

OUTER APPROXIMATION FOR GENERALIZED CONVEX MIXED-INTEGER NONLINEAR ROBUST OPTIMIZATION PROBLEMS

MARTINA KUCHLBAUER


ABSTRACT. We consider nonlinear robust optimization problems with mixed-integer decisions as well as nonconvexities. In detail, we consider cases where objective and constraint functions can be nonsmooth and generalized convex, i.e., f° -quasiconvex or f° -pseudoconvex.

We propose an algorithm for such robust optimization problems that does not require a certain structure of the adversarial problem but only requires that approximate worst cases are available. As a result, our algorithm finds a robust optimal solution up to a tolerance. Our method integrates a bundle method into an outer approximation approach where the bundle method is used for the arising continuous subproblems. We rely on methods from the literature, namely a bundle method for nonlinear and nonconvex robust optimization problems and outer approximation approaches for quasiconvex settings. Our contribution is to combine them to one convergent robust optimization method that can cope with inexactness of worst-case evaluations.

Further, we propose the gas transport under uncertainties as a relevant application and demonstrate that generalized convexity is fulfilled for a type of a network structure.

1. INTRODUCTION

Robust optimization problems become very challenging when we allow for nonlinearities and non-convexities in both, decision variables and uncertain parameters, as well as mixed-integer decisions. So far, there are no general solution approaches for this setting. Our motivation for the present work is the fact that we have at hand several tools, which we can combine to one convergent method to tackle this type of problem under some assumptions. The basis of our approach is to interpret a robust optimization problem as a nonsmooth optimization problem: Let $V(x, u) \leq 0$ be a constraint that is affected by uncertainty in u , against which we want to be robustly protected, and \mathcal{U} an uncertainty set, which is guaranteed to contain u . The constraint $V(x, u) \leq 0$ can then be written as infinitely many constraints, $V(x, u) \leq 0$ for all $u \in \mathcal{U}$, and eventually as one nonsmooth constraint, $\max_{u \in \mathcal{U}} V(x, u) \leq 0$. Motivated by this, we tackle our rather general class of robust optimization problems by nonsmooth methods. Further, to address non-convexities in the inner adversarial problem, i.e., $\max_{u \in \mathcal{U}} V(x, u)$, we require only inexact solutions to it. This results in inexact function evaluations in the outer nonsmooth optimization problem. With this, we can dispense with stronger linearity or convexity-type assumptions, which we would need for reformulation approaches such as those presented in, e.g., [4].

(M. Kuchlbauer ) UNIVERSITY OF TECHNOLOGY NUREMBERG, GERMANY; ULMENSTRASSE 52H, 90443 NÜRNBERG, GERMANY

E-mail address: martina.kuchlbauer@utn.de.

2020 *Mathematics Subject Classification.* 90C17, 90C30, 90C11, 90C47.

Key words and phrases. robust optimization, mixed-integer nonlinear optimization, generalized convexity, outer approximation, bundle method.

Also relying on nonsmooth optimization methods, [16, 15] proposed a bundle method for nonlinear robust optimization with non-convexities, inexact worst-case evaluations, and continuous decisions. With respect to that, we extend the setting here by mixed-integer decisions, but have to restrict the type of non-convexity. To handle mixed-integer decisions and hence the resulting nonsmooth mixed-integer nonlinear problem (MINLP), we can use an outer approximation method. In detail, [7] proposed for the continuous nonsmooth subproblems a combined method that incorporates a bundle method into an outer approximation method. In [17, 15], the method by [7] was extended to robust optimization, also with inexact worst-case evaluations. Since there are outer approximation methods as well as bundle methods for non-convex settings, a next step is to generalize the method by [7] to the case of non-convexities, as pointed out in [10]. From [17, 15], we have the extension of [7] to the setting of robust problems with inexact worst-case evaluations, which can deal also with non-convex continuous subproblems. Here, we extend this method to generalized convex settings by incorporating an according outer approximation method. The types of non-convexities, for which we allow, are f° -pseudo- and f° -quasiconvexity. We rely for the outer approximation method on the algorithms by [5] and [12]. In [5], the authors propose an outer approximation approach, combined with a feasibility pump, for problems with convex feasibility sets. The method by [12] is a single-tree outer approximation method with alternative feasibility cuts.

Our contribution is to propose the resulting combined method, to prove its convergence, to identify and analyze the applicable problem class with respect to the type of non-convexities and inexactness, and to propose a relevant and suitable application to gas transport on a certain type of network structure. We contribute to the field of mixed-integer nonlinear optimization as well as robust optimization. First, our contribution is to propose an outer approximation method with a concrete method for the continuous subproblems for a non-convex setting. We hence extend the approach by [7], where a bundle method is incorporated into an outer approximation method, to a non-convex setting. Second, our approach is more general in the sense that it applies also to nonlinear robust optimization problems. Hence, it extends the robust optimization method and setting from [17] to a, in possibly all decision variables, non-convex setting.

We conclude the introduction by preliminary definitions from the literature, on which we rely later on. As mentioned above, we use for our nonsmooth and generalized convex setting the concepts of f° -pseudo- and f° -quasiconvexity. We use the following definition of the generalized directional derivative.

Definition 1.1 ([6]). *Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be locally Lipschitz continuous. The generalized directional derivative of f at x in direction $d \in \mathbb{R}^n$ is defined by*

$$f^\circ(x; d) = \limsup_{y \rightarrow x, t \downarrow 0} \frac{f(y + td) - f(y)}{t}.$$

With this, we can define the following types of generalized convexity, which are relaxations of convexity.

Definition 1.2 ([3]). *A function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is f° -quasiconvex, if it is locally Lipschitz continuous and for all $x, y \in \mathbb{R}^n$*

$$f(y) \leq f(x) \Rightarrow f^\circ(x; y - x) \leq 0. \quad (1)$$

Definition 1.3 ([3]). *A function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is f° -pseudoconvex, if it is locally Lipschitz continuous and for all $x, y \in \mathbb{R}^n$*

$$f(y) < f(x) \Rightarrow f^\circ(x; y - x) < 0. \quad (2)$$

2. NON-CONVEX ROBUST MINLPS VIA OUTER APPROXIMATION

We consider a robust optimization problem with continuous decision variables $x \in \mathbb{R}^{n_x}$ and discrete decision variables $y \in \mathbb{Z}^{n_y}$. The constraints of the problem additionally depend on an uncertain parameter u that is known to lie in an uncertainty set \mathcal{U} . We aim for a solution that is robustly protected against this uncertainty and write the resulting robust optimization problem as

$$\begin{aligned} \min_{x,y} \quad & C(x, y) \\ \text{s. t.} \quad & V_i(x, y; u) \leq 0, i \in [n], u \in \mathcal{U} \\ & x \in X \\ & y \in Y \cap \mathbb{Z}^{n_y}, \end{aligned} \tag{RO}$$

where bounds and further linear constraints on the continuous variables are modeled by $X \subset \mathbb{R}^{n_x}$ and $Y \subseteq \mathbb{R}^{n_y}$, and Y is compact. We require that the functions C and V_i , for every $i \in [n]$, are locally Lipschitz continuous and lower C^1 . Now we can interpret this robust optimization problem as a nonsmooth optimization problem: we reformulate the constraints $V_i(x, y; u) \leq 0$, which have to be fulfilled for every $i \in [n]$ and $u \in \mathcal{U}$, to one nonsmooth constraint $H(x, y) \leq 0$ with

$$H(x, y) := \max_{u \in \mathcal{U}} V(x, y; u), \quad V(x, y; u) := \max_{i \in [n]} V_i(x, y; u). \tag{3}$$

We call this inner problem, i.e., the evaluation of this optimal value function H , the adversarial problem. This problem can be challenging, e.g., when it involves non-convexities. We allow for this and do not make any further assumptions on its structure. We only require that for every $x, y \in \mathbb{R}^{n_x+n_y}$ and every given tolerance ε_H , we have access to an approximate worst case u with $V(x, y; u) \geq H(x, y) - \varepsilon_H$.

With this nonsmooth optimal value function H , we write the robust optimization problem (RO) equivalently as the nonsmooth optimization problem

$$\begin{aligned} \min_{x,y} \quad & C(x, y) \\ \text{s. t.} \quad & H(x, y) \leq 0 \\ & x \in X \\ & y \in Y \cap \mathbb{Z}^{n_y}. \end{aligned} \tag{P}$$

Efficient approaches for such MINLPs are outer approximation methods, which were originally proposed by [8, 11] for the case of smooth and convex functions, and extended to nonsmooth settings in [9, 21, 22]. They further have been generalized to quasiconvex settings in [5, 12].

We follow [12, 22] for an outer approximation method in a non-convex and nonsmooth setting and aim to solve the problem (P) based on their methods. To prepare this, we concretize here the arising subproblems and master problem. In the course of an outer approximation approach as in [12], we encounter the following subproblems.

First, in every outer approximation iteration K , we solve for a fixed integer solution \hat{y}_K the NLP-relaxation

$$\begin{aligned} \min_x \quad & C(x, \hat{y}_K) \\ \text{s. t.} \quad & H(x, \hat{y}_K) \leq 0 \\ & x \in X. \end{aligned} \tag{NLP}(\hat{y}_K)$$

Second, if the NLP-relaxation is infeasible, we solve the following projection problem for the fixed integer solution \hat{y}_K .

$$\begin{aligned} \min_{x,y} \quad & d_{\hat{y}_K}(y) \\ \text{s.t.} \quad & H(x,y) \leq 0 \\ & x \in X, y \in Y, \end{aligned} \quad (\text{ProjNLP}(\hat{y}_K))$$

with $d_{\hat{y}_K}(y) = \|y - \hat{y}_K\|_1$. Both problems have only continuous variables, but are nonlinear. We denote the resulting subproblem solution in an outer approximation iteration K by (x_K, y_K) .

As a master problem, we iteratively solve the following mixed-integer linear problem, which is a relaxation of the original problem (P) (see [12]).

$$\begin{aligned} \min_{x,y} \quad & C(x,y) \\ \text{s.t.} \quad & \begin{pmatrix} \xi_L \\ \eta_L \end{pmatrix}^T \begin{pmatrix} x - x_L \\ y - y_L \end{pmatrix} \leq 0 \quad \forall (x_K, y_K) \in S^K \\ & x \in X, y \in Y \cap \mathbb{Z}^{n_y}, \end{aligned} \quad (\text{MP}^K)$$

where $S^K := \{(x_L, y_L) \mid L \leq K\}$ is the set of encountered subproblem solutions, and ξ_L, η_L , with $L \leq K$, define linearized constraints generated at subproblem solutions.

In the following, in Section 3, we elaborate how the subproblems can, despite inexact worst-case evaluations, be tackled by an existing bundle method. Then, in Section 4, we concretize our resulting inexact outer approximation approach that incorporates the bundle method.

3. A BUNDLE METHOD FOR THE SUBPROBLEMS

In every outer approximation iteration, we solve the subproblem $(\text{NLP}(\hat{y}_K))$ and/or the subproblem $(\text{ProjNLP}(\hat{y}_K))$ for a fixed \hat{y}_K . These subproblems are nonlinear robust optimization problems with continuous decision variables and we face inexactness in worst-case evaluations. For this setting, we can use the adaptive bundle method from [16]. The problems have the general form

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ \text{s.t.} \quad & v_i(x) \leq 0 \quad \forall i \in [n], \end{aligned} \quad (\text{NLP})$$

where $g(x) = \max_{u \in \mathcal{U}} v(x, u) = \max_{i \in [n]} v_i(x, u)$ and all functions are locally Lipschitz. We can apply the adaptive bundle method only to unconstrained problems, so that we consider the above problem by its penalty formulation $\min_{x \in \mathbb{R}^n} f(x) + \psi[g(x)]^+$. In detail, the penalty formulation of the NLP-relaxation $(\text{NLP}(\hat{y}_K))$ reads

$$\min_{x \in X} \quad C(x, \hat{y}_K) + \psi H(x, \hat{y}_K), \quad (\text{NLP}_\psi(\hat{y}_K))$$

and the penalty formulation of the projection problem $(\text{ProjNLP}(\hat{y}_K))$ reads

$$\min_{x \in X, y \in Y} \quad d_{\hat{y}_K}(y) + \psi H(x, y) \quad (\text{ProjNLP}_\psi(\hat{y}_K))$$

with $d_{\hat{y}_K}(y) = \|y - \hat{y}_K\|_1$. The linear constraints that are modeled by X and Y can also be integrated into the function H , by modeling them as further constraint functions V_i , or be handled via an additional penalty term with penalty parameter ψ_X .

We have assumed that the involved functions are lower C^1 and that we have access to an approximate worst case for every given tolerance (see Section 2). As a natural consequence, we assume that for every such approximate worst case $u_{(x,y)}$, we have access to the according function value $V(x, y, u_{(x,y)})$ and a subgradient $g_{(x,y)} \in \partial_{(x,y)} V(x, y, u_{(x,y)})$, where $\partial_{(x,y)}$ denotes a joint subdifferential. When we deal with the first subproblem, $(NLP_\psi(\hat{y}_K))$, it suffices for the bundle method to have access to a subgradient with respect to x . With this, we can apply the adaptive bundle method and rely on the convergence results in [16].

To avoid troubles in the outer approximation algorithm as, e.g., not recognizing feasibility or infeasibility and then cycling between solving the NLP-relaxation and the projection problem, we are interested in the relation between solutions to the original subproblems and to the penalty subproblems. To ensure that solving these penalty problems leads to solutions of the original problems, we assume f° -pseudoconvexity of the constraints and that a constraint qualification holds.

First, we assume that the Cottle constraint qualification holds, which is defined as follows.

Definition 3.1 (CCQ, see, e.g., [3]). *The Cottle constraint qualification is fulfilled by the problem (NLP) at a feasible point x , if either $0 \notin \partial g(x)$ or $g(x) < 0$.*

We note that it implies the Slater CQ (see [3]), whereas the converse does not necessarily hold in the non-convex case.

Assumption 3.2. *If the problem $(NLP(\hat{y}_K))$ is feasible, then it fulfills the CCQ. In the case of infeasibility, the problem $(ProjNLP(\hat{y}_K))$ fulfills the CCQ.*

With this assumption, we can make use of the following result.

Theorem 3.3 ([3]). *Let (NLP) satisfy the CCQ at x^* with $g(x^*) \leq 0$, $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be f° -pseudoconvex and $g : \mathbb{R}^n \rightarrow \mathbb{R}$ be quasiconvex and subdifferentially regular at x^* . Then, x^* is a global optimum of (NLP) if and only if there exists a multiplier $\lambda \geq 0$ such that $\lambda g(x^*) = 0$ and*

$$0 \in \partial f(x^*) + \lambda \partial g(x^*).$$

Further, we make the following, stronger assumption on our constraint functions.

Assumption 3.4. *The constraint functions $V_i(x, y, u)$, for every $i \in [n], u \in \mathcal{U}$, are f° -pseudoconvex. Further, for every $\psi > 0$, the objective functions of the penalized problems are f° -pseudoconvex.*

For example, with the assumption of additively strict monotonicity, $f(\cdot) + \lambda \bar{g}(\cdot)$ is f° -pseudoconvex ([3, Theorem 5.7]). Another example would be, for the projection problem, that the generalized directional derivative of the constraint function has always a norm of 0 or greater than 1.

We obtain in Theorem 3.6 for this case of f° -pseudoconvex functions an equivalence result of penalized and original problems. For this, we rely on the following equivalence of critical points and global minima for f° -pseudoconvex functions.

Theorem 3.5 ([3, Theorem 5.2]). *If a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is f° -pseudoconvex, then it has a global minimum at $x^* \in \mathbb{R}^n$ if and only if*

$$0 \in \partial f(x^*).$$

Theorem 3.6. *Let the problem (NLP) be such that it fulfills the CCQ, f is f° -pseudoconvex and g is subdifferentially regular and f° -pseudoconvex. Further, for every $\Lambda > 0$, the function $f(x) + \Lambda[g(x)]^+$ is f° -pseudoconvex. Then, there is a λ such that, for all $\Lambda > \lambda$, a point*

x^* solves (NLP) if and only if it solves

$$\min_{x \in \mathbb{R}^n} f(x) + \Lambda[g(x)]^+. \quad (4)$$

Proof. As f° -pseudoconvexity implies quasiconvexity [3, Theorem 5.19, Corollary 5.2], the assertion of Theorem 3.3 also holds for the setting of constraint functions that are not only quasiconvex but also f° -pseudoconvex.

Let x^* be a feasible and optimal solution to (NLP). Then it also is feasible and optimal to the problem

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ \text{s. t.} \quad & g(x) \leq 0. \end{aligned} \quad (5)$$

The function $g(x)$ is again f° -pseudoconvex [3, Theorem 5.6], and due to Theorem 3.3, there is a $\lambda \geq 0$ such that $0 \in \partial f(x^*) + \lambda \partial g(x^*)$ and $\lambda g(x^*) = 0$. Further, for the function $\bar{g}(x) := [g(x)]^+$, we have that, if $g(x) = 0$, then $\partial \bar{g}(\cdot) = \text{conv}\{0, \partial g(x)\}$. With this, we get $0 \in \partial f(x^*) + \lambda \partial \bar{g}(x^*)$.

From Theorem 3.5, we obtain that x^* globally solves the problem $\min_{x \in \mathbb{R}^n} f(x) + \lambda \bar{g}(x)$. So does every other optimal solution to (NLP), and as $\bar{g}(x) \geq 0$ for every $x \in \mathbb{R}^n$, this also holds if we increase the value of λ .

Now let \hat{x} be an optimal solution to (4) for $\Lambda > \lambda$ with λ chosen as before. Hence, we have $f(\hat{x}) + \Lambda \bar{g}(\hat{x}) = f(x^*)$. If $\bar{g}(\hat{x}) > 0$, we obtain

$$f(\hat{x}) + \Lambda \bar{g}(\hat{x}) > f(\hat{x}) + \lambda \bar{g}(\hat{x}) \geq f(x^*) + \lambda \bar{g}(x^*) = f(x^*), \quad (6)$$

which leads to a contradiction, so that $\bar{g}(\hat{x}) = 0$. Thus, \hat{x} is an optimal solution to (5) and hence to (NLP). \square

Output of the adaptive bundle method We choose approximate values $\tilde{H}(\cdot, \cdot)$ of H at a final bundle iterate x^l (and y^l , in the case of the projection problem) such that $\tilde{H}(x^l, \hat{y}_K) = V(x^l, \hat{y}_K, u^l)$ or $\tilde{H}(x^l, y^l) = V(x^l, y^l, u^l)$, respectively. As solution, we output $x_K := x^l$ and, in the case of the projection problem, $y_K := y^l$.

For (NLP) $_\psi(\hat{y}_K)$, we choose within the bundle method subgradients (s_x^l, s_y^l) that fulfill the property of subgradients in an f° -quasiconvex setting (see Theorem 1.2), i.e.,

$$x \in Y, y \in Y \cap \mathbb{Z}^{n_y}, H(x, y) \leq 0 \Rightarrow \begin{pmatrix} s_x^l \\ s_y^l \end{pmatrix}^T \begin{pmatrix} x - x^l \\ y - \hat{y}_K \end{pmatrix} \leq 0. \quad (7)$$

For the output of the adaptive bundle method, we proceed as in [17]. If we have a solution that we label as feasible, then we obtain an output with $\tilde{H}(x^l, \hat{y}_K) = H(x^l, \hat{y}_K) - \varepsilon_H^p \leq 0$ and cluster points $\alpha_K \in \partial C(x^l, \hat{y}_K)$, $\tilde{\xi}_K \in \tilde{\partial}_a H(x^l, \hat{y}_K)$, $\zeta_K \in N(X, x^l)$ with $\alpha_K + \psi \tilde{\xi}_K + \psi_X \zeta_K = 0$ for penalty parameters ψ, ψ_X .

For (ProjNLP) $_\psi(\hat{y}_K)$, we choose within the bundle method subgradients (s_x^l, s_y^l) that fulfill, analogously to (8),

$$x \in Y, y \in Y \cap \mathbb{Z}^{n_y}, H(x, y) \leq 0 \Rightarrow \begin{pmatrix} s_x^l \\ s_y^l \end{pmatrix}^T \begin{pmatrix} x - x_K \\ y - y_K \end{pmatrix} \leq 0. \quad (8)$$

Again, we choose the output of the adaptive bundle method analogously to [17]. We obtain an output with $\tilde{H}(x_K, y_K) = H(x_K, y_K) - \varepsilon_H^p \leq 0$ and cluster points $\alpha_K \in \partial d_{\hat{y}_K}(y_K)$, $(\tilde{\xi}_K, \tilde{\eta}_K) \in \tilde{\partial}_a H(x_K, y_K)$, $\zeta_K \in N(X \times Y, (x_K, y_K))$ with $(0, \alpha_K)^T + \psi(\tilde{\xi}_K, \tilde{\eta}_K)^T + \psi_X \zeta_K = 0$ for penalty parameters ψ, ψ_X .

When the projection problem is solved with a positive objective value, then we know from the above equivalence result that the NLP-relaxation is indeed infeasible. Also, when the NLP-relaxation is solved with a feasible point in the sense that $\tilde{H}(x^l, y^l) \leq 0$, then we directly obtain that the NLP-relaxation is indeed feasible. For the case that the NLP-relaxation is not solved with a feasible point but the projection problem is solved with an objective value of zero, we can use the strategies as proposed in [17]: We can label the solution as feasible with an according tolerance or we can choose the approximation strategy and tolerances in the bundle method such that an exact solution is obtained.

4. AN INEXACT OUTER APPROXIMATION ALGORITHM

We now incorporate the adaptive bundle method from [16] into an outer approximation method that can handle non-convexities and applies to the problem (P) . In detail, we elaborate assumptions on the output of a subproblem solver and then present the outer approximation method. We will see that the assumptions are fulfilled by the output of the adaptive bundle method as specified in Section 3. This integration of the output of the bundle method works similar to the integration for the case with stronger convexity assumptions that has been treated in [17].

In an outer approximation approach, we need to generate cutting planes that are valid for feasible points of (P) . Instead of convexity assumptions, we require the following property for approximate subgradients $(\tilde{\xi}_K, \tilde{\eta}_K)$ at iterates (x_K, y_K) with $H(x_K, y_K) \geq 0$, where K is an outer approximation iteration.

$$x \in Y, y \in Y \cap \mathbb{Z}^{n_y}, H(x, y) \leq 0 \Rightarrow \begin{pmatrix} \tilde{\xi}_K \\ \tilde{\eta}_K \end{pmatrix}^T \begin{pmatrix} x - x_K \\ y - y_K \end{pmatrix} \leq 0. \quad (9)$$

This property (9) is fulfilled, e.g., for the following f° -quasiconvex setting.

Example 4.1. Let $V_i(\cdot, \cdot, u)$ be f° -quasiconvex for every $i \in [n]$ and every $u \in \mathcal{U}$. Then also $V(\cdot, \cdot, u)$ for every $u \in \mathcal{U}$, and $H(\cdot, \cdot)$, as defined in (3), are f° -quasiconvex. With this, subgradients of $H(\cdot, \cdot)$ fulfill the property (9) due to f° -quasiconvexity (see, e.g., [3]). Now let u_K be for an iterate (x_K, y_K) an approximate worst case with

$$\tilde{H}(x_K, y_K) = V(x_K, y_K, u_K) \geq 0, \quad (10)$$

where $(x_K, y_K, u_K) \in X \times Y \times \mathcal{U}$. For every $(x, y) \in X \times Y$, we have $V(x, y, u_K) \leq H(x, y)$. If $H(x, y) \leq 0$, then we have $V(x, y, u_K) \leq 0$ and hence $V(x, y, u_K) \leq V(x_K, y_K, u_K)$. From f° -quasiconvexity of $V(\cdot, \cdot, u_K)$, it follows that every subgradient $(\tilde{\xi}_K, \tilde{\eta}_K) \in \partial_{x,y} V(x_K, y_K, u_K)$ fulfills the property (9).

As discussed in the previous section, we can expect from the adaptive bundle method only inexact solutions to the continuous subproblems. For such an output of a subproblem solution, we make the following assumption. Later in this section, we prove that we can ensure with this assumption that the outer approximation approach terminates finitely.

Assumption 4.2. At an arbitrary iteration K , the problem $(NLP(\hat{y}_K))$ can be solved with output $(x_K, \varepsilon_H^K, \tilde{H}(x_K, y_K), \tilde{\xi}_K, \tilde{\eta}_K, \zeta_K)$ such that $x_K \in X$ and

$$\exists 0 \leq \varepsilon_H^p \leq \varepsilon_H^K : \tilde{H}(x_K, \hat{y}) = H(x_K, \hat{y}) - \varepsilon_H^p \leq 0, \quad (11a)$$

$$(\tilde{\xi}_K, \tilde{\eta}_K) \text{ fulfills (9), } \zeta_K \in N(X, x_K), \quad (11b)$$

$$\exists \psi, \psi_X > 0 : 0 = c_x + \psi \tilde{\xi}_K + \psi_X \zeta_K. \quad (11c)$$

If $(NLP(\hat{y}_K))$ is feasible, then it can be solved with $\tilde{H}(x_K, \hat{y}) = 0$.

Further, the problem $(ProjNLP(\hat{y}_K))$ can be solved with output $((x_K, y_K), \varepsilon_H^K, \tilde{H}(x_K, y_K), \alpha_K, \tilde{\xi}_K, \tilde{\eta}_K, \zeta_K)$ such that $(x_K, y_K) \in X \times Y$ and

$$\exists 0 \leq \varepsilon_H^p \leq \varepsilon_H^K : \tilde{H}(x_K, y_K) = H(x_K, y_K) - \varepsilon_H^p \leq 0, \quad (12a)$$

$$\alpha_K \in \partial d_{\hat{y}_K}(y_K), (\tilde{\xi}_K, \tilde{\eta}_K) \text{ fulfills (9)}, \zeta_K \in N(X \times Y, (x_K, y_K)), \quad (12b)$$

$$\exists \psi, \psi_X > 0 : 0 = \begin{pmatrix} 0 \\ \alpha_K \end{pmatrix} + \psi \begin{pmatrix} \tilde{\xi}_K \\ \tilde{\eta}_K \end{pmatrix} + \psi_X \zeta_K. \quad (12c)$$

We use the output of the adaptive bundle method to generate cutting planes that refine the relaxation of the original problem by the master problem. In an outer approximation iteration K , we then obtain the following inexact master problem.

$$\begin{aligned} \min_{x, y} \quad & C(x, y) \\ \text{s. t.} \quad & \begin{pmatrix} \tilde{\xi}_L \\ \tilde{\eta}_L \end{pmatrix}^T \begin{pmatrix} x - x_L \\ y - y_L \end{pmatrix} \leq 0 \quad \forall (x_L, y_L) \in S^K \\ & x \in X, y \in Y \cap \mathbb{Z}^{n_y}. \end{aligned} \quad (\widetilde{MP}^K)$$

Due to Theorem 4.2, every outer approximation iterate (x_K, y_K) fulfills property (9), and we directly infer the following result.

Lemma 4.3. *If (x, y) is feasible for (P) , then it is feasible for (\widetilde{MP}^K) .*

Now we can state the resulting outer approximation method in Algorithm 1.

In order to show finite convergence of the outer approximation method, we now show that every integer part \hat{y}_K of a solution to a master problem (\widetilde{MP}^K) , with K an outer approximation iteration, can occur only once.

Lemma 4.4. *In an iteration K of Algorithm 1, if a point (x, \hat{y}_L) with $L \leq K$ and $x \in X$ is feasible for (\widetilde{MP}^K) , then $C(x, \hat{y}_L) > \Theta^K - \varepsilon_{oa}$.*

Proof. We distinguish the two cases that $(x_L, y_L) \in S^K$ has been output as a solution to $(NLP(\hat{y}_L))$ with $\tilde{H}(x_L, y_L) \leq 0$ and that it has been output as a solution to $(ProjNLP(\hat{y}_L))$.

i) Let x_L be the solution output for $(NLP(\hat{y}_L))$, fulfilling (11) with $\tilde{H}(x_L, y_L) \leq 0$. Then, the cutting plane

$$\begin{pmatrix} \tilde{\xi}_L \\ \tilde{\eta}_L \end{pmatrix}^T \begin{pmatrix} x - x_L \\ y - y_L \end{pmatrix} \leq 0 \quad (13)$$

is a constraint of (\widetilde{MP}^K) and $\hat{y}_L = y_L$. If (x, \hat{y}_L) is feasible for (\widetilde{MP}^K) , then $\tilde{\xi}_L^T(x - x_L) \leq 0$. As $x \in X$, we have $\zeta_K^T(x - x_L) \leq 0$. From (11c), it thus follows that $c_x^T(x - x_L) \geq 0$, and hence $c^T(x, \hat{y}_L)^T \geq c^T(x_L, y_L)^T$. The claim follows with $c^T(x_L, \hat{y}_L)^T \geq \Theta^K$.

ii) Let $(NLP(\hat{y}_L))$ be detected to be infeasible. Then, we have in Algorithm 1 a point (x_L, y_L) as a solution to $(ProjNLP(\hat{y}_K))$, fulfilling (12). If (x, \hat{y}_L) is feasible for (\widetilde{MP}^K) , then

$$\begin{pmatrix} \tilde{\xi}_L \\ \tilde{\eta}_L \end{pmatrix}^T \begin{pmatrix} x - x_L \\ \hat{y}_L - y_L \end{pmatrix} \leq 0 \quad (14)$$

Algorithm 1 OA for non-convex robust MINLP with inexactness

```

1: Fix parameter  $\varepsilon_{oa} > 0$ .
2: Choose initial values:  $y^1 \in Y, S^0 := \emptyset, \theta^0 := 0, \Theta^0 := \infty, K := 0$ .
3: while  $\widetilde{MP}^K$  feasible and  $\theta^K \leq \Theta^K - \varepsilon_{oa}$  do
4:   Increase  $K$  by 1.
5:   Solve  $(NLP_\psi(\hat{y}_K))$ , fulfilling Theorem 4.2 and denote the output by
       $(x_K, \varepsilon_H^K, \tilde{H}(x_K, y_K), \tilde{\xi}_K, \tilde{\eta}_K, \zeta_K)$ . ▷ bundle
6:   if  $\tilde{H}(x_K, \hat{y}_K) \leq 0$  then
7:      $y_K \leftarrow \hat{y}_K$ .
8:      $S^K := S^{K-1} \cup \{(x_K, y_K)\}$ .
9:      $\Theta^K := \min\{\Theta^{K-1}, C(x_K, y_K)\}$ .
10:  else
11:    Solve  $(ProjNLP_\psi(\hat{y}_K))$ , fulfilling Theorem 4.2 and denote the output by
         $((x_K, y_K), \varepsilon_H^K, \tilde{H}(x_K, y_K), \alpha_K, \tilde{\xi}_K, \tilde{\eta}_K, \zeta_K)$ . ▷ bundle
12:    if  $d_{\hat{y}_K}(y_K) > 0$  then
13:       $S^K := S^{K-1} \cup \{(x_K, y_K)\}$ .
14:       $\Theta^K := \Theta^{K-1}$ .
15:    else
16:      Go to line 5 and enforce  $\tilde{H}(x_K, \hat{y}) = 0$ .
17:    end if
18:  end if
19:  Solve  $(\widetilde{MP}^K)$  with value  $\theta^K$  and denote the solution's integer part by  $\hat{y}_{K+1}$ .
20: end while
21: if  $\Theta^K = \infty$  then
22:   Stop and claim infeasibility.
23: end if
24: Set  $K^* \in \{J \mid J \leq K, \Theta^K = C(x_K, y_K)\}$ .
25: Return  $(x_{K^*}, y_{K^*})$ .

```

We make use of (12b)-(12c). From $\alpha_L \in \partial d_{\hat{y}_L}(y_L)$, we have

$$\alpha_K^T \begin{pmatrix} x - x_L \\ \hat{y}_L - y_L \end{pmatrix} = -\|y_L - \hat{y}_L\| < 0. \quad (15)$$

Further, from $\zeta_L \in N(X \times Y, (x_L, y_L))$ and $(x, \hat{y}_L) \in X \times Y$, we have $\zeta^T(x - x_L, \hat{y}_L - y_L)^T \leq 0$. Together with (12c), this results in a contradiction to (14). The point (x, \hat{y}_L) is not feasible for (\widetilde{MP}^K) and the claim holds for this case. \square

From the results above, we can infer finite convergence of our outer approximation algorithm.

Theorem 4.5. *The outer approximation algorithm Algorithm 1 converges finitely. It either correctly detects infeasibility of (P) or it outputs a solution that is $\varepsilon_H^{K^*}$ -feasible and ε_{oa} -optimal.*

Proof. Follows directly from Theorem 4.3 and Theorem 4.4. \square

5. AN EXAMPLE FROM GAS TRANSPORT

We demonstrate in this section the practical relevance of our setting by the example of robust gas transport. In detail, we consider the stationary isothermal gas transport problem with discrete-continuous decision variables [13] under robustly treated uncertainties. In detail, we allow the demand and physical parameters such as pressure loss coefficients to be uncertain. Similar and equivalent settings have been treated, e.g., by [17, 15, 16, 2, 1]. Our model is the same as used in [17], but we impose different assumptions on the network and obtain a setting that is in a sense complementary: In [17], it is assumed that there are no compressors to control on cycles, which leads to a problem of convex type. Here, we allow for a compressor on a cycle, but not for any other active elements in the network, and show that this fits into our pseudoconvex setting. This is not a more general setting than handled in [17], but we demonstrate that exactly the type of constraint that is missing there can be covered by our pseudoconvex setting. However, it remains for future research to combine the two settings. The direct combination of the settings is not covered by the class of generalized convex problems considered here.

We model a gas network by a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{A})$ with n nodes, m arcs and an incidence matrix $A \in \{-1, 0, 1\}^{n \times m}$. Pipes and compressors are modeled by the arcs. We further denote by state variables $q \in \mathbb{R}^m$ and $\pi \in \mathbb{R}^n$ the flow on arcs and the pressure at nodes, respectively. Further, we assume the pressure to be fixed at one root node $r \in \mathcal{V}$. The pressure loss on pipes and compressors is modeled by the Weymouth equation, i.e., $\lambda_a q_a |q_a|$, with pressure loss coefficients λ , and a linear compressor model with control variables Δ (cf., e.g., [1]), respectively. With these assumptions, the state variable values are uniquely determined once the uncertain parameters (demand d , pressure loss coefficients λ) and control variables are fixed [20]. This relies on the following cycle equation, which has a unique solution. For this, we divide the set of arcs in basic arcs $B \subseteq \mathcal{A}$ that define a spanning tree and the remaining nonbasic arcs $N \subseteq \mathcal{A}$ —one for each fundamental cycle. The flow q is decomposed accordingly into the sum of basic flow, q_B , which is induced by the demand on the spanning tree and the nonbasic flow, q_N , which is the remaining flow on cycles. Once the nonbasic flow is determined, the basic flow can easily be determined by solving a linear equation and we write it as $q_B(q_N)$. The cycle equation now describes the pressure loss on the arcs of fundamental cycles, which has to sum up to 0 for every cycle. Denoting the pressure loss on an arc a by F_a , this reads, for a cycle $\mathcal{C} \subseteq \mathcal{A}$ with $\tilde{a} \in \mathcal{C} \cap N$, with all arcs pointing in the same cyclic direction,

$$\sum_{a \in \mathcal{C} \setminus \tilde{a}} F_a(\Delta, q_B(q_N)) = -F_{\tilde{a}}(\Delta, q_N). \tag{16}$$

In the gas transport problem of controlling the compressor variables, we encounter constraints of the form $\underline{\pi}_v \leq \pi_v(\Delta; d, \lambda) \leq \bar{\pi}_v$ for every node $v \in \mathcal{V}$. Under robustly treated uncertainties, this is desired to hold for every possible realization of demands and pressure loss coefficients. Hence, we encounter in the robust gas transport problem a constraint function of the form

$$H(\Delta) := \max_{(d, \lambda) \in \mathcal{U}} \max_{v \in \mathcal{V}} \max\{\underline{\pi}_v - \pi_v(\Delta; d, \lambda), \pi_v(\Delta; d, \lambda) - \bar{\pi}_v\} \leq 0. \tag{17}$$

We make the following assumption on the network, under which we then show that the constraint function H is f° -pseudoconvex.

Assumption 5.1. *There is one compressor in the network and it is on a cycle.*

We can write the pressure at a fixed node $v \in \mathcal{V}$ as the sum of pressure losses on the way from the root node r to v ,

$$\pi_v(\Delta; d, \lambda) = \pi_r + \left((\hat{A}_B^{-1})^T F_B(\Delta, q_B(q_N(\Delta; d, \lambda))) \right)_v. \quad (18)$$

The only nonbasic flow value that depends on the value of Δ is the one on the nonbasic arc that corresponds to the cycle with a compressor. We denote this arc by $\tilde{a} = (\tilde{u}, \tilde{v})$ and the set of arcs on this cycle by \tilde{C} . Without loss of generality, we choose for the root node $r = \tilde{u}$ and can assume that \tilde{a} is the compressor. Further, we can assume that the arcs in \tilde{C} are oriented such that they point into the same cyclic direction, i.e., no two arcs have the same vertex as head or as tail, respectively. For every node v , we see from (18) and Theorem 5.1 that the pressure π_v depends linearly and monotonically increasing on the flow value $q_{\tilde{a}}$.

The flow $q_{\tilde{a}}$ is computed by the cycle equation

$$\sum_{a \in \tilde{C} \setminus \tilde{a}} F_a(\Delta, q_B(q_N)) = -F_{\tilde{a}}(\Delta, q_N) \quad (19)$$

$$\Leftrightarrow \sum_{a \in \tilde{C} \setminus \tilde{a}} \lambda_a q_a(q_{\tilde{a}}) |q_a(q_{\tilde{a}})| = \Delta_{\tilde{a}}. \quad (20)$$

For fixed $\Delta_{\tilde{a}}$, this equation then has a unique solution, as mentioned above. Further, due to the linearity of $q_B(q_{\tilde{a}})$ and the strict monotonicity of the left-hand side with respect to q_B , it follows that $q_{\tilde{a}}$ is strictly monotonically increasing in $\Delta_{\tilde{a}}$. It follows that $\pi_v(\Delta; d, \lambda)$ is either constant or strictly monotonically increasing in $\Delta_{\tilde{a}}$. Using [3, Theorem 5.3], we infer that $\pi_v(\Delta_{\tilde{a}}; d, \lambda)$ and $-\pi_v(\Delta_{\tilde{a}}; d, \lambda)$ are f° -pseudoconvex in Δ . With [3, Theorem 5.6], it follows that

$$\max_{(d, \lambda) \in \mathcal{U}} \max_{v \in V} \max \{ \bar{\pi}_v - \pi_v(\Delta_{\tilde{a}}; d, \lambda), \pi_v(\Delta_{\tilde{a}}; d, \lambda) - \bar{\pi}_v \} \quad (21)$$

is f° -pseudoconvex. With this, the constraint (21) and hence (17) fits into our considered generalized convex setting, which exemplary demonstrates that our setting covers relevant real-world applications.

6. CONCLUSION

We proposed in this work a convergent algorithm for mixed-integer nonlinear and non-convex robust optimization problems, for which no general methods exist. We relied on assumptions of generalized convexity with respect to the decision variables. Concerning uncertain parameters and uncertainty sets, we could dispense with any restrictive assumptions such as concavity but relied on the assumption that approximate worst-case evaluations are available. We formulated the resulting robust MINLP as a nonsmooth MINLP with inexactness. For this, we presented an outer approximation approach with an integrated bundle method and proved convergence of the resulting algorithm. This combined method also contributes to MINLP methods, since it applies to generalized convex settings and explicitly addresses the solution of the subproblems.

There are numerous possibilities for further research. Especially, the performance of the proposed method could be evaluated and improved. A first step would be to identify suitable test instances and to evaluate the method in practice. Then, the method's performance could be improved by various strategies such as, e.g., a single-tree approach for the outer approximation, recycling of cutting planes or regularization strategies (see, e.g., [19, 18, 7, 14]).

STATEMENTS AND DECLARATIONS

The author has no conflicts or competing interests to declare that are relevant to the content of this article.

DATA AVAILABILITY STATEMENT

This manuscript has no associated data.

ACKNOWLEDGMENTS

The author thanks Frauke Liers and Michael Stingl for all the helpful discussions on this topic. Further, the author thanks the DFG for their support within Project B06 in CRC TRR 154.

REFERENCES

- [1] Aßmann, D. “Exact Methods for Two-Stage Robust Optimization with Applications in Gas Networks”. PhD thesis. FAU University Press, 2019. DOI: 10.25593/978-3-96147-234-5.
- [2] Aßmann, D., Liers, F., Stingl, M., and Vera, J. C. “Deciding robust feasibility and infeasibility using a set containment approach: an application to stationary passive gas network operations”. In: *SIAM J. Optim.* 28.3 (2018), pp. 2489–2517. ISSN: 1052-6234. DOI: 10.1137/17M112470X.
- [3] Bagirov, A., Karmitsa, N., and Mäkelä, M. M. *Introduction to nonsmooth optimization*. Theory, practice and software. Springer, Cham, 2014, pp. xviii+372. ISBN: 978-3-319-08113-7; 978-3-319-08114-4. DOI: 10.1007/978-3-319-08114-4.
- [4] Ben-Tal, A., Hertog, D. den, and Vial, J.-P. “Deriving robust counterparts of nonlinear uncertain inequalities”. In: *Math. Program.* 149.1-2, Ser. A (2015), pp. 265–299. ISSN: 0025-5610. DOI: 10.1007/s10107-014-0750-8.
- [5] Bonami, P., Cornuéjols, G., Lodi, A., and Margot, F. “A Feasibility Pump for mixed integer nonlinear programs”. In: *Math. Program.* 119.2, Ser. A (2009), pp. 331–352. ISSN: 0025-5610. DOI: 10.1007/s10107-008-0212-2.
- [6] Clarke, F. H. *Optimization and nonsmooth analysis*. Second. Vol. 5. Classics in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1990, pp. xii+308. ISBN: 0-89871-256-4. DOI: 10.1137/1.9781611971309.
- [7] Delfino, A. and Oliveira, W. de. “Outer-approximation algorithms for nonsmooth convex MINLP problems”. In: *Optimization* 67.6 (2018), pp. 797–819. ISSN: 0233-1934. DOI: 10.1080/02331934.2018.1434173.
- [8] Duran, M. A. and Grossmann, I. E. “An outer-approximation algorithm for a class of mixed-integer nonlinear programs”. In: *Math. Program.* 36.3 (1986), pp. 307–339. ISSN: 0025-5610. DOI: 10.1007/BF02592064.
- [9] Eronen, V.-P., Mäkelä, M. M., and Westerlund, T. “On the generalization of ECP and OA methods to nonsmooth convex MINLP problems”. In: *Optimization* 63.7 (2014), pp. 1057–1073. ISSN: 0233-1934. DOI: 10.1080/02331934.2012.712118.
- [10] Eronen, V.-P., Westerlund, T., and Mäkelä, M. M. “On Mixed Integer Nonsmooth Optimization”. In: *Numerical Nonsmooth Optimization: State of the Art Algorithms*. Ed. by A. M. Bagirov, M. Gaudioso, N. Karmitsa, M. M. Mäkelä, and S. Taheri. Springer, Cham, 2020, pp. 549–578. ISBN: 978-3-030-34910-3. DOI: 10.1007/978-3-030-34910-3_16.

- [11] Fletcher, R. and Leyffer, S. “Solving mixed integer nonlinear programs by outer approximation”. In: *Math. Program.* 66.3, Ser. A (1994), pp. 327–349. ISSN: 0025-5610. DOI: 10.1007/BF01581153.
- [12] Hamzei, M. and Luedtke, J. “Linearization-based algorithms for mixed-integer nonlinear programs with convex continuous relaxation”. In: *J. Global Optim.* 59.2-3 (2014), pp. 343–365. ISSN: 0925-5001. DOI: 10.1007/s10898-014-0172-4.
- [13] T. Koch, B. Hiller, M. E. Pfetsch, and L. Schewe, eds. *Evaluating gas network capacities*. Vol. 21. MOS-SIAM Series on Optimization. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2015, pp. xv+364. ISBN: 978-1-611973-68-6. DOI: 10.1137/1.9781611973693.
- [14] Kronqvist, J., Bernal, D. E., and Grossmann, I. E. “Using regularization and second order information in outer approximation for convex MINLP”. In: *Math. Program.* 180.1-2, Ser. A (2020), pp. 285–310. ISSN: 0025-5610. DOI: 10.1007/s10107-018-1356-3.
- [15] Kuchlbauer, M. “Mixed-integer nonlinear robust optimization via nonsmooth methods”. Doctoral thesis. Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), 2023.
- [16] Kuchlbauer, M., Liers, F., and Stingl, M. “Adaptive Bundle Methods for Nonlinear Robust Optimization”. In: *INFORMS Journal on Computing* 34.4 (2022), pp. 2106–2124. DOI: 10.1287/ijoc.2021.1122.
- [17] Kuchlbauer, M., Liers, F., and Stingl, M. “Outer approximation for mixed-integer nonlinear robust optimization”. In: *J. Optim. Theory Appl.* 195.3 (2022), pp. 1056–1086. ISSN: 0022-3239. DOI: 10.1007/s10957-022-02114-y.
- [18] Oliveira, W. de. “Regularized optimization methods for convex MINLP problems”. In: *TOP* 24.3 (2016), pp. 665–692. ISSN: 1134-5764. DOI: 10.1007/s11750-016-0413-4.
- [19] Quesada, I. and Grossmann, I. E. “An LP/NLP based branch and bound algorithm for convex MINLP optimization problems”. In: *Computers & chemical engineering* 16.10-11 (1992), pp. 937–947. DOI: 10.1016/0098-1354(92)80028-8.
- [20] Stangl, C. “Modelle, Strukturen und Algorithmen für stationäre Flüsse in Gasnetzen”. PhD thesis. Universitätsbibliothek Duisburg-Essen, 2014.
- [21] Wei, Z. and Ali, M. M. “Convex mixed integer nonlinear programming problems and an outer approximation algorithm”. In: *J. Global Optim.* 63.2 (2015), pp. 213–227. ISSN: 0925-5001. DOI: 10.1007/s10898-015-0284-5.
- [22] Wei, Z., Ali, M. M., Xu, L., Zeng, B., and Yao, J.-C. “On solving nonsmooth mixed-integer nonlinear programming problems by outer approximation and generalized Benders decomposition”. In: *J. Optim. Theory Appl.* 181.3 (2019), pp. 840–863. ISSN: 0022-3239. DOI: 10.1007/s10957-019-01499-7.