The Cost of Not Knowing Enough: Mixed-Integer Optimization with Implicit Lipschitz Nonlinearities

Martin Schmidt, Mathias Sirvent, Winnifried Wollner

Abstract. It is folklore knowledge that nonconvex mixed-integer nonlinear optimization problems can be notoriously hard to solve in practice. In this paper we go one step further and drop analytical properties that are usually taken for granted in mixed-integer nonlinear optimization. First, we only assume Lipschitz continuity of the nonlinear functions and additionally consider multivariate implicit constraint functions that cannot be solved for any parameter analytically. For this class of mixed-integer problems we propose a novel algorithm based on an approximation of the feasible set in the domain of the nonlinear function—in contrast to an approximation of the graph of the function considered in prior work. This method is shown to compute global optimal solutions in finite time and we also provide a worst-case iteration bound. However, first numerical experiences reveal that a lot of work is still to be done for this highly challenging class of problems and we thus finally propose some possible directions of future research.

1. Introduction

Mixed-integer nonlinear optimization problems (MINLPs) are one of the most important classes of models in mathematical optimization. This is due to its capability of modeling decisions via incorporating discrete aspects as well as the possibility of modeling nonlinear phenomena. However, the combination of these two aspects also makes these problems very hard to solve [10, 24]. For a general overview about MINLP we refer to [1].

Both from a theoretic and algorithmic point of view it is important to classify MINLPs further. Probably the most important distinction is to be made between convex and nonconvex MINLPs. Since gradients yield valid cuts in the convex case, outer approximations of the feasible sets of the convex nonlinearities can be derived and exploited in algorithms [2, 5, 7]. In the nonconvex case this is not possible. Here, one typically needs to derive convex underestimators and concave overestimators that yield (piecewise) convex relaxations of the nonconvex nonlinearities [25, 43, 44].

The latter usually exploits known analytical properties of the nonconvex nonlinear functions, which is obviously not possible if these properties, like, e.g., differentiability, are not known or even knowledge about the explicit representation is missing. In this case, one typically tries to resort to Lipschitz assumptions about the nonlinearities, which leads to the field of global Lipschitz optimization; see, e.g., [19, 20, 22, 31–33, 45] to name only a few. For a more detailed overview about this field see the textbook [21] and the references therein.

In this paper, we focus on a specific setting that can be observed in a lot of applications (see below for some problem-specific references), namely that the Lipschitzian MINLP under consideration can be decomposed into a mixed-integer

Date: April 26, 2018.
2010 *Mathematics Subject Classification.* 90-08, 90C11, 90C26, 90C30, 90C90.
Key words and phrases. Mixed-Integer Nonlinear Optimization, Global Optimization, Lipschitz Optimization, Gas Networks.

linear (MILP) part and a nonlinear part. Our working hypothesis in this and also the preceding works [16, 17, 37, 41] is that the MILP part can be solved comparably fast and reliable whereas the nonlinearity really hampers the solution process—at least in combination with the MILP part of the problem. See also [3, 12, 13, 42], where this working hypothesis is followed as well. At this point we remark that the cited literature usually tries to get rid of the nonlinear functions by replacing or, so to say, re-modeling them using MILP-representable approximations. In this paper, we consider the case in which this re-modeling is not possible without adding additional unknowns and constraints to the problem.

One specific field of application that fits into the above discussion and that has been studied very successfully in recent years is mixed-integer nonlinear optimization of gas transport networks [8, 14, 15, 17, 30]; see the recent book [26] and the survey [34] for more references. This application will also be studied later in our numerical case study. MINLPs from gas transport optimization are defined on graphs that represent the transport network. Mass balances and simple physical as well as technical bounds yield linear constraints in this context and controllable elements are typically modeled by mixed-integer (non)linear sets of variables and constraints. Here, the complicating nonlinearity is the system of differential equations modeling the gas flow through pipes. In its full beauty, these equations yield implicit and highly nonlinear constraints that can be evaluated only by simulation and that do not possess additional and desired analytical properties like convexity; see [9, 23, 28, 38–40]. In almost all contributions of the literature cited above the authors make, depending on their special focus, tailored assumptions that allow to develop very effective solution approaches. In this paper, we follow the way paved by the paper [37], where general methods have been developed (and tested on gas transport problems) that only require Lipschitzian, 1-dimensional, and explicitly given nonlinearities. In this paper, we show that the approach can be extended to the case where the nonlinearity is only given implicitly. To this end, we relax the assumptions used in [37] as follows:

(1) We consider constraints that are only implicitly given, i.e., constraints of type $F(x) = 0$ with $x \in \mathbb{R}^n$ such that a representation $F(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) = x_i$ cannot be achieved analytically for any $i \in \{1, \dots, n\}$.
(2) We consider multi-dimensional constraint functions $F(x) = 0$ with $x \in \mathbb{R}^n$ and $n \geq 2$.

Regarding (1) it is of course possible to re-model the implicit constraint as $F(x) = y$ and by adding additional linear constraints $y = 0$. We, however, refrain from such a re-modeling in this paper and consider the case were such a re-modeling, which also enlarges the problem, is not appropriate. It will turn out that both aspects drastically accentuate the computational hardness of the problem.

Our contribution is the following. We formalize the problem class in Section 2 and the corresponding assumptions sketched above (Section 3). Based on this, we state an algorithm and prove its correctness, i.e., that it computes globally optimal solutions (or proves the infeasibility of the problem) in finite time. Moreover, we prove a worst-case iteration bound for the algorithm. Finally, we present a numerical case study in Section 4, which reveals the considerable computational hardness of the considered class of problems: Only very small gas transport networks can be solved in reasonable time under the stated very weak assumptions. This is why we consider the addressed problem class as an open computational challenge for which we state possible directions of future research in Section 5.

2. Problem Definition

We consider problems of the form

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

$$\text{s.t.} \quad Ax \geq b, \quad x \in [\underline{x}, \bar{x}], \quad x \in \mathbb{R}^n \times \mathbb{Z}^m, \tag{1b}$$

$$F_i(x_{\mathrm{c}(i)}) = 0, \quad i \in [p], \tag{1c}$$

where $A \in \mathbb{R}^{\ell \times (n+m)}$, $b \in \mathbb{R}^\ell$, $h \in \mathbb{R}^{n+m}$, and $[p] := \{1, \dots, p\}$. The p constraints $F_i : \mathbb{R}^{n_i} \to \mathbb{R}$, $i \in [p]$, comprise all nonlinearities of the problem. Here, and in what follows, we use the splitting $x = (x_\mathrm{c}, x_\mathrm{d})$ of the entire variable vector with $x_\mathrm{c} \in \mathbb{R}^n$ and $x_\mathrm{d} \in \mathbb{Z}^m$. That is, x_c are all continuous and x_d are all discrete variables of the problem and the nonlinear functions $F_i, i \in [p]$, only depend on the continuous variables $x_{\mathrm{c}(i)}$, i.e., $x_{\mathrm{c}(i)}$ is a sub-vector of the vector x_c of all continuous variables. We remark that the assumption that the nonlinearities only depend on the continuous variables is only made to simplify the technical discussions later on—it can always be formally satisfied by introducing auxiliary continuous variables.

Instead of optimizing the objective of (1) over the feasible set $\mathcal{F}$ given by (1b)–(1c) we replace $\mathcal{F}$ by an approximating sequence $\mathcal{F}_k \approx \mathcal{F}$ and globally optimize the sequence of problems

$$\min \{h^\top x \colon x \in \mathcal{F}_k\}. \tag{2}$$

The iteration can then be stopped once a solution x^k of (2) is close enough to the original feasible set $\mathcal{F}$. To this end let

$$\begin{aligned} \min_{x} \quad & h^\top x \\ \text{s.t.} \quad & Ax \geq b, \quad x \in [\underline{x}, \bar{x}], \quad x \in \mathbb{R}^n \times \mathbb{Z}^m, \\ & |F_i(x_{\mathrm{c}(i)})| \leq \varepsilon, \quad i \in [p], \end{aligned} \tag{3}$$

be the ε-relaxed version of the original problem (1). Note that we only relax the nonlinearities whereas all other constraints stay as they are. The precise choice of the approximate sets $\mathcal{F}_k$ will be detailed later in Section 3.

Definition 2.1 (ε-feasibility)**.** We call a point ε-feasible if it is feasible for Problem (3).

3. Algorithm

The main idea of our algorithm for solving Problem (1) to global optimality is to split the problem into its mixed-integer linear and its nonlinear part. The mixed-integer linear part is solved in the so-called master problem, which additionally contains a successively tightened approximation of the zero sets of the nonlinearities F_i, $i \in [p]$. Let x^k denote the master problem's solution in iteration k. Then, for every master problem solution, we check the feasibility of the solution w.r.t. the nonlinearities F_i. In case of feasibility, we have found a global optimal solution of the original problem; and in case of infeasibility, we construct a tighter approximation of the zero sets for the next master problem based on the information obtained by evaluating the nonlinearities F_i at the current master solution x^k. For the latter, we need (i) the function values $F_i(x^k_{\mathrm{c}(i)})$, $i \in [p]$, as well as (ii) the global Lipschitz constants L_i of the F_i.

Assumption 1. *We have an oracle that evaluates $F_i(x_{\mathrm{c}(i)})$ for all $i \in [p]$ and all F_i are globally Lipschitz continuous on $x_{\mathrm{c}(i)} \in [\underline{x}_{\mathrm{c}(i)}, \bar{x}_{\mathrm{c}(i)}]$ with known global Lipschitz constant L_i.*

We now show how to construct the successively tightened approximations of the zero sets in case of Assumption 1. Let Ω_i^k denote this approximation of the zero set of function F_i in iteration k of the algorithm. Initially, we start with the box $\Omega_i^0 := [\underline{x}_{\mathrm{c}(i)}, \bar{x}_{\mathrm{c}(i)}]$ defined by the original bounds of the problem (or after presolving; see Section 4). Assume now that the evaluation of the nonlinearity yields $|F_i(x^k_{\mathrm{c}(i)})| \geq \varepsilon$ for a prescribed tolerance $\varepsilon > 0$. We then exclude the box

$$B_i^k := \left\{ x \in \mathbb{R}^{n_i} \colon \|x - x^k_{\mathrm{c}(i)}\|_\infty < \frac{|F_i(x^k_{\mathrm{c}(i)})|}{L_i} \right\} \tag{4}$$

in the next iteration. If $|F_i(x^k_{\mathrm{c}(i)})| < \varepsilon$ holds, we set $B_i^k = \emptyset$. Putting these boxes together, we obtain the nonconvex approximation

$$\Omega_i^k = [\underline{x}_{\mathrm{c}(i)}, \bar{x}_{\mathrm{c}(i)}] \setminus \bigcup_{j \in [k-1]} B_i^j, \quad k = 1, 2, \dots \tag{5}$$

With this at hand, we can now formulate the master problem:

$$\min_x \quad h^\top x \tag{6a}$$

$$\text{s.t.} \quad Ax \geq b, \quad x \in [\underline{x}, \bar{x}], \quad x \in \mathbb{R}^n \times \mathbb{Z}^m, \tag{6b}$$

$$x_{\mathrm{c}(i)} \in \Omega_i^k, \quad i \in [p]. \tag{6c}$$

Before we proof correctness of the algorithm, i.e., finite termination of the algorithm at global optimal points, we first state and prove some properties of the master problem that will be used later and formally state the algorithm.

Proposition 3.1. *Assume that Assumption 1 holds. Then, it holds*

$$\Omega_i^k \supseteq (\ker(F_i) \cap [\underline{x}_{\mathrm{c}(i)}, \bar{x}_{\mathrm{c}(i)}])$$

for all $i \in [p]$ and all k. Thus, the master problem (6) is a relaxation of (1) for all k.

Proof. The proposition follows directly from the construction (4) and (5) and the Lipschitz continuity of F_i. □

The next lemma shows that we can use state-of-the-art MILP software for solving the master problems.

Lemma 3.2. *The nonconvex master problem (6) can be modeled as a mixed-integer linear problem. The number of additional variables and constraints required to formulate (6c) for an $i \in [p]$ is in $\mathcal{O}(k\,|c(i)|)$.*

Proof. The constraints $Ax \geq b$, $x \in [\underline{x}, \bar{x}]$, and $x \in \mathbb{R}^n \times \mathbb{Z}^m$ are obviously mixed-integer linear constraints. Thus, it remains to prove that $x_{c(i)} \in \Omega_i^k$, $i \in [p]$, can be formulated with mixed-integer linear constraints as well. We define the index set

$$\mathcal{I}_i^k := \{ j \in [k-1] \colon |F_i(x^j_{\mathrm{c}(i)})| \geq \varepsilon \}, \quad i \in [p],$$

for the boxes to exclude. Moreover, we define

$$\underline{x}^j_{\mathrm{c}(i)} := x^j_{\mathrm{c}(i)} - 1 \frac{|F_i(x^j_{\mathrm{c}(i)})|}{L_i} \quad \text{and} \quad \bar{x}^j_{\mathrm{c}(i)} := x^j_{\mathrm{c}(i)} + 1 \frac{|F_i(x^j_{\mathrm{c}(i)})|}{L_i} \tag{7}$$

for all $i \in [p]$ and $j \in \mathcal{I}_i^k$; see Figure 1 for an illustration. In (7) we use the notation 1 for the vector of ones in appropriate dimension. To model the gray area in Figure 1,

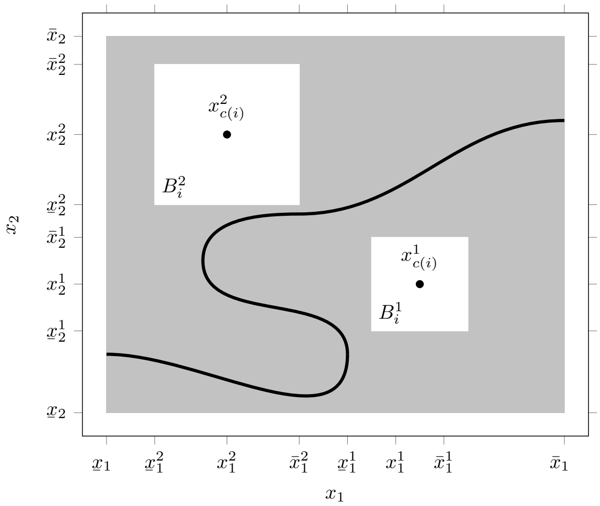

FIGURE 1. Set Ω_i^2 with $\mathrm{c}(i) = \{1, 2\}$.

we have to exclude the white rectangles B_i^j for all $j \in \mathcal{I}_i^k$. This can be done in the following way:

$$\Omega_i^k = \Big\{ x_{\mathrm{c}(i)} : z_l^{j,1}, z_l^{j,2} \in \{0,1\} \quad \text{for all } l \in \mathrm{c}(i), j \in \mathcal{I}_i^k, \tag{8a}$$

$$x_l \le \bar{x}_l^j + z_l^{j,1}(\bar{x}_l - \bar{x}_l^j) \quad \text{for all } l \in \mathrm{c}(i), j \in \mathcal{I}_i^k, \tag{8b}$$

$$x_l \ge \underline{x}_l + z_l^{j,1}(\bar{x}_l^j - \underline{x}_l) \quad \text{for all } l \in \mathrm{c}(i), j \in \mathcal{I}_i^k, \tag{8c}$$

$$x_l \le \underline{x}_l^j + (1 - z_l^{j,2})(\bar{x}_l - \underline{x}_l^j) \quad \text{for all } l \in \mathrm{c}(i), j \in \mathcal{I}_i^k, \tag{8d}$$

$$x_l \ge \underline{x}_l + (1 - z_l^{j,2})(\underline{x}_l^j - \underline{x}_l) \quad \text{for all } l \in \mathrm{c}(i), j \in \mathcal{I}_i^k, \tag{8e}$$

$$\sum_{l \in \mathrm{c}(i)} (z_l^{j,1} + z_l^{j,2}) \ge 1 \quad \text{for all } j \in \mathcal{I}_i^k \Big\}. \tag{8f}$$

We remark that (8b) and (8c) check if $x_l \le \bar{x}_l^j$. If this holds true, then $z_l^{j,1} = 0$. On the other hand, (8d) and (8e) check if $x_l \ge \underline{x}_l^j$. If this holds true, then $z_l^{j,2} = 0$. As we want to ensure that B_i^j is excluded, we have to add the inequality (8f). One can easily see that System (8) requires $\mathcal{O}(k\,|c(i)|)$ additional variables and constraints for every $i \in [p]$. □

We additionally remark that after having solved the master problem, all F_i can be evaluated in parallel in every iteration. The overall algorithm is formally stated in Algorithm 1. We now prove correctness of the algorithm, i.e., that the algorithm terminates after a finite number of iterations with a global optimal solution or with an indication of infeasibility of the original problem.

Theorem 3.3. *Suppose that Assumption 1 holds. Then, Algorithm 1 terminates after a finite number of iterations at an approximate globally optimal point x^k of Problem (1) or with an indication that Problem (1) is infeasible.*

Proof. We assume that the algorithm does not terminate after a finite number of iterations. That is, there exists an $i \in [p]$ and a corresponding subsequence (indexed by l) of iterates with

$$|F_i(x_{\mathrm{c}(i)}^l)| > \varepsilon. \tag{9}$$

Algorithm 1 Global Optimization of MINLPs with Implicit Nonlinearities

Require: Problem (1) and $\varepsilon > 0$.
Ensure: Returns an globally optimal point for Problem (1) or an indication of infeasibility.

```
Set k ← 0 and initialize Ω_i^0 ← [x̲_c(i), x̄_c(i)] for all i ∈ [p].
while true do
    Solve the master problem (6) to global optimality.
    if (6) is infeasible then return "Problem (1) is infeasible".
    Let x^k denote the optimal solution of (6).
    Evaluate F_i(x^k_c(i)) for all i ∈ [p].
    if |F_i(x^k_c(i))| ≤ ε for all i ∈ [p] then return x^k.
    for i ∈ [p] do
        if |F_i(x^k_c(i))| > ε then set Ω_i^{k+1} ← Ω_i^k \ B_i^k else set Ω_i^{k+1} ← Ω_i^k.
    end for
    Increase k ← k + 1.
end while
```

We investigate the master problem's solutions $x^l_{\mathrm{c}(i)}$. Since all variables are bounded, the subsequence $(x^l_{\mathrm{c}(i)})$ has a convergent subsequence $(x^\mu_{\mathrm{c}(i)})$. Thus, we can write

$$\|x^\alpha_{\mathrm{c}(i)} - x^\beta_{\mathrm{c}(i)}\|_2 < \delta \tag{10}$$

for all sufficiently large indices α and β of the μ-subsequence and arbitrarily small $\delta > 0$. However, all iterates $x^l_{\mathrm{c}(i)}$ are excluded from the subsequent feasible set, see (4), and together with (9) we can write

$$\|x^\alpha_{\mathrm{c}(i)} - x^\beta_{\mathrm{c}(i)}\|_2 \geq \frac{|F_i(x^k_{\mathrm{c}(i)})|}{L_i} > \frac{\varepsilon}{L_i},$$

for all α, β. This contradicts (10). □

Next, we establish a worst-case bound for required number of iterations.

Theorem 3.4. *For given $i \in [p]$, let $n_i = |c(i)|$, i.e., $x_{c(i)} \in \mathbb{R}^{n_i}$. Furthermore, let $\sigma^i := \bar{x}_{c(i)} - \underline{x}_{c(i)} \in \mathbb{R}^{n_i}$. Then, Algorithm 1 terminates after a maximum number of*

$$\sum_{i\in[p]} \prod_{j=1}^{n_i} \left(\left\lfloor \sigma^i_j \frac{L_i}{\varepsilon} + 1 \right\rfloor \right)$$

iterations.

Proof. We show that for each constraint $F_i(x_{\mathrm{c}(i)}) = 0$, $i \in [p]$, there are at most

$$\prod_{j=1}^{n_i} \left(\left\lfloor \sigma^i_j \frac{L_i}{\varepsilon} + 1 \right\rfloor \right)$$

iterations k for which $\Omega_i^{k+1} \neq \Omega_i^k$ holds. Since in each iteration at least one of the sets Ω_i^k needs to be changed this proves the assertion. To see the claim, we notice, that the hypercube $[\underline{x}_{\mathrm{c}(i)}, \bar{x}_{\mathrm{c}(i)}] \subset \mathbb{R}^{n_i}$ can be covered by

$$N = \prod_{j=1}^{n_i} \left(\left\lfloor \sigma^i_j \frac{L_i}{\varepsilon} + 1 \right\rfloor \right)$$

hypercubes H_ε^k, $k \in [N]$, with side-length ε/L_i. To see this, we note that each of the intervals $[\underline{x}_j, \bar{x}_j]$ for $j \in \mathrm{c}(i)$ can be decomposed as

$$\underline{x}_j = \underline{x}_j + 0 \cdot \frac{\varepsilon}{L_i} < \underline{x}_j + 1 \cdot \frac{\varepsilon}{L_i} < \cdots < \underline{x}_j + \left\lfloor \sigma_j^i \frac{L_i}{\varepsilon} \right\rfloor \frac{\varepsilon}{L_i} \leq \bar{x}_j$$

into $\lfloor \sigma_j^i L_i/\varepsilon + 1 \rfloor$ subintervals of length lower or equal ε/L_i. Taking the product of these intervals gives the desired cover of the hypercube.

Now, by definition, in an iteration k during which Ω_i^k is changed, a point $x_{\mathrm{c}(i)}$ together with its neighborhood B_i^k is excluded from Ω_i^k. In order for this to happen,

$$\frac{|F_i(x_{\mathrm{c}(i)})|}{L_i} > \frac{\varepsilon}{L_i}$$

needs to hold. Consequently, the neighborhood B_i^k contains a hypercube with side-length $2\varepsilon/L_i$. As a consequence, if $x_{\mathrm{c}(i)}$ is contained in H_ε^k, then no future iterate can be placed in this H_ε^k. This proves the claimed bound on the iterations. □

The last theorem shows that the maximum number of required iterations is bounded above by $\mathcal{O}((L^{\max}/\varepsilon)^{\bar{n}})$, where $\bar{n} = \max\{\mathrm{c}(i) \colon i \in [p]\}$ is the maximum number of arguments of the nonlinear functions F_i and $L^{\max} = \max\{L_i \colon i \in [p]\}$ is the maximum Lipschitz constant. Moreover, the asymptotic bounds derived in [29, 46] indicate that, in general, no better worst-case iteration bound can be expected.

Remark 3.5. During the course of the algorithm, the excluded boxes B_i^k are determined based on the master problem's solution. To tighten the feasible set of the initial master problem, we first equidistantly sample points from the initial box Ω_i^0. Afterward, we sort these sampling points by the excluded box volume, see (4), in descending order and add them to the initial master problem if they are not excluded by a previous sampling point. By doing so, we obtain a tighter feasible region from the beginning on at the cost of a larger MILP formulation.

4. Numerical Case Study

In this section, we present computational results of Algorithm 1 applied to stationary gas transport optimization. We briefly describe the model in Section 4.1 and discuss the results in Section 4.2.

4.1. **Stationary Gas Transport Optimization.** One central task in the gas industry is to transport prescribed supplied and discharged flows at minimum costs. Gas mainly flows from higher to lower pressures and in order to transport gas over large distances through pipeline systems, it is required to increase the gas pressure. This is done by compressors that add discrete aspects to the problem. We consider the stationary case, briefly present a mixed-integer model, and show how 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, the set of exit nodes V_-, where gas is discharged, and the set of inner nodes V_0. The set of arcs consists of pipes A_{pi} and compressors A_{cm}. We associate positive gas flow on arcs $a = (u, v)$ with mass 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. Moreover, mass flow variables q_a are bounded from below and above, i.e., $q_a \in [\underline{q}_a, \bar{q}_a]$. The sets $\delta^{\mathrm{in}}(u) := \{a \in A : a = (v, u)\}$ and $\delta^{\mathrm{out}}(u) := \{a \in A : a = (u, v)\}$ are the sets of in- and outgoing arcs for node $u \in V$. Thus, we model mass conservation by

$$\sum_{a \in \delta^{\text{out}}(u)} q_a - \sum_{a \in \delta^{\text{in}}(u)} q_a = q_u \begin{cases} \geq 0, & u \in V_+, \\ \leq 0, & u \in V_-, \\ = 0, & u \in V_0, \end{cases} \quad \text{for all } u \in V, \tag{11}$$

where q_u denotes the supplied or discharged flows. In addition, we need bounded pressure variables $p_u \in [\underline{p}_u, \bar{p}_u]$ for each node $u \in V$.

Pipes $a \in A_{\text{pi}}$ are used to transport the gas through the network. We consider their length L_a, their diameter D_a, their cross-sectional area A_a, their slope s_a, and their friction factor λ_a, which we model using the implicit formula of Prandtl–Colebrook; see, e.g., [4] or [35, Chap. 9]. Gas flow in networks is described by a system of partial differential equations—the Euler equations for compressible fluids [6]. In what follows, we consider the stationary case and assume small velocities, constant temperature $\hat{T}$, and constant compressibility factor $\hat{z}$. This leads to the so-called Weymouth equation, see [9, 12], that describes the relation between pressure and mass flow on a pipe via

$$p_v^2 - p_u^2 + \Lambda_a |q_a| q_a = 0 \quad \text{for all } a = (u, v) \in A_{\text{pi}}, \quad \Lambda_a = \frac{L_a \lambda_a R_{\text{s}} \hat{z} \hat{T}}{A_a^2 D_a}. \tag{12}$$

We remark that R_{s} and g denote the specific gas constant and the gravitational constant, respectively.

Finally, we describe our model of compressors $a \in A_{\text{cm}}$. They are used to increase the inflow gas pressure to a higher outflow pressure. In general, a compressor can be active, in bypass mode, or closed. Closed compressors simply block the gas flow ($q_a = 0$) and thus decouple the in- and outflow pressure. If compressors are open, they can operate in bypass mode, yielding equal pressures $p_u = p_v$. Finally, if activated, they are able to increase the outflow pressure. We only consider so-called turbo compressors that are typically modeled by characteristic diagrams, which determine the feasible range of an active compressor depending on p_u, p_v, and q_a; see, e.g. [38, 40] for a detailed description of turbo compressor models. 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

$$l_a(p_u, p_v, q_a, P_a, y_a) \geq 0 \quad \text{for all } a = (u, v) \in A_{\text{cm}} \tag{13}$$

of the operating ranges. In (13), P_a stands for the power required for compression and the variables y_a are additional auxiliary variables required to formulate the specific outer approximation model; see [9, 12] for the details.

We now collect all component models and obtain the mixed-integer optimization problem

$$\begin{aligned} \min \quad & \sum_{a \in A_{\text{cm}}} P_a \\ \text{s.t.} \quad & \text{pressure and flow bounds}, \quad && \text{mass conservation: (11)}, \\ & \text{pipe model: (12)}, && \text{compressor model: (13)}. \end{aligned}$$

Note that the only nonlinearity of the model is given by the pipe model (12). Consequently, we compute the global Lipschitz constant analytically and obtain (4) by evaluating (12).

4.2. **Results.** Our test instances are the networks GasLib-4 and GasLib-4-Tree; see Figure 2 left and right, respectively. These instances will be included in the next version of the publicly available GasLib test set; see [36].

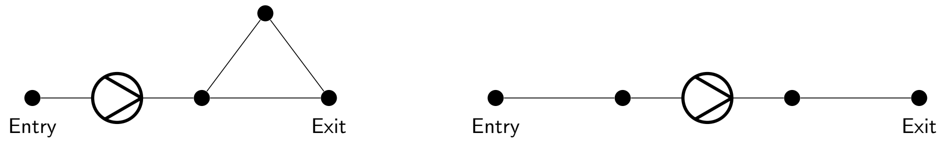

FIGURE 2. GasLib-4 (left) and GasLib-4-Tree (right).

TABLE 1. Overview of the results of Algorithm 1

	q_u^{nom}	k	t	obj.	abs.	rel.
GasLib-4	220	113	467.45	580.30	3.364	0.797
	230	96	413.09	613.06	5.794	0.582
GasLib-4-Tree	170	11	0.21	218.52	0.006	4×10^{-4}
	180	13	0.25	401.85	0.011	8×10^{-4}

Let us first note that we incorporate preprocessing techniques that are problem-specific bound tightenings for flow and pressure variables; see [11]. This yields both smaller initial boxes Ω_i^0 and potentially improved Lipschitz constants. Moreover, we a-priorily insert up to 20 sampling points in the case of GasLib-4-Tree and up to 150 sampling points in the case of GasLib-4; see Remark 3.5.

We consider two different kinds of termination criteria. On the one hand, we check the maximum absolute error as stated in Algorithm 1. On the other hand, we check the maximum relative error. If one of both criteria is fulfilled for all nonlinear functions, we stop at the current solution. Otherwise, if one of both criteria is violated, we use the master problem's solution to tighten its feasible set according to Line 9 of Algorithm 1. The maximum absolute error and the maximum relative error allowed is 0.1 bar and 0.001 % in the case of GasLib-4-Tree and 1 bar and 1 % in the case of GasLib-4, respectively.

Our algorithm and the model are implemented using the C++ framework LaMaTTO++; see [27]. We solve the MILPs with Gurobi 6.5.0 using all available 4 threads; see [18]. All computations were performed on an Intel© Core™i5-3360M CPU with 4 cores and 2.8 GHz each and 4 GB RAM.

Table 1 gives an overview of the results. There, the results are grouped by GasLib-4 and GasLib-4-Tree with nomination values (q_u^{nom}; in $1000\text{Nm}^3/\text{h}$), iteration numbers (k), solution times (t; in seconds), objective values (obj.; in kW), maximum absolute errors (abs.; in bar^2), and maximum relative errors (rel.). All four instances are solved to global optimality. It is obvious that iteration numbers and solution times for GasLib-4 are much higher than for GasLib-4-Tree. Figure 3 (left) shows the amount of binary variables while Figure 3 (right) shows the solution times in a semi-log plot. Note that we truncate both plots at iteration 50 and that we selected the instance with $q_u = 230$ for GasLib-4 and with $q_u = 180$ for GasLib-4-Tree. The high solution times and iteration numbers in the case of GasLib-4 compared to GasLib-4-Tree mainly result from one reason. The bound strengthening fixes the flow values in the entire network and the dimension of the domain of the nonlinearity effectively reduces to two. Due to its cycle, this is not possible for the GasLib-4 network and, thus, the domain of the nonlinearity is of higher dimension. This leads to the need for more sampling points in the second case (150 vs. 20), which significantly increases the number of initial binary variables; see Figure 3 (left). Consequently, solution times are higher as well; see Figure 3 (right). Furthermore, the additional dimension leads to a three-dimensional domain Ω_i of the nonlinearity,

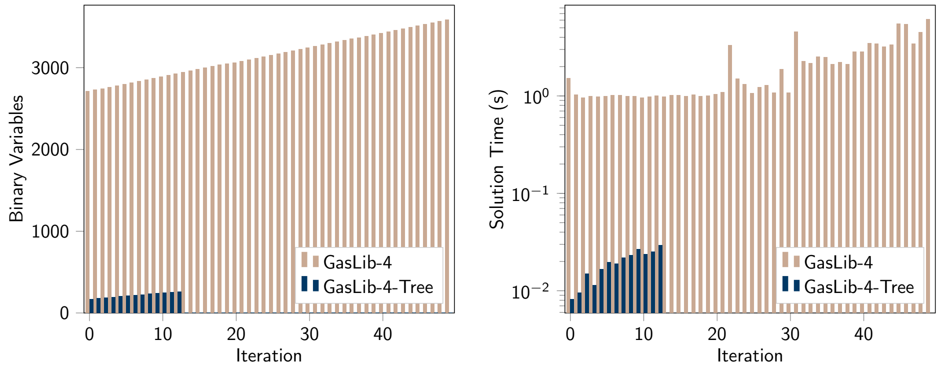

FIGURE 3. Binary variables (top) and solution times (bottom) in every iteration for GasLib-4 with nomination $q_u = 230$ and GasLib-4-Tree with nomination $q_u = 180$ up to iteration 50.

which requires significantly more boxes to obtain an accurate representation of the feasible set. This, in turn, leads to a higher iteration numbers, although the termination criteria are less strict for the GasLib-4 network.

5. Conclusion

Without any doubt, mixed-integer nonlinear optimization is hard. Fortunately, in many practically relevant situations effective solution approaches exist. In this paper, we considered especially hard cases of mixed-integer nonlinear programming with implicit constraints that have the adverse property that evaluating the constraint does not give a feasible point and that they may be multi-dimensional. The numerical results are somehow sobering—although the class of considered problems is of high importance in many fields of applications.

Thus, there is need for computational improvement for this class of problems if they are tackled using MILP technology as a working horse. At this point, we suggest three possible directions of future research.

(1) Is it possible to exploit local Lipschitz information in order to increase the volume of excluded infeasible regions during the course of the iteration?
(2) Is it possible to create iterates that both yield larger excluded regions for the next iteration but that are also near the feasible region of the original problem?
(3) Presolve and sampling seem to be of utmost computational importance. Is it possible to incorporate these techniques in the course of the iterations of the proposed algorithm?

Acknowledgments

The research of the first author has been performed as part of the Energie Campus Nürnberg and is supported by funding of the Bavarian State Government. All authors thank the DFG for their support within projects A05, A08, and B08 in CRC TRR 154. Finally, we greatly thank our colleagues Alexander Martin and Günter Leugering. A few years ago, their research vision pointed out the direction of study that we follow here and lead to the first ideas and results that paved the way for this paper.

References

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

[2] P. Bonami, L. T. Biegler, A. R. Conn, G. Cornuéjols, I. E. Grossmann, C. D. Laird, J. Lee, A. Lodi, F. Margot, N. Sawaya, and A. Wächter. "An algorithmic framework for convex mixed integer nonlinear programs." In: *Discrete Optimization* 15 (2008), pp. 186–204. DOI: 10.1016/j.disopt.2006.10.011.

[3] R. Burlacu, B. Geißler, and L. Schewe. *Solving Mixed-Integer Nonlinear Programs using Adaptively Refined Mixed-Integer Linear Programs*. Tech. rep. FAU Erlangen-Nürnberg, 2017. URL: https://opus4.kobv.de/opus4-trr154/frontdoor/index/index/docId/151.

[4] 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 (1939), pp. 133–156. DOI: 10.1680/ijoti.1939.13150.

[5] 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.

[6] 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.

[7] 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.

[8] A. Fügenschuh, B. Geißler, R. Gollmer, C. Hayn, R. Henrion, B. Hiller, J. Humpola, T. Koch, T. Lehmann, A. Martin, Radoslava Mirkov, A. Morsi, J. Rövekamp, L. Schewe, M. Schmidt, R. Schultz, R. Schwarz, J. Schweiger, C. Stangl, M. C. Steinbach, and B. M. Willert. "Mathematical optimization for challenging network planning problems in unbundled liberalized gas markets." In: *Energy Systems* 5.3 (2014), pp. 449–473. DOI: 10.1007/s12667-013-0099-8.

[9] A. Fügenschuh, B. Geißler, R. Gollmer, A. Morsi, M. E. Pfetsch, J. Rövekamp, M. Schmidt, Klaus Spreckelsen, and M. C. Steinbach. "Physical and technical fundamentals of gas networks." In: *Evaluating Gas Network Capacities*. Ed. by T. Koch, B. Hiller, M. E. Pfetsch, and L. Schewe. SIAM-MOS series on Optimization. SIAM, 2015. Chap. 2, pp. 17–44. DOI: 10.1137/1.9781611973693.ch2.

[10] M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. New York, NY, USA: W. H. Freeman & Co., 1979.

[11] B. Geißler. "Towards Globally Optimal Solutions for MINLPs by Discretization Techniques with Applications in Gas Network Optimization." PhD thesis. FAU Erlangen-Nürnberg, 2011. URL: http://www.dr.hut-verlag.de/9783843901680.html.

[12] 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.

[13] 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. Vol. 154. The IMA Volumes in Mathematics and its Applications. 2012, pp. 287–314. DOI: 10.1007/978-1-4614-1927-3_10.

[14] B. Geißler, A. Morsi, L. Schewe, and M. Schmidt. "Solving Highly Detailed Gas Transport MINLPs: Block Separability and Penalty Alternating Direction Methods." In: *INFORMS Journal on Computing* 30.2 (2018), pp. 1–15. DOI: 10.1287/ijoc.2017.0780. Forthcoming.

[15] B. Geißler, A. Morsi, L. Schewe, and M. Schmidt. "Solving power-constrained gas transportation problems using an MIP-based alternating direction method." In: *Computers & Chemical Engineering* 82 (2015), pp. 303–317. DOI: 10.1016/j.compchemeng.2015.07.005.

[16] M. Gugat, G. Leugering, A. Martin, M. Schmidt, M. M. Sirvent, and D. Wintergerst. "Towards Simulation Based Mixed-Integer Optimization with Differential Equations." In: *Networks* (2018). DOI: 10.1002/net.21812.

[17] M. Gugat, G. Leugering, A. Martin, M. Schmidt, M. Sirvent, and D. Wintergerst. "MIP-Based Instantaneous Control of Mixed-Integer PDE-Constrained Gas Transport Problems." In: *Computational Optimization and Applications* 70.1 (2018), pp. 267–294. DOI: 10.1007/s10589-017-9970-1.

[18] Gurobi. *Gurobi Optimization Inc.* URL: http://www.gurobi.com.

[19] R. Horst. "Deterministic global optimization with partition sets whose feasibility is not known: Application to concave minimization, reverse convex constraints, DC-programming, and Lipschitzian optimization." In: *Journal of Optimization Theory and Applications* 58.1 (1988), pp. 11–37. DOI: 10.1007/BF00939768.

[20] R. Horst and Ng V. Thoai. "Branch-and-bound methods for solving systems of Lipschitzian equations and inequalities." In: *Journal of Optimization Theory and Applications* 58.1 (1988), pp. 139–145. DOI: 10.1007/BF00939776.

[21] R. Horst and H. Tuy. *Global Optimization*. 2nd. Berlin: Springer Verlag, 1996. DOI: 10.1007/978-3-662-03199-5.

[22] R. Horst and H. Tuy. "On the convergence of global methods in multiextremal optimization." In: *Journal of Optimization Theory and Applications* 54.2 (1987), pp. 253–271. DOI: 10.1007/BF00939434.

[23] I. Joormann, M. Schmidt, M. C. Steinbach, and B. M. Willert. "What does "feasible" mean?" In: *Evaluating Gas Network Capacities*. Ed. by T. Koch, B. Hiller, M. E. Pfetsch, and L. Schewe. SIAM-MOS series on Optimization. SIAM, 2015. Chap. 11, pp. 211–232. DOI: 10.1137/1.9781611973693.ch11.

[24] R. Kannan and C. L. Monma. "On the Computational Complexity of Integer Programming Problems." In: *Optimization and Operations Research: Proceedings of a Workshop Held at the University of Bonn, October 2–8, 1977*. Ed. by R. Henn, B. Korte, and W. Oettli. Berlin, Heidelberg: Springer Berlin Heidelberg, 1978, pp. 161–172. DOI: 10.1007/978-3-642-95322-4_17.

[25] F. A. Al-Khayyal and H. D. Sherali. "On Finitely Terminating Branch-and-Bound Algorithms for Some Global Optimization Problems." In: *SIAM Journal on Optimization* 10.4 (2000), pp. 1049–1057. DOI: 10.1137/S105262349935178X.

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

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

[28] V. Mehrmann, M. Schmidt, and J. J. Stolwijk. *Model and Discretization Error Adaptivity within Stationary Gas Transport Optimization*. Tech. rep. Eingereicht beim Vietnam Journal of Mathematics. FAU Erlangen-Nürnberg; TU Berlin, Dec. 2017. URL: http://www.optimization-online.org/DB_HTML/2017/12/6365.html. Submitted.

[29] A. S. Nemirovsky and D. B. Yudin. *Problem complexity and method effciency in optimization*. Wiley-Interscience Series in Discrete Mathematics. New York: John Wiley & Sons, 1983.

[30] 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.

[31] J. Pintér. "Branch- and bound algorithms for solving global optimization problems with Lipschitzian structure." In: *Optimization* 19.1 (1988), pp. 101–110. DOI: 10.1080/02331938808843322.

[32] J. Pintér. "Globally convergent methods for n-dimensional multiextremal optimization." In: *Optimization* 17.2 (1986), pp. 187–202. DOI: 10.1080/02331938608843118.

[33] J. D. Pintér. *Global Optimization in Action (Continuous and Lipschitz Optimization: Algorithms, Implementations and Applications)*. Springer Verlag, 1996. DOI: 10.1007/978-1-4757-2502-5.

[34] 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.

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

[36] M. Schmidt, D. Aßmann, R. Burlacu, J. Humpola, I. Joormann, N. Kanelakis, T. Koch, D. Oucherif, M. E. Pfetsch, L. Schewe, R. Schwarz, and M. Sirvent. "GasLib—A Library of Gas Network Instances." In: *Data* 2.4 (2017). DOI: 10.3390/data2040040.

[37] M. Schmidt, M. Sirvent, and W. Wollner. *A Decomposition Method for MINLPs with Lipschitz Continuous Nonlinearities*. Tech. rep. FAU Erlangen-Nürnberg; TU Darmstadt, 2017. URL: http://www.optimization-online.org/DB_HTML/2017/07/6130.html. Submitted.

[38] 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.

[39] M. Schmidt, M. C. Steinbach, and B. M. Willert. "High detail stationary optimization models for gas networks: validation and results." In: *Optimization and Engineering* 17.2 (2016), pp. 437–472. DOI: 10.1007/s11081-015-9300-3.

[40] M. Schmidt, M. C. Steinbach, and B. M. Willert. "The precise NLP model." In: *Evaluating Gas Network Capacities*. Ed. by T. Koch, B. Hiller, M. E. Pfetsch, and L. Schewe. SIAM-MOS series on Optimization. SIAM, 2015. Chap. 10, pp. 181–210. DOI: 10.1137/1.9781611973693.ch10.

[41] M. Sirvent. "Incorporating Differential Equations into Mixed-Integer Programming for Gas Transport Optimization." PhD thesis. FAU Erlangen-Nürnberg, 2018.

[42] M. Sirvent, N. Kanelakis, B. Geißler, and P. Biskas. "Linearized model for optimization of coupled electricity and natural gas systems." In: *Journal of Modern Power Systems and Clean Energy* 5.3 (May 2017), pp. 364–374. DOI: 10.1007/s40565-017-0275-2.

[43] E. M. B. Smith and C. C. Pantelides. "Global optimisation of nonconvex MINLPs." In: *Computers & Chemical Engineering* 21 (1997), S791–S796. DOI: 10.1016/S0098-1354(97)87599-0.

[44] M. Tawarmalani and N. V. Sahinidis. *Convexification and Global Optimization in Continuous and Mixed-Integer Nonlinear Programming. Theory, Algorithms, Software, and Applications*. Vol. 65. Springer Science & Business Media, 2002. DOI: 10.1007/978-1-4757-3532-1.

[45] H. Tuy and R. Horst. "Convergence and restart in branch-and-bound algorithms for global optimization. Application to concave minimization and D.C. Optimization problems." In: *Mathematical Programming* 41.1 (1988), pp. 161–183. DOI: 10.1007/BF01580762.

[46] S. A. Vavasis. *Nonlinear Optimization: Complexity Issues*. New York, NY, USA: Oxford University Press, Inc., 1991.

(M. Schmidt) (A) Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Discrete Optimization, Cauerstr. 11, 91058 Erlangen, Germany; (B) Energie Campus Nürnberg, Fürther Str. 250, 90429 Nürnberg, Germany;
E-mail address: mar.schmidt@fau.de

(M. Sirvent) Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Discrete Optimization, Cauerstr. 11, 91058 Erlangen, Germany;
E-mail address: mathias.sirvent@fau.de

(W. Wollner) Technische Universität Darmstadt, Fachbereich Mathematik, Dolivostr. 15, 64293 Darmstadt, Germany;
E-mail address: wollner@mathematik.tu-darmstadt.de