CONFLUENTES MATHEMATICI

Serge NICAISE, Ismael MERABET, and Rihana REZZAG BARA

A priori and a posteriori error analysis for a hybrid formulation of a prestressed shell model

Tome 14, nº 2 (2022), p. 53–86.
https://doi.org/10.5802/cml.87

Confluentes Mathematici est membre du
Centre Mersenne pour l'édition scientifique ouverte
http://www.centre-mersenne.org/
e-ISSN : 1793-7434

Confluentes Math.
14, 2 (2022) 53-86

A PRIORI AND A POSTERIORI ERROR ANALYSIS FOR A HYBRID FORMULATION OF A PRESTRESSED SHELL MODEL

SERGE NICAISE, ISMAEL MERABET, AND RIHANA REZZAG BARA

Abstract. This work deals with the finite element approximation of a prestressed shell model in the case of isometrically deformed shell. Using a new formulation where the unknowns (the displacement and the rotation of fibers normal to the midsurface) are described in Cartesian and local covariant basis respectively. Due to the constraint involved in the definition of the functional space, a penalized version is then considered. We obtain a non robust a priori error estimate of this penalized formulation, but a robust one is obtained for its mixed formulation. Moreover, we present a reliable and efficient a posteriori error estimator of the penalized formulation. Numerical tests are included that confirm the efficiency of our residual a posteriori estimator.

1. Introduction

Shell models can be 3D or 2D models, i. e., systems of PDEs involving three or two independent variables [9, 7]. To reduce the numerical costs, two dimensional models are often preferred to the three-dimensional ones when they represent a "good" approximation of them. For 2D models, the unknowns are given on the mid-surface of the shell and they could be described in local (curvilinear) coordinates system (see [9]) or global (Cartesian) coordinate system (see [4]). Another possibility of describing shell models is to use a hybrid formulation, i.e., use global and local coordinates system to describe some measurable physical quantities (displacement, rotation, stress, ...), arising in response to given loads and boundary conditions.

Prestressing of a structure is the intentional creation of permanent stresses in the structure for the purpose of improving its performance under various service conditions. In our days this concept is widely employed in the design of buildings, towers, bridges, etc. Consequently the numerical analysis of such models is of great importance.

In this work we are then performing some error analysis of a prestressed (two-dimensional) shell model which was introduced for the first time in [17]. This model is the same as the one of a parametrized shell up to the addition of a prestressed energy term. This term (as well as the flexural one) is derived from the Kirchhoff model of the bending of the nonlinear elastic plate (obtained as a limit of three-dimensional nonlinear elasticity). The unknown of the problem is the couple (u, r), where u is the displacement from the reference configuration and r is the infinitesimal rotation of the cross section of the shell. In [17] both u and r are described in Cartesian coordinates and they are sought in the Sobolev space H^1 (each one is a vector field with three components). However, the bilinear form describing the model involves the first order derivative of the components $u_i, i = 1, 2, 3$ and $r_\alpha, \alpha = 1, 2$, whereas, it does not use any derivative for the component r_3. This

2020 *Mathematics Subject Classification:* 74S05, 65N50.
Keywords: finite element, adaptive methods, penalty method, hybrid formulation.

causes a loss of coercivity of the bilinear form on the space H^1. In order to solve this issue, a larger Hilbert space was considered in [20], where the third component r_3 is sought in the L^2 space.

A hybrid formulation is considered here, i.e., the unknowns (the displacement and the rotation to the shell midsurface are described respectively in Cartesian and local covariant basis). The use of a hybrid formulation in a similar spirit of the present paper was used in Blouza [3] for Naghdi's shell model. The aim of using hybrid formulation in [3] was to reduce the number of the unknowns (from six to five because $s \cdot a_3 = 0$) and to get rid of the tangency constraint for the rotation which was presented by Blouza and al [4]. Hybrid formulation allows to use conforming finite element methods on unconstrained functional space with a smaller number of degrees of freedom. Another hybrid formulation of general shell element involving three incremental displacements corresponding to the stationary global coordinate directions and two rotations described in a local coordinates system was used in [1].

The purpose of this work is to provide a non robust a priori error analysis and a robust a posteriori error estimator of the penalized version of the hybrid formulation (but a robust a priori error is obtained for its mixed formulation). These a posteriori estimators yield global upper and local lower bounds for the error (the distance between exact solution and its approximated solution in the energy norm). When the error estimator provides an upper bound for the error, this means, that our estimator is "reliable" and it is called (locally) efficient if it provides a (local) lower bound for the (local) error apart from (local) data resolution. Different types of a posteriori estimators are available in the literature. Here we perform an a posteriori analysis of residual type of the penalized version and prove upper and lower bounds for the error, with an explicit dependancy of the involved constants with respect to the penalization parameter. Based on those estimators, adapted meshes can be constructed allowing to compute an approximated solution with a given accuracy using a smaller number of degrees of freedom than uniform meshes.

The a posteriori indicators are computable, since they depend on known quantities such as the size of the mesh cells, the problem data, and the approximate solution. Efficient a posteriori estimators are successfully used for adaptive algorithms that involve local mesh refinements. A lot of works concerning a rigorous mathematical justification of the convergence of adaptive finite element methods can be found. The basic idea is to prove a contraction property of the errors between two consecutive adaptive meshes, but most of these works are concerned with simple model problems. We do not give a rigorous mathematical justification of such result for our model. However, the error indicators that we propose show good convergence results. Finally different numerical tests are presented in order to show that the use of adaptive meshes speeds up the convergence and to compare the penalized formulation and the mixed one.

For plates and shell models, there already exist several a posteriori error estimation approaches. We refer to [5, 6, 15, 14] for the pioneering works concerning plate models. Up to our knowledge, the first a posteriori estimate concerning shell models formulated in global coordinate system was done in [2] for Naghdi's shell model.

An outline of the paper is as follows.

- In section 2 we recall some geometrical preliminaries of surfaces and the prestressed model presented in [20].
- In section 3 we present a hybrid formulation of a prestressed shell model where the unknowns (the displacement and the rotation of fibers normal to the midsurface) are described in Cartesian and local covariant basis respectively, we study the existence and uniqueness of the solution. We then present a penalized version for the new variational formulation and prove its well-posedness.
- Section 4 is devoted to the finite element approximation for the penalized problem and we prove the existence and uniqueness of the discrete solution, we derive a priori error estimates.
- The strong formulation of the penalized problem is detailed in section 5.
- In section 6 we derive a posteriori estimates and prove the reliability and efficiency of our a posteriori error estimator.
- Numerical experiments are presented in section 7.

Let us finish this introduction with some notation used in the paper. The usual norm and semi-norm of $W^{s,p}(\Omega,\mathbb{R}^\ell)$ (with $s\geqslant 0$, $p\in[1,\infty]$ and $\ell\in\mathbb{N}$) are denoted by $\|\cdot\|_{s,p,\Omega}$ and $|\cdot|_{s,p,\Omega}$, respectively (the scalar or vector-valued character of the involved functions is not specified because there is no possible ambiguity). For $s=0$ (resp. $p=2$), we drop the index s (resp. p). The notation $A\lesssim B$ is used for the estimate $A\leqslant C\,B$, where C is a generic constant that does not depend on A and B, in particular this constant does not depend on the penalization parameter ϵ and the mesh size h, but it may depend on the thickness of the shell t which is supposed to be a strictly positive constant. Further for two vectors v and w of $\mathbb{R}^n$ (with $n=2$ or 3, written in column), we denote by $v\cdot w=v^\top w$ their euclidean inner product. Similarly for two 2×2 matrices $M=(m_{\alpha\beta})_{1\leqslant\alpha,\beta\leqslant 2}$ and $N=(n_{\alpha\beta})_{1\leqslant\alpha,\beta\leqslant 2}$ with real or vector valued coefficients, we set

$$M\colon N=\sum_{\alpha,\beta=1,2} m_{\alpha\beta}\cdot n_{\alpha\beta}.$$

2. Presentation of the model.

2.1. **Geometry of the shell midsurface.** As standard in the present context, Greek indices and exponents take their values in the set $1,2$ and Latin indices and exponents take their values in the set $1,2,3$. Unless otherwise specified, the summation convention for repeated indices and exponents according to this set of values is assumed.

For a given domain ω of $\mathbb{R}^2$ with a Lipschitz boundary, we consider a shell whose middle surface S is given by

$$S=\boldsymbol{\varphi}\left(\bar{\omega}\right),$$

where $\boldsymbol{\varphi}\in W^{2,\infty}\left(\omega,\mathbb{R}^3\right)$ is supposed to be a one-to-one mapping and isometric, i.e., $\boldsymbol{\varphi}\in\mathcal{A}_d$, where,

$$\mathcal{A}_d=\{\psi\in W^{2,2}(\omega,\mathbb{R}^3);|\partial_1\psi|=|\partial_2\psi|=1,\partial_1\psi\cdot\partial_2\psi=0\}. \tag{2.1}$$

Following [17] and [20] the assumption that the chart $\boldsymbol{\varphi}$ is isometric is crucial to obtain the considered prestressed model. Indeed, if ψ is an isometric deformation of ω obtained by applying some external forces, then the local basis of ψ is the product of an orthogonal matrix (of determinant 1) times the local basis of $\boldsymbol{\varphi}$.

In other words, the new local basis is just a rotation of the old one. Linear or nonlinear shell models are usually expressed in terms of the components of the change of metric tensor and those of the change of curvature tensor between two deformed surfaces (see Ciarlet [9] for instance). Using the fact that both deformed surfaces are isometric simplifies significantly the expression of those components (see [20, p. 3]). Of course, prestressed model can be derived without using the isometric deformation assumption (see Chapelle and Bathe [7, p. 216] for instance), but other a priori mechanical assumptions have to be added. Those assumptions are very hard to justify mathematically. Moreover, its well-posedness may need some additional assumptions.

Now, we define the local basis $a_i, i = 1, 2, 3$ by

$$a_\alpha = \partial_\alpha \varphi, \ \alpha = 1, 2 \text{ and } a_3 = a_1 \times a_2. \tag{2.2}$$

Using the fact that φ is isometric we have

$$a_i \cdot a_j = \delta_i^j, \tag{2.3}$$

δ_i^j being the Kronecker symbol. The contravariant basis $\{a^i\}_{i=1,2,3}$ is then equal to the covariant basis $\{a_i\}_{i=1,2,3}$. Consequently, the covariant and contravariant components of the metric (or the first fundamental form) are equal to the identity matrix

$$(a_{\alpha\beta}) = (a_\alpha \cdot a_\beta) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ and } (a^{\alpha\beta}) = (a_{\alpha\beta}), \quad a(x) = \det(a_{\alpha\beta}) = 1. \tag{2.4}$$

The second fundamental forms of the surface is given in covariant components by

$$b_{\alpha\beta} = b_{\beta\alpha} = a_3 \cdot \partial_\beta a_\alpha = -a_\alpha \cdot \partial_\beta a_3.$$

The Christoffel symbols of the surface $\Gamma^\rho_{\alpha\beta}$ take the form

$$\Gamma^\rho_{\alpha\beta} = \Gamma^\rho_{\beta\alpha} = a^\rho \cdot \partial_\beta a_\alpha = a^\rho \cdot \partial_\alpha a_\beta.$$

The linearized strain tensor is given by

$$\gamma(u) = (\gamma_{\alpha\beta}(u))_{1 \leqslant \alpha, \beta \leqslant 2},$$

where the components $\gamma_{\alpha\beta}(u)$ are defined by

$$\gamma_{\alpha\beta}(u) = \frac{1}{2}\left(\partial_\alpha u \cdot a_\beta + \partial_\beta u \cdot a_\alpha\right), \tag{2.5}$$

and the symmetrized linearized second fundamental form is defined by

$$\Pi(s) = \begin{pmatrix} \partial_1 s \cdot a_2 & \frac{1}{2}(\partial_2 s \cdot a_2 - \partial_1 s \cdot a_1) \\ \frac{1}{2}(\partial_2 s \cdot a_2 - \partial_1 s \cdot a_1) & -\partial_2 s \cdot a_1 \end{pmatrix}. \tag{2.6}$$

2.2. **The variational formulation in cartesian coordinates.** The unknown of the problem is the couple (u, r) where u is the displacement from the reference configuration ω and r is the infinitesimal rotation of the cross-section of the shell. We assume that the shell is fixed on a part Γ_0 of positive measure of the boundary of ω. According to [17, 20], the model takes the following variational form

$$\begin{cases} \text{Find } U = (u, r) \in \mathbb{V} \text{ such that} \\ \mathbf{a}(U, V) + \boldsymbol{a}_p(U, V) = \mathcal{L}(V), \quad \forall V = (v, s) \in \mathbb{V}, \end{cases} \tag{2.7}$$

where

$$\mathbb{V} = \{(v,s) \in H^1(\omega,\mathbb{R}^3) \times L^2(\omega,\mathbb{R}^3) \ : \ s\cdot a_\alpha \in H^1(\omega), \text{ and}$$
$$s\cdot a_3 = \tilde{\gamma}_{12}(v) = \frac{1}{2}(\partial_1 v\cdot\partial_2\varphi - \partial_2 v\cdot\partial_1\varphi), \quad v|_{\Gamma_0} = (s\cdot a_\alpha)|_{\Gamma_0} = 0\}, \tag{2.8}$$

and

$$\boldsymbol{a}(U,V) = ta_m(u,v) + ta_s((u,r),(v,s)) + \frac{t^3}{12}a_f(r,s),$$

t being the thickness of the shell assumed to be constant and positive. The bilinear forms $a_m(\cdot,\cdot), a_f(\cdot,\cdot)$ and $a_s(\cdot,\cdot)$ respectively corresponding to the membrane, flexural, and the transverse shear energies are given by

$$a_m(u,v) = \frac{4\lambda\mu}{\lambda+2\mu}\int_\omega \mathrm{tr}\gamma(u)\mathrm{tr}\gamma(v)\ \ dx + 4\mu\int_\omega \gamma(u)\colon\gamma(v)\ \ dx, \tag{2.9}$$

$$a_f(r,s) = \frac{2\lambda\mu}{\lambda+2\mu}\int_\omega \mathrm{tr}\Pi(r)\mathrm{tr}\Pi(s)\ \ dx\ + 2\mu\int_\omega \Pi(r)\colon\Pi(s)\ \ dx, \tag{2.10}$$

$$a_s((u,r),(v,s)) = \mu\int_\omega a_3^\top(\nabla u - r\times\nabla\varphi)\left[a_3^\top(\nabla v - s\times\nabla\varphi)\right]^\top\ dx, \tag{2.11}$$

where μ and λ are the Lamé moduli of the homogeneous and isotropic material that constitutes the shell. As usual ∇v is the jacobian matrix of v, namely

$$\nabla v = (\partial_1 v, \partial_2 v) = \begin{pmatrix} \partial_1 v_1 & \partial_2 v_1 \\ \partial_1 v_2 & \partial_2 v_2 \\ \partial_1 v_3 & \partial_2 u_3 \end{pmatrix}.$$

Furthermore as in [17], we have $s\times\nabla\varphi = (s\times a_1, s\times a_2)$.

The contribution of the prestressed term is represented by

$$\boldsymbol{a}_p(r,s) = \frac{t^3}{12}\left(2\mu\int_\omega \mathrm{tr}\left((II_0 + II_0^t)\tau(r,s)\right)\ dx + \frac{4\lambda\mu}{2\mu+\lambda}\int_\omega \mathrm{tr}II_0\mathrm{tr}\tau(r,s)\ dx\right), \tag{2.12}$$

where

$$\tau(r,s) = \theta(r)(s\cdot a_3) + \theta(s)(r\cdot a_3) \tag{2.13}$$

with

$$\theta(s) = \frac{1}{2}\begin{pmatrix} -\gamma_{11}(s) & \tilde{\gamma}_{12}(s) \\ \tilde{\gamma}_{12}(s) & \gamma_{22}(s) \end{pmatrix}, \tag{2.14}$$

and

$$II_0 = (\nabla\varphi)^\top\cdot\nabla a_3 = \begin{pmatrix} \partial_1\varphi\cdot\partial_1 a_3 & \partial_1\varphi\cdot\partial_2 a_3 \\ \partial_2\varphi\cdot\partial_1 a_3 & \partial_2\varphi\cdot\partial_2 a_3 \end{pmatrix}.$$

Note that II_0 is symmetric and therefore in (2.12) the factor $II_0 + II_0^t$ may be replaced by $2II_0$. Note further that the prestressed term $\boldsymbol{a}_p(r,r)$ is not necessarily positive for an arbitrary element $(v,r)\in\mathbb{V}$.

In our previous paper [20] we have stressed on the importance and the effect of the prestressed term $\boldsymbol{a}_p$. We recall that the nonpositive character of the prestressed term may break the coercivity of the bilinear form on the space of admissible test functions space $\mathbb{V}$ even if the bilinear form without prestressed term is $\mathbb{V}$-elliptic. On the other hand, by mean of the considered model formulated in global coordinates system, we have observed that the solution with prestress term is more stiff than

the corresponding model without prestressing which confirms a naturel phenomenon that can be observed in several applications.

The linear form $\mathcal{L}$ is given by

$$\mathcal{L}(v,s) = \int_\omega f \cdot v \, dx,$$

with $f \in L^2(\omega, \mathbb{R}^3)$ that represents a given resultant force density.

3. A hybrid formulation

Let us introduce the space $\mathbb{W}$ such that the displacement and the rotation are described in Cartesian and local covariant or contravariant basis respectively.

$$\mathbb{W} = \{(v, s = \sum_{i=1}^{3} s_i a_i) \in H^1\left(\omega, \mathbb{R}^3\right) \times (L^2(\omega))^3 \mid s_\alpha \in H^1(\omega), \text{ and}$$
$$s_3 = \tilde{\gamma}_{12}(v) = \frac{1}{2}(\partial_1 v \cdot \partial_2 \varphi - \partial_2 v \cdot \partial_1 \varphi), \text{ a.e. in } \omega, \quad v|_{\Gamma_0} = s_\alpha|_{\Gamma_0} = 0\}, \tag{3.1}$$

equipped with the norm

$$\|(v,s)\|_{\mathbb{X}} = \Big(\|v\|^2_{H^1(\omega,\mathbb{R}^3)} + \sum_{\alpha=1,2} \|s_\alpha\|^2_{H^1(\omega)} + \|s_3\|^2_{L^2(\omega)} \Big)^{\frac{1}{2}}. \tag{3.2}$$

The difference between the definition of $\mathbb{W}$ and $\mathbb{V}$ is that the regularity of the rotation variable r and the constraint is expressed in curvilinear variables instead of cartesian ones. Let us now show that the definitions are equivalent. Indeed if $r = (r_1^{\text{ca}}, r_2^{\text{ca}}, r_3^{\text{ca}})$ is the expression of the rotation in cartesian coordinates, then it can also be written as

$$r = \sum_{i=1}^{3} r_i a_i,$$

where $r_i, i = 1, 2, 3$ are its curvilinear coordinates. Then by the properties (2.3), we get

$$r_i = r \cdot a_i.$$

This simply means that $\mathbb{W}$ coincides with $\mathbb{V}$, and therefore the bilinear forms $\mathbf{a}$ and $\boldsymbol{a}_p$ are well defined (and continuous with respect to the norm (3.2)) on $\mathbb{W}$.

Before going, we want to emphasize that from now on for $(u, r) \in \mathbb{W}$, r_i always mean the curvilinear coordinates of r.

Lemma 3.1. — *The space $\mathbb{W}$ equipped with the norm (3.2) is a Hilbert space.*

Proof. — We remark that $\mathbb{W}$ is a closed subspace of

$$\mathbb{X} = \{(v, s = \sum_{i=1}^{3} s_i a_i) \in H^1\left(\omega, \mathbb{R}^3\right) \times (L^2(\omega))^3 \mid s_\alpha \in H^1(\omega),\, v|_{\Gamma_0} = s_\alpha|_{\Gamma_0} = 0\}, \tag{3.3}$$

equipped with the norm (3.2) because $\mathbb{W}$ is simply the kernel of the linear and continuous operator $\mathcal{Q}$ defined by

$$\mathcal{Q} : \mathbb{X} \longrightarrow L^2(\omega) \ : (v,s) \longmapsto s_3 - \tilde{\gamma}_{12}(v). \qquad \square$$

Then, the new variational formulation reads

$$\begin{cases} \text{Find } U=(u,r)\in\mathbb{W} \text{ such that} \\ \mathbf{a}(U,V)+\boldsymbol{a}_p(U,V)=\mathcal{L}(V), \quad \forall V=(v,s)\in\mathbb{W}. \end{cases} \tag{3.4}$$

Since the bilinear form $\mathbf{a}+\boldsymbol{a}_p$ and the form $\mathcal{L}$ are clearly continuous on $\mathbb{W}$, the well-posedness of problem (3.4) will be guaranteed if $\mathbf{a}+\boldsymbol{a}_p$ is coercive on $\mathbb{W}$. For that purpose, we need the following lemmata.

LEMMA 3.2. — *Suppose that $\varphi\in H^2(\omega,\mathbb{R}^3)$ and that $\varphi(\Gamma_0)$ is not included into a straight line. Let $V=(v,s)\in\mathbb{W}$. Then $\boldsymbol{a}(V,V)=0$ if and only if $V=0$.*

LEMMA 3.3. — *Under the assumptions of Lemma 3.2, the bilinear form $\boldsymbol{a}(\cdot,\cdot)$ is coercive on $\mathbb{W}$.*

The proofs are fully similar to those given in Lemmata 2 and 3 from [20] and are then omitted.

THEOREM 3.4. — *If $\|\nabla a_3\|_{L^\infty(\omega)}$ is small enough problem (3.4) admits a unique solution. Moreover, this solution satisfies*

$$\|U\|_{\mathbb{X}}\lesssim\|\mathcal{L}\|. \tag{3.5}$$

Proof. — If $\|\nabla a_3\|_{L^\infty(\omega)}$ is small enough, the bilinear form $\boldsymbol{a}(\cdot,\cdot)+\boldsymbol{a}_p(\cdot,\cdot)$ remains coercive on $\mathbb{W}$. Hence, the well-posedness of (3.4) follows from the Lax-Milgram lemma. □

3.1. **A penalized version of problem (3.4).** In this subsection, we present a penalized version for the prestressed model (3.4). The approach used here consists in adding a penalized term in (3.4) used to reformulate the original constrained problem as an unconstrained one, set on the variational space $\mathbb{X}$ defined by (3.3) and equipped with the norm (3.2).

For a real number $\epsilon\in(0,1)$, we consider the following variational problem:

$$\begin{cases} \text{Find } U_\epsilon=(u_\epsilon,r_\epsilon)\in\mathbb{X} \text{ such that} \\ \boldsymbol{a}(U_\epsilon,V)+\boldsymbol{a}_p(r_\epsilon,s)+\epsilon^{-1}b(U_\epsilon,V)=\mathcal{L}(V),\forall V=(v,s)\in\mathbb{X}. \end{cases} \tag{3.6}$$

For $W=(w,t),V=(v,s)\in\mathbb{X}$, the bilinear form $b(\cdot,\cdot)$ reads

$$b(W,V)=\int_\omega \mathcal{Q}(W)\mathcal{Q}(V)dx \tag{3.7}$$

where,

$$\mathcal{Q}(V)=s\cdot a_3-\tilde{\gamma}_{12}(v), \qquad \text{for any } V=(v,s)\in\mathbb{X}.$$

LEMMA 3.5. — *Under the assumption of Lemma 3.2, we have*

$$\boldsymbol{a}(V,V)+\frac{1}{\epsilon}b(V,V)\gtrsim\|V\|^2_{\mathbb{X}}, \quad \forall V=(v,s)\in\mathbb{X} \tag{3.8}$$

Proof. — Since $b(U,U)\geqslant 0$ for any $U\in\mathbb{X}$, the coercivity of $\boldsymbol{a}+\frac{1}{\epsilon}b$ on $\mathbb{X}$ (with a coercivity constant independent of ϵ) follows from Lemma 3.3. □

THEOREM 3.6. — *Under the assumptions of Lemma 3.2 and Theorem 3.4, the variational problem (3.6) has a unique solution U_ϵ in $\mathbb{X}$ that satisfies*

$$\|U_\epsilon\|_{\mathbb{X}}\lesssim\|f\|_\omega. \tag{3.9}$$

Proof. — The existence and uniqueness of U_ϵ directly follows from the Lax-Milgram Lemma. Concerning the bound (3.9), we simply take $V = U_\epsilon = (u_\epsilon, r_\epsilon)$ in (3.6) to get

$$\boldsymbol{a}(U_\epsilon, U_\epsilon) + \boldsymbol{a}_p(r_\epsilon, r_\epsilon) + \epsilon^{-1} b(U_\epsilon, U_\epsilon) = \mathcal{L}(U_\epsilon).$$

By the coerciveness property mentioned before and Cauchy-Schwarz's inequality we have

$$\|U_\epsilon\|_{\mathbb{X}}^2 \lesssim \|f\|_\omega \|u_\epsilon\|_\omega.$$

By Poincaré's type inequality we have

$$\|u_\epsilon\|_\omega \lesssim \|u_\epsilon\|_{H^1(\omega,\mathbb{R}^3)},$$

which leads to (3.9). □

Proposition 3.7. — *Let $U := (u, r)$ be the solution of the problem (3.4) and $U_\epsilon := (u^\epsilon, r^\epsilon)$ be the solution of problem (3.6) and let us assume that the assumptions of Theorem 3.6 are satisfied. Then*

$$\|r^\epsilon \cdot a_3 - \tilde{\gamma}_{12}(u^\epsilon)\|_{L^2(\omega)} \lesssim \sqrt{\epsilon} \tag{3.10}$$

$$\lim_{\epsilon \to 0} \|U_\epsilon - U\|_{\mathbb{X}} = 0. \tag{3.11}$$

Proof. — See [20, Theorem 8] . □

Remark 3.8. — Note that the results of Proposition 3.7 can be improved. First we recall that

$$b(U, V) = (\mathcal{Q}(U), \mathcal{Q}(V)) \tag{3.12}$$

with $\mathcal{Q}$ (defined above) is a bounded operator from $\mathbb{X}$ into $L^2(\omega)$. But this operator $\mathcal{Q}$ has a closed range $\mathcal{R}(\mathcal{Q})$, because $\mathcal{Q}$ is surjective, namely $\mathcal{R}(\mathcal{Q}) = L^2(\omega)$.

Indeed, let $f \in L^2(\omega)$ then there exists $g := (g_1, g_2) \in H_0^1(\omega) \times H_0^1(\omega)$ such that $\operatorname{div} g = f - \bar{f}$, with $\bar{f} = \frac{1}{|\omega|} \int_\omega f \, dx$. Then we set $v = \frac{1}{2}(g_2 a_1 - g_1 a_2)$ and $s = \big(\bar{f} + \frac{1}{2}(g_2 a_1 - g_1 a_2) \cdot (\partial_2 a_1 - \partial_1 a_2)\big) a_3$. We now readily check that $V = (v, s_3 a_3)$ belongs to $\mathbb{X}$ and satisfies $\mathcal{Q}(V) = f$. Then following [18, Corollary 2.4] (see also [19, Corollary 3.5]), the next estimates

$$\|r^\epsilon \cdot a_3 - \tilde{\gamma}_{12}(u^\epsilon)\|_{L^2(\omega)} \lesssim \epsilon \quad \text{and} \quad \|U_\epsilon - U\|_{\mathbb{X}} \lesssim \epsilon \tag{3.13}$$

hold.

4. Finite element approximation of the penalized problem.

As we have mentioned, the constrained problem (3.4) cannot be approximated by robust conforming methods for a general shell, hence we here propose an approximation of its penalized version (3.6). Let $(\mathcal{T}_h)_{h>0}$ be a regular affine family of triangulations which covers the domain ω. We introduce the finite dimensional space

$$\mathbb{X}_h = \{V_h = (v_h, s_h = \sum_{i=1}^{3} s_{ih} a_i) \in \mathbb{X} \,|\, v_{h|T} \in \mathbb{P}_k(T)^3, s_{ih} \in \mathbb{P}_k(T), \ \forall T \in \mathcal{T}_h, k \geqslant 1\}, \tag{4.1}$$

and consider the following discrete problem:

$$\begin{cases} \text{Find } U_h = (u_h, r_h) \in \mathbb{X}_h \text{ such that} \\ \boldsymbol{a}(U_h, V_h) + \boldsymbol{a}_p(U_h, V_h) + \epsilon^{-1} b(U_h, V_h) = \mathcal{L}(V_h), \forall V_h = (v_h, s_h) \in \mathbb{X}_h. \end{cases} \tag{4.2}$$

4.1. **A priori error analysis of the penalized problem.** In this subsection we derive a non robust a priori error analysis of the penalized problem (3.6).

PROPOSITION 4.1. — *Under the assumptions of Theorem 3.6, problem (4.2) has a unique solution $U_h \in \mathbb{X}_h$ that satisfies*

$$\|U_h\|_{\mathbb{X}} \lesssim \|f\|_\omega, \tag{4.3}$$

Furthermore if we assume that the solution U_ϵ of the problem (3.6) belongs to $[H^2(\omega;\mathbb{R}^3)] \times [H^2(\omega)]^2 \times [H^1(\omega)]$, then the following a priori error estimate holds

$$\|U_\epsilon - U_h\|_{\mathbb{X}} \lesssim \frac{h}{\epsilon}\Big(\|u^\epsilon\|_{H^2(\omega;\mathbb{R}^3)} + \sum_{\alpha=1,2} \|r_\alpha^\epsilon\|_{H^2(\omega)} + \|r_3^\epsilon\|_{H^1(\omega)}\Big). \tag{4.4}$$

Proof. — Since $\mathbb{X}_h \subset \mathbb{X}$, the existence of U_h and the a priori bound (4.3) follow from the that the bilinear form $\boldsymbol{a}+\boldsymbol{a}_p+\epsilon^{-1}b$ has an ellipticity constant that behaves like 1, see the proof of Theorem 3.6. On the other hand as its continuity constant behaves like $\frac{1}{\epsilon}$, Céa's lemma and standard interpolation error estimates directly yield (4.4). □

Remark 4.2. — It is clear that the estimate provided by Proposition 4.1, is not robust as ϵ goes to zero unless $h = o(\epsilon)$.

Remark 4.3. — Note that under the same assumptions of proposition 4.1, the estimate (4.4) can be improved to get a constant which behaves like $\frac{h}{\epsilon^{1/2}}$ instead of $\frac{h}{\epsilon}$, if one uses the following ϵ-dependent norm

$$\|V\|_\epsilon^2 := \boldsymbol{a}(V,V) + \boldsymbol{a}_p(s,s) + \varepsilon^{-1} b(V,V). \tag{4.5}$$

With this norm the analogous of (4.3) remains true since we have

$$\|U_h\|_\epsilon \lesssim \|f\|_\omega.$$

Further with the new norm (4.5), the continuity constant behaves like 1 and therefore, by Céa's lemma, we get

$$\|U_\epsilon - U_h\|_\epsilon \lesssim \inf_{V_h \in \mathbb{X}_h} \|U_\epsilon - V_h\|_\epsilon \lesssim \|U_\epsilon - \mathcal{C}_h U_\epsilon\|_\epsilon,$$

where $\mathcal{C}_h$ is for instance a Clément type interpolant.

Now for $\|U_\epsilon - \mathcal{C}_h U_\epsilon\|_{\mathbb{X}}$, the properties of Clément interpolant (see below) yield

$$\|U_\epsilon - U_h\|_{\mathbb{X}} \lesssim h\Big(\|u^\epsilon\|_{H^2(\omega;\mathbb{R}^3)} + \sum_{\alpha=1,2} \|r_\alpha^\epsilon\|_{H^2(\omega)} + \|r_3^\epsilon\|_{H^1(\omega)}\Big),$$

while for $\epsilon^{-\frac{1}{2}} b(U_\epsilon - \mathcal{C}_h U_\epsilon, U_\epsilon - \mathcal{C}_h U_\epsilon)^{\frac{1}{2}}$, we have,

$$\epsilon^{-\frac{1}{2}} \|\mathcal{Q}(U_\epsilon - \mathcal{C}_h U_\epsilon)\|_\omega \lesssim \epsilon^{-\frac{1}{2}} h \left(\|u^\epsilon\|_{H^2(\omega;\mathbb{R}^3)} + \|r_3^\epsilon\|_{H^1(\omega)}\right).$$

Hence, (4.4) becomes

$$\|U_\epsilon - U_h\|_\epsilon \lesssim \frac{h}{\epsilon^{\frac{1}{2}}}\Big(\|u^\epsilon\|_{H^2(\omega;\mathbb{R}^3)} + \sum_{\alpha=1,2} \|r_\alpha^\epsilon\|_{H^2(\omega)} + \|r_3^\epsilon\|_{H^1(\omega)}\Big). \tag{4.6}$$

4.2. **A priori error analysis of the mixed formulation of the penalized problem.** In order to obtain a uniform a priori estimate, we use a mixed formulation of the penalized problem (3.6) (as in [19, sec.4]). Let us first introduce the following new unknown

$$\psi_\epsilon := \frac{\mathcal{Q}(U_\epsilon)}{\epsilon},$$

and the functional space $\mathbb{M} = L^2(\omega)$. Then we rewrite the continuous penalized problem (3.6) as

$$\begin{cases} \text{Find } (U_\epsilon, \psi_\epsilon) \in \mathbb{X}\times\mathbb{M} \quad \text{such that} \\ \tilde{\boldsymbol{a}}(U_\epsilon, V) + (\psi_\epsilon, \mathcal{Q}(V)) = \mathcal{L}(V), \quad \forall V \in \mathbb{X}, \\ (\mathcal{Q}(U_\epsilon), \phi) - \epsilon(\psi_\epsilon, \phi) = 0, \qquad \forall \phi \in \mathbb{M}, \end{cases} \tag{4.7}$$

where $\tilde{\boldsymbol{a}}(\cdot,\cdot) = \boldsymbol{a}(\cdot,\cdot) + \boldsymbol{a}_p(\cdot,\cdot)$ and consider its discrete version:

$$\begin{cases} \text{Find } (U_h, \psi_h) \in \mathbb{X}_h\times\mathbb{M}_h \quad \text{such that} \\ \tilde{\boldsymbol{a}}(U_h, V_h) + (\psi_h, \mathcal{Q}(V_h)) = \mathcal{L}(V_h), \quad \forall V_h \in \mathbb{X}_h, \\ (\mathcal{Q}(U_h), \phi_h) - \epsilon(\psi_h, \phi_h) = 0, \qquad \forall \phi_h \in \mathbb{M}_h, \end{cases} \tag{4.8}$$

where

$$\mathbb{M}_h = \{\phi_h \in \mathbb{M} \mid \phi_{h|T} \in \mathbb{P}_k(T), \forall T \in \mathcal{T}_h, k \geqslant 0\}. \tag{4.9}$$

THEOREM 4.4. — *Let $(U_\epsilon, \psi_\epsilon)$ be the solution of (4.7) and let (U_h, ψ_h) be the solution of problem (4.8). Then we have the following error estimate*

$$\|U_\epsilon - U_h\|_{\mathbb{X}} + \sqrt{\epsilon}\|\psi_\epsilon - \psi_h\|_{\mathbb{M}} \lesssim \frac{1}{\epsilon^{1/2}} \inf_{W_h\in\mathbb{X}_h} \|U_\epsilon - W_h\|_h + \inf_{\varphi_h\in\mathbb{M}_h} \|\psi_\epsilon - \varphi_h\|_{\mathbb{M}}. \tag{4.10}$$

Proof. — Let $\tilde{U} \in \mathbb{X}_h$, and $\tilde{\psi} \in \mathbb{M}_h$. Then

$$\tilde{\boldsymbol{a}}(U_h - \tilde{U}, V_h) + (\mathcal{Q}(V_h), \psi_h - \tilde{\psi}) = \tilde{\boldsymbol{a}}(U_\epsilon - \tilde{U}, V_h) + (\mathcal{Q}(V_h), \psi_\epsilon - \tilde{\psi}), \quad \forall V_h \in \mathbb{X}_h, \tag{4.11}$$

$$(\mathcal{Q}(U_h - \tilde{U}), \phi_h) - \epsilon(\psi_h - \tilde{\psi}, \phi_h) = (\mathcal{Q}(U_\epsilon - \tilde{U}), \phi_h) - \epsilon(\psi_\epsilon - \tilde{\psi}, \phi_h), \quad \forall \phi_h \in \mathbb{M}_h. \tag{4.12}$$

By taking $V_h = U_h - \tilde{U}$, and $\phi_h = \psi_h - \tilde{\psi}$ and subtracting (4.12) from (4.11), we get

$$\begin{aligned} \|U_h - \tilde{U}\|_{\mathbb{X}}^2 + \epsilon\|\psi_h - \tilde{\psi}\|_{\mathbb{M}}^2 &\lesssim \tilde{\boldsymbol{a}}(U_\epsilon - \tilde{U}, U_h - \tilde{U}) + (\mathcal{Q}(U_h - \tilde{U}), \psi_\epsilon - \tilde{\psi}) \\ &\quad - (\mathcal{Q}(U_\epsilon - \tilde{U}), \psi_h - \tilde{\psi}) + \epsilon(\psi_\epsilon - \tilde{\psi}, \psi_h - \tilde{\psi}) \\ &\lesssim \|U_\epsilon - \tilde{U}\|_{\mathbb{X}}\|U_h - \tilde{U}\|_{\mathbb{X}} + \|U_h - \tilde{U}\|_{\mathbb{X}}\|\psi_\epsilon - \tilde{\psi}\|_{\mathbb{M}} \\ &\quad + \|U_\epsilon - \tilde{U}\|_{\mathbb{X}}\|\psi_h - \tilde{\psi}\|_{\mathbb{M}} + \epsilon\|\psi_\epsilon - \tilde{\psi}\|_{\mathbb{M}}\|\psi_h - \tilde{\psi}\|_{\mathbb{X}}. \end{aligned} \tag{4.13}$$

According to Young's inequality we deduce that

$$\begin{aligned} \|U_h - \tilde{U}\|_{\mathbb{X}} + \sqrt{\epsilon}\|\psi_h - \tilde{\psi}\|_{\mathbb{M}} &\lesssim \frac{1}{\sqrt{\epsilon}}\|U_\epsilon - \tilde{U}\|_{\mathbb{X}} + \|\psi_\epsilon - \tilde{\psi}\|_{\mathbb{M}} + \sqrt{\epsilon}\|\psi_\epsilon - \tilde{\psi}\|_{\mathbb{M}} \\ &\lesssim \frac{1}{\sqrt{\epsilon}}\|U_\epsilon - \tilde{U}\|_{\mathbb{X}} + \|\psi_\epsilon - \tilde{\psi}\|_{\mathbb{M}}. \end{aligned}$$

□

Remark 4.5. — Again, the estimate provided by Theorem 4.4 is not uniform in ϵ.

In order to get a uniform estimate in ϵ we first need to the following uniform discrete inf-sup condition.

LEMMA 4.6. — *For $\mathbb{X}_h$ defined in (4.1) and $\mathbb{M}_h$ given by (4.9), we have the following inf-sup condition:*

$$\forall \phi_h \in \mathbb{M}_h, \quad \sup_{V_h \in \mathbb{X}_h} \frac{(\mathcal{Q}(V_h), \phi_h)}{\|V_h\|_{\mathbb{X}}} \gtrsim \|\phi_h\|_{\mathbb{M}}. \tag{4.14}$$

Proof. — Let $\phi_h \in \mathbb{M}_h$, then by choosing $V_h = (v_h, s_h = \sum_i s_{hi} a_i)$ with $v_h = 0, s_{h\alpha} = 0, \alpha = 1, 2$ and $s_{h3} = \phi_h$ we get

$$\frac{(\mathcal{Q}(V_h), \phi_h)}{\|V_h\|_{\mathbb{X}}} \geqslant \|\phi_h\|_{\mathbb{M}}. \qquad \square$$

THEOREM 4.7. — *Let $(U_\epsilon, \psi_\epsilon)$ be the solution of (4.7) and let (U_h, ψ_h) be the solution of problem (4.8). Then for ϵ small enough, we have the following error estimate*

$$\|U_\epsilon - U_h\|_{\mathbb{X}} + \|\psi_\epsilon - \psi_h\|_{\mathbb{M}} \lesssim \inf_{W_h \in \mathbb{X}_h} \|U_\epsilon - W_h\|_{\mathbb{X}} + \inf_{\varphi_h \in \mathbb{M}_h} \|\psi_\epsilon - \varphi_h\|_{\mathbb{M}}. \tag{4.15}$$

Proof. — We use the same choice of test functions as in the proof of Theorem 4.4, but treating the term

$$(\mathcal{Q}(U_\epsilon - \tilde{U}), \psi_h - \tilde{\psi})$$

differently. Indeed, form (4.11) and (4.14) we have

$$\|\psi_h - \tilde{\psi}\|_{\mathbb{M}} \lesssim \|U_h - \tilde{U}\|_{\mathbb{X}} + \|U_\epsilon - \tilde{U}\|_{\mathbb{X}} + \|\psi_\epsilon - \tilde{\psi}\|_{\mathbb{M}}.$$

Exploiting this estimate in (4.13), we get

$$\begin{aligned}
\|U_h - \tilde{U}\|_{\mathbb{X}}^2 + \|\psi_h - \tilde{\psi}\|_{\mathbb{M}}^2 + \epsilon \|\psi_h - \tilde{\psi}\|_{\mathbb{M}}^2 \\
&\lesssim \|U_\epsilon - \tilde{U}\|_{\mathbb{X}} \|U_h - \tilde{U}\|_{\mathbb{X}} + \|U_h - \tilde{U}\|_{\mathbb{X}} \|\psi_\epsilon - \tilde{\psi}\|_{\mathbb{M}} \\
&\quad + \|U_\epsilon - \tilde{U}\|_{\mathbb{X}} \left(\|U_h - \tilde{U}\|_{\mathbb{X}} + \|U_\epsilon - \tilde{U}\|_{\mathbb{X}} + \|\psi_\epsilon - \tilde{\psi}\|_{\mathbb{M}} \right) \\
&\quad + \epsilon \|\psi_\epsilon - \tilde{\psi}\|_{\mathbb{M}} \left(\|U_h - \tilde{U}\|_{\mathbb{X}} + \|U_\epsilon - \tilde{U}\|_{\mathbb{X}} + \|\psi_\epsilon - \tilde{\psi}\| \right) \\
&\quad + \left(\|U_h - \tilde{U}\|_X + \|U_\epsilon - \tilde{U}\|_X + \|\psi_\epsilon - \tilde{\psi}\|_{\mathbb{M}} \right)^2.
\end{aligned}$$

Then using Young's inequality we obtain the desired estimate. $\square$

COROLLARY 4.8. — *Let $(U_\epsilon, \psi_\epsilon)$ be the solution of (4.7) and let (U_h, ψ_h) be the solution of problem (4.8). Assume that $U_\epsilon = (u_\epsilon, r_\epsilon)$ satisfies $u_\epsilon \in H^2(\omega, \mathbb{R}^3)$, $r_\epsilon \cdot a_\alpha \in H^2(\omega)$ and $r_\epsilon \cdot a_3 \in H^1(\omega)$. Then for ϵ small enough, it holds*

$$\|U_\epsilon - U_h\|_{\mathbb{X}} + \|\psi_\epsilon - \psi_h\|_{\mathbb{M}} \lesssim h(\|u_\epsilon\|_{2,\omega} + \sum_{\alpha=1,2} \|r_\epsilon \cdot a_\alpha\|_{2,\omega} + \|r_\epsilon \cdot a_3\|_{1,\omega}). \tag{4.16}$$

Proof. — Using (4.7), we find

$$\tilde{\boldsymbol{a}}(U_\epsilon, V) + (\psi_\epsilon, \mathcal{Q}(V)) - (\mathcal{Q}(U_\epsilon), \phi) + \epsilon(\psi_\epsilon, \phi) = \mathcal{L}(V), \quad \forall V \in \mathbb{X}, \forall \phi \in \mathbb{M}. \tag{4.17}$$

Take $\phi = 0$ and $V = (v, s = \sum_i s_i a_i)$, with $v = 0$, $s_\alpha = 0, \alpha = 1, 2$ and $s_3 \in L^2(\omega)$ in (4.17) to get, for all $s_3 \in L^2(\omega)$,

$$(\psi_\epsilon, s_3) = -t a_t((u^\epsilon, r^\epsilon), (0, 0, s_3)) - \frac{t^3}{12} a_f(r^\epsilon, (0, 0, s_3)) - \frac{t^3}{12} a_p(r^\epsilon, (0, 0, s_3)).$$

Then the regularity of U_ϵ and the form of the bilinear form $\tilde{\boldsymbol{a}}(\cdot,\cdot)$ amount to write

$$(\psi_\epsilon, s_3) = (\tilde{f}, s_3), \quad \forall s_3 \in L^2(\omega),$$

with $\tilde{f} \in H^1(\omega)$ which implies that $\psi_\epsilon = \tilde{f}$ belongs to $H^1(\omega)$ with the estimate

$$\|\psi_\epsilon\|_{H^1(\omega)} \lesssim \|u_\epsilon\|_{2,\omega} + \sum_{\alpha=1,2} \|r_\epsilon \cdot a_\alpha\|_{2,\omega} + \|r_\epsilon \cdot a_3\|_{1,\omega}.$$

Taking in (4.15), $(W_h, \varphi_h) = \mathcal{C}_h(U_\epsilon, \psi_\epsilon)$, where $\mathcal{C}_h$ is the Clément interpolation operator and using a standard interpolation estimate (see below), the conclusion follows by using the previous estimates in (4.15). □

5. The strong formulation (PDEs form) of the penalized problem.

Usually, a posteriori estimator is computed by element-wise integration by parts starting from the classical formulation or the PDE form of the problem. Hence in this section we give the strong formulation of problem (3.6). As before we use the covariant representation of the unknowns, i.e, in the following $s = \sum_{i=1}^3 s_i a_i$, which makes it easier to obtain the PDEs form. We use also the following notation $\hat{s} = (s \cdot a_1, s \cdot a_2)^T$. We recall that the elasticity coefficients in local coordinates are given by

$$a^{\alpha\beta\rho\sigma} = 2\mu(a^{\alpha\rho}a^{\beta\sigma} + a^{\alpha\sigma}a^{\beta\rho}) + \frac{4\lambda\mu}{\lambda+2\mu}a^{\alpha\beta}a^{\rho\sigma}.$$

Let us then denote by $\mathbb{A}$ the elasticity tensor whose components are $a^{\alpha\beta\rho\sigma} \in L^\infty(\omega)$ and define

$$T(u) := t\ \mathbb{A}\gamma(u),$$

that is a 2×2 matrix with coefficients in $\mathbb{R}^3$. Note that the property (2.4) implies that

$$\mathbb{A}M : N = 4\mu M : N + \frac{4\lambda\mu}{\lambda+2\mu}\mathrm{tr}M\mathrm{tr}N, \tag{5.1}$$

for all symmetric 2×2 matrices M and N. According to (2.9), using these definitions, and this last property, we have

$$a_m(u,v) = \int_\omega \mathbb{A}\gamma(u) : \gamma(v)\ dx, \tag{5.2}$$

and hence

$$\begin{aligned} a_m(u,v) &= \int_\omega T^{\alpha\beta}(u) \cdot \gamma_{\alpha\beta}(v) dx \\ &= \int_\omega T^{\alpha\beta}(u)\partial_\alpha v \cdot a_\beta\, dx. \end{aligned}$$

Hence if u is smooth enough, by Green's formula we have

$$\begin{aligned} t a_m(u,v) &= -\int_\omega \partial_\alpha(T^{\alpha\beta}(u)a_\beta) \cdot v\, dx + \int_{\partial\omega} T^{\alpha\beta}(u)n_\alpha a_\beta \cdot v\, d\sigma(x) \\ &= -\int_\omega \mathrm{Div}\,(T(u)A) \cdot v\, dx + \int_{\Gamma_1} nT(u)A \cdot v\, d\sigma(x), \end{aligned} \tag{5.3}$$

where $d\sigma$ is the surface measure on the boundary $\partial\omega$ of ω, $n = (n_1, n_2)$ is the unit outward normal vector (written in line) along $\partial\omega$, $A = (a_1, a_2)^\top$ is 2×3

matrix and here and below for a 2×3 matrix valued function $M = (m_{\alpha i})_{\alpha,i}$, $\mathrm{Div}\, M = (\sum_\alpha \partial_\alpha m_{\alpha i})_{i=1,2,3}$ (hence is a column vector valued function).

Let us now consider the contribution of the bilinear form $a_t(\cdot,\cdot)$. For that purpose, recalling that $\nabla\varphi = (a_1, a_2), a_1 \times a_3 = -a_2$ and $a_2 \times a_3 = a_1$, we remark that

$$\begin{aligned} a_3^\top(\nabla v - s \times \nabla\varphi) &= a_3^\top(\partial_1 v, \partial_2 v) - (a_3^\top \cdot (s \times a_1), a_3^\top \cdot (s \times a_2)) \\ &= (a_3^\top \cdot \partial_1 v + s \cdot (a_1 \times a_3), \quad a_3^\top \cdot \partial_2 v + s \cdot (a_2 \times a_3)) \\ &= (a_3^\top \cdot \partial_1 v + s_2, \quad a_3^\top \cdot \partial_2 v - s_1). \end{aligned}$$

Hence if we set

$$J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

we have

$$a_3^\top(\nabla v - s \times \nabla\varphi) = a_3^\top \nabla v + \hat{s}^\top J^\top.$$

This expression in (2.11) yields

$$a_s((u,r),(v,s)) = \mu \int_\omega (a_3^\top \nabla v + \hat{s}^\top J^\top)((\nabla u)^\top a_3 + J\hat{r})\, dx. \tag{5.4}$$

We therefore introduce the 2×1 vector valued function

$$S(u,r) := t\ \mu((\nabla u)^\top a_3 + J\hat{r}).$$

Using this notation and (5.4), we get

$$\begin{aligned} t a_s((u,r),(v,s)) &= \int_\omega (a_3^\top \nabla v + \hat{s}^\top J^\top) S(u,r)\, dx \\ &= \int_\omega \big(a_3 \cdot \partial_\alpha v S^\alpha(u,r) + \hat{s}^\top J^\top S(u,r)\big)\, dx, \end{aligned}$$

where $S^\alpha(u,r)$ are the two components of $S(u,r)$. As before if $S^\alpha(u,r)$ is smooth enough, by Green's formula we will obtain

$$\begin{aligned} t a_s((u,r),(v,s)) &= \int_\omega (-\partial_\alpha(S^\alpha(u,r)a_3) \cdot v dx + \int_{\Gamma_1} S^\alpha(u,r) n_\alpha a_3 \cdot v\, d\sigma(x) \\ &\quad + \int_\omega J^\top S(u,r) \cdot \hat{s}\, dx \\ &= -\int_\omega \mathrm{Div}\,(S(u,r)a_3) \cdot v\, dx + \int_{\Gamma_1} n S(u,r) a_3 \cdot v\, d\sigma(x) \\ &\quad + \int_\omega J^\top S(u,r) \cdot \hat{s}\, dx. \end{aligned} \tag{5.5}$$

Next we consider the bilinear form $a_f(r,s)$. Due to (2.6) and the definition of the tensor $\mathbb{A}$, we may write

$$a_f(r,s) = \frac{1}{2} \int_\omega \mathbb{A}\Pi(r) \colon \Pi(s)\, dx. \tag{5.6}$$

Hence if we set

$$M(r) := \frac{t^3}{24} \mathbb{A}\Pi(r) = \frac{t^3}{24} (a^{\alpha\beta\rho\sigma} \Pi_{\rho\sigma}(r))_{\alpha,\beta},$$

we obtain

$$\frac{t^3}{12}a_f(r,s)=\int_\omega M(r)\colon \Pi(s)\ dx. \tag{5.7}$$

We now need to transform $\Pi(s)$. For that purpose, recalling (2.6), by setting

$$\bar{s}=\begin{pmatrix} s_2\\ -s_1\end{pmatrix}=J\begin{pmatrix} s_1\\ s_2\end{pmatrix},$$

using the property (see [10, Theorem 2.6-1])

$$\partial_\alpha s\cdot a_\beta=\partial_\alpha s_\beta-\Gamma^\rho_{\alpha\beta}s_\rho-b_{\alpha\beta}s_3, \tag{5.8}$$

we get

$$\Pi(s)=e(\bar{s})-\bar{\ell}(s), \tag{5.9}$$

where $e(\cdot)$ is the usual deformation tensor of the two dimensional elasticity, i.e

$$e\begin{pmatrix} w_1\\ w_2\end{pmatrix}=\begin{pmatrix} \partial_1 w_1 & \frac{1}{2}(\partial_1 w_2+\partial_2 w_1)\\ \frac{1}{2}(\partial_1 w_2+\partial_2 w_1) & \partial_2 w_2\end{pmatrix},$$

and $\bar{\ell}(\cdot)$ is an operator of order zero which acts on any three dimensional vector field s as follows

$$\begin{aligned}\bar{\ell}(s)=\bar{\Gamma}^\rho s_\rho+\bar{B}s_3&=\begin{pmatrix} \Gamma^\rho_{12} & \frac{1}{2}(\Gamma^\rho_{22}-\Gamma^\rho_{11})\\ \frac{1}{2}(\Gamma^\rho_{22}-\Gamma^\rho_{11}) & -\Gamma^\rho_{21}\end{pmatrix}s_\rho\\ &\quad+\begin{pmatrix} b_{12} & \frac{1}{2}(b_{22}-b_{11})\\ \frac{1}{2}(b_{22}-b_{11}) & -b_{12}\end{pmatrix}s_3.\end{aligned}$$

The splitting (5.9) into (5.6) and (5.7) yields

$$a_f(r,s)=\frac{1}{2}\int_\omega \mathbb{A}(e(\bar{r})-\bar{\ell}(r))\colon (e(\bar{s})-\bar{\ell}(s))\ dx, \tag{5.10}$$

and

$$\frac{t^3}{12}a_f(r,s)=\int_\omega M(r)\colon (e(\bar{s})-\bar{\ell}(s))\ dx,$$

and if $M(r)$ is smooth enough by Green's formula we obtain

$$\begin{aligned}\frac{t^3}{12}a_f(r,s)&=-\int_\omega \mathrm{Div}\, M(r)\cdot\bar{s}\ dx+\int_{\partial\omega} nM(r)\bar{s}\,d\sigma(x)-\int_\omega M(r)\colon\bar{\ell}(s)\ dx\\ &=-\int_\omega J^T\ \mathrm{Div}\, M(r)\cdot\hat{s}\ dx+\int_{\partial\omega} J^TM(r)n^\top\cdot\hat{s}\,d\sigma(x)-\int_\omega M(r)\colon\bar{\ell}(s)\ dx.\end{aligned}$$

Finally using the above expression of $\bar{\ell}(s)$

$$\begin{aligned}\frac{t^3}{12}a_f(U,V)&=-\int_\omega J^\top Div(M(r))\cdot\hat{s}dx+\int_{\Gamma_1} J^TM(r)n^\top\cdot\hat{s}\,d\sigma(x)\\ &\quad-\int_\omega\left(\begin{pmatrix} M(r)\colon\bar{\Gamma}^1\\ M(r)\colon\bar{\Gamma}^2\end{pmatrix}\cdot\hat{s}+\left(\bar{B}\colon M(r)\right)s_3\right)dx.\end{aligned} \tag{5.11}$$

Now we give the contribution of the prestressed term $\boldsymbol{a}_p(\cdot,\cdot)$. First as II_0 and $\tau(r,s)$ are symmetric, we directly check that

$$\frac{1}{2}\mathrm{tr}((II_0+II_0^t)\tau(r,s))=\mathrm{tr}(II_0\tau(r,s))=II_0\colon\tau(r,s),$$

furthermore by (2.13), we have

$$\mathrm{tr}\tau(r,s) = (s\cdot a_3)\mathrm{tr}\ \theta(r) + (r\cdot a_3)\mathrm{tr}\ \theta(s).$$

Hence we have

$$\begin{aligned}2\mu\mathrm{tr}\left((II_0 + II_0^t)\tau(r,s)\right) + \frac{4\lambda\mu}{2\mu+\lambda}\mathrm{tr}II_0\mathrm{tr}\tau(r,s) = \\ (s\cdot a_3)\Big(4\mu II_0 \colon \theta(r) + \frac{4\lambda\mu}{\lambda+2\mu}\mathrm{tr}\ II_0\mathrm{tr}\ \theta(r)\Big) \\ + (r\cdot a_3)\Big(4\mu II_0 \colon \theta(s) + \frac{4\lambda\mu}{\lambda+2\mu}\mathrm{tr}\ II_0\mathrm{tr}\ \theta(s)\Big). \\ = (s\cdot a_3)\mathbb{A}II_0 \colon \theta(r) + (r\cdot a_3)\mathbb{A}II_0 \colon \theta(s),\end{aligned}$$

this last identity following from (5.1). Accordingly, $\boldsymbol{a}_p(r,s)$ takes the equivalent form

$$\boldsymbol{a}_p(r,s) = \frac{t^3}{12}\int_\omega (s_3\mathbb{A}II_0 \colon \theta(r) + r_3\mathbb{A}II_0 \colon \theta(s))\ dx. \tag{5.12}$$

Now setting

$$\begin{aligned}P(r) &= \frac{t^3}{12}\mathbb{A}II_0 r_3, \\ \kappa(r) &= \frac{t^3}{12}(II_0 \colon \mathbb{A}\theta(r)),\end{aligned}$$

we deduce that

$$\boldsymbol{a}_p(r,s) = \int_\omega P(r) \colon \theta(s)\ dx + \int_\omega \kappa(r)s_3\ dx. \tag{5.13}$$

At this stage we need to transform the matrix $\theta(s)$. First using (5.8), we check that

$$\begin{aligned}-\gamma_{11}(s) &= -\partial_1 s_1 + \Gamma_{11}^\rho s_\rho + b_{11}s_3, \\ \tilde{\gamma}_{12}(s) &= \frac{\partial_1 s_2 - \partial_2 s_1}{2}, \\ \gamma_{22}(s) &= \partial_2 s_2 - \Gamma_{22}^\rho s_\rho - b_{22}s_3.\end{aligned}$$

Hence introducing $\tilde{s} = \tilde{J}\hat{s}$ with

$$\tilde{J} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

and the operator of order zero $\tilde{\ell}$ which acts on any three dimensional vector field s as follows

$$\tilde{\ell}(s) = \tilde{\Gamma}^\rho s_\rho + \tilde{B}s_3 = \begin{pmatrix} \Gamma_{11}^\rho & 0 \\ 0 & -\Gamma_{22}^\rho \end{pmatrix} s_\rho + \begin{pmatrix} b_{11} & 0 \\ 0 & -b_{22} \end{pmatrix} s_3,$$

we obtain

$$\theta(s) = \frac{1}{2}\left(e(\tilde{s}) + \tilde{\ell}(s)\right). \tag{5.14}$$

This expression in (5.13) yields

$$\boldsymbol{a}_p(r,s) = \frac{1}{2}\int_\omega P(r) \colon (e(\tilde{s}) + \tilde{\ell}(s))\ dx + \int_\omega \kappa(r)s_3\ dx.$$

Again if r is smooth enough, we can apply Green's formula and find

$$\begin{aligned}\boldsymbol{a}_p(r,s) = &-\int_\omega \frac{1}{2}\tilde{J}\ \mathrm{Div}\,(P(r))\cdot\hat{s}\ dx + \int_{\Gamma_1}\frac{1}{2}\tilde{J}P(r)n^\top\cdot\hat{s}\,d\sigma(x)\\ &+\int_\omega(\kappa(r)+\frac{1}{2}\tilde{B}\colon P(r))\ s_3\ dx + \int_\omega\frac{1}{2}\begin{pmatrix} P(u)\colon\tilde{\Gamma}^1\\ P(u)\colon\tilde{\Gamma}^2\end{pmatrix}\cdot\hat{s}\ dx. \end{aligned}\tag{5.15}$$

For the bilinear form $b(\cdot,\cdot)$, as $\tilde{\gamma}_{12}(v)=\frac{1}{2}(\partial_1 v\cdot\partial_2\varphi-\partial_2 v\cdot\partial_1\varphi)$, if $\mathcal{Q}(U)$ is sufficiently regular we find

$$\begin{aligned}\frac{1}{\epsilon}b(U,V) = \frac{1}{\epsilon}\int_\omega\mathcal{Q}(U)(s_3-\tilde{\gamma}_{12}(v))\,dx = \frac{1}{2\epsilon}\int_\omega\mathrm{Div}\,(\mathcal{Q}(U)JA)\cdot v\,dx\\ -\frac{1}{2\epsilon}\int_{\Gamma_1}\mathcal{Q}(U)A^\top Jn^\top\cdot v\,d\sigma(x)+\frac{1}{\epsilon}\int_\omega\mathcal{Q}(U)s_3\,dx.\end{aligned}\tag{5.16}$$

Using the identities (5.3), (5.5), (5.11), (5.15), (5.16), we see that the solution $U_\epsilon=(u_\epsilon,r_\epsilon)\in\mathbb{X}$ of problem (3.6) satisfies

$$\left.\begin{array}{lll}
-\mathrm{Div}\,(T(u_\epsilon)A)-\mathrm{Div}\,(S(U_\epsilon)a_3)+\frac{1}{2\epsilon}\mathrm{Div}\,(\mathcal{Q}(U_\epsilon)JA) &= f & \text{in } \omega,\\
-J^\top\mathrm{Div}\,M(r_\epsilon)-\begin{pmatrix} M(r_\epsilon):\bar{\Gamma}^1\\ M(r_\epsilon):\bar{\Gamma}^2\end{pmatrix}+J^\top S(U_\epsilon) & & \\
-\frac{1}{2}\tilde{J}\mathrm{Div}\,(P(r_\epsilon))+\frac{1}{2}\begin{pmatrix} P(u_\epsilon):\tilde{\Gamma}^1\\ P(u_\epsilon):\tilde{\Gamma}^2\end{pmatrix} &= 0 & \text{in } \omega,\\
-\big(\bar{B}\colon M(r_\epsilon)\big)+\kappa(r_\epsilon)+\frac{1}{2}\tilde{B}\colon P(r_\epsilon)+\frac{1}{\epsilon}\mathcal{Q}(U_\epsilon) &= 0 & \text{in } \omega,\\
u_\epsilon = r_\alpha^\epsilon &= 0 & \text{on } \Gamma_0,\\
nT(u_\epsilon)A+nS(U_\epsilon)a_3-\frac{1}{2\epsilon}\mathcal{Q}(U_\epsilon)A^\top Jn^\top &= 0 & \text{on } \Gamma_1,\\
\frac{1}{2}\tilde{J}P(r_\epsilon)n^\top+J^\top M(r_\epsilon)n^\top &= 0 & \text{on } \Gamma_1.
\end{array}\right\}\tag{5.17}$$

Note that by taking test functions in $\mathcal{D}(\omega)^6$ in (5.3), (5.5), (5.11), (5.15), (5.16), we find that the three first identities are valid in the distributional sense. This means that the left-hand side of this identities belongs to $L^2(\omega)^3, L^2(\omega)^2$, and $L^2(\omega)$ respectively. Using the following Green's formula

$$\int_\omega(J\cdot\nabla\phi+\phi\mathrm{div}\ \mathrm{J})\ dx = \langle J\cdot n^\top,\phi\rangle_{H^{-\frac{1}{2}}(\partial\omega)-H^{\frac{1}{2}}(\partial\omega)},$$

valid for $\phi\in H^1(\Omega)$ and $J\in H(\mathrm{div},\omega)$ (see [13, (I.2.17)]), we deduce that the boundary conditions in Γ_1 holds in $(\tilde{H}^{\frac{1}{2}}(\Gamma_1)^3)'$ and $(\tilde{H}^{\frac{1}{2}}(\Gamma_1)^2)'$ respectively.

6. A posteriori error estimate of the penalized problem.

As we have mentioned in the introduction of this paper, we focus only on residual a posteriori estimate. For the problem (3.6), the residual $\mathcal{R}_{U_h}(\cdot)$ is then defined as follows

$$\begin{aligned}\mathcal{R}_{U_h} &= \boldsymbol{a}(U^\epsilon-U_h,V)+\boldsymbol{a}_p(U^\epsilon-U_h,V)+\epsilon^{-1}b(U^\epsilon-U_h,V)\\ &= \mathcal{L}(V-V_h)-\boldsymbol{a}(U_h,V-V_h)-\boldsymbol{a}_p(U_h,V-V_h)-\epsilon^{-1}b(U_hV-V_h),\end{aligned}\tag{6.1}$$

for an arbitrary $V_h\in\mathbb{X}_h$. From the fact that $\boldsymbol{a}(\cdot,\cdot)+\boldsymbol{a}_p(\cdot,\cdot)+\epsilon^{-1}b(\cdot,\cdot)$ is coercive with a coercivity constant equivalent to 1, we infer that

$$\|U^\epsilon-U_h\|_{\mathbb{X}}\lesssim\|\mathcal{R}_{U_h}\|_{\mathbb{X}'}.$$

We first observe that the bilinear forms $\boldsymbol{a}(\cdot,\cdot), \boldsymbol{a}_p(\cdot,\cdot)$ and $b(\cdot,\cdot)$ have variable coefficients. In such a case, in order to construct error indicators we need to approximate the data and the coefficients by piecewise polynomials, see [2].

6.1. **Approximation of the data and coefficients.** We introduce the approximation spaces $\tilde{\mathbb{M}}_h^{(\ell)}$, with $\ell \in \mathbb{N}$ and $\mathbb{Z}_h$ as follows

$$\begin{aligned}\tilde{\mathbb{M}}_h^{(\ell)} &= \left\{\chi_h \in L^2(\omega); \forall T \in \mathcal{T}_h, \chi_h|_T \in \mathbb{P}_\ell(T)\right\},\\ \mathbb{Z}_h &= \left\{g_h \in L^2(\omega)^3; \forall T \in \mathcal{T}_h, g_h|_T \in \mathbb{P}_0(T)^3\right\},\end{aligned}$$

and consider an approximation f_h of f in $\mathbb{Z}_h$ and an approximation $b_{\alpha\beta}^h$ of the coefficient $b_{\alpha\beta}$ in $\tilde{\mathbb{M}}_h^{(1)}$ (as $b_{12} = b_{21}$, we assume that $b_{12}^h = b_{21}^h$). Similarly, we consider approximations a_k^h of the vectors a_k and $d_{\alpha\beta}^h$ of $\partial_\alpha a_\beta$ in $(\tilde{\mathbb{M}}_h^{(2)})^3$ and $(\tilde{\mathbb{M}}_h^{(1)})^3$ respectively. Obviously we assume that these approximated coefficients are uniformly bounded (with respect to the L^∞-norm) in h. We introduce the approximations $\boldsymbol{a}_h(\cdot,\cdot), \boldsymbol{a}_p^h(\cdot,\cdot)$ and $b_h(\cdot,\cdot)$ of the bilinear forms $\boldsymbol{a}(\cdot,\cdot), \boldsymbol{a}_p(\cdot,\cdot)$ and $b(\cdot,\cdot)$ respectively where a_i, $\partial_\alpha a_\beta$, and $b_{\alpha\beta}$ are replaced by their approximations. More precisely, for $U = (u, \sum_i r_i a_i) \in \mathbb{X}$, we set (compare with (2.5), (5.14), and (5.9))

$$\begin{aligned}\gamma_{\alpha\beta}^h(u) &= \frac{1}{2}\left(\partial_\alpha u \cdot a_\beta^h + \partial_\beta u \cdot a_\alpha^h\right),\\ \tilde{\gamma}_{12}^h(u) &= \frac{1}{2}\left(\partial_1 u \cdot a_2^h - \partial_2 u \cdot a_1^h\right),\\ \Pi^h(s) &= e(\bar{s}) - \bar{\ell}^h(s),\\ \theta^h(s) &= \frac{1}{2}\left(e(\tilde{s}) + \tilde{\ell}^h(s)\right),\\ II_0^h &= -\begin{pmatrix} b_{11}^h & b_{12}^h \\ b_{12}^h & b_{22}^h \end{pmatrix},\\ \mathcal{Q}^h(U) &= r_3 - \tilde{\gamma}_{12}^h(u),\end{aligned}$$

where $\bar{\ell}^h(s)$ and $\tilde{\ell}^h(s)$ are defined as $\bar{\ell}(s)$ and $\tilde{\ell}(s)$, the coefficients $b_{\alpha\beta}$ and $\Gamma_{\alpha\beta}^\rho$ being replaced by $b_{\alpha\beta}^h$ and $a_\rho^h \cdot d_{\alpha\beta}^h$ respectively. Then we set (compare with (5.2), (5.10), (5.4) and (5.12))

$$\begin{aligned}a_m^h(u,v) &= \int_\omega \mathbb{A}\gamma^h(u) : \gamma^h(v)\, dx,\\ a_f^h(r,s) &= \frac{1}{2}\int_\omega \mathbb{A}(e(\bar{r}) - \bar{\ell}^h(r)) : (e(\bar{s}) - \bar{\ell}^h(s))\, dx,\\ a_s^h((u,r),(v,s)) &= \mu\int_\omega ((a_3^h)^\top \nabla v + \hat{s}^\top J^\top)((\nabla u)^\top a_3^h + J\hat{r})\, dx,\\ \boldsymbol{a}_p^h(r,s) &= \frac{t^3}{12}\int_\omega (s_3 \mathbb{A} II_0^h : \theta^h(r) + r_3 \mathbb{A} II_0^h : \theta^h(s))\, dx,\end{aligned}$$

and finally

$$\begin{aligned}\boldsymbol{a}_h(U,V) &= t a_m^h(u,v) + t a_t^h((u,r),(v,s)) + \frac{t^3}{12} a_f^h(r,s),\\ b_h(U,V) &= \int_\omega \mathcal{Q}^h(U)\mathcal{Q}^h(V) dx.\end{aligned}$$

We also introduce the approximation $\mathcal{L}_h$ of the linear form $\mathcal{L}$, namely,

$$\mathcal{L}_h(V) = \int_\omega f_h v \, dx.$$

Then for any $V_h \in \mathbb{X}_h$, we may write the residual as

$$\begin{aligned}\mathcal{R}_{U_h} =&\mathcal{L}(V - V_h) - \boldsymbol{a}(U_h, V - V_h) - \boldsymbol{a}_p(U_h, V - V_h) - \frac{1}{\epsilon} b(U_h, V - V_h)\\ =&(\mathcal{L} - \mathcal{L}_h)(V - V_h) - (\boldsymbol{a} - \boldsymbol{a}_h)(U_h, V - V_h)\\ &-(\boldsymbol{a}_p - \boldsymbol{a}_p^h)(U_h, V - V_h) - \frac{1}{\epsilon}(b - b_h)(U_h, V - V_h)\\ &-\boldsymbol{a}_h(U_h, V - V_h) - \boldsymbol{a}_p(U_h, V - V_h) - \frac{1}{\epsilon} b_h(U_h, V - V_h) + \mathcal{L}_h(V - V_h).\end{aligned} \tag{6.2}$$

We again recall the properties of the Clément operator $\mathcal{C}_h$ [11], for $0 \leqslant m \leqslant l \leqslant 1$

$$\forall h, \forall T \in \mathcal{T}_h, \forall w \in H^l(\omega) \quad \|w - \mathcal{C}_h w\|_{m,T} \lesssim h_T^{l-m} \|w\|_{l,\Delta(T)}, \tag{6.3}$$

$$\forall h, \forall e \in \mathcal{E}_h, \forall w \in H^l(\omega) \quad \|w - \mathcal{C}_h w\|_{m,e} \lesssim h_e^{l-m-\frac{1}{2}} \|w\|_{l,\Delta(e)}, \tag{6.4}$$

where $\Delta(T) = \cup_{T' \in \mathcal{T}_h : T' \cap T \neq \emptyset} T'$ (resp. $\Delta(e) = \cup_{T' \in \mathcal{T}_h : e \subset T'} T'$) is the patch associated with the element T (resp. the edge e) and $\mathcal{E}_h$ is the set of edges of the triangulation.

Lemma 6.1. — *Let $V = (v, \sum_i s_i a_i)$ and $V_h = (v_h, s_h) = (\mathcal{C}_h v, \sum_i (\mathcal{C}_h s_i) a_i)$, then we have the following estimate*

$$\begin{aligned}\big|(\mathcal{L} - \mathcal{L}_h)(V - V_h) - (\boldsymbol{a} - \boldsymbol{a}_h)(U_h, V - V_h) - (\boldsymbol{a}_p - \boldsymbol{a}_p^h)(U_h, V - V_h)\\ -\epsilon^{-1}(b - b_h)(U_h, V - V_h)\big| \lesssim \left(\varepsilon_h^d + \varepsilon_h^c\right) \|V\|_{\mathbb{X}},\end{aligned}$$

where

$$\begin{aligned}\varepsilon_h^c =& (\varepsilon^{-1} \max_{k=1,2,3} \|a_k - a_k^h\|_{\infty,\omega} + \max_{\alpha,\beta=1,2} \|\partial_\alpha a_\beta - d_{\alpha\beta}^h\|_{\infty,\omega}\\ &+ \max_{\rho,\sigma=1,2} \|b_{\rho\sigma} - b_{\rho\sigma}^h\|_{\infty,\omega}) \, \|f\|_\omega,\\ \varepsilon_T^d =& h_T \|f - f_h\|_T,\end{aligned}$$

and

$$\varepsilon_h^d = (\sum_T \left(\varepsilon_T^d\right)^2)^{\frac{1}{2}}.$$

Proof. — First one estimates the term $(\mathcal{L} - \mathcal{L}_h)(V - V_h)$. As we have

$$\begin{aligned}(\mathcal{L} - \mathcal{L}_h)(V - V_h) &= \int_\omega f \cdot (v - \mathcal{C}_h v) dx - \int_\omega f_h \cdot (v - \mathcal{C}_h v) dx\\ &= \int_\omega (f - f_h) \cdot (v - \mathcal{C}_h v) dx\\ &= \sum_{T \in \mathcal{T}_h} \int_T (f - f_h) \cdot (v - \mathcal{C}_h v) dx,\end{aligned}$$

Cauchy-Schwarz's inequality and the property (6.3) of $\mathcal{C}_h$ yield

$$|(\mathcal{L} - \mathcal{L}_h)(V - V_h)| \leqslant \varepsilon_h^d \|V\|_{\mathbb{X}}.$$

Secondly we estimate

$$(\boldsymbol{a}-\boldsymbol{a}_h)(U_h,V-V_h)+(\boldsymbol{a}_p-\boldsymbol{a}_p^h)(U_h,V-V_h)+\epsilon^{-1}(b-b_h)(U_h,V-V_h).$$

We only give an abridged proof of this technical result. We first estimate

$$\begin{aligned}(\boldsymbol{a}-\boldsymbol{a}_h)(U_h,V-V_h)&=t(a_m-a_m^h)(u_h,v-v_h)+t(a_t-a_t^h)(U_h,V-V_h)\\&\quad+\frac{t^3}{12}(a_f-a_f^h)(r_h,s-s_h).\end{aligned}$$

To estimate the term $(a_m-a_m^h)(U_h,V-V_h)$, we typically have to estimate a term like

$$A_h(u_h,v-v_h):=\int_\omega(\gamma_{11}(u_h)\gamma_{11}(v-v_h)-\gamma_{11}^h(u_h)\gamma_{11}^h(v-v_h))\,dx.$$

That we transform as

$$\begin{aligned}&A_h(u_h,v-v_h)\\&\quad=\int_\omega(\gamma_{11}(u_h)(\gamma_{11}(v-v_h)-\gamma_{11}^h(v-v_h))+(\gamma_{11}(u_h)-\gamma_{11}^h(u_h))\gamma_{11}^h(v-v_h))\,dx.\end{aligned}$$

For the first term, we use the identity $\gamma_{11}(u)-\gamma_{11}^h(u)=\partial_1 u\cdot(a_1-a_1^h)$, and apply Cauchy-Schwarz's inequality and (4.3) to get

$$\left|\int_\omega\gamma_{11}(u_h)(\gamma_{11}(v-v_h)-\gamma_{11}^h(v-v_h))\,dx\right|\lesssim\|f\|_\omega\|\partial_1(v-v_h)\cdot(a_1-a_1^h)\|_\omega.$$

As

$$\|\partial_1(v-v_h)\cdot(a_1-a_1^h)\|_\omega\leqslant\|a_1-a_1^h\|_{L^\infty(\omega)}\|\partial_1(v-v_h)\|_\omega,$$

by the property (6.3), we deduce that

$$\left|\int_\omega\gamma_{11}(u_h)(\gamma_{11}(v-v_h)-\gamma_{11}^h(v-v_h))\,dx\right|\lesssim\varepsilon_h^c\|f\|_\omega\|V\|_{\mathbb{X}}.$$

The second term is estimated in the same manner, which leads to

$$|A_h(u_h,v-v_h)|\lesssim\varepsilon_h^c\|f\|_\omega\|V\|_{\mathbb{X}}.$$

The same techniques on the remaining terms of $\boldsymbol{a}-\boldsymbol{a}_h$ and on all terms of $\boldsymbol{a}_p-\boldsymbol{a}_p^h$ yield

$$\begin{aligned}|(\boldsymbol{a}-\boldsymbol{a}_h)(u_h,v-v_h)|&\lesssim\varepsilon_h^c\|f\|_\omega\|V\|_{\mathbb{X}},\\|(\boldsymbol{a}_p-\boldsymbol{a}_p^h)(r_h,s-s_h)|&\lesssim\varepsilon_h^c\|f\|_\omega\|V\|_{\mathbb{X}}.\end{aligned}$$

The last term $\varepsilon^{-1}(b-b_h)$ requires a more specific attention. First it is split up as follows

$$\begin{aligned}\varepsilon^{-1}(b-b_h)(U_h,V-V_h)&=\varepsilon^{-1}\int_\omega\left(\mathcal{Q}(U_h)\mathcal{Q}(V-V_h)-\mathcal{Q}^h(U_h)\mathcal{Q}^h(V-V_h)\right)\,dx\\&=\varepsilon^{-1}\int_\omega\mathcal{Q}(U_h)(\mathcal{Q}(V-V_h)-\mathcal{Q}^h(V-V_h))dx\\&\quad+\varepsilon^{-1}\int_\omega\mathcal{Q}^h(V-V_h)(\mathcal{Q}(U_h)-\mathcal{Q}^h(U_h))dx.\end{aligned}$$

Hence using Cauchy-Schwarz's inequality, and the property

$$\mathcal{Q}(u,r)-\mathcal{Q}^h(u,r)=-\frac{1}{2}((a_2-a_2^h)\partial_1 u-(a_1-a_1^h)\partial_2 u),$$

we find

$$\varepsilon^{-1}|(b-b_h)(U_h,V-V_h)| \lesssim \varepsilon^{-1} \sup_{k=1,2,3} \|a_i - a_i^h\|_{L^\infty(\omega)} \ \|U_h\|_{\mathbb{X}} \|V-V_h\|_{\mathbb{X}}.$$

Using the bound (4.3) and the estimate (6.3), we find

$$\varepsilon^{-1}|(b-b_h)(U_h,V-V_h)| \lesssim \varepsilon^{-1} \sup_{k=1,2,3} \|a_i - a_i^h\|_{L^\infty(\omega)} \ \|f\|_{\omega} \|V\|_{\mathbb{X}}.$$

The previous estimates yield the conclusion. □

Now we need to estimate the term

$$\mathcal{L}_h(V-V_h) - \boldsymbol{a}_h(U_h,V-V_h) - \boldsymbol{a}_p^h(U_h,V-V_h) - \frac{1}{\epsilon} b_h(U_h,V-V_h).$$

In order to define appropriately the indicators, we introduce

$$\begin{aligned}
T_h(u) &= t\ \mathbb{A}\gamma^h(u),\\
A_h &= (a_1^h, a_2^h)^\top,\\
S_h(u,r) &= t\ \mu((\nabla u)^\top a_3^h + J\hat{r}),\\
M_h(r) &= \frac{t^3}{24}\mathbb{A}\Pi^h(r),\\
P_h(r) &= \frac{t^3}{12}\mathbb{A}II_0^h r_3,\\
\kappa_h(r) &= \frac{t^3}{12}(II_0^h : \mathbb{A}\theta^h(r)).
\end{aligned}$$

Now for all $T \in \mathcal{T}_h$, we can define the following indicators (compare with problem (5.17))

$$\begin{aligned}
\eta_T^{(1)} &= h_T\|f_h + \mathrm{Div}\,(T_h(u_h)A_h) + \mathrm{Div}\,(S_h(U_h)a_3^h) - \frac{1}{2\epsilon}\mathrm{Div}\,(\mathcal{Q}^h(U_h)JA_h)\|_{L^2(T,\mathbb{R}^3)}\\
&+ \sum_{e\in\mathcal{E}_h^i\cap\partial T} \frac{1}{2} h_e^{\frac{1}{2}} \|[nT_h(u_h)A_h + nS_h(U_h)a_3^h - \frac{1}{2\epsilon}\mathcal{Q}^h(U_h)A_h^\top Jn^\top]_e\|_{L^2(e,\mathbb{R}^3)}\\
&+ \sum_{e\in\mathcal{E}_h^b\cap\bar{\Gamma}_1\cap\partial T} h_e^{\frac{1}{2}} \|nT_h(u_h)A_h + nS_h(U_h)a_3^h - \frac{1}{2\epsilon}\mathcal{Q}^h(U_h)A_h^\top Jn^\top\|_{L^2(e,\mathbb{R}^3)},\\
\eta_T^{(2)} &= h_T\|J^\top \mathrm{Div}\, M_h(r_h) + \begin{pmatrix} M_h(r_h) : \bar{\Gamma}_h^1 \\ M_h(r_h) : \bar{\Gamma}_h^2 \end{pmatrix} - J^\top S_h(U_h)\\
&+ \frac{1}{2}\tilde{J}\mathrm{Div}\,(P_h(r_h)) - \frac{1}{2}\begin{pmatrix} P_h(u_h) : \tilde{\Gamma}_h^1 \\ P_h(u_h) : \tilde{\Gamma}_h^2 \end{pmatrix} \|_{L^2(T)^2}\\
&+ \sum_{e\in\mathcal{E}_h^i\cap\partial T} h_e^{\frac{1}{2}} \|[\frac{1}{2}\tilde{J}P_h(r_h)n^\top + J^\top M_h(r_h)n^\top]_e\|_{L^2(e)^2}\\
&+ \sum_{e\in\mathcal{E}_h^b\cap\bar{\Gamma}_1\cap\partial T} h_e^{\frac{1}{2}} \|\frac{1}{2}\tilde{J}P_h(r_h)n^\top + J^\top M_h(r_h)n^\top\|_{L^2(e)^2},\\
\eta_T^{(3)} &= \|\bar{B}_h\colon M_h(r_h) - \kappa_h(r_h) - \frac{1}{2}\tilde{B}_h\colon P_h(r_h) - \frac{1}{\epsilon}\mathcal{Q}^h(U_h)\|_{L^2(T)},
\end{aligned}$$

where $\mathcal{E}_h^b$ is the set of edges of the triangulation included into the boundary of ω, while $\mathcal{E}_h^i = \mathcal{E}_h \setminus \mathcal{E}_h^b$. We further introduce the local indicator

$$\eta_T = \eta_T^{(1)} + \eta_T^{(2)} + \eta_T^{(3)},$$

and the global one

$$\eta_h = \Big(\sum_{T\in\mathcal{T}_h} \eta_T^2 \Big)^{\frac{1}{2}}. \tag{6.5}$$

PROPOSITION 6.2. — *Let $V = (v, \sum_i s_i a_i) \in \mathbb{X}$ and let $V_h = (\mathcal{C}_h v, \sum_i (\mathcal{C}_h s_i) a_i)$ be the Clément interpolant of V, then*

$$|\boldsymbol{a}_h(U_h, V-V_h) + \boldsymbol{a}_p^h(U_h, V-V_h) + \epsilon^{-1} b_h(U_h, V-V_h) - \mathcal{L}_h(V-V_h)| \lesssim \eta_h \|V\|_{\mathbb{X}}. \tag{6.6}$$

Proof. — We split up the left-hand side of (6.6) in three terms as follows

$$\begin{aligned}\mathcal{L}_h(V-V_h) - \boldsymbol{a}_h(U_h, V-V_h) - \boldsymbol{a}_p^h(U_h, V-V_h) - \epsilon^{-1} b_h(U_h, V-V_h)\\ = A_1(U_h, V-V_h) + A_2(U_h, V-V_h) + A_3(U_h, V-V_h),\end{aligned}$$

where

$$\begin{aligned}
A_1(U_h, V-V_h) &= \mathcal{L}_h(v - \mathcal{C}_h v) - \boldsymbol{a}_h(U_h, (v-\mathcal{C}_h v, 0)) - \epsilon^{-1} b_h(U_h, (v-\mathcal{C}_h v, 0)),\\
A_2(U_h, V-V_h) &= -\boldsymbol{a}_h(U_h, (0, \sum_\alpha (s_\alpha - \mathcal{C}_h s_\alpha) a_\alpha)) - \boldsymbol{a}_p^h(U_h, (0, \sum_\alpha (s_\alpha - \mathcal{C}_h s_\alpha) a_\alpha))\\
&\quad - \epsilon^{-1} b_h(U_h, (0, \sum_\alpha (s_\alpha - \mathcal{C}_h s_\alpha) a_\alpha)),\\
A_3(U_h, V-V_h) &= -\boldsymbol{a}_h(U_h, (0, (s_3 - \mathcal{C}_h s_3) a_3)) - \boldsymbol{a}_p^h(U_h, (0, (s_3 - \mathcal{C}_h s_3) a_3))\\
&\quad - \epsilon^{-1} b_h(U_h, (0, (s_3 - \mathcal{C}_h s_3) a_3)).
\end{aligned}$$

For the first term, by elementwise Green's formula we directly have

$$\begin{aligned}
A_1(U_h, V-V_h) = &\sum_{T\in\mathcal{T}_h}\\
&\int_T (f_h + \mathrm{Div}\,(T_h(u_h) A_h) + \mathrm{Div}\,(S_h(U_h) a_3^h) - \tfrac{1}{2\epsilon} \mathrm{Div}\,(\mathcal{Q}^h(U_h) J A_h) \cdot (v - \mathcal{C}_h v)\, dx\\
&+ \sum_{T\in\mathcal{T}_h} \sum_{e\in\bar{\Gamma}_1\cap\partial T} \int_e (\tfrac{1}{2\epsilon} \mathcal{Q}^h(U_h) A_h^T J n^\top - n T_h(u_h) A_h - n S_h(U_h) a_3^h) \cdot (v - \mathcal{C}_h v)\, d\sigma(x).
\end{aligned} \tag{6.7}$$

Cauchy-Schwarz' inequality and the properties of the Clément interpolant $\mathcal{C}_h$ yield

$$|A_1(U_h, V-V_h)| \lesssim \Big(\sum_{T\in\mathcal{T}_h} \big(\eta_T^{(1)}\big)^2 \Big)^{\frac{1}{2}} \|V\|_{\mathbb{X}}.$$

In a fully similar manner, we have

$$|A_2(U_h, V-V_h)| \lesssim \Big(\sum_{T} \big(\eta_T^{(2)}\big)^2 \Big)^{\frac{1}{2}} \|V\|_{\mathbb{X}}.$$

Finally we directly check that

$$A_3(U_h, V-V_h) = \sum_T \int_T (\bar{B}_h : M_h(r_h) - \kappa(r_h) - \tfrac{1}{2} \tilde{B}_h : P_h(r_h) - \tfrac{1}{\epsilon} \mathcal{Q}^h(U_h))(s_3 - \mathcal{C}_h s_3) dx, \tag{6.8}$$

hence using (6.3), we directly get

$$|A_3(U_h, V - V_h)| \lesssim \Big(\sum_T \big(\eta_T^{(3)}\big)^2 \Big)^{\frac{1}{2}} \|V\|_{\mathbb{X}}.$$

The estimates on $|A_i(U_h, V - V_h)|$ directly yield the conclusion. □

6.2. **Upper and lower error bounds.** At this stage we are able to prove the following robust upper bound.

THEOREM 6.3. — *The following a posteriori error estimate holds between the solution U_ϵ of problem (3.6) and the solution U_h of problem (4.2)*

$$\|U_\epsilon - U_h\|_{\mathbb{X}} \lesssim \eta_h + \varepsilon_h^d + \varepsilon_h^c. \tag{6.9}$$

Proof. — The estimate (6.9) follows from the fact that $\boldsymbol{a}(\cdot,\cdot) + \boldsymbol{a}_p(\cdot,\cdot) + \epsilon^{-1} b(\cdot,\cdot)$ is coercive with a coercivity constant equivalent to 1, by using the identity (6.2), Lemma 6.1 and Proposition 6.2. □

Let us go with the lower bound.

THEOREM 6.4. — *Let U_ϵ be the solution of problem (3.6) and U_h the solution of problem (4.2). Then we have the following bound*

$$\eta_T^{(i)} \lesssim \epsilon^{-1} \|U_\epsilon - U_h\|_{\mathbb{X}(\omega_T)} + \varepsilon_{\omega_T}^d + \varepsilon_{\omega_T}^c, \quad i = 1, 2, 3, \tag{6.10}$$

where the index ω_T means that the quantity is taken only in ω_T and the norm $\mathbb{X}(\omega_T)$ means the norm of $\mathbb{X}$ with integrals restricted to ω_T.

Proof. — The proof is quite standard and is based on standard inverse inequaltiy, see [22] for instance. We will only prove the inequality (6.10) for $\eta_T^{(1)}$ since it is fully similar for $\eta_T^{(2)}$ and $\eta_T^{(3)}$. For shortness, we write $\eta_T^{(1)}$ in the following compact form

$$\eta_T^{(1)} = h_T \|F_h\|_T + \sum_{e \in \mathcal{E}_h^i \cap \partial T} h_e^{\frac{1}{2}} \|[G_h]_e\|_e + \sum_{e \in \mathcal{E}_h^b \cap \partial T} h_e^{\frac{1}{2}} \|G_h\|_e.$$

First of all, let us fix the standard bubble function ψ_T associated with T and set

$$v = \begin{cases} F_h \psi_T \text{ in } T, \\ 0 \quad \text{in } \omega \setminus T. \end{cases} \tag{6.11}$$

By the definition of ψ_T, we may notice that $v \in H_0^1(\omega, \mathbb{R}^3)$ and hence $(v, 0)$ belongs to $\mathbb{X}$. It follows from (6.7) with $V_h = 0$ that

$$\begin{aligned} \mathcal{L}_h(v,0) - \boldsymbol{a}_h(U_h,(v,0)) - \epsilon^{-1} b_h(U_h,(v,0)) &= \int_T \big(f_h + \mathrm{Div}\,(T_h(u_h)A_h) + \mathrm{Div}\,(S_h(U_h)a_3^h) - \frac{1}{2\epsilon}\mathrm{Div}\,(\mathcal{Q}^h(U_h)JA_h)\big) \cdot v\, dx \\ &= \|F_h \psi_T^{\frac{1}{2}}\|^2_{L^2(T)^3}. \end{aligned}$$

Using the identity (6.2), we may write

$$\begin{aligned} \boldsymbol{a}(U^\epsilon - U_h,(v,0)) + \epsilon^{-1} b(U^\epsilon - U_h,(v,0)) &= (\mathcal{L} - \mathcal{L}_h)((v,0)) - (\boldsymbol{a} - \boldsymbol{a}_h)(U_h,(v,0)) \\ &\quad - \frac{1}{\epsilon}(b - b_h)(U_h,(v,0)) - \boldsymbol{a}_h(U_h,(v,0)) \\ &\quad - \frac{1}{\epsilon} b_h(U_h,(v,0)) + \mathcal{L}_h((v,0)). \end{aligned}$$

Hence

$$\begin{aligned}\mathcal{L}_h(v,0) - \boldsymbol{a}_h(U_h,(v,0)) - \epsilon^{-1}b_h(U_h,(v,0)) = \\ \boldsymbol{a}(U^\epsilon - U_h,(v,0)) + \epsilon^{-1}b(U^\epsilon - U_h,(v,0)) - (\mathcal{L}-\mathcal{L}_h)((v,0)) \\ + (\boldsymbol{a}-\boldsymbol{a}_h)(U_h,(v,0)) - \frac{1}{\epsilon}(b-b_h)(U_h,(v,0)).\end{aligned}$$

By the previous identities, we get

$$\begin{aligned}\|F_h\psi_T^{\frac{1}{2}}\|^2_{L^2(T)^3} =& \boldsymbol{a}(U^\epsilon - U_h,(v,0)) + \epsilon^{-1}b(U^\epsilon - U_h,(v,0)) \\ &- (\mathcal{L}-\mathcal{L}_h)(v,0) + (\boldsymbol{a}-\boldsymbol{a}_h)(U_h,(v,0)) \\ &- \frac{1}{\epsilon}(b-b_h)(U_h,(v,0)).\end{aligned}$$

So by Cauchy-Schwarz's inequality and the arguments of Lemma 6.1, we find

$$\|F_h\psi_T^{\frac{1}{2}}\|^2_{L^2(T,\mathbb{R}^3)} \lesssim \left(\epsilon^{-1}\|U_\epsilon - U_h\|_{\mathbb{X}(T)} + \varepsilon_T^d + \varepsilon_h^c\right)\|v\|_{H^1(T,\mathbb{R}^3)}. \tag{6.12}$$

Using the following inverse inequality

$$\|v\|_{H^1(T,\mathbb{R}^3)} \lesssim h_T^{-1}\|v\|_{L^2(T,\mathbb{R}^3)}, \tag{6.13}$$

and using that the function ψ_T takes it values between 0 and 1, we deduce

$$\|v\|_{H^1(T,\mathbb{R}^3)} \lesssim h_T^{-1}\|F_h\|_{L^2(T)^3}. \tag{6.14}$$

In addition we have

$$\|F_h\|_{L^2(T)^3} \leqslant c\|F_h\psi_T^{\frac{1}{2}}\|_{L^2(T)^3}. \tag{6.15}$$

Combining (6.12), (6.14) and (6.15) we get

$$h_T\|F_h\|_{L^2(T)^3} \lesssim \epsilon^{-1}\|U_\epsilon - U_h\|_{\mathbb{X}(T)} + \varepsilon_T^d + \varepsilon_T^c.$$

The second step is to bound the second term of $\eta_T^{(1)}$, for all edges e of T shared with the element T'. In this case we choose the function v in (6.7) as follows

$$v = \begin{cases}\mathcal{M}_{e,\kappa}([G_h]_e)\psi_e \text{ for } \kappa \in \{T,T'\}, \\ 0 \quad \text{in } \omega \setminus (T \cup T'),\end{cases} \tag{6.16}$$

where ψ_e is the standard edge bubble function associated with e and $\mathcal{M}_{e,\kappa}(q)$ is an extension operator that sends a polynomial q in the edge coordinate of e to a polynomial in cartesian coordinates in κ. As before we see that

$$\begin{aligned}\|[G_h]_e\psi_e\|^2_{L^2(e)^2} = \boldsymbol{a}_h(U_h,(v,0,0,0)) + \epsilon^{-1}b_h(U_h,(v,0,0,0)) - \mathcal{L}_h(v,0,0,0) \\ + \int_{\Delta(e)}(f_h + \mathrm{Div}\,(T_h(u_h)A_h) + \mathrm{Div}\,(S_h(U_h)a_3^h) - \frac{1}{2\epsilon}\mathrm{Div}\,(\mathcal{Q}^h(U_h)JA_h)\cdot v dx.\end{aligned}$$

Using the identity (6.2) and the arguments of Lemma 6.1, we then have

$$\begin{aligned}\|[G_h]_e\psi_e\|^2_e \lesssim \epsilon^{-1}\|U_\epsilon - U_h\|_{\mathbb{X}(\Delta(e))}\|v\|_{\mathbb{X}(\Delta(e))} + \left(\varepsilon^d_{\Delta(e)} + \varepsilon^c_{\Delta(e)}\right)\|v\|_{\mathbb{X}(\Delta(e))} \\ + \|F_h\|_{\Delta(e)}\|v\|_{\Delta(e)}.\end{aligned}$$

By a standard inverse inequality, we conclude

$$h_e^{\frac{1}{2}}\|[G_h]_e\psi_e\|^2_e \lesssim \epsilon^{-1}\sum_{\kappa\in\{T,T'\}}\|U_\epsilon - U_h\|_{\mathbb{X}(\Delta(e))} + \varepsilon^d_{\Delta(e)} + \varepsilon^c_{\Delta(e)}.$$

The third term is bounded in the same manner than the second one. The proof is therefore complete. □

Remark 6.5. — As usual, if f is more regular than simply $L^2(\omega,\mathbb{R}^3)$, namely if f belongs to $H^s(\omega,\mathbb{R}^3)$ for some $s>0$, and if f_h is chosen to be the $L^2(\omega,\mathbb{R}^3)$ projection of f into $\mathbb{Z}_h$, then ε_h^d is a higher order term (h.o.t.) with respect to the optimal order of convergence h expected for the error between U_ϵ and U_h (since the boundary of ω may be non smooth). Indeed by a standard interpolation estimate we have

$$\|f-f_h\|_T \lesssim h^s |f|_{s,T},$$

which yields

$$\varepsilon_T^d \lesssim h^{1+s}.$$

Since s is positive, we deduce that $\frac{\varepsilon^d}{h}=o(1)$, in other words ε_h^d is a h.o.t.

Similarly if the coefficients a_α are smoother, namely if $a_\alpha \in W^{3,p}(\omega,\mathbb{R}^3)$, for some $p>2$ (which is guaranteed by the regularity $\varphi \in W^{4,p}(\omega,\mathbb{R}^3)$), if $h\lesssim\epsilon$, and if a_k^h (resp. $d_{\alpha\beta}^h$ and $b_{\alpha\beta}^h$) is chosen to be the $L^2(\omega)$ projection of a_k into $(\tilde{\mathbb{M}}_h^{(2)})^3$ (resp. of $\partial_\alpha a_\beta$ into $(\tilde{\mathbb{M}}_h^{(1)})^3$ and of $b_{\alpha\beta}$ into $\tilde{\mathbb{M}}_h^{(1)}$), then ε_h^c is a h.o.t. Indeed under the above assumptions, using [8, Theorem 3.1.4], we have

$$\begin{aligned}
\|a_k-a_k^h\|_{L^\infty(T)^3} &\lesssim h^{3-\frac{2}{p}}|a_k|_{3,p,\omega},\\
\|\partial_\alpha a_\beta - d_{\alpha\beta}^h\|_{L^\infty(T)^3} &\lesssim h^{2-\frac{2}{p}}|\partial_\alpha a_\beta|_{2,p,\omega},\\
\|b_{\alpha\beta}-b_{\alpha\beta}^h\|_{L^\infty(T)} &\lesssim h^{2-\frac{2}{p}}|b_{\alpha\beta}|_{2,p,\omega}.
\end{aligned}$$

With the assumption $h\lesssim\epsilon$, we deduce that

$$\varepsilon_h^c \lesssim h^{2-\frac{2}{p}} \max_{\alpha=1,2}\|a_\alpha\|_{3,p,\omega}.$$

We then deduce that $\frac{\varepsilon_h^c}{h}=h^{1-\frac{2}{p}}$ which is $o(1)$ as soon as $p>2$.

7. Numerical experiments

We now describe how the error indicators exhibited in section 5 can be used to adapt the mesh for the discrete problem (4.2). We use Dörfler [12] marking strategy, which is a practical procedure to estimate and equidistribute the local error. An efficient indicator identifies the parts of the domain that induces large errors and use this information to locally refine and if necessary repeat the finite element computation. We start with an initial coarse triangulation $\mathcal{T}_h$ followed by an iterative loops procedure of the form:

SOLVE → ESTIMATE → MARK → REFINE

The numerical experiments that we now present have been performed using the finite element code FreeFem++ [16]. Note that Freefem++ contains an anisotropic mesh generator (BAMG[1]), thus the mesh is refined automatically, hence the adapted mesh is not necessarily quasi uniform. The obtained results will be used to test the reliability of the anisotropic adaptive mesh procedure.

Numerical computations are made using the scheme (4.2) with P_3-Lagrange elements for the displacement and P_2-Lagrange element for the rotation, as well as

[1]Bidimensional Anisotropic Mesh Generator

the scheme (4.8) with P_2 elements for displacement, P_1 for the rotation and P_1 for the Lagrange multiplier. It is well known that condition numbers of penalized methods are larger than their non-penalized counterparts, especially for very small values of the thickness t. In the context of the linear Koiter shell model, the relationship between the thickness t, the mesh size h, the parameter ϵ and the corresponding condition number can be found in [19, Sec.6]. Even for a fixed value of the thickness, taking "very small" values of the parameter ϵ leads to systems with very large condition number and therefore the solutions are unstable with respect to small changes in data. However, the strategy (of choosing the parameter) that we have made for this model is the same than the one we have used for other classical shell models (Naghdi, Koiter ...) without the prestressed term. Indeed, we have observed that when the penalization parameter ϵ^{-1} was $10 \times \frac{E}{2(1+\nu)}$ we obtain excellent results for the constraint $r^\epsilon \cdot a_3 - \tilde{\gamma}_{12}(u^\epsilon)$ and also good results in terms of the reference value. Note also that in our considered numerical experiments we have observed that the range of values of ϵ that can be used is between 10^{-6} and 10^{-16} and that the results are robust with respect to ϵ since they are very close for $\epsilon^{-1} = 10^9$ and $\epsilon^{-1} = 7.6 \times 10^{11}$. For shortness, we have only presented the results for this last value.

7.1. **First example.** In the first example, we consider a cylindrical shell (see Figure 7.1). We take E = 200 GPa for the Young modulus and $\nu = 0.3$ for the Poisson ratio of the material. Note that this gives the values

$$\frac{4\lambda\mu}{\lambda+2\mu} = 1.31868\times10^{11},\ 4\mu = 3.07692\times10^{11} \text{ and } \epsilon^{-1} = 10\times\frac{E}{2(1+\nu)} = 7.6\times10^{11},$$

where,

$$\lambda = \frac{\mathrm{E}\ \nu}{(1+\nu)(1-2\nu)}, \quad \mu = \frac{\mathrm{E}}{2(1+\nu)}$$

The radius $R = 3/2$, the length $L = 2R$, and the angle $\alpha = 40°$, hence the width of the rectangle is $2R_0$, with $R_0 = R\sin\alpha = 0.6427$. In other words, the domain ω is defined to be the rectangle

$$\omega = \{(x,y) \in \mathbb{R}^2; \quad -R_0 \leqslant x \leqslant R_0 \quad \text{and} \quad 0 \leqslant y \leqslant 2L\}.$$

The middle surface S can be parametrized by the chart φ, with

$$\varphi(x,y) = (R\sin(x/R), y, R\cos(x/R)).$$

Then the covariant basis is

$$a_1 = (\cos(x/R), 0, -\sin(x/R))^\top, \quad a_2 = (0,1,0)^\top, \quad a_3 = (\sin(x/R), 0, \cos(x/R))^\top,$$

and

$$(b_{\alpha\beta})_{1\leqslant\alpha,\beta\leqslant2} = \begin{pmatrix} -\frac{1}{R} & 0 \\ 0 & 0 \end{pmatrix}.$$

The asymptotic directions are curves of the form $x = Cte$. As loading we chose f consistent with the flexural regime, namely,

$$f = t^3 \times q \times \cos(2y) a_3, \quad q = -5 \times 10^7,$$

while the thickness t of the shell is fixed to be $t = 0.01R$. Note that the numerical values of the constraint $r^\epsilon \cdot a_3 - \tilde{\gamma}_{12}(u^\epsilon)$ vary between $\pm 4 \times 10^{-8}$.

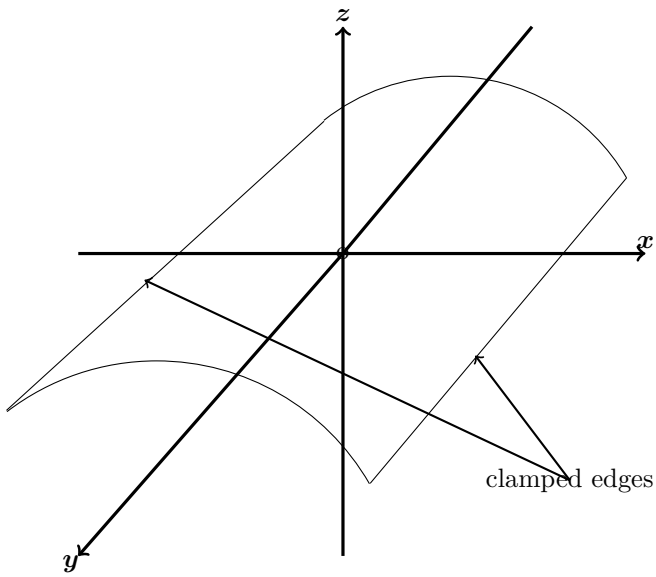


FIGURE 7.1. The shell geometry

Since the values of t are relatively small in our tests, we avoid the use of the P2-P1 schemes because it is shown in [20, Table 1. p. 16] that this scheme suffers from locking phenomena.

In Table 7.1 we present the convergence results for P3-P2 elements for uniformly refined meshes. We have used seven uniformly refined meshes of the form $2N \times N$, $N = 10, 15, 20, 25, 30, 35, 40$, where N denotes the number of points of the triangulation on the shorter side of the rectangular domain. Table 7.2 presents the convergence results for P3-P2 elements using adaptive method based on our a posteriori error estimator. Note that since we do not know the exact solution for the considered test, the reference solution is obtained using P3-P2 elements on a very fine mesh corresponding to $N = 80$ (which leads to a linear system with 502806 degrees of freedom) and we present the relative L_2 error in term of the number of degrees of freedom or the mesh size. This allows us to avoid the dependency with respect to the coefficient E and the parameter ϵ^{-1} which are rather larger for our numerical tests.

Using the residual a posteriori estimator with 73363 degrees of freedom we obtain good results for the error with respect to the L_2 norm, while 126606 degrees of freedom are needed for the uniform adaptive method to get an error of the same order. This means that the use of adaptive meshes speeds up the convergence. Note that we may observe that the error slightly oscillates for the adaptive method, this phenomena was also observed for other classical thin shell models.

Table 7.3 presents the values of $\max_{T\in\mathcal{T}_h} \eta_T^i$ for $i = 1, 2, 3$ against the number of degrees of freedom from step 1 to step 6. We notice that their values decrease and converge to zero, which confirm the effectiveness of our estimator. Note that the values of the components $\max\limits_{T\in\mathcal{T}_h} \eta_T^i$ using the P3-P2 elements reveal that for the

N	h	Number of degrees of freedom	$\frac{\|u-u^{ref}\|}{\|u^{ref}\|}$	$\frac{\|r-r^{ref}\|}{\|r^{ref}\|}$
10	0.244	8256	0.002478	0.000179
15	0.162	18231	0.000418	4.9×10^{-5}
20	0.122	32106	0.000151	1.32×10^{-5}
25	0.097	49881	8.10×10^{-5}	3.76×10^{-6}
30	0.081	71556	5.07×10^{-5}	1.29×10^{-6}
35	0.069	79131	3.23×10^{-5}	4.62×10^{-7}
40	0.061	126606	2.36×10^{-5}	3.25×10^{-7}

TABLE 7.1. Convergence results using uniform refined meshes for the penalized version

Iteration	Number of degrees of freedom	$\frac{\|u-u^{ref}\|}{\|u^{ref}\|}$	$\frac{\|r-r^{ref}\|}{\|r^{ref}\|}$
1	3438	0.000872	0.006916
2	7293	0.000174	0.005271
3	15450	0.000819	0.001484
4	33615	0.000527	4.14×10^{-5}
5	73362	1.44×10^{-5}	6.31×10^{-5}

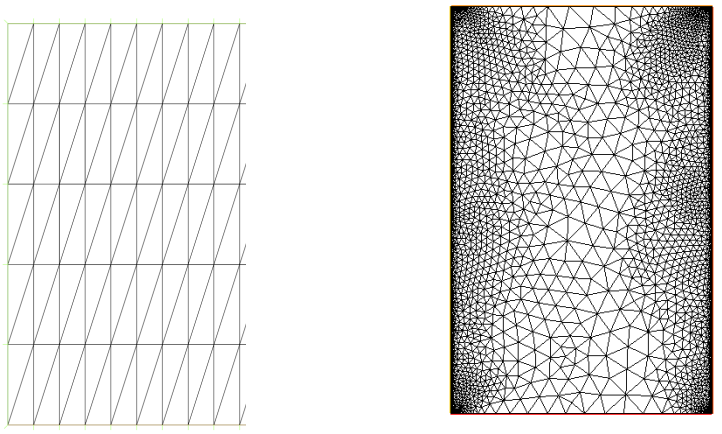
TABLE 7.2. Convergence results using adaptive method for the penalized version

Iteration	Degrees of freedom	$\max_{T\in\mathcal{T}_h}\eta_T^{(1)}$	$\max_{T\in\mathcal{T}_h}\eta_T^{(2)}$	$\max_{T\in\mathcal{T}_h}\eta_T^{(3)}$	$\frac{\ln\left(\frac{\eta(i)}{\eta(i+1)}\right)}{\ln\left(\frac{N(i)}{N(i+1)}\right)}$	$\frac{\ln\left(\frac{\eta(i)}{\eta(i+1)}\right)}{\ln\left(\frac{h(i)}{h(i+1)}\right)}$
1	3438	58.51	0.44	1.36	1.6291	4.2327
2	7293	14.94	0.12	1.66	1.0280	3.6051
3	15450	7.49	0.10×10^{-5}	0.75	1.0924	3.3954
4	33615	3.26	0.05×10^{-5}	0.29	0.9484	3.0442
5	73362	1.95	0.008×10^{-5}	0.31	1.1064	2.9388
6	145116	0.80	0.007×10^{-6}	0.11		

TABLE 7.3. Convergence results for the error indicators defined in Section 5 for the penalized version

considered example the maximal value of the indicator η_T^2 is indeed much smaller than η_T^1 and η_T^3. This is due to the fact that the indicators η_T^1 and η_T^3 depend on the parameter ϵ^{-1} while η_T^2 does not (see the proof of Proposition 6.2). In the sixth column of Table 7.3, we show the computed value of the ratio $\ln\left(\frac{\eta(i)}{\eta(i+1)}\right) / \ln\left(\frac{N(i)}{N(i+1)}\right)$, where $\eta(i)$ is global indicator η_h (defined by (6.5)) in the step $i = 1, \ldots, 5$ and $N(i)$ is the number of degrees of freedom for the same step. We observe that the value of this ratio is close to 1 which suggests the relation $\eta_h \approx Cte\ N^{-1}$. The last column of Table 7.3 shows the relationship between the global indicator and the mesh size h, from those results, we may conjecture that η_h is proportional to h^α, with $\alpha \approx 3$.

We then conclude that the error indicators that we propose show good convergence results for the considered example.



(A) Initial mesh (B) Adapted mesh after six iterations

FIGURE 7.2. Initial and adapted mesh

Figure 7.2 represents the initial coarse mesh and the refined mesh after six iterations, we notice that the number of triangles is dense only in the vicinity of the clamped edge, and get decreased whenever we go for away from the clamped boundary. This is due to the boundary layer effect (see [21]).

In order to compare the two discrete formulations, namely the penalized and the mixed ones, we have computed the scaled energy

$$E_{sc} = t^{-2}a_m(u,u) + t^{-2}a_{sE_s}((u,r),(u,r)) + \frac{1}{12}a_f(r,r) + \frac{1}{12}a_p(r,r)$$

for different values of t, i. e., $0.01R, 0.001R$ and $0.0001R$ (for a fixed fine mesh). We have observed that the bending energy is dominant even when we use the P2-P1-P1 [2] elements for $t = 0.01R$. But the scaled energy decreases when t decreases and the bending energy is not dominant (see Table 7.5). Whereas when we use P3-P2-P1 elements or P4-P3-P2 elements the scaled energy remains very close to a constant value even when t decreases to zero (see Tables 7.6 and 7.7). Note that the same phenomena was observed for the penalized formulation, see Tables 7.8 to 7.10. From our results, we may observe that the convergence results for the mixed formulation are slightly better than that of the penalized formulation. But in any case, we can conclude that both methods are locking free if polynomials of sufficiently high order are used.

Table 7.4 represents the results for the first example when we use the mixed formulation with P3-P2-P2 elements, uniform adapted meshes and $t = 0.01R$. Note that using the same number of triangles as that for the penalized version leads to a slightly larger number of degrees of freedom. We observe also that the convergence

[2] P_2 elements for displacement, P_1 for the rotation and P_1 for the Lagrange multiplier

results for the mixed formulation are slightly better than that of the penalized formulation, compare with Table 7.1.

N	h	Number of degrees of freedom	$\frac{\|u-u^{ref}\|}{\|u^{ref}\|}$	$\frac{\|r-r^{ref}\|}{\|r^{ref}\|}$
10	0.244	9117	0.002112	0.000126
15	0.162	20122	0.000367	3.3×10^{-5}
20	0.122	35427	0.000127	1.12×10^{-5}
25	0.097	55032	7.02×10^{-5}	2.34×10^{-6}
30	0.0814	78937	4.79×10^{-5}	1.03×10^{-6}
35	0.069	107142	2.77×10^{-5}	2.34×10^{-7}
40	0.061	139647	1.54×10^{-5}	1.23×10^{-7}

TABLE 7.4. Convergence results using uniform refined meshes for the mixed FE

$k = 2, 1, 1$	$\frac{t}{R} = 0.01$	$\frac{t}{R} = 0.001$	$\frac{t}{R} = 0.0001$
E_m	5.778	6.827	0.140
E_s	0.178	0.028	0.00061
E_f	30.622	3.032	0.00062
E_p	-0.00013	-1.7×10^{-6}	-8×10^{-10}
E_{sc}	36.578	9.88	0.142

TABLE 7.5. Energy values for the mixed method.

$k = 3, 2, 1$	$\frac{t}{R} = 0.01$	$\frac{t}{R} = 0.001$	$\frac{t}{R} = 0.0001$
E_m	5.255	0.397	0.106
E_s	0.168	0.004	0.154
E_f	33.520	33.52	31.577
E_p	-0.00014	-3.437×10^{-7}	-2.495×10^{-9}
E_{sc}	38.299	33.922	31.838

TABLE 7.6. Energy values for the mixed method.

$k = 4, 3, 2$	$\frac{t}{R} = 0.01$	$\frac{t}{R} = 0.001$	$\frac{t}{R} = 0.0001$
E_m	5.089	0.0616	0.0042
E_s	0.168	0.0018	2.4×10^{-5}
E_f	34.28	34.288	34.5278
E_p	-0.000157	-5×10^{-7}	7×10^{-9}
E_{sc}	39.644	34.35	34.269

TABLE 7.7. Energy values for the mixed method.

$k=2,1$	$\frac{t}{R}=0.01$	$\frac{t}{R}=0.001$	$\frac{t}{R}=0.0001$
E_m	5.75962	6.81387	0.140034
E_s	0.176713	0.0257213	0.000509613
E_f	30.5737	3.00906	0.000619921
E_p	-9.90932×10^{-5}	-9.18144×10^{-8}	-3.03847×10^{-12}
E_{sc}	36.5099	9.84865	0.141163

TABLE 7.8. Energy values for the penalized version.

$k=3,2$	$\frac{t}{R}=0.01$	$\frac{t}{R}=0.001$	$\frac{t}{R}=0.0001$
E_m	5.12205	0.410782	0.0914724
E_s	0.169229	0.0034947	0.0002709
E_f	34.6118	33.7115	32.1919
E_p	-0.000150652	-1.39831×10^{-7}	-6.95845×10^{-11}
E_{sc}	39.9029	34.1258	32.2836

TABLE 7.9. Energy values for the penalized version.

$k=4,3$	$\frac{t}{R}=0.01$	$\frac{t}{R}=0.001$	$\frac{t}{R}=0.0001$
E_m	5.12074	0.0609105	0.00700619
E_s	0.168618	0.00196899	4.09061×10^{-5}
E_f	34.6458	34.5506	34.5278
E_p	-0.000154659	-3.1463×10^{-7}	-1.84628×10^{-8}
E_{sc}	39.935	34.6135	34.5349

TABLE 7.10. Energy values for the penalized version.

7.2. **Second example.** In the second example, we are concerned with another issue, namely, the propagation of singularities along the characteristic curves, which in the context of the linear Koiter shell model was rigorously justified and numerically illustrated in [21]. Note that the results proved in [21] claim that although the Koiter model for fixed thickness is an elliptic problem, the limit problem when $t\to 0$ has the same nature of that the considered surface and therefore the limit problem may exhibits propagation of singularities along the asymptotic directions, if the considered surface is parabolic or the characteristic curves if the surface is hyperbolic. Those singularities are much more important then the boundary layers due to the boundary conditions. We consider the same shell but we consider the edge $\{y=0\}$ as the clamped edge. In our test we use the same loading f than in the previous test but it applied only on a part of the shell ▲ defined as follows (see Figure 7.3)

$$\blacktriangle=\{(x,y)\in\omega : 0\leqslant y\leqslant \min\{\frac{x}{2R_0},-\frac{x}{2R_0}\}+\frac{1}{2},\ \text{for}\ -R_0\leqslant x\leqslant R_0\}.$$

So, the loading f is defined as follows:

$$f=\begin{cases} t^3\times q\times\cos(2y)a_3,\ \text{if}\ (x,y)\in\ \blacktriangle\ , \\ 0 \qquad \text{elsewhere.}\end{cases} \tag{7.1}$$

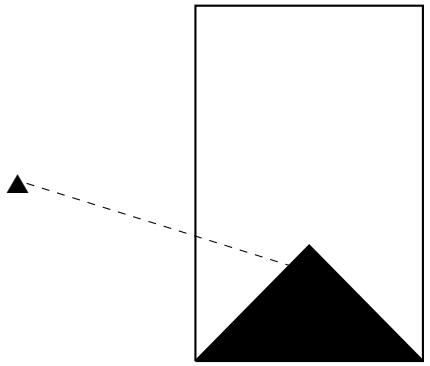


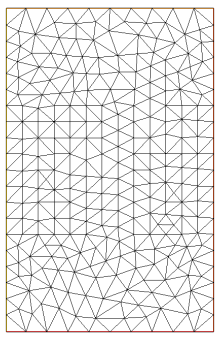
FIGURE 7.3. The region ▲

This kind of loading will generate singularities along the curves:

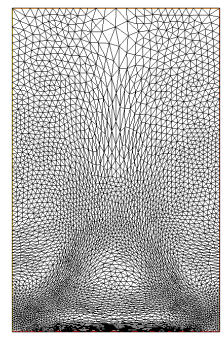
$$y = \frac{x}{2R_0} + \frac{1}{2}; \quad 0 \leqslant y \leqslant 1/2 \quad \text{and} \quad y = -\frac{x}{2R_0} + \frac{1}{2}; \quad 0 \leqslant y \leqslant 1/2,$$

which may imply the appearance of internal layers (in the interior of the domain) for small values of the thickness. So, there are three regions of particular interest, the clamped edge, the lines $\partial\blacktriangle \cap \omega$ and the asymptotic directions. Note that the asymptotic directions for the considered surface are curves of the form $x = Cte$. In this test we consider the values of thickness $t = 0.01$, $t = 0.001$, $t = 0.0001$ and $t = 0.00001$. Our objective is to compare the internal and the boundary layers for the considered example when $t \to 0$. Note that for the Koiter shell model it is shown in [21] that internal layers are more important then boundary layers for very small values of the thickness.

For the value of thickness $t = 0.01$, after nine iterations the obtained adapted mesh is shown in Figure 7.4 (B). We observe that, for this value of t, the internal



(A) Initial mesh

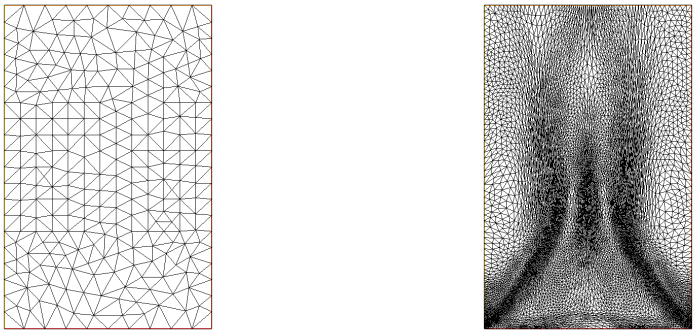


(B) Adapted mesh after nine iterations

FIGURE 7.4. $t = 0.01$

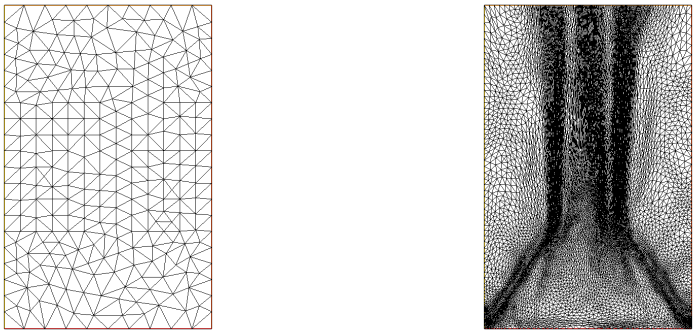
layers on $\partial\blacktriangle \cap \omega$ are less than the boundary layers on the clamped edge and

we do not observe any concentration of the energy on the asymptotic directions. For $t = 0.001$ the internal layers ($\partial\blacktriangle \cap \omega$ and the asymptotic directions) and the boundary layers are relatively more important then that on the clamped edge . The energy is concentrated in the vicinity of $\partial\blacktriangle \cap \omega$ (see Figure 7.5). Whereas, for the value of thickness $t = 0.0001$, after nine iterations we observe in Figure 7.6 that the internal layers on the asymptotic directions are clearly more important than the boundary layers and also than the internal boundary layers on $\partial\blacktriangle \cap \omega$. This may explain that the elliptic nature of the problem for a fixed t may be influenced by the type of the surface, which is parabolic for the considered example, as t tends to 0.



(A) Initial mesh (B) Adapted mesh after nine iterations

FIGURE 7.5. $t = 0.001$



(A) Initial mesh (B) Adapted mesh after nine iterations

FIGURE 7.6. $t = 0.0001$

Acknowledgements

The authors acknowledge the support of the LMA Laboratory of Ouargla University, 30000 Algeria, and Directorate-General of Scientic Research and Technological Development (DGRSDT) who supports the Laboratory. Our acknowledgement goes also to the Editor and the Referee of the journal for their illuminative guidance.

References

[1] K.-J. Bathe and E. Dvorkin. A formulation of general shell elements - the use of mixed interpolation of tensorial components. *International Journal for Numerical Methods in Engineering*, 22, 1986.

[2] C. Bernardi, A. Blouza, F. Hecht, and H. Le Dret. A posteriori analysis of finite element discretizations of a Naghdi shell model. *IMA J. Numer. Anal.*, 33(1):190–211, 2013.

[3] A. Blouza. Une formulation hybride du modèle de coque de Naghdi. *Comptes Rendus Mathematique*, 351(7-8):317–321, 2013.

[4] A. Blouza, F. Hecht, and H. Le Dret. Two finite element approximations of Naghdi's shell model in Cartesian coordinates. *SIAM J. Numer. Anal.*, 44(2):636–654, 2006.

[5] C. Carstensen. Residual-based a posteriori error estimate for a nonconforming reissner-mindlin plate finite element. *SIAM J. Numer. Anal.*, 39(6):2034–2044, 2002.

[6] C. Carstensen and J. Schöberl. Residual-based a posteriori error estimate for a mixed reissner-mindlin plate finite element method. *Numerische Mathematik*, 103(2):225–250, 2006.

[7] D. Chapelle and K. J. Bathe. *The finite element analysis of shells-fundamentals*. Computational Fluid and Solid Mechanics. Springer, 2nd edition, 2011.

[8] P. Ciarlet. *The finite element method for elliptic problems*. North–Holland, Amsterdam, 1978.

[9] P. G. Ciarlet. *Mathematical elasticity. Vol. III*, volume 29 of *Studies in Mathematics and its Applications*. North-Holland Publishing Co., Amsterdam, 2000. Theory of shells.

[10] P. G. Ciarlet. An introduction to differential geometry with application to elasticity. *J. Elasticity*, 78/79(1-3):iv+215, 2005.

[11] P. Clément. Approximation by finite element functions using local regularization. *Rev. Française Automat. Informat. Recherche Opérationnelle Sér.*, 9(R-2):77–84, 1975.

[12] W. Dörfler. A convergent adaptive algorithm for poisson's equation. *SIAM J. Numer. Anal.*, 33, 06 1996.

[13] V. Girault and R. P.-A. *Finite element methods for Navier-Stokes equations*, volume 5 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 1986. Theory and algorithms.

[14] T. Grätsch and K.-J. Bathe. Influence functions and goal-oriented error estimation for finite element analysis of shell structures. *International journal for numerical methods in engineering*, 63(5):709–736, 2005.

[15] T. Grätsch and K.-J. Bathe. A posteriori error estimation techniques in practical finite element analysis. *Computers and structures*, 83(4-5):235–265, 2005.

[16] F. Hecht. New development in freefem++. *Journal of numerical mathematics*, 20(3-4):251–266, 2012.

[17] M. Marohnić and J. Tambača. On a model of a flexural prestressed shell. *Math. Methods Appl. Sci.*, 38(18):5231–5241, 2015.

[18] B. Maury. Numerical analysis of finite element/volume penalty method. *SIAM J. Numer. Anal.*, 47:1126–1148, 2009.

[19] I. Merabet and S. Nicaise. A penalty method for a linear Koiter shell model. *ESAIM Math. Model. Numer. Anal.*, 51(5):1783–1803, 2017.

[20] R. Rezzag Bara, S. Nicaise, and I. Merabet. Finite element approximation of a prestressed shell model. *Math. Methods Appl. Sci.*, 43(7):4336–4352, 2020.

[21] E. Sanchez-Palencia, O. Millet, and F. Béchet. *Singular Problems in Shell Theory Anisotropic Error Estimates in the Layers*, volume 45 of *Lecture Notes in Applied and Computational Mechanics*. Springer, 2010.

[22] R. Verfürth. *A review of a posteriori error estimation and adaptive mesh-refinement techniques*. Wiley–Teubner Series Advances in Numerical Mathematics. Wiley–Teubner, Chichester, Stuttgart, 1996.

Manuscript received 17th August 2021,
revised 20th May 2022,
accepted 5th November 2022.

Serge NICAISE
Université polytechnique Hauts-de-France, LAMAV, FR CNRS 2037, F-59313 - Valenciennes Cedex 9 France

Serge.Nicaise@univ-valenciennes.fr (Corresponding author)

Ismael MERABET
Laboratoire de Mathématiques Appliquées, Université Kasdi Merbah, BP 511, Ouargla 30000, Algeria

merabetsmail@gmail.com

Rihana REZZAG BARA
Laboratoire de Mathématiques Appliquées, Université Kasdi Merbah, BP 511, Ouargla 30000, Algeria

rihana.rezzag@gmail.com