Deutsche Forschungsgemeinschaft
Priority Programme 1962

A Switching Cost Aware Rounding Method for Relaxations of Mixed-Integer Optimal Control Problems

Felix Bestehorn, Christoph Hansknecht, Christian Kirches, Paul Manns

SPP 1962
Non-smooth and Complementarity-based Distributed Parameter Systems: Simulation and Hierarchical Optimization

Preprint Number SPP1962-124

received on October 15, 2019

Edited by
SPP1962 at Weierstrass Institute for Applied Analysis and Stochastics (WIAS)
Leibniz Institute in the Forschungsverbund Berlin e.V.
Mohrenstraße 39, 10117 Berlin, Germany
E-Mail: spp1962@wias-berlin.de

World Wide Web: http://spp1962.wias-berlin.de/

A switching cost aware rounding method for relaxations of mixed-integer optimal control problems

Felix Bestehorn
Christoph Hansknecht
Christian Kirches
Paul Manns
Institute for Mathematical Optimization
Technische Universität Carolo-Wilhelmina zu Braunschweig, Germany
{f.bestehorn,c.hansknecht,c.kirches,paul.manns}@tu-braunschweig.de

***Abstract*— This article investigates a class of Mixed-Integer Optimal Control Problems (MIOCPs) with switching costs. We introduce the problem class of Minimal-Switching-Cost Optimal Control Problems (MSCP) with an objective function that consists of two summands, a continuous term depending on the state vector and an encoding of the discrete switching costs. State vectors of Mixed-Integer Optimal Control problems can be approximated by means of sequences of roundings of appropriate relaxations, which often result in a switching cost blow-up. We reformulate the problem such that trading convergence of the state vector against increasing switching costs is possible, which then allows to conserve known convergence properties of previous approaches for Mixed-Integer Optimal Control approximations. To demonstrate the findings and applicability, we present validating numerical results and the trade-off capability of our approach for a benchmark problem.**

I. INTRODUCTION

Due to their wide range of practical applications Mixed-Integer Optimal Control Problems (MIOCPs, also known as switched or hybrid systems) have been gaining considerable attention in recent years. The scope of applications ranges from the shifting of gears in a car [4], [11], supply chain networks [12], chemical engineering [5], to renewable energy [2]. A survey has been provided by Zhu and Antsaklis [25].

In this article we are interested in investigating the control of dynamic processes associated with discrete switches and switching costs. Several frameworks to approach MIOCPs with discrete switches have been proposed in recent years, using e.g. outer convexification [8], time transformation [17] or mixed-integer programming (MIP) approaches [23]. In the aforementioned publications, costs for switching have not been included into the objective functional of the MIOCP. Kirches et al. [9] recently proposed a framework, which penalizes switches through an L^1 regularization term. In this article, we generalize switching costs by allowing different values for different states that are switched on and off as well as different costs for switching them off and on, thus allowing a penalization of switching costs in a more flexible way.

A. The model problem

In this article we will study the prototypical *Minimal-Switching-Cost Optimal Control Problem (MSCP)*, given by

$$\min_{\boldsymbol{y},\boldsymbol{v}} J(\boldsymbol{y}) + C(\boldsymbol{v}) \tag{MSCP}$$

$$\text{s.t.}\quad \begin{aligned} \dot{\boldsymbol{y}} &= f(\boldsymbol{y},\boldsymbol{v}) \\ \boldsymbol{y}(0) &= \boldsymbol{y}_0 \\ \boldsymbol{v}(t) &\in \{v_1,\dots,v_M\} \text{ for } t \in [0,T]. \end{aligned}$$

The state vector trajectory solving the considered dynamics is denoted by $\boldsymbol{y}$ and vector of the discrete-valued control trajectory that enters the dynamics on the right hand side is denoted by $\boldsymbol{v}$. We note that the following considerations are amenable to having additional continuously-valued control vector trajectories. However, for the sake of simplicity, we consider only discrete-valued controls in this article.

Regarding the initial value problem constraining (MSCP), we assume that f is Lipschitz continuous in the first argument, which yields the existence of an absolutely continuous solution $\boldsymbol{y}$ of the state equation, see [3]. The objective is the sum of two terms. The first term J depends on the state vector and does not play an important role in the remainder. The second term C will encode switching costs as they arise in the aforementioned applications. We will assume a certain structure of C for our rounding algorithm in Section III-A.

B. Contribution

We propose a novel approach towards MIOCPs with switching costs. This extends current modeling frameworks for MIOCPs and allows to generate solutions trading approximation of optimal state vector trajectory of a relaxed problem against reducing switching costs. We demonstrate this behavior by comparing our method with an established rounding method for approximation of state vectors of MIOCPs.

C. Structure of the remainder

We continue with a summary of the approximation methodology, which we use as the starting point for our considerations. Then, we introduce our rounding method and show how it fits in the described approximation framework.

In Section IV, we describe and analyze the results of a computational example, which we use to evaluate our rounding method. We close with Section V, in which we draw a conclusion and summarize the method's benefits and drawbacks, which we identified in the results.

II. RELATED WORK

Our method is developed along the ideas of Sager et al. [6], [10], [20], [22]. The authors of the aforementioned publications solve Mixed-Integer Optimal Control Problems (MIOCPs) by executing the following steps:

1) Reformulate the problem by means of partial outer convexification.
2) Solve a continuous relaxation of the MIOCP.
3) Compute a rounding on some discretization grid to obtain a discrete-valued control trajectory from the continuously-valued one.

The first step serves to replace the integer-valued variables by discrete-valued variables that are designed to switch on and off the different realizations of the dynamics $f(\cdot, v_i)$ encoded by the v_i. The equivalent partial outer convexification reformulation of (MSCP) reads

$$\min_{\boldsymbol{y},\boldsymbol{\omega}} J(\boldsymbol{y}) + C\left(\sum_{i=1}^{M} \boldsymbol{\omega}_i v_i\right) \tag{BC}$$

$$\text{s.t.} \quad \begin{aligned} \dot{\boldsymbol{y}} &= \textstyle\sum_{i=1}^{M} \boldsymbol{\omega}_i f(\boldsymbol{y}, v_i) \\ \boldsymbol{y}(0) &= \boldsymbol{y}_0 \\ \boldsymbol{\omega}(t) &\in \{0,1\}^M \text{ for } t \in [0,T] \\ \textstyle\sum_{i=1}^{M} \boldsymbol{\omega}_i(t) &= 1 \text{ for } t \in [0,T]. \end{aligned}$$

Quite naturally, the continuous relaxation arises by relaxing the last two constraints to convex combinations, specifically we obtain

$$\min_{\boldsymbol{y},\boldsymbol{\alpha}} J(\boldsymbol{y}) + C\left(\sum_{i=1}^{M} \boldsymbol{\alpha}_i v_i\right) \tag{RC}$$

$$\text{s.t.} \quad \begin{aligned} \dot{\boldsymbol{y}} &= \textstyle\sum_{i=1}^{M} \boldsymbol{\alpha}_i f(\boldsymbol{y}, v_i) \\ \boldsymbol{y}(0) &= \boldsymbol{y}_0 \\ \boldsymbol{\alpha}(t) &\in [0,1]^M \text{ for } t \in [0,T] \\ \textstyle\sum_{i=1}^{M} \boldsymbol{\alpha}_i(t) &= 1 \text{ for } t \in [0,T]. \end{aligned}$$

The key approximation property given in the literature above is summarized in the following proposition.

Proposition 2.1 (see [15]): Let $f(\cdot, v_i) : \mathbb{R}^{n_y} \to \mathbb{R}^{n_y}$ be Lipschitz continuous for all $i \in \{1, \ldots, M\}$. Let $\boldsymbol{\alpha} \in L^\infty((0,T), \mathbb{R}^M)$ be given and $(\boldsymbol{\omega}^{(h)})_h \subset L^\infty((0,T), \mathbb{R}^M)$ satisfy the convergence property

$$\sup_{t\in(0,T)} \left\| \int_0^t \boldsymbol{\alpha} - \boldsymbol{\omega}^{(h)} \right\|_\infty \to 0 \tag{1}$$

for $h \to 0$. Then,

$$\boldsymbol{y}^{(h)} \to \boldsymbol{y}$$

if $\boldsymbol{y}$ denotes the solution of the IVP in (RC) and the $\boldsymbol{y}^{(h)}$ denote the solutions of the IVPs in (BC). ■

Proposition 2.1 has severe consequences. In particular the following corollary is of high value.

Corollary 2.2 (see [15]): Let the assumptions of Proposition 2.1 hold. Let J be continuous with respect to $\boldsymbol{y}$. Then,

$$J(\boldsymbol{y}^{(h)}) \to J(\boldsymbol{y}) \text{ for } h \to 0.$$

■

Thus, if $(\boldsymbol{y}, \boldsymbol{\alpha})$ was a minimizer of (RC), $J(\boldsymbol{y})$ can be approximated arbitrarily well by means of binary-valued controls $\boldsymbol{\omega}^{(h)}$. Moreover, if the second term in the objective is not present, the minimum of (RC) and the infimum of (BC) coincide.

We note that the result is constructive, since there are rounding algorithms satisfying this property, such as *Sum-Up Rounding*, (see [10], [20]), *Next-Forced Rounding* (see [7]), or the family of *Combinatorial Integral Approximation* algorithms (see [22]).

We will compare our results against one of them, namely the rounding algorithm (SUR) that has been analyzed in [10], [20]. It has been used successfully for practical applications in the past, see [11], [24], and is introduced in the definition below.

Definition 2.3 (Alg. (SUR)*):* Let $\boldsymbol{\alpha} \in L^\infty((0,T), \mathbb{R}^M)$ satisfy the last two constraints of (RC). Let $t_0 < \ldots < t_N$ be a grid discretizing $(0,T)$ with $0 < t_k - t_{k-1} \le h$ for all $k \in \{1, \ldots, N\}$.

Then, the binary-valued step function

$$\boldsymbol{\omega} : [0,T] \to \{0,1\}^M,$$
$$\boldsymbol{\omega}_i(t) := \begin{cases} 1, & \text{if } i = i^*(k) \\ 0, & \text{else} \end{cases} \quad \text{for all } t \in [t_{k-1}, t_k)$$

is constructed iteratively for $1 \le k \le N$ by the following rule to determine the rounding index $i^*(k)$ for the interval $[t_{k-1}, t_k)$:

$$\begin{aligned} i^*(k) &:= \operatorname{argmax}_{i\in\{1,\ldots,M\}} \{\gamma_{k,i}\}, \\ \gamma_{k,i} &:= \int_0^{t_k} \boldsymbol{\alpha}_i(t)\,\mathrm{d}t - \int_0^{t_{k-1}} \boldsymbol{\omega}_i(t)\,\mathrm{d}t. \end{aligned} \tag{SUR}$$

Algorithm (SUR) indeed satisfies (1), which is proven in [10] for the case of equidistant grids. This is stated in the following proposition.

Proposition 2.4: Let the assumptions of Definition 2.3 hold. Let $\boldsymbol{\omega}^{(h)}$ be constructed from $\boldsymbol{\alpha}$ by means of (SUR) for an equidistant discretization of $[t_0, t_\mathrm{f}]$ with $h = (t_\mathrm{f} - t_0)/N$. Then it holds that

$$\sup_{t\in[0,T]} \left\| \int_0^t \boldsymbol{\alpha}(s) - \boldsymbol{\omega}^{(h)}(s)\,\mathrm{d}s \right\|_\infty \le \sum_{\ell=2}^{M} \frac{1}{\ell} h.$$

In particular, (1) is satisfied. ■

III. A SWITCH-COST AWARE ROUNDING ALGORITHM

Equipped with the findings from Section II, we make the following observation. Instead of minimizing the left side of (1), which is e.g. achieved by means of a mixed-integer linear program (MILP) in [22], we can write the left hand side as a constraint into an optimization problem and insert an objective that minimizes the costs C for a given h. By means of this methodology, we can conserve the convergence property (1), but reduce the costs C in practice. Of course, the costs C can still be unbounded as the satisfaction of (1) may necessitate infinitely many switches when driving $h \to 0$. However, we will obtain a means to trade the convergence off against the increase of the switching costs. This result is particularly useful for the applications mentioned in Section I, seeing as mechanical systems often do not allow for h to become arbitrarily small in practice.

A. Preparations

Now, we set forth and formalize this idea. Let $0 < t_0 < \ldots < t_N = T$ be a grid discretizing $(0,T)$ with maximum grid coarseness $h \coloneqq \max_{1\le k\le N} t_k - t_{k-1}$ and let $\boldsymbol{\alpha} \in L^\infty((0,T),\mathbb{R}^M)$ satisfy the last two constraints of (RC). We introduce the following quantities and variables

$$
\begin{aligned}
h_k &\coloneqq t_k - t_{k-1},\\
\boldsymbol{\alpha}_k &\coloneqq \frac{1}{h_k}\int_{t_{k-1}}^{t_k} \boldsymbol{\alpha}(t)\,\mathrm{d}t \in [0,1]^M,\\
\boldsymbol{\omega}_k &\in \{0,1\}^M,\\
\boldsymbol{\epsilon}_k &\in \{0,1\}^M,\\
\boldsymbol{\lambda}_k &\in \{0,1\}^M,
\end{aligned}
$$

for $k \in \{1,\ldots,N\}$ for h_k, $\boldsymbol{\alpha}_k$ and $\boldsymbol{\omega}_k$ and $k \in \{1,\ldots,N-1\}$ for $\boldsymbol{\epsilon}_k$ and $\boldsymbol{\lambda}_k$. Here, $\boldsymbol{\alpha}_k$ denotes the value of $\boldsymbol{\alpha}$ averaged over the k-th interval, $\boldsymbol{\omega}_k$ is the desired output of the rounding to indicate which realization v_i of the discrete states is switched on in which interval, $\boldsymbol{\epsilon}_{k,i}$ will indicate a *switch on* of the i-th state from interval $k-1$ to k and $\boldsymbol{\lambda}_k$ *switch off* of the i-th state. Clearly, we can reconstruct the function $\boldsymbol{\omega}$ from the $\boldsymbol{\omega}_k$ as $\boldsymbol{\omega} = \sum_{k=1}^M \chi_{[t_{k-1},t_k)}\boldsymbol{\omega}_k$, where χ_A denotes the characteristic function for the set A. Furthermore, we take a discrete view on C and assume that it can be written as

$$
C\left(\sum_{i=1}^M\sum_{k=1}^{N-1}\chi_{[t_{k-1},t_k)}\boldsymbol{\omega}_{k,i}v_i\right) = \sum_{k=1}^{N-1}\sum_{i=1}^M c_i\boldsymbol{\epsilon}_{k,i} + d_i\boldsymbol{\lambda}_{k,i},
$$

with $c_i, d_i \ge 0$, which corresponds to an integration of the *switch on/off* costs of the predefined states with respect to Dirac measures at the jumps.

B. The ILP for rounding

Now, we can state the switch-cost aware rounding heuristic in the ILP *Switching-Cost Aware Rounding Problem* (SCARP) below.

$$
\begin{aligned}
\min_{\boldsymbol{\omega}_{k,i},\boldsymbol{\epsilon}_{k,i},\boldsymbol{\lambda}_{k,i}}\ & \sum_{k=1}^{N-1}\sum_{i=1}^M c_i\boldsymbol{\epsilon}_{k,i} + d_i\boldsymbol{\lambda}_{k,i} \qquad \text{(SCARP)}\\
& + \sum_{i=1}^M c_i\boldsymbol{\omega}_{1,i} + \sum_{i=1}^M d_i\boldsymbol{\omega}_{N,i}\\
\text{s.t. } & \sum_{i=1}^M \boldsymbol{\omega}_{k,i} = 1 \text{ for all } k \in \{1,\ldots,N\}\\
& -Kh \le \sum_{\ell=1}^k h_\ell(\boldsymbol{\alpha}_{\ell,i} - \boldsymbol{\omega}_{\ell,i}) \le Kh\\
& \quad \text{for all } k \in \{1,\ldots,N\}, i \in \{1,\ldots,M\}\\
& \boldsymbol{\omega}_{k+1,i} - \boldsymbol{\omega}_{k,i} \le \boldsymbol{\epsilon}_{k,i}\\
& \quad \text{for all } k \in \{1,\ldots,N-1\}, i \in \{1,\ldots,M\}\\
& \boldsymbol{\omega}_{k,i} - \boldsymbol{\omega}_{k+1,i} \le \boldsymbol{\lambda}_{k,i}\\
& \quad \text{for all } k \in \{1,\ldots,N-1\}, i \in \{1,\ldots,M\}\\
& \boldsymbol{\omega}_{k,i}, \boldsymbol{\epsilon}_{k,i}, \boldsymbol{\lambda}_{k,i} \in \{0,1\} \text{ for all } i,k
\end{aligned}
$$

We immediately obtain the following proposition that guarantees the convergence of the corresponding state vector sequences with the theory summarized in Section II.

Proposition 3.1: Let $K \ge 1$. Let $\boldsymbol{\alpha} \in L^\infty((0,T),\mathbb{R}^M)$ satisfy the last two constraints of (RC). Let $t_0 < \ldots < t_N$ be a grid discretizing $(0,T)$ with $h \coloneqq \max_{1\le k\le N}(t_k - t_{k-1})$. Then, (SCARP) has a solution. Consider the function $\omega^{(h)} \coloneqq \sum_{k=1}^M \chi_{[t_{k-1},t_k)}\boldsymbol{\omega}_k^{(h)}$ with the $\omega_{k,i}^{(h)}$ solving (SCARP). Then,

$$
\sup_{t\in(0,T)}\left\|\int_0^t \boldsymbol{\alpha}(s) - \boldsymbol{\omega}(s)\,\mathrm{d}s\right\|_\infty \le Kh
$$

. In particular, (1) holds true.

Proof: We note that the sup in the desired estimate is actually a max as the integral is an absolutely continuous operation. As $\boldsymbol{\omega}$ is a binary-valued piecewise constant function on the intervals (t_{k-1}, t_k) and $\boldsymbol{\alpha}$ is positive and entrywise bounded by 1, the functions $t \mapsto \int_{t_{k-1}}^t \boldsymbol{\alpha}_i - \boldsymbol{\omega}_i$ are monotone for $t \in (t_{k-1}, t_k)$. Consequently, the supremum (maximum) of the left hand side of (1) is attained at the grid points t_k. Furthermore, we observe

$$
\int_0^{t_k} \boldsymbol{\alpha}_i - \boldsymbol{\omega}_i = \sum_{\ell=1}^k h_\ell(\boldsymbol{\alpha}_{\ell,i} - \boldsymbol{\omega}_{\ell,i})
$$

for all i. Combining these two observations gives

$$
\sup_t\left\|\int_0^t \boldsymbol{\alpha} - \boldsymbol{\omega}\right\|_\infty = \max_{k,i}\left|\sum_{\ell=1}^k h_\ell(\boldsymbol{\alpha}_{\ell,i} - \boldsymbol{\omega}_{\ell,i})\right|
$$

Thus, by the bound $\pm Kh$ on the right term in (SCARP), any feasible point of (SCARP) implies the desired bound for the reconstructed step function. It remains to show that a feasible point exist. The algorithm Next-Forced Rounding, see [7], produces some $(\boldsymbol{\omega}_{k,i})_{1\le k\le N, 1\le i\le M}$ that satisfies the constraint with constant $K = 1$. For the other constraints observe that setting $\boldsymbol{\epsilon}_{k,i} = \boldsymbol{\lambda}_{k,i} = 1$ for all i, k does the job. Hence, a feasible point of (SCARP) exists and as only finitely many feasible points for (SCARP) may exist, (SCARP) has a solution. ■

In the proof, we use the results for Next-Forced Rounding, despite the fact that the algorithm is not very widely used, because its results for the bound with $K = 1$ are asymptotically (for $M \to \infty$) tight.

C. *Interpretation of* (SCARP)

As mentioned before, (SCARP) allows to trade the state vector approximation off against the switching costs. Usually, the maximal frequency for switching is subject to some physical (mechanical or electrical) constraints, which will determine h. Thus, from the setup of (SCARP), it is clear that the parameter governing the trade-off is K.

A high value of K leaves more room for the ILP solver to find a solution with low switching costs while a value close to 1 will not leave much room for switching cost minimization and most likely result in a high value of the second term in the objective.

Note that for high values of K and N, we expect (SCARP) to become prohibitively hard to compute as we assume it can be reduced to a knapsack problem, which is known to be weakly NP-hard. We defer the analysis to future work.

IV. COMPUTATIONAL EXAMPLE

A. *Setup*

To illustrate the effect of introducing discrete switching costs and to compare the proposed modeling approach with an established method for MIOCP, we use the Lotka-Volterra benchmark fishing problem introduced in [21], which can be found in a benchmark library[1] for MIOCPs [19]. The problem without the additional discrete switching cost summand in the objective reads

$$\min_{\boldsymbol{y},\boldsymbol{\omega}} \int_{t_0}^{t_f} (\boldsymbol{y}_0(t)-1)^2 + (\boldsymbol{y}_1(t)-1)^2 \mathrm{d}t \qquad \text{(LV)}$$

$$\text{s.t.} \quad \begin{array}{rcl} \dot{\boldsymbol{y}}_0 &=& \boldsymbol{y}_0 - \boldsymbol{y}_0\boldsymbol{y}_1 - c_0\boldsymbol{y}_0 \sum_{i=1}^M \boldsymbol{\omega}_i v_i, \\ \dot{\boldsymbol{y}}_1 &=& -\boldsymbol{y}_1 + \boldsymbol{y}_0\boldsymbol{y}_1 - c_1\boldsymbol{y}_1 \sum_{i=1}^M \boldsymbol{\omega}_i v_i, \\ \boldsymbol{y}(t_0) &=& (0.5,\, 0.7,\, 0)^T, \\ \boldsymbol{\omega}(t) &\in& \{0,\, 1\}^M \text{ for } t \in [t_0,\, t_f], \\ \sum_{i=1}^M \boldsymbol{\omega}_i(t) &=& 1 \text{ for } t \in [t_0,\, t_f]. \end{array}$$

We have already phrased the dynamics in convexified form. For our computations, we used the values $c_0 = 0.4$, $c_1 = 0.2$ and $[t_0,\, t_f] = [0,\, 12], M = 3$. The problem is particularly suited as an example for our study as the relaxed solution of (LV) contains a singular arc [18] and was thoroughly investigated under different aspects [18], [21], [22]. The discrete control realizations are $v_1 = 1$, $v_2 = 0.2$ and $v_3 = 0.0$.

The addition of switching costs to the objective of (LV) could be interpreted as the necessity of a fisher to rent equipment, i.e. for $\boldsymbol{\omega}_2(t) = 1$ the fisher uses a fishing rod, while for $\boldsymbol{\omega}_1(t) = 1$ he uses a fishing net. We modeled the switching costs as described in Section III and for our computations, which are evaluated below, (SCARP) was solved with switch on costs $c = (2\ 1\ 0)^T$, switch off costs $d = (0.1\ 0.1\ 0)^T$ and $K = \frac{5}{6}$, the constant from Proposition 2.4.

B. *Computational setup*

We solved the continuous relaxation of (LV) by means of a *first discretize, then optimize* methodology, using direct multiple shooting to discretize the dynamics (see [1] for details on the topic) employing the solver MUSCOD-II, see [13]. We solved (SCARP) using version 8.1 of the GUROBI optimizer, see [14], All computations were conducted on an Intel Core i7-965 clocked at 3.20 Ghz.

C. *Validation against* (SUR)

We computed the solution of (RC) disregarding the switching costs for the setup from Section IV-A for two tracking-type objectives, one with L^1-regularized control J_{L^1}, and one without regularization J_{NO}, in formulas:

$$J_{NO}(\boldsymbol{y}) = \int_0^T (\boldsymbol{y}_0(t)-1)^2 + (\boldsymbol{y}_1(t)-1)^2 \,\mathrm{d}t,$$

$$J_{L^1}(\boldsymbol{y}) = J_{NO}(\boldsymbol{y}) + \eta \left\| \sum_{i=1}^M \boldsymbol{\alpha}_i v_i \right\|_{L^1}.$$

In our computations, the regularizing constant was chosen as $\eta = 0.01$. This yielded the trajectories $\boldsymbol{\alpha}_{L^1}$ and $\boldsymbol{\alpha}_{NO}$. Then, we computed roundings $\boldsymbol{\omega}^{(h)}_{\mathcal{R},\mathcal{A}}$ from $\boldsymbol{\alpha}_{\mathcal{R}}$ for the rounding algorithms $\mathcal{A} \in \{\text{(SUR)}, \text{(SCARP)}\}$, the relaxations $\mathcal{R} \in \{L^1, NO\}$ and equidistant discretizations of (t_0, t_f) with $h \in \{2^{-1}(t_f - t_0), \ldots, 2^{-10}(t_f - t_0)\}$.

The state equation has been solved for these $\boldsymbol{\omega}^{(h)}_{\mathcal{R},\mathcal{A}}$ and compared to the state trajectory for the corresponding $\boldsymbol{\alpha}_{\mathcal{R}}$ using the relative error

$$d^{(h)}_{\mathcal{R},\mathcal{A}} = \frac{\sup_t \left\| \boldsymbol{y}(\boldsymbol{\omega}^{(h)}_{\mathcal{R},\mathcal{A}}(t)) - \boldsymbol{y}(\boldsymbol{\alpha}_{\mathcal{R}})(t)) \right\|_2}{\sup_t \left\| \boldsymbol{y}(\boldsymbol{\alpha}_{\mathcal{R}})(t)) \right\|_2}.$$

For both regularizations, the accuracy of the state vector approximation by means of (SUR) is very similar to the one of (SCARP). There is almost no difference with respect to the different regularizations. This is visualized in Figure 1.

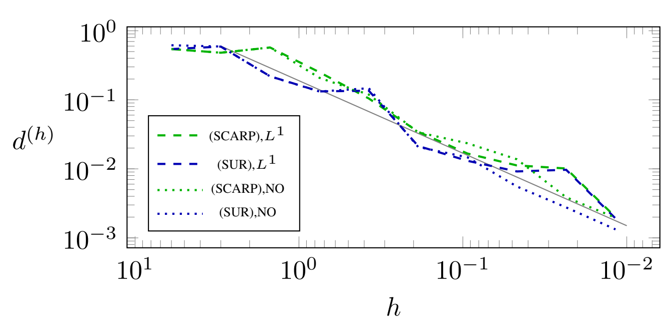

Fig. 1. Validation of theoretical results regarding the approximations of the state vectors of (RC) for (SCARP) (green) and (SUR) (blue) with (dashed) and without (dotted) L^1 regularization. The grey line visualizes the desired linear decrease of the error.

[1]The library consisting of problems along with their descriptions may be found at www.mintoc.de

As the solution operator of the Lotka-Volterra IVP is nonlinear, we cannot expect monotone decrease of the error for all iterations.

We observe a similar behavior for the objectives, where we compute the relative error as

$$e^{(h)}_{\mathcal{R},\mathcal{A}} = \frac{|J_{\mathcal{R}}(\boldsymbol{y}(\boldsymbol{\omega}^{(h)}_{\mathcal{R},\mathcal{A}})) - J_{\mathcal{R}}(\boldsymbol{y}(\boldsymbol{\alpha}_{\mathcal{R}}))|}{|J_{\mathcal{R}}(\boldsymbol{y}(\boldsymbol{\alpha}_{\mathcal{R}}))|}.$$

This is visualized in Figure 2.

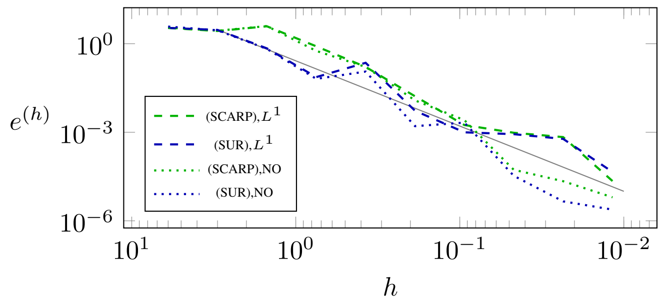

Fig. 2. Validation of theoretical results regarding the approximations of the objective of (RC) for (SCARP) (green) and (SUR) (blue) with (dashed) and without (dotted) L^1 regularization.

Note that for L^1-regularization, the objective converges although it depends on $\boldsymbol{\alpha}$, which is not covered by our statements on (MSCP) and its reformulation and relaxation. This follows easily from Proposition 2.4 in our setting, but is not true in general, see [16].

We also compute the switching costs

$$C^{(h)}_{\mathcal{R},\mathcal{A}} = C\left(\sum_{i=1}^{M} \boldsymbol{\omega}^{(h)}_{\mathcal{R},\mathcal{A}}\right)$$

and observe that the switching costs of the roundings computed with (SCARP) are a lower bound for the switching costs of the roundings computed with (SUR) although (SUR) is already optimal in some cases. For both algorithms, the switching costs increase with the grid refinements. This is visualized in Figure 3.

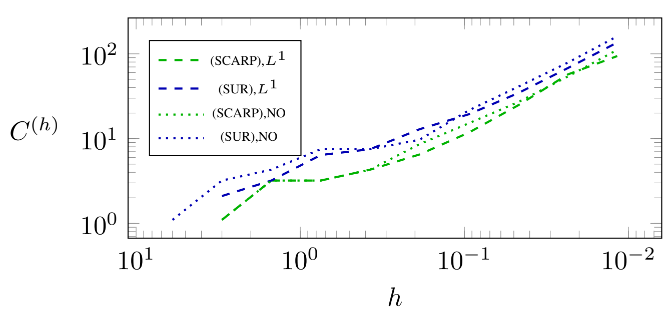

Fig. 3. Switching costs for (SCARP) (green) and (SUR) (blue) with (dashed) and without (dotted) L^1 regularization.

D. Influence of the tradeoff parameter K

In the setting from Section IV-C, we also computed solutions of (SCARP) for relaxed bounds on the approximation constraint, in particular, we used the choices $K \in \frac{5}{6} \cdot \{1, 1.5, 2\}$. The switching costs in the optimal solution sink with the increase of K, which is visualized in Figure 4, while the state vector approximation is less accurate, which is visualized in Figure 5. Increasing the parameter

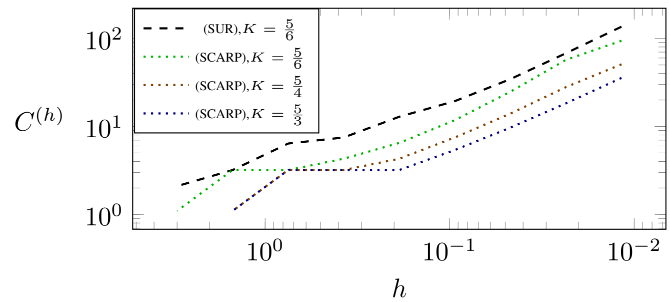

Fig. 4. Switching costs (L^1-regularized case) for (SUR) (dashed) and (SCARP) (dotted) for $K = \frac{5}{6}, \frac{5}{4}, \frac{5}{3}$.

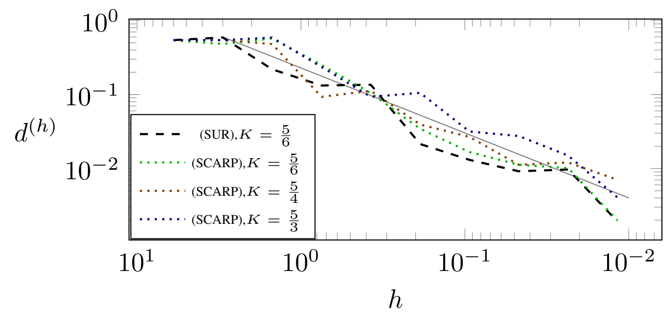

Fig. 5. Approximations of the state vectors (L^1-regularized case) of (RC) for (SUR) (dashed) and (SCARP) (dotted) for $K = \frac{5}{6}, \frac{5}{4}, \frac{5}{3}$.

K increases the computational effort GUROBI has to spend significantly. This is visualized in Figure 6 for the case of the L^1 regularized relaxation. The results are similar for the unregularized relaxation.

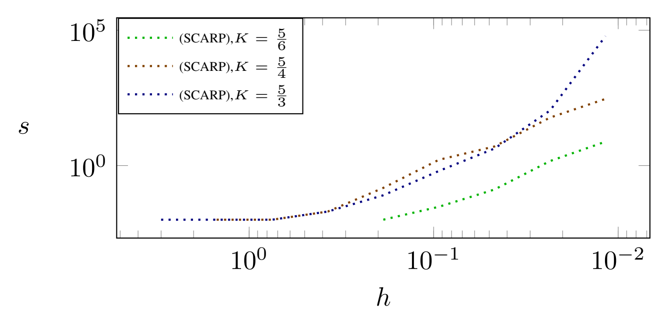

Fig. 6. Compute time of (SCARP) in case of the L^1 regularization.

V. CONCLUSION

We presented an alternative rounding approach that allows make a trade-off between approximation of the optimal state vector trajectory of a relaxed problem that ignores the switching costs and the possible requirement of reducing the switching costs.

The numerical results validate our theoretical findings in several regards. As solutions of (SCARP) will exploit the boundaries of the approximation constraint to reduce the switching costs, it seems likely that (SUR) yields a better approximation of the state vector $\boldsymbol{y}$ and objective term J obtained by solving (RC) without switching costs. As (SUR) computes a feasible point for (SCARP), its switching costs have to exceed the optimal switching costs computed by solving (SCARP). Both considerations can be observed in the plots in Section IV-C. Furthermore, the switching costs increase very fast when the grid is refined. This is also expected as the fact that $\sup_t \left\|\int_0^t \boldsymbol{\alpha} - \boldsymbol{\omega}\right\|$ has to be driven to zero can only occur by more and more frequent switching in intervals where $\boldsymbol{\alpha}(t) \notin \{0, 1\}^M$.

The results in Section IV-D show that we can indeed trade in a decrease in state vector approximation quality (factor of 2 when doubling K for $N = 1024$) for much better switching costs (factor of 4 when doubling K for $N = 1024$).

Finally, we highlight that (SUR) is an $\mathcal{O}(N)$ algorithm that can easily be used for real-time applications, which is not that easy for (SCARP) because we have to solve an ILP in this case, which can have much higher computational costs. As the number of feasible points grows with the increase of K, we have to invest higher compute times to obtain the solutions of (SCARP).

ACKNOWLEDGMENTS

C. Kirches and P. Manns acknowledge funding by Deutsche Forschungsgemeinschaft through Priority Programme 1962 “Non-smooth and Complementarity-based Distributed Parameter Systems: Simulation and Hierarchical Optimization”. C. Kirches and C. Hansknecht were supported by the German Federal Ministry of Education and Research, grant n° 61210304-ODINE. C. Kirches was supported by the German Federal Ministry of Education and Research, grants n° 05M17MBA-MoPhaPro and 05M18MBA-MOReNet.

REFERENCES

[1] H.G. Bock and K.-J. Plitt. A multiple shooting algorithm for direct solution of optimal control problems. *IFAC Proceedings Volumes*, 17(2):1603–1608, 1984.

[2] A. Bürger, C. Zeile, A. Altmann-Dieses, S. Sager, and M. Diehl. An algorithm for mixed-integer optimal control of solar thermal climate systems with MPC-capable runtime. 07 2018.

[3] F. Clarke. *Functional analysis, calculus of variations and optimal control*, volume 264. Springer Science & Business Media, 2013.

[4] M. Gerdts. Solving mixed-integer optimal control problems by branch&bound: a case study from automobile test-driving with gear shift. *Optimal Control Applications and Methods*, 26:1–18, 01 2005.

[5] I.E. Grossmann, P.A. Aguirre, and M. Barttfeld. Optimal synthesis of complex distillation columns using rigorous models. In A. Barbosa-Póvoa and H. Matos, editors, *European Symposium on Computer-Aided Process Engineering-14*, volume 18 of *Computer Aided Chemical Engineering*, pages 53–74. Elsevier, 2004.

[6] F.M. Hante and S. Sager. Relaxation methods for mixed-integer optimal control of partial differential equations. *Computational Optimization and Applications*, 55(1):197–225, 2013.

[7] M. Jung. *Relaxations and approximations for mixed-integer optimal control*. PhD thesis, Heidelberg University, 2013.

[8] M.N. Jung, C. Kirches, S. Sager, and S. Sass. Computational approaches for mixed integer optimal control problems with indicator constraints. *Vietnam Journal of Mathematics*, 46(4):1023–1051, December 2018.

[9] C. Kirches, E. Kostina, A. Meyer, and M. Schlöder. Numerical solution of optimal control problems with switches, switching costs and jumps. *Optimization Online*, 6888, October 2018.

[10] C. Kirches, F. Lenders, and P. Manns. Approximation properties and tight bounds for constrained mixed-integer optimal control. *Optimization Online Preprint*, 5404, April 2016. (submitted).

[11] C. Kirches, S. Sager, H.G. Bock, and J.P. Schlöder. Time-optimal control of automobile test drives with gear shifts. *Optimal Control Applications and Methods*, 31:137–153, 03 2010.

[12] C. Kirchner, M. Herty, S. Göttlich, and S. Klar. Optimal control for continuous supply network models. *Networks and hetegenerous media*, 1:675–688, 2006.

[13] D.B. Leineweber, I. Bauer, H.G. Bock, and J.P. Schlöder. An efficient multiple shooting based reduced SQP strategy for large-scale dynamic process optimization. Part 1: theoretical aspects. *Computers & Chemical Engineering*, 27(2):157–166, 2003.

[14] Gurobi Optimization LLC. Gurobi optimizer reference manual, 2018.

[15] P. Manns and C. Kirches. Improved regularity assumptions for partial outer convexification of MIPDECOs. *Optimization Online Preprint*, 6585, April 2018. (submitted).

[16] P. Manns and C. Kirches. Multidimensional sum-up rounding for elliptic control systems. *DFG Preprint SPP1962-080*, 2018. (submitted).

[17] M. Ringkamp, S. Ober-Blöbaum, and S. Leyendecker. Discrete mechanics and mixed integer optimal control of dynamical systems. *PAMM*, 14, 12 2014.

[18] S. Sager. *Numerical Methods for Mixed-Integer Optimal Control Problems*. Der Andere Verlag, 2005.

[19] S. Sager. A benchmark library of mixed-integer optimal control problems. In Jon Lee and Sven Leyffer, editors, *Mixed Integer Nonlinear Programming*, pages 631–670, New York, NY, 2012. Springer New York.

[20] S. Sager, H.G. Bock, and M. Diehl. The integer approximation error in mixed-integer optimal control. *Mathematical Programming*, 133(1–2):1–23, 2012.

[21] S. Sager, H.G. Bock, M. Diehl, G. Reinelt, and J.P. Schlöder. *Numerical methods for optimal control with binary control functions applied to a Lotka-Volterra type fishing problem*. Springer, 2005.

[22] S. Sager, M. Jung, and C. Kirches. Combinatorial Integral Approximation. *Mathematical Methods of Operations Research*, 73(3):363–380, 2011.

[23] O. Stursberg and S. Engell. Optimal control of switched continuous systems using mixed-integer programming. *IFAC Proceedings Volumes*, 35(1):433–438, 2002. 15th IFAC World Congress.

[24] D. Telen, F. Logist, E. Van Derlinden, I. Tack, and J. Van Impe. Optimal experiment design for dynamic bioprocesses: a multi-objective approach. *Chemical Engineering Science*, 78:82–97, 2012.

[25] F. Zhu and P. J. Antsaklis. Optimal control of hybrid switched systems: A brief survey. *Discrete Event Dynamic Systems*, 25(3):345–364, Sep 2015.