GLOBAL OPTIMIZATION OF MIXED-INTEGER ODE CONSTRAINED NETWORK PROBLEMS USING THE EXAMPLE OF STATIONARY GAS TRANSPORT*

OLIVER HABECK†, MARC E. PFETSCH†, AND STEFAN ULBRICH†

Abstract. In this paper we propose a new approach for finding global solutions for a class of mixed-integer nonlinear optimization problems with ordinary differential equation (ODEs) constraints on networks. Instead of using a discretize-then-optimize approach, we combine spatial and variable branching with appropriate discretizations of the differential equations to derive relaxations of the original problem. To construct the relaxations we derive sufficient conditions under which appropriate numerical discretization schemes yield lower and upper bounds on the ODE solutions. Moreover, we derive conditons that ensure the convexity or concavity of the obtained under- and overestimators. Thereby, we make use of the underlying network structure, where the solutions of the ODEs only need to be known at a finite number of points, that is, at the junctions of the network. This property enables us to adaptively refine the discretization and relaxation without introducing new variables in the mixed-integer optimization problem. The incorporation into a spatial branch-and-bound process allows to compute global ε-optimal solutions or to decide infeasibility. We prove that this algorithm terminates finitely under some natural assumptions. We then show how this approach works for the example of stationary gas transport and provide some illustrative computational examples.

1. Introduction. In this paper we develop algorithms to globally solve a class of mixed-integer nonlinear optimization problems with ordinary differential equation (ODE) constraints and an underlying network structure. More precisely, we consider problems of the form

$$\begin{aligned} \min \quad & C(x, y^0, y^S, z) \\ \text{s.t.} \quad & G(x, y^0, y^S, z) \leq 0, \\ & \partial_s y(s) = f\big(s, x, y(s)\big), && s \in [0, S], \\ & y^0 = y(0),\ y^S = y(S), \\ & x \in X,\ y^0 \in Y^0,\ y^S \in Y^S,\ z \in Z, \end{aligned} \tag{1}$$

where $X \subset \mathbb{R}^k$ and $Y^0, Y^S \subset \mathbb{R}^n$ are polytopes and $Z \subset \mathbb{Z}^m$ is bounded. The objective function $C\colon X \times Y^0 \times Y^S \times Z \to \mathbb{R}$ as well as the constraints $G\colon X \times Y^0 \times Y^S \times Z \to \mathbb{R}^l$ can be nonlinear. The variables $y(s)$ are functions that solve the ODEs specified by the function $f\colon \mathbb{R} \times \mathbb{R}^k \times \mathbb{R}^n \to \mathbb{R}^n$. Moreover, continuous variables x and integer variables z are present.

The distinguishing feature of (1) is that y only needs to be known at a finite number of positions, namely 0 and S. The objective function and further constraints of (1) only depend on the corresponding values y^0 and y^S but not on the ODE solution $y(s)$ at some intermediate point $s \in (0, S)$. Note that for notational simplicity, we assume that the ODEs are defined on the same interval $[0, S]$; this can be assured by reparametrization. Moreover, we assume that C, G, and f are continuously differentiable.

*Funding: This work was supported by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) Research Grant CRC/Transregio 154 within Project A01. Moreover, S. Ulbrich was supported by the German Research Foundation within GSC 1070 Darmstadt Graduate School of Energy Science and Engineering.

†Research Group Optimization, Department of Mathematics, TU Darmstadt, Germany ({habeck,pfetsch,ulbrich}@opt.tu-darmstadt.de)

The particular structure of (1) is motivated by stationary gas or water networks. In this case, the differential equations are composed of n one-dimensional ODEs $\partial_s y_i(s) = f_i\big(s, x, y_i(s)\big)$ for $i = 1, \dots, n$, one for each of the n connections (pipelines) in the corresponding network. The relevant values are the flows x (constant along each connection) and the pressures y^0 and y^S at the nodes, which are coupled by $G(x, y^0, y^S, z) \leq 0$, representing flow conservation and further network components (compressors/pumps, resistors, valves, ...). The integer variables are used to open or close valves or to turn compressors/pumps on or off. The objective often minimizes the energy for operating compressors/pumps. We will use stationary gas transport to illustrate the approach of this paper, but it can also be applied to stationary water networks.

This paper develops a method to globally solve a class of optimization problems of the form (1). The general approach is to use branch-and-bound to handle the integer variables z and spatial branching for handling nonlinearities. Both approaches are standard in mixed-integer nonlinear programming (MINLP), see, for example the books by Horst and Tuy [19], by Lee and Leyffer [22] and Locatelli and Schoen [25] or the overview articles of Floudas and Gounaris [9] and Hemmecke et al. [16]. Spatial branching refers to the technique in which the domain of a (continuous) variable is split into two (nonempty) parts, creating two new child nodes in the branch-and-bound tree. Since the bounds on the variable are tighter in each child node, the hope is that this can be used by other solver components to further tighten bounds. This process tends to produce better relaxations in the child nodes and results in a converging solution algorithm under appropriate conditions.

The new contribution of this paper concerns the handling of the ODE constraints. We consider convex underestimators and concave overestimators for the (nonlinear) functions that map one boundary value $y(0)$ or $y(S)$ to the other. In Section 2, we show finite convergence of the corresponding method under natural assumptions. Moreover, a key point of our approach is that these estimators can be easily constructed by using basic discretization schemes (e.g., midpoint and trapezoidal rule) which produce a lower or upper bound on the ODE solution if their local truncation error is signed, i.e., either nonnegative or nonpositive, see Section 3. This can be guaranteed for particular discretization rules under convexity/concavity requirements. To demonstrate the approach, we apply the algorithm and estimation techniques to stationary gas transport in Section 4. We have also implemented our approach and provide computational experiments in Section 5.

1.1. Literature Review. The topic of this paper has clear connections to optimization with ODE or partial differential equation (PDE) constraints, see, e.g., [18] for a starting point. Due to its analytical and computational complexity, the focus in this area often lies on the computation of local optima without the consideration of discrete decisions. However, in recent years there has been some effort to approach global and discrete decisions. We review some of the literature in this direction.

A very natural approach is to use the discretize-then-optimize approach, i.e., to discretize the state space (usually time and spatial directions) in order to obtain an MINLP, which one typically tries to solve to global optimality; discrete decisions are then often handled by branch-and-bound. A number of articles use this approach – a partial list is as follows. Čižniar et al. [6] proposed a method that uses a time discretization and a polynomial basis to represent the solutions between the time points. Sager et al. [35] developed a convexification method to handle specific discrete decisions over time that switch the right-hand sides of the differential equations (e.g., gear

shifting) and show how to efficiently compute feasible solutions; if the corresponding continuous relaxation is solved to global optimality, then such solutions converge to a global optimal solution while refining the discretization. Extending this approach, Sager et al. [36] and Jung et al. [20], develop a solution algorithm for the so-called combinatorial integral approximation problem. An open source implementation of a general discretize-then-optimize approach is available at [48], which uses relaxations based on piecewise-linearization, see Fügenschuh and Vierhaus [10] for a description. Bock et al. [3] consider problems with implicit and explicit switches. Based on discretize-then-optimize, they provide a reformulation as a nonlinear program with vanishing constraints, which they solve numerically. Using the αBB approach (introduced by Adjiman et al. [2, 1]), Diedam and Sager [7] develop a method to globally solve the nonlinear programs (NLPs) arising from a multiple-shooting discretization for optimal control problems without integer decisions.

All these discretize-then-optimize approaches use a fixed discretization, i.e., the solutions only provide an approximation of the solutions of the ODEs with respect to an a priori fixed accuracy. Thus, the discretization error is ignored or it is implicitly assumed that the discretization is refined if an a posteriori accuracy check fails. Note that the corresponding MINLPs become very large for high precision.

Esposito and Floudas [8] developed an approach based on a fixed discretization of the control and αBB relaxations of the solution operator (control $\to$ ODE solution).

Several publications approached the goal to also control the discretization error. For instance, Nedialkov et al. [27] review methods for enclosing solutions of initial value problems using interval arithmetic. Taylor approximations in combination with interval arithmetic for parametric ODEs have been developed by Neher et al. [28], Lin et al. [23, 24], Sahlodin and Chachuat [37, 38] and Villanueva et al. [49]. These methods can be used to perform bound propagation, but to the best of our knowledge have not been used in a global optimization approach.

We now review global optimization approaches. Papamichail and Adjiman [31, 32] consider parametric ODEs and construct approximations via the αBB approach. They propose an NLP-based spatial branch-and-bound algorithm in which ODEs have to be solved within the solution of the NLPs. Singer and Barton [46] construct lower and upper bounds of parametric ODE solutions as follows: using a linearization of f at a solution for a particular (continuous) parameter, two new right-hand sides are constructed with the property that solutions of this auxiliary ODE system provide lower and upper bounds on the original solution for all parameters and the solutions are concave, respectively, convex in the parameters. Using these bounds, Singer and Barton [47] developed a global optimization approach for ODE constrained optimization problems by spatial branching on the parameters. Chachuat et al. [5] use an outer-approximation algorithm to extend this work to the mixed-integer case. The construction of lower and upper bounds has been improved by using McCormick relaxations by Scott et al. [45] and Scott and Barton [43]. Further applications can be found in Scott et al. [44] for spatial branch-and-bound and Scott and Barton [42] for differential algebraic equations. Note that all these approaches require the exact solution of ODEs, but do not explicitly have to add variables corresponding to the discretization, i.e., they are "sequential" in the local dynamics terminology.

Buchheim et al. [4] present a global approach for solving particular semilinear elliptic mixed-integer PDE problems with distributed and boundary control using outer approximation. In [15], Hante and Sager extend the convexification approach of [35] to mixed-integer PDE problems and derive a relaxation.

As mentioned earlier, we will use stationary gas transport as example in this arti-

cle. We refer to [33, 21], Ríos-Mercado and Borraz-Sánchez [34], Hante et al. [14] for general information on modeling of and solution methods for gas transport problems.

A related approach for solving mathematical optimization problems with ODEs in the context of gas transport is described in Gugat et al. [13]. Here, a global decomposition approach is described if the underlying network is a tree. Since in every iteration a mixed-integer linear master optimization problem is solved, this amounts to a "multi-tree" approach, while the method described in this paper works as a "single-tree". Related is the approach of Schmidt et al. [40], who consider the solution of MINLPs with equality constraints using univariate Lipschitz continuous functions for which the constants are known or approximated and the function evaluations may be approximate; this is applied to a stationary gas transport problem in which the underlying network is a tree. Moreover, Gugat et al. [12] present an instantaneous control approach for solving instationary gas transport problems, where a mixed-integer linear problem needs to be solved for each time step.

The technique that we present in this paper is distinct from the approaches mentioned above in the following way: We adaptively refine the discretization, which is not done in the approaches based on discretize-then-optimize. Moreover, our method to derive lower and upper bounds exploits the particular network and ODE structure and is different from the general-purpose approximations for ODEs and the convexifications mentioned above.

2. Solution Method. In this section we introduce our solution method. We begin with a natural basic assumption on the differential equations, which we assume to hold throughout the paper.

ASSUMPTION 1. *The initial value problem*

$$y(0) = y^0, \quad \partial_s y(s) = f\big(s, x, y(s)\big), \quad s \in [0, S] \tag{2}$$

is uniquely solvable for all $x \in X,\ y^0 \in Y^0$.

This assumption can be guaranteed, for example, if f is Lipschitz continuous w.r.t. y. We then denote the solution operator by

$$F\colon X \times Y^0 \to \mathbb{R}^n, \quad \big(x, y^0\big) \mapsto y(S),$$

the unique solution of the initial value problem (2) for every $x \in X$ and $y^0 \in Y^0$. Replacing the ODE constraints by $y^S - F\big(x, y^0\big) = 0$ yields the equivalent problem

$$\begin{aligned} \min \quad & C(x, y^0, y^S, z) \\ \text{s.t.} \quad & G(x, y^0, y^S, z) \le 0, \\ & y^S - F\big(x, y^0\big) = 0, \\ & x \in X,\ y^0 \in Y^0,\ y^S \in Y^S,\ z \in Z. \end{aligned} \tag{3}$$

Since X, Y^0, Y^S are polytopes, Z is bounded and C is continuous, the problem has an optimal solution if the feasible set is nonempty. If there is an analytical formula for F, then we could in principle use spatial branch-and-bound to solve (3). In the following, we assume that this is not the case or that the formula is hard to evaluate.

Our idea is to construct under- and overestimators of $F(x, y^0)$ by the right choice of suitable numerical methods. That is, we choose one-step methods which provably yield lower and upper bounds for $y(s)$, respectively. We will later see an example for this approach. For now, we assume the existence of under- and overestimators.

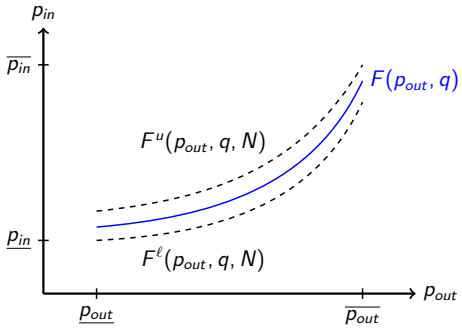

FIGURE 1. *Example from stationary gas transport, see Section 4. Here, for one pipe, the inflow pressure is a convex function $F(p_{out}, q)$ of the outflow pressure p_{out} and mass flow q (solid line, shown for fixed flow), and we can define F^ℓ and F^u (dashed lines) by suitable discretization methods.*

ASSUMPTION 2. *There exist functions $F^\ell \colon X \times Y^0 \times \mathbb{N}^n \to \mathbb{R}^n$ and $F^u \colon X \times Y^0 \times \mathbb{N}^n \to \mathbb{R}^n$, which fulfill the inequality*

$$F^\ell\big(x, y^0, N\big) \le F\big(x, y^0\big) \le F^u\big(x, y^0, N\big)$$

for all $x \in X$ and $y^0 \in Y^0$. Furthermore, we assume that on the polytopes X, Y^0 the functions F_i^ℓ and F_i^u converge uniformly to F_i for $N_i \to \infty$, $i = 1, \dots, n$.

For an example of this assumption see Figure 1. This figure illustrates our application to gas transport, where the stationary isothermal Euler equation defines the relation between the pressure at the ends of a pipe and the flow. Given the pressure at the end and the mass flow, one can compute lower and upper bounds on the pressure at the start. Thereby, N corresponds to the number of grid points in the discretization.

We then relax the constraint $y^S = F(x, y^0)$ of the problem (3) by means of the functions F^ℓ and F^u. In this way we derive the relaxation

$$\begin{aligned} \min \quad & C(x, y^0, y^S, z) \\ \text{s.t.} \quad & G(x, y^0, y^S, z) \le 0, \\ & F^\ell\big(x, y^0, N\big) \le y^S \le F^u\big(x, y^0, N\big), \\ & x \in X,\ y^0 \in Y^0,\ y^S \in Y^S,\ z \in Z. \end{aligned} \tag{4}$$

This is a relaxation of (3), since every feasible point of (3) is feasible for the new constraint

$$F^\ell\big(x, y^0, N\big) \le y^S \le F^u\big(x, y^0, N\big),$$

and the objective function is the same. Note that this constraint and thus (4) depends on $N \in \mathbb{N}^n$. Again, the optimal value of this problem is bounded from below, because X, Y^0, Y^S are polytopes, Z is bounded, and C is continuous.

In order to solve (4) with spatial branch-and-bound, we need a convex relaxation of the feasible set. Thus, we suppose that we can construct a convex underestimator $\check{F}^\ell$ of F^ℓ and a concave overestimator $\hat{F}^u$ of F^u for every $N \in \mathbb{N}^n$. In addition, let $\check{G}$ and $\check{C}$ be convex underestimators of G and C, respectively, for example, obtained by the αBB approach [2, 1] or McCormick relaxations [26]. Then we obtain the following

convex relaxation of (4):

$$
\begin{aligned}
\min\quad & \alpha \\
\text{s.t.}\quad & \check{C}(x, y^0, y^S, z) - \alpha \le 0, \\
& \check{G}(x, y^0, y^S, z) \le 0, \\
& \check{F}^\ell(x, y^0, N) \le y^S \le \hat{F}^u(x, y^0, N), \\
& x \in X,\ y^0 \in Y^0,\ y^S \in Y^S,\ z \in \operatorname{conv}(Z).
\end{aligned} \tag{5}
$$

Spatial branch-and-bound will enable us to compute so called (ε, δ)-*optimal solutions* of the relaxation (4). For a vector $y \in \mathbb{R}^n$ we denote with $(y)_+$ the vector of the componentwise maxima of y_i and 0.

DEFINITION 2.1. *We say that a vector $(x, y^0, y^S, z) \in X \times Y^0 \times Y^S \times Z$ is a* δ-feasible solution *of (3) if the following condition holds:*

$$
\max\left\{ \left\| \left(G(x, y^0, y^S, z)\right)_+ \right\|_\infty,\ \left\| y^S - F(x, y^0) \right\|_\infty \right\} \le \delta.
$$

Analogously, we call $(x, y^0, y^S, z) \in X \times Y^0 \times Y^S \times Z$ a δ-feasible solution *of (4) if*

$$
\begin{aligned}
\max\{ \left\| \left(G\left(x, y^0, y^S, z\right)\right)_+ \right\|_\infty,\ & \left\| \left(F^\ell(x, y^0, N) - y^S\right)_+ \right\|_\infty, \\
& \left\| \left(y^S - F^u\left(x, y^0, N\right)\right)_+ \right\|_\infty \} \le \delta
\end{aligned}
$$

holds. Furthermore, we call $(x, y^0, y^S, z) \in X \times Y^0 \times Y^S \times Z$ an (ε, δ)-optimal solution *of (3) or (4) if it is δ-feasible and the objective function satisfies $C(x, y^0, y^S, z) \le C^* + \varepsilon$, where $C^* > -\infty$ is the optimal value of (3) or (4), or $C^* = \infty$ if their respective feasible set is empty.*

Note that this definition is consistent with the definition in the literature, e.g., Locatelli and Schoen [25]. Since our goal is to find (ε, δ)-optimal solutions of (3) by approximatively solving (4), we now show how their respective (ε, δ)-optimal solutions are related.

LEMMA 2.2. *Let $(x, y^0, y^S, z) \in X \times Y^0 \times Y^S \times Z$ be an (ε, δ_1)-optimal solution of (4) for some $N \in \mathbb{N}^n$. Additionally, for $\delta_2 \ge 0$, let the condition*

$$
\left\| F^u(x, y^0, N) - F^\ell(x, y^0, N) \right\|_\infty \le \delta_2 \tag{6}
$$

be satisfied. Then (x, y^0, y^S, z) is an (ε, δ)-optimal solution of (3) for all $\delta \ge \delta_1 + \delta_2$.

Proof. First, we prove that (x, y^0, y^S, z) is δ-feasible for (3). Because of the δ_1-feasibility for (4), we know that $\left\| \left(G(x, y^0, y^S, z)\right)_+ \right\|_\infty \le \delta_1 \le \delta$ as well as

$$
\left(y^S - F^u(x, y^0, N)\right)_+ + \left(F^\ell(x, y^0, N) - y^S\right)_+ \le \delta_1.
$$

Here, we used that $F^u \ge F^\ell$ holds and for all $i = 1, \dots, n$ only one of $y_i^S - F_i^u\left(x, y^0, N\right)$

and $F_i^\ell(x, y^0, N) - y_i^S$ can be positive. Thus,

$$\begin{aligned}
\left|y_i^S - F_i(x,y^0)\right| &= \left(y_i^S - F_i(x,y^0)\right)_+ + \left(F_i(x,y^0) - y_i^S\right)_+ \\
&\le \left(y_i^S - F_i^u(x,y^0,N)\right)_+ + \left(F_i^u(x,y^0,N) - F_i(x,y^0)\right)_+ \\
&\quad + \left(F_i(x,y^0) - F_i^\ell(x,y^0,N)\right)_+ + \left(F_i^\ell(x,y^0,N) - y_i^S\right)_+ \\
&= \left(y_i^S - F_i^u(x,y^0,N)\right)_+ + F_i^u(x,y^0,N) \\
&\quad - F_i^\ell(x,y^0,N) + \left(F_i^\ell(x,y^0,N) - y_i^S\right)_+ \le \delta_1 + \delta_2 \le \delta.
\end{aligned}$$

That is, (x, y^0, y^S, z) is δ-feasible for (3). Let $(\bar{x}, \bar{y}^0, \bar{y}^S, \bar{z})$ be an optimal solution of (3). As (4) is a relaxation of (3), this solution is feasible for (4). Hence, there exists an optimal solution $(\tilde{x}, \tilde{y}^0, \tilde{y}^S, \tilde{z})$ of (4) with $C(\tilde{x}, \tilde{y}^0, \tilde{y}^S, \tilde{z}) \le C(\bar{x}, \bar{y}^0, \bar{y}^S, \bar{z})$. Since (x, y^0, y^S, z) is an (ε, δ_1)-optimal solution of the relaxation (4), we can derive

$$C(x, y^0, y^S, z) \le C(\tilde{x}, \tilde{y}^0, \tilde{y}^S, \tilde{z}) + \varepsilon \le C(\bar{x}, \bar{y}^0, \bar{y}^S, \bar{z}) + \varepsilon,$$

that is, (x, y^0, y^S, z) is an (ε, δ)-optimal solution of (3). Otherwise, if (3) is infeasible, the condition $C(x, y^0, y^S, z) \le C^* + \varepsilon = \infty$ is obviously satisfied. ∎

Algorithm 1 Spatial branch-and-bound for (4)

Input: Problem (4), N, $\delta > 0$ and $\varepsilon > 0$.
Output: (ε, δ)-optimal solution $(\bar{x}, \bar{y}^0, \bar{y}^S, \bar{z})$ of (4) or "infeasible".

```
Upper bound 𝒰 ← ∞
List of active nodes ℒ ← {X × Y^0 × Y^S × Z}
While ℒ ≠ ∅ do
    choose a node X̃ × Ỹ^0 × Ỹ^S × Z̃ ∈ ℒ and set ℒ ← ℒ \ {X̃ × Ỹ^0 × Ỹ^S × Z̃}.
    Build the convex relaxation (5) w.r.t. X̃ × Ỹ^0 × Ỹ^S × Z̃.
    If (5) is feasible then
        let (α̃, x̃, ỹ^0, ỹ^S, z̃) be an optimal solution of (5).
        If z̃ ∈ ℤ^m then
            If (x̃, ỹ^0, ỹ^S, z̃) is δ-feasible for (4) and C(x̃, ỹ^0, ỹ^S, z̃) < 𝒰 then
                set 𝒰 ← C(x̃, ỹ^0, ỹ^S, z̃) and (x̄, ȳ^0, ȳ^S, z̄) ← (x̃, ỹ^0, ỹ^S, z̃).
        If α̃ < 𝒰 − ε then
            choose z̃_i ∉ ℤ or x̃_i, ỹ_i^0, ỹ_i^S appearing in a δ-violated constraint or in the
            ε-violated constraint C(x̃, ỹ^0, ỹ^S, z̃) − α̃ > ε.
            Branch w.r.t. the variable x̃_i, ỹ_i^0, ỹ_i^S or z̃_i and add nodes to ℒ.
```

Lemma 2.2 shows how to generate an (ε, δ)-optimal solution of (3). Since the functions F^ℓ and F^u uniformly converge to F, we can choose N such that condition (6) is satisfied for all $(x, y^0) \in X \times Y^0$. If the under- and overestimators fulfill certain technical conditions, spatial branch-and-bound can compute an (ε, δ_1)-optimal solution of (4) and therefore an (ε, δ)-optimal solution of (3). This idea yields Algorithm 1.

For proving that Algorithm 1 terminates, we require the following conditions. Suppose the algorithm produces (through branching) an infinite nested sequence of nodes $\mathcal{F}_k = X_k \times Y_k^0 \times Y_k^S \times Z_k$ with $\mathcal{F}_{k+1} \subseteq \mathcal{F}_k$ for all $k \in \mathbb{N}_0$. Then the branching rules have to satisfy the condition

$$\lim_{k\to\infty} \operatorname{diam}(\mathcal{F}_k) = 0, \tag{7}$$

where diam is the diameter of a set U, i.e., $\operatorname{diam}(U) \coloneqq \max_{u,u' \in U} \|u - u'\|_2$. Then for every $\mathcal{F}_k$ we need to be able to construct the convex underestimators $\check{C}$, $\check{G}$, $\check{F}^\ell$ and the concave overestimator $\hat{F}^u$ over $\mathcal{F}_k$. We denote the dependency on $\mathcal{F}_k$ by the index k, e.g., $\check{C}_k$. Furthermore, if (7) holds, for $k \to \infty$ the estimators have to satisfy

$$\max_{(x,y^0,y^S,z)\in\mathcal{F}_k} \Big\{ \|G(x,y^0,y^S,z) - \check{G}_k(x,y^0,y^S,z)\|_\infty,\ \|F^\ell(x,y^0,N) - \check{F}^\ell_k(x,y^0,N)\|_\infty, \\ \|\hat{F}^u_k(x,y^0,N) - F^u(x,y^0,N)\|_\infty,\ |C(x,y^0,y^S,z) - \check{C}_k(x,y^0,y^S,z)| \Big\} \to 0. \tag{8}$$

Under these conditions, Algorithm 1 terminates finitely, which follows from Locatelli and Schoen [25, Theorem 5.26]; here, we assume for simplicity that the relaxations can be evaluated exactly, i.e., without rounding errors, otherwise a further approximation error would have to be handled.

THEOREM 2.3. *Let $\varepsilon > 0$, $\delta > 0$ hold. Suppose that the Conditions (7) and (8) are satisfied. Then Algorithm 1 terminates after a finite number of iterations and either returns an (ε,δ)-optimal solution of (4) or that (4) is infeasible.*

Note that there can exist (ε,δ)-optimal solutions even if (4) is infeasible. In this case, both results are possible. It can happen that Algorithm 1 finds an (ε,δ)-optimal solution or that δ-feasible solutions of (4) are infeasible for (5) and the algorithm returns "infeasible". This is due to the fact that under- and overestimators are usually tight at some points and cut off δ-feasible solutions. For example, the McCormick estimators for the product of two variables over a square are exact in the corners.

Choosing N big enough, we can now compute (ε,δ)-optimal solutions of (3).

COROLLARY 2.4. *Let $\varepsilon > 0$, $\delta > 0$, and suppose that Conditions (7) and (8) hold. Then we can compute an (ε,δ)-optimal solution of (3) in finite time, or establish the infeasibility of (3).*

Proof. By Assumption 2 we know that F^ℓ and F^u converge uniformly to F for $N_i \to \infty$, $i = 1,\dots,n$. Therefore, we can choose N such that

$$\big\|F^u(x,y^0,N) - F^\ell(x,y^0,N)\big\|_\infty \le \frac{\delta}{2}$$

holds for all $x \in X$, $Y^0 \in Y^0$. By Theorem 2.3 the spatial branch-and-bound algorithm with parameters $\frac{\delta}{2} > 0$ and $\varepsilon > 0$ returns an $(\varepsilon, \frac{\delta}{2})$-optimal solution of (4) or that it is infeasible.

Since (4) is a relaxation of (3), if the algorithm returns "infeasible", there is no feasible solution of (3). Otherwise, the algorithm returns an $(\varepsilon, \frac{\delta}{2})$-optimal solution of (4) and Lemma 2.2 states that this solution is an (ε,δ)-optimal solution of (3). ∎

2.1. Adaptive Spatial Branch-and-Bound. A disadvantage of Algorithm 1 is that we have to choose N in advance, such that Condition (6) holds on the whole feasible set. Hence, we might have to select N larger than necessary. This leads to more computational effort, for example, when N corresponds to the number of grid points of a discretization. To circumvent this problem, we replace Line 5 of Algorithm 1 by the following adaptive procedure, see Algorithm 2.

We start with constructing $\tilde{C}$ and $\tilde{G}$ on the current node by standard methods, e.g., McCormick [26] or αBB relaxations, see Adjiman et al. [2, 1]. We then pick some initial convex relaxation of F^ℓ and F^u, e.g., the relaxation of the parent node during the branch-and-bound process or $X \times Y^0 \times Y^S \times \operatorname{conv}(Z)$ in the root node. Then we solve the convex relaxation. If the relaxation is infeasible, so is the corresponding

original problem and we are done. Otherwise, let $(\tilde{\alpha}, \tilde{x}, \tilde{y}^0, \tilde{y}^S, \tilde{z})$ be the solution of the relaxation. We then possibly increase N until (6) holds in $(\tilde{x}, \tilde{y}^0)$. Note that we do not need to solve the relaxation again, since we do not update the relaxation when increasing N. If $(\tilde{x}, \tilde{y}^0, \tilde{y}^S, \tilde{z})$ is a δ_1-feasible solution of

$$F^\ell(\tilde{x}, \tilde{y}^0, N) \leq \tilde{y}^S \leq F^u(\tilde{x}, \tilde{y}^0, N), \tag{9}$$

we stop with the current solution, otherwise we pick the most violated constraint, see Line 10 of Algorithm 2. Then either $\tilde{y}_i^S > F_i^u(\tilde{x}, \tilde{y}^0, N) + \delta_1$, or $\tilde{y}_i^S < F_i^\ell(\tilde{x}, \tilde{y}^0, N) - \delta_1$ holds. Subsequently, we improve the under- or overestimator by cutting off the current solution. If this is not possible, we stop with the current solution and have to perform branching to resolve the infeasibility.

Algorithm 2 Adaptive convex relaxation

Input: Node of the branch-and-bound tree $\tilde{X} \times \tilde{Y}^0 \times \tilde{Y}^S \times \tilde{Z}$, $\delta_1, \delta_2 > 0$, $N \in \mathbb{N}^n$, and convex underestimators $\check{C}$, $\check{G}$.

Output: A δ_1-feasible solution of (9) satisfying (6), "infeasible" or "branch".

1: Choose relaxation of (9), e.g., $\tilde{X} \times \tilde{Y}^0 \times \tilde{Y}^S \times \tilde{Z}$ or relaxation of parent node.
2: **For** $k = 1, 2, \ldots$ **do**
3: **If** (5) is feasible **then**
4: Let $(\tilde{\alpha}, \tilde{x}, \tilde{y}^0, \tilde{y}^S, \tilde{z})$ be a solution of the relaxation.
5: **While** $\left\|F^u(\tilde{x}, \tilde{y}^0, N) - F^\ell(\tilde{x}, \tilde{y}^0, N)\right\|_\infty > \delta_2$ **do**
6: increase N_i for all i with $\left|F_i^u(\tilde{x}, \tilde{y}^0, N_i) - F_i^\ell(\tilde{x}, \tilde{y}^0, N_i)\right| > \delta_2$.
7: **If** $(\tilde{x}, \tilde{y}^0, \tilde{y}^S, \tilde{z})$ is δ_1-feasible for (9) **then**
8: **return** solution $(\tilde{\alpha}, \tilde{x}, \tilde{y}^0, \tilde{y}^S, \tilde{z})$
9: **else**
10: choose "most violated" constraint i, i.e.,
 $i \in \arg\max_{j=1,\ldots,n} \max\left\{F_j^\ell(\tilde{x}, \tilde{y}^0, N_j) - \tilde{y}_j^S,\ \tilde{y}_j^S - F_j^u(\tilde{x}, \tilde{y}^0, N_j)\right\}$.
11: **If** $\tilde{y}_i^S > F_i^u(\tilde{x}, \tilde{y}^0, N_i) + \delta_1$ **then**
12: "improve overestimator" or return "branch"
13: **else if** $\tilde{y}_i^S < F_i^\ell(\tilde{x}, \tilde{y}^0, N_i) - \delta_1$ **then**
14: "improve underestimator" or return "branch"
15: **else**
16: **return** "infeasible".

Lines 12 and 14 of Algorithm 2 require to improve the estimators, which depends on the particular problem. For example, one could add a linear inequality, which cuts off the current solution of the relaxation and is feasible for (4). This is possible in (convex) outer-approximation, where current solutions can be cut off by gradient cuts. Another possibility is to add an estimator dynamically, instead of adding all inequalities at once. Thus, if an over- or underestimator consists of multiple inequalities, we only add an inequality if it cuts off the current solution.

Incorporating Algorithm 2 into the spatial branch-and-bound algorithm results in Algorithm 3. The main change is of course that N need not be constant any more. Note that a δ_1-feasible solution of (4) for N might not be a δ_1-feasible solution of (4) for $N' \geq N$, but still is a $(\delta_1 + \delta_2)$-feasible solution of (3) if it fulfills Condition (6) for N. Therefore, Algorithm 3 solves (3) and not (4). Another big difference is that we do not have to reconstruct $\check{F}^\ell$ and $\hat{F}^u$ in every node, but instead refine them only if needed. Besides this, the algorithm is almost the same as before.

Algorithm 3 Adaptive spatial branch-and-bound for (3)

Input: Problem (3), $N = N_0 \in \mathbb{N}^n$, $\delta_1, \delta_2 > 0$ and $\varepsilon > 0$.
Output: $(\varepsilon, \delta_1 + \delta_2)$-optimal solution $\big(\bar{x}, \bar{y}^0, \bar{y}^S, \bar{z}\big)$ or "infeasible".

1: Upper bound $\mathcal{U} \leftarrow \infty$
2: List of active nodes $\mathcal{L} \leftarrow \big\{X \times Y^0 \times Y^S \times Z\big\}$
3: **While** $\mathcal{L} \neq \emptyset$ **do**
4: choose a node $\tilde{X} \times \tilde{Y}^0 \times \tilde{Y}^S \times \tilde{Z} \in \mathcal{L}$ and set $\mathcal{L} \leftarrow \mathcal{L} \setminus \{\tilde{X} \times \tilde{Y}^0 \times \tilde{Y}^S \times \tilde{Z}\}$.
5: Construct underestimators $\check{C}$, $\check{G}$.
6: Run Algorithm 2.
7: **If** Algorithm 2 stops with a solution of the relaxation **then**
8: let $\big(\tilde{\alpha}, \tilde{x}, \tilde{y}^0, \tilde{y}^S, \tilde{z}\big)$ be the solution.
9: **If** the solution is δ_1-feasible for (9) **then**
10: **If** $\big(\tilde{x}, \tilde{y}^0, \tilde{y}^S, \tilde{z}\big)$ is δ_1-feasible for (4) and $C\big(\tilde{x}, \tilde{y}^0, \tilde{y}^S, \tilde{z}\big) < \mathcal{U}$ **then**
11: set $\mathcal{U} \leftarrow C\big(\tilde{x}, \tilde{y}^0, \tilde{y}^S, \tilde{z}\big)$ and $\big(\bar{x}, \bar{y}^0, \bar{y}^S, \bar{z}\big) \leftarrow \big(\tilde{x}, \tilde{y}^0, \tilde{y}^S, \tilde{z}\big)$.
12: **If** $\tilde{\alpha} < \mathcal{U} - \varepsilon$ **then**
13: choose $\tilde{z}_i \notin \mathbb{Z}$ or $\tilde{x}_i, \tilde{y}^0_i, \tilde{y}^S_i$ appearing in a δ_1-violated constraint $G \leq 0$ or in the possibly ε-violated constraint $C\big(\tilde{x}, \tilde{y}^0, \tilde{y}^S, \tilde{z}\big) - \tilde{\alpha} > \varepsilon$ or in the "most violated" constraint chosen in the last iteration of Algorithm 2.
14: Branch w.r.t. the chosen variable $\tilde{x}_i$, $\tilde{y}^0_i$, $\tilde{y}^S_i$ or $\tilde{z}_i$ and add nodes to $\mathcal{L}$.

Note that Line 10 contains a hidden integrality check for $\tilde{z}$, since by definition the solution of the relaxation is δ_1-feasible for (4) if and only if $\tilde{z}$ is integral. Furthermore, if Algorithm 2 cannot resolve infeasibility by improving an under- or overestimator and returns "branch", we can in Line 13 still choose to perform branching with respect to some integral variable or due to another violated constraint first.

The crucial point for showing that Algorithm 3 terminates, is that Algorithm 2 terminates after a finite number of iterations. As this cannot be proved in general, but only for a given construction method of $\check{F}^\ell$ and $\hat{F}^u$, we need the following assumption.

ASSUMPTION 3. *If Algorithm 2 keeps N fixed then it terminates after finitely many iterations.*

Note that we do not suppose that the algorithm stops with a δ_1-feasible solution, it only has to stop with either a solution, "infeasible" or "branch". The next Lemma shows that this assumption is enough such that Algorithm 2 terminates after finitely many iterations.

LEMMA 2.5. *If Assumption 3 holds, then Algorithm 2 terminates finitely.*

Proof. Assume that the algorithm does not terminate. Then it produces a sequence of points which are feasible solutions of the convex relaxation but not δ_1-feasible for (9). We denote with $K \subset \mathbb{N}$ the iterations where N has to be increased.

Since F^ℓ and F^u converge uniformly to F w.r.t. N and $\tilde{X} \times \tilde{Y}^0$ is bounded, there exists an $N_0 \in \mathbb{N}^n$ such that

$$\big\|F^u\big(x, y^0, N\big) - F^\ell\big(x, y^0, N\big)\big\|_\infty \leq \delta_2$$

is satisfied for all $(x, y^0) \in \tilde{X} \times \tilde{Y}^0$ and all $N \geq N^0$. That is, each N_i can only be increased a finite number of times until $N_i \geq N^0_i$ holds. Hence, K is either empty or a finite set. Then N is fixed either from the beginning or after the last iteration $k \in K$. Due to Assumption 3 the algorithm stops after another finite number of iterations. ∎

We can now prove that Algorithm 3 terminates finitely. Again, we consider an infinite nested sequence of nodes $\mathcal{F}_k = X_k \times Y_k^0 \times Y_k^S \times Z_k$ with $\mathcal{F}_{k+1} \subseteq \mathcal{F}_k$ and corresponding grid sizes N^k for all $k \geq 0$ produced by Algorithm 3. The branching rules still have to satisfy the condition $\lim_{k\to\infty} \operatorname{diam}(\mathcal{F}_k) = 0$. Since Algorithm 2 only improves the estimators if (9) is δ_1-violated in the current relaxation solution, it might happen that an estimator $\hat{F}^u$ or $\check{F}^\ell$ does not change although

$$\max_{(x,y^0,y^S,z)\in\mathcal{F}_k} \Big\{ \|F^\ell(x,y^0,N^k) - \check{F}_k^\ell(x,y^0,N^k)\|_\infty, \\ \|\hat{F}_k^u(x,y^0,N^k) - F^u(x,y^0,N^k)\|_\infty \Big\} > \delta_1$$

holds. Thus, (8) cannot hold either. Instead, if (7) holds, the following has to hold for convex underestimators of C and G for $k \to \infty$:

$$\max_{(x,y^0,y^S,z)\in\mathcal{F}_k} \Big\{ \|G(x,y^0,y^S,z) - \check{G}_k(x,y^0,y^S,z)\|_\infty, \\ |C(x,y^0,y^S,z) - \check{C}_k(x,y^0,y^S,z)| \Big\} \to 0. \tag{10}$$

We then assume that for every sequence of nodes, which satisfies (7), there exists an iteration $k_0 \in \mathbb{N}$ such that the optimal solutions $(\tilde{\alpha}^k, \tilde{x}^k, \tilde{y}^{0,k}, \tilde{y}^{S,k}, \tilde{z}^k)_{k\in\mathbb{N}}$ of the relaxation (5) over $\mathcal{F}_k$ satisfy

$$\max \Big\{ \big\| \big(F^\ell(\tilde{x}^k, \tilde{y}^{0,k}, N^k) - \tilde{y}^{S,k}\big)_+ \big\|_\infty, \big\| \big(\tilde{y}^{S,k} - F^u(\tilde{x}^k, \tilde{y}^{0,k}, N^k)\big)_+ \big\|_\infty \Big\} \leq \delta_1 \tag{11}$$

for $k \geq k_0$. We now have all the conditions and assumptions we need.

THEOREM 2.6. *Suppose that Conditions (7), (10) and (11), and the Assumptions 1, 2 and 3 hold. Then Algorithm 3 terminates with an $(\varepsilon, \delta_1 + \delta_2)$-optimal solution of (3) or "infeasible" after a finite number of nodes.*

Proof. Suppose that Algorithm 3 does not terminate. Then it produces at least one infinite nested sequence of nodes $\mathcal{F}_k = X_k \times Y_k^0 \times Y_k^S \times Z_k$ and a sequence $\big(\tilde{\alpha}^k, \tilde{x}^k, \tilde{y}^{0,k}, \tilde{y}^{S,k}, \tilde{z}^k\big)_{k\in\mathbb{N}}$, where each element is the solution of (5) over $\mathcal{F}_k$ w.r.t. N^k during the last iteration of Algorithm 2. Note that the relaxation has to be feasible for every node, otherwise, the node would be pruned and the sequence $\mathcal{F}_k$ ends finitely.

We show that there exists a $K \in \mathbb{N}$ such that $\big(\tilde{\alpha}^K, \tilde{x}^K, \tilde{y}^{0,K}, \tilde{y}^{S,K}, \tilde{z}^K\big)$ is a $(\delta_1 + \delta_2)$-feasible solution of (3). By the Conditions (7) and (10) there exists an iteration $k_0 \in \mathbb{N}$ such that $|C(x,y^0,y^S,z) - \check{C}_k(x,y^0,y^S,z)| < \varepsilon$ and $\|G(x,y^0,y^S,z) - \check{G}_k(x,y^0,y^S,z)\|_\infty < \delta_1$ holds for all $\big(x,y^0,y^S,z\big) \in \mathcal{F}_k$ and all nodes $k \geq k_0$.

For the ODE-relaxation we again notice that there is a $N_0 \in \mathbb{N}^n$ such that Condition (6) is satisfied on the whole domain $X \times Y^0 \times Y^S \times Z$ and all $N \geq N_0$, because of the assumption that F^ℓ and F^u converge uniformly to F with respect to N. Therefore, at some iteration $k_1 \in \mathbb{N}$, N is increased for the last time. Then after $\max\{k_0, k_1+1\}$ nodes the only constraint which can be violated is

$$F^\ell\big(x,y^0,N\big) - \delta_1 \leq y^S \leq F^u\big(x,y^0,N\big) + \delta_1.$$

But by Condition (11), there is an iteration $k_2 \in \mathbb{N}$ such that this condition holds for all solutions $\{(\tilde{\alpha}^k, \tilde{x}^k, \tilde{y}^{0,k}, \tilde{y}^{S,k}, \tilde{z}^k)\}_{k\geq k_2}$ of (5) produced by Algorithm 2. Hence, $(\tilde{x}^K, \tilde{y}^{0,K}, \tilde{y}^{S,K}, \tilde{z}^K)$ with $K = \max\{k_0, k_1, k_2\}$ is a $(\delta_1 + \delta_2)$-feasible solution of (3). Thus, the upper bound $\mathcal{U}$ will be updated if $C\big(\tilde{x}^K, \tilde{y}^{0,K}, \tilde{y}^{S,K}, \tilde{z}^K\big) < \mathcal{U}$ holds and the node $\mathcal{F}_K$ will get fathomed, because

$$\tilde{\alpha}^K > C\big(\tilde{x}^K, \tilde{y}^{0,K}, \tilde{y}^{S,K}, \tilde{z}^K\big) - \varepsilon \geq \mathcal{U} - \varepsilon$$

is satisfied for $K \geq k_0$. That is, the algorithm does not produce an infinite sequence of nodes and, therefore, terminates finitely.

It remains to show that the output of the algorithm is correct. Suppose Algorithm 3 terminates with upper bound $\mathcal{U} = \infty$. This only happens if every node was fathomed because the relaxations are infeasible. Since the leaf nodes define a partition of the feasible set and the relaxations are infeasible, so has to be the original problem.

Suppose the algorithm terminates with a solution $(\bar{x}, \bar{y}^0, \bar{y}^S, \bar{z})$. By construction of the algorithm and Lemma 2.2, it is clear that the solution is $(\delta_1 + \delta_2)$-feasible for (3). We distinguish two cases:

1. There is an optimal solution of (3) with optimal value $C^* < \infty$.
2. The feasible set of (3) is empty, i.e., $C^* = \infty$.

In the second case, clearly $C(\bar{x}, \bar{y}^0, \bar{y}^S, \bar{z}) - \varepsilon \leq C^*$ holds and $(\bar{x}, \bar{y}^0, \bar{y}^S, \bar{z})$ is ε-optimal. In the first case, let $\mathcal{F}_k = X_k \times Y_k^0 \times Y_k^S \times Z_k$ denote all nodes of the branch-and-bound tree which are fathomed due to $\alpha^k \geq \mathcal{U}^k - \varepsilon$ with optimal solution value α^k of the relaxation and current upper bound $\mathcal{U}^k$. Then $\bigcup_k \mathcal{F}_k$ defines a partition of the feasible set and $\min_k \alpha^k$ is a lower bound for C^*. With $C(\bar{x}, \bar{y}^0, \bar{y}^S, \bar{z}) \leq \mathcal{U}^k$ we can derive $C(\bar{x}, \bar{y}^0, \bar{y}^S, \bar{z}) - \varepsilon \leq \mathcal{U}^k - \varepsilon \leq \alpha^k$ and therefore the inequality

$$C(\bar{x}, \bar{y}^0, \bar{y}^S, \bar{z}) - \varepsilon \leq \min_k \ \alpha^k \leq C^*,$$

is true, i.e., $(\bar{x}, \bar{y}^0, \bar{y}^S, \bar{z})$ is ε-optimal. ∎

The presented algorithm for ODE constrained problems works under the assumption that the under- and overestimators F^ℓ and F^u exist, see Assumption 2, but their construction has not yet been specified. One possibility is to use the results of Scott and Barton [43] and Scott et al. [44, 45] for parametric ODEs of the form

$$\partial_s y(s,x) = f\big(s, x, y(s,x)\big), \qquad y(0,x) = y_0(x),$$

where the initial value is given by a continuous function $y_0(x)$ depending on the parameter x. In this setting, the solutions of the auxiliary ODE system

$$\begin{aligned} \partial_s c(s,x) &= u\big(s, x, c(s,x), C(s,x)\big), & c(0,x) &= c_0(x), \\ \partial_s C(s,x) &= o\big(s, x, c(s,x), C(s,x)\big), & C(0,x) &= C_0(x), \end{aligned}$$

satisfy $c(s,x) \leq y(s,x) \leq C(s,x)$ for all s and suitable right-hand side functions u and o. Furthermore, the solutions c and C are convex, respectively, concave w.r.t. the parameter x. Moreover, this relaxation can be used in a spatial branch-and-bound algorithm.

These relaxations may define our functions F^ℓ and F^u. However, the ODE system has to be solved with arbitrary precision, i.e., we loose the adaptivity of our approach. Hence, in the following section we present a different approach to derive adaptive under- and overestimators for solutions of ODEs that fulfill particular conditions.

3. Existence of Bounding Schemes. In the previous section, the existence of functions F^ℓ and F^u that satisfy Assumption 2 remained open. In this section, we will investigate how to define such functions based on suitable numerical methods for differential equations. Therefore, we consider a one-dimensional parameter dependent initial value problem

$$y(0) = y^0, \quad \partial_s y(s) = f(s, x, y(s)), \quad s \in [0, S]$$

with parameters x in some polytope X and possibly implicit one-step methods which can be written in the form

$$y_0 = y^0, \quad y_{i+1} = y_i + h_i\, f_h(s_i, h_i, x, y_i, y_{i+1}), \quad \forall i = 0, \dots, N-1, \tag{12}$$

given a discretization $0 = s_0 < s_1 < \cdots < s_N = S$. Define $h_i := s_{i+1} - s_i$ for all $i = 0, \dots, N-1$. The idea is to define F^ℓ and F^u via the execution of a method of the form (12), i.e., $F^\ell \colon (x, y^0) \mapsto y_N$ or $F^u \colon (x, y^0) \mapsto y_N$.

The goal is to derive lower and upper bounds on the exact solution $y(S)$ in this way. The *global truncation error* $e_N = y(S) - y_N$ might be a good indicator, but usual techniques only yield estimates on the absolute value of e_N. Therefore, we use the *local truncation error* $\tau(s,h) = y(s+h) - y(s) - h\, f_h\big(s, h, x, y(s), y(s+h)\big)$.

EXAMPLE 3.1. *We consider an explicit method with nonnegative local truncation error, i.e.,*

$$y(s_{i+1}) - y(s_i) - h_i\, f_h\big(s_i, h_i, x, y(s_i)\big) \geq 0 \tag{13}$$

holds for all $i = 0, \dots, N-1$. From (13), we can immediately derive

$$y(s_1) - y_1 = y(s_1) - y(s_0) - h_0\, f_h\big(s_0, h_0, x, y(s_0)\big) \geq 0.$$

Nevertheless, this does not guarantee $y(s_2) - y_2 \geq 0$. For example, let $\partial_s y(s) = -y(s)$ and $y(0) = 1$. For this ODE, the explicit Euler method (i.e., $f_h(s_i, h_i, x, y_i, y_{i+1}) = f(s_i, y_i)$) with equidistant step size has nonnegative truncation error and produces the solution $y_i = (1-h)^i$ for all i. Thus, with $h = 2$ we have $y_{2i} \leq y(s_{2i})$ and $y(s_{2i+1}) \leq y_{2i+1}$ for all i, whereas $y_i \leq y(s_i)$ holds for all i if $0 < h \leq 1$.

This example suggests that a signed local truncation error and small step sizes are sufficient for producing lower and upper bounds. In fact, by (13) the inequality

$$y(s_i) + h\, f_h\big(s_i, h, x, y(s_i)\big) \geq y_i + h\, f_h\big(s_i, h, x, y_i\big) = y_{i+1}$$

holds if $y(s_i) \geq y_i$ and $y + h\, f_h(s, h, x, y)$ is nondecreasing w.r.t. y, which typically is true for small step sizes. Thus, we can derive $y(s_{i+1}) \geq y_{i+1}$.

LEMMA 3.2. *Consider a method of the form (12) for a scalar ODE, i.e., $y(s) \in \mathbb{R}$, and let the local truncation error of the method be nonnegative, i.e., the inequality*

$$y(s+h) - y(s) - h\, f_h\big(s, h, x, y(s), y(s+h)\big) \geq 0$$

holds for all $s \in [0, S]$ and $h \geq 0$ with $s + h \leq S$ and all parameters $x \in X$. Suppose the derivatives satisfy

$$\partial_y f_h(s, h, x, y, \tilde{y}) \geq b \quad \text{and} \quad \partial_{\tilde{y}} f_h(s, h, x, y, \tilde{y}) \leq B$$

for constants b, $B \in \mathbb{R}$. Then if

$$0 < h_i \leq h_{max} = \begin{cases} \infty, & \text{if } b \geq 0 \text{ and } B \leq 0, \\ \frac{1}{\max\{-b, B\}}, & \text{otherwise,} \end{cases}$$

for all $i = 0, 1, \dots, N-1$, the one-step method produces for all $x \in X$ a lower bound on the solution $y(s)$, i.e., $y_i \leq y(s_i)$ for all $i = 1, \dots, N$.

Otherwise, if the local truncation error of the method is nonpositive, i.e.,

$$y(s+h) - y(s) - h\, f_h\big(s, h, x, y(s), y(s+h)\big) \leq 0$$

holds for all $s \in [0, S]$ and $h \geq 0$ with $s + h \leq S$ and all parameters $x \in X$, then we obtain under the same assumptions $y_i \geq y(s_i)$ for all $i = 1, \dots, N$.

Proof. We prove the lemma by induction on the number of grid points. Consider the function

$$R(s,h,x,y,\tilde{y}) = \tilde{y} - y - h\, f_h(s,h,x,y,\tilde{y}).$$

Then one step of (12) is given by $R(s_i,h_i,x,y_i,y_{i+1}) = 0$. By the assumption on the derivatives of f_h, we get

$$\partial_{\tilde{y}} R(s,h,x,y,\tilde{y}) = 1 - h\,\partial_{\tilde{y}} f_h(s,h,x,y,\tilde{y}) \geq 1 - h\,B.$$

Obviously, R is nondecreasing w.r.t. $\tilde{y}$ if $B \leq 0$ holds, or if all step sizes satisfy $h \leq \frac{1}{B}$. Since the local truncation error is nonnegative and $y_0 = y(0)$, we can derive the inequality

$$R\big(0,h_0,x,y(0),y(s_1)\big) \geq 0 = R\big(0,h_0,x,y_0,y_1\big).$$

Therefore, we gain the inequality $y(s_1) \geq y_1$ if either $B \leq 0$ or $h_0 \leq \frac{1}{B}$ holds. For induction we assume that $y(s_i) \geq y_i$ is satisfied. Since $\partial_y f_h$ is bounded below by b, we know

$$\partial_y R(s,h,x,y,\tilde{y}) = -1 - h\,\partial_y f_h(s,h,x,y,\tilde{y}) \leq -1 - h\,b$$

holds. Thus, R is nonincreasing w.r.t. y if either $b \geq 0$ or $h \leq -\frac{1}{b}$ holds. Again, using that the local truncation error is nonnegative we derive

$$R\big(s_i,h_i,x,y(s_i),y(s_{i+1})\big) \geq 0 = R\big(s_i,h_i,x,y_i,y_{i+1}\big).$$

Furthermore, the monotonicity w.r.t. the third argument and $y(s_i) \geq y_i$ results in

$$R\big(s_i,h_i,x,y_i,y(s_{i+1})\big) \geq R\big(s_i,h_i,x,y(s_i),y(s_{i+1})\big) \geq R\big(s_i,h_i,x,y_i,y_{i+1}\big),$$

and consequently $y(s_{i+1}) \geq y_{i+1}$ holds. Then induction yields that the one-step method produces a lower bound on $y(S)$.

The case of nonpositive truncation error can be treated in the same way. ∎

Note that for most one-step methods like Runge-Kutta methods, assuming that the right-hand side f of the ODE is Lipschitz-continuous already ensures that the partial derivatives $\partial_y f_h$ and $\partial_{\tilde{y}} f_h$ are bounded.

REMARK 3.3. *If we consider an explicit one-step method, i.e., $f_h(s,h,x,y,\tilde{y})$ is independent of $\tilde{y}$, that is $y_{i+1} = y_i - h_i\, f_h(s_i,h_i,x,y_i)$, then the previous lemma yields that we can choose*

$$h_{\max} = \begin{cases} \infty, & \text{if } b \geq 0, \\ -\frac{1}{b}, & \text{else.} \end{cases}$$

REMARK 3.4. *If we consider an "end value problem" instead of an initial value problem, that is, $\partial_s y(s) = f(s,x,y(s))$ holds for $s \in [0,S]$ and $y(S) = y^S$, then Lemma 3.2 still holds true with the modification that the bounds are now reversed, i.e., positive truncation errors now yield upper bounds and negative truncation errors now yield lower bounds.*

REMARK 3.5. *In the autonomous case (f_h is independent of s), the sign condition of Lemma 3.2 for the local truncation error is also necessary in the following sense. If there exist s, $y(s)$, such that there is no $\bar{h} > 0$ where (13) holds for all $0 < h \leq \bar{h}$, then the scheme (12) does not produce lower bounds for all initial data y^0. In fact, in the autonomous case without restrictions on the initial value y^0, we can choose $y(0) = y^0$ such that there exists $h_0 > 0$ with nonpositive local truncation error. Then a strict version of the second assertion of Lemma 3.2 is applicable, yielding $y_1 > y(s_1)$, i.e., we obtain an upper instead of a lower bound.*

For the case of a system of ODEs, we have the following analogue of Lemma 3.2 that can be proven in a similar way.

LEMMA 3.6. *Consider a method of the form (12) for a system of ODEs with $f : \mathbb{R} \times \mathbb{R}^k \times \mathbb{R}^n \to \mathbb{R}^n$. Let the local truncation error of the method be nonnegative, i.e., the inequality*

$$y(s+h) - y(s) - h\, f_h\big(s, h, x, y(s), y(s+h)\big) \geq 0$$

holds for all $s \in [0, S]$ and $h \geq 0$ with $s + h \leq S$. Define the mean value derivatives

$$\hat{D}_y f_h(s,h,x,y,\tilde{y},z,\tilde{z}) \coloneqq \int_0^1 \partial_y f_h(s,h,x,y+\tau(z-y),\tilde{y}+\tau(\tilde{z}-\tilde{y}))\,\mathrm{d}\tau,$$
$$\hat{D}_{\tilde{y}} f_h(s,h,x,y,\tilde{y},z,\tilde{z}) \coloneqq \int_0^1 \partial_{\tilde{y}} f_h(s,h,x,y+\tau(z-y),\tilde{y}+\tau(\tilde{z}-\tilde{y}))\,\mathrm{d}\tau.$$

Suppose there are $h_{max} > 0$ and $d_{max} > 0$ such that

$$\Big(I - h\,\hat{D}_{\tilde{y}} f_h(s,h,x,y(s),y(s+h),z,\tilde{z})\Big)^{-1}\Big(I + h\,\hat{D}_y f_h(s,h,x,y(s),y(s+h),z,\tilde{z})\Big)$$

has nonnegative entries for all $0 < h \leq h_{max}$, $s \in [0, S-h]$, $\|z - y(s)\| \leq d_{max}$, and $\|\tilde{z} - y(s+h)\| \leq d_{max}$. Then for all $0 < h \leq h_{max}$ such that the solution of the scheme (12) satisfies $\|y_i - y(s_i)\| \leq d_{max}$, $i = 0, \ldots, N$, one has $y_i \leq y(s_i)$ for all $i = 1, \ldots, N$.

Otherwise, if the local truncation error is nonpositive, then we obtain $y_i \geq y(s_i)$ for all $i = 1, \ldots, N$, under the same assumptions.

Coming back to the scalar case, given a specific numerical one-step method we can use Lemma 3.2 to characterize conditions under which the method produces lower or upper bounds for a scalar ODE. For example, consider the explicit midpoint method with step size h_i given by

$$y_0 = y^0, \quad y_{i+1} = y_i + h_i\, f\big(s_i + \tfrac{h_i}{2}, x, y_i + \tfrac{h_i}{2} f(s_i, x, y_i)\big), \quad \forall i = 0, \ldots, N-1. \tag{14}$$

COROLLARY 3.7. *Let $\partial_y f(s,x,y)$ be nonnegative. If both $y(s)$ and $\partial_s y(s)$ are convex then the explicit midpoint method produces lower bounds on $y(s_i)$ for all $i = 1, \ldots, N$. If both $y(s)$ and $\partial_s y(s)$ are concave then the method produces upper bounds.*

Otherwise, let $\partial_y f(s,x,y)$ be nonpositive and bounded, i.e., $b \leq \partial_y f(s,x,y) \leq 0$ holds for some $b \in \mathbb{R}$, and let the step sizes h_i satisfy the condition $0 < h_i \leq -\frac{1}{b}$ for all $i = 0, \ldots, N-1$. Then y_i is a lower bound on the ODE solution $y(s_i)$ for all $i = 1, \ldots, N$ if $y(s)$ is concave and $\partial_s y(s)$ is convex. On the other hand, y_i is an upper bound if $y(s)$ is convex and $\partial_s y(s)$ is concave.

Proof. Let $\partial_y f(s,x,y)$ be nonnegative and $y(s)$ and $\partial_s y(s)$ be convex. In order to use Lemma 3.2 we have to show that the local truncation error $y(s+h) - y(s) - h\, f\big(s + \frac{h}{2}, y(s) + \frac{h}{2} f(s, y(s))\big)$ is nonnegative.

As $\partial_s y(s)$ is convex, we obtain

$$\begin{aligned} y(s+h) - y(s) &= \int_s^{s+h} \partial_s y(\tilde{s})\,\mathrm{d}\tilde{s} \geq \int_s^{s+h} \partial_s y(s+\tfrac{h}{2}) + \partial_{ss} y(s+\tfrac{h}{2})\big(\tilde{s} - (s+\tfrac{h}{2})\big)\,\mathrm{d}\tilde{s} \\ &= h\,\partial_s y(s+\tfrac{h}{2}) + \partial_{ss} y(s+\tfrac{h}{2}) \cdot 0 = h\, f\big(s+\tfrac{h}{2}, x, y(s+\tfrac{h}{2})\big). \end{aligned} \tag{15}$$

By assumption $y(s)$ is also convex, thus, the inequality

$$y(s+\tfrac{h}{2}) \geq y(s) + \tfrac{h}{2}\partial_s y(s) \tag{16}$$

holds. Together with $\partial_y f(s,x,y) \geq 0$ the inequalities (15) and (16) yield

$$y(s+h) - y(s) - h\, f\big(s+\tfrac{h}{2}, x, y(s) + \tfrac{h}{2} f(s,y(s))\big) \geq 0,$$

thus, by Lemma 3.2 the explicit midpoint method produces lower bounds on the solution of the ODE.

The other cases are similar. ∎

Note that one does not have to know the exact solution of an ODE to determine whether the assumptions of Corollary 3.7 are satisfied. Since the second derivative of the solution is given by

$$\partial_{ss} y(s) = \partial_s f\big(s,x,y(s)\big) + \partial_y f\big(s,x,y(s)\big)\cdot f\big(s,x,y(s)\big), \tag{17}$$

it often suffices to analyze the right-hand side to check whether the solution $y(s)$ is convex or concave. Analoguously, we can check the condition on $\partial_s y(s)$.

Corollary 3.7 shows that in some cases we can use the explicit midpoint method to define one of the functions $F^\ell(x,y^0,N)$ and $F^u(x,y^0,N)$ of Assumption 2 by using the mapping $(x,y_0,N) \mapsto y_N$. Furthermore, we can state sufficient conditions for this mapping to be convex or concave.

LEMMA 3.8. *Let $\partial_y f(s,x,y)$ be nonnegative and $f(s,x,y)$ be convex in (x,y) for all s. Then the mapping $F^{em}\colon X \times Y^0 \times \mathbb{N} \to \mathbb{R}$ defined by $(x,y^0,N) \mapsto y_N$ through computing the explicit midpoint method (14) with parameter x, initial value y^0 and N discretization steps is continuously differentiable and convex w.r.t. (x,y^0).*

If $f(s,x,y)$ is concave instead of convex, then F^{em} is concave w.r.t. (x,y^0).

Proof. We consider the mapping which describes a single step of the explicit midpoint method:

$$y^{em}(s,h,x,y) = y + h\, f\big(s+\tfrac{h}{2}, x, y + \tfrac{h}{2} f(s,x,y)\big),$$

i.e., we can write $y_{i+1} = y^{em}(s_i,h_i,x,y_i)$.

Let $\tilde{y} = \tau y + (1-\tau)y'$ and $\tilde{x} = \tau x + (1-\tau)x'$ with $\tau \in [0,1]$. Then if f is convex w.r.t. (x,y), the inequality

$$\tilde{y} + \tfrac{h}{2} f(s,\tilde{x},\tilde{y}) \leq \tau[y + \tfrac{h}{2} f(s,x,y)] + (1-\tau)[y' + \tfrac{h}{2} f(s,x',y')]$$

holds. Together with $\partial_y f(s,x,y)$ being nonnegative, we can derive

$$\begin{aligned}
y^{em}(s,h,\tilde{x},\tilde{y}) &= \tilde{y} + h\, f\big(s+\tfrac{h}{2}, \tilde{x}, \tilde{y} + \tfrac{h}{2} f(s,\tilde{x},\tilde{y})\big) \\
&\leq \tilde{y} + h\, f\big(s+\tfrac{h}{2}, \tilde{x}, \tau[y + \tfrac{h}{2} f(s,x,y)] + (1-\tau)[y' + \tfrac{h}{2} f(s,x',y')]\big) \\
&\leq \tilde{y} + \tau h\, f\big(s+\tfrac{h}{2}, x, y + \tfrac{h}{2} f(s,x,y)\big) + (1-\tau) h\, f\big(s+\tfrac{h}{2}, x', y' + \tfrac{h}{2} f(s,x',y')\big) \\
&= \tau y^{em}(s,h,x,y) + (1-\tau) y^{em}(s,h,x',y'),
\end{aligned}$$

i.e., y^{em} is convex w.r.t. (x,y). Additionally, y^{em} is continuously differentiable if f is continuously differentiable and, since $\partial_y f(s,x,y)$ is nonnegative, we can derive that $\partial_y y^{em}(s,h,x,y) \geq 1$.

Since the composition of an increasing convex function with a convex function is convex and y^{em} is continuously differentiable, we can inductively show that y_N is continuously differentiable and convex w.r.t. parameter and initial value. Analogously, we can see that it is concave if f is concave. ∎

In the case of $\partial_y f(s,x,y) \le 0$ the above proof ideas do not work, but y^{em} might still be convex or concave.

Under weaker assumptions, which are motivated by the application to gas flow in pipes with positive slope, analogous bounding properties hold for the second order Taylor scheme

$$y_0 = y^0,\ y_{i+1} = y_i + h_i\, f(s_i,x,y_i) + \tfrac{h_i^2}{2}(\partial_y f\, f + \partial_s f)(s_i,x,y_i),\ \forall i = 0,\dots,N-1. \tag{18}$$

COROLLARY 3.9. *Let $b \le 0$ and $b_1 \ge 0$ with $b \le \partial_y f(s,x,y)$ and $-b_1 \le (\partial_{yy} f\, f + \partial_y f^2 + \partial_{sy} f)(s,x,y)$ for all $x \in X$. If $\partial_s y(s)$ is convex, then the second order Taylor scheme (18) produces lower bounds on $y(s_i)$ for all $i = 1,\dots,N$ if*

$$0 < h_i \le \begin{cases} \infty & \text{if } b = b_1 = 0, \\ \frac{2}{-b+\sqrt{b^2+2b_1}} & \text{otherwise.} \end{cases} \tag{19}$$

If $\partial_s y(s)$ is concave, then under the above condition on h_i the second order Taylor scheme (18) produces upper bounds on $y(s_i)$ for all $i = 1,\dots,N$.

Proof. Suppose that $\partial_s y(s)$ is convex. Then using (17)

$$\begin{aligned} y(s+h) - y(s) &= \int_s^{s+h} \partial_s y(\tilde{s})\,\mathrm{d}\tilde{s} \ge \int_s^{s+h} \partial_s y(s) + \partial_{ss} y(s)\cdot(\tilde{s}-s)\,\mathrm{d}\tilde{s} \\ &= h\, f_h(s,h,x,y(s)), \end{aligned}$$

where $f_h(s,h,x,y) \coloneqq f(s,x,y) + \frac{h}{2}(\partial_y f\, f + \partial_s f)(s,x,y)$. Thus, (18) has a nonnegative local truncation error.

In order to apply Lemma 3.2, we additionally have to show that (19) implies

$$\tfrac{b-\sqrt{b^2+2b_1}}{2} \le \partial_y f_h(s,h,x,y) = \partial_y f(s,x,y) + \tfrac{h}{2}(\partial_{yy} f\, f + \partial_y f^2 + \partial_{sy} f)(s,x,y).$$

This is easily verified by inserting the lower bounds b and $-b_1$ and the upper bound for h according to (19).

If $\partial_s y(s)$ is concave, then the numerical method has a nonpositive local truncation error. Thus, it produces upper bounds. ∎

We now state conditions that ensure the convexity or concavity of the mapping $(x, y^0, N) \mapsto y_N$ defined by the second order Taylor method.

LEMMA 3.10. *Let $f_h(s,h,x,y) = f(s,x,y) + \frac{h}{2}(\partial_y f\, f + \partial_s f)(s,x,y)$ and let f be three times continuously differentiable. Assume that there exists $\bar{h} > 0$ satisfying (19) such that f_h is convex w.r.t. (x,y) for all s and $0 < h \le \bar{h}$. Then for $0 < h \le \bar{h}$ the mapping $F^{ta}\colon X \times Y^0 \times \mathbb{N} \to \mathbb{R}$ defined by $(x, y^0, N) \mapsto y_N$ through the Taylor method (18) is continuously differentiable and convex w.r.t. (x, y^0).*

Alternatively, let $f(s,x,y)$ be uniformly convex w.r.t. (x,y) and its derivatives w.r.t. y and x up to order three be bounded. Then there exists $\bar{h} > 0$ such that $F^{ta}\colon X \times Y^0 \times \mathbb{N} \to \mathbb{R}$ is convex w.r.t. (x, y^0) for $0 < h_i \le \bar{h}$.

If convex is replaced by concave in the assumptions, F^{ta} is concave w.r.t. (x, y^0).

Proof. Consider the function $y^{ta}(s,x,y) \coloneqq y + h\, f_h(s,h,x,y)$ defining one step of the Taylor scheme (18). Then we can write (18) as

$$y_0 = y^0, \quad y_{i+1} = y^{ta}(s_i,x,y_i), \quad \forall i = 0,\dots,N-1.$$

Hence, if y^{ta} is nondecreasing w.r.t. y and convex w.r.t. (x, y), we can inductively derive that F^{ta} is convex w.r.t. (x, y).

By assumption there exists $\bar{h} > 0$ satisfying (19) such that f_h is convex for $0 < h \leq \bar{h}$. Then $y^{ta}(s, x, y)$ is convex. Moreover, as shown in Corollary 3.9, $\partial_y y^{ta}(s, x, y) \geq 0$ if the step size h satisfies (19).

In the alternative case, assume that $f(s, x, y)$ is uniformly convex w.r.t. (x, y). Then $D^2 f(s, x, y)$ is uniformly positive definite. As before, (19) ensures that we have $\partial_y y^{ta}(s, x, y) \geq 0$. Furthermore, we have

$$D^2 y^{ta}(s, x, y) = hD^2 f(s, y, x) + \tfrac{h^2}{2} D^2(\partial_y f\, f + \partial_s f)(s, x, y).$$

Since $D^2 f(s, y, x)$ is by assumption uniformly positive definite and the derivatives in the second term are bounded, we can choose $\bar{h}$ satisfying (19) small enough such that $D^2 y^{ta}$ is positive definite for all $0 < h \leq \bar{h}$. Then $y^{ta}(s, x, y)$ is convex in (x, y) and monotone increasing w.r.t. y and thus F^{ta} is convex w.r.t. (x, y).

The concave case can be handled analogously. ∎

To obtain opposite bounds we consider the trapezoidal rule, which is given by

$$y_0 = y^0, \quad y_{i+1} = y_i + \tfrac{h}{2}\big[f(s_i, x, y_i) + f(s_{i+1}, x, y_{i+1})\big], \quad \forall i = 0, \dots, N-1 \tag{20}$$

with the discretization $0 = s_0 < s_1 < \cdots < s_N = S$. For simplicity we assume that the discretization is equidistant with step size h – the results proven in the following can easily be extended to the non-equidistant case. Again, by using Lemma 3.2 we can derive the following corollary.

COROLLARY 3.11. *Let $\partial_s y(s)$ be convex and $b \leq \partial_y f(s, x, y) \leq B$ be bounded by some constants $b, B \in \mathbb{R}$ for all $x \in X$. Furthermore, suppose the step size satisfies the condition $h \cdot \max\{-b, B\} \leq 2$ and a solution to (20) for all $i = 1, \dots, N$ exists. Then y_i is an upper bound on the ODE solution $y(s_i)$ for all $i = 1, \dots, N$.*

If $\partial_s y(s)$ is concave instead of convex, then y_i is a lower bound for all i.

Proof. We only discuss the case in which $\partial_s y(s)$ is convex, the other case works analogously. By Lemma 3.2 we have to show that the local truncation error is nonpositive. Since $\partial_s y(s) = f\big(s, x, y(s)\big)$ is convex, the equation

$$f\big(\tilde{s}, x, y(\tilde{s})\big) \leq f\big(s, x, y(s)\big) + \tfrac{1}{h}\Big[f\big(s+h, x, y(s+h)\big) - f\big(s, x, y(s)\big)\Big](\tilde{s} - s)$$

holds for all $\tilde{s} \in [s, s+h]$. With this we can derive

$$y\big(s+h\big) - y\big(s\big) = \int_s^{s+h} f\big(\tilde{s}, x, y(\tilde{s})\big)\, \mathrm{d}\tilde{s} = \tfrac{h}{2}\Big[f\big(s, x, y(s)\big) + f\big(s+h, x, y(s+h)\big)\Big]$$

and, thus, the local truncation error is nonpositive.

By assumption the derivative of $\partial_y f(s, x, y)$ is bounded, which implies the boundedness condition of Lemma 3.2. Also the upper bound on the step sizes immediately follows from the formula in Lemma 3.2. Therefore, the trapezoidal rule produces upper bounds on the solution $y(s_i)$ for all $i = 1, \dots, N$. ∎

As for the explicit midpoint method, we now consider the function defined by computing the trapezoidal rule and mapping parameter x and initial value y_0 to y_N.

LEMMA 3.12. *Let $b \leq \partial_y f(s, x, y) \leq B$ for some constants b, $B \in \mathbb{R}$ and let $f(s, x, y)$ be convex in (x, y) for all s. Furthermore, suppose there is a solution to (20)*

for all $i = 1, \dots, N$ if the step size satisfies the condition $h \cdot \max\{-b, B\} < 2$. Then the mapping $F^{tr} \colon X \times Y^0 \times \mathbb{N} \to \mathbb{R}$ defined by $(x, y^0, N) \mapsto y_N$ through computing the trapezoidal rule (20) with parameter x and initial value y^0 and N discretization steps is continuously differentiable and convex w.r.t. (x, y^0).

If $f(s, x, y)$ is concave instead of convex, then F^{tr} is concave w.r.t. (x, y^0).

Proof. We consider the following function, which is defined by a single step of the trapezoidal rule:

$$R(s, x, y, y^{tr}) = y^{tr} - y - \frac{h}{2}\big[f(s, x, y) + f(s + h, x, y^{tr})\big].$$

By assumption, there exists a solution y_{i+1} of $R(s_i, x, y_i, y^{tr}) = 0$ w.r.t. y^{tr} for all $i = 0, 1, \dots, N - 1$. Furthermore, the inequality

$$\partial_{y^{tr}} R(s, x, y, y^{tr}) = 1 - \tfrac{h}{2}\, \partial_y f(s + h, x, y^{tr}) \geq 1 - \tfrac{h}{2}\, B > 0$$

holds, if $B \leq 0$ or $h < \frac{2}{B}$ is satisfied, i.e., the assumptions of the implicit function theorem are satisfied. Thus, there exists a continuously differentiable function $y^{tr}(s, x, y)$ with

$$y^{tr}(s, x, y) - \tfrac{h}{2}\, f\big(s + h, x, y^{tr}(s, x, y)\big) = y + \tfrac{h}{2}\, f\big(s, x, y\big). \tag{21}$$

Analogously to the proof of Lemma 3.8, we will use this function to inductively derive that y_N is a convex function w.r.t. (x, y^0) assuming that $f(s, x, y)$ is a convex function w.r.t. (x, y).

Let $\tilde{y} = \tau y + \tau' y'$ and $\tilde{x} = \tau x + \tau' x'$ for some $\tau \in [0, 1]$ and $\tau' = 1 - \tau$. By the definition of y^{tr}, we can derive the inequality

$$\begin{aligned}
&y^{tr}(s, \tilde{x}, \tilde{y}) - \tfrac{h}{2}\, f\big(s + h, \tilde{x}, y^{tr}(s, \tilde{x}, \tilde{y})\big) = \tilde{y} + \tfrac{h}{2}\, f\big(s, \tilde{x}, \tilde{y}\big) \\
&\leq \tilde{y} + \tfrac{h}{2}\big[\tau f(s, x, y) + \tau' f(s, x', y')\big] \\
&= \tau\big[y^{tr}(s, x, y) - \tfrac{h}{2}\, f\big(s + h, x, y^{tr}(s, x, y)\big)\big] \\
&\qquad + \tau'\big[y^{tr}(s, x', y') - \tfrac{h}{2}\, f\big(s + h, x', y^{tr}(s, x', y')\big)\big] \\
&\leq \tau y^{tr}(s, x, y) + \tau' y^{tr}(s, x', y') - \tfrac{h}{2}\, f\big(s + h, \tilde{x}, \tau y^{tr}(s, x, y) + \tau' y^{tr}(s, x', y')\big).
\end{aligned}$$

From this we can derive the convexity of $y^{tr}(s, x, y)$ if $y - \frac{h}{2}\, f(s, x, y)$ is nondecreasing w.r.t. y, that is, if $1 - \frac{h}{2}\partial_y f(s, x, y) \geq 0$ holds. As we have seen before, this is true due to the choice of $h \leq \frac{2}{B}$ if $B > 0$. Thus, $y^{tr}(s, x, y)$ is convex.

Next, we show that $y^{tr}(s, x, y)$ is nondecreasing w.r.t. y. By differentiating (21) we derive

$$\partial_y y^{tr}(s, x, y)\big[1 - \tfrac{h}{2}\, \partial_y f\big(s + h, x, y^{tr}(s, x, y)\big)\big] = 1 + \tfrac{h}{2}\, \partial_y f\big(s, x, y\big).$$

The right-hand side is nonnegative if either $b \geq 0$ or $h < \frac{2}{-b}$ is satisfied. Together with $1 - \frac{h}{2}\, \partial_y f\big(s + h, x, y^{tr}(s, x, y)\big) > 0$ if either $B \leq 0$ or $h < \frac{2}{B}$, this yields that $y^{tr}(s, x, y)$ is nondecreasing.

By interpreting the approximations y_i for $i = 1, \dots, N$ of the ODE solution as a function of parameter and initial value, we can use the following representation

$$y_{i+1}(x, y^0) = y^{tr}\big(s_i, x, y_i(x, y^0)\big).$$

Since $y^{tr}(s,x,y)$ is continuously differentiable and nondecreasing, and the composition of a convex function with a nondecreasing convex function is convex, we can inductively derive that F^{tr} is continuously differentiable and convex w.r.t. (x,y). Furthermore, the condition on the minimal number of discretization steps N directly follows from the condition on the step size. ∎

Convex relaxations $\check{F}_i^\ell(x,y^0,N) \le y_i^S \le \hat{F}_i^u(x,y^0,N)$, $i \in \{1,\dots,n\}$, on $X \times Y^0$, see (5), can be constructed using the results of this section as follows. If the lower bound $\check{F}_i^\ell$ is convex, one can either directly use it in an NLP solver or generate gradient cuts. If the upper bound $\hat{F}_i^u$ is concave, it can again be used directly. If $\hat{F}_i^u$ is convex, then the best concave overestimator is piecewise-linear and can be constructed using the values of $\hat{F}_i^u$ at the extreme points of $X \times Y^0$ in a standard fashion, see Horst and Tuy [19, Theorem IV.6].

In the context of bound propagation, the conditions stated in Corollaries 3.7, 3.9 and 3.11 are rather strict, since they require the derivative of the solution of the scalar differential equation to be convex or concave (in Corollary 3.7 also the solution itself). These requirements can be checked by analyzing the right-hand side, see the example of gas transport in the next section.

For example, the second derivative of the solution of an autonomous ODE is $\partial_{ss} y(s) = \partial_y f\big(x,y(s)\big) f\big(x,y(s)\big)$. Thus, if bounds on the possible solutions are known, which is often the case, one can check whether the right-hand side and its derivatives change their signs in this interval. Therefore, the set of possible solution values can be partitioned in such a way that the solution is either convex or concave on each part. Constucting under- and overestimators for every part provides under- and overestimators for the whole solution. In Subsection 4.4 we analyze an example, where in one case the right-hand side changes its sign and in a second case the derivative of the right-hand side changes its sign.

4. Application – Stationary Gas Transport. Within this section we show that our solution method can be applied to the example of stationary gas transport.

4.1. The Model. Let a gas network be given by a directed graph $\mathcal{G} = \big(\mathcal{V},\mathcal{A}\big)$, where the nodes in $\mathcal{V}$ are entries, exits and junctions of the network, and $\mathcal{A}$ are the network elements like valves, resistors, compressors and pipes. As a basis for the gas flow in pipes we consider the stationary isothermal Euler-equation, which is a one-dimensional ordinary differential equation in space. In the following we concentrate on the differential equation, whereas we refer to [21, Chapter 6] for the models of the other network elements.

In terms of pressure p and mass flow rate q, the Euler equation reads

$$\partial_x p(x) \left(1 - \frac{c^2 q^2}{A^2 p(x)^2}\right) = -\frac{\lambda c^2}{2DA^2} q|q| \frac{1}{p(x)} - \frac{g}{c^2}\sigma\, p(x), \quad 0 < x < L. \tag{22}$$

This ODE describes the pressure of gas flowing along a single cylindrical pipe of length L. Note that in stationary gas transport, the mass flow rate is constant along each pipe, but within our spatial branch-and-bound framework the mass flow rates for each pipe are variables. The constants are the speed of sound c, the cross-sectional area of the pipe A, the friction coefficient λ, the diameter of the pipe D, the gravitational acceleration g, and the slope $\sigma \in [-1,1]$ of the pipe. Note that we could also use non-constant models for the friction factor, see Remark 4.12

In the following, we assume the pipe to be horizontal, i.e., $\sigma = 0$, and in Subsection 4.4 we extend the method to the general case. Assuming that the gas flows with

subsonic velocity v, i.e., $\frac{v}{c} < 1$, we can use the relation $q = A\,\rho\,v$ of mass flow rate and velocity with density ρ and the relation $\rho\,c^2 = p$ to derive $\frac{v}{c} = \frac{c|q|}{Ap} < 1$. With these assumptions, we can rewrite (22) to

$$\partial_x p(x) = \varphi(p(x), q), \quad 0 < x < L, \quad \varphi(p, q) \coloneqq -\frac{\lambda\, c^2\, q|q|\, p}{2D\big(A^2p^2 - c^2q^2\big)}.$$

In the following, we will need the derivative $\partial_p\varphi$ to be bounded, which is implied by an upper bound on $\frac{v}{c}$, strictly smaller than 1. For instance, we can require $\frac{v}{c} \leq 0.8$, which is often used in aerodynamics to distinguish subsonic velocity from transonic velocity, i.e., velocity near the speed of sound. In practice this is no limitation, since the gas velocity is typically much smaller than the speed of sound. See also Remark 4.5.

With this, we can formulate our optimization model:

$$\begin{aligned}
\min \quad & C(p, q, z) \\
\text{s.t.} \quad & G(p, q, z) \leq 0, \\
& \partial_x p_a(x) = \varphi_a\big(p_a(x), q_a\big) \quad \forall\, a \in \mathcal{A}_{pipe} \subseteq \mathcal{A}, \\
& p_u = p_a(0),\ p_v = p_a(L_a) \quad \forall\, a = (u, v) \in \mathcal{A}_{pipe}, \\
& p \in P,\ q \in Q,\ z \in Z.
\end{aligned} \tag{23}$$

Here, we have pressure variables p_v for all nodes $v \in \mathcal{V}$, functions $p_a(x)$ for each pipe $a \in \mathcal{A}_{pipe} \subseteq \mathcal{A}$, flow variables q_a for all edges $a \in \mathcal{A}$, and binary variables $z \in \{0, 1\}^m$ for the discrete decisions like to open or close a valve, or turn a compressor on or off. The sets $P \subset \mathbb{R}^{\mathcal{V}}$ and $Q \subset \mathbb{R}^{\mathcal{A}}$ are given by variable bounds $0 < \underline{p}_v \leq p_v \leq \overline{p}_v$ for all $v \in \mathcal{V}$ and $\underline{q}_a \leq q_a \leq \overline{q}_a$ for all $a \in \mathcal{A}$, respectively. Just like before, the ODEs only define a coupling between the pressure and flow variables for a single pipe. Furthermore, note that φ_a depends on $a \in \mathcal{A}_{pipe}$, since the pipes might differ in length, diameter or friction coefficient.

The constraint $G(p, q, z) \leq 0$ represents the models for the different network elements, as well as the flow conservation given by

$$\sum_{a \in \delta^+(v)} q_a - \sum_{a \in \delta^-(v)} q_a = q_v^{\pm} \qquad \forall\, v \in \mathcal{V}, \tag{24}$$

where $\delta^+(v)$ denotes the outgoing arcs of v, $\delta^-(v)$ denotes the ingoing arcs and $q_v^{\pm}$ is the in- or outflow at node v. Note that the flow is negative if it flows in the opposite direction of the arc. Furthermore, $G(p, q, z) \leq 0$ contains the inequalities

$$5\,c\,q_a - 4A\,p_v \leq 0, \quad -5\,c\,q_a - 4A\,p_u \leq 0$$

for all pipes $a = (u, v)$ such that the subsonic flow condition is satisfied. Note that it is sufficient to demand $5\,c\,|q| \leq 4A\,p$ at the ends of the pipe, since the pressure drops in the direction of the flow.

As in the abstract problem setting, we suppose that in the branch-and-bound framework $C(p, q, z)$ and $G(p, q, z)$ can be treated by standard techniques, like the α-BB method [1, 2], such that we can focus on the differential equations.

4.2. Lower and Upper Bounds for the Inflow Pressure. Within this subsection we interpret the differential equation as an initial value problem with the pressure at the end of the pipe as "initial" value, see Remark 4.6 for an explanation.

We will derive lower and upper bounds on the inflow pressure. Therefore, we assume from now on that the flow is nonnegative, i.e., $q \geq 0$. During the branch-and-bound process we can ensure this by branching on the flow w.r.t. $q = 0$.

Let $\mathcal{D} \coloneqq \{(p,q) \in \mathbb{R}^2 \mid 0 < \underline{p} \leq p \leq \overline{p},\ 0 \leq \underline{q} \leq q \leq \overline{q},\ 5c\,q \leq 4A\,p\}$ be the domain of φ given by variable bounds and the subsonic flow condition. Simple differentiating and computing the eigenvalues of the Hessian yields the following properties of φ.

LEMMA 4.1. *The function $\varphi\colon \mathcal{D} \to \mathbb{R}$ is nonpositive, nondecreasing in p, nonincreasing in q and concave in $(p,q) \in \mathcal{D}$. Its second derivatives satisfy $\partial_{pp}\varphi(p,q) \leq 0$, $\partial_{pq}\varphi(p,q) \geq 0$ and $\partial_{qq}\varphi(p,q) < 0$. Furthermore, we have that $\varphi(p,q)$, $\partial_p\varphi(p,q)$, $\partial_q\varphi(p,q)$, $\partial_{pp}\varphi(p,q)$ or $\partial_{pq}\varphi(p,q)$ are 0 if and only if $q = 0$.*

This leads to the following properties of the differential equation.

COROLLARY 4.2. *The ordinary differential equation*

$$p(L) = p^0, \quad \partial_x p(x) = \varphi\big(p(x), q\big), \quad 0 \leq x \leq L \tag{25}$$

has a unique solution $p(x)$ for all $(p^0, q) \in \mathcal{D}$. Furthermore, $p(x)$, as well as $\partial_x p(x)$, is nonincreasing and concave.

Proof. For $q = 0$ the ODE has the only solution $p(x) = p^0$, since $\varphi(p,0) = 0$. For fixed $q > 0$, the right-hand side $\varphi\big(p(x), q\big)$ is negative, i.e., $p(x)$ is nonincreasing and the pressure is bounded from below by p^0. Thus, $\partial_p\varphi(p,q)$ is bounded by $\partial_p\varphi\big(\frac{5cq}{4A}, q\big) \geq \partial_p\varphi(p,q) > 0$, i.e., φ is Lipschitz-continuous w.r.t. p. Hence, there is a unique solution of the ODE. Finally, $\partial_{xx}p(x) = (\partial_p\varphi\,\varphi)(p(x),q)$, $\partial_{xxx}p(x) = (\partial_{pp}\varphi\,\varphi^2 + (\partial_p\varphi)^2\,\varphi)(p(x),q)$ and the properties of φ in Lemma 4.1 yield the remaining statements. ∎

In the next step, we apply the explicit midpoint method and the implicit trapezoidal rule opposite to the flow direction. Then we deduce by Corollaries 3.7 and 3.11 that the methods define lower and upper bounds on $p(0)$, as illustrated in Figure 1. Therefore, let a discretization $0 = x_N < \ldots < x_1 < x_0 = L$ be given. For simplicity, we assume that the discretization is equidistant with step size $h = \frac{L}{N}$. Then the explicit midpoint method is defined by

$$p_0^\ell = p^0, \quad p_{i+1}^\ell = p_i^\ell - h\,\varphi\big(p_i^\ell - \tfrac{h}{2}\varphi(p_i^\ell, q), q\big), \qquad \forall i = 0, \ldots, N-1, \tag{26a}$$

and the implicit trapezoidal rule is

$$p_0^u = p^0, \quad p_{i+1}^u = p_i^u - \tfrac{h}{2}\left[\varphi(p_i^u, q) + \varphi(p_{i+1}^u, q)\right], \quad \forall i = 0, \ldots, N-1. \tag{26b}$$

COROLLARY 4.3. *The explicit midpoint method (26a) with $(p^0, q) \in \mathcal{D}$ and step size $0 < h \leq \frac{81}{328}\frac{D}{\lambda}$ defines a lower bound on the solution $p(0)$ of (25).*

Proof. Since (25) is an end value problem, we apply Corollary 3.7 to the transformed differential equation given by $\tilde{p}(x) = p(L - x)$ with

$$\tilde{p}(0) = p^0, \quad \partial_x\tilde{p}(x) = -\varphi\big(\tilde{p}(x), q\big), \quad 0 \leq x \leq L,$$

which is an ODE of the form considered in the previous section. By Corollary 4.2 we can derive that $\tilde{p}(x)$ is concave and $\partial_p\tilde{p}(x)$ is convex. Furthermore, the right-hand side satisfies

$$0 \geq -\partial_p\varphi\big(p, q\big) \geq -\partial_p\varphi\big(\tfrac{5cq}{4A}, q\big) = -\tfrac{328}{81}\tfrac{\lambda}{D} =: b$$

since $-\partial_{pp}\varphi(p,q) \geq 0$ holds. Hence, by Corollary 3.7 the explicit midpoint method produces a lower bound on $\tilde{p}(L) = p(0)$. ∎

Analogously to this Corollary, we can immediatly derive the following from Corollary 3.11.

COROLLARY 4.4. *The trapezoidal rule* (26b) *with step size* $0 < h \leq \frac{81}{164}\frac{D}{\lambda}$ *defines an upper bound on the solution* $p(0)$ *of* (25).

REMARK 4.5. *Note that for Corollaries 4.3 and 4.4 to hold, it is essential that* $\frac{v}{c} = \frac{c\,q}{A\,p}$ *has an upper bound, which is strictly smaller than* 1. *Otherwise, the derivative* $\partial_p\varphi(p,q)$ *would not be bounded, and we could not apply Corollaries 3.7 and 3.11.*

REMARK 4.6. *Instead of applying the methods* (26a) *and* (26b) *in opposite direction of the flow, we could also use them to compute bounds in the direction of the flow. Then the explicit midpoint method would produce upper bounds and the trapezoidal rule would produce lower bounds, if there is a solution* p_i *for all grid points* x_i.

If we start with a small input pressure, then we cannot guarantee that there is a solution with $5c\,q \leq 4A\,p_i$ *for all* i. *If both schemes fail to produce solutions, which fulfill this bound, then we can deduce that the input pressure is infeasible. Otherwise, if only the trapezoidal rule fails to produce a lower bound, e.g., if we choose the input pressure such that for the exact solution* $p(L) = \frac{5c\,q}{4A}$ *holds, then we cannot decide whether the discretization is too coarse or the input pressure is infeasible. Therefore, we compute the schemes in opposite direction of the flow.*

With these two schemes in mind, we define two functions $p^\ell, p^u \colon \mathcal{D} \times \mathbb{N} \to \mathbb{R}$ through the computation of (26a) and (26b). That is

$$p^\ell(p,q,N) \coloneqq p_N^\ell \qquad \text{and} \qquad p^u(p,q,N) \coloneqq p_N^u,$$

where $p_0^\ell = p_0^u = p$. We can derive the following properties for p^ℓ and p^u.

LEMMA 4.7. *Let* N *be big enough such that the condition* $h = \frac{L}{N} \leq 0.16\frac{D}{\lambda}$ *holds. Then the functions* p^ℓ *and* p^u *are nondecreasing, continuously differentiable and convex in* (p,q). *Furthermore, every solution* $\bar{p}(x)$ *of the differential equation* $\partial_x p(x) = \varphi\big(p(x),q\big)$ *with* $q \geq 0$ *satisfies the inequality*

$$p^\ell\big(\bar{p}(L),q,N\big) \leq \bar{p}(0) \leq p^u\big(\bar{p}(L),q,N\big). \tag{27}$$

Proof. The inequalities in (27) follow from Corollaries 4.3 and 4.4. Furthermore, the properties of p^u follow directly from Lemma 3.12, whereas we cannot apply Lemma 3.8, since the assumption of $\partial_y f$ being nonnegative is not satisfied. Thus, it remains to show differentiability, monotonicity and convexity of $p^\ell \colon \mathcal{D} \times \mathbb{N} \to \mathbb{R}$.

With $p_0^\ell = p_0^\ell(p,q) = p$, we can write (26a) as

$$p_{i+1}^\ell(p,q) = F^\ell\big(p_i^\ell(p,q),q,h\big) \coloneqq p_i^\ell(p,q) - h\,\varphi\big(p_i^\ell(p,q) - \tfrac{h}{2}\varphi(p_i^\ell(p,q),q),q\big),$$

for $i = 0,\dots,N-1$. Differentiating yields $\partial_p\, p_0^\ell(p,q) = 1$, $\partial_q\, p_0^\ell(p,q) = 0$ and

$$\begin{aligned}
\partial_p\, p_{i+1}^\ell(p,q) &= \partial_p F^\ell\big(p_i^\ell(p,q),q,h\big)\ \partial_p\, p_i^\ell(p,q),\\
\partial_q\, p_{i+1}^\ell(p,q) &= \partial_p F^\ell\big(p_i^\ell(p,q),q,h\big)\ \partial_q\, p_i^\ell(p,q) + \partial_q F^\ell\big(p_i^\ell(p,q),q,h\big),
\end{aligned}$$

where $\partial_p F^\ell$ and $\partial_q F^\ell$ denotes the partial derivative of F^ℓ with respect to the first and second argument, respectively. Moreover, we have $D^2 p_0^\ell(p,q) = 0$ and

$$\begin{aligned}
D^2 p_{i+1}^\ell(p,q) = {} & D\begin{pmatrix} p_i^\ell(p,q) \\ q \end{pmatrix}^\top D^2 F^\ell\big(p_i^\ell(p,q),q,h\big) D\begin{pmatrix} p_i^\ell(p,q) \\ q \end{pmatrix}\\
& + \partial_p F^\ell\big(p_i^\ell(p,q),q,h\big) D^2 p_i^\ell(p,q).
\end{aligned}$$

Hence, we obtain by induction that $\partial_p p^\ell_{i+1}(p,q) \geq 0$ and $D^2 p^\ell_{i+1}(p,q)$ is positive semidefinite, if $\partial_p F^\ell\big(p^\ell_i(p,q),q,h\big) \geq 0$ and $D^2 F^\ell\big(p^\ell_i(p,q),q,h\big)$ is positive semidefinite. If additionally $\partial_q F^\ell\big(p^\ell_i(p,q),q,h\big) \geq 0$ holds then also $\partial_q p^\ell_{i+1}(p,q) \geq 0$ follows.

Since $\varphi(p,q) \leq 0$ on $\mathcal{D}$, we obtain by (26a) that $p^\ell_{i+1}(p,q) \geq p^\ell_i(p,q)$ and thus $(p^\ell_i(p,q),q) \in \mathcal{D}$ for $i = 0,\dots,N$. Moreover, (26a) yields

$$\partial_p F^\ell(p,q,h) = 1 - h\,\partial_p\,\varphi\big(p - \tfrac{h}{2}\varphi(p,q),q\big)\big(1 - \tfrac{h}{2}\partial_p\,\varphi(p,q)\big).$$

By Lemma 4.1 and its proof we have $\partial_p\varphi \geq 0$, $\partial_{pp}\varphi \leq 0$ on $\mathcal{D}$ and $1 - \frac{h}{2}\partial_p\,\varphi(p^\ell_i,q) \geq 0$ for $0 < h \leq \frac{81D}{164\lambda}$. This shows that

$$\partial_p F^\ell(p^\ell_i,q,h) \geq 1 - h\,\partial_p\,\varphi\big(p^\ell_i - \tfrac{h}{2}\varphi(p^\ell_i,q),q\big) \geq 1 - h\,\partial_p\,\varphi(p^\ell_i,q) \geq 0 \quad \forall\, 0 < h \leq \tfrac{81D}{328\lambda}.$$

Moreover, one can verify that $D^2F^\ell(p^\ell_i,q)$ is singular and is thus positive semidefinite on $\mathcal{D}$ if $\partial_{pp}F^\ell(p^\ell_i,q) \geq 0$ on $\mathcal{D}$.

To show the latter, we observe that $\partial_{pp}F^\ell(p^\ell_i,q,h)$ is a rational function in p^ℓ_i and h with positive denominator on $\mathcal{D}$. The numerator is a polynomial in p^ℓ_i whose value and all its derivatives are nonnegative at $p^\ell_i = \frac{5cq}{4A}$ for all $0 < h \leq 0.16\frac{D}{\lambda}$. Hence, the numerator – and thus $\partial_{pp}F^\ell(p^\ell_i,q,h)$ – is nonnegative for all $(p^\ell_i,q) \in \mathcal{D}$. As already observed, this implies by induction that $\partial_p p^\ell_{i+1}(p,q) \geq 0$ and that $D^2 p^\ell_{i+1}(p,q)$ is positive semidefinite for all $(p,q) \in \mathcal{D}$.

Finally, $\partial_q F^\ell(p,0,h) = 0$ and $\partial_{qq}F^\ell(p,q,h) \geq 0$ on $\mathcal{D}$, since $D^2F^\ell(p,q,h)$ is positive semidefinite on $\mathcal{D}$. Thus, also $\partial_q F^\ell(p^\ell_i(p,q),q,h) \geq 0$ holds and we deduce $\partial_q p^\ell_{i+1}(p,q) \geq 0$. ∎

Remark 4.8. *Our numerical results, see Section 5, motivate to use a smaller bound on* $\frac{v}{c} = \frac{cq}{Ap}$*, since the maximal ratio we observed in any (optimal) solution is less than* 0.1*. Therefore, if we would use* $\frac{cq}{Ap} \leq 0.2$ *or a more conservative bound* 0.4 *instead of* 0.8*, the upper bounds on the step sizes in Corollary 4.3, Corollary 4.4 and Lemma 4.7 increases drastically.*

Using the conservative bound 0.4*, Lemma 4.7 holds if the step size satisfies the condition* $h \leq 4.925\frac{D}{\lambda}$*. The bound* 0.2 *leads to an upper bound* $h \leq 29.15\frac{D}{\lambda}$*.*

The functions p^ℓ and p^u can be used to relax the ODE constraints, see (4). For deriving a convex relaxation we only have to construct a concave overestimator of p^u, since p^ℓ is already convex. The domain $\mathcal{D}$ of p^ℓ and p^u is given by the intersection of the box $[\underline{p},\overline{p}] \times [\underline{q},\overline{q}]$ with nonnegative lower bounds and the inequality $5c\,q \leq 4A\,p$. The resulting polytope has at most five vertices. Since the concave envelope of a convex function over a polytope is defined by the vertices of the polytope (see, e.g., Horst and Tuy [19, Theorem IV.6]), it consists of at most three linear inequalities.

Consequently, choosing the number of grid points N_a for each pipe $a \in \mathcal{A}$ sufficiently big such that the condition $|p^u_a(p,q,N_a) - p^\ell_a(p,q,N_a)| \leq \delta_1$ is fulfilled (compare with Lemma 2.2), we can apply Algorithm 1 and produce (ε,δ)-optimal solutions for the relaxation of problem (23).

4.3. Adaptive LP-Relaxation. In this section, we discuss how to compute valid linear inequalities for an LP-based branch-and-bound method. We note that p^ℓ and p^u are given by an iterative scheme instead of a single explicit formula. Therefore, we designed an adaptive approach, which has the advantage that the evaluation of p^ℓ and p^u is only needed "on demand". Furthermore, the number of grid points need not satisfy the condition (6) up front for the whole domain, but only in specific points. Moreover, we use an outer-approximation approach, which dynamically adds inequalities if they cut off the current solution of the relaxation.

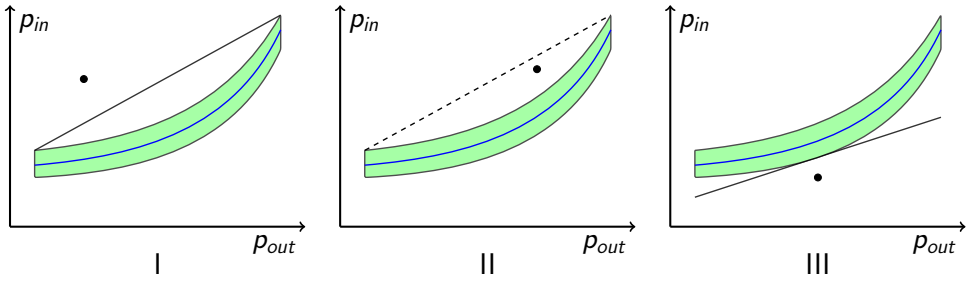

FIGURE 2. *Three different cases of infeasibility of a pair* (p_{in}, p_{out}) *for fixed mass flow rate.*

We proceed as in Algorithm 2, but adjusted to the gas model, see Algorithm 4. In the root node of the branch-and-bound tree we start with the minimal number of grid points N_a, such that the condition $\frac{L_a}{N_a} \leq 0.16\frac{D_a}{\lambda_a}$ is satisfied for all pipes $a \in \mathcal{A}_{pipe}$, and the variable bounds as the initial relaxation of the differential equation. In every other node, we use the relaxation of the parent node. Next, we solve the LP-relaxation and obtain a triple $(\tilde{p}_{in}, \tilde{p}_{out}, \tilde{q})$ for every pipe $a \in \mathcal{A}_{pipe}$. We then compute $p_a^\ell(\tilde{p}_{out}, \tilde{q}, N_a)$ and $p_a^u(\tilde{p}_{out}, \tilde{q}, N_a)$. If the difference does not satisfy (6), we increase N_a and recompute p_a^ℓ, p_a^u until they do (Lines 7 and 8). We then check whether the inequality

$$p_a^\ell(\tilde{p}_{out}, \tilde{q}, N_a) \leq \tilde{p}_{in} \leq p_a^u(\tilde{p}_{out}, \tilde{q}, N_a) \tag{28}$$

is satisfied for all pipes. If all triples are feasible, Algorithm 4 returns the current solution to the branch-and-bound process. In the case that at least one triple is not feasible, we pick the pipe a with the largest deviation of $\tilde{p}_{in}$ from the next bound p_a^ℓ, or p_a^u, see Line 11. If the flow direction of this pipe is not fixed, we perform branching with respect to $q_a = 0$ to fix the flow direction (Line 13); this step is particular to our context here. Otherwise, if the flow direction is already fixed, we try to cut off the solution. Thereby we distinguish three different cases, as shown in Figure 2.

In the first case (Line 15), $\tilde{p}_{in}$ is larger than the concave envelope of p_a^u. Here, we add a linear inequality which separates $\tilde{p}_{in}$ from the feasible region. In the second case, we have $p_a^u(\tilde{p}_{out}, \tilde{q}, N_a) \leq \tilde{p}_{in} \leq \hat{p}_a^u(\tilde{p}_{out}, \tilde{q}, N_a)$, i.e., we cannot cut off the current solution with the concave envelope. Instead, we have to resolve the infeasibility by branching. In the last case (Line 17), when $\tilde{p}_{in}$ is less than $p_a^\ell(\tilde{p}_{out}, \tilde{q}, N_a)$, we use the convexity of p_a^ℓ and cut off the solution with a gradient cut

$$p_{in} \geq p_a^\ell(\tilde{p}_{out}, \tilde{q}, N_a) + \nabla p_a^\ell(\tilde{p}_{out}, \tilde{q}, N_a)^\top \begin{pmatrix} p_{out} - \tilde{p}_{out} \\ q - \tilde{q} \end{pmatrix}. \tag{29}$$

In the first and last case, we then iterate and again solve the relaxation. In the second case, Algorithm 4 stops and instructs the branch-and-bound process to branch. After the convex relaxation algorithm terminates with a solution, we carry on like in the spatial branch-and-bound Algorithm 3.

PROPOSITION 4.9. *Algorithm 4 terminates after a finite number of iterations.*

Proof. We show that Assumption 3 holds for this example: Suppose that the vector of grid points $N \in \mathbb{N}^{\mathcal{A}_{pipe}}$ stays constant during the execution of Algorithm 4, i.e., the condition

$$\|p^u(p_{out}^k, q^k, N) - p^\ell(p_{out}^k, q^k, N)\|_\infty < \delta_2$$

Algorithm 4 Adaptive convex relaxation of gas flow

Input: Node of branch-and-bound tree, $\delta_1, \delta_2 > 0$ and $N = N_0 \in \mathbb{N}^{\mathcal{A}_{pipe}}$.
Output: δ_1-feasible solution of the Euler-equation, "infeasible" or "branch".

Choose initial convex relaxation: In the root node take the box $P \times Q$, else take the relaxation of the parent node.
For $k = 1, 2, \ldots$ **do**
  Solve the convex relaxation.
  If the relaxation is feasible **then**
    for each pipe $a \in \mathcal{A}_{pipe}$ let $\left(p^k_{in}, p^k_{out}, q^k\right)_a$ be the solution of the relaxation.
    for all $a \in \mathcal{A}_{pipe}$ **do**
      While $\left|p^u_a\left(p^k_{out}, q^k, N_a\right) - p^\ell_a\left(p^k_{out}, q^k, N_a\right)\right| > \delta_2$ **do**
        increase N_a.
    If all $\left(p^k_{in}, p^k_{out}, q^k\right)_a$ are δ_1-feasible for (28) **then**
      return the solution $(p^k_{in}, p^k_{out}, q^k)_a$.
    Choose "most infeasible" pipe $a \in \mathcal{A}_{pipe}$, i.e.,
    $a \in \arg\max_{a \in \mathcal{A}_{pipe}} \max\left\{p^k_{in} - p^u_a\left(p^k_{out}, q^k, N_a\right),\ p^\ell_a\left(p^k_{out}, q^k, N_a\right) - p^k_{in}\right\}$.
    If $\underline{q}_a < 0 < \overline{q}_a$ **then**
      fix orientation of flow on pipe a via branching.
    If $p^k_{in} > p^u_a\left(p^k_{out}, q^k, N_a\right)$ **then**
      try to cut off the solution with one inequality of the concave envelope; if this fails, then return "branch"
    else if $p^k_{in} < p^\ell_a\left(p^k_{out}, q^k, N_a\right)$ **then**
      add a gradient cut,
  else
    return "infeasible".

is satisfied for all produced pairs $\left(p^k_{out}, q^k\right)$. By Lemma 2.5 this shows the statement.

It suffices to only consider a single pipe. One can straightforwardly extend this proof to an arbitrary number of pipes.

Suppose that the algorithm does not terminate, that is, it produces an infinite sequence of points which are feasible for the convex relaxation but not δ_2-feasible for (28). Since the algorithm does not terminate, the orientation of flow already has to be fixed. Hence, we can distinguish between input pressure and output pressure.

Let $\left(p^k_{in}, p^k_{out}, q^k\right)_{k \in \mathbb{N}}$ denote the sequence of solutions produced by the algorithm. We divide the iterations into two sets. With $\mathcal{O} \subseteq \mathbb{N}$ we denote the set of iterations with $p^k_{in} > p^u(p^k_{out}, q^k, N)$, and with $\mathcal{L} = \mathbb{N} \setminus \mathcal{O}$ we denote the set of iterations with $p^k_{in} < p^\ell(p^k_{out}, q^k, N)$. We will show that both sets have to be finite and, therefore, the algorithm terminates after a finite number of iterations.

We first consider the subsequence $\mathcal{O}$. Since the function $p^u\left(p_{out}, q\right)$ is convex, the concave envelope over the feasible set

$$\mathcal{F} := \left[\underline{p}_{out}, \overline{p}_{out}\right] \times \left[\underline{q}, \overline{q}\right] \cap \left\{(p_{out}, q) \,\middle|\, 5c\, q \le 4A\, p_{out}\right\}$$

consists of at most three linear inequalities. Thus, after at most three iterations $k \in \mathcal{O}$ the concave envelope is fully added to the convex relaxation. Any further point $\left(p^k_{in}, p^k_{out}, q^k\right)$ with $k \in \mathcal{O}$ cannot be separated from p^u with a linear inequality. Thus, the algorithm would terminate with the instruction to branch. Therefore, $\mathcal{O}$ can have at most three elements (in general, three times the number of pipes).

Next, we show that the sequence $\mathcal{L}$ is finite, too. In every iteration $k \in \mathcal{L}$ we add

an inequality of the form (29) to the relaxation. Since p^ℓ is convex and continuously differentiable, it is Lipschitz continuous on the compact set $\mathcal{F}$. Hence, there is a radius r such that the inequality

$$0 \le p^\ell\big(p_{out}, q, N\big) - p^\ell\big(p^k_{out}, q^k, N\big) - \nabla p^\ell\big(p^k_{out}, q^k, N\big)^\top \begin{pmatrix} p_{out} - p^k_{out} \\ q - q^k \end{pmatrix} < \delta_1$$

holds for all $k \in \mathcal{L}$ and for all points $\big(p_{out}, q\big) \in B_r\big(p^k_{out}, q^k\big) \cap \mathcal{F}$, where $B_r\big(p^k_{out}, q^k\big)$ is the open ball around $\big(p^k_{out}, q^k\big)$ with radius r. That is, any point $\big(p_{out}, q\big) \in B_r\big(p^k_{out}, q^k\big) \cap \mathcal{F}$, which satisfies the gradient cut for iteration k, is δ_1-feasible for (28).

Now, suppose that $\mathcal{L}$ is not finite. Then there exists a subset $(k_i)_{i\in\mathbb{N}} \subset \mathcal{L}$ such that the sequence $(p^{k_i}_{out}, q^{k_i})_{i\in\mathbb{N}}$ converges to $(\bar{p}_{out}, \bar{q}) \in \mathcal{F}$. Let k_j be an element of the sequence, such that $(\bar{p}_{out}, \bar{q}) \in B_r\big(p^{k_j}_{out}, q^{k_j}\big)$ holds. Then, since the sequence converges to $(\bar{p}_{out}, \bar{q})$, there exists an index $n > j$ such that also $\big(p^{k_n}_{out}, q^{k_n}\big) \in B_r\big(p^{k_j}_{out}, q^{k_j}\big)$ holds. By the above argument this implies that $\big(p^{k_n}_{in}, p^{k_n}_{out}, q^{k_n}\big)$ is a δ_1-feasible for (28), i.e., the algorithm stops. Thus, $\mathcal{L}$ has to be finite. ∎

In order to apply Algorithm 3 and Theorem 2.6, it remains to show that Condition (11) holds.

PROPOSITION 4.10. *Let Conditions (7) and (10) be true. Suppose that Algorithm 3 applied to problem (23) produces an infinite nested sequence of nodes. Then the solutions of the convex relaxation produced by Algorithm 4 satisfy Condition (11).*

Proof. Again, we only consider a single pipe $a = (u, v)$. Suppose that Algorithm 3 produces an infinite nested sequence of bounding boxes

$$\mathcal{F}_k = \big[\underline{p}^k_u, \overline{p}^k_u\big] \times \big[\underline{p}^k_v, \overline{p}^k_v\big] \times \big[\underline{q}^k_a, \overline{q}^k_a\big]$$

and let (p^k_u, p^k_v, q^k_a) be the last solution of the relaxation produced by Algorithm 4 for node k. Since our first priority is to fix the direction of the flow, we can assume that q_a is restricted to nonnegative values. Then p_u is the inflow-pressure and p_v is the outflow-pressure.

We have to prove that there is a $\tilde{k} \in \mathbb{N}$ such that the condition

$$\max\Big\{\big(p^k_u - p^u_a(p^k_v, q^k_a, N_a)\big)_+,\ \big(p^\ell_a(p^k_v, q^k_a, N_a) - p^k_u\big)_+\Big\} \le \delta_1$$

holds for all $k \ge \tilde{k}$. Note that the algorithm only returns a point (p^k_u, p^k_v, q^k_a) with $p^\ell_a(p^k_v, q^k_a, N_a) - p^k_u > \delta_1$ if there is another pipe for which infeasibility cannot be resolved by adding a cut. Otherwise, by construction of Algorithm 4 a gradient cut, which cuts off the current solution, would be added. Hence, it suffices to show that $\tilde{k} \in \mathbb{N}$ with $p^k_u - p^u_a(p^k_v, q^k_a, N_a) \le \delta_1$ for all $k \ge \tilde{k}$ exists.

We will show that $\tilde{k} \in \mathbb{N}$ exists such that the concave envelope of p^u_a cuts off all δ_1-infeasible points. Since the relaxation contains the constraint $5c\,q_a \le 4A\,p_v$, we assume for simplicity of notation that $5c\,\underline{q}^k_a \le 4A\,\underline{p}^k_v$ and $5c\,\overline{q}^k_a \le 4A\,\overline{p}^k_v$ is true for all nodes $\mathcal{F}_k$. Then due to monotonicity, the inequality

$$p^u_a\big(\underline{p}^k_v, \underline{q}^k_a, N_a\big) \le p^u_a\big(p_v, q_a, N_a\big) \le p^u_a\big(\overline{p}^k_v, \overline{q}^k_a, N_a\big)$$

is fulfilled for all nodes k and all points (p_v, q_a), which are feasible for the relaxation. Therefore, the concave envelope $\hat{p}^u_a$ of p^u_a satisfies the inequality $\hat{p}^u_a\big(p_v, q_a\big) \le p^u_a\big(\overline{p}^k_v, \overline{q}^k_a, N_a\big)$. Hence, if

$$p^u_a\big(\overline{p}^k_v, \overline{q}^k_a, N_a\big) - p^u_a\big(\underline{p}^k_v, \underline{q}^k_a, N_a\big) \le \delta_1 \tag{30}$$

holds, δ_1-infeasible points can be cut off by the concave envelope, i.e., (p_u^k, p_v^k, q_a^k) is δ_1-feasible.

By construction of p_a^u it is continuous and converges to the solution of the ODE for $N_a \to \infty$, therefore N_a is only increased a finite number of times. Thus, by the continuity of p_a^u and Condition (7), i.e., $\lim_{k\to\infty} \operatorname{diam} \mathcal{F}_k = 0$, we can derive that an index $\tilde{k} \in \mathbb{N}$ exists such that inequality (30) holds for all $k \geq \tilde{k}$. Therefore, Condition (11) holds for all $k \geq \tilde{k}$. ∎

Lemma 4.9 and Proposition 4.10 show that our construction of under- and overestimators for the Euler-equation satisfies the necessary requirements of Theorem 2.6.

COROLLARY 4.11. *Suppose that Conditions (7) and (10) hold. Then for $\varepsilon > 0$ and $\delta_1, \delta_2 > 0$, Algorithm 3 applied to problem (23) terminates after a finite number of iterations with an $(\varepsilon, \delta_1 + \delta_2)$-optimal solution of (23) or the conclusion that the problem is infeasible.*

This Corollary shows that our approach and Algorithm 3 works for the example of stationary gas transport.

REMARK 4.12. *In our model we assumed that the friction factor is constant. However, it is possible to adapt Algorithm 4 such that we can handle a nonconstant friction factor, e.g., when using a model in which λ depends on q. Indeed, we can show that the functions $p^\ell(p, q, N)$ and $p^u(p, q, N)$ are nondecreasing with respect to λ. Hence, we can add additional constraints for the friction model and by assuming that λ is bounded, we can still derive lower and upper bounds on the input pressure by using the lower and upper bounds on λ. Then we can use the under- and overestimators as before to cut off infeasible solutions and use spatial branching on λ as well as the pressure and flow variables to improve the relaxations.*

4.4. Extension to Pipes with Height-Differences. So far we have considered horizontal pipes. In this case $p(x)$ is concave, see Corollary (4.2), the mapping $(p(L), q) \mapsto p(0)$ is convex and under the conditions of Lemma 4.7 the explicit midpoint rule (26a) yields a convex lower bound and the trapezoidal rule (26b) a convex upper bound.

We discuss now the more general case of nonzero slope. We will show that the Taylor scheme (18) and the trapezoidal rule (20) can still be used to obtain convex upper and lower bounds, although $p(x)$ is not necessarily convex in this case.

Recall that the Euler equation (22) is given by

$$\partial_x p(x) \left(1 - \frac{c^2 q^2}{A^2 p(x)^2}\right) = -\frac{\lambda c^2}{2DA^2} q|q| \frac{1}{p(x)} - \frac{g}{c^2} \sigma\, p(x), \quad 0 < x < L,$$

with a constant slope $\sigma \in [-1, 1]$. We still assume that $q \geq 0$ and $5c\,q \leq 4A\,p$ holds. Then we can rewrite the equation to

$$\partial_x p(x) = \varphi_\sigma\big(p(x), q\big), \quad 0 < x < L, \quad \varphi_\sigma(p, q) \coloneqq -\frac{1}{2} \frac{p}{c^2 D} \frac{2Dg\sigma A^2 p^2 + \lambda c^4 q^2}{A^2 p^2 - c^2 q^2}.$$

In the case of a nonnegative slope $\sigma \geq 0$, the right-hand side is always negative. Otherwise, if $\sigma < 0$, the right-hand side has the root $p_r(q, \sigma) \coloneqq \frac{c^2 q}{A} \sqrt{\frac{-\lambda}{2Dg\sigma}}$, and $\varphi_\sigma(p, q) \geq 0$ for $p \geq p_r(q, \sigma)$ and $\varphi_\sigma(p, q) \leq 0$ for $p \leq p_r(q, \sigma)$ holds.

Figure 3 shows the change in pressure of gas flowing along a pipe with positive slope on the left and a pipe with negative slope on the right. For $\sigma > 0$ the gas has

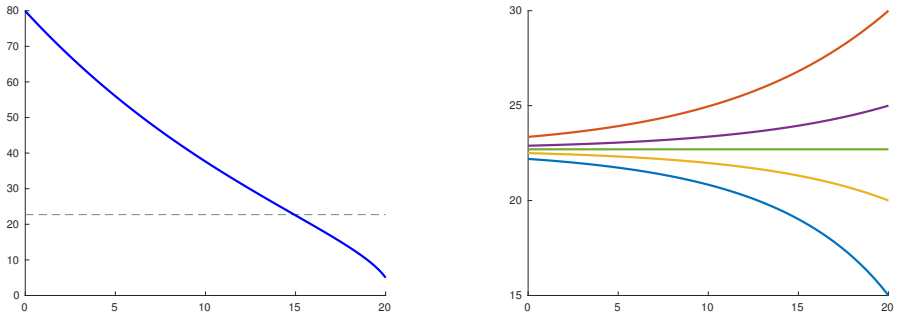

FIGURE 3. *The figure shows the pressure $p(x)$ in* bar *along a* 20 km *pipe. The left figure depicts $p(x)$ for one initial value and positive slope, and the right figure shows $p(x)$ for different initial values and negative slope.*

to compensate friction and gravitation, which results in an increased pressure drop in comparison with the case $\sigma = 0$. In the case $\sigma < 0$, the gravitation works contrary to the friction, such that there is no pressure drop or increase if the input pressure equals $p_r(q, \sigma)$ (see the middle line on the right of Figure 3). If the input pressure is less than the root, the pressure drop is less than in the case of $\sigma = 0$. Furthermore, if the input pressure is larger than the root, the pressure actually increases in flow direction.

Analogously to the proof of Corollary 4.3, we consider the ODE

$$\tilde{p}(0) = p^0, \quad \partial_x \tilde{p}(x) = -\varphi_\sigma\big(\tilde{p}(x), q\big), \quad 0 \le x \le L$$

to show that we can compute lower and upper bounds on the input pressure $p(0) = \tilde{p}(L)$. Application of the Taylor method yields

$$p_0^\ell = p^0, \qquad p_{i+1}^\ell = p_i^\ell - h\big(\varphi_\sigma(p_i^\ell, q) - \tfrac{h}{2}(\partial_p \varphi_\sigma\, \varphi_\sigma)(p_i^\ell, q)\big), \qquad \forall i = 0, \dots, N-1,$$

and application of the trapezoidal rule is given by (26b) with φ_σ instead of φ.

LEMMA 4.13. *Let $c^2\lambda \ge -2Dg\sigma$. Then $-\varphi_\sigma(p, q)$ is convex on $\mathcal{D}$. Moreover, if $c^2\lambda \ge 2Dg\sigma$ then $g_h(p, q) := -\varphi_\sigma(p, q) + \frac{h}{2}(\partial_p \varphi_\sigma \varphi_\sigma)(p, q)$ is convex on $\mathcal{D}$ for $0 < h \le 0.078\frac{D}{\lambda}$ if $\sigma \le 0$ and $0 < h \le 0.035\frac{D}{\lambda}$ if $\sigma > 0$.*

Proof. $D^2\varphi_\sigma(p, q)$ is singular and it is easy to check that $-\partial_{pp}\varphi_\sigma(p, q) \ge 0$ on $\mathcal{D}$ if $c^2\lambda \ge -2Dg\sigma$. Hence, $-D^2\varphi_\sigma(p, q)$ is positive semidefinite on $\mathcal{D}$.

Also $D^2 g_h(p, q)$ is singular. Moreover, $\partial_{pp} g_h(p, q)$ is a rational function with positive denominator on $\mathcal{D}$. The numerator is a linear function in h and is nonnegative for $0 < h \le 0.078\frac{D}{\lambda}$ if $\sigma \le 0$ and $0 < h \le 0.035\frac{D}{\lambda}$ if $\sigma > 0$. ∎

We analyze the following three cases:

1. negative slope $\sigma < 0$, with $p^0 \le p_r(q, \sigma)$, or $\sigma = 0$,
2. negative slope $\sigma < 0$, with $p^0 \ge p_r(q, \sigma)$,
3. positive slope $\sigma > 0$.

LEMMA 4.14.

1. *If $\sigma \le 0$ with $p^0 \le p_r(q, \sigma)$, or $\sigma = 0$, and $\lambda c^2 \ge -2Dg\sigma$, then for $0 < h \le \frac{81}{164}\frac{D}{\lambda}$ the trapezoidal rule produces a convex upper bound $p^u(p, q, N)$ and for $0 < h \le 0.078\frac{D}{\lambda}$ the Taylor method yields a convex lower bound $p^\ell(p, q, N)$ on $\mathcal{D}$.*

2. *If* $\sigma < 0$, $p^0 \geq p_r(q,\sigma)$ *and* $\lambda c^2 \geq -3Dg\sigma$ *then for* $0 < h \leq \frac{D}{\lambda}$ *the trapezoidal rule produces a convex lower bound* $p^\ell(p,q,N)$ *and for* $0 < h \leq 0.078\frac{D}{\lambda}$ *the Taylor method yields a convex upper bound* $p^u(p,q,N)$ *on* $\mathcal{D}$.
3. *If* $\sigma > 0$ *and* $\lambda c^2 \geq 2Dg\sigma$ *then for* $0 < h \leq \frac{162D}{1231\lambda}$ *the trapezoidal rule produces a convex upper bound* $p^u(p,q,N)$ *and for* $0 < h \leq 0.035\frac{D}{\lambda}$ *the Taylor method yields a convex lower bound* $p^\ell(p,q,N)$ *on* $\mathcal{D}$.

Proof. We have

$$\partial_{xx}\tilde{p}(x) = (\partial_p\varphi_\sigma\,\varphi_\sigma)(\tilde{p}(x),q), \quad \partial_{xxx}\tilde{p}(x) = -(\partial_{pp}\varphi_\sigma\,\varphi_\sigma^2 + (\partial_p\varphi_\sigma)^2\,\varphi_\sigma)(\tilde{p}(x),q).$$

In all cases $-\varphi_\sigma(p,q)$ is convex on $\mathcal{D}$ and $-\varphi_\sigma(p,q)+\frac{h}{2}(\partial_p\varphi_\sigma\varphi_\sigma)(p,q)$ is convex under the above bounds $0 < h \leq 0.078\frac{D}{\lambda}$ or $0 < h \leq 0.035\frac{D}{\lambda}$.

Ad 1.): This case is analogous to the analysis of $\sigma = 0$. We have that $\varphi_\sigma(p,q) \leq 0$, $\partial_p\varphi_\sigma(p,q) \geq 0$ and $\partial_{pp}\varphi_\sigma(p,q) \leq 0$ hold, therefore $\partial_{xx}\tilde{p}(x) \leq 0$ and $\partial_{xxx}\tilde{p}(x) \geq 0$ also holds, i.e., $\tilde{p}(x)$ is concave and $\partial_x\tilde{p}(x)$ is convex. As in Corollary 4.3 we have $0 \geq -\partial_p\varphi_\sigma\big(p,q\big) \geq -\partial_p\varphi_\sigma\big(\frac{5cq}{4A},q\big) = -\frac{328}{81}\frac{\lambda}{D} =: b$. Hence, the trapezoidal rule produces by Corollary 3.11 and Lemma 3.12 a convex upper bound for $0 < h \leq \frac{81}{164}\frac{D}{\lambda}$. Moreover, since $g(p,q) := (\partial_{pp}\varphi_\sigma\,\varphi_\sigma + (\partial_p\varphi_\sigma)^2)(p,q) \geq 0$, Corollary 3.9 holds with $b_1 = 0$ and yields together with Lemma 3.10 that the Taylor method gives a convex lower bound for $0 < h \leq 0.078\frac{D}{\lambda}$.

Ad 2.): Here, one has $\partial_p\varphi_\sigma(p,q) \geq 0$, $\partial_{pp}\varphi_\sigma(p,q) \leq 0$ and $\varphi_\sigma(p,q) \geq 0$. Hence, $\tilde{p}(x)$ is convex and $\partial_p\varphi_\sigma$ is bounded by $0 \geq -\partial_p\varphi_\sigma\big(p,q\big) \geq -\partial_p\varphi_\sigma\big(p_r(q,\sigma),q\big) = \frac{2\lambda g\sigma}{2Dg\sigma+\lambda c^2} \geq -\frac{2\lambda}{D}$ for $\lambda c^2 \geq -3Dg\sigma$. Furthermore, $\partial_x\tilde{p}(x)$ is concave. In fact, since $\varphi_\sigma(p,q) \geq 0$ it suffices to show $g(p,q) \geq 0$. It is easy to verify that $g(p,q) > 0$ for all $p > 0$ large enough. On the other hand, $\partial_p g(p,q)$ is a rational function with positive denominator on $\mathcal{D}$ and negative numerator if $\lambda c^2 \geq 2Dg\sigma$. This shows that $\partial_p g(p,q) \leq 0$ and thus $g(p,q) \geq 0$. Hence, the trapezoidal rule yields by Corollary 3.11 and Lemma 3.12 a convex lower bound for $0 < h \leq \frac{D}{\lambda}$. Moreover, since $g(p,q) \geq 0$ Corollary 3.9 holds with $b_1 = 0$ and yields together with Lemma 3.10 that the Taylor method produces a convex upper bound for $0 < h \leq 0.078\frac{D}{\lambda}$.

Ad 3.): Then $\varphi_\sigma(p,q) < 0$ and $\partial_{pp}\varphi_\sigma(p,q) \leq 0$ and thus $\partial_{xxx}\tilde{p}(x) \geq 0$, i.e., $\partial_x\tilde{p}(x)$ is convex. However, $\partial_p\varphi_\sigma(p,q)$ may change sign and thus $\tilde{p}(x)$ may change form concave to convex, see the left of Figure 3. From the second derivatives we can derive that $\partial_p\varphi_\sigma(p,q)$ is bounded by

$$-\tfrac{1231\lambda}{162D} \leq -\tfrac{328\lambda}{81D} - \tfrac{575g\sigma}{81c^2} = -\partial_p\varphi_\sigma\big(\tfrac{5cq}{4A},q\big) \leq -\partial_p\varphi_\sigma(p,q) \leq -\partial_p\varphi_\sigma(p,0) \leq \tfrac{g\sigma}{c^2} \leq \tfrac{\lambda}{2D}.$$

Hence, the trapezoidal rule produces by Corollary 3.11 and Lemma 3.12 a convex upper bound for $0 < h \leq \frac{162D}{1231\lambda}$. Moreover, since $g(p,q) \geq 0$ Corollary 3.9 holds with $b_1 = 0$ and shows together with Lemma 3.10 that the Taylor method produces a convex lower bound for $0 < h \leq 0.035\frac{D}{\lambda}$. ∎

5. Numerical Results. In this section, we demonstrate the behavior of the methods discussed in this paper on two examples of stationary gas networks without heights. To this end, we implemented our approach using the LP-based branch-and-cut framework SCIP version 5.0, see [11, 41]. We keep the description of the implementation extremely short and refer to [33, 21] for a detailed discussion of modeling issues in stationary gas transport.

Recall from Section 4 that a gas network is given by the graph $\mathcal{G} = \big(\mathcal{V},\mathcal{A}\big)$. The mass flows q_a, $a \in \mathcal{A}$, have to satisfy flow conservation (24). Moreover, flow and

pressures have to satisfy given lower and upper bounds ($\underline{q}_a \le q_a \le \overline{q}_a$, $\underline{p}_v \le p_v \le \overline{p}_v$). We consider several different network elements, which are handled as follows:

- Pipes are handled in the way described in Section 4. We determine the friction coefficient λ with the formula of Nikuradse [29, 30], i.e.,

$$\lambda = \left(2\log_{10}\left(\frac{D}{k}\right) + 1.138\right)^{-2},$$

  where k is the roughness of the pipe.
- Valves are modeled using binary variables. For open valves, the pressures on both sides are equal. For closed valves, the flow is 0 and the pressures are decoupled.
- Compressors allow to increase gas pressure. Compressor stations consist of several compressors that are connected by piping and valves. We use a rather simple model in which we approximate the operation states by a polyhedron, see Hiller et al. [17]. Two binary variables are used to decide whether the compressor is turned on/off or if its bypass is open/closed; the bypass is used to allow flow bypassing the compressor as well as using the opposite direction. The possible values are: compressor on/bypass closed, compressor off/bypass open and compressor off/bypass closed.
- Control valves and resistors are modeled as described in [33, 21].

As objective function, we consider the following options:

- Minimize the number of compressors running. This can be seen as a proxy for the energy used. Note that we cannot express the consumed energy to run the compressors with the currently used compressor model.
- Maximize the sum of all pressures.
- Minimize the power loss, i.e., we minimize the function

$$\sum_{a=(u,v)\in\mathcal{A}} (p_u - p_v)\, q_a = \sum_{v\in\mathcal{V}} \left(\sum_{a=(v,w)\in\mathcal{A}} q_a - \sum_{a=(u,v)\in\mathcal{A}} q_a \right) p_v = \sum_{v\in\mathcal{V}} q_v^{\pm}\, p_v.$$

  Note that the change in pressure times the flow is proportional to the change in energy.

Furthermore, we included binary variables, which indicate the flow directions, in our computational model. That is, for every pipe $a \in \mathcal{A}$, we introduced variables $s_a^-, s_a^+ \in \{0,1\}$, and coupled them with the flow variables, i.e., we added the inequality

$$\underline{q}_a \cdot s_a^- \le q_a \le \overline{q}_a \cdot s_a^+.$$

We use these variables to strengthen the initial relaxation by coupling the flow direction with the pressure difference of the incident nodes, i.e., if the flow on pipe $a = (u, v)$ is nonnegative, then the difference $p_u - p_v$ has to be nonnegative as well. Additionally, due to the gas physics it is not possible that gas flows in a cycle, unless there is a compressor involved. Hence, we compute a cycle basis of the graph and we strengthen the initial relaxation by adding inequalities that forbid flow in a cycle, on those cycles in the basis, which do not contain compressors.

We implemented bound propagation based on the numerical methods (26a) and (26b). Since the input pressure is nondecreasing in output pressure and mass flow, we can derive an upper bound on the input pressure by computing $p_{in}^u(\overline{p}_{out}, \overline{q}_a)$. Similarly, $p_{in}^\ell(\underline{p}_{out}, \underline{q}_a)$ defines a lower bound on the input pressure, if $\underline{q}_a \ge 0$ holds. Note that we can also apply the methods in the direction of the flow and try to compute lower and

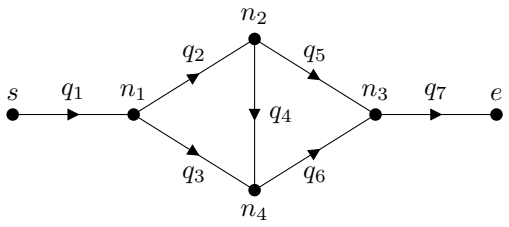

FIGURE 4. *A small network with one entry (s) on the left, one exit (e) on the right and two overlapping circles.*

upper bounds on the output pressure. Thereby, the bounds produced by the methods are reversed, i.e., the explicit midpoint method now produces upper bounds and the trapezoidal rule produces lower bounds. During this propagation step, we also check whether $p_i = p_i^\ell$ or $p_i = p_i^u$ violate the inequality $5c\,\overline{q}_a \le 4A\,p_i$. In the worst case, when propagating the lower input pressure, it can happen that $c\,\overline{q}_a > A\,p_i^\ell$ holds and p_N^u is a feasible output pressure. Then $p_{i+1}^\ell > p_i^\ell$ follows, although the pressure is nonincreasing! In this case, we cannot strengthen bounds and neither decide whether $(\underline{p}_{in}, \overline{q}_a)$ is feasible. This may happen in the beginning of the algorithm, since the bounds are large and the inequality $5c\,q \le 4A\,p$ may have a nonempty intersection with the current feasible region; in the final solution, however, the slack of this inequality will be large.

All computations are performed with a precision of $\delta_1 = \varepsilon = 10^{-6}$ (SCIP default values) and $\delta_2 = 10^{-4}$, see Corollary 4.11. The feasibility tolerance δ_1 is used for all constraints and therefore is dimensionless. In contrast, δ_2 is used to check if the lower and upper bounds for the input pressure are within a range of 10^{-4} bar.

5.1. Diamond Graph. In the first example, the graph has a diamond shape, see Figure 4. It consists of one entry s, one exit e, four other nodes $n_1, \dots, n_4$ and seven pipes. The pipes vary in diameter and length. The diameter is either 1 m or 1.30 m and the length varies between 14 km and 40 km. The roughness is $k = 0.01$ mm. The pipes were discretized into at least 735 and at most 1893 steps, which led to step sizes from 19.8 m to 26.9 m.

During presolving, the flows q_1 and q_7 are fixed and the bounds of the other flows are improved, in particular, the flow directions of q_2 and q_3 are fixed. The flows q_3 and q_6 are replaced by $q_1 - q_2$ and $q_7 - q_5$, which can be done due to the flow conservation in the nodes n_1 and n_3 and because the flows q_1 and q_7 are already determined by the in- and outflow at s and e. Furthermore, all lower pressure bounds are improved by the propagation process described above. In contrast, the upper pressure bounds except on the entry s are not improved, because the bounds of $q_2, \dots, q_6$ are too large, and positive as well as negative flow is possible in the beginning. Nevertheless, the propagation process induced a lot of bound changes, when maximizing pressures, a pressure bound was improved 798 times, while 536 times when minimizing the power loss.

The branch-and-bound tree had 17 nodes with 8 leaves when maximizing the pressure and only 5 nodes with 3 leaves while minimizing the power loss. In the first case, branching mostly occurred for fixing the flow directions. It took only two branching steps on a pressure variable and one branching step on nonzero flow. The ODEs on the pipes are approximated by 27 gradient cuts of the form (29), varying from 0 on pipe 1 to 10 on pipe 2, and 29 overestimators, varying from 0 to 14 on each

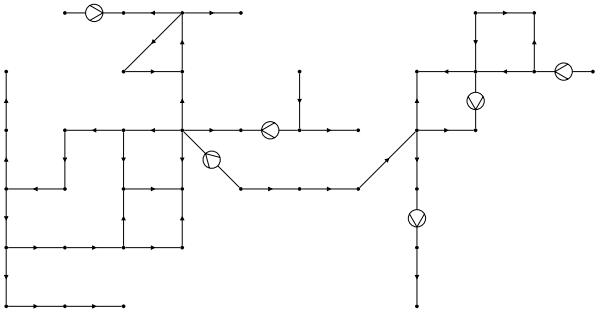

FIGURE 5. *Picture of the GasLib-40, where the compressor stations are depicted as a circle with $>$ inside.*

pipe. In the second case, branching occurred only for fixing the flow directions. Here, 14 gradient and 6 overestimating cuts were applied. Overall, this example was solved in less than one second.

5.2. GasLib-40. The second example is the GasLib-40 instance [39], see Figure 5. This instance consists of 40 nodes, 39 pipes and 6 compressors. Thus, it contains 12 binary variables that determine whether a compressor is turned on or is in bypass. For this instance, the roughness varies from 0.012 mm to 0.05 mm, the length varies from 3 km to 86 km and the diameter varies between 40 cm and 1 m.

Table 1 shows some statistics on the work of Algorithm 3 for the three different objective functions. All instances can be solved in roughly one minute. As one can see, branching on the flow directions, i.e., branching with respect to $q \leq 0$ and $q \geq 0$, seems to be most important, while branching on pressure and nonzero flow only accounts for a small portion of all branchings. This can somewhat be explained since over- and underestimators can only be added when the flow direction is already fixed. Then bound propagation, over- and underestimators yield a good approximation of the ODE solutions. This is corroborated by the noticeable high number of bound propagations when minimizing the power loss, where no branching on pressure variables or nonzero flow occurred. Furthermore, branching on binary variables amounts only to a small part of the branching steps, because branching on them is preferred over continuous variables. Therefore, branching on the binary variables happens early in the tree and will not be repeated in the deeper regions.

Bound propagation in the tree plays a much more important role than in the example before. Here, in some cases a node can be cut off due to bound propagation, e.g., when propagating an upper pressure bound yields a value lower than the lower bound.

The diamond graph and the GasLib-40 with minimization of the power loss suggest that adding underestimators is more important than adding overestimators. This is contradicted by the GasLib-40 with the other objectives. So it is not clear which cuts have more impact on the performance.

The discretizations, used to compute (26a) and (26b), are initialized with a minimum number of 179 grid points and at most 7961 grid points on a single pipe. These numbers correspond to step sizes between 5.12 m and 19.2 m. The minimal step size encountered during the solving process was 3.79 m. The mean step size of the final discretizations was 9.91 m. Overall, the network was subdivided by 112 250 grid points, when minimizing the power loss. That is, 112 211 additional pressure variables would have been needed in a discretize-then-optimize approach with the same precision.

TABLE 1
Statistics on the GasLib-40 instance for the three different objective functions.

objective: minimize	$-\sum_{v\in\mathcal{V}} p_v$	$\sum_{v\in\mathcal{V}} q_v^{\pm} p_v$	$\sum_{a\in\mathcal{A}_{cs}} z_a$
solving time (seconds)	59.15	30.23	64.41
processed Nodes	79	16	182
branchings on flow (%)	7.18	0.0	25.95
branchings on flow directions (%)	72.23	74.75	47.84
branchings on pressure (%)	3.43	0.0	10.94
branchings on binary variables (%)	17.16	25.25	15.27
leaves	32	7	72
cut offs by propagation	29	13	45
bound changes by propagation	7160	17 528	9152
added overestimators	321	119	768
added underestimators	218	366	179

Note that we can use much coarser discretizations if we use a smaller bound on the ratio $\frac{v}{c} = \frac{cq}{Ap}$, see Remark 4.5 and Remark 4.8. If we use $\frac{cq}{Ap} \leq 0.4$, then the initial step sizes would vary between 150 m and 570 m and the number of discretization steps per pipe would vary between 6 and 259. The bound $\frac{cq}{Ap} \leq 0.2$, which is also sufficiently big for our applications, would lead to step sizes between 875 m and 3.4 km.

6. Outlook. In this paper, we have investigated an adaptive method to construct relaxations for mixed-integer optimization problems with ODE constraints and its integration in a branch-and-bound framework. The method relies on the fact that the values of the ODE solution are only required at a priori fixed positions. If the ODE has favorable properties – like in water or gas transport – we have derived an effective way to produce lower and upper bounds based on discretization methods, leading to finite convergence.

For gas transport, the interplay of branching on integer variables and spatial branching needs further investigation. Finally, the implementation can be extended by considering more elaborate compressor models and additional presolving techniques. The performance could be improved by additional branching rules and primal heuristics.

It would be interesting to extend the approach to instationary network problems, which leads to mixed-integer PDE-constrained optimal control problems.

Acknowledgement. We thank Benjamin Hiller and Tom Walther for computing the polyhedral approximations of compressor stations and two anonymous reviewers whose comments helped to improve the presentation and quality of this paper.

REFERENCES

[1] C. S. ADJIMAN, I. P. ANDROULAKIS, AND C. A. FLOUDAS, *A global optimization method, αBB, for general twice differentiable NLPs – II. Implementation and computational results*, Comput. Electr. Eng., 22 (1998), pp. 1159–1179.

[2] C. S. ADJIMAN, S. DALLWIG, C. A. FLOUDAS, AND A. NEUMAIER, *A global optimization method, αBB, for general twice differentiable NLPs – I. Theoretical advances*, Comput. Electr. Eng., 22 (1998), pp. 1137–1158.

[3] H. G. BOCK, C. KIRCHES, A. MEYER, AND A. POTSCHKA, *Numerical solution of optimal control problems with explicit and implicit switches*, Optimization Methods and Software,

33 (2018), pp. 450–474, https://doi.org/10.1080/10556788.2018.1449843.

[4] C. Buchheim, R. Kuhlmann, and C. Meyer, *Combinatorial optimal control of semilinear elliptic PDEs*, Computational Optimization and Applications, 70 (2018), pp. 641–675, https://doi.org/10.1007/s10589-018-9993-2.

[5] B. Chachuat, A. B. Singer, and P. I. Barton, *Global mixed-integer dynamic optimization*, AIChE Journal, 51 (2005), pp. 2235–2253.

[6] M. Čižniar, M. Podmajerský, T. Hirmajer, M. Fikar, and A. M. Latifi, *Global optimization for parameter estimation of differential-algebraic systems*, Chemical Papers, 63 (2009), pp. 274–283, https://doi.org/10.2478/s11696-009-0017-7.

[7] H. Diedam and S. Sager, *Global optimal control with the direct multiple shooting method*, Optim. Control Appl. Methods, (2017), pp. 1–22.

[8] W. R. Esposito and C. A. Floudas, *Deterministic global optimization in nonlinear optimal control problems*, Journal of Global Optimization, 17 (2000), pp. 97–126, https://doi.org/10.1023/A:1026578104213.

[9] C. A. Floudas and C. E. Gounaris, *A review of recent advances in global optimization*, J. Glob. Optim., 45 (2008), pp. 3–38.

[10] A. Fügenschuh and I. Vierhaus, *A global approach to the optimal control of system dynamics models*, in Proc. 31st Inter. Conf. System Dynamics Society, 2013.

[11] A. Gleixner, L. Eifler, T. Gally, G. Gamrath, P. Gemander, R. L. Gottwald, G. Hendel, C. Hojny, T. Koch, M. Miltenberger, B. Müller, M. E. Pfetsch, C. Puchert, D. Rehfeldt, F. Schlösser, F. Serrano, Y. Shinano, J. M. Viernickel, S. Vigerske, D. Weninger, J. Witt, and J. Witzig, *The SCIP Optimization Suite 5.0*, tech. report, Optimization Online, 2017, http://www.optimization-online.org/DB_HTML/2017/12/6385.html.

[12] 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*, Computational Optimization and Applications, 70 (2018), pp. 267–294, https://doi.org/10.1007/s10589-017-9970-1.

[13] M. Gugat, G. Leugering, A. Martin, M. Schmidt, M. Sirvent, and D. Wintergerst, *Towards simulation based mixed-integer optimization with differential equations*, Networks, 72 (2018), pp. 60–83, https://doi.org/10.1002/net.21812.

[14] F. M. Hante, G. Leugering, A. Martin, L. Schewe, and M. Schmidt, *Challenges in optimal control problems for gas and fluid flow in networks of pipes and canals: From modeling to industrial applications*, in Industrial Mathematics and Complex Systems: Emerging Mathematical Models, Methods and Algorithms, P. Manchanda, R. Lozi, and A. H. Siddiqi, eds., Springer, Singapore, 2017, pp. 77–122, https://doi.org/10.1007/978-981-10-3758-0_5.

[15] F. M. Hante and S. Sager, *Relaxation methods for mixed-integer optimal control of partial differential equations*, Comput. Optim. Appl., 55 (2013), pp. 197–225.

[16] R. Hemmecke, M. Köppe, J. Lee, and R. Weismantel, *Nonlinear integer programming*, in 50 Years of Integer Programming 1958–2008 – From the Early Years to the State-of-the-Art, M. Jünger, T. M. Liebling, D. Naddef, G. L. Nemhauser, W. R. Pulleyblank, G. Reinelt, G. Rinaldi, and L. A. Wolsey, eds., Springer, 2010, pp. 561–618.

[17] B. Hiller, R. Saitenmacher, and T. Walther, *Analysis of operating modes of complex compressor stations*, in Operations Research Proceedings 2016, A. Fink, A. Fügenschuh, and M. J. Geiger, eds., Cham, 2018, Springer, pp. 251–257, https://doi.org/10.1007/978-3-319-55702-1_34.

[18] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich, *Optimization with PDE constraints*, vol. 23 of Mathematical Modelling: Theory and Applications, Springer, New York, 2009.

[19] R. Horst and H. Tuy, *Global Optimization: Deterministic Approaches*, Springer, Berlin, 3 ed., 1996.

[20] M. N. Jung, G. Reinelt, and S. Sager, *The Lagrangian relaxation for the combinatorial integral approximation problem*, Optim. Methods Softw., 30 (2015), pp. 54–80.

[21] T. Koch, B. Hiller, M. E. Pfetsch, and L. Schewe, eds., *Evaluating Gas Network Capacities*, SIAM, Philadelphia, PA, 2015.

[22] J. Lee and S. Leyffer, eds., *Mixed Integer Nonlinear Programming*, vol. 154 of IMA Volumes in Mathematics and its Applications, Springer-Verlag, New York, 2012.

[23] Y. Lin, J. A. Enszer, and M. A. Stadtherr, *Enclosing all solutions of two-point boundary value problems for ODEs*, Comput. Electr. Eng., 32 (2008), pp. 1714–1725.

[24] Y. Lin and M. A. Stadtherr, *Validated solutions of initial value problems for parametric ODEs*, Appl. Numer. Math., 57 (2007), pp. 1145–1162.

[25] M. Locatelli and F. Schoen, *Global Optimization*, SIAM, Philadelphia, PA, 2013.

[26] G. P. McCormick, *Computability of global solutions to factorable nonconvex programs: Part I — convex underestimating problems*, Math. Program., 10 (1976), pp. 147–175.
[27] N. Nedialkov, K. Jackson, and G. Corliss, *Validated solutions of initial value problems for ordinary differential equations*, Appl. Math. Comput., 105 (1999), pp. 21–68.
[28] M. Neher, K. R. Jackson, and N. S. Nedialkov, *On Taylor model based integration of ODEs*, SIAM J. Numer. Anal., 45 (2007), pp. 236–262.
[29] J. Nikuradse, *Strömungsgesetze in rauhen Rohren*, Forschungsheft auf dem Gebiete des Ingenieurwesens, VDI-Verlag, Düsseldorf, 1933.
[30] J. Nikuradse, *Laws of Flow in Rough Pipes*, vol. Technical Memorandum 1292, National Advisory Committee for Aeronautics Washington, 1950.
[31] I. Papamichail and C. S. Adjiman, *A rigorous global optimization algorithm for problems with ordinary differential equations*, J. Glob. Optim., 24 (2002), pp. 1–33, https://doi.org/10.1023/A:1016259507911.
[32] I. Papamichail and C. S. Adjiman, *Proof of convergence for a global optimization algorithm for problems with ordinary differential equations*, Journal of Global Optimization, 33 (2005), pp. 83–107, https://doi.org/10.1007/s10898-004-6100-2.
[33] 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*, Optim. Methods Softw., 30 (2015), pp. 15–53.
[34] R. Z. Ríos-Mercado and C. Borraz-Sánchez, *Optimization problems in natural gas transportation systems: A state-of-the-art review*, Applied Energy, 147 (2015), pp. 536–555.
[35] S. Sager, H. G. Bock, and G. Reinelt, *Direct methods with maximal lower bound for mixed-integer optimal control problems*, Math. Program., 118 (2009), pp. 109–149.
[36] S. Sager, M. Jung, and C. Kirches, *Combinatorial integral approximation*, Math. Methods Oper. Res., 73 (2011), pp. 363–380.
[37] A. Sahlodin and B. Chachuat, *Convex/concave relaxations of parametric ODEs using Taylor models*, Comput. Electr. Eng., 35 (2011), pp. 844–857.
[38] A. M. Sahlodin and B. Chachuat, *Discretize-then-relax approach for convex/concave relaxations of the solutions of parametric ODEs*, Appl. Numer. Math., 61 (2011), pp. 803–820.
[39] 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*, Data, 2 (2017), https://doi.org/10.3390/data2040040. Article 40.
[40] M. Schmidt, M. Sirvent, and W. Wollner, *A decomposition method for MINLPs with Lipschitz continuous nonlinearities*, Mathematical Programming, (2018), https://doi.org/10.1007/s10107-018-1309-x. To appear.
[41] SCIP, *Solving Constraint Integer Programs*. http://scip.zib.de/.
[42] J. K. Scott and P. I. Barton, *Convex and concave relaxations for the parametric solutions of semi-explicit index-one differential-algebraic equations*, J. Optim. Theory Appl., 156 (2013), pp. 617–649.
[43] J. K. Scott and P. I. Barton, *Improved relaxations for the parametric solutions of ODEs using differential inequalities*, J. Glob. Optim., 57 (2013), pp. 143–176.
[44] J. K. Scott, B. Chachuat, and P. I. Barton, *Nonlinear convex and concave relaxations for the solutions of parametric ODEs*, Optim. Control Appl. Methods, 34 (2013), pp. 145–163.
[45] J. K. Scott, M. D. Stuber, and P. I. Barton, *Generalized McCormick relaxations*, Journal of Global Optimization, 51 (2011), pp. 569–606.
[46] A. B. Singer and P. I. Barton, *Bounding the solutions of parameter dependent nonlinear ordinary differential equations*, SIAM J. Sci. Comput., 27 (2006), pp. 2167–2182.
[47] A. B. Singer and P. I. Barton, *Global optimization with nonlinear ordinary differential equations*, J. Glob. Optim., 34 (2006), pp. 159–190.
[48] I. Vierhaus and R. Gottwald, *SD-SCIP – system dynamics SCIP: A SCIP plug-in for solving system dynamics optimization problems*. http://sdscip.zib.de/, 2017.
[49] M. E. Villanueva, B. Houska, and B. Chachuat, *Unified framework for the propagation of continuous-time enclosures for parametric nonlinear ODEs*, J. Glob. Optim., 62 (2015), pp. 575–613.