Information Processing Letters 118 (2017) 110–116

Contents lists available at ScienceDirect

Information Processing Letters

www.elsevier.com/locate/ipl

Inexact feasibility pump for mixed integer nonlinear programming

M. Li [*,1], Q. Liu

School of Mathematical Sciences, Shandong Normal University, Ji'nan, China

ARTICLE INFO

Article history:
Received 16 December 2015
Received in revised form 15 September 2016
Accepted 27 October 2016
Available online 3 November 2016
Communicated by J.L. Fiadeiro

Keywords:
Mixed integer nonlinear programming
Feasibility pump
Algorithms
Inexactness
Heuristic

ABSTRACT

The mixed integer nonlinear programming (MINLP) problem as an optimization problem involves both continuous and discrete variables. Moreover, at least one of the functions defining the objective function or the constraints must be nonlinear. Because of its complexity, it is very difficult to obtain the exact optimal solution. Therefore, the heuristic methods for getting a feasible solution of MINLPs are very important in practice. The feasibility pump is one of the famous heuristic methods, which alternates between solving nonlinear programming (NLP) problems and mixed integer linear programming (MILP) relaxed master problems.

In this paper, we will extend the feasibility pump to the case where the NLP problems are solved inexactly and propose the convergence of this method under some conditions. Moreover, we present the study of inexactness of the Lagrange multipliers (which are returned negative) of the NLP subproblems.

1. Introduction

The mixed integer nonlinear programming (MINLP) problem is an optimization problem with both continuous and discrete variables. Moreover, at least one of the functions defining the objective function or the constraints must be nonlinear. We denote the MINLP problems in the following form:

$$\text{(MINLP)} \begin{cases} \min & F(x, y) \\ \text{s.t.} & G(x, y) \le 0, \\ & x \in X \cap \mathbb{Z}^{n_1},\ y \in Y, \end{cases}$$

where X is assumed to be a bounded polyhedral subset of $\mathbb{R}^{n_1}$ and Y a polyhedral subset of $\mathbb{R}^{n_2}$. The functions $F : X \times Y \longrightarrow \mathbb{R}$ and $G : X \times Y \longrightarrow \mathbb{R}^m$ are supposed continuously differentiable. Moreover, the constraint functions $G_j,\ j = 1, \ldots, m$ are assumed to be convex.

When the discrete variables are relaxed, it is easy to see that a MINLP problem is indeed a nonlinear programming (NLP). Furthermore, if all the constraints and the objective function are linear, the MINLPs are mixed integer linear programming (MILP) problems. Therefore, the research on the solution of the MINLPs is more complex, compared to that of NLPs or MILPs. In fact, most of the algorithms about MINLPs are based on the iterations of NLPs and MILPs. Since the MINLP problem is NP-hard, it may take a lot of effort before its first feasible solution appears. Therefore, the heuristic methods which do not provide a guarantee of optimality at last, and thus the incumbent or best point found so far is only feasible, are very important in practice.

There are several kinds of heuristics which have been successfully used for MILP problems: diving heuristics

* Corresponding author.
E-mail addresses: liminEmily@sdnu.edu.cn (M. Li), lq_qsh@163.com (Q. Liu).

[1] Support for this author was provided by the Project Sponsored by the Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry (Grant No. 20151098), and Shandong Provincial Natural Science Foundation, China (Grant No. ZR2014J L001).

http://dx.doi.org/10.1016/j.ipl.2016.10.009
0020-0190/

(see [3]), the relaxation induced neighborhood search (see [10]), the feasibility pump technique (see [11,13]), and so on. Later, Bonami and Gonçalves [8] extended some of these heuristics to MINLP problems. Here, we will mainly study one of theses heuristic methods, the feasibility pump, denoted by FP for short. FP was researched by Fischetti et al. [11] to obtain a good feasible candidate point $(\hat{x}, \hat{y})$ for MILPs, where the point $(\hat{x}, \hat{y}) \in (X \cap \mathbb{Z}^{n_1}) \times Y$ was such that $G(\hat{x}, \hat{y}) \leq 0$ possibly did not have a very high objective value $F(\hat{x}, \hat{y})$. However, the functions F and G were assumed to be affine. This method was successfully applied to MILPs with binary discrete variables, but not so well to ones with general discrete variables. In 2007, FP was improved both by Achterberg and Berthold [1] and Bertacco et al. [5], for obtaining a better MILP feasible candidate point with lower objective value and for being used to solve more general scenario, respectively. Moreover, Achterberg and Berthold [1] first introduced the *objective feasibility pump* (denoted by OFP), which was recently extended to convex MINLPs as a multi-objective problem by Sharma et al. [15]. Bonami et al. [7] applied FP to be adapted to MINLP problems in 2009, which alternated between solving NLP problems and MILP relaxed master problems. Moreover, they solved the NLPs exactly and constructed the cuts in the MILPs by the outer approximation method. Recently, the FP method still is researched actively for both convex and non-convex MINLPs in [2,4,6,9,12,14]. Meanwhile, Li and Vicente [16] have generalized the decomposition methods for solving MINLPs, such as the outer approximation method and the generalized Benders decomposition method, to the inexact case when the respective NLP subproblems were solved inexactly. In this paper, we will investigate the effect of FP as a heuristic method for solving MINLPs when the NLPs are solved inexactly. This work is mainly based on [7] and [16], and we will extend the result of basic feasibility pump in [7] by using outer approximation method when the constraint functions are convex.

First, the continuous relaxation of (MINLP) is obtained by moving the integer control in the following way:

$$\text{(R-MINLP)} \begin{cases} \min & F(x, y) \\ \text{s.t.} & G(x, y) \leq 0, \\ & x \in X, \ y \in Y. \end{cases}$$

The feasible region of (R-MINLP) is denoted by S:

$$S = \{(x, y) \in X \times Y : G(x, y) \leq 0\}$$

Similarly, we define the integer region by T:

$$T = \{(x, y) \in X \times Y : x \in \mathbb{Z}^{n_1}\}$$

Next, for $k \geq 0$, the two sequences $\{(\bar{x}^k, \bar{y}^k)\}$ and $\{(\tilde{x}^k, \tilde{y}^k)\}$ will be constructed with the following properties.

The sequence $(\bar{x}^k, \bar{y}^k)$ is generated by solving MILP problems, where the points $(\bar{x}^k, \bar{y}^k) \in T$ do not necessarily satisfy $G(\bar{x}^k, \bar{y}^k) \leq 0$.

The sequence $(\tilde{x}^k, \tilde{y}^k)$ is generated by inexactly solving NLP problems, where the points $(\tilde{x}^k, \tilde{y}^k)$ belong to S but $\tilde{x}^k \notin X \cap \mathbb{Z}^{n_1}$, i.e., $(\tilde{x}^k, \tilde{y}^k) \notin T$.

The procedure between obtaining the above two sequences by solving MILP master problem and inexactly solving NLP subproblems is called *inexact Feasibility Pump*, denoted by *inexact* FP. As a result of being an extension of the feasibility pump and outer approximation method, *inexact* FP is developed for convex MINLPs.

We organize the paper in the following way. In Sect. 2, we extend FP to relax the exact solution of the NLP subproblems, by using inexact residual information to redefine the cuts in the MILP relaxed master problems. In Sect. 3, we will report some computational results to show the effectiveness of *inexact* FP. In Sect. 4, we will consider the case that the approximate Lagrange multipliers returned may be negative, by a small residual amount, when the NLP subproblems are solved inexactly. Sect. 5 concludes the paper.

2. Inexact feasibility pump when the Lagrange multipliers are nonnegative

Given any $\bar{x}^i$ $(i = 0, 1, \ldots)$ belongs to $X \cap \mathbb{Z}^{n_1}$, one wants to find a point (x, y) in the set S satisfying that x is closest to $\bar{x}^i$. It can be realized by solving the following NLP problems:

$$\text{(IFP-NLP)}^i \begin{cases} \min & \|x - \bar{x}^i\|_2 \\ \text{s.t.} & G(x, y) \leq 0, \\ & x \in X, \ y \in Y. \end{cases}$$

If the problem (MINLP) has a finite optimal value, it is easy to see that these NLP problems are feasible, i.e. for each (IFP-NLP)^i, there exists $(\hat{x}^i, \hat{y}^i) \in X \times Y$ such that $G(\hat{x}^i, \hat{y}^i) \leq 0$. The case that the NLP problems are exactly solved has been discussed and used in [7].

Now, by inexactly solving (IFP-NLP)^k, where $k \in \{0, 1, \ldots, i\}$, we can obtain its approximate optimal solution $(\tilde{x}^k, \tilde{y}^k)$ with the inexact Lagrange multipliers μ^k. If the constraint qualification for (IFP-NLP)^k holds at $(\tilde{x}^k, \tilde{y}^k)$, the first-order necessary conditions may not be satisfied because of the inexact solutions. Therefore, we define the inexact KKT conditions at $(\tilde{x}^k, \tilde{y}^k, \mu^k)$ for NLP problem (IFP-NLP)^k as follows:

$$\frac{\tilde{x}^k - \bar{x}^k}{\|\tilde{x}^k - \bar{x}^k\|_2} + \sum_{j=1}^{m} \mu_j^k \nabla_x G_j(\tilde{x}^k, \tilde{y}^k) = u^k, \tag{1}$$

$$\sum_{j=1}^{m} \mu_j^k \nabla_y G_j(\tilde{x}^k, \tilde{y}^k) = v^k, \tag{2}$$

$$\mu_j^k G_j(\tilde{x}^k, \tilde{y}^k) = w_j^k, \quad j = 1, \ldots, m, \tag{3}$$

$$\mu^k \geq 0, \tag{4}$$

where, $u^k \in \mathbb{R}^{n_1}$, $v^k \in \mathbb{R}^{n_2}$, and $w^k \in \mathbb{R}^m$ are called the residuals of the NLP problem (IFP-NLP)^k. The methodology is borrowed from [16], while the forms here are more complex for two reasons. On one hand, the residuals involve both x and y variables, not only y. On the other hand, both variables involved by the residuals appear nonlinearly.

Now for $i \geq 0$, we define the perturbed master problems based on [7] and [16], whose constraints involve the residuals appeared by inexactly solving the NLP problems.

$$
(\text{IFP-OA})^{i+1} \begin{cases} \min & \|x - \bar{x}^i\|_1 \\ \text{s.t.} & \begin{pmatrix} \nabla_x G_j(\tilde{x}^k, \tilde{y}^k) - \frac{1}{\|\mu^k\|_1} u^k \\ \nabla_y G_j(\tilde{x}^k, \tilde{y}^k) - \frac{1}{\|\mu^k\|_1} v^k \end{pmatrix}^\top \\ & \quad \times \begin{pmatrix} x - \tilde{x}^k \\ y - \tilde{y}^k \end{pmatrix} + G_j(\tilde{x}^k, \tilde{y}^k) \leq b_j^k, \\ & j = 1, \dots, m \text{ and } k = 0, 1, \dots, i, \\ & x \in X \cap \mathbb{Z}^{n_1}, \ y \in Y, \end{cases}
$$

where, for $j = 1, \dots, m$ and $k = 0, \dots, i$,

$$
b_j^k = \begin{cases} \frac{w_j^k}{\mu_j^k}, & \text{if } \mu_j^k > 0, \\ 0, & \text{if } \mu_j^k = 0, \end{cases} \tag{5}
$$

and its optimal solution will be denoted by $(\bar{x}^{i+1}, \bar{y}^{i+1})$. One can see that if all the residuals from (1) to (4) are zeros, the cuts in current master problem are the same as the ones in [7], which also are famous as outer approximation of the constraints. Furthermore, even if we borrowed the methodology from [16], as we explained above, the current cuts are more complex, since the residuals u^k perturbed of the gradient with respect to x are nonlinear. Moreover, when the number of iterations increases, the number of the cuts in the master MILPs also grows larger by a factor of m (the number of the constraints in the original problem (MINLP)).

The *inexact* FP iterates between inexactly solving $(\text{IFP-NLP})^i$ and $(\text{IFP-OA})^{i+1}$ (where $i \geq 0$) until either an inexact feasible solution of MINLP is found or $(\text{IFP-OA})^{i+1}$ becomes infeasible.

Even if we permit the existence of the residuals during inexactly solving the NLPs, their size cannot be arbitrarily large. Now, we assume that the residuals satisfy the following condition.

Assumption 2.1. For any different iteration steps k and i (without loss of generality, we assume that $k < i$), the inexact solutions and corresponding Lagrange multipliers of $(\text{IFP-NLP})^k$ and of $(\text{IFP-NLP})^i$ are denoted by $(\tilde{x}^k, \tilde{y}^k, \mu^k)$ and $(\tilde{x}^i, \tilde{y}^i, \mu^i)$, respectively, which satisfies the inexact KKT conditions. We assume that the following inequality is satisfied for some $\sigma \in (0, 1)$:

$$
\begin{aligned} & \left\|u^k\right\|_2 + \left\|v^k\right\|_2 + \left\|u^i\right\|_2 + \left\|v^i\right\|_2 + \frac{1}{\mathrm{M}} \left\|w^i\right\|_1 \\ & \leq \frac{\sigma}{\mathrm{M}} \cdot \frac{\min\left\{\|\mu^k\|_1, \|\mu^i\|_1\right\}}{\max\left\{\|\mu^k\|_1, \|\mu^i\|_1\right\}} \cdot \left\|\bar{x}^i - \tilde{x}^i\right\|_2, \end{aligned} \tag{6}
$$

where $\mathrm{M} = \|\bar{x}^i - \tilde{x}^i\|_2 + \|\bar{y}^i - \tilde{y}^i\|_2$.

Algorithm 2.1 *(Inexact feasibility pump).*

Initialization

Let x^0 be given. Set $p = 0$ and $\mathrm{F} = +\infty$.

REPEAT

1. Inexactly solve the NLP problem $(\text{IFP-NLP})^p$. Let $(\tilde{x}^p, \tilde{y}^p)$ be an approximate optimal solution and $\tilde{F}^p$ be the corresponding inexact optimal value. At the same time, obtain the corresponding inexact Lagrange multipliers μ^p, and evaluate the residuals u^p, v^p, and w^p.
2. Linearize the constraints at $(\tilde{x}^p, \tilde{y}^p)$.
3. Solve the relaxed master problem $(\text{IFP-OA})^{p+1}$, obtaining a new discrete assignment x^{p+1} to be tested in the algorithm and the optimal value Q^{p+1}. Set $F = \min\{F, Q^{p+1}\}$. Then Increment p by one unit.

UNTIL ($(\text{IFP-OA})^p$ is infeasible or $F = 0$ or $\tilde{F}^p = 0$).

Given any vector $v \in \mathbb{R}^n$, for the convenience of the following explanation, we define index set N_v to denote its negative components,

$$
N_v = \{i : v_i < 0, i = 1, \dots, n\}, \tag{7}
$$

and P_v its complement,

$$
P_v = \{1, \dots, n\} \setminus N_v. \tag{8}
$$

Therefore, one can find that $N_{\mu^k} = \emptyset$ because of the condition (4).

Theorem 2.1. *In the inexact* FP, *let* $(\tilde{x}^i, \tilde{y}^i)$ *be an approximate optimal solution of* $(\text{IFP-NLP})^i$ *and* $(\bar{x}^i, \bar{y}^i)$ *(where* $(\bar{x}^0, \bar{y}^0)$ *is assumed to be a random point in the set T) be an optimal solution of* $(\text{IFP-OA})^i$*. If the inexact* KKT *conditions are satisfied at every* $(\tilde{x}^i, \tilde{y}^i)$ *and Assumption 2.1 is satisfied, then* $\bar{x}^i \neq \tilde{x}^k$ *for all* $k = 0, \dots, i-1$.

Proof. If not, suppose that there is some $k \leq i-1$ such that $\bar{x}^i = \tilde{x}^k$, i.e., $(\tilde{x}^k, \tilde{y}^k)$ is also an approximate optimal solution of $(\text{IFP-NLP})^i$. Then it satisfies the inexact KKT conditions with the Lagrange multipliers μ^i as follows:

$$
\frac{\tilde{x}^k - \bar{x}^i}{\|\tilde{x}^k - \bar{x}^i\|_2} + \sum_{j=1}^m \mu_j^i \nabla_x G_j(\tilde{x}^k, \tilde{y}^k) = u^i, \tag{9}
$$

$$
\sum_{j=1}^m \mu_j^i \nabla_y G_j(\tilde{x}^k, \tilde{y}^k) = v^i, \tag{10}
$$

$$
\mu_j^i G_j(\tilde{x}^k, \tilde{y}^k) = w_j^i, \quad j = 1, \dots, m, \tag{11}
$$

$$
\mu^i \geq 0. \tag{12}
$$

Secondly, $(\tilde{x}^k, \tilde{y}^k)$ as an approximate minimizer of problem $(\text{IFP-NLP})^k$ satisfies the inexact KKT conditions (1)–(4). So, by conditions (3) and (11), we can obtain the following equalities:

$$
\mu_j^i w_j^k = \mu_j^k w_j^i, \quad j = 1, \dots, m. \tag{13}
$$

Thirdly, by the stopping criteria that a feasible solution of MINLP is found, we have that $\tilde{x}^i \neq \bar{x}^i$.

Moreover, since $(\bar{x}^i, \bar{y}^i)$ is an optimal solution of $(\text{IFP-OA})^i$, it is obviously feasible. Then, for the above

$k \le i-1$, $(\bar{x}^i, \bar{y}^i)$ satisfies the inexact outer approximate of the constraints at $(\tilde{x}^k, \tilde{y}^k)$, that is, for $j = 1, \dots, m$,

$$\begin{pmatrix} \nabla_x G_j(\tilde{x}^k, \tilde{y}^k) - \frac{1}{\|\mu^k\|_1} u^k \\ \nabla_y G_j(\tilde{x}^k, \tilde{y}^k) - \frac{1}{\|\mu^k\|_1} v^k \end{pmatrix}^\top \begin{pmatrix} \bar{x}^i - \tilde{x}^k \\ \bar{y}^i - \tilde{y}^k \end{pmatrix} + G_j(\tilde{x}^k, \tilde{y}^k) \le b_j^k, \tag{14}$$

where b_j^k is defined as in (5).

Multiplying the inequalities in (14) by the nonnegative multipliers $\mu_1^i, \dots, \mu_m^i$, and summing them up, one has

$$\begin{pmatrix} \sum_{j=1}^m \mu_j^i \nabla_x G_j(\tilde{x}^k, \tilde{y}^k) - \frac{\|\mu^i\|_1}{\|\mu^k\|_1} u^k \\ \sum_{j=1}^m \mu_j^i \nabla_y G_j(\tilde{x}^k, \tilde{y}^k) - \frac{\|\mu^i\|_1}{\|\mu^k\|_1} v^k \end{pmatrix}^\top \begin{pmatrix} \bar{x}^i - \tilde{x}^k \\ \bar{y}^i - \tilde{y}^k \end{pmatrix} + \sum_{j=1}^m \mu_j^i G_j(\tilde{x}^k, \tilde{y}^k) \le \sum_{j=1}^m \mu_j^i b_j^k. \tag{15}$$

Then by conditions (9)–(12), (15) can be written as:

$$\begin{aligned} &\|\bar{x}^i - \tilde{x}^i\|_2 + \left(u^i - \frac{\|\mu^i\|_1}{\|\mu^k\|_1} u^k \right)^\top (\bar{x}^i - \tilde{x}^k) \\ &\quad + \left(v^i - \frac{\|\mu^i\|_1}{\|\mu^k\|_1} v^k \right)^\top (\bar{y}^i - \tilde{y}^k) + \sum_{j=1}^m w_j^i - \sum_{j=1}^m \mu_j^i b_j^k \\ &\quad \le 0. \end{aligned} \tag{16}$$

Replacing $(\tilde{x}^k, \tilde{y}^k)$ in (16) by $(\tilde{x}^i, \tilde{y}^i)$, we have that:

$$\begin{aligned} &\|\bar{x}^i - \tilde{x}^i\|_2 + \left(u^i - \frac{\|\mu^i\|_1}{\|\mu^k\|_1} u^k \right)^\top (\bar{x}^i - \tilde{x}^i) \\ &\quad + \left(v^i - \frac{\|\mu^i\|_1}{\|\mu^k\|_1} v^k \right)^\top (\bar{y}^i - \tilde{y}^i) + \sum_{j=1}^m w_j^i - \sum_{j=1}^m \mu_j^i b_j^k \\ &\quad \le 0. \end{aligned} \tag{17}$$

Now, we start to discuss a part of inequality in (17) by the definition of b_j^k in (5).

$$\begin{aligned} \sum_{j=1}^m w_j^i - \sum_{j=1}^m \mu_j^i b_j^k &= \sum_{j \in P_{\mu^k}} \left(w_j^i - \frac{\mu_j^i}{\mu_j^k} w_j^k\right) + \sum_{j \in N_{\mu^k}} w_j^i \\ &= \sum_{j \in P_{\mu^k}} \frac{\mu_j^k w_j^i - \mu_j^i w_j^k}{\mu_j^k} + \sum_{j \in N_{\mu^k}} w_j^i \\ &= \sum_{j \in N_{\mu^k}} w_j^i \end{aligned} \tag{18}$$

where the definitions of index sets N_{μ^k} and P_{μ^k} have been introduced in (7) and (8), respectively, and the last equation is hold because of the result in (13). Moreover, by the definition of w_j^i in (11) and because of the feasibility of $(\tilde{x}^k, \tilde{y}^k)$ in (IFP-NLP)^i, it is easy to obtain the following property:

$$\sum_{j \in \bar{J}} w_j^i \le 0. \tag{19}$$

Then, by the results in (18) and (19), the inequality in (17) can imply the following one:

$$\begin{aligned} &-\left[\|\mu^k\|_1 \begin{pmatrix} u^i \\ v^i \end{pmatrix} - \|\mu^i\|_1 \begin{pmatrix} u^k \\ v^k \end{pmatrix} \right]^\top \begin{pmatrix} \bar{x}^i - \tilde{x}^i \\ \bar{y}^i - \tilde{y}^i \end{pmatrix} \\ &\quad - \|\mu^k\|_1 \sum_{j \in \bar{J}} w_j^i \ge \|\mu^k\|_1 \cdot \|\bar{x}^i - \tilde{x}^i\|_2. \end{aligned} \tag{20}$$

While on the other hand, we can get the following discussion in terms of the properties of the norm for vectors.

$$\begin{aligned} &-\left[\|\mu^k\|_1 \begin{pmatrix} u^i \\ v^i \end{pmatrix} - \|\mu^i\|_1 \begin{pmatrix} u^k \\ v^k \end{pmatrix} \right]^\top \begin{pmatrix} \bar{x}^i - \tilde{x}^i \\ \bar{y}^i - \tilde{y}^i \end{pmatrix} \\ &\le \left\| \|\mu^k\|_1 \begin{pmatrix} u^i \\ v^i \end{pmatrix} - \|\mu^i\|_1 \begin{pmatrix} u^k \\ v^k \end{pmatrix} \right\|_2 \cdot \left\| \begin{pmatrix} \bar{x}^i - \tilde{x}^i \\ \bar{y}^i - \tilde{y}^i \end{pmatrix} \right\|_2 \\ &\le \max\left\{ \left\|\mu^k\right\|_1, \left\|\mu^i\right\|_1 \right\} \left(\left\|u^i\right\|_2 + \left\|v^i\right\|_2 + \left\|u^k\right\|_2 \right. \\ &\quad \left. + \left\|v^k\right\|_2 \right) \cdot \left(\left\|\bar{x}^i - \tilde{x}^i\right\|_2 + \left\|\bar{y}^i - \tilde{y}^i\right\|_2 \right) \end{aligned}$$

Therefore, the left hand in (20) can be enlarged in the following way by (6) in the Assumption 2.1:

$$\begin{aligned} &-\left[\|\mu^k\|_1 \begin{pmatrix} u^i \\ v^i \end{pmatrix} - \|\mu^i\|_1 \begin{pmatrix} u^k \\ v^k \end{pmatrix} \right]^\top \begin{pmatrix} \bar{x}^i - \tilde{x}^i \\ \bar{y}^i - \tilde{y}^i \end{pmatrix} \\ &\quad - \|\mu^k\|_1 \sum_{j \in \bar{J}} w_j^i \\ &\le \max\left\{ \left\|\mu^k\right\|_1, \left\|\mu^i\right\|_1 \right\} \cdot \left(\left\|u^i\right\|_2 + \left\|v^i\right\|_2 + \left\|u^k\right\|_2 \right. \\ &\quad \left. + \left\|v^k\right\|_2 \right) \cdot \left(\left\|\bar{x}^i - \tilde{x}^i\right\|_2 + \left\|\bar{y}^i - \tilde{y}^i\right\|_2 \right) \\ &\quad + \max\left\{ \left\|\mu^k\right\|_1, \left\|\mu^i\right\|_1 \right\} \cdot \left\|w^i\right\|_1 \\ &= \max\left\{ \left\|\mu^k\right\|_1, \left\|\mu^i\right\|_1 \right\} \cdot \mathrm{M} \cdot \left(\left\|u^i\right\|_2 + \left\|v^i\right\|_2 \right. \\ &\quad \left. + \left\|u^k\right\|_2 + \left\|v^k\right\|_2 + \frac{1}{\mathrm{M}} \cdot \left\|w^i\right\|_1 \right) \\ &\le \max\left\{ \left\|\mu^k\right\|_1, \left\|\mu^i\right\|_1 \right\} \cdot \mathrm{M} \cdot \frac{\sigma}{\mathrm{M}} \\ &\quad \cdot \frac{\min\left\{ \|\mu^k\|_1, \|\mu^i\|_1 \right\}}{\max\left\{ \|\mu^k\|_1, \|\mu^i\|_1 \right\}} \cdot \left\|\bar{x}^i - \tilde{x}^i\right\|_2 \\ &< \|\mu^k\|_1 \cdot \|\bar{x}^i - \tilde{x}^i\|_2 \end{aligned}$$

which is contradicts to the right hand in (20). □

Note that the result in Theorem 2.1 extends that of [7] when the residuals all are zeros. Moreover, the proof of the former follows closely the lines of the latter. However, we should point out that there exists difference between them. Especially, when the case that two NLP problems have the same optimal solution but with different Lagrange multipliers occurs. In [7], since the NLP problems were solved exactly, neither their solutions nor the MILP problems needed to involve the residuals at all, then there is no influence on this supposed case. However, when we solve the NLPs inexactly, it will become a difficult issue

Table 1
The number of variables and constraints, and the optimal values of all tested problems. The number of constraints include linear equalities and inequalities and nonlinear inequalities.

Problem	n_1	n_2	# of constraints
batch	24	22	69
trimloss2	31	6	24
trimloss4	85	20	64
trimloss5	161	30	90
trimloss6	173	42	120
optprloc	25	5	29
CLay0203H	18	72	132
CLay0203M	18	12	54
CLay0204H	32	132	234
CLay0204M	32	20	90
CLay0205M	50	210	365
CLay0303H	21	78	150
CLay0303M	21	12	66
CLay0304H	36	140	258
Syn10H	10	67	112
Syn10M	10	25	54
Syn40H	40	282	466
Syn40M	40	90	226
Synthehs21	5	6	13
Synthehs312	8	9	23
Syn05M02H	20	84	151
Syn05M02M	20	40	101
Syn05M03H	30	126	249
Syn10M02H	40	154	294
Syn10M02M	40	70	198
FLay03H	12	110	144
FLay03M	12	14	24
FLay04M	24	18	42

because both the corresponding Lagrange multipliers and the constraints of the MILP problems are perturbed by the different residuals.

3. Numerical experiments

In this part, some of the practical features of *inexact* FP algorithm will be displayed by reporting numerical results, where the test problems taken from AMPL are introduced in Table 1. Moreover, the AMPL code for problems batch, trimloss2, trimloss4, trimloss5, trimloss6, and optprloc was taken from the MacMINLP collection [18] and for the others from the Open Source CMU-IBM Project [17].

The implementation and testing of Algorithm 2.1 was run in MATLAB (version 7.9.0, R2009b). The NLP subproblems were solved by the MATLAB function `fmincon`, and the MILP problems were solved by function `cplexslqmilp` from CPLEX [19] (version 12.4 called from MATLAB). As explained in [7] it is not necessary to obtain an optimal solution $(\bar{x}^i, \bar{y}^i)$ of (IFP-OA)^i, so we replace the 1-norm of objective function by 2-norm in our implementation. Furthermore, the linear equality constraints possibly present in the original problems (as well as the bounds in the variables) were kept in the MILP master problems.

We will give the results not only about the algorithms FP (denoted by **FP**) and *inexact* FP (denoted by **IFP**), but also about the case **IFP (exact cuts)**, where the cuts of MILPs are defined exactly but the NLPs are solved inexactly. In fact, the default tolerances for function values are set to 10^{-6} in `fmincon`. To ensure the achievement of feasibility in the solution of the NLPs, we keep the tolerance to 10^{-6} corresponding to the constraint violation (`TolCon` in `fmincon`). But `TolFun` and `TolX` used in `fmincon` are set to 10^{-2} when the NLPs are solved inexactly.

In the table of results, we use N to denote the number of iterations. It should be pointed out that we did not apply the Assumption 2.1 in our implementation to the variant **IFP**. But we gave the number C of inequalities of this assumption that were violated by more than 10^{-8}. The columns labeled "Time" show the CPU time in seconds rounded to the number with two digits. If the CPU time is more than 1 hour, it will be denoted by "—".

The stopping criteria of this algorithm consisted of the corresponding master program being infeasible, or the objective value of the NLP problems being zero, or the number of iterations exceeding 50.

Table 2 summarizes the application of *inexact* FP (Algorithm 2.1) to our test set, compared to the exact FP. For problems trimloss2, trimloss5, CLay0203H, CLay0304H, Syn05M02H, and Syn05M03H, the numbers of iterations for the variant **IFP (exact cuts)** are more than 50. They indicate that the inexactness of NLPs indeed influenced the convergence of the FP, and we can try to avoid this issue by involving the inexactness of cuts in MILPs. Moreover, we also observe that *inexact* FP converged in most of the cases (except the problems trimloss4 and CLay0204M) while neglecting the imposition of the inequalities of Assumption 2.1.

4. Inexact feasibility pump when some Lagrange multipliers are negative

During the above discussion, the condition (4) that the approximate Lagrange multipliers should be non-negative played a more key role. However, when the NLP subproblems are solved inexactly, depending on the solver chosen, the approximate Lagrange multipliers returned may be negative, by a small residual amount (such as 0 possibly is approximated by -10^{-6}). Now, we will generalize the method of *inexact* FP to allow inexactness in the non-negativity of the inexact Lagrange multipliers of the NLP subproblems.

When every NLP subproblem (IFP-NLP)^k is solved inexactly, we also use y^k to denote its approximate optimal solution, which satisfies an inexact form of the corresponding KKT conditions. Here, we assume that there exists a vector of inexact multipliers $\mu^k \in \mathbb{R}^m$ (not necessarily non-negative) and vectors of residuals $u^k \in \mathbb{R}^{n_1}$, $v^k \in \mathbb{R}^{n_2}$ and $w^k \in \mathbb{R}^m$ such that the following equations hold:

$$\frac{\tilde{x}^k - \bar{x}^k}{||\tilde{x}^k - \bar{x}^k||_2} + \sum_{j=1}^{m} \mu_j^k \nabla_x G_j(\tilde{x}^k, \tilde{y}^k) = u^k, \tag{21}$$

$$\sum_{j=1}^{m} \mu_j^k \nabla_y G_j(\tilde{x}^k, \tilde{y}^k) = v^k, \tag{22}$$

$$\mu_j^k G_j(\tilde{x}^k, \tilde{y}^k) = w_j^k, \quad j = 1, \ldots, m. \tag{23}$$

Now, we define the new inexact multipliers in following way:

Table 2
The table reports the number N of iterations taken, the number C of inequalities found to violate Assumption 2.1 and the Time of CPU time taken in seconds.

Problem	FP		IFP (exact cuts)		IFP		
	N	Time	N	Time	N	Time	C
batch	2	0.66	2	0.51	4	1.31	3
trimloss2	5	173.00	>50	129.00	18	270.00	153
trimloss4	10	460.10	>50	791.20	2(Infeasible)	33.95	0
trimloss5	22	1846.00	>50	1743.00	9	150.00	15
trimloss6	>50	–	>50	–	6	291.20	15
optprloc	1	0.39	1	0.08	1	0.08	0
CLay0203H	3	2.89	>50	436.50	2	1.86	1
CLay0203M	2	2.13	2	2.14	9	20.64	28
CLay0204H	2	25.77	2	31.56	2	16.52	1
CLay0204M	2	6.68	2	6.78	2(Infeasible)	14.48	0
CLay0205M	2	11.28	2	9.07	3	12.39	1
CLay0303H	>50	530.60	>50	288.30	2	4.23	1
CLay0303M	2	2.50	3	3.42	19	18.29	153
CLay0304H	2	15.28	>50	2693.00	2	13.05	1
Syn10H	1	5.25	1	4.99	1	5.17	0
Syn10M	1	0.19	1	0.13	1	0.13	0
Syn40H	1	74.90	1	76.88	1	79.75	0
Syn40M	1	0.30	1	0.30	1	0.30	0
synthes21	1	1.00	1	0.04	1	0.03	0
synthes312	1	0.03	1	0.03	1	0.03	0
Syn05M02H	6	159.10	>50	480.70	14	38.19	70
Syn05M02M	2	1.65	2	0.61	1	14.28	1
Syn05M03H	19	1577.00	>50	–	9	54.26	36
Syn10M02H	40	–	3	49.67	12	329.50	47
Syn10M02M	2	477.70	2	14.20	2	23.58	1
FLay03H	1	23.63	2	8.72	2	12.03	1
FLay03M	1	2.48	2	0.40	28	8.04	351
FLay04M	1	3.37	2	0.54	5	1.19	6

$$\bar{\mu}_j^k = \begin{cases} \mu_j^k, & \text{if } j \in P_{\mu^k}, \\ -\mu_j^k, & \text{if } j \in N_{\mu^k}, \end{cases} \tag{24}$$

for $j = 1, \ldots, m$, where the index sets N_{μ^k} and P_{μ^k} are defined as in (7) and (8). And the new residuals are defined as

$$\bar{u}^k = u^k - 2 \sum_{j \in N_{\mu^k}} \mu_j^k \nabla_x G_j(\tilde{x}^k, \tilde{y}^k), \tag{25}$$

$$\bar{v}^k = v^k - 2 \sum_{j \in N_{\mu^k}} \mu_j^k \nabla_y G_j(\tilde{x}^k, \tilde{y}^k), \tag{26}$$

$$\bar{w}_j^k = \begin{cases} w_j^k, & \text{if } j \in P_{\mu^k}, \\ -w_j^k, & \text{if } j \in N_{\mu^k}, \end{cases}$$

for $j = 1, \ldots, m$. By using these new inexact multipliers and residuals, the inexact KKT conditions (21)–(23) can be rewritten equivalently in the following form,

$$\frac{\tilde{x}^k - \bar{x}^k}{||\tilde{x}^k - \bar{x}^k||_2} + \sum_{j=1}^m \bar{\mu}_j^k \nabla_x G_j(\tilde{x}^k, \tilde{y}^k) = \bar{u}^k, \tag{27}$$

$$\sum_{j=1}^m \bar{\mu}_j^k \nabla_y G_j(\tilde{x}^k, \tilde{y}^k) = \bar{v}^k, \tag{28}$$

$$\bar{\mu}_j^k G_j(\tilde{x}^k, \tilde{y}^k) = \bar{w}_j^k, \quad j = 1, \ldots, m. \tag{29}$$

Here, $\bar{\mu}^k$ defined as in (24) is now a non-negative vector in $\mathbb{R}^m$. Thus, note that the conditions (27)–(29) have exactly the same form and properties as the inexact KKT conditions (1)–(3). Thus, the approach of Section 2 carry out to the setting of the current part, by simply replacing u^k, v^k, and w^k by $\bar{u}^k$, $\bar{v}^k$, and $\bar{w}^k$.

Now, we discuss how to control the new residuals, where we only consider $\bar{u}^k$ and $\bar{v}^k$, since $\|\bar{w}^k\| = \|w^k\|$. Firstly, we can find easily that the new residuals can be controlled as follows by (25) and (26):

$$\|\bar{u}^k\| \le \|u^k\| + 2 \sum_{j \in N_{\mu^k}} |\mu_j^k| \left\| \nabla_x G_j(\tilde{x}^k, \tilde{y}^k) \right\|,$$

$$\|\bar{v}^k\| \le \|v^k\| + 2 \sum_{j \in N_{\mu^k}} |\mu_j^k| \left\| \nabla_y G_j(\tilde{x}^k, \tilde{y}^k) \right\|$$

Then, if the quantity

$$\max_{j \in N_{\mu^k}} \left\{ \left\| \nabla_x G_j(\tilde{x}^k, \tilde{y}^k) \right\|, \left\| \nabla_y G_j(\tilde{x}^k, \tilde{y}^k) \right\| \right\}$$

is assumed to be bounded, we can obtain $\bar{u}^k$ (and $\bar{v}^k$) as small as we want by reducing the size of the residual u^k (and v^k) and the size of the negative elements in the inexact Lagrange multipliers,

$$\max_{j \in N_{\mu^k}} |\mu_j^k|,$$

possibly by resolving the NLP subproblem under tighter tolerances.

5. Conclusions and final remarks

In this paper, we have studied the effect of inexactness when the NLP problems are solved inexactly in the feasibility pump, for finding a feasible solution to MINLPs. We first gave some conditions to the residuals for making sure the convergence of *inexact* FP. In our test problems, we ignored these conditions, while some feasible MINLP problems became infeasible. Moreover, we can see that the cuts for *inexact* FP should be changed depending on the residuals. Otherwise, the number of iterations may become bigger. Furthermore, we extended the *inexact* FP to the case that the Lagrange multipliers returned may be negative.

How to find a feasible solution is our main work, but we cannot explain if it is with good objective value. So, we will continue studying the OFP in inexact case. Also, our work is based on the conditions that all the constraint functions of MINLPs are convex. The non-convex MINLP problems still have wide research background and also are studied actively [2,9]. We want to continue studying this respect using augmented Lagrange functions by adding the twice differentiable conditions on the constraint functions.

References

[1] T. Achterberg, T. Berthold, Improving the feasibility pump, Discrete Optim. 4 (2007) 77–86.
[2] C. D'Ambrosio, A. Frangioni, L. Liberti, A. Lodi, A storm of feasibility pumps for nonconvex MINLP, Math. Program., Ser. B 136 (2012) 375–462.
[3] T. Berthold, Primal Heuristics for Mixed Integer Programs, master's thesis, Technical University of Berlin, 2006.
[4] T. Berthold, A.M. Gleixner, Undercover: a primal MINLP heuristic exploring a largest sub-MIP, Math. Program., Ser. A 144 (2014) 315–346.
[5] L. Bertacco, M. Fischetti, A. Lodi, A feasibility pump heuristic for general mixed-integer problems, Discrete Optim. 4 (2007) 63–76.
[6] N.L. Boland, A.C. Eberhard, F.G. Engineer, M. Fischetti, M.W.P. Savelsbergh, A. Tsoukalas, Boosting the feasibility pump, Math. Program. Comput. 6 (2014) 255–279.
[7] P. Bonami, G. Cornuéjols, A. Lodi, F. Margot, A feasibility pump for mixed integer nonlinear programs, Math. Program. 119 (2009) 331–352.
[8] P. Bonami, J.P.M. Gonçalves, Heuristics for convex mixed integer nonlinear programs, Comput. Optim. Appl. 51 (2012) 729–747.
[9] S. Burer, Adam N. Letchford, Non-convex mixed-integer nonlinear programming: a survey, Surv. Oper. Res. Manag. Sci. 17 (2012) 97–106.
[10] E. Danna, E. Rothberg, C.L. Pape, Exploring relaxation induced neighborhoods to improve MIP solutions, Math. Program. 102 (2005) 71–90.
[11] M. Fischetti, F. Glover, A. Lodi, The feasibility pump, Math. Program. 104 (2005) 91–104.
[12] M. Fischetti, M. Monaci, Proximity search for 0-1 mixed-integer convex programming, J. Heuristics 20 (2014) 709–731.
[13] M. Fischetti, D. Salvagnin, Feasibility pump 2.0, Math. Program. Comput. 1 (2009) 201–222.
[14] J. Humpola, A. Fügenschuh, T. Lehmann, A primal heuristic for optimizing the topology of gas networks based on dual information, EURO J. Comput. Optim. 3 (2015) 53–78.
[15] S. Sharma, B.R. Knudsen, B. Grimstad, Towards an objective feasibility pump for convex MINLPs, Comput. Optim. Appl. 63 (2016) 737–753.
[16] M. Li, L.N. Vicente, Inexact solution of NLP subproblems in MINLP, J. Glob. Optim. 55 (2013) 877–899.
[17] CMU/IBM MINLP Project, http://egon.cheme.cmu.edu/ibm/page.htm.
[18] S. Leyffer, MacMINLP, http://wiki.mcs.anl.gov/leyffer/index.php/MacMINLP.
[19] IBM ILOG CPLEX, http://www-01.ibm.com/software/integration/optimization/cplex-optimizer, 2012, version 12.4.