



# On mixed-integer optimal control with constrained total variation of the integer control

Sebastian Sager<sup>1</sup> · Clemens Zeile<sup>1</sup>

Received: 21 October 2019 / Accepted: 5 November 2020 / Published online: 7 December 2020  
© The Author(s) 2020

## Abstract

The *combinatorial integral approximation* (CIA) decomposition suggests solving mixed-integer optimal control problems by solving one continuous nonlinear control problem and one mixed-integer linear program (MILP). Unrealistic frequent switching can be avoided by adding a constraint on the total variation to the MILP. Within this work, we present a fast heuristic way to solve this CIA problem and investigate in which situations optimality of the constructed feasible solution is guaranteed. In the second part of this article, we show tight bounds on the integrality gap between a relaxed continuous control trajectory and an integer feasible one in the case of two controls. Finally, we present numerical experiments to highlight the proposed algorithm's advantages in terms of run time and solution quality.

**Keywords** Mixed-integer linear programming · Optimal control · Discrete approximations · Switched dynamic systems · Approximation methods and heuristics

**Mathematics Subject Classification** 49M20 · 49M27 · 90B35 · 90C11 · 90C59

## 1 Introduction

Mixed-Integer Optimal Control has been established as a useful tool for modeling real-world problems [6, 10, 22, 33]. In practice, only an optimal control policy is realistic that avoids frequent switching between the system modes. However, it remains an open research question as to how switching costs or a limited number of switches can be efficiently incorporated into the optimization problem.

---

✉ Clemens Zeile  
clemens.zeile@ovgu.de

Sebastian Sager  
sager@ovgu.de

<sup>1</sup> Department of Mathematics, Otto-von-Guericke-Universität Magdeburg, Magdeburg 39106, Germany

In this article, we follow a *first-discretize-then-optimize* approach because in contrast to *indirect methods* and *dynamic programming*, a more generic problem class can be solved for which efficient numerical methods, in particular, a decomposition approach used in this article, are available. By this approach, the control problem is discretized via, e.g., Direct Multiple Shooting [5] or Direct Collocation [29], which leads to a mixed-integer nonlinear program (MINLP). This problem class is NP-hard in general, so that it has been proposed to reduce complexity by solving first the relaxed problem with dropped integrality constraint, which is a nonlinear program (NLP), before approximating relaxed controls in a second step with binary controls as part of a mixed-integer linear program (MILP). The second problem is usually referred to as *combinatorial integral approximation* (CIA) problem [27], whereas the whole algorithm is called CIA decomposition [31]. It is common to use the fast Sum-Up Rounding (SUR) heuristic [24] to find a feasible approximative solution for the CIA problem, so that this second step is also named rounding. However, standard SUR does not consider time-coupled combinatorial constraints, which is why the use of the CIA problem is necessary in this case. The *variable time transformation* [2, 11, 21] method also avoids solving a MINLP by assuming a given sequence of the system modes so that only their durations as part of an NLP have to be computed. Recently, De Marchi extended this approach by including switching costs with sparse optimization methods [9]. As part of the CIA decomposition, Sager [24], Kirches [16], and Rieck [20] proposed to add penalty terms to the control problem objective in order to account for switching costs or to reduce the number of switches between active controls. Recently, Kirches et al. [17] investigated this approach for the setting with implicit switches and where jumps of the differential state values may occur. One issue of the penalization approach is the appropriate choice of the penalty factor since heavy penalization can in some instances [16] attract solutions involving frequent switching. Bestehorn et al. [3, 4] presented an idea to incorporate switching costs into the problem by fixing a small control deviation tolerance in the rounding problem and minimizing switching costs subject to this deviation tolerance.

This work builds on [27], where it has been proposed to solve the CIA problem with constraints that limit the total variation (TV) of the integer control. By this, the original optimal control objective remains untouched, and we require only the number of switches to be less than a desired threshold. Our idea is to use generally a Branch & Bound (BNB) scheme [7] for solving this CIA problem, but, as we will show in this article, this can be replaced for certain instances by a sequence of rounding scheme evaluations. Furthermore, we evaluate upper bounds on the CIA objective subject to these discrete TV constraints. It has been shown that the integer approximation error, i.e., the difference of relaxed optimal control objective value and CIA rounded objective value, can be driven to zero under mild conditions and if the grid length is driven to zero [19,

27]. This result does not hold anymore if discrete TV constraints are included. Still, we expect our approach to yield feasible solutions of the MIOCP with an a priori integer gap.

### 1.1 Contribution

To accelerate the CIA solving process, we propose the Maximum Dwell Rounding (MDR) scheme, which is a fast rounding heuristic. It is based on the idea to activate a chosen control mode as long as possible without violating a desired integrality gap  $\theta$ , i.e., the accumulated deviation of relaxed and binary controls, and then perform this with the next promising mode. We apply it iteratively as part of the Adaptive Maximum Dwell Rounding (AMDR) algorithm for finding binary controls that satisfy time-coupled combinatorial constraints such as a TV bound and derive optimality conditions of the obtained binary control function with respect to the CIA problem. Based on this scheme, we prove the tightest possible upper bound on the integrality gap for equidistant discretization and the case of two binary controls, which reads

$$\theta \leq \frac{N + \sigma_{\max} + 1}{3 + 2\sigma_{\max}} \bar{\Delta},$$

where  $N$  denotes the number of intervals,  $\sigma_{\max}$  the TV bound and  $\bar{\Delta}$  the maximum grid length. We are going to establish further bounds for the situation of non-equidistant grids or more than two binary controls.

### 1.2 Outline

We give a problem definition of the MIOCP of interest in Sect. 2 and describe the proposed CIA decomposition algorithm with (CIA) as subproblem in Sect. 3. Next, we introduce auxiliary CIA problems and derive a lower bound for these problems in Sect. 4. We define the MDR scheme in Sect. 5 and show its usefulness with respect to solving the CIA problem subject to discrete TV constraints. We continue by analyzing the worst-case integrality gap for  $n_\omega = 2$  in Sect. 6, respectively for  $n_\omega > 2$  in Sect. 7. Finally, we present numerical experiments in Sect. 8 and conclusions in Sect. 9.

### 1.3 Notations

Let  $[n] := \{1, \dots, n\}$ ,  $[n]_0 := \{0\} \cup [n]$ , for  $n \in \mathbb{N}$ . We use Gauss' bracket notation, i.e.  $\lfloor x \rfloor := \max\{k \in \mathbb{Z} \mid k \leq x\}$ ,  $x \in \mathbb{R}$ , and analogously for  $\lceil x \rceil$ . We indicate by  $\lceil x \rceil^{0.5}$  the rounding up of  $x \in \mathbb{R}$  to the next multiple of 0.5:

$$\lceil x \rceil^{0.5} := \min\{y \mid y = n \cdot 0.5, n \in \mathbb{N}, y \geq x\}.$$

We write "for a.e.  $t \in \mathcal{T}$ " for abbreviating for all  $t \in \mathcal{T} \subset \mathbb{R}$  except on a set of measure zero. Moreover, we write *control* to abbreviate a control realization  $\omega_i(\cdot), i \in [n]$ , of a control function  $\omega(\cdot) = (\omega_1(\cdot), \dots, \omega_n(\cdot))^T$ .

## 2 Mixed-integer optimal control problem

Mixed-integer optimal control problems (MIOCPs) can be equivalently reformulated into problems with affinely entering binary controls in the right-hand side of the ordinary differential equation via the (*partial*) *outer convexification* method, see [24] for further details. Therefore, we declare this reformulated MIOCP as the problem of interest and provide the corresponding definitions in this section.

We consider problems on a given time horizon  $\mathcal{T} := [t_0, t_f] \subset \mathbb{R}$ . Throughout this paper, we assume a problem involving  $n_\omega \geq 2$  binary control realizations. We introduce the binary control function after defining the TV of a function and its associated space.

**Definition 1 (Total Variation of a function and BV space)** The TV of a function  $\omega : \mathcal{T} \rightarrow \mathbb{R}^{n_\omega}$  is defined to be the quantity

$$TV(\omega) := \sup_{P \in \mathcal{P}} \left\{ \frac{1}{2} \sum_{i \in [n_\omega]} \sum_{j \in [n_p]} |\omega_i(t_j) - \omega_i(t_{j-1})| \right\}, \tag{2.1}$$

where  $P = (t_0, \dots, t_{n_p})$  is a partition out of the set of all partitions  $\mathcal{P}$  of the interval  $\mathcal{T}$  and  $n_p$  denotes the partition specific number of time points.

We group the functions  $\omega$  with finite TV into the space of *bounded variation BV*:

$$BV(\mathcal{T}, \mathbb{R}^{n_\omega}) := \{ \omega : \mathcal{T} \rightarrow \mathbb{R}^{n_\omega} \mid TV(\omega) < \infty \}.$$

**Definition 2 (Binary  $\omega$  and relaxed control functions  $\alpha$ )** Let the vector of binary controls  $\omega$  on the simplex and its corresponding vector of relaxed controls  $\alpha$  defined by their function space domains

$$\begin{aligned} \Omega &:= \left\{ \omega \in BV(\mathcal{T}, \{0, 1\}^{n_\omega}) \mid \sum_{i \in [n_\omega]} \omega_i(t) = 1, \text{ for a.e. } t \in \mathcal{T} \right\}, \\ \mathcal{A} &:= \left\{ \alpha \in BV(\mathcal{T}, [0, 1]^{n_\omega}) \mid \sum_{i \in [n_\omega]} \alpha_i(t) = 1, \text{ for a.e. } t \in \mathcal{T} \right\}. \end{aligned}$$

**Definition 3 (Problems (MIOCP) and (OCP))** Let a maximum number of switches  $\sigma_{\max} \in \mathbb{N}$  be given. We refer to the following general problem class as (**MIOCP**)

$$\begin{aligned} & \min_{x, \omega \in \Omega} \Phi(x(t_f)) \\ \text{s. t.} \quad & \dot{x} = f_0(x(t)) + \sum_{i \in [n_\omega]} \omega_i(t) f_i(x(t)), \quad \text{for a.e. } t \in \mathcal{T}, \end{aligned} \tag{2.2}$$

$$x(t_0) = x_0, \tag{2.3}$$

$$TV(\omega) \leq \sigma_{\max} \tag{2.4}$$

We minimize a Mayer term functional  $\Phi \in C^1(\mathbb{R}^{n_x}, \mathbb{R})$  over the binary controls  $\omega_i$  and differential states  $x \in W^{1,\infty}(\mathcal{T}, \mathbb{R}^{n_x})$  with fixed initial values  $x_0 \in \mathbb{R}^{n_x}$ . Constraint (2.2) expresses the dynamical system as a switched system in *partial outer convexified* form, i.e., as a sum of a drift term  $f_0$  and control specific functions  $f_i$  both out of  $C^0(\mathbb{R}^{n_x}, \mathbb{R}^{n_x})$ . We assume that there exists a solution  $x$  for the above problem; for this we may assume that a uniform Lipschitz estimate on  $f_0$  and  $f_i$  exists so that the theorem by Picard–Lindelöf is applicable. We limit the number of switches between active modes to be at most  $\sigma_{\max}$  in the TV constraint (2.4). Finally, we define (OCP) as the canonical relaxation of problem (MIOCP) where we optimize over  $\alpha \in \mathcal{A}$  instead of  $\omega \in \Omega$ .

**Remark 1** When we write that control  $i$  is *active*, we indicate that  $\omega_i(t) = 1$ . In fact, we count the switches in (2.4), respectively in (2.1), twice since we sum up the control that has just been deactivated with the one that has just been activated. That explains the factor of one-half in (2.1). We remark that our study assumes no mode-specific switching limits, but we may also impose them by splitting up (2.4) into  $n_\omega$  inequalities and by dropping the first sum in (2.1).

Without loss of generality, we omit further constraints and continuous control functions  $u \in L^\infty(\mathcal{T}, \mathbb{R}^{n_u})$  in our problem definition. See [28] how to cope with these and further extensions.

### 3 Combinatorial integral approximation decomposition

We propose to solve (MIOCP) with the CIA decomposition [27, 31], which relies on the use of direct methods (*first discretize, then optimize* approach). This section explains and defines the problem’s temporal discretization and the subproblems that constitute the decomposition algorithm.

**Definition 4** ( $\mathcal{G}_N, \Delta$ ) Let the ordered set  $\mathcal{G}_N := \{t_0 < \dots < t_N = t_f\}$  denote a time grid with  $N$  intervals and lengths  $\Delta_j := t_j - t_{j-1}$  for  $j \in [N]$ ,  $\bar{\Delta} := \max_{j \in [N]} \Delta_j$  as well as  $\underline{\Delta} := \min_{j \in [N]} \Delta_j$ .

Next, we define the matrix sets of the discretized binary and relaxed binary control functions  $\Omega_N, \mathcal{A}_N$ .

**Definition 5 (Convex combination constraint (Conv) and  $\Omega_N, \mathcal{A}_N$ )** Let  $N \in \mathbb{N}$ . We express the requirement that the columns of a matrix  $(m_{i,j}) \in [0, 1]^{n_\omega \times N}$  sum up to one by

$$\sum_{i \in [n_\omega]} m_{i,j} = 1, \quad \text{for } j \in [N], \tag{Conv}$$

and call it *convex combination constraint (Conv)* in the remainder. Based on this constraint, we define

$$\Omega_N := \{w \in \{0, 1\}^{n_\omega \times N} \mid w \text{ satisfies (Conv)}\}, \quad \mathcal{A}_N := \{a \in [0, 1]^{n_\omega \times N} \mid a \text{ satisfies (Conv)}\}.$$

We introduce a discretized version of the TV constraint (2.4) that relies on  $w$ :

**Definition 1 (Discretized version of (2.4))** Let  $\mathcal{G}_N$  and  $\sigma_{\max} \in \mathbb{N}$  be given. We use auxiliary variables  $\sigma_{i,j} \in \mathbb{N}$  for introducing a discretized version of the TV constraint that reads

$$\sigma_{\max} \geq \frac{1}{2} \sum_{i \in [n_\omega]} \sum_{j \in [N]} \sigma_{i,j}, \tag{3.1}$$

$$\sigma_{i,j} \geq \pm (w_{i,j} - w_{i,j-1}), \quad i \in [n_\omega], j \in [N]. \tag{3.2}$$

In order to solve the upcoming subproblems efficiently, we have deliberately formulated the above constraints without an absolute value term. This and replacing  $w$  with  $a$  in (3.2) results in a differentiable TV constraint. We define the discretizations of (MIOCP) and (OCP) below.

**Definition 6 (( $NLP_{\text{rel}}$ ), ( $NLP_{\text{bin}}$ ))** Consider (OCP) with the following modifications:

- We discretize (2.2) with  $\mathcal{G}_N$  and by using direct collocation or direct multiple shooting together with an appropriate integrator function (e.g. Runge-Kutta methods [18, 23]).
- The controls are piecewise constant functions on  $\mathcal{G}_N$  with  $(a_{i,j}) \in \mathcal{A}_N$ :

$$\alpha_i(t) := a_{i,j}, \quad \text{for } i \in [n_\omega], t \in [t_{j-1}, t_j], j \in [N], t_j \in \mathcal{G}_N.$$

- The TV constraint (2.4) is replaced by the constraints (3.1) and (3.2). In (3.2), we replace  $w$  with  $a$ .

We denote the resulting discretized optimization problem with relaxed binary control functions  $a \in \mathcal{A}_N$  by ( $NLP_{\text{rel}}$ ) and by ( $NLP_{\text{bin}}$ ) for fixed binary control functions  $w \in \Omega_N$ .

**Definition 7 [(CIA) problem,  $\theta(w)$ ]** Let  $a \in \mathcal{A}_N$  be given. Then, we define the problem (CIA) to be

$$\min_{\sigma, \theta, w \in \Omega_N} \theta \tag{3.3}$$

$$\text{s. t. } \theta \geq \pm \sum_{l \in [j]} (a_{i,l} - w_{i,l}) \Delta_l, \quad \text{for } i \in [n_\omega], j \in [N] \tag{3.4}$$

TV constraints (3.1) and (3.2).

We denote with  $\theta(\mathbf{w})$  the (CIA) objective value for a feasible solution  $\mathbf{w} \in \Omega_N$ .

With these subproblem definitions we are able to summarize the CIA decomposition in Algorithm 1. We first solve the relaxed problem (NLP<sub>rel</sub>) and approximate the resulting relaxed binary controls with binary values in the (CIA) problem. The last step consists of evaluating (NLP<sub>bin</sub>) with a fixed binary control function  $\mathbf{w}_{\text{CIA}}$  in order to obtain the objective value of (MIOCP).

We remark that the TV constraints (3.1)–(3.2) in (NLP<sub>rel</sub>) and (NLP<sub>bin</sub>) may be replaced by other TV reformulations, such as the ones presented in [17] or even dropped because (CIA) guarantees feasibility with

---

**Algorithm 1:** CIA decomposition algorithm for error-controlled solution of (MIOCP)

---

- Input** : Discretized (MIOCP) instance with time grid  $\mathcal{G}_N$ .
  - 1 Solve (NLP<sub>rel</sub>)  $\rightarrow \mathbf{x}, \mathbf{a}, \Phi_{\text{rel}}$
  - 2 Solve (CIA) for  $\mathbf{a} \rightarrow \mathbf{w}_{\text{CIA}}$ ;
  - 3 Evaluate (NLP<sub>bin</sub>) with  $\mathbf{w} = \mathbf{w}_{\text{CIA}}$  fixed  $\rightarrow \mathbf{x}, \Phi_{\text{bin}}$ ;
  - 4 **return**:  $(\mathbf{x}^*, \mathbf{w}^*, \Phi^*) = (\mathbf{x}, \mathbf{w}, \Phi)$ ;
- 

respect to bounded TV in any case. The algorithmic focus of this article lies on the (CIA) step, so that we omit further considerations of TV reformulations.

We stress that this algorithm solves two problems of which each one is less hard to solve than the original MINLP, which denotes the discretized (MIOCP). It yields only an approximation of the optimal solution of (MIOCP), but Sager et al. [26] showed that—without TV constraints and under certain regularity assumptions—the difference between the differential states based on relaxed and binary control values depends linearly on the difference of the integrals of their corresponding control function. In particular, they proved that the so called *integrality gap* with feasible solutions constructed by (CIA) is linearly bounded by the maximum grid length  $\bar{\Delta}$  in the sense of

$$\sup_{t \in \mathcal{T}} \left\| \int_{t_0}^t \boldsymbol{\alpha}(\tau) - \boldsymbol{\omega}(\tau) \, d\tau \right\|_\infty \leq C(n_\omega) \cdot \bar{\Delta}, \tag{3.5}$$

where  $C(n_\omega)$  is a constant depending on the number of controls. This implies that the differential state trajectories of (MIOCP) and (OCP)—*without* the TV

constraint—are arbitrarily close with vanishing grid length  $\bar{\Delta}$  and by assumed Lipschitz continuity of the objective the same holds also for the objective values.<sup>1</sup>

In Sects. 6–7, we are going to show that in the presence of discrete TV constraints (3.1)–(3.2) the rounding error does not vanish in general with grid length going to zero. More precisely, Theorem 4 and Corollary 7 will present concrete results on the rounding error.

#### 4 (CIA- $\bar{\theta}$ ), (CIA- $\bar{\theta}$ -init) and an associated lower bound

In this section, we address a problem that minimizes the used switches subject to a given approximation error  $\bar{\theta}$  that shall not be exceeded by the accumulated control deviation. Afterward, we aim for a lower bound on its objective that will be useful in the next section and introduce useful auxiliary variables and definitions for this.

**Definition 8 ((CIA- $\bar{\theta}$ ), (CIA- $\bar{\theta}$ -init))** For given  $\mathbf{a} \in \mathcal{A}_N$ ,  $\bar{\theta} > 0$  and initial active control  $i_0 \in [n_\omega]$  the problem **(CIA- $\bar{\theta}$ -init)** is defined to be

$$\sigma^* := \min_{\mathbf{w} \in \Omega_N} \frac{1}{2} \sum_{l=1}^{N-1} \sum_{i=1}^{n_\omega} |w_{i,l+1} - w_{i,l}| \tag{4.1}$$

$$\text{s. t.} \quad \bar{\theta} \geq \pm \sum_{l=1}^j (a_{i,l} - w_{i,l}) \Delta_j \quad i \in [n_\omega], j \in [N], \tag{4.2}$$

$$w_{i_0,1} = 1. \tag{4.3}$$

We define the problem **(CIA- $\bar{\theta}$ )** to be **(CIA- $\bar{\theta}$ -init)** without the constraint (4.3).

Taking the fixed initial active control in (4.3) aside, the problems (CIA- $\bar{\theta}$ ) and (CIA) from Definition 7 are closely connected with each other because the TV constraints (3.1) and (3.2) are reinterpreted as objective function subject to a fixed approximation error  $\bar{\theta}$ . This justifies the naming. We will introduce the MDR algorithm in Sect.5 to (heuristically) solve (CIA- $\bar{\theta}$ -init). By applying this algorithm to all  $i \in [n_\omega]$  as initial active controls, we exploit this relationship to solve (CIA- $\bar{\theta}$ ) as well, which will be used then as part of a bisection algorithm to solve (CIA).

We stress that fixing the initial active control  $i_0$  may seem odd, though this fixing reduces the problem complexity, which later yields in Theorem 1, Section 5, an optimality result of the solution constructed by the MDR algorithm concerning (CIA- $\bar{\theta}$ -init).

We notice that (CIA- $\bar{\theta}$ -init) is very similar to (SCARP) from [3, 4]. The latter problem aims at minimizing the switching costs, representing a generalized

<sup>1</sup> For the sake of completeness we mention that an improved approximation result for this context has recently been proven by Manns and Kirches [19].



objective function of (CIA- $\bar{\theta}$ -init), whereas in (CIA- $\bar{\theta}$ -init) the initial active control is fixed.

**Remark 2** (Link to scheduling theory) On an equidistant grid, (CIA- $\bar{\theta}$ ) can be reformulated into the following, equivalent, scheduling problem: On a single machine, minimize the total setup costs (TSC) until the  $N$ th processed job,  $N \leq n$ , so that  $n$  jobs  $(f, k)$  are processed within  $f \in [n_\omega]$  job families subject to release times  $r_{f,k}$ , deadlines  $d_{f,k}$ , equal processing times  $\bar{\Delta}$  and sequence-independent setup costs, which can be summarized in scheduling notation [13] as

$$(1|r_{f,k}, d_{f,k}, SC_{si,b} = 1, p_{f,k} = \bar{\Delta}|TSC|_1^N).$$

In the following we will revert to scheduling-like concepts, but explicitly dispense with its notation to not distract the reader from the usual MIOCP notation.

Next, we need some definitions to derive a lower bound for (CIA- $\bar{\theta}$ -init) at the end of this section. We stress that we establish our results on an equidistant grid but will sometimes drop this assumption in definitions used in later sections.

**Definition 9 (Activations, release  $r_{i,k}$  and deadline intervals  $d_{i,k}$ )** For each control  $i \in [n_\omega]$  on an equidistant grid  $\mathcal{G}_N$ , we introduce the number of possible activations  $n_i$  as

$$n_i := \max \left\{ k \in \mathbb{N} \mid \sum_{l=1}^N a_{i,l} - k \geq -\bar{\theta}/\bar{\Delta} \right\}.$$

Each activation  $k \in [n_i]$  is associated with a release and deadline interval, which are defined by:

$$r_{i_0,1} := 1, \quad d_{i_0,1} := 1, \tag{4.4}$$

$$r_{i,k} := \min \left\{ j \geq r_{i,k-1} + 1 \mid \sum_{l=1}^j a_{i,l} - k \geq -\bar{\theta}/\bar{\Delta} \right\}, \text{ with } r_{i,0} := 0, \tag{4.5}$$

$$d_{i,k} := \begin{cases} \infty, & \text{if } \sum_{l=1}^N a_{i,l} - k \leq \bar{\theta}/\bar{\Delta}, \\ \max \left\{ j \mid \sum_{l=1}^j a_{i,l} - k \leq \bar{\theta}/\bar{\Delta} \right\}, & \text{else.} \end{cases} \tag{4.6}$$

Finally, we call the  $k$ th activation of control  $i$  necessary, if  $d_{i,k} < \infty$ .

**Definition 10** (Switch, activation block) Consider  $w \in \Omega_N$ . If we have on interval  $j \geq 2$  and for any  $i \in [n_\omega]$

$$w_{i,j-1} = 0, \quad w_{i,j} = 1,$$

then we say  $w$  switches on  $j$ . We introduce the set of switches

$$S := \{j \in \{2, \dots, N\} \mid w \text{ switches on } j\},$$

and set  $n_s := |S|$ . We denote by  $\tau_j \in [N]$  the corresponding interval of the  $j$ th switch of  $w$ , where we set  $\tau_0 := 0$ ,  $\tau_{n_s+1} := N$ . On an equidistant grid and if  $i \in [n_\omega]$  is active between two consecutive switches or one switch and the first/last interval, we define the set of activations of  $i$  between these switches as an *activation block*  $B \subset [n_i]$ . On a general grid, we further define the length of the  $j$ th activation block between the  $(j - 1)$ st switch on, i.e.  $\tau_{j-1}$ , and before the  $j$ th switch, i.e.  $\tau_j - 1$ , via the auxiliary variable  $\delta_j = \sum_{l=\tau_{j-1}}^{\tau_j-1} \Delta_l$  for  $j \in [n_s + 1]$ .

We notice that the switches actually occur on the grid points; however, we have indexed the variables  $w_{i,j}$  according to the intervals, and therefore, for simplicity, we refer to switches on intervals. In the following, we will sometimes abbreviate activation block with block. In order to keep the number of used switches small and when deciding to set up a new block, it is highly relevant to know how many activations could be at most included in this block beginning with activation  $k$ . An activation  $j > k$  cannot be included in the block if its release interval begins later than the deadline interval of activation  $k$  plus the number of activations between  $k$  and  $j$ . We give a definition that formalizes these deadlines for initial activation-dependent deadlines of blocks. Based on these block deadlines, it is straightforward to introduce the notion of a *block deadline feasible* partition of activations into blocks. The constraint (4.3) imposes that the control  $i_0$ 's first activation has to be executed on the first interval, for which we introduce the definition of *fixed initial active control* feasibility.

**Definition 11 (*db<sub>i,k</sub>*, **block deadline and fiac feasible partition**)** Consider an equidistant grid. The deadline of a block for  $i \in [n_\omega]$  that begins with the  $k$ th activation,  $k \in [n_i]$ , is defined by

$$db_{i,k} := d_{i,l}, \quad \text{where } l := \max\{j \geq k \mid r_{i,j} \leq d_{i,k} + j - k\}. \tag{4.7}$$

Let  $P_i$  denote a partition of all activations  $[n_i]$  for  $i \in [n_\omega]$ . We call  $P_i$  *block deadline feasible* if for all subsets  $B \in P_i$ , i.e., all blocks, hold:

$$r_{i,\max\{k \in B\}} \leq d_{i,\min\{k \in B\}} + |B| - 1.$$

Furthermore, we refer to  $P_i$  as a *fixed initial active control (fiac) feasible* partition if for all  $k \in B_1$  hold

$$r_{i,k} = k,$$

where  $B_1 \in P_i$  denotes the first activation block of  $P_i$ .

In the last definition, we provided the concept of a control specific partition of all activations.

The  $k$ th activation of control  $i \in [n_\omega]$  does generally not coincide with the  $k$ th interval. The following example illustrates the introduced concepts and, in particular,

that there may be in total *more possible* but *less necessary* activations than intervals  $N$ .

**Example 1** Let the following matrices  $\mathbf{a} \in \mathcal{A}_N$  and  $\mathbf{w} \in \Omega_N$  for equidistant discretization given:

$$\mathbf{a} := \begin{pmatrix} 1 & 1 & 0.8 & 0 & 0 & 0 & 0 & 0.5 \\ 0 & 0 & 0.2 & 0 & 0.1 & 0.8 & 1 & 0.5 \\ 0 & 0 & 0 & 1 & 0.9 & 0.2 & 0 & 0 \end{pmatrix}, \quad \mathbf{w} := \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \end{pmatrix},$$

where  $n_\omega = 3, N = 9$ . Consider  $i = 1$  to be the fixed initial active control and a rounding threshold of  $\bar{\theta} = 1\bar{\Delta}$ . Then, we deal with in total eleven possible activations with their release and deadline intervals:

$$\begin{aligned} i = 1 : & \quad [r_{1,k}, d_{1,k}] = [1, 1], [2, 3], [3, 9], [9, \infty], \\ i = 2 : & \quad [r_{2,k}, d_{2,k}] = [1, 6], [6, 7], [7, 8], [8, \infty], \\ i = 3 : & \quad [r_{3,k}, d_{3,k}] = [1, 5], [4, 6], [6, \infty]. \end{aligned}$$

There are 4, 3, and 2 activations in  $\mathbf{w}$  for the controls  $i = 1, 2$ , and 3, respectively. These activations are grouped into in total 4 activation blocks so that  $\mathbf{w}$  uses 3 switches. For instance, the first block of control  $i = 1$  has a length of  $\delta_1 = 3\bar{\Delta}$  and its deadline is  $db_{1,1} = d_{1,3} = 9$ . The partition  $P_1 = \{\{1, 2, 3\}, \{4\}\}$  is *fiac feasible* for  $i = 1$ . For control  $i = 3$ , the partitions  $P_3 = \{\{1, 2, 3\}\}, \{\{1, 2\}, \{3\}\}$  are amongst others *block deadline feasible*.

As illustrated in Example 1, a feasible solution  $\mathbf{w}$  of (CIA- $\bar{\theta}$ -init) may not use all possible activations. To this end, we define an extension of the set of blocks of  $\mathbf{w}$  to become a partition of  $[n_i]$  for all  $i \in [n_\omega]$  in the following lemma. The extension may seem arbitrary, but is necessary to compare any  $\mathbf{w} \in \Omega_N$  with partitions of  $[n_i]$ . Thereby, we establish a connection between the above feasibility concepts and a feasible solution  $\mathbf{w}$  of (CIA- $\bar{\theta}$ -init).

**Lemma 1** *For an equidistant grid, let  $\mathbf{w} \in \Omega_N$  be feasible for (CIA- $\bar{\theta}$ -init) and let  $P'_i$  denote the set of blocks of  $\mathbf{w}$  for control  $i \in [n_\omega]$ . We define  $\bar{P}'_i := \{k \in [n_i] \mid \nexists B \in P'_i : k \in B\}$  and  $P_i := P'_i \cup_{k \in \bar{P}'_i} \{k\}$ . Then,  $P_i$  is a block deadline feasible partition and if  $i = i_0$ ,  $P_i$  is also fiac feasible.*

**Proof** We first argue that  $P_i$  is by definition a partition of  $[n_i]$ . We need to prove that these partitions are *block deadline feasible*, respectively *fiac feasible*. If for  $i \in [n_\omega]$  and an activation block  $B \in P_i$  holds

$$r_{i, \max\{k \in B\}} > d_{i, \min\{k \in B\}} + |B| - 1,$$

this would imply that the  $(\max\{k \in B\})$ th activation of  $i$  has been processed before its release interval because  $B$  can not be interrupted by activations from other controls. Therefore, the above inequality does not hold and *block deadline feasibility*

is established. We apply the same argument for confirming *fiac feasibility*. By constraint (4.3), the first activation of  $i_0$  is scheduled on the first interval. Hence, all activations  $k \in B_1$  of the first block  $B_1$  must be processed on the  $k$ th interval and require therefore a release interval that is no later than  $k$ .  $\square$

**Remark 3** (*Necessary condition for feasibility of (CIA- $\bar{\theta}$ -init)*) The formation of activations into *block deadline* and for  $i_0$  *fiac feasible* partitions is a *necessary* feasibility criterion of  $w \in \Omega_N$  for (CIA- $\bar{\theta}$ -init) by virtue of Lemma 1. Nevertheless, it is not a *sufficient* criterion since the order of the processing of blocks is not clarified. In particular, one might order the blocks to contain an activation whose release interval is later than its executed interval.

Next, we formalize specific partitions of the control  $i$ 's possible activations  $n_i$  whose blocks are constructed to include as many activations as possible without violating their block deadlines. These quantities serve as tools to derive a lower bound of necessary blocks per control independent of the other control's blocks. This will result in a lower bound for (CIA- $\bar{\theta}$ -init) in Proposition 2. We distinguish between the case that  $i$  is the fixed initial active control, i.e.,  $i = i_0$ , or not.

**Definition 12** Consider an equidistant grid and the controls  $i_0, i \in [n_\omega]$ . Let

$$k_1 := \max\{j \leq n_i \mid d_{i,j} \leq db_{i,1}\}, \quad B_{i,1} := \{1, \dots, k_1\}, \tag{4.8}$$

$$k_1^{\text{init}} := \max\{j \leq n_i \mid r_{i_0,j} = j\}, \quad B_{i_0,1}^{\text{init}} := \{1, \dots, k_1^{\text{init}}\}. \tag{4.9}$$

We write  $(\cdot)_i^{(\text{init})}$  to indicate that equations or inequalities each apply to the parameters  $(\cdot)_i$  and  $(\cdot)_{i_0}^{\text{init}}$ . We define the blocks  $B_{i,l}^{(\text{init})}$  recursively for  $l \geq 2$  and while  $k_l^{(\text{init})} < n_i$  by

$$k_l^{(\text{init})} := \max\left\{j \leq n_i \mid d_{i,j} \leq db_{i,k_{l-1}^{(\text{init})}+1}\right\}, \quad B_{i,l}^{(\text{init})} := \left\{k_{l-1}^{(\text{init})} + 1, \dots, k_l^{(\text{init})}\right\}. \tag{4.10}$$

Let  $nb_{i,\min}^{(\text{init})}$  denote the number of blocks  $B_{i,l}^{(\text{init})}$  and  $P_{i,\min}^{(\text{init})}$  the partitions of  $[n_i]$  constructed by the latter:

$$P_{i,\min}^{(\text{init})} := \left\{B_{i,l}^{(\text{init})} \mid l \in [nb_{i,\min}^{(\text{init})}]\right\}.$$

The MDR scheme from the next section creates switches that resemble the above  $k_l^{(\text{init})}$  terms. The latter, though, only expresses the grouping of activations, while the switches explicitly specify the corresponding intervals as well. It turns out that the partitions  $P_{i,\min}$  and  $P_{i_0,\min}^{\text{init}}$  are minimal in the number of blocks as indicated in the following proposition.

**Proposition 1** For  $i_0, i \in [n_\omega]$  let the partitions  $P_{i,\min}, P_{i_0,\min}^{\text{init}}$  be given as in Definition 12. For any partition  $P_i$  of  $[n_i]$  with  $i = i_0$  included, we define its restriction to the first  $\tilde{n}_i \leq n_i$  activations as

$$P_i|_{\tilde{n}_i} := \{B \cap [\tilde{n}_i] \mid B \in P_i\}.$$

Then, the partition  $P_{i,\min}$  respectively  $P_{i_0,\min}^{\text{init}}$  consists for any  $\tilde{n}_i \leq n_i$  of a minimal number of blocks on the first  $\tilde{n}_i$  activations compared with all other block deadline feasible, respectively both block deadline and fiac feasible, partitions  $P_i$ :

$$\left| P_{i,\min}^{(\text{init})}|_{\tilde{n}_i} \right| \leq \left| P_i|_{\tilde{n}_i} \right|. \tag{4.11}$$

**Proof** We consider first  $P_{i,\min}|_{\tilde{n}_i}$ . It is *block deadline feasible* because the deadline of the last activation for each block is defined in (4.8) and (4.10) to be less or equal than the corresponding block deadline. Assume there is a *block deadline feasible* partition  $P_i$  for the control  $i \in [n_\omega]$  with  $\left| P_i|_{\tilde{n}_i} \right| < \left| P_{i,\min}|_{\tilde{n}_i} \right|$ . In other words, there exists a subset of the first  $j$  blocks of  $P_i|_{\tilde{n}_i}$  that includes more activations than the ones included into the first  $j$  blocks of  $P_{i,\min}|_{\tilde{n}_i}$ . We consider the minimal number of blocks  $j$  with this property:

$$j := \min \left\{ l \in [nb_{i,\min}] \mid B_{i,l} \in P_{i,\min}|_{\tilde{n}_i}, B'_{i,l} \in P_i|_{\tilde{n}_i} : \max\{k \in B_{i,l}\} < \max\{k \in B'_{i,l}\} \right\}. \tag{4.12}$$

The block index  $j$  is unique since the association of activations to blocks is monotonically increasing, meaning that there are no  $k_1$ th,  $k_2$ th activations,  $k_1 < k_2$ , with  $k_1 \in B_{i,l_1}, k_2 \in B_{i,l_2}$  and  $l_1 > l_2$ . We conclude

$$\min\{k \in B'_{i,j}\} \leq \min\{k \in B_{i,j}\}, \quad B'_{i,j} \in P_i|_{\tilde{n}_i}, B_{i,j} \in P_{i,\min}|_{\tilde{n}_i}, \tag{4.13}$$

so that block  $B'_{i,j}$ 's first activation  $k'$  is smaller or equal than  $k$  which marks the earliest activation of  $B_{i,j}$ . The definition of release intervals (4.54.6) implies  $r_{i,k'} \leq r_{i,k}$  for  $k' \leq k$ . Similarly, the definition of block deadlines (4.7) implies  $db_{i,k'} \leq db_{i,k}$  for  $r_{i,k'} \leq r_{i,k}$  and we find with (4.13) in particular

$$db_{i,\min\{k \in B'_{i,j}\}} \leq db_{i,\min\{k \in B_{i,j}\}}. \tag{4.14}$$

On the other hand, the definition of  $P_{i,\min}$  in (4.10) implies

$$db_{i,k_{j-1}+1} = \max\{k \in B_{i,j}\}. \tag{4.15}$$

Then, the definition of  $j$  yields

$$db_{i,\min\{k \in B_{i,j}\}} \stackrel{(4.10)}{=} db_{i,k_{j-1}+1} \stackrel{(4.15)}{=} \max\{k \in B_{i,j}\} < \max\{k \in B'_{i,j}\} \leq db_{i,\min\{k \in B'_{i,j}\}}, \tag{4.16}$$

where the last inequality must hold due to the assumption of  $P_i$  being *block deadline feasible*. Inequality (4.14) contradicts inequality (4.16), or equivalently there is no such partition  $P_i$  and  $P_{i,\min}$  uses indeed a minimal number of blocks on any  $[\tilde{n}_i] \subset [n_i]$ .

The same argumentation for  $j \geq 2$  in equation (4.12) can be applied in order to prove the result for  $P_{i_0,\min}^{\text{init}}$  as  $P_{i_0}$  is also assumed to be *block deadline feasible* in this case and the same holds for  $P_{i_0,\min}^{\text{init}}$  from the second block on. We just need to take care of the case when  $j = 1$ , i.e., if  $B_{i_0,j}$ , respectively  $B'_{i_0,j}$ , is the first block of the control  $i_0$ . Here,  $\max\{k \in B_{i_0,1}\} < \max\{k \in B'_{i_0,1}\}$  cannot appear, since  $P_{i_0}$  is assumed to be *fiac feasible* and the construction of the first block of  $P_{i_0,\min}^{\text{init}}$  implies that no further activation can be added to  $B_{i_0,1}$  without violating *fiac feasibility*. Thus,  $j = 1$  is impossible in (4.12) and  $P_{i_0,\min}^{\text{init}}$  is also minimal in the number of blocks. □

**Corollary 1** Consider the setting of Proposition 1 and the controls  $i_0, i \in [n_\omega]$ . We define

$$\tilde{n}_{i,N} := \max\{k \mid d_{i,k} \leq N\}, \quad nb_{i,\min}^N := \left| P_{i,\min} \Big|_{\tilde{n}_{i,N}} \right|, \quad nb_{i_0,\min}^{N,\text{init}} := \left| P_{i_0,\min}^{\text{init}} \Big|_{\tilde{n}_{i_0,N}} \right|.$$

There is no block deadline feasible partition, respectively block deadline and fiac feasible partition, that uses less than  $nb_{i,\min}^N$  blocks on  $[\tilde{n}_{i,N}]$ , respectively  $nb_{i_0,\min}^{N,\text{init}}$  blocks on  $[\tilde{n}_{i_0,N}]$ .

**Proof** The result follows directly from Proposition 1 with  $\tilde{n}_i = \tilde{n}_{i,N}$  and  $\tilde{n}_{i_0} = \tilde{n}_{i_0,N}$ . □

As a final result for this section, we establish a lower bound for (CIA- $\bar{\theta}$ -init) that will be useful in Theorem 1.

**Proposition 2 (Lower bound for (CIA- $\bar{\theta}$ -init))** Let  $\sigma^*$  be the objective of (CIA- $\bar{\theta}$ -init) with equidistant discretization and  $i_0$  the fixed initial active control as defined in Definition 8. Let  $nb_{i,\min}^N$  for all  $i \neq i_0$  and  $nb_{i_0,\min}^{N,\text{init}}$  be given as in Corollary 1. It results

$$\sum_{i \in [n_\omega], i \neq i_0} nb_{i,\min}^N + nb_{i_0,\min}^{N,\text{init}} - 1 \leq \sigma^*. \tag{4.17}$$

**Proof** By virtue of Lemma 1, a feasible solution of (CIA- $\bar{\theta}$ -init) satisfies the necessary condition of generating only *block deadline feasible* partitions  $P_i$  and if  $i = i_0$ , the activation partition  $P_i$  is also *fiac feasible*. Moreover, all activations are executed no later than their deadline interval. That holds particularly for those that are due no later than  $N$ . Hence, we can apply Corollary 1 and conclude the minimum number of blocks of a feasible solution until the in total  $N$ th activation is  $nb_{i,\min}^N$ , respectively

$nb_{t_0, \min}^{N, \text{init}}$ . Finally, we obtain the claim (4.17) by summing up over all controls and using that the setup of the first block does not count as switch.  $\square$

## 5 Maximum dwell rounding

This section is dedicated to solving (CIA). Generally, we recommend a tailored BNB algorithm that has been proposed by Jung et al. [15, 27] and implemented in the open-source software package `pycombin` [7]. The BNB algorithm outperforms standard MILP solvers in a case study [14] by three orders of magnitude. However, in some instances, the algorithm struggles to find the optimal solution quickly because the node relaxation can be quite weak [8]. We, therefore, present a polynomial-time algorithm that constructs good initial guesses for BNB and, in some situations, solves (CIA) even to optimality. We proceed by giving the necessary definitions of the algorithm itself and its auxiliary variables in the first subsection and investigate beneficial properties in the second subsection.

### 5.1 Definition of the algorithm

**Definition 13 (Accumulated control deviation  $\theta_{i,j}, \gamma_{i,j}$ )** Let  $\mathbf{a} \in \mathcal{A}_N$  and  $\mathbf{w} \in \Omega_N$ . For control  $i \in [n_\omega]$  and interval  $j \in [N]$  we define the accumulated control deviation variables as

$$\theta_{i,j} := \sum_{l=1}^j (a_{i,l} - w_{i,l}) \Delta_l, \quad \gamma_{i,j} := \sum_{l=1}^j a_{i,l} \Delta_l - \sum_{l=1}^{j-1} w_{i,l} \Delta_l,$$

and set  $\theta_{i,0} := 0$ .

The following lemma is useful for Proposition 3 on page 11 and Lemma 7 on page 17.

**Lemma 2** Consider  $\mathbf{a} \in \mathcal{A}_N$  and  $\mathbf{w} \in \Omega_N$ , for each  $j \in [N]$  holds

$$\sum_{i \in [n_\omega]} \theta_{i,j} = 0, \quad \sum_{i \in [n_\omega]} \gamma_{i,j} = \Delta_j.$$

**Proof** These equations follow directly from the definition of  $\theta$  and  $\gamma$  as well as from the convexity property of  $\mathbf{a}$  and  $\mathbf{w}$ .  $\square$

**Definition 14 (Inadmissible, next forced and forced activation)** Consider a rounding threshold  $\bar{\theta} > 0$  and  $\mathbf{a} \in \mathcal{A}_N$ . Let the values of  $\mathbf{w} \in \Omega_N$  be given until interval  $j - 1$ , with  $j \geq 2$ . The choice  $w_{i,j} = 1$  for  $i \in [n_\omega], j \in [N]$  is *admissible* if we have that

$$\theta_{i,j} \geq -\bar{\theta}$$

and we call the control  $i$  otherwise *inadmissible*. Similarly, the choice  $w_{i,j} = 1$  is *forced* if we have that

$$\gamma_{i,j} > \bar{\theta}.$$

Let further  $\mathcal{N}_j(i) \in \{j, \dots, N\}$  denote the next interval on which control  $i$  would become forced without activation after interval  $j - 1$ :

$$\mathcal{N}_j(i) := \begin{cases} \operatorname{argmin}_{k=j, \dots, N} \{ \theta_{i,j-1} + \sum_{l=j}^k a_{i,l} \Delta_l > \bar{\theta} \}, & \text{if } \theta_{i,j-1} + \sum_{l=j}^N a_{i,l} \Delta_l > \bar{\theta}, \\ \infty, & \text{else.} \end{cases}$$

Then, we define a control  $i^* \in [n_{\omega}]$  on interval  $j$  to be *next forced* if and only if

$$\mathcal{N}_j(i^*) = \min_{i \in [n_{\omega}]} \mathcal{N}_j(i) \quad \text{and} \quad \mathcal{N}_j(i^*) < \infty.$$

The above definition allows more than one control to be *next forced* or *forced* for an arbitrary interval  $j \in [N]$ . This is supposedly not the standard case in our discussion but will also be taken into account in our considerations. The guiding idea behind the above control activations is that we include more and more summands of  $\mathbf{w}$  into the computation of  $\theta$  and can choose the next row of  $\mathbf{w}$  accordingly. With this definition we have introduced necessary activation properties of feasible solutions for (CIA- $\bar{\theta}$ -init), but neglected so far the fixed initial active control constraint (4.3). The following definition fills this gap.

**Definition 15 (Initially admissible control)** We define a control  $i \in [n_{\omega}]$  to be *initially admissible* if it is *admissible* on the first interval and if there is no other control  $i_1 \neq i$  that is *forced* on the first interval.

Now, we can define the MDR in Algorithm 2.

The MDR algorithm assumes a given initial control  $i_0$  and activates it until it becomes *inadmissible* or until there is another *forced* control. We require the control  $i_0$  to be *initially admissible* because otherwise  $\mathbf{w}^{\text{MDR}}$  would violate the control accumulation constraint (4.2). Otherwise, the control  $i$  with the maximum forward control deviation  $\gamma_{i,j}$  is set active and remains so until it becomes inadmissible or another control becomes forced. This procedure is performed forward in time until the end of the time horizon  $N$  is reached. We named the algorithm “maximum dwell rounding” because it tries to stay in the current mode as long as possible without violating the given rounding threshold.

The AMDR is defined in Algorithm 3 and can be described as a *bisection* method. We initialize it with a trivial lower bound  $LB$  and upper bound  $UB$  for (CIA). The algorithm runs MDR iteratively with different threshold  $\bar{\theta}$  and initially admissible control as long as the difference of



---

**Algorithm 2:** Maximum Dwell Rounding

---

**Input** : Relaxed control values  $\mathbf{a} \in \mathcal{A}_N$ , rounding threshold  $\bar{\theta}$ , initially admissible control  $i_0 \in [n_\omega]$ .

```

1 Initialize  $\mathbf{w} = 0$ ,  $w_{i_0,1} = 1$  and  $i = i_0$ .
2 for  $j = 2, \dots, N$  do
3   if  $i$  is inadmissible or there is a forced control  $i_f \neq i$  then
4     Set  $i = \operatorname{argmax}_{k \in [n_\omega]} \gamma_{k,j}$  (ties may be broken arbitrarily);
5   end
6   Set  $w_{i,j} = 1$ ;
7 end
8 return:  $\mathbf{w}^{\text{MDR}} = \mathbf{w}$ ;
```

---



---

**Algorithm 3:** Adaptive Maximum Dwell Rounding

---

**Input** : Relaxed control values  $\mathbf{a} \in \mathcal{A}_N$ , optimum tolerance  $TOL > 0$ , allowed number of switches  $\sigma_{max}$ .

```

1 Initialize  $LB = 0$ ;  $UB = \bar{\theta} = t_f - t_0$ .
2 while  $UB - LB > TOL$  do
3   for  $i = 1, \dots, n_\omega$  do
4     if  $i$  is an initially admissible control then
5       Run MDR with  $i$  as initial control and set  $\mathbf{w} = \mathbf{w}^{\text{MDR}(i)}$ ;
6       if  $\mathbf{w}$  satisfies TV constraints (3.1),(3.2) and  $\theta(\mathbf{w}) < UB$  then
7          $UB = \theta(\mathbf{w})$ ;
8          $\bar{\theta} = UB - 0.5 \cdot (UB - LB)$ ;
9          $\mathbf{w}^{\text{AMDR}} = \mathbf{w}$ ;
10        break;
11      else if  $i = n_\omega$  then
12         $LB = \bar{\theta}$ ;
13         $\bar{\theta} = LB + 0.5 \cdot (UB - LB)$ .
14      end
15    end
16 end
17 return:  $(\mathbf{w}^{\text{AMDR}}, UB)$ ;
```

---

lower and upper bound is larger than the chosen tolerance  $TOL$  (lines 2–5). If the computed control function satisfies the TV constraint and exhibits a (CIA) objective value that is smaller than the current  $UB$ , we update  $UB$ , reset the rounding threshold  $\bar{\theta}$  via interval halving of  $UB - LB$  and save the current best solution (lines 6-10). The evaluation  $\theta(\mathbf{w})$  is necessary since MDR may construct a control function with a rounding gap larger than the desired gap  $\bar{\theta}$ , as will be discussed in the next subsection. If no computed control function  $\mathbf{w}$  with given initial control and  $\bar{\theta}$  fulfills the TV constraint, then we increase the  $LB$  (lines 11-14).

**5.2 Solution quality and properties of MDR (Algorithm 2)**

Although the MDR algorithm may seem simple, it generates optimal solutions  $\mathbf{w}$  for (CIA- $\bar{\theta}$ -init) under certain conditions, for which we need the following definition.

**Definition 16 (Canonical switch)** We define a switch  $j \in \mathcal{S}$  as defined in Definition 10 to be *canonical*, if on interval  $\tau_j$  holds: exactly one control  $i_1$  is *inadmissible* and exactly one control  $i_2 \neq i_1$  is *forced*.

We build our theoretical results of this section mainly on the following assumption.

**Assumption 1 (MDR uses only canonical switches)** Suppose  $w^{\text{MDR}} \in \Omega_N$  has been generated by MDR. We assume that all switches of  $w^{\text{MDR}}$  are *canonical*.

### 5.2.1 Properties of the MDR algorithm

Assumption 1 may seem restrictive, although it is satisfied anyway under certain conditions.

**Proposition 3 (MDR with  $n_\omega = 2$  and  $\bar{\theta} \geq \frac{1}{2}\bar{\Delta}$  uses canonical switches)** Consider  $n_\omega = 2$ ,  $a \in \mathcal{A}_N$  and any grid  $\mathcal{G}_N$ . If we choose  $\bar{\theta} \geq \frac{1}{2}\bar{\Delta}$ , then the control function  $w^{\text{MDR}}$  constructed by the MDR scheme uses only canonical switches.

**Proof** We have to prove:

1. If control  $i_1$  is forced on interval  $j \geq 2$ , then it is admissible.
2. For all intervals  $j \geq 2$  hold: control  $i_1$  is inadmissible, if and only if  $i_2 \neq i_1$  is forced.

1. follows from the definition of forced activation and from  $\bar{\theta} \geq \frac{1}{2}\bar{\Delta}$ :

$$\theta_{i_1,j} = \theta_{i_1,j-1} + a_{i_1,j}\Delta_j - \Delta_j > \bar{\theta} - \Delta_j \geq -\frac{1}{2}\bar{\Delta} \geq -\bar{\theta}.$$

For proving 2. let us assume  $i_1$  is forced on  $j \in [N]$ , i.e.,  $\gamma_{i_1,j} > \bar{\theta}$ . By virtue of Lemma 2 for  $\gamma_{i_2,j}$  we derive

$$\theta_{i_2,j-1} + a_{i_2,j}\Delta_j = \gamma_{i_2,j} < -\bar{\theta} + \Delta_j,$$

which means  $i_2$  is inadmissible on  $j$ . Conversely, if  $i_1$  is inadmissible on  $j$ , we conclude from  $\theta_{i_1,j-1} + (a_{i_1,j} - 1)\Delta_j < -\bar{\theta}$  and from the equation for  $\theta$  in Lemma 2  $\gamma_{i_2,j} = \theta_{i_2,j-1} + a_{i_2,j}\Delta_j > \bar{\theta}$ . Therefore,  $i_2$  is forced. □

**Remark 4** Assumption 1 is not necessarily true for a control problem that involves more than two binary controls. It may, however, hold for special cases of such a problem. For instance, if the relaxed values are of bang-bang type, i.e.,  $a_{i,j} \in \{0, 1\}$ , and  $\bar{\theta}$  is chosen smaller than the smallest activation block, then the situation resembles the case  $n_\omega = 2$  and Assumption 1 may hold (without proof). On the other hand, Example 3 is going to demonstrate that this assumption can indeed be quite restrictive.

Assumption 1 allows us to prove strong properties of control functions obtained by MDR and AMDR. The first result expresses that the MDR scheme produces indeed control functions which exhibit a (CIA) objective value smaller or equal than  $\bar{\theta}$ .

**Lemma 3 (MDR solution satisfies  $\bar{\theta}$  bound)** *Let Assumption 1 hold and let  $w^{MDR} \in \Omega_N$  be constructed by MDR with given threshold  $\bar{\theta}$ . Then, we obtain  $\theta(w^{MDR}) \leq \bar{\theta}$ .*

**Proof** As soon as the activated control becomes inadmissible or there is a forced control on interval  $j \geq 2$ ,  $w^{MDR}$  has a switch by the definition of MDR. By Assumption 1, the newly activated control is both forced and admissible, hence  $\theta_{i,j} \geq -\bar{\theta}$ , and there is also no other forced control on  $j$ , thus  $\theta_{i,j} \leq \bar{\theta}$ .  $\square$

The following example demonstrates that  $\theta(w^{MDR}) > \bar{\theta}$  may generally appear without Assumption 1.

**Example 2** Consider an equidistant discretization and  $a \in \mathcal{A}_N$  with the first values given as  $a_{1,1} = 1, a_{2,1} = 0, a_{1,2} = 0.5, a_{2,2} = 0.5$ . For this relaxed value, let  $w^{MDR} \in \Omega_N$  be the corresponding binary control function computed by MDR with given threshold  $\bar{\theta} = 0.4\bar{\Delta}$  and initial control  $i = 1$ . Then,  $w_{1,2}^{MDR} = 0, w_{2,2}^{MDR} = 1$  holds since the second control becomes forced on the second interval. At the same time, control  $i = 2$  is inadmissible on the second interval, hence Assumption 1 is violated, and it results  $\theta_{2,2} = -0.6\bar{\Delta} < -\bar{\theta}$ .

We reuse concepts from the previous chapter, especially *activations* and their grouping into *blocks*.

**Theorem 1 (Least switches property of MDR)** *Let Assumption 1 hold. For given  $a \in \mathcal{A}_N$  and an equidistant grid, let  $w^{MDR}$  be constructed by MDR with  $i$  as initial control and any  $\theta > 0$ , where we assume that  $i$  is initially admissible. Let  $\sigma(w^{MDR})$  denote the number of switches used by  $w^{MDR}$ . Then, for the optimal objective value  $\sigma^*$  of (CIA- $\bar{\theta}$ -init) with  $i_0 = i$  as initial control holds*

$$\sigma^* = \sigma(w^{MDR}). \tag{5.1}$$

**Proof** We can conclude  $w^{MDR}$  is a feasible solution of (CIA- $\bar{\theta}$ -init) by Lemma 3. Combining this with Proposition 2 yields

$$\sum_{i \in [n_\omega], i \neq i_0} nb_{i,\min}^N + nb_{i_0,\min}^{N,\text{init}} - 1 \leq \sigma^* \leq \sigma(w^{MDR}). \tag{5.2}$$

For proving optimality we note that  $w^{MDR}$  constructs partitions of the activations  $[n_i], i \in [n_\omega]$  that are due no later than  $N$  and let  $P_i^{MDR}$  denote these partitions. With the notation from Corollary 1 we want to show that these partitions coincide with the partitions constructed in Definition 12

$$P_{i_0}^{MDR} = P_{i_0,\min}^{\text{init}} \Big|_{\tilde{n}_{i_0,N}}, \quad P_i^{MDR} = P_{i,\min} \Big|_{\tilde{n}_{i,N}}, \text{ for } i \neq i_0, \tag{5.3}$$

because then Corollary 1 would imply  $w^{MDR}$  has  $\sum_{i \in [n_\omega], i \neq i_0} nb_{i,\min}^N + nb_{i_0,\min}^{N,\text{init}}$  activation blocks or equivalently

$$\sigma(w^{\text{MDR}}) = \sum_{i \in \{n_\omega\}, i \neq i_0} nb_{i,\min}^N + nb_{i_0,\min}^{N,\text{init}} - 1,$$

and the claim follows from inequality (5.2). Consider the first blocks  $B_1 \in P_{i_0}^{\text{MDR}}$  and  $B_1^{\text{init}} \in P_{i_0,\min}^{\text{init}}$ . By Assumption 1, the MDR algorithm activates  $i_0$  until it becomes *inadmissible* on the interval  $\tau_1$  of the first switch:

$$\theta_{i_0,\tau_1} = \sum_{l=1}^{\tau_1} a_{i_0,l} - w_{i_0,l} < -\bar{\theta}/\bar{\Delta}.$$

We compare this inequality with the Definition 9 of release intervals and notice that either the next activation  $\tau_1$  of control  $i_0$  has a release interval that is later than  $\tau_1$  or there is no further possible activation. So, if the  $\tau_1$ th activation exists, then its release interval has not yet been reached:

$$r_{i_0,\tau_1} = r_{i_0,\max\{k \in B_1\}+1} > \max\{k \in B_1\} + 1.$$

By the definition of  $P_{i_0,\min}^{\text{init}}$  we conclude  $B_1 = B_1^{\text{init}}$ . Let us now consider the  $j$ th blocks  $B_j \in P_{i_0}^{\text{MDR}}$  and  $B_j^{\text{init}} \in P_{i_0,\min}^{\text{init}} \Big|_{\bar{n}_{i_0,N}}$ , where  $j \geq 2$ . Again by the definition of MDR and Assumption 1,  $i_0$  is *forced* on interval  $\tau_j$ , which is equivalent to  $d_{i_0,\min\{k \in B_j\}} = \tau_j$ . The MDR scheme activates  $i_0$  either until  $N$  (then trivially  $B_j = B_j^{\text{init}}$ ) or until it becomes *inadmissible* on interval  $\tau_{j+1}$  (by Assumption 1). With the argumentation for  $j = 1$ , inadmissible means hereby the  $(\max\{k \in B_j\} + 1)$ th activation has a release interval greater than  $\tau_{j+1}$ . Using that the block's first activation  $\tau_j$  is processed on its deadline interval  $d_{i_0,\min\{k \in B_j\}}$ , this yields

$$r_{i_0,\max\{k \in B_j\}+1} > d_{i_0,\min\{k \in B_j\}} + |B_j| - 1.$$

The above inequality expresses that  $B_j$  contains as many activations as possible without violating its block deadline  $db_{i_0,\min\{k \in B_j\}}$  and by construction of  $P_{i_0,\min}^{\text{init}}$  this is equivalent to  $B_j = B_j^{\text{init}}$ . This settles the case  $i = i_0$  in (5.3). We can reuse the above arguments about *forced* and *inadmissible* activation for  $j \geq 2$  in order to analogously prove the case  $i \neq i_0$  in (5.3). □

**Remark 5** Theorem 1 is predicated on the assumption of an equidistant grid. We stress that after grid refinement of the optimal control problem, i.e., after several rounds of applying the CIA decomposition, this might be a restriction.

The following corollary establishes a way to find the optimum of (CIA- $\bar{\theta}$ -init) in the setting of Theorem 1.

**Corollary 2 (Using MDR to find a control function with minimum number of switches)** Consider the setting of Theorem 1. A control function  $w^*$  that uses a minimum number of switches, i.e.,  $\sigma(w^*) = \sigma^*$ , can be found by running MDR.

**Proof** Let  $i$  be the initial control of  $\mathbf{w}^*$ . Execute MDR with  $i$  as initial control so that the result follows directly from Theorem 1. □

It is not clear which control is the optimal initial active one in order to minimize switches. In practice, MDR must be executed one after the other for all controls  $i \in [n_\omega]$  as initial active control. This expresses the following corollary.

**Corollary 3 (Link between MDR and (CIA- $\bar{\theta}$ ))** Consider the setting of Theorem 1. We assume that the MDR algorithm constructs for all  $i \in [n_\omega]$  as initial active controls the control functions  $\mathbf{w}^{\text{MDR}}$  that use only canonical switches. Then, there is a minimizing control  $\mathbf{w}^* \in \Omega_N$  for (CIA- $\bar{\theta}$ ) that only uses canonical switches. Moreover, there exists  $i_0 \in [n_\omega]$  such that running MDR with  $i_0$  as initial control produces  $\mathbf{w}^{\text{MDR}} \in \Omega_N$  that minimizes (CIA- $\bar{\theta}$ ).

**Proof** If the MDR algorithm produces  $\mathbf{w}^{\text{MDR}}$  that only uses canonical switches,  $\mathbf{w}^{\text{MDR}}$  is optimal by Theorem 1 for (CIA- $\bar{\theta}$ -init) with the corresponding initial control fixed. Then, the result follows from the fact that the optimal solution of (CIA- $\bar{\theta}$ ) is contained in the set of optimal solutions for the set of problems (CIA- $\bar{\theta}$ -init) with each control  $i \in [n_\omega]$  initially fixed. □

**Lemma 4** Consider  $\mathbf{a} \in \mathcal{A}_N$  on an equidistant grid and assume  $n_\omega = 2$ . Let  $\mathbf{w}^{\text{MDR}}$  denote the control function constructed by MDR with  $\bar{\theta} > 0$  and given initial control. If  $\theta(\mathbf{w}^{\text{MDR}}) > \bar{\theta}$ , then there is no control function  $\mathbf{w} \in \Omega_N$  with the same initial active control and  $\theta(\mathbf{w}) \leq \bar{\theta}$ .

**Proof** We consider the first interval  $j$  on which the accumulated control deviation of  $\mathbf{w}^{\text{MDR}}$  is greater than  $\bar{\theta}$ . Let control  $i_1$  be active on  $j$ . By definition of the MDR scheme,  $|\theta_{i_1,j}| > \bar{\theta}$  or  $|\theta_{i_2,j}| > \bar{\theta}$  can only appear if there is a switch on interval  $j$  and

1.  $i_1$  is on interval  $j$  both forced and inadmissible or
2. both  $i_1$  and  $i_2$  are inadmissible on interval  $j$ .

Proposition 3 establishes that  $\mathbf{w}^{\text{MDR}}$  uses only canonical switches for  $\bar{\theta} \geq \frac{1}{2}\bar{\Delta}$  and thus the above cases cannot appear for  $\bar{\theta} \geq \frac{1}{2}\bar{\Delta}$ . Let us focus on  $\bar{\theta} < \frac{1}{2}\bar{\Delta}$ . In order to create a control function  $\mathbf{w}$  that does fulfill  $\theta(\mathbf{w}) \leq \bar{\theta}$ , we need to change at least one activation of  $\mathbf{w}^{\text{MDR}}$  on an earlier interval  $l < j$ . However, we recognize that any earlier change of activation is not possible:

- We cannot extend an activation block at its end, since the active control is inadmissible.
- If the active control  $i_1$  is admissible on  $l$ , then the other control  $i_2$  is not forced on  $l$  – otherwise it would be active in the MDR scheme. This means  $\theta_{i_2,l-1} + a_{i_2,l}\bar{\Delta} \leq \bar{\theta}$ . Activating  $i_2$  on  $l$  results in

$$\theta_{i_2,l} = \theta_{i_2,l-1} + (a_{i_2,l} - 1)\bar{\Delta} \leq \bar{\theta} - \bar{\Delta} < -\frac{1}{2}\bar{\Delta} < -\bar{\theta},$$

where we applied  $\bar{\theta} < \frac{1}{2}\bar{\Delta}$ . This indicates the (CIA) objective value of  $\mathbf{w}$  is again greater than  $\bar{\theta}$ .

Hence, no previous activation  $\mathbf{w}^{\text{MDR}}$  can be changed so that there is no  $\mathbf{w}$  with  $\theta(\mathbf{w}) \leq \bar{\theta}$ . □

### 5.2.2 Properties of the AMDR algorithm

Theorem 2 states that the AMDR Algorithm is able to find the optimal solution of (CIA) for  $n_\omega = 2$  and equidistant discretization. Otherwise, strict assumptions are required for optimality, and in general, the found feasible solution represents only a promising upper bound.

**Theorem 2 (Properties of Algorithm 3)** *Algorithm 3 terminates for given  $\mathbf{a} \in \mathcal{A}_N$ ,  $TOL > 0$  and  $\sigma_{\max} \in \mathbb{N}$  after a finite number of iterations. Furthermore, consider an equidistant grid  $\mathcal{G}_N$ . Let  $\mathbf{w}^{\text{AMDR}}$  denote the solution constructed by Algorithm 3. It follows:*

1.  $\mathbf{w}^{\text{AMDR}}$  is a feasible solution of (CIA).
2. (a) If  $n_\omega = 2$ , we have for the optimum  $\theta^*$  of (CIA):  $\theta(\mathbf{w}^{\text{AMDR}}) \leq \theta^* + TOL$ .  
 (b) Let  $n_\omega > 2$ . We assume the MDR scheme uses in every run only canonical switches. Furthermore, suppose we have : If the MDR scheme constructs a solution with  $\theta(\mathbf{w}^{\text{MDR}}) > \bar{\theta}$ , then there is no control function  $\mathbf{w} \in \Omega_N$  with the same initial active control and  $\theta(\mathbf{w}) \leq \bar{\theta}$ . With these assumptions we obtain for the optimum  $\theta^*$  of (CIA):  $\theta(\mathbf{w}^{\text{AMDR}}) \leq \theta^* + TOL$ .
3. ADMR has time complexity  $\mathcal{O}(n_\omega \cdot C_{\text{MDR}} \cdot \log_2(\lceil (t_f - t_0)/TOL \rceil))$ , where  $C_{\text{MDR}} \in \mathcal{O}(N)$  denotes the time complexity of the MDR scheme.

**Proof** AMDR is a bisection algorithm that either decreases *UB* (line 7-8) or increases *LB* (line 12-13) by at least one half of  $(UB - LB)$  in every `while` iteration (line 2). From this and because of  $TOL > 0$ , we conclude that the `while` loop and AMDR as a whole terminate after finitely many iterations.

1. The objective of (CIA) cannot be greater than  $t_f - t_0$ , even with no switches allowed, i.e.,  $\sigma_{\max} = 0$ . Since we initialize the AMDR algorithm with  $UB = t_f - t_0$ , it finds in any case a feasible solution.
2. Every in line 5 by MDR generated  $\mathbf{w}^{\text{MDR}}$  that satisfies the TV constraints together with  $\theta(\mathbf{w}^{\text{MDR}}) < UB$  represents an upper bound on  $\theta^*$ , i.e.,  $UB = \theta(\mathbf{w}^{\text{MDR}}) \geq \theta^*$ . For proving that AMDR constructs valid lower bounds *LB* on  $\theta^*$  we exploit that  $\mathbf{w}^{\text{MDR}}$  only uses canonical switches for  $n_\omega = 2$  and by assumption for  $n_\omega > 2$  so that Corollary 3 is applicable. If there is no initial control  $i \in [n_\omega]$  for which

MDR produces  $w^{MDR}$  for given  $\bar{\theta}$  that uses less or equal switches than by the TV constraints required, we conclude by Corollary 3 that there exists no such  $w \in \Omega_N$  for this specific threshold  $\bar{\theta}$  and, hence,  $LB = \bar{\theta} \leq \theta^*$  is a true lower bound on the optimal (CIA) objective value. Moreover, if MDR constructs for a given  $\bar{\theta}$  and all initial controls  $i \in [n_\omega]$  control functions  $w^{MDR}$  with  $\theta(w^{MDR}) > \bar{\theta}$ , Lemma 4 and the assumption in (b) guarantee that this  $\bar{\theta}$  is also a true lower bound on the optimal (CIA) objective value. Altogether, AMDR iteratively generates valid lower  $LB$  and upper bounds  $UB$  for  $\theta^*$  and produces a feasible solution that is optimal up to the chosen tolerance  $TOL$ .

- MDR runs forward in time and computes solely the accumulated control deviation  $\gamma$  and  $\theta$  for all intervals  $j \in [N]$ , therefore  $C_{MDR} \in \mathcal{O}(N)$ . The interval halving in AMDR ensures that we execute the `while` loop a maximum of  $\log_2((t_f - t_0)/TOL)$  times. Inside this loop, we need to run the MDR scheme in the worst case with all  $n_\omega$  controls as initial controls. Combining these findings yields the asserted complexity. □

**Remark 6** Several meaningful modifications for the AMDR algorithm are available. We may use it also for finding control functions fulfilling other combinatorial constraints such as minimum dwell time constraints by checking them together with the TV constraint in line 5. As part of the MDR scheme, the control with maximum forward control deviation  $\gamma$  is activated if the previously active control is *inadmissible*. Instead, one may choose a less greedy variant. For instance, we could activate the *next forced* and *admissible* control. Lastly, the initial upper bound  $UB$  can be reduced, as we will point out in the next section.

**Remark 7** If we drop the TV constraint on  $w$ , the AMDR scheme finds a control function with the same objective value as the one obtained by the control function of Sum-Up Rounding [24] (without proof).

Most results of this section are based on the assumption of an equidistant discretization, which is common in practice. However, the assumption of dealing only with *canonical* switches in the produced control function is critical. The following example illustrates that a control function generated by MDR with non-canonical switches may use more switches than needed or may not satisfy the rounding bound  $\bar{\theta}$ .

**Example 3** Consider an equidistant grid. Let the following two relaxed values  $a^1, a^2 \in \mathcal{A}_N$  be defined as

$$(a^1_{i,j})_{i \in [3], j \in [3]} := \begin{pmatrix} 1 & 0.25 & 0 \\ 0 & 0.375 & 0.5 \\ 0 & 0.375 & 0.5 \end{pmatrix}, \quad (a^2_{i,j})_{i \in [3], j \in [3]} := \begin{pmatrix} 1 & 0.2 & 0 \\ 0 & 0.4 + \epsilon & 0 \\ 0 & 0.4 - \epsilon & 1 \end{pmatrix}, \quad 0 < \epsilon < 0.4.$$

Then, MDR with  $i = 1$  as initial control and  $\bar{\theta}^1 = 0.75\bar{\Delta}$ , respectively  $\bar{\theta}^2 = (0.6 + \epsilon)\bar{\Delta}$ , constructs the following control functions:

$$(w_{ij}^{\text{MDR},1})_{i \in [3], j \in [3]} := \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad (w_{ij}^{\text{MDR},2})_{i \in [3], j \in [3]} := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

In the first example, two controls are simultaneously *forced* on the third interval; the (CIA) objective value would be smaller if  $w_{3,2} = 1$  was chosen. The MDR constructs in the second example a control function that uses two switches, although activating the third control on the second interval would result in only one switch with almost the same (CIA) objective value. Both examples have in common that *non-canonical* switches are used. Hence, the improved control functions would be

$$(w_{ij}^{\text{OPT},1})_{i \in [3], j \in [3]} := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (w_{ij}^{\text{OPT},2})_{i \in [3], j \in [3]} := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

### 6 Upper bounds on (CIA) with $n_\omega = 2$

In this section, we use the MDR algorithm and previous results to deduce bounds on (CIA). We consider a given (CIA) problem with grid  $\mathcal{G}_N$ , relaxed value  $\mathbf{a} \in \mathcal{A}_N$  and maximum number of switches  $\sigma_{\max} > 0$ . The idea in the following is to construct a control function  $w^{\text{MDR}}$  that bounds the objective of (CIA). For finding an appropriate initial active control for the MDR scheme, we introduce an auxiliary grid  $\tilde{\mathcal{G}}_N$  which ends at  $\tilde{t}_f$  and has  $\tilde{N}$  intervals:

$$\tilde{\mathcal{G}}_N := \mathcal{G}_N \cap \left[ t_0, t_0 + \frac{5(t_f - t_0)}{3 + 2\sigma_{\max}} \right], \quad \tilde{N} := |\tilde{\mathcal{G}}_N| - 1, \quad \tilde{t}_f := \max\{t_j \mid t_j \in \tilde{\mathcal{G}}_N\}.$$

In the definition of  $\tilde{\mathcal{G}}_N$  we intersect two sets because we consider given  $\mathcal{G}_N$  and  $\sigma_{\max}$ . To specify the rounding of a value  $t_0 \leq t$  down to the next grid point, we utilize the following brackets notation

$$\lfloor t \rfloor_{\mathcal{G}_N} := \max\{t_j \in \mathcal{G}_N \mid t_j \leq t\}.$$

Depending on whether we deal with an equidistant grid or not, we can prove a sharp bound for (CIA). We are going to distinguish between these two cases in the upcoming results and introduce the following constant

$$C_1 := \begin{cases} \frac{1}{3+2\sigma_{\max}}, & \text{if } \mathcal{G}_N \text{ equidistant,} \\ 0, & \text{else.} \end{cases} \tag{6.1}$$

We propose to apply the rounding threshold

$$\bar{\theta} := \frac{t_f - t_0}{3 + 2\sigma_{\max}} + \frac{1}{2} \bar{\Delta} - \frac{C_1}{2} \bar{\Delta} \tag{6.2}$$



in the MDR scheme and claim that this choice will be later beneficial for proving upper bounds on (CIA). Next, we establish useful properties of the rounding  $\lfloor \cdot \rfloor_{\mathcal{G}_N}$  to the next grid point.

**Lemma 5 (Distance to next grid points)** *Consider  $\sigma_{\max} > 0$  and the rounding threshold  $\bar{\theta}$  defined as above. The following holds true:*

1.  $\lfloor t_0 + j\bar{\theta} \rfloor_{\mathcal{G}_N} \geq t_0 + j\bar{\theta} - \bar{\Delta} + C_1\bar{\Delta}, \quad j \in [2],$
2.  $\left\lfloor t_0 + \frac{5(t_f - t_0)}{3 + 2\sigma_{\max}} \right\rfloor_{\mathcal{G}_N} \geq t_0 + \frac{5(t_f - t_0)}{3 + 2\sigma_{\max}} - \bar{\Delta} + C_1\bar{\Delta}.$

**Proof**

1. Let us first consider the non-equidistant case. If  $t_0 + j\bar{\theta} \leq t_f$ , we deduce that the maximum distance of  $t_0 + j\bar{\theta}$  to the next smaller or equal grid point is  $\bar{\Delta}$ . If  $t_0 + j\bar{\theta} > t_f$ , we have  $\lfloor t_0 + j\bar{\theta} \rfloor_{\mathcal{G}_N} = t_f$  and obtain

$$t_0 + j\bar{\theta} \leq t_0 + 2\bar{\theta} \leq t_0 + 2 \frac{t_f - t_0}{3 + 2 \cdot 1} + \bar{\Delta} = \frac{3}{5}t_0 + \frac{2}{5}t_f + \bar{\Delta} < t_f + \bar{\Delta}.$$

This settles the non-equidistant case:  $\lfloor t_0 + j\bar{\theta} \rfloor_{\mathcal{G}_N} \geq t_0 + j\bar{\theta} - \bar{\Delta}$ . For the equidistant case, we observe

$$\bar{\theta} = \frac{t_f - t_0}{3 + 2\sigma_{\max}} + \frac{1}{2}\bar{\Delta} - \frac{1}{2(3 + 2\sigma_{\max})}\bar{\Delta} = \frac{N\bar{\Delta} + (\sigma_{\max} + 1)\bar{\Delta}}{3 + 2\sigma_{\max}} = \frac{(N + \sigma_{\max} + 1)\bar{\Delta}}{3 + 2\sigma_{\max}}.$$

We look at the right fraction and notice that the numerator consists of a product of an integer and  $\bar{\Delta}$ , whereas the denominator is the integer  $3 + 2\sigma_{\max}$ . Thus, the maximum cut-off by rounding down to the closest grid point is  $\frac{3 + 2\sigma_{\max} - 1}{3 + 2\sigma_{\max}}\bar{\Delta}$ , which is equal to  $\bar{\Delta} - C_1\bar{\Delta}$  and proves the claim.

2. This follows from a similar argumentation as for the claim “1.”. For the non-equidistant case we need only to consider  $t_0 + \frac{5(t_f - t_0)}{3 + 2\sigma_{\max}} \leq t_f$ , for the equidistant case we take again advantage of  $t_f - t_0 = N\bar{\Delta}$ . □

We continue with a lemma that quantifies the length of activation blocks in  $\mathfrak{w}^{\text{MDR}}$ .

**Lemma 6 (Length of activation blocks  $\delta_l$ )** *Consider a feasible control solution for (CIA- $\bar{\theta}$ ) that only uses canonical switches. Then, for the length of its activation block  $\delta_l, 2 \leq l \leq \sigma_{\max}$ , follows:*

$$\delta_l \geq 2\bar{\theta} - \bar{\Delta} + C_1\bar{\Delta}.$$

**Proof** Let  $i$  be the active control on activation block  $l$ . We are using the assumption regarding canonical switches twice. First,  $i$  is forced for the earlier switch  $l - 1$ :

$$\theta_{i,\tau_{l-1}} + a_{i,\tau_{l-1}} \Delta_{\tau_{l-1}} > \bar{\theta} \tag{6.3}$$

and second, it is inadmissible on interval  $\tau_j$ :

$$\theta_{i,\tau_{l-1}} + \sum_{j=\tau_{l-1}}^{\tau_l-1} (a_{i,j} - 1)\Delta_j + (a_{i,\tau_l} - 1)\Delta_{\tau_l} = \theta_{i,\tau_l} + (a_{i,\tau_l} - 1)\Delta_{\tau_l} < -\bar{\theta}. \tag{6.4}$$

By definition of activation blocks we have  $\delta_l = \sum_{j=\tau_{l-1}}^{\tau_l-1} \Delta_j$  so that we obtain by rearranging (6.4):

$$\delta_l > \theta_{i,\tau_{l-1}} + \bar{\theta} + \sum_{j=\tau_{l-1}}^{\tau_l-1} a_{i,j}\Delta_j + (a_{i,\tau_l} - 1)\Delta_{\tau_l}.$$

Plugging (6.3) into the above inequality yields

$$\delta_l > 2\bar{\theta} - a_{i,\tau_{l-1}}\Delta_{\tau_{l-1}} + \sum_{j=\tau_{l-1}}^{\tau_l-1} a_{i,j}\Delta_j + (a_{i,\tau_l} - 1)\Delta_{\tau_l} = 2\bar{\theta} + \sum_{j=\tau_{l-1}+1}^{\tau_l} a_{i,j}\Delta_j - \Delta_{\tau_l} \geq 2\bar{\theta} - \bar{\Delta}$$

which settles the non-equidistant case. For an equidistant grid, we compute

$$2\bar{\theta} - \bar{\Delta} = \frac{2N\bar{\Delta}}{3 + 2\sigma_{\max}} + \bar{\Delta} - C_1\bar{\Delta} - \bar{\Delta} = \frac{2N - 1}{3 + 2\sigma_{\max}}\bar{\Delta},$$

and because  $\delta_l$  is a multiple of  $\bar{\Delta}$ , it follows from  $\delta_l > 2\bar{\theta} - \bar{\Delta}$

$$\delta_l \geq \frac{2N - 1}{3 + 2\sigma_{\max}}\bar{\Delta} + \frac{1}{3 + 2\sigma_{\max}}\bar{\Delta} = 2\bar{\theta} - \bar{\Delta} + C_1\bar{\Delta}.$$

□

Next, we propose in Algorithm 4 a specification of the initial active control  $i_0$ . We observe that a small number of switches on  $\tilde{\mathcal{G}}_N$  in terms of  $\bar{\theta}$  is sufficient, as quantified in the following lemma.

---

**Algorithm 4:** Detecting the initial active control for MDR that results in at most one switch on  $\tilde{\mathcal{G}}_N$ .

---

```

Input : Relaxed control values  $\mathbf{a} \in \mathcal{A}_N$ , where  $n_\omega = 2$ , rounding threshold  $\bar{\theta}$  as in Equation (6.2).
1 if there is a control  $i_1$  with  $\sum_{j=1}^{\bar{N}} a_{i_1,j}\Delta_j \leq \bar{\theta}$  then
2   | Set  $i_0 = i_2 \neq i_1$ ;
3 else if there is a control  $i_1$  with  $\sum_{j=1}^{\bar{N}} a_{i_1,j}\Delta_j \leq 2\bar{\theta} - \bar{\Delta} + C_1\bar{\Delta}$  then
4   | Set  $i_0 = i_1$ ;
5 else
6   | Set  $i_0 = i_1$ , where  $i_1$  is a next forced control on interval  $j = 1$ ;
7 end
8 return:  $i_0$  as initial control;
    
```

---

**Lemma 7** *The MDR algorithm applied to the auxiliary grid  $\tilde{\mathcal{G}}_N$  with rounding threshold  $\bar{\theta}$ ,  $n_\omega = 2$  and  $i_0$  from Algorithm 4 as initial control constructs a control function  $\mathbf{w}^{\text{MDR}}$  that uses at most one switch on  $\tilde{\mathcal{G}}_N$ .*

**Proof** We distinguish between the three possibilities of the initial control in Algorithm 4.

1. If MDR is initialized with  $i_2$  and for  $i_1$  holds  $\sum_{j=1}^{\tilde{N}} a_{i_1,j} \Delta_j \leq \bar{\theta}$ , the latter does not become *forced* on  $\tilde{\mathcal{G}}_N$ . For this reason there is no switch.
2. If  $i_1$  with  $\sum_{j=1}^{\tilde{N}} a_{i_1,j} \Delta_j \leq 2\bar{\theta} - \bar{\Delta} + C_1 \bar{\Delta}$  is the initial active control, a switch has to occur in case  $i_2$  is *forced* on some interval  $\tau_1 \in [\tilde{N}]$ . We need to prove that  $i_1$  does not become *forced* after the first switch, which is equivalent to  $i_2$  does not become inadmissible due to  $n_\omega = 2$  and Lemma 2, because then there is no other switch. For this, we derive a lower bound on the length of the first activation block  $\delta_1$ , where  $i_1$  is active. The control  $i_1$  becomes at the earliest *inadmissible* when it has been active on intervals  $j$  with  $a_{i_1,j} = 0$  whose lengths sum up to be more than  $\bar{\theta}$ , i.e.,  $[t_0 + \bar{\theta}]_{\mathcal{G}_N} - t_0$ . With this observation and Lemma 5.1 we derive

$$\delta_1 = \sum_{j=1}^{\tau_1-1} \Delta_j \geq [t_0 + \bar{\theta}]_{\mathcal{G}_N} - t_0 \stackrel{\text{Lemma 5.1}}{\geq} \bar{\theta} - \bar{\Delta} + C_1 \bar{\Delta}.$$

Note that  $\gamma_{i_1,j}$  is monotonically increasing with increasing interval  $j > \tau_1$  as long as  $i_1$  is inactive, i.e.,  $w_{i_1,j-1} = 0$ . Hence, if we are able to prove  $\gamma_{i_1,\tilde{N}} \leq \bar{\theta}$  in case of  $w_{i_1,j} = 0$  for  $j > \tau_1$ , we also have that  $\gamma_{i_1,j} \leq \bar{\theta}$  for any interval  $j > \tau_1$  meaning there is no second switch. Altogether, we get with the above inequality

$$\gamma_{i_1,\tilde{N}} = \sum_{j=1}^{\tilde{N}} a_{i_1,j} \Delta_j - \delta_1 \leq 2\bar{\theta} - \bar{\Delta} + C_1 \bar{\Delta} - (\bar{\theta} - \bar{\Delta} + C_1 \bar{\Delta}) \leq \bar{\theta}$$

so that  $w^{\text{MDR}}$  switches no more than once on  $\tilde{\mathcal{G}}_N$ .

3. Otherwise we have

$$\sum_{j=1}^{\tilde{N}} a_{i,j} \Delta_j > 2\bar{\theta} - \bar{\Delta} + C_1 \bar{\Delta}, \quad i = i_1, i_2, \tag{6.5}$$

in the *else* case. We can argue similarly as in the previous case, which is why we only have to prove  $\gamma_{i_1,\tilde{N}} \leq \bar{\theta}$ . Since  $i_1$  is a *next forced control* on the first interval, there is an interval  $l \leq \tau_1$  with  $\sum_{j=1}^l a_{i_1,j} \Delta_j > \bar{\theta}$ . This implies the interval  $\tau_1$  of the earliest possible switch is given by

$$\tau_1 = \operatorname{argmin}_{l \in [N]} \left\{ \sum_{j=1}^l (a_{i_1,j} - 1) \Delta_j < -\bar{\theta} \mid \sum_{j=1}^l a_{i_1,j} \Delta_j > \bar{\theta} \right\}$$

from which we find  $\sum_{j=1}^{\tau_1} \Delta_j > 2\bar{\theta}$ . We conclude for the grid point  $t_{\tau_1} = [t_0 + \sum_{j=1}^{\tau_1} \Delta_j]_{\mathcal{G}_N} > [t_0 + 2\bar{\theta}]_{\mathcal{G}_N}$ , which implies  $t_{\tau_1-1} = [t_0 + \sum_{j=1}^{\tau_1-1} \Delta_j]_{\mathcal{G}_N} \geq [t_0 + 2\bar{\theta}]_{\mathcal{G}_N}$ . This is equivalent to  $\sum_{j=1}^{\tau_1-1} \Delta_j \geq [t_0 + 2\bar{\theta}]_{\mathcal{G}_N} - t_0$  and so

$$\delta_1 = \sum_{j=1}^{\tau_1-1} \Delta_j \geq [t_0 + 2\bar{\theta}]_{\mathcal{G}_N} - t_0 \stackrel{\text{Lemma 5.1}}{\geq} 2\bar{\theta} - \bar{\Delta} + C_1 \bar{\Delta}. \tag{6.6}$$

Using the (Conv) property yields  $\sum_{j=1}^{\tilde{N}} a_{i_1,j} \Delta_j = \sum_{j=1}^{\tilde{N}} \Delta_j - \sum_{j=1}^{\tilde{N}} a_{i_2,j} \Delta_j$ , and therefore

$$\begin{aligned} \gamma_{i_1, \tilde{N}} &\leq \sum_{j=1}^{\tilde{N}} a_{i_1,j} \Delta_j - \delta_1 \stackrel{(6.5)}{\leq} \sum_{j=1}^{\tilde{N}} \Delta_j - \sum_{j=1}^{\tilde{N}} a_{i_2,j} \Delta_j - (2\bar{\theta} - \bar{\Delta} + C_1 \bar{\Delta}) \\ &< \tilde{t}_f - t_0 - (2\bar{\theta} - \bar{\Delta} + C_1 \bar{\Delta}) - (2\bar{\theta} - \bar{\Delta} + C_1 \bar{\Delta}) \\ &= \left[ t_0 + \frac{5(t_f - t_0)}{3 + 2\sigma_{\max}} \right]_{\mathcal{G}_N} - t_0 - \frac{4(t_f - t_0)}{3 + 2\sigma_{\max}} \\ &\leq t_0 + \frac{5(t_f - t_0)}{3 + 2\sigma_{\max}} - t_0 - \frac{4(t_f - t_0)}{3 + 2\sigma_{\max}} \\ &= \frac{(t_f - t_0)}{3 + 2\sigma_{\max}} \\ &< \bar{\theta}, \end{aligned}$$

where we used  $\tilde{t}_f = \left[ t_0 + \frac{5(t_f - t_0)}{3 + 2\sigma_{\max}} \right]_{\mathcal{G}_N}$  in the third equation. To conclude, there is also at most one switch. □

The above three lemmata are crucial for the following theorem, which provides an upper bound on (CIA).

**Theorem 3** Consider any grid  $\mathcal{G}_N$ , relaxed values  $\mathbf{a} \in \mathcal{A}_N$  and a maximum number of switches  $\sigma_{\max} > 0$ . The objective of (CIA) is bounded by

$$\theta \leq \frac{N}{3 + 2\sigma_{\max}} \bar{\Delta} + \frac{1}{2} \bar{\Delta} - \frac{C_1}{2} \bar{\Delta}.$$

**Proof** We want to prove that the control function  $\mathbf{w}^{\text{MDR}}$  constructed by MDR with rounding threshold  $\bar{\theta}$  from (6.2) and initial control from Algorithm 4 is feasible and satisfies the claimed bound. We observe  $\bar{\theta} \geq \frac{1}{2} \bar{\Delta}$  from its definition in (6.2) and the definition of  $C_1$  in (6.1). Thus, we can apply Proposition 3 in connection with Lemma 3 so that  $\mathbf{w}^{\text{MDR}}$  fulfills indeed the claimed bound:

$$\theta(\mathbf{w}^{\text{MDR}}) \leq \bar{\theta} = \frac{t_f - t_0}{3 + 2\sigma_{\max}} + \frac{1}{2} \bar{\Delta} - \frac{C_1}{2} \bar{\Delta}.$$

What remains to be shown is that  $\mathbf{w}^{\text{MDR}}$  is a feasible solution for (CIA- $\bar{\theta}$ ), i.e., it does not use more than  $\sigma_{\max}$  switches. In the sequel, we write  $n = \sigma_{\max}$  in variable indices to improve readability of the latter. We assume there are already  $\sigma_{\max}$  switches taken in  $\mathbf{w}^{\text{MDR}}$  and calculate the maximum length of the possibly last activation block, i.e.,  $\delta_{n+1} = t_f - t_{\tau_n-1}$ . In Lemma 7 we have derived that at most one

switch is used on the reduced grid  $\tilde{\mathcal{G}}_N$  until  $\tilde{t}_f$ , but there may follow another switch shortly afterwards, i.e.,  $\tau_2 \geq \tilde{N} + 1$ . For the remaining  $\sigma_{\max} - 2$  activation blocks until  $t_{\tau_n-1}$  we can apply Lemma 6, since Proposition 3 states that MDR uses *canonical* switches for  $n_\omega = 2$ . Lemma 6 states

$$\delta_l \geq 2\theta - \bar{\Delta} + C_1\bar{\Delta}, \quad \text{for } 3 \leq l \leq \sigma_{\max}.$$

Combining these findings and using Lemma 5.2 results in

$$\begin{aligned} t_f - t_{\tau_n-1} &= t_f - \sum_{j=1}^{\sigma_{\max}} \delta_j \leq t_f - (\tilde{t}_f + (\sigma_{\max} - 2)(2\bar{\theta} - \bar{\Delta} + C_1\bar{\Delta})) \\ &\leq t_f - \left\lfloor t_0 + \frac{5(t_f - t_0)}{3 + 2\sigma_{\max}} \right\rfloor_{\tilde{\mathcal{G}}_N} - (\sigma_{\max} - 2)(2\bar{\theta} - \bar{\Delta} + C_1\bar{\Delta}) \\ &\leq t_f - (t_0 + \frac{5(t_f - t_0)}{3 + 2\sigma_{\max}} - \bar{\Delta} + C_1\bar{\Delta}) - (\sigma_{\max} - 2)(2\bar{\theta} - \bar{\Delta} + C_1\bar{\Delta}) \\ &= \frac{(3 + 2\sigma_{\max})(t_f - t_0)}{3 + 2\sigma_{\max}} - \frac{5(t_f - t_0)}{3 + 2\sigma_{\max}} + \bar{\Delta} - C_1\bar{\Delta} - (\sigma_{\max} - 2) \left( \frac{2(t_f - t_0)}{3 + 2\sigma_{\max}} \right) \\ &= \frac{2(t_f - t_0)}{3 + 2\sigma_{\max}} + \bar{\Delta} - C_1\bar{\Delta}. \end{aligned} \tag{6.7}$$

Let  $i$  denote the control that is active after the  $\sigma_{\max}$ th switch of  $\mathbf{w}^{\text{MDR}}$ . Note that  $\theta_{i,j}$  is monotonically decreasing with increasing interval  $j \geq \tau_n$  since  $i$  is chosen to be active on interval  $j$ . Hence, if we are able to show that  $i$  is *admissible* on interval  $N$ , then it is also *admissible* on earlier intervals. We want to prove the admissibility of control  $i$  on interval  $N$  as in this case there will be no further switch until  $N$ . For this, let us assume  $i$  is *inadmissible* on interval  $N$ . We obtain

$$\begin{aligned} -\bar{\theta} &> \theta_{i,N-1} + (a_{i,N} - 1)\Delta_N \\ &= \underbrace{\theta_{i,\tau_n-1} + a_{i,\tau_n}\Delta_{\tau_n}}_{> \bar{\theta}} - \Delta_{\tau_n} + \sum_{j=\tau_n+1}^{N-1} (a_{i,j} - 1)\Delta_j + (a_{i,N} - 1)\Delta_N \\ &> \bar{\theta} + \sum_{j=\tau_n+1}^N a_{i,j}\Delta_j - \sum_{j=\tau_n}^N \Delta_j \\ &\geq \bar{\theta} + 0 - (t_f - t_{\tau_n-1}) \\ &\stackrel{(6.7)}{\geq} \bar{\theta} - \left( \frac{2(t_f - t_0)}{3 + 2\sigma_{\max}} + \bar{\Delta} - C_1\bar{\Delta} \right) \\ &= - \left( \frac{t_f - t_0}{3 + 2\sigma_{\max}} \right) - \frac{1}{2}\bar{\Delta} + \frac{C_1}{2}\bar{\Delta} \\ &= -\bar{\theta}. \quad \quad \quad \zeta \end{aligned}$$

In the second inequality we used that control  $i$  is *forced* on the interval  $\tau_n$  of the  $n$ th, respectively  $\sigma_{\max}$ th, switch. With this contradiction, there cannot be a further switch after  $\tau_n$ ; in other words,  $\mathbf{w}^{\text{MDR}}$  uses at most  $\sigma_{\max}$  switches and is a feasible solution of (CIA). This completes the proof.  $\square$

In the sequel, we are going to elaborate how sharp the upper bound from Theorem 3 is, and we will thereby exclude the case  $\sigma_{\max} \geq N - 1$ ; otherwise the TV constraint would be no longer restrictive. Before presenting the main result in this context, we need a technical lemma again.

**Lemma 8** For  $N, \sigma_{\max} \in \mathbb{N}$ , where  $1 \leq \sigma_{\max} \leq N - 2$ , let  $R \in \mathbb{Q}$  be defined by

$$R := \frac{N}{3 + 2\sigma_{\max}}.$$

We have that

$$2[R]^{0.5} - 1 \leq \left\lfloor \frac{N - [R]}{\sigma_{\max} + 1} \right\rfloor, \quad (6.8)$$

where we indicate by  $[x]^{0.5}$  the rounding up of  $x \in \mathbb{R}$  to the next multiple of 0.5 as defined in Sect. 1.3.

**Proof** Since  $R$  is a rational number with  $3 + 2\sigma_{\max}$  in the denominator, we have

$$[R]^{0.5} \leq \frac{N}{3 + 2\sigma_{\max}} + 0.5 \left( 1 - \frac{1}{3 + 2\sigma_{\max}} \right). \quad (6.9)$$

Moreover, using basic properties of floor and ceiling functions yields

$$[R] \leq [R]^{0.5} + 0.5, \quad (6.10)$$

$$\frac{N - [R]}{\sigma_{\max} + 1} - \frac{\sigma_{\max}}{\sigma_{\max} + 1} \leq \left\lfloor \frac{N - [R]}{\sigma_{\max} + 1} \right\rfloor. \quad (6.11)$$

Next, we calculate

$$\begin{aligned}
 2(\lceil R \rceil^{0.5} - 1) &= \frac{(2\lceil R \rceil^{0.5} - 1)(\sigma_{\max} + 1) - 1}{(\sigma_{\max} + 1)} - \frac{\sigma_{\max}}{\sigma_{\max} + 1} \\
 &= \frac{(2\lceil R \rceil^{0.5} - 1)(\sigma_{\max} + 1) + \lceil R \rceil^{0.5} - 1}{(\sigma_{\max} + 1)} - \frac{\lceil R \rceil^{0.5}}{(\sigma_{\max} + 1)} - \frac{\sigma_{\max}}{\sigma_{\max} + 1} \\
 &= \frac{\lceil R \rceil^{0.5}(3 + 2\sigma_{\max}) - (\sigma_{\max} + 1) - 1}{(\sigma_{\max} + 1)} - \frac{\lceil R \rceil^{0.5}}{(\sigma_{\max} + 1)} - \frac{\sigma_{\max}}{\sigma_{\max} + 1} \\
 (6.9) \quad &\leq \frac{\left(\frac{N}{3+2\sigma_{\max}} + 0.5 - \frac{1}{2(3+2\sigma_{\max})}\right)(3 + 2\sigma_{\max}) - (\sigma_{\max} + 2)}{(\sigma_{\max} + 1)} - \frac{\lceil R \rceil^{0.5}}{(\sigma_{\max} + 1)} - \frac{\sigma_{\max}}{\sigma_{\max} + 1} \\
 &= \frac{N - 0.5}{(\sigma_{\max} + 1)} - \frac{\lceil R \rceil^{0.5}}{(\sigma_{\max} + 1)} - \frac{\sigma_{\max}}{\sigma_{\max} + 1} - \frac{1}{2(\sigma_{\max} + 1)} \\
 (6.10) \quad &\leq \frac{N}{(\sigma_{\max} + 1)} - \frac{\lceil R \rceil}{(\sigma_{\max} + 1)} - \frac{\sigma_{\max}}{\sigma_{\max} + 1} - \frac{1}{2(\sigma_{\max} + 1)} \\
 &= \frac{N - \lceil R \rceil}{(\sigma_{\max} + 1)} - \frac{\sigma_{\max}}{\sigma_{\max} + 1} - \frac{1}{2(\sigma_{\max} + 1)} \\
 (6.11) \quad &\leq \left\lfloor \frac{N - \lceil R \rceil}{\sigma_{\max} + 1} \right\rfloor - \frac{1}{2(\sigma_{\max} + 1)}, \\
 &< \left\lfloor \frac{N - \lceil R \rceil}{\sigma_{\max} + 1} \right\rfloor.
 \end{aligned}$$

Both  $2(\lceil R \rceil^{0.5} - 1) \in \mathbb{Z}$  and  $\left\lfloor \frac{N - \lceil R \rceil}{\sigma_{\max} + 1} \right\rfloor \in \mathbb{Z}$  are valid so that we can deduce from the above inequality

$$2(\lceil R \rceil^{0.5} - 1) \leq \left\lfloor \frac{N - \lceil R \rceil}{\sigma_{\max} + 1} \right\rfloor - 1.$$

□

**Theorem 4** For  $N, \sigma_{\max} \in \mathbb{N}$ , where  $1 \leq \sigma_{\max} \leq N - 2$ , there is an equidistant grid  $\mathcal{G}_N$  and an  $\mathbf{a} \in \mathcal{A}_N$  so that (CIA) has an objective value of

$$\theta \geq \left\lfloor \frac{N}{3 + 2\sigma_{\max}} \right\rfloor^{0.5} \bar{\Delta}. \tag{6.12}$$

**Proof** If  $\sigma_{\max} + 2 \leq N < 3 + 2\sigma_{\max}$ , then we define  $\mathbf{a}$  by specifying the values of control  $i_1$  for the intervals  $j \in [N]$  by

$$a_{i_1,j} = \begin{cases} 1, & \text{if } j \text{ odd,} \\ 0, & \text{if } j \text{ even.} \end{cases}$$

Since there are more intervals  $N$  than the maximum number of switches  $\sigma_{\max}$  plus one, there is an interval  $j$  on which the optimal solution  $\mathbf{w}$  of (CIA) has the value  $w_{i_1,j} = 1$ , while  $a_{i_1,j} = 0$  holds. This results in  $\theta \geq 1\bar{\Delta} = \left\lceil \frac{N}{3+2\sigma_{\max}} \right\rceil^{0.5} \bar{\Delta}$ .

Otherwise, if  $N \geq 3 + 2\sigma_{\max}$ , we proceed as follows

1. We construct a specific matrix  $\mathbf{a}$  that depends on the choice of  $\sigma_{\max}$  and  $N$ .
2. We prove that the MDR scheme constructs for both initial active controls, for this  $\mathbf{a}$  value and with a rounding threshold of

$$\bar{\theta} := \left\lceil \frac{N}{3 + 2\sigma_{\max}} \right\rceil^{0.5} \bar{\Delta} - \epsilon, \quad \text{for any } 0 < \epsilon < \left\lceil \frac{N}{3 + 2\sigma_{\max}} \right\rceil^{0.5} \bar{\Delta}, \tag{6.13}$$

control functions  $\mathbf{w}^{\text{MDR}}$  that use more than  $\sigma_{\max}$  switches. Then, we can come back to the idea of the AMDR scheme and Theorem 2.1., which states that  $\mathbf{w}^{\text{AMDR}}$  is feasible for (CIA), i.e. uses at most  $\sigma_{\max}$  switches, resulting in

$$\left\lceil \frac{N}{3 + 2\sigma_{\max}} \right\rceil^{0.5} \bar{\Delta} \leq \theta(\mathbf{w}^{\text{AMDR}}).$$

Theorem 2 provides in 2. (a) also a statement about the relation to the optimal solution of (CIA):

$$\theta(\mathbf{w}^{\text{AMDR}}) \leq \theta^* + \text{TOL}.$$

Because the tolerance  $\text{TOL}$  can be arbitrarily small, we conclude the optimal solution of (CIA) involves an objective value of at least  $\left\lceil \frac{N}{3+2\sigma_{\max}} \right\rceil^{0.5} \bar{\Delta}$ .

1. We reuse the notation of  $R$  from Lemma 8 and introduce the auxiliary constant  $n_l \in \mathbb{N}$ :

$$R := \frac{N}{3 + 2\sigma_{\max}}, \quad n_l := \left\lfloor \frac{N - [R]}{\sigma_{\max} + 1} \right\rfloor.$$

Next, we are interested in designing a specific  $\mathbf{a} \in \mathcal{A}_N$  with the property to enforce an improper covering by any  $\mathbf{w} \in \Omega_N$  that satisfies a (CIA) objective of at most  $\bar{\theta}$ . By improper covering we indicate  $\mathbf{w} \in \Omega_N$  has to use more than  $\sigma_{\max}$  switches in order to yield the desired (CIA) objective value of at most  $\bar{\theta}$ . We create sets of consecutive intervals for  $\mathbf{a}$  on which either  $\mathbf{a}_{i_1}$  or  $\mathbf{a}_{i_2}$  is set to one (and the other control is thereby set to zero). We call these sets of consecutive intervals with the same value here *index sections*. We generate  $\sigma_{\max} + 2$  index sections, where the two controls are alternately set to one in  $\mathbf{a}$ , with the idea that a feasible solution  $\mathbf{w}$  of (CIA) with at most  $\sigma_{\max}$  switches shall contain at most  $\sigma_{\max} + 1$  activation blocks. The first index section will include  $[R]$  intervals, followed by index sections with  $n_l$  intervals, and the last index section arises from the remaining intervals until  $N$  is reached. After conveying some intuition of the specific  $\mathbf{a} \in \mathcal{A}_N$ , we continue with a technical definition of the index set  $\mathcal{J}$  that specifies the index sections on which  $\mathbf{a}_{i_1}$  is set to one:

$$\begin{aligned} \mathcal{J}_{\text{even}} &:= [ [R] ] \cup \{ j \mid [R] + (2k - 1)n_l + 1 \leq j \leq [R] + 2kn_l, k \in [ \lfloor \sigma_{\max}/2 \rfloor ] \}, \\ \mathcal{J} &:= \begin{cases} \mathcal{J}_{\text{even}}, & \text{if } \sigma_{\max} \text{ is even,} \\ \mathcal{J}_{\text{even}} \cup \{ j \mid [R] + (2\lfloor \sigma_{\max}/2 \rfloor + 1)n_l + 1 \leq j \leq N \}, & \text{if } \sigma_{\max} \text{ is odd.} \end{cases} \end{aligned}$$

With these definitions we introduce  $\mathbf{a}$  by fixing the values of control  $i_1$ .



$$a_{i_1,j} = \begin{cases} 1, & \text{if } j \in \mathcal{J}, \\ 0.5, & \text{if } j = \lceil R \rceil, \text{ and } \lfloor R \rfloor < R \leq \lfloor R \rfloor + 0.5, \\ 1, & \text{if } j = \lceil R \rceil, \text{ and } R > \lfloor R \rfloor + 0.5, \\ 0, & \text{else.} \end{cases} \tag{6.14}$$

The value of  $a_{i_1}$  on the  $(\lceil R \rceil)$ th interval in the second and third case may seem unintuitive. The idea of this construction is that it results  $\theta_{i_1, \lceil R \rceil} = \lceil R \rceil^{0.5}$  if control  $i_1$  is neither active on the first index section, nor on the  $(\lceil R \rceil)$ th interval. In this way, control  $i_1$  needs to be already active on the first index section in order to maintain a (CIA) objective value of at most  $\bar{\theta}$ .

2. We want to prove that the MDR scheme with the rounding threshold from (6.13) and with  $\mathbf{a}$  defined in (6.14) constructs a control function that uses more than  $\sigma_{\max}$  switches, independent of the initial active control. For this, we are going to establish the following claim:

- a) If  $i_1$  is the initial active control, the  $k$ th switch of  $\mathbf{w}^{\text{MDR}}$  happens before the  $(\lceil R \rceil + kn_I)$ th interval, where  $k \in [\sigma_{\max} + 1]$ .
- b) If  $i_2$  is the initial active control, the  $k$ th switch of  $\mathbf{w}^{\text{MDR}}$  happens before the  $(\lceil R \rceil + (k - 1)n_I)$ th interval, where  $k \in [\sigma_{\max} + 1]$ .

Assuming the claim is true,  $\mathbf{w}^{\text{MDR}}$  uses indeed more than  $\sigma_{\max}$  switches because the  $(\lceil R \rceil + (\sigma_{\max} + 1)n_I)$ th interval exists, i.e., is smaller than or equal to  $N$ :

$$\lceil R \rceil + (\sigma_{\max} + 1)n_I = \lceil R \rceil + (\sigma_{\max} + 1) \left\lfloor \frac{N - \lceil R \rceil}{\sigma_{\max} + 1} \right\rfloor \leq \lceil R \rceil + (\sigma_{\max} + 1) \frac{N - \lceil R \rceil}{\sigma_{\max} + 1} = N.$$

The inequality above shows that there are indeed  $\sigma_{\max} + 2$  index sections for  $\mathbf{a}$  as described above. With this information we deduce that  $\bar{\theta} < \frac{1}{2}\bar{\Delta}$  results directly in more than  $\sigma_{\max}$  switches or in control solutions that does not satisfy the claimed optimal (CIA) objective value from (6.12) anyway:

- If  $\mathbf{a}$  consists only of zeros and ones and  $\bar{\theta} < \frac{1}{2}\bar{\Delta}$ , the MDR algorithm creates switches on all intervals  $j$  for which  $\mathbf{a}_{\cdot,j} \neq \mathbf{a}_{\cdot,j-1}$  holds true. Thus, the activation blocks of  $\mathbf{w}^{\text{MDR}}$  would match the index sections of  $\mathbf{a}$ , i.e.  $\mathbf{w}^{\text{MDR}} = \mathbf{a}$ . As we derived  $\sigma_{\max} + 2$  index sections for  $\mathbf{a}$ , there are  $\sigma_{\max} + 2$  blocks for  $\mathbf{w}^{\text{MDR}}$  and therefore  $\sigma_{\max} + 1$  switches.
- If  $a_{i_1, \lceil R \rceil} = 0.5$ , then there is no  $\mathbf{w}$  with  $\theta(\mathbf{w}) < \frac{1}{2}\bar{\Delta}$  regardless of which control is active on interval  $\lceil R \rceil$  since  $\mathbf{a}$  is either zero or one on all other intervals. Hence, we can exclude the case  $\bar{\theta} < \frac{1}{2}\bar{\Delta}$  from further consideration.

Thus, we are left with the case  $\bar{\theta} \geq \frac{1}{2}\bar{\Delta}$ . In this case, we can apply Proposition 3 and conclude that we deal only with canonical switches. We return to prove the claim, and we proceed via induction.

Base case

- (a) We consider  $k = 1$  and conclude from  $N \geq 3 + 2\sigma_{\max}$  that  $\lceil R \rceil^{0.5} \geq 1$  holds. Plugging this into inequality (6.8) from Lemma 8 results in  $\lceil R \rceil^{0.5} < n_l$ , and thus

$$\bar{\theta} < n_l \bar{\Delta}. \tag{6.15}$$

By construction of  $\mathbf{a}$ , the values  $a_{i_1,j}$  are equal to one for  $1 \leq j \leq \lceil R \rceil$ . The value  $a_{i_1,\lceil R \rceil}$  is either 0.5 or 1. Therefore,  $-0.5 \leq \theta_{i_1,\lceil R \rceil} \leq 0$  holds for the accumulated control deviation of  $\mathbf{w}^{\text{MDR}}$  with  $i_1$  as initial active control. After the  $(\lceil R \rceil)$ th interval  $n_l$  intervals follow on which  $a_{i_1,j}$  is zero. We conclude  $i_1$  becomes inadmissible by (6.15) before interval  $\lceil R \rceil + n_l$  and hence, the first switch appears before this interval.

- (b) We show the claim for the first two switches because we take an interest in a switch that occurs after interval  $\lceil R \rceil$  in the induction step. Let  $k = 1$ . Since

$$\sum_{j=1}^{\lceil R \rceil} (a_{i_2,j} - 1) \bar{\Delta} = (\lceil R \rceil - \lceil R \rceil^{0.5} - \lceil R \rceil) \bar{\Delta} < -\lceil R \rceil^{0.5} \bar{\Delta} + \epsilon = -\bar{\theta},$$

we conclude control  $i_2$  becomes inadmissible the latest on interval  $\lceil R \rceil$  when being the initial active control and, equivalently,  $\mathbf{w}^{\text{MDR}}$  has a switch on interval  $\lceil R \rceil$  at the latest. This is equivalent to at least one activation of  $i_1$  up to and including interval  $\lceil R \rceil$ , which we use for proving the assertion in case of  $k = 2$ . Let us assume the second switch happens on or after interval  $\lceil R \rceil + n_l$ . This implies  $i_1$  would be admissible on that interval and we derive

$$\begin{aligned} -\bar{\theta} \leq \theta_{i_1,\lceil R \rceil+n_l} &= \theta_{i_1,\lceil R \rceil} + \sum_{l=\lceil R \rceil+1}^{\lceil R \rceil+n_l} (0-1) \bar{\Delta} \\ &\stackrel{\text{Case } k=1}{\leq} \lceil R \rceil^{0.5} \bar{\Delta} - \bar{\Delta} - n_l \bar{\Delta} \\ &\stackrel{\text{Lemma 8}}{\leq} \lceil R \rceil^{0.5} \bar{\Delta} - \bar{\Delta} - (2\lceil R \rceil^{0.5} - 1) \bar{\Delta} \\ &= -\lceil R \rceil^{0.5} \bar{\Delta} < -\bar{\theta}. \quad \zeta \end{aligned}$$

Consequently, the second switch happens before the  $(\lceil R \rceil + n_l)$ th interval.

**Induction step**

Assume the assertion holds for  $k - 1 \leq \sigma_{\max}$ , we show that it is also true for  $k$ . At first, we prove an auxiliary result. For  $i \in [2]$  and  $j \geq \lceil R \rceil$  we have that

$$\theta_{i,j} = \lceil R \rceil^{0.5} \bar{\Delta} + z \bar{\Delta}, \quad \text{for some } z \in \mathbb{Z}. \tag{6.16}$$

We prove the equation (6.16) by computing the accumulated control deviation:

$$\theta_{i,j} = \bar{\Delta} \left( \sum_{l=1}^{\lceil R \rceil} a_{i,l} + \sum_{l=1+\lceil R \rceil}^j a_{i,l} - \sum_{l=1}^j w_{i,l} \right) = \lceil R \rceil^{0.5} \bar{\Delta} + \left( \sum_{l=1+\lceil R \rceil}^j a_{i,l} - \sum_{l=1}^j w_{i,l} \right) \bar{\Delta}.$$

For  $j > \lceil R \rceil$  we have defined  $a_{i_1,j} \in \{0, 1\}$  so that (6.16) holds with  $z = \left( \sum_{l=1+\lceil R \rceil}^j a_{i_1,l} - \sum_{l=1}^j w_{i_1,l} \right)$ . On the other hand, for the other control  $i_2$  holds

$$\theta_{i_2,j} = \bar{\Delta} \left( \sum_{l=1}^{\lceil R \rceil} a_{i_2,l} + \sum_{l=1+\lceil R \rceil}^j a_{i_2,l} - \sum_{l=1}^j w_{i_2,l} \right) = (\lceil R \rceil - \lceil R \rceil^{0.5})\bar{\Delta} + \left( \sum_{l=1+\lceil R \rceil}^j a_{i_2,l} - \sum_{l=1}^j w_{i_2,l} \right) \bar{\Delta}$$

and therefore (6.16) is satisfied with  $z = \left( \lceil R \rceil - 2\lceil R \rceil^{0.5} + \sum_{l=1+\lceil R \rceil}^j a_{i_2,l} - \sum_{l=1}^j w_{i_2,l} \right)$ .

In order to make use of the established auxiliary result, we need to argue that the  $(k - 1)$ st switch happens after the interval  $\lceil R \rceil$ . In case a) the MDR algorithm will not deactivate  $i_1$  due to  $a_{i_1,j} = 1$  before the  $\lceil R \rceil$ th interval. So it does on the  $\lceil R \rceil$ th interval if  $a_{i_1,\lceil R \rceil} = 0.5$  because we have established  $\bar{\theta} \geq \frac{1}{2}\bar{\Delta}$ . In case b) we use the base case for the second switch. We consider the interval  $\tau_1$  of the first switch of case a) and compare the two accumulated control deviations for the two cases a) and b) on  $\tau_1$  and obtain  $\theta_{i_1,\tau_1}(b) \geq \theta_{i_1,\tau_1}(a)$  because  $i_2$  has already been activated in case b) in contrast to case a). Since  $\tau_1 > \lceil R \rceil$ , we are done.

Now, without loss of generality, let  $i_1$  be the active control after the switch on interval  $\tau_{k-1}$ . We know that  $i_2$  is active and thus admissible on interval  $\tau_{k-1} - 1$ :

$$-\bar{\theta} \leq \theta_{i_2,\tau_{k-1}-1},$$

which implies by Lemma 2 for the control  $i_1$

$$\theta_{i_1,\tau_{k-1}-1} \leq \bar{\theta} = \lceil R \rceil^{0.5} \bar{\Delta} - \epsilon,$$

and by equation (6.16) we have for some  $z_{i_1} \geq 1$

$$\theta_{i_1,\tau_{k-1}-1} = \lceil R \rceil^{0.5} \bar{\Delta} - z_{i_1} \bar{\Delta} \leq (\lceil R \rceil^{0.5} - 1) \bar{\Delta}. \tag{6.17}$$

The control  $i_2$  is inadmissible on interval  $\tau_{k-1}$  as there are only canonical switches. If  $a_{i_2,\tau_{k-1}} = 1$  would be true, then  $i_2$  would already have been inadmissible on interval  $\tau_{k-1} - 1$ . Also,  $a_{i_2,\tau_{k-1}} = 0.5$  is not possible because we derived  $\tau_{k-1} > \lceil R \rceil$ . We conclude  $a_{i_2,\tau_{k-1}} = 0$ . From this and the induction hypothesis, which states that the  $(k - 1)$ st switch appears before the  $(\lceil R \rceil + (k - 1)n_l)$ th interval, follows  $a_{i_1,j} = 1$  for the intervals  $j$  between  $\tau_{k-1}$  and  $(\lceil R \rceil + (k - 1)n_l)$ . Hence,  $\theta_{i_1,\lceil R \rceil+(k-1)n_l} \leq (\lceil R \rceil^{0.5} - 1)\bar{\Delta}$  holds due to (6.17). Finally, we assume  $i_1$  can stay active up to and including interval  $\lceil R \rceil + kn_l$  without becoming *inadmissible*. This and  $a_{i_1,j} = 0$  for  $\lceil R \rceil + (k - 1)n_l + 1 \leq j \leq \lceil R \rceil + kn_l$  imply

$$\begin{aligned}
 -\bar{\theta} &\leq \theta_{i_1, \lceil R \rceil + kn_I} \\
 &= \theta_{i_1, \lceil R \rceil + (k-1)n_I} + \sum_{l=\lceil R \rceil + (k-1)n_I + 1}^{\lceil R \rceil + kn_I} (a_{i_1, l} - 1)\bar{\Delta} \\
 &\leq (\lceil R \rceil^{0.5} - 1)\bar{\Delta} + 0\bar{\Delta} - n_I\bar{\Delta} \\
 &\stackrel{\text{Lemma 8}}{\leq} (\lceil R \rceil^{0.5} - 1)\bar{\Delta} + 0 - (2\lceil R \rceil^{0.5} - 1)\bar{\Delta} \\
 &< -\bar{\theta} \quad \not\leq.
 \end{aligned}$$

Thus,  $i_1$  is *not* active until the  $(\lceil R \rceil + kn_I)$ th interval, respectively with an analogous computation for case b)  $i_1$  is *not* active until the  $(\lceil R \rceil + (k - 1)n_I)$ th interval. Thereby, we showed that indeed the assertion holds for  $k$ . Altogether, the constructed control function  $w^{\text{MDR}}$  uses more than  $\sigma_{\max}$  switches for the chosen rounding threshold  $\bar{\theta}$  so that the (CIA) objective value is at least  $\bar{\theta}$  and we conclude the claimed theorem is true. □

We complete this section by drawing a conclusion from the Theorems 3 and 4.

**Corollary 4** Consider an equidistant grid  $\mathcal{G}_N$ ,  $\mathbf{a} \in \mathcal{A}_N$  and  $1 \leq \sigma_{\max} \leq N - 2$ . The objective of (CIA) is bounded by

$$\theta \leq \frac{N + \sigma_{\max} + 1}{3 + 2\sigma_{\max}} \bar{\Delta}, \tag{6.18}$$

which is the tightest possible bound.

**Proof** The inequality (6.18) is achieved by Theorem 3 applied to the equidistant case and rearranging terms:

$$\theta \leq \left( \frac{N}{3 + 2\sigma_{\max}} + \frac{1}{2} - \frac{1}{2(3 + 2\sigma_{\max})} \right) \bar{\Delta} = \frac{N + \sigma_{\max} + 1}{3 + 2\sigma_{\max}} \bar{\Delta}.$$

It is the tightest possible bound by Theorem 4 and the case  $N = k(3 + 2\sigma_{\max}) + 2 + \sigma_{\max}$ ,  $k \in \mathbb{N}_0$ :

$$\theta \geq \left[ \frac{k(3 + 2\sigma_{\max}) + 2 + \sigma_{\max}}{3 + 2\sigma_{\max}} \right]^{0.5} \bar{\Delta} = (k + 1)\bar{\Delta} = \frac{N + \sigma_{\max} + 1}{3 + 2\sigma_{\max}} \bar{\Delta}.$$

□

### 7 Upper bounds on (CIA) with $n_\omega > 2$

Deriving bounds for the (CIA) problem with more than two controls is more difficult compared to the last chapter as the number of possibilities increases significantly. Let  $\theta^{\max}$  denote the maximum possible objective value of (CIA) for any given  $\mathbf{a} \in \mathcal{A}_N$  and  $\mathcal{G}_N$  in this section. We will first use known results to derive lower and upper bounds for  $\theta^{\max}$ . Then, we dedicate ourselves to the continuous relaxation of (CIA), which allows us to prove a sharper lower bound. Based on this, we state a conjecture about the actual value of  $\theta^{\max}$ .

**Corollary 5** *Let  $1 \leq \sigma_{\max} \leq N - 2$  and  $n_\omega > 2$ . We have that  $\theta^{\max} \geq \frac{N + \sigma_{\max} + 1}{3 + 2\sigma_{\max}} \bar{\Delta}$ .*

**Proof** This bound has been established in Theorem 4 and Corollary 4 for the case  $n_\omega = 2$ . The provided example in the proof of Theorem 4 can be also applied to the case  $n_\omega > 2$  by setting the values of the relaxed controls  $\mathbf{a}_i$ , for all  $i \in [n_\omega]$ ,  $i > 2$ , to zero. □

**Corollary 6** *Let  $1 \leq \sigma_{\max} \leq N - 2$  and  $n_\omega > 2$ . We have that  $\theta^{\max} \leq \frac{2n_\omega - 3}{2n_\omega - 2} \left( \frac{t_f - t_0}{\sigma_{\max} + 1} + \bar{\Delta} \right)$ .*

**Proof** For the (CIA) problem without TV constraints, but with minimum up time constraints the sharp bound

$$\theta \leq \frac{2n_\omega - 3}{2n_\omega - 2} (C_U + \bar{\Delta})$$

is proven in Theorem 2 in [32], where the constant  $C_U \geq 0$  represents the given minimum up time. If we require for (CIA) that an activated control remains active for a time period of at least  $\frac{t_f - t_0}{\sigma_{\max} + 1}$ , at most  $\sigma_{\max}$  switches take place. Thus, the TV constraint serves as a relaxation of the minimum up time constraint. □

We tighten the above results by investigating the continuous version of the (CIA) problem.

**Definition 17 (CCIA)** Let  $\alpha \in \mathcal{A}$  and  $\sigma_{\max} \in \mathbb{N}$  be given. Then, we define the continuous CIA problem to be

$$\min_{\theta, \omega \in \Omega} \theta \tag{7.1}$$

$$\text{s. t. } \theta \geq \pm \int_{t_0}^t (\alpha_i(s) - \omega_i(s)) \, ds, \quad \text{for all } i \in [n_\omega], t \in \mathcal{T}, \tag{7.2}$$

$$\sigma_{\max} \geq TV(\omega). \tag{7.3}$$

Obviously, the problem (CCIA) is a reformulation of

$$\min_{\omega \in \Omega} \max_{t \in \mathcal{T}} \left\| \int_{t_0}^t (\alpha(s) - \omega(s)) \, ds \right\|, \quad \text{s.t. TV constraint(7.3).}$$

We stress that the given data  $\alpha$  lives for (CCIA) in  $\mathcal{A}$  not as usually in  $\mathcal{A}_N$  and analogously we try to find a binary control function  $\omega \in \Omega$ , not  $w \in \Omega_N$ . We obtain a lower bound for the maximum objective  $\theta^{\max}$  of (CCIA) over all  $\alpha \in \mathcal{A}$  by constructing a specific instance, as indicated in the following theorem.

**Proposition 4** *The following lower bound holds true for the maximum objective  $\theta^{\max}$  of (CCIA):*

$$\theta^{\max} \geq \begin{cases} \frac{t_f - t_0}{\sigma_{\max} + 2}, & \text{if } \sigma_{\max} \leq n_{\omega} - 2, \\ \frac{t_f - t_0}{2\sigma_{\max} + 4 - n_{\omega}}, & \text{else.} \end{cases} \tag{7.4}$$

**Proof** Let us first prove the bound for  $\sigma_{\max} \leq n_{\omega} - 2$ . We abbreviate  $\tilde{t} := \frac{t_f - t_0}{\sigma_{\max} + 2}$  and construct the following instance

$$\alpha_i(t) := \begin{cases} 1, & \text{for } t \in [t_0 + (i - 1) \cdot \tilde{t}, t_0 + i \cdot \tilde{t}), \\ 0, & \text{else,} \end{cases} \quad \text{for all } i \in [\sigma_{\max} + 2],$$

$$\alpha_i(t) := 0 \quad \text{for all } i = \sigma_{\max} + 3, \dots, n_{\omega}.$$

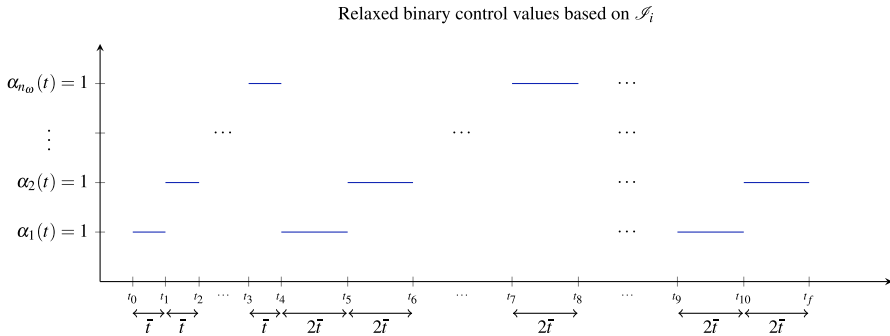
We conclude  $\alpha \in \mathcal{A}$  since the convex combination constraint is satisfied on the whole time horizon  $\mathcal{T}$ . It results  $\int_{\mathcal{T}} \alpha_i(s) \, ds = \tilde{t}$  for all  $i \in [\sigma_{\max} + 2]$ . Thus, we would need to activate each control  $\omega_i$  for some time to achieve a lower objective than  $\tilde{t}$ . But, because the number of switches is restricted to be at most  $\sigma_{\max}$ , this is not possible. Hence,  $\theta \geq \tilde{t}$ .

Next, we consider  $\sigma_{\max} > n_{\omega} - 2$  and we will again construct a specific instance with the claimed objective value of  $\theta \geq \frac{t_f - t_0}{2\sigma_{\max} + 4 - n_{\omega}}$ . We abbreviate  $\tilde{t} := \frac{t_f - t_0}{2\sigma_{\max} + 4 - n_{\omega}}$ . Similar to the above example, we first let the relaxed controls  $\alpha_i$  be active one after the other for a period of  $\tilde{t}$ . After all relaxed controls have been active once and the end of the time horizon has not yet been reached, we activate each control for a period of  $2\tilde{t}$  according to the ascending index  $i \in [n_{\omega}]$  until we reach the end of the time horizon. We express this idea by introducing the domains of activation for all  $i \in [n_{\omega}]$ :

$$\mathcal{I}_i := [t_0 + (i - 1) \cdot \tilde{t}, t_0 + i \cdot \tilde{t}) \cup \{ [t_0 + (n_{\omega} + 2(j - 1))\tilde{t}, t_0 + (n_{\omega} + 2j)\tilde{t}) \mid j \in [\sigma_{\max} + 2 - n_{\omega}], j \equiv i \pmod{n_{\omega}} \}.$$

Based on these domains we define the functions  $\alpha_i(t)$  via

$$\alpha_i(t) := \begin{cases} 1, & \text{for } t \in \mathcal{I}_i, \\ 0, & \text{else,} \end{cases} \quad \text{for all } i \in [n_{\omega}].$$



**Fig. 1** Exemplary visualization of the relaxed binary control function  $\alpha$  from the proof to Proposition 4 resulting in a (CCIA) objective value of  $\bar{t}$ . Since  $i = 2$  is in this example the last activated control, the maximum number of allowed switches is  $\sigma_{\max} = kn_{\omega} + 2 - 2 = kn_{\omega}$ , for  $k \geq 2$

Fig. 1 gives a visualization of this specifically defined control function  $\alpha$ . We have

$$n_{\omega} \bar{t} + (\sigma_{\max} + 2 - n_{\omega})2\bar{t} = t_f - t_0$$

so that  $\cup_{i \in [n_{\omega}]} \mathcal{I}_i = \mathcal{T}$  follows and because the intervals  $\mathcal{I}_i$  are all disjoint, we obtain  $\alpha_i(t) = 1$  for exactly one control  $i$  and for all  $t \in \mathcal{T}$ . Hence,  $\alpha \in \mathcal{A}$ . The next observation about  $\alpha$  is that it consists of  $n_{\omega} + (\sigma_{\max} + 2 - n_{\omega}) = \sigma_{\max} + 2$  activation blocks (interpreted in this continuous setting), meaning there are  $\sigma_{\max} + 1$  changes of the active control. Now, let us assume we can approximate  $\alpha$  with a binary control function  $\omega \in \Omega$  resulting in an (CCIA) objective value of less than  $\bar{t}$ . We have

$$\int_{t_0}^{t_0 + i\bar{t}} \alpha_i(s) \, ds = \bar{t}, \quad \text{for all } i \in [n_{\omega}].$$

So, each control  $\omega_i$ ,  $i \in [n_{\omega}]$  needs to be active for some time until  $t_0 + i\bar{t}$  resulting in at least  $n_{\omega} - 1$  switches up to and including  $t_0 + n_{\omega} \bar{t}$ . Then, we have that

$$\int_{t_0}^{t_0 + n_{\omega} \bar{t}} \alpha_i(s) - \omega_i(s) \, ds < \bar{t}, \quad \text{for all } i \in [n_{\omega}].$$

This, and using that the next activation blocks of  $\alpha$  last for a period of  $2\bar{t}$ , imply each control  $\omega_i$  needs to be activated again up to and including  $t_0 + (n_{\omega} + 2i)\bar{t}$ . If it were possible for some  $i \in [n_{\omega}]$  to skip the activation of  $\omega_i$  without violating the control deviation bound  $\bar{t}$ , this would result in

$$\left| \int_{t_0}^{t_0+(n_\omega+2i)\bar{t}} \alpha_i(s) - \omega_i(s) \, ds \right| < \bar{t}$$

and at the same time, it would hold

$$\int_{t_0+(n_\omega+2(i-1))\bar{t}}^{t_0+(n_\omega+2i)\bar{t}} \omega_i(s) \, ds = 0,$$

which implies

$$\left| \int_{t_0}^{t_0+(n_\omega+2(i-1))\bar{t}} \alpha_i(s) - \omega_i(s) \, ds \right| > \bar{t}$$

because of  $\int_{t_0+(n_\omega+2(i-1))\bar{t}}^{t_0+(n_\omega+2i)\bar{t}} \alpha_i(s) \, ds = 2\bar{t}$ . We apply this argument for all activation blocks of  $\alpha$  until  $t_f$  and conclude that  $\omega$  must use at least one switch for each activation block of  $\alpha$  after  $t_0 + n_\omega\bar{t}$ , i.e., it must use at least  $(\sigma_{\max} + 2 - n_\omega)$  switches. Overall, there are at least  $n_\omega - 1 + (\sigma_{\max} + 2 - n_\omega) = \sigma_{\max} + 1$  switches. Therefore, any  $\omega \in \Omega$  that uses at most  $\sigma_{\max}$  switches involves an (CCIA) objective value of at least  $\bar{t}$ , which settles the lower bound for the case  $\sigma_{\max} > n_\omega - 2$ .  $\square$

(CIA) can be interpreted as a discretized version of (CCIA). Thereby, we can deduce the following corollary.

**Corollary 7** *Let  $1 \leq \sigma_{\max} \leq N - 2$  and  $n_\omega > 2$ . We obtain for the maximum optimal objective value  $\theta^{\max}$  of (CIA):*

$$\theta^{\max} \geq \begin{cases} \frac{t_f - t_0}{\sigma_{\max} + 2}, & \text{if } \sigma_{\max} \leq n_\omega - 2, \\ \frac{t_f - t_0}{2\sigma_{\max} + 4 - n_\omega}, & \text{else.} \end{cases} \tag{7.5}$$

**Proof** (CCIA) is a relaxation of (CIA) since every feasible solution of (CIA) corresponds to a feasible solution of (CCIA). Thus, the claim follows from Proposition 4.  $\square$

The lower bound in Corollary 7 is generally sharp in the sense that there are combinations of  $n_\omega$ ,  $\sigma_{\max}$  and  $\mathcal{G}_N$  so that  $\theta^{\max}$  equals the claimed lower bound. The following example illustrates this relationship.

**Example 4** Let the grid be equidistant with  $N = 3$  and  $n_\omega = 3$ . Consider the following two instances:



$$(a^1_{i,j})_{i \in [3], j \in [3]} := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (a^2_{i,j})_{i \in [3], j \in [3]} := \begin{pmatrix} 1 & 0.5 & 0 \\ 0 & 0.25 & 0.5 \\ 0 & 0.25 & 0.5 \end{pmatrix}.$$

Consider  $\sigma_{\max} = 1$  in the first example. Then,  $\theta = \bar{\Delta}$  and  $\theta^{\max} \geq \bar{\Delta}$  follows from the above corollary. Any asymmetric modification of  $(a^1_{i,j})$  with unequal control accumulation  $\sum_{j=1}^3 a^1_{i_1,j} \neq \sum_{j=1}^3 a^1_{i_2,j}$  would result in a binary control function  $w^{\text{OPT}}$  that activates the controls with highest control accumulation and hence  $\theta < \bar{\Delta}$ . We conclude that the claimed bound is sharp, i.e.,  $\theta^{\max} = \bar{\Delta}$ .

Let us assume  $\sigma_{\max} = 2$  for the second instance. Then,  $w^{\text{OPT}}_{i,j} = 1$  for  $(i,j) = (1, 1), (1, 2), 2, 3)$  and thus  $\theta = 0.75\bar{\Delta}$ . Therefore, the bound in Corollary 7, which amounts to  $\theta^{\max} \geq \frac{3}{5}\bar{\Delta}$ , is not tight for this instance.

Finding the exact value of  $\theta^{\max}$  is difficult due to the nonconvex objective  $\max_a \min_w \max_{i \in [n_\omega], j \in [N]}$  and the tremendously increased number of different  $\omega \in \Omega$  when  $n_\omega > 2$ , but we conjecture that the lower bound in Proposition 4 cannot be improved. We recognize the symmetry of the constructed  $\alpha$  in the proof: Any modification of  $\alpha$  that alters the length of its activation blocks would result either in less than  $\sigma_{\max} + 2$  activation blocks or in at least one block with a smaller length compared with the previous length. The latter block length would be smaller than  $\frac{t_f - t_0}{2\sigma_{\max} + 4 - n_\omega}$  if the block is the first control’s activation, respectively smaller than  $2 \cdot \frac{t_f - t_0}{2\sigma_{\max} + 4 - n_\omega}$  else. With the argumentation from the proof of Proposition 4, this would allow us to choose a control function  $\omega \in \Omega$  with a (CCIA) objective value smaller than  $\frac{t_f - t_0}{2\sigma_{\max} + 4 - n_\omega}$ . Furthermore, we argue that the optimal objective value of (CCIA) is at most by  $\frac{1}{2}\bar{\Delta}$  smaller than the one of (CIA) because the switching times of the optimal  $\omega \in \Omega$  differ at most by one half of the maximum grid length from the optimal  $w \in \Omega_N$ . We close this section by summarizing these thoughts in the following conjecture.

**Conjecture 1** *Let  $1 \leq \sigma_{\max} \leq N - 2$  and  $n_\omega > 2$ . We obtain for (CIA)*

$$\theta^{\max} = \begin{cases} \frac{t_f - t_0}{\sigma_{\max} + 2} + \frac{1}{2}\bar{\Delta}, & \text{if } \sigma_{\max} \leq n_\omega - 2, \\ \frac{t_f - t_0}{2\sigma_{\max} + 4 - n_\omega} + \frac{1}{2}\bar{\Delta}, & \text{else.} \end{cases} \tag{7.6}$$

### 8 Numerical experiments

We test the proposed algorithm with a benchmark example from the <https://mintO.C.de> library [25], with a real-world adsorption cooling machine problem [7] and with generic data. We use the CIA decomposition in order to solve these problems, where we applied CasADi v3.4.5 [1] to parse the nonlinear program (NLP) with

efficient derivative calculation to the solver Ipopt 3.12.3 [30]. We implemented the AMDR algorithm into an add-on as part of the open-source software package `pycombina`<sup>2</sup> [7] and used its BNB solver for benchmarking reasons. The BNB scheme is based on the idea to branch forward in time and exploits that an evaluation of the objective function up to the current grid point yields a valid lower bound that is extremely cheap to compute, see [15, 27] for further details. We set the tolerance parameter of the AMDR algorithm to  $TOL = 0.0001$ . All computational experiments are executed on a workstation with 4 Intel i5-4210U CPUs (1.7 GHz) and 7.7 GB RAM.

## 8.1 Multimode MIOCP

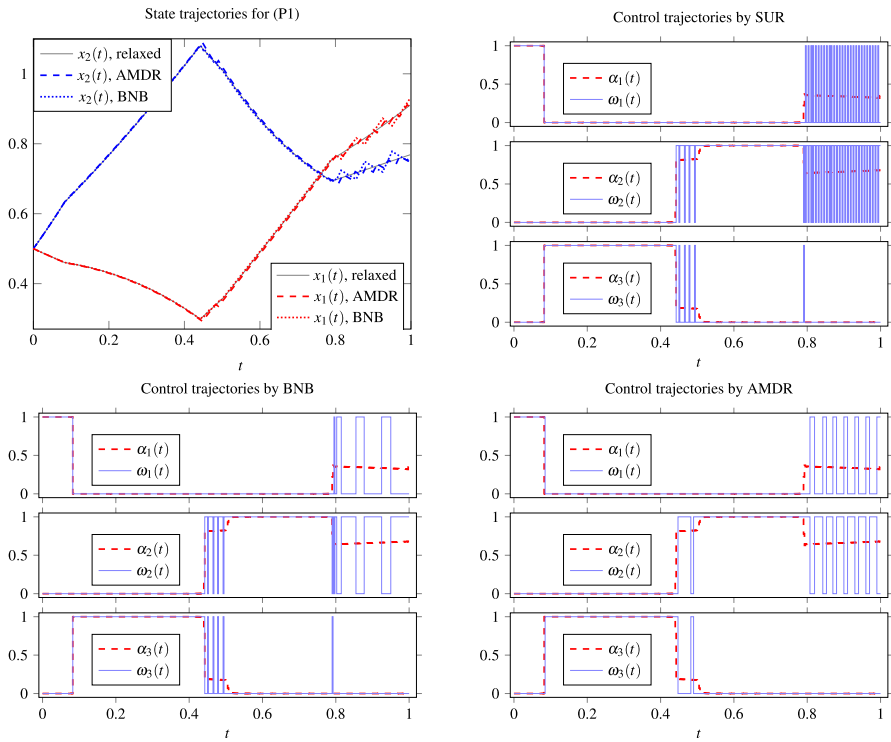
We consider the following MIOCP, which is a modified version of the *Egerstedt standard problem* from <https://mintOC.de>:

$$\begin{aligned} \min_{x, \omega \in \Omega} \quad & x_1(t_f)^2 + x_2(t_f)^2 \\ \text{s.t.} \quad & \text{for a.e. } t \in [0, 1] : \\ & \dot{x}_1(t) = -x_1(t)\omega_1(t) + (x_1(t) + x_2(t))\omega_2(t) + (x_1(t) - x_2(t))\omega_3(t), \\ & \dot{x}_2(t) = (x_1(t) + 2x_2(t))\omega_1(t) + (x_1(t) - 2x_2(t))\omega_2(t) + (x_1(t) + x_2(t))\omega_3(t), \\ & \mathbf{x}(0) = \mathbf{x}_0. \end{aligned} \tag{P1}$$

Obviously, we deal with 3 different modes, i.e.,  $n_\omega = 3$ . We use as initial values  $\mathbf{x}_0 := (0.5, 0.5)^T$ . Furthermore, we add the TV constraints (3.1)–(3.2) to (P1), with varying maximum number of switches  $\sigma_{\max}$ . Fig. 2 illustrates the differential state and control trajectories for  $\sigma_{\max} = 20$  and with relaxed binary controls as well as binary controls based on SUR, BNB and AMDR. We remark that the control function constructed by SUR uses 70 switches and is therefore infeasible with respect to  $\sigma_{\max} = 20$ . The relaxed control values are greater than zero and less than one around  $t \approx 0.45$  and for  $t \geq 0.8$  so that the corresponding approximated state trajectories of BNB and AMDR are slightly different from the relaxed one from  $t \approx 0.45$  on. We set the BNB iteration limit to  $5 \cdot 10^6$  so that it stopped after 15.3s with (CIA) objective value  $\theta = 9.1 \cdot 10^{-3}$  and  $\Phi = 0.991855$  as (P1) objective value. The execution of AMDR took 0.2s and resulted in the improved objective values  $\theta = 4.6 \cdot 10^{-3}$ , respectively  $\Phi = 0.991509$ , which can be explained by more uniformly distributed switches compared with the BNB solution.

Table 1 shows that the BNB algorithm constructs for small instances, e.g.  $N = 200$ , better (CIA) objective values than AMDR if enough time is available. If the BNB scheme finds a good solution, it will usually do so after a few million iterations. While the  $\theta$  values of AMDR are close to the ones from BNB for  $N = 200$ , they are clearly outperforming the latter for bigger instances. Its run time is only slightly increasing with a grid's refinement, from about 0.1 seconds to at most 0.6

<sup>2</sup> see <https://github.com/adbuenger/pycombina>



**Fig. 2** Differential state and control trajectories for the test problem (P1). The problem has been discretized with Multiple Shooting and  $N = 400$  intervals. The state trajectories based on SUR are very similar to the relaxed one (i.e., based on  $\alpha$ ), so that we skip their presentation

seconds. A C++ implementation could still improve the run time as we used so far a prototype implementation in python. It appears that selecting the next-forced control rather than the one with a maximum  $\gamma$  value is beneficial as part of the AMDR algorithm and tends to yield the solution with the smallest (CIA) objective value.

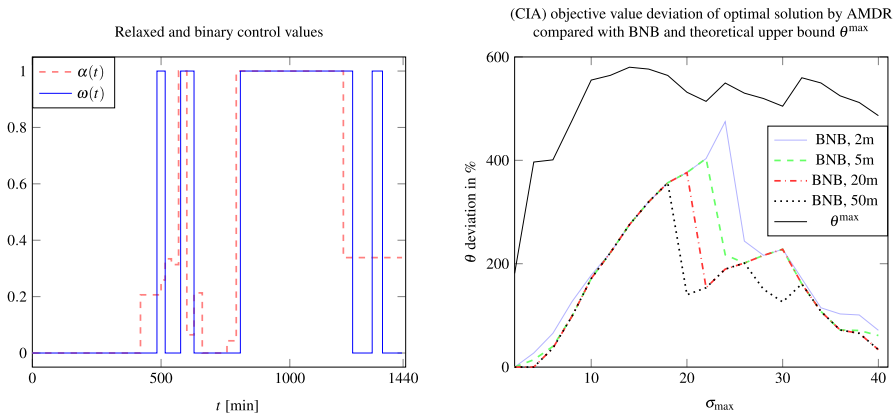
### 8.2 Dualmode adsorption cooling machine problem

In [6, 7], a complex renewable energy system in the form of a solar thermal climate system with nonlinear system behavior is introduced as an MIOCP. The system’s core is an adsorption cooling machine, which can be switched on to intensify the cooling down of ambient temperature. The goal is to control the room temperature in a comfort zone and at the same time to minimize the energy costs. We skip a detailed system’s description but refer to [7], and consider the relaxed binary control values as given, as illustrated in Fig. 3 in the left plot. We assume two modes of the adsorption cooling machine, i.e.,  $n_\omega = 2$ , and a whole day time horizon with control adjustment every four minutes, i.e.,  $N = 360$ .

**Table 1** Comparison of (CIA) objective values and run time of (P1) for different solving methods and varying  $\sigma_{\max}$ .

$\sigma_{\max}$	AMDR		AMDR-NF		BNB, 5 million iter.		BNB, 50 million iter.		Bounds	
	$\theta$	CPU [s]	$\theta$	CPU [s]	$\theta$	CPU [s]	$\theta$	CPU [s]	$\theta^*$	$\theta^{\max}$
<i>N</i> = 200										
5	0.017899	0.11	0.017899	0.26	<b>0.015212</b>	14.35	<b>0.015212</b>	145.22	0.015212	0.09341
10	0.007622	0.12	0.007622	0.23	0.010297	14.29	<b>0.007254</b>	144.99	0.007254	0.05012
20	<b>0.004368</b>	0.16	<b>0.004368</b>	0.22	<b>0.004368</b>	14.73	<b>0.004368</b>	15.03	0.004368	0.02689
30	<b>0.003828</b>	0.1	<b>0.003828</b>	0.15	<b>0.003828</b>	0.58	<b>0.003828</b>	0.56	0.003828	0.01889
40	<b>0.002991</b>	0.13	<b>0.002991</b>	0.17	<b>0.002991</b>	0.0	<b>0.002991</b>	0.0	0.002991	0.01485
<i>N</i> = 400										
5	0.016816	0.21	0.016816	0.27	0.015164	14.67	<b>0.014283</b>	148.12	0.014283	0.09216
10	<b>0.007562</b>	0.28	<b>0.007562</b>	0.35	0.011264	14.34	0.008186	144.9	0.007562	0.04887
20	0.004607	0.18	<b>0.004226</b>	0.26	0.006600	14.28	0.005554	143.56	0.004202	0.02564
30	0.004607	0.18	<b>0.003248</b>	0.29	0.005212	14.36	0.004039	145.64	0.003075	0.01764
40	0.002604	0.31	<b>0.002578</b>	0.24	0.003764	14.56	0.003208	152.65	0.002351	0.01360
<i>N</i> = 800										
5	0.015952	0.58	0.015952	0.58	0.020405	15.21	<b>0.013830</b>	147.94	0.013830	0.09153
10	<b>0.007535</b>	0.47	<b>0.007535</b>	0.56	0.018007	14.87	0.010561	149.56	0.006933	0.04824
20	0.005405	0.47	<b>0.005025</b>	0.54	0.018007	14.96	0.010561	147.74	0.004116	0.02502
30	0.003315	0.4	<b>0.002620</b>	0.47	0.008500	14.69	0.006292	145.04	0.002620	0.01702
40	0.002018	0.57	<b>0.002018</b>	0.59	0.008390	14.39	0.005850	146.8	0.001914	0.01297

AMDR corresponds to Algorithm 3, while AMDR-NF represents a modification in which the *admissible* and *next-forced* control is selected to be active in line 4 of Algorithm 2. By BNB we refer to pycombina's BNB algorithm with depth-first node selection strategy, where we set up an iteration limit to 5 and 50 million nodes. We highlight the best objective values in bold. The last two columns show the optimal objective values and the upper bounds from Conjecture 1

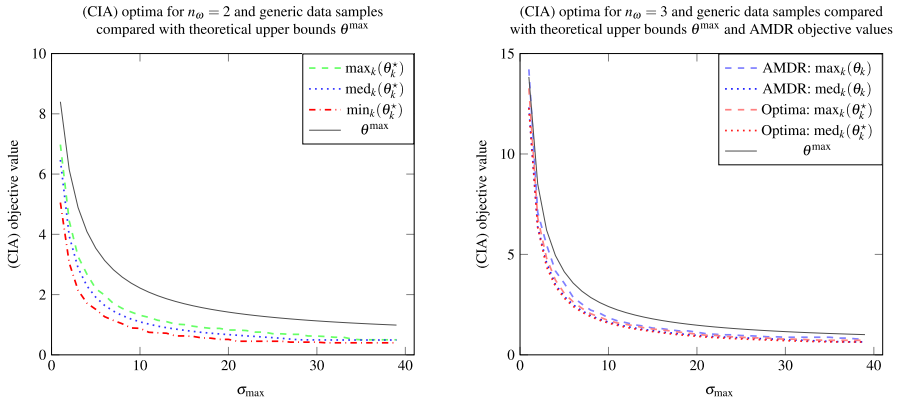


**Fig. 3** Left: Relaxed binary control values  $\alpha$  for the adsorption cooling machine problem for the whole time horizon of a day and exemplary approximated binary values  $\omega$  obtained by AMDR and with  $\sigma_{\max} = 8$ . Right: Comparison of the (CIA) objective values based on the upper bound  $\theta^{\max}$  from Corollary 4 and BNB solutions with the optimal solutions constructed by AMDR and varying  $\sigma_{\max}$  values. We report the deviation in percent. The number next to BNB indicates the maximum number of iterations, e.g. 2 million

We use the AMDR scheme to calculate a candidate solution of the (CIA) problem depending on  $\sigma_{\max}$ , which is optimal by virtue of Theorem 2.2. (a). The right plot in Fig. 3 compares these optimal solutions with the (CIA) objective values of BNB solutions with increasing iteration limit. For a small and large number of allowed switches, the deviation of the BNB solutions is small. One explanation for this is the limited degree of freedom for a small  $\sigma_{\max}$ , so that the width of the BNB tree is very limited. With a large  $\sigma_{\max}$ , on the other hand, solutions with a small  $\theta$  value can be found quickly, with which many nodes can be pruned. The deviation from the optimal solution is particularly striking for medium-sized  $\sigma_{\max}$ . For some instances, especially for  $10 \leq \sigma_{\max} \leq 20$ , an increase of the iteration limit hardly leads to an improvement because the BNB algorithm seems to remain in a suboptimal branch. We also compare the optimal solution of (CIA) with the upper bound from Corollary 4 and see that the latter appears between 200 and 600 percent larger.

### 8.3 Comparison of optima for (CIA) with upper bounds based on generic data

The two investigated MIOCPs showed a relatively large deviation of the optimal (CIA) objective value compared with the derived upper bounds. Therefore, we generated uniformly distributed random values  $\mathbf{a} \in \mathcal{A}_N$  for  $N = 40$  equidistant intervals,  $n_\omega = 2, 3$  controls, and examined how the ratio of these two values results here. We illustrate this comparison in Fig. 4, where we use the upper bound from Corollary 4 for  $n_\omega = 2$  and the one from Conjecture 1 for  $n_\omega = 3$ . The objective values  $\theta, \theta^*$  and bounds  $\theta^{\max}$  decrease logarithmically with the increase of  $\sigma_{\max}$ , as expected. In contrast to the above MIOCPs, the (CIA) objective values come close to the upper bounds, particularly for small  $\sigma_{\max}$ , but a relevant gap remains for larger  $\sigma_{\max}$ . This gap may be further reduced utilizing a larger sample size; we considered here only



**Fig. 4** Optimal objective values  $\theta_k^*$  of (CIA) for randomly generated values  $\mathbf{a} \in \mathcal{A}_N$  with  $k \in [1000]$  samples,  $N = 40$  and varying  $\sigma_{\max}$  values compared with derived upper bounds  $\theta^{\max}$ . We display the maximal, median and minimal objective values of the samples for each given  $\sigma_{\max}$ . Left: The optima are computed with the AMDR algorithm for the case  $n_\omega = 2$ . Right: Comparison of the values constructed by the AMDR algorithm with the optimal values obtained by BNB

1000 (CIA) instances per  $\sigma_{\max}$  value. We also note that the values generated by the AMDR algorithm are very close to the optimal ones.

### 8.4 Discussion

As expected by the polynomial run time complexity, our prototype implementation of AMDR constructs (CIA) feasible solutions very quickly. Their  $\theta$  values are mostly outperforming the ones obtained by the BNB algorithm or are at least close to the latter for a problem with more than two binary controls. Consequently, the AMDR solution is itself a promising (CIA) feasible solution or is a fast option to initialize the BNB with a competitive upper bound. As stated in Remark 6, the AMDR algorithm may also be used to include combinatorial constraints other than the TV constraints.

For comparison with the BNB method, we restrict that we only used the depth-first node selection strategy and could have tuned it a bit more to achieve more competitive feasible solutions of (CIA). Besides, the BNB algorithm can include a variety of combinatorial conditions of the (CIA) problem, so it is generally advantageous.

We also note that our calculations mainly examine the (CIA) objective value because it correlates with the (MIOCP) objective value. With very similar or large (CIA) objective values, however, the smaller value may lead to a worse (MIOCP) objective value—and vice versa. There may be several binary control functions with the same (CIA) objective value but different (MIOCP) objective values. In some instances, we observed that the AMDR algorithm generates a control function with suboptimal (MIOCP) objective value since its switches are structurally delayed compared to the switches on bang-bang-arcs of the relaxed binary values. In this case, we tested, as a heuristic, shifting the AMDR binary values backward in time

by  $[\theta/\bar{\Delta}]$  intervals so that the control function is more similar to the relaxed binary values, which worked well.

## 9 Conclusions

In this paper, we have devised a fast rounding method for the MIOCP with constrained TV of the integer control. The proposed algorithm constructs under certain assumptions, e.g.,  $n_\omega = 2$ , an optimal solution of the (CIA) subproblem. Based on this, we have proven bounds on the integrality gap of (CIA) for the constrained TV case. Our numerical results have shown that the computed control function's quality outperforms in many cases the BNB solution, for which an iteration limit has been set up. Due to the very short run time, we recommend the proposed method, especially for the mixed-integer model predictive control setting or for instances with a vast number of binary variables. In the future, this algorithmic proposal could be compared with a penalty alternating direction method [12] or extended to switching costs as in [3].

**Acknowledgements** The authors thank Felix Bestehorn for his helpful comments on the paper and for all the fruitful discussions on the MDR algorithm. We thank the anonymous reviewers for their constructive comments on earlier drafts of the manuscript which served to improve and clarify the paper. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No 647573) and from Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—314838170, GRK 2297 MathCoRe and SPP 1962.

**Funding** Open Access funding enabled and organized by Projekt DEAL.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Andersson, J.A., Gillis, J., Horn, G., Rawlings, J.B., Diehl, M.: Casadi: a software framework for nonlinear optimization and optimal control. *Math. Program. Comput.* **11**(1), 1–36 (2019)
2. Axelsson, H., Wardi, Y., Egerstedt, M., Verriest, E.: Gradient descent approach to optimal mode scheduling in hybrid dynamical systems. *J. Optim. Theory Appl.* **136**(2), 167–186 (2008)
3. Bestehorn, F., Hansknecht, C., Kirches, C., Manns, P.: A switching cost aware rounding method for relaxations of mixed-integer optimal control problems (2019). [http://www.optimization-online.org/DB\\_FILE/2019/03/7099.pdf](http://www.optimization-online.org/DB_FILE/2019/03/7099.pdf). Accessed 25 Nov 2020
4. Bestehorn, F., Hansknecht, C., Kirches, C., Manns, P.: Mixed-integer optimal control problems with switching costs: A shortest path approach. submitted to *Mathematical Programming B* (2020). [http://www.optimization-online.org/DB\\_HTML/2020/02/7630.html](http://www.optimization-online.org/DB_HTML/2020/02/7630.html). Accessed 25 Nov 2020

5. Bock, H., Plitt, K.: A Multiple Shooting algorithm for direct solution of optimal control problems. In: Proceedings of the 9th IFAC World Congress, pp. 242–247. Pergamon Press, Budapest (1984). <http://www.iwr.uni-heidelberg.de/groups/agbock/FILES/Bock1984.pdf>
6. Buerger, A., Zeile, C., Altmann-Dieses, A., Sager, S., Diehl, M.: An algorithm for mixed-integer optimal control of solar thermal climate systems with MPC-capable runtime (2018). <https://ieeexplore.ieee.org/document/8550424>
7. Buerger, A., Zeile, C., Altmann-Dieses, A., Sager, S., Diehl, M.: Design, implementation and simulation of an mpc algorithm for switched nonlinear systems under combinatorial constraints. *Process Control* **81**, 15–30 (2019). <https://mathopt.de/PUBLICATIONS/Buerger2019.pdf>. Accessed 25 Nov 2020
8. Bürger, A., Zeile, C., Hahn, M., Altmann-Dieses, A., Sager, S., Diehl, M.: pycombina: An open-source tool for solving combinatorial approximation problems arising in mixed-integer optimal control (2020). Submitted
9. De Marchi, A.: On the mixed-integer linear-quadratic optimal control with switching cost. *IEEE Control Syst. Lett.* **3**(4), 990–995 (2019)
10. Doban, A.I., Lazar, M.: A switched systems approach to cancer therapy. In: 2015 European Control Conference (ECC), pp. 2718–2724. IEEE (2015)
11. Gerdt, M.: A variable time transformation method for mixed-integer optimal control problems. *Optim. Control Appl. Methods* **27**(3), 169–182 (2006)
12. Göttlich, S., Hante, F.M., Potschka, A., Schewe, L.: Penalty alternating direction methods for mixed-integer optimal control with combinatorial constraints. arXiv preprint [arXiv:1905.13554](https://arxiv.org/abs/1905.13554) (2019)
13. Graham, R.L., Lawler, E.L., Lenstra, J.K., Kan, A.R.: Optimization and approximation in deterministic sequencing and scheduling: a survey. *Ann. Discrete Math.* **5**, 287–326 (1979)
14. Jung, M.: Relaxations and Approximations for Mixed-Integer Optimal Control. Ph.D. thesis, University Heidelberg (2013). <http://www.ub.uni-heidelberg.de/archiv/16036>. Accessed 25 Nov 2020
15. Jung, M., Reinelt, G., Sager, S.: The lagrangian relaxation for the combinatorial integral approximation problem. *Optim. Methods Softw.* **30**(1), 54–80 (2015)
16. Kirches, C.: Fast numerical methods for mixed-integer nonlinear model-predictive control. Ph.D. thesis, Ruprecht-Karls-Universität Heidelberg (2010). <http://www.ub.uni-heidelberg.de/archiv/11636/>. Accessed 25 Nov 2020
17. Kirches, C., Kostina, E., Meyer, A., Schlöder, M.: Numerical solution of optimal control problems with switches, switching costs and jumps [http://www.optimization-online.org/DB\\_HTML/2018/10/6888.html](http://www.optimization-online.org/DB_HTML/2018/10/6888.html). Accessed 25 Nov 2020
18. Kutta, M.: Beitrag zur näherungsweise Integration totaler Differentialgleichungen. *Zeitschrift für Mathematik und Physik* **46**, 435–453 (1901)
19. Manns, P., Kirches, C.: Improved regularity assumptions for partial outer convexification of mixed-integer pde-constrained optimization problems. *ESAIM: Control, Optimisation and Calculus of Variations* (2019). [http://www.optimization-online.org/DB\\_HTML/2018/04/6585.html](http://www.optimization-online.org/DB_HTML/2018/04/6585.html). Accessed 18 Mar 2020
20. Rieck, R.M.: Discrete controls and constraints in optimal control problems. Ph.D. thesis, Technische Universität München (2017)
21. Ringkamp, M., Ober-Blöbaum, S., Leyendecker, S.: On the time transformation of mixed integer optimal control problems using a consistent fixed integer control function. *Math. Program.* **161**(1), 551–581 (2017). <https://doi.org/10.1007/s10107-016-1023-5>
22. Robuschi, N., Zeile, C., Sager, S., Braghin, F., Cheli, F.: Multiphase mixed-integer nonlinear optimal control of hybrid electric vehicles. *Automatica* **123**, 109325 (2021). <https://doi.org/10.1016/j.automatica.2020.109325>
23. Runge, C.D.T.: Über die numerische Auflösung von Differentialgleichungen. *Math. Ann.* **46**(2), 167–178 (1895)
24. Sager, S.: Numerical methods for mixed-integer optimal control problems. Der andere Verlag, Töning, Lübeck, Marburg (2005). <https://mathopt.de/PUBLICATIONS/Sager2005.pdf>. Accessed 18 Mar 2020
25. Sager, S.: A benchmark library of mixed-integer optimal control problems. In: J. Lee, S. Leyffer (eds.) *Mixed Integer Nonlinear Programming*, pp. 631–670. Springer (2012). <https://mathopt.de/PUBLICATIONS/Sager2012b.pdf>. Accessed 18 Mar 2020
26. Sager, S., Bock, H., Diehl, M.: The Integer Approximation Error in Mixed-Integer Optimal Control. *Mathematical Programming A* **133**(1–2), pp. 1–23 (2012). <https://mathopt.de/PUBLICATIONS/Sager2012a.pdf>. Accessed 18 Mar 2020



27. Sager, S., Jung, M., Kirches, C.: Combinatorial Integral Approximation. *Mathematical Methods of Operations Research* **73**(3): 363–380 (2011). <https://doi.org/10.1007/s00186-011-0355-4>. <https://mathopt.de/PUBLICATIONS/Sager2011a.pdf>. Accessed 18 Mar 2020
28. Sager, S., Reinelt, G., Bock, H.: Direct Methods With Maximal Lower Bound for Mixed-Integer Optimal Control Problems. *Mathematical Programming* **118**(1), 109–149 (2009). <https://mathopt.de/PUBLICATIONS/Sager2009.pdf>. Accessed 18 Mar 2020
29. Tsang, T., Himmelblau, D., Edgar, T.: Optimal control via collocation and non-linear programming. *Int. J. Control* **21**, 763–768 (1975)
30. Wächter, A., Biegler, L.: On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Math. Program.* **106**(1), 25–57 (2006)
31. Zeile, C., Rauwolf, T., Schmeisser, A., Mizerski, J. K., Braun-Dullaues, R. C., Sager, S.: A Personalized Switched Systems Approach for the Optimal Control of Ventricular Assist Devices based on Atrioventricular Plane Displacement (2020). <https://doi.org/10.1101/2020.05.27.119149>
32. Zeile, C., Robuschi, N., Sager, S.: Mixed-integer optimal control under minimum dwell time constraints. *Math. Program.* (2020). <https://doi.org/10.1007/s10107-020-01533-x>
33. Zeile, C., Weber, T., Sager, S.: Combinatorial integral approximation decompositions for mixed-integer optimal control. submitted to *Optimization Methods and Software* (2018). [http://www.optimization-online.org/DB\\_HTML/2018/02/6472.html](http://www.optimization-online.org/DB_HTML/2018/02/6472.html). Submitted

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.