Proceedings of the 9th International Symposium on
Dynamics and Control of Process Systems (DYCOPS 2010),
Leuven, Belgium, July 5-7, 2010
Mayuresh Kothare, Moses Tade, Alain Vande Wouwer, Ilse Smets (Eds.)

Combining Direct and Indirect Methods for Optimal Control – a Case Study ⋆

Ralf Hannemann [*,**] **Wolfgang Marquardt** [*,1]

[*] *Aachener Verfahrenstechnik – Process Systems Engineering, RWTH Aachen University, Turmstr. 46, 52064 Aachen, Germany*
[**] *German Research School for Simulation Sciences GmbH, 52425 Jülich, Germany*

Abstract: Adaptive control vector parameterization for the solution of optimal control problems approximates the original infinite-dimensional optimal control problem by a set of finite-dimensional nonlinear programs (NLPs) whose control grids are iteratively refined. The refinement is stopped by a heuristic stopping criterion. The Hessians of the Lagrangian of these NLPs can be efficiently computed by the technique of composite adjoints as recently proposed by the authors. By means of a case study, namely the optimal control of the Williams-Otto semi-batch reactor, we show how to interpret composite adjoints as estimates for the continuous adjoints referred to by Pontryagin's Minimum Principle. Thus, these composite adjoints can be used to (i) construct a novel and mathematical sound stopping criterion for the iterative refinement of the control grid and to (ii) setup an indirect multiple shooting method the solution of which verifies and improves the approximate solution to the exact one.

Keywords: optimal control, direct single shooting, indirect multiple shooting, multipoint boundary value problem, Pontryagin, adjoints

1. INTRODUCTION

Optimal control problems arise in many engineering applications. Originally dealing with problems in aerospace engineering, nowadays optimal control and especially nonlinear model predictive control (NMPC) are a current research topics in process control. The authors' research group develops and maintains DyOS (**Dy**namic **O**ptimization **S**oftware) for the numerical solution of large-scale optimal control problems in real time. DyOS has been successfully applied to a number of challenging open-loop and closed-loop optimal control problems (Hartwich and Marquardt, 2010; Würth et al., 2009). One strength of DyOS lies in its adaptive control vector parameterization (Schlegel et al., 2005; Schlegel and Marquardt, 2006), a variant of the single shooting approach (Sargent and Sullivan, 1978).

DyOS starts with a coarse parameterization of the control vector and iteratively eliminates or inserts points into the control grid to find a problem-tailored parameterization. Currently, the refinement is stopped by monitoring the change in the objective function: if the objective function value does not improve in some subsequent grid refinement iterations, the algorithm stops (Schlegel et al., 2005). Though this approach works well in practice it suffers from two drawbacks. Firstly, the algorithm performs at least one "redundant" iteration, since the objective function does not improve significantly from the penultimate to the last iteration. Secondly, we actually do not know to what extent the solution satisfies the necessary conditions of optimality in the sense of Pontryagin's Minimum Principle (Pontryagin et al., 1962), i.e. we do not know how close or far away we are from the true solution. We are searching for methodologies to resolve these issues by extending the direct solution approach by some elements of an indirect approach.

This contribution shows by means of a case study how information of the single shooting NLP, especially composite adjoints (Hannemann and Marquardt, 2010), can be interpreted as continuous adjoints in the sense of Pontryagin's Minimum Principle and thus can be used to decide whether the control grid is sufficiently refined. Further, we utilize them to initialize an indirect method to detect the distance from the true solution. We discuss how the information about the continuous adjoints extracted from the NLP can be incorporated in a new adaptive control vector parameterization algorithm in the future.

This paper is organized as follows. In Section 2 the class of optimal control problems under consideration is presented and in Section 3, we recall the necessary conditions of optimality. The discretization of optimal control problems by means of single shooting leads to nonlinear programs which is presented in Section 4. The interpretation of NLP information, especially the meaning of composite adjoints, is discussed in Section 5. A case study is investigated in Section 6. The results of the case studies give rise to the modified adaptive control vector parameterization presented in Section 7. Finally, Section

⋆ This work was supported by the German Research Foundation (DFG) under grant MA 1188/28-1.
[1] Corresponding author: wolfgang.marquardt@avt.rwth-aachen.de

8 presents our conclusion and gives an outlook on future research directions.

2. PROBLEM FORMULATION

We consider exemplarily an optimal control problem with one pure state constraint and several control constraints where the controls appear linearly, motivated by the optimal control of the Williams-Otto semi-batch reactor as introduced by Forbes (1994), given by

$$\min_{u(t)} \quad \Phi(x(t_f)) \tag{1}$$

$$\text{s. t.} \quad \dot{x}(t) = a(x(t)) + b(x(t))\, u(t), \tag{2}$$

$$x(t_0) = x_0 \in \mathbb{R}^{n_x}\,, \tag{3}$$

$$s(x(t)) \leq 0 \ \ \forall\, t \in [t_0, t_f], \tag{4}$$

$$h(x(t_f)) = 0\,, \tag{5}$$

$$u(t) \in U := [u_{\min}, u_{\max}] \subset \mathbb{R}^{n_u}. \tag{6}$$

Here, $\Phi, s, h : \mathbb{R}^{n_x} \to \mathbb{R}$, $a : \mathbb{R}^{n_x} \to \mathbb{R}^{n_x}$, $b : \mathbb{R}^{n_x} \to \mathbb{R}^{n_x \times n_u}$ are smooth functions. The independent variable t is time, $x(t) \in \mathbb{R}^{n_x}$ is the time-dependent state vector and $u(t) \in \mathbb{R}^{n_u}$ is the time-dependent control vector. We assume that the optimal control is a piecewise continuous function with a finite number of discontinuities. The state vector $x(t)$ is continuous and piecewise continuous differentiable with respect to time. Eq. (2) holds except for the points of discontinuities of u. The detailed presentation of the formulation of the Williams-Otto semi-batch reactor benchmark optimization problem considered later in a case study can be found in Hannemann and Marquardt (2010).

An extension to different classes of optimal control problems (controls appearing nonlinearly, several pure state constraints, mixed control-state constraints, etc.) is possible whenever suitable optimality conditions are available. Hartl et al. (1995) provide a survey of such optimality conditions. However, the authors are aware of the fact that not all possible formulations are covered and that we are relying on the assumption that the optimal control profiles exhibit only a finite number of discontinuities, which may be violated in some cases.

3. PONTRYAGIN'S MINIMUM PRINCIPLE

There are a couple of proven or informal minimum principles for optimal control problems with state constraints (Hartl et al., 1995). We restrict our investigation on the so-called direct adjoining approach. For the discussion of the minimum principle we need some notational prerequisites which are introduced next.

Let $\tau_1 < \tau_2$ be real numbers. A subinterval $(\tau_1, \tau_2) \subset [t_0, t_f]$ is called an *interior interval* of the trajectory $x(\cdot)$ if $s(x(t)) < 0$ for all $t \in (\tau_1, \tau_2)$. In contrast, a subinterval $[\tau_1, \tau_2] \subset [t_0, t_f]$ is called a *boundary interval* if $s(x(t)) = 0$ for all $t \in [\tau_1, \tau_2]$.

Let $t_0 < \tau_1 < \tau_2 < t_f$ and $[\tau_1, \tau_2]$ be a maximal boundary interval. Then, τ_1 is called *entry time* and τ_2 is called *exit time*, taken together they are called *junction times*. We will not discuss the case of a *boundary point* τ (Jacobsen et al., 1971) where $s(x(\tau)) = 0$ but $x(\cdot)$ is in the interior just before and just after τ since this case is not relevant for the Williams-Otto reactor.

To compute the optimal control within a boundary interval $[\tau_1, \tau_2]$ one has to introduce the concept of order of the state constraint. We assume that the additional equation $s(x(t)) = 0$ uniquely determines one control variable, say u_1, on $[\tau_1, \tau_2]$. Since $s(x(t)) = 0$ does not explicitly depend on u_1 we have to differentiate this equation recursively with respect to t. We define

$$s^0 := s,$$

$$s^{i+1} := \frac{\partial s^i}{\partial x}\left(a(\cdot) + b(\cdot)u\right),\ i = 0, 1, \ldots.$$

We say $s(x(t)) \leq 0$ has the order $p \geq 1$, iff

$$\frac{\partial s^{p-1}}{\partial u_1} \equiv 0 \text{ but } \frac{\partial s^p}{\partial u_1} \neq 0.$$

If the order of the state constraint is p, then the equation $s^p(x, u) = 0$ can be used to compute the control on a boundary interval. The order of the state constraint of the Williams-Otto reactor is $p = 1$.

To formulate the Minimum Principle of Pontryagin or, in other words, the necessary conditions of optimality we introduce the extended Hamiltonian

$$H(x, \lambda, \eta, u) := \lambda^T \left(a(x) + b(x)\, u\right) + \eta\, s(x), \tag{7}$$

where $x, \lambda \in \mathbb{R}^{n_x}$, $u \in \mathbb{R}^{n_u}$, $\eta \in \mathbb{R}$. Let $x(t), u(t)$ be an optimal solution of eqs. (1) – (6). Further, let the equation $s^p(x, u) = 0$ be solvable for $u_1 = u_1(x)$, and let $u_1(t) \in (u_{\min 1}, u_{\max 1})$ for all boundary intervals $[t_1, t_2]$. Then, there exists

(a) a multiplier $\rho \in R$,
(b) a piecewise continuous and piecewise continuously adjoint function $\lambda : [t_0, t_f] \to \mathbb{R}^{n_x}$,
(c) a piecewise continuous multiplier function $\eta : [t_0, t_f] \to \mathbb{R}$,
(d) multipliers $0 \leq \nu(t_i) \in \mathbb{R}$ for every junction time t_i,

such that the following statements hold:

(1) $u(t) = \arg\min_{u \in U} H(x(t), \lambda(t), \eta(t), u)$ for all points of continuity $t \in [t_0, t_f]$ of $u(t)$,
(2) $\dot{\lambda}(t)^T = -H_x(x(t), \lambda(t), \eta(t), u(t))$
$\lambda(t_f)^T = \Phi_x(x(t_f)) + \rho\, h_x(x(t_f))$,
(3) $\eta(t) \geq 0$ and $\eta(t)\, s(x(t)) = 0$,
(4) for each junction time t_i the following junction conditions hold:
$\lambda(\tau_i^+)^T = \lambda(\tau_i^-)^T + \nu(\tau_i)\, s_x(x(\tau_i))$, where $\lambda(\tau_i^+) = \lim_{\varepsilon \downarrow 0} \lambda(\tau_i + \varepsilon)$, $\lambda(\tau_i^-) = \lim_{\varepsilon \downarrow 0} \lambda(\tau_i - \varepsilon)$.

In our setup, the Hamiltonian is linear in u. Hence, if $H_{u_i} \neq 0$ holds for an optimal solution, the control u_i stays at its lower or upper bound. We introduce the switching functions

$$\sigma_i(x, \lambda) = H_{u_i}(x, \lambda, \eta, u),\ i = 1, \ldots, n_u. \tag{8}$$

For an optimal solution (x^*, u^*) with corresponding adjoint functions λ^*, we simply write for convenience

$$\sigma_i(t) := \sigma_i(x^*(t), \lambda^*(t)),\ i = 1, \ldots, n_u, \tag{9}$$

such that the optimal control can be characterized by

$$\left\{ \begin{array}{ll} u_i^*(t) = u_{\max i}, & \text{if } \sigma_i(t) < 0 \\ u_i^*(t) \in [u_{\min i}, u_{\max i}], & \text{if } \sigma_i(t) = 0 \\ u_i^*(t) = u_{\min i}, & \text{if } \sigma_i(t) > 0 \end{array} \right\}. \tag{10}$$

Because of (10) we have

$$\sigma_i \equiv 0 \quad \text{on } [\tau_1, \tau_2], \tag{11}$$

for a boundary interval $[\tau_1, \tau_2]$ of the control u_i.

4. DIRECT SINGLE SHOOTING

So far, we have characterized the optimal solution of the continuous (infinite-dimensional) optimal control problem. At this point we state how to approximate optimal solutions by means of direct single shooting.

The basic idea is to substitute the control vector $u(t)$ by an approximation $\tilde{u}(t,p)$ employing parameters p_{ij} and basis functions $\phi_{ij}(t)$, i.e.

$$\tilde{u}_i(t,p) := \sum_{j=1}^{P_i} p_{ij}\,\phi_{ij}(t), \quad p_{ij} \in \mathbb{R},\ i = 1,\dots,n_u\,. \tag{12}$$

Typically, the functions $\phi_{ij}(t)$ are constant, linear or cubic B-splines. The vector $p = (p_{11},\dots,p_{n_u P_{n_u}})^T \in \mathbb{R}^{n_p}$ concatenates of all degrees of freedom. The state vector depends on the parameter p and can be computed by solving the initial value problem

$$\dot{\tilde{x}}(t,p) = f(\tilde{x}(t,p), \tilde{u}(t,p)), \tag{13}$$

$$\tilde{x}(0,p) = x_0, \tag{14}$$

for example by Runge-Kutta methods.

The path constraint (4) is relaxed by defining a grid $t_0 < t_1 < t_2 < \cdots < t_N = t_f$ to result in

$$s(\tilde{x}(t_k,p)) \le 0,\ k = 0,\dots N.$$

The infinite-dimensional optimal control problem is approximated by the finite-dimensional nonlinear program

$$\begin{aligned} \min_{p} \quad & \Phi(\tilde{x}(t_N,p)) \\ \text{s. t.} \quad & s(\tilde{x}(t_k,p)) \le 0, \quad k = 0,\dots,N \\ & h(\tilde{x}(t_N,p)) = 0\,. \end{aligned} \tag{OC-NLP}$$

The Lagrangian of (OC-NLP) can be stated as

$$\begin{aligned} \mathcal{L}(p,\tilde{\eta},\tilde{\rho}) =& \Phi(\tilde{x}(t_N,p)) + \left(\sum_{k=0}^{N} \tilde{\eta}_k\, s(\tilde{x}(t_k,p))\right) \\ & + \tilde{\rho}\, h(\tilde{x}(t_N,p))) \end{aligned} \tag{15}$$

with Lagrange multipliers $\tilde{\eta}_k \in \mathbb{R}, k = 0,\dots,N$ and $\tilde{\rho} \in \mathbb{R}$.

The well-known Karush-Kuhn-Tucker necessary conditions characterize the optimal solution of (OC-NLP). Let p^* be an optimal solution of (OC-NLP). Under mild assumptions, there exist unique Lagrange multipliers $\tilde{\eta}_k \in \mathbb{R}, k = 0,\dots,N$ and $\tilde{\rho} \in \mathbb{R}$, such that

(1) $\mathcal{L}_p(p^*,\tilde{\eta},\tilde{\rho}) = 0$,
(2) The constraints of (OC-NLP) are satisfied, and
(3) $\tilde{\eta}_i \ge 0,\ \tilde{\eta}_i\, s(x(t_i,p^*)) = 0,\ i = 0,\dots,N$.

5. COMPOSITE ADJOINTS

Composite adjoints refer to a recently introduced technique to efficiently compute the Hessian of the Lagrangian in eq. (15). They are an extension of the first- and second-order adjoint sensitivity analysis for multipoint-evaluated ODE-embedded functionals. For a detailed discussion of composite adjoints we refer to Hannemann and Marquardt (2010).

According to Büskens and Maurer (2000, p. 92, eq. (30)), the first-order composite adjoints $\tilde{\lambda}(t,p)$ are a good approximation of the adjoint function from Pontryagin's Minimum Principle [2]:

$$\tilde{\lambda}(t,p^*) = \frac{\partial \mathcal{L}}{\partial x(t,p)} \approx \lambda(t). \tag{16}$$

Furthermore, we have the relation

$$\tilde{\rho} \approx \rho. \tag{17}$$

There exists also a relation between the multipliers $\tilde{\eta}_i$ and the continuous multiplier function $\eta(t)$ and the junction time multipliers $\nu(\tau_i)$. Let $[\tau_1,\tau_2]$ be a maximal boundary interval with entry time τ_1 and exit time τ_2, $K < L$ integers, and $t_K < t_{K+1} < \cdots < t_L$ a finite sequence of times with $\tau_1 < t_{K-1} < t_{L+1} < \tau_2$. Then, we can state the conjecture that

$$\int_{t_K}^{t_L} \eta(t)\,dt \approx \sum_{i=K}^{L} \tilde{\eta}_i - \frac{\tilde{\eta}_K + \tilde{\eta}_L}{2}, \tag{18}$$

and additionally, if $t_K \approx \tau_1, t_L \approx \tau_2, t_{K-1} < \tau_1, \tau_2 < t_{L+1}$:

$$\int_{t_K}^{t_L} \eta(t)\,dt + \nu(\tau_1) + \nu(\tau_2) \approx \sum_{i=K}^{L} \tilde{\eta}_i. \tag{19}$$

This (not yet proved) conjecture is a generalization of the results reported by Büskens and Maurer (2000) and has been confirmed by numerical experiments. The relations (16) – (19) can be used to correlate information from (OC-NLP) with the statements of Pontryagin's Minimum Principle. We use them to set up and initialize the corresponding multipoint boundary problem to further improve the approximate solution resulting from (OC-NLP) to finally obtain an exact solution of the optimal control problem.

6. WILLIAMS-OTTO SEMI-BATCH REACTOR

The Williams-Otto semi-batch reactor, as introduced by Forbes (1994), is optimized using

(1) direct single shooting applying an equidistant piecewise constant parameterization with
    (a) a coarse grid (20 parameters for each control),
    (b) a fine grid (200 parameters for each control), and
(2) indirect multiple shooting in order to obtain a highly accurate numerical solution.

In the reactor, the reactions $A + B \longrightarrow C$, $C + B \longrightarrow P + E$ and $P + C \longrightarrow G$ take place. Reactant A is already in the reactor at initial time, whereas reactant B is fed continuously into the reactor during operation. The products P and E as well as the side-product G are formed. The heat generated through the exothermic reaction is removed by a cooling jacket, which is controlled by manipulating the cooling water temperature T_W. At the end of the batch, the conversion to the desired products P and E should be maximized. We have $n_x = 9$ states and $n_u = 2$ controls. The manipulated control variables

[2] In the past, there were different approaches to utilize NLP information of direct methods for estimates of the continuous adjoints. Von Stryk and Bulirsch (1992) use a full discretization approach based on collocation and provide estimates for the adjoints variables. Grimm and Markl (1997) provide estimates for the adjoints based on direct multiple shooting, as introduced by Bock and Plitt (1984). Büskens and Maurer (2000) implement the single shooting technique to provide an a-posteriori estimate of the adjoints.

of this process are $u_1(t) = T_W(t)$ and the flow rate of B $u_2(t) = F_{B,in}(t)$. The batch time is 1000 seconds.

The economic objective is to maximize the yield of the main products at the end of the batch. The dynamic optimization problem is of the form (1) – (6) where the pure state constraint (4) refers to the reactor temperature.

We present the results of the numerical computations by direct single and indirect multiple shooting. The direct single shooting method has been implemented using a fixed-stepsize fourth-order Runge-Kutta method with equidistant piecewise constant parameterization of the controls. Indirect multiple shooting is performed by the Fortran routine BNDSCO of Oberle and Grimm (1989).

Though in practice, we started with the direct method, we first discuss the exact solution computed by the indirect method. The exact control u_1 and its switching function σ_1 are sketched in Figure 3. Note that the control stays continuous when entering the boundary arc at τ_2. This behavior does not contradict the theory of necessary conditions but is not observed often in practice. The control u_2 has a bang-bang structure (not shown). All in all we have six switching points $\tau_1, \ldots, \tau_6$. Setting $\tau_0 = t_0 = 0$ and $\tau_7 = t_f = 1000$ we can identify seven intervals $I_k = [\tau_{k-1}, \tau_k], k = 1, \ldots, 7$. The structure of the optimal solution is shown in Table 1. The bold '**l**' and '**u**' mark controls which are at its **l**ower and **u**pper bound, respectively. The bold **b** is reserved for **b**oundary intervals, where the state constraint $s(x(t)) \leq 0$ is active. Here, I_3 is the only boundary interval. The switching times itself are found in Table 2. As seen in Fig. 3, the control u_1 is continuous at the entry time of the boundary interval I_3 and discontinuous at its exit time. According to Maurer and Heidemann (1975), this behavior results in a nonzero multiplier $\nu(\tau_2)$ in the entry and a zero multiplier $\nu(\tau_3)$ in the exit time, as it is confirmed by the numerical values given in Table 3.

The approximated control $\tilde{u}_1$ computed by means of the coarse control grid is presented in Fig. 4. In accordance with eqs. (8) and (9), we introduce the discrete switching functions

$$\tilde{\sigma}_i(t) = \sigma_i(\tilde{x}(t,p), \tilde{\lambda}(t,p)), \; i = 1, \ldots, n_u, \qquad (20)$$

which will serve as measures for the solution quality of the control grid. Fig. 4 also presents $\tilde{\sigma}_1$ for the coarse grid. If we analyze the trajectories of $\tilde{\sigma}_1$ and $\tilde{u}_1$ with respect to eqs. (10) and (11), we discover inconsistencies: on the interval $[\tau_2, \tau_3]$ the control $\tilde{u}_1$ seems to be on a boundary interval but $\tilde{\sigma}_1$ does not vanish. Hence the chosen control grid is not suited to reflect the true solution.

In contrast, visual inspection of Fig. 5 reveals that $\tilde{u}_1$ of the fine control grid is a good approximation to the "true" solution and that eqs. (8) and (9) are at least satisfied approximately. In the following, we provide some further investigations of the relations between the exact and the approximated solution on the fine control grid. These relations were used to setup the right MPBVP for the computation of the exact solution by means of indirect multiple shooting.

Figs. 1 and 2 illustrate the relation of the discrete and continuous multipliers $\tilde{\eta}$ and $\eta(t)$. Though the correspondence of the multipliers is not visible at first sight because of the different scalings and numerical artifacts, relation (19) is verified in Table 4. Note that the "jump" of $\tilde{\eta}_k$ at $k = 58$ in Fig. 1 approximates the multiplier $\nu(\tau_2)$ given in Table 3 and is used as an initial guess for the indirect method. Table 6 presents the multipliers for the endpoint condition (5), which are in accordance with eq. (17). The initial values of the composite and exact adjoints are given in Table 5. They confirm eq. (16) and show that first-order composite adjoints constitute a good approximation of the adjoints of the continuous problem.

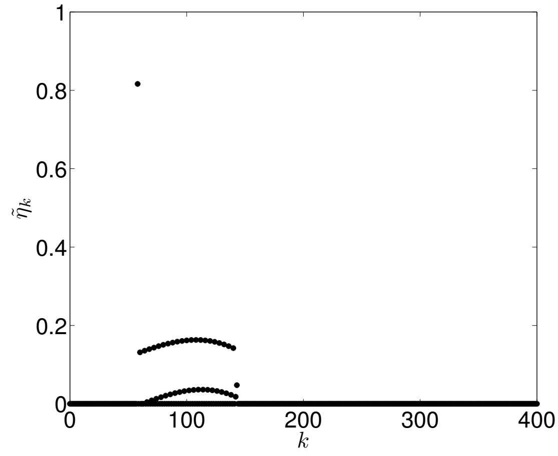

Fig. 1. Discrete multipliers $\tilde{\eta}_k$ (fine grid), $k = 0, 1, \ldots, 400$

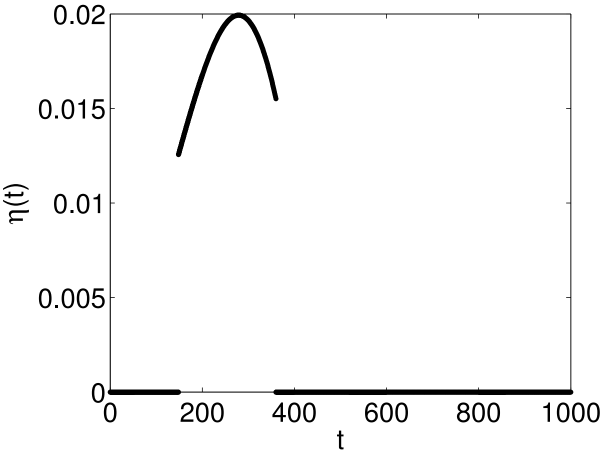

Fig. 2. Exact multiplier function $\eta(t)$

Table 1. Structure of the optimal control

	I_1	I_2	I_3	I_4	I_5	I_6	I_7
$u_1(t)$	**l**	**u**	**b**	**u**	**u**	**l**	**u**
$u_2(t)$	**u**	**u**	**u**	**u**	**l**	**l**	**l**
l: $u_i = u_{\min i}$, **u**: $u_i = u_{\max i}$, **b**: $u_1 = u_1(x)$							

Table 2. Switching times

τ_1	0.52337586E+02
τ_2	0.14825782E+03
τ_3	0.35986598E+03
τ_4	0.51867220E+03
τ_5	0.53868488E+03
τ_6	0.85879617E+03

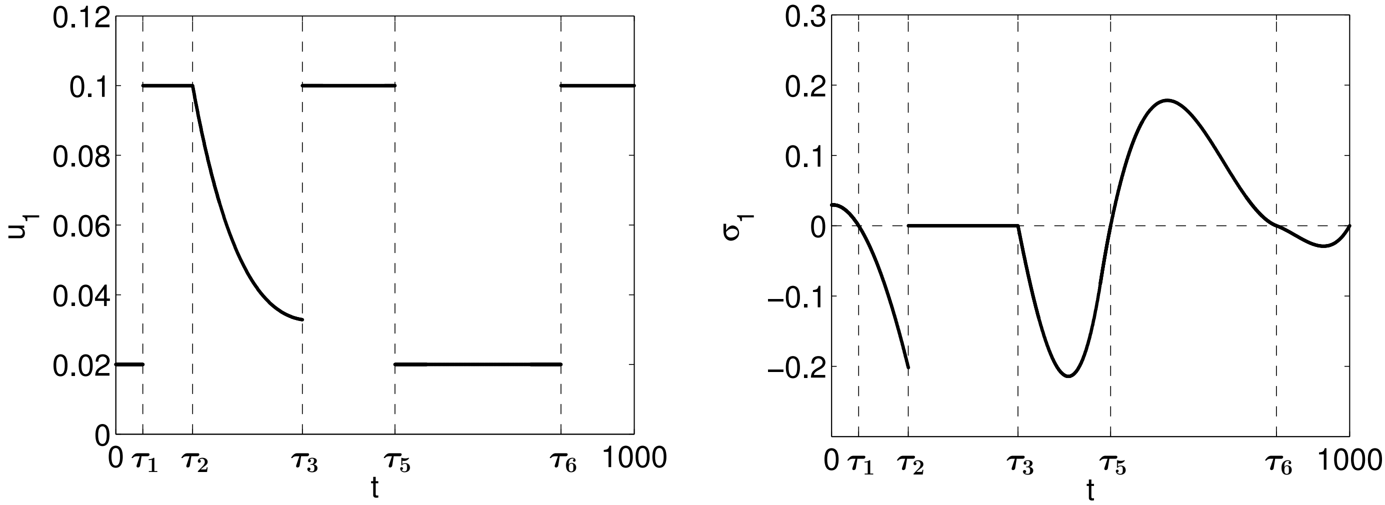

Fig. 3. Exact optimal control u_1 (left) and its switching function (right) computed by the indirect method

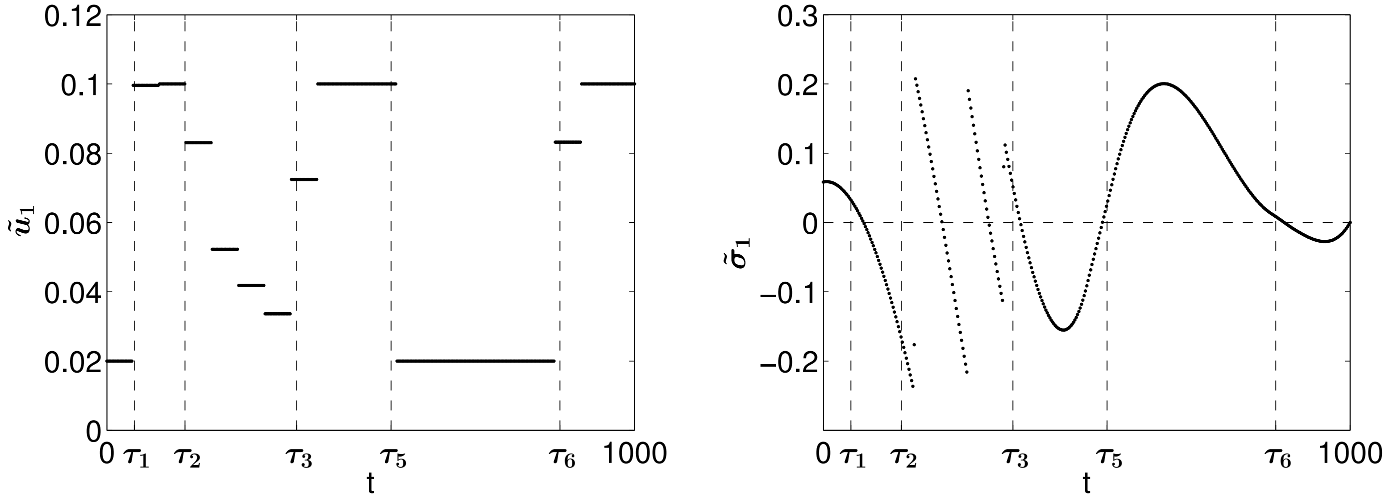

Fig. 4. Piecewise constant approximation of control u_1 with 20 parameters (left) and its discrete switching function (right)

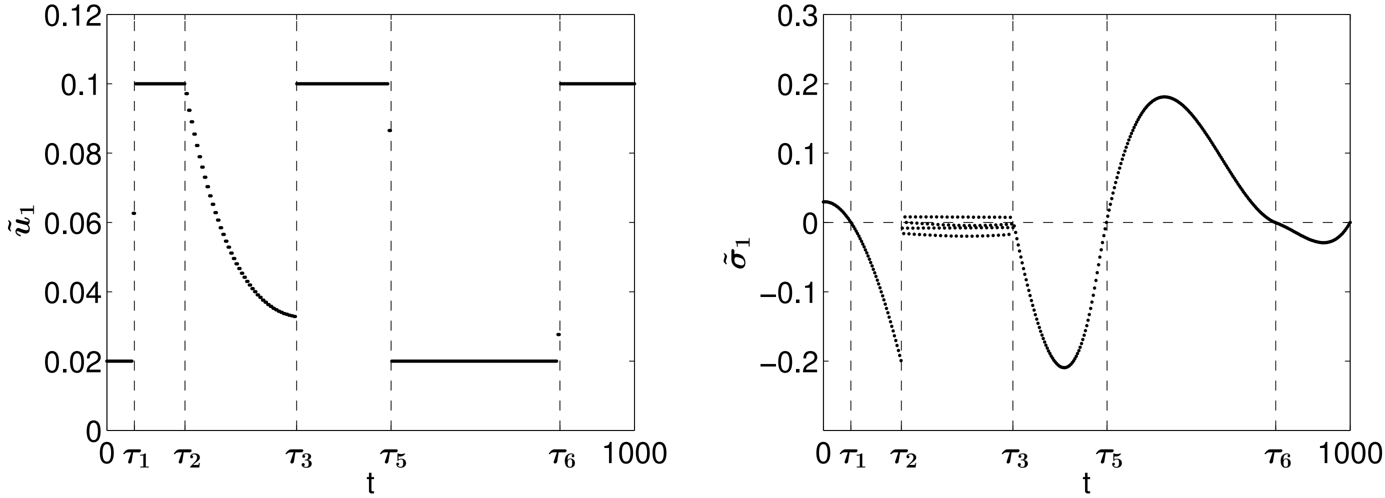

Fig. 5. Piecewise constant approximation of control u_1 with 200 parameters (left) and its discrete switching function (right)

Table 3. Multiplier for junction times

$\nu(\tau_2)$	0.82948591E+00
$\nu(\tau_3)$	0

Table 4. Multiplier functions

indirect	direct (fine grid)
$\int_{\tau_2}^{\tau_3} \eta(t)\, dt + \nu(\tau_2) + \nu(\tau_3)$ $= 4.57382$	$\sum_{i=K}^{L} \tilde{\eta}_i$ $= 4.58857$

Table 5. Composite (direct, fine grid) and exact (indirect) adjoints

i	indirect $\lambda_i(t_0)$	direct $\tilde{\lambda}_i(t_0)$
1	-0.468420E+03	-0.467945E+03
2	-0.313515E+04	-0.313509E+04
3	-0.172955E+04	-0.172936E+04
4	0.516053E+04	0.516086E+04
5	0.164392E-29	0.000000E+00
6	-0.258480E-28	0.000000E+00
7	0.121051E+00	0.121705E+00
8	0.119475E+04	0.119625E+04
9	0.100000E+01	0.100000E+01

Table 6. Endpoint constraint multipliers

indirect ρ	direct $\tilde{\rho}$ (fine grid)
0.135189E+04	0.135325E+04

7. NOVEL STOPPING CRITERION

The results of the following section give rise to a modified adaptive control vector parameterization which is sketched in Table 7. By now the decision in step 5, whether the solution quality is sufficient, is taken manually but our group works on the automatization of the stopping criterion.

Table 7. Modified adaptive control vector parameterization

(1) Create an initial control grid.
(2) Solve (OC-NLP) on the current control grid.
(3) Check based on the discrete switching functions in eq. (20) whether the relations (8),(9) are (nearly) fulfilled. If yes then goto 5. else goto 4.
(4) Apply the grid refinement algorithm according to Schlegel et al. (2005) and goto 2.
(5) STOP. Solution quality is sufficient.
(6) If desired setup the proper MPBVP and solve it for example by indirect multiple shooting to verify the approximate and to compute the exact solution.

8. CONCLUSIONS

We computed the optimal control of the Williams-Otto semi-batch reactor by means of single shooting. We showed that the dual information of the (OC-NLP) of direct single shooting, especially the information of the composite adjoints, corresponds one to one to the dual information of the continuous solution. We utilized this insight to propose a modified adaptive control vector parameterization with a new stopping criterion which further allows to set up the correct MPBVP to facilitate the computation of the exact solution.

REFERENCES

Bock, H. and Plitt, K. (1984). A multiple shooting algorithm for direct solution of optimal control problems. In *9th IFAC World Congress Budapest*, 242–247. Pergamon Press.

Büskens, C. and Maurer, H. (2000). 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(1), 85–108.

Forbes, J. (1994). *Model Structure and Adjustable Parameter Selection for Operations Optimizations.* Ph.D. thesis, McMaster University, Hamilton, Canada.

Grimm, W. and Markl, A. (1997). Adjoint estimation from a direct multiple shooting method. *Journal of Optimization Theory and Applications*, 92(2), 263–283.

Hannemann, R. and Marquardt, W. (2010). Continuous and discrete composite adjoints for the hessian of the lagrangian in shooting algorithms for dynamic optimization. *SIAM Journal on Scientific Computing*, 31(6), 4675–4695.

Hartl, R., Sethi, S., and Vickson, R. (1995). A survey of the maximum principles for optimal control problems with state constraints. *SIAM Review*, 37(2), 181–218.

Hartwich, A. and Marquardt, W. (2010). Dynamic optimization of the load change of a large-scale chemical plant by adaptive single shooting. *Computers & Chemical Engineering*, In Press, Corrected Proof.

Jacobsen, D.H., Lele, M.M., and Speyer, J.L. (1971). New necessary conditions of optimality for control problems with state-variable inequality constraints. *Journal of Mathematical Analysis and Applications*, 35, 255–284.

Maurer, H. and Heidemann, U. (1975). Optimale Steuerprozesse mit Zustandsbeschränkungen. URL http://dx.doi.org/10.1007/BFb0079180.

Oberle, H.J. and Grimm, W. (1989). BNDSCO-a program for the numerical solution of optimal control problems. *Institute for Flight Systems Dynamics, DLR, Oberpfaffenhofen, Germany, Internal Report*, 22.

Pontryagin, L., Boltyanski, V., Gamkrelidze, R., and Misčenko, E. (1962). *The Mathematical Theory of Optimal Processes.* Wiley, New York.

Sargent, R. and Sullivan, G. (1978). The development of an efficient optimal control package. In J. Stoer (ed.), *Proceedings of the 8th IFIP Conference on Optimization Technique*, 158–168. Springer-Verlag.

Schlegel, M. and Marquardt, W. (2006). Detection and exploitation of the control switching structure in the solution of dynamic optimization problems. *Journal of Process Control*, 16(3), 275–290.

Schlegel, M., Stockmann, K., Binder, T., and Marquardt, W. (2005). Dynamic optimization using adaptive control vector parameterization. *Computers & Chemical Engineering*, 29(8).

Von Stryk, O. and Bulirsch, R. (1992). Direct and indirect methods for trajectory optimization. *Ann. Oper. Res.*, 37(1-4), 357–373.

Würth, L., Hartwich, A., Bausa, J., Salau, N., and Marquardt, W. (2009). Nichtlineare modellprädiktive Regelung eines Polymerisationsreaktors im Labor. In *Jahrestreffen der Fachgemeinschaft Prozess-, Apparate- und Anlagentechnik*. Weinheim.