

Hiller, Benjamin; Saitenmacher, René; Walther, Tom

Article — Published Version

Improved models for operation modes of complex compressor stations

Mathematical Methods of Operations Research

Provided in Cooperation with:

Springer Nature

Suggested Citation: Hiller, Benjamin; Saitenmacher, René; Walther, Tom (2021) : Improved models for operation modes of complex compressor stations, Mathematical Methods of Operations Research, ISSN 1432-5217, Springer, Berlin, Heidelberg, Vol. 94, Iss. 2, pp. 171-195,
<https://doi.org/10.1007/s00186-021-00745-x>

This Version is available at:

<https://hdl.handle.net/10419/286721>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>



Improved models for operation modes of complex compressor stations

Benjamin Hiller^{1,2}  · René Saitenmacher^{1,3} · Tom Walther^{1,4}

Received: 3 May 2018 / Revised: 26 April 2021 / Accepted: 9 June 2021 / Published online: 11 September 2021
© The Author(s) 2021

Abstract

We study combinatorial structures in large-scale mixed-integer (nonlinear) programming problems arising in gas network optimization. We propose a preprocessing strategy exploiting the observation that a large part of the combinatorial complexity arises in certain subnetworks. Our approach analyzes these subnetworks and the combinatorial structure of the flows within these subnetworks in order to provide alternative models with a stronger combinatorial structure that can be exploited by off-the-shelf solvers. In particular, we consider the modeling of operation modes for complex compressor stations (i.e., ones with several in- or outlets) in gas networks. We propose a refined model that allows to precompute tighter bounds for each operation mode and a number of model variants based on the refined model exploiting these tighter bounds. We provide a procedure to obtain the refined model from the input data for the original model. This procedure is based on a nontrivial reduction of the graph representing the gas flow through the compressor station in an operation mode. We evaluate our model variants on reference benchmark data, showing that they reduce the average running time between 10% for easy instances and 46% for hard instances. Moreover, for three of four considered networks, the average number of search tree nodes is at least halved, showing the effectivity of our model variants to guide the solver's search.

Keywords Multi-way compressor station · Bound tightening · Model reformulation · Flow patterns

This paper is the long version of the extended abstract (Hiller et al. 2016).

The authors thank the Deutsche Forschungsgemeinschaft for their support within project A04 in CRC TRR154 and the BMBF Research Campus Modal (fund number 05M14ZAM) and ICT COST Action TD1207 for additional support.

Extended author information available on the last page of the article

1 Introduction

Gas transmission networks are a crucial part of the European energy supply infrastructure. The gas flow is driven by pressure potentials. To maintain the necessary pressure levels and control the routing of the gas in the network, compressor stations are used. In the German transmission network compressor stations usually interconnect two or more pipeline systems. They often have a complex internal structure, allowing them to realize different routing patterns and compression levels (Koch et al. 2015). An example of such a complex compressor station is shown in Fig. 1. In particular, some or all boundary nodes of the compressor station may serve as inlet or outlet, depending on the requirements of the surrounding network.

In this paper, we propose improved variants of the compressor station modeling introduced in Pfetsch et al. (2015); Koch et al. (2015). This model combines a network containing so-called *compressor groups*, valves, control valves, and a set of switching states for these elements to describe all feasible operation modes of a compressor station. According to Koch et al. (2015), a *compressor group* is a modeling element that represents a set of compressor units (i.e., the machines doing the gas compression) that may be used in different *configurations* (e.g., in serial or parallel) as a single network arc. Note that a compressor group may already be sufficient to model a (simple) compressor station – one with exactly one inlet and one outlet. However, only the more general structure of a complex compressor station consisting of a network with several compressor groups allows to capture the complexity of real-world compressor stations from the meshed transmission networks found e.g., in Germany.

The constraints describing the technical capability of a compressor unit may be nonlinear and nonconvex. Combining this with the combinatorics for switching active network elements yields hard-to-solve MINLP models for many gas network optimization problems.

We illustrate the relevance of compressor stations for gas network optimization problems with some details about the compressor station model shown in Fig. 1. The corresponding network contains 4 compressor groups, 2 control valves, and 19 valves, whereas the entire network contains 7 compressor groups, 26 control valves, and 33 valves. In other words: This subnetwork contains more than a third of the switchable network elements of the entire network and hence accounts for a large part of the combinatorial complexity of the overall optimization problem. Thus proper handling of the feasible operation modes, i.e., joint switching states of all active network elements within the compressor station subnetwork, is important to effectively solve gas network optimization problems.

We propose model refinements that aim to help the solution process to detect unsuitable operation modes early. For instance, to improve the model a natural idea is to precompute, for each operation mode, bounds on the minimum and maximum flow and pressure that can be handled and to include this information in the model. However, the modeling of an operation mode from Koch et al. (2015) and the data available in the GasLib (Humpola et al. 2017) do not specify whether a compressor group is actively compressing or bypassed, which are very distinct behaviors. Thus no non-trivial flow bounds may be obtained for an operation mode as specified in the original modeling. Moreover, the solver has no way to discard all the operation modes where

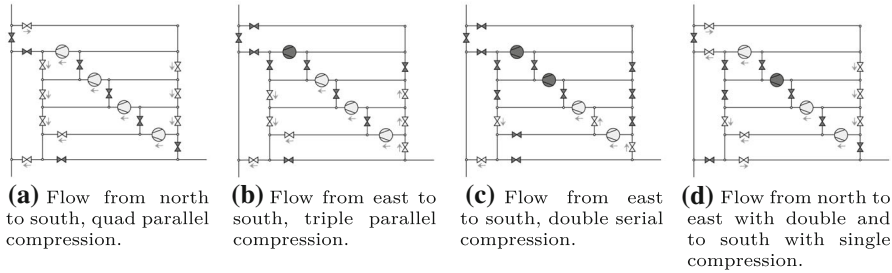


Fig. 1 Four operation modes of a large compressor station. Compressor groups are represented as \odot , valves are represented as \bowtie . Arrows denote a flow direction. Elements colored dark are *closed* and do not admit gas flow

a certain controllable element is e.g., in bypass although active operation is required because it is implied by the remaining model. Hence the solver has to spend substantial additional effort to enumerate the solution space.

We develop techniques for analyzing the original representation of operation modes to generate a more detailed representation where each operation mode prescribes for each compressor group whether it is actively compressing or bypassed. In addition to enable the branching possibilities mentioned above, this allows to compute tight bounds for the pressure/flow combinations that can be handled by each such *fully specified* operation mode. There may be many fully specified operation modes arising from a single original operation mode. To keep the number of fully specified operation modes to consider small, a crucial ingredient is a method to obtain a reduced representation of a fully specified operation mode. This reduced representation is used to identify a minimal set of fully specified operation modes that is equivalent to the set of original operation modes. Examples of such reduced representations are shown in Fig. 2. Moreover, these reduced representations facilitate the classification of operation modes according to what we call flow patterns. A *flow pattern* characterizes which boundary nodes of the compressor station act as sources or sinks for the gas flow in the compressor station. Flow patterns group related fully specified operation modes, which may also be useful for effective branching. This intuition is confirmed by the observation from our computational results that models taking the flow patterns into account usually improve over the variant without the flow pattern.

As the number of fully specified operation modes may still be rather large, model size becomes a concern. To address this, we propose an alternative MILP formulation of the respective constraints which is substantially sparser and where the number of additional constraints is independent of the number of operation modes.

Related work We briefly mention some related papers and refer to Ríos-Mercado and Borraz-Sánchez (2015) for a comprehensive overview. Most of the work on the optimization of compressor stations has focused on simple compressor stations that are compressing in a predetermined direction from a single inlet to a single outlet. Neglecting the fact that a compressor station usually features several (often distinct) compressor units the entire compressor station is often modeled in an aggregated way, like a range for the power required by the compression process (Martin et al. 2006;

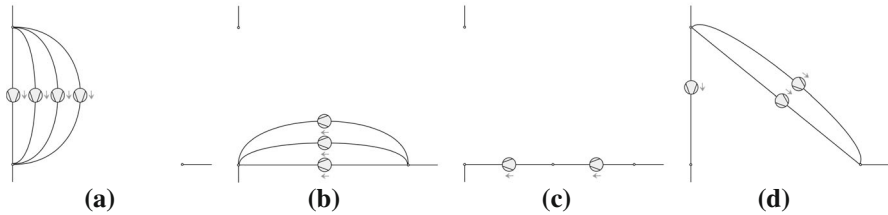


Fig. 2 Reduced representation of the four operation modes shown in Fig. 1. All compressor group symbols correspond to active compressor groups, i.e., ones that are not bypassed

de Wolf and Bakhouya 2012), box constraints for flows and pressures (Carter 1996), or a polyhedral model relating inlet pressure, outlet pressure, and gas flow (Wu et al. 2000; van der Hoeven 2004; Humpola et al. 2014).

Papers using a detailed model for the operation of a single compressor unit usually assume that a compressor station consists of several identical compressor units that are operated in parallel (Wu et al. 2000), the only discrete decision being the number of compressor units switched on. A recent exception is the work of Rose et al. (2016) which considers the selection of a configuration of compressor units that consists of serial stages of compressor units used in parallel. Complex multi-way compressor stations with multiple operation modes are only considered in Koch et al. (2015) and related work, using the basic model that this paper improves upon as outlined above.

The remaining paper is structured as follows. Section 2 recalls the general model for complex compressor stations introduced in Koch et al. (2015), the original MILP model from Geißler et al. Geißler et al. (2015a) and presents our novel smaller and sparser formulation. In Sect. 3, we propose a method that reduces the description of a single operation mode to a kind of “normal form”. This is used in Sect. 4 to detect redundancy and generate a set of redundancy-free operation modes that, as a set, are equivalent to the original operation modes. Section 5 explains the classification of operation modes according to flow patterns and how this can be incorporated in the model. Finally, we report on some computational results in Sect. 6 and conclude in Sect. 7.

2 Model for complex compressor stations

We use the following model for compressor stations that closely follows the modeling proposed in Koch et al. (2015).

Network We assume that a (complex) compressor station is represented as a network where the arcs correspond to different types of network elements. Some of these network elements are controllable (valves, compressor groups, control valves) and some are not controllable or passive (pipes, resistors, short pipes). We assume that the only passive network elements within compressor station networks are short pipes, i.e., there are no pipes or resistors. In the following we briefly explain the different types of network elements.

A compressor group as introduced in Pfetsch et al. (2015); Koch et al. (2015) consists of one or more compressor units which may be operated in one of a given set of serial-parallel configurations. Compressor groups are treated by us as abstract entities that facilitate pressure increase and that may operate in one of the states *closed* (no gas flow), *active* (compressing), and *bypass* (gas flow without compression).

In contrast, control valves are used to facilitate the decrease of pressure, and they may operate in the same set of three states as compressor groups with the difference that an active control valve reduces the pressure of the gas flowing through it. Valves are used to control the route of gas through the compressor station, and they can be *open* or *closed*. Finally, short pipes are connection elements that allow gas flow without pressure drop.

Using the set of valves A_{va} , the set of compressor groups A_{cg} , the set of control valves A_{cv} , and the set of short pipes A_{sp} , we represent a compressor station as a directed graph (V, A) with $A := A_{va} \cup A_{cg} \cup A_{cv} \cup A_{sp}$. Moreover, we partition the set of nodes V into boundary nodes V_{\pm} and inner nodes V_0 .

For each node $u \in V$ we introduce a variable for the pressure p_u with non-negative lower and upper bounds \underline{p}_u and \bar{p}_u , i.e.,

$$0 \leq \underline{p}_u \leq p_u \leq \bar{p}_u \quad \text{for all } u \in V. \tag{1}$$

For each arc $a \in A$ there is a variable for the mass flow q_a with lower and upper bounds \underline{q}_a and \bar{q}_a , i.e.,

$$\underline{q}_a \leq q_a \leq \bar{q}_a \quad \text{for all } a \in A. \tag{2}$$

Positive mass flow values indicate flow in the direction of the arc, whereas negative values represent flow in the opposite direction. The precise values of the bounds depend on the type and state of an element and are discussed below. For each node $u \in V$ we denote the sets of incoming and outgoing arcs by $\delta^-(u)$ and $\delta^+(u)$, respectively. We define b_u , the inflow at node u , whereby at inner nodes the mass flow is conserved, i.e.,

$$b_u := \sum_{a \in \delta^-(u)} q_a - \sum_{a \in \delta^+(u)} q_a \quad \text{for all } u \in V, \tag{3}$$

$$b_u = 0 \quad \text{for all } u \in V_0. \tag{4}$$

Short pipes For short pipes the following holds:

$$\underline{q}_a = -\infty, \bar{q}_a = \infty \quad \text{for all } a \in A_{sp}, \tag{5a}$$

$$p_u = p_v \quad \text{for all } (u, v) \in A_{sp}. \tag{5b}$$

Valves For each valve $a \in A_{va}$, a binary variable s_a distinguishes between the states *open* and *closed* where $s_a = 1$ corresponds to *open*, $s_a = 0$ corresponds to *closed*.

$$s_a = 0 \implies q_a = 0 \quad \text{for all } a \in A_{va}, \quad (6a)$$

$$s_a = 1 \implies \begin{cases} \underline{q}_a = -\infty, \bar{q}_a = \infty, \\ p_u = p_v \end{cases} \quad \text{for all } a = (u, v) \in A_{va}. \quad (6b)$$

Compressor groups and control valves For each compressor group $a \in A_{cg}$ or control valve $a \in A_{cv}$, binary variables s_a, s_a^{ac}, s_a^{bp} distinguish between the states *active*, *bypass* and *closed* where $s_a^{ac} = 1$ corresponds to *active*, $s_a^{bp} = 1$ corresponds to *bypass* and $s_a = 0$ corresponds to *closed*. We model the set of feasible operation points of an active compressor group or control valve a by an abstract set $P_a \subseteq \mathbb{R}_{\geq 0}^3$ of feasible inlet pressure, outlet pressure, and mass flow. Thus our methods apply to all variants of compressor groups and control valve models. In total, compressor groups and control valves are modeled by the constraints

$$s_a = s_a^{ac} + s_a^{bp} \quad \text{for all } a \in (A_{cg} \cup A_{cv}), \quad (7a)$$

$$s_a = 0 \implies q_a = 0 \quad \text{for all } a \in (A_{cg} \cup A_{cv}), \quad (7b)$$

$$s_a^{bp} = 1 \implies \begin{cases} \underline{q}_a = -\infty, \bar{q}_a = \infty, \\ p_u = p_v \end{cases} \quad \text{for all } a = (u, v) \in (A_{cg} \cup A_{cv}), \quad (7c)$$

$$s_a^{ac} = 1 \implies (p_u, p_v, q_a) \in P_a \subseteq \mathbb{R}_{\geq 0}^3 \quad \text{for all } a = (u, v) \in (A_{cg} \cup A_{cv}). \quad (7d)$$

The constraints describing the capability set P_a of a compressor unit may be nonlinear and nonconvex, leading to hard-to-solve MINLPs for large-scale networks (Koch et al. 2015; Rose et al. 2016).

In the remainder of the paper, compressor groups and control valves will be handled exactly the same, unless mentioned otherwise. For simplicity, we will only talk about compressor groups with the understanding that the techniques also work for control valves.

Operation modes An operation mode specifies the switching state of each controllable element (valves, control valves, compressor groups) and thus the route of the gas flow through the compressor station. Operation modes are modeled in Geißler et al. (2015a, sect.6.1.8) by a triple $(A_{ctrl}, \mathcal{M}, d)$, where $A_{ctrl} = A_{va} \cup A_{cg} \cup A_{cv}$ is the set of controllable elements. The set $\mathcal{M} \subseteq \{0, 1\}^{A_{ctrl}}$ describes each operation mode $m \in \mathcal{M}$ by stating whether a controllable element a is *open* ($m_a = 1$) or *closed* ($m_a = 0$). In the case of an open compressor group it is not yet specified whether this compressor group is *active* or in *bypass*. Finally, the function $d : A_{ctrl} \times \mathcal{M} \rightarrow \{-1, 0, 1\}$ describes whether the flow direction for a controllable element $a = (u, v)$ is restricted

or not, using the convention

$$d(a, m) = \begin{cases} -1, & \text{if gas flows from } v \text{ to } u \text{ in operation mode } m, \\ 0, & \text{if the flow direction on } a \text{ is unrestricted in operation mode } m, \\ 1, & \text{if gas flows from } u \text{ to } v \text{ in operation mode } m. \end{cases}$$

Binary decision variables $s_m \in \{0, 1\}$ for all $m \in \mathcal{M}$ are used to formulate the constraint that *exactly one* operation mode has to be chosen, i.e.,

$$\sum_{m \in \mathcal{M}} s_m = 1. \tag{8a}$$

Further, an element $a \in A_{\text{ctrl}}$ may only be used as specified by the elements of \mathcal{M} :

$$\sum_{m \in \mathcal{M}} m_a s_m \geq s_a \quad \text{for all } a \in A_{\text{ctrl}}, \tag{8b}$$

$$\sum_{m \in \mathcal{M}} (1 - m_a) s_m \geq (1 - s_a) \quad \text{for all } a \in A_{\text{ctrl}}. \tag{8c}$$

Finally, the flow directions specified by an operation mode (if any) need to be enforced. This is achieved by adding the constraints

$$\underline{q}_a s_m + q_a \geq \underline{q}_a \quad \text{for all } a \in A_{\text{ctrl}}, m \in \mathcal{M} \text{ with } d(a, m) = 1, \tag{8d}$$

$$\bar{q}_a s_m + q_a \leq \bar{q}_a \quad \text{for all } a \in A_{\text{ctrl}}, m \in \mathcal{M} \text{ with } d(a, m) = -1. \tag{8e}$$

Observe that (8b) and (8c) imply that there are $2|A_{\text{ctrl}}|$ additional constraints with $(|\mathcal{M}| + 2)|A_{\text{ctrl}}|$ nonzero elements to model the operation modes. For an active element $a \in A_{\text{ctrl}}$, let m^{fwd} (m^{bwd}) denote the number of operation modes that specify a forward (backward) flow direction. Then (8d) and (8e) yield $m^{\text{fwd}} + m^{\text{bwd}}$ additional constraints with in total $2(m^{\text{fwd}} + m^{\text{bwd}})$ nonzero elements. Hence it is important to control the number of the operation modes as their number has a severe impact on overall model size.

In the following we describe our suggested improvements to the given model. As mentioned in the introduction, the fact that the representation of operation modes does not specify whether an open compressor group is *active* or running in *bypass* precludes us from obtaining tight bounds for flows and pressures obtainable by an operation mode. Moreover, the solver has no way to discard all the operation modes where a certain controllable element is e.g., in bypass although active operation is required because it is implied by the remaining model. We thus propose a more detailed representation where each operation mode prescribes for each compressor group whether it is *active* or in *bypass*. To obtain this representation from the original one we enumerate all *active/bypass* combinations for each operation mode. Since this leads to a large number of mostly redundant operation modes, we apply the methods described in Sect. 3 to obtain a smaller, yet complete set of *fully specified operation modes*. These are described by a triple $(A_{\text{ctrl}}, \mathcal{M}', d')$. The set $\mathcal{M}' \subseteq \{0, 1\}^{A_{\text{ctrl}}} \times \{0, 1\}^{A_{\text{ctrl}}}$ specifies for

each operation mode $(m^{ac}, m^{bp}) \in \mathcal{M}'$ whether a valve a is *open* ($m_a^{ac} = 0, m_a^{bp} = 1$) or *closed* ($m_a^{ac} = 1, m_a^{bp} = 0$) and whether a compressor group b is in *bypass* ($m_b^{ac} = 0, m_b^{bp} = 1$), *active* ($m_b^{ac} = 1, m_b^{bp} = 0$) or *closed* ($m_b^{ac} = 0, m_b^{bp} = 0$). The function $d' : A_{ctrl} \times \mathcal{M}' \rightarrow \{-1, 0, 1\}$ describes restrictions on the flow direction for controllable elements, using the same convention as the function d in the original formulation. The new model can then be formulated analogously to (8a)–(8e) with $\mathcal{M} := \{m^{ac} + m^{bp} \mid (m^{ac}, m^{bp}) \in \mathcal{M}'\}$, $d := d'$ and two additional constraints enforcing the full state for all compressor groups and control valves as specified by the elements of \mathcal{M}' :

$$s_m \leq s_a^{ac} \quad \text{for all } a \in (A_{cg} \cup A_{cv}), m_a = (1, 0) \in \mathcal{M}', \quad (9a)$$

$$s_a^{ac} + s_m \leq 1 \quad \text{for all } a \in (A_{cg} \cup A_{cv}), m_a = (0, 1) \in \mathcal{M}'. \quad (9b)$$

Note that (9a) and (9b) yield one constraint per fully specified operation mode with 2 nonzero elements each.

The number of fully specified operation modes is at worst exponential in the number of controllable elements in the network representing the compressor station. Hence it is crucial to control the size of the submodel for fully specified operation modes. With the operation mode constraints (8b)–(9b), the model size increases drastically if there are many fully specified operation modes. In particular, the constraint pairs (8d), (8e) and (9a), (9b) each yield several rows per fully specified operation mode.

There are alternative sparser formulations, one of which we sketch in the following. The constraint that an active element a has to be open (closed) if an operation mode specifying it to be open (closed) is selected, can be expressed as

$$\sum_{m \in \mathcal{M}' : m_a \neq (0,0)} s_m = s_a \quad \text{or} \quad (10a)$$

$$\sum_{m \in \mathcal{M}' : m_a = (0,0)} s_m = 1 - s_a \quad (10b)$$

depending on which index set is smaller. In comparison to (8b) and (8c) this yields only one instead of two constraints per active element and at least halves the number of nonzero elements. The flow direction of an active element a can be enforced according to the operation mode via

$$\left(1 - \sum_{m \in \mathcal{M}' : d(a, m)=1} s_m \right) \underline{q}_a \leq q_a, \quad (10c)$$

$$\left(1 - \sum_{m \in \mathcal{M}' : d(a, m)=-1} s_m \right) \bar{q}_a \geq q_a \quad (10d)$$

Recall that m^{fwd} and m^{bwd} denote the number of operation modes that specify a forward or backward flow direction, respectively. This formulation yields (at most) 2

constraints with $m^{fwd} + m^{bwd} + 2$ nonzero elements compared to the formulation given by (8d) and (8e) which yields $m^{fwd} + m^{bwd}$ constraints with $2(m^{fwd} + m^{bwd})$ nonzero elements. Finally, we can use the same trick as in (10a) and (10b) to enforce that an active element is in state active or bypass according to the selected operation mode by having a single constraint enforce this state, choosing the one which occurs in fewer operation modes:

$$\sum_{m \in \mathcal{M}' : m_a=(1,0)} s_m = s_a^{ac} \quad \text{or} \quad (10e)$$

$$\sum_{m \in \mathcal{M}' : m_a=(0,1)} s_m = s_a^{bp} \quad (10f)$$

In comparison to (9a) and (9b) this yields one constraint instead of two, again at least halving the number of nonzero elements. All in all, in this formulation the number of constraints does not depend on the number of operation modes anymore and there are substantially fewer nonzero elements.

For each fully specified operation mode we can now compute tight pressure and inflow bounds by solving the optimization problem given by

- the model (1)–(7) for the network elements,
- the decision variables fixed according to the operation mode, and
- suitable objective functions, i.e., minimizing/maximizing the inflow/pressure at the boundary nodes.

Then, with $\underline{p}_u(m)$, $\bar{p}_u(m)$ and $\underline{b}_u(m)$, $\bar{b}_u(m)$ denoting the pressure bounds and inflow bounds, respectively, for node u in operation mode m , the following inequalities are valid:

$$\sum_{m \in \mathcal{M}'} \underline{p}_u(m) s_m \leq p_u \leq \sum_{m \in \mathcal{M}'} \bar{p}_u(m) s_m \quad \text{for all } u \in V_{\pm}, \quad (11a)$$

$$\sum_{m \in \mathcal{M}'} \underline{b}_u(m) s_m \leq b_u \leq \sum_{m \in \mathcal{M}'} \bar{b}_u(m) s_m \quad \text{for all } u \in V_{\pm}. \quad (11b)$$

Aiming again for sparse constraints, we can improve constraints (11a) and (11b) by just enforcing bounds that are tighter than the variable bounds:

$$\underline{p}_u + \sum_{\substack{m \in \mathcal{M}' : \\ \underline{p}_u(m) > \underline{p}_u}} (\underline{p}_u(m) - \underline{p}_u) s_m \leq p_u \leq \bar{p}_u + \sum_{\substack{m \in \mathcal{M}' : \\ \bar{p}_u(m) > \bar{p}_u}} (\bar{p}_u(m) - \bar{p}_u) s_m, \quad (12a)$$

$$\underline{b}_u + \sum_{\substack{m \in \mathcal{M}' : \\ \underline{b}_u(m) > \underline{b}_u}} (\underline{b}_u(m) - \underline{b}_u) s_m \leq b_u \leq \bar{b}_u + \sum_{\substack{m \in \mathcal{M}' : \\ \bar{b}_u(m) > \bar{b}_u}} (\bar{b}_u(m) - \bar{b}_u) s_m, \quad (12b)$$

for all $u \in V_{\pm}$.

If no bound is tighter, there is no additional row in the LP.

An active compressor group has much tighter flow bounds than one in bypass. We can exploit the fully specified operation modes to benefit from this by tightening the

flow bounds for the operation modes where a compressor group is active. Denoting by $\underline{q}_a^{\text{ac}}$ and \bar{q}_a^{ac} the lower and upper flow bounds if compressor group $a \in A_{\text{cg}}$ is active, we can follow the above sparsification scheme and add the valid inequalities

$$\begin{aligned} &\underline{q}_a + \sum_{m \in \mathcal{M}': m_a=(1,0), \underline{q}_a^{\text{ac}} > \underline{q}_a} (\underline{q}_a^{\text{ac}} - \underline{q}_a) s_m \\ &\leq q_a \leq \bar{q}_a + \sum_{m \in \mathcal{M}': m_a=(1,0), \bar{q}_a^{\text{ac}} < \bar{q}_a} (\bar{q}_a^{\text{ac}} - \bar{q}_a) s_m \quad \text{for all } a \in A_{\text{cg}}. \end{aligned} \tag{12c}$$

From now on, we call the formulation using the original operation modes the *compact model* and the one using the above sparse model (10a)–(10f) for fully specified operation modes the *extended model*. Moreover, the *extended model* enriched with additional bounds for each operation mode, e.g., the above bounds (12a)–(12c), will be called *bounded extended model*.

3 Topology simplification for a single operation mode

Our goal in this section is to simplify the topology of a single fully specified operation mode of a compressor station to obtain a reduced representation suitable for comparing operation modes via graph isomorphism detection.

We consider the network N^m corresponding to a fully specified operation mode m derived from the compressor station network as follows.

1. We remove all closed elements. Note that this step may render the network disconnected.
2. We remove all weakly connected components that do not contain a boundary node or an active compressor group. (Note that in a component without boundary nodes and active compressor groups, there cannot be a positive flow. In case there are no boundary nodes but active compressor groups, there may be a cycle flow.)
3. We replace all short pipes by two opposing short pipes with lower flow bound equal to zero.
4. We replace all open valves and compressor groups that are in *bypass*. If they have a fixed flow direction, they are equivalently replaced by a single short pipe with lower flow bound equal to zero. Otherwise, they are replaced by two opposing short pipes with lower flow bound equal to zero.

Note that our compressor station network is based on a simple, directed graph, and these transformations do not introduce any loops or parallel edges.

However, in order to cope with subsequent reduction steps we model the network N^m corresponding to operation mode m as a directed multigraph network $N^m = (V^m, A^m, \underline{q}^m, \bar{q}^m, \underline{p}^m, \bar{p}^m, t^m, h^m)$ where

- V^m and A^m are a node and a arc set, respectively,
- $\underline{q}^m, \bar{q}^m: A^m \rightarrow \mathbb{R}$ are maps providing lower and upper bounds for the flow along each arc,
- $\underline{p}^m, \bar{p}^m: V^m \rightarrow \mathbb{R}$ are maps providing lower and upper bounds for the pressure at each node, and

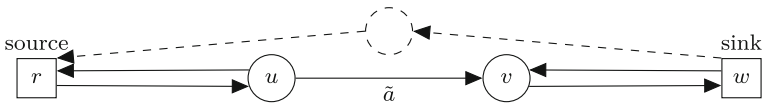


Fig. 3 All arcs a in this network are short pipes with $\underline{q}_a^m = 0$ and $\bar{q}_a^m = \infty$ according to (13d). If we consider the network induced by $\{r, u, v, w\}$ (all solid arcs), then all short pipes except \bar{a} can be contracted since they do not contain information about constraints to the flow direction. The short pipe \bar{a} may not be contracted as it carries the information that no flow is permitted from w to r . However, \bar{a} could be contracted if the network already contained some other directed path of short pipes from w to r , e.g., the dashed path

– $t^m, h^m : A^m \rightarrow V^m$ are maps indicating the tail node and head node of each arc, respectively.

We remark that the described transformations are not only valid with respect to network topology and admissible flows, but also with respect to admissible pressures since the constraints for open valves or bypassed compressor groups are equivalent to those of short pipes. Hence, the arc set A^m consists only of short pipes with non-negative flow and active compressor groups. Thus, the model for a single operation mode is given by

$$0 \leq \underline{p}_u^m \leq p_u \leq \bar{p}_u^m \quad \text{for all } u \in V^m, \tag{13a}$$

$$0 \leq \underline{q}_a^m \leq q_a \leq \bar{q}_a^m \quad \text{for all } a \in A^m, \tag{13b}$$

$$b_u = 0 \quad \text{for all } u \in V_0^m, \tag{13c}$$

$$\underline{q}_a^m = 0, \bar{q}_a^m = \infty \quad \text{for all } a \in A_{sp}^m, \tag{13d}$$

$$p_{t_a^m} = p_{h_a^m} \quad \text{for all } a \in A_{sp}^m, \tag{13e}$$

$$(p_{t_a^m}, p_{h_a^m}, q_a) \in P_a \subseteq \mathbb{R}_{\geq 0}^3 \quad \text{for all } a \in (A_{cg}^m \cup A_{cv}^m). \tag{13f}$$

However, the resulting network may be unnecessarily complex, as short pipes are often redundant (cf. Proposition 1 below). In order to obtain minimal networks that are useful for detecting redundancies as described in Sect. 4 we can reduce the size of the network by contracting a short pipe $a \in A_{sp}^m$ as follows. We remove a from the graph, identify the incident nodes of a , only keeping h_a^m , and update the pressure bounds of the remaining node h_a^m to be the intersection of the pressure intervals for the original two nodes t_a^m and h_a^m . If there are any other arcs between the two nodes, we do keep them as self-loops. If a itself is a self-loop, we simply remove it from the graph. But we need to be careful when applying this contraction since the remaining short pipes sometimes do carry important information on the topology of feasible flows. An example of when contracting a short pipe would lead to a loss of information is shown in Fig. 3.

We now devise a criterion for safely removing short pipes. For this, we assume the operation mode m to be fixed in the following. We consider the short pipe subgraph of N^m, G^{sp} , its set of sources V_+^{sp} , its set of sinks V_-^{sp} and for all sources $w \in V_+^{sp}$ the set $R_{N^m}(w) \subseteq V_-^{sp}$ of sinks reachable using only short pipes:

$$G^{sp} := (V^m, A_{sp}^m), \tag{14}$$

$$V_+^{sp} := V_{\pm} \cup \{h_a^m \mid \forall a \in (A_{cg}^m \cup A_{cv}^m)\}, \tag{15}$$

$$V_-^{SP} := V_{\pm} \cup \{t_a^m \mid \forall a \in (A_{cg}^m \cup A_{cv}^m)\}, \tag{16}$$

$$R_{N^m}(w) := \{u \mid u \in V_-^{SP} \wedge \exists w - u\text{-path in } G^{SP}\} \quad \text{for all } w \in V_+^{SP}. \tag{17}$$

Proposition 1 Consider a short pipe \tilde{a} from $u := t_a^m$, $u \in V^m \setminus \{V_+^{SP}\}$ to $v := h_a^m$ and the network N' arising from N^m when contracting \tilde{a} to v . If this contraction does not affect the reachability from sources to sinks in the short pipe subgraph of the network, i.e., if

$$R_{N^m}(w) = R_{N'}(w) \quad \text{for all } w \in V_+^{SP} \tag{18}$$

then for every admissible flow-pressure combination (p', q') for N' there exists an admissible flow-pressure combination (p, q) for N^m such that

$$q'_a = q_a \quad \text{for all } a \in (A_{cg}^m \cup A_{cv}^m), \tag{19}$$

$$b'_w = b_w \quad \text{for all } w \in V_{\pm}, \tag{20}$$

$$p'_w = p_w \quad \text{for all } w \in V_{\pm}, \tag{21}$$

and vice versa.

Proof We show that given a flow-pressure combination (p', q') for N' we can construct a flow-pressure combination (p, q) for N^m that fulfills (19), (20) and (21). This is straight-forward with respect to the pressure p' and, using a flow decomposition of q' , for all flow paths and cycles which do not contain v . For all flow paths and cycles that do contain v , we show that there exist corresponding flow paths and cycles as part of the flow decomposition of q such that the proposition holds.

Consider a flow-pressure combination (p', q') for N' . Define the pressure vector p by

$$p_w := \begin{cases} p'_w & w \in V^m \setminus \{u\}, \\ p'_v & w = u. \end{cases} \tag{22}$$

Then, p, p' trivially satisfy (21). Also, p is admissible pressure: By definition, it satisfies (13e) for \tilde{a} . It satisfies (13a) for u and v since by our definition of contraction the pressure bounds on v in N' are the intersection of the pressure intervals for u and v in N^m and p' is admissible. Also, it satisfies (13a), (13e) and (13f) (with respect to pressure) for all other nodes and arcs, because p' does.

We obtain q by constructing a flow decomposition of it based on a flow decomposition of q' . Consider a flow decomposition of q' into paths and cycles. Let $P' = v_1, \dots, v, \dots, v_k$ be a path with $v_1, v_k \in V_{\pm}$ and let κ be the flow value on P' in the flow decomposition of q' . We now construct a corresponding path P in N^m with identical endpoints and flow value, but as part of the flow decomposition of q . If P' does not contain v , we simply set $P := P'$. Otherwise, the idea is to partition P into short pipe subpaths that go between compressor groups or boundary nodes. Exactly

one of these subpaths contains v and is thus possibly not in N^m . But, by (18) we can replace this short pipe subpath in N' with a short pipe subpath in N^m that has the same endpoints.

Consider the partition of P' into short pipe subpaths $P'_1, \dots, P'_n, \dots, P'_r$ by removal of all compressor groups and let $P'_n = v_i, \dots, v, \dots, v_j$. By (15) and (16) we have $v_i \in V_+^{sp}$ and $v_j \in V_-^{sp}$, and by (17) we have $v_j \in R_{N'}(v_i)$. Therefore, by (18) we have $v_j \in R_{N^m}(v_i)$. Let $P_n = v_i, \dots, v_j$ be a short pipe path in N^m and define P to be P' with P'_n replaced by P_n .

Similarly, let C' be a cycle in N' with flow value κ' in the flow decomposition of q' . If C' contains a compressor group, we can employ the same reasoning as before to obtain a cycle C in N^m with identical flow value κ' , but as part of the flow decomposition of q . Otherwise, if C consists of short pipes only, we let the flow value on it as part of the flow decomposition of q be zero.

Altogether, this gives a flow decomposition of q . Since by construction from q' to q only the flow on short pipes may be affected, (19), (13b) and (13f) (with respect to flow) hold. Also, for every path in the flow decomposition of q' there exists exactly one corresponding path in the flow decomposition of q and that path has identical endpoints and flow value. By this, (20) and (13c) hold. Therefore, q is both admissible and satisfies the proposition.

Analogously, the same reasoning can be employed in the other direction to construct a flow-pressure combination (p', q') for N' from a flow-pressure combination (p, q) for N^m . □

We remark that the proposition can be stated and proved similarly without the premise that $u \in V^m \setminus \{V_+^{sp}\}$, but it requires some technical refinements in the statement and the proof of our proposition in order to account for the case that $u \notin N'$.

4 Detection of redundant operation modes

In the previous section we have presented all that is needed for simplifying the network topology for a single operation mode. We now briefly summarize how we use this in order to obtain reduced network representations for all operation modes. Following this we explain how these reduced network representations allow us to detect redundant operation modes.

We reduce the network representation of a single operation mode by iteratively contracting all short pipes which are incident to at most one boundary node and which can be contracted in accordance with Proposition 1. On the reduced network we tighten all flow and pressure bounds by solving the optimization problem given by (13a)–(13f) with suitable objective functions. All operation modes turning out to be infeasible are discarded.

In a last step, and this is the contribution of this section, we detect redundancy between the remaining feasible operation modes by an extended isomorphism test between their reduced network representations. Intuitively, an operation mode m^2 is “at least as good as” another operation mode m^1 if there is a mapping between the operation modes’ networks such that every solution for m^1 is represented by a solution

for m^2 . In particular, every feasible flow and pressure value for an element of N^1 needs to be feasible for the elements' image in N^2 . We capture this formally in the following definition.

Definition 1 For operation modes m^1 and m^2 with reduced network representations $N^1 = (V^1, A^1, \underline{q}^1, \bar{q}^1, \underline{p}^1, \bar{p}^1, t^1, h^1)$ and $N^2 = (V^2, A^2, \underline{q}^2, \bar{q}^2, \underline{p}^2, \bar{p}^2, t^2, h^2)$, respectively, we say that m^1 is dominated by m^2 if all of the following hold:

1. There exists a graph isomorphism defined by the bijections $\phi : V^1 \rightarrow V^2$ and $\psi : A^1 \rightarrow A^2$, i.e.,

$$t_{\psi(a)}^2 = \phi(t_a^1), \quad h_{\psi(a)}^2 = \phi(h_a^1) \quad \text{for all } a = (t_a^1, h_a^1) \in A^1. \tag{23}$$

2. All boundary nodes are fixed points of ϕ :

$$\phi(u) = u \quad \text{for all } u \in V_{\pm}^1. \tag{24}$$

3. All compressor groups and control valves are fixed points of ψ :

$$\psi(a) = a \quad \text{for all } a \in (A_{cg}^1 \cup A_{cv}^1). \tag{25}$$

4. Short pipes are mapped to short pipes by ψ :

$$a \in A_{sp}^1 \iff \psi(a) \in A_{sp}^2 \quad \text{for all } a \in A^1. \tag{26}$$

5. The node pressure bounds in N^2 are not tighter than their pre-image ones in N^1 :

$$\underline{p}_u^1 \geq \underline{p}_{\phi(u)}^2, \quad \bar{p}_u^1 \leq \bar{p}_{\phi(u)}^2 \quad \text{for all } u \in V^1. \tag{27}$$

Observe that (25) trivially implies that feasible flow/pressure values for a compressor group or control valve in m^1 are feasible in m^2 , too. Hence the reduction technique we propose is valid for every compressor group and control valve model (in particular, for different models of compressor units represented by the compressor group).

Proposition 2 Let two operation modes m^1 and m^2 be given as in Definition 1, with m^1 being dominated by m^2 . Consider a gas network optimization problem such that

- the objective function does not depend on the flow over short pipes, and
- the flow over short pipes is only constrained by flow conservation.¹

Then the gas network optimization problem arising by dropping operation mode m^1 is equivalent to the original one.

Proof We first observe that a solution that is feasible for the reduced network N^1 is also feasible for the reduced network N^2 . N^1 and N^2 are topologically equivalent by (23)–(26). Requirement (27) ensures that the pressure values of the solution for N^1

¹ Observe that there are no flow bounds for short pipes in our model, see (5a).

are also feasible for N^2 . The analogous feasibility of flow values follows from the fact that all element types are conserved by the arc map ψ according to (25) and (26).

The equivalence of the operation modes m^1 and m^1 w.r.t. feasibility follows from the equivalence of the reduced networks and Proposition 1.

Finally, to establish the equivalence w.r.t. to the objective function, we observe that in the construction in the proof of Proposition 1 only flow values of short pipes are possibly changed. By assumption, this does not change the objective value.

Due to Proposition 2, we can iteratively discard dominated operation modes such that for our new, redundancy-free set of operation modes \mathcal{M}' no operation mode $m^1 \in \mathcal{M}'$ is dominated by another operation mode $m^2 \in \mathcal{M}'$.

We note that the extended isomorphism test from Definition 1 can be performed on the basis of any general graph isomorphism test, which generates all isomorphisms between N^1 and N^2 (with V_{\pm}^1 as fixed points), and by checking if any of these isomorphisms fulfills (24)–(27). In practice, we use an existing implementation (Hagberg et al. 2008) of the VF2 graph isomorphism algorithm (Cordella et al. 2001, 2004) to efficiently generate all isomorphisms between N^1 and N^2 that fulfill (24)–(26) and then check whether (27) holds.

5 Classification of operation modes by compressor station flow patterns

Because arcs in our network model can admit flow in either direction, most boundary nodes cannot generally be classified as either source or sink of the compressor station. Instead, they may take on either role, depending on the general network situation and the selected operation mode of the station. Thus, compressor stations often admit multiple *flow patterns* where a *flow pattern* describes a partition of the boundary nodes into sinks and sources. This corresponds to an inflow bound of the form $b_u \leq 0$ (if u is a sink) or $b_u \geq 0$ (if u is a source) for every boundary node u . We say that a compressor station admits a flow pattern if it admits a flow which satisfies the flow pattern.

However, when its operation mode is fixed, a compressor station usually admits very few or even just a single flow pattern, which is mostly due to the fixed flow direction of active compressor groups. Therefore, we suggest as an additional modeling step the classification of operation modes by the flow patterns they admit in conjunction with introducing binary variables for the different flow patterns. The idea is that, when solving a gas network optimization task, branching on the flow patterns should quickly identify the viable flow patterns and thereby reduce the set of operation modes that need to be considered.

We incorporate this into our model as follows. For every flow pattern F_i , $i \in I$, that a compressor station admits, let \mathcal{F}_i be the set of operation modes of the station that admit this flow pattern. Then, every feasible operation mode is contained in at least one of the sets \mathcal{F}_i for some $i \in I$. For all $i \in I$ let $f_i \in \{0, 1\}$ be a binary variable that is 0 if flow pattern F_i is not chosen, i.e., if the selected operation mode does not

admit F_i . This is expressed by the constraints

$$f_i = \sum_{m \in \mathcal{F}_i} s_m \quad \text{for all } i \in I. \quad (28)$$

Note that due to the fact that a single operation mode can admit multiple flow patterns $f_i = 1$ does not generally imply that flow pattern F_i is chosen. It just means that the selected operation mode also admits this flow pattern. This modeling choice was made to avoid splitting the operation modes further to achieve a one-to-one correspondence between operation modes and flow patterns, which would inflate the model. Hence, the model only requires that at least one flow pattern is chosen:

$$\sum_{i \in I} f_i \geq 1. \quad (29)$$

Finally, we couple the choice of the flow pattern with flow bounds for the boundary nodes according to the flow pattern:

$$\sum_{i \in I} \underline{b}_u(F_i) f_i \leq b_u \leq \sum_{i \in I} \bar{b}_u(F_i) f_i \quad \text{for all } u \in V_{\pm} \quad (30)$$

with

$$\underline{b}_u(F_i) := \begin{cases} \min_{m \in \mathcal{F}_i} \underline{b}_u(m) & \text{if } u \text{ is a sink for } F_i \\ 0 & \text{otherwise} \end{cases} \quad \text{for all } i \in I, u \in V_{\pm},$$

$$\bar{b}_u(F_i) := \begin{cases} \max_{m \in \mathcal{F}_i} \bar{b}_u(m) & \text{if } u \text{ is a source for } F_i \\ 0 & \text{otherwise} \end{cases} \quad \text{for all } i \in I, u \in V_{\pm}.$$

6 Computational results

When applied to a single compressor station with a linear model for the compressor groups, earlier computational results in Hiller et al. (2016) showed some benefit of the proposed model extensions for detecting infeasibility early. However, presumably due to the simplified nature of this setting, the overall impact was negligible.

We therefore evaluated the effectiveness of our approach on entire gas networks with several compressor stations. For this, we consider the openly available gas network benchmark instances GasLib-582 (Humpola et al. 2017) and GasLib-4197 (Schmidt et al. 2017), complemented by real network data from Open Grid Europe GmbH, Germany's biggest gas network operator. For the latter, we considered the two network variants HN-SN and HN-AB studied in Hiller et al. (2015a, p.234). These two variants differ slightly in their topology and more in the properties of the corresponding *nominations*, i.e., particular gas flow situations, that are considered. The nominations for HN-SN are generated based on a model for network usage (Hiller et al. 2015b), whereas the nominations of HN-AB represent expert scenarios. In both cases, the nominations

Table 1 Overview of the sizes of the considered networks

Element type	GasLib-582	HN-SN	<i>HN – AB</i>	GasLib-4197
Compressor group	5	6	7	12
Control valve	23	23	26	120
Valve	26	34	33	426
Pipe	278	452	498	3537
Short pipe	269	97	116	343
Resistor	8	9	9	28
# Compressor stations	2 (1)	2 (1)	2 (2)	6 (3)
# Nominations	418	57	43	207

The row “# compressor stations” gives the number of *complex* compressor stations consisting of at least two compressor group arcs. The number of simple compressor stations consisting of a compressor group only is given in parentheses

are designed to use the network at its limits. As there are very many nominations for each GasLib instance, we selected every tenth nomination from GasLib-582 and GasLib-4197 in order to limit the computational effort. The selection was done based on the numbering of the nominations encoded in the filename; we just used the nominations ending with a 0. Moreover, we ignored nominations for which infeasibility was detected during the preprocessing of the model generation. An overview of the size of the networks and the number of considered nominations is given in Table 1.

Unfortunately, the original network data does not provide the subnetworks that model compressor stations. However, they do provide sets of operation modes for groups of related active elements. We derive suitable subnetworks via the following heuristic. First, we remove all pipes longer than 2.0 km. For each remaining component with at least one compressor group arc, we determine the set of active elements involved in the operation modes for the compressor arcs to form a set of *unremovable arcs*. We then successively remove leaves from the component as long as no unremovable arc is deleted. This step is intended to provide a “compact” subnetwork. Although this heuristic provides reasonable subnetworks, we believe it is preferable to obtain the subnetworks as part of the network data and the network modeling process. In particular, one of the subnetworks determined for GasLib-4197 (“station 1” below) is unusually large and has as many as 20 boundary nodes. This is not an artifact of our heuristic, but a result of the original GasLib-4197 data which specifies operation modes for groups of active elements that are unusually far apart from each other.

Note that the subnetworks for a compressor station may contain passive elements like resistors or pipes with small length. The reason is that the piping in real-world compressor stations yields a certain pressure drop, so network modelers sometimes include passive elements to account for this. To avoid model inconsistencies, we replace these elements by short pipes in the original network which we use as the basis of our preprocessing and to construct all MILP models.

Table 2 shows details on the number of operation modes after the different preprocessing steps proposed in the preceding sections. For instance, for the first compressor station of HN-SN the original data describes 53 operation modes which, as explained

in Sect. 2, for all controllable elements specify one of the states *open* or *closed* and possibly a flow direction. But they do not specify for open compressor groups (or control valves) whether these are *active* or in *bypass*. Enumerating all combinations of *active* and *bypass* leads to 1249 fully specified operation modes, 194 of which remain after filtering out those that are invalid with respect to computed flow and pressure bounds on the elements or have been detected as redundant by the method presented in Sect. 4. These 194 remaining operation modes are then used for the formulation of our *extended model* and *bounded extended models*. The total duration of all preprocessing steps combined is at most 3 minutes per network for all networks except GasLib-4197, for which it took 6216 seconds. Note that this computation needs to be performed only once for each network and not for each nomination, and in practice we are interested in computing many nominations (which vary) for a gas network (which does not vary). Therefore this preprocessing time is not included in the running times reported hereafter.

For the first station of GasLib-4197, the very high number of 2304 fully specified operation modes arising from only 10 original operation modes is striking, in particular since the 53 original operation modes for HN-AB and HN-SN yield only 1249 fully specified operation modes. A closer investigation reveals that several of the 10 original operation modes have 9 active elements in state *open*, corresponding to 512 ways to obtain a fully specified operation mode. In contrast, the number of open active elements for an operation mode in HN-AB and HN-SN is at most 6.

For each nomination (consisting essentially of a gas demand vector) we use the lamatto framework (LaMaTTO 2014) that has also been used in Geißler et al. (2012), [7], Geißler et al. (2015b, 2018) to generate the compact MILP model. We chose this model as it is a standard MILP model and thus suitable for experiments with standard MIP solvers. The pressure bounds $\underline{p}_u(m)$ and $\bar{p}_u(m)$ for boundary node $u \in V_{\pm}$ in operation mode m and lower and upper flow bounds $\underline{q}_a^{\text{ac}}$ and \bar{q}_a^{ac} for an active compressor group $a \in A_{\text{cg}}$ have been obtained as described in Walther and Hiller (2017). We extended lamatto to generate the sparse constraints (10a)–(10f) and (12a)–(12c) for our *extended model*. The file with the *extended model* is then the basis for adding the bound constraints to obtain the *bounded extended models*.

In total, we compared 9 different model variants: the *compact model* (C), the *extended model* (E) and the *bounded extended models* which consist of the *extended model* enriched with flow bounds (F) as in (12b) and (12c), pressure bounds (P) as in (12a), flow pattern constraints (M) as in (28)–(30) and all combinations of the former (FP, FM, PM, FPM). For Gaslib-4197, we did not consider the variants with the pattern constraints as they blow up the model size due to the many boundary nodes. All computations were performed with Gurobi 8.1.0 (Gurobi 2020) single-threaded on an Intel i7-4790 CPU with 3.60GHz and with a time limit of four hours. We performed all computations three times with the same three different random seeds and averaged the performance over these three runs. The time limit was only exceeded for GasLib-4197 (C: 57 times, E: 15, F: 21, P: 14, FP: 27). We remark that some nominations are inherently numerically unstable, so we ignored them in the analysis. More precisely, we consider a nomination to be numerically unstable if Gurobi found conflicting solutions for the three runs of at least one model variant. In all cases where instability was observed for one of our extended models the other seeds provided the

Table 2 Overview on the compressor station subnetworks, in particular the number of operation modes after the different processing steps

	# of Boundary nodes	# of operation modes			
		Original data	After full specification	After removal of invalid ones	After removal of redundant ones
Gaslib-582					
Station 1	5	4	12	12	10
Station 2	3	19	225	81	52
HN-SN					
Station 1	3	53	1249	409	194
Station 2	4	6	8	8	6
Station 3	2	3	5	5	5
HN-AB					
Station 1	3	53	1249	405	198
Station 2	4	9	11	11	9
Station 3	2	3	5	5	5
Station 4	2	3	5	5	5
Gaslib-4197					
Station 1	20	10	2304	2304	1824
Station 2	3	4	10	8	6
Station 3	7	4	6	6	6
Station 4	4	4	6	4	4
Station 5	5	2	3	3	3
Station 6	6	3	3	3	3

The operation modes in the left-most column are used in the *compact model*. The operation modes in the right-most column are used in the *extended model* and the *bounded extended models*

same solution value as the *compact model*, indicating that the instability is not due to the reformulated model per se. Overall, instability was detected for one nomination in the original model and for 3 additional nominations in one of our extended models (HN-AB: 1 nomination, GasLib-582: 1, GasLib-4197: 2).

Table 3 and Table 4 show the distribution of running times and number of nodes (averaged over the three random seeds), respectively, for all models over all nominations. As we can see, all our models improve upon the original *compact model C*. A significant part of the improvement comes from using (reduced) fully specified operation modes obtained through our preprocessing method described in Sect. 4 as evidenced by the performance of the *extended model E*. The addition of constraints in the *bounded extended models* further improves performance, except for GasLib-4197, with the best performing variant varying for the considered instance sets. While the improvement on GasLib-582 is marginal, suggesting that the benchmark might be too easy for our method as overall running times are also comparatively low, our models show their true potential on the real gas networks HN-SN, HN-AB, and GasLib-4197. For these instance sets, all our models reduce the average number of nodes by at

Table 3 Distribution of running times over all nominations

	Running time in seconds				
	Mean	0.5 quantile	0.75 quantile	0.9 quantile	Max
Gaslib-582					
C	61.6	59.3	77.8	96.6	463.5
E	56.4	56.3	74.7	87.4	237.2
F	56.7	55.1	72.1	88.8	257.0
M	55.5	55.1	71.8	87.3	291.7
P	57.7	58.1	74.3	89.3	364.6
FM	57.6	58.1	74.9	92.7	262.9
FP	57.3	57.7	74.5	91.8	238.0
PM	56.3	56.2	73.7	87.2	289.5
FPM	58.0	54.6	74.8	91.6	305.7
HN-SN					
C	150.5	131.8	172.7	265.7	454.4
E	96.2	93.6	130.9	157.8	234.7
F	103.3	83.6	114.7	187.7	634.4
M	95.2	89.0	122.5	156.5	359.4
P	84.4	76.8	106.6	133.6	257.5
FM	87.8	77.3	109.7	148.4	313.4
FP	85.8	75.2	114.7	137.4	257.8
PM	81.7	73.9	97.5	127.1	397.6
FPM	81.6	71.6	103.4	125.5	257.2
HN-AB					
C	118.0	72.4	161.9	322.3	487.7
E	89.2	45.0	136.3	217.8	316.5
F	86.7	60.7	125.0	210.2	249.2
M	84.9	42.6	143.0	203.8	320.8
P	87.0	46.4	140.7	207.2	281.8
FM	85.4	46.1	138.8	203.8	302.6
FP	79.4	42.7	129.6	174.7	254.6
PM	82.0	45.5	129.4	209.7	247.1
FPM	78.3	47.9	121.9	186.2	249.8
Gaslib-4197					
C	2741.6	1098.2	3337.7	7775.1	14400.1
E	1477.4	697.0	1462.9	2840.0	14400.1
F	1637.4	711.9	1579.7	4150.1	13307.2
P	1556.5	698.8	1379.4	3670.7	14400.1
FP	1803.0	728.5	1726.0	5400.2	14400.1

The best value for each metric (quantile or maximum) is highlighted in bold

Table 4 Distribution of number of search tree nodes over all nominations

	Number of search tree nodes				
	Mean	0.5 quantile	0.75 quantile	0.9 quantile	Max
Gaslib-582					
C	703.3	623.0	887.3	1262.6	10274.0
E	631.8	569.0	824.3	1091.3	5783.7
F	611.2	550.0	725.0	1049.0	4046.3
M	602.7	554.3	763.7	975.8	4420.0
P	662.3	625.3	844.7	1100.0	8882.0
FM	624.9	557.0	769.3	1125.1	4299.3
FP	633.7	599.3	790.7	1049.3	4587.3
PM	635.1	605.3	808.7	1168.7	4170.0
FPM	625.6	584.7	781.3	1049.7	3796.3
HN-SN					
C	2212.1	1948.7	2502.3	3334.2	10075.3
E	1072.6	1031.0	1432.7	1797.8	3140.7
F	1787.3	950.0	1403.7	2045.6	39850.3
M	1066.8	952.3	1525.0	1732.3	2853.3
P	992.5	942.7	1436.3	1786.9	2109.7
FM	969.9	924.3	1307.7	1712.7	3005.3
FP	1044.5	979.3	1389.7	1894.7	3019.7
PM	895.7	874.0	1158.0	1540.3	3425.7
FPM	901.9	811.3	1137.0	1627.5	4305.3
HN-AB					
C	1706.5	1229.2	2127.1	4032.5	7448.7
E	863.5	935.7	1326.3	1990.3	2513.3
F	858.3	964.7	1357.4	1793.1	2166.0
M	740.4	850.3	1163.3	1542.1	1994.0
P	855.7	827.3	1528.3	1766.7	2140.7
FM	812.0	905.8	1319.0	1700.2	2387.7
FP	767.5	837.5	1233.6	1641.4	1994.3
PM	736.7	780.7	1195.9	1527.3	2045.0
FPM	708.6	765.7	1145.0	1436.2	1753.3
Gaslib-4197					
C	20312.7	5389.3	21211.7	59587.5	295297.0
E	6354.4	2443.3	4849.3	14029.6	76750.7
F	9117.5	2401.0	5649.0	25748.3	177179.3
P	8123.2	2394.0	4917.3	14964.3	257718.7
FP	9999.1	2383.3	6447.7	35315.7	121797.3

The best value for each metric (quantile or maximum) is highlighted in bold

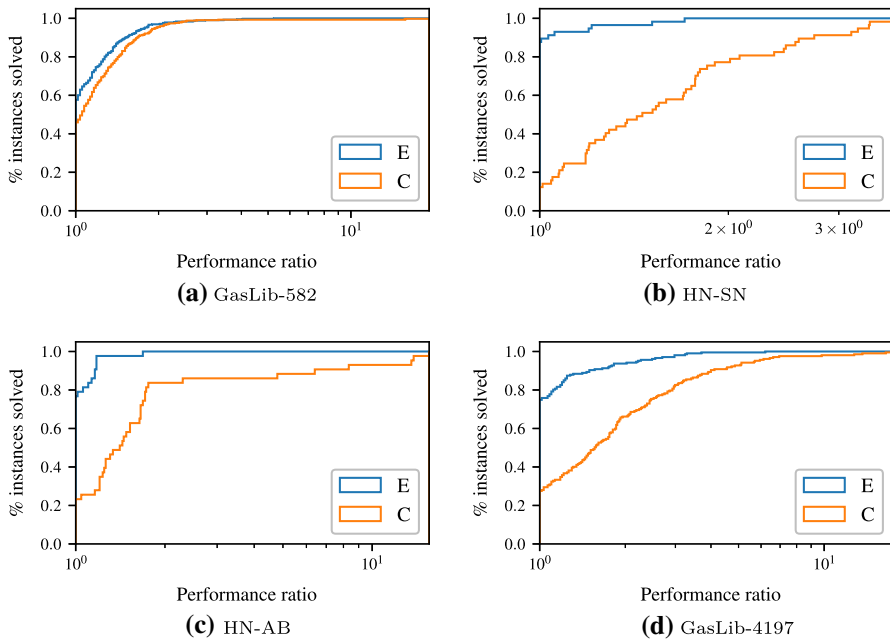


Fig. 4 Performance profiles for the running time on the different networks

least half compared to the *compact model C*, and the average running time is reduced by 46%, 31%, and 37%, respectively, for the best performing model variant. The impact for the 0.9-quantile is even more pronounced (reduction by 53%, 46%, and 63%, respectively), indicating that the alternative models particularly help with the hard instances. An interesting observation is that models taking the flow patterns into account usually improve over the variant without the flow pattern. This suggests being able to branch on the flow pattern is frequently useful.

Figure 4 presents the performance profiles (Dolan and Moré 2002) of the running time for the *extended model E* and the *compact model C* in comparison, clearly showing the superior performance of the former. The performance profile of a model essentially provides the probability that the running time to solve this model is within a factor of τ of the running time for the better of the two models compared. As we can see, the probability that the *extended model E* requires a shorter running time is always higher than that probability for the *compact model C*.

In summary, we can conclude that for all networks the *extended model E* using fully specified operation modes yields a substantial reduction both in running time and in the number of solving nodes needed. Except for GasLib-4197, a stronger reduction is offered by the *bounded extended models*, but the best variant varies between the instance sets. We believe that the fact that the *bounded extended models* provide no improvement but a poorer performance than the *extended model* is related to the mentioned issues with the subnetwork having very many boundary nodes.

7 Conclusion

We proposed models based on fully specified operation modes in order to address the issues of the existing model for operation modes outlined in the introduction. The evaluation of our model variants on reference benchmark data for gas networks shows that the models reduce the average time required by the state-of-the-art solver Gurobi by 10–46%, depending on the instance set. Our model variants help in particular to reduce the running times for instances that are hard to solve for the original model.

Our methods may be extended in several ways, both to be more realistic and also to detect further redundancies. For our computational experiments we modeled all passive elements in compressor station subnetworks by short pipes, which of course is a coarse approximation. This can be avoided by handling these elements like compressor groups and control valves, at the possible expense of having more irredundant operation modes. If several elements in a subnetwork are actually (technically) identical (as is likely the case for control valves), this symmetry yields redundancies that are currently not detected. This can be resolved e.g., by relaxing the fixed point requirement (25) to allow mapping to different but identical elements. Finally, it is possible to “expand” compressor groups to incorporate their *configurations* into the operation modes and apply the method to this expanded subnetwork. If distinct compressor groups contain identical compressor units, this symmetry can be detected and turned into a reduction of (irredundant) operation modes to consider. However, the number of irredundant operations modes will also increase noticeably.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Carter R (1996) Compressor station optimization: computational accuracy and speed. Technical Report PSIG 9605, Pipeline Simulation Interest Group, 1996
- Cordella LP, Foggia P, Sansone C, Vento M (2001) An improved algorithm for matching large graphs. In: 3rd IAPR-TC15 workshop on graph-based representations in pattern recognition, pp 149–159
- Cordella LP, Foggia P, Sansone C, Vento M (2004) A (sub) graph isomorphism algorithm for matching large graphs. *IEEE Trans Pattern Anal Mach Intell* 26(10):1367–1372
- de Wolf D, Bakhouya B (2012) Optimal dimensioning of pipe networks: The new situation when the distribution and the transportation functions are disconnected. In: Klatte D, Lüthi H-J, Schmedders K, (eds) *Operations research proceedings 2011: selected papers of the international conference on operations research (OR 2011)*, pp 369–374. Springer
- Dolan ED, Moré JJ (2002) Benchmarking optimization software with performance profiles. *Math Program* 91(2):201–213
- Geißler B, Martin A, Morsi A, Schewe L (2012) Using piecewise linear functions for solving MINLPs. In: Lee J, Leyffer S (eds) *Mixed integer nonlinear programming*, vol 154. The IMA volumes in mathematics and its applications. Springer, New York, pp 287–314

- Geißler B, Martin A, Morsi A, Schewe L (2015a) The MILP-relaxation approach. In: Koch T, Hiller B, Pfetsch ME, Schewe L (eds) Evaluating gas network capacities. SIAM-MOS series on optimization, ch 6. SIAM, Philadelphia, pp 103–122. <https://doi.org/10.1137/1.9781611973693.ch6>
- Geißler B, Morsi A, Schewe L, Schmidt M (2015b) Solving power-constrained gas transportation problems using an MIP-based alternating direction method. *Comput Chem Eng* 82:303–317
- Geißler B, Morsi A, Schewe L, Schmidt M (2018) Solving highly detailed gas transport MINLPs: block separability and penalty alternating direction methods. *INFORMS J Comput* 30(2):309–323
- Gurobi L (2020) Optimization. Gurobi optimizer reference manual
- Hagberg AA, Schult DA, Swart PJ (2008) Exploring network structure, dynamics, and function using networkx. Technical report, Los Alamos National Lab.(LANL), Los Alamos, NM (United States)
- Hiller B, Hayn C, Heitsch H, Henrion R, Leövey H, Möller A, Römisch W (2015a) Methods for verifying booked capacities. In: Koch T (ed) Evaluating gas network capacities. SIAM, pp 291–315
- Hiller B, Hayn C, Heitsch H, Henrion R, Leövey H, Möller A, Römisch W (2015b) Methods for verifying booked capacities. In: Koch T, Hiller B, Pfetsch ME, Schewe L (eds) Evaluating gas network capacities. SIAM-MOS series on optimization. SIAM. <https://doi.org/10.1137/1.9781611973693.ch12>
- Hiller B, Saitenmacher R, Walther T (2016) Analysis of operating modes of complex compressor stations. In: Fink A, Fügenschuh A, Geiger M, (eds) Operations research proceedings 2016, Operations Research Proceedings, pp 251–257
- Humpola J, Fügenschuh A, Hiller B, Koch T, Lehmann T, Lenz R, Schwarz R, Schweiger J (2014) The specialized MINLP approach. In: Koch T, Hiller B, Pfetsch ME, Schewe L (eds) Evaluating gas network capacities. SIAM-MOS series on optimization. MOS-SIAM Series on Optimization, Chap 7. SIAM, Philadelphia
- Humpola J, Joormann I, Kanelakis N, Oucherif D, Pfetsch ME, Schewe L, Schmidt M, Schwarz R, Sirvent M (2017) GasLib: a library of gas network instances. Technical report
- Koch T, Hiller B, Pfetsch M, Schewe L (eds) (2015) Evaluating gas network capacities. SIAM, MOS-SIAM Series on Optimization. SIAM
- LAMATTO++: a framework for modeling and solving mixed-integer nonlinear programming problems on networks., 2014. <https://en.www.math.fau.de/edom/projects-edom/mixed-integer-programming/lamatto/>
- Martin A, Möller M, Moritz S (2006) Mixed integer models for the stationary case of gas network optimization. *Math Program* 105(2):563–582
- Pfetsch ME, Fügenschuh A, Geißler B, Geißler N, Gollmer R, Hiller B, Humpola J, Koch T, Lehmann T, Martin A, Morsi A, Rövekamp J, Schewe L, Schmidt M, Schultz R, Schwarz R, Schweiger J, Stangl C, Steinbach MC, Vigerske S, Willert BM (2015) Validation of nominations in gas network optimization: models, methods, and solutions. *Optim Methods Softw* 30(1):15–53
- Rose D, Schmidt M, Steinbach MC, Willert BM (2016) Computational optimization of gas compressor stations: MINLP models versus continuous reformulations. *Math Methods Oper Res* 83(3):409–444
- Ríos-Mercado RZ, Borraz-Sánchez C (2015) Optimization problems in natural gas transportation systems: a state-of-the-art review. *Appl Energy* 147:536–555
- Schmidt M, Alßmann D, Burlacu R, Humpola J, Joormann I, Kanelakis N, Koch T, Oucherif D, Pfetsch ME, Schewe L, Schwarz R, Sirvent M (2017) GasLib: a library of gas network instances. *Data* 2(4)
- van der Hoeven T (2004) Math in gas and the art of linearization. Ph.D. thesis, Rijksuniversiteit Groningen
- Walther T, Hiller B (2017) Modelling compressor stations in gas networks. ZIB Report 17-67, Zuse Institute Berlin
- Wu S, Ríos-Mercado RZ, Boyd EA, Scott LR (2000) Model relaxations for the fuel cost minimization of steady-state gas pipeline networks. *Math Comput Model* 31(2):197–220

Affiliations

Benjamin Hiller^{1,2}  · **René Saitenmacher**^{1,3} · **Tom Walther**^{1,4}

✉ Benjamin Hiller
hiller@zib.de; hiller@atesio.de

René Saitenmacher
r.saitenmacher@tu-berlin.de

¹ Zuse Institute Berlin, Takustraße 7, 14195 Berlin, Germany

² Present Address: atesio GmbH, Bundesallee 89, 12161 Berlin, Germany

³ Present Address: Machine Learning Group, Technische Universität Berlin, Marchstr. 23, 10587 Berlin, Germany

⁴ Present Address: Gnosis Ltd, Berlin, Germany