IMA Journal of Numerical Analysis (2018) **38**, 921–954
doi: 10.1093/imanum/drx024
Advance Access publication on June 19, 2017

Residual-based *a posteriori* error estimation for contact problems approximated by Nitsche's method

FRANZ CHOULY
Laboratoire de Mathématiques de Besançon - UMR CNRS 6623, Université Bourgogne Franche Comté, 16 route de Gray, 25030 Besançon Cedex, France
franz.chouly@univ-fcomte.fr

MATHIEU FABRE*
EPFL SB MATHICSE (Bt. MA), Station 8, CH 1015 Lausanne, Switzerland and Istituto di Matematica Applicata e Tecnologie, Informatiche 'E. Magenes' del CNR, via Ferrata 1, 27100 Pavia, Italy
*Corresponding author: mathieu.fabre@epfl.ch

PATRICK HILD
Institut de Mathématiques de Toulouse - UMR CNRS 5219, Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse Cedex 9, France
patrick.hild@math.univ-toulouse.fr

AND

JÉRÔME POUSIN AND YVES RENARD
Université de Lyon, CNRS, INSA-Lyon, ICJ UMR5208, LaMCoS UMR5259, F-69621 Villeurbanne, France
jerome.pousin@insa-lyon.fr yves.renard@insa-lyon.fr

[Received on 27 May 2015; revised on 24 March 2017]

We introduce a residual-based *a posteriori* error estimator for contact problems in two- and three-dimensional linear elasticity, discretized with linear and quadratic finite elements and Nitsche's method. Efficiency and reliability of the estimator are proved under a saturation assumption. Numerical experiments illustrate the theoretical properties and the good performance of the estimator.

Keywords: unilateral contact; finite elements; Nitsche's method; *a posteriori* error estimates; residuals.

1. Introduction

The computations of contact problems between deformable bodies are usually obtained with the finite element method (Laursen, 2003; Wriggers, 2006). An important aspect for the user is to quantify the quality of the simulations by evaluating the discretization errors coming from the finite element approximation. This quantification requires the definition of *a posteriori* error estimators that can be of different types (residual based, equilibrated fluxes, smoothing of the stress fields, etc.). The main aim of the estimators is to furnish some information on the local error in order to adapt or refine the mesh and to reduce the computational costs.

Among the finite element discretizations for contact problems, a recent effort was devoted to Nitsche's method, which can be seen as a consistent penalty formulation with only one primal unknown (like the variational inequality formulation or like the penalty method): the displacement field. In contrast to Nitsche's method, the Lagrange (stabilized or standard) methods admit the contact pressure as a supplementary unknown. Nitsche's method was introduced for the frictionless unilateral contact problem in a simple (symmetric) form in Chouly & Hild (2013), while investigated in a generalized and numerical form in Chouly *et al.* (2015). In the latter references, the theoretical results deal with well-posedness of the discrete problems and *a priori* error estimates in two- and three-dimensional spaces with linear and quadratic finite elements. A generalization to frictional contact problems is carried out in Chouly (2014) (numerical analysis for Tresca friction) and Renard (2013) (numerical study for Coulomb friction). To our knowledge, the *a posteriori* quantification of the discretization errors committed by the Nitsche finite element approximation has not been considered for unilateral contact problems up to now.

Nevertheless, there are several studies concerning *a posteriori* error analyses for frictionless or frictional contact problems in Buscaglia *et al.* (2001), Carstensen *et al.* (1999), Fernández & Hild (2010), Lee & Oden (1994) and Wriggers & Scherf (1998) (residual approach using a penalization of the contact condition or the normal compliance law); in Coorevits *et al.* (2000, 2001), Louf *et al.* (2003), Weiss & Wohlmuth (2009) and Wohlmuth (2007) (equilibrated residual method); in Banz & Stephan (2015), Eck & Wendland (2003) and Maischak & Stephan (2005) (residual approach for BEM-discretizations); in Blum & Suttmeier (2000) and Schröder & Rademacher (2011) (error technique measure developed in Becker & Rannacher, 1996). Moreover, a residual-type estimator for the Signorini problem in its common formulation (variational inequality or mixed method) can be found in Bostan & Han (2006), Hild & Nicaise (2005, 2007) and Schröder (2012) and in the recent work by Krause *et al.* (2015).

Finally, we mention that only a few works are devoted to *a posteriori* error estimates for Nitsche's method, and all concern linear boundary/interface conditions. For interface conditions and elliptic problems, Hansbo & Hansbo (2002) introduce a residual-type estimator for a Nitsche unfitted treatment of the interface condition. They prove an upper bound on a linear functional of the error, in the spirit of Becker & Rannacher (2001). Note as well an early work of Becker (2002) in the context of optimal control for Navier–Stokes equations, with a Nitsche treatment of Dirichlet boundary conditions and an *a posteriori* error estimate for the functional to minimize. Residual error estimates are introduced as well by Becker, Hansbo and Stenberg in Becker *et al.* (2003) for a Nitsche-based domain decomposition with nonmatching meshes. Upper bounds in both the H^1- and L^2-norms are established, with the help of a saturation assumption (as in Wohlmuth, 1999) for the H^1-norm. In the context of composite grids, two variants of residual-based error estimates are proposed by Hansbo *et al.* (2003). Upper bounds in the H^1-norm without any saturation assumption are proposed for both of them. Later, Juntunen & Stenberg (2008) provide a residual-based error estimator for the stabilized Bassi–Rebay discontinuous Galerkin method that relies on Nitsche's treatment of continuity. Upper and lower bounds are proved for this method. The same authors in Juntunen & Stenberg (2009) introduce a Nitsche method for a general boundary condition and an associated residual error estimator. They prove an upper bound in the H^1-norm under a saturation assumption (as in Braess & Verfürth, 1996), and they establish a lower bound too. Finally, let us mention two recent papers on the Brinkman problem by Juntunen & Stenberg (2010) and Könnö & Stenberg (2011).

The paper is outlined as follows. In Section 2, the Nitsche finite element discretization for contact problems in linear elasticity is described and results dealing with well-posedness are recalled from Chouly *et al.* (2015). In Section 3, a residual *a posteriori* error estimator is introduced, and we prove its reliability and efficiency. In Section 4, numerical experiments in two- and three-dimensional spaces illustrate the theoretical results and allow us to assess the quality of the estimator for different values of the numerical parameters.

Let us introduce some useful notation. In what follows, bold letters like $\mathbf{u}, \mathbf{v}$ indicate vector- or tensor-valued quantities, while the capital ones (e.g., $\mathbf{V}, \mathbf{K}$) represent functional sets involving vector fields. As usual, we denote by $(H^s(\cdot))^d$, $s \in \mathbb{R}, d = 1, 2, 3$, the Sobolev spaces in one, two or three space dimensions (see Adams, 1975), with the convention $H^0 = L^2$. The usual norm (respectively seminorm) of $(H^s(D))^d$ is denoted by $\|\cdot\|_{s,D}$ (respectively $|\cdot|_{s,D}$), and we keep the same notation for any $d = 1, 2, 3$. In the sequel, the symbol $|\cdot|$ will either denote the Euclidean norm in $\mathbb{R}^d$ or the measure of a domain in $\mathbb{R}^d$. The letter C stands for a generic constant, independent of the discretization parameters.

For two scalar quantities a and b, the notation $a \lesssim b$ means there exists a constant C, independent of the mesh-size parameters and of the Nitsche parameter γ_0 (see Section 2.2), such that $a \le Cb$. Moreover, $a \sim b$ means that $a \lesssim b$ and $b \lesssim a$.

2. Setting

2.1 *The unilateral contact problem*

We consider an elastic body whose reference configuration is represented by the polygonal or polyhedral domain Ω in $\mathbb{R}^d$, with $d = 2$ or $d = 3$. Small strain assumptions are made, as well as plane strain, when $d = 2$. The boundary $\partial\Omega$ of Ω consists of three nonoverlapping parts Γ_{D}, Γ_{N} and the (potential) contact boundary Γ_{C}, with meas$(\Gamma_{\mathrm{D}}) > 0$ and meas$(\Gamma_{\mathrm{C}}) > 0$. The (potential) contact boundary is supposed to be a straight line segment when $d = 2$ or a planar polygon when $d = 3$ to simplify. The unit outward normal vector on $\partial\Omega$ is denoted $\mathbf{n}$. In its initial stage, the body shows a gap on Γ_{C} with a rigid foundation (the extension to two elastic bodies in contact can be easily made, at least for small strain models). The non-negative gap function is denoted by g, and we assume that g is continuous on $\overline{\Gamma}_{\mathrm{C}}$. We suppose that the unknown final contact zone after deformation will be included in Γ_{C}. The body is clamped on Γ_{D} for the sake of simplicity. It is subjected to volume forces $\mathbf{f} \in (L^2(\Omega))^d$ and to surface loads $\mathbf{F} \in (L^2(\Gamma_{\mathrm{N}}))^d$.

The unilateral contact problem in linear elasticity consists of finding the displacement field $\mathbf{u} : \Omega \to \mathbb{R}^d$ verifying the equations and conditions (2.1)–(2.2):

$$\begin{aligned} \mathbf{div}\,\boldsymbol{\sigma}(\mathbf{u}) + \mathbf{f} &= \mathbf{0} && \text{in } \Omega, \\ \boldsymbol{\sigma}(\mathbf{u}) &= \mathbf{A}\,\boldsymbol{\varepsilon}(\mathbf{u}) && \text{in } \Omega, \\ \mathbf{u} &= \mathbf{0} && \text{on } \Gamma_{\mathrm{D}}, \\ \boldsymbol{\sigma}(\mathbf{u})\mathbf{n} &= \mathbf{F} && \text{on } \Gamma_{\mathrm{N}}, \end{aligned} \tag{2.1}$$

where $\boldsymbol{\sigma} = (\sigma_{ij}),\ 1 \le i, j \le d$ stands for the stress tensor field and $\mathbf{div}$ denotes the divergence operator of tensor-valued functions. The notation $\boldsymbol{\varepsilon}(\mathbf{v}) = (\boldsymbol{\nabla}\mathbf{v} + \boldsymbol{\nabla}\mathbf{v}^{\mathrm{T}})/2$ represents the linearized strain tensor field and $\mathbf{A}$ is the fourth-order symmetric elasticity tensor having the usual uniform ellipticity and boundedness property. For any displacement field $\mathbf{v}$ and for any density of surface forces $\boldsymbol{\sigma}(\mathbf{v})\mathbf{n}$ defined on $\partial\Omega$, we adopt the following notation:

$$\mathbf{v} = v_n\mathbf{n} + \mathbf{v}_{\mathbf{t}} \quad \text{and} \quad \boldsymbol{\sigma}(\mathbf{v})\mathbf{n} = \sigma_n(\mathbf{v})\mathbf{n} + \boldsymbol{\sigma}_{\mathbf{t}}(\mathbf{v}),$$

where $\mathbf{v_t}$ (respectively $\boldsymbol{\sigma}_{\mathbf{t}}(\mathbf{v})$) is the tangential component of $\mathbf{v}$ (respectively $\boldsymbol{\sigma}(\mathbf{v})\mathbf{n}$). The conditions describing unilateral contact without friction on Γ_{C} are

$$\begin{aligned} &\text{(i)} & u_n &\leq g, \\ &\text{(ii)} & \sigma_n(\mathbf{u}) &\leq 0, \\ &\text{(iii)} & \sigma_n(\mathbf{u})\,(u_n - g) &= 0, \\ &\text{(iv)} & \sigma_{\mathbf{t}}(\mathbf{u}) &= 0. \end{aligned} \tag{2.2}$$

We introduce the Hilbert space $\mathbf{V}$ and the convex cone $\mathbf{K}$ of admissible displacements that satisfy non-interpenetration on the contact zone Γ_{C}:

$$\mathbf{V} := \left\{\mathbf{v} \in (H^1(\Omega))^d \;:\; \mathbf{v} = \mathbf{0} \text{ on } \Gamma_{\mathrm{D}}\right\}, \quad \mathbf{K} := \{\mathbf{v} \in \mathbf{V} \;:\; v_n = \mathbf{v}\cdot\mathbf{n} \leq g \text{ on } \Gamma_{\mathrm{C}}\}.$$

We define as well

$$a(\mathbf{u},\mathbf{v}) := \int_{\Omega} \boldsymbol{\sigma}(\mathbf{u}) : \boldsymbol{\varepsilon}(\mathbf{v})\,\mathrm{d}\Omega, \quad L(\mathbf{v}) := \int_{\Omega} \mathbf{f}\cdot\mathbf{v}\,\mathrm{d}\Omega + \int_{\Gamma_{\mathrm{N}}} \mathbf{F}\cdot\mathbf{v}\,\mathrm{d}\Gamma$$

for any $\mathbf{u}$ and $\mathbf{v}$ in $\mathbf{V}$. From the previous assumptions, we deduce that $a(\cdot,\cdot)$ is bilinear, symmetric, $\mathbf{V}$-elliptic and continuous on $\mathbf{V}\times\mathbf{V}$. Likewise we observe that $L(\cdot)$ is a continuous linear form on $\mathbf{V}$. The weak formulation of problem (2.1)–(2.2), as a variational inequality (see Fichera, 1963/1964; Kikuchi & Oden, 1988; Haslinger *et al.*, 1996), reads

$$\begin{cases} \text{find } \mathbf{u} \in \mathbf{K} \text{ such that} \\ a(\mathbf{u}, \mathbf{v}-\mathbf{u}) \geq L(\mathbf{v}-\mathbf{u}) \qquad \forall\, \mathbf{v} \in \mathbf{K}. \end{cases} \tag{2.3}$$

Stampacchia's theorem ensures that problem (2.3) admits a unique solution.

2.2 *Finite element setting and Nitsche-based method*

To approximate problem (2.3), we fix a family of meshes $(T_h)_{h>0}$, regular in Ciarlet's sense (see Ciarlet, 1991), made of closed elements and assumed to be subordinated to the decomposition of the boundary $\partial\Omega$ into Γ_{D}, Γ_{N} and Γ_{C}. For $K \in T_h$, we recall that h_K is the diameter of K and $h := \max_{K\in T_h} h_K$. The regularity of the mesh implies notably that for any edge (or face when $d = 3$) E of K one has $h_E := |E| \sim h_K$.

Let us define E_h as the set of edges (or faces when $d = 3$) of the triangulation and define $E_h^{\mathrm{int}} = \{E \in E_h : E \subset \Omega\}$ as the set of interior edges/faces of T_h (the edges/faces are supposed to be relatively open). We denote by $E_h^{\mathrm{N}} = \{E \in E_h : E \subset \Gamma_{\mathrm{N}}\}$ the set of boundary edges/faces that correspond to Neumann conditions, and similarly, $E_h^{\mathrm{C}} = \{E \in E_h : E \subset \Gamma_{\mathrm{C}}\}$ is the set of boundary edges/faces included in the contact boundary.

For an element K, we denote by E_K the set of edges/faces of K and according to the above notation, we set $E_K^{\mathrm{int}} = E_K \cap E_h^{\mathrm{int}}$, $E_K^{\mathrm{N}} = E_K \cap E_h^{\mathrm{N}}$, $E_K^{\mathrm{C}} = E_K \cap E_h^{\mathrm{C}}$. For an edge/face E of an element K, introduce $\boldsymbol{\nu}_{K,E}$, the unit outward normal vector to K along E. Furthermore, for each edge/face E, we fix one of the two normal vectors and denote it by $\boldsymbol{\nu}_E$. The jump of some vector-valued function $\boldsymbol{v}$ across an edge/face $E \in E_h^{\mathrm{int}}$ at a point $\boldsymbol{y} \in E$ is defined as

$$[\![\boldsymbol{v}]\!]_E(\boldsymbol{y}) = \lim_{\alpha\to 0^+} \boldsymbol{v}(\boldsymbol{y} + \alpha\boldsymbol{\nu}_E) - \boldsymbol{v}(\boldsymbol{y} - \alpha\boldsymbol{\nu}_E).$$

Note that the sign of $[\![\boldsymbol{v}]\!]_E$ depends on the orientation of $\boldsymbol{\nu}_E$. Finally, we will need local subdomains (also called patches). As usual, let ω_K be the union of all elements having a nonempty intersection with K. Similarly for a node $\boldsymbol{x}$ and an edge/face E, let $\omega_x = \cup_{K:x\in K} K$ and $\omega_E = \cup_{x\in \overline{E}} \omega_x$.

The chosen finite element space $\mathbf{V}^h \subset \mathbf{V}$ involves standard Lagrange finite elements of degree k, with $k = 1$ or $k = 2$ (see Ciarlet, 1991; Ern & Guermond, 2004; Brenner & Scott, 2007), i.e.,

$$\mathbf{V}^h := \left\{ \mathbf{v}^h \in (\mathscr{C}^0(\overline{\Omega}))^d : \mathbf{v}^h_{|K} \in (P_k(K))^d \; \forall K \in T_h, \mathbf{v}^h = \mathbf{0} \text{ on } \Gamma_{\mathrm{D}} \right\}.$$

Let us introduce the notation $[\cdot]_+$ for the positive part of a scalar quantity $a \in \mathbb{R}$: $[a]_+ = a$ if $a > 0$ and $[a]_+ = 0$ otherwise. The monotonicity property below holds:

$$([a]_+ - [b]_+)(a - b) \geq ([a]_+ - [b]_+)^2 \geq 0. \tag{2.4}$$

Note that condition (2.4) can be straightforwardly extended to real-valued functions.

Let γ be a positive piecewise constant function on the contact interface Γ_{C}, that satisfies

$$\gamma\big|_{K\cap\Gamma_{\mathrm{C}}} = \gamma_0 h_K$$

for every K that has a nonempty intersection of dimension $d-1$ with Γ_{C}, and where γ_0 is a positive given constant that we call the ‘Nitsche parameter’. Note that the value of γ on element intersections has no influence.

We introduce the discrete linear operator

$$P_\gamma : \begin{array}{ccc} \mathbf{V}^h & \to & L^2(\Gamma_{\mathrm{C}}), \\ \mathbf{v}^h & \mapsto & v^h_n - \gamma\, \sigma_n(\mathbf{v}^h), \end{array}$$

and the bilinear form where $\theta \in \mathbb{R}$ is a fixed parameter:

$$A_{\theta\gamma}(\mathbf{u}^h, \mathbf{v}^h) := a(\mathbf{u}^h, \mathbf{v}^h) - \int_{\Gamma_{\mathrm{C}}} \theta\gamma\, \sigma_n(\mathbf{u}^h)\sigma_n(\mathbf{v}^h)\, \mathrm{d}\Gamma.$$

Our Nitsche-based method then reads

$$\begin{cases} \text{find } \mathbf{u}^h \in \mathbf{V}^h \text{ such that} \\ A_{\theta\gamma}(\mathbf{u}^h, \mathbf{v}^h) + \displaystyle\int_{\Gamma_{\mathrm{C}}} \frac{1}{\gamma}\, [P_\gamma(\mathbf{u}^h) - g]_+ P_{\theta\gamma}(\mathbf{v}^h)\, \mathrm{d}\Gamma = L(\mathbf{v}^h) \quad \forall\, \mathbf{v}^h \in \mathbf{V}^h. \end{cases} \tag{2.5}$$

We consider the quasi-interpolation (regularization) operator introduced in, e.g., Bernardi & Girault (1998, formula (4.11)) and its straightforward extension to the vectorial case, which we denote $R^h : \mathbf{V} \to \mathbf{V}^h$. This operator has the following approximation and stability properties.

LEMMA 2.1 For any $\boldsymbol{v} \in \mathbf{V} \cap (H^l(\Omega))^d$, $1 \leq l \leq k+1$, the following estimates hold:

$$\|\boldsymbol{v} - R^h\boldsymbol{v}\|_{0,K} \lesssim h_K^l |\boldsymbol{v}|_{l,\omega_K} \quad \forall\, K \in T_h, \tag{2.6}$$

$$\|\boldsymbol{v} - R^h\boldsymbol{v}\|_{0,E} \lesssim h_E^{l-1/2} |\boldsymbol{v}|_{l,\omega_E} \quad \forall\, E \in E_h. \tag{2.7}$$

Moreover, R^h is stable in the H^1-norm, i.e.,

$$\|R^h \boldsymbol{v}\|_{1,\Omega} \lesssim \|\boldsymbol{v}\|_{1,\Omega} \quad \forall\, \boldsymbol{v} \in \mathbf{V}. \tag{2.8}$$

Proof. Estimates (2.6) and (2.7) are provided in Bernardi & Girault (1998, Theorem 4.8, Remark 8). The stability of R^h in the H^1-norm is proved in Bernardi & Girault (1998, Theorem 4.4) (in all cases, it suffices to apply the results of Bernardi & Girault, 1998 componentwise). □

We next define a convenient mesh-dependent norm that is in fact a weighted $L^2(\Gamma_{\mathrm{C}})$-norm (since $\gamma/\gamma_0 = h_K$).

Definition 2.2 For any $v \in L^2(\Gamma_{\mathrm{C}})$, we set

$$\|v\|_{-1/2,h,\Gamma_{\mathrm{C}}} := \left\| \left(\frac{\gamma}{\gamma_0}\right)^{1/2} v \right\|_{0,\Gamma_{\mathrm{C}}}.$$

We end this subsection with a discrete trace inequality that will be useful for the analysis (for the proof, see, e.g., Chouly *et al.*, 2015).

Lemma 2.3 For any $\mathbf{v}^h \in \mathbf{V}^h$, we have

$$\|\sigma_n(\mathbf{v}^h)\|_{-1/2,h,\Gamma_{\mathrm{C}}} \lesssim \|\mathbf{v}^h\|_{1,\Omega}. \tag{2.9}$$

2.3 *Consistency and well-posedness of the Nitsche-based method*

We recall two theoretical properties for the Nitsche-based method (2.5): consistency and well-posedness. These properties, together with optimal *a priori* error estimates in the $H^1(\Omega)$-norm, are proved in Chouly *et al.* (2015) in the particular case of a zero gap function (i.e., $g = 0$).

Like Nitsche's method for second-order elliptic problems with Dirichlet boundary conditions or domain decomposition (Becker *et al.*, 2003), our Nitsche-based formulation (2.5) for unilateral contact is consistent.

Lemma 2.4 The Nitsche-based method for contact is consistent: suppose that the solution $\mathbf{u}$ of (2.1)–(2.2) lies in $(H^{3/2+\nu}(\Omega))^d$ with $\nu > 0$ and $d = 2, 3$. Then $\mathbf{u}$ is also a solution of

$$A_{\theta\gamma}(\mathbf{u}, \mathbf{v}^h) + \int_{\Gamma_{\mathrm{C}}} \frac{1}{\gamma}\, [P_\gamma(\mathbf{u}) - g]_+ P_{\theta\gamma}(\mathbf{v}^h)\, \mathrm{d}\Gamma = L(\mathbf{v}^h) \quad \forall\, \mathbf{v}^h \in \mathbf{V}^h.$$

Problem (2.5) is well posed in the following sense and under the assumptions below.

Theorem 2.5 Suppose that either $\theta \neq -1$ and $\gamma_0 > 0$ is sufficiently small, or $\theta = -1$ and $\gamma_0 > 0$. Then problem (2.5) admits one unique solution $\mathbf{u}^h$ in $\mathbf{V}^h$.

Lemma 2.4 and Theorem 2.5 are obtained exactly as in Chouly *et al.* (2015), for any gap function g, by noting that the contact conditions (2.2) (*i*)–(*iii*) are equivalent to

$$\sigma_n(\mathbf{u}) = -\frac{1}{\gamma}[u_n - g - \gamma\sigma_n(\mathbf{u})]_+.$$

REMARK 2.6 When γ_0 is large and $\theta \neq -1$, we can conclude neither to uniqueness nor to existence of a solution. In reference Chouly *et al.* (2015), there are some simple explicit examples of nonexistence and nonuniqueness of solutions.

3. *A posteriori* error analysis

3.1 *Definition of the residual error estimator*

The element residual of the equilibrium equation in (2.1) is defined by

$$\mathbf{div}\,\boldsymbol{\sigma}(\mathbf{u}^h) + \mathbf{f} \quad \text{in } K.$$

REMARK 3.1 For linear elements ($k = 1$), the term $\mathbf{div}\,\boldsymbol{\sigma}(\mathbf{u}^h)$ vanishes.

As usual, this element residual can be replaced by some finite-dimensional approximation, called an approximate element residual (see, e.g., Ainsworth & Oden, 2000),

$$\mathbf{div}\,\boldsymbol{\sigma}(\mathbf{u}^h) + \mathbf{f}_K, \quad \mathbf{f}_K \in (P_l(K))^d, \quad l \geq 0.$$

A current choice is to take $\mathbf{f}_K = \int_K \mathbf{f}(\boldsymbol{x})\,/|K|\,\mathrm{d}\boldsymbol{x}$ since for $\mathbf{f} \in (H^1(\Omega))^d$, scaling arguments yield $\|\mathbf{f} - \mathbf{f}_K\|_{0,K} \lesssim h_K\|\mathbf{f}\|_{1,K}$, and it is then negligible with respect to the estimator η defined hereafter. In the same way, $\mathbf{F}$ is approximated by a computable quantity denoted $\mathbf{F}_E$ on any $E \in E_h^{\mathrm{N}}$ and the gap g is computed using an approximation denoted g_{C}.

DEFINITION 3.2 The local error estimators η_K and the global estimator η are defined by

$$\eta_K = \left(\sum_{i=1}^{4}\eta_{iK}^2\right)^{1/2},$$

$$\eta_{1K} = h_K\|\mathbf{div}\,\boldsymbol{\sigma}(\mathbf{u}^h) \;+\; \mathbf{f}_K\|_{0,K},$$

$$\eta_{2K} = h_K^{1/2}\left(\sum_{E\in E_K^{\mathrm{int}}\cup E_K^{\mathrm{N}}}\|J_{E,n}(\mathbf{u}^h)\|_{0,E}^2\right)^{1/2},$$

$$\eta_{3K} = h_K^{1/2}\left(\sum_{E\in E_K^{\mathrm{C}}} \|\boldsymbol{\sigma}_{\mathbf{t}}(\mathbf{u}^h)\|_{0,E}^2\right)^{1/2},$$

$$\eta_{4K} = h_K^{1/2}\left(\sum_{E\in E_K^{\mathrm{C}}} \left\|\frac{1}{\gamma}[P_\gamma(\mathbf{u}^h) - g_{\mathrm{C}}]_+ + \sigma_n(\mathbf{u}^h)\right\|_{0,E}^2\right)^{1/2},$$

$$\eta = \left(\sum_{K\in T_h} \eta_K^2\right)^{1/2},$$

where $J_{E,n}(\mathbf{u}^h)$ means the constraint jump of $\mathbf{u}^h$ in the normal direction, i.e.,

$$J_{E,n}(\mathbf{u}^h) = \begin{cases} [\![\boldsymbol{\sigma}(\mathbf{u}^h)\boldsymbol{\nu}_E]\!]_E & \forall E \in E_h^{\mathrm{int}}, \\ \boldsymbol{\sigma}(\mathbf{u}^h)\boldsymbol{\nu}_E - \mathbf{F}_E & \forall E \in E_h^{\mathrm{N}}. \end{cases} \tag{3.1}$$

The local and global approximation terms are given by

$$\zeta_K = \left(h_K^2 \sum_{K'\subset\omega_K} \|\mathbf{f} - \mathbf{f}_{K'}\|_{0,K'}^2 + h_K \sum_{E\subset E_K^{\mathrm{N}}} \|\mathbf{F} - \mathbf{F}_E\|_{0,E}^2 + \frac{1}{\gamma_0^2 h_K} \sum_{E\subset E_K^{\mathrm{C}}} \|\mathbf{g} - \mathbf{g}_{\mathrm{C}}\|_{0,E}^2\right)^{1/2},$$

$$\zeta = \left(\sum_{K\in T_h} \zeta_K^2\right)^{1/2}.$$

3.2 *Upper error bound*

First, we state a ‘saturation’ assumption that we need in order to prove the estimate (see also Becker *et al.*, 2003 in the case of Nitsche for domain decomposition, and Wohlmuth, 1999 for mortar methods).

ASSUMPTION 3.3 The solution $\mathbf{u}$ of (2.3) and the discrete solution $\mathbf{u}^h$ of (2.5) are such that

$$\left\|\sigma_n(\mathbf{u} - \mathbf{u}^h)\right\|_{-1/2,h,\Gamma_{\mathrm{C}}} \lesssim \|\mathbf{u} - \mathbf{u}^h\|_{1,\Omega}. \tag{3.2}$$

REMARK 3.4 Note that for a Nitsche treatment of (linear) interface conditions, an upper bound for a residual-based estimator has been derived without such an assumption in Hansbo *et al.* (2003). Similarly for some classes of mixed nonconforming finite element approximations, an assumption such as Assumption 3.3 has been revealed to be superfluous; see, e.g., Carstensen (1997) and Kim (2007). However, for method (2.5), the derivation of an upper bound without such a saturation assumption remains an open issue.

The following statement guarantees the reliability of the *a posteriori* error estimator given in Definition 3.2.

THEOREM 3.5 Let $\mathbf{u}$ be the solution to the variational inequality (2.3), with $\mathbf{u} \in (H^{3/2+\nu}(\Omega))^d$ ($\nu > 0$ and $d = 2, 3$), and let $\mathbf{u}^h$ be the solution to the corresponding discrete problem (2.5). Assume that for $\theta \neq -1$, γ_0 is sufficiently small, and otherwise that $\gamma_0 > 0$ for $\theta = -1$. Assume that the saturation assumption (3.2) holds as well. Then, we have

$$\|\mathbf{u} - \mathbf{u}^h\|_{1,\Omega} + \left\|\sigma_n(\mathbf{u}) + \frac{1}{\gamma}[P_\gamma(\mathbf{u}^h) - g]_+\right\|_{-1/2,h,\Gamma_C} + \|\sigma_n(\mathbf{u}) - \sigma_n(\mathbf{u}^h)\|_{-1/2,h,\Gamma_C} \lesssim (1 + \gamma_0)(\eta + \zeta).$$

Proof. Let $\mathbf{v}^h \in \mathbf{V}^h$. To lighten the notation, we define $\mathbf{e} := \mathbf{u} - \mathbf{u}^h$. We first use the $\mathbf{V}$-ellipticity of $a(\cdot, \cdot)$, together with the Green formula, equations (2.1) and (2.5) to obtain

$$\begin{aligned}
\alpha\|\mathbf{e}\|^2_{1,\Omega} &\leq a(\mathbf{u} - \mathbf{u}^h, \mathbf{u} - \mathbf{u}^h) \\
&= a(\mathbf{u}, \mathbf{u} - \mathbf{u}^h) - a(\mathbf{u}^h, \mathbf{u} - \mathbf{v}^h) - a(\mathbf{u}^h, \mathbf{v}^h - \mathbf{u}^h) \\
&= L(\mathbf{u} - \mathbf{u}^h) + \int_{\Gamma_C} \sigma_n(\mathbf{u})(u_n - u_n^h)\,\mathrm{d}\Gamma - a(\mathbf{u}^h, \mathbf{u} - \mathbf{v}^h) \\
&\quad - L(\mathbf{v}^h - \mathbf{u}^h) + \int_{\Gamma_C} \frac{1}{\gamma}[P_\gamma(\mathbf{u}^h) - g]_+ P_{\theta\gamma}(\mathbf{v}^h - \mathbf{u}^h)\,\mathrm{d}\Gamma - \theta \int_{\Gamma_C} \gamma\, \sigma_n(\mathbf{u}^h)\sigma_n(\mathbf{v}^h - \mathbf{u}^h)\,\mathrm{d}\Gamma \\
&= \mathcal{T}_1 + \mathcal{T}_2,
\end{aligned} \tag{3.3}$$

where α is the $\mathbf{V}$-ellipticity constant of $a(\cdot, \cdot)$ and

$$\begin{aligned}
\mathcal{T}_1 &:= L(\mathbf{u} - \mathbf{v}^h) - a(\mathbf{u}^h, \mathbf{u} - \mathbf{v}^h) + \int_{\Gamma_C} \frac{1}{\gamma}[P_\gamma(\mathbf{u}^h) - g]_+ (v_n^h - u_n)\,\mathrm{d}\Gamma, \\
\mathcal{T}_2 &:= \int_{\Gamma_C} \sigma_n(\mathbf{u})(u_n - u_n^h)\,\mathrm{d}\Gamma + \int_{\Gamma_C} \frac{1}{\gamma}[P_\gamma(\mathbf{u}^h) - g]_+ P_{\theta\gamma}(\mathbf{u} - \mathbf{u}^h)\,\mathrm{d}\Gamma \\
&\quad - \theta \int_{\Gamma_C} \frac{1}{\gamma}[P_\gamma(\mathbf{u}^h) - g]_+ \gamma \sigma_n(\mathbf{v}^h - \mathbf{u})\,\mathrm{d}\Gamma - \theta \int_{\Gamma_C} \gamma\, \sigma_n(\mathbf{u}^h)\sigma_n(\mathbf{v}^h - \mathbf{u}^h)\,\mathrm{d}\Gamma.
\end{aligned}$$

The quantity $\mathcal{T}_1$ is an expression that is handled hereafter in a classical way. Namely, by integrating by parts on each triangle K, using the definition of $J_{E,n}(\mathbf{u}^h)$ in (3.1) and splitting the integrals on Γ_C into normal and tangential components we get

$$\begin{aligned}
\mathcal{T}_1 &= \sum_{K \in T_h} \int_K (\mathbf{div}\,\boldsymbol{\sigma}(\mathbf{u}^h) + \mathbf{f}) \cdot (\mathbf{u} - \mathbf{v}^h)\,\mathrm{d}\Gamma \\
&\quad + \sum_{E \in E_h^C} \int_E \left(\frac{1}{\gamma}[P_\gamma(\mathbf{u}^h) - g]_+ + \sigma_n(\mathbf{u}^h)\right)(v_n^h - u_n)\,\mathrm{d}\Gamma
\end{aligned}$$

$$
+ \sum_{E \in E_h^{\mathrm{C}}} \int_E \boldsymbol{\sigma}_{\mathbf{t}}(\mathbf{u}^h) \cdot (\mathbf{v}_{\mathbf{t}}^h - \mathbf{u}_{\mathbf{t}})\, \mathrm{d}\Gamma - \sum_{E \in E_h^{\mathrm{int}} \cup E_h^{\mathrm{N}}} \int_E J_{E,n}(\mathbf{u}^h) \cdot (\mathbf{u} - \mathbf{v}^h)\ \, \mathrm{d}\Gamma
$$

$$
+ \sum_{E \in E_h^{\mathrm{N}}} \int_E (\mathbf{F} - \mathbf{F}_E) \cdot (\mathbf{u} - \mathbf{v}^h)\, \mathrm{d}\Gamma. \tag{3.4}
$$

We now need to estimate each term of this right-hand side. For that purpose, we take

$$
\mathbf{v}^h = \mathbf{u}^h + R^h(\mathbf{u} - \mathbf{u}^h), \tag{3.5}
$$

where R^h is the quasi-interpolation operator defined in Section 2.2.

We start with the integral term on elements K. The Cauchy–Schwarz inequality implies

$$
\sum_{K \in T_h} \int_K (\mathbf{div}\, \boldsymbol{\sigma}(\mathbf{u}^h)\ +\ \mathbf{f}) \cdot (\mathbf{u} - \mathbf{v}^h)\, \mathrm{d}\Gamma \le \sum_{K \in T^h} \|\mathbf{div}\, \boldsymbol{\sigma}(\mathbf{u}^h)\ +\ \mathbf{f}\|_{0,K} \|\mathbf{u} - \mathbf{v}^h\|_{0,K},
$$

and it suffices to estimate $\|\mathbf{u} - \mathbf{v}^h\|_{0,K}$ for any triangle K. From the definition of $\mathbf{v}^h$ and (2.6), we get

$$
\|\mathbf{u} - \mathbf{v}^h\|_{0,K} = \|\mathbf{e} - R^h \mathbf{e}\|_{0,K} \lesssim h_K \|\mathbf{e}\|_{1,\omega_K}.
$$

As a consequence,

$$
\left| \int_\Omega (\mathbf{div}\, \boldsymbol{\sigma}(\mathbf{u}^h)\ +\ \mathbf{f}) \cdot (\mathbf{u} - \mathbf{v}^h)\, \mathrm{d}\Gamma \right| \lesssim (\eta + \zeta) \|\mathbf{e}\|_{1,\Omega}.
$$

We now consider the interior and Neumann boundary terms in (3.4). As we previously noticed, the application of the Cauchy–Schwarz inequality leads to

$$
\left| \sum_{E \in E_h^{\mathrm{int}} \cup E_h^{\mathrm{N}}} \int_E J_{E,n}(\mathbf{u}^h) \cdot (\mathbf{u} - \mathbf{v}^h)\, \mathrm{d}\Gamma \right| \le \sum_{E \in E_h^{\mathrm{int}} \cup E_h^{\mathrm{N}}} \|J_{E,n}(\mathbf{u}^h)\|_{0,E} \|\mathbf{u} - \mathbf{v}^h\|_{0,E}.
$$

Therefore, using expression (3.5) and estimate (2.7), we obtain

$$
\|\mathbf{u} - \mathbf{v}^h\|_{0,E} = \|\mathbf{e} - R^h \mathbf{e}\|_{0,E} \lesssim h_E^{1/2} \|\mathbf{e}\|_{1,\omega_E}.
$$

Inserting this estimate into the previous one we deduce that

$$
\left| \sum_{E \in E_h^{\mathrm{int}} \cup E_h^{\mathrm{N}}} \int_E J_{E,n}(\mathbf{u}^h) \cdot (\mathbf{u} - \mathbf{v}^h)\, \mathrm{d}\Gamma \right| \lesssim \eta \|\mathbf{e}\|_{1,\Omega}.
$$

Moreover,

$$
\left| \sum_{E \in E_h^{\mathrm{N}}} \int_E (\mathbf{F} - \mathbf{F}_E) \cdot (\mathbf{u} - \mathbf{v}^h)\, \mathrm{d}\Gamma \right| \lesssim \zeta \|\mathbf{e}\|_{1,\Omega}.
$$

The two following terms are handled in a similar way as previously. Using the inequality $[a+b]_+ \leq [a]_+ + |b|$ for $a, b \in \mathbb{R}$, we bound

$$\begin{aligned}
&\left| \sum_{E \in E_h^{\mathrm{C}}} \int_E \left(\frac{1}{\gamma} [P_\gamma(\mathbf{u}^h) - g]_+ + \sigma_n(\mathbf{u}^h) \right) (v_n^h - u_n) \, \mathrm{d}\Gamma \right| \\
&\quad \leq \sum_{E \in E_h^{\mathrm{C}}} \int_E \left| \frac{1}{\gamma} [P_\gamma(\mathbf{u}^h) - g_{\mathrm{C}}]_+ + \sigma_n(\mathbf{u}^h) \right| |v_n^h - u_n| \, \mathrm{d}\Gamma + \sum_{E \in E_h^{\mathrm{C}}} \int_E \frac{1}{\gamma} |g - g_{\mathrm{C}}| \, |v_n^h - u_n| \, \mathrm{d}\Gamma \\
&\quad \lesssim (\eta + \zeta) \|\mathbf{e}\|_{1,\Omega}.
\end{aligned}$$

Moreover, there holds

$$\left| \sum_{E \in E_h^{\mathrm{C}}} \int_E \boldsymbol{\sigma}_{\mathbf{t}}(\mathbf{u}^h) \cdot (\mathbf{v}_{\mathbf{t}}^h - \mathbf{u}_{\mathbf{t}}) \, \mathrm{d}\Gamma \right| \lesssim \eta \|\mathbf{e}\|_{1,\Omega}.$$

Collecting the previous results, we deduce

$$\mathcal{T}_1 \lesssim (\eta + \zeta) \|\mathbf{e}\|_{1,\Omega}. \tag{3.6}$$

The first two terms in $\mathcal{T}_2$ are split using the definitions of $P_\gamma(\cdot)$ and $P_{\theta\gamma}(\cdot)$, and the last one is split using relationship $\sigma_n(\mathbf{v}^h - \mathbf{u}^h) = \sigma_n((\mathbf{v}^h - \mathbf{u}) + (\mathbf{u} - \mathbf{u}^h))$. This leads to

$$\begin{aligned}
\mathcal{T}_2 = &\int_{\Gamma_{\mathrm{C}}} \sigma_n(\mathbf{u}) P_\gamma(\mathbf{u} - \mathbf{u}^h) \, \mathrm{d}\Gamma + \int_{\Gamma_{\mathrm{C}}} \sigma_n(\mathbf{u}) \gamma \sigma_n(\mathbf{u} - \mathbf{u}^h) \, \mathrm{d}\Gamma \\
&+ \int_{\Gamma_{\mathrm{C}}} \frac{1}{\gamma} [P_\gamma(\mathbf{u}^h) - g]_+ P_\gamma(\mathbf{u} - \mathbf{u}^h) \, \mathrm{d}\Gamma + (1-\theta) \int_{\Gamma_{\mathrm{C}}} \frac{1}{\gamma} [P_\gamma(\mathbf{u}^h) - g]_+ \gamma \sigma_n(\mathbf{u} - \mathbf{u}^h) \, \mathrm{d}\Gamma \\
&- \theta \int_{\Gamma_{\mathrm{C}}} \frac{1}{\gamma} \left([P_\gamma(\mathbf{u}^h) - g]_+ + \sigma_n(\mathbf{u}^h) \right) \gamma \sigma_n(\mathbf{v}^h - \mathbf{u}) \, \mathrm{d}\Gamma - \theta \int_{\Gamma_{\mathrm{C}}} \gamma \, \sigma_n(\mathbf{u}^h) \sigma_n(\mathbf{u} - \mathbf{u}^h) \, \mathrm{d}\Gamma.
\end{aligned}$$

Then, we split the second term in the above expression using $1 = \theta + (1-\theta)$ and we gather the resulting terms:

$$\begin{aligned}
\mathcal{T}_2 = &\int_{\Gamma_{\mathrm{C}}} \left(\frac{1}{\gamma} [P_\gamma(\mathbf{u}^h) - g]_+ + \sigma_n(\mathbf{u}) \right) P_\gamma(\mathbf{u} - \mathbf{u}^h) \, \mathrm{d}\Gamma \\
&+ (1-\theta) \int_{\Gamma_{\mathrm{C}}} \frac{1}{\gamma} \left(\sigma_n(\mathbf{u}) + [P_\gamma(\mathbf{u}^h) - g]_+ \right) \gamma \sigma_n(\mathbf{u} - \mathbf{u}^h) \, \mathrm{d}\Gamma \\
&- \theta \int_{\Gamma_{\mathrm{C}}} \left(\frac{1}{\gamma} [P_\gamma(\mathbf{u}^h) - g]_+ + \sigma_n(\mathbf{u}^h) \right) \gamma \sigma_n(\mathbf{v}^h - \mathbf{u}) \, \mathrm{d}\Gamma + \theta \|\gamma^{1/2} \sigma_n(\mathbf{u} - \mathbf{u}^h)\|_{0,\Gamma_{\mathrm{C}}}^2.
\end{aligned}$$

Now we substitute $\sigma_n(\mathbf{u})$ using the reformulation of contact conditions (2.2) (i)–(iii) as $\sigma_n(\mathbf{u}) = -\frac{1}{\gamma} [P_\gamma(\mathbf{u}) - g]_+$ (see, e.g., Alart & Curnier, 1988; Chouly & Hild, 2013). This reformulation makes sense in

$L^2(\Gamma_C)$ due to the regularity assumption $\mathbf{u} \in (H^{3/2+\nu}(\Omega))^d$. Afterward, we apply the bound (2.4) in the first term as well as the Cauchy–Schwarz inequality in the second one:

$$
\begin{aligned}
\mathcal{T}_2 \leq & - \left\| \gamma^{1/2}(\sigma_n(\mathbf{u}) + \frac{1}{\gamma}[P_\gamma(\mathbf{u}^h) - g]_+) \right\|_{0,\Gamma_C}^2 \\
& + |\theta - 1| \left\| \gamma^{1/2}(\sigma_n(\mathbf{u}) + \frac{1}{\gamma}[P_\gamma(\mathbf{u}^h) - g]_+) \right\|_{0,\Gamma_C} \|\gamma^{1/2}\sigma_n(\mathbf{u} - \mathbf{u}^h)\|_{0,\Gamma_C} \\
& - \theta \int_{\Gamma_C} \left(\frac{1}{\gamma}[P_\gamma(\mathbf{u}^h) - g]_+ + \sigma_n(\mathbf{u}^h) \right) \gamma \sigma_n(\mathbf{v}^h - \mathbf{u})\, \mathrm{d}\Gamma + \theta \|\gamma^{1/2}\sigma_n(\mathbf{u} - \mathbf{u}^h)\|_{0,\Gamma_C}^2.
\end{aligned}
$$

The expression $ab \leq a^2 + b^2/4$ yields, for any $\beta > 0$,

$$
\begin{aligned}
\mathcal{T}_2 \leq & \frac{|\theta - 1|^2}{4} \|\gamma^{1/2}\sigma_n(\mathbf{u} - \mathbf{u}^h)\|_{0,\Gamma_C}^2 \\
& - \theta \int_{\Gamma_C} \left(\frac{1}{\gamma}[P_\gamma(\mathbf{u}^h) - g]_+ + \sigma_n(\mathbf{u}^h) \right) \gamma \sigma_n(\mathbf{v}^h - \mathbf{u})\, \mathrm{d}\Gamma + \theta \|\gamma^{1/2}\sigma_n(\mathbf{u} - \mathbf{u}^h)\|_{0,\Gamma_C}^2 \\
= & \frac{(\theta + 1)^2}{4} \|\gamma^{1/2}\sigma_n(\mathbf{u} - \mathbf{u}^h)\|_{0,\Gamma_C}^2 - \theta \int_{\Gamma_C} \left(\frac{1}{\gamma}[P_\gamma(\mathbf{u}^h) - g]_+ + \sigma_n(\mathbf{u}^h) \right) \gamma \sigma_n(\mathbf{v}^h - \mathbf{u})\, \mathrm{d}\Gamma \\
\leq & \frac{(\theta + 1)^2}{4} \|\gamma^{1/2}\sigma_n(\mathbf{u} - \mathbf{u}^h)\|_{0,\Gamma_C}^2 + |\theta| \gamma_0^{1/2}(\eta + \zeta) \|\gamma^{1/2}\sigma_n(\mathbf{v}^h - \mathbf{u})\|_{0,\Gamma_C} \\
\leq & \frac{(\theta + 1)^2}{4} \|\gamma^{1/2}\sigma_n(\mathbf{u} - \mathbf{u}^h)\|_{0,\Gamma_C}^2 + \beta\theta^2\gamma_0(\eta + \zeta)^2 \\
& + \frac{1}{2\beta} \left(\|\gamma^{1/2}\sigma_n(\mathbf{v}^h - \mathbf{u}^h)\|_{0,\Gamma_C}^2 + \|\gamma^{1/2}\sigma_n(\mathbf{u}^h - \mathbf{u})\|_{0,\Gamma_C}^2 \right) \\
= & \left(\frac{1}{2\beta} + \frac{(\theta + 1)^2}{4} \right) \gamma_0 \|\sigma_n(\mathbf{u} - \mathbf{u}^h)\|_{-1/2,h,\Gamma_C}^2 + \beta\theta^2\gamma_0(\eta + \zeta)^2 + \frac{\gamma_0}{2\beta} \|\sigma_n(\mathbf{v}^h - \mathbf{u}^h)\|_{-1/2,h,\Gamma_C}^2.
\end{aligned}
$$

Using (2.9) and the H^1-stability of R^h (see (2.8) in Lemma 2.1), we bound

$$
\|\sigma_n(\mathbf{v}^h - \mathbf{u}^h)\|_{-1/2,h,\Gamma_C} \leq C\|\mathbf{v}^h - \mathbf{u}^h\|_{1,\Omega} = C\|R^h(\mathbf{u} - \mathbf{u}^h)\|_{1,\Omega} \leq C\|\mathbf{u} - \mathbf{u}^h\|_{1,\Omega}.
$$

We combine this last bound with the saturation assumption (3.2) and get

$$
\mathcal{T}_2 \leq C\gamma_0 \left(\frac{(\theta + 1)^2}{4} + \frac{1}{\beta} \right) \|\mathbf{u} - \mathbf{u}^h\|_{1,\Omega}^2 + \beta\theta^2\gamma_0(\eta + \zeta)^2. \tag{3.7}
$$

Now we combine estimates (3.3), (3.6) and (3.7):

$$
\alpha\|\mathbf{e}\|_{1,\Omega}^2 \leq C(\eta + \zeta)\|\mathbf{e}\|_{1,\Omega} + C\gamma_0 \left(\frac{(\theta + 1)^2}{4} + \frac{1}{\beta} \right) \|\mathbf{e}\|_{1,\Omega}^2 + \beta\theta^2\gamma_0(\eta + \zeta)^2.
$$

We treat the first term on the right-hand side with Young's inequality and obtain

$$\left(\frac{\alpha}{2} - C\gamma_0\left(\frac{(\theta+1)^2}{4} + \frac{1}{\beta}\right)\right)\|\mathbf{e}\|^2_{1,\Omega} \le \frac{C}{\alpha}(\eta^2 + \zeta^2) + \beta\theta^2\gamma_0(\eta+\zeta)^2.$$

When $\theta \neq -1$, we choose γ_0 sufficiently small, and for $\theta = -1$, we can choose, e.g., $\beta = 4C\gamma_0/\alpha$ (for a fixed value of $\gamma_0 > 0$, which does not need to be small in this case). We obtain the upper bound on the error in natural norm,

$$\|\mathbf{e}\|_{1,\Omega} \lesssim (1+\gamma_0)(\eta+\zeta).$$

The saturation assumption (3.2) provides directly a bound on the contact stress error:

$$\|\sigma_n(\mathbf{u}-\mathbf{u}^h)\|_{-1/2,h,\Gamma_C} \lesssim \|\mathbf{e}\|_{1,\Omega}.$$

For the contact error, we make use of the triangle inequality and of the above inequality:

$$\begin{aligned}
&\left\|\sigma_n(\mathbf{u}) + \frac{1}{\gamma}[P_\gamma(\mathbf{u}^h) - g]_+\right\|_{-1/2,h,\Gamma_C} \\
&\quad \le \|\sigma_n(\mathbf{u}-\mathbf{u}^h)\|_{-1/2,h,\Gamma_C} + \left\|\sigma_n(\mathbf{u}^h) + \frac{1}{\gamma}[P_\gamma(\mathbf{u}^h) - g]_+\right\|_{-1/2,h,\Gamma_C} \\
&\quad \lesssim \|\mathbf{e}\|_{1,\Omega} + \eta + \zeta.
\end{aligned}$$

Collecting the three previous results allows us to prove the theorem. □

3.3 *Lower error bound*

We now consider the local lower error bounds of the discretization error terms.

THEOREM 3.6 For all elements $K \in T_h$, the following local lower error bounds hold:

$$\eta_{1K} \lesssim \|\mathbf{u}-\mathbf{u}^h\|_{1,K} + \zeta_K, \tag{3.8}$$

$$\eta_{2K} \lesssim \|\mathbf{u}-\mathbf{u}^h\|_{1,\omega_K} + \zeta_K. \tag{3.9}$$

For all elements K, such that $K \cap E_h^C \neq \emptyset$, the following local lower error bounds hold:

$$\eta_{3K} \lesssim \|\mathbf{u}-\mathbf{u}^h\|_{1,K} + \zeta_K, \tag{3.10}$$

$$\eta_{4K} \lesssim \sum_{E \in E_K^C} h_K^{1/2}\left(\left\|\sigma_n(\mathbf{u}) + \frac{1}{\gamma}[P_\gamma(\mathbf{u}^h) - g]_+\right\|_{0,E} + \left\|\sigma_n(\mathbf{u}-\mathbf{u}^h)\right\|_{0,E}\right) + \zeta_K. \tag{3.11}$$

Proof. The estimates of η_{1K}, η_{2K} in (3.8)–(3.9) are standard (see, e.g., Verfürth, 1999). The estimate η_{3K} in (3.10) is handled in a standard way, as in Hild & Nicaise (2007).

The estimate of η_{4K} in (3.11) is obtained from Definition 3.2 by using triangular inequalities. □

REMARK 3.7 Note that, from Theorem 3.6, optimal convergence rates of order $O(h^{\min(k,1/2+\nu)})$ are expected for the estimator of Definition 3.2.

REMARK 3.8 An extension of the above analysis for the Tresca friction case is sketched in the appendix.

REMARK 3.9 We carried out the analysis considering a straight contact boundary Γ_C for the sake of simplicity. However, some numerical tests (see Section 4.3) illustrate that this analysis might be extended to the case of curved boundaries. For contact problems with a curved boundary, one can refer to Hlaváček *et al.* (1988, Theorem 3.3, p. 149) and Wang (2000) for *a priori* error estimates, and to Schröder (2012) and Banz & Stephan (2015) for *a posteriori* error estimates.

4. Numerical experiments

We illustrate numerically the theoretical properties of the error estimator η given in Definition 3.2 and compute its convergence order when h vanishes. To study separately the global contributions of each component of η, we introduce the notation

$$\eta_i = \left(\sum_{K \in T_h} \eta_{iK}^2 \right)^{1/2}, \quad 1 \le i \le 4,$$

where the expressions of η_{iK} are provided in Definition 3.2. In all the examples below, Hooke's law is considered: E and ν_P will denote respectively Young's modulus and Poisson's ratio. Moreover, a dimensional analysis allows us to deduce that γ_0 is the inverse of a stiffness parameter. Consequently, we choose in our discussion $\gamma_0 = C/E$, where C is a constant that does not depend on E. The finite element method (2.5), as well as the residual estimator η, are implemented under the open source finite element library GetFEM++.[1] For details on numerical solution, we refer to Chouly *et al.* (2015) and Renard (2013).

To measure the quality of the estimator η, we introduce the effectivity index:

$$\mathrm{Eff}_E = \frac{\eta}{E \left\| \mathbf{u} - \mathbf{u}^h \right\|_{1,\Omega}}.$$

As in Hild & Nicaise (2007), this index has been normalized with respect to Young's modulus E. Indeed, we remark that if $\mathbf{u}(E)$ denotes the solution of a (linear) Lamé system with Young's modulus E then $\mathbf{u}(mE) = \mathbf{u}(E)/m$, whereas $\sigma(\mathbf{u}(mE)) = \sigma(\mathbf{u}(E))$. Thus, the error estimator η is independent of E (for η_4 this property comes from the scaling $\gamma_0 = C/E$). In contrast, there holds $\|\mathbf{u}(mE) - \mathbf{u}^h(mE)\|_{1,\Omega} = \|\mathbf{u}(E) - \mathbf{u}^h(E)\|_{1,\Omega}/m$, which becomes independent of E for the choice $m = 1/E$.

REMARK 4.1 Note that the parameter γ_0 scales as the inverse of an elastic coefficient. In the case of an isotropic or heterogeneous material, a possible option would be to set γ_0 relatively to the greater stiffness. This ensures at least well-posedness and correct behavior of both the approximation and the *a posteriori* error estimate. Improving this choice remains an open issue (see, e.g., Stein & Ohnimus, 1999 for elasticity without contact).

[1] see http://getfem.org

4.1 *First example: a square with slip and separation*

4.1.1 *Description.* We first consider a test case taken from Hild & Nicaise (2007) (see also Hild & Lleras, 2009; Lleras, 2009 in the frictional case). We consider the domain $\Omega = (0,1) \times (0,1)$ with material characteristics $E = 10^6$ and $\nu_P = 0.3$. A homogeneous Dirichlet condition on $\Gamma_{\mathrm{D}} = \{0\} \times (0,1)$ is prescribed to clamp the body. The body is potentially in contact on $\Gamma_{\mathrm{C}} = \{1\} \times (0,1)$ with a rigid obstacle, and $\Gamma_{\mathrm{N}} = (0,1) \times (\{0\} \cup \{1\})$ is the location of a homogeneous Neumann condition. There is no initial gap between the body and the rigid obstacle ($g = 0$). The body Ω is acted on by a vertical volume density of force $\mathbf{f} = (0, f_2)$ with $f_2 = -76518$, such that there is coexistence of a slip zone and a separation zone with a transition point between both zones. For error computations, since we do not have a closed-form solution, a reference solution is computed with Lagrange P_2 elements, $h = 1/160$, $\gamma_0 = 1/E$ and $\theta = -1$.

First of all, we illustrate in Fig. 1 the difference between uniform and adaptive refinement. For the latter, we refine only the mesh elements K in which the local estimator η_K is below a given threshold $s = 2.5 \times 10^{-3}$. The minimal (respectively maximal) size of the adaptive mesh is equal to $1/160$ (respectively $h = 1/40$). As expected, the rate of convergence with respect to the number of degrees of freedom is far better in the case of adaptive refinement than with uniform refinement.

The solution obtained with adaptive refinement and $\theta = -1$ is depicted in Fig. 2. We observe that the error is concentrated at both left corners (transition between Dirichlet and Neumann conditions) and near the transition point between contact and separation. As expected, we observe that all the nodes on Γ_{C} have a negative tangential displacement and that Γ_{C} is divided into two parts: the upper part where the body remains in contact (slipping nodes) and the lower part where it is separated, with a transition point near $(1, 0.685)$. The value is close to the transition point $(1, 0.69 \pm 0.01)$ found in Hild & Nicaise (2007) and $(1, 0.65)$ found in Lleras (2009). The slight difference with Lleras (2009) should be due to Coulomb friction.

REMARK 4.2 Note that the solution in the case $\theta = 1$ (see Fig. 3) has an error estimator on the contact zone, which is larger than in the case $\theta = -1$. In the case $\theta = -1$, the discrete solution is less dependent on the parameter γ_0 than for the other methods (see Chouly *et al.*, 2015) and we obtain a better approximation of the problem on the contact boundary Γ_{C}.

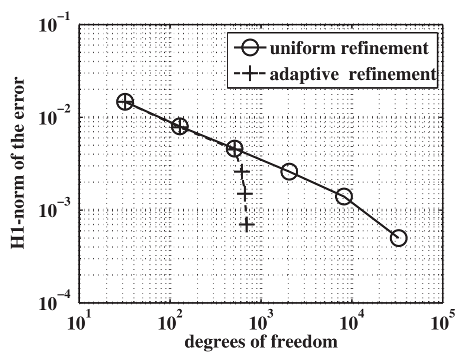


FIG. 1. Rate of convergence for uniform and adaptive refinement methods. Parameters $\gamma_0 = 1/E$, $\theta = -1$ and Lagrange P_2 elements.

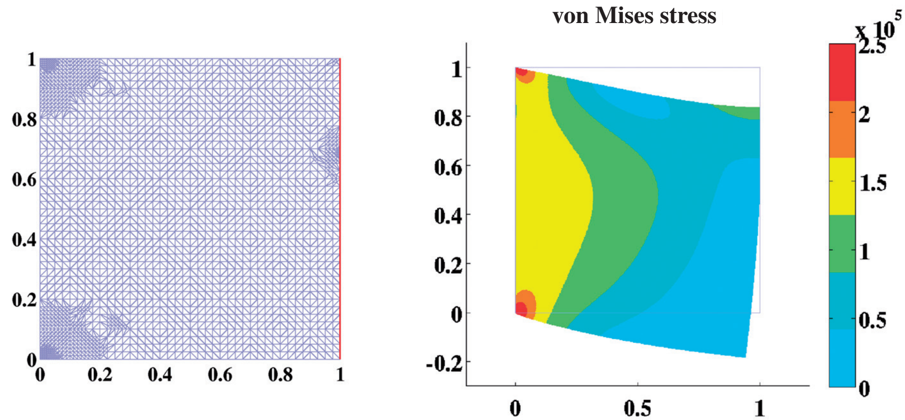


FIG. 2. Left panel: mesh with adaptive refinement and contact boundary on the right. Right panel: plot of von Mises stress. Parameters $\gamma_0 = 1/E$, $\theta = -1$ and Lagrange P_2 elements.

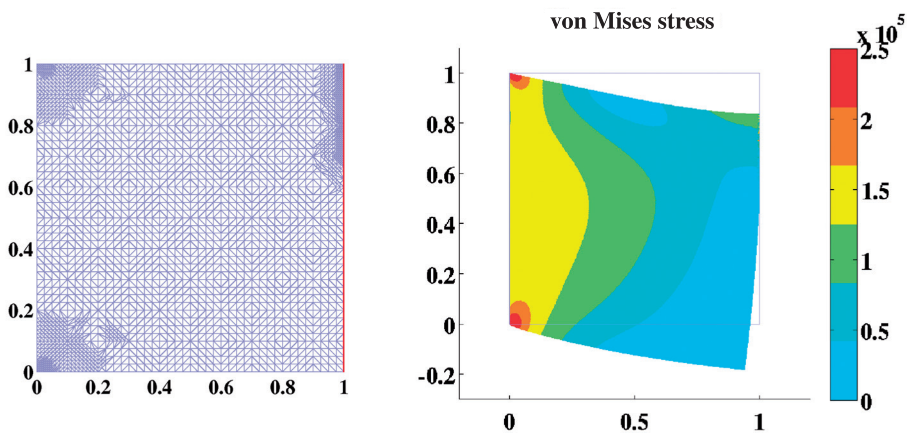
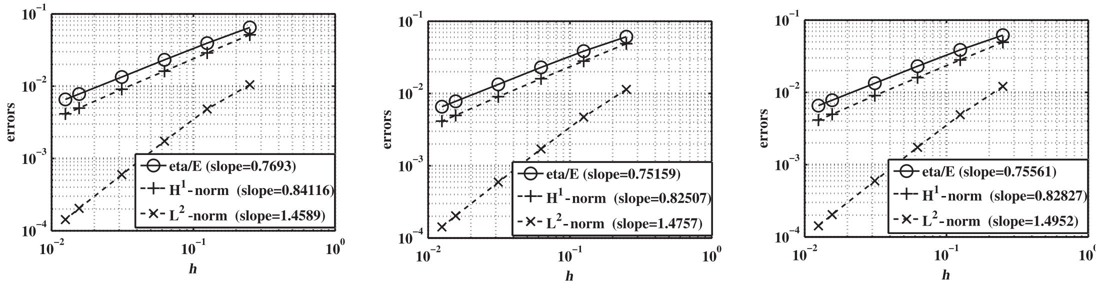


FIG. 3. Left panel: mesh with adaptive refinement and contact boundary on the right. Right panel: plot of von Mises stress. Parameters $\gamma_0 = 1/E$, $\theta = 1$ and Lagrange P_2 elements.

4.1.2 *Numerical convergence.* We perform a numerical convergence study for three variants of method (2.5) corresponding to $\theta = 1$, $\theta = 0$ and $\theta = -1$. The Nitsche parameter γ_0 is fixed to $1/E$, which should ensure well-posedness and optimal convergence in each case. Lagrange P_1 finite elements are chosen. The reference solution for error computations corresponds to the one described in Section 4.1.1 and depicted in Fig. 2 (P_2 finite elements, $\theta = -1$ and adaptive finest mesh). No mesh adaptation is carried out anymore and only uniform refinement is imposed, with a sequence of decreasing mesh sizes h.

First, the estimator η, the L^2- and the H^1-norms of the error $\mathbf{u} - \mathbf{u}^h$ are depicted in Fig. 4. One can note a suboptimality of the convergence rate in the L^2- and H^1-norms of the error. They are caused by the Neumann–Dirichlet transition in the left corners of Ω (the same observation has been reported in Fabre *et al.*, 2016).



$\theta = 1$ $\theta = 0$ $\theta = -1$

FIG. 4. First example. Convergence curves of the error estimator η, the L^2- and H^1-norms of the error $\mathbf{u} - \mathbf{u}^h$, for $\gamma_0 = 1/E$.

TABLE 1 *First example,* $\theta = 1$ *and* $\gamma_0 = 1/E$

Mesh size h	1/4	1/8	1/16	1/32	1/64	1/80	Slope
Degrees of freedom	32	128	512	2048	8192	12800	
$\left\Vert \mathbf{u} - \mathbf{u}^h \right\Vert_{0,\Omega}$ $(\times 10^{-4})$	104.7551	48.2436	17.3689	5.9666	2.0366	1.4262	1.4589
$\left\Vert \mathbf{u} - \mathbf{u}^h \right\Vert_{1,\Omega}$ $(\times 10^{-3})$	51.3896	28.8563	16.1335	9.0627	4.9777	4.1489	0.8412
η_1	16719.8	8359.9	4179.95	2089.97	1044.99	835.99	1.0000
η_2	60779.5	38076.7	22698	13222.3	7724.01	6507.89	0.7522
η_3	7626.32	3209.18	1207.19	427.694	157.242	118.467	1.4107
η_4	13501	4604.89	1395.58	370.912	100.73	77.2	1.7646
η	64916.4	39385.6	23153.3	13398.4	7796.61	6562.89	0.7779
Effectivity index Eff_E	1.2632	1.3649	1.4351	1.4784	1.5661	1.5816	

Then the different contributions of η are reported in Tables 1–3. The convergence rate of η_1 is strictly equal to 1 since, for piecewise linear finite elements, the expression of this estimator reduces to $\eta_{1K} = h_K \|\mathbf{f}_K\|_{0,K}$. More generally, all the estimators η_i converge toward zero as h vanishes, and they behave identically whatever the value of θ is (this is due to the low value of γ_0). Moreover, the convergence rate of η_2 is slightly less than that of the H^1-norms of the error, whereas the convergence rates of η_3 and η_4 are far greater and higher than 1 (we do not have a clear interpretation of this). In all cases, we obtain an effectivity index between 1.2 and 1.6 (the average is close to 1.45 and the standard deviation is close to 0.12). These overall results are quite similar to those presented in Hild & Nicaise (2007) and Lleras (2009).

Figure 5 shows the numerical experiment performed for a larger parameter $\gamma_0 = 1000/E$. In the case $\theta = 1$ and in the case $\theta = 0$, the convergence rate is degraded compared with the case $\gamma_0 = 1/E$. Conversely, in the case $\theta = -1$, the convergence is not deteriorated, which confirms the theoretical results obtained in both the *a priori* analysis in Chouly *et al.* (2015) and the *a posteriori* analysis in Section 3 (see Theorem 3.5).

4.1.3 *The case of a very large* γ_0. Additionally, we present a numerical convergence study for $\theta = 1, 0, -1$ and for a very large value of the parameter $\gamma_0 = 10^6/E$, far from its reference value of $1/E$. In this case, for $\theta = 1$ and $\theta = 0$, there is no longer a guarantee of well-posedness and optimal convergence

TABLE 2 *First example, $\theta = 0$ and $\gamma_0 = 1/E$*

Mesh size h	1/4	1/8	1/16	1/32	1/64	1/80	Slope
Degrees of freedom	32	128	512	2048	8192	12800	
$\|\mathbf{u}-\mathbf{u}^h\|_{0,\Omega}$ $(\times 10^{-4})$	113.6807	47.1350	17.0780	5.9262	2.0312	1.4229	1.4757
$\|\mathbf{u}-\mathbf{u}^h\|_{1,\Omega}$ $(\times 10^{-3})$	48.8181	28.0213	15.9877	9.0359	4.9716	4.1459	0.8251
η_1	16719.8	8359.9	4179.95	2089.97	1044.99	835.99	1.0000
η_2	57305.3	37374.7	22547.2	13200.7	7720.86	6505.24	0.7356
η_3	3938.22	1852.35	720.951	256.135	95.0474	71.047	1.3686
η_4	11946.5	4002.56	1154.11	324.915	89.6552	61.026	1.7809
η	61005.6	38551.4	22971.7	13371.5	7792.35	6559.4	0.7779
Effectivity index Eff_E	1.2496	1.3758	1.4368	1.4798	1.5672	1.5819	

TABLE 3 *First example, $\theta = -1$ and $\gamma_0 = 1/E$*

Mesh size h	1/4	1/8	1/16	1/32	1/64	1/80	Slope
Degrees of freedom	32	128	512	2048	8192	12800	
$\|\mathbf{u}-\mathbf{u}^h\|_{0,\Omega}$ $(\times 10^{-4})$	120.9371	48.9718	17.3613	5.9619	2.0360	1.4255	1.4952
$\|\mathbf{u}-\mathbf{u}^h\|_{1,\Omega}$ $(\times 10^{-3})$	49.3705	28.1269	16.0087	9.0385	4.9714	4.1467	0.8283
η_1	16719.8	8359.9	4179.95	2089.97	1044.99	835.99	1.0000
η_2	58846.3	37649.9	22607.7	13213.2	7723.58	6506.99	0.7428
η_3	2690.5	1464.81	558.637	192.194	70.7559	53.7733	1.3544
η_4	9202.06	2854.93	832.228	229.683	62.842	44.0949	1.8004
η	61922.2	38700.1	23012.7	13380.8	7794.52	6560.84	0.7779
Effectivity index Eff_E	1.2542	1.3759	1.4375	1.4804	1.5677	1.5820	

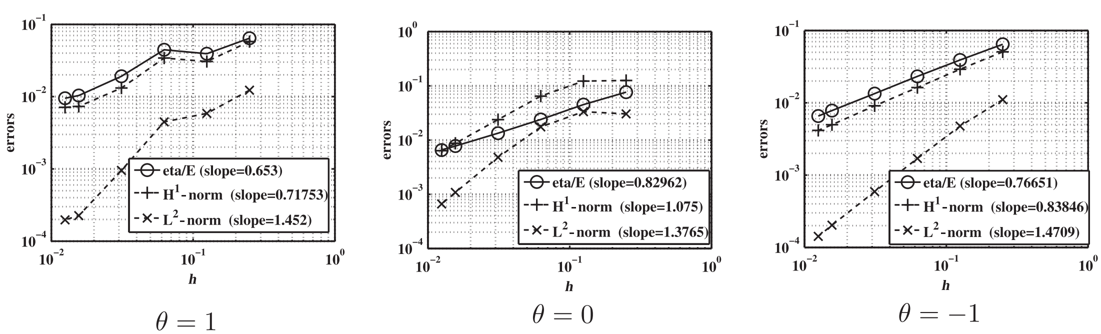
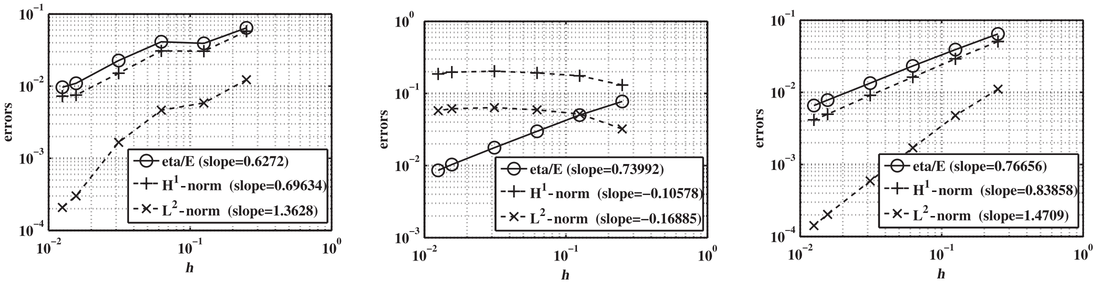


FIG. 5. First example. Convergence curves of the error estimator η, the L^2- and H^1-norms of the error $\mathbf{u}-\mathbf{u}^h$, for $\gamma_0 = 1000/E$.



$\theta = 1$ $\theta = 0$ $\theta = -1$

FIG. 6. First example. Convergence curves of the error estimator η, the L^2- and H^1-norms of the error $\mathbf{u} - \mathbf{u}^h$, for $\gamma_0 = 10^6/E$.

TABLE 4 *First example, $\theta = 1$ and $\gamma_0 = 10^6/E$*

Mesh size h	1/4	1/8	1/16	1/32	1/64	1/80	Slope
Degrees of freedom	32	128	512	2048	8192	12800	
$\left\Vert \mathbf{u} - \mathbf{u}^h \right\Vert_{0,\Omega}$ $(\times 10^{-4})$	122.6500	58.4959	46.4511	16.6143	3.0112	2.0808	1.3628
$\left\Vert \mathbf{u} - \mathbf{u}^h \right\Vert_{1,\Omega}$ $(\times 10^{-3})$	57.7770	30.5558	30.8275	15.1381	7.5190	7.2669	0.6963
η_1	16719.8	8359.9	4179.95	2089.97	1044.99	835.99	1.0000
η_2	62073.4	38335.4	41033.3	22552.4	10916.7	9635.48	0.6172
$\eta_3 (\times 10^{-2})$	2.13709	1.01961	0.63268	0.768462	0.414401	0.42499	0.4878
$\eta_4 (\times 10^{-2})$	2.52415	0.571842	2.55605	1.38494	0.525521	0.446467	0.4177
η	64285.7	39236.3	41245.6	22649	10966.6	9671.68	0.6272
Effectivity index Eff_E	1.1127	1.2841	1.3380	1.4962	1.4564	1.3266	

(see Chouly *et al.*, 2015). The error estimator η, the L^2- and H^1-norms of the error $\mathbf{u} - \mathbf{u}^h$ are plotted in Fig. 6, while Tables 4–6 present the different contributions of η.

For the method $\theta = 0$, the solution does not converge, while the effectivity index (Eff_E) tends to 0. This is consistent with our theoretical results since Theorem 3.5 is no longer applicable and no upper bound is guaranteed. The estimator η converges, though the term η_4 is slightly increasing (but remark that η_4 is very small in comparison to η). For the method $\theta = 1$, even though γ_0 is large, the method converges in L^2- and H^1-norms of the error with an acceptable effectivity index, but with a deteriorated convergence rate. Conversely, for the method $\theta = -1$, both convergence and the effectivity index are optimal and are not deteriorated compared with the case $\gamma_0 = 1/E$. This supports its theoretical property of robustness with respect to γ_0.

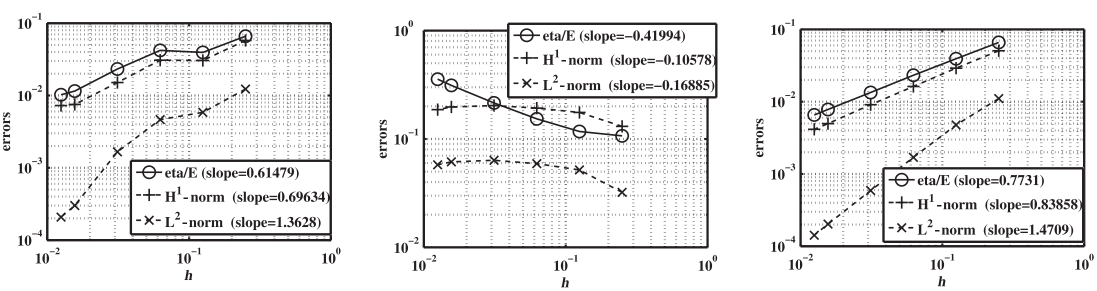
The previous experiment for $\theta = 0$ reveals the bad behavior of η for very large γ_0. An heuristic to recover a meaningful estimator is to decouple the value of γ_0 for problem (2.5) and for the estimator η_4. A final experiment performed, as shown in Fig. 7, shows the convergence curves in the same case $\gamma_0 = 10^6/E$, yet with an error estimator that makes use of a Nitsche parameter $\tilde{\gamma}_0 = 1/E$. For the methods $\theta = 1$ and $\theta = -1$, this has no visible influence on the effectivity index Eff_E. For the method $\theta = 0$, a better effectivity index is obtained: at least the estimator does not tend to zero for a nonconvergent solution, in contrast with what happens in Fig. 6. To summarize, this study for large γ_0 confirms the

TABLE 5 *First example, $\theta = 0$ and $\gamma_0 = 10^6/E$*

Mesh size h	1/4	1/8	1/16	1/32	1/64	1/80	Slope
Degrees of freedom	32	128	512	2048	8192	12800	
$\left\Vert \mathbf{u}-\mathbf{u}^h \right\Vert_{0,\Omega}$ $(\times 10^{-4})$	321.9223	518.2042	592.2560	635.8190	615.1016	577.5225	−0.1689
$\left\Vert \mathbf{u}-\mathbf{u}^h \right\Vert_{1,\Omega}$ $(\times 10^{-3})$	130.6775	175.6706	192.4574	203.4624	197.2235	186.4225	−0.1058
η_1	16719.8	8359.9	4179.95	2089.97	1044.99	835.99	1.0000
η_2	75562.4	49342.9	29582.9	17606.7	10284.5	8546.95	0.7339
η_3	2468.7	908.957	316.586	107.411	38.2804	28.2145	1.5045
η_4 $(\times 10^{-1})$	2.00548	3.65123	5.58251	8.52101	1.38006	1.37438	−0.6465
η	77429.5	50054.3	29878.5	17730.7	10337.5	8587.78	0.7399
Effectivity index Eff_E $(\times 10^{-1})$	5.9252	2.8493	1.5525	0.8714	0.5242	0.4607	

TABLE 6 *First example, $\theta = -1$ and $\gamma_0 = 10^6/E$*

Mesh size h	1/4	1/8	1/16	1/32	1/64	1/80	Slope
Degrees of freedom	32	128	512	2048	8192	12800	
$\left\Vert \mathbf{u}-\mathbf{u}^h \right\Vert_{0,\Omega}$ $(\times 10^{-4})$	110.5852	47.6266	16.9809	5.9093	2.0290	1.4216	1.4709
$\left\Vert \mathbf{u}-\mathbf{u}^h \right\Vert_{1,\Omega}$ $(\times 10^{-3})$	50.6403	29.1195	16.2386	9.0861	4.9803	4.1565	0.8386
η_1	16719.8	8359.9	4179.95	2089.97	1044.99	835.99	1.0000
η_2	62292.2	38204	22809.5	13249.4	7732.01	6512.98	0.7582
$\eta_3(\times 10^{-4})$	143.671	83.7405	26.7592	8.72031	3.05775	0.0236947	1.4405
η_4 $(\times 10^{-4})$	168.808	67.5774	14.2866	3.96942	1.17445	0.948677	1.8030
η	64497	39108	23189.4	13413.2	7802.31	6566.42	0.7666
Effectivity index Eff_E	1.2736	1.3430	1.4280	1.4762	1.5663	1.5796	



$\theta = 1$ $\theta = 0$ $\theta = -1$

FIG. 7. First example. Convergence curves of the error estimator η, the L^2- and H^1-norms of the error $\mathbf{u} - \mathbf{u}^h$, with $\gamma_0 = 10^6/E$ in Nitsche's method and $\tilde{\gamma}_0 = 1/E$ in the error estimator η_4.

analysis provided in Section 3, which requires a sufficiently small parameter γ_0 to obtain a reliable and an efficient *a posteriori* estimator when $\theta \neq -1$.

All the variants of Nitsche's method ($\theta = -1, 0, 1$) compare well whenever γ_0 is small enough. The symmetric version $\theta = 1$ has the advantage that it leads to a symmetric tangent matrix when the problem is solved with a generalized Newton algorithm. Also the nonsymmetric version $\theta = 0$ involves fewer terms in the weak formulation and may be preferred for this reason. Nevertheless, and as already observed in, e.g., Chouly *et al.* (2015), the skew-symmetric variant $\theta = -1$ appears to be robust in the sense that it preserves optimal convergence for a wide range of values for γ_0. Because of this property, we keep the choice $\theta = -1$ in the remaining part of the paper.

4.2 *Second example: a square/cube with softer singularities*

4.2.1 *Description.* We study another example, with softer singularities, inspired by Hild & Lleras (2009). We consider the domain $\Omega = (0, 1) \times (0, 1)$ with material characteristics $E = 10^4$, $\nu_P = 0.2$ and no body force ($\mathbf{f} = \mathbf{0}$). We adopt symmetry conditions

$$u_n = 0, \qquad \boldsymbol{\sigma}_{\mathbf{t}}(\mathbf{u}) = \mathbf{0}$$

on the boundary $\Gamma_S := \{0\} \times (0, 1)$. The contact with the rigid obstacle is on $\Gamma_C = (0, 1) \times \{0\}$. There is no initial gap between the body and the rigid obstacle ($g = 0$). On the remaining part of the boundary Γ_N, we impose a Neumann boundary condition, with the following expression for the surface force:

$$\mathbf{F}(x, y) = \begin{cases} (-y + 0.5, 0) & \text{if } x = 1 \text{ and } 0.5 \leq y \leq 1, \\ (0, -0.5 + x) & \text{if } y = 1 \text{ and } 0 \leq x \leq 0.5, \\ (0, 0) & \text{otherwise.} \end{cases}$$

This means that the force $\mathbf{F}$ is applied inward the body at the top and on the right side. Since there is no Dirichlet boundary condition ($\Gamma_D = \emptyset$) this corresponds to the $\mathbf{K}$-elliptic case (Haslinger *et al.*, 1996, Theorem 6.3). For error computations, since we do not have a closed-form solution, a reference solution is computed with Lagrange P_2 elements, $h = 1/100$, $\gamma_0 = 1/(100E)$ and $\theta = -1$. The reference solution is depicted in Fig. 8, with a displacement that is amplified by factor 2000. We recover a solution close to Hild & Lleras (2009), with a separation on the contact boundary at the bottom.

4.2.2 *Numerical convergence in two-dimensional.* We first investigate the numerical convergence of the error estimator. Error curves are depicted in Fig. 9. As in the previous example, convergence in the L^2- and H^1-norms is observed, as well as convergence of the error estimator η itself, with however a lower rate. The detailed behavior of η is provided in Table 7. Note that the effectivity index is close to 0.4.

4.2.3 *Adaptive refinement in two-dimensional.* We carry out adaptive refinement. At each refinement iteration, all the elements K for which $\eta_K \geq 0.5\eta_{\text{MAX}}$ are refined, where η_{MAX} denotes the maximum value of η_K over all the elements K in the mesh T^h. Solutions with adaptive refinement, as well as the error map (value of η_K at each element K), are depicted in Fig. 10. Note that both the highest values of the error estimator and the refinement are concentrated (i) at the symmetry–Neumann and Neumann–Neumann transitions near the top and right edges and (ii) near the separation on the contact boundary to resolve

FIG. 8. Reference solution for the second example. Plot of von Mises stress. Displacement amplified by 2000. Parameters $\gamma_0 = 1/(100E)$, $\theta = -1$ and Lagrange P_2 elements.

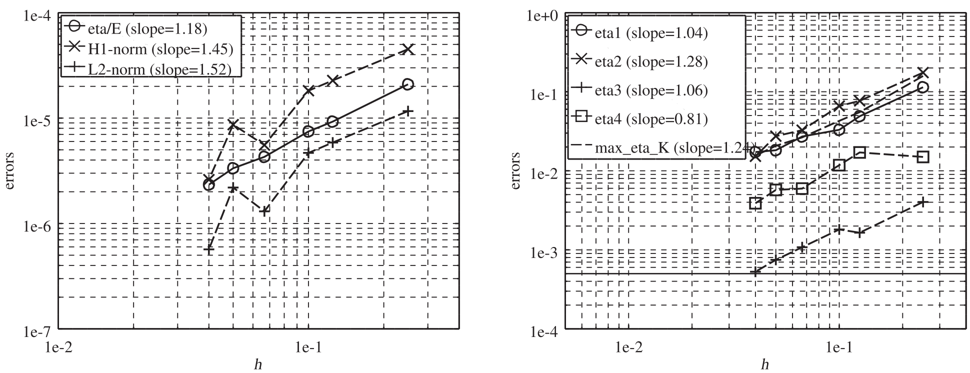


FIG. 9. Second example (two dimensions). Error curves for uniform refinement. Left: global estimator η and L^2- and the H^1-norms of the error. Right: separated components of the estimator $\eta_1, \ldots, \eta_4$, and maximum value of η_K.

the transition between contact and noncontact. After the fifth refinement iteration, no more significant evolution of the mesh is observed.

Error curves are depicted in Fig. 11, which allow us to assess that the error in the L^2- and H^1-norms decreases after each refinement, as well as η and all of its components $\eta_1, \ldots, \eta_4$. Finally, Fig. 12 shows how the error is reduced compared to the degrees of freedom, both for uniform and for adaptive refinement strategies.

TABLE 7 *Second example (two dimensions). Uniform refinement. Lagrange P_2 finite elements*

Mesh size h	1/4	1/8	1/10	1/15	1/20	1/25	Slope
Degrees of freedom	64	256	400	900	1600	2500	
$\lVert \mathbf{u} - \mathbf{u}^h \rVert_{0,\Omega}$ $(\times 10^{-7})$	116.47	58.61	46.75	13.04	21.93	5.68	1.5235
$\lVert \mathbf{u} - \mathbf{u}^h \rVert_{1,\Omega}$ $(\times 10^{-6})$	45.06	22.74	18.19	5.55	8.67	2.61	1.4500
η_1	0.113959	0.0489525	0.0329965	0.027139	0.0181964	0.0172498	1.0410
η_2	0.174624	0.0766977	0.0658897	0.0326302	0.0274335	0.0152579	1.2823
η_3	0.00403177	0.00164762	0.00180977	0.00107956	0.000746811	0.000528401	1.0609
η_4	0.0149697	0.017064	0.0117852	0.00595715	0.00572391	0.00388521	0.8093
η	0.209095	0.0925893	0.0746484	0.0428708	0.033422	0.0233609	1.1808
Effectivity index Eff_E	0.46401	0.40723	0.41044	0.77291	0.38570	0.89638	

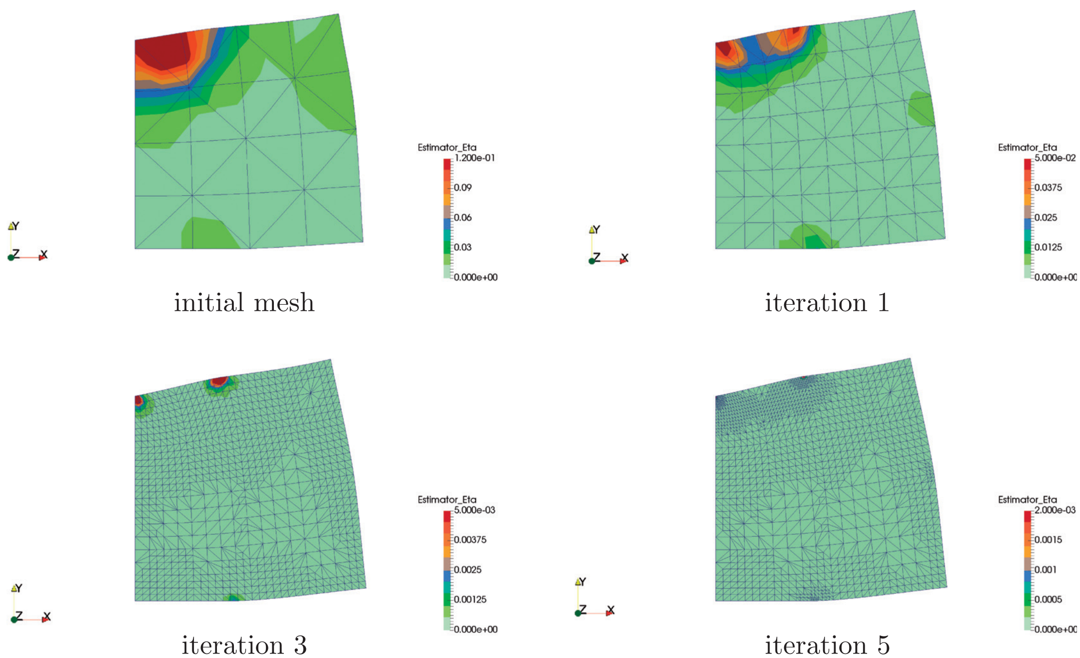


FIG. 10. Second example (two dimensions). Error map and refined mesh: initial guess for a coarse mesh and refinement iterations 1, 3 and 5.

4.2.4 *Adaptive refinement in three-dimensional.* We carry out a test to assess the performance of the error estimator η in the three-dimensional case, and its capability to resolve contact conditions even in three dimensions. We consider this time a cube $\Omega = (0,1) \times (0,1) \times (0,1)$ with the same material characteristics as in two dimensions. Symmetry conditions are imposed on the boundary $\Gamma_S := \{0\} \times (0,1) \times (0,1) \cup (0,1) \times \{0\} \times (0,1)$. The contact with the rigid obstacle is on $\Gamma_C = (0,1) \times (0,1) \times \{0\}$. There is still no initial gap between the body and the rigid obstacle ($g = 0$). On Γ_N the expression for the

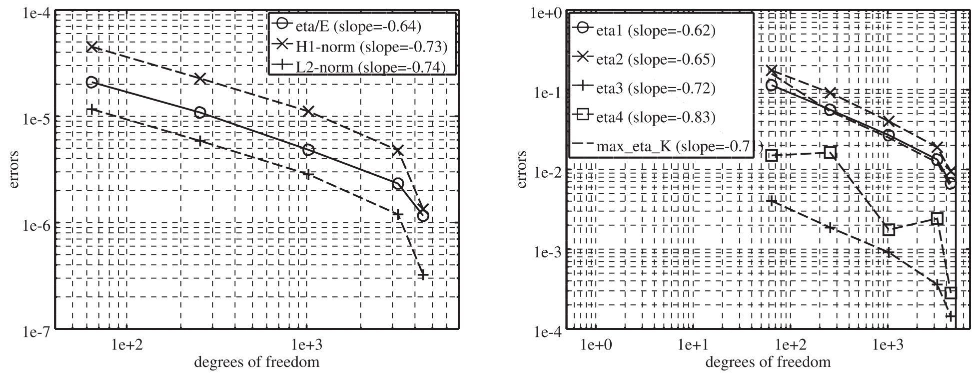


FIG. 11. Second example (two dimensions). Error curves for adaptive refinement. Left: global estimator η and L^2- and H^1-norms of the error. Right: separated components of the estimator $\eta_1, \ldots, \eta_4$, and maximum value of η_K.

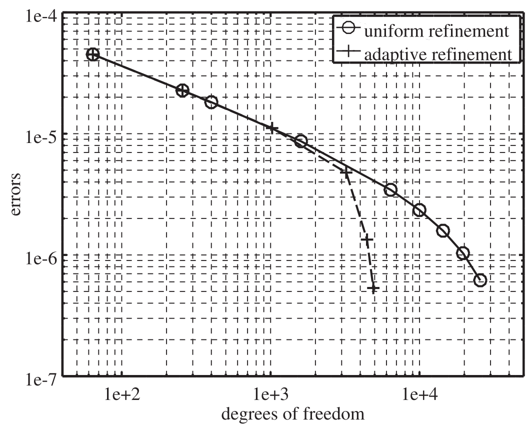


FIG. 12. Second example (two dimensions). Rate of convergence for uniform and adaptive refinement methods.

surface force is now

$$\mathbf{F}(x,y,z) = \begin{cases} (-0.5(z-0.5), 0, 0) & \text{if } x = 1 \text{ and } 0.5 \leq z \leq 1, \\ (0, -0.5(z-0.5), 0) & \text{if } y = 1 \text{ and } 0.5 \leq z \leq 1, \\ (0, 0, -4(0.5-x)(0.5-y)) & \text{if } z = 1 \text{ and } 0 \leq x, y \leq 0.5, \\ (0,0,0) & \text{otherwise.} \end{cases}$$

Refinement is still carried out with a relative threshold $\eta_K \geq 0.4\,\eta_{\mathrm{MAX}}$. Figure 13 depicts the convergence behavior of the error estimator, which decreases at each refinement iteration. Note, moreover, the better performance of the adaptative refinement, compared to uniform refinement, with a lower value and a slightly better slope. Figure 14 depicts the solution on the initial mesh and on the final mesh after 6 refinement iterations. Note that, as in the two-dimensional case, refinement occurs near the Neumann–Neumann (and symmetry–Neumann) transitions, as well as near the transition between contact and noncontact.

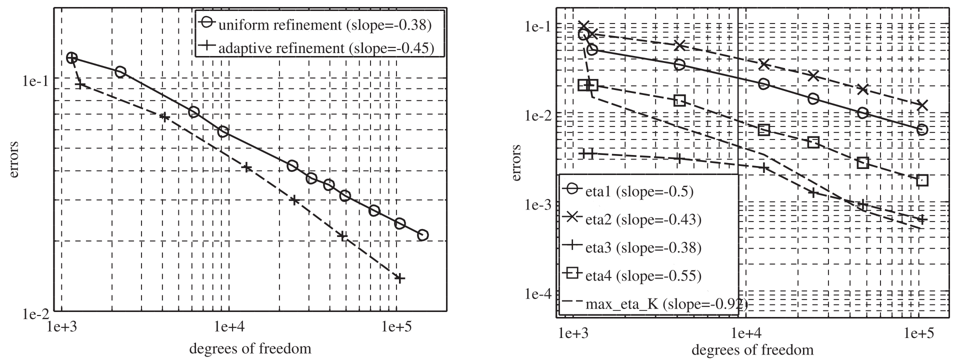


FIG. 13. Second example (three dimensions). Error curves for refinement in three dimensions. Left: global estimator η (uniform vs. adaptive). Right: separated components of the estimator $\eta_1, \ldots, \eta_4$, and maximum value of η_K (adaptive refinement only).

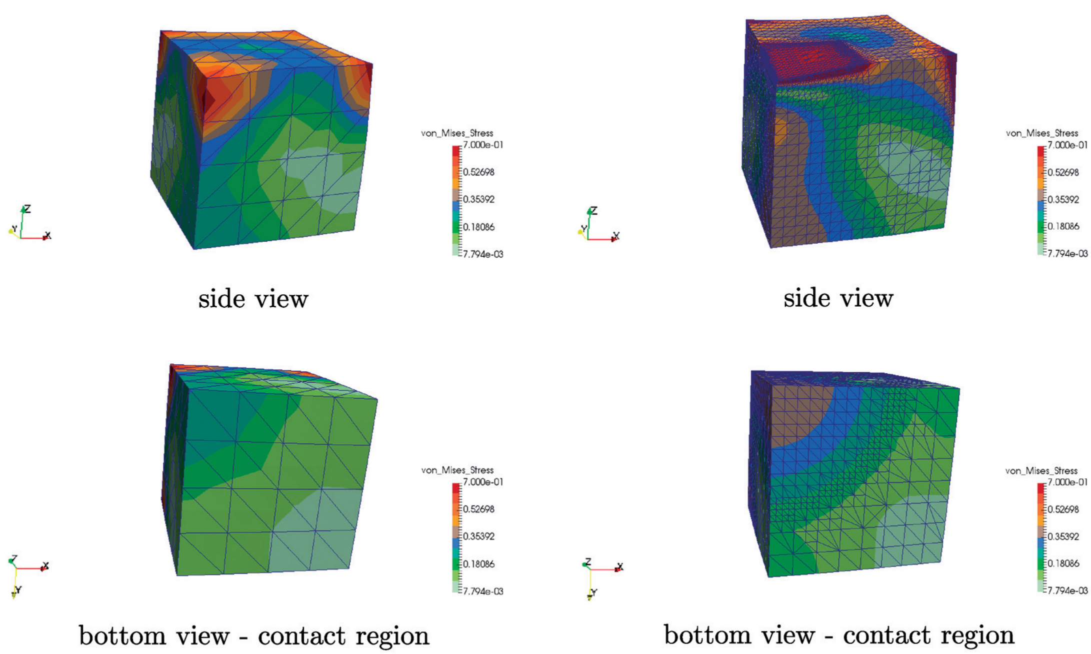


FIG. 14. Second example (three dimensions). Initial mesh (left panel) and final mesh (right panel) after 6 refinement iterations. The deformation is amplified by factor 2000.

4.3 *Third example: Hertz's contact*

4.3.1 *Description.* We consider Hertz's contact problems of a disk/a sphere with a plane rigid foundation (see, e.g., the numerical examples in Hild & Nicaise, 2007; Chouly *et al.*, 2015). The parameters have been fixed as $\theta = -1$ and $\gamma_0 = 10^{-3}/E$.

The disk (respectively the sphere) is of center $(0, 20)$ (respectively of center $(0, 0, 20)$) and radius 20. The lower part of the boundary Γ_C is potentially in contact with the rigid support. The remaining

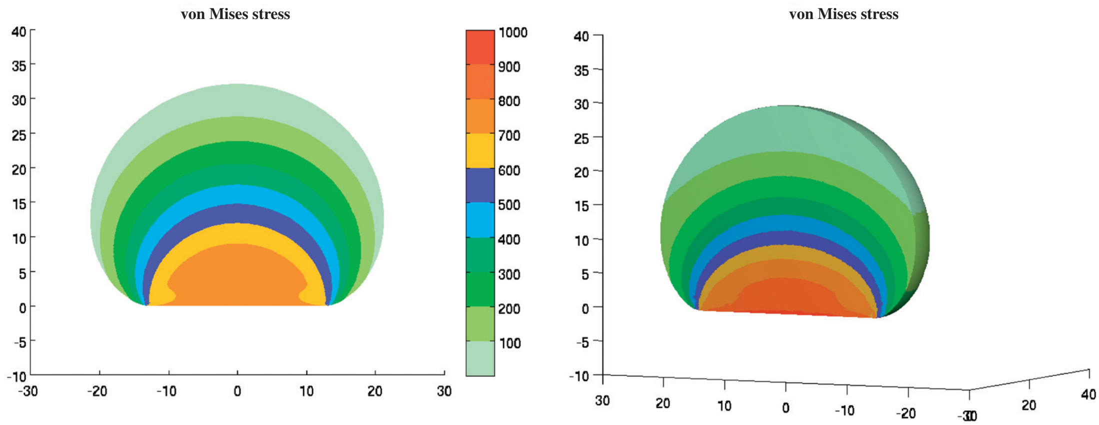


FIG. 15. Reference solutions with von Mises stresses, in two dimensions (left) and three dimensions (right).

(upper part) of the boundary Γ_N is subjected to a homogenous Neumann condition. To overcome the nondefiniteness coming from free rigid motions, the horizontal displacement is prescribed to be zero at the two points with coordinates $(0, 20)$ and $(0, 25)$ (respectively the horizontal displacement components u_1 and u_2 at the point (0, 0, 20), the component u_1 at the point (0, 5, 20) and the component u_2 at the point (5, 0, 20)): this blocks horizontal translation and rigid rotation. Young's modulus is fixed at $E = 25$ and Poisson's ratio is $\nu_P = 0.25$. A vertical density of volume forces of intensity 20 is applied in Ω. The reference solutions are depicted in Fig. 15. There are uniformly refined solutions with an average mesh size $h = 0.10$ for the disk (respectively $h = 1.27$ for the sphere), Lagrange P_2 elements, $\theta = -1$ and $\gamma = 10^{-3}/E$.

The initial gap between Γ_C and the obstacle is computed as $g(\mathbf{x}) := \mathbf{x} \cdot \mathbf{n}_{obs}$, where $\mathbf{x} \in \Gamma_C$ and with $\mathbf{n}_{obs}$ the unit outward normal vector on the boundary of the plane obstacle. In such a simple situation, we can take $g = g_C$, so that there is no approximation error associated with the gap.

4.3.2 *Numerical convergence in two-dimensional.* The error curves in the two-dimensional case are depicted in Fig. 16, for both linear and quadratic finite elements. In the case of P_1 finite elements, and as in Chouly *et al.* (2015), a slight super convergence is observed in the H^1-norm of the error (1.5 instead of 1). This behavior is not recovered by the error estimator η, which converges with a rate close to 1. The origin of this difference is unknown. For P_2 finite elements, the agreement between η and the error in the H^1-norm is better: for the H^1-norm, the convergence rate is close to 1.7, while approximately 1.5 for η. We observe the same results for the variants $\theta = 0, 1$.

In Table 8, the contribution of each component η_i of η is detailed. Each term of the error estimator converges toward zero when h becomes smaller. Note, however, the increasing values of the effectivity index, due to the super convergence in the H^1-norm and the convergence rate of the contribution η_4, which is close to 1.5.

4.3.3 *Numerical convergence in three-dimensional.* The error curves in the three-dimensional case are depicted in Fig. 17, for both linear and quadratic finite elements. For P_1 finite elements, the convergence rates for η and for the error in the H^1-norm are close (around 1.3) and slightly above the expected rate of

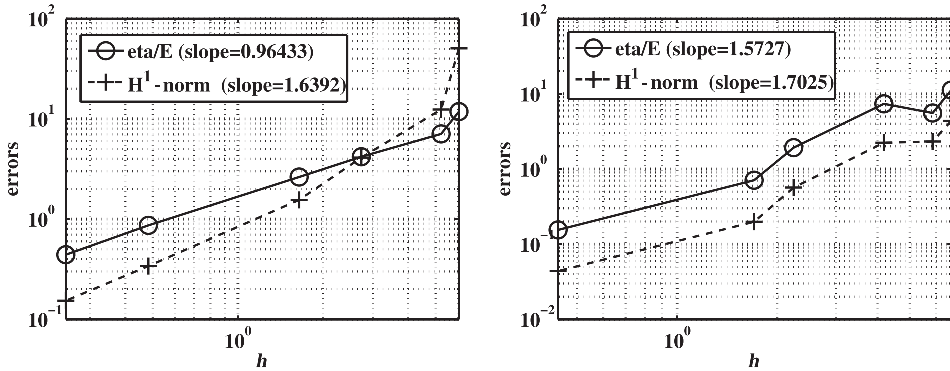


FIG. 16. Hertz's contact in two dimensions. Error estimator η and the H^1-norm of the error $\mathbf{u} - \mathbf{u}^h$, for Lagrange P_1 (left) and P_2 (right) finite elements.

TABLE 8 *Hertz's contact in two dimensions,* $\theta = -1$, $\gamma_0 = 10^{-3}/E$ *and Lagrange* P_1 *elements*

Mesh size h	6.04766	5.23002	2.7327	1.64637	0.482414	0.246359	Slope
$\lVert \mathbf{u} - \mathbf{u}^h \rVert_{1,\Omega}$	50.5984	12.3996	4.1083	1.5561	0.3399	0.1534	1.6392
η_1	7781.84	7376.92	4066.11	2379.44	728.236	359.817	0.9715
η_2	18000.7	12350.7	9279.79	5866.83	2009.88	1029.86	0.8525
η_3	2523.15	1055.64	852.542	458.121	90.2956	38.1934	1.2132
η_4	21999.5	10276.6	2537.53	1735.77	321.871	152.501	1.4597
η	29579.2	17711.1	10479.2	6580.59	2163.73	1102.18	0.9643
Effectivity index Eff_E	0.2338	0.5713	1.0203	1.6916	2.5467	2.8735	

1. For P_2 finite elements, we observe a suboptimality of the error estimator η, which converges but with a rate of 1, while the error in the H^1-norm remains optimal, with a convergence rate around 1.5.

The contribution of each component η_i of η is detailed in Tables 9 and 10 for linear and quadratic finite elements, respectively. For P_1 finite elements, the effectivity index remains close to 0.7 and the error estimator η_4 of the contact condition converges faster than the others. For P_2 finite elements, such behavior is not recovered, and η_4 converges with a rate of 1 approximately. The effectivity index is lower than for P_1 finite elements and remains around 0.3.

REMARK 4.3 Note that this test case is not fully covered by the theoretical analysis of Section 3 since the contact boundary is curved. We use isoparametric Lagrange P_1 and P_2 elements that provide affine and quadratic approximations, respectively, of the curved boundary. The numerical results presented in this section show that the theoretical bounds of Theorems 3.5 and 3.6 may be extended to a setting with a curved contact boundary.

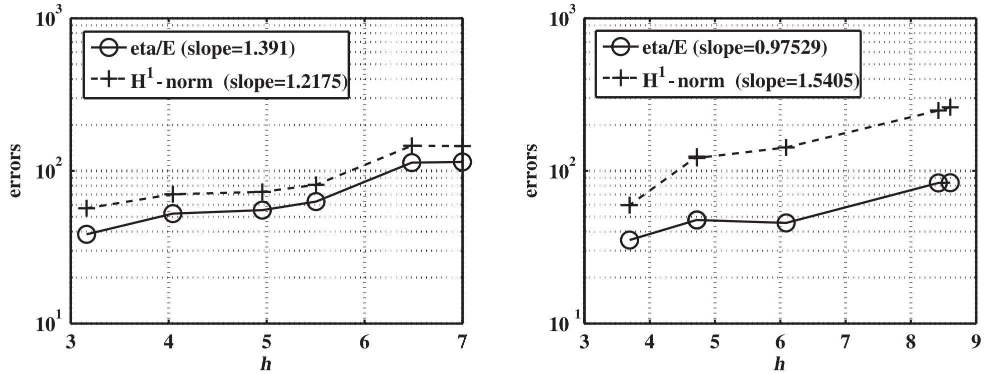


FIG. 17. Hertz's contact in three dimensions. Error estimator η and the H^1-norm of the error $\mathbf{u}-\mathbf{u}^h$, for Lagrange P_1 (left) and P_2 (right) finite elements.

TABLE 9 *Hertz's contact in three dimensions, $\theta=-1$, $\gamma_0=1/1000E$, Lagrange P_1 elements*

Mesh size h	6.99992	6.48188	5.50504	4.95584	4.04204	3.16207	Slope
$\left\|\mathbf{u}-\mathbf{u}^h\right\|_{1,\Omega}$	145.5261	146.1777	80.9980	72.8654	70.3068	56.9181	1.2175
η_1	63195.6	63786.9	58218	50778.4	47260.1	35050.5	0.7432
η_2	130398	130152	107003	96631.9	101191	80077.5	0.5911
η_3	9404.94	9276.75	15015.9	7973.6	6642.56	5246.94	0.8664
η_4	246585	242981	98555	84586.1	68588.5	39314.2	2.3401
η	286164	283080	157409	138328	131231	95990	1.3910
Effectivity index Eff_E	0.7866	0.7746	0.7773	0.7594	0.7466	0.6746	

TABLE 10 *Hertz's contact in three dimensions, $\theta=-1$, $\gamma_0=1/1000E$, Lagrange P_2 elements*

Mesh size h	8.60341	8.42192	6.09033	4.72471	4.72145	3.69153	Slope
$\left\|\mathbf{u}-\mathbf{u}^h\right\|_{1,\Omega}$	261.2041	248.7847	142.4049	121.5570	124.0673	59.6301	1.5405
η_1	80200.8	78955.6	58608	40699.3	40816.2	28237.3	1.2090
η_2	98066.1	98430	78979.6	65528.8	65304.7	54911.2	0.6951
η_3	5824.05	5734.21	6074.19	3196.26	3159.32	2468.21	1.0548
η_4	166077	165873	57204.1	90815.3	91107.4	62615.7	1.0431
η	208960	208493	113938	119198	119336	87973.9	0.9753
Effectivity index Eff_E	0.3200	0.3352	0.3200	0.3922	0.3847	0.5901	

Acknowledgements

Région de Franche-Comté is acknowledged for partial funding ('Convention Région 2015C-4991. Modèles mathématiques et méthodes numériques pour l'élasticité non-linéaire'). The authors thank

Roland Becker, Daniela Capatina and Robert Luce for helpful discussions and comments, as well as the anonymous referee for his suggestions.

References

ADAMS, R. A. (1975) *Sobolev Spaces*. Pure and Applied Mathematics, vol. 65. New York, London: Academic Press.

AINSWORTH, M. & ODEN, J. T. (2000) *A Posteriori Error Estimation in Finite Element Analysis*. New York: Wiley-Interscience.

ALART, P. & CURNIER, A. (1988) A generalized Newton method for contact problems with friction. *J. Mech. Theor. Appl.*, **7**, 67–82.

BANZ, L. & STEPHAN, E. P. (2015) On *hp*-adaptive BEM for frictional contact problems in linear elasticity. *Comput. Math. Appl.*, **69**, 559–581.

BECKER, R. (2002) Mesh adaptation for Dirichlet flow control via Nitsche's method. *Comm. Numer. Methods Eng.*, **18**, 669–680.

BECKER, R., HANSBO, P. & STENBERG, R. (2003) A finite element method for domain decomposition with non-matching grids. *Math. Model. Numer. Anal.*, **37**, 209–225.

BECKER, R. & RANNACHER, R. (1996) A feed-back approach to error control in finite element methods: basic analysis and examples. *J. Numer. Math.*, **4**, 237–264.

BECKER, R. & RANNACHER, R. (2001) An optimal control approach to a posteriori error estimation in finite element methods. *Acta Numer.*, **10**, 1–102.

BERNARDI, C. & GIRAULT, V. (1998) A local regularization operator for triangular and quadrilateral finite elements. *SIAM J. Numer. Anal.*, **35**, 1893–1916.

BLUM, H. & SUTTMEIER, F. (2000) An adaptive finite element discretization for a simplified Signorini problem. *Calcolo*, **37**, 65–77.

BOSTAN, V. & HAN, W. (2006) A posteriori error analysis for finite element solutions of a frictional contact problem. *Comput. Methods Appl. Mech. Eng.*, **195**, 1252–1274.

BRAESS, D. & VERFÜRTH, R. (1996) A posteriori error estimators for the Raviart-Thomas element. *SIAM J. Numer. Anal.*, **33**, 2431–2444.

BRENNER, S.-C. & SCOTT, L.-R. (2007) *The Mathematical Theory of Finite Element Methods*. Texts in Applied Mathematics, vol. 15. New York: Springer.

BUSCAGLIA, G. C., DURÁN, R., FANCELLO, E. A., FEIJÓO, R. A. & PADRA, C. (2001) An adaptive finite element approach for frictionless contact problems. *Internat. J. Numer. Methods Eng.*, **50**, 395–418.

CARSTENSEN, C. (1997) A posteriori error estimate for the mixed finite element method. *Math. Comp.*, **66**, 465–476.

CARSTENSEN, C., SCHERF, O. & WRIGGERS, P. (1999) Adaptive finite elements for elastic bodies in contact. *SIAM J. Sci. Comput.*, **20**.

CHOULY, F. (2014) An adaptation of Nitsche's method to the Tresca friction problem. *J. Math. Anal. Appl.*, **411**, 329–339.

CHOULY, F. & HILD, P. (2013) A Nitsche-based method for unilateral contact problems: numerical analysis. *SIAM J. Numer. Anal.*, **51**, 1295–1307.

CHOULY, F., HILD, P. & RENARD, Y. (2015) Symmetric and non-symmetric variants of Nitsche's method for contact problems in elasticity: theory and numerical experiments. *Math. Comp.*, **84**, 1089–1112.

CIARLET, P. G. (1991) The finite element method for elliptic problems. *Handbook of Numerical Analysis*, vol. II, chap. 1 (P. G. Ciarlet & J. L. Lions eds). North Holland, pp. 17–352.

COOREVITS, P., HILD, P. & HJIAJ, M. (2001) A posteriori error control of finite element approximations for Coulomb's frictional contact. *SIAM J. Sci. Comput.*, **23**, 976–999.

COOREVITS, P., HILD, P. & PELLE, J.-P. (2000) A posteriori error estimation for unilateral contact with matching and nonmatching meshes. *Comput. Methods Appl. Mech. Eng.*, **186**, 65–83.

DÖRSEK, P. & MELENK, J. M. (2010) Adaptive *hp*-FEM for the contact problem with Tresca friction in linear elasticity: the primal–dual formulation and a posteriori error estimation. *Appl. Numer. Math.*, **60**, 689–704.

ECK, C. & WENDLAND, W. (2013) A residual-based error estimator for BEM-discretizations of contact problems. *Numer. Math.*, **95**, 253–282.

ERN, A. & GUERMOND, J.-L. (2004) *Theory and Practice of Finite Elements*. Applied Mathematical Sciences, vol. 159. New York: Springer.

FABRE, M., POUSIN, J. AND RENARD, Y. (2016) A fictitious domain method for frictionless contact problems in elasticity using Nitsche's method. *SMAI J. Comput. Math.*, **2**, 19–50.

FERNÁNDEZ, J. R. & HILD, P. (2010) A posteriori error analysis for the normal compliance problem. *Appl. Numer. Math.*, **60**, 64–73.

FICHERA, G. (1963–1964) Problemi elastostatici con vincoli unilaterali: Il problema di Signorini con ambigue condizioni al contorno. *Atti Accad. Naz. Lincei Mem. Cl. Sci. Fis. Mat. Natur. Sez. Ia*, **7**, 91–140.

HANSBO, A. & HANSBO, P. (2002) An unfitted finite element method, based on Nitsche's method, for elliptic interface problems. *Comput. Methods Appl. Mech. Eng.*, **191**, 5537–5552.

HANSBO, A., HANSBO, P. & LARSON, M. G. (2003) A finite element method on composite grids based on Nitsche's method. *Math. Model. Numer. Anal.*, **37**, 495–514.

HASLINGER, J., HLAVÁČEK, I. & NEČAS, J. (1996) Numerical methods for unilateral problems in solid mechanics. *Handbook of Numerical Analysis* (P. G. Ciarlet & J. L. Lions eds), vol. IV, chap. 2. North Holland, pp. 313–385.

HILD, P. & LLERAS, V. (2009) Residual error estimators for Coulomb friction, *SIAM J. Numer. Anal.*, **47**, 3550–3583.

HILD, P. & NICAISE, S. (2005) A posteriori error estimations of residual type for Signorini's problem. *Numer. Math.*, **101**, 523–549.

HILD, P. & NICAISE, S. (2007) Residual a posteriori error estimators for contact problems in elasticity. *Math. Model. Numer. Anal.*, **41**, 897–923.

HLAVÁČEK, I., HASLINGER, J., NEČAS, J. & LOVÍŠEK, J. (1988) *Solution of Variational Inequalities in Mechanics*. Applied Mathematical Sciences, vol. 66. New York: Springer. Translated from the Slovak by J. Jarník.

JUNTUNEN, M. & STENBERG, R. (2008) On a mixed discontinuous Galerkin method. *Electron. Trans. Numer. Anal.*, **32**, 17–32.

JUNTUNEN, M. & STENBERG, R. (2009) Nitsche's method for general boundary conditions. *Math. Comp.*, **78**, 1353–1374.

JUNTUNEN, M. & STENBERG, R. (2010) Analysis of finite element methods for the Brinkman problem. *Calcolo*, **47**, 129–147.

KIKUCHI, N. & ODEN, J. T. (1988) *Contact Problems in Elasticity: A Study of Variational Inequalities and Finite Element Methods*. SIAM Studies in Applied Mathematics, vol. 8. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM).

KIM, K. Y. (2007) A posteriori error analysis for locally conservative mixed methods. *Math. Comp.*, **76**, 43–66.

KÖNNÖ, J. & STENBERG, R. (2011) $H(\mathrm{div})$-conforming finite elements for the Brinkman problem. *Math. Models Methods Appl. Sci.*, **21**, 2227–2248.

KRAUSE, R., VEESER, A. & WALLOTH, M. (2015) An efficient and reliable residual-type a posteriori error estimator for the Signorini problem. *Numer. Math.*, **130**, 151–197.

LAURSEN, T. A. (2003) *Computational Contact and Impact Mechanics*. Berlin: Springer.

LEE, C.-Y. & ODEN, J.-T. (1994) A posteriori error estimation of *h-p* finite element approximations of frictional contact problems. *Comput. Methods Appl. Mech. Eng.*, **113**, 11–45.

LLERAS, V. (2009) Modélisation, analyse et simulation de problèmes de contact en mécanique des solides et des fluides. *Ph.D. Thesis*, Besançon.

LOUF, F., COMBE, J.-P. & PELLE, J.-P. (2003) Constitutive error estimator for the control of contact problems involving friction. *Comput. Struct.*, **81**, 1759–1772.

MAISCHAK, M. & STEPHAN, E. (2005) Adaptive hp-versions of BEM for Signorini problems. *Appl. Numer. Math.*, **54**, 425–449.

RENARD, Y. (2013) Generalized Newton's methods for the approximation and resolution of frictional contact problems in elasticity. *Comp. Methods Appl. Mech. Eng.*, **256**, 38–55.

SCHRÖDER, A. (2012) A posteriori error estimates of higher-order finite elements for frictional contact problems. *Comput. Methods Appl. Mech. Eng.*, **249/252**, 151–157.
SCHRÖDER, A. & RADEMACHER, A. (2011) Goal-oriented error control in adaptive mixed FEM for Signorini's problem. *Comp. Methods Appl. Mech. Eng.*, **200**, 345–355.
STEIN, E. & OHNIMUS, S. (1999) Anisotropic discretization- and model-error estimation in solid mechanics by local Neumann problems. *Comput. Methods Appl. Mech. Eng.*, **176**, 363–385.
VERFÜRTH, R. (1999) A review of a posteriori error estimation techniques for elasticity problems. *Comput. Methods Appl. Mech. Eng.*, **176**, 419–440.
WANG, L.-H. (2000) On the error estimate of linear finite element approximation to the elastic contact problem with curved contact boundary. *J. Comput. Math.*, **18**, 561–566.
WEISS, A. & WOHLMUTH, B. I. (2009) A posteriori error estimator and error control for contact problems. *Math. Comp.*, 1237–1267.
WOHLMUTH, B. I. (1999) A residual based error estimator for mortar finite element discretizations. *Numer. Math.*, **84**, 143–171.
WOHLMUTH, B. I. (2007) An a posteriori error estimator for two body contact problems on non-matching meshes. *J. Sci. Comput.*, **33**, 25–45.
WRIGGERS, P. (2006) *Computational Contact Mechanics*, 2nd edn. Springer-Verlag, Heidelberg, Berlin: Wiley.
WRIGGERS, P. & SCHERF, O. (1998) Different a posteriori error estimators and indicators for contact problems. *Math. Comput. Model.*, **28**, 1605–1626.

Appendix. Extension to the Tresca friction case

We extend, in this appendix, the analysis of Section 3 to the case of unilateral contact with Tresca friction. For *a posteriori* error estimates for the Tresca friction problem one may refer, e.g., to Dörsek & Melenk (2010) and references therein.

Setting and Nitsche-based finite element method for Tresca

Let $s \in L^2(\Gamma_C)$, $s \geq 0$ be a given threshold. The Tresca friction problem with unilateral contact consists in finding the displacement field $\mathbf{u} : \Omega \to \mathbb{R}^d$ verifying the equations and conditions (2.1)–(2.2 (*i*,*ii*,*iii*))–(A.1), with (A.1) given by

$$\begin{cases} |\boldsymbol{\sigma}_{\mathbf{t}}(\mathbf{u})| \leq s & \text{if } \mathbf{u_t} = \mathbf{0}, \quad \text{(i)} \\ \boldsymbol{\sigma}_{\mathbf{t}}(\mathbf{u}) = -s\dfrac{\mathbf{u_t}}{|\mathbf{u_t}|} & \text{otherwise}, \quad \text{(ii)} \end{cases} \tag{A.1}$$

where $|\cdot|$ stands for the Euclidean norm in $\mathbb{R}^{d-1}$.

For any $\alpha \in \mathbb{R}^+$, we introduce the notation $[\cdot]_\alpha$ for the orthogonal projection onto $\mathscr{B}(\mathbf{0}, \alpha) \subset \mathbb{R}^{d-1}$, where $\mathscr{B}(\mathbf{0}, \alpha)$ is the closed ball centered at the origin $\mathbf{0}$ and of radius α. The following property holds for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{d-1}$:

$$(\mathbf{y} - \mathbf{x}) \cdot ([\mathbf{y}]_\alpha - [\mathbf{x}]_\alpha) \geq |[\mathbf{y}]_\alpha - [\mathbf{x}]_\alpha|^2, \tag{A.2}$$

where $\cdot$ is the Euclidean scalar product in $\mathbb{R}^{d-1}$.

Let us introduce the discrete linear operator $\mathbf{P}^{\mathbf{t}}_\gamma : \mathbf{v}^h \mapsto \mathbf{v}^h_{\mathbf{t}} - \gamma\, \boldsymbol{\sigma}_{\mathbf{t}}(\mathbf{v}^h)$ and the bilinear form $A_{\theta\gamma}(\mathbf{u}^h, \mathbf{v}^h) := a(\mathbf{u}^h, \mathbf{v}^h) - \displaystyle\int_{\Gamma_C} \theta\gamma\, \boldsymbol{\sigma}(\mathbf{u}^h)\mathbf{n} \cdot \boldsymbol{\sigma}(\mathbf{v}^h)\mathbf{n}\, \mathrm{d}\Gamma$. The extension of our Nitsche-based method for unilateral contact with Tresca friction then reads

$$\begin{cases} \text{find } \mathbf{u}^h \in \mathbf{V}^h \text{ such that} \\ A_{\theta\gamma}(\mathbf{u}^h, \mathbf{v}^h) + \displaystyle\int_{\Gamma_C} \frac{1}{\gamma} \, [P_\gamma(\mathbf{u}^h) - g]_+ P_{\theta\gamma}(\mathbf{v}^h) \, \mathrm{d}\Gamma \\ + \displaystyle\int_{\Gamma_C} \frac{1}{\gamma} \, \left[\mathbf{P}^{\mathbf{t}}_\gamma(\mathbf{u}^h)\right]_{\gamma s} \cdot \mathbf{P}^{\mathbf{t}}_{\theta\gamma}(\mathbf{v}^h) \, \mathrm{d}\Gamma = L(\mathbf{v}^h) \qquad \forall \, \mathbf{v}^h \in \mathbf{V}^h. \end{cases} \tag{A.3}$$

Consistency, well-posedness and *a priori* error estimates for method (A.3) are established in Chouly (2014).

Residual error estimator, upper and lower bound

Definition 3.2 still holds for problem (A.3), except for η_{3K} whose expression is now

$$\eta_{3K} = h_K^{1/2} \left(\sum_{E \in E_K^C} \left\| \frac{1}{\gamma} \left[\mathbf{P}^{\mathbf{t}}_\gamma(\mathbf{u}^h)\right]_{\gamma s} + \boldsymbol{\sigma}_{\mathbf{t}}(\mathbf{u}^h) \right\|_{0,E}^2 \right)^{1/2}.$$

First, we provide counterparts to Assumption 3.3 and to the discrete trace inequality of Lemma 2.3.

ASSUMPTION A.1 The solution $\mathbf{u}$ of (2.1)–(2.2 (i,ii,iii))–(A.1) and the discrete solution $\mathbf{u}^h$ of (A.3) are such that

$$\left\|\sigma_n(\mathbf{u} - \mathbf{u}^h)\right\|_{-1/2,h,\Gamma_C} + \left\|\boldsymbol{\sigma}_{\mathbf{t}}(\mathbf{u} - \mathbf{u}^h)\right\|_{-1/2,h,\Gamma_C} \lesssim \|\mathbf{u} - \mathbf{u}^h\|_{1,\Omega}. \tag{A.4}$$

LEMMA A.2 For any $\mathbf{v}^h \in \mathbf{V}^h$, we have

$$\|\sigma_n(\mathbf{v}^h)\|_{-1/2,h,\Gamma_C} + \|\boldsymbol{\sigma}_{\mathbf{t}}(\mathbf{v}^h)\|_{-1/2,h,\Gamma_C} \lesssim \|\mathbf{v}^h\|_{1,\Omega}. \tag{A.5}$$

For contact with Tresca friction, the following statement guarantees the reliability of the *a posteriori* error estimator.

THEOREM A.3 Let $\mathbf{u}$ be the solution to (2.1)–(2.2 (i,ii,iii))–(A.1), with $\mathbf{u} \in (H^{3/2+\nu}(\Omega))^d$ ($\nu > 0$ and $d = 2, 3$), and let $\mathbf{u}^h$ be the solution to the corresponding discrete problem (A.3). Assume that, for $\theta \neq -1$, γ_0 is sufficiently small, and otherwise that $\gamma_0 > 0$ for $\theta = -1$. Assume that the saturation assumption (A.4) holds as well. Then we have

$$\|\mathbf{u} - \mathbf{u}^h\|_{1,\Omega} + \left\|\sigma_n(\mathbf{u}) + \frac{1}{\gamma}[P_\gamma(\mathbf{u}^h) - g]_+\right\|_{-1/2,h,\Gamma_C} + \left\|\boldsymbol{\sigma}_{\mathbf{t}}(\mathbf{u}) + \frac{1}{\gamma}\left[\mathbf{P}^{\mathbf{t}}_\gamma(\mathbf{u}^h)\right]_{\gamma s}\right\|_{-1/2,h,\Gamma_C}$$
$$+ \|\sigma_n(\mathbf{u}) - \sigma_n(\mathbf{u}^h)\|_{-1/2,h,\Gamma_C} + \|\boldsymbol{\sigma}_{\mathbf{t}}(\mathbf{u}) - \boldsymbol{\sigma}_{\mathbf{t}}(\mathbf{u}^h)\|_{-1/2,h,\Gamma_C} \lesssim (1 + \gamma_0)(\eta + \zeta).$$

Proof. The proof is a direct adaptation of Theorem 3.5. Let $\mathbf{v}^h \in \mathbf{V}^h$. To lighten the notation, we define $\mathbf{e} := \mathbf{u} - \mathbf{u}^h$. We start as in Theorem 3.5 and get

$$\alpha\|\mathbf{e}\|_{1,\Omega}^2 \le \mathcal{T}_1 + \mathcal{T}_2,$$

where α is the $\mathbf{V}$-ellipticity constant of $a(\cdot,\cdot)$ and

$$\begin{aligned}
\mathcal{T}_1 &:= L(\mathbf{u}-\mathbf{v}^h) - a(\mathbf{u}^h, \mathbf{u}-\mathbf{v}^h)\\
&\quad + \int_{\Gamma_C} \frac{1}{\gamma}[P_\gamma(\mathbf{u}^h) - g]_+ (v_n^h - u_n)\,\mathrm{d}\Gamma + \int_{\Gamma_C} \frac{1}{\gamma}\left[\mathbf{P}_\gamma^{\mathbf{t}}(\mathbf{u}^h)\right]_{\gamma s} \cdot (\mathbf{v}_{\mathbf{t}}^h - \mathbf{u}_{\mathbf{t}})\,\mathrm{d}\Gamma,\\
\mathcal{T}_2 &:= \int_{\Gamma_C} \boldsymbol{\sigma}(\mathbf{u})\mathbf{n}\cdot(\mathbf{u}-\mathbf{u}^h)\,\mathrm{d}\Gamma + \int_{\Gamma_C} \frac{1}{\gamma}[P_\gamma(\mathbf{u}^h) - g]_+ P_{\theta\gamma}(\mathbf{u}-\mathbf{u}^h)\,\mathrm{d}\Gamma\\
&\quad + \int_{\Gamma_C} \frac{1}{\gamma}\left[\mathbf{P}_\gamma^{\mathbf{t}}(\mathbf{u}^h)\right]_{\gamma s} \cdot \mathbf{P}_{\theta\gamma}^{\mathbf{t}}(\mathbf{u}-\mathbf{u}^h)\,\mathrm{d}\Gamma\\
&\quad - \theta\int_{\Gamma_C} \frac{1}{\gamma}[P_\gamma(\mathbf{u}^h) - g]_+ \gamma\sigma_n(\mathbf{v}^h - \mathbf{u})\,\mathrm{d}\Gamma - \theta\int_{\Gamma_C} \frac{1}{\gamma}\left[\mathbf{P}_\gamma^{\mathbf{t}}(\mathbf{u}^h)\right]_{\gamma s} \cdot \gamma\boldsymbol{\sigma}_{\mathbf{t}}(\mathbf{v}^h - \mathbf{u})\,\mathrm{d}\Gamma\\
&\quad - \theta\int_{\Gamma_C} \gamma\,\boldsymbol{\sigma}(\mathbf{u}^h)\mathbf{n}\cdot\boldsymbol{\sigma}(\mathbf{v}^h - \mathbf{u}^h)\mathbf{n}\,\mathrm{d}\Gamma.
\end{aligned}$$

The quantity $\mathcal{T}_1$ is bounded almost exactly as in Theorem 3.5, except for the new Tresca friction term, that is bounded as

$$\left|\sum_{E\in E_h^C} \int_E \left(\frac{1}{\gamma}\left[\mathbf{P}_\gamma^{\mathbf{t}}(\mathbf{u}^h)\right]_{\gamma s} + \boldsymbol{\sigma}_{\mathbf{t}}(\mathbf{u}^h)\right)\cdot(\mathbf{v}_{\mathbf{t}}^h - \mathbf{u}_{\mathbf{t}})\,\mathrm{d}\Gamma\right| \lesssim \eta\|\mathbf{e}\|_{1,\Omega}.$$

Note that the remaining terms in $\mathcal{T}_2$ can be split as

$$\mathcal{T}_2 = \mathcal{T}_2^{\mathrm{C}} + \mathcal{T}_2^{\mathrm{T}},$$

where $\mathcal{T}_2^{\mathrm{C}}$ represents the contact terms and $\mathcal{T}_2^{\mathrm{T}}$ contains the Tresca friction terms. The contact terms $\mathcal{T}_2^{\mathrm{C}}$ are handled as in Theorem 3.5. Moreover, we can bound the friction terms $\mathcal{T}_2^{\mathrm{T}}$ in a similar fashion, following step by step the proof of Theorem 3.5 and using the bound (A.2). We get finally for any $\beta > 0$

$$\mathcal{T}_2^{\mathrm{T}} \le \left(\frac{1}{2\beta} + \frac{(\theta+1)^2}{4}\right)\gamma_0\|\boldsymbol{\sigma}_{\mathbf{t}}(\mathbf{u}-\mathbf{u}^h)\|_{-1/2,h,\Gamma_C}^2 + \beta\gamma_0\theta^2\eta^2 + \frac{\gamma_0}{2\beta}\|\boldsymbol{\sigma}_{\mathbf{t}}(\mathbf{v}^h - \mathbf{u}^h)\|_{-1/2,h,\Gamma_C}^2.$$

Using (A.5) and the H^1-stability of R^h (see (2.8) in Lemma 2.1) we bound

$$\|\boldsymbol{\sigma}_{\mathbf{t}}(\mathbf{v}^h - \mathbf{u}^h)\|_{-1/2,h,\Gamma_C} \le C\|\mathbf{v}^h - \mathbf{u}^h\|_{1,\Omega} = C\|R^h(\mathbf{u}-\mathbf{u}^h)\|_{1,\Omega} \le C\|\mathbf{u}-\mathbf{u}^h\|_{1,\Omega}.$$

We combine this last bound with the saturation assumption (A.4) and get (remembering the result that holds for the contact terms $\mathcal{T}_2^{\mathrm{C}}$)

$$\mathcal{T}_2 \leq C\gamma_0 \left(\frac{(\theta+1)^2}{4} + \frac{1}{\beta} \right) \|\mathbf{u} - \mathbf{u}^h\|_{1,\Omega}^2 + \beta\gamma_0\theta^2(\eta+\zeta)^2.$$

From now on the proof is exactly the same as in Theorem 3.5. $\square$

REMARK A.4 An extension of Theorem 3.6 holds as well for problem (A.3) and similar local lower error bounds can be derived following the same method. The only difference is that the term η_{3K} is bounded as

$$\eta_{3K} \lesssim \sum_{E \in E_K^{\mathrm{C}}} h_K^{1/2} \left(\left\| \boldsymbol{\sigma}_{\mathbf{t}}(\mathbf{u}) + \frac{1}{\gamma} \left[\mathbf{P}_\gamma^{\mathbf{t}}(\mathbf{u}^h) \right]_{\gamma s} \right\|_{0,E} + \left\| \boldsymbol{\sigma}_{\mathbf{t}}(\mathbf{u} - \mathbf{u}^h) \right\|_{0,E} \right).$$