

Computational Parametric Sensitivity Analysis of Perturbed PDE Optimal Control Problems with State and Control Constraints

Christof Büskens* Roland Griesse†

July 15, 2005

Abstract

We study parametric optimal control problems governed by a system of time-dependent partial differential equations (PDE) and subject to additional control and state constraints. An approach is presented to compute optimal control functions and so-called *sensitivity differentials* of the optimal solution with respect to perturbations. This information plays an important role in the analysis of optimal solutions as well as in real-time optimal control.

The method of lines is used to transform the perturbed PDE system into a large system of ordinary differential equations. A subsequent discretization is discussed that transcribes parametric ODE optimal control problems into perturbed nonlinear programming problems (NLP) which can be solved efficiently by SQP methods.

Second order sufficient conditions can be checked numerically, and we propose to apply an NLP-based approach for robust computation of sensitivity differentials of optimal solutions with respect to the perturbation parameters.

The advertised numerical method is illustrated by the optimal control and sensitivity analysis of the Burgers equation. This example demonstrates the general ability of the algorithm to efficiently and robustly calculate an accurate numerical solution.

Keywords: perturbed optimal control problems; control-state constraints; nonlinear programming methods; partial differential equations; parametric sensitivity analysis

1 Introduction

This paper is concerned with the numerical solution and sensitivity analysis of perturbed optimal control problems governed by time-dependent partial differential equations and subject to control and state constraints. Such control problems play an

*Professor für Mathematik, Zentrum für Technomathematik, Universität Bremen, D-28334 Bremen, Germany

†Senior Scientist, Johann Radon Institute for Computational and Applied Mathematics (RICAM), Austrian Academy of Sciences, Altenbergerstraße 69, A-4040 Linz, Austria

important role in the natural sciences and other disciplines, where numerous real-life applications exist. A growing interest in optimization techniques has been stimulated by theoretical and numerical investigation of these problems. In practice, control problems are often subject to disturbances or perturbations in the system data. In mathematical terms, perturbations can be expressed by means of parameters appearing in the dynamics, boundary conditions or in control and state constraints. *Stability* and *sensitivity analysis* are concerned with the behavior of optimal solutions under parameter perturbations. The so-called parametric sensitivity derivatives are a helpful tool in the analysis and assessment of optimal solutions for practitioners. In addition, these derivatives are being widely used in real-time optimal control applications, see e.g. Pesch (Ref. 1) and Büskens and Maurer (Refs. 2–5).

There exist many recent papers on the numerical solution and theoretical treatment of optimal control problems subject to partial differential equations of different type. Numerical aspects of solutions to unconstrained problems by SQP methods are studied, e.g., by Kupfer and Sachs (Ref. 6), and Heinkenschloss (Ref. 7). Problems with additional constraints are discussed in Heinkenschloss and Sachs (Ref. 8), Tröltzsch (Ref. 9), Goldberg and Tröltzsch (Ref. 10), Casas (Ref. 11), Casas, Tröltzsch and Unger (Refs. 12, 13), Ito and Kunisch (Refs. 14, 15) and Volkwein (Ref. 16), Maurer and Mittelmann (Refs. 17, 18) to name only some references. The quantity of articles occupied with this interesting field of mathematical optimization demonstrates the growing interest in this class of problems.

In the past, numerical algorithms were often in use before the related theory was completely developed. For instance, direct methods for the solution of optimal control problems have been used successfully since the end of the 1960s. However, first results in convergence theory of the discretized optimal control problem towards the continuous solution did not appear until the 1990s, see e.g. Dontchev and Hager (Ref. 19) and Malanowski, Büskens and Maurer (Ref. 20). The present paper is written in this spirit, aware of the fact there exists a gap between the computational and theoretical aspects of parametric sensitivity analysis of optimal control problems with PDEs.

In this paper the method of lines is used to discretize the partial differential equations, transforming the original system into a system of ordinary differential equations. We obtain a perturbed optimal control problem for an ODE system with constraints that can be solved by well-known standard techniques.

Malanowski (Ref. 21), Malanowski and Maurer (Ref. 22), Maurer and Pesch (Ref. 23) have studied *differential properties* of optimal solutions. The theoretical framework in these papers rests on *indirect* methods using boundary value techniques. Following the general philosophy of discretization, it is straightforward to develop techniques to compute sensitivity differentials via nonlinear optimization methods applied to a subsequent discretization of the ODE control problem instead of working with boundary value methods. This leads to so-called *direct optimization methods* which have been studied extensively in the last 20 years. Direct optimization methods have proved to be powerful tools for solving ODE optimal control problems; cf., e.g., Büskens (Ref. 24), Büskens and Maurer (Refs. 2–5) and the references cited therein. The basic idea of direct optimization methods is to discretize the control problem and to apply nonlinear programming techniques to the resulting finite-dimensional optimization problem. These methods use only control and state variables as opti-

mization variables and completely dispense with adjoint variables. The latter can eventually be obtained by a post-optimal calculation using the Lagrange multipliers of the resulting nonlinear optimization problem.

Second order sufficient conditions (SSC) for *continuous* control problems represent an essential prerequisite for sensitivity analysis and convergence of discretized problems and are usually difficult to verify. Alt (Ref. 25), Dontchev and Hager (Ref. 19) and Malanowski, Büskens and Maurer (Ref. 20) discuss the question of convergence of the discretized problem to the continuous solution for optimal control problems involving *ordinary* differential equations. For general optimal control problems involving *partial* differential equations, these results do not yet exist. Hence, we tacitly assume the convergence to the presumed unique continuous solution as the mesh size tends to zero. However, SSC for the *discretized* control problem can be easily tested using well-known linear algebra techniques for nonlinear optimization problems. This paper follows the ideas of Büskens (Ref. 24), Büskens and Maurer (Refs. 2–5) where sensitivity analysis of perturbed ODE optimal control problems is discussed.

To the authors' knowledge there do not exist any papers which deal with the *numerical computation* of sensitivity differentials for optimal control problems for PDEs at all. A reason might be found in the fact that for a closed theoretical approach to solution differentiability and sensitivity analysis second order sufficient conditions are needed for which research is just at the beginning, cf. the book by Bonnans and Shapiro (Ref. 26), Malanowski (Ref. 27), Malanowski and Tröltzsch (Ref. 28), Mittelmann (Ref. 29) and Raymond and Tröltzsch (Ref. 30). Nevertheless the proposed NLP-based method is capable of computing approximations to sensitivity differentials for the state, control and adjoint variables in parametric PDE optimal control problems.

The general mathematical structure of perturbed PDE optimal control problems is outlined in section 2. Discretization details for the PDE can be found in section 3. Section 4 is concerned with the solution of ODE optimal control problems by direct methods, which in turn are approximated by a suitable discretization to obtain a perturbed nonlinear optimization problem. A short overview of basic results on sensitivity analysis and solution differentiability for perturbed finite-dimensional nonlinear optimization problems is offered in section 5. By re-transforming the numerical solution and sensitivity quantities, information about the optimal solution and the sensitivity of the original PDE optimal control problem can be gained. This will be covered in section 6. Finally, section 7 presents two numerical examples for PDE optimal control problems with constraints. The numerical methods presented in sections 3–6 lead to a complete numerical solution including states, controls, adjoints, and their sensitivity differentials.

2 Parametric PDE optimal control problems

Let $n_y, n_x, n_u, n_p \in \mathbb{N}^+$ be given positive numbers. Let $t \in [t_0, t_f] = \Omega_t \subset \mathbb{R}$ and $x \in \Omega_x \subset \mathbb{R}^{n_x}$ of dimension n_x denote the *time* and *spatial variables*, respectively. Here, Ω_x is a bounded domain with piecewise smooth boundary $\Gamma_x = \partial\Omega_x$. Let $\Omega = \Omega_x \times \Omega_t$ and $\Gamma = \Gamma_x \times \Omega_t$. Moreover, let $y : \Omega \rightarrow \mathbb{R}^{n_y}$ be a vector function of

dimension n_y of which $y_x(x, t)$, $y_{xx}(x, t)$ are the first and second partial derivatives of the components of y w.r.t. the spatial variables x . Likewise, $y_t(x, t)$ denotes the first derivative w.r.t. the time variable t . In the sequel, the *perturbations* are characterized by a parameter vector $p \in \Omega_p \subset \mathbb{R}^{n_p}$. The vector function u of dimension n_u has components defined either on $\Omega_x \times \Omega_t$ (*distributed control*) or on the boundary $\Gamma_x \times \Omega_t$ (*boundary control*).

We consider the following general coupled time-dependent PDE system:

$$y_t(x, t) = f(y(x, t), y_x(x, t), y_{xx}(x, t), u(x, t), x, t, p). \quad (1)$$

A solution of (1) depends on the spatial variable x , the time variable t , the control function u and the fixed parameter vector p .

While we mainly have a parabolic PDE system in mind, hyperbolic and PDE systems of mixed type can be considered, too, but the method of lines (Section 3) may have to be replaced by other discretization techniques, transforming the PDE-optimal control problem into a finite-dimensional nonlinear programming problem (22). We point out again that—here and in the sequel—questions of existence, numerical solution and its convergence for PDEs are not the focus of this paper. Instead, we aim at presenting a general method to compute sensitivity differentials.

Initial conditions of (1) with respect to time may depend on the perturbation parameter p and are defined by

$$y(x, t_0) = y_0(x, p) \quad \text{for } x \in \Omega_x. \quad (2)$$

In principle, terminal conditions like

$$\int_{\Omega_x} \psi(y(x, t_f), p) dx = 0$$

can be taken into account by our method.

On the boundary Γ_x , the partial differential equation (1) has to satisfy Dirichlet or Neumann conditions of the form

$$\begin{aligned} y(x, t) &= y_D(u(x, t), x, t, p) & \text{for } x \in \Gamma_x, t \in \Omega_t \\ \partial_\nu y(x, t) &= y_N(y(x, t), u(x, t), x, t, p) & \text{for } x \in \Gamma_x, t \in \Omega_t \end{aligned} \quad (3)$$

for given functions $y_D : \mathbb{R}^{n_u} \times \Gamma_x \times \Omega_t \times \Omega_p \rightarrow \mathbb{R}^{n_y}$ or $y_N : \mathbb{R}^{n_y} \times \mathbb{R}^{n_u} \times \Gamma_x \times \Omega_t \times \Omega_p \rightarrow \mathbb{R}^{n_y}$. The derivative in the direction of the outward unit normal ν of Γ_x is denoted by ∂_ν in (3). Combinations of Dirichlet and Neumann boundary conditions are also admitted, always provided that the resulting PDE system in (1) is well-defined and uniquely solvable for given u and p .

In addition, inequality constraints on the control and state subject to disturbances have to be observed: In case of distributed control,

$$\begin{aligned} u_{min}(p) \leq u(x, t) \leq u_{max}(p) & \quad \text{for } (x, t) \in \overline{\Omega_x} \times \Omega_t, \\ C(y(x, t), u(x, t), x, t, p) \leq 0 & \quad \text{for } (x, t) \in \overline{\Omega_x} \times \Omega_t \end{aligned} \quad (4)$$

for suitable vectors $u_{min}(p), u_{max}(p) \in \mathbb{R}^{n_u} \cup \{-\infty, \infty\}$ and a vector function $C : \mathbb{R}^{n_y} \times \mathbb{R}^{n_u} \times \overline{\Omega_x} \times \Omega_t \times \Omega_p \rightarrow \mathbb{R}^{n_c}$, $n_c \geq 0$. The values $-\infty$ and ∞ for u_{min} or

u_{max} characterize the unconstrained cases. We point out that in the presence of state constraints, the analysis of the optimal control problem is severely more difficult. Nevertheless, the proposed method can, in principle, take this type of constraints into account.

The problem is to determine a control vector — containing boundary or distributed control elements or both — that minimizes the functional

$$F(y, u, p) = \int_{\Gamma} f_1(y(x, t), u(x, t), x, t, p) dx dt + \int_{\Omega} f_2(y(x, t), u(x, t), x, t, p) dx dt. \quad (5)$$

Note that (5) can be extended to contain more general terms, for example of Mayer form which evaluate the state only at the final time t_f .

Hence the time-dependent partial differential equation optimal control problem **POCP**(\mathbf{p}) with constraints is defined by

$$\mathbf{POCP}(\mathbf{p}) : \begin{array}{ll} \text{Minimize} & F(y, u, p) \\ \text{subject to} & (1)-(4). \end{array} \quad (6)$$

Problem **POCP**(\mathbf{p}) is of very general form. Hence a precise description of how to calculate the solution, perform a sensitivity analysis and compute the sensitivity differentials as advertised in the introduction is a complex task. To facilitate the presentation, it is appropriate to perform some simplifications on problem (6). Please note that these are stimulated only by conceptual reasons. The basic principle of sensitivity analysis by NLP methods described in the following sections works as well for the original problem (6). Let us assume in a first step that the pure control constraint in (4) are picked up by the constraints C .

3 Discretization of PDE control problems

The underlying idea is to transform the partial differential equation into a system of ordinary differential equations by discretizing all functions with respect to the spatial variable x . This approach is known as the method of lines. Let us emphasize that the methods presented in this and the following section are not the only ones possible to transform a PDE control problem into a finite-dimensional nonlinear programming problem. However, the results obtained in Section 5 and 6 will still hold, independent of the discretization technique used.

The discussion is restricted to Dirichlet conditions. As an example, we present discretization by finite differences. The finite element method and others are also possible choices.

Let $\tilde{\Omega}_x$ and $\tilde{\Gamma}_x$ be a finite difference grid approximation of $\bar{\Omega}_x$ and Γ_x , respectively. For simplicity, we restrict ourselves to the case $\Omega_x = (0, 1)^{n_x}$:

Let $N_x \geq 2$ be a natural number and $h_x := \frac{1}{N_x-1} \in \mathbb{R}^+$. A discrete approximation $\tilde{\Omega}_x$ of $\bar{\Omega}_x$ and $\tilde{\Gamma}_x$ of Γ_x is defined by

$$\begin{aligned} \tilde{\Omega}_x &:= \{x = h_x \cdot (i^1, \dots, i^{n_x}) \in \bar{\Omega}_x \mid i^1, \dots, i^{n_x} \in \{0, 1, \dots, N_x-1\}\}, \\ \tilde{\Gamma}_x &:= \Omega_x \setminus \Omega_x. \end{aligned} \quad (7)$$

To enumerate all $x \in \tilde{\Omega}_x$ and $x \in \tilde{\Gamma}_x$, the index sets

$$\begin{aligned}\tilde{\Omega}_x^I &:= \left\{ (i^1, \dots, i^{n_x}) \in \mathbb{N}^{n_x} \mid \exists x \in \tilde{\Omega}_x, x = h_x \cdot (i^1, \dots, i^{n_x}) \right\}, \\ \tilde{\Gamma}_x^I &:= \left\{ (i^1, \dots, i^{n_x}) \in \tilde{\Omega}_x^I \mid \exists x \in \tilde{\Gamma}_x, x = h_x \cdot (i^1, \dots, i^{n_x}) \right\}\end{aligned}\quad (8)$$

will be used in the sequel. Instead of calculating the continuous solution on Ω_x for the partial differential equation in (1), the PDE is solved on the discretized grid $\tilde{\Omega}_x$. The spatial variable $x \in \Omega_x$ is replaced by $x \in \tilde{\Omega}_x$. Hence the partial derivatives for y in (1) with respect to x have to be substituted by their finite difference approximation. For reasons of clarity let $e_i \in \mathbb{R}^{n_x}$ denote the i -th unit vector and define

$$\begin{aligned}x_l &:= h_x \cdot (i^1, \dots, i^{n_x}), \quad l := (i^1, \dots, i^{n_x}) \in \tilde{\Omega}_x^I, \\ \bar{l}^k &:= l + e_k = (i^1, \dots, i^{k-1}, i^k + 1, i^{k+1}, \dots, i^{n_x}), \\ \underline{l}^k &:= l - e_k = (i^1, \dots, i^{k-1}, i^k - 1, i^{k+1}, \dots, i^{n_x}), \\ \bar{l}^{k,j} &:= \bar{l}^k + e_j = l + e_k + e_j, \\ \underline{l}^{k,j} &:= \underline{l}^k - e_j = l - e_k - e_j, \\ \bar{\underline{l}}^{k,j} &:= \bar{l}^k - e_j = l + e_k - e_j = \underline{l}^j + e_k.\end{aligned}\quad (9)$$

Let $l \in \tilde{\Omega}_x^I$. Then $y_x(x_l, t) = (y_{x^1}(x_l, t), \dots, y_{x^{n_x}}(x_l, t))$ is approximated by finite differences, e.g.,

$$y_{x^k}(x_l, t) \approx \frac{y(x_{\bar{l}^k}, t) - y(x_{\underline{l}^k}, t)}{2h_x}, \quad \text{whenever } l, \bar{l}^k, \underline{l}^k \in \tilde{\Omega}_x^I. \quad (10)$$

Second order derivatives can be approximated in the same manner, provided that $l, \bar{l}^k, \underline{l}^k, \bar{l}^{k,j}, \underline{l}^{k,j}, \bar{\underline{l}}^{j,k}, \underline{\underline{l}}^{k,j} \in \tilde{\Omega}_x^I$:

$$\begin{aligned}y_{x^k x^k}(x_l, t) &\approx \frac{y(x_{\bar{l}^k}, t) - 2y(x_{l^k}, t) + y(x_{\underline{l}^k}, t)}{(h_x)^2}, \\ y_{x^k x^j}(x_l, t) &\approx \frac{y(x_{\bar{l}^{k,j}}, t) - y(x_{\underline{l}^{k,j}}, t) - y(x_{\bar{\underline{l}}^{j,k}}, t) + y(x_{\underline{\underline{l}}^{k,j}}, t)}{(h_x)^2}, \quad k \neq j.\end{aligned}\quad (11)$$

If boundary values are required in equations (10) or (11), the Dirichlet condition in (3) can be substituted directly into the finite difference approximations (10) and (11). In the sequel, approximations of $y_x(x_l, t)$ and $y_{xx}(x_l, t)$ for $x_l \in \tilde{\Omega}_x \setminus \tilde{\Gamma}_x$ using formulas (10) and (11) are denoted by $\tilde{y}_x(x_l, t)$ and $\tilde{y}_{xx}(x_l, t)$. Hence applying the conventions in (9) for $l := (i^1, \dots, i^{n_x}) \in \tilde{\Omega}_x^I$, an approximation of problem (6) on $\Omega_x = (0, 1)^{n_x}$

with Dirichlet conditions only is given by the following semi-discretized formulation:

$$\begin{aligned}
& \text{Minimize} \\
F(y, u, p) &= \int_{t_0}^{t_f} h_x^{n_x-1} \sum_{k \in \tilde{\Gamma}_x^I} f_1(y(x_k, t), u(x_k, t), x_k, t, p) dt \\
&+ \int_{t_0}^{t_f} h_x^{n_x} \sum_{l \in \tilde{\Omega}_x^I} f_2(y(x_l, t), u(x_l, t), x_l, t, p) dt \tag{12}
\end{aligned}$$

subject to

$$\begin{aligned}
y_t(x_l, t) &= f(y(x_l, t), \tilde{y}_x(x_l, t), \tilde{y}_{xx}(x_l, t), u(x_l, t), x_l, t, p) \quad \forall l \in \tilde{\Omega}_x^I \setminus \tilde{\Gamma}_x^I \\
y(x_l, t_0) &= y_0(x_l, p) \quad \text{for } x_l \in \tilde{\Omega}_x \\
y(x_l, t) &= y_D(u(x_l, t), x_l, t, p) \quad \text{for } x_l \in \tilde{\Gamma}_x^I, t \in \Omega_t, \\
0 &\geq C(y(x_l, t), u(x_l, t), x_l, t, p) \quad \text{for } x_l \in \tilde{\Omega}_x, t \in \Omega_t.
\end{aligned}$$

Note that instead of the simple Euler discretization in the objective of (12) higher order approximations can be used as well. Due to the discretization, x no longer represents a continuous variable. Furthermore, the discretized Dirichlet condition in (12) can be inserted directly into (12) wherever needed. Hence the expressions for $y_t(x_l, t)$ represent a system of ordinary differential equations that simplifies to

$$y_t(x_{l^-}, t) = \dot{y}_{l^-}(t) = f_{l^-}(y_{l^-}(t), u_l(t), t, p) \tag{13}$$

where the subscripts $l \in \tilde{\Omega}_x^I$ and $l^- \in \tilde{\Omega}_x^I \setminus \tilde{\Gamma}_x^I$ denote the dependency on the spatial variable x_l . The other expressions in (12) can be treated in the same manner. The state vector y_{l^-} is of dimension $n_y \cdot (N_x - 2)^{n_x}$ while the dimension of the control vector u_l is at most $n_u \cdot (N_x)^{n_x}$ and depends on the particular choice of boundary and/or distributed control functions. One finds that problem (12) in fact reduces to

$$\begin{aligned}
& \text{Minimize} && F(y_{l^-}, u_l, p) \quad (\text{defined by (12)}) \\
& \text{subject to} && \dot{y}_{l^-}(t) = f_{l^-}(y_{l^-}(t), u_l(t), t, p), \\
& && y_{l^-}(t_0) = y_0^{l^-}(p), \\
& && 0 \geq C(y_{l^-}(t), u_l(t), t, p), \tag{14}
\end{aligned}$$

$\forall l \in \tilde{\Omega}_x^I, l^- \in \tilde{\Omega}_x^I \setminus \tilde{\Gamma}_x^I$ and $t \in \Omega_t$, for appropriately defined functions f_{l^-} and $y_0^{l^-}$. Equation (14) represents a perturbed optimal ODE control problem whose standard form is given by (15). Please note that the state variable on the boundary $\tilde{\Gamma}_x^I$ is not included in the state vector $y_{l^-}(t)$. However, it can be retrieved from the Dirichlet condition if required.

4 Numerical solution

The numerical solution of optimal ODE control problems by nonlinear programming techniques is well developed and there exist a number of excellent methods. In principle these methods can be divided into two classes: The first type of methods is characterized by a discretization of the state *and* control variables which results in

the ODE system being approximated by a huge number of equality constraints in the resulting NLP problem. This leads to a high-dimensional NLP problem which has a sparse structure in the Jacobian of the constraints and the Hessian of the Lagrangian. Because of the typically high dimension of the ODE system for $y_{l-}(t)$ in (14), a method is preferred which gives rise to smaller NLP problems. This motivates the second type of NLP methods for the numerical solution of optimal ODE control problems. Here only the control functions are discretized while the state variables are calculated autonomously by integrating the ODE system using suitable numerical solvers. In contrast to the first approach, one obtains a small but dense NLP problem. For the numerical solution of (14) the code NUDOCCCS (Refs. 24, 31) is used which features post-optimal calculation of the adjoint variables to high precision. In addition, a parametric sensitivity analysis of the optimal solution can be performed by the code SENSIA (Ref. 24). The results including the sensitivity information are subsequently transferred to the original perturbed optimal PDE control problem (6). In the sequel, a short summary of the underlying idea is given. For a more detailed discussion please refer to Büskens (Ref. 24) and Büskens and Maurer (Refs. 2–5).

We consider the following perturbed optimal ODE control problem (**OCP(p)**):

$$\begin{aligned}
\text{Minimize} \quad & F(w, v, p) = g(w(t_f), v(t_f), t_f, p) + \int_{t_0}^{t_f} f_0(w(t), v(t), t, p) dt \\
\text{subject to} \quad & \dot{w}(t) = f(w(t), v(t), t, p), \\
& w(t_0) = w_0(p), \\
& C(w(t), v(t), t, p) \leq 0, \quad t \in [t_0, t_f].
\end{aligned} \tag{15}$$

Here, $w(t) \in \mathbb{R}^n$ and $v(t) \in \mathbb{R}^m$ denote the state of the system and the control in a given time interval $[t_0, t_f] = \Omega_t$, respectively. Data perturbations in the system are modeled by a parameter $p \in \Omega_p$. The functions $g : \mathbb{R}^{n+m+1} \times \Omega_p \rightarrow \mathbb{R}$, $f_0 : \mathbb{R}^{n+m+1} \times \Omega_p \rightarrow \mathbb{R}$, $f : \mathbb{R}^{n+m+1} \times \Omega_p \rightarrow \mathbb{R}^n$, $w_0 : \Omega_p \rightarrow \mathbb{R}^n$, and $C : \mathbb{R}^{n+m+1} \times \Omega_p \rightarrow \mathbb{R}^k$ are assumed to be sufficiently smooth on appropriate open sets. The admissible class of control functions is the class of piecewise continuous controls. The final time t_f can be either fixed or free. Note that the formulation of mixed state-control constraints $C(w(t), v(t), p) \leq 0$ in (15) may contain *pure control constraints* $C(v(t), p) \leq 0$ as well as *pure state constraints* $C(w(t), p) \leq 0$. By identifying $w(t) = y_{l-}(t)$ and $v(t) = u_{l-}(t)$ and adjusting the other components of (14), one finds that (14) is of the form (15).

For a natural number N_t , let $\tau^i \in \Omega_t$ be the points of the time grid $\tilde{\Omega}_t = \{\tau^i \mid i = 1, \dots, N_t\}$ in ascending order:

$$t_0 = \tau^1 < \dots < \tau^{N_t-1} < \tau^{N_t} = t_f. \tag{16}$$

Moreover, to simplify notation, we assume that the discretization in (16) is equidistant:

$$h_t := \frac{t_f - t_0}{N_t - 1}, \quad \tau^i = t_0 + (i - 1) \cdot h_t, \quad i = 1, \dots, N_t. \tag{17}$$

Let the vectors $w^i \in \mathbb{R}^n$ and $v^i \in \mathbb{R}^m$, $i = 1, \dots, N_t$, be approximations at the grid points of the state variable $w(\tau^i)$ and the control variable $v(\tau^i)$, respectively. For

every choice of the discrete control variables

$$\eta := (v^1, v^2, \dots, v^{N_t-1}, v^{N_t}) \in \mathbb{R}^{N_\eta}, \quad N_\eta := m \cdot N_t, \quad (18)$$

the state variables $w(\tau^i)$ can be computed recursively, e.g., by the explicit Euler approximation

$$w^{i+1} = w^i + h_t \cdot f(w^i, v^i, \tau^i, p), \quad i = 1, \dots, N_t - 1. \quad (19)$$

The integral in the objective in (15) can be approximated by

$$\int_{t_0}^{t_f} f_0(w(t), v(t), t, p) dt \approx h_t \sum_{i=1}^{N_t-1} f_0(w^i, v^i, t^i, p). \quad (20)$$

Instead of Euler's method incorporated into the relations (19) and (20), one can use higher order integration methods combined with higher order control approximations. Either way, it follows that the state variables can be understood as functions of the control variables with initial condition $w^1 = w_0(p)$ taken from (15):

$$w^i = w^i(\eta, p) := w^i(v^1, \dots, v^{i-1}, p). \quad (21)$$

This leads to the following NLP problem (**NLP(p)**):

$$\begin{aligned} \text{Minimize} \quad & \tilde{F}(\eta, p) = \tilde{F}(w^1(\eta, p), \dots, w^{N_t}(\eta, p), \eta, p), \\ \text{subject to} \quad & C^i(\eta, p) = C(w^i(\eta, p), v^i, \tau^i, p) \leq 0, \quad i = 1, \dots, N_t, \end{aligned} \quad (22)$$

where \tilde{F} denotes an approximation of the objective in (15) according to (20).

Problems of the form (22) can be solved efficiently using sequential quadratic programming (SQP) methods; see, e.g., the codes mentioned in the survey article (Ref. 32). All calculations described in the sequel were performed by the code NUDOCCCS of Büskens (Refs. 24, 31) which provides implementations of various higher order explicit and implicit approximations for state and control variables. For the examples presented in section 7, the SQP solver E04UCF from the NAG library was used in NUDOCCCS.

Recently, the convergence of solutions discretized via Euler's method to solutions of the continuous control problem has been proved in Malanowski, Büskens and Maurer (Ref. 33). By solving the NLP problem (22) we obtain an estimate of the *continuous* control and state variables (u, y) of (6) at appropriate $x_l \in \tilde{\Omega}_x$, $\tau^j \in \tilde{\Omega}_t$, from

$$\begin{aligned} v(\tau^j) &= u_l(\tau^j) \approx u(x_l, \tau^j) \\ w(\tau^j) &= y_l(\tau^j) \approx y(x_l, \tau^j). \end{aligned} \quad (23)$$

Likewise, all other variables and functions of the continuous problem (6) can be determined approximately from the quantities in problem (22).

5 Parametric sensitivity analysis for perturbed non-linear optimization problems

In sections 3 and 4, a method to transform a perturbed control problem into a parametric NLP problem has been discussed. As was mentioned, the results hereafter do not depend on the discretization technique used there.

The Lagrangian of this problem (cf. (22)) is of the form

$$L(\eta, \zeta, p) := \tilde{F}(\eta, p) + \zeta^T G(\eta, p), \quad (24)$$

where $G(\eta, p) := (C^1(\eta, p), \dots, C^{N_t}(\eta, p))^T$ denotes the collection of all constraints in (22) and ζ are the corresponding Lagrange multipliers. At a *reference parameter* $p = p_0$, let η_0 and ζ_0 denote an optimal solution for (22) satisfying first order necessary (KKT-)optimality conditions. Define $I^a(p_0) := \{i \in \{1, \dots, r + N_t \cdot k\} \mid G_i(\eta_0, p_0) = 0\}$ as the set of active indices of size $m^a := \#I^a(p_0)$. Multipliers corresponding to active constraints $G^a := (G_i)_{i \in I^a(p_0)}$ are denoted by $\zeta_0^a \in \mathbb{R}^{m^a}$. Then the following strong second order sufficient conditions can be formulated, cf. Fiacco (Ref. 34):

Theorem: (Strong Second Order Sufficient Conditions)

Let \tilde{F}, G be twice continuously differentiable with respect to η and p , and let $G_\eta^a(\eta_0, p_0)$ have maximal rank m_a . Assume that $\zeta_0^a > 0$ and

$$\xi^T L_{\eta\eta}(\eta_0, \zeta_0, p_0) \xi > 0 \quad \forall \xi \in \text{Ker}(G_\eta^a(\eta_0, p_0)), \quad \xi \neq 0. \quad (25)$$

Then η_0 is a local minimum for (22) at $p = p_0$.

The main difficulty verifying these SSC numerically consists in establishing the positive definiteness criterion of the Hessian in (25). This can be done by evaluating the Hessian projected onto $\text{Ker}(G_\eta^a)$ and computing its eigenvalues, cf. (Ref. 24, 4).

SSCs form the basis of the sensitivity analysis for parametric optimization problems which started in the mid-seventies with the work of Fiacco (Refs. 34, 35) and Robinson (Ref. 36). They independently suggested to use the classical implicit function theorem to show differentiability of solutions to finite-dimensional parametric mathematical programs.

Theorem: (Solution Differentiability and Sensitivity)

Assume that the optimal solution (η_0, ζ_0) satisfies the strong second order sufficient conditions for the nominal problem $\mathbf{NLP}(\mathbf{p}_0)$. Then for p near p_0 , the unperturbed solution (η_0, ζ_0) can be embedded into a C^1 -family of perturbed optimal solutions $(\eta(p), \zeta(p))$ for $\mathbf{NLP}(\mathbf{p})$ such that $(\eta(p_0), \zeta(p_0)) = (\eta_0, \zeta_0)$. The active sets $I^a(p)$ coincide with $I^a(p_0)$, and hence it follows that $\zeta_i(p) = 0$ for all $i \notin I^a(p_0)$. The sensitivity differentials of the optimal solutions and Lagrange multipliers are given by the formula

$$\begin{pmatrix} \frac{d\eta}{dp}(p_0) \\ \frac{d\zeta^a}{dp}(p_0) \end{pmatrix} = - \begin{pmatrix} L_{\eta\eta}(\eta_0, \zeta_0, p_0) & G_\eta^a(\eta_0, p_0)^T \\ G_\eta^a(\eta_0, p_0) & 0 \end{pmatrix}^{-1} \begin{pmatrix} L_{\eta p}(\eta_0, \zeta_0, p_0) \\ G_p^a(\eta_0, p_0) \end{pmatrix}. \quad (26)$$

Moreover, the sensitivity of the objective function is obtained from

$$\frac{d\tilde{F}}{dp}(\eta(p_0), p_0) = L_p(\eta_0, \zeta_0, p_0). \quad (27)$$

A second order sensitivity derivative for the objective function is given by the $n_p \times n_p$ matrix

$$\frac{d^2\tilde{F}}{dp^2}[p_0] = \left(\frac{d\eta}{dp}(p_0)\right)^T L_{\eta\eta}[p_0] \left(\frac{d\eta}{dp}(p_0)\right) + 2 \left(L_{p\eta}[p_0] \frac{d\eta}{dp}(p_0)\right)^T + L_{pp}[p_0], \quad (28)$$

where the notation $[p_0]$ stands for all respective nominal arguments.

In practice, the first and second order derivatives in the right hand side of (26) can be approximated by finite differences (Ref. 24) or evaluated using automatic differentiation (Ref. 37). Note that the so-called Kuhn–Tucker matrix on the right hand side of (26) is *regular* since second order sufficient conditions (25) are assumed to hold. In the following section these results are applied to the optimal control problems defined in Section 2.

6 Sensitivity analysis for perturbed PDE optimal control problems

The numerical examples, e.g., in Malanowski and Maurer (Refs. 22, 33) and Pesch (Ref. 1) show that, already for ordinary differential equations, a numerical sensitivity analysis becomes rather tedious using indirect methods. Hence, the purpose of this section is to develop a robust direct NLP method for the computation of sensitivity derivatives of optimal solutions with respect to parameters. The method is based on the formulas (26)–(28) applied to a discretization of control problem (15), and it allows to compute approximate sensitivity differentials for state, control and adjoint variables as well as for the objective.

By $(y(p), u(p))$, we denote the solution of our original problem **POCP**(\mathbf{p}) for a given parameter $p \in \Omega_p$, and $\lambda(p)$ is the corresponding adjoint variable. Recall that the quantities of interest are the gradients $\frac{d}{dp}$ of the maps $p \mapsto y(p)$, $p \mapsto u(p)$, $p \mapsto \lambda(p)$, and $p \mapsto F(y(p), u(p), p)$, evaluated at the nominal parameter $p = p_0$. They are denoted by $\frac{dF}{dp}(y(p_0), u(p_0), p_0)$, etc. The respective notation is used for the terms in **OCP**(\mathbf{p}) and **NLP**(\mathbf{p}).

Equations (27) and (28) yield sensitivity differential approximations of first and second order for the objective function $F(y, u, p)$ in (5) by means of

$$\begin{aligned} \frac{dF}{dp}(y(p_0), u(p_0), p_0) &\approx \frac{dF}{dp}(w(p_0), v(p_0), p_0) \approx \frac{d\tilde{F}}{dp}(\eta(p_0), p_0), \\ \frac{d^2F}{dp^2}(y(p_0), u(p_0), p_0) &\approx \frac{d^2F}{dp^2}(w(p_0), v(p_0), p_0) \approx \frac{d^2\tilde{F}}{dp^2}(\eta(p_0), p_0). \end{aligned} \quad (29)$$

Note that initial value perturbations in (2), e.g.,

$$y_0(x, p) = y_0(x) + p \quad \text{for } x \in \Omega_x \quad (30)$$

lead to a simplification of (29):

$$\frac{dF}{dp}(y(p_0), u(p_0), p_0) \approx \lambda(x, t_0) \quad \text{for } x \in \Omega_x \quad (31)$$

where $\lambda(x, t_0)$ denotes the adjoint variable at $t = t_0$ belonging to the state y computed post-optimally by the code NUDOCCCS. Equation (31) represents an approximation of the well-known marginal interpretation of the adjoint variable.

Applying the expression for $\frac{d\eta}{dp}(p_0)$ from formula (26) to the discretized controls $v(\tau^j) = v^j$ provides an approximation of the sensitivity differentials of the perturbed optimal solution's control component at all mesh points (x_l, τ^j) , $x_l \in \tilde{\Omega}_x$, $\tau^j \in \tilde{\Omega}_t$, namely,

$$\frac{du}{dp}(x_l, \tau^j; p_0) \approx \frac{du_l}{dp}(\tau^j; p_0) = \frac{dv}{dp}(\tau^j; p_0) \approx \frac{dv^j}{dp}(p_0). \quad (32)$$

Likewise the sensitivity differential $\frac{dt_f}{dp}(p_0)$ for a free terminal time t_f can be calculated from equation (26) since it is handled as an additional optimization variable in (22).

The state sensitivities $\frac{dy}{dp}(x, t; p_0)$ can be approximated by

$$\frac{dy}{dp}(x_l, \tau^j; p_0) \approx \frac{dy_l}{dp}(\tau^j; p_0) = \frac{dw}{dp}(\tau^j; p_0) \approx \frac{dw^j}{dp}(\eta_0, p_0), \quad (33)$$

if we take the recursive expression (19) into account and differentiate the control–state relation (21) with respect to the parameter:

$$\frac{dw^j}{dp}(\eta_0, p_0) = \frac{\partial w^j}{\partial \eta}(\eta_0, p_0) \frac{d\eta}{dp}(p_0) + \frac{\partial w^j}{\partial p}(\eta_0, p_0). \quad (34)$$

The quantity $\frac{d\eta}{dp}(p_0)$ is taken from (26).

The sensitivities of the Lagrange multipliers $\frac{d\zeta^a}{dp}(p_0)$ in (26) not considered up to now can be used to approximate the sensitivity differentials of the adjoint variables, but this will not be discussed here.

7 Examples

We shall present a numerical example to illustrate the performance of the algorithms presented in sections 2–6. All computations were performed on a 1GHz PIII personal computer using the code NUDOCCCS.

7.1 Example 1: The Burgers Equation

We consider tracking-type optimal control of the non-linear one-dimensional Burgers equation with additional control constraints. For a detailed analysis of this problem

we refer to Volkwein (Refs. 38, 39):

$$\begin{aligned}
& \text{Minimize} \\
F(y, u, p) &= \frac{1}{2} \int_{t_0}^{t_f} \int_{\Omega} [y(x, t) - 0.035]^2 dx dt + \frac{\sigma}{2} \int_{t_0}^{t_f} [u_1(t)^2 + u_2(t)^2] dt, \\
& \text{subject to} \\
y_t(x, t) &= p_1 \nu y_{xx}(x, t) - y(x, t) y_x(x, t) + p_4 t, \\
y(x, t_0) &= p_2 x^2 (p_3 - x)(1 - x), \\
y_x(0, t) &= u_1(t), \quad y_x(1, t) = u_2(t), \\
u_{min} &\leq u_i(t) \leq u_{max}, \quad i = 1, 2,
\end{aligned} \tag{35}$$

with $n_x = n_y = 1$, $n_u = 2$, $n_p = 4$, $\Omega_t = [0, 1]$, $\Omega_x = (0, 1)$, $\sigma = 0.01$, $\nu = 0.1$, $-u_{min} = u_{max} = 0.015$. The solution of the state equation is known to exist and be unique in $W(0, t_f) \cap C(\bar{\Omega})$ where (Refs. 38, 39)

$$W(0, t_f) = \{\varphi \in L^2(0, t_f; H^2(\Omega_x) \cap H_0^1(\Omega_x)) : \varphi_t \in L^2(0, t_f; H^{-1}(\Omega_x))\}.$$

Recent research shows that the differentiability of the optimal solution depends on a certain property of the first order necessary conditions, called strong regularity, see Malanowski (Ref. 27) and Griesse (Ref. 40). Interestingly, strong regularity is also a prerequisite in proving convergence of the generalized Newton method, see Tröltzsch and Volkwein (Ref. 41). The latter authors also prove strong regularity for a distributed control problem similar to our example.

For the nominal perturbation parameter $p_0 = (1, 1, 1, 0)^T$, the optimal solution is obtained from the discretized formulation (22). All computations are performed with a discretization of $N_x = 34$ for the spatial variable x which leads to a system of 32 first order ordinary differential equations if the Neumann conditions are directly inserted into all functions. The state variables on the boundaries $y(0, t)$ and $y(1, t)$ — which are used on the right hand side of the PDE in (35) to approximate $y_x(x, t)$ and $y_{xx}(x, t)$ — are calculated by the second order approximations

$$\begin{aligned}
y(0, t) &\approx \frac{4y(h_x, t) - y(2h_x, t) - 2h_x u_1(t)}{3}, \\
y(1, t) &\approx \frac{4y(1 - h_x, t) - y(1 - 2h_x, t) + 2h_x u_2(t)}{3}.
\end{aligned} \tag{36}$$

In order to be able to use an explicit higher order Runge–Kutta scheme, it is necessary to choose the number of time grid points sufficiently high. The discretized time interval has $N_t = 151$ grid points, but due to printing resolution reasons only every fifth point in time is included in the three-dimensional plots (Figures 1, 3–6). An explicit fourth order Runge–Kutta scheme and a linear interpolation of the control variable is used for the integration in time. Similar results can be achieved using an implicit Runge–Kutta method with fewer time steps, at the expense of additional numerical cost. The spatial integral in the objective is approximated by the trapezoidal method. After about 9.5 seconds of computational time, the optimal nominal solution is obtained with an objective value $F(y(p_0), u(p_0), p_0) \approx 2.5073 \cdot 10^{-5}$. The unperturbed optimal control functions with two boundary arcs are depicted in Figure 2.

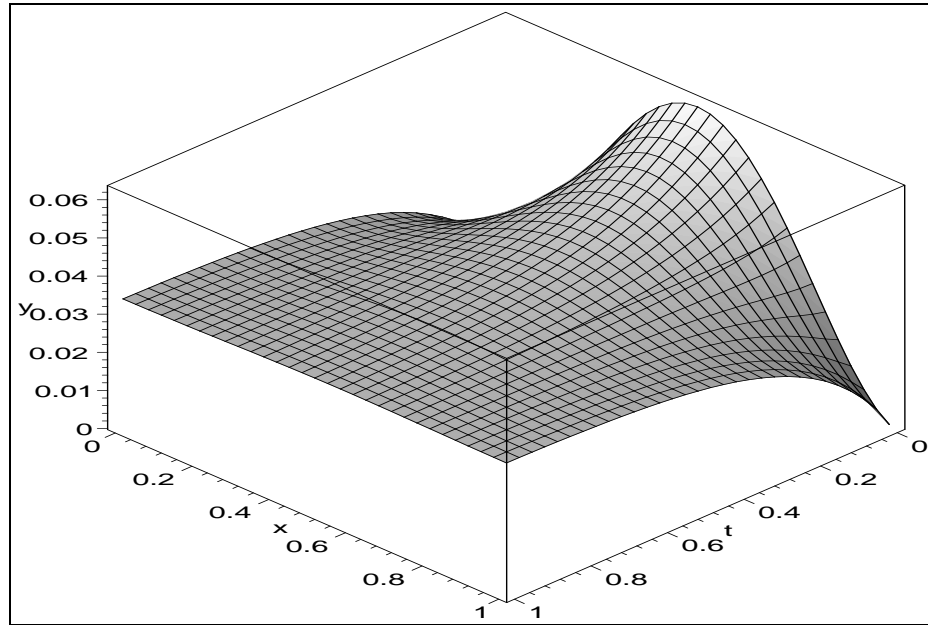


Figure 1: Optimal nominal solution $y(x, t; p_0)$.

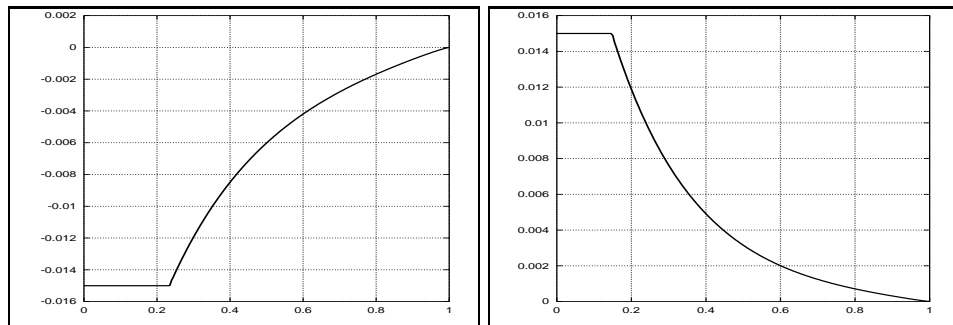


Figure 2: Nominal control $u_1(p_0)$ (left) and $u_2(p_0)$ (right).

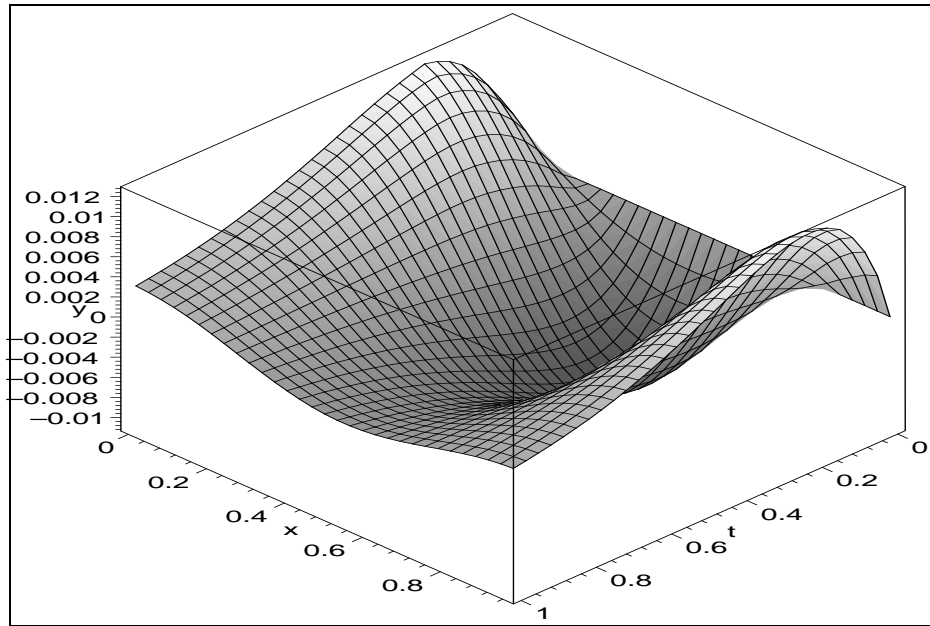


Figure 3: Sensitivities $\frac{dy}{dp_1}(x,t;p_0)$

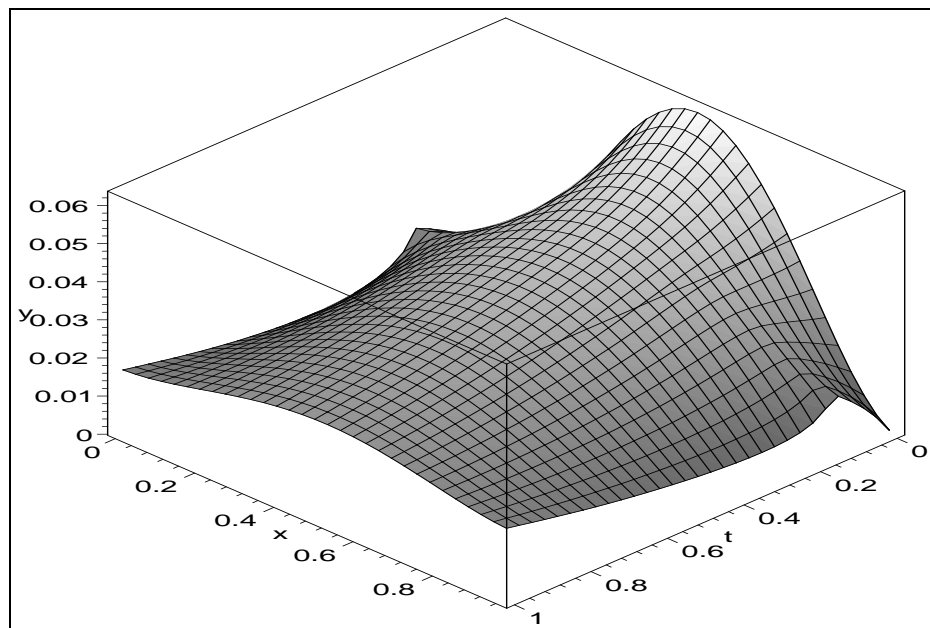


Figure 4: Sensitivities $\frac{dy}{dp_2}(x,t;p_0)$

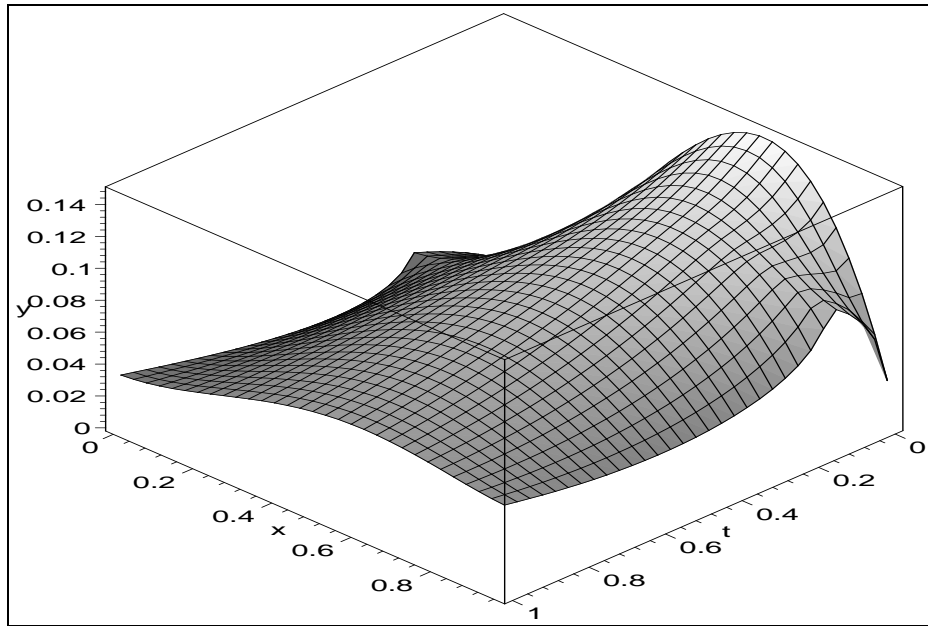


Figure 5: Sensitivities $\frac{dy}{dp_3}(x, t; p_0)$

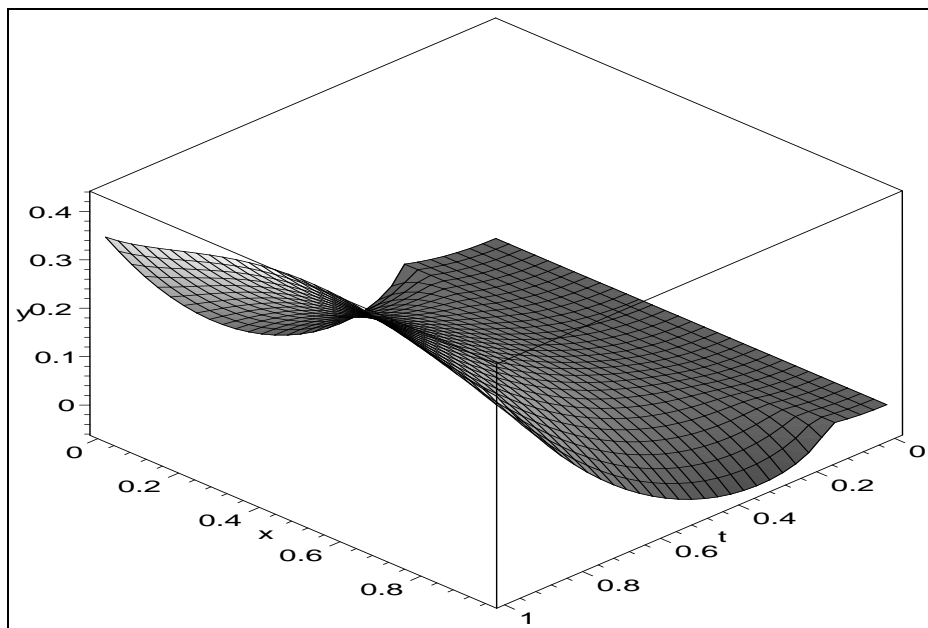


Figure 6: Sensitivities $\frac{dy}{dp_4}(x, t; p_0)$

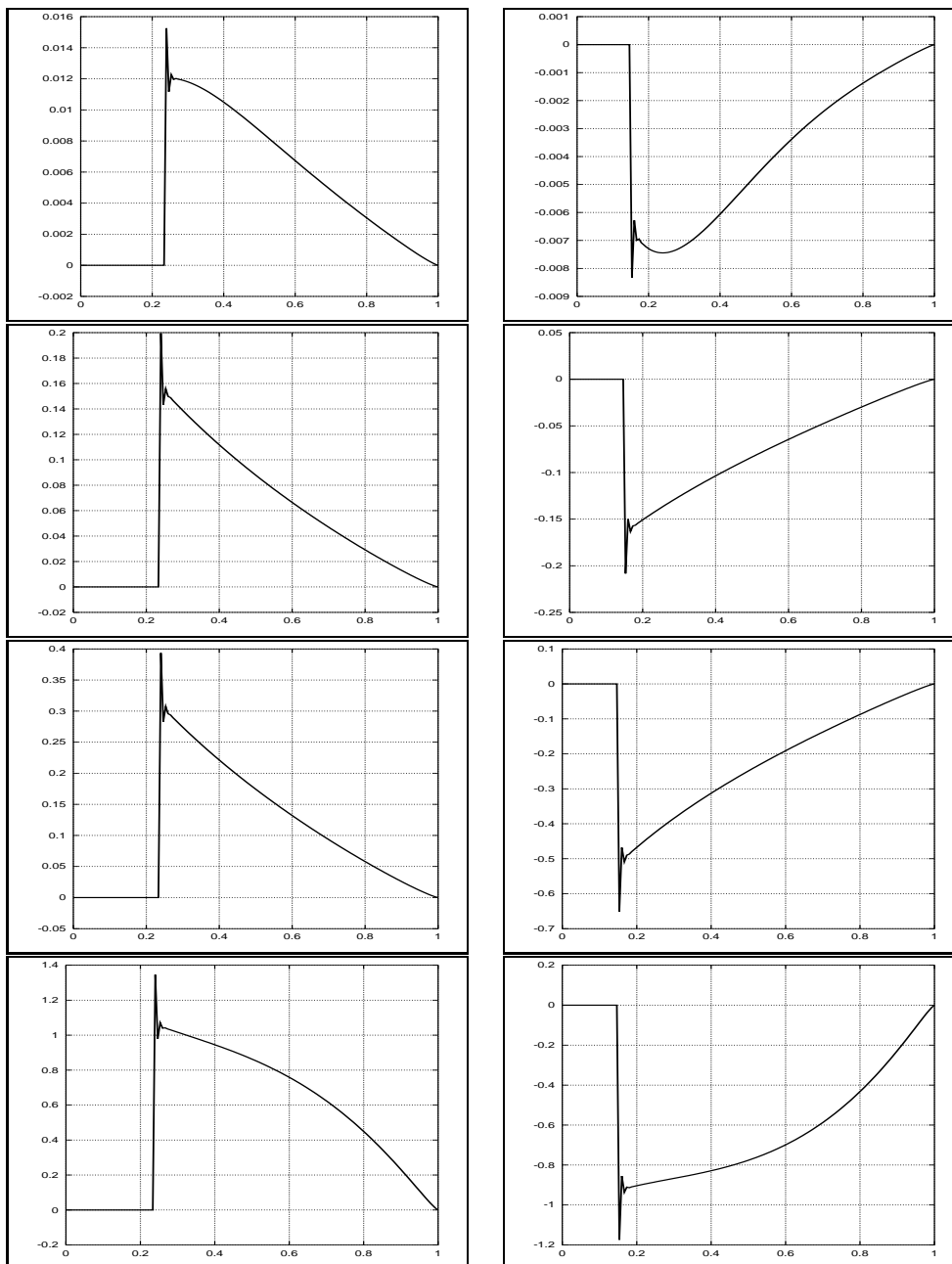


Figure 7: Sensitivities of the control functions $\frac{du_1}{dp_j}(p_0)$ (left) and $\frac{du_2}{dp_j}(p_0)$ (right).

All assumptions of the theorem on solution differentiability and sensitivity have been carefully checked numerically for the discretized problem, and the Hessian of the Lagrangian has been found to be positive definite on the kernel of the Jacobian of the active constraints. Hence the sensitivity differentials of the control and state variables and of the objective can be obtained from the expressions (26)–(34). Figures 3–6 display the respective sensitivity differentials.

Note that the sensitivities of the controls in figure 7 are zero on the boundary arcs. Moreover, from the theory known in the case of ODEs (Refs. 22, 33) one expects that the control sensitivities are discontinuous at each junction point of the control constraints. Overshooting at these points in figure 7 results from the linear interpolation of the control variables.

An investigation of the optimality conditions for problem (35) shows that since there is no observation of the terminal state in the objective, the adjoint state is zero at t_f . From the control law (where \mathcal{P} denotes projection onto the admissible interval $[u_{min}, u_{max}]$)

$$\begin{aligned} u_1(t) &= \mathcal{P} \left(\frac{\lambda(0, t)}{\sigma} \right) \\ u_2(t) &= \mathcal{P} \left(\frac{\lambda(1, t)}{\sigma} \right) \end{aligned}$$

it follows that $u_i(t_f) = 0$ holds, see figure 2. As the sensitivity quantities satisfy a linearized version of these optimality conditions, one infers that all control sensitivities will have the property that $du_i/dp_j(p_0)(t_f) = 0$, as is clearly shown in figure 7.

First and second order sensitivity derivatives of the objective function can be computed from the equations (27)–(29) which yield

$$\begin{aligned} \frac{dF}{dp}[p_0] &\approx 10^{-5} \cdot (-2.50, 4.25, 6.56, -5.15) \\ \frac{d^2F}{dp^2}[p_0] &\approx \begin{pmatrix} 4.98 \cdot 10^{-5} & -4.87 \cdot 10^{-5} & -8.38 \cdot 10^{-5} & 3.67 \cdot 10^{-5} \\ -4.87 \cdot 10^{-5} & 9.49 \cdot 10^{-4} & 2.37 \cdot 10^{-3} & 3.77 \cdot 10^{-3} \\ -8.38 \cdot 10^{-5} & 2.37 \cdot 10^{-3} & 5.96 \cdot 10^{-3} & 9.26 \cdot 10^{-3} \\ 3.67 \cdot 10^{-5} & 3.77 \cdot 10^{-3} & 9.26 \cdot 10^{-3} & 3.44 \cdot 10^{-2} \end{pmatrix}. \end{aligned} \quad (37)$$

Note that while the absolute values of the sensitivities appear small, they have to be interpreted in relation to the absolute value of the objective $F \approx 2.5073 \cdot 10^{-5}$. The sign of the entries in the first order sensitivity of the objective provide information about the direction of change of the objective under parameter changes. This follows from first-order Taylor expansion. For example, the objective will decrease when the parameter p_1 increases. However, second-order information should be included whenever possible as can be seen in the case of p_4 : $\frac{d^2F}{dp_4^2}[p_0] \approx 3.44 \cdot 10^{-2}$ has opposite sign and is large compared to $\frac{dF}{dp_4}[p_0] \approx -5.15 \cdot 10^{-5}$. A discussion of the second-order Taylor approximation reveals the influence of perturbations of different size.

The perturbation parameter p_1 is connected with the viscosity coefficient ν . Figure 3 shows the sensitivity profile for changes in $p_1\nu$. Perturbations in the initial values are studied by means of the parameters p_2 and

p_3 where the latter causes unsymmetry which is being propagated with time. The corresponding sensitivities are depicted in figure 4 and 5. The sharp corners in these quantities near the boundaries are caused by the control constraint.

Finally, p_4 produces a forcing term on the right hand side which grows with time.

7.2 Example 2: The Wave Equation

We now consider the optimal control of the one-dimensional wave equation with additional control and state constraints:

$$\begin{aligned}
& \text{Minimize} \\
F(y, u, p) &= \frac{1}{2} p_2 \int_{\Omega} y(x, t)^2 dx dt + \frac{\sigma}{2} \int_{t_0}^{t_f} u(t)^2 dt, \\
& \text{subject to} \\
y_{tt}(x, t) &= p_1 y_{xx}(x, t), \\
y(x, t_0) &= \begin{cases} \cos(2\pi(2x - 1)) + 1, & \text{if } x \in [0.25, 0.75], \\ 0, & \text{else,} \end{cases} \\
y_t(x, t_0) &= 0 \\
y(0, t) &= 0, \\
y(1, t) &= u(t), \\
0 &\leq at + b - \frac{1}{2} \int_{t_0}^t \int_{\Omega_x} y(x, t)^2 dx dt, \\
u_{min} &\leq u(t) \leq u_{max},
\end{aligned} \tag{38}$$

with $n_u = n_x = n_y = 1$, $n_p = 2$, $\Omega_t = [0, 5]$, $\Omega_x = (0, 1)$, $\sigma = 10$, $a = 0.06$, $b = 0.15$, $-u_{min} = u_{max} = 0.25$,

The goal of this control problem is to eliminate the initial wave while taking control cost into account. At $x = 1$, the control allows to absorb parts of the wave. The boundary condition at $x = 0$ causes reflection of any incoming wave.

At this time, no results are available concerning the existence of parametric sensitivity functions for optimal control problems involving the wave equation, much less in the presence of state constraints. Here, clearly the suggested numerical method precedes the theoretical investigations.

For the nominal perturbation parameter $p_0 = (1, 1)^T$, the optimal solution is obtained from the discretized formulation (22). All computations are performed with a discretization of $N_x = 34$ for the spatial variable x which leads to a system of 32 second order ordinary differential equations (or 64 first order ordinary differential equations) if the Dirichlet conditions are directly inserted into all functions. The time interval is split into $N_t = 151$ discrete points to obtain the three-dimensional plots (Figures 8, 10 and 11). Figures 9 and 12 were obtained from a discretization of $N_t = 351$ discrete points of time to highlight the subtle structures of the control and its sensitivities. An explicit fourth order Runge–Kutta scheme and a linear interpolation of the control variable is used for the integration in time. The spatial integrals in the objective and in the state constraint are approximated by the trape-

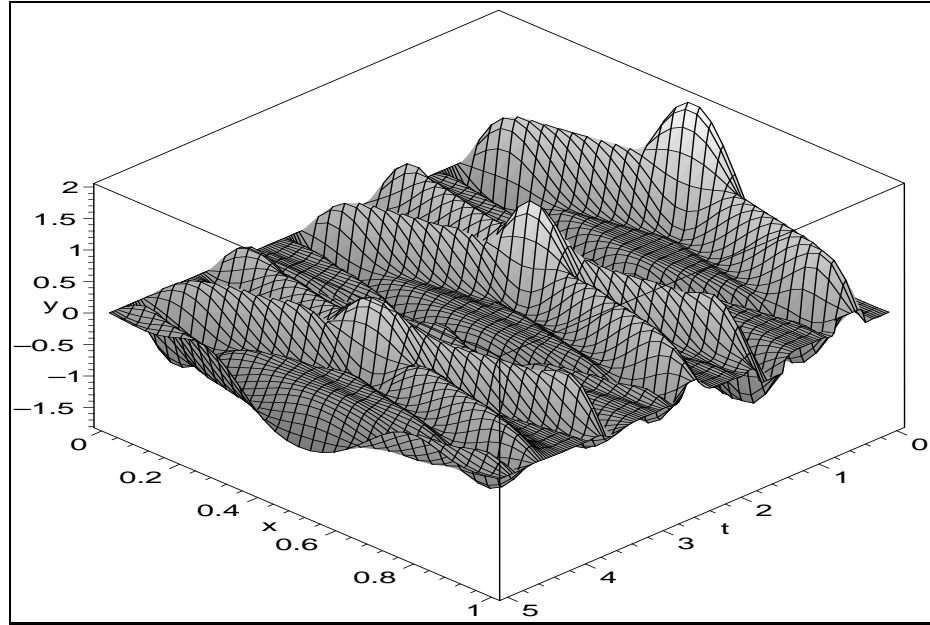


Figure 8: Optimal nominal solution $y(x, t; p_0)$.

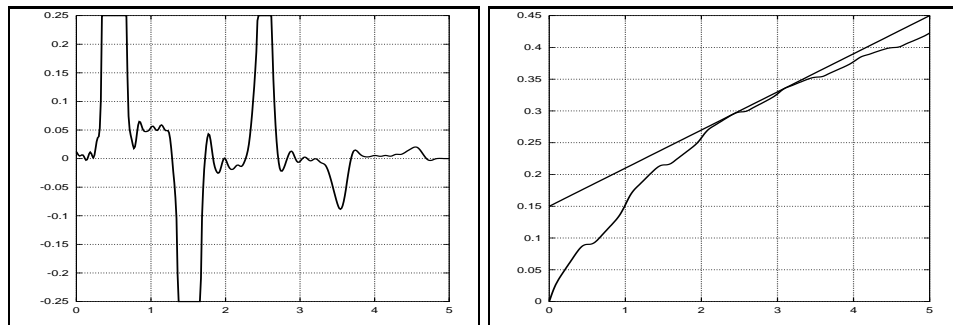


Figure 9: Nominal control $u(p_0)$ (left) and state constraint (right).

zoidal method. After about 4.5 seconds of computational time the optimal nominal solution is obtained with $F(y(p_0), u(p_0), p_0) \approx 0.7225$.

One can easily recognize the characteristic lines illustrating the typical behavior of hyperbolic PDEs propagating information. Since the control acts only on one of the boundaries one part of the initial wave is only attenuated after its reflection and return to the control boundary. Moreover, the damping of the remaining wave in time can be observed.

The unperturbed optimal control function with three boundary arcs is depicted on the left hand side of figure 9. There exist two contact points $t_1 \approx 2.43$ and $t_2 \approx 3.10$ where the state constraint in (38) becomes active, cf. the right hand side of figure 9.

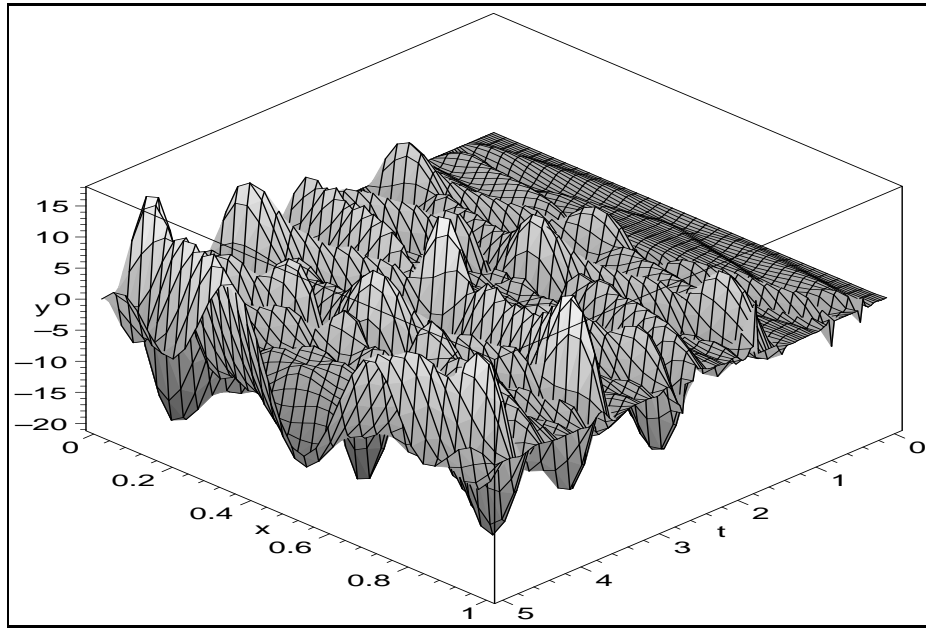


Figure 10: Sensitivities $\frac{dy}{dp_1}(x,t;p_0)$.

Again all assumptions of the theorem on solution differentiability and sensitivity have been carefully checked numerically for the discretized problem, and the Hessian of the Lagrangian has been verified to be positive definite on the kernel of the Jacobian of the active constraints. Hence the sensitivity differentials of the control and state variables and of the objective can be obtained from the expressions (26)–(34). Figures 10–12 display the respective sensitivity differentials. As expected, the characteristic lines also appear in the state sensitivities.

Note that the sensitivities of the controls in figure 12 are zero on the boundary arcs and that the overshooting at each junction point of the control constraints results from the linear interpolation of the control variables.

First and second order sensitivity derivatives of the objective function can be

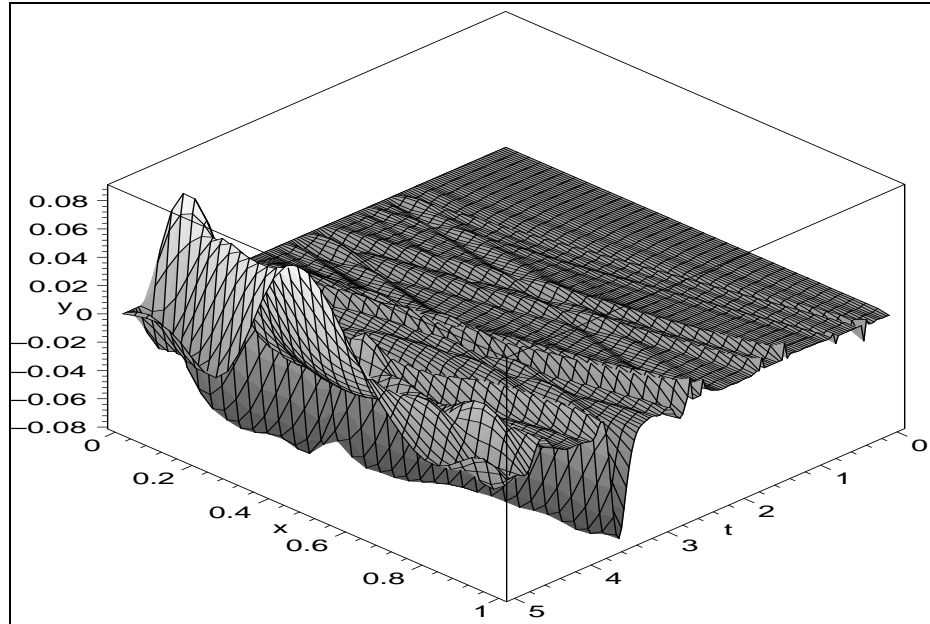


Figure 11: Sensitivities $\frac{dy}{dp_2}(x,t;p_0)$.

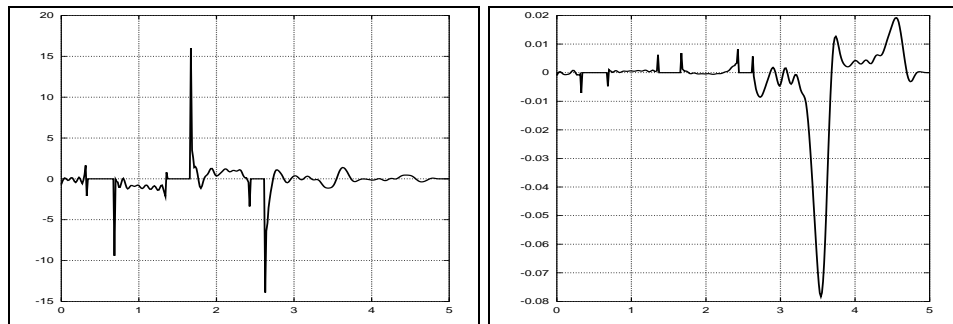


Figure 12: Sensitivities of the control function $\frac{du}{dp_1}(p_0)$ (left) and $\frac{du}{dp_2}(p_0)$ (right).

computed from the equations (27)–(29) which yield

$$\begin{aligned} \frac{dF}{dp}[p_0] &\approx (-0.8666, 0.4235) \\ \frac{d^2F}{dp^2}[p_0] &\approx \begin{pmatrix} 13.1883 & 0.0746 \\ 0.0746 & -0.0115 \end{pmatrix}. \end{aligned} \tag{39}$$

The parameter p_1 influences the wave propagation speed. Figure 10 shows the corresponding state sensitivity.

The tracking-type cost depends on the parameter p_2 . In figure 11, one can find the sensitivity whose amplitude grows with time. Note that the parameter p_2 does not appear directly in the PDE. So perturbations in p_2 will only alter the boundary control function whose domain of influence does not include the upper-left corner of the state in the (x, t) picture. Therefore, the sensitivity $\frac{dy}{dp_2}(p_0)$ equals zero in the corner mentioned.

The control sensitivities in figure 12 feature very fine structures. While they are in fact smooth functions except at the junction points of the control constraints, they appear jagged due to scaling reasons.

The discussion of objective sensitivities can be carried out like in the previous example.

8 Conclusion

We have proposed a method to compute parametric sensitivities of perturbed optimal control problems for time-dependent PDEs with control and state constraints. Under convergence assumptions to the presumed unique continuous solution, the original control problem has been discretized at various stages to finally yield a perturbed NLP problem for the discrete control variables. For this finite-dimensional problem, we are able to compute the sensitivity differentials of the optimal solution with respect to the perturbation parameters. These sensitivities represent an approximation of the corresponding continuous variables of the original PDE optimal control problem.

The examples show that this technique can be used for different types of PDEs and demonstrates its general ability to calculate accurate solutions in a robust way.

The nominal solution and its sensitivities can be used for real-time optimal control strategies, cf. (Refs. 2–5) for ODE optimal control problems. These ideas will be adapted to real-time PDE optimal control strategies in an upcoming paper.

References

1. H.J. Pesch. Real-time computation of feedback controls for constrained optimal control problems, Part 2: A correction method based on multiple shooting. *Optimal Control Applications & Methods* **10**, p. 147–171, 1989.
2. C. Büskens and H. Maurer. Sensitivity analysis and real-time control of nonlinear optimal control systems via nonlinear programming methods. In W.H. Schmidt,

- K. Heier, L. Bittner, and R. Bulirsch, editors, *Variational Calculus, Optimal Control and Applications*, volume 124 of *International Series of Numerical Mathematics*, p. 185–196. Birkhäuser Verlag, Basel, 1998.
3. C. Büskens and H. Maurer. Real-time control of robots with initial value perturbations via nonlinear programming methods. *Optimization* **47**, p. 383–405, 2000.
 4. C. Büskens and H. Maurer. SQP-methods for solving optimal control problems with control and state constraints: adjoint variables, sensitivity analysis and real-time control. *Journal of Computational and Applied Mathematics* **120**, p. 85–108, 2000.
 5. C. Büskens and H. Maurer. Nonlinear programming methods for real-time control of an industrial robot. *Journal of Optimization Theory and Applications* **107**, p. 505–527, 2000.
 6. S.F. Kupfer and E.W. Sachs. Numerical solution of a nonlinear parabolic control problem by a reduced SQP method. *Computational Optimisation and Applications* **1**, p. 113–135, 1992.
 7. M. Heinkenschloss. The numerical solution of a control problem governed by a phase field model. *Optimization Methods and Software* **7**, p. 211–263, 1997.
 8. M. Heinkenschloss and E.W. Sachs. Numerical solution of a constrained control problem for a phase field model. *Control and Estimation of Distributed Parameter Systems, Int. Ser. Num. Math* **118**, p. 171–188, 1994.
 9. F. Tröltzsch. An SQP method for optimal control of a nonlinear heat equation. *Control and Cybernetics* **23**, p. 267–288, 1994.
 10. H. Goldberg and F. Tröltzsch. On a Lagrange–Newton method for a nonlinear parabolic boundary control problem. *Optimization Methods & Software* **8**, p. 225–247, 1998.
 11. E. Casas. Boundary control with pointwise state constraints. *SIAM Journal on Control Optimization* **31**, p. 993–1006, 1993.
 12. E. Casas, F. Tröltzsch, and A. Unger. Second order sufficient optimality conditions for a class of elliptic control problems. *Journal for Analysis and its Applications* **15**, p. 687–707, 1996.
 13. E. Casas, F. Tröltzsch, and A. Unger. Second order sufficient optimality conditions for some state–constrained control problems of semilinear elliptic equations. *SIAM Journal on Control and Optimization* **38**, p. 1369–1391, 2000.
 14. K. Ito and K. Kunisch. Augmented Lagrangian–SQP methods for nonlinear optimal control problems of tracking type. *SIAM Journal on Control Optimization* **34**, p. 874–891, 1996.

15. K. Ito and K. Kunisch. Augmented Lagrangian–SQP methods in Hilbert spaces and application to control in the coefficients problem. *SIAM Journal on Optimization* **6**, p. 96–125, 1996.
16. K. Kunisch and S. Volkwein. Augmented Lagrangian–SQP techniques and their approximations. *Contemporary Mathematics* **209**, p. 147–159, 1997.
17. H. Maurer and H.D. Mittelmann. Optimization techniques for solving elliptic control problems with control and state constraints. Part 1: Boundary control. *Computational Optimization and Applications* **16**, p. 29–55, 2000.
18. H. Maurer and H.D. Mittelmann. Optimization techniques for solving elliptic control problems with control and state constraints. Part 2: Distributed control. *Computational Optimization and Applications* **18**, p. 141–160, 2001.
19. A.L. Dontchev and W.W. Hager Lipschitzian stability in nonlinear control and optimization *SIAM Journal on Control and Optimization* **31**, p. 569–603, 1993.
20. K. Malanowski, C. Büskens and H. Maurer Convergence of approximations to nonlinear optimal control problems In A.V. Fiacco, editor, *Mathematical Programming with Data Perturbations*, Lecture Notes in Pure and Applied Mathematics, Volume 195, p. 253–284, Marcel Dekker Inc., 1998.
21. K. Malanowski. Stability and sensitivity of solutions to nonlinear optimal control problems. *Applied Mathematics and Optimization* **32**, p. 111–141, 1995.
22. K. Malanowski and H. Maurer. Sensitivity analysis for parametric control problems with control–state constraints. *Computational Optimization and Applications* **5**, p. 253–283, 1996.
23. H. Maurer, H.J. Pesch. Solution differentiability for parametric nonlinear control problems. *SIAM Journal on Control and Optimization* **32**, p. 1542–1554, 1994.
24. C. Büskens. *Optimierungsmethoden und Sensitivitätsanalyse für optimale Steuerprozesse mit Steuer- und Zustands- Beschränkungen*. Dissertation, Institut für Numerische Mathematik, Universität Münster, Münster, Germany, 1998.
25. W. Alt Stability of solutions and the Lagrange-Newton method for nonlinear optimization and optimal control problems. Habilitationsschrift, Universität Bayreuth, 1990.
26. J.F. Bonnans and A. Shapiro. *Perturbation Analysis of Optimization Problems* Springer Series in Operations Research, Springer, Berlin, 2000.
27. K. Malanowski. Sensitivity analysis for parametric optimal control of semilinear parabolic equations. *Journal of Convex Analysis* **9**, p. 543–561, 2002.
28. K. Malanowski and F. Tröltzsch. Lipschitz stability of solutions to parametric optimal control problems for parabolic equations. *Zeitschrift für Analysis und ihre Anwendungen* **18**, p. 469–489, 1999.

29. H.D. Mittelmann. Verification of second-order sufficient optimality conditions for semilinear and parabolic control problems. *Computational Optimization and Applications* **20**, p. 93–110, 2001.
30. J.-P. Raymond and F. Tröltzsch. Second order sufficient optimality conditions for nonlinear parabolic control problems with state constraints. *Discrete and Continuous Dynamical Systems* **6**, p. 431–450, 2000.
31. C. Büskens. Direkte Optimierungsmethoden zur numerischen Berechnung optimaler Steuerungen. Diploma thesis, Institut für Numerische Mathematik, Universität Münster, Münster, Germany, 1993.
32. A. Barclay, Ph.E. Gill, J.B. Rosen. SQP methods and their application to numerical optimal control. In W.H. Schmidt, K. Heier, L. Bittner, R. Bulirsch, editors, *Variational Calculus, Optimal Control and Applications*, p. 207–222, Birkhäuser Verlag, Basel, 1998.
33. K. Malanowski and H. Maurer. Sensitivity analysis for state constrained control problems. *Discrete and Continuous Dynamical Systems* **4**, p. 241–272, 1998.
34. A.V. Fiacco. Sensitivity analysis for nonlinear programming using penalty methods. *Mathematical Programming* **10**, p. 287–311, 1976.
35. A.V. Fiacco. *Introduction to Sensitivity and Stability Analysis in Nonlinear Programming*, volume 165 of *Mathematics in Science and Engineering*. Academic Press, New York, 1983.
36. S.M. Robinson. Perturbed Kuhn–Tucker points and rate of convergence for a class of nonlinear programming algorithms. *Mathematical Programming* **7**, p. 1–16, 1974.
37. R. Griesse and A. Walther. Parametric sensitivity analysis for optimal control problems using automatic differentiation. *Optimal Control Applications and Methods* **24**, p. 297–314, 2003.
38. S. Volkwein. *Mesh-Independence of an Augmented Lagrangian-SQP Method in Hilbert Spaces and Control Problems for the Burgers Equation*. Dissertation, Fachbereich Mathematik, Technische Universität Berlin, 1997.
39. S. Volkwein. Lagrange sqp techniques for the control constrained optimal boundary control for the burgers equation. *Computational Optimization and Applications* **26**, p. 253–284, 2003.
40. R. Griesse. *Parametric Sensitivity Analysis for Control-Constrained Optimal Control Problems Governed by Systems of Parabolic Partial Differential Equations*. Dissertation, Fakultät für Mathematik und Physik, Universität Bayreuth, 2003.
41. F. Tröltzsch and S. Volkwein. The SQP method for control constrained optimal control of the Burgers equation. *ESAIM: Control, Optimisation and Calculus of Variations* **6**, p. 649–674, 2001.