Submitted to *Mathematical and Computer Modelling of Dynamical Systems*
Vol. 00, No. 00, month 20xx, 1–34

Optimal Control of an Elastic Crane-Trolley-Load System – A Case Study for Optimal Control of Coupled ODE-PDE Systems – (Extended version with two appendices)

S.-J. Kimmerle[a*], M. Gerdts[b], and R. Herzog[c]

[a]*Institut für Mathematik und Bauinformatik, Universität der Bundeswehr München, Neubiberg/München, Germany;* [b]*Institut für Mathematik und Rechneranwendung, Universität der Bundeswehr München, Neubiberg/München, Germany;* [c]*Fakultät für Mathematik, Technische Universität Chemnitz, Germany;*

(*v0.9 released July 2016*)

We present a mathematical model of a crane-trolley-load model, where the crane beam is subject to the partial differential equation (PDE) of static linear elasticity and the motion of the load is described by the dynamics of a pendulum that is fixed to a trolley moving along the crane beam. The resulting problem serves as a case study for optimal control of fully coupled partial and ordinary differential equations (ODEs). This particular type of coupled systems arises from many applications involving mechanical multi-body systems.

We motivate the coupled ODE-PDE model, show its analytical well-posedness locally in time and examine the corresponding optimal control problem numerically by means of a projected gradient method with BFGS update.

Keywords: crane-trolley-load system; ODE-PDE-constrained optimization; optimal control of ODE; coupled ODE-PDE system; projected gradient method; FEM for moving boundary conditions

AMS Subject Classification: 90C11; 49M25; 49L20

1. Introduction

In this article we consider a crane model, where the crane arm is modelled by an elastic beam Ω fixed at one end to the crane tower. The load is represented as a mathematical pendulum that is fixed to a trolley which moves along the crane beam. The mechanical displacement $\boldsymbol{u}$ of the crane beam is determined by the elliptic partial differential equation of static linear elasticity. The trolley and load states $\boldsymbol{q} = (q_1, q_2)$, that is, the position of the trolley and the angle of the load are subject to an ordinary differential equation. It turns out that all differential equations are fully coupled with one another. The goal is to transport the load by means of the trolley along the crane beam from a given initial position $\boldsymbol{q}^0$ and initial velocity $\boldsymbol{v}^0$ to a designated terminal position $\boldsymbol{q}^f$ with terminal velocity $\boldsymbol{v}^f$, while minimizing vibrations as well as the total time T. Here the control is given by the acceleration force U of the trolley. For the relevance of this optimal control problem

*Corresponding author. Email: sven-joachim.kimmerle@unibw.de

in engineering, see [1, 2]. The optimal control of a gantry crane has been examined by Biswas [3] who considers additionally the displacement of a non-rigid cable fixing the load, while the rails are considered as unflexible. This different model yields a coupled ODE-PDE system as well, but with a one-dimensional PDE.

For the geometry of our model, we refer to Fig. 1. Our optimal control problem (OCP) is to find

$$\min_{U,T,\boldsymbol{q}} \; J(\boldsymbol{q}, \dot{\boldsymbol{q}}, U, T) \tag{1}$$

with the objective

$$\begin{aligned} J(\boldsymbol{q}, \dot{\boldsymbol{q}}, U, T) = \nu_1 T + \frac{\nu_2}{2} \|\dot{q}_2\|^2_{L^2(0,T)} + \frac{\nu_3}{2} \|U\|^2_{L^2(0,T)} \\ + \sum_{i=0}^{1} \frac{\nu_{4+i}}{2} |q_i(T) - q_i^f|^2 + \sum_{i=0}^{1} \frac{\nu_{6+i}}{2} |\dot{q}_i(T)|^2 \end{aligned} \tag{2}$$

subject to the elliptic PDE

$$-\operatorname{div} \sigma(\boldsymbol{u}) = \boldsymbol{H} \quad \text{in } \Omega \times [0,T] \tag{3}$$

and the ODE system

$$\mathbf{M}(\boldsymbol{q}, \bar{\boldsymbol{u}})\, \ddot{\boldsymbol{q}} = \boldsymbol{F}(\boldsymbol{q}, \dot{\boldsymbol{q}}, \bar{\boldsymbol{u}}, U) \quad \text{in } [0,T], \tag{4}$$

and the PDE and ODE are completed by boundary and initial conditions, respectively. Here $\boldsymbol{\nu} \in \mathbb{R}^7$ denotes a non-zero vector of non-negative weights. The state $\boldsymbol{q}(t)$ denotes the vector of generalized coordinates of the rigid bodies at time t (see Sect. 2.3). The right hand side $\boldsymbol{H}$ encodes the gravitational forces due to the weight of the beam itself, $\mathbf{M}$ the mass matrix, and $\boldsymbol{F}$ comprises generalized Coriolis forces and external forces. We mention as a particular feature of our model that only mean values $\bar{\boldsymbol{u}}$ of the beam's displacement $\boldsymbol{u}$, averaged over the surface $\Gamma_C(q_1) = q_1 + \Gamma_C(q_1^0)$ connecting the trolley to the beam (for the precise definition see Fig. 1 and Sect. 2.1), enter into the ODE (4). The differential equations are completed by boundary and initial conditions. Time-optimal control is realized by scaling to a fixed time interval, yielding the total time T only as a control parameter. The scaled full problem is stated precisely in Section 2.

Our problem class cannot be solved directly by standard software owing to the strong coupling of the constraints. In this study we prove the local-in-time well-posedness of the coupled dynamical problem and present an algorithm for solving the OCP, based on a first-discretize-then-optimize (FDTO) approach. The non-standard Algorithms 4.1–4.4 are designed specifically for our problem and are based on a projected gradient method without and with BFGS update, respectively. We emphasize that it is not clear whether Newton-based methods as in Chen and Gerdts [4, 5] could be applied in our situation. On the one hand it is not obvious whether our problem exhibits the smoothness needed for Newton methods, and on the other hand this approach requires the solution of the necessary optimality conditions that turns out to be challenging. Finally we present numerical results for the optimal control of the crane-trolley-load system. The open-source finite element software FEniCS [6] is employed to solve the PDEs of linear elasticity. Special attention has

to be given to the elements at surfaces where forces are applied, that is, at the free boundary between moving trolley and crane beam. For ease of presentation we focus in this article on a crane that does not rotate.

In the modelling of multi-body systems we often encounter the following situation. We find ODEs representing the interactions between the centers of mass, algebraic equations from constraining forces and elliptic PDEs modelling mechanical deformations within the bodies. Here we focus on the crane-trolley-load system serving as a case study. Other applications include, for example, (a) a quarter-car model, where an elastic wheel-tyre-damper system with free road contact is controlled [7], (b) the heat-optimal ascent and re-entry into atmosphere of a hypersonic spacecraft [8], and (c) a truck or plane with a tank filled with fluid [9, 10]. In the latter example the PDEs are the St-Venant equations yielding the height and velocity of the fluid. However, it should be emphasized that the class of optimal control problems subject to ODE and PDE constraints is heterogeneous, since different types of PDE require already different theories and methods. Furthermore the methods for this problem class depend on the particular coupling structure between ODE and PDE.

On optimal control of PDE and optimal control of differential-algebraic equations (DAE) alone there exists a wide variety of results and numerical approaches, see for example, the overview article [11] or the textbooks [12–15] for control of PDEs and [16, 17] for optimal control of ODE and DAE, respectively.

However, only few results about *combined* ODE-PDE constrained optimal control exist so far. Biswas *et al.* [18, 19] examine control and optimal control of a large flexible space structure, e.g. a satellite, subject to PDE modelling vibrations and to ODE describing the dynamics. Chudej *et al.* [8] consider optimal control for coupling of the heat equation and equations of motions. Similar as in our situation, the coupling from ODE to PDE is effected by means of a boundary condition, but the controls arise in the ODE system only and the PDE is considered in one space dimension. In our problem, the controls arise in the ODE as well as in the boundary condition. Our control acts on the PDE by means of a Neumann boundary control. In our problem we encounter control constraints, too. For a particular class of ODE-PDE-problem, a so-called hypersonic rocket car problem, including the problem of [8], some new phenomena have been discovered by Pesch *et al.* [20]. In [21] this OCP is reformulated as a state constrained optimal control problem for PDE and necessary optimality conditions for it are derived.

Our study is organized as follows. In Section 2 we derive the mathematical model. The local-in-time well-posedness of our model is shown in Section 3. Due to the complexity of the resulting coupled ODE-PDE problem, we apply here the direct FDTO method, and follow a sensitivity-based approach. The discretization for the numerical simulation, the projected gradient method, and the quasi-Newton method BFGS for the optimal control are described in Section 4. In Section 5, numerical results are presented. Finally we close with a discussion of our results and give, in particular, an outlook on open questions and future work.

2. Mathematical model of the elastic crane

2.1. *Geometry*

For the geometry of a crane in the undeformed configuration, see Fig. 1. In this study we consider no rotation of the crane beam, yielding that the pendulum is

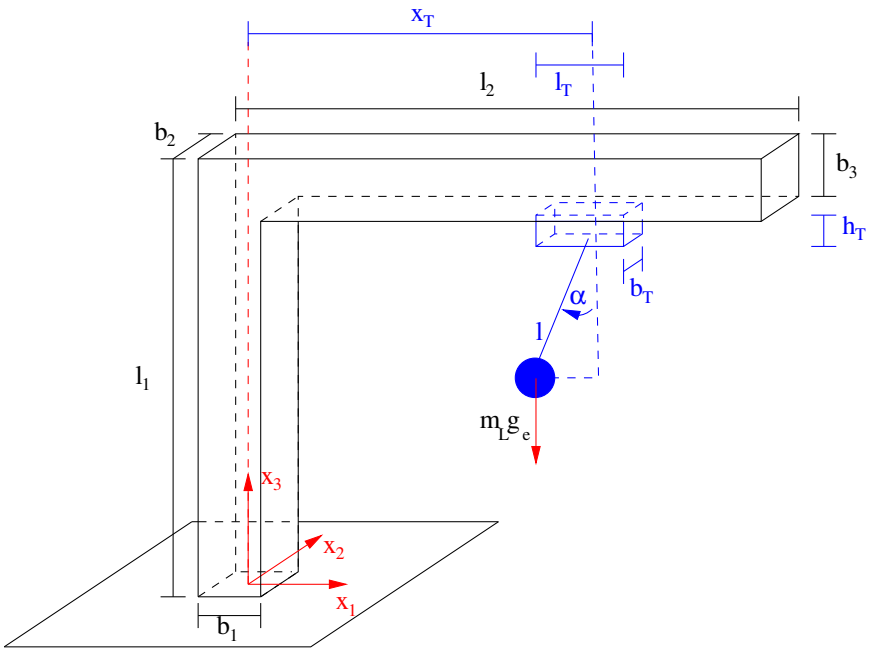

Figure 1. Configuration of the elastic crane (within Lagrangian coordinates).

restricted to the plane defined by $x_2 \equiv 0$. The full 3d model with a rotating crane is presented in the study [23].

The deformations of the crane beam are here considered in a full 3d model avoiding the issue, how to replace the crane beam by a lower dimensional model sensitively. The domain

$$\Omega = \{\boldsymbol{x} = (x_1, x_2, x_3) \in \mathbb{R}^3 \,|\, b_1/2 \le x_1 \le l_2 - b_1/2,\ |x_2| \le b_2/2,\ l_1 - b_3 \le x_3 \le l_1\},$$

where $b_1 < l_2$, $b_3 < l_1$, describes the undeformed extension arm of the crane, which is considered elastic and attached to the rigid vertical bar. On $\Gamma_D = \partial\Omega \cap \{\boldsymbol{x} \in \mathbb{R}^3 \,|\, x_1 = b_1/2\}$ the extension arm is fixed.

A force is applied on the trolley at its center of mass, located at

$$\boldsymbol{r}_{\mathbf{M}}(t) = \left(x_{\mathrm{T}}(t),\ 0,\ l_1 - b_3 - \frac{h_{\mathrm{T}}}{2}\right)^{\top} \tag{5}$$

in the undeformed state. This load is a pendulum of length l with a point mass m_{L}. The trolley, with mass m_{T}, may move on rails, modelled by translating the x_1-position $x_{\mathrm{T}}(t) \in ((b_1+l_T)/2, l_2-(b_1+l_T)/2)$ of its center of mass. The application of a force U, e.g., by means of a neglectable cable along the beam, at the trolley's center of gravity allows to control the trolley. Alternatively, if the trolley was controlled by a motor within the trolley, moving a wheel that is in contact with rails at the crane beam, this would yield the same model, but with a slightly different area of support Γ_{C}, now being the wheel's area of support. On the boundary $\Gamma_{\mathrm{C}}(x_{\mathrm{T}}) = \{\boldsymbol{x} \in \partial\Omega \,|\, x_{\mathrm{T}}(t)-l_{\mathrm{T}}/2 \le x_1 \le x_{\mathrm{T}}(t)+l_{\mathrm{T}}/2,\ |x_2| \le b_{\mathrm{T}}/2,\ x_3 = l_1-b_3\} = \Gamma_{\mathrm{C}}(x_{\mathrm{T}}(0))+x_{\mathrm{T}}\,\boldsymbol{e}_1$ varying with time, the trolley exerts a force on the extension arm. As usual $\boldsymbol{e}_1$ denotes the unit vector in x_1-direction. In particular, the definition of Γ_{C} implies

that the surface area of Γ_{C}, denoted by $|\Gamma_{\mathrm{C}}|$, remains constant. The application of a moment U_{M} at the trolley's center of gravity allows to rotate the crane by an angle β. We assume $\alpha \in (-\pi/2, \pi/2)$.

2.2. *Strains and stresses in the beam*

Within the domain Ω of the crane we aim to solve for the mechanical displacement field. For a realistic crane we may assume small displacement gradients and, hence, model the elastic deformations within the structure of the crane by linear elasticity. As usual in linear elasticity, we do not distingiush between the *reference* (undeformed) configuration Ω (in Lagrangian coordinates $\boldsymbol{x}$) and the deformed configuration $\Omega' \subset \mathbb{R}^3$ (in Eulerian coordinates $\hat{\boldsymbol{x}}$). As reference configuration we consider the crane's extension arm in the absence of strains or stresses. For the interaction between the deformed beam and the trolley, the deformation of the beam is not neglected in our study. In this context, we consider u in the reference configuration. For the deformed configuration the formula (19) would look different, though leading to the same results within the approximation of small displacement gradients. The deformation depends on time through the control. The system is considered on the compact time interval $[0, T] \subset \mathbb{R}$ and thus the mechanical displacement field reads $\boldsymbol{u} : \Omega \times [0, T] \to \Omega'$, $(\boldsymbol{x}, t) \mapsto \boldsymbol{u}(\boldsymbol{x}, t) = \hat{\boldsymbol{x}}(\boldsymbol{x}, t) - \boldsymbol{x}$. The symmetrized strain associated with the displacement field $\boldsymbol{u}$ is $\epsilon(\boldsymbol{u}) = (\boldsymbol{\nabla u} + \boldsymbol{\nabla u}^\top)/2$ and as a constitutive assumption we work with the Cauchy stress tensor

$$\sigma(\boldsymbol{u}) = \lambda \operatorname{trace}(\epsilon(\boldsymbol{u}))\mathbf{1} + 2\,\mu\,\epsilon(\boldsymbol{u}), \tag{6}$$

where $\mathbf{1}$ denotes the unit matrix and $\mu > 0$, $\lambda > -2\mu/3$ are the Lamé constants scaled by $1/(l\, m_{\mathrm{L}})$. For our purposes it suffices to consider $\lambda > 0$. We assume that we may neglect the deformation of the trolley, since typically $h_{\mathrm{T}} \ll b_3$ holds.

Terms of higher order in $\|\boldsymbol{\nabla u}\|$ are neglected, since $\|\boldsymbol{\nabla u}\| \ll 1$. (This is emphasized in the following by the $\approx$ symbol.) For more details see [22, Ch. 3].

2.3. *Governing equations for the trolley and the load*

We introduce the generalized coordinates $\boldsymbol{q} = (x_{\mathrm{T}}, \alpha)^\top$, see Fig. 1. The trolley's center of gravity at time t is given by

$$\boldsymbol{r}_{\mathrm{T}}(q_1, \boldsymbol{u}, t) = \boldsymbol{r}_{\mathrm{M}}(q_1) + \boldsymbol{u}(r_{\mathrm{M}}(q_1), t), \tag{7}$$

where $\boldsymbol{r}_{\mathrm{M}}$ is defined by (5). The position of the load, considered as point mass, at time t is

$$\boldsymbol{r}_{\mathrm{L}}(\boldsymbol{q}, \boldsymbol{u}, t) = \boldsymbol{r}_{\mathrm{T}}(q_1, \boldsymbol{u}, t) + \begin{pmatrix} l \sin q_2 \\ 0 \\ -l \cos q_2 \end{pmatrix},$$

where $q_2 = \alpha$ is positive for a counterclockwise rotation around the x_2-axis, see Fig. 1. The kinetic energy of the mechanical system written for the generalized

coordinates $\boldsymbol{q} = (x_\mathrm{T}, \alpha)^\top$ and scaled by $1/(l\, m_\mathrm{L})$ is given by

$$\mathcal{T}(\boldsymbol{q}, \dot{\boldsymbol{q}}, \boldsymbol{u}, t) = \frac{1}{2\, l\, m_\mathrm{L}} \left(m_\mathrm{T} \|\dot{\boldsymbol{r}}_\mathrm{T}(q_1, \boldsymbol{u}, t)\|^2 + m_\mathrm{L} \|\dot{\boldsymbol{r}}_\mathrm{L}(\boldsymbol{q}, \boldsymbol{u}, t)\|^2 \right), \tag{8}$$

where we neglect the moment of inertia of the trolley and the load. Furthermore, we have neglected velocities

$$\dot{\boldsymbol{u}} \approx \boldsymbol{0} \tag{9}$$

in (8), which is motivated by a short dimensional analysis (see Subsection 2.4) and is consistent as we will see in our numerics (see, e.g., Fig. 3, bottom left). For brevity, we set $m := (m_\mathrm{T} + m_\mathrm{L})/(l\, m_\mathrm{L})$. The scaled, generalized potential $\mathcal{V}$, uniquely determined up to a constant, is $\mathcal{V}(\boldsymbol{q}, U) = -U q_1 - g_\mathrm{e} \cos q_2$, where U is the control (divided by $l\, m_\mathrm{L}$) and g_e is the gravity acceleration. From $\boldsymbol{E} = -\boldsymbol{\nabla}_{\boldsymbol{q}} \mathcal{V}$ we obtain the vector of applied generalized forces

$$\boldsymbol{E}(\boldsymbol{q}) = \begin{pmatrix} U \\ -g_\mathrm{e} \sin q_2 \end{pmatrix}, \tag{10}$$

that is, the control acting along the deformed rail of the trolley and gravitation acting on the deformed system. The Lagrange function is $\mathcal{L}(\boldsymbol{q}, \dot{\boldsymbol{q}}, \boldsymbol{u}, U) = \mathcal{T}(\boldsymbol{q}, \dot{\boldsymbol{q}}, \boldsymbol{u}) - \mathcal{V}(\boldsymbol{q}, U)$.

By means of the Euler-Lagrange equation $\frac{\mathrm{d}}{\mathrm{d}t} \boldsymbol{\nabla}_{\dot{\boldsymbol{q}}} \mathcal{L} = \boldsymbol{\nabla}_{\boldsymbol{q}} \mathcal{L}$ for a given control we derive the equations of motion within our approximation of small displacement gradients. Neglecting $\dot{\boldsymbol{u}}$, we obtain from (7) the time derivative

$$\dot{\boldsymbol{r}}_\mathrm{T} = \mathbf{F}_\mathrm{D}(\boldsymbol{\nabla} \boldsymbol{u})^\top \dot{q}_1 \boldsymbol{e}_1,$$

where $\mathbf{F}_\mathrm{D}(\boldsymbol{\nabla} \boldsymbol{u}) = \mathbf{1} + \boldsymbol{\nabla} \boldsymbol{u}$ is the deformation gradient. In our approximation of small displacement gradients we have

$$F^2_{D,11}(\boldsymbol{\nabla} \boldsymbol{u}) + F^2_{D,21}(\boldsymbol{\nabla} \boldsymbol{u}) + F^2_{D,31}(\boldsymbol{\nabla} \boldsymbol{u}) = 1 + 2\partial_1 u_1 + \sum_{i=1}^{3} |\partial_i u_1|^2 \approx 1 + 2\partial_1 u_1.$$

Furthermore, $\dot{\boldsymbol{r}}_\mathrm{L} = \dot{\boldsymbol{r}}_\mathrm{T} - l\, (\cos q_2, 0, \sin q_2)^\top \dot{q}_2$ and the scaled kinetic energy reads

$$\begin{aligned} \mathcal{T}(\boldsymbol{q}, \dot{\boldsymbol{q}}, \boldsymbol{u}, t) = {} & \frac{m}{2} \left(1 + 2\partial_1 u_1(\boldsymbol{r}_\mathrm{M}(q_1), t)\right) \dot{q}_1^2 + \frac{l}{2} \dot{q}_2^2 \\ & + \left((1 + \partial_1 u_1(\boldsymbol{r}_\mathrm{M}(q_1), t)) \cos q_2 + \partial_1 u_3(\boldsymbol{r}_\mathrm{M}(q_1), t) \sin q_2 \right) \dot{q}_1 \dot{q}_2. \end{aligned}$$

For brevity, we introduce the notation $\partial_i := \partial_{x_i}$ and the operator D_1 defined by

$$D_1 u_i := \partial^2_{1,1} u_i + \partial^2_{2,1} u_i + \partial^2_{3,1} u_i, \quad i = 1, 2, 3, \tag{11}$$

and abbreviate

$$\begin{aligned} \Phi_1(q_2, \mathbf{D}^2 \boldsymbol{u}) &= D_1 u_1 \cos q_2 + D_1 u_3 \sin q_2, \\ \Phi_2(q_2, \mathbf{D} \boldsymbol{u}) &= -(1 + \partial_1 u_1) \sin q_2 + \partial_1 u_3 \cos q_2. \end{aligned}$$

We compute

$$\partial_{\boldsymbol{q}}\mathcal{T}(\boldsymbol{q},\dot{\boldsymbol{q}},\boldsymbol{u}) = \begin{pmatrix} m\,D_1u_1\,\dot{q}_1^2 + \Phi_1(q_2,\mathbf{D}^2\boldsymbol{u})\,\dot{q}_1\dot{q}_2 \\ \Phi_2(q_2,\mathbf{D}\boldsymbol{u})\,\dot{q}_1\dot{q}_2 \end{pmatrix}$$

and

$$\partial_{\dot{\boldsymbol{q}},\boldsymbol{q}}\mathcal{T}\,\dot{\boldsymbol{q}} = \begin{pmatrix} 2m\,D_1u_1\,\dot{q}_1^2 + \Phi_1(q_2,\mathbf{D}^2\boldsymbol{u})\,\dot{q}_1\dot{q}_2 + \Phi_2(q_2,\mathbf{D}\boldsymbol{u})\dot{q}_2^2 \\ \Phi_1(q_2,\mathbf{D}^2\boldsymbol{u})\dot{q}_1^2 + \Phi_2(q_2,\mathbf{D}\boldsymbol{u})\dot{q}_1\dot{q}_2 \end{pmatrix}.$$

Thus the generalized Coriolis forces are

$$G(\boldsymbol{q},\dot{\boldsymbol{q}},\boldsymbol{u}) = \partial_{\boldsymbol{q}}\mathcal{T}(\boldsymbol{q},\dot{\boldsymbol{q}},\boldsymbol{u}) - \partial^2_{\dot{\boldsymbol{q}},\boldsymbol{q}}\mathcal{T}(\boldsymbol{q},\dot{\boldsymbol{q}},\boldsymbol{u})\,\dot{\boldsymbol{q}} = \begin{pmatrix} -mD_1u_1\dot{q}_1^2 - \Phi_2(q_2,\mathbf{D}\boldsymbol{u})\dot{q}_2^2 \\ -\Phi_1(q_2,\mathbf{D}^2\boldsymbol{u})\dot{q}_1^2 \end{pmatrix}.$$

This yields as equation of motion for $\boldsymbol{q}$,

$$\tilde{\mathbf{M}}(\boldsymbol{q},\boldsymbol{u})\,\ddot{\boldsymbol{q}} = \boldsymbol{G}(\boldsymbol{q},\dot{\boldsymbol{q}},\boldsymbol{u}) + \boldsymbol{E}(\boldsymbol{q},U) \tag{12}$$

where the symmetric and positive definite mass matrix reads

$$\tilde{\mathbf{M}}(\boldsymbol{q},\boldsymbol{u}) = \begin{pmatrix} m(1+2\partial_1u_1) & (1+\partial_1u_1)\cos q_2 + \partial_1u_3\sin q_2 \\ * & l \end{pmatrix}, \tag{13}$$

and where the generalized Coriolis forces are

$$\boldsymbol{G}(\boldsymbol{q},\dot{\boldsymbol{q}},\boldsymbol{u}) = \begin{pmatrix} -m\,D_1u_1\,\dot{q}_1^2 + ((1+\partial_1u_1)\,\sin q_2 - \partial_1u_3\cos q_2)\,\dot{q}_2^2) \\ -(D_1u_1\cos q_2 + D_1u_3\sin q_2)\,\dot{q}_1^2 \end{pmatrix}. \tag{14}$$

The ODE system for itself has a unique solution locally in time, since $\tilde{\mathbf{M}}$ is invertible for $\|\boldsymbol{\nabla u}\| \ll 1$.

2.4. *Governing equations for the crane beam*

For the domain of the crane's extension arm we consider the standard model of linear elasticity [22]. We do not observe significant vibrations in the beam in our simulations. Elastic waves within the crane beam may be safely neglected due to different typical time scales, see [23] for the full elastodynamical problem for the displacement field and its dimensional analysis. Indeed, the speed of shear waves is $\sqrt{l\mu/\rho} \approx 4.73$ km/s and the speed of longitudinal waves, i.e. the speed of sound, is of the same order. This elastostatic problem, where time enters as a parameter, reads

$$-\operatorname{div}\sigma(\boldsymbol{u}) = \boldsymbol{H} \qquad \text{in } \Omega\times[0,T], \tag{15}$$

$$\boldsymbol{u} = \mathbf{0} \qquad \text{on } \Gamma_{\mathrm{D}}\times[0,T], \tag{16}$$

$$-\sigma(\boldsymbol{u}).\boldsymbol{n} - \tilde{\boldsymbol{g}}(\boldsymbol{q},\boldsymbol{u},U) = \mathbf{0} \qquad \text{on } \Gamma_{\mathrm{N}}\times[0,T]. \tag{17}$$

where $\tilde{\boldsymbol{g}} : \Gamma_{\mathrm{N}}\times[0,T] \to \mathbb{R}^3$ is a boundary force (scaled by $1/(l\,m_{\mathrm{L}})$), Γ_{D} is the part of $\partial\Omega$ where the Dirichlet boundary condition (b.c.) holds and $\Gamma_{\mathrm{N}} := \partial\Omega \setminus \overline{\Gamma_{\mathrm{D}}}$ the

boundary with the Neumann b.c. The term

$$\boldsymbol{H} = -\rho\, g\, \boldsymbol{e}_3 \tag{18}$$

in (15) models the gravity of the crane cantilever. Here ρ is the mass density divided by m_{L} and $g := g_{\mathrm{e}}/l$ is the reduced gravity acceleration.

Note that we work within the approximation of small displacement gradients and, thus, for consistency the linear deformation of the crane beam has to be modelled for the trolley-load system as well, only higher order terms in $\boldsymbol{\nabla u}$ may be safely neglected. According to (10), the scaled forces $\boldsymbol{E}$ applied to the crane beam read in Cartesian coordinates

$$\boldsymbol{E}_{\mathrm{c}}(\boldsymbol{q}, U) = \begin{pmatrix} U + g \sin q_2 \\ 0 \\ -g \cos q_2 \end{pmatrix}$$

acting in the *deformed* configuration. We assume that this force is realized by a constant pressure acting on the contact surface Γ_{C} between the trolley and the beam. In order to transform a surface integral from the deformed to the reference configuration we use [22, Th. 1.7-1], also called Nanson's formula. The deformation gradient $\mathbf{F}_{\mathrm{D}}(\boldsymbol{\nabla u}) = \mathbf{1} + \boldsymbol{\nabla u}$ is invertible for $\|\boldsymbol{\nabla u}\| \ll 1$. Note that $\partial_2 u_i \approx 0$, $i = 1, 2, 3$, on Γ_{C}. In our situation the scaled gravity and control forces by the load and the trolley onto the crane beam yield $\boldsymbol{g}_0 : \mathbb{R} \times \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}^3$ by

$$\begin{aligned} &\det(\mathbf{F}_{\mathrm{D}}(\boldsymbol{\nabla u})^{-1})\, \mathbf{F}_{\mathrm{D}}^{\top}(\boldsymbol{\nabla u})\, (\boldsymbol{E}_{\mathrm{c}}(\boldsymbol{q}, U) - \eta\, g\, \boldsymbol{e}_3) \\ &\approx \big((1 - \operatorname{trace}(\boldsymbol{\nabla u}))\, \mathbf{1} + \boldsymbol{\nabla u}^{\top}\big)\, \big(U + g \sin q_2, 0, -g(\cos q_2 + \eta)\big)^{\top} =: \boldsymbol{g}_0(\boldsymbol{q}, \boldsymbol{u}, U), \end{aligned}$$

where $\eta = m_{\mathrm{T}}/m_{\mathrm{L}}$ abbreviates the trolley/load mass ratio. This pressure is distributed on the surface $\partial\Omega \cap \{x_3 = l_1 - b_3\}$ as follows

$$\tilde{\boldsymbol{g}}(\boldsymbol{x}, \boldsymbol{q}(t), \boldsymbol{u}(\boldsymbol{x}, t), U(t)) := \begin{cases} \frac{1}{|\Gamma_{\mathrm{C}}|} \boldsymbol{g}_0(\boldsymbol{q}(t), \boldsymbol{u}(\boldsymbol{x}, t), U(t)); & \boldsymbol{x} \in \Gamma_{\mathrm{C}}(q_1(t)), \\ \mathbf{0} \qquad\qquad\qquad\qquad\quad ; & \text{otherwise.} \end{cases} \tag{19}$$

The elliptic PDE problem (15)–(17) for $\boldsymbol{u}$ depends on the trolley/load states $\boldsymbol{q}$ by means of the contact pressure $\tilde{\boldsymbol{g}}$, while the ODE system (12) for $\boldsymbol{q}$ is coupled as it depends on derivatives of $\boldsymbol{u}$. Since the important coupling effect that we would like to consider takes place on a small part Γ_{C} of the surface, we introduce the mean values

$$\bar{\boldsymbol{u}}(t) := \begin{pmatrix} \bar{u}_1(t) \\ \bar{u}_2(t) \\ \bar{u}_3(t) \\ \bar{u}_4(t) \\ \bar{u}_5(t) \\ \bar{u}_6(t) \end{pmatrix} = \begin{pmatrix} \frac{1}{|\Gamma_{\mathrm{C}}|} \int_{\Gamma_{\mathrm{C}}(q_1(t))} \partial_1 u_1(\boldsymbol{x}, t)\, \mathrm{d}\boldsymbol{x} \\ \frac{1}{|\Gamma_{\mathrm{C}}|} \int_{\Gamma_{\mathrm{C}}(q_1(t))} \partial_1 u_3(\boldsymbol{x}, t)\, \mathrm{d}\boldsymbol{x} \\ \frac{1}{|\Gamma_{\mathrm{C}}|} \int_{\Gamma_{\mathrm{C}}(q_1(t))} D_1 u_1(\boldsymbol{x}, t)\, \mathrm{d}\boldsymbol{x} \\ \frac{1}{|\Gamma_{\mathrm{C}}|} \int_{\Gamma_{\mathrm{C}}(q_1(t))} D_1 u_3(\boldsymbol{x}, t)\, \mathrm{d}\boldsymbol{x} \\ \frac{1}{|\Gamma_{\mathrm{C}}|} \int_{\Gamma_{\mathrm{C}}(q_1(t))} \partial_3 u_1(\boldsymbol{x}, t)\, \mathrm{d}\boldsymbol{x} \\ \frac{1}{|\Gamma_{\mathrm{C}}|} \int_{\Gamma_{\mathrm{C}}(q_1(t))} \partial_3 u_3(\boldsymbol{x}, t)\, \mathrm{d}\boldsymbol{x} \end{pmatrix}, \tag{20}$$

in the coupling terms of the ODE and, consistently, in $\tilde{\boldsymbol{g}}$. For the definition of the operator D_1 see (11). This also implies that we do not have to solve for every point

in space $\boldsymbol{x}$ another ODE system. We will consider $\bar{\boldsymbol{u}}$ as an independent state variable in the following.

2.5. *Objective function*

Our aim is to transport a load at rest ($\boldsymbol{v}^0 = \mathbf{0}$) from an initial state $\boldsymbol{q}^0$ to a terminal position q_1^f with angle $q_2^f = 0$, where the load should be at rest, that is, $\boldsymbol{v}^f = \mathbf{0}$. We would like to achieve this in minimal time, while also minimizing the swing of the load. Now the objective function (2), consisting of a time-minimal term, a (kinetic) energy-minimal term, possibly a regularization term, and terms penalizing the violation of terminal conditions, reads:

$$\begin{aligned} J(\boldsymbol{q}, \dot{\boldsymbol{q}}, \boldsymbol{U}, T) = \nu_1 T + \frac{\nu_2}{2} \|\dot{q}_2\|^2_{L^2(0,T)} + \frac{\nu_3}{2} \|U\|^2_{L^2(0,T)} + \frac{\nu_4}{2} |q_1(T) - q_1^f|^2 \\ + \frac{\nu_5}{2} |q_2(T)|^2 + \frac{\nu_6}{2} |\dot{q}_1(T)|^2 + \frac{\nu_7}{2} |\dot{q}_2(T)|^2. \end{aligned} \tag{21}$$

Different choices for the weights $\nu_1 > 0$, $\nu_j \geq 0$, $j = 2, \ldots, 7$ are discussed in Section 5. Except for the first term, the objective function J exhibits only quadratic terms and J is positive. Note that the displacements $\boldsymbol{u}$ or $\bar{\boldsymbol{u}}$ do not enter explicitly into the objective function. However, since $\bar{\boldsymbol{u}}$ enters into the ODE for $\boldsymbol{q}$ by means of the mass matrix $\mathbf{M}$ and the right-hand side $\mathbf{F}$, finding an optimal control U implies that vibrations of $\bar{\boldsymbol{u}}$ are damped out as well.

2.6. *Optimal control problem for 2d crane*

We solve the second-order ODE (12) for $\boldsymbol{q}$ as a system of coupled ODEs for $\boldsymbol{q}$ and $\boldsymbol{v} := \dot{\boldsymbol{q}}$, where $\boldsymbol{v}$ is considered as an independent state.

Our optimal control problem reads: Find states $\boldsymbol{q} \in [C^2([0,T])]^2$, $\boldsymbol{v} \in [C^1([0,T])]^2$, $\boldsymbol{u} \in C^0([0,T]; H^1(\Omega; \mathbb{R}^3))$, a control $U \in L^\infty(0,T)$ and a parameter $T \geq T_{\min} > 0$ (without loss of generality $T_{\min}$ arbitrarily small) such that the reduced cost function

$$\mathcal{F}(U,T) = J(\boldsymbol{q}(U,T), \boldsymbol{v}(U,T), U, T) \tag{22}$$

is minimized under the following constraints:

- the ODE system

$$\mathbf{M}(\boldsymbol{q}, \bar{\boldsymbol{u}})\, \dot{\boldsymbol{v}} = \boldsymbol{F}(\boldsymbol{q}, \boldsymbol{v}, \bar{\boldsymbol{u}}, U) \tag{23}$$

$$\dot{\boldsymbol{q}} = \boldsymbol{v} \qquad \text{in } (0,T), \tag{24}$$

resulting from (12) having employed the mean values $\bar{\boldsymbol{u}}$ in (13), yielding

$$\mathbf{M}(\boldsymbol{q}, \bar{\boldsymbol{u}}) := \begin{pmatrix} m(1 + 2\bar{u}_1) & (1 + \bar{u}_1)\cos q_2 + \bar{u}_2 \sin q_2 \\ (1 + \bar{u}_1)\cos q_2 + \bar{u}_2 \sin q_2 & l \end{pmatrix}, \tag{25}$$

and analgously in (14) and in (10), yielding

$$\boldsymbol{F}(\boldsymbol{q}, \boldsymbol{v}, \bar{\boldsymbol{u}}, U) := \begin{pmatrix} -m\,\bar{u}_3\, v_1^2 + ((1 + \bar{u}_1)\sin q_2 - \bar{u}_2 \cos q_2)\, v_2^2 + U \\ -(\bar{u}_3 \cos q_2 + \bar{u}_4 \sin q_2)\, v_1^2 - g_{\mathrm{e}} \sin q_2 \end{pmatrix}, \tag{26}$$

together with initial conditions

$$\boldsymbol{q}(0) = \boldsymbol{q}^0, \quad \boldsymbol{v}(0) = \boldsymbol{0}, \tag{27}$$

- the PDE (15)–(17)

$$\begin{aligned} -\operatorname{div} \sigma(\boldsymbol{u}) &= \boldsymbol{H} && \text{in } \Omega \times [0,T], && (28)\\ \boldsymbol{u} &= \boldsymbol{0} && \text{on } \Gamma_{\mathrm{D}} \times [0,T], && (29)\\ -\sigma(\boldsymbol{u}).\boldsymbol{n} &= \boldsymbol{R}(\boldsymbol{q}, \bar{\boldsymbol{u}}, U) && \text{on } \Gamma_{\mathrm{N}} \times [0,T], && (30) \end{aligned}$$

with the gravity term $\boldsymbol{H}$ from (18) and with averaging over Γ_{C} in (19) yielding

$$\boldsymbol{R}(\boldsymbol{q}, \bar{\boldsymbol{u}}, U) := \begin{cases} \frac{1}{|\Gamma_{\mathrm{C}}|} \boldsymbol{R}_0(\boldsymbol{q}, \bar{\boldsymbol{u}}, U); & \boldsymbol{x} \in \Gamma_{\mathrm{C}}(q_1), \\ 0 \qquad\qquad\quad ; & \text{otherwise}, \end{cases} \tag{31}$$

where

$$\boldsymbol{R}_0(\boldsymbol{q}, \bar{\boldsymbol{u}}, U) := \begin{pmatrix} (1-\bar{u}_6)(U + g\sin q_2) - \bar{u}_2 g(\cos q_2 + \eta) \\ 0 \\ \bar{u}_5(U + g\sin q_2) - (1-\bar{u}_1) g(\cos q_2 + \eta) \end{pmatrix}, \tag{32}$$

- the state equation (20) for $\bar{\boldsymbol{u}}$, and
- the control constraints

$$U_{\min} \le U(t) \le U_{\max} \quad \text{point-wise for all } t \in [0,T]. \tag{33}$$

In this paper we consider a smoothed version $\chi^{\varepsilon}_{\Gamma_{\mathrm{C}}}$, see (47), of the characteristic function $\chi_{\Gamma_{\mathrm{C}}}$ (being 1 on the set Γ_{C} and 0 otherwise) entering in (31). The solvability of this coupled problem for a given control is discussed in Section 3. We check that in case of neglected elastic deformations we recover in (23)–(27) and (33) the standard ODE-problem for an inelastic overhead gantry crane [4, 5]. As in the model in [4, 5] we consider no damping term and, furthermore, neglect the moment of inertia.

The terminal conditions $\boldsymbol{q}(T) = \boldsymbol{q}^f$ and $\boldsymbol{v}(T) = \boldsymbol{0}$ are realized approximately by the penalty terms corresponding to the weights ν_i, $i = 4, \dots 7$, within the objective function (21). We could also require state constraints, as $(b_1 + l_{\mathrm{T}})/2 < q_1 < l_2 - (b_1 + l_{\mathrm{T}})/2$, modelling that neither the trolley touches the crane tower nor that it jumps off the rails at the end of the crane beam, and $|q_2| < q_2^{max}$, reflecting that the pendant cord may not be tight for angles larger than a certain q_2^{max}. State constraints are neglected as in [4, 5], since we see in our numerical experiments that for typical initial data a control is found such that the state constraints are safely guaranteed.

3. Local-in-time existence and uniqueness of the coupled states

3.1. *Smoothed computational domain*

Let us consider a sufficiently small $T > 0$. We emphasize that, without loss of generality, Ω has a C^3 boundary, that could be obtained by smoothing out the corners, since corners are not relevant for the trolley-load system. We introduce

$\Omega_{\tilde{\varepsilon}} = \{\boldsymbol{x} \in \Omega \,|\, x_1 > b_1/2 + \tilde{\varepsilon}\}$ where $\tilde{\varepsilon} \in (0, \min\{(l_2 - b_1)/2, b_2/2, b_3/2\})$ may be chosen arbitrarily small. In order to avoid technical regularity issues, we focus in this section on

$$\Omega^* := \Omega_{\tilde{\varepsilon}} \cup \mathcal{B}_{\tilde{\varepsilon}}(\{\boldsymbol{x} \in \Omega \,|\, x_1 = b_1/2 + \tilde{\varepsilon}, |x_2| \leq b_2/2 - \tilde{\varepsilon}, l_1 - b_3 + \tilde{\varepsilon} \leq x_3 \leq l_1 - \tilde{\varepsilon}\})$$

where $\mathcal{B}_{\tilde{\varepsilon}}(S) = \cup_{\boldsymbol{x} \in S} B_{\tilde{\varepsilon}}(\boldsymbol{x})$ is a usual $\tilde{\varepsilon}$-neighborhood of a closed set S. Ω^* is constructed smoothly such that the contact lines $\overline{\Gamma_{\mathrm{D}}} \cap \overline{\Gamma_{\mathrm{N}}}$ between Dirichlet and Neumann boundary, where singularities might occur (compare [25, Sect. 3]), are taken out of Ω. Let $\boldsymbol{u}^*$ denote for the moment the solution on Ω^*. On $B^* := \partial\Omega^* \setminus \partial\Omega$ we prescribe $a^*\boldsymbol{u} + (1 - a^*)\sigma(\boldsymbol{u}^*).\boldsymbol{n} = \boldsymbol{0}$ as boundary condition, where a^* is a smooth spatial function on B^* such that $a^*(\boldsymbol{x}) = 1$ for $x_1 = b_1/2$ and $a^*(\boldsymbol{x}) = 0$ for $x_3 = l_1 - b_3$. According to [25, Sect. 3] and the references therein, $\boldsymbol{u}$ can be decomposed into a regular and singular parts. For the intersection $\{\boldsymbol{x} \in \Omega \,|x_1 = b_1/2, |x_2| = b_2/2$ or $|x_3 - l_1 - b_3/2| = b_3/2\}$ between homogeneous Dirchlet and Neumann b.c., the usual regularity results hold for the regular part of $\boldsymbol{u}$, while the singular parts (smooth except for the singularity) are bounded by $\tilde{r}^{-\tilde{m}}$, $\tilde{r}$ being the distance to the respective singularity, where $0 < \tilde{m} < 3$ depends on the interior angle and the Lamé parameters [26, Sect. 2]. Thus the norm of the whole singular part is arbitrarily small on Ω^* for suitable $\tilde{\varepsilon}$. Furthermore, our numerical simulations do not exhibit a singular behaviour in the corners or at the edges or in a neighborhood of these, thereby reconfirming that the corner and edge smoothing is only an auxiliary construction in order to avoid technical details in the proof concerning the interplay between crane and trolley.

Furthermore, we assume that the prescribed control $U \in L^\infty(0,T)$. For suitable data U, ρ, η, and g and sufficiently small T we may assume $\|\boldsymbol{\nabla u}\|_{L^2(\Omega)}, |\bar{u}_j| \ll 1$, $j = 1, 2, 5, 6$, for all $t \in [0,T]$, meaning that the assumption of small displacement gradients is justified (see also our numerics in Sect. 5). The latter implies *inter alia* $1 - 2\bar{u}_1 > 0$. As usual we denote, e.g., $H^k = W^{k,2}$ or $L^pH^k = L^p(0,T;H^k(\Omega^*;\mathbb{R}^3))$ for the Bochner space with the norm $\|\boldsymbol{f}\|^p_{L^pH^k} = \int_0^T \|\boldsymbol{f}(t)\|^p_{H^k(\Omega^*;\mathbb{R}^3)}\,\mathrm{d}t$.

3.2. *Standard results for uncoupled differential equations*

We summarize standard results for the ODE system and the PDE considered as stand alone equations. For given $\bar{u}_j \in C^0([0,T])$, $j = 1, 2, 3, 4$, $U \in C^0([0,T])$ there exists a unique solution $\boldsymbol{q} \in [C^2([0,T])]^2$ of the ODE system for sufficiently small T by the theorem of Picard-Lindelöf, since $\mathbf{M}^{-1}\boldsymbol{F}$ is Lipschitz in $\boldsymbol{q}$ and $\boldsymbol{v} = \dot{\boldsymbol{q}}$. For given $\bar{u}_j \in L^\infty(0,T)$, $j = 1, 2, 3, 4$, $U \in L^\infty(0,T)$ there exists a unique solution $\boldsymbol{q} \in [W^{2,\infty}(0,T)]^2$ (that is, C^2 for almost all t) of the ODE system for sufficiently small T. We have the standard estimate for ODE

$$\|q_j\|_{H^2(0,T)} \leq C\|(\mathbf{M}^{-1}\boldsymbol{F})(\boldsymbol{q}, \dot{\boldsymbol{q}}, \bar{\boldsymbol{u}}, U)\|_{L^2(0,T)}.$$

Let $\lambda, \mu > 0$ in (6). For given $\boldsymbol{q} \in [W^{\theta,\kappa}(0,T)]^2$, $U \in L^\kappa(0,T)$ and $\bar{u}_j \in L^\kappa(0,T)$, $j = 1, 2, 5, 6$, there exists a unique solution $\boldsymbol{u} \in L^\kappa(0,T;W^{1,p}(\Omega))$ of the PDE problem for $0 \leq \theta < \infty$, $6/5 \leq p < \infty$, and $1 \leq \kappa \leq \infty$ [22, Sect. 6.3], where the time regularity carries over from the data. For almost all t, there holds the estimate

$$\|\boldsymbol{u}(\cdot,t)\|_{H^1(\Omega)} \leq c\,(\|\boldsymbol{R}^{\varepsilon}(\boldsymbol{q}, \bar{\boldsymbol{u}}, U)\|_{L^2(\Gamma_{\mathrm{N}})} + 1), \tag{34}$$

where

$$\boldsymbol{R}^{\varepsilon}(\boldsymbol{q},\bar{\boldsymbol{u}},U) := \frac{1}{|\Gamma_{\mathrm{C}}|}\boldsymbol{R}_0(\boldsymbol{q},\bar{\boldsymbol{u}},U)\,\chi^{\varepsilon}_{\Gamma_{\mathrm{C}}}(\boldsymbol{x}) \tag{35}$$

is a smoothed version of $\boldsymbol{R}$ defined in (31), $\varepsilon > 0$ being another sufficiently small smoothing parameter. Note that the estimate (34) holds on Ω (as well as for $\boldsymbol{u}^*$ on Ω^* with the corresponding boundary conditions) and, therefore, $\boldsymbol{u} \in H^{1/2}(B^*)$. Now testing on Ω^* (instead of Ω) yields an additional integral term $\|a^*\boldsymbol{u}\|_{L^2(B^*)}$ in the estimates for $\boldsymbol{u}^*$.

We need the regularity result from [22, Sect. 6.3] combined with [25, Sect. 3] that for $\partial\Omega^* \in C^3$ we have $\boldsymbol{u}^* \in W^{3,p}(\Omega^*)$ for any $p \geq 6/5$, since $\boldsymbol{R}^{\varepsilon}$ is $W^{1-1/r,r}(\Gamma_{\mathrm{N}})$, $6/5 \leq r \leq \infty$. Thus together with the Sobolev embedding for $W^{3,p}(\Omega^*)$ for $p > 3/(1-\delta)$, $0 < \delta < 1$ [27, Th. 10.13, 2)], we get

$$\boldsymbol{u}^* \in L^{\kappa}(0,T;W^{3,p}(\Omega^*)) \hookrightarrow L^{\kappa}(0,T;C^{2,\delta}(\overline{\Omega^*})). \tag{36}$$

We consider only Ω^* in the following and, for ease of notation, we write again $\boldsymbol{u}$ instead of $\boldsymbol{u}^*$. (36) shows that the derivatives entering in the definition (20) of $\bar{\boldsymbol{u}}$ are well-defined. We remark that using mean values is crucial to obtain the stated regularity for $\boldsymbol{u}$ and $\bar{\boldsymbol{u}}$.

3.3. *Local-in-time existence for coupled differential equations*

We cannot expect an existence and uniqueness result for an arbitrary right hand side $\mathbf{M}^{-1}\boldsymbol{F}$ in the ODE (23). We will need the following local Lipschitz estimate.

Lemma 3.1: (Estimate for the right-hand-side of the ODE) *Within our approximation $\|\boldsymbol{\nabla u}\| \ll 1$, there holds for two different state pairs $(\boldsymbol{q}^{(i)},\bar{\boldsymbol{u}}^{(i)})$, $i = 1,2$,*

$$\begin{aligned}&\|(\mathbf{M}^{-1}\boldsymbol{F})(\boldsymbol{q}^{(1)},\dot{\boldsymbol{q}}^{(1)},\bar{\boldsymbol{u}}^{(1)},U) - (\mathbf{M}^{-1}\boldsymbol{F})(\boldsymbol{q}^{(2)},\dot{\boldsymbol{q}}^{(2)},\bar{\boldsymbol{u}}^{(2)},U)\|_{L^2(0,T)}\\ &\quad\leq \mathrm{Const}(k,K)\left(\|\boldsymbol{q}^{(1)}-\boldsymbol{q}^{(2)}\|_{H^1(0,T)} + \|\bar{\boldsymbol{u}}^{(1)}-\bar{\boldsymbol{u}}^{(2)}\|_{L^2(0,T)}\right),\end{aligned} \tag{37}$$

when $\sup_j \|\dot{q}_j^{(i)}\|_{L^\infty(0,T)} \leq k$ *and* $\sup_l \|\bar{u}_l^{(i)}\|_{L^\infty(0,T)} \leq K$ *for all* $i = 1,2$.

Proof. The proof relies on computing $\mathbf{M}^{-1}$ explicitly by Cramer's rule. For further details, see [24, A]. Then we exploit that for the quadratic terms

$$\begin{aligned}\|(\dot{q}_j^{(1)})^2 - (\dot{q}_j^{(2)})^2\|_{L^2(0,T)} &\leq \|(\dot{q}_j^{(1)} + \dot{q}_j^{(2)})(\dot{q}_j^{(1)} - \dot{q}_j^{(2)})\|_{L^2(0,T)}\\ &\leq 2k\|\dot{q}_j^{(1)} - \dot{q}_j^{(2)}\|_{L^2(0,T)},\end{aligned}$$

that for the mixed terms

$$\begin{aligned}\|\bar{u}_l^{(1)}\dot{q}_j^{(1)} - \bar{u}_l^{(2)}\dot{q}_j^{(2)}\|_{L^2(0,T)} &\leq \|\bar{u}_l^{(1)}\dot{q}_j^{(1)} - \bar{u}_l^{(1)}\dot{q}_j^{(2)} + \bar{u}_l^{(1)}\dot{q}_j^{(2)} - \bar{u}_l^{(2)}\dot{q}_j^{(2)}\|_{L^2(0,T)}\\ &\leq K\|\dot{q}_j^{(1)} - \dot{q}_j^{(2)}\|_{L^2(0,T)} + k\|\bar{u}_l^{(1)} - \bar{u}_l^{(2)}\|_{L^2(0,T)},\end{aligned}$$

and that $|\sin q_1|, |\cos q_2| \leq 1$ and sin, cos are Lipschitz, where $j = 1, 2$, $l = 1, 2, 3, 4$. □

Theorem 3.2: (Local-in-time well-posedness of the dynamics) *Let Ω^* be a C^3 domain and let $\lambda, \mu > 0$ be the Lamé parameters. Suppose that $U \in L^\infty(0, \infty)$ and that for $\mathbf{M}^{-1}\boldsymbol{F}$ the estimate (37) holds. Then for each pair (k, K) of positive numbers, there exists $T > 0$ and an area of the contact surface $|\Gamma_C| > 0$, such that the coupled ODE-PDE problem (23) –(32), with (20) and the smoothing (35), has a unique solution $\boldsymbol{q} \in [H^2(0, T)]^2$ and $\boldsymbol{u} \in L^\infty(0, T; W^{3,p}(\Omega^*))$ for any $p > 3$.*

Our strategy is inspired by Algorithm 4.1: We solve alternately for $\boldsymbol{u}$ and the mean values $\bar{\boldsymbol{u}}$, then the result is used for the right-hand side of the ODE. The ODE solution $\boldsymbol{q}$ is inserted into the PDE and a fixed point iteration is invoked. The idea of proof, relying on the Banach fixed point theorem and the estimate (38) yielding a factor proportional to $\sqrt{T}$ in the contraction constant, has been described for example by Niethammer [28] for a coupled ODE-Laplace PDE problem. A similar proof as needed for our problem is given in [29] for a coupled problem consisting of a single ODE for a free boundary, a quasilinear diffusion PDE, and the PDE of linear elasticity. Both cited proofs consider free boundary problems with a time-dependent domain and require beforehand a transformation to a fixed domain that is not needed here.

Proof. We would like to apply the Banach fixed point theorem in the space

$$\mathcal{M} = \mathcal{M}_T^k \times \mathcal{M}_T^K,$$

where

$$\mathcal{M}_T^k := \left\{ \boldsymbol{q} \in [H^2(0, T)]^2 \;\middle|\; \sup_j \|\dot{q}_j\|_{L^\infty(0,T)} \leq k \right\},$$
$$\mathcal{M}_T^K := \left\{ \boldsymbol{u} \in L^\infty H^2 \;\middle|\; \sup_l \|\bar{u}_l\|_{L^\infty(0,T)} \leq K \right\},$$

to the map

$$\begin{aligned} &\boldsymbol{G} : \mathcal{M} \to [H^2(0, T)]^2 \times L^\infty H^2, \\ &\quad (\boldsymbol{q}, \boldsymbol{u}) \mapsto (\boldsymbol{q}^+, \boldsymbol{u}^+) := (\mathcal{L}_1(\mathbf{M}^{-1}\boldsymbol{F})(\boldsymbol{q}, \dot{\boldsymbol{q}}, \bar{\boldsymbol{u}}^+, U),\ \mathcal{L}_2 \boldsymbol{R}^\varepsilon(\boldsymbol{q}, \bar{\boldsymbol{u}}, U)), \end{aligned}$$

where the operators $\mathcal{L}_1 : L^2(0, T) \to H^2(0, T)$ and $\mathcal{L}_2 : W^{2,\infty}(\Gamma_N) \to H^2(\Omega)$ map onto the solution of the ODE with right-hand side $\mathbf{M}^{-1}\boldsymbol{F}$ and onto the solution of the elasticity problem with right-hand side $\boldsymbol{R}^\varepsilon$ on Γ_N, respectively.

I. Strict contraction:
We consider two pairs of time trajectories $(\boldsymbol{q}^{(1)}, \boldsymbol{u}^{(1)})$ and $(\boldsymbol{q}^{(2)}, \boldsymbol{u}^{(2)})$. Define $\boldsymbol{q}_\Delta = \boldsymbol{q}^{(1)} - \boldsymbol{q}^{(2)}$ and $\boldsymbol{u}_\Delta = \boldsymbol{u}^{(1)} - \boldsymbol{u}^{(2)}$. The main ingredient is the following Poincaré inequality:

$$\|q_j - q_j^0\|_{L^\infty(0,T)} = \left\| \int_0^T \dot{q}_j \, \mathrm{d}t \right\|_{L^\infty(0,T)} \leq \sqrt{T} \|\dot{q}_j\|_{L^2(0,T)} \tag{38}$$

that follows by applying Hölder's inequality. Thus

$$\|q_j - q_j^0\|_{L^2(0,T)} = \sqrt{\int_0^T (q_j - q_j^0)^2 \,\mathrm{d}t} \le \sqrt{T}\,\|q_j - q_j^0\|_{L^\infty(0,T)},$$
$$\|q_j - q_j^0\|_{H^1(0,T)} \le \sqrt{T^2+1}\,\|\dot q_j\|_{L^2(0,T)}.$$

This procedure is repeated:

$$\|\dot q_j\|_{L^2(0,T)} = \sqrt{\int_0^T \dot q_j^2 \,\mathrm{d}t} \le \sqrt{T}\|\dot q_j\|_{L^\infty(0,T)} \le T\|\ddot q_j\|_{L^2(0,T)}.$$

Thus

$$\begin{aligned}\|q_j - q_j^0\|_{H^1(0,T)} &\le T\,\sqrt{T^2+1}\,\|\ddot q_j\|_{L^2(0,T)},\\ \|q_j - q_j^0\|_{H^2(0,T)} &\le \sqrt{T^4+T^2+1}\,\|\ddot q_j\|_{L^2(0,T)}.\end{aligned} \tag{39}$$

We write $\boldsymbol{q}_\Delta^+ = \boldsymbol{q}^{(1),+} - \boldsymbol{q}^{(2),+}$ and $\boldsymbol{u}_\Delta^+ = \boldsymbol{u}^{(1),+} - \boldsymbol{u}^{(2),+}$. Applying this to the map $\mathcal{G}$ yields

$$\begin{aligned}&\|q_{\Delta,j}^+\|_{H^2(0,T)} \le \sqrt{T^4+T^2+1}\,\times\\ &\quad\times \|(\mathbf{M}^{-1}\boldsymbol{F})(\boldsymbol{q}^{(1)},\dot{\boldsymbol{q}}^{(1)},\bar{\boldsymbol{u}}^{(1),+},U) - (\mathbf{M}^{-1}\boldsymbol{F})(\boldsymbol{q}^{(2)},\dot{\boldsymbol{q}}^{(2)},\bar{\boldsymbol{u}}^{(2),+},U)\|_{L^2(0,T)},\end{aligned}$$

and together with Lemma 3.1 we have

$$\|q_{\Delta,j}^+\|_{H^2(0,T)} \le \sqrt{T^4+T^2+1}\;\mathrm{Const}(k,K)\,\left(\|\boldsymbol{q}_\Delta\|_{H^1(0,T)} + \|\bar{\boldsymbol{u}}_\Delta^+\|_{L^2(0,T)}\right). \tag{40}$$

For the PDE we estimate for fixed ε and $\tilde\varepsilon$

$$\begin{aligned}&\|\boldsymbol{R}^\varepsilon(\boldsymbol{q}^{(1)},\bar{\boldsymbol{u}}^{(1)},U) - \boldsymbol{R}^\varepsilon(\boldsymbol{q}^{(2)},\bar{\boldsymbol{u}}^{(2)},U)\|_{L^2(\Gamma_\mathrm{N})}\\ &\le \mathrm{const}(k,K)\,|\Gamma_\mathrm{C}|^{\frac12}\left(\sum_{j=1,2,5,6}|\bar u_j^{(1)} - \bar\mu_j^{(2)}| + |\boldsymbol{q}^{(1)} - \boldsymbol{q}^{(2)}|\right)\end{aligned}$$

where, for instance, $\bar\mu_1^{(2)} = |\Gamma_\mathrm{C}|^{-1}\int_{\Gamma_\mathrm{C}(q_1^{(1)})} \mathfrak{d}_1 u_1^{(2)}\,\mathrm{d}\boldsymbol{x}$. The other $\bar\mu_j$ are defined analogously. By using $\bar u_1 = |\Gamma_\mathrm{C}|^{-1}\int_{\Gamma_\mathrm{C}} \mathfrak{d}_1 u_1\,\mathrm{d}\boldsymbol{x}$, the Hölder inequality and the trace theorem [30, Th. 5.22] in 3d

$$\begin{aligned}|\Gamma_\mathrm{C}(q_1^{(1)})|^{\frac12}|\bar u_1^{(1)} - \bar\mu_1^{(2)}| &\le |\Gamma_\mathrm{C}|^{-\frac12}\int_{\Gamma_\mathrm{C}(q_1^{(1)})}|\mathfrak{d}_1 u_1^{(1)} - \mathfrak{d}_1 u_1^{(2)}|\,\mathrm{d}\boldsymbol{x}\\ &\le |\Gamma_\mathrm{C}|^{-\frac12+\frac56}\|\boldsymbol{\nabla}\boldsymbol{u}^{(1)} - \boldsymbol{\nabla}\boldsymbol{u}^{(2)}\|_{L^6(\Gamma_\mathrm{C}(q_1^{(1)}))}\\ &\le |\Gamma_\mathrm{C}|^{\frac13}C(\Omega^*)\,\|\boldsymbol{u}^{(1)} - \boldsymbol{u}^{(2)}\|_{H^2(\Omega^*;\mathbb{R}^3)},\end{aligned} \tag{41}$$

where the constant does not depend on $|\Gamma_\mathrm{C}|$. Analogously the proof follows for $j = 2,5,6$. Using the estimate corresponding to (34) for the difference $\boldsymbol{u}_\Delta^+$

and for $\partial_i \boldsymbol{u}_\Delta^+$, $i = 1, 2, 3$, we end up with

$$\|\boldsymbol{u}_\Delta^+\|_{L^\infty H^2} \leq \tilde{C}(k, K, \Omega^*) \left(|\Gamma_C|^{\frac{1}{3}} \|\boldsymbol{u}_\Delta\|_{L^\infty H^2} + \|\boldsymbol{q}_\Delta\|_{L^\infty(0,T)} \right). \tag{42}$$

We use the regularity result (36) for $\boldsymbol{u}$ in order to get that $\bar{u}_3, \bar{u}_4$ are well-defined, then we combine the estimates (38) and (42), yielding

$$\|\boldsymbol{u}_\Delta^+\|_{L^\infty H^2} \leq \hat{C}(k, K, \Omega^*) \left(|\Gamma_C|^{\frac{1}{3}} \|\boldsymbol{u}_\Delta\|_{L^\infty H^2} + \sqrt{T} \|\boldsymbol{q}_\Delta\|_{H^2(0,T)} \right).$$

This estimate is inserted into (40) and we use (39) and (41), yielding

$$\begin{aligned} \|\boldsymbol{q}_\Delta^+\|_{H^2(0,T)} \leq{}& \sqrt{T^4 + T^2 + 1}\, \check{C}(k, K, \Omega^*) \times \\ &\times \left(\left(T\sqrt{T^2+1} + |\Gamma_C|^{-\frac{1}{6}} \sqrt{T} \right) \|\boldsymbol{q}_\Delta\|_{H^2(0,T)} + |\Gamma_C|^{\frac{1}{6}} \|\boldsymbol{u}_\Delta\|_{L^\infty H^2} \right). \end{aligned}$$

Now choosing $|\Gamma_C|$ and T (depending on $|\Gamma_C|$) sufficiently small guarantees that $\mathcal{G}$ is a strict contraction. From the strict contraction we also get directly that the local-in-time solution is unique.

II. Self-mapping:
In order to check that $\mathcal{G}$ maps $\mathcal{M}$ into itself, we use that the estimates from Part I carry over. So the self-mapping property follows analogously for sufficiently small T and $|\Gamma_C|$.

This shows that there exists a unique solution $(\boldsymbol{q}, \boldsymbol{u})$ with the stated regularity. From the estimates in Part I of the proof, we see that the solution depends continuously on the data. □

Since we do not know from the last lemma whether the time T guaranteed by the fixed-point method is larger then the total time obtained by the optimal control, global existence of a solution for our problem is not guaranteed and cannot be expected for arbitrary data and control. However, when minimizing our objective function we may hope that the control U is determined in such a way that no blow up for q_1 or q_2 may happen in finite time.

4. Discretized problem and optimal control method

Due to the complicated model we follow here a first-discretize-then-optimize (FDTO) approach for solving our optimal control problem. Furthermore we work with a sensitivity-based approach since the adjoint optimality system cannot be derived by standard methods. The reason for this is the averaging over $\boldsymbol{u}$ that appears in the integral equation (20).

Since in our numerical example we consider $\nu_1 > 0$ in (21), we perform a time transformation onto a fixed time interval, mapping $t \in [0, T]$ to $\tau \in [0, 1]$. This provides an easy way to determine the control parameter T in the sequel. According to the time transformation, time derivatives have to be scaled with the factor T. We consider from now on time-transformed functions but denote them again with the same symbol in order to keep the notation simple.

4.1. *Time integration*

The PDE problem for $\boldsymbol{u}$ and $\bar{\boldsymbol{u}}$ is solved by a finite element method (FEM). In order to avoid locking effects [31, Ch. VI, §3] we use Lagrange $\mathbb{P}_2$-elements and a vertically refined layered mesh. The spatially discretized problem may be considered as a semi-explicit DAE of index 1 with $\boldsymbol{u}$ being the algebraic variable. We solve in time by means of the explicit Heun method with respect to $\boldsymbol{q} = (x_{\mathrm{T}}, \alpha)^\top$ and $\boldsymbol{v} = \dot{\boldsymbol{q}}$. This second-order method allows to solve the 2-dimensional ODE system accurately enough.

By dividing the time interval $[0, 1]$ into $N \in \mathbb{N}$ intervals of length $h := 1/N$, we define the time steps $\tau_k := kh$, $k = 0, \dots, N$. We abbreviate $\boldsymbol{q}_{(k)} := \boldsymbol{q}(\tau_k)$, $\boldsymbol{v}_{(k)} := \boldsymbol{v}(\tau_k)$, $\bar{\boldsymbol{u}}_{(k)} := \bar{\boldsymbol{u}}(\tau_k)$, $\boldsymbol{u}_{(k)}(\cdot) := \boldsymbol{u}(\cdot, \tau_k)$, $U_{(k)} = U(\tau_k)$, and for the predictor step of the Heun method we introduce the values $\tilde{\boldsymbol{q}}_{(k)}$ and $\tilde{\boldsymbol{v}}_{(k)}$, $k = 1, \dots, N$. Furthermore we write, for instance,

$$\begin{aligned}
\mathbf{M}_{(k)} &= \mathbf{M}(\boldsymbol{q}_{(k)}, \bar{\boldsymbol{u}}_{(k)}), \quad \boldsymbol{F}_{(k)} = \boldsymbol{F}(\boldsymbol{q}_{(k)}, \boldsymbol{v}_{(k)}, \bar{\boldsymbol{u}}_{(k)}, U_{(k)}),\\
\tilde{\mathbf{M}}_{(k+1)} &= \mathbf{M}(\tilde{\boldsymbol{q}}_{(k+1)}, \bar{\boldsymbol{u}}_{(k)}), \quad \tilde{\boldsymbol{F}}_{(k+1)} = \boldsymbol{F}(\tilde{\boldsymbol{q}}_{(k+1)}, \tilde{\boldsymbol{v}}_{(k+1)}, \bar{\boldsymbol{u}}_{(k)}, U_{(k)}),\\
\boldsymbol{R}_{(k)} &= \boldsymbol{R}^{\varepsilon}(\boldsymbol{q}_{(k)}, \bar{\boldsymbol{u}}_{(k)}, U_{(k)}).
\end{aligned}$$

The discretized version of the set of admissible controls is

$$\mathcal{U}_{ad} := \{\boldsymbol{V} \in \mathbb{R}^{N+1} \,\big|\, V_{(j)} \in U_{ad} := [U_{\min}, U_{\max}] \;\forall\, j = 0, \dots, N\}.$$

Thanks to Theorem 3.2 we may expect that the following algorithm, including a fixed-point iteration for $\boldsymbol{u}$ and $\bar{\boldsymbol{u}}$, works out well.

Algorithm 4.1: (Simulation procedure with Heun method in time)

(0) Init: Let $T = T^{(0)}$ be given, $k := 0$, $\bar{\boldsymbol{u}}_{-1} \equiv 0$. Let control input $U_{(\cdot)} \in \mathcal{U}_{ad}$ and initial values $\boldsymbol{q}_0 = \boldsymbol{q}^0$ and $\boldsymbol{v}_0 = \boldsymbol{v}^0$ be given.

(1) (o) $\bar{\boldsymbol{u}}_{(k)} = \bar{\boldsymbol{u}}_{(k-1)}$.

(i) At time τ_k solve

$$\begin{aligned}
-\operatorname{div} \sigma(\boldsymbol{u}_{(k)}) &= \boldsymbol{H} && \text{in } \Omega \times \{\tau_k\},\\
\boldsymbol{u}_{(k)} &= \boldsymbol{0} && \text{on } \Gamma_{\mathrm{D}} \times \{\tau_k\},\\
-\sigma(\boldsymbol{u}_{(k)}).\boldsymbol{n} &= \boldsymbol{R}_{(k)} && \text{on } \Gamma_{\mathrm{N}} \times \{\tau_k\}.
\end{aligned}$$

(ii) $\bar{\boldsymbol{u}}_{(k)}^{old} := \bar{\boldsymbol{u}}_{(k)}$. Compute $\bar{\boldsymbol{u}}_{(k)}$.

(iii) If $\|\bar{\boldsymbol{u}}_{(k)} - \bar{\boldsymbol{u}}_{(k)}^{old}\| > err_0$ for a suitable norm and given error tolerance err_0 go to (i).

(2) Set

$$\mathbf{M}_{(k)}\tilde{\boldsymbol{v}}_{(k+1)} = \mathbf{M}_{(k)}\boldsymbol{v}_{(k)} + hT\boldsymbol{F}_{(k)}, \tag{43}$$

$$\tilde{\boldsymbol{q}}_{(k+1)} = \boldsymbol{q}_{(k)} + hT\boldsymbol{v}_{(k)}, \tag{44}$$

$$\tilde{\mathbf{M}}_{(k+1)}\boldsymbol{v}_{(k+1)} = \tilde{\mathbf{M}}_{(k+1)}\frac{1}{2}(\boldsymbol{v}_{(k)} + \tilde{\boldsymbol{v}}_{(k+1)}) + \frac{hT}{2}\tilde{\boldsymbol{F}}_{(k+1)}, \tag{45}$$

$$\boldsymbol{q}_{(k+1)} = \frac{1}{2}\left(\boldsymbol{q}_{(k)} + \tilde{\boldsymbol{q}}_{(k+1)}\right) + \frac{hT}{2}\tilde{\boldsymbol{v}}_{(k+1)}. \tag{46}$$

(3) Set $\tau_{k+1} = \tau_k + h$, $k := k + 1$.
If $\tau_k < 1$ go to (1), otherwise Stop.

In order to avoid technical issues with the size of finite elements, we replace the characteristic function χ_{Γ_C}, appearing in the definition (31) of $\boldsymbol{R}$, by an approximation with a version $\tilde{\chi}^{\varepsilon}_{\Gamma_C}$ smoothed in x_1-direction, where $\varepsilon > 0$ is a small, but fixed parameter:

$$\tilde{\chi}^{\varepsilon}_{\Gamma_C}(\boldsymbol{x}) := 1 - \tanh^2\left(((x_1 - q_1)/\delta_1)^L\right) \tag{47}$$

with $L \in \mathbb{N}$, $\delta_1 > 0$ such that

$$|\chi_{\Gamma_C}(\boldsymbol{x}) - \tilde{\chi}^{\varepsilon}_{\Gamma_C}(\boldsymbol{x})| \leq \begin{cases} 1/2\,; & \text{for } |x_1 - q_1 \mp l_T/2| < \varepsilon, \\ \varepsilon\ ; & \text{otherwise.} \end{cases}$$

Abbreviating $A_1 := \ln(\operatorname{atanh}(\sqrt{1-\varepsilon}))$, $A_2 := \ln(\operatorname{atanh}(\sqrt{\varepsilon}))$ and $B_{1/2} = \ln(l_T/2 \pm \varepsilon)$, we have $L = \lceil (A_1 - A_2)/(2(B_1 - B_2)) + 1 \rceil$ ($\lceil \cdot \rceil$ denoting the ceiling function) and $\delta_1 = \exp((A_1 B_2 - A_2 B_1)/(A_1 - A_2))$. For our numerics we work with $\varepsilon = 0.1$, yielding $L = 4$ and $\delta_1 \approx 0.314$, whereas we do not consider any corner and edge smoothing for the simulations.

For small $\eta = m_T/m_L$, which is realistic in applications, the matrix $\mathbf{M}$ is ill-conditioned. But this issue may be overcome by a left preconditioning of the equations (43) and (45), respectively, with the matrix $\mathbf{M}(\boldsymbol{q} = \boldsymbol{q}_{pc}, \bar{\boldsymbol{u}} = \bar{\boldsymbol{u}}_{pc})^{-1}$, where, for example, $\boldsymbol{q}_{pc} = (x^0_T, -1/(2\pi))^\top$ and $\bar{\boldsymbol{u}}_{pc} = (0, -0.1, 0, 0, 0, 0)^\top$.

We remark that in Algorithm 4.1, Step (1)(o), we could set alternatively $\bar{\boldsymbol{u}}_{(k)} \equiv 0$, but the choice above turns out to yield fewer iterations in Step (1).

4.2. *Projected gradient method with BFGS update*

We consider a projected gradient method [11, Algorithm 2.2] using a sensitivity-based approach (cf. [14, p. 58] within the context of Banach spaces). We prefer to project onto the set of admissible controls within the line search, instead of projecting first and then performing a line search. We discretize the integrals appearing in the reduced objective function by the trapezoidal rule, according to our second-order time integration. Then we calculate a discretized reduced gradient (cf. [17, Ch. 4] for a full discretization approach).

Note that due to our modelling ansatz $\boldsymbol{u}$ and $\bar{\boldsymbol{u}}$ do not enter the reduced objective function. Thus our OCP comes down to finding a time-optimal parameter T and control U of a discretized optimal control problem for a DAE with index 1 that may be solved with respect to $\boldsymbol{u}$. It is an open question whether for this class of problems we may expect the existence of optimal controls (U, T). The objective exhibits a quadratic term in U and additional nonlinear terms with respect to T and U.

We use the notation $\boldsymbol{q}^{(n)}_{(k)} = \boldsymbol{q}(U^{(n)}, T^{(n)})(t_k)$, $\mathcal{F}^{(n)} = \mathcal{F}(U^{(n)}, T^{(n)})$ and $J^{(n)} = J(\boldsymbol{q}^{(n)}, \boldsymbol{v}^{(n)}, U^{(n)}, T^{(n)})$ to indicate the dependence on the control $(U^{(n)}, T^{(n)})$ in the optimization iteration $n \in \mathbb{N}$. For brevity we write $S_{(i)} = 1$ for $i = 1, \ldots, N-1$ and $S_{(i)} = 1/2$ for $i = 0$ or N. The discretized time-scaled version of the objective

function (21) reads according to the trapezoidal rule

$$\begin{aligned}\mathcal{F}^{(n)} = \nu_1 T^{(n)} + \frac{\nu_2}{2} h T^{(n)} \sum_{k=0}^{N} S_{(i)} |v^{(n)}_{(k),2}|^2 + \frac{\nu_3}{2} h T^{(n)} \sum_{i=0}^{N} S_{(i)} |U^{(n)}_{(i)}|^2 \\ + \frac{\nu_4}{2} |q^{(n)}_{(N),1} - q^f_1|^2 + \frac{\nu_5}{2} |q^{(n)}_{(N),2}|^2 + \frac{\nu_6}{2} |v^{(n)}_{(N),1}|^2 + \frac{\nu_7}{2} |v^{(n)}_{(N),2}|^2. \end{aligned} \tag{48}$$

As usual we write $\delta_{U^{(n)}_{(i)}} \boldsymbol{q}^{(n)}_{(i)}$ for the sensitivity of the discretized states $\boldsymbol{q}$ with respect to the discretized control U (at time step i and optimization iteration n). Other sensitivities are defined analogously.

Algorithm 4.2: (Projected gradient method with BFGS update, sensitivity-based approach)

(i) Initialize $U^{(0)}_{(k)} \in U_{ad}$, $k = 0, \dots N$, $T^{(0)} \geq T_{\min}$, $\mathbf{H}^{(0)} = \mathbf{1}$, $n = 0$.

(ii) For $k = 0, \dots, N-1$ solve the state equations for $\boldsymbol{q}^{(n)}_{(k+1)}, \boldsymbol{u}^{(n)}_{(k)}, \bar{\boldsymbol{u}}^{(n)}_{(k)}$ by Algorithm 4.1, given a control $U^{(n)}_{(k)}$ and a parameter $T^{(n)}$.

(iii) For $k = 0, \dots, N-1$ compute the sensitivities for $\boldsymbol{u}^{(n)}_{(k)}$, $\bar{\boldsymbol{u}}^{(n)}_{(k)}$ by iterations, then solve the sensitivity equations for $\boldsymbol{q}^{(n)}_{(k+1)}$, $\boldsymbol{v}^{(n)}_{(k+1)}$, the latter are,

$$\begin{aligned}\mathbf{M}^{(n)}_{(k)} \delta_{U^{(n)}_{(i)}} \boldsymbol{v}^{(n)}_{(k+1)} &= \mathbf{M}^{(n)}_{(k)} \delta_{U^{(n)}_{(i)}} \boldsymbol{v}^{(n)}_{(k)} + h T^{(n)} \delta_{U^{(n)}_{(i)}} \boldsymbol{F}^{(n)}_{(k)} \\ &\quad - \delta_{U^{(n)}_i} \mathbf{M}^{(n)}_{(k)} (\boldsymbol{v}^{(n)}_{(k+1)} - \boldsymbol{v}^{(n)}_{(k)}), \quad i = 0, \dots, N, \end{aligned} \tag{49}$$

$$\begin{aligned}\mathbf{M}^{(n)}_{(k)} \delta_{T^{(n)}} \boldsymbol{v}^{(n)}_{(k+1)} &= \mathbf{M}^{(n)}_{(k)} \delta_{T^{(n)}} \boldsymbol{v}^{(n)}_{(k)} + h \left(T^{(n)} \delta_{T^{(n)}} \boldsymbol{F}^{(n)}_{(k)} + \boldsymbol{F}^{(n)}_{(k)} \right) \\ &\quad - \delta_{T^{(n)}} \mathbf{M}^{(n)}_{(k)} (\boldsymbol{v}^{(n)}_{(k+1)} - \boldsymbol{v}^{(n)}_{(k)}), \end{aligned} \tag{50}$$

and set

$$\delta_{U^{(n)}_{(i)}} \boldsymbol{q}^{(n)}_{(k+1)} = \delta_{U^{(n)}_{(i)}} \boldsymbol{q}^{(n)}_{(k)} + h T^{(n)} \delta_{U^{(n)}_{(i)}} \boldsymbol{v}^{(n)}_{(k)}, \quad i = 0, \dots, N, \tag{51}$$

$$\delta_{T^{(n)}} \boldsymbol{q}^{(n)}_{(k+1)} = \delta_{T^{(n)}} \boldsymbol{q}^{(n)}_{(k)} + h \left(T^{(n)} \delta_{T^{(n)}} \boldsymbol{v}^{(n)}_{(k)} + \boldsymbol{v}^{(n)}_{(k)} \right). \tag{52}$$

(iv) Determine $\boldsymbol{D}^{(n)}$ as quasi-Newton direction, solving

$$\mathbf{H}^{(n)} \boldsymbol{D}^{(n)} = -\boldsymbol{\delta}_{(U^{(n)}, T^{(n)})} \mathcal{F}^{(n)},$$

where the approximated Hessian $\mathbf{H}^{(n)}$ is determined by the modified BFGS update [32] and the anti-gradient of $\mathcal{F}$ with respect to the scalar product in $L^2(0,T) \times \mathbb{R}$ is calculated using the sensitivities:

$$\begin{aligned}\delta_{U^{(n)}_{(i)}} \mathcal{F}^{(n)} &= \boldsymbol{\delta}_{\boldsymbol{q}^{(n)}} J^{(n)} \cdot \delta_{U^{(n)}_{(i)}} \boldsymbol{q}^{(n)} + \boldsymbol{\delta}_{\boldsymbol{v}^{(n)}} J^{(n)} \cdot \delta_{U^{(n)}_{(i)}} \boldsymbol{v}^{(n)} + \delta_{U^{(n)}_{(i)}} J^{(n)}, \\ &\qquad\qquad i = 0, \dots, N, \\ \delta_{T^{(n)}} \mathcal{F}^{(n)} &= \boldsymbol{\delta}_{\boldsymbol{q}^{(n)}} J^{(n)} \cdot \delta_{T^{(n)}} \boldsymbol{q}^{(n)} + \boldsymbol{\delta}_{\boldsymbol{v}^{(n)}} J^{(n)} \cdot \delta_{T^{(n)}} \boldsymbol{v}^{(n)} + \delta_{T^{(n)}} J^{(n)}. \end{aligned}$$

(v) If a stopping condition is fulfilled, then Stop.
(vi) Determine the step size $s^{(n)}$ by an Armijo line search (see Algorithm 4.3)

$$\mathcal{F}(\mathrm{P}_A((U^{(n)},T^{(n)})^\top + s^{(n)}\boldsymbol{D}^{(n)})) = \min_{s\in(0,1]} \mathcal{F}(\mathrm{P}_A((U^{(n)},T^{(n)})^\top + s\boldsymbol{D}^{(n)})),$$

where P_A is the Euclidean projection onto the set $A := U_{ad} \times \{T \geq T_{\min}\}$.
(vii) Update control and parameter $(U^{(n+1)},T^{(n+1)})^\top = \mathrm{P}_A((U^{(n)},T^{(n)})^\top + s^{(n)}\boldsymbol{D}^{(n)})$.
(viii) Set $n := n+1$ and go to Step (i).

For the ease of presentation we have stated the last algorithm with Euler steps in (49)–(52), instead of Heun steps. Actually, we implemented a Heun method being consist with Algorithm 4.1. For further information and the formulas for derivatives of $\boldsymbol{F}^{(n)}_{(k)}$ and $\mathbf{M}^{(n)}_{(k)}$ with respect to controls and for the derivatives $\boldsymbol{\delta}_{(\boldsymbol{q}^{(n)},\boldsymbol{v}^{(n)},U^{(n)},T^{(n)})}J^{(n)}$, see [24, B]. Since $\bar{\boldsymbol{u}}$ and $\boldsymbol{u}$ do not appear in the objective function, the sensitivities for $\bar{\boldsymbol{u}}$ only enter in $\delta_{U^{(n)}_{(i)}}\boldsymbol{F}^{(n)}_{(k)}$, $\delta_{T^{(n)}}\boldsymbol{F}^{(n)}_{(k)}$, $\delta_{U^{(n)}_{(i)}}\mathbf{M}^{(n)}_{(k)}$, and $\delta_{T^{(n)}}\mathbf{M}^{(n)}_{(k)}$.

The Armijo line search in Step (vi) of the Algorithm 4.2 is performed as follows

Algorithm 4.3: (Projected Armijo line search)

(i) Let $\beta_A \in (0,1)$, $\sigma_A \in (0,1/2)$, $l \in \mathbb{N}^*$ and let $\phi(s) = \mathcal{F}(\mathrm{P}_A((U^{(n)},T^{(n)}) + s\boldsymbol{D}^{(n)}))$ be given.
(ii) Find maximal $s = \beta_A^l$ such that

$$\phi(s) \leq \phi(0) + \sigma_A\, s\, \phi'(0) \tag{53}$$

with $\phi'(0) = \boldsymbol{\delta}_{(U^{(n)},T^{(n)})}\mathcal{F}^{(n)} \cdot \tilde{\boldsymbol{D}}^{(n)}$, where for $i = 0,\dots,N$

$$\tilde{\boldsymbol{D}}^{(n)}_{(i)} = \begin{cases} 0 \quad ; U^{(n)}_{(i)} \notin U_{ad} = [U_{\min}, U_{\max}] \text{ or} \\ \qquad (U^{(n)}_{(i)} = U_{\max} \wedge D^{(n)}_{(i)} \geq 0) \text{ or } (U^{(n)}_{(i)} = U_{\min} \wedge D^{(n)}_{(i)} \leq 0), \\ D^{(n)}_{(i)} \text{ ; else,} \end{cases}$$

and

$$\tilde{\boldsymbol{D}}^{(n)}_{(N+1)} = \begin{cases} 0 \quad ; T^{(n)} < T_{\min} \text{ or } (T^{(n)} = T_{\min} \wedge D^{(n)}_{(N+1)} \leq 0), \\ D^{(n)}_{(N+1)} \text{ ; else.} \end{cases}$$

(iii) Set $s^{(n)} := s$.

For the choice $\beta_A = 0.9$ and $\sigma_A = 10^{-4}$ we observe a good performance of the line search within our numerical experiments. For the Armijo line search it turns out to be crucial to solve for the new states, when computing $\phi(s)$ in the Armijo condition (53), to sufficient precision, requiring a finite element (FE) solution of the PDE for every β_A^l. Otherwise the algorithm might terminate since an admissible step size cannot be found.

Our reduced gradient reads,

$$\begin{aligned}\delta_{U_{(i)}^{(n)}}\mathcal{F}^{(n)} = \nu_2\, hT^{(n)} \sum_{k=i+1}^{N} S_{(k)} v_{(k),2}^{(n)}\, \delta_{U_{(i)}} v_{(k),2}^{(n)} + \nu_3\, hT^{(n)} S_{(i)} U_{(i)}^{(n)} \\ + \nu_4\, (q_{(N),1}^{(n)} - q_1^f)\, \delta_{U_{(i)}} q_{(N),1}^{(n)} + \nu_5\, q_{(N),2}^{(n)}\, \delta_{U_{(i)}} q_{(N),2}^{(n)} \\ + \nu_6\, v_{(N),1}^{(n)}\, \delta_{U_{(i)}} v_{(N),1}^{(n)} + \nu_7\, v_{(N),2}^{(n)}\, \delta_{U_{(i)}} v_{(N),2}^{(n)}\end{aligned}$$

for the components $i = 0, \ldots, N$, and

$$\begin{aligned}\delta_{T^{(n)}}\mathcal{F}^{(n)} = \nu_1 + \frac{\nu_2}{2} h \sum_{k=0}^{N} S_{(k)} |v_{(k),2}^{(n)}|^2 + \nu_2\, hT^{(n)} \sum_{k=1}^{N} S_{(k)} v_{(k),2}^{(n)}\, \delta_{T^{(n)}} v_{(k),2}^{(n)} \\ + \frac{\nu_3}{2} h \sum_{k=0}^{N} S_{(k)} |U_{(k)}^{(n)}|^2 + \nu_4\, (q_{(N),1}^{(n)} - q_1^f)\, \delta_{T^{(n)}} q_{(N),1}^{(n)} + \nu_5\, q_{(N),2}^{(n)}\, \delta_{T^{(n)}} q_{(N),2}^{(n)} \\ + \nu_6\, v_{(N),1}^{(n)}\, \delta_{T^{(n)}} v_{(N),1}^{(n)} + \nu_7\, v_{(N),2}^{(n)}\, \delta_{T^{(n)}} v_{(N),2}^{(n)}.\end{aligned}$$

Clearly, for $i \geq k$ we find $\delta_{U_{(i)}^{(n)}} \boldsymbol{q}_{(k)}^{(n)} = 0$ and $\delta_{U_{(i)}^{(n)}} \boldsymbol{v}_{(k)}^{(n)} = 0$. Thus we can simplify $\delta_{U_{(N)}^{(n)}}\mathcal{F}^{(n)} = \nu_3 (hT^{(n)}/2) U_{(N)}^{(n)}$. We conclude that if the control cost parameter $\nu_3 = 0$, then $U_{(N)}^{(n)}$ is arbitrary. For uniqueness, we set $U_{(N)}^{(n)} = 0$.

For small trolley/load mass ratios $\eta = m_{\mathrm{T}}/m_{\mathrm{L}}$, our ODE system turns out to be stiff and a semi-explicit method [33] is applied by using on the right-hand side of (44) $\tilde{\boldsymbol{v}}_{(k+1)}$ instead of $\boldsymbol{v}_{(k)}$ and on the right-hand side of (46) $\boldsymbol{v}_{(k+1)}$ instead of $\tilde{\boldsymbol{v}}_{(k+1)}$.

A suitable algorithm for determining the weights ν_i, $i = 4, \ldots, 7$, is crucial in order to avoid a breakdown of the Armijo line search and to obtain feasible computing times. Within this framework our objective function can be interpreted as an exact penalty function [34, Sect. 5.4]. The optimal penalty weights ν_i, $i = 4, \ldots, 7$ are unknown, but we may expect that we approximate the optimal weights numerically sufficiently well.

Algorithm 4.4: (Penalty method)

(i) Set initial weights $\nu_i = \nu_i^0$, $i = 4, \ldots, 7$.
(ii) Run Algorithm 4.2.
(iii) If $|q_{(N),j}^{(n)} - q_j^f|, |v_{(N),j}^{(n)}| < err_1$ for $j = 1, 2$ and a given error tolerance err_1, then Stop.
(iv) Increase the weights $\nu_j := \tilde{\rho}\, \nu_j$ (where $\tilde{\rho} > 1$) for indices j corresponding to violated terminal conditions. Go to (ii).

Typically we consider the factor $\tilde{\rho} = 10$. We remark that the initial value of ν_4 has to be chosen such that the trolley remains within the feasible area of the crane beam for the first run of Algorithm 4.2. We start with $\nu_4^0 = 1/(q_1^f - q_1^0)^2$, $\nu_5^0 = 10\, \nu_4^0$, $\nu_6^0 = 5000\, \nu_7^0$, and $\nu_7^0 = 1/(T^{(0)})^2$.

5. Numerical results

We solve our optimal control problem by means of Algorithm 4.2. This algorithm has been implemented in the open-source software package FEniCS v1.4 (API Python 2.7.3 with PETSc v3.2 as linear algebra package). It has been executed on a workstation, equipped with Intel(R) Xeon(R) CPU E5640 @2.67GHz × 16 processors and a memory of 23.6 GiB, under Ubuntu Linux. For data considered in our simulations, see Table 1.

Table 1. Typical data for a large crane (see, e.g., [35]) and further used parameters.

Description	Symbol	Value	Unit
Width crane beam	b_2	0.80	m
Height crane beam	$b_3 = b_2$	0.80	m
Length crane beam	l_2	45.80	m
Width trolley	b_T	0.80	m
Height trolley	h_T	0.10	m
Length trolley	l_T	0.60	m
Mass trolley	m_T	150	kg
Mass load	m_L	3340	kg
Length pendulum	l	17.5	m
Scaled maximal accelerating force	U_max	0.006	s^{-2}
Scaled minimal accelerating force	U_min	-0.006	s^{-2}
Scaled mass density crane beam	ρ	0.0104	m^{-3}
Scaled Lamé parameter 1	λ	$1.76 \cdot 10^6$	N kg^{-1} m^{-3}
Scaled Lamé parameter 2	μ	$1.33 \cdot 10^6$	N kg^{-1} m^{-3}
Standard gravity earth	g_e	9.81	N kg^{-2}
Initial angle trolley	q_2^0	0	rad
Terminal angle trolley	q_2^f	0	rad
Initial velocities	$\boldsymbol{v}^0$	$\mathbf{0} = (0,0)^\top$	(m s^{-1}, rad s^{-1})$^\top$
Terminal velocities	$\boldsymbol{v}^f$	$\mathbf{0} = (0,0)^\top$	(m s^{-1}, rad s^{-1})$^\top$

In the following we present the results for a mesh that has been refined adaptively on $\{\Gamma_\mathrm{C}(q_1) \,|\, q_1^0 \le q_1 \le q_1^f\}$ beforehand by solving an auxiliary Poisson problem and refining recursively cells with a residual error larger than 10^{-4}. This procedure yielded about 10 500 vertices and 43 200 3d-cells. We consider $N = 100$ time steps on the normalized time interval $[0, 1]$, the error tolerance $err_0 = 2.0 \cdot 10^{-5}$ for the $\boldsymbol{u}$-iteration, the error tolerance $err_1 = 5.0 \cdot 10^{-2}$ for the terminal conditions, and the relative error tolerance $err = 10^{-8}$ for the optimization. As stopping condition we work with

$$\frac{\mathcal{F}^{(n)} - \mathcal{F}^{(n+1)}}{1 + \mathcal{F}^{(n+1)}} < err. \tag{54}$$

As initial guess we start with

$$U_0(t) = \begin{cases} U_\mathrm{max} & ; \quad 0 \le t < 0.4, \\ -U_\mathrm{max}; & 0.4 \le t < 0.8, \\ 0 & ; \quad 0.8 \le t \le 1, \end{cases}$$

for the control, motivated by the intuitive strategy to (i) accelerate, then (ii) brake, and (iii) wait until the system swings out. As initial guess for the total time we take $T^{(0)} = 18.5$ [s].

For the weights ν we consider $\nu_1 = 10/T^{(0)}$, $\nu_2 = 500$ and we focus here on the situation $\nu_3 = 0$, though the algorithm clearly can handle $\nu_3 > 0$ as well. According to Algorithm 4.4, a suitable choice for the weights ν_i, $i = 4, \ldots, 7$ is obtained by

successively increasing the weights according to violated terminal conditions and restarting with the control and final time determined so far, until the terminal conditions are fulfilled to sufficient accuracy. It turns out to be crucial that the weights are scaled such that the reduced objective function is approximately of order 1.

5.1. *Numerical time-optimal control*

In a first study we consider a projected gradient method as obtained from Algorithm 4.2 by fixing $\mathbf{H}^{(n)} = \mathbf{H}^{(0)}$. Fig. 2 (left) shows the descent in the reduced objective function versus optimization iterations. Tab. 2 shows the stages of the penalty method and the violation of terminal conditions. The optimization of the final time T is illustrated in Tab. 2. Here we consider as initial position of the trolley $q_1^0 = 11.7$ [m] and as terminal position $q_1^f = 17.0$ [m]. The resulting control and states are depicted in Fig. 2, right, and in Fig. 3, where the final penalty weights are $\nu_4 = 10^4\nu_4^0$, $\nu_5 = \nu_5^0$, $\nu_6 = 10^2\nu_6^0$, and $\nu_7 = 10^4\nu_7^0$.

Table 2. Numerical results for the number of optimization iterations, violation of terminal conditions and final time $T^{(n)}$ for each stage of the penalty method.

Stage	# It.	$\lvert q^{(n)}_{(N),1} - q_1^f \rvert$	$\lvert q^{(n)}_{(N),2} \rvert$	$\lvert v^{(n)}_{(N),1} \rvert$	$\lvert v^{(n)}_{(N),2} \rvert$	$T^{(n)}$
1	18	0.4710181849	0.0342489553	0.0112566293	0.1530321823	18.49935459
2	2	0.4710181849	0.0342489553	0.0112566293	0.1530321823	18.49935459
3	184	0.0564336136	0.0166066948	0.0018213617	0.1608312973	18.49840484
4	2	0.0564336136	0.0166066948	0.0018213616	0.1608312973	18.49840484
5	998	0.0475493891	0.0451017689	0.0046039452	0.0354990300	18.49836286

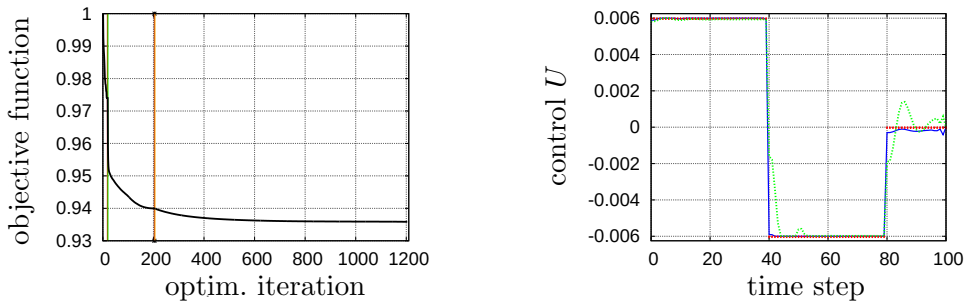

Figure 2. Reduced objective function $\mathcal{F}(U,T)$ (left), and optimization of the control U (right). On the left-hand side we show the decrease of $\mathcal{F}$ over all 5 stages (separated by vertical lines) of the penalty method yielding 1204 optimization iterations in total. $\mathcal{F}$ is normalized to the initial value or to the last value of the preceeding stage of the penalty method, respectively. On the right-hand side we see the initial control (red dotted line) versus time steps, the computed control after 1 stage (blue dashed line) and the computed optimal control (green continuous line) after 5 stages. The objective $\mathcal{F}$ is dimensionless and the control U is given in s^{-2}.

From this simulation, we find as optimal total time $T = 18.49836$ [s]. In Fig. 2 (right) we see that the initial guess of bang-bang type is quite good already, but the obtained optimal control for the final values of the weights is not of bang-bang type, though we omit a regularization term for the control by setting $\nu_3 = 0$. If we

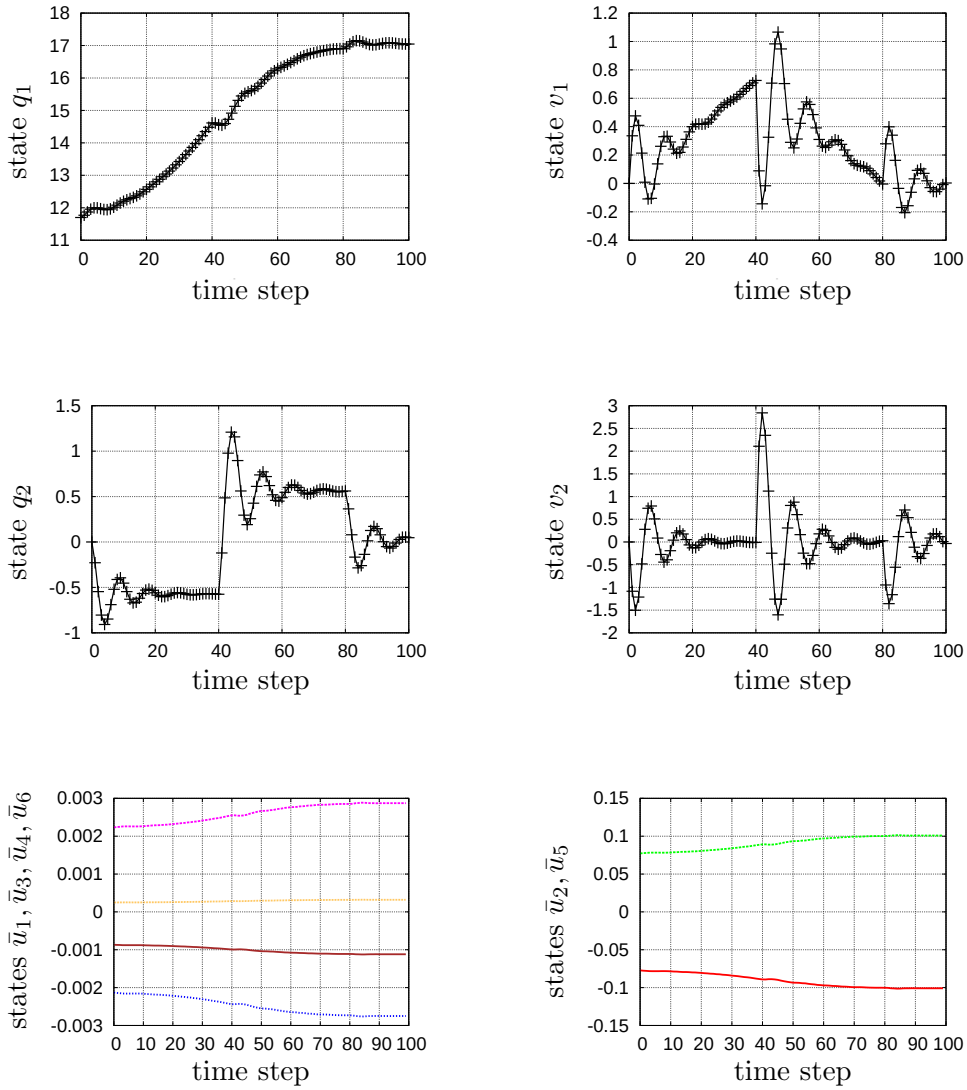

Figure 3. States $\boldsymbol{q} = (x_{\mathrm{T}}, \alpha)^{\top}$ (top/center left), $\boldsymbol{v} = \dot{\boldsymbol{q}}$ (top/center right), $\bar{\boldsymbol{u}}$ (bottom left/right) versus time step number on the interval $[0, T]$ with $T = 18.49836$ [s]; bottom left: $\bar{u}_1$ brown continuous line, $\bar{u}_3$ magenta upper dashed line, $\bar{u}_4$ blue lower dashed line, $\bar{u}_6$ orange central dashed line; bottom right: $\bar{u}_2$ red continuous line; $\bar{u}_5$ green dashed line. The quantities q_1, v_1 are given in m, q_2, v_2 in degree, $\bar{u}_j$, $j = 1, 2, 5, 6$, are dimensionless and $\bar{u}_3$, $\bar{u}_4$ are given in m^{-1}.

start with an arbitrary initial guess, then the computing times are increased and not for any initial guess convergence is obtained. By Fig. 3 we convince ourselves that the terminal conditions $q_1^0 = 17.0, q_2^f = v_1^f = v_2^f = 0.00$ are respected within good approximation. In Fig. 3, bottom left, we see $\bar{u}_2 + \bar{u}_5 = |\Gamma_{\mathrm{C}}|^{-1} \int_{\Gamma_{\mathrm{C}}} \partial_1 u_3 + \partial_3 u_1 \approx 0$, this reflects the observation that the (orthogonal) shear stress $\tau_{13} = \tau_{31} = \mu(\partial_1 u_3 + \partial_3 u_1)$ vanishes on the trolley-beam contact surface, representing a mainly isotropic compression or tension on the surface. Fig. 3, bottom right, shows $\bar{u}_3 + \bar{u}_4 \approx 0$, corresponding to $\partial_1 \operatorname{trace}(\boldsymbol{\nabla u}) \approx 0$, that is, the average pressure on Γ_{C} is constant

in x_1-direction. Keep in mind that displacements in x_2-direction are very small due to the non-rotating crane.

By increasing the fineness of the mesh, we checked that our choice of Lagrange $\mathbb{P}_2$-elements and the applied spatial discretization avoid the well-known locking effects for Timoshenko beams [31, Ch. VI, §3]. For a mesh with about twice the number of cells, we obtain almost identical results. By the penalty method, Algorithm 4.4, we are able to obtain a lower value of the objective $\mathcal{F}^{(n)}$ than by solving directly for a fixed choice of weights. Furthermore, using the Euler method for time discretization requires several thousands of iterations while by the Heun method our algorithm terminates within several hundreds of iterations.

5.2. *Numerical time-optimal control using a modified BFGS update*

It turns out that time-optimal control of our problem works faster, when applying a projected quasi-Newton method relying on the BFGS update. However, decreasing optimality and feasibility tolerances further for time-optimal control, yields Armijo steps close to the computing precision and long computing times. Therefore we focus on the case of a fixed terminal time in the following.

5.3. *Numerical optimal control for fixed terminal time*

Our numerical optimal control presented in Subsection 5.1 turns out to require large computing times for a decreased feasibility tolerance err_1. We examine this phenomenon by considering the optimal control of the problem but now with $\nu_1 = 0$, $\nu_3 = 0.1$, a fixed terminal time $T = 19.0$ [s], and $N = 500$ time steps. Our results are depicted in Fig. 4 for same data as in Subsection 5.1, except for $err_1 = 10^{-8}$, and absolute error tolerance $err_0 = 10^{-7}$ for the optimality, where as stopping criterion we use

$$|T^{(n+1)} - T^{(n)}| + \|U^{(n+1)} - U^{(n)}\|_\infty < err_0$$

instead of (54). Contrary to the last subsection, we consider $q_1^0 = 12.0$ [m] and $q_1^f = 18.0$ [m], in order to demonstrate that our algorithm does not work only for particular initial and terminal conditions of the trolley. Since the computed states and control show a qualitative behavior similar to Fig. 3, we omit them here.

We see that our combined Algorithms 4.2 and 4.4 perform more accurately in case of a fixed terminal time. This suggests that the convergence issues in Subsection 5.1 are due to the nonlinear time-optimal control. Other explanations for this observation could be the non-existence of a time-optimal control or that a further time grid refinement, turning out to be a computational challenge, is required in case of a free terminal time. Typically the precise resolution of switching points of controls is numerically expensive.

5.4. *Observations from simulations*

For the solution of the linear 3d FE systems of moderate size (43 200 cells, 21 100 degrees of freedom) for this initial study, it turned out to be sufficient to employ a direct method which re-uses the factorization of the stiffness matrix.

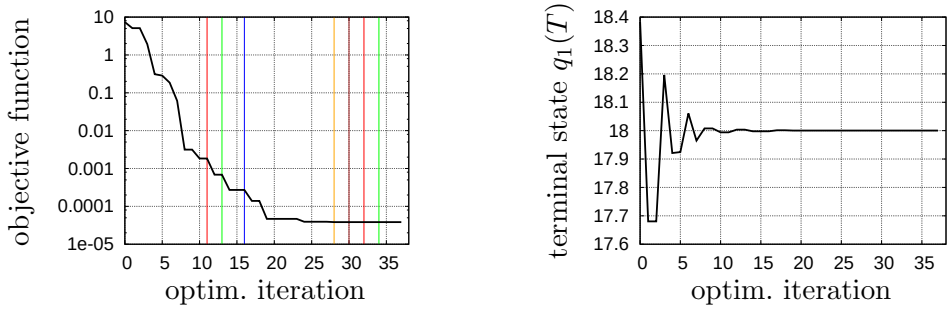

Figure 4. Reduced objective function $\mathcal{F}(U,T)$ (left), and convergence of $q_1(T)$ vs. q_1^f (right) for fixed terminal time T. On the left-hand side we show the decrease of $\mathcal{F}$ over all 8 stages (separated by vertical lines) of the penalty method yielding 38 optimization iterations in total. $\mathcal{F}$ is normalized to the initial value or to the last value of the preceeding stage of the penalty method, respectively. On the right-hand side we see the approach to the terminal trolley position q_1^f vs. the optimization iteration. The objective $\mathcal{F}$ is dimensionless and the terminal trolley position $q_1(T)$ is given in m.

From further simulations we noticed the following coupling effects and dependencies:

- For heavier crane beams, that is, larger values of the mass density ρ, the trolley position q_1 moves faster to the free end and it may happen that the mass matrix $\mathbf{M}$ becomes singular.
- The longer the pendulum, that is, the larger the value of l, the smaller $\max_\Omega \|\boldsymbol{u}\|$. Then the faster q_1, the more q_2 deviates from 0 and $\det(\mathbf{M})$ tends to zero.
- For a crane beam of half the length, $\max_\Omega \|\boldsymbol{u}\|$ becomes smaller and $\det(\mathbf{M})$ is very close to 1.
- The typical ratio of trolley and load masses is $\eta = m_\mathrm{T}/m_\mathrm{L} \approx 0.05$. For $\eta \to \infty$ while the scaled total mass m remains constant, the values of $\dot{q}_1$ decrease and $\max_\Omega \|\boldsymbol{u}\|$ becomes smaller. In the opposite case, for lower values of η such as 0.01, q_1 is faster and larger angles q_2 appear.
- The inclusion of mechanical displacements has a significant impact on the speed of the trolley, for example, considering $\bar{\boldsymbol{u}}$ terms yields a slowdown of the trolley of about 10% compared to an inelastic ($\bar{\boldsymbol{u}} \equiv \mathbf{0}$) trolley system.
- Adding a damping term to the ODEs for the trolley and for the load, respectively, yields that the amplitudes of the oscillations decay faster, but the qualitative effect is neglectable on the optimal control of the whole coupled system (see [23]).

In particular, the latter observation underlines the importance of taking elastic deformations into account, represented by the terms involving $\bar{\boldsymbol{u}}$ in the ODE.

5.5. *Discussion*

With the beam fixed on both sides, that is, homogeneous Dirichlet b.c. for $\boldsymbol{u}$ on Γ_D as well as on $\Gamma_{\mathrm{D,r}} = \partial\Omega \cap \{x_1 = l_2 - b_1/2\}$, our situation can be thought of as an overhead gantry crane as used frequently in environments as different as high rise

racks or seaports. For $m_{\mathrm{L}} \to 0$ one might think of a train or truck ("trolley") on a very long beam bridge ("elastic structure") that might easily tend to vibrate, see [36] for the model and simulations. The considerations of this article could be easily applied to these elastic crane bridge-trolley-load or elastic bridge-vehicle situations.

Our results of Section 5.4 underline the necessity to incorporate elastic deformations into the standard trolley-load system. We compare with the optimal controls obtained by Chen and Gerdts [4, 5] for such a trolley-load system without elasticity. They have applied smoothed Newton methods for the optimal control. Most data is of same order as in our case, but they simulate for a significantly larger mass ratio $\eta = 0.6$ and for the case $\nu_1 = \nu_2 = 0$ and $\nu_3 > 0$. The numerical results resemble our figures, but we observe more oscillations after a change in the control. This might be due to mechanical effects and to the lower value for η. In particular, when no control costs enter the objective function, we do not necessarily end up with a bang-bang control (see Fig. 2, right). It is not clear so far whether the "bang-bang-principle", known for optimal control of ODE and of semilinear elliptic PDE apart, does not hold for this kind of ODE-PDE-constrained OCP or whether our observation is only a numerical artefact. A possible explanation for this new phenomenon might be that the crane beam constitutes an infinite-dimensional system of states. From the work of Pesch *et al.* [8, 20, 21], we see that they encounter also controls that are not of bang-bang type. Please note that our situation differs from the Neumann boundary control of semilinear elliptic PDE, where a bang-bang result holds [37], since we consider a control in time acting by a shift on the Neumann boundary condition.

We give some reflections over coupled ODE-PDE control problems. Although this class of problems has many real-life applications, only few results exist as discussed in the introduction. This study shows a typical case study for this class of OCP, yielding a richer variety of controls than OCP with ODE, DAE or PDE constraints alone. We remark that the coupled ODE-PDE system could be regarded as a partial differential algebraic equation system (PDAE) with differential time index 1 and differential space index infinity [38]. However, the literature on PDAE, even on basic definitions, seems to be small.

6. Conclusion & Outlook

We have formulated an OCP for a coupled elastic crane-trolley-load system, proved analytically the local-in-time well-posedness of the coupled ODE-PDE problem, presented a solution algorithm and computed a first numerical optimal control for typical data. The following challenges had to be overcome:

- the complexity of the model and the involved scales,
- a special algorithm had to be developed for this non-standard problem, involving ODE *and* PDE constraints and possibly a lack of differentiability of $\boldsymbol{u}$ with respect to q_1,
- the trolley-load ODE system resembles a double pendulum system, exhibiting i) exponential reaction on perturbations and ii) possibly chaotic behaviour,
- the solution times for the PDE (yielding a large number of degrees of freedom) are an issue.

However, the computation of the numerical optimal control could still be improved with respect to computing times and resolution (the latter could be improved, e.g., by use of suitable variational integrators), in particular in the case of time-optimal

control. Provided an improved algorithm, the optimal control problem of a crane rotating as well could be tackled numerically.

Finally, we would like to close with an outlook. It might be useful to employ automatic differentiation for the calculation of sensitivities. Work-in-progress includes (i) modelling and numerics of the pendulum in 3d, (ii) possibly faster and more accurate first-optimize-then-discretize methods, and (iii) to examine whether we might guarantee terminal conditions by Newton-type methods instead of penalty techniques. However, it is not obvious whether a Newton method can be applied to our problem as in Chen and Gerdts [4, 5]. This depends on the smoothness of the necessary optimality conditions that is not clear, since we need the Fréchet differentiabilty of $\boldsymbol{u}$ with respect to moving boundary conditions. The corresponding Fréchet derivative turns out to involve line measures, see [23]. Furthermore, it cannot be guaranteed that a Newton-based method is at all faster than a projected gradient method combined with exact penalty techniques, because the coupling of the problem might yield a dense Hessian matrix. In particular, open questions comprise theoretical results for ODE-PDE constrained optimal control.

References

[1] E.P. Hofer, O. Sawodny, H. Aschemann, S. Lahres, and A. Hartmann, *Verfahren zur Bahnregelung von Kranen und Vorrichtung zum bahngenauen Verfahren einer Last*, German patent application No. 199 07 989.7, 24.02.1999.

[2] H. Aschemann, *Passivity-Based Trajectory Control of an Overhead Crane by Interconnection and Damping Assignment*, in H. Ulbrich and L. Ginzinger (eds.), *Motion and Vibration Control*, Springer Science + Business Media B.V., 2009, pp. 21–30.

[3] S.K. Biswas, *Optimal control of gantry crane for minimum payload oscillations*, in *International Conference on Dynamic Systems and Applications*, Atlanta, May 20–24, 2003, Proc. of Dynam. Systems Appl. 4 (2004).

[4] J. Chen and M. Gerdts, *Smoothing technique of nonsmooth Newton methods for control-state constrained optimal control problems*, SIAM J. Numer. Anal. 50 (2012), pp. 1982–2011.

[5] J. Chen and M. Gerdts, *Numerical solution of control-state constrained optimal control problems with an inexact smoothing Newton method*, IMA J. Numer. Anal.31 (2011), pp. 1598–1624.

[6] A. Logg, K.-A. Mardal, G.N. Wells (eds.), *Automated Solution of Differential Equations by the Finite Element Method, The FEniCS Book*, Springer, Heidelberg, 2012.

[7] S.-J. Kimmerle and R. Moritz, *Optimal control of an elastic tyre-damper system with road contact*, Proc. Apply. Math. Mech. 14 (2014), pp. 875–876.

[8] K. Chudej, H.J. Pesch, M. Wächter, G. Sachs, and F. Le Bras, *Instationary heat-constrained trajectory optimization of a hypersonic space vehicle by ODE-PDE-constrained optimal control*, in G. Buttazzo and A. Frediani, eds., *Variational Analysis and Aerospace Engineering*, Springer Optimization and Its Applications Vol. 33, Springer, New York (2009), pp. 127–144.

[9] J.-M. Coron, *Control and Nonlinearity*, AMS, Providence, 2009.

[10] M. Gerdts and S.-J. Kimmerle, *Numerical optimal control of a coupled ODE-PDE model of a truck with a fluid basin*, Discrete Contin. Dyn. Syst. 2015, Suppl., pp. 515–524.

[11] R. Herzog and K. Kunisch, *Algorithms for PDE-constrained optimization*, in *GAMM-Mitteilungen* vol. 33, Wiley-VCH Verlag, Weinheim, 2010, pp. 163–176.

[12] I. Lasiecka and R. Triggiani, *Control theory for partial differential equations: Continuous and approximation theories, Vols. I & II*, Encyclopedia of Mathematics and its Applications Vol. 74, Cambridge University Press, Cambridge, 2000.

[13] J.L. Lions, *Contrôle optimal de systeèmes gouvernés par des équations aux dérivées partielles*, Dunod and Gauthier-Villars, Paris, 1968.
[14] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich, *Optimization with PDE Constraints*, Mathematical Modelling: Theory and Applications Vol. 23, Springer, New York, 2009.
[15] F. Tröltzsch, *Optimal Control of Partial Differential Equations, Theory, Methods and Applications*, Graduate Studies in Mathematics Vol. 112, AMS, Providence, 2010.
[16] J.T. Betts, *Practical Methods for Optimal Control and Estimation Using Nonlinear Programming*, 2nd ed., Advances in Design and Control Vol. 19, SIAM, Philadelphia, 2010.
[17] M. Gerdts, *Optimal Control of ODEs and DAEs*, De Gruyter, Berlin, 2012.
[18] S.K. Biswas and N.U. Ahmed, *Stabilization of a class of hybrid systems arising in flexible spacecraft*, J. Optim. Theory Appl. 50 (1986), pp. 83–108.
[19] S.K. Biswas and N.U. Ahmed, *Optimal control of large space structures governed by a coupled system of ordinary and partial differential equations*, Math. Control Signals Systems 2 (1989), pp. 1–18.
[20] H.J. Pesch, A. Rund, W. Von Wahl, and S. Wendl, *On some new phenomena in state-constrained optimal control if ODEs as well as PDEs are involved*, Control Cybern. 39 (2010), pp. 647–660.
[21] S. Wendl, A. Rund, and H.J. Pesch, *On a state-constrained PDE optimal control problem arising from ODE-PDE optimal control*, in M. Diehl, F. Glineur, and W. Michiels, eds., *Recent Advances in Optimization and its Applications in Engineering*, Springer, Berlin/Heidelberg, 2010, pp. 429–438.
[22] P.-G. Ciarlet, *Mathematical Elasticity Vol. 1: Three Dimensional Elasticity*, Studies in Mathematics and its Applications Vol. 20, Elsevier, Amsterdam, 1998.
[23] S.-J. Kimmerle, M. Gerdts, and R. Herzog, *An optimal control problem for a rotating elastic crane-trolley-load system,* preprint, submitted 2017. Available at `http://www.unibw.de/sven-joachim.kimmerle` under "Preprints".
[24] S.-J. Kimmerle, M. Gerdts, and R. Herzog, *Optimal control of an elastic crane-trolley-load system. A case study for optimal control of coupled ODE-PDE systems*, online version with an additional appendix. Available at `http://www.unibw.de/sven-joachim.kimmerle` under "Refereed publications".
[25] S. Nicaise, *About the Lamé system in a polygonal or a polyhedral domain and a coupled problem between the Lamé system and the plate equation. I: regularity of the solutions*, Annali della Scuola Normale Superiore di Pisa, Classe di Scienze 4^e série 19 (1992), pp. 327–361.
[26] T. Apel, A.-M. Sändig, J. R. Whiteman, *Graded Mesh Refinement and Error Estimates for Finite Element Solutions of Elliptic Boundary Value Problems in Non-smooth Domains*, Math. Methods Appl. Sci. 19 (1996), pp. 63–85.
[27] H.W. Alt, *Linear Functional Analysis*, Springer, Berlin/Heidelberg, 2016.
[28] B. Niethammer, *Derivation of the LSW theory for Ostwald ripening by homogenization methods*, Arch. Ration. Mech. Anal. 147 (1999), pp. 119–178.
[29] S.-J. Kimmerle, *Well-posedness of a coupled quasilinear parabolic and elliptic free boundary problem from a model for precipitation in crystalline solids*, preprint.
[30] R. Adams and J. Fournier, *Sobolev Spaces*, 2nd ed., Academic Press, New York, 2003.
[31] D. Braess, *Finite Elements. Theory, Fast Solvers, and Applications in Solid Mechanics*, 3rd ed., Cambridge University Press, Canbridge, UK, 2007.
[32] M.J.D. Powell, *A fast algorithm for nonlinearily constrained optimization calculation*, in G.A. Watson, ed., *Numerical Analysis*, Lecture Notes in Mathematics Vol. 630, Springer, New York, 1978.
[33] M. Arnold, B. Burgermeister, C. Führer, G. Hippmann, and G. Rill, *Numerical methods in vehicle system dynamics: state of the art and current developments*, Vehicle Syst. Dyn. (2011), pp. 1159–1207.
[34] C. Geiger and C. Kanzow, *Theorie und Numerik restringierter Optimierungsaufgaben*, Springer, Berlin, 2002.

[35] Liebherr Turmdrehkran/Tower Crane 71 EC, Technical Data sheet, Liebherr-Werk, Biberach, 2004.
[36] S.-J. Kimmerle, *Modelling, simulation and optimization of an elastic structure under moving loads*, Proc. Appl. Math. Mech. 16 (2016), pp. 697–698.
[37] H. Maurer and H.D. Mittelmann, *Optimization Techniques for Solving Elliptic Control Problems with Control and State Constraints: Part 1. Boundary Control*, Comput. Optim. Appl. 16 (2000), pp. 29–55.
[38] W.S. Martinson and P.I. Barton, *A differentiation index for partial differential-algebraic equations*, SIAM J. Sci. Comput. 21 (1999), pp. 2295–2315.

Appendix A. Right-hand side of ODE system

Within the proof of Th. 3.2 we have to estimate the following term, appearing on the right-hand side of the ODE after inverting $\mathbf{M}$,

$$\mathbf{M}^{-1}\boldsymbol{F} = \frac{1}{\det(\mathbf{M})} \begin{pmatrix} l, & -((1+\bar{u}_1)\cos q_2 + \bar{u}_2 \sin q_2) \\ *, & m(1+2\bar{u}_1) \end{pmatrix} \times \begin{pmatrix} -m\bar{u}_3\dot{q}_1^2 + ((1+\bar{u}_1)\sin q_2 - \bar{u}_2\cos q_2)\,\dot{q}_2{}^2 + U \\ -(\bar{u}_3\cos q_2 + \bar{u}_4 \sin q_2)\dot{q}_1^2 - g_{\mathrm{e}} \sin q_2 \end{pmatrix}$$

with

$$\begin{aligned} \det(\mathbf{M}) &= lm(1+2\bar{u}_1) - ((1+\bar{u}_1)\cos q_2 + \bar{u}_2 \sin q_2)^2 \\ &\approx lm(1+2\bar{u}_1) - ((1+2\bar{u}_1)\cos q_2 + 2\bar{u}_2 \sin q_2)\cos q_2 \\ &= (lm - \cos^2 q_2)\left(1 + 2\bar{u}_1 - 2\frac{\bar{u}_2 \sin q_2 \cos q_2}{lm - \cos^2 q_2}\right). \end{aligned}$$

Note that $lm = (m_{\mathrm{T}} + m_{\mathrm{L}})/m_{\mathrm{L}} > \cos^2 q_2$ if $m_{\mathrm{T}} > 0$ or $|q_2| < q_2^{max} \leq \pi/2$. Thus

$$\frac{1}{\det(\mathbf{M})} \approx \frac{1}{lm - \cos^2 q_2}\left(1 - 2\bar{u}_1 + 2\frac{\bar{u}_2 \sin q_2 \cos q_2}{lm - \cos^2 q_2}\right)$$

and we get

$$\begin{aligned} \mathbf{M}^{-1}\boldsymbol{F} \approx \frac{1}{\det(\mathbf{M})}\Bigg(& \begin{pmatrix} -\bar{u}_3(lm - \cos^2 q_2) + \bar{u}_4 \sin q_2 \cos q_2 \\ m(\bar{u}_3(1 - \cos q_2) - \bar{u}_4 \sin q_2) \end{pmatrix} \dot{q}_1^2 \\ & + \begin{pmatrix} l\,((1+\bar{u}_1)\sin q_2 - \bar{u}_2 \cos q_2) \\ -(1+2\bar{u}_1)\cos q_2 \sin q_2 + \bar{u}_2 \cos(2q_2) \end{pmatrix} \dot{q}_2^2 \\ & + \begin{pmatrix} l \\ -(1+\bar{u}_1)\cos q_2 - \bar{u}_2 \sin q_2 \end{pmatrix} U + \begin{pmatrix} (1+\bar{u}_1)\cos q_2 + \bar{u}_2 \sin q_2 \\ -m(1+2\bar{u}_1) \end{pmatrix} g_{\mathrm{e}} \sin q_2 \Bigg). \end{aligned}$$

Putting in $1/\det(\mathbf{M})$, we end up with

$$\begin{aligned}
\mathbf{M}^{-1}\boldsymbol{F} \approx & \frac{1}{lm-\cos^2 q_2}\begin{pmatrix} -\bar{u}_3(lm-\cos^2 q_2)+\bar{u}_4\sin q_2\cos q_2 \\ m(\bar{u}_3(1-\cos q_2)-\bar{u}_4\sin q_2)\end{pmatrix}\dot{q}_1^2 \\
& + \frac{1}{lm-\cos^2 q_2}\begin{pmatrix} l\left((1-\bar{u}_1+2\frac{\bar{u}_2\sin q_2\cos q_2}{lm-\cos^2 q_2})\sin q_2-\bar{u}_2\cos q_2\right) \\ -(1+2\frac{\bar{u}_2\sin q_2\cos q_2}{lm-\cos^2 q_2})\cos q_2\sin q_2+\bar{u}_2\cos(2q_2)\end{pmatrix}\dot{q}_2^2 \\
& + \frac{1}{lm-\cos^2 q_2}\begin{pmatrix} l(1-2\bar{u}_1+2\frac{\bar{u}_2\sin q_2\cos q_2}{lm-\cos^2 q_2}) \\ -(1-\bar{u}_1+2\frac{\bar{u}_2\sin q_2\cos q_2}{lm-\cos^2 q_2})\cos q_2-\bar{u}_2\sin q_2\end{pmatrix}U \\
& + \frac{1}{lm-\cos^2 q_2}\begin{pmatrix} (1-\bar{u}_1+2\frac{\bar{u}_2\sin q_2\cos q_2}{lm-\cos^2 q_2})\cos q_2+\bar{u}_2\sin q_2 \\ -m(1+2\frac{\bar{u}_2\sin q_2\cos q_2}{lm-\cos^2 q_2})\end{pmatrix}g_{\mathrm{e}}\sin q_2
\end{aligned}$$

within our approximation of small displacement gradients.

Appendix B. Calculation of sensitivities and gradient for the projected gradient method

For completeness, we state here how the sensitivity equations are solved by the implemented Heun method. The first Heun step yields

$$\begin{aligned}
\mathbf{M}^{(n)}_{(k)}\delta_{U^{(n)}_{(i)}}\tilde{\boldsymbol{v}}^{(n)}_{(k+1)} &= \mathbf{M}^{(n)}_{(k)}\delta_{U^{(n)}_{(i)}}\boldsymbol{v}^{(n)}_{(k)} + hT^{(n)}\delta_{U^{(n)}_{(i)}}\boldsymbol{F}^{(n)}_{(k)} \\
&\quad - \delta_{U^{(n)}_i}\mathbf{M}^{(n)}_{(k)}\left(\tilde{\boldsymbol{v}}^{(n)}_{(k+1)}-\boldsymbol{v}^{(n)}_{(k)}\right), \quad i=0,\dots,N, \\
\mathbf{M}^{(n)}_{(k)}\delta_{T^{(n)}}\tilde{\boldsymbol{v}}^{(n)}_{(k+1)} &= \mathbf{M}^{(n)}_{(k)}\delta_{T^{(n)}}\boldsymbol{v}^{(n)}_{(k)} + h\left(T^{(n)}\delta_{T^{(n)}}\boldsymbol{F}^{(n)}_{(k)}+\boldsymbol{F}^{(n)}_{(k)}\right) \\
&\quad - \delta_{T^{(n)}}\mathbf{M}^{(n)}_{(k)}\left(\tilde{\boldsymbol{v}}^{(n)}_{(k+1)}-\boldsymbol{v}^{(n)}_{(k)}\right), \\
\delta_{U^{(n)}_{(i)}}\tilde{\boldsymbol{q}}^{(n)}_{(k+1)} &= \delta_{U^{(n)}_{(i)}}\boldsymbol{q}^{(n)}_{(k)} + hT^{(n)}\delta_{U^{(n)}_{(i)}}\boldsymbol{v}^{(n)}_{(k)}, \quad i=0,\dots,N, \\
\delta_{T^{(n)}}\tilde{\boldsymbol{q}}^{(n)}_{(k+1)} &= \delta_{T^{(n)}}\boldsymbol{q}^{(n)}_{(k)} + h\left(T^{(n)}\delta_{T^{(n)}}\boldsymbol{v}^{(n)}_{(k)}+\boldsymbol{v}^{(n)}_{(k)}\right)
\end{aligned}$$

and the second Heun step reads

$$\begin{aligned}
\tilde{\mathbf{M}}^{(n)}_{(k+1)}\delta_{U^{(n)}_{(i)}}\boldsymbol{v}^{(n)}_{(k+1)} &= \tilde{\mathbf{M}}^{(n)}_{(k+1)}\frac{1}{2}\left(\delta_{U^{(n)}_{(i)}}\boldsymbol{v}^{(n)}_{(k)}+\delta_{U^{(n)}_{(i)}}\tilde{\boldsymbol{v}}^{(n)}_{(k+1)}\right)+\frac{hT^{(n)}}{2}\delta_{U^{(n)}_{(i)}}\tilde{\boldsymbol{F}}^{(n)}_{(k+1)} \\
&\quad - \delta_{U^{(n)}_i}\tilde{\mathbf{M}}^{(n)}_{(k+1)}\left(\boldsymbol{v}^{(n)}_{(k+1)}-\frac{1}{2}\left(\boldsymbol{v}^{(n)}_{(k)}+\tilde{\boldsymbol{v}}^{(n)}_{(k+1)}\right)\right), \\
&\qquad\qquad i=0,\dots,N, \\
\tilde{\mathbf{M}}^{(n)}_{(k+1)}\delta_{T^{(n)}}\boldsymbol{v}^{(n)}_{(k+1)} &= \tilde{\mathbf{M}}^{(n)}_{(k+1)}\frac{1}{2}\left(\delta_{T^{(n)}}\boldsymbol{v}^{(n)}_{(k)}+\delta_{T^{(n)}}\tilde{\boldsymbol{v}}^{(n)}_{(k+1)}\right)+\frac{hT^{(n)}}{2}\delta_{T^{(n)}}\tilde{\boldsymbol{F}}^{(n)}_{(k+1)} \\
&\quad + \frac{h}{2}\tilde{\boldsymbol{F}}^{(n)}_{(k)} - \delta_{T^{(n)}}\tilde{\mathbf{M}}^{(n)}_{(k+1)}\left(\boldsymbol{v}^{(n)}_{(k+1)}-\frac{1}{2}\left(\boldsymbol{v}^{(n)}_{(k)}+\tilde{\boldsymbol{v}}^{(n)}_{(k+1)}\right)\right),
\end{aligned}$$

and

$$\delta_{U_{(i)}^{(n)}} \boldsymbol{q}_{(k+1)}^{(n)} = \frac{1}{2}\left(\delta_{U_{(i)}^{(n)}} \boldsymbol{q}_{(k)}^{(n)} + \delta_{U_{(i)}^{(n)}} \tilde{\boldsymbol{q}}_{(k+1)}^{(n)}\right) + \frac{hT^{(n)}}{2} \delta_{U_{(i)}^{(n)}} \tilde{\boldsymbol{v}}_{(k+1)}^{(n)}, \quad i = 0, \ldots, N,$$
$$\delta_{T^{(n)}} \boldsymbol{q}_{(k+1)}^{(n)} = \frac{1}{2}\left(\delta_{T^{(n)}} \boldsymbol{q}_{(k)}^{(n)} + \delta_{T^{(n)}} \tilde{\boldsymbol{q}}_{(k+1)}^{(n)}\right) + \frac{h}{2}\left(T^{(n)} \delta_{T^{(n)}} \tilde{\boldsymbol{v}}_{(k+1)}^{(n)} + \tilde{\boldsymbol{v}}_{(k+1)}^{(n)}\right).$$

Within Algorithm 4.2 the following terms appear in the ODEs for sensitivities, derivatives of $\boldsymbol{F}_{(k)}^{(n)}$

$$\delta_{U_{(i)}^{(n)}} \boldsymbol{F}_{(k)}^{(n)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 2 \begin{pmatrix} -m\bar{u}_{(k),3}^{(n)} v_{(k),1}^{(n)} \\ -\left(\bar{u}_{(k),3}^{(n)} \cos q_{(k),2}^{(n)} + \bar{u}_{(k),4}^{(n)} \sin q_{(k),2}^{(n)}\right) v_{(k),1}^{(n)} \end{pmatrix} \delta_{U_{(i)}^{(n)}} v_{(k),1}^{(n)}$$
$$+ \begin{pmatrix} 2\left((1+\bar{u}_{(k),1}^{(n)}) \sin q_{(k),2}^{(n)} - \bar{u}_{(k),2}^{(n)} \cos q_{(k),2}^{(n)}\right) v_{(k),2}^{(n)} \\ 0 \end{pmatrix} \delta_{U_{(i)}^{(n)}} v_{(k),2}^{(n)}$$
$$+ \begin{pmatrix} \left((1+\bar{u}_{(k),1}^{(n)}) \cos q_{(k),2}^{(n)} + \bar{u}_{(k),2}^{(n)} \sin q_{(k),2}^{(n)}\right) |v_{(k),2}^{(n)}|^2 \\ \left(\bar{u}_{(k),3}^{(n)} \sin q_{(k),2}^{(n)} - \bar{u}_{(k),4}^{(n)} \cos q_{(k),2}^{(n)}\right) |v_{(k),1}^{(n)}|^2 - g_{\mathrm{e}} \cos q_{(k),2}^{(n)} \end{pmatrix} \delta_{U_{(i)}^{(n)}} q_{(k),2}^{(n)},$$

$$\delta_{T^{(n)}} \boldsymbol{F}_{(k)}^{(n)} = 2 \begin{pmatrix} -m\bar{u}_{(k),3}^{(n)} v_{(k),1}^{(n)} \\ -\left(\bar{u}_{(k),3}^{(n)} \cos q_{(k),2}^{(n)} + \bar{u}_{(k),4}^{(n)} \sin q_{(k),2}^{(n)}\right) v_{(k),1}^{(n)} \end{pmatrix} \delta_{T^{(n)}} v_{(k),1}^{(n)}$$
$$+ \begin{pmatrix} 2\left((1+\bar{u}_{(k),1}^{(n)}) \sin q_{(k),2}^{(n)} - \bar{u}_{(k),2}^{(n)} \cos q_{(k),2}^{(n)}\right) v_{(k),2}^{(n)} \\ 0 \end{pmatrix} \delta_{T^{(n)}} v_{(k),2}^{(n)}$$
$$+ \begin{pmatrix} \left((1+\bar{u}_{(k),1}^{(n)}) \cos q_{(k),2}^{(n)} + \bar{u}_{(k),2}^{(n)} \sin q_{(k),2}^{(n)}\right) |v_{(k),2}^{(n)}|^2 \\ \left(\bar{u}_{(k),3}^{(n)} \sin q_{(k),2}^{(n)} - \bar{u}_{(k),4}^{(n)} \cos q_{(k),2}^{(n)}\right) |v_{(k),1}^{(n)}|^2 - g_{\mathrm{e}} \cos q_{(k),2}^{(n)} \end{pmatrix} \delta_{T^{(n)}} q_{(k),2}^{(n)},$$

and the symmetric derivatives of the mass matrices

$$\delta_X \mathbf{M}_{(k)}^{(n)} = \begin{pmatrix} 0, & -(1+\bar{u}_{(k),1}^{(n)}) \sin q_{(k),2}^{(n)} + \bar{u}_{(k),2}^{(n)} \cos q_{(k),2}^{(n)} \\ *, & 0 \end{pmatrix} \delta_X q_{(k),2}^{(n)}$$

where $X \in \{U_{(i)}^{(n)}, T^{(n)}\}$. Note that for ease of presentation, we have neglected terms of the form $\delta_X \bar{\boldsymbol{u}}_{(k)}^{(n)}$ here (corresponding to (9)). The additional terms read:

$$\delta_X \boldsymbol{F}_{(k)}^{(n)} = \ldots + \begin{pmatrix} \sin q_{(k),2}^{(n)} |v_{(k),2}^{(n)}|^2 \\ 0 \end{pmatrix} \delta_X \bar{u}_{(k),1}^{(n)} - \begin{pmatrix} \cos q_{(k),2}^{(n)} |v_{(k),2}^{(n)}|^2 \\ 0 \end{pmatrix} \delta_X \bar{u}_{(k),2}^{(n)}$$
$$+ \begin{pmatrix} -m|v_{(k),1}^{(n)}|^2 \\ -\cos q_{(k),2}^{(n)} |v_{(k),1}^{(n)}|^2 \end{pmatrix} \delta_X \bar{u}_{(k),3}^{(n)} + \begin{pmatrix} 0 \\ -\sin q_{(k),2}^{(n)} |v_{(k),1}^{(n)}|^2 \end{pmatrix} \delta_X \bar{u}_{(k),4}^{(n)},$$

and

$$\delta_X \mathbf{M}^{(n)}_{(k)} = \ldots + \begin{pmatrix} 2m,\ \cos q^{(n)}_{(k),2} \\ *\ , \qquad 0 \end{pmatrix} \delta_X \bar{u}^{(n)}_{(k),1} + \begin{pmatrix} 0,\ \sin q^{(n)}_{(k),2} \\ *, \qquad 0 \end{pmatrix} \delta_X \bar{u}^{(n)}_{(k),2}.$$

For computing $\delta_{U^{(n)}_{(i)}} \bar{\boldsymbol{u}}^{(n)}_{(k)}$ and $\delta_{T^{(n)}} \bar{\boldsymbol{u}}^{(n)}_{(k)}$ it is required to solve partial differential equations for sensitivities of $\partial_1 u^{(n)}_{(k)}$, $\partial_3 \boldsymbol{u}^{(n)}_{(k)}$, and $D_1 \boldsymbol{u}^{(n)}_{(k)}$:

$$\begin{aligned} -\operatorname{div} \sigma(\delta_X \partial_1 \boldsymbol{u}^{(n)}_{(k)}) &= 0 && \text{in } \Omega \times \{\tau_k\}, \\ \delta_X \partial_1 \boldsymbol{u}^{(n)}_{(k)} &= 0 && \text{on } \Gamma_{\mathrm{D}} \times \{\tau_k\}, \\ -\sigma(\delta_X \partial_1 \boldsymbol{u}^{(n)}_{(k)}).n &= \delta_X \partial_1 \boldsymbol{R}^{(n)}_{(k)} && \text{on } \Gamma_{\mathrm{N}} \times \{\tau_k\}, \end{aligned}$$

where according to (47)

$$\delta_X \partial_1 \boldsymbol{R}^{(n)}_{(k)} = \frac{1}{|\Gamma_{\mathrm{C}}|} \left(\delta_X \boldsymbol{R}^{(n)}_{0,(k)}\ \partial_1 \tilde{\chi}^{(n)}_{\Gamma_{\mathrm{C}},(k)} + \boldsymbol{R}^{(n)}_{0,(k)}\ \left(\delta_{q^{(n)}_{(k),1}}\ \partial_1 \tilde{\chi}^{(n)}_{\Gamma_{\mathrm{C}},(k)} \right) \delta_X q^{(n)}_{(k),1} \right), \quad \text{(B1)}$$

with

$$\begin{aligned} &\delta_{U^{(n)}_{(i)}} \boldsymbol{R}^{(n)}_{0,(k)} \\ &= \begin{pmatrix} (1-\bar{u}^{(n)}_{(k),6})(1+g\cos q^{(n)}_{(k),2}\,\delta_{U^{(n)}_{(i)}} q^{(n)}_{(k),2}) + \bar{u}^{(n)}_{(k),2}\, g \sin q^{(n)}_{(k),2}\,\delta_{U^{(n)}_{(i)}} q^{(n)}_{(k),2}) \\ 0 \\ \bar{u}^{(n)}_{(k),5}(1+g\cos q^{(n)}_{(k),2}\,\delta_{U^{(n)}_{(i)}} q^{(n)}_{(k),2}) + (1-\bar{u}^{(n)}_{(k),1})\, g \sin q^{(n)}_{(k),2}\,\delta_{U^{(n)}_{(i)}} q^{(n)}_{(k),2}) \end{pmatrix} \\ &\quad + \begin{pmatrix} -\delta_{U^{(n)}_{(i)}} \bar{u}^{(n)}_{(k),6}(U^{(n)} + g \sin q^{(n)}_{(k),2}) - \delta_{U^{(n)}_{(i)}} \bar{u}^{(n)}_{(k),2}\, g(\cos q^{(n)}_{(k),2} + \eta) \\ 0 \\ \delta_{U^{(n)}_{(i)}} \bar{u}^{(n)}_{(k),5}(U^{(n)} + g \sin q^{(n)}_{(k),2}) + \delta_{U^{(n)}_{(i)}} \bar{u}^{(n)}_{(k),1}\, g(\cos q^{(n)}_{(k),2} + \eta) \end{pmatrix}, \end{aligned} \quad \text{(B2)}$$

$$\begin{aligned} &\delta_{T^{(n)}} \boldsymbol{R}^{(n)}_{0,(k)} \\ &= \begin{pmatrix} (1-\bar{u}^{(n)}_{(k),6}) g \cos q^{(n)}_{(k),2}\,\delta_{T^{(n)}} q^{(n)}_{(k),2} + \bar{u}^{(n)}_{(k),2}\, g \sin q^{(n)}_{(k),2}\,\delta_{T^{(n)}} q^{(n)}_{(k),2}) \\ 0 \\ \bar{u}^{(n)}_{(k),5}\, g \cos q^{(n)}_{(k),2}\,\delta_{T^{(n)}} q^{(n)}_{(k),2} + (1-\bar{u}^{(n)}_{(k),1})\, g \sin q^{(n)}_{(k),2}\,\delta_{T^{(n)}} q^{(n)}_{(k),2}) \end{pmatrix} \\ &\quad + \begin{pmatrix} -\delta_{T^{(n)}} \bar{u}^{(n)}_{(k),6}(U^{(n)} + g \sin q^{(n)}_{(k),2}) - \delta_{T^{(n)}} \bar{u}^{(n)}_{(k),2}\, g(\cos q^{(n)}_{(k),2} + \eta) \\ 0 \\ \delta_{T^{(n)}} \bar{u}^{(n)}_{(k),5}(U^{(n)} + g \sin q^{(n)}_{(k),2}) + \delta_{T^{(n)}} \bar{u}^{(n)}_{(k),1}\, g(\cos q^{(n)}_{(k),2} + \eta) \end{pmatrix}, \end{aligned} \quad \text{(B3)}$$

and

$$\partial_1 \tilde{\chi}^{(n)}_{\Gamma_C,(k)} = -\frac{2L}{\delta_1}\left(\frac{x_1 - q^{(n)}_{(k),1}}{\delta_1}\right)^{L-1} \tanh\left(\left(\frac{x_1 - q^{(n)}_{(k),1}}{\delta_1}\right)^{L}\right) \times \\ \times \left(1 - \tanh^2\left(\left(\frac{x_1 - q^{(n)}_{(k),1}}{\delta_1}\right)^{L}\right)\right).$$

We see that $\delta_X \partial_3 \boldsymbol{u}^{(n)}_{(k)} \equiv 0$, yielding $\delta_X \bar{u}^{(n)}_{(k),5} = \delta_X \bar{u}^{(n)}_{(k),6} \equiv 0$. This allows to simplify (B2) and (B3).
We observe $\delta_X D_1 \boldsymbol{u}^{(n)}_{(k)} = \delta_X \partial^2_{1,1} \boldsymbol{u}^{(n)}_{(k)}$, thus

$$\begin{aligned} -\operatorname{div} \sigma(\delta_X \partial_1 \boldsymbol{u}^{(n)}_{(k)}) &= 0 && \text{in } \Omega \times \{\tau_k\}, \\ \delta_X \partial_1 \boldsymbol{u}^{(n)}_{(k)} &= 0 && \text{on } \Gamma_D \times \{\tau_k\}, \\ -\sigma(\delta_X \partial_1 \boldsymbol{u}^{(n)}_{(k)}).n &= \delta_X \partial_{1,1} \boldsymbol{R}^{(n)}_{(k)} && \text{on } \Gamma_N \times \{\tau_k\} \end{aligned}$$

where according to (B1)

$$\delta_X \partial_{1,1} \boldsymbol{R}^{\varepsilon}_{(k)} = \frac{1}{|\Gamma_C|}\left(\delta_X \boldsymbol{R}^{(n)}_{0,(k)} \partial_{1,1} \tilde{\chi}^{(n)}_{\Gamma_C,(k)} + \boldsymbol{R}^{(n)}_{0,(k)} \left(\delta_{q^{(n)}_{(k),1}} \partial_{1,1} \tilde{\chi}^{(n)}_{\Gamma_C,(k)}\right) \delta_X q^{(n)}_{(k),1}\right).$$

We have

$$\partial_{1,1} \tilde{\chi}^{(n)}_{\Gamma_C,(k)} = -\frac{2L^2}{\delta_1^2}\left(\frac{x_1 - q^{(n)}_{(k),1}}{\delta_1}\right)^{L-2} \left[\frac{L-1}{L} \tanh\left(\left(\frac{x_1 - q^{(n)}_{(k),1}}{\delta_1}\right)^{L}\right) + \right. \\ \left. + \left(\frac{x_1 - q^{(n)}_{(k),1}}{\delta_1}\right)^{L} \left(1 - 3\tanh^2\left(\left(\frac{x_1 - q^{(n)}_{(k),1}}{\delta_1}\right)^{L}\right)\right)\right] \times \\ \times \left(1 - \tanh^2\left(\left(\frac{x_1 - q^{(n)}_{(k),1}}{\delta_1}\right)^{L}\right)\right).$$

Both involved PDE for sensitivities with respect to $\bar{\boldsymbol{u}}^{(n)}_{(k)}$ are solved by a fixed point iteration over $\delta_X \bar{\boldsymbol{u}}^{(n)}_{(k)}$.

When computing the gradient of the reduced objective function (48) the following intermediate terms appear. We get as derivatives with respect to states $q^{(n)}_{(k)}$

$$\delta_{q^{(n)}_{(N),1}} J^{(n)} = \nu_4 \,(q^{(n)}_{(N),1} - q^f_1), \quad \partial_{q^{(n)}_{(N),2}} J^{(n)} = \nu_5 \, q^{(n)}_{(N),2},$$

with respect to velocities $\boldsymbol{v}_{(k)}^{(n)}$

$$\delta_{v_{(N),1}^{(n)}} J^{(n)} = \nu_6\, v_{(N),1}^{(n)}, \quad \delta_{v_{(k),2}^{(n)}} J^{(n)} = \nu_2\, hT^{(n)} v_{(k),2}^{(n)}, \quad k = 1, \dots N-1,$$
$$\delta_{v_{(N),2}^{(n)}} J^{(n)} = \nu_2\, hT^{(n)} \frac{1}{2} v_{(N),2}^{(n)} + \nu_7\, v_{(N),2}^{(n)},$$

and with respect to controls

$$\delta_{U_{(i)}^{(n)}} J^{(n)} = \nu_3\, hT^{(n)} S_{(i)} U_{(i)}^{(n)}, \quad i = 0, \dots N,$$
$$\delta_{T^{(n)}} J^{(n)} = \nu_1 + \frac{\nu_2}{2} h \sum_{k=0}^{N} S_{(k)} |v_{(k),2}^{(n)}|^2 + \frac{\nu_3}{2} h \sum_{i=0}^{N} S_{(i)} |U_{(i)}^{(n)}|^2.$$