IMA Journal of Numerical Analysis (2018) **00**, 1–35
doi:10.1093/imanum/dry050

Optimal error estimates of the unilateral contact problem in a curved and smooth boundary domain by the penalty method

IBRAHIMA DIONE
Département de mathématiques et de statistique, Pavillon Vachon, 1045 Avenue de la Médecine, Université Laval, Québec, G1V 0A6, Canada
ibrahima.dione.1@ulaval.ca

[Received on 31 May 2017; revised on 20 June 2018]

We consider linear finite elements to approximate the elasticity equations with unilateral contact boundary conditions, in a bounded two- or three-dimensional domain with curved and smooth boundary. We use the penalty method to weakly impose these boundary conditions. We establish an error estimate in the energy norm with respect to the mesh size h and the penalty parameter ε. Assuming $\boldsymbol{H}^{\frac{3}{2}+\nu}(\Omega)$ regularity of the solution, $0 < \nu \leq \frac{1}{2}$, we obtain an $\mathcal{O}(h^{\frac{1}{2}+\nu} + \varepsilon^{\frac{1}{2}+\nu})$ convergence rate. Therefore, if the penalty parameter is chosen as $\varepsilon(h) := ch^{\theta}$ with $0 < \theta \leq 1$, we obtain an $\mathcal{O}(h^{\theta(\frac{1}{2}+\nu)})$ convergence rate. Thus, the optimal linear convergence rate is obtained when ε behaves like h (that is, $\theta = 1$) and $\nu = \frac{1}{2}$. We present a numerical example to illustrate the theoretical analysis.

Keywords: *a priori* error estimates; curved and smooth boundary; frictionless contact; linear finite elements; penalty method; polygonal or tetrahedral triangulations; unilateral contact.

1. Introduction

Let $\Omega \subset \mathbb{R}^d, d = 2, 3$, be a bounded domain with curved and smooth boundary representing the configuration of a linearly elastic body, which satisfies the equilibrium equations

$$\begin{aligned} -\nabla \cdot \boldsymbol{\sigma}(\boldsymbol{u}) &= \boldsymbol{f} && \text{in } \Omega, \\ \boldsymbol{\sigma}(\boldsymbol{u}) &= \boldsymbol{A}\boldsymbol{\varepsilon}(\boldsymbol{u}) && \text{in } \Omega, \end{aligned} \tag{1.1}$$

where $\boldsymbol{u}$ is the displacement field, $\boldsymbol{f} \in (L^2(\Omega))^d$ is a body force and $\boldsymbol{\sigma} := (\sigma_{ij})_{1\leq i,j\leq d}$ is the stress tensor field. The tensor $\boldsymbol{A}$ is the fourth-order symmetric elasticity tensor satisfying the usual uniform ellipticity and boundedness properties. The strain–displacement tensor ε is defined by

$$\varepsilon(\boldsymbol{u}) := \frac{1}{2}\left(\nabla \boldsymbol{u} + \nabla \boldsymbol{u}^{\mathrm{T}}\right).$$

We associate to (1.1) the so-called unilateral contact conditions defined on the part Γ_c of the boundary $\partial\Omega$,

$$\boldsymbol{u}\cdot\boldsymbol{n} - \ell \leq 0, \quad \boldsymbol{\sigma}_n(\boldsymbol{u}) \leq 0, \quad (\boldsymbol{u}\cdot\boldsymbol{n} - \ell)\,\boldsymbol{\sigma}_n(\boldsymbol{u}) = 0, \ \text{ on } \Gamma_c, \tag{1.2}$$

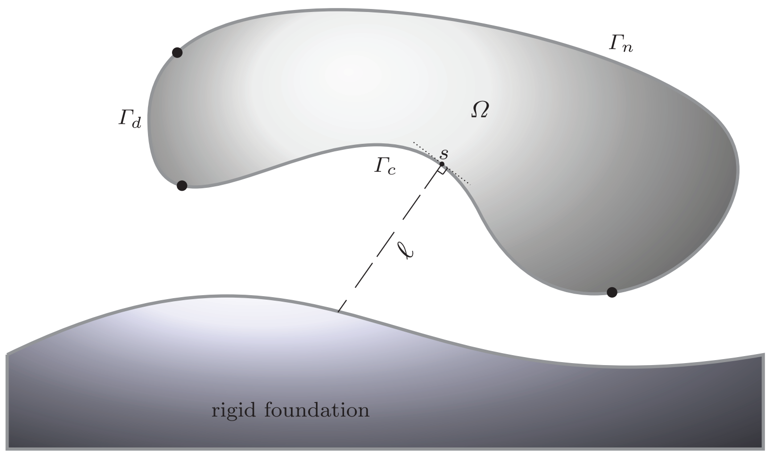

FIG. 1. A deformable bounded body Ω of curved and smooth boundary becoming in contact with a rigid foundation.

where the vector $\boldsymbol{n}$ is the unit outward normal to the boundary $\partial\Omega$ and $\ell : \Gamma_c \longrightarrow \mathbb{R}^+$ is a continuous mapping associating every point $\boldsymbol{s} \in \Gamma_c$ with its normal distance to the rigid foundation (see Fig. 1). The contact pressure $\boldsymbol{\sigma}_n(\boldsymbol{u})$ is the normal component of the surface force $\boldsymbol{\sigma}(\boldsymbol{u})\boldsymbol{n}$ and is defined by

$$\boldsymbol{\sigma}_n(\boldsymbol{u}) := (\boldsymbol{\sigma}(\boldsymbol{u})\boldsymbol{n}) \cdot \boldsymbol{n}.$$

We introduce the tangential component of the surface force $\boldsymbol{\sigma}(\boldsymbol{u})\boldsymbol{n}$ as

$$\boldsymbol{\sigma}_t(\boldsymbol{u}) := (\boldsymbol{\sigma}(\boldsymbol{u})\boldsymbol{n}) \cdot \boldsymbol{t},$$

and we assume there is no friction effect on the contact surface Γ_c by imposing the following frictionless boundary condition (which completes the unilateral boundary conditions in (1.2)):

$$\boldsymbol{\sigma}_t(\boldsymbol{u}) = 0 \text{ on } \Gamma_c. \tag{1.3}$$

We complete problem (1.1)–(1.3) by applying Dirichlet and Neumann boundary conditions, respectively, on the remaining boundary parts Γ_d and Γ_n of the entire boundary $\partial\Omega$,

$$\boldsymbol{u} = \boldsymbol{0} \quad \text{on} \quad \Gamma_d, \tag{1.4}$$

$$\boldsymbol{\sigma}(\boldsymbol{u})\boldsymbol{n} = \boldsymbol{g} \quad \text{on} \quad \Gamma_n, \tag{1.5}$$

where $\boldsymbol{g} \in \boldsymbol{L}^2(\Gamma_n)$ is a surface load. The bounded domain Ω is Lipschitz and its boundary (which is curved and smooth) is partitioned into the three non-overlapping parts Γ_d, Γ_n and Γ_c ($\partial\Omega = \overline{\Gamma}_d \cup \overline{\Gamma}_n \cup \overline{\Gamma}_c$) satisfying meas$(\Gamma_d) > 0$, meas$(\Gamma_c) > 0$ (see Fig. 1).

The unilateral contact problem then consists in finding the displacement field $\boldsymbol{u} : \overline{\Omega} \longrightarrow \mathbb{R}^d$ satisfying (1.1)–(1.5).

We denote by $\boldsymbol{x}$ and $\boldsymbol{s}$ generic points in Ω and on the boundary $\partial\Omega$, respectively. We recall the classical Sobolev spaces $H^m(\Omega)$, $m \geq 0$ and $L^2(\Omega) := H^0(\Omega)$, equipped with the usual norm $\|\cdot\|_{m,\Omega}$ (Adams & Fournier, 2003). We denote in bold the Sobolev spaces for vectorial functions:

$$\boldsymbol{H}^m(\Omega) := \left(H^m(\Omega)\right)^d \text{ and } \boldsymbol{L}^2(\Omega) := \left(L^2(\Omega)\right)^d.$$

Instead of working under the constraint induced by the non-inter-penetration condition $\boldsymbol{u}\cdot\boldsymbol{n} - \ell \leq 0$ on the contact zone Γ_c, techniques more favorable for a computational purpose may be needed. One of the classical methods to circumvent this constraint is the penalty method. The penalty technique is a classical method for the numerical treatment of constrained problems (see Kikuchi & Oden, 1988 and Li, 1998) and is readily applicable in most numerical codes. Nevertheless, this method remains an approximation technique since the solution of the penalized problem (which we denote by $\boldsymbol{u}_\varepsilon$) is expected to approximate as accurately as possible the solution $\boldsymbol{u}$ of the problem (1.1)–(1.5), when the penalty parameter ε tends to zero.

We apply in this paper the penalty method to weakly impose the non-inter-penetration condition and construct finite element approximations of the resulting penalized problem. Using the penalty technique to study this problem has been the subject of various studies, both in a polygonal boundary domain (Chernov *et al.*, 2007, Chouly & Hild, 2013, Dione, 2018) and in a curved and smooth boundary domain (Kikuchi & Song, 1981, Kikuchi & Oden, 1988) as is the case here. We denote by h the size of elements constituting the mesh over the domain Ω or the polygonal approximation domain Ω_h of Ω. Denoting by $\boldsymbol{u}_{\varepsilon h}$ the solution of the penalty finite element approximation problem, it is expected to achieve the *a priori* error estimate

$$\|\boldsymbol{u} - \boldsymbol{u}_{\varepsilon h}\|_{1,\Omega} = \mathcal{O}(h), \tag{1.6}$$

if the finite element approximation spaces $\boldsymbol{V}_h$ consist of continuous and piecewise affine functions and if the penalty parameter is taken in the form $\varepsilon(h) = ch^\theta$, with a suitable value of θ. However, it is not an easy task to obtain the optimal convergence estimate (1.6), which is not even proven in all cases. For instance in the context of a polygonal boundary $\partial\Omega$, where the domain Ω can be exactly partitioned into a union of triangles (resulting thus in conforming approximations $\boldsymbol{V}_h \subset \boldsymbol{V}$), the best error estimate obtained without any specific operator and numerical integration rule (for the treatment of the discrete penalty term) is established by Chouly & Hild (2013, Theorem 3.2). They obtain the *a priori* estimate

$$\begin{aligned}\|\boldsymbol{u} - \boldsymbol{u}_{\varepsilon h}\|_{1,\Omega} &+ \varepsilon^{\frac{1}{2}} \left\|\boldsymbol{\sigma}_n(\boldsymbol{u}) + \frac{1}{\varepsilon}[\boldsymbol{u}_{\varepsilon h}\cdot\boldsymbol{n}]^+\right\|_{0,\Gamma_c} \\ &\leq c \begin{cases} \left(h^{\frac{1}{2}+\frac{\nu}{2}+\nu^2} + h^\nu \varepsilon^{\frac{1}{2}} + h^{\nu-\frac{1}{2}}\varepsilon\right)\|\boldsymbol{u}\|_{\frac{3}{2}+\nu,\Omega}, & \text{if } 0 < \nu < \frac{1}{2}, \\ \left(h\,|\ln(h)|^{\frac{1}{2}} + (h\varepsilon)^{\frac{1}{2}} + \varepsilon\right)\|\boldsymbol{u}\|_{2,\Omega}, & \text{if } \nu = \frac{1}{2}, \end{cases}\end{aligned} \tag{1.7}$$

where for recovering the best convergence rate, namely $\mathcal{O}(h\,|\ln(h)|^{\frac{1}{2}})$, the penalty parameter is to be fixed as $\varepsilon(h) = ch$. Recently, Dione (2018) improved the estimate (1.7) by establishing, for $0 < \nu \leq \frac{1}{2}$, the *a priori* estimate

$$\|\boldsymbol{u} - \boldsymbol{u}_{\varepsilon h}\|_{1,\Omega} + \left(\varepsilon^{\frac{1}{2}} - ch^{\frac{1}{2}}\right)\left\|\boldsymbol{\sigma}_n(\boldsymbol{u}) + \frac{1}{\varepsilon}[\boldsymbol{u}_{\varepsilon h}\cdot\boldsymbol{n}]^+\right\|_{0,\Gamma_c} \leq c\left(h^{\frac{1}{2}+\nu} + \varepsilon^{\frac{1}{2}+\nu}\right)\|\boldsymbol{u}\|_{\frac{3}{2}+\nu,\Omega}. \tag{1.8}$$

Estimate (1.8) provides an $\mathcal{O}\,(h^{\theta(\frac{1}{2}+\nu)})$ convergence rate under the regularity $\boldsymbol{u} \in \boldsymbol{H}^{\frac{3}{2}+\nu}(\Omega)$ and if the penalty parameter is taken in the form $\varepsilon(h) = ch^{\theta}$, with $0 < \theta \leq 1$. The consequence of this result is that, the optimal linear convergence order (1.6) is achieved when $\nu = \frac{1}{2}$ and $\theta = 1$.

In the present work, we consider a domain Ω with curved and smooth boundary (of regularity at least $\mathscr{C}^{1,1}$) in which two classes of meshing were proposed in the literature. For the first class, the domain Ω is a subset of a discrete domain D_h consisting of a union of triangular or tetrahedral elements. The integrations involved in the variational formulation are performed exactly over Ω and on its boundary $\partial\Omega$. This class of meshes was considered by Kikuchi & Song (1981) in the context of penalty approximation of the unilateral contact problem. By the 9-node isoparametric quadrilateral (Q_2) elements, they define an operator of numerical integration and consider Simpson's numerical integration rule to treat the discrete penalty term. Under the regularities $\boldsymbol{u} \in \boldsymbol{H}^3(\Omega)$ and $\boldsymbol{\sigma}_n(\boldsymbol{u}) \in H^{\frac{3}{2}}(\Gamma_c)$, they establish the *a priori* estimates

$$\|\boldsymbol{u} - \boldsymbol{u}_{\varepsilon h}\|_{1,\Omega} \leq c\left(h^2 + h^{-\frac{1}{4}}\varepsilon^{\frac{1}{2}}\right), \tag{1.9}$$

$$\left\|\boldsymbol{\sigma}_n(\boldsymbol{u}) + \frac{1}{\varepsilon}[\boldsymbol{u}_{\varepsilon h} \cdot \boldsymbol{n} - \ell]^+\right\|_{0,\Gamma_c} \leq c\left(h^{\frac{3}{2}} + h^{-\frac{3}{4}}\varepsilon^{\frac{1}{2}}\right). \tag{1.10}$$

On the other hand, Kikuchi & Oden (1988, Chaper 6, Theorem 6.8 and inequalities 6.102 and 6.104) improve result (1.9) by establishing the convergence estimate

$$\|\boldsymbol{u} - \boldsymbol{u}_{\varepsilon h}\|_{1,\Omega} \leq c\left(h^k + h^{-\frac{1}{2}}\varepsilon^{\frac{1}{2}}\right), \tag{1.11}$$

with Q_k-quadrilateral (and $\mathscr{P}_1$-triangular) elements and assuming that $\boldsymbol{u} \in \boldsymbol{H}^{k+1}(\Omega)$, where $k = 1, 2$. For the second class of meshes, polygonal or polyhedral approximations $\widetilde{\Omega}_h$ of the smooth domain Ω are built (see Fig. 2). The integrations involved in the variational formulation are performed over $\widetilde{\Omega}_h$ and on its boundary $\partial\widetilde{\Omega}_h$. In this paper we consider this latter case in which numerical integrations involved in the finite element formulation are perturbed (a case of a variational crime in which $\boldsymbol{V}_h \not\subseteq \boldsymbol{V}$). To the best of our knowledge, the finite element analysis of the penalized unilateral contact problem in this context of variational crime has never been explored. However, this type of meshing, detailed for instance in Section 3, is a well-known approach and has already been used in the context of slip boundaries (Bänsch & Deckelnick, 1999; Dione & Urquiza, 2015) or in the case of Dirichlet boundary conditions (Barrett & Elliott, 1986; Dione, 2016). Greatly inspired by the work done in the case of a polygonal boundary in Dione (2018), we establish without any specific numerical integration rule (for treating the discrete penalty term) the following *a priori* estimate, for $0 < \nu \leq \frac{1}{2}$:

$$\begin{aligned}\|\boldsymbol{u} - \overline{\boldsymbol{u}}_{\varepsilon h}\|_{1,\Omega} + \left(\varepsilon^{\frac{1}{2}} - c_o h^{\frac{1}{2}}\right)\left\|\boldsymbol{\sigma}_n(\boldsymbol{u}) + \frac{1}{\varepsilon}[\overline{\boldsymbol{u}}_{\varepsilon h} \cdot \boldsymbol{n} - \ell]^+\right\|_{0,\Gamma_c}& \\ \leq c\left(h^{\frac{1}{2}+\nu} + \varepsilon^{\frac{1}{2}+\nu}\right)\left(\|\boldsymbol{u}\|_{\frac{3}{2}+\nu,\Omega} + \|\boldsymbol{f}\|_{0,\Omega} + \|\boldsymbol{g}\|_{0,\Gamma_n}\right).&\end{aligned} \tag{1.12}$$

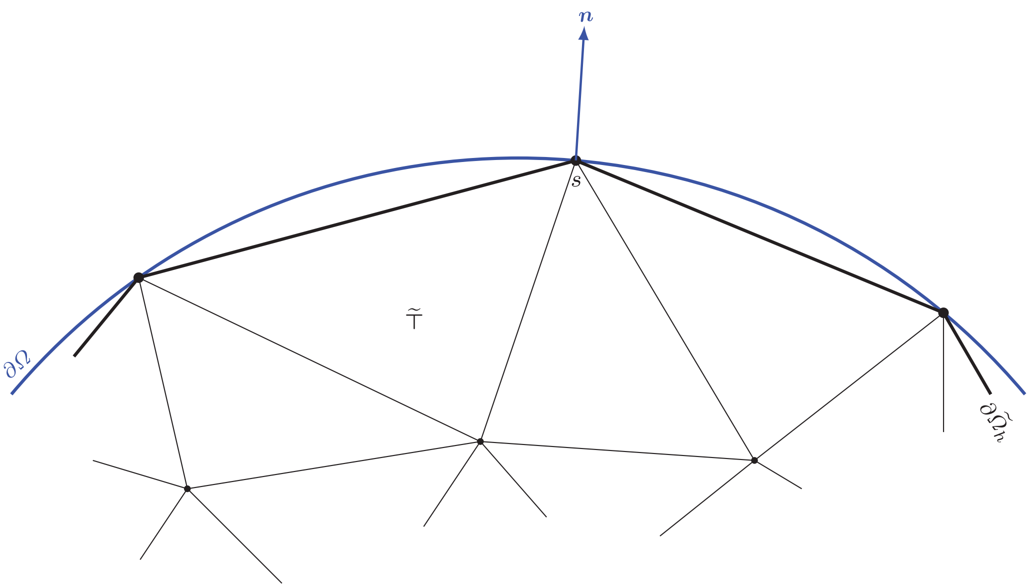

FIG. 2. Polygonal approximation $\partial\widetilde{\Omega}_h$ of the curved and smooth boundary $\partial\Omega$ of the domain $\Omega \subset \mathbb{R}^2$.

The first advantage of estimate (1.12) is that it does not involve a negative power in h, as is the case in estimates (1.7), (1.9), (1.10) and (1.11). Hence, we do not need conditions on the penalty parameter as underlined in Chouly & Hild (2013) for estimate (1.7) (see Remark 3.6), to ensure the convergence of the error term in the left-hand side of estimate (1.12). Moreover, estimate (1.12) satisfies the (optimal) expected *a priori* estimate (1.6) if the penalty parameter is chosen in the form $\varepsilon(h) = (c+1)^2\, h$ and by fixing $\nu = \frac{1}{2}$.

The paper is organized as follows: we present in Section 2 the weak formulation of the unilateral contact problem (1.1)–(1.5) and its penalty weak formulation. Section 3 is devoted to the finite element approximation of the penalized problem, where we consider polygonal (or polyhedral in the three-dimensional case) triangulations of the domain. We present in Section 4 the main results of the paper by providing the *a priori* estimate of problem (1.1)–(1.5). We propose a first step of the *a priori* estimate of the problem through Theorem 4.1, from which we derive the complete form of such an estimate in Theorem 4.2. A numerical example is presented in Section 5 to illustrate our theoretical results.

2. Penalty formulation of the unilateral contact problem

Let us introduce the set of admissible displacements $\boldsymbol{K}$, satisfying the non-inter-penetration condition $\boldsymbol{u}\cdot\boldsymbol{n} - \ell \le 0$ on the contact zone Γ_c:

$$\boldsymbol{K} := \Big\{\boldsymbol{v} \in \boldsymbol{V} : \boldsymbol{v}\cdot\boldsymbol{n} - \ell \le 0 \ \text{ on } \ \Gamma_c\Big\},$$

where

$$\boldsymbol{V} := \Big\{\boldsymbol{v} \in \boldsymbol{H}^1(\Omega) : \boldsymbol{v} = \boldsymbol{0} \ \text{ on } \ \Gamma_d\Big\}.$$

The weak formulation of the frictionless unilateral problem (1.1)–(1.5) is defined by

$$\begin{cases} \text{find } \boldsymbol{u} \in \boldsymbol{K} \text{ such that} \\ A(\boldsymbol{u}, \boldsymbol{v}-\boldsymbol{u}) \geq F(\boldsymbol{v}-\boldsymbol{u}) \quad \forall\, \boldsymbol{v} \in \boldsymbol{K}, \end{cases} \tag{2.1}$$

where $A(\cdot,\cdot)$ and $F(\cdot)$ are bilinear and linear forms respectively, defined for any $\boldsymbol{v}$, $\boldsymbol{w} \in \boldsymbol{H}^1(\Omega)$ by

$$A(\boldsymbol{v},\boldsymbol{w}) := \int_\Omega \boldsymbol{\sigma}(\boldsymbol{v}) : \varepsilon(\boldsymbol{w})\,\mathrm{d}\boldsymbol{x},$$

$$F(\boldsymbol{v}) := \int_\Omega \boldsymbol{f}\cdot\boldsymbol{v}\,\mathrm{d}x + \int_{\Gamma_n} \boldsymbol{g}\cdot\boldsymbol{v}\,\mathrm{d}\boldsymbol{s}.$$

The bilinear form $A(\cdot,\cdot)$ is continuous and elliptic from $\boldsymbol{V}\times\boldsymbol{V}$ into $\mathbb{R}$, that is, for $\boldsymbol{v}$, $\boldsymbol{w} \in \boldsymbol{V}$,

$$m\|\boldsymbol{v}\|^2_{1,\Omega} \leq A(\boldsymbol{v},\boldsymbol{v}), \tag{2.2}$$

$$A(\boldsymbol{v},\boldsymbol{w}) \leq M\|\boldsymbol{v}\|_{1,\Omega}\|\boldsymbol{w}\|_{1,\Omega}, \tag{2.3}$$

where M and m are positive constants. The mathematical analysis of problem (2.1) such as existence, uniqueness and regularity of its solution can be found in Kikuchi & Oden (1988). It admits a unique solution according to Stampacchia's theorem.

In order to relax the constraint condition $\boldsymbol{u}\cdot\boldsymbol{n} - \ell \leq 0$ characterizing the set $\boldsymbol{K}$, we are interested in the following penalty weak formulation of the variational inequality problem (2.1):

$$\begin{cases} \text{find } \boldsymbol{u}_\varepsilon \in \boldsymbol{V} \text{ such that} \\ A(\boldsymbol{u}_\varepsilon, \boldsymbol{v}) + \frac{1}{\varepsilon}\int_{\Gamma_c} \left[\boldsymbol{u}_\varepsilon\cdot\boldsymbol{n} - \ell\right]^+ \left(\boldsymbol{v}\cdot\boldsymbol{n}\right)\mathrm{d}\boldsymbol{s} = F(\boldsymbol{v}) \quad \forall\, \boldsymbol{v} \in \boldsymbol{V}, \end{cases} \tag{2.4}$$

where the notation $[\,\cdot\,]^+$, for a scalar quantity $a \in \mathbb{R}$ stands for

$$[a]^+ := \begin{cases} a & \text{if } a > 0, \\ 0 & \text{otherwise.} \end{cases}$$

Let us recall the following two useful properties:

$$a \leq [a]^+, \quad a\,[a]^+ = [a]^+\,[a]^+, \quad \forall\, a \in \mathbb{R}, \tag{2.5}$$

from which one can deduce the monotonicity property (Chouly & Hild, 2013, proof of Theorem 2.2)

$$\left([a]^+ - [b]^+\right)\left(a-b\right) \geq \left([a]^+ - [b]^+\right)^2. \tag{2.6}$$

Well-posedness and uniform convergence of the penalized problem (2.4) have already been studied by Kikuchi & Song (1981, Theorems 3.1 and 3.2) and by Kikuchi & Oden (1988, Theorems 3.15 and 6.6). Uniform convergence of the penalized solution is studied in Chernov *et al.* (2007, Theorem 6). Recently, in the context of a polygonal boundary domain, Chouly & Hild (2013, Theorem 3.1) tackle these questions again. Since we have to deal with a curved domain, we establish in the following theorem the uniform convergence of the penalized solution by relying on Chouly & Hild (2013, Theorem 3.1).

THEOREM 2.1 Let $\Omega \subset \mathbb{R}^d, d = 2, 3$, be a bounded domain with a curved and smooth boundary ($\partial\Omega \in \mathscr{C}^{1,1}$), $\boldsymbol{u}$ and $\boldsymbol{u}_\varepsilon$ the solutions of problems (2.1) and (2.4), respectively. If $\boldsymbol{u}$ belongs to $\boldsymbol{H}^{\frac{3}{2}+\nu}(\Omega)$, where $\nu \in \left(0, \frac{1}{2}\right]$, then we obtain the *a priori* estimates

$$\|\boldsymbol{u} - \boldsymbol{u}_\varepsilon\|_{1,\Omega} + \varepsilon^{\frac{1}{2}} \left\|\boldsymbol{\sigma}_n(\boldsymbol{u}) + \tfrac{1}{\varepsilon}\left[\boldsymbol{u}_\varepsilon \cdot \boldsymbol{n} - \ell\right]^+\right\|_{0,\Gamma_c} \leq c\varepsilon^{\frac{1}{2}+\nu}\|\boldsymbol{u}\|_{\frac{3}{2}+\nu,\Omega}, \tag{2.7}$$

$$\left\|\boldsymbol{\sigma}_n(\boldsymbol{u}) + \tfrac{1}{\varepsilon}\left[\boldsymbol{u}_\varepsilon \cdot \boldsymbol{n} - \ell\right]^+\right\|_{-\nu,\Gamma_c} \leq c\varepsilon^{2\nu}\|\boldsymbol{u}\|_{\frac{3}{2}+\nu,\Omega}, \tag{2.8}$$

where $c > 0$ is a constant, independent of the penalty parameter ε and the solution $\boldsymbol{u}$.

Proof. Multiplying equation (1.1) by test functions $\boldsymbol{v} \in \boldsymbol{H}^1(\Omega)$ and integrating over Ω by taking into account Green's formula and boundary conditions (1.3), (1.4) and (1.5), we obtain

$$A(\boldsymbol{u}, \boldsymbol{v}) - \int_{\Gamma_c} \boldsymbol{\sigma}_n(\boldsymbol{u})(\boldsymbol{v} \cdot \boldsymbol{n})\, \mathrm{d}s = F(\boldsymbol{v}). \tag{2.9}$$

If $\boldsymbol{u} \in \boldsymbol{V}$, equation (2.9) may have no meaning because of lack of regularity. Then we suppose that $\boldsymbol{u} \in \boldsymbol{H}^{\frac{3}{2}+\nu}(\Omega)$, where $\nu > 0$ to justify this calculation by obtaining $\boldsymbol{\sigma}_n(\boldsymbol{u}) \in H^\nu(\Gamma_c)$.

Using the coercivity relation (2.2) of $A(\cdot,\cdot)$, equation (2.9) and problem (2.4), we obtain

$$\begin{aligned}
m\|\boldsymbol{u} - \boldsymbol{u}_\varepsilon\|_{1,\Omega}^2 &\leq A(\boldsymbol{u} - \boldsymbol{u}_\varepsilon, \boldsymbol{u} - \boldsymbol{u}_\varepsilon) \\
&= A(\boldsymbol{u}, \boldsymbol{u} - \boldsymbol{u}_\varepsilon) - A(\boldsymbol{u}_\varepsilon, \boldsymbol{u} - \boldsymbol{u}_\varepsilon) \\
&= \int_{\Gamma_c} \left(\boldsymbol{\sigma}_n(\boldsymbol{u}) + \frac{1}{\varepsilon}\,[\boldsymbol{u}_\varepsilon \cdot \boldsymbol{n} - \ell]^+\right)(\boldsymbol{u} \cdot \boldsymbol{n} - \boldsymbol{u}_\varepsilon \cdot \boldsymbol{n})\, \mathrm{d}s \\
&= \int_{\Gamma_c} \left(\boldsymbol{\sigma}_n(\boldsymbol{u}) + \frac{1}{\varepsilon}\,[\boldsymbol{u}_\varepsilon \cdot \boldsymbol{n} - \ell]^+\right)((\boldsymbol{u} \cdot \boldsymbol{n} - \ell) - (\boldsymbol{u}_\varepsilon \cdot \boldsymbol{n} - \ell))\, \mathrm{d}s \\
&= \int_{\Gamma_c} \boldsymbol{\sigma}_n(\boldsymbol{u})\,(\boldsymbol{u} \cdot \boldsymbol{n} - \ell)\, \mathrm{d}s + \int_{\Gamma_c} \frac{1}{\varepsilon}\,[\boldsymbol{u}_\varepsilon \cdot \boldsymbol{n} - \ell]^+\,(\boldsymbol{u} \cdot \boldsymbol{n} - \ell)\, \mathrm{d}s \\
&\quad - \int_{\Gamma_c} \boldsymbol{\sigma}_n(\boldsymbol{u})\,(\boldsymbol{u}_\varepsilon \cdot \boldsymbol{n} - \ell)\, \mathrm{d}s - \int_{\Gamma_c} \frac{1}{\varepsilon}\,[\boldsymbol{u}_\varepsilon \cdot \boldsymbol{n} - \ell]^+\,(\boldsymbol{u}_\varepsilon \cdot \boldsymbol{n} - \ell)\, \mathrm{d}s.
\end{aligned} \tag{2.10}$$

We first remark, thanks to the unilateral contact condition (1.2), the following two estimates:

$$\begin{aligned}\int_{\Gamma_c} \boldsymbol{\sigma}_n(\boldsymbol{u})(\boldsymbol{u}\cdot\boldsymbol{n}-\ell)\,\mathrm{d}\boldsymbol{s} &= 0,\\ \int_{\Gamma_c} \frac{1}{\varepsilon}\,[\boldsymbol{u}_\varepsilon\cdot\boldsymbol{n}-\ell]^+\,(\boldsymbol{u}\cdot\boldsymbol{n}-\ell)\,\mathrm{d}\boldsymbol{s} &\le 0.\end{aligned} \tag{2.11}$$

Using property (2.5) and the same condition (1.2) we obtain

$$\begin{aligned}-\int_{\Gamma_c} \boldsymbol{\sigma}_n(\boldsymbol{u})\big(\boldsymbol{u}_\varepsilon\cdot\boldsymbol{n}-\ell\big)\,\mathrm{d}\boldsymbol{s} &\le -\int_{\Gamma_c} \boldsymbol{\sigma}_n(\boldsymbol{u})\,[\boldsymbol{u}_\varepsilon\cdot\boldsymbol{n}-\ell]^+\,\mathrm{d}\boldsymbol{s},\\ -\int_{\Gamma_c} \frac{1}{\varepsilon}\big[\boldsymbol{u}_\varepsilon\cdot\boldsymbol{n}-\ell\big]^+\big(\boldsymbol{u}_\varepsilon\cdot\boldsymbol{n}-\ell\big)\,\mathrm{d}\boldsymbol{s} &= -\int_{\Gamma_c} \frac{1}{\varepsilon}\big[\boldsymbol{u}_\varepsilon\cdot\boldsymbol{n}-\ell\big]^+\big[\boldsymbol{u}_\varepsilon\cdot\boldsymbol{n}-\ell\big]^+\,\mathrm{d}\boldsymbol{s}.\end{aligned} \tag{2.12}$$

Then taking into account relations (2.11), (2.12) and the Young inequality, estimate (2.10) becomes

$$\begin{aligned}m\|\boldsymbol{u}-\boldsymbol{u}_\varepsilon\|^2_{1,\Omega} &\le -\int_{\Gamma_c}\left(\boldsymbol{\sigma}_n(\boldsymbol{u})+\frac{1}{\varepsilon}\,[\boldsymbol{u}_\varepsilon\cdot\boldsymbol{n}-\ell]^+\right)[\boldsymbol{u}_\varepsilon\cdot\boldsymbol{n}-\ell]^+\,\mathrm{d}\boldsymbol{s}\\ &\le -\varepsilon\int_{\Gamma_c}\left(\boldsymbol{\sigma}_n(\boldsymbol{u})+\frac{1}{\varepsilon}\,[\boldsymbol{u}_\varepsilon\cdot\boldsymbol{n}-\ell]^+\right)\left(\boldsymbol{\sigma}_n(\boldsymbol{u})-\boldsymbol{\sigma}_n(\boldsymbol{u})+\frac{1}{\varepsilon}\,[\boldsymbol{u}_\varepsilon\cdot\boldsymbol{n}-\ell]^+\right)\mathrm{d}\boldsymbol{s}\\ &\le -\varepsilon\left\|\boldsymbol{\sigma}_n(\boldsymbol{u})+\frac{1}{\varepsilon}\,[\boldsymbol{u}_\varepsilon\cdot\boldsymbol{n}-\ell]^+\right\|^2_{0,\Gamma_c}+\varepsilon\int_{\Gamma_c}\left(\boldsymbol{\sigma}_n(\boldsymbol{u})+\frac{1}{\varepsilon}\,[\boldsymbol{u}_\varepsilon\cdot\boldsymbol{n}-\ell]^+\right)\boldsymbol{\sigma}_n(\boldsymbol{u})\,\mathrm{d}\boldsymbol{s}\\ &\le -\varepsilon\left\|\boldsymbol{\sigma}_n(\boldsymbol{u})+\frac{1}{\varepsilon}\,[\boldsymbol{u}_\varepsilon\cdot\boldsymbol{n}-\ell]^+\right\|^2_{0,\Gamma_c}+\varepsilon^{\delta}\left\|\boldsymbol{\sigma}_n(\boldsymbol{u})+\frac{1}{\varepsilon}\,[\boldsymbol{u}_\varepsilon\cdot\boldsymbol{n}-\ell]^+\right\|_{-\nu,\Gamma_c}\varepsilon^{1-\delta}\,\|\boldsymbol{\sigma}_n(\boldsymbol{u})\|_{\nu,\Gamma_c}\\ &\le -\varepsilon\left\|\boldsymbol{\sigma}_n(\boldsymbol{u})+\frac{1}{\varepsilon}\,[\boldsymbol{u}_\varepsilon\cdot\boldsymbol{n}-\ell]^+\right\|^2_{0,\Gamma_c}+\frac{\varepsilon^{2\delta}}{2\beta}\left\|\boldsymbol{\sigma}_n(\boldsymbol{u})+\frac{1}{\varepsilon}\,[\boldsymbol{u}_\varepsilon\cdot\boldsymbol{n}-\ell]^+\right\|^2_{-\nu,\Gamma_c}\\ &\quad+\frac{\beta\varepsilon^{2-2\delta}}{2}\,\|\boldsymbol{\sigma}_n(\boldsymbol{u})\|^2_{\nu,\Gamma_c},\end{aligned} \tag{2.13}$$

with $\delta\in[0,1]$ and $\beta>0$. Using then estimate (A.3) in Lemma A1, we take again estimate (2.13) and obtain

$$\begin{aligned}m\|\boldsymbol{u}-\boldsymbol{u}_\varepsilon\|^2_{1,\Omega} &\le -\varepsilon\left(1-c\frac{\varepsilon^{2(\delta+\nu)-1}}{\beta}\right)\left\|\boldsymbol{\sigma}_n(\boldsymbol{u})+\frac{1}{\varepsilon}\big[\boldsymbol{u}_\varepsilon\cdot\boldsymbol{n}-\ell\big]^+\right\|^2_{0,\Gamma_c}+c\frac{\varepsilon^{2(\delta+\nu)-1}}{\beta}\|\boldsymbol{u}-\boldsymbol{u}_\varepsilon\|^2_{1,\Omega}\\ &\quad+\frac{\beta\varepsilon^{2-2\delta}}{2}\,\|\boldsymbol{\sigma}_n(\boldsymbol{u})\|^2_{\nu,\Gamma_c}.\end{aligned} \tag{2.14}$$

Hence, making the same choice as in (Chouly & Hild, 2013, proof of Theorem 3.1) of the parameters $\delta = \frac{1}{2} - \nu$, $\beta = 2c \max\left(1, m^{-1}\right)$ and taking into account the estimate $\|\boldsymbol{\sigma}_n(\boldsymbol{u})\|_{\nu,\Gamma_c} \leq c\|\boldsymbol{u}\|_{\frac{3}{2}+\nu,\Omega}$, we obtain from (2.14) the following estimate which established (2.7):

$$\left(m - \frac{1}{2\max\left(1, m^{-1}\right)}\right)\|\boldsymbol{u} - \boldsymbol{u}_\varepsilon\|_{1,\Omega}^2 + \varepsilon\left(1 - \frac{1}{2\max\left(1, m^{-1}\right)}\right)\left\|\boldsymbol{\sigma}_n(\boldsymbol{u}) + \frac{1}{\varepsilon}\left[\boldsymbol{u}_\varepsilon \cdot \boldsymbol{n} - \ell\right]^+\right\|_{0,\Gamma_c}^2$$
$$\leq c \max\left(1, m^{-1}\right) \varepsilon^{1+2\nu} \|\boldsymbol{u}\|_{\frac{3}{2}+\nu,\Omega}^2.$$

To establish result (2.8), we take again estimate (A.3) in Lemma A1 and we use result (2.7). □

Uniform convergence of the penalized solution $\boldsymbol{u}_\varepsilon$ having thus been acquired from Theorem 2.1, one can exclusively focus on the finite element approximations of problem (2.4) in order to establish an *a priori* estimate of its solution. This is the approach followed by Chernov *et al.* (2007), which constitutes a fundamental difference from most works on the penalty method of the unilateral contact problem (namely Kikuchi & Song, 1981, Kikuchi & Oden, 1988 and Chouly & Hild, 2013).

3. Finite element approximations of the penalty formulation

We follow the process described in Bänsch & Deckelnick (1999) for the construction of finite element approximations. Let us denote by $\widetilde{\mathcal{T}}_h$ a finite set of straight and closed d-simplices which mesh a domain $\widetilde{\Omega}_h$,

$$\overline{\widetilde{\Omega}}_h := \bigcup_{\widetilde{\mathrm{T}} \in \widetilde{\mathcal{T}}_h} \widetilde{\mathrm{T}} \tag{3.1}$$

in such a way that all vertices on its boundary $\partial\widetilde{\Omega}_h$ also lie on $\partial\Omega$ (see Fig. 2). We denote by $h_{\widetilde{\mathrm{T}}}$ and $\rho_{\widetilde{\mathrm{T}}}$ the diameter and the radius of the largest ball inscribed in $\widetilde{\mathrm{T}}$, respectively, and define the mesh size h by

$$h := \max_{\widetilde{\mathrm{T}} \in \widetilde{\mathcal{T}}_h} h_{\widetilde{\mathrm{T}}}. \tag{3.2}$$

Moreover, let us assume the usual *shape regularity* condition

$$\sup_h \max_{\widetilde{\mathrm{T}} \in \widetilde{\mathcal{T}}_h} \frac{h_{\widetilde{\mathrm{T}}}}{\rho_{\widetilde{\mathrm{T}}}} \leq \kappa < \infty.$$

For every element $\widetilde{\mathrm{T}} \in \widetilde{\mathcal{T}}_h$ there exists an invertible affine function $\widetilde{\mathbb{F}}_{\widetilde{\mathrm{T}}}$, mapping the standard reference triangle $\widehat{\mathrm{T}}$ onto $\widetilde{\mathrm{T}}$, which is locally defined by

$$\widetilde{\mathbb{F}}_{\widetilde{\mathrm{T}}} : \begin{array}{l} \mathbb{R}^d \longrightarrow \mathbb{R}^d, \\ \widehat{\boldsymbol{x}} \longmapsto \mathbf{A}_{\widetilde{\mathrm{T}}}\widehat{\boldsymbol{x}} + \boldsymbol{b}_{\widetilde{\mathrm{T}}}, \end{array}$$

where $\mathbf{A}_{\widetilde{\mathrm{T}}}$ is an endomorphism in $\mathbb{R}^d$ and the vector $\boldsymbol{b}_{\widetilde{\mathrm{T}}} \in \mathbb{R}^d$.

Besides the straight triangulation $\widetilde{\mathcal{T}}_h$ defining the computational domain $\widetilde{\Omega}_h$, which we will use to define the discrete weak formulation of problem (2.4), we introduce the exact triangulation $\mathcal{T}_h$ of the smooth domain Ω (see Bernardi, 1989; Lenoir, 1986),

$$\overline{\Omega} := \bigcup_{\mathsf{T}\in\mathcal{T}_h} \mathsf{T}, \tag{3.3}$$

where for every $\widetilde{\mathsf{T}} \in \widetilde{\mathcal{T}}_h$, there is a mapping $\phi_{\widetilde{\mathsf{T}}} \in \mathscr{C}^3(\widehat{\mathsf{T}}; \mathbb{R}^d)$ such that the function

$$\mathbb{F}_{\widetilde{\mathsf{T}}} := \widetilde{\mathbb{F}}_{\widetilde{\mathsf{T}}} + \phi_{\widetilde{\mathsf{T}}}$$

maps $\widehat{\mathsf{T}}$ onto a curved d-simplex $\mathsf{T} \subset \overline{\Omega}$.

While the set $\widetilde{\mathcal{T}}_h$ is made of straight triangles (or tetrahedrons in the three-dimensional domain) building the polygonal (or polyhedral) approximation $\widetilde{\Omega}_h$ of the domain Ω, the exact triangulation $\mathcal{T}_h$ of Ω is made of curved d-simplices along its boundary $\partial\Omega$ and of d-simplices with straight edges in its interior (see Fig. 2). Hence, the mapping $\mathbb{G}_h$ (see Fig. 3), locally defined by

$$\mathbb{G}_{h/\widetilde{\mathsf{T}}} := \mathbb{F}_{\widetilde{\mathsf{T}}\circ}\widetilde{\mathbb{F}}_{\widetilde{\mathsf{T}}}^{-1} \equiv I + \phi_{\widetilde{\mathsf{T}}\circ}\widetilde{\mathbb{F}}_{\widetilde{\mathsf{T}}}^{-1},$$

is a homeomorphism from $\widetilde{\Omega}_h$ to Ω, where by construction the function $\phi_{\widetilde{\mathsf{T}}}$ vanishes in any triangle $\widetilde{\mathsf{T}}$ having at most one vertex on $\partial\widetilde{\Omega}_h$ (which implies that $\mathbb{G}_h \equiv I$ on all triangles which are disjoint from $\partial\widetilde{\Omega}_h$). Moreover, we have the following estimates (see Lenoir, 1986, Propositions 2 and 3):

$$\sup_{\boldsymbol{x}\in\widetilde{\mathsf{T}}} \left| \left|\det\left(D\mathbb{G}_{h/\widetilde{\mathsf{T}}}\right)(\boldsymbol{x})\right| - 1 \right| \le ch, \quad \sup_{\boldsymbol{x}\in\mathsf{T}} \left| \left|\det\left(D\mathbb{G}_{h/\widetilde{\mathsf{T}}}^{-1}\right)(\boldsymbol{x})\right| - 1 \right| \le ch, \tag{3.4}$$

$$c_1\left|\det\left(\mathbf{A}_{\widetilde{\mathsf{T}}}\right)\right| \le \left|\det\left(D\mathbb{F}_{\widetilde{\mathsf{T}}}\left(\widehat{\boldsymbol{x}}\right)\right)\right| \le c_2\left|\det\left(\mathbf{A}_{\widetilde{\mathsf{T}}}\right)\right|, \quad \widehat{\boldsymbol{x}} \in \widehat{\mathsf{T}}, \tag{3.5}$$

where $|\det(D\mathbb{G}_{h/\widetilde{\mathsf{T}}})(\boldsymbol{x})|$ stands for the absolute value of the Jacobian determinant of $\mathbb{G}_h$ at point $\boldsymbol{x}$. As a consequence, we notably have $\boldsymbol{v} \in \mathbf{H}^m(\Omega)$ if and only if $\boldsymbol{v}_\circ\mathbb{G}_h \in \mathbf{H}^m(\widetilde{\Omega}_h)$, $m = 0, 1$, and there exist $c_1, c_2 > 0$ such that the following estimates hold:

$$c_1 \|\boldsymbol{v}\|_{m,\Omega} \le \|\boldsymbol{v}_\circ\mathbb{G}_h\|_{m,\widetilde{\Omega}_h} \le c_2 \|\boldsymbol{v}\|_{m,\Omega}. \tag{3.6}$$

Let us denote by $\widetilde{\Gamma}_c^h$, $\widetilde{\Gamma}_d^h$ and $\widetilde{\Gamma}_n^h$ the polygonal boundary parts of $\partial\widetilde{\Omega}_h$ that correspond to the curved and smooth boundary parts Γ_c, Γ_d and Γ_n of $\partial\Omega$, respectively. Each of these parts is obtained by joining vertices of $\partial\widetilde{\Omega}_h$ that lie on Γ_c, Γ_d and Γ_n, respectively.

We define the following finite-dimensional space of standard continuous and piecewise affine functions

$$V_h := \left\{ v_h \in \mathscr{C}^0(\overline{\widetilde{\Omega}}_h) : v_{h|\widetilde{\mathsf{T}}} \in \mathscr{P}_1(\widetilde{\mathsf{T}}) \quad \forall\, \widetilde{\mathsf{T}} \in \widetilde{\mathcal{T}}_h,\ v_h = 0 \text{ on } \widetilde{\Gamma}_d^h \right\},$$

where $\mathscr{P}_1(\widetilde{\mathsf{T}})$ is the space of polynomials on $\widetilde{\mathsf{T}}$ of order less than or equal to 1. We thus consider the finite-dimensional vector space $\boldsymbol{V}_h := (V_h)^d$ and also introduce the space of traces on $\widetilde{\Gamma}_c^h$ for functions in V_h as

$$W_h(\widetilde{\Gamma}_c^h) := \left\{ \mu_h \in \mathscr{C}^0(\overline{\widetilde{\Gamma}_c^h});\ \exists\, v_h \in V_h,\ v_{h|\widetilde{\Gamma}_c^h} = \mu_h \right\}.$$

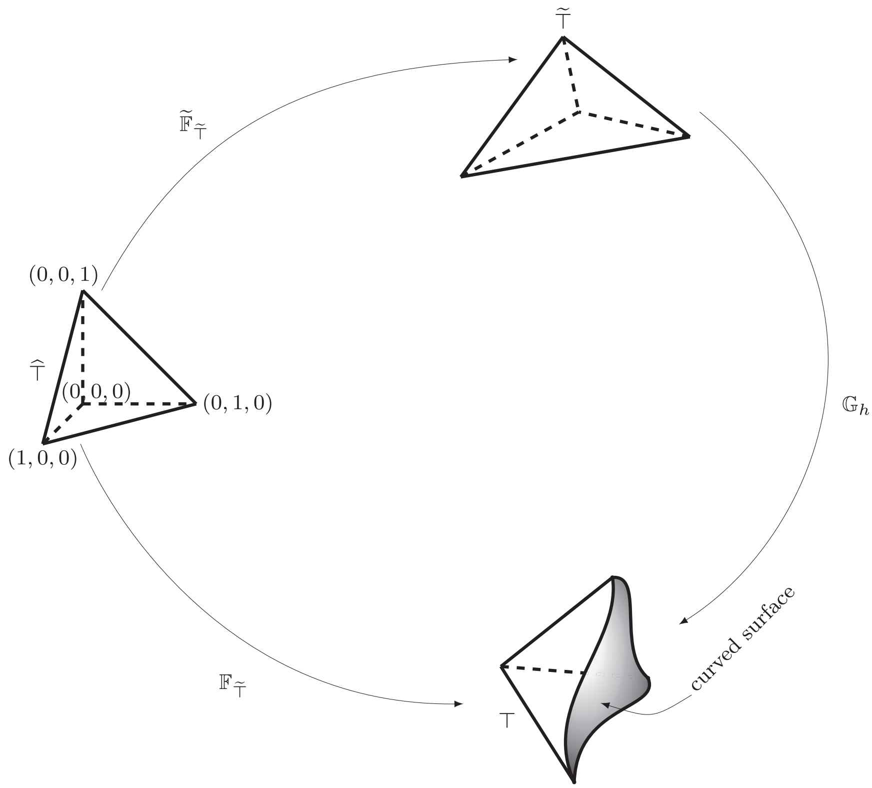

FIG. 3. The standard d-simplex $\widehat{\mathbb{T}}$ and the tetrahedrons $\widetilde{\mathbb{T}}$, $\mathbb{T}$ at their respective boundaries (see Bänsch & Deckelnick, 1999, Fig. 1).

Supposing the end points (in a two-dimensional space) or the border (in a three-dimensional space) of $\widetilde{\Gamma}_c^h$ belong to $\widetilde{\Gamma}_n^h$, we assume that the mesh on $\widetilde{\Gamma}_c^h$ induced by $\widetilde{\mathcal{T}}_h$ is quasi-uniform. This implies, in the sense of Bramble *et al.* (2001), local quasi-uniformity of the mesh on $\widetilde{\Gamma}_c^h$.

Let us recall the stability and interpolation properties of the $L^2(\widetilde{\Gamma}_c^h)$-projection operator onto $W_h(\widetilde{\Gamma}_c^h)$, denoted by $\mathcal{P}^h : L^2(\widetilde{\Gamma}_c^h) \longrightarrow W_h(\widetilde{\Gamma}_c^h)$. These properties, proofs of which can be found in Bernardi *et al.* (1994) and Bramble *et al.* (2001), are stated in the following lemma.

LEMMA 3.1 Suppose that the mesh associated to $W_h(\widetilde{\Gamma}_c^h)$ is locally quasi-uniform. For all $s \in [0, 1]$ and all $v \in H^s(\widetilde{\Gamma}_c^h)$, we have the stability estimate

$$\left\| \mathcal{P}^h(v) \right\|_{s,\widetilde{\Gamma}_c^h} \leq c \|v\|_{s,\widetilde{\Gamma}_c^h}. \tag{3.7}$$

The following interpolation estimate also holds:

$$\left\| v - \mathcal{P}^h(v) \right\|_{0,\widetilde{\Gamma}_c^h} \leq c h^s \|v\|_{s,\widetilde{\Gamma}_c^h} \tag{3.8}$$

for all $v \in H^s(\widetilde{\Gamma}_c^h)$, where the constant $c > 0$ in both cases is independent of v and the mesh size h.

We recall another lemma, proven in Bjorstad & Widlund (1986) (see also Dominguez & Sayas, 2003), which concerns the existence of a discrete bounded lifting from the contact boundary $\widetilde{\Gamma}_c^h$ to the domain $\widetilde{\Omega}_h$.

LEMMA 3.2 Suppose that the mesh on the contact boundary $\widetilde{\Gamma}_c^h$ is quasi-uniform. There exist an extension operator $\mathcal{R}^h : W_h(\widetilde{\Gamma}_c^h) \longrightarrow V_h$ and $c > 0$, such that

$$\mathcal{R}^h(\mu_h)_{|\widetilde{\Gamma}_c^h} = \mu_h, \quad \left\|\mathcal{R}^h(\mu_h)\right\|_{1,\widetilde{\Omega}_h} \le c\|\mu_h\|_{\frac{1}{2},\widetilde{\Gamma}_c^h} \quad \forall\, \mu_h \in W_h(\widetilde{\Gamma}_c^h). \tag{3.9}$$

REMARK 3.3 Let us denote by $\boldsymbol{\mathcal{P}}^h$ and $\boldsymbol{\mathcal{R}}^h$ the vector forms of the operators $\mathcal{P}^h$ and $\mathcal{R}^h$, respectively, which are defined for any vector $\boldsymbol{w} := (w_i)_{1\le i\le d}$ as

$$\boldsymbol{\mathcal{P}}^h(\boldsymbol{w}) := \left(\mathcal{P}^h(w_i)\right)_{1\le i\le d} \quad \text{and} \quad \boldsymbol{\mathcal{R}}^h(\boldsymbol{w}) := \left(\mathcal{R}^h(w_i)\right)_{1\le i\le d}. \tag{3.10}$$

It is an easy task to see that the operators $\boldsymbol{\mathcal{P}}^h$ and $\boldsymbol{\mathcal{R}}^h$ also satisfy the stability and interpolation properties (3.7), (3.8) and (3.9).

We thus propose the following penalty finite element approximation of the penalized problem (2.4):

$$\begin{cases} \text{given } \varepsilon > 0, \text{ find } \boldsymbol{u}_{\varepsilon h} \in \boldsymbol{V}_h \text{ such that} \\ A_h(\boldsymbol{u}_{\varepsilon h}, \boldsymbol{v}_h) + \frac{1}{\varepsilon}\int_{\widetilde{\Gamma}_c^h} \left|\det(D\mathbb{G}_h)\right| \left[\boldsymbol{u}_{\varepsilon h}\cdot \boldsymbol{n}_\circ\mathbb{G}_h - \ell_\circ\mathbb{G}_h\right]^+ \left(\boldsymbol{v}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h\right) \mathrm{d}\boldsymbol{s} = F_h(\boldsymbol{v}_h) \quad \forall\, \boldsymbol{v}_h \in \boldsymbol{V}_h, \end{cases} \tag{3.11}$$

where the bilinear and linear functionals $A_h(\cdot,\cdot)$ and $F_h(\cdot)$ are defined, for $\boldsymbol{v}_h, \boldsymbol{w}_h \in \boldsymbol{H}^1(\widetilde{\Omega}_h)$, by

$$A_h(\boldsymbol{v}_h, \boldsymbol{w}_h) := \int_{\widetilde{\Omega}_h} \boldsymbol{\sigma}(\boldsymbol{v}_h) : \boldsymbol{\varepsilon}(\boldsymbol{w}_h)\,\mathrm{d}\boldsymbol{x},$$

$$F_h(\boldsymbol{v}_h) := \int_{\widetilde{\Omega}_h} \boldsymbol{f}_\circ\mathbb{G}_h\cdot\boldsymbol{v}_h\,\mathrm{d}\boldsymbol{x} + \int_{\widetilde{\Gamma}_n^h} \boldsymbol{g}_\circ\mathbb{G}_h\cdot\boldsymbol{v}_h\,\mathrm{d}\boldsymbol{s}.$$

In the forthcoming analysis, in order to remedy the problem that the functions $\boldsymbol{u}_\varepsilon$ and $\boldsymbol{u}_{\varepsilon h}$ are respectively defined in different domains Ω and $\widetilde{\Omega}_h$, we assign to each $\boldsymbol{v}_h \in \boldsymbol{V}_h$ the function $\overline{\boldsymbol{v}}_h$ defined by

$$\overline{\boldsymbol{v}}_h := \boldsymbol{v}_{h\circ}\mathbb{G}_h^{-1}, \tag{3.12}$$

which belongs to $\boldsymbol{V}$ (since it vanishes on Γ_d). Otherwise, using the the transformation rule to migrate from Ω to $\widetilde{\Omega}_h$, problem (2.4) can be rewritten for any $\boldsymbol{v} \in \boldsymbol{V}$ as

$$\begin{aligned}\int_{\widetilde{\Omega}_h} \left|\det(D\mathbb{G}_h)\right| \boldsymbol{\sigma}(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h) : \boldsymbol{\varepsilon}(\boldsymbol{v}_\circ\mathbb{G}_h)\,\mathrm{d}\boldsymbol{x} + \frac{1}{\varepsilon}\int_{\widetilde{\Gamma}_c^h} \left|\det(D\mathbb{G}_h)\right| \left[\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h - \ell_\circ\mathbb{G}_h\right]^+ (\boldsymbol{v}_\circ\mathbb{G}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h)\mathrm{d}\boldsymbol{s} \\ = \int_{\widetilde{\Omega}_h} \left|\det(D\mathbb{G}_h)\right| \boldsymbol{f}_\circ\mathbb{G}_h\cdot\boldsymbol{v}_\circ\mathbb{G}_h\,\mathrm{d}\boldsymbol{x} + \int_{\widetilde{\Gamma}_n^h} \left|\det(D\mathbb{G}_h)\right| \boldsymbol{g}_\circ\mathbb{G}_h\cdot\boldsymbol{v}_\circ\mathbb{G}_h\,\mathrm{d}\boldsymbol{s}. \end{aligned} \tag{3.13}$$

We establish well-posedness of the penalty finite element problem (3.11) using an argument proposed by Brezis (1968) for M-type pseudo-monotone operators (see also Lions, 1969 and Kikuchi & Song, 1981).

THEOREM 3.4 For all $\varepsilon > 0$ and $0 < h \le h_0$ with h_0 small enough, problem (3.11) admits a unique solution $\boldsymbol{u}_{\varepsilon h}$ in $\boldsymbol{V}_h$.

Proof. Let us recall the following assumption established by Lenoir (1986, Subsection 6.1):

$$(\mathbb{A})\qquad \begin{gathered}\text{There exists a continuous function } C(h)\text{, which vanishes with } h \text{ and satisfies}\\ \left|A(\overline{\boldsymbol{v}}_h, \overline{\boldsymbol{w}}_h) - A_h(\boldsymbol{v}_h, \boldsymbol{w}_h)\right| \le C(h)\|\overline{\boldsymbol{v}}_h\|_{1,\Omega}\|\overline{\boldsymbol{w}}_h\|_{1,\Omega} \quad \forall\, \boldsymbol{v}_h,\, \boldsymbol{w}_h \in \boldsymbol{V}_h.\end{gathered} \tag{3.14}$$

Combining assumption $(\mathbb{A})$ (whose proof is inferred from Lenoir, 1986, Lemma 8) and the coercivity of the bilinear form $A(\cdot,\cdot)$, we obtain the coercivity of the discrete bilinear form $A_h(\cdot,\cdot)$ for sufficiently small h. In fact, from assumption $(\mathbb{A})$ and the coercivity of the bilinear form $A(\cdot,\cdot)$ we obtain

$$(m - C(h))\|\overline{\boldsymbol{v}}_h\|_{1,\Omega}\|\overline{\boldsymbol{w}}_h\|_{1,\Omega} \le A(\overline{\boldsymbol{v}}_h, \overline{\boldsymbol{w}}_h) - C(h)\|\overline{\boldsymbol{v}}_h\|_{1,\Omega}\|\overline{\boldsymbol{w}}_h\|_{1,\Omega} \le A_h(\boldsymbol{v}_h, \boldsymbol{w}_h).$$

Since $C(h)$ vanishes with h, then there exists a small enough $h_0 > 0$ such that $C(h) < m$ for all $h \le h_0$. Moreover, taking into account relation (3.6), we obtain the coercivity of the bilinear form $A_h(\cdot,\cdot)$ as

$$\beta\|\boldsymbol{v}_h\|_{1,\widetilde{\Omega}_h}\|\boldsymbol{w}_h\|_{1,\widetilde{\Omega}_h} \le (m - C(h))\, \|\overline{\boldsymbol{v}}_h\|_{1,\Omega}\|\overline{\boldsymbol{w}}_h\|_{1,\Omega} \le A_h(\boldsymbol{v}_h, \boldsymbol{w}_h)\ \forall\, \boldsymbol{v}_h, \boldsymbol{w}_h \in \boldsymbol{V}_h. \tag{3.15}$$

With the help of the coercivity property (3.15), we establish well-posedness using the same arguments as those employed in Chouly & Hild (2013, proof of Theorem 2.2). Let us define, thanks to the Riesz representation theorem, the nonlinear operator $B_h : \boldsymbol{V}_h \longrightarrow \boldsymbol{V}_h$,

$$\begin{aligned}(B_h(\boldsymbol{v}_h),\, \boldsymbol{w}_h)_{1,\widetilde{\Omega}_h} &:= A_h(\boldsymbol{v}_h, \boldsymbol{w}_h)\\ &\quad + \frac{1}{\varepsilon}\int_{\widetilde{\Gamma}_c^h} \left|\det(D\mathbb{G}_h)\right| \left[\boldsymbol{v}_h \cdot \boldsymbol{n}_\circ\mathbb{G}_h - \ell_\circ\mathbb{G}_h\right]^+ \left(\boldsymbol{w}_h \cdot \boldsymbol{n}_\circ\mathbb{G}_h\right) \mathrm{d}s \quad \forall\, \boldsymbol{v}_h,\, \boldsymbol{w}_h \in \boldsymbol{V}_h,\end{aligned}$$

where $(\cdot\,,\,\cdot)_{1,\widetilde{\Omega}_h}$ denotes the inner product in $\boldsymbol{H}^1(\widetilde{\Omega}_h)$. To prove the well-posedness of problem (3.11), we establish that the operator B_h is one-to-one by proving its coercivity and hemicontinuity.

Using property (2.6) by taking into account the coercivity of the bilinear form $A_h(\cdot,\cdot)$ in (3.15), we obtain

$$\begin{aligned}
(B_h(\boldsymbol{v}_h) - B_h(\boldsymbol{w}_h),\, \boldsymbol{v}_h - \boldsymbol{w}_h)_{1,\widetilde{\Omega}_h} &= A_h(\boldsymbol{v}_h - \boldsymbol{w}_h, \boldsymbol{v}_h - \boldsymbol{w}_h) \\
&+ \frac{1}{\varepsilon}\int_{\widetilde{\Gamma}_c^h} \left|\det\left(D\mathbb{G}_h\right)\right| \left([\boldsymbol{v}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h - \ell_\circ\mathbb{G}_h]^+ - [\boldsymbol{w}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h - \ell_\circ\mathbb{G}_h]^+\right)((\boldsymbol{v}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h - \ell_\circ\mathbb{G}_h) \\
&- (\boldsymbol{w}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h - \ell_\circ\mathbb{G}_h))\,\mathrm{d}\boldsymbol{s} \\
&\geq \beta\|\boldsymbol{v}_h - \boldsymbol{w}_h\|^2_{1,\widetilde{\Omega}_h} + \frac{1}{\varepsilon}\int_{\widetilde{\Gamma}_c^h} \left|\det\left(D\mathbb{G}_h\right)\right| \left([\boldsymbol{v}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h - \ell_\circ\mathbb{G}_h]^+ - [\boldsymbol{w}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h - \ell_\circ\mathbb{G}_h]^+\right)^2\,\mathrm{d}\boldsymbol{s} \\
&\geq \beta\|\boldsymbol{v}_h - \boldsymbol{w}_h\|^2_{1,\widetilde{\Omega}_h}.
\end{aligned} \tag{3.16}$$

Otherwise, to prove hemicontinuity of the operator B_h, we introduce the real-valued function φ_h,

$$[0,\,1] \ni t \longrightarrow \varphi_h(t) := \left(B_h(\boldsymbol{v}_h - t\boldsymbol{w}_h),\, \boldsymbol{w}_h\right)_{1,\widetilde{\Omega}_h}, \forall\, \boldsymbol{v}_h \quad \boldsymbol{w}_h \in \boldsymbol{V}_h,$$

and prove that the function φ_h is Lipschitz. Indeed, for any $s, t \in [0,\,1]$, we obtain

$$\begin{aligned}
|\varphi_h(t) - \varphi_h(s)| &= \left|(B_h(\boldsymbol{v}_h - t\boldsymbol{w}_h) - B_h(\boldsymbol{v}_h - s\boldsymbol{w}_h),\, \boldsymbol{w}_h)_{1,\widetilde{\Omega}_h}\right| \\
&\leq |A_h(\boldsymbol{v}_h - t\boldsymbol{w}_h, \boldsymbol{w}_h) - A_h(\boldsymbol{v}_h - s\boldsymbol{w}_h, \boldsymbol{w}_h)| \\
&\quad + \frac{1}{\varepsilon}\left|\int_{\widetilde{\Gamma}_c^h} |\det\left(D\mathbb{G}_h\right)|([\boldsymbol{v}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h - t\boldsymbol{w}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h - \ell_\circ\mathbb{G}_h]^+\right. \\
&\quad \left. - [\boldsymbol{v}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h - s\boldsymbol{w}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h - \ell_\circ\mathbb{G}_h]^+)\,(\boldsymbol{w}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h)\,\mathrm{d}\boldsymbol{s}\right| \\
&\leq |s-t|\,A_h(\boldsymbol{w}_h, \boldsymbol{w}_h) + \frac{1}{\varepsilon}\int_{\widetilde{\Gamma}_c^h} |\det\left(D\mathbb{G}_h\right)|\,|\,[\boldsymbol{v}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h - t\boldsymbol{w}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h - \ell_\circ\mathbb{G}_h]^+ \\
&\quad - [\boldsymbol{v}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h - s\boldsymbol{w}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h - \ell_\circ\mathbb{G}_h]^+\,||\boldsymbol{w}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h|\,\mathrm{d}\boldsymbol{s}.
\end{aligned}$$

Moreover, with the help of the inequality $\left|[a]^+ - [b]^+\right| \leq |a-b|$ (which may be obtained from inequality (2.6) by applying the Young inequality), we obtain

$$\begin{aligned}
|\varphi_h(t) - \varphi_h(s)| &\leq |s-t|\,A_h(\boldsymbol{w}_h, \boldsymbol{w}_h) + \frac{1}{\varepsilon}\int_{\widetilde{\Gamma}_c^h} |s-t|\,|\det\left(D\mathbb{G}_h\right)|\,(\boldsymbol{w}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h)^2\,\mathrm{d}\boldsymbol{s} \\
&= |s-t|\left(A_h(\boldsymbol{w}_h, \boldsymbol{w}_h) + \frac{1}{\varepsilon}\left\||\det\left(D\mathbb{G}_h\right)|^{\frac{1}{2}}\,(\boldsymbol{w}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h)\right\|^2_{0,\widetilde{\Gamma}_c^h}\right),
\end{aligned}$$

establishing that the function φ_h is Lipschitz.

Hence, the function φ_h is continuous and thus the operator B_h is hemicontinuous. From Brezis (1968, Corollary 15), inequality (3.16) and hemicontinuity suffice to conclude that the operator B_h is one-to-one. □

4. *A priori* estimate in terms of the penalty parameter ε and the mesh size h

In this section, we exclusively focus on the convergence of the penalty finite element approximation solution $(\overline{\boldsymbol{u}}_{\varepsilon h}, \frac{1}{\varepsilon}[\overline{\boldsymbol{u}}_{\varepsilon h}\cdot\boldsymbol{n}-\ell]^+)$ towards $(\boldsymbol{u}_\varepsilon, \frac{1}{\varepsilon}[\boldsymbol{u}_\varepsilon\cdot\boldsymbol{n}-\ell]^+)$ but not towards the solution $(\boldsymbol{u}, \boldsymbol{\sigma}_n(\boldsymbol{u}))$ of the constrained problem (2.1). In fact, we establish the *a priori* estimate (4.1) in terms of the interpolation errors, the penalty approximation errors and the errors induced by polygonal or polyhedral approximations of the curved and smooth boundary domain Ω. This approach is the fundamental difference between this study and most of the works on the penalty method of the unilateral contact problem (see for instance Kikuchi & Song, 1981, Kikuchi & Oden, 1988 and Chouly & Hild, 2013).

THEOREM 4.1 Let $\Omega\subset\mathbb{R}^d, d=2,3$ be a bounded domain with a curved and smooth boundary ($\partial\Omega\in\mathscr{C}^{1,1}$), $\boldsymbol{u}_\varepsilon$ and $\boldsymbol{u}_{\varepsilon h}$ the solutions of problems (2.4) and (3.11), respectively. Let us assume that $\boldsymbol{f}\in\boldsymbol{L}^2(\Omega)$ and $\boldsymbol{g}\in\boldsymbol{L}^2(\Gamma_n)$; then for $\varepsilon>0$ and $0<h\le h_0$ with h_0 small enough, we obtain the following *a priori* estimate:

$$\begin{aligned}
&\|\boldsymbol{u}_\varepsilon-\overline{\boldsymbol{u}}_{\varepsilon h}\|_{1,\Omega}+\left(\varepsilon^{\frac12}-c_o h^{\frac12}\right)\left\|\frac{1}{\varepsilon}[\boldsymbol{u}_\varepsilon\cdot\boldsymbol{n}-\ell]^+-\frac{1}{\varepsilon}[\overline{\boldsymbol{u}}_{\varepsilon h}\cdot\boldsymbol{n}-\ell]^+\right\|_{0,\Gamma_c}\\
&\quad\le c\Bigg(\|\boldsymbol{u}-\overline{\boldsymbol{v}}_h\|_{1,\Omega}+\|\boldsymbol{u}-\boldsymbol{u}_\varepsilon\|_{1,\Omega}+\big|A_h(\boldsymbol{u}_{\varepsilon h},\boldsymbol{v}_h-\boldsymbol{u}_{\varepsilon h})\\
&\qquad-A(\overline{\boldsymbol{u}}_{\varepsilon h},\overline{\boldsymbol{v}}_h-\overline{\boldsymbol{u}}_{\varepsilon h})\big|^{\frac12}+\big|F_h(\boldsymbol{v}_h-\boldsymbol{u}_{\varepsilon h})-F(\overline{\boldsymbol{v}}_h-\overline{\boldsymbol{u}}_{\varepsilon h})\big|^{\frac12}\\
&\qquad+\sup_{\boldsymbol{x}\in\overline{\widetilde{\Omega}}_h}\Big|\big|\det(D\mathbb{G}_h)\big|(\boldsymbol{x})-1\Big|\left(\|\boldsymbol{u}_\varepsilon\|_{1,\Omega}+\|\boldsymbol{f}\|_{0,\Omega}+\|\boldsymbol{g}\|_{0,\Gamma_n}\right)\Bigg)
\end{aligned}\tag{4.1}$$

for all $\boldsymbol{v}_h\in\boldsymbol{V}_h$, where c, c_o are positive constants independent of the solution $\boldsymbol{u}$, the mesh size h and the penalty parameter ε.

Proof. For all $\varepsilon>0$ and $0<h\le h_0$ with h_0 small enough, problem (3.11) has a unique solution $\boldsymbol{u}_{\varepsilon h}$ (see Theorem 3.4). Thus, taking test functions $\boldsymbol{v}_h-\boldsymbol{u}_{\varepsilon h}\in\boldsymbol{V}_h$ with $\boldsymbol{v}_h\in\boldsymbol{V}_h$, we have

$$\begin{aligned}
&A_h(\boldsymbol{u}_{\varepsilon h},\boldsymbol{v}_h-\boldsymbol{u}_{\varepsilon h})+\frac{1}{\varepsilon}\int_{\widetilde{\Gamma}_c^h}\big|\det(D\mathbb{G}_h)\big|\big[\boldsymbol{u}_{\varepsilon h}\cdot\boldsymbol{n}\circ\mathbb{G}_h-\ell\circ\mathbb{G}_h\big]^+\big(\boldsymbol{v}_h\cdot\boldsymbol{n}\circ\mathbb{G}_h-\boldsymbol{u}_{\varepsilon h}\cdot\boldsymbol{n}\circ\mathbb{G}_h\big)\,\mathrm{d}\boldsymbol{s}\\
&\qquad=F_h(\boldsymbol{v}_h-\boldsymbol{u}_{\varepsilon h}).
\end{aligned}\tag{4.2}$$

In problem (2.4), taking test functions $\overline{\boldsymbol{v}}_h-\overline{\boldsymbol{u}}_{\varepsilon h}\in\boldsymbol{V}$ with $\boldsymbol{v}_h\in\boldsymbol{V}_h$, we obtain

$$A(\boldsymbol{u}_\varepsilon,\overline{\boldsymbol{v}}_h-\overline{\boldsymbol{u}}_{\varepsilon h})+\frac{1}{\varepsilon}\int_{\Gamma_c}\big[\boldsymbol{u}_\varepsilon\cdot\boldsymbol{n}-\ell\big]^+\big(\overline{\boldsymbol{v}}_h\cdot\boldsymbol{n}-\overline{\boldsymbol{u}}_{\varepsilon h}\cdot\boldsymbol{n}\big)\,\mathrm{d}\boldsymbol{s}=F(\overline{\boldsymbol{v}}_h-\overline{\boldsymbol{u}}_{\varepsilon h}).\tag{4.3}$$

Due to the ellipticity and the continuity of the bilinear form $A(\cdot,\cdot)$ (described in (2.2) and (2.3)) and taking into account equality (4.3), we obtain for all $\boldsymbol{v}_h\in\boldsymbol{V}_h$,

$$\begin{aligned}
m\|\boldsymbol{u}_\varepsilon-\overline{\boldsymbol{u}}_{\varepsilon h}\|_{1,\Omega}^2&\le A(\boldsymbol{u}_\varepsilon-\overline{\boldsymbol{u}}_{\varepsilon h},\boldsymbol{u}_\varepsilon-\overline{\boldsymbol{u}}_{\varepsilon h})\\
&\le A(\boldsymbol{u}_\varepsilon-\overline{\boldsymbol{u}}_{\varepsilon h},\boldsymbol{u}_\varepsilon-\overline{\boldsymbol{v}}_h)+A(\boldsymbol{u}_\varepsilon-\overline{\boldsymbol{u}}_{\varepsilon h},\overline{\boldsymbol{v}}_h-\overline{\boldsymbol{u}}_{\varepsilon h})
\end{aligned}$$

$$
\begin{aligned}
&\leq A(\boldsymbol{u}_\varepsilon - \overline{\boldsymbol{u}}_{\varepsilon h}, \boldsymbol{u}_\varepsilon - \overline{\boldsymbol{v}}_h) + A(\boldsymbol{u}_\varepsilon, \overline{\boldsymbol{v}}_h - \overline{\boldsymbol{u}}_{\varepsilon h}) - A(\overline{\boldsymbol{u}}_{\varepsilon h}, \overline{\boldsymbol{v}}_h - \overline{\boldsymbol{u}}_{\varepsilon h})\\
&\leq M\|\boldsymbol{u}_\varepsilon - \overline{\boldsymbol{u}}_{\varepsilon h}\|_{1,\Omega}\|\boldsymbol{u}_\varepsilon - \overline{\boldsymbol{v}}_h\|_{1,\Omega} + F(\overline{\boldsymbol{v}}_h - \overline{\boldsymbol{u}}_{\varepsilon h}) - A(\overline{\boldsymbol{u}}_{\varepsilon h}, \overline{\boldsymbol{v}}_h - \overline{\boldsymbol{u}}_{\varepsilon h})\\
&\quad - \frac{1}{\varepsilon}\int_{\Gamma_c}\left[\boldsymbol{u}_\varepsilon \cdot \boldsymbol{n} - \ell\right]^+\left(\overline{\boldsymbol{v}}_h \cdot \boldsymbol{n} - \overline{\boldsymbol{u}}_{\varepsilon h} \cdot \boldsymbol{n}\right)\mathrm{d}\boldsymbol{s}.
\end{aligned}
\tag{4.4}
$$

Adding the terms in (4.2) into the right-hand side of inequality (4.4), using the Young inequality and the Cauchy–Schwarz inequality, we obtain the estimate

$$
\begin{aligned}
m\|\boldsymbol{u}_\varepsilon - \overline{\boldsymbol{u}}_{\varepsilon h}\|^2_{1,\Omega} \leq\ & \frac{1}{2\alpha}\|\boldsymbol{u}_\varepsilon - \overline{\boldsymbol{u}}_{\varepsilon h}\|^2_{1,\Omega} + \frac{\alpha M^2}{2}\|\boldsymbol{u}_\varepsilon - \overline{\boldsymbol{v}}_h\|^2_{1,\Omega} + \{A_h(\boldsymbol{u}_{\varepsilon h}, \boldsymbol{v}_h - \boldsymbol{u}_{\varepsilon h}) - A(\overline{\boldsymbol{u}}_{\varepsilon h}, \overline{\boldsymbol{v}}_h - \overline{\boldsymbol{u}}_{\varepsilon h})\}\\
&+ \{F(\overline{\boldsymbol{v}}_h - \overline{\boldsymbol{u}}_{\varepsilon h}) - F_h(\boldsymbol{v}_h - \boldsymbol{u}_{\varepsilon h})\} - \frac{1}{\varepsilon}\int_{\Gamma_c}[\boldsymbol{u}_\varepsilon \cdot \boldsymbol{n} - \ell]^+\,(\overline{\boldsymbol{v}}_h \cdot \boldsymbol{n} - \overline{\boldsymbol{u}}_{\varepsilon h} \cdot \boldsymbol{n})\ \mathrm{d}\boldsymbol{s}\\
&+ \frac{1}{\varepsilon}\int_{\widetilde{\Gamma}^h_c}\left|\det(D\mathbb{G}_h)\right|[\boldsymbol{u}_{\varepsilon h} \cdot \boldsymbol{n}_\circ\mathbb{G}_h - \ell_\circ\mathbb{G}_h]^+\,(\boldsymbol{v}_h \cdot \boldsymbol{n}_\circ\mathbb{G}_h - \boldsymbol{u}_{\varepsilon h} \cdot \boldsymbol{n}_\circ\mathbb{G}_h)\ \mathrm{d}\boldsymbol{s}\\
\leq\ & \frac{1}{2\alpha}\|\boldsymbol{u}_\varepsilon - \overline{\boldsymbol{u}}_{\varepsilon h}\|^2_{1,\Omega} + \alpha M^2\left(\|\boldsymbol{u}_\varepsilon - \boldsymbol{u}\|^2_{1,\Omega} + \|\boldsymbol{u} - \overline{\boldsymbol{v}}_h\|^2_{1,\Omega}\right)\\
&+ \{A_h(\boldsymbol{u}_{\varepsilon h}, \boldsymbol{v}_h - \boldsymbol{u}_{\varepsilon h}) - A(\overline{\boldsymbol{u}}_{\varepsilon h}, \overline{\boldsymbol{v}}_h - \overline{\boldsymbol{u}}_{\varepsilon h})\} + \{F(\overline{\boldsymbol{v}}_h - \overline{\boldsymbol{u}}_{\varepsilon h}) - F_h(\boldsymbol{v}_h - \boldsymbol{u}_{\varepsilon h})\}\\
&+ \frac{1}{\varepsilon}\int_{\widetilde{\Gamma}^h_c}\left|\det(D\mathbb{G}_h)\right|[\boldsymbol{u}_{\varepsilon h} \cdot \boldsymbol{n}_\circ\mathbb{G}_h - \ell_\circ\mathbb{G}_h]^+\,(\boldsymbol{v}_h \cdot \boldsymbol{n}_\circ\mathbb{G}_h - \boldsymbol{u}_{\varepsilon h} \cdot \boldsymbol{n}_\circ\mathbb{G}_h)\ \mathrm{d}\boldsymbol{s}\\
&- \frac{1}{\varepsilon}\int_{\Gamma_c}[\boldsymbol{u}_\varepsilon \cdot \boldsymbol{n} - \ell]^+\,(\overline{\boldsymbol{v}}_h \cdot \boldsymbol{n} - \overline{\boldsymbol{u}}_{\varepsilon h} \cdot \boldsymbol{n})\ \mathrm{d}\boldsymbol{s}.
\end{aligned}
\tag{4.5}
$$

In order to estimate the last two terms in the right-hand side of inequality (4.5), since the two integrals are defined in different domains, we use the transformation rule to migrate from Γ_c to $\widetilde{\Gamma}^h_c$ before introducing the term $\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h \cdot \boldsymbol{n}_\circ\mathbb{G}_h$ defined on $\widetilde{\Gamma}^h_c$:

$$
\begin{aligned}
&\frac{1}{\varepsilon}\int_{\widetilde{\Gamma}^h_c}\left|\det(D\mathbb{G}_h)\right|[\boldsymbol{u}_{\varepsilon h}\cdot\boldsymbol{n}_\circ\mathbb{G}_h - \ell_\circ\mathbb{G}_h]^+\,(\boldsymbol{v}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h - \boldsymbol{u}_{\varepsilon h}\cdot\boldsymbol{n}_\circ\mathbb{G}_h)\mathrm{d}\boldsymbol{s} - \frac{1}{\varepsilon}\int_{\Gamma_c}[\boldsymbol{u}_\varepsilon\cdot\boldsymbol{n} - \ell]^+\,(\overline{\boldsymbol{v}}_h\cdot\boldsymbol{n} - \overline{\boldsymbol{u}}_{\varepsilon h}\cdot\boldsymbol{n})\mathrm{d}\boldsymbol{s}\\
&= \int_{\widetilde{\Gamma}^h_c}\left|\det(D\mathbb{G}_h)\right|\left(\frac{1}{\varepsilon}[\boldsymbol{u}_{\varepsilon h}\cdot\boldsymbol{n}_\circ\mathbb{G}_h - \ell_\circ\mathbb{G}_h]^+ - \frac{1}{\varepsilon}[\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h - \ell_\circ\mathbb{G}_h]^+\right)(\boldsymbol{v}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h - \boldsymbol{u}_{\varepsilon h}\cdot\boldsymbol{n}_\circ\mathbb{G}_h)\mathrm{d}\boldsymbol{s}\\
&= \int_{\widetilde{\Gamma}^h_c}\left|\det(D\mathbb{G}_h)\right|\left(\frac{1}{\varepsilon}[\boldsymbol{u}_{\varepsilon h}\cdot\boldsymbol{n}_\circ\mathbb{G}_h - \ell_\circ\mathbb{G}_h]^+ - \frac{1}{\varepsilon}[\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h - \ell_\circ\mathbb{G}_h]^+\right)(\boldsymbol{v}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h - \boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h)\mathrm{d}\boldsymbol{s}\\
&\quad + \int_{\widetilde{\Gamma}^h_c}\left|\det(D\mathbb{G}_h)\right|\left(\frac{1}{\varepsilon}[\boldsymbol{u}_{\varepsilon h}\cdot\boldsymbol{n}_\circ\mathbb{G}_h - \ell_\circ\mathbb{G}_h]^+ - \frac{1}{\varepsilon}[\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h - \ell_\circ\mathbb{G}_h]^+\right)(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h - \boldsymbol{u}_{\varepsilon h}\cdot\boldsymbol{n}_\circ\mathbb{G}_h)\mathrm{d}\boldsymbol{s}.
\end{aligned}
\tag{4.6}
$$

Using the monotonicity property (2.6) by taking into account the identity $\left|\det\left(D\mathbb{G}_h^{-1}\right)\right|\left|\det\left(D\mathbb{G}_h\right)\right|=1$, the last term in the right-hand side of equality (4.6) is estimated as

$$\begin{aligned}
&\int_{\widetilde{\Gamma}_c^h}\left|\det\left(D\mathbb{G}_h\right)\right|\left(\frac{1}{\varepsilon}\left[\boldsymbol{u}_{\varepsilon h}\cdot\boldsymbol{n}_\circ\mathbb{G}_h-\ell_\circ\mathbb{G}_h\right]^+-\frac{1}{\varepsilon}\left[\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h-\ell_\circ\mathbb{G}_h\right]^+\right)\left(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h-\boldsymbol{u}_{\varepsilon h}\cdot\boldsymbol{n}_\circ\mathbb{G}_h\right)\,\mathrm{d}s\\
&=-\varepsilon\int_{\Gamma_c}\left|\det\left(D\mathbb{G}_h\right)\right|\left|\det\left(D\mathbb{G}_h^{-1}\right)\right|\left(\frac{1}{\varepsilon}\left[\boldsymbol{u}_\varepsilon\cdot\boldsymbol{n}-\ell\right]^+-\frac{1}{\varepsilon}\left[\overline{\boldsymbol{u}}_{\varepsilon h}\cdot\boldsymbol{n}-\ell\right]^+\right)\\
&\quad\times\left(\frac{1}{\varepsilon}\left(\boldsymbol{u}_\varepsilon\cdot\boldsymbol{n}-\ell\right)-\frac{1}{\varepsilon}\left(\overline{\boldsymbol{u}}_{\varepsilon h}\cdot\boldsymbol{n}-\ell\right)\right)\mathrm{d}s\\
&=-\varepsilon\int_{\Gamma_c}\left(\frac{1}{\varepsilon}\left[\boldsymbol{u}_\varepsilon\cdot\boldsymbol{n}-\ell\right]^+-\frac{1}{\varepsilon}\left[\overline{\boldsymbol{u}}_{\varepsilon h}\cdot\boldsymbol{n}-\ell\right]^+\right)\left(\frac{1}{\varepsilon}\left(\boldsymbol{u}_\varepsilon\cdot\boldsymbol{n}-\ell\right)-\frac{1}{\varepsilon}\left(\overline{\boldsymbol{u}}_{\varepsilon h}\cdot\boldsymbol{n}-\ell\right)\right)\mathrm{d}s\\
&\leq-\varepsilon\left\|\frac{1}{\varepsilon}\left[\boldsymbol{u}_\varepsilon\cdot\boldsymbol{n}-\ell\right]^+-\frac{1}{\varepsilon}\left[\overline{\boldsymbol{u}}_{\varepsilon h}\cdot\boldsymbol{n}-\ell\right]^+\right\|_{0,\Gamma_c}^2. \qquad (4.7)
\end{aligned}$$

For the first term in the right-hand side of equality (4.6), we introduce the term $\mathcal{P}^h\left(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h-\boldsymbol{v}_h\right)\cdot\boldsymbol{n}_\circ\mathbb{G}_h$ before using the Cauchy–Schwarz inequality, relation (3.10) and the equality in (3.9) to obtain the estimate

$$\begin{aligned}
&\int_{\widetilde{\Gamma}_c^h}\left|\det\left(D\mathbb{G}_h\right)\right|\left(\frac{1}{\varepsilon}\left[\boldsymbol{u}_{\varepsilon h}\cdot\boldsymbol{n}_\circ\mathbb{G}_h-\ell_\circ\mathbb{G}_h\right]^+-\frac{1}{\varepsilon}\left[\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h-\ell_\circ\mathbb{G}_h\right]^+\right)\left(\boldsymbol{v}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h-\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h\right)\mathrm{d}s\\
&=\int_{\widetilde{\Gamma}_c^h}\left|\det\left(D\mathbb{G}_h\right)\right|\left(\frac{1}{\varepsilon}\left[\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h-\ell_\circ\mathbb{G}_h\right]^+-\frac{1}{\varepsilon}\left[\boldsymbol{u}_{\varepsilon h}\cdot\boldsymbol{n}_\circ\mathbb{G}_h-\ell_\circ\mathbb{G}_h\right]^+\right)\\
&\quad\times\left(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h-\boldsymbol{v}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h\right)\mathrm{d}s\\
&=\int_{\widetilde{\Gamma}_c^h}\left|\det\left(D\mathbb{G}_h\right)\right|\left(\frac{1}{\varepsilon}\left[\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h-\ell_\circ\mathbb{G}_h\right]^+-\frac{1}{\varepsilon}\left[\boldsymbol{u}_{\varepsilon h}\cdot\boldsymbol{n}_\circ\mathbb{G}_h-\ell_\circ\mathbb{G}_h\right]^+\right)\\
&\quad\times\left(\left(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h-\boldsymbol{v}_h\right)\cdot\boldsymbol{n}_\circ\mathbb{G}_h-\mathcal{P}^h\left(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h-\boldsymbol{v}_h\right)\cdot\boldsymbol{n}_\circ\mathbb{G}_h\right)\mathrm{d}s+\int_{\widetilde{\Gamma}_c^h}\left|\det\left(D\mathbb{G}_h\right)\right|\\
&\quad\times\left(\frac{1}{\varepsilon}\left[\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h-\ell_\circ\mathbb{G}_h\right]^+-\frac{1}{\varepsilon}\left[\boldsymbol{u}_{\varepsilon h}\cdot\boldsymbol{n}_\circ\mathbb{G}_h-\ell_\circ\mathbb{G}_h\right]^+\right)\mathcal{P}^h\left(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h-\boldsymbol{v}_h\right)\cdot\boldsymbol{n}_\circ\mathbb{G}_h\,\mathrm{d}s\\
&\leq c\left\|\frac{1}{\varepsilon}\left[\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h-\ell_\circ\mathbb{G}_h\right]^+-\frac{1}{\varepsilon}\left[\boldsymbol{u}_{\varepsilon h}\cdot\boldsymbol{n}_\circ\mathbb{G}_h-\ell_\circ\mathbb{G}_h\right]^+\right\|_{0,\widetilde{\Gamma}_c^h}\\
&\quad\times\left\|\left(\left(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h-\boldsymbol{v}_h\right)-\mathcal{P}^h\left(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h-\boldsymbol{v}_h\right)\right)\cdot\boldsymbol{n}_\circ\mathbb{G}_h\right\|_{0,\widetilde{\Gamma}_c^h}+\int_{\widetilde{\Gamma}_c^h}\left|\det\left(D\mathbb{G}_h\right)\right|\\
&\quad\times\left(\frac{1}{\varepsilon}\left[\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h-\ell_\circ\mathbb{G}_h\right]^+-\frac{1}{\varepsilon}\left[\boldsymbol{u}_{\varepsilon h}\cdot\boldsymbol{n}_\circ\mathbb{G}_h-\ell_\circ\mathbb{G}_h\right]^+\right)\mathcal{R}^h\left(\mathcal{P}^h\left(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h-\boldsymbol{v}_h\right)\right)\cdot\boldsymbol{n}_\circ\mathbb{G}_h\,\mathrm{d}s,
\end{aligned}$$
$$(4.8)$$

where we have used the estimates in (3.5).

Using the interpolation estimate (3.8) (by taking $s = \frac{1}{2}$), the Young inequality, the continuity of the trace operator and estimate (3.6), we bound the first term in the right-hand side of inequality (4.8) as

$$
\begin{aligned}
&\left\| \frac{1}{\varepsilon}\big[\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h\cdot\boldsymbol{n}_{\circ}\mathbb{G}_h-\ell_{\circ}\mathbb{G}_h\big]^{+}-\frac{1}{\varepsilon}\big[\boldsymbol{u}_{\varepsilon h}\cdot\boldsymbol{n}_{\circ}\mathbb{G}_h-\ell_{\circ}\mathbb{G}_h\big]^{+}\right\|_{0,\widetilde{\Gamma}_c^h}\left\|\Big(\big(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h-\boldsymbol{v}_h\big)-\boldsymbol{\mathcal{P}}^h\big(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h-\boldsymbol{v}_h\big)\Big)\cdot\boldsymbol{n}_{\circ}\mathbb{G}_h\right\|_{0,\widetilde{\Gamma}_c^h}\\
&\quad\le ch^{\frac{1}{2}}\left\| \frac{1}{\varepsilon}\big[\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h\cdot\boldsymbol{n}_{\circ}\mathbb{G}_h-\ell_{\circ}\mathbb{G}_h\big]^{+}-\frac{1}{\varepsilon}\big[\boldsymbol{u}_{\varepsilon h}\cdot\boldsymbol{n}_{\circ}\mathbb{G}_h-\ell_{\circ}\mathbb{G}_h\big]^{+}\right\|_{0,\widetilde{\Gamma}_c^h}\|\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h-\boldsymbol{v}_h\|_{\frac{1}{2},\widetilde{\Gamma}_c^h}\\
&\quad\le \frac{ch}{2}\left\| \frac{1}{\varepsilon}\big[\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h\cdot\boldsymbol{n}_{\circ}\mathbb{G}_h-\ell_{\circ}\mathbb{G}_h\big]^{+}-\frac{1}{\varepsilon}\big[\boldsymbol{u}_{\varepsilon h}\cdot\boldsymbol{n}_{\circ}\mathbb{G}_h-\ell_{\circ}\mathbb{G}_h\big]^{+}\right\|^2_{0,\widetilde{\Gamma}_c^h}+\frac{c}{2}\|\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h-\boldsymbol{v}_h\|^2_{\frac{1}{2},\widetilde{\Gamma}_c^h}\\
&\quad\le ch\left\| \frac{1}{\varepsilon}\big[\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h\cdot\boldsymbol{n}_{\circ}\mathbb{G}_h-\ell_{\circ}\mathbb{G}_h\big]^{+}-\frac{1}{\varepsilon}\big[\boldsymbol{u}_{\varepsilon h}\cdot\boldsymbol{n}_{\circ}\mathbb{G}_h-\ell_{\circ}\mathbb{G}_h\big]^{+}\right\|^2_{0,\widetilde{\Gamma}_c^h}+c\,\|\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h-\boldsymbol{v}_h\|^2_{1,\widetilde{\Omega}_h}\\
&\quad\le ch\left\| \frac{1}{\varepsilon}\big[\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h\cdot\boldsymbol{n}_{\circ}\mathbb{G}_h-\ell_{\circ}\mathbb{G}_h\big]^{+}-\frac{1}{\varepsilon}\big[\boldsymbol{u}_{\varepsilon h}\cdot\boldsymbol{n}_{\circ}\mathbb{G}_h-\ell_{\circ}\mathbb{G}_h\big]^{+}\right\|^2_{0,\widetilde{\Gamma}_c^h}\\
&\qquad+c\Big(\|\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h-\boldsymbol{u}_{\circ}\mathbb{G}_h\|^2_{1,\widetilde{\Omega}_h}+\|\boldsymbol{u}_{\circ}\mathbb{G}_h-\boldsymbol{v}_h\|^2_{1,\widetilde{\Omega}_h}\Big)\\
&\quad\le ch\left\| \frac{1}{\varepsilon}\big[\boldsymbol{u}_{\varepsilon}\cdot\boldsymbol{n}-\ell\big]^{+}-\frac{1}{\varepsilon}\big[\overline{\boldsymbol{u}}_{\varepsilon h}\cdot\boldsymbol{n}-\ell\big]^{+}\right\|^2_{0,\Gamma_c}+c\Big(\|\boldsymbol{u}_{\varepsilon}-\boldsymbol{u}\|^2_{1,\Omega}+\|\boldsymbol{u}-\overline{\boldsymbol{v}}_h\|^2_{1,\Omega}\Big).
\end{aligned}
\tag{4.9}
$$

Otherwise, using equation (3.13) and problem (3.11), we estimate the last term in the right-hand side of inequality (4.8) as

$$
\begin{aligned}
&\int_{\widetilde{\Gamma}_c^h}\big|\det(D\mathbb{G}_h)\big|\left(\frac{1}{\varepsilon}\big[\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h\cdot\boldsymbol{n}_{\circ}\mathbb{G}_h-\ell_{\circ}\mathbb{G}_h\big]^{+}-\frac{1}{\varepsilon}\big[\boldsymbol{u}_{\varepsilon h}\cdot\boldsymbol{n}_{\circ}\mathbb{G}_h-\ell_{\circ}\mathbb{G}_h\big]^{+}\right)\mathcal{R}^h\left(\mathcal{P}^h\left(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h-\boldsymbol{v}_h\right)\right)\cdot\boldsymbol{n}_{\circ}\mathbb{G}_h\,\mathrm{d}\boldsymbol{s}\\
&\qquad= A_h\left(\boldsymbol{u}_{\varepsilon h},\boldsymbol{\mathcal{R}}^h\left(\boldsymbol{\mathcal{P}}^h\left(\boldsymbol{u}_{\varepsilon}-\boldsymbol{v}_h\right)\right)\right)-F_h\left(\boldsymbol{\mathcal{R}}^h\left(\boldsymbol{\mathcal{P}}^h(\boldsymbol{u}_{\varepsilon}-\boldsymbol{v}_h)\right)\right)\\
&\qquad\quad-\int_{\widetilde{\Omega}_h}\big|\det(D\mathbb{G}_h)\big|\boldsymbol{\sigma}(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h):\boldsymbol{\varepsilon}\left(\boldsymbol{\mathcal{R}}^h\left(\boldsymbol{\mathcal{P}}^h(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h-\boldsymbol{v}_h)\right)\right)\mathrm{d}\boldsymbol{x}\\
&\qquad\quad+\int_{\widetilde{\Omega}_h}\big|\det(D\mathbb{G}_h)\big|\boldsymbol{f}_{\circ}\mathbb{G}_h\cdot\boldsymbol{\mathcal{R}}^h\left(\boldsymbol{\mathcal{P}}^h(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h-\boldsymbol{v}_h)\right)\mathrm{d}\boldsymbol{x}\\
&\qquad\quad+\int_{\widetilde{\Gamma}_n^h}\big|\det(D\mathbb{G}_h)\big|\boldsymbol{g}_{\circ}\mathbb{G}_h\cdot\boldsymbol{\mathcal{R}}^h\left(\boldsymbol{\mathcal{P}}^h(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h-\boldsymbol{v}_h)\right)\mathrm{d}\boldsymbol{s}\\
&\qquad= A_h\left(\boldsymbol{u}_{\varepsilon h},\boldsymbol{\mathcal{R}}^h\left(\boldsymbol{\mathcal{P}}^h\left(\boldsymbol{u}_{\varepsilon}-\boldsymbol{v}_h\right)\right)\right)-F_h\left(\boldsymbol{\mathcal{R}}^h\left(\boldsymbol{\mathcal{P}}^h(\boldsymbol{u}_{\varepsilon}-\boldsymbol{v}_h)\right)\right)\\
&\qquad\quad+\int_{\widetilde{\Omega}_h}\Big(1-\big|\det(D\mathbb{G}_h)\big|\Big)\boldsymbol{\sigma}(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h):\boldsymbol{\varepsilon}\left(\boldsymbol{\mathcal{R}}^h\left(\boldsymbol{\mathcal{P}}^h(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h-\boldsymbol{v}_h)\right)\right)\mathrm{d}\boldsymbol{x}
\end{aligned}
$$

$$
\begin{aligned}
&+ \int_{\widetilde{\Omega}_h} \Big(\big|\det(D\mathbb{G}_h)\big| - 1\Big) \boldsymbol{f}_\circ\mathbb{G}_h \cdot \boldsymbol{\mathcal{R}}^h\left(\boldsymbol{\mathcal{P}}^h(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h - \boldsymbol{v}_h)\right)\,\mathrm{d}\boldsymbol{x} \\
&+ \int_{\widetilde{\Gamma}_n^h} \Big(\big|\det(D\mathbb{G}_h)\big| - 1\Big) \boldsymbol{g}_\circ\mathbb{G}_h \cdot \boldsymbol{\mathcal{R}}^h\left(\boldsymbol{\mathcal{P}}^h(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h - \boldsymbol{v}_h)\right)\,\mathrm{d}\boldsymbol{s} \\
&+ F_h\left(\boldsymbol{\mathcal{R}}^h\left(\boldsymbol{\mathcal{P}}^h(\boldsymbol{u}_\varepsilon - \boldsymbol{v}_h)\right)\right) - A_h\left(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h, \boldsymbol{\mathcal{R}}^h\left(\boldsymbol{\mathcal{P}}^h(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h - \boldsymbol{v}_h)\right)\right) \\
\leq\; & A_h\left(\boldsymbol{u}_{\varepsilon h} - \boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h, \boldsymbol{\mathcal{R}}^h\left(\boldsymbol{\mathcal{P}}^h(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h - \boldsymbol{v}_h)\right)\right) + \sup_{\boldsymbol{x}\in\widetilde{\Omega}_h}\Big|1 - \big|\det(D\mathbb{G}_h(\boldsymbol{x}))\big|\Big| \\
&\times \int_{\widetilde{\Omega}_h} \left(\left|\boldsymbol{\sigma}\left(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h\right) : \boldsymbol{\varepsilon}\left(\boldsymbol{\mathcal{R}}^h\left(\boldsymbol{\mathcal{P}}^h(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h - \boldsymbol{v}_h)\right)\right)\right| + \left|\boldsymbol{f}_\circ\mathbb{G}_h \cdot \boldsymbol{\mathcal{R}}^h\left(\boldsymbol{\mathcal{P}}^h(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h - \boldsymbol{v}_h)\right)\right|\right)\mathrm{d}\boldsymbol{x} \\
&+ \sup_{\boldsymbol{s}\in\widetilde{\Gamma}_n^h}\Big|\big|\det(D\mathbb{G}_h(\boldsymbol{s}))\big| - 1\Big| \int_{\widetilde{\Gamma}_n^h} \left|\boldsymbol{g}_\circ\mathbb{G}_h \cdot \boldsymbol{\mathcal{R}}^h\left(\boldsymbol{\mathcal{P}}^h(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h - \boldsymbol{v}_h)\right)\right|\,\mathrm{d}\boldsymbol{s}.
\end{aligned} \tag{4.10}
$$

To estimate the first term in the right-hand side of inequality (4.10), we use the transformation rule, the continuity property (2.3) and the Young inequality to get

$$
\begin{aligned}
A_h&\left(\boldsymbol{u}_{\varepsilon h} - \boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h, \boldsymbol{\mathcal{R}}^h\left(\boldsymbol{\mathcal{P}}^h(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h - \boldsymbol{v}_h)\right)\right) \\
&= \int_{\widetilde{\Omega}_h} \boldsymbol{\sigma}\left(\boldsymbol{u}_{\varepsilon h} - \boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h\right) : \boldsymbol{\varepsilon}\left(\boldsymbol{\mathcal{R}}^h\left(\boldsymbol{\mathcal{P}}^h(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h - \boldsymbol{v}_h)\right)\right)\mathrm{d}\boldsymbol{x} \\
&= \int_{\Omega} \big|\det(D\mathbb{G}_h^{-1})\big| \boldsymbol{\sigma}\left(\overline{\boldsymbol{u}}_{\varepsilon h} - \boldsymbol{u}_\varepsilon\right) : \boldsymbol{\varepsilon}\left(\boldsymbol{\mathcal{R}}^h\left(\boldsymbol{\mathcal{P}}^h(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h - \boldsymbol{v}_h)\right)_\circ \mathbb{G}_h^{-1}\right)\mathrm{d}\boldsymbol{x} \\
&\leq c\|\boldsymbol{u}_\varepsilon - \overline{\boldsymbol{u}}_{\varepsilon h}\|_{1,\Omega} \left\|\boldsymbol{\mathcal{R}}^h\left(\boldsymbol{\mathcal{P}}^h(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h - \boldsymbol{v}_h)\right)_\circ \mathbb{G}_h^{-1}\right\|_{1,\Omega} \\
&\leq c\|\boldsymbol{u}_\varepsilon - \overline{\boldsymbol{u}}_{\varepsilon h}\|_{1,\Omega} \left\|\boldsymbol{\mathcal{R}}^h\left(\boldsymbol{\mathcal{P}}^h(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h - \boldsymbol{v}_h)\right)\right\|_{1,\widetilde{\Omega}_h} \\
&\leq \frac{1}{2\alpha}\|\boldsymbol{u}_\varepsilon - \overline{\boldsymbol{u}}_{\varepsilon h}\|_{1,\Omega}^2 + \frac{\alpha c^2}{2}\left\|\boldsymbol{\mathcal{R}}^h\left(\boldsymbol{\mathcal{P}}^h(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h - \boldsymbol{v}_h)\right)\right\|_{1,\widetilde{\Omega}_h}^2.
\end{aligned} \tag{4.11}
$$

For the second term in the right-hand side of inequality (4.10), we use the transformation rule, the continuity property (2.3), the Cauchy–Schwarz inequality, the trace theorem and the Young inequality:

$$
\begin{aligned}
&\sup_{\boldsymbol{x}\in\widetilde{\Omega}_h}\Big|1 - \big|\det(D\mathbb{G}_h(\boldsymbol{x}))\big|\Big| \int_{\widetilde{\Omega}_h} \left|\boldsymbol{\sigma}\left(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h\right) : \boldsymbol{\varepsilon}\left(\boldsymbol{\mathcal{R}}^h\left(\boldsymbol{\mathcal{P}}^h(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h - \boldsymbol{v}_h)\right)\right)\right|\mathrm{d}\boldsymbol{x} \\
&\qquad + \sup_{\boldsymbol{x}\in\widetilde{\Omega}_h}\Big|1 - \big|\det(D\mathbb{G}_h(\boldsymbol{x}))\big|\Big| \int_{\widetilde{\Omega}_h} \left|\boldsymbol{f}_\circ\mathbb{G}_h \cdot \boldsymbol{\mathcal{R}}^h\left(\boldsymbol{\mathcal{P}}^h(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h - \boldsymbol{v}_h)\right)\right|\mathrm{d}\boldsymbol{x}
\end{aligned}
$$

$$
\begin{aligned}
&\le c \sup_{\boldsymbol{x}\in\widetilde{\Omega}_h} \left|1-\left|\det(D\mathbb{G}_h(\boldsymbol{x}))\right|\right| \left(\|\boldsymbol{u}_\varepsilon\|_{1,\Omega}+\|\boldsymbol{f}\|_{0,\Omega}\right) \left\|\boldsymbol{\mathcal{R}}^h\left(\boldsymbol{\mathcal{P}}^h(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h-\boldsymbol{v}_h)\right)\right\|_{1,\widetilde{\Omega}_h}\\
&\le \frac{c^2}{2} \sup_{\boldsymbol{x}\in\widetilde{\Omega}_h} \left|1-\left|\det(D\mathbb{G}_h(\boldsymbol{x}))\right|\right|^2 \left(\|\boldsymbol{u}_\varepsilon\|_{1,\Omega}+\|\boldsymbol{f}\|_{0,\Omega}\right)^2+\frac{1}{2}\left\|\boldsymbol{\mathcal{R}}^h\left(\boldsymbol{\mathcal{P}}^h(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h-\boldsymbol{v}_h)\right)\right\|^2_{1,\widetilde{\Omega}_h}. \qquad (4.12)
\end{aligned}
$$

The third term in the right-hand side of inequality (4.10) is estimated by using the Cauchy–Schwarz inequality, the trace theorem and the transformation rule:

$$
\begin{aligned}
&\sup_{\boldsymbol{s}\in\widetilde{\Gamma}_n^h}\left|\left|\det(D\mathbb{G}_h(\boldsymbol{s}))\right|-1\right| \int_{\widetilde{\Gamma}_n^h}\left|\boldsymbol{g}_\circ\mathbb{G}_h\cdot\boldsymbol{\mathcal{R}}^h\left(\boldsymbol{\mathcal{P}}^h(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h-\boldsymbol{v}_h)\right)\right|\,\mathrm{d}\boldsymbol{s}\\
&\quad\le \sup_{\boldsymbol{s}\in\widetilde{\Gamma}_n^h}\left|\left|\det(D\mathbb{G}_h(\boldsymbol{s}))\right|-1\right| \left\|\boldsymbol{g}_\circ\mathbb{G}_h\right\|_{0,\widetilde{\Gamma}_n^h}\left\|\boldsymbol{\mathcal{R}}^h\left(\boldsymbol{\mathcal{P}}^h(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h-\boldsymbol{v}_h)\right)\right\|_{0,\widetilde{\Gamma}_n^h}\\
&\quad\le c\sup_{\boldsymbol{s}\in\widetilde{\Gamma}_n^h}\left|\left|\det(D\mathbb{G}_h(\boldsymbol{s}))\right|-1\right| \|\boldsymbol{g}\|_{0,\Gamma_n}\left\|\boldsymbol{\mathcal{R}}^h\left(\boldsymbol{\mathcal{P}}^h(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h-\boldsymbol{v}_h)\right)\right\|_{1,\widetilde{\Omega}_h}\\
&\quad\le \frac{c^2}{2}\sup_{\boldsymbol{s}\in\widetilde{\Gamma}_n^h}\left|\left|\det(D\mathbb{G}_h(\boldsymbol{s}))\right|-1\right|^2 \|\boldsymbol{g}\|^2_{0,\Gamma_n}+\frac{1}{2}\left\|\boldsymbol{\mathcal{R}}^h\left(\boldsymbol{\mathcal{P}}^h(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h-\boldsymbol{v}_h)\right)\right\|^2_{1,\widetilde{\Omega}_h}. \qquad (4.13)
\end{aligned}
$$

From estimates (4.11), (4.12) and (4.13), the left-hand term of inequality (4.10) is bounded by

$$
\begin{aligned}
&\int_{\widetilde{\Gamma}_c^h}\left|\det(D\mathbb{G}_h)\right|\left(\frac{1}{\varepsilon}\left[\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h-\ell_\circ\mathbb{G}_h\right]^+-\frac{1}{\varepsilon}\left[\boldsymbol{u}_{\varepsilon h}\cdot\boldsymbol{n}_\circ\mathbb{G}_h-\ell_\circ\mathbb{G}_h\right]^+\right)\boldsymbol{\mathcal{R}}^h\left(\boldsymbol{\mathcal{P}}^h(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h-\boldsymbol{v}_h)\right)\cdot\boldsymbol{n}_\circ\mathbb{G}_h\,\mathrm{d}\boldsymbol{s}\\
&\quad\le c\sup_{\boldsymbol{x}\in\overline{\widetilde{\Omega}}_h}\left|1-\left|\det(D\mathbb{G}_h(\boldsymbol{x}))\right|\right|^2\left(\|\boldsymbol{u}_\varepsilon\|_{1,\Omega}+\|\boldsymbol{f}\|_{0,\Omega}+\|\boldsymbol{g}\|_{0,\Gamma_n}\right)^2\\
&\quad\quad+\left(1+\frac{\alpha c^2}{2}\right)\left\|\boldsymbol{\mathcal{R}}^h\left(\boldsymbol{\mathcal{P}}^h(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h-\boldsymbol{v}_h)\right)\right\|^2_{1,\widetilde{\Omega}_h}+\frac{1}{2\alpha}\|\boldsymbol{u}_\varepsilon-\overline{\boldsymbol{u}}_{\varepsilon h}\|^2_{1,\Omega}.
\end{aligned}
$$

By the stability properties (3.7)–(3.9), the transformation rule, the triangle inequality and the Young inequality, the previous estimate finally becomes

$$
\begin{aligned}
&\int_{\widetilde{\Gamma}_c^h}\left|\det(D\mathbb{G}_h)\right|\left(\frac{1}{\varepsilon}\left[\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h-\ell_\circ\mathbb{G}_h\right]^+-\frac{1}{\varepsilon}\left[\boldsymbol{u}_{\varepsilon h}\cdot\boldsymbol{n}_\circ\mathbb{G}_h-\ell_\circ\mathbb{G}_h\right]^+\right)\boldsymbol{\mathcal{R}}^h\left(\boldsymbol{\mathcal{P}}^h(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h-\boldsymbol{v}_h)\right)\cdot\boldsymbol{n}_\circ\mathbb{G}_h\,\mathrm{d}\boldsymbol{s}\\
&\quad\le c\sup_{\boldsymbol{x}\in\overline{\widetilde{\Omega}}_h}\left|1-\left|\det(D\mathbb{G}_h(\boldsymbol{x}))\right|\right|^2\left(\|\boldsymbol{u}_\varepsilon\|_{1,\Omega}+\|\boldsymbol{f}\|_{0,\Omega}+\|\boldsymbol{g}\|_{0,\Gamma_n}\right)^2\\
&\quad\quad+\left(1+\frac{\alpha c^2}{2}\right)\left\|\boldsymbol{\mathcal{P}}^h(\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h-\boldsymbol{v}_h)\right\|^2_{\frac{1}{2},\widetilde{\Gamma}_c^h}+\frac{1}{2\alpha}\|\boldsymbol{u}_\varepsilon-\overline{\boldsymbol{u}}_{\varepsilon h}\|^2_{1,\Omega}
\end{aligned}
$$

$$
\begin{aligned}
&\le c \sup_{\boldsymbol{x}\in\overline{\widetilde{\Omega}}_h} \Big|1-\big|\det(D\mathbb{G}_h(\boldsymbol{x}))\big|\Big|^2 \left(\|\boldsymbol{u}_\varepsilon\|_{1,\Omega}+\|\boldsymbol{f}\|_{0,\Omega}+\|\boldsymbol{g}\|_{0,\Gamma_n}\right)^2 \\
&\quad+\left(1+\frac{\alpha c^2}{2}\right)\|\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h-\boldsymbol{v}_h\|^2_{\frac12,\widetilde{\Gamma}_c^h}+\frac{1}{2\alpha}\|\boldsymbol{u}_\varepsilon-\overline{\boldsymbol{u}}_{\varepsilon h}\|^2_{1,\Omega} \\
&\le c \sup_{\boldsymbol{x}\in\overline{\widetilde{\Omega}}_h} \Big|1-\big|\det(D\mathbb{G}_h(\boldsymbol{x}))\big|\Big|^2 \left(\|\boldsymbol{u}_\varepsilon\|_{1,\Omega}+\|\boldsymbol{f}\|_{0,\Omega}+\|\boldsymbol{g}\|_{0,\Gamma_n}\right)^2 \\
&\quad+c_1\left(1+\frac{\alpha c^2}{2}\right)\left(\|\boldsymbol{u}_\varepsilon-\boldsymbol{u}\|^2_{1,\Omega}+\|\boldsymbol{u}-\overline{\boldsymbol{v}}_h\|^2_{1,\Omega}\right)+\frac{1}{2\alpha}\|\boldsymbol{u}_\varepsilon-\overline{\boldsymbol{u}}_{\varepsilon h}\|^2_{1,\Omega}. \qquad (4.14)
\end{aligned}
$$

Finally, taking into account estimates (4.9) and (4.14), estimate (4.8) becomes

$$
\begin{aligned}
&\int_{\widetilde{\Gamma}_c^h}\big|\det(D\mathbb{G}_h)\big|\left(\frac{1}{\varepsilon}\big[\boldsymbol{u}_{\varepsilon h}\cdot\boldsymbol{n}_\circ\mathbb{G}_h-\ell_\circ\mathbb{G}_h\big]^+-\frac{1}{\varepsilon}\big[\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h-\ell_\circ\mathbb{G}_h\big]^+\right)\left(\boldsymbol{v}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h-\boldsymbol{u}_{\varepsilon\circ}\mathbb{G}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h\right)\mathrm{d}\boldsymbol{s} \\
&\le\frac{1}{2\alpha}\|\boldsymbol{u}_\varepsilon-\overline{\boldsymbol{u}}_{\varepsilon h}\|^2_{1,\Omega}+c_2\left(1+\frac{\alpha c^2}{2}\right)\left(\|\boldsymbol{u}_\varepsilon-\boldsymbol{u}\|^2_{1,\Omega}+\|\boldsymbol{u}-\overline{\boldsymbol{v}}_h\|^2_{1,\Omega}\right) \\
&\quad+ch\left\|\frac{1}{\varepsilon}[\boldsymbol{u}_\varepsilon\cdot\boldsymbol{n}-\ell]^+-\frac{1}{\varepsilon}[\overline{\boldsymbol{u}}_{\varepsilon h}\cdot\boldsymbol{n}-\ell]^+\right\|^2_{0,\Gamma_c} \\
&\quad+c\sup_{\boldsymbol{x}\in\overline{\widetilde{\Omega}}_h}\Big|1-\big|\det(D\mathbb{G}_h(\boldsymbol{x}))\big|\Big|^2\left(\|\boldsymbol{u}_\varepsilon\|_{1,\Omega}+\|\boldsymbol{f}\|_{0,\Omega}+\|\boldsymbol{g}\|_{0,\Gamma_n}\right)^2. \qquad (4.15)
\end{aligned}
$$

Combining inequalities (4.7) and (4.15), inequality (4.6) finally becomes

$$
\begin{aligned}
&\frac{1}{\varepsilon}\int_{\widetilde{\Gamma}_c^h}\big|\det(D\mathbb{G}_h)\big|\big[\boldsymbol{u}_{\varepsilon h}\cdot\boldsymbol{n}_\circ\mathbb{G}_h-\ell_\circ\mathbb{G}_h\big]^+\left(\boldsymbol{v}_h\cdot\boldsymbol{n}_\circ\mathbb{G}_h-\boldsymbol{u}_{\varepsilon h}\cdot\boldsymbol{n}_\circ\mathbb{G}_h\right)\mathrm{d}s-\frac{1}{\varepsilon}\int_{\Gamma_c}\big[\boldsymbol{u}_\varepsilon\cdot\boldsymbol{n}-\ell\big]^+\left(\overline{\boldsymbol{v}}_h\cdot\boldsymbol{n}-\overline{\boldsymbol{u}}_{\varepsilon h}\cdot\boldsymbol{n}\right)\mathrm{d}\boldsymbol{s} \\
&\le\frac{1}{2\alpha}\|\boldsymbol{u}_\varepsilon-\overline{\boldsymbol{u}}_{\varepsilon h}\|^2_{1,\Omega}+(ch-\varepsilon)\left\|\frac{1}{\varepsilon}[\boldsymbol{u}_\varepsilon\cdot\boldsymbol{n}-\ell]^+-\frac{1}{\varepsilon}[\overline{\boldsymbol{u}}_{\varepsilon h}\cdot\boldsymbol{n}-\ell]^+\right\|^2_{0,\Gamma_c}+c_2\left(1+\frac{\alpha c^2}{2}\right)\Big(\|\boldsymbol{u}_\varepsilon-\boldsymbol{u}\|^2_{1,\Omega} \\
&\quad+\|\boldsymbol{u}-\overline{\boldsymbol{v}}_h\|^2_{1,\Omega}\Big)+c\sup_{\boldsymbol{x}\in\overline{\widetilde{\Omega}}_h}\Big|1-\big|\det(D\mathbb{G}_h(\boldsymbol{x}))\big|\Big|^2\left(\|\boldsymbol{u}_\varepsilon\|_{1,\Omega}+\|\boldsymbol{f}\|_{0,\Omega}+\|\boldsymbol{g}\|_{0,\Gamma_n}\right)^2. \qquad (4.16)
\end{aligned}
$$

Rewriting estimate (4.5) by taking into account inequality (4.16), we establish the desired result:

$$
\begin{aligned}
&\left(m-\frac{1}{\alpha}\right)\|\boldsymbol{u}_\varepsilon-\overline{\boldsymbol{u}}_{\varepsilon h}\|^2_{1,\Omega}+(\varepsilon-ch)\left\|\frac{1}{\varepsilon}\big[\boldsymbol{u}_\varepsilon\cdot\boldsymbol{n}-\ell\big]^+-\frac{1}{\varepsilon}\big[\overline{\boldsymbol{u}}_{\varepsilon h}\cdot\boldsymbol{n}-\ell\big]^+\right\|^2_{0,\Gamma_c} \\
&\quad\le c_3\left(2+\alpha M^2+\frac{\alpha c^2}{2}\right)\left(\|\boldsymbol{u}-\overline{\boldsymbol{v}}_h\|^2_{1,\Omega}+\|\boldsymbol{u}_\varepsilon-\boldsymbol{u}\|^2_{1,\Omega}\right)
\end{aligned}
$$

$$+\, c \sup_{\boldsymbol{x}\in\overline{\widetilde{\Omega}}_h} \left|1-\left|\det(D\mathbb{G}_h(\boldsymbol{x}))\right|\right|^2 \left(\|\boldsymbol{u}_\varepsilon\|_{1,\Omega}+\|\boldsymbol{f}\|_{0,\Omega}+\|\boldsymbol{g}\|_{0,\Gamma_n}\right)^2$$

$$+\left|A_h(\boldsymbol{u}_{\varepsilon h},\boldsymbol{v}_h-\boldsymbol{u}_{\varepsilon h})-A(\overline{\boldsymbol{u}}_{\varepsilon h},\overline{\boldsymbol{v}}_h-\overline{\boldsymbol{u}}_{\varepsilon h})\right|+\left|F_h(\boldsymbol{v}_h-\boldsymbol{u}_{\varepsilon h})-F(\overline{\boldsymbol{v}}_h-\overline{\boldsymbol{u}}_{\varepsilon h})\right|$$

for $\alpha > \frac{1}{m}$, where c, c_3 are positive constants which are independent of $\boldsymbol{u}$, the penalty parameter ε and the mesh size h. □

We now devote ourselves to bounding the error $\|\boldsymbol{u}-\overline{\boldsymbol{u}}_{\varepsilon h}\|_{1,\Omega}+(\varepsilon^{\frac{1}{2}}-ch^{\frac{1}{2}})\|\boldsymbol{\sigma}_n(\boldsymbol{u})+\frac{1}{\varepsilon}[\overline{\boldsymbol{u}}_{\varepsilon h}\cdot\boldsymbol{n}-\ell]^+\|_{0,\Gamma_c}$ with respect to the mesh size h and the penalty parameter ε, by estimating each term on the right-hand side of estimate (4.1) before we conclude, thanks to Theorem 2.1, by using the triangle inequality.

THEOREM 4.2 Under the assumptions of Theorem 4.1, if the solution $\boldsymbol{u}$ of problem (2.1) belongs to $\boldsymbol{H}^{\frac{3}{2}+\nu}(\Omega)$ with $0<\nu\le\frac{1}{2}$, then there exists $0<h_*$ such that for $0<h\le h_*$ and for $\varepsilon>0$, we obtain the following *a priori* estimate:

$$\|\boldsymbol{u}-\overline{\boldsymbol{u}}_{\varepsilon h}\|_{1,\Omega}+\left(\varepsilon^{\frac{1}{2}}-c_o h^{\frac{1}{2}}\right)\left\|\boldsymbol{\sigma}_n(\boldsymbol{u})+\frac{1}{\varepsilon}[\overline{\boldsymbol{u}}_{\varepsilon h}\cdot\boldsymbol{n}-\ell]^+\right\|_{0,\Gamma_c}$$
$$\le c\left(h^{\frac{1}{2}+\nu}+\varepsilon^{\frac{1}{2}+\nu}\right)\left(\|\boldsymbol{u}\|_{\frac{3}{2}+\nu,\Omega}+\|\boldsymbol{f}\|_{0,\Omega}+\|\boldsymbol{g}\|_{0,\Gamma_n}\right), \tag{4.17}$$

where c, c_o are positive constants independent of the solution $\boldsymbol{u}$, the mesh size h and the penalty parameter ε.

Proof. We need to bound each term in the right-hand side of estimate (4.1). We make the choice $\boldsymbol{v}_h := \mathcal{I}_h^1(\boldsymbol{u}_\circ\mathbb{G}_h)$, where $\mathcal{I}_h^1(\cdot)$ is the Lagrange interpolation operator of degree 1 mapping onto $\boldsymbol{V}_h$. Clearly, the following approximation property holds (Brenner & Scott, 2002, Sect. 4.4) for $-\frac{1}{2}<\nu\le\frac{1}{2}$:

$$\|\boldsymbol{u}_\circ\mathbb{G}_h-\boldsymbol{v}_h\|_{1,\widetilde{\mathrm{T}}}\le ch_{\widetilde{\mathrm{T}}}^{\frac{1}{2}+\nu}\|\boldsymbol{u}_\circ\mathbb{G}_h\|_{\frac{3}{2}+\nu,\widetilde{\mathrm{T}}}\quad\forall\,\widetilde{\mathrm{T}}\in\widetilde{\mathcal{T}}_h. \tag{4.18}$$

In view of the transformation rule, the local estimate (4.18) and relation (3.2), we transfer to Ω as

$$\begin{aligned}\|\boldsymbol{u}-\overline{\boldsymbol{v}}_h\|_{1,\Omega}=\left(\sum_{\mathrm{T}\in\mathcal{T}_h}\|\boldsymbol{u}-\overline{\boldsymbol{v}}_h\|_{1,\mathrm{T}}^2\right)^{\frac{1}{2}}&\le c\left(\sum_{\widetilde{\mathrm{T}}\in\widetilde{\mathcal{T}}_h}\|\boldsymbol{u}_\circ\mathbb{G}_h-\boldsymbol{v}_h\|_{1,\widetilde{\mathrm{T}}}^2\right)^{\frac{1}{2}}\\&\le ch^{\frac{1}{2}+\nu}\left(\sum_{\widetilde{\mathrm{T}}\in\widetilde{\mathcal{T}}_h}\|\boldsymbol{u}_\circ\mathbb{G}_h\|_{\frac{3}{2}+\nu,\widetilde{\mathrm{T}}}^2\right)^{\frac{1}{2}}\\&\le ch^{\frac{1}{2}+\nu}\|\boldsymbol{u}\|_{\frac{3}{2}+\nu,\Omega}.\end{aligned} \tag{4.19}$$

Taking $\boldsymbol{v}:=\boldsymbol{u}_\varepsilon$ in problem (2.4) by using the equality in relation (2.5), we obtain

$$A(\boldsymbol{u}_\varepsilon,\boldsymbol{u}_\varepsilon)+\frac{1}{\varepsilon}\int_{\Gamma_c}\left(\left[\boldsymbol{u}_\varepsilon\cdot\boldsymbol{n}\right]^+\right)^2\,\mathrm{d}\boldsymbol{s}=F(\boldsymbol{u}_\varepsilon).$$

Using the fact that the bilinear form $A(\cdot,\cdot)$ is elliptic, the continuity of the linear form $F(\cdot)$, the Cauchy–Schwarz inequality and the trace theorem, we obtain

$$m\|\boldsymbol{u}_\varepsilon\|_{1,\Omega}^2 \leq A(\boldsymbol{u}_\varepsilon,\boldsymbol{u}_\varepsilon) + \frac{1}{\varepsilon}\int_{\Gamma_c}\left(\left[\boldsymbol{u}_\varepsilon\cdot\boldsymbol{n}\right]^+\right)^2 \,\mathrm{d}\boldsymbol{s} = F(\boldsymbol{u}_\varepsilon) \leq c\big(\|\boldsymbol{f}\|_{0,\Omega} + \|\boldsymbol{g}\|_{0,\Gamma_n}\big)\|\boldsymbol{u}_\varepsilon\|_{1,\Omega},$$
$$m\|\boldsymbol{u}_\varepsilon\|_{1,\Omega} \leq c\big(\|\boldsymbol{f}\|_{0,\Omega} + \|\boldsymbol{g}\|_{0,\Gamma_n}\big). \tag{4.20}$$

The terms $|A_h(\boldsymbol{u}_{\varepsilon h},\boldsymbol{v}_h-\boldsymbol{u}_{\varepsilon h}) - A(\overline{\boldsymbol{u}}_{\varepsilon h},\overline{\boldsymbol{v}}_h-\overline{\boldsymbol{u}}_{\varepsilon h})|$ and $|F(\overline{\boldsymbol{v}}_h-\overline{\boldsymbol{u}}_{\varepsilon h}) - F_h(\boldsymbol{v}_h-\boldsymbol{u}_{\varepsilon h})|$ have already been treated in Bänsch & Deckelnick (1999, proof of Theorem 3.4). We will start with their estimates in order to adapt them to this context of the penalization method. Let us rename these terms

$$T_{1,h} := \left|A_h(\boldsymbol{u}_{\varepsilon h},\boldsymbol{v}_h-\boldsymbol{u}_{\varepsilon h}) - A(\overline{\boldsymbol{u}}_{\varepsilon h},\overline{\boldsymbol{v}}_h-\overline{\boldsymbol{u}}_{\varepsilon h})\right|,$$
$$T_{2,h} := \left|F(\overline{\boldsymbol{v}}_h-\overline{\boldsymbol{u}}_{\varepsilon h}) - F_h(\boldsymbol{v}_h-\boldsymbol{u}_{\varepsilon h})\right|.$$

From Bänsch & Deckelnick (1999, Lemma 3.2), there exists $h_1>0$ such that for all $0<h\leq h_1$,

$$T_{1,h} \leq ch\left(\sum_{\mathsf{T}\cap\partial\Omega\neq\emptyset}\|\overline{\boldsymbol{u}}_{\varepsilon h}\|_{1,\mathsf{T}}^2\right)^{\frac12}\left(\sum_{\mathsf{T}\cap\partial\Omega\neq\emptyset}\|\overline{\boldsymbol{v}}_h-\overline{\boldsymbol{u}}_{\varepsilon h}\|_{1,\mathsf{T}}^2\right)^{\frac12}$$
$$\leq ch\|\overline{\boldsymbol{u}}_{\varepsilon h}\|_{1,\Omega}\|\overline{\boldsymbol{v}}_h-\overline{\boldsymbol{u}}_{\varepsilon h}\|_{1,\Omega}. \tag{4.21}$$

Using the triangle inequality on the terms in the right-hand side of inequality (4.21) we obtain

$$\begin{aligned}T_{1,h} &\leq ch\big(\|\boldsymbol{u}_\varepsilon\|_{1,\Omega} + \|\boldsymbol{u}_\varepsilon-\overline{\boldsymbol{u}}_{\varepsilon h}\|_{1,\Omega}\big)\big(\|\boldsymbol{u}_\varepsilon-\overline{\boldsymbol{u}}_{\varepsilon h}\|_{1,\Omega} + \|\boldsymbol{u}_\varepsilon-\overline{\boldsymbol{v}}_h\|_{1,\Omega}\big)\\ &\leq ch\big(\|\boldsymbol{u}_\varepsilon\|_{1,\Omega}\|\boldsymbol{u}_\varepsilon-\overline{\boldsymbol{u}}_{\varepsilon h}\|_{1,\Omega} + \|\boldsymbol{u}_\varepsilon\|_{1,\Omega}\|\boldsymbol{u}_\varepsilon-\overline{\boldsymbol{v}}_h\|_{1,\Omega} + \|\boldsymbol{u}_\varepsilon-\overline{\boldsymbol{u}}_{\varepsilon h}\|_{1,\Omega}^2\\ &\quad + \|\boldsymbol{u}_\varepsilon-\overline{\boldsymbol{u}}_{\varepsilon h}\|_{1,\Omega}\|\boldsymbol{u}_\varepsilon-\overline{\boldsymbol{v}}_h\|_{1,\Omega}\big).\end{aligned}$$

Finally, by the Young inequality and using estimate (4.20) we bound the term $T_{1,h}$ as

$$\begin{aligned}T_{1,h} &\leq \left(ch_1+\frac{1}{\alpha}\right)\|\boldsymbol{u}_\varepsilon-\overline{\boldsymbol{u}}_{\varepsilon h}\|_{1,\Omega}^2 + c\left(\frac{1+\alpha h_1^2}{2}\right)\|\boldsymbol{u}_\varepsilon-\overline{\boldsymbol{v}}_h\|_{1,\Omega}^2 + c\left(\frac{1+\alpha}{2}\right)h^2\|\boldsymbol{u}_\varepsilon\|_{1,\Omega}^2\\ &\leq \left(ch_1+\frac{1}{\alpha}\right)\|\boldsymbol{u}_\varepsilon-\overline{\boldsymbol{u}}_{\varepsilon h}\|_{1,\Omega}^2 + c\left(1+\alpha h_1^2\right)\left(\|\boldsymbol{u}_\varepsilon-\boldsymbol{u}\|_{1,\Omega}^2 + \|\boldsymbol{u}-\overline{\boldsymbol{v}}_h\|_{1,\Omega}^2\right)\\ &\quad + c\left(\frac{1+\alpha}{2}\right)h^2\left(\|\boldsymbol{f}\|_{0,\Omega} + \|\boldsymbol{g}\|_{0,\Gamma_n}\right)^2\end{aligned}$$

for all $\alpha > 0$ and $0 < h \le h_1$. Hence, taking into account estimates (2.7) and (4.19), we obtain

$$T_{1,h}^{\frac{1}{2}} = \left|A_h(\boldsymbol{u}_{\varepsilon h}, \boldsymbol{v}_h - \boldsymbol{u}_{\varepsilon h}) - A(\overline{\boldsymbol{u}}_{\varepsilon h}, \overline{\boldsymbol{v}}_h - \overline{\boldsymbol{u}}_{\varepsilon h})\right|^{\frac{1}{2}}$$
$$\le \left(ch_1 + \frac{1}{\alpha}\right)^{\frac{1}{2}} \|\boldsymbol{u}_\varepsilon - \overline{\boldsymbol{u}}_{\varepsilon h}\|_{1,\Omega} + c\left(h^{\frac{1}{2}+\nu} + \varepsilon^{\frac{1}{2}+\nu}\right)\left(\|\boldsymbol{u}\|_{\frac{3}{2}+\nu,\Omega} + \|\boldsymbol{f}\|_{0,\Omega} + \|\boldsymbol{g}\|_{0,\Gamma_n}\right). \tag{4.22}$$

In order to estimate the second term $T_{2,h}$, we first remark that

$$T_{2,h} \le \left|\int_\Omega \boldsymbol{f}\cdot(\overline{\boldsymbol{v}}_h - \overline{\boldsymbol{u}}_{\varepsilon h})\,\mathrm{d}\boldsymbol{x} - \int_{\widetilde{\Omega}_h} \boldsymbol{f}_\circ\mathbb{G}_h\cdot(\boldsymbol{v}_h - \boldsymbol{u}_{\varepsilon h})\,\mathrm{d}\boldsymbol{x}\right|$$
$$+ \left|\int_{\Gamma_n} \boldsymbol{g}\cdot(\overline{\boldsymbol{v}}_h - \overline{\boldsymbol{u}}_{\varepsilon h})\,\mathrm{d}\boldsymbol{s} - \int_{\widetilde{\Gamma}_n^h} \boldsymbol{g}_\circ\mathbb{G}_h\cdot(\boldsymbol{v}_h - \boldsymbol{u}_{\varepsilon h})\,\mathrm{d}\boldsymbol{s}\right|. \tag{4.23}$$

The first term in the right-hand side of inequality (4.23) is bounded in Bänsch & Deckelnick (1999, proof of Theorem 3.4) with the regularity $\boldsymbol{f} \in \boldsymbol{H}^1(\Omega)$. But regarding the required regularity of the solution, namely $\boldsymbol{u} \in \boldsymbol{H}^{\frac{3}{2}+\nu}(\Omega)$ with $0 < \nu \le \frac{1}{2}$, less regularity of the function $\boldsymbol{f}$ (namely $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega)$) is needed. We will take this estimate again with the regularity $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega)$. Using the transformation rule, the Cauchy–Schwarz inequality and estimate (3.4), we obtain

$$\left|\int_\Omega \boldsymbol{f}\cdot(\overline{\boldsymbol{v}}_h - \overline{\boldsymbol{u}}_{\varepsilon h})\,\mathrm{d}\boldsymbol{x} - \int_{\widetilde{\Omega}_h} \boldsymbol{f}_\circ\mathbb{G}_h\cdot(\boldsymbol{v}_h - \boldsymbol{u}_{\varepsilon h})\,\mathrm{d}\boldsymbol{x}\right| = \left|\sum_{\widetilde{\mathrm{T}}\cap\widetilde{\Omega}_h\neq\emptyset}\int_{\widetilde{\mathrm{T}}}\Big(\left|\det(D\mathbb{G}_h)\right| - 1\Big)\boldsymbol{f}_\circ\mathbb{G}_h\cdot\big(\boldsymbol{v}_h - \boldsymbol{u}_{\varepsilon h}\big)\,\mathrm{d}\boldsymbol{x}\right|$$
$$\le ch\sum_{\widetilde{\mathrm{T}}\cap\widetilde{\Omega}_h\neq\emptyset}\|\boldsymbol{f}_\circ\mathbb{G}_h\|_{0,\widetilde{\mathrm{T}}}\|\boldsymbol{v}_h - \boldsymbol{u}_{\varepsilon h}\|_{0,\widetilde{\mathrm{T}}}$$
$$\le ch\left(\sum_{\widetilde{\mathrm{T}}\cap\widetilde{\Omega}_h\neq\emptyset}\|\boldsymbol{f}_\circ\mathbb{G}_h\|_{0,\widetilde{\mathrm{T}}}^2\right)^{\frac{1}{2}}\left(\sum_{\widetilde{\mathrm{T}}\cap\widetilde{\Omega}_h\neq\emptyset}\|\boldsymbol{v}_h - \boldsymbol{u}_{\varepsilon h}\|_{0,\widetilde{\mathrm{T}}}^2\right)^{\frac{1}{2}}.$$

By the transformation rule, the triangle inequality and the Young inequality, we obtain

$$\left|\int_\Omega \boldsymbol{f}\cdot\big(\overline{\boldsymbol{v}}_h - \overline{\boldsymbol{u}}_{\varepsilon h}\big)\,\mathrm{d}\boldsymbol{x} - \int_{\widetilde{\Omega}_h} \boldsymbol{f}_\circ\mathbb{G}_h\cdot(\boldsymbol{v}_h - \boldsymbol{u}_{\varepsilon h})\,\mathrm{d}\boldsymbol{x}\right|$$
$$\le ch\|\boldsymbol{f}\|_{0,\Omega}\|\overline{\boldsymbol{v}}_h - \overline{\boldsymbol{u}}_{\varepsilon h}\|_{1,\Omega}$$
$$\le ch\|\boldsymbol{f}\|_{0,\Omega}\Big(\|\boldsymbol{u} - \overline{\boldsymbol{v}}_h\|_{1,\Omega} + \|\boldsymbol{u} - \boldsymbol{u}_\varepsilon\|_{1,\Omega} + \|\boldsymbol{u}_\varepsilon - \overline{\boldsymbol{u}}_{\varepsilon h}\|_{1,\Omega}\Big)$$
$$\le \frac{1}{2\alpha}\|\boldsymbol{u}_\varepsilon - \overline{\boldsymbol{u}}_{\varepsilon h}\|_{1,\Omega}^2 + \frac{c}{2}\Big(\|\boldsymbol{u} - \overline{\boldsymbol{v}}_h\|_{1,\Omega}^2 + \|\boldsymbol{u} - \boldsymbol{u}_\varepsilon\|_{1,\Omega}^2\Big) + c\left(1 + \frac{\alpha}{2}\right)h^2\,\|\boldsymbol{f}\|_{0,\Omega}^2 \tag{4.24}$$

for all $\alpha > 0$ and $\boldsymbol{v}_h \in \boldsymbol{V}_h$.

To estimate the second term in the right-hand side of inequality (4.23), we proceed in the same way using the Cauchy–Schwarz inequality, the estimate (3.4) and the transformation rule

$$\begin{aligned}
&\left|\int_{\Gamma_n} \boldsymbol{g}\cdot\left(\overline{\boldsymbol{v}}_h-\overline{\boldsymbol{u}}_{\varepsilon h}\right)\mathrm{d}\boldsymbol{s}-\int_{\widetilde{\Gamma}_n^h}\boldsymbol{g}_{\circ}\mathbb{G}_h\cdot\left(\boldsymbol{v}_h-\boldsymbol{u}_{\varepsilon h}\right)\mathrm{d}\boldsymbol{s}\right|\\
&\qquad=\left|\sum_{\widetilde{\mathrm{T}}\cap\widetilde{\Gamma}_n^h\neq\emptyset}\int_{\widetilde{\mathrm{T}}\cap\widetilde{\Gamma}_n^h}\Big(\left|\det(D\mathbb{G}_h)\right|-1\Big)\boldsymbol{g}_{\circ}\mathbb{G}_h\cdot\left(\boldsymbol{v}_h-\boldsymbol{u}_{\varepsilon h}\right)\mathrm{d}\boldsymbol{s}\right|\\
&\qquad\le h\sum_{\widetilde{\mathrm{T}}\cap\widetilde{\Gamma}_n^h\neq\emptyset}\|\boldsymbol{g}_{\circ}\mathbb{G}_h\|_{0,\widetilde{\mathrm{T}}\cap\widetilde{\Gamma}_n^h}\|\boldsymbol{v}_h-\boldsymbol{u}_{\varepsilon h}\|_{0,\widetilde{\mathrm{T}}\cap\widetilde{\Gamma}_n^h}\\
&\qquad\le ch\|\boldsymbol{g}\|_{0,\Gamma_n}\|\overline{\boldsymbol{v}}_h-\overline{\boldsymbol{u}}_{\varepsilon h}\|_{1,\Omega}.
\end{aligned}$$

By the triangle and Young inequalities, we finally have the bound

$$\begin{aligned}
&\left|\int_{\Gamma_n} \boldsymbol{g}\cdot\left(\overline{\boldsymbol{v}}_h-\overline{\boldsymbol{u}}_{\varepsilon h}\right)\mathrm{d}\boldsymbol{s}-\int_{\widetilde{\Gamma}_n^h}\boldsymbol{g}_{\circ}\mathbb{G}_h\cdot\left(\boldsymbol{v}_h-\boldsymbol{u}_{\varepsilon h}\right)\mathrm{d}\boldsymbol{s}\right|\\
&\qquad\le\frac{1}{2\alpha}\|\boldsymbol{u}_\varepsilon-\overline{\boldsymbol{u}}_{\varepsilon h}\|_{1,\Omega}^2+\frac{c}{2}\Big(\|\boldsymbol{u}-\overline{\boldsymbol{v}}_h\|_{1,\Omega}^2+\|\boldsymbol{u}-\boldsymbol{u}_\varepsilon\|_{1,\Omega}^2\Big)+c\left(1+\frac{\alpha}{2}\right)h^2\|\boldsymbol{g}\|_{0,\Gamma_n}^2 \qquad (4.25)
\end{aligned}$$

for all $\alpha>0$ and for $\boldsymbol{v}_h\in\boldsymbol{V}_h$.

From inequalities (4.24) and (4.25), the term $T_{2,h}$ is finally estimated as

$$T_{2,h}\le\frac{1}{\alpha}\|\boldsymbol{u}_\varepsilon-\overline{\boldsymbol{u}}_{\varepsilon h}\|_{1,\Omega}^2+c\Big(\|\boldsymbol{u}-\overline{\boldsymbol{v}}_h\|_{1,\Omega}^2+\|\boldsymbol{u}-\boldsymbol{u}_\varepsilon\|_{1,\Omega}^2\Big)+c\left(1+\frac{\alpha}{2}\right)h^2\Big(\|\boldsymbol{f}\|_{0,\Omega}^2+\|\boldsymbol{g}\|_{0,\Gamma_n}^2\Big). \qquad (4.26)$$

Finally, taking into account estimates (2.7) and (4.19), we obtain from (4.26) the estimate

$$\begin{aligned}
T_{2,h}^{\frac{1}{2}}&=\left|F(\overline{\boldsymbol{v}}_h-\overline{\boldsymbol{u}}_{\varepsilon h})-F_h(\boldsymbol{v}_h-\boldsymbol{u}_{\varepsilon h})\right|^{\frac{1}{2}}\\
&\le\frac{1}{\alpha^{\frac{1}{2}}}\|\boldsymbol{u}_\varepsilon-\overline{\boldsymbol{u}}_{\varepsilon h}\|_{1,\Omega}+c\left(h^{\frac{1}{2}+\nu}+\varepsilon^{\frac{1}{2}+\nu}\right)\left(\|\boldsymbol{u}\|_{\frac{3}{2}+\nu,\Omega}+\|\boldsymbol{f}\|_{0,\Omega}+\|\boldsymbol{g}\|_{0,\Gamma_n}\right). \qquad (4.27)
\end{aligned}$$

Arranging estimates (4.19), (4.20), (4.22), (4.27) and the estimate (3.4), we rewrite estimate (4.1) as

$$\begin{aligned}
&\left(1-\left((ch_1)^{\frac{1}{2}}+\frac{2}{\alpha^{\frac{1}{2}}}\right)\right)\|\boldsymbol{u}_\varepsilon-\overline{\boldsymbol{u}}_{\varepsilon h}\|_{1,\Omega}+\left(\varepsilon^{\frac{1}{2}}-c_o h^{\frac{1}{2}}\right)\left\|\frac{1}{\varepsilon}[\boldsymbol{u}_\varepsilon\cdot\boldsymbol{n}-\ell]^+-\frac{1}{\varepsilon}[\overline{\boldsymbol{u}}_{\varepsilon h}\cdot\boldsymbol{n}-\ell]^+\right\|_{0,\Gamma_c}\\
&\qquad\le c\left(h^{\frac{1}{2}+\nu}+\varepsilon^{\frac{1}{2}+\nu}\right)\left(\|\boldsymbol{u}\|_{\frac{3}{2}+\nu,\Omega}+\|\boldsymbol{f}\|_{0,\Omega}+\|\boldsymbol{g}\|_{0,\Gamma_n}\right) \qquad (4.28)
\end{aligned}$$

for $0 < h \leq h_1$ with $0 < \nu \leq \frac{1}{2}$, where α and h_1 are suitably chosen such that $1-\left((ch_1)^{\frac{1}{2}}+2/\alpha^{\frac{1}{2}}\right) > 0$, and the penalty parameter chosen as $\varepsilon^{\frac{1}{2}} > c_o h^{\frac{1}{2}}$.

Taking into account the relation $\varepsilon^{\frac{1}{2}} > \varepsilon^{\frac{1}{2}} - c_o h^{\frac{1}{2}} > 0$, we obtain from estimate (2.7) the inequality

$$\|\boldsymbol{u}-\boldsymbol{u}_\varepsilon\|_{1,\Omega} + \left(\varepsilon^{\frac{1}{2}} - c_o h^{\frac{1}{2}}\right)\left\|\boldsymbol{\sigma}_n(\boldsymbol{u}) + \frac{1}{\varepsilon}\left[\boldsymbol{u}_\varepsilon\cdot\boldsymbol{n}-\ell\right]^+\right\|_{0,\Gamma_c} \leq c\varepsilon^{\frac{1}{2}+\nu}\|\boldsymbol{u}\|_{\frac{3}{2}+\nu,\Omega}. \tag{4.29}$$

Combining (4.28) and (4.29) by using the triangle inequality, we establish (4.17) for all $0 < h \leq h_* := \min\{h_0, h_1\}$. □

REMARK 4.3 From estimate (4.17) we obtain various error estimates depending on how we choose the penalty parameter ε with respect to the mesh size h.

- If the penalty parameter ε behaves like the mesh size h, that is, $\varepsilon(h) := (c_o+1)^2 h$, we obtain from estimate (4.17) the following *a priori* estimate, for $0 < \nu \leq \frac{1}{2}$:

$$\|\boldsymbol{u}-\overline{\boldsymbol{u}}_{\varepsilon h}\|_{1,\Omega} + h^{\frac{1}{2}}\left\|\boldsymbol{\sigma}_n(\boldsymbol{u}) + \frac{1}{\varepsilon}\left[\overline{\boldsymbol{u}}_{\varepsilon h}\cdot\boldsymbol{n}-\ell\right]^+\right\|_{0,\Gamma_c} \leq ch^{\frac{1}{2}+\nu}\left(\|\boldsymbol{u}\|_{\frac{3}{2}+\nu,\Omega} + \|\boldsymbol{f}\|_{0,\Omega} + \|\boldsymbol{g}\|_{0,\Gamma_n}\right). \tag{4.30}$$

  The main consequence of result (4.30) is the $\mathcal{O}(h^{\frac{1}{2}+\nu})$ convergence rate obtained under $\boldsymbol{H}^{\frac{3}{2}+\nu}(\Omega)$ regularity, without any additional assumptions on the contact boundary Γ_c. Thus, we obtain the same optimal error estimate as the direct finite element approximation of the variational inequality (2.1) established by Drouet & Hild (2015) in the case of a polygonal or polyhedral domain. In particular, we have optimal linear convergence if $\nu = \frac{1}{2}$.

- If the penalty parameter is defined as $\varepsilon(h) := c_o^2 h^\theta, 0 < \theta < 1$, we obtain from (4.17) the following suboptimal estimate of the $\mathcal{O}(h^{\theta(\frac{1}{2}+\nu)})$ convergence rate:

$$\begin{aligned}\|\boldsymbol{u}-\overline{\boldsymbol{u}}_{\varepsilon h}\|_{1,\Omega} + c_o h^{\frac{\theta}{2}}\left(1-h^{\frac{1-\theta}{2}}\right)\left\|\boldsymbol{\sigma}_n(\boldsymbol{u}) + \frac{1}{\varepsilon}\left[\overline{\boldsymbol{u}}_{\varepsilon h}\cdot\boldsymbol{n}-\ell\right]^+\right\|_{0,\Gamma_c}\\ \leq ch^{\theta\left(\frac{1}{2}+\nu\right)}\left(\|\boldsymbol{u}\|_{\frac{3}{2}+\nu,\Omega} + \|\boldsymbol{f}\|_{0,\Omega} + \|\boldsymbol{g}\|_{0,\Gamma_n}\right).\end{aligned} \tag{4.31}$$

  Since $(1-h^{\frac{1-\theta}{2}}) \longrightarrow 1$ if $h \longrightarrow 0$, then in the forthcoming numerical experiments we will compute the error $h^{\frac{\theta}{2}}\|\boldsymbol{\sigma}_n(\boldsymbol{u}) + \frac{1}{\varepsilon}\left[\overline{\boldsymbol{u}}_{\varepsilon h}\cdot\boldsymbol{n}-\ell\right]^+\|_{0,\Gamma_c}$ with respect to the mesh size h.

REMARK 4.4 Another important point to underline is that the result obtained in Theorem 4.2 should also be valid for quadratic finite elements (since the standard Lagrange interpolation estimate (4.19) remains valid), if the continuous estimate (2.7) remains true for a more regular solution (that is, for $\boldsymbol{u} \in \boldsymbol{H}^{\frac{3}{2}+\nu}(\Omega), 0 < \nu \leq \frac{3}{2}$).

5. Numerical experiments

It is not an easy task to find an analytical solution to the contact problem (1.1)–(1.5) in this case of a curved domain. Nevertheless, we build here an analytic reference solution which does not entirely cover all the given unilateral frictionless boundary conditions (1.2), (1.3), but in which the setting is even more complicated. Let us suppose our body occupies the following domain, defined by an annulus (see Fig. 4)

$$\Omega := \Big\{ (x, y) \text{ such that } 1 \leq x^2 + y^2 \leq 2^2 \Big\},$$

where the outer boundary represents the contact boundary part Γ_c, which is defined by

$$\Gamma_c := \Big\{ (x, y) \text{ such that } x^2 + y^2 = 2^2 \Big\},$$

and where the unit outward normal to the boundary Γ_c is given by

$$\boldsymbol{n}\,(x, y) := \left(\frac{x}{\sqrt{x^2 + y^2}}, \frac{y}{\sqrt{x^2 + y^2}} \right) \quad \forall\, x, y \in \Gamma_c.$$

Otherwise, the inner boundary part where a Dirichlet boundary condition is applied is defined by

$$\Gamma_d := \Big\{ (x, y) \text{ such that } x^2 + y^2 = 1 \Big\}.$$

Finally, we consider the following two-dimensional exact solution of problem (1.1)–(1.5):

$$\boldsymbol{u}\,(x, y) := \left(2x^3y^2,\ 4x^2y\left(2^2 - \tfrac{3}{2}x^2 - y^2\right)\right). \tag{5.1}$$

Thus, the right-hand sides of the unilateral contact problem (1.1)–(1.5) are fixed accordingly. Let us note that the normal component $\boldsymbol{u} \cdot \boldsymbol{n}$ satisfies

$$\boldsymbol{u}\,(x, y) \cdot \boldsymbol{n}\,(x, y) := -\frac{4x^2y^2\left(x^2 + y^2 - 2^2\right)}{\sqrt{x^2 + y^2}} = 0 \quad \forall\, x, y \in \Gamma_c.$$

Otherwise, the normal component of the surface force $\boldsymbol{\sigma}(\boldsymbol{u})\boldsymbol{n}$ satisfies $\boldsymbol{\sigma}_n(\boldsymbol{u}) < 0$. Hence, on the boundary part Γ_c, the following unilateral boundary condition is satisfied:

$$\boldsymbol{u} \cdot \boldsymbol{n} = 0, \quad \boldsymbol{\sigma}_n(\boldsymbol{u}) < 0, \quad (\boldsymbol{u} \cdot \boldsymbol{n})\, \boldsymbol{\sigma}_n(\boldsymbol{u}) = 0.$$

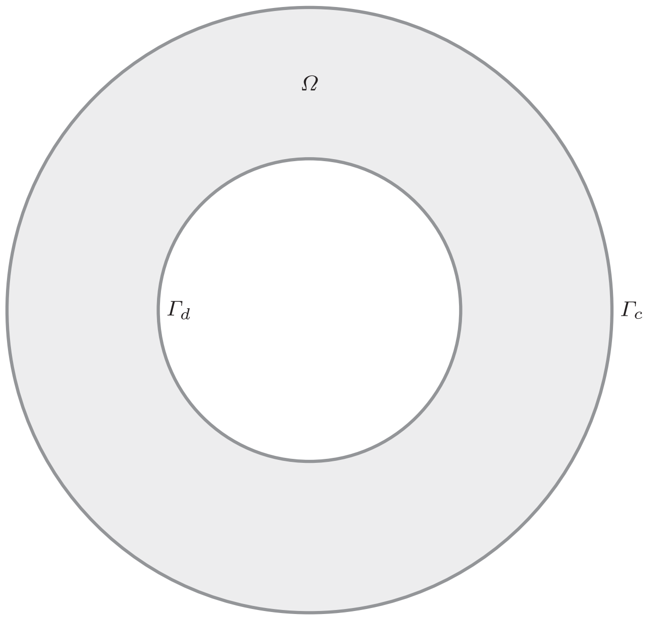

FIG. 4. Smooth domain $\Omega \subset \mathbb{R}^2$ with a curved boundary.

However, there exists a frictional effect on the contact surface Γ_c, since the tangential component of the surface force $\boldsymbol{\sigma}(\boldsymbol{u})\boldsymbol{n}$ does not vanish:

$$\boldsymbol{\sigma}_t(\boldsymbol{u}) \neq 0.$$

We consider a body with a Young modulus equal to $E := 70\,\text{GPa}$ and a Poisson's ratio defined by $\eta := 0.35$. Thus, the associated Lamé coefficients are given by $\lambda := 60.49\,\text{GPa}$ and $\mu := 25.93\,\text{GPa}$.

According to Remark 4.3, we define the penalty parameter ε as a function of the mesh size,

$$\varepsilon(h) := ch^{\theta},$$

where $c > 0$ is a constant and $\theta > 0$ is a fixed power. For the computations, we work on a series of six meshes by starting from a regular mesh and obtaining other meshes by subdividing element sides into two new sides. On each mesh, we solve the finite element problem (3.11) and compare the obtained numerical solution to the manufactured analytical solution defined in (5.1) by computing the errors $\|\nabla \boldsymbol{u} - \nabla \overline{\boldsymbol{u}}_{\varepsilon h}\|_{0,\Omega}$ and $h^{\frac{\theta}{2}}\|\boldsymbol{\sigma}_n(\boldsymbol{u}) + \frac{1}{\varepsilon}[\overline{\boldsymbol{u}}_{\varepsilon h} \cdot \boldsymbol{n}]^+\|_{0,\Gamma_c}$. Hence, convergence orders, with respect to the discrete parameter h, can be easily computed. Our goal is to verify that the numerical convergence orders are at least as good as the theoretical orders established in the previous section.

For the first round of computations, we set $\theta = 1$ and make a series of calculations for different constants, namely $c = 1$, $c = 10^{-1}$, $c = 10^{-2}$ and $c = 10^{-3}$. In addition to the theoretical expected rate which is represented in the light green curve of slope 1, we present in Fig. 5 the errors $\|\nabla \boldsymbol{u} - \nabla \overline{\boldsymbol{u}}_{\varepsilon h}\|_{0,\Omega}$

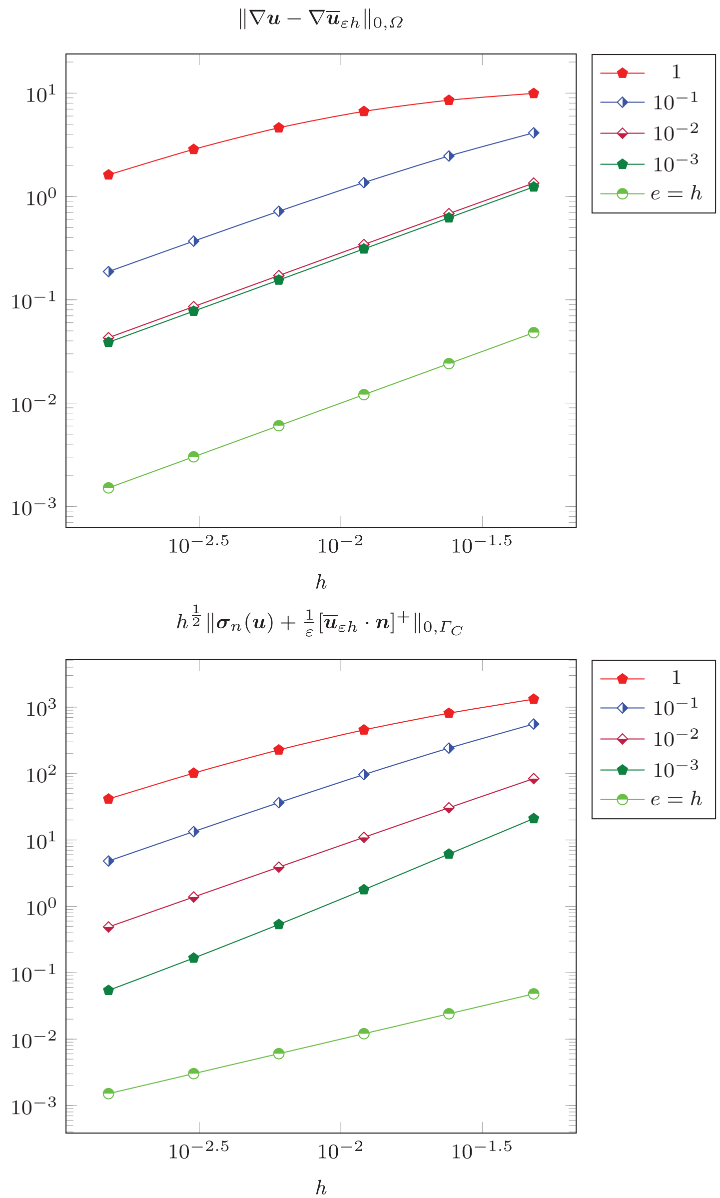

FIG. 5. Errors $\|\nabla \boldsymbol{u} - \nabla \overline{\boldsymbol{u}}_{\varepsilon h}\|_{0,\Omega}$ and $h^{\frac{1}{2}}\|\boldsymbol{\sigma}_n(\boldsymbol{u}) + \frac{1}{\varepsilon}[\overline{\boldsymbol{u}}_{\varepsilon h} \cdot \boldsymbol{n}]^+\|_{0,\Gamma_C}$ obtained by choosing the penalty parameter as $\varepsilon(h) = ch$ for different values of the constant c, namely $c = 1, 10^{-1}, 10^{-2}, 10^{-3}$ and the slope $e = h$ representing the expected convergence rate.

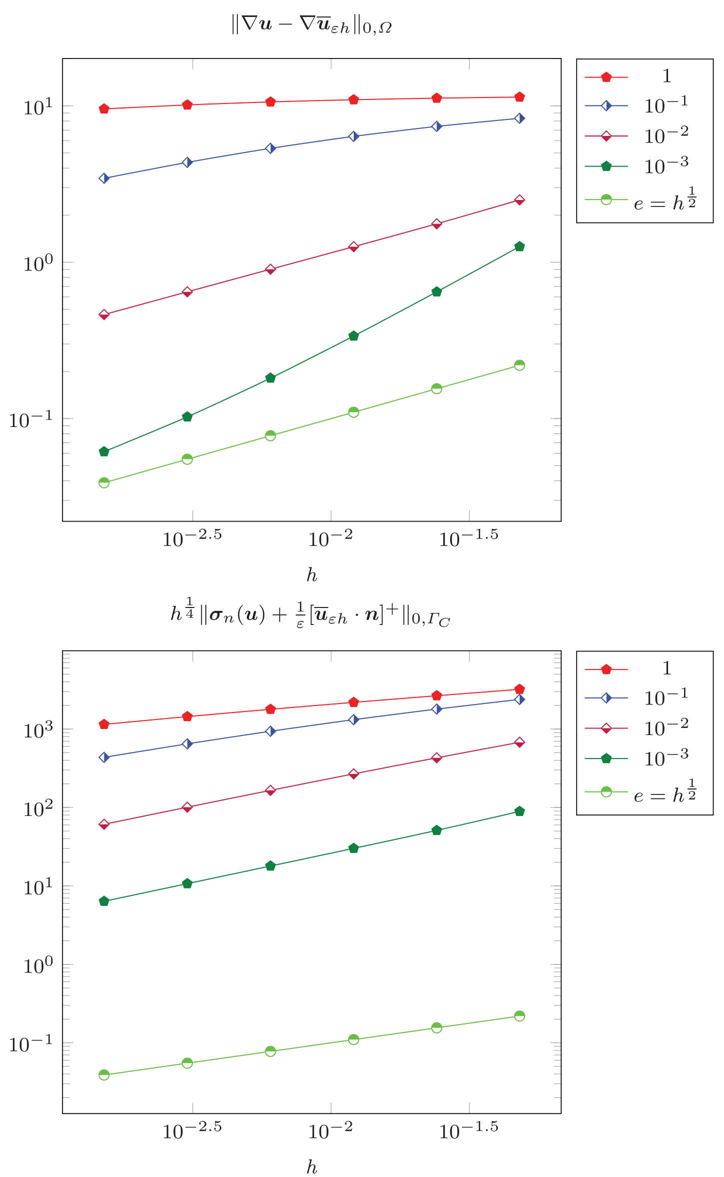

FIG. 6. Errors $\|\nabla \boldsymbol{u} - \nabla \overline{\boldsymbol{u}}_{\varepsilon h}\|_{0,\Omega}$ and $h^{\frac{1}{4}}\|\boldsymbol{\sigma}_n(\boldsymbol{u}) + \frac{1}{\varepsilon}[\overline{\boldsymbol{u}}_{\varepsilon h} \cdot \boldsymbol{n}]^{+}\|_{0,\Gamma_C}$ obtained by choosing the penalty parameter as $\varepsilon(h) = ch^{\frac{1}{2}}$ for different values of the constant c, namely $c = 1, 10^{-1}, 10^{-2}, 10^{-3}$, and the slope $e = h^{\frac{1}{2}}$ representing the expected convergence rate.

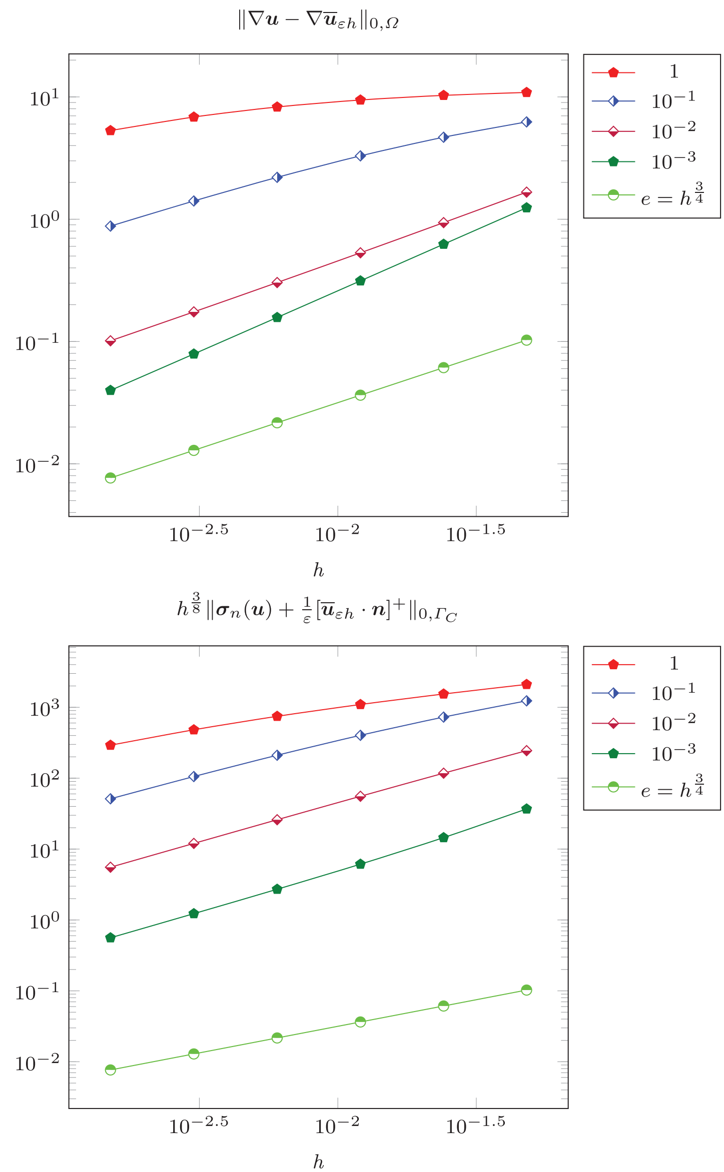

FIG. 7. Errors $\|\nabla \boldsymbol{u} - \nabla \overline{\boldsymbol{u}}_{\varepsilon h}\|_{0,\Omega}$ and $h^{\frac{3}{8}}\|\boldsymbol{\sigma}_n(\boldsymbol{u}) + \frac{1}{\varepsilon}[\overline{\boldsymbol{u}}_{\varepsilon h} \cdot \boldsymbol{n}]^{+}\|_{0,\Gamma_C}$ obtained by choosing the penalty parameter as $\varepsilon(h) = ch^{\frac{3}{4}}$ for different values of the constant c, namely $c = 1, 10^{-1}, 10^{-2}, 10^{-3}$ and the slope $e = h^{\frac{3}{4}}$ representing the expected convergence rate.

and $h^{\frac{1}{2}}\|\boldsymbol{\sigma}_n(\boldsymbol{u}) + \frac{1}{\varepsilon}[\overline{\boldsymbol{u}}_{\varepsilon h}\cdot\boldsymbol{n}]^+\|_{0,\Gamma_c}$. We thus observe a linear convergence rate corresponding to the theoretical rate predicted in estimate (4.30) (noting that our manufactured solution in (5.1) is sufficiently regular and the parameter ν is equal to $\frac{1}{2}$). On the other hand, we can note that our results are asymptotic knowing that the same convergence rate is observed for different constants (even if the choice of the best constant c is a difficult problem, which is beyond the scope of this paper).

For a second round of computations, we first take $\theta = \frac{1}{2}$ and make the same series of computations as the previous one by taking $c = 1$, $c = 10^{-1}$, $c = 10^{-2}$ and $c = 10^{-3}$. We thus present in Fig. 6 the errors $\|\nabla\boldsymbol{u} - \nabla\overline{\boldsymbol{u}}_{\varepsilon h}\|_{0,\Omega}$ and $h^{\frac{1}{4}}\|\boldsymbol{\sigma}_n(\boldsymbol{u}) + \frac{1}{\varepsilon}[\overline{\boldsymbol{u}}_{\varepsilon h}\cdot\boldsymbol{n}]^+\|_{0,\Gamma_c}$, to which we added a line in light green of slope $\frac{1}{2}$ representing the theoretical rate. We then took $\theta = \frac{3}{4}$ and repeated the same computations, presenting in Fig. 7 the errors $\|\nabla\boldsymbol{u} - \nabla\overline{\boldsymbol{u}}_{\varepsilon h}\|_{0,\Omega}$ and $h^{\frac{3}{8}}\|\boldsymbol{\sigma}_n(\boldsymbol{u}) + \frac{1}{\varepsilon}[\overline{\boldsymbol{u}}_{\varepsilon h}\cdot\boldsymbol{n}]^+\|_{0,\Gamma_c}$. We observe in Figs. 6 and 7 convergence rates that are at least as good as the theoretical ones of $\mathcal{O}(h^{\frac{1}{2}})$ and $\mathcal{O}(h^{\frac{3}{4}})$, respectively. These results thus correspond to the predicted rate in the estimate (4.31).

References

Adams, R. A. & Fournier, J. J. F. (2003) *Sobolev Spaces*. Amsterdam: Elsevier/Academic Press.

Bänsch, E. & Deckelnick, K. (1999) Optimal error estimates for the Stokes and Navier-Stokes equations with slip-boundary condition. *Math. Model. Numer. Anal.*, **33**, 923–938.

Barrett, J. W. & Elliott C. M. (1986) Finite element approximation of the Dirichlet problem using the boundary penalty method. *Numer. Math.*, **49**, 343–366.

Bernardi, C. (1989) Optimal finite element interpolation on curved domains. *SIAM J. Numer. Anal.*, **26**, 212–234.

Bernardi, C., Maday, Y. & Patera, A. T. (1994) A new nonconforming approach to domain decomposition: the mortar element method. *Nonlinear Partial Differential Equations and their Applications.* Collège de France Seminars, vol. XI, Pitman Res. Notes Math. Ser., vol. 299, Harlow: Longman Sci. Tech, pp. 13–51,

Bjorstad, P. E. & Widlund O. B. (1986) Iterative methods for the solution of elliptic problems on regions partitioned into substructures. *SIAM J. Numer. Anal.*, **23**, 1097–1120.

Bramble J. H., Pasciak, J. E. & Steinbach, O. (2001) On the stability of the L^2 projection in $H^1(\Omega)$. *Math. Comput.*, **71**, 147–156.

Brenner, C. S. & Scott, R. L. (2002) *The Mathematical Theory of Finite Element Methods*. New York: Springer.

Brezis, H. (1968) Equations et inéquations non linéaires dans les espaces vectoriels en dualité. *Ann. Inst. Fourier*, **18**, 115–175.

Chernov, A., Maischak, M. & Stephan, E. P. (2007) A priori error estimates for *hp* penalty BEM for contact problems in elasticity. *Comput. Methods Appl. Mech. Engrg.*, **196**, 3871–3880.

Chouly, F. & Hild, P. (2013) On convergence of the penalty method for unilateral contact problems. *Appl. Numer. Math.*, **65**, 27–40.

Dione, I. (2016) Towards optimal finite element error estimates for the penalized Dirichlet problem in a domain with curved boundary. *Comput. Math. Appl.*, **71**, 76–84.

Dione, I. (2018) Optimal convergence analysis of the unilateral contact problem with and without Tresca friction conditions by the penalty method (submitted to *Journal of Mathematical Analysis and Applications*).

Dione, I. & Urquiza, J. M. (2015) Penalty: finite element approximation of Stokes equations with slip boundary conditions. *Numer. Math.*, **129**, 587–610.

Dominguez, V. & Sayas, F. J. (2003) Stability of discrete liftings. *C. R. Math. Acad. Sci. Paris*, **337**, 805–808.

Drouet, G. & Hild, P. (2015) Optimal convergence for discrete variational inequality modelling Signorini contact in 2D and 3D without additional assumptions on the unknown contact set. *SIAM J. Numer. Anal.*, **53**, 1488–1507.

Kikuchi, N. & Oden, J. T. (1988) *Contact Problems in Elasticity: A Study of Variational Inequalities and Finite Element Methods*. Society for Industrial and Applied Mathematics, Philadelphia: PA.

Kikuchi, N. & Song, Y. J. (1981) Penalty-finite element approximation of a class of unilateral problems in linear elasticity. *Quart. Appl. Math.* **XXXIX**, 1–22

Lenoir, M. (1986) Optimal isoparametric finite elements and errors estimates for domains involving curved boundaries. *SIAM J. Numer. Anal.*, **23**, 562–580.

Li, Z. C. (1998) *Combined Methods for Elliptic Equations with Singularities, Interfaces and Infinities*. Dordrecht: Kluwer Academic Publishers.

Lions, J.-L. (1969) *Quelques méthodes de résolution des problèmes aux limites non linéaires*. Paris: Dunod.

Appendix

Let us introduce the exact triangulation $\mathcal{T}_\varepsilon$ (see Fig. A1) composed of a finite set of closed d-simplices T meshing the domain Ω, with mesh size ε:

$$\overline{\Omega} := \bigcup_{\mathsf{T}\in\mathcal{T}_\varepsilon} \mathsf{T}. \tag{A.1}$$

We introduce, as explained in Chouly & Hild (2013, proof of Lemma 3.9), the fictitious finite element space

$$\boldsymbol{V}_\varepsilon := \left\{ \boldsymbol{v}_\varepsilon \in \left(\mathscr{C}^0(\overline{\Omega})\right)^d : \boldsymbol{v}_{\varepsilon|\mathsf{T}} \in (\mathscr{P}_1(\mathsf{T}))^d \quad \forall\, \mathsf{T} \in \mathcal{T}_\varepsilon,\, \boldsymbol{v}_\varepsilon = 0 \text{ on } \Gamma_d \right\}.$$

We also consider the fictitious discrete trace space as defined in Chouly & Hild (2013)

$$W_\varepsilon(\Gamma_c) := \left\{ \mu_\varepsilon \in \mathscr{C}^0(\overline{\Gamma_c}) : \exists\, \boldsymbol{v}_\varepsilon \in \boldsymbol{V}_\varepsilon,\, \boldsymbol{v}_\varepsilon \cdot \boldsymbol{n} = \mu_\varepsilon \right\}.$$

We thus introduce, as we have also detailed in the finite element Section 3, the $L^2(\Gamma_c)$-projection $\mathcal{P}^\varepsilon : L^2(\Gamma_c) \longrightarrow W_\varepsilon(\Gamma_c)$ satisfying estimates (3.7) and (3.8). Supposing that the mesh on the contact boundary Γ_c is quasi-uniform, we also introduce the discrete lifting operator $\mathcal{R}^\varepsilon : W_\varepsilon(\Gamma_c) \longrightarrow \boldsymbol{V}_\varepsilon$ in the sense defined in Chouly & Hild (2013, Lemma 3.8) such that

$$\mathcal{R}^\varepsilon(\mu_\varepsilon)_{|\Gamma_c} \cdot \boldsymbol{n} = \mu_\varepsilon, \quad \left\|\mathcal{R}^\varepsilon(\mu_\varepsilon)\right\|_{1,\Omega} \leq c\|\mu_\varepsilon\|_{\frac{1}{2},\Gamma_c} \quad \forall\, \mu_\varepsilon \in W_\varepsilon(\Gamma_c). \tag{A.2}$$

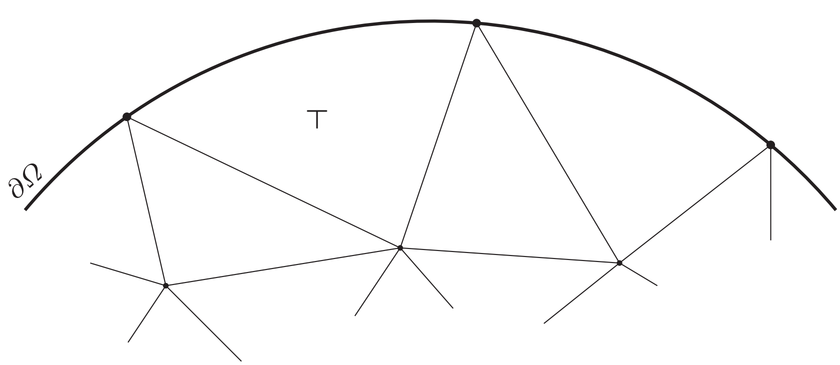

Fig. A1. Exact triangulation $\mathcal{T}_\varepsilon$ of the curved and smooth boundary $\partial\Omega$ of the domain $\Omega \subset \mathbb{R}^2$.

LEMMA A1 Let $\Omega \subset \mathbb{R}^d, d = 2, 3$ be a bounded domain with a curved and smooth boundary ($\partial\Omega \in \mathscr{C}^{1,1}$), $\boldsymbol{u}$ and $\boldsymbol{u}_\varepsilon$ the solutions of problems (2.1) and (2.4), respectively. If $\boldsymbol{u}$ belongs to $\boldsymbol{H}^{\frac{3}{2}+\nu}(\Omega)$, where $\nu \in \left(0, \frac{1}{2}\right]$, then we obtain the *a priori* estimate

$$\left\|\boldsymbol{\sigma}_n(\boldsymbol{u}) + \tfrac{1}{\varepsilon}\left[\boldsymbol{u}_\varepsilon \cdot \boldsymbol{n} - \ell\right]^+\right\|_{-\nu,\Gamma_c} \leq c\left(\varepsilon^\nu \left\|\boldsymbol{\sigma}_n(\boldsymbol{u}) + \tfrac{1}{\varepsilon}\left[\boldsymbol{u}_\varepsilon \cdot \boldsymbol{n} - \ell\right]^+\right\|_{0,\Gamma_c} + \varepsilon^{\nu-\frac{1}{2}}\|\boldsymbol{u} - \boldsymbol{u}_\varepsilon\|_{1,\Omega}\right), \quad \text{(A.3)}$$

where $c > 0$ is a constant, independent of the parameter ε and the solution $\boldsymbol{u}$.

Proof. By definition of the norm $\|\cdot\|_{-\nu,\Gamma_c}$ and using the projection operator $\mathcal{P}^\varepsilon$, the Cauchy–Schwarz inequality and the interpolation estimate (3.8) applied to the context here, we obtain

$$\begin{aligned}
&\left\|\boldsymbol{\sigma}_n(\boldsymbol{u}) + \frac{1}{\varepsilon}\left[\boldsymbol{u}_\varepsilon \cdot \boldsymbol{n} - \ell\right]^+\right\|_{-\nu,\Gamma_c} \\
&\quad := \sup_{v\in H^\nu(\Gamma_c)} \frac{\left\langle \boldsymbol{\sigma}_n(\boldsymbol{u}) + \frac{1}{\varepsilon}\left[\boldsymbol{u}_\varepsilon \cdot \boldsymbol{n} - \ell\right]^+, v\right\rangle_{\Gamma_c}}{\|v\|_{\nu,\Gamma_c}} \\
&\quad = \sup_{v\in H^\nu(\Gamma_c)} \frac{\left\langle \boldsymbol{\sigma}_n(\boldsymbol{u}) + \frac{1}{\varepsilon}\left[\boldsymbol{u}_\varepsilon \cdot \boldsymbol{n} - \ell\right]^+, v - \mathcal{P}^\varepsilon(v)\right\rangle_{\Gamma_c}}{\|v\|_{\nu,\Gamma_c}} \\
&\qquad + \sup_{v\in H^\nu(\Gamma_c)} \frac{\left\langle \boldsymbol{\sigma}_n(\boldsymbol{u}) + \frac{1}{\varepsilon}\left[\boldsymbol{u}_\varepsilon \cdot \boldsymbol{n} - \ell\right]^+, \mathcal{P}^\varepsilon(v)\right\rangle_{\Gamma_c}}{\|v\|_{\nu,\Gamma_c}} \\
&\quad \leq \left\|\boldsymbol{\sigma}_n(\boldsymbol{u}) + \frac{1}{\varepsilon}\left[\boldsymbol{u}_\varepsilon \cdot \boldsymbol{n} - \ell\right]^+\right\|_{0,\Gamma_c} \sup_{v\in H^\nu(\Gamma_c)} \frac{\|v - \mathcal{P}^\varepsilon(v)\|_{0,\Gamma_c}}{\|v\|_{\nu,\Gamma_c}} \\
&\qquad + \sup_{v\in H^\nu(\Gamma_c)} \frac{\left\langle \boldsymbol{\sigma}_n(\boldsymbol{u}) + \frac{1}{\varepsilon}\left[\boldsymbol{u}_\varepsilon \cdot \boldsymbol{n} - \ell\right]^+, \mathcal{P}^\varepsilon(v)\right\rangle_{\Gamma_c}}{\|v\|_{\nu,\Gamma_c}} \\
&\quad \leq c\varepsilon^\nu \left\|\boldsymbol{\sigma}_n(\boldsymbol{u}) + \frac{1}{\varepsilon}\left[\boldsymbol{u}_\varepsilon \cdot \boldsymbol{n} - \ell\right]^+\right\|_{0,\Gamma_c} + \sup_{v\in H^\nu(\Gamma_c)} \frac{\left\langle \boldsymbol{\sigma}_n(\boldsymbol{u}) + \frac{1}{\varepsilon}\left[\boldsymbol{u}_\varepsilon \cdot \boldsymbol{n} - \ell\right]^+, \mathcal{P}^\varepsilon(v)\right\rangle_{\Gamma_c}}{\|v\|_{\nu,\Gamma_c}}, \qquad \text{(A.4)}
\end{aligned}$$

where $\langle\cdot,\cdot\rangle_{\Gamma_c}$ represents the duality product between $H^\nu(\Gamma_c)$ and its topological dual. Otherwise, using the discrete lifting operator (A.2), equations (2.4), (2.9) and the boundedness (2.3) of $A(\cdot,\cdot)$, we obtain

$$\begin{aligned}
\left\langle \boldsymbol{\sigma}_n(\boldsymbol{u}) + \tfrac{1}{\varepsilon}\left[\boldsymbol{u}_\varepsilon \cdot \boldsymbol{n} - \ell\right]^+, \mathcal{P}^\varepsilon(v)\right\rangle_{\Gamma_c} &= \left\langle \boldsymbol{\sigma}_n(\boldsymbol{u}) + \tfrac{1}{\varepsilon}\left[\boldsymbol{u}_\varepsilon \cdot \boldsymbol{n} - \ell\right]^+, \mathcal{R}^\varepsilon\left(\mathcal{P}^\varepsilon(v)\right) \cdot \boldsymbol{n}\right\rangle_{\Gamma_c} \\
&= A\left(\boldsymbol{u} - \boldsymbol{u}_\varepsilon, \mathcal{R}^\varepsilon\left(\mathcal{P}^\varepsilon(v)\right)\right) \\
&= M\|\boldsymbol{u} - \boldsymbol{u}_\varepsilon\|_{1,\Omega}\left\|\mathcal{R}^\varepsilon\left(\mathcal{P}^\varepsilon(v)\right)\right\|_{1,\Omega} \\
&\leq cM\|\boldsymbol{u} - \boldsymbol{u}_\varepsilon\|_{1,\Omega}\left\|\mathcal{P}^\varepsilon(v)\right\|_{\frac{1}{2},\Gamma_c}. \qquad \text{(A.5)}
\end{aligned}$$

Since we have supposed a quasi-uniform mesh property on Γ_c, then the following inverse property holds:

$$\left\|\mathcal{P}^{\varepsilon}(v)\right\|_{\frac{1}{2},\Gamma_c} \leq c\varepsilon^{\nu-\frac{1}{2}}\left\|\mathcal{P}^{\varepsilon}(v)\right\|_{\nu,\Gamma_c}. \tag{A.6}$$

Combining estimates (3.7) and (A.6) by taking into account (A.5), we bound the last term in the right-hand side of (A.4) as

$$\begin{aligned}\sup_{v\in H^{\nu}(\Gamma_c)} \frac{\left\langle \boldsymbol{\sigma}_n(\boldsymbol{u}) + \frac{1}{\varepsilon}\left[\boldsymbol{u}_\varepsilon\cdot\boldsymbol{n}-\ell\right]^+, \mathcal{P}^{\varepsilon}(v)\right\rangle_{\Gamma_c}}{\|v\|_{\nu,\Gamma_c}} &\leq \sup_{v\in H^{\nu}(\Gamma_c)} \frac{cM\varepsilon^{\nu-\frac{1}{2}}\|\boldsymbol{u}-\boldsymbol{u}_\varepsilon\|_{1,\Omega}\|v\|_{\nu,\Gamma_c}}{\|v\|_{\nu,\Gamma_c}}\\ &= cM\varepsilon^{\nu-\frac{1}{2}}\|\boldsymbol{u}-\boldsymbol{u}_\varepsilon\|_{1,\Omega}.\end{aligned} \tag{A.7}$$

Thus, taking again estimate (A.4) by using inequality (A.7), we obtain the desired estimate (A.3).

□