A Gauss-Newton-based Decomposition Algorithm for Nonlinear Mixed-Integer Optimal Control Problems ⋆

Adrian Bürger [a,b] , Clemens Zeile [c] , Angelika Altmann-Dieses [a] , Sebastian Sager [c] , Moritz Diehl [b,d]

[a] *Institute of Refrigeration, Air-Conditioning, and Environmental Engineering (IKKU), Karlsruhe University of Applied Sciences, Moltkestraße 30, 76133 Karlsruhe, Germany.*

[b] *Systems Control and Optimization Laboratory, Department of Microsystems Engineering (IMTEK), University of Freiburg, Georges-Koehler-Allee 102, 79110 Freiburg im Breisgau, Germany.*

[c] *Institute for Mathematical Optimization, Faculty of Mathematics, Otto von Guericke University Magdeburg, Universitätsplatz 2, 39106 Magdeburg, Germany.*

[d] *Department of Mathematics, University of Freiburg, Ernst-Zermelo-Straße 1, 79104 Freiburg im Breisgau, Germany.*

Abstract

For the fast approximate solution of Mixed-Integer Non-Linear Programs (MINLPs) arising in the context of Mixed-Integer Optimal Control Problems (MIOCPs) a decomposition algorithm exists that solves a sequence of three comparatively less hard subproblems to determine an approximate MINLP solution. In this work, we propose a problem formulation for the second algorithm stage that is a convex approximation of the original MINLP and relies on the Gauss-Newton approximation. We analyze the algorithm in terms of approximation properties and establish a first-order consistency result. Then, we investigate the proposed approach considering an illustrative numerical example of Mixed-Integer Optimal Control (MIOC) of a simple nonlinear and unstable system and considering a more complex application that is a numerical case study of MIOC of a renewable energy system. The investigation shows that the proposed formulation can yield an improved integer solution regarding the objective of the original MINLP compared with the Combinatorial Integral Approximation (CIA) algorithm.

Key words: Mixed-integer optimal control; switched nonlinear systems; algorithms and software

1 Introduction

1.1 Background and motivation

Using Mixed-Integer Optimal Control Problems (MIOCPs) in real-time control applications requires the fast solution of Mixed-Integer Non-Linear Programs (MINLPs), cf. Kirches (2011). While general MINLP solvers are usually not suitable for this task (Sager, 2009), approximate fast solutions of such MINLPs can be obtained using a decomposition approach which relies on the solution of a sequence of three comparatively less hard subproblems, cf. Sager (2009); Sager et al. (2011, 2012). The approach has been implemented successfully in various applications such as the control of renewable energy systems (Bürger et al., 2019) and hybrid electric vehicles (Robuschi et al., 2021).

In the second stage of this algorithm, an integer solution for the discrete inputs of the controlled system is computed based on a corresponding relaxed solution determined in the first stage. This is often realized by solving a so-called Combinatorial Integral Approximation (CIA) problem (Sager et al., 2011), which yields an integer approximation with minimum distance to the determined relaxed solution as measured by a dedicated norm. The formulation of this problem, which is a Mixed-Integer Linear Program (MILP) of a special structure, originally relies only on the relaxed integer solution and possibly additional combinatorial constraints for the integer so-

⋆ This paper was not presented at any IFAC meeting. Corresponding author M. Diehl. Tel. +49 761 203 67852

Email addresses: adrian.buerger@h-ka.de (Adrian Bürger), clemens.zeile@ovgu.de (Clemens Zeile), angelika.altmann-dieses@h-ka.de (Angelika Altmann-Dieses), sager@ovgu.de (Sebastian Sager), moritz.diehl@imtek.uni-freiburg.de (Moritz Diehl).

lution (e. g. maximum permissible number of switches (Sager and Zeile, 2021) or minium dwell times (Zeile et al., 2021)), but does not explicitly consider further information of previous algorithm stages, e. g., the effects of the approximated discrete inputs on the system state development.

1.2 Relevant literature

With regard to the CIA decomposition, the inclusion of combinatorial constraints for MIOCPs has been investigated via multibang or total variation regularization (Leyffer and Manns, 2021; Manns, 2021; Sager and Zeile, 2021) and using a penalty alternating direction method (Göttlich et al., 2021). To obtain state constraint feasibility of the rounded solution, a forward simulation of differential states as part of the CIA integer approximation problem was proposed in Lilienthal et al. (2020). Moreover, the steps in the CIA decomposition can be modified and iterated similarly as in the Feasibility Pump algorithm (Fischetti et al., 2005) for MIOCPs for which the feasibility of solutions is difficult to achieve.

Apart from the three-step CIA decomposition, a bilevel approach has been proposed to solve MIOCPs, where the switching sequence and switching times are optimized on the upper and lower level, respectively (Bemporad et al., 2002). The idea of switching time optimization (Axelsson et al., 2008; Gerdts, 2006) is to optimize over the durations of the system modes, which results in a Non-Linear Program (NLP). As part of this approach, switching costs can be included (De Marchi, 2019).

Approaches that approximate MINLPs with Mixed-Integer Quadratic Programs (MIQPs) are computationally advantageous and rely, e.g., on Sequential Quadratic Programming (SQP)-type algorithms (Exler et al., 2012) or quadratic outer approximation (Fletcher and Leyffer, 1994). These approaches are closely related to the one presented in this work.

1.3 Contribution

We present a problem formulation for the second stage of the decomposition algorithm that relies on linearization of the original MINLP. In particular, we derive a convex approximation of the original problem that relies on the Gauss-Newton approximation. The solution of this problem, which is an MIQP, can yield an improved integer solution in terms of the objective and feasibility of the original MINLP as it allows for an explicit consideration of information from the first algorithm stage and how the integer approximation affects the system state. We discuss the approximation properties of the proposed approach and demonstrate the possible advantages first considering an illustrative numerical example of Mixed-Integer Optimal Control (MIOC) of a simple nonlinear and unstable system and then considering a more complex application that is a numerical case study of MIOC of a renewable energy system.

1.4 Outline

The remainder of this work is organized as follows: Section 2 formally introduces the class of MIOCPs considered in this work. Section 3 provides a description of the decomposition algorithm. Section 4 introduces different distance functions for use within the second stage of the decomposition algorithm, and a novel Gauss-Newton (GN)-based distance function. The proposed convex approximation is described in Section 5 and its properties are discussed in Section 6. Sections 7 and 8 state the potential advantages of the proposed approach considering the numerical case studies. Section 9 concludes the study and provides suggestions for future work.

2 Mixed-integer optimal control problems

This section introduces the MIOCP and the derived MINLP formulation considered in this work. For this, we follow the notation and presentation of Rawlings et al. (2020). The concatenation of a number of vectors $a \in \mathbb{R}^{n_a}$ as in

$$\mathbf{a} := \left[a(1)^\top, \dots, a(k)^\top\right]^\top \in \mathbb{R}^{k \cdot n_a} \tag{1}$$

is in the following denoted by the expression

$$\mathbf{a} := \big(a(1), \dots, a(k)\big). \tag{2}$$

2.1 Problem formulation

The algorithms and approaches presented in this work are based on a careful formulation of the MIOCP into an MINLP with a special structure. The MINLP is similar to a standard Optimal Control Problem (OCP) in discrete time; however, the inputs comprise continuous inputs u and integer inputs i, such that the system is described by

$$x^+ = f(x, u, i). \tag{3}$$

Without loss of generality, we restrict ourselves to integers $i \in \mathbb{Z}^{n_i}$ inside a bounded convex polyhedron $P \subset \mathbb{R}^{n_i}$. The polyhedral constraint $i \in P$ allows us to exclude some combinations, e. g., if two machines cannot be operated simultaneously. The polyhedron P can and should be chosen as the convex hull of the admissible integer values in each time step. Also, note that an integer program in which all integer variables are constrained so that the possible choices for each integer variable define a finite set, can be equivalently formulated as a binary program using partial outer convexification, cf. Kirches (2011).

We might have additional combinatorial constraints that couple different time steps with each other. We express these coupling constraints in the form of another polyhedron $\mathbf{P}_{\text{cpl}}$ in the space of integer control trajectories, which are denoted by

$$\mathbf{i} \coloneqq \big(i(0), i(1), \dots, i(N-1)\big) \in \mathbb{R}^{N \cdot n_i}, \tag{4}$$

where N denotes the number of discretization intervals. The overall polyhedron describing all combinatorial constraints imposed on the integer control trajectory is denoted by

$$\mathbf{P} = P^N \cap \mathbf{P}_{\text{cpl}}. \tag{5}$$

Note that the set of integer feasible trajectories is bounded and given by

$$\mathbf{P} \cap \mathbb{Z}^{N \cdot n_i}, \tag{6}$$

i. e., the intersection of $\mathbf{P}$ with the integers. For notational convenience, we also introduce the sequences of predicted states $\mathbf{x}$ and continuous control inputs $\mathbf{u}$ as

$$\mathbf{x} \coloneqq \big(x(0), x(1), \dots, x(N)\big) \in \mathbb{R}^{(N+1) \cdot n_x}, \tag{7}$$
$$\mathbf{u} \coloneqq \big(u(0), u(1), \dots, u(N-1)\big) \in \mathbb{R}^{N \cdot n_u}. \tag{8}$$

Additionally, we introduce a sequence of vectors of slack variables

$$\mathbf{s} \coloneqq \big(s(0), s(1), \dots, s(N)\big) \in \mathbb{R}^{(N+1) \cdot n_s}, \tag{9}$$

which can be utilized to preserve feasibility of an optimization problem in the case of small constraint violations, which can be relevant for practical applications as outlined in Rawlings et al. (2020) and shown later in this work.

We assume that the objective terms consist of a nonlinear least squares term $\frac{1}{2}\left\|F_1(\cdot)\right\|_2^2$ and a nonlinear term $F_2(\cdot)$, i.e., they can be written as

$$\ell(x, u, i, s) = \frac{1}{2}\left\|F_{\ell,1}(x, u, i, s)\right\|_2^2 + F_{\ell,2}(x, u, i, s), \tag{10}$$
$$V_{\text{f}}(x, s) = \frac{1}{2}\left\|F_{\text{f},1}(x, s)\right\|_2^2 + F_{\text{f},2}(x, s). \tag{11}$$

The inequality constraints h and h_{f} are assumed to be generally nonlinear. Then, the MINLP under investigation can be formulated as

$$\min_{\mathbf{x},\mathbf{u},\mathbf{i},\mathbf{s}} \sum_{k=0}^{N-1} \ell\big(x(k), u(k), i(k), s(k)\big) + V_{\text{f}}\big(x(N), s(N)\big) \tag{12a}$$
$$\text{s.t. } x(0) = x_0, \tag{12b}$$
$$x(k+1) = f\big(x(k), u(k), i(k)\big), \quad k = 0, \dots, N-1, \tag{12c}$$
$$h\big(x(k), u(k), i(k), s(k)\big) \leq 0, \quad k = 0, \dots, N-1, \tag{12d}$$
$$h_{\text{f}}\big(x(N), s(N)\big) \leq 0, \tag{12e}$$
$$\mathbf{i} \in \mathbf{P}, \tag{12f}$$
$$\mathbf{i} \in \mathbb{Z}^{N \cdot n_i}. \tag{12g}$$

In the following, we assume that the problem has a feasible solution for any fixed control trajectory $\mathbf{i} \in \mathbf{P}$, considering slack variables for inequality constraint relaxation if necessary.

Without the last constraint (12g), problem (12) would be a standard NLP with optimal control structure complicated only by additional couplings between integer controls that enter via the coupling constraints expressed within the polyhedron $\mathbf{P}$.

2.2 Notational simplifications

To simplify the notation, we collect all continuous variables in the vector

$$z = (\mathbf{x}, \mathbf{u}, \mathbf{s}) \in \mathbb{R}^{n_z} \tag{13}$$

with $n_z = (N+1)n_x + Nn_u + (N+1)n_s$ and the integer variables in the vector

$$y = \mathbf{i} \in \mathbb{R}^{n_y} \tag{14}$$

with $n_y = Nn_i$. In these summarized variable vectors, problem (12) can be compactly written as:

$$\min_{y,z} \; F(y, z) \tag{15a}$$
$$\text{s.t. } G(y, z) = 0 \tag{15b}$$
$$H(y, z) \leq 0 \tag{15c}$$
$$y \in \mathbf{P} \tag{15d}$$
$$y \in \mathbb{Z}^{n_y}. \tag{15e}$$

The objective includes the differentiable nonlinear least squares term $\frac{1}{2}\left\|F_1(\cdot)\right\|_2^2$ and the differentiable nonlinear term F_2 as expressed in

$$F(y, z) = \frac{1}{2}\left\|F_1(y, z)\right\|_2^2 + F_2(y, z). \tag{16}$$

The functions G and H contain differentiable nonlinear equality and inequality constraints, respectively. Although the polyhedron $\mathbf{P}$ is convex, the overall problem (15) is nonconvex owing to the nonlinear least squares term in the objective and constraint functions.

3 Three-step decomposition algorithm

This section provides a brief general description of the decomposition algorithm for fast approximate solution of MINLPs used in this work.

3.1 Preliminaries

We assume that there exists a globally optimal solution of MINLP (15) which we denote by (y°, z°). In most practical applications, the main difficulty is to determine the optimal integer choice y°. If y° would be known, one could just solve an NLP in the remaining variables z to determine the corresponding value z°. This statement can be cleanly expressed if we define the optimal objective function for a fixed $y \in \mathbf{P}$, which we might call the *fixed integer value function*, as follows: $J_{\mathrm{NLP}} : \mathbf{P} \to \mathbb{R}$, $y \mapsto J_{\mathrm{NLP}}(y)$ with

$$\begin{aligned} J_{\mathrm{NLP}}(y) := \min_{z \in \mathbb{R}^{n_z}} \; & F(y,z) && \text{(17a)} \\ \text{s.t. } & G(y,z) = 0, && \text{(17b)} \\ & H(y,z) \leq 0. && \text{(17c)} \end{aligned}$$

The integer part y° of the MINLP solution is the optimal combinatorially feasible value of this function J_{NLP}, i.e.,

$$y^\circ = \arg\min_y J_{\mathrm{NLP}}(y) \text{ s.t. } y \in \mathbf{P} \cap \mathbb{Z}^{n_y}. \tag{18}$$

Note that a lower bound of the optimal MINLP solution can be obtained by the continuous relaxation

$$J_{\mathrm{NLP}}(y^*) = \min_y J_{\mathrm{NLP}}(y) \text{ s.t. } y \in \mathbf{P}. \tag{19}$$

3.2 Description of the decomposition algorithm

The lower bound given by (19) inspires the first step of the three-step decomposition algorithm presented in Sager (2009); Sager et al. (2011, 2012). In the following, a general description of the three algorithm steps S1, S2, and S3 to determine a good feasible solution of MINLP (12) resp. (15) is provided.

S1: In a first step, a *relaxed* version of problem (15), i.e., the NLP that arises if we omit the integer constraint (15e), is solved. The result is a relaxed optimal solution (y^*, z^*) with $y^* \in \mathbf{P}$ but possibly $y^* \notin \mathbf{Z}^{n_y}$.

S2: Afterward, an integer trajectory $y^{**} \in \mathbf{P} \cap \mathbf{Z}^{n_y}$ is heuristically obtained by minimizing some distance function $d(y, y^*)$, i.e., we set

$$y^{**} := \arg\min_{y \in \mathbf{P} \cap \mathbf{Z}^{n_y}} d(y, y^*). \tag{20}$$

S3: Finally, the integer controls are fixed to y^{**} and the restricted NLP (17) is solved to obtain a solution z^{***}.

3.3 Properties of the obtained solution

The result of the algorithm is typically a feasible but suboptimal solution (y^{**}, z^{***}) for the original MINLP. Owing to optimality, the optimal objective values after steps S1 and S3 equal $J_{\mathrm{NLP}}(y^*)$ and $J_{\mathrm{NLP}}(y^{**})$, respectively. In addition, the value of step S1 equals the relaxed minimum of J_{NLP}, i.e., $J_{\mathrm{NLP}}(y^*) = \min_{y \in \mathbf{P}} J_{\mathrm{NLP}}(y)$. Thus, by construction, steps S1 and S3 provide a lower and upper bound of the true MINLP minimum, i.e.,

$$F(y^*, z^*) \leq F(y^\circ, z^\circ) \leq F(y^{**}, z^{***}). \tag{21}$$

In terms of the nonlinear function J_{NLP}, the same fact can be equivalently stated as

$$J_{\mathrm{NLP}}(y^*) \leq J_{\mathrm{NLP}}(y^\circ) \leq J_{\mathrm{NLP}}(y^{**}). \tag{22}$$

The choice of the distance function in step S2 affects both solution quality and computational complexity, which are typically conflicting objectives. In the remainder of this work, we will discuss different choices for step S2, while steps S1 and S3 will be identical in different variants of the algorithm.

Note that the integer gap between the optimal objective function values of (12) and its canonical relaxation, i. e., without (12g), was shown to vanish asymptotically for some $\mathbf{P}$ with the maximal control discretization grid size for affinely entering binary inputs in the dynamics (12c). Thus, any problem (12) can be reformulated via partial outer convexification (Sager, 2009). In the following, we assume that this is the case in (12).

4 Distance functions

In this section, we discuss different choices for the distance function $d(y, y^*)$ in step S2 and introduce a novel Gauss-Newton-based distance function.

4.1 Exact optimal NLP distance function

The perfect choice of distance measure in step S2 would be the suboptimality of $J_{\mathrm{NLP}}(y)$ with respect to the relaxed solution $J_{\mathrm{NLP}}(y^*)$, which we denote by

$$d_{\mathrm{NLP}}(y, y^*) := J_{\mathrm{NLP}}(y) - J_{\mathrm{NLP}}(y^*). \tag{23}$$

This function is nonnegative for all $y \in \mathbf{P}$ and zero for $y = y^*$. The minimizer

$$y^{**}_{\mathrm{NLP}} := \arg\min_{y \in \mathbf{P} \cap \mathbf{Z}^{n_y}} d_{\mathrm{NLP}}(y, y^*) \tag{24}$$

would equal the optimal integer solution y° of the original MINLP. However, the minimization of this distance function is computationally as expensive as solving the original MINLP.

4.2 Norm-based distance functions

A usually much cheaper choice is to use a norm as distance measure. For a given norm $\|\cdot\|_{\text{nrm}}$, one would set

$$d_{\text{nrm}}(y, y^*) := \|y - y^*\|_{\text{nrm}}. \tag{25}$$

Particularly useful are Linear Programming (LP) representable norms such as the L^1 and L^∞ norm, because the problem

$$y^{**}_{\text{nrm}} := \arg \min_{y \in \mathbf{P} \cap \mathbf{Z}^{n_y}} d_{\text{nrm}}(y, y^*) \tag{26}$$

would be an MILP. A particularly successful and widely used norm from this class is the so-called Combinatorial Integral Approximation (CIA) norm that can be defined as

$$\|y - y^*\|_{\text{CIA}} := \|U(y - y^*)\|_\infty \tag{27}$$

for a given y^*, where U is a specially chosen invertible matrix. Each row of this matrix computes an integral over one component of the integer control trajectory from the start to an intermediate time point. In the case of an equidistant grid and a single integer control trajectory, i.e., if $n_i = 1$, the matrix U is an upper triangular matrix with all ones. The MILP to be solved in step S2 of the decomposition algorithm would be given by

$$\min_{\theta \in \mathbb{R}, y \in \mathbb{Z}^{n_y}} \quad \theta \tag{28a}$$

$$\text{s.t.} \; -\mathbf{1}\theta \leq U(y - y^*) \leq \mathbf{1}\theta, \tag{28b}$$

$$y \in \mathbf{P}, \tag{28c}$$

where $\mathbf{1}$ is a vector of all ones and $\mathbf{P}$ would contain a one-hot encoding constraint with regard to the binary controls (Sager et al., 2011). The specific structure of this problem referred to as the CIA problem can be exploited in tailored solution algorithms, cf. Sager et al. (2011), some of which are available in the open-source tool pycombina (Bürger et al., 2020). Several surrogate rounding algorithms have been devised instead of solving the CIA problem, e. g., Sum-Up Rounding (Sager, 2009) and SCARP (Bestehorn et al., 2020). Note that one could define a function

$$J_{\text{CIA}}(y; y^*) := J_{\text{NLP}}(y^*) + \gamma \|y - y^*\|_{\text{CIA}} \tag{29}$$

with arbitrary $\gamma > 0$, which one could interpret as the CIA-approximation of J_{NLP}.

Norm-based distance functions can be advantageous in terms of computational efficiency as the corresponding optimization problems can be solved comparatively fast. However, integer approximations are usually computed based only on y^* as found in step S1 (and possibly additional combinatorial constraints), with limited possibility to consider further information from step S1 or the effects of the integer approximation on the system state.

4.3 Novel Gauss-Newton-based distance function

As an intermediate approach that lies between the aforementioned two in terms of ease of computation and approximation accuracy of J_{NLP}, we propose here to use linearizations of the original MINLP founding at (y^*, z^*). In summary, we will define a convex approximation of J_{NLP} that we denote by $J_{\text{GN}}(y, z; y^*, z^*)$ and which relies on the GN approximation, cf. Rawlings et al. (2020). The resulting optimization problem is an MIQP, for which we define a nonnegative distance function (to be minimized in step S2) as follows:

$$d_{\text{GN}}(y, y^*) := J_{\text{GN}}(y; y^*, z^*) - J_{\text{NLP}}(y^*). \tag{30}$$

The proposed approximation is detailed in the next section.

5 Gauss-Newton integer approximation

To formulate the proposed approximation, we first define the linearization of function G at $(\bar{y}, \bar{z})$ as

$$\begin{aligned} G_{\text{L}}(y, z; \bar{y}, \bar{z}) := G(\bar{y}, \bar{z}) \\ + \frac{\partial G}{\partial (y, z)}(\bar{y}, \bar{z}) \, ((y, z) - (\bar{y}, \bar{z})), \end{aligned} \tag{31}$$

i. e., its first-order Taylor series expansion. Similarly, we denote the linearization of F and H around $(\bar{y}, \bar{z})$ by $F_{\text{L}}(y, z; \bar{y}, \bar{z})$ and $H_{\text{L}}(y, z; \bar{y}, \bar{z})$, respectively. We define the quadratic approximation of F as

$$\begin{aligned} F_{\text{QP}}(y, z; \bar{y}, \bar{z}, B) := F_{\text{L}}(y, z; \bar{y}, \bar{z}) \\ + \frac{1}{2} \begin{bmatrix} y - \bar{y} \\ z - \bar{z} \end{bmatrix}^\top B \begin{bmatrix} y - \bar{y} \\ z - \bar{z} \end{bmatrix}, \end{aligned} \tag{32}$$

where B is a positive semidefinite Hessian approximation. An interesting and favorable positive semidefinite Hessian approximation is given by the squared Jacobian of the function F_1 in the objective, which we call the Gauss-Newton Hessian approximation. This special case of B is given by

$$B_{\text{GN}}(\bar{y}, \bar{z}) := \frac{\partial F_1}{\partial (y, z)}(\bar{y}, \bar{z}) \left(\frac{\partial F_1}{\partial (y, z)}(\bar{y}, \bar{z}) \right)^\top. \tag{33}$$

We define F_{GN} as a special case of F_{QP}, in which we set $B = B_{\text{GN}}$. For a given linearization point $(\bar{y}, \bar{z})$, and matrix B we define the following MIQP approximation of the original MINLP problem that is given by

$$\min_{y,z} \; F_{\text{QP}}(y, z; \bar{y}, \bar{z}, B) \tag{34a}$$

$$\text{s.t.} \; G_{\text{L}}(y, z; \bar{y}, \bar{z}) = 0 \tag{34b}$$

$$H_{\text{L}}(y, z; \bar{y}, \bar{z}) \leq 0 \tag{34c}$$

$$y \in \mathbf{P} \tag{34d}$$
$$(y \in \mathbb{Z}^{n_y}). \tag{34e}$$

This MIQP is solved in step S2 of the decomposition algorithm, with the linearization point chosen to equal the relaxed NLP solution, i.e., with $(\bar{y}, \bar{z}) = (y^*, z^*)$.

The MIQP (34) can be relaxed to a convex Quadratic Program (QP) if we drop the integer constraint (34e). This relaxed QP is an approximation of the relaxed NLP, which comes along with a QP approximation of the function J_{NLP} that we denote by J_{QP} and which is defined by $J_{\text{QP}} : \mathbf{P} \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_z} \times \mathbb{R}^{(n_y+n_z)\times(n_y+n_z)} \to \mathbb{R}, y \mapsto J_{\text{QP}}(y; \bar{y}, \bar{z}, B)$ with

$$J_{\text{QP}}(y; \bar{y}, \bar{z}, B) := \min_{z \in \mathbb{R}^{n_z}} \; F_{\text{QP}}(y, z; \bar{y}, \bar{z}, B) \tag{35a}$$
$$\text{s.t. } G_{\text{L}}(y, z; \bar{y}, \bar{z}) = 0 \tag{35b}$$
$$H_{\text{L}}(y, z; \bar{y}, \bar{z}) \leq 0. \tag{35c}$$

Given the convexity of the relaxed problem (34) in its variables (y, z), the function J_{QP} is convex in y. In fact, being the optimal value function of a parametric QP, it is a continuous and piecewise quadratic function, cf. Bank et al. (1983); Borrelli et al. (2017). As a particular outcome of the QP, we define the Gauss-Newton approximation as

$$J_{\text{GN}}(y; \bar{y}, \bar{z}) := J_{\text{QP}}(y; \bar{y}, \bar{z}, B_{\text{GN}}(\bar{y}, \bar{z})). \tag{36}$$

In step S2, we require $y \in \mathbf{P} \cap \mathbb{Z}^{n_y}$.

6 Finite-dimensional approximation properties

In this section, we illustrate the approximation properties of the proposed approach. For this, we first review classical results of parametric optimization of Chapter 5.2 in Bonnans and Shapiro (2013), which will be utilized afterward.

6.1 Results from parametric optimization

We reconsider the parametric NLP (17) at y, with y being the parameter vector. We consider the linearization of the parametric problem at the optimal solution (y^*, z^*) in the direction $y - y^*$ and define this problem as (LP):

$$J_{\text{LP}}(y; y^*, z^*) := \min_{z \in \mathbb{R}^{n_z}} \; F_{\text{L}}(y, z; y^*, z^*) \tag{37a}$$
$$\text{s.t. } G_{\text{L}}(y, z; y^*, z^*) = 0 \tag{37b}$$
$$S_{\mathbb{A}(y^*, z^*)} H_{\text{L}}(y, z; y^*, z^*) \leq 0, \tag{37c}$$

where $\mathbb{A}(y^*, z^*)$ is the set of active inequality constraints at (y^*, z^*) and $S_{\mathbb{A}(y^*, z^*)}$ selects the active inequality constraints of H_{L} with entries equal to 1 for active constraints and 0 otherwise. The theory in Bonnans and Shapiro (2013) is based on a related LP which we denote as (PL). We define (PL) as (LP) but with $F_{\text{L}}(y, z; y^*, z^*) - F(y^*, z^*)$ as the objective function. Hence, for the corresponding value function holds:

$$J_{\text{PL}}(y - y^*) := J_{\text{LP}}(y; y^*, z^*) - F(y^*, z^*). \tag{38}$$

Let $L(y, \lambda, z)$ be the Lagrangian function associated with NLP (17) at y and let $\Lambda(y, z)$ be the set of feasible Lagrange multipliers. Then, the dual of the linearized problem (PL) can be written as

$$\max_{\lambda \in \Lambda(y^*, z^*)} \; \frac{\partial L}{\partial y}(y^*, \lambda, z^*)(y - y^*) \tag{39}$$

and we use $\mathcal{S}(\text{DL})$ as the set of its optimal solutions. Finally, we need a strong form of a second-order sufficient optimality condition (in the direction $y - y^*$), which states that for all $z - z^* \in C(z^*) \setminus \{0\}$ holds

$$\sup_{\lambda \in \mathcal{S}(\text{DL})} \; \frac{\partial^2 L}{\partial^2 z}(y^*, \lambda, z^*)(z - z^*, z - z^*) > 0, \tag{40}$$

where $C(z^*)$ is the critical cone associated with z^*. With these definitions, we can state the key result from Bonnans and Shapiro (2013), that we need for this paper.

Theorem 1. (Bonnans and Shapiro, 2013). Let $\|\cdot\|$ denote a vector norm. Suppose that

(1) the unperturbed NLP with $y = y^*$ has a unique optimal solution z^*,
(2) the Mangasarian-Fromovitz constraint qualification holds at the point z^*,
(3) the set of feasible Lagrange multipliers $\Lambda(y^*, z^*)$ is nonempty,
(4) the second-order sufficient condition (40) is satisfied,
(5) for all $\|y - y^*\|$ sufficiently small, the feasible set of NLP (17) at y is nonempty and uniformly bounded.

Then, for any optimal solution z of NLP (17) at y holds:

$$\|z - z^*\| = O(\|y - y^*\|), \tag{41}$$
$$J_{\text{NLP}}(y) = J_{\text{LP}}(y; y^*, z^*) + O(\|y - y^*\|^2). \tag{42}$$

Proof. The results follow from Theorem 5.53, (a) and (b), in Bonnans and Shapiro (2013). In comparison with Bonnans and Shapiro (2013), we replaced Gollan's condition with the Mangasarian-Fromovitz condition, which is a stronger condition based on Theorem 5.50, (v), in Bonnans and Shapiro (2013). Moreover, we dropped parameterization with $t \geq 0$ of $y(t)$ and $z(t)$ by considering the directions $y - y^*$ and $z - z^*$ in the optimal solution y^* and z^*. Finally, we express (42) with J_{LP} instead of

J_{PL}, which are equivalent up to the term $F(y^*, z^*)$. Note that we define (PL) here with the constant constraint terms $G(y^*, z^*)$ and $H(y^*, z^*)$ in contrast to Bonnans and Shapiro (2013). As these terms are zero in the optimum (y^*, z^*), the constraints of (PL) reduce to the first derivative parts and, thus, are equivalent to the constraints of the linear problem described in Bonnans and Shapiro (2013). □

6.2 Implications for the proposed algorithm

The following corollary applies the results of Theorem 1 to the QP (35) and the NLP (17).

Corollary 1. Let the assumptions of Theorem 1 hold for the parametric NLP (17) at $y = y^*$ and assume that the parametric QP (35) at $y = \bar{y} = y^*$ has a unique solution and satisfies the second order condition (40). Then, it holds that

$$J_{\mathrm{QP}}(y; y^*, z^*, B) = J_{\mathrm{LP}}(y; y^*, z^*) + O(\|y - y^*\|^2). \quad (43)$$

Furthermore, there is a $C_1 > 0$ and a small $\epsilon_1 > 0$ so that for all $y \in \mathbf{P}_{\epsilon_1} := \{y : \ \|y - y^*\| < \epsilon_1\}$ holds

$$|J_{\mathrm{NLP}}(y) - J_{\mathrm{QP}}(y; y^*, z^*, B)| \leq C_1 \|y^* - y\|^2. \quad (44)$$

Proof. While Theorem 1 was so far formulated for the original NLP (17), denoted by NLP(y), at the point $y = y^*$, we can also apply Theorem 1 to the QP (35), denoted by QP$(y; \bar{y}, \bar{z})$, which depends on two sets of parameters, on y and on $(\bar{y}, \bar{z})$, but where we assume $(\bar{y}, \bar{z})$ as fixed to (y^*, z^*). Thus, QP$(y; y^*, z^*)$ is a parametric problem of the same form as NLP(y) if we only regard its parametric dependence on y. We regard this parametric QP problem again at the point $y = y^*$. To apply Theorem 1 to the QP$(y; y^*, z^*)$, we use the fact that all first order terms of the NLP(y^*) and of the QP$(y^*; y^*, z^*)$ coincide, so assumptions (2), (3), and (5) in Theorem 1 hold because they hold for the original NLP, and by the explicit assumptions also (1) and (4) hold. Also, the corresponding LPs (37) are identical. Thus, (43) follows from the application of Theorem 1 to the QP$(y; y^*, z^*)$ at $y = y^*$.

The inequality (44) follows from the first claim (43) and Theorem 1, (42), by combining these results and expressing the $\mathcal{O}$-notation with the constant C_1. □

We remark that the $\mathcal{O}$ notation in the context of integers can be considered as conflicting as the integers y might not be arbitrarily close to y^*. However, we conjecture this could be achieved by applying a finer discretization grid. Based on the aforementioned corollary, we can state the main theoretical result for our algorithm, which states that the constructed solution y^{**} is correct up to first-order terms with regard to the size of the deviation of the integer controls.

Theorem 2. Let the assumptions of Theorem 1 hold for the parametric NLP (17), with the optimal solution (y^*, z^*) of the relaxed NLP. Consider the constant C_1 from Corollary 1 for the set $\mathbf{P}_{\epsilon_1}$. Let $y^{**}, y^\circ \in \mathbf{P}_{\epsilon_1}$ be the optimal solutions for y of the MIQP (34), with $\bar{y} = y^*, \ \bar{z} = z^*$, and the MINLP (15), respectively. Then, we have

$$\begin{aligned} &|J_{\mathrm{NLP}}(y^{**}) - J_{\mathrm{NLP}}(y^\circ)| \\ &\quad \leq C_1 \left(\|y^* - y^{**}\|^2 + \|y^* - y^\circ\|^2 \right). \end{aligned} \quad (45)$$

Proof. We abbreviate in the following $J_{\mathrm{QP}}(y) := J_{\mathrm{QP}}(y; y^*, z^*, B)$. First, note that owing to optimality of y^{**} and y° in the MIQP (34) and MINLP (15), respectively, it holds that

$$J_{\mathrm{QP}}(y^{**}) - J_{\mathrm{QP}}(y^\circ) \leq 0. \quad (46)$$
$$J_{\mathrm{NLP}}(y^{**}) - J_{\mathrm{NLP}}(y^\circ) \geq 0. \quad (47)$$

We use this to derive the result:

$$\begin{aligned} |J_{\mathrm{NLP}}(y^{**}) - J_{\mathrm{NLP}}(y^\circ)| &\overset{(47)}{=} J_{\mathrm{NLP}}(y^{**}) - J_{\mathrm{NLP}}(y^\circ) \\ &= J_{\mathrm{NLP}}(y^{**}) - J_{\mathrm{NLP}}(y^\circ) + J_{\mathrm{QP}}(y^{**}) \\ &\quad - J_{\mathrm{QP}}(y^{**}) + J_{\mathrm{QP}}(y^\circ) - J_{\mathrm{QP}}(y^\circ) \\ &\overset{(46)}{\leq} J_{\mathrm{NLP}}(y^{**}) - J_{\mathrm{QP}}(y^{**}) + J_{\mathrm{NLP}}(y^\circ) - J_{\mathrm{QP}}(y^\circ) \\ &\leq |J_{\mathrm{NLP}}(y^{**}) - J_{\mathrm{QP}}(y^{**})| + |J_{\mathrm{NLP}}(y^\circ) - J_{\mathrm{QP}}(y^\circ)| \\ &\leq C_1 \left(\|y^* - y^{**}\|^2 + \|y^* - y^\circ\|^2 \right), \end{aligned}$$

where we apply in the last inequality Corollary 1 twice for y^{**} and y°. □

The above result holds for any positive semidefinite matrix B and in particular for the GN-MIQP. Using a suitable vector norm, we may have consistency of order one in the control grid size. By iteratively refining the control grid size, one could possibly establish an asymptotic convergence to the optimal solution, similar to the CIA decomposition (Jung et al., 2015) but here with quadratic convergence in the grid size. In fact, numerical experiments confirm that the difference from inequality (44) with y^{**} used for y becomes quadratically smaller with finer control discretization, as illustrated in Appendix C. The asymptotic convergence theory can be considered for future studies.

6.3 On the choice of the Hessian approximation

The theoretical results mentioned in the previous section hold regardless of the choice of the Hessian approxima-

tion B, as long as it is chosen to be positive semidefinite. However, in practice, the choice is quite relevant. The exact Hessian matrix of the Lagrangian would possibly ensure second-order consistency, however, at the price of expensive computation and possibly indefiniteness. This is in contrast to the pure linearization with $B = 0$, which is inexpensive, but often leads to unfavorably large steps as numerical experiments show. The choice of the GN-approximation appears as a good compromise solution, which is computationally relatively cheap and represents a good approximation of the Hessian matrix.

7 MIOC of a nonlinear unstable system

In the following, potential advantages of the proposed approach are illustrated within a numerical case study of MIOC of a simple nonlinear and unstable system (Rawlings et al., 2020).

7.1 Setup and implementation

We consider a simple MIOCP of the form (12) for a nonlinear unstable system with one state $x \in \mathbb{R}$ and one binary control $b \in \{0, 1\}$. The continuous time system is described by

$$\dot{x} = x^3 - b \tag{48}$$

and transformed to a discrete time system

$$x^+ = f(x, b) \tag{49}$$

using one Runge-Kutta (RK)-4 step of step length $h = 0.05$. The aim is to track a reference $x_{\text{ref}} = 0.7$ starting from the initial value $x_0 = 0.8$ on a horizon of length $N = 30$, resulting in the following MINLP:

$$\min_{\mathbf{x},\mathbf{b}} \ \Phi(\mathbf{x}) = \frac{1}{2}\sum_{k=0}^{N}(x(k) - x_{\text{ref}})^2 \tag{50a}$$

$$\text{s.t. } x(0) = x_0, \tag{50b}$$

$$x(k+1) = f(x(k), b(k)), \quad k = 0, \dots, N-1, \tag{50c}$$

$$\mathbf{b} \in \mathbf{B} \cap \mathbb{Z}^N. \tag{50d}$$

The combinatorial constraint set $\mathbf{B}$ imposes a minimum uptime constraint that requires that b remains active for at least three consecutive time steps, i.e., we have

$$\begin{aligned}\mathbf{B} = \{\mathbf{b} \in [0,1]^N \mid \\ b(k) \geq b(k-1) - b(k-2), \\ b(k) \geq b(k-1) - b(k-3), \\ k = 0, \dots, N-1\}.\end{aligned} \tag{51}$$

The required previous values $b(-1)$, $b(-2)$, $b(-3)$ are all set to zero. Further details on minimum dwell time constraints in the MIOCP context can be found in Zeile et al. (2021).

Table 1
Objective values of (50) for different solution approaches.

Solution approach	Objective value
Relaxed MINLP	$\Phi(\mathbf{x}^*) = 8.97 \cdot 10^{-3}$
Exact integer solution	$\Phi(\mathbf{x}^\circ) = 2.07 \cdot 10^{-2}$
CIA	$\Phi(\mathbf{x}^{***}_{\text{CIA}}) = 1.32 \cdot 10^{-1}$
GN-MIQP	$\Phi(\mathbf{x}^{***}_{\text{GN}}) = 2.07 \cdot 10^{-2}$

Then, the described problem is solved by applying the decomposition algorithm once using the CIA-MILP and once using the GN-MIQP. For this, the MINLP is implemented using CasADi (Andersson et al., 2019), which provides the possibility for the simplified generation of the derivatives of the MINLP components required for solving the NLP and setting up the GN-MIQP. The NLP stage in step S1 of the decomposition algorithm is solved using Ipopt (Wächter and Biegler, 2006). Like in Bürger et al. (2021), we neglect combinatorial constraints in step S1. The CIA problem is solved using a tailored branch-and-bound algorithm available in pycombina (Bürger et al., 2020), the GN-MIQP is solved using Gurobi (Gurobi Optimization, LLC, 2021). Owing to the absence of continuous controls, step S3 for this example just amounts to a system simulation. Moreover, the problem is solved for comparison using a branch-and-bound style simulation procedure described in Appendix A which has been implemented in Python and allows to determine the globally optimal solution.

7.2 Numerical results

Table 1 lists the objective values for (50) obtained using the described solution approaches and the objective value of the relaxed problem after step S1. It shows that for this setup, the GN-MIQP approach was able to achieve an improved solution compared with the CIA approach in terms of the MINLP objective. The solution obtained by the GN-MIQP approach here even is a globally optimal solution as the obtained objective value corresponds to the exact integer solution obtained by the branch-and-bound style simulation.

The optimized state and control obtained by the GN-MIQP and CIA approach are given in Fig. 1. The reference value x_{ref} of the state is indicated by a dashed grey line, the (identical) relaxed solution (x^*, b^*) after step S1 by a blue solid line, and the integer solutions $(x^{***}_{\text{GN}}, b^{**}_{\text{GN}})$ and $(x^{***}_{\text{CIA}}, b^{**}_{\text{CIA}})$ after step S3 by dashed green lines and dash-dotted orange lines, respectively. For the relaxed solution, it can be observed that higher control effort is applied at the beginning of the control horizon to drive the system state towards the reference value, and afterward the state reference value is perfectly tracked by constantly applying a control input of $b^*_k \approx 0.34$, $k = 4, \dots, N-1$.

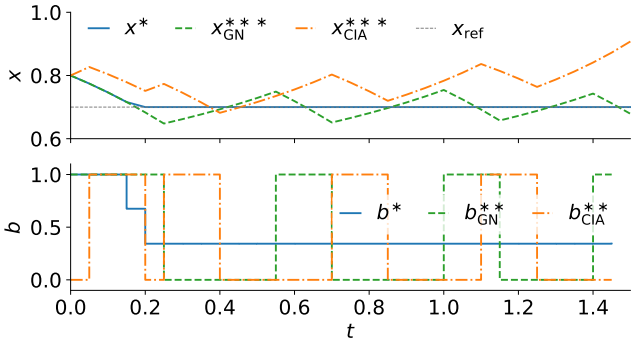

Fig. 1. Relaxed and binary feasible solution for the GN-MIQP and the CIA approach.

By comparing the state and control trajectories shown in Fig. 1, the influence of the obtained integer approximation on the state trajectory development can be observed. While the integer solution obtained using the CIA approach is a globally optimal solution of the approximation problem in terms of the objective function (28a), an application of b^{**}_{CIA} for control of the unstable system causes the system state to diverge from its reference x_{ref}. In contrast, the integer solution obtained using the GN-MIQP approach can maintain the state in a region around the reference value.

Fig. 2 shows the values of $J_{\mathrm{NLP}}\big(\bar{b}(\alpha_{\mathrm{GN}}, \alpha_{\mathrm{CIA}})\big)$, of $J_{\mathrm{QP}}\big(\bar{b}(\alpha_{\mathrm{GN}}, \alpha_{\mathrm{CIA}}); b^*, z^*, B_{\mathrm{GN}}\big)$ which uses the GN Hessian, of $J_{\mathrm{QP}}\big(\bar{b}(\alpha_{\mathrm{GN}}, \alpha_{\mathrm{CIA}}); b^*, z^*, 0\big)$ which uses an all-zero Hessian, as well as the values of $J_{\mathrm{CIA}}\big(\bar{b}(\alpha_{\mathrm{GN}}, \alpha_{\mathrm{CIA}}); b^*\big)$ for (50), with

$$\begin{aligned}\bar{b}(\alpha_{\mathrm{GN}}, \alpha_{\mathrm{CIA}}) = b^* &+ \alpha_{\mathrm{GN}}(b^{**}_{\mathrm{GN}} - b^*) \\ &+ \alpha_{\mathrm{CIA}}(b^{**}_{\mathrm{CIA}} - b^*)\end{aligned} \tag{52}$$

evaluated for $(\alpha_{\mathrm{GN}} \in [0,1], \alpha_{\mathrm{CIA}} = 0)$, $(\alpha_{\mathrm{GN}} = 0, \alpha_{\mathrm{CIA}} \in [0,1])$, and $\alpha_{\mathrm{GN}} + \alpha_{\mathrm{CIA}} = 1$. The plot illustrates the piecewise linear approximation yielded by the CIA approach and the piecewise quadratic approximation yielded by the GN-MIQP approach.

Fig. 3 compares the runtimes and relative deviation of $J_{\mathrm{NLP}}(b^{**})$ and $J_{\mathrm{NLP}}(b^*)$ achieved by the CIA and GN-MIQP approach for a varying number of discretization intervals N (as annotated next to the data points) with minimum uptime constraints and step length adjusted to the corresponding number of discrete time steps. For comparability, the objective function is for all scenarios evaluated at the original $N + 1 = 31$ discrete time points. It can be observed that the runtime of the GN-MIQP approach is always higher than the CIA runtime and increases faster. However, the graph also shows that for this case, the optimality of the results achieved using the CIA approach is of higher variability than the results obtained using the GN-MIQP approach, which also exhibits less deviation for the considered problem

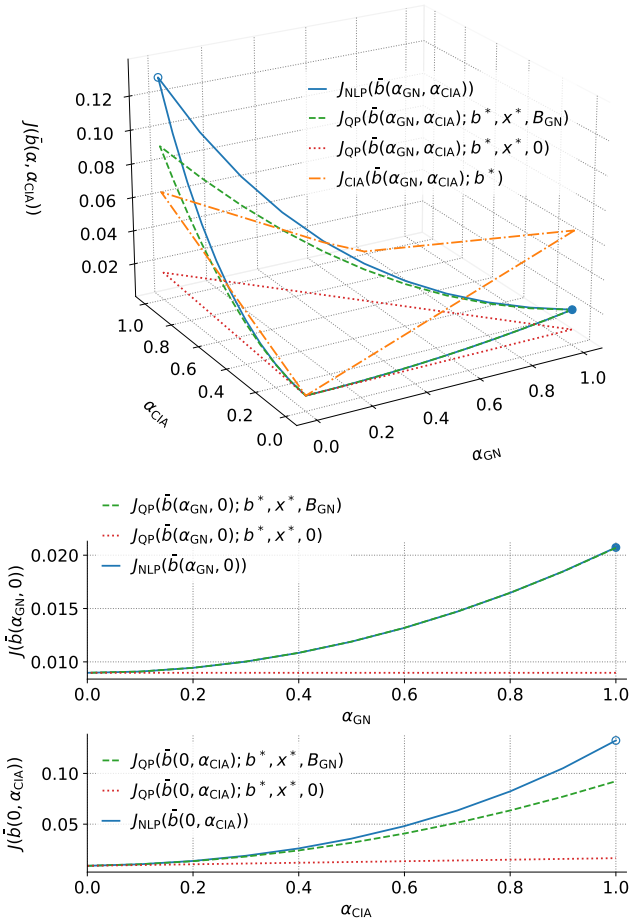

Fig. 2. Evaluation of the functions $J_{\mathrm{NLP}}\big(\bar{b}(\alpha_{\mathrm{GN}}, \alpha_{\mathrm{CIA}})\big)$, $J_{\mathrm{QP}}\big(\bar{b}(\alpha_{\mathrm{GN}}, \alpha_{\mathrm{CIA}}); b^*, x^*, B_{\mathrm{GN}}\big)$, $J_{\mathrm{QP}}\big(\bar{b}(\alpha_{\mathrm{GN}}, \alpha_{\mathrm{CIA}}); b^*, x^*, 0\big)$, and $J_{\mathrm{CIA}}\big(\bar{b}(\alpha_{\mathrm{GN}}, \alpha_{\mathrm{CIA}}); b^*\big)$. The value of $J_{\mathrm{NLP}}(b^{**}_{\mathrm{GN}})$ is indicated by the symbol $\bullet$, the value of $J_{\mathrm{NLP}}(b^{**}_{\mathrm{CIA}})$ by the symbol $\circ$.

instances. Although one would expect the deviation to decrease for the CIA approach when the grid is refined, cf. Sager (2009); Sager et al. (2011), note that the optimal solution of a CIA problem might not be unique, so that different solutions that yield the same optimal CIA objective might results in different objective values for the original MINLP. This is especially important in the presence of constraints that limit the switching possibilities of the optimal solution, such as minimum uptime constraints, which can play a significant role regarding the optimal objective value of a CIA problem.

8 MIOC of a renewable energy system

In this section, the proposed approach is applied to a more complex application that is a numerical case study of MIOC of a renewable energy system.

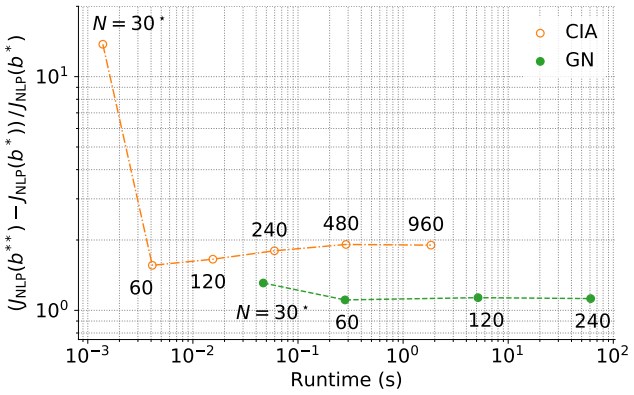

Fig. 3. Runtime and relative deviation between $J_{\mathrm{NLP}}(b^{**})$ and $J_{\mathrm{NLP}}(b^*)$ for the CIA and GN-MIQP approaches for a varying number of discrete intervals N, as annotated at the corresponding data points. The values of $N = 30$ (annotated by the symbol $\star$) correspond to the values displayed in Fig. 2.

8.1 Setup of the case study

The system considered in this study is a solar thermal climate system located at Karlsruhe University of Applied Sciences. It comprises an Adsorption Cooling Machine (ACM) which can produce cooling energy to cover heat loads of a university building in two different operation modes: in Adsorption Cooling (AC) mode ($b_{\mathrm{ac}} = 1$), solar thermal heat can be used to drive the ACM; in Free Cooling (FC) mode ($b_{\mathrm{fc}} = 1$), the recooler of the ACM can be used to directly cool down the medium at the ambient in times of low ambient temperature. Heating and cooling energy can be stored in dedicated stratified storages. A detailed system description is given in Bürger et al. (2021).

An experimental operation of the system using MIOC and CIA was carried out in previous studies (Bürger, 2020; Bürger et al., 2021). It showed that insufficient quality of approximated binary controls can lead to less efficient system operation, e. g., through activation of ACM operation modes at unsuitable time periods (Bürger, 2020). In this case study, the GN-MIQP approach is applied to solve an optimal control problem for the system to achieve efficient system operation in accordance with defined constraints regarding the system's temperature states. The results are compared to those obtained using the CIA approach with regard to solution quality and runtime. For this, we use a variant of the existing model and associated constraints presented in Bürger et al. (2021) as outlined in the following section.

8.2 System modeling

The model used in this study is described by a set of Ordinary Differential Equations (ODEs) as a switched nonlinear system

$$\begin{aligned}\frac{\mathrm{d}x(t)}{\mathrm{d}t} &= f_0\big(x(t), u(t), c(t)\big) \\ &\quad + \sum_{i=1}^{n_b} b_i \cdot f_i\big(x(t), u(t), c(t)\big)\end{aligned} \tag{53}$$

with differential states $x(t) \in \mathbb{R}^{n_x}$, continuous controls $u(t) \in \mathbb{R}^{n_u}$, binary controls $b(t) \in \{0,1\}^{n_b}$, and time-varying parameters $c(t) \in \mathbb{R}^{n_c}$, with $n_x = 20$, $n_u = 5$, $n_b = 2$, and $n_c = 4$. A schematic is depicted in Fig. 4. For simplicity, the building part of the model in Bürger et al. (2021) is replaced by a simplified load model that facilitates the application of a specific cooling load profile $\dot{Q}_{\mathrm{lc}}$ to the system. In particular, the temperature T_{lc} of the medium returning from the fan coil units and corresponding water mass flow $\dot{m}_{\mathrm{lc}}$ are calculated as

$$T_{\mathrm{lc}}(t) = T_{\mathrm{lts},n_{\mathrm{lts}}}(t) - \Delta T_{\mathrm{lc}}, \tag{54}$$

$$\dot{m}_{\mathrm{lc}}(t) = \frac{\dot{Q}_{\mathrm{lc}}(t)}{c_{\mathrm{w}} \Delta T_{\mathrm{lc}}}, \tag{55}$$

with c_{w} being the medium's specific heat capacity and ΔT_{lc} being an assumed constant temperature difference for the medium returning from the fan coil units.

8.3 Operational constraints

The system is subject to different operational constraints such as state and control boundaries, minimum activations times of the machinery, and others, summarized in the inequality constraints vector h as

$$h\big(x(t), u(t), b(\cdot), c(t), s(t)\big) \leq 0 \tag{56}$$

where we write $b(\cdot)$ to indicate that the constraints for the binary controls can be coupled over time, cf. Bürger et al. (2020). The vector $s \in \mathbb{R}^{n_s}, n_s \in \mathbb{N}$ contains slack variables that can be used to relax certain conditions in h, e. g., soft constraints for the system states, to preserve the feasibility of an optimization problem if necessary. With regard to the introduction of (54) and (55) in the model, the upper and lower limits of the temperature states of the Low-Temperature Storage (LTS) are set to $T_{\mathrm{lts,max}} = 18\,^\circ\mathrm{C}$ and $T_{\mathrm{lts,min}} = 8\,^\circ\mathrm{C}$, respectively.

8.4 Optimal control problem formulation

The aim of the optimal control problem for the system is to compensate the cooling load of the building while keeping the system states (e. g. the storage temperatures) within suitable bounds and reduce the electrical energy consumption of the system. Formulated in con-

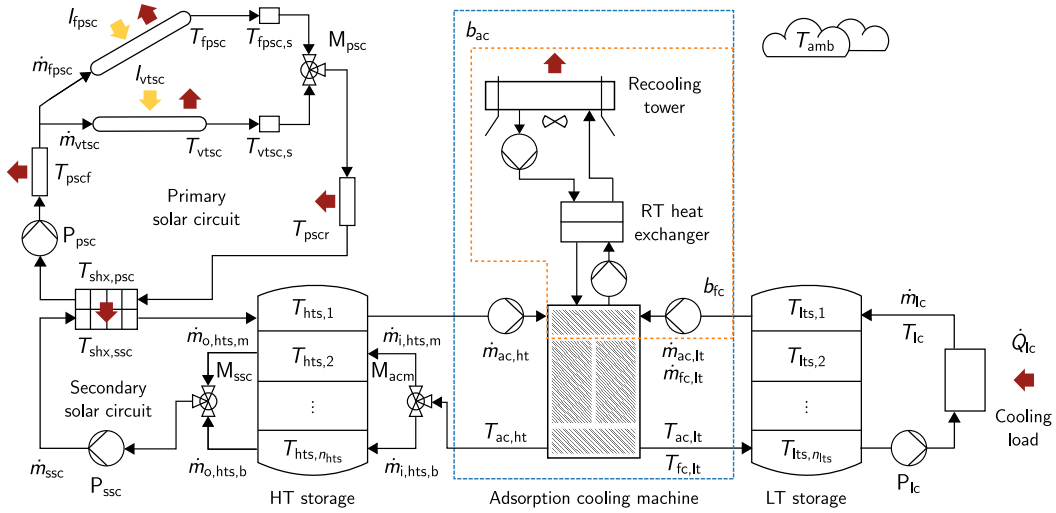

Fig. 4. Schematic of the system subject to this study.

tinuous time, the utilized MIOCP reads as

$$\min_{\substack{x(\cdot),u(\cdot),\\ b(\cdot),s(\cdot)}} \frac{1}{2}\int_{t_0}^{t_\mathrm{f}} L_1\left(x(t),u(t),b(t),s(t)\right)\,\mathrm{d}t + \int_{t_0}^{t_\mathrm{f}} L_2\left(x(t),u(t),b(t),s(t)\right)\,\mathrm{d}t \tag{57a}$$

$$\text{s.t. for } t\in[t_0,t_\mathrm{f}]: \quad \frac{\mathrm{d}x(t)}{\mathrm{d}t} = f_0\big(x(t),u(t),c(t)\big) + \sum_{i=1}^{n_b} b_i\cdot f_i\big(x(t),u(t),c(t)\big), \tag{57b}$$

$$h\big(x(t),u(t),b(\cdot),c(t),s(t)\big)\le 0, \tag{57c}$$

$$\sum_{i=1}^{n_b} b_i \le 1, \tag{57d}$$

$$x(t_0)=\hat{x}. \tag{57e}$$

The objective (57a) consists of two Lagrangian terms L_1 and L_2, where L_1 contains the nonlinear least squares terms and L_2 the linear terms of the objective function. The system model and constraints are contained in (57b) and (57c), respectively. Eq. (57d) ensures that at most one operation mode of the ACM is active at a time and (57e) is the initial state constraint.

8.5 Discretization and implementation

The MIOCP is formulated as an MINLP using *first discretize, then optimize* approaches. For steps S1 and S3 of the decomposition approach, an MINLP is derived using direct collocation (Tsang et al., 1975). For generating the components of the GN-MIQP, an MINLP is derived using direct multiple shooting (Bock and Plitt, 1984) to obtain a matrix G_L that is comparatively smaller than a corresponding matrix derived using direct collocation, with the idea to facilitate a faster solution of the MIQP.

The implementation of the MINLPs is carried out using CasADi. For generation of the initial guess via simulation and for setting up the multiple shooting discretization, CVODES (Hindmarsh et al., 2005) is used as the numerical integrator. The NLPs are solved using Ipopt with linear solver MA57 (HSL, 2019). Again, we neglect the combinatorial constraints in step S1. The MIQP is solved using Gurobi and the CIA problem is solved using the tailored branch-and-bound algorithm implemented in pycombina.

In contrast to the problem discussed in Section 7, where we were able to compute the globally optimal solution of (50) for comparison using a dedicated simulation procedure, herein, we cannot compute the globally optimal solution of (57).

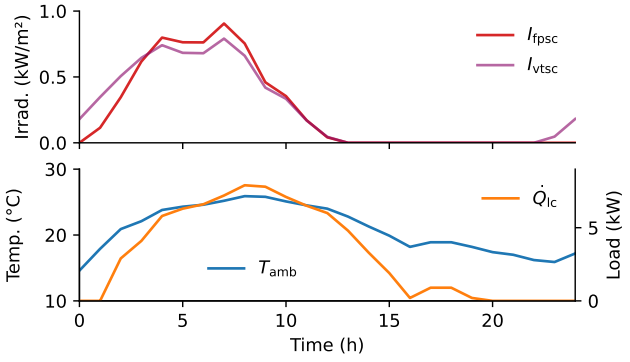

Fig. 5. Depiction of the time-varying parameters entering the MIOCP (57): solar irradiation on the collector arrays I_{fpsc} and I_{vtsc}, ambient temperature T_{amb}, and cooling load $\dot{Q}_{\text{lc}}$.

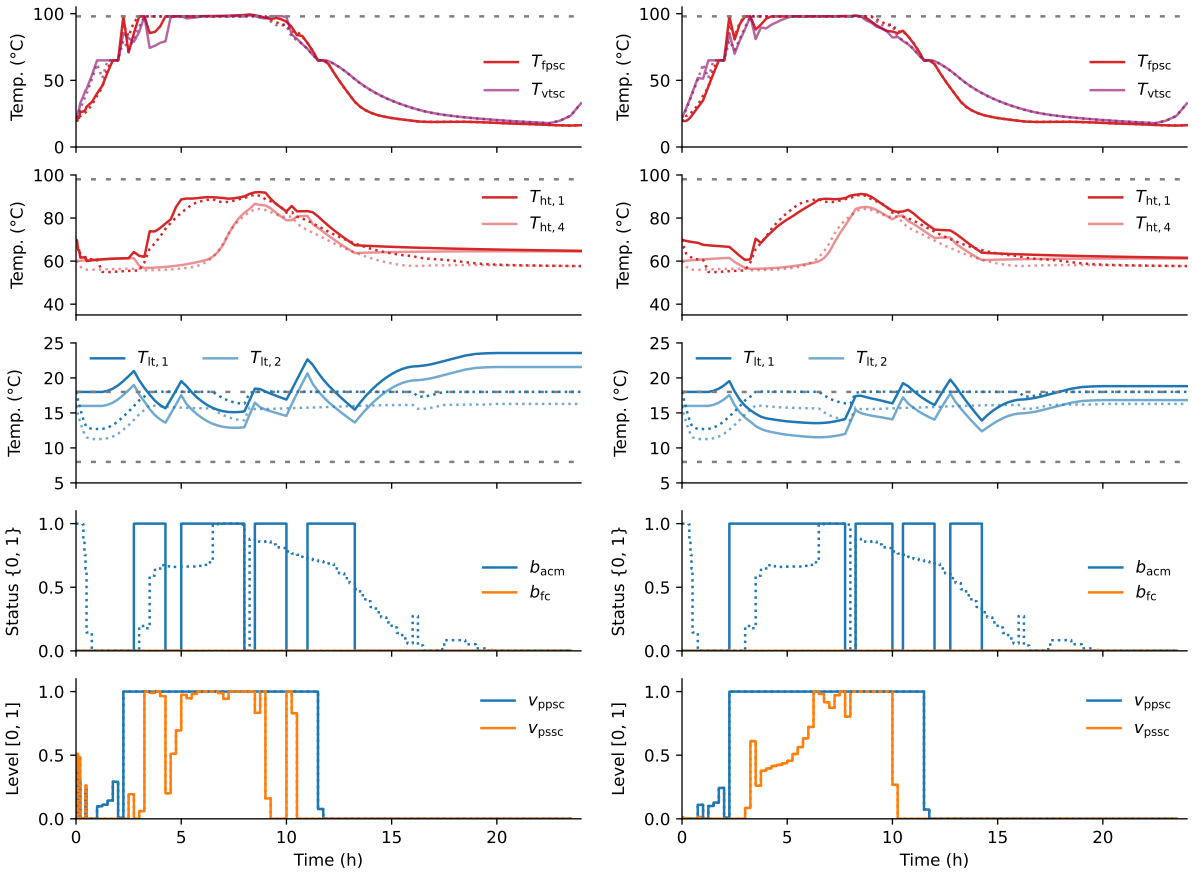

Fig. 6. Results of the CIA (left) and GN-MIQP (right) approaches after step S3 of the decomposition algorithm.

8.6 Numerical results

The CIA and GN-MIQP approaches are used to solve the MIOCP (57) for an exemplary operation scenario. The values of the time-varying parameters c are depicted in Fig. 5. Solar irradiation and ambient temperature are taken from an exemplary Test Reference Year (TRY) dataset (Deutscher Wetterdienst, 2014) and depict a summer day in August. The solar irradiation on the oriented FPSC arrays was calculated using pvlib-python (Holmgren et al., 2018), the cooling load is a generic load profile. Fig. 6 depicts the results obtained using the CIA approach in the left plots and the solution obtained using the GN-MIQP approach in the right plots. The (identical) solution after step S1 is denoted by dotted lines and the solution after step S3 is denoted by solid lines. The state boundaries considered in the form of soft constraints are indicated using dashed grey lines. It can be observed that the state boundaries of the solar collector temperatures T_{fpsc} and T_{vtsc} as well as the High-Temperature Storage (HTS) temperatures $T_{\mathrm{hts},\{1\ldots4\}}$ are widely met (only the collector temperature limits are slightly violated using the CIA approach). However, this is not the case for the LTS temperatures $T_{\mathrm{lts},1}$ and $T_{\mathrm{lts},2}$ depicted in the center plots.

The LTS temperatures are directly influenced by the binary approximation b_{ac}^{**} and b_{fc}^{**} obtained in step S2 and at the same time widely operated at their upper boundaries in the solution of step S1. Therefore, an application of the CIA approach which does not consider the development of the system states but only the relaxed binary solution within step S2, can easily violate those temperature boundaries. In contrast, it can be observed that violations of LTS temperature constraints occur less frequent and to a lesser extent while using the GN-MIQP approach. Different from the LTS temperatures, the solar collector temperatures are additionally influenced by continuous controls, e. g., the speed of the pumps $\mathrm{P}_{\mathrm{psc}}$ and $\mathrm{P}_{\mathrm{ssc}}$ and the positions of the valves $\mathrm{M}_{\mathrm{psc}}$ and $\mathrm{M}_{\mathrm{ssc}}$, so that their operation can still be adjusted to a certain extent in step S3 according to the binary approximation obtained in step S2. The HTS temperatures are, in addition to the collector operation, influenced by the ACM operation, and with this, by the quality of the obtained binary approximation. However, the HTS is not operated at its upper temperature boundary in this scenario; thus, violations are less likely to occur when an approximated binary solution is applied.

Analogously to Fig. 2, Fig. 7 displays the values

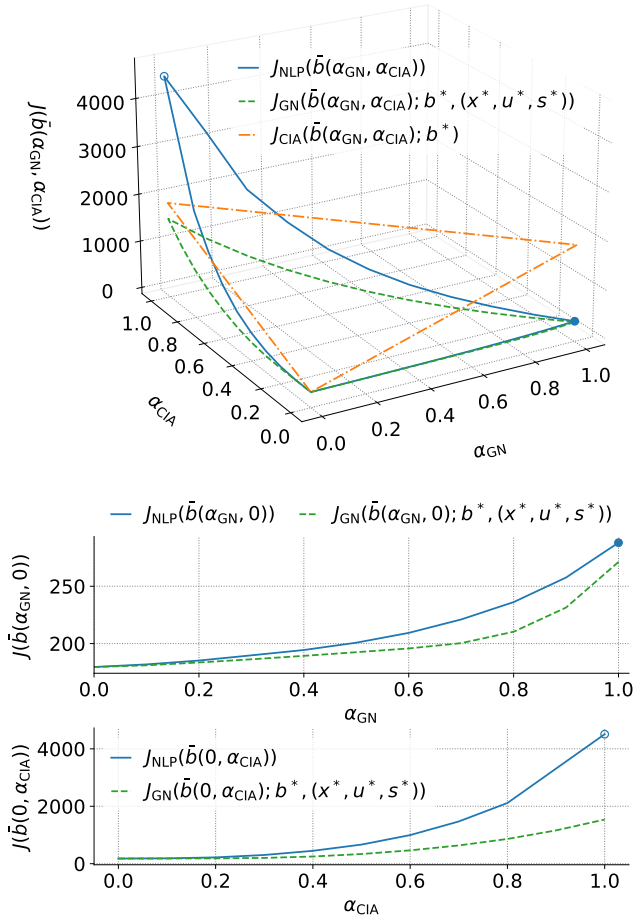

Fig. 7. Evaluation of the functions $J_{\mathrm{NLP}}\big(\bar{b}(\alpha_{\mathrm{GN}}, \alpha_{\mathrm{CIA}})\big)$, $J_{\mathrm{GN}}\big(\bar{b}(\alpha_{\mathrm{GN}}, \alpha_{\mathrm{CIA}}); b^*, (x^*, u^*, s^*)\big)$, and $J_{\mathrm{CIA}}\big(\bar{b}(\alpha_{\mathrm{GN}}, \alpha_{\mathrm{CIA}}); b^*\big)$. The value of $J_{\mathrm{NLP}}(b^{**}_{\mathrm{GN}})$ is indicated by the symbol •, the value of $J_{\mathrm{NLP}}(b^{**}_{\mathrm{CIA}})$ by the symbol ∘.

of $J_{\mathrm{NLP}}\big(\bar{b}(\alpha_{\mathrm{GN}}, \alpha_{\mathrm{CIA}})\big)$, $J_{\mathrm{CIA}}\big(\bar{b}(\alpha_{\mathrm{GN}}, \alpha_{\mathrm{CIA}}); b^*\big)$, and $J_{\mathrm{GN}}\big(\bar{b}(\alpha_{\mathrm{GN}}, \alpha_{\mathrm{CIA}}); b^*, (x^*, u^*, s^*)\big)$ for (57). It shows that J_{NLP} can be well approximated by J_{GN} locally around the linearization point $b^*, (x^*, u^*, s^*)$.

A comparison of the runtimes of each solution step for both methods is listed in Table 2. The solution times in step S1 are identical since the same problem is solved here. For step S2, the solution of the CIA problem, which was computed using a dedicated solution algorithm (Sager et al., 2011; Bürger et al., 2020), is obtained several orders of magnitude faster compared with the solution of the GN-MIQP (to full optimality). The total solution time of 330.2 s for the GN-MIQP in Table 2 includes the time for setting up the matrix G_{L} (21.8 s) and solve the problem (308.4 s).

An additional advantage of our new GN-MIQP approach is that linear constraints, which arise from many

Table 2
Runtime of the solution steps for (57).

Step	Runtime CIA (s)	Runtime GN (s)
S1	$2.46 \cdot 10^1$	$2.46 \cdot 10^1$
S2	$1.91 \cdot 10^{-2}$	$3.30 \cdot 10^2$
S3	$1.22 \cdot 10^1$	$2.26 \cdot 10^1$

practical applications, can be exactly represented in the GN-MIQP. A study in which some constraints of (56) are represented as linear reformulations for the solution of MIOCP (57), resulting in additional control performance improvements, is given in Appendix B.

9 Conclusion and future work

We presented a novel MIQP problem formulation for the second stage of an existing decomposition algorithm for fast solution of MIOCPs. First, we established a first-order consistency approximation result for the proposed algorithm. Then, we achieved improved solutions in terms of the original MINLP objective functions and overall control performance for a simple case and a more complex application but at the cost of increased solution times.

Future work could examine techniques to achieve GN-MIQP solutions faster, e. g., if warm-starting based on solutions obtained using the CIA approach can have a beneficial effect on the runtime. Moreover, it could be investigated if suboptimal solutions of the GN-MIQP can already yield an improved control performance, especially in the context of gradually improving an existing CIA result.

Conflict of interest

There is no conflict of interest.

Acknowledgements

A. Bürger, A. Altmann-Dieses, and M. Diehl acknowledge funding from INTERREG V Upper Rhine, project ACA-MODES. S. Sager and C. Zeile acknowledge funding from Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) via GRK 2297 MathCoRe (314838170) and SPPs 1962 and 2331. M. Diehl acknowledges funding by DFG via Research Unit FOR 2401 and project 424107692 and by the EU via ELO-X 953348.

A Branch-and-bound style simulation procedure for globally optimal solution of (50)

To obtain a globally optimal solution of the MINLP (50), we can consider the fact that there exists only one binary control b per discrete time interval but no other

Algorithm 1: Branch-and-bound style simulation procedure

```
Initialize an empty queue Q ;
Initialize best node n* ← ∅ and upper bound
  Φ* ← ∞ ;
Initialize i ← 0, j ← 0, n° ← ∅, j° ← 0, b° ← 0,
  x° ← x_0, Φ° ← (x_0 − x_ref)^2 ;
foreach b ∈ {0, 1} do
    j ← j° ; i ← N_mut − 1 ;
    if (b° ≠ b) ∧ (b = 1) then
        // Consider minimum uptimes
        i ← 0 ;
    while i < N_mut do
        x° ← f(x°, b) ;
        Φ° ← Φ° + (x° − x_ref)^2 ;
        i ← i + 1 ; j ← j + 1 ;
    n ← [n°, j, b, x°, Φ°] ;
    Add n to Q ;
while Q is not empty do
    n° ← next node selected from Q with the
      highest depth j; if multiple nodes with the
      same depth j exist, select the one with the
      highest objective value Φ° ;
    Unpack [n°_p, j°, b°, x°, Φ°] ← n°;
    if j° = N then
        if Φ° < Φ* then
            // Solution update
            n* ← n° ; Φ* ← Φ° ;
    else
        foreach b ∈ {0, 1} do
            j ← j° ; i ← N_mut − 1 ;
            if (b° ≠ b) ∧ (b = 1) then
                i ← 0 ;
            while i < N_mut do
                x° ← f(x°, b) ;
                Φ° ← Φ° + (x° − x_ref)^2 ;
                i ← i + 1 ; j ← j + 1 ;
            n ← [n°, j, b, x°, Φ°] ;
            Add n to Q ;
Reconstruct b*, x* starting from n* by recursively
  following the parent node references until a root
  node is reached ;
return b*, x*, Φ*
```

optimization variables, e. g., continuous controls or slack variables. As the only decision that can be made per discrete time interval is whether the binary control b is either 0 or 1, the state trajectory resulting from a given sequence of binary controls $\mathbf{b}$ can be obtained via forward simulation. Moreover, the objective function (50a) when evaluated up to a discrete time point j as in

$$\Phi_j(\mathbf{x}) = \frac{1}{2}\sum_{k=0}^{j}(x(k) - x_{\mathrm{ref}})^2 \tag{A.1}$$

with $j \in \{0, \dots, N\}$ yields a lower bound for all objective values $\Phi_l(\mathbf{x})$ with $l > j$, so that

$$\Phi_j(\mathbf{x}) \leq \Phi_l(\mathbf{x}). \tag{A.2}$$

These insights motivate the simulation procedure depicted in Algorithm 1. This procedure, which has been used in Rawlings et al. (2020), allows us to compute the globally optimal solution of MINLP (50) and obtain the optimal trajectories of binary controls $\mathbf{b}^*$ and states $\mathbf{x}^*$ and the optimal objective value Φ^* . Each node $n = [n_{\mathrm{p}}^{\circ}, j, b^{\circ}, x^{\circ}, \Phi^{\circ}]$ used within this branch-and-bound style procedure contains information on the parent node n_{p}° (while root nodes reference empty nodes as parent nodes, as denoted by $\emptyset$), corresponding discrete time point j, applied binary control b°, state x°, and objective value Φ° at the discrete time point. Parameter N_{mut} is the number of consecutive time steps that a binary control must remain active after activation, with regard to the minimum uptime constraints (51).

B Solution of MIOCP (57) using partially linearized constraint formulations

With regard to the GN-MIQP, constraints that are present in the original MIOCP formulation can be explicitly considered within the binary approximation step. Linear constraints can even be exactly represented in the GN-MIQP and not only in an approximate form as it is the case for nonlinear constraints. Therefore, we solve a variant of MIOCP (57) where the nonlinear and nonconvex constraints

$$b_{\mathrm{ac}}(t)\left(\begin{pmatrix} T_{\mathrm{ac,ht,min}} - T_{\mathrm{hts},1}(t) \\ T_{\mathrm{ac,lt,min}} - T_{\mathrm{lts},1}(t) \\ T_{\mathrm{ac,amb,min}} - T_{\mathrm{amb}}(t) \end{pmatrix} - s_{\mathrm{ac,lb}}(t)\right) \leq \epsilon_{\mathrm{ac,lb}}, \tag{B.1a}$$

$$b_{\mathrm{ac}}(t)\left(\begin{pmatrix} T_{\mathrm{hts},1}(t) - T_{\mathrm{ac,ht,max}} \\ T_{\mathrm{amb}}(t) - T_{\mathrm{ac,amb,max}} \end{pmatrix} - s_{\mathrm{ac,ub}}(t)\right) \leq \epsilon_{\mathrm{ac,ub}}, \tag{B.1b}$$

from Bürger et al. (2021), which are introduced to achieve that the ACM is only operated under suitable conditions with regard to HTS, LTS, and ambient temperature, are replaced by the following linear reformulations:

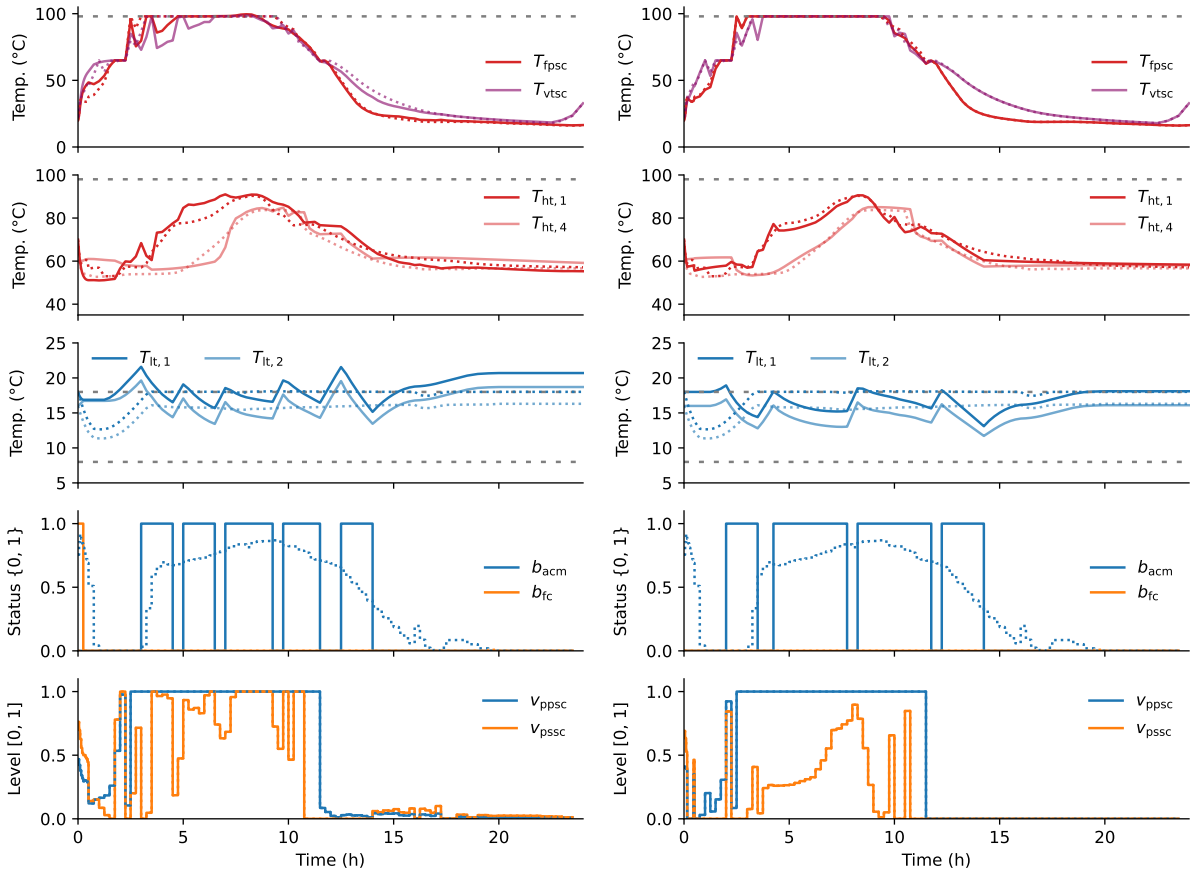

Fig. B.1. Results of the CIA (left) and GN-MIQP (right) approaches after step S3 of the decomposition algorithm using the linearized reformulation (B.2) instead of the original constraint formulation (B.1) of Bürger et al. (2021).

$$b_{\mathrm{ac}}(t)\left(\begin{pmatrix}T_{\mathrm{ac,lt,min}}\\T_{\mathrm{ac,ht,min}}\\T_{\mathrm{amb,min}}\end{pmatrix}-T_{\mathrm{min}}\right)+T_{\mathrm{min}}-s_{\mathrm{ac,lb}}(t)\leq\begin{pmatrix}T_{\mathrm{lts,1}}(t)\\T_{\mathrm{hts,1}}(t)\\T_{\mathrm{amb}}(t)\end{pmatrix}, \tag{B.2a}$$

$$b_{\mathrm{ac}}(t)\left(\begin{pmatrix}T_{\mathrm{ac,ht,max}}\\T_{\mathrm{amb,max}}\end{pmatrix}-T_{\mathrm{max}}\right)+T_{\mathrm{max}}+s_{\mathrm{ac,ub}}(t)\geq\begin{pmatrix}T_{\mathrm{hts,1}}(t)\\T_{\mathrm{amb}}(t)\end{pmatrix}. \tag{B.2b}$$

Fig. B.1 shows the solution of this variant of MINLP (57), which has been achieved using the GN-MIQP and CIA approaches. It can be observed that although the LTS storage temperature constraint violations have been reduced for both approaches, they occur less frequently and to a much lesser extent for the GN-MIQP approach.

As an additional remark, it can be observed in Fig. B.1 that the solution obtained using the CIA approach contains a short activation of the free cooling mode b_{fc} at the beginning of the control horizon, whereas b_{fc} occurs in the solution of step S1 only to a negligible extent. Similar to the effects described in Bürger (2020), this can be explained by the circumstance that the optimal solution of a CIA problem might not be unique. In the case presented here, the short activation of b_{fc} is not the dominating term of the CIA objective function, which is formulated to minimize the maximum accumulated difference between the relaxed and binary solution of all binary controls from the beginning of the control horizon up to all time points. Although the optimal solution of the CIA problem (28) for this example yields an objective value of $\theta^* = 1935.6$, the activation of b_{fc} only yields an accumulated difference of 900.0, which is not the maximum accumulated difference. In the solution obtained using the GN-MIQP approach, such activation does not occur.

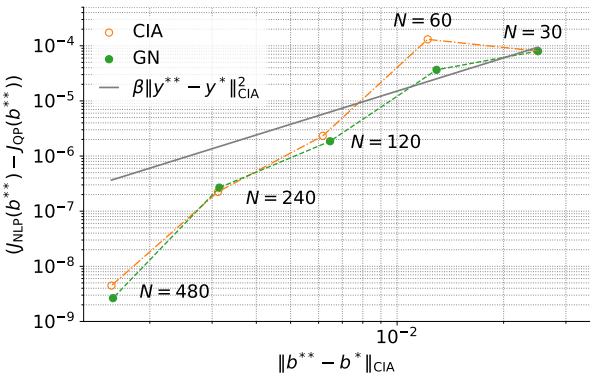

Fig. C.1. Quadratically shrinking difference between $J_{\text{NLP}}(b_N^{**})$ and $J_{\text{QP}}(b_N^{**}; b_{480}^*, x_{480}^*, B_{\text{GN},480})$ for an increase in the number of discretization intervals N, as annotated at the corresponding data points, for Problem (50) solved using the GN-MIQP and CIA approaches. The solid gray line depicts a quadratic function fitted to the data points using nonlinear least squares parameter estimation.

C Quadratic convergence for finer control discretization

To illustrate the conjectured quadratically shrinking difference between $J_{\text{NLP}}(y^{**})$ and $J_{\text{QP}}(y^{**}; y^*, z^*, B)$ in (43) with an increase in the number of control intervals mentioned in Section 6.2, we solve the MINLP (50) with combinatorial constraints $\mathbf{B}$ dropped for an increase in the number of intervals $N \in \{30, 60, 120, 240, 480\}$ using the GN-MIQP and CIA approaches. The binary solutions b_N^{**} obtained this way are used to evaluate $J_{\text{NLP}}(b_N^{**})$ and $J_{\text{QP}}(b_N^{**}; b_{480}^*, x_{480}^*, B_{\text{GN},480})$ for the finest discretization $N = 480$. The results are shown in Figure C.1 with the difference of b^{**} and b^* in the CIA-norm used as reference for the x-axis, along with a quadratic function $\beta\|b^{**} - b^*\|^2_{\text{CIA}}$ with $\beta = 0.15071$ fitted to the data points using nonlinear least squares parameter estimation. The numerical results confirm the Inequality (44) with y^{**} used for y, i.e., that the difference between J_{NLP} and J_{QP} is bounded by the quadratically normed difference of y^{**} and y^*, which can be decreased by refining the control grid size.

References

J. A. E. Andersson, J. Gillis, G. Horn, J. B. Rawlings, and M. Diehl. CasADi – A software framework for nonlinear optimization and optimal control. *Mathematical Programming Computation*, 11(1):1–36, 2019.

H. Axelsson, Y. Wardi, M. Egerstedt, and E. Verriest. Gradient descent approach to optimal mode scheduling in hybrid dynamical systems. *Journal of Optimization Theory and Applications*, 136(2):167–186, 2008.

B. Bank, J. Guddat, D. Klatte, B. Kummer, and K. Tammer. *Non-Linear Parametric Optimization*. Birkhäuser Verlag, 1983.

A. Bemporad, A. Giua, and C. Seatzu. A master-slave algorithm for the optimal control of continuous-time switched affine systems. In *Proceedings of the 41st IEEE Conference on Decision and Control, 2002.*, volume 2, pages 1976–1981. IEEE, 2002.

F. Bestehorn, C. Hansknecht, C. Kirches, and P. Manns. Mixed-integer optimal control problems with switching costs: a shortest path approach. *Mathematical Programming*, pages 1–32, 2020.

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

J. F. Bonnans and A. Shapiro. *Perturbation analysis of optimization problems*. Springer Science & Business Media, 2013.

F. Borrelli, A. Bemporad, and M. Morari. *Predictive Control for Linear and Hybrid Systems*. Cambridge University Press, 1st edition, 2017.

A. Bürger. *Nonlinear mixed-integer model predictive control of renewable energy systems*. PhD thesis, University of Freiburg, 2020.

A. Bürger, C. Zeile, A. Altmann-Dieses, S. Sager, and M. Diehl. Design, implementation and simulation of an MPC algorithm for switched nonlinear systems under combinatorial constraints. *Journal of Process Control*, 81:15–30, 2019.

A. Bürger, C. Zeile, M. Hahn, A. Altmann-Dieses, S. Sager, and M. Diehl. pycombina: An open-source tool for solving combinatorial approximation problems arising in mixed-integer optimal control. In *IFAC-PapersOnLine*, volume 53, pages 6502–6508, 2020. 21st IFAC World Congress.

A. Bürger, D. Bull, P. Sawant, M. Bohlayer, A. Klotz, D. Beschütz, A. Altmann-Dieses, M. Braun, and M. Diehl. Experimental operation of a solar-driven climate system with thermal energy storages using mixed-integer nonlinear model predictive control. *Optimal Control Applications and Methods*, 42(5):1293–1319, 2021.

A. De Marchi. On the mixed-integer linear-quadratic optimal control with switching cost. *IEEE Control Systems Letters*, 3(4):990–995, 2019.

Deutscher Wetterdienst. Testreferenzjahre von Deutschland für mittlere, extreme und zukünftige Witterungsverhältnisse. Deutscher Wetterdienst, Offenbach, 2014.

O. Exler, T. Lehmann, and K. Schittkowski. A comparative study of SQP-type algorithms for nonlinear and nonconvex mixed-integer optimization. *Mathematical Programming Computation*, 4(4):383–412, 2012.

M. Fischetti, F. Glover, and A. Lodi. The feasibility pump. *Mathematical Programming*, 104(1):91–104, 2005.

R. Fletcher and S. Leyffer. Solving mixed integer nonlinear programs by outer approximation. *Mathematical programming*, 66(1):327–349, 1994.

M. Gerdts. A variable time transformation method for mixed-integer optimal control problems. *Optimal Control Applications and Methods*, 27(3):169–182,

2006.

S. Göttlich, F. M. Hante, A. Potschka, and L. Schewe. Penalty alternating direction methods for mixed-integer optimal control with combinatorial constraints. *Mathematical Programming*, 188:599–619, 2021.

Gurobi Optimization, LLC. Gurobi optimizer reference manual. `http://www.gurobi.com`, 2021. Last accessed March 12, 2021.

A. C. Hindmarsh, P. N. Brown, K. E. Grant, S. L. Lee, R. Serban, D. E. Shumaker, and C. S. Woodward. SUNDIALS: Suite of nonlinear and differential/algebraic equation solvers. *ACM Transactions on Mathematical Software*, 31(3):363–396, 2005.

W. F. Holmgren, C. W. Hansen, and M. A. Mikofski. pvlib python: a python package for modeling solar energy systems. *Journal of Open Source Software*, 3(29):884, 2018. Software version doi:10.5281/zenodo.5366883.

HSL. A collection of Fortran codes for large scale scientific computation. `http://www.hsl.rl.ac.uk/`, 2019. Last accessed May 17, 2019.

M. Jung, G. Reinelt, and S. Sager. The Lagrangian relaxation for the combinatorial integral approximation problem. *Optimization Methods and Software*, 30(1): 54–80, 2015.

C. Kirches. *Fast Numerical Methods for Mixed-Integer Nonlinear Model-Predictive Control*. Vieweg+Teubner Verlag, 2011.

S. Leyffer and P. Manns. Sequential linear integer programming for integer optimal control with total variation regularization. *arXiv preprint arXiv:2106.13453*, 2021.

P. Lilienthal, M. Tetschke, E. Schalk, T. Fischer, and S. Sager. Optimized and personalized phlebotomy schedules for patients suffering from polycythemia vera. *Frontiers in physiology*, 11:328, 2020.

P. Manns. Relaxed multibang regularization for the combinatorial integral approximation. *SIAM Journal on Control and Optimization*, 59(4):2645–2668, 2021.

J. B. Rawlings, D. Q. Mayne, and M. M. Diehl. *Model Predictive Control: Theory, Computation, and Design*. Nob Hill, 2nd edition, 2020. 3rd printing.

N. Robuschi, C. Zeile, S. Sager, and F. Braghin. Multiphase mixed-integer nonlinear optimal control of hybrid electric vehicles. *Automatica*, 123:109325, 2021.

S. Sager. Reformulations and algorithms for the optimization of switching decisions in nonlinear optimal control. *Journal of Process Control*, 19(8):1238–1247, 2009.

S. Sager and C. Zeile. On mixed-integer optimal control with constrained total variation of the integer control. *Computational Optimization and Applications*, 78(2): 575–623, 2021.

S. Sager, M. Jung, and C. Kirches. Combinatorial integral approximation. *Mathematical Methods of Operations Research*, 73(3):363, 2011.

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

T. H. Tsang, D. M. Himmelblau, and T. F. Edgar. Optimal control via collocation and non-linear programming. *International Journal of Control*, 21(5):763–768, 1975.

A. Wächter and L. T. Biegler. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical Programming*, 106(1):25–57, 2006.

C. Zeile, N. Robuschi, and S. Sager. Mixed-integer optimal control under minimum dwell time constraints. *Mathematical Programming*, 188:653–694, 2021.