

Algorithms for Optimal Control of Elastic Contact Problems with Finite Strain

Anton Schiela and Matthias Stöcklein

Abstract. Optimal control of hyperelastic contact problems in the regime of finite strains combines various severe theoretical and algorithmic difficulties. Apart from being large scale, the main source of difficulties is the high non-linearity and non-convexity of the elastic energy functional which precludes uniqueness of solutions and simple local sensitivity results. In addition, the contact conditions add non-smoothness to the overall problem.

In this paper, we discuss algorithmic approaches to address these issues. In particular, the non-smoothness is tackled by a path-following approach, whose theoretical properties are reviewed. The subproblems are highly non-linear optimal control problems, which can be solved by an affine invariant composite step method. For increased robustness and efficiency this method has to be adapted to the particular problem, taking into account its large scale nature, its function space structure and its non-convexity.

Mathematics Subject Classification (2000). Primary 49M37; Secondary 90C55.

Keywords. nonlinear elasticity, optimal control, contact problem.

1. Introduction

The analysis and simulation of elastic contact problems, in particular for small deformations are a classical subject of applied mathematics. Already in their simplest form, the Signorini problem [23, 5, 12], they are intrinsically non-smooth and lead to a variational inequality on an appropriate Sobolev space. Nevertheless, even in this simple, convex setting, the simulation of linearly elastic contact may be challenging [12], depending on the geometric configuration.

Nonlinearly elastic contact problems combine these difficulties with those that arise in nonlinear hyperelasticity. Solutions of hyperelastic problems

This work was supported by the DFG grant SCHI 1379/2-1 within the priority programme SPP 1962 (Non-smooth and Complementarity-based Distributed Parameter Systems: Simulation and Hierarchical Optimization).

can be modeled as energy minimizers [4]. Due to non-convexity of the energy, minimizers do not have to be unique. Also, owing to the high nonlinearity of the problem, local minimizers do not have to satisfy the weak form of the equilibrium equation in general.

Here we consider optimization problems in the context of nonlinear elasticity and contact. This connects to the works [13, 14], where first steps into the topic were taken. In particular, existence of optimal solutions to such kinds of problems was shown. A composite step method [15] was developed for the numerical solution of these problems. A similar topic was considered in [9], where optimal control problems in the context of biological models were investigated and solved by a quasi-Newton approach. Optimal control of linear contact problems has been considered in [2, 17, 25].

In the last few years, additional progress has been made for this class of problems, and the aim of this paper is to report on this progress, both concerning theoretical results and algorithmic concepts. We build upon and extend the results from [15], [22] and [21], where additional details can be found.

This work consists of two main parts: The first part is a concise recapitulation of the available theoretical results for optimal control of finite strain contact problems. It is mainly based on [22], and, after fixing the framework, describes analytic results on a path-following approach for the solution of these problems. The second part deals with the algorithmic development, that has taken place recently. Here a number of inherent numerical and practical challenges are described, and algorithmic ways to deal with them are presented. First, elastic problems in three spatial dimensions yield large scale systems after discretization. Hence, efficient iterative solvers for the computation of steps have to be used. To this end, in [21] an algorithmic framework for inexact step computations was developed, but our class of problems demands further advance in this direction. Second, an appropriate choice of functional analytic framework is discussed. It turns out that the choice of norms has a decisive impact on the performance of the algorithms. Finally, we consider the treatment of the inherent non-convexity in the problem, both in the objective and in the energy functional. Our discussions are illustrated by numerical examples.

2. Contact Problems in Hyperelasticity

Generally speaking, we study the deformation of a nonlinear elastic body made of a hyperelastic material. The body is considered to be under stress from an external boundary force which causes the deformation. Additionally, deformations are constrained by an obstacle which the body cannot penetrate. In the context of hyperelasticity, computing such deformations corresponds to solving an energy minimization problem.

In this section, we introduce the setting and review the central results for nonlinear elastic contact problems. In particular, we give an overview of the existence

theory in nonlinear elasticity and we address a suitable regularization approach for the contact constraints.

Nonlinear elasticity. In our setting, the nonlinear elastic body is represented by a domain $\Omega \subset \mathbb{R}^3$ which is required to be Lipschitz continuous. In addition, its boundary is divided into three subsets as follows:

$$\Gamma = \overline{\Gamma_D \cup \Gamma_N \cup \Gamma_C},$$

where each subset has non-zero boundary measure. Here, Γ_D and Γ_N denote the parts where Dirichlet and Neumann boundary conditions hold, respectively. Further, $\Gamma_U \subset \Gamma_N$ and Γ_C denote the parts where the boundary force acts and where the contact constraints are enforced, respectively.

Next, we denote by

$$y : \overline{\Omega} \rightarrow \mathbb{R}^3, \text{ and } u : \Gamma_U \rightarrow \mathbb{R}^3$$

the deformation of the body and the boundary force, respectively.

For simplicity, we consider the following contact constraint:

$$y_3 \geq 0 \text{ a.e. on } \Gamma_C.$$

This describes a setting where the body has to stay above the plane that is spanned by the first two canonical basis vectors.

As deformation space, we choose the Sobolev space $W^{1,p}(\Omega; \mathbb{R}^3)$ with $p \geq 2$. Correspondingly, as space for the boundary force we choose $L^2(\Gamma_U, \mathbb{R}^3)$. If there is no risk of ambiguity, we will skip the notation for the image space in all vector-valued spaces. In the setting of optimal control, deformations will act as the state and boundary forces will act as the control. Accordingly, we introduce the notation $Y = W^{1,p}(\Omega)$ and $U = L^2(\Gamma_U)$. Further, let $id : \Omega \rightarrow \Omega$ be the identity mapping, and let \mathbb{M}_+^3 denote the space of invertible 3×3 matrices with positive determinant. Lastly, if not stated otherwise, we define the matrix norm $\|M\| := \sqrt{\text{tr } M^T M}$ and we denote by $\text{Cof } M := \det(M)M^{-T}$ the cofactor matrix for $M \in \mathbb{M}_+^3$.

For hyperelastic materials, computing the deformation of a body subjected to an external load is equivalent to finding a respective energy minimizer. The corresponding total energy functional $I : Y \times U \rightarrow \mathbb{R}$ can be defined by

$$I(y, u) := \int_{\Omega} \hat{W}(\omega, \nabla y(\omega)) d\omega - \int_{\Gamma_U} yu ds,$$

with the common splitting

$$I_{\text{strain}}(y) = \int_{\Omega} \hat{W}(\omega, \nabla y(\omega)) d\omega \quad \text{and} \quad I_{\text{out}}(y, u) = \int_{\Gamma_U} yu ds.$$

Here, $\hat{W} : \Omega \times \mathbb{M}_+^3 \rightarrow \mathbb{R}$ denotes the stored energy function which depends on the material. For detailed discussion of the specific choice of \hat{W} we refer to [3]. For the further analysis, we require the following assumptions, which are standard in hyperelasticity:

Assumption 2.1. Let $\hat{W} : \Omega \times \mathbb{M}_+^3 \rightarrow \mathbb{R}$ be the stored energy function. We assume that the following properties hold:

1. *Polyconvexity:* For almost all $\omega \in \Omega$, there is a convex function $\mathbb{W}(\omega, \cdot, \cdot, \cdot) : \mathbb{M}^3 \times \mathbb{M}^3 \times]0, +\infty[\rightarrow \mathbb{R}$ such that

$$\hat{W}(\omega, M) = \mathbb{W}(\omega, M, \text{Cof } M, \det M), \quad \text{for all } M \in \mathbb{M}_+^3.$$

The function $\mathbb{W}(\cdot, M, \text{Cof } M, \det M) : \Omega \rightarrow \mathbb{R}$ is measurable for all $M \in \mathbb{M}_+^3$.

2. For almost all $\omega \in \Omega$, the implication $\det M \rightarrow 0^+ \Rightarrow \hat{W}(\omega, M) \rightarrow \infty$ holds.
3. The sets of admissible deformations defined by

$$\begin{aligned} \mathcal{A} &:= \{y \in W^{1,p}(\Omega), \text{Cof } \nabla y \in L^s(\Omega), \det \nabla y \in L^r(\Omega), \\ &\quad y = \text{id a.e. on } \Gamma_D, \det \nabla y > 0 \text{ a.e. in } \Omega\}, \\ \mathcal{A}_c &:= \{y \in \mathcal{A} : y_3 \geq 0 \text{ a.e. on } \Gamma_c\}, \end{aligned}$$

for $p \geq 2$, $s \geq \frac{p}{p-1}$, $r > 1$ are non-empty.

4. *Coercivity:* There exist $a \in \mathbb{R}$, and $b > 0$, such that

$$\hat{W}(\omega, M) \geq a + b(\|M\|^p + \|\text{Cof } M\|^s + |\det M|^r).$$

5. The identity $\text{id} : \bar{\Omega} \rightarrow \bar{\Omega}$ satisfies $I_{\text{strain}}(\text{id}) = 0$, $\text{id}_3 \geq 0$ a.e. on Γ_C .

With this at hand, computing the deformation of a body constrained by an obstacle can be described by the optimization problem

$$y \in \underset{v \in \mathcal{A}_c}{\text{argmin}} I(v, u). \quad (1)$$

Existence of energy minimizers has been established in [4, Theorem 4.2], extending techniques from [1].

Regularization of contact constraints. Contact constraints add non-smoothness to an already highly nonlinear and non-convex problem. Therefore, we will introduce a suitable regularization approach.

Here, we apply the normal compliance regularization used in [18, 16]. In this context, we introduce the penalty functional $P : Y \rightarrow \mathbb{R}_0^+$ defined by

$$P(v) := \frac{1}{k} \int_{\Gamma_C} [-v_3]_+^k ds, \quad k \in \mathbb{N}, k > 1, \quad v \in Y,$$

which measures the violation of the constraints. We add the scaled penalty function P to the total energy functional I

$$I_\gamma(y, u) := I(y, u) + \gamma P(y) \quad \gamma > 0.$$

This approach allows us to drop the contact constraints. As a result, we obtain the regularized minimization problem:

$$y \in \underset{v \in \mathcal{A}}{\text{argmin}} I_\gamma(v, u). \quad (2)$$

The well-posedness of the regularized problem (2) and a convergence result that links (1) to (2) have been proven in [22, Theorem 2.3, Proposition 2.1].

Theorem 2.1. *Let $\gamma > 0$ be a fixed penalty parameter and $u \in U$ be some fixed boundary force. Then, under Assumption 2.1, the regularized total energy functional $I_\gamma(\cdot, u)$ has at least one minimizer in \mathcal{A} .*

3. Optimal Control of Nonlinear Elastic Contact Problems

In the optimal control setting, we aim at minimizing an objective functional

$$J : Y \times U \rightarrow \mathbb{R},$$

subject to the constraint that an optimal state y_* is a minimizer of the total energy functional i.e.

$$y_* \in \operatorname{argmin}_{v \in \mathcal{A}_c} I(v, u_*),$$

where u_* is the corresponding optimal control. We restrict ourselves here to a tracking type functional of the form

$$J(y, u) := \frac{1}{2} \|y - y_d\|_{L^2(\Omega)}^2 + \frac{\alpha}{2} \|u\|_{L^2(\Gamma_U)}^2,$$

for $y_d \in L^2(\Omega)$, $\alpha > 0$ and $U = L^2(\Gamma_U)$. Accordingly, the optimal control problem reads as follows:

$$\min_{(y, u) \in Y \times U} J(y, u) \quad \text{s.t.} \quad y \in \operatorname{argmin}_{v \in \mathcal{A}_c} I(v, u). \quad (3)$$

Based on the analysis in [14], the following existence result was derived in [22, Theorem 4.1]:

Theorem 3.1. *Problem (3) has at least one optimal solution.*

Solving those kinds of optimal control problems numerically is already a challenging task, even without contact constraints. Therefore, we are going to apply the previously introduced normal compliance regularization in order to avoid dealing with the contact constraints numerically. As a result, we obtain the regularized optimal control problem:

$$\min_{(y, u) \in Y \times U} J(y, u) \quad \text{s.t.} \quad y \in \operatorname{argmin}_{v \in \mathcal{A}} I_\gamma(v, u), \quad (4)$$

for some fixed parameter $\gamma > 0$.

Analogously to above, we can show the existence of optimal solutions.

Theorem 3.2. *For each $\gamma > 0$ problem (4) has at least one optimal solution.*

Proof. See [22, Proof of Theorem 4.2]. □

Convergence of Solutions of the Regularized Problem. With the regularized optimal control problem at hand, we now have to verify that solutions of the regularized problem (4) approach solutions of the original control problem (3).

However, concerning the regularization (4), one corner case precludes the desired result: it may happen that the energy minimization problems in (3) admit a larger set of energy minimizers than can be approximated by solutions of (4). So the desired convergence theory for (4) can only be established under an assumption that rules this case out, as discussed in [22]. However, a modified regularization can solve this problem.

We introduce the following alternative regularized problem:

$$\mathcal{E}_\gamma(y, u) := I_\gamma(y, u) + \varphi(\gamma) \frac{1}{2} \|y - y_d\|_{L^2(\Omega)}^2, \quad (5)$$

where $\varphi : [0, \infty[\rightarrow]0, \infty[$ is a positive function in γ , that is monotonically decreasing, such that

$$\lim_{\gamma \rightarrow \infty} \varphi(\gamma) = 0.$$

Again, the well-posedness and a convergence result for this new regularization were established in [22].

For the convergence analysis, it has to be ensured that the regularization function φ does not approach zero too quickly. This is necessary to guarantee that the minimization of a fraction of the objective functional J is sufficiently weighted at all times. This property is specified in the following assumption.

Assumption 3.1. *Let $u \in U$ be fixed. Assume that*

$$\lim_{\gamma \rightarrow \infty} \frac{\min_{v \in \mathcal{A}_c} I(v, u) - \min_{v \in \mathcal{A}} I_\gamma(v, u)}{\varphi(\gamma)} = 0.$$

With this in mind, we can show convergence for this approach.

Theorem 3.3. *Let $\gamma_n \rightarrow \infty$ be a positive and monotonically increasing sequence of penalty parameters. Furthermore, let (y_*, u_*) denote an optimal solution to problem (3). In addition, let $(y_n, u_n) \subset \mathcal{A} \times U$ be a sequence of optimal solutions to the corresponding regularized problems, where the regularization function φ satisfies Assumption 3.1 w.r.t. u_* . Then,*

$$\lim_{n \rightarrow \infty} J(y_n, u_n) = \min_S J.$$

Furthermore, there exists a subsequence (y_{n_k}, u_{n_k}) and a pair $(\bar{y}, \bar{u}) \in \mathcal{A}_c \times U$ such that we obtain the weak convergence $y_{n_k} \rightharpoonup \bar{y}$ in Y and the strong convergence $u_{n_k} \rightarrow \bar{u}$ in $L^2(\Gamma_U)$. Additionally, (\bar{y}, \bar{u}) solves the original problem (3).

Proof. See [22, Proof Theorem 5.4]. □

In [22] additional results are established that allow an a-priori choice of φ , depending on the regularity of the geometric configuration.

Formal KKT Conditions. Let us define

$$c_\gamma(y, u)v = \partial_y \mathcal{E}_\gamma(y, u)v \quad \forall v \in P$$

where P is a reflexive space of test functions. This mapping corresponds to equilibrium conditions of our hyperelastic problem. We thus have a nonlinear mapping:

$$c_\gamma : Y \times U \rightarrow P^*$$

that can be split additively as follows:

$$c_\gamma(y, u) = A_\gamma(y) - Bu$$

into a nonlinear operator $A_\gamma : Y \rightarrow P^*$ and a linear operator $B : U \rightarrow P^*$. Then, formally, the KKT-conditions at a minimizer x_* state the existence of an adjoint state p such that:

$$\begin{aligned} J'(y_*, u_*) + c'_\gamma(y_*, u_*)^* p &= 0 \\ c_\gamma(y_*, u_*) &= 0. \end{aligned}$$

We stress that a rigorous derivation of these conditions seems to be out of reach at the moment. The main reason is the lack of local sensitivity results of solutions of hyperelasticity with respect to perturbations of u .

4. Numerical Optimization Algorithms

In order to algorithmically approach this problem, we formally replace the energy minimizing constraint by its first order optimality condition. Then, the reformulated problem reads as follows:

$$\min_{(y, u) \in Y \times U} J(y, u) \quad s.t. \quad c_\gamma(y, u) = 0. \quad (6)$$

As result, we obtain an equality constrained optimization problem for each parameter $\gamma > 0$. This formulation allows the application of solution algorithms for equality constraints. Nevertheless a couple of intrinsic difficulties have to be considered. To overcome them, measures have to be taken that go beyond standard equality constrained optimization:

- The problem is posed in function space, and even after discretization (which is done here by a displacement formulation), this inherent structure should be taken into account by the algorithm.
- In three dimensional elasticity, the use of direct solvers limits severely the resolution of discretizations. Thus, all arising linear systems have to be solved by iterative methods, preferably of conjugate gradient type. In this context, the issue of finding appropriate preconditioners and termination criteria arises.
- Although the elastic energy minimization problem has been replaced by its equilibrium conditions, the goal remains to compute stable solutions, i.e., energy minimizers. Hence, our algorithm should have built-in preference towards energy decreasing search directions. This issue arises due to non-convexity, in particular, if the problem of linearized elasticity yields a Hessian

matrix that is not positive definite. This case also precludes the direct application of a conjugate gradient method.

- The high nonlinearity of the energy functional also includes a singularity near $\det \nabla y = 0$, while local self-penetration, i.e., $\det \nabla y < 0$ is infeasible.

While we concentrated [22] on the construction of a path-following method for the regularization of the contact constraints, the aim of this section is to propose algorithmic measures to tackle the above problems, which are sometimes specific to nonlinear elasticity, and to illustrate their numerical performance.

For the following numerical computations we always employ the tracking type functional, described above and a nonlinear material that is used for modelling soft biological tissue. They differ in terms of geometric configuration and in terms of the desired deformed state y_d .

4.1. An affine covariant composite step method

For brevity of notation we set $x := (y, u)$ and $X = Y \times U$, which yields the formulation

$$\min_{x \in X} J(x) \quad \text{s.t.} \quad c_\gamma(x) = 0. \quad (7)$$

The space of iterates is equipped with an appropriately chosen scalar-product $\langle \cdot, \cdot \rangle$ that gives rise to a Riesz-isomorphism $M : X \rightarrow X^*$, i.e., $\langle v, w \rangle = (Mv)(w)$. As usual, we define the Lagrangian function $L : X \times P \rightarrow \mathbb{R}$ via $L(x, p) := J(x) + pc_\gamma(x) = J(x) + p \circ c_\gamma(x)$.

For the solution of the minimization problem (7), we apply a composite step algorithm based on the preceding work [15], from which we recapitulate the main ideas.

The idea of composite step methods is to split the Newton update δx into a normal step δn and a tangential step δt for a precise treatment of optimality and feasibility. A normal step δn satisfies $\delta n \in \ker c'(x)^\perp$ and aims for feasibility. It can be computed via the augmented system:

$$\begin{pmatrix} M & c'_\gamma(x)^* \\ c'_\gamma(x) & 0 \end{pmatrix} \begin{pmatrix} \delta n \\ q \end{pmatrix} + \begin{pmatrix} 0 \\ c_\gamma(x) \end{pmatrix} = 0, \quad (8)$$

which corresponds to the minimization problem:

$$\min_v \frac{1}{2} \langle v, v \rangle_M \quad \text{s.t.} \quad c'_\gamma(x)v + c_\gamma(x) = 0.$$

If necessary, a damping factor $\nu \in]0, 1]$ is applied. The tangential step δt satisfies $\delta t \in \ker c'(x)$ and aims for a decrease of the functional value. It can be computed by solving the problem:

$$\begin{pmatrix} L_{xx}(x, p) & c'_\gamma(x)^* \\ c'_\gamma(x) & 0 \end{pmatrix} \begin{pmatrix} \delta t \\ q \end{pmatrix} + \begin{pmatrix} L_x(x, p) + L_{xx}(x, p)\nu\delta n \\ 0 \end{pmatrix} = 0. \quad (9)$$

This problem corresponds to the following minimization problem for $\delta x = \nu\delta n + \delta t$ with fixed $\nu\delta n$:

$$\min_{\delta t} f'(x)\delta x + \frac{1}{2}L_{xx}(x, p)(\delta x, \delta x) \quad \text{s.t.} \quad c'(x)\delta t = 0.$$

The Lagrange multiplier p , which is needed for this step, is computed before as follows:

$$\begin{pmatrix} M & c'(x)^* \\ c'(x) & 0 \end{pmatrix} \begin{pmatrix} g \\ p \end{pmatrix} + \begin{pmatrix} f'(x) \\ 0 \end{pmatrix} = 0. \quad (10)$$

In contrast to (8), the upper left block in (9) need not be positive definite. The best, we can hope for is that $L_{xx}(x, p)$ is positive definite on $\ker c'_\gamma(x)$ if x is close to a strict minimizer of the problem.

Adding δn and δt directly results in a full Lagrange-Newton step, while damping can be used to construct a globalization procedure, e.g., or the form:

$$\delta x := \nu \delta n + \tau \delta t,$$

where $\nu \in]0, 1]$ and $\tau > 0$ are damping and step size parameters. This class of methods is popular in equality constrained optimization and optimal control and there are various realizations of this principal idea available [24, 19, 20, 10, 27].

The motivation of the approach in [15] is due to functional analytic considerations. Most methods for equality constraint optimization use residual norms of the form $\|c_\gamma(x)\|_{P^*}$, e.g., within a merit function or a filter. If $c_\gamma(x)$ models a partial differential equation, appropriate norms are dual norms and thus not easy to evaluate. The use of a simple norm at this position may degrade the performance of the globalization procedure considerably.

The concept of affine covariance [7] allows to dispense with the evaluation of residuum norms. Instead, a simplified Newton step δs is used, which is defined as a minimum norm solution of a simplified Newton equation:

$$\begin{pmatrix} M & c'_\gamma(x)^* \\ c'_\gamma(x) & 0 \end{pmatrix} \begin{pmatrix} \delta s \\ q \end{pmatrix} + \begin{pmatrix} 0 \\ c_\gamma(x + \delta x) - c_\gamma(x) - c'_\gamma(x)\delta x \end{pmatrix} = 0 \quad (11)$$

By computing the ratio $\Theta := \|\delta s\|_X / \|\delta x\|_X$, it is possible to estimate the Newton contraction towards the feasible manifold. If $\Theta \ll 1$, we are in the region of fast local convergence of Newton's method for the solution of the underdetermined problem $c_\gamma(x) = 0$. This globalization idea is combined with an appropriate decrease criterion. Details are elaborated in [15]. There it is also shown that δs helps to overcome the Maratos effect, since $J(x + \delta x + \delta s)$ is approximated better by the quadratic model of the Lagrange Newton step than $J(x + \delta x)$ is. Thus, δs plays the role of a second order correction and thus is beneficial in two ways.

To enforce the non-self-penetration condition $\det \nabla y > 0$, additional damping is applied if the trial iterate violates this condition.

4.2. Computation of steps by iterative solvers

We observe that the systems (8)-(11) all have a common structure, which, after splitting of $X = Y \times U$ can be written as follows:

$$\begin{pmatrix} H & C^* \\ C & 0 \end{pmatrix} \begin{pmatrix} v \\ q \end{pmatrix} + \begin{pmatrix} c_{1,2} \\ c_3 \end{pmatrix} = 0 \Leftrightarrow \begin{pmatrix} H_y & 0 & A^* \\ 0 & H_u & -B^* \\ A & -B & 0 \end{pmatrix} \begin{pmatrix} v_y \\ v_u \\ q \end{pmatrix} + \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = 0 \quad (12)$$

Here we used the observation that $J(y, u) = J_1(y) + J_2(u)$, so that H is block diagonal and set $Cx = Ay - Bu$. Since $J_2(u) = \alpha/2\|u\|^2$, we know that H_u is positive definite. However, H may not always be positive definite on $\ker C$ in case of the tangential step (9). This can only be expected close to a minimizer. Moreover, in our context of hyperelasticity the block $A = A'_\gamma(y) = \partial_{yy}\mathcal{E}_\gamma(y, u)$ is symmetric, but not always positive definite (and may be singular) due to non-convexity of the elastic energy. However, as we will discuss below, it is possible to modify A and H , such that the modified operators are positive definite, so that, in particular, A is invertible. Finally, we observe that $c_3 = 0$ in (10) and (9), while by solving the system $Av_{y,0} + c_3 = 0$ we can reduce (8) and (11) to a problem, where $c_3 = 0$ holds as well.

In this setting a conjugate gradient method on $\ker C$ can be applied to (12). For its implementation a constraint preconditioner is needed that guarantees that iterates remain in $\ker C$, as long as $c_3 = 0$. We obtain a projected conjugate gradient method, cf. e.g. [8]. Taking into account the block structure of (12), we use the following block lower triangular preconditioner P :

$$P := \begin{pmatrix} 0 & 0 & A^* \\ 0 & \tilde{H}_u & -B^* \\ A & -B & 0 \end{pmatrix},$$

dropping H_y and replacing H_u by a preconditioner \tilde{H}_u , e.g., if H_u is a mass matrix, its diagonal. This decouples (12) into three equations, which can be solved sequentially:

$$A^*q = -c_1 \quad \rightarrow q \quad (13)$$

$$\tilde{H}_u v_u = B^*q - c_2 \quad \rightarrow v_u \quad (14)$$

$$Av_y = Bv_u \quad \rightarrow v_y. \quad (15)$$

The main computational effort is spent solving (13) and (15), which are problems of linearized elasticity in 3d. For coarse discretizations a sparse direct solver can be used to factorize $A = A^*$ and solve (13) and (15). In that case, the preconditioned cg-method can be applied directly to (12) despite its saddle point structure.

Inexact constraint preconditioning. For fine discretizations, A cannot be factorized directly, and we have to resort to preconditioned conjugate gradients. Here, we use a multigrid preconditioner of BPX-type, equipped with a block Jacobi smoother that uses the diagonal of 3×3 blocks of A , respecting the vector valued nature of the problem.

Some care has to be taken, when this method is implemented. If (13) and (15) are solved only inexactly, the iterates are not contained in $\ker C$ any longer. Instead of (13) and (15), one actually solves near-by problems

$$\widetilde{A}^*q = -c_1 \quad (16)$$

$$\widetilde{A}v_y = Bv_u \quad (17)$$

with $\tilde{A} \approx \tilde{A}^* \approx A = A^*$, changing slightly from step to step. If the cg method applies the operator in (12) as the forward operator, this may lead to spurious occurrence of directions of negative curvature, unless the accuracy of solution of (13) and (15) is very high.

Hence, we have to avoid application of the A and A^* blocks in (12). This can be done by a simple auxiliary recursion within the projected cg-method. Although it is known that the convergence theory of conjugate gradients requires that preconditioners are linear mappings that do not change during the iteration, the cg-method then tolerates small errors in the application of the preconditioners. It can be observed, however, that loose tolerances in the solution of (16) and (17) are detrimental for the speed of convergence of the outer cg-iteration.

Remark 4.1. *It is desirable to reduce the accuracy requirement for (16) and (17) even further. Then a linear iteration scheme, such as a preconditioned Chebychev semi-iteration with a fixed number of steps, has to be employed and we end up with a solution of the perturbed problem:*

$$\begin{pmatrix} H_y & 0 & \tilde{A}^* \\ 0 & H_u & -B^* \\ \tilde{A} & -B & 0 \end{pmatrix} \begin{pmatrix} \tilde{v}_y \\ \tilde{v}_u \\ q \end{pmatrix} + \begin{pmatrix} c_1 \\ c_2 \\ 0 \end{pmatrix} = 0, \quad (18)$$

where in contrast to before, \tilde{A} is a linear operator, so that the outer cg-iteration really solves a well defined linear problem.

Solving (15) with \tilde{v}_u on the right hand side (by conjugate gradients) yields solutions in $\ker C$ again. A linear solver, based on this idea has been tested with promising results for optimal control of linear elliptic problems. Application to optimal control of nonlinear elasticity is under current investigation.

Remark 4.2. *Within the approach of Byrd-Omojokun composite step methods, inexact system solvers were considered in [11, 10, 20]. Here an alternative route is taken. A GMRES method is used to solve normal steps inexactly, allowing for loose tolerances in the evaluation of A and A^* . The solver for this problem also serves as a preconditioner for the tangential step, for which a projected cg with re-orthogonalization is used.*

Accuracy matching. For the efficiency of the overall method it is important that the steps are computed with neither too tight tolerances, which renders each step too expensive, nor with too loose tolerances, which may lead to loss of robustness and increase of the number of outer iterations. Setting fixed tolerances is usually not the best way to cope with this problem, since each step of the outer iteration has a different characteristic. For example, if the tangential step is dominant (which often happens close to the optimal solution), then normal and simplified normal step can be computed with low relative accuracy. A strategy for accuracy matching has been proposed and tested in [21], where in particular the impact of inexact normal and simplified normal steps on the outer iteration was considered and adaptive termination criteria were derived.

4.3. Choice of functional analytic framework

A very important issue for our problem is a good choice of a Hilbert space norm on $X = Y \times U$ that measures the step-lengths and also defines what *normal* means, when normal steps are computed. Our numerical results show that this has considerable impact on the performance of our algorithm. Recall that the Riesz-operator $M : X \rightarrow X^*$ enters the definition of normal and simplified normal step, as well as the computation of the Lagrange multiplier in (8), (11), and (10).

For certain classes of optimal control problems with mildly nonlinear PDEs (e.g. semilinear equations) it is possible to find a Hilbert space norm on X that allows a rigorous (local) convergence theory in function space, based in the corresponding functional analytic setting. To obtain analogous results for nonlinear elasticity is illusory. Even the solution of the forward problem is prone to a two-norm discrepancy: differentiability of the energy functional cannot be expected in a space less regular than $W^{1,\infty}(\Omega)$, while the energy space is only $W^{1,2}(\Omega)$. We thus have a norm-gap that is hard to bridge.



FIGURE 1. Problem of pushing down a plate. Left: Undeformed domain. Middle: Desired deformation. Right: Optimal deformation. Upper horizontal boundary: Γ_U , colour codes intensity of force. Lower horizontal boundary: $\Gamma_N \setminus \Gamma_U$. Vertical boundaries: Γ_D .

As a consequence mesh-dependent behaviour of solvers has to be expected, at least, if difficult problems with large strains are solved. If the problem is not too hard, however, additional regularity of the steps can usually be observed, which alleviates the difficulty in practice. This favourable effect depends on the concrete configuration and is hard to grasp a-priori in a mathematical theory.

Our numerical observations confirm our considerations. In the following consider two alternative norms:

$$\begin{aligned} \|(y, u)\|_{M_0}^2 &:= \frac{1}{2} \|y\|_{L_2(\Omega)}^2 + \frac{\alpha}{2} \|u\|_{L_2(\Gamma)}^2, \\ \|(y, u)\|_{M_1}^2 &:= \frac{1}{2} \|y\|_{H^1(\Omega)}^2 + \frac{\alpha}{2} \|u\|_{L_2(\Gamma)}^2. \end{aligned}$$

Clearly, $\|\cdot\|_{M_0}$ is in close correspondence with the objective functional to be minimized, but does not take into account the regularity requirements of the non-linearity, at all. In contrast $\|\cdot\|_{M_1}$ promotes smoother states, so, although not

guaranteeing $W^{1,\infty}(\Omega)$ regularity, is certainly considerably closer to the ideal situation.

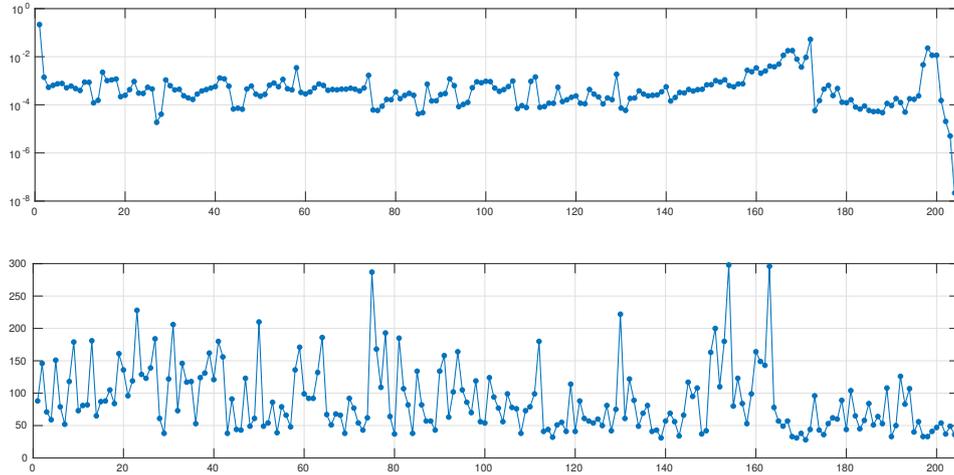


FIGURE 2. Iteration history for problem from Figure 1, using M_0 : Top: norms of steps, taken by the algorithm, Bottom: total number of outer cg-Iterations in each step

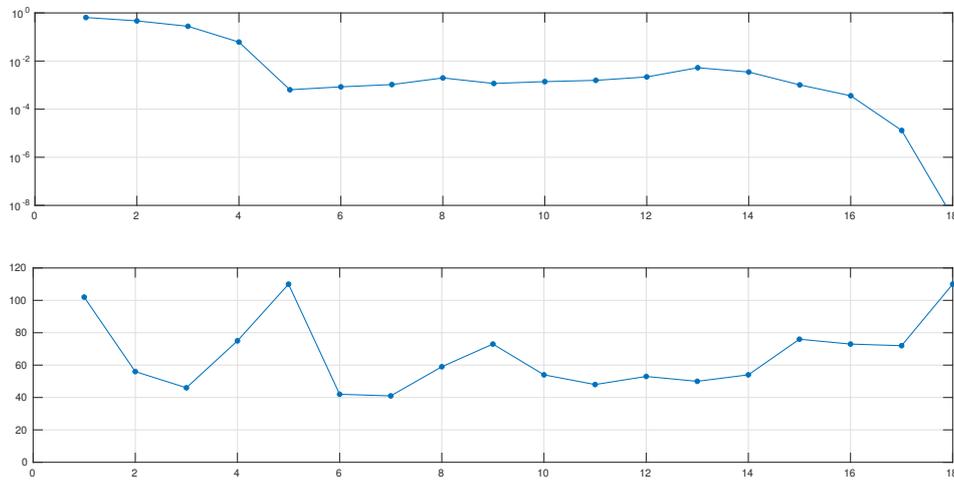


FIGURE 3. Iteration history for problem from Figure 1, using M_1 : Top: norms of steps, taken by the algorithm, Bottom: total number of outer cg-Iterations in each step

We test these two alternatives at a problem, described in Figure 1. The results of a numerical comparison are depicted in Figure 2 and Figure 3. We observe a marked difference, although the computed final solutions, as inspection showed, are the same. Equipped with $\|\cdot\|_{M_0}$ our algorithm takes about 200 steps and shows quite irregular behaviour. If $\|\cdot\|_{M_1}$ is used, we observe from Figure 3 that the behaviour of our algorithm is much faster and also much more regular, concerning choice of damping factors.

4.4. Non-convexity of Objective and Energy

A major difficulty in the considered class of problems is the occurrence of non-convexity, not only in the objective functional, but also in the energy functional. The difficulties, introduced by non-convexity are well known: we usually obtain non-unique local minimizers, additional stationary points, and quadratic models used in SQP methods are not positive definite anymore. As a consequence, Lagrange-Newton methods, even if equipped with some damping are usually not appropriate for finding minimizers of non-convex problems. Various algorithmic techniques for non-convex optimization have been developed in the last few decades (cf. e.g. [6]).

Two popular techniques for large scale problems are truncated conjugate gradients, where the cg method is performed, until a direction of negative curvature is detected and hessian modification, where a positive definite term is added to the hessian, such that the sum is positive definite. In our context, the latter strategy yields markedly more robust behaviour, if applied appropriately. For this it is important that the regularization term is chosen adequately, taking into account the underlying functional analytic structure.

While truncated cg methods typically yield cheaper steps, the computed search directions are often very irregular (an effect that is not present in \mathbb{R}^n) and thus yield very small damping parameters and many outer iterations. The reason is that steps result from an algebraic computation that is just terminated at a point where things became particularly difficult with often irregular cg-iterates.

On the contrary, an appropriate hessian modification yields a well defined problem in function space. For example, a regularized elastic problem is still an elastic problem, however with a different, stiffer material. Hence, solutions of regularized problems typically have better regularity properties than an arbitrary element of the energy space. This is in agreement with the above observation that additional regularity of solutions helps to bridge the problem inherent norm-gap.

Non-convexity of the Objective. If $L_{xx}(x, p)$ is not positive definite on $\ker c'_\gamma(x)$, then (9) does not correspond to a quadratic minimization problem and descent of tangential steps is not guaranteed, unless appropriate modifications are made. Just as described above, we use a hessian modification approach. Instead of solving (9) we solve the modified problem:

$$\begin{pmatrix} L_{xx}(x, p) + \lambda M & c'_\gamma(x)^* \\ c'_\gamma(x) & 0 \end{pmatrix} \begin{pmatrix} \delta t \\ q \end{pmatrix} + \begin{pmatrix} L_x(x, p) + L_{xx}(x, p)\delta n \\ 0 \end{pmatrix} = 0, \quad (19)$$

where M is the Riesz-isomorphism that is used to define normal steps and $\lambda \geq 0$ is an algorithmic parameter that is chosen large enough to render $L_{xx}(x, p) + \lambda M$ positive definite on $\ker c'_\gamma(x)$.

This system corresponds to the following minimization problem

$$\min_{\delta t} f'(x)\delta x + \frac{1}{2}(L_{xx}(x, p) + \lambda \tilde{M})(\delta x, \delta x) \quad \text{s.t.} \quad c'(x)\delta t = 0,$$

where $\tilde{M} = M$ on $\ker c'_\gamma(x)$, but $\tilde{M} = 0$ on $(\ker c'_\gamma(x))^\perp$. Our quadratic model is thus mainly modified on $\ker c'_\gamma(x)$ but not on its orthogonal complement.

In our current implementation each step is started with $\lambda = 0$. If non-convexity is encountered, $\lambda > 0$ is chosen and a new attempt to compute a step is made. Subsequently, λ is increased, until directions of negative curvature are no longer encountered.

Non-convexity of the Energy. If models for elastic materials are intended to be realistic for large deformations, they have to be non-convex. A classical and practically relevant example, caused by non-convexity is buckling. It can be observed, if compressive forces act on the opposite ends of a slim body. For small forces the body is compressed in the direction of force, but as forces increase, this state becomes unstable and energy is decreased if the body is bent in some direction. Depending on the symmetries of the body the new energy minimizers may be non-unique, and they may differ dramatically from the previous solution. Seemingly stable structures collapse suddenly, if a certain critical force is exceeded.

If such a behaviour is encountered during the course of solution of an optimal control problem, which is the case for the problem, described in Figure 4, a couple of numerical difficulties arise. First of all, the operator A in (15) is likely to be indefinite, so (15) does not correspond to a quadratic energy minimization problem and also cannot be solved by a cg-method. Furthermore, solutions of the nonlinear

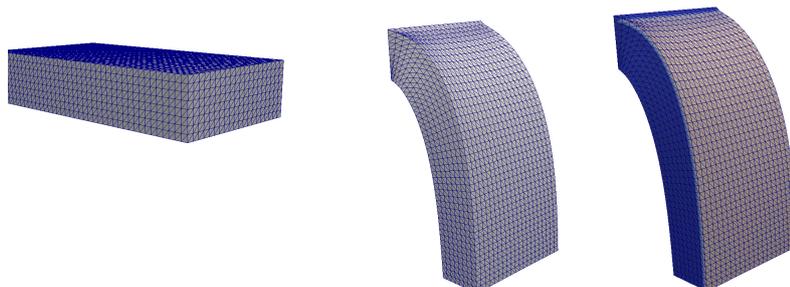


FIGURE 4. Problem of bending down a horizontal cantilever. Left: Undeformed domain. Middle: Desired deformation y_d . Right: Optimal deformation. Upper/front horizontal boundary: Γ_U , colour codes intensity of applied forces. Rear vertical boundary: Γ_D . All other boundaries: $\Gamma_N \setminus \Gamma_U$.

equation $c_\gamma(y, u) = 0$ cease to be energy minimizers or change rapidly, if u is perturbed.

In analogy to the regularization of the objective function we apply a regularization term to the energy functional. This has to be done in such a way that δn , δt and δs can be computed in a consistent way.

Assume that at the iterate (y_k, u_k) the operator $A = A'_\gamma(y_k)$ is not positive definite. This can be detected during the attempt to solve (13) or (15) by a cg-method. In that case, we choose a regularization factor $\lambda > 0$, and define a regularized energy functional as follows:

$$\hat{\mathcal{E}}_\gamma(y, u) := \mathcal{E}_\gamma(y, u) + \frac{\lambda}{2}q(y - y_k),$$

where q is a quadratic, positive definite energy. For our computations we have chosen $q(v) = \langle \nabla v, \nabla v \rangle_{L_2}$. For this modification we compute:

$$\begin{aligned} \partial_y \hat{\mathcal{E}}_\gamma(y, u) &= \partial_y \mathcal{E}_\gamma(y, u) + \lambda q'(y - y_k) \Rightarrow \partial_y \hat{\mathcal{E}}_\gamma(y_k, u_k) = \partial_y \mathcal{E}_\gamma(y_k, u_k), \\ \hat{A} &:= \partial_{yy}^2 \hat{\mathcal{E}}_\gamma(y_k, u_k) = \partial_{yy}^2 \mathcal{E}_\gamma(y_k, u_k) + \lambda q''(0). \end{aligned}$$

The effect of this regularization is threefold: first of all, if λ is sufficiently large to render \hat{A} is positive definite, then the solution of (13) or (15) with A replaced by \hat{A} is a minimization problem. Thus conjugate gradients can be applied. Second, due to ellipticity of \hat{A} , normal steps are shifted towards descent for $\mathcal{E}_\gamma(y, u)$ at

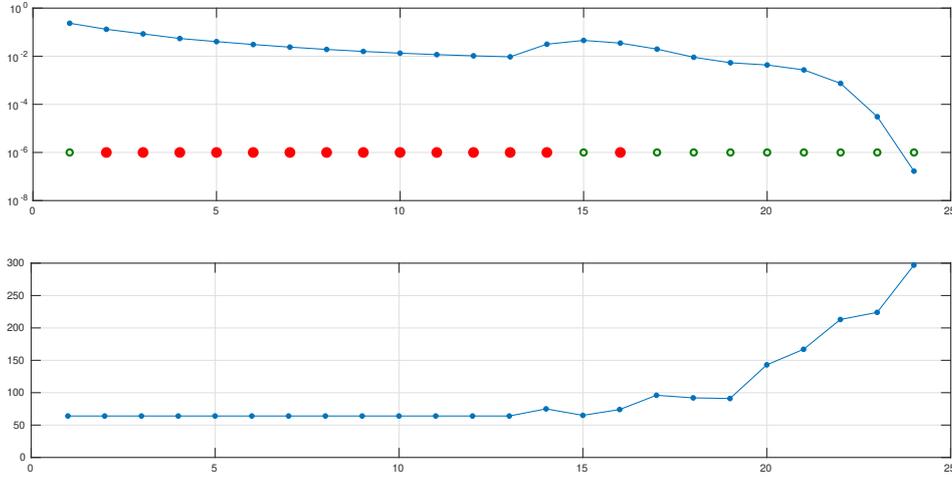


FIGURE 5. Iteration history for problem from Figure 4: Top: norms of steps, taken by the algorithm. Green circles: convex energy, red dots: non-convex energy. Bottom: total number of outer cg-Iterations in each step

(y_k, u_k) , because the linearized constraint imposed on δn reads:

$$\hat{A}\delta n_y - B\delta n_u + \partial_y \mathcal{E}_\gamma(y_k, u_k) = 0.$$

Since $\mathcal{E}_\gamma(y_k, u_k)$ is linear in u , we conclude:

$$\partial_y \mathcal{E}_\gamma(y_k, u_k + \delta n_u)\delta n_y = (\partial_y \mathcal{E}_\gamma(y_k, u_k) - B\delta n_u)\delta n_y = -(\hat{A}\delta n_y)\delta n_y < 0.$$

So δn_y is a descent direction for the total energy at the point $(y_k, u_k + \delta n_u)$. Thus, finding energy minimizers is promoted. Third, long steps are penalized, which results in a more stable behaviour of the optimization algorithm in the presence of an instability of the elastic problem.

As a numerical example we consider the problem, described in Figure 4. In Figure 6 we see some of the iterates taken by our optimization algorithm for a problem, where non-convex behaviour of the energy functional is encountered. In the beginning of the iteration, a buckling type non-convexity is encountered. We observe that the applied forces in the early phase of the algorithm are rather intense, since the material resists the applied compressive forces. Also the necessary regularization of the energy adds some artificial stiffness to the material. After

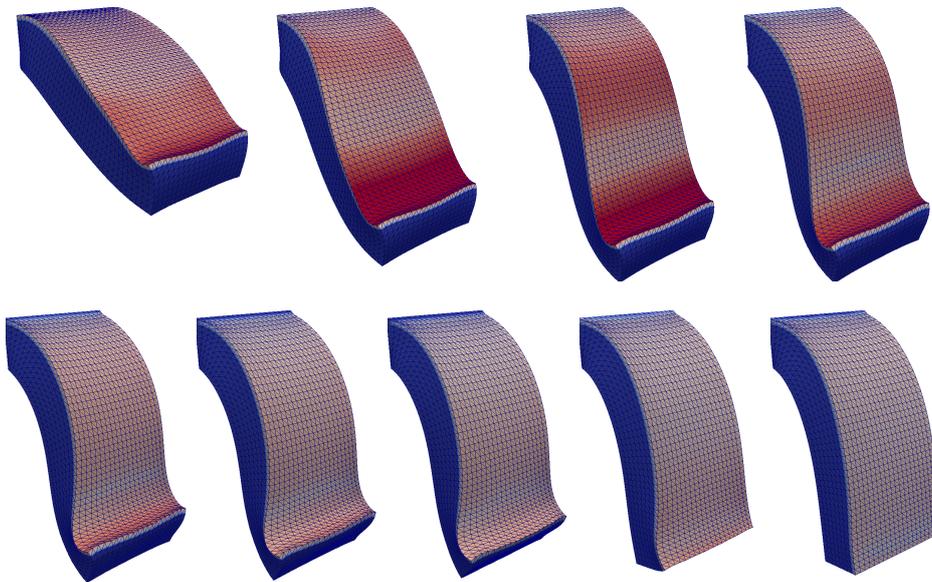


FIGURE 6. Iterates taken by the composite step method for problem from Figure 4 without contact. The colour codes the intensity of the forces. Top row: iterates 1,2,3,4. Bottom row: iterate 7,10,13,16,19.

some bending has taken place, the algorithm finds the optimal solution using comparably small forces.

An interesting effect is the S-shape shown by some intermediate deformations. This is a consequence of the nonlinearity of the problem, which occurs in particular for boundary forces. A tangential traction force tries to push the body towards the desired deformation. Due to the nonlinearity of the problem, however, the body is bent in upward direction, due to the moment introduced by this force.

Remark 4.3. *An alternative to the presented idea was also tested. If during the algorithm non-convexity of the energy was encountered at (y_k, u_k) , the control u_k was kept fixed and an energy minimization algorithm, based on the ideas of [26], was applied to compute a minimizer of $\mathcal{E}_\gamma(\cdot, u_k)$ was computed. So we temporarily switched to a black-box method. At least for our test case, the performance of this variant was not satisfactory. It showed a rather unstable and erratic behaviour. The reason for this seems to be that buckling can occur during such an algorithm.*

4.5. Path-following

With a robust solver for (6) at hand, we can now use a path-following method to approximate solutions of the original optimal control problem with contact. We use a simple approach, where after (6) has been solved for some γ_k , the regularization parameter is multiplied by some fixed factor $s > 1$. A choice of $s = 10$ has proven quite appropriate.

We added contact constraints to the problem, described in Figure 4. An illustration of the path-following procedure is given in Figure 7. Obviously, the regularization procedure works as intended. For moderate γ , the contact constraint is clearly violated, but γ becomes larger this violation gradually vanishes. We also observe that our method is well capable to deal with large deformations and strains.

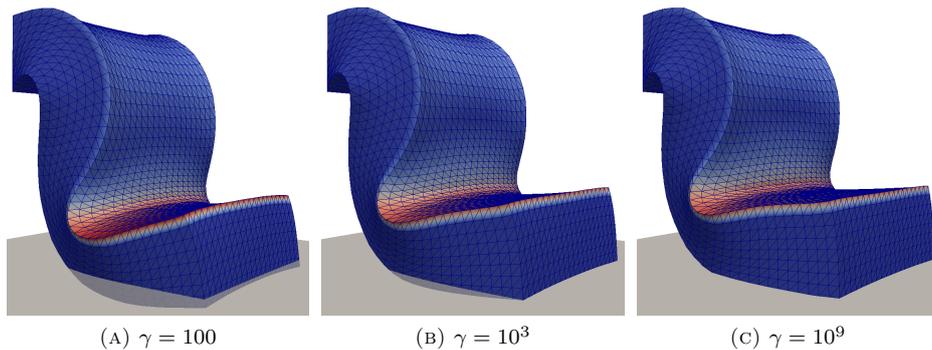


FIGURE 7. Optimal deformations with penalty parameter γ for problem from Figure 4 with contact. Colour codes intensity of the forces.

A detailed discussion of the problem from Figure 1 with added contact constraints, including convergence plots, can be found in [22].

5. Conclusion and Outlook

We conclude that optimal optimal control problems with finite strain hyperelastic materials and contact offer a broad range of challenges, concerning both theoretical and algorithmic aspects.

The main theoretical challenge is that not much analytic structure is available to build a theory upon. The main result of polyconvexity, a weak lower semi-continuity property of the energy functional could be exploited to conclude results about the path of regularized solutions. However, satisfactory, stronger results were only possible by employing refined techniques. Still there are many questions that remain open, most importantly a local sensitivity result that could permit a rigorous derivation of optimality conditions for the optimal control problem.

From a numerical view point this class of problems combines high nonlinearity and non-convexity with large scale. To obtain efficient and robust solution algorithms, significant advances had to be made, compared to generic optimization methods. Decisive ingredients are a good choice of functional analytic framework, a sound concept for inexact computation of steps by iterative solvers, and a proper treatment of non-convexities, both in the objective and in the energy. In this paper we concentrated on these computations aspects. An observed key ingredient to efficient algorithmic behaviour is to produce regular steps where possible.

A couple of algorithmic concepts are subject to current work. First, as pointed out in Remark 4.1 new ideas the use of iterative solvers for the A -block in (12) are currently under investigation. Second, the nonlinearity of finite deformation problems exhibits some very interesting geometrical structure. The analysis hints on using nonlinear updates, instead of the usual linear ones. Currently, promising numerical results that go into this direction are available for solving the energy minimization problem. The application of this concept to optimal control problems is still subject to current research and is planned to be published in a forthcoming paper.

As a future perspective, the algorithmic solution of our class of problems has to be extended to real world applications. A particular example are inverse problems in the context of biomechanics, where elastic contact problems occur in joints. In addition to the described difficulties, an envisioned solution algorithm will have to deal with complicated contact geometries.

References

- [1] John M. Ball. Convexity conditions and existence theorems in nonlinear elasticity. *Archive for Rational Mechanics and Analysis*, 63(4):337–403, 1977.

- [2] Thomas Betz. *Optimal control of two variational inequalities arising in solid mechanics*. PhD thesis, 2015.
- [3] P.G. Ciarlet. *Mathematical Elasticity: Three-dimensional elasticity*. Number Bd. 1. North-Holland, 1994.
- [4] Philippe G. Ciarlet and Jindřich Nečas. Unilateral problems in nonlinear, three-dimensional elasticity. *Archive for Rational Mechanics and Analysis*, 87(4):319–338, 1985.
- [5] Marius Cocu. Existence of solutions of signorini problems with friction. *International journal of engineering science*, 22(5):567–575, 1984.
- [6] A.R. Conn, N.I.M. Gould, and P.L. Toint. *Trust-Region Methods*. SIAM, 2000.
- [7] Peter Deuffhard. *Newton Methods for Nonlinear Problems: Affine Invariance and Adaptive Algorithms*. Springer Publishing Company, Incorporated, 2011.
- [8] N.I.M. Gould, M.E. Hribar, and J. Nocedal. On the solution of equality constrained quadratic programming problems arising in optimization. *SIAM Journal on Scientific Computing*, 23(4):1376–1395, 2001.
- [9] Andreas Günnel and Roland Herzog. Optimal control problems in finite-strain elasticity by inner pressure and fiber tension. *Frontiers in Applied Mathematics and Statistics*, 2:4, 2016.
- [10] M. Heinkenschloss and D. Ridzal. A matrix-free trust-region SQP method for equality constrained optimization. *SIAM J. Optim.*, 24(3):1507–1541, 2014.
- [11] M. Heinkenschloss and L.N. Vicente. Analysis of inexact trust-region SQP algorithms. *SIAM J. Optim.*, 12(2):283–302, 2001/02.
- [12] Noboru Kikuchi and John Tinsley Oden. *Contact problems in elasticity: a study of variational inequalities and finite element methods*, volume 8. siam, 1988.
- [13] Lars Lubkoll. *An Optimal Control Approach to Implant Shape Design : Modeling, Analysis and Numerics*. PhD thesis, Bayreuth, 2015.
- [14] Lars Lubkoll, Anton Schiela, and Martin Weiser. An optimal control problem in polyconvex hyperelasticity. *SIAM J. Control Opt.*, 52(3):1403 – 1422, 2014.
- [15] Lars Lubkoll, Anton Schiela, and Martin Weiser. An affine covariant composite step method for optimization with pdes as equality constraints. *Optimization Methods and Software*, 32:1132–1161, 2017.
- [16] J.A.C. Martins and J.T. Oden. Existence and uniqueness results for dynamic contact problems with nonlinear normal and friction interface laws. *Nonlinear Analysis: Theory, Methods and Applications*, 11(3):407 – 428, 1987.
- [17] Georg Müller and Anton Schiela. On the control of time discretized dynamic contact problems. *Computational Optimization and Applications*, 68(2):243–287, Nov 2017.
- [18] J.T. Oden and J.A.C. Martins. Models and computational methods for dynamic friction phenomena. *Computer Methods in Applied Mechanics and Engineering*, 52(1):527 – 634, 1985.
- [19] E. O. Omojokun. *Trust Region Algorithms for Optimization with Nonlinear Equality and Inequality Constraints*. PhD thesis, Boulder, CO, USA, 1989. UMI Order No: GAX89-23520.
- [20] D. Ridzal. *Trust-region SQP methods with inexact linear system solves for large-scale optimization*. ProQuest LLC, Ann Arbor, MI, 2006. Thesis (Ph.D.)–Rice University.

- [21] Manuel Schaller, Anton Schiela, and Matthias Stöcklein. A composite step method with inexact step computations for PDE constrained optimization. Preprint SPP1962-098, 10 2018.
- [22] Anton Schiela and Matthias Stöcklein. Optimal control of static contact in finite strain elasticity. Preprint SPP1962-097, 10 2018.
- [23] Antonio Signorini. Sopra alcune questioni di elastostatica. *Atti della Societa Italiana per il Progresso delle Scienze*, 1933.
- [24] A. Vardi. A trust region algorithm for equality constrained minimization: convergence properties and implementation. *SIAM J. Numer. Anal.*, 22(3):575–591, 1985.
- [25] Jan Christoph Wehrstedt. *Formoptimierung mit Variationsungleichungen als Nebenbedingung und eine Anwendung in der Kieferchirurgie*. PhD thesis, 2007.
- [26] M. Weiser, P. Deuffhard, and B. Erdmann. Affine conjugate adaptive Newton methods for nonlinear elastomechanics. *Opt. Meth. Softw.*, 22(3):414–431, 2007.
- [27] J.C. Ziemis and S. Ulbrich. Adaptive multilevel inexact SQP methods for PDE-constrained optimization. *SIAM J. Optim.*, 21(1):1–40, 2011.

Anton Schiela
Mathematisches Institut
Universität Bayreuth
95440 Bayreuth
Germany
e-mail: anton.schiela@uni-bayreuth.de

Matthias Stöcklein
Mathematisches Institut
Universität Bayreuth
95440 Bayreuth
Germany
e-mail: matthias.stoecklein@uni-bayreuth.de