GasLib – A Library of Gas Network Instances

Martin Schmidt[1,2], Denis Assmann[1], Robert Burlacu[1], Jesco Humpola[3], Imke Joormann[4], Nikolaos Kanelakis[5] Thorsten Koch[3], Djamal Oucherif[6], Marc E. Pfetsch[7], Lars Schewe[1], Robert Schwarz[3], Mathias Sirvent[1]

Abstract. The development of mathematical simulation and optimization models and algorithms for solving gas transport problems is an active field of research. In order to test and compare these models and algorithms, gas network instances together with demand data are needed. The goal of GasLib is to provide a set of publicly available gas network instances that can be used by researchers in the field of gas transport. The advantages are that researchers save time by using these instances and that different models and algorithms can be compared on the same specified test sets. The library instances are encoded in an XML format. In this paper, we explain this format and present the instances that are available in the library.

1. Summary

The mathematical simulation and optimization of gas transport through pipeline systems is an important field of research with a large practical impact. Over the last decades many different mathematical models on different levels of accuracy for different components of gas networks have been developed. Based on these models, several simulation and optimization algorithms have been proposed. We refer to Koch et al. (2015), Pfetsch et al. (2015), and Ríos-Mercado and Borraz-Sánchez (2015) and the references therein for more information. With GasLib we provide a set of network instances that can be used to test and compare such models and the algorithms for solving them.

Gas networks are used to transport gas from entry nodes, where gas is supplied, to exit nodes, where gas is discharged from the network. The set of supplied and discharged flows, together with additional specifications on the supplied and discharged gas pressure and the chemical composition of the supplied gas, is called a nomination. In stationary models, it is typically assumed that the sum of supplied flows equals the sum of the discharged flows, i.e., the nominations have to be balanced. Gas flows, except for situations of large slopes, from higher to lower pressure and the pressure drop is caused by friction at the rough inner pipe walls. This makes it necessary to increase the gas pressure in compressor stations in order to transport gas over large distances. On the other hand, gas pressure has to be reduced in so-called control valve stations at the transition to regional distribution networks. Moreover, gas transport networks comprise other elements like valves (e.g., to control the direction of the gas flow) or resistors (e.g., to model pressure drops caused by internal station piping or partly closed valves). We refer to Sect. 2.2 for the description of all network elements that are part of the GasLib networks and the chapter Fügenschuh et al. (2015) in the book Koch et al. (2015) for a physical

Date: November 18, 2017.

2010 *Mathematics Subject Classification.* 90-08, 90C90, 90B10.

Key words and phrases. Gas Transport, Networks, Problem Instances, Mixed-Integer Nonlinear Optimization, GasLib.

and technical description of the network elements as well as for several mathematical models.

From a mathematical point of view, the resulting problems lead to large-scale simulation instances and hard mixed-integer nonlinear and nonconvex optimization or feasibility problems. Typically, the solution of the optimization problems are out of reach for state-of-the-art general-purpose solvers for large-scale instances. Therefore, more research is needed in order to provide accurate and fast optimization methods for practical gas transport instances. Unfortunately, gas network data is only seldomly publicly available, except for the Belgian gas network published in De Wolf and Smeers (2000). Moreover, the wealth of gas transport models leads to the problem that they are usually not compared or not on the same networks. The gas network library GasLib tries to fix these problems by providing publicly available small to large-scale academic and real-world networks. It currently contains stationary instances only. The contained real-world data originates from the completed industrial research project ForNe, the project "Investigation of the technical capacities of gas networks" (2009–2012) that was funded by the German Federal Ministry of Economic Affairs and Energy, and from the collaborative research center TRR 154 "Mathematical Modelling, Simulation and Optimization using the Example of Gas Networks", which is funded by the German research foundation. Since some of the original data from our former industry partner Open Grid Europe GmbH[1] and from the Greek network operator DESFA is confidential, the corresponding GasLib instances are distorted versions of the original transport networks. See Sect. 2.1 for a more detailed description of the instances. The complete library is freely available at http://gaslib.zib.de.

The GasLib is useful for different target groups: First, researchers that develop new mathematical models and simulation or optimization algorithms can use the GasLib to test and evaluate these algorithms on network instances of different sizes. Second, developers of optimization and simulation solvers can use GasLib to improve the performance and robustness of their solvers using the example of challenging models from a real-world problem.

The network instances within GasLib are composed of different data that can be used to set up simulation or optimization models for stationary gas transport problems. First, every instance includes a structural and technical description of the network and all contained network elements. Second, several sets of nominations are provided. More detailed information on the provided data and the data format is given in Sect. 2.2.

For some instances, GasLib additionally provides GAMS (2017) models of specific mixed-integer nonlinear models based on the corresponding network instances. This collection of models can be used to test and improve MINLP solvers. The corresponding models are feasibility problems, i.e., no objective function is given, as described in Koch et al. (2015), and the modeling decisions are the same as for the MINLP model described in Geißler et al. (2015). For instance, the MINLP instances arising from the GasLib-582 network have roughly 2200 variables (of which roughly 250 are binary) and 3700 constraints (of which roughly 900 are nonlinear). Numerical results on an earlier version of these models can be found in Pfetsch et al. (2015).

Finally, we like to mention that this paper does not contain descriptions of methods for solving the discussed simulation or optimization models. However, there are already some published research papers that used GasLib instances. We give the corresponding references in Sect. 2.1, where we present the instances in detail.

[1] http://www.open-grid-europe.com

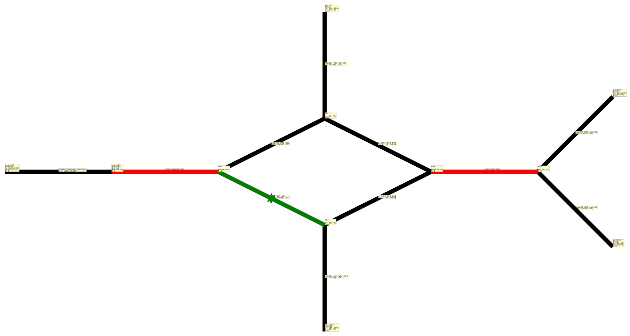

FIGURE 1. Schematic plot of the GasLib-11 network

2. Data Description

In this section we describe the data contained in the GasLib and explain the different data formats.

2.1. **Overview over the Instances.** GasLib currently contains seven instances: four handmade networks with 11, 24, 40, and 135 nodes intended for testing purposes, one medium-scale real-world network with 134 nodes, and two large real-world networks with 582 and 4197 nodes, respectively. In the following, we describe all networks in detail. Note that the number of nodes is always part of the name of the network.

2.1.1. *GasLib-11.* This artificial small-scale network is provided for testing purposes. Its small size allows for fast computations as well as for testing and comparing different models of network elements. The network consists of 3 entries, 3 exits, 8 pipes, 2 compressor stations, and 1 valve. It has the special property that closing the valve yields a tree-structured network topology; see Figure 1. In this and all other network plots different colors are used to denote different network elements: Black arcs are pipes or short pipes, red arcs are compressor stations, resistors are cyan, and control valves are orange. See Sect. 2.2.1 for a more detailed description of these elements. The single nomination provided requires to route 300 Nm^3/h from the entries to the exits.

2.1.2. *GasLib-24.* The small-scale network GasLib-24 is also provided for testing purposes. It is an extended variant of the network used in Ehrhardt and Steinbach (2005); see the description in the cited paper for more details. Its small size again allows for fast computations as well as for testing and comparing different models of network elements. The network consists of 3 entries, 5 exits, 19 pipes, 3 compressor stations, and 1 control valve, short pipe, and resistor, each; see Fig. 2. The single nomination requires to route 544.32 Nm^3/h units of flow from the entries to the exits. This network has already been used in the papers Gugat et al. (2017) and Hante and Schmidt (2017).

2.1.3. *GasLib-40.* This third small-scale network is also provided for testing purposes. Its size also allows for fast computations as well as for testing and comparing different models of network elements. The network consists of 3 entries, 29 exits, 39 pipes, and 6 compressor stations; see Fig. 3. It roughly represents a part of the German low-calorific gas transport network in the Rhine-Main-Ruhr area. A single nomination is provided in which the same amount of gas (75 Nm^3/h) is discharged at every exit node. The entry nodes equally provide the gas, i.e., the inflow at every entry is 725 Nm^3/h. Numerical results for this instance can be found in Borraz-Sánchez

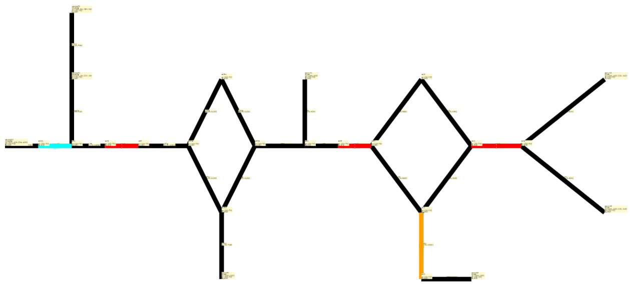

FIGURE 2. Schematic plot of the GasLib-24 network

FIGURE 3. Schematic plot of the GasLib-40 network

et al. (2016), Groß et al. (2017), Habeck et al. (2017), Humpola (2017), and Mak et al. (2015).

2.1.4. *GasLib-134.* The medium-scale network GasLib-134 is a close approximation of the real-world tree-like gas transport network of Greece and consists of a main pipeline connecting the Greek-Bulgarian border with southern Greece and additional 16 branches, comprising an overall length of 1412 km. The network consists of 3 entries, 45 exits, 86 inner nodes, 86 pipes, 45 short pipes, 1 turbo compressor machine, and 1 control valve with preset pressure; see Fig. 4. Due to the confidential nature of the compressor's technical data, it has been replaced by a random example that roughly approximates its main characteristics, namely a maximum mechanical power of 7500 kW and an adiabatic efficiency of 0.84. The associated 1234 daily nominations have been taken from the TSO website DESFA (2017) and processed so that they only represent balanced samples between the entries and the exits. Moreover, all nomination data have been converted from heat power to volumetric units by utilizing a conversion with a single calorific value, thus neglecting the gas mixing within the network. This instance has been used in the papers Gugat et al. (2016), Schmidt et al. (2017), and Sirvent et al. (2017).

FIGURE 4. Schematic plot of the GasLib-134 network

2.1.5. *GasLib-135.* The medium-scale network GasLib-135 is also provided for testing purposes. It allows to test models on a larger network. The network consists of 6 entries, 99 exits, 30 inner nodes, 141 pipes, and 29 compressor stations. It roughly represents a part of the German high-calorific gas transport network; see Fig. 5. Note that this network also contains cycles, which is not the case for the almost equally sized GasLib-134 network. This aspect typically makes the network harder to solve than the GasLib-134. Again, a single nomination is provided in which the same amount of gas ($40\,\mathrm{Nm}^3/\mathrm{h}$) is discharged at every exit. The entries equally provide the gas, i.e., the inflow at every entry is $660\,\mathrm{Nm}^3/\mathrm{h}$. Numerical results for this instances can be found in Borraz-Sánchez et al. (2016), Groß et al. (2017), Humpola (2017), and Mak et al. (2015).

2.1.6. *GasLib-582.* The large-scale network GasLib-582 is based on a real-world gas transport network, which covers approximately one fourth of the area of Germany. Since the original network data from Open Grid Europe GmbH are confidential, we carefully distorted them such that no direct conclusion regarding the true operational capability of the network can be drawn, while still maintaining a realistic behavior. In particular, feasibility of the edited instances is comparable to the original ones.

The network GasLib-582 contains 31 entries, 129 exits, and 422 inner nodes, which are connected by 278 pipes, 5 compressor stations, 23 control valves, 8 resistors, and 269 short pipes; see Fig. 6. Note that this network again contains cycles. The compressor stations contain 8 turbo compressors and 1 piston compressor in total.

The associated 4227 nominations are obtained in a two-stage process. First, the load on some nodes (typically exits) are sampled from distributions that are estimated from historical data. The load for the remaining nodes is then determined such that the load is balanced and a random linear objective is maximized. See Hiller et al. (2015) for a detailed description of the entire process. Since the solution of these nominations can be rather challenging, we additionally provide a set of

Figure 5. Schematic plot of the GasLib-135 network

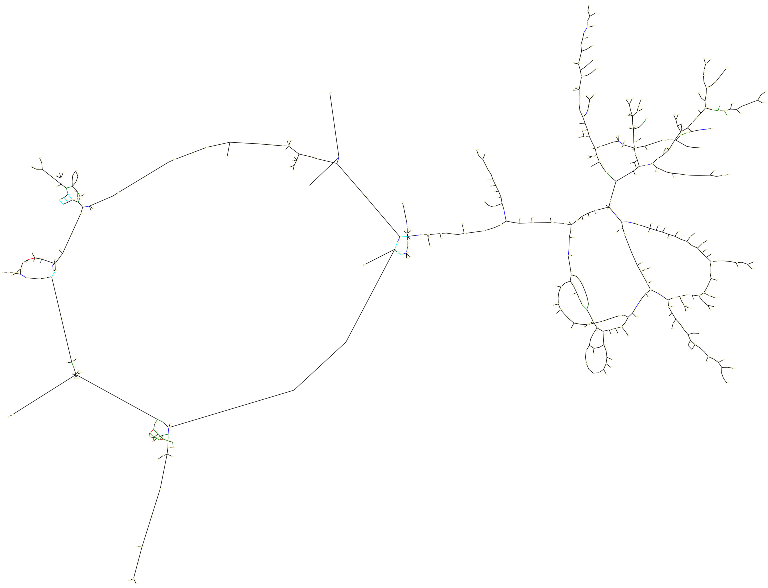

Figure 6. Schematic plot of the GasLib-582 network

nominations that are scaled to 95 % of the flow amount, i.e., the supplied (discharged) amount of gas at every entry (exit) is scaled by 0.95.

The provided network data are meant to test whether and how models and solution methods can handle practical data in realistic sizes. Computational results for the GasLib-582 instances are given in Borraz-Sánchez et al. (2016), Burlacu et al. (2017), Hennig and Schwarz (2016), Koch et al. (2015), Pfetsch et al. (2015), Rose et al. (2016), and Schmidt et al. (2016).

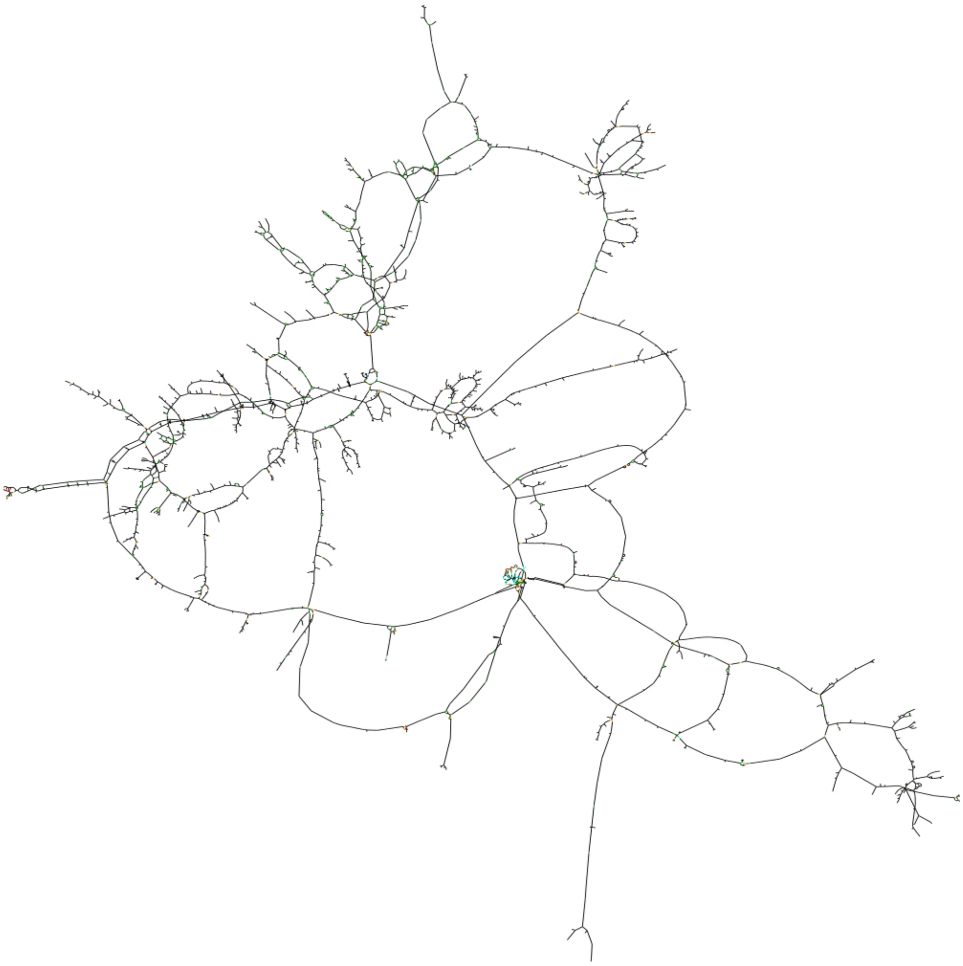

FIGURE 7. Schematic plot of the GasLib-4197 network

2.1.7. *GasLib-4197.* The network GasLib-4197 is based on a real-world gas transport network spanning a third of Germany. Since the original network data from Open Grid Europe GmbH again are confidential, we carefully distorted them as described for the GasLib-582 network.

We expect the feasibility of the edited instances to be comparable to the original ones, but it has to be noted that given the size of the network, we are not yet able to compute feasible solutions to all scenarios yet. Thus, while we assume them to be feasible in reality, there is no proof for this so far.

The network contains 11 entry nodes, 1009 exit nodes, and 3177 inner nodes as well as 3537 pipes, 12 compressor stations, 426 valves, 120 control valves, 28 resistors, and 343 short pipes; see Fig. 7. The compressor stations contain 22 turbo compressors in total. The associated 2859 nominations (i.e., fixed and balanced inflow-outflow-situations with pressure bounds) are obtained as sampled statistical load scenarios of the original network similar to GasLib-582; see Hiller et al. (2015) for a detailed description of the sampling process.

The provided network data are meant to analyze how models and solution methods perform on practical data in realistic sizes. Computational results for the GasLib-4197 instances are given in Geißler et al. (2015, 2017).

Table 1. Exemplary units that are supported by GasLib

Physical/technical quantity	Units
Gas flow	$m^3\,s^{-1}$, $m^3\,h^{-1}$, $1000\,m^3\,h^{-1}$, ...
Gas temperature	°C, K, ...
Length	mm, m, cm, km, ...

2.2. **Description of the Data Files.** All files contained in GasLib are XML files. XML is a language for describing data in a hierarchical format; see XML (2017). A multitude of software to read and write XML files on different platforms is available, rendering XML useful for collecting and publishing data that should be easy to read, to parse, and to write.

We use XSD (XML Schema Definition) files to specify the format of the XML files and the types of the data therein; see XSD (2017) for more details on XSD. The schemas given in the XSD files can be used by software to validate a given XML file against its specification or even to generate XML parsers automatically. XSD files are written in XML as well.

A GasLib network is described using different files. First, the network topology itself is specified in a `net` file. The contained compressor machines are described in a `cs` file and the interplay of different controllable elements of the network is given in a `cdf` file. Finally, a nomination is specified by a separate `scn` file. Thus, the full set of files required for specifying a single stationary demand scenario (nomination), for a gas transport network needs four files. However, the `cdf` file is optional and does not need to be given. Note further that there might be arbitrarily many `scn` files but only one `cs` and one `cdf` file for a given network description (`net`) file.

Most of the data given in the mentioned files are physical or technical quantities. Thus, in order to increase the readability (both for humans and computers), every physical or technical quantity is specified by a value and a unit in which the value has to be interpreted. GasLib provides the usage of many different units for the same physical or technical quantity. Examples are given in Table 1 and a full list can be found in the documentation on the GasLib website or in the XSD files.

2.2.1. *The `net` File.* A gas network is given as a directed graph in which the nodes represent junctions and the arcs represent the elements of the network. This topology is described in the `net` file together with the technical parameters of the network elements.

There are three types of nodes: entry nodes (`source`), exit nodes (`sink`), and junctions (`innode`). Common to all types of nodes are bounds for the pressure, but `source` and `sink` nodes have additional bounds on the supplied and discharged amount of flow at the node. Positive flow at a `source` node means injection of gas, while positive flow at a `sink` stands for the discharge of gas. The `source` nodes have additional parameters, describing the composition of the supplied gas. For example, for source nodes we have to state the values of, e.g., the calorific value or the pseudocritical temperature; see Fig. 8. The flow bounds at the nodes in the `net` file are usually very wide. Fixed values for the nominated flows at nodes are given in the `scn` file that complements the `net` file; see Sect. 2.2.4.

All nodes in the network have schematic and geographic coordinates, which can be used, e.g., for visualization. The `height` value of the nodes is needed to compute the slope of pipes, which typically has an impact on the pressure losses in the network. See Fig. 8 for examples of the three node types.

There are six types of arcs that can occur in a `net` file: pipes (`pipe`), short pipes (`shortPipe`), resistors (`resistor`), valves (`valve`), control valve stations

```
<source geoWGS84Long="10.0667004121" alias="" y="6691.6" x="12108"
        geoWGS84Lat="48.448723929" id="source_1">
  <height unit="m" value="7"/>
  <pressureMin unit="bar" value="1.01325"/>
  <pressureMax unit="bar" value="121.01325"/>
  <flowMin unit="1000m_cube_per_hour" value="0"/>
  <flowMax unit="1000m_cube_per_hour" value="10000"/>
  <gasTemperature unit="Celsius" value="15"/>
  <calorificValue unit="MJ_per_m_cube" value="41.342270292"/>
  <normDensity unit="kg_per_m_cube" value="0.82"/>
  <coefficient-A-heatCapacity value="31.61010551"/>
  <coefficient-B-heatCapacity value="-0.004284754861"/>
  <coefficient-C-heatCapacity value="8.019089e-05"/>
  <molarMass unit="kg_per_kmol" value="18.0488790169"/>
  <pseudocriticalPressure unit="bar" value="46.7020607"/>
  <pseudocriticalTemperature unit="K" value="202.4395142"/>
</source>
<sink geoWGS84Long="10.0667004121" alias="" y="6794.3" x="12090"
      geoWGS84Lat="48.448723929" id="sink_1">
  <height unit="m" value="7"/>
  <pressureMin unit="bar" value="1.01325"/>
  <pressureMax unit="bar" value="121.01325"/>
  <flowMin unit="1000m_cube_per_hour" value="0"/>
  <flowMax unit="1000m_cube_per_hour" value="10000"/>
</sink>
<innode geoWGS84Long="7.92474003681" alias="" y="6324.6" x="5389.9"
        geoWGS84Lat="48.3578033109" id="innode_1">
  <height unit="m" value="77"/>
  <pressureMin unit="bar" value="2.01325"/>
  <pressureMax unit="bar" value="86.01325"/>
</innode>
```

FIGURE 8. Examples of `source`, `sink`, and `innode` in the `net` file

(`controlValve`), and compressor stations (`compressorStation`). All arcs have bounds restricting the flow through the arc (`flowMin`, `flowMax`) and specify their incident nodes as well as their orientation in the network. In addition, every arc type has its specific set of parameters.

Pipes are specified by the additional parameters `length`, `diameter`, and `roughness`; see Fig. 9. They are used to compute the friction factor in the pressure loss equation. The `pressureMax` element specifies the maximum pressure that the material of the pipe can withstand. It is typically used to also bound the pressure on the incident nodes. The `heatTransferCoefficient` specifies the heat transfer coefficient of the material of the pipe's wall and, finally, the `speedLimit` specifies the maximum velocity of the gas in the pipe. The latter parameter is optional. A typical value is $50\,\mathrm{m\,s^{-1}}$.

Short pipes can be interpreted as pipes with zero length and thus do not induce any pressure loss. This especially means that tail and head node of a short pipe have the same location. A `shortPipe` has no further attributes; see Fig. 10 for an example.

A resistor is an artificial network elements that serves as a surrogate for elements like measurement devices or filters. These elements are typically not captured exactly in simulation or optimization models but lead to a pressure loss that should be modeled. Two types of resistors appear in the GasLib instances. One results in a constant pressure loss (in the direction of the flow), which is specified by the

```
<pipe alias="" from="sink_2" id="pipe_1" to="innode_15">
  <flowMin unit="1000m_cube_per_hour" value="-10000"/>
  <flowMax unit="1000m_cube_per_hour" value="10000"/>
  <length unit="km" value="39.7474810299"/>
  <diameter unit="mm" value="1300.0"/>
  <roughness unit="mm" value="0.01"/>
  <pressureMax unit="bar" value="102.0"/>
  <heatTransferCoefficient unit="W_per_m_square_per_K" value="2"/>
  <speedLimit unit="m_per_s" value="50"/>
</pipe>
```

FIGURE 9. Example of `pipe` in the `net` file

```
<shortPipe alias="" from="sink_118" id="shortPipe_1" to="source_1">
  <flowMin unit="1000m_cube_per_hour" value="-10000"/>
  <flowMax unit="1000m_cube_per_hour" value="10000"/>
</shortPipe>
```

FIGURE 10. Example of `shortPipe` in the `net` file

```
<resistor alias="" from="innode_149" id="resistor_1" to="innode_164">
  <flowMin unit="1000m_cube_per_hour" value="-15000"/>
  <flowMax unit="1000m_cube_per_hour" value="15000"/>
  <pressureLoss unit="bar" value="0.5"/>
</resistor>
<resistor alias="" from="innode_30" id="resistor_1" to="innode_29">
  <flowMin unit="1000m_cube_per_hour" value="-10000"/>
  <flowMax unit="1000m_cube_per_hour" value="10000"/>
  <dragFactor value="63.50999832"/>
  <diameter unit="mm" value="1000"/>
</resistor>
```

FIGURE 11. Example of two `resistor` types in the `net` file

```
<valve alias="" from="innode_10" id="valve_1" to="innode_14">
  <flowMin unit="1000m_cube_per_hour" value="-10000"/>
  <flowMax unit="1000m_cube_per_hour" value="10000"/>
  <pressureDifferentialMax unit="bar" value="120"/>
</valve>
```

FIGURE 12. Example of `valve` in the `net` file

element `pressureLoss`. The other type induces a pressure loss that continuously depends on the flow rate through the resistor as specified by the tags `dragFactor` and `diameter`. An example for both types of resistors is given in Fig. 11.

The elements specified by the type `valve` can be open or closed. In the open state, the pressures at the end nodes attain the same value, while in the closed state, the flow rate is zero. They only have one specific parameter `pressureDifferentialMax` that restricts the difference of pressures between the `from` and `to` node when the valve is closed; see Fig. 12 for an example.

There are two variants of `controlValve` elements, as shown in Fig. 13. These elements can be set to three states: active, bypass mode, and closed. In the active state, gas pressure can be reduced (in the direction of the orientation of the arc) and the parameters `pressureDifferentialMin` and `pressureDifferentialMax` define the range of possible pressure reductions. The bypass and closed states are equivalent

```
<controlValve from="innode_8" alias="" gasPreheaterExisting="0"
              to="innode_404" internalBypassRequired="0"
              id="controlValve_1">
  <flowMin unit="1000m_cube_per_hour" value="-10000"/>
  <flowMax unit="1000m_cube_per_hour" value="10000"/>
  <pressureDifferentialMin unit="bar" value="0"/>
  <pressureDifferentialMax unit="bar" value="120"/>
  <pressureInMin unit="bar" value="1.01325"/>
  <pressureOutMax unit="bar" value="86.01325"/>
  <pressureLossIn unit="bar" value="0.75"/>
  <pressureLossOut unit="bar" value="0.75"/>
</controlValve>
<controlValve from="innode_1021" alias="" gasPreheaterExisting="0"
              to="innode_1019" internalBypassRequired="1"
              id="controlValve_42">
  <flowMin unit="1000m_cube_per_hour" value="-0.001"/>
  <flowMax unit="1000m_cube_per_hour" value="15000"/>
  <pressureDifferentialMin unit="bar" value="0"/>
  <pressureDifferentialMax unit="bar" value="120"/>
  <pressureInMin unit="bar" value="21.01325"/>
  <pressureOutMax unit="bar" value="80.00325"/>
  <dragFactorIn value="0"/>
  <diameterIn unit="mm" value="900"/>
  <dragFactorOut value="0"/>
  <diameterOut unit="mm" value="900"/>
</controlValve>
```

FIGURE 13. Example of two types of `controlValve` in the `net` file

to the corresponding states of valves. This means that a control valve in bypass mode may has nonzero flow and does not influence the pressures, whereas a closed control valve blocks the gas flow. The attribute `internalBypassRequired` specifies whether the bypass state needs to be modeled (if `1`) or whether a bypass valve is already given in the network topology. The `controlValve` elements have resistors at their in- and outlet, which are described with the same data as stand-alone resistors. The resistors only have to be considered if the `controlValve` is active. Otherwise, they do not have any impact on the gas flow. Finally, the attribute `gasPreheaterExisting` specifies if a so-called gas preheater is present at the control valve station.

Compressor stations are also given in two variants depending on the type of the included resistors; see Fig. 14 for the examples. The attribute `fuelGasVertex` specifies the node from which the fuel gas is taken and the attribute `gasCoolerExisting` specifies whether a so-called gas cooler is installed at the station or not. Note that the description of the `compressorStation` in the `net` file is not complete. The hosted machines are the most complicated elements of the considered networks, which is the reason why they are described in a separate `cs` file; see Sect. 2.2.2.

2.2.2. *The `cs` File.* The full specification of the compressor stations listed in a `net` file is given in a `cs` file. These files describe all compressor machines and their drives. In addition, a list of so-called configurations is given that defines the set of possible interconnections of the machines in the station.

Most of the machines that are hosted in compressor stations can be described by so-called characteristic diagrams that are achieved by least-squares fits of (bi-)quadratic functions to a given set of measurements. In `cs` files, both a parameterization of the fitted functions as well as the measurements are given.

```xml
<compressorStation from="innode_14" alias="" gasCoolerExisting="1"
                   fuelGasVertex="innode_14" to="innode_389"
                   internalBypassRequired="0"
                   id="compressorStation_1">
  <flowMin unit="1000m_cube_per_hour" value="0"/>
  <flowMax unit="1000m_cube_per_hour" value="10000"/>
  <pressureLossIn unit="bar" value="0.8000000119"/>
  <pressureLossOut unit="bar" value="0.200000003"/>
  <pressureInMin unit="bar" value="21.01325"/>
  <pressureOutMax unit="bar" value="86.01325"/>
</compressorStation>
<compressorStation from="innode_401" alias="" gasCoolerExisting="1"
                   fuelGasVertex="innode_401" to="innode_402"
                   internalBypassRequired="1"
                   id="compressorStation_5">
  <flowMin unit="1000m_cube_per_hour" value="-10000"/>
  <flowMax unit="1000m_cube_per_hour" value="10000"/>
  <dragFactorIn value="18"/>
  <diameterIn unit="mm" value="900"/>
  <dragFactorOut value="16"/>
  <diameterOut unit="mm" value="900"/>
  <pressureInMin unit="bar" value="40.00325168"/>
  <pressureOutMax unit="bar" value="84.11324847"/>
</compressorStation>
```

FIGURE 14. Example of two types of `compressorStation` in the `net` file

In order to describe the XML format below, we briefly state the ansatz functions for the fits: For quadratic functions f_b we have $f_b(\mathbf{x}) = b^\top \mathbf{x}$ where $b \in \mathbb{R}^3$ and $\mathbf{x} = (1, x, x^2)^\top$. For biquadratic functions $f_A(\mathbf{x}, \mathbf{y})$ we have $f_A(\mathbf{x}, \mathbf{y}) = \mathbf{x}^\top A \mathbf{y}$ with $A \in \mathbb{R}^{3\times 3}$, $\mathbf{x} = (1, x, x^2)^\top$, and $\mathbf{y} = (1, y, y^2)^\top$. More details about the technical and physical background as well as on the mathematical modeling can be found in Fügenschuh et al. (2015) or Rose et al. (2016).

Every compressor station hosts a certain number of compressor machines. GasLib distinguishes two types of compressor machines: turbo and piston compressors. Turbo compressors are described by characteristic diagrams in the $(Q, H_{\text{ad}}, n, \eta_{\text{ad}})$-space, where Q is the volumetric flow rate through the compressor, H_{ad} is the adiabatic change in specific enthalpy, n is the speed of the machine, and η_{ad} denotes the adiabatic efficiency of the compression process. The meaning of this quadruple can be summarized as follows: the turbo compressor is able to compress a volumetric flow rate Q with an adiabatic change in specific enthalpy H_{ad} using a compressor speed of n and yielding an efficiency of η_{ad}; see Fig. 16.

The quadruples given for a specific machine are obtained by measurements that are given in the `characteristicDiagramMeasurements` element of the turbo compressor; see Fig. 15. The biquadratic fits are given in the `cs` file by the coefficients `n_isoline_coeff_1`,..., `n_isoline_coeff_9` for the biquadratic isolines of speed of the machine (solid lines in Fig. 16) and the coefficients `eta_ad_isoline_coeff_1`,..., `eta_ad_isoline_coeff_9` for the biquadratic isolines of adiabatic efficiency (dashed lines in Fig. 16). The upper and lower bounds of the characteristic diagram are given by the isolines of speed together with the minimum speed `speedMin` as well as the maximum speed `speedMax`, respectively. The left and right border, i.e., the surge- and the chokeline of the machine are given separately as quadratic functions (of the volumetric flow rate Q) with coefficients `surgeline_coeff_1`,..., `surgeline_coeff_3` and `chokeline_coeff_1`,...,

```
<turboCompressor drive="drive_1" id="compressor_1">
  <speedMin value="4700" unit="per_min"/>
  <speedMax value="6500" unit="per_min"/>
  <n_isoline_coeff_1 value="0"/>
  <!-- ... remaining speed isoline coefficients ... -->
  <eta_ad_isoline_coeff_1 value="0.85558"/>
  <!-- ... remaining efficiency isoline coefficients ... -->
  <surgeline_coeff_1 value="39.67013"/>
  <!-- ... remaining surgeline coefficients ... -->
  <chokeline_coeff_1 value="-43.00857"/>
  <!-- ... remaining chokeline coefficients ... -->
  <efficiencyOfChokeline value="0.5"/>
  <surgelineMeasurements>
    <measurement>
      <speed value="4700" unit="per_min"/>
      <adiabaticHead value="63.62545" unit="kJ_per_kg"/>
      <volumetricFlowrate value="0.20223" unit="m_cube_per_s"/>
    </measurement>
    <!-- ... other surgeline measurements ... -->
  </surgelineMeasurements>
  <characteristicDiagramMeasurements>
    <adiabaticEfficiency value="0.82">
      <measurement>
        <speed value="4700" unit="per_min"/>
        <adiabaticHead value="61.96803" unit="kJ_per_kg"/>
        <volumetricFlowrate value="0.6449" unit="m_cube_per_s"/>
      </measurement>
      <!-- ... other char. diagram measurements ... -->
    </adiabaticEfficiency>
    <!-- ... other adiabaticEfficiency elements ... -->
  </characteristicDiagramMeasurements>
</turboCompressor>
```

FIGURE 15. Exemplary definition of a turbo compressor in a `cs` file

`chokeline_coeff_3`, respectively. The chokeline efficiency is given explicitly in the element `efficiencyOfChokeline`, whereas the surgeline is specified by its measurements contained in the element `surgelineMeasurements`. These measurements are given by triples of `speed`, `adiabaticHead`, and `volumetricFlowrate`. All other measurements of the characteristic diagram are given in the element `characteristicDiagramMeasurements`, which lists sets of such triples for every measured `adiabaticEfficiency`.

The data format for piston compressors is easier; see Fig. 17. Piston compressors are mainly characterized by the volume (`operatingVolume`), which fits into their compression cylinder and the minimal speed (`speedMin`) and maximal speed (`speedMax`) they can be operated in. Additionally, the maximal torque (`maximalTorque`) of the piston compressor's crankshaft, the maximal compression ratio (`maximalCompressionRatio`), and the adiabatic efficiency (`adiabaticEfficiency`) of the compression process are specified. The volumetric flow rate through the piston compressor depends on the rotational speed of the crankshaft. Operating the piston compressor at its minimal speed therefore gives a lower bound on the volumetric flow rate. However, for some piston compressors it is required to decrease the volumetric flow by technical means to a level that is lower than the aforementioned bound. If this is the case, the reducing factor is given in the element `additionalReductionVolFlow`.

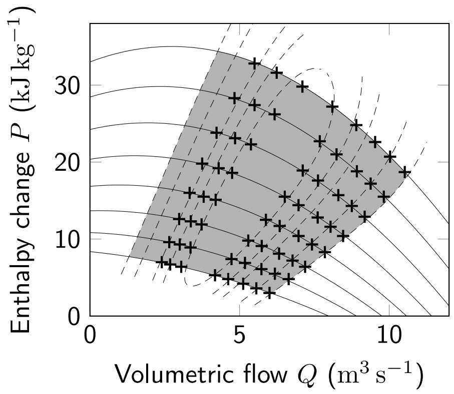

FIGURE 16. Characteristic diagram of an exemplary turbo compressor

```
<pistonCompressor drive="drive_2" id="compressor_2">
  <speedMin unit="per_min" value="165"/>
  <speedMax unit="per_min" value="350"/>
  <operatingVolume unit="m_cube" value="0.5"/>
  <maximalTorque unit="kNm" value="0"/>
  <maximalCompressionRatio value="2"/>
  <adiabaticEfficiency value="0.95"/>
  <additionalReductionVolFlow value="0.35"/>
</pistonCompressor>
```

FIGURE 17. Exemplary specification of a piston compressor in a `cs` file

Every compressor is mounted on a drive that delivers the compressor with the required power. Drives are described by their specific energy consumption rate that is given as a quadratic fit with coefficients `energy_rate_fun_coeff_1`,..., `energy_rate_fun_coeff_3` and the maximum power function. It depends on the specific drive type if the latter is modeled by a quadratic function of the compressor's speed (with coefficients `power_fun_coeff_1`,..., `power_fun_coeff_3`) or by a biquadratic function of the compressor's speed and the ambient temperature at the station (with coefficients `power_fun_coeff_1`,..., `power_fun_coeff_9`). For exemplary XML specifications; see Fig. 18. For further details on modeling drives we refer the reader to Fügenschuh et al. (2015) and Schmidt et al. (2015, 2016).

Finally, the `cs` file contains a list of so-called configurations for every compressor station; see Fig. 19. A configuration is a serial arrangement of parallel combinations of compressor machines and the element `configurations` lists all possible configurations for the corresponding station.

2.2.3. *The `cdf` File.* The `cdf` file optionally complements the `net` file by describing additional restrictions for the switching of certain sets of controllable elements (`valve`, `controlValve`, and `compressorStation`).

Besides compressor machines there are many other elements in real-world compressor or control valve stations that are used to increase the flexibility of the network operator, e.g., for routing the gas near the stations in different directions. The interplay of these sets of network elements is described in the `cdf` file. We call the grouped sets of controllable elements "subnetworks" and the restrictions of their interplay are called "combined decisions". A subnetwork is modeled as a `decisionGroup` in the data format and the corresponding set of possible switching decisions for the elements of the subnetwork are modeled as a list of the element `decision`; see Fig. 20

```
<drives>
  <gasTurbine id="drive_1">
    <energy_rate_fun_coeff_1 value="0.01"/>
    <!-- ... remaining coefficients ... -->
    <power_fun_coeff_1 value="-1496.222462872491"/>
    <!-- ... remaining coefficients ... -->
  </gasTurbine>
  <gasDrivenMotor id="drive_2">
    <energy_rate_fun_coeff_1 value="2629"/>
    <!-- ... remaining coefficients ... -->
    <power_fun_coeff_1 value="230.514705882"/>
    <!-- ... remaining coefficients ... -->
    <specificEnergyConsumptionMeasurements>
      <measurement>
        <compressorPower unit="kW" value="5250"/>
        <fuelConsumption unit="kW" value="15997"/>
      </measurement>
      <!-- ... other measurements ... -->
    </specificEnergyConsumptionMeasurements>
    <maximalPowerMeasurements>
      <measurement>
        <speed unit="per_min" value="165"/>
        <maximalPower unit="kW" value="4375"/>
      </measurement>
      <!-- ... other measurements ... -->
    </maximalPowerMeasurements>
  </gasDrivenMotor>
</drives>
```

FIGURE 18. Example of the specification of a gas turbine and a gas driven motor in a `cs` file

```
<configurations>
  <configuration nrOfSerialStages="1" confId="config_1">
    <stage nrOfParallelUnits="1" stageNr="1">
      <compressor nominalSpeed="350" id="compressor_2"/>
    </stage>
  </configuration>
  <configuration nrOfSerialStages="1" confId="config_3">
    <stage nrOfParallelUnits="2" stageNr="1">
      <compressor nominalSpeed="350" id="compressor_2"/>
      <compressor nominalSpeed="6500" id="compressor_1"/>
    </stage>
  </configuration>
</configurations>
```

FIGURE 19. Example of two configurations in a `cs` file

for examples. One `decision` must be chosen, with implications for the arcs in the `decisionGroup`: For a `value` of `0`, the corresponding network element is closed. For a `value` of `1`, it is open and maybe active or in bypass mode. Optionally, for each arc a `flowDirection` (`0` or `1` for backward or forward flow) and a `mode` (`active` or `bypass`) may be given. If a `decisionGroup` contains only a single `decision`, this effectively fixes the switching decisions for the affected arcs. This is used, e.g., in the GasLib-4197 instance to fix certain `valve` elements that are not remote-controlled.

```
<decisionGroup id="dg1">
  <decision id="d1">
    <controlValve id="controlValve_1" value="0"/>
    <controlValve flowDirection="0" id="controlValve_2" value="1"/>
    <compressorStation id="compressorStation_1" value="0"/>
    <compressorStation id="compressorStation_2" value="0"/>
  </decision>
  <decision id="d2">
    <controlValve id="controlValve_1" value="1" mode="active"/>
    <controlValve id="controlValve_2" value="0"/>
    <compressorStation id="compressorStation_1" value="1"
                       mode="bypass"/>
    <compressorStation id="compressorStation_2" value="1"/>
  </decision>
</decisionGroup>
<decisionGroup id="dg2">
  <decision id="d1">
    <valve id="valve_1" value="1"/>
  </decision>
</decisionGroup>
```

FIGURE 20. Two examples for a `decisionGroup` of a `cdf` file

2.2.4. *The `scn` File.* The `scn` file specifies a demand and supply scenario for a network. Thus, a gas transport instance is only given completely by the combination of a network (with a `net`, a `cs`, and an optional `cdf` file) together with an `scn` file. GasLib contains a single scenario file for the four small- to medium-scale networks GasLib-11, GasLib-24, GasLib-40, and GasLib-135 and many different `scn` files for the real-world networks GasLib-134, GasLib-582, and GasLib-4197.

The supplied and discharged amounts of gas at entries and exits can be specified in mass flow q ($\mathrm{kg\,s^{-1}}$), normal volumetric flow Q_0 ($\mathrm{m^3\,h^{-1}}$), or power P (kW). The power can be computed from the mass flow q and the calorific value H_c via $P = qH_\mathrm{c}$. The mass flow is computed from the normal volumetric flow Q_0 and the density under normal conditions ρ_0 via $q = Q_0\rho_0$. Conservation of mass at the nodes of the network is typically stated in mass flow q. However, it can also be expressed in the normal volumetric flow if a constant density under normal conditions is assumed throughout the network. All flow bounds in the `net` and `scn` files of GasLib are expressed in normal volumetric flow since it is the standard unit in gas transport.

In stationary models of gas flow, the supplied and discharged values must be in balance, i.e., the sum of the supplies at the `source` nodes equals the sum of the discharged flows at the `sink` nodes. The `flow` values in all `scn` files of GasLib are fixed to values that satisfy this balance constraint. Thus, all provided `scn` files define so-called nominations as described in the introduction.

For every `node` in the `scn` file, a `type` is specified: `entry` or `exit`. Typically, nodes that are a `source` (`sink`) in the `net` file should be an `entry` (`exit`) in the `scn` file. Otherwise, the sign of the corresponding flow value has to be changed.

Finally, `pressure` bounds may be specified for the nodes. These bounds should be intersected with the bounds from the `net` file, i.e., the tightest bound holds. Typically, lower pressure bounds are given for exits and upper pressure bounds are given for entries. Optionally, `contractPressure` bounds may be specified for further restricting the allowed pressure values.

See Figure 21 for an exemplary specification of pressure bounds and a fixation of the flow amount in an `scn` file.

```xml
<node type="entry" id="source_1">
  <pressure unit="bar" bound="lower" value="2.0133"/>
  <pressure unit="bar" bound="upper" value="86.013"/>
  <flow unit="1000m_cube_per_hour" bound="both" value="472.636"/>
</node>
```

FIGURE 21. An exemplary node specification from an scn file

Acknowledgements

This work has partially been supported by the German Federal Ministry of Economics and Technology owing to a decision of the German Bundestag. We are very grateful to our former industry partner Open Grid Europe GmbH for providing us with real-world network and nomination data as well as for numerous discussions on that topic. Finally, we thank all collaborators of the ForNe project, i.e., all associated researchers from Zuse Institute Berlin, Friedrich-Alexander-Universität Erlangen-Nürnberg, Leibniz Universität Hannover, University of Duisburg-Essen, Humboldt University zu Berlin, TU Darmstadt, and WIAS Institute in Berlin. We further acknowledge funding through the DFG SFB/Transregio 154, subprojects A01, A05, A07, B06, B07, B08, Z01, Z02.

3. Outlook

The next steps in the further development of GasLib are to add more and even larger instances. Moreover, we also plan to provide transient data in the near future. In any case, GasLib will only be successful if it is used. Moreover, we are always grateful for constructive remarks and, of course, for providing new data.

References

Borraz-Sánchez, C., R. Bent, S. Backhaus, H. Hijazi, and P. van Hentenryck (2016). "Convex Relaxations for Gas Expansion Planning." In: *INFORMS Journal on Computing* 28.4, pp. 645–656. DOI: 10.1287/ijoc.2016.0697.

Burlacu, R., B. Geißler, and L. Schewe (2017). *Solving Mixed-Integer Nonlinear Programs using Adaptively Refined Mixed-Integer Linear Programs*. Tech. rep. FAU Erlangen-Nürnberg. URL: http://www.optimization-online.org/DB_HTML/2017/05/6029.html. Submitted.

De Wolf, D. and Y. Smeers (2000). "The Gas Transmission Problem Solved by an Extension of the Simplex Algorithm." In: *Management Science* 46.11, pp. 1454–1465. DOI: 10.1287/mnsc.46.11.1454.12087.

DESFA (2017). *DESFA*. http://www.desfa.gr. Accessed: 2017-11-01. DESFA.

Ehrhardt, K. and M. C. Steinbach (2005). "Nonlinear Optimization in Gas Networks." In: *Modeling, Simulation and Optimization of Complex Processes*. Ed. by H. G. Bock, E. Kostina, H. X. Phu, and R. Rannacher. Springer: Berlin, pp. 139–148. DOI: 10.1007/3-540-27170-8_11.

XML (2017). *Extensible Markup Language (XML)*. https://www.w3.org/XML/. Accessed: 2017-11-16. W3C.

Fügenschuh, A., B. Geißler, R. Gollmer, A. Morsi, M. E. Pfetsch, J. Rövekamp, M. Schmidt, K. Spreckelsen, and M. C. Steinbach (2015). "Physical and technical fundamentals of gas networks." In: *Evaluating Gas Network Capacities*. Ed. by T. Koch, B. Hiller, M. E. Pfetsch, and L. Schewe. SIAM-MOS series on Optimization. SIAM. Chap. 2, pp. 17–43. DOI: 10.1137/1.9781611973693.ch2.

Geißler, B., A. Martin, A. Morsi, and L. Schewe (2015). "The MILP-relaxation approach." In: *Evaluating Gas Network Capacities*. Ed. by T. Koch, B. Hiller, M. E. Pfetsch, and L. Schewe. SIAM-MOS series on Optimization. SIAM. Chap. 6, pp. 103–122. DOI: 10.1137/1.9781611973693.ch6.

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

Geißler, B., A. Morsi, L. Schewe, and M. Schmidt (2017). "Solving Highly Detailed Gas Transport MINLPs: Block Separability and Penalty Alternating Direction Methods." In: *INFORMS Journal on Computing*. DOI: 10.1287/ijoc.2017.0780. URL: http://www.optimization-online.org/DB_HTML/2016/06/5523.html. Forthcoming.

GAMS (2017). *General Algebraic Modeling System*. http://gams.com. Accessed: 2017-11-01. GAMS Development Corporation.

Groß, M., M. E. Pfetsch, L. Schewe, M. Schmidt, and M. Skutella (2017). *Algorithmic Results for Potential-Based Flows: Easy and Hard Cases*. Tech. rep. URL: http://www.optimization-online.org/DB_HTML/2017/08/6185.html. Submitted.

Gugat, M., G. Leugering, A. Martin, M. Schmidt, M. Sirvent, and D. Wintergerst (2016). *Towards Simulation Based Mixed-Integer Optimization with Differential Equations*. Tech. rep. FAU Erlangen-Nürnberg. URL: http://www.optimization-online.org/DB_HTML/2016/07/5542.html. Submitted.

Gugat, M., G. Leugering, A. Martin, M. Schmidt, M. Sirvent, and D. Wintergerst (2017). *MIP-Based Instantaneous Control of Mixed-Integer PDE-Constrained Gas Transport Problems*. Tech. rep. FAU Erlangen-Nürnberg. URL: http://www.optimization-online.org/DB_HTML/2017/04/5955.html. Submitted.

Habeck, O., M. E. Pfetsch, and S. Ulbrich (2017). *Global optimization of ODE constrained network problems on the example of stationary gas transport*. Preprint. http://www.optimization-online.org/DB_HTML/2017/10/6288.html. Optimization Online.

Hante, F. M. and M. Schmidt (2017). *Complementarity-Based Nonlinear Programming Techniques for Optimal Mixing in Gas Networks*. Tech. rep. FAU Erlangen-Nürnberg.

Hennig, K. and R. Schwarz (2016). *Using Bilevel Optimization to find Severe Transport Situations in Gas Transmission Networks*. eng. Tech. rep. 16-68. Takustr.7, 14195 Berlin: ZIB.

Hiller, B., C. Hayn, H. Heitsch, R. Henrion, H. Leövey, A. Möller, and W. Römisch (2015). "Methods for verifying booked capacities." In: *Evaluating Gas Network Capacities*. Ed. by T. Koch, B. Hiller, M. E. Pfetsch, and L. Schewe. SIAM-MOS series on Optimization. SIAM. Chap. 14, pp. 291–315. DOI: 10.1137/1.9781611973693.ch14.

Humpola, J. (2017). "Gas Network Optimization by MINLP." PhD thesis. Berlin, Germany: Technische Universität Berlin.

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

Mak, T. W. K., P. Van Hentenryck, A. Zlotnik, H. Hijazi, and R. Bent (2015). *Efficient Dynamic Compressor Optimization in Natural Gas Transmission Systems*. Tech. rep.

Pfetsch, M. E., 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 (2015). "Validation of nominations in gas network optimization: models, methods, and solutions." In: *Optimization Methods and Software* 30.1, pp. 15–53. DOI: 10.1080/10556788.2014.888426.

Ríos-Mercado, R. Z. and C. Borraz-Sánchez (2015). "Optimization Problems in Natural Gas Transportation Systems: A State-of-the-Art Review." In: *Applied Energy* 147, pp. 536–555. DOI: 10.1016/j.apenergy.2015.03.017.

Rose, D., M. Schmidt, M. C. Steinbach, and B. M. Willert (2016). "Computational optimization of gas compressor stations: MINLP models versus continuous reformulations." In: *Mathematical Methods of Operations Research* 83.3, pp. 409–444. DOI: 10.1007/s00186-016-0533-5.

Schmidt, M., M. Sirvent, and W. Wollner (2017). *A Decomposition Method for MINLPs with Lipschitz Continuous Nonlinearities*. Tech. rep. URL: http://www.optimization-online.org/DB_HTML/2017/07/6130.html. Submitted.

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

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

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

XSD (2017). *W3C XML Schema Definition Language (XSD)*. https://www.w3.org/TR/xmlschema11-1/. Accessed: 2017-11-16. W3C.

[1] Friedrich-Alexander-Universität Erlangen-Nürnberg, Discrete Optimization, Cauerstr. 11, 91058 Erlangen, Germany; [2] Energie Campus Nürnberg, Fürther Str. 250, 90429 Nürnberg, Germany; [3] Konrad-Zuse-Zentrum für Informationstechnik Berlin, Takustr. 7, 14195 Berlin, Germany; [4] Technische Universität Braunschweig, Institut für Mathematische Optimierung, Universitätsplatz 2, 38106 Braunschweig, Germany; [5] Aristotle University of Thessaloniki, School of Electrical and Computer Engineering, 54124 Thessaloniki, Greece; [6] Leibniz Universität Hannover, Institut für Angewandte Mathematik, Welfengarten 1, 30167 Hannover, Germany; [7] Technische Universität Darmstadt, Department of Mathematics, Dolivostr. 15, 64293 Darmstadt, Germany