Towards Simulation Based Mixed-Integer Optimization with Differential Equations

Martin Gugat[1], Günter Leugering[1], Alexander Martin[2], Martin Schmidt[2,3], Mathias Sirvent[2], David Wintergerst[1]

Abstract. We propose a decomposition based method for solving mixed-integer nonlinear optimization problems with "black-box" nonlinearities, where the latter, e.g., may arise due to differential equations or expensive simulation runs. The method alternatingly solves a mixed-integer linear master problem and a separation problem for iteratively refining the mixed-integer linear relaxation of the nonlinear equalities. The latter yield nonconvex feasible sets for the optimization model but we have to restrict ourselves to convex and monotone constraint functions. Under these assumptions, we prove that our algorithm finitely terminates with a global optimal solution of the mixed-integer nonlinear problem. Additionally, we show the applicability of our approach for three applications from optimal control with integer variables, from the field of pressurized flows in pipes with elastic walls, and from steady-state gas transport. For the latter we also present promising numerical results of our method applied to real-world instances that particularly show the effectiveness of our method for problems defined on networks.

1. Introduction

Mixed-integer nonlinear models (MINLPs) form an important class of optimization problems because they combine the possibility of modeling nonlinear aspects and discrete decisions. However, since MINLPs allow for both nonlinearities and discrete variables they also inherit the theoretical and practical difficulties of nonlinear and mixed-integer optimization; see, e.g., [1] for a recent survey of MINLPs. This is the reason why general-purpose MINLP algorithms often have their limits when it comes to solving large-scale instances from real-world applications. Thus, typically problem tailored approaches must be developed to solve application-specific subclasses of MINLPs.

One main branch of MINLP techniques tries to rely on mixed-integer linear (MIP) solvers as working horse. These approaches approximate or relax the nonlinearities by (piecewise) affine-linear constraints. Moreover, in certain cases the resulting linearization error can be controlled such that multiple MIPs are solved with adaptively chosen linearization quality, finally leading to a global MIP solution not violating the original nonlinearities by more than a prescribed tolerance.

All these approaches suffer from an inherent drawback, namely that they require to re-model the original nonlinearities to obtain a mixed-integer linear surrogate model. This re-modeling is typically done by the modeler before the resulting MIP is solved. Furthermore, there are many situations in which such a re-modeling is not possible because the original nonlinearity is not known explicitly. Typical examples are nonlinearities described by differential equations for which no analytic solutions are known or nonlinearities that are the result of expensive simulation runs.

Date: September 25, 2017.

2010 *Mathematics Subject Classification.* 65K05, 90-08, 90B10, 90C11, 90C35, 90C90.

Key words and phrases. Mixed-Integer Optimization, Simulation Based Optimization, Optimization with Differential Equations, Decomposition Method, Gas Transport Networks.

In our contribution, we propose a decomposition based method for solving mixed-integer nonlinear problems with "black-box" nonlinearities. The term "black-box" is meant with respect to (w.r.t.) the actual optimization problem that is solved within our solution method. Here, the "black-box" nonlinearity is not explicitly part of the model but is only evaluated in a separate simulation step. The main assumption of the algorithm is that the addressed nonlinear equalities are convex and monotone functions. Although the resulting feasible set is nonconvex it is possible to construct a mixed-integer linear relaxation of the nonlinear equality using a mix of separating hyperplanes and piecewise linear approximations. These hyperplanes and piecewise linearizations are obtained using the simulation step—i.e., by solving, e.g., an initial value problem for a differential equation—and are then used to refine the linear relaxation of the original MINLP feasible set. By doing so, we construct a MIP surrogate model for the original MINLP without explicitly re-modeling the nonlinearity.

The contribution of this paper is threefold. The first goal is to formally state the problem class under consideration and the algorithm. Second, we prove correctness of our method. In the context of the presented algorithm, this means that it terminates after a finite number of iterations with a globally optimal point (w.r.t. prescribed tolerances) or with an indication of infeasibility. The third goal is to illustrate the applicability of our modeling approach by three case studies from the fields of stationary gas transport, of pressurized flows in pipes with elastic walls, and of optimal control with integer variables. In all these three cases, the nonlinearity arises due to the presence of certain differential equations and thus leads to optimization problems that are typically hard to solve for state-of-the-art general-purpose MINLP solvers. We show that in all cases, our analytical assumptions mentioned above are satisfied such that our tailored method can be applied. For the first case we also present detailed computational results of our algorithm applied to the Greek gas transport network that particularly show the effectiveness of our method for problems defined on networks.

This paper is organized as follows. In Section 2 we define the class of optimization problems we consider, discuss a tailored decomposition of these problems, and finally present the decomposition based solution algorithm. Moreover, we relate our solution approach to other methods from the literature. In Section 3 we then prove that the algorithm always terminates with a globally optimal solution or an indication of infeasibility. Section 4 presents some exemplary applications that can be tackled with our method and Section 5 presents numerical results that show the applicability of our approach. The paper ends with a summary and some notes on future work in Section 6.

2. Problem Statement and Algorithm

In this section we present the type of problem we address in the following, state a tailored decomposition, and finally the decomposition based solution algorithm.

The problems that we consider are discrete-continuous optimization models. Let $\mathcal{C}$ and $\mathcal{I}$ be finite index sets of continuous and integer variables, i.e., the entire variable vector reads $x := (x_{\mathcal{C}}^\top, x_{\mathcal{I}}^\top)^\top \in \mathbb{R}^{|\mathcal{C}|} \times \mathbb{Z}^{|\mathcal{I}|}$ and is assumed to be constrained by bounds $\underline{x} \leq x \leq \bar{x}$. All non-trivial linear constraints are denoted in LP form $Ax \geq b$ with $A \in \mathbb{R}^{m \times n}, n = n_{\mathcal{C}} + n_{\mathcal{I}} = |\mathcal{C}| + |\mathcal{I}|$, and $b \in \mathbb{R}^m$. We assume that the objective function is linear with coefficients $c \in \mathbb{R}^n$. For the remaining constraints we choose a rather abstract setting. Let $x_d = (x_{d_1}, x_{d_2})$ be a pair of variables for all variable index pairs $d = (d_1, d_2) \in \mathcal{D}$ with $d_1, d_2 \in \{1, \ldots, n_{\mathcal{C}}\}$ and a finite set $\mathcal{D}$. Moreover, let (x_{d_1}, x_{d_2}) be coupled by a function $f_d : \mathbb{R} \to \mathbb{R}$, i.e., $x_{d_2} = f_d(x_{d_1})$ holds for all $d \in \mathcal{D}$. For instance, think of $\mathcal{D}$ as the set of node pairs (u, v) associated to

arcs $a = (u, v)$ of a graph. In this situation, x_{d_1} and x_{d_2} are node variables that are coupled by the arc-related function f_d. Throughout the paper, we make the following assumption:

Assumption 1. *The functions $f_d, d \in \mathcal{D}$, are strictly monotonic, strictly concave or convex, and differentiable with a bounded first derivative.*

Moreover, we note that our method does not require explicit knowledge of the nonlinear functions f_d and that the assumption of differentiability might also be dropped. For the latter case, f_d is differentiable almost everywhere and we can use sub- or supergradients at the points of non-differentiability. We also remark that all of our results hold for all four combinations of strictly increasing or decreasing and strictly concave or convex functions. For the ease of presentation, we fix the setting to strictly increasing and strictly concave functions in the following. Finally consider the case in which it is possible to construct DC decompositions $f = \varphi - \psi$ of general nonlinear equality constraints, where φ and ψ satisfy Assumption 1; see, e.g., [41]. In this case, a much more general class of problems can be tackled by our method.

With these notations, we can state the problem under consideration:

$$\min_{x} \quad c^\top x \tag{1a}$$

$$\text{s.t.} \quad \underline{x} \le x \le \bar{x}, \quad x_{\mathcal{C}} \in \mathbb{R}^{|\mathcal{C}|}, \quad x_{\mathcal{I}} \in \mathbb{Z}^{|\mathcal{I}|}, \quad Ax \ge b, \tag{1b}$$

$$x_{d_2} = f_d(x_{d_1}) \quad \text{for all } d \in \mathcal{D}. \tag{1c}$$

Note that Problem (1) is an MINLP equipped with "black-box" constraint functions f_d that yield a nonconvex feasible set. We now highlight some exemplary situations in which Assumption 1 is satisfied. Detailed real-world applications are discussed in Section 4.

Example 1. *Assume an initial value problem with an ordinary differential equation (ODE)*

$$y' = g(x, y(x)), \quad y(0) = y_0, \quad x \in [0, L],$$

is given and denote the solution by $y = y(x; y_0)$. If the function $f : \mathbb{R} \to \mathbb{R}$ that maps initial values onto the solutions of the initial value problem, i.e., $f(y_0) = y(L; y_0)$, satisfies Assumption 1, this initial value problem solution mapping is a possible function f that can be used in Problem (1).

Example 2. *Sufficient conditions to ensure the satisfaction of Assumption 1 can be found if the functions f_d are given as the solutions of a separable ordinary differential equation with initial value $y_0 \in \mathbb{R}$,*

$$y' = g(y)h(x), \quad y(0) = y_0, \tag{2}$$

with $g \in \mathcal{C}^1(\mathbb{R}, \mathbb{R})$ and $h \in \mathcal{C}^0(\mathbb{R}, \mathbb{R})$. Assume that $g(y) \neq 0$ in some interval $[\underline{y}, \bar{y}]$. Note that the solution is constant if $g(y) = 0$ holds. Application of separation of variables yields that the solution of (2) *satisfies*

$$\int_{y_0}^{y} \frac{1}{g(z)} \, \mathrm{d}z = \int_0^x h(s) \, \mathrm{d}s.$$

Defining the primitive of $1/g$ as $G(y) = \int_0^y 1/g(z) \, \mathrm{d}z$ yields

$$G(y) - G(y_0) = \int_0^x h(s) \, \mathrm{d}s. \tag{3}$$

For $y \in [\underline{y}, \bar{y}]$, the function G is strictly monotonic as $g(y) \neq 0$ for $y \in [\underline{y}, \bar{y}]$. Thus, its inverse exists and the solution of (2) *is given by*

$$y(x) = G^{-1}\left(G(y_0) + \int_0^x h(s) \, \mathrm{d}s\right).$$

The derivative of the solution w.r.t. the initial value can be obtained by differentiating (3)*:*

$$G'(y)\partial_{y_0} y - G'(y_0) = 0$$
$$\iff \partial_{y_0} y = \frac{G'(y_0)}{G'(y)} = \frac{g(y)}{g(y_0)} \neq 0 \quad \text{for } y \in [\underline{y}, \bar{y}].$$

The second derivative fulfills

$$\partial^2_{y_0} y = \frac{g(y)}{g(y_0)^2}(g'(y) - g'(y_0)).$$

Consequently, Assumption 1 with $y_0 = x_{d_1}$ *and* $f_d(x_{d_1}) = y(y_0)$ *is fulfilled if* $g(y) \neq 0$ *on* $[\underline{y}, \underline{y}]$*, which ensures monotonicity, and* g *is either strictly convex or strictly concave.*

Example 3. *A more abstract setting than the case of separable ODEs has been discussed in [28], where it is shown ([28, Theorem 2.4]) that the dependence of a solution of the initial value problem*

$$y' = g(x, y), \quad y(0) = y_0$$

is convex with respect to the initial value, if

- *the function* g *is continuous, bounded, and the derivative* g_y *is bounded;*
- *for each* $x \in [0, L]$ *the function* $g(x, \cdot)$ *is quasimonotone nondecreasing; and*
- *for each* $x \in [0, L]$ *the function* $g(x, \cdot)$ *is convex.*

Our driving hypothesis about Problem (1) is that optimizing over the full constraint set is very expensive, whereas solely optimizing over Constraints (1b) as well as the sole evaluation of the functions f_d is also expensive but tractable in practice. This gives us a direct recipe for decomposing the problem: We split up (1) into a master problem and a subproblem by separating the "easy" mixed-integer linear part (1b) and the Constraints (1c) involving nonlinear black-box functions. Thus, the master problem is a MIP relaxation of the original problem and the subproblem consists of the handling of (1c). The main algorithmic idea is that these problems are solved alternately and that the solution of the subproblems are used to generate refined (piecewise) linear relaxations of the feasible set of Constraints (1c) by only exploiting local knowledge about the functions f_d that are delivered by the subproblem. By doing so, the master problem iteratively gets a better approximation of the original problem and finally yields an global optimal solution of Problem (1) or a proof of infeasibility.

We now give a formal description of the master problem and subproblem and start with the latter. To this end, let d be fixed and $y_d, d \in \mathcal{D}$, be part of a solution of the master problem. The subproblem determines the closest point w.r.t. the given solution y_d on the graph $(x_{d_1}, f_d(x_{d_1}))$, i.e., we solve the following minimum ℓ_2 distance problem over the bounded feasible set of Constraint (1c):

$$\chi(y_d) := \min_{x_d} \left\{ ||x_d - y_d||_2^2 \,:\, x_{d_2} = f_d(x_{d_1}),\ \underline{x}_d \leq x_d \leq \bar{x}_d \right\}. \tag{4}$$

We choose the ℓ_2 norm for the subproblem since it can be directly used in a continuous optimization problem like (4); the Euclidean norm is only squared to obtain a differentiable objective. The solutions of (4) are denoted by z_d. Moreover, we assume that solving the subproblems also delivers the values $f_d(z_{d_1})$ and $f'_d(z_{d_1})$.

The master problem is a relaxation of Problem (1), in which we neglect Constraint (1c) but take a piecewise linear relaxation of its feasible set

$$\mathrm{gr}(f_d) := \left\{ x_d \in \mathbb{R}^2 \,:\, x_{d_2} = f_d(x_{d_1}) \right\}$$

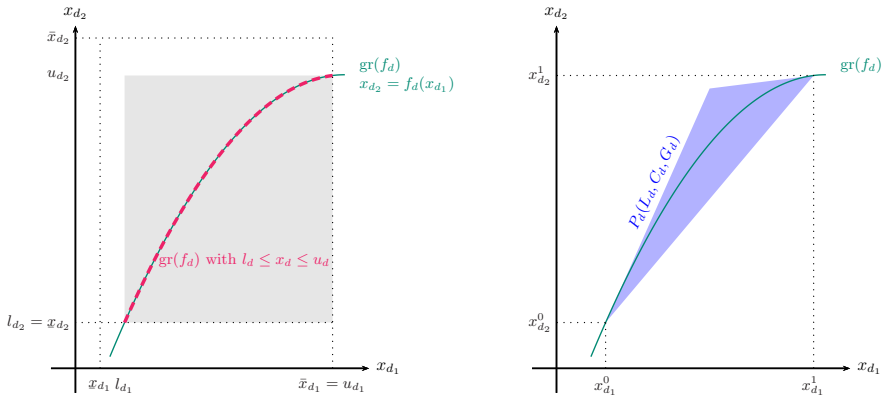

FIGURE 1. Left: Bound tightening for x_d; the gray area represents the tightened box. Right: Linear relaxation $P_d(L_d, C_d, G_d)$ with $L_d = \{x^0_{d_1}, x^1_{d_1}\}$, $C_d = \{x^0_{d_2}, x^1_{d_2}\}$, and $G_d = \{f'_d(x^0_{d_1}), f'_d(x^1_{d_1})\}$.

into account. Initially, Assumption 1 is combined with the bounds $\underline{x}_d \le x_d \le \bar{x}_d$ to strengthen the original formulation to $l_d \le x_d \le u_d$ with

$$\begin{aligned} l_{d_1} &= \max\left\{\underline{x}_{d_1}, f_d^{-1}(\underline{x}_{d_2})\right\}, & u_{d_1} &= \min\left\{\bar{x}_{d_1}, f_d^{-1}(\bar{x}_{d_2})\right\}, \\ l_{d_2} &= \max\left\{\underline{x}_{d_2}, f_d(\underline{x}_{d_1})\right\}, & u_{d_2} &= \min\left\{\bar{x}_{d_2}, f_d(\bar{x}_{d_1})\right\}, \end{aligned}$$

see Figure 1 (left) for an illustration. The piecewise linear relaxation of $\mathrm{gr}(f_d) \cap [l_d, u_d]$ is obtained by a combination of the incremental method [32] and classical outer approximations [10, 12]. To this end, we define the sets L_d, C_d, and G_d containing linearization points, corresponding function values, and corresponding derivatives of f_d, respectively. That is, we have

$$\begin{aligned} L_d &:= \left\{l_{d_1} = x^0_{d_1}, x^1_{d_1}, \dots, x^{r-1}_{d_1}, x^r_{d_1} = u_{d_1}\right\}, \\ C_d &:= \left\{l_{d_2} = x^0_{d_2}, x^1_{d_2}, \dots, x^{r-1}_{d_2}, x^r_{d_2} = u_{d_2}\right\}, \\ G_d &:= \left\{f'_d(x^0_{d_1}), f'_d(x^1_{d_1}), \dots, f'_d(x^r_{d_1})\right\}, \end{aligned}$$

with $x^i_{d_1} < x^{i+1}_{d_1}$ for all $i \in [r-1] := \{0, \dots, r-1\}$. As f_d is strictly increasing and strictly concave, $x^i_{d_2} < x^{i+1}_{d_2}$ and $f'_d(x^i_{d_1}) > f'_d(x^{i+1}_{d_1})$ hold for all $i \in [r-1]$ as well. With this notation, the master problem reads

$$\begin{aligned}
\min_x \quad & c^\top x && && \text{(5a)}\\
\text{s.t.} \quad & Ax \ge b, \quad x_{\mathcal{C}} \in \mathbb{R}^{|\mathcal{C}|}, \quad x_{\mathcal{I}} \in \mathbb{Z}^{|\mathcal{I}|}, && && \text{(5b)}\\
& \underline{x} \le x \le \bar{x}, && && \text{(5c)}\\
& x_{d_2} \le x^i_{d_2} + f'_d(x^i_{d_1})(x_{d_1} - x^i_{d_1}) && \text{for all } d \in \mathcal{D}, i \in [r], && \text{(5d)}\\
& x_{d_1} = x^0_{d_1} + \sum_{i=1}^{r} (x^i_{d_1} - x^{i-1}_{d_1})\delta^i_d && \text{for all } d \in \mathcal{D}, && \text{(5e)}\\
& x_{d_2} \ge x^0_{d_2} + \sum_{i=1}^{r} (x^i_{d_2} - x^{i-1}_{d_2})\delta^i_d && \text{for all } d \in \mathcal{D}, && \text{(5f)}\\
& \delta^i_d \ge w^i_d \ge \delta^{i+1}_d && \text{for all } d \in \mathcal{D}, i \in [r-1] \setminus \{0\}, && \text{(5g)}\\
& \delta^i_d \in [0,1] && \text{for all } d \in \mathcal{D}, i \in [r] \setminus \{0\}, && \text{(5h)}\\
& w^i_d \in \{0,1\} && \text{for all } d \in \mathcal{D}, i \in [r-1] \setminus \{0\}. && \text{(5i)}
\end{aligned}$$

The linear overestimators (5d) are valid outer approximation cuts and Constraints (5e)–(5g) are piecewise linear underestimators by Assumption 1. The relaxed version of the incremental method uses the auxiliary variables in (5h) and (5i). The variables and constraints in (5d)–(5i) thus yield a piecewise linear relaxation of the feasible set of the original constraints (1c). We denote the feasible set of (5c)–(5i) by $P_d = P_d(L_d, C_d, G_d)$ for every $d \in \mathcal{D}$; see Figure 1 (right).

We are now able to formally state the decomposition algorithm for solving Problem (1); see Algorithm (1). In the following section we then prove that this

Algorithm 1 Decomposition Method

Input: Problem (1) and $\varepsilon > 0$.

Output: If there is an ε-feasible point for Problem (1), the algorithm returns an globally optimal ε-feasible solution y. Otherwise, it returns an indication of infeasibility.

Set $L_d := \{l_{d_1}, u_{d_1}\}$, $C_d := \{l_{d_2}, u_{d_2}\}$, and $G_d := \{f'(l_{d_1}), f'(u_{d_1})\}$ for all $d \in \mathcal{D}$.
for $k = 0, 1, 2, \ldots$ **do**
  Solve master problem (5) with L_d, C_d, and G_d.
  if master problem (5) is infeasible **return** "Problem (1) is infeasible".
  Denote the solution of master problem (5) by y^k.
  Solve subproblem (4) for all $d \in \mathcal{D}$ yielding solutions z_d^k and objective values $\chi(y_d^k)$.
  if $\chi(y_d^k) \leq \varepsilon$ for all $d \in \mathcal{D}$ **then**
    return globally optimal ε-feasible solution y^k
  else
    Set $L_d \leftarrow L_d \cup \{z_{d_1}^k\}, C_d \leftarrow C_d \cup \{z_{d_2}^k\}$, and $G_d \leftarrow G_d \cup \{f'(z_{d_1}^k)\}$ for all $d \in \mathcal{D}$ with $\chi(y_d^k) > \varepsilon$.
  end if
end for

algorithm is correct in the sense that it terminates at globally optimal ε-feasible points of Problem (1):

Definition 1 (ε-feasibility)**.** *A solution of the master problem* (5) *is called* ε-feasible *if* $\chi(y_d) \leq \varepsilon$ *for all* $d \in \mathcal{D}$*. Otherwise it is called* ε*-infeasible.*

The main ideas of the proof are to show that the subproblem solutions yield an improved piecewise linear relaxation of the original problem that cut off the last master problem solution and that a finite number of these improvements suffices to achieve an ε-feasible point. Thus, the solution of the subproblem can also be seen as a separation oracle that yields new cutting planes for the last infeasible (master problem) solution.

Before we prove the theoretical properties of Algorithm 1 in the next section we first discuss the relation to other algorithms of mixed-integer nonlinear optimization. For a much broader overview over MINLPs see the recent survey [1] and the references therein.

As a result of Algorithm 1 the nonlinearities of the MINLP are replaced with (piecewise) linear relaxations that are iteratively tightened by the algorithm. This idea is related to the MIP based solution techniques for MINLPs discussed in [14, 17] that have also been applied to problems from the field of gas transport; see [16, 18, 34] and Section 4.3. The algorithms published in the given citations mainly differ in two facts. First, they rely on the fact that the nonlinear functions are given in closed form, and second, they construct the relaxations a-priori in dependence of some given tolerance and not on demand as in our case. An additional minor difference is the use of the ℓ_1 norm in [14, 17] whereas we use an ℓ_2 projection. To model the piecewise linear relaxations the incremental method [32] is used in both approaches. We refrain from reviewing the large literature on piecewise linear

approximation of nonlinear functions but refer the reader to the references in the hitherto cited papers.

Moreover, our method has similarities to Generalized Benders Decomposition [2, 19]; see [4] for a textbook version. However, it is known that the standard version of Generalized Benders Decomposition does not converge for nonconvex problems [37]—a problem that our algorithm does not suffer from.

At last we want to highlight that the class of problems that we consider is related to problems from the fields of mixed-integer optimal control (MIOCP) and mixed-integer dynamic optimization (MIDO), where systems governed by ODEs or PDEs are optimized over a control input; see, e.g., [1, Chap. 7], [36] as well as [7], and the references therein. Also related are approaches for mixed-integer optimization with differential equations based on linearization; see, e.g., [13, 15]. However, MIOCP addresses integers distributed in time or space, which is not considered in this paper.

Very recently, an approach for mixed-integer optimal control with semilinear elliptic PDEs and static integer controls has been proposed in [6]. This approach has some similarities with our approach since they also decouple the entire problem into a master and subproblem. The former addresses all mixed-integer aspects whereas the latter addresses the differential equation and delivers outer approximation cutting planes for the next master problem. However, the focus of [6] is more on the specific type of differential equation that yields convex feasible sets whereas our approach abstracts from specific differential equations but also considers nonconvex feasible sets from nonlinear equality constraints.

3. Convergence Results

In order to prove the correctness of Algorithm 1 we show that an ε-infeasible solution y^k of the kth master problem is cut off in the next iteration $k+1$ due to suitable extensions of the sets L_d, C_d, and G_d, which are augmented for all $d \in \mathcal{D}$ with $\chi(y_d^k) > \varepsilon$. This is proven in Theorem 1. The correctness is then stated in Theorem 2. Throughout the section we use the notations $\operatorname{epi}(f_d)$ and $\operatorname{hyp}(f_d)$ for the epi- and hypograph of the function as well as the notations $\operatorname{epi_S}(f_d)$ and $\operatorname{hyp_S}(f_d)$ for their strict versions.

We first prove an auxiliary result for which we need the following definition:

Definition 2. *Let $y_d \in P_d$. For $y_d \in \operatorname{epi}(f_d)$ we define*

$$\bar{R}(y_d) := \{x_d \in \operatorname{gr}(f_d) : x_{d_1} \geq y_{d_1}, x_{d_2} \leq y_{d_2}\}$$

and for $y_d \in \operatorname{hyp}(f_d)$ we define

$$\underline{R}(y_d) := \{x_d \in \operatorname{gr}(f_d) : x_{d_1} \leq y_{d_1}, x_{d_2} \geq y_{d_2}\}.$$

Lemma 1. *Let $y_d \in P_d(L_d, C_d, G_d)$ and let z_d be the solution of the subsequent subproblem (4). Then the following holds:*

(a) If there exists an index $j \in [r]$ with $(x_{d_1}^j, x_{d_2}^j) = (y_{d_1}, y_{d_2})$ then $z_d = y_d$ and $\chi(y_d) = 0$.

(b) Suppose that $\chi(y_d) > \varepsilon > 0$ holds. Then there exists an index $j \in [r-1]$ with

$$x_{d_1}^j < y_{d_1} < x_{d_1}^{j+1}, \quad x_{d_2}^j < y_{d_2} < x_{d_2}^{j+1}, \tag{6a}$$

$$x_{d_1}^j < z_{d_1} < x_{d_1}^{j+1}, \quad x_{d_2}^j < z_{d_2} < x_{d_2}^{j+1}. \tag{6b}$$

Moreover, $z_d \in \bar{R}(y_d)$ if $y_d \in \operatorname{epi_S}(f_d)$ or $z_d \in \underline{R}(y_d)$ if $y_d \in \operatorname{hyp_S}(f_d)$.

Proof. (a) Since $y_d \in P_d(L_d, C_d, G_d)$ holds it also fulfills the Constraints (5d)–(5i). As the linearization points $(x_{d_1}^i, x_{d_2}^i)$ are chosen such that $x_{d_2}^i = f_d(x_{d_1}^i)$ holds for

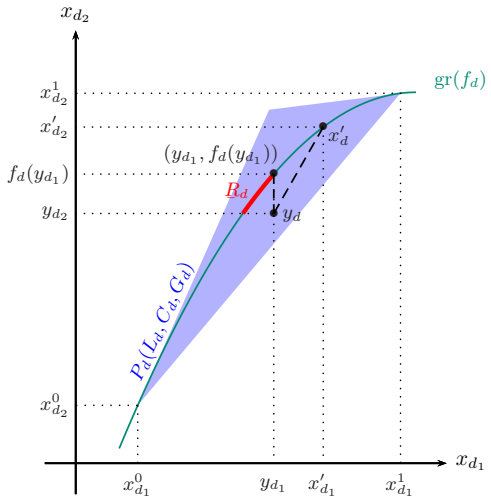

FIGURE 2. Illustration of the proof of (6b) of Lemma 1 for the case $y_d \in \mathrm{hyp_S}(f_d)$ and $L_d = \{x^0_{d_1}, x^1_{d_1}\}$, $C_d = \{x^0_{d_2}, x^1_{d_2}\}$, $G_d = \{f'_d(x^0_{d_1}), f'_d(x^1_{d_1})\}$

all $i \in [r]$, we have $f(y_{d_1}) = f(x^j_{d_1}) = x^j_{d_2} = y_{d_2}$. Thus, y_d is feasible for (4) with optimal value $\chi(y_d) = 0$ and $z_d = y_d$.

(b) Because $\chi(y_d) > \varepsilon > 0$ and (a), there exists an index $j \in [r-1]$ with

$$x^j_{d_1} < y_{d_1} < x^{j+1}_{d_1}.$$

Together with (5e) this implies $\delta^1_d = \cdots = \delta^j_d = 1$, $\delta^{j+1}_d \in (0,1)$, and $\delta^{j+2}_d = \cdots = \delta^r_d = 0$. Thus, (5f) and monotonicity of f_d yield

$$x^j_{d_2} < y_{d_2} < x^{j+1}_{d_2}.$$

We only prove the case $y_{d_2} \in \mathrm{hyp_S}(f_d)$; see Figure 2 for an illustration. The case $y_{d_2} \in \mathrm{epi_S}(f_d)$ can be shown analogously. We first prove $z_d \in \underline{R}(y_d)$. Let $x'_d \in \mathrm{gr}(f_d) \setminus \underline{R}(y_d)$ with $y_{d_1} < x'_{d_1}$. Strict monotonicity of f_d and $y_{d_2} < f_d(y_{d_1})$ imply $y_{d_2} < f_d(y_{d_1}) < f_d(x'_{d_1}) = x'_{d_2}$ and we obtain

$$||y_d - (y_{d_1}, f_d(y_{d_1}))||_2^2 < ||y_d - x'_d||_2^2.$$

This implies that the subproblem's objective at $(y_{d_1}, f_d(y_{d_1})) \in \underline{R}(y_d)$ is smaller than at $x'_d \notin \underline{R}(y_d)$. The same holds for every $x'_d \in \mathrm{gr}(f_d) \setminus \underline{R}(y_d)$ with $x'_{d_1} < f_d^{-1}(y_{d_2})$. This shows $z_d \in \underline{R}(y_d)$, which then yields

$$z_{d_1} \leq y_{d_1} \quad \text{and} \quad y_{d_2} \leq z_{d_2}.$$

From (6a) now follows

$$z_{d_1} < x^{j+1}_{d_1} \quad \text{and} \quad x^j_{d_2} < z_{d_2}$$

and strict monotonicity of f_d finally implies

$$x^j_{d_1} < z_{d_1} < x^{j+1}_{d_1} \quad \text{and} \quad x^j_{d_2} < z_{d_2} < x^{j+1}_{d_2}. \qquad \square$$

Part (a) of Lemma 1 states that $\chi(y_d) = 0$ holds if the solution of the master problem is a linearization point in C_d. In this case, we do not have to cut off the solution of the master problem. Thus, we only consider the case $\chi(y_d) > \varepsilon > 0$ in the following and exploit the situation given by Lemma 1(b).

Proposition 1. *Let $y_d \in P_d$ be a solution of the master problem* (5) *with $y_d \in \mathrm{epi}(f_d)$ and let z_d be the solution of the subsequent subproblem* (4) *with $\chi(y_d) > \varepsilon > 0$. Then z_d is the orthogonal projection of y_d onto $\mathrm{gr}(f_d)$, respectively on the tangent*

$$x_{d_2} = f_d(z_{d_1}) + f_d'(z_{d_1})(x_{d_1} - z_{d_1}).$$

We now prove that a solution y_d of the master problem with $\chi(y_d) > \varepsilon > 0$ is always cut off in the formulation of the subsequent master problems.

Theorem 1. *Let $y_d \in P_d(L_d, C_d, G_d)$ be a solution of the master problem* (5) *and let z_d be a solution of the subsequent subproblem* (4) *with $\chi(y_d) > \varepsilon > 0$. Then $y_d \notin P_d(L_d', C_d', G_d')$ with $L_d' = L_d \cup \{z_{d_1}\}$, $C_d' = C_d \cup \{z_{d_2}\}$, and $G_d' = G_d \cup \{f'(z_{d_1})\}$.*

Proof. We distinguish the two cases $y_d \in \mathrm{epi_S}(f_d)$ and $y_d \in \mathrm{hyp_S}(f_d)$ and start with the former. The outer approximation cut for f_d at z_d as given in Constraint (5d) can be written as $\bar{n}_d^\top x_d \leq \bar{\xi}_d$ with $\bar{n}_d = (-f_d'(z_{d_1}), 1)^\top$ and $\bar{\xi}_d = z_{d_2} - f_d'(z_{d_1}) z_{d_1}$, where $\bar{n}_d$ is the normal and $\bar{n}_d^\top z_d = \bar{\xi}_d$. Proposition 1 yields that z_d is the unique orthogonal projection of y_d onto $\mathrm{gr}(f_d)$. Thus, there is some $c_d \in \mathbb{R}_{>0}$ with $y_d - z_d = c_d \bar{n}_d$. Multiplication of the tangent with c_d yields $n_d^\top x_d \leq \xi_d$ with $n_d = y_d - z_d$ and $\xi_d = c_d \bar{\xi}_d$. Obviously, $\mathrm{hyp}(f_d)$ is strictly convex since f_d is strictly concave. Hence, we have $(z_d - y_d)^\top (z_h - z_d) \geq 0$ for all $z_h \in \mathrm{hyp}(f_d)$ because $y_d \notin \mathrm{hyp}(f_d)$. This implies

$$n_d^\top z_h \leq n_d^\top z_d = c_d \bar{n}_d^\top z_d = c_d \bar{\xi}_d = \xi_d.$$

Thus, $n_d^\top x_d \leq \xi_d$ is a valid cut for $\mathrm{gr}(f_d)$ as $\mathrm{gr}(f_d) \subseteq \mathrm{hyp}(f_d)$. On the other hand,

$$n_d^\top y_d = n_d^\top z_d + n_d^\top (y_d - z_d) = \xi_d + ||n_d||_2^2 > \xi_d$$

holds. Hence, $n_d^\top x_d \leq \xi_d$ separates y_d and we have shown $y_d \notin P_d(L_d', C_d', G_d')$. See Figure 3 (left) for an illustration.

Consider now the case $y_d \in \mathrm{hyp_S}(f_d)$ and define $P_d(L_d', C_d', G_d')$ by integrating z_d into the linearization. Lemma 1(b) yields

$$L_d' = \{x_{d_1}^0, \ldots, x_{d_1}^j, z_{d_1}, x_{d_1}^{j+1}, \ldots, x_{d_1}^{n+1}\}, \tag{7a}$$

$$C_d' = \{x_{d_2}^0, \ldots, x_{d_2}^j, z_{d_2}, x_{d_2}^{j+1}, \ldots, x_{d_2}^{n+1}\}, \tag{7b}$$

$$G_d' = \{f_d'(x_{d_1}^0), \ldots, f_d'(x_{d_1}^j), f_d'(z_{d_1}), f_d'(x_{d_1}^{j+1}), \ldots, f_d'(x_{d_1}^{n+1})\}. \tag{7c}$$

We define

$$W(y_d) := \{x_d \in \mathrm{gr}(f_d) : y_d \notin P_d(\hat{L}_d, \hat{C}_d, \hat{G}_d), x_d^j \leq x_d \leq x_d^{j+1}\},$$
$$\bar{W}(y_d) := \{x_d \in \mathrm{gr}(f_d) \setminus W(y_d) : x_d^j \leq x_d \leq x_d^{j+1}\},$$

with $\hat{L}_d := L_d \cup \{x_{d_1}\}, \hat{C}_d := C_d \cup \{x_{d_2}\}$, and $\hat{G}_d := G_d \cup \{f_d'(x_{d_1})\}$. The set $W(y_d)$ represents all solutions z_d of the subproblem (4), for which y_d is cut off in the next master problem; see Figure 3 (right) for an illustration. Thus, we have to show $z_d \in W(y_d)$. Since Lemma 1(b) ensures $z_d \in \underline{R}(y_d)$ it is sufficient to prove $\underline{R}(y_d) \subseteq W(y_d)$. Assume $z_d \notin W(y_d)$, i.e., $z_d \in \bar{W}(y_d)$. This implies $y_d \in P_d(L_d', C_d', G_d')$. We now consider the Constraints (5e) and (5f) with $\delta_d^i = 1$ for all $i \leq j$, $\delta_d^{j+1} \in [0, 1]$, and $\delta_d^i = 0$ for all $i \geq j + 2$;

$$y_{d_1} = x_{d_1}^j + (z_{d_1} - x_{d_1}^j)\delta_d^{j^*} + (x_{d_1}^{j+1} - z_{d_1})\delta_d^{j+1}, \tag{8}$$

$$y_{d_2} \geq x_{d_2}^j + (z_{d_2} - x_{d_2}^j)\delta_d^{j^*} + (x_{d_2}^{j+1} - z_{d_2})\delta_d^{j+1}. \tag{9}$$

Here, j^* is the new index of the incremental method after adding z_d to L_d', C_d', and G_d'; see (7). Finally, we apply a case analysis:

(i) $\delta_d^{j^*} = \delta_d^{j+1} = 0$. Equation (8) yields $y_{d_1} = x_{d_1}^j$, which contradicts (6a) of Lemma 1.

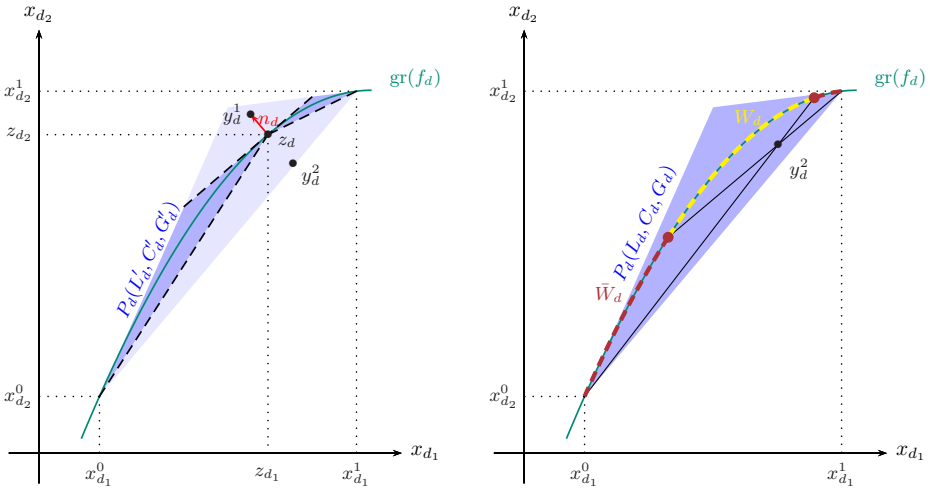

FIGURE 3. Illustration of the proof of Theorem 1; Left: first case $y^1_d \in \mathrm{epi_S}(f_d)$ and second case $y^2_d \in \mathrm{hyp_S}(f_d)$ with $L'_d = \{x^0_{d_1}, z_{d_1}, x^1_{d_1}\}$, $C'_d = \{x^0_{d_2}, z_{d_2}, x^1_{d_2}\}$, and $G'_d = \{f'_d(x^0_{d_1}), f'_d(z_{d_1}), f'_d(x^1_{d_1})\}$, Right: The proof of the second case uses the sets $W(y_d)$ and $\bar{W}(y_d)$.

(ii) $\delta^{j^*}_d \in (0,1), \delta^{j+1}_d = 0$. Thus, (8) yields $x^j_{d_1} < y_{d_1} < z_{d_1}$; a contradiction because $z_d \in \underline{R}(y_d)$ implies $z_{d_1} \leq y_{d_1}$.

(iii) $\delta^{j^*}_d = 1, \delta^{j+1}_d = 0$. Thus, (8) yields $y_{d_1} = z_{d_1}$ respectively $f_d(y_{d_1}) = z_{d_2}$ and (9) yields $y_{d_2} \geq z_{d_2}$. However, $y_d \in \mathrm{hyp_S}(f_d)$ implies $y_{d_2} < f_d(y_{d_1}) = z_{d_2}$; again a contradiction.

(iv) $\delta^{j^*}_d = 1, \delta^{j+1}_d \in (0,1)$. Thus, (8) yields $z_{d_1} < y_{d_1} < x^{j+1}_{d_1}$ and (9) yields $y_{d_2} \geq z_{d_2} + (x^{j+1}_{d_2} - z_{d_2})\delta^{j+1}_d > z_{d_2}$. Additionally, $z_d \in \underline{R}(y_d)$ implies $y_{d_2} \leq z_{d_2}$; a contradiction.

(v) $\delta^{j^*}_d = \delta^{j+1}_d = 1$. In this case, (8) yields $y_{d_1} = x^{j+1}_{d_1}$, which contradicts (6a).

This shows $z_d \in W(y_d)$ and thus $y_d \notin P_d(L'_d, C'_d, G'_d)$; see Figure 3 (left). □

We remark that the first case also follows from a suitable minimum principle; see, e.g., [31, Theorem 9.3.3]. See also [3] for an application of the principle in a closely related context.

Theorem 2. *Algorithm 1 terminates after a finite number of iterations at an globally optimal ε-feasible solution of* (1) *or with an indication that* (1) *is infeasible.*

Proof. We first consider a single $d \in \mathcal{D}$. Assume that the algorithm does not terminate after a finite number of iterations. That is, there exists a subsequence (indexed by ℓ) of the iterates with

$$\chi(y^\ell_d) > \varepsilon \quad \text{for all } \ell.$$

Thus, there exists a further subsequence indexed by m such that all m-iterates satisfy

$$y^m_d \in \mathrm{epi_S}(f_d) \quad \text{or} \quad y^m_d \in \mathrm{hyp_S}(f_d).$$

We show that both cannot happen and start with the first case.

Let y^m_d be the corresponding master problem solutions. Since all variables are bounded, this subsequence of iterates is bounded as well and we thus have a

convergent subsequence (y_d^μ). Thus, for every $\delta > 0$ we have

$$||y_d^\alpha - y_d^\beta||_2 < \delta \tag{10}$$

for all sufficiently large indices α and β of the μ-subsequence. On the other hand, all μ-iterates are excluded by an outer approximation cut from the feasible set of the master problem in iteration $m+1$; see Theorem 1. This cut excludes at least the ε-ball $B_\varepsilon(y_d^m)$, which yields

$$||y_d^\alpha - y_d^\beta||_2 > \varepsilon > 0$$

for all α, β. This contradicts (10).

Next, we consider the case $y_d^m \in \mathrm{hyp_S}(f_d)$ for all iterates in the subsequence indexed by m. Strict monotonicity of f_d implies that the rectangle spanned by the points y_d^m and z_d^m is cut off from the feasible set of the master problem in iteration $m+1$ by the Constraints (5e) and (5f). Let $\delta > \varepsilon > 0$ be the length of the diagonal of this rectangle and let a and b denote its sides. Since z_d^m is an orthogonal projection of y_d^m on f_d, the equality

$$\frac{a}{b} = f_d'(z_{d_1}^m)$$

holds. Using the Pythagorean theorem, a and b can be written depending on δ and $f_d'(z_{d_1}^m)$ and so can the area of the rectangle:

$$ab = \frac{\delta f_d'(z_{d_1}^m)}{\sqrt{1 + f_d'(z_{d_1}^m)^2}} \frac{\delta}{\sqrt{1 + f_d'(z_{d_1}^m)^2}} = \frac{\delta^2 f_d'(z_{d_1}^m)}{1 + f_d'(z_{d_1}^m)^2} > \delta^2\gamma > \varepsilon^2\gamma$$

for some $\gamma > 0$ since f_d' is both bounded above and bounded away from zero. However, the area of

$$\mathrm{hyp_S}(f_d) \cap [l_d, u_d]$$

is finite, yielding a contradiction. □

We close this section by some concluding remarks. First, if Problem (1) is infeasible, the master problem can still have ε-feasible solutions (see Definition 1) that then will be found by Algorithm 1 according to Theorem 2. Second, let x^* be the solution of Problem (1) and let $\hat{x}$ be the solution returned by Algorithm 1. Then, $c^\top \hat{x} \leq c^\top x^*$ holds because Algorithm 1 always considers a relaxation of the original problem (1). Third, we remark that we have not made any statements about the speed of convergence. It is possible to give worst-case iteration bounds based on the geometrical ideas of the proof of Theorem 2. However, pathological examples might exist that lead to arbitrary large worst-case estimations on the required number of iterations. See, e.g., Figure 4, where the area that is cut off in every iteration is very small and the subsequent master problem solutions are very close to each other. Finally, we remark that one may also think of using other distance measures like the ℓ_2 norm in the subproblem. We expect that it is also possible to prove a theorem analogue to Theorem 2 when using, e.g., the ℓ_1 norm. However, the presented proofs rely on the ℓ_2 norm and thus do not carry over directly.

4. Applications

The assumptions that are required for proving the correctness of our method are quite strong. However, there exist important applications that satisfy the assumptions and that can thus be solved by our algorithm. In this section we present three different examples where this is the case. First, we discuss a simple mixed-integer optimal control problem for a car fleet in Section 4.1. It is shown that the model can be reformulated such that it is of type (1) and that it satisfies Assumption 1. Afterward, we show that the system of differential equations for

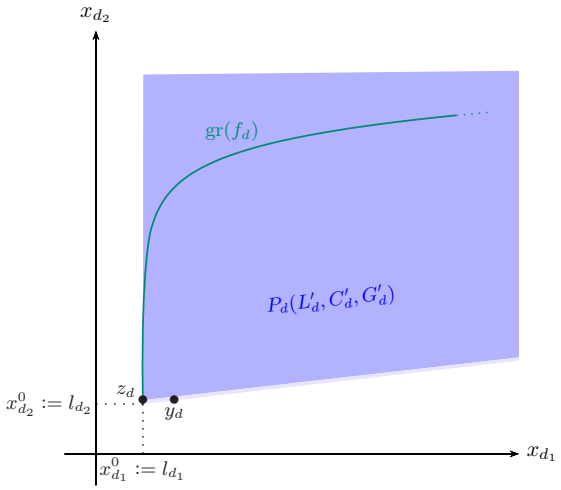

FIGURE 4. A pathological example with slow convergence

modeling pressurized flow in pipes with elastic walls also satisfies the claimed analytical properties in Section 4.2. Finally, we present a detailed mixed-integer nonlinear model of steady-state gas flow in pipeline networks in Section 4.3 and show that Assumption 1 is satisfied.

In order to illustrate the practical behavior of our method on real-world instances, we present numerical results of our decomposition method applied to the latter problem in the next section.

4.1. **Optimal Control of a Fleet of Rocket Cars.** The control of a rocket car is a classical example in the optimal control literature; see, e.g., Example 1 in [30]. Although the system dynamics for a single rocket car are linear, with a quadratic objective function (e.g., for measuring the control cost in the L^2 norm) this classical example gives rise to a nonlinear optimal control problem.

In order to illustrate how discrete and continuous control decisions are combined in the problems that we consider in this paper, we extend this problem to the problem to control a fleet of rocket cars. The cars have to be chosen in such a way that the overall control cost satisfies an inequality constraint and, at the same time, the objective function is minimized. In the objective function, the terminal time T appears at which the cars have reached the target state. The corresponding control cost (e.g., in \$) depends on T in a nonlinear way.

Let a fleet of $N \in \mathbb{N}$ rocket cars be given. For $i \in [N]$ and $t \in (0, T)$, the ith rocket car is modeled by the linear system

$$x_i(0) = 0, \qquad x_i'(0) = 0, \tag{11a}$$

$$x_i''(t) = d_i u_i(t), \quad t \in (0, T), \tag{11b}$$

$$x_i(T) = \tau, \qquad x_i'(T) = 0, \tag{11c}$$

where $d_i > 0$ is a real number that models the efficiency of the car; see, e.g., [24]. Here, the target state is described by a real number $\tau \neq 0$. Initially the cars start at zero. The units of the physical quantities are s for T and t, m for x_i and τ, $\mathrm{m\,s^{-1}}$ for x_i', and $\mathrm{m\,s^{-2}}$ for x_i'' and u_i. Moreover, in the optimal control problem we introduce the additional coefficient $\nu > 0$ (in \$/s) in order to obtain consistent units for all objective function terms.

The aim is to choose $M \leq N$ rocket cars from the fleet that can be steered to the target state in such a way that an overall fuel constraint is satisfied. To define this constraint, let some bound $\gamma > 0$ for the sum of the control costs be given. The objective function that is to be minimized is the sum of the time T when all M cars have arrived at the target and the costs $c_i > 0$ of using the different cars. Usually, the most efficient cars have the highest cost.

We consider the problem of time-optimal control of a selection of M out of N rocket cars with a constraint on the total fuel consumption

$$\begin{aligned}
\min \quad & \nu\, T + \sum_{i=1}^{N} c_i \sigma_i && \text{(12a)}\\
\text{s.t.} \quad & \sigma_i \in \{0,1\} \quad \text{for all } i \in [N], && \text{(12b)}\\
& \sum_{i=1}^{N} \sigma_i \|u_i\|^2_{L^2(0,T)} \leq \gamma, \quad \sum_{i=1}^{N} \sigma_i = M, && \text{(12c)}\\
& x_i(0) = 0, x_i'(0) = 0 \quad \text{for all } i \in [N], && \text{(12d)}\\
& x_i''(t) = \sigma_i d_i u_i(t) \quad \text{for all } i \in [N], t \in (0,T), && \text{(12e)}\\
& x_i(T) = \sigma_i \tau, x_i'(T) = 0 \quad \text{for all } i \in [N]. && \text{(12f)}
\end{aligned}$$

The variable vector contains $T \geq 0, u_i \in L^2(0,T)$, and $\sigma_i \in \{0,1\}$ with $i \in [N]$. Note that both Constraint (12c) and (12e) introduce nonlinearities to the problem. To rewrite this problem in the form (1), for $i \in [N]$ and $T \geq 0$ we define the strictly decreasing differentiable convex function

$$\begin{aligned}
f_i(T) = \inf \quad & \|u_i\|^2_{L^2(0,T)}\\
\text{s.t.} \quad & u_i \in L^2(0,T) \text{ and (11)}.
\end{aligned}$$

For an optimal fleet of cars, all fuel resources will be used to reach the target state as fast as possible. Therefore we may assume that for an optimal fleet of cars, the inequality constraint in (12c) is active. Thus, we can replace it by the corresponding equality constraint. To obtain a linear model we introduce auxiliary variables $z_i, y_i \in \mathbb{R}$ and additional linear inequality constraints. In fact, (12) is equivalent to

$$\begin{aligned}
\min \quad & \nu T + \sum_{i=1}^{N} c_i \sigma_i\\
\text{s.t.} \quad & \sigma_i \in \{0,\, 1\} \quad \text{for all } i \in [N],\\
& \sum_{i=1}^{N} y_i = \gamma, \quad \sum_{i=1}^{N} \sigma_i = M,\\
& z_i - \gamma(1-\sigma_i) \leq y_i \leq \gamma \sigma_i,\\
& 0 \leq y_i \leq z_i, \quad z_i = f_i(T).
\end{aligned}$$

with variables $T \geq 0, \sigma_i \in \{0,1\}$, and $z_i, y_i \in \mathbb{R}$ for all $i \in [N]$, where the auxiliary variables y_i are used to linearize the nonlinear products $\sigma_i z_i$. This model is exactly of type (1) and satisfies Assumption 1. Thus, it can be tackled by our proposed algorithm.

4.2. **Pressurized Flow in Pipes with Elastic Walls.** We now consider the following system of hyperbolic balance equations describing fluid flow in an extensible pipe system. Such systems are adequate for describing blood flow in arteries or flow in high pressure pipe systems, in particular under high temperature; see, e.g., [5]

for a recent survey. Especially for the latter case, the design of high pressure steam pipes in, e.g., power plants, requires the incorporation of discrete aspects; see [39].

The system under consideration can be described as follows. Let A_0 be the stress-free reference cross section, e.g., at the beginning, of the given straight pipe and $A(x,t)$ the cross section at the spatial point $x \in [0,L]$ at time $t \in [0,T]$. The averaged velocity of the fluid is denoted by $m(x,t)$, ρ denotes the fluid's density which is assumed to be constant for simplicity, ν is the viscosity, and $\alpha > 1$ is a parameter representing the Coriolis force. The system of equations consisting of the continuity and momentum equation reads

$$\begin{aligned}
&\frac{\partial}{\partial t}A + \frac{\partial}{\partial x}m = 0,\\
&\frac{\partial}{\partial t}m + \frac{\partial}{\partial x}\left(\frac{\alpha m^2}{A}\right) + \frac{A}{\rho}\frac{\partial}{\partial x}p(A) = -\frac{2\alpha}{\alpha-1}\nu\frac{m}{A}.
\end{aligned}$$

Let the wall stiffness $\beta \geq 1$ be given. Then the pressure law is given by

$$p(A) = G_0\left(\left(\frac{A}{A_0}\right)^{\frac{\beta}{2}} - 1\right).$$

The case $\beta = 1$ is a commonly studied situation, while $\beta \to \infty$ represents the case when the wall stiffness tends to ∞ and, thus, no extension takes place. We, therefore, have the following closed system

$$\begin{aligned}
&\frac{\partial}{\partial t}A + \frac{\partial}{\partial x}m = 0,\\
&\frac{\partial}{\partial t}m + \frac{\partial}{\partial x}\left(\frac{\alpha m^2}{A} + \frac{G_0\beta A_0}{\rho(\beta+2)}\left(\frac{A}{A_0}\right)^{\frac{\beta}{2}+1}\right) = -\frac{2\alpha}{\alpha-1}\nu\frac{m}{A}.
\end{aligned}$$

We define $a := A/A_0 > 1$,

$$F(a,m) := \frac{\alpha m^2}{A_0}\frac{1}{a} + \frac{G_0\beta A_0}{\rho(\beta+2)}a^{\frac{\beta}{2}+1},$$

and assume A_0 to be a constant, as well as that the process is stationary, i.e.,

$$\frac{\partial}{\partial t}A = 0, \quad \frac{\partial}{\partial t}m = 0, \quad (x,t) \in [0,L]\times[0,T].$$

This implies that $m(x,t) = m_0$ for all $(x,t) \in [0,L]\times[0,T]$. The equation to solve for a is then given by

$$\frac{\partial}{\partial x}F(a,m) = -\frac{2\alpha}{\alpha-1}\frac{\nu m}{A_0}\frac{1}{a}.$$

Separation of variables leads to the implicit equation

$$\mathcal{F}(a,m)(x) = \mathcal{F}(a,m)(x_0) - \frac{2\alpha}{\alpha-1}\nu m(x-x_0)$$

for a, where $x_0 = 0$ indicates the left end of the pipe and

$$\mathcal{F}(a,m)(x) = -\alpha\frac{m^2}{A_0}\ln(a(x)) + \frac{G_0\beta A_0}{\rho(\beta+4)}a(x)^{\frac{\beta}{2}+2}.$$

We assume positive pressure in this particular pipe and obtain

$$\frac{\partial}{\partial a}\mathcal{F}(a,m) = -\frac{\alpha m^2}{a} + \frac{G_0\beta}{2\rho}a^{\frac{\beta}{2}+1}.$$

Clearly, for suitable parameters, $\partial_a\mathcal{F}(a,m) > 0$ and

$$\frac{\partial^2}{\partial a^2}\mathcal{F}(a,m) = \frac{\alpha m^2}{A_0}\frac{1}{a^2} + \frac{G_0\beta A_0}{4\rho}(\beta+2)a^{\frac{\beta}{2}+1} > 0$$

holds. Thus, as a function of a, $\mathcal{F}(a,m)$ is, on $(1,\infty)$, strictly monotonically increasing and convex. Hence, the inverse is concave and, as a function of $a(x_0)$, $a(\cdot)$ is concave because minus the inverse of $\mathcal{F}(\cdot,m)$ is convex and $\mathcal{F}(\cdot,m)$ is convex as well as strictly monotonically increasing in the given range. A detailed analysis as in the previous example also reveals this fact. In summary, models of type (1) that include the discussed system satisfy Assumption 1 and can thus be tackled by the proposed method.

4.3. **Stationary Gas Transport Optimization.** The optimization of gas transport networks is currently a highly active field of research of applied optimization; for an overview of the literature see the recent book [26] and the survey article [35] as well as the references therein.

One of the main questions in gas transport is to decide feasibility of nominations, i.e., of prescribed supply and discharge flows together with additional restrictions like bounds for the gas pressures etc. Gas mainly flows from higher to lower pressures. Thus, in order to transport gas over large distances through pipeline systems it is required to increase the gas pressure by compression. This is realized by compressors that can be, among other network devices, controlled by the dispatcher and thus add discrete aspects to the problem. In combination with highly nonlinear gas physics, entire gas transport models are mixed-integer nonlinear problems governed by differential equations for modeling gas physics.

We focus on two special cases: First, we consider the stationary case, i.e., the network is in an equilibrium. Second, we restrict ourselves to transport networks that are trees—a case that is frequently discussed in the literature; see, e.g., [35] for a recent survey and, e.g., [21] for a theoretical study. This case leads to a-priorily known flows in the network. In what follows we describe the considered mixed-integer model with ODEs and show why it can be tackled with the algorithm presented above.

We model a gas network as a directed graph $G = (V, A)$ with node set V and arc set A. The set of nodes is partitioned into the set of entry nodes V_+, where gas is supplied, and the set of exit nodes V_-, where gas is discharged from the network, and the set of inner nodes V_0. Entry nodes model, e.g., points where liquefied natural gas (LNG) carriers arrive and feed in gas, whereas exit nodes may be power plants or municipal utilities. The set of arcs consist of pipes A_{pi}, control valves A_{cv}, and compressor machines A_{cm}, which all are described in detail below.

Gas flow in networks is mainly described by mass flow q and the three gas state quantities pressure p, temperature T, and density ρ. These state quantities are coupled by an equation of state, which we choose to be the thermodynamical standard equation for real gases

$$\rho R_{\mathrm{s}} z T = p, \tag{13}$$

where R_{s} is the specific gas constant and z is the compressibility factor that we model using the formula

$$z(p,T) = 1 + \alpha p, \quad \alpha = 0.257\frac{1}{p_{\mathrm{c}}} - 0.533\frac{T_{\mathrm{c}}}{p_{\mathrm{c}} T}$$

of the American Gas Association; see, e.g., [27]. The quantities p_{c} and T_{c} are the pseudocritical pressure and temperature, respectively, which we assume to be constant. We further assume isothermal gas flow, i.e., we fix the gas temperature T at a suitable constant value. This is a usual assumption in gas transport optimization since gas pipelines often lie under ground and gas coolers and heaters exist at compressor stations and control valves to control the outflow temperature. We associate positive gas flow on arcs $a = (u, v)$ with flow in arc direction, i.e., $q_a > 0$ if gas flows from u to v and $q_a < 0$ if gas flows from v to u. The sets $\delta^{\mathrm{in}}(u) := \{a \in$

$A : a = (v, u)\}$ and $\delta^{\text{out}}(u) := \{a \in A : a = (u, v)\}$ are the sets of in- and outgoing arcs for node $u \in V$.

For each node $u \in V$, we assume lower and upper values p_u^- and p_u^+ to be given that bound the corresponding pressure variable p_u, i.e.,

$$p_u \in [p_u^-, p_u^+] \quad \text{for all } u \in V. \tag{14}$$

In addition, we model mass conservation by

$$\sum_{a \in \delta^{\text{out}}(u)} q_a - \sum_{a \in \delta^{\text{in}}(u)} q_a = q_u \begin{cases} \geq 0, & u \in V_+, \\ \leq 0, & u \in V_-, \\ = 0, & u \in V_0, \end{cases} \quad \text{for all } u \in V. \tag{15}$$

For every arc $a \in A$ there is a mass flow variable q_a that is bounded from below and above, i.e.,

$$q_a \in [q_a^-, q_a^+] \quad \text{for all } a \in A. \tag{16}$$

The remaining model of an arc depends on the specific arc type.

Pipes $a \in A_{\text{pi}}$ are used to transport the gas through the network. They typically outnumber all other network elements. A pipe is specified by its length L_a, its diameter D_a, and its friction factor λ_a, which we model using the empirical formula of Prandtl–Colebrook; see, e.g., [8] or [38, Chap. 9]. We assume each pipe to be cylindrically shaped and horizontal. In this situation, isothermal gas flow through a pipe is described by a system of partial differential equations—the Euler equations for compressible fluids [11]—consisting of the continuity and the momentum equation that form a quasilinear system of hyperbolic balance laws. In what follows, we only consider the case in which the gas network is in an equilibrium. In this stationary case the continuity equation states constant mass flow for every pipe, justifying the choice for a single flow variable q_a for all arcs $a \in A$. Thus, we are left with the stationary variant

$$\partial_x \left(p_a + \frac{\bar{q}_a^2}{\rho_a} \right) = -\frac{1}{2} \theta_a \frac{\bar{q}_a |\bar{q}_a|}{\rho_a}, \quad \bar{q}_a = q_a / A_a, \quad \theta_a = \frac{\lambda_a}{D_a}, \quad a \in A_{\text{pi}}, \tag{17}$$

of the momentum equation, coupling density $\rho_a = \rho_a(x)$ and pressure $p_a = p_a(x)$ with mass flow q_a along the arc, i.e., $x \in [0, L_a]$. In words, the momentum equation (17) describes the pressure loss in a pipe due to ram pressure and frictional forces. Note that the density ρ in (17) can be eliminated using the equation of state (13). The coupling of the pressure solution of (17) with node pressure variables of (14) is given by

$$p_u = p_a(0), \quad p_v = p_a(L_a) \quad \text{for all } a = (u, v) \in A_{\text{pi}}. \tag{18}$$

We remark that we also consider so-called short pipes that are used as auxiliary network elements, e.g., to model multiple entries at a single entry node. They can be seen as pipes $a = (u, v)$ with length 0, yielding $p_u = p_v$.

Control valves $a \in A_{\text{cv}}$ are used to decrease gas pressure. This is mainly required at transition points between large transport pipelines and regional substructures that are not able to handle high pressure levels. These elements involve some discrete aspects since they can be operated in different modes: They can be active, in bypass mode, or closed. Closed control valves simply block the gas flow ($q_a = 0$) and thus decouple the in- and outflow pressure. If they are open, control valves can operate in bypass mode, yielding equal pressures $p_u = p_v$. Finally, if activated, control valves are able to decrease the inflow pressure by a controllable amount $\Delta_a \in [\Delta_a^-, \Delta_a^+]$.

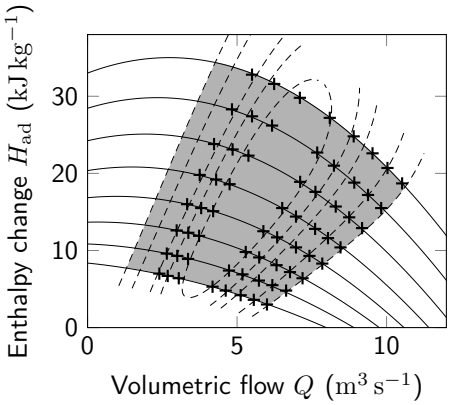

FIGURE 5. Characteristic diagrams of a turbo compressor. Feasible operating ranges are marked gray.

In summary, the complete description reads

$$a \text{ is active} \implies p_v = p_u - \Delta_a, \quad \Delta_a \in [\Delta_a^-, \Delta_a^+], \tag{19a}$$

$$a \text{ is in bypass mode} \implies p_v = p_u, \tag{19b}$$

$$a \text{ is closed} \implies q_a = 0. \tag{19c}$$

More detailed information about control valves and specific MIP models can be found in [16].

Finally, we describe our model of compressor machines $a \in A_{\mathrm{cm}}$. They are used to increase the inflow gas pressure to a higher outflow pressure in order to transport gas over large distances. In general, a compressor machine can be in the same three modes as control valves, see (19). However, the active state is much more complicated. We only consider so-called turbo compressors that are typically modeled by characteristic diagrams; see Figure 5 for an example. The gray area denotes the feasible operating range of the machine and the additionally given isolines allow to compute the required power P_a for the compression process. It turns out that the model of a turbo compressor is highly nonlinear and nonconvex. Since our focus here does not lie on detailed compressor modeling we use known mixed-integer linear outer approximations

$$c_a(p_u, p_v, q_a, P_a, y_a) \geq 0 \quad \text{for all } a = (u, v) \in A_{\mathrm{cm}} \tag{20}$$

of the operating ranges. The variables y_a are additional auxiliary variables required to formulate the specific outer approximation model; see [16] for the details.

We now collect all component models and obtain the entire optimization problem

$$\min \quad \sum_{a \in A_{\mathrm{cm}}} P_a \tag{21a}$$

$$\text{s.t.} \quad \text{pressure and flow bounds: (14), (16);} \tag{21b}$$

$$\text{mass conservation: (15);} \tag{21c}$$

$$\text{pipe model: (17), (18);} \tag{21d}$$

$$\text{control valve model: (19);} \tag{21e}$$

$$\text{compressor model: (20).} \tag{21f}$$

This model is a nonconvex mixed-integer optimization problem that contains the ODE (17) for every pipe. Typically, the ODE is discretized or replaced by an

approximation of its solution in order to obtain a finite-dimensional problem; see [26, 40] and the references therein. We, however, follow a different approach and decompose (21) as described in Section 2.

To this end, we have to prove that the described model can be written as an optimization problem of type (1) with concave and strictly increasing functions f. In our setting, the functions f_d correspond to the relation between in- and outflow pressures of the pipes, which is given by (21d), i.e., $\mathcal{D} = A_{\mathrm{pi}}$. We now prove that the relation between in- and outflow pressures is concave and strictly increasing. Under the assumption of subsonic flow, i.e., for the squared mach number η satisfying

$$\eta_a(x) := \frac{\bar{q}_a^2 \hat{T}}{p_a(x)^2} < 1, \quad x \in [0, L_a], \quad \hat{T} := R_{\mathrm{s}} T \tag{22}$$

and a positive compressibility factor

$$z(p_a(x)) > 0 \quad \text{for all } x \in [0, L_a], \tag{23}$$

the pressure loss along pipe $a = (u, v) \in A_{\mathrm{pi}}$ for given inflow pressure $p_u = p_a(0)$ and mass flow $\bar{q}_a$ reads

$$p_a(x; p_u, \bar{q}_a) = F_a^{-1}\left(F_a(p_u) - \frac{1}{2}\hat{T}\bar{q}_a|\bar{q}_a|\theta_a x\right), \quad x \in [0, L_a], \tag{24}$$

where

$$F_a(p_a) := \frac{1}{\alpha}p_a + \left(\bar{q}_a^2\hat{T} - \frac{1}{\alpha^2}\right)\ln(|1 + \alpha p_a|) - \bar{q}_a^2\hat{T}\ln(p_a)$$

holds; see [23] for the case of real gas and [22] for the case of ideal gas. The derivatives of F_a are given by

$$F_a'(p_a) = \frac{p_a^2 - \bar{q}_a^2\hat{T}}{p_a(1 + \alpha p_a)}, \quad F_a''(p_a) = \frac{p_a^2 + \bar{q}_a^2\hat{T}(1 + 2\alpha p_a)}{p_a^2(1 + \alpha p_a)^2}, \tag{25}$$

and satisfy

$$F_a'(p_a) > 0, \quad F_a''(p_a) > 0. \tag{26}$$

Thus, the use of the inverse in (24) is justified because F_a is strictly increasing in p_a for subsonic flow. From now on, we omit the parameters $p_u, \bar{q}_a$ of the function $p_a(x; p_u, \bar{q}_a)$ and assume $x \in (0, L_a]$. The first derivatives of the pressure solution are given by

$$\partial_{p_u} p_a(x) = \frac{F_a'(p_u)}{F_a'(p_a(x))}$$

and

$$\partial_{\bar{q}_a} p_a(x) = \frac{2\bar{q}_a\hat{T}p_a(x)z_a(p_a(x))}{p_a(x)^2 - \bar{q}_a^2\hat{T}}\left[\ln\left(\frac{z_a(p_u)}{z_a(p_a(x))}\right) - \ln\left(\frac{p_u}{p_a(x)}\right) - \frac{1}{2}\operatorname{sign}(\bar{q}_a)\theta_a x\right].$$

For positive flow they satisfy $\partial_{p_u} p_a(x) > 0$ and $\partial_{\bar{q}_a} p_a(x) < 0$. Having these preliminary results at hand, we can now state and prove concavity of $p_a(x)$ as a function of p_u.

Theorem 3. *Suppose that assumptions* (22), (23), *and* $\bar{q}_a > 0$ *hold. Then*

$$\partial_{p_u}^2 p_a(x) < 0 \quad \textit{for all } x \in (0, L_a].$$

Proof. For the ease of presentation, we now also drop the argument x and the arc index a of $p_a(x)$. A straightforward calculation gives

$$\partial_{p_u}^2 p = \frac{F'(p)F''(p_u) - F'(p_u)F''(p)\partial_{p_u} p}{F'(p)^2}$$

and substituting the first derivative yields

$$\partial_{p_u}^2 p = \frac{F''(p_u)}{F'(p)} - \frac{F'(p_u)^2 F''(p)}{F'(p)^3}.$$

We multiply by $F'(p)$ and obtain

$$F'(p)\partial_{p_u}^2 p = F''(p_u) - \left(\frac{F'(p_u)}{F'(p)}\right)^2 F''(p),$$

i.e., we see that because of (26) only the right-hand side determines the sign of $\partial_{p_u}^2 p$. Substituting (25) yields

$$\begin{aligned} & F'(p)\partial_{p_u}^2 p \\ &= \frac{p_u^2 + \bar{q}^2\hat{T}(1+2\alpha p_u)}{p_u^2(1+\alpha p_u)^2} - \left(\frac{(p_u^2 - \bar{q}^2\hat{T})\ p(1+\alpha p)}{(p^2 - \bar{q}^2\hat{T})\ p_u(1+\alpha p_u)}\right)^2 \frac{p^2 + \bar{q}^2\hat{T}(1+2\alpha p)}{p^2(1+\alpha p)^2} \\ &= \frac{p_u^2 + \bar{q}^2\hat{T}(1+2\alpha p_u)}{p_u^2(1+\alpha p_u)^2} - \frac{(p_u^2 - \bar{q}^2\hat{T})^2\ (p^2 + \bar{q}^2\hat{T}(1+2\alpha p))}{(p^2 - \bar{q}^2\hat{T})^2\ p_u^2(1+\alpha p_u)^2}. \end{aligned}$$

We take the common denominator and get

$$F'(p)\partial_{p_u}^2 p = \frac{(p_u^2 + \bar{q}^2\hat{T}(1+2\alpha p_u))(p^2 - \bar{q}^2\hat{T})^2 - (p_u^2 - \bar{q}^2\hat{T})^2(p^2 + \bar{q}^2\hat{T}(1+2\alpha p))}{p_u^2(1+\alpha p_u)^2(p^2 - \bar{q}^2\hat{T})^2}.$$

The sign is completely determined by the numerator. Using the squared mach numbers as defined in (22) we see that this numerator equals

$$p_u^2(1+\eta_0(1+2\alpha p_u))p^4(1-\eta)^2 - p_u^2(1+\eta_0(1+2\alpha p))p_u^4(1-\eta_0)^2,$$

where $\eta_0 = \eta_a(0)$ is used. We compare the relevant terms one by one. For positive flow the pressure is decreasing in x; see [23]. Thus, $p^4 < p_u^4$ holds and the squared mach numbers satisfy $\eta_0 < \eta$, yielding $(1-\eta)^2 < (1-\eta_0)^2$. Using $\alpha < 0$ we get $\alpha p > \alpha p_u$, which implies

$$1 + \eta_0(1+2\alpha p_u) < 1 + \eta_0(1+2\alpha p).$$

It remains to prove the positivity of $1+\eta_0(1+2\alpha p)$ and $1+\eta_0(1+2\alpha p_u)$. Combining (22) and (23) we get

$$1 + \eta_0(1+2\alpha p) = 1 - \eta_0 + 2\eta_0(1+\alpha p) > 0.$$

The term $1+\eta_0(1+2\alpha p_u)$ can be treated analogously. Finally, putting all together yields that the above given numerator is strictly negative and the claim $\partial_{p_u}^2 p_a(x) < 0$ follows. □

5. Numerical Results

We now present and discuss numerical results of Algorithm 1 applied to Problem (21). Our real-world test instance is the Greek natural gas transport network that is operated by the Greek natural gas transmission system operator DESFA; see [9]. Figure 6 visualizes the network topology and Table 1 gives some statistical information about the network. The network data is also part of the GasLib library of publicly available gas transport networks, i.e., all data required for the computations is available at the GasLib website[1]. Note that the considered network is a tree. Since detailed information about the compressor are not publicly available, we use the data of the compressor `compressor_1` of the publicly available instance GasLib-135; see [25]. Daily nomination data ranging from 11/01/2011 to 02/17/2016 is available on the DESFA website. 336 days in this range have imbalanced nominations or

[1]The network data can be downloaded as the GasLib-134 instance at `http://gaslib.zib.de`.

FIGURE 6. The Greek gas network with entries (triangle), one compressor machine (circle), and one control valve (square). The large number of exits (45) are distributed over entire network.

TABLE 1. Data of the Greek gas network with 134 nodes (above) and 133 arcs (below). One of the entry nodes is of LNG type. Eleven of the exit nodes are power plants.

Type	Quantity
Entries	3
Exits	45
Inner nodes	86
Pipes	86
Short pipes	45
Control valves	1
Compressors	1

other invalid data, yielding an instance set of 1234 remaining days. We additionally increase all supplied and discharged flows by a factor of 2 in order to increase the nonlinearity and thus the overall hardness of the problem.

Algorithm 1 iteratively solves an MIP and multiple NLP models. We implemented the MIP models using the C++ software framework LaMaTTO++ [29] and used Gurobi 6.5.0 [20] for solving the MIPs. All NLP models are implemented in Matlab R2015b [33] and solved using fmincon with default settings. That is, we use the internal SQP solver with given first- and second-order information and a constraint violation and termination tolerance of 10^{-6}. All computations have been performed on an Intel© Core™i5-3360M CPU with 4 cores and 2.8 GHz each and

TABLE 2. Basic overview of the results; separated for feasible and infeasible instances. The number of iterations is denoted by k and all runtimes are given in seconds.

Solution status	#	∅k	∅ Runtime	∅ Runtime (id.)
optimal	710	4.20	14.99	1.09
infeasible	524	1.00	0.59	0.06

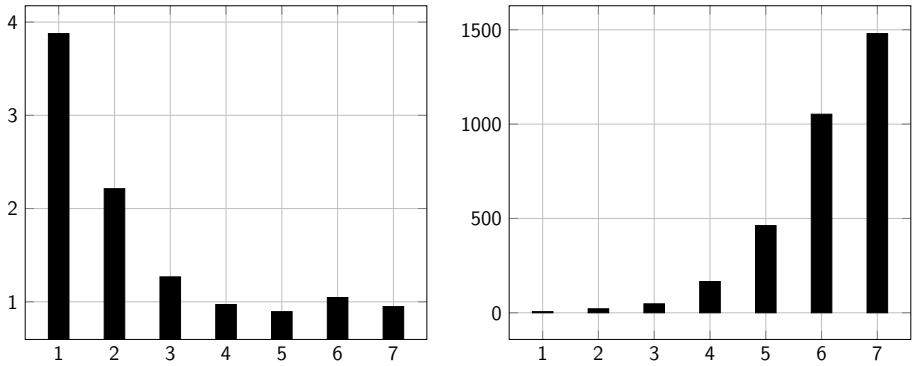

FIGURE 7. Average fmincon iterations for all subproblems (left) and Gurobi simplex iterations for all master problems (right) for all instances in iteration k of the decomposition method.

4 GB RAM. We choose $\varepsilon = 10^{-5}$ (scaled to denote error in bar for the pressure loss on pipes) as the tolerance of Algorithm 1.

Table 2 gives a basic overview of the results. All instances have been solved by our algorithm: Approximately 58 % of the instances are solved to global optimality whereas ∼42 % are proven to be infeasible by our method. It can be seen that infeasibility is detected faster than global optimality is proven (∼1 vs. 4 iterations). The last two columns state the average runtime of our current implementation (∅ Runtime)—in which we solve all 86 subproblem NLPs (one for every pipe) sequentially—and the idealized runtime (∅ Runtime (id.)) that we obtain if all independent subproblems would have been solved in parallel.

For a more detailed analysis of the results and the behavior of our algorithm we have randomly chosen 30 instances and state the results in Table 3. Here, 66 % of the considered instances are solved to optimality while the remaining 33 % are proven to be infeasible. Feasible instances can be further distinguished into nominations that require compression (i.e., the objective value is larger than zero) and those that do not require compression. It is noteworthy, that ∼71 % of all feasible instances require compression yielding an optimal objective greater than 0. Regarding the method itself it can be seen again that infeasibility is detected quite fast (0 or 1 iteration), while finding a global optimal solution requires more iterations (3 to 5).

The required iteration numbers are quite low for the considered class of gas transport problems. The reason is that the relation of in- and outflow pressures on tree-like gas transport networks has rather mild nonlinearities. This can be seen in Figure 8, where the refined relaxations of this nonlinearity is shown over the course of the iterations. The initial relaxation after preprocessing is displayed in the top-left figure. On a large scale this relaxation already seems to be quite tight. It will later turn out that the optimal solution will lie around an inflow pressure of approximately 48 bar. A zoom into this region is shown in the top-right figure.

TABLE 3. Detailed results for 30 randomly chosen instances. Solution status (Stat.), optimal objective function value (Obj.), required number of iterations (k) and total runtimes (Total), idealized total runtimes (Total (id.)) in case of subproblem parallelization, aggregated runtimes of all master problems (Master), of all subproblems (Sub), and of all subproblems in case of parallelization (Sub (id.)). All runtimes are given in seconds.

				Runtimes				
Instance	Stat.	Obj.	k	Total	Total (id.)	Master	Sub	Sub (id.)
11/05/2011	inf.	—	1	0.28	0.02	0.01	0.27	0.01
11/08/2011	inf.	—	1	0.29	0.02	0.01	0.28	0.01
12/23/2011	opt.	525.81	4	12.14	0.88	0.06	12.08	0.82
04/03/2012	opt.	362	4	12.28	0.83	0.09	12.19	0.74
04/19/2012	opt.	314.23	4	13.29	0.95	0.14	13.15	0.81
05/29/2012	opt.	510.19	4	11.94	0.83	0.06	11.89	0.77
07/29/2012	opt.	455.45	4	20.63	1.38	0.11	20.51	1.27
09/19/2012	opt.	0	5	14.99	1.04	0.18	14.81	0.86
10/08/2012	opt.	497.92	4	12.59	0.88	0.08	12.51	0.80
03/06/2013	inf.	—	1	0.26	0.01	0.01	0.25	0.01
03/16/2013	inf.	—	1	0.30	0.03	0.01	0.29	0.02
04/24/2013	opt.	644.42	4	11.94	0.84	0.06	11.88	0.78
05/27/2013	opt.	332.4	4	12.31	0.96	0.07	12.24	0.88
08/24/2013	opt.	334.23	4	12.46	0.90	0.09	12.37	0.82
01/25/2014	inf.	—	1	0.40	0.05	0.01	0.40	0.04
05/15/2014	opt.	0	5	15.30	1.14	0.20	15.10	0.94
07/04/2014	inf.	—	1	0.32	0.03	0.01	0.32	0.03
07/23/2014	opt.	0	5	16.75	1.26	0.23	16.51	1.03
08/12/2014	opt.	281.57	4	12.27	0.87	0.08	12.19	0.79
09/04/2014	opt.	441.71	4	12.75	0.84	0.06	12.69	0.78
09/07/2014	opt.	250.44	4	13.03	1.00	0.12	12.91	0.88
11/06/2014	inf.	—	0	0.00	0.00	0.00	0.00	0.00
11/14/2014	inf.	—	0	0.00	0.00	0.00	0.00	0.00
11/25/2014	opt.	440.35	5	18.56	1.45	0.17	18.40	1.28
03/21/2015	opt.	0	3	9.54	0.79	0.07	9.47	0.72
07/19/2015	opt.	405.8	5	17.16	1.18	0.19	16.97	0.99
08/27/2015	opt.	282.92	4	12.50	0.92	0.10	12.39	0.82
09/06/2015	opt.	397.69	4	14.42	0.98	0.07	14.35	0.91
11/06/2015	inf.	—	1	0.26	0.03	0.01	0.26	0.02
12/14/2015	inf.	—	1	0.28	0.01	0.00	0.28	0.01

It can be seen that the possible error is ~ 0.5 bar to 0.75 bar, which is still large compared to the tolerance of 1×10^{-5} bar. However, the next iterations directly give very tight relaxations (see bottom-left and -right figure) that yield the reported low iteration numbers.

Finally, we discuss the complexity of the overall method that is determined by the complexity of the master and subproblems as well as the number of iterations of the method itself. Applied to the gas transport problem, the number of (independent) NLPs to be solved in the subproblem is constant over all iterations and equals the number of pipes in the network. Thus, the complexity of all subproblems stays the same over the course of the iterations. Nevertheless, the NLPs become easier

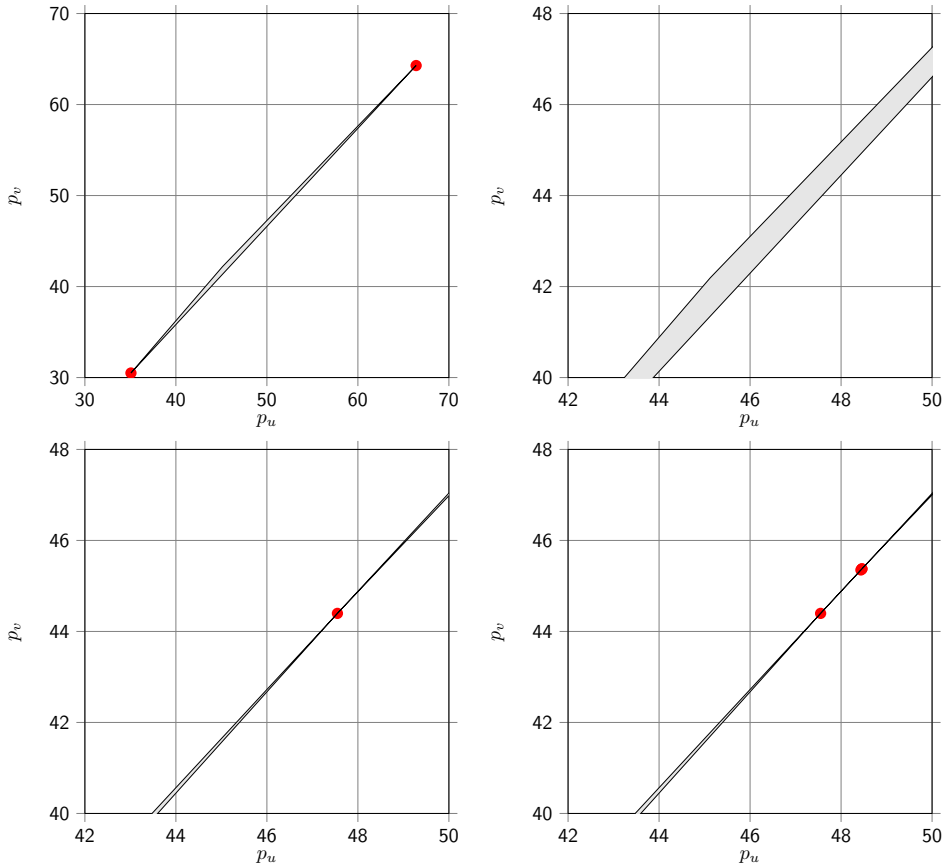

FIGURE 8. Tightened relaxation of the in- and outflow pressure relation over the course of the iterations

to solve as the master problems provide better initial solutions from iteration to iteration. This can be seen in Figure 7 (left), where the respective average fmincon iterations for all 1234 instances are plotted. Note further that there are even average fmincon iteration values less than 1. In these cases, the master problem's solution is often already feasible for the subproblem. The development of the complexity of the MIPs of the master problem over the course of the iterations is different. These MIPs grow from iteration to iteration due to added cut inequalities and additional binary variables. Since we never add more than one discretization point per iteration and pipe, this growth is bounded from above and the size of the MIP (in terms of constraints and variables) grows linearly over the course of the iterations. This linear growth translates into an exponential increase in simplex iterations; see Figure 7 (right). However, this fact cannot be avoided when solving MIPs since it is due to the exponential worst case complexity of MIPs themselves. As a consequence, our method is especially tailored for instances that yield low iteration counts of the overall method as it is the case for the gas transport problem at hand.

6. Conclusions

In this paper we proposed a decomposition based method for solving mixed-integer nonlinear problems, where the nonlinearity is not explicitly known. Those "black-box" nonlinearities may arise, e.g., due to differential equations for which no analytic solution is known or due to expensive simulation runs. By assuming certain

analytic properties of the nonlinearities we can prove that our algorithm finitely terminates with a globally optimal solution or with an indication of infeasibility. The applicability of our method is illustrated by three case studies from the field of steady-state gas transport, of pressurized flows in pipes with elastic walls, and from the field of optimal control. It is shown that all these applications satisfy the analytical assumptions on the nonlinearities. For the gas transport problem we also present promising numerical results on instances of the Greek natural gas transport network. We are convinced that comparable results could be obtained on similar problems that are also defined on a network like, e.g., water or power flow networks.

Nonetheless, there is still room for improvement and there are more complicated cases to consider. For instance, one focus of our future work will be to extend the algorithmic ideas in this paper to the case of mixed-integer nonlinear models where the nonlinearities are given by time-dependent partial differential equations.

Acknowledgements

We acknowledge funding through the DFG SFB/Transregio 154, Subproject A05, B07, B08, and C03. This research has been performed as part of the Energie Campus Nürnberg and supported by funding through the "Aufbruch Bayern (Bavaria on the move)" initiative of the state of Bavaria. We are also grateful to Nikolaos Kanelakis for his help in preparing the data of the Greek natural gas transport network. Finally, we thank Robert Burlacu, Falk Hante, Fabian Rüffler, and Lars Schewe for many fruitful discussions on the topic of the paper.

References

[1] P. Belotti, C. Kirches, S. Leyffer, J. Linderoth, J. Luedtke, and A. Mahajan. "Mixed-integer nonlinear optimization." In: *Acta Numerica* 22 (May 2013), pp. 1–131. DOI: 10.1017/S0962492913000032.

[2] J. F. Benders. "Partitioning procedures for solving mixed-variables programming problems." In: *Numerische Mathematik* 4.1 (1962), pp. 238–252. DOI: 10.1007/BF01386316.

[3] P. Bonami, G. Cornuéjols, A. Lodi, and F. Margot. "A Feasibility Pump for mixed integer nonlinear programs." In: *Mathematical Programming* 119.2 (2009), pp. 331–352. DOI: 10.1007/s10107-008-0212-2.

[4] J. F. Bonnans, J. C. Gilbert, C. Lemaréchal, and C. A. Sagastizábal. *Numerical Optimization: Theoretical and Practical Aspects*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2006. DOI: 10.1007/978-3-540-35447-5.

[5] A. Bressan, S. Čanić, M. Garavello, M. Herty, and B. Piccoli. "Flows on networks: recent results and perspectives." In: *EMS Surveys in Mathematical Sciences* 1.1 (2014), pp. 47–111. DOI: 10.4171/EMSS/2.

[6] C. Buchheim, C. Meyer, and R. Schäfer. *Combinatorial Optimal Control of Semilinear Elliptic PDEs*. Tech. rep. TU Dortmund, Oct. 2015. URL: http://www.optimization-online.org/DB_HTML/2015/10/5161.html.

[7] B. Chachuat, A. B. Singer, and P. I. Barton. "Global mixed-integer dynamic optimization." In: *AIChE Journal* 51.8 (2005), pp. 2235–2253. DOI: 10.1002/aic.10494.

[8] C. F. Colebrook. "Turbulent flow in pipes with particular reference to the transition region between smooth and rough pipe laws." In: *Journal of the Institution of Civil Engineers* 11 (4 Feb. 1939), pp. 133–156. DOI: 10.1680/ijoti.1939.13150.

[9] *DESFA*. 2016. URL: http://www.desfa.gr (visited on 05/03/2016).

[10] M. A. Duran and I. E. Grossmann. "An outer-approximation algorithm for a class of mixed-integer nonlinear programs." In: *Mathematical Programming* 36.3 (1986), pp. 307–339. DOI: 10.1007/BF02592064.

[11] M. M. Feistauer. *Mathematical methods in fluid dynamics*. Pitman monographs and surveys in pure and applied mathematics. Harlow, Essex, England: Longman Scientific & Technical New York, 1993. URL: http://opac.inria.fr/record=b1084590.

[12] R. Fletcher and S. Leyffer. "Solving mixed integer nonlinear programs by outer approximation." In: *Mathematical Programming* 66.1 (1994), pp. 327–349. DOI: 10.1007/BF01581153.

[13] A. Fügenschuh, B. Geißler, A. Martin, and A. Morsi. "The Transport PDE and Mixed-Integer Linear Programming." In: *Models and Algorithms for Optimization in Logistics*. Ed. by C. Barnhart, U. Clausen, U. Lauther, and R. H. Möhring. Dagstuhl Seminar Proceedings 09261. Dagstuhl, Germany: Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, Germany, 2009. URL: http://drops.dagstuhl.de/opus/volltexte/2009/2167.

[14] B. Geißler. "Towards Globally Optimal Solutions for MINLPs by Discretization Techniques with Applications in Gas Network Optimization." PhD thesis. Friedrich-Alexander-Universität Erlangen-Nürnberg, 2011.

[15] B. Geißler, O. Kolb, J. Lang, G. Leugering, A. Martin, and A. Morsi. "Mixed integer linear models for the optimization of dynamical transport networks." In: *Mathematical Methods of Operations Research* 73.3 (2011), pp. 339–362. DOI: 10.1007/s00186-011-0354-5.

[16] B. Geißler, A. Martin, A. Morsi, and L. Schewe. "The MILP-relaxation approach." In: *Evaluating Gas Network Capacities*. SIAM-MOS series on Optimization. SIAM, 2015. Chap. 6, pp. 103–122. DOI: 10.1137/1.9781611973693.ch6.

[17] B. Geißler, A. Martin, A. Morsi, and L. Schewe. "Using Piecewise Linear Functions for Solving MINLPs." In: *Mixed Integer Nonlinear Programming*. Ed. by J. Lee and S. Leyffer. New York, NY: Springer New York, 2012, pp. 287–314. DOI: 10.1007/978-1-4614-1927-3_10.

[18] B. Geißler, A. Morsi, and L. Schewe. "A New Algorithm for MINLP Applied to Gas Transport Energy Cost Minimization." In: *Facets of Combinatorial Optimization: Festschrift for Martin Grötschel*. Ed. by M. Jünger and G. Reinelt. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 321–353. DOI: 10.1007/978-3-642-38189-8_14.

[19] A. M. Geoffrion. "Generalized Benders decomposition." In: *Journal of Optimization Theory and Applications* 10.4 (1972), pp. 237–260. DOI: 10.1007/BF00934810.

[20] Z. Gu, E. Rothberg, and R. Bixby. *Gurobi Optimizer Reference Manual, Version 6.5.0*. Houston, Texas, USA: Gurobi Optimization Inc., 2015.

[21] M. Gugat, M. Dick, and G. Leugering. "Gas Flow in Fan-Shaped Networks: Classical Solutions and Feedback Stabilization." In: *SIAM Journal on Control and Optimization* 49.5 (2011), pp. 2101–2117. DOI: 10.1137/100799824.

[22] M. Gugat, F. M. Hante, M. Hirsch-Dick, and G. Leugering. "Stationary states in gas networks." In: *Networks and Heterogeneous Media* 10.2 (2015), pp. 295–320. DOI: 10.3934/nhm.2015.10.295.

[23] M. Gugat, R. Schultz, and D. Wintergerst. *Networks of pipelines for gas with nonconstant compressibility factor: stationary states*. 2016, pp. 1–32. DOI: 10.1007/s40314-016-0383-z. URL: http://dx.doi.org/10.1007/s40314-016-0383-z.

[24] L. M. Hocking. *Optimal control: an introduction to the theory with applications*. Oxford University Press, 1991. DOI: 10.1002/acs.4480070107.

[25] J. Humpola, I. Joormann, N. Kanelakis, D. Oucherif, M. E. Pfetsch, L. Schewe, M. Schmidt, R. Schwarz, and M. Sirvent. *GasLib – A Library of Gas Network Instances*. Sept. 2017. URL: http://www.optimization-online.org/DB_HTML/2015/11/5216.html.

[26] T. Koch, B. Hiller, M. E. Pfetsch, and L. Schewe. *Evaluating Gas Network Capacities*. SIAM-MOS series on Optimization. SIAM, 2015. xvii + 364. DOI: 10.1137/1.9781611973693.

[27] J. Králik, P. Stiegler, Z. Vostrý, and J. Záworka. *Dynamic Modeling of Large–Scale Networks with Application to Gas Distribution*. Vol. 6. Studies in Automation and Control. New York: Elsevier Sci. Publ., 1988.

[28] V. Lakshmikantham, N. Shahzad, and W. Walter. "Convex dependence of solutions of differential equations in a Banach space relative to initial data." In: *Nonlinear Analysis: Theory, Methods & Applications* 27.12 (1996), pp. 1351–1354. DOI: 10.1016/0362-546X(95)00138-L.

[29] *LaMaTTO++: A Framework for Modeling and Solving Mixed-Integer Nonlinear Programming Problems on Networks*. 2017. URL: http://www.mso.math.fau.de/edom/projects/lamatto.html (visited on 04/13/2017).

[30] J. Macki and A. Strauss. *Introduction to optimal control theory*. Springer New York, 1982. DOI: 10.1007/978-1-4612-5671-7.

[31] O. L. Mangasarian. *Nonlinear Programming*. Society for Industrial and Applied Mathematics, 1994. DOI: 10.1137/1.9781611971255.

[32] H. M. Markowitz and A. S. Manne. "On the Solution of Discrete Programming Problems." In: *Econometrica* 25.1 (1957), pp. 84–110. URL: http://www.jstor.org/stable/1907744.

[33] *MATLAB and Statistics Toolbox Release 2015b*. Natick, Massachusetts, United States: The MathWorks Inc., 2015.

[34] M. E. Pfetsch, A. Fügenschuh, B. Geißler, N. Geißler, R. Gollmer, B. Hiller, J. Humpola, T. Koch, T. Lehmann, A. Martin, A. Morsi, J. Rövekamp, L. Schewe, M. Schmidt, R. Schultz, R. Schwarz, J. Schweiger, C. Stangl, M. C. Steinbach, S. Vigerske, and B. M. Willert. "Validation of nominations in gas network optimization: models, methods, and solutions." In: *Optimization Methods and Software* 30.1 (2015), pp. 15–53. DOI: 10.1080/10556788.2014.888426.

[35] R. Z. Ríos-Mercado and C. Borraz-Sánchez. "Optimization problems in natural gas transportation systems: A state-of-the-art review." In: *Applied Energy* 147 (2015), pp. 536–555. DOI: 10.1016/j.apenergy.2015.03.017.

[36] S. Sager. "Numerical methods for mixed-integer optimal control problems." PhD thesis. Universität Heidelberg, 2006. URL: http://mathopt.de/PUBLICATIONS/Sager2005.pdf.

[37] N. V. Sahinidis and I. E. Grossmann. "Convergence properties of generalized benders decomposition." In: *Computers & Chemical Engineering* 15.7 (1991), pp. 481–491. DOI: 10.1016/0098-1354(91)85027-R.

[38] J. Saleh. *Fluid Flow Handbook*. McGraw-Hill Handbooks. McGraw-Hill, 2002.

[39] J. Schelbert. "Discrete Approaches for Optimal Routing of High Pressure Pipes." PhD thesis. Friederich-Alexander-Universität Erlangen-Nürnberg, 2015.

[40] M. Schmidt, M. C. Steinbach, and B. M. Willert. "High detail stationary optimization models for gas networks." In: *Optimization and Engineering* 16.1 (2015), pp. 131–164. DOI: 10.1007/s11081-014-9246-x.

[41] H. Tuy. *Convex analysis and global optimization*. Vol. 22. Springer Science & Business Media, 2013. DOI: 10.1007/978-1-4757-2809-5.

[1]Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Lehrstuhl für Angewandte Mathematik 2, Cauerstr. 11, 91058 Erlangen, Germany, [2]Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Discrete Optimization, Cauerstr. 11, 91058 Erlangen, Germany, [3]Energie Campus Nürnberg, Fürther Str. 250, 90429 Nürnberg, Germany