M}\|\boldsymbol{\lambda}-\boldsymbol{\lambda}_i\|_{\mathbf{M}} + \frac{1}{\beta_M}\|\mathbf{u}-\mathbf{u}_h\|_{E,b}.
\end{aligned} \tag{A.14}$$

For estimating $\boldsymbol{\psi}-\boldsymbol{\psi}_h$ we proceed in an identical manner: we just have to use Assumption 2 with $\boldsymbol{\varphi} = \boldsymbol{\psi}_i-\boldsymbol{\psi}_h$. We start using (5.7) and (4.4), then

$$\begin{aligned}
&\beta_\Phi \|\boldsymbol{\psi}_i-\boldsymbol{\psi}_h\|_M \leq c(\boldsymbol{\mu}_\varphi, \boldsymbol{\psi}_i-\boldsymbol{\psi}_h) = \text{ (a-s } \boldsymbol{\psi}) \\
&= c(\boldsymbol{\mu}_\varphi, \boldsymbol{\psi}_i-\boldsymbol{\psi}) + c(\boldsymbol{\mu}_\varphi, \boldsymbol{\psi}-\boldsymbol{\psi}_h) = \text{ (use (A.3))} \\
&= c(\boldsymbol{\mu}_\varphi, \boldsymbol{\psi}_i-\boldsymbol{\psi}) + b(\mathbf{u}_h-\mathbf{u}, \boldsymbol{\mu}_\varphi) \leq \text{ (use (2.29) and (2.31))} \\
&\leq \|\boldsymbol{\psi}-\boldsymbol{\psi}_i\|_{\mathbf{\Phi}} \|\boldsymbol{\mu}_\varphi\|_{\mathbf{M}} + \|\mathbf{u}-\mathbf{u}_h\|_{E,b} \|\boldsymbol{\mu}_\varphi\|_{\mathbf{M}} = \text{ (use } \|\boldsymbol{\mu}_\varphi\|_{\mathbf{M}} = 1) \\
&= \|\boldsymbol{\psi}-\boldsymbol{\psi}_i\|_{\mathbf{M}} + \|\mathbf{u}-\mathbf{u}_h\|_{E,b}.
\end{aligned} \tag{A.15}$$

As in (A.14) we can combine the triangle inequality with (A.15) and obtain

$$\begin{aligned}
&\|\boldsymbol{\psi}-\boldsymbol{\psi}_h\|_{\mathbf{\Phi}} \leq \|\boldsymbol{\psi}-\boldsymbol{\psi}_i\|_{\mathbf{\Phi}} + \|\boldsymbol{\psi}_i-\boldsymbol{\psi}_h\|_{\mathbf{\Phi}} \\
&\leq \|\boldsymbol{\psi}-\boldsymbol{\psi}_i\|_{\mathbf{\Phi}} + \frac{1}{\beta_\Phi}(\|\boldsymbol{\psi}-\boldsymbol{\psi}_i\|_{\mathbf{\Phi}} + \|\mathbf{u}-\mathbf{u}_h\|_{E,b}) \\
&\leq \frac{\beta_\Phi+1}{\beta_\Phi}\|\boldsymbol{\psi}-\boldsymbol{\psi}_i\|_{\mathbf{\Phi}} + \frac{1}{\beta_\Phi}\|\mathbf{u}-\mathbf{u}_h\|_{E,b}.
\end{aligned} \tag{A.16}$$

At this point we can come back to our estimate (A.7). We set $E_u := \|\mathbf{u} - \mathbf{u}_h\|_{E,b}$, $E_\lambda := \|\boldsymbol{\lambda} - \boldsymbol{\lambda}_h\|_{\mathbf{M}}$, $E_\psi := \|\boldsymbol{\psi} - \boldsymbol{\psi}_h\|_{\boldsymbol{\Phi}}$, then $I_u := \|\mathbf{u} - \mathbf{u}_i\|_{E,b}$, $I_\lambda := \|\boldsymbol{\lambda} - \boldsymbol{\lambda}_i\|_{\mathbf{M}}$, and finally $I_\psi := \|\boldsymbol{\psi} - \boldsymbol{\psi}_i\|_{\boldsymbol{\Phi}}$. Inserting the estimates (A.8)-(A.12) for I-V into (A.7) we have first

$$E_u^2 \leq E_u I_u + I_u E_\lambda + E_u I_\lambda + I_\lambda E_\psi + E_\lambda I_\psi \tag{A.17}$$

Inequalities (A.14) and (A.16), recalling that $\beta_M \leq 1$ and $\beta_\Phi \leq 1$, become now:

$$E_\lambda \leq \frac{2}{\beta_M} I_\lambda + \frac{1}{\beta_M} E_u \quad \text{and} \quad E_\psi \leq \frac{2}{\beta_\Phi} I_\psi + \frac{1}{\beta_\Phi} E_u, \tag{A.18}$$

that inserted in (A.17) give

$$E_u^2 \leq \Big(\frac{2}{\beta_M} + \frac{2}{\beta_\Phi}\Big)\Big(E_u(I_u + I_\lambda + I_\psi) + I_u I_\lambda + I_\lambda I_\psi\Big), \tag{A.19}$$

At this point we just need a suitable amount of arithmetic-geometric mean inequalities (valid for all real numbers a, b, and for all $\varepsilon > 0$)

$$2ab \leq \varepsilon a^2 + b^2/\varepsilon \qquad \text{(which is just telling that } (a\sqrt{\varepsilon} - b/\sqrt{\varepsilon})^2 \geq 0), \tag{A.20}$$

to reach the form

$$E_u^2 \leq C^2 (I_u^2 + I_\lambda^2 + I_\psi^2), \tag{A.21}$$

with C bounded by $1/\beta_M + 1/\beta_\Phi$. From (A.21) and (A.18) the desired estimate (A.1) follows easily. □

B Appendix: proof of Theorem 6.1

We present here the proof of Theorem 6.1, whose statement is recalled for convenience of the reader.

Theorem B.1 *Assume that our choices of finite element spaces for $\mathbf{V}_h$, $\mathbf{M}_h$ and $\boldsymbol{\Phi}_h$ satisfy Assumptions $1h$ and $2h$. If $(\mathbf{u}, \boldsymbol{\lambda}, \boldsymbol{\psi})$ and $(\mathbf{u}_h, \boldsymbol{\lambda}_h, \boldsymbol{\psi}_h)$ are the solutions of the elasticity problem* (2.9) *(as given in Theorem 2.1) and of the discretized problem* (5.1)-(5.3)*, respectively, then we have the following error estimate:*

$$\begin{aligned} &\|\mathbf{u} - \mathbf{u}_h\|_{E,b} + \beta_M^* \|\boldsymbol{\lambda} - \boldsymbol{\lambda}_h\|_{\mathbf{M}} \\ &\leq C_B \Big(\inf_{\mathbf{v}\in\mathbf{V}_h} \|\mathbf{u} - \mathbf{v}\|_{E,b} + (h_{\mathcal{K}}^{1/2} + h_{\mathcal{S}}^{1/2}) \inf_{\boldsymbol{\mu}\in\mathbf{M}_h} \|\boldsymbol{\lambda} - \boldsymbol{\mu}\|_{0,M} + \inf_{\boldsymbol{\varphi}\in\boldsymbol{\Phi}_h} \|\boldsymbol{\psi} - \boldsymbol{\varphi}\|_{\boldsymbol{\Phi}} \Big), \end{aligned} \tag{B.1}$$

where C_B has the form $C_B = c_B \Big(\frac{1}{\beta_M^*} + \frac{1}{(\beta_\Phi^*)^2}\Big)$ and c_B depends only on the shape of the elements in $\mathcal{S}_h$ and of those elements in $\mathcal{K}_h$ that have at least a vertex on Σ.

Before proving the new error estimates we need some additional notation and a few crucial lemmata. In (A.6) we introduced the interpolants of $\mathbf{u}$, $\boldsymbol{\lambda}$, and $\boldsymbol{\psi}$ in $\mathbf{V}_h$, $\mathbf{M}_h$ and $\boldsymbol{\Phi}_h$, respectively. In the sequel, we shall need to use the interpolants of displacements (as $\mathbf{v}$ or $\boldsymbol{\varphi}$) in the space $\mathbf{M}_h$. As this, dimensionally, could make people uneasy, we shall actually interpolate

$\kappa\mathbf{v}$ (or $\kappa\boldsymbol{\varphi}$), with κ defined in (6.7). Hence, for $\mathbf{v} \in \mathbf{V}$ or $\boldsymbol{\varphi} \in \boldsymbol{\Phi}$, we define $\mathbf{v}_{i,M}$ and $\boldsymbol{\varphi}_{i,M}$ in $\mathbf{M}_h$ as

$$\|\mathbf{v}_{i,M} - \kappa\mathbf{v}\|_{0,M} = \inf_{\boldsymbol{\mu}\in\mathbf{M}_h} \|\kappa\mathbf{v} - \boldsymbol{\mu}\|_{0,M}; \quad \|\boldsymbol{\varphi}_{i,M} - \kappa\boldsymbol{\varphi}\|_{0,M} = \inf_{\boldsymbol{\mu}\in\mathbf{M}_h} \|\kappa\boldsymbol{\varphi} - \boldsymbol{\mu}\|_{0,M}. \quad \text{(B.2)}$$

We recall, without proof (we refer, for instance, to [24]), the following approximation results:

$$\|\kappa\mathbf{v} - \mathbf{v}_{i,M}\|_{0,M} \leq C_{\mathcal{K}}\, h_{\mathcal{K}}^{1/2} \|\mathbf{v}\|_{\mathbf{D}}, \quad \text{and} \quad \|\kappa\boldsymbol{\varphi} - \boldsymbol{\varphi}_{i,M}\|_{0,M} \leq C_{\mathcal{K}}\, h_{\mathcal{K}}^{1/2} \|\boldsymbol{\varphi}\|_{\boldsymbol{\Phi}} \quad \text{(B.3)}$$

where $C_{\mathcal{K}}$ depends only on the shape of the elements in $\mathcal{K}_{h|\Sigma}$. We shall also need to interpolate on the *natural grid* with respect to the L^2-types norms. In particular we set

$$\|\boldsymbol{\lambda} - \boldsymbol{\lambda}_I\|_{0,M} = \inf_{\boldsymbol{\mu}\in\mathbf{M}_h} \|\boldsymbol{\lambda} - \boldsymbol{\mu}\|_{0,M}; \quad \|\boldsymbol{\varphi} - \boldsymbol{\varphi}_I\|_{0,D} = \inf_{\boldsymbol{\chi}\in\boldsymbol{\Phi}_h} \|\boldsymbol{\varphi} - \boldsymbol{\chi}\|_{0,D}. \quad \text{(B.4)}$$

Comparing (6.1)-(6.2) with (6.7)-(6.8), and using (6.11), we easily obtain

$$\|\boldsymbol{\lambda} - \boldsymbol{\lambda}_I\|_{\mathbf{M}_h} \leq h_{\mathcal{K}}^{1/2} \|\boldsymbol{\lambda} - \boldsymbol{\lambda}_I\|_{0,M}. \quad \text{(B.5)}$$

We also recall (see always [24]), that:

$$\|\boldsymbol{\varphi} - \boldsymbol{\varphi}_I\|_{0,D} \leq C_{\mathcal{S}}\, h_{\mathcal{S}}^{1/2} \|\boldsymbol{\varphi}\|_{\boldsymbol{\Phi}}, \quad \text{(B.6)}$$

where $C_{\mathcal{S}}$ depends only on the shape of the elements in $\mathcal{S}_h$. We consider now the following lemma.

Lemma B.2 *Under Assumption $2h$, for every $\boldsymbol{\lambda} = (\boldsymbol{\lambda}^1, \boldsymbol{\lambda}^2)$ in, say, $(L^2(\Sigma))^3 \times (L^2(\Sigma))^3$ there exist a unique $\boldsymbol{\lambda}_i^*$ in $\mathbf{M}_h$ and a unique $\boldsymbol{\zeta}^*$ in $\boldsymbol{\Phi}_h$ such that*

$$(\boldsymbol{\lambda} - \boldsymbol{\lambda}_i^*, \boldsymbol{\mu})_{0,M} = (\boldsymbol{\mu}, \boldsymbol{\zeta}^*)_{0,\Sigma} \quad \forall \boldsymbol{\mu} \in \mathbf{M}_h, \quad \text{(B.7)}$$

$$(\boldsymbol{\lambda} - \boldsymbol{\lambda}_i^*, \boldsymbol{\varphi})_{0,\Sigma} = 0 \quad \forall \boldsymbol{\varphi} \in \boldsymbol{\Phi}_h. \quad \text{(B.8)}$$

Moreover we have the estimates

$$\|\boldsymbol{\lambda} - \boldsymbol{\lambda}_i^*\|_{0,M} \leq \frac{2}{\beta_{\Phi}^*} \|\boldsymbol{\lambda} - \boldsymbol{\lambda}_I\|_{0,M}, \quad \beta_{\Phi^*} \|\boldsymbol{\zeta}^*\|_{0,M} \leq \|\boldsymbol{\lambda} - \boldsymbol{\lambda}_i^*\|_{0,M}, \quad \text{(B.9)}$$

$$\|\boldsymbol{\lambda} - \boldsymbol{\lambda}_i^*\|_{\mathbf{M}} \leq C_1 \Big(C_{\mathcal{K}}\, h_{\mathcal{K}}^{1/2} + C_{\mathcal{S}}\, h_{\mathcal{S}}^{1/2} \Big) \|\boldsymbol{\lambda} - \boldsymbol{\lambda}_I\|_{0,M}, \quad \text{(B.10)}$$

where $C_{\mathcal{K}}$ and $C_{\mathcal{S}}$ are the constants in (B.3) *and* (B.6)*, respectively, and C_1 is given by*

$$C_1 = \frac{10}{\gamma(\beta_{\Phi}^*)^2}, \quad \text{(B.11)}$$

where γ is the constant appearing in (2.32)*.*

P r o o f. As we did for Theorem 5.1 we shall prove directly the estimates (B.9) and (B.10) *for any possible solution* of (B.7)-(B.8). Since for $\boldsymbol{\lambda} = 0$ we obviously have $\boldsymbol{\lambda}_I = 0$ as well, then (B.9) will imply that for $\boldsymbol{\lambda} = 0$ the problem (B.7)-(B.8) can only have the zero solution: hence the determinant of the corresponding matrix is different from zero, hence for any given $\boldsymbol{\lambda}$ the system will have a unique solution. Let us prove, therefore, the estimates (B.9) and (B.10). We first estimate $\boldsymbol{\zeta}^*$ in terms of $\boldsymbol{\lambda}-\boldsymbol{\lambda}_i^*$ in the $\|\cdot\|_{0,M}$ norm. For this we use Assumption $2h$. Let $\boldsymbol{\mu}^*$ be the $\boldsymbol{\mu}_\varphi \in \mathbf{M}_h$ that we obtain from it taking $\boldsymbol{\varphi}=\boldsymbol{\zeta}^*$: we use (6.13), then (B.7), then the usual Cauchy-Schwarz inequality and the fact that $\|\boldsymbol{\mu}^*\|_{0,M}=1$ to obtain

$$\beta_\Phi^*\|\boldsymbol{\zeta}^*\|_{0,D} \le (\boldsymbol{\mu}^*,\boldsymbol{\zeta}^*)_{0,\Sigma} = (\boldsymbol{\lambda}-\boldsymbol{\lambda}_i^*,\boldsymbol{\mu}^*)_{0,M} \le \|\boldsymbol{\lambda}-\boldsymbol{\lambda}_i^*\|_{0,M}. \tag{B.12}$$

We now pass to the estimate of $\boldsymbol{\lambda}-\boldsymbol{\lambda}_i^*$ in the norm $\|\cdot\|_{0,M}$:

$$\begin{aligned}
&\|\boldsymbol{\lambda}-\boldsymbol{\lambda}_i^*\|_{0,M}^2 = (\boldsymbol{\lambda}-\boldsymbol{\lambda}_i^*,\boldsymbol{\lambda}-\boldsymbol{\lambda}_i^*)_{0,M} = \text{ (a-s } \boldsymbol{\lambda}_I)\\
&= \underbrace{(\boldsymbol{\lambda}-\boldsymbol{\lambda}_i^*,\boldsymbol{\lambda}-\boldsymbol{\lambda}_I)_{0,M}}_{I} + (\boldsymbol{\lambda}-\boldsymbol{\lambda}_i^*,\boldsymbol{\lambda}_I-\boldsymbol{\lambda}_i^*)_{0,M} = \text{ (use (B.7))}\\
&= I + (\boldsymbol{\zeta}^*,\boldsymbol{\lambda}_I-\boldsymbol{\lambda}_i^*)_{0,\Sigma} = \text{ (a-s } \boldsymbol{\lambda})\\
&= I + \underbrace{(\boldsymbol{\zeta}^*,\boldsymbol{\lambda}_I-\boldsymbol{\lambda})_{0,\Sigma}}_{II} + (\boldsymbol{\zeta}^*,\boldsymbol{\lambda}-\boldsymbol{\lambda}_i^*)_{0,\Sigma} = \text{ (use (B.8))} = I + II,
\end{aligned}$$

which implies, using (6.10), and then (B.12) for estimating $\boldsymbol{\zeta}^*$:

$$\|\boldsymbol{\lambda}-\boldsymbol{\lambda}_i^*\|_{0,M}^2 \le \|\boldsymbol{\lambda}-\boldsymbol{\lambda}_i^*\|_{0,M}\,\|\boldsymbol{\lambda}-\boldsymbol{\lambda}_I\|_{0,M} + \frac{1}{\beta_\Phi^*}\,\|\boldsymbol{\lambda}-\boldsymbol{\lambda}_i^*\|_{0,M}\,\|\boldsymbol{\lambda}-\boldsymbol{\lambda}_I\|_{0,M}, \tag{B.13}$$

that using $\beta_\Phi^* \le 1$, gives easily (B.9).

We now estimate $\kappa\boldsymbol{\zeta}^*$ in the norm $\mathbf{M}$. We remember that $\kappa\boldsymbol{\zeta}^*$ is single valued on Σ. Hence we can use (2.32), recalling that we obviously have $\gamma \le 1$. Let $\overline{\boldsymbol{\varphi}} \in \boldsymbol{\Phi}$ be the element that realizes the supremum in (2.32), that is

$$\gamma\,\|\kappa\boldsymbol{\zeta}^*\|_{\mathbf{M}} \le \sup_{\boldsymbol{\varphi}\in\boldsymbol{\Phi}} \frac{(\kappa\boldsymbol{\zeta}^*,\boldsymbol{\varphi})_{0,\Sigma}}{\|\boldsymbol{\varphi}\|_{\boldsymbol{\Phi}}} = \frac{(\kappa\boldsymbol{\zeta}^*,\overline{\boldsymbol{\varphi}})_{0,\Sigma}}{\|\overline{\boldsymbol{\varphi}}\|_{\boldsymbol{\Phi}}}. \tag{B.14}$$

Then we have

$$\begin{aligned}
&(\kappa\boldsymbol{\zeta}^*,\overline{\boldsymbol{\varphi}})_{0,\Sigma} = (\boldsymbol{\zeta}^*,\kappa\overline{\boldsymbol{\varphi}})_{0,\Sigma} = \text{ (a-s } \overline{\boldsymbol{\varphi}}_{i,M})\\
&= \underbrace{(\boldsymbol{\zeta}^*,\kappa\overline{\boldsymbol{\varphi}}-\overline{\boldsymbol{\varphi}}_{i,M})_{0,\Sigma}}_{I} + (\boldsymbol{\zeta}^*,\overline{\boldsymbol{\varphi}}_{i,M})_{0,\Sigma} = \text{ (use (B.7) with } \boldsymbol{\mu}=\overline{\boldsymbol{\varphi}}_{i,M})\\
&= I + (\boldsymbol{\lambda}-\boldsymbol{\lambda}_i^*,\overline{\boldsymbol{\varphi}}_{i,M})_{0,M} = \text{ (a-s } \kappa\overline{\boldsymbol{\varphi}})\\
&= I + \underbrace{(\boldsymbol{\lambda}-\boldsymbol{\lambda}_i^*,\overline{\boldsymbol{\varphi}}_{i,M}-\kappa\overline{\boldsymbol{\varphi}})_{0,M}}_{II} + (\boldsymbol{\lambda}-\boldsymbol{\lambda}_i^*,\kappa\overline{\boldsymbol{\varphi}})_{0,M} = \text{ (a-s } \kappa\overline{\boldsymbol{\varphi}}_I)\\
&= I + II + (\boldsymbol{\lambda}-\boldsymbol{\lambda}_i^*,\kappa\overline{\boldsymbol{\varphi}}-\kappa\overline{\boldsymbol{\varphi}}_I)_{0,M} + (\boldsymbol{\lambda}-\boldsymbol{\lambda}_i^*,\kappa\overline{\boldsymbol{\varphi}}_I)_{0,M} = \text{ (use (6.9))}\\
&= I + II + \underbrace{(\boldsymbol{\lambda}-\boldsymbol{\lambda}_i^*,\overline{\boldsymbol{\varphi}}-\overline{\boldsymbol{\varphi}}_I)_{0,\Sigma}}_{III} + (\boldsymbol{\lambda}-\boldsymbol{\lambda}_i^*,\overline{\boldsymbol{\varphi}}_I)_{0,\Sigma} = \text{ (use (B.8) with } \boldsymbol{\varphi}=\overline{\boldsymbol{\varphi}}_I)\\
&= I + II + III \le \text{ (use (6.10), with (B.3) for } I \text{ and } II \text{ and with (B.6) for } III)\\
&\le \left(C_{\mathcal{K}}\,h_{\mathcal{K}}^{1/2}\,\|\boldsymbol{\zeta}^*\|_{0,D} + C_{\mathcal{K}}\,h_{\mathcal{K}}^{1/2}\,\|\boldsymbol{\lambda}-\boldsymbol{\lambda}_i^*\|_{0,M} + C_{\mathcal{S}}\,h_{\mathcal{S}}^{1/2}\,\|\boldsymbol{\lambda}-\boldsymbol{\lambda}_i^*\|_{0,M}\right)\|\overline{\boldsymbol{\varphi}}\|_{\boldsymbol{\Phi}},
\end{aligned}$$

which combined with (B.12) and (B.14) gives immediately

$$\gamma \|\kappa\boldsymbol{\zeta}^*\|_{\mathbf{M}} \leq \Big(\frac{\beta^*_\Phi + 1}{\beta^*_\Phi} C_{\mathcal{K}}\, h_{\mathcal{K}}^{1/2} + C_{\mathcal{S}}\, h_{\mathcal{S}}^{1/2}\Big) \|\boldsymbol{\lambda} - \boldsymbol{\lambda}^*_i\|_{0,M}. \tag{B.15}$$

We finally estimate $\boldsymbol{\lambda} - \boldsymbol{\lambda}^*_i$ in the $\mathbf{M}$ norm. We remember that $\boldsymbol{\lambda} - \boldsymbol{\lambda}^*_i$ is double valued on the interface Σ. Let $\overline{\mathbf{v}}$ the element in $\mathbf{V}$ that realizes the supremum for the dual norm $\mathbf{M}$ (see (2.28)), that is

$$\|\boldsymbol{\lambda} - \boldsymbol{\lambda}^*_i\|_{\mathbf{M}} \equiv \sup_{\mathbf{v}\in\mathbf{V}} \frac{(\boldsymbol{\lambda} - \boldsymbol{\lambda}^*_i, \mathbf{v})_{0,\Sigma}}{\|\mathbf{v}\|_{\mathbf{D}}} = \frac{(\boldsymbol{\lambda} - \boldsymbol{\lambda}^*_i, \overline{\mathbf{v}})_{0,\Sigma}}{\|\overline{\mathbf{v}}\|_{\mathbf{D}}}. \tag{B.16}$$

Then, using (6.9), adding and subtracting $\overline{\mathbf{v}}_{i,M}$, using (B.7) , then adding and subtracting $\kappa\overline{\mathbf{v}}$, then using (6.10) and (B.3) we have

$$\begin{aligned}
&(\boldsymbol{\lambda} - \boldsymbol{\lambda}^*_i, \overline{\mathbf{v}})_{0,\Sigma} = (\boldsymbol{\lambda} - \boldsymbol{\lambda}^*_i, \kappa\overline{\mathbf{v}})_{0,M}\\
&= \underbrace{(\boldsymbol{\lambda} - \boldsymbol{\lambda}^*_i, \kappa\overline{\mathbf{v}} - \overline{\mathbf{v}}_{i,M})_{0,M}}_{I} + (\boldsymbol{\lambda} - \boldsymbol{\lambda}^*_i, \overline{\mathbf{v}}_{i,M})_{0,M}\\
&= I + (\boldsymbol{\zeta}^*, \overline{\mathbf{v}}_{i,M})_{0,\Sigma} = I + (\boldsymbol{\zeta}^*, \overline{\mathbf{v}}_{i,M} - \kappa\overline{\mathbf{v}})_{0,\Sigma} + (\kappa\boldsymbol{\zeta}^*, \overline{\mathbf{v}})_{0,\Sigma}\\
&\leq C_{\mathcal{K}}\, h_{\mathcal{K}}^{1/2}\, \|\overline{\mathbf{v}}\|_{\mathbf{D}} \,(\|\boldsymbol{\lambda} - \boldsymbol{\lambda}^*_i\|_{0,M} + \|\boldsymbol{\zeta}^*\|_{0,D}) + \|\kappa\boldsymbol{\zeta}^*\|_{\mathbf{M}}\, \|\overline{\mathbf{v}}\|_{\mathbf{D}},
\end{aligned} \tag{B.17}$$

which joined with (B.16) and (B.12) gives immediately

$$\|\boldsymbol{\lambda} - \boldsymbol{\lambda}^*_i\|_{\mathbf{M}} \leq C_{\mathcal{K}}\, h_{\mathcal{K}}^{1/2}\, \frac{\beta^*_\Phi + 1}{\beta^*_\Phi} \|\boldsymbol{\lambda} - \boldsymbol{\lambda}^*_i\|_{0,M} + \|\kappa\boldsymbol{\zeta}^*\|_{\mathbf{M}}. \tag{B.18}$$

Combining (B.18) with (B.15), and recalling that γ and β^*_Φ are ≤ 1, we have then

$$\|\boldsymbol{\lambda} - \boldsymbol{\lambda}^*_i\|_{\mathbf{M}} \leq \frac{1}{\gamma}\Big(\frac{4}{\beta^*_\Phi} C_{\mathcal{K}}\, h_{\mathcal{K}}^{1/2} + C_{\mathcal{S}}\, h_{\mathcal{S}}^{1/2}\Big) \|\boldsymbol{\lambda} - \boldsymbol{\lambda}^*_i\|_{0,M}. \tag{B.19}$$

This, joined with (B.9) (already proved) easily implies (B.10) with (B.11). □

Before tackling the final estimate, we need two additional results. To state them, however, it will be convenient to introduce, as an additional notation, the set $\widetilde{\mathcal{K}_h}$ defined as follows

$$\widetilde{\mathcal{K}_h} := \{K \in \mathcal{K}_h \mid K \text{ has at least a vertex on } \Sigma\} \tag{B.20}$$

We have then the following lemmata.

Lemma B.3 *Let $\mathbf{v} \in \mathbf{V}_h$ be a vector valued finite element pair that vanishes at all degrees of freedom that are not on Σ. Then*

$$\|\mathbf{v}\|_{E,b} \leq C_*\, \|\mathbf{v}\|_{\mathbf{D}_h}, \tag{B.21}$$

where the norm $\mathbf{D}_h$ is the one defined in (6.4)*, and C_* is a constant that depends only on the shape of the elements in $\widetilde{\mathcal{K}_h}$.*

P r o o f. The property is known (see e.g. [9]). It can also be easily checked by direct computation on the reference element and suitable scaling arguments. □

Lemma B.4 *There exists a constant C^*, depending only on the shape of the elements of $\widetilde{\mathcal{K}}_h$, such that: for every $\mathbf{v} \in \mathbf{V}$ we can find a $\mathbf{v}_h \in \mathbf{V}_h$ such that*

$$\|\mathbf{v} - \mathbf{v}_h\|_{\mathbf{D}_h} \leq C^* \|\mathbf{v}\|_{\mathbf{D}} \qquad \|\mathbf{v}_h\|_{E,b} \leq C^* \|\mathbf{v}\|_{\mathbf{D}}. \tag{B.22}$$

P r o o f. This is also a known property (see e.g. [9]). You have to define first $\mathbf{v}_h$ on Σ as a Clément interpolant of $\mathbf{v}$. The first inequality follows from local interpolation estimates plus the known fact that the sum of the squares of the $1/2$-seminorms on the elements is bounded by the $1/2$-seminorm on the whole domain. The second inequality follows from an extension theorem, by constructing a suitable prolongation of $\mathbf{v}_h$ to Ω (see e.g. [38] or [9]). □

We are finally ready for the error estimates on the three-field formulation, under the Assumptions $1h$ and $2h$. We follow exactly the procedure (A.7), using however $\boldsymbol{\lambda}_i^*$ instead of $\boldsymbol{\lambda}_i$. Let us see how the different terms I to V (appearing in (A.7)) can be estimated in the present situation.

$$I \equiv a(\mathbf{u} - \mathbf{u}_h, \mathbf{u} - \mathbf{u}_i) \leq \|\mathbf{u} - \mathbf{u}_h\|_{E,b} \|\mathbf{u} - \mathbf{u}_i\|_{E,b}. \tag{B.23}$$

$$II \equiv b(\mathbf{u} - \mathbf{u}_i, \boldsymbol{\lambda} - \boldsymbol{\lambda}_h) \leq \|\mathbf{u} - \mathbf{u}_i\|_{E,b} \|\boldsymbol{\lambda} - \boldsymbol{\lambda}_h\|_{\mathbf{M}}. \tag{B.24}$$

$$III \equiv b(\mathbf{u}_h - \mathbf{u}, \boldsymbol{\lambda} - \boldsymbol{\lambda}_i^*) \leq \|\mathbf{u} - \mathbf{u}_h\|_{E,b} \|\boldsymbol{\lambda} - \boldsymbol{\lambda}_i^*\|_{\mathbf{M}}. \tag{B.25}$$

$$\begin{aligned} IV &\equiv c(\boldsymbol{\lambda}_i^* - \boldsymbol{\lambda}, \boldsymbol{\psi} - \boldsymbol{\psi}_h) = \text{ (use (B.8))} \\ &= c(\boldsymbol{\lambda}_i^* - \boldsymbol{\lambda}, \boldsymbol{\psi} - \boldsymbol{\psi}_i) \leq \|\boldsymbol{\lambda} - \boldsymbol{\lambda}_i^*\|_{\mathbf{M}} \|\boldsymbol{\psi} - \boldsymbol{\psi}_i\|_{\boldsymbol{\Phi}}. \end{aligned} \tag{B.26}$$

$$V \equiv c(\boldsymbol{\lambda} - \boldsymbol{\lambda}_h, \boldsymbol{\psi} - \boldsymbol{\psi}_i) \leq \|\boldsymbol{\lambda} - \boldsymbol{\lambda}_h\|_{\mathbf{M}} \|\boldsymbol{\psi} - \boldsymbol{\psi}_i\|_{\boldsymbol{\Phi}}. \tag{B.27}$$

The $\mathbf{M}$-norm of $\boldsymbol{\lambda} - \boldsymbol{\lambda}_i^*$ has been estimated already. We need to estimate that of $\boldsymbol{\lambda} - \boldsymbol{\lambda}_h$. Proceeding as in (A.13) we use Assumption $1h$ with $\boldsymbol{\mu} = \boldsymbol{\lambda}_I - \boldsymbol{\lambda}_h$, and we remark that we are using the corresponding $\mathbf{v}_\mu$ only on Σ. This implies that we are allowed to assume that all its degrees of freedom that are not on Σ are set to zero, so that the property (B.21) holds. Then, keeping also in mind that $\|\mathbf{v}_\mu\|_{\mathbf{D}_h} = 1$ we have:

$$\begin{aligned} &\beta_M^* \|\boldsymbol{\lambda}_h - \boldsymbol{\lambda}_I\|_{\mathbf{M}_h} \leq \text{ (use (6.12) and (4.3))} \\ &\leq b(\mathbf{v}_\mu, \boldsymbol{\lambda}_I - \boldsymbol{\lambda}_h) = \text{ (a-s } \boldsymbol{\lambda}) \\ &= b(\mathbf{v}_\mu, \boldsymbol{\lambda}_I - \boldsymbol{\lambda}) + b(\mathbf{v}_\mu, \boldsymbol{\lambda} - \boldsymbol{\lambda}_h) \text{ (use (A.2))} \\ &= b(\mathbf{v}_\mu, \boldsymbol{\lambda}_I - \boldsymbol{\lambda}) + a(\mathbf{u}_h - \mathbf{u}, \mathbf{v}_\mu) \leq \text{ (use (6.6) and (2.16))} \\ &\leq \|\boldsymbol{\lambda} - \boldsymbol{\lambda}_I\|_{\mathbf{M}_h} \|\mathbf{v}_\mu\|_{\mathbf{D}_h} + \|\mathbf{u} - \mathbf{u}_h\|_{E,b} \|\mathbf{v}_\mu\|_{E,b} \leq \text{ (use (B.21))} \\ &\leq \|\boldsymbol{\lambda} - \boldsymbol{\lambda}_I\|_{\mathbf{M}_h} + C_* \|\mathbf{u} - \mathbf{u}_h\|_{E,b} \leq \text{ (use(B.5))} \\ &\leq h_{\mathcal{K}}^{1/2} \|\boldsymbol{\lambda} - \boldsymbol{\lambda}_I\|_{0,M} + C_* \|\mathbf{u} - \mathbf{u}_h\|_{E,b}. \end{aligned} \tag{B.28}$$

A simple use of the triangle inequality together with (B.5) and (B.28) gives then

$$\|\boldsymbol{\lambda} - \boldsymbol{\lambda}_h\|_{\mathbf{M}_h} \leq \frac{\beta_M^* + 1}{\beta_M^*} h_{\mathcal{K}}^{1/2} \|\boldsymbol{\lambda} - \boldsymbol{\lambda}_I\|_{0,M} + \frac{C_*}{\beta_M^*} \|\mathbf{u} - \mathbf{u}_h\|_{E,b}. \tag{B.29}$$

Finally we use Lemma B.4 to estimate the $\mathbf{M}$ norm of $\boldsymbol{\lambda} - \boldsymbol{\lambda}_h$: let first, as in (B.16), $\overline{\mathbf{v}}$ be the element in $\mathbf{V}$ that realizes the supremum for the dual norm $\mathbf{M}$, that is

$$\|\boldsymbol{\lambda} - \boldsymbol{\lambda}_h\|_{\mathbf{M}} \equiv \sup_{\mathbf{v} \in \mathbf{V}} \frac{(\boldsymbol{\lambda} - \boldsymbol{\lambda}_h, \mathbf{v})_{0,M}}{\|\mathbf{v}\|_{\mathbf{D}}} = \frac{(\boldsymbol{\lambda} - \boldsymbol{\lambda}_h, \overline{\mathbf{v}})_{0,M}}{\|\overline{\mathbf{v}}\|_{\mathbf{D}}}, \tag{B.30}$$

and let $\overline{\mathbf{v}}_h$ be the element in $\mathbf{V}_h$ described in Lemma B.4. Then we have:

$$\begin{aligned}
&(\boldsymbol{\lambda}-\boldsymbol{\lambda}_h,\overline{\mathbf{v}})_{0,\Sigma} = \text{ (a-s } \overline{\mathbf{v}}_h)\\
&= (\boldsymbol{\lambda}-\boldsymbol{\lambda}_h,\overline{\mathbf{v}}-\overline{\mathbf{v}}_h)_{0,\Sigma} + (\boldsymbol{\lambda}-\boldsymbol{\lambda}_h,\overline{\mathbf{v}}_h)_{0,\Sigma} = \text{ (use (A.2))}\\
&= (\boldsymbol{\lambda}-\boldsymbol{\lambda}_h,\overline{\mathbf{v}}-\overline{\mathbf{v}}_h)_{0,\Sigma} + a(\mathbf{u}-\mathbf{u}_h,\overline{\mathbf{v}}_h) \le \text{ (use (6.6) and (2.16))}\\
&\le \|\boldsymbol{\lambda}-\boldsymbol{\lambda}_h\|_{\mathbf{M}_h}\|\overline{\mathbf{v}}-\overline{\mathbf{v}}_h\|_{\mathbf{D}_h} + \|\mathbf{u}-\mathbf{u}_h\|_{E,b}\,\|\overline{\mathbf{v}}_h\|_{E,b} \text{ (use (B.22))}\\
&\le (\|\boldsymbol{\lambda}-\boldsymbol{\lambda}_h\|_{\mathbf{M}_h} + \|\mathbf{u}-\mathbf{u}_h\|_{E,b})\,C^*\,\|\overline{\mathbf{v}}\|_{\mathbf{D}},
\end{aligned} \tag{B.31}$$

which joined with (B.30) and then using (B.29) gives

$$\begin{aligned}
\|\boldsymbol{\lambda}-\boldsymbol{\lambda}_h\|_{\mathbf{M}} &\le C^*(\|\boldsymbol{\lambda}-\boldsymbol{\lambda}_h\|_{\mathbf{M}_h} + \|\mathbf{u}-\mathbf{u}_h\|_{E,b})\\
&\le \frac{1}{\beta_M^*}\Big(2C^*\,h_{\mathcal{K}}^{1/2}\|\boldsymbol{\lambda}-\boldsymbol{\lambda}_I\|_{0,M} + C^*\,(C_*+1)\|\mathbf{u}-\mathbf{u}_h\|_{E,b}\Big).
\end{aligned} \tag{B.32}$$

We can now proceed as in the previous section. We set $E_u := \|\mathbf{u}-\mathbf{u}_h\|_{E,b}$, $E_\lambda := \|\boldsymbol{\lambda}-\boldsymbol{\lambda}_h\|_{\mathbf{M}}$, then $I_u := \|\mathbf{u}-\mathbf{u}_i\|_{E,b}$, $I_\lambda^* := \|\boldsymbol{\lambda}-\boldsymbol{\lambda}_i^*\|_{\mathbf{M}}$, $I_\lambda := (h_{\mathcal{K}}^{1/2}+h_{\mathcal{S}}^{1/2})\|\boldsymbol{\lambda}-\boldsymbol{\lambda}_I\|_{0,M}$, and finally $I_\psi := \|\boldsymbol{\psi}-\boldsymbol{\psi}_i\|_{\boldsymbol{\Phi}}$. Inserting the estimates (B.23)-(B.27) for I-V into (A.7) we have this time

$$E_u^2 \le E_u\,I_u + I_u\,E_\lambda + E_u\,I_\lambda^* + I_\lambda^*\,I_\psi + E_\lambda\,I_\psi \tag{B.33}$$

Inequalities (B.10) and (B.32) become now:

$$I_\lambda^* \le \frac{C_2}{(\beta_\Phi^*)^2}\,I_\lambda \quad \text{ and } \quad E_\lambda \le \frac{C_3}{\beta_M^*}\,(I_\lambda + E_u), \tag{B.34}$$

with constants C_2 and C_3 depending only on γ, $C_{\mathcal{K}}$ and $C_{\mathcal{S}}$ (hence on the geometry of Ω and on the shapes of the elements in $\widetilde{\mathcal{K}_h}$ and $\mathcal{S}_h$). Inserting (B.34) into (B.33) gives

$$E_u^2 \le C_4\left(\frac{1}{\beta_M^*} + \frac{1}{(\beta_\Phi^*)^2}\right)\{E_u(I_u+I_\lambda+I_\psi) + I_u\,I_\lambda + I_\lambda\,I_\psi\}, \tag{B.35}$$

where again C_4 depends only on the geometry of Ω and on the shapes of the elements in $\widetilde{\mathcal{K}_h}$ and $\mathcal{S}_h$. With a few arithmetic-geometric mean inequalities (A.20), this can be reduced to the form (A.21), and the result of Theorem B.1 follows, using also (B.32) one more time.

With a *more complicated* argument, involving duality estimates and using quasi-uniformity of the meshes, we could also prove an estimate for $\boldsymbol{\psi}-\boldsymbol{\psi}_h$. (*"That's nothing to what I could say if I chose," the Duchess replied, in a pleased tone* [21].)

Acknowledgements We acknowledge the support of MIUR project PRIN-2003.

References

[1] M.A. Aminpour, J.B. Ransom, and S.L. Mc Cleary, *Coupled Analysis of Independently Modeled Finite Element Subdomains*, presented at the AIAA/ASME/ASCE/AHS/ASC 33rd Structures, Structural Dynamics, and Materials Conference, Dallas, Texas, April 13-15, (1992).

[2] M.A. Aminpour, J.B Ransom, and S.L. Mc Cleary, *A coupled analysis method for structures with independently modeled finite-element subdomains*, Int. J. Numer. Meth. Engrg., **3**8, 3695-3718 (1995).
[3] T. Arbogast, L. Cowsar, M.F. Wheeler and I. Yotov, *Mixed finite element methods on nonmatching multiblock grids*. SIAM J. Numer. Anal. **3**7 1295–1315 (2000).
[4] S.N Atluri, T. Nishioka, and M. Nakagaki, *Numerical Modeling of Dynamic and Nonlinear Crack Propagation in Finite Bodies by Moving Singular Elements*, in Nonlinear and Dynamic Fracture Mechanics, edited by N. Perrone and S.N. Atluri, Vol. 35, (ASME 1979), p. 66.
[5] I. Babuška, *The finite element method with lagrangian multipliers*, Numer.Math., **2**0, 179-192 (1973).
[6] C. Baiocchi, F. Brezzi, and L. D. Marini, *Stabilization of Galerkin methods and applications to domain decomposition*, in Future Tendencies in Computer Science, Control and Applied Mathematics, edited by A. Bensoussan et al., (Springer Lecture Notes in Computer Science vol. 653 1992) p. 345.
[7] H.J.C. Barbosa and T.J.R. Hughes, *Boundary Lagrange multipliers in finite element methods: error analysis in natural norms*, Numer.Math., **62**, 1-16 (1992).
[8] Ch. Bernardi, Y. Maday and A. Patera, *Domain decomposition by the mortar element method*, in Asymptotic and numerical methods for partial differential equations with critical parameters (Beaune, 1992), (NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci., 384, Kluwer Acad. Publ., Dordrecht, 1993) p. 269.
[9] Ch. Bernardi, Y. Maday and F. Rapetti, *Discrétisations variationnelles de problèmes aux limites elliptiques*, Springer Series on Mathématiques et Applications , Vol. 45 (Springer, Berlin, 2004)
[10] S. Bertoluzza, *Analysis of a stabilized three-fields domain decomposition method*, Numer. Math., **9**3, 611-634 (2003).
[11] S. Bertoluzza, *Substructuring Preconditioners for the Three Fields Domain Decomposition Method*, Math. Comput., **7**3, 659-689 (2004)
[12] S. Beroluzza and A. Kunoth, *Wavelet stabilization and preconditioning for domain decomposition*, IMA J. Numer. Anal., **2**0, 533-559 (2000).
[13] P.E. Bjørstad and O.B. Widlund, *Iterative methods for the solution of elliptic problems on regions partitioned into substructures*, SIAM J. Numer. Anal., **2**3, 1093–1120 (1986).
[14] F. Bouillault, A. Buffa, Y. Maday, and F. Rapetti *The mortar edge element method in three dimensions: application to magnetostatics*, SIAM J. Sci. Comput., **2**4, 1303–1327 (2003).
[15] J.H. Bramble, J. Pasciak and A. Schatz, *The construction of preconditioners for elliptic problems by substructuring, IV*, Math. Comput., **5**3, 1–24 (1989).
[16] F. Brezzi, L. P. Franca, L. D. Marini, and A. Russo, *Stabilization techniques for domain decomposition methods with non-matching grids*, In Domain Decomposition Methods in Science and Engineering, ($9th$ International Conference, Bergen, Norway, 1996, edited by P. Biørstad et al. (Domain Decomposition Press, Bergen, 1998) p. 1.
[17] F. Brezzi and M. Fortin, *Mixed and hybrid finite element methods*, (Springer-Verlag, New York, 1991).
[18] F. Brezzi and L. D. Marini, *Macro hybrid elements and domain decomposition methods*, in Optimization et Contrôle, Meeting in honour of J. Céa, April 8-10 (1992), edited by J. Desideri et al., (CÉPADUÈs-Editions,Toulouse, 1993) p. 89.
[19] F. Brezzi and L. D. Marini *Error estimates for the three-field formulation with bubble stabilization*, Math. Comput., **7**0 , 911-934 (2000), .
[20] A. Buffa, *Error estimate for a stabilized domain decomposition method with nonmatching grids*. Numer. Math., **9**0, 617-640 (2002).
[21] L. Carroll, *Alice's adventures in Wonderland*, Chapter IX, The Mock Turtle's Story, (Oxford, 1865)
[22] T.F. Chan and T.P. Mathew, *Domain decomposition algorithms*, Acta Numerica, 61–144, (1994).
[23] D. Chapelle and K.J. Bathe, *The inf-sup test*, Comput. & Structures, **4**7, 537–545 (1993).
[24] P. G. Ciarlet, *The finite element method for elliptic problems*, (North-Holland Publishing Company, Amsterdam 1980).

[25] C. Eck and B. Wohlmuth, *Convergence of a Contact-Neumann Iteration for the Solution of Two-Body Contact Problems*, Math. Mod. Meth. Appl. Sci., **1**3, 1103-1118 (2003)
[26] C. Farhat, M. Lesoinne, P. LeTallec, K. Pierson and D. Rixen, *FETI-DP: a dual-primal unified FETI method. I. A faster alternative to the two-level FETI method*, Int. J. Numer. Meths. Engrg., **5**0, 1523-1544 (2001).
[27] S. Gaiffe, R. Glowinski and R. Masson, *Domain decomposition and splitting methods for mortar mixed finite element approximations to parabolic equations*, Numer. Math., **9**3, 53-75 (2002).
[28] R. Glowinski, W.K. Kinton and M.F. Wheeler, *Acceleration of domain decomposition algorithms for mixed finite elements by multilevel methods*, in Third Int. Symp. on Domain decomposition methods for partial differential equations, edited by T. Chan et al., (SIAM, Philadelphia, 1990).
[29] A. Iosilevich, K.J. Bathe, F. Brezzi, *On evaluating the Inf-Sup condition for plate bending elements*, Int. J. Num. Meth. in Eng., **4**0 3639-3663 (1997).
[30] Yu. A. Kuznetsov and M.F. Wheeler, *Optimal order substructuring preconditioners for mixed finite element methods on nonmatching grids*, East-West Journal of Numerical Mathematics, **3**, 127–144 (1995).
[31] P. Le Tallec, *Domain decomposition methods in computational mechanics*, Comput. Mech. Adv. **1** (1994), 121–220.
[32] Q. Liu, L.K. Li and J.G. Huang, *On mortar-type TRUNC element method for a plate bending problem*, Math. Mod. Meth. Appl. Sci., **1**3, 545-571 (2003)
[33] A. Quarteroni and A. Valli, *Domain decomposition methods for partial differential equations*, Numerical Mathematics and Scientific Computation, (Oxford University Press, New York, 1999).
[34] G. Rapin and G. Lubet, *A stabilized scheme for the lagrange multiplier method for advection-diffusion equations*, Math Mod. Meth. Appl. Sci., **1**4, 1035-1060 (2004)
[35] B.F Smith and O.B. Widlund, *A domain decomposition algorithm using a hierarchical basis*, SIAM J. Sci. Comput., **1**1, 1212–1220 (1990).
[36] P. Tong, *New displacement hybrid finite element model for solid continua*, Int. J. Num. Meths. Engrg., **2**, 73-83 (1970).
[37] C.H. Tong, T.F. Chan and C.C.J. Kuo, *A domain decomposition preconditioner based on a change to a multilevel nodal basis*, SIAM J. Sci. Comput., **1**2, 1486–1495 (1991).
[38] O.B. Widlund, *An extension theorem for finite element spaces with three applications*, in Numerical Techniques in Continuum Mechanics, Proceedings of the Second GAMM Seminar, edited by W. Hackbush and K. Witsch, (Kiel, 1986)
[39] B. Wohlmuth, *Discretization Techniques and Iterative Solvers Based on Domain Decomposition*, Lectures Notes in Computational Science and Engineering, **1**7 (Springer, Berlin, 2001).
[40] J. Xu, *Theory of multilevel methods*, PhD thesis, Cornell University (1989).