Journal of Optimization Theory and Applications (2019) 181:840–863
https://doi.org/10.1007/s10957-019-01499-7

On Solving Nonsmooth Mixed-Integer Nonlinear Programming Problems by Outer Approximation and Generalized Benders Decomposition

Zhou Wei[1] · M. Montaz Ali[2] · Liang Xu[3] · Bo Zeng[3] · Jen-Chih Yao[4,5]

Received: 23 March 2018 / Accepted: 6 March 2019 / Published online: 19 March 2019
© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract
In this paper, we mainly study nonsmooth mixed-integer nonlinear programming problems and solution algorithms by outer approximation and generalized Benders decomposition. Outer approximation and generalized Benders algorithms are provided to solve these problems with nonsmooth convex functions and with conic constraint, respectively. We illustrate these two algorithms by providing detailed procedure of solving several examples. The numerical examples show that outer approximation and generalized Benders decomposition provide a feasible alternative for solving such problems without differentiability.

✉ Jen-Chih Yao
yaojc@mail.cmu.edu.tw

Zhou Wei
wzhou@ynu.edu.cn

M. Montaz Ali
Montaz.Ali@wits.ac.za

Liang Xu
lix21@pitt.edu

Bo Zeng
bzeng@pitt.edu

1 Department of Mathematics, Yunnan University, Kunming 650091, People's Republic of China

2 School of Computer Science and Applied Mathematics, University of the Witwatersrand, Wits, Johannesburg 2050, South Africa

3 Department of Industrial Engineering, University of Pittsburgh, Pittsburgh, PA 15261, USA

4 Department of Mathematics, Zhejiang Normal University, Jinhua 321004, People's Republic of China

5 Research Center for Interneural Computing, China Medical University Hospital, China Medical University, Taichung 40402, Taiwan

Keywords Mixed-integer nonlinear programming · Outer approximation · Generalized Benders decomposition · Subgradient · Master program

Mathematics Subject Classification 90C11 · 90C25 · 90C30

1 Introduction

Many practical optimization problems have both continuous and integer variables, and these problems are generally modeled as mixed-integer nonlinear programming problems (MINLPs). Such a problem is known as convex MINLP if the objective and constraint functions are convex; otherwise, it is a nonconvex MINLP.

Since MINLPs provide a powerful framework for mathematically modeling many optimization problems, they have been used in various applications such as electricity transmission and electric power system, design of water distribution networks, process industry, and finance and scheduling problems and production planning and control. We refer readers to bibliographies [1–9] for more details on applications. The study on MINLPs as well as solution algorithms has been an active focus of research over the past decades. These algorithms are mainly based on branch-and-bound (B&B), extended cutting plane, Lagrangian relaxation, outer approximation (OA) and generalized Benders decomposition (GBD). Readers are invited to consult references [10–23] for more details on these solution methods.

One focus of this paper is OA. Duran and Grossmann [15] first introduced OA to solve MINLPs in which subproblems are affine for fixed integer variables and convex in continuous variables. Fletcher and Leyffer [16] improved the OA to solve convex MINLPs with continuously differentiable functions and considered the nonsmooth extension in the context of penalty functions therein. Kesavan et al. [17] applied OA to separable nonconvex MINLPs and proposed outer approximation algorithms to solve several relaxed master problems and nonlinear programming problems. Drewes and Ulbrich [24] used OA to solve a mixed-integer second-order cone programming problem with the linear objective function and the affine constraint function. Eronen and Mäkelä [25] applied the generalized OA to nonsmooth convex MINLPs. Based on the above references, we study the OA for one MINLP problem defined by nonsmooth convex functions and use subgradients to reformulate the problem as an equivalent mixed-integer linear program (MILP). An outer approximation algorithm is given for solving this MILP, and several illustrative examples are computed by this algorithm. Compared with aforementioned references, one part of main work in this paper is to show the existence of subgradients that are applicable to the reformulation and provide a complete proof on the equivalence between MINLP and the reformulated problem.

The other focus of this paper is GBD. It is known that Benders [20] first introduced this decomposition for solving mixed-integer linear programming problems and exploiting the structure of complicated problems. Geoffrion [21] generalized this decomposition method to deal with MINLPs where the subproblems are nonlinear programming. Hooker and Ottosson [8] studied logic-based Benders decomposition and applied it to planning and scheduling problems (cf. [6,7] and references therein). As the other part of main work in this paper, we study GBD and a generalized Benders

algorithm for one MINLP with conic constraint. The main idea of this algorithm is given and then several examples as well as their solutions are provided by the generalized Benders algorithm.

The remainder of this paper is organized as follows. In Sect. 2, we give some definitions and preliminaries to be used in this paper. In Sect. 3, we consider OA and the outer approximation algorithm and use this algorithm to solve nonsmooth convex MINLP examples. Section 4 is devoted to the study of GBD and the generalized Benders algorithm. It is shown that several convex MINLP examples with conic constraint can be solved by this algorithm. Section 5 is devoted to the proofs for results in Sects. 3 and 4. Conclusions of this paper are given in Sect. 6.

2 Preliminaries

Let $\|\cdot\|$ be the norm of an Euclidean space $\mathbb{R}^n$ and denote by $\langle\cdot,\cdot\rangle$ the inner product between two elements of $\mathbb{R}^n$. For any subset $C \subset \mathbb{R}^n$, we denote by $\mathrm{int}(C)$ and $\mathrm{cl}(C)$ the interior and the closure of C, respectively.

Let A be a closed convex set of $\mathbb{R}^n$ and $x \in A$. We denote by

$$T(A, x) := \mathrm{cl}(\mathbb{R}_+(A - x))$$

the *contingent cone* of A at x. Thus, $v \in T(A, x)$ if and only if there exist $v_k \to v$ and $t_k \to 0^+$ such that $x + t_k v_k \in A$ for all $k \in \mathbb{N}$, where $\mathbb{N}$ denotes the set of all natural numbers.

We denote by

$$N(A, x) := \{\gamma \in \mathbb{R}^n : \langle\gamma, y - x\rangle \leq 0 \ \text{ for all } y \in A\}$$

the *normal cone* of A at x. It is known that $N(A, x)$ is the dual of $T(A, x)$; that is,

$$N(A, x) = \{\gamma \in \mathbb{R}^n : \langle\gamma, v\rangle \leq 0 \ \text{ for all } v \in T(A, x)\}.$$

Let $\varphi : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a proper lower semicontinuous convex function. We denote by $\mathrm{dom}(\varphi) := \{u \in \mathbb{R}^n : \varphi(u) < +\infty\}$ the *domain* of φ. Let $x \in \mathrm{dom}(\varphi)$. We denote by

$$\partial\varphi(x) := \{\alpha \in \mathbb{R}^n : \langle\alpha, y - x\rangle \leq \varphi(y) - \varphi(x) \ \text{ for all } y \in \mathbb{R}^n\} \tag{1}$$

the *subdifferential* of φ at x. Each vector in $\partial\varphi(x)$ is said to be a *subgradient* of φ at x.

Let $K \subset \mathbb{R}^m$ be a closed and convex cone with nonempty interior. The *partial order* in $\mathbb{R}^m$ by K is defined as follows: for any $z_1, z_2 \in \mathbb{R}^m$,

$$z_1 \preceq_K z_2 \Longleftrightarrow z_2 - z_1 \in K \text{ and } z_1 \prec_K z_2 \Longleftrightarrow z_2 - z_1 \in \mathrm{int}(K). \tag{2}$$

For a function $\psi : \mathbb{R}^n \to \mathbb{R}^m$, recall that ψ is said to be *K-convex*, if

$$\psi(\lambda x_1 + (1-\lambda)x_2) \preceq_K \lambda\psi(x_1) + (1-\lambda)\psi(x_2) \ \ \forall x_1, x_2 \in \mathbb{R}^n \text{ and } \forall \lambda \in [0, 1].$$

For the case when $\psi : \mathbb{R}^n \to \mathbb{R}$ and $K := [0, +\infty[$, K-convexity of ψ reduces to the convexity of real-valued function ψ.

The following lemma, proved in [19, Proposition 2.3], will be used in our analysis.

Lemma 2.1 *Let $\phi : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}$ be a continuous and convex function and suppose that $(\bar{x}, \bar{y}) \in \mathbb{R}^n \times \mathbb{R}^p$. Then for any $\alpha \in \partial\phi(\cdot, \bar{y})(\bar{x})$, there exists $\beta \in \mathbb{R}^p$ such that $(\alpha, \beta) \in \partial\phi(\bar{x}, \bar{y})$.*

We conclude this section with the following lemma (cf. [26, Theorem 2.4.18]).

Lemma 2.2 *Let $\varphi : \mathbb{R}^n \to \mathbb{R}$ be a continuous convex function. Define*

$$\varphi_+(x) := \max\{\varphi(x), 0\}, \quad \forall x \in \mathbb{R}^n.$$

Then φ_+ is a convex and continuous function and

$$\partial\varphi_+(x) = [0, 1]\partial\varphi(x)$$

holds for all $x \in \mathbb{R}^n$ with $\varphi(x) = 0$, where $[0, 1]\partial\varphi(x) := \{t\gamma : t \in [0, 1]$ and $\gamma \in \partial\varphi(x)\}$ for any $x \in \mathbb{R}^n$.

3 OA for MINLPs with Nonsmooth Convex Functions

In this section, we study OA and an outer approximation algorithm for MINLPs with nonsmooth convex functions. To illustrate this algorithm, we compute nonsmooth convex MINLP examples with detailed computing procedure, including the choice of nonnegative multipliers and subgradients that satisfy the KKT conditions and the computation of normal cone.

3.1 Overview of an Outer Approximation Algorithm

This subsection is devoted to an overview of OA for MINLPs with nonsmooth convex functions. The main idea of this method is to reformulate MINLP as an equivalent mixed-integer linear programming (MILP) problem and then construct an outer approximation algorithm for solving a sequence of relaxed MILPs using several solvers or softwares.

Suppose $f, g_i : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}$ $(i = 1, \ldots, m)$ are continuous convex functions (not necessarily differentiable), $X \subset \mathbb{R}^n$ is a nonempty compact convex set and $Y \subset \mathbb{Z}^p$ is a set of integers. A generalized nonsmooth MINLP is generally defined as follows:

$$(\mathrm{P})\begin{cases}\min\limits_{x,y} & f(x,y)\\ \text{s.t.} & g_i(x,y)\le 0, i=1,\dots,m,\\ & x\in X, y\in Y \text{ integer}.\end{cases} \tag{3}$$

For any fixed $y \in Y$, we consider the following subproblem $P(y)$:

$$P(y)\begin{cases}\min\limits_{x} & f(x,y)\\ \text{s.t.} & g_i(x,y)\le 0, i=1,\dots,m,\\ & x\in X.\end{cases}$$

We divide Y into two disjoint subsets:

$$T := \{y \in Y : P(y) \text{ is feasible}\} \text{ and } S := \{y \in Y : P(y) \text{ is infeasible}\}. \tag{4}$$

We assume that problem (P) satisfies Slater constraint qualification:

(A1) For any $y \in T$, there is $\hat{x} \in X$ satisfying $g_i(\hat{x}, y) < 0$ for all $i = 1, \dots, m$.

Let $y \in Y$ be given. Then exactly one of following two cases occurs:

Case 1 $y \in T$. We assume that $y = y_j$ for some $y_j \in T$.

Solve subproblem $P(y_j)$ and obtain an optimal solution $x_j \in X$. By assumption (A1) and KKT conditions (cf. [27]), there exist $\lambda_{j,1}, \dots, \lambda_{j,m} \ge 0$ such that

$$\sum_{i=1}^{m}\lambda_{j,i}g_i(x_j,y_j)=0 \text{ and } 0\in\partial f(\cdot,y_j)(x_j)+\sum_{i=1}^{m}\lambda_{j,i}\partial g_i(\cdot,y_j)(x_j)+N(X,x_j). \tag{5}$$

This and Lemma 2.1 imply that there exist $\alpha_j \in \partial f(\cdot, y_j)(x_j)$, $\xi_{j,i} \in \partial g_i(\cdot, y_j)(x_j)$ and $\beta_j \in \mathbb{R}^p$, $\eta_{j,i} \in \mathbb{R}^p$ such that $(\alpha_j, \beta_j) \in \partial f(x_j, y_j)$, $(\xi_{j,i}, \eta_{j,i}) \in \partial g_i(x_j, y_j)$ $(i = 1, \dots, m)$ and

$$-\alpha_j-\sum_{i=1}^{m}\lambda_{j,i}\xi_{j,i}\in N(X,x_j). \tag{6}$$

Case 2 $y \notin T$. We assume that $y = y_l$ for some $y_l \in S$.

Let J_l be one subset of $\{1, \dots, m\}$ such that there is some $\tilde{x} \in X$ satisfying

$$g_i(\tilde{x}, y_l) < 0, \quad \forall i \in J_l.$$

Denote by $I_l := \{1, \dots, m\}\backslash J_l$ the complement of J_l and consider the following subproblem $F(y_l)$:

$$F(y_l)\begin{cases}\min\limits_{x} & \sum\limits_{i\in I_l}[g_i(x,y_l)]_+ := \min\limits_{x}\sum\limits_{i\in I_l}\max\{g_i(x,y_l),0\}\\ \text{s.t.} & g_i(x,y_l)\le 0 \ \ \forall i\in J_l,\\ & x\in X.\end{cases} \tag{7}$$

Solve subproblem $F(y_l)$ and obtain the optimal solution x_l. By KKT conditions, there exist $\lambda_{l,i} \geq 0$ for all $i \in J_l$ such that

$$\sum_{i\in J_l} \lambda_{j,i} g_i(x_j, y_j) = 0 \text{ and } 0 \in \sum_{i\in I_l} \partial[g_i(\cdot, y_l)]_+(x_l) + \sum_{i\in J_l} \lambda_{l,i} \partial g_i(\cdot, y_l)(x_l) + N(X, x_l). \tag{8}$$

This together with Lemmas 2.1 and 2.2 implies that there exist $\lambda_{l,i} \in [0, 1]$ for all $i \in I_l$ and $\xi_{l,i} \in \partial g_i(\cdot, y_l)(x_l)$, $\eta_{l,i} \in \mathbb{R}^p$ $(\forall i \in I_l \cup J_l)$ such that $(\xi_{l,i}, \eta_{l,i}) \in \partial g_i(x_l, y_l)$ for all $i \in I_l \cup J_l = \{1, \ldots, m\}$ and

$$\begin{aligned} &-\sum_{i\in I_l\cup J_l} \lambda_{l,i}\xi_{l,i} \in N(X, x_l),\\ &\lambda_{l,i} = 0, \text{ if } i \in I_l \text{ with } g_i(x_l, y_l) < 0,\\ &\lambda_{l,i} = 1, \text{ if } i \in I_l \text{ with } g_i(x_l, y_l) > 0. \end{aligned} \tag{9}$$

The following proposition provides constraints to exclude integers producing infeasible subproblems. The proof is given in Sect. 5.

Proposition 3.1 *Consider the following constraint:*

$$\begin{aligned} &g_i(x_l, y_l) + (\xi_{l,i}, \eta_{l,i})^T \begin{pmatrix} x - x_l \\ y - y_l \end{pmatrix} \leq 0, i = 1, \ldots, m,\\ &x \in X, y \in Y. \end{aligned} \tag{10}$$

Then for any $(\hat{x}, \hat{y}) \in X \times Y$ feasible to the constraints in (10), one has $\hat{y} \neq y_l$.

By virtue of subgradients given by (5),(6) (8) and (9), one can establish the following MILP master program (MP):

$$\text{(MP)} \begin{cases} \min\limits_{x,y,\theta} \ \theta \\ \text{s.t. } f(x_j, y_j) + (\alpha_j, \beta_j)^T \begin{pmatrix} x - x_j \\ y - y_j \end{pmatrix} \leq \theta \quad \forall y_j \in T, \\ \quad g_i(x_j, y_j) + (\xi_{j,i}, \eta_{j,i})^T \begin{pmatrix} x - x_j \\ y - y_j \end{pmatrix} \leq 0 \ \forall y_j \in T, i = 1, \ldots, m, \\ \quad g_i(x_l, y_l) + (\xi_{l,i}, \eta_{l,i})^T \begin{pmatrix} x - x_l \\ y - y_l \end{pmatrix} \leq 0 \ \forall y_l \in S, i = 1, \ldots, m, \\ \quad x \in X, y \in Y \text{ integer}. \end{cases} \tag{11}$$

The following proposition shows the equivalence between the master program (MP) and the MINLP problem (P). The reader is invited to consult Sect. 5 for the proof.

Proposition 3.2 *The master program (MP) of (11) is equivalent to problem (P) in the sense that both problems have the same optimal value and the optimal solution $(\bar{x}, \bar{y})$ to problem (P) corresponds to the optimal solution $(\bar{x}, \bar{y}, \bar{\theta})$ to (MP) with $\bar{\theta} = f(\bar{x}, \bar{y})$.*

Remark 3.1 The KKT conditions play an important role in the equivalent reformulation of MINLP problem (P), and Example 3.2 shows that the subgradients, if not satisfying KKT conditions, may be insufficient for this equivalent reformulation. □

Based on Proposition 3.2, the outer approximation algorithm can be constructed to solve a relaxed MP of (11).

At iteration k, subsets T^k and S^k are defined as follows:

$$T^k := T \cap \{y_1, \ldots, y_k\} \text{ and } S^k := S \cap \{y_1, \ldots, y_k\}.$$

Exactly one of the following cases occurs:

(a) $y_k \in T^k$: Solve subproblem $P(y_k)$ and denote the optimal solution by x_k. Solve the following optimization problem $OP(x_k, y_k)$:

$$\textit{Find} \begin{bmatrix} (\alpha_k, \beta_k) \\ (\xi_{k,i}, \eta_{k,i}) \\ \lambda_{k,i} \in \mathbb{R} \end{bmatrix} \textit{such that} \begin{cases} -\alpha_k - \sum\limits_{i=1}^{m} \lambda_{k,i}\xi_{k,i} \in N(X, x_k), \\ (\alpha_k, \beta_k) \in \partial f(x_k, y_k), \\ (\xi_{k,i}, \eta_{k,i}) \in \partial g_i(x_k, y_k), i = 1, \ldots, m, \\ \lambda_{k,i} \geq 0, i = 1, \ldots, m. \end{cases} \tag{12}$$

Obtain subgradients (α_k, β_k) and $(\xi_{k,i}, \eta_{k,i})$ for all $i = 1, \ldots, m$.

(b) $y_k \in S^k$: Take $I_k, J_k \subset \{1, \ldots, m\}$ and create subproblem $F(y_k)$ as in (7). Solve $F(y_k)$ and denote the solution by x_k. Solve the following optimization problem *Sub-OP*(x_k, y_k):

$$\textit{Find} \begin{bmatrix} (\xi_{k,i}, \eta_{k,i}) \\ \lambda_{k,i} \in \mathbb{R} \end{bmatrix} \textit{such that} \begin{cases} -\sum\limits_{i=1}^{m} \lambda_{k,i}\xi_{k,i} \in N(X, x_k), \\ (\xi_{k,i}, \eta_{k,i}) \in \partial g_i(x_k, y_k), \ i = 1, \ldots, m, \\ \lambda_{k,i} \geq 0, \ i = 1, \ldots, m, \\ \lambda_{k,i} = 0, \quad i \in I_k \textit{ with } g_i(x_k, y_k) < 0, \\ \lambda_{k,i} = 1, \quad i \in I_k \textit{ with } g_i(x_k, y_k) > 0, \\ \lambda_{k,i} \in [0, 1], i \in I_k \textit{ with } g_i(x_k, y_k) = 0. \end{cases} \tag{13}$$

Obtain subgradients $(\xi_{k,i}, \eta_{k,i})$ for all $i = 1, \ldots, m$.

Let $UBD^k := \min\{f(x_j, y_j) : y_j \in T^k\}$. Consider the following relaxed master program MP^k:

$$MP^k \begin{cases} \min\limits_{x,y,\theta} \ \theta \\ \text{s.t.} \ \ \theta < UBD^k \\ \qquad f(x_j, y_j) + (\alpha_j, \beta_j)^T \begin{pmatrix} x - x_j \\ y - y_j \end{pmatrix} \leq \theta \ \ \forall y_j \in T^k, \\ \qquad g_i(x_j, y_j) + (\xi_{j,i}, \eta_{j,i})^T \begin{pmatrix} x - x_j \\ y - y_j \end{pmatrix} \leq 0 \ \ \forall y_j \in T^k, i = 1, \ldots, m, \\ \qquad g_i(x_l, y_l) + (\xi_{l,i}, \eta_{l,i})^T \begin{pmatrix} x - x_l \\ y - y_l \end{pmatrix} \leq 0 \ \ \forall y_l \in S^k, i = 1, \ldots, m, \\ \qquad x \in X, y \in Y \text{ integer}. \end{cases} \tag{14}$$

The outer approximation algorithm for solving the relaxed master problems MP^k is stated as follow.

Algorithm 1 (Outer Approximation Algorithm)

1: Initialization. Given an initial $y_0 \in Y$, set $T^0 = S^0 := \emptyset$, $UBD^0 := \infty$ and let $k := 1$
2: **for** $k = 1, 2, \cdots$, **do**
3: Check subproblem $P(y_k)$
4: **if** $y_k \in T$ is feasible **then**
5: Solve subproblem $P(y_k)$ and denote the solution by x_k
Solve $OP(x_k, y_k)$ in (12) and obtain subgradients (α_k, β_k), $(\xi_{k,i}, \eta_{k,i})$
Set $T^k := T^{k-1} \cup \{y_k\}$, $S^k := S^{k-1}$ and $UBD^k := \min\{UBD^{k-1}, f(x_k, y_k)\}$
6: **else**
7: Create subproblem $F(y_k)$. Solve $F(y_k)$ and denote the solution by x_k
Solve *Sub-OP*(x_k, y_k) in (13) and obtain subgradients $(\xi_{k,i}, \eta_{k,i})$
Set $S^k := S^{k-1} \cup \{y_k\}$, $T^k := T^{k-1}$ and $UBD^k := UBD^{k-1}$
8: **end if**
9: Solve the relaxation MP^k and obtain a new integer y_{k+1}
Set $k := k + 1$ and go back to line 3
10: **end for**

The following proposition is on the convergence of Algorithm 1. The proof is given in Sect. 5.

Proposition 3.3 *Suppose that the cardinality of Y is finite. Then either problem (P) is infeasible or Algorithm 1 terminates in a finite number of steps at an optimal value of problem (P).*

3.2 Illustrative Examples

In this subsection, Algorithm 1 is tested by numerical examples and the results show that Algorithm 1 is feasible and applicable to solve MINLPs with nonsmooth convex functions. All computational tests for these examples are implemented in MATLAB version R2014a and solved by BARON version 16.10.6 (cf. [12]).

Example 3.1 Let $X = \{(u, v) : -1 \le u + v \le 1 \text{ and } -1 \le u - v \le 1\}$ and $Y = \{-1, 1, 2\}$. Consider the following convex MINLP:

$$\begin{aligned} \min_{x,y} \ & f(x, y) = \max\{-u + y, v^2 + y - 1\} \\ \text{s.t.} \ & g_1(x, y) = \max\{u - y, v + y - 1\} \le 0, \\ & g_2(x, y) = \max\{2u - 4 + y, v - y\} \le 0, \\ & x = (u, v) \in X, y \in Y. \end{aligned} \tag{15}$$

Choose the tolerance $\varepsilon = 0.005$. Let $y_1 = 2$ be the initial point.

At the first iteration, check subproblem $P(y_1)$ that is feasible. Solve subproblem $P(y_1)$ and denote the optimal solution by $x_1 = (0, -1)$. Solve optimization problem

$OP(x_1, y_1)$ as in (12) and obtain multipliers $\lambda_{1,1} = 2, \lambda_{1,2} = 0.5$ and subgradients $(\alpha_1, \beta_1) = ((-1, 0), 1)$, $(\xi_{1,1}, \eta_{1,1}) = ((0, 1), 1)$, $(\xi_{1,2}, \eta_{1,2}) = ((2, 0), 1)$. Solve the following relaxed MILP problem:

$$MP^1 \begin{cases} \min\limits_{x,y,\theta} \ \theta \\ \text{s.t. } \theta \le f(x_1, y_1) - \varepsilon, \\ \qquad f(x_1, y_1) + (\alpha_1, \beta_1)^T \begin{pmatrix} x - x_1 \\ y - y_1 \end{pmatrix} \le \theta, \\ \qquad g_i(x_1, y_1) + (\xi_{1,i}, \eta_{1,i})^T \begin{pmatrix} x - x_1 \\ y - y_1 \end{pmatrix} \le 0, i = 1, 2, \\ \qquad x \in X, y \in Y. \end{cases}$$

Denote the optimal solution by $(\hat{x}, \hat{y}, \hat{\theta}) = ((0, 1), -1, -2)$.

In the second iteration, we set $y_2 = \hat{y} = -1$ and check subproblem $P(y_2)$ by Algorithm 1; $P(y_2)$ is infeasible. Consider the following subproblem $F(y_2)$:

$$F(y_2) \begin{cases} \min\limits_{x} \ \max\{g_1(x, y_2), 0\} \\ \text{s.t. } g_2(x, y_2) = \max\{2u - 5, v + 1\} \le 0, \\ \qquad x = (u, v) \in X. \end{cases}$$

Solve subproblem $F(y_2)$ and denote the optimal solution by $x_2 = (0, -1)$. Solve optimization problem *Sub-OP*(x_2, y_2) as in (13) and obtain subgradients $(\xi_{2,1}, \eta_{2,1}) = ((1, 0), -1)$ and $(\xi_{2,2}, \eta_{2,2}) = ((0, -1), -1)$. According to Algorithm 1, consider the following constraints:

$$g_i(x_2, y_2) + (\xi_{2,i}, \eta_{2,i})^T \begin{pmatrix} x - x_2 \\ y - y_2 \end{pmatrix} \le 0, i = 1, 2.$$

Add these two constraints into MP^1 and obtain relaxed MILP problem MP^2. Solve MP^2 and denote the optimal solution by $(\hat{x}, \hat{y}, \hat{\theta}) = ((-1, 0), 1, 0)$. The second iteration is concluded.

In the third iteration, we set $y_3 = \hat{y} = 1$ and check subproblem $P(y_3)$, which is feasible. Solve subproblem $P(y_3)$ and denote the optimal solution by $x_3 = (1, 0)$. Solve optimization problem $OP(x_3, y_3)$ as in (12) and obtain multipliers $\lambda_{3,1} = \lambda_{3,2} = 0$ and subgradients $(\alpha_3, \beta_3) = ((0, 0), 1)$, $(\xi_{3,1}, \eta_{3,1}) = ((1, 0), -1)$, $(\xi_{3,2}, \eta_{3,2}) = ((0, 1), -1)$.

By Algorithm 1, consider the following constraints:

$$\begin{aligned} &\theta \le f(x_3, y_3) - \varepsilon, \\ &f(x_3, y_3) + (\alpha_3, \beta_3)^T \begin{pmatrix} x - x_3 \\ y - y_3 \end{pmatrix} \le \theta, \\ &g_i(x_3, y_3) + (\xi_{3,i}, \eta_{3,i})^T \begin{pmatrix} x - x_3 \\ y - y_3 \end{pmatrix} \le 0, i = 1, 2. \end{aligned}$$

Add these four constraints into MP^2 and get relaxed MILP problem MP^3. As MP^3 is infeasible, the algorithm stops and an ε-optimal solution is obtained. □

The following example shows that the chosen subgradients, if not satisfying KKT conditions, may be not sufficient for solving nonsmooth convex MINLPs.

Example 3.2 Let $X = [0, 2]$ and $Y = \{1, 2, 3\}$. Consider the following convex MINLP:

$$\text{(P)} \begin{cases} \min\limits_{x,y} \ f(x, y) = x + y \\ \text{s.t.} \ g_1(x, y) = \max\{-x + y + 1, x - y + 1\} \le 0, \\ \qquad g_2(x, y) = x - y \le 0, \\ \qquad x \in X, y \in Y. \end{cases} \tag{16}$$

It is not hard to verify that (P) of (16) is infeasible.

Let $y_1 = 1$ be the initial point and check subproblem $P(y_1)$. As $P(y_1)$ is infeasible, we consider the following subproblem $F(y_1)$:

$$F(y_1) \begin{cases} \min\limits_{x} \ \max\{g_1(x, 1), 0\} \\ \text{s.t.} \ g_2(x, y_1) = x - 1 \le 0, \\ \qquad x \in X. \end{cases}$$

Solve subproblem $F(y_1)$ and denote the optimal solution by $x_1 = 1$. Solve optimization problem $Sub - OP(x_1, y_1)$ as in (12) and obtain subgradients $\xi_{1,1} = -1$, $\xi_{1,2} = 1$, multiplier $\lambda_{1,2} = 1$ and numbers $\eta_{1,1} = 1$, $\eta_{1,2} = -1$ such that

$$\xi_{1,1} + \lambda_{1,2}\xi_{1,2} \in N(X, x_1) \ \text{ and } \ (\xi_{1,i}, \eta_{1,i}) \in \partial g_i(x_1, y_1), i = 1, 2.$$

According to Algorithm 1, consider the following constraints:

$$g_1(x_1, y_1) + (\xi_{1,1}, \eta_{1,2})^T \begin{pmatrix} x - x_1 \\ y - y_1 \end{pmatrix} \le 0,$$
$$g_2(x_1, y_1) + (\xi_{2,1}, \eta_{2,2})^T \begin{pmatrix} x - x_1 \\ y - y_1 \end{pmatrix} \le 0.$$

Solve this constraint problem and then the problem is infeasible which shows that problem (P) of (16) is infeasible.

However, if taking

$$(\xi^*_{1,1}, \eta^*_{1,1}) = (1, -1) \in \partial g_1(x_1, y_1) \ \text{ and } \ (\xi^*_{1,2}, \eta^*_{1,2}) = (1, -1) \in \partial g_2(x_1, y_1),$$

then one can verify that these two subgradients do not satisfy KKT conditions; that is,

$$\nexists \lambda_{1,2} \ge 0 \ \text{ satisfying } \ \xi^*_{1,1} + \lambda_{1,2}\xi^*_{1,2} \in N(X, x_1) = \{0\}.$$

Consider the following constraints:

$$g_1(x_1, y_1) + (\xi^*_{1,1}, \eta^*_{1,2})^T \begin{pmatrix} x - x_1 \\ y - y_1 \end{pmatrix} \leq 0,$$
$$g_2(x_1, y_1) + (\xi^*_{2,1}, \eta^*_{2,2})^T \begin{pmatrix} x - x_1 \\ y - y_1 \end{pmatrix} \leq 0,$$

which has a feasible point $(\hat{x}, \hat{y}) = (0, 1)$. This means that OA for this MINLP may generate an infinite loop between point (x_1, y_1) and point $(0, 1)$. Thus OA is invalid for MINLP problem of (16). □

4 GBD for MINLPs with Conic Constraint

In this section, we study GBD for solving MINLPs with the vector conic constraint. Several MINLP examples with vector conic constraint are solved, and the computing procedure which includes the choice of optimal Lagrange multipliers of subproblems is also provided.

4.1 Overview of the Generalized Benders Algorithm

This subsection is devoted to an overview of the generalized Benders algorithm for solving MINLPs with vector conic constraint. The main idea of this algorithm is to separate the original MINLP into many subproblems, then reformulate MINLP as an equivalent master program by optimal Lagrange multipliers (of subproblems) and finally construct an algorithm for solving a sequence of relaxed master program problems.

Let $X \subset \mathbb{R}^n$ be a closed bounded convex set, $Y \subset \mathbb{Z}^p$ be a set with integers and $K \subset \mathbb{R}^m$ be a closed convex cone with a nonempty interior which specifies a partial order $\preceq_K$ on $\mathbb{R}^m$ as (2). The MINLP problem is defined as follows:

$$\text{(VOP)} \begin{cases} \min\limits_{x,\, y} \ f(x, y) \\ \text{s.t.} \ \ g(x, y) \preceq_K 0, \\ \qquad x \in X,\, y \in Y \text{ integer}, \end{cases} \tag{17}$$

where $f : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}$ and $g : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^m$ satisfy that $f(\cdot, y)$ is convex and $g(\cdot, y)$ is K-convex on X for any fixed $y \in Y$.

We denote by

$$K^+ := \{z^* \in \mathbb{R}^m : \langle z^*, z\rangle \geq 0 \text{ for all } z \in K\} \tag{18}$$

the dual cone of K.

Let $y \in Y$. We consider the following subproblem $P(y)$:

$$P(y) \begin{cases} \min\limits_{x} \ f(x, y) \\ \text{s.t.} \ \ g(x, y) \preceq_K 0, \\ \qquad x \in X. \end{cases}$$

Recall that the perturbation function $v_y(\cdot)$ associated with subproblem $P(y)$ is defined by

$$v_y(z) : \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}, z \mapsto v_y(z) := \inf_{x \in X} \{f(x, y) : g(x, y) \preceq_K z\}, \tag{19}$$

here we use the convention that infimum over the empty set is $+\infty$.

Recall that a linear continuous functional $\bar{u}^* \in \mathbb{R}^m$ is said to be an optimal Lagrange multiplier for problem $P(y)$ if there exists $\bar{x} \in X$ such that $(\bar{x}, \bar{u}^*)$ satisfies the following optimality conditions:

$$\begin{aligned} &\text{(i) } f(\bar{x}, y) + \langle \bar{u}^*, g(\bar{x}, y)\rangle = \min_{x \in X} \left\{ f(x, y) + \langle \bar{u}^*, g(x, y)\rangle \right\}, \\ &\text{(ii) } \langle \bar{u}^*, g(\bar{x}, y)\rangle = 0, \\ &\text{(iii) } \bar{u}^* \in K^+, \\ &\text{(iv) } g(\bar{x}, y) \preceq_K 0. \end{aligned} \tag{20}$$

The following condition is assumed to hold for (VOP) in (17):

(A2) Suppose that functions $f(\cdot, y), g(\cdot, y)$ are continuous on X for any $y \in Y$ and the Slater constraint qualification

$$g(\hat{x}, y) \prec_K 0 \ \textit{for some} \ \hat{x} \in X \tag{21}$$

holds for any $y \in Y$ such that subproblem $P(y)$ is feasible.

The following proposition establishes properties of the perturbation function $v_y(\cdot)$ and of the optimal Lagrange multipliers for subproblem $P(y)$. The proof is given in Sect. 5.

Proposition 4.1 *The following statements hold:*

(i) *For any $y \in Y$, perturbation function $v_y(\cdot)$ is convex, lower semicontinuous and monotone nonincreasing (by $\preceq_K$) on $\mathbb{R}^m$, and dom(v_y) is convex and closed.*
(ii) *For any $y \in Y$ satisfying subproblem $P(y)$ is feasible, $v_y(\cdot)$ is continuous at $0 = (0, \ldots, 0) \in \mathbb{R}^m$ and $-\partial v_y(0)$ is the set of all optimal Lagrange multipliers of subproblem $P(y)$.*

The following proposition is a key result for reformulating MINLP (VOP) as an equivalent master program. The reader is invited to consult Sect. 5 for the proof.

Proposition 4.2 *The following statements hold:*

(i) *Let $y \in Y$. Then subproblem $P(y)$ is feasible if and only if*

$$\inf_{x \in X} \langle u^*, g(x, y) \rangle \leq 0 \ \ \forall u^* \in K^+ \text{ with } \|u^*\| = 1. \tag{22}$$

(ii) *Let $y \in Y$ be such that subproblem $P(y)$ is feasible. Then*

$$v_y(0) = \max_{u^* \in K^+} \left\{ \inf_{x \in X} \left\{ f(x, y) + \langle u^*, g(x, y) \rangle \right\} \right\}. \tag{23}$$

Based on Proposition 4.2, we have the following master program (MP) which is equivalent to MINLP problem (VOP) in (17):

$$\text{(MP)} \begin{cases} \min\limits_{y,\theta} \ \theta \\ \text{s.t} \ \ \inf\limits_{x \in X} \{ f(x, y) + \langle u^*, g(x, y) \rangle \} \leq \theta, \ \ \forall u^* \in K^+ \\ \quad \inf\limits_{x \in X} \langle z^*, g(x, y) \rangle \leq 0, \ \ \forall z^* \in K^+ \text{ with } \|z^*\| = 1, \\ \quad y \in Y, \ \theta \in \mathbb{R}. \end{cases} \tag{24}$$

To solve the equivalent master program (MP) in (24), we consider the following relaxation of (MP) and optimization problems $L\text{-}P(y_k)$ and $In\text{-}P(y_k)$ referring to subproblem $P(y_k)$.

For any subsets $T \subset K^+$, $S \subset K^+$, the relaxed master program RMP(T, S) *is defined as follows:*

$$\text{RMP}(T, S) \begin{cases} \min\limits_{y,\theta} \ \theta \\ \text{s.t.} \ \inf\limits_{x \in X} \{ f(x, y) + \langle u^*, g(x, y) \rangle \} \leq \theta, \ \ \forall u^* \in T, \\ \quad \inf\limits_{x \in X} \langle z^*, g(x, y) \rangle \leq 0, \ \ \forall z^* \in S, \\ \quad y \in Y, \ \theta \in \mathbb{R}. \end{cases} \tag{25}$$

For any $y_k \in Y$ such that $P(y_k)$ is feasible, optimization problem $L - P(y_k)$ is defined as follows:

$$\textit{Find} \begin{bmatrix} x_k \in X \\ u_k^* \in \mathbb{R}^m \end{bmatrix} \ \textit{such that} \ \begin{cases} x_k \in \underset{x \in X}{argmin} \{ f(x, y_k) + \langle u_k^*, g(x, y_k) \rangle \}, \\ \langle u_k^*, g(x_k, y_k) \rangle = 0, \\ u_k^* \in K^+, \\ g(x_k, y_k) \preceq_K 0. \end{cases} \tag{26}$$

For any $y_k \in Y$ such that $P(y_k)$ is infeasible, optimization problem In-$P(y_k)$ is defined as follows:

$$\textit{Find} \ \left[z_k^* \in \mathbb{R}^m \right] \ \textit{such that} \ \begin{cases} \inf\limits_{x \in X} \langle z_k^*, g(x, y_k) \rangle > 0, \\ z_k^* \in K^+, \\ \|z_k^*\| = 1. \end{cases} \tag{27}$$

The generalized Benders algorithm is stated as follows. This algorithm can be used to find the optimum of MINLP problem (VOP).

Algorithm 2 (Generalized Benders Algorithm)

1: Initialization. Take $y_1 \in Y$ such that $P(y_1)$ is feasible and $z_1^* \in K^+$ with $\|z_1^*\| = 1$
Solve $L\text{-}P(y_1)$ as in (26) and obtain the optimal Lagrange multiplier u_1^*
Set $T^1 := \{u_1^*\}$, $S^1 := \{z_1^*\}$ and $\text{UBD}^1 := v_{y_1}(0)$
Select the convergence tolerance parameter $\varepsilon > 0$ and set $k := 1$
2: **for** $k = 1, 2, \cdots,$ **do**
3: Solve RMP(T^k, S^k) and denote the optimal solution by (y_{k+1}, θ_{k+1})
4: **if** $\text{UBD}^k \le \theta_{k+1} + \varepsilon$ **then**
5: Stop the Algorithm
6: **else**
7: Check subproblem $P(y_{k+1})$
8: **if** $P(y_{k+1})$ is feasible **then**
9: **if** $v_{y_{k+1}}(0) \le \theta_{k+1} + \varepsilon$ **then**
10: Stop the Algorithm
11: **else**
12: Solve subproblem $L\text{-}P(y_{k+1})$ as in (26) and obtain u_{k+1}^*
Set $T^{k+1} := T^k \cup \{u_{k+1}^*\}$, $S^{k+1} := S^k$, $\text{UBD}^{k+1} := \min\{\text{UBD}^k, v_{y_{k+1}}(0)\}$
13: **end if**
14: **else**
15: Solve subproblem $In\text{-}P(x_{k+1})$ as in (27) and obtain z_{k+1}^*
Set $T^{k+1} := T^k$, $S^{k+1} := S^k \cup \{z_{k+1}^*\}$, $\text{UBD}^{k+1} := \text{UBD}^k$
16: **end if**
17: **end if**
18: Set $k := k + 1$ and go back to line 3
19: **end for**

The following proposition shows the termination of Algorithm 2 after a finite number of steps. The detailed proof is presented in Sect. 5.

Proposition 4.3 *Suppose that the cardinality of Y is finite. Then for any given $\varepsilon > 0$, Algorithm 2 terminates in a finite number of steps.*

4.2 Illustrative Examples

In this subsection, MINLP examples are solved by Algorithm 2. The algorithm is implemented in MATLAB version R2014a, and examples are all solved by BARON version 16.10.6 (cf. [12]). The implementation of Algorithm 2 also demonstrates that GBD is feasible and applicable to solve MINLPs with conic constraint.

Example 4.1 Let $K := \{(u, v) \in \mathbb{R}^2 : u \ge 2|v|, u \ge 0\}$, $Y = \{-1, 0, 1, 2\}$ and $X = \{(u, v) \in \mathbb{R}^2 : -1 \le u + v \le 1 \textit{ and } -1 \le u - v \le 1\}$. Consider the following MINLP with vector conic constraint:

$$\begin{aligned} \min_{x,y} \ & f(x,y) \\ \text{s.t.} \ & g(x,y) \preceq_K 0, \\ & x = (u,v) \in X, \\ & y \in Y, \end{aligned} \tag{28}$$

where $f(x,y) = \max\{-u+y, v^2+y-1\}$, $g(x,y) = (g_1(x,y), g_2(x,y))$ with

$$g_1(x,y) = u^2 + v^2 - \frac{1}{2}y \ \ \textit{and} \ \ g_2(x,y) = -\frac{1}{2}v^2 + \frac{1}{4}y.$$

It is easy to verify that $g(\cdot, y)$ is K-convex. We solve MINLP in (28) by Algorithm 2.

Choose the tolerance $\varepsilon = 0.005$. Let $y_1 = -1$ be the initial point and $z_1^* = (1,0) \in K^+$ with $\|z_1^*\| = 1$. Solve optimization problem $L\text{-}P(y_1)$ as in (26) and obtain an optimal Lagrange multiplier $u_1^* = (0.991, 0.01)$. Set $T^1 = \{u_1^*\}$, $S^1 = \{z_1^*\}$ and $\text{UBD}^1 = v_{y_1}(0) = 2$. Solve the relaxed master program $\text{RMP}(T^1, S^1)$ and denote the optimal solution by $(y_2, \theta_2) = (1, 0.2547)$. Since the termination criterion is not satisfied, we check subproblem $P(y_2)$ which is feasible and obtain optimal Lagrange multiplier $u_2^* = (0,0)$ by solving $L-P(y_2)$ as in (26). Set $T^2 = T^1 \cup \{u_2^*\}$, $S^2 = S^1$ and $\text{UBD}^2 = \min\{\text{UBD}^1, v_{y_2}(0)\} = 0$. Solve the relaxed master program $\text{RMP}(T^2, S^2)$ and denote the optimal solution by $(y_3, \theta_3) = (1, 0)$. Since $\text{UBD}^2 \leq \theta_3 + \varepsilon$, it follows that the termination criterion is satisfied and the algorithm stops. This means that $(x_3, y_3) = ((1,0), 1)$ is an ε-tolerance optimal solution of problem (28). □

Example 4.2 Let $K := \{(u,v) \in \mathbb{R}^2 : -u \geq 2|v|, u \leq 0\}$, $Y = \{-2, -1, 0, 1\}$ and $X = \{(u,v) : u^2 + v^2 \leq 2\}$. Consider the following MINLP with vector conic constraint:

$$\begin{aligned} \min_{x,y} \ & f(x,y) \\ \text{s.t.} \ & g(x,y) \preceq_K 0, \\ & x = (u,v) \in X, \\ & y \in Y. \end{aligned} \tag{29}$$

where $f(x,y) = -2u + v + y$, $g(x,y) = (g_1(x,y), g_2(x,y))$ with

$$g_1(x,y) = -u^2 - v^2 + y \ \ \textit{and} \ \ g_2(x,y) = -\frac{1}{2}u^2 + \frac{1}{4}v^2 + \frac{1}{2}y.$$

Choose the tolerance $\varepsilon = 0.005$. Let $y_1 = 1$ be the initial point and $z_1^* = (-1, 0) \in K^+$ with $\|z_1^*\| = 1$. Solve optimization problem $L\text{-}P(y_1)$ as in (26) and obtain the optimal Lagrange multiplier $u_1^* = (-0.001, 0.002)$. Set $T^1 = \{u_1^*\}$, $S^1 = \{z_1^*\}$ and $\text{UBD}^1 = v_{y_1}(0) = -1.001$. Solve the relaxed master program $\text{RMP}(T^1, S^1)$ and denote the optimal solution by $(y_2, \theta_2) = (-2, -5.1697)$. Noting that the termination criterion is not satisfied, we check subproblem $P(y_2)$ (that is feasible) and get the optimal value $v_{y_2}(0) = -5.1623$ (of $P(y_2)$). This implies that the termination criterion is

satisfied and then the algorithm is stopped. Thus $(x_2, y_2) = ((1.2649, -0.6325), -2)$ is an ε-tolerance optimal solution of MINLP problem (28). □

Let us turn back to Example 3.1 in Sect. 3. We next use generalized Benders decomposition procedure in Algorithm 2 to solve this example.

For this case, one has that $K = \{(u, v) : u \geq 0, v \geq 0\}$ and $K^+ = K$.

Let $y_1 = 2$ be the initial point and $z_1^* = (0, 1) \in K^+$ with $\|z_1^*\| = 1$. Obtain optimal Lagrange multiplier $u_1^* = (1, 0.001)$ by solving the optimization problem $L - P(y_1)$ as in (26). Set $T^1 = \{u_1^*\}$, $S^1 = \{z_1^*\}$ and $\mathrm{UBD}^1 = v_{y_1}(0) = 2$. Solve the relaxed master program $\mathrm{RMP}(T^1, S^1)$ and denote the optimal solution by $(y_2, \theta_2) = (1, -0.002)$. Since $v_{y_2}(0) \leq \theta_2 + \varepsilon$, the termination criterion is satisfied and then the algorithm stops. Thus $(x_2, y_2) = ((1, 0), 1)$ is an ε-tolerance optimal solution of MINLP problem (15). □

5 Proofs of Results in Sects. 3 and 4

This section is devoted to proofs of Propositions 3.1–3.3 and 4.1–4.3.

Proof of Proposition 3.1 From $y_l \in S$, one has $\sum_{i \in I_l} [g_i(x_l, y_l)]_+ > 0$. Suppose on the contrary that there exists $\hat{x} \in X$ such that $(\hat{x}, y_l)$ is feasible to the constraint of (10). Then

$$g_i(x_l, y_l) + \langle \xi_{l,i}, \hat{x} - x_l \rangle \leq 0, \quad \forall i \in I_l \cup J_l = \{1, \ldots, m\}.$$

This implies that

$$\sum_{i=1}^{m} (\lambda_{l,i} g_i(x_l, y_l) + \langle \lambda_{l,i} \xi_{l,i}, \hat{x} - x_l \rangle) \leq 0$$

and by (8), one has

$$\sum_{i \in I_l} \lambda_{l,i} g_i(x_l, y_l) + \sum_{i=1}^{m} \langle \lambda_{l,i} \xi_{l,i}, \hat{x} - x_l \rangle \leq 0. \tag{30}$$

Noting that $\hat{x} - x_l \in T(X, x_l)$ by the convexity of X, it follows from (9) and (30) that

$$\sum_{i \in I_l} \lambda_{l,i} g_i(x_l, y_l) \leq 0.$$

By (9), one can verify that

$$\sum_{i \in I_l} [g_i(x_l, y_l)]_+ = \sum_{i \in I_l} \lambda_{l,i} g_i(x_l, y_l) \leq 0,$$

which contradicts $\sum_{i \in I_l} [g_i(x_l, y_l)]_+ > 0$. The proof is complete. □

For the proof of Proposition 3.2, we need the following lemma.

Lemma 5.1 *Let x_j solve subproblem $P(y_j)$ and subgradients (α_j, β_j), $(\xi_{j,i}, \eta_{j,i})$ be given as said in the case that $y_j \in T$. Then*

$$\langle \alpha_j, x - x_j \rangle \geq 0, \ \forall x \in X \text{ with } g_i(x_j, y_j) + \langle \xi_{j,i}, x - x_j \rangle \leq 0, \forall i = 1, \ldots, m. \tag{31}$$

Proof Let $x \in X$ be such that

$$g_i(x_j, y_j) + \langle \xi_{j,i}, x - x_j \rangle \leq 0, \forall i = 1, \ldots, m.$$

This and (5) imply that

$$\sum_{i=1}^{m} \langle \lambda_{j,i} \xi_{j,i}, x - x_j \rangle = \sum_{i=1}^{m} (\lambda_{j,i} g_i(x_j, y_j) + \langle \lambda_{j,i} \xi_{j,i}, x - x_j \rangle) \leq 0. \tag{32}$$

Noting that $x - x_j \in T(X, x_j)$ by the convexity of X, it follows from (6) and (32) that

$$\langle \alpha_j, x - x_j \rangle = \left\langle \alpha_j + \sum_{i=1}^{m} \lambda_{j,i} \xi_{j,i}, x - x_j \right\rangle - \left\langle \sum_{i=1}^{m} \lambda_{j,i} \xi_{j,i}, x - x_j \right\rangle \geq 0.$$

Hence (31) holds. The proof is complete. □

Proof of Proposition 3.2. Suppose that $(\hat{x}, \hat{y}, \hat{\theta})$ is an optimal solution of master program (MP) in (11) and (x_{j_0}, y_{j_0}) solves problem (P) in (3). We need to prove that $\hat{\theta} = f(x_{j_0}, y_{j_0})$.

Note that $(\hat{x}, \hat{y}, \hat{\theta})$ is feasible to (MP) and thus $(\hat{x}, \hat{y}, \hat{\theta})$ satisfies constraints in (MP). By Proposition 3.1, one has $\hat{y} \notin S$ and thus we can assume that $\hat{y} = y_{j_1}$ for some $y_{j_1} \in T$. Then

$$f(x_{j_1}, y_{j_1}) + (\alpha_{j_1}, \beta_{j_1})^T \begin{pmatrix} \hat{x} - x_{j_1} \\ 0 \end{pmatrix} \leq \hat{\theta} \tag{33}$$

and

$$g_i(x_{j_1}, y_{j_1}) + (\xi_{j_1,i}, \eta_{j_1,i})^T \begin{pmatrix} \hat{x} - x_{j_1} \\ 0 \end{pmatrix} \leq 0, i = 1, \ldots, m.$$

This and Lemma 5.1 imply that $\langle \alpha_{j_1}, \hat{x} - x_{j_1} \rangle \geq 0$ and by (33), one has

$$\hat{\theta} \geq f(x_{j_1}, y_{j_1}) \geq f(x_{j_0}, y_{j_0}).$$

On the other side, (x_{j_0}, y_{j_0}) is feasible to problem (P) and thus $g_i(x_{j_0}, y_{j_0}) \leq 0$ for all $i = 1, \ldots, m$. Noting that $(\alpha_j, \beta_j) \in \partial f(x_j, y_j)$, $(\xi_{j,i}, \eta_{j,i}) \in \partial g_i(x_j, y_j)$ for all $y_j \in T$ and $(\xi_{l,i}, \eta_{l,i}) \in \partial g_i(x_l, y_l)$ for all $y_l \in S$, it follows that

$$
\begin{aligned}
&f(x_j, y_j) + (\alpha_j, \beta_j)^T \begin{pmatrix} x_{j_0} - x_j \\ y_{j_0} - y_j \end{pmatrix} \leq f(x_{j_0}, y_{j_0}) \quad \forall y_j \in T, \\
&g_i(x_j, y_j) + (\xi_{j,i}, \eta_{j,i})^T \begin{pmatrix} x_{j_0} - x_j \\ y_{j_0} - y_j \end{pmatrix} \leq g_i(x_{j_0}, y_{j_0}) \leq 0 \ \ \forall y_j \in T, i = 1, \ldots, m, \\
&g_i(x_l, y_l) + (\xi_{l,i}, \eta_{l,i})^T \begin{pmatrix} x_{j_0} - x_l \\ y_{j_0} - y_l \end{pmatrix} \leq g_i(x_{j_0}, y_{j_0}) \leq 0 \ \ \forall y_l \in S, i = 1, \ldots, m.
\end{aligned}
$$

This implies that $(x_{j_0}, y_{j_0}, f(x_{j_0}, y_{j_0}))$ is feasible to (MP) and consequently

$$\hat{\theta} \leq f(x_{j_0}, y_{j_0}).$$

Hence $\hat{\theta} = f(x_{j_0}, y_{j_0})$. The proof is complete. □

Proof of Proposition 3.3. We prove that no integer variable will be generated more than once. Granting this, Algorithm 1 terminates in a finite number of steps by the finite cardinality of Y.

At iteration k, suppose that $(\hat{x}, \hat{y}, \hat{\theta})$ is an optimal solution to the relaxed master program MP^k. We need to show that $\hat{y} \notin T^k \cup S^k$.

By Proposition 3.1, one has $\hat{y} \notin S^k$ and thus it suffices to prove that $\hat{y} \notin T^k$. Suppose on the contrary that $\hat{y} = y_{j_k}$ for some $y_{j_k} \in T^k$. Then $(\hat{x}, y_{j_k}, \hat{\theta})$ is feasible to the relaxed master program MP^k and

$$
\begin{aligned}
&\hat{\theta} < UBD^k \leq f(x_{j_k}, y_{j_k}), \\
&f(x_{j_k}, y_{j_k}) + \langle \alpha_{j_k}, \hat{x} - x_{j_k} \rangle \leq \hat{\theta}, \\
&g_i(x_{j_k}, y_{j_k}) + \langle \xi_{j_k,i}, \hat{x} - x_{j_k} \rangle \leq 0, i = 1, \ldots, m.
\end{aligned} \tag{34}
$$

By Lemma 5.1, one has

$$\langle \alpha_{j_k}, \hat{x} - x_{j_k} \rangle \geq 0.$$

This and (34) imply that $f(x_{j_k}, y_{j_k}) \leq \hat{\theta}$, which contradicts $\hat{\theta} < f(x_{j_k}, y_{j_k})$.

Now suppose that Algorithm 1 terminate at k-th step for some k; that is, the relaxed master program MP^k is infeasible. To complete the proof, we need to show that the optimal value of problem (P) is attainable.

Let r^* denote the optimal value of problem (P) in (3). If there exists $y_j \in T^{k-1}$ such that $f(x_j, y_j) = r^*$, then the conclusion holds. Next, we consider the case that $f(x_j, y_j) > r^*$ for all $y_j \in T^{k-1}$. Then $UBD^{k-1} > r^*$ and $y_k \in T^k$ (otherwise, $y_k \in S^k$, $UBD^k = UBD^{k-1}$ by Algorithm 1 and consequently MP^k is feasible, a contradiction). This implies that subproblem $P(y_k)$ is feasible and thus $f(x_k, y_k) \geq r^*$. We claim that $f(x_k, y_k) = r^*$. (Otherwise $f(x_k, y_k) > r^*$. Then $r^* < UBD^k$ and thus one can verify that MP^k is feasible, a contradiction). The proof is complete. □

Proof of Proposition 4.1. (i) Let $y \in Y$. From the convexity of $f(\cdot, y)$ and K-convexity of $g(\cdot, y)$, one can verify that $v_y(\cdot)$ is a convex and monotone nonincreasing and thus $\mathrm{dom}(v_y)$ is convex. Since X is compact and $f(\cdot, y)$, $g(\cdot, y)$ are continuous, it is easy to verify that $\mathrm{dom}(v_y)$ is closed.

We next prove the lower semicontinuity of $v_y(\cdot)$. Let $z \in \mathrm{dom}(v_y)$ and $z_k \to z$ with $z_k \in \mathrm{dom}(v_y)$. Then for any $k \in \mathbb{N}$, there exists $x_k \in X$ such that $g(x_k, y) \preceq_K z_k$ and

$$v_y(z_k) \geq f(x_k, y) - \frac{1}{k}.$$

Noting that X is compact, without loss of generality, we can assume that $x_k \to \bar{x} \in X$ (considering subsequence if necessary). Thus $g(\bar{x}, y) \preceq_K z$ by the continuity of $g(\cdot, y)$. This implies that

$$\liminf_{k\to\infty} v_y(z_k) \geq \liminf_{k\to\infty}(f(x_k, y) - \frac{1}{k}) \geq f(\bar{x}, y) \geq v_y(z).$$

(ii) Let $y \in Y$ be such that $P(y)$ is feasible. By the Slater constraint qualification in (21), one has $-g(\hat{x}, y) \in \mathrm{int}(K)$ and then there exists $\delta > 0$ such that $-g(\hat{x}, y) + B(0, \delta) \subset K$. This implies that $0 \in \mathrm{int}(\mathrm{dom}(v_y))$ and it follows from [28, Proposition 3.3] that v_y is continuous at 0.

Let $-\bar{u}^* \in \partial v_y(0)$ and suppose that $\bar{x}$ is an optimal solution of subproblem $P(y)$. We next prove that $(\bar{x}, \bar{u}^*)$ satisfies the optimality conditions in (20).

For any $z \in K$, by $-\bar{u}^* \in \partial v_y(0)$, one has

$$\langle \bar{u}^*, z\rangle \geq v_y(0) - v_y(z) \geq 0$$

as $v_y(\cdot)$ is monotone nonincreasing. This implies that $\bar{u}^* \in K^+$. Noting that $\bar{x}$ solves $P(y)$, it follows that

$$g(\bar{x}, y) \preceq_K 0,\ \ v_y(g(\bar{x}, y)) = v_y(0) \text{ and } \langle \bar{u}^*, g(\bar{x}, y)\rangle \geq v_y(0) - v_y(g(\bar{x}, y)) = 0. \tag{35}$$

From $\bar{u}^* \in K^+$, one has $\langle \bar{u}^*, g(\bar{x}, y)\rangle \leq 0$ and so $\langle \bar{u}^*, g(\bar{x}, y)\rangle = 0$ by (35) .

For any $x \in X$, by $-\bar{u}^* \in \partial v_y(0)$, one has

$$\begin{aligned} f(x, y) \geq v_y(g(x, y)) \geq v_y(0) - \langle \bar{u}^*, g(x, y)\rangle = f(\bar{x}, y) \\ + \langle \bar{u}^*, g(\bar{x}, y)\rangle - \langle \bar{u}^*, g(x, y)\rangle. \end{aligned}$$

This implies that $f(\bar{x}, y) + \langle \bar{u}^*, g(\bar{x}, y)\rangle = \min_{x\in X}\left\{f(x, y) + \langle \bar{u}^*, g(x, y)\rangle\right\}$. Hence $\bar{u}^*$ is the optimal Lagrange multiplier of $P(y)$.

Conversely, let $\bar{u}^*$ be an optimal Lagrange multiplier of $P(y)$. Then there exists $\bar{x} \in X$ such that $(\bar{x}, \bar{u}^*)$ satisfies optimality conditions in (20). Let $z \in \mathrm{dom}(v_y)$ and $x \in X$ satisfying $g(x, y) \preceq_K z$. Then $\langle \bar{u}^*, z\rangle \geq \langle \bar{u}^*, g(x, y)\rangle$ and

$$f(x, y) + \langle \bar{u}^*, z\rangle \geq f(x, y) + \langle \bar{u}^*, g(x, y)\rangle \geq f(\bar{x}, y) + \langle \bar{u}^*, g(\bar{x}, y)\rangle = v_y(0).$$

This implies that $v_y(z) \geq v_y(0) - \langle \bar{u}^*, z\rangle$ and thus $-\bar{u}^* \in \partial v_y(0)$. The proof is complete. □

Proof of Proposition 4.2 (i) Note that if $P(y)$ is feasible, then there is $x \in X$ such that $g(x, y) \preceq_K 0$ and thus the necessity part follows.

The sufficiency part. By (22), one has

$$\sup_{u^* \in K^+, \|u^*\|=1} \inf_{x \in X} \langle u^*, g(x, \bar{y})\rangle \leq 0.$$

This implies that

$$\max_{u^* \in K^+} \left\{ \inf_{x \in X} \langle u^*, g(x, \bar{y})\rangle \right\} = 0. \tag{36}$$

Note that $P(y)$ is feasible if and only if $0 \in \mathrm{dom}(v_y)$ and it suffices to prove that

$$0 \in \mathrm{dom}(v_y).$$

Suppose on the contrary that $0 \notin \mathrm{dom}(v_y)$. Since $\mathrm{dom}(v_y)$ is convex and closed by Proposition 4.1(i), it follows from separation theorem that there exist $z^* \in \mathbb{R}^m$ with $\|z^*\| = 1$ and $r \in \mathbb{R}$ such that

$$\inf\{\langle z^*, z\rangle : z \in \mathrm{dom}(v_y)\} > r > 0. \tag{37}$$

We first show that $z^* \in K^+$. Let $z \in K$ and $k \in \mathbb{N}$. Take any $x \in X$. Then $kz \in K$ and $g(x, y) + kz \in \mathrm{dom}(v_y)$. By (37), one has $\langle z^*, g(x, y) + kz\rangle > r > 0$ and thus

$$\frac{1}{k}\langle z^*, g(x, y)\rangle + \langle z^*, z\rangle > \frac{r}{k} > 0.$$

By taking limits as $k \to \infty$, one has $\langle z^*, z\rangle \geq 0$. Hence $z^* \in K^+$.

Noting that $f(\cdot, y)$ is continuous and X is compact, it follows that

$$\inf_{x \in X} f(x, y) > -\infty$$

and for any $t > 0$, one has $tz^* \in K^+$ and

$$\inf_{x \in X}\{f(x, y) + \langle tz^*, g(x, y)\rangle\} \geq \inf_{x \in X} f(x, y) + t \inf_{x \in X}\langle z^*, g(x, y)\rangle > \inf_{x \in X} f(x, y) + tr.$$

By taking limits as $t \to +\infty$, one has

$$\max_{u^* \in K^+} \left\{ \inf_{x \in X} \left\{ f(x, y) + \langle u^*, g(x, y)\rangle \right\} \right\} = +\infty,$$

which contradicts (36).

(ii) Using the weak duality, one has

$$\max_{u^*\in K^+}\left\{\inf_{x\in X}\{f(x,y)+\langle u^*,g(x,y)\rangle\}\right\}\le \inf_{x\in X}\{f(x,y):g(x,y)\preceq_K 0\}=v_y(0).$$

On the other side, suppose that $-\bar{u}^*\in\partial v_y(0)$. By Proposition 4.1(ii), $\bar{u}^*\in K^+$ and for any $x\in X$, one has

$$f(x,y)\ge v_y(g(x,y))\ge v_y(0)-\langle\bar{u}^*,g(x,y)\rangle$$

and thus

$$\inf_{x\in X}\{f(x,y)+\langle\bar{u}^*,g(x,y)\rangle\}\ge v_y(0).$$

This means that

$$\max_{u^*\in K^+}\left\{\inf_{x\in X}\{f(x,y)+\langle u^*,g(x,y)\rangle\}\right\}\ge v_y(0).$$

Hence (23) holds. The proof is complete. □

For the proof of Proposition 4.3, we need the following lemma.

Lemma 5.2 *Let $L:Y\times K^+\to\mathbb{R}$ be defined as*

$$L(y,u^*):=\inf_{x\in X}\left\{f(x,y)+\langle u^*,g(x,y)\rangle\right\},\quad\forall(y,u^*)\in Y\times K^+. \tag{38}$$

Then for any $\bar{y}\in Y$, one has that $L(\bar{y},\cdot)$ is continuous on K^+.

Proof Let $\bar{y}\in Y$ be fixed and $\bar{u}^*\in K^+$. We first prove that $L(\bar{y},\cdot)$ is lower semicontinuous at $\bar{u}^*$. To do this, take any $u_k^*\to\bar{u}^*$ with $u_k^*\in K^+$. Then for any $k\in\mathbb{N}$, there exists $x_k\in X$ such that

$$f(x_k,\bar{y})+\langle u_k^*,g(x_k,\bar{y})\rangle=L(\bar{y},u_k^*) \tag{39}$$

as X is compact and $f(\cdot,\bar{y}),g(\cdot,\bar{y})$ are continuous. Without loss of generality, we can assume that $x_k\to\bar{x}\in X$ (considering subsequence if necessary). By (39), one has

$$L(\bar{y},\bar{u}^*)\le f(\bar{x},\bar{y})+\langle\bar{u}^*,g(\bar{x},\bar{y})\rangle\le\lim_k(f(x_k,\bar{y})+\langle u_k^*,g(x_k,\bar{y})\rangle).$$

This implies that $L(\bar{y},\cdot)$ is lower semicontinuity at $\bar{u}^*$.

We next show that $L(\bar{y},\cdot)$ is upper semicontinuous at $\bar{u}^*$. For any $x\in X$, one has

$$f(x,\bar{y})+\langle\bar{u}^*,g(x,\bar{y})\rangle=\lim_k(f(x,\bar{y})+\langle u_k^*,g(x,\bar{y})\rangle)\ge\limsup_k L(\bar{y},u_k^*).$$

This implies that

$$L(\bar{y}, \bar{u}^*) = \inf_{x \in X} (f(x, \bar{y}) + \langle \bar{u}^*, g(x, \bar{y}) \rangle) \geq \limsup_k L(\bar{y}, u_k^*)$$

and consequently $L(\bar{y}, \cdot)$ is upper semicontinuous at $\bar{u}^*$. Hence $L(\bar{y}, \cdot)$ is continuous at $\bar{u}^*$. The proof is complete. □

Proof of Proposition 4.3. Suppose on the contrary that there exists $\varepsilon_0 > 0$ such that Algorithm 2 can generate a sequence $\{(y_k, \theta_k)\}$ satisfying subproblem $P(y_k)$ is feasible and $\theta_k \in \mathbb{R}$ for all k. Then for any k, one has $u_k^* \in K^+$ is the optimal Lagrange multiplier of subproblem $P(y_k)$ and thus $u_k^* \in -\partial v_{y_k}(0)$ by Proposition 4.1(ii). From the relaxed master program RMP(T^k, S^k), one can verify that $\{\theta_k\}$ is nondecreasing and bounded above. By the finite cardinality of Y, there exist $\bar{y} \in Y$ and a subsequence $\{k_i : i = 1, 2, \ldots\}$ such that $y_{k_i} \equiv \bar{y}$ for all $i \in \mathbb{N}$. Using Proposition 4.1(ii), one has that $v_{\bar{y}}(\cdot)$ is continuous at 0 and it follows from [28, Proposition 1.11] that $\partial v_{\bar{y}}(0)$ is closed and bounded. This implies that $\{u_{k_i}^*\}$ is bounded. Thus, without loss of generality, we can assume that $(u_{k_i}^*, \theta_{k_i}) \to (\bar{u}^*, \bar{\theta}) \in -\partial v_{\bar{y}}(0) \times \mathbb{R}$.

By solving relaxed master program RMP(T^{k_i}, S^{k_i}) in Algorithm 2, one has

$$\theta_{k_{i+1}} \geq L(y_{k_{i+1}}, u_{k_i}^*) = L(\bar{y}, u_{k_i}^*).$$

By Lemma 5.2, $L(\bar{y}, \cdot)$ is continuous at $\bar{u}^*$ and it follows by taking limits as i that

$$\bar{\theta} \geq L(\bar{y}, \bar{u}^*) = v_{\bar{y}}(0).$$

This implies that when i is sufficiently large, one has

$$\theta_{k_i} + \frac{\varepsilon_0}{2} > \bar{\theta} \geq v_{\bar{y}}(0) = v_{y_{k_i}}(0) > v_{y_{k_i}}(0) - \frac{\varepsilon_0}{2}$$

and consequently

$$v_{y_{k_i}}(0) \leq \theta_{k_i} + \varepsilon_0.$$

Then the termination criterion in Algorithm 2 is satisfied and thus Algorithm 2 will terminate at k_i-step, a contradiction. The proof is complete. □

6 Conclusions

Compared with smooth functions, nonsmooth functions appear more frequently and generally in optimization problems; for example, for the function space consisting of all continuous functions on closed interval $[a, b]$, the complement of the subset with functions that fail to be differentiable everywhere is the first category. This means that the number of nonsmooth functions is far more than that of smooth functions. Meanwhile, over the past three decades, nonsmooth analysis has been recognized as a

broad spectrum of mathematical theory that has grown in connection with many aspects of mathematical programming and optimization. This paper is motivated by the idea of applying techniques in nonsmooth analysis to dealing with nonsmooth optimization problems. By the subgradients, we study two algorithms by OA and GBD for dealing with MINLPs that are lack of differentiability data. The further work may be done along two lines: one is to consider how to solve more complicated nonsmooth MINLPs that are with a larger scale and difficult to solve by differentiability techniques, and the other is to tackle the nonsmooth MINLPs by relaxing or dropping the convexity assumption.

Acknowledgements This research of the first author was supported by the National Natural Science Foundations of China (Nos. 11826204, 11826206 and 11771384), the Natural Science Foundation of Yunnan Province of China (No. 2018FB004) and the Scientific Research Foundation of Yunnan University under grant No. 2018YDJQ010, and by Joint Key Project of Yunnan Provincial Science and Technology Department and Yunnan University (No. 2018FY001(-014)) and IRTSTYN. The research work of Bo Zeng was supported by National Science Foundation CMMI-1436452. The research work of M. Montaz Ali was supported by the funding from NRF of South Africa (No. 114846). The research work of Jen-Chih Yao was partially supported by the Grant MOST 105-2221-E-039-009-MY3.

References

1. Bonami, P., Biegler, L., Conn, A.R., Cornuéjols, G., Grossmann, I.E., Laird, C., Lee, J., Lodi, A., Margot, F., Sawaya, N., Wächter, A.: An algorithmic framework for convex mixed integer nonlinear programs. Discrete Optim. **5**(2), 186–204 (2008)
2. Boukouvala, F., Misener, R., Floudas, C.A.: Global optimization advances in mixed-integer nonlinear programming, MINLP, and constrained derivative-free optimization, CDFO. Eur. J. Oper. Res. **252**(3), 701–727 (2016)
3. Castillo, I., Westerlund, J., Emet, S., Westerlund, T.: Optimization of block layout design problems with unequal areas: a comparison of MILP and MINLP optimization methods. Comput. Chem. Eng. **30**, 54–69 (2005)
4. Floudas, C.A., Ciric, A.R.: Strategies for overcoming uncertainties in heat exchanger network synthesis. Comput. Chem. Eng **13**, 1133–1152 (1989)
5. Grossmann, I.E.: Review of nonlinear mixed-integer and disjunctive programming techniques. Optim. Eng. **3**, 227–252 (2002)
6. Hooker, J.N.: Logic-Based Methods for Optimization: Combining Optimization and Constraint Satisfaction. Wiley, New York (2000)
7. Hooker, J.N.: Planning and scheduling by logic-based Benders decomposition. Oper. Res. **55**, 588–602 (2007)
8. Hooker, J.N., Ottosson, G.: Logic-based Benders decomposition. Math. Program. **96**, 33–60 (2003)
9. Tawarmalani, M., Sahinidis, N.V.: Convexification and Global Optimization in Continuous and Mixed-Integer Nonlinear Programming: Theory, Algorithms, Software, and Applications. Kluwer Academic Publishers, Alphen aan den Rijn (2002)
10. Gupta, O.K., Ravindran, V.: Branch and bound experiments in convex nonlinear integer programming. Manag. Sci. **31**(12), 1533–1546 (1985)
11. Leyffer, S.: Integrating SQP and branch-and-bound for mixed integer nonlinear programming. Comput. Optim. Appl. **18**, 295–309 (2001)
12. Tawarmalani, M., Sahinidis, N.V.: A polyhedral branch-and-cut approach to global optimization. Math. Program. **103**(2), 225–249 (2005)
13. Kelley, J.E.: The cutting-plane method for solving convex programs. J. SIAM **8**, 703–712 (1960)
14. Westerlund, T., Pettersson, F.: An extended cutting plane method for solving convex MINLP problems. Comput. Chem. Eng. **19**, 131–136 (1995)
15. Duran, M., Grossmann, I.E.: An Outer approximation algorithm for a class of mixed-integer nonlinear programs. Math. Program. **36**, 307–339 (1986)

16. Fletcher, R., Leyffer, S.: Solving mixed-integer nonlinear programs by outer approximation. Math. Program. **66**, 327–349 (1994)
17. Kesavan, P., Allgor, R.J., Gatzke, E.P., Barton, P.I.: Outer approximation algorithms for separable nonconvex mixed-integer nonlinear programs. Math. Program. **100**(3), 517–535 (2004)
18. Wei, Z., Ali, M.M.: Outer approximation algorithm for one class of convex mixed-integer nonlinear programming problems with partial differentiability. J. Optim. Theory Appl. **167**(2), 644–652 (2015)
19. Wei, Z., Ali, M.M.: Convex mixed integer nonlinear programming problems and an outer approximation algorithm. J. Glob. Optim. **63**(2), 213–227 (2015)
20. Benders, J.F.: Partitioning procedures for solving mixed-variables programming problems. Numer. Math. **4**, 238–252 (1962)
21. Geoffrion, A.M.: Generalized Benders decomposition. J. Optim. Theory Appl. **10**(4), 237–260 (1972)
22. Li, X., Tomasgard, A., Barton, P.I.: Nonconvex generalized Benders decomposition for stochastic separable mixed-integer nonlinear programs. J. Optim. Theory Appl. **151**(3), 425 (2011)
23. Wei, Z., Ali, M.M.: Generalized Benders decomposition for one class of MINLPs with vector conic constraint. SIAM J. Optim. **25**(3), 1809–1825 (2015)
24. Drewes, S., Ulbrich, S.: Subgradient based outer approximation for mixed integer second order cone programming. In: Mixed Integer Nonlinear Programming, The IMA Volumes in Mathematics and Its Applications, vol. 154, pp. 41–59 (2012)
25. Eronen, V.-P., Mäkelä, M.M., Westerlund, T.: On the generalization of ECP and OA methods to nonsmooth convex MINLP problems. Optimization **63**, 1057–1073 (2014)
26. Zălinescu, C.: Convex Analysis in General Vector Spaces. World Scientific, Singapore (2002)
27. Rockafellar, R.T.: Convex Analysis. Princeton University Press, Princeton (1996)
28. Phelps, R.R.: Convex functions, Monotone operators and Differentiability. Lecture Notes in Mathematics, vol. 1364. Springer, New York (1989)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.