ESAIM: M2AN 53 (2019) 173–195
https://doi.org/10.1051/m2an/2018047

ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

AUGMENTED LAGRANGIAN FINITE ELEMENT METHODS FOR CONTACT PROBLEMS

Erik Burman[1,*], Peter Hansbo[2] and Mats G. Larson[3]

Abstract. We propose two different Lagrange multiplier methods for contact problems derived from the augmented Lagrangian variational formulation. Both the obstacle problem, where a constraint on the solution is imposed in the bulk domain and the Signorini problem, where a lateral contact condition is imposed are considered. We consider both continuous and discontinuous approximation spaces for the Lagrange multiplier. In the latter case the method is unstable and a penalty on the jump of the multiplier must be applied for stability. We prove the existence and uniqueness of discrete solutions, best approximation estimates and convergence estimates that are optimal compared to the regularity of the solution.

Mathematics Subject Classification. 65N12, 65N30, 74M15, 74S05.

Received July 21, 2017. Accepted July 18, 2018.

1. Introduction

We consider the Signorini problem, find u and λ such that

$$\begin{aligned} -\Delta u &= f \text{ in } \Omega \\ u &= 0 \text{ on } \Gamma_D \\ u \leq 0,\ \lambda \leq 0,\ u\,\lambda &= 0 \text{ on } \Gamma_C, \end{aligned} \tag{1.1}$$

or the obstacle problem

$$\begin{aligned} -\Delta u - \lambda &= f \text{ in } \Omega \\ u &= 0 \text{ on } \partial\Omega \\ u \leq 0,\ \lambda \leq 0,\ u\,\lambda &= 0 \text{ in } \Omega. \end{aligned} \tag{1.2}$$

Here $\Omega \subset \mathbb{R}^d$, $d = 2,3$ is a bounded polyhedral (polygonal) domain and $f \in L^2(\Omega)$. It is well known that these problems admit unique solutions $u \in H^1(\Omega)$. This follows from the theory of Stampacchia applied to the corresponding variational inequality (see for instance [24]). For the discussion below we will also assume the additional regularity $u \in H^{1+s}(\Omega)$, $\lambda \in H^{1-s}(\Gamma_C)$, $s > 1/2$, for the Signorini problem (1.1) (see [3]) and $u \in H^{1+s}(\Omega)$, $\lambda \in H^{1-s}(\Omega)$, $s \geq 1$, for the obstacle problem (1.2) (see [15]).

Keywords and phrases. Signorini problem, obstacle problem, finite element method, Lagrange mutlipliers, augmented Lagrangian, error estimates.

[1] Department of Mathematics, University College London, Gower Street, UK–WC1E 6BT London, UK.

[2] Department of Mechanical Engineering, Jönköping University, 55111 Jönköping, Sweden.

[3] Department of Mathematics and Mathematical Statistics, Umeå University, 901 87 Umeå, Sweden.

*Corresponding author: e.burman@ucl.ac.uk

From a mechanical point of view, these equations model the deflection of a membrane in isotropic tension under the load f, assuming small deformations. The membrane is either in contact with an obstacle on part of the boundary, (1.1), or in the interior of the membrane, (1.2), preventing positive displacements u. In both cases the Lagrange multiplier has the interpretation of a distributed reaction force enforcing the contact condition $u \leq 0$. We present the numerical analysis in the framework of the simplified model above, but there are no conceptual differences when working with more realistic models of elasticity (in which case friction can also be considered, *cf.* [16,29]). Extensions to adhesive contact models are given in [13].

2. Finite element discretization

Our aim in this paper is to design a consistent penalty method for contact problems that can easily be included in a standard Lagrange-multiplier method, without having to resort to the solution of variational inequalities. We consider two different choices for the multiplier spaces, either a stable choice or an unstable choice where a stabilization term is needed to ensure the stability of the formulation. In the latter case we add a penalty on the jump of the multiplier over element faces in the spirit of [11,12].

There exists a large body of litterature treating finite element methods for contact problems [5–7,9,19,22,28, 35,36]. Discretization of (1.1) is usually performed on the variational inequality or using a penalty method. The first case however leads to some nontrivial choices in the construction of the discretization spaces in order to satisfy the nonpenetration condition and associated inf-sup conditions and until recently it has proved difficult to obtain optimal error estimates [21,26]. The latter case, on the other hand leads to the usual consistency and conditioning problems of penalty methods. Another approach proposed by Hild and Renard [25] is to use a stabilized Lagrange-multiplier in the spirit of Barbosa and Hughes [4]. As a further development one may use the reformulation of the contact condition

$$\lambda = -\gamma^{-1}[u - \gamma\lambda]_+ \tag{2.1}$$

where $[x]_+ = \max(0, x)$, introduced by Alart and Curnier [1] in an augmented Lagrangian framework. Using the close relationship between the Barbosa–Hughes method and Nitsche's method [30] discussed by Stenberg [32], this method was then further developed in the elegant Nitsche-type formulation for the Signorini problem introduced by Chouly, Hild and Renard [18,20]. In these works optimal error estimates for the above model problem were obtained for the first time. For an overview, see [17].

Using the notation $\langle u, v\rangle_C$ for the L^2 inner product over C we have in the case of the Signorini problem (1.1) that C corresponds to Γ_C, the boundary part where the contact conditions hold and

$$\langle u, v\rangle_C := \int_{\Gamma_C} uv \, \mathrm{d}s,$$

while for the obstacle problem (1.2) $C \equiv \Omega$ and

$$\langle u, v\rangle_C := \int_{\Omega} uv \, \mathrm{d}x.$$

Finally, we define $\|v\|_C := \langle v, v\rangle_C^{1/2}$. With this notation, the augmented Lagrangian multiplier seeks the stationary point to the functional

$$\mathfrak{F}(u, \lambda) := \frac{1}{2}a(u, u) + \frac{1}{2\gamma}\|[u - \gamma\lambda]_+\|_C^2 - \frac{\gamma}{2}\|\lambda\|_C^2, \tag{2.2}$$

where γ is a positive parameter, *cf.* Alart and Curnier [1], and $a(u, v) := (\nabla u, \nabla v)_\Omega$. Observe that formally the stationary point is given by (u, λ) such that

$$\begin{aligned} a(u, v) + \langle \gamma^{-1}[u - \gamma\lambda]_+, v\rangle_C &= (f, v)_\Omega \\ \langle \gamma\lambda + [u - \gamma\lambda]_+, \mu\rangle_C &= 0 \end{aligned} \tag{2.3}$$

for all (v, μ), or by substituting the second equation in the first

$$\begin{aligned} a(u,v) - \langle \lambda, v\rangle_C &= (f,v)_\Omega \\ \langle \gamma\lambda + [u-\gamma\lambda]_+, \mu\rangle_C &= 0. \end{aligned} \tag{2.4}$$

It follows that under our regularity assumptions any solution to (1.1) or (1.2) also solves the Euler-Lagrange equations (2.3) and (2.4). Observing now that the contact condition equally well can be written on the primal variable as $u = -[\gamma\lambda - u]_+$ we get by adding and subtracting u in the second equation of (2.4)

$$\begin{aligned} a(u,v) - \langle \lambda, v\rangle_C &= (f,v)_\Omega \\ \langle u + [\gamma\lambda - u]_+, \mu\rangle_C &= 0. \end{aligned} \tag{2.5}$$

In this paper we consider two different methods, resulting from this approach. The first formulation is the straightforward discretization of (2.3) resulting in a method that gives the stationary point of the functional (2.2) over the discrete spaces. The second formulation is a discretization of (2.5) that is chosen for its closeness to the standard Lagrange multiplier method for the imposition of Dirichlet boundary conditions.

The augmented Lagrangian approach is well known as a solution procedure for variational inequalities, see for instance [27], however our objective herein is to show that it may also be considered as a discretization method that yields optimally convergent approximations to contact problems for sufficiently smooth solutions.

We consider discretization either with a choice of approximation spaces that results in a stable approximation, or a choice that is stable only with an added stabilizing term. Here we consider stabilization based on the interior penalty stabilized Lagrange multiplier method introduced by Burman and Hansbo [12] for solving elliptic interface problems. The appeal of this latter approach is that we may use the lowest order approximation spaces where the displacement is piecewise linear and the multiplier constant per element (or element side). When considering the Signorini problem (1.1) these spaces match the regularity of the physical problem perfectly and therefore in some sense is the most economical choice.

For an alternative stabilization method of Barbosa–Hughes type in the augmented Lagrangian setting, see Hansbo, Rashid, and Salomonsson [23].

We assume for simplicity that $\{\mathcal{T}\}_h$ is a family of quasiuniform meshes of Ω (the extension to locally quasi uniform meshes is straightforward), such that the mesh is fitted to the zone C. That is, for the Signorini problem, $C \subset \Gamma_C$ is a subset of boundary element faces of simplices K such that $K \cap \Gamma_C \neq \emptyset$, $F := \partial K \cap \Gamma_C$ $\mathcal{T}_C := \{F\}$, $C := \cup_{F \in \mathcal{T}_C}$ with $C \subset \mathbb{R}^{d-1}$. For the obstacle problem C is defined by Ω and hence $\cup_{K\in\mathcal{T}} =: C \subset \mathbb{R}^d$ and $\mathcal{T}_C \equiv \mathcal{T}$. Below we will denote the elements of $\mathcal{T}_C$ by K in both cases. We define V_h to be the space of H^1-conforming functions on $\mathcal{T}$, satisfying the homogeneous boundary condition of Γ_D.

$$V_h^k := \{v_h \in H^1(\Omega) : v|_{\Gamma_D} = 0; v|_K \in \mathbb{P}_k(K),\ \forall K \in \mathcal{T}\},$$

where $\mathbb{P}_k(K)$ denotes the set of polynomials of order less than or equal to k on the simplex K. Whenever the superscript is dropped we refer to the generic space of order k. For the multipliers we introduce the space Λ_h defined as the space piecewise polynomials of order less than or equal to l defined on C.

$$\Lambda_h := \{\mu_h \in L^2(C) : \mu_h|_K \in \mathbb{P}_{k-1}(K), \forall K \in \mathcal{T}_C\}.$$

Whenever $l = k-1$ the superscript is dropped. We will detail the case of discontinuous multipliers, but all arguments below are valid also in case the Lagrange multiplier is approximated in the space of continuous functions, $\Lambda_h^l \cap C^0(C)$, $l \geq 1$, in this case no stabilization is necessary. The differences in the analysis will be outlined.

Both formulations that we consider herein take the form: Find $(u_h, \lambda_h) \in V_h \times \Lambda_h$ such that

$$a(u_h, v_h) + b[(u_h,\lambda_h);(v_h,\mu_h)] = (f, v_h)_\Omega \quad \forall (v_h, \mu_h) \in V_h \times \Lambda_h \tag{2.6}$$

where $(\cdot,\cdot)_\Omega$ denotes the standard L^2-inner product, and the methods are distinguished by the definition of the form $b[\cdot;\cdot]$ that acts only in the zone where contact may occur. The stabilization term will be included in the form $b[\cdot;\cdot]$. As already pointed out this term is necessary only when the choice $V_h \times \Lambda_h$, does not satisfy the inf-sup condition. In our framework, this is the case when the multiplier is discontinuous over element faces. In this paper we will focus on a stabilization using a penalty on the jumps over element faces of the multiplier variable in the spirit of [11, 12],

$$s(\lambda_h, \mu_h) := \sum_{F \in \mathcal{F}_C} \delta\gamma \int_F h[\![\lambda_h]\!][\![\mu_h]\!] \, \mathrm{d}s, \tag{2.7}$$

where $\delta > 0$ is a parameter, $[\![x]\!]|_F$ denotes the jump of the quantity x over the face F and $\mathcal{F}_C$ denotes the set of interior element faces of the elements in $\mathcal{T}_C$. The semi-norm associated with the stabilization operator will be defined as $|\cdot|_s := s(\cdot,\cdot)^{\frac{1}{2}}$.

We will also below use the compact notation

$$A_h[(u_h, \lambda_h), (v_h, \mu_h)] := a(u_h, v_h) + b[(u_h, \lambda_h); (v_h, \mu_h)]$$

and the associated formulation, find $(u_h, \lambda_h) \in V_h \times \Lambda_h$ such that

$$A_h[(u_h, \lambda_h), (v_h, \mu_h)] = (f, v_h)_\Omega, \text{ for all } (v_h, \mu_h) \in V_h \times \Lambda_h. \tag{2.8}$$

We will now specify two different choices of $b[\cdot;\cdot]$ leading to two different Lagrange-multiplier methods.

Formulation 1: In the first formulation we use the original formula for the contact condition proposed by Alart and Curnier, $\lambda = -\gamma^{-1}[u - \gamma\lambda]_+$

$$\begin{aligned} b[(u_h, \lambda_h); (v_h, \mu_h)] := & \left\langle \gamma^{-1}[u_h - \gamma\lambda_h]_+, v_h \right\rangle_C \\ & + \left\langle \gamma^{-1}[u_h - \gamma\lambda_h]_+, \gamma\mu_h \right\rangle_C \\ & + \langle \gamma\lambda_h, \mu_h \rangle_C + s(\lambda_h, \mu_h) \end{aligned} \tag{2.9}$$

or, changing the sign of μ we may write the nonlinearity as the derivative of a quadratic form. Here, using the notation $P_{\gamma\pm}(u_h, \lambda_h) := \pm(u_h - \gamma\lambda_h)$

$$\begin{aligned} b[(u_h, \lambda_h); (v_h, \mu_h)] := & \left\langle \gamma^{-1}[P_{\gamma+}(u_h, \lambda_h)]_+, P_{\gamma+}(v_h, \mu_h) \right\rangle_C \\ & - \langle \gamma\mu_h, \lambda_h \rangle_C - s(\lambda_h, \mu_h), \end{aligned} \tag{2.10}$$

with $\gamma > 0$ a parameter to determine. In this case the finite element formulation corresponds to the approximate solutions of (2.3) in the finite element space.

Formulation 2: In the second formulation we use a reformulation of the contact condition on the displacement variable, $u = -[\gamma\lambda - u]_+$ to obtain the semi-linear form

$$\begin{aligned} b[(u_h, \lambda_h); (v_h, \mu_h)] := & - \langle \lambda_h, v_h \rangle_C + \langle \mu_h, u_h \rangle_C \\ & + \langle \mu_h, [P_{\gamma-}(u_h, \lambda_h)]_+ \rangle_C + s(\lambda_h, \mu_h), \end{aligned} \tag{2.11}$$

with $\gamma > 0$ a parameter to determine. In this case the finite element formulation corresponds to the approximate solutions of (2.5) in the finite element space.

2.1. Alternative formulations

In both Formulation 1 and 2 above it is possible to derive an alternative formulation of the same method using the relation

$$[P_{\gamma-}(u_h,\lambda_h)]_+ = [P_{\gamma+}(u_h,\lambda_h)]_+ - P_{\gamma+}(u_h,\lambda_h).$$

Considering the form (2.10) and adding and subtracting $P_{\gamma+}(u_h,\lambda_h)$ in the nonlinear term we have the alternative form (omitting the stabilization term)

$$\begin{aligned} b[(u_h,\lambda_h);(v_h,\mu_h)] &= \left\langle \gamma^{-1}[P_{\gamma+}(u_h,\lambda_h)]_+, P_{\gamma+}(v_h,\mu_h)\right\rangle_C - \langle \gamma\mu_h,\lambda_h\rangle_C \\ &= \left\langle \gamma^{-1}([P_{\gamma+}(u_h,\lambda_h)]_+ - P_{\gamma+}(u_h,\lambda_h)), P_{\gamma+}(v_h,\mu_h)\right\rangle_C \\ &\quad + \left\langle \gamma^{-1}P_{\gamma+}(u_h,\lambda_h), P_{\gamma+}(v_h,\mu_h)\right\rangle_C - \langle \gamma\mu_h,\lambda_h\rangle_C \\ &= -\langle \lambda_h, v_h\rangle_C - \langle \mu_h, u_h\rangle_C + \gamma^{-1}\langle u_h, v_h\rangle_C \\ &\quad + \left\langle \gamma^{-1}([P_{\gamma-}(u_h,\lambda_h)]_+, P_{\gamma+}(v_h,\mu_h)\right\rangle_C. \end{aligned} \tag{2.12}$$

Similarly for Formulation 2 we obtain in (2.11) omitting for simplicity the stabilization term

$$b[(u_h,\lambda_h);(v_h,\mu_h)] = -\langle \lambda_h, v_h\rangle_C + \gamma\langle \mu_h,\lambda_h\rangle_C + \langle \mu_h, [P_{\gamma+}(u_h,\lambda_h)]_+\rangle_C. \tag{2.13}$$

We see that this semi-linear form corresponds to a discretization of (2.4).

The methods defined by (2.11) and (2.13) or (2.10) and (2.12) respectively are equivalent, but if during the solution process the linear and nonlinear parts are separated in the nonlinear solver, one can expect the different formulations to have different behavior and give rise to different sequences of approximations in the iterative procedure.

3. Technical results

Here we will collect some useful elementary results. We will frequently make use of the notation $a \lesssim b$ for $a \le cb$ with c some positive constant. First recall the following inverse inequalities and trace inequalities (for a proof see, *e.g.*, [2])

$$\|\nabla u_h\|_K \le c_I h^{-1}\|u_h\|_K, \quad \forall u_h \in V_h \tag{3.1}$$

$$\|u\|_{\partial K} \le c_T(h^{-\frac12}\|u\|_K + h^{\frac12}\|\nabla u\|_K), \quad \forall u \in H^1(K) \tag{3.2}$$

$$\|u_h\|_{\partial K} \le c_T h^{-\frac12}\|u_h\|_K, \quad \forall u_h \in V_h. \tag{3.3}$$

Similar inequalities hold for functions in Λ_h and we will use them without making any distinction between the two cases. We let $\pi_0 : L^2(C) \to \Lambda_h^0$ denote the standard L^2 projection onto Λ_h^0 and we observe that there holds, by standard approximation properties of the projection onto constants (and a trace inequality in the case of lateral contact),

$$\|(1-\pi_0)v_h\|_C \le c_0 h^s \|\nabla v_h\|_\Omega \tag{3.4}$$

with $s = 1$ for the Obstacle problem where $C \subset \Omega$ and $s = \frac12$ for the Signorini problem where $C \subset \partial\Omega$. Similarly we define $\pi_l : L^2(C) \to \Lambda_h^l \cap C^0(\bar C)$ and note that the corresponding inequality holds for π_1

$$\|(1-\pi_1)v_h\|_C \le c_1 h^s \|\nabla v_h\|_\Omega. \tag{3.5}$$

We also observe for future reference that $\|u\|_C \lesssim \|u\|_{H^1(\Omega)}$ in both cases.

For the analysis below it is useful to introduce an indicator function for the contact domain C defined on the space V_h. Let ξ_h denote a finite element function such that $\xi_h \in V_h^1$ with $\xi_h(x) = 0$ for nodes in $(\bar\Omega \setminus \bar C) \cup \bar\Gamma_D$, that is nodes outside the contact zone. For all other nodes $x_i \in K$ with $K \subset \mathcal{T}_C$, $x_i \notin \bar{\Gamma}_D$, $\xi_h(x_i) = 1$. The following bound is well known, see for instance [14]

$$\exists c_\xi \in \mathbb{R}^+ \text{ such that } c_\xi \|\mu_h\|_C \leq \|\xi_h^{\frac{1}{2}} \mu_h\|_C, \quad \forall \mu_h \in \Lambda_h^l, \quad l \geq 0. \tag{3.6}$$

Stability of the method will rely on the satisfaction of the following assumption:

Assumption 3.1. *There exists* $c_D \in [0, 1)$ *such that for all* $\mu_h \in \Lambda_h$ *there holds*

$$\|(1 - \xi_h)\mu_h\|_C \leq c_D \|\mu_h\|_C.$$

The assumption holds whenever there exists a quadrature rule on the simplex, with positive weights and only interior quadrature points, that is exact for polynomials of order less than or equal to $l+1$. This is easily shown by observing that since $(1 - \xi_h)\mu_h|_K \in \mathbb{P}_{l+1}(K)$,

$$\begin{aligned}
\|(1 - \xi_h)\mu_h\|_C^2 &= \sum_{K \in \mathcal{T}_C} \sum_{i \in \mathcal{Q}_K} (1 - \xi_h(x_i))^2 \mu_h(x_i)^2 \omega_i \\
&\leq \max_{K \in \mathcal{T}_C} (\max_{i \in \mathcal{Q}_K} (1 - \xi_h(x_i))^2 \sum_{K \in \mathcal{T}} \sum_{i \in \mathcal{Q}_K} (\mu_h(x_i))^2 \omega_i \\
&= c_D^2 \|\mu_h\|_C^2
\end{aligned}$$

where $\mathcal{Q}_K$ is a set of integers indexing the quadrature points in K, ω_i is the weight associated to the point $x_i \in \mathcal{Q}_K$ and

$$c_D \equiv \max_{K \in \mathcal{T}_C} (\max_{i \in \mathcal{Q}_K} (1 - \xi_h(x_i))^2.$$

Since $1 - \xi_h$ is zero only on the boundary of C and no points $x_i \in \mathcal{Q}_K$ are on the boundary we conclude that $c_D < 1$.

This is a very mild condition, on triangles it has been shown to hold for the integration of polynomials of degree at least up to 23, see [33, 37]. It follows that for the Signorini problem in three dimensions and the obstacle problem in two space dimensions the analysis holds at least up to $k = 12$. For the lowest order case where the multipliers are constant per element it is straightforward to show that $c_D \leq 1/2$ if $C \subset \mathbb{R}^2$ and $c_D \leq \frac{1}{3}$ if $C \subset \mathbb{R}^3$.

Lemma 3.2. *Let* $a, b \in \mathbb{R}$*; then there holds*

$$([a]_+ - [b]_+)^2 \leq ([a]_+ - [b]_+)(a - b),$$

$$|[a]_+ - [b]_+| \leq |a - b|.$$

Proof. Expanding the left hand side of the expression we have

$$[a]_+^2 + [b]_+^2 - 2[a]_+[b]_+ \leq [a]_+ a + [b]_+ b - a[b]_+ - [a]_+ b = ([a]_+ - [b]_+)(a - b).$$

For the proof of the second claim, this is trivially true in case both a and b are positive or negative. If a is negative and b positive then

$$|[a]_+ - [b]_+| = |b| \leq |b - a|$$

and similarly if b is negative and a positive

$$|[a]_+ - [b]_+| = |a| \leq |b - a|.$$

□

Lemma 3.3. *(Continuity of $b[\cdot;\cdot]$) The forms (2.11) and (2.10) satisfy*

$$
\begin{aligned}
&|b[(u_1,\lambda_1);(v,\mu)] - b[(u_2,\lambda_2);(v,\mu)]| \\
&\qquad \le (\gamma^{-\frac{1}{2}}\|(u_1-u_2)\|_{H^1(\Omega)} + \gamma^{\frac{1}{2}}\|\lambda_1-\lambda_2\|_C)(\gamma^{-\frac{1}{2}}\|v\|_C + \gamma^{\frac{1}{2}}\|\mu\|_C) \\
&\qquad\qquad + |\lambda_1-\lambda_2|_s|\mu|_s.
\end{aligned}
$$

Proof. Immediate by the definitions of $b[\cdot;\cdot]$, the second inequality of Lemma 3.2, the Cauchy–Schwarz inequality and the assumptions on $\|\cdot\|_C$. $\square$

Next we define the local averaging interpolation operator $I_{cf} : \Lambda_h \to \Lambda_h \cap C^0(C)$ such that for every Lagrangian node $x_i \in \mathcal{T}_C$, with the associated set $\Omega_i := \{K \subset \mathcal{T}_C : x_i \in K\}$

$$
I_{cf}\lambda_h(x_i) = \kappa_i^{-1} \sum_{K\in\Omega_i} \lambda_h(x_i)|_K,
$$

where κ_i denotes the cardinality of the set Ω_i. Observe that since $\xi_h \in V_h^1$, for any $\mu_h \in \Lambda_h$ there are functions R_μ in V_h such that $R_\mu|_C = I_{cf}\xi_h\mu_h$. We recall the following interpolation result between discrete spaces:

Proposition 3.4. *For all $\mu_h \in \Lambda_h$ there holds, for positive c_s and c_{cf},*

$$
\|\xi_h\mu_h - I_{cf}(\xi_h\mu_h)\|_C \le c_s\|h^{\frac{1}{2}}[\![\mu_h]\!]\|_{\mathcal{F}_C}, \quad \|I_{cf}\mu_h\|_C \le c_{cf}\|\mu_h\|_C.
$$

Proof. For a proof we refer to Lemma 5.3 of [10]. $\square$

Lemma 3.5. *Let $r_h \in \Lambda_h \cap C^0(C)$, then there exists $R_h \in V_h$ such that $R_h|_C = \xi_h r_h$ and $\|R_h\|_{H^1(\Omega)} + \|R_h\|_C \le c_R h^{-s}\|r_h\|_C$, with $s = 1/2$ when C is a subset of $\partial\Omega$ and $s = 1$ when C is a subset of Ω.*

Proof. Define R_h so that $R_h(x) = \xi_h r_h(x)$ for all nodes x in $\mathcal{T}_C$ and $R_h(x) = 0$ for all other nodes x in the mesh. First consider the case when C is a subset of the bulk domain Ω. Then, using an inverse inequality,

$$
\|\nabla R_h\|_\Omega \le c_I h^{-1}\|R_h\|_\Omega = c_I h^{-1}\|r_h\|_\Omega = c_I h^{-1}\|r_h\|_C.
$$

In the case C is a subset of the boundary of Ω we observe that

$$
\|\nabla R_h\|_\Omega = \left(\sum_{K\subset\mathcal{T}:\partial K\cap C\neq\emptyset} \|\nabla R_h\|_K^2\right)^{\frac{1}{2}} \lesssim \left(\sum_{K\subset\mathcal{T}:\partial K\cap C\neq\emptyset} h^{-2}\|R_h\|_K^2\right)^{\frac{1}{2}}.
$$

Using that R_h is defined by the nodes in C, combined with the shape regularity of the mesh, we may use the following inverse trace inequality ([10], Lem. 3.1) on every $K : \partial K \cap C \neq \emptyset$,

$$
\|R_h\|_K \lesssim h^{\frac{1}{2}}\|R_h\|_{\partial K\cap C}.
$$

It follows, since $R_h|_C = \xi_h r_h$, that

$$
\|\nabla R_h\|_\Omega \lesssim h^{-1/2}\|R_h\|_C \lesssim h^{-1/2}\|r_h\|_C.
$$

$\square$

4. Existence of unique discrete solution

In the previous works on Nitsche's method for contact problems [18, 20] existence and uniqueness has been proven by using the monotonicity and hemi-continuity of the operator. Here we propose a different approach where we use Brouwer's fixed point theorem to establish existence and the monotonicity of the nonlinearity for uniqueness. To this end we introduce the finite dimensional nonlinear system corresponding to the formulation (2.6).

Let $M := N_V + N_\Lambda$, where N_V and N_Λ denote the number of degrees of freedom of V_h and Λ_h respectively. Then define $U, V \in \mathbb{R}^M$, where $U = \{u_i\}_{i=1}^{N_V} \cup \{\lambda_i\}_{i=1}^{N_\Lambda}$, $V = \{v_i\}_{i=1}^{N_V} \cup \{\mu_i\}_{i=1}^{N_\Lambda}$, where $\{u_i\}, \{v_i\}$ and $\{\lambda_i\}, \{v_i\}$ denote the vectors of unknowns associated to the basis functions of V_h and Λ_h respectively.

Consider the mapping $G : \mathbb{R}^M \mapsto \mathbb{R}^M$ defined by

$$(G(U), V)_{\mathbb{R}^M} := A_h[(u_h, \lambda_h), (v_h, \mu_h)] - (f, v_h)_\Omega.$$

Existence and uniqueness of a solution to (2.6) is equivalent to showing that there exists a unique $U \in \mathbb{R}^M$ such that $G(U) = 0$.

We start by showing some positivity results and *a priori* bounds

Lemma 4.1. *There exists $\alpha > 0$ and an associated constant $c_\alpha > 0$ so that with the form b defined by (2.10), $\delta > 0$ and $\gamma = \gamma_0^{-1} h^{2s}$ with $\gamma_0 > 0$ there holds, for all $(u_h, \lambda_h) \in V_h \times \Lambda_h$*

$$\|\nabla u_h\|_\Omega^2 + \gamma \|\lambda_h + \gamma^{-1}[P_{\gamma+}(u_h, \lambda_h)]_+\|_C^2 + c_\alpha \|\gamma^{\frac{1}{2}} \lambda_h\|_C^2 \lesssim A_h[(u_h, \lambda_h), (u_h - \alpha R_h, \lambda_h)], \tag{4.1}$$

where $R_h \in V_h$ is defined in Lemma 3.5, such that $R_h|_C := \gamma \xi_h I_{cf} \lambda_h$.

There exists $\alpha > 0$ and an associated constant $c_\alpha > 0$ so that with the form b defined by (2.11), $k \geq 2$ and $\gamma = \gamma_0 h^{2s}$ with $\gamma_0 > 0$, γ_0 sufficiently large, and $\delta > 0$ there holds, for all $(u_h, \lambda_h) \in V_h \times \Lambda_h$

$$\|\nabla u_h\|_\Omega^2 + \gamma^{-1} \|u_h + [P_{\gamma-}(u_h, \lambda_h)]_+\|_C^2 + c_\alpha \|\gamma^{\frac{1}{2}} \lambda_h\|_C^2 \lesssim A_h[(u_h, \lambda_h), (u_h + \alpha R_h, \lambda_h + \gamma^{-1} \pi_0 u_h)], \tag{4.2}$$

with R_h as before. In case $k = 1$ (4.2) holds under the additional that $0 < \delta \leq (c_0 c_T)^2 \gamma_0^{-1}$.

Under the same conditions on the parameters as above, for both formulations there also holds, for (u_h, λ_h) solution of (2.8),

$$\|\nabla u_h\|_\Omega + \|\gamma^{\frac{1}{2}} \lambda_h\|_C \lesssim \|f\|_\Omega. \tag{4.3}$$

The hidden constants are independent of h.

Remark 4.2. For $k \geq 2$ and continuous multiplier space the parameter δ and the term $|\lambda_h|_s^2$ can be dropped above.

Proof. First consider the claims for formulation 1. By testing in (2.8), using (2.10), with $v_h = u_h$ and $\mu_h = -\lambda_h$

$$A_h((u_h, \lambda_h), (u_h, -\lambda_h)) = \|\nabla u_h\|_\Omega^2 + \langle \gamma^{-1} P_{\gamma+}(u_h, \lambda_h), u_h + \gamma \lambda_h \rangle_C + \|\gamma^{\frac{1}{2}} \lambda_h\|_C^2 + |\lambda_h|_s^2$$

and observing that

$$\begin{aligned} &\left\langle \gamma^{-1}[P_{\gamma+}(u_h, \lambda_h)]_+, \gamma \lambda_h \right\rangle_C + \|\gamma^{\frac{1}{2}} \lambda_h\|_C^2 \\ &\qquad = \|\gamma^{\frac{1}{2}} \lambda_h\|_C^2 + \left\langle \gamma^{-\frac{1}{2}}[P_{\gamma+}(u_h, \lambda_h)]_+, -\gamma^{\frac{1}{2}} \lambda_h \right\rangle_C + 2 \left\langle \gamma^{-\frac{1}{2}}[P_{\gamma+}(u_h, \lambda_h)]_+, \gamma^{\frac{1}{2}} \lambda_h \right\rangle_C \end{aligned}$$

implies

$$\langle \gamma^{-1} P_{\gamma+}(u_h, \lambda_h), u_h + \gamma \lambda_h \rangle_C + \|\gamma^{\frac{1}{2}} \lambda_h\|_C^2 = \|\gamma^{\frac{1}{2}} \lambda_h\|_C^2 + \|\gamma^{-\frac{1}{2}}[P_{\gamma+}(u_h, \lambda_h)]_+\|_C^2 + 2 \left\langle \gamma^{-\frac{1}{2}}[P_{\gamma+}(u_h, \lambda_h)]_+, \gamma^{\frac{1}{2}} \lambda_h \right\rangle_C$$

we obtain the relation

$$\|\nabla u_h\|_\Omega^2 + \|\gamma^{-\frac{1}{2}}[P_{\gamma+}(u_h,\lambda_h)]_+ + \gamma^{\frac{1}{2}}\lambda_h\|_C^2 + |\lambda_h|_s^2 = A_h[(u_h,\lambda_h),(u_h,\lambda_h)]. \tag{4.4}$$

Testing (2.8) with $v_h = -\alpha R_h$, with $r_h = \gamma I_{cf}(\xi_h \lambda_h)$ and $\mu_h = 0$,

$$\begin{aligned} A_h[(u_h,\lambda_h),(-\alpha R_h,0)] &= a(u_h,-\alpha R_h) + \left\langle \gamma^{-1}[P_{\gamma+}(u_h,\lambda_h)]_+, -\gamma I_{cf}(\xi_h\lambda_h)\right\rangle_C \\ &= a(u_h,-\alpha R_h) + \left\langle \gamma^{-1}[P_{\gamma+}(u_h,\lambda_h)]_+ + \lambda_h, -\gamma I_{cf}(\xi_h\lambda_h)\right\rangle_C \\ &\quad - \langle \lambda_h, \gamma(\xi_h\lambda_h - I_{cf}(\xi_h\lambda_h))\rangle_C + \langle \lambda_h, \gamma\xi_h\lambda_h\rangle_C . \end{aligned} \tag{4.5}$$

For the last term in the right hand side we have by the inequality (3.6), $c_\xi^2\|\gamma^{\frac{1}{2}}\lambda_h\|_C^2 \le (\gamma\lambda_h, \xi_h\lambda_h)_C$. The second to last term of the right hand side, which is zero for continuous multiplier spaces, can be bounded using Proposition 3.4

$$(\gamma\lambda_h, \xi_h\lambda_h - I_{cf}(\xi_h\lambda_h))_C \le c_\xi^2 \frac{1}{4}\|\gamma^{\frac{1}{2}}\lambda_h\|_C^2 + c_s^2 c_\xi^{-2}\delta^{-1}|\lambda_h|_s^2. \tag{4.6}$$

The second term is bounded using a Cauchy–Schwarz inequality and the stability of I_{cf},

$$\left\langle \gamma^{-1}[P_{\gamma+}(u_\lambda,\lambda_h)]_+ + \lambda_h, \gamma I_{cf}(\xi_h\lambda_h)\right\rangle_C \le \frac{1}{2}(c_{cf}c_\xi)^{-2}\|\gamma^{-\frac{1}{2}}[P_{\gamma+}(u_\lambda,\lambda_h)]_+ + \gamma^{\frac{1}{2}}\lambda_h\|_C^2 + \frac{1}{4}c_\xi^2\|\gamma^{\frac{1}{2}}\lambda_h\|_C^2 \tag{4.7}$$

for the first term we use the Cauchy–Schwarz inequality followed by the stability of R_h, Lemma 3.5, and of I_{cf} to obtain

$$a(u_h, R_h) \le c_R^2 h^{-2s}\gamma c_{cf}^2 c_\xi^{-2}\|\nabla u_h\|_\Omega^2 + c_\xi^2\frac{1}{4}\|\gamma^{\frac{1}{2}}\lambda_h\|_C^2. \tag{4.8}$$

Applying the inequalities (4.6)–(4.8) to (4.5) we have

$$\begin{aligned} &c_\xi^2\alpha\|\gamma^{\frac{1}{2}}\lambda_h\|_C^2 - c_R^2 h^{-2s}\gamma c_{cf}^2 c_\xi^{-2}\alpha\|\nabla u_h\|_\Omega \\ &\qquad - \frac{1}{2}(c_{cf}c_\xi)^{-2}\alpha\|\gamma^{-\frac{1}{2}}[P_{\gamma+}(u_\lambda,\lambda_h)]_+ + \gamma^{\frac{1}{2}}\lambda_h\|_C^2 \\ &\qquad - c_s^2 c_\xi^{-2}\delta^{-1}\alpha|\lambda_h|_s^2 \\ &\quad \le A_h[(u_h,\lambda_h),(-\alpha R_h,0)]. \end{aligned} \tag{4.9}$$

We conclude that (4.1) holds, by observing that $h^{-2s}\gamma = O(1)$ and by combining the bounds (4.4) and (4.9) with α small enough. The *a priori* estimate (4.3) follows noting that for (u_h,λ_h) solution of (2.6) there holds using the Poincaré inequality and the properties of R_h,

$$A_h[(u_h,\lambda_h),(u_h-\alpha R_h,0)] = (f, u_h - \alpha R_h) \lesssim \|f\|_\Omega(\|\nabla u_h\|_\Omega + \|\gamma^{\frac{1}{2}}\lambda_h\|_C).$$

To prove (4.2) we start by testing in the left hand side of (2.8) with $v_h = u_h$ and $\mu_h = \lambda_h + \gamma^{-1}\pi_i u_h = \gamma^{-1}P_{\gamma-}(u_h,\lambda_h) + \gamma^{-1}(u_h + \pi_i u_h)$, where $i=0$ if $k=1$ and $i=1$ for $k\ge 2$. Observing this time that, using the definition (2.11) and adding and subtracting u_h at suitable places the following equality holds

$$\begin{aligned} b[(u_h,\lambda_h),(u_h,\lambda_h+\gamma^{-1}\pi_i u_h)] &= \gamma^{-1}\langle u_h,u_h\rangle_C - \gamma^{-1}\|\pi_i u_h - u_h\|_C^2 \\ &\quad + \gamma^{-1}\|[P_{\gamma-}(u_h,\lambda_h)]_+\|_C^2 \\ &\quad + 2\left\langle \gamma^{-1}u_h, [P_{\gamma-}(u_h,\lambda_h)]_+\right\rangle_C \\ &\quad + \gamma^{-1}\langle \pi_i u_h - u_h, [P_{\gamma-}(u_h,\lambda_h)]_+\rangle_C \\ &\quad + |\lambda_h|_s^2 + s(\lambda_h, \gamma^{-1}(\pi_i u_h - u_h)). \end{aligned}$$

This results in

$$\|\nabla u_h\|_\Omega^2 + \gamma^{-1}\|u_h + [P_{\gamma-}(u_h,\lambda_h)]_+\|_C^2 + |\lambda_h|_s^2 - \gamma^{-1}\|\pi_i u_h - u_h\|_C^2 \\ + \gamma^{-1}\left\langle \pi_i u_h - u_h, u_h + [P_{\gamma-}(u_h,\lambda_h)]_+\right\rangle_C + s(\lambda_h, \gamma^{-1}(\pi_i u_h - u_h)) \\ = A_h[(u_h,\lambda_h),(v_h,\mu_h)].$$

We now bound the three last terms on the left hand side. First by the properties of π_i we have

$$\gamma^{-1}\|\pi_i u_h - u_h\|_C^2 \le c_i^2 h^{2s}\gamma^{-1}\|\nabla u_h\|_\Omega^2, \quad i = 0,1. \tag{4.10}$$

Using a Cauchy–Schwarz inequality, the previous result and an arithmetic-geometric inequality we have

$$\gamma^{-1}\left\langle \pi_i u_h - u_h, u_h + [P_{\gamma-}(u_h,\lambda_h)]_+\right\rangle_C \le \frac{1}{2}c_i^2 h^{2s}\gamma^{-1}\|\nabla u_h\|_\Omega^2 + \frac{1}{2}\gamma^{-1}\|u_h + [P_{\gamma-}(u_h,\lambda_h)]_+\|_C^2. \tag{4.11}$$

Finally for $k = 1$ we have for the last term

$$s(\lambda_h, \gamma^{-1}(\pi_0 u_h - u_h)) \le \frac{1}{2}|\lambda_h|_s^2 + \gamma^{-1}\delta c_T^2\|\pi_0 u_h - u_h\|_C^2 \le \frac{1}{2}|\lambda_h|_s^2 + \frac{1}{2}\frac{\delta c_0^2 c_T^2}{\gamma_0}\|\nabla u_h\|_\Omega^2$$

and for $k \ge 2$, $s(\lambda_h, \gamma^{-1}(\pi_1 u_h - u_h)) = 0$. Collecting the results above we obtain for $k = 1$

$$(1 - 3/2c_0^2\gamma_0^{-1} - 1/2\delta c_0^2 c_T^2\gamma_0^{-1})\|\nabla u_h\|_\Omega^2 + \frac{1}{2}\gamma^{-1}\|u_h + [P_{\gamma-}(u_h,\lambda_h)]_+\|_C^2 + \frac{1}{2}|\lambda_h|_s^2 \\ \lesssim A_h[(u_h,\lambda_h),(v_h,\mu_h)]. \tag{4.12}$$

We see that the factor $(1 - 3/2c_0^2\gamma_0^{-1} - 1/2\delta c_0^2 c_T^2\gamma_0^{-1})$ is positive under the assumptions on γ_0 and δ. The corresponding inequality for $k \ge 2$ is obtained by omitting the term with δ and replacing c_0 with c_1. Observe that by using $v_h = R_h$ with $r_h = -\gamma I_{cf}\xi_h\lambda_h$ and $\mu_h = 0$ we have using similar arguments as above

$$\gamma\|\xi_h^{\frac{1}{2}}\lambda_h\|_C^2 - \left\langle \lambda_h, \gamma(\xi_h\lambda_h - I_{cf}\xi_h\lambda_h)\right\rangle_C + a(u_h, R_h) = A_h[(u_h,\lambda_h),(R_h,0)]. \tag{4.13}$$

Using once again (4.6) and (4.8)

$$\frac{1}{2}c_\xi\|\gamma^{\frac{1}{2}}\lambda_h\|_C^2 - c_R^2 h^{-2s}\gamma c_{cf}^2 c_\xi^{-2}\|\nabla u_h\|_\Omega^2 - c_s^2 c_\xi^{-2}\delta^{-1}|\lambda_h|_s^2 \le A_h[(u_h,\lambda_h),(R_h,0)] \tag{4.14}$$

where the stabilization contribution can be dropped whenever continuous approximation is used for the multiplier space. We conclude as in the previous case by combining the bounds (4.14) and (4.12). The *a priori* estimate (4.3) also follows as before. □

Proposition 4.3. *The formulation* (2.8) *using the contact operators* (2.11) *or* (2.10)*, and the same assumptions on the parameters* δ*,* γ *as in Lemma 4.1, admits a unique solution.*

Proof. By the positivity results (4.1) and (4.2) of Lemma 4.1 we have for each method that there exists a linear mapping $B : \mathbb{R}^M \mapsto \mathbb{R}^M$ such that $b_1|U| < |BU| \le b_2|U|$ for some $0 < b_1 \le b_2$ and that for U sufficiently big

$$0 < (G(U), BU). \tag{4.15}$$

We give details regarding the construction of B only in the case of formulation 2 with $k = 1$. The argument for $k \ge 2$, and that for formulation 1, are similar. Let the positive constants $\underline{c}_h$ and $\overline{c}_h$ denote the smallest and the largest eigenvalues respectively of the block diagonal matrix in $\mathbb{R}^{M\times M}$ with diagonal blocks given by $(\nabla\varphi_i, \nabla\varphi_j)_\Omega + \left\langle\gamma^{-1}\varphi_i, \varphi_j\right\rangle_C$, $1 \le i,j \le N_V$ where φ_i, denotes the basis functions for the space V_h and

$\frac{1}{2}\gamma(\psi_i,\psi_j)_C$ where ψ_i, denotes the basis functions for the space Λ_h, $1 \le i,j \le N_\Lambda$ such that, with $\|u_h\|^2_{1,h} := \|\nabla u_h\|^2_\Omega + \|\gamma^{-\frac{1}{2}} u_h\|^2_C$

$$\underline{c}_h |U|^2_{\mathbb{R}^M} \le \|u_h\|^2_{1,h} + \frac{1}{4}\gamma\|\lambda_h\|^2_C \le \bar{c}_h |U|^2_{\mathbb{R}^M}.$$

Recalling the *a priori* bound (4.2), let B denote the transformation matrix such that the finite element function corresponding to the vector BU is the function $(u_h + \alpha R_h, \lambda_h + \gamma^{-1}\pi_0 u_h)$, with R_h defined in Lemma 4.1. First we show that for α sufficiently small, there are constants b_1 and b_2 such that $b_1|U|_{\mathbb{R}^M} \le |BU|_{\mathbb{R}^M} \le b_2|U|_{\mathbb{R}^M}$. This can be seen by adding and subtracting $\alpha R_h$and $\gamma^{-1}\pi_0 u_h$ and using $(a+b)^2 \le 2a^2 + 2b^2$,

$$\begin{aligned}
\|u_h\|^2_{1,h} + \frac{1}{4}\gamma\|\lambda_h\|^2_C &\le 2\|u_h + \alpha R_h\|^2_{1,h} + \frac{1}{2}\gamma\|\lambda_h + \gamma^{-1}\pi_0 u_h\|^2_C \\
&\quad + \frac{1}{2}\|\gamma^{-1/2}\pi_0 u_h\|^2_C + 2\|\alpha R_h\|^2_{1,h} \\
&\le 2\bar{c}_h|BU|^2_{\mathbb{R}^M} + \frac{1}{2}\|u_h\|^2_{1,h} + c\alpha\gamma^2 h^{-2s}\|\lambda_h\|^2_C
\end{aligned}$$

where we have used the properties of R_h from Lemma 3.5. It follows, for α small enough, recalling that $\gamma = O(h^{2s})$, that

$$\frac{1}{2}\underline{c}_h|U|^2_{\mathbb{R}^M} \le (1 - \frac{1}{2})\|u_h\|^2_{1,h} + (\frac{1}{4} - c\alpha)\gamma\|\lambda_h\|^2_C \le 2\bar{c}_h|BU|^2_{\mathbb{R}^M}.$$

Similarly we may prove the upper bound: using that by the properties of R_h and $\pi_0 u_h$ we have

$$\begin{aligned}
\underline{c}_h|BU|^2_{\mathbb{R}^M} &\le \|u_h + \alpha R_h\|^2_{1,h} + \frac{1}{4}\gamma\|\lambda_h + \gamma^{-1}\pi_0 u_h\|^2_C \\
&\le 2\|u_h\|^2_{1,h} + 2\|\alpha R_h\|^2_{1,h} + \frac{1}{2}\gamma\|\lambda_h\|^2_C + \frac{1}{2}\|\gamma^{-1/2}\pi_0 u_h\|^2_C \\
&\lesssim \left(\|u_h\|^2_{1,h} + \frac{1}{4}\gamma\|\lambda_h\|^2_C\right).
\end{aligned}$$

We have shown that (4.15) holds and we observe that, by Lemma 3.3, $G(U)$ is continuous. Existence of a solution to the nonlinear system is then a consequence of Brouwer's fixed point theorem using standard arguments, see for instance ([34], Lem. 1.4, Chap. 2).

Uniqueness is consequence of the positivity results of Lemma 4.1 and the monotonicity of Lemma 3.2. Considering first formulation 1, where the form $b[\cdot;\cdot]$ is given by (2.10), we have

$$\begin{aligned}
\|\nabla(u_1 - u_2)\|^2_\Omega = &-\gamma^{-1}\left\langle [P_{\gamma+}(u_1,\lambda_1)]_+ - [P_{\gamma+}(u_2,\lambda_2)]_+, u_1 - u_2 + \gamma(\lambda_1 - \lambda_2)\right\rangle_C \\
&- \gamma\|\lambda_1 - \lambda_2\|^2_C - |\lambda_1 - \lambda_2|^2_s.
\end{aligned}$$

It follows that, defining

$$|||u,\lambda|||^2 := \|\nabla u\|^2_\Omega + |\lambda|^2_s,$$

$$\begin{aligned}
|||u_1 - u_2, \lambda_1 - \lambda_2|||^2 = &-\gamma\|\lambda_1 - \lambda_2\|^2_C \\
&- \gamma^{-1}\left\langle [P_{\gamma+}(u_1,\lambda_1)]_+ - [P_{\gamma+}(u_2,\lambda_2)]_+, P_{\gamma+}(u_1 - u_2, \lambda_1 - \lambda_2)\right\rangle_C \\
&- \gamma^{-1}\left\langle [P_{\gamma+}(u_1,\lambda_1)]_+ - [P_{\gamma+}(u_2,\lambda_2)]_+, 2\gamma(\lambda_1 - \lambda_2)\right\rangle_C.
\end{aligned}$$

Then, using the monotonicity of Lemma 3.2 we deduce

$$\begin{aligned}
&|||u_1 - u_2, \lambda_1 - \lambda_2|||^2 + \gamma\|\lambda_1 - \lambda_2\|^2_C + \gamma^{-1}\|[P_{\gamma+}(u_1,\lambda_1)]_+ - [P_{\gamma+}(u_2,\lambda_2)]_+\|^2_C \\
&\quad \le -\left\langle [P_{\gamma+}(u_1,\lambda_1)]_+ - [P_{\gamma+}(u_2,\lambda_2)]_+, 2\gamma(\lambda_1 - \lambda_2)\right\rangle_C.
\end{aligned}$$

Therefore

$$|||u_1 - u_2, \lambda_1 - \lambda_2|||^2 + \gamma^{-1}\|\gamma(\lambda_1 - \lambda_2) + [P_{\gamma+}(u_1, \lambda_1)]_+ - [P_{\gamma+}(u_2, \lambda_2)]_+\|_C^2 = 0 \tag{4.16}$$

and $u_1 = u_2$. Repeating the arguments leading to (4.9) on $\lambda_1 - \lambda_2$ and using (4.16) allows us to conclude that $\lambda_1 = \lambda_2$.

In the case of formulation 2 we only give the details for $k \geq 2$, the case $k = 1$ is similar, but we need to handle an additional stabilization term. Assume that (u_1, λ_1) and (u_2, λ_2) solves (2.8) with the contact conditions defined by (2.11).

$$\begin{aligned}\|\nabla(u_1 - u_2)\|_\Omega^2 &= \langle \lambda_1 - \lambda_2, u_1 - u_2\rangle_C \\ &= -\gamma \langle \lambda_1 - \lambda_2, [P_{\gamma-}(u_1, \lambda_1)]_+ - [P_{\gamma-}(u_2, \lambda_2)]_+\rangle_C \\ &\quad - |\lambda_1 - \lambda_2|_s^2.\end{aligned}$$

Observing that with $\mu_h = \gamma^{-1}\pi_1(u_1 - u_2)$ we also have

$$\gamma^{-1}\|\pi_1(u_1 - u_2)\|_C^2 + \gamma^{-1} \langle \pi_1(u_1 - u_2), [P_{\gamma-}(u_1, \lambda_1)]_+ - [P_{\gamma-}(u_2, \lambda_2)]_+\rangle_C = 0$$

and therefore we can write

$$\begin{aligned}&\|\nabla(u_1 - u_2)\|_\Omega^2 + \gamma^{-1}\|u_1 - u_2 + [P_{\gamma-}(u_1, \lambda_1)]_+ - [P_{\gamma-}(u_2, \lambda_2)]_+\|_C^2 + |\lambda_1 - \lambda_2|_s^2 \\ &\quad = \gamma^{-1} \langle (1 - \pi_1)(u_1 - u_2), u_1 - u_2 + [P_{\gamma-}(u_1, \lambda_1)]_+ - [P_{\gamma-}(u_2, \lambda_2)]_+\rangle_C .\end{aligned}$$

By splitting the term in the right hand side using the arithmetic-geometric inequality and using the approximation properties of π_1,

$$\|(1 - \pi_1)(u_1 - u_2)\|_C \leq c_1 h^s \|\nabla(u_1 - u_2)\|_\Omega$$

we may conclude that

$$\begin{aligned}&(1 - \gamma^{-1}c_1^2 h^{2s})\|\nabla(u_1 - u_2)\|_\Omega^2 + \frac{1}{2}\gamma^{-1}\|u_1 - u_2 + [P_{\gamma-}(u_1, \lambda_1)]_+ - [P_{\gamma-}(u_2, \lambda_2)]_+\|_C^2 \\ &\quad + |\lambda_1 - \lambda_2|_s^2 \leq 0.\end{aligned}$$

As a consequence $u_1 = u_2$ when γ_0 is sufficiently large. That $\lambda_1 = \lambda_2$ is immediate from (4.13) since the first equation of (2.4) is linear. □

5. Error estimates

In this section we will prove the main results of the paper which are error estimates for the two methods given by (2.6) with the two contact formulations (2.11) and (2.10). The idea of the proof is to combine the uniqueness argument with a Galerkin type perturbation analysis. Since this result is central to the present work we give full detail for both formulations.

Theorem 5.1. *(Formulation 1) Assume that $u \in H^1(\Omega)$ and $\lambda \in L^2(C)$ satisfies (2.3) and (u_h, λ_h) the solution to (2.6) with (2.9) and $0 < \gamma = \gamma_0^{-1} h^{2s}$, where $s = 1/2$ for the Signorini problem and $s = 1$ for the Obstacle problem. Also assume that $\gamma_0 \in \mathbb{R}^+$ is sufficiently large and $\delta \in \mathbb{R}^+$ sufficiently large. Then there holds for all $(v_h, \mu_h) \in V_h \times \Lambda_h$*

$$\begin{aligned}&\alpha\|u - u_h\|_{H^1(\Omega)}^2 + \gamma\|(\lambda - \lambda_h)\|_C^2 + \gamma\|(\lambda + \gamma^{-1}[P_{\gamma+}(u_h, \lambda_h)]_+)\|_C^2 \\ &\quad \lesssim \frac{1}{\alpha}\|u - v_h\|_{H^1(\Omega)}^2 + \gamma\|(\lambda - \mu_h)\|_C^2 + \gamma^{-1}\|(u - v_h)\|_C^2 + |\mu_h|_s^2.\end{aligned}$$

Proof. Using the coercivity of $a(\cdot,\cdot)$ we may write

$$\begin{aligned}\alpha\|u-u_h\|^2_{H^1(\Omega)} &\le a(u-u_h,u-u_h)\\ &= a(u-u_h,u-v_h)+a(u-u_h,v_h-u_h)\\ &\le \frac{\alpha}{4}\|u-u_h\|^2_{H^1(\Omega)}+\frac{1}{\alpha}\|u-v_h\|^2_{H^1(\Omega)}+a(u-u_h,v_h-u_h).\end{aligned}$$

It follows, using Galerkin orthogonality, that

$$\begin{aligned}a(u-u_h,v_h-u_h) &= \left\langle \gamma^{-1}[P_{\gamma+}(u_h,\lambda_h)]_+ - \gamma^{-1}[P_{\gamma+}(u,\lambda)]_+, v_h-u_h\right\rangle_C\\ &= \left\langle \gamma^{-1}[P_{\gamma+}(u_h,\lambda_h)]_+ - \gamma^{-1}[P_{\gamma+}(u,\lambda)]_+, v_h-u_h+\gamma(\mu_h-\lambda_h)\right\rangle_C\\ &\quad + \langle\gamma(\lambda_h-\lambda),(\mu_h-\lambda_h)\rangle + s(\lambda_h,\mu_h-\lambda_h).\end{aligned} \tag{5.1}$$

First observe that

$$\begin{aligned}\langle\gamma(\lambda_h-\lambda),(\mu_h-\lambda_h)\rangle = &-\|\gamma^{\frac12}(\mu_h-\lambda_h)\|^2_C\\ &+\|\gamma^{\frac12}(\mu_h-\lambda)\|_C\|\gamma^{\frac12}(\mu_h-\lambda_h)\|_C\\ \le &\ (\varepsilon_1-1)\|\gamma^{\frac12}(\mu_h-\lambda_h)\|^2_C+\frac{1}{4\varepsilon_1}\|\gamma^{\frac12}(\mu_h-\lambda)\|^2_C\end{aligned}$$

where we see that the first term can be made negative by choosing ε_1 small enough. Similarly

$$s(\lambda_h,\mu_h-\lambda_h) = -|\mu_h-\lambda_h|^2_s + s(\mu_h,\mu_h-\lambda_h) \le (\varepsilon_2-1)|\mu_h-\lambda_h|^2_s+\frac{1}{4\varepsilon_2}|\mu_h|^2_s$$

where once again the first term on the right hand side can be made negative by choosing ε_2 small. Considering the first term on the right hand side of equation (5.1) we may write

$$\begin{aligned}&\left\langle \gamma^{-1}[P_{\gamma+}(u_h,\lambda_h)]_+ - \gamma^{-1}[P_{\gamma+}(u,\lambda)]_+, v_h-u_h+\gamma(\mu_h-\lambda_h)\right\rangle_C\\ &\quad = \underbrace{\left\langle \gamma^{-1}[P_{\gamma+}(u_h,\lambda_h)]_+ - \gamma^{-1}[P_{\gamma+}(u,\lambda)]_+, P_{\gamma+}(v_h-u,\mu_h-\lambda)\right\rangle_C}_{I}\\ &\qquad + \underbrace{\left\langle \gamma^{-1}[P_{\gamma+}(u_h,\lambda_h)]_+ - \gamma^{-1}[P_{\gamma+}(u,\lambda)]_+, P_{\gamma+}(u-u_h,\lambda-\lambda_h)\right\rangle_C}_{II}\\ &\qquad + \underbrace{\left\langle \gamma^{-1}[P_{\gamma+}(u_h,\lambda_h)]_+ - \gamma^{-1}[P_{\gamma+}(u,\lambda)]_+, 2\gamma(\mu_h-\lambda_h)\right\rangle_C}_{III}\\ &\quad = I+II+III\end{aligned}$$

The term I may be bounded using the Cauchy–Schwarz inequality followed by the arithmetic geometric inequality

$$I \le \varepsilon_3\|\gamma^{\frac12}\lambda+\gamma^{-\frac12}[P_{\gamma+}(u_h,\lambda_h)]_+\|^2_C+\frac{1}{4\varepsilon_3}\|\gamma^{-\frac12}P_{\gamma+}(v_h-u,\mu_h-\lambda)\|^2_C.$$

For the term II we use the monotonicity property $([a]_+-[b]_+)(b-a)\le-([a]_+-[b]_+)^2$ to deduce that

$$II \le -\|\gamma^{\frac12}(\lambda+\gamma^{-1}[P_{\gamma+}(u_h,\lambda_h)]_+)\|^2_C$$

Finally to estimate term III, let R_h be defined by Lemma 3.5 with the associated $r_h := I_{cf}(2\xi_h\gamma(\mu_h-\lambda_h))$ and set $\zeta_h = 1-\xi_h$. Using that

$$a(u-u_h,R_h)-\left\langle \gamma^{-1}[P_{\gamma+}(u_h,\lambda_h)]_+ - \gamma^{-1}[P_{\gamma+}(u,\lambda)]_+, r_h\right\rangle_C = 0$$

and adding and subtracting $2\xi_h(\mu_h-\lambda_h)$ in the right slot, we may write

$$\begin{aligned}
&\big\langle \gamma^{-1}[P_{\gamma+}(u_h,\lambda_h)]_+ - \gamma^{-1}[P_{\gamma+}(u,\lambda)]_+, 2\gamma(\mu_h-\lambda_h)\big\rangle_C \\
&= \underbrace{\big\langle \gamma^{-1}[P_{\gamma+}(u_h,\lambda_h)]_+ - \gamma^{-1}[P_{\gamma+}(u,\lambda)]_+, 2\zeta_h\gamma(\mu_h-\lambda_h)\big\rangle_{\Gamma_C}}_{IIIa} \\
&\quad + \underbrace{\big\langle \gamma^{-1}[P_{\gamma+}(u_h,\lambda_h)]_+ - \gamma^{-1}[P_{\gamma+}(u,\lambda)]_+, 2\xi_h\gamma(\mu_h-\lambda_h) - r_h\big\rangle_{\Gamma_C}}_{IIIb} \\
&\quad + \underbrace{a(u-u_h,R_h)}_{IIIc} \\
&= IIIa + IIIb + IIIc.
\end{aligned}$$

We estimate $IIIa$–$IIIc$ term by term. For $IIIa$ we use the Assumption 3.1

$$IIIa \le c_D\|\gamma^{-\frac12}[P_{\gamma+}(u_h,\lambda_h)]_+ + \gamma^{\frac12}\lambda\|_C^2 + c_D\|\gamma^{\frac12}(\mu_h-\lambda_h)\|_C^2$$

As a consequence of Proposition 3.4 we get the following bound of term $IIIb$

$$IIIb \le \varepsilon_4\|\gamma^{-\frac12}[P_{\gamma+}(u_h,\lambda_h)]_+ + \gamma^{\frac12}\lambda\|_C^2 + \frac{c_s^2\gamma}{\varepsilon_4}\|h^{\frac12}[\![\mu_h-\lambda_h]\!]\|_{\mathcal{F}}^2.$$

For the third term we observe that by the continuity of a and Lemma 3.5 we have

$$\begin{aligned}
IIIc &\le \|u-u_h\|_{H^1(\Omega)}\|R_h\|_{H^1(\Omega)} \le c_R\|u-u_h\|_{H^1(\Omega)}h^{-s}\|r_h\|_C \\
&\le c_R\|u-u_h\|_{H^1(\Omega)}h^{-s}\gamma^{\frac12}\|\gamma^{\frac12}(\lambda_h-\mu_h)\|_C \\
&\le \frac{\alpha}{4}\|u-u_h\|_{H^1(\Omega)}^2 + c_R^2h^{-2s}\gamma\alpha^{-1}\|\gamma^{\frac12}(\lambda_h-\mu_h)\|_C^2.
\end{aligned}$$

Collecting the above bounds and recalling that by definition

$$s(\lambda_h,\lambda_h) = \delta\gamma\|h^{\frac12}[\![\lambda_h]\!]\|_{\mathcal{F}_C}^2,$$

we have

$$\begin{aligned}
&\frac{\alpha}{2}\|u-u_h\|_{H^1(\Omega)}^2 + (1-\varepsilon_3-\varepsilon_4-c_D)\|\gamma^{-\frac12}(\gamma\lambda+[P_{\gamma+}(u_h,\lambda_h)]_+)\|_C^2 \\
&\qquad + (1-\varepsilon_1-c_D-c_R^2\gamma_0^{-1}/\alpha)\|\gamma^{\frac12}(\mu_h-\lambda_h)\|_C^2 \\
&\qquad + (1-\varepsilon_2-c_s^2/(\delta\varepsilon_4))|\mu_h-\lambda_h|_s^2 \\
&\quad \le \frac{1}{\alpha}\|u-v_h\|_{H^1(\Omega)}^2 + \frac{1}{4\varepsilon_3}\|\gamma^{-\frac12}P_{\gamma+}(v_h-u,\mu_h-\lambda)\|_C^2 \\
&\qquad + \frac{1}{4\varepsilon_1}\|\gamma^{\frac12}(\mu_h-\lambda)\|_C^2 + \frac{1}{4\varepsilon_2}|\mu_h|_s^2
\end{aligned}$$

Observe that, as usual when a continuous multiplier space is used, all terms and coefficients associated to the jump operator may be omitted.

Fixing $\varepsilon_1,\varepsilon_3,\varepsilon_4$ and γ_0 sufficiently large so that

$$\varepsilon_1 + c_R^2/(\gamma_0\alpha) = \varepsilon_3+\varepsilon_4 = (1-c_D)/2,$$

and ε_2 sufficiently small and δ sufficiently large so that $\varepsilon_2 + c_s^2/(\delta\varepsilon_4) < 1$, then there holds

$$\begin{aligned}
&\alpha\|u-u_h\|_{H^1(\Omega)}^2 + \|\gamma^{-\frac12}[P_{\gamma+}(u_h,\lambda_h)]_+ + \gamma^{\frac12}\lambda\|_C^2 + \|\gamma^{\frac12}(\mu_h-\lambda_h)\|_C^2 + |\mu_h-\lambda_h|_s^2 \\
&\quad \lesssim \frac{1}{\alpha}\|u-v_h\|_{H^1(\Omega)}^2 + \|\gamma^{-\frac12}P_{\gamma+}(v_h-u,\mu_h-\lambda)\|_C^2 + \|\gamma^{\frac12}(\mu_h-\lambda)\|_C^2 + |\mu_h|_s^2.
\end{aligned}$$

The triangle inequality $\|\gamma^{\frac12}(\lambda-\lambda_h)\|_C^2 \lesssim \|\gamma^{\frac12}(\mu_h-\lambda_h)\|_C^2 + \|\gamma^{\frac12}(\mu_h-\lambda)\|_C^2$ concludes the proof. $\square$

Corollary 5.2. *Assume that* $u \in H^r(\Omega)$, $1+s < r \leq k+1$ *and* $\lambda \in H^{r-1-s}(C)$, *with* $s = 1/2$ *for the Signorini problem and* $s = 1$ *for the Obstacle problem. Let* (u_h, λ_h) *be the solution of* (2.6) *with the contact operator defined by* (2.9). *Under the same conditions on the parameters as in Theorem 5.1 there holds*

$$\begin{aligned}&\alpha\|u-u_h\|_{H^1(\Omega)} + \gamma^{\frac{1}{2}}\|(\lambda-\lambda_h)\|_C + \gamma^{\frac{1}{2}}\|(\lambda+\gamma^{-1}[P_{\gamma+}(u_h,\lambda_h)]_+)\|_C \\ &\quad\lesssim h^{r-1}(|u|_{H^r(\Omega)} + |\lambda|_{H^{r-1-s}(C)}),\end{aligned}$$

Proof. Let $v_h = i_h u$ where i_h denotes the standard nodal interpolant and let $\mu_h = \pi_l \lambda$ where π_l denotes the L^2-projection. Using standard approximation estimates and the trace inequality (3.2) we may then bound the right hand side of the estimate of Theorem 5.1,

$$\|u-v_h\|^2_{H^1(\Omega)} \lesssim h^{2(r-1)}|u|^2_{H^r(\Omega)},$$

$$\gamma^{-1}\|(u-v_h)\|^2_C \lesssim \gamma^{-1}h^{2(r+s)}|u|^2_{H^r(\Omega)} \lesssim h^{2(r-1)}|u|^2_{H^r(\Omega)},$$

$$\gamma^{\frac{1}{2}}\|(\lambda-\mu_h)\|^2_C \lesssim \gamma^{\frac{1}{2}}h^{2(r-1-s)}|\lambda|_{H^{r-1-s}(C)} \lesssim h^{2(r-1)}|\lambda|_{H^{r-1-s}(C)}.$$

Finally we have,

$$\begin{aligned}|\mu_h|^2_s = s(\mu_h,\mu_h) &= s(\pi_l\lambda-\mu_h,\pi_l\lambda-\mu_h) \lesssim \gamma h^{-1}\|\pi_l\lambda-\mu_h\|^2_C \\ &\lesssim \gamma(\|\pi_l\lambda-\lambda\|^2 + \|\lambda-\mu_h\|^2_C) \lesssim h^{2s+2(r-1-s)}|\lambda|^2_{H^{r-1-s}(C)} \\ &\lesssim h^{2(r-1)}|\lambda|^2_{H^{r-1-s}(C)}\end{aligned}$$

and we conclude by taking square roots. □

Theorem 5.3. *(Formulation 2) Assume that* $u \in H^1(\Omega)$ *and* $\lambda \in L^2(C)$ *satisfies* (2.3) *and that* (u_h, λ_h) *is the solution to* (2.8) *with* (2.11) *and* $\gamma = \gamma_0 h^{2s}$, *where* $s = 1/2$ *for the Signorini problem and* $s = 1$ *for the Obstacle problem. If* γ_0 *sufficiently large and* $\delta > 0$ *sufficiently small, then there holds for all* $(v_h, \mu_h) \in V_h \times \Lambda_h$

$$\begin{aligned}&\alpha\|u-u_h\|^2_{H^1(\Omega)} + \gamma\|(\lambda-\lambda_h)\|^2_C \\ &\qquad + \gamma^{-1}\|(u-u_h) + [P_{\gamma-}(u,\lambda)]_+ - [P_{\gamma-}(u_h,\lambda_h)]_+\|^2_C \\ &\quad\lesssim \frac{1}{\alpha}\|u-v_h\|^2_{H^1(\Omega)} + \ \gamma\|\mu_h-\lambda\|^2_C + |\mu_h|^2_s + \gamma^{-1}\|v_h-u\|^2_C.\end{aligned} \tag{5.2}$$

Proof. Using the coercivity of $a(\cdot,\cdot)$ we may write

$$\begin{aligned}\alpha\|u-u_h\|^2_{H^1(\Omega)} &\leq a(u-u_h,u-u_h) = a(u-u_h,u-v_h) + a(u-u_h,v_h-u_h) \\ &\leq \frac{\alpha}{4}\|u-u_h\|^2_{H^1(\Omega)} + \frac{1}{\alpha}\|u-v_h\|^2_{H^1(\Omega)} \\ &\quad + a(u-u_h,v_h-u_h).\end{aligned} \tag{5.3}$$

By Galerkin orthogonality and by adding and subtracting suitable quantities it follows that

$$\begin{aligned}a(u-u_h,v_h-u_h) &= \langle\lambda-\lambda_h,v_h-u_h\rangle_C \\ &= \langle\lambda-\lambda_h,v_h-u_h\rangle_C \\ &\quad - \langle\mu_h-\lambda_h,u-u_h\rangle_C + s(\lambda_h,\mu_h-\lambda_h) \\ &\quad - \langle\mu_h-\lambda_h,[P_{\gamma-}(u,\lambda)]_+ - [P_{\gamma-}(u_h,\lambda_h)]_+\rangle_C.\end{aligned}$$

Then we proceed by adding and subtracting u in the right slot of the first term on the right hand side, λ in the left slot of the second term, μ_h in the left slot of the third term and finally $\lambda + \gamma^{-1}(u - u_h)$ in the left slot of the fourth term in the right hand side, leading to

$$\begin{aligned} a(u-u_h, v_h-u_h) = {} & \langle \lambda - \lambda_h, v_h - u\rangle_C - \langle \mu_h - \lambda, u - u_h\rangle_C \\ & - \langle \mu_h - \lambda, [P_{\gamma-}(u,\lambda)]_+ - [P_{\gamma-}(u_h,\lambda_h)]_+\rangle_C \\ & - \gamma^{-1}\langle P_{\gamma-}(u,\lambda) - P_{\gamma-}(u_h,\lambda_h), [P_{\gamma-}(u,\lambda)]_+ - [P_{\gamma-}(u_h,\lambda_h)]_+\rangle_C \\ & - \left\langle \gamma^{-1}(u-u_h), [P_{\gamma-}(u,\lambda)]_+ - [P_{\gamma-}(u_h,\lambda_h)]_+\right\rangle_C \\ & - |\mu_h - \lambda_h|_s^2 + s(\mu_h, \mu_h - \lambda_h). \end{aligned}$$

We may then apply the monotonicity of Lemma 3.2 to obtain the bound

$$\begin{aligned} a(u-u_h, v_h-u_h) \le {} & \langle \lambda - \lambda_h, v_h - u\rangle_C \\ & - \langle \mu_h - \lambda, (u-u_h) + [P_{\gamma-}(u,\lambda)]_+ - [P_{\gamma-}(u_h,\lambda_h)]_+\rangle_C \\ & - \gamma^{-1}\|[P_{\gamma-}(u,\lambda)]_+ - [P_{\gamma-}(u_h,\lambda_h)]_+\|_C^2 \\ & - \gamma^{-1}\langle (u-u_h), [P_{\gamma-}(u,\lambda)]_+ - [P_{\gamma-}(u_h,\lambda_h)]_+\rangle_C \\ & - |\mu_h - \lambda_h|_s^2 + s(\mu_h, \mu_h - \lambda_h). \end{aligned} \tag{5.4}$$

Summarizing (5.3) and (5.4) we have

$$\begin{aligned} & \frac{3}{4}\alpha\|u-u_h\|_{H^1(\Omega)}^2 + \frac{3}{4}|\mu_h - \lambda_h|_s^2 + \gamma^{-1}\|[P_{\gamma-}(u,\lambda)]_+ - [P_{\gamma-}(u_h,\lambda_h)]_+\|_C^2 \\ & \qquad + \gamma^{-1}\langle (u-u_h), [P_{\gamma-}(u,\lambda)]_+ - [P_{\gamma-}(u_h,\lambda_h)]_+\rangle_C \\ & \quad \le \frac{1}{\alpha}\|u - v_h\|_{H^1(\Omega)}^2 + \langle \lambda - \lambda_h, v_h - u\rangle_C \\ & \qquad - \langle \mu_h - \lambda, (u-u_h) + [P_{\gamma-}(u,\lambda)]_+ - [P_{\gamma-}(u_h,\lambda_h)]_+\rangle_C + |\mu_h|_s^2. \end{aligned} \tag{5.5}$$

To control the fourth term on the left hand side we need to obtain some control of $\gamma^{-1}\|u-u_h\|_C^2$. To this end first observe that the following Galerkin orthogonality holds

$$\langle \mu_h, u - u_h\rangle_C - s(\lambda_h, \mu_h) + \langle \mu_h, [P_{\gamma-}(u,\lambda)]_+ - [P_{\gamma-}(u_h,\lambda_h)]_+\rangle_C, \quad \forall \mu_h \in \Lambda_h. \tag{5.6}$$

Defining $\bar{e} = \pi_i(u - u_h)$, with $i = 0$ for $k = 1$ and $i = 1$ for $k \ge 2$ and taking $\mu_h = \gamma^{-1}\bar{e}$ in (5.6) we may write (the stabilization term is present only for $k = 1$)

$$\begin{aligned} 0 &= \gamma^{-1}\langle \bar{e}, [P_{\gamma-}(u,\lambda)]_+ - [P_{\gamma-}(u_h,\lambda_h)]_+\rangle_C + \gamma^{-1}\|\bar{e}\|_C^2 + s(\lambda_h, \gamma^{-1}\bar{e}) \\ &\ge \gamma^{-1}\langle \bar{e}, [P_{\gamma-}(u,\lambda)]_+ - [P_{\gamma-}(u_h,\lambda_h)]_+\rangle_C + \frac{3}{4}\gamma^{-1}\|\bar{e}\|_C^2 - c\delta|\lambda_h|_s^2. \end{aligned}$$

In the last inequality we used the Cauchy–Schwarz inequality and (3.3), to obtain the bound

$$s(\lambda_h, \gamma^{-1}\bar{e}) \le c\delta|\lambda_h|_s^2 + \frac{1}{4}\gamma^{-1}\|\bar{e}\|_C^2$$

It follows that

$$\frac{3}{4}\gamma^{-1}\|\bar{e}\|_C^2 - c\delta|\lambda_h|_s^2 \le -\gamma^{-1}\langle \bar{e}, [P_{\gamma-}(u,\lambda)]_+ - [P_{\gamma-}(u_h,\lambda_h)]_+\rangle_C.$$

We recall that by the L^2-orthogonality there holds, with $e = u - u_h$, $\|\bar{e}\|_C^2 = \|e\|_C^2 - \|e - \bar{e}\|_C^2$ and therefore

$$\gamma^{-1}\|e\|_C^2 \le \gamma^{-1}\|\bar{e}\|_C^2 + c\gamma^{-1}h^{2s}\|e\|_{H^1(\Omega)}^2.$$

For the Signorini problem this follows by using approximation $\|e-\bar{e}\|_C \le ch^{\frac12}\|e\|_{H^{\frac12}(C)} \le ch^{\frac12}\|e\|_{H^1(\Omega)}$. Consequently using also that $\|\bar{e}\|_C \le \|e\|_C$, there exists a constant c independent of γ and h such that

$$\frac{1}{2}\gamma^{-1}\|e\|_C^2 - \frac{1}{2}\gamma^{-1}\|[P_{\gamma-}(u,\lambda)]_+ - [P_{\gamma-}(u_h,\lambda_h)]_+\|_C^2 \\ - 2c\delta|\mu_h-\lambda_h|_s^2 - c(\delta+1)\gamma^{-1}h^{2s}\|u-u_h\|_{H^1(\Omega)}^2 \le 2c\delta|\mu_h|_s^2 \quad (5.7)$$

Collecting the results of equations (5.5), and (5.7) we have

$$\begin{aligned}\left(\frac{3}{4}\alpha - c(\delta+1)\gamma_0^{-1}\right)\|u-u_h\|_{H^1(\Omega)}^2 + \left(\frac{3}{4}-2c\delta\right)|\mu_h-\lambda_h|_s^2 \\ + \frac{1}{2}\gamma^{-1}\|e\|_C^2 + \frac{1}{2}\gamma^{-1}\|[P_{\gamma-}(u,\lambda)]_+ - [P_{\gamma-}(u_h,\lambda_h)]_+\|_C^2 \\ + \gamma^{-1}\left\langle (u-u_h), [P_{\gamma-}(u,\lambda)]_+ - [P_{\gamma-}(u_h,\lambda_h)]_+\right\rangle_C \\ \le \frac{1}{\alpha}\|u-v_h\|_{H^1(\Omega)}^2 + \langle \lambda-\lambda_h, v_h-u\rangle_C \\ - \langle \mu_h-\lambda, u-u_h + [P_{\gamma-}(u,\lambda)]_+ - [P_{\gamma-}(u_h,\lambda_h)]_+\rangle_C + (2c\delta+1)|\mu_h|_s^2.\end{aligned}$$

Assuming that γ_0 is large enough so that $c(\delta+1)\gamma^{-1}h^{2s} \le \frac{1}{4}\alpha$ and δ small so that $2c\delta \le 1/4$, using that $\frac12 a^2 + \frac12 b^2 + ab = \frac12(a+b)^2$ and the Cauchy–Schwarz inequality followed by the arithmetic-geometric inequality in the second to last term in the right hand side we obtain

$$\begin{aligned}\alpha\|u-u_h\|_{H^1(\Omega)}^2 + |\mu_h-\lambda_h|_s^2 \\ + \frac{1}{2}\gamma^{-1}\|(u-u_h) + [P_{\gamma-}(u,\lambda)]_+ - [P_{\gamma-}(u_h,\lambda_h)]_+\|_C^2 \\ \le \frac{2}{\alpha}\|u-v_h\|_{H^1(\Omega)}^2 + 2\langle \lambda-\lambda_h, v_h-u\rangle_C + 2\gamma\|\mu_h-\lambda\|_C^2 + 4(c\delta+1)|\mu_h|_s^2.\end{aligned} \quad (5.8)$$

It remains to control the Lagrange multiplier. Observe that taking $v_h = R_h$ as defined in Lemma 3.5 with $r_h = -\delta\gamma I_{cf}\xi_h(\mu_h-\lambda_h)$ we may use Galerkin orthogonality to obtain

$$\delta\gamma\|\xi_h^{\frac12}(\mu_h-\lambda_h)\|_C^2 - \delta\gamma(\mu_h-\lambda_h, (1-I_{cf})\xi_h(\mu_h-\lambda_h))_C \\ + \delta\gamma(\lambda-\mu_h, I_{cf}(\xi_h(\mu_h-\lambda_h)))_C + a(u-u_h,R_h) = 0.$$

Applying the bound (3.6), $c_\xi^2\|\mu_h-\lambda_h\|_C^2 \le \|\xi_h^{\frac12}(\mu_h-\lambda_h)\|_C^2$ in the first term and using the Cauchy–Schwarz inequality followed by Proposition 3.4 in the second and the third terms term of the left hand side leads to

$$\frac{c_\xi^2}{2}\delta\gamma\|\mu_h-\lambda_h\|_C^2 - c_s^2c_\xi^{-2}|\mu_h-\lambda_h|_s^2 + a(u-u_h,R_h) \lesssim \delta\gamma\|\lambda-\mu_h\|_C^2.$$

Applying now the triangle inequality $\|(\lambda-\lambda_h)\|_C^2 \le 2\|\mu_h-\lambda_h\|_C^2 + 2\|\mu_h-\lambda\|_C^2$ we obtain, with $c_{s,\xi} = c_s^2c_\xi^{-2}$

$$\frac{c_\xi^2}{4}\delta\gamma\|(\lambda-\lambda_h)\|_C^2 - c_{s,\xi}|\mu_h-\lambda_h|_s^2 + a(u-u_h,R_h) \lesssim \delta\gamma\|\lambda-\mu_h\|_C^2.$$

Recall that by Lemma 3.5 we have

$$a(u-u_h,R_h) \ge -\frac{c_\xi^2}{8}\delta\gamma\|(\lambda-\lambda_h)\|_C^2 - \frac{c_\xi^2}{8}\delta\gamma\|(\mu_h-\lambda)\|_C^2 - \frac{4c_R^2\gamma\delta}{c_\xi^2h^{2s}}\|\nabla(u-u_h)\|_\Omega^2$$

from which we deduce that there exists a constant $c_\lambda > 0$ depending on γ_0, c_R, $c_\xi$such that, (recalling the assumption, $\frac{\gamma}{h^{2s}} = \mathcal{O}(1)$)

$$c_\lambda \delta\gamma \|(\lambda - \lambda_h)\|_C^2 - |\mu_h - \lambda_h|_s^2 - \delta \|u - u_h\|_{H^1(\Omega)}^2 \lesssim \delta\gamma \|\lambda - \mu_h\|_C^2. \tag{5.9}$$

Multiplying both sides of (5.9) by $\frac{1}{2}$ taking $\delta < \alpha$, and adding it to (5.8) leads to the inequality

$$\begin{aligned} \alpha \|u - u_h\|_{H^1(\Omega)}^2 + |\mu_h - \lambda_h|_s^2 + \gamma \|(\lambda - \lambda_h)\|_C^2 \\ + \gamma^{-1} \|(u - u_h) + [P_{\gamma-}(u,\lambda)]_+ - [P_{\gamma-}(u_h,\lambda_h)]_+\|_C^2 \\ \lesssim \frac{1}{\alpha}\|u - v_h\|_{H^1(\Omega)}^2 + \gamma\|\mu_h - \lambda\|_C^2 + |\mu_h|_s^2 + \langle \lambda - \lambda_h, v_h - u\rangle_C . \end{aligned} \tag{5.10}$$

Finally splitting the last term on the right hand side

$$c\,\langle \lambda - \lambda_h, v_h - u\rangle_C \le \frac{1}{2}\gamma\|(\lambda - \lambda_h)\|_C^2 + \frac{1}{2}c^2\gamma^{-1}\|v_h - u\|_C^2$$

we conclude that

$$\begin{aligned} \alpha \|u - u_h\|_{H^1(\Omega)}^2 + |\mu_h - \lambda_h|_s^2 + \frac{1}{2}\gamma \|(\lambda - \lambda_h)\|_C^2 \\ + \gamma^{-1} \|(u - u_h) + [P_{\gamma-}(u,\lambda)]_+ - [P_{\gamma-}(u_h,\lambda_h)]_+\|_C^2 \\ \lesssim \frac{1}{\alpha}\|u - v_h\|_{H^1(\Omega)}^2 + \gamma\|\mu_h - \lambda\|_C^2 + |\mu_h|_s^2 + \gamma^{-1}\|v_h - u\|_C^2. \end{aligned} \tag{5.11}$$

□

Corollary 5.4. *Assume that $u \in H^r(\Omega)$, $s + \frac{1}{2} < r \le k+1$ and $\lambda \in H^{r-1-s}(C)$, with$s = 1/2$ for the Signorini problem and $s = 1$ for the Obstacle problem. Let (u_h, λ_h) be the solution of (2.6) with the contact operator defined by (2.11). Under the same conditions on the parameters as in Theorem 5.3 there holds*

$$\begin{aligned} \alpha\|u - u_h\|_{H^1(\Omega)} + \gamma\|(\lambda - \lambda_h)\|_C \\ + \gamma^{-1/2}\|(u - u_h) + [P_{\gamma-}(u,\lambda)]_+ - [P_{\gamma-}(u_h,\lambda_h)]_+\|_C \\ \lesssim h^{r-1}(|u|_{H^r(\Omega)} + |\lambda|_{H^{r-1-s}(C)}). \end{aligned}$$

Proof. Similar to that of Corollary 5.2. □

Remark 5.5. Note that the assumption $\lambda \in L^2(C)$ hides an assumption on the primal variable u. Formally identifying λ with the exterior normal derivative $\partial_n u$ for the Signorini problem we have $\partial_n u \in L^2(\Gamma_C)$ and similarly for the obstacle problem $\Delta u \in L^2(\Omega)$.

6. Numerical examples

In the numerical examples below, we define $h = 1/\sqrt{\mathrm{NNO}}$, where NNO denotes the number of nodes in a uniformly refined mesh. We use the formulation (2.8) with the nonlinear term defined by (2.10). Numerical experiments with the formulation (2.11), not reported here, resulted in very similar results. For the spaces we choose piecewise linear finite elements for the primal variable and piecewise constants for the Lagrange multipliers, constant per element for the obstacle problem, and constant per element edge on the Signorini boundary for the Signorini problem.

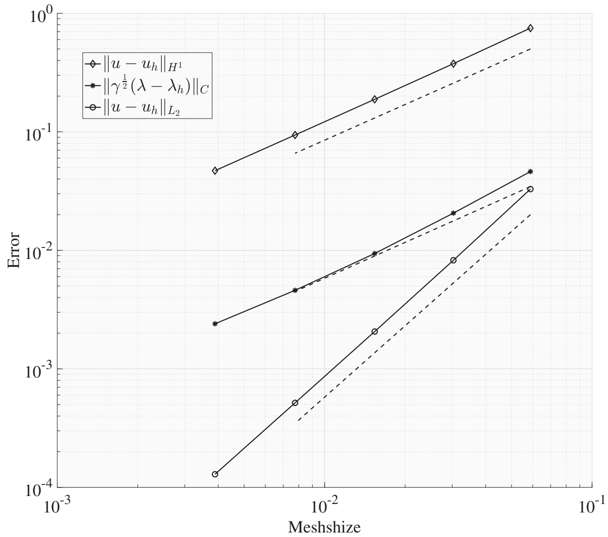

FIGURE 1. Convergence for the smooth obstacle. Dashed lines indicate expected convergence for smooth solutions.

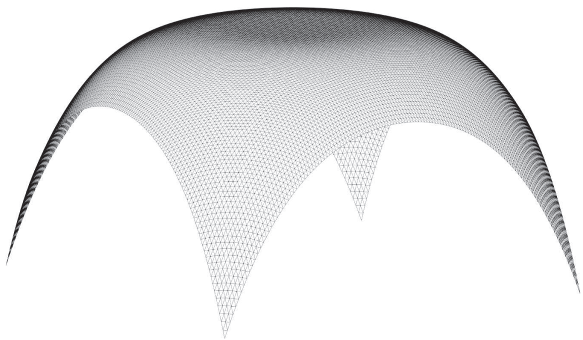

FIGURE 2. Elevation of the discrete solution, smooth obstacle.

6.1. Smooth obstacle problem

Our smooth obstacle example, adapted from [31], is posed on the square $\Omega = (-1,1) \times (-1,1)$ with

$$f(r) = \begin{cases} 8r_0^2(1-(r^2-r_0^2)) & \text{if } r \le r_0, \\ 8(r^2+(r^2-r_0^2)) & \text{if } r > r_0, \end{cases}$$

where $r = \sqrt{x^2+y^2}$ and $r_0 = 1/4$, and with Dirichlet boundary conditions taken from the corresponding exact solution

$$u = -[r^2 - r_0^2]_+^2.$$

We choose $\gamma = h^2/\gamma_0$ with $\gamma_0 = 10^{-1}$ and show the convergence in the L^2– and H^1–norms in Figure 1. An elevation of the computed solution on one of the meshes in a sequence is given in Figure 2. We note the optimal convergence of $O(h^2)$ in L^2 (dashed line has inclination 2:1) and $O(h)$ in H^1 as well as for a weighted norm of the multiplier error (dashed line has inclination 1:1).

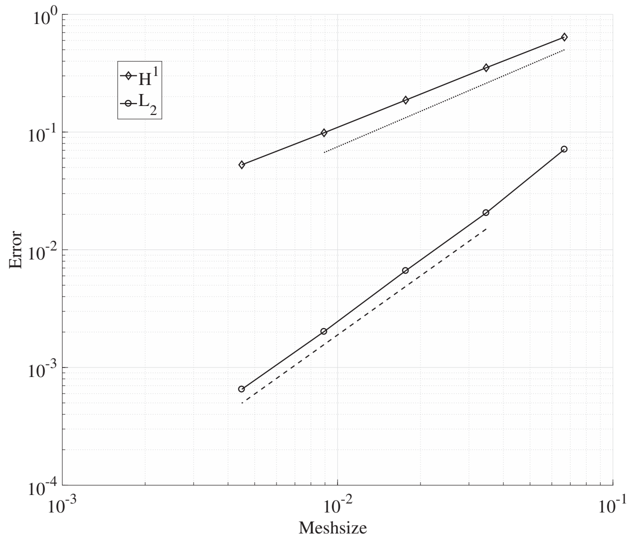

FIGURE 3. Convergence for the nonsmooth obstacle. Dotted line has inclination 1:1, dashed line has inclination 1:5/3.

6.2. Nonsmooth obstacle problem

This example was proposed by Braess *et al.* [8]. The domain is $\Omega = (-2,2) \times (-2,2) \setminus [0,2) \times (-2,0]$ with

$$f(r,\varphi) = r^{2/3} \sin(2\varphi/3)(g_1'(r)/r + g_1''(r)) + \frac{4}{3} r^{-1/3} g_1'(r) \sin(2\varphi/3) + g_2(r)$$

where, with $\hat{r} = 2(r - 1/4)$,

$$g_1(r) = \begin{cases} 1, & \hat{r} < 0 \\ -6\hat{r}^5 + 15\hat{r}^4 - 10\hat{r}^3 + 1, & 0 \leq \hat{r} < 1 \\ 0, & \hat{r} \geq 1, \end{cases}$$

$$g_2(r) = \begin{cases} 0, & r \leq 5/4, \\ 1 & \text{elsewhere}. \end{cases}$$

with Dirichlet boundary conditions taken from the corresponding exact solution

$$u(r,\varphi) = -r^{2/3} g_1(r) \sin(2\varphi/3)$$

which belongs to $H^{5/3-\varepsilon}(\Omega)$ for arbitrary $\varepsilon > 0$.

For this example we plot, in Figure 3, the error on consecutive refined meshes. We note the suboptimal convergence in L^2 which agrees with the regularity of the exact solution. In Figure 4 we show an elevation of the approximate solution on one of the meshes used to compute convergence.

6.3. Signorini problem

The Signorini problem is posed on the unit square $(0,1) \times (0,1)$ with homogeneous Dirichlet boundary conditions at $y = 1$, homogeneous Neumann boundary conditions at $x = 0$ and $x = 1$, and a Signorini boundary at $y = 0$. The load is $f = -2\pi \sin 2\pi x$ (following [6]), and we set $\gamma_0 = 10$. No explicit solution is available and we instead use an overkill solution, using 66049 nodes (corresponding to $h \approx 4 \times 10^{-3}$) to estimate the error. In

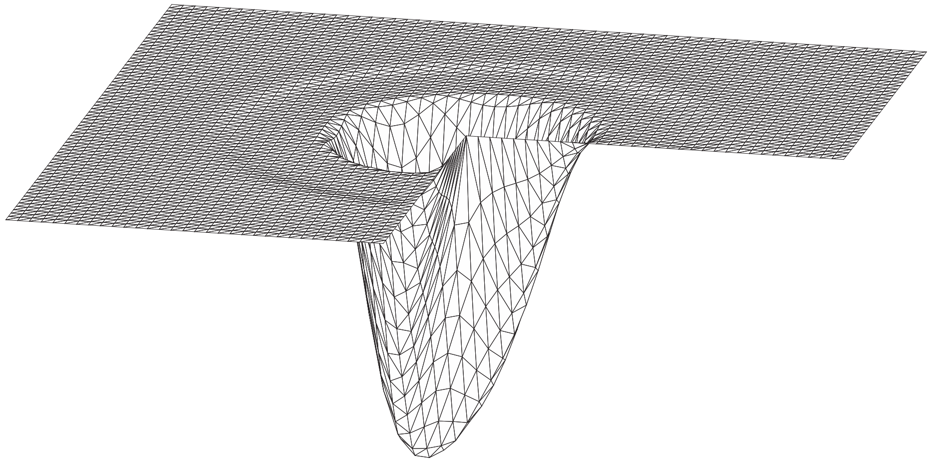

FIGURE 4. Elevation of the discrete solution, nonsmooth obstacle.

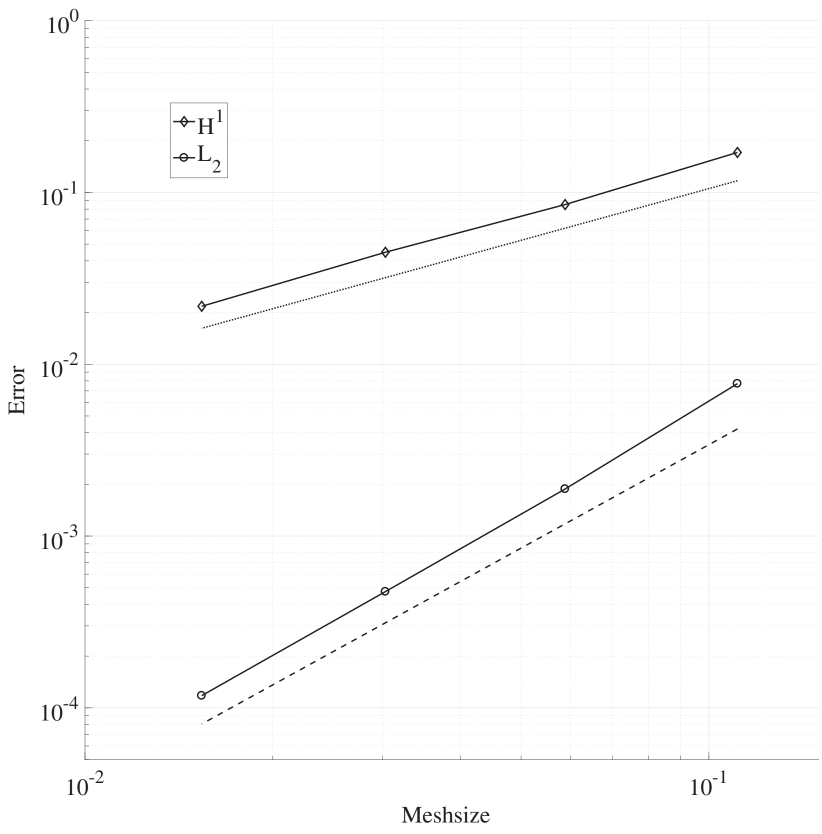

FIGURE 5. Convergence for the Signorini case.

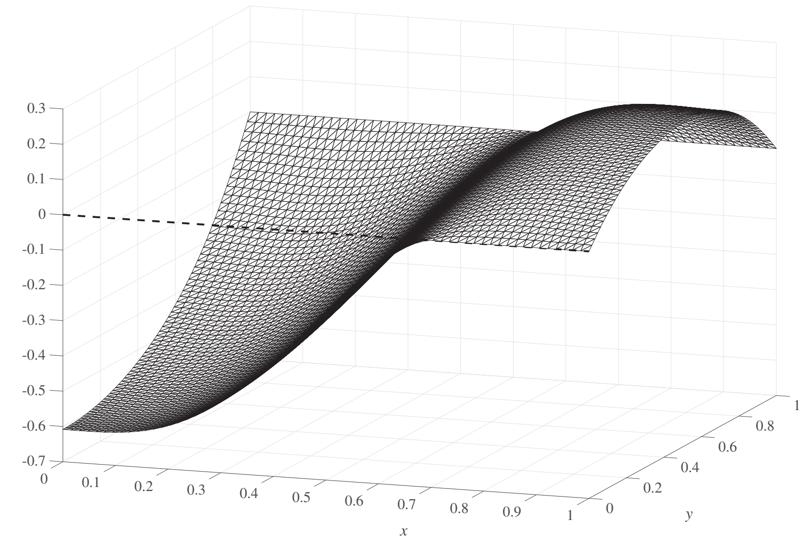

FIGURE 6. Elevation of the discrete solution, Signorini case. The Signorini boundary is indicated by the dashed line.

Figure 5 we show the convergence in the L^2– and H^1–norms and again we observe optimal convergence of $O(h^2)$ in L^2 (dashed line has inclination 2:1) and $O(h)$ in H^1 (dotted line has inclination 1:1). Finally, in Figure 6 we show an elevation of the computed solution.

Acknowledgements. This research was supported in part by EPSRC, UK (EP/J002313/1), the Swedish Foundation for Strategic Research (AM13-0029), the Swedish Research Council (2011-4992, 2013-4708), and the Swedish Strategic Research Program Essence. The first author wishes to thank Dr. Franz Chouly, Prof. Patrick Hild and Prof. Yves Renard for interesting discussions on Nitsche's method for contact problems and the augmented Lagrangian method. The authors also wish to thank the anonymous reviewers who helped improve the manuscript through their many constructive comments.

REFERENCES

[1] P. Alart and A. Curnier, A mixed formulation for frictional contact problems prone to Newton like solution methods. *Comput. Methods Appl. Mech. Eng.* **92** (1991) 353–375.

[2] D.N. Arnold, An interior penalty finite element method with discontinuous elements. *SIAM J. Numer. Anal.* **19** (1982) 742–760.

[3] I. Athanasopoulos and L.A. Caffarelli, Optimal regularity of lower dimensional obstacle problems. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)* **310** (2004) 49–66.

[4] 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) 1–15.

[5] Z. Belhachmi and F. Ben Belgacem, Quadratic finite element approximation of the Signorini problem. *Math. Comp.* **72** (2003) 83–104.

[6] F. Ben Belgacem, Numerical simulation of some variational inequalities arisen from unilateral contact problems by the finite element methods. *SIAM J. Numer. Anal.* **37** (2000) 1198–1216.

[7] F. Ben Belgacem and Y. Renard, Hybrid finite element methods for the Signorini problem. *Math. Comp.* **72** (2003) 1117–1145.

[8] D. Braess, C. Carstensen and R.H.W. Hoppe, Convergence analysis of a conforming adaptive finite element method for an obstacle problem. *Numer. Math.* **107** (2007) 455–471.

[9] F. Brezzi, W.W. Hager and P.-A. Raviart, Error estimates for the finite element solution of variational inequalities. II. Mixed methods. *Numer. Math.* **31** (1978/1979) 1–16.

[10] E. Burman and A. Ern, Continuous interior penalty *hp*-finite element methods for advection and advection-diffusion equations. *Math. Comp.* **76** (2007) 1119–1140.

[11] E. Burman and P. Hansbo, Fictitious domain finite element methods using cut elements: I. A stabilized Lagrange multiplier method. *Comput. Methods Appl. Mech. Eng.* **199** (2010) 2680–2686.

[12] E. Burman and P. Hansbo, Interior-penalty-stabilized Lagrange multiplier methods for the finite-element solution of elliptic interface problems. *IMA J. Numer. Anal.* **30** (2010) 870–885.

[13] E. Burman and P. Hansbo, Deriving robust unfitted finite element methods from augmented Lagrangian formulations, edited by S. Bordas, E. Burman, M. Larson, M. Olshanskii. In: Geometrically Unfitted Finite Element Methods and Applications. In Vol. 121 of *Lecture Notes in Computational Science and Engineering*. Springer, Cham. (2017).

[14] E. Burman and F. Schieweck, Local CIP stabilization for composite finite elements. *SIAM J. Numer. Anal.* **54** (2016) 1967–1992.

[15] L.A. Caffarelli, The obstacle problem revisited. *J. Fourier Anal. Appl.* **4** (1998) 383–402.

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

[17] 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, edited by S. Bordas, E. Burman, M. Larson and M. Olshanskii. In: Geometrically Unfitted Finite Element Methods and Applications. In Vol. 121 of *Lecture Notes in Computational Science and Engineering*. Springer, Cham. (2017).

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

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

[20] F. Chouly, P. Hild and Y. Renard, Symmetric and non-symmetric variants of Nitsche's method for contact problems in elasticity: theory and numerical experiments. *Math. Comp.* **84** (2015) 1089–1112.

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

[22] R. Glowinski and P. Le Tallec, Augmented Lagrangian and operator-splitting methods in nonlinear mechanics. In: Vol. 9 of *SIAM Studies in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA (1989).

[23] P. Hansbo, A. Rashid and K. Salomonsson, Least-squares stabilized augmented Lagrangian multiplier method for elastic contact. *Finite Elem. Anal. Des.* **116** (2016) 32–37.

[24] J. Haslinger, I. Hlaváček and J. Nečas, Numerical methods for unilateral problems in solid mechanics. In *Handbook of Numerical Analysis*. IV. North-Holland, Amsterdam (1996) 313–485

[25] P. Hild and Y. Renard, A stabilized Lagrange multiplier method for the finite element approximation of contact problems in elastostatics. *Numer. Math.* **115** (2010) 101–129.

[26] P. Hild and Y. Renard, An improved a priori error analysis for finite element approximations of Signorini's problem. *SIAM J. Numer. Anal.* **50** (2012) 2400–2419.

[27] K. Ito and K. Kunisch, An augmented Lagrangian technique for variational inequalities. *Appl. Math. Optim.* **21** (1990) 223–241.

[28] N. Kikuchi and J.T. Oden, Contact problems in elasticity: a study of variational inequalities and finite element methods. In: Vol. 8 of *Studies in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA (1988).

[29] R. Mlika, Y. Renard and F. Chouly, An unbiased Nitsche's formulation of large deformation frictional contact and self-contact. *Comput. Methods Appl. Mech. Eng.* **325** (2017) 265–288.

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

[31] R.H. Nochetto, K.G. Siebert and A. Veeser, Pointwise a posteriori error control for elliptic obstacle problems. *Numer. Math.* **95** (2003) 163–195.

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

[33] M.A. Taylor, B.A. Wingate and L.P. Bos, Several new quadrature formulas for polynomial integration in the triangle. Preprint ArXiv: math/0501496 (2005).

[34] R. Temam, Navier-Stokes Equations. Theory and Numerical Analysis. Reprint of the 1984 edn. AMS Chelsea Publishing, Providence, RI (2001).

[35] B.I. Wohlmuth, A. Popp, M.W. Gee and W.A. Wall, An abstract framework for a priori estimates for contact problems in 3D with quadratic finite elements. *Comput. Mech.* **49** (2012) 735–747.

[36] P. Wriggers and G. Zavarise, A formulation for frictionless contact problems using a weak form introduced by Nitsche. *Comput. Mech.* **41** (2007) 407–420.

[37] L. Zhang, T. Cui and H. Liu, A set of symmetric quadrature rules on triangles and tetrahedra. *J. Comput. Math.* **27** (2009) 89–96.