Deutsche Forschungsgemeinschaft

Priority Programme 1962

Matching Algorithms and Complexity Results for Constrained Mixed-Integer Optimal Control with Switching Costs

Felix Bestehorn, Christian Kirches

SPP 1962

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

Preprint Number SPP1962-164

received on March 3, 2021

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/

MATCHING ALGORITHMS AND COMPLEXITY RESULTS FOR CONSTRAINED MIXED-INTEGER OPTIMAL CONTROL WITH SWITCHING COSTS[*]

FELIX BESTEHORN[†] AND CHRISTIAN KIRCHES[‡]

Abstract. We extend recent work on the performance of the combinatorial integral approximation decomposition approach for Mixed-Integer Optimal Control Problems (MIOCPs) in the presence of combinatorial constraints or switching costs on an equidistant grid. For the time discretized problem, we reformulate the emerging rounding problem in the decomposition approach as a matching problem on a bipartite graph and show that a feasible integral control can be obtained by computing a maximal matching. Given a relaxed solution of an MIOCP with combinatorial constraints the approach allows the computation of a control with minimal integrated control deviation θ within $\mathcal{O}(NM\sqrt{N+M}\log(N+M))$, where M is the number of binary controls and N denotes the number of discretized time intervals. We derive a tight bound on θ for decomposition methods in the presence of combinatorial constraints and show that the worst-case approximation bound for this problem class is larger than for problems without combinatorial constraints. Two modeling extensions involving the switching costs are then considered. We provide a polynomial runtime algorithm for the class of rounding problems containing (switching) costs independent of previous rounding choices and give $\mathcal{NP}$-hardness and inapproximability proofs for the class of rounding problems with (switching) cost functions dependent on previous rounding choices.

Key words. Optimal Control, Mixed-Integer Programming, Combinatorial Optimization, Computational Complexity, Discrete approximations in optimal control

AMS subject classifications. 34H05, 90C11, 90C27, 90C60, 49M25

1. Introduction.

Mixed-Integer Optimal Control Problems (MIOCPs) present a powerful tool to model and solve many real-world problems and have been gaining considerable attention in recent years. The scope of applications ranges from the shifting of gears in a truck [14, 26], chemical engineering [2, 6], biology [18, 19], renewable energy and heating [7, 27] to switched dynamic systems [24, 38].

In this work we consider MIOCPs constrained by ordinary differential equations (ODEs). Additionally, an inequality constraint on the system states and the binary control (CC) and an additional switching cost term $C(\cdot)$ is present. The general form of the considered mixed-integer switching cost problem with combinatorial constraints on a fixed and finite time horizon $[0,T] \subset \mathbb{R}$ for some $T > 0$ is given by

$$\begin{aligned}
&\inf_{y,v} \; J(y) + C(v) \\
&\text{(MSCP-CC)} \qquad \text{s.t.} \; \dot{y}(t) = f(y(t), v(t)) \text{ for (a. a.) } t \in [0,T], \; y(0) = y_0, \\
&\qquad v(t) \in \{v_1, \ldots, v_M\} \text{ for (a. a.) } t \in [0,T], \\
&\text{(CC)} \qquad 0 \le c(y(t), v(t)) \text{ for (a. a.) } t \in [0,T].
\end{aligned}$$

The measurable function $v : [0,T] \to \{v_1, \ldots, v_M\} \subset \mathbb{R}^{n_V}$ is the discrete-valued

[*]Submitted to the editors DATE.

Funding: C. Kirches and F. Bestehorn acknowledge funding by Deutsche Forschungsgemeinschaft through Priority Programme 1962 “Non-smooth and Complementarity-based Distributed Parameter Systems: Simulation and Hierarchical Optimization”. C. Kirches has received funding by the German Federal Ministry of Education and Research through grants n° 61210304-ODINE, 05M18MBA-MOReNet and 05M20MBA-LEOPLAN.

[†]Institute for Mathematical Optimization, Technische Universität Braunschweig, Braunschweig, Germany (f.bestehorn@tu-bs.de)

[‡]Institute for Mathematical Optimization, Technische Universität Braunschweig, Braunschweig, Germany (c.kirches@tu-bs.de)

control input with choices $\{v_1, \dots, v_M\} \in \mathbb{R}^{n_V}$. The function $y : [0,T] \to \mathbb{R}^{n_Y}$ is the state vector of an initial value problem (IVP). For the functions f and c we assume Lipschitz continuity in the first argument.

The continuous function $C : L^\infty((0,T), \mathbb{R}^{n_V}) \to \mathbb{R}$ assigns cost to the control function v. We leave its regularity unspecified here but introduce a distinguishing property of switching cost functions in Section 2. A natural choice is a measure of switching or fixed costs, e.g. the total variation or cost of operating in a certain state.

Several different approaches to solve MIOCPs have been investigated. Reformulation methods, e.g. *variable time-transformations* [9, 10, 15] as well as *bilevel optimization* [8], *switching time optimization* [37, 38], *dual dynamic integer programming* [29] and *branch & bound* [14, 35] methods have been proposed. These methods are either challenging to apply in practice due to their limited applicability to only selected problem classes, necessitate the sometimes difficult evaluation of the costate, or require an a-priori fixed maximal number of switchings. The presented approach can be seen as complementary to the aforementioned methods and follows the decomposition approach originating from [33]. It consists of reformulating the problem by means of a partial outer convexification, determining a relaxed control $\widetilde{\alpha}$ and computing a discrete-valued control $\widetilde{\omega}$ by application of a rounding algorithm. For this methodology to work, the rounded control on a grid with maximum interval length $\overline{\Delta}$ has to satisfy the approximation property

$$\sup_{t\in[0,T]} \left\| \int_0^t [\widetilde{\alpha}(s) - \widetilde{\omega}(s)] \, \mathrm{d}s \right\|_\infty \leq \overline{\Delta}\theta,$$

which will be defined rigorously in Subsection 2.2. Therefore, one aim is to minimize the *integral control deviation* θ, while quickly obtaining a rounded binary control $\widetilde{\omega}$. Two well known approaches for MIOCPs without (CC) constraints and switching costs are Sum-up Rounding (SUR) [33] and Next-forced Rounding (NFR) [22]. Unfortunately, in the presence of (CC) constraints, both methods do not perform well [32, Example 3.2] or even break down [30, Section 5.4]. The variant (SUR-VC) of (SUR) proposed in [25] is able to handle vanishing constraints, but its integral control deviation bound is still dependent on the number M of control choices. As *switching costs* are difficult to handle for (SUR),(SUR-VC), and (NFR), a switching cost aware rounding method (SCARP) based on shortest paths in a directed acyclic graph was proposed in [4]. By construction, this approach can handle (CC) constraints as well as switching costs, but has a runtime exponential in the number of controls choices M. The current capabilities of the decomposition approach in terms of established bounds on θ as well as the capability to handle (CC) constraints or switching costs (SC) are summarized in Table 1.

1.1. Contribution. We extend the theoretical and algorithmic foundations of the partial outer convexification approach addressing MIOCPs with constraints or switching costs on the binary control on an equidistant grid.

We provide an algorithmic framework for MIOCPs with constraints on the binary control based on graph theoretical matchings. The presented approach allows the calculation of a binary control with an upper bound of 1 on the integral control deviation in $\mathcal{O}(\sqrt{N+M}NM)$. It is shown that the bound is tight for the problem class and that the presented approach can also be used to calculate a binary control within the tight bound of $\theta = \frac{2M-3}{2M-2}$ for the combinatorial integral approximation problem [40], thus improving the worst-case runtime reported there.

We generalize the problem further by adding switching costs $C(v)$ to MIOCPs

TABLE 1
Comparison of known algorithms for the decomposition methodology. A + indicates that the method is suited for handling problems with this feature, while − indicates the opposite. The label Parameter *indicates that θ can be chosen for the method (an infeasibility certificate is generated if the chosen θ disallows a solution) or can be optimized.*

	θ	CC	SC	Runtime
SUR [25]	$\sum_{i=2}^{M} \frac{1}{M}$	−	−	$\mathcal{O}(NM)$
SUR-VC [32]	$\lfloor M/2 \rfloor$	+	−	$\mathcal{O}(NM)$
NFR [22]	1	−	−	$\mathcal{O}(N^2M)$
SCARP [4]	Parameter	+	+	$\mathcal{O}\left(\frac{N(\lfloor 2\theta \rfloor+3)^{2M}}{M\left((\lfloor 2\theta \rfloor+3)^2-1\right)}\right)$

with (CC) constraints and adapt the previously developed approach. We prove that if the switching costs adhere to an easily verifiable condition regarding sequence independency, optimal solutions can still be obtained in a worst-case runtime subcubic in N. In total, we propose four rounding schemes which are summarized in Table 2.

TABLE 2
Rounding via Matching (RVM) approaches for the decomposition methodology introduced in this article. A + indicates that the method is suited for handling problems with this feature, while − indicates the opposite. The indication θ^* indicates that a feasible binary control with minimal integral control deviation θ will be computed. ISC *indicate sequence independent switching costs, Def. 2.6*

	θ	CC	ISC	Runtime
RVM (Cor. 4.10)	$\frac{2M-3}{2M-2}$	−	−	$\mathcal{O}(NM\sqrt{N+M})$
RVM-VC (Thm. 4.4)	1	+	−	$\mathcal{O}(NM\sqrt{N+M})$
RVM-θ (Cor. 5.2)	θ^*	+	−	$\mathcal{O}(NM\sqrt{(N+M)\log{(N+M)}})$
RVM-ISC (Thm. 5.4)	$\frac{2M-3}{2M-2}$ or 1 (VC-case)	+	+	$\mathcal{O}\left(N^2M\lfloor 2\theta+1\rfloor + N^2\log(N)\right)$

Furthermore, we investigate the general setting for sequence dependent switching costs and prove that calculating a solution within a given integral control deviation θ constitutes a strongly $\mathcal{NP}$-hard problem. Additionally, we prove that the general setting is as hard to approximate as TSP and thus does not even admit a polynomial time approximation within any bound, unless $\mathcal{P} = \mathcal{NP}$.

1.2. Structure of the Remainder. The remainder of this article is structured in the following way. Section 2 introduces the reformulation and relaxation of (MSCP-CC). A short introduction into the decomposition methodology for MIOCPs is given subsequently and the rounding problems investigated in this article are formally introduced. In Section 3 we establish the equivalence of some of the aforementioned rounding problems and matchings in bipartite graphs. The matching framework is used in Section 4 and provides the means to compute an integer feasible control for rounding problems with vanishing controls and with an integral deviation gap solely dependent on the grid refinement in subquadratic time. An extension of the problem class together with an algorithmic approach for sequence independent switching costs is considered in Section 5. Section 6 carries out the proof of strong $\mathcal{NP}$-hardness as well as inapproximability for sequence dependent switching functions. We conclude with a brief summary of the article in Section 7.

1.3. Notation. For $N \in \mathbb{N}$ we denote the set of numbers $\{1, \dots, N\}$ by $[N]$. A row slice of row s until row k of a matrix $A \in \mathbb{R}^{N\times M}$ is denoted by

$A_{s:k} \in \mathbb{R}^{((k-s+1)\times M)}$. Similarly, a column slice of column s until column k of row ℓ of matrix A is denoted by $A_{\ell,s:k} \in \mathbb{R}^{(k-s+1)}$.

2. Approximating (MSCP-CC). Following the decomposition methodology developed along the ideas of Sager et al. [3, 17, 25, 34, 35, 40] we derive an equivalent reformulation in conjunction with two relaxations of (MSCP-CC). We state the consistency property, Definition 2.2, which is required for the essential result, Proposition 2.3 and Proposition 2.4 of the decomposition methodology with and without vanishing constraints. The consistency property and the underlying results motivate the desire to improve rounding algorithms.

2.1. Reformulation and Relaxation of (MSCP-CC) and its variants. We reformulate (MSCP-CC) by replacing the control function v by a $\{0,1\}^M$ valued function ω that models a *special ordered set of type 1* (SOS-1) activation of the different control realizations $v_1, v_2, \ldots, v_M$. This reformulation is called a partial outer convexification as it allows to move the control activations in the second argument of f to the different right hand sides of the ODE. The reformulated problem without the combinatorial constraint, i.e. $c(y(t), v(t)) \equiv 0$, was given by the authors in [3, 4], while the variant including the combinatorial constraint, but without the control cost function, i.e. $C(v) \equiv 0$, is discussed in [25, 32]. We note that relaxing the ODE as well as the combinatorial constraint in this way is motivated in [25] and was first presented in [34]. Combining the reformulations from [4, 25] we arrive at the following partial outer convexification of (MSCP-CC):

$$
\begin{aligned}
& \inf_{y,\alpha}\ J(y) + C\left(\sum_{i=1}^{M} \omega_i v_i\right) \\
& \text{s.t.}\quad \dot{y}(t) = \sum_{i=1}^{M} \omega_i(t) f(y(t), v_i) \text{ for (a. a.) } t \in [0,T],\ y(0) = y_0, \\
& \text{(BC}_\delta\text{-VC)} \qquad \omega_i(t) \in \{0,1\}^M \text{ for (a. a.) } t \in [0,T], \\
& \qquad \sum_{i=1}^{M} \omega_i(t) = 1 \\
& \text{(BCC)} \qquad -\delta \le \omega_i(t) c(y(t), v_i) \text{ for (a. a.) } t \in [0,T] \text{ and } 1 \le i \le M.
\end{aligned}
$$

The reformulation is equivalent to (MSCP-CC) for $\delta = 0$ and a relaxation if $\delta > 0$, see [25]. Another natural relaxation of (MSCP-CC) arises by relaxation of the (SOS-1) constraint to convex coefficients for $\delta = 0$:

$$
\begin{aligned}
& \min_{y,\omega}\ J(y) + C\left(\sum_{i=1}^{M} \alpha_i v_i\right) \\
& \text{s.t.}\quad \dot{y}(t) = \sum_{i=1}^{M} \alpha_i(t) f(y(t), v_i) \text{ for (a.a.) } t \in [0,T],\ y(0) = y_0, \\
& \text{(VC-RC)} \qquad \alpha_i(t) \in [0,1]^M \text{ for (a.a.) } t \in [0,T], \\
& \text{(SOS-1)} \qquad \sum_{i=1}^{M} \alpha_i(t) = 1 \\
& \text{(RCC)} \qquad 0 \le \alpha_i(t) c(y(t), v_i) \text{ for (a.a.) } t \in [0,T] \text{ and } 1 \le i \le M.
\end{aligned}
$$

Throughout the article we tacitly assume that (VC-RC) admits a solution, which is indicated by changing the infimum in the formulation (BC_δ-VC) to a minimum in (VC-RC). Furthermore we note that the proposed method is applicable to all feasible points of (VC-RC) and is not restricted to optimal solutions. We also highlight that this natural relaxation should only be used, if the switching cost function can also be relaxed in a natural way, e.g. if the switching cost function is *sequence independent*, see Definition 2.6 below.

In the absence of the (RCC) or (BCC) constraints we will denote the problems (VC-RC) or (BC_δ-VC) as (RC) or (BC) respectively. Following the terminology from [31] we call a function $\widetilde{\omega} \in L^\infty((0,T),\mathbb{R}^M)$ satisfying the second and third constraint of (BC_δ-VC) a *binary control*. Similarly, a function $\alpha \in L^\infty((0,T),\mathbb{R}^M)$ satisfying the second and third constraint of (VC-RC) is a *relaxed control*. We will denote the binary and relaxed controls that are functions in $L^\infty((0,T),\mathbb{R}^M)$ by $\widetilde{\ }$. Discretized functions using piecewise constant discretizations are denoted by matrices without $\widetilde{\ }$.

2.2. Approximation of binary controls through rounding algorithms. Of course, binary controls $\widetilde{\omega}$ are piecewise constant functions. An algorithm deriving a binary control from a relaxed control, therefore naturally operates on a grid, which we call rounding grid from now on and define formally below together with the rounding algorithms themselves.

Definition 2.1 (Rounding grid). *Let $0 = t_0 < \ldots < t_n = T$ be a grid discretizing $[0,T]$ into $[N]$ intervals. The set $\{[t_0,t_1),[t_1,t_2),\ldots,[t_{N-2},t_{N-1}),[t_{N-1},t_N]\}$ is called a rounding grid. If the* length of an interval *is denoted by $\Delta_k := t_k - t_{k-1}$, then $\overline{\Delta} := \max_{k\in[N]} \Delta_k$ is called the* mesh size.

We call an algorithm a *rounding algorithm* if it determines a binary control that is constant per interval from a relaxed control and a rounding grid. The following consistency property facilitates the relationship between (BC_δ-VC) and (VC-RC) and its variants.

Definition 2.2 (Consistency property). *A rounding algorithm is called* consistent *if there exists a constant $\theta > 0$ such that for all relaxed controls $\widetilde{\alpha}$ and all rounding grids, the determined binary control $\widetilde{\omega}$ satisfies*

$$d(\widetilde{\omega},\widetilde{\alpha}) \leq \theta\overline{\Delta},$$

where d is the pseudo-metric given as

$$d(\widetilde{\omega},\widetilde{\alpha}) := \sup_{t\in[0,T]} \left\| \int_0^t [\widetilde{\alpha}(s) - \widetilde{\omega}(s)]\,\mathrm{d}s \right\|_\infty . \tag{2.1}$$

The consistency property takes a central role in obtaining the proofs for the approximation relationship between (BC_δ-VC) and (VC-RC), which is stated in Proposition 2.3 and forms the theoretical basis for the decomposition methodology.

Proposition 2.3 ([25, Theorem 3.6],[32, Theorem 2.1]). *Let $f(\,\cdot\,,v_i) : \mathbb{R}^{n_y} \to \mathbb{R}^{n_y}$ and $c(\,\cdot\,,v_i) : \mathbb{R}^{n_y} \to \mathbb{R}^{n_y}$ be Lipschitz continuous for all $i \in [M]$. Let $\widetilde{\alpha} \in L^\infty((0,T),\mathbb{R}^M)$ be a relaxed control and let y denote the solutions of the IVP in (VC-RC) for $\widetilde{\alpha}$. Let $\overline{\Delta}^{(n)}$ be a sequence of mesh sizes that satisfies $\overline{\Delta}^{(n)} \to 0$. Let $(\widetilde{\omega}^{(n)})_n \subset L^\infty((0,T),\mathbb{R}^M)$ be a corresponding sequence of binary controls such that $\widetilde{\omega}^{(n)}$ and $\widetilde{\alpha}$ satisfy $d(\widetilde{\omega}^{(n)},\widetilde{\alpha}) \leq \theta\overline{\Delta}^{(n)}$ for all $n \in \mathbb{N}$ and $\operatorname{supp}\widetilde{\omega}_i^{(n)} \subset \operatorname{supp}\widetilde{\alpha}_i$ for all $i \in [M]$ and all $n \in \mathbb{N}$, where* supp *denotes the essential support. Then,* *a)* $\widetilde{\omega}^{(n)} \rightharpoonup^* \widetilde{\alpha}$ *in* $L^\infty((0,T),\mathbb{R}^M)$.

b) $y^n \to \widetilde{y}$ *in* $C([0,T],\mathbb{R}^{n_y})$ *and* $\dot{y}^{(n)} \rightharpoonup^* \dot{\widetilde{y}}$ *in* $L^\infty((0,T),\mathbb{R}^M)$, *where* $y^{(n)}$ *and* $\widetilde{y}$ *are the solutions of the state equations of* (BC$_\delta$-VC) *and* (VC-RC) *for* $\widetilde{\omega}^{(n)}$ *and* $\widetilde{\alpha}$.

c) If $0 \le \widetilde{\alpha}_i(t)c(y(t),v_i)$ *a.e. then* $-\delta_i^{(n)} \le \widetilde{\omega}_i^{(n)}(t)c(y^{(n)}(t),v_i)$ *a.e. for all* $n \in \mathbb{N}$ *and* $\delta_i^{(n)} \to 0$.

d) If J *is continuous, then*

$$\begin{aligned} &\min\{J(y) : (\widetilde{\alpha}, y) \textit{ feasible for } (\text{VC-RC})\} \\ =&\inf\{J(y) : (\widetilde{\omega}, y) \textit{ feasible for } (\text{BC}_\delta\text{-VC})\}. \end{aligned}$$

As observed in [32, Example 3.2] and [16, Section 3], the partial outer convexification of the (CC) constraint in (MSCP-CC) leads to a vanishing constraint in (BC$_\delta$-VC) and a rounding algorithm, which, when applied to a problem of this form, should uphold the following condition (VC) as otherwise the solution may constantly violate the (CC) constraint, that is 2.3.c) does not hold.

$$\text{(VC)} \qquad \omega_{t,i} = 1 \Rightarrow \alpha_{t,i} > 0 \text{ for all } t \in [N].$$

The condition guarantees that a rounding solution will violate the (CC) constraint by no more than a small mesh size dependent value $\delta^{(n)} \ge 0$, see [25, Section 5]. The following sections utilize (VC) to obtain a binary control such that the result from Proposition 2.3 holds. In absence of the (CC) constraint, the consistency property allows to prove the following result for (BC) and (RC).

PROPOSITION 2.4 ([31, Theorem 5.1]). *Let* $f(\cdot, v_i) : \mathbb{R}^{n_y} \to \mathbb{R}^{n_y}$ *be Lipschitz continuous for all* $i \in [M]$. *Let* $\widetilde{\alpha} \in L^\infty((0,T),\mathbb{R}^M)$ *be a relaxed control, and let* $\widetilde{y}$ *denote the solution of the IVP in* (RC) *for* $\widetilde{\alpha}$. *Let* $(\widetilde{\omega}^{(n)})_n \subset L^\infty((0,T),\mathbb{R}^M)$ *be a sequence of binary controls such that*

$$d(\widetilde{\omega}^{(n)}, \widetilde{\alpha}) \to 0.$$

Then, the solutions $y^{(n)}$ *of the IVPs in* (BC) *for the control inputs* $\widetilde{\omega}^{(n)}$ *satisfy*

$$y^{(n)} \to \widetilde{y} \textit{ in } C([0,T],\mathbb{R}^{n_Y}).$$

Furthermore, let J *be continuous. Then,*

$$\min\{J(y) : (\widetilde{\alpha}, y) \textit{ feasible for } (\text{RC})\} = \inf\{J(y) : (\widetilde{\omega}, y) \textit{ feasible for } (\text{BC})\}.$$

We note that algorithms Sum-up Rounding (SUR) [33], Sum-up Rounding for Vanishing Constraints (SUR-VC) [25], Next-forced Rounding (NFR) [22, 40], the CIA [35] and the shortest path approach [4] satisfy Definition 2.2 and consequently Proposition 2.4. In the case of (SUR-VC) and the shortest path approach, Proposition 2.3 holds as well.

For ease of further reference we conclude this section by distinguishing the different rounding problems, which will be discussed during the remainder of this article.

DEFINITION 2.5 (Rounding problems (CIA) and (CIA-VC)). *Let* $C(v) \equiv 0$ *for all* v, *let* $\overline{\Delta}$ *be the mesh size and* $\theta \in \mathbb{R}$ *be the maximally allowed integral control deviation.*

Let $\alpha \in [0,1]^{N \times M}$ *be a solution of* (RC), *then the rounding problem (CIA) is defined as finding a feasible binary control* $\omega \in \{0,1\}^{N \times M}$ *with* $d(\omega, \alpha) \le \theta\overline{\Delta}$ *for* (BC).

Let $\alpha \in [0,1]^{N\times M}$ *be a solution of* (VC-RC)*, then the rounding problem (CIA-VC) is defined as finding a feasible binary control* $\omega \in \{0,1\}^{N\times M}$ *with* $d(\omega,\alpha) \leq \theta\overline{\Delta}$ *and*

$$\omega_{t,i} = 1 \Rightarrow \alpha_{t,i} > 0 \text{ for all } t \in [N] \tag{VC}$$

for (BC_δ-VC)*.*

We introduce two natural extensions of the (CIA) and (CIA-VC) problems, where costs are added to the switches, i.e. $C(\cdot) \neq 0$. In the first extension the switching costs are sequence independent, while the second extension generalizes the concept and makes no assumption on the sequence independence of controls.

Definition 2.6 (Sequence independent switching costs). *The switching cost function* $C(\omega)$ *is called* sequence independent *if for all* $t \in [N]$ *and all* $i \in [M]$ *the switching costs for* $C(\omega_{t,i}, v_i)$ *with* $\omega_{t,i} = 1$ *can be determined without computing* $\omega_{t-1,1:M} \in \{0,1\}^{1\times M}$*.*

We will show that if the switching cost function is sequence independent then the underlying rounding problem is significantly simpler.

Definition 2.7 (Rounding problems (ISC-CIA) and (DSC-CIA)). *Let* $\alpha \in [0,1]^{N\times M}$ *be a solution of* (RC)*, let* $\overline{\Delta}$ *be the mesh size and* $\theta \in \mathbb{R}$ *be the largest allowed integral control deviation.*

Let the switching costs $C(\cdot)$ *be* sequence independent*, then the rounding problem (ISC-CIA) is defined as the problem of finding a feasible binary control* $\omega \in \{0,1\}^{N\times M}$ *with minimal switching costs and* $d(\omega,\alpha) \leq \theta\overline{\Delta}$ *for* (BC)*.*

Let $C(\cdot)$ *be an arbitrary switching cost function, then the rounding problem (DSC-CIA) is defined as the problem of finding a feasible binary control* $\omega \in \{0,1\}^{N\times M}$ *for* (BC) *with minimal switching costs and* $d(\omega,\alpha) \leq \theta\overline{\Delta}$*.*

Analogously to the pair of (CIA) *and* (CIA-VC)*, combinatorial constraints can be present and therefore the* (VC) *condition in the partial outer convexified variant can be added to* (ISC-CIA) *and* (DSC-CIA)*, giving rise to the rounding problems (ISC-CIA-VC) and (DSC-CIA-VC).*

The problem classes (ISC-CIA) and (DSC-CIA) along with their vanishing constraint counterparts generalize the rounding problems (CIA) and (CIA-VC). Solving these problems efficiently allows a higher modeling capability of the decomposition approach as more information of the original (MSCP-CC) problem can be included into the rounding process, while maintaining the favorable properties of rounding algorithms.

3. Rounding via matching (RVM). This section introduces the graph construction, which will be used in the following two sections for the purpose of solving the rounding problems (CIA),(CIA-VC), (ISC-CIA) and (ISC-CIA-VC). A first idea for this procedure was given by the authors in the short proceedings paper [5].

For the following construction we let $\widetilde{\alpha}$ be a relaxed control and given some $\theta \in \mathbb{R}^+$, we seek an approximating binary control $\widetilde{\omega} \in L^\infty((0,T),\mathbb{R}^M)$ adhering to the (VC) condition, such that the integrated control deviation $d(\widetilde{\omega},\widetilde{\alpha})$ does not exceed $\theta\overline{\Delta}$. From Proposition 2.3 we know that instead of determining $\widetilde{\omega}$ we can also determine $\omega \in \{0,1\}^{N\times M}$ with ω satisfying the (SOS-1) and (VC) conditions. Under the assumption of an equidistant grid we propose to reformulate the rounding problem as a matching problem on a bipartite graph and show that computing a maximal matching covering certain vertices is equivalent to determining a control $\widetilde{\omega} \in L^\infty((0,T),\mathbb{R}^M)$ satisfying the (VC) condition as well as $d(\omega,\alpha) \leq \theta\overline{\Delta}$.

Assumption 3.1. *The rounding grid is equidistant, i.e.* $\Delta_t = \overline{\Delta}$ *for all* $t \in [N]$.

On the one hand Assumption 3.1 allows to factor $d(\alpha,\omega)$ by $\overline{\Delta}$, which we do throughout the proofs of the remainder of the article, and on the other hand it allows to define the number of *necessary activations* and *maximally permissible* activations for all controls $i \in [M]$ and all grid points $t \in [N]$.

Definition 3.2 (Necessary and maximally permissible activations). *Let Assumption 3.1 hold, let* $\theta \in \mathbb{R}^+$ *and let* $\alpha \in [0,1]^{N\times M}$ *be a relaxed control adhering to* (SOS-1). *Then the number of* necessary activations *for control* $i \in [M]$ *is defined as*

$$p^{(i)} := \max\left\{0, \left\lceil \sum_{t=1}^{N} \alpha_{t,i} - \theta \right\rceil\right\}$$

and the number of maximally permissible activations *for control* $i \in [M]$ *is defined by*

$$q^{(i)} := \left\{0, \left\lfloor \sum_{t=1}^{N} \alpha_{t,i} + \theta \right\rfloor\right\}.$$

The motivation stems from the fact that for $d(\omega,\alpha) \leq \theta$ to hold it is necessary that

$$p^{(i)} \leq \sum_{t=1}^{N} \omega_{t,i} \leq q^{(i)} \text{ for all } i \in [M]. \tag{3.1}$$

Considering the whole time grid, Definition 3.2 inherently gives rise to the definition of sets of grid points in which a certain activation of a control has to happen, as otherwise the bound θ on the integral control deviation cannot be satisfied.

Definition 3.3 (Control set). *Let Assumption 3.1 hold, let* $\theta \in \mathbb{R}^+$ *and let* $\alpha \in [0,1]^{N\times M}$ *be a relaxed control. The* control set *for control* $i \in [M]$ *with* $p \in \mathbb{N}\backslash\{0\}, p \leq q^{(i)}$, *is*

$$\begin{aligned} S_{i,p} := &\left\{ t \in [N] \,\middle|\, \left\lceil \sum_{k=1}^{t} \alpha_{k,i} - \theta \right\rceil \leq p-1,\, p \leq \left\lfloor \sum_{k=1}^{t} \alpha_{k,i} + \theta \right\rfloor \right\} \\ &\bigcup \min_{t\in[N]} \left\{ t \in [N] \,\middle|\, \left\lceil \sum_{k=1}^{t-1} \alpha_{k,i} - \theta \right\rceil \leq p-1 < \left\lceil \sum_{k=1}^{t} \alpha_{k,i} - \theta \right\rceil \right\}. \end{aligned} \tag{3.2}$$

The superset of control sets *is the set*

$$S := \left\{S_{1,1}, S_{1,2}, \ldots, S_{1,q^{(1)}}, S_{2,1}, \ldots, S_{M,1}, \ldots, S_{M,q^{(M)}}\right\}$$

and the subset of necessary controls *is defined as*

$$S_p := \left\{S_{1,1}, S_{1,2}, \ldots, S_{1,p^{(1)}}, S_{2,1}, \ldots, S_{M,1}, \ldots, S_{M,p^{(M)}}\right\}.$$

Thus, a control $i \in [M]$ with $p \leq q^{(i)}$ and with control set $S_{i,p} = \{t_1,t_2,t_3\}$, $t_1,t_2,t_3 \in [N]$ must have been activated at least p times at grid point t_3 and must not be activated p times before grid point t_1.

Control sets enable the definition of bipartite graphs on control sets and grid points, Definition 3.4, which can then be used to determine appropriate controls $\omega \in \{0,1\}^{N \times M}$. For ease of notation, we will, in the remainder of this article, access vertices through their corresponding labels, e.g. a vertex representing the control set $S_{i,p}$ will be denoted by $v_{S_{i,p}}$ and a vertex representing the grid point $t \in [N]$ by v_t.

DEFINITION 3.4 (Bipartite (vanishing constraint) control graph). *Let Assumption 3.1 hold, let $\theta \in \mathbb{R}$ and let $\alpha \in [0,1]^{N \times M}$ be a relaxed control. Let S be the superset of controls from Definition 3.3, $V(S)$ be the vertices representing S and $V(T)$ be the vertices representing the rounding grid. Then, the* bipartite control graph *$G(\theta, \alpha) := (V, E)$ is defined through the vertex set*

$$V := \{V(S) \cup V(T)\}$$

and the edge set

$$E := \left\{ (v_{S_{i,p}}, v_{t_k}) \in S \times T \,|\, t_k \in S_{i,p} \right\}.$$

The bipartite vanishing control graph *$G_{VC}(\theta, \alpha) := (V, E_{VC})$ consists of the same vertex set V as $G(\theta, \alpha)$, but has only a subset of E as edge set*

$$E_{VC} := \left\{ (v_{S_{i,p}}, v_{t_k}) \in E \,|\, t_k \in S_{i,p}, \alpha_{t_k,i} > 0 \right\}.$$

$$\underset{\alpha^T}{\begin{bmatrix} 6/21 & 0 & 0 & 15/21 \\ 5/21 & 8/21 & 0 & 6/21 \\ 5/21 & 7/21 & 10/2 & 0 \\ 5/21 & 6/21 & 11/2 & 0 \end{bmatrix}} \quad \Rightarrow$$

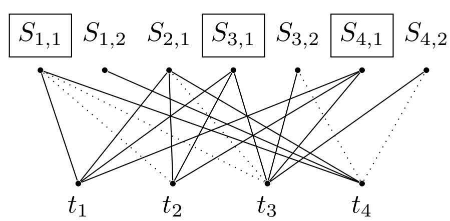


FIG. 1. *A relaxed control α with the corresponding graph $G(\alpha, \theta)$ from Definition 3.4 for $\theta = 1$. Dotted edges satisfy where $\alpha_{t,i} = 0$. The boxed vertices denote the subset of necessary controls $S_p = \{S_{1,1}, S_{3,1}, S_{4,1}\}$.*

The graphs from Definition 3.4 are bipartite by construction. Furthermore, we observe that an assignment of edges in the form of a maximal matching covering vertices $V(T)$ and $V(S_p)$ in the control graph spanned by α and θ is equivalent to a binary control ω with integral control deviation smaller than θ. This equivalence relation allows a different approach towards rounding problems, which will be shown in the next section.

PROPOSITION 3.5. *Let Assumption 3.1 hold, let $\theta \in \mathbb{R}$ and let $\alpha \in [0,1]^{N \times M}$ be a relaxed control. Then,*

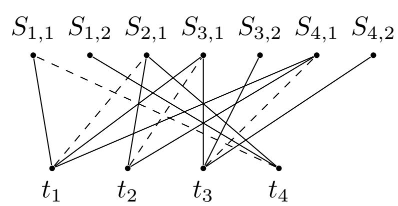
- *there exists a maximal matching M on $G_{VC}(\alpha, \theta)$, where all elements from $V(S_p) \cup V(T) \subseteq V$ are matched if and only if there exists an $\omega \in \{0,1\}^{N \times M}$ with $d(\omega, \alpha) \leq \theta \overline{\Delta}$ adhering to the (SOS-1) and (VC) conditions.*
- *there exists a maximal matching M on $G(\alpha, \theta)$, where all elements from $V(S_p) \cup V(T) \subseteq V$ are matched if and only if there exists an $\omega \in \{0,1\}^{N \times M}$ with $d(\omega, \alpha) \leq \theta \overline{\Delta}$ adhering to the (SOS-1) condition.*

Proof. First let $\mathcal{M}$ be a maximal matching $\mathcal{M}$, where all elements from $V(S_p) \cup V(T) \subseteq V$ are matched. We construct $\omega \in \{0,1\}^{N\times M}$ by setting $\omega_{t,i} = 1$ if the vertex v_t representing t is matched to a vertex $v_{S_{i,p}}, p \in P_i$. It remains to show that the constructed ω adheres to (SOS-1), condition (VC) and that $d(\omega,\alpha) \leq \theta\overline{\Delta}$ holds. Because $\omega_{t,i} = 1$ if v_t was matched to a vertex $v_{S_{i,p}}, p \in P_i$, it holds that $\sum_{i=1}^{M} \omega_{t,i} = 1$ as $\mathcal{M}$ was a maximal matching and $T \leq S$ holds by construction. Additionally, if there exists an edge $e = (v_t, v_{Si,p}), t \in [N], i \in [M], p \in P_i$, then $\alpha_{t,i} > 0$. Thus, ω adheres to (SOS-1) and the condition (VC). Furthermore, an edge $e = (v_t, v_{S_{i,p}}), t \in [N], i \in [M], p \in P_i$ only exists if $\left\lceil \sum_{k=1}^{t} \alpha_{t,i} - \theta \right\rceil \leq p \leq \left\lfloor \sum_{k=1}^{t} \alpha_{t,i} + \theta \right\rfloor$ holds. This implies $\max_{i\in[M]} |\sum_{k=1}^{t} \alpha_{t,i} - \omega_{t,i}| \leq \theta$ for all $t \in [N]$ and $i \in [M]$. Together with Assumption 3.1, $d(\omega,\alpha) \leq \theta\overline{\Delta}$ follows.

Now, let $\omega \in \{0,1\}^{N\times M}$ with $d(\omega,\alpha) \leq \theta\overline{\Delta}$ such that the (SOS-1) and (VC) conditions hold. Let $\mathcal{M} = \emptyset$ and traverse the matrix $\omega \in \{0,1\}^{N\times M}$ row-wise, adding the edge $(v_t, v_{S_{i,p}})$ to the matching $\mathcal{M}$ if $\omega_{t,i} = 1$ and $\sum_{k=1}^{t-1} \omega_{k,i} = p - 1$. It remains to show that $\mathcal{M}$ is maximal and all elements from $V(S_p) \cup V(T) \subseteq V$ are matched. As ω fulfilled the (SOS-1) condition, $\sum_{i=1}^{M} \omega_{t,i} = 1$ holds for all $t \in [N]$ and therefore all elements from $V(T)$ are matched. This additionally implies maximality of $\mathcal{M}$ as $G_{VC}(\alpha,\theta)$ was bipartite and one of the two partitions was $V(T)$. Assuming that vertex $v_{S_{i,p}}, i \in [M], p \in P_i$ is not matched in $\mathcal{M}$ it follows immediately that $d(\omega,\alpha) \leq \theta\overline{\Delta}$ did not hold, which contradicts the assumption made with respect to ω. Thus $\mathcal{M}$ is maximal and all elements from $V(S_p) \cup V(T) \subseteq V$ are matched.

For the second claim we observe that $G_{VC}(\alpha,\theta)$ is a subgraph of the bipartite graph $G(\alpha,\theta)$ and thus the second claim follows analogously to the first claim. ∎

$$\underbrace{\begin{bmatrix} 6/21 & 0 & 0 & 15/21 \\ 5/21 & 8/21 & 0 & 6/21 \\ 5/21 & 7/21 & 10/2 & 0 \\ 5/21 & 6/21 & 11/2 & 0 \end{bmatrix}}_{\alpha^T} \Rightarrow$$



$$\Rightarrow \underbrace{\begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{\omega^T}$$

FIG. 2. *A relaxed control α and the corresponding graph $G_{VC}(\alpha,\theta)$ from Definition 3.4 for $\theta = 1$. Dashed edges are edges of a maximal matching used to construct a binary control ω adhering to (VC) with integral control deviation $\max_{t\in[N]} \|\sum_{k=1}^{t} \alpha_k - \omega_k\|_\infty = 16/21$ in $t \in \{1\}$ for $i = 2$. Dashed edges are edges of a maximal matching used to construct a binary control ω adhering to (VC) with integral control deviation $\max_{t\in[N]} \|\sum_{k=1}^{t} \alpha_k - \omega_k\|_\infty = 16/21$ in $t \in \{1\}$ for $i = 2$. Note that for this instance the (SUR) algorithm would compute a control with $\omega_{1,1} = \omega_{2,2} = \omega_{3,3} = \omega_{4,4} = 1$, while the (SUR-VC) algorithm would construct $\omega_{1,1} = \omega_{2,2} = \omega_{3,3} = \omega_{4,1} = 1$. The deviation for both algorithms would be $\frac{22}{21}$ in $t = 3, i = 4$ and (SUR) would violate the (VC) in $t = 4, i = 4$.*

Proposition 3.5 allows to obtain solutions to (BC_δ-VC) and (BC) by performing a maximal matching on a bipartite graph in such a way that certain vertices are matched. For example, the vertices that have to be matched in Figure 2 are the vertices corresponding to T and to $S_p = \{S_{1,1}, S_{3,1}, S_{4,1}\}$.

The following section investigates how large θ has to be chosen for a matching satisfying the conditions to exist, and the runtime necessary to calculate such a matching.

4. Calculating controls with tight integral control deviation for MIOCP and MPVC. It is known from the previous section that for a relaxed solution α and an allowed integral deviation value θ an integral control ω adhering to the (SOS-1) and (VC) condition can be generated by determining a maximal matching satisfying the additional property that all vertices stemming from S_p and T are covered in the matching. We now show that this matching always exists for $\theta \geq 1$. Having established the existence of such a matching, we analyze the maximal cardinalities of $|V|$ and $|E|$ to establish a worst-case runtime estimate (dependent on N, M and θ) for calculating such a matching. We then continue by showing that the bound on θ is tight and, using a result from [40], we show that the matching approach can also be used to solve the (CIA) rounding problem for $\theta = \frac{2M-3}{2M-2}$ efficiently.

As a first step we use the graph construction from Definition 3.4 to show that the neighborhood of the vertices belonging to T and S_p is always sufficiently large for the choice $\theta = 1$ because they fulfill Hall's condition [36, Theorem 22.1].

LEMMA 4.1. *Let Assumption 3.1 hold, let $\alpha \in [0,1]^{N\times M}$ be relaxed control and let $\theta = 1$. Then, for $G_{VC}(\alpha, \theta)$ constructed as in Definition 3.4, we have that*
a) for all $Q \subseteq V(T)$ it holds that $|Q| \leq |N(Q)|$ and
b) for all $R \subseteq V(S_p)$ it holds that $|R| \leq |N(R)|$.

Proof. For the first claim we let $Q \subseteq V(T)$ with $m \coloneqq |Q|$. The (SOS-1) constraint yields that for all $v_t \in Q$ it holds that $\sum_{i=1}^{M} \alpha_{t,i} = 1$. Thus $\sum_{t\in Q}\sum_{i=1}^{M} \alpha_{t,i} = m$. We proceed by distinguishing two cases.

In case that for every $v_{S_{i,q}} \in N(Q), i \in [M], 0 < p \leq q^{(i)}$ it yields that $\sum_{t\in Q\cap S_{i,p}} \alpha_{i,t} \geq 1$ holds, then $|N(Q)| \geq m$ follows by construction of the graph. This would conclude the proof for this direction.

For the second case assume that there exists an $v_{S_{i,q}} \in N(Q), i \in [M], 0 < p \leq q^{(i)}$ with $\sum_{t\in Q\cap S_{i,p}} \alpha_{i,t} < 1$. Application of the pidgeonhole principle then yields that for every such $v_{S_{i,q}}$ there must exist a $v_{S_{j,q}} \in N(Q), j \in [M], 0 < q \leq q^{(j)}$ such that $\sum_{t\in Q\cap S_{j,q}} \alpha_{j,t} > 1$. Therefore, a pair $j \in [M], 0 < q \leq q^{(j)}$ exists such that $\sum_{t\in Q\cap S_{j,q}} \alpha_{j,t} > 1$. By the defining chain of inequalities for control sets (3.2) and as $\theta = 1$ this implies that either $v_{S_{j,q-1}}$ or $v_{S_{j,q+1}}$ are also in the neighborhood of Q. As this holds for all such pairs (j,q) with $\sum_{t\in Q\cap S_{j,q}} \alpha_{j,t} > 1$, and as every pair can only be counted once in the neighborhood, we have $|N(Q)| \geq m$.

For the second claim we introduce the notation $S_{j_\ell,q_\ell}(L)$ and $S_{j_\ell,q_\ell}(L-1)$, where $S_{j_\ell,q_\ell}(L)$ denotes the grid point index $t_k \in [N]$ with highest subindex k belonging to S_{j_ℓ,q_ℓ}, while $S_{j_\ell,q_\ell}(L-1)$ denotes the grid point index $t_k \in [N]$ with the second highest subindex k belonging to S_{j_ℓ,q_ℓ}. Now we let $R \subseteq V(S_p)$ with $r \coloneqq |R|$. Definition 3.3 implies that there exist r different pairs $S_{j_\ell,q_\ell}, j_\ell \in [M], q_\ell \in P_{j_\ell}, \ell \in [r]$. As $\theta = 1$, we have for all $q_\ell \geq 1$ that

$$\sum_{k=1}^{S_{j_\ell,q_\ell}(L)} \alpha_{k,j_\ell} - \max\left\{0, \sum_{k=1}^{S_{j_\ell,q_\ell-1}(L-1)} \alpha_{k,j_\ell}\right\} \geq 1, \tag{4.1}$$

where the max-term is necessary to distinguish between the cases $q_\ell > 1$ and $q_\ell = 1$. From inequality (4.1) it follows that $\sum_{t\in S_{j_\ell,q_\ell}} \alpha_{t,j_\ell} \geq 1$ for all $S_{j_\ell,q_\ell} \in R$. A summation over all r different pairs contained in R leads to

$$\sum_{(j_\ell,q_\ell)\in R} \left(\sum_{t\in S_{j_\ell,q_\ell}} \alpha_{t,j_\ell} \right) \geq r.$$

The (SOS-1) constraint now yields that at least r different vertices belonging to $V(T)$

have to be adjacent to the vertices from R and therefore $|R| \leq |N(R)|$ holds. ∎

Lemma 4.1 allows to apply the following generalization of a well known matching theorem of König for bipartite graphs on the graph $G_{VC}(\alpha, \theta)$.

PROPOSITION 4.2 ([20],[36, Theorem 16.8]). *Let $G = (V, E)$ be a bipartite graph with color classes A and B. Let $W \subseteq V$. Then there exists a matching $\mathcal{M}$ covering W if and only if there exists a matching $\mathcal{M}_A$ covering $W \cap A$ and a matching $\mathcal{M}_B$ covering $W \cap B$.*

LEMMA 4.3. *Let Assumption 3.1 hold and let $\alpha \in [0,1]^{N\times M}$ be a relaxed vanishing constraint control. Let $\theta := 1$ then there exists a feasible solution $\omega \in \{0,1\}^{N\times M}$ for the rounding problem (CIA-VC) with*

$$d(\omega, \alpha) \leq \theta\overline{\Delta}. \tag{4.2}$$

Proof. Let $\alpha \in [0,1]^{N\times M}$ and $\theta = 1$. We construct a corresponding bipartite graph $G_{VC}(\alpha, \theta) = (V, E)$ in accordance with Definition 3.4. We define $W := S_p \cup T$. Combining Lemma 4.1 and Hall's matching theorem now states that for S_p and T there exist matchings $\mathcal{M}_A$ covering $W \cap S_p$ and $\mathcal{M}_B$ covering $W \cap T$. Applying Proposition 4.2 shows that there exists a matching $\mathcal{M}$ covering W. Therefore, all elements from S_p and T are matched in $\mathcal{M}$. Proposition 3.5 now yields an $\omega \in \{0,1\}^{N\times M}$, which adheres to the (SOS-1) constraint, the (VC) condition. By construction of the graph it follows that $d(\omega, \alpha) \leq \theta\overline{\Delta}$ holds as well, which concludes the proof. ∎

Naturally, existence of a binary control for $\theta = 1$, established in Lemma 4.3, implies the existence of binary controls for $\theta > 1$. Additionally the proof already gives a constructive way of determining the binary control by using a bipartite matching algorithm. One possibility is the Hopcroft-Karp algorithm [21], which has a runtime of $\mathcal{O}(\sqrt{|V|}|E|)$ on a bipartite graph $G = (V, E)$. To accomplish feasibility of the maximal matching with respect to Proposition 3.5 one first matches the elements of S_p and then the remaining elements of T, which by Lemma 4.1 leads to a maximal matching. To assess the practical computational feasibility of the approach we have to estimate the cardinalities of $|V|$ and $|E|$.

THEOREM 4.4. *Let $\alpha \in [0,1]^{N\times M}$ be a solution of (VC-RC) and let Assumption 3.1 hold. Then, for a chosen value of $\theta \geq 1$ one can determine a solution $\omega \in \{0,1\}^{N\times M}$ of (BC$_\delta$-VC) with $d(\omega, \alpha) \leq \theta\overline{\Delta}$ in $\mathcal{O}(\sqrt{2N + M\lfloor\theta + 1\rceil}NM\lfloor 2\theta\rceil)$. In particular, a solution $\omega \in \{0,1\}^{N\times M}$ of (BC$_\delta$-VC) for the choice of $\theta = 1$ adhering to $d(\omega, \alpha) \leq \overline{\Delta}$ exists and can be calculated in $\mathcal{O}(\sqrt{N + M}NM)$.*

Proof. Let $\alpha \in [0,1]^{N\times M}$ be a solution of (VC-RC) and $\theta \geq 1$ be the allowed integral control deviation. With Lemma 4.3 there exists a binary feasible solution for (BC$_\delta$-VC) with $\theta = 1$ and can be calculated by first matching all elements of $V(S_p)$ and then matching the remaining elements of T to some elements of $V(S) \backslash V(S_p)$. Following the procedure outlined in Section 3 we can construct a corresponding bipartite graph $G_{VC}(\alpha, \theta) = (V, E)$. By construction V contains exactly one vertex v_t for every $t \in [N]$, thus $|V(T)| = N$. Further, by construction of $G_{VC}(\alpha, \theta)$ the number of maximally allowed activations $q^{(i)}$ per control $i \in [M]$ implies the existence of corresponding vertices $v_{S_{i,1}}, \ldots, v_{S_{i,q(i)}}$. Therefore, from Definitions 3.2 – 3.4 it follows for $i \in [M]$ that

$$\sum_{\ell=1}^{q^{(i)}} v_{S_{i,\ell}} \underset{\text{Def. 3.4}}{\overset{\text{Def. 3.3}}{=}} q^{(i)} \overset{\text{Def. 3.2}}{=} \left\lfloor \sum_{t=1}^{N} \alpha_{t,i} + \theta \right\rceil. \tag{4.3}$$

Summation over all controls and using that the (SOS-1) constraint implies $\sum_{i=1}^{M}\sum_{t=1}^{N}\alpha_{t,i}=N$ allows to estimate the number of vertices $V(S)$ by

$$\sum_{i=1}^{M}\left\lfloor\sum_{t=1}^{N}\alpha_{t,i}+\theta\right\rfloor \overset{\lfloor\cdot\rfloor}{\leq} \sum_{i=1}^{M}\left\lfloor\sum_{t=1}^{N}\alpha_{t,i}\right\rfloor + M\lfloor\theta\rfloor + M \tag{4.4}$$

$$\overset{\lfloor\cdot\rfloor}{\leq} \left\lfloor\sum_{i=1}^{M}\sum_{t=1}^{N}\alpha_{t,i}\right\rfloor + M\lfloor\theta+1\rfloor \overset{\text{(SOS-1)}}{=} N + M\lfloor\theta+1\rfloor. \tag{4.5}$$

In total we have $|V|\leq 2N+M\lfloor\theta+1\rfloor$. It remains to estimate the number of edges $|E|$. We pick an arbitrary vertex $v_t, t\in[N]$ and let $v_{S_{i,p}}, i\in[M], p\in P_i$ be adjacent to v_t. Then, Definition 3.4 yields

$$\left\lceil\sum_{k=1}^{t}\alpha_{t,i}-\theta\right\rceil \leq p \leq \left\lfloor\sum_{k=1}^{t}\alpha_{t,i}+\theta\right\rfloor. \tag{4.6}$$

As $P_i\subset\mathbb{N}$ it follows immediately that for p_i satisfying (4.6) at most $\lfloor 2\theta\rfloor+1$ elements $p\in P_i$ satisfy (4.6) at the same grid point $t\in[N]$. By Definition of E_{VC} at most $\lfloor 2\theta\rfloor$ additional vertices originating from control i can therefore be adjacent to v_t. Observing that v_{t_1} may only be adjacent to vertices $v_{S_{i,1}}$ for $i\in[M]$ and using the fact that v_t was chosen arbitrarily and $G_{VC}(\alpha,\theta)$ was bipartite, we get

$$|E| = \sum_{t=1}^{N} N(v_t) < N\max_{t\in[N]} N(v_t) = NM\left\lfloor 2\theta+1\right\rfloor. \tag{4.7}$$

Inserting $|V|$ and $|E|$ into the runtime estimation for the Hopcroft-Karp algorithm and using $\theta\geq 1$ now yields the first claim. The second claim for $\theta=1$ follows immediately by insertion of $\theta=1$ in (4.4) and (4.7). ∎

Remark 4.5. *It was shown in [1] that finding a maximal matching the Hopcroft-Karp algorithm will, with high probability, need a runtime of only* $\mathcal{O}\left(\log\left(2N+M\lfloor\theta+1\rfloor\right)NM\left\lfloor 2\theta\right\rfloor\right)$. *It was additionally shown that this probability rises if the average degree is large. Following the proof of Theorem 4.4 and especially inequality (4.4) this occurs more often when θ is larger, as the average degree of $G_{VC}(V,E)$ is strongly dependent on the value of θ.*

In general one cannot guarantee existence of a solution for (CIA-VC) for the choice of $\theta<1$. That is because given $\theta<1$ and $M\geq 3$, one can construct an $\alpha\in[0,1]^{N\times M}$ satisfying (SOS-1) such that there exists no $\omega\in\{0,1\}^{N\times M}$ satisfying (SOS-1) and the (VC) condition with $d(\omega,\alpha)\leq\theta\overline{\Delta}$, as follows:

Theorem 4.6. *Let Assumption 3.1 hold and let $\epsilon>0$. Let $\theta=1-\epsilon$. Then for $M=3$ there exists an $N\in\mathbb{N}$ and an $\alpha\in[0,1]^{N\times M}$ satisfying (SOS-1) such that there does not exists an $\omega\in\{0,1\}^{N\times M}$ satisfying (SOS-1) and the (VC) condition with $d(\omega,\alpha)<\theta\overline{\Delta}$.*

Proof. Let $\theta=1-\epsilon$ with $\epsilon>0$ then there exists $C\in\mathbb{N}$ such that

$$\theta<\theta^* := \frac{CM-3}{CM-2}. \tag{4.8}$$

Let $N \coloneqq 2CM - 8$ and construct $\alpha \in [0,1]^{N \times M}$ as follows.

$$\alpha_{t,i} \coloneqq \begin{cases} \theta^* & \text{if } i \equiv 1 \text{ and } t \pmod 2 = 1, \\ \theta^* & \text{if } i \equiv 2 \text{ and } t \pmod 2 = 0, \\ 1-\theta^* & \text{if } i \equiv 2 \text{ and } t \pmod 2 = 1 \text{ and } \quad t > 1, \\ 1-\theta^* & \text{if } i \equiv 3 \text{ and } t \pmod 2 = 0 \text{ or } \quad t = 1, \\ 0 & \text{else.} \end{cases} \tag{4.9}$$

$$\begin{bmatrix} 6/7 & 0 & 6/7 & 0 & 6/7 & 0 & 6/7 & 0 & 6/7 & 0 \\ 0 & 6/7 & 1/7 & 6/7 & 1/7 & 6/7 & 1/7 & 6/7 & 1/7 & 6/7 \\ 1/7 & 1/7 & 0 & 1/7 & 0 & 1/7 & 0 & 1/7 & 0 & 1/7 \end{bmatrix}$$

$$\alpha^T$$

FIG. 3. *The relaxed control α constructed using (4.9) for $\theta^* = \frac{3M-3}{3M-2}$, where $M \equiv 3$.*

Clearly, α satisfies the (SOS-1) condition. It remains to show that there exists no $\omega \in \{0,1\}^{N \times M}$ satisfying (SOS-1) and the (VC) condition with $d(\omega, \alpha) < \theta h$. Because of the definition of C and θ^* in (4.8) it suffices to show that the lowest possible bound for an ω satisfying (SOS-1) and the (VC) condition is $d(\omega, \alpha) \geq \theta^*$. We first note that for any ω satisfying (VC), it is necessary that control $i = 1$ is only activated at uneven grid point indices, and control $i = 3$ can only be activated at even grid point indices or at $t = 1$.

We distinguish between the two choices possible at grid index 1. Setting $\omega_{1,3} = 1$ is the first possibility. Because $\alpha_{1,3} = 1 - \theta^*$ this would lead to $d(\omega, \alpha) \geq \theta^* > \theta$. Therefore, the other choice is the only remaining one, and we have to set $\omega_{1,1} = 1$. Similarly, we have to choose between controls $i = 2$ and $i = 3$ at grid index $t = 2$. Now, control 2 has to be chosen as otherwise $\sum_{k=1}^{2} \alpha_{k,2} = \theta^* > \theta$. The two choices made at grid index $t = 1$ and $t = 2$ now set the stage for the remaining $2CM - 6$ grid point indices. Because $\alpha_{t,3} = 0$ for all $t \pmod 2 = 1, t > 1$, we can only choose between controls $i = 1$ and $i = 2$ at uneven grid point indices. Additionally, we have that control 2 cannot be chosen at any uneven grid index because it was chosen at the even grid indices before, beginning with $t = 2$ and has to be chosen at all remaining even grid indices until grid index $2CM - 7$, as otherwise the integral deviation error from the even grid indices and the uneven grid index before that would sum up to θ^*, which would yield $d(\omega, \alpha) > \theta$. Thus, to avoid $d(\omega, \alpha) = \theta^*$ up to and including grid index $2CM - 7$, all controls have to be chosen by alternating between controls $i = 1$ and $i = 2$.

Since all rounding decisions made up to and including grid point index $2CM - 7$ were forced, the control deviations before choosing the last control can be computed from (2.1):

$$\sum_{k=1}^{N} \omega_{k,1} - \alpha_{k,1} = CM - 4 - \theta^*(CM-4) = (CM-4)\left(\frac{1}{CM-2}\right) = \frac{CM-4}{CM-2} < \theta^*,$$
$$\sum_{k=1}^{N} \alpha_{k,2} - \omega_{k,2} = CM - 5 + \theta^* - (CM-5) = \theta^* \text{ and}$$
$$\sum_{k=1}^{N} \alpha_{k,3} = \frac{CM-4}{CM-2} + \frac{1}{CM-2} = \theta^*.$$

Therefore, avoiding $d(\omega, \alpha) = \theta^*$ is only possible at grid index $2CM - 8$ if any choice before that grid index was done differently, which in turn would imply $d(\omega, \alpha) > \theta$. Thus, all possible choices from the first grid index onwards are either forced and lead to an integral control deviation of θ^* at the first or last grid index, or an integral control deviation of at least θ^* was obtained in between. This concludes the proof, as $\theta < \theta^*$ and no binary vanishing constraint feasible control with $d(\omega, \alpha) \leq \theta$ was possible for the constructed α satisfying (SOS-1). ∎

Still, the presented approach can be applied even for $\theta < 1$, as an infeasibility certificate can be generated by returning the number of unmatched vertices from S_p and T. If either number is larger than 0, then by Proposition 3.5 and Lemma 4.1 there exists no control ω adhering to $d(\omega, \alpha) \leq \theta$.

Theorem 4.6 shows that even for $M = 3$ the bound proven in Lemma 4.3 is tight and can be generalized straightforwardly for $M > 3$, as one can simply set $\alpha_{t,i} = 0$ for all $i > 3$. Additionally, we note that removing Assumption 3.1 generalizes the problem, which yields the following corollary.

COROLLARY 4.7. *Let $\epsilon > 0$ and $\theta = 1 - \epsilon$. Then, for $M \geq 3$ there exists an $N \in \mathbb{N}$ and an $\alpha \in [0,1]^{N \times M}$ satisfying* (SOS-1) *such that there exists no $\omega \in \{0,1\}^{N \times M}$ satisfying* (SOS-1) *and the* (VC) *condition with $d(\omega, \alpha) < \theta\overline{\Delta}$.*

REMARK 4.8. *The* (VC) *condition was introduced in [25] to prevent violation of the* (CC) *constraint by the rounded binary control. The condition can also be formulated in a less restrictive way, allowing more rounding options such that Proposition 2.3 still holds, and the constructed rounded binary vanishing constraint control ω still does not constantly violate the* (CC) *constraint. The rounding would then have to obey the following condition instead of* (VC) *for all $i \in [M]$ and $t \in [N]$:*

$$\begin{aligned} \text{(4.10)} \qquad & \omega_{t,i} = 1 \Rightarrow \alpha_{t,i} > 0 \textit{ or given} \\ & \qquad \omega_{1:t,:} \in \{0,1\}^{t \times M} \textit{ the condition } c(y(t), v_i(t)) \geq 0 \textit{ has to hold.} \end{aligned}$$

This condition allows more rounding choices and is only dependent on the previous rounding choices. Unfortunately it does not allow for a smaller θ bound, because the proof of Theorem 4.6 can be modified to accommodate the additional flexibility by adding the constraint $c(y(t), v_3(t)) = -1$, if $\omega_{t-1,1} = 0$. It would then be allowed to switch into control 3 only if control 1 was activated at the previous grid point index. This leads to a control which then either violates the bound on the integrated control deviation or the combinatorial constraint.

Formulation (4.10) *has the additional disadvantage that the rounding choices become sequence dependent and thus the necessary calculations are more complex. The implications of sequence dependency will be inspected in more detail in Section 6.*

Before extending the problem classes in the next section by adding sequence independent costs to the problem we observe that the matching approach is applicable to the rounding problem (CIA) as it is a special case of (CIA-VC). Using the following result from [40], which guarantees existence of a solution for $\theta = \frac{2M-3}{2M-2}$ for (CIA), one can extend the result of Theorem 4.4 for the problem (CIA).

PROPOSITION 4.9 ([40, Corollary 1]). *Let α be a relaxed solution of* (RC)*, then there exists an $\omega \in \{0,1\}^{N\times M}$ satisfying* (SOS-1) *with integrated control deviation*

$$d(\alpha,\omega) \leq \frac{2M-3}{2M-2}\overline{\Delta}.$$

Proposition 4.9 guarantees that a control ω satisfying $d(\omega,\alpha) \leq \theta\overline{\Delta}$ exists. Now, Proposition 3.5 yields the existence of a matching in $G(\alpha,\theta)$, which matches all elements of T and S_p. Thus, we have the following corollary.

COROLLARY 4.10. *Given a relaxed control $\alpha \in [0,1]^{N\times M}$ that solves* (RC) *on an equidistant grid, one can find a binary control $\omega \in \{0,1\}^{N\times M}$ satisfying* (SOS-1) *with $d(\omega,\alpha) \leq \frac{2M-3}{2M-2}\overline{\Delta}$ in $\mathcal{O}(\sqrt{N+M}NM)$.*

Proof. Proposition 4.9 yields the existence of $\omega \in \{0,1\}^{N\times M}$ for $\theta = \frac{2M-3}{2M-2}$, which by Proposition 3.5 induces the existence of a corresponding matching, that can be found by application of the Hopcroft-Karp algorithm on $G(\alpha,\theta)$. The runtime estimation now follows directly from the proof of Theorem 4.4. ∎

Corollary 4.10 shows that computing a binary control with error of at most $\frac{2M-3}{2M-2}$ for (CIA) is possible within a worst case runtime subquadratic in N and M.

We also observe that the graph based approach allows to set the integrated control deviation θ to different values θ_i for the controls $i \in [M]$, thus enabling a simple way to incorporate knowledge of the systems dynamics into the problem without further computational effort. We proceed by showing that solving (CIA) and (CIA-VC) in polynomial time with minimal integral control deviation θ is possible, but necessitates slightly more computational effort. Furthermore, we generalize the problem class to include sequence independent switching costs and present an algorithmic approach.

5. Solving rounding problems with sequence independent switching costs.

Previously, we were interested in the existence of feasible binary solutions and rounding procedures to obtain these solutions within a given integral control deviation θ. We will now extend the scope of our investigation by introducing sequence independent switching costs.

We begin with a special case of sequence independent switching costs. Namely, we show that the pseudo-metric used for the integral control deviation induces a sequence independent switching cost function and that given a relaxed solution α the corresponding rounding problems can be solved up to minimal integral control deviation on an equidistant grid.

LEMMA 5.1. *Let Assumption 3.1 hold, let α be a relaxed solution of* (RC) *or* (VC-RC)*. Let θ be the maximally allowed integral control deviation. Then, the pseudo-metric* (2.1) *induces a sequence independent switching cost function for bottleneck assignment problems on the corresponding bipartite (vanishing constraint) control graph.*

Proof. Let α be a relaxed solution and θ be the maximally allowed integral control deviation. Following Definition 3.4 we construct a bipartite (vanishing constraint) control graph $G(\alpha, \theta) := (V, E)$. Picking any edge $e = (v_{S_{i,p}}, v_t) \in E$ we can define a cost function on the edges

$$C_E(e) := \max\left(-\left(p - \sum_{k=1}^{t} \alpha_{k,i} \right), p - \sum_{k=1}^{t} \alpha_{k,i} \right). \tag{5.1}$$

By associating the edge $e = (v_{S_{i,p}}, v_t)$ with $\omega_{t,i} \equiv 1$ and $\sum_{k=1}^{t} \omega_{t,i} = p$ we have constructed the sequence independent switching cost function

$$\max_{e \in E} C_E(e). \tag{5.2}$$

By Proposition 3.5 existence of a maximal matching $\mathcal{M}$ implies the existence of a feasible control ω with $d(\omega, \alpha) \leq \theta\overline{\Delta}$ and by definition of C_E and the reconstruction of ω from $\mathcal{M}$ we have that $C_E(\mathcal{M}) = d(\omega, \alpha)$. Thus, we have a bottleneck assignment problem on the bipartite (vanishing constraint) control graph G. ∎

The previous Lemma implies that solving problems (ISC-CIA) or (ISC-CIA-VC) with switching cost function (5.2) leads to a control with minimal integral control deviation with respect to α for (CIA) or (CIA-VC). Using known results for the bottleneck assignment problem [12] (weighted bipartite matching with bottleneck cost function) and the estimation on graph sizes from Section 4 we arrive at the following.

Corollary 5.2. *Let Assumption 3.1 hold, let α be a relaxed solution of* (RC) *or* (VC-RC)*. A binary feasible control ω for the problems* (ISC-CIA) *and* (ISC-CIA-VC) *with minimal integral control deviation $d(\omega, \alpha)$ can be computed in*

$$\mathcal{O}(NM\sqrt{(N+M)\log(N+M)}).$$

Proof. Let α be a relaxed solution. Depending on the presence of vanishing constraints we choose θ according to Proposition 4.9 or Lemma 4.3 and construct a bipartite (vanishing control) graph $G(\alpha, \theta) = (V, E)$. Using the sequence independent cost function from Lemma 5.1, we have a bottleneck assignment problem, which can be solved in $\mathcal{O}(|E|\sqrt{|V|\log(|V|)})$, [12]. We obtain the estimations for $|V|$ and $|E|$ from (4.4) and (4.7) for $\theta \leq 1$, which proves the claim. ∎

It will become clear in the remainder of this section that Corollary 5.2 allows to treat the special case of sequence independent switching cost functions, which exhibit the special bottleneck structure in a more efficient way than (ISC-CIA) problems with other sequence independent switching cost functions.

Using the graph construction from Section 3 as well as the results from Section 4 we proceed to show that for any θ one can determine a feasible binary control with minimal switching costs in polynomial runtime, if the switching cost function is sequence independent. For that we use the Hungarian algorithm [39] implemented with Fibonacci heaps [11].

Proposition 5.3 ([11, Section 4]). *Let $G = (V, E)$ be an edge-weighted bipartite graph with edge-weight function $h : E \to \mathbb{R}$. The Hungarian (or Kuhn–Munkres) algorithm implemented with Fibonacci heaps solves the assignment problem (weighted bipartite matching) within $\mathcal{O}(r|E| + r^2 \log(r))$, where r is the size of a maximal matching on G.*

Combining the results of Proposition 5.3 and Section 4 leads to the following Theorem. Note that if the choice of θ, e.g. $\theta < 1$, is infeasible, then an infeasibility certificate is generated by the Hungarian algorithm containing all unassigned vertices from S_p and T.

THEOREM 5.4. *Let α be a solution of* (RC)*, let $\overline{\Delta}$ be the mesh size and $\theta \in \mathbb{R}$ be the maximally allowed integral control deviation. An optimal solution for the class of rounding problems* (ISC-CIA) *and* (ISC-CIA-VC) *for the choice of θ can be computed in*

$$\mathcal{O}\left(N^2 M \lfloor 2\theta + 1 \rfloor + N^2 \log(N)\right). \tag{5.3}$$

The proof of Theorem 5.4 follows along the ideas presented in the previous two sections. Thus, we first construct a bipartite control graph, then transform the switching costs into edge costs and finally estimate the sizes of the vertex and edge sets.

Proof. Let α be a relaxed solution of (RC), where the switching costs are sequence independent. Let $\theta > 0$ be given. Using Definition 3.4, we can construct a bipartite control graph $G(\alpha, \theta) = (V, E)$. Because the cost function $C(\cdot)$ was assumed to be sequence independent, we can assign the corresponding edge costs $h_{t,i} \coloneqq C(\omega_{t,i}, v_i)$ for all $t \in [N], i \in [M]$. Thus, we have constructed an assignment problem on an edge-weighted bipartite graph, which can be solved up to optimality by using the Hungarian Method. Matching the vertices corresponding to elements from S_p first and then the remaining elements of T in combination with the construction of $G(\alpha, \theta)$ yields the feasibility of the calculated control with respect to $d(\omega, \alpha)$ if a matching covering all vertices of S_p and T exists. Existence of such a matching is guaranteed for $\theta \geq 1$ if vanishing constraints are present by Lemma 4.3 and otherwise for $\theta \geq \frac{2M-3}{2M-2}$ by Proposition 4.9. Application of the Hungarian algorithm for the assignment problem guarantees the optimality of the switching cost term.

Thus, estimating the runtime remains. By Proposition 5.3 it remains to estimate r, the size of a maximal matching in $G(\alpha, \theta)$. Following the estimations made in the proof of Theorem 4.4, especially inequalities (4.4) and (4.7), we get $r = \min\{N, N + M \lfloor \theta + 1 \rfloor\} = N$ and $|E| \leq NM \lfloor 2\theta + 1 \rfloor$. Insertion into the runtime estimation of the Hungarian algorithm now yields the claim.

The correctness and runtime claim for the rounding problem (ISC-CIA-VC) follow analogously by using $G_{VC}(\alpha, \theta)$ from Definition 3.4 instead of $G(\alpha, \theta)$. ∎

Thus, rounding problems with sequence independent switching costs (ISC-CIA) and (ISC-CIA-VC) can be solved up to optimality in subcubic or, in the case of bottleneck functions, even subquadratic time. Therefore, they are polynomially solvable, which distinguishes them from problems with sequence dependent switching costs, to be considered in the Section 6. Before investigating this next extension of rounding problems, some remarks are in order.

REMARK 5.5.

- a) *The construction of the bipartite graphs from Definition 3.4 and the proof of Theorem 5.4 as well as the statements from previous sections yield that sequence independent constraints and problems with both sequence independent switching costs and sequence independent constraints can be solved by constructing the appropriate bipartite graph and using an algorithm for weighted assignment problems, e.g. the Hungarian algorithm.*
- b) *The parameter θ allows a tradeoff between approximation quality and switching costs, because choosing larger θ values usually leads to smaller switching costs as more costly controls can be avoided more often.*
- c) *Similar to the previous section, θ can also be chosen as a vector in $\mathbb{R}^M$, which allows different integral control deviations for the controls and thus leads to possibly better solutions with respect to the switching costs.*

6. MIOCPs and rounding problems with sequence-dependent switching costs. This section addresses the most general class of rounding problems investigated in this article, which consist of the rounding problems with sequence dependent switching costs.

By reducing the well-known strongly $\mathcal{NP}$-hard symmetric TSP problem, see e.g. [28, Theorem 15.43] or [13, Problem ND22], to the decision problem devised from (DSC-CIA) we show that finding a binary feasible control is a strongly $\mathcal{NP}$-complete problem when sequence-dependent switching costs are present. This holds even for the smallest integral control deviation guaranteeing existence of a binary control, $\theta = \frac{2M-3}{2M-2}$ from Proposition 4.9. This affirms that there exists no (fully) polynomial-time approximation scheme for (DSC-CIA). Additionally, we show that there does not even exist a β-approximation algorithm for any $\beta \geq 1$ for (DSC-CIA) as this would imply solvability of the Hamiltonian circuit problem [13, Problem GT37] in polynomial time. We highlight that these results are tighter than just proving $\mathcal{NP}$-hardness and inapproximability of (MSCP-CC), as they show that computing a binary feasible solution within vicinity of a relaxed feasible solution is already both $\mathcal{NP}$-hard and inapproximable.

We proceed by formulating the optimization problem DSC-CIA as a decision problem.

Definition 6.1. *(DEC-DSC-CIA) Let Assumption 3.1 hold. Let $\alpha \in [0,1]^{N\times M}$ be a solution of* (RC)*, let $\overline{\Delta}$ be the mesh size, let $\theta \in \mathbb{R}$ be the maximally allowed integral control deviation and let $K \in \mathbb{Q}$ be the cost threshold for C. The decision problem* (DEC-DSC-CIA) *is now to decide whether there exists a feasible binary control ω with $d(\omega,\alpha) \leq \theta\overline{\Delta}$ and switching costs $C(\omega) \leq K$.*

Note that if the decision problem (DEC-DSC-CIA) could be solved in polynomial time, the corresponding optimization problem (DSC-CIA) could be solved in polynomial time as well as one can iterate over the threshold variable K to find a feasible binary control with minimal switching costs. Unfortunately, the well-known symmetric traveling salesman problem can be reduced to this decision problem.

Theorem 6.2. *Let Assumption 3.1 hold. Then* (DEC-DSC-CIA) *is strongly $\mathcal{NP}$-hard for $M > 2$ with $\theta = \frac{2M-3}{2M-2}$.*

The following proof is divided into two parts. We first devise the reduction and show that it is strongly polynomial in the input size of the TSP instance. This is done by constructing a relaxed solution α adhering to the (SOS-1) constraints and a cost function, which inherits the cost values from the TSP instance. Afterwards we show correctness of the reduction and that if problem (DEC-DSC-CIA) could be decided in polynomial time then TSP could be solved in polynomial time as well.

Because of Assumption 3.1 the mesh size is factored out and can constantly assumed to be 1 in the proof. It is therefore left out for ease of presentation.

Proof. We first note that $\theta = \frac{2M-3}{2M-2}$ can be encoded polynomially in $\mathcal{O}(M)$. Thus, membership of the decision problem (DEC-DSC-CIA) in $\mathcal{NP}$ is confirmed, as verifying the feasibility of a solution is possible in polynomial time with respect to the input sizes N and M.

Now consider an undirected complete graph $G = (V,E)$ with $V \coloneqq \{v_1,\ldots,v_n\}$ and a cost function $h : E \to \mathbb{N}$ as an instance (G,h) of the symmetric TSP problem. Let $K \in \mathbb{Q}$ be the cost threshold value of the symmetric TSP. We define $W \coloneqq n \max_{i,j\in[n],i\neq j} h_{i,j} + 1$ and observe immediately that any non-trivial TSP instance admits an optimal tour R with costs smaller than W. Hence, w.l.o.g. we can assume that $K < W$, which will be used in the following construction.

Besides the decision value K, the (DEC-DSC-CIA) problem needs matrix $\alpha \in [0,1]^{N\times M}$, a cost function C_h and the maximally allowed integral deviation θ as inputs, which we construct within polynomial time and space bounds from the input (G,h) in the following.

We set the number of grid points as $N = n+1$ and the number of controls as $M = n$. Setting $\theta \coloneqq \frac{2M-3}{2M-2} = \frac{2n-3}{2n-2}$ allows an encoding in space $\mathcal{O}(n^3)$. A matrix α is defined for all $t \in [n+1]$ and $i \in [n]$ as follows:

$$\alpha_{t,i} \coloneqq \begin{cases} \frac{1}{M} & \text{if } t < n+1, \\ 1 & \text{if } t \equiv n+1 \text{ and } i \equiv 1, \\ 0 & \text{else.} \end{cases} \tag{6.1}$$

It holds by construction that α satisfies the (SOS-1) condition and that the matrix α of size $(n+1)\times(n)$ can be constructed and encoded in time and space within $\mathcal{O}(n^3)$. The sequence dependent cost function C_h can now be derived from the TSP input for all $t \in [n+1]$ and $i \in [n]$:

$$C_h(t,i,\omega_{1:t-1,:}) \coloneqq \begin{cases} 0 & \text{if } t = 1 \text{ and } i = 1, \\ h_{j,i} & \text{if } \omega_{t-1,j} = 1, \\ W & \text{if } \displaystyle\sum_{k=1}^{t-1} \omega_{k,i} = 1 \text{ and either } i \neq 1 \text{ or } i \equiv 1 \text{ and } t \neq n+1, \\ W & \text{else.} \end{cases} \tag{6.2}$$

Note that given $\omega_{1:t-1,:}$ every evaluation of the cost function is possible in $\mathcal{O}(n^2)$ for any $t \in [n+1]$ and $i \in [n]$.

Thus, the construction of a (DSC-CIA) instance from a symmetric TSP instance (G,h) is possible in time and space polynomial to the input size of (G,h) and if the input size of (G,h) is bounded in space by a polynomial in n, then the construction of (DSC-CIA) is also bounded in space by a polynomial in n.

It remains to show that, given a decision threshold K, the TSP decision problem (G,h) admits a tour R with costs $h(R) \leq K$ if and only if the constructed (DSC-CIA) instance admits a solution ω with switching costs $C_h(\omega) \leq K$.

Assume first that there exists a tour R such that

$$h(R) = X \leq K$$

holds. As any complete tour is a Hamiltonian cycle, we can w.l.o.g. assume that it begins and ends at vertex v_1. Setting $\omega_{t,i} = 1$ if vertex v_i was traversed at the t-th position in R leads to an $\omega \in \{0,1\}^{N\times M}$ matrix adhering to the (SOS-1) constraint with the property that $\omega_{1,1} = \omega_{n+1,1} = 1$. By construction of C_h in (6.2) we have

$$C_h(\omega) = h(R) \leq K.$$

It remains to show that $d(\omega,\alpha) \leq \theta$ holds as well. By (6.1) we have that all $\alpha_{t,i}$ are equal for all $t \in [n]$ and $i \in [n]$. Because $\alpha_{t,i} \in [0,1]$, $i \in [M]$, $t \in [N]$ we only have to consider the first grid point as well as the last two grid points for the purpose of verifying $d(\omega,\alpha) \leq \theta$.

For the first grid point and $i = 1$ it holds that

$$\alpha_{1,1} + \theta = \frac{1}{M} + \frac{2M-3}{2M-2} = \frac{2M^2 - M - 2}{2M^2 - 2M} \overset{M>2}{>} 1. \tag{6.3}$$

Thus, setting $\omega_{1,1} = 1$ does not violate $d(\omega, \alpha) \leq \theta$ in $t = 1$. As R was a tour and by construction of ω all controls have been activated exactly once at grid point n, thus,

$$\sum_{t=1}^{n} \alpha_{t,i} - \theta = 1 - \frac{2M-3}{2M-2} = \frac{1}{2M-2} > 0 \text{ for all } i \in [M] \tag{6.4}$$

is not violated. Additionally no control $i \in [M]$ besides the first one can be activated at grid point $n + 1$ as

$$\sum_{t=1}^{n+1} \alpha_{t,i} + \theta = 1 + \frac{2M-3}{2M-2} < 2 \text{ for all } i \in [M]\backslash\{1\}. \tag{6.5}$$

Thus, $d(\omega, \alpha) \leq \theta$ holds and ω is a solution to the decision problem (DEC-DSC-CIA) for K.

For the reverse implication we assume that ω is a solution of (DSC-CIA) with switching costs

$$C_h(\omega) = X \leq K. \tag{6.6}$$

Note that the first control must be activated at $t = 1$ as otherwise $C_h(\omega) \geq W > K$ holds. By construction of α every control must be activated exactly once at grid point n as

$$\left\lceil \sum_{t=1}^{n} \alpha - \theta \right\rceil = \left\lceil \frac{2M-2}{2M-2} - \frac{2M-3}{2M-2} \right\rceil = 1. \tag{6.7}$$

Furthermore, we have by construction that only the first control can be activated for the second time at grid point $n+1$. Thus, the first control was activated at grid points 1 and $n + 1$ and traversing the vertices $v_1, \ldots, v_n$ in the order in which the controls were activated along the grid yields a TSP tour R with costs $h(R) = C_h(\omega) \leq K$. Hence, the TSP instance (G, h) admits a solution with value less or equal to K if and only if the (DSC-CIA) instance with $\theta = \frac{2M-3}{2M-2}$ and switching cost function C_h admits a solution ω with $C_h(\omega) \leq K$. ∎

Slight modification of this proof yields an inapproximability result as any approximation algorithm for (DSC-CIA) would solve the Hamiltonian cycle problem, see [23] or [13, Problem GT37].

Corollary 6.3. *Let Assumption 3.1 hold and let $M \geq 3$. Unless $\mathcal{P} = \mathcal{NP}$, there exists no β-approximation algorithm for any $\beta \geq 1$ with polynomial runtime for the rounding problem* (DSC-CIA) *with $\theta = \frac{2M-3}{2M-2}$.*

Proof. We can change the cost function for the TSP-instance such that any algorithm approximating (DSC-CIA) for a fixed $\beta \geq 1$ solves the Hamiltonian cycle problem. For this let $G := (V, E)$ be an instance of the Hamiltonian cycle problem. Then, let the undirected complete graph $\overline{G} = (V, \overline{E})$ be the corresponding TSP-instance with edge costs

$$c_{i,j} := \begin{cases} 1 & \text{if } (i,\, j) \in E, \\ \beta n + 1 & \text{otherwise.} \end{cases} \tag{6.8}$$

We immediately note that G is Hamiltonian if and only if $\overline{G}$ admits a tour with cost of n. Furthermore, by the construction of C in the proof of Theorem 6.2 any solution with cost strictly less than $\beta n + n + 1$ must consist of a sequence of controls corresponding to edges forming a tour in $\overline{G}$ of cost n and thus shows that G is Hamiltonian. As $\beta \geq 1$ was arbitrary but fixed, this implies that approximating (DSC-CIA) in polynomial runtime for any $\beta \geq 1$ would decide the Hamiltonian cycle problem in polynomial time, which contradicts the $\mathcal{NP}$-hardness of the Hamiltonian cycle problem [23]. ◻

Theorem 6.2 and Corollary 6.3 show that (DSC-CIA) is $\mathcal{NP}$-hard and even hard to approximate for $\theta = \frac{2M-3}{2M-2}$. Examining the definition of the cost function (6.2) one immediately notes that the results extend to all choices of $\theta \geq \frac{2M-3}{2M-2}$. This is natural, as polynomial decidability of an instance with larger θ would imply polynomial decidability for instance with smaller θ in the sense that either a solution can be computed in polynomial time or an infeasibility certificate is generated.

COROLLARY 6.4. *Let Assumption 3.1 hold. Then,* (DEC-DSC-CIA) *is strongly $\mathcal{NP}$-hard for $M > 2$ with $\theta \geq \frac{2M-3}{2M-2}$.*

One also observes that the results from Theorem 6.2 and Corollaries 6.3 and 6.4 extend to (DSC-CIA-VC). A proof follows analogously by setting $\theta = 1$ as the statements made in equations (6.3)–(6.7) still hold and the stated consequences are still valid by construction of α and the (SOS-1) constraint.

COROLLARY 6.5. *Let Assumption 3.1 hold. Then the decision problem of* (DSC-CIA-VC) *is strongly $\mathcal{NP}$-hard for $M > 2$ with $\theta \geq 1$.*

The previous Theorem 6.2 also implies that the runtime results proven in [4] are asymptotically optimal for the problem class of (DSC-CIA) as $M, \theta \ll N$ is almost certainly the case for optimal control problems.

Furthermore, we have for the original problem (MSCP-CC) that computing a binary control ω that is optimal with respect to sequence dependent switching costs and within the smallest radius around a relaxed control α for which a solution exists (Proposition 4.9 without (CC) constraints and Lemma 4.3 in combination with Theorem 4.6 otherwise) is $\mathcal{NP}$-hard and also not approximable within polynomial time.

7. Conclusion. On equidistant grids the presented graph based matching approach, Rounding via Matching (RVM), derives a feasible control for the rounding problems associated with the decomposition approach for MIOCPs with state-control inequality constraints dependent on the integer control. The proposed RVM method solves the rounding problem within the tightest possible bound for the (CIA-VC) problem, which was shown to be $\overline{\Delta}$. Additionally, the presented RVM approach is able to handle the rounding problems from MIOCPs without inequality constraints and solves them up to a supplied bound or to optimality with respect to the integrated control deviation. In comparison to prior investigated algorithms it either holds that the approximation quality of RVM is superior or that the computational costs are smaller. The modeling and solving capabilities of the decomposition approximation approach were extended by introducing the concept of sequence independent switching costs, which were shown to be solvable in $\mathcal{O}\left(N^2 M \lfloor \theta + 1 \rfloor + N^2 \log(N)\right)$. Finally, we showed that introduction of sequence dependent switching costs leads to a strongly $\mathcal{NP}$-hard and, unless $\mathcal{P} = \mathcal{NP}$, polynomially inapproximable problem.

REFERENCES

[1] H. Bast, K. Mehlhorn, G. Schäfer, and H. Tamaki, *Matching Algorithms Are Fast in Sparse Random Graphs*, in STACS 2004, V. Diekert and M. Habib, eds., Berlin, Heidelberg, 2004, Springer, pp. 81–92, https://doi.org/10.1007/978-3-540-24749-4_8.

[2] B. T. Baumrucker, J. G. Renfro, and L. T. Biegler, *MPEC problem formulations and solution strategies with chemical engineering applications*, Computers & Chemical Engineering, 32 (2008), pp. 2903 – 2913, https://doi.org/10.1016/j.compchemeng.2008.02.010.

[3] F. Bestehorn, C. Hansknecht, C. Kirches, and P. Manns, *A switching cost aware rounding method for relaxations of mixed-integer optimal control problems*, in 2019 IEEE 58th Conference on Decision and Control (CDC), 2019, pp. 7134–7139, https://doi.org/10.1109/CDC40024.2019.9030063.

[4] F. Bestehorn, C. Hansknecht, C. Kirches, and P. Manns, *Mixed-Integer Optimal Control Problems with switching costs: A shortest path approach*, Optimization Online Preprint 7630, (2020), http://www.optimization-online.org/DB_HTML/2020/02/7630.html. (submitted to Mathematical Programming Series B).

[5] F. Bestehorn and C. Kirches, *The integrated control deviation of a combinatorial framework for mixed-integer optimal control problems with vanishing constraints*, in Proceedings in Applied Mathematics and Mechanics, 2020. (submitted).

[6] H. G. Bock, D. H. Cebulla, C. Kirches, and A. Potschka, *Mixed-Integer Optimal Control for Multimodal Chromatography*, Optimization Online Preprint 7856, (2020), http://www.optimization-online.org/DB_HTML/2020/06/7856.html. (submitted to Computers & Chemical Engineering).

[7] A. Bürger, M. Bohlayer, S. Hoffmann, A. Altmann-Dieses, M. Braun, and M. Diehl, *A whole-year simulation study on nonlinear mixed-integer model predictive control for a thermal energy supply system with multi-use components*, Applied Energy, 258 (2019), p. 114064, https://doi.org/10.1016/j.apenergy.2019.114064.

[8] A. De Marchi, *On the Mixed-Integer Linear-Quadratic Optimal Control With Switching Cost*, IEEE Control Systems Letters, 3 (2019), pp. 990–995, https://doi.org/10.1109/LCSYS.2019.2920425.

[9] X. C. Ding, Y. Wardi, D. Taylor, and M. Egerstedt, *Optimization of Switched-Mode Systems with Switching Costs*, in 2008 American Control Conference, 2008, pp. 3965–3970, https://doi.org/10.1109/ACC.2008.4587112.

[10] M. Egerstedt, Y. Wardi, and H. Axelsson, *Transition-Time Optimization for Switched-Mode Dynamical Systems*, IEEE Transactions on Automatic Control, 51 (2006), pp. 110–115, https://doi.org/10.1109/TAC.2005.861711.

[11] M. L. Fredman and R. E. Tarjan, *Fibonacci Heaps and Their Uses in Improved Network Optimization Algorithms*, Journal of the Association for Computing Machinery, 34 (1987), p. 596–615, https://doi.org/10.1145/28869.28874.

[12] H. N. Gabow and R. E. Tarjan, *Algorithms for Two Bottleneck Optimization Problems*, Journal of Algorithms, 9 (1988), pp. 411 – 417, http://www.sciencedirect.com/science/article/pii/0196677488900314.

[13] M. R. Garey and D. S. Johnson, *Computers and Intractability; A Guide to the Theory of NP-Completeness*, W. H. Freeman & Co., USA, 1990.

[14] 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 (2005), pp. 1–18, https://doi.org/10.1002/oca.751.

[15] M. Gerdts, *A variable time transformation method for mixed-integer optimal control problems*, Optimal Control Applications and Methods, 27 (2006), pp. 169–182, https://onlinelibrary.wiley.com/doi/abs/10.1002/oca.778.

[16] M. Hahn, C. Kirches, P. Manns, S. Sager, and C. Zeile, *Decomposition and approximation for PDE-constrained mixed-integer optimal control*, in SPP1962 Special Issue, Michael Hintermüller et al., ed., Birkhäuser, 2019. (accepted).

[17] F. M. Hante and S. Sager, *Relaxation methods for mixed-integer optimal control of partial differential equations*, Computational Optimization and Applications, 55 (2013), pp. 197–225, https://doi.org/10.1007/s10589-012-9518-3.

[18] C. A. Henao and C. T. Maravelias, *Surrogate-based superstructure optimization framework*, AIChE Journal, 57 (2011), pp. 1216–1232, https://aiche.onlinelibrary.wiley.com/doi/abs/10.1002/aic.12341.

[19] T. Hirmajer, E. Balsa-Canto, and J. R. Banga, *Mixed-integer non-linear optimal control in systems biology and biotechnology: numerical methods and a software toolbox*, IFAC Proceedings Volumes, 43 (2010), pp. 314–319, https://doi.org/https://doi.org/10.3182/20100705-3-BE-2011.00052. 9th IFAC Symposium on Dynamics and Control of Process Systems.

[20] A. J. Hoffman, *Generalization of a theorem of König*, Journal of the Washington Academy of Science, 46 (1956), pp. 211–212.

[21] J. E. Hopcroft and R. M. Karp, *An $n^{5/2}$ Algorithm for Maximum Matchings in Bipartite Graphs*, SIAM Journal on Computing, 2 (1973), pp. 225–231, https://doi.org/10.1137/0202019.

[22] M. Jung, *Relaxations and Approximations for Mixed-Integer Optimal Control*, Dissertation, Heidelberg University, 2013, https://doi.org/10.11588/heidok.00016036.

[23] R. M. Karp, *Reducibility among Combinatorial Problems*, Springer US, Boston, MA, 1972, pp. 85–103, https://doi.org/10.1007/978-1-4684-2001-2_9.

[24] C. Kirches, E. A. Kostina, A. Meyer, and M. Schlöder, *Generation of Optimal Walking-Like Motions Using Dynamic Models with Switches, Switch Costs, and State Jumps*, in IEEE 58th Conference on Decision and Control (CDC), 2019, pp. 1538–1543, https://doi.org/10.1109/CDC40024.2019.9029577.

[25] C. Kirches, F. Lenders, and P. Manns, *Approximation Properties and Tight Bounds for Constrained Mixed-Integer Optimal Control*, SIAM Journal on Control and Optimization, 58 (2020), pp. 1371–1402, https://doi.org/10.1137/18M1182917.

[26] 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 (2010), pp. 137–153, https://doi.org/10.1002/oca.892.

[27] M. Koller and R. Hofmann, *Mixed-Integer Linear Programming Formulation of Combined Heat and Power Units for the Unit Commitment Problem*, Journal of Sustainable Development of Energy, Water and Environment Systems, 6 (2018), https://doi.org/10.13044/j.sdewes.d6.0207.

[28] B. H. Korte and J. Vygen, *Combinatorial Optimization: Theory and Algorithms*, Springer Publishing Company, Incorporated, New York, NY, 5th ed., 2012, https://doi.org/10.1007/978-3-642-24488-9.

[29] R. Kumar, M. J. Wenzel, M. N. ElBsat, M. J. Risbeck, K. H. Drees, and V. M. Zavala, *Dual Dynamic Programming for Multi-Scale Mixed-Integer MPC*, 2020, https://arxiv.org/abs/2007.10149.

[30] P. Manns, *Approximation Properties of Sum-Up Rounding*, Dissertation, Technische Universität Carolo-Wilhelmina zu Braunschweig, 2019, https://doi.org/10.24355/dbbs.084-202001071053-0.

[31] P. Manns and C. Kirches, *Multi-dimensional sum-up rounding for elliptic control systems*, SIAM Journal on Numerical Analysis, (2020), https://spp1962.wias-berlin.de/preprints/080r.pdf. (accepted, DOI assignment pending).

[32] P. Manns, C. Kirches, and F. Lenders, *A Linear Bound on the Integrality Gap for Sum-up Rounding in the Presence of Vanishing Constraints*. Preprint SPP1962-077r, 4 2020, https://spp1962.wias-berlin.de/preprints/077r.pdf.

[33] S. Sager, *Numerical methods for mixed-integer optimal control problems*, Der Andere Verlag Tönning, Lübeck, Marburg, 2005, https://mathopt.de/PUBLICATIONS/Sager2005.pdf.

[34] S. Sager, H.-G. Bock, and M. Diehl, *The Integer Approximation Error in Mixed-Integer Optimal Control*, Mathematical Programming, Series A, 133 (2012), pp. 1–23, https://doi.org/10.1007/s10107-010-0405-3.

[35] S. Sager, M. Jung, and C. Kirches, *Combinatorial Integral Approximation*, Mathematical Methods of Operations Research, 73 (2011), pp. 363–380, https://doi.org/10.1007/s00186-011-0355-4.

[36] A. Schrijver, *Combinatorial optimization. Polyhedra and efficiency (3 volumes).*, vol. 24, Berlin: Springer, 2003.

[37] B. Stellato, S. Ober-Blöbaum, and P. J. Goulart, *Optimal control of switching times in switched linear systems*, in 2016 IEEE 55th Conference on Decision and Control (CDC), 2016, pp. 7228–7233, https://doi.org/10.1109/CDC.2016.7799384.

[38] B. Stellato, S. Ober-Blöbaum, and P. J. Goulart, *Second-order switching time optimization for switched dynamical systems*, IEEE Transactions on Automatic Control, 62 (2017), pp. 5407–5414, https://doi.org/10.1109/TAC.2017.2697681.

[39] R. E. Tarjan, *Data Structures and Network Algorithms*, Society for Industrial and Applied Mathematics, USA, 1983, https://doi.org/10.1137/1.9781611970265.

[40] C. Zeile, N. Robuschi, and S. Sager, *Mixed-Integer Optimal Control under Minimum Dwell Time Constraints*, Mathematical Programming B, (2020), https://doi.org/10.1007/s10107-020-01533-x.