On Perspective Functions and Vanishing Constraints in Mixed-Integer Nonlinear Optimal Control

Michael N. Jung, Christian Kirches, and Sebastian Sager

Abstract Logical implications appear in a number of important mixed-integer nonlinear optimal control problems (MIOCPs). Mathematical optimization offers a variety of different formulations that are equivalent for boolean variables, but result in different relaxations. In this article we give an overview over a variety of different modeling approaches, including outer versus inner convexification, generalized disjunctive programming, and vanishing constraints. In addition to the tightness of the respective relaxations, we also address the issue of constraint qualification and the behavior of computational methods for some formulations. As a benchmark, we formulate a truck cruise control problem with logical implications resulting from gear-choice specific constraints. We provide this benchmark problem in AMPL format along with different realistic scenarios. Computational results for this benchmark are used to investigate feasibility gaps, integer feasibility gaps, quality of local solutions, and well-behavedness of the presented reformulations of the benchmark problem. Vanishing constraints give the most satisfactory results.

1 Introduction

Sebastian Sager (in 2006, [56]) and Christian Kirches (in 2010, [39]) received their Ph.D.s from Heidelberg University under the supervision of Gerhard Reinelt, thus

M.N. Jung · C. Kirches
Interdisziplinäres Zentrum für Wissenschaftliches Rechnen, Ruprecht-Karls-Universität Heidelberg, Im Neuenheimer Feld 368, 69120 Heidelberg, Germany

M.N. Jung
e-mail: michael.jung@iwr.uni-heidelberg.de

C. Kirches
e-mail: christian.kirches@iwr.uni-heidelberg.de

S. Sager (✉)
Institut für Mathematische Optimierung, Otto-von-Guericke-Universität Magdeburg, Universitätsplatz 2, 02-224, 39106 Magdeburg, Germany
e-mail: sager@ovgu.de

becoming grand-sons of Martin Grötschel. Michael Jung is a Ph.D. student at Heidelberg University supervised by Reinelt and Sager, and expects to become both a grandson and a grand-grandson of Grötschel in 2013. Another interesting link is the scientific and personal socialization of all three authors in the realm of Hans Georg Bock, who, like Grötschel, is a protagonist of the University of Augsburg's golden age of applied mathematics in the late eighties.

Mixed-integer optimal control problems (MIOCPs) have been gaining significantly increased interest over the last years. This is due to the fact that the underlying processes have a high potential for optimization, while at the same time they are hard to assess manually due to their combinatorial, nonlinear, and dynamic nature. Typical examples are the choice of gears in automotive control [26, 40], water or gas networks [14, 49], traffic flow [23, 32], supply chain networks [29], distributed autonomous systems [1], and processes in chemical engineering that involve valves [38, 65]. The truck benchmark problem we present in this article is motivated by work of [35, 39, 68] on heavy duty trucks. See [59] for an open benchmark library for MIOCPs. In this article, we are especially interested in switching decisions that imply constraints. In the interest of readability we focus on the specific problem class of MIOCPs in ordinary differential equations (ODEs) of the following form.

Definition 1 (MIOCP) In this contribution a mixed-integer optimal control problem (MIOCP) refers to the switched system on a fixed time horizon $[0, t_f]$ given by

$$
\begin{aligned}
\min_{x(\cdot), u(\cdot), Y(\cdot)} \quad & e\big(x(t_f)\big) \\
\text{s.t.} \quad & \bigvee_{i \in \{1, \ldots, n_\omega\}} \begin{bmatrix} Y_i(t) \\ \dot{x}(t) = f(x(t), u(t), v^i) \\ 0 \leq c(x(t), u(t), v^i) \end{bmatrix} \quad \forall t \in [0, t_f], \\
& x(0) = x_0, \\
& 0 \leq d\big(x(t), u(t)\big) \quad \forall t \in [0, t_f].
\end{aligned}
\tag{1}
$$

The logical operator $\bigvee_{i \in \{1, \ldots, n_\omega\}}$ implies that at all times $t \in [0, t_f]$ exactly one of the n_ω possible modes is chosen and is represented here by time-dependent logical literals $Y_i(\cdot)$, $1 \leq i \leq n_\omega$. The optimal control $u : [0, t_f] \to \mathbb{R}^{n_u}$ is assumed to be measurable and of bounded variation, the differential states $x : [0, t_f] \to \mathbb{R}^{n_x}$ to be uniquely determined once a switching regime $Y(\cdot)$ and the controls $u(\cdot)$ are fixed. The vectors $v^i \in \mathbb{R}^{n_v}$ comprise constant values that are specific for the given mode. We write $\Omega := \{v^1, v^2, \ldots, v^{n_\omega}\}$. The objective function $e : \mathbb{R}^{n_x} \to \mathbb{R}$ of Mayer type and the constraint functions $c : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_v} \to \mathbb{R}^{n_c}$ and $d : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}^d$ are assumed to be sufficiently often continuously differentiable.

For example, in Sect. 2 the vectors v^i will denote degrees of efficiency and engine speed bounds for a particular choice i of the gear. Note that problem class (1) can

be generalized in many directions to include different objective functionals, multi-point constraints, algebraic variables, more general hybrid systems, and the likes, compare [57] for further details and references.

Although the first MIOCPs, namely the optimization of subway trains that are equipped with discrete acceleration stages, were already solved in the early eighties for the city of New York [12], the so-called *indirect methods* used there do not seem appropriate for generic optimal control problems with underlying nonlinear dynamic systems and mixed path-control constraints. The discussion of advantages and disadvantages of indirect approaches is ongoing since the 1980s and beyond the scope of this paper. See [57] for a discussion of the applicability of dynamic programming and global maximum principles to MIOCP and further references.

Direct and all-at-once methods, such as Bock's direct multiple shooting [13, 44] and direct collocation [5, 9, 10] have emerged as the methods of choice for the computational solution of many purely continuous optimal control problems. They are also at the heart of our approach to MIOCP. These approaches have in common that they result in highly structured nonlinear programs (NLPs). As the focus of this article is on properties of relaxed MIOCPs, we do not go into details on how these transformations are carried out, but rather refer the interested reader to [11, 44]. Nevertheless, we need some NLP definitions to illustrate and transfer important concepts.

Definition 2 (NLP) A nonlinear programming problem (NLP) is given by

$$\min_{y \in \mathbb{R}^{n_y}} E(y) \quad \text{s.t.} \quad C(y) = 0, \qquad D(y) \geq 0 \tag{2}$$

with twice continuously differentiable functions $E : \mathbb{R}^{n_y} \to \mathbb{R}$, $C : \mathbb{R}^{n_y} \to \mathbb{R}^{n_C}$, and $D : \mathbb{R}^{n_y} \to \mathbb{R}^{n_D}$.

Note that we distinguish between control problems and nonlinear programs by using lower and upper case letters respectively, and make use of the unusual E for the objective function to avoid confusion with the right hand side function $f(\cdot)$.

Definition 3 (LICQ) A nonlinear programming problem is said to satisfy the *linear independence constraint qualification* (LICQ) in a feasible point $\bar{y} \in \mathbb{R}^{n_y}$ if for all $d \neq 0 \in \mathbb{R}^{n_y}$ it holds that

$$\begin{aligned}
\nabla C_i(\bar{y})^T d &\neq 0, \quad 1 \leq i \leq n_C, \\
\nabla D_i(\bar{y})^T d &\neq 0, \quad i \in \left\{ 1 \leq j \leq n_D : D_j(\bar{y}) = 0 \right\}.
\end{aligned} \tag{3}$$

In Sect. 4 we will see that certain constraint formulations do not enjoy the LICQ property. Thus, we discuss mathematical programs with complementarity structure.

Definition 4 (MPCC, MPVC) A nonlinear programming problem

$$\min_{y \in \mathbb{R}^{n_y}} E(y)$$

$$\text{s.t.} \quad G(y) \geq 0, \qquad H(y) \geq 0, \qquad G_i(y) \cdot H_i(y) = 0, \quad 1 \leq i \leq n_{GH}, \tag{4}$$

with twice continuously differentiable functions E, G, and H is called a *Mathematical Program with Complementarity Constraints*, e.g., [7]. An NLP

$$\min_{y \in \mathbb{R}^{n_y}} E(y) \quad \text{s.t.} \quad H(y) \geq 0, \qquad G_i(y) \cdot H_i(y) \leq 0, \quad 1 \leq i \leq n_{GH}, \tag{5}$$

with twice continuously differentiable functions F, G, and H is called a *Mathematical Program with Vanishing Constraints*, see e.g., [2].

MPCCs and MPVCs are known to possess critical points that violate LICQ. Their computational solution with standard NLP software is prone to numerical difficulties and will often terminate in suboptimal points that possess trivial descent directions and are not local minimizers of (4) or (5).

Remark 1 (MINLP view of MPVCs) One way to see the difficulty is a look at the mixed-integer nonlinear program given by

$$\min_{y \in \mathbb{R}^{n_y}, z \in \{0,1\}} E(y) \quad \text{s.t.} \quad G(y) \leq zb \tag{6}$$

with G such that $\{y \in \mathbb{R}^{n_y} | G(y) \leq 0\} = \{0\}$. It is closely related to problem (5) as $H_i(y)$ toggles the constraint on $H_i(y) \cdot G_i(y)$ in (5) like z does for the constraint on $G(y) - b$ in (6):

$$z = 1, G(y) - b \leq 0 \quad \Longleftrightarrow \quad H_i(y) > 0, H_i(y) \cdot G_i(y) \leq 0,$$

$$z = 0, G(y) = 0 \quad \Longleftrightarrow \quad H_i(y) = 0, H_i(y) \cdot G_i(y) = 0.$$

In this article, we investigate different approaches to reformulate problem (1). Here, reformulations are optimal control problems that have the same feasible set and hence the same optimal solution if integrality is required, but have a possibly different and larger feasible set if the discrete decisions are relaxed. In Sect. 3 we discuss formulations of the dynamic equations with further references to the literature. With respect to formulations of the inequalities in the logical disjunctions in (1), not only tightness of the relaxation is crucial. Also, constraint qualifications may or may not hold and homotopies might be needed to obtain convergence to local solutions. We refer to this as "well-behavedness" of a relaxation for our benchmark problem.

There are several reformulations of logical relationships in the literature. *Generalized Disjunctive Programming* results directly from a logical modeling paradigm. It generalizes the disjunctive programming approach of [4]. Logical variables are

usually either incorporated by means of big-M constraints or via a convex hull formulation, see [31, 50, 51, 67]. From a different point of view, disjunctive programming formulations can be interpreted as the result of *reformulation-linearization technique* (RLT) steps [64]. For both, the convex hull relaxation uses *perspective functions*. Based on this, the use of perspective cuts to strengthen convex MINLP relaxations has been proposed in various articles, for example [15, 22, 33]. We refer the reader to [8] for a recent survey on MINLP techniques.

Complementarity (4) *and Vanishing Constraints* (5) are another way to look at logical implications. The general concept of nonlinear optimization over nonconvex structures is discussed in [62, 63]. For the comparatively young problem class of MPVCs we refer to [2, 36]. Due to the lack of constraint qualification, various approaches for the computational solution of MPCCs and MPVCs have been devised and include regularizations [36, 54, 66], smoothing [16, 36], and combinations thereof; see [24] for an overview. Nonlinear programming for MPCCs is addressed in [3, 21, 45, 47]. Active set methods tailored to the nonconvex structure are discussed in [17, 39, 43]. Formulations of MPCCs and MPVCs in optimal control can be found in [7, 39, 43, 52, 53].

The remainder of this article is organized as follows. In Sect. 2 we describe a heavy-duty truck cruise control problem based on a dynamic vehicle model that includes gear shift decisions. This is a prototypical MIOCP with constraints based on logical implications. In Sect. 3 we explain the *partial outer convexification* approach for MIOCPs. We discuss extensions of this approach and different reformulations arising for the heavy-duty truck model, and numerically assess the advantage of partial outer convexification over an inner convexification. In Sect. 4 we investigate reformulations for logical implication constraints. We discuss merits and drawbacks as well as the consequences arising for underlying NLP solvers. Numerical results for the heavy-duty truck control problem are presented in Sect. 5.

2 A Cruise Control Problem for a Heavy-Duty Truck

In this section we describe a cruise control problem for a heavy-duty truck, and model it as a mixed-integer tracking control problem on a prediction horizon.

Typical heavy-duty trucks feature from $n_\mu = 8$ to $n_\mu = 24$ gears with different transmission ratios and efficiencies. The decision on energy-optimal gear shift strategies under real-time constraints usually requires extensive training on the driver's side, is a subject of intense scientific research, and bears considerable potential for savings especially in view of future hybrid engines, e.g., [35, 39, 68].

Controls and Dynamic System   We describe a basic ODE truck model as introduced in [68] that is used for all computations. The ODE system of the truck model comprises three input controls: the indicated engine torque M_{ind}, the engine brakes torque M_{EB}, and the integer gear choice μ.

The vehicle model involves two differential states, velocity v and accumulated fuel consumption Q. Traveled distance s is chosen as the independent variable and

we consider the interval $[0, s_f]$. The first state $v(s)$ denotes velocity (in m/s) and is determined from the sum of directed torques,

$$mv(s)\dot{v}(s) = \big(M_{acc}(s) - M_{brk}(s)\big)i_A/r_{stat} - M_{air}(s) - M_{road}(s), \qquad (7)$$

depending on the effective accelerating and braking torques M_{acc}, M_{brk} as well as on turbulent and static friction torques M_{air}, M_{road} (all in Nm). Further, i_A is the fixed axle transmission ratio and r_{stat} (in m) is the static tire radius. The second state, fuel consumption $Q(s)$ (in l/s), is computed from integration over a consumption map

$$v(s)\dot{Q}(s) = Q_{fuel}\big(n_{eng}\big(s, \mu(s)\big), M_{ind}(s)\big), \qquad (8)$$

depending on engine speed n_{eng} and torque M_{ind}. Several terms are computed from algebraic formulas. The transmitted engine speed n_{eng} (in 1/min) depends on the selected gear μ and is obtained from velocity,

$$n_{eng}\big(s, \mu(s)\big) := v(s)i_A i_T\big(\mu(s)\big)60/(2\pi r_{stat}).$$

The accelerating torque M_{acc} is computed from the ratio $i_T(\mu)$ and efficiency $\eta_T(\mu)$ associated with the selected gear μ. Braking torques M_{brk} are due to controlled engine brake torque M_{EB} and internal engine friction torque M_{fric},

$$M_{acc}(s) := i_T\big(\mu(s)\big)\eta_T\big(\mu(s)\big)M_{ind}(s),$$
$$M_{brk}(s) := M_{EB}(s) + i_T\big(\mu(s)\big)M_{fric}\big(n_{eng}\big(s, \mu(s)\big)\big).$$

Additional braking torques due to turbulent friction M_{air}, and due to static road conditions M_{road}, are taken into account,

$$M_{air}(s) := \frac{1}{2}c_w A\rho_{air}v^2(s), \qquad M_{road}(s) := mg\big(\sin\gamma(s) + f_r\cos\gamma(s)\big).$$

Here c_w is the shape coefficient, A denotes the flow surface (in m^2), and ρ_{air} the air density (in kg/m^3). Further, m is the vehicle's mass (in kg), gravity is denoted by g (in m/s^2), $\gamma(s)$ denotes the road's slope, and f_r is a rolling friction coefficient.

Objective   The cost criterion to be minimized on the prediction horizon $[0, s_f]$ is composed of a weighted sum of three different objectives. To simplify notation we transform the Lagrange terms into a Mayer term by introducing artificial differential states. First, the deviation of the truck's velocity from the desired one is penalized,

$$v(s)\dot{\Phi}_{dev}(s) := \big(v(s) - v_{des}(s)\big)^2 \qquad (9)$$

Second, the fuel consumption is found from a fuel consumption rate map Q, see [39], and depends on the integer gear choice $\mu(s)$,

$$v(s)\dot{\Phi}_{fuel} := Q\big(n_{eng}\big(s, \mu(s)\big), M_{ind}(s)\big). \qquad (10)$$

Third, rapid changes of the engine and brake torques degrade driving comfort,

$$v(s)\dot{\Phi}_{\text{comf}} := \dot{M}_{\text{ind}}^2(s) + \dot{M}_{\text{brk}}^2(s), \tag{11}$$

where we approximate the torque derivatives by one-sided finite differences.

Constraints   On the prediction horizon mechanical constraints and velocity limits need to be respected. Beside the bounds on the truck input controls and on the system's states for $s \in [0, s_{\text{f}}]$,

$$0 \le M_{\text{ind}}(s), \qquad 0 \le M_{\text{brk}}(s) \le M_{\text{brk,max}}, \tag{12}$$

the truck's velocity $v(s)$ is subject to velocity limits imposed by law,

$$v(s) \le v_{\text{law}}(s), \quad s \in [0, s_{\text{f}}]. \tag{13}$$

The indicated torque must respect state-dependent upper limits as specified by the engine characteristics for $s \in [0, s_{\text{f}}]$:

$$M_{\text{ind}}(s) \le M_{\text{ind,max}}\big(n_{\text{eng}}(s, \mu(s))\big). \tag{14}$$

In addition, the transmitted engine speed $n_{\text{eng}}(s, \mu(s))$ must stay within prescribed limits according to the engine's specification,

$$n_{\text{eng,min}} \le n_{\text{eng}}\big(s, \mu(s)\big) \le n_{\text{eng,max}}, \quad s \in [0, s_{\text{f}}]. \tag{15}$$

Problem Formulation   The MIOC problem formulation for the heavy-duty truck control problem on the prediction horizon $s \in [0, s_{\text{f}}]$ reads

$$\min_{x(\cdot),u(\cdot),\mu(\cdot)} \lambda_{\text{dev}}\Phi_{\text{dev}}(s_{\text{f}}) + \lambda_{\text{fuel}}\Phi_{\text{fuel}}(s_{\text{f}}) + \lambda_{\text{comf}}\Phi_{\text{comf}}(s_{\text{f}})$$

$$\text{s.t.} \quad \bigvee_{i \in \{1,\dots,n_\mu\}} \begin{bmatrix} \mu(s) = i \\ \text{ODE system (7), (8)} \\ \text{Constraints (14), (15)} \end{bmatrix} \quad \forall s \in [0, s_{\text{f}}], \tag{16}$$

$$x(0) = x_0,$$

Constraints (12), (13) $\forall s \in [0, s_{\text{f}}]$,

with state vector $x(s) = (v, Q, \Phi_{\text{dev}}, \Phi_{\text{fuel}}, \Phi_{\text{comf}})(s)$, continuous control vector $u(s) = (M_{\text{ind}}, M_{\text{brk}})(s)$, integer controls $\mu(s)$, and initial state information x_0. Note that problem (16) is a MIOCP of form (1), with the gear choice $\mu(s)$ that causes switches in the right hand side as well as in the constraints and the associated vectors $v^i = (i_{\text{T}}(i), \eta_{\text{T}}(i))$ that contain gear-specific values for transmission ratio and efficiency.

Logical Implications   The modeling and optimization with logical variables has gained increasing interest, [31, 50]. The main reason is that it provides a modeling paradigm that is both generic and intuitive, and allows for tailored relaxations and algorithms.

Logical literals $Y \in \{\text{false, true}\}$ can of course be identified with binary variables $z_i \in \{0, 1\}$, or can be used implicitly, as we did in formulation (16),

$$Y_i(t) = \text{true} \quad \Leftrightarrow \quad \mu(t) = i.$$

It is well known that logical relations and Boolean expressions can be represented as linear (in)equalities. For example, the implication is represented by

$$Y_1 \Rightarrow Y_2 \quad \Longleftrightarrow \quad \neg Y_1 \vee Y_2 \quad \Longleftrightarrow \quad 1 - z_1 + z_2 \geq 1.$$

For time-dependent logical literals $Y(\cdot)$, systems of differential equations, and implied constraints the representation is, however, neither straightforward nor unique.

In Sect. 3 we discuss two different formulations for the ODE system and review recent results. Constraints (14) and (15) are of particular interest in the context of this article. They represent the logical implication

When gear i is active, gear-specific constraints must hold.

In Sect. 4, we address different mathematical reformulations of this implication and investigate the tightness of the induced relaxations.

3 Inner and Outer Convexification in MIOC

In this section we investigate two different approaches to reformulate problem (1) when $n_c = 0$, i.e., when no mode-specific constraints are present and only the right hand side function of the dynamic system depends on the logical mode choice at time t. We denote the two approaches by *inner convexification* and *(partial) outer convexification*, depending on how the variables that represent the logical choice enter the right hand side.

The goal in both cases is identical: a formulation that allows the relaxation of discrete decisions. The relaxed MIOCPs are purely continuous control problems that can be solved using, for example, Bock's direct multiple shooting method or direct collocation.

Using the truck control problem from Sect. 2, we compare both approaches numerically and explain qualitatively why the outer convexification formulation performs so much better than the inner convexification. We close this section by referring to related work and extensions.

Inner Convexification   For some control problems of type (1) it is possible to reformulate the time-dependent disjunctions by means of a function $g : [1, n_\omega] \rightarrow \mathbb{R}^{n_v}$ that can be inserted into the right hand side function $f(\cdot)$ and has the property

$g(i) = v^i$ for $i \in \{1, \ldots, n_\omega\}$. Possibilities are a piecewise linear representation of the form

$$g(i + \xi_{i+1}) = \xi_i v^i + \xi_{i+1} v^{i+1} \tag{17}$$

with special ordered set type 2 (SOS-2) variables

$$\xi_i \in [0, 1], \qquad \sum_i \xi_i = 1, \qquad \xi_i \neq 0 \quad \Rightarrow \quad \xi_j = 0 \quad \forall j \neq i, i+1,$$

a convex combination

$$g\left(\sum_{i=1}^{n_\omega} \xi_i i\right) = \sum_{i=1}^{n_\omega} \xi_i v^i \tag{18}$$

with special ordered set type 1 (SOS-1) variables $\xi_i \in [0, 1]$, $\sum_i \xi_i = 1$, or fitted smooth convex functions $g(\cdot)$ as suggested in [25]. Either way, this approach allows a reformulation of (1) into

Definition 5 (MIOCP after IC Reformulation)

$$\min_{x(\cdot), u(\cdot), \psi(\cdot)} e\big(x(t_f)\big)$$

$$\text{s.t.} \quad \dot{x}(t) = f\big(x(t), u(t), g(\psi)\big) \quad \forall t \in [0, t_f],$$

$$0 \leq d\big(x(t), u(t)\big) \quad \forall t \in [0, t_f], \tag{19}$$

$$x(0) = x_0,$$

$$\psi(t) \in \{1, \ldots, n_\omega\} \quad \forall t \in [0, t_f].$$

By construction, problem (19) can be relaxed toward $\psi(t) \in [1, n_\omega]$.

Outer Convexification   The outer convexification approach (or *partial* outer convexification, because the convexification applies to the integer controls only, and not to the rest of the control problem) has been investigated in the context of optimal control in [56, 60, 61]. It consists of an evaluation of all possible right hand sides, their multiplication with convex multipliers, and the summation of the products. We introduce control functions $\omega : [0, t_f] \to \{0, 1\}^{n_\omega}$ as convex multipliers and obtain:

Definition 6 (MIOCP after OC Reformulation)

$$\min_{x(\cdot), u(\cdot), \omega(\cdot)} e\big(x(t_f)\big)$$

$$\text{s.t.} \quad \dot{x}(t) = \sum_{i=1}^{n_\omega} \omega_i(t) f\big(x(t), u(t), v^i\big) \quad \forall t \in [0, t_f],$$

$$0 \leq d\big(x(t), u(t)\big) \quad \forall t \in [0, t_f], \tag{20}$$

$$x(0) = x_0$$

$$1 = \sum_{i=1}^{n_\omega} \omega_i(t), \quad \omega(t) \in \{0, 1\}^{n_\omega} \ \forall t \in [0, t_f].$$

Problem (20) can be relaxed toward $\omega(t) \in [0, 1]^{n_\omega}$. We use the notation $\alpha(t) \in [0, 1]^{n_\omega}$ to highlight the difference between the original problem and its continuous relaxation. In the following we need the notation of

Definition 7 (Fractionality of solutions) The *fractionality* of binary control functions $\omega(\cdot)$ on a time horizon $[0, t_f]$ is given by

$$\frac{1}{n_\omega} \sum_{i=1}^{n_\omega} \int_0^{t_f} \left(0.5 - \left\|\omega_i(\tau) - 0.5\right\|\right) d\tau.$$

Discussion of Inner Versus Outer Convexification   Relaxations of reformulation (19) may be faster to solve, as we only have one control function $\psi : [0, t_f] \to [1, n_\omega]$ that enters the control problem instead of n_ω functions $\omega_i : [0, t_f] \to [0, 1]$ in (20). Hence, there are less derivatives to be computed and the subproblems arising in iterations of an interior point or SQP approach are cheaper to solve. Moreover in (20), the aggregated right hand side function $\sum_i f(\cdot, v^i)$ may become more expensive to evaluate. As n_ω reflects the number of possible combinations and switches, this number may get large. Note that the linear equality constraint may be used for elimination of one function at the cost of losing some sparsity.

However, depending on the separability properties of $f(\cdot)$, integer controls often decouple, leading to a reduced number n_ω of admissible choices, e.g., [30].

Example 1 (Outer Convexification and Separability) Assume we have

$$\dot{x}(t) = f_1\big(\cdot, v_1(t)\big) + f_2\big(\cdot, v_2(t)\big), \quad v_1(t) \in \Omega_1, v_2(t) \in \Omega_2.$$

Then an equivalent reformulation leading to $n_\omega = n_{\omega_1} + n_{\omega_2}$ controls instead of $n_\omega = n_{\omega_1} n_{\omega_2}$ is given for $t \in [t_0, t_f]$ by

$$\dot{x}(t) = \sum_{i=1}^{n_{\omega_1}} f_1\big(\cdot, v_1^i\big)\omega_{1,i}(t) + \sum_{i=1}^{n_{\omega_2}} f_2\big(\cdot, v_2^i\big)\omega_{2,i}(t),$$

$$\sum_{i=1}^{n_{\omega_1}} \omega_{1,i}(t) = 1, \quad \omega_1 \in \{0, 1\}^{n_{\omega_1}}, \qquad \sum_{i=1}^{n_{\omega_2}} \omega_{2,i}(t) = 1, \quad \omega_2 \in \{0, 1\}^{n_{\omega_2}}.$$

In most practical applications the binary control functions enter linearly (such as valve controls that indicate whether a certain flow term is present or not), or n_ω increases linearly with the number of choices (e.g., the gears), or integer controls decouple. Hence, one can expect a modest (linear) increase in the number of control

functions. This increase is usually more than outweighed by important advantages of (20) over (19):

- Not in all cases can a meaningful inner convexification $g(\cdot)$ be found. An example are black-box simulators that can only be evaluated for certain modes (integer values), but not in between. In general, using $g(\cdot)$ for a relaxation may lead to problems such as divisions by zero, or index changes in the DAE case [6]. Evaluating the model only for vectors v^i avoids these problems;
- The integer gap between the optimal solutions of (19) and its relaxation may become arbitrarily large [56], whereas the integer gap between the optimal solutions of (20) and its relaxation is bounded by a multiple of the control discretization grid size Δt [61].

Finding a tight relaxation of (20) is vitally important for the computational solution of (1). It allows a decoupling of the MIOCP into a continuous OCP and a mixed-integer linear programming problem with a huge potential for computational savings and a posteriori bounds on the gap to the best possible MIOCP solution [37, 61]. Moreover, the relaxed control problems often have optimal integer solutions, which implies almost arbitrary computational savings when compared to an OCP-based branch&bound approach to solve (19) to optimality. This has first been compared in [40] for a benchmark problem posed in [25]. While identical solutions were obtained, a speedup of several orders of magnitude was observed for the outer convexification approach. Similar behavior can be observed when comparing to MINLP-based branch&bound. We use the truck control example from Sect. 2 to reproduce and explain this qualitative result.

Inner and Outer Convexification for Truck Control   Optimal solutions by their very nature tend to exploit constraints as much as possible. In the special case of vehicle operation and for a fixed acceleration, it is natural that the gear is chosen that provides the largest torque when compared to all other gears.
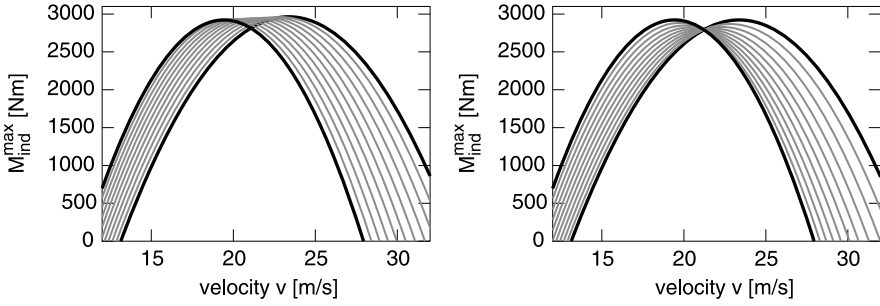
In Fig. 1 we show the maximum indicated engine torque $M_{ind,max}$ depending on the velocity v for two adjacent gear choices $\mu(\cdot) = i$ and $\mu(\cdot) = i + 1$. Figure 1 (left) shows

$$M_{ind,max}^{IC}(v) = M_{ind,max}\big(n_{eng}\big(v(s), \xi i + (1 - \xi)(i + 1)\big)\big) \qquad (21)$$

for $\xi = 0.0, 0.1, \ldots, 1.0$, while Fig. 1 (right) shows

$$M_{ind,max}^{OC}(v) = \alpha M_{ind,max}\big(n_{eng}\big(v(s), i\big)\big) + (1 - \alpha) M_{ind,max}\big(n_{eng}\big(v(s), i + 1\big)\big) \quad (22)$$

for $\alpha = 0.0, 0.1, \ldots, 1.0$. Apparently, in the particular case of truck control, the relaxation of (19) comprises combinations of the state variable $v(\cdot)$ and non-integral $\xi \in (0, 1)$ that yield an unphysical, larger value of $M_{ind,max}^{IC}$. Conversely, the relaxation of (20) is by construction maximal only for $\alpha \in \{0, 1\}$. Thus, one can expect that optimal solutions of the relaxed version of (20) are integral, whereas solutions of the relaxation of (19) may have non-integral solutions $\xi \in (0, 1)$. The two approaches would only coincide if both n_{eng} and $M_{ind,max}$ were linear functions

Fig. 1 Maximum indicated engine torque $M_{\mathrm{ind,max}}$ depending on the velocity v for two adjacent gear choices $\mu(\cdot) = i$ and $\mu(\cdot) = i+1$. *Left*: inner convexification (19). *Right*: outer convexification (20). Observe the maximal indicated engine torques for non-integral values of ξ (*gray*) on *the left hand side*, while non-integral values for α (*gray*) on *the right hand side* are never better than integral values (*black*)

Table 1 *Fractionality* as in Definition 7, *objective* function value, and *CPU* time obtained with SNOPT/IPOPT for the solution of relaxations of (19) and (20) with different control grids

# Control intervals	Inner convexification (19)			Outer convexification (20)		
	Fractionality	Objective	CPU [sec]	Fractionality	Objective	CPU [sec]
40	0.101895	1.01687	1/*	0.073163	1.03315	1/4
80	0.093632	0.99866	6/2	0.010639	1.01265	2/85
160	0.098872	0.98878	28/8	0.000409	1.00030	3/51
320	0.095983	0.98325	*/20	0.000820	0.99283	53/365

of v^i. The special case of truck control is prototypical for many (in particular time-optimal) MIOCPs, as often bang-bang solutions are optimal. However, this is not guaranteed in the general case.

To evaluate this effect, we apply both inner and outer convexification to the truck control problem (16), however for the time being without the constraints (14)–(15). For convenience, we use formulation (18) in the following. The effect is strongest for time- and energy-optimal driving, hence we consider the case where $\lambda_{\mathrm{fuel}} = 1$, $\lambda_{\mathrm{dev}} = \lambda_{\mathrm{comf}} = 0$ in the uphill setting of Fig. 2.

Table 1 shows numerical results for the cruise control of the heavy-duty truck. Both the inner and the outer convexification formulations have been relaxed and solved to local optimality with the active-set solver SNOPT [28] and the interior point solver IPOPT [69]. The most important findings can be summarized as follows:

- For inner convexification, the fractionality of the optimal relaxed gear choices $\alpha_j(\cdot)$ does not improve for finer discretizations. For the outer convexification however, it goes to zero as the control discretization grid is equidistantly refined. The reason is that the optimal control $\alpha(\cdot)$ of (20) is of bang-bang type;
- There is a gap between the lower bound from the relaxed IC solutions and the best possible integer solution, which is almost attained by the OC formulation;

- The interior point code IPOPT seems to perform better on inner convexification, the SQP code SNOPT performs better on outer convexification;
- Both codes run into problems with inner convexification instances.

Although on single instances it may be faster to compute the relaxation of the inner convexification formulation, it is clear that outer convexification is the method of choice within a branch and bound framework. This is due to the tighter relaxation (better bound) in combination with the feature that relaxed solutions are usually already almost integer feasible if the control discretization grid is chosen adequately.

Extending the Outer Convexification Approach   The advantages of outer convexification over inner convexification have stimulated additional research into the formulation (20). Here, we briefly review significant developments for the reader's convenience. First, to fully exploit the beneficial integrality property that has been exemplified above and that is related to bang–bang solutions in optimal control, direct methods need to be equipped with adaptivity in the control grid discretization and follow-up transformations into a switching time optimization, [56, 60]. Second, so-called sensitivity-seeking or path-constrained arcs (i.e., time periods during which $\alpha(\cdot)$ is not binary) can be treated with tailored *sum up rounding* strategies. This rounding strategy is a constructive part in the proof for the dependence of the integer gap on the control discretization grid size Δt. Furthermore, it is the optimal solution to a MILP that minimizes the deviation of a binary control $\omega(\cdot)$ from a relaxed one $\alpha(\cdot)$ with respect to

$$\max_{t \in [0, t_f]} \left\| \int_0^t \omega(\tau) - \alpha(\tau) \mathrm{d}\tau \right\|.$$

If additional constraints like a maximum number of switches need to be fulfilled, tailored branching algorithms can be applied to efficiently solve the constrained MILP [37]. Third, the numerical algebra to cope with the specific structures induced by the control functions $\alpha(\cdot)$ can be exploited in the context of SQP approaches [41, 42]. Finally, an extension to more general MIOCPs is possible, e.g., to multi-objective problems [48], to differential-algebraic systems [27], and to certain partial differential equations [34]. For an overview see [57, 58].

4 Constraint Formulations

In this section, we study reformulations for logically implied constraints $c(\cdot)$, i.e., $n_c > 0$ in (1). Again, we consider reformulations that are equivalent for integer control functions and discuss properties of their relaxations. We do this both in general and for the special case of the constraints (14)–(15) of the heavy-duty truck model from Sect. 2 that result in

$$c^{\text{truck}}(v, M_{\text{ind}}, \mu) = \begin{pmatrix} M_{\text{ind,max}}(n_{\text{eng}}(v, \mu)) - M_{\text{ind}} \\ n_{\text{eng,max}} - n_{\text{eng}}(v, \mu) \\ n_{\text{eng}}(v, \mu) - n_{\text{eng,min}} \end{pmatrix}. \tag{23}$$

In the interest of notational simplicity we omit the lower linear engine speed bound $n_{eng}(v, \mu) - n_{eng,min}$ in the rest of this section. We note that the function $c^{truck}(\cdot)$ has components that are quadratic and linear in $i_T(\mu)$,

$$c^{truck}(v, M_{ind}, \mu) = \begin{pmatrix} 3000 - (3i_A i_T(\mu)v/(\pi r_{stat}) - 125)^2 - M_{ind} \\ \pi r_{stat} n_{eng,max}/(30 i_A i_T(\mu)) - v \end{pmatrix}. \qquad (24)$$

4.1 Inner Convexification of the Constraints

The inner convexification formulation (19) can be augmented in a straightforward way with constraints

$$0 \le c\big(x(t), u(t), g(\psi)\big)$$

that guarantee that integer feasible solutions $\psi \in \{1, \ldots, n_\omega\}$ of (19) are feasible solutions of (1). Using (18) for $g(\cdot)$ we obtain the following reformulation of the constraints (14), (15),

$$0 \le 3000 - \left(3i_A \left(\sum_{i=1}^{n_\omega} \alpha_i i_T(i)\right) v/(\pi r_{stat}) - 125\right)^2 - M_{ind}, \qquad (25a)$$

$$0 \le \pi r_{stat} n_{eng,max} \Big/ \left(30 i_A \left(\sum_{i=1}^{n_\omega} \alpha_i i_T(i)\right)\right) - v. \qquad (25b)$$

Just as in Sect. 3, the evaluation of convex combinations within a nonlinear function may give rise to optimality of feasible fractional values, while neighboring integer values may not be optimal. Hence, the inner convexification approach can be expected to potentially yield weak relaxations.

4.2 Outer Convexification/One Row Formulation of the Constraints

The outer convexification reformulation (20) can be applied to the constraint expression as well. Residuals are evaluated for all possible choices, and the constraint is imposed on the convex combination of residuals resulting in

$$0 \le \sum_{i=1}^{n_\omega} \alpha_i(t) c\big(x(t), u(t), v^i\big). \qquad (26)$$

This reformulation avoids the problem of evaluation in fractional choices, and ensures that all feasible integer points are feasible points of the original MIOCP.

Remark 2 (LICQ for Outer Convexification) Let $(\bar{x}, \bar{u}, \bar{\alpha})$ be a feasible solution of the relaxation of problem (20), (26) and let the matrix

$$\begin{bmatrix} \sum_{i=1}^{n_\omega} \bar{\alpha}_i(t)\nabla c(\cdot, v^i) \\ \nabla d(\cdot) \end{bmatrix} := \frac{\partial}{\partial(x,u)} \begin{bmatrix} \sum_{i=1}^{n_\omega} \bar{\alpha}_i(t)c(\bar{x}, \bar{u}, v^i) \\ d(\bar{x}, \bar{u}) \end{bmatrix}$$

of active constraints have full row rank. Drop the upper bounds $\alpha_i(\cdot) \leq 1$, which are implicitly implied by $\alpha_i(\cdot) \geq 0$ and $\sum_i \alpha_i(\cdot) = 1$. Then LICQ is satisfied for the relaxation of problem (20), (26).

Proof We look at the constraint matrix of all (active) constraints in $(\cdot, \bar{\alpha})$, given by

$$\frac{\partial}{\partial(\alpha_1, \ldots, \alpha_{n_\omega}, (x,u))} \begin{bmatrix} \sum_{i=1}^{n_\omega} \alpha_i(t)c(\cdot, v^i) \\ d(\cdot) \\ \sum_{i=1}^{n_\omega} \alpha_i(t) - 1 \\ \alpha_1(t) \\ \cdots \\ \alpha_{n_\omega}(t) \end{bmatrix}$$

$$= \begin{bmatrix} c(\cdot, v^1) & \cdots & c(\cdot, v^{n_{n_\omega}}) & \sum_{i=1}^{n_{n_\omega}} \bar{\alpha}_i(t)\nabla c(\cdot, v^i) \\ 0 & \cdots & 0 & \nabla d(\cdot) \\ 1 & \cdots & 1 & 0 \\ 1 & & & 0 \\ & \ddots & & \vdots \\ & & 1 & 0 \end{bmatrix}$$

Due to feasibility of $\bar{\alpha}$ and the SOS-1 constraint there exists at least one index $1 \leq i \leq n_\omega$ for which $\alpha_i(t) \geq 0$ is not active. Hence, the bottom n_ω rows are linearly independent. The first block of rows may contribute an entry in column i, but by assumption is linearly independent from the other rows due to the right-most block of columns, which concludes the proof. $\square$

The same holds true if we eliminate one of the multipliers α_i using the SOS-1 constraint, which becomes an inequality constraint. For the heavy-duty truck model, the torque constraint (14) and the engine speed constraint (15) are reformulated to read

$$0 \leq \sum_{i=1}^{n_\mu} \alpha_i(t)M_{\text{ind,max}}\big(n_{\text{eng}}(v(s), i)\big) - M_{\text{ind}}(s), \tag{27a}$$

$$0 \leq n_{\text{eng,max}} - \sum_{i=1}^{n_\mu} \alpha_i(t)n_{\text{eng}}\big(v(s), i\big). \tag{27b}$$

Note that (25b) and (27b) are identical, as $i_T(\mu)$ enters linearly.

As the constraints are summed up, *compensatory effects* may lead to feasible residuals for fractional values of the convex multipliers in both cases, as observed in [39]. For example, a first gear choice leading to an engine speed violating the upper bound $n_{\text{eng,max}}$ and a second gear choice in violation of the lower bound $n_{\text{eng,min}}$ may form a feasible convex combination in (27b).

4.3 Complementarity Formulation

To address the problem of compensatory effects, [39] proposes to enforce feasibility individually for each possible choice of the integer control via

$$0 \le \alpha_i(t)c\big(x(t), u(t), v^i\big), \quad 1 \le i \le n_{n_\omega}. \tag{28}$$

For the heavy-duty truck model, torque and engine speed constraints are

$$0 \le \alpha_i(t)\big(M_{\text{ind,max}}\big(n_{\text{eng}}\big(v(s), i\big)\big) - M_{\text{ind}}(s)\big), \tag{29a}$$

$$0 \le \alpha_i(t)\big(n_{\text{eng,max}} - n_{\text{eng}}\big(v(s), i\big)\big) \tag{29b}$$

for $1 \le i \le n_\mu$. It is obvious that optimal solutions with nonzero convex multiplier $\alpha_i(t) > 0$ are now feasible for $\alpha_i(t) = 1$ as well. Compared to the outer convexification formulation, the number of constraints has increased from 4 to $4n_\mu$, though. More important, due to the structure of the constraints (29a)–(29b), the NLP obtained from discretization of the relaxed convexified MIOCP now is a MPVC in the form of (5). In the case of equality constraints, we obtain a MPCC, see (4).

4.4 Addressing the Complementarity Formulation

As mentioned in the introduction, MPCCs and MPVCs lose constraint qualification. This can also be seen directly.

Remark 3 (No LICQ for Complementarity) Let $(\bar{x}, \bar{u}, \bar{\alpha})$ be a feasible solution of the relaxation of problem (20), (28) with

$$\bar{\alpha}_i(t) = 0, \qquad c\big(\bar{x}, \bar{u}, v^i\big) = 0$$

for at least one $1 \le i \le n_\omega$. Then LICQ is not satisfied.

Proof We look at the constraint matrix of all constraints in $(\cdot, \bar{\alpha})$, given by

$$
\frac{\partial}{\partial(\alpha_1, \ldots, \alpha_{n_\omega}, (x, u))}
\begin{bmatrix}
\alpha_1(t)c(\cdot, v^1) \\
\cdots \\
\alpha_{n_\omega}(t)c(\cdot, v^{n_\omega}) \\
d(\cdot) \\
\sum_{i=1}^{n_\omega} \alpha_i(t) - 1 \\
\alpha_1(t) \\
\cdots \\
\alpha_{n_\omega}(t)
\end{bmatrix}
$$

$$
=
\begin{bmatrix}
c(\cdot, v^1) & \cdots & 0 & \bar{\alpha}_1(t)\nabla c(\cdot, v^1) \\
& \ddots & & \vdots \\
0 & \cdots & c(\cdot, v^{n_{n_\omega}}) & \bar{\alpha}_{n_\omega}(t)\nabla c(\cdot, v^{n_\omega}) \\
0 & \cdots & 0 & \nabla d(\cdot) \\
1 & \cdots & 1 & 0 \\
1 & & & 0 \\
& \ddots & & \vdots \\
& & 1 & 0
\end{bmatrix}
$$

and note that the rows that correspond to $\bar{\alpha}_i(t)c(\cdot, v^i) \geq 0$ contain only zeros, leading to a trivially linear dependent constraint system. □

The constraint system also violates the weaker Mangasarian-Fromovitz constraint qualification. However, some cone-based constraint qualifications are satisfied, e.g., [2]. Thus, local minimizers are still KKT points.

Complementarity Pivoting Techniques   Computational approaches for solving MPCCs and MPVCs directly are mentioned in the introduction, but generally rely on dedicated implementations of complementarity solvers. If one intends to use existing numerical software, then the complementarity reformulation (28) needs to be addressed by regularization or smoothing techniques that recover LICQ.

Regularization and Smoothing Techniques   These techniques follow the principle of reformulation of the MPCC or MPVC in violation of LICQ using a parameter $\varepsilon > 0$ in a way that recovers constraint qualification for $\varepsilon \in (0, \overline{\varepsilon})$. Then, convergence of the sequence of local minimizers to a limit point is established for $\varepsilon \to 0$. The assessment of the type of stationarity obtained for the limit point is crucial. For MIOCP, we find it essential to ask for *Bouligand stationarity*. Lesser stationarity concepts allow for termination in *spurious stationary points* that possess trivial descent directions, thus giving rise to "missed" switches of the integer control.

In the following, an overview over different possible reformulations is given, and their general form is presented and applied to the particular truck problem.

Regularization Reformulations   Conversely, regularization formulations relax the feasible set using $\varepsilon > 0$ by requiring

$$-\varepsilon \le \alpha_i(t) \cdot c\big(x(t), u(t), v^i\big), \quad 1 \le i \le n_{n_\omega} \quad \text{(MPVC)},$$

or, in what is sometimes referred to as the *lumped* formulation,

$$-\varepsilon \le \sum_{i=1}^{n_{n_\omega}} \alpha_i(t) \cdot c\big(x(t), u(t), v^i\big). \tag{30}$$

For MIOC and in particular for the truck model, the lumped formulation will satisfy LICQ but amounts to applying outer convexification to the original constraint, see Sect. 4.2. For the truck model, the first formulation reads

$$-\varepsilon \le \alpha_i(t) \cdot \big(M_{\text{ind,max}}\big(n_{\text{eng}}(v(s), i)\big) - M_{\text{ind}}(s)\big) \tag{31a}$$

$$-\varepsilon \le \alpha_i(t) \cdot \big(n_{\text{eng,max}} - n_{\text{eng}}(v(s), i)\big) \tag{31b}$$

NCP Function Reformulations   A function $\phi : \mathbb{R}^2 \to \mathbb{R}$ is called *NCP*-function if

$$\phi(a, b) = 0 \quad \Rightarrow \quad a \ge 0, b \ge 0, ab = 0, \quad \text{(MPCC)},$$

$$\phi(a, b) = 0 \quad \Rightarrow \quad b \ge 0, ab \le 0, \quad \text{(MPVC)},$$

see, e.g., [18] for a survey. For MPCCs the Fischer-Burmeister function

$$\phi^{\text{FB}}(a, b) := a + b - \sqrt{a^2 + b^2}$$

may be used, [19]. For MPVC, [36] uses the NCP function

$$\phi^{\text{VC}}(a, b) = \frac{1}{2}\big(ab + \sqrt{a^2 b^2} + \sqrt{b^2} - b\big). \tag{32}$$

Piecewise smooth complementarity functions to replace the complementarity constraint in a MPCC are developed in [46], and a one-to-one correspondence between strongly stationary points and KKT points of the reformulated MPCC is established. The non-smoothness is shown to not impede the convergence of SQP methods. The approach however cannot avoid convergence of SQP solvers to spurious stationary points in the degenerate case. The NCP formulation for the truck model is

$$0 \ge \phi^{\text{VC}}\big(M_{\text{ind}}(s) - M_{\text{ind,max}}\big(n_{\text{eng}}(v(s), i)\big), \alpha_i(t)\big), \tag{33a}$$

$$0 \ge \phi^{\text{VC}}\big(n_{\text{eng}}(v(s), i) - n_{\text{eng,max}}, \alpha_i(t)\big). \tag{33b}$$

Smoothing Reformulations Smoothing reformulations have first been developed for MPCCs and work by appropriately including a smoothing parameter $\varepsilon > 0$:

$$0 \geq \phi_\varepsilon\big(-c\big(x(t), u(t), v^i\big), \alpha_i(t)\big), \quad 1 \leq i \leq n_{n_\omega} \quad \text{(MPVC)}, \tag{34}$$

with NCP function ϕ, e.g., $\phi_\varepsilon^{\text{FB}}(a, b) := a + b - \sqrt{a^2 + b^2 + \varepsilon}$ for MPCC and

$$0 \geq \phi_\varepsilon^{\text{VC}}(a, b) := \frac{1}{2}\big(ab + \sqrt{a^2 b^2 + \varepsilon^2} + \sqrt{b^2 + \varepsilon^2} - b\big) \tag{35}$$

for MPVC. For MPCC, smoothing under an LICQ-type assumption and a second order condition has been shown to yield B-stationarity of the limit point. For MPVC, smoothing is shown in [36] to not yield a satisfactory framework for solving MPVCs. For the truck model, the smoothing reformulation reads

$$0 \geq \phi_\varepsilon^{\text{VC}}\big(M_{\text{ind}}(s) - M_{\text{ind,max}}\big(n_{\text{eng}}\big(v(s), i\big)\big), \alpha_i(t)\big), \tag{36a}$$

$$0 \geq \phi_\varepsilon^{\text{VC}}\big(n_{\text{eng}}\big(v(s), i\big) - n_{\text{eng,max}}, \alpha_i(t)\big). \tag{36b}$$

Smoothing-Regularization Reformulations Smoothing-regularization approaches combine the concepts of smoothing and regularization to form

$$\varepsilon \geq \phi_\varepsilon\big(-c\big(x(t), u(t), v^i\big), \alpha_i(t)\big), \quad 1 \leq i \leq n_{n_\omega} \quad \text{(MPVC)}. \tag{37}$$

For MPVC, using (32), the smoothing-relaxation formulation reads $\varepsilon \geq \phi_\varepsilon^{\text{VC}}(a, b)$. Convergence for $\varepsilon \to 0$ to a B-stationary limit point is established under a weak LICQ-type constraint qualification, assuming existence and asymptotic nondegeneracy of a sequence $\{x_\varepsilon\}_{\varepsilon \to 0}$ of feasible points for the sequence NLP(ε). LICQ is shown to be satisfied for all NLP(ε) with $\varepsilon \in (0, \overline{\varepsilon})$.

For the truck problem, the smoothing-regularization formulation reads

$$\varepsilon \geq \phi_\varepsilon^{\text{VC}}\big(M_{\text{ind}}(s) - M_{\text{ind,max}}\big(n_{\text{eng}}\big(v(s), i\big)\big), \alpha_i(t)\big), \tag{38a}$$

$$\varepsilon \geq \phi_\varepsilon^{\text{VC}}\big(n_{\text{eng}}\big(v(s), i\big) - n_{\text{eng,max}}, \alpha_i(t)\big). \tag{38b}$$

4.5 Generalized Disjunctive Programming

Based on the work of Balas for integer linear programs, Grossmann and coworkers developed Generalized Disjunctive Programming (GDP) for mixed-integer nonlinear programs [31]. The problem formulation (1) uses disjunctive notation, hence it is natural to take the GDP point of view. It is motivated by disjunctions of the form

$$\min_{y,Y} e(y) + \sum_{i \in I} e_i$$

$$\text{s.t.} \quad \begin{bmatrix} Y_i \\ g_i(y) \leq 0 \\ e_i = \gamma_i \end{bmatrix} \vee \begin{bmatrix} \neg Y_i \\ B_i y = 0 \\ e_i = 0 \end{bmatrix} \quad \forall i \in I \tag{39}$$

$$0 \leq y \leq U, \qquad \Phi(Y) = \text{true}, \quad Y \in \{\text{false, true}\}^{|I|}$$

with application in process synthesis where Y_i represents presence or absence of units and y a vector of continuous variables. If a corresponding unit is not used, the equation $B_i y = 0$ eliminates variables and $e_i = 0$ sets the costs to zero.

As stated above, logical relations $\Phi(Y) = \text{true}$ can be translated into constraints with binary variables $z \in \{0, 1\}^{|I|}$. An interesting question, however, is how the disjunctions are formulated. We consider two approaches.

Big-M   Using large enough constants M is a well-known technique to model logical relations in combinatorial optimization. For (39), this yields

$$g_i(y) \leq M(1 - z_i), \qquad -M z_i \leq B_i y \leq M z_i, \qquad e_i = z_i \gamma_i.$$

Convex Hull Reformulation   A possibly tighter relaxation can be obtained from the nonlinear convex hull reformulation. It makes use of the perspective of a function.

Definition 8 (Perspective function) The *perspective* of a function $f : \mathbb{R}^n \to \mathbb{R}$ is the function $\hat{f} : \mathbb{R}^{n+1} \to \mathbb{R}$ defined by

$$\hat{f}(\lambda, y) = \begin{cases} \lambda f(y/\lambda) & \text{if } \lambda > 0, \\ 0 & \text{if } \lambda = 0, y = 0, \\ \infty & \text{otherwise.} \end{cases}$$

An important property is that if $f(\cdot)$ is convex, then $\hat{f}(\cdot)$ is also convex. Perspectives have been used for strong formulations of MINLPs [15] and have been gaining increasing interest lately also for the derivation of *perspective cuts*, e.g., [33]. Perspectives can be used to obtain the nonlinear convex hull of a feasible set as follows:

Definition 9 (Nonlinear Convex Hull) Problem (6) can be equivalently restated as minimization over all (y, λ) of the nonconvex function

$$\Phi(y, \lambda) := \begin{cases} 0 & \text{if } \lambda = 0, y = 0, \\ F(y) & \text{if } \lambda = 1, G(y) \leq b, \\ \infty & \text{else.} \end{cases} \tag{40}$$

wherein we have assumed w.l.o.g. that $\Phi(0, \lambda) = 0$. The closure $\overline{co}\Phi(y, \lambda)$ of the convex hull of Φ is given by

$$\overline{co}\Phi(y, \lambda) = \begin{cases} 0 & \text{if } \lambda = 0, y = 0, \\ \lambda F(y/\lambda) & \text{if } \lambda \in (0, 1], G(y) \leq \lambda b, \\ \infty & \text{else.} \end{cases} \tag{41}$$

The convex hull representation of the GDP (39) is then obtained by introduction of additional variables $0 \leq \lambda_{ij} \leq 1$, $\lambda_{i1} + \lambda_{i0} = 1$ and $0 \leq v_{ij} \leq \lambda_{ij} U$, $y = v_{i1} + v_{i0}$ and constraints $0 \geq \lambda_{i1} g_i(v_{i1}/\lambda_{i1})$, $B_i v_{i0} = 0$, $f_i = \lambda_{i1} \gamma_i$.

Numerical Difficulties For constraint formulations using perspectives usually LICQ holds. The reason is that unlike in Remark 3, the derivative $\lambda \nabla_v g(v/\lambda) \neq 0$ for $\lambda = 0$. E.g., for linear constraints $g(v) = av + b$ we have $\lambda \nabla_v g(v/\lambda) = a$.

However, for nonlinear $g(\cdot)$ there are computational challenges due to division by zero or near-zero values. For $\varepsilon > 0$ sufficiently small $\lambda f(y/\lambda)$ can be approximated by $(\lambda + \varepsilon) f(y/(\lambda + \varepsilon))$, [31], or by $(\lambda) f(y/(\lambda + \varepsilon))$, [50], or as proposed in the Ph.D. thesis of Nicolas Sawaya by $((1 - \varepsilon)\lambda + \varepsilon) f(y/((1 - \varepsilon)\lambda + \varepsilon))$. The results may depend heavily on the choice of ε and be computationally challenging.

4.6 Generalized Disjunctive Programming for MIOCP

First, we consider the Big-M formulation for GDP. For general constraints $c(\cdot)$ and SOS-1 variables $\alpha(\cdot)$ we obtain for $1 \leq i \leq n_{n_\omega}$

$$M(\alpha_i(t) - 1) \leq c(x(t), u(t), v^i) \tag{42}$$

and thus for the truck constraints (23) for $1 \leq i \leq n_\mu$

$$M(\alpha_i(t) - 1) \leq (M_{\text{ind,max}}(n_{\text{eng}}(v(s), i)) - M_{\text{ind}}(s)), \tag{43a}$$

$$M(\alpha_i(t) - 1) \leq (n_{\text{eng,max}} - n_{\text{eng}}(v(s), i)). \tag{43b}$$

This formulation does not cause LICQ problems and good values for M can be determined from bounds on $v(\cdot)$. Still, it is expected to give weak relaxations.

Second, we apply the GDP convex hull technique to the MIOCP (1). [50] proposed to lift all variables (controls, differential states, horizon lengths). We describe one of the many possibilities that stem from a modeler's freedom to lift only some of the variables and to use different formulations for the disjunctive constraints.

One first observation is that we can use the convex multiplier functions $\alpha_i(\cdot)$ from outer convexification in place of λ, and let $y(\cdot) := (x(\cdot), u(\cdot))$. The perspective formulation for the constraints $c(\cdot)$ then yields

$$0 \leq \alpha_i(t) c(x(t)/\alpha_i(t), u(t)/\alpha_i(t), v^i), \quad t \in [0, t_\text{f}], i = 1, \ldots, n_\omega. \tag{44}$$

To properly obtain the convex hull, we also apply this procedure to the dynamic constraint for $t \in [0, t_\text{f}]$, $i = 1, \ldots, n_\omega$:

$$0 = \alpha_i(t)\big(\dot{x}(t)/\alpha_i(t) - f\big(x(t)/\alpha_i(t), u(t)/\alpha_i(t), v^i\big)\big) \tag{45a}$$

$$= \dot{x}(t) - \alpha_i(t) f\big(x(t)/\alpha_i(t), u(t)/\alpha_i(t), v^i\big). \tag{45b}$$

We introduce disaggregated state derivatives $\dot{x}^i(\cdot)$, resulting states $x^i(\cdot)$, and disaggregated control decision variables $u^i(\cdot)$ for all nonlinearly entering controls:

$$\dot{x}(t) = \sum_{i=1}^{n_\omega} \dot{x}^i(t), \qquad x(t) = \sum_{i=1}^{n_\omega} x^i(t), \qquad u(t) = \sum_{i=1}^{n_\omega} u^i(t).$$

We consider a time discretization of the states of problem (1),

$$x(t_k) := s_k, \qquad \dot{x}(t) = f\big(x(t), u(t), v(t)\big), \quad t \in [t_k, t_{k+1}], \tag{46a}$$

$$s_{k+1} = x\big(t_{k+1}; t_k, s_k, u(\cdot), v(\cdot)\big), \quad 0 \le k \le N-1, \tag{46b}$$

which yields the MIOCP

$$\min_{x(\cdot), u(\cdot), Y} e(s_N)$$

$$\text{s.t.} \quad \bigvee_{i \in \{1,\dots,n_\omega\}} \begin{bmatrix} Y_{ik} \\ \dot{x}(t) = f(x(t), u(t), v^i) \\ x(t_k) = s_k \\ s_{k+1} = x(t_{k+1}; t_k, s_k, u(\cdot), v^i) \\ 0 \le c(x(t), u(t), v^i) \end{bmatrix} \quad \forall t \in [t_k, t_{k+1}] \tag{47}$$

$$0 \le d\big(x(t), u(t)\big) \quad \forall t \in [t_k, t_{k+1}]$$

with $k = 0, \dots, N-1$, for which the convex hull formulation is, with piecewise constant controls $\alpha_i(t) = \alpha_{ik}$ and $t \in [t_k, t_{k+1}]$ for $k = 0, \dots, N-1$,

$$\min_{s, u, \alpha} e(s_N)$$

$$\text{s.t.} \quad \dot{x}^i(t) = \alpha_{ik} f\big(x^i(t)/\alpha_{ik}, u^i(t)/\alpha_{ik}, v^i\big)$$

$$x^i(t_k) = \alpha_{ik} s_k$$

$$s_{k+1} = \sum_{i=1}^{n_\omega} x^i\big(t_{k+1}; t_k, s_k, u^i(\cdot)/\alpha_{ik}, v^i\big) \tag{48}$$

$$0 \le \alpha_{ik} c\big(x^i(t)/\alpha_{ik}, u^i(t)/\alpha_{ik}, v^i\big)$$

$$0 \le d\left(\sum_{i=1}^{n_\omega} x^i(t), \sum_{i=1}^{n_\omega} u^i(t)\right)$$

$$\sum_{i=1}^{n_\omega} \alpha_{ik} = 1, \qquad 0 \le x^i(t) \le \alpha_{ik} M^s, \qquad 0 \le u^i(t) \le \alpha_{ik} M^u.$$

Note that for the purpose of the disjunctive term controlled by Y_{ik}, s_k is a constant when forming the perspective of the initial value constraint. Moreover, as the s_{k+1} enter linearly, we do not disaggregate them but instead aggregate the IVP results. Next, we enforce path constraints $0 \le x^i(\cdot) \le \alpha_{ik}M^s$, $c(\cdot)$ and $d(\cdot)$ in the discretization points $t = t_k$ only, as is customary in direct optimal control. The respective constraints of (48) are replaced by

$$0 \le s_k \le M^s \tag{49a}$$

$$0 \le \alpha_{ik}c\big(s_k, u^i(t_k)/\alpha_{ik}, v^i\big) \tag{49b}$$

$$0 \le d\left(s_k, \sum_{i=1}^{n_\omega} u^i(t_k)\right). \tag{49c}$$

We further substitute u by $\bar{u}_i(t) = u_i(t)/\alpha_{ik}$. This poses no problem for $\alpha_{ik} = 0$ due to the bound $u_i(t) \le \alpha_{ik}M_k^u$, and yields lifted controls $\bar{u}_i(t)$ with $0 \le \bar{u}^i(t_k) \le M^u$:

$$\dot{x}^i(t) = \alpha_{ik}f\big(x^i(t)/\alpha_{ik}, \bar{u}^i(t_k), v^i\big) \tag{50a}$$

$$0 \le \alpha_{ik}c\big(s_k, \bar{u}^i(t_k), v^i\big) \tag{50b}$$

$$0 \le d\left(s_k, \sum_{i=1}^{n_\omega} \alpha_{ik}\bar{u}^i(t_k)\right) \tag{50c}$$

Since this formulation still poses numerical difficulties, we go one step further and modify the problem by aggregating the states over all time steps and not only during the shooting intervals. This idea comes from the observation that, except for the ODE constraint, the $x^i(\cdot)$ enter only via their convex combination. Hence, we replace $x^i(t)$ by $\alpha_{ik}x(t)$ and obtain $\sum_{i=1}^{n_\omega} x^i(t_{k+1}) = \sum_{i=1}^{n_\omega} \alpha_{ik}x(t_{k+1}) = x(t_{k+1})$ and thus (51)

$$\min_{s,\bar{u},\alpha} e(s_N)$$

$$\text{s.t.} \quad \dot{x}(t) = \sum_{i=1}^{n_\omega} \alpha_{ik}f\big(x(t), \bar{u}^i(t), v^i\big)$$

$$x(t_k) = s_k$$

$$s_{k+1} = x\big(t_{k+1}; t_k, s_k, \bar{u}(\cdot)\big) \tag{51}$$

$$0 \le \alpha_{ik}c\big(s_k, \bar{u}^i(t_k), v^i\big)$$

$$0 \le d\left(s_k, \sum_{i=1}^{n_\omega} \alpha_{ik}\bar{u}^i(t)\right)$$

$$\sum_{i=1}^{n_\omega} \alpha_{ik} = 1, \qquad 0 \le s_k \le M^s, \qquad 0 \le \bar{u}^i(t) \le M^u$$

Table 2 Comparison of constraint reformulations for $\alpha_i = 1 \Rightarrow \gamma_i \leq v \leq \Gamma_i$ and their properties

Name	Formulation	Advantages	Drawbacks
Inner convexification Section 4.1, Eqs. (25a)–(25b)	$\gamma(\psi) \leq v \leq \Gamma(\psi)$ $\gamma(i) = \gamma_i, \Gamma(i) = \Gamma_i$	Easy to formulate very fast to solve	Integer infeasible, fractional evaluations, very weak relaxation
Outer convexification Section 4.2, Eqs. (27a)–(27b)	$\sum_i \alpha_i \gamma_i \leq v \leq \sum_i \alpha_i \Gamma_i$	Evaluation in integer points, fast to solve	Compensation effects
MPVC Section 4.3, Eqs. (29a)–(29b)	$\alpha_i \gamma_i \leq \alpha_i v \leq \alpha_i \Gamma_i$	Guarantees integer feasibility	CQs violated, needs tailored methods
MPVC reformulations Section 4.4, Eqs. (31a)–(31b), (36a)–(36b), (38a)–(38b)	Regularization, smoothing, combined	Use existing NLP solvers	Sequence of problems, ill-conditioned, ε-relaxed solution
GDP Big-M Section 4.5, Eqs. (43a)–(43b)	$\gamma_i - M(1 - \alpha_i) \leq v$ $\leq \Gamma_i + M(1 - \alpha_i)$	CQs, fast	Weak relaxation
GDP Convex Hull Section 4.6, Eq. (51) with Sect. 4.4, Eqs. (31a)–(31b)	Disaggregated controls constraints as MPVC (reformulation)	More degrees of freedom	More variables, in addition to the drawbacks of MPVC (reformulations)

is partial outer convexification of the ODE, uses the vanishing constraint formulation for the constraints, but differs in the disaggregation of the controls $\bar{u}^i(\cdot)$, which gives the system more freedom to potentially find a better solution. Note that this particular way of making use of a GDP formulation needs to be smoothened again, e.g., using (38a)–(38b). For the application at hand this approach was superior to a full lifting and the difficult numerical treatment of the constraints that are quadratic in $1/\alpha_i(\cdot)$.

Table 2 summarizes the reformulations for constraints directly depending on an integer control that have been discussed in this section.

5 Numerical Results

In this section we illustrate above concepts by applying the different formulations to the engine speed constraints of the heavy duty truck.

Reproducible Benchmark Implementations We aim to make the heavy-duty truck cruise control problem available to the community as a MIOCP benchmark problem. A complete description comprises scenario data for $\gamma(s)$ and $v_{\text{law}}(s)$ that characterizes the route to be traveled, with positioning information s assumed to be

---

Algorithm 1: Homotopy method for problems presented in Sect. 4.4 and Sect. 4.6

---

1 $\delta = 0.6, \varepsilon^* = 10^5$
2 while $\varepsilon^* > 1e - 3$ **do**
3 $\varepsilon = \delta \cdot \varepsilon^*$
4   Solve the problem corresponding to ε starting from last solution σ^*
5   **if** terminal point infeasible, or cannot restore feasibility of the initial point
      then $\delta = 1.6 \cdot \delta$ **else** store solution as $\sigma^*, \varepsilon^* = \varepsilon, \delta = \delta/1.2$
6 end

---

available. It also comprises $M_{\text{ind,max}} := 3000 - (n_{\text{eng}}(s, \mu(s)) - 1250)^2/100$, M_{fric}, Q_{fuel}, $n_{\text{eng,min}}$, and $n_{\text{eng,max}}$ as vehicle- and engine-specific nonlinear data sets. This data as well as other parameter values, functions, and initial values is available on the web page [55] and in the thesis [39]. Moreover, this web page includes AMPL models for all reformulations we use throughout this article.

In the following calculations, we discretize all control problems using an implicit Euler method with 40 equidistant steps on the horizon [0, 1000].

Computational Setup All results were computed on a single core of an Intel Core i7-2600 CPU with 3.40 GHz and 8 GB memory. IPOPT 3.10.0 and SNOPT 7.2-8 were run with the standard solver options, invoked from AMPL version 20120505. For the homotopy methods, IPOPT warm start options were added. The homotopy method starts with a big value for ε to find a solution for the control problem without vanishing constraints. The last solution σ^* is stored together with the corresponding ε^*. ε^* is then adjusted by a factor δ to obtain a new ε as in Algorithm 1.

Note that also $\delta > 1$ is allowed, since sometimes the last solution cannot be restored even with the same ε due to numerical difficulties.

Discussion of Scenarios We present numerical results for two selected scenarios in Fig. 2 and Fig. 3. The track's height profile is shown on top, followed by eight plots of the relaxed (local) optimal solutions identified by IPOPT for the formulations listed in Table 2. From both scenarios, a clear picture emerges. Inner convexification solutions suffer from significant infeasibility. Outer convexification is computationally the fastest formulation, and yields reasonable approximations that in parts suffer from compensatory effects. In contrast, the three vanishing constraint formulations succeed in yielding feasible solutions for both scenarios. For scenario 2, these solutions are even integer feasible, while for scenario 1 there are at most two control intervals with fractional gear choices. For the feasible vanishing constraint solutions, we can also compare the resulting objective functions. Here, the simple regularized formulations performs best, with the smoothing-regularization and the plain vanishing constraint formulation tied in second place. The GDP Big-M formulation does not yield feasible solutions, while our convex hull variant of GDP performs slightly better than outer convexification also in terms of integer feasibility. It suffers from a high computational runtime, though. The combination of integer

Scenario 1
$\lambda_{\text{fuel}} = 25$
$\lambda_{\text{dev}} = 1$
$\lambda_{\text{comf}} = 1$

Unconstrained
Constraints   -
Objective   582.784
Fractionality 0.1
Infeasibility $3e+03$
Runtime   2.82 sec

Inner Convexification
Constraints   (25a)-(25b)
Objective   54417.3
Fractionality 0.17
Infeasibility $6e+02$
Runtime   0.852 sec

Outer Convexification
Constraints   (27a)-(27b)
Objective   58144.4
Fractionality 0.24
Infeasibility 1.3
Runtime   4.61 sec

Vanishing Constraints
Constraints   (29a)-(29b)
Objective   70210.7
Fractionality 0.041
Infeasibility 0
Runtime   35.8 sec

Relaxed VCs
Constraints   (31a)-(31b)
Objective   60222.7
Fractionality 0.0077
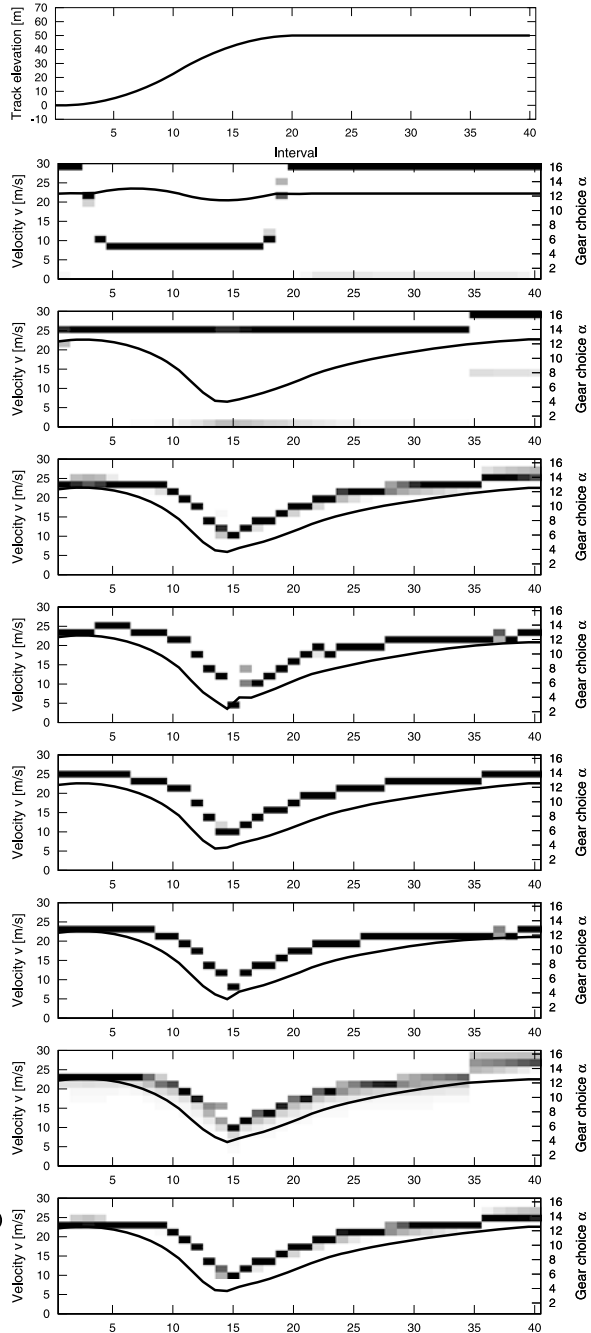Infeasibility $1.6e-05$
Runtime   18.8 sec

Smoothed VCs
Constraints   (36a)-(36b)
Objective   61364
Fractionality 0.022
Infeasibility $1e-06$
Runtime   40.6 sec

GDP Big-*M*
Constraints   (43a)-(43b)
Objective   52976.1
Fractionality 0.62
Infeasibility 51
Runtime   8.35 sec

GDP Convex Hull
Constraints   (51)+(38a)-(38b)
Objective   58241.3
Fractionality 0.19
Infeasibility 0.96
Runtime   357 sec

Fig. 2 Relaxed gear choices α_j (reflected by *the intensity of gray*) and corresponding velocities

Scenario 2
$\lambda_{\text{fuel}} = 10$
$\lambda_{\text{dev}} = 100$
$\lambda_{\text{comf}} = 1$

Unconstrained
Constraints   -
Objective   17352.1
Fractionality 0.31
Infeasibility $5.8e + 03$
Runtime   3.42 sec

Inner Convexification
Constraints   (25a)-(25b)
Objective   3748470
Fractionality 0.35
Infeasibility $7e + 02$
Runtime   1.34 sec

Outer Convexification
Constraints   (27a)-(27b)
Objective   3960800
Fractionality 0.33
Infeasibility 6
Runtime   5.15 sec

Vanishing Constraints
Constraints   (29a)-(29b)
Objective   774479
Fractionality 0.064
Infeasibility 0
Runtime   11.8 sec

Relaxed VCs
Constraints   (31a)-(31b)
Objective   4089720
Fractionality 0
Infeasibility $1.4e - 05$
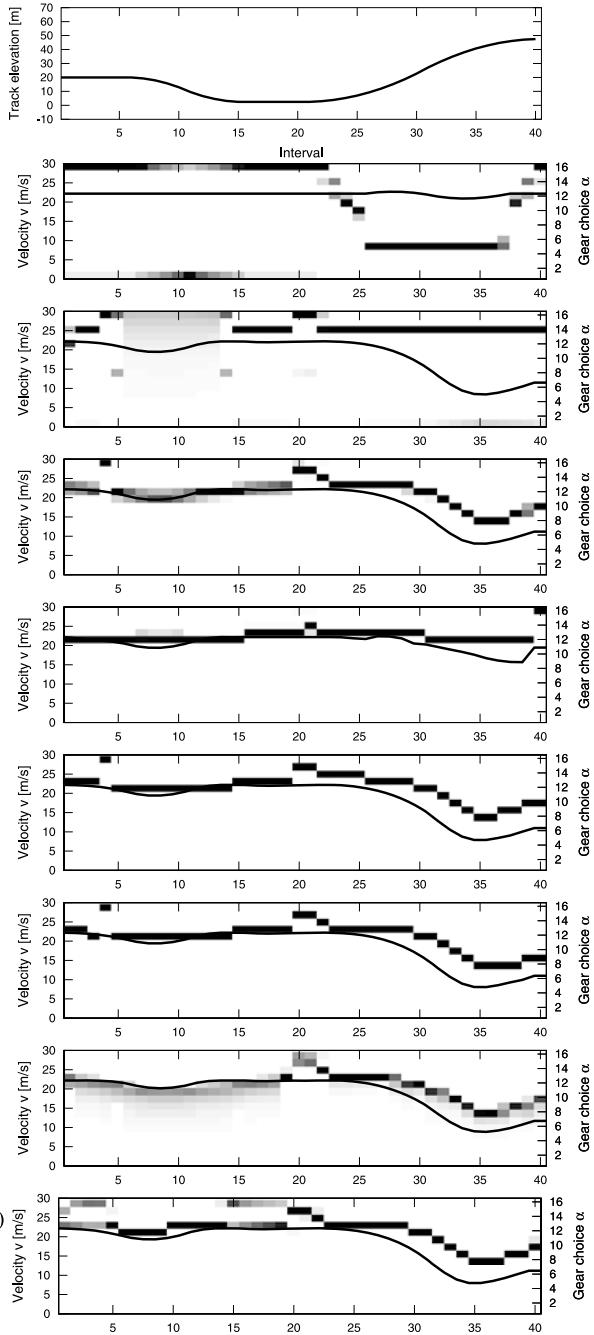Runtime   19.2 sec

Smoothed VCs
Constraints   (36a)-(36b)
Objective   4014800
Fractionality $1e - 06$
Infeasibility $1e - 06$
Runtime   32.4 sec

GDP Big-*M*
Constraints   (43a)-(43b)
Objective   3507270
Fractionality 0.69
Infeasibility $2e + 02$
Runtime   4.68 sec

GDP Convex Hull
Constraints   (51)+(38a)-(38b)
Objective   4002530
Fractionality 0.19
Infeasibility 0.62
Runtime   368 sec

Fig. 3 Relaxed gear choices α_j (reflected by *the intensity of gray*) and corresponding velocities

control functions, nonlinear constraints, and nonlinear differential equations allows for a variety of different GDP implementations within a direct multiple shooting or direct collocation framework and may allow for significant better performance.

6 Future Developments

From the computational results, some evidence can be read in favor of vanishing constraint formulations. For the case of equality constraints depending on the integer control, complementarity constraints are obtained. In view of the computational inefficiency and the remaining infeasibility of homotopy based approaches, it will be desirable to treat MPCCs and MPVCs in their rigorous formulation of Sect. 4.3, Eqs. (29a)–(29b) in the future. SLP-EQP methods, first described by [20], can do so efficiently when the LP subproblem is replaced by the corresponding LPCC [17] or LPVC. Moreover, they can be shown to identify B-stationary points of such problems [47]. The efficiency improvements one would gain from such a computational setup would also make the application to closed-loop mixed-integer control [39] viable.

It will be interesting to study all constraints in the context of a Branch&Cut framework. Furthermore a comparison of global solutions would be helpful. As generic global solvers like COUENNE cannot yet solve the benchmark, a specific dynamic programming approach seems the best choice.

Overcoming the numerical problems associated with perspective functions and a better exploitation of the numerical features of the resulting optimal control problems are yet other challenges.

References

1. Abichandani, P., Benson, H., Kam, M.: Multi-vehicle path coordination under communication constraints. In: American Control Conference, pp. 650–656 (2008)
2. Achtziger, W., Kanzow, C.: Mathematical programs with vanishing constraints: optimality conditions and constraint qualifications. Math. Program., Ser. A **114**, 69–99 (2008)
3. Anitescu, M., Tseng, P., Wright, S.: Elastic-mode algorithms for mathematical programs with equilibrium constraints: global convergence and stationarity properties. Math. Program., Ser. A **110**, 337–371 (2007)
4. Balas, E.: Disjunctive programming and a hierarchy of relaxations for discrete optimization problems. SIAM J. Algebr. Discrete Methods **6**, 466–486 (1985)
5. Bär, V.: Ein Kollokationsverfahren zur numerischen Lösung allgemeiner Mehrpunktrandwert aufgaben mit Schalt- und Sprungbedingungen mit Anwendungen in der optimalen Steuerung und der Parameteridentifizierung. Diploma thesis, Rheinische Friedrich-Wilhelms-Universität zu Bonn (1983)
6. Barton, P.: The modelling and simulation of combined discrete/continuous processes. Ph.D. thesis, Department of Chemical Engineering, Imperial College of Science, Technology and Medicine, London (1992)
7. Baumrucker, B., Biegler, L.: MPEC strategies for optimization of a class of hybrid dynamic systems. J. Process Control **19**(8), 1248–1256 (2009)

8. Belotti, P., Kirches, C., Leyffer, S., Linderoth, J., Luedtke, J., Mahajan, A.: Mixed-integer non-linear optimization. In: Iserles, A. (ed.) Acta Numerica 2013, vol. 22. Cambridge University Press, Cambridge (2013). www.optimization-online.org/DB_HTML/2012/12/3698.html
9. Betts, J.: Practical Methods for Optimal Control Using Nonlinear Programming. SIAM, Philadelphia (2001)
10. Biegler, L.: Solution of dynamic optimization problems by successive quadratic programming and orthogonal collocation. Comput. Chem. Eng. **8**, 243–248 (1984)
11. Biegler, L.: Nonlinear Programming: Concepts, Algorithms, and Applications to Chemical Processes. Series on Optimization. SIAM, Philadelphia (2010)
12. Bock, H., Longman, R.: Computation of optimal controls on disjoint control sets for minimum energy subway operation. Adv. Astronaut. Sci. **50**, 949–972 (1985)
13. 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)
14. Burgschweiger, J., Gnädig, B., Steinbach, M.: Nonlinear programming techniques for operative planning in large drinking water networks. Open Appl. Math. J. **3**, 1–16 (2009)
15. Ceria, S., Soares, J.: Convex programming for disjunctive optimization. Math. Program. **86**, 595–614 (1999)
16. Facchinei, F., Jiang, H., Qi, L.: A smoothing method for mathematical programs with equilibrium constraints. Math. Program. **85**, 107–134 (1999)
17. Fang, H., Leyffer, S., Munson, T.: A pivoting algorithm for linear programs with complementarity constraints. Optim. Methods Softw. **87**, 89–114 (2012)
18. Ferris, M., Kanzow, C.: Complementarity and related problems: a survey (1998)
19. Fischer, A.: A special Newton-type optimization method. Optimization **24**, 269–284 (1992)
20. Fletcher, R., de la Maza, E.S.: Nonlinear programming and nonsmooth optimization by successive linear programming. Math. Program. **43**(3), 235–256 (1989)
21. Fletcher, R., Leyffer, S.: Solving mathematical programs with complementarity constraints as nonlinear programs. Optim. Methods Softw. **19**(1), 15–40 (2004)
22. Frangioni, A., Gentile, C.: Perspective cuts for a class of convex 0–1 mixed integer programs. Math. Program., Ser. A **106**, 225–236 (2006)
23. Fügenschuh, A., Herty, M., Klar, A., Martin, A.: Combinatorial and continuous models for the optimization of traffic flows on networks. SIAM J. Optim. **16**(4), 1155–1176 (2006)
24. Fukushima, M., Qi, L. (eds.): Reformulation: Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods. Kluwer Academic, Dordrecht (1999)
25. Gerdts, M.: Solving mixed-integer optimal control problems by Branch&Bound: a case study from automobile test-driving with gear shift. Optim. Control Appl. Methods **26**, 1–18 (2005)
26. Gerdts, M.: A variable time transformation method for mixed-integer optimal control problems. Optim. Control Appl. Methods **27**(3), 169–182 (2006)
27. Gerdts, M., Sager, S.: Mixed-integer DAE optimal control problems: necessary conditions and bounds. In: Biegler, L., Campbell, S., Mehrmann, V. (eds.) Control and Optimization with Differential-Algebraic Constraints, pp. 189–212. SIAM, Philadelphia (2012)
28. Gill, P., Murray, W., Saunders, M.: SNOPT: an SQP algorithm for large-scale constrained optimization. SIAM J. Optim. **12**, 979–1006 (2002)
29. Göttlich, S., Herty, M., Kirchner, C., Klar, A.: Optimal control for continuous supply network models. Netw. Heterog. Media **1**(4), 675–688 (2007)
30. Gräber, M., Kirches, C., Bock, H., Schlöder, J., Tegethoff, W., Köhler, J.: Determining the optimum cyclic operation of adsorption chillers by a direct method for periodic optimal control. Int. J. Refrig. **34**(4), 902–913 (2011)
31. Grossmann, I.: Review of nonlinear mixed-integer and disjunctive programming techniques. Optim. Eng. **3**, 227–252 (2002)
32. Gugat, M., Herty, M., Klar, A., Leugering, G.: Optimal control for traffic flow networks. J. Optim. Theory Appl. **126**(3), 589–616 (2005)
33. Günlük, O., Linderoth, J.: Perspective reformulations of mixed integer nonlinear programs with indicator variables. Math. Program. **124**(1–2), 183–205 (2010)

34. Hante, F., Sager, S.: Relaxation methods for mixed-integer optimal control of partial differential equations. Comput. Optim. Appl. **55**(1), 197–225 (2013)
35. Hellström, E., Ivarsson, M., Aslund, J., Nielsen, L.: Look-ahead control for heavy trucks to minimize trip time and fuel consumption. Control Eng. Pract. **17**, 245–254 (2009)
36. Hoheisel, T.: Mathematical programs with vanishing constraints. Ph.D. thesis, Julius-Maximilians-Universität Würzburg (2009)
37. Jung, M.N., Reinelt, G., Sager, S.: The Lagrangian relaxation for the combinatorial integral approximation problem. Optim. Methods Softw. (2012, submitted). www.optimization-online.org/DB_HTML/2012/02/3354.html
38. Kawajiri, Y., Biegler, L.: A nonlinear programming superstructure for optimal dynamic operations of simulated moving bed processes. Ind. Eng. Chem. Res. **45**(25), 8503–8513 (2006)
39. Kirches, C.: Fast Numerical Methods for Mixed-Integer Nonlinear Model-Predictive Control. Advances in Numerical Mathematics. Springer Vieweg, Wiesbaden (2011)
40. Kirches, C., Sager, S., Bock, H., Schlöder, J.: Time-optimal control of automobile test drives with gear shifts. Optim. Control Appl. Methods **31**(2), 137–153 (2010)
41. Kirches, C., Bock, H., Schlöder, J., Sager, S.: Block structured quadratic programming for the direct multiple shooting method for optimal control. Optim. Methods Softw. **26**(2), 239–257 (2011)
42. Kirches, C., Bock, H., Schlöder, J., Sager, S.: A factorization with update procedures for a KKT matrix arising in direct optimal control. Math. Program. Comput. **3**(4), 319–348 (2011)
43. Kirches, C., Wirsching, L., Bock, H., Schlöder, J.: Efficient direct multiple shooting for nonlinear model predictive control on long horizons. J. Process Control **22**(3), 540–550 (2012)
44. Leineweber, D., Bauer, I., Schäfer, A., Bock, H., Schlöder, J.: An efficient multiple shooting based reduced SQP strategy for large-scale dynamic process optimization (Parts I and II). Comput. Chem. Eng. **27**, 157–174 (2003)
45. Leyffer, S.: Complementarity constraints as nonlinear equations: theory and numerical experience. In: Optimization with Multivalued Mappings: Theory, Applications, and Algorithms, pp. 169–208. Springer, Berlin (2006)
46. Leyffer, S., López-Calva, G., Nocedal, J.: Interior methods for mathematical programs with complementarity constraints. SIAM J. Optim. **17**(1), 52–77 (2006)
47. Leyffer, S., Munson, T.: A globally convergent filter method for MPECs. Preprint ANL/MCS-P1457-0907, Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, IL, USA (2007)
48. Logist, F., Sager, S., Kirches, C., van Impe, J.: Efficient multiple objective optimal control of dynamic systems with integer controls. J. Process Control **20**(7), 810–822 (2010)
49. Martin, A., Möller, M., Moritz, S.: Mixed integer models for the stationary case of gas network optimization. Math. Program. **105**, 563–582 (2006)
50. Oldenburg, J., Marquardt, W.: Disjunctive modeling for optimal control of hybrid systems. Comput. Chem. Eng. **32**(10), 2346–2364 (2008)
51. Prata, A., Oldenburg, J., Kroll, A., Marquardt, W.: Integrated scheduling and dynamic optimization of grade transitions for a continuous polymerization reactor. Comput. Chem. Eng. **32**, 463–476 (2008)
52. Raghunathan, A., Biegler, L.: Mathematical programs with equilibrium constraints (MPECs) in process engineering. Comput. Chem. Eng. **27**, 1381–1392 (2003)
53. Raghunathan, A., Diaz, M., Biegler, L.: An MPEC formulation for dynamic optimization of distillation operations. Comput. Chem. Eng. **28**, 2037–2052 (2004)
54. Ralph, D., Wright, S.J.: Some properties of regularization and penalization schemes for MPECs. Optim. Methods Softw. **19**, 527–556 (2004)
55. Sager, S.: MIOCP benchmark site. mintoc.de
56. Sager, S.: Numerical Methods for Mixed-Integer Optimal Control Problems. Der Andere Verlag, Tönning (2005)
57. Sager, S.: Reformulations and algorithms for the optimization of switching decisions in nonlinear optimal control. J. Process Control **19**(8), 1238–1247 (2009)

58. Sager, S.: On the integration of optimization approaches for mixed-integer nonlinear optimal control. Habilitation, University of Heidelberg (2011)
59. Sager, S.: A benchmark library of mixed-integer optimal control problems. In: Lee, J., Leyffer, S. (eds.) Mixed Integer Nonlinear Programming, pp. 631–670. Springer, Berlin (2012)
60. Sager, S., Reinelt, G., Bock, H.: Direct methods with maximal lower bound for mixed-integer optimal control problems. Math. Program. **118**(1), 109–149 (2009)
61. Sager, S., Bock, H., Diehl, M.: The integer approximation error in mixed-integer optimal control. Math. Program., Ser. A **133**(1–2), 1–23 (2012)
62. Scholtes, S.: Convergence properties of a regularization scheme for mathematical programs with complementarity constraints. SIAM J. Optim. **11**, 918–936 (2001)
63. Scholtes, S.: Nonconvex structures in nonlinear programming. Oper. Res. **52**(3), 368–383 (2004)
64. Sherali, H.: RLT: a unified approach for discrete and continuous nonconvex optimization. Ann. Oper. Res. **149**, 185–193 (2007)
65. Sonntag, C., Stursberg, O., Engell, S.: Dynamic optimization of an industrial evaporator using graph search with embedded nonlinear programming. In: Proceedings of the 2nd IFAC Conference on Analysis and Design of Hybrid Systems (ADHS), pp. 211–216 (2006)
66. Stein, O., Oldenburg, J., Marquardt, W.: Continuous reformulations of discrete-continuous optimization problems. Comput. Chem. Eng. **28**(10), 3672–3684 (2004)
67. Stubbs, R., Mehrotra, S.: Generating convex polynomial inequalities for mixed 0–1 programs. J. Glob. Optim. **24**, 311–332 (2002)
68. Terwen, S., Back, M., Krebs, V.: Predictive powertrain control for heavy duty trucks. In: Proceedings of IFAC Symposium in Advances in Automotive Control, Salerno, Italy, pp. 451–457 (2004)
69. 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)