SHAPE OPTIMIZATION IN CONTACT PROBLEMS WITH COULOMB FRICTION AND A SOLUTION-DEPENDENT FRICTION COEFFICIENT[*]

P. BEREMLIJSKI[†], J. HASLINGER[‡], J. V. OUTRATA[§], AND R. PATHÓ[¶]

Abstract. The present paper deals with shape optimization in discretized two-dimensional (2D) contact problems with Coulomb friction, where the coefficient of friction is assumed to depend on the unknown solution. Discretization of the continuous state problem leads to a system of finite-dimensional implicit variational inequalities, parametrized by the so-called design variable, that determines the shape of the underlying domain. It is shown that if the coefficient of friction is Lipschitz and sufficiently small in the $C^{0,1}$-norm, then the discrete state problems are uniquely solvable for all admissible values of the design variable (the admissible set is assumed to be compact), and the state variables are Lipschitzian functions of the design variable. This facilitates the numerical solution of the discretized shape optimization problem by the so-called implicit programming approach. Our main results concern sensitivity analysis, which is based on the well-developed generalized differential calculus of B. Mordukhovich and generalizes some of the results obtained in this context so far. The derived subgradient information is then combined with the bundle trust method to compute several model examples, demonstrating the applicability and efficiency of the presented approach.

Key words. shape optimization, contact problems, Coulomb friction, solution-dependent coefficient of friction, mathematical programs with equilibrium constraints

AMS subject classifications. 49Q10, 74M10, 74S05

DOI. 10.1137/130948070

1. Introduction. Contact shape optimization is a branch of optimal control theory in which the control variables, called in this context the design variables, are linked to the geometry of elastic bodies that are in mutual contact. By changing their shapes only, one strives to achieve the best possible or some a priori given properties of the system. A design is evaluated by the so-called cost functional that is subject to minimization. Common examples include minimizing the normal stress along the contact surface (related to the minimization of the potential energy) or attaining a given contact stress profile; see, e.g., [1]. Physical quantities, subject to

[†]Centre of Excellence, IT4Innovations and Department of Applied Mathematics, VŠB-Technical University of Ostrava, 17. listopadu 15/2172, 708 33 Ostrava-Poruba, Czech Republic (petr.beremlijski@vsb.cz). This author was supported by the European Regional Development Fund in the Centre of Excellence project IT4Innovations (CZ.1.05/1.1.00/02.0070) and by the project SPOMECH - Creating a multidisciplinary R&D team for reliable solution of mechanical problems, reg. no. CZ.1.07/2.3.00/20.0070 within Operational Programme "Education for competitiveness" funded by Structural Funds of the European Union and state budget of the Czech Republic.
[‡]Department of Numerical Mathematics, Charles University in Prague, Sokolovská 83, 186 75 Praha 8, Czech Republic (hasling@karlin.mff.cuni.cz).
[§]Institute of Theory of Information and Automation of the AS CR, Pod Vodárenskou věží 4, 182 08 Praha 8, Czech Republic, and Graduate School of Information Technology and Mathematical Sciences, University of Ballarat, Ballarat VIC 3350, Australia (outrata@utia.cas.cz). This author was supported by the ARC project DP110102011.
[¶]Department of Numerical Mathematics, Charles University in Prague, Sokolovská 83, 186 75 Praha 8, Czech Republic (patho@karlin.mff.cuni.cz). This author was supported by GAUK no. 719912.

optimization, are related to the design variables via a state relation, usually modeled by partial differential equations or inequalities, whose complexity largely depends on the physical phenomena involved. However, a common feature of various contact shape optimization problems is their nonsmooth nature, which stems from the fact that the respective control-to-state map, the mapping assigning to a design variable the corresponding state variable, is typically nondifferentiable. This is a major source of complications for sensitivity analysis as well as for numerical realization; see, e.g., [25, 10, 13] and the references therein.

Shape optimization problems similar to the type in this paper have been considered in [3, 2] and [11]. Whereas [3] and [2] were concerned with the standard Coulomb friction model with a fixed friction coefficient, [11] investigates the Tresca model with a coefficient of friction that may depend on the unknown solution. The model from our paper seems to be even more involved; indeed, it is the model from [3] where, in addition, the friction coefficient depends also on the solution. So the implicit relationship, characteristic for the Coulomb friction, is here substantially more complicated.

Our main workhorse is the *implicit programming approach* (ImP) which proved its efficiency both in deriving optimality conditions for various problems of this type and in their numerical solutions (e.g., [19, 21, 20]). This tool has already been successfully used in [3, 2] and [11] in cooperation with a reliable bundle trust method of nonsmooth optimization from [24]. The computation of the (sub)gradient information, required by this method, can be conducted essentially in two different ways:

(i) If the underlying control-to-state map is piecewise C^1 (PC^1), it is convenient to describe it via the generalized Jacobians of Clarke [5] and to obtain a desired (sub)gradient completely within the generalized differential calculus of Clarke. This way has been applied, e.g., in [3].

(ii) If one has to work with nonpolyhedral multifunctions, and hence the PC^1 nature of the control-to-state mapping cannot be guaranteed, it seems reasonable to perform the sensitivity analysis via the generalized differential calculus of Mordukhovich [18], which is richer concerning specialized calculus rules.

Since in the considered model we are concerned with rather complicated nonsmooth and set-valued mappings, we have chosen the second approach. However, even some upper estimates from the Mordukhovich calculus can sometimes be tightened if one takes into account possible additional structural properties. This is very important, because the computed (sub)gradient should be as precise as possible. We make use of these possibilities in two cases in section 4.

The outline of the paper is as follows. We conclude this introductory section with a review of the notation to be used, and we recall some definitions from variational analysis that we will extensively use in section 4. Section 2 is devoted to the state problem, which is first formulated in its continuous, infinite-dimensional setting. The shape optimization problem is presented, as well. A discretization of the state problem then leads to a system of parametrized, implicit, algebraic variational inequalities. In section 3 we study its structural properties, including existence and uniqueness of solutions, independently of the design parameter. The discrete state problem is then reformulated as a generalized equation, and we show that it is strongly regular in the sense of Robinson. Section 4 concerns sensitivity analysis; i.e., we compute a Clarke's subgradient of the composite cost functional using techniques from variational analysis and the theory of generalized differentiation. This subgradient information

is then supplied to the chosen bundle method for the numerical solution of the shape optimization problem in section 5, where several numerical examples conclude the paper.

We use the following notation: the symbol $H^k(\Omega)$ ($k \geq 0$ integer) stands for the Sobolev space of functions which are, together with their derivatives, up to order k square integrable in Ω, i.e., elements of $L^2(\Omega)$ (we set $H^0(\Omega) := L^2(\Omega)$). The norm in $H^k(\Omega)$ will be denoted by $\|\cdot\|_{k,\Omega}$. Vector-valued functions and the respective spaces of vector-valued functions will be denoted by bold characters. Bold characters will also be used for vectors $\boldsymbol{u} = (u_1, \ldots, u_n)^\top, \boldsymbol{v} = (v_1, \ldots, v_n)^\top \in \mathbb{R}^n$, with the Euclidean scalar product $\boldsymbol{u} \cdot \boldsymbol{v} := (\boldsymbol{u}, \boldsymbol{v})_n = \sum_{i=1}^n u_i v_i$ and corresponding norm $\|\boldsymbol{u}\|_n = \sqrt{\boldsymbol{u} \cdot \boldsymbol{u}}$. The open ball centered with center $\boldsymbol{x} \in \mathbb{R}^n$ and radius $R > 0$ will be denoted by $\mathbb{B}_R(\boldsymbol{x})$. Given two matrices $\boldsymbol{A} = (a_{ij}), \boldsymbol{B} = (b_{ij}) \in \mathbb{R}^{n \times n}$, their scalar product is denoted by $\boldsymbol{A} : \boldsymbol{B} = \sum_{i,j=1}^n a_{ij} b_{ij}$. For a set $A \subset X$, $\overline{A}$ stands for the *closure* of A with respect to the topology of X. For $X = \mathbb{R}^n$ and $\bar{\boldsymbol{x}} \in A$ we denote by $\widehat{N}_A(\bar{\boldsymbol{x}})$ the *regular (Fréchet) normal cone* to A at $\bar{\boldsymbol{x}}$:

$$\widehat{N}_A(\bar{\boldsymbol{x}}) := \left\{ \boldsymbol{x}^* \in \mathbb{R}^n \,\middle|\, \limsup_{\boldsymbol{x} \xrightarrow{A} \bar{\boldsymbol{x}}} \frac{(\boldsymbol{x}^*, \boldsymbol{x} - \bar{\boldsymbol{x}})_n}{\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|_n} \leq 0 \right\},$$

whereas the *limiting (Mordukhovich) normal cone* to A at $\bar{\boldsymbol{x}}$ will be denoted by $N_A(\bar{\boldsymbol{x}})$:

$$N_A(\bar{\boldsymbol{x}}) := \operatorname{Lim\,sup}_{\boldsymbol{x} \xrightarrow{A} \bar{\boldsymbol{x}}} \widehat{N}_A(\boldsymbol{x}).$$

Here the symbol "Lim sup" stands for the Kuratowski–Painlevé outer limit of sets (cf. [23]). We say that A is *normally regular* at $\bar{\boldsymbol{x}}$, provided $\widehat{N}_A(\bar{\boldsymbol{x}}) = N_A(\bar{\boldsymbol{x}})$.

On the basis of these notions, local behavior of multifunctions may be investigated as follows. First, given a multifunction $Q : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, let us denote its *graph* by $\operatorname{Gr} Q := \{(\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{R}^n \times \mathbb{R}^m \mid \boldsymbol{y} \in Q(\boldsymbol{x})\}$. The *regular coderivative* of the closed-graph multifunction Q at the point $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}) \in \operatorname{Gr} Q$ is given by the multifunction $\widehat{D}^* Q(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}) : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$, which is defined as follows:

$$\widehat{D}^* Q(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}})(\boldsymbol{y}^*) := \{\boldsymbol{x}^* \in \mathbb{R}^n \mid (\boldsymbol{x}^*, -\boldsymbol{y}^*) \in \widehat{N}_{\operatorname{Gr} Q}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}})\}.$$

Analogously, the multifunction $D^* Q(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}) : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$, defined by

$$D^* Q(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}})(\boldsymbol{y}^*) := \{\boldsymbol{x}^* \in \mathbb{R}^n \mid (\boldsymbol{x}^*, -\boldsymbol{y}^*) \in N_{\operatorname{Gr} Q}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}})\},$$

is called the *limiting (Mordukhovich) coderivative* of Q at $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}})$. In the case when $Q(\bar{\boldsymbol{x}})$ is a singleton, one simply writes $\widehat{D}^* Q(\bar{\boldsymbol{x}})$ and $D^* Q(\bar{\boldsymbol{x}})$, respectively. Moreover, if Q happens to be continuously differentiable around $\bar{\boldsymbol{x}}$, then both coderivative mappings are single-valued and linear and amount to the adjoint Jacobian $(\nabla Q(\bar{\boldsymbol{x}}))^\top$.

Finally, we will need the notion of *strong regularity* [6], which in the case of a variational inequality goes back to Robinson [22]. Toward this end, let us consider a *generalized equation* (GE) of the form

$$(1) \qquad\qquad \boldsymbol{0} \in \boldsymbol{G}(\boldsymbol{x}, \boldsymbol{y}) + Q(\boldsymbol{y}),$$

where $\boldsymbol{G} : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ is continuously differentiable and $Q : \mathbb{R}^m \rightrightarrows \mathbb{R}^m$ is a closed-graph multifunction. Let (1) be satisfied for a pair $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}})$. We say that (1) is

strongly regular at $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}})$, provided there exist neighborhoods $\mathcal{U}$ of $\boldsymbol{0}$ and $\mathcal{V}$ of $\bar{\boldsymbol{y}}$ such that the mapping

$$\boldsymbol{\xi} \mapsto \{\boldsymbol{y} \in \mathcal{V} \mid \boldsymbol{\xi} \in \boldsymbol{G}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}) + \nabla_y \boldsymbol{G}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}})(\boldsymbol{y} - \bar{\boldsymbol{y}}) + Q(\boldsymbol{y})\}$$

is single-valued and Lipschitz on $\mathcal{U}$.

2. Problem formulation and discretization. The aim of this section is to present a class of optimal shape design problems in solid mechanics whose discretized form leads to the algebraic model analyzed in the following sections of this paper.

Let us consider a plane elastic body represented by a bounded domain $\Omega \subset \mathbb{R}^2$, with the boundary decomposed into three nonoverlapping parts: $\partial\Omega = \overline{\Gamma}_u \cup \overline{\Gamma}_p \cup \overline{\Gamma}_c$. On Γ_u the body is fixed; surface tractions of density $\boldsymbol{P} \in (L^2(\Gamma_p))^2$ act on Γ_p. The body is unilaterally supported by a rigid smooth foundation $\mathcal{H}$ along the portion Γ_c. In addition to nonpenetration conditions prescribed on Γ_c, we will take into account the influence of friction. Friction is involved in the model through the use of the local Coulomb law, whose coefficient $\mathcal{F}$ depends on the solution itself. Finally, Ω is subject to body forces of density $\boldsymbol{F} \in (L^2(\Omega))^2$. The equilibrium state of Ω is given by a displacement vector $\boldsymbol{u} : \Omega \to \mathbb{R}^2$, which satisfies the following system of equations and boundary conditions:

- *equilibrium equations*:

$$\text{(2)} \qquad \operatorname{div} \sigma(\boldsymbol{u}) + \boldsymbol{F} = \boldsymbol{0} \ \text{ in } \ \Omega;$$

- *Hooke's law*:

$$\text{(3)} \qquad \sigma(\boldsymbol{u}) = \mathcal{C}\varepsilon(\boldsymbol{u}), \quad \varepsilon(\boldsymbol{u}) := \frac{1}{2}\left(\nabla\boldsymbol{u} + (\nabla\boldsymbol{u})^T\right);$$

- *prescribed displacements*:

$$\text{(4)} \qquad \boldsymbol{u} = \boldsymbol{0} \ \text{ on } \ \Gamma_u;$$

- *prescribed tractions*:

$$\text{(5)} \qquad \boldsymbol{T}(\boldsymbol{u}) := \sigma(\boldsymbol{u})\boldsymbol{n} = \boldsymbol{P} \ \text{ on } \ \Gamma_p;$$

- *nonpenetration conditions*:

$$\text{(6)} \qquad u_n := \boldsymbol{u} \cdot \boldsymbol{n} \le d, \ T_n(\boldsymbol{u}) := \boldsymbol{T}(\boldsymbol{u}) \cdot \boldsymbol{n} \le 0, \ T_n(\boldsymbol{u})(u_n - d) = 0 \ \text{ on } \ \Gamma_c;$$

- *Coulomb's law of friction*:

$$\text{(7)} \qquad |T_t(\boldsymbol{u})| \le -\mathcal{F}(|u_t|)T_n(\boldsymbol{u}), \ u_t := \boldsymbol{u} \cdot \boldsymbol{t}, \ T_t(\boldsymbol{u}) := \boldsymbol{T}(\boldsymbol{u}) \cdot \boldsymbol{t} \ \text{ on } \ \Gamma_c;$$

$$u_t(x) \ne 0 \Rightarrow T_t(\boldsymbol{u})(x) = \mathcal{F}(|u_t(x)|)T_n(\boldsymbol{u}(x))\frac{u_t(x)}{|u_t(x)|}, \ x \in \Gamma_c.$$

The meanings of the symbols used in (2)–(7) are the following: $\sigma(\boldsymbol{u}) : \Omega \to \mathbb{R}^{2\times2}$ is the symmetric stress tensor corresponding to a displacement vector $\boldsymbol{u}$; $\varepsilon(\boldsymbol{u}) : \Omega \to \mathbb{R}^{2\times2}$ is the respective linearized strain tensor; $\mathcal{C}$ is the fourth order elasticity tensor; $\boldsymbol{n}$, $\boldsymbol{t}$ are the unit outward normal and tangential vector to $\partial\Omega$, respectively; $\boldsymbol{T}(\boldsymbol{u})$ is the stress vector on $\partial\Omega$; $T_n(\boldsymbol{u})$, $T_t(\boldsymbol{u})$ are the normal and tangential component of $\boldsymbol{T}(\boldsymbol{u})$, respectively (and similarly for $\boldsymbol{u}$, u_n, u_t). Finally, d is the so-called gap function characterizing the distance between Γ_c and $\mathcal{H}$.

We call any $\boldsymbol{u}$ satisfying (2)–(7) a *classical solution* to our problem. For the mathematical analysis of this problem we use its weak form. Toward this end we introduce the following spaces and sets of functions defined in Ω:

$$V = \{v \in H^1(\Omega) \mid v = 0 \ \text{ on } \ \Gamma_u\},$$
$$\boldsymbol{V} = V \times V,$$
$$\boldsymbol{K} = \{\boldsymbol{v} \in \boldsymbol{V} \mid v_n - d \leq 0 \ \text{ on } \ \Gamma_c\}.$$

DEFINITION 2.1. *We call any $\boldsymbol{u} \in \boldsymbol{K}$ satisfying the following implicit variational inequality a* weak solution *to* (2)–(7):

$$(\mathcal{P}) \qquad a(\boldsymbol{u}, \boldsymbol{v} - \boldsymbol{u}) + j(\boldsymbol{u}, \boldsymbol{v}) - j(\boldsymbol{u}, \boldsymbol{u}) \geq L(\boldsymbol{v} - \boldsymbol{u}) \quad \forall \boldsymbol{v} \in \boldsymbol{K}.$$

The bilinear form a, the linear form L, and the convex, nonsmooth functional j are defined as follows:

$$(8) \qquad \begin{cases} a(\boldsymbol{u}, \boldsymbol{v}) := \displaystyle\int_\Omega \mathcal{C}\varepsilon(\boldsymbol{u}) : \varepsilon(\boldsymbol{v}) \, d\boldsymbol{x}, \\[2mm] L(\boldsymbol{v}) := \displaystyle\int_\Omega \boldsymbol{F} \cdot \boldsymbol{v} \, d\boldsymbol{x} + \int_{\Gamma_P} \boldsymbol{P} \cdot \boldsymbol{v} \, ds, \\[2mm] j(\boldsymbol{u}, \boldsymbol{v}) := - \displaystyle\int_{\Gamma_c} \mathcal{F}(|u_t|) T_n(\boldsymbol{u}) |v_t| \, ds. \end{cases}$$

Remark 2.2. Let us mention that the definition of problem $(\mathcal{P})$ is presented only in a formal manner; for the rigorous setting we refer the reader to [7]. Applying Green's formula to integrals in $(\mathcal{P})$, we derive (2)–(7).

In addition to the assumptions on $\boldsymbol{F}$ and $\boldsymbol{P}$ we suppose that the coefficient of friction $\mathcal{F}$ is represented by a nonnegative, continuous, and bounded function in $\mathbb{R}^1_+$:

$$(9) \qquad \exists \mathcal{F}_{max} > 0: \ 0 \leq \mathcal{F} \leq \mathcal{F}_{max} \ \text{ in } \ \mathbb{R}^1_+, \quad \mathcal{F} \in C\left(\mathbb{R}^1_+\right).$$

We will need also a stronger property; namely, $\mathcal{F}$ is globally Lipschitz in $\mathbb{R}^1_+$:

$$(10) \qquad \exists C_{lip} > 0: \ |\mathcal{F}(x_1) - \mathcal{F}(x_2)| \leq C_{lip} |x_1 - x_2| \quad \forall x_1, x_2 \in \mathbb{R}^1_+.$$

Finally, the elasticity tensor $\mathcal{C}$ is assumed to be symmetric and positive definite resulting in the $\boldsymbol{V}$-ellipticity of the bilinear form a on $\boldsymbol{V}$:

$$\exists \gamma > 0: \ a(\boldsymbol{v}, \boldsymbol{v}) \geq \gamma \|\boldsymbol{v}\|^2_{1,\Omega} \quad \forall \boldsymbol{v} \in \boldsymbol{V}.$$

In order to release the unilateral constraint $\boldsymbol{v} \in \boldsymbol{K}$ we use the duality approach. Toward this end we introduce the following spaces and sets defined on Γ_c:

$$X = \{\varphi \in L^2(\Gamma_c) \mid \exists v \in V : \ v = \varphi \ \text{ on } \ \Gamma_c\},$$
$$X_+ = \{\varphi \in X \mid \varphi \geq 0 \ \text{ on } \ \Gamma_c\},$$
$$X'_+ = \{\mu \in X' \mid \langle \mu, \varphi \rangle \geq 0 \ \forall \varphi \in X_+\},$$

where X' stands for the topological dual of X.

Next we suppose that Ω is sufficiently smooth so that $v_n, v_t \in X$ for any $\boldsymbol{v} \in \boldsymbol{V}$. Finally, the symbol $\langle \cdot, \cdot \rangle$ stands for a duality pairing between X and X' with the property

$$\boldsymbol{v} \in \boldsymbol{K} \ \Leftrightarrow \ \boldsymbol{v} \in \boldsymbol{V} \ \text{ and } \ \langle \mu, v_n - d \rangle \leq 0 \quad \forall \mu \in X'_+.$$

DEFINITION 2.3. *By a mixed formulation of* $(\mathcal{P})$ *we mean a problem of finding a pair* $(\boldsymbol{u}, \lambda) \in \boldsymbol{V} \times X'_+$ *satisfying*

$$
(\mathcal{M}) \qquad
\begin{cases}
a(\boldsymbol{u}, \boldsymbol{v} - \boldsymbol{u}) + j(\boldsymbol{u}, \boldsymbol{v}) - j(\boldsymbol{u}, \boldsymbol{u}) \geq L(\boldsymbol{v} - \boldsymbol{u}) - \langle \lambda, v_n - u_n \rangle \quad \forall \boldsymbol{v} \in \boldsymbol{V}, \\
\qquad\qquad \langle \mu - \lambda, u_n - d \rangle \leq 0 \quad \forall \mu \in X'_+.
\end{cases}
$$

It can be shown (cf. [8]) that the first component $\boldsymbol{u}$ of the solution to $(\mathcal{M})$ is the displacement vector which solves $(\mathcal{P})$, while λ is the Lagrange multiplier releasing the unilateral constraint $u_n - d \leq 0$ on Γ_c. Moreover, $\lambda = -T_n(\boldsymbol{u})$. A suitable discretization of $(\mathcal{M})$ will play the key role in the sensitivity analysis of discretized shape optimization problems.

Now we pass to shape optimization problems in which $(\mathcal{M})$ is used as the state relation. Here and in what follows we will suppose that *only* the contact part Γ_c is the object of optimization. By U_{ad} we denote the set of all admissible contact parts Γ_c and by O_{ad} the admissible set of the corresponding Ω's. Next we suppose that the data $\boldsymbol{F}$, $\boldsymbol{P}$, and $\mathcal{C}$ in Ω are the restrictions of functions $\widehat{\boldsymbol{F}}$, $\widehat{\boldsymbol{P}}$, and $\widehat{\mathcal{C}}$, respectively, defined in a hold all domain $\widehat{\Omega}$, i.e., $\boldsymbol{F} = \widehat{\boldsymbol{F}}|_\Omega$, $\boldsymbol{P} = \widehat{\boldsymbol{P}}|_{\Gamma_p}$, $\mathcal{C} = \widehat{\mathcal{C}}|_\Omega$, $\Omega \in O_{ad}$. To emphasize that the previous spaces depend on a particular choice of $\Gamma_c \in U_{ad}$, and hence on $\Omega \in O_{ad}$, we write $V(\Omega)$, $\boldsymbol{V}(\Omega)$, $\boldsymbol{K}(\Omega)$, and for the trace spaces $X(\Gamma_c)$, $X'(\Gamma_c)$, etc. in what follows. Similarly, $(\mathcal{M}(\Omega))$ stands for the problem $(\mathcal{M})$ defined on Ω, and $(\boldsymbol{u}(\Omega), \lambda(\Gamma_c))$ stands for its solution.

Since the state problem $(\mathcal{M}(\Omega))$ may have more than one solution, the respective *control-to-state mapping* $\Phi : \Omega \mapsto (\boldsymbol{u}(\Omega), \lambda(\Gamma_c))$ is multivalued in general. Denote by $\mathcal{G}$ its graph.

Let $J : \Delta \to \mathbb{R}^1$, where $\Delta = \{(\Omega, \boldsymbol{y}, \mu) \mid \Omega \in O_{ad}, \boldsymbol{y} \in \boldsymbol{V}(\Omega), \mu \in X'(\Gamma_c)\}$, be an objective functional.

We define the following shape optimization problem:

$$
(\mathbb{P}') \qquad
\begin{cases}
\text{Find } (\Omega^*, \boldsymbol{u}(\Omega^*), \lambda(\Gamma_c^*)) \in \mathcal{G} \text{ such that} \\
(\Omega^*, \boldsymbol{u}(\Omega^*), \lambda(\Gamma_c^*)) \in \operatorname{argmin} \left\{ J(\Omega, \boldsymbol{u}(\Omega), \lambda(\Gamma_c)) \mid (\Omega, \boldsymbol{u}(\Omega), \lambda(\Gamma_c)) \in \mathcal{G} \right\}.
\end{cases}
$$

If Φ is single-valued, then the previous problem can be written as

$$
(\mathbb{P}) \qquad
\begin{cases}
\text{Find } \Omega^* \in O_{ad} \text{ such that} \\
\Omega^* \in \operatorname{argmin} \left\{ \mathcal{J}(\Omega) \mid \Omega \in O_{ad} \right\},
\end{cases}
$$

where $\mathcal{J}(\Omega) := J(\Omega, \boldsymbol{u}(\Omega), \lambda(\Gamma_c))$. This transformation is at the heart of ImP, and we use this approach for numerical solution of the shape optimization problem.

Now we pass to a discretization of $(\mathbb{P}')$ and $(\mathbb{P})$. It consists of a discretization of the set U_{ad} and of the state problem. We suppose that the shape of any $\Gamma_c \in U_{ad}$ is uniquely determined by s parameters $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_s)$, which belong to a compact set $\mathcal{U}_{ad} \subset \mathbb{R}^s$, and s does not depend on $\Gamma_c \in U_{ad}$. For instance, the vector $\boldsymbol{\alpha}$ can be formed by the coordinates of control points of Bézier curves. The number s then determines the degree of polynomials or is related to the number of segments if piecewise Bézier curves are used. If $\boldsymbol{\alpha} \in \mathcal{U}_{ad}$, then the corresponding domain will be denoted by $\Omega(\boldsymbol{\alpha})$ and termed the *discrete design* domain. Computations of the state problem are usually not done directly on $\Omega(\boldsymbol{\alpha})$ but on the so-called *discrete computational* domain denoted as $\Omega_h(\boldsymbol{\alpha})$ with the contact part $\Gamma_{ch}(\boldsymbol{\alpha})$. Indeed, in most cases we use a finite element method. Since any $\Omega(\boldsymbol{\alpha})$ is still a domain with

the curved boundary, computations of the state problem are usually performed on polygonal approximations $\Omega_h(\boldsymbol{\alpha})$ of $\Omega(\boldsymbol{\alpha})$. Then h is related to the norm of the triangulation of $\overline{\Omega}_h(\boldsymbol{\alpha})$.

A discretization of our state problem will be based on the formulation $(\mathcal{M}(\Omega(\boldsymbol{\alpha})))$, $\boldsymbol{\alpha} \in \mathcal{U}_{ad}$. Toward this end one has to use appropriate finite-dimensional approximations of $\boldsymbol{V}(\Omega)$, $X(\Gamma_c)$, $X'(\Gamma_c)$, $X'_+(\Gamma_c)$ denoted by $\boldsymbol{V}_h(\boldsymbol{\alpha})$, $X_h(\boldsymbol{\alpha})$, $X'_h(\boldsymbol{\alpha})$, and $X'_{h+}(\boldsymbol{\alpha})$, respectively. These spaces contain functions defined on $\Omega_h(\boldsymbol{\alpha})$ and $\Gamma_{ch}(\boldsymbol{\alpha})$, respectively. Let us note that $X'_h(\boldsymbol{\alpha}) \subseteq (X_h(\boldsymbol{\alpha}))'$; i.e., $X'_h(\boldsymbol{\alpha})$ is a subspace of the dual space to $X_h(\boldsymbol{\alpha})$ in general.

The approximation of $(\mathcal{M}(\Omega(\boldsymbol{\alpha})))$ reads as follows:

$$(\mathcal{M}_h(\boldsymbol{\alpha})) \quad \begin{cases} \text{Find } (\boldsymbol{u}_h(\boldsymbol{\alpha}), \lambda_h(\boldsymbol{\alpha})) \in \boldsymbol{V}_h(\boldsymbol{\alpha}) \times X'_{h+}(\boldsymbol{\alpha}) \text{ such that} \\ a(\boldsymbol{u}_h(\boldsymbol{\alpha}), \boldsymbol{v}_h - \boldsymbol{u}_h(\boldsymbol{\alpha})) + j_h(\boldsymbol{u}_h(\boldsymbol{\alpha}), \boldsymbol{v}_h) - j_h(\boldsymbol{u}_h(\boldsymbol{\alpha}), \boldsymbol{u}_h(\boldsymbol{\alpha})) \\ \qquad \geq L(\boldsymbol{v}_h - \boldsymbol{u}_h(\boldsymbol{\alpha})) - \langle \lambda_h(\boldsymbol{\alpha}), v_{hn} - u_{hn}(\boldsymbol{\alpha}) \rangle_h \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h(\boldsymbol{\alpha}), \\ \langle \mu_h - \lambda_h(\boldsymbol{\alpha}), u_{hn}(\boldsymbol{\alpha}) - d \rangle_h \leq 0 \quad \forall \mu_h \in X'_{h+}(\boldsymbol{\alpha}). \end{cases}$$

The forms a and L are defined by (8) with $\Omega := \Omega_h(\boldsymbol{\alpha})$, while

$$(11) \qquad j_h(\boldsymbol{u}_h(\boldsymbol{\alpha}), \boldsymbol{v}_h) = \int_{\Gamma_{ch}(\boldsymbol{\alpha})} \mathcal{F}(|u_{ht}(\boldsymbol{\alpha})|) \lambda_h(\boldsymbol{\alpha}) |v_{ht}| \, ds,$$

and $\langle \cdot, \cdot \rangle_h$ stands for a bilinear form on $X'_h(\boldsymbol{\alpha}) \times X_h(\boldsymbol{\alpha})$.

The discrete optimization problem is defined in a similar way as $(\mathbb{P}')$ and $(\mathbb{P})$, but using $\mathcal{U}_{ad}$ and $(\mathcal{M}_h(\boldsymbol{\alpha}))$ instead of U_{ad} and $(\mathcal{M}(\Omega))$, respectively. Assuming that $(\mathcal{M}_h(\boldsymbol{\alpha}))$ has a unique solution for any $\boldsymbol{\alpha} \in \mathcal{U}_{ad}$, the discrete form of $(\mathbb{P})$ reads as follows:

$$(\mathbb{P}_h) \qquad \begin{cases} \text{Find } \boldsymbol{\alpha}^* \in \mathcal{U}_{ad} \text{ such that} \\ \boldsymbol{\alpha}^* \in \text{argmin} \{ \mathcal{J}(\boldsymbol{\alpha}) \mid \boldsymbol{\alpha} \in \mathcal{U}_{ad} \}, \end{cases}$$

where $\mathcal{J}(\boldsymbol{\alpha}) := J(\Omega_h(\boldsymbol{\alpha}), \boldsymbol{u}_h(\boldsymbol{\alpha}), \lambda_h(\boldsymbol{\alpha}))$.

We conclude this section by presenting the algebraic form of $(\mathcal{M}_h(\boldsymbol{\alpha}))$. Suppose that $\dim \boldsymbol{V}_h(\boldsymbol{\alpha}) = m$, $\dim X_h(\boldsymbol{\alpha}) = p$, and $\dim X'_h(\boldsymbol{\alpha}) = q$, where m, p, q do not depend on $\boldsymbol{\alpha} \in \mathcal{U}_{ad}$. These spaces can be identified with $\mathbb{R}^m$, $\mathbb{R}^p$, and $\mathbb{R}^q$, respectively, and X'_{h+} with $\mathbb{R}^q_+$. The algebraic counterpart of a and L is obvious. For the evaluation of j_h and $\langle \cdot, \cdot \rangle_h$ given by integrals, we use appropriate integration formulas.

Let $[\cdot, \cdot]_h : \mathbb{R}^q \times \mathbb{R}^p \to \mathbb{R}^1$ be a bilinear form which approximates $\langle \cdot, \cdot \rangle_h$:

$$(12) \qquad \langle \mu_h, v_{hn} \rangle_h \sim [\boldsymbol{\mu}, \boldsymbol{v}_n]_h,$$

where $(\boldsymbol{\mu}, \boldsymbol{v}_n) \in \mathbb{R}^q \times \mathbb{R}^p$ are the coordinates of $(\mu_h, v_{hn}) \in X'_h(\boldsymbol{\alpha}) \times X_h(\boldsymbol{\alpha})$ with respect to a basis of $X'_h(\boldsymbol{\alpha}) \times X_h(\boldsymbol{\alpha})$. In what follows we suppose that this basis and $[\cdot, \cdot]_h$ are chosen in such a way that

$$(13) \qquad [\boldsymbol{\mu}, \boldsymbol{v}_n]_h = 0 \quad \forall \boldsymbol{v}_n \in \mathbb{R}^p \iff \boldsymbol{\mu} = \boldsymbol{0} \text{ in } \mathbb{R}^q$$

and

$$(14) \qquad [\boldsymbol{\mu}, \boldsymbol{v}_n]_h \leq 0 \quad \forall \boldsymbol{\mu} \in \mathbb{R}^q_+ \iff \boldsymbol{v}_n \leq \boldsymbol{0} \text{ in } \mathbb{R}^p.$$

From (13)–(14) it easily follows that $p = q$ so that $X_h'(\boldsymbol{\alpha}) = (X_h(\boldsymbol{\alpha}))'$. As a consequence of (14), the nonpenetration condition $\langle \mu_h, v_{hn} - d \rangle_h \leq 0 \ \forall \mu_h \in X_+'(\boldsymbol{\alpha})$ from $(\mathcal{M}_h(\boldsymbol{\alpha}))$ can be expressed in the componentwise form for $\boldsymbol{v}_n \in \mathbb{R}^p$:

$$(15) \qquad\qquad \boldsymbol{v}_n - \boldsymbol{\psi}(\boldsymbol{\alpha}) \leq \boldsymbol{0} \ \text{ in } \ \mathbb{R}^p,$$

where $\boldsymbol{\psi} = (\psi_1, \ldots, \psi_p) : \mathbb{R}^s \to \mathbb{R}^p$ is a new discrete gap-function which depends on the discrete design variable $\boldsymbol{\alpha} \in \mathcal{U}_{ad}$ and the distance function d. The frictional term j_h will be approximated as follows:

$$j_h(\boldsymbol{u}_h(\boldsymbol{\alpha}), \boldsymbol{v}_h) = \int_{\Gamma_{ch}(\boldsymbol{\alpha})} \mathcal{F}(|u_{ht}(\boldsymbol{\alpha})|)\lambda_h(\boldsymbol{\alpha})|v_{ht}| \, ds$$

$$(16) \qquad\qquad \approx \sum_{i=1}^p \omega_i(\boldsymbol{\alpha}) \mathcal{F}(|(\boldsymbol{u}_t(\boldsymbol{\alpha}))_i|)\boldsymbol{\lambda}_i(\boldsymbol{\alpha})|(\boldsymbol{v}_t)_i|,$$

where $\boldsymbol{v}_t \in \mathbb{R}^p$, $\boldsymbol{\lambda}(\boldsymbol{\alpha}) \in \mathbb{R}^p$ is the vector of coordinates of $v_{ht} \in X_h(\alpha)$, and $\lambda_h(\boldsymbol{\alpha}) \in X_{h+}'(\boldsymbol{\alpha})$, with respect to the corresponding basis. Further, $(\boldsymbol{a})_i$ stands for the ith coordinate of $\boldsymbol{a} \in \mathbb{R}^p$, and $\omega_i(\boldsymbol{\alpha})$, $i = 1, \ldots, p$, are weights of the used integration formula. We will suppose that $\omega_i(\boldsymbol{\alpha}) > 0 \ \forall i = 1, \ldots, p, \ \forall \boldsymbol{\alpha} \in \mathcal{U}_{ad}$.

To simplify notation, the last term in (16) will be written as $(\omega(\boldsymbol{\alpha}) \bullet \mathcal{F}(|\boldsymbol{u}_t(\boldsymbol{\alpha})|) \bullet \boldsymbol{\lambda}(\boldsymbol{\alpha}), |\boldsymbol{v}_t|)_p$, where $\boldsymbol{a} \bullet \boldsymbol{b} := (a_1 b_1, \ldots, a_p b_p) \in \mathbb{R}^p$, $\boldsymbol{a} = (a_1, \ldots, a_p)$, $\boldsymbol{b} = (b_1, \ldots, b_p)$.

On the basis of (12)–(16) the algebraic form of $(\mathcal{M}_h(\boldsymbol{\alpha}))$ reads as follows:

$$(\tilde{\mathcal{M}}(\boldsymbol{\alpha})) \quad \begin{cases} \text{Find } (\boldsymbol{u}(\boldsymbol{\alpha}), \boldsymbol{\lambda}(\boldsymbol{\alpha}), \boldsymbol{\lambda}_{alg}(\boldsymbol{\alpha})) \in \mathbb{R}^m \times \mathbb{R}_+^p \times \mathbb{R}_+^p : \\ (\boldsymbol{A}(\boldsymbol{\alpha})\boldsymbol{u}(\boldsymbol{\alpha}), \boldsymbol{v} - \boldsymbol{u}(\boldsymbol{\alpha}))_m + (\omega(\boldsymbol{\alpha}) \bullet \mathcal{F}(|\boldsymbol{u}_t(\boldsymbol{\alpha})|) \bullet \boldsymbol{\lambda}(\boldsymbol{\alpha}), |\boldsymbol{v}_t| - |\boldsymbol{u}_t(\boldsymbol{\alpha})|)_p \\ \qquad \geq (\boldsymbol{L}(\boldsymbol{\alpha}), \boldsymbol{v} - \boldsymbol{u}(\boldsymbol{\alpha}))_m + (\boldsymbol{\lambda}_{alg}(\boldsymbol{\alpha}), \boldsymbol{v}_n - \boldsymbol{u}_n(\boldsymbol{\alpha}))_p \quad \forall \boldsymbol{v} \in \mathbb{R}^m, \\ (\boldsymbol{\mu} - \boldsymbol{\lambda}_{alg}(\boldsymbol{\alpha}), \boldsymbol{u}_n(\boldsymbol{\alpha}) - \boldsymbol{\psi}(\boldsymbol{\alpha}))_p \leq 0 \quad \forall \boldsymbol{\mu} \in \mathbb{R}_+^p, \end{cases}$$

where $\boldsymbol{\lambda}_{alg}(\boldsymbol{\alpha}) \in \mathbb{R}_+^p$ is the algebraic Lagrange multiplier releasing the unilateral constraint (15). In the present form, problem $(\tilde{\mathcal{M}}(\boldsymbol{\alpha}))$ has too many unknowns. On the other hand, there exists a relation between $\boldsymbol{\lambda}_{alg}(\boldsymbol{\alpha})$ and $\boldsymbol{\lambda}(\boldsymbol{\alpha})$ which enables us to eliminate one of them. If, for example, $X_h'(\boldsymbol{\alpha})$ consists of piecewise constant functions on a specific partition of $\Gamma_{ch}(\boldsymbol{\alpha})$, then it can be shown that $\omega_i(\boldsymbol{\alpha})\boldsymbol{\lambda}_i(\boldsymbol{\alpha}) = (\boldsymbol{\lambda}_{alg}(\boldsymbol{\alpha}))_i$ (see [12]). Using this result, problem $(\tilde{\mathcal{M}}(\boldsymbol{\alpha}))$ can be expressed in terms of $(\boldsymbol{u}(\boldsymbol{\alpha}), \boldsymbol{\lambda}_{alg}(\boldsymbol{\alpha}))$. For simplicity of notation we write $\boldsymbol{\lambda}(\boldsymbol{\alpha})$ instead of $\boldsymbol{\lambda}_{alg}(\boldsymbol{\alpha})$. Thus the new problem reads as follows:

$$(\mathcal{M}(\boldsymbol{\alpha})) \quad \begin{cases} \text{Find } (\boldsymbol{u}, \boldsymbol{\lambda}) := (\boldsymbol{u}(\boldsymbol{\alpha}), \boldsymbol{\lambda}(\boldsymbol{\alpha})) \in \mathbb{R}^m \times \mathbb{R}_+^p : \\ (\boldsymbol{A}(\boldsymbol{\alpha})\boldsymbol{u}, \boldsymbol{v} - \boldsymbol{u})_m + (\mathcal{F}(|\boldsymbol{T}(\boldsymbol{\alpha})\boldsymbol{u}|) \bullet \boldsymbol{\lambda}(\boldsymbol{\alpha}), |\boldsymbol{T}(\boldsymbol{\alpha})\boldsymbol{v}| - |\boldsymbol{T}(\boldsymbol{\alpha})\boldsymbol{u}|)_p \\ \qquad \geq (\boldsymbol{L}(\boldsymbol{\alpha}), \boldsymbol{v} - \boldsymbol{u})_m + (\boldsymbol{\lambda}, \boldsymbol{N}(\boldsymbol{\alpha})\boldsymbol{v} - \boldsymbol{N}(\boldsymbol{\alpha})\boldsymbol{u})_p \quad \forall \boldsymbol{v} \in \mathbb{R}^m, \\ (\boldsymbol{\mu} - \boldsymbol{\lambda}, \boldsymbol{N}(\boldsymbol{\alpha})\boldsymbol{u} - \boldsymbol{\psi}(\boldsymbol{\alpha}))_p \leq 0 \quad \forall \boldsymbol{\mu} \in \mathbb{R}_+^p, \end{cases}$$

where $\boldsymbol{A}(\boldsymbol{\alpha}) \in \mathbb{R}^{m \times m}$ is a symmetric, positive definite matrix, $\boldsymbol{N}(\boldsymbol{\alpha}), \boldsymbol{T}(\boldsymbol{\alpha}) \in \mathbb{R}^{p \times m}$ are matrices representing the linear mappings $\boldsymbol{v} \mapsto \boldsymbol{v}_n$ and $\boldsymbol{v} \mapsto \boldsymbol{v}_t$, respectively, and $\boldsymbol{L}(\boldsymbol{\alpha})$ is a discretization of L on $\Omega_h(\boldsymbol{\alpha})$, $\boldsymbol{\alpha} \in \mathcal{U}_{ad}$.

3. Analysis of the discrete state problem $(\mathcal{M}(\boldsymbol{\alpha}))$. In this section we study the existence and uniqueness of the solution to $(\mathcal{M}(\boldsymbol{\alpha}))$ and its properties as a function of the discrete design variable $\boldsymbol{\alpha} \in \mathcal{U}_{ad}$.

In addition to (9) and (10) we will need the following assumptions on the data of $(\mathcal{M}(\boldsymbol{\alpha}))$:

The matrices $\boldsymbol{A}(\boldsymbol{\alpha})$ are *symmetric* $\forall \boldsymbol{\alpha} \in \mathcal{U}_{ad}$ and *uniformly* positive definite with respect to $\boldsymbol{\alpha} \in \mathcal{U}_{ad}$:

(17) $\exists \gamma > 0 : \ (\boldsymbol{A}(\boldsymbol{\alpha})\boldsymbol{x}, \boldsymbol{x})_m \geq \gamma \|\boldsymbol{x}\|_m^2 \quad \forall \boldsymbol{x} \in \mathbb{R}^m \, \forall \boldsymbol{\alpha} \in \mathcal{U}_{ad};$

(18)   the mappings $\boldsymbol{\alpha} \mapsto Z(\boldsymbol{\alpha}) \in \{\boldsymbol{A}(\boldsymbol{\alpha}), \boldsymbol{L}(\boldsymbol{\alpha}), \boldsymbol{\psi}(\boldsymbol{\alpha}), \boldsymbol{N}(\boldsymbol{\alpha}), \boldsymbol{T}(\boldsymbol{\alpha})\}$ are
       continuously differentiable in $\mathcal{U}_{ad}$;

(19)   the discrete gap function $\boldsymbol{\psi}$ is *nonnegative* in $\mathcal{U}_{ad}$;

(20) $\exists \beta > 0 : \ \sup\limits_{\boldsymbol{v} \in \mathbb{R}^m} \dfrac{(\boldsymbol{\mu}, \boldsymbol{N}(\boldsymbol{\alpha})\boldsymbol{v})_p}{\|\boldsymbol{v}\|_m} \geq \beta \|\boldsymbol{\mu}\|_p \quad \forall \boldsymbol{\mu} \in \mathbb{R}^p \, \forall \boldsymbol{\alpha} \in \mathcal{U}_{ad};$

(21)   the Euclidean norm of each row of $\boldsymbol{N}(\boldsymbol{\alpha})$ and $\boldsymbol{T}(\boldsymbol{\alpha})$ is 1 $\forall \boldsymbol{\alpha} \in \mathcal{U}_{ad};$

(22)   each column of $\boldsymbol{N}(\boldsymbol{\alpha})$ and $\boldsymbol{T}(\boldsymbol{\alpha})$ contains at most one nonzero element
       for every $\boldsymbol{\alpha} \in \mathcal{U}_{ad};$

(23) $\boldsymbol{N}(\boldsymbol{\alpha})(\mathbb{R}^m) = \boldsymbol{N}(\boldsymbol{\alpha})(\ker \boldsymbol{T}(\boldsymbol{\alpha})) \ \forall \boldsymbol{\alpha} \in \mathcal{U}_{ad}.$

Let $\boldsymbol{\alpha} \in \mathcal{U}_{ad}$ be fixed. To prove the existence of a solution to $(\mathcal{M}(\boldsymbol{\alpha}))$ we use a fixed-point approach. For any $(\boldsymbol{\varphi}, \boldsymbol{g}) \in \mathbb{R}_+^p \times \mathbb{R}_+^p$ we define the following auxiliary problem (since $\boldsymbol{\alpha} \in \mathcal{U}_{ad}$ is fixed, it will be omitted in arguments of functions):

$(\mathcal{M}(\boldsymbol{\varphi}, \boldsymbol{g}))$
$$\begin{cases} \text{Find } (\boldsymbol{u}, \boldsymbol{\lambda}) := (\boldsymbol{u}(\boldsymbol{\varphi}, \boldsymbol{g}), \boldsymbol{\lambda}(\boldsymbol{\varphi}, \boldsymbol{g})) \in \mathbb{R}^m \times \mathbb{R}_+^p \text{ such that} \\ (\boldsymbol{A}\boldsymbol{u}, \boldsymbol{v} - \boldsymbol{u})_m + (\mathcal{F}(\boldsymbol{\varphi}) \bullet \boldsymbol{g}, |\boldsymbol{T}\boldsymbol{v}| - |\boldsymbol{T}\boldsymbol{u}|)_p \\ \qquad \geq (\boldsymbol{L}, \boldsymbol{v} - \boldsymbol{u})_m + (\boldsymbol{\lambda}, \boldsymbol{N}\boldsymbol{v} - \boldsymbol{N}\boldsymbol{u})_p \quad \forall \boldsymbol{v} \in \mathbb{R}^m, \\ (\boldsymbol{\mu} - \boldsymbol{\lambda}, \boldsymbol{N}\boldsymbol{u} - \boldsymbol{\psi})_p \leq 0 \quad \forall \boldsymbol{\mu} \in \mathbb{R}_+^p. \end{cases}$$

It is well known that $(\mathcal{M}(\boldsymbol{\varphi}, \boldsymbol{g}))$ has a unique solution for any $(\boldsymbol{\varphi}, \boldsymbol{g}) \in \mathbb{R}_+^p \times \mathbb{R}_+^p$. In addition, $\boldsymbol{u} \in \mathcal{K}$, and it satisfies the variational inequality

(24) $(\boldsymbol{A}\boldsymbol{u}, \boldsymbol{v} - \boldsymbol{u})_m + (\mathcal{F}(\boldsymbol{\varphi}) \bullet \boldsymbol{g}, |\boldsymbol{T}\boldsymbol{v}| - |\boldsymbol{T}\boldsymbol{u}|)_p \geq (\boldsymbol{L}, \boldsymbol{v} - \boldsymbol{u})_m \quad \forall \boldsymbol{v} \in \mathcal{K},$

where

(25) $\mathcal{K} = \{\boldsymbol{x} \in \mathbb{R}^m \mid \boldsymbol{N}\boldsymbol{x} - \boldsymbol{\psi} \leq \boldsymbol{0}\}.$

The vector $\boldsymbol{\lambda} \in \mathbb{R}_+^p$ in $(\mathcal{M}(\boldsymbol{\varphi}, \boldsymbol{g}))$ is the Lagrange multiplier releasing the inequality constraint in $\mathcal{K}$.

Define a mapping $\Xi : \mathbb{R}_+^p \times \mathbb{R}_+^p \to \mathbb{R}_+^p \times \mathbb{R}_+^p$ by

$$\Xi(\boldsymbol{\varphi}, \boldsymbol{g}) := (|\boldsymbol{T}\boldsymbol{u}|, \boldsymbol{\lambda}) \quad \forall (\boldsymbol{\varphi}, \boldsymbol{g}) \in \mathbb{R}_+^p \times \mathbb{R}_+^p,$$

where $(\boldsymbol{u}, \boldsymbol{\lambda})$ solves $(\mathcal{M}(\boldsymbol{\varphi}, \boldsymbol{g}))$. Comparing the definitions of $(\mathcal{M}(\boldsymbol{\varphi}, \boldsymbol{g}))$ and $(\mathcal{M}(\boldsymbol{\alpha}))$ we see that the solution to $(\mathcal{M}(|\boldsymbol{T}\boldsymbol{u}|, \boldsymbol{\lambda}))$ is a solution of $(\mathcal{M}(\boldsymbol{\alpha}))$, too. In other words, $(|\boldsymbol{T}\boldsymbol{u}|, \boldsymbol{\lambda})$ is a fixed-point of Ξ in $\mathbb{R}_+^p \times \mathbb{R}_+^p$. For the proof of the existence and uniqueness of a fixed-point we use the Brouwer and Banach fixed-point theorems, respectively.

In [2] it has been shown that if (9), (17), (20), (21), and (23) are satisfied, then the solutions $(\boldsymbol{u}, \boldsymbol{\lambda})$ of $(\mathcal{M}(\boldsymbol{\varphi}, \boldsymbol{g}))$ are bounded:

(26) $\|\boldsymbol{u}\|_m \leq \dfrac{1}{\gamma} \|\boldsymbol{L}\|_m, \quad \|\boldsymbol{\lambda}\|_p \leq \dfrac{1}{\beta} \left(\dfrac{\|\boldsymbol{A}\|}{\gamma} + 1 \right) \|\boldsymbol{L}\|_m,$

where $\|\boldsymbol{A}\| := \sup\{\|\boldsymbol{A}(\boldsymbol{\alpha})\| \mid \boldsymbol{\alpha} \in \mathcal{U}_{ad}\}$, and γ and β are the constants from (17) and (20), respectively. From (21) and (26) it follows that

$$(27) \qquad \|\Xi(\boldsymbol{\varphi}, \boldsymbol{g})\|_{p+p} \leq \|\boldsymbol{u}\|_m + \|\boldsymbol{\lambda}\|_p \leq \left[\frac{1}{\gamma} + \frac{1}{\beta}\left(\frac{\|\boldsymbol{A}\|}{\gamma} + 1\right)\right]\|\boldsymbol{L}\|_m =: R.$$

Remark 3.1. Since γ and β do not depend on $\boldsymbol{\alpha} \in \mathcal{U}_{ad}$, neither does the constant R in (27).

It is readily seen that the mapping Ξ is continuous in $\mathbb{R}_+^p \times \mathbb{R}_+^p$. From this, (27), and the Brouwer fixed-point theorem, we obtain the following existence result.

THEOREM 3.2. *Let* (9), (17), (20), (21), *and* (23) *be satisfied. Then problem* $(\mathcal{M}(\boldsymbol{\alpha}))$ *has at least one solution for any* $\boldsymbol{\alpha} \in \mathcal{U}_{ad}$. *In addition, all solutions* $(\boldsymbol{u}, \boldsymbol{\lambda})$ *lie in* $(\mathbb{R}^m \times \mathbb{R}_+^p) \cap \overline{\mathbb{B}}_R(\boldsymbol{0})$, *where* R *is defined as in* (27) *and does not depend on* $\boldsymbol{\alpha} \in \mathcal{U}_{ad}$.

Next we show that if, in addition to (9), $\mathcal{F}$ satisfies (10), then Ξ is Lipschitz in $(\mathbb{R}_+^p \times \mathbb{R}_+^p) \cap \overline{\mathbb{B}}_R(\boldsymbol{0})$ and the modulus of Lipschitz continuity will be estimated.

Let $(\boldsymbol{\varphi}^{(i)}, \boldsymbol{g}^{(i)}) \in (\mathbb{R}_+^p \times \mathbb{R}_+^p) \cap \overline{\mathbb{B}}_R(\boldsymbol{0})$ be given and denote by $(\boldsymbol{u}^{(i)}, \boldsymbol{\lambda}^{(i)})$ the solutions of $(\mathcal{M}(\boldsymbol{\varphi}^{(i)}, \boldsymbol{g}^{(i)}))$, $i = 1, 2$. From (24) we know that $\boldsymbol{u}^{(i)} \in \mathcal{K}$ solves the following variational inequality:

$$(\mathcal{A}^{(i)}) \qquad \begin{cases} \left(\boldsymbol{A}\boldsymbol{u}^{(i)}, \boldsymbol{v}^{(i)} - \boldsymbol{u}^{(i)}\right)_m + \left(\mathcal{F}(\boldsymbol{\varphi}^{(i)}) \bullet \boldsymbol{g}^{(i)}, |\boldsymbol{T}\boldsymbol{v}| - |\boldsymbol{T}\boldsymbol{u}^{(i)}|\right)_p \\ \qquad\qquad\qquad \geq \left(\boldsymbol{L}, \boldsymbol{v} - \boldsymbol{u}^{(i)}\right)_m \quad \forall \boldsymbol{v} \in \mathcal{K}, \end{cases}$$

$i = 1, 2$. Inserting $\boldsymbol{v} = \boldsymbol{u}^{(2)}$ into $(\mathcal{A}^{(1)})$ and $\boldsymbol{v} = \boldsymbol{u}^{(1)}$ into $(\mathcal{A}^{(2)})$, adding these inequalities, and using (17), we obtain

$$(28) \qquad \gamma\|\boldsymbol{u}^{(1)} - \boldsymbol{u}^{(2)}\|_m^2 \leq \left(\mathcal{F}(\boldsymbol{\varphi}^{(1)}) \bullet \boldsymbol{g}^{(1)} - \mathcal{F}(\boldsymbol{\varphi}^{(2)}) \bullet \boldsymbol{g}^{(2)}, |\boldsymbol{T}\boldsymbol{u}^{(2)}| - |\boldsymbol{T}\boldsymbol{u}^{(1)}|\right)_p.$$

Adding and subtracting the term $\mathcal{F}(\boldsymbol{\varphi}^{(1)}) \bullet \boldsymbol{g}^{(2)}$ on the right of (28) and using (9), (10), the Cauchy–Schwarz inequality, and (21), we get

$$\gamma\|\boldsymbol{u}^{(1)} - \boldsymbol{u}^{(2)}\|_m^2 \leq \left(\mathcal{F}_{max}\|\boldsymbol{g}^{(2)} - \boldsymbol{g}^{(1)}\|_p + C_{lip}\|\boldsymbol{g}^{(2)}\|_\infty\|\boldsymbol{\varphi}^{(1)} - \boldsymbol{\varphi}^{(2)}\|_p\right)\|\boldsymbol{u}^{(1)} - \boldsymbol{u}^{(2)}\|_m.$$

Thus

$$\|\boldsymbol{T}\boldsymbol{u}^{(1)} - \boldsymbol{T}\boldsymbol{u}^{(2)}\|_p \leq \|\boldsymbol{u}^{(1)} - \boldsymbol{u}^{(2)}\|_m$$

$$\leq \frac{1}{\gamma}\left(\mathcal{F}_{max}\|\boldsymbol{g}^{(1)} - \boldsymbol{g}^{(2)}\|_p + C_{lip}R\|\boldsymbol{\varphi}^{(1)} - \boldsymbol{\varphi}^{(2)}\|_p\right)$$

$$(29) \qquad \leq \frac{1}{\gamma}\max\{\mathcal{F}_{max}, C_{lip}R\}\|(\boldsymbol{\varphi}^{(1)} - \boldsymbol{\varphi}^{(2)}, \boldsymbol{g}^{(1)} - \boldsymbol{g}^{(2)})\|_{p+p}.$$

It remains to estimate $\|\boldsymbol{\lambda}^{(1)} - \boldsymbol{\lambda}^{(2)}\|_p$. From the definition of $(\mathcal{M}(\boldsymbol{\varphi}^{(i)}, \boldsymbol{g}^{(i)}))$ it follows that

$$\left(\boldsymbol{A}\boldsymbol{u}^{(i)}, \boldsymbol{v}\right)_m = (\boldsymbol{L}, \boldsymbol{v})_m + \left(\boldsymbol{\lambda}^{(i)}, \boldsymbol{N}\boldsymbol{v}\right)_p \quad \forall \boldsymbol{v} \in \ker \boldsymbol{T}, \quad i = 1, 2,$$

and consequently,

$$\left(\boldsymbol{\lambda}^{(1)} - \boldsymbol{\lambda}^{(2)}, \boldsymbol{N}\boldsymbol{v}\right)_p = \left(\boldsymbol{A}(\boldsymbol{u}^{(1)} - \boldsymbol{u}^{(2)}), \boldsymbol{v}\right)_m \quad \forall \boldsymbol{v} \in \ker \boldsymbol{T}.$$

From this and (23) we get

$$\beta\|\boldsymbol{\lambda}^{(1)} - \boldsymbol{\lambda}^{(2)}\|_p \leq \|\boldsymbol{A}\|\|\boldsymbol{u}^{(1)} - \boldsymbol{u}^{(2)}\|_m.$$

This, and (29) give the following estimate:

$$(30) \quad \|\Xi(\boldsymbol{\varphi}^{(1)}, \boldsymbol{g}^{(1)}) - \Xi(\boldsymbol{\varphi}^{(2)}, \boldsymbol{g}^{(2)})\|_{p+p} \leq \|\boldsymbol{u}^{(1)} - \boldsymbol{u}^{(2)}\|_m + \|\boldsymbol{\lambda}^{(1)} - \boldsymbol{\lambda}^{(2)}\|_p$$

$$\leq \frac{\beta + \|\boldsymbol{A}\|}{\beta\gamma} \max\{\mathcal{F}_{max}, C_{lip}R\}\|(\boldsymbol{\varphi}^{(1)} - \boldsymbol{\varphi}^{(2)}, \boldsymbol{g}^{(1)} - \boldsymbol{g}^{(2)})\|_{p+p}.$$

THEOREM 3.3. *In addition to the assumptions of Theorem 3.2, let* (10) *be satisfied. Then the mapping* Ξ *is Lipschitz in* $(\mathbb{R}_+^p \times \mathbb{R}_+^p) \cap \overline{\mathbb{B}}_R(\boldsymbol{0})$ *with the Lipschitz modulus equal to* $\dfrac{\beta + \|\boldsymbol{A}\|}{\beta\gamma} \max\{\mathcal{F}_{max}, C_{lip}R\}$, *where* $\mathcal{F}_{max}$, C_{lip} *and* R *are as in* (9), (10) *and* (27), *respectively.*

A direct consequence of this theorem is the following uniqueness result.

THEOREM 3.4. *Let all the assumptions of Theorem 3.3 be satisfied. If*

$$(31) \qquad\qquad \max\{\mathcal{F}_{max}, C_{lip}R\} < \frac{\beta\gamma}{\beta + \|\boldsymbol{A}\|},$$

then Ξ *is contractive in* $(\mathbb{R}_+^p \times \mathbb{R}_+^p) \cap \overline{\mathbb{B}}_R(\boldsymbol{0})$. *Consequently* Ξ *has a unique fixed-point or, equivalently,* $(\mathcal{M}(\boldsymbol{\alpha}))$ *has a unique solution for any* $\boldsymbol{\alpha} \in \mathcal{U}_{ad}$. *Moreover, the upper bound* (31) *does not depend on* $\boldsymbol{\alpha} \in \mathcal{U}_{ad}$.

Let us comment briefly on the assumptions (17)–(23). Since $\boldsymbol{A}(\boldsymbol{\alpha})$ is positive definite for any $\boldsymbol{\alpha} \in \mathcal{U}_{ad}$, the matrix function $\boldsymbol{A}$ is continuous in $\mathcal{U}_{ad}$, and $\mathcal{U}_{ad}$ is compact, we get (17). Next we will suppose that the finite-dimensional space $\boldsymbol{V}_h(\boldsymbol{\alpha})$ consists of piecewise polynomial functions constructed over a triangulation $\mathcal{T}_h(\boldsymbol{\alpha})$ of $\overline{\Omega}_h(\boldsymbol{\alpha})$. In addition to the requirement that $\dim \boldsymbol{V}_h(\boldsymbol{\alpha})$ does not depend on $\boldsymbol{\alpha} \in \mathcal{U}_{ad}$, we suppose that the position of the nodes of $\mathcal{T}_h(\boldsymbol{\alpha})$ depends solely on the position of $\boldsymbol{\alpha}$ in a smooth way and that the nodes themselves do not change their neighbors when changing $\boldsymbol{\alpha}$. If this is so, then (18) is satisfied (cf. [9]). The condition $\boldsymbol{\psi} \geq \boldsymbol{0}$ in $\mathcal{U}_{ad}$ will be included as a constraint in the definition of $\mathcal{U}_{ad}$. Hence (19) is fulfilled. The remaining assumptions (20)–(23) will be easily satisfied when the Lagrange type spaces $\boldsymbol{V}_h(\boldsymbol{\alpha})$ are used. In this case the vector $\boldsymbol{v}$ of the coordinates of $\boldsymbol{v}_h \in \boldsymbol{V}_h(\boldsymbol{\alpha})$ consists of the values of $\boldsymbol{v}_h$ at the nodal points (see [4]). Then each row of the matrices $\boldsymbol{N}(\boldsymbol{\alpha})$, $\boldsymbol{T}(\boldsymbol{\alpha})$ contains at least one and at most two nonzero elements $\forall \boldsymbol{\alpha} \in \mathcal{U}_{ad}$, namely the coordinates of $\boldsymbol{n}$, and $\boldsymbol{t}$, respectively, at the corresponding contact node, i.e., the node of $\mathcal{T}_h(\boldsymbol{\alpha})$ lying on $\overline{\Gamma}_{ch}(\boldsymbol{\alpha}) \setminus \overline{\Gamma}_u(\boldsymbol{\alpha})$. From this, (21) follows. Moreover, the inequality constraint (15) can be split into m inequality constraints imposed just on two components of $\boldsymbol{v}$ which represent the displacement at the individual contact nodes. Thus (22) is satisfied. Clearly, $\boldsymbol{N}(\boldsymbol{\alpha})$ has the full row rank $\forall \boldsymbol{\alpha} \in \mathcal{U}_{ad}$. From this (20) follows, making use of the compactness of $\mathcal{U}_{ad}$. Finally, it is also readily seen that (23) is satisfied.

In the rest of the paper we will assume that all assumptions of Theorem 3.4 are satisfied, ensuring in particular that the state problems $(\mathcal{M}(\boldsymbol{\alpha}))$, $\boldsymbol{\alpha} \in \mathcal{U}_{ad}$ are uniquely solvable. In the next part of this section we introduce a reduced version of $(\mathcal{M}(\boldsymbol{\alpha}))$, formulated as a generalized equation, and show that its solution is a Lipschitz function of the design variable $\boldsymbol{\alpha} \in \mathcal{U}_{ad}$. Although this fact could be proven directly, in our presentation it will be a consequence of the strong regularity condition that will play an important role in sensitivity analysis, too.

We start with an auxiliary result. Keeping $\boldsymbol{\alpha} \in \mathcal{U}_{ad}$ fixed, we show in the next theorem that the solution of $(\mathcal{M}(\boldsymbol{\alpha}))$ is a locally Lipschitzian function of the *load vector* $\boldsymbol{L}$. Since the domain, corresponding to the design vector $\boldsymbol{\alpha} \in \mathcal{U}_{ad}$, will be

fixed and $\boldsymbol{L}$ will be variable, let us relabel the problem $(\mathcal{M}(\boldsymbol{\alpha}))$ as $(\mathcal{M}(\boldsymbol{L}))$ and the auxiliary problems $(\mathcal{M}(\boldsymbol{\varphi}, \boldsymbol{g}))$ as $(\mathcal{M}(\boldsymbol{L}, \boldsymbol{\varphi}, \boldsymbol{g}))$. Also, $\boldsymbol{\alpha}$ in the argument of functions will be omitted.

Denote the modulus of Lipschitz continuity of Ξ (cf. (30)) by

$$\delta(\|\boldsymbol{L}\|_m) := \frac{\beta + \|\boldsymbol{A}\|}{\beta\gamma} \max\{\mathcal{F}_{max}, R(\|\boldsymbol{L}\|_m)C_{lip}\},$$

where (cf. (27))

$$(32) \qquad R(\|\boldsymbol{L}\|_m) = \left[\frac{1}{\gamma} + \frac{1}{\beta}\left(\frac{\|\boldsymbol{A}\|}{\gamma} + 1\right)\right] \|\boldsymbol{L}\|_m =: \kappa\|\boldsymbol{L}\|_m.$$

Recall that the function δ and the constant κ do not depend on $\boldsymbol{\alpha} \in \mathcal{U}_{ad}$. Using this notation, we may rewrite Proposition 3.5 from [2] as follows.

LEMMA 3.5. *Let* $(\boldsymbol{u}^{(i)}, \boldsymbol{\lambda}^{(i)})$ *be the solutions of* $(\mathcal{M}(\boldsymbol{L}^{(i)}, \boldsymbol{\varphi}, \boldsymbol{g}))$, $i = 1, 2$. *Then*

$$(33) \qquad \|(\boldsymbol{u}^{(1)}, \boldsymbol{\lambda}^{(1)}) - (\boldsymbol{u}^{(2)}, \boldsymbol{\lambda}^{(2)})\|_{m+p} \le \kappa\|\boldsymbol{L}^{(1)} - \boldsymbol{L}^{(2)}\|_m$$

with κ *from* (32).

Remark 3.6. It is worth noticing that (33) holds for any $(\boldsymbol{\varphi}, \boldsymbol{g}) \in \mathbb{R}_+^p \times \mathbb{R}_+^p$ and any $\boldsymbol{\alpha} \in \mathcal{U}_{ad}$. Now we may prove the announced Lipschitz continuity result.

PROPOSITION 3.7. *Let* $\boldsymbol{\alpha} \in \mathcal{U}_{ad}$ *be fixed and let the assumptions of Theorem 3.4 be satisfied, i.e.,* $\delta(\|\boldsymbol{L}\|_m) < 1$ *for some* $\boldsymbol{L} \in \mathbb{R}^m$. *Then there exist positive constants* ϵ *and* $K := K(\boldsymbol{L}, \epsilon)$ *such that*

$$\|(\overline{\boldsymbol{u}}, \overline{\boldsymbol{\lambda}}) - (\widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{\lambda}})\|_{m+p} \le K\|\overline{\boldsymbol{L}} - \widetilde{\boldsymbol{L}}\|_m \quad \forall \overline{\boldsymbol{L}}, \widetilde{\boldsymbol{L}} \in \mathbb{B}_\epsilon(\boldsymbol{L}),$$

where $(\overline{\boldsymbol{u}}, \overline{\boldsymbol{\lambda}})$, $(\widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{\lambda}})$ *denote the unique solutions of* $(\mathcal{M}(\overline{\boldsymbol{L}}))$ *and* $(\mathcal{M}(\widetilde{\boldsymbol{L}}))$, *respectively.*

Proof. Existence of $\epsilon > 0$ satisfying

$$(34) \qquad \delta(\|\boldsymbol{L}'\|_m) < 1 \quad \forall \boldsymbol{L}' \in \overline{\mathbb{B}}_\epsilon(\boldsymbol{L})$$

follows immediately by continuity of the function $\delta : \mathbb{R}_+^1 \to \mathbb{R}_+^1$. We choose such ϵ and denote $r := \max\{\delta(\|\boldsymbol{L}'\|_m) \mid \boldsymbol{L}' \in \overline{\mathbb{B}}_\epsilon(\boldsymbol{L})\}$. From (34) it holds that $0 < r < 1$. Further, let $\overline{\boldsymbol{L}}, \widetilde{\boldsymbol{L}} \in \mathbb{B}_\epsilon(\boldsymbol{L})$ and $(\boldsymbol{\varphi}, \boldsymbol{g}) \in \mathbb{R}_+^p \times \mathbb{R}_+^p$ be arbitrary. Then we build recurrently the following sequences:

- Let $(\overline{\boldsymbol{u}}^{(0)}, \overline{\boldsymbol{\lambda}}^{(0)})$ denote the solution of the auxiliary problem $(\mathcal{M}(\overline{\boldsymbol{L}}, \boldsymbol{\varphi}, \boldsymbol{g}))$. For each $k \in \mathbb{N}$ then denote by $(\overline{\boldsymbol{u}}^{(k)}, \overline{\boldsymbol{\lambda}}^{(k)}) \in \mathbb{R}^m \times \mathbb{R}_+^p$ the unique solution to the problem $(\mathcal{M}(\overline{\boldsymbol{L}}, |\boldsymbol{T}\overline{\boldsymbol{u}}^{(k-1)}|, \overline{\boldsymbol{\lambda}}^{(k-1)}))$.
- Let $(\widetilde{\boldsymbol{u}}^{(0)}, \widetilde{\boldsymbol{\lambda}}^{(0)})$ stand for the solution of $(\mathcal{M}(\widetilde{\boldsymbol{L}}, \boldsymbol{\varphi}, \boldsymbol{g}))$. Analogously, for each $k \in \mathbb{N}$ denote by $(\widetilde{\boldsymbol{u}}^{(k)}, \widetilde{\boldsymbol{\lambda}}^{(k)}) \in \mathbb{R}^m \times \mathbb{R}_+^p$ the solution of the contact problem $(\mathcal{M}(\widetilde{\boldsymbol{L}}, |\boldsymbol{T}\widetilde{\boldsymbol{u}}^{(k-1)}|, \widetilde{\boldsymbol{\lambda}}^{(k-1)}))$.
- Finally, for every $k \in \mathbb{N}$ let $(\boldsymbol{U}^{(k)}, \Lambda^{(k)})$ stand for the unique solution of problem $(\mathcal{M}(\overline{\boldsymbol{L}}, |\boldsymbol{T}\widetilde{\boldsymbol{u}}^{(k-1)}|, \widetilde{\boldsymbol{\lambda}}^{(k-1)}))$.

It follows from the contractivity of Ξ that the sequences $\{(|\boldsymbol{T}\overline{\boldsymbol{u}}^{(k)}|, \overline{\boldsymbol{\lambda}}^{(k)})\}_{k\in\mathbb{N}}$ and $\{(|\boldsymbol{T}\widetilde{\boldsymbol{u}}^{(k)}|, \widetilde{\boldsymbol{\lambda}}^{(k)})\}_{k\in\mathbb{N}}$ of elements from $\mathbb{R}_+^p \times \mathbb{R}_+^p$ tend to the unique fixed-points of Ξ in $\mathbb{R}_+^p \times \mathbb{R}_+^p$; i.e., the sequences $\{(\overline{\boldsymbol{u}}^{(k)}, \overline{\boldsymbol{\lambda}}^{(k)})\}_{k\in\mathbb{N}}$ and $\{(\widetilde{\boldsymbol{u}}^{(k)}, \widetilde{\boldsymbol{\lambda}}^{(k)})\}_{k\in\mathbb{N}}$ converge

to the unique solutions $(\overline{\boldsymbol{u}}, \overline{\boldsymbol{\lambda}})$ and $(\widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{\lambda}})$ of $(\mathcal{M}(\overline{\boldsymbol{L}}))$ and $(\mathcal{M}(\widetilde{\boldsymbol{L}}))$, respectively. Now, one has

$$
\begin{aligned}
\|(\overline{\boldsymbol{u}}^{(k)}, &\overline{\boldsymbol{\lambda}}^{(k)}) - (\widetilde{\boldsymbol{u}}^{(k)}, \widetilde{\boldsymbol{\lambda}}^{(k)})\|_{m+p} \\
&\leq \|(\overline{\boldsymbol{u}}^{(k)}, \overline{\boldsymbol{\lambda}}^{(k)}) - (\boldsymbol{U}^{(k)}, \Lambda^{(k)})\|_{m+p} + \|(\boldsymbol{U}^{(k)}, \Lambda^{(k)}) - (\widetilde{\boldsymbol{u}}^{(k)}, \widetilde{\boldsymbol{\lambda}}^{(k)})\|_{m+p} \\
&\leq \delta(\|\overline{\boldsymbol{L}}\|_m)\|(|\boldsymbol{T}\overline{\boldsymbol{u}}^{(k-1)}|, \overline{\boldsymbol{\lambda}}^{(k-1)}) - (|\boldsymbol{T}\widetilde{\boldsymbol{u}}^{(k-1)}|, \widetilde{\boldsymbol{\lambda}}^{(k-1)})\|_{p+p} + \kappa\|\overline{\boldsymbol{L}} - \widetilde{\boldsymbol{L}}\|_m \\
&\leq r\|(\overline{\boldsymbol{u}}^{(k-1)}, \overline{\boldsymbol{\lambda}}^{(k-1)}) - (\widetilde{\boldsymbol{u}}^{(k-1)}, \widetilde{\boldsymbol{\lambda}}^{(k-1)})\|_{m+p} + \kappa\|\overline{\boldsymbol{L}} - \widetilde{\boldsymbol{L}}\|_m,
\end{aligned}
$$

making use of (30) and (33). Since the above estimate holds for every $k \in \mathbb{N}$, we obtain by induction

$$
\begin{aligned}
\|(\overline{\boldsymbol{u}}^{(k)}, &\overline{\boldsymbol{\lambda}}^{(k)}) - (\widetilde{\boldsymbol{u}}^{(k)}, \widetilde{\boldsymbol{\lambda}}^{(k)})\|_{m+p} \\
&\leq r^k\|(\overline{\boldsymbol{u}}^{(0)}, \overline{\boldsymbol{\lambda}}^{(0)}) - (\widetilde{\boldsymbol{u}}^{(0)}, \widetilde{\boldsymbol{\lambda}}^{(0)})\|_{m+p} + (r^{k-1} + \cdots + r + 1)\kappa\|\overline{\boldsymbol{L}} - \widetilde{\boldsymbol{L}}\|_m \\
&\leq r^k\kappa\|\overline{\boldsymbol{L}} - \widetilde{\boldsymbol{L}}\|_m + (r^{k-1} + \cdots + r + 1)\kappa\|\overline{\boldsymbol{L}} - \widetilde{\boldsymbol{L}}\|_m \\
&\leq \frac{\kappa}{1-r}\|\overline{\boldsymbol{L}} - \widetilde{\boldsymbol{L}}\|_m.
\end{aligned}
$$

Finally, letting $k \to \infty$ we finish the proof.   □

Next, we present a reduced version of the state problem $(\mathcal{M}(\boldsymbol{\alpha}))$, $\boldsymbol{\alpha} \in \mathcal{U}_{ad}$, that involves only variables defined on the contact boundary, i.e., $\boldsymbol{u}_t(\boldsymbol{\alpha}) = \boldsymbol{T}(\boldsymbol{\alpha})\boldsymbol{u}(\boldsymbol{\alpha})$, $\boldsymbol{u}_n(\boldsymbol{\alpha}) = \boldsymbol{N}(\boldsymbol{\alpha})\boldsymbol{u}(\boldsymbol{\alpha})$, and $\boldsymbol{\lambda}(\boldsymbol{\alpha})$. The following generalizes the procedure described in section 3 of [3] (see also [2]).

Let $\boldsymbol{\alpha} \in \mathcal{U}_{ad}$ be fixed. Observe that conditions (21) and (22) yield matrices $\boldsymbol{N}(\boldsymbol{\alpha})$ and $\boldsymbol{T}(\boldsymbol{\alpha})$ with orthonormal rows, respectively. Moreover, it is easy to check that (23) implies $\boldsymbol{N}(\boldsymbol{\alpha})^\top \boldsymbol{T}(\boldsymbol{\alpha}) = \boldsymbol{0} \in \mathbb{R}^{p \times p}$. In other words, assuming (21)–(23), the rows of $\boldsymbol{N}(\boldsymbol{\alpha})$ and $\boldsymbol{T}(\boldsymbol{\alpha})$ form a system of $2p$ orthonormal vectors for each $\boldsymbol{\alpha} \in \mathcal{U}_{ad}$. We complete this set by $(m-2p)$ vectors into an orthonormal basis and define the matrix $\boldsymbol{I}(\boldsymbol{\alpha}) \in \mathbb{R}^{(m-2p) \times m}$ as the one containing these vectors in its rows. Thus, for any displacement field $\boldsymbol{v} \in \mathbb{R}^m$, the vector $\boldsymbol{v}_{int}(\boldsymbol{\alpha}) := \boldsymbol{I}(\boldsymbol{\alpha})\boldsymbol{v} \in \mathbb{R}^{(m-2p) \times m}$ contains the nodal values of $\boldsymbol{v}$ at the *noncontact* nodes of $\mathcal{T}_h(\boldsymbol{\alpha})$, and the following decomposition holds true:

$$
(35) \qquad \boldsymbol{v} = \boldsymbol{I}^\top(\boldsymbol{\alpha})\boldsymbol{v}_{int} + \boldsymbol{T}^\top(\boldsymbol{\alpha})\boldsymbol{v}_t + \boldsymbol{N}^\top(\boldsymbol{\alpha})\boldsymbol{v}_n.
$$

Further, let us define the matrices (to unburden the notation, we skip the argument $\boldsymbol{\alpha}$) $\boldsymbol{A}_{II} := \boldsymbol{I A I}^\top$, $\boldsymbol{A}_{IT} := \boldsymbol{I A T}^\top$, $\boldsymbol{A}_{IN} := \boldsymbol{I A N}^\top$, $\boldsymbol{A}_S := \boldsymbol{A} - \boldsymbol{I}^\top \boldsymbol{A}_{II}^{-1} \boldsymbol{I A}$, $\boldsymbol{A}_{tn} := \boldsymbol{T A}_S \boldsymbol{N}^\top$ (analogously for $\boldsymbol{A}_{tt}$, $\boldsymbol{A}_{nt}$, $\boldsymbol{A}_{nn}$), and the vectors $\boldsymbol{L}_S := \boldsymbol{L} - \boldsymbol{A I}^\top \boldsymbol{A}_{II}^{-1} \boldsymbol{I L}$, $\boldsymbol{L}_t := \boldsymbol{T L}_S$, $\boldsymbol{L}_n := \boldsymbol{N L}_S$. Employing this notation and inserting appropriate test vectors $\boldsymbol{v} \in \mathbb{R}^m$ into $(\mathcal{M}(\boldsymbol{\varphi}, \boldsymbol{g}))$ $((\boldsymbol{\varphi}, \boldsymbol{g}) \in \mathbb{R}^p_+ \times \mathbb{R}^p_+$ fixed), one may easily derive the following equivalent system of equations and generalized equations for the new unknowns $(\boldsymbol{u}_{int}, \boldsymbol{u}_t, \boldsymbol{u}_n, \boldsymbol{\lambda}) := (\boldsymbol{u}_{int}(\boldsymbol{\alpha}), \boldsymbol{u}_n(\boldsymbol{\alpha}), \boldsymbol{u}_t(\boldsymbol{\alpha}), \boldsymbol{\lambda}(\boldsymbol{\alpha}))$:

$$
(36) \qquad \boldsymbol{A}_{II}(\boldsymbol{\alpha})\boldsymbol{u}_{int} = \boldsymbol{I}(\boldsymbol{\alpha})\boldsymbol{L}(\boldsymbol{\alpha}) - \boldsymbol{A}_{IT}(\boldsymbol{\alpha})\boldsymbol{u}_t - \boldsymbol{A}_{IN}(\boldsymbol{\alpha})\boldsymbol{u}_n
$$

and

$$
(37) \qquad \begin{cases} \boldsymbol{0} \in \boldsymbol{A}_{tt}(\boldsymbol{\alpha})\boldsymbol{u}_t + \boldsymbol{A}_{tn}(\boldsymbol{\alpha})\boldsymbol{u}_n - \boldsymbol{L}_t(\boldsymbol{\alpha}) + \partial j(\boldsymbol{u}_t), \\ \boldsymbol{0} = \boldsymbol{A}_{nt}(\boldsymbol{\alpha})\boldsymbol{u}_t + \boldsymbol{A}_{nn}(\boldsymbol{\alpha})\boldsymbol{u}_n - \boldsymbol{\lambda} - \boldsymbol{L}_n(\boldsymbol{\alpha}), \\ \boldsymbol{0} \in -\boldsymbol{u}_n + \boldsymbol{\psi}(\boldsymbol{\alpha}) + N_{\mathbb{R}^p_+}(\boldsymbol{\lambda}), \end{cases}
$$

where

$$j(\boldsymbol{u}_t) := (\mathcal{F}(\boldsymbol{\varphi}) \bullet \boldsymbol{g}, |\boldsymbol{u}_t|)_p$$

and the symbols ∂, $N_{\mathbb{R}^p_+}$ denote the standard convex subdifferential and normal cone mapping, respectively. Finally, dropping (36) and inserting the fixed-point of Ξ into (37), we arrive at the *reduced state problem*

(38)
$$\begin{cases} \text{Find } (\boldsymbol{u}_n, \boldsymbol{u}_t, \boldsymbol{\lambda}) := (\boldsymbol{u}_n(\boldsymbol{\alpha}), \boldsymbol{u}_t(\boldsymbol{\alpha}), \boldsymbol{\lambda}(\boldsymbol{\alpha})) \in \mathbb{R}^{3p} \text{ such that} \\ \boldsymbol{0} \in \boldsymbol{A}_{tt}(\boldsymbol{\alpha})\boldsymbol{u}_t + \boldsymbol{A}_{tn}(\boldsymbol{\alpha})\boldsymbol{u}_n - \boldsymbol{L}_t(\boldsymbol{\alpha}) + Q_t(\boldsymbol{u}_t, \boldsymbol{\lambda}), \\ \boldsymbol{0} = \boldsymbol{A}_{nt}(\boldsymbol{\alpha})\boldsymbol{u}_t + \boldsymbol{A}_{nn}(\boldsymbol{\alpha})\boldsymbol{u}_n - \boldsymbol{\lambda} - \boldsymbol{L}_n(\boldsymbol{\alpha}), \\ \boldsymbol{0} \in -\boldsymbol{u}_n + \boldsymbol{\psi}(\boldsymbol{\alpha}) + N_{\mathbb{R}^p_+}(\boldsymbol{\lambda}). \end{cases}$$

Here the multivalued mapping $Q_t : \mathbb{R}^p \times \mathbb{R}^p \rightrightarrows \mathbb{R}^p$ takes the following form:

$$\big(Q_t(\boldsymbol{x}, \boldsymbol{z})\big)_i := \mathcal{F}(|x_i|)z_i\partial|x_i| \quad \forall i = 1, \ldots, p, \ \forall \boldsymbol{x}, \boldsymbol{z} \in \mathbb{R}^p.$$

As we already know, this system has a unique solution $(\boldsymbol{u}_t, \boldsymbol{u}_n, \boldsymbol{\lambda})$ which, together with $\boldsymbol{u}_{int} := \boldsymbol{u}_{int}(\boldsymbol{\alpha})$ computed from (36), gives the unique solution to $(\mathcal{M}(\boldsymbol{\alpha}))$, provided the assumptions of Theorem 3.4 are satisfied.

Introducing the *state vector* $\boldsymbol{y} := (\boldsymbol{u}_t, \boldsymbol{u}_n, \boldsymbol{\lambda})$ we may rewrite (38) in a more compact form

(39)
$$\begin{cases} \text{Find } \boldsymbol{y} \in \mathbb{R}^{3p} \text{ such that} \\ \boldsymbol{0} \in \boldsymbol{F}(\boldsymbol{\alpha}, \boldsymbol{y}) + Q(\boldsymbol{y}), \end{cases}$$

where for each $\boldsymbol{\alpha} \in \mathcal{U}_{ad}$ and $\boldsymbol{z} = (\boldsymbol{z}_1, \boldsymbol{z}_2, \boldsymbol{z}_3) \in (\mathbb{R}^p)^3$,

$$\boldsymbol{F}(\boldsymbol{\alpha}, \boldsymbol{z}) = \begin{bmatrix} \boldsymbol{A}_{tt}(\boldsymbol{\alpha}) & \boldsymbol{A}_{tn}(\boldsymbol{\alpha}) & \boldsymbol{0} \\ \boldsymbol{A}_{nt}(\boldsymbol{\alpha}) & \boldsymbol{A}_{nn}(\boldsymbol{\alpha}) & -\boldsymbol{E} \\ \boldsymbol{0} & -\boldsymbol{E} & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{z}_1 \\ \boldsymbol{z}_2 \\ \boldsymbol{z}_3 \end{bmatrix} - \begin{bmatrix} \boldsymbol{L}_t(\boldsymbol{\alpha}) \\ \boldsymbol{L}_n(\boldsymbol{\alpha}) \\ -\boldsymbol{\psi}(\boldsymbol{\alpha}) \end{bmatrix}, \quad Q(\boldsymbol{z}) = \begin{bmatrix} Q_t(\boldsymbol{z}_1, \boldsymbol{z}_3) \\ \boldsymbol{0} \\ N_{\mathbb{R}^p_+}(\boldsymbol{z}_3) \end{bmatrix},$$

and $\boldsymbol{E} \in \mathbb{R}^{p \times p}$ stands for the $(p \times p)$ identity matrix.

With the generalized equation (39) we associate the control-to-state mapping

$$S : \mathcal{U}_{ad} \ni \boldsymbol{\alpha} \mapsto \{\boldsymbol{y} \in \mathbb{R}^{3p} \mid \boldsymbol{0} \in \boldsymbol{F}(\boldsymbol{\alpha}, \boldsymbol{y}) + Q(\boldsymbol{y})\}$$

and denote by $\text{Gr } S$ its graph. By Theorem 3.4 we know that S is single-valued in $\mathcal{U}_{ad}$. Its Lipschitz continuity follows from the next result.

PROPOSITION 3.8. *Let the assumptions of Theorem 3.4 be satisfied. Then the generalized equation (39) is strongly regular at each* $(\boldsymbol{\alpha}, \boldsymbol{y}) \in \text{Gr } S$.

Proof. Let a pair $(\boldsymbol{\alpha}, \boldsymbol{y}) \in \text{Gr } S$ and a parameter $\boldsymbol{\xi} \in \mathbb{R}^{3p}$ be fixed and observe that $\boldsymbol{F}$ is linear in the second variable. Hence we need to verify that the mapping

(40)
$$\boldsymbol{\xi} \mapsto \{\boldsymbol{y}' \in \mathbb{R}^{3p} \mid \boldsymbol{\xi} \in \boldsymbol{F}(\boldsymbol{\alpha}, \boldsymbol{y}') + Q(\boldsymbol{y}')\}$$

is single-valued and Lipschitz in some neighborhood of $\boldsymbol{0}$. The generalized equation in (40), written componentwise for $\boldsymbol{y}' = (\boldsymbol{u}'_t, \boldsymbol{u}'_n, \boldsymbol{\lambda}')$, $\boldsymbol{\xi} = (\boldsymbol{\xi}_t, \boldsymbol{\xi}_n, \boldsymbol{\xi}_\lambda)$, reads

(41)
$$\begin{cases} \boldsymbol{\xi}_t \in \boldsymbol{A}_{tt}(\boldsymbol{\alpha})\boldsymbol{u}'_t + \boldsymbol{A}_{tn}(\boldsymbol{\alpha})\boldsymbol{u}'_n - \boldsymbol{L}_t(\boldsymbol{\alpha}) + Q_t(\boldsymbol{u}'_t, \boldsymbol{\lambda}'), \\ \boldsymbol{\xi}_n = \boldsymbol{A}_{nt}(\boldsymbol{\alpha})\boldsymbol{u}'_t + \boldsymbol{A}_{nn}(\boldsymbol{\alpha})\boldsymbol{u}'_n - \boldsymbol{\lambda}' - \boldsymbol{L}_n(\boldsymbol{\alpha}), \\ \boldsymbol{\xi}_\lambda \in -\boldsymbol{u}'_n + \boldsymbol{\psi}(\boldsymbol{\alpha}) + N_{\mathbb{R}^p_+}(\boldsymbol{\lambda}'). \end{cases}$$

Now, we rewrite the system (41) as follows:

(42)
$$\begin{cases} \text{Find } (\boldsymbol{u}'_t, \boldsymbol{u}'_n, \boldsymbol{\lambda}') \in \mathbb{R}^{3p} \text{ such that} \\ \boldsymbol{0} \in \boldsymbol{A}_{tt}(\boldsymbol{\alpha})\boldsymbol{u}'_t + \boldsymbol{A}_{tn}(\boldsymbol{\alpha})(\boldsymbol{u}'_n + \boldsymbol{\xi}_\lambda) - (\boldsymbol{L}_t(\boldsymbol{\alpha}) + \boldsymbol{\xi}_t + \boldsymbol{A}_{tn}(\boldsymbol{\alpha})\boldsymbol{\xi}_\lambda) + Q_t(\boldsymbol{u}'_t, \boldsymbol{\lambda}'), \\ \boldsymbol{0} = \boldsymbol{A}_{nt}(\boldsymbol{\alpha})\boldsymbol{u}'_t + \boldsymbol{A}_{nn}(\boldsymbol{\alpha})(\boldsymbol{u}'_n + \boldsymbol{\xi}_\lambda) - \boldsymbol{\lambda}' - (\boldsymbol{L}_n(\boldsymbol{\alpha}) + \boldsymbol{\xi}_n + \boldsymbol{A}_{nn}(\boldsymbol{\alpha})\boldsymbol{\xi}_\lambda), \\ \boldsymbol{0} \in -(\boldsymbol{u}'_n + \boldsymbol{\xi}_\lambda) + \boldsymbol{\psi}(\boldsymbol{\alpha}) + N_{\mathbb{R}^p_+}(\boldsymbol{\lambda}'). \end{cases}$$

Observe that (42) represents the Signorini problem with Coulomb friction and a solution-dependent coefficient of friction on the domain given by $\boldsymbol{\alpha} \in \mathcal{U}_{ad}$ and with load vector

(43)
$$\boldsymbol{L}_\xi = \begin{bmatrix} \boldsymbol{L}_t(\boldsymbol{\alpha}) + \boldsymbol{\xi}_t + \boldsymbol{A}_{tn}(\boldsymbol{\alpha})\boldsymbol{\xi}_\lambda \\ \boldsymbol{L}_n(\boldsymbol{\alpha}) + \boldsymbol{\xi}_n + \boldsymbol{A}_{nn}(\boldsymbol{\alpha})\boldsymbol{\xi}_\lambda \\ -\boldsymbol{\psi}(\boldsymbol{\alpha}) \end{bmatrix}$$

having the solution $\boldsymbol{y}' = (\boldsymbol{u}'_t, \boldsymbol{u}'_n + \boldsymbol{\xi}_\lambda, \boldsymbol{\lambda}')$. From Proposition 3.7 we know that for sufficiently small $\epsilon > 0$ and $\boldsymbol{\xi} \in \mathcal{V} := \mathbb{B}_\epsilon(\boldsymbol{0})$ contact problem $(\mathcal{M}(\boldsymbol{\alpha}))$ with the load vector $\boldsymbol{L}_\xi$ has exactly one solution, i.e., (41) is uniquely solvable. Thus the mapping defined in (40) is single-valued. To see that it is Lipschitz on $\mathcal{V}$ as well, let $\boldsymbol{\xi}^{(1)}, \boldsymbol{\xi}^{(2)} \in \mathcal{V}$ be arbitrary and denote the corresponding solutions of (41) by $\boldsymbol{y}^{(1)}$, $\boldsymbol{y}^{(2)}$. Then, employing Proposition 3.7 (here, $c > 0$ stands for a generic constant independent of $\boldsymbol{\xi}^{(i)}$, $\boldsymbol{y}^{(i)}$) we have

$$\begin{aligned} \|\boldsymbol{y}^{(1)} - \boldsymbol{y}^{(2)}\|_{3p} &\leq \|\boldsymbol{u}_t^{(1)} - \boldsymbol{u}_t^{(2)}\|_p + \|(\boldsymbol{u}_n^{(1)} + \boldsymbol{\xi}_\lambda^{(1)}) - (\boldsymbol{u}_n^{(2)} + \boldsymbol{\xi}_\lambda^{(2)})\|_p \\ &\quad + \|\boldsymbol{\lambda}^{(1)} - \boldsymbol{\lambda}^{(2)}\|_p + \|\boldsymbol{\xi}_\lambda^{(1)} - \boldsymbol{\xi}_\lambda^{(2)}\|_p \\ &\leq c\|\boldsymbol{L}_{\xi^{(1)}} - \boldsymbol{L}_{\xi^{(2)}}\|_{3p} + \|\boldsymbol{\xi}_\lambda^{(1)} - \boldsymbol{\xi}_\lambda^{(2)}\|_p \\ &\leq c\|\boldsymbol{\xi}^{(1)} - \boldsymbol{\xi}^{(2)}\|_{3p}, \end{aligned}$$

and the proof is complete. □

The next statements are immediate consequences of Proposition 3.8.

COROLLARY 3.9. *Let the assumptions of Theorem* 3.4 *be satisfied. Then*
(i) *S is Lipschitz in $\mathcal{U}_{ad}$;*
(ii) *for every continuous cost functional $J : \mathcal{U}_{ad} \times \mathbb{R}^{3p} \to \mathbb{R}$, the shape optimization problem*

$(\hat{\mathbb{P}})$ $$\begin{cases} \text{Find } \boldsymbol{\alpha}^* \in \mathcal{U}_{ad} \text{ such that} \\ \boldsymbol{\alpha}^* \in \operatorname{argmin}\{\mathcal{J}(\boldsymbol{\alpha}) := J(\boldsymbol{\alpha}, S(\boldsymbol{\alpha})) \mid \boldsymbol{\alpha} \in \mathcal{U}_{ad}\} \end{cases}$$

has at least one solution.

Proof. Assertion (i) follows from Theorem 2.1 in [6]; assertion (ii) is an easy consequence of (i) and the compactness of $\mathcal{U}_{ad}$. □

4. Sensitivity analysis. In what follows, we assume that a *continuously differentiable*[1] cost functional $J : \mathcal{U}_{ad} \times \mathbb{R}^{3p} \to \mathbb{R}$ is given. The composite cost functional $\mathcal{J}$, resulting from ImP is only locally Lipschitz. Therefore the minimization problem

---

[1]This smoothness assumption is unnecessarily strong and is imposed only to avoid unimportant technical issues in the presentation. From a theoretical point of view, a *locally Lipschitz J* would work as well.

$(\hat{\mathbb{P}})$ has to be solved by an algorithm of nondifferentiable optimization. Each evaluation of the function $\mathcal{J}$ is costly because it involves solving the contact problem (39). Hence, derivative-free methods are not preferred. To the bundle methods, however, which proved to be rather efficient in this context, one has to supply a subgradient $\boldsymbol{\xi} \in \overline{\partial}\mathcal{J}(\bar{\boldsymbol{\alpha}})$ at each step $\bar{\boldsymbol{\alpha}} \in \mathcal{U}_{ad}$. This section is devoted to its computation.

Let us start with the following observation. From the chain rule in [5, Theorem 2.6.6], Lipschitz continuity of S, and [17, Corollary 3.3.2], the following holds:

$$\overline{\partial}\mathcal{J}(\bar{\boldsymbol{\alpha}}) = \nabla_{\alpha}J(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}}) + \left(\overline{\partial}S(\bar{\boldsymbol{\alpha}})\right)^T \nabla_y J(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}})$$
$$\supset \nabla_{\alpha}J(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}}) + D^*S(\bar{\boldsymbol{\alpha}})(\nabla_y J(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}})).$$

Therefore, it is sufficient to determine an element $\boldsymbol{v}^* \in D^*S(\bar{\boldsymbol{\alpha}})(\nabla_y J(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}}))$ and to set $\boldsymbol{\xi} := \nabla_{\alpha}J(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}}) + \boldsymbol{v}^*$, which yields a desired Clarke's subgradient. The computation of one such $\boldsymbol{v}^*$ is described in the next theorem.

THEOREM 4.1. *Let $(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}}) \in \mathrm{Gr}\, S$ be given. Then for each $\boldsymbol{v}^* \in D^*S(\bar{\boldsymbol{\alpha}})(\nabla_y J(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}}))$ there exists an adjoint variable $\boldsymbol{p}^* \in \mathbb{R}^{3p}$ such that*

$$(44) \qquad\qquad\qquad \boldsymbol{v}^* = \nabla_{\alpha}F(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}})^T \boldsymbol{p}^*$$

and $\boldsymbol{p}^*$ is a solution to the adjoint GE *(AGE):*

$$(45) \qquad \boldsymbol{0} \in \nabla_y J(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}}) + \nabla_y F(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}})^T \boldsymbol{p}^* + D^*Q(\bar{\boldsymbol{y}}, -F(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}}))(\boldsymbol{p}^*).$$

Proof. Due to the strong regularity condition (see Proposition 3.8) the assumptions of [14, Theorem 5] are satisfied. See also [2, Theorem 4.1]. ∎

Note that Theorem 4.1, in general, provides only an upper approximation of $\overline{\partial}\mathcal{J}(\bar{\boldsymbol{\alpha}})$ since the vector $\boldsymbol{v}^*$ constructed using (44) and (45) may lie outside of $D^*S(\bar{\boldsymbol{\alpha}})(\nabla_y J(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}}))$. However, this can happen only at points where $\mathrm{Gr}\, Q$ is *not* normally regular, and even if it does happen (at a nonregular point), the used bundle method may not inevitably collapse. In this case a recovery step has to be made in which the bundle method is provided with a *correct* subgradient. Nevertheless, computational experience shows that this occurs very rarely; therefore we will rely on the construction of subgradients via the AGE (45) as described in Theorem 4.1.

The rest of this section is devoted to expressing the coderivative D^*Q in terms of the problem data, as D^*Q is the only unknown quantity remaining in (45). In doing so, we follow closely [2] and begin with reordering the equations in (39) so that $\boldsymbol{y} \in (\mathbb{R}^3)^p$, with $\boldsymbol{y}_i = ((\boldsymbol{u}_t)_i, (\boldsymbol{u}_n)_i, \lambda_i)^T$ comprising all state variables associated with the ith contact node ($i = 1, \ldots, p$). This way the multifunction Q takes the form

$$(46) \qquad\qquad\qquad Q(\boldsymbol{y}) = \begin{bmatrix} \Phi(\boldsymbol{y}_1) \\ \Phi(\boldsymbol{y}_2) \\ \vdots \\ \Phi(\boldsymbol{y}_p) \end{bmatrix},$$

where $\Phi : \mathbb{R}^2 \times \mathbb{R}_+ \rightrightarrows \mathbb{R}^3$ is defined as

$$(47) \qquad\qquad \Phi(\boldsymbol{a}) := \begin{bmatrix} \mathcal{F}(|a_1|)a_3 \partial |a_1| \\ 0 \\ N_{\mathbb{R}_+}(a_3) \end{bmatrix} \qquad \forall \boldsymbol{a} \in \mathbb{R}^2 \times \mathbb{R}_+.$$

Due to the above reordering (46) and [23, Example 6.10], one has for every $(\bar{\boldsymbol{y}}, \bar{\boldsymbol{q}}) \in$ $\operatorname{Gr} Q$ and $\boldsymbol{p}^* \in (\mathbb{R}^3)^p$

$$(48) \qquad D^* Q(\bar{\boldsymbol{y}}, \bar{\boldsymbol{q}})(\boldsymbol{p}^*) = \begin{bmatrix} D^* \Phi(\bar{\boldsymbol{y}}_1, \bar{\boldsymbol{q}}_1)(\boldsymbol{p}_1^*) \\ D^* \Phi(\bar{\boldsymbol{y}}_2, \bar{\boldsymbol{q}}_2)(\boldsymbol{p}_2^*) \\ \vdots \\ D^* \Phi(\bar{\boldsymbol{y}}_p, \bar{\boldsymbol{q}}_p)(\boldsymbol{p}_p^*) \end{bmatrix}.$$

Therefore, we will consider arbitrary $(\bar{\boldsymbol{a}}, \bar{\boldsymbol{b}}) \in \operatorname{Gr} \Phi$, $\boldsymbol{b}^* \in \mathbb{R}^3$ and compute the coderivative $D^* \Phi(\bar{\boldsymbol{a}}, \bar{\boldsymbol{b}})(\boldsymbol{b}^*)$ according to the position of $(\bar{\boldsymbol{a}}, \bar{\boldsymbol{b}})$ as given by the following partition of $\operatorname{Gr} \Phi$:

$$(49) \qquad \operatorname{Gr} \Phi = L \cup M_1 \cup M_2 \cup M_3^+ \cup M_3^- \cup M_4,$$

where the sets on the right-hand side of (49) are defined as in Table 4.1. From a mechanical point of view, partition (49) represents all possible contact and sliding modes of a point on the contact boundary.

TABLE 4.1
Contact and sliding modes of $(\boldsymbol{a}, \boldsymbol{b}) \in \operatorname{Gr} \Phi$.

	No contact: $a_3 = 0, \ b_3 < 0$	Weak contact: $a_3 = 0, \ b_3 = 0$	Strong contact: $a_3 > 0, \ b_3 = 0$
Sliding: $a_1 \neq 0,$ $b_1 = \operatorname{sgn}(a_1)\mathcal{F}(\|a_1\|)a_3$		M_2	M_1
Weak sticking: $a_1 = 0,$ $\|b_1\| = \mathcal{F}(0)a_3$	L	M_4	M_3^-
Strong sticking: $a_1 = 0,$ $\|b_1\| < \mathcal{F}(0)a_3$	$\times \times \times$	$\times \times \times$	M_3^+

As easily seen from their definitions, the sets L, M_1, and M_3^+ are open in the relative topology of $\operatorname{Gr} \Phi$; i.e., each $\Sigma \in \{L, M_1, M_3^+\}$ satisfies

$$(50) \qquad \forall (\bar{\boldsymbol{a}}, \bar{\boldsymbol{b}}) \in \Sigma \ \exists \text{ neighborhood } \mathcal{O}: \ \operatorname{Gr} \Phi \cap \mathcal{O} \subset \Sigma.$$

This makes the analysis in these cases substantially easier, since

$$(51) \qquad N_{\operatorname{Gr} \Phi}(\bar{\boldsymbol{a}}, \bar{\boldsymbol{b}}) = N_\Sigma(\bar{\boldsymbol{a}}, \bar{\boldsymbol{b}}) = \operatorname*{Lim\,sup}_{(\boldsymbol{a},\boldsymbol{b}) \xrightarrow{\Sigma} (\bar{\boldsymbol{a}},\bar{\boldsymbol{b}})} \widehat{N}_\Sigma(\boldsymbol{a}, \boldsymbol{b}),$$

as will be used frequently below.

PROPOSITION 4.2. *Let $(\bar{\boldsymbol{a}}, \bar{\boldsymbol{b}}) \in L$ and $\boldsymbol{b}^* \in \mathbb{R}^3$ be given. Then*

$$(52) \qquad D^* \Phi(\bar{\boldsymbol{a}}, \bar{\boldsymbol{b}})(\boldsymbol{b}^*) = \begin{cases} \{0\} \times \{0\} \times \mathbb{R} & \text{if } b_3^* = 0, \\ \emptyset & \text{otherwise.} \end{cases}$$

Proof. Let $(\boldsymbol{a}, \boldsymbol{b}) \in L$ be arbitrary. Then there exists a neighborhood $\mathcal{O}$ of $(\boldsymbol{a}, \boldsymbol{b})$ such that

$$\operatorname{Gr} \Phi \cap \mathcal{O} = (\mathbb{R} \times \mathbb{R} \times \{0\}) \times (\{0\} \times \{0\} \times \mathbb{R}) \cap \mathcal{O}.$$

Therefore,

$$\widehat{N}_{\mathrm{Gr}\,\Phi}(\boldsymbol{a},\boldsymbol{b}) = \big(\{0\}\times\{0\}\times\mathbb{R}\big)\times\big(\mathbb{R}\times\mathbb{R}\times\{0\}\big), \tag{53}$$

and the assertion follows directly from (51) and the definition of $D^*\Phi$. □

PROPOSITION 4.3. *Let* $(\bar{\boldsymbol{a}},\bar{\boldsymbol{b}})\in M_3^+$ *and* $\boldsymbol{b}^*\in\mathbb{R}^3$ *be given. Then*

$$D^*\Phi(\bar{\boldsymbol{a}},\bar{\boldsymbol{b}})(\boldsymbol{b}^*) = \begin{cases} \mathbb{R}\times\{0\}\times\{0\} & \text{if } b_1^* = 0, \\ \emptyset & \text{otherwise.} \end{cases} \tag{54}$$

Proof. In this case, for every $(\boldsymbol{a},\boldsymbol{b})\in M_3^+$ one can find a suitable neighborhood $\mathcal{O}$ such that

$$\mathrm{Gr}\,\Phi\cap\mathcal{O} = \big(\{0\}\times\mathbb{R}\times\mathbb{R}\big)\times\big(\mathbb{R}\times\{0\}\times\{0\}\big)\cap\mathcal{O},$$

whence

$$\widehat{N}_{\mathrm{Gr}\,\Phi}(\boldsymbol{a},\boldsymbol{b}) = \big(\mathbb{R}\times\{0\}\times\{0\}\big)\times\big(\{0\}\times\mathbb{R}\times\mathbb{R}\big). \tag{55}$$

The rest follows again from (51) and the definition of the coderivative. □

Convention. For convenience, in what follows $\mathcal{F}$ will signify the *even extension* of the coefficient of friction to the whole $\mathbb{R}$, i.e., $\mathcal{F}(x) := \mathcal{F}(-x)\ \forall x < 0$, so that $\mathcal{F}(|x|) = \mathcal{F}(x)\ \forall x\in\mathbb{R}$. Clearly, $\mathcal{F}$ is (globally) Lipschitz in $\mathbb{R}$.

PROPOSITION 4.4. *Let* $(\bar{\boldsymbol{a}},\bar{\boldsymbol{b}})\in M_1$ *and* $\boldsymbol{b}^*\in\mathbb{R}^3$ *be given. Then*

$$D^*\Phi(\bar{\boldsymbol{a}},\bar{\boldsymbol{b}})(\boldsymbol{b}^*) = \begin{bmatrix} D^*\mathcal{F}(\bar{a}_1)(\mathrm{sgn}(\bar{a}_1)\bar{a}_3 b_1^*) \\ 0 \\ \mathrm{sgn}(\bar{a}_1)\mathcal{F}(\bar{a}_1)b_1^* \end{bmatrix}. \tag{56}$$

Proof. There exists a neighborhood $\tilde{\mathcal{O}}$ of $\bar{\boldsymbol{a}}$ such that Φ is single-valued on $\tilde{\mathcal{O}}$ and equals

$$\Phi(\boldsymbol{a}) = \begin{bmatrix} \mathrm{sgn}(\bar{a}_1)\mathcal{F}(a_1)a_3 \\ 0 \\ 0 \end{bmatrix} \quad \forall\boldsymbol{a}\in\tilde{\mathcal{O}}.$$

From the definition of the regular coderivative,

$$\widehat{N}_{\mathrm{Gr}\,\Phi}(\boldsymbol{a},\Phi(\boldsymbol{a})) = \{(\boldsymbol{a}^*,\boldsymbol{b}^*)\in\mathbb{R}^3\times\mathbb{R}^3\ |$$
$$\langle\boldsymbol{a}^*,\boldsymbol{x}-\boldsymbol{a}\rangle_3 + \langle\boldsymbol{b}^*,\Phi(\boldsymbol{x})-\Phi(\boldsymbol{a})\rangle_3 \leq o(\|\boldsymbol{x}-\boldsymbol{a}\|)\ \forall\boldsymbol{x}\},$$

employing the Lipschitz continuity of $\mathcal{F}$. A straightforward calculation yields

$$\widehat{N}_{\mathrm{Gr}\,\Phi}(\boldsymbol{a},\Phi(\boldsymbol{a})) = \{(\boldsymbol{a}^*,\boldsymbol{b}^*)\ |\ a_2^* = 0,\ a_3^* = -\,\mathrm{sgn}(\bar{a}_1)\mathcal{F}(a_1)b_1^*, \tag{57}$$
$$(a_1^*,\mathrm{sgn}(\bar{a}_1)b_1^*a_3)\in\widehat{N}_{\mathrm{Gr}\,\mathcal{F}}(a_1,\mathcal{F}(a_1))\}.$$

Hence (see (51))

$$N_{\mathrm{Gr}\,\Phi}(\bar{\boldsymbol{a}},\bar{\boldsymbol{b}}) = \{(\boldsymbol{a}^*,\boldsymbol{b}^*)\ |\ a_2^* = 0,\ a_3^* = -\,\mathrm{sgn}(\bar{a}_1)\mathcal{F}(\bar{a}_1)b_1^*,$$
$$(a_1^*,\mathrm{sgn}(\bar{a}_1)b_1^*\bar{a}_3)\in N_{\mathrm{Gr}\,\mathcal{F}}(\bar{a}_1,\mathcal{F}(\bar{a}_1))\}$$

and the proof is complete. □

Remark 4.5.

(i) *If $\mathcal{F}$ happens to be smooth around $\bar{a}_1$, then Φ is smooth in $\tilde{\mathcal{O}}$ and (56) reduces to the adjoint Jacobian of Φ, as expected:*

$$D^*\Phi(\bar{a}, \bar{b})(b^*) = \operatorname{sgn}(\bar{a}_1) \begin{bmatrix} \mathcal{F}'(\bar{a}_1)\bar{a}_3 & 0 & 0 \\ 0 & 0 & 0 \\ \mathcal{F}(\bar{a}_1) & 0 & 0 \end{bmatrix} b^*.$$

(ii) *It can be seen from the proofs of Propositions 4.2 and 4.3 that $\operatorname{Gr}\Phi$ is normally regular at each point of L and M_3^+. It is normally regular at those points $(\bar{a}, \bar{b}) \in M_1$ for which $\operatorname{Gr}\mathcal{F}$ is normally regular at $(\bar{a}_1, \mathcal{F}(\bar{a}_1))$. In particular, if $\mathcal{F}$ is smooth, then $\operatorname{Gr}\Phi$ is normally regular also on M_1.*

Unfortunately, the situation becomes more involved when dealing with the sets M_2 and M_3^-, since they lie on the boundary of two open sets,

$$(58) \qquad M_2 = \operatorname{relint}(\partial L \cap \partial M_1) \quad \text{and} \quad M_3^- = \operatorname{relint}(\partial M_1 \cap \partial M_3^+),$$

where $\operatorname{relint}(A)$ denotes the relative interior of the set A.

In order to compute $D^*\Phi$ at points belonging to M_2, we will use a slightly generalized version of [2, Lemma 4.6]. In particular, we show that its assertion holds with equality under less restrictive conditions.

LEMMA 4.6. *Consider a multifunction $F : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^o \rightrightarrows \mathbb{R}^l \times \mathbb{R}^p$ given by*

$$F(x, y, z) = \begin{bmatrix} G(x, y) \\ H(y, z) \end{bmatrix},$$

where $G : \mathbb{R}^n \times \mathbb{R}^m \rightrightarrows \mathbb{R}^l$, $H : \mathbb{R}^m \times \mathbb{R}^o \rightrightarrows \mathbb{R}^p$ are closed-graph multifunctions. Assume that the point $(\bar{x}, \bar{y}, \bar{z}, \bar{f}_1, \bar{f}_2)$ belongs to $\operatorname{Gr} F$ and that the qualification condition

$$(59) \qquad \left. \begin{array}{l} \begin{bmatrix} 0 \\ w_2 \end{bmatrix} \in D^*G(\bar{x}, \bar{y}, \bar{f}_1)(0) \\[2mm] \begin{bmatrix} -w_2 \\ 0 \end{bmatrix} \in D^*H(\bar{y}, \bar{z}, \bar{f}_2)(0) \end{array} \right\} \Rightarrow w_2 = 0$$

holds. Then one has

$$(60) \quad D^*F(\bar{x}, \bar{y}, \bar{z}, \bar{f}_1, \bar{f}_2)(d_1^*, d_2^*) \subset \{(u_1, u_2 + v_1, v_2) \mid$$
$$(u_1, u_2) \in D^*G(\bar{x}, \bar{y}, \bar{f}_1)(d_1^*), \ (v_1, v_2) \in D^*H(\bar{y}, \bar{z}, \bar{f}_2)(d_2^*)\}.$$

Assume, in addition, that for each sequence $y^{(i)} \to \bar{y}$ and each $\eta \in D^*G(\bar{x}, \bar{y}, \bar{f}_1)(d_1^*)$ there exist sequences $\{x^{(i)}\}$, $\{f_1^{(i)}\}$ such that $(x^{(i)}, y^{(i)}, f_1^{(i)}) \xrightarrow{\operatorname{Gr} G} (\bar{x}, \bar{y}, \bar{f}_1)$ and $d_1^{*(i)} \to d_1^*$ such that

$$(61) \qquad \eta \in \operatorname*{Lim\,sup}_{i \to \infty} \widehat{D}^*G(x^{(i)}, y^{(i)}, f_1^{(i)})(d_1^{*(i)}).$$

Then (60) holds as equality.

Proof. The first assertion has already been proved in [2]. To prove the second, let η be an element of the right-hand side of (60), i.e.,

$$\eta = (u_1, u_2 + v_1, v_2),$$

for some $(\boldsymbol{u}_1, \boldsymbol{u}_2) \in D^* G(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}, \bar{\boldsymbol{f}}_1)(\boldsymbol{d}_1^*)$ and $(\boldsymbol{v}_1, \boldsymbol{v}_2) \in D^* H(\bar{\boldsymbol{y}}, \bar{\boldsymbol{z}}, \bar{\boldsymbol{f}}_2)(\boldsymbol{d}_2^*)$. Thus, there exist sequences $(\boldsymbol{y}^{(i)}, \boldsymbol{z}^{(i)}, \boldsymbol{f}_2^{(i)}) \xrightarrow{\mathrm{Gr}\, H} (\bar{\boldsymbol{y}}, \bar{\boldsymbol{z}}, \bar{\boldsymbol{f}}_2)$, $\boldsymbol{d}_2^{*(i)} \to \boldsymbol{d}_2^*$, $(\boldsymbol{v}_1^{(i)}, \boldsymbol{v}_2^{(i)}) \to (\boldsymbol{v}_1, \boldsymbol{v}_2)$ such that $(\boldsymbol{v}_1^{(i)}, \boldsymbol{v}_2^{(i)}) \in \widehat{D}^* H(\boldsymbol{x}^{(i)}, \boldsymbol{y}^{(i)}, \boldsymbol{f}_2^{(i)})(\boldsymbol{d}_2^{*(i)})$. By virtue of our additional assumption, there are sequences $\boldsymbol{x}^{(i)} \to \bar{\boldsymbol{x}}$, $\boldsymbol{f}_1^{(i)} \to \bar{\boldsymbol{f}}_1$, $\boldsymbol{d}_1^{*(i)} \to \boldsymbol{d}_1^*$, and $(\boldsymbol{u}_1^{(i)}, \boldsymbol{u}_2^{(i)}) \in \widehat{D}^* G(\boldsymbol{x}^{(i)}, \boldsymbol{y}^{(i)}, \boldsymbol{f}_1^{(i)})(\boldsymbol{d}_1^{*(i)})$ such that

$$(\boldsymbol{u}_1^{(i)}, \boldsymbol{u}_2^{(i)}) \to (\boldsymbol{u}_1, \boldsymbol{u}_2).$$

It follows from [23, Theorem 10.40] that for all $i \in \mathbb{N}$,

$$(\boldsymbol{u}_1^{(i)}, \boldsymbol{u}_2^{(i)} + \boldsymbol{v}_1^{(i)}, \boldsymbol{v}_2^{(i)}) \in \widehat{D}^* F(\boldsymbol{x}^{(i)}, \boldsymbol{y}^{(i)}, \boldsymbol{z}^{(i)}, \boldsymbol{f}_1^{(i)}, \boldsymbol{f}_2^{(i)})(\boldsymbol{d}_1^{*(i)}, \boldsymbol{d}_2^{*(i)}),$$

and consequently $\boldsymbol{\eta} \in D^* F(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}, \bar{\boldsymbol{z}}, \bar{\boldsymbol{f}}_1, \bar{\boldsymbol{f}}_2)(\boldsymbol{d}_1^*, \boldsymbol{d}_2^*)$. $\qquad\square$

Next we show that the second assumption of Lemma 4.6, ensuring equality in (60), is fulfilled in the following case:

$$(62) \qquad \boldsymbol{G} : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^l \quad \text{is given by} \quad \boldsymbol{G}(\boldsymbol{x}, \boldsymbol{y}) = f(\boldsymbol{x}) \boldsymbol{g}(\boldsymbol{y}),$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is *locally Lipschitz* and $\boldsymbol{g} : \mathbb{R}^m \to \mathbb{R}^l$ is *continuously differentiable*.

LEMMA 4.7. *Let $\boldsymbol{G}$ be as in* (62) *with $f : \mathbb{R}^n \to \mathbb{R}$ and $\boldsymbol{g} : \mathbb{R}^m \to \mathbb{R}^l$ Lipschitz around $\bar{\boldsymbol{x}} \in \mathbb{R}^n$ and $\bar{\boldsymbol{y}} \in \mathbb{R}^m$, respectively. Then*

$$(63) \qquad \widehat{D}^* \boldsymbol{G}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}})(\boldsymbol{d}^*) = \begin{bmatrix} \widehat{D}^* f(\bar{\boldsymbol{x}})(\boldsymbol{g}(\bar{\boldsymbol{y}})^T \boldsymbol{d}^*) \\ \widehat{D}^* \boldsymbol{g}(\bar{\boldsymbol{y}})(f(\bar{\boldsymbol{x}})\boldsymbol{d}^*) \end{bmatrix} \quad \forall \boldsymbol{d}^* \in \mathbb{R}^l.$$

Proof. From the definition of the regular coderivative we have

$$\widehat{D}^* \boldsymbol{G}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}})(\boldsymbol{d}^*) = \{(\boldsymbol{x}^*, \boldsymbol{y}^*) \in \mathbb{R}^n \times \mathbb{R}^m \mid$$
$$\langle \boldsymbol{x}^*, \boldsymbol{x} - \bar{\boldsymbol{x}} \rangle_n + \langle \boldsymbol{y}^*, \boldsymbol{y} - \bar{\boldsymbol{y}} \rangle_m - \langle \boldsymbol{d}^*, f(\boldsymbol{x})\boldsymbol{g}(\boldsymbol{y}) - f(\bar{\boldsymbol{x}})\boldsymbol{g}(\bar{\boldsymbol{y}}) \rangle_l$$
$$\leq o(\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|_n + \|\boldsymbol{y} - \bar{\boldsymbol{y}}\|_m) \quad \forall(\boldsymbol{x}, \boldsymbol{y})\}.$$

In particular, for $(\boldsymbol{x}, \bar{\boldsymbol{y}})$ and $(\bar{\boldsymbol{x}}, \boldsymbol{y})$ we get the following two relations:

$$(64) \qquad \langle \boldsymbol{x}^*, \boldsymbol{x} - \bar{\boldsymbol{x}} \rangle_n - \langle \boldsymbol{d}^*, (f(\boldsymbol{x}) - f(\bar{\boldsymbol{x}}))\boldsymbol{g}(\bar{\boldsymbol{y}}) \rangle_l \leq o(\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|_n) \quad \forall \boldsymbol{x},$$
$$(65) \qquad \langle \boldsymbol{y}^*, \boldsymbol{y} - \bar{\boldsymbol{y}} \rangle_m - \langle \boldsymbol{d}^*, f(\bar{\boldsymbol{x}})(\boldsymbol{g}(\boldsymbol{y}) - \boldsymbol{g}(\bar{\boldsymbol{y}})) \rangle_l \leq o(\|\boldsymbol{y} - \bar{\boldsymbol{y}}\|_m) \quad \forall \boldsymbol{y},$$

yielding the inclusion $\subset$ in (63). To prove the converse inclusion, let $\boldsymbol{x}^* \in \mathbb{R}^n$ and $\boldsymbol{y}^* \in \mathbb{R}^m$ satisfy (64) and (65), respectively. Summing both equations, one has

$$\langle \boldsymbol{x}^*, \boldsymbol{x} - \bar{\boldsymbol{x}} \rangle_n + \langle \boldsymbol{y}^*, \boldsymbol{y} - \bar{\boldsymbol{y}} \rangle_m - \langle \boldsymbol{d}^*, f(\boldsymbol{x})\boldsymbol{g}(\boldsymbol{y}) - f(\bar{\boldsymbol{x}})\boldsymbol{g}(\bar{\boldsymbol{y}}) \rangle_l$$
$$\leq \langle \boldsymbol{d}^*, (f(\boldsymbol{x}) - f(\bar{\boldsymbol{x}}))(\boldsymbol{g}(\boldsymbol{y}) - \boldsymbol{g}(\bar{\boldsymbol{y}})) \rangle_l + o(\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|_n) + o(\|\boldsymbol{y} - \bar{\boldsymbol{y}}\|_m).$$

It remains to realize that the right-hand side is $o(\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|_n + \|\boldsymbol{y} - \bar{\boldsymbol{y}}\|_m)$. Let us only show how its first term can be estimated:

$$\frac{\langle \boldsymbol{d}^*, (f(\boldsymbol{x}) - f(\bar{\boldsymbol{x}}))(\boldsymbol{g}(\boldsymbol{y}) - \boldsymbol{g}(\bar{\boldsymbol{y}})) \rangle_l}{\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|_n + \|\boldsymbol{y} - \bar{\boldsymbol{y}}\|_m}$$
$$\leq \|\boldsymbol{d}^*\|_l \frac{|f(\boldsymbol{x}) - f(\bar{\boldsymbol{x}})|}{\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|_n} \frac{\|\boldsymbol{g}(\boldsymbol{y}) - \boldsymbol{g}(\bar{\boldsymbol{y}})\|_l}{\|\boldsymbol{y} - \bar{\boldsymbol{y}}\|_m} \frac{\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|_n}{\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|_n + \|\boldsymbol{y} - \bar{\boldsymbol{y}}\|_m} \|\boldsymbol{y} - \bar{\boldsymbol{y}}\|_m.$$

Exploiting the Lipschitz continuity of f and $\boldsymbol{g}$ we conclude that the first four factors in the expression on the right-hand side are bounded, whereas the last term on the right converges to zero if $(\boldsymbol{x}, \boldsymbol{y}) \to (\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}})$. $\quad\square$

LEMMA 4.8. *Let the assumptions of Lemma 4.7 hold, with $\boldsymbol{g} : \mathbb{R}^m \to \mathbb{R}^l$ continuously differentiable around $\bar{\boldsymbol{y}} \in \mathbb{R}^m$. Then $\boldsymbol{G}$ satisfies (61), i.e., $\forall \boldsymbol{\eta} \in D^* \boldsymbol{G}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}})(\boldsymbol{d}^*)$ $\forall \boldsymbol{y}^{(i)} \to \bar{\boldsymbol{y}} \; \exists \boldsymbol{x}^{(i)} \to \bar{\boldsymbol{x}} \; \exists \boldsymbol{d}^{(i)} \to \boldsymbol{d}^* \; \exists \boldsymbol{\eta}^{(i)} \in \widehat{D}^* \boldsymbol{G}(\boldsymbol{x}^{(i)}, \boldsymbol{y}^{(i)})(\boldsymbol{d}^{(i)}) : \boldsymbol{\eta}^{(i)} \to \boldsymbol{\eta}$.*

Proof. Let $\boldsymbol{\eta} \in D^* \boldsymbol{G}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}})(\boldsymbol{d}^*)$ and $\boldsymbol{y}^{(i)} \to \bar{\boldsymbol{y}}$ be arbitrary. From the scalarization formula and [18, Corollary 1.111(i)] it follows easily that

$$(66) \qquad \boldsymbol{\eta} = \begin{bmatrix} \boldsymbol{\pi} \\ f(\bar{\boldsymbol{x}})\nabla\boldsymbol{g}(\bar{\boldsymbol{y}})^T \boldsymbol{d}^* \end{bmatrix} \quad \text{for some} \quad \boldsymbol{\pi} \in D^* f(\bar{\boldsymbol{x}})(\boldsymbol{g}(\bar{\boldsymbol{y}})^T \boldsymbol{d}^*).$$

By the definition of the (limiting) coderivative,

$$(67) \qquad \exists \boldsymbol{x}^{(i)} \to \bar{\boldsymbol{x}} \; \exists r^{(i)} \to \boldsymbol{g}(\bar{\boldsymbol{y}})^T \boldsymbol{d}^* \; \exists \boldsymbol{\pi}^{(i)} \in \widehat{D}^* f(\boldsymbol{x}^{(i)})(r^{(i)}) : \quad \boldsymbol{\pi}^{(i)} \to \boldsymbol{\pi}.$$

Let us distinguish between the following two situations.

(i) $\boldsymbol{g}(\bar{\boldsymbol{y}})^T \boldsymbol{d}^* \neq 0$. Then, clearly, $\boldsymbol{g}(\boldsymbol{y}^{(i)}) \neq \boldsymbol{0}$ for i sufficiently large. For these indices we may select any sequence $\{\boldsymbol{d}^{(i)}\}$ satisfying the conditions

$$\boldsymbol{d}^{(i)} \to \boldsymbol{d}^* \quad \text{and} \quad \boldsymbol{g}(\boldsymbol{y}^{(i)})^T \boldsymbol{d}^{(i)} = r^{(i)}.$$

Observe that such choice of $\{\boldsymbol{d}^{(i)}\}$ is always possible. By Lemma 4.7,

$$(68) \qquad \boldsymbol{\eta}^{(i)} := \begin{bmatrix} \boldsymbol{\pi}^{(i)} \\ f(\boldsymbol{x}^{(i)})\nabla\boldsymbol{g}(\boldsymbol{y}^{(i)})^T \boldsymbol{d}^{(i)} \end{bmatrix} \in \widehat{D}^* \boldsymbol{G}(\boldsymbol{x}^{(i)}, \boldsymbol{y}^{(i)})(\boldsymbol{d}^{(i)})$$

and so the assertion follows.

(ii) $\boldsymbol{g}(\bar{\boldsymbol{y}})^T \boldsymbol{d}^* = 0$. It follows that $\boldsymbol{\pi} = \boldsymbol{0}$, since $D^* f(\bar{\boldsymbol{x}})(\boldsymbol{g}(\bar{\boldsymbol{y}})^T \boldsymbol{d}^*) = \{\boldsymbol{0}\}$ by virtue of the Mordukhovich criterion [23, Theorem 9.40]. Consider now arbitrary sequences $\boldsymbol{x}^{(i)} \to \bar{\boldsymbol{x}}$, $\boldsymbol{d}^{(i)} \to \boldsymbol{d}^*$, and $\boldsymbol{\pi}^{(i)} \in \widehat{D}^* f(\boldsymbol{x}^{(i)})(\boldsymbol{g}(\boldsymbol{y}^{(i)})^T \boldsymbol{d}^{(i)}) = D^* f(\boldsymbol{x}^{(i)})(\boldsymbol{g}(\boldsymbol{y}^{(i)})^T \boldsymbol{d}^{(i)}) \neq \emptyset$. Such sequences do exist because f is differentiable on a dense subset of its domain (Rademacher theorem) and at these points

$$\widehat{D}^* f(\boldsymbol{x}^{(i)})(r) = D^* f(\boldsymbol{x}^{(i)})(r) \neq \emptyset \quad \forall r \in \mathbb{R}.$$

Clearly, $\boldsymbol{\pi}^{(i)} \to \boldsymbol{0}$ by the outer semicontinuity of the limiting coderivative, and the statement follows again from Lemma 4.7. $\quad\square$

PROPOSITION 4.9. *Let $(\bar{\boldsymbol{a}}, \bar{\boldsymbol{b}}) \in M_2$ and $\boldsymbol{b}^* \in \mathbb{R}^3$ be given. Then*

$$(69) \qquad D^*\Phi(\bar{\boldsymbol{a}}, \bar{\boldsymbol{b}})(\boldsymbol{b}^*) = \left\{ \begin{bmatrix} 0 \\ 0 \\ \text{sgn}(\bar{a}_1)\mathcal{F}(\bar{a}_1)b_1^* + w \end{bmatrix} \middle| \, w \in \begin{cases} \mathbb{R} & \text{if } b_3^* = 0, \\ \mathbb{R}_- & \text{if } b_3^* < 0, \\ \{0\} & \text{if } b_3^* > 0. \end{cases} \right\}$$

Proof. Consider a reference point $(\bar{\boldsymbol{a}}, \bar{\boldsymbol{b}}) = (\bar{a}_1, \bar{a}_2, 0, 0, 0, 0) \in M_2$, where $\bar{a}_1 \neq 0$ by the definition of M_2. Then Φ attains the form

$$\Phi(\boldsymbol{a}) = \begin{bmatrix} \text{sgn}(\bar{a}_1)\mathcal{F}(a_1)a_3 \\ 0 \\ N_{\mathbb{R}_+}(a_3) \end{bmatrix} \quad \forall \boldsymbol{a} \in \tilde{\mathcal{O}}$$

for a sufficiently small neighborhood $\tilde{\mathcal{O}}$ of $\bar{a}$. Defining the function $G(x, y) := \mathcal{F}(x)g(y)$, where $g(y) := \mathrm{sgn}(\bar{a}_1)y$ and the closed-graph multifunction $H(y) = N_{\mathbb{R}_+}(y)$, Lemma 4.6 yields

$$(70) \qquad D^*\Phi(\bar{a}, \bar{b})(b^*) = \{(u_1, 0, u_2 + v) \mid (u_1, u_2) \in D^*G(\bar{a}_1, 0)(b_1^*),$$
$$v \in D^*H(0, 0)(b_3^*)\}$$

because G is of the form (62), and thus the second assumption of Lemma 4.6 is satisfied. Since $g(0) = 0$ and $g'(0) = \mathrm{sgn}(\bar{a}_1)$, it follows from (66) that

$$(71) \qquad D^*G(\bar{a}_1, 0)(b_1^*) = \left\{ \begin{bmatrix} 0 \\ \mathrm{sgn}(\bar{a}_1)\mathcal{F}(\bar{a}_1)b_1^* \end{bmatrix} \right\}.$$

For the coderivative of the normal cone mapping H at $(0, 0) \in \mathrm{Gr}\, H$ one has

$$(72) \qquad D^*H(0, 0)(b_3^*) = \begin{cases} \mathbb{R} & \text{if } b_3^* = 0, \\ \mathbb{R}_- & \text{if } b_3^* < 0, \\ \{0\} & \text{if } b_3^* > 0. \end{cases}$$

Finally, the assertion follows by collecting (70), (71), and (72). $\qquad\blacksquare$

In order to give a formula for the coderivative $D^*\Phi$ at points in M_3^- we will, in addition, assume that the coefficient of friction $\mathcal{F}$ is *weakly semismooth* at 0 (cf. [16]), implying that

$$(73) \qquad \exists \mathcal{F}_+'(0) \in \mathbb{R} \quad \text{and} \quad \limsup_{x \to 0_+} \overline{\partial}\mathcal{F}(x) = \{\mathcal{F}_+'(0)\},$$

where $\mathcal{F}_+'$ stands for the right-hand derivative of $\mathcal{F}$. Now the following result holds.

PROPOSITION 4.10. *Let* $(\bar{a}, \bar{b}) \in M_3^-$ *and* $b^* \in \mathbb{R}^3$ *be given. Then*

$$(74) \quad D^*\Phi(\bar{a}, \bar{b})(b^*) = \left\{ \begin{bmatrix} \mathcal{F}_+'(0)\bar{a}_3 b_1^* + w \\ 0 \\ \mathrm{sgn}(\bar{b}_1)\mathcal{F}(0)b_1^* \end{bmatrix} \middle| \, w \in \begin{cases} \mathbb{R} & \text{if } b_1^* = 0, \\ \mathrm{sgn}(\bar{b}_1)\mathbb{R}_+ & \text{if } b_1^* \, \mathrm{sgn}(\bar{b}_1) < 0, \\ \{0\} & \text{otherwise.} \end{cases} \right\}$$

Proof. Let $(\bar{a}, \bar{b}) \in M_3^-$ be given, i.e., $(\bar{a}, \bar{b}) = (0, \bar{a}_2, \bar{a}_3, \bar{b}_1, 0, 0) \in \mathbb{R}^3 \times \mathbb{R}^3$, where $\bar{a}_3 > 0$ and $|\bar{b}_1| = \mathcal{F}(0)\bar{a}_3$. It can be easily seen that there exists a neighborhood $\mathcal{O}$ of $(\bar{a}, \bar{b})$ such that

$$(75) \qquad \mathrm{sgn}(b_1) = \mathrm{sgn}(\bar{b}_1) \quad \text{and} \quad \mathrm{sgn}(a_1)\,\mathrm{sgn}(\bar{b}_1) \geq 0 \quad \forall (a, b) \in \mathrm{Gr}\, \Phi \cap \mathcal{O}.$$

Moreover (cf. (58) and Table 4.1)

$$(76) \qquad N_{\mathrm{Gr}\, \Phi}(\bar{a}, \bar{b}) = \mathcal{N}_1 \cup \mathcal{N}_2 \cup \mathcal{N}_3,$$

where

$$\mathcal{N}_1 := \limsup_{(a,b) \xrightarrow{M_1} (\bar{a}, \bar{b})} \widehat{N}_{M_1}(a, b),$$

$$\mathcal{N}_2 := \limsup_{(a,b) \xrightarrow{M_3^+} (\bar{a}, \bar{b})} \widehat{N}_{M_3^+}(a, b),$$

$$\mathcal{N}_3 := \limsup_{(a,b) \xrightarrow{M_3^-} (\bar{a}, \bar{b})} \widehat{N}_{\mathrm{Gr}\, \Phi}(a, b).$$

Let us first calculate $\mathcal{N}_1$. From (57), (75), and the definition of the regular coderivative it follows that

$$(77) \quad \widehat{N}_{M_1}(\boldsymbol{a}, \boldsymbol{b}) = \{(\boldsymbol{x}^*, \boldsymbol{y}^*) \mid x_2^* = 0, \ x_3^* = -\operatorname{sgn}(\bar{b}_1)\mathcal{F}(0)y_1^*,$$
$$x_1^* \in \widehat{D}^*\mathcal{F}(a_1)(-\operatorname{sgn}(\bar{b}_1)a_3y_1^*)\}$$

for each $(\boldsymbol{a}, \boldsymbol{b}) \in M_1$. Using the scalarization formula and [17, Corollary 3.3.2] we get

$$(78) \quad \widehat{D}^*\mathcal{F}(a_1)(-\operatorname{sgn}(\bar{b}_1)a_3y_1^*) \subset D^*\mathcal{F}(a_1)(-\operatorname{sgn}(\bar{b}_1)a_3y_1^*)$$
$$= \partial(-\operatorname{sgn}(\bar{b}_1)a_3y_1^*\mathcal{F})(a_1) \subset -\operatorname{sgn}(\bar{b}_1)a_3y_1^*\overline{\partial}\mathcal{F}(a_1).$$

Note, that $\mathcal{N}_1$ is nonempty (it follows easily from the Lipschitz continuity of $\mathcal{F}$ and the Rademacher theorem). In light of this fact, (77) and (78), together with the semismoothness assumption (73) and (75), yield

$$(79) \quad \mathcal{N}_1 = \{(\boldsymbol{a}^*, \boldsymbol{b}^*) \mid a_2^* = 0, \ a_3^* = -\operatorname{sgn}(\bar{b}_1)\mathcal{F}(0)b_1^*,$$
$$a_1^* = -\mathcal{F}'_+(0)\bar{a}_3b_1^*\}.$$

Concerning $\mathcal{N}_2$, from (55) one has immediately

$$(80) \quad \mathcal{N}_2 = (\mathbb{R} \times \{0\} \times \{0\}) \times (\{0\} \times \mathbb{R} \times \mathbb{R}).$$

However, the computation of the cone $\mathcal{N}_3$ is more involved. In particular, let $(\boldsymbol{a}, \boldsymbol{b}) \in M_3^-$ be given, and observe that $\operatorname{Gr} \Phi$ locally around $(\boldsymbol{a}, \boldsymbol{b})$ can be written as the union of the following two disjoint sets (cf. Table 4.1 and (75)):

$$G_1 := \{(\boldsymbol{x}, \boldsymbol{y}) \mid \operatorname{sgn}(x_1) = \operatorname{sgn}(\bar{b}_1), \ x_3 > 0, \ y_1 = \operatorname{sgn}(\bar{b}_1)\mathcal{F}(x_1)x_3, \ y_2 = y_3 = 0\},$$
$$G_2 := \{(\boldsymbol{x}, \boldsymbol{y}) \mid x_1 = 0, \ x_3 > 0, \ \operatorname{sgn}(\bar{b}_1)y_1 \leq \mathcal{F}(0)x_3, \ y_2 = y_3 = 0\}.$$

This way one has

$$(81) \quad T_{\operatorname{Gr} \Phi}(\boldsymbol{a}, \boldsymbol{b}) = T_{G_1}(\boldsymbol{a}, \boldsymbol{b}) \cup T_{G_2}(\boldsymbol{a}, \boldsymbol{b}),$$

and hence,

$$(82) \quad \widehat{N}_{\operatorname{Gr} \Phi}(\boldsymbol{a}, \boldsymbol{b}) = (T_{\operatorname{Gr} \Phi}(\boldsymbol{a}, \boldsymbol{b}))^0 = \widehat{N}_{G_1}(\boldsymbol{a}, \boldsymbol{b}) \cap \widehat{N}_{G_2}(\boldsymbol{a}, \boldsymbol{b}).$$

The contingent cone to G_1 can be determined as follows:

$$T_{G_1}(\boldsymbol{a}, \boldsymbol{b}) = \{(\boldsymbol{h}, \boldsymbol{k}) \mid \exists \boldsymbol{h}^{(i)} \to \boldsymbol{h}, \ \boldsymbol{k}^{(i)} \to \boldsymbol{k}, \ \lambda^{(i)} \to 0_+, \ \forall i :$$
$$(\boldsymbol{a} + \lambda^{(i)}\boldsymbol{h}^{(i)}, \boldsymbol{b} + \lambda^{(i)}\boldsymbol{k}^{(i)}) \in G_1\}$$
$$= \{(\boldsymbol{h}, \boldsymbol{k}) \mid \exists \boldsymbol{h}^{(i)} \to \boldsymbol{h}, \ \boldsymbol{k}^{(i)} \to \boldsymbol{k}, \ \lambda^{(i)} \to 0_+, \ \forall i :$$
$$\operatorname{sgn}(\lambda^{(i)}h_1^{(i)}) = \operatorname{sgn}(\bar{b}_1), \ a_3 + \lambda^{(i)}h_3^{(i)} > 0,$$
$$\operatorname{sgn}(\bar{b}_1)\mathcal{F}(0)a_3 + \lambda^{(i)}k_1^{(i)} = \operatorname{sgn}(\bar{b}_1)\mathcal{F}(\lambda^{(i)}h_1^{(i)})(a_3 + \lambda^{(i)}h_3^{(i)}),$$
$$\lambda^{(i)}k_2^{(i)} = 0, \ \lambda^{(i)}k_3^{(i)} = 0\},$$

from which

$$k_1^{(i)} = \operatorname{sgn}(\bar{b}_1)\frac{\mathcal{F}(\lambda^{(i)}h_1^{(i)}) - \mathcal{F}(0)}{\lambda^{(i)}h_1^{(i)}}h_1^{(i)}a_3 + \operatorname{sgn}(\bar{b}_1)\mathcal{F}(\lambda^{(i)}h_1^{(i)})h_3^{(i)}$$
$$= \frac{\mathcal{F}(\lambda^{(i)}|h_1^{(i)}|) - \mathcal{F}(0)}{\lambda^{(i)}|h_1^{(i)}|}h_1^{(i)}a_3 + \operatorname{sgn}(\bar{b}_1)\mathcal{F}(\lambda^{(i)}h_1^{(i)})h_3^{(i)}$$
$$\longrightarrow \mathcal{F}'_+(0)h_1a_3 + \operatorname{sgn}(\bar{b}_1)\mathcal{F}(0)h_3 \quad \text{for} \quad i \to \infty,$$

as follows from (73). Thus we get

$$
(83) \qquad T_{G_1}(\boldsymbol{a}, \boldsymbol{b}) = \{(\boldsymbol{h}, \boldsymbol{k}) \mid \operatorname{sgn}(\bar{b}_1)h_1 \geq 0, \ k_2 = k_3 = 0,
$$
$$
k_1 = \mathcal{F}'_+(0)a_3 h_1 + \operatorname{sgn}(\bar{b}_1)\mathcal{F}(0)h_3\}.
$$

An analogous computation yields

$$
(84) \qquad T_{G_2}(\boldsymbol{a}, \boldsymbol{b}) = \{(\boldsymbol{h}, \boldsymbol{k}) \mid h_1 = 0, \ k_2 = k_3 = 0, \ \operatorname{sgn}(\bar{b}_1)k_1 \leq \mathcal{F}(0)h_3\}.
$$

Now, the negative polars to the cones (83), (84) can be easily calculated as

$$
\widehat{N}_{G_1}(\boldsymbol{a}, \boldsymbol{b}) = \{(\boldsymbol{x}^*, \boldsymbol{y}^*) \mid (x_1^* + \mathcal{F}'_+(0)a_3 y_1^*)\operatorname{sgn}(\bar{b}_1) \leq 0,
$$
$$
x_2^* = 0, \ x_3^* = -\operatorname{sgn}(\bar{b}_1)\mathcal{F}(0)y_1^*\}
$$

and

$$
\widehat{N}_{G_2}(\boldsymbol{a}, \boldsymbol{b}) = \{(\boldsymbol{x}^*, \boldsymbol{y}^*) \mid x_2^* = 0, \ x_3^* = -\operatorname{sgn}(\bar{b}_1)\mathcal{F}(0)y_1^*, \ y_1^*\operatorname{sgn}(\bar{b}_1) \geq 0\},
$$

so that

$$
(85) \qquad \widehat{N}_{G_1}(\boldsymbol{a}, \boldsymbol{b}) \cap \widehat{N}_{G_2}(\boldsymbol{a}, \boldsymbol{b}) = \{(\boldsymbol{x}^*, \boldsymbol{y}^*) \mid (x_1^* + \mathcal{F}'_+(0)a_3 y_1^*)\operatorname{sgn}(\bar{b}_1) \leq 0,
$$
$$
x_2^* = 0, \ x_3^* = -\operatorname{sgn}(\bar{b}_1)\mathcal{F}(0)y_1^*, \ y_1^*\operatorname{sgn}(\bar{b}_1) \geq 0\}.
$$

Finally, from (82) and (85) we get

$$
(86) \qquad \mathcal{N}_3 = \{(\boldsymbol{a}^*, \boldsymbol{b}^*) \mid (a_1^* + \mathcal{F}'_+(0)\bar{a}_3 b_1^*)\operatorname{sgn}(\bar{b}_1) \leq 0,
$$
$$
a_2^* = 0, \ a_3^* = -\operatorname{sgn}(\bar{b}_1)\mathcal{F}(0)b_1^*, \ b_1^*\operatorname{sgn}(\bar{b}_1) \geq 0\}.
$$

The assertion of the proposition follows now from (79), (80), (86), and the definition of the coderivative. □

5. Numerical results. The theoretical results on sensitivity analysis proven in the previous section now will be used for solving two model examples. Cost functionals will be continuously differentiable so that the resulting composite function $\mathcal{J}$ to be minimized is locally Lipschitz. Therefore one can use the implicit programming approach (cf. [21]) to solve the shape optimization problem $(\hat{\mathbb{P}})$. Minimization itself was realized by the MATLAB implementation of the bundle trust method due to Schramm and Zowe [24]. Each step of this method requires the value of $\mathcal{J}$ and one arbitrary Clarke's subgradient at the current point (for details see [5]). The latter has been discussed in the foregoing section. Let us briefly comment on the former issue. To get the value of $\mathcal{J}$ one has to solve the state problem $(\mathcal{M}(\boldsymbol{\alpha}))$ for each admissible $\boldsymbol{\alpha}$. Since solutions to this problem are defined by means of fixed-points of a certain mapping (see section 3), the method of successive approximations is used as a natural approach. Each iterative step is represented by the Signorini problem with given friction and a given coefficient of friction updated from the previous step. This is equivalent to the minimization problem for a quadratic function augmented by a nondifferentiable sublinear term over a convex set defined by linear inequality constraints. Computations were performed by the MatSol library [15] developed in the MATLAB environment.

The setting of the state problem is the same for both model examples. The geometry is shown in Figure 1: the body Ω is represented by a "rectangle" having one curved side Γ_c along which Ω is supported by the half plane $\mathcal{H} := \mathbb{R}^2_- = \{(x_1, x_2) \mid$
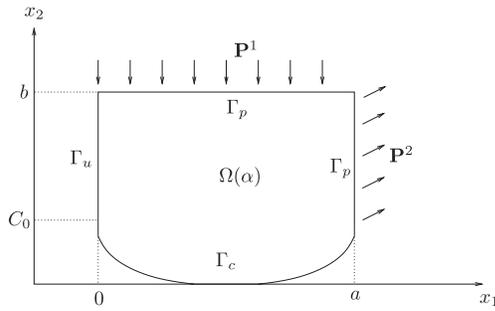
FIG. 1. *The elastic body and applied loads.*
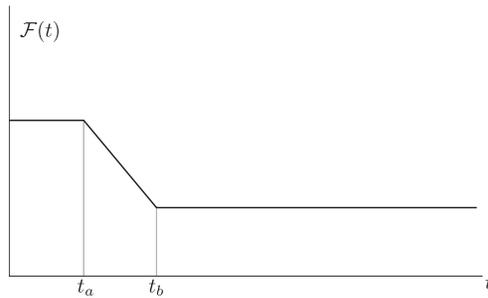
FIG. 2. *The definition of the friction coefficient.*

$x_2 \leq 0$}. We suppose that Γ_c is the graph of a nonnegative Lipschitz function α : $[0, a] \rightarrow \mathbb{R}^1_+$ which determines the final shape of Ω. To emphasize that Γ_c and Ω depend on a particular choice of α, we will write $\Gamma_c(\alpha)$ and $\Omega(\alpha)$, where

$$\Gamma_c(\alpha) = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_2 = \alpha(x_1), \ x_1 \in (0, a)\},$$
$$\Omega(\alpha) = \{(x_1, x_2) \in \mathbb{R}^2 \mid \alpha(x_1) < x_2 < b, \ x_1 \in (0, a)\}$$

and $0 < a$, $0 < b$ are given. In our examples we use $a = 2$, $b = 1$. The body is fixed along the left vertical side Γ_u, while surface tractions of density $\boldsymbol{P}^1$, $\boldsymbol{P}^2$ act on the top and the right vertical side, respectively. Their exact specifications will be done in each example. The body forces are neglected. Finally, Ω is made of a homogeneous, isotropic material characterized by the Young modulus $E = 1$ GPa, the Poisson constant $\sigma = 0.3$. The coefficient of friction $\mathcal{F}$ (see Figure 2) is defined by

$$\mathcal{F}(t) := \begin{cases} 0.25 & \text{for } t \in [0, 0.01], \\ 0.25 \cdot (-60t + 1.6) & \text{for } t \in (0.01, 0.02), \\ 0.1 & \text{for } t \in [0.02, \infty). \end{cases}$$

Due to the particular shape of the foundation $\mathcal{H}$ and the parametrization of $\Gamma_c(\alpha)$, the nonpenetration conditions (6) can be written in the following form:

$$u_2 \geq -\alpha, \quad T_2(u) \geq 0, \quad T_2(u)(u_2 + \alpha) = 0 \quad \text{on } \Gamma_c(\alpha),$$

where $\boldsymbol{u}_i$, $T_i(\boldsymbol{u})$, $i = 1, 2$, are the ith component of the displacement vector $\boldsymbol{u}$, and the stress vector $\boldsymbol{T}(\boldsymbol{u})$, respectively. In a similar way one can modify the friction

conditions (7) by setting $u_t := u_1$, $T_t(\boldsymbol{u}) := \boldsymbol{T}_1(\boldsymbol{u})$. This means that instead of u_n, u_t, $T_n(\boldsymbol{u})$, $T_t(\boldsymbol{u})$ which depend explicitly on α, we use u_i, $T_i(\boldsymbol{u})$, $i = 1, 2$, which do not. This fact simplifies computations.

Now we describe the discretization of the shape optimization problem ($\mathbb{P}$). Contact part Γ_c, which is the object of optimization, is modeled by Bézier functions ξ_α of order d on $[0, a]$, i.e., $\xi_\alpha = \sum_{i=0}^{d} \alpha_i B_d^{(i)}$, where $B_d^{(i)}$, $i = 0, \ldots, d$, are the Bernstein polynomials of order d:

$$B_d^{(i)}(x) = \frac{1}{a^d} \binom{d}{i} x^i (a - x)^{d-i}, \quad i = 0, \ldots, d, \quad x \in [0, a],$$

and the coefficients α_i of the linear combination are the x_2-coordinates of control points $\{A_i\}_{i=0}^d$ of ξ_α, $A_i = (ih, \alpha_i)$, $i = 0, 1, \ldots, d$, $h = a/d$. The shape of Γ_c (and hence of Ω) is uniquely determined by the vector $\boldsymbol{\alpha} = (\alpha_0, \ldots, \alpha_d)$. The discretization of the state problem uses a polygonal approximation $\Omega_h(\boldsymbol{\alpha})$ of $\Omega(\boldsymbol{\alpha})$ obtained by a piecewise linear approximation ξ_α^h of ξ_α on a partition of $[0, a]$. The finite-dimensional space $\boldsymbol{V}_h(\boldsymbol{\alpha})$ consists of Q_1-isoparametric elements considered on a partition of $\Omega_h(\boldsymbol{\alpha})$ into quadrilaterals.

The admissible set $\mathcal{U}_{ad}$ of the discrete design variables is defined as follows:

$$\begin{aligned}
\mathcal{U}_{ad} = \{\boldsymbol{\alpha} \in \mathbb{R}^{d+1} \mid\ & 0 \leq \alpha_i \leq C_0, \ i = 0, 1, \ldots, d; \\
& |\alpha_{i+1} - \alpha_i| \leq C_1 h, \ i = 0, 1, \ldots, d - 1; \\
& |\alpha_{i+1} - 2\alpha_i + \alpha_{i-1}| \leq C_2 h^2, \ i = 1, 2, \ldots, d - 1; \\
& C_{31} \leq \operatorname{meas} \Omega_h(\boldsymbol{\alpha}) \leq C_{32}\},
\end{aligned}$$

where C_0, C_1, C_2, C_{31}, and C_{32} are given positive constants chosen in such a way that $\mathcal{U}_{ad} \neq \emptyset$. The second and third inequality constraints control the first and second derivatives of ξ_α. To avoid trivial solutions, the volume constraint imposed on $\Omega_h(\boldsymbol{\alpha})$ is present. In computations, the constants in $\mathcal{U}_{ad}$ are specified as follows: $C_0 = 0.75$, $C_1 = 3$, $C_2 = 10$, $C_{31} = 1.8$, $C_{32} = 2$. The total number of the nodes of $\mathcal{T}_h$ $\forall \boldsymbol{\alpha} \in \mathcal{U}_{ad}$ is 1800, including 60 on the contact part. Finally, the Bézier functions of order $d = 20$ are used.

Example 1. Our aim is to find a shape of the contact part which minimizes peaks of normal contact stresses represented by the vector of the Lagrange multipliers $\boldsymbol{\lambda}(\boldsymbol{\alpha})$ or, equivalently, to find $\boldsymbol{\alpha} \in \mathcal{U}_{ad}$ minimizing the max-norm of $\boldsymbol{\lambda}(\boldsymbol{\alpha})$. Since this norm is not continuously differentiable, we will use the p-norm $\|\cdot\|_p$ of vectors instead, taking p large enough ($p = 6$ in our case). The density of surface tractions is set as follows: $\boldsymbol{P}^1 = (0; -60 \text{ MPa})$ on $(0, 1.8) \times \{1\}$ and zero on $(1.8, 2) \times \{1\}$, $\boldsymbol{P}^2 = (50 \text{ MPa}; 30 \text{ MPa})$ on $\{2\} \times (0, 1)$. We consider the shape optimization problem

$$(\hat{\mathbb{P}}_1) \qquad \begin{aligned} &\text{minimize} \quad \|\boldsymbol{\lambda}(\boldsymbol{\alpha})\|_6^6 \ . \\ &\boldsymbol{\alpha} \in \mathcal{U}_{ad} \end{aligned}$$

Let us observe that if $\boldsymbol{\alpha} \in \mathcal{U}_{ad}$ is such that $\boldsymbol{u}(\boldsymbol{\alpha}) \in \operatorname{int} \mathcal{K}$, where $\mathcal{K}$ is defined by (25), i.e., all linear inequality constraints are inactive, then $\boldsymbol{\lambda} = \boldsymbol{0}$ and such $\boldsymbol{\alpha}$ automatically solves ($\hat{\mathbb{P}}_1$). To prevent this trivial case, the admissible set $\mathcal{U}_{ad}$ involves also the volume constraint imposed on the computational domains with an appropriate choice of the constants C_{31} and C_{32}. Figure 3 shows the initial shape $\Omega(\boldsymbol{\alpha}_{in})$ with a finite element partition before (left) and after (right) the deformation. The same is depicted for the computed optimal shape $\Omega(\boldsymbol{\alpha}_{opt})$ in Figure 4. The solid line in
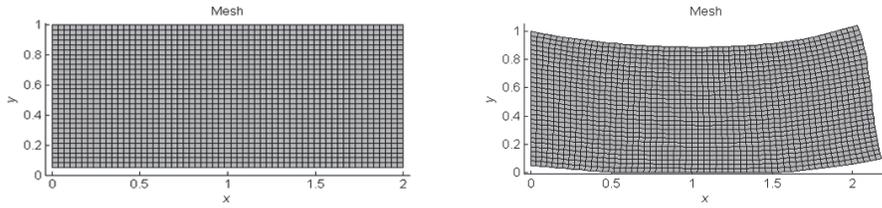
FIG. 3. *Example* 1, *initial design.*
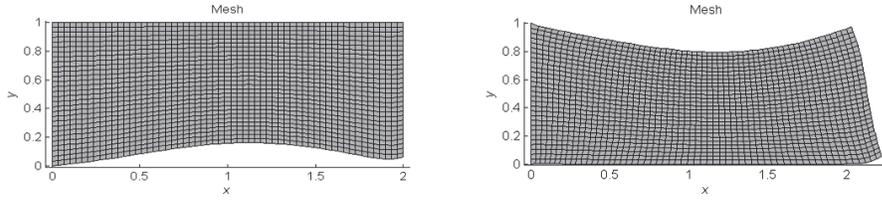
FIG. 4. *Example* 1, *optimal design.*
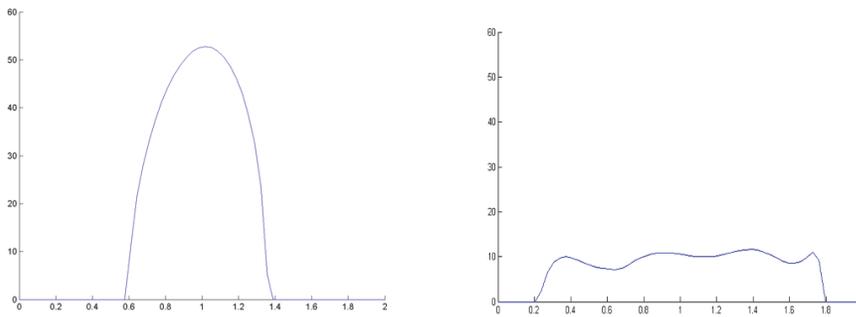
FIG. 5. *Example* 1, *normal stress distribution for initial (left) and optimal (right) design.*

Figure 5 shows the distribution of the vector of normal contact stresses $\boldsymbol{\lambda}(\boldsymbol{\alpha})$ along the contact part of $\Omega(\boldsymbol{\alpha}_{in})$ (left) and $\Omega(\boldsymbol{\alpha}_{opt})$ (right), respectively. It is readily seen from here that the peak of stresses is considerably suppressed and, in addition, $\boldsymbol{\lambda}(\boldsymbol{\alpha}_{opt})$ is evenly distributed along the contact part. The value of the objective function at $\boldsymbol{\alpha}_{in}$ is $2.1159 \cdot 10^{11}$ and $4.5492 \cdot 10^{7}$ at $\boldsymbol{\alpha}_{opt}$.

Example 2. The purpose of this shape optimization problem is to find a shape of the contact part such that the vector $\boldsymbol{\lambda}(\boldsymbol{\alpha})$ is as close as possible to a given target vector $\boldsymbol{\lambda}_{tar}$. The density of surface tractions $\boldsymbol{P}^1$ acting on the top of the body is the same as in Example 1, while $\boldsymbol{P}^2 = (30 \text{ MPa}; 10 \text{ MPa})$ on $\{2\} \times (0,1)$. The shape optimization problem reads as follows:

$$(\hat{\mathbb{P}}_2) \qquad \begin{aligned} &\text{minimize} && \|\boldsymbol{\lambda}(\boldsymbol{\alpha}) - \boldsymbol{\lambda}_{tar}\|_6^6 \ . \\ &\boldsymbol{\alpha} \in \mathcal{U}_{ad} \end{aligned}$$

Figures 6 and 7 are analogous to Figures 3 and 4. The dotted lines in Figure 8 correspond to the distribution of the target $\boldsymbol{\lambda}_{tar}$, while the solid lines represent the distribution of $\boldsymbol{\lambda}(\boldsymbol{\alpha}_{in})$ (left) and $\boldsymbol{\lambda}(\boldsymbol{\alpha}_{opt})$ (right) along the corresponding contact
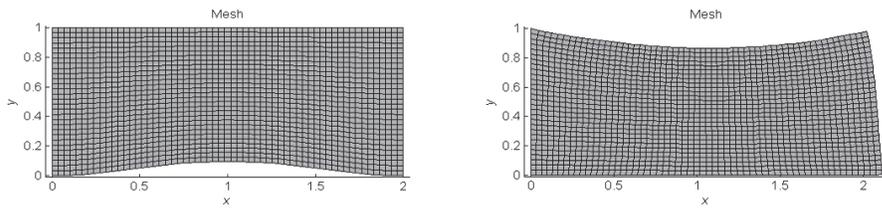
FIG. 6. *Example 2, initial design.*
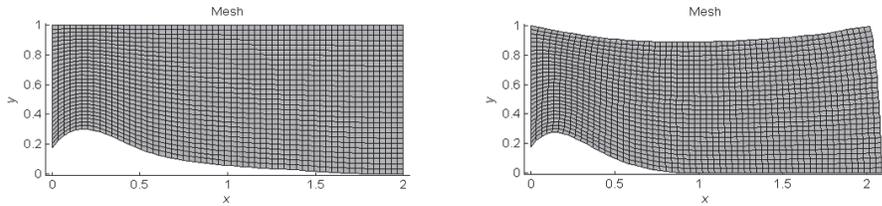
FIG. 7. *Example 2, optimal design.*
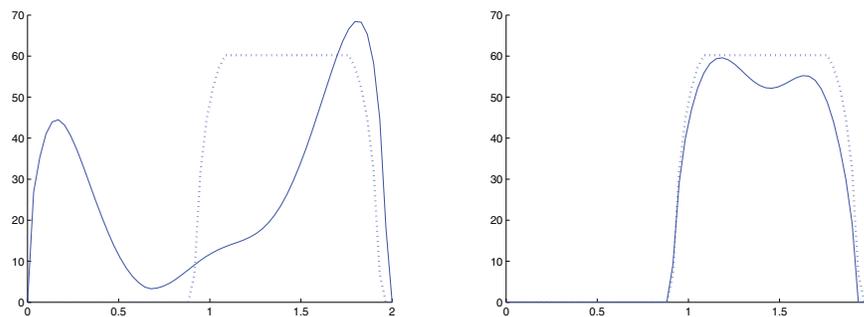
FIG. 8. *Example 2; normal stress distribution for initial (left) and optimal (right) design.*

parts. Again we see a considerable improvement. The value of the objective function at $\boldsymbol{\alpha}_{in}$ is $1.0746 \cdot 10^{14}$ and $4.7879 \cdot 10^{9}$ at $\boldsymbol{\alpha}_{opt}$.

Example 3. To show that the final result strongly depends on the choice of the model of friction, Example 2 was recomputed using the state problem with a simpler Tresca model of friction whose coefficient depends on the solution (for detailed analysis we refer the reader to [11]). Unlike Coulomb's law of friction, the unknown contact stress $T_n(\boldsymbol{u})$ in (7) is now replaced by a slip bound g given a priori. We set $g = 100$. Starting from the same initial configuration $\Omega(\boldsymbol{\alpha}_{in})$ as in Example 2, we arrive at the optimal shape $\Omega(\bar{\boldsymbol{\alpha}}_{opt})$ depicted in Figure 9. The resulting normal stress distribution along $\Gamma_c(\bar{\boldsymbol{\alpha}}_{\text{opt}})$ is shown in Figure 10 (left). Then $\Omega(\bar{\boldsymbol{\alpha}}_{opt})$ was used as the computational domain for the direct contact problem with Coulomb friction yielding the normal stress as in Figure 10 (right). Comparing Figures 10 and 8 (right) one can see that the optimal shape $\Omega(\bar{\boldsymbol{\alpha}}_{opt})$ for the Tresca model of friction is not optimal for friction obeying Coulomb's law.

Conclusion. In the present paper we have considered shape optimization in discretized two-dimensional contact problems with Coulomb friction and a solution-
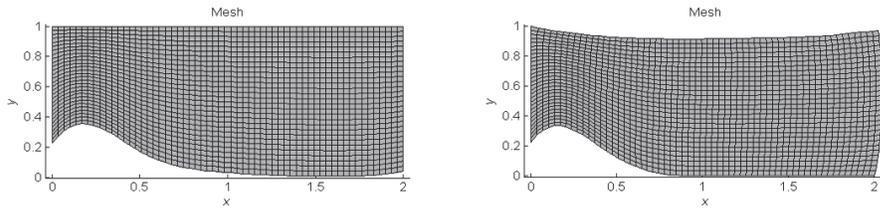
FIG. 9. *Example 3 for the Tresca model; optimal design* $\Omega(\bar{\boldsymbol{\alpha}}_{opt})$.
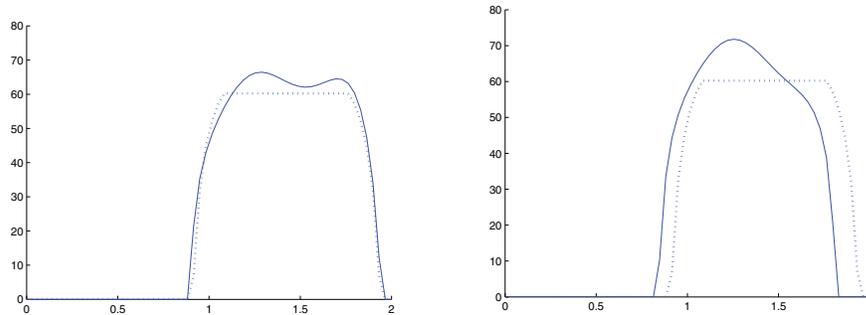
FIG. 10. *Example 3; normal stress distribution on* $\Gamma_c(\bar{\boldsymbol{\alpha}}_{opt})$.

dependent coefficient of friction. It was shown that if the coefficient of friction is sufficiently small in the $C^{0,1}$-norm, then the state problems are uniquely solvable and their solutions are Lipschitzian with respect to the design variable. Sensitivity analysis was carried out using the generalized differential calculus of B. Mordukhovich, enabling us to solve the shape optimization problems via the implicit programming approach. The obtained results were illustrated in several examples.

Extending the present results into the three-dimensional setting seems a challenge and is currently being investigated. From the theoretical point of view, it would be beneficial to show convergence of the bundle trust method in our context. This is linked to the semismoothness of the control-to-state mapping and will be the subject of future research.

REFERENCES

[1] R. L. BENEDICT AND J. E. TAYLOR, *Optimal design for elastic bodies in contact*, in Optimization of Distributed Parameter Structures, Part II, E. J. Haug and J. Céa, eds., Sijthof & Noorddhoff, Alphen aan den Rijn, The Netherlands, 1981, pp. 1553–1599.

[2] P. BEREMLIJSKI, J. HASLINGER, M. KOČVARA, R. KUČERA, AND J. V. OUTRATA, *Shape optimization in three-dimensional contact problems with Coulomb friction*, SIAM J. Optim., 20 (2009), pp. 416–444.

[3] P. BEREMLIJSKI, J. HASLINGER, M. KOČVARA, AND J. OUTRATA, *Shape optimization in contact problems with Coulomb friction*, SIAM J. Optim., 13 (2002), pp. 561–587.

[4] P. G. CIARLET, *The Finite Element Method for Elliptic Problems*, Stud. Math. Appl. 4, North–Holland, Amsterdam, 1978. Reprinted as Classics Appl. Math. 40, SIAM, Philadelphia, 2002.

[5] F. F. CLARKE, *Optimization and Nonsmooth Analysis*, John Wiley & Sons, New York, 1983.

[6] A. L. Dontchev and W. W. Hager, *Implicit functions, Lipschitz maps and stability in optimization*, Math. Oper. Res., 19 (1994), pp. 753–768.

[7] C. Eck, J. Jarušek, and M. Krbec, *Unilateral Contact Problems: Variational Methods and Existence Theorems*, Pure Appl. Math. 270, Chapman & Hall/CRC, Taylor & Francis Group, Boca Raton, FL, 2005.

[8] J. Haslinger, I. Hlaváček, and J. Nečas, *Numerical methods for unilateral problems in solid mechanics*, in Handbook of Numerical Analysis, Vol. IV, Part 2, P. G. Ciarlet and J.-L. Lions, eds., North–Holland, Amsterdam, 1996, pp. 313–491.

[9] J. Haslinger and R. A. E. Mäkinen, *Introduction to Shape Optimization: Theory, Approximation and Computation*, Adv. Des. Control 7, SIAM, Philadelphia, 2003.

[10] J. Haslinger and P. Neittaanmäki, *Finite Element Approximation for Optimal Shape, Material and Topology Design*, 2nd ed., John Wiley & Sons, New York, 1996.

[11] J. Haslinger, J. V. Outrata, and R. Pathó, *Shape optimization in $2D$ contact problems with given friction and a solution-dependent coefficient of friction*, Set-Valued Var. Anal., 20 (2012), pp. 31–59.

[12] J. Haslinger and O. Vlach, *Approximation and numerical realization of $2D$ contact problems with Coulomb friction and a solution-dependent coefficient of friction*, J. Comput. Appl. Math., 197 (2006), pp. 421–436.

[13] D. Hilding, A. Klarbring, and J. Petersson, *Optimization of structures in unilateral contact*, ASME Appl. Mech. Rev., 52 (1999), pp. 139–160.

[14] M. Kočvara and J. V. Outrata, *Optimization problems with equilibrium constraints and their numerical solution*, Math. Program. B, 101 (2004), pp. 119–150.

[15] MatSol, http://industry.it4i.cz/en/products/matsol/ (2014).

[16] R. Mifflin, *Semismooth and semiconvex functions in constrained optimization*, SIAM J. Control Optim., 15 (1977), pp. 957–972.

[17] B. S. Mordukhovich, *Approximation Methods in Problems of Optimization and Control*, Nauka, Moscow, 1988 (in Russian).

[18] B. S. Mordukhovich, *Variational Analysis and Generalized Differentiation, I: Basic Theory. II: Applications*, Grundlehren Math. Wiss. 330 and 331, Springer-Verlag, Berlin, Heidelberg, 2006.

[19] J. V. Outrata, *Mathematical programs with equilibrium constraints: Theory and numerical methods*, in Nonsmooth Mechanics of Solids, J. Haslinger and G. E. Stavroulakis, eds., CISM 485, Springer, Vienna, 2006, pp. 221–274.

[20] J. V. Outrata and M. Červinka, *On the implicit programming approach in a class of mathematical programs with equilibrium constraints*, Control Cybernet., 38 (2009), pp. 1557–1574.

[21] J. Outrata, M. Kočvara, and J. Zowe, *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints: Theory, Applications and Numerical Results*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998.

[22] S. M. Robinson, *Strongly regular generalized equations*, Math. Oper. Res., 5 (1980), pp. 43–62.

[23] R. T. Rockafellar and R. Wets, *Variational Analysis*, Springer-Verlag, Berlin, 1998.

[24] H. Schramm and J. Zowe, *A version of the bundle idea for minimizing a nonsmooth function: Conceptual idea, convergence analysis, numerical results*, SIAM J. Optim., 2 (1992), pp. 121–152.

[25] J. Sokolowski and J.-P. Zolesio, *Introduction to Shape Optimization: Shape Sensitivity Analysis*, Springer Ser. Comput. Math. 16, Springer-Verlag, Berlin, 1992.