A PRIMAL HEURISTIC FOR NONSMOOTH MIXED INTEGER NONLINEAR OPTIMIZATION

MARTIN SCHMIDT, MARC C. STEINBACH, BERNHARD M. WILLERT

Abstract. Complex real-world optimization tasks often lead to mixed-integer nonlinear problems (MINLPs). However, current MINLP algorithms are not always able to solve the resulting large-scale problems. One remedy is to develop problem specific primal heuristics that quickly deliver feasible solutions. This paper presents such a primal heuristic for a certain class of MINLP models. Our approach features a clear distinction between nonsmooth but continuous and genuinely discrete aspects of the model. The former are handled by suitable smoothing techniques; for the latter we employ reformulations using complementarity constraints. The resulting mathematical programs with equilibrium constraints (MPEC) are finally regularized to obtain MINLP-feasible solutions with general purpose NLP solvers.

1. Introduction

Mixed-integer nonlinear optimization is a highly versatile tool for modeling application problems in many areas: it covers both discrete aspects of decision making and nonlinear real-world phenomena. However, state-of-the-art algorithms for mixed-integer nonlinear problems (MINLPs) are still far from offering the reliability, performance and robustness of solvers for mixed-integer linear problems (MIPs) or nonlinear optimization problems (NLPs). As a consequence, hard MINLPs that cannot be solved directly are frequently tackled by one of the following approaches:

(1) *MIP-based approach*: Nonlinearities are replaced by local linearizations or by piecewise linear global approximations. This yields MIP models for which the number of discrete variables is often drastically increased due to the linearization techniques; see, e.g., [10, 25, 29, 43] and the references therein.
(2) *NLP-based approach*: Discrete aspects are reformulated with continuous variables and constraints or approximated to obtain an NLP model; see Sect. 2.2 for a brief literature review.

Both approaches offer specific advantages and suffer from inherent drawbacks. The MIP approach shows its strength if the MINLP is dominated by discrete variables and incorporates only a few nonlinear constraints. This will typically lead to a slight increase of the number of discrete variables only. Moreover, standard MIP solvers deliver global solutions (of the linearized problem). The NLP approach is often superior for MINLPs with only few discrete variables but a large number of nonlinear constraints, but in general it delivers only local minima. From a numerical point of view, the MIP approach tends to be more robust in terms of starting points, scaling of the problem, etc., whereas the NLP approach can be very fast.

Of course, ultimately one would like to solve hard MINLP models directly. An essential and generally difficult subtask of dedicated MINLP algorithms consists in finding (approximately) feasible solutions to obtain upper bounds. In this paper we present a primal heuristic for that purpose. It is inspired by the NLP-based

Date: April 2, 2013.

approach sketched above, and we will refer to it as the *MPEC-based approach*. The MINLP models that can be handled have the key property that their (moderate number of) discrete aspects possess equivalent reformulations with problem-specific *complementarity constraints*, yielding nonlinear *mathematical programs with equilibrium constraints* (MPEC). The MPECs are finally regularized by standard techniques so that they can be solved by NLP algorithms. The goal is to deliver approximate MINLP-feasible solutions quickly (or to fail quickly).

Our approach aims at large-scale optimization models arising in real-life applications. Such models frequently involve nonsmooth constraints (continuous but only piecewise C^2) that can be converted to smooth (C^2-)constraints with artificial discrete variables to identify the smoothness domains. Solving MINLP models of this type tends to be extremely hard: algorithms have been observed to spend a lot of time (or fail) to get the artificial discrete "decisions" right, without making any progress toward a minimum. An important aspect of the approach proposed here lies in avoiding this (inappropriate) modeling of nonsmooth constraints. Instead we consider C^0-MINLPs and employ suitable smoothing techniques to obtain C^2-constraints. A precursor of this approach has been successfully applied in operative planning of water networks [6, 7]. In the application part of this paper (Sect. 4) we demonstrate that the proposed MPEC-based approach delivers approximately MINLP-feasible solutions and hence can be used as a primal heuristic in MINLP solver frameworks.

The paper is organized as follows. Sect. 2 gives a formal definition of the problem classes arising in the reformulation of nonsmooth mixed-integer nonlinear problems as smoothed and regularized NLP models. In Sect. 3 the relations between these models are discussed. As a proof of concept the real-world application of validation of nominations in gas transport networks is presented in Sect. 4. Finally we give a brief summary in Sect. 5.

2. A Hierarchy of Optimization Models

Here the reformulation of a given nonsmooth MINLP is presented step by step in order to discuss certain properties of the models and to explain some model transition techniques. The sequence of reformulations finally leads to an NLP model.

In all models we denote constraints by c. A vector of constraints is indexed with a corresponding index set. For instance, $c_E := (c_i)_{i\in E}$, is the vector of all equality constraints; c_I the vector of inequality constraints. Frequently we use superindices to indicate the semantics of constraints. A superindex d marks nonsmooth constraints, s marks smoothed constraints and r marks constraints that result from regularization techniques for MPECs. Constraints without superindex are always assumed to be twice continuously differentiable. Continuous variables are referred to as x and discrete ones as z. Objective functions are denoted by f.

2.1. **Standard Mixed-Integer Nonlinear Problems.**

The general MINLP model is given by

$$\min_{x,z} \quad f(x,z) \tag{1a}$$

$$\text{s.t.} \quad c_E(x,z) = 0, \quad c_I(x,z) \geq 0, \tag{1b}$$

$$x \in \mathbb{R}^{n_x}, \quad z \in \mathbb{Z}^{n_z}. \tag{1c}$$

Instead of $z \in \mathbb{Z}^{n_z}$ we may have $z \in \{0,1\}^{n_z}$, i.e., z is further restricted to be binary. The objective f and constraints $c = (c_E, c_I)$ are assumed to be twice continuously differentiable.

As discussed in the introduction, many applications do not satisfy the smoothness assumption. While jump discontinuities are properly handled by mixed-integer techniques involving artificial discrete variables and additional (big-M) constraints, this approach typically yields unnecessarily hard MINLPs when applied to nonsmooth but continuous functions. Therefore we consider the refined problem class C^0-MINLP defined by

$$\min_{x,z} \quad f(x,z) \tag{2a}$$

$$\text{s.t.} \quad c_E(x,z) = 0, \quad c_I(x,z) \geq 0, \tag{2b}$$

$$c_E^d(x,z) = 0, \quad c_I^d(x,z) \geq 0, \tag{2c}$$

$$x \in \mathbb{R}^{n_x}, \quad z \in \mathbb{Z}^{n_z}. \tag{2d}$$

Here we split the constraints into smooth and nonsmooth ones: $c = (c_E, c_I) \in C^2$, $c^d = (c_E^d, c_I^d) \in C^0$ and piecewise C^2. For the objective we still assume $f \in C^2$ without loss of generality: nonsmooth terms can always be moved to the constraints c^d. Problem (2) will be the basis of the following sequence of reformulations.

2.2. **From C^0-MINLP to C^0-MPEC: Complementarity Constraints.** The primary difficulties in the C^0-MINLP (2) are the discrete variables z and the nonsmooth constraints c^d. In this section we reformulate (2) with continuous variables and additional (smooth) constraints to obtain an equivalent problem without discrete variables. The original nonsmooth constraints c^d will be kept in this step.

Typical reformulation approaches [23, 40] make use of so called *NCP-functions* (see [42] for an overview). In particular, the *Fischer–Burmeister* function is used to restrict a continuous variable $x \in \mathbb{R}$ to $B := \{0,1\}$ or to $\tilde{B} := \{0\} \cup [1,\infty)$. However, since NCP-functions are nonsmooth we prefer an alternative approach that works directly with complementarity constraints [2]. Ultimately it is based on the trivial fact that

$$x\,(x-1) = 0 \iff x \in B. \tag{3}$$

Since this formulation is very ill-behaved numerically, a lifted version with an additional continuous variable y is usually preferred, yielding the standard MPEC formulation

$$xy = 0, \quad x, y \geq 0. \tag{4}$$

Here the cases $x = 0$, $y > 0$ and $x > 0$, $y = 0$ correspond to $x = 0$ and $x = 1$, respectively, but the improved numerical behavior comes at the price of an undecided third state, $x = y = 0$.

Fortunately, many applications with discrete alternatives share a property that can be exploited in a more useful way: the alternatives can be represented as subsets of a space of *continuous* variables. Formally, let A be some model aspect with a finite set of states $A_1, \ldots, A_a$ that correspond to constraint sets for some vector $x_A \in \mathbb{R}^{n_A}$,

$$c_{E,A_i}(x_A) = 0, \quad c_{I,A_i}(x_A) \geq 0, \quad i = 1, \ldots, a. \tag{5}$$

Then aspect A has a generic MINLP formulation as part of (1) or (2), using binary variables

$$z_A = (z_{A_i})_{i=1}^a \in \{0,1\}^a \tag{6}$$

together with big-M and SOS-1 constraints,

$$M_{E,A_i}(1 - z_{A_i}) - c_{E,A_i}(x_A) \geq 0, \quad i = 1, \dots, a, \tag{7a}$$

$$M_{E,A_i}(1 - z_{A_i}) + c_{E,A_i}(x_A) \geq 0, \quad i = 1, \dots, a, \tag{7b}$$

$$M_{I,A_i}(1 - z_{A_i}) + c_{I,A_i}(x_A) \geq 0, \quad i = 1, \dots, a, \tag{7c}$$

$$\sum_{i=1}^{a} z_{A_i} = 1, \quad z_{A_i} \in \{0, 1\}, \quad i = 1, \dots, a. \tag{7d}$$

An equivalent general disjunctive programming formulation [19, 33] is given by

$$\bigvee_{i=1}^{a} \begin{pmatrix} z_{A_i} = 1 \\ c_{E,A_i}(x_A) = 0 \\ c_{I,A_i}(x_A) \geq 0 \end{pmatrix}. \tag{8}$$

For the following we need the key concept of *non-disjunctive* states: states whose constraint sets *overlap*. The formal definition involves characteristic functions.

Definition 1. *Let A be a model aspect with states A_i, $i = 1, \dots, a$, represented by variables and constraints as in* (5) *and* (6)*.*

(1) *A function $\chi_{A_i} : \mathbb{R}^{n_A} \to \mathbb{R}$ is called a* characteristic function *of state A_i if*

$$\begin{aligned} &\chi_{A_i}(x) = 0 \quad \text{if } c_{E,A_i}(x) = 0 \text{ and } c_{I,A_i}(x) \geq 0, \\ &\chi_{A_i}(x) > 0 \quad \text{else.} \end{aligned} \tag{9}$$

(2) *Two states A_i and A_j are called* non-disjunctive *if there exists $x \in \mathbb{R}^{n_A}$ such that*

$$\chi_{A_i}(x) = \chi_{A_j}(x) = 0. \tag{10}$$

In what follows we only consider C^0-MINLP models where all discrete aspects have *two* non-disjunctive states and refer to this class as *2-state-C^0-MINLP*. As a direct consequence of the above definition, we can then state the following Lemma.

Lemma 1. *Let A_1, A_2 be non-disjunctive states of a model aspect A that is modeled with variables (x_A, z_A) and constraint sets $c_{E,A_i}, c_{I,A_i}, i = 1, 2$. Let χ_{A_i} denote corresponding characteristic functions. Then the MINLP model of A can be equivalently replaced by the MPEC model*

$$\chi_{A_1}(x_A)\chi_{A_2}(x_A) = 0. \tag{11}$$

Proof. Let x_A^* be a solution of the reformulated MPEC model and let $\chi_{A_1}(x_A^*) = 0$. Then it follows with (9) that $c_{E,A_1}(x_A^*) = 0$ and $c_{I,A_1}(x_A^*) \geq 0$. By setting $z_{A_1} = 1$ and $z_{A_2} = 0$ we have constructed a feasible solution to (7). The case $\chi_{A_2}(x_A^*) = 0$ and the reverse direction are analogous. □

An equivalent reformulation of the general 2-state-C^0-MINLP model as an C^0-MPEC model according to Lemma 1 can now be written

$$\min_{x \in \mathbb{R}^{n_x}} \quad f(x) \tag{12a}$$

$$\text{s.t.} \quad c_E(x) = 0, \quad c_I(x) \geq 0, \tag{12b}$$

$$c_E^d(x) = 0, \quad c_I^d(x) \geq 0, \tag{12c}$$

$$\phi_i(x)\psi_i(x) = 0, \quad i = 1, \dots, p, \tag{12d}$$

$$\phi_i(x), \psi_i(x) \geq 0, \quad i = 1, \dots, p. \tag{12e}$$

Here $\phi_i, \psi_i : \mathbb{R}^{n_x} \to \mathbb{R}$ are the complementarity constraint pairings constructed from characteristic functions and p is the number of 2-state model aspects A_i,

$$\phi_i = \chi_{A_{i,1}}, \quad \psi_i = \chi_{A_{i,2}}, \quad i = 1, \dots, p. \tag{13}$$

Note that the "undecided third state" of (4) does not pose a problem here: the crucial property of non-disjunctive states is that their continuous variables can be identical.

2.3. **From C^0-MPEC to C^2-MPEC: Smoothing.** This section addresses the remaining major difficulty of (12): the nonsmooth constraints c^d. Quite often the nonsmoothness arises from the absolute value function or from functions that can be expressed in terms of it, like $\min(x, y)$ and $\max(x, y)$.

As already mentioned, first-order discontinuities should not be modeled by artificial discrete variables in the current context; the constraints c^d should rather be replaced with sufficiently smooth approximations c^s. We use approximations that depend on a smoothing parameter, $c^s(x; \tau) \approx c^d(x)$ with $\tau > 0$, satisfying at least a *pointwise approximation property:*

$$\forall x\colon \lim_{\tau \to 0} c^s(x; \tau) = c^d(x). \tag{14}$$

This provides control over the approximation quality by adjusting the smoothing parameter τ. For the nonsmooth functions mentioned above we actually have uniformly convergent approximations,

$$|x| \approx v(x; \tau) = \sqrt{x^2 + \tau}, \tag{15}$$

$$\min(x, y) \approx y - \tfrac{1}{2}(v(x - y; \tau) - (x - y)), \tag{16}$$

$$\max(x, y) \approx y + \tfrac{1}{2}(v(x - y; \tau) + (x - y)). \tag{17}$$

Of course, the value of τ should not be chosen too small because this would introduce numerical instabilities and ill-conditioning. Unfortunately there is no general rule for choosing parameters like τ; they have to be tuned separately for each model. Moreover, one often needs problem specific smoothing techniques. We will give an example for an application in gas network optimization in Sect. 4.

The smoothed C^0-MPEC model (12) will be referred to as C^2-MPEC; it reads

$$\min_{x \in \mathbb{R}^{n_x}} \quad f(x) \tag{18a}$$

$$\text{s.t.} \quad c_E(x) = 0, \quad c_I(x) \geq 0, \tag{18b}$$

$$c_E^s(x; \tau) = 0, \quad c_I^s(x; \tau) \geq 0, \tag{18c}$$

$$\phi_i(x)\psi_i(x) = 0, \quad i = 1, \ldots, p, \tag{18d}$$

$$\phi_i(x), \psi_i(x) \geq 0, \quad i = 1, \ldots, p. \tag{18e}$$

2.4. **From C^2-MPEC to C^2-NLP: Regularization.** We now have to solve the smooth MPEC model (18). It is well-known that standard NLP constraint qualifications, such as the Mangasarian–Fromowitz constraint qualification (MFCQ) or the linear independence constraint qualification (LICQ), do not hold at *any* feasible point of the MPEC. To apply standard NLP algorithms without losing their good convergence properties, various regularization schemes have therefore been developed. There are basically three groups of existing schemes:

(1) relaxation schemes,
(2) penalization schemes,
(3) smoothing schemes.

The common idea of all regularization schemes is to replace the MPEC constraints (18d) and (18e) with NLP constraints that depend on a regularization parameter μ. Then one solves a sequence NLP(μ_k) with $\mu_k \to 0$ whose solutions converge to an MPEC solution under suitable assumptions.

One of the first regularization schemes is the relaxation scheme of Scholtes [39], which relaxes the complementarity constraints (18d) to

$$\phi_i(x)\psi_i(x) \le \mu, \quad i = 1, \dots, p. \tag{19}$$

This yields a regularized problem NLP(μ) with feasible set $F(\mu)$ such that the original MPEC-feasible set is $F(0)$ and $F(\mu_0) \subset F(\mu)$ for all $\mu > \mu_0 \ge 0$. If NLP(μ) is to be solved by an interior point method, the relaxation (19) has the drawback that it lacks strict interior points in the limit. DeMiguel et al. address this problem in [12] by additionally relaxing the nonnegativity constraints (18e) to

$$\phi_i(x), \psi_i(x) \ge -\theta, \quad i = 1, \dots, p. \tag{20}$$

They propose a method that drives either θ or μ to zero in the limit but not both.

Penalization schemes remove the complementarity constraints completely from the constraints set and introduce a weighted penalty term in the objective instead. Thus (18d) is dropped from (18), and (18a) is replaced with

$$f(x) + \frac{1}{\mu}\Pi(\phi(x), \psi(x)). \tag{21}$$

Theoretical results for a general class of penalty functions Π can be found in [21]. In particular, these results include the concrete instance

$$\Pi(\phi(x), \psi(x)) = \sum_{i=1}^{p} \phi_i(x)\psi_i(x) \tag{22}$$

that is most frequently used in practice.

Other regularization approaches use nonsmooth reformulations of the complementarity constraints $\phi_i(x)\psi_i(x) = 0$, such as

$$\min\{\phi_i(x), \psi_i(x)\} = 0, \tag{23}$$

and employ nonsmooth optimization techniques to solve the resulting problem.

Finally there are smoothing techniques using modified NCP-functions like the perturbed Fischer–Burmeister function (first proposed in [16]),

$$\zeta(\phi, \psi; \tau) = \phi + \psi - \sqrt{\phi^2 + \psi^2 + \tau} = 0. \tag{24}$$

In what follow we concentrate on relaxation and penalization schemes since these performed best on the application problem presented below. Both approaches of regularizing the MPEC model (18) yield a C^2-NLP that satisfies a standard constraint qualification and can be written in the general form

$$\min_{x \in \mathbb{R}^{n_x}} \quad f(x) + g(x; \mu) \tag{25a}$$

$$\text{s.t.} \quad c_E(x) = 0, \quad c_I(x) \ge 0, \tag{25b}$$

$$c_E^s(x; \tau) = 0, \quad c_I^s(x; \tau) \ge 0, \tag{25c}$$

$$c_E^r(x; \mu) = 0, \quad c_I^r(x; \mu) \ge 0, \tag{25d}$$

$$\phi_i(x), \psi_i(x) \ge -\theta, \quad i = 1, \dots, p. \tag{25e}$$

For the relaxation scheme (19), possibly extended by (20), c_E^r vanishes and we have $g = 0$, $\theta \ge 0$, $c_I^r(x; \mu) = (\phi_i(x)\psi_i(x) - \mu)_{i=1}^p$. The penalization scheme (21) has $g(x; \mu) = \frac{1}{\mu}\Pi(\phi(x), \psi(x))$, $\theta = 0$, and $c_E^r(x; \mu), c_I^r(x; \mu)$ are not needed. Defining $\bar{c}_E := (c_E, c_E^s, c_E^r)$, $\bar{c}_I := (c_I, c_I^s, c_I^r, \phi, \psi)$ and $\bar{f} := f + g$ now yields the (parameterized) standard NLP formulation

$$\min_{x \in \mathbb{R}^{n_x}} \quad \bar{f}(x; \mu) \quad \text{s.t.} \quad \bar{c}_E(x; \tau, \mu) = 0, \quad \bar{c}_I(x; \tau, \mu) \ge 0. \tag{26}$$

In the context of our primal heuristic a final remark is in order. Since we are primarily interested in finding feasible solutions of the original MINLP quickly rather than solving the approximating MPEC with high accuracy, we do not attempt to solve an entire sequence NLP(μ_k) with $\mu_k \to 0$. Instead we take a more aggressive approach and try to solve only a single instance NLP(μ) where the parameter $\mu > 0$ is fixed at a carefully chosen problem-specific value.

3. Relations of the Model Classes

This section briefly highlights basic theoretical relations between the models presented in the last section. We focus on feasible points and not on optimality and stationarity because our main topic here is the primal heuristic.

The first result gives a relation between feasible points of the original model 2-state-C^0-MINLP and its first reformulation C^0-MPEC.

Lemma 2. *Let P be a 2-state-C^0-MINLP in the form* (2) *and let Q be a reformulation as C^0-MPEC in the form* (12)*. Then for every Q-feasible point x_Q^* there exists a P-feasible point (x_P^*, z_P^*). Conversely, if there is no Q-feasible point, P is also infeasible. Thus, there is a one-to-one correspondence of feasible points between 2-state-C^0-MINLP models and their C^0-MPEC reformulations.*

Proof. The claim follows directly from Def. 1 and Lemma 1. □

To obtain binary decisions z_P^* from x_Q^* the complementarity constraints of C^0-MPEC are evaluated at the solution x_Q^*. Then the states are determined according to Def. 1 and (4). If a complementarity constraint is biactive, i.e. both characteristic functions evaluate to zero, the discrete state is arbitrary.

The second transition step, from C^0-MPEC to C^2-MPEC, has a genuinely heuristic flavor: pointwise convergence of the smoothing functions for $\tau \to 0$ does not necessarily imply any useful convergence of C^2-MPEC-feasible sets to C^0-MPEC-feasible sets. Moreover, the smoothing parameter τ is not driven to zero but has to be fixed at some positive value. For complex application problems one will typically try problem-specific smoothings and parameter tuning anyway, and the smoothing error is often smaller than other model inaccuracies. If the overall heuristic still fails, one simply has to rely on computationally more expensive rigorous methods.

The properties of the transition from C^2-MPEC to C^2-NLP depend on the regularization scheme being used. We will discuss penalization and relaxation in the following. To this end, some basic MPEC theory is needed [21, 36, 39] and we consider the standard MPEC formulation:

$$\min_x \quad f(x) \tag{27a}$$

$$\text{s.t.} \quad c_E(x) = 0, \quad c_I(x) \geq 0, \tag{27b}$$

$$\phi_i(x)\psi_i(x) = 0, \quad i = 1, \dots, p, \tag{27c}$$

$$\phi_i(x), \psi_i(x) \geq 0, \quad i = 1, \dots, p. \tag{27d}$$

Definition 2. *We say that the* MPEC linear independence constraint qualification *holds for an MPEC-feasible point x if and only if the standard LICQ holds for the entire constraints system with the exception of complementarity constraints* (27c)*:*

$$c_E(x) = 0, \quad c_I(x) \geq 0, \quad \phi_i(x) \geq 0, \quad \psi_i(x) \geq 0. \tag{28}$$

For the following, we define several sets of active indices,

$$\mathcal{A}_c(x) = \{i \in I \colon c_i(x) = 0\}, \tag{29a}$$

$$\mathcal{A}_\phi(x) = \{i \in \{1, \dots, p\} \colon \phi_i(x) = 0\}, \tag{29b}$$

$$\mathcal{A}_\psi(x) = \{i \in \{1, \dots, p\} \colon \psi_i(x) = 0\}. \tag{29c}$$

The next theorem extends standard first-order KKT conditions for NLP to MPEC. A proof can be found in [36].

Theorem 1. *Let x^* be a minimizer of* (27) *and let MPEC-LICQ hold at x^*. Then there exist dual variables $\lambda_E^* \in \mathbb{R}^{|E|}, \lambda_I^* \in \mathbb{R}^{|I|}$ and $\gamma_\phi^*, \gamma_\psi^* \in \mathbb{R}^p$ so that*

$$\nabla f^* - \nabla c_E^{*\,T} \lambda_E^* - \nabla c_I^{*\,T} \lambda_I^* - \nabla \phi^{*\,T} \gamma_\phi^* - \nabla \psi^{*\,T} \gamma_\psi^* = 0, \tag{30a}$$

$$c_E^* = 0, \quad c_I^* \geq 0, \quad \phi_i^* \geq 0, \quad \psi_i^* \geq 0, \tag{30b}$$

$$\phi_i^* = 0 \quad \textit{or} \quad \psi_i^* = 0, \quad i = 1, \dots, p, \tag{30c}$$

$$c_i^* \lambda_{I_i}^* = 0, \quad i \in I, \quad \phi_i^* \gamma_{\phi_i}^* = 0 \quad \textit{and} \quad \psi_i^* \gamma_{\psi_i}^* = 0, \quad i = 1, \dots, p, \tag{30d}$$

$$\lambda_{I_i}^* \geq 0, \quad i = 1, \dots, p, \tag{30e}$$

$$\gamma_{\phi_i}^* \geq 0, \quad \gamma_{\psi_i}^* \geq 0, \quad i \in \mathcal{A}_\phi^* \cap \mathcal{A}_\psi^*. \tag{30f}$$

The superindex * indicates function evaluation at x^*. Condition (30a) corresponds to standard dual feasibility, (30b) and (30c) cover primal feasibility of (27), and (30d) is the standard complementarity of inequalities and their multipliers. Finally (30e) and (30f) correspond to nonnegativity of the multipliers of inequality constraints. Note that the last condition is only required for so called *corner pairings* [24], i.e. complementarity pairings satisfying $\phi_i(x^*) = \psi_i(x^*) = 0$.

Theorem 1 is the basis of MPEC stationarity concepts [36]:

Definition 3. *Let x^* be MPEC-feasible, i.e.* (30b) *and* (30c) *hold, and assume that there exist dual variables $\lambda_E^*, \lambda_I^*, \gamma_\phi^*, \gamma_\psi^*$ satisfying* (30a)–(30e). *Then x^* is called*

(1) strongly stationary *if in addition* (30f) *holds,*
(2) C-stationary *if in addition $\gamma_{\phi_i}^* \gamma_{\psi_i}^* \geq 0$.*

After these preparations we can discuss the main convergence results of MPEC regularization schemes for $\mu \to 0$.

For the *relaxation* scheme (19) it is shown in [39] that the sequence of stationary points of the relaxed MPECs converge to C-stationary points if the MPEC-LICQ condition holds in the limit. In [20] it is shown that this scheme in fact converges to C-stationary points under the milder assumption of MPEC-MFCQ.

Theoretical results for the *penalization* scheme can be found in [21]. For the penalty objective (21) in particular, convergence to C-stationary points is obtained if MPEC-LICQ holds in the limit.

Stronger convergence results can be proved under stronger assumptions such as the *weak second order necessary condition* or *upper level strict complementarity.* In particular, both schemes deliver MPEC-feasible accumulation points in these cases, which is sufficient for our purpose of constructing a primal heuristic.

4. Application: Gas Network Optimization

The techniques just presented are now applied as a primal heuristic for a planning problem in gas transport. We model the gas network as a directed graph $G = (\mathbb{V}, \mathbb{A})$. The node set $\mathbb{V}$ consists of entries $\mathbb{V}_+$, exits $\mathbb{V}_-$ and junctions $\mathbb{V}_0$, and the arc set $\mathbb{A}$ consists of pipes $\mathbb{A}_{\text{pi}}$, resistors $\mathbb{A}_{\text{re}}$, valves $\mathbb{A}_{\text{va}}$, control valves $\mathbb{A}_{\text{cv}}$ and compressor groups $\mathbb{A}_{\text{cg}}$. A *nomination* defines the amounts of all flows that are supplied or discharged by entry and exit customers. In addition, a nomination determines bounds for the gas pressure as well as specific values of certain gas quality parameters.

We address a problem referred to as *validation of nominations (NoVa)*, which is to decide whether a given nomination can be served by some feasible stationary network operation. Because of discrete decisions at active elements (status of valves, control valves and compressor groups) and the mostly nonlinear and partly nonsmooth physical models of network elements this task leads to a nonsmooth MINLP *feasibility* problem. If successful, our heuristic will thus deliver solutions directly.

Various related problems of the gas transport industry have been addressed in the literature. In [8, 30, 45] one finds first (often heuristic) attempts at mixed-integer nonlinear optimization, addressing single compressor groups under fixed operating conditions. Later research incorporates more detailed physical models [5, 46], and more recently also additional discrete aspects and network elements [9]. For large-scale real-world network models MIP-driven approaches have been developed in [27, 28, 32, 11] together with problem-specific heuristic enhancements [26]. NLP-oriented investigations include [4, 34, 35] for stationary optimization, [14, 15, 41] for the transient case and [13] where specific MIP and NLP approaches are compared. The work presented here results from the large industry project ForNe that aims at developing mathematical methods for all kinds of network planning problems. Publications in preparation related to the ForNe project include [17, 22, 31, 38, 37]. ForNe is funded by Open Grid Europe GmbH. The scientific project partners are Friedrich-Alexander Universität Erlangen-Nürnberg, Konrad Zuse Zentrum für Informationstechnik Berlin (ZIB), Universität Duisburg-Essen, Weierstraß Institut für Angewandte Analysis und Stochastik (WIAS), Humboldt Universität zu Berlin, Technische Universität Darmstadt and Leibniz Universität Hannover.

4.1. **Model.** In this section the C^0-MINLP model of the NoVa problem is presented along with the smoothed MPEC reformulation C^2-MPEC. As we wish to obtain results quickly, we use a reasonably simplified model rather than the highly detailed model developed in [37]. For instance, the model presented is isothermal, i.e. all temperatures are considered constant. If an approximate solution with correct discrete decisions is found, the accuracy of the continuous variables can still be increased by an extra optimization run with a refined model. On the other hand, real-world instances may contain discrete aspects of global nature like interdependencies of decisions that can currently not be handled by the model.

We introduce every network element model separately and show that the discrete decisions lead to a 2-state-C^0-MINLP in the form (2). The notation is similar to the previous sections except that subindices now refer to network elements or sets thereof. For instance, x_i denotes the variables of the component model of node i, and $\mathbf{c}_{\mathbb{A}_{\mathrm{pi}}}$ denotes the constraints of the component models of all pipes.

4.1.1. *Nodes.* Every node $i \in \mathbb{V}$ has a gas pressure variable with simple bounds, $p_i \in [p_i^-, p_i^+]$. The flows at node i satisfy a mass balance equation

$$0 = c_i^{\mathrm{flow}}(x) = \sum_{a \in \delta_i^-} q_a - \sum_{a \in \delta_i^+} q_a + d_i, \tag{31}$$

where d_i is the externally supplied flow:

$$d_i \geq 0 \text{ for } i \in \mathbb{V}_+, \quad d_i = 0 \text{ for } i \in \mathbb{V}_0, \quad d_i \leq 0 \text{ for } i \in \mathbb{V}_-. \tag{32}$$

The complete (smooth) node model reads

$$0 = \mathbf{c}_i(x) = c_i^{\mathrm{flow}}(x), \quad x_i = p_i. \tag{33}$$

4.1.2. *Pipes.* Gas dynamics in pipes $a = ij \in \mathbb{A}_{\mathrm{pi}}$ are properly described by the Euler equations for compressible fluids: a PDE system involving mass, momentum and energy balances. Consider a cylindrical pipe with diameter D, cross-sectional area A, roughness k and slope $s \in [-1, +1]$ (the tangent of the inclination angle). For the isothermal and stationary case considered here, the mass balance (continuity equation) yields constant mass flow q along the pipe, the energy equation is not needed, and we are left with the *stationary momentum equation*

$$\frac{\partial p}{\partial x} + \frac{q^2}{A^2} \partial_x \frac{1}{\rho} + g\rho s + \lambda(q) \frac{|q|q}{2A^2 D\rho} = 0. \tag{34}$$

Here g denotes gravitational acceleration, and the *friction coefficient* $\lambda(q)$ is given in terms of the *Reynolds number* $\mathrm{Re}(q)$: for laminar flow by the law of Hagen–Poiseuille,

$$\lambda^{\mathrm{HP}}(q) = \frac{64}{\mathrm{Re}(q)}, \quad \mathrm{Re}(q) = \frac{D}{A\eta}|q|, \quad |q| \le q_{\mathrm{crit}} \tag{35}$$

and for turbulent flow by the empirical implicit model of Prandtl–Colebrook,

$$\frac{1}{\sqrt{\lambda^{\mathrm{PC}}(q)}} = -2\log_{10}\left(\frac{2.51}{\mathrm{Re}(q)\sqrt{\lambda^{\mathrm{PC}}(q)}} + \frac{k}{3.71D}\right), \quad |q| > q_{\mathrm{crit}}. \tag{36}$$

The state quantities pressure p, density ρ and temperature T in (34) are coupled by an *equation of state*; we use the thermodynamical standard equation

$$\rho = \rho(p,T) = \frac{p}{R_s z(p,T)T}, \tag{37}$$

where R_s is the specific gas constant. The *compressibility factor* $z(p,T)$ is given by an empirical model; here we use a formula of the American Gas Association (AGA),

$$z(p,T) = 1 + 0.257\frac{p}{p_c} - 0.533\frac{p/p_c}{T/T_c}, \tag{38}$$

where p_c and T_c denote the pseudocritical gas pressure and temperature.

The ODE (34) essentially yields the pressure loss along pipe a for which various approximation formulas exist. We use a quadratic approximation of Weymouth type,

$$0 = c_a^{\text{p-loss}}(x) = p_j^2 - \left(p_i^2 - \Lambda_a z_{a,m}\lambda_a q_a|q_a|\frac{e^{S_a}-1}{S_a}\right)e^{-S_a}, \tag{39}$$

$$0 = c_a^{\text{slope}}(x) = S_a z_{a,m} - \frac{2L_a g}{R_s T}s_a. \tag{40}$$

The coefficient Λ_a and inclination variable S_a depend on pipe data like length L_a and slope s_a, and on an approximate mean value $z_{a,m}$ of the compressibility factor,

$$0 = c_a^{\text{z-mean}}(x) = z_{a,m} - z(p_{a,m},T), \tag{41}$$

$$0 = c_a^{\text{p-mean}}(x) = p_{a,m} - \frac{2}{3}\left(p_i + p_j - \frac{p_j p_j}{p_i + p_j}\right). \tag{42}$$

The friction variable λ_a in $c_a^{\text{p-loss}}$ (39) has to satisfy the nonsmooth constraint

$$0 = c_a^{\mathrm{HPPC}}(x) = \lambda_a - \begin{cases}\lambda^{\mathrm{HP}}(q_a), & q_a \le q_{\mathrm{crit}},\\ \lambda^{\mathrm{PC}}(q_a), & q_a > q_{\mathrm{crit}}.\end{cases} \tag{43}$$

The complete pipe model then reads

$$0 = \mathbf{c}_a(x) = \begin{pmatrix} c_a^{\text{p-loss}}(x)\\ c_a^{\text{p-mean}}(x)\\ c_a^{\text{z-mean}}(x)\\ c_a^{\mathrm{HPPC}}(x)\\ c_a^{\text{slope}}(x)\end{pmatrix}, \quad x_a = \begin{pmatrix} q_a\\ z_{a,m}\\ p_{a,m}\\ \lambda_a\\ S_a\end{pmatrix}. \tag{44}$$

4.1.3. *Pipe Model Reformulation: Smoothing.* The pipe model is discontinuous at $q_a = q_{\text{crit}}$ due to c_a^{HPPC}, and second-order discontinuous at $q_a = 0$ due to the term $q_a|q_a|$ in $c_a^{\text{p-loss}}$. We replace the term $\lambda_a q_a|q_a|$ in (39) and constraint (43) by a new variable ϕ_a defined by a smooth constraint,

$$0 = c_a^{\text{p-loss-s}}(x) = p_j^2 - \left(p_i^2 - \Lambda_a z_{a,m}\phi_a \frac{e^{S_a} - 1}{S_a}\right) e^{-S_a}, \tag{45}$$

$$0 = c_a^{\text{HPPC-s}}(x) = \phi_a - r_a q_a \left(\sqrt{q_a^2 + e_a^2} + b_a + \frac{c_a}{\sqrt{q_a^2 + d_a^2}}\right). \tag{46}$$

This provides an asymptotically correct second-order approximation of $\lambda_a q_a|q_a|$ if the parameters r_a, b_a, c_a, d_a, e_a are suitably chosen [6, 37]. In summary, we obtain the smooth pipe model

$$0 = \mathbf{c}_a^{\text{smooth}}(x) = \begin{pmatrix} c_a^{\text{p-loss-s}}(x) \\ c_a^{\text{p-mean}}(x) \\ c_a^{\text{z-mean}}(x) \\ c_a^{\text{HPPC-s}}(x) \\ c_a^{\text{slope}}(x) \end{pmatrix}, \quad x_a^{\text{smooth}} = \begin{pmatrix} q_a \\ z_{a,m} \\ p_{a,m} \\ \phi_a \\ S_a \end{pmatrix}. \tag{47}$$

4.1.4. *Resistors.* Resistors $a = ij \in \mathbb{A}_{\text{rs}}$ are fictitious network elements modeling the approximate pressure loss across gadgets, partly closed valves, filters, etc. The pressure loss has the same sign as the mass flow and is either assumed to be (piecewise) constant,

$$0 = c_a^{\text{p-loss-lin}}(x) = p_i - p_j - \xi_a \operatorname{sign}(q_a), \tag{48}$$

or (piecewise) quadratic according to the law of Darcy–Weisbach,

$$0 = c_a^{\text{p-loss-nl}}(x) = p_i - p_j - \frac{8\zeta_a}{\pi^2 D_a^4} \frac{q_a|q_a|}{\rho_{a,k}}. \tag{49}$$

Here ζ_a is the resistance coefficient and $\rho_{a,k}$ is the inflow gas density according to the equation of state (37),

$$0 = c_a^{\text{dens-in}}(x) = \rho_{a,k} - \rho(p_k, T) \quad \text{with} \quad k := \begin{cases} i, & q_a \geq 0, \\ j, & q_a < 0. \end{cases} \tag{50}$$

The compressibility factor z has to be evaluated at the inflow node as well,

$$0 = c_a^{\text{z-in}}(x) = z_{a,k} - z(p_k, T). \tag{51}$$

In summary, the piecewise constant resistor model ($a \in \mathbb{A}_{\text{lin-rs}}$) reads

$$0 = \mathbf{c}_a(x) = c_a^{\text{p-loss-lin}}(x), \quad x_a = q_a, \tag{52}$$

and the piecewise quadratic resistor model ($a \in \mathbb{A}_{\text{nonlin-rs}}$) reads

$$0 = \mathbf{c}_a(x) = \begin{pmatrix} c_a^{\text{p-loss-nl}}(x) \\ c_a^{\text{dens-in}}(x) \\ c_a^{\text{z-in}}(x) \end{pmatrix}, \quad x_a = \begin{pmatrix} q_a \\ z_{a,k} \\ \rho_{a,k} \end{pmatrix}. \tag{53}$$

4.1.5. *Resistor Model Reformulation: Smoothing.* The resistor models (52) and (53) are nonsmooth because of three reasons:

(1) the discontinuous sign function in (48),
(2) the second-order discontinuous term $|q_a|q_a$ in (49),
(3) the direction dependence of the inflow gas density $\rho_{a,k}$ in (49).

Note that items 1 and 3 violate the assumptions made so far. However, resistors play a minor role in the NoVa problem and it suffices to include a coarse approximation in the model, so we just proceed with a problem-specific smoothing.

In the piecewise constant resistor model, we use the identity $\operatorname{sign}(x) = x/|x|$ together with the standard smoothing of $|x|$. For $a \in \mathbb{A}_{\text{lin-rs}}$ this yields

$$0 = \mathbf{c}_a^{\text{smooth}}(x) = c_a^{\text{p-loss-lin-s}}(x) = p_i - p_j - \xi_a \frac{q_a}{\sqrt{q_a^2 + \tau}}. \tag{54}$$

The same approximation of the absolute value function is applied to the piecewise quadratic resistor model (49):

$$0 = c_a^{\text{p-loss-nl-s}}(x) = p_i - p_j - \frac{8\zeta_a}{\pi^2 D_a^4} \frac{q_a \sqrt{q_a^2 + \tau}}{\rho_{a,k}}. \tag{55}$$

Finally, the direction dependence of the inflow gas density $\rho_{a,k}$ is addressed by using the mean density

$$0 = c_a^{\text{dens-mean}}(x) = \rho_{a,m} - \frac{1}{2}\left(\rho_{a,\text{in}} + \rho_{a,\text{out}}\right). \tag{56}$$

As a consequence, we need to evaluate the equation of state and the compressibility factor at both nodes i and j,

$$c_a^{\text{dens-in}} = \rho_{a,\text{in}} - \rho(p_i, T), \qquad c_a^{\text{dens-out}} = \rho_{a,\text{out}} - \rho(p_j, T), \tag{57}$$

$$c_a^{\text{z-in}} = z_{a,\text{in}} - z(p_i, T), \qquad c_a^{\text{z-out}} = z_{a,\text{out}} - z(p_j, T). \tag{58}$$

This yields for $a \in \mathbb{A}_{\text{nonlin-rs}}$ the smoothed model

$$0 = \mathbf{c}_a^{\text{smooth}}(x) = \begin{pmatrix} c_a^{\text{p-loss-nl-s}}(x) \\ c_a^{\text{dens-in}}(x) \\ c_a^{\text{dens-out}}(x) \\ c_a^{\text{dens-mean}}(x) \\ c_a^{\text{z-in}}(x) \\ c_a^{\text{z-out}}(x) \end{pmatrix}, \quad x_a^{\text{smooth}} = \begin{pmatrix} q_a \\ z_{a,\text{in}} \\ z_{a,\text{out}} \\ \rho_{a,\text{in}} \\ \rho_{a,\text{out}} \\ \rho_{a,m} \end{pmatrix}. \tag{59}$$

4.1.6. *Valves.* Valves $a = ij \in \mathbb{A}_{\text{vl}}$ have two discrete states: *open* and *closed*. Across open valves, the pressures are identical and the flow is arbitrary within its technical bounds,

$$p_j = p_i, \quad q_a \in [q_a^-, q_a^+]. \tag{60}$$

Closed valves block the gas flow and the pressures are arbitrary within their bounds,

$$q_a = 0, \quad p_i \in [p_i^-, p_i^+], \quad p_j \in [p_j^-, p_j^+]. \tag{61}$$

This behavior can be modeled with one binary variable $z_a \in \{0, 1\}$ together with big-M constraints:

$$0 \le c_a^{\text{flow-lb}}(x, z) = q_a - z_a q_a^-, \tag{62}$$

$$0 \le c_a^{\text{flow-ub}}(x, z) = -q_a + z_a q_a^+, \tag{63}$$

$$0 \le c_a^{\text{p-coupl-1}}(x, z) = M_{a,1}(1 - z_a) - p_j + p_i, \tag{64}$$

$$0 \le c_a^{\text{p-coupl-2}}(x, z) = M_{a,2}(1 - z_a) - p_i + p_j. \tag{65}$$

The resulting valve MINLP model reads

$$0 \le \mathbf{c}_a(x, z) = \begin{pmatrix} c_a^{\text{flow-lb}}(x, z) \\ c_a^{\text{flow-ub}}(x, z) \\ c_a^{\text{p-coupl-1}}(x, z) \\ c_a^{\text{p-coupl-2}}(x, z) \end{pmatrix}, \quad x_a = q_a. \tag{66}$$

4.1.7. *Valve Model Reformulation: Complementarity Constraints.* It is easily seen that (66) directly fits into the concept of 2-state model aspects. Here, the model aspect valve has the two states $A_1 =$ *open* and $A_2 =$ *closed*. They are non-disjunctive if $0 \in [q_a^-, q_a^+]$ and $[p_i^-, p_i^+] \cap [p_j^-, p_j^+] \neq \emptyset$. The characteristic functions are

$$\chi_a^{\text{open}}(x) = p_j - p_i, \quad \chi_a^{\text{closed}}(x) = q_a. \tag{67}$$

According to Sect. 2, the 2-state-MINLP model (66) can be equivalently reformulated using a complementarity constraint:

$$0 = \mathbf{c}_a^{\text{mpec}}(x) = c_a^{\text{vl-state}}(x) = \chi_a^{\text{open}}(x)\chi_a^{\text{closed}}(x), \quad x_a^{\text{mpec}} = q_a. \tag{68}$$

It offers two advantages: no binary variables are required and the number of constraints reduces from four to one.

4.1.8. *Control Valves.* Control valves $a = ij \in \mathbb{A}_{\text{cv}}$ are used to decrease the gas pressure in a technically prescribed direction (which we define as the graph direction $i \to j$). They possess three discrete states: *active*, *bypass* and *closed*. An *active* control valve reduces the inflow pressure by a certain amount,

$$p_j = p_i - \Delta p_a, \quad \Delta p_a \in [\Delta p_a^-, \Delta p_a^+], \quad q_a \in [q_a^-, q_a^+] \cap \mathbb{R}_+. \tag{69}$$

A *closed* control valve acts like a closed regular valve, leading to the simple state model (61). A control valve in *bypass* mode acts like an open regular valve, with arbitrary flow direction and without decreasing the pressure, cf. (60). Our complete mixed-integer linear model is based on the variable vector $x_a = (q_a, \Delta p_a)^T$ and $z_a = (z_{1,a}, z_{2,a})^T$, where $z_{1,a}$ defines if the control valve is open ($z_{1,a} = 1$) or closed ($z_{1,a} = 0$) and $z_{2,a}$ defines if it is active ($z_{2,a} = 1$) or not ($z_{2,a} = 0$). In terms of the constraints

$$0 \le c_a^{\text{flow-lb-open}}(x, z) = q_a - z_{a,1} q_a^-, \tag{70a}$$

$$0 \le c_a^{\text{flow-ub-open}}(x, z) = -q_a + z_{a,1} q_a^+, \tag{70b}$$

$$0 \le c_a^{\text{flow-lb-active}}(x, z) = q_a - (1 - z_{a,2})\, q_a^-, \tag{70c}$$

$$0 \le c_a^{\text{p-coupl-1}}(x, z) = M_{a,1}(1 - z_{a,1}) + \Delta p_a^+ z_{a,2} - (p_i - p_j), \tag{70d}$$

$$0 \le c_a^{\text{p-coupl-2}}(x, z) = M_{a,2}(1 - z_{a,1}) - \Delta p_a^- z_{a,2} - (p_j - p_i), \tag{70e}$$

$$0 \le c_a^{\text{consistent-states}}(x, z) = z_{a,1} - z_{a,2}, \tag{70f}$$

the resulting mixed-integer model then becomes

$$0 \le \mathbf{c}_a(x, z) = \begin{pmatrix} c_a^{\text{flow-lb-open}}(x, z) \\ c_a^{\text{flow-ub-open}}(x, z) \\ c_a^{\text{flow-lb-active}}(x, z) \\ c_a^{\text{p-coupl-1}}(x, z) \\ c_a^{\text{p-coupl-2}}(x, z) \\ c_a^{\text{consistent-states}}(x, z) \end{pmatrix}, \quad x_a = \begin{pmatrix} q_a \\ \Delta p_a \end{pmatrix}, \quad z_a = \begin{pmatrix} z_{1,a} \\ z_{2,a} \end{pmatrix}. \tag{71}$$

4.1.9. *Control Valve Model Reformulation: Complementarity Constraints.* For our reformulation, we require $\Delta p_a^- = 0$. However, this appears to be a moderate restriction in practice: it holds in all cases we have encountered. With this, we can model control valves as a model aspect with two non-disjunctive states $A_1 =$ *open* and $A_2 =$*closed* and the characteristic functions

$$\chi_a^{\text{open}}(x) = p_j - p_i + \Delta p_a, \quad \chi_a^{\text{closed}}(x) = q_a. \tag{72}$$

The state *open* can then be distinguished in *active* or *bypass* depending on the value of Δp_a. Thus, we have the MPEC reformulation

$$0 = c_a^{\text{cv-state}}(x) = \chi_a^{\text{open}}(x)\chi_a^{\text{closed}}(x). \tag{73}$$

In addition, the restriction to nonnegative flows in the active state is modeled as

$$0 \le c_a^{\text{cv-active-flow}}(x) = \Delta p_a q_a. \tag{74}$$

The complete MPEC type control valve model now reads

$$\begin{aligned} &0 = \mathbf{c}_{a,E}^{\text{mpec}}(x) = c_a^{\text{cv-state}}(x), \quad 0 \le \mathbf{c}_{a,I}^{\text{mpec}}(x) = c_a^{\text{cv-active-flow}}(x), \\ &x_a = \begin{pmatrix} q_a \\ \Delta p_a \end{pmatrix}. \end{aligned} \tag{75}$$

4.1.10. *Compressor Groups.* Compressor groups $a = ij \in \mathbb{A}_{\text{cg}}$ typically consist of several compressor units of different types that can be combined in various configurations to increase the gas pressure; see [37] and the upcoming publications [17, 22, 31].

For our primal heuristic we use a substantially simplified model where compressor groups can work in the same states as control valves: *open*, *closed* and *active*. The only difference is the sign of the pressure control variable Δp_a in the characteristic function (72). Thus we have sign changes in (70d) and (70e), yielding adapted mixed-integer and MPEC formulations corresponding to (71) and (75), respectively.

4.2. **Model Summary.** In the preceding sections we have described components of gas transport networks, both as nonsmooth nonlinear mixed-integer models and, if necessary, as smooth MPEC reformulations. Now we combine the components into complete models.

The complete feasibility problem in C^0-MINLP form reads

$$\exists?\ (x, z): \quad \mathbf{c}_E(x) = 0, \quad \mathbf{c}_I(x, z) \ge 0, \tag{76}$$

where

$$\mathbf{c}_E(x, z) = \begin{pmatrix} \mathbf{c}_{\mathbb{V}}(x) \\ \mathbf{c}_{\mathbb{A}_{\text{pi}}}(x) \\ \mathbf{c}_{\mathbb{A}_{\text{lin-rs}}}(x) \\ \mathbf{c}_{\mathbb{A}_{\text{nonlin-rs}}}(x) \end{pmatrix}, \quad \mathbf{c}_I(x, z) = \begin{pmatrix} \mathbf{c}_{\mathbb{A}_{\text{va}}}(x, z) \\ \mathbf{c}_{\mathbb{A}_{\text{cv}}}(x, z) \\ \mathbf{c}_{\mathbb{A}_{\text{cg}}}(x, z) \end{pmatrix} \tag{77}$$

and

$$x = (x_{\mathbb{V}}, x_{\mathbb{A}_{\text{pi}}}, x_{\mathbb{A}_{\text{lin-rs}}}, x_{\mathbb{A}_{\text{nonlin-rs}}}, x_{\mathbb{A}_{\text{va}}}, x_{\mathbb{A}_{\text{cv}}}, x_{\mathbb{A}_{\text{cg}}}), \quad z = (z_{\mathbb{A}_{\text{va}}}, z_{\mathbb{A}_{\text{cv}}}, z_{\mathbb{A}_{\text{cg}}}). \tag{78}$$

Note that the equality constraints do not contain any discrete aspects in our case. Here nonsmooth aspects arise in all passive elements: $\mathbf{c}_{\mathbb{A}_{\text{pi}}}(x)$, $\mathbf{c}_{\mathbb{A}_{\text{lin-rs}}}(x)$, $\mathbf{c}_{\mathbb{A}_{\text{nonlin-rs}}}(x)$. Discrete decisions (with "genuine" binary variables) arise in the active elements: $\mathbf{c}_{\mathbb{A}_{\text{va}}}(x, z), \mathbf{c}_{\mathbb{A}_{\text{cv}}}(x, z), \mathbf{c}_{\mathbb{A}_{\text{cg}}}(x, z)$. The node model $\mathbf{c}_{\mathbb{V}}(x)$ is smooth and will be kept in its original form.

Collecting all smoothed and complementarity constrained components yields the following C^2-MPEC model:

$$\exists?\ x: \quad \text{s.t.} \quad \mathbf{c}_E(x) = 0, \quad \mathbf{c}_I(x) \ge 0, \tag{79}$$

where

$$\mathbf{c}_E(x) = \begin{pmatrix} \mathbf{c}_{\mathbb{V}}(x) \\ \mathbf{c}^{\text{smooth}}_{\mathbb{A}_{\text{pi}}}(x) \\ \mathbf{c}^{\text{smooth}}_{\mathbb{A}_{\text{lin-rs}}}(x) \\ \mathbf{c}^{\text{smooth}}_{\mathbb{A}_{\text{nonlin-rs}}}(x) \\ \mathbf{c}^{\text{mpec}}_{\mathbb{A}_{\text{va}}}(x) \\ \mathbf{c}^{\text{mpec}}_{\mathbb{A}_{\text{cv}},E}(x) \\ \mathbf{c}^{\text{mpec}}_{\mathbb{A}_{\text{cg}},E}(x) \end{pmatrix}, \quad \mathbf{c}_I(x) = \begin{pmatrix} \mathbf{c}^{\text{mpec}}_{\mathbb{A}_{\text{cv}},I}(x) \\ \mathbf{c}^{\text{mpec}}_{\mathbb{A}_{\text{cg}},I}(x) \end{pmatrix} \tag{80}$$

and

$$x = \left(x_{\mathbb{V}}, x^{\text{smooth}}_{\mathbb{A}_{\text{pi}}}, x_{\mathbb{A}_{\text{lin-rs}}}, x^{\text{smooth}}_{\mathbb{A}_{\text{nonlin-rs}}}, x^{\text{mpec}}_{\mathbb{A}_{\text{va}}}, x^{\text{mpec}}_{\mathbb{A}_{\text{cv}}}, x^{\text{mpec}}_{\mathbb{A}_{\text{cg}}}\right). \tag{81}$$

Finally the C^2-MPEC model (79) is regularized by any of the techniques from Sect. 2. The reformulation is generic except for one aspect. Some of the complementarity constraint pairings in $\mathbf{c}_E(x)$ do not have to be nonnegative as in the standard MPEC (27). For instance, this can happen for flow variables with a negative lower bound. In this case we square the corresponding functions so that condition (9) in Def. 1 is satisfied. We denote the regularization of (79) by C^2-NLP.

The primal heuristic for our concrete application actually involves additional problem-specific steps. As already mentioned in Sect. 4.1.10, we use an idealized compressor group model that disregards individual compressor units. Solutions of the above C^2-NLP (*stage-1*) are therefore refined by solving a second NLP (*stage-2*) that incorporates discrete decisions of individual compressor units by a special convexification approach; see [22] for details. If both stages are successful, we finally check the stage-2 feasible solution with a highly accurate validation NLP [37] to decide whether it is sufficiently accurate to be used in practice. Full details of these and other aspects of the application problem will be given in the future papers [22, 31, 17, 37, 38].

4.3. **Numerical Results.** We have tested the primal heuristic on the northern high-calorific gas network of our industry partner Open Grid Europe GmbH. The network model contains 452 pipes, 9 resistors, 35 valves, 23 control valves and 6 compressor groups. Gas is supplied at 31 entry nodes and discharged at 129 exit nodes. The intermediate C^2-MPEC models are regularized by penalization; the resulting NLP models are formulated with the modeling language GAMS v23.8.2 [18] and solved with the interior point code Ipopt v3.10 [44] on a Desktop PC with an Intel i7 920 CPU and 12 GiB RAM. The C++ software framework LaMaTTO++ [1] is used to implement the models and to interface the problem data. LaMaTTO++ is a framework for modeling and solving mixed-integer nonlinear programming problems on networks. It was originally developed by the working groups of Jens Lang and Alexander Martin and is now being used and extended within the ForNe project.

The test set includes some 12 000 NoVa instances of four different types: the sets SN $i, i = 2, 3, 4$, and Expert. SN $i, i = 2, 3, 4$, are automatically generated nominations. The generation process depends on the current set of contracts with supplying and discharging customers and historical data about nominated entry and exit capacities. The three sets of nominations mainly differ in how the contracts are modeled within the generation process. The sets SN i are of increasing degree of difficulty (see [22] for full details). The set Expert contains 40 manually designed nominations from our industry partner that are intended to represent *hard* instances.

The success rates are given in Table 1. The first column states the name of the NoVa test set and the second column gives the number of instances in the set. The

Test set	Size	Success rate (%)		
		stage-1	stage-2	NLP
SN2	3 882	100.0	90.34	72.1
SN3	4 077	100.0	85.45	65.0
SN4	4 227	99.97	88.24	59.0
Expert	40	100.0	47.50	30.0

Table 1. Success rate of primal heuristic on NoVa test sets.

Test set	stage-1			stage-2			NLP		
	min	max	avg	min	max	avg	min	max	avg
SN2	1.8	26.3	7.5	0.3	65.5	1.0	0.8	11.7	1.1
SN3	2.0	27.6	10.3	0.2	51.6	1.1	0.5	4.0	1.4
SN4	2.3	28.6	10.5	0.3	50.4	1.2	0.9	5.3	2.1
Expert	5.0	40.8	11.4	0.9	4.7	2.0	1.1	1.7	1.3

Table 2. Min, max, and average Ipopt CPU time (seconds) of primal heuristic on NoVa test sets.

Test set	stage-1			stage-2			NLP		
	min	max	avg	min	max	avg	min	max	avg
SN2	74	913	263.4	11	2 345	39.0	22	256	31.2
SN3	78	969	365.2	6	3 000	44.3	11	62	33.6
SN4	90	996	371.0	13	1 700	46.7	25	68	34.4
Expert	157	692	322.1	29	137	61.9	28	40	31.4

Table 3. Min, max, and average Ipopt iterations of primal heuristic on NoVa test sets.

following three columns show the percentage of instances that have successfully passed stage-1, stage-2, and the final validation NLP, respectively. Table 2 offers statistics of the CPU times of successful runs. For both stages and the validation NLP, the minimum, maximum and average CPU times of the different test sets are displayed. Similarly, Table 3 shows the minimum, maximum and average numbers of iterations of the successful instances. The maximum allowed number of iterations was set to 3 000.

In addition, the profiles in Fig. 1 and Fig. 2 display the distribution of the required iterations and CPU time. More formally, if P denotes one of the test sets and if $t_p, p \in P$, is the considered performance measure for problem p (here: the number of iterations or the CPU time), the plots show the graph $\tau \mapsto 100|\{p \in P \colon t_p \le \tau\}|/|P|$. Thus, the graphs display the percentage of feasible instances that need at most τ iterations (or τ seconds) to be solved. The displayed data represents stage-1, stage-2 and the validation NLP on test set SN3. The profiles of the remaining test sets look essentially similar.

We see that stage-1 is successful on all instances but one, in spite of the smoothings, MPEC regularization and other approximations that are involved. This is primarily due to the model simplifications employed here, in particular the idealized compressor group model. With the detailed compressor group model in stage two, discrete decisions for individual compressor units are found for 85% to 90% of the statistical nominations and for less than half of the expert nominations. Finally,

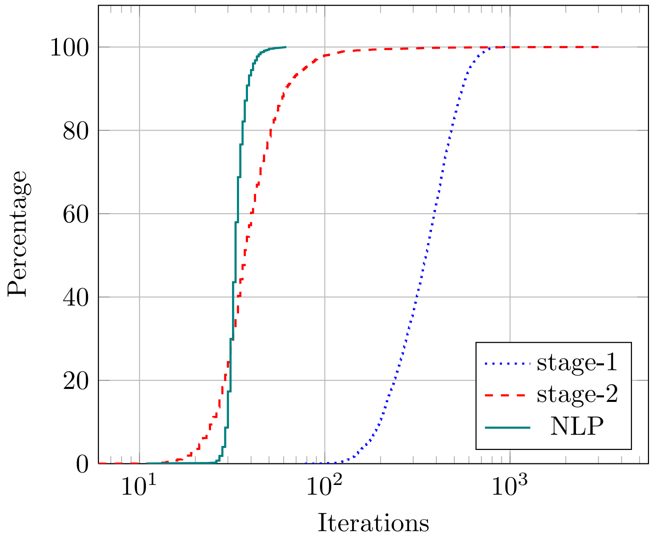

FIGURE 1. Profiles of Ipopt iterations in stages 1, 2, and validation NLP on test set SN3.

59% to 72% of all statistical nominations and 30% of the expert nominations pass the high-accuracy validation NLP.

It is apparent how the success rate slowly decreases with increasing difficulty of the test sets (see Table 1). In particular, the expert nominations really prove to be hard from stage-2 on. This is because of the central role of the compressor units: in hard cases they have to be operated close to their limits, and a highly accurate model is required to distinguish feasible and infeasible operating points. To lower the risk of missing feasible solutions early on, we therefore use increasingly restrictive compressor models in the successive stages.

A further reason of failure can occur in case of approximately biactive complementarity constraints, i.e. complementarity constraints with small non-zero values of both characteristic functions. Infeasible discrete decisions may then be deduced; for instance, a valve may be considered to be closed although a very small flow is actually required.

A comparison of the CPU times and iterations shows that most of the computational effort is spent on stage-1, taking about 10 s, while stage-2 and the validation NLP roughly require another second each. This indicates that the simplified stage-1 model is still reasonably hard (it encompasses the major difficulties) and that stage-2 can actually be considered as a refinement step. The profile plots support this interpretation, since the stage-1 curves are located significantly to the right of both the stage-2 and the validation NLP curves.

5. Summary

We have proposed a general reformulation technique as a primal heuristic for a certain class of nonsmooth mixed-integer nonlinear optimization problems. Our approach explicitly distinguishes discrete aspects and nonsmooth but continuous aspects. The former are handled with complementarity constraints; the latter are handled by generic or problem-specific smoothing techniques. Additional regularizations are applied to obtain a smooth and regular nonlinear optimization problem that can be solved by standard NLP solvers to produce (approximately) feasible solutions of the underlying nonsmooth MINLP efficiently. As a proof of concept, we

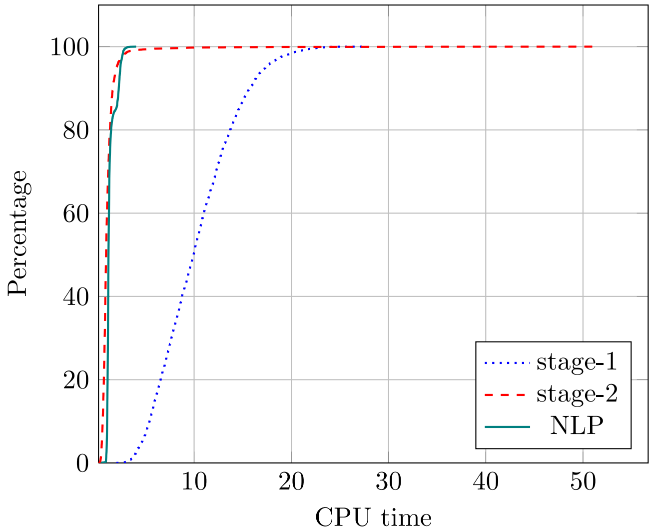

FIGURE 2. Profiles of Ipopt CPU time (seconds) in stages 1, 2, and validation NLP on test set SN3.

have successfully applied our heuristic to the problem of validation of nominations in real-life gas networks.

Acknowledgements

This work has been funded by the Federal Ministry of Economics and Technology owing to a decision of the German Bundestag. We would next like to thank our industry partner Open Grid Europe GmbH and the project partners in the ForNe consortium.

References

[1] *LaMaTTO++*. `http://www.mso.math.fau.de/edom/projects/lamatto.html`.
[2] B. T. Baumrucker, J. G. Renfro, and L. T. Biegler, *MPEC problem formulations and solution strategies with chemical engineering applications*, Computers & Chemical Engineering, (2008), pp. 2903–2913.
[3] H. G. Bock, E. Kostina, H. X. Phu, and R. Rannacher, eds., *Modeling, Simulation and Optimization of Complex Processes*, Springer, Berlin, 2005.
[4] C. Borraz-Sánchez and R. Z. Ríos-Mercado, *A Hybrid Meta-Heuristic Approach for Natural Gas Pipeline Network Optimization*, in Hybrid Metaheuristics, M. Blesa, C. Blum, A. Roli, and M. Sampels, eds., vol. 3636 of Lecture Notes in Computer Science, Springer, 2005, pp. 54–65.
[5] E. A. Boyd, L. R. Scott, and S. Wu, *Evaluating the quality of pipeline optimization algorithms*, in PSIG 29th Annual Meeting, Tucson, Arizona, Pipeline Simulation Interest Group, 1997. Paper 9709.
[6] J. Burgschweiger, B. Gnädig, and M. C. Steinbach, *Optimization models for operative planning in drinking water networks*, Optim. Eng., 10 (2009), pp. 43–73. Published online 2008.
[7] J. Burgschweiger, B. Gnädig, and M. C. Steinbach, *Nonlinear programming techniques for operative planning in large drinking water networks*, The Open Appl. Math. J., 3 (2009), pp. 14–28.
[8] R. G. Carter, *Compressor station optimization: Computational accuracy and speed*, in 28th Annual Meeting, Pipeline Simulation Interest Group, 1996. Paper 9605.
[9] D. Cobos-Zaleta and R. Z. Ríos-Mercado, *A MINLP model for a problem of minimizing fuel consumption on natural gas pipeline networks*, in Proc. XI Latin-Ibero-American Conference on Operations Research, 2002, pp. 1–9. Paper A48-01.
[10] G. B. Dantzig, *On the significance of solving linear programming problems with some integer variables*, Econometrica, 28 (1960), pp. pp. 30–44.

[11] D. de Wolf and Y. Smeers, *The gas transmission problem solved by an extension of the simplex algorithm*, Management Sci., 46 (2000), pp. 1454–1465.
[12] A.-V. DeMiguel, M. P. Friedlander, F. J. Nogales, and S. Scholtes, *A two-sided relaxation scheme for mathematical programs with equilibrium constraints*, SIAM J. Optim., 16 (2005), pp. 587–609.
[13] P. Domschke, B. Geissler, O. Kolb, J. Lang, A. Martin, and A. Morsi, *Combination of nonlinear and linear optimization of transient gas networks*, INFORMS Journal of Computing, 23 (2011), pp. 605–617.
[14] K. Ehrhardt and M. C. Steinbach, *KKT systems in operative planning for gas distribution networks*, Proc. Appl. Math. Mech., 4 (2004), pp. 606–607.
[15] ———, *Nonlinear optimization in gas networks*, in Bock et al. [3], pp. 139–148.
[16] A. Fischer, *A special Newton-type optimization method*, Optimization, 24 (1992), pp. 269–284.
[17] A. Fügenschuh, B. Geissler, R. Gollmer, C. Hayn, R. Henrion, B. Hiller, J. Humpola, T. Koch, T. Lehmann, A. Martin, R. Mirkov, A. Morsi, W. Römisch, J. Rövekamp, L. Schewe, M. Schmidt, R. Schwarz, R. Schultz, J. Schweiger, C. Stangl, M. C. Steinbach, and B. M. Willert, *Mathematical optimization for challenging network planning problems in unbundled liberalized gas markets*. In preparation.
[18] *GAMS – A Users Guide*, Redwood City, 1988.
[19] I. E. Grossmann, *Review of nonlinear mixed-integer and disjunctive programming techniques*, Optim. Eng., 3 (2002), pp. 227–252. Special issue on mixed-integer programming and its applications to engineering.
[20] T. Hoheisel, C. Kanzow, and A. Schwartz, *Theoretical and numerical comparison of relaxation methods for mathematical programs with complementarity constraints*, Preprint 299, Institute of Mathematics, University of Würzburg, Sept. 2010.
[21] X. M. Hu and D. Ralph, *Convergence of a penalty method for mathematical programming with complementarity constraints*, J. Optim. Theory Appl., 123 (2004), pp. 365–390.
[22] T. Koch, D. Bargmann, M. Ebbers, A. Fügenschuh, B. Geissler, N. Geissler, R. Gollmer, U. Gotzes, C. Hayn, H. Heitsch, R. Henrion, B. Hiller, J. Humpola, I. Joormann, V. Kühl, T. Lehmann, H. Leövey, A. Martin, R. Mirkov, A. Möller, A. Morsi, D. Oucherif, A. Pelzer, M. E. Pfetsch, L. Schewe, W. Römisch, J. Rövekamp, M. Schmidt, R. Schultz, R. Schwarz, J. Schweiger, K. Spreckelsen, C. Stangl, M. C. Steinbach, A. Steinkamp, I. Wegner-Specht, B. M. Willert, and S. Vigerske, *From simulation to optimization: Evaluating gas network capacities*. In preparation.
[23] K. Kraemer and W. Marquardt, *Continuous reformulation of MINLP problems*, in Recent Advances in Optimization and its Applications in Engineering, M. Diehl, F. Glineur, E. Jarlebring, and W. Michiels, eds., Springer Berlin Heidelberg, 2010, pp. 83–92.
[24] S. Leyffer, G. López-Calva, and J. Nocedal, *Interior methods for mathematical programs with complementarity constraints*, SIAM J. Optim., 17 (2004), pp. 52–77.
[25] H. M. Markowitz and A. S. Manne, *On the solution of discrete programming problems*, Econometrica, 25 (1957), pp. pp. 84–110.
[26] A. Martin, D. Mahlke, and S. Moritz, *A simulated annealing algorithm for transient optimization in gas networks*, Math. Methods Oper. Res., 66 (2007), pp. 99–115.
[27] A. Martin and M. Möller, *Cutting planes for the optimization of gas networks*, in Bock et al. [3], pp. 307–329.
[28] A. Martin, M. Möller, and S. Moritz, *Mixed integer models for the stationary case of gas network optimization*, Math. Program., 105 (2006), pp. 563–582.
[29] R. R. Meyer, *Mixed integer minimization models for piecewise-linear functions of a single variable*, Discrete Mathematics, 16 (1976), pp. 163 – 171.
[30] A. Osiadacz, *Nonlinear programming applied to the optimum control of a gas compressor station*, Int. J. Numer. Methods Eng., 15 (1980), pp. 1287–1301.
[31] M. E. Pfetsch, A. Fügenschuh, B. Geissler, N. Geissler, R. Gollmer, B. Hiller, J. Humpola, T. Koch, T. Lehmann, A. Martin, A. Morsi, J. Rövekamp, L. Schewe, M. Schmidt, R. Schultz, R. Schwarz, J. Schweiger, C. Stangl, M. C. Steinbach, S. Vigerske, and B. M. Willert, *Validation of nominations in gas network optimization: Models, methods, and solutions*. Submitted.
[32] K. F. Pratt and J. G. Wilson, *Optimization of the operation of gas transmission systems*, Transactions of the Institute of Measurement and Control, 6 (1984), pp. 261–269.
[33] R. Raman and I. E. Grossmann, *Modeling and computational techniques for logic based integer programming*, Comput. Chem. Eng., 18 (1994), pp. 563–578.

[34] R. Z. Ríos-Mercado, S. Kim, and A. E. Boyd, *Efficient operation of natural gas transmission systems: A network-based heuristic for cyclic structures*, Computers & Operations Research, 33 (2006), pp. 2323–2351.

[35] R. Z. Ríos-Mercado, S. Wu, L. R. Scott, and A. E. Boyd, *A reduction technique for natural gas transmission network optimization problems*, Ann. Oper. Res., 117 (2002), pp. 217–234.

[36] H. Scheel and S. Scholtes, *Mathematical programs with complementarity constraints: Stationarity, optimality and sensitivity*, Math. Oper. Res., 25 (2000), pp. 1–22.

[37] M. Schmidt, M. C. Steinbach, and B. M. Willert, *High detail stationary optimization models for gas networks — Part 1: Model components.*, IfAM Preprint 94, Inst. of Applied Mathematics, Leibniz Universität Hannover, 2012. Submitted.

[38] ———, *High detail stationary optimization models for gas networks — Part 2: Validation and results.* In preparation, 2013.

[39] S. Scholtes, *Convergence properties of a regularization scheme for mathematical programs with complementarity constraints*, SIAM J. Optim., 11 (2001), pp. 918–936.

[40] O. Stein, J. Oldenburg, and W. Marquardt, *Continuous reformulations of discrete-continuous optimization problems*, Comput. Chem. Eng., 28 (2004), pp. 1951–1966.

[41] M. C. Steinbach, *On PDE solution in transient optimization of gas networks*, J. Comput. Appl. Math., 203 (2007), pp. 345–361.

[42] D. Sun and L. Qi, *On NCP-functions*, Comput. Optim. Appl., 13 (1999), pp. 201–220. Computational Optimization - a Tribute to Olvi Mangasarian, Part II.

[43] J. P. Vielma and G. L. Nemhauser, *Modeling disjunctive constraints with a logarithmic number of binary variables and constraints*, in Integer Programming and Combinatorial Optimization, A. Lodi, A. Panconesi, and G. Rinaldi, eds., vol. 5035 of Lecture Notes in Computer Science, Springer Berlin Heidelberg, 2008, pp. 199–213.

[44] A. Wächter and L. T. Biegler, *On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming*, Math. Program., 106 (2006), pp. 25–57.

[45] S. Wright, M. Somani, and C. Ditzel, *Compressor station optimization*, in 30th Annual Meeting, Pipeline Simulation Interest Group, 1998. Paper 9805.

[46] S. Wu, R. Z. Ríos-Mercado, A. E. Boyd, and L. R. Scott, *Model relaxations for the fuel cost minimization of steady-state gas pipeline networks*, Tech. Rep. TR-99-01, University of Chicago, January 1999.

Martin Schmidt, Marc C. Steinbach, Bernhard M. Willert, Leibniz Universitt Hannover, Institute for Applied Mathematics, Welfengarten 1, 30167 Hannover, Germany

E-mail address: {mschmidt,mcs,willert}@ifam.uni-hannover.de

URL: www.ifam.uni-hannover.de/~{mschmidt,steinbach,willert}