SIAM J. SCI. COMPUT.
Vol. 42, No. 2, pp. B425–B446

ON NITSCHE'S METHOD FOR ELASTIC CONTACT PROBLEMS*

TOM GUSTAFSSON[†], ROLF STENBERG[†], AND JUHA VIDEMAN[‡]

Abstract. We show quasi-optimality and a posteriori error estimates for the frictionless contact problem between two elastic bodies with a zero-gap function. The analysis is based on interpreting Nitsche's method as a stabilized finite element method for which the error estimates can be obtained with minimal regularity assumptions and without the saturation assumption. We present three different Nitsche's mortaring techniques for the contact boundary, each corresponding to a different stabilizing term. Our numerical experiments show the robustness of Nitsche's method and corroborate the efficiency of the a posteriori error estimators.

Key words. Nitsche's method, elastic contact, variational inequality

AMS subject classification. 65N30

DOI. 10.1137/19M1246869

1. Introduction. In this paper, we analyze the Nitsche method for elastic contact problems. Over the last decade, this method has been studied by a number of authors (see, e.g., [9, 6, 7, 10]) and shown to be a robust and efficient method. The advantages are an easy implementation based on the displacement variables only and, when compared to mixed methods with Lagrange multipliers, the absence of an "inf-sup" stability condition which renders a symmetric positive definite system instead of one with a saddle point structure.

From a theoretical point of view, the previously mentioned works suffer from two shortcomings. First, for the problem posed in H^1, the solution is typically assumed to be in H^s with $s > 3/2$. Second, the a posteriori error analyses are often based on a nonrigorous saturation assumption.

We have addressed these issues in our recent articles; cf. [12, 13]. Our approach dates back to [23] where different ways to enforce weakly the Dirichlet boundary conditions were discussed in the context of the so-called stabilized mixed methods [2, 3] wherein the bilinear form of the original mixed finite element method is augmented with a properly weighted residual term to ensure stability. In [23], it was shown that the local elimination of the Lagrange multiplier leads essentially to a method introduced by Nitsche in the early age of the finite element analysis [22]. Since Nitsche's method is straightforward both to analyze (under the additional smoothness assumption) and to implement, we started to advocate it, in particular for contact problems; cf. [24, 4].

What we have realized recently is that one should take full advantage of the relation between Nitsche's and stabilized method when analyzing the former. In fact, we were able to get rid of both the smoothness and the saturation assumption for the membrane obstacle problem in [12]. In this paper, we will continue on this path

*Submitted to the journal's Computational Methods in Science and Engineering section February 25, 2019; accepted for publication (in revised form) January 6, 2020; published electronically March 23, 2020.
https://doi.org/10.1137/19M1246869
Funding: This work was supported by the Portuguese government through FCT (Fundacao para a Ciencia e a Tecnologia), I.P., under the projects PTDC/MAT-PUR/28686/2017 and UTAP-EXPL/MAT/0017/2017.
[†]Department of Mathematics and Systems Analysis, Aalto University, 00076 Aalto, Finland (tom.gustafsson@alumni.aalto.fi, rolf.stenberg@aalto.fi).
[‡]CAMGSD/Departamento de Matemática, Universidade de Lisboa, Universidade de Lisboa, 1049-001 Lisbon, Portugal (jvideman@math.tecnico.ulisboa.pt).

and perform an error analysis, both quasi-optimality and a posteriori, for a simplified two-body contact problem without friction. Besides the theoretical improvements, we present three versions of the Nitsche's method where the changes in the material parameters between the bodies are taken into account. The simplest is a typical "master-slave" approach where the contact surface of the stiffer body is chosen as the master part and the slave surface is then mortared by the Nitsche's technique. In the two other variants, the material parameters appear as weights in the Nitsche formulation so that the methods decide by themselves which part is the master and which is the slave. In order to simplify the notation, analysis, and implementation of the adaptive methods, we assume that the elastic bodies are initially in full contact (see, e.g., [17]) and leave the case with a nonvanishing initial gap between the elastic bodies for a future work.

Although our analysis is built upon our earlier works (cf. [12, 14]), we will present proofs of all the main theorems. We also note that the elastic contact problem literature is vast, and therefore we only refer to the review paper [27], and to all the references therein, for the analysis and application of finite element methods arising from mixed formulations and to [20, 8], and to all the references therein, for the a posteriori error analyses of contact problems. We end the paper by presenting results of our computational experiments.

2. The contact problem. Let $\Omega_i \subset \mathbb{R}^d$, $i = 1, 2$, $d \in \{2, 3\}$, denote two elastic bodies in their reference configuration, and assume that the bodies are initially in contact. Moreover, assume that Ω_i are polygonal (polyhedral) domains, and denote by $\Gamma = \partial\Omega_1 \cap \partial\Omega_2$ their common boundary. The boundary $\partial\Omega_i$ is split into three disjoint sets $\Gamma_{D,i}, \Gamma_{N,i}$, and $\Gamma_{C,i}$, with $\Gamma_{D,i}$ denoting the part where homogeneous Dirichlet data is given, $\Gamma_{N,i}$ the part with a Neumann boundary condition, and $\Gamma_{C,i}$ the part where contact can occur; see Figure 1.

Letting $\boldsymbol{u}_i : \Omega_i \to \mathbb{R}^d$, $i = 1, 2$, be the displacement of the body Ω_i, the infinitesimal strain tensor is defined as

$$\boldsymbol{\varepsilon}(\boldsymbol{u}_i) = \frac{1}{2}\Big(\nabla \boldsymbol{u}_i + (\nabla \boldsymbol{u}_i)^T\Big). \tag{2.1}$$

We assume homogenous isotropic bodies and a plain strain problem in the two dimensional case. The stress tensor is thus given by

$$\boldsymbol{\sigma}_i(\boldsymbol{u}_i) = 2\mu_i\, \boldsymbol{\varepsilon}(\boldsymbol{u}_i) + \lambda_i \operatorname{tr} \boldsymbol{\varepsilon}(\boldsymbol{u}_i)\boldsymbol{I}, \tag{2.2}$$

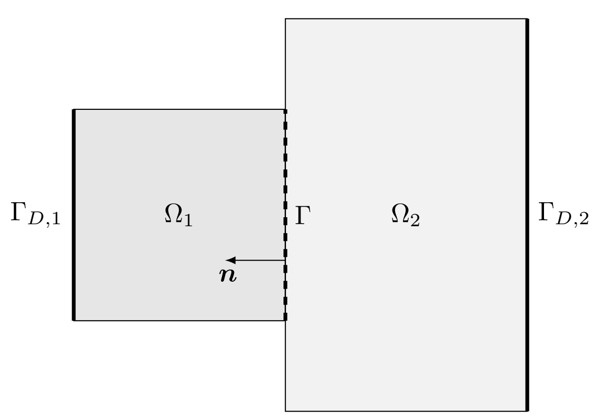


FIG. 1. *Notation for the elastic contact problem.*

where $\mu_i > 0$ is the shear modulus and λ_i the second Lamé parameter of the body Ω_i and $\boldsymbol{I}$ denotes the d-dimensional identity tensor. We will exclude the possibility that the materials are nearly incompressible and hence it holds that $\lambda_i \lesssim \mu_i$. (For nearly incompressible materials the standard approach of reformulating the problem in mixed form [5] should be used.)

By $\boldsymbol{n}_i \in \mathbb{R}^d$ we denote the outward unit normal to $\partial\Omega_i$, and define $\boldsymbol{n} = \boldsymbol{n}_1 = -\boldsymbol{n}_2$. In what follows, $\boldsymbol{t}$ denotes any unit vector that satisfies $\boldsymbol{n} \cdot \boldsymbol{t} = 0$.

We decompose the traction vector on $\partial\Omega_i$, $\boldsymbol{\sigma}_i(\boldsymbol{u}_i)\boldsymbol{n}_i$, into its normal and tangential parts, viz.,

$$\boldsymbol{\sigma}_i(\boldsymbol{u}_i)\boldsymbol{n}_i = \boldsymbol{\sigma}_{i,n}(\boldsymbol{u}_i) + \boldsymbol{\sigma}_{i,t}(\boldsymbol{u}_i). \tag{2.3}$$

For the scalar normal tractions we use the sign convention

$$\sigma_{1,n}(\boldsymbol{u}_1) = \boldsymbol{\sigma}_{1,n}(\boldsymbol{u}_1) \cdot \boldsymbol{n}_1 \tag{2.4}$$

and

$$\sigma_{2,n}(\boldsymbol{u}_2) = -\boldsymbol{\sigma}_{2,n}(\boldsymbol{u}_2) \cdot \boldsymbol{n}_2 \tag{2.5}$$

and note that on Γ these tractions are either both zero or continuous and compressive, i.e., it holds that

$$\sigma_{1,n}(\boldsymbol{u}_1) = \sigma_{2,n}(\boldsymbol{u}_2), \quad \sigma_{i,n}(\boldsymbol{u}_i) \le 0,\ i = 1, 2. \tag{2.6}$$

The physical nonpenetration constraint on Γ reads as

$$\boldsymbol{u}_1 \cdot \boldsymbol{n}_1 + \boldsymbol{u}_2 \cdot \boldsymbol{n}_2 \le 0, \tag{2.7}$$

which, defining

$$u_n = -(\boldsymbol{u}_1 \cdot \boldsymbol{n}_1 + \boldsymbol{u}_2 \cdot \boldsymbol{n}_2), \tag{2.8}$$

can be written as

$$[\![u_n]\!] \ge 0, \tag{2.9}$$

where $[\![\cdot]\!]$ denotes the jump over Γ.

We thus have the following problem.

PROBLEM 1 (strong formulation). *Find* $\boldsymbol{u}_i : \Omega_i \to \mathbb{R}^d$, $i = 1, 2$, $d \in \{2, 3\}$, *such that*

$$\begin{aligned}
-\mathbf{div}\,\boldsymbol{\sigma}_i(\boldsymbol{u}_i) &= \boldsymbol{f}_i && \text{in } \Omega_i,\\
\boldsymbol{u}_i &= \boldsymbol{0} && \text{on } \Gamma_{D,i},\\
\boldsymbol{\sigma}_i(\boldsymbol{u}_i)\boldsymbol{n}_i &= \boldsymbol{0} && \text{on } \Gamma_{N,i},\\
\boldsymbol{\sigma}_{i,t}(\boldsymbol{u}_i) &= \boldsymbol{0} && \text{on } \Gamma,\\
\sigma_{1,n}(\boldsymbol{u}_1) - \sigma_{2,n}(\boldsymbol{u}_2) &= 0 && \text{on } \Gamma,\\
[\![u_n]\!] &\ge 0 && \text{on } \Gamma,\\
\sigma_{i,n}(\boldsymbol{u}_i) &\le 0 && \text{on } \Gamma,\\
[\![u_n]\!]\,\sigma_{i,n}(\boldsymbol{u}_i) &= 0 && \text{on } \Gamma,
\end{aligned} \tag{2.10}$$

where $\boldsymbol{f}_i \in [L^2(\Omega_i)]^d$ *denotes the volume force on* Ω_i.

Letting $\lambda = -\sigma_{1,n}(\boldsymbol{u}_1) = -\sigma_{2,n}(\boldsymbol{u}_2)$ denote a Lagrange multiplier associated with the contact constraint, we obtain an equivalent mixed formulation in which the normal traction on the contact surface is an independent unknown.

Problem 2 (mixed formulation). *Find* $\boldsymbol{u}_i : \Omega_i \to \mathbb{R}^d$, $i = 1, 2$, $d \in \{2, 3\}$, *and* $\lambda : \Gamma \to \mathbb{R}$, *such that*

$$\begin{aligned} -\mathbf{div}\,\boldsymbol{\sigma}_i(\boldsymbol{u}_i) &= \boldsymbol{f}_i && \text{in } \Omega_i,\\ \boldsymbol{u}_i &= \boldsymbol{0} && \text{on } \Gamma_{D,i},\\ \boldsymbol{\sigma}_i(\boldsymbol{u}_i)\boldsymbol{n}_i &= \boldsymbol{0} && \text{on } \Gamma_{N,i},\\ \boldsymbol{\sigma}_{i,t}(\boldsymbol{u}_i) &= \boldsymbol{0} && \text{on } \Gamma,\\ \lambda + \sigma_{1,n}(\boldsymbol{u}_1) &= 0, && \text{on } \Gamma,\\ \lambda + \sigma_{2,n}(\boldsymbol{u}_2) &= 0, && \text{on } \Gamma,\\ [\![u_n]\!] &\geq 0 && \text{on } \Gamma,\\ \lambda &\geq 0 && \text{on } \Gamma,\\ [\![u_n]\!]\,\lambda &= 0 && \text{on } \Gamma. \end{aligned} \tag{2.11}$$

To present a variational formulation for Problem 1, we introduce function spaces for the displacements

$$\boldsymbol{V}_i = \{\boldsymbol{w}_i \in [H^1(\Omega_i)]^d : \boldsymbol{w}_i|_{\Gamma_{D,i}} = \boldsymbol{0}\} \tag{2.12}$$

and equip them with the usual norms $\|\cdot\|_{1,\Omega_i}$. Moreover, we write $\boldsymbol{V} = \boldsymbol{V}_1 \times \boldsymbol{V}_2$ and assume that Γ is a compact subset of $\partial\Omega_i \setminus \Gamma_{D,i}$ for $i = 1, 2$. Thus the normal components of the displacement traces on the contact zone are in $H^{\frac{1}{2}}(\Gamma)$ with the intrinsic norm in $H^{\frac{1}{2}}(\Gamma)$ defined by (cf., e.g., [25])

$$\|w\|^2_{\frac{1}{2},\Gamma} = \|w\|^2_{0,\Gamma} + \int_\Gamma \int_\Gamma \frac{|w(x) - w(y)|^2}{|x - y|^d}\,\mathrm{d}x\,\mathrm{d}y. \tag{2.13}$$

The inequality constraint on Γ is imposed by the Lagrange multiplier which belongs to $H^{-\frac{1}{2}}(\Gamma)$, the topological dual of $H^{\frac{1}{2}}(\Gamma)$, i.e., $H^{-\frac{1}{2}}(\Gamma) = H^{\frac{1}{2}}(\Gamma)'$. The duality pairing is denoted by $\langle\cdot,\cdot\rangle : H^{\frac{1}{2}}(\Gamma) \times H^{-\frac{1}{2}}(\Gamma) \to \mathbb{R}$, and the norm is then

$$\|\xi\|_{-\frac{1}{2},\Gamma} = \sup_{w \in W} \frac{\langle w, \xi\rangle}{\|w\|_{\frac{1}{2},\Gamma}}. \tag{2.14}$$

Moreover, we define the positive part of $H^{-\frac{1}{2}}(\Gamma)$ as

$$\varLambda = \{\xi \in H^{-\frac{1}{2}}(\Gamma) : \langle w, \xi\rangle \geq 0 \ \ \forall w \in H^{\frac{1}{2}}(\Gamma),\ w \geq 0 \text{ a.e. on } \Gamma\} \tag{2.15}$$

and introduce the bilinear and linear forms

$$\mathcal{B}(\boldsymbol{w}, \xi; \boldsymbol{v}, \eta) = \sum_{i=1}^{2} (\boldsymbol{\sigma}_i(\boldsymbol{w}_i), \boldsymbol{\varepsilon}(\boldsymbol{v}_i))_{\Omega_i} - \langle [\![v_n]\!], \xi\rangle - \langle [\![w_n]\!], \eta\rangle \tag{2.16}$$

and

$$\mathcal{L}(\boldsymbol{v}) = \sum_{i=1}^{2} (\boldsymbol{f}_i, \boldsymbol{v}_i)_{\Omega_i}. \tag{2.17}$$

The variational problem now reads as follows.

PROBLEM 3 (weak formulation). *Find* $(\boldsymbol{u},\lambda)\in\boldsymbol{V}\times\mathit{\Lambda}$ *such that*

$$\mathcal{B}(\boldsymbol{u},\lambda;\boldsymbol{v},\eta-\lambda)\leq\mathcal{L}(\boldsymbol{v})\quad\forall(\boldsymbol{v},\eta)\in\boldsymbol{V}\times\mathit{\Lambda}. \tag{2.18}$$

We refer to [16, 15] for the derivation of weak formulation from Problem 1 and for the proof of existence and uniqueness of solutions to Problem 3.

3. Finite element method. Let the bodies $\Omega_i\subset\mathbb{R}^d$ be separately divided into sets of nonoverlapping simplices $\mathcal{C}_h^i$, $i=1,2$. The $d-1$ dimensional facets of the elements in $\mathcal{C}_h^i$ are further divided into the set of interior facets $\mathcal{E}_h^i$, the set of facets on the contact boundary $\mathcal{G}_h^i$, and the set of facets on the Neumann boundary $\mathcal{N}_h^i$. We denote by $\mathcal{G}_h^{12}$ the boundary mesh on Γ which is obtained by intersecting the facets of $\mathcal{G}_h^1$ and $\mathcal{G}_h^2$. In particular, each $E\in\mathcal{G}_h^{12}$ corresponds to a pair $(E_1,E_2)\in\mathcal{G}_h^1\times\mathcal{G}_h^2$ such that $E=E_1\cap E_2$. The finite element subspaces are

$$\boldsymbol{V}_{i,h}=\{\boldsymbol{v}_{i,h}\in\boldsymbol{V}_i:\boldsymbol{v}_{i,h}|_K\in[P_p(K)]^d\ \forall K\in\mathcal{C}_h^i\}, \tag{3.1}$$

$$\boldsymbol{V}_h=\boldsymbol{V}_{1,h}\times\boldsymbol{V}_{2,h}, \tag{3.2}$$

$$Q_h=\{\eta_h\in H^{-\frac{1}{2}}(\Gamma):\eta_h|_E\in P_p(E)\ \forall E\in\mathcal{G}_h^{12}\}, \tag{3.3}$$

where $P_p(K)$ denotes the polynomials of degree p on K. Moreover, we introduce a subset of $\mathit{\Lambda}$, denoted by $\mathit{\Lambda}_h$, as the positive part of Q_h, i.e.,

$$\mathit{\Lambda}_h=\{\eta_h\in Q_h:\eta_h\geq 0\}. \tag{3.4}$$

Now, defining a stabilized bilinear form $\mathcal{B}_h$ through

$$\mathcal{B}_h(\boldsymbol{w}_h,\xi_h;\boldsymbol{v}_h,\eta_h)=\mathcal{B}(\boldsymbol{w}_h,\xi_h;\boldsymbol{v}_h,\eta_h)-\alpha\mathcal{S}_h(\boldsymbol{w}_h,\xi_h;\boldsymbol{v}_h,\eta_h), \tag{3.5}$$

where $\alpha>0$ is a stabilization parameter and

$$\mathcal{S}_h(\boldsymbol{w}_h,\xi_h;\boldsymbol{v}_h,\eta_h)=\sum_{i=1}^2\sum_{E\in\mathcal{G}_h^i}\frac{h_E}{\mu_i}\Big(\xi_h+\sigma_{i,n}(\boldsymbol{w}_{i,h}),\eta_h+\sigma_{i,n}(\boldsymbol{v}_{i,h})\Big)_E, \tag{3.6}$$

we arrive at the following finite element formulation which is an extension of the mortar method introduced in [19, 14].

PROBLEM 4 (stabilized discrete formulation). *Find* $(\boldsymbol{u}_h,\lambda_h)\in\boldsymbol{V}_h\times\mathit{\Lambda}_h$ *such that*

$$\mathcal{B}_h(\boldsymbol{u}_h,\lambda_h;\boldsymbol{v}_h,\eta_h-\lambda_h)\leq\mathcal{L}(\boldsymbol{v}_h)\quad\forall(\boldsymbol{v}_h,\eta_h)\in\boldsymbol{V}_h\times\mathit{\Lambda}_h. \tag{3.7}$$

We will now derive an equivalent formulation wherein the Lagrange multiplier is not explicitly present. To this end, we start by defining $L^2(\Gamma)$-functions $\mathcal{h}_i$ through

$$\mathcal{h}_i|_E=h_E\quad\forall E\in\mathcal{G}_h^i,\ i=1,2, \tag{3.8}$$

and introduce the notation

$$\{\!\!\{\sigma_n(\boldsymbol{u}_h)\}\!\!\}=\frac{\mathcal{h}_1\mu_2}{\mathcal{h}_1\mu_2+\mathcal{h}_2\mu_1}\,\sigma_{1,n}(\boldsymbol{u}_{1,h})+\frac{\mathcal{h}_2\mu_1}{\mathcal{h}_1\mu_2+\mathcal{h}_2\mu_1}\,\sigma_{2,n}(\boldsymbol{u}_{2,h}), \tag{3.9}$$

i.e., a convex combination of the discrete normal tractions. Furthermore, we let

$$l_h(\boldsymbol{u}_h)=-\{\!\!\{\sigma_n(\boldsymbol{u}_h)\}\!\!\}-\beta_h\,[\![u_{h,n}]\!], \tag{3.10}$$

where

$$\beta_h = \frac{\mu_1\mu_2}{\alpha(\mathscr{h}_1\mu_2 + \mathscr{h}_2\mu_1)}. \tag{3.11}$$

Next, we will show that the discrete Lagrange multiplier λ_h can be eliminated locally (i.e., element by element). This leads to a Nitsche formulation with the displacements as sole unknowns. Choosing $\boldsymbol{v}_h = \boldsymbol{0}$ in the variational inequality (3.7) gives

$$- \langle [\![u_{h,n}]\!], \eta_h - \lambda_h \rangle - \alpha \sum_{i=1}^{2} \sum_{E \in \mathcal{G}_h^i} \frac{h_E}{\mu_i} \big(\lambda_h + \sigma_{i,n}(\boldsymbol{u}_{i,h}), \eta_h - \lambda_h\big)_E \leq 0, \tag{3.12}$$

which, in view of the notation defined above, can be written as

$$\langle \lambda_h - l_h(\boldsymbol{u}_h), \eta_h - \lambda_h \rangle \leq 0 \quad \forall \eta_h \in \Lambda_h. \tag{3.13}$$

Let then $E \in \mathcal{G}_h^{12}$ be an element on which $\lambda_h|_E > 0$, and denote by ϕ_E one of the basis functions of $Q_h|_E$. Moreover, choose a test function η_h in (3.13) in such a way that it vanishes at $\Gamma \setminus E$ and $\eta_h|_E = \lambda_h \pm \epsilon\phi_E$, with $\epsilon > 0$ chosen small enough so that $\eta_h|_E > 0$. It follows that

$$0 = \langle \lambda_h - l_h(\boldsymbol{u}_h), \phi_E \rangle = \int_E \big(\lambda_h - l_h(\boldsymbol{u}_h)\big)\phi_E \, ds, \tag{3.14}$$

and, since

$$\big(\lambda_h - l_h(\boldsymbol{u}_h)\big)|_E \in Q_h|_E, \tag{3.15}$$

we conclude that

$$\big(\lambda_h - l_h(\boldsymbol{u}_h)\big)|_E = 0. \tag{3.16}$$

This shows that

$$\lambda_h = (l_h(\boldsymbol{u}_h))_+ , \tag{3.17}$$

where $(a)_+ = \max(0, a)$ denotes the positive part of a.

The discrete contact region, defined as

$$\Gamma_c(\boldsymbol{u}_h) = \{\, \boldsymbol{x} \in \Gamma : \lambda_h(\boldsymbol{x}) > 0 \,\}, \tag{3.18}$$

can now, in view of (3.17), be written as

$$\Gamma_c(\boldsymbol{u}_h) = \{\boldsymbol{x} \in \Gamma : l_h(\boldsymbol{u}_h(\boldsymbol{x})) > 0\}. \tag{3.19}$$

On the other hand, testing with $\boldsymbol{v}_h$ in (3.7) and using (3.17) yields

$$\begin{aligned} &\sum_{i=1}^{2} (\boldsymbol{\sigma}_i(\boldsymbol{u}_{i,h}), \boldsymbol{\varepsilon}(\boldsymbol{v}_{i,h}))_{\Omega_i} - \langle [\![v_{h,n}]\!], (l_h(\boldsymbol{u}_h))_+ \rangle \\ &\qquad - \alpha \sum_{i=1}^{2} \sum_{E \in \mathcal{G}_h^i} \frac{h_E}{\mu_i} \Big((l_h(\boldsymbol{u}_h))_+ + \sigma_{i,n}(\boldsymbol{u}_{i,h}),\ \sigma_{i,n}(\boldsymbol{v}_{i,h}) \Big)_E \\ &\qquad = \sum_{i=1}^{2} (\boldsymbol{f}_i, \boldsymbol{v}_{i,h})_{\Omega_i} \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h. \end{aligned} \tag{3.20}$$

It follows from (3.10) that

$$\begin{aligned} &-\langle [\![v_{h,n}]\!], (l_h(\boldsymbol{u}_h))_+\rangle \\ &\quad = \Big(\{\!\{\sigma_n(\boldsymbol{u}_h)\}\!\}, [\![v_{h,n}]\!] \Big)_{\Gamma_c(\boldsymbol{u}_h)} + \Big(\beta_h [\![u_{h,n}]\!], [\![v_{h,n}]\!]\Big)_{\Gamma_c(\boldsymbol{u}_h)}, \end{aligned} \tag{3.21}$$

and on $\Gamma_c(\boldsymbol{u}_h)$ it holds that

$$(l_h(\boldsymbol{u}_h))_+ + \sigma_{1,n}(\boldsymbol{u}_1) = \frac{\hbar_2\mu_1}{\hbar_1\mu_2 + \hbar_2\mu_1}\big(\sigma_{1,n}(\boldsymbol{u}_1) - \sigma_{2,n}(\boldsymbol{u}_2)\big) - \beta_h [\![u_{h,n}]\!], \tag{3.22}$$

$$(l_h(\boldsymbol{u}_h))_+ + \sigma_{2,n}(\boldsymbol{u}_2) = \frac{\hbar_1\mu_2}{\hbar_1\mu_2 + \hbar_2\mu_1}\big(\sigma_{2,n}(\boldsymbol{u}_2) - \sigma_{1,n}(\boldsymbol{u}_2)\big) - \beta_h [\![u_{h,n}]\!]. \tag{3.23}$$

Therefore, defining the jump

$$[\![\sigma_n(\boldsymbol{u}_h)]\!] = \sigma_{2,n}(\boldsymbol{u}_1) - \sigma_{1,n}(\boldsymbol{u}_2) \tag{3.24}$$

and the $L^2(\Gamma)$-function

$$\gamma_h = \frac{\alpha \hbar_1 \hbar_2}{\hbar_1\mu_2 + \hbar_2\mu_1}, \tag{3.25}$$

and substituting the above five expressions into (3.20), we obtain after rearranging terms the following Nitsche's formulation for Problem 4 with $\boldsymbol{u}_h$ as the sole unknown.

NITSCHE FORMULATION 1. *Find $\boldsymbol{u}_h \in \boldsymbol{V}_h$ such that*

$$\begin{aligned} &\sum_{i=1}^{2}(\sigma_i(\boldsymbol{u}_{i,h}), \boldsymbol{\varepsilon}(\boldsymbol{v}_{i,h}))_{\Omega_i} + \Big(\beta_h [\![u_{h,n}]\!], [\![v_{h,n}]\!]\Big)_{\Gamma_c(\boldsymbol{u}_h)} \\ &\qquad + \Big(\{\!\{\sigma_n(\boldsymbol{u}_h)\}\!\}, [\![v_{h,n}]\!] \Big)_{\Gamma_c(\boldsymbol{u}_h)} + \Big(\{\!\{\sigma_n(\boldsymbol{v}_h)\}\!\}, [\![u_{h,n}]\!] \Big)_{\Gamma_c(\boldsymbol{u}_h)} \\ &\qquad - \Big(\gamma_h [\![\sigma_n(\boldsymbol{u}_h)]\!], [\![\sigma_n(\boldsymbol{v}_h)]\!]\Big)_{\Gamma_c(\boldsymbol{u}_h)} \\ &\qquad - \alpha \sum_{i=1}^{2} \left(\frac{\hbar_i}{\mu_i}\sigma_{i,n}(\boldsymbol{u}_{i,h}),\ \sigma_{i,n}(\boldsymbol{v}_{i,h})\right)_{\Gamma\setminus\Gamma_c(\boldsymbol{u}_h)} \\ &\quad = \sum_{i=1}^{2}(\boldsymbol{f}_i, \boldsymbol{v}_{i,h})_{\Omega_i} \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h. \end{aligned} \tag{3.26}$$

Remark 3.1. Since $\sigma_n(\boldsymbol{u}_i)$ vanishes on $\Gamma \setminus \Gamma_c(\boldsymbol{u}_h)$, this set can be reinterpreted as being part of $\Gamma_{N,i}$, $i = 1, 2$. Consequently, the term

$$\alpha \sum_{i=1}^{2} \left(\frac{\hbar_i}{\mu_i}\sigma_{i,n}(\boldsymbol{u}_{i,h}),\ \sigma_{i,n}(\boldsymbol{v}_{i,h})\right)_{\Gamma\setminus\Gamma_c(\boldsymbol{u}_h)}$$

can be dropped.

Next we present two other variants of Nitsche's method. The first is the so-called master-slave formulation.

Assume that the material parameters satisfy $\mu_1 \geq \mu_2$. The body Ω_1 is the master part, Ω_2 the slave, and the mortaring at the contact surface is only done for the latter, less rigid body, i.e., the stabilizing term is now

$$\mathcal{S}_h(\boldsymbol{w}_h, \xi_h; \boldsymbol{v}_h, \eta_h) = \sum_{E \in \mathcal{G}_h^2} \frac{h_E}{\mu_2} \Big(\xi_h + \sigma_{2,n}(\boldsymbol{w}_{2,h}), \eta_h + \sigma_{2,n}(\boldsymbol{v}_{2,h}) \Big)_E. \tag{3.27}$$

Repeating the steps above, we obtain $\lambda_h = (l_h(\boldsymbol{u}_h))_+$ with

$$l_h(\boldsymbol{u}_h) = -\sigma_{2,n}(\boldsymbol{u}_{2,h}) - \frac{\mu_2}{\alpha \hbar_2} [\![u_{h,n}]\!] . \tag{3.28}$$

The contact region $\Gamma_c(\boldsymbol{u}_h)$ is given by (3.19) with $l_h(\boldsymbol{u}_h)$ taken from (3.28), and we have the following method.

NITSCHE FORMULATION 2. *Find $\boldsymbol{u}_h \in \boldsymbol{V}_h$ such that*

$$\begin{aligned} &\sum_{i=1}^{2} (\boldsymbol{\sigma}_i(\boldsymbol{u}_{i,h}), \boldsymbol{\varepsilon}(\boldsymbol{v}_{i,h}))_{\Omega_i} + \left(\frac{\mu_2}{\alpha \hbar_2} [\![u_{h,n}]\!] , [\![v_{h,n}]\!] \right)_{\Gamma_c(\boldsymbol{u}_h)} \\ &\quad + \Big(\sigma_{2,n}(\boldsymbol{u}_{2,h}), [\![v_{h,n}]\!] \Big)_{\Gamma_c(\boldsymbol{u}_h)} + \Big(\sigma_{2,n}(\boldsymbol{v}_{2,h}), [\![u_{h,n}]\!] \Big)_{\Gamma_c(\boldsymbol{u}_h)} \\ &\quad - \alpha \left(\frac{\hbar_2}{\mu_2} \sigma_{2,n}(\boldsymbol{u}_{2,h}), \sigma_{2,n}(\boldsymbol{v}_{2,h}) \right)_{\Gamma \setminus \Gamma_c(\boldsymbol{u}_h)} \\ &= \sum_{i=1}^{2} (\boldsymbol{f}_i, \boldsymbol{v}_{i,h})_{\Omega_i} \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h. \end{aligned} \tag{3.29}$$

Again, the term

$$\alpha \left(\frac{\hbar_2}{\mu_2} \sigma_{2,n}(\boldsymbol{u}_{2,h}), \sigma_{2,n}(\boldsymbol{v}_{2,h}) \right)_{\Gamma \setminus \Gamma_c(\boldsymbol{u}_h)}$$

can be dropped; see Remark 3.1.

In the third alternative, we follow [18] and define the stabilizing term through

$$\alpha \mathcal{S}_h(\boldsymbol{w}_h, \xi_h; \boldsymbol{v}_h, \eta_h) = \Big(\beta_h^{-1} (\xi_h + \{\!\{ \sigma_n(\boldsymbol{w}_h) \}\!\}), \eta_h + \{\!\{ \sigma_n(\boldsymbol{v}_h) \}\!\} \Big)_\Gamma. \tag{3.30}$$

Repeating once more the above computations, we arrive at the following method.

NITSCHE FORMULATION 3. *Find $\boldsymbol{u}_h \in \boldsymbol{V}_h$ such that*

$$\begin{aligned} &\sum_{i=1}^{2} (\boldsymbol{\sigma}_i(\boldsymbol{u}_{i,h}), \boldsymbol{\varepsilon}(\boldsymbol{v}_{i,h}))_{\Omega_i} + \Big(\beta_h [\![u_{h,n}]\!] , [\![v_{h,n}]\!] \Big)_{\Gamma_c(\boldsymbol{u}_h)} \\ &\quad + \Big(\{\!\{ \sigma_n(\boldsymbol{u}_h) \}\!\} , [\![v_{h,n}]\!] \Big)_{\Gamma_c(\boldsymbol{u}_h)} + \Big(\{\!\{ \sigma_n(\boldsymbol{v}_h) \}\!\} , [\![u_{h,n}]\!] \Big)_{\Gamma_c(\boldsymbol{u}_h)} \\ &\quad - \Big(\beta_h^{-1} (\{\!\{ \sigma_n(\boldsymbol{u}_h) \}\!\}), \{\!\{ \sigma_n(\boldsymbol{v}_h) \}\!\} \Big)_{\Gamma \setminus \Gamma_c(\boldsymbol{u}_h)} \\ &= \sum_{i=1}^{2} (\boldsymbol{f}_i, \boldsymbol{v}_{i,h})_{\Omega_i} \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h, \end{aligned} \tag{3.31}$$

with $\Gamma_c(\boldsymbol{u}_h)$ given by (3.19) (*and $l_h(\boldsymbol{u}_h)$ as in* (3.17)).

Also here the term

$$\Big(\beta_h^{-1}(\{\!\{\sigma_n(\boldsymbol{u}_h)\}\!\}), \{\!\{\sigma_n(\boldsymbol{v}_h)\}\!\}\Big)_{\Gamma\setminus\Gamma_c(\boldsymbol{u}_h)} \tag{3.32}$$

can be dropped.

4. Error analysis. The energy norm for the problem is

$$\sum_{i=1}^{2}(\boldsymbol{\sigma}_i(\boldsymbol{w}_i), \boldsymbol{\varepsilon}(\boldsymbol{w}_i))_{\Omega_i}. \tag{4.1}$$

Since we exclude nearly incompressible materials, it holds that $\lambda_i \lesssim \mu_i$, and hence with our choice of boundary conditions the Korn inequality is valid in both regions, and we have the norm equivalence

$$\sum_{i=1}^{2}(\boldsymbol{\sigma}_i(\boldsymbol{w}_i), \boldsymbol{\varepsilon}(\boldsymbol{w}_i))_{\Omega_i} \approx \sum_{i=1}^{2}\mu_i\|\boldsymbol{w}\|_{1,\Omega_i}^2. \tag{4.2}$$

The error estimate will be given in the continuous norm

$$|\!|\!|(\boldsymbol{w},\xi)|\!|\!|^2 = \sum_{i=1}^{2}\Big(\mu_i\|\boldsymbol{w}\|_{1,\Omega_i}^2 + \frac{1}{\mu_i}\|\xi\|_{-\frac12,\Gamma}^2\Big), \tag{4.3}$$

but in the analysis we will also use the following mesh dependent norm:

$$|\!|\!|(\boldsymbol{w}_h,\xi_h)|\!|\!|_h^2 = |\!|\!|(\boldsymbol{w}_h,\xi_h)|\!|\!|^2 + \sum_{i=1}^{2}\sum_{E\in\mathcal{G}_h^i}\frac{h_E}{\mu_i}\|\xi_h\|_{0,E}^2. \tag{4.4}$$

THEOREM 4.1 (continuous stability). *For every $(\boldsymbol{w},\xi)\in\boldsymbol{V}\times Q$ there exists $\boldsymbol{v}\in\boldsymbol{V}$ such that*

$$\mathcal{B}(\boldsymbol{w},\xi;\boldsymbol{v},-\xi) \gtrsim |\!|\!|(\boldsymbol{w},\xi)|\!|\!|^2 \tag{4.5}$$

and

$$\|\boldsymbol{v}\|_V \lesssim |\!|\!|(\boldsymbol{w},\xi)|\!|\!|. \tag{4.6}$$

Proof. It is well known that the inf-sup condition

$$\sup_{\boldsymbol{z}_i\in\boldsymbol{V}_i}\frac{\langle -\boldsymbol{z}_i\cdot\boldsymbol{n}_i,\xi\rangle}{\|\nabla\boldsymbol{z}_i\|_{0,\Omega_i}} \ge C_i\|\xi\|_{-\frac12,\Gamma} \qquad \forall\xi\in Q \tag{4.7}$$

holds in both subdomains Ω_i (cf. [1]). Therefore

$$\sup_{\boldsymbol{z}=(\boldsymbol{z}_1,\boldsymbol{z}_2)\in\boldsymbol{V}}\frac{\langle [\![z_n]\!],\xi\rangle}{(\sum_{i=1}^{2}\mu_i\|\nabla\boldsymbol{z}_i\|_{0,\Omega_i}^2)^{1/2}} \ge C\left(\frac{1}{\mu_1}+\frac{1}{\mu_2}\right)^{1/2}\|\xi\|_{-\frac12,\Gamma} \qquad \forall\xi\in Q\,. \tag{4.8}$$

Assume then that $(\boldsymbol{w},\xi)\in\boldsymbol{V}\times Q$ is given, and let $\boldsymbol{v}_i=\boldsymbol{w}_i-\boldsymbol{q}_i$, where $\boldsymbol{q}_i\in\boldsymbol{V}_i$ solves the problem

$$(\boldsymbol{\sigma}_i(\boldsymbol{q}_i),\boldsymbol{\varepsilon}(\boldsymbol{z}_i))_{\Omega_i} = \langle -\boldsymbol{z}_i\cdot\boldsymbol{n}_i,\xi\rangle \quad \forall\boldsymbol{z}_i\in\boldsymbol{V}_i\,,\quad i=1,2\,.$$

Choosing $\boldsymbol{z}_i = \boldsymbol{q}_i$ above, we obtain after summing

$$\sum_{i=1}^{2} (\boldsymbol{\sigma}_i(\boldsymbol{q}_i), \boldsymbol{\varepsilon}(\boldsymbol{q}_i))_{\Omega_i} = \langle [\![q_n]\!], \xi \rangle \,.$$

Moreover, from (4.7), it follows that

$$\|\xi\|_{-\frac{1}{2},\Gamma} \lesssim \sup_{\boldsymbol{z}_i \in \boldsymbol{V}_i} \frac{\langle -\boldsymbol{z}_i \cdot \boldsymbol{n}_i, \xi \rangle}{\|\nabla \boldsymbol{z}_i\|_{0,\Omega_i}} = \sup_{\boldsymbol{z}_i \in \boldsymbol{V}_i} \frac{(\boldsymbol{\sigma}_i(\boldsymbol{q}_i), \boldsymbol{\varepsilon}(\boldsymbol{z}_i))_{\Omega_i}}{\|\nabla \boldsymbol{z}_i\|_{0,\Omega_i}} \lesssim \mu_i \|\boldsymbol{q}_i\|_{1,\Omega_i}$$

and thus

$$\left(\frac{1}{\mu_1} + \frac{1}{\mu_2} \right)^{1/2} \|\xi\|_{-\frac{1}{2},\Gamma} \lesssim \left(\sum_{i=1}^{2} \mu_i \|\boldsymbol{q}_i\|^2_{1,\Omega_i} \right)^{1/2} .$$

Now, it is easy to see that

$$\begin{aligned} \mathcal{B}(\boldsymbol{w}, \xi; \boldsymbol{v}, -\xi) &= \sum_{i=1}^{2} \Big\{ (\boldsymbol{\sigma}_i(\boldsymbol{w}_i), \boldsymbol{\varepsilon}(\boldsymbol{w}_i))_{\Omega_i} - (\boldsymbol{\sigma}_i(\boldsymbol{w}_i), \boldsymbol{\varepsilon}(\boldsymbol{q}_i))_{\Omega_i} \Big\} + \langle [\![q_n]\!], \xi \rangle \\ &\gtrsim \sum_{i=1}^{2} \mu_i \|\boldsymbol{w}_i\|^2_{1,\Omega_i} - \frac{1}{2} \sum_{i=1}^{2} \mu_i \|\boldsymbol{w}_i\|^2_{1,\Omega_i} - \frac{1}{2} \sum_{i=1}^{2} \mu_i \|\boldsymbol{q}_i\|^2_{1,\Omega_i} \\ &\quad + \sum_{i=1}^{2} (\boldsymbol{\sigma}_i(\boldsymbol{q}_i), \boldsymbol{\varepsilon}(\boldsymbol{q}_i))_{\Omega_i} \\ &\gtrsim \sum_{i=1}^{2} \mu_i \|\boldsymbol{w}_i\|^2_{1,\Omega_i} + \left(\frac{1}{\mu_1} + \frac{1}{\mu_2} \right) \|\xi\|^2_{-\frac{1}{2},\Gamma} = |\!|\!|(\boldsymbol{w}, \xi)|\!|\!|^2 \end{aligned}$$

and that $\|\boldsymbol{v}\|_V = \|\boldsymbol{w} - \boldsymbol{q}\|_V \lesssim |\!|\!|(\boldsymbol{w}, \xi)|\!|\!|$. ∎

Above and in the following we write $a \gtrsim b$ (or $a \lesssim b$) when $a \geq Cb$ (or $a \leq Cb$) for some positive constant C independent of the finite element mesh.

To derive the discrete stability estimate, we need the following discrete trace inequality, easily shown by a scaling argument.

LEMMA 4.1 (discrete trace estimate). *There exists $C_I > 0$, independent of the mesh parameter h, such that*

$$C_I \sum_{E \in \mathcal{G}_h^i} \frac{h_E}{\mu_i} \|\sigma_{i,n}(\boldsymbol{v}_{i,h})\|^2_{0,E} \leq \mu_i \|\boldsymbol{v}_{i,h}\|^2_{1,\Omega_i} \quad \forall \boldsymbol{v}_{i,h} \in \boldsymbol{V}_i, \quad i = 1, 2. \tag{4.9}$$

THEOREM 4.2 (discrete stability). *Suppose that $0 < \alpha < C_I$. Then, for every $(\boldsymbol{w}_h, \xi_h) \in \boldsymbol{V}_h \times Q_h$, there exists $\boldsymbol{v}_h \in \boldsymbol{V}_h$ such that*

$$\mathcal{B}_h(\boldsymbol{w}_h, \xi_h; \boldsymbol{v}_h, -\xi_h) \gtrsim |\!|\!|(\boldsymbol{w}_h, \xi_h)|\!|\!|_h^2 \tag{4.10}$$

and

$$\|\boldsymbol{v}_h\|_V \lesssim |\!|\!|(\boldsymbol{w}_h, \xi_h)|\!|\!|_h. \tag{4.11}$$

Proof. From the discrete trace estimate it follows that

$$\mathcal{B}_h(\boldsymbol{w}_h, \xi_h; \boldsymbol{w}_h, -\xi_h) \geq \left(1 - \frac{\alpha}{C_I} \right) \sum_{i=1}^{2} \mu_i \|\boldsymbol{w}_{i,h}\|^2_{1,\Omega_i} + \alpha \sum_{i=1}^{2} \sum_{E \in \mathcal{G}_h} \frac{h_E}{\mu_i} \|\xi_h\|^2_{0,E} \,,$$

which proves the result in the mesh-dependent norm of ξ_h for $0 < \alpha < C_I$.

On the other hand, the continuous inf-sup condition (4.8) implies that for any $\xi_h \in Q_h$ there exists $\boldsymbol{v} \in \boldsymbol{V}$ such that

$$\frac{\langle [\![v_n]\!] , \xi_h \rangle}{(\sum_{i=1}^2 \mu_i \|\nabla \boldsymbol{v}_i\|^2_{0,\Omega_i})^{1/2}} \geq C_1 \left(\frac{1}{\mu_1} + \frac{1}{\mu_2} \right)^{1/2} \|\xi_h\|_{-\frac{1}{2},\Gamma} .$$

This means that (cf. the proof of Lemma 3.2 in [12])

$$\langle [\![(I_h v)_n]\!] , \xi_h \rangle \geq C_2 \left(\frac{1}{\mu_1} + \frac{1}{\mu_2} \right) \|\xi_h\|^2_{-\frac{1}{2},\Gamma} - C_3 \sum_{i=1}^2 \sum_{E \in \mathcal{G}_h} \frac{h_E}{\mu_i} \|\xi_h\|^2_{0,E} \tag{4.12}$$

$$\sum_{i=1}^2 \mu_i \|I_h \boldsymbol{v}_i\|_{1,\Omega_i} \leq C_4 \left(\frac{1}{\mu_1} + \frac{1}{\mu_2} \right) \|\xi_h\|^2_{-\frac{1}{2},\Gamma}, \tag{4.13}$$

where C_2, C_3, C_4 are positive constants and $I_h \boldsymbol{v} \in \boldsymbol{V}_h$ is the Clément interpolant of $\boldsymbol{v}$. Using again the discrete trace estimate and inequalities (4.12) and (4.13), we then obtain

$$\begin{aligned}
\mathcal{B}_h(\boldsymbol{w}_h, \xi_h; -I_h \boldsymbol{v}, 0) = &- \sum_{i=1}^2 (\boldsymbol{\sigma}_i(\boldsymbol{w}_{i,h}), \boldsymbol{\varepsilon}(I_h \boldsymbol{v}_i))_{\Omega_i} + \langle [\![(I_h v)_n]\!] , \xi_h \rangle \\
&- \sum_{i=1}^2 \sum_{E \in \mathcal{G}_h^i} \frac{h_E}{\mu_i} \Big(\xi_h + \sigma_{i,n}(\boldsymbol{w}_{i,h}), \sigma_{i,n}(I_h \boldsymbol{v}_i) \Big)_E, \\
\geq &\, C_5 \left(\frac{1}{\mu_1} + \frac{1}{\mu_2} \right) \|\xi_h\|^2_{-\frac{1}{2},\Gamma} - C_6 \sum_{i=1}^2 \mu_i \|\boldsymbol{w}_{i,h}\|_{1,\Omega_i} \\
&- C_7 \sum_{i=1}^2 \sum_{E \in \mathcal{G}_h} \frac{h_E}{\mu_i} \|\xi_h\|^2_{0,E}.
\end{aligned}$$

Now, it is straightforward to show (cf. [14]) that there exists $\delta > 0$ such that

$$\mathcal{B}_h(\boldsymbol{w}_h, \xi_h; \boldsymbol{w}_h - \delta I_h \boldsymbol{v}, -\xi_h) \gtrsim |\!|\!| (\boldsymbol{w}_h, \xi_h) |\!|\!|_h^2$$

and that $\|\boldsymbol{w}_h - \delta I_h \boldsymbol{v}\|_V \lesssim |\!|\!| (\boldsymbol{w}_h, \xi_h) |\!|\!|_h$. ∎

In our improved error analysis, we use techniques from the a posteriori error analysis. Let $\boldsymbol{f}_{i,h} \in \boldsymbol{V}_{i,h}$ be the $[L^2(\Omega_i)]^d$ projection of $\boldsymbol{f}_i$, define on any $K \in \mathcal{C}_h^i$ the oscillation of $\boldsymbol{f}_i$ by

$$\operatorname{osc}_K(\boldsymbol{f}_i) = h_K \|\boldsymbol{f}_i - \boldsymbol{f}_{i,h}\|_{0,K} , \qquad i = 1, 2,$$

and, for each $E \in \mathcal{G}_h^i$, let $K(E) \in \mathcal{G}_h^i$ denote the element such that $\partial K(E) \cap E = E$.

LEMMA 4.2. *For any $(\boldsymbol{v}_h, \eta_h) \in \boldsymbol{V}_h \times Q_h$, it holds that*

$$\begin{aligned}
&\left(\sum_{i=1}^2 \sum_{E \in \mathcal{G}_h^i} \frac{h_E}{\mu_i} \|\eta_h + \sigma_{i,n}(\boldsymbol{v}_{i,h})\|^2_{0,E} \right)^{1/2} \\
&\qquad \leq |\!|\!| (\boldsymbol{u} - \boldsymbol{v}_h, \lambda - \eta_h) |\!|\!| + \left(\sum_{i=1}^2 \mu_i^{-1} \sum_{E \in \mathcal{G}_h^i} \operatorname{osc}_{K(E)}(\boldsymbol{f}_i)^2 \right)^{1/2}.
\end{aligned} \tag{4.14}$$

Proof. We follow the reasoning presented for the mortar method in [14]. It is clearly enough to prove the result in Ω_1. Thus, let $b_E \in P_d(E)$, $E \in \mathcal{G}_h^1$, be the usual edge/facet bubble function, and define τ_E on $K(E) \in \mathcal{C}_h^1$ through

$$\tau_E\big|_E = \frac{h_E b_E}{\mu_1}\Big(\eta_h + \sigma_{1,n}(\boldsymbol{v}_{1,h})\Big) \quad \text{and} \quad \tau_E\big|_{\partial K(E)\setminus E} = 0,$$

where $K(E)$ is such that $\overline{K(E)} \cap E = E$. It follows that

$$\frac{h_E}{\mu_1}\Big\|\eta_h + \sigma_{1,n}(\boldsymbol{v}_{1,h})\Big\|_{0,E}^2 \lesssim \Big(\eta_h + \sigma_{1,n}(\boldsymbol{v}_{1,h}), \tau_E\Big)_E. \tag{4.15}$$

Next, defining $\boldsymbol{\tau} \in \boldsymbol{V}_{1,h}$ in such a way that $\tau_n := -\boldsymbol{\tau}\cdot\boldsymbol{n} = \sum_{E\in\mathcal{G}_h^1} \tau_E$ and testing Problem 3 with $(\boldsymbol{v}_1, \boldsymbol{v}_2, \eta) = (-\boldsymbol{\tau}, 0, \lambda)$, we obtain

$$0 \leq (\boldsymbol{\sigma}_1(\boldsymbol{u}_1), \boldsymbol{\varepsilon}(\boldsymbol{\tau}))_{\Omega_1} - \langle \tau_n, \lambda\rangle - (\boldsymbol{f}_1, \boldsymbol{\tau})_{\Omega_1}.$$

Summing (4.15) over the edges in $\mathcal{G}_h^1$ gives then

$$\begin{aligned}
&\sum_{E\in\mathcal{G}_h^1} \frac{h_E}{\mu_1}\Big\|\eta_h + \sigma_{1,n}(\boldsymbol{v}_{1,h})\Big\|_{0,E}^2 \\
&\quad\lesssim \langle \tau_n, \eta_h - \lambda\rangle + (\boldsymbol{\sigma}_1(\boldsymbol{u}_1), \boldsymbol{\varepsilon}(\boldsymbol{\tau}))_{\Omega_1} - (\boldsymbol{f}_1, \boldsymbol{\tau})_{\Omega_1} + \sum_{E\in\mathcal{G}_h^1} (\sigma_{1,n}(\boldsymbol{v}_{1,h}), \tau_E)_E \\
&\quad= \langle \tau_n, \eta_h - \lambda\rangle + (\boldsymbol{\sigma}_1(\boldsymbol{u}_1), \boldsymbol{\varepsilon}(\boldsymbol{\tau}))_{\Omega_1} - (\boldsymbol{f}_1, \boldsymbol{\tau})_{\Omega_1} \\
&\qquad - (\mathbf{div}\,\boldsymbol{\sigma}_1(\boldsymbol{v}_{1,h}), \boldsymbol{\tau})_{\Omega_1} - (\boldsymbol{\sigma}_1(\boldsymbol{v}_{1,h}), \boldsymbol{\varepsilon}(\boldsymbol{\tau})_{\Omega_1} \\
&\quad= \langle \tau_n, \eta_h - \lambda\rangle + (\boldsymbol{\sigma}_1(\boldsymbol{u}_1) - \boldsymbol{\sigma}_1(\boldsymbol{v}_{1,h}), \boldsymbol{\varepsilon}(\boldsymbol{\tau}))_{\Omega_1} - (\mathbf{div}\,\boldsymbol{\sigma}_1(\boldsymbol{v}_{1,h}) + \boldsymbol{f}_1, \boldsymbol{\tau})_{\Omega_1}.
\end{aligned}$$

Inverse estimates imply that

$$\mu_1\|\boldsymbol{\tau}\|_{1,\Omega_1}^2 \lesssim \mu_1 \sum_{E\in\mathcal{G}_h^1} h_E^{-2}\|\tau_E\|_{0,K(E)}^2 \lesssim \sum_{E\in\mathcal{G}_h^1} \frac{h_E}{\mu_1}\|\eta_h + \sigma_{1,n}(\boldsymbol{v}_{1,h})\|_{0,E}^2. \tag{4.16}$$

Now, one readily sees, using trace inequalities and the norm equivalence (4.2), that

$$\begin{aligned}
&\sum_{E\in\mathcal{G}_h^1} \frac{h_E}{\mu_1}\Big\|\eta_h + \sigma_{1,n}(\boldsymbol{v}_{1,h})\Big\|_{0,E}^2 \\
&\quad\lesssim \mu_1^{-1/2}\|\eta_h - \lambda\|_{-\frac{1}{2},\Gamma}\, \mu_1^{1/2}\|\boldsymbol{\tau}\|_{1,\Omega_1} + \mu_1^{1/2}\|\boldsymbol{u}_1 - \boldsymbol{v}_{1,h}\|_{1,\Omega_1}\, \mu_1^{1/2}\|\boldsymbol{\tau}\|_{1,\Omega_1} \\
&\qquad + \left(\sum_{E\in\mathcal{G}_h^1} \frac{h_E^2}{\mu_1}\|\mathbf{div}\,\boldsymbol{\sigma}_1(\boldsymbol{v}_{1,h}) + \boldsymbol{f}_1\|_{0,E}^2\right)^{1/2} \left(\mu_1 \sum_{E\in\mathcal{G}_h^1} h_E^{-2}\|\tau_E\|_{0,K(E)}^2\right)^{1/2},
\end{aligned}$$

from which, using the standard estimates for interior residuals (cf. [26]) and the inverse estimate (4.16) to bound the last term, it follows that

$$\left(\sum_{E\in\mathcal{G}_h^1} \frac{h_E}{\mu_1}\Big\|\eta_h + \sigma_{1,n}(\boldsymbol{v}_{1,h})\Big\|_{0,E}^2\right)^{1/2} \lesssim |||(\boldsymbol{u} - \boldsymbol{v}_h, \lambda - \eta_h)||| + \left(\mu_1^{-1} \sum_{E\in\mathcal{G}_h^1} \mathrm{osc}_{K(E)}(\boldsymbol{f}_1)^2\right)^{1/2},$$

which concludes the proof. ∎

We can now establish the quasi-optimality of the method.

THEOREM 4.3. *For $0 < \alpha < C_I$ it holds that*

$$\begin{aligned}|||(\boldsymbol{u}-\boldsymbol{u}_h, \lambda-\lambda_h)||| \lesssim \inf_{(\boldsymbol{v}_h,\eta_h)\in \boldsymbol{V}_h\times\Lambda_h} \Big(|||(\boldsymbol{u}-\boldsymbol{v}_h, \lambda-\eta_h)||| + \sqrt{\langle [\![u_n]\!], \eta_h\rangle}\Big) \\ + \left(\sum_{i=1}^{2} \mu_i^{-1} \sum_{E\in\mathcal{G}_h^i} \mathrm{osc}_{K(E)}(\boldsymbol{f}_i)^2\right)^{1/2}.\end{aligned} \tag{4.17}$$

Proof. On account of the discrete stability estimate, there exists $\boldsymbol{w}_h \in \boldsymbol{V}_h$ such that

$$\|\boldsymbol{w}_h\|_V \lesssim |||(\boldsymbol{u}_h-\boldsymbol{v}_h, \lambda_h-\eta_h)|||_h \tag{4.18}$$

and

$$|||(\boldsymbol{u}_h-\boldsymbol{v}_h, \lambda_h-\eta_h)|||_h^2 \lesssim \mathcal{B}_h(\boldsymbol{u}_h-\boldsymbol{v}_h, \lambda_h-\eta_h; \boldsymbol{w}_h, \eta_h-\lambda_h). \tag{4.19}$$

Using the bilinearity and (3.7), we obtain

$$\begin{aligned}&\mathcal{B}_h(\boldsymbol{u}_h-\boldsymbol{v}_h, \lambda_h-\eta_h; \boldsymbol{w}_h, \eta_h-\lambda_h)\\ &\quad = \mathcal{B}_h(\boldsymbol{u}_h, \lambda_h; \boldsymbol{w}_h, \eta_h-\lambda_h) - \mathcal{B}_h(\boldsymbol{v}_h, \eta_h; \boldsymbol{w}_h, \eta_h-\lambda_h)\\ &\quad \lesssim \mathcal{L}(\boldsymbol{w}_h) - \mathcal{B}_h(\boldsymbol{v}_h, \eta_h; \boldsymbol{w}_h, \eta_h-\lambda_h)\\ &\quad = \mathcal{B}(\boldsymbol{u}-\boldsymbol{v}_h, \lambda-\eta_h; \boldsymbol{w}_h, \eta_h-\lambda_h) + \mathcal{L}(\boldsymbol{w}_h)\\ &\qquad - \mathcal{B}(\boldsymbol{u}, \lambda; \boldsymbol{w}_h, \eta_h-\lambda_h) + \alpha\mathcal{S}_h(\boldsymbol{v}_h, \eta_h; \boldsymbol{w}_h, \eta_h-\lambda_h).\end{aligned} \tag{4.20}$$

The terms above can be estimated as follows. First, continuity of the bilinear form $\mathcal{B}$ and inequality (4.18) yield

$$\mathcal{B}(\boldsymbol{u}-\boldsymbol{v}_h, \lambda-\eta_h; \boldsymbol{w}_h, \eta_h-\lambda_h) \lesssim |||(\boldsymbol{u}-\boldsymbol{v}_h, \lambda-\eta_h)|||\,|||(\boldsymbol{u}_h-\boldsymbol{v}_h, \lambda_h-\eta_h)|||. \tag{4.21}$$

Next, using Problem 3 and the fact that $[\![u_n]\!] \geq 0$ and $\lambda_h \geq 0$, we obtain

$$\mathcal{L}(\boldsymbol{w}_h) - \mathcal{B}(\boldsymbol{u}, \lambda; \boldsymbol{w}_h, \eta_h-\lambda_h) = \langle [\![u_n]\!], \eta_h-\lambda_h\rangle \leq \langle [\![u_n]\!], \eta_h\rangle. \tag{4.22}$$

Finally, from the discrete trace estimate (4.9) it follows that

$$\begin{aligned}&\alpha\mathcal{S}_h(\boldsymbol{v}_h, \eta_h; \boldsymbol{w}_h, \eta_h-\lambda_h)\\ &\quad \lesssim \left(\sum_{i=1}^{2}\sum_{E\in\mathcal{G}_h^i} \frac{h_E}{\mu_i}\|\eta_h+\sigma_{i,n}(\boldsymbol{u}_{i,h})\|_{0,E}^2\right)^{1/2} |||(\boldsymbol{u}_h-\boldsymbol{v}_h, \lambda_h-\eta_h)|||_h.\end{aligned} \tag{4.23}$$

Using Lemma 4.2 and collecting the above estimates, we arrive at the asserted error estimate. ∎

Remark 4.1. We refrain from giving an a priori error estimate assuming a regular solution. The reasons are twofold. Firstly, contact singularities are inevitable and essential in contact problems. Secondly, to derive an a priori bound, one would need to estimate the term $\sqrt{\langle [\![u_n]\!], \eta_h\rangle}$ with η_h being the interpolant to λ. Besides, and perhaps most importantly, one of the main results of this paper is the fact that we do not need to assume that the solution belongs to H^s with $s > 3/2$.

For the a posteriori error analysis, we define the local estimators

$$\eta_K^2 = \frac{h_K^2}{\mu_i}\|\mathbf{div}\,\boldsymbol{\sigma}_i(\boldsymbol{u}_{i,h}) + \boldsymbol{f}_i\|_{0,K}^2, \quad K \in \mathcal{C}_h^i, \tag{4.24}$$

$$\eta_{E,\Omega}^2 = \frac{h_E}{\mu_i}\,\|[\![\boldsymbol{\sigma}_i(\boldsymbol{u}_{i,h})\boldsymbol{n}]\!]\|_{0,E}^2, \quad E \in \mathcal{E}_h^i, \tag{4.25}$$

$$\begin{aligned}\eta_{E,\Gamma}^2 = \frac{h_E}{\mu_i}&\left\{\|\lambda_h + \sigma_{i,n}(\boldsymbol{u}_{i,h})\|_{0,E}^2 + \|\boldsymbol{\sigma}_{i,t}(\boldsymbol{u}_{i,h})\|_{0,E}^2\right\}\\ &+ \frac{\mu_i}{h_E}\|([\![u_{h,n}]\!])_-\|_{0,E}^2, \quad E \in \mathcal{G}_h^i,\end{aligned} \tag{4.26}$$

$$\eta_{E,\Gamma_N}^2 = \frac{h_E}{\mu_i}\,\|\boldsymbol{\sigma}_i(\boldsymbol{u}_{i,h})\boldsymbol{n}\|_{0,E}^2, \quad E \in \mathcal{N}_h^i, \tag{4.27}$$

with $i = 1, 2$. The corresponding global estimator η is then defined as

$$\eta^2 = \sum_{i=1}^{2}\left\{\sum_{K\in\mathcal{C}_h^i}\eta_K^2 + \sum_{E\in\mathcal{E}_h^i}\eta_{E,\Omega}^2 + \sum_{E\in\mathcal{G}_h^i}\eta_{E,\Gamma}^2 + \sum_{E\in\mathcal{N}_h^i}\eta_{E,\Gamma_N}^2\right\}. \tag{4.28}$$

In addition, we need an estimator S defined only globally as

$$S^2 = \big(([\![u_{h,n}]\!])_+, \lambda_h\big)_\Gamma. \tag{4.29}$$

THEOREM 4.4 (a posteriori error estimate). *It holds that*

$$|\!|\!|(\boldsymbol{u} - \boldsymbol{u}_h, \lambda - \lambda_h)|\!|\!| \lesssim \eta + S. \tag{4.30}$$

Proof. In view of the continuous stability estimate, there exists $\boldsymbol{v} \in \boldsymbol{V}$ with

$$\|\boldsymbol{v}\|_V \lesssim |\!|\!|(\boldsymbol{u} - \boldsymbol{u}_h, \lambda - \lambda_h)|\!|\!|, \tag{4.31}$$

and

$$|\!|\!|(\boldsymbol{u} - \boldsymbol{u}_h, \lambda - \lambda_h)|\!|\!|^2 \lesssim \mathcal{B}(\boldsymbol{u} - \boldsymbol{u}_h, \lambda - \lambda_h; \boldsymbol{v}, \lambda_h - \lambda). \tag{4.32}$$

Let $\tilde{\boldsymbol{v}} \in \boldsymbol{V}_h$ be the Clément interpolant of $\boldsymbol{v}$. From (3.7), it follows that

$$0 \le -\mathcal{B}(\boldsymbol{u}_h, \lambda_h; \tilde{\boldsymbol{v}}, 0) + \alpha\mathcal{S}_h(\boldsymbol{u}_h, \lambda_h, -\tilde{\boldsymbol{v}}, 0) - \mathcal{L}(\tilde{\boldsymbol{v}}). \tag{4.33}$$

Using Problem 3, this gives

$$\begin{aligned}&\mathcal{B}(\boldsymbol{u} - \boldsymbol{u}_h, \lambda - \lambda_h; \boldsymbol{v}, \lambda_h - \lambda)\\ &\quad\lesssim \mathcal{L}(\boldsymbol{v} - \tilde{\boldsymbol{v}}) - \mathcal{B}(\boldsymbol{u}_h, \lambda_h; \boldsymbol{v} - \tilde{\boldsymbol{v}}, \lambda_h - \lambda) + \alpha\mathcal{S}_h(\boldsymbol{u}_h, \lambda_h, -\tilde{\boldsymbol{v}}, 0).\end{aligned} \tag{4.34}$$

Integrating by parts, we obtain for the first two terms above

$$\begin{aligned}&\mathcal{L}(\boldsymbol{v} - \tilde{\boldsymbol{v}}) - \mathcal{B}(\boldsymbol{u}_h, \lambda_h; \boldsymbol{v} - \tilde{\boldsymbol{v}}, \lambda_h - \lambda)\\ &= \sum_{i=1}^{2}\sum_{K\in\mathcal{C}_h^i}(\mathbf{div}\,\boldsymbol{\sigma}_i(\boldsymbol{u}_{i,h}) + \boldsymbol{f}_i, \boldsymbol{v}_i - \tilde{\boldsymbol{v}}_i)_K\\ &\quad - \sum_{i=1}^{2}\sum_{E\in\mathcal{E}_h^i}([\![\boldsymbol{\sigma}_i(\boldsymbol{u}_{i,h})\boldsymbol{n}]\!], \boldsymbol{v}_i - \tilde{\boldsymbol{v}}_i)_E\\ &\quad - \sum_{i=1}^{2}\sum_{E\in\mathcal{N}_h^i}(\boldsymbol{\sigma}_i(\boldsymbol{u}_{i,h})\boldsymbol{n}, \boldsymbol{v}_i - \tilde{\boldsymbol{v}}_i)_E - \sum_{i=1}^{2}\sum_{E\in\mathcal{G}_h^i}(\boldsymbol{\sigma}_{i,t}(\boldsymbol{u}_i), (\boldsymbol{v}_{i,t} - \tilde{\boldsymbol{v}}_{i,t}))_E\\ &\quad - \sum_{i=1}^{2}\sum_{E\in\mathcal{G}_h^i}(\lambda_h + \sigma_{i,n}(\boldsymbol{u}_{i,h}), (\boldsymbol{v}_i - \tilde{\boldsymbol{v}}_i)\cdot\boldsymbol{n})_E + \langle[\![u_{h,n}]\!], \lambda_h - \lambda\rangle.\end{aligned} \tag{4.35}$$

Moreover, using an inverse inequality for the $H^{1/2}(\Gamma)$-norm (cf. [11]) we get

$$\begin{aligned}\langle [\![u_{h,n}]\!], \lambda_h - \lambda\rangle &\le \big(([\![u_{h,n}]\!])_+, \lambda_h\big)_\Gamma + \big\langle ([\![u_{h,n}]\!])_-, \lambda_h - \lambda\big\rangle \\ &\lesssim \big(([\![u_{h,n}]\!])_+, \lambda_h\big)_\Gamma \\ &\quad + |\!|\!|(\boldsymbol{u}-\boldsymbol{u}_h, \lambda-\lambda_h)|\!|\!| \left((\mu_1+\mu_2)\|([\![u_{h,n}]\!])_-\|^2_{1/2,\Gamma}\right)^{1/2} \\ &\lesssim \big(([\![u_{h,n}]\!])_+, \lambda_h\big)_\Gamma \\ &\quad + |\!|\!|(\boldsymbol{u}-\boldsymbol{u}_h, \lambda-\lambda_h)|\!|\!| \left(\sum_{i=1}^2 \sum_{E\in\mathcal{G}_h^i} \frac{\mu_i}{h_E}\|([\![u_{h,n}]\!])_-\|^2_{0,E}\right)^{1/2}.\end{aligned} \tag{4.36}$$

Finally, using the discrete trace estimate (4.9) and the standard bounds for the Clément interpolant, and recalling (4.31), we obtain for the stabilizing term

$$\begin{aligned}&|\mathcal{S}_h(\boldsymbol{u}_h, \lambda_h, -\tilde{\boldsymbol{v}}, 0)| \\ &\quad\lesssim \left(\sum_{i=1}^2 \sum_{E\in\mathcal{G}_h^i} \frac{h_E}{\mu_i}\|\lambda_h + \sigma_{i,n}(\boldsymbol{u}_{i,h})\|^2_{0,E}\right)^{1/2} |\!|\!|(\boldsymbol{u}-\boldsymbol{u}_h, \lambda-\lambda_h)|\!|\!|.\end{aligned} \tag{4.37}$$

Estimate (4.30) follows from collecting the above bounds. ∎

The estimator η bounds the error from below. For the proof of the following theorem we refer to [12].

Theorem 4.5 (a posteriori estimate—efficiency). *It holds that*

$$\eta \lesssim |\!|\!|(\boldsymbol{u}-\boldsymbol{u}_h, \lambda-\lambda_h)|\!|\!|. \tag{4.38}$$

The analysis of Nitsche Formulations 2 and 3 is analogous. In the a posteriori estimates the term

$$\sum_{i=1}^2 \sum_{E\in\mathcal{G}_h^i} \frac{h_E}{\mu_i}\|\lambda_h + \sigma_{i,n}(\boldsymbol{u}_{i,h})\|^2_{0,E}$$

is replaced by

$$\sum_{E\in\mathcal{G}_h^2} \frac{h_E}{\mu_2}\|\lambda_h + \sigma_{2,n}(\boldsymbol{u}_{2,h})\|^2_{0,E} \tag{4.39}$$

and

$$\|\beta_h^{-1/2}(\lambda_h + \{\!\{\sigma_n(\boldsymbol{u}_h)\}\!\})\|^2_{0,\Gamma} \tag{4.40}$$

for Nitsche Formulations 2 and 3, respectively.

5. Computational experiments. All computations presented in this section were obtained using Nitsche Formulation 3 with the term (3.32) dropped. Had we considered other formulations, the results would have been practically identical. We also note that since the stabilized/Nitsche's method is variationally conforming (as a mortaring method) it passes the patch test of [21, p. 425]. This was confirmed numerically up to machine accuracy.

We consider the geometry given by

$$\Omega_1 = [0.5, 1.0]\times[0.25, 0.75], \quad \Omega_2 = [1, 1.6]\times[0,1] \tag{5.1}$$

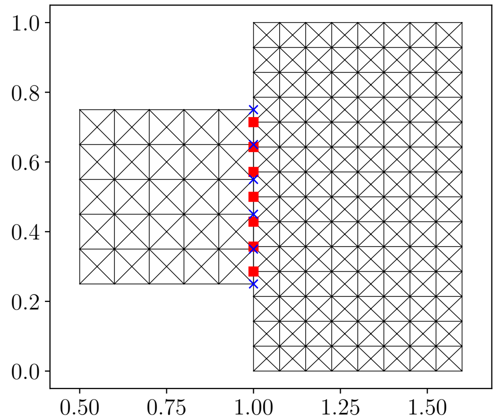


FIG. 2. *A finite element mesh and the vertices belonging to* Γ.

and define the boundary conditions on the following subsets:

$$\Gamma_{D,1} = \{(x,y) \in \partial\Omega_1 : x = 0.5\}, \quad \Gamma_{N,1} = \partial\Omega_1 \setminus (\Gamma_{D,1} \cup \Gamma), \tag{5.2}$$

$$\Gamma_{D,2} = \{(x,y) \in \partial\Omega_2 : x = 1.6\}, \quad \Gamma_{N,2} = \partial\Omega_2 \setminus (\Gamma_{D,2} \cup \Gamma). \tag{5.3}$$

Thus, the geometry is the one given in Figure 1. A nonmatching discretization of the geometry is depicted in Figure 2. Initially, the material parameters are $E_1 = E_2 = 1$ and $\nu_1 = \nu_2 = 0.3$, and the loading is

$$\boldsymbol{f}_1 = (x - 0.5, 0), \quad \boldsymbol{f}_2 = (0, 0). \tag{5.4}$$

For this loading, the displacement is constrained on $\Gamma_{D,i}$, $i = 1, 2$, only in the horizontal direction which minimizes the effect of the singularities—other than the ones related to the contact boundary—on the rates of convergence. We consider both linear and quadratic elements with $\alpha = 10^{-2}$ and $\alpha = 10^{-3}$, respectively.

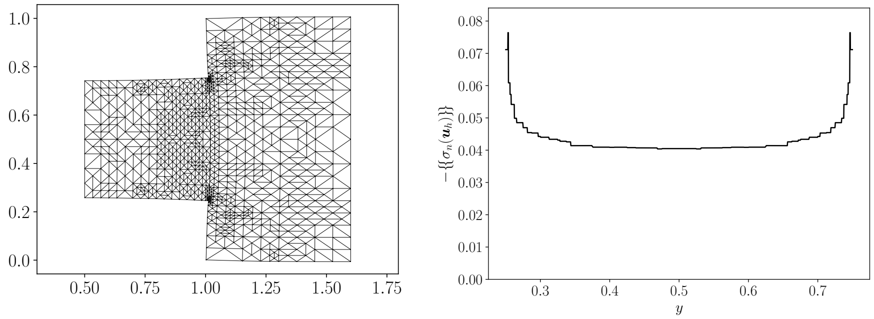
The adaptively refined meshes are shown in Figure 3(a) and (b), and the global error estimator $\eta + S$ is plotted as a function of the number of degrees-of-freedom N in Figure 3(c). Since $\eta + S$ is an upper bound for the total error, the results suggest that the total error of the quadratic solution is limited to $\mathcal{O}(N^{-0.5})$ when using uniform refinements and that adaptivity successfully improves the order of the discretization error to $\mathcal{O}(N^{-1})$.

Next we fix also the vertical displacement on $\Gamma_{D,i}$, $i = 1, 2$, and consider the loading

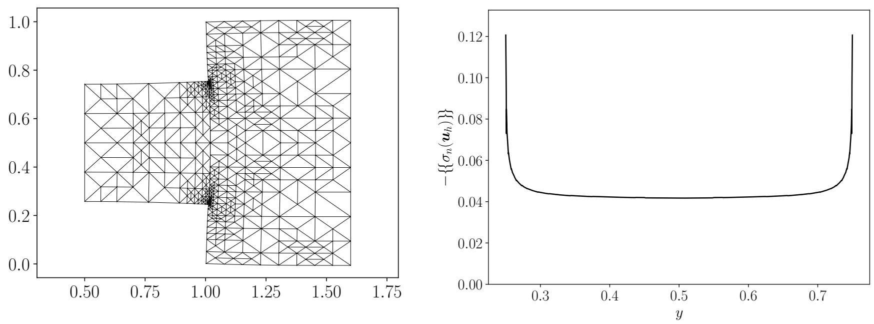
$$\boldsymbol{f}_1 = (0, -0.05), \quad \boldsymbol{f}_2 = (0, 0), \tag{5.5}$$

which causes the left block to bend slightly downwards and, as a consequence, the active contact region is a nontrivial subset of Γ. The active contact region is found via an iterative solution of the linearized problem; cf. [12]. See Figure 4(a) and (b) for the final meshes and contact stresses and Figure 4(c) for the convergence rates. We observe that the singularity at the upper corner of the contact region is properly resolved by the adaptive meshing strategy and that the convergence is similar albeit less idealized as in the first example.

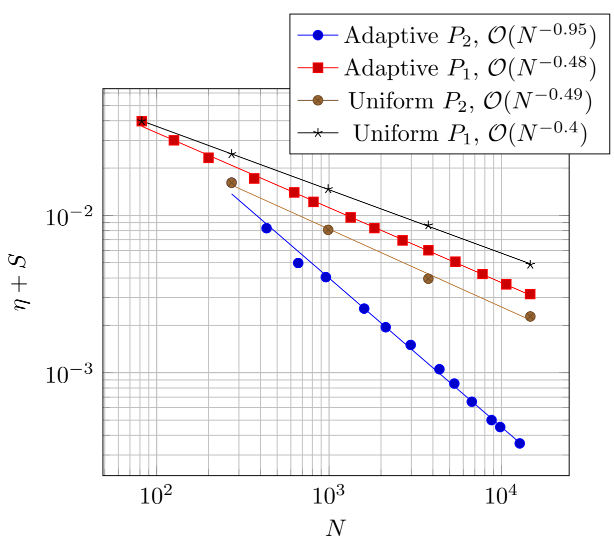
In Figure 5, we demostrate how the improved convergence rates can be obtained for P_2 elements even if the value of the Young's modulus changes significantly over the contact boundary. In Figure 6, we demonstrate that the effect of the stabilization



(a) P_1 after 8 adaptive refinements.

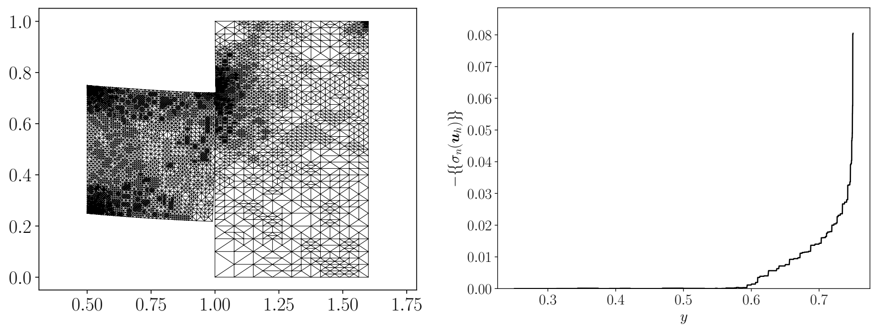


(b) P_2 after 8 adaptive refinements.

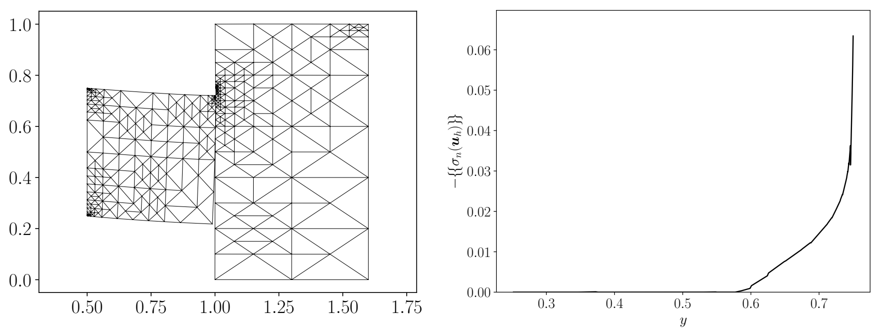


(c) The convergence rates of the total error estimator $\eta + S$ as a function of the number of degrees-of-freedom N.

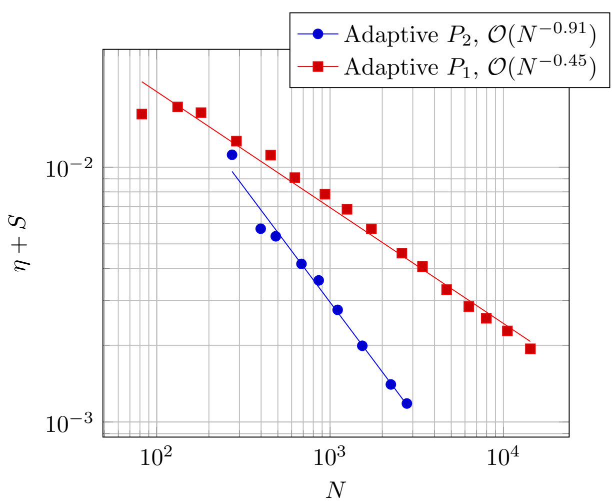
FIG. 3. *Block against a block example.*



(a) P_1 after 15 adaptive refinements.

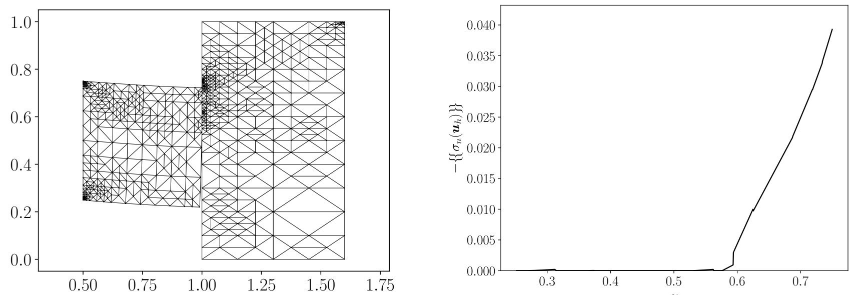


(b) P_2 after 8 adaptive refinements.

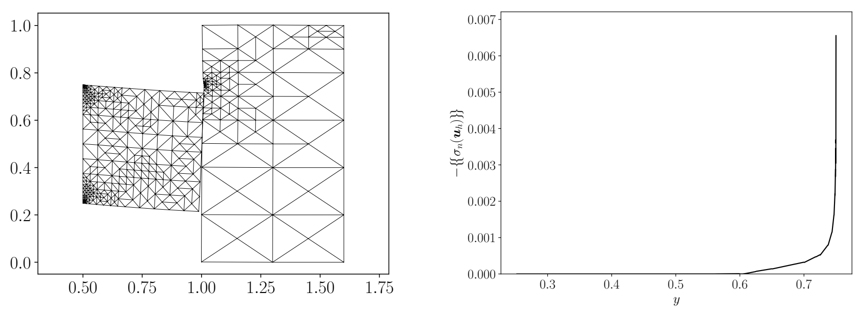


(c) The convergence rates of the total error estimator $\eta + S$ as a function of the number of degrees-of-freedom N.

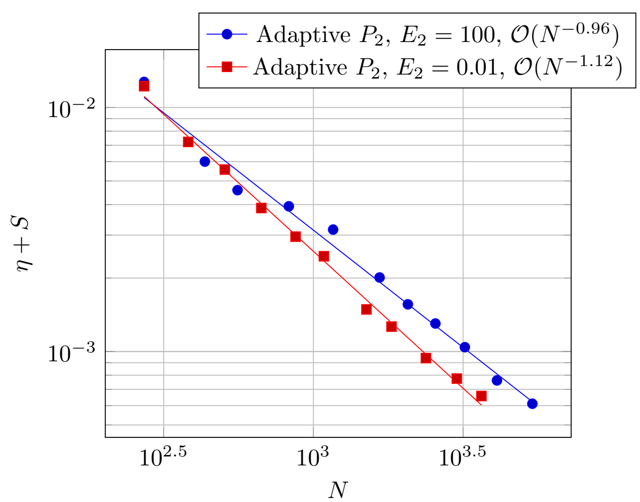
FIG. 4. *Downward bending block example.*



(a) P_2 after 10 adaptive refinements with $E_2 = 100$.

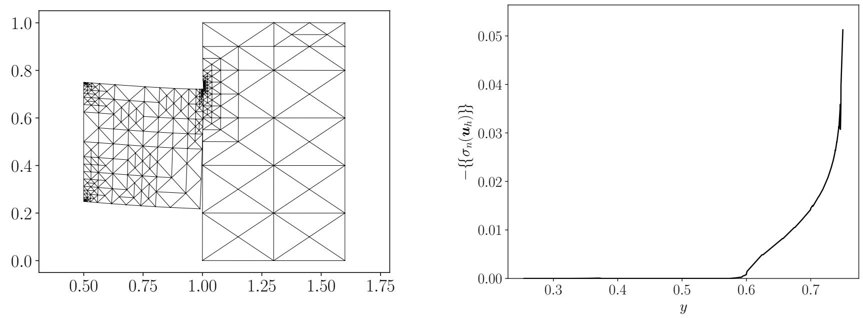


(b) P_2 after 10 adaptive refinements with $E_2 = 0.01$.

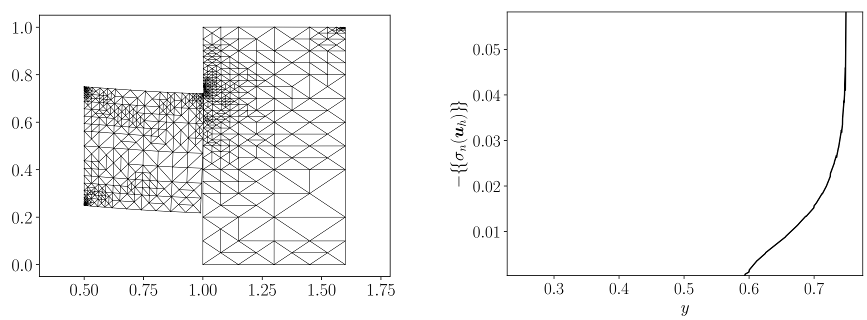


(c) The convergence rates of the total error estimator $\eta + S$ as a function of the number of degrees-of-freedom N.

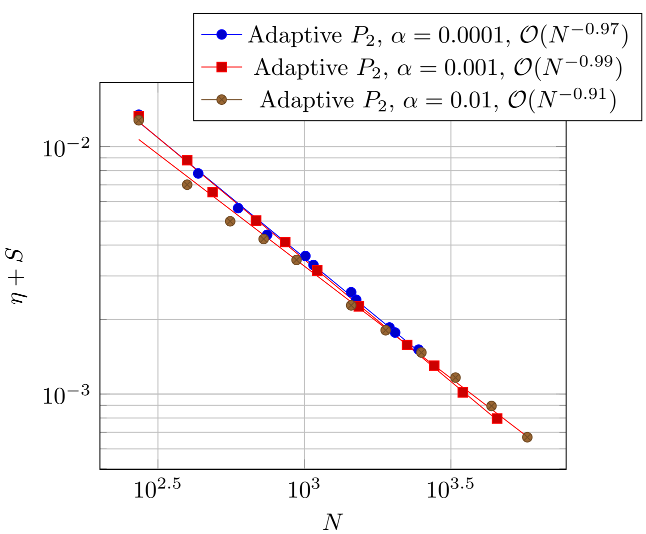
FIG. 5. *Effect of a jump in the Young's modulus.*



(a) P_2 after 10 adaptive refinements with $\alpha = 0.0001$.



(b) P_2 after 10 adaptive refinements with $\alpha = 0.01$.



(c) The convergence rates of the total error estimator $\eta + S$ as a function of the number of degrees-of-freedom N.

FIG. 6. *Effect of changing the stabilization parameter.*

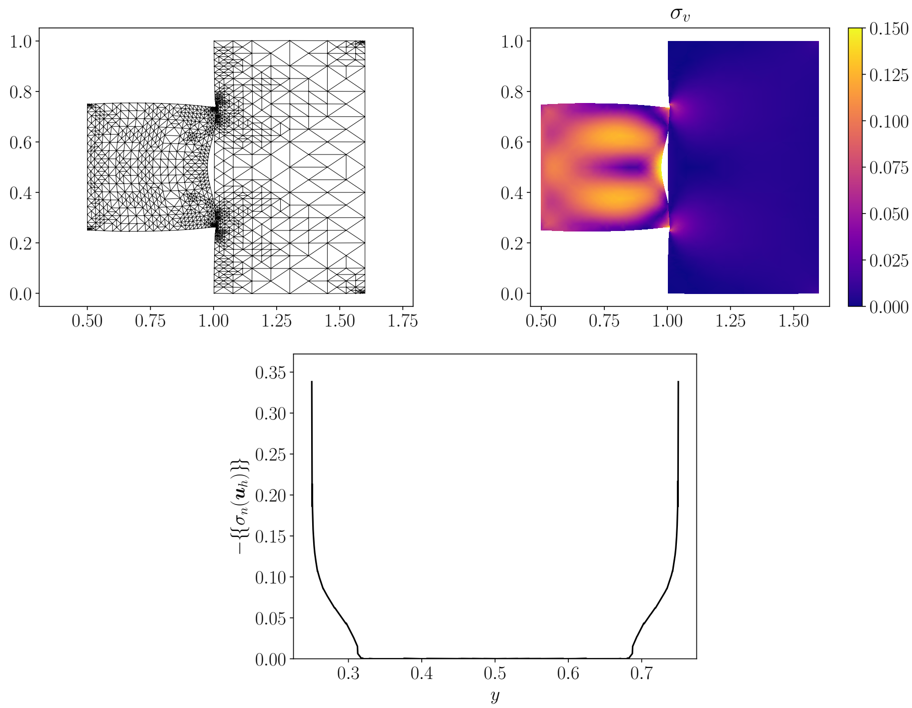


FIG. 7. *Example with a contact boundary consisting of two disjoint active sets with von Mises stress σ_v plotted in the top right figure.*

parameter is small in the asymptotic limit. Finally, in Figure 7, we consider the loading

$$\boldsymbol{f}_1 = (-\cos(4\pi(y - 0.5)), 0), \quad \boldsymbol{f}_2 = (0, 0), \tag{5.6}$$

which results in an active contact boundary consisting of two disjoint parts and a perfectly symmetric contact stress.

REFERENCES

[1] I. Babuška, *The finite element method with Lagrangian multipliers*, Numer. Math., 20 (1973), pp. 179–192.

[2] H. J. C. Barbosa and T. J. R. Hughes, *The finite element method with Lagrange multipliers on the boundary: Circumventing the Babuška-Brezzi condition*, Comput. Methods Appl. Mech. Engrg., 85 (1991), pp. 109–128.

[3] H. J. C. Barbosa and T. J. R. Hughes, *Boundary Lagrange multipliers in finite element methods: Error analysis in natural norms*, Numer. Math., 62 (1992), pp. 1–15.

[4] R. Becker, P. Hansbo, and R. Stenberg, *A finite element method for domain decomposition with non-matching grids*, ESAIM Math. Model. Numer. Anal., 37 (2003), pp. 209–225.

[5] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, Springer Ser. Comput. Math. 15, Springer-Verlag, New York, 1991.

[6] F. Chouly, M. Fabre, P. Hild, R. Mlika, J. Pousin, and Y. Renard, *An overview of recent results on Nitsche's method for contact problems*, in Geometrically Unfitted Finite Element Methods and Applications, S. Bordas, E. Burman, M. Larson, and M. Olshanskii, eds., Lect. Notes Comput. Sci. Eng. 121, Springer, New York, 2017, pp. 93–141.

[7] F. Chouly, M. Fabre, P. Hild, J. Pousin, and Y. Renard, *Residual-based a posteriori error estimation for contact problems approximated by Nitsche's method*, IMA J. Numer. Anal., 38 (2018), pp. 921–954.

[8] F. Chouly, M. Fabre, P. Hild, J. Pousin, and Y. Renard, *Residual-based a posteriori error estimation for contact problems approximated by Nitsche's method*, IMA J. Numer. Anal., 38 (2018), pp. 921–954.

[9] F. Chouly and P. Hild, *A Nitsche-based method for unilateral contact problems: Numerical analysis*, SIAM J. Numer. Anal., 51 (2013), pp. 1295–1307.

[10] F. Chouly, R. Mlika, and Y. Renard, *An unbiased Nitsche's approximation of the frictional contact between two elastic structures*, Numer. Math., 139 (2018), pp. 593–631.

[11] W. Dahmen, B. Faermann, I. G. Graham, W. Hackbusch, and S. A. Sauter, *Inverse inequalities on non-quasi-uniform meshes and application to the mortar element method*, Math. Comp., 73 (2004), pp. 1107–1138.

[12] T. Gustafsson, R. Stenberg, and J. Videman, *Mixed and stabilized finite element methods for the obstacle problem*, SIAM J. Numer. Anal., 55 (2017), pp. 2718–2744.

[13] T. Gustafsson, R. Stenberg, and J. Videman, *On finite element formulations for the obstacle problem – mixed and stabilised methods*, Comput. Methods Appl. Math, 17 (2017), pp. 413–429.

[14] T. Gustafsson, R. Stenberg, and J. Videman, *Error analysis of Nitsche's mortar method*, Numer. Math., 142 (2019), pp. 973–994.

[15] J. Haslinger, I. Hlaváček, and J. Nečas, *Numerical methods for unilateral problems in solid mechanics*, in Finite Element Methods (Part 2), Numerical Methods for Solids (Part 2), P. Ciarlet and J. Lions, eds., Handb. Numer. Anal. 4, North-Holland, Amsterdam, 1996, pp. 313–485.

[16] I. Hlaváček, J. Haslinger, J. Nečas, and J. Lovíšek, *Solution of Variational Inequalities in Mechanics*, Springer, New York, 1988.

[17] S. Hüeber and B. Wohlmuth, *An optimal a priori error estimate for nonlinear multibody contact problems*, SIAM J. Numer. Anal., 43 (2005), pp. 156–173.

[18] M. Juntunen, *On the connection between the stabilized Lagrange multiplier and Nitsche's methods*, Numer. Math., 131 (2015), pp. 453–471.

[19] M. Juntunen and R. Stenberg, *Nitsche's method for discontinuous material parameters*, in Proceedings of the 25th Nordic Seminar on Computational Mechanics, K. Persson, J. Revstedt, G. Sandberg, and M. Wallin, eds., Lund University, 2012, pp. 95–98.

[20] R. Krause, A. Veeser, and M. Walloth, *An efficient and reliable residual-type a posteriori error estimator for the Signorini problem*, Numer. Math., 130 (2015), pp. 151–197.

[21] T. A. Laursen, *Computational Contact and Impact Mechanics*, Springer, New York, 2003.

[22] J. Nitsche, *Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind*, Abh. Math. Sem. Univ. Hambg., 36 (1971), pp. 9–15.

[23] R. Stenberg, *On some techniques for approximating boundary conditions in the finite element method*, J. Comput. Appl. Math., 63 (1995), pp. 139–148.

[24] R. Stenberg, *Mortaring by a method of J. A. Nitsche*, in Computational Mechanics – New Trends and Applications, S. Idelsohn, E. Oñate, and E. Dvorkin, eds., CIMNE, Barcelona, 1998.

[25] L. Tartar, *An Introduction to Sobolev Spaces and Interpolation Spaces*, Lect. Notes Unione Mat. Ital. 3, Springer, Berlin, 2007.

[26] R. Verfürth, *A Posteriori Error Estimation Techniques for Finite Element Methods*, Oxford University Press, Oxford, 2013.

[27] B. Wohlmuth, *Variationally consistent discretization schemes and numerical algorithms for contact problems*, Acta Numer., 20 (2011), pp. 569–734.