**No. 571** **June 2017**

# Adaptive optimal control of the signorini's problem

**A. Rademacher, K. Rosin**

**ISSN: 2190-1767**

# ADAPTIVE OPTIMAL CONTROL OF THE SIGNORINI'S PROBLEM

A. RADEMACHER AND K. ROSIN

ABSTRACT. In this article, we present a-posteriori error estimations in context of optimal control of contact problems; in particular of Signorini's problem. Due to the contact side-condition, the solution operator of the underlying variational inequality is not differentiable, yet we want to apply Newton's method. Therefore, the non-smooth problem is regularized by penalization and afterwards discretized by finite elements. We derive optimality systems for the regularized formulation in the continuous as well as in the discrete case. This is done explicitly for Signorini's contact problem, which covers linear elasticity and linearized surface contact conditions. The latter creates the need for treating trace-operations carefully, especially in contrast to obstacle contact conditions, which exert in the domain. Based on the dual weighted residual method and these optimality systems, we deduce error representations for the regularization, discretization and numerical errors. Those representations are further developed into error estimators. The resulting error estimator for regularization error is defined only in the contact area. Therefore its computational cost is especially low for Signorini's contact problem. Finally, we utilize the estimators in an adaptive refinement strategy balancing regularization and discretization errors. Numerical results substantiate the theoretical findings. We present different examples concerning Signorini's problem in two and three dimensions.

## 1. INTRODUCTION

Investigations of industrial production processes are commonly based on partial differential equations (PDEs). We mention exemplarily the incremental rolling of a surface in order to grade the surface texture. In such a process, the goal is to achieve a particular structure, which is modeled via the deformation. We realize this by introducing an optimal control problem. In this context we call the deformation state and the control is given by a volume or Neumann force. Our quantity of interest is a tracking-type goal functional depending on the difference between the actual and wanted states as well as a Tikhonov term for the control. Additionally to the control force; the roll exerts on the workpiece, leading to surface contact conditions. In general, we also have to consider time-dependent elasto-plastic material laws providing lasting deformations. However in this work, we focus on a subproblem. We study only the static case consisting in semi-linear variational inequalitys (VIs). For instance, we consider the combination of linear, elastic material laws with the surface contact conditions; known as Signorini's problem. Additionally, we formulate the goal functional for the theoretical analysis more generally, since the complexity does not increase by doing so. Thus we arrive at an optimal control problem governed by static contact problems. The analysis of such problems is done for instance in [6, 37].

In order to gain an efficient algorithm for solving the optimal control problem, we want to apply Newton's method. However due to the non-differentiability of the solution operator associated with the VIs, the reduced cost functional is not differentiable. This is the very same challenge, as introduced for the optimal control of the obstacle problem, see [24]. To circumvent this difficulty, we regularize the problem via penalization and examine the sequence of penalized problems. This common approach is performed for instance in [2, 18, 21] and [30].

Obviously, the solution of a regularized problem differs in general from the solution of the original optimal control problem. Hence, we want to study the introduced regularization error as well as a discretization error, which comes into play, when we discretize the system with bilinear finite elements (FEs). The numerical treatment of the non-linear system produces another error. The error in the cost functional, is split into three parts corresponding to those errors.

Our goal is an efficient, adaptive algorithm balancing the regularization and discretization error, while decreasing the numerical error to a negligible level. For this purpose, we develop a-posteriori error estimators based on the dual weighted residual (DWR) method. There are many contributions regarding similar estimates. Recent overviews for a-posteriori error estimators are presented in [35]. Error estimators concerning the obstacle problem itself, without optimal control, are treated for instance in [9, 32] with the DWR method. Far less contributions are dedicated to Signorini type contact problems, e. g. [26]. For residual based error estimations, we refer to e. g. [11, 13, 15, 19, 34] and in particular to [16], where residual type error estimates for a penalized obstacle problem were derived.

In the literature, error estimates are also developed for optimization problems subject to PDEs. We refer especially to [36] and the survey [27]. Following the pioneering work of [3], we derive DWR error estimates; cf. also [1, 4] Optimization problems subject to VIs are considered in, e. g. [14]. A-priori estimates for optimal control problems are covered for instance in [23] and residual-based a-posteriori estimates in [17].

A short insight into early studies is granted in [8]. Therein, we consider optimal control problems subject to elasto-plastic contact problems and use two different kinds of regularization techniques. On the one hand a local smoothing for elasto-plasticity and on the other hand the global penalty for the contact situation. However, the regularization error estimators in [8] are based on the Richardson extrapolation. Another quasi-linear problem to a different regularization parameter is solved for each estimation. This approach is obviously very time-consuming.

A more efficient error estimator for the regularization error is developed in [22]; a paper dedicated to optimal control problems subject to standard obstacle problems. Therein, the extrapolation is based on Taylor expansion with respect to the regularization parameter. With only one additional Newton step, it can be evaluated. Afterwards, the DWR method is applied to gain an error estimator for the discretization and numerical errors. However, a difficulty arises, namely the dependency of the DWR-residual on the regularization parameter. Even though that residual is of higher order for a fixed parameter, there is no guarantee that it is negligible, when the penalty parameter approaches infinity.

Yet we follow the basic idea of utilizing Taylor expansion in this paper. First of all, we consider the more difficult contact problems, for instance Signorini's problem. While the obstacle problem uses the obstacle itself in its constraints; in Signorini's problem a so-called initial gap function is used, leading to a different sign in the penalty term. Furthermore, the contact situation happens on the surface, which brings the trace operator into play. Therefore, we need to re-investigate the problems carefully.

However the more important difference is that we switch the order of applying the DWR method and the Taylor expansion. This leads to the major advantage that the residual depends no longer on neither the penalty parameter nor the discretization. For tracking-type goal functionals it even vanishes completely. Furthermore, the resulting regularization error estimator is only the dual pairing of a slack variable and the adjoint state, which is evaluated only on the contact surface for constraints of Signorini type.

In the next section, we introduce Signorini's problem and sketch the simplified version. Afterwards, the optimal control problem governed by contact problems is formally defined. In the third section, its regularization and discretization are discussed. Here, optimality systems for the original, the regularized, and the discrete regularized problem are derived. Section 4 is addressed to the analysis of the error. The error is split into three parts as mentioned above. We derive a-posteriori error estimators for each part and utilize them in an adaptive algorithm. In Section 5, we apply the algorithm to different examples in two or three dimensions. The numerical results confirm that the performed error splitting and the respective error estimation are well-posed. Additionally, we see that the error balancing takes place even though a concrete formulation of an assumption, namely a sign condition for the Lagrange multiplier, is not fulfilled. We close the paper with a brief conclusion and outlook.

## 2. Problem formulation

In this section we introduce the optimal control problem governed by VIs arising from contact problems. We start by presenting Signorini's problem and a brief comparison to other contact problems. Next we incorporate them into an optimization problem.

2.1. **Signorini's problem.** Signorini's problems cover linear, elastic material laws restricted by surface contact conditions. A major difference between the Signorini's problem and the obstacle one is implied by the contact area itself. In contrast to the obstacle problem, where the contact area is the whole domain $\Omega$, this area is here part of the boundary $\Gamma_C \subset \partial\Omega$. Hence, the trace operator tr becomes very important in this context.

Let $\Omega$ be a Lipschitz domain, with a partition of the boundary $\partial\Omega$ into $\Gamma_C, \Gamma_D$ and $\Gamma_N$. We presume that the contact boundary $\Gamma_C$ and the Dirichlet boundary $\Gamma_D$ have a strict positive distance, i. e. $\bar{\Gamma}_C \cap \bar{\Gamma}_D = \emptyset$. Hence, the Neumann boundary $\Gamma_N$ is nonempty. The state $u$ lives in the Cartesian product of the scalar function space

$$V := \mathscr{H}_D(\Omega) := \left(H_D(\Omega)\right)^d := \left\{v \in H^1(\Omega) \mid \mathrm{tr}(v) = 0 \text{ on } \Gamma_D\right\}^d.$$

We use the parametrization of the contact surface $\phi$ and of the rigid obstacle $\tilde{\phi}$. Either of them must be smooth enough such that the respective normal is well posed. Together they give the normalized, initial gap function

$$\psi(\alpha) = \frac{\tilde{\phi}(\alpha) - \phi(\alpha)}{\sqrt{1 + (\nabla_\alpha \phi(\alpha))^2}},$$

which is a linear approximation of the gap between the obstacle and the domain at a given parameter point $\alpha \in \mathbb{R}^{d-1}$ on the boundary $\Gamma_C$ in the deformed configuration, cf. [25, Ch. 2.3]. This gap function shall be in the factorial boundary space $H^{1/2}(\Gamma_C)$, see [25, Ch. 5.3].

The operator $\tau$ is the trace operator in normal direction restricted to the contact boundary, i. e. $\tau := \mathrm{tr} \cdot n\,|_{\Gamma_C}$. We know that $\tau$ is a linear, continuous operator $V \to L^2(\Gamma_C)$. Since every displacement $u$ is restricted by the obstacle, we get the feasible set of states $K$ by means of the approximated gap $\psi$

$$K = \{v \in V : \tau(v) \leq \psi \text{ a. e. on } \Gamma_C\}.$$

The state $u$ must fulfill an adequate state of equilibrium, which is in the context of Signorini's problem given by a linear elastic material law. The necessary notations are introduced in the following. The dual pairing between $V$ and its dual space $V'$ is denoted by $\langle \cdot, \cdot \rangle$ and the standard scalar product on $V$ by $(\cdot, \cdot)$. Moreover, the bilinear form $a : V \times V$ is given by the following second-order elliptic operator $A : V \to V'$

$$\langle Au, v\rangle = a(u, v) = \int_\Omega \sigma(u) : \varepsilon(v)\,\mathrm{d}x \tag{1}$$

where $\varepsilon$ denotes the linearized strain tensor,

$$\varepsilon(u) := \frac{1}{2}\left(\nabla u + \nabla^\top u\right)$$

and $\sigma$ the second order stress tensor. The latter is given by applying the fourth order tensor $\mathbb{C}$, which is known as the elastic tensor, onto the linearized strain tensor, i. e. $\sigma = \mathbb{C}\varepsilon$. For $d = 3$, it holds

$$\mathbb{C}\varepsilon := \lambda_m \operatorname{trace}(\varepsilon)\mathrm{I} + 2\mu_m \varepsilon$$

with Lamé constants $\lambda_m > -\frac{2}{3}\mu_m$ and $\mu_m > 0$. The two dimensional case can be treated in different manners. For instance one could use plane stress, plane strain or an academical modeling. In the last approach, neither stress nor strain are related to realistic physical behavior, but the strain and stress tensors are simply the two dimensional analogs of the three dimensional case. More details to the first two approaches can be found for instance in [12, Ch. VI, §5]. We only define Signorini's problem in the weak form and we use the notation $V^0 = L^2(\Omega)^d$, which will come in handy later on.

**Definition 2.1** (Signorini's problem)**.** *Let $\Omega$ be a given Lipschitz domain with the boundary parts $\Gamma_C$, $\Gamma_D$ and $\Gamma_N$. For a known volume force or control $q \in V^0$ find a displacement or state $u \in K$ such that*

$$a(u,\ v-u) \geq (q,\ v-u) \qquad \forall v \in K. \tag{2}$$

We introduce the function $G(u) := \psi - \tau(u)$ to unify the notation and for comparison to other contact problems. In respect to our convex set $K$ and the operator $\tau$, we define $r : \mathbb{R}^+ \times V \to V'$, such that for any arbitrary $v \in V$

$$r(\gamma;u)(v) = -\int_{\Gamma_C} \left[\max\left(\gamma(-G(u)),0\right)\right]^3 \tau(v)\ \mathrm{d}o. \tag{3}$$

The penalty term $r$ is derived from a bi-quadratic penalization of the respective energy functional associated with (2) and its Gâteaux-derivatives read

$$\partial_u r(\gamma;u)(\varphi,v) = \int_{\Gamma_C} 3\,\gamma\left[\max\left(\gamma(\tau(u)-\psi),0\right)\right]^2 \tau(\varphi)\tau(v)\,\mathrm{d}o \tag{4}$$

and

$$\partial^2_{uu} r(\gamma;u)(\chi,\varphi,v) = \int_{\Gamma_C} 6\,\gamma^2\left[\max\left(\gamma(\tau(u)-\psi),0\right)\right] \tau(\chi)\tau(\varphi)\tau(v)\,\mathrm{d}o.$$

We demand that the contact boundary $\Gamma_C$ is $\mathscr{C}^{1,1}$, i. e. the normal, which is defined by derivatives evaluated on the boundary part $\Gamma_C$, is Lipschitz continuous, too. For later use, we demand for a sequence of deformations $\{u_\gamma\} \subset V$ the following result.

**Lemma 2.1.** *There shall be a subdomain $\Omega' \subset \Omega$, such that*

$$\Omega' \in \mathscr{C}^{1,1}, \qquad \mathrm{dist}(\Gamma_D,\Omega') > 0 \qquad \textit{and} \qquad \partial\Omega' \cap \partial\Omega \supseteq \Gamma_C.$$

*Furthermore, for each deformation $u_\gamma$ there is a corresponding function* $\max\left(\tau(u_\gamma)-\psi,0\right)$. *We demand of that very surface function to be in $H^{1/2}_{00}(\Gamma_C)$ - the special subspace given by Lions and Magenes, cf.* [25, Chapter 5.3]

$$H^{1/2}_{00}(\Gamma_C) := \left\{ v \in H^{1/2}(\Gamma_C) \mid v \textit{ vanishes rapidly enough at } \partial\Gamma_C \right\}.$$

*Then a function $U_{0,\gamma} \in V$ and a constant $C$ exist, such that*

$$\tau(U_{0,\gamma}) = \max\left(\tau(u_\gamma)-\psi,0\right)$$

*and*

$$\|U_{0,\gamma}\|_V \leq C\|\max\left(\tau(u_\gamma)-\psi,0\right)\|_{\Gamma_C}.$$

*If* $\max\left(\tau(u_\gamma)-\psi,0\right) \to 0$ *in* $H^{1,2}(\Gamma_C)$ *for $\gamma \to \infty$, then $U_{0,\gamma} \to 0$ in $\Omega$.*

*Proof.* Additionally to the normal trace $\tau = \mathrm{tr} \cdot n$, there exists a surjective linear continuous mapping $\mathrm{tr}_T : V \to H^{1/2}_T(\partial\Omega')$ with the tangential space

$$H^{1/2}_T(\partial\Omega') = \left\{ v \in H^{1/2}(\partial\Omega')^d \mid v \cdot n = 0 \right\}.$$

For any $g \in H^{1/2}(\partial\Omega')$ and $h \in H^{1/2}_T(\partial\Omega')$ there exists a $v \in \mathscr{H}(\Omega')$ and a constant $C$ such that

$$\tau(v) = g \quad \text{and} \quad \mathrm{tr}_T(v) = h \qquad \text{with} \quad \|v\|_{1,\Omega'} \leq C\big(\|g\|_{1/2,\partial\Omega'} + \|h\|_{1/2,\partial\Omega'}\big). \tag{5}$$

This is because $\Omega'$ is smooth enough, see [25, Theorem 5.5]. We choose $h = 0$ and $g = \max\left(\tau(u_\gamma)-\psi,0\right)$. Since we presumed that $g \in H^{1/2}_{00}(\Gamma_C)$ and because every function in $H^{1/2}_{00}(\Gamma_C)$ can be continued trivially onto $H^{1/2}(\partial\Omega')$, we get the existence of a function $v$ with

$$\|v\|_{1,\Omega'} \leq C\|\max\left(\tau(u_\gamma)-\psi,0\right)\|_{\Gamma_C}.$$

Each function in $\mathscr{H}_D(\Omega')$ can be continued trivially onto $V$. We conclude the proof by calling this continuation $U_{0,\gamma}$ and using the relation $\|v\|_{1,\Omega'} = \|U_{0,\gamma}\|_{1,\Omega}$. □

For further studies of Signorini's contact problem we refer to [25].

2.2. **Comparison between contact problems.** Here, we show the differences in the nature of the investigated contact problems with the aid of Tables 1 and 2. Next to the common obstacle problem and the Signorini's problem, there is the simplified Signorini's problem, in which Poisson's equation is subject to a boundary contact condition.

In the following tables we summarize the necessary function spaces, the functions that describe the obstacle or the approximated gap $\psi$, and the areas, where the contact is expected $\Sigma$. We include the bilinear forms $a(\cdot,\,\cdot)$ for the respective partial differential inequality as well. By adding the index $D$ to the dual spaces,

TABLE 1. Function spaces

| | $V$ | $V^0$ | $V'$ | $W$ | $W'$ |
|---|---|---|---|---|---|
| Obstacle | $H^1_D(\Omega)$ | $L^2(\Omega)$ | $H^{-1}_D(\Omega)$ | $V$ | $V'$ |
| Simplified Signorini | $H^1_D(\Omega)$ | $L^2(\Omega)$ | $H^{-1}_D(\Omega)$ | $H^{1/2}(\Gamma_C)$ | $H^{-1/2}(\Gamma_C)$ |
| Signorini | $H^1_D(\Omega)^d$ | $L^2(\Omega)^d$ | $(H^{-1}_D(\Omega))^d$ | $H^{1/2}(\Gamma_C)$ | $H^{-1/2}(\Gamma_C)$ |

we want to emphasize the difference from our dual spaces to $H^{-1}(\Omega)$. Only in case of the obstacle problem homogeneous Dirichlet boundary data can be postulated on the whole boundary, i. e. $\partial\Omega=\Gamma_D$. Then $H^1_D(\Omega)$ equals the common space $H^1_0(\Omega)$ and its dual space is given by the standard notation $H^{-1}(\Omega)=H^{-1}_D(\Omega)$. We also bear in mind that $L^2(\Omega)\supset H^1_D(\Omega)\supset H^1_0(\Omega)$ and $L^2(\Omega)\subset H^{-1}_D(\Omega)\subset H^{-1}(\Omega)$. The space $W$ corresponds to the respective distance function $G(\cdot)$.

TABLE 2. Functions, bilinear forms and contact areas $\Sigma$

| | $\psi$ | $\tau$ | $G(u)$ | $a(\cdot,\,\cdot)$ | $\Sigma$ | d$s$ |
|---|---|---|---|---|---|---|
| Obstacle | $\tilde{\phi}$ | $\mathrm{Id}(\cdot)$ | $u-\psi$ | $(\nabla\cdot,\,\nabla\cdot)$ | $\Omega$ | d$x$ |
| Simplified Signorini | $\tilde{\phi}$ | $\mathrm{tr}\,\vert_{\Gamma_C}(\cdot)$ | $\tau(u)-\psi$ | $(\nabla\cdot,\,\nabla\cdot)$ | $\Gamma_C$ | d$o$ |
| Signorini | $\frac{\tilde{\phi}-\phi}{\sqrt{1+\Delta\phi}}$ | $\mathrm{tr}\,\vert_{\Gamma_C}(\cdot)n$ | $\psi-\tau(u)$ | $(\mathbb{C}\varepsilon(\cdot),\,\varepsilon(\cdot))$ | $\Gamma_C$ | d$o$ |

Note that for any $W'$ there is a subset $W'_+$ defined as

$$W'_+ := \left\{\lambda\in W' \mid \langle\lambda,\,w\rangle\geq 0 \quad \forall\, w\in W, w\geq 0\right\}. \tag{6}$$

Since $\tau: V\to W$ and $\lambda\in W'$, we can declare an element $\tau^*\lambda\in V'$ by

$$\langle\tau^*\lambda,\,\cdot\rangle_{V',V}=\langle\lambda,\,\tau(\cdot)\rangle_{W',W}. \tag{7}$$

Using the notations given in Tab. 1 and Tab. 2, we arrive at the following feasible sets and penalty terms:

For the *classical obstacle conditions*

$$K:=\{v\in V: v\geq\ \psi \text{ a. e. in } \Omega\},$$
$$r(\gamma,u)(\cdot):=-\int_\Omega(\max\{\gamma(\psi-\tau(u)),0\})^3\tau(\cdot)\,\mathrm{d}x,$$

for *simplified Signorini's conditions*

$$K:=\{v\in V:\tau(v)\geq\ \psi \text{ a. e. on } \Gamma_C\},$$
$$r(\gamma,u)(\cdot):=-\int_{\Gamma_C}(\max\{\gamma(\psi-\tau(u)),0\})^3\tau(\cdot)\,\mathrm{d}o,$$

and for *Signorini's conditions*

$$K := \{v \in V : \tau(v) \leq \psi \text{ a. e. on } \Gamma_C\}$$
$$r(\gamma,u)(\cdot) := \int_{\Gamma_C} (\max\{\gamma(\tau(u)-\psi),0\})^3 \tau(\cdot)\,\mathrm{d}o.$$

We can use a common description of the convex sets of feasible states using the function $G(u)$ and the contact area $\Sigma$

$$K = \{v \in V : G(v) \geq 0 \text{ a. e. on } \Sigma\}.$$

The penalty term reads in general

$$r(\gamma,u)(\cdot) = \int_{\Sigma} \max\{-\gamma G(u),0\}^3 \tau(\cdot)\,\mathrm{d}s.$$

2.3. **Optimization problem.** From this point on, the contact problems appear as side conditions in an optimal control setting. This means that we consider the following optimal control problem

$$\left.\begin{aligned} \min_{q\in V^0} \quad & J(q,u) \\ \text{s.t.} \quad & a(u,v-u) \geq (q,\ v-u) \quad \forall v \in K \\ & u \in K \end{aligned}\right\} \tag{P}$$

where $\Omega \subset \mathbb{R}^d$ for $d = 2,3$ is a Lipschitz domain and K is the convex set defined by the contact conditions. The bilinear form $a$ is elliptic, i.e.,

$$a(u,u) \geq \beta \, \|u\|_V^2 \quad \forall u \in V \tag{8}$$

with a constant $\beta > 0$. Hence the corresponding operator $A$, where $\langle Au,\ v\rangle = a(u,\ v)$, is coercive. The goal functional $J$ is given by

$$J(q,u) = j(u) + g(q), \tag{9}$$

where $g : V^0 \to \mathbb{R}$ and $j : V \to \mathbb{R}$ are supposed to be three times continuously differentiable. Moreover, $j$ is assumed to be bounded from below. Whereas, we demand of $g$ that there exists a constant $\alpha > 0$ such that

$$g''(u)(h,h) \geq \alpha \, \|h\|_{V^0}^2 \quad \forall u,h \in V^0. \tag{10}$$

Hence, $g$ is convex. It is well known that the VI in (P), i.e.

$$u \in K, \quad a(u,v-u) \geq (q,\ v-u) \quad \forall v \in K, \tag{11}$$

can equivalently be reformulated by a complementarity system. Consequently, the optimal control problem (P) is equal to

$$\left.\begin{aligned} \min_{q\in V^0} \quad & J(q,u) \\ \text{s.t.} \quad & Au = q + \tau^*\lambda \\ & G(u) \geq 0, \quad \lambda \in W'_+, \quad \langle \lambda,\ G(u)\rangle_{W',W} = 0 \end{aligned}\right\} \tag{$P'$}$$

The corresponding Lagrange multiplier is $\lambda \in W'_+$. It is well known that the side condition holds a unique solution $u$ for any given load $q$, see for instance [25, Ch. 6.3]. Furthermore, it is easily seen that the corresponding solution operator $S : V' \ni q \mapsto u \in V$ is globally Lipschitz continuous with the constant $L = 1/\beta$, where $\beta$ is the constant in (8). The existence of at least one globally optimal solution can be shown by using standard arguments, but the uniqueness cannot be expected, because the problem (P) is not convex in general due to the non-linearity of $S$. Furthermore, $S$ is not Gâteaux-differentiable, since the directional derivative at $q$ in direction $h$ is in turn a solution of a VI of first kind, as shown in [24] and [7] for the constraints defined by the obstacle problem or Signorini's problem respectively.

## 3. Regularization and discretization

Our goal is a fast and robust algorithm for solving the optimization problem ($P'$). Since the solution operator $S$ is not Gâteaux-differentiable as mentioned above, we regularize the very problem and apply the finite element method (FEM).

3.1. **Regularization: Known and preliminary results.** We formulate the regularized problem

$$\left.\begin{aligned} \min_{q_\gamma \in V^0} \quad & J(q_\gamma, u_\gamma) \\ \text{s.t.} \quad & Au_\gamma + r(\gamma; u_\gamma) = q_\gamma. \end{aligned}\right\} \qquad (\mathrm{P}_\gamma)$$

Applying the same arguments as for the variational inequality (11), we get the existence and uniqueness of a solution to the semi-linear PDE in ($\mathrm{P}_\gamma$), i.e.

$$Au_\gamma + r(\gamma; u_\gamma) = q_\gamma \tag{12}$$

for every $\gamma > 0$. The associated solution operator is denoted by $S_\gamma : V' \to V$ and is Lipschitz continuous with the same Lipschitz constant as $S$, hence independent of $\gamma$. The solution for ($\mathrm{P}_\gamma$) exists as an immediate consequence of the same arguments that were applied for the unregularized problem.

Owing to the monotonicity of $r(\gamma; .)$, we can apply Stampacchia's classical technique, cf. [20], to prove the following

**Lemma 3.1.** *For every $q_\gamma \in V^0$ the unique solution $u_\gamma$ of* (12) *is essentially bounded.*

Following standard techniques, we deduct the Fréchet-differentiability of $r(\gamma; \cdot)$ from $L^\infty(\Sigma)$ to $V'$ and we derive first-order necessary optimality conditions for the regularized problems, see, e.g., [33, Ch. 4]. In this way we obtain the following result:

**Proposition 3.2.** *Let $q_\gamma$ be a local optimum of* ($\mathrm{P}_\gamma$) *with associated state $u_\gamma = S_\gamma(q_\gamma)$. Then there exist $\lambda_\gamma, \mu_\gamma \in L^2(\Sigma) \subset W'$ and $p_\gamma, \theta_\gamma \in V$ such that*

$$Au_\gamma = q_\gamma + \tau^*(\lambda_\gamma) \tag{13a}$$

$$\tau^*(\lambda_\gamma) + r(\gamma; u_\gamma) = 0 \tag{13b}$$

$$A^* p_\gamma = \nabla j(u_\gamma) - \tau^*(\mu_\gamma), \tag{13c}$$

$$p_\gamma + \nabla g(q_\gamma) = 0, \tag{13d}$$

$$p_\gamma - \theta_\gamma = 0, \tag{13e}$$

$$\tau^*(\mu_\gamma) - \partial_u r(\gamma; u_\gamma)(\theta_\gamma) = 0. \tag{13f}$$

*Note that $p_\gamma$ and thus $\theta_\gamma$ and $\mu_\gamma$ are uniquely defined by* (13d).

The variable $p_\gamma$ is known as the adjoint state. Note further that $\mu_\gamma$ and $\theta_\gamma$ can be eliminated directly from the system, but we introduced them for reasons of comparison with later optimality systems. We now address the convergence for $\gamma \to \infty$. Concerning the state equation, the following approximation result holds true:

**Lemma 3.3.** *Let $q \in V^0$ be given and denote by $u, u_\gamma \in V$ the solutions to* (11) *and* (12)*, respectively. Then $u_\gamma \to u$ strongly in $V$ as $\gamma \to \infty$.*

The following first-order necessary optimality conditions for (P) can be concluded from the above results.

**Theorem 3.4.** *We assume that the requirements of Lemma 2.1 for all $u_\gamma$, $\psi$ and $\Omega$ hold.*

(1) *For every $\gamma > 0$ there is a globally optimal solution of* ($\mathrm{P}_\gamma$), *denoted by $q_\gamma$. If $\gamma \to \infty$, then every sequence $\{q_\gamma\}$ of global minimizers of* ($\mathrm{P}_\gamma$) *admits a weak accumulation point $\bar{q} \in V^0$. Every weak accumulation point is also a strong accumulation point, i. e., $q_\gamma \to \bar{q}$ strongly in $V^0$, and each of these accumulation points is a global minimizer of* (P).

(2) *If $q_\gamma \to \bar{q}$ in $V^0$, then the associated sequence of solutions to* (13) *fulfills*

$$u_\gamma \to \bar{u} \qquad \text{in } V, \tag{14a}$$

$$\lambda_\gamma \to \bar{\lambda} \qquad \text{in } W', \tag{14b}$$

$$\tau^*\lambda_\gamma \to \tau^*\bar{\lambda} \qquad \text{in } V', \tag{14c}$$

$$p_\gamma \rightharpoonup \bar{p} \qquad \text{in } V, \tag{14d}$$

$$\theta_\gamma \rightharpoonup \bar{\theta} \qquad \text{in } V, \tag{14e}$$

$$\mu_\gamma \rightharpoonup \bar{\mu} \qquad \text{in } W', \tag{14f}$$

$$\tau^*\mu_\gamma \rightharpoonup \tau^*\bar{\mu} \qquad \text{in } V', \tag{14g}$$

*and the limit satisfies the following optimality system:*

$$A\bar{u} = \bar{q} + \tau^*\bar{\lambda}, \tag{15a}$$

$$G(\bar{u}) \geq 0 \text{ a.e. on } \Gamma_C, \quad \bar{\lambda} \geq 0 \text{ in } W', \quad \langle \bar{\lambda},\, G(\bar{u})\rangle_{W',W} = 0, \tag{15b}$$

$$A^*\bar{p} = \nabla j(\bar{u}) - \tau^*(\bar{\mu}), \tag{15c}$$

$$\bar{p} + \nabla g(\bar{q}) = 0, \tag{15d}$$

$$\bar{p} - \bar{\theta} = 0, \tag{15e}$$

$$\langle \bar{\lambda},\, \tau(\bar{\theta})\rangle_{W',W} = 0, \quad \langle \bar{\mu},\, G(\bar{u})\rangle_{W',W} = 0, \quad \langle \bar{\mu},\, \tau(\bar{\theta})\rangle_{W',W} \geq 0, \tag{15f}$$

We present only the proof in case of Signorini's contact condition. It was already proven for the obstacle case in [22, Thm. 3.4.]. The case of simplified Signorini's contact condition is omitted, since it is a simple combination of both of them.

The proof follows by standard arguments known from other types of regularization, cf. e.g. [22, 30]. However, we have to consider the different sign in the penalty term and the normal trace $\tau$.

*Proof.* Using standard arguments, we can prove most of the parts. For the convenience of the reader, we add the complete proof in the appendix. However, we present the details of the trickiest part, namely the very first complementarity relation in (15f)

$$\langle \bar{\lambda},\, \tau(\bar{\theta})\rangle_{W',W} = 0.$$

We cannot adopt the techniques presented for instance in [22] straight forwardly, since $V \neq W$ in the contact problem at hand. We tackle this problem by analyzing auxiliary problems.

For this purpose, we construct an adequate family of test functions $m_\gamma$ in $V = H^1_D(\Omega)^d$. Let $m_\gamma \in V$ be the solution of the mixed boundary problem

$$\left.\begin{aligned} Am_\gamma &= q_\gamma && \text{in } \Omega,\\ m_\gamma &= 0 && \text{on } \Gamma_D,\\ \sigma(m_\gamma)\cdot n &= 0 && \text{on } \Gamma_N,\\ m_\gamma \cdot n &= \max\{\tau(u_\gamma) - \psi, 0\} && \text{on } \Gamma_C,\\ m_\gamma \cdot t_i &= 0 && \text{on } \Gamma_C. \end{aligned}\right\} \tag{$P_{\text{aux}}$}$$

Here, $n$ denotes the normal vector and $t_1$ the corresponding tangential vector for $\Omega \subset \mathbb{R}^2$. For $\Omega \subset \mathbb{R}^3$, there are two orthogonal, tangential vectors $t_1, t_2$. We recall the relation for the stress $\sigma = \mathbb{C}\varepsilon$, cf. 1. The boundary conditions on $\Gamma_C$ are equivalent to inhomogeneous Dirichlet data

$$\tau(U_{0,\gamma}) = \max\{\tau(u_\gamma) - \psi, 0\} \qquad \text{on } \Gamma_C,$$

for a function $U_{0,\gamma} \in V$ due to Lemma 2.1. Note, by construction $U_{0,\gamma} \cdot t_i$ is zero. Problem ($P_{\text{aux}}$) is a standard, static, linear elastic problem with mixed boundary conditions. It is well known that this problem has a unique solution, see, e. g. [31, Ch. 25.3]. Therefore, the description of $m_\gamma$ is well posed. Furthermore, the limit case for $\gamma \to \infty$ is well posed, too, because

$$\left.\begin{aligned} A\bar{m} &= \bar{q} && \text{in } \Omega \\ \bar{m} &= 0 && \text{on } \Gamma_D \cup \Gamma_C \\ \sigma_n(\bar{m}) &= 0 && \text{on } \Gamma_N, \end{aligned}\right\} \tag{$\bar{P}_{\text{aux}}$}$$

has also a unique solution $\bar{m}$ in $V$. Additionally, it holds $m_\gamma \to \bar{m}$ in $V$; which is proven next. Solving the substitute problem for $v_\gamma = m_\gamma - U_{0,\gamma}$

$$\left.\begin{aligned} Av_\gamma &= q_\gamma - AU_{0,\gamma} && \text{in } \Omega, \\ v_\gamma &= 0 && \text{on } \Gamma_D \cup \Gamma_C, \\ \sigma_n(v_\gamma) &= 0 && \text{on } \Gamma_N \end{aligned}\right\} \tag{$P'_{\text{aux}}$}$$

instead of ($P_{\text{aux}}$), we achieve

$$\begin{aligned} \beta\|v_\gamma - \bar{m}\|_V^2 &\le a(v_\gamma - \bar{m}, v_\gamma - \bar{m}) = \big(q_\gamma - \bar{q},\ v_\gamma - \bar{m}\big) + a(-U_{0,\gamma}, v_\gamma - \bar{m}) \\ &\le c\|q_\gamma - \bar{q}\|_{V^0}\, \|v_\gamma - \bar{m}\|_V + c\|U_{0,\gamma}\|_V\, \|v_\gamma - \bar{m}\|_V. \end{aligned}$$

We use again Lemma 2.1 and conclude

$$\begin{aligned} \|m_\gamma - \bar{m}\|_V &\le \frac{c}{\beta}\|q_\gamma - \bar{q}\| + \left(\frac{c}{\beta} + 1\right)\|U_{0,\gamma}\| \\ &\le \frac{c}{\beta}\|q_\gamma - \bar{q}\| + \left(\frac{c}{\beta} + 1\right) C\|\max(\tau(u_\gamma) - \psi, 0)\|_{\Gamma_C} \\ &\to 0 \quad \text{as } \gamma \to \infty\,. \end{aligned}$$

One can show that

$$\begin{aligned} \int_{\Gamma_C} \big[\max\big(\gamma(\tau(u_\gamma) - \psi), 0\big)\,\tau(p_\gamma)\big]^2\, \mathrm{d}o &\le \frac{1}{3\gamma}\big(\nabla j(u_\gamma),\ p_\gamma\big) \\ &\le \frac{1}{3\gamma\beta}\|\nabla j(u_\gamma)\|_{V'}^2 \ \to\ 0 \qquad \text{as } \gamma \to \infty\,. \end{aligned} \tag{16}$$

The details are presented in the appendix. Finally, combining inequality (16) with

$$\begin{aligned} &\|\max\big(\gamma(\tau(u_\gamma) - \psi), 0\big)\|_{L^4(\Gamma_C)}^4 \\ &\quad = -\int_{\Gamma_C} -\Big(\max\big(\gamma(\tau(u_\gamma) - \psi), 0\big)^3 \cdot \gamma\max\big(\tau(u_\gamma) - \psi, 0\big)\Big)\, \mathrm{d}o \\ &\quad = -\gamma\big(\lambda_\gamma,\ \tau(m_\gamma)\big)_{\Gamma_C} \end{aligned}$$

leads to

$$\begin{aligned}
|\big(\lambda_\gamma,\ \tau(p_\gamma)\big)|^2 &= \Big|\int_{\Gamma_C} -\big[\max\big(\gamma(\tau(u_\gamma)-\psi),0\big)\big]^3\,\tau(p_\gamma)\,\mathrm{d}o\Big|^2 \\
&\le \|\max\big(\gamma(\tau(u_\gamma)-\psi),0\big)\|^4_{L^4(\Gamma_C)}\|\max\big(\gamma(\tau(u_\gamma)-\psi),0\big)\tau(p_\gamma)\|^2_{L^2(\Gamma_C)} \\
&\le \gamma\big|\big(\lambda_\gamma,\ \tau(m_\gamma)\big)\big|\frac{1}{3\gamma}\big|\big(\nabla j(u_\gamma),\ p_\gamma\big)\big| \\
&= \frac{1}{3}\big|\big(\nabla j(u_\gamma),\ p_\gamma\big)\big|\big|\big(\lambda_\gamma,\ \tau(m_\gamma)\big)\big| \quad \to \quad C\big|\big(\bar\lambda,\ \tau(\bar m)\big)\big| = 0,
\end{aligned}$$

since $\tau(\bar m)=0$. Due to (15e), the first complementarity condition in (15f) is shown.

The other complementarity conditions follow easily. □

After regularizing the optimal control problem ($\mathrm{P}_\gamma$) and establishing the optimality systems (13) and (15), we deduce a discrete analog next.

3.2. **Discretization.** The discretization is performed by means of the FEM with $d$-linear Ansatz-functions. We presume that the domain $\Omega$ is polygonal. There is a conform triangulation $\mathbb{T}_h=\{\mathscr{T}\}$ of quadrilaterals such that $\bar\Omega=\cup\mathscr{T}$. Hence, we use finite element spaces $V_h\subset V$ and $W_h\subset W$, which are boundary conform. In case of the optimal control problem subject to Signorini's problem, those functions spaces are defined as

$$V_h := \Big\{v_h\in V \mid \forall\ \mathscr{T}\in\mathbb{T}_h : v_{h|\mathscr{T}}\in\mathscr{Q}_1(\mathscr{T};\mathbb{R}^d)\Big\}\subset C(\bar\Omega,\Gamma_D)^d$$

and

$$W_h := \big\{w_h\in W \mid \forall\ \mathscr{E}\in\mathbb{E}_h : w_{h|\mathscr{E}}\in\mathscr{P}_1(\mathscr{E};\mathbb{R})\big\}\subset C(\bar\Omega,\Gamma_D),$$

where $\mathbb{E}_h=\{\mathscr{T}\cap\Gamma_C\}$ and $\mathscr{P}_1$, $\mathscr{Q}_1$ are spaces of linear or $d$-linear functions. The finite element dimension is denoted by $N$. Furthermore, we require the mesh to have patch structure and restrict the number of hanging nodes to one per hyperplane.

In view of this discretization we investigate the discretized optimal control problem, i. e. find $u_{\gamma,h}\in V_h$ and $q_{\gamma,h}\in V_h$, which solve

$$\begin{aligned}
\min\quad & J(q_{\gamma,h},u_{\gamma,h}) \\
\text{s.t.}\quad & Au_{\gamma,h}+r(\gamma;u_{\gamma,h})=q_{\gamma,h}.
\end{aligned} \tag{$\mathrm{P}_{\gamma,h}$}$$

Applying the very same arguments as in the continuous case, we get the following discrete optimality conditions.

**Theorem 3.5.** *Let $q_{\gamma,h}$ be a local unique optimum of the discrete regularized problem* ($\mathrm{P}_{\gamma,h}$) *with the associated state $u_{\gamma,h} = S_{\gamma,h}(q_{\gamma,h})$. Then there exist $\lambda_{\gamma,h},\mu_{\gamma,h}\in W_h$ and $p_{\gamma,h},\theta_{\gamma,h}\in V_h$ such that for any $\varphi\in V_h$*

$$a\big(u_{\gamma,h},\ \varphi_h\big) = \big(q_{\gamma,h},\ \varphi_h\big)+\big(\lambda_{\gamma,h},\ \tau(\varphi_h)\big)_{\Gamma_C}, \tag{17a}$$

$$\big(\lambda_{\gamma,h},\ \tau(\varphi_h)\big)_{\Gamma_C}+r(\gamma;u_{\gamma,h})(\varphi_h)=0, \tag{17b}$$

$$a\big(\varphi_h,\ p_{\gamma,h}\big) = \nabla j(u_{\gamma,h})(\varphi)-\big(\mu_{\gamma,h},\ \tau(\varphi_h)\big)_{\Gamma_C}, \tag{17c}$$

$$\big(p_{\gamma,h},\ \varphi_h\big)+\nabla g(q_{\gamma,h})(\varphi_h)=0, \tag{17d}$$

$$\big(p_{\gamma,h}-\theta_{\gamma,h},\ \varphi_h\big)=0, \tag{17e}$$

$$\big(\mu_{\gamma,h},\ \tau(\varphi_h)\big)_{\Gamma_C}-\partial_u r(\gamma;u_{\gamma,h})(\theta_{\gamma,h},\varphi_h)=0. \tag{17f}$$

*Note that $p_{\gamma,h}$ and thus $\theta_{\gamma,h}$ and $\mu_{\gamma,h}$ are uniquely defined by* (17d).

The gradient $\nabla g:\ V^0 \ \to\ V^0$ is a strongly monotone and continuous operator due to (10). Hence, (13d)

$$p_\gamma + \nabla g(q_\gamma) = 0$$

can be resolved for $q_\gamma \in V^0$, i.e., there is a mapping $Q_c : V^0 \to V^0$ such that

$$p_\gamma + \nabla g(q_\gamma) = 0 \quad \Longleftrightarrow \quad q_\gamma = Q_c(p_\gamma).$$

In combination with the $L^2$-projection onto $V_h$, we get a mapping $Q : V_h \to V_h$ such that the relation

$$\big(p_{\gamma,h},\ \varphi_h\big) + \nabla g(q_{\gamma,h})(\varphi_h) = 0 \qquad \forall \varphi_h \in V_h$$

is equivalent to

$$\big(q_{\gamma,h},\ \varphi_h\big) = \big(Q(p_{\gamma_h}),\ \varphi_h\big) \qquad \forall \varphi_h \in V_h.$$

Therefore, the discrete system (17) can be reduced to a discrete, nonlinear system with only two unknowns

$$Au_{\gamma,h} = Q(p_{\gamma,h}) - r(\gamma; u_{\gamma,h}), \tag{18a}$$
$$A^* p_{\gamma,h} = \nabla j(u_{\gamma,h}) - \partial_u r(\gamma; u_{\gamma,h})(p_{\gamma,h}). \tag{18b}$$

## 4. An a-posteriori estimation of the total error

In this section, we discuss the errors after the discretization of the regularized problem.

4.1. **Error representation.** Based on the DWR-method, we analyze the error between a local solution of (P) and the approximated solution of ($\mathrm{P}_{\gamma,h}$) w.r.t. the objective, i. e. $J(\bar{q},\bar{u}) - J(\tilde{q}_{\gamma,h},\tilde{u}_{\gamma,h})$. We begin by defining the following MPEC-Lagrangian:

$$\begin{aligned} &\mathscr{L} : V^0 \times V \times V' \times V \times W' \times V \to \mathbb{R};\\ &\mathscr{L}(q,u,\lambda,p,\mu,\theta) := J(q,u) - \langle Au - q - \lambda,\ p\rangle + \langle \mu,\ \psi - \tau(u)\rangle_{W',W} - \langle \lambda,\ \theta\rangle. \end{aligned} \tag{19}$$

To ease the notation we do not distinguish between $\lambda$ in $W'$ and $\tau^*\lambda \in V'$ anymore, i. e. for $v \in V$ we use the notation $\langle \lambda,\ v\rangle$ even so we actually perform $\langle \lambda,\ \tau(v)\rangle_{W',W}$. In the following, we abbreviate the collection of variables $\xi := (q,u,\lambda,p,\mu,\theta)$ and their approximations $\tilde{\xi}_{\gamma,h} := (\tilde{q}_{\gamma,h},\tilde{u}_{\gamma,h},\tilde{\lambda}_{\gamma,h},\tilde{p}_{\gamma,h},\tilde{\mu}_{\gamma,h},\tilde{\theta}_{\gamma,h})$.

**Definition 4.1.** *The dual residual is given by*

$$\tilde{\rho}^*(\cdot) := \rho^*(\tilde{q}_{\gamma,h},\tilde{u}_{\gamma,h},\tilde{p}_{\gamma,h})(\cdot) := \nabla j(\tilde{u}_{\gamma,h})(\cdot) - a(\cdot,\tilde{p}_{\gamma,h}) - \partial_u r(\gamma;\tilde{u}_{\gamma,h})(\tilde{p}_{\gamma,h},\cdot),$$

*the control residual by*

$$\tilde{\rho}^q(\cdot) := \rho^q(\tilde{q}_{\gamma,h},\tilde{p}_{\gamma,h})(\cdot) := \nabla g(\tilde{q}_{\gamma,h})(\cdot) - (\tilde{p}_{\gamma,h},\cdot),$$

*and the primal residual by*

$$\tilde{\rho}(\cdot) := \rho(\tilde{q}_{\gamma,h},\tilde{u}_{\gamma,h})(\cdot) := -a(\tilde{u}_{\gamma,h},\cdot) + (\tilde{q}_{\gamma,h},\cdot) - r(\gamma;\tilde{u}_{\gamma,h})(\cdot).$$

The approximated solution $\tilde{\xi}_{\gamma,h}$ is determined by solving (18) using an inexact method. Hence, as an $L^2$-projection $\tilde{q}_{\gamma,h}$ equals $Q(\tilde{p}_{\gamma,h})$ in $V_h$ but not in $V$. This implies that we cannot ignore the control residual and must not substitute $\tilde{q}_{\gamma,h}$ in the primal residual. Using all three residuals, we get the following theorem concerning the error identity.

**Theorem 4.1.** *Let* $(\bar{q},\bar{u})$ *be the solution of problem* (P) *and* $(\tilde{q}_{\gamma,h},\tilde{u}_{\gamma,h})$ *be an approximation of the solution of the regularized and discretized solution* $(\mathrm{P}_{\gamma,h})$. *Then holds the error identity.*

$$\begin{aligned} J(\bar{q},\bar{u}) &- J(\tilde{q}_{\gamma,h},\tilde{u}_{\gamma,h}) \\ &= \left\langle \frac{1}{2}(\bar{\mu}+\tilde{\mu}_{\gamma,h}), \tau(\bar{u})-\tau(\tilde{u}_{\gamma,h}) \right\rangle + \left\langle \bar{\lambda}-\tilde{\lambda}_{\gamma,h}, \frac{1}{2}(\bar{\theta}+\tilde{\theta}_{\gamma,h}) \right\rangle \\ &\quad + \frac{1}{2}\Big[\tilde{\rho}(\bar{p}-\tilde{p}_{\gamma,h}) + \tilde{\rho}^{q}(\bar{q}-\tilde{q}_{\gamma,h}) + \tilde{\rho}^{*}(\bar{u}-\tilde{u}_{\gamma,h})\Big] \\ &\quad + \tilde{\rho}(\tilde{p}_{\gamma,h}) + \mathscr{R}^{(3)}_{\mathrm{reg}} \end{aligned} \tag{20}$$

*with a higher order remainder term*

$$\mathscr{R}^{(3)}_{\mathrm{reg}} := \frac{1}{2}\int_0^1 J'''(\tilde{\xi}_{\gamma,h}+t(\bar{\xi}-\tilde{\xi}_{\gamma,h}))(\bar{\xi}-\tilde{\xi}_{\gamma,h})^3\, t(t-1)\,dt.$$

*Proof.* Due to (15a) and (15f), we have $J(\bar{q},\bar{u}) = \mathscr{L}(\bar{\xi})$. We apply the trapezoidal rule such that (17a) yields

$$\begin{aligned} J(\bar{q},\bar{u}) &- J(\tilde{q}_{\gamma,h},\tilde{u}_{\gamma,h}) \\ &= \mathscr{L}(\bar{\xi}) - \mathscr{L}(\tilde{\xi}_{\gamma,h}) + \big(\tilde{\mu}_{\gamma,h},\ \psi-\tau(\tilde{u}_{\gamma,h})\big) - \Big(\tilde{\lambda}_{\gamma,h},\ \tilde{\theta}_{\gamma,h}\Big) \\ &\quad - \Big(A\tilde{u}_{\gamma,h}-\tilde{q}_{\gamma,h}-\tilde{\lambda}_{\gamma,h},\ \tilde{p}_{\gamma,h}\Big) \\ &= \frac{1}{2}\mathscr{L}'(\bar{\xi})(\bar{\xi}-\tilde{\xi}_{\gamma,h}) + \frac{1}{2}\mathscr{L}'(\tilde{\xi}_{\gamma,h})(\bar{\xi}-\tilde{\xi}_{\gamma,h}) + \mathscr{R}'_{\mathrm{reg}} \\ &\quad + \big(\tilde{\mu}_{\gamma,h},\ \psi-\tau(\tilde{u}_{\gamma,h})\big) - \Big(\tilde{\lambda}_{\gamma,h},\ \tilde{\theta}_{\gamma,h}\Big) - \Big(A\tilde{u}_{\gamma,h}-\tilde{q}_{\gamma,h}-\tilde{\lambda}_{\gamma,h},\ \tilde{p}_{\gamma,h}\Big) \end{aligned}$$

with

$$\begin{aligned} \mathscr{R}'_{\mathrm{reg}} &:= \frac{1}{2}\int_0^1 \mathscr{L}'''(\tilde{\xi}_{\gamma,h}+t(\bar{\xi}-\tilde{\xi}_{\gamma,h}))(\bar{\xi}-\tilde{\xi}_{\gamma,h})^3\, t(t-1)\,dt \\ &= \frac{1}{2}\int_0^1 \Big( j'''(\tilde{u}_{\gamma,h}+t(\bar{u}-\tilde{u}_{\gamma,h}))(\bar{u}-\tilde{u}_{\gamma,h})^3 \\ &\qquad + g'''(\tilde{q}_{\gamma,h}+t(\bar{q}-\tilde{q}_{\gamma,h}))(\bar{q}-\tilde{q}_{\gamma,h})^3\Big) t(t-1)\,dt \\ &= \frac{1}{2}\int_0^1 J'''\Big(\tilde{\xi}_{\gamma,h}+t(\bar{\xi}-\tilde{\xi}_{\gamma,h})\Big)\Big(\bar{\xi}-\tilde{\xi}_{\gamma,h}\Big)^3 t(t-1)\,dt \\ &= \mathscr{R}^{(3)}_{\mathrm{reg}}. \end{aligned}$$

In view of (19), (13), and (15), we arrive at

$$\begin{aligned} J(\bar{q},\bar{u}) &- J(\tilde{q}_{\gamma,h},\tilde{u}_{\gamma,h}) \\ &= \left\langle \frac{1}{2}(\bar{\mu}+\tilde{\mu}_{\gamma,h}), \tau(\bar{u})-\tau(\tilde{u}_{\gamma,h}) \right\rangle + \left\langle \bar{\lambda}-\tilde{\lambda}_{\gamma,h}, \frac{1}{2}(\bar{\theta}+\tilde{\theta}_{\gamma,h}) \right\rangle \\ &\quad + \frac{1}{2}\Big[ -\Big(A\tilde{u}_{\gamma,h}-\tilde{q}_{\gamma,h}-\tilde{\lambda}_{\gamma,h},\ \bar{p}-\tilde{p}_{\gamma,h}\Big) + \big(\nabla g(\tilde{q}_{\gamma,h})+\tilde{p}_{\gamma,h},\ \bar{q}-\tilde{q}_{\gamma,h}\big) \\ &\qquad + \big(\nabla j(\tilde{u}_{\gamma,h}) - A^{*}\tilde{p}_{\gamma,h}-\tilde{\mu}_{\gamma,h},\ \bar{u}-\tilde{u}_{\gamma,h}\big)\Big] \\ &\quad - \Big(A\tilde{u}_{\gamma,h}-\tilde{q}_{\gamma,h}-\tilde{\lambda}_{\gamma,h},\ \tilde{p}_{\gamma,h}\Big) + \frac{1}{2}\Big(\bar{\lambda}-\tilde{\lambda}_{\gamma,h},\ \tilde{p}_{\gamma,h}-\tilde{\theta}_{\gamma,h}\Big) + \mathscr{R}^{(3)}_{\mathrm{reg}}, \end{aligned} \tag{21}$$

after using the complementarity conditions in (15f) to switch the sign of $\big(\bar{\lambda},\ \bar{\theta}\big)$ and $(\bar{\mu},\ \psi-\tau(\bar{u}))$. Note that $p_{\gamma,h}$ equals $\theta_{\gamma,h}$ almost everywhere, hence we drop the term with $\tilde{p}_{\gamma,h}-\tilde{\theta}_{\gamma,h}$. Using these information

(21) can be rewritten as

$$\begin{aligned} J(\bar{q},\bar{u}) - J(\tilde{q}_{\gamma,h},\tilde{u}_{\gamma,h}) & \\ &= \left\langle \frac{1}{2}(\bar{\mu}+\tilde{\mu}_{\gamma,h}), \tau(\bar{u})-\tau(\tilde{u}_{\gamma,h}) \right\rangle + \left\langle \bar{\lambda}-\tilde{\lambda}_{\gamma,h}, \frac{1}{2}(\bar{\theta}+\tilde{\theta}_{\gamma,h}) \right\rangle \\ &\quad + \frac{1}{2}\left[\tilde{\rho}(\bar{p}-\tilde{p}_{\gamma,h}) + \tilde{\rho}^{q}(\bar{q}-\tilde{q}_{\gamma,h}) + \tilde{\rho}^{*}(\bar{u}-\tilde{u}_{\gamma,h})\right] \\ &\quad + \tilde{\rho}(\tilde{p}_{\gamma,h}) + \mathscr{R}_{\text{reg}}^{(3)}, \end{aligned} \tag{22}$$

which concludes the proof. □

The goal functional $J$ is three times continuously Fréchet-differentiable due to the assumptions on $j$ and $g$, cf. (9) and (10). Therefore, the Lagrangian $\mathscr{L}$ is three times continuously Fréchet-differentiable, too. If we assume further that the trilinear forms $j'''(u)$ and $g'''(q)$ are uniformly bounded by constants $c_j$ and $c_g$, then

$$|\mathscr{R}_{\text{reg}}^{(3)}| \leq \frac{1}{12}\left(c_j \|\tilde{u}_{\gamma,h}-\bar{u}\|_{H^1(\Omega)}^3 + c_g \|\tilde{q}_{\gamma,h}-\bar{q}\|_{L^2(\Omega)}^3\right)$$

such that $\mathscr{R}_{\text{reg}}^{(3)}$ can be neglected in a neighborhood of $(\bar{u},\bar{q})$. Trivially, it holds $\mathscr{R}_{\text{reg}}^{(3)} = 0$, if $J$ is of tracking type, i.e., if $j$ and $g$ are squared norms.

However, there is no guarantee that $\tilde{q}_{\gamma,h} \to \bar{q}$ and $\tilde{u}_{\gamma,h} \to \bar{u}$ in general. After all, only the side condition (11) is a convex problem, but not the optimal control problem (P) as a whole, and thus multiple local minima can occur, cf. Theorem 3.4.

4.2. **Towards estimation: Differentiation with respect to the regularization parameter.** In the end, we want to evaluate the error $\xi - \tilde{\xi}_{\gamma,h}$. However, we typically do not know the exact solution and need some kind of approximation. We start by performing Taylor expansion with respect to the regularization parameter. [1] Here, we differentiate the system (17) with respect to the regularization parameter $\gamma$ and exploit the result in the next section.

Since $g''$ is coercive, cf. (10), we get that $g''(q_\gamma) : V^0 \times V^0 \to \mathbb{R}$ is a homeomorphism. Applying the implicit function theorem yields that the mapping $Q$ is continuously Fréchet-differentiable and

$$Q'(p_{\gamma,h})v_h = -g''(Q(p_{\gamma,h}))^{-1} v_h \qquad \forall v_h \in V_h^0. \tag{23}$$

With a major assumption, similar to the assumption for instance in [22], we see that the differentiation is well posed.

**Assumption 4.1.1.** *We assume that $u_{\gamma,h}$ and $p_{\gamma,h}$ are such that for every $v_h$ in $V_h$*

$$\int_\Omega j''(u_{\gamma,h})(v_h,v_h)\,\mathrm{d}x - \int_\Sigma 6\gamma^2 \max(\gamma(-G(u_{\gamma,h})),0)\tau(p_{\gamma,h})\tau(v_h)\tau(v_h)\mathrm{d}s \geq 0.$$

This assumption holds true for instance, when

(1) The first part $j$ of the objective, acting on the state $u_{\gamma,h}$, is supposed to be convex in $V_h$, i. e. for every $v_h$ in $V_h$ it holds $j''(u_{\gamma,h})(v_h,v_h) \geq 0$ a.e. in $\Omega$.
(2) We assume that $u_{\gamma,h}$ and $p_{\gamma,h}$ are such that

$$\max(\gamma(-G(u_{\gamma,h})),0)\tau(p_{\gamma,h}) \leq 0 \text{ a.e. on } \Sigma.$$

---

[1] If the expansion is well posed, this leads to a general approximation of a function $f$ depending on $1/\gamma$

$$f(0) = f(1/\gamma) - 1/\gamma f'(1/\gamma) + o\left(1/\gamma^2\right) \qquad \text{and for } g(x) = f(1/x) \text{ with } g'(x) = -f'(1/x)/x^2$$
$$g(\infty) = g(\gamma) + \gamma g'(\gamma) + o\left(1/\gamma^2\right).$$

Obviously, tracking type goal functionals fulfill the condition for $j$. If we choose such a functional

$$J(q,u) := \frac{1}{2}\|u-u_D\|_{V^0}^2 + \frac{\alpha}{2}\|q-q_D\|_{V^0}^2,$$

the relation between the control variable $q$ and the adjoint state $p$ also simplifies and it holds

$$q := Q(p) := -\frac{1}{\alpha}p + q_D.$$

Taking this into account, we arrive at the condition for each $v_h$ in $V_h$

$$6\gamma^3 \int_\Sigma \max(-G(u_{\gamma,h})),0)\alpha\tau(q_D - q_{\gamma,h})\tau(v_h)^2 \mathrm{d}s \le \|v_h\|_{V^0}^2. \tag{24}$$

We could for example assume the following for Signorini's contact condition. Where $u_{\gamma,h}$ is not feasible (on $\Gamma_C$), we require that $q_{\gamma,h}\cdot n \le q_D \cdot n$ with the outer normal $n$; i. e. $p_{\gamma,h}\cdot n \le 0$. The left hand side of (24) is negative and hence Assumption 4.1.1 is fulfilled. However, numerical examinations show that we cannot expect such behavior in general.

**Theorem 4.2.** *Under Assumption 4.1.1 the system*

$$\begin{aligned} A\dot{u}_{\gamma,h} - Q'(p_{\gamma,h})\dot{p}_{\gamma,h} + \partial_u r(\gamma;u_{\gamma,h})\dot{u}_{\gamma,h} &= z_1 \\ A^*\dot{p}_{\gamma,h} - j''(u_{\gamma,h})\dot{u}_{\gamma,h} + \partial_{u,u}^2 r(\gamma;u_{\gamma,h})(p_{\gamma,h},\dot{u}_{\gamma,h}) + \partial_u r(\gamma;u_{\gamma,h})\dot{p}_{\gamma,h} &= z_2 \end{aligned} \tag{25}$$

*has a unique solution $(\dot{u}_{\gamma,h},\dot{p}_{\gamma,h})$ for any $(z_1,z_2)$ in $V_h^0\times V_h^0$.*

The proof requires the following theorem, which addresses saddle point problems, cf. [5, Theorem 3.4].

**Theorem 4.3.** *The block matrix*

$$\mathscr{A} = \begin{pmatrix} A & B \\ B & -C \end{pmatrix}$$

*is given. Assume that $H$, the symmetric part of $A$, is positive semi-definite, $B$ has full rank, and $C$ is symmetric positive semi-definite (possibly zero). Then*

$$\ker(H)\cap\ker(B) = \{0\} \qquad \textit{implies} \qquad \mathscr{A} \textit{ invertible}. \tag{26}$$

With this Theorem we can proof the previous Theorem 4.2.

*Proof of Theorem 4.2.* Let $V_h$ be equipped with the basis $\{\varphi_1,\dots,\varphi_N\}$. Using this discretization and defining the respective vectors $\vec{d}_u,\vec{d}_p$ and $\vec{z}_{1,2}$ in $\mathbb{R}^N$, and matrices in $\mathbb{R}^{N\times N}$ with the components

$$\begin{aligned} (A_\mathrm{M})_{i,j} &:= a(\varphi_i,\varphi_j), & (R_{\mathrm{M},1})_{i,j} &:= \left(\partial_u r(\gamma;u_{\gamma,h})(\varphi_i),\ \varphi_j\right), \\ (J_\mathrm{M})_{i,j} &:= \left(j''(u_{\gamma,h})\varphi_i,\ \varphi_j\right), & (R_{\mathrm{M},2})_{i,j} &:= \left(\partial_u^2 r(\gamma;u_{\gamma,h})(p_{\gamma,h},\varphi_i),\ \varphi_j\right), \\ (Q_\mathrm{M})_{i,j} &:= \left(Q'(p_{\gamma,h})\varphi_i,\ \varphi_j\right), \end{aligned}$$

leads to the linear system

$$\begin{pmatrix} J_\mathrm{M} - R_{\mathrm{M},2} & -A_\mathrm{M} - R_{\mathrm{M},1} \\ -A_\mathrm{M} - R_{\mathrm{M},1} & Q_\mathrm{M} \end{pmatrix} \begin{pmatrix} \vec{d}_u \\ \vec{d}_p \end{pmatrix} = \begin{pmatrix} -\vec{z}_2 \\ -\vec{z}_1 \end{pmatrix}. \tag{27}$$

When the saddle point problem (27) has a unique solution $(\vec{d}_u,\vec{d}_p)$ for any $(\vec{z}_2,\vec{z}_1)$ in $\mathbb{R}^{N+N}$, the theorem is proven. Under the given Assumption 4.1.1 and (4), it obviously yields that $J_\mathrm{M} - R_{\mathrm{M},2}$ is symmetric and positive semi-definite. Whereas, $Q_\mathrm{M}$ is negative semi-definite, because of testing (23) with $Q'(p_{\gamma,h})v_h$ holds

$$\left(g''(Q(p_{\gamma,h}))Q'(p_{\gamma,h}),\ Q'(p_{\gamma,h})v_h\right)_{V_h^0} = -\left(Q'(p_{\gamma,h})v_h,\ v_h\right)_{V_h^0}$$

for any $v_h \neq 0$. Due to (10), this is equivalent to

$$-\left(Q'(p_{\gamma,h})v_h,\ v_h\right)_{V_h^0} \ge \alpha\|Q'(p_{\gamma,h})v_h\|_{V_h^0} \ge 0.$$

At last, we need to show that $-A_{\mathrm{M}} - R_{\mathrm{M},1}$ has full rank and that

$$\ker(J_M - R_{M,2}) \cap \ker(-A_{\mathrm{M}} - R_{\mathrm{M},1}) = \{0\}.$$

We know that for any $\varphi_i$ in $V_h$

$$\begin{aligned} a(\varphi_i,\varphi_i) &+ \partial_u r(\gamma;u_\gamma)(\varphi_i,\varphi_i) \\ &= a(\varphi_i,\varphi_i) + 3\gamma \int_\Sigma \max\{\gamma(\psi - \tau(u_{\gamma,h})),0\}^2 \tau(\varphi_i)^2 \,\mathrm{d}s \\ &\geq c\|\varphi_i\|_V. \end{aligned}$$

Hence, $A_{\mathrm{M}} + R_{\mathrm{M},1}$ is positive definite, i. e. $-A_{\mathrm{M}} - R_{\mathrm{M},1}$ is negative definite. The kernels of a positive and a negative definite matrix have only the zero in common. Under consideration of these properties, the system fulfills the requirements of Theorem 4.3. Therefore, the block matrix is invertible and the saddle point problem (27) has a unique solution.

□ □

Assumption 4.1.1 guarantees that the first block in system (27) has the properties that are needed for the proof. However, we just need that there is a (locally) unique solution to the nonlinear system (25). Hence we could simply presume that the Newton matrix is regular.

An immediate consequence of Theorem 4.2 is that the implicit function theorem is applicable. This gives the following Corollaries.

**Corollary 4.4.** *Under Assumption 1, the mapping $\mathbb{R} \to V_h^0 \times V_h \times V_h^0 \times V \times L^2(\Omega) \times V_h$, which projects the regularization parameter to the solution of* (17)*, i. e. $\gamma \mapsto \xi_{\gamma,h}$ is continuously differentiable with respect to $\gamma$. This $\gamma$-derivative can be obtained by solving* (27)*, where the right hand side $(\vec{z}_1,\vec{z}_2) \in \mathbb{R}^N \times \mathbb{R}^N$ is given by*

$$\begin{aligned} (\vec{z}_1)_i &= -\partial_\gamma r(\gamma;u_{\gamma,h})(\varphi_{h,i}), \\ (\vec{z}_2)_i &= -\partial^2_{\gamma,u} r_\gamma(\gamma;u_{\gamma,h})(p_{\gamma,h},\,\varphi_{h,i}). \end{aligned}$$

If we consider all variables, we get the MPEC-version.

**Corollary 4.5.** *Let $(q_{\gamma,h},u_{\gamma,h})$ with corresponding multipliers $(p_{\gamma,h},\lambda_{\gamma,h},\theta_{\gamma,h})$ satisfy the optimality system* (17)*. If Assumption 4.1.1 is fulfilled, then the solution tuple $(q_{\gamma,h},u_{\gamma,h},p_{\gamma,h},\lambda_{\gamma,h},\theta_{\gamma,h})$ is locally unique. Furthermore, the solution variables are differentiable with respect to the penalty parameter $\gamma$. The derivatives $(\dot q_{\gamma,h},\dot u_{\gamma,h},\dot p_{\gamma,h},\dot\lambda_{\gamma,h},\dot\theta_{\gamma,h})$ solve the following system of linearized equations:*

$$A\dot u_{\gamma,h} = \dot q_{\gamma,h} + \dot\lambda_{\gamma,h}, \tag{28a}$$

$$\dot\lambda_{\gamma,h} + \partial_\gamma r(\gamma;u_{\gamma,h}) + \partial_u r(\gamma;u_{\gamma,h})(\dot u_{\gamma,h}) = 0, \tag{28b}$$

$$A^*\dot p_{\gamma,h} = j''(u_{\gamma,h})\dot u_{\gamma,h} - \dot\mu_{\gamma,h}, \tag{28c}$$

$$\dot p_{\gamma,h} + g''(q_{\gamma,h})\dot q_{\gamma,h} = 0, \tag{28d}$$

$$\dot p_{\gamma,h} - \dot\theta_{\gamma,h} = 0, \tag{28e}$$

$$\dot\mu_{\gamma,h} - \partial_u r(\gamma;u_{\gamma,h})(\dot\theta_{\gamma,h}) - \partial^2_{u,\gamma} r(\gamma;u_{\gamma,h})(\theta_{\gamma,h}) - \partial^2_{u,u} r(\gamma;u_{\gamma,h})(\dot u_{\gamma,h},p_{\gamma,h}) = 0. \tag{28f}$$

Following the same reasoning, we get a second derivative with respect to $\gamma$.

**Lemma 4.6.** *If Assumption 4.1.1 is fulfilled, the second derivative is the solution of*

$$\begin{aligned} A\ddot u_{\gamma,h} - Q'(p_{\gamma,h})\ddot p_{\gamma,h} + \partial_u r(\gamma;u_{\gamma,h})\ddot u_{\gamma,h} &= z_1 \\ A^*\ddot p_{\gamma,h} - j''(u_{\gamma,h})\ddot u_{\gamma,h} + \partial_u^2 r(\gamma;u_{\gamma,h})(p_{\gamma,h},\ddot u_{\gamma,h}) + \partial_u r(\gamma;u_{\gamma,h})\ddot p_{\gamma,h} &= z_2 \end{aligned}$$

*with the right hand side* $(z_1,z_2)$ *in* $V_h^0\times V_h^0$, *where*

$$\begin{aligned}
z_1 &= Q''(p_{\gamma,h})(\dot p_{\gamma,h},\dot p_{\gamma,h}) - \partial^2_{u,u} r(\gamma;u_{\gamma,h})(\dot u_{\gamma,h},\dot u_{\gamma,h},\cdot)\\
&\quad - 2\partial^2_{\gamma,u} r(\gamma;u_{\gamma,h})(\dot u_{\gamma,h},\cdot) - \partial^2_{\gamma,\gamma} r(\gamma;u_{\gamma,h})(\cdot),\\
z_2 &= j^{(3)}(u_{\gamma,h})(\dot u_{\gamma,h},\dot u_{\gamma,h},\cdot) - \partial^3_{u,u,u} r(\gamma;u_{\gamma,h})(\dot u_{\gamma,h},p_{\gamma,h},\dot u_{\gamma,h},\cdot)\\
&\quad - 2\partial^2_{u,u} r(\gamma;u_{\gamma,h})(\dot p_{\gamma,h},\dot u_{\gamma,h},\cdot) - 3\partial^2_{\gamma,u} r(\gamma;u_{\gamma,h})(\dot p_{\gamma,h},\cdot)\\
&\quad - \partial^3_{\gamma,\gamma,u} r(\gamma;u_{\gamma,h})(p_{\gamma,h},\cdot).
\end{aligned}$$

4.3. **Estimation.** Based on the error identity in Theorem 4.1, we introduce in this subsection estimates up to terms of higher order. We aim for estimators, which have a distinct role such that

$$J(\bar q,\bar u) - J(\tilde q_{\gamma,h},\tilde u_{\gamma,h}) \approx \eta_\gamma + \eta_h + \eta_{\text{it}}$$

and each estimator $\eta_*$ corresponds respectively to the regularization, discretization, or numerical error.

We begin to analyze the errors with the last two terms of (20). The error $\mathscr{R}_{\text{reg}}$ introduced by the DWR method is at least of higher order. In case of a tracking type goal functional, it actually equals zero. For this reason $\mathscr{R}_{\text{reg}}$ is dropped. The estimator $\eta_{\text{it}}$ for the numerical error vanishes, if $\tilde{\xi}_{\gamma,h}\to\xi_{\gamma,h}$, since under consideration of (17a) the residuum

$$\eta_{\text{it}} := \tilde\rho(\tilde p_{\gamma,h}) = -a(\tilde u_{\gamma,h},\tilde p_{\gamma,h}) + (\tilde q_{\gamma,h},\tilde p_{\gamma,h}) - r(\gamma;\tilde u_{\gamma,h})(\tilde p_{\gamma,h})$$

has the limit

$$\rho(p_{\gamma,h}) = -a(u_{\gamma,h},p_{\gamma,h}) + (q_{\gamma,h},p_{\gamma,h}) - r(\gamma;u_{\gamma,h})(p_{\gamma,h}) = 0.$$

The remaining terms of $J(\bar q,\bar u) - J(\tilde q_{\gamma,h},\tilde u_{\gamma,h})$, which consist of

$$\begin{aligned}
&\text{the regularization error} && \left(\frac12(\bar\mu+\tilde\mu_{\gamma,h}),\ \tau(\bar u)-\tau(\tilde u_{\gamma,h})\right) + \left(\bar\lambda-\tilde\lambda_{\gamma,h},\ \frac12(\bar\theta+\tilde\theta_{\gamma,h})\right)\\
&\text{and the discretization error} && \frac12\Big[\tilde\rho(\bar p-\tilde p_{\gamma,h}) + \tilde\rho^q(\bar q-\tilde q_{\gamma,h}) + \tilde\rho^*(\bar u-\tilde u_{\gamma,h})\Big],
\end{aligned}$$

depend on the unknown exact solution. The crucial part of the regularization error is the accurate determination of the difference. Therefore, we approximate the average of the analytical and discrete variables by the discrete ones.We know that the mapping $\gamma\mapsto\xi_\gamma$ is twice continuously differentiable under Assumption 4.1.1, because of Theorem 4.4 and Lemma 4.6. In the next step, we perform the well posed Taylor expansion for the variables $u$ and $\lambda$, which leads to

$$\tilde u_{\infty,h} = \tilde u_{\gamma,h} + \gamma\dot{\tilde u}_{\gamma,h} + o(\gamma^{-2}), \tag{29a}$$

$$\tilde\lambda_{\infty,h} = \tilde\lambda_{\gamma,h} + \gamma\dot{\tilde\lambda}_{\gamma,h} + o(\gamma^{-2}). \tag{29b}$$

Hence, $\tilde u_{\infty,h}$ and $\tilde\lambda_{\infty,h}$ are better approximations of $\bar u$ and $\bar\lambda$ than $\tilde u_{\gamma,h}$ and $\tilde\lambda_{\gamma,h}$. We substitute $\bar\xi$ with $\tilde\xi_{\infty,h}$ in the regularization error so that

$$\begin{aligned}
&\left\langle\frac12(\bar\mu+\tilde\mu_{\gamma,h}),\tau(\bar u)-\tau(\tilde u_{\gamma,h})\right\rangle + \left\langle\bar\lambda-\tilde\lambda_{\gamma,h},\frac12(\bar\theta+\tilde\theta_{\gamma,h})\right\rangle\\
&\approx \left\langle\tilde\mu_{\gamma,h},\tau(\bar u)-\tau(\tilde u_{\gamma,h})\right\rangle + \left\langle\bar\lambda-\tilde\lambda_{\gamma,h},\tilde\theta_{\gamma,h}\right\rangle\\
&\approx \gamma\Big(\left\langle\tilde\mu_{\gamma,h},\tau(\dot{\tilde u}_{\gamma,h})\right\rangle + \left\langle\dot{\tilde\lambda}_{\gamma,h},\tilde\theta_{\gamma,h}\right\rangle\Big) =: \circledast\ .
\end{aligned}$$

Afterwards, we exploit the identities given by (17f) and (28b).

$$
\begin{aligned}
\circledast &= \gamma\left[\partial_u r(\gamma;\tilde{u}_{\gamma,h})(\dot{\tilde{u}}_{\gamma,h},\tilde{p}_{\gamma,h}) - \partial_\gamma r(\gamma;\tilde{u}_{\gamma,h})(\tilde{p}_{\gamma,h}) - \partial_u r(\gamma;\tilde{u}_{\gamma,h})(\dot{\tilde{u}}_{\gamma,h},\tilde{p}_{\gamma,h})\right]\\
&= -\gamma\partial_\gamma\left(\int_\Sigma \max\{\gamma(-G(\tilde{u}_{\gamma,h})),0\}^3\,\tau(\tilde{p}_{\gamma,h})\,\mathrm{d}s\right)\\
&= -3\gamma\left(\int_\Sigma \max\{\gamma(-G(\tilde{u}_{\gamma,h})),0\}^2\,\tau(\tilde{p}_{\gamma,h})\left(-G(\tilde{u}_{\gamma,h})\right)\,\mathrm{d}s\right)\\
&= -3\left(\int_\Sigma \max\{\gamma(-G(\tilde{u}_{\gamma,h})),0\}^3\,\tau(\tilde{p}_{\gamma,h})\,\mathrm{d}s\right)\\
&= 3\left(\tilde{\lambda}_{\gamma,h},\ \tilde{p}_{\gamma,h}\right) =: \eta_\gamma.
\end{aligned}
$$

Thus, we arrive at the very compact formulation of the regularization error estimator. Note, that we do not evaluate the $\gamma$-derivative for this error estimator, but need its existence for the derivation. Furthermore, we deduced an error estimator that would be a natural choice, too. The violation of the complementarity condition in the continuous optimality system is measured, cf.(15f).

In the next step, we want to eliminate the exact solution variables in the discretization error terms, i. e. in

$$
\frac{1}{2}\Big[\tilde{\rho}(\bar{q}-\tilde{p}_{\gamma,h}) + \tilde{\rho}^q(\bar{q}-\tilde{q}_{\gamma,h}) + \tilde{\rho}^*(\bar{u}-\tilde{u}_{\gamma,h})\Big].
$$

As each summand vanishes for any $\varphi_h$ in $V_h$, we cannot apply the same approximation as described above in (29a), but need a further improved approximation. For this task, we consider the interpolation operator $I := I^{(2)}_{2h} : V_h \to V^{(2)}_{2h}$ and the function space $V^{(2)}_{2h}$, which reads as follows in case of Signorini's problem

$$
V^{(2)}_{2h} := \left\{ v_h \in C(\bar{\Omega},\Gamma_D)^d \mid \forall\, \mathscr{T} \in \mathbb{T}_{2h} : v_{h|\mathscr{T}} \in \mathscr{Q}_2(\mathscr{T};\mathbb{R}^d) \right\},
$$

which contains the interpolating functions. Thus, $\tilde{u}_{\infty,h} \in V_h$ is replaced by its interpolant $I(\tilde{u}_{\infty,h})$. Likewise, are $\tilde{q}_{\infty,h} := \tilde{q}_{\gamma,h} + \gamma\dot{\tilde{q}}_{\gamma,h}$ and $\tilde{p}_{\infty,h} := \tilde{p}_{\gamma,h} + \gamma\dot{\tilde{p}}_{\gamma,h}$. This standard substitution is substantiated by numerical experience, e. g. [10] and [28]; as the theoretical aspect is still an open question for adaptively refined meshes or non-smooth solutions. However, it is well known that this approximation is of higher order for smooth solutions on uniformly refined meshes, see [1, Ch. 5.2.ii]. The final discretization error estimator reads as

$$
\eta_h := \frac{1}{2}\Big[\tilde{\rho}\left(I(\tilde{p}_{\infty,h}) - \tilde{p}_{\gamma,h}\right) + \tilde{\rho}^q\left(I(\tilde{q}_{\infty,h}) - \tilde{q}_{\gamma,h}\right) + \tilde{\rho}^*\left(I(\tilde{u}_{\infty,h}) - \tilde{u}_{\gamma,h}\right)\Big].
$$

Furthermore, if the wanted control $q_D$ is indeed an element of $V_h$, s.t. $\tilde{\rho}^q = 0$, we can actually describe $J(\bar{q},\bar{u}) - J(\tilde{q}_{\gamma,h},\tilde{u}_{\gamma,h})$ depending only on $u$ and $p$.

## 5. NUMERICAL RESULTS

In this section we substantiate the theoretical findings by numerical experiments. First, we introduce briefly the canonical adaptive algorithm. Afterwards two examples based on the Signorini's problem in two dimensions and one three-dimensional example are presented.

5.1. **Adaptive algorithm.** The different estimators have distinct roles, which can be utilized in an $h,\gamma$-adaptive algorithm. The goal is an error balancing of the regularization error and the discretization error. Therefore, the numerical error should be less than either of them by a safety factor of at least $\varkappa = 0.01$. Furthermore, we demand of the underlying globalized Newton's method that a damping factor equal to one is achieved. This implies that the current approximation is in or at least near the neighborhood of quadratic convergence. The balancing itself is realized with an equilibration factor $c_e = 5$.

**Algorithm 5.1.** *The quantity* tol *denotes the exit tolerance of the algorithm.*

(1) *Perform few Newton steps and get* $(\tilde{u}_{\gamma,h}, \tilde{p}_{\gamma,h})$.
(2) *Solve the same system with another right hand side to get* $(\dot{\tilde{u}}_{\gamma,h}, \dot{\tilde{p}}_{\gamma,h})$.
(3) *Evaluate* $\eta_h$ *and* $\eta_{\text{it}}$.
(4) *If* $\eta_{\text{it}} > \varkappa \eta_h$, *go to step 1.*
(5) *Evaluate* $\eta_\gamma$.
(6) *If* $|\eta| \leq \text{tol}$, *then stop.*
(7) *If* $|\eta_h| > c_e |\eta_\gamma|$, *use an* $h-$*adaptive refinement strategy. Go to step 1.*
(8) *If* $|\eta_\gamma| > c_e |\eta_h|$, *increase the penalization parameter* $\gamma$ *by an constant factor. Go to step 1.*
(9) *Else (if* $c_e^{-1} \leq \left|\eta_h/\eta_\gamma\right| \leq c_e$*), perform alternately either step 7 or step 8.*

In step 9 of the Algorithm 5.1, we chose to perform alternately either a local refinement or an amplification of the penalization. This variation is more robust than the performance of both actions at the same time, yet the path following algorithm is slowed down.

The local refinement is based on the filtering techniques presented in [10] and the optimal mesh strategy developed in [29].

5.2. **Optimal control of Signorini's problem in two and three dimensions.** Although we derived our estimator and the adaptive algorithm for general optimal control problems restricted by different contact problems, we only focus on one specific type to prevent a cluttered section. We consider optimal control problems with tracking type goal functional, e. g.

$$\begin{aligned} \min_{q\in V^0} \quad & J(q,u) = \frac{1}{2}\|u-u_D\|^2 + \frac{\alpha}{2}\|q-q_D\|^2 \\ \text{s.t.} \quad & a(u,v-u) \geq (q,\ v-u) \quad \forall v \in V, \end{aligned}$$

where the side condition is given by the Signorini's problem in two or three dimensions. The Tikhonov parameter is set to $\alpha = 10^{-6}$.

The tracking type goal functional meets many of the theoretical requirements; for example additive splitting of in the goal functional (9) and the convexity of its control part (10). Furthermore, the higher order remainder of the error identity (20) vanishes as mentioned previously in Section 4.1. Note, that all integrations are performed with a two point Gaussian formula, which might not be exact for the wanted states $q_D$ or initial gap functions $\psi$. However, numerical tests proved that the quadrature error is negligible for the momentary study, hence we use this variant reducing the computational costs.

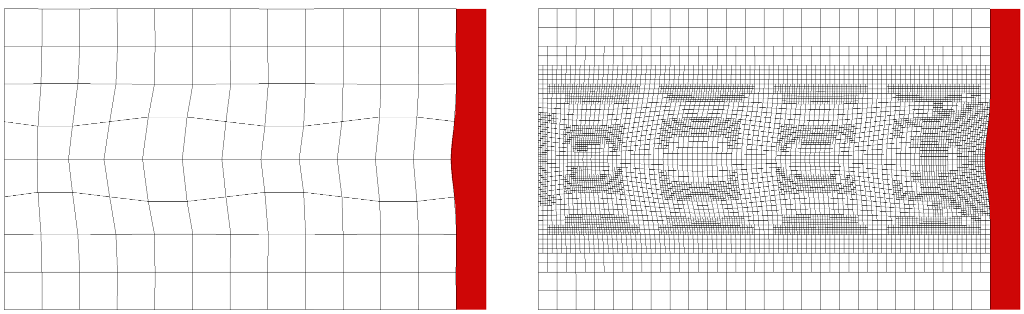
*Signorini with known solution.*
We start with a known solution, cf. [26], $(\bar{u},\bar{q})$ on the rectangular domain $\Omega = (-3,0)\times(-1,1)$. The left boundary $\Gamma_D = \{-3\}\times(-1,1)$ is fixed and contact occurs only on the right boundary $\Gamma_C = \{0\}\times(-1,1)$.

We presume homogeneous Neumann data on $\Gamma_N = \partial\Omega \setminus \Gamma_D \cup \Gamma_C$. The exact state $\bar{u} = (u_1, u_2)^\top$ is given by

$$u_1(x,y) := \begin{cases} -(x+3)^2 \left(y - \frac{x^2}{18} - \frac{1}{2}\right)^4 \left(y + \frac{x^2}{18} + \frac{1}{2}\right)^4, & \text{if } |y| < \frac{x^2}{18} + \frac{1}{2}, \\ 0, & \text{else}, \end{cases}$$

$$u_2(x,y) := \begin{cases} \frac{27}{\pi} \sin\left(\frac{4\pi(x+3)}{3}\right) \left[(y-\frac{1}{2})^3(y+\frac{1}{2})^4 + (y-\frac{1}{2})^4(y+\frac{1}{2})^3\right], & \text{if } |y| < \frac{1}{2}, \\ 0, & \text{else} \end{cases}$$

and the exact control is the volume force $\bar{q} = -\operatorname{div}(\sigma(\bar{u}))$ with a plain strain model for the two dimensional stress $\sigma$. The material parameters are $E = 10.0$ and $\nu = 0.3$. The deformation on the contact boundary $\Gamma_C$ defines the initial gap function $\psi(y) := u(0,y) \times (1,0)^\top = u_1(0,y)$. We aim to reproduce the analytic solution and hence choose $u_D = \bar{u}$ and $q_D = \bar{q}$. The limit $\bar{p}$ of the adjoint states $\tilde{p}_{\gamma,h}$ equals consequently zero.



(a) $l = 0, \gamma = 10^2$ (b) $l = 7, \gamma = 10^3$

FIGURE 1. First example: deformation and adaptive refinement with the obstacle

Figure 1 shows the deformation on a coarse mesh and the adaptive refinement. We see that the refinement is mostly done in the interior and not at the contact zone. At first glance this seems to contradict the usual behavior of contact problems, where refinement occurs mostly at the transition area between contact and non-contact. However, since we exert a force $q_D$, which shapes the domain into the right form, the influence of the contact conditions is lessened. The wanted state $u_D$ does not intersect with the obstacle and so does the discrete state $u_h$, except for numerical reasons.

The $L^2$-error $E = J(\bar{q},\bar{u}) - J(\tilde{q}_{\gamma,h}, \tilde{u}_{\gamma,h})$ and its estimator $\eta$ is shown in Tab. 3. They decrease steadily in each iteration. After the second iteration both values are close to each other, which is depicted by an efficiency index $I_{\text{eff}} := E/\eta$ of nearly one. The respective parts of the estimators are also presented in Tab. 3 and plotted in Fig. 2. The iterations, in which the numerical error is reduced, are skipped for a clearer display. The figure is also clipped to a lower bound of $\frac{1}{8} \cdot 10^{-16}$ for a better visualization and hence some marks of the numerical error estimator are cut off. The expected balancing of the discretization and regularization error estimates is observed.

In Figure 3 the development of the regularization parameter $\gamma$ and the mesh size is presented. The error splitting in (20) and the corresponding assignment of the parts to the regularization or the discretization error is reflected by the behavior of the estimators. For example, the regularization parameter $\gamma$ is increased between iteration 5 and 6. That leads to a decrease of the estimated regularization error, whereas the estimated discretization error is almost constant. On the other hand, when the mesh is refined, like from iteration 6 to 7, the discretization error estimate decreases clearly, while the regularization error estimate almost stagnates.

We observe further that the interaction of regularization and discretization actually happens. The estimated regularization error decreases in the first iterations, although the parameter $\gamma$ remains the same. This

TABLE 3. First example: error, estimators and efficiency index

| Iter | $\eta_h$ | $\eta_{\text{it}}$ | $\eta_\gamma$ | $\eta$ | $E$ | $I_{\text{eff}}$ |
|---|---|---|---|---|---|---|
| 0 | $-1.44{\cdot}10^{-5}$ | $-1.59{\cdot}10^{-13}$ | $-2.28{\cdot}10^{-7}$ | $-1.46{\cdot}10^{-5}$ | $-1.37{\cdot}10^{-4}$ | 9.35 |
| 1 | $-9.55{\cdot}10^{-6}$ | $3.51{\cdot}10^{-21}$ | $-1.60{\cdot}10^{-8}$ | $-9.57{\cdot}10^{-6}$ | $-7.17{\cdot}10^{-5}$ | 7.49 |
| 2 | $-2.82{\cdot}10^{-6}$ | $-1.52{\cdot}10^{-10}$ | $-2.04{\cdot}10^{-8}$ | $-2.84{\cdot}10^{-6}$ | $-5.73{\cdot}10^{-6}$ | 2.02 |
| 3 | $-3.02{\cdot}10^{-7}$ | $-8.28{\cdot}10^{-20}$ | $-1.07{\cdot}10^{-8}$ | $-3.13{\cdot}10^{-7}$ | $-4.10{\cdot}10^{-7}$ | 1.31 |
| 4 | $-3.39{\cdot}10^{-8}$ | $-4.39{\cdot}10^{-11}$ | $-9.71{\cdot}10^{-9}$ | $-4.36{\cdot}10^{-8}$ | $-1.28{\cdot}10^{-7}$ | 2.94 |
| 5 | $-3.30{\cdot}10^{-8}$ | $-8.63{\cdot}10^{-19}$ | $-5.51{\cdot}10^{-8}$ | $-8.81{\cdot}10^{-8}$ | $-8.36{\cdot}10^{-8}$ | 0.95 |
| 6 | $-6.31{\cdot}10^{-9}$ | $-1.97{\cdot}10^{-11}$ | $-5.40{\cdot}10^{-8}$ | $-6.04{\cdot}10^{-8}$ | $-4.83{\cdot}10^{-8}$ | 0.80 |
| 7 | $-5.93{\cdot}10^{-9}$ | $-1.89{\cdot}10^{-11}$ | $-1.36{\cdot}10^{-8}$ | $-1.96{\cdot}10^{-8}$ | $-1.15{\cdot}10^{-8}$ | 0.59 |
| 8 | $-1.58{\cdot}10^{-9}$ | $-1.74{\cdot}10^{-12}$ | $-1.29{\cdot}10^{-8}$ | $-1.45{\cdot}10^{-8}$ | $-8.37{\cdot}10^{-9}$ | 0.58 |
| 9 | $-1.54{\cdot}10^{-9}$ | $1.46{\cdot}10^{-21}$ | $-1.63{\cdot}10^{-9}$ | $-3.16{\cdot}10^{-9}$ | $-1.96{\cdot}10^{-9}$ | 0.62 |
| 10 | $-2.63{\cdot}10^{-10}$ | $-9.74{\cdot}10^{-13}$ | $-1.63{\cdot}10^{-9}$ | $-1.89{\cdot}10^{-9}$ | $-1.00{\cdot}10^{-9}$ | 0.53 |
| 11 | $-2.63{\cdot}10^{-10}$ | $1.47{\cdot}10^{-21}$ | $-1.75{\cdot}10^{-10}$ | $-4.38{\cdot}10^{-10}$ | $-2.50{\cdot}10^{-10}$ | 0.57 |
| 12 | $-7.78{\cdot}10^{-11}$ | $-1.11{\cdot}10^{-13}$ | $-1.75{\cdot}10^{-10}$ | $-2.53{\cdot}10^{-10}$ | $-1.38{\cdot}10^{-10}$ | 0.54 |
| 13 | $-7.74{\cdot}10^{-11}$ | $-4.13{\cdot}10^{-13}$ | $-2.52{\cdot}10^{-11}$ | $-1.03{\cdot}10^{-10}$ | $-5.85{\cdot}10^{-11}$ | 0.57 |
| 14 | $-1.24{\cdot}10^{-11}$ | $-7.17{\cdot}10^{-14}$ | $-1.83{\cdot}10^{-11}$ | $-3.08{\cdot}10^{-11}$ | $-1.62{\cdot}10^{-11}$ | 0.53 |

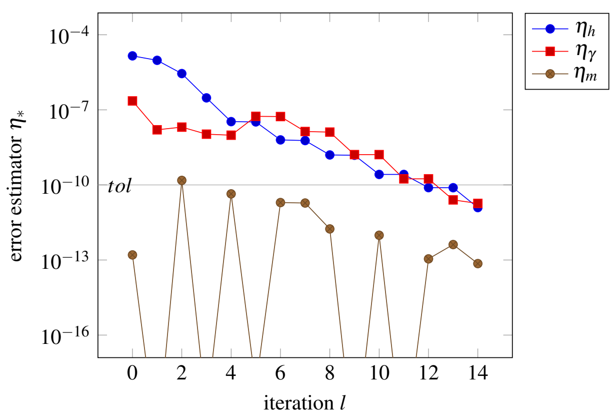



FIGURE 2. First example: absolute value of the estimator in the $l$-th iteration

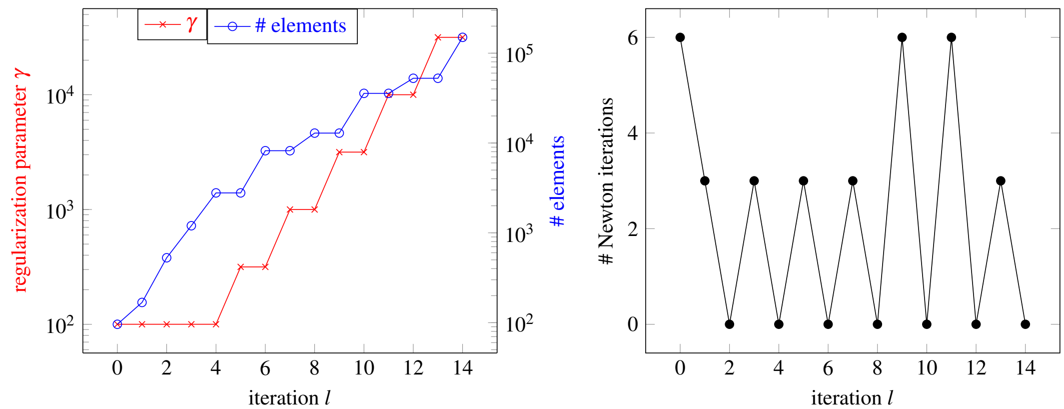
interaction is a direct result of the higher order dependency of between the regularization and discretization errors, cf. (29a).

Taking Figure 2 into account, we see another expected effect. There is a correlation between regularization parameter and the difficulty of the nonlinear problem, measured by the number of necessary Newton iterations during a global iteration. By increasing the penalty parameter the condition number worsens leading to more Newton steps (and even more damping steps), while after a mesh-refinement the non-linear problem is treated like a linear one. The solution is computed by only one Newton step.

Another investigation is addressed to the assumption (24). We evaluated the term

$$6\gamma^3 Q_E\left(\max(-G(u_{\gamma,h})),0)\alpha\tau(q_D-q_{\gamma,h})\tau(v_h)^2\right)-\|v_h\|^2_{L^2(T)}=:a(v_h). \tag{30}$$

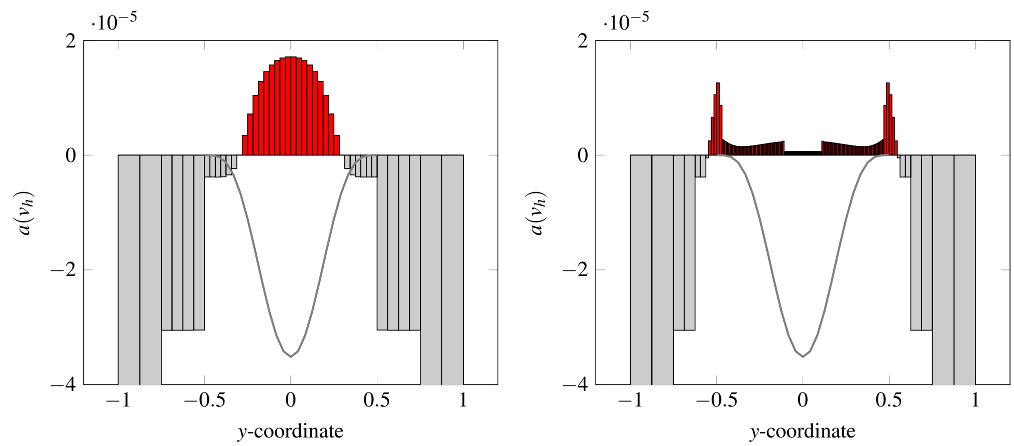
for each standard base function $v_h$ on every face $E\subset\Gamma_C$. Obviously, the respective cell $T$ is uniquely determined as the cell containing the face $E$. The normal trace of such a function is either zero or equals either of both standard linear 1D-base functions. In Figure 4 we plotted the maximum value of $a(v_h)$ for




(a) regularization parameter $\gamma$ and mesh size (b) Newton iterations

FIGURE 3. First example: development of simulation parameters and number of Newton iterations per global iteration

each edge $E \subset \Gamma_C$. If and only if the maximum is negative on each and every face, assumption (4.1.1) holds true. Hence, we see clearly a violation in the contact area, yet the adaptive algorithm operates as wanted.




(a) $l = 5$, $\gamma = 10^{5/2}$ (b) $l = 13$, $\gamma = 10^{9/2}$

FIGURE 4. First example: evaluation of (30) corresponding to the assumption (4.1.1). Additionally, the initial gap function is scaled and plotted for an easier comparison to Fig. 1.

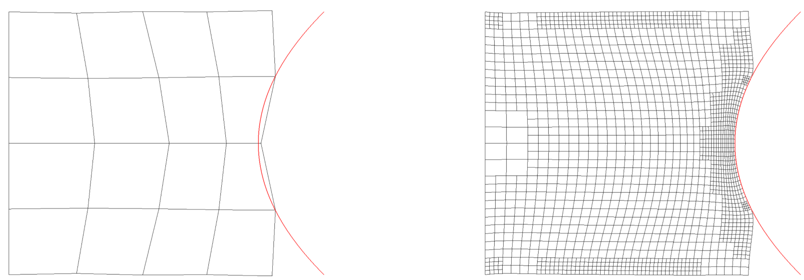
*Signorini without known solution.*
In contrast to the previous example, the analytical solution is in the following example unknown. Here, we aim for a state of the form

$$u_D(x) = \begin{pmatrix} -0.1\sin(\pi x_1)\sin(\pi x_2) \\ 0 \end{pmatrix}$$

on the domain $\Omega = [0,1]^2$, which is fixed on the left edge $\Gamma_D = \{0\} \times [0,1]$. Whereas, the contact boundary is the right edge $\Gamma_C = \{1\} \times [0,1]$. The initial gap function $g : \Gamma_C \to \mathbb{R}^2$ reads

$$g(x) = -0.05 + (x_2 - 0.5)^2.$$

The corresponding obstacle prevents the state from achieving its wanted form. Therefore, the goal functional does not vanish and we get a more interesting optimal control problem. In addition, this leads to an intersection of the wanted state and the obstacle and a distinctive refinement at the contact zone, especially at the transition of contact to non-contact. Figure 5 shows the deformation and refinement.



(a) $l = 0$, $\gamma = 100$ (b) $l = 9$, $\gamma = 10^4$

FIGURE 5. Second example: refinement on the deformed mesh and the obstacle

The adaptive simulation terminated, because the new mesh size exceeded a size limit of 200,000 cells. In Figure 6, we see as expected a similar trend for the errors as in Fig. 2. The equilibrium behaves congruently to the development of the mesh and the regularization, which is shown in Fig. 7. The Table 4 corresponds to the total estimator.

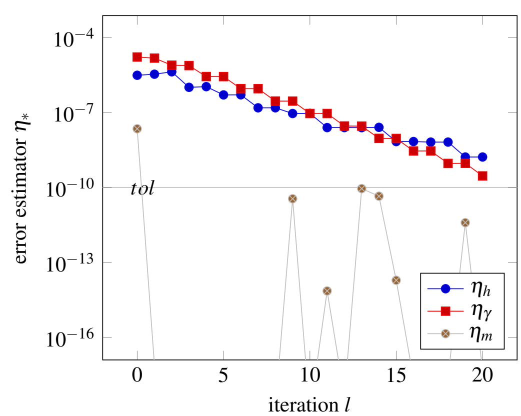
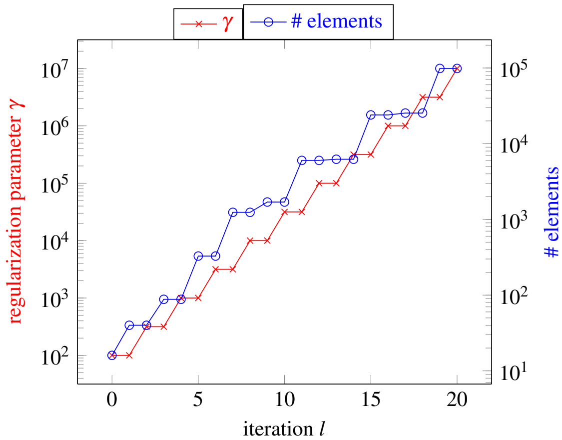
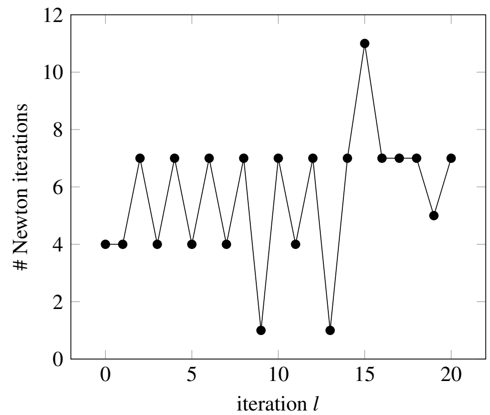



FIGURE 6. Second example: absolute value of the estimator in the $l$-th iteration




(a) regularization parameter $\gamma$ and mesh size




(b) Newton iterations

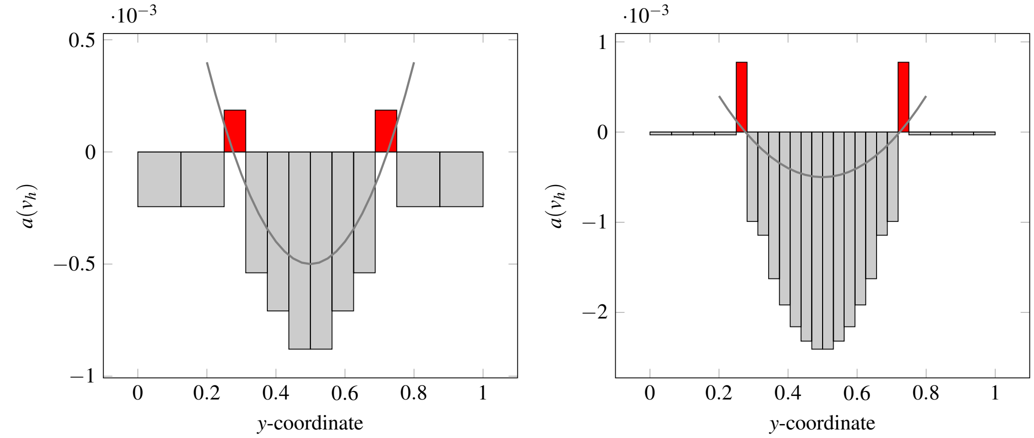
FIGURE 7. Second example: development of simulation parameters and number of Newton iterations per global iteration

An approximation of the analytical goal value $J_{\text{ref}} = 2.7000744835864975 \cdot 10^{-5}$ was calculated on a uniformly refined mesh with 262,144 cells. Hence, we get $E = J_{\text{ref}} - J(\tilde{q}_{\gamma,h}, \tilde{u}_{\gamma,h})$ for an approximated error and an approximated efficiency index $I_{\text{eff}} = E/\eta$. Table 4 shows in almost every step an index close to one. The outliers in iteration 10 and 14 are results of switching signs in the column describing the total error estimator $\eta$.

TABLE 4. Second example: error, estimators and efficiency index

| Iter | $\eta_h$ | $\eta_{\text{it}}$ | $\eta_\gamma$ | $\eta$ | $E$ | $I_{\text{eff}}$ |
|---|---|---|---|---|---|---|
| 0 | $-1.44\cdot10^{-5}$ | $-1.59\cdot10^{-13}$ | $-2.28\cdot10^{-7}$ | $-1.46\cdot10^{-5}$ | $-1.37\cdot10^{-4}$ | 9.35 |
| 1 | $-9.55\cdot10^{-6}$ | $3.51\cdot10^{-21}$ | $-1.60\cdot10^{-8}$ | $-9.57\cdot10^{-6}$ | $-7.17\cdot10^{-5}$ | 7.49 |
| 2 | $-2.82\cdot10^{-6}$ | $-1.52\cdot10^{-10}$ | $-2.04\cdot10^{-8}$ | $-2.84\cdot10^{-6}$ | $-5.73\cdot10^{-6}$ | 2.02 |
| 3 | $-3.02\cdot10^{-7}$ | $-8.28\cdot10^{-20}$ | $-1.07\cdot10^{-8}$ | $-3.13\cdot10^{-7}$ | $-4.10\cdot10^{-7}$ | 1.31 |
| 4 | $-3.39\cdot10^{-8}$ | $-4.39\cdot10^{-11}$ | $-9.71\cdot10^{-9}$ | $-4.36\cdot10^{-8}$ | $-1.28\cdot10^{-7}$ | 2.94 |
| 5 | $-3.30\cdot10^{-8}$ | $-8.63\cdot10^{-19}$ | $-5.51\cdot10^{-8}$ | $-8.81\cdot10^{-8}$ | $-8.36\cdot10^{-8}$ | 0.95 |
| 6 | $-6.31\cdot10^{-9}$ | $-1.97\cdot10^{-11}$ | $-5.40\cdot10^{-8}$ | $-6.04\cdot10^{-8}$ | $-4.83\cdot10^{-8}$ | 0.80 |
| 7 | $-5.93\cdot10^{-9}$ | $-1.89\cdot10^{-11}$ | $-1.36\cdot10^{-8}$ | $-1.96\cdot10^{-8}$ | $-1.15\cdot10^{-8}$ | 0.59 |
| 8 | $-1.58\cdot10^{-9}$ | $-1.74\cdot10^{-12}$ | $-1.29\cdot10^{-8}$ | $-1.45\cdot10^{-8}$ | $-8.37\cdot10^{-9}$ | 0.58 |
| 9 | $-1.54\cdot10^{-9}$ | $1.46\cdot10^{-21}$ | $-1.63\cdot10^{-9}$ | $-3.16\cdot10^{-9}$ | $-1.96\cdot10^{-9}$ | 0.62 |
| 10 | $-2.63\cdot10^{-10}$ | $-9.74\cdot10^{-13}$ | $-1.63\cdot10^{-9}$ | $-1.89\cdot10^{-9}$ | $-1.00\cdot10^{-9}$ | 0.53 |
| 11 | $-2.63\cdot10^{-10}$ | $1.47\cdot10^{-21}$ | $-1.75\cdot10^{-10}$ | $-4.38\cdot10^{-10}$ | $-2.50\cdot10^{-10}$ | 0.57 |
| 12 | $-7.78\cdot10^{-11}$ | $-1.11\cdot10^{-13}$ | $-1.75\cdot10^{-10}$ | $-2.53\cdot10^{-10}$ | $-1.38\cdot10^{-10}$ | 0.54 |
| 13 | $-7.74\cdot10^{-11}$ | $-4.13\cdot10^{-13}$ | $-2.52\cdot10^{-11}$ | $-1.03\cdot10^{-10}$ | $-5.85\cdot10^{-11}$ | 0.57 |
| 14 | $-1.24\cdot10^{-11}$ | $-7.17\cdot10^{-14}$ | $-1.83\cdot10^{-11}$ | $-3.08\cdot10^{-11}$ | $-1.62\cdot10^{-11}$ | 0.53 |

*Signorini without known solution - three dimensions.*
The three dimensional example is motivated by a rolling process on a thin plate. The domain $\Omega$ is a quadratic section with the dimensions $[-1,1]\times[-1,1]\times[0;0.25]$. We assume that the plate is somehow fixed from below, which implies homogeneous Dirichlet-data on $\Gamma_D = [-1,1]\times[-1,1]\times\{0\}$. The contact surface is on




(a) $l = 3$, $\gamma = 10^{5/2}$ (b) $l = 6$, $\gamma = 10^{7/2}$

FIGURE 8. Second example: evaluation of (30) corresponding to the assumption (4.1.1). Additionally, the initial gap function is scaled and plotted for an easier comparison to Fig. 5.

the opposite boundary $\Gamma_C = [-1,1] \times [-1,1] \times \{0.25\}$. The Neumann boundary is given by the remaining faces.

Here, we want as little change of the body as possible implying the wanted state $q_D \equiv 0$. The roll-obstacle is given by the lower half of a sphere surface, where the sphere is $B_{1.0}(0,0,1.125)$. The linearized gap function is

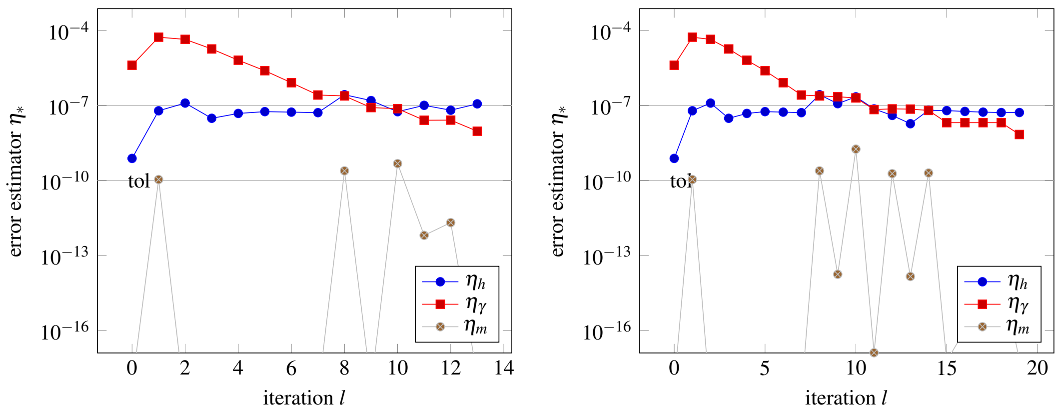
$$g(x) = \begin{cases} -\sqrt{1 - x_1^2 - x_2^2} + 0.875, & \text{if } x_1^2 + x_2^2 \leq 1.0 \\ 0.875 & \text{else.} \end{cases}$$

We observe the very same behavior in the three dimensional example as in the previous ones. However the higher order dependency is even more prominent as depicted in Fig. 9. Another striking feature of the 3D-example is that mostly regularization steps are performed, see Fig. 10. That leads to another problem, namely the worsening condition number of the Newton matrix. While the algorithm in the two dimensional examples aborted, because the error tolerance of $10^{-10}$ was achieved or the mesh size exceeded a limit which we set to $10^5$ elements; the algorithm in the present example terminated, because Newton's algorithm did not converge within over 100 steps. The simulation ended, before the discretization error could be reduced. However, the regularization error decreases steadily and a small downward trend in step 8 and 11 is observable.

We modified the algorithm 5.1 in order to resolve that problem. In the $h,\gamma$-adaptive step 9, we choose originally to perform alternately either an adaptive mesh refinement or an increase of the penalty parameter $\gamma$. Since the problems in this example originates from a $\gamma$ that became too fast too large, we change the relation of $h$- or $\gamma$-adaptive steps that are performed instead of an $h,\gamma$-adaptive one. Previously the relation was 1 : 1; now it is manually set to 3 : 1.

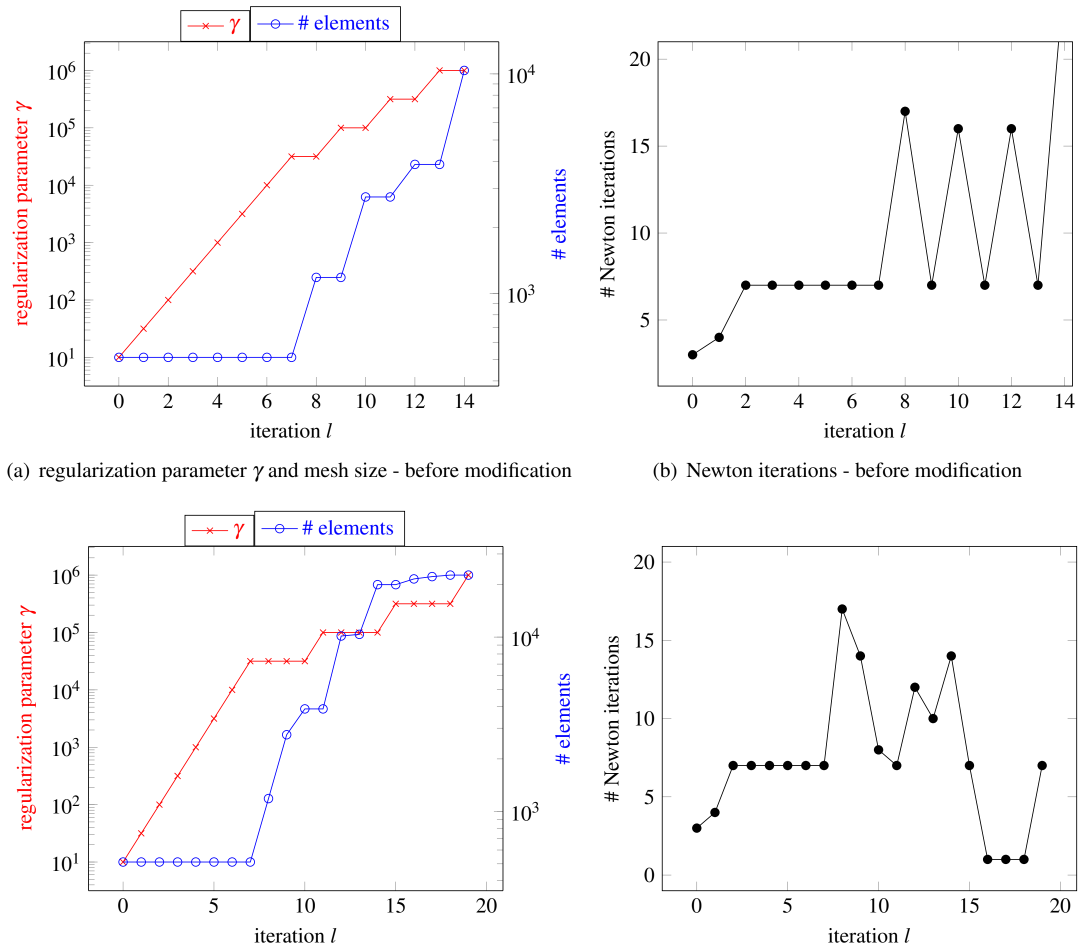
The example still proves itself difficult, because a mesh refinement does not guarantee a reduction of the discretization error. This can be seen in the global iteration step 7 (before and after modification) as well as in steps 9 and 13. Step 13 is in particular an interesting one, since only few cells are additionally refined, but the discretization error increases quite much. This behavior can be explained by the signed sum of the cell contributions. While the absolute maximum decreases always, terms might cancel each other out in the signed sum; e.g. leading to switching signs between step 7 and 8.

Note that once again Assumption (4.1.1) is not fulfilled.




(a) before modification (b) after modification

FIGURE 9. Third example: absolute value of the estimator in the $l$-th iteration




(a) regularization parameter $\gamma$ and mesh size - before modification

(b) Newton iterations - before modification

(c) regularization parameter $\gamma$ and mesh size - after modification

(d) Newton iterations - after modification

FIGURE 10. Third example: development of simulation parameters and number of Newton iterations per global iteration

## 6. Conclusions and outlook

In this paper, we proposed an a-posteriori error estimator for discrete, regularized optimal control problems governed by contact problems. The error between the solution of the original optimal control problem and the solution of the discrete, regularized problem is measured by the evaluation of the goal functional. We used the DWR method to prove an error identity, which consists of three parts. Each one is assigned to one error source; the regularization, the discretization or the numerical approximation.

The numerical results undermine that the performed error splitting as well as the respective estimators are well posed; even though a higher order dependency is noticeable. The adaptive path following algorithm provides a balancing strategy for the regularization and discretization error estimates. As expected after an adaptive mesh refinement the discretization error decreases and the regularization error almost stagnates; while after a $\gamma$-adaptive step, increasing the penalty parameter, the regularization error is reduced and the discretization error almost stagnates.

Note that for the (simplified) Signorini's problem the error estimator for the regularization error is evaluated only on the contact boundary leading to a highly efficient computation. The main costs are introduced by the discretization error estimator, while the numerical error estimator is its by-product with a negligible overhead.

Even though the balancing strategy works fine, the three dimensional example proves itself quite difficult. At first, the regularization parameter became too quickly too large such that the adaptive algorithm aborted because of non-converging Newton's method. After modifying the implementation of algorithm, the problem was shifted to the mesh size. The simulation terminated due to the memory limits. We were able to reduce the estimated error at least slightly. Hence, further considerations regarding an automatic parameter fitting of the algorithm are advisable. For instance an automatic choice of the factor, which is used to increase the penalty parameter.

Other expansions should be considered as well; in particular, the simulation of elasto-plastic contact problems. The additional nonlinearity transforms the semi-linear regularized problem into a quasi-linear one. We expect that the error estimator for the plasticity error is, analogously to the regularization estimator at hand, the dual paring of the corresponding slack and adjoint variables.

Another natural yet non-trivial extension is the inclusion of friction on the contact surface.

## APPENDIX A. PROOF OF THEOREM 3.4

As stated above, we present the complete proof of Theorem 3.4 here.

*Proof.*
*Convergence of the primal variables:*
The first part of the theorem is a collection of the above results. We have already noted that every regularized problem ($P_\gamma$) has a globally optimal solution $q_\gamma$. The sequence $\{q_\gamma\}$ of global minimizers has a weak accumulation point $\bar{q} \in V^0$, because of the boundedness of the goal functional $J$. Taking the obvious relation

$$J(q_\gamma, S_\gamma(q_\gamma)) \leq J(q, S_\gamma(q))$$

for any $q \in V^0$ and Lemma 3.3 into account, we arrive at

$$\begin{aligned} J(\bar{q},\bar{u}) = J(\bar{q},S(\bar{q})) &\leq \liminf_{\gamma\to\infty} J(q_\gamma, S_\gamma(q_\gamma)) \leq \limsup_{\gamma\to\infty} J(q_\gamma, S_\gamma(q_\gamma)) \\ &\leq \lim_{\gamma\to\infty} J(q, S_\gamma(q)) = J(q, S(q)). \end{aligned}$$

Hence, the accumulation point $\bar{q}$ is a global minimizer of (P) and with $\bar{q} = q$ we achieve the convergence of the objective. The strong convergence results from the following considerations. First of all, the convergence in the dual norm

$$\|q_\gamma - \bar{q}\|_{V'} = \sup_{v\in V\setminus\{0\}} \frac{|\langle q_\gamma - \bar{q},\ v\rangle|}{\|v\|} \quad \to \quad 0$$

implying the strong convergence of the state variable

$$\|u_\gamma - \bar{u}\|_V = \|S_\gamma(q_\gamma) - S(\bar{q})\|_V \leq L\|q_\gamma - \bar{q}\|_{V'} + \|S_\gamma(\bar{q}) - S(\bar{q})\|_V \quad \to \quad 0 \tag{31}$$

is an immediate consequence of the compact embedding $V^0 \hookrightarrow V^{-1}$, of the Lipschitz continuity of $S_\gamma$ and of Lemma 3.3. As the objective converges, we further know

$$\begin{aligned} &\Big| \big|g(q_\gamma) - g(\bar{q})\big| - \big|j(S(q_\gamma)) - j(S(\bar{q}))\big| \Big| \leq \Big|J(q_\gamma, S_\gamma(q_\gamma)) - J(\bar{q}, S\bar{q})\Big| && \to \quad 0 \\ &\lim_{\gamma\to\infty} \big|g(q_\gamma) - g(\bar{q})\big| = \lim_{\gamma\to\infty} \big|j(S(q_\gamma)) - j(S(\bar{q}))\big| \leq \lim_{\gamma\to\infty} C\|S_\gamma(q_\gamma) - S(\bar{q})\|_V && \to \quad 0, \end{aligned}$$

because of the regularity of $j$, which gives a constant $C$, and (31). The mapping $g$ is presumed to be convex, and twice continuously differentiable. The first derivative $g'(\bar{q})$ is an element of the dual space $(V^0)^*$. Therefore there exists a $t \in [0,1]$ so that together with (10)

$$\begin{aligned} g(q_\gamma) - g(\bar{q}) &= g'(\bar{q})(q_\gamma - \bar{q}) + \frac{1}{2} g''\left(\bar{q} + t(q_\gamma - \bar{q})\right)(q_\gamma - \bar{q})^2 \\ &\geq g'(\bar{q})(q_\gamma - \bar{q}) + \frac{\alpha}{2}\|q_\gamma - \bar{q}\|^2_{V^0} \quad \to \quad 0. \end{aligned}$$

This concludes the proof of the first claim.

*Convergence of the dual variables.* The convergence of the slack variables is an easy consequence of the continuity of $A : V \to V^*$.

$$\tau^*\lambda_\gamma = Au_\gamma - q_\gamma \quad \to \quad A\bar{u} - \bar{q} = \tau^*\bar{\lambda} \qquad \text{in } V^*$$

and $\lambda_\gamma \to \bar{\lambda}$ in $W'$ because of the continuity of $\tau$. As $\bar{u} = S(\bar{q})$ is the solution of (11), $\bar{\lambda}$ is the associated slack variable fulfilling the complementarity system in (15b). To prove the weak convergence of the adjoint state, insert $\theta_\gamma$ and $\mu_\gamma$ in (13c) to obtain

$$A^* p_\gamma + \partial_u r(\gamma; u_\gamma)(p_\gamma) = \nabla j(u_\gamma).$$

Testing this equation with $p_\gamma$ itself yields

$$\|p_\gamma\|_V \leq \frac{1}{\beta}\,\|\nabla j(u_\gamma)\|_{V'} \tag{32}$$

due to the coercivity of the bilinear form $a(\cdot,\ \cdot)$. Additionally, we get

$$\begin{aligned}\int_{\Gamma_C} \left[\max\left(\gamma(\tau(u_\gamma)-\psi),0\right)\tau(p_\gamma)\right]^2 \,\mathrm{d}o &\leq \frac{1}{3\gamma}\left(\nabla j(u_\gamma),\ p_\gamma\right)\\ &\leq \frac{1}{3\gamma\beta}\|\nabla j(u_\gamma)\|_{V'}^2 \ \to\ 0 \qquad \text{as } \gamma\to\infty,\end{aligned} \tag{33}$$

where we used (4) and (32). The boundedness of $\{u_\gamma\}$ in $V$ is exploited for the passage to the limit. Note, that we presume that $j$ is three times continuously differentiable. From (32) we infer the existence of a subsequence, weakly converging in $V$ to $\bar{p}$. For simplicity we denote this subsequence by $\{p_\gamma\}$, too. This proves equation (14d). Moreover, the convergence of $\mu_\gamma$ and (15c) follow from

$$\mu_\gamma = \nabla j(u_\gamma) - A^* p_\gamma \rightharpoonup \nabla j(\bar{u}) - A^*\bar{p} =: \bar{\mu} \quad \text{in } V'.$$

*Complementarity relations.* It remains to verify the complementarity relations in (15f). For this purpose, we construct an adequate family of test functions $m_\gamma$ in $V = H^1_D(\Omega)^d$. Let $m_\gamma \in V$ be the solution of the mixed boundary problem

$$\left.\begin{aligned} Am_\gamma &= q_\gamma && \text{in } \Omega,\\ m_\gamma &= 0 && \text{on } \Gamma_D,\\ \sigma_n(m_\gamma) &= 0 && \text{on } \Gamma_N,\\ m_\gamma\cdot n &= \max\{\tau(u_\gamma)-\psi,0\} && \text{on } \Gamma_C,\\ m_\gamma\cdot t_i &= 0 && \text{on } \Gamma_C.\end{aligned}\right\} \tag{$P_{\text{aux}}$}$$

Here, $n$ denotes the normal vector and $t_1$ the corresponding tangential vector for $\Omega\subset\mathbb{R}^2$. For $\Omega\subset\mathbb{R}^3$, there are two orthogonal, tangential vectors $t_{1,2}$. We recall the relation for the stress $\sigma = \mathbb{C}\varepsilon$, cf. 1. The boundary conditions on $\Gamma_C$ are equivalent to inhomogeneous Dirichlet data

$$\tau(U_{0,\gamma}) = \max\{\tau(u_\gamma)-\psi,0\} \qquad \text{on } \Gamma_C,$$

for a function $U_{0,\gamma}\in V$ due to Lemma 2.1. Note, by construction $U_{0,\gamma}\cdot t_i$ is zero. Problem ($P_{\text{aux}}$) is a standard, static, linear elastic problem with mixed boundary conditions. It is well known that this problem has a unique solution, see, e. g. [31]. Therefore, the description of $m_\gamma$ is well posed. Furthermore, the limit case for $\gamma\to\infty$ is well posed, too, because

$$\left.\begin{aligned} A\bar{m} &= \bar{q} && \text{in } \Omega\\ \bar{m} &= 0 && \text{on } \Gamma_D\cup\Gamma_C\\ \sigma_n(\bar{m}) &= 0 && \text{on } \Gamma_N,\end{aligned}\right\} \tag{$\bar{P}_{\text{aux}}$}$$

has also a unique solution $\bar{m}$ in $V$. Additionally, $m_\gamma\to\bar{m}$ in $V$. To show this, we solve the substitute problem for $v_\gamma = m_\gamma - U_{0,\gamma}$

$$\left.\begin{aligned} Av_\gamma &= q_\gamma - AU_{0,\gamma} && \text{in } \Omega,\\ v_\gamma &= 0 && \text{on } \Gamma_D\cup\Gamma_C,\\ \sigma_n(v_\gamma) &= 0 && \text{on } \Gamma_N\end{aligned}\right\} \tag{$P'_{\text{aux}}$}$$

instead of ($P_{\text{aux}}$) and get

$$\begin{aligned}\beta\|v_\gamma-\bar{m}\|_V^2 &\leq a(v_\gamma-\bar{m},v_\gamma-\bar{m}) = \left(q_\gamma-\bar{q},\ v_\gamma-\bar{m}\right) + a(-U_{0,\gamma},v_\gamma-\bar{m})\\ &\leq c\|q_\gamma-\bar{q}\|_{V^0}\,\|v_\gamma-\bar{m}\|_V + c\|U_{0,\gamma}\|_V\,\|v_\gamma-\bar{m}\|_V.\end{aligned}$$

We use again Lemma 2.1 and conclude

$$\|m_\gamma - \bar{m}\|_V \le \frac{c}{\beta}\|q_\gamma - \bar{q}\| + \left(\frac{c}{\beta}+1\right)\|U_{0,\gamma}\| \to 0 \quad \text{as } \gamma \to \infty.$$

Finally, combining the inequality (33) with

$$\begin{aligned}
&\|\max\big(\gamma(\tau(u_\gamma)-\psi),0\big)\|^4_{L^4(\Gamma_C)} \\
&\quad = -\int_{\Gamma_C} -\Big(\max\big(\gamma(\tau(u_\gamma)-\psi),0\big)^3 \cdot \gamma \max\big(\tau(u_\gamma)-\psi,0\big)\Big)\, \mathrm{d}o \\
&\quad = -\gamma\big(\lambda_\gamma,\ \tau(m_\gamma)\big)_{\Gamma_C}
\end{aligned}$$

leads to

$$\begin{aligned}
|\big(\lambda_\gamma,\ \tau(p_\gamma)\big)|^2 &= \Big|\int_{\Gamma_C} -\big[\max\big(\gamma(\tau(u_\gamma)-\psi),0\big)\big]^3 \tau(p_\gamma)\, \mathrm{d}o\Big|^2 \\
&\le \|\max\big(\gamma(\tau(u_\gamma)-\psi),0\big)\|^4_{L^4(\Gamma_C)} \|\max\big(\gamma(\tau(u_\gamma)-\psi),0\big)\tau(p_\gamma)\|^2_{L^2(\Gamma_C)} \\
&\le \gamma\big|\big(\lambda_\gamma,\ \tau(m_\gamma)\big)\big| \frac{1}{3\gamma}\big|\big(\nabla j(u_\gamma),\ p_\gamma\big)\big| \\
&= \frac{1}{3}\big|\big(\nabla j(u_\gamma),\ p_\gamma\big)\big|\big|\big(\lambda_\gamma,\ \tau(m_\gamma)\big)\big| \quad \to \quad C\big|\big(\bar{\lambda},\ \tau(\bar{m})\big)\big| = 0,
\end{aligned}$$

since $\tau(\bar{m}) = 0$. Due to (15e), the first complementarity condition in (15f) is shown.

To derive the second equation of the complementarity conditions, we observe that the definition of $\mu_\gamma$ in (13f) implies

$$\begin{aligned}
\big(\mu_\gamma,\ \tau(u_\gamma)-\psi\big) &= 3\int_{\Gamma_C} \big[\max\big(\gamma(\tau(u_\gamma)-\psi),0\big)\big]^2\, \tau(p_\gamma)\, \gamma(\tau(u_\gamma)-\psi))\, \mathrm{d}o \\
&= 3\int_{\Gamma_C} \big[\max\big(\gamma(\tau(u_\gamma)-\psi),0\big)\big]^3\, \tau(p_\gamma)\, \mathrm{d}o \\
&= -3\,\big(\lambda_\gamma,\ \tau(p_\gamma)\big) \to 0.
\end{aligned}$$

Since $\mu_\gamma \rightharpoonup \bar{\mu}$ in $W'$ and $u_\gamma \to \bar{u}$ in $V$, this gives the claim. In order to prove the sign condition in (15f), we test (13c) and (15c) each with $p_\gamma - \bar{p}$ and subtract the arising equations to obtain

$$\begin{aligned}
\big(\mu_\gamma - \bar{\mu},\ \tau(p_\gamma - \bar{p})\big) &= \big(\nabla j(u_\gamma) - \nabla j(\bar{u}),\ p_\gamma - \bar{p}\big) - a(p_\gamma - \bar{p}, p_\gamma - \bar{p}) \\
&\le \big(\nabla j(u_\gamma) - \nabla j(\bar{u}),\ p_\gamma - \bar{p}\big).
\end{aligned}$$

Employing again the definition of $\mu_\gamma$ in (13f), we find

$$\big(\mu_\gamma,\ \tau(p_\gamma)\big) = 3\int_\Omega \gamma\big[\max\big(\gamma(\tau(u_\gamma)-\psi),0\big)\big]^2 \tau(p_\gamma)^2\, dx \ge 0 \quad \forall \gamma > 0.$$

Thus we arrive at

$$\begin{aligned}
\langle\bar{\mu},\ \tau(\bar{p})\rangle &= \langle\mu_\gamma - \bar{\mu},\ \tau(p_\gamma - \bar{p})\rangle - \big(\mu_\gamma,\ \tau(p_\gamma)\big) + \big(\mu_\gamma,\ \tau(\bar{p})\big) + \langle\bar{\mu},\ \tau(p_\gamma)\rangle \\
&\le \big(\nabla j(u_\gamma) - \nabla j(\bar{u}),\ \tau(p_\gamma - \bar{p})\big) + \big(\mu_\gamma,\ \tau(\bar{p})\big) + \langle\bar{\mu},\ \tau(p_\gamma)\rangle.
\end{aligned}$$

Because of $u_\gamma \to \bar{u}$ in $V$, $p_\gamma \rightharpoonup \bar{p}$ in $V$, and $\mu_\gamma \rightharpoonup \bar{\mu}$ in $V'$, the right hand side converges to $2\langle\bar{\mu},\ \tau(\bar{p})\rangle$, which gives the desired sign condition. □

## REFERENCES


[1] W. Bangerth and R. Rannacher. *Adaptive finite element methods for differential equations*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2003.

[2] V. Barbu. *Optimal control of variational inequalities*, volume 100 of *Research Notes in Mathematics*. Pitman (Advanced Publishing Program), Boston, MA, 1984.

[3] R. Becker and R. Rannacher. A feed-back approach to error control in finite element methods: basic analysis and examples. *East-West J. Numer. Math.*, 4(4):237–264, 1996.

[4] R. Becker and R. Rannacher. An optimal control approach to a posteriori error estimation in finite element methods. *Acta Numerica*, 10:1–102, 2001.

[5] M. Benzi, G. H. Golub, and J. Liesen. Numerical solution of saddle point problems. *Acta Numer.*, 14:1–137, 2005.

[6] A. Bermúdez and C. Saguez. Optimal control of a Signorini problem. *SIAM J. Control Optim.*, 25(3):576–582, 1987.

[7] T. Betz. *Optimal control of two variational inequalities arising in solid mechanics*. PhD thesis, Technische Universität Dortmund, 2015.

[8] T. Betz, C. Meyer, A. Rademacher, and K. Rosin. Adaptive optimal control of elastoplastic contact problems. Technical report, Fakultät für Mathematik, TU Dortmund, 2014.

[9] H. Blum and F.-T. Suttmeier. Weighted error estimates for finite element solutions of variational inequalities. *Computing*, 65(2):119–134, 2000.

[10] M. Braack and A. Ern. A posteriori control of modeling errors and discretization errors. *Multiscale Model. Simul.*, 1(2):221–238, 2003.

[11] D. Braess. A posteriori error estimators for obstacle problems - another look. *Numer. Math.*, 101(3):415–421, 2005.

[12] D. Braess. *Finite Elemente*. Springer Spektrum, 5 edition, 2013.

[13] 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(3):455–471, 2007.

[14] C. Brett, C. M. Elliott, M. Hintermüller, and C. Löbhard. Mesh adaptivity in optimal control of elliptic variational inequalities with point-tracking of the state. *Interfaces Free Bound.*, 17(1):21–53, 2015.

[15] Z. Chen and R. H. Nochetto. Residual type a posteriori error estimates for elliptic obstacle problems. *Numer. Math.*, 84(4):527–548, 2000.

[16] D. A. French, S. Larsson, and R. H. Nochetto. Pointwise a posteriori error analysis for an adaptive penalty finite element method for the obstacle problem. *Comput. Methods Appl. Math.*, 1(1):18–38, 2001.

[17] A. Gaevskaya, M. Hintermüller, R. H. W. Hoppe, and C. Löbhard. Adaptive finite elements for optimally controlled elliptic variational inequalities of obstacle type. In *Optimization with PDE constraints*, volume 101 of *Lect. Notes Comput. Sci. Eng.*, pages 95–150. Springer, Cham, 2014.

[18] M. Hintermüller. Inverse coefficient problems for variational inequalities: optimality conditions and numerical realization. *M2AN Math. Model. Numer. Anal.*, 35(1):129–152, 2001.

[19] C. Johnson. Adaptive finite element methods for the obstacle problem. *Math. Models Methods Appl. Sci.*, 2(4):483–487, 1992.

[20] D. Kinderlehrer and G. Stampacchia. *An Introduction to Variational Inequalities and Their Applications*. Academic Press, New York, 1980.

[21] K. Kunisch and D. Wachsmuth. Sufficient optimality conditions and semi-smooth Newton methods for optimal control of stationary variational inequalities. *ESAIM Control Optim. Calc. Var.*, 18(2):520–547, 2012.

[22] C. Meyer, A. Rademacher, and W. Wollner. Adaptive optimal control of the obstacle problem. *Siam J. Sci. Comput.*, 37(2):A918–A945, 2015.

[23] C. Meyer and O. Thoma. A priori finite element error analysis for optimal control of the obstacle problem. *SIAM Journal on Numerical Analysis*, 51(1):605–628, 2013.

[24] F. Mignot. Contrôle dans les inéquations variatonelles elliptiques. *Journal of Functional Analysis*, 22(2):130–185, 1976.

[25] J. T. Oden and N. Kikuchi. *Contact problems in elasticity*. SIAM, 1988.

[26] A. Rademacher. NCP Function-Based Dual Weighted Residual Error Estimators for Signorini's Problem. *SIAM J. Scientific Computing*, 38(3), 2016.

[27] R. Rannacher, B. Vexler, and W. Wollner. A posteriori error estimation in PDE-constrained optimization with pointwise inequality constraints. In *Constrained optimization and optimal control for partial differential equations*, volume 160 of *Internat. Ser. Numer. Math.*, page 349–373. Birkhäuser/Springer Basel AG, Basel, 2012.

[28] R. Rannacher and J. Vihharev. Adaptive finite element analysis of nonlinear problems: balancing of discretization and iteration errors. *Journal on Numerical Mathematics*, 21(1):23–61, 2013.

[29] T. Richter. *Parallel Multigrid Method for Adaptive Finite Elements with Application to 3D Flow Problems*. PhD thesis, Ruprecht-Karls-Universität Heidelberg, (Diss.), 2005.

[30] A. Schiela and D. Wachsmuth. Convergence analysis of smoothing methods for optimal control of stationary variational inequalities. *ESAIM Math. Model. Numer. Anal.*, 47(3):771–787, 2013.

[31] B. Schweizer. *Partielle Differentialgleichungen*. Springer-Lehrbuch Masterclass. Springer Spektrum, 2013.

[32] F.-T. Suttmeier. General approach for a posteriori error estimates for finite element solutions of variational inequalities. *Comput. Mech.*, 27(4):317–323, 2001.

[33] F. Tröltzsch. *Optimal Control of Partial Differential Equations*, volume 112 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, 2010. Theory, methods and applications, Translated from the 2005 German original by Jürgen Sprekels.

[34] A. Veeser. Efficient and reliable a posteriori error estimators for elliptic obstacle problems. *SIAM J. Numer. Anal.*, 39(1):146–167, 2001.

[35] R. Verfürth. *A posteriori error estimation techniques for finite element methods*. Numerical Mathematics and Scientific Computation. Oxford University Press, Oxford, 2013.

[36] B. Vexler and W. Wollner. Adaptive finite elements for elliptic optimization problems with control constraints. *SIAM J. Control Optim.*, 47(1):509–534, 2008.

[37] G. Wachsmuth. Strong stationarity for optimal control of the obstacle problem with control constraints. *SIAM J. Optim.*, 24(4):1914–1932, 2014.

FACULTY OF MATHEMATICS, TECHNISCHE UNIVERSITÄT DORTMUND, VOGELPOTHSWEG 87, 44227 DORTMUND, GERMANY

*E-mail address*: andreas.rademacher@tu-dortmund.de

*E-mail address*: korinna.rosin@tu-dortmund.de