Pré-Publicações do Departamento de Matemática
Universidade de Coimbra
Preprint Number 11–12

# INEXACT SOLUTION OF NLP SUBPROBLEMS IN MINLP

M. LI AND L. N. VICENTE

ABSTRACT: In the context of convex mixed-integer nonlinear programming (MINLP), we investigate how the outer approximation method and the generalized Benders decomposition method are affected when the respective NLP subproblems are solved inexactly. We show that the cuts in the corresponding master problems can be changed to incorporate the inexact residuals, still rendering equivalence and finiteness in the limit case. Some numerical results will be presented to illustrate the behavior of the methods under NLP subproblem inexactness.

KEYWORDS: Mixed integer nonlinear programming, outer approximation, generalized Benders decomposition, inexactness, convexity.
AMS SUBJECT CLASSIFICATION (2010): 90C11, 90C30.

## 1. Introduction

Recently, mixed integer nonlinear programming (MINLP) has become again a very active research area [1, 2, 4, 5, 6, 7, 14, 16]. Benders [3] developed in the 60's a technique for solving linear mixed-integer problems, later called Benders decomposition. Geoffrion [11] extended it to MINLP in 1972, in what become known as generalized Benders decomposition (GBD). Much later, in 1986, Duran and Grossmann [8] derived a new outer approximation (OA) method to solve a particular class of MINLP problems, which become widely used in practice. Although the authors shown finiteness of the OA algorithm, their theory was restricted to problems where the discrete variables appear linearly and the functions involving the continuous variables are convex. Both OA and GBD are iterative schemes requiring at each iteration the solution of a (feasible or infeasible) NLP subproblem and one mixed-integer linear programming (MILP) master problem.

For these particular MINLP problems, Quesada and Grossmann [15] then proved that the cuts in the master problem of OA imply the cuts in the master problem of GBD, showing that the GBD algorithm provides weaker lower bounds and generally requires more major iterations to converge. Fletcher

*Date*: April 15, 2011.

Support for this work was provided by FCT under the scholarship SFRH/BD/33369/2008 and the grant PTDC/MAT/098214/2008.

and Leyffer [9] generalized the OA method of Duran and Grossmann [8] into a wider class of problems where nonlinearities in the discrete variables are allowed as long as the corresponding functions are convex in these variables. They also introduced a new and simpler proof of finiteness of the OA algorithm. The relationship between OA and GBD was then addressed, again, by Grossmann [12] in this wider context of MINLP problems, showing once more that the lower bound predicted by the relaxed master problem of OA is greater than or equal to the one predicted by the relaxed master problem of GBD (see also Flippo and Rinnooy Kan [10] for the relationship between the two techniques). Recently, Bonami et al. [4] suggested a different OA algorithm using linearizations of both the objective function and the constraints, independently of being taken at the feasible or infeasible NLP subproblem, to build the MILP master problem. This technique is, in fact, different from the traditional OA (see [9]), where the cuts in the master MILP problems do not involve linearizations of the objective function in the infeasible case.

Westerlund and Pettersson [18] generalized the cutting plane method [13] from convex NLP to convex MINLP, in what is known as the extended cutting plane (ECP) method (see also [19, 20]). While OA and GBD alternate between the solution of MILP and NLP subproblems, the ECP relies only on the solution of MILP problems.

In the above mentioned OA and GBD approaches, the NLP subproblems are solved exactly, at least for the derivation of the theoretical properties, such as equivalence between original and master problem and finite termination of the corresponding algorithms. In this paper we investigate the effect of NLP subproblem inexactness in these two techniques. We show how the cuts in the master problems can be changed to incorporate the inexact residuals of the first order necessary conditions of the NLP subproblems, in a way that still renders the equivalence and finiteness properties, as long as the size of these residuals allow inferring the cuts from convexity properties.

In this paper, we will adopt the MINLP formulation

$$\mathrm{P}\left\{\begin{array}{ll} \min & f(x,y) \\ \text{s.t.} & g(x,y) \ \leq \ 0, \\ & x \in X \cap \mathbb{Z}^{n_d}, y \in Y, \end{array}\right.$$

where $X$ is a bounded polyhedral subset of $\mathbb{R}^{n_d}$ and $Y$ a polyhedral subset of $\mathbb{R}^{n_c}$. The functions $f : X \times Y \longrightarrow \mathbb{R}$ and $g : X \times Y \longrightarrow \mathbb{R}^m$ are assumed continuously differentiable. We will also assume that P is convex, i.e., that $f$ and $g$ are convex functions.

Let $x^j$ be any element of $X \cap \mathbb{Z}^{n_d}$. Consider, then, the (convex) subproblem

$$\mathrm{NLP}(x^j) \begin{cases} \min & f(x^j, y) \\ \text{s.t.} & g(x^j, y) \ \leq \ 0, \\ & y \in Y, \end{cases}$$

and suppose it is feasible. In this case, $y^j$ will represent an approximated optimal solution of $\mathrm{NLP}(x^j)$. For an $x^k$ in $X \cap \mathbb{Z}^{n_d}$ for which $\mathrm{NLP}(x^k)$ is infeasible, $y^k$ is instead defined as an approximated optimal solution of the following feasibility (convex) subproblem

$$\mathrm{NLPF}(x^k) \begin{cases} \min & u \\ \text{s.t.} & g_i(x^k, y) \ \leq \ u, \ i = 1, \ldots, m, \\ & y \in Y, u \in \mathbb{R}, \end{cases}$$

where one minimizes the $\ell_\infty$-norm of the measure of infeasibility of subproblem $\mathrm{NLP}(x^k)$.

For a matter of simplification, and without loss of generality, we suppose that the constraints $y \in Y$ are part of the constraints $g(x^j, y) \leq 0$ and $g_i(x^k, y) \leq u, \ i = 1, \ldots, m$, in the subproblems $\mathrm{NLP}(x^j)$ and $\mathrm{NLPF}(x^k)$, respectively. In addition, let us assume that the approximated optimal solutions of the NLP subproblems satisfy an inexact form of the corresponding first order necessary Karush-Kuhn-Tucker (KKT) conditions. More particularly, in the case of $\mathrm{NLP}(x^j)$, we assume the existence of $\lambda^j \in \mathbb{R}^m_+, r^j \in \mathbb{R}^{n_c}$, and $s^j \in \mathbb{R}^m$, such that

$$\nabla_y f(x^j, y^j) + \sum_{i=1}^{m} \lambda_i^j \nabla_y g_i(x^j, y^j) \ = \ r^j, \tag{1}$$

$$\lambda_i^j g_i(x^j, y^j) \ = \ s_i^j, \quad i = 1, \ldots, m. \tag{2}$$

When NLP$(x^k)$ is infeasible, we assume, for NLPF$(x^k)$, the existence of $\mu^k \in \mathbb{R}^m_+$, $z^k \in \mathbb{R}^m$, $w^k \in \mathbb{R}$, and $v^k \in \mathbb{R}^{n_c}$, such that

$$\sum_{i=1}^{m} \mu_i^k \nabla_y g_i(x^k, y^k) \;=\; v^k, \tag{3}$$

$$1 - \sum_{i=1}^{m} \mu_i^k \;=\; w^k, \tag{4}$$

$$\mu_i^k (g_i(x^k, y^k) - u^k) \;=\; z_i^k, \quad i = 1, \ldots, m. \tag{5}$$

Points satisfying the inexact KKT conditions can be seen as solutions of appropriate perturbed subproblems (see the Appendix). The following two sets will then be used to index these two sets of approximated optimal solutions:

$$T \;=\; \{j : x^j \in X \cap \mathbb{Z}^{n_d}, \text{NLP}(x^j) \text{ is feasible and } y^j \text{ appr. solves } \text{NLP}(x^j)\}$$

and

$$S \;=\; \{k : x^k \in X \cap \mathbb{Z}^{n_d}, \text{NLP}(x^k) \text{ is infeasible and } y^k \text{ appr. solves } \text{NLPF}(x^k)\}.$$

The inexact versions of OA and GBD studied in this paper will attempt to find the best pair among all of the form $(x^j, y^j)$ corresponding to $j \in T$. Implicitly, we are thus redefining a perturbed version of problem P and will denote it by $\mathcal{P}$:

$$\mathcal{P} \quad \min_{j \in T} f(x^j, y^j). \tag{6}$$

This problem is well defined if $T \neq \emptyset$ which in turn can be assumed when the original MINLP problem P has a finite optimal value.

We use the superscripts $l$, $p$, and $q$ to denote the iteration count, superscript $j$ to index the feasible NLP subproblems defined above, and $k$ to indicate infeasible subproblems. The following notation is adopted to distinguish between function values and functions. $f^l = f(x^l, y^l)$ denotes the value of $f$ evaluated at the point $(x^l, y^l)$, similarly, $\nabla f^l = \nabla f(x^l, y^l)$ is the value of the gradient of $f$ at the point $(x^l, y^l)$, $\nabla_x f^l = \nabla_x f(x^l, y^l)$ is the value of the gradient of $f$ with respect to $x$ at the point $(x^l, y^l)$, and $\nabla_y f^l = \nabla_y f(x^l, y^l)$ is the value of the gradient of $f$ with respect to $y$ at the point $(x^l, y^l)$. Moreover, the same conventions apply for all other functions.

We organize the paper in the following way. In Section 2, we extend OA for the inexact solution of the NLP subproblems, rederiving the corresponding background theory and main algorithm. In Section 3 we proceed similarly for GBD, also discussing the relationship between the inexact forms of OA and GBD. Section 4 describes a set of preliminary numerical experiments, reported to better understand some of the theoretical features encountered in our study of inexactness in MINLP.

## 2. Inexact outer approximation

**2.1. Equivalence between perturbed and master problems for OA.** OA relies on the fact that the original problem P is equivalent to a MILP (master problem) formed by minimizing the least of the linearized forms of $f$ for indices in $T$ subject to the linearized forms of $g$ for indices in $S$ and $T$. When the NLP subproblems are solved inexactly, one has to consider perturbed forms of such cuts or linearized forms in order to keep an equivalence, this time to the perturbed problem $\mathcal{P}$. In turn, these inexact cuts lead to a different, perturbed MILP (master problem) given by

$$\mathcal{P}^{\mathrm{OA}} \left\{ \begin{array}{ll} \min & \alpha \\ \text{s.t.} & \begin{pmatrix} \nabla_x f(x^j, y^j) \\ \nabla_y f(x^j, y^j) - r^j \end{pmatrix}^\top \begin{pmatrix} x - x^j \\ y - y^j \end{pmatrix} + f(x^j, y^j) \ \leq \ \alpha, \\ & \nabla g(x^j, y^j)^\top \begin{pmatrix} x - x^j \\ y - y^j \end{pmatrix} + g(x^j, y^j) \ \leq \ t^j, \ \ \forall j \in T, \\ & \begin{pmatrix} \nabla_x g_i(x^k, y^k) \\ \nabla_y g_i(x^k, y^k) - \frac{1}{1-w^k} v^k \end{pmatrix}^\top \begin{pmatrix} x - x^k \\ y - y^k \end{pmatrix} + g_i(x^k, y^k) \ \leq \ a_i^k, \\ & i = 1, \ldots, m, \ \forall k \in S, \\ & x \in X \cap \mathbb{Z}^{n_d}, y \in Y, \alpha \in \mathbb{R}, \end{array} \right.$$

where, for $i = 1, \ldots, m$,

$$t_i^j \ = \left\{ \begin{array}{ll} \frac{s_i^j}{\lambda_i^j}, & \text{if } \lambda_i^j \ > \ 0, \\ 0, & \text{if } \lambda_i^j \ = \ 0, \end{array} \right. \tag{7}$$

and

$$a_i^k \ = \left\{ \begin{array}{ll} \frac{m z_i^k - w^k u^k}{m \mu_i^k}, & \text{if } \mu_i^k \ > \ 0, \\ 0, & \text{if } \mu_i^k \ = \ 0. \end{array} \right. \tag{8}$$

Note that when $r$, $s$, $v$, $w$, and $z$ are zero, we obtain the well-known master problem in OA. Also, optionally, one could have added the cuts

$$\nabla f(x^k, y^k)^\top \begin{pmatrix} x - x^k \\ y - y^k \end{pmatrix} + f(x^k, y^k) \ \leq \ \alpha, \ \ \forall k \in S, \tag{9}$$

corresponding to linearizations of the objective function in the infeasible cases, as suggested in [4].

From the convexity and continuous differentiability of $f$ and $g$, we know that, for any $(x^l, y^l) \in \mathbb{R}^{n_d} \times \mathbb{R}^{n_c}$,

$$f(x,y) \;\geq\; f(x^l,y^l) + \nabla f(x^l,y^l)^\top \begin{pmatrix} x - x^l \\ y - y^l \end{pmatrix}, \tag{10}$$

$$g(x,y) \;\geq\; g(x^l,y^l) + \nabla g(x^l,y^l)^\top \begin{pmatrix} x - x^l \\ y - y^l \end{pmatrix}. \tag{11}$$

In addition, when $y^j$ is a feasible point of NLP$(x^j)$, we obtain from (11) and $g(x^j, y^j) \leq 0$ that

$$0 \;\geq\; g(x^l,y^l) + \nabla g(x^l,y^l)^\top \begin{pmatrix} x^j - x^l \\ y^j - y^l \end{pmatrix}. \tag{12}$$

The inexact OA method reported in this section as well as the GBD method of the next section require the residuals of the inexact KKT conditions to satisfy the bounds given in the next two assumptions, in order to validate the equivalence between perturbed and master problems, and to ensure finiteness of the respective algorithms. Essentially, these bounds will ensure that the above convexity properties will still imply the inexact cuts at the remaining points. We first give the bounds on the residuals $r$ and $s$ for the feasible case.

**Assumption 2.1.** *Given any $l, j \in T$, with $l \neq j$, assume that*

$$\|r^l\| \;\leq\; \frac{-\tau\left[(\nabla f^l)^\top \begin{pmatrix} x^j - x^l \\ y^j - y^l \end{pmatrix} + f^l - f^j\right]}{\|y^j - y^l\|},$$

*for some $\tau \in [0,1)$, and*

$$|s_i^l| \;\leq\; -\sigma_i \lambda_i^l\left[(\nabla g_i^l)^\top \begin{pmatrix} x^j - x^l \\ y^j - y^l \end{pmatrix} + g_i^l\right],$$

*for some $\sigma_i \in [0,1]$, $i = 1, \ldots, m$.*

Now, we state the bounds for the residuals $v$, $w$, and $z$ in the infeasible case.

**Assumption 2.2.** *Given any $j \in T$ and any $k \in S$, and for all $i \in \{1,\ldots,m\}$, if $\mu_i^k \neq 0$, assume that*

$$\frac{1}{1-w^k}\|v^k\|\,\|y^j-y^k\| + \frac{1}{\mu_i^k}|z_i^k| + \frac{u^k}{m\mu_i^k}|w^k| \;\leq\; -\beta_i[(\nabla g_i^k)^\top \begin{pmatrix} x^j-x^k \\ y^j-y^k \end{pmatrix} + g_i^k],$$

*for some $\beta_i \in [0,1]$, otherwise, assume that*

$$\frac{1}{1-w^k}\|v^k\|\,\|y^j-y^k\| \;\leq\; -\eta_i[(\nabla g_i^k)^\top \begin{pmatrix} x^j-x^k \\ y^j-y^k \end{pmatrix} + g_i^k],$$

*for some $\eta_i \in [0,1]$.*

We are now in a position to state the equivalence between the original, perturbed MINLP problem and the MILP master problem $\mathcal{P}^{\mathrm{OA}}$.

**Theorem 2.1.** *Let* P *be a convex* MINLP *problem and $\mathcal{P}$ be its perturbed problem as defined in the Introduction. Assume that* P *is feasible with a finite optimal value and that the residuals of the* KKT *conditions of the* NLP *subproblems satisfy Assumptions 2.1 and 2.2. Then $\mathcal{P}^{\mathrm{OA}}$ and $\mathcal{P}$ have the same optimal value.*

*Proof*: The proof follows closely the lines of the proof of [4, Theorem 1]. Since problem P has a finite optimal value it follows that, for every $x \in X \cap \mathbb{Z}^{n_d}$, either problem NLP$(x)$ is feasible with a finite optimal value or it is infeasible, that the sets $T$ and $S$ are well defined, and that the set $T$ is nonempty. Now, given any $x^l \in X \cap \mathbb{Z}^{n_d}$ with $l \in T \cup S$, let $\mathcal{P}^{\mathrm{OA}}_{x^l}$ denote the problem in $\alpha$ and $y$ obtained from $\mathcal{P}^{\mathrm{OA}}$ when $x$ is fixed to $x^l$. First we will prove that problem $\mathcal{P}^{\mathrm{OA}}_{x^k}$ is infeasible for every $k \in S$.

**Part I. Establishing infeasibility of $\mathcal{P}^{\mathrm{OA}}_{x^k}$ for $k \in S$.**

In this case, problem NLP$(x^k)$ is infeasible and $y^k$ is an approximated optimal solution of NLPF$(x^k)$ with corresponding inexact nonnegative Lagrange multipliers $\mu^k$. When we set $x = x^k$, the corresponding constraints in $\mathcal{P}^{\mathrm{OA}}$ will result in

$$(\nabla_y g_i(x^k,y^k) - \frac{1}{1-w^k}v^k)^\top(y-y^k) + g_i(x^k,y^k) \;\leq\; a_i^k, \tag{13}$$

for $i = 1,\ldots,m$. Multiplying the inequalities in (13) by the nonnegative multipliers $\mu_1^k,\ldots,\mu_m^k$, and summing them up, one obtains

$$(\sum_{i=1}^m \mu_i^k \nabla_y g_i(x^k,y^k) - v^k)^\top(y-y^k) \;\leq\; \sum_{i=1}^m (z_i^k - \mu_i^k g_i(x^k,y^k)) - w^k u^k. \tag{14}$$

By using (3), one can see that the left hand side of the inequality in (14) is equal to 0. On the other hand, by using equation (5), the right hand side of the inequality in (14) results in $\sum_{i=1}^{m}(z_i^k - \mu_i^k g_i(x^k, y^k)) - w^k u^k = -(\Sigma_{i=1}^{m}\mu_i^k + w^k)u^k$, which is equal to $-u^k$ by (4). Since $\mathrm{NLP}(x^k)$ is infeasible, $-u^k$ must be strictly negative. We have thus proved that the inequality (14) has no solution $y$.

This derivation implies that the minimum value of $\mathcal{P}^{\mathrm{OA}}$ should be found as the minimum value of $\mathcal{P}^{\mathrm{OA}}_{x^j}$ over all $x^j \in X \cap \mathbb{Z}^{n_d}$ with $j \in T$. We prove in the next two separate subparts that, for every $j \in T$, the optimal value $\bar{\alpha}^j$ of $\mathcal{P}^{\mathrm{OA}}_{x^j}$ coincides with the approximated optimal value of $\mathrm{NLP}(x^j)$.

**Part II. Establishing that $\mathcal{P}^{\mathrm{OA}}_{x^j}$ has the same objective value as the perturbed $\mathrm{NLP}(x^j)$ for $j \in T$.**

We will show next that $(y^j, f(x^j, y^j))$ is a feasible solution of $\mathcal{P}^{\mathrm{OA}}_{x^j}$, and therefore that $f(x^j, y^j)$ is an upper bound on the optimal value $\bar{\alpha}^j$ of $\mathcal{P}^{\mathrm{OA}}_{x^j}$.

**Part II–A. Establishing that $f(x^j, y^j)$ is an upper bound for the optimal value of $\mathcal{P}^{\mathrm{OA}}_{x^j}$ for $j \in T$.**

In this case, it is easy to see that $\mathcal{P}^{\mathrm{OA}}_{x^j}$ contains all the constraints indexed by $l \in T$

$$\begin{pmatrix} \nabla_x f(x^l, y^l) \\ \nabla_y f(x^l, y^l) - r^l \end{pmatrix}^\top \begin{pmatrix} x^j - x^l \\ y - y^l \end{pmatrix} + f(x^l, y^l) \ \leq \ \alpha, \tag{15}$$

$$\nabla g(x^l, y^l)^\top \begin{pmatrix} x^j - x^l \\ y - y^l \end{pmatrix} + g(x^l, y^l) \ \leq \ t^l, \tag{16}$$

where, for $i = 1, \ldots, m$,

$$t_i^l \ = \ \begin{cases} \frac{s_i^l}{\lambda_i^l}, & \text{if } \lambda_i^l \ > \ 0, \\ 0, & \text{if } \lambda_i^l \ = \ 0, \end{cases}$$

as well as all the constraints indexed by $k \in S$ and $i \in \{1, \ldots, m\}$

$$\begin{pmatrix} \nabla_x g_i(x^k, y^k) \\ \nabla_y g_i(x^k, y^k) - \frac{1}{1-w^k} v^k \end{pmatrix}^\top \begin{pmatrix} x^j - x^k \\ y - y^k \end{pmatrix} + g_i(x^k, y^k) \ \leq \ a_i^k, \tag{17}$$

where $a_i^k$ is given as in (8).

First take any $l \in T$ and assume that $y^l$ is an approximated optimal solution of $\mathrm{NLP}(x^l)$ with corresponding inexact nonnegative Lagrange multipliers $\lambda^l$. If $l = j$, it is easy to verify that $(y^j, f(x^j, y^j))$ satisfies (15) and (16). Assume then that $l \neq j$. From Assumption 2.1, we know that, for some $\tau \in [0,1)$,

$$-(r^l)^\top (y^j - y^l) \;\leq\; \|r^l\|\,\|y^j - y^l\| \;\leq\; -\tau[(\nabla f^l)^\top \begin{pmatrix} x^j - x^l \\ y^j - y^l \end{pmatrix} + f^l - f^j].$$

Thus,

$$\begin{aligned}&[(\nabla f^l)^\top \begin{pmatrix} x^j - x^l \\ y^j - y^l \end{pmatrix} + f^l - f^j] - (r^l)^\top (y^j - y^l) \\ &\qquad \leq\; (1-\tau)[(\nabla f^l)^\top \begin{pmatrix} x^j - x^l \\ y^j - y^l \end{pmatrix} + f^l - f^j] \;\leq\; 0,\end{aligned}$$

where the last inequality comes from $1-\tau > 0$ and (10) with $(x,y) = (x^j, y^j)$. We then see that (15) is satisfied with $\alpha = f(x^j, y^j)$ and $y = y^j$.

Now, from Assumption 2.1, one has for some $\sigma_i \in [0,1], i = 1, \ldots, m$,

$$\begin{aligned}\lambda_i^l[(\nabla g_i^l)^\top \begin{pmatrix} x^j - x^l \\ y^j - y^l \end{pmatrix} + g_i^l] - s_i^l \;&\leq\; \lambda_i^l[(\nabla g_i^l)^\top \begin{pmatrix} x^j - x^l \\ y^j - y^l \end{pmatrix} + g_i^l] \\ &\qquad - \sigma_i \lambda_i^l[(\nabla g_i^l)^\top \begin{pmatrix} x^j - x^l \\ y^j - y^l \end{pmatrix} + g_i^l] \\ &\leq\; (1-\sigma_i)\lambda_i^l[(\nabla g_i^l)^\top \begin{pmatrix} x^j - x^l \\ y^j - y^l \end{pmatrix} + g_i^l] \\ &\leq\; 0,\end{aligned}$$

where the last inequality is justified by (12) and $\sigma_i \in [0,1]$. Thus,

$$\lambda_i^l[(\nabla g_i^l)^\top \begin{pmatrix} x^j - x^l \\ y^j - y^l \end{pmatrix} + g_i^l] \;\leq\; s_i^l,\; i = 1, \ldots, m. \tag{18}$$

If $\lambda_i^l$ is equal to 0, so is $t_i^l$ by its definition and we see that $(y^j, f(x^j, y^j))$ satisfies the constraints (16) with $y = y^j$. If $\lambda_i^l \neq 0$, then (18) can be written as:

$$\nabla g_i(x^l, y^l)^\top \begin{pmatrix} x^j - x^l \\ y^j - y^l \end{pmatrix} + g_i(x^l, y^l) \;\leq\; \frac{s_i^l}{\lambda_i^l} \;=\; t_i^l,$$

which also shows that the constraints (16) hold with $y = y^j$.

Finally, we take any $k \in S$ and assume that $y^k$ is an approximate optimal solution of NLP$(x^k)$ with corresponding inexact Lagrange multipliers $\mu^k$. For every $i \in \{1, \ldots, m\}$, if $\mu_i^k \;\neq\; 0$, from Assumption 2.2, we have for some

$\beta_i \in [0,1]$, that

$$-\frac{1}{1-w^k}(v^k)^\top(y^j - y^k) - \frac{1}{\mu_i^k} z_i^k + \frac{u^k}{m\mu_i^k} w^k \;\le\; -\beta_i[(\nabla g_i^k)^\top \begin{pmatrix} x^j - x^k \\ y^j - y^k \end{pmatrix} + g_i^k],$$

i.e.,

$$-\frac{1}{1-w^k}(v^k)^\top(y^j - y^k) - a_i^k \;\le\; -\beta_i[(\nabla g_i^k)^\top \begin{pmatrix} x^j - x^k \\ y^j - y^k \end{pmatrix} + g_i^k]$$

by the definition of $a_i^k$. Thus, the constraints (17) are satisfied with $y = y^j$. When $\mu_i^k = 0$, it results that $a_i^k = 0$ by its definition and, also by Assumption 2.2, we have that, for some $\eta_i \in [0,1]$,

$$\begin{aligned} &(\nabla g_i^k)^\top \begin{pmatrix} x^j - x^k \\ y^j - y^k \end{pmatrix} + g_i^k - \frac{1}{1-w^k}(v^k)^\top(y^j - y^k) \\ &\le\; (1-\eta_i)[(\nabla g_i^k)^\top \begin{pmatrix} x^j - x^k \\ y^j - y^k \end{pmatrix} + g_i^k] \;\le\; 0. \end{aligned}$$

This also shows that the constraints (17) hold with $y = y^j$.

We can therefore say that $(y^j, f(x^j, y^j))$ is a feasible point of $\mathcal{P}_{x^j}^{\mathrm{OA}}$, and thus $\bar{\alpha}^j \le f(x^j, y^j)$. Next, we will prove that $f(x^j, y^j)$ is also a lower bound, i.e., $\bar{\alpha}^j \ge f(x^j, y^j)$.

**Part II–B. Establishing that $f(x^j, y^j)$ is a lower bound for the optimal value of $\mathcal{P}_{x^j}^{\mathrm{OA}}$ for $j \in T$.**

Recall that $y^j$ is an approximated optimal solution of NLP$(x^j)$ satisfying the inexact KKT conditions (1) and (2). By construction, any solution of $\mathcal{P}_{x^j}^{\mathrm{OA}}$ has to satisfy the inexact outer-approximation constraints:

$$(\nabla_y f(x^j, y^j) - r^j)^\top (y - y^j) + f(x^j, y^j) \;\le\; \alpha, \tag{19}$$

$$\nabla_y g(x^j, y^j)^\top (y - y^j) + g(x^j, y^j) \;\le\; t^j. \tag{20}$$

Multiplying the inequalities (20) by the nonnegative multipliers $\lambda_1^j, \ldots, \lambda_m^j$ and summing them together with (19), one obtains

$$\begin{aligned} &(\nabla_y f(x^j, y^j) - r^j)^\top (y - y^j) + f(x^j, y^j) \\ &+ \sum_{i=1}^m \lambda_i^j (\nabla_y g_i(x^j, y^j)^\top (y - y^j) + g_i(x^j, y^j) - s_i^j) \le\; \alpha. \end{aligned} \tag{21}$$

The left hand side of the inequality (21) can be rewritten as:

$$(\nabla_y f(x^j,y^j)+\sum_{i=1}^{m}\lambda_i^j\nabla_y g_i(x^j,y^j)-r^j)^\top(y-y^j)+\sum_{i=1}^{m}(\lambda_i^j g_i(x^j,y^j)-s_i^j)+f(x^j,y^j).$$

By using (1) and (2), this quantity is equal to $f(x^j,y^j)$, and it follows that inequality (21) is equivalent to $f(x^j,y^j)\le\alpha$.

In conclusion, for any $x^j\in X\cap\mathbb{Z}^{n_d}$ with $j\in T$, problems $\mathcal{P}_{x^j}^{\mathrm{OA}}$ and perturbed NLP($x^j$) have the same optimal value. In other words, the MILP problem $\mathcal{P}^{\mathrm{OA}}$ has the same optimal value as the perturbed problem $\mathcal{P}$ given by (6). ■

Since in the exact case, all the KKT residuals are zero, it results from Theorem 2.1 what is well known for OA:

**Corollary 2.1.** *Let* P *be a convex* MINLP *problem. Assume that* P *is feasible with a finite optimal value and the residuals of the* KKT *conditions of the* NLP *subproblems are zero. Then* $\mathcal{P}^{\mathrm{OA}}$ *and* P *have the same optimal value.*

In the paper [4], the feasibility NLP subproblem is stated as

$$\mathrm{P}_{x^k}^F\begin{cases}\min & \sum_{i=1}^m u_i\\ \text{s.t.} & g(x^k,y)\ \le\ u,\\ & u\ \ge\ 0,\\ & y\in Y, u\in\mathbb{R}^m.\end{cases}$$

Note that one could easily rederive a result similar to [4, Theorem 1] replacing their $\mathrm{P}_{x^k}^F$ by our NLPF($x^k$). In fact, the argument needed here is essentially Part I of the proof of Theorem 2.1 with $v$, $w$, and $z$ set to zero. In their approach, the cuts (9) are included in the master problem, but one can also see that the proof of Theorem 2.1 remains true in this case (it would suffice to observe that (9) is satisfied trivially in the convex case when $y=y^j$ and $\alpha=f(x^j,y^j)$).

**2.2. Inexact-OA algorithm.** One knows that the outer approximation algorithm terminates finitely in the convex case and when the optimal solutions of the NLP subproblems satisfy the first order KKT conditions (see [9]). In this section, we will extend the outer approximation algorithm to the inexact solution of the NLP subproblems by incorporating the corresponding residuals in the cuts of the master problems.

As in the exact case, at each step of the inexact OA algorithm, one tries to solve a subproblem NLP($x^p$), where $x^p$ is chosen as a new discrete assignment. Two results can then occur: either NLP($x^p$) is feasible and an approximated optimal solution $y^p$ can be given, or this subproblem is found infeasible and another NLP subproblem, NLPF($x^p$), is solved, yielding an approximated optimal solution $y^p$. In the algorithm, the sets $T$ and $S$ defined in the Introduction will be replaced by:

$$\begin{aligned} T^p \;=\; & \{j : j \le p,\; x^j \in X \cap \mathbb{Z}^{n_d}, \mathrm{NLP}(x^j) \text{ is feasible and } y^j \text{ appr. solves} \\ & \mathrm{NLP}(x^j)\} \end{aligned}$$

and

$$\begin{aligned} S^p \;=\; & \{k : k \le p,\; x^k \in X \cap \mathbb{Z}^{n_d}, \mathrm{NLP}(x^k) \text{ is infeasible and } y^k \text{ appr. solves} \\ & \mathrm{NLPF}(x^k)\}. \end{aligned}$$

In order to prevent any $x^j$, $j \in T^p$, from becoming the solution of the relaxed master problem to be solved at the $p$–iteration, one needs to add the constraint

$$\alpha \;<\; \mathrm{UBD}^p,$$

where

$$\mathrm{UBD}^p \;=\; \min_{j \,\le\, p, j \in T^p} f(x^j, y^j).$$

Then we define the following inexact relaxed MILP master problem

$$(\mathcal{P}^{\mathrm{OA}})^p \left\{ \begin{array}{ll} \min & \alpha \\ \text{s.t.} & \alpha \;<\; \mathrm{UBD}^p, \\ & \begin{pmatrix} \nabla_x f(x^j, y^j) \\ \nabla_y f(x^j, y^j) - r^j \end{pmatrix}^\top \begin{pmatrix} x - x^j \\ y - y^j \end{pmatrix} + f(x^j, y^j) \;\le\; \alpha, \\ & \nabla g(x^j, y^j)^\top \begin{pmatrix} x - x^j \\ y - y^j \end{pmatrix} + g(x^j, y^j) \;\le\; t^j, \;\; \forall j \in T^p, \\ & \begin{pmatrix} \nabla_x g_i(x^k, y^k) \\ \nabla_y g_i(x^k, y^k) - \frac{1}{1 - w^k} v^k \end{pmatrix}^\top \begin{pmatrix} x - x^k \\ y - y^k \end{pmatrix} + g_i(x^k, y^k) \;\le\; a_i^k, \\ & i = 1, \ldots, m, \;\; \forall k \in S^p, \\ & x \in X \cap \mathbb{Z}^{n_d}, y \in Y, \alpha \in \mathbb{R}, \end{array} \right.$$

where $t^j$ and $a_i^k$ were defined in (7) and (8), respectively. The presentation of the inexact OA algorithm (given next) and the proof of its finiteness in Theorem 2.2 follows the lines in [9].

**Algorithm 2.1 (Inexact Outer Approximation).**

**Initialization**

Let $x^0$ be given. Set $p = 0, T^{-1} = \emptyset, S^{-1} = \emptyset$, and UBD $= +\infty$.

**REPEAT**

(1) Inexactly solve the subproblem NLP($x^p$), or the feasibility subproblem NLPF($x^p$) provided NLP($x^p$) is infeasible, and let $y^p$ be an approximated optimal solution. At the same time, obtain the corresponding inexact Lagrange multipliers $\lambda^p$ of NLP($x^p$) (resp. $\mu^p$ of NLPF($x^p$)). Evaluate the residuals $r^p$ and $s^p$ of NLP($x^p$) (resp. $v^p$, $w^p$, and $z^p$ of NLPF($x^p$)).

(2) Linearize the objective functions and constraints at $(x^p, y^p)$. Renew $T^p = T^{p-1} \cup \{p\}$ or $S^p = S^{p-1} \cup \{p\}$.

(3) If (NLP($x^p$) is feasible and $f^p <$ UBD), then update current best point by setting $\bar{x} = x^p, \bar{y} = y^p$, and UBD $= f^p$.

(4) Solve the relaxed master problem $(\mathcal{P}^{\mathrm{OA}})^p$, obtaining a new discrete assignment $x^{p+1}$ to be tested in the algorithm. Increment $p$ by one unit.

**UNTIL** ($(\mathcal{P}^{\mathrm{OA}})^p$ is infeasible).

If termination occurs with UBD $= +\infty$ , then the algorithm visited every discrete assignment $x \in X \cap \mathbb{Z}^{n_d}$ but did not obtain a feasible point for the original MINLP problem P, or perturbed version $\mathcal{P}$. In this case, the MINLP is declared infeasible. Next, we will show that the inexact OA algorithm also terminates in a finite number of steps.

**Theorem 2.2.** *Let* P *be a convex* MINLP *problem and* $\mathcal{P}$ *be its perturbed problem as defined in the Introduction. Assume that either* P *has a finite optimal value or is infeasible, and that the residuals of the* KKT *conditions of the* NLP *subproblems satisfy Assumptions 2.1 and 2.2. Then Algorithm 2.1 terminates in a finite number of steps at an optimal solution of* $\mathcal{P}$ *or with an indication that* $\mathcal{P}$ *is infeasible.*

*Proof*: Since the set $X$ is bounded by assumption, finite termination of Algorithm 2.1 will be established by proving that no discrete assignment is generated twice by the algorithm.

Let $q \leq p$. If $q \in S^p$, it has been shown in Part I of the proof of Theorem 2.1 that the corresponding constraint in $\mathcal{P}^{\mathrm{OA}}_{x^p}$, derived from the feasibility problem NLPF($x^q$), cannot be satisfied, showing that $x^q$ cannot be feasible for $(\mathcal{P}^{\mathrm{OA}})^p$.

We will now show that $x^q$ cannot be feasible for $(\mathcal{P}^{\mathrm{OA}})^p$ when $q \in T^p$. For this purpose, let us assume that $x^q$ is feasible in $(\mathcal{P}^{\mathrm{OA}})^p$ and try to reach a contradiction. Let $y^q$ be an approximated optimal solution of NLP($x^q$) satisfying the inexact KKT conditions, that is, there exist $\lambda^q \in \mathbb{R}^m_+$, $r^q \in \mathbb{R}^{n_c}$, and $s^q \in \mathbb{R}^m$, such that

$$\begin{aligned} \nabla_y f^q + \sum_{i=1}^m \lambda_i^q \nabla_y g_i(x^q, y^q) &= r^q, \quad (22)\\ \lambda_i^q g_i(x^q, y^q) &= s_i^q, \quad i = 1, \ldots, m. \quad (23) \end{aligned}$$

If $x^q$ would be feasible for $(\mathcal{P}^{\mathrm{OA}})^p$ it would satisfy the following set of inequalities for some $y$:

$$\begin{aligned} \alpha^p \;<\; \mathrm{UBD}^p &\leq f^q, \quad (24)\\ \begin{pmatrix} \nabla_x f^q \\ \nabla_y f^q - r^q \end{pmatrix}^\top \begin{pmatrix} 0 \\ y - y^q \end{pmatrix} + f^q &\leq \alpha^p, \quad (25)\\ (\nabla g^q)^\top \begin{pmatrix} 0 \\ y - y^q \end{pmatrix} + g^q &\leq t^q, \quad (26) \end{aligned}$$

where, for $i = 1, \ldots, m$,

$$t_i^q = \begin{cases} \frac{s_i^q}{\lambda_i^q}, & \text{if } \lambda_i^q > 0, \\ 0, & \text{if } \lambda_i^q = 0. \end{cases}$$

Multiplying the rows in (26) by the Lagrange multipliers $\lambda_i^q \geq 0$, $i = 1, \ldots, m$, and adding (25), we obtain that

$$\begin{aligned} (\nabla_y f^q - r^q)^\top (y - y^q) + f^q + \sum_{i=1}^m \lambda_i^q \nabla_y g_i(x^q, y^q)^\top (y - y^q) + \sum_{i=1}^m \lambda_i^q g_i^q \\ \leq \alpha^p + \sum_{i=1}^m \lambda_i^q t_i^q, \end{aligned}$$

which, by the definition of $t^q$, is equivalent to

$$(\nabla_y f^q - r^q)^\top (y - y^q) + f^q + \sum_{i=1}^{m} \lambda_i^q \nabla_y g_i(x^q, y^q)^\top (y - y^q) + \sum_{i=1}^{m} (\lambda_i^q g_i^q - s_i^q) \\ \leq \alpha^p.$$

The left hand side of this inequality can be written as:

$$[\nabla_y f^q - r^q + \sum_{i=1}^{m} \lambda_i^q \nabla_y g_i(x^q, y^q)]^\top (y - y^q) + \sum_{j=1}^{m} (\lambda_i^q g_i^q - s_i^q) + f^q.$$

Using (22) and (23), this is equal to $f^q$ and therefore we obtain the inequality

$$f^q \leq \alpha^p,$$

which contradicts (24).

The rest of the proof is exactly as in [9, Theorem 2] but we repeat here for completeness and possible changes in notation. Finally, we will show that Algorithm 2.1 always terminates at a solution of $\mathcal{P}$ or with an indication that $\mathcal{P}$ is infeasible (which occurs when UBD $= +\infty$ at the exit). If $\mathcal{P}$ is feasible, then let $(x^*, y^*)$ be an optimal solution of $\mathcal{P}$ with optimal value $f^*$. Without loss of generality, we will not distinguish between $(x^*, y^*)$ and any other optimal solution with the same objective value $f^*$. Note that from Theorem 2.1, $(x^*, y^*, f^*)$ is also an optimal solution of $\mathcal{P}^{\mathrm{OA}}$. Now assume that the algorithm terminates indicating a nonoptimal point $(x', y')$ with $f' > f^*$. In such a situation, the previous relaxation of the master problem $\mathcal{P}^{\mathrm{OA}}$ after adding the constraints at the point $(x', y', f')$, called $(\mathcal{P}^{\mathrm{OA}})^p$, is infeasible, causing the above mentioned termination. We will get a contradiction by showing that $(x^*, y^*, f^*)$ is feasible for $(\mathcal{P}^{\mathrm{OA}})^p$. First, by the assumption that UBD $= f' > f^*$, the first constraint $\alpha = f^* <$ UBD of $(\mathcal{P}^{\mathrm{OA}})^p$ holds. Secondly, since $(x^*, y^*, f^*)$ is an optimal solution to $\mathcal{P}^{\mathrm{OA}}$, it must be feasible for all other constraints of $(\mathcal{P}^{\mathrm{OA}})^p$. Therefore, the algorithm could not terminate at $(x', y')$ with UBD $= f'$. ∎

## 3. Inexact generalized Benders decomposition

### 3.1. Equivalence between perturbed and master problems for GBD.

In the generalized Benders decomposition (GBD), the MILP master problem involves only the discrete variables. When considering the inexact case, the master problem of GBD is the following:

$$
\mathcal{P}^{\text{GBD}} \begin{cases} \min & \alpha \\ \text{s.t.} & f(x^j, y^j) + \nabla_x f(x^j, y^j)^\top (x - x^j) \\ & + \sum_{i=1}^m \lambda_i^j \nabla_x g_i(x^j, y^j)^\top (x - x^j) \ \leq \ \alpha, \ \forall j \in T, \\ & \sum_{i=1}^m \mu_i^k [g_i(x^k, y^k) + \nabla_x g_i(x^k, y^k)^\top (x - x^k)] + w^k u^k \\ & - \sum_{i=1}^m z_i^k \ \leq \ 0, \ \forall k \in S, \\ & x \in X \cap \mathbb{Z}^{n_d}, \alpha \in \mathbb{R}. \end{cases}
$$

One can easily recognize the classical form of (exact) GBD master problem when $w^k = 0$ and $z^k = 0$. Moreover, as we show in the Appendix, this MILP can also be derived in the inexact case from a perturbed duality representation of the original, perturbed problem.

A proof similar to the one of exact GBD and exact and inexact OA (Theorem 2.1) allows us to establish the desired equivalence between the original, perturbed MINLP problem and the MILP master problem $\mathcal{P}^{\text{GBD}}$.

**Theorem 3.1.** *Let* P *be a convex* MINLP *problem and* $\mathcal{P}$ *be its perturbed problem as defined in the Introduction. Assume that* P *is feasible with a finite optimal value and that the residuals of the* KKT *conditions of the* NLP *subproblems satisfy Assumptions 2.1 and 2.2. Then* $\mathcal{P}^{\text{GBD}}$ *and* $\mathcal{P}$ *have the same optimal value.*

*Proof*: Given any $x^l \in X \cap \mathbb{Z}^{n_d}$ with $l \in T \cup S$, let $\mathcal{P}^{\text{GBD}}_{x^l}$ denote the problem in $\alpha$ obtained from $\mathcal{P}^{\text{GBD}}$ when $x$ is fixed to $x^l$. First we will prove that problem $\mathcal{P}^{\text{GBD}}_{x^k}$ is infeasible for every $k \in S$. When we set $x = x^k$, in the corresponding constraint of $\mathcal{P}^{\text{GBD}}$, one obtains

$$
\sum_{i=1}^m \mu_i^k g_i(x^k, y^k) + w^k u^k - \sum_{i=1}^m z_i^k \ \leq \ 0
$$

From (4) and (5), it results that $u^k \leq 0$, but one knows that $u^k$ is strictly positive when NLP$(x^k)$ is infeasible.

Next, we will prove that for each $x^j \in X \cap \mathbb{Z}^{n_d}$, with $j \in T$, $\mathcal{P}^{\text{GBD}}_{x^j}$ has the same optimal value as the perturbed NLP$(x^j)$. First, we will prove that the following constraints of $\mathcal{P}^{\mathrm{GBD}}_{x^j}$

$$\begin{aligned} f(x^l,y^l) + \nabla_x f(x^l,y^l)^\top (x^j - x^l) + \sum_{i=1}^m \lambda_i^l \nabla_x g_i(x^l,y^l)^\top (x^j - x^l) \;&\le\; \alpha, \\ \forall l \in T, \qquad & (27) \\ \sum_{i=1}^m \mu_i^k [g_i(x^k,y^k) + \nabla_x g_i(x^k,y^k)^\top (x^j - x^k)] + w^k u^k - \sum_{i=1}^m z_i^k \;&\le\; 0, \\ \forall k \in S. \qquad & (28) \end{aligned}$$

are satisfied with $\alpha = f(x^j,y^j)$. Under Assumptions 2.1 and Assumption 2.2, we know from the proof of Theorem 2.1 (Part II–A) that the following hold: (15) with $y = y^j$ and $\alpha = f(x^j,y^j)$, (16) with $y = y^j$, and (17) with $y = y^j$.

When $l \in T$, multiplying the inequalities (16) with $y = y^j$ by the nonnegative multipliers $\lambda_1^j, \ldots, \lambda_m^j$ and summing them together with (15) with $y = y^j$ and $\alpha = f(x^j,y^j)$, one obtains

$$\begin{aligned} & f(x^l,y^l) + \nabla_x f(x^l,y^l)^\top (x^j - x^l) + \sum_{i=1}^m \lambda_i^l \nabla_x g_i(x^l,y^l)^\top (x^j - x^l) \\ \le\;& f(x^j,y^j) - [\nabla_y f(x^l,y^l) + \sum_{i=1}^m \lambda_i^l \nabla_y g_i(x^l,y^l) - r^l]^\top (y^j - y^l) \\ & - \sum_{i=1}^m \lambda_i^l g(x^l,y^l) + \sum_{i=1}^m \lambda_i^l t_i^l. \end{aligned}$$

The right hand side is equal to $f(x^j,y^j)$ by the definitions of $r^l, s^l$, and $t^l$, showing that (27) holds with $\alpha = f(x^j,y^j)$.

When $k \in S$, multiplying the inequalities in (17) with $y = y^j$ by the nonnegative multipliers $\mu_1^k, \ldots, \mu_m^k$, and summing them up, one obtains using (3) and (4)

$$\sum_{i=1}^m \mu_i^k \nabla_x g_i(x^k,y^k)^\top (x^j - x^k) + \sum_{i=1}^m \mu_i^k g_i(x^k,y^k) \;\le\; \sum_{i=1}^m \mu_i^k a_i^k,$$

which, by the definition of $a^k$, is the same as (28).

Thus, $f(x^j,y^j)$ is a feasible point of $\mathcal{P}^{\mathrm{GBD}}_{x^j}$, and therefore $f(x^j,y^j)$ is an upper bound on the optimal value $\bar{\alpha}^j$ of $\mathcal{P}^{\mathrm{OA}}_{x^j}$. To show that is also a lower bound, i.e., that $\bar{\alpha}^j \geq f(x^j, y^j)$, note that from (27), when $l = j$, $\mathcal{P}_{x^j}^{\mathrm{GBD}}$ contains the constraint:

$$f(x^j, y^j) \; \leq \; \alpha.$$

We have thus proved that for any $x^j \in X \cap \mathbb{Z}^{n_d}$, with $j \in T$, problems $\mathcal{P}_{x^j}^{\mathrm{GBD}}$ and perturbed NLP$(x^j)$ have the same optimal value, which concludes the proof. ■

When all the KKT residuals are zero we obtain as a corollary the known equivalence result in GBD:

**Corollary 3.1.** *Let* P *be a convex* MINLP *problem. Assume that* P *is feasible with a finite optimal value and the residuals of the* KKT *conditions of the* NLP *subproblems are zero. Then* $\mathcal{P}^{\mathrm{GBD}}$ *and* P *have the same optimal value.*

**Remark 3.1.** *It is well known that the constraints of the* GBD *master problem can be derived from the corresponding ones of the* OA *master problem, in the convex, exact case (see* [15]*). The same happens naturally in the inexact case. In fact, from the proof of Theorem 3.1 above, we can see that the constraints in* $\mathcal{P}_{x^j}^{\mathrm{OA}}$*, for* $j \in T$*, imply the corresponding ones in* $\mathcal{P}_{x^j}^{\mathrm{GBD}}$*. Moreover, one can easily see that any of the constraints in* $\mathcal{P}^{\mathrm{OA}}$ *imply the corresponding ones in* $\mathcal{P}^{\mathrm{GBD}}$*.*

*Thus, one can also say in the inexact case that the lower bounds produced iteratively by the* OA *algorithm are stronger than the ones provided by the corresponding* GBD *algorithm (given next).*

**3.2. Inexact GBD algorithm.** As we know for exact GBD, it is possible to derive an algorithm for the inexact case, terminating finitely, by solving at each iteration a relaxed MILP formed by the cuts collected so far. The definitions of UBD$^p$, $T^p$, and $S^p$ are the same as those in Section 2.2. The relaxed MILP to be solved at each iteration is thus given by

$$
(\mathcal{P}^{\mathrm{GBD}})^p \left\{
\begin{array}{ll}
\min & \alpha \\
\text{s.t.} & \alpha \ < \ \mathrm{UBD}^{\mathrm{p}} \\
& f(x^j, y^j) + \nabla_x f(x^j, y^j)^\top (x - x^j) \\
& + \sum_{i=1}^m \lambda_i^j \nabla_x g_i(x^j, y^j)^\top (x - x^j) \ \leq \ \alpha, \ \forall j \in T^p \\
& \sum_{i=1}^m \mu_i^k [g_i(x^k, y^k) + \nabla_x g_i(x^k, y^k)^\top (x - x^k)] + w^k u^k \\
& - \sum_{i=1}^m z_i^k \ \leq \ 0, \ \forall k \in S^p \\
& x \in X \cap \mathbb{Z}^{n_d}, \alpha \in \mathbb{R}.
\end{array}
\right.
$$

The inexact GBD algorithm is given next (and follows the presentation in [9] for OA).

**Algorithm 3.1** (**Inexact GBD Approximation).**

**Initialization**

Let $x^0$ be given. Set $p = 0, T^{-1} = \emptyset, S^{-1} = \emptyset$, and $\mathrm{UBD} = +\infty$.

**REPEAT**

(1) Inexactly solve the subproblem NLP($x^p$), or the feasibility subproblem NLPF($x^p$) provided NLP($x^p$) is infeasible, and let $y^p$ be an approximated optimal solution. At the same time, obtain the corresponding inexact Lagrange multipliers $\lambda^p$ of NLP($x^p$) (resp. $\mu^p$ of NLPF($x^p$)). Evaluate the residuals $r^p$ and $s^p$ of NLP($x^p$) (resp. $v^p$, $w^p$, and $z^p$ of NLPF($x^p$)).

(2) Linearize the objective functions and constraints at $x^p$. Renew $T^p = T^{p-1} \cup \{p\}$ or $S^p = S^{p-1} \cup \{p\}$.

(3) If (NLP($x^p$) is feasible and $f^p < \mathrm{UBD}$), then update current best point by setting $\bar{x} = x^p, \bar{y} = y^p$, and $\mathrm{UBD} = f^p$.

(4) Solve the relaxed master problem $(\mathcal{P}^{\mathrm{GBD}})^p$, obtaining a new discrete assignment $x^{p+1}$ to be tested in the algorithm. Increment $p$ by one unit.

**UNTIL** ($(\mathcal{P}^{\mathrm{GBD}})^p$ is infeasible).

Similarly as in Theorem 2.2 for OA, one can establish that the above inexact GBD algorithm terminates in a finite number of steps.

**Theorem 3.2.** *Let* P *be a convex* MINLP *problem and* $\mathcal{P}$ *be its perturbed problem as defined in the Introduction. Assume that either* P *has a finite optimal value or is infeasible, and that the residuals of the* KKT *conditions of the* NLP *subproblems satisfy Assumptions 2.1 and 2.2. Then Algorithm 3.1 terminates in a finite number of steps at an optimal solution of* $\mathcal{P}$ *or with an indication that* $\mathcal{P}$ *is infeasible.*

## 4. Numerical experiments

We will illustrate some of the practical features of inexact OA and GBD algorithms by reporting numerical results on three test problems: Example 1, taken from [8, test problem no. 1], has 3 discrete variables and 3 continuous variables; Example 2, taken from [8, test problem no. 2], has 5 discrete variables and 6 continuous variables; Example 3, taken from [8, test problem no. 3], and has 8 discrete variables and 9 continuous variables. All the three examples are convex and linear in the discrete variables, and consist of simplified versions of process synthesis problems. In the third example we found a point better than the one in [8]: $x^* = (0, 2, 0.46782, 0.58477, 2, 0, 0, 0.26667, 1.25144)^\top$, $y^* = (0, 1, 0, 1, 0, 1, 0, 1)^\top$ with corresponding optimal value $f^* = 44.6764$.

The implementation and testing of Algorithms 2.1 and 3.1 was made in MATLAB (version 7.11.0, R2010b). We used `fmincon` (from MATLAB) to solve the NLP subproblems and `ip1` [17] to solve the MILP problems, arising in both algorithms. The linear equality constraints possibly present in the original problems were kept in the MILP master problems.

For both methods, we report results for two variants, depending on the form of the cuts. In a first variant the subproblems are solved inexactly (with tolerances varying from $10^{-6}$ to $10^{-1}$) but the cuts are the exact ones. When the tolerance is set to $10^{-6}$ we are essentially running exact OA and GBD. The second variant also incorporates inexact solution of NLP subproblems (again with tolerances varying from $10^{-6}$ to $10^{-1}$) but the cuts are now the inexact ones.

In the tables of results we report the number $N$ of iterations taken by Algorithms 2.1 and 3.1. We also report, in the tables corresponding to the second variant, the number $C$ of constraint inequalities of Assumptions 2.1 and 2.2 that were violated by more than $10^{-8}$. The stopping criteria of both algorithms consisted of the corresponding master program being infeasible or the number of iterations exceeding 50 or the solution of the MILP master

TABLE 1. Application of OA (inexact solution of NLP subproblems and exact cuts) to Example 1. The table reports the number $N$ of iterations taken.

| | Tolerances | | | | | |
|---|---|---|---|---|---|---|
| initial point | 1e-6 | 1e-5 | 1e-4 | 1e-3 | 1e-2 | 1e-1 |
| $(0,0,0)^\top$ | 2 | 2 | 2 | 2 | 2 | NPC |
| $(1,0,0)^\top$ | 2 | 2 | 2 | 2 | 2 | NPC |
| $(0,1,0)^\top$ | 2 | 2 | 2 | 2 | 2 | NPC |
| $(1,0,1)^\top$ | 2 | 2 | 2 | 2 | 2 | NPC |
| $(0,1,1)^\top$ | 3 | 3 | 3 | 3 | 3 | NPC |
| $(0,0,1)^\top$ | 2 | 2 | 2 | 2 | 2 | NPC |

NPC stands for no proper convergence (repeated solution of MILP master problem).

program coinciding with a previous one (these third cases were marked with NPC, standing for no proper convergence). We note that in all the NPC cases found, the repeated integer solution of the MILP master problem was indeed the solution of the original MINLP.

**4.1. Results for inexact OA method.** Tables 1–6 summarize the application of inexact OA (Algorithm 2.1) (variant inexact solution of NLP subproblems and exact cuts, and variant inexact solution of NLP subproblems and inexact cuts) to Examples 1–3. Comparing Tables 1 and 2, one can see little difference between using exact or inexact cuts for the first, smaller example. However, looking at Tables 3 and 4 for the second example and Tables 5 and 6 for the third example, one can see, for larger values of the tolerances, that the inexact case with exact cuts has more tendency for unproper convergence (i.e., the MILP is incapable of either provide a new integer solution or render infeasible), while the variant incorporating the inexactness in the cuts does not. We also observe that inexact OA converged even neglecting the imposition of the inequalities of Assumptions 2.1 and 2.2.

**4.2. Results for inexact GBD method.** Tables 7–12 summarize the application of inexact GBD (Algorithm 3.1) (variant inexact solution of NLP subproblems and exact cuts, and variant inexact solution of NLP subproblems and inexact cuts) to Examples 1–3. One can observe that there is little

TABLE 2. Application of OA (inexact solution of NLP subproblems and inexact cuts) to Example 1. The table reports the number $N$ of iterations taken as well as the number $C$ of inequalities violated in Assumptions 2.1 and 2.2.

| | Tolerances | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1e-6 | | 1e-5 | | 1e-4 | | 1e-3 | | 1e-2 | | 1e-1 | |
| initial point | $N$ | $C$ | $N$ | $C$ | $N$ | $C$ | $N$ | $C$ | $N$ | $C$ | $N$ | $C$ |
| $(0,0,0)^\top$ | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 | NPC | 0 |
| $(1,0,0)^\top$ | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 | NPC | 0 |
| $(0,1,0)^\top$ | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 | NPC | 0 |
| $(1,0,1)^\top$ | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 | NPC | 0 |
| $(0,1,1)^\top$ | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0 | NPC | 0 |
| $(0,0,1)^\top$ | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 | NPC | 0 |

The maximum number for $C$ is $13t(t-1)/2 + 12st$, where $t = |T|$, $s = |S|$, and $s + t = N$.

NPC stands for no proper convergence (repeated solution of MILP master problem).

difference between the two variants since the 'exact' cuts in the first variant already incorporate inexact information coming from the inexact Lagrange multipliers. One observes that inexact GBD takes more iterations than inexact OA in these examples, which, according to Remark 3.1, is expected since inexact GBD yields weaker lower bounds and hence generally requires more major iterations to converge than inexact OA. The number of inequalities of Assumptions 2.1 and 2.2 violated in inexact GBD is also higher than the one in inexact OA.

## 5. Conclusions and final remarks

In this paper we have attempted to gain a better understanding of the effect of inexactness when solving NLP subproblems in two well known decomposition techniques for Mixed Integer Nonlinear Programming (MINLP), the outer approximation (OA) and the generalized Benders decomposition (GBD).

As pointed out to us by I. E. Grossmann, solving the NLP subproblems inexactly in OA positions this approach somewhere in between exact OA and the extended cutting plane method [18].

TABLE 3. Application of OA (inexact solution of NLP subproblems and exact cuts) to Example 2. The table reports the number $N$ of iterations taken.

| initial point | Tolerances 1e-6 | 1e-5 | 1e-4 | 1e-3 | 1e-2 | 1e-1 |
|---|---|---|---|---|---|---|
| $(1,0,0,0,0)^\top$ | 3 | 3 | 3 | NPC | NPC | NPC |
| $(0,1,0,0,0)^\top$ | 2 | 2 | 2 | NPC | NPC | NPC |
| $(1,0,1,0,0)^\top$ | 3 | 3 | 3 | NPC | NPC | NPC |
| $(1,0,0,1,0)^\top$ | 2 | 2 | 2 | NPC | NPC | NPC |
| $(1,0,0,0,1)^\top$ | 3 | 3 | 3 | NPC | NPC | NPC |
| $(0,1,1,0,0)^\top$ | 2 | 2 | 2 | NPC | NPC | NPC |
| $(0,1,0,1,0)^\top$ | 3 | 3 | 3 | NPC | NPC | NPC |
| $(0,1,0,0,1)^\top$ | 2 | 2 | 2 | NPC | NPC | NPC |
| $(1,0,1,1,0)^\top$ | 3 | 3 | 3 | NPC | NPC | NPC |
| $(1,0,1,0,1)^\top$ | 3 | 3 | 3 | NPC | NPC | NPC |
| $(0,1,1,1,0)^\top$ | 2 | 2 | 2 | NPC | NPC | NPC |
| $(0,1,1,0,1)^\top$ | 2 | 2 | 2 | NPC | NPC | NPC |

NPC stands for no proper convergence (repeated solution of MILP master problem).

Although all the conditions required on the residuals of the inexact KKT conditions can be imposed, one can see from Assumptions 2.1 and 2.2 that the complete satisfaction of all those inequalities would ask for repeated NLP subproblem solution for all the previous addressed discrete assignments. Such requirement would then undermine the practical purpose of saving computational effort aimed by the NLP subproblem inexactness. In our preliminary numerical tests we disregarded the conditions of Assumptions 2.1 and 2.2 and verified, after terminating each run of inexact OA or GBD, how many of them were violated. The results indicated that proper convergence can be achieved without imposing Assumptions 2.1 and 2.2, and that the number of violated inequalities was relatively low. The results also seem to indicate that the cuts in OA and GBD must be changed accordingly when the corresponding NLP subproblems are solved inexactly. Testing these inexact approaches in a wider test set of larger problems is out of the scope of this paper, although it seems a necessary step to further validate these indications.

TABLE 4. Application of OA (inexact solution of NLP subproblems and inexact cuts) to Example 2. The table reports the number $N$ of iterations taken as well as the number $C$ of inequalities violated in Assumptions 2.1 and 2.2.

| | Tolerances | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1e-6 | | 1e-5 | | 1e-4 | | 1e-3 | | 1e-2 | | 1e-1 | |
| initial point | $N$ | $C$ | $N$ | $C$ | $N$ | $C$ | $N$ | $C$ | $N$ | $C$ | $N$ | $C$ |
| $(1,0,0,0,0)^\top$ | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0 |
| $(0,1,0,0,0)^\top$ | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 |
| $(1,0,1,0,0)^\top$ | 3 | 1 | 3 | 1 | 3 | 1 | 3 | 1 | 3 | 1 | 3 | 1 |
| $(1,0,0,1,0)^\top$ | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 |
| $(1,0,0,0,1)^\top$ | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0 |
| $(0,1,1,0,0)^\top$ | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| $(0,1,0,1,0)^\top$ | 3 | 1 | 3 | 1 | 3 | 1 | 3 | 1 | 3 | 1 | 3 | 1 |
| $(0,1,0,0,1)^\top$ | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 |
| $(1,0,1,1,0)^\top$ | 3 | 1 | 3 | 1 | 3 | 1 | 3 | 1 | 3 | 1 | 3 | 1 |
| $(1,0,1,0,1)^\top$ | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0 |
| $(0,1,1,1,0)^\top$ | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 |
| $(0,1,1,0,1)^\top$ | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 |

The maximum number for $C$ is $27t(t-1)/2 + 26st$, where $t = |T|$, $s = |S|$ and $s + t = N$.

Our study was performed under the assumption of convexity of the functions involved. Moreover, we also assumed that the approximated optimal solutions of the NLP subproblems were feasible in these subproblems, and that the corresponding inexact Lagrange multipliers were nonnegative. Relaxing these assumptions introduces another layer of difficulty but certainly deserves proper attention in the future.

# Appendix A

**A.1. Inexact KKT conditions and perturbed problems.** As we said in the Introduction of the paper, the point $y^j$ satisfying the inexact KKT conditions (1)–(2) of the subproblem NLP$(x^j)$ can be interpreted as a solution of a perturbed NLP subproblem, which has the form

$$\text{perturbed NLP}(x^j) \begin{cases} \min & f(x^j, y) - (r^j)^\top (y - y^j) \\ \text{s.t.} & g(x^j, y) - t^j \ \le \ 0, \\ & y \in Y, \end{cases}$$

where $t^j$ is given by (7). The data of this perturbed subproblem depends, however, on the approximated optimal solution $y^j$ and inexact Lagrange multipliers $\lambda^j$. Similarly, the

Table 5. Application of OA (inexact solution of NLP subproblems and exact cuts) to Example 3. The table reports the number $N$ of iterations taken.

| initial point | Tolerances 1e-6 | 1e-5 | 1e-4 | 1e-3 | 1e-2 | 1e-1 |
|---|---|---|---|---|---|---|
| $(1,0,0,0,0,0,0,0)^\top$ | 3 | 3 | 3 | NPC | NPC | NPC |
| $(1,0,0,0,0,0,0,1)^\top$ | 2 | 2 | 2 | NPC | NPC | NPC |
| $(1,0,1,0,0,0,0,1)^\top$ | 1 | 1 | 1 | NPC | NPC | NPC |
| $(1,0,0,0,1,0,0,0)^\top$ | 3 | 3 | 3 | NPC | NPC | NPC |
| $(1,0,0,0,1,0,0,1)^\top$ | 2 | 2 | 2 | NPC | NPC | NPC |
| $(1,0,1,0,1,0,0,1)^\top$ | 1 | 1 | 1 | NPC | NPC | NPC |
| $(1,0,0,1,0,1,0,0)^\top$ | 3 | 3 | 3 | NPC | NPC | NPC |
| $(1,0,0,1,0,1,0,1)^\top$ | 2 | 2 | 2 | NPC | NPC | NPC |
| $(1,0,0,1,0,0,1,0)^\top$ | 3 | 3 | 3 | NPC | NPC | NPC |
| $(1,0,0,1,0,0,1,1)^\top$ | 1 | 1 | 1 | NPC | NPC | NPC |
| $(0,1,0,0,0,0,0,0)^\top$ | 3 | 3 | 3 | NPC | NPC | NPC |
| $(0,1,0,0,0,0,0,1)^\top$ | 2 | 2 | 2 | NPC | NPC | NPC |
| $(0,1,1,0,0,0,0,1)^\top$ | 2 | 2 | 2 | NPC | NPC | NPC |
| $(0,1,0,0,1,0,0,0)^\top$ | 3 | 3 | 3 | NPC | NPC | NPC |
| $(0,1,0,0,1,0,0,1)^\top$ | 2 | 2 | 2 | NPC | NPC | NPC |
| $(0,1,1,0,1,0,0,1)^\top$ | 2 | 2 | 2 | NPC | NPC | NPC |
| $(0,1,0,1,0,1,0,0)^\top$ | 3 | 3 | 3 | NPC | NPC | NPC |
| $(0,1,0,1,0,1,0,1)^\top$ | 1 | 1 | 1 | NPC | NPC | NPC |
| $(0,1,1,1,0,1,0,1)^\top$ | 2 | 2 | 2 | NPC | NPC | NPC |
| $(0,1,0,1,0,0,1,0)^\top$ | 3 | 3 | 3 | NPC | NPC | NPC |
| $(0,1,0,1,0,0,1,1)^\top$ | 2 | 2 | 2 | NPC | NPC | NPC |
| $(1,0,1,1,0,1,0,1)^\top$ | 2 | 2 | 2 | NPC | NPC | NPC |

NPC stands for no proper convergence (repeated solution of MILP master problem).

point $y^k$ satisfying the inexact KKT conditions (3)–(5) of the subproblem NLPF($x^k$) can be interpreted as a solution of the following perturbed NLP subproblem

$$\text{perturbed NLPF}(x^k)\begin{cases} \min & u - w^k(u-u^k) - (v^k)^\top(y-y^k) \\ \text{s.t.} & g_i(x^k,y) - u - c_i^k \;\leq\; 0,\; i=1,\dots,m, \\ & y \in Y, u \in \mathbb{R}, \end{cases}$$

TABLE 6. Application of OA (inexact solution of NLP subproblems and inexact cuts) to Example 3. The table reports the number $N$ of iterations taken as well as the number $C$ of inequalities violated in Assumptions 2.1 and 2.2.

| | Tolerances | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1e-6 | | 1e-5 | | 1e-4 | | 1e-3 | | 1e-2 | | 1e-1 | |
| initial point | $N$ | $C$ | $N$ | $C$ | $N$ | $C$ | $N$ | $C$ | $N$ | $C$ | $N$ | $C$ |
| $(1,0,0,0,0,0,0,0)^\top$ | 3 | 1 | 3 | 1 | 3 | 1 | 3 | 2 | 3 | 2 | 3 | 1 |
| $(1,0,0,0,0,0,0,1)^\top$ | 2 | 0 | 2 | 0 | 2 | 1 | 2 | 2 | 2 | 3 | 2 | 2 |
| $(1,0,1,0,0,0,0,1)^\top$ | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 2 | 2 |
| $(1,0,0,0,1,0,0,0)^\top$ | 3 | 1 | 3 | 1 | 3 | 1 | 3 | 1 | 3 | 2 | 3 | 1 |
| $(1,0,0,0,1,0,0,1)^\top$ | 2 | 0 | 2 | 0 | 2 | 1 | 2 | 2 | 2 | 3 | 2 | 2 |
| $(1,0,1,0,1,0,0,1)^\top$ | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 2 | 2 |
| $(1,0,0,1,0,1,0,0)^\top$ | 3 | 1 | 3 | 1 | 3 | 1 | 3 | 2 | 3 | 2 | 3 | 1 |
| $(1,0,0,1,0,1,0,1)^\top$ | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | 3 | 3 |
| $(1,0,0,1,0,0,1,0)^\top$ | 3 | 1 | 3 | 1 | 3 | 1 | 3 | 2 | 3 | 2 | 3 | 1 |
| $(1,0,0,1,0,0,1,1)^\top$ | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 2 | 3 |
| $(0,1,0,0,0,0,0,0)^\top$ | 3 | 1 | 3 | 1 | 3 | 1 | 3 | 1 | 3 | 1 | 4 | 4 |
| $(0,1,0,0,0,0,0,1)^\top$ | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 4 | 3 | 7 |
| $(0,1,1,0,0,0,0,1)^\top$ | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 1 | 3 | 3 |
| $(0,1,0,0,1,0,0,0)^\top$ | 3 | 1 | 3 | 1 | 3 | 1 | 3 | 1 | 3 | 1 | 4 | 4 |
| $(0,1,0,0,1,0,0,1)^\top$ | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 4 | 3 | 7 |
| $(0,1,1,0,1,0,0,1)^\top$ | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 1 | 3 | 3 |
| $(0,1,0,1,0,1,0,0)^\top$ | 3 | 1 | 3 | 1 | 3 | 1 | 3 | 1 | 3 | 1 | 4 | 4 |
| $(0,1,0,1,0,1,0,1)^\top$ | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 3 |
| $(0,1,1,1,0,1,0,1)^\top$ | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 3 | 2 | 3 | 2 | 3 |
| $(0,1,0,1,0,0,1,0)^\top$ | 3 | 1 | 3 | 1 | 3 | 1 | 3 | 1 | 3 | 1 | 4 | 4 |
| $(0,1,0,1,0,0,1,1)^\top$ | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 3 | 5 |
| $(1,0,1,1,0,1,0,1)^\top$ | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 | 3 | 3 |

The maximum number for $C$ is $21t(t-1)+41st$, where $t=|T|$, $s=|S|$, and $s+t=N$.

where, for $i=1,\ldots,m$,

$$c_i^k = \begin{cases} \frac{z_i^k}{\mu_i^k}, & \text{if } \mu_i^k > 0, \\ 0, & \text{if } \mu_i^k = 0. \end{cases}$$

**A.2. Derivation of the master problem for inexact GBD.** As in the exact case, the MILP master problem $\mathcal{P}^{\rm GBD}$ can be derived from a more general master problem closer to

TABLE 7. Application of GBD (inexact solution of NLP subproblems and exact cuts) to Example 1. The table reports the number $N$ of iterations taken.

| | Tolerances | | | | | |
|---|---|---|---|---|---|---|
| initial point | 1e-6 | 1e-5 | 1e-4 | 1e-3 | 1e-2 | 1e-1 |
| $(0,0,0)^\top$ | 3 | 3 | 3 | 3 | 3 | 3 |
| $(1,0,0)^\top$ | 3 | 3 | 3 | 3 | 3 | 3 |
| $(0,1,0)^\top$ | 3 | 3 | 3 | 3 | 3 | 3 |
| $(1,0,1)^\top$ | 3 | 3 | 3 | 3 | 3 | 3 |
| $(0,1,1)^\top$ | 4 | 4 | 4 | 4 | 4 | 4 |
| $(0,0,1)^\top$ | 3 | 3 | 3 | 3 | 3 | 3 |

TABLE 8. Application of GBD (inexact solution of NLP subproblems and inexact cuts) to Example 1. The table reports the number $N$ of iterations taken as well as the number $C$ of inequalities violated in Assumptions 2.1 and 2.2.

| | Tolerances | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1e-6 | | 1e-5 | | 1e-4 | | 1e-3 | | 1e-2 | | 1e-1 | |
| initial point | $N$ | $C$ | $N$ | $C$ | $N$ | $C$ | $N$ | $C$ | $N$ | $C$ | $N$ | $C$ |
| $(0,0,0)^\top$ | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0 |
| $(1,0,0)^\top$ | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0 |
| $(0,1,0)^\top$ | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0 |
| $(1,0,1)^\top$ | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0 |
| $(0,1,1)^\top$ | 4 | 0 | 4 | 0 | 4 | 0 | 4 | 0 | 4 | 0 | 4 | 1 |
| $(0,0,1)^\top$ | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0 |

The maximum number for $C$ is as in Table 2.

the original duality motivation of GBD:

$$\mathcal{P}^{\text{GBD}\,1}\left\{\begin{array}{ll} \min & \alpha \\ \text{s.t.} & \inf_{y\in Y}\{f(x,y)+(\lambda^j)^\top g(x,y)-(r^j)^\top(y-y^j)\}-\sum_{i=1}^m s_i^j \;\leq\; \alpha,\ \forall j\in T, \\ & \inf_{y\in Y}\{(\mu^k)^\top g(x,y)-(v^k)^\top(y-y^k)\}+w^k u^k-\sum_{i=1}^m z_i^k \;\leq\; 0,\ \forall k\in S, \\ & x\in X\cap\mathbb{Z}^{n_d}, \alpha\in\mathbb{R}. \end{array}\right.$$

In fact, we will show next that the constraints in problem $\mathcal{P}^{\text{GBD}\,1}$ imply those of $\mathcal{P}^{\text{GBD}}$.

When $l\in T$, one knows that NLP$(x^l)$ has an approximated optimal solution $y^l$, satisfying the corresponding inexact KKT conditions with inexact Lagrange multipliers $\lambda^l$. By the

TABLE 9. Application of GBD (inexact solution of NLP subproblems and exact cuts) to Example 2. The table reports the number $N$ of iterations taken.

| | Tolerances | | | | | |
|---|---|---|---|---|---|---|
| initial point | 1e-6 | 1e-5 | 1e-4 | 1e-3 | 1e-2 | 1e-1 |
| $(1,0,0,0,0)^\top$ | 8 | 8 | 8 | 8 | 8 | 8 |
| $(0,1,0,0,0)^\top$ | 7 | 7 | 7 | 7 | 9 | 9 |
| $(1,0,1,0,0)^\top$ | 5 | 5 | 5 | 5 | 5 | 5 |
| $(1,0,0,1,0)^\top$ | 6 | 6 | 6 | 6 | 6 | 6 |
| $(1,0,0,0,1)^\top$ | 7 | 7 | 7 | 6 | 6 | 6 |
| $(0,1,1,0,0)^\top$ | 6 | 6 | 6 | 6 | 6 | 6 |
| $(0,1,0,1,0)^\top$ | 5 | 5 | 5 | 5 | 5 | 5 |
| $(0,1,0,0,1)^\top$ | 7 | 7 | 7 | 7 | 6 | 6 |
| $(1,0,1,1,0)^\top$ | 8 | 8 | 8 | 8 | 9 | 9 |
| $(1,0,1,0,1)^\top$ | 5 | 5 | 5 | 5 | 5 | 5 |
| $(0,1,1,1,0)^\top$ | 8 | 8 | 8 | 8 | 8 | 8 |
| $(0,1,1,0,1)^\top$ | 6 | 6 | 6 | 6 | 6 | 6 |

convexity of $f$ and $g$ (see (10) and (11)),

$$
\begin{aligned}
& f(x,y) + (\lambda^l)^\top g(x,y) - (r^l)^\top (y-y^l) \\
\geq\ & f(x^l,y^l) + \nabla_x f(x^l,y^l)^\top (x-x^l) + \nabla_y f(x^l,y^l)^\top (y-y^l) \\
& + \sum_{i=1}^m \lambda_i^l [g_i(x^l,y^l) + \nabla_x g_i(x^l,y^l)^\top (x-x^l) + \nabla_y g_i(x^l,y^l)^\top (y-y^l)] \\
& - (r^l)^\top (y-y^l).
\end{aligned}
$$

Thus, using the inexact KKT conditions (1),

$$
\begin{aligned}
\alpha \ \geq\ & \inf_{y\in Y} \{ f(x,y) + (\lambda^l)^\top g(x,y) - (r^l)^\top (y-y^l) \} - \sum_{i=1}^m s_i^l \\
\geq\ & \inf_{y\in Y} \{ f(x^l,y^l) + \nabla_x f(x^l,y^l)^\top (x-x^l) + \sum_{i=1}^m \lambda_i^l \nabla_x g_i(x^l,y^l)^\top (x-x^l) + \sum_{i=1}^m \lambda_i^l g_i(x^l,y^l) \} \\
& - \sum_{i=1}^m s_i^l
\end{aligned}
$$

TABLE 10. Application of GBD (inexact solution of NLP subproblems and inexact cuts) to Example 2. The table reports the number $N$ of iterations taken as well as the number $C$ of inequalities violated in Assumptions 2.1 and 2.2.

| | Tolerances | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1e-6 | | 1e-5 | | 1e-4 | | 1e-3 | | 1e-2 | | 1e-1 | |
| initial point | $N$ | $C$ | $N$ | $C$ | $N$ | $C$ | $N$ | $C$ | $N$ | $C$ | $N$ | $C$ |
| $(1,0,0,0,0)^\top$ | 8 | 1 | 8 | 1 | 8 | 1 | 8 | 1 | 8 | 1 | 8 | 1 |
| $(0,1,0,0,0)^\top$ | 7 | 2 | 7 | 2 | 7 | 2 | 7 | 2 | 9 | 3 | 9 | 3 |
| $(1,0,1,0,0)^\top$ | 5 | 3 | 5 | 3 | 5 | 3 | 5 | 3 | 5 | 3 | 5 | 3 |
| $(1,0,0,1,0)^\top$ | 6 | 0 | 6 | 0 | 6 | 0 | 6 | 0 | 6 | 0 | 6 | 0 |
| $(1,0,0,0,1)^\top$ | 7 | 2 | 7 | 2 | 7 | 2 | 6 | 2 | 6 | 2 | 6 | 2 |
| $(0,1,1,0,0)^\top$ | 6 | 4 | 6 | 4 | 6 | 4 | 6 | 4 | 6 | 4 | 6 | 4 |
| $(0,1,0,1,0)^\top$ | 5 | 0 | 5 | 0 | 5 | 0 | 5 | 0 | 5 | 0 | 5 | 0 |
| $(0,1,0,0,1)^\top$ | 7 | 2 | 7 | 2 | 7 | 2 | 7 | 2 | 6 | 1 | 6 | 1 |
| $(1,0,1,1,0)^\top$ | 8 | 3 | 8 | 3 | 8 | 3 | 8 | 3 | 9 | 3 | 9 | 3 |
| $(1,0,1,0,1)^\top$ | 5 | 0 | 5 | 0 | 5 | 0 | 5 | 0 | 5 | 0 | 5 | 0 |
| $(0,1,1,1,0)^\top$ | 8 | 1 | 8 | 1 | 8 | 1 | 8 | 1 | 8 | 1 | 8 | 1 |
| $(0,1,1,0,1)^\top$ | 6 | 0 | 6 | 0 | 6 | 0 | 6 | 0 | 6 | 0 | 6 | 0 |

The maximum number for $C$ is as in Table 4.

$$
\begin{aligned}
&= f(x^l,y^l)+\nabla_x f(x^l,y^l)^\top(x-x^l)+\sum_{i=1}^m \lambda_i^l\nabla_x g_i(x^l,y^l)^\top(x-x^l)+\sum_{i=1}^m \lambda_i^l g_i(x^l,y^l)\\
&\quad -\sum_{i=1}^m s_i^l\\
&= f(x^l,y^l)+\nabla_x f(x^l,y^l)^\top(x-x^l)+\sum_{i=1}^m \lambda_i^l\nabla_x g_i(x^l,y^l)^\top(x-x^l).
\end{aligned}
$$

The last equality holds due to (2).

When $l\in S$, we know that NLPF$(x^l)$ has an approximated optimal solution $y^l$ satisfying the corresponding inexact KKT conditions with inexact Lagrange multipliers $\mu^l$. Also by the convexity of $g$ (see (11)), we have that

$$
\begin{aligned}
&(\mu^l)^\top g(x,y)-(v^l)^\top(y-y^l)\\
&\geq (\mu^l)^\top[g(x^l,y^l)+\nabla_x g(x^l,y^l)^\top(x-x^l)]+(\sum_{i=1}^m \mu_i^l\nabla_y g_i(x^l,y^l)-v^l)^\top(y-y^l).
\end{aligned}
$$

TABLE 11. Application of GBD (inexact solution of NLP subproblems and exact cuts) to Example 3. The table reports the number $N$ of iterations taken.

| | Tolerances | | | | | |
|---|---|---|---|---|---|---|
| initial point | 1e-6 | 1e-5 | 1e-4 | 1e-3 | 1e-2 | 1e-1 |
| $(1,0,0,0,0,0,0,0)^\top$ | 10 | 10 | 10 | 10 | 10 | 10 |
| $(1,0,0,0,0,0,0,1)^\top$ | 12 | 12 | 12 | 11 | 11 | 11 |
| $(1,0,1,0,0,0,0,1)^\top$ | 9 | 9 | 9 | 9 | 9 | 9 |
| $(1,0,0,0,1,0,0,0)^\top$ | 10 | 10 | 10 | 10 | 10 | 10 |
| $(1,0,0,0,1,0,0,1)^\top$ | 12 | 12 | 12 | 11 | 11 | 11 |
| $(1,0,1,0,1,0,0,1)^\top$ | 9 | 9 | 9 | 9 | 9 | 9 |
| $(1,0,0,1,0,1,0,0)^\top$ | 9 | 9 | 9 | 9 | 9 | 9 |
| $(1,0,0,1,0,1,0,1)^\top$ | 9 | 9 | 9 | 9 | 9 | 9 |
| $(1,0,0,1,0,0,1,0)^\top$ | 10 | 10 | 10 | 9 | 9 | 9 |
| $(1,0,0,1,0,0,1,1)^\top$ | 9 | 9 | 9 | 9 | 9 | 9 |
| $(0,1,0,0,0,0,0,0)^\top$ | 9 | 9 | 9 | 8 | 8 | 8 |
| $(0,1,0,0,0,0,0,1)^\top$ | 9 | 9 | 10 | 10 | 10 | 9 |
| $(0,1,1,0,0,0,0,1)^\top$ | 11 | 11 | 11 | 11 | 11 | 11 |
| $(0,1,0,0,1,0,0,0)^\top$ | 9 | 9 | 9 | 8 | 8 | 8 |
| $(0,1,0,0,1,0,0,1)^\top$ | 9 | 9 | 10 | 10 | 10 | 9 |
| $(0,1,1,0,1,0,0,1)^\top$ | 11 | 11 | 11 | 11 | 11 | 11 |
| $(0,1,0,1,0,1,0,0)^\top$ | 9 | 9 | 9 | 8 | 8 | 8 |
| $(0,1,0,1,0,1,0,1)^\top$ | 10 | 10 | 10 | 9 | 9 | 9 |
| $(0,1,1,1,0,1,0,1)^\top$ | 9 | 9 | 9 | 9 | 9 | 9 |
| $(0,1,0,1,0,0,1,0)^\top$ | 10 | 10 | 10 | 9 | 9 | 9 |
| $(0,1,0,1,0,0,1,1)^\top$ | 9 | 9 | 9 | 8 | 8 | 8 |
| $(1,0,1,1,0,1,0,1)^\top$ | 9 | 9 | 9 | 9 | 9 | 9 |

Then, using the inexact KKT conditions (4),

$$\begin{aligned} 0 &\geq \inf_{y\in Y}\{(\mu^l)^\top g(x,y) - (v^l)^\top(y-y^l)\} + w^l u^l - \sum_{i=1}^m z_i^l \\ &\geq \inf_{y\in Y}\{\sum_{i=1}^m \mu_i^l[g_i(x^l,y^l) + \nabla_x g_i(x^l,y^l)^\top(x-x^l)]\} + w^l u^l - \sum_{i=1}^m z_i^l \\ &= \sum_{i=1}^m \mu_i^l[g_i(x^l,y^l) + \nabla_x g_i(x^l,y^l)^\top(x-x^l)] + w^l u^l - \sum_{i=1}^m z_i^l. \end{aligned}$$

TABLE 12. Application of GBD (inexact solution of NLP subproblems and inexact cuts) to Example 3. The table reports the number $N$ of iterations taken as well as the number $C$ of inequalities violated in Assumptions 2.1 and 2.2.

| | Tolerances | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1e-6 | | 1e-5 | | 1e-4 | | 1e-3 | | 1e-2 | | 1e-1 | |
| initial point | $N$ | $C$ | $N$ | $C$ | $N$ | $C$ | $N$ | $C$ | $N$ | $C$ | $N$ | $C$ |
| $(1,0,0,0,0,0,0,0)^\top$ | 10 | 6 | 10 | 6 | 10 | 6 | 10 | 8 | 10 | 13 | 10 | 12 |
| $(1,0,0,0,0,0,0,1)^\top$ | 12 | 20 | 12 | 20 | 12 | 19 | 11 | 15 | 11 | 17 | 11 | 14 |
| $(1,0,1,0,0,0,0,1)^\top$ | 9 | 10 | 9 | 9 | 9 | 7 | 9 | 10 | 9 | 13 | 9 | 11 |
| $(1,0,0,0,1,0,0,0)^\top$ | 10 | 6 | 10 | 6 | 10 | 6 | 10 | 8 | 10 | 13 | 10 | 12 |
| $(1,0,0,0,1,0,0,1)^\top$ | 12 | 20 | 12 | 20 | 12 | 19 | 11 | 15 | 11 | 17 | 11 | 14 |
| $(1,0,1,0,1,0,0,1)^\top$ | 9 | 10 | 9 | 9 | 9 | 7 | 9 | 10 | 9 | 13 | 9 | 11 |
| $(1,0,0,1,0,1,0,0)^\top$ | 9 | 7 | 9 | 7 | 9 | 6 | 9 | 8 | 9 | 13 | 9 | 13 |
| $(1,0,0,1,0,1,0,1)^\top$ | 9 | 4 | 9 | 4 | 9 | 4 | 9 | 5 | 9 | 10 | 9 | 8 |
| $(1,0,0,1,0,0,1,0)^\top$ | 10 | 14 | 10 | 14 | 10 | 12 | 9 | 10 | 9 | 12 | 9 | 8 |
| $(1,0,0,1,0,0,1,1)^\top$ | 9 | 7 | 9 | 7 | 9 | 6 | 9 | 8 | 9 | 13 | 9 | 13 |
| $(0,1,0,0,0,0,0,0)^\top$ | 9 | 13 | 9 | 13 | 9 | 12 | 8 | 7 | 8 | 8 | 8 | 8 |
| $(0,1,0,0,0,0,0,1)^\top$ | 9 | 7 | 9 | 7 | 10 | 10 | 10 | 11 | 10 | 20 | 9 | 14 |
| $(0,1,1,0,0,0,0,1)^\top$ | 11 | 7 | 11 | 7 | 11 | 6 | 11 | 8 | 11 | 13 | 11 | 13 |
| $(0,1,0,0,1,0,0,0)^\top$ | 9 | 13 | 9 | 13 | 9 | 12 | 8 | 7 | 8 | 8 | 8 | 8 |
| $(0,1,0,0,1,0,0,1)^\top$ | 9 | 7 | 9 | 7 | 10 | 10 | 10 | 11 | 10 | 20 | 9 | 14 |
| $(0,1,1,0,1,0,0,1)^\top$ | 11 | 7 | 11 | 7 | 11 | 6 | 11 | 8 | 11 | 13 | 11 | 13 |
| $(0,1,0,1,0,1,0,0)^\top$ | 9 | 19 | 9 | 19 | 9 | 18 | 8 | 14 | 8 | 16 | 8 | 13 |
| $(0,1,0,1,0,1,0,1)^\top$ | 10 | 15 | 10 | 15 | 10 | 13 | 9 | 11 | 9 | 13 | 9 | 9 |
| $(0,1,1,1,0,1,0,1)^\top$ | 9 | 6 | 9 | 6 | 9 | 6 | 9 | 8 | 9 | 13 | 9 | 12 |
| $(0,1,0,1,0,0,1,0)^\top$ | 10 | 15 | 10 | 15 | 10 | 16 | 9 | 11 | 9 | 11 | 9 | 9 |
| $(0,1,0,1,0,0,1,1)^\top$ | 9 | 14 | 9 | 14 | 9 | 13 | 8 | 8 | 8 | 9 | 8 | 9 |
| $(1,0,1,1,0,1,0,1)^\top$ | 9 | 6 | 9 | 6 | 9 | 5 | 9 | 7 | 9 | 10 | 9 | 9 |

The maximum number for $C$ is as in Table 6.

In summary we have the following property.

**Property A.1.** *Given some sets $T$ and $S$, the lower bound predicted by the master problem $\mathcal{P}^{\mathrm{GBD\,1}}$ is greater than or equal to the one predicted by the master problem $\mathcal{P}^{\mathrm{GBD}}$.*

# References


[1] A. Ahlatçıoğlu and M. Guignard. Convex hull relaxation (CHR) for convex and nonconvex MINLP problems with linear constraints. 2011.

[2] P. Belotti, J. Lee, L. Liberti, F. Margot, and A. Wächter. Branching and bounds tightening techniques for non-convex MINLP. *Optim. Methods Softw.*, 24:597–634, 2009.

[3] J. F. Benders. Partitioning procedures for solving mixed-variables programming problems. *Numerische Mathematik*, 4:238–252, 1962.

[4] P. Bonami, L. T. Biegler, A. R. Conn, G. Cornuéjols, I. E. Grossmann, C. D. Laird, J. Lee, A. Lodi, F. Margot, N. Sawaya, and A. Wächter. An algorithmic framework for convex mixed integer nonlinear programs. *Discrete Optim.*, 5:186–204, 2008.

[5] P. Bonami, G. Cornuéjols, A. Lodi, and F. Margot. A feasibility pump for mixed integer nonlinear programs. *Math. Program.*, 119:331–352, 2009.

[6] P. Bonami and J. P. M. Gonçalves. Heuristics for convex mixed integer nonlinear programs. *Comput. Optim. Appl.*, pages 1–19, 2010.

[7] I. Castillo, J. Westerlund, S. Emet, and T. Westerlund. Optimization of block layout design problems with unequal areas: A comparison of MILP and MINLP optimization methods. *Computers and Chemical Engineering*, 30:54–69, 2005.

[8] M. A. Duran and I. E. Grossmann. An out-approximation algorithm for a class of mixed-integer nonlinear programs. *Math. Program.*, 36:307–339, 1986.

[9] R. Fletcher and S. Leyffer. Solving mixed integer nonlinear programs by outer approximation. *Math. Program.*, 66:327–349, 1994.

[10] O. E. Flippo and A. H. G. Rinnooy Kan. Decomposition in general mathematical programming. *Math. Program.*, 60:361–382, 1993.

[11] A. M. Geoffrion. Generalized Benders decomposition. *J. Optim. Theory Appl.*, 10:237–260, 1972.

[12] I. E. Grossmann. Review of nonlinear mixed-integer and disjunctive programming techniques. *Optim. Eng.*, 3:227–252, 2002.

[13] J. E. Kelley. The cutting-plane method for solving convex programs. *J. Soc. Indust. Appl. Math.*, 8:703–712, 1960.

[14] G. Nannicini, P. Belotti, J. Lee, J. Linderoth, F. Margot, and A. Wächter. A probing algorithm for MINLP with failure prediction by SVM. Technical Report RC25103 (W1101-112), IBM Research, Computer Science, 2011.

[15] I. Quesada and I. E. Grossmann. An LP/NLP based branch and bound algorithm for convex MINLP optimization problems. *Computers and Chemical Engineering*, 16:937–947, 1992.

[16] M. Tawarmalani and N. V. Sahinidis. Global optimization of mixed-integer nonlinear programs: A theoretical and computational study. *Math. Program.*, 99:563–591, 2004.

[17] S. A. Tawfik. `http://www.mathworks.com/matlabcentral/fx\_files/6990/1/IP.m`.

[18] T. Westerlund and F. Pettersson. An extended cutting plane method for solving convex MINLP problems. *Computers and Chemical Engineering*, 19:131–136, 1995.

[19] T. Westerlund and R. Pörn. Solving pseudo-convex mixed integer optimization problems by cutting plane techniques. *Optim. Eng.*, 3:253–280, 2002.

[20] T. Westerlund, H. Skrifvars, I. Harjunkoski, and R. Pörn. An extended cutting plane method for a class of non-convex MINLP problems. *Computers and Chemical Engineering*, 22:357–365, 1998.



M. Li
Department of Mathematics, University of Coimbra, 3001-454 Coimbra, Portugal (limin@mat.uc.pt).

L. N. Vicente
CMUC, Department of Mathematics, University of Coimbra, 3001-454 Coimbra, Portugal (lnv@mat.uc.pt).