



Bounds on the Objective Value of Feasible Roundings

Christoph Neumann¹ · Oliver Stein¹  · Nathan Sudermann-Merx²

Received: 19 July 2019 / Accepted: 7 January 2020 / Published online: 18 March 2020
© The Author(s) 2020

Abstract

For mixed-integer linear and nonlinear optimization problems we study the objective value of feasible points which are constructed by the feasible rounding approaches from Neumann et al. (Comput. Optim. Appl. **72**, 309–337, 2019; J. Optim. Theory Appl. **184**, 433–465, 2020). We provide a-priori bounds on the deviation of such objective values from the optimal value and apply them to explain and quantify the positive effect of finer grids of integer feasible points on the performance of the feasible rounding approaches. Computational results for large scale knapsack problems illustrate our theoretical findings.

Keywords Rounding · Granularity · Inner parallel set · Global error bound

Mathematics Subject Classification (2010) 90C11 · 90C10 · 90C31 · 90C30

1 Introduction

In this paper we study the quality of feasible points of mixed-integer linear and nonlinear optimization problems (MILPs and MINLPs, resp.) as they are constructed by the feasible rounding approaches from [19, 20]. These approaches are based on a property of the feasible set which we call granularity and which states that a certain inner parallel set of the continuously relaxed feasible set is nonempty. The main effect of granularity is that it relaxes the difficulties imposed by the integrality conditions and hence, under suitable assumptions, provides a setting in which feasible points of MINLPs may be generated at low computational cost.

Dedicated to Marco Antonio López-Cerdá at the occasion of his 70th birthday.

✉ Oliver Stein
stein@kit.edu

Christoph Neumann
christoph.neumann@kit.edu

Nathan Sudermann-Merx
nathan.sudermann-merx@basf.com

¹ Institute of Operations Research, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

² Advanced Business Analytics, BASF SE, Ludwigshafen, Germany

Our present analysis is motivated by the successful application of this algorithmic approach to mixed-integer linear and nonlinear problems in [19] and [20], respectively. In these papers computational studies of problems from the MIPLIB and MINLPLib libraries show that granularity may be expected and exploited in various real world applications. Moreover, the practical performance is observed to improve for optimization problems with finer grids of integer feasible points, that is, with ‘more integer feasible points relative to the size of the continuously relaxed feasible set’. This positive effect does not only refer to the applicability of the granularity concept, but also to the quality of the generated feasible points in terms of their objective values.

In fact, in applications a small deviation of this objective value from the minimal value may lead to the decision to accept the feasible point as ‘close enough to optimal’. Otherwise, a feasible point with low objective value may be used to initialize an appropriate branch-and-cut method with a small upper bound on the optimal value, or to start a local search heuristic there (cf., e.g., [4]).

The remainder of this article is structured as follows. Section 2 recalls some preliminaries from [19, 20], such as the main construction needed for the definition of granularity, an explicit description of a subset of the appearing inner parallel set, and enlargement ideas which promote the performance of the resulting feasible rounding approach. Section 3 provides a-priori bounds on the deviation of the objective value of the generated feasible point from the optimal value, before Section 4 applies these bounds to explain and quantify the positive effect of fine integer grids on the performance of the feasible rounding approach. In Section 5 we illustrate our theoretical findings by computational results for large scale knapsack problems, and some conclusions and final remarks end the article in Section 6.

2 Preliminaries

We study mixed-integer nonlinear optimization problems of the form

$$\text{MINLP} : \quad \min_{(x,y) \in \mathbb{R}^n \times \mathbb{Z}^m} f(x,y) \quad \text{s.t.} \quad g_i(x,y) \leq 0, \quad i \in I, \quad (x,y) \in D$$

with real-valued functions $f, g_i, i \in I$, defined on $\mathbb{R}^n \times \mathbb{R}^m$, a finite index set $I = \{1, \dots, q\}, q \in \mathbb{N}$, and a nonempty polyhedral set

$$D = \{(x,y) \in \mathbb{R}^n \times \mathbb{R}^m \mid Ax + By \leq b\}$$

with some (p, n) -matrix A , (p, m) -matrix B and $b \in \mathbb{R}^p, p \in \mathbb{N}$.

To verify granularity of a given problem *MINLP* we will impose additional Lipschitz assumptions for the functions $f, g_i, i \in I$, on the set D , and to state the a-priori bounds on objective values we shall further require convexity of the functions $f, g_i, i \in I$. However, these additional assumptions will only be introduced where necessary, since the granularity concept also covers various nonlinear instances which, in particular, go beyond the case of mixed-integer convex optimization problems (MICPs). On the other hand, we shall specify our general results to mixed-integer linear optimization problems (MILPs) when appropriate (cf. Example 2 and Corollary 1) and, in particular, we will illustrate them in Section 5 along some MILP. Note that the presented results are novel not only for MINLPs, but also for MILPs. While the purely integer case ($n = 0$) is included in our analysis, we will assume $m, p > 0$ throughout this article.

2.1 Granularity

In the following let us recall some constructions which were presented in [19, 20] for the case of mixed-integer linear optimization problems (MILPs) and mixed-integer nonlinear optimization problems (MINLPs), respectively. We shall denote the feasible set of the NLP relaxation \widehat{MINLP} of $MINLP$ by

$$\widehat{M} = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m \mid (x, y) \in D, g(x, y) \leq 0\},$$

where g denotes the vector of functions $g_i, i \in I$. Moreover, for any point $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$ we call (\check{x}, \check{y}) a rounding if

$$\check{x} = x \quad \text{and} \quad \check{y} \in \mathbb{Z}^m, \quad |\check{y}_j - y_j| \leq \frac{1}{2}, \quad j = 1, \dots, m,$$

hold, that is, y is rounded componentwise to a point in the mesh \mathbb{Z}^m , and x remains unchanged. Note that a rounding does not have to be unique.

With the sets

$$B_\infty\left(0, \frac{1}{2}\right) := \{y \in \mathbb{R}^m \mid \|y\|_\infty \leq \frac{1}{2}\} \quad \text{and} \quad K := \{0\} \times B_\infty\left(0, \frac{1}{2}\right)$$

any rounding of (x, y) obviously satisfies

$$(\check{x}, \check{y}) \in ((x, y) + K) \cap (\mathbb{R}^n \times \mathbb{Z}^m). \tag{1}$$

The central object of our technique is the *inner parallel set* of \widehat{M} with respect to K ,

$$\widehat{M}^- := \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m \mid (x, y) + K \subseteq \widehat{M}\}.$$

Any rounding of any point $(x, y) \in \widehat{M}^-$ must lie in the feasible set M of $MINLP$ since in view of (1) it satisfies

$$(\check{x}, \check{y}) \in ((x, y) + K) \cap (\mathbb{R}^n \times \mathbb{Z}^m) \subseteq \widehat{M} \cap (\mathbb{R}^n \times \mathbb{Z}^m) = M.$$

Hence, if the inner parallel set \widehat{M}^- is nonempty, then also M is nonempty. Of course, this observation is only useful if the inner parallel set is nonempty, which gives rise to the following definition.

Definition 1 We call the set M *granular* if the inner parallel set \widehat{M}^- of \widehat{M} is nonempty. Moreover, we call a problem $MINLP$ granular if its feasible set M is granular.

We remark that [19, 20] provide several examples for granular problems in the linear as well as in the nonlinear case.

In the terminology of Definition 1 our above observation states that any granular problem $MINLP$ is consistent. Firstly, this gives rise to a feasibility test for MINLPs and, secondly, for any granular $MINLP$ one may aim at the explicit computation of some feasible point. For a discussion of the former aspect we refer to [19, 20] and rather focus on the latter in the present paper.

To this end we need to compute at least a subset T^- of \widehat{M}^- explicitly which, like the set \widehat{M}^- , is not restricted by integrality constraints. The general idea of the *feasible rounding approaches* from [19, 20] is to minimize f over T^- and round any optimal point to a point in M . This employment of the objective function f aims at obtaining a feasible point with a reasonably good objective value. In Section 3 we shall quantify how ‘bad’ this objective value may be in the worst case.

2.2 A Functional Description for the Inner Parallel Set

To obtain a functional description of some set $T^- \subseteq \widehat{M}^-$ observe that with the abbreviation

$$G := \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m \mid g(x, y) \leq 0\}$$

we may write $\widehat{M} = D \cap G$ and that the inner parallel set of \widehat{M} thus satisfies $\widehat{M}^- = D^- \cap G^-$. From [25, Lemma 2.3] we know the closed form expression for the inner parallel set of D ,

$$D^- = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m \mid Ax + By + \|\beta\|_1/2 \leq b\}, \tag{2}$$

where $\beta_i^T, i = 1, \dots, p$, denote the rows of the matrix B and, by a slight abuse of notation, $\|\beta\|_1$ stands for the vector $(\|\beta_1\|_1, \dots, \|\beta_p\|_1)^T$.

Moreover, the definition of the set G^- yields its semi-infinite description

$$\begin{aligned} G^- &= \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m \mid (x, y) + K \subseteq G\} \\ &= \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m \mid g(x, y + \eta) \leq 0 \ \forall \eta \in B_\infty(0, \frac{1}{2})\}. \end{aligned}$$

For the derivation of some algorithmically tractable inner approximation of G^- we employ global Lipschitz conditions with respect to y uniformly in x for the functions $g_i, i \in I$, on the set D . This distinction between the roles of x and y is caused by the definition of inner parallel sets, whose geometric construction only depends on the discrete variable y .

In fact, for any $x \in \mathbb{R}^n$ we define the set $D(x) := \{y \in \mathbb{R}^m \mid (x, y) \in D\}$ and denote by $\text{pr}_x D := \{x \in \mathbb{R}^n \mid D(x) \neq \emptyset\}$ the parallel projection of D to the ‘ x -space’ \mathbb{R}^n . Then the functions $g_i, i \in I$, are assumed to satisfy Lipschitz conditions with respect to the ℓ_∞ -norm on the fibers $\{x\} \times D(x)$, independently of the choice of $x \in \text{pr}_x D$.

Assumption 1 *For all $i \in I$ there exists some $L_\infty^i \geq 0$ such that for all $x \in \text{pr}_x D$ and all $y^1, y^2 \in D(x)$ we have*

$$|g_i(x, y^1) - g_i(x, y^2)| \leq L_\infty^i \|(x, y^1) - (x, y^2)\|_\infty = L_\infty^i \|y^1 - y^2\|_\infty.$$

Some problem classes for which the Lipschitz constants from Assumption 1 can be calculated are discussed in [20]. In particular, if the set D is bounded and the functions $g_i, i \in I$, are continuously differentiable with respect to y , one may choose

$$L_\infty^i = \max_{(x,y) \in D} \|\nabla_y g_i(x, y)\|_1, \quad i \in I,$$

which allows to compute such Lipschitz constants for many test instances from the MINLPLib [20].

Under Assumption 1, and with L_∞ denoting the vector of Lipschitz constants $L_\infty^i, i \in I$, one may define the set

$$T^- := \{(x, y) \in D^- \mid g(x, y) + L_\infty/2 \leq 0\}$$

and show the desired inclusion $T^- \subseteq \widehat{M}^-$ [20]. Recall that thus any rounding (\check{x}, \check{y}) of any point $(x, y) \in T^-$ lies in M .

2.3 Enlargements

We point out that it may depend on the geometry of the relaxed feasible set \widehat{M} whether the feasible set M of MINLP is granular or not. In particular, for MINLPs with binary variables the standard formulation of \widehat{M} would usually lead to an empty inner parallel set \widehat{M}^- [19, 20]. On the other hand, the set \widehat{M} may often be replaced by a set \widetilde{M} in such a way that the corresponding new inner parallel set \widetilde{M}^- is larger than \widehat{M}^- , without losing the

property that any rounding of any of its elements lies in M . This admits to exploit granularity also for many MINLPs with binary variables [19, 20].

The main idea for the construction of such enlargements is a preprocessing step for the functional description of $MINLP$. It first enlarges the relaxed feasible set \widehat{M} of M to some set $\widetilde{M} \supseteq \widehat{M}$ for which the feasible set M of $MINLP$ can still be written as

$$M = \widetilde{M} \cap (\mathbb{R}^n \times \mathbb{Z}^m). \tag{3}$$

Then we call the inner parallel set

$$\widetilde{M}^- = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m \mid (x, y) + K \subseteq \widetilde{M}\}$$

of \widetilde{M} an *enlarged inner parallel set* of \widehat{M} since the relation $\widehat{M} \subseteq \widetilde{M}$ implies $\widehat{M}^- \subseteq \widetilde{M}^-$. Depending on the functional description of \widetilde{M} and, in particular, the appearing Lipschitz constants, the inner approximation T^- of \widetilde{M}^- may then be larger than T^- .

While there are several options for the construction of enlargements of the set \widehat{M} [20], in the following let us focus on those resulting from constant additive relaxations $Ax + By \leq b + \sigma$ and $g(x, y) \leq \tau$ of its constraints with appropriately chosen vectors $\sigma, \tau \geq 0$. Note that this approach maintains algorithmically attractive properties like the polyhedrality of D and differentiability or convexity of the functions $g_i, i \in I$. We set

$$D_\sigma := \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m \mid Ax + By \leq b + \sigma\}$$

as well as

$$G_\tau := \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m \mid g(x, y) \leq \tau\}.$$

Clearly, for each $\rho := (\sigma, \tau) \geq 0$ the set

$$\widehat{M}_\rho := D_\sigma \cap G_\tau \tag{4}$$

satisfies $\widehat{M} \subseteq \widehat{M}_\rho$. If we denote the appropriate choices of ρ for (3) by

$$R := \{\rho \in \mathbb{R}^p \times \mathbb{R}^q \mid \rho \geq 0, M = \widehat{M}_\rho \cap (\mathbb{R}^n \times \mathbb{Z}^m)\}$$

then for each $\rho \in R$ any rounding of any element of \widehat{M}_ρ^- lies in M . Furthermore, we have $\widehat{M}^- \subseteq \widehat{M}_\rho^-$, so that \widehat{M}_ρ^- is more likely to be nonempty than \widehat{M}^- . In fact, after preprocessing \widehat{M} to \widehat{M}_ρ for some $\rho \in R$, according to Definition 1 the set M and the problem $MINLP$ are granular and, thus, consistent if the enlarged inner parallel set \widehat{M}_ρ^- is nonempty. Note that due to (4) we may write $\widehat{M}_\rho^- = D_\sigma^- \cap G_\tau^-$ with

$$D_\sigma^- = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m \mid Ax + By + \|\beta\|_1/2 \leq b + \sigma\}.$$

Moreover,

$$T_\rho^- = \{(x, y) \in D_\sigma^- \mid g(x, y) + L_\infty/2 \leq \tau\}$$

is an inner approximation of \widehat{M}_ρ^- , where the entries of the vector L_∞ are Lipschitz constants of the functions $g_i(x, y) - \tau_i, i \in I$, on D_σ in the sense of Assumption 1. Observe that, while these Lipschitz constants do not depend on τ , they may well depend on σ . This leads to the undesirable issue that for $\rho \in R$, despite the inclusion $\widehat{M}^- \subseteq \widehat{M}_\rho^-$, the corresponding inner approximations T^- and T_ρ^- do not necessarily satisfy $T^- \subseteq T_\rho^-$. In the present paper we do not further discuss this problem, but refer to [20] for its treatment by the alternative concept of pseudo-granularity.

As a consequence, here we shall study the following version of the *feasible rounding approach by shrink-optimize-round (FRA-SOR)* [19, 20]: For a given problem $MINLP$ compute enlargement parameters $\rho = (\sigma, \tau) \in R$ and corresponding Lipschitz constants of

the functions $g_i(x, y) - \tau_i, i \in I$, on D_σ in the sense of Assumption 1. Then compute an optimal point (x_ρ^s, y_ρ^s) of f over T_ρ^- , that is, of the problem

$$P_\rho^s : \min_{(x,y) \in \mathbb{R}^n \times \mathbb{R}^m} f(x, y) \quad \text{s.t. } Ax + By + \|\beta\|_1/2 \leq b + \sigma, \\ g(x, y) + L_\infty^\sigma/2 \leq \tau,$$

and round it to $(\check{x}_\rho^s, \check{y}_\rho^s) \in M$.

Due to rounding effects as well as due to the necessary modifications on the transition from M to T_ρ^- , the generated point $(\check{x}_\rho^s, \check{y}_\rho^s)$ cannot be expected to be optimal for $MINLP$. Yet, in the next section we shall show that one can use the close relation of the sets T_ρ^- and \widehat{M}_ρ to derive an upper bound on the objective value of $(\check{x}_\rho^s, \check{y}_\rho^s)$ that merely depends on the problem data.

3 Bounds on the Objective Value

Since any point $(\check{x}_\rho^s, \check{y}_\rho^s)$ generated by FRA-SOR is feasible for $MINLP$, its objective value $\check{v}_\rho^s := f(\check{x}_\rho^s, \check{y}_\rho^s)$ exceeds the optimal value v of $MINLP$. Unfortunately, simple examples illustrate that the gap between \check{v}_ρ^s and v may actually be arbitrarily large [20]. The main aim of this section is to state an upper bound for the gap $\check{v}_\rho^s - v$ in terms of the problem data.

As v is unknown, we bound it in terms of the optimal value \widehat{v} of the continuously relaxed problem \widehat{MINLP} by

$$0 \leq \check{v}_\rho^s - v \leq \check{v}_\rho^s - \widehat{v}. \tag{5}$$

This bound may be computed explicitly. In fact, after the solution of P_ρ^s , in addition only the relaxed problem \widehat{MINLP} has to be solved. Note that in this bound we propose to use the optimal value \widehat{v} of f over \widehat{M} without any enlargement constructions, rather than the optimal value of f over \widehat{M}_ρ , since this leads to a tighter bound.

While such an a-posteriori bound can be achieved at low computational cost under suitable assumptions, it is not useful for investigations with regard to the dependence of the achieved objective value on the problem data. Hence, next we shall derive an a-priori bound for $\check{v}_\rho^s - v$ which does not depend on the solution of some auxiliary optimization problem, but merely on the data of $MINLP$.

In particular, in Section 4 we will be interested in the behavior of $(\check{v}_\rho^s - v)/|v|$ for different degrees of integer grid fineness in $MINLP$, as this will not only confirm but also quantify the empirically observed fact from [19, 20] that finer grids lead to smaller relative deviations between \check{v}_ρ^s and v .

For the derivation of the main results in this section let us temporarily ignore the possibility of enlargements by some $\rho \in R$, but consider the original problem corresponding to $\rho = 0$. For our analysis we will assume $T^- \neq \emptyset$, since otherwise FRA-SOR does not provide a point $(\check{x}^s, \check{y}^s)$. Subsequently

$$\text{dist}((\widehat{x}, \widehat{y}), T^-) := \inf_{(x,y) \in T^-} \|(x, y) - (\widehat{x}, \widehat{y})\| \tag{6}$$

shall denote the distance of some point $(\widehat{x}, \widehat{y}) \in \mathbb{R}^n \times \mathbb{R}^m$ to the set T^- with respect to some norm $\|\cdot\|$ on $\mathbb{R}^n \times \mathbb{R}^m$. In addition to the uniform Lipschitz continuity of the functions $g_i, i \in I$, with respect to the ℓ_∞ -norm from Assumption 1, in the following we will also need Lipschitz continuity of f with respect to the norm from (6).

Assumption 2 *There exists some $L^f \geq 0$ such that for all $(x^1, y^1), (x^2, y^2) \in D$ we have*

$$|f(x^1, y^1) - f(x^2, y^2)| \leq L^f \|(x^1, y^1) - (x^2, y^2)\|.$$

Furthermore, in the following $L^f_\infty \geq 0$ will denote a Lipschitz constant with respect to y uniformly in x for f on D with respect to the ℓ_∞ -norm. Under Assumption 2 a possible, but not necessarily tight, choice is $L^f_\infty := \kappa L^f$ with some norm constant $\kappa > 0$ such that for all $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$ we have $\|(x, y)\| \leq \kappa \|(x, y)\|_\infty$. In particular, the existence of L^f_∞ follows from Assumption 2. However, the following example indicates a better choice for linear objective functions.

Example 1 For a linear objective function $f(x, y) = c^T x + d^T y$ the best possible Lipschitz constant on $\mathbb{R}^n \times \mathbb{R}^m$, that is, the Lipschitz modulus

$$\sup_{(x^1, y^1) \neq (x^2, y^2)} \frac{|f(x^1, y^1) - f(x^2, y^2)|}{\|(x^1, y^1) - (x^2, y^2)\|},$$

is easily seen to coincide with the dual norm of (c, d) , so that we may set

$$L^f := \|(c, d)\|^* := \max\{c^T x + d^T y \mid \|(x, y)\| \leq 1\}.$$

Moreover, we may choose

$$L^f_\infty := \|d\|_\infty^* = \|d\|_1.$$

Lemma 1 *Let Assumptions 1 and 2 hold, let (\hat{x}^*, \hat{y}^*) denote any optimal point of \widehat{MINLP} , and let $(\check{x}^s, \check{y}^s)$ denote any rounding of any optimal point (x^s, y^s) of P^s . Then the value $\check{v}^s = f(\check{x}^s, \check{y}^s)$ satisfies*

$$0 \leq \check{v}^s - v \leq L^f_\infty/2 + L^f \text{dist}((\hat{x}^*, \hat{y}^*), T^-).$$

Proof As above, the first inequality stems from the feasibility of $(\check{x}^s, \check{y}^s)$ for $MINLP$. For the proof of the second inequality note that, with any projection (x^π, y^π) of (\hat{x}^*, \hat{y}^*) onto the set T^- with respect to $\|\cdot\|$, the upper bound $\check{v}^s - \widehat{v}$ from (5) may be written as $\check{v}^s - \widehat{v} = (f(\check{x}^s, \check{y}^s) - f(x^s, y^s)) + (f(x^s, y^s) - f(x^\pi, y^\pi)) + (f(x^\pi, y^\pi) - f(\hat{x}^*, \hat{y}^*))$.

Due to $\check{x}^s = x^s$, the first term satisfies

$$f(\check{x}^s, \check{y}^s) - f(x^s, y^s) \leq L^f_\infty \|(x^s, \check{y}^s) - (x^s, y^s)\|_\infty = L^f_\infty \|\check{y}^s - y^s\|_\infty \leq L^f_\infty/2.$$

Since (x^s, y^s) is an optimal point of P^s , while (x^π, y^π) is a feasible point, for the second term we obtain

$$f(x^s, y^s) - f(x^\pi, y^\pi) \leq 0.$$

Finally, as the distance is the optimal value of the corresponding projection problem, the third term can be bounded by

$$f(x^\pi, y^\pi) - f(\hat{x}^*, \hat{y}^*) \leq L^f \|(x^\pi, y^\pi) - (\hat{x}^*, \hat{y}^*)\| = L^f \text{dist}((\hat{x}^*, \hat{y}^*), T^-),$$

and the assertion is shown. □

It remains to bound the expression $\text{dist}((\hat{x}^*, \hat{y}^*), T^-)$ from the upper bound in Lemma 1 in terms of the problem data. To this end, we will employ a global error bound for the system of inequalities describing T^- . For the statement of this global error bound we construct the penalty function

$$\|(Ax + By - b + \|\beta\|_1/2)^+, (g(x, y) + L_\infty/2)^+\|_\infty$$

of the set T^- where $a^+ := (\max\{0, a_1\}, \dots, \max\{0, a_N\})^\top$ denotes the componentwise positive-part operator for vectors $a \in \mathbb{R}^N$. A global error bound relates the geometric distance to the (consistent) set T^- with the evaluation of its penalty function by stating the existence of a constant $\gamma > 0$ such that for all $(\hat{x}, \hat{y}) \in \mathbb{R}^n \times \mathbb{R}^m$ we have

$$\text{dist}((\hat{x}, \hat{y}), T^-) \leq \gamma \|(A\hat{x} + B\hat{y} - b + \|\beta\|_1/2)^+, (g(\hat{x}, \hat{y}) + L_\infty/2)^+\|_\infty. \tag{7}$$

As Hoffman showed the existence of such a bound for any linear system of inequalities in his seminal work [9], γ is also called a *Hoffman constant*, and the error bound (7) is known as a *Hoffman error bound*. Short proofs of this result for the polyhedral case can be found in [7, 11]. For global error bounds of broader problem classes see, for example, [1, 6, 12–18, 24], and [2, 3, 22] for surveys. These references also contain sufficient conditions for the existence of global error bounds. To cite an early result for the nonlinear case from [24], if for convex functions $g_i, i \in I$, the set T^- is bounded and satisfies Slater’s condition, then a global error bound holds.

The next result simplifies the error bound (7) for points $(\hat{x}, \hat{y}) \in \widehat{M}$. It was used analogously in [26, Theorem 3.3] and follows from the subadditivity of the max operator, the monotonicity of the ℓ_∞ -norm, as well as $(A\hat{x} + B\hat{y} - b)^+ = 0$ and $g^+(\hat{x}, \hat{y}) = 0$ for any $(\hat{x}, \hat{y}) \in \widehat{M}$. Furthermore, we use that the appearing term $\|(\|\beta\|_1)\|_\infty$ coincides with the maximal absolute row sum $\|B\|_\infty$ of the matrix B .

Lemma 2 *Let Assumption 1 hold, and let the error bound (7) hold with some $\gamma > 0$. Then all $(\hat{x}, \hat{y}) \in \widehat{M}$ satisfy*

$$\text{dist}((\hat{x}, \hat{y}), T^-) \leq \gamma \max\{\|B\|_\infty, \|L_\infty\|_\infty\}/2.$$

The combination of Lemma 1 and Lemma 2 yields the main result of this section.

Theorem 3 *Let Assumptions 1 and 2 hold, and let the error bound (7) hold with some $\gamma > 0$. Then the objective value \check{v}^s of any rounding of any optimal point of P^s satisfies*

$$0 \leq \check{v}^s - v \leq \left(L_\infty^f + L^f \gamma \max\{\|B\|_\infty, \|L_\infty\|_\infty\} \right) / 2.$$

Example 2 For a mixed-integer linear problem *MILP* the nonlinear function g is absent (i.e., $q = 0$), and f has the form from Example 1. Furthermore, from [9] it is known that for polyhedral constraints a global error bound always holds so that this assumption may be dropped from Theorem 3. In view of Example 1 this results in the a-priori bound

$$0 \leq \check{v}^s - v \leq (\|d\|_1 + \|(c, d)\|_* \gamma \|B\|_\infty) / 2.$$

The result of Theorem 3 remains valid for any enlargement vector $\rho \in R$ with minor modifications. In fact, if γ satisfies the error bound estimate

$$\text{dist}((\hat{x}, \hat{y}), T_\rho^-) \leq \gamma \|(A\hat{x} + B\hat{y} - b - \sigma + \|\beta\|_1/2)^+, (g(\hat{x}, \hat{y}) - \tau + L_\infty/2)^+\|_\infty$$

instead of (7), then also the objective value \check{v}_ρ^s of any rounding of any optimal point of P_ρ^s satisfies the estimates from Theorem 3. We remark that for the proof of this result Lemma 1 needs to consider an optimal point $(\hat{x}_\rho^*, \hat{y}_\rho^*)$ of \widehat{MINLP}_ρ .

4 The Effect of the Integer Grid Fineness on the Objective Value

As mentioned before, we are particularly interested in a-priori bounds which explain the behavior of $(\check{v}_\rho^s - v)/|v|$ for different degrees of integer grid fineness in $MINLP$. To motivate our subsequent model for this effect, let us start by considering the purely integer linear problem

$$ILP_t : \quad \min_{y \in \mathbb{Z}^m} d^T y \quad \text{s.t.} \quad By \leq b + t\bar{b}$$

with a parameter $t > 0$ and a fixed vector $\bar{b} > 0$ which perturb the right-hand side vector b . Increasing values of t increase the number of feasible integer points and, in relation to the relaxed feasible sets $\widehat{M}(t)$, the grid becomes finer.

For increasing values of t we wish to analyze the quality of the feasible point generated by FRA-SOR for ILP_t , that is, the deviation of its objective value \check{v}_t^s from v_t . However, since also $|v_t|$ increases with t , instead of the absolute deviation we consider the behavior of the relative deviation $(\check{v}_t^s - v_t)/|v_t|$.

To make such an analysis also applicable to nonlinear constraints, let us focus on the relative effect of the integer grid fineness on the objective value. This may be modeled by a parameter $h > 0$ which scales the variables (x, y) to (hx, hy) and leads to the parametric problem

$$MINLP_h : \quad \min_{(x,y) \in \mathbb{R}^n \times \mathbb{Z}^m} f(hx, hy) \quad \text{s.t.} \quad g(hx, hy) \leq 0, \quad (hx, hy) \in D$$

with optimal value v_h . The above effect for increasing t is now translated to decreasing $h > 0$ and, due to $hy \in h\mathbb{Z}^m$, for $h \searrow 0$ the grid $h\mathbb{Z}^m$ indeed becomes finer. A similar construction is mentioned in [25], but there it is neither used explicitly, nor does [25] provide bounds on the objective values of feasible points.

To apply the results on a-priori bounds from Section 3 to $MINLP_h$, for simplicity let us again ignore enlargement constructions, and let us define the functions $f^h(x, y) := f(hx, hy)$ and $g^h(x, y) := g(hx, hy)$ as well as the matrices $A^h := hA$ and $B^h := hB$. We may then rewrite $MINLP_h$ as

$$\min_{(x,y) \in \mathbb{R}^n \times \mathbb{Z}^m} f^h(x, y) \quad \text{s.t.} \quad A^h x + B^h y \leq b, \quad g^h(x, y) \leq 0,$$

and the application of FRA-SOR consists in rounding an optimal point $(\widehat{x}_h^s, \widehat{y}_h^s)$ of

$$P_h^s : \quad \min_{(x,y) \in \mathbb{R}^n \times \mathbb{R}^m} f^h(x, y) \quad \text{s.t.} \quad A^h x + B^h y + \|\beta^h\|_1/2 \leq b, \quad g^h(x, y) + L_\infty^h/2 \leq 0,$$

to $(\check{x}_h^s, \check{y}_h^s)$ with objective value $\check{v}_h^s := f^h(\check{x}_h^s, \check{y}_h^s)$. Here, the vector $\|\beta^h\|_1$ coincides with $h\|\beta\|_1$, and for each $i \in I$ the i -th entry of the vector L_∞^h denotes the Lipschitz constant of g_i^h on $\mathbb{R}^n \times \mathbb{R}^m$ with respect to y and uniformly in x . It is easily seen to coincide with hL_∞^i , if the functions $g_i, i \in I$, satisfy Assumption 1 with Lipschitz constants $L_\infty^i, i \in I$.

In this notation, the above mentioned empirical observation from [19, 20] is that the relative bound $(\check{v}_h^s - v_h)/|v_h|$ seems to tend to zero for $h \searrow 0$. In the remainder of this section we shall prove this conjecture and quantify the corresponding rate of decrease.

In the subsequent result we will use that the optimal value \widehat{v} of the continuously relaxed problem \widehat{MINLP}_h does not depend on h , since in the relaxed problem the substitution of (hx, hy) by (x, y) is just a (scaling) transformation of coordinates. We will also assume $\widehat{v} > 0$, which may always be attained by adding a suitable constant to the objective function f .

Furthermore, we shall use the set

$$\tilde{T}_h^- := \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m \mid Ax + By + h\|\beta\|_1/2 \leq b, g(x, y) + hL_\infty/2 \leq 0\}$$

and assume that it satisfies the error bound

$$\text{dist}((\hat{x}, \hat{y}), \tilde{T}_h^-) \leq \gamma_h \|(A\hat{x} + B\hat{y} - b + h\|\beta\|_1/2)^+, (g(\hat{x}, \hat{y}) + hL_\infty/2)^+\|_\infty \tag{8}$$

for all $(\hat{x}, \hat{y}) \in \mathbb{R}^n \times \mathbb{R}^m$ with the Hoffman constant $\gamma_h > 0$.

Lemma 3 *Let $\hat{v} > 0$, let f and g satisfy Assumptions 2 and 1, respectively, for $h > 0$ let \tilde{T}_h^- be nonempty, and let the error bound (8) hold with some $\gamma_h > 0$. Then the objective value \check{v}_h^s of any rounding of any optimal point of P_h^s satisfies*

$$0 \leq \frac{\check{v}_h^s - v_h}{v_h} \leq h \frac{L_\infty^f + L^f \gamma_h \max\{\|B\|_\infty, \|L_\infty\|_\infty\}}{2\hat{v}}. \tag{9}$$

Proof Since the denominator v_h is bounded below by the optimal value $\hat{v} > 0$ of the continuous relaxation, we obtain

$$0 \leq \frac{\check{v}_h^s - v_h}{v_h} \leq \frac{\check{v}_h^s - v_h}{\hat{v}}.$$

Moreover, the numerator $\check{v}_h^s - v_h$ may be bounded above by applying Theorem 3 to the problem $MINLP_h$. In fact, it is easy to see that Assumption 2 for f with the Lipschitz constant L^f implies that f^h satisfies Assumption 2 with the Lipschitz constant hL^f . Analogously, hL_∞^f and hL_∞^i are Lipschitz constants with respect to y uniformly in x for f^h and $g_i^h, i \in I$, respectively, with respect to the ℓ_∞ -norm.

Since a point (x, y) lies in

$$\begin{aligned} T_h^- &:= \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m \mid A^h x + B^h y + h\|\beta\|_1/2 \leq b, g^h(x, y) + hL_\infty/2 \leq 0\} \\ &= \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m \mid A(hx) + B(hy) + h\|\beta\|_1/2 \leq b, g(hx, hy) + hL_\infty/2 \leq 0\} \end{aligned}$$

if and only if $(\xi, \eta) := (hx, hy)$ lies in \tilde{T}_h^- , the assumption $\tilde{T}_h^- \neq \emptyset$ implies that also T_h^- is nonempty. Furthermore, as an error bound for the set T_h^- we obtain for all $(\hat{x}, \hat{y}) \in \mathbb{R}^n \times \mathbb{R}^m$

$$\begin{aligned} \text{dist}((\hat{x}, \hat{y}), T_h^-) &= \inf_{(x, y) \in T_h^-} \|(x, y) - (\hat{x}, \hat{y})\| = \inf_{(hx, hy) \in \tilde{T}_h^-} \|(x, y) - (\hat{x}, \hat{y})\| \\ &= \inf_{(\xi, \eta) \in \tilde{T}_h^-} \|(\xi, \eta)/h - (\hat{x}, \hat{y})\| = \inf_{(\xi, \eta) \in \tilde{T}_h^-} \|(\xi, \eta) - (h\hat{x}, h\hat{y})\|/h \\ &= \text{dist}((h\hat{x}, h\hat{y}), \tilde{T}_h^-)/h \\ &\leq \gamma_h \|(A(h\hat{x}) + B(h\hat{y}) - b + h\|\beta\|_1/2)^+, (g(h\hat{x}, h\hat{y}) + hL_\infty/2)^+\|_\infty/h \\ &= \gamma_h \|(A^h\hat{x} + B^h\hat{y} - b + h\|\beta\|_1/2)^+, (g^h(\hat{x}, \hat{y}) + hL_\infty/2)^+\|_\infty/h, \end{aligned}$$

that is, γ_h/h serves as a Hoffman constant for the system of inequalities describing T_h^- . Theorem 3 thus yields

$$\begin{aligned} \check{v}_h^s - v_h &\leq \left((hL_\infty^f) + (hL^f)(\gamma_h/h) \max\{h\|B\|_\infty, \|(hL_\infty)\|_\infty\} \right) / 2 \\ &= h \left((L_\infty^f) + L^f \gamma_h \max\{\|B\|_\infty, \|L_\infty\|_\infty\} \right) / 2. \end{aligned}$$

The combination of the estimates for the numerator and the denominator now yields the assertion. □

Theorem 4 *Let $\widehat{v} > 0$, let f and g satisfy Assumptions 2 and 1, respectively, let the functions $g_i, i \in I$, be convex on $\mathbb{R}^n \times \mathbb{R}^m$, and let the set $\widehat{M} = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m \mid Ax + By \leq b, g(x, y) \leq 0\}$ be bounded and satisfy Slater’s condition. Then for all sufficiently small $h > 0$ the set T_h^- is nonempty, the error bound (8) holds with some $\gamma_h > 0$, the relative bound $(\check{v}_h^s - v_h)/v_h$ satisfies the relations (9), and it tends to zero at least linearly with $h \searrow 0$.*

Proof By Slater’s condition for the set \widehat{M} there is a point (\bar{x}, \bar{y}) with $A\bar{x} + B\bar{y} < b$ and $g(\bar{x}, \bar{y}) < 0$. Then for all sufficiently small $h > 0$ the point (\bar{x}, \bar{y}) is also a Slater point of T_h^- . Firstly, this implies that also T_h^- is nonempty for these h . Secondly, since T_h^- is bounded as a subset of the bounded set \widehat{M} , by [24] the error bound (8) holds with some $\gamma_h > 0$. The relations (9) thus follow from Lemma 3. Finally, the linear decrease of the relative bounds with $h \searrow 0$ is due to the fact that the Hoffman constants γ_h remain bounded for all sufficiently small $h > 0$. This is shown in [21] for convex problems under mild assumptions and may be applied to the present setting along the lines of the proof of [26, Corollary 3.6]. □

We point out that the assumptions of Theorem 4 may be significantly relaxed in the MILP case from Example 2. Then not only Assumptions 2 and 1 hold with $L^f = \|(c, d)\|^*$ and $L_\infty = \|\beta\|_1$ as well as $L_\infty^f = \|d\|_1$, but from [9] it is also known that for polyhedral constraints the error bound (8) is satisfied without further assumptions, and that the corresponding Hoffman constant γ may even be chosen independently of the right-hand side vector and, thus, in our case independently of h . Altogether, this shows the following result.

Corollary 1 *For an MILP let $\widehat{v} > 0$, and let the set $\widehat{M} = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m \mid Ax + By \leq b\}$ satisfy Slater’s condition. Then for all sufficiently small $h > 0$ the set T_h^- is nonempty, the error bound (8) holds with some $\gamma > 0$, the relative bound $(\check{v}_h^s - v_h)/v_h$ satisfies the relations*

$$0 \leq \frac{\check{v}_h^s - v_h}{v_h} \leq h \frac{\|d\|_1 + \|(c, d)\|^* \gamma \|B\|_\infty}{2\widehat{v}}$$

and, thus, it tends to zero at least linearly with $h \searrow 0$.

5 An Application to Bounded Knapsack Problems

The following computational study comprises results for the bounded knapsack problem which was introduced in [5] and is known to be an NP-hard MILP (cf. [10, pp. 483–491]). In its original formulation, which is also called the *0-1 knapsack problem*, all decision variables are binary. The *bounded knapsack problem* (BKP) is a generalization of the 0-1 knapsack problem where it is possible to pick more than one piece per item, that is, the integer decision variables may not be binary. A possible numerical approach to bounded knapsack problems is to transform them into equivalent 0-1 knapsack problems for which solution techniques exist that perform very well in practical applications. In contrast to this approach we exploit granularity of the BKP and obtain very good feasible points by applying FRA-SOR to test instances of the bounded knapsack problem.

In the bounded knapsack problem we have $m \in \mathbb{N}$ item types and denote the value and weight of item $j \in \{1, \dots, m\}$ by v_j and w_j , respectively. Further, there are at most $b_j \in \mathbb{N}$ units of item j available and the capacity of the knapsack is given by $c > 0$. By maximizing

Table 1 Relative optimality gap of FRA-SOR for different choices of U and m

$U \backslash m$	5	10	100	1000	10000
100	4.35e-01	2.27e-01	2.43e-02	2.45e-03	2.45e-04
1000	4.62e-01	2.50e-01	2.70e-02	2.72e-03	2.72e-04
10000	4.46e-01	2.40e-01	2.58e-02	2.60e-03	2.60e-04
100000	4.48e-01	2.41e-01	2.60e-02	2.62e-03	2.62e-04
1000000	4.47e-01	2.40e-01	2.59e-02	2.61e-03	2.61e-04

the total value of all items in the knapsack we arrive at the purely integer optimization problem

$$BKP : \quad \max_{y \in \mathbb{Z}^m} v^\top y \quad \text{s.t.} \quad w^\top y \leq c, \quad 0 \leq y \leq b.$$

In order to obtain hard test examples of the BKP we create so-called *strongly correlated instances* (cf. [23] for an analogous treatment in the context of 0-1 knapsack problems), that is, the weights w_j are uniformly distributed in the interval $[1, 10000]$ and we have $v_j = w_j + 1000$. Furthermore, $b_j, j \in \{1, \dots, m\}$, is uniformly distributed within the set $\{0, \dots, U\}$ for an integer upper bound $U \in \mathbb{N}$ and, in order to avoid trivial solutions, we set $c = \delta w^\top b$ for some $\delta \in (0, 1)$.

Note that the integer grid fineness of the BKP is controlled by the randomly chosen data b and w as well as $\delta \in (0, 1)$. The *expected* value of b is $(U/2)e$, where e denotes the all ones vector. At least for a fixed vector of weights w the expected value of c then is $\delta(U/2)w^\top e$. For the expected test instances the parameter U thus plays the role of the parameter t from the problem ILP_t in Section 4 and controls the grid fineness.

According to (2) the inner parallel set of $D = \{y \in \mathbb{R}^m \mid w^\top y \leq c, \quad 0 \leq y \leq b\}$ is

$$T^- = D^- = \{y \in \mathbb{R}^m \mid w^\top y \leq c - \frac{1}{2}\|w\|_1, \quad -y \leq -\frac{1}{2}e, \quad y \leq b - \frac{1}{2}e\}.$$

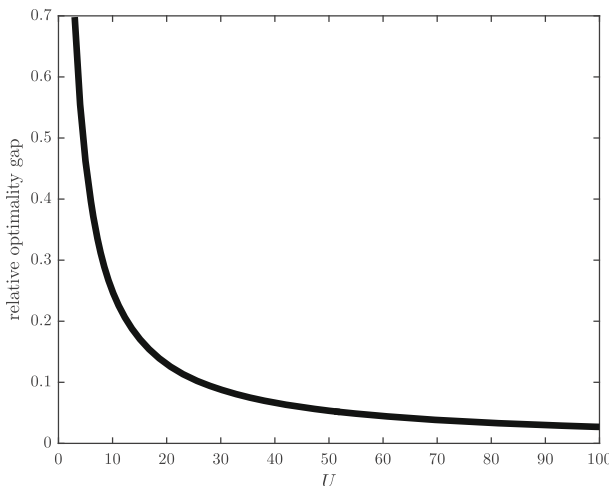


Fig. 1 Relative optimality gap for $m = 1000$ and different choices of U

Table 2 Computing time in seconds for FRA-SOR (left) and Gurobi (right) for different choices of U and m

$U \backslash m$	5	10	100	1000	10000	100000	1000000			
100	0.015	0.016	0.002	0.031	0.001	0.000	0.001	0.000	0.002	0.000
1000	0.004	0.016	0.004	0.016	0.002	0.000	0.002	0.016	0.002	0.016
10000	0.045	0.141	0.035	0.125	0.024	0.125	0.051	0.125	0.028	0.125
100000	0.580	2.063	0.654	2.109	0.546	2.047	0.574	2.031	0.548	1.984
1000000	7.546	35.484	7.845	13.766	8.014	33.219	8.154	33.078	8.254	33.594

Using the enlargement technique from Section 2.3 with $\sigma = (0, \frac{1}{2}e, \frac{1}{2}e)$ yields the enlarged inner parallel set

$$T_{\sigma}^{-} := \{y \in \mathbb{R}^m \mid w^T y \leq c - \frac{1}{2}\|w\|_1, 0 \leq y \leq b\}.$$

We see that T_{σ}^{-} is nonempty if and only if $c - \frac{1}{2}\|w\|_1 \geq 0$ holds. For our specific choice of c the latter is equivalent to

$$w^T(\delta b - \frac{1}{2}e) \geq 0.$$

In particular, T_{σ}^{-} may be empty for small values of δ and $b_j, j \in \{1, \dots, m\}$. In the remainder of this section we set $\delta = 1/3$ and use different values of $U \geq 5$. Then the expected values of the terms $\delta b_j - 1/2, j = 1, \dots, m$, exceed $1/3$, so that the enlarged inner parallel sets may be expected to be nonempty.

In fact, the enlarged inner parallel set T_{σ}^{-} turns out to be nonempty in all created test instances, so that all test problems are granular in the sense of Definition 1. In particular, no further enlargement of the inner parallel set is necessary.

FRA-SOR is implemented in MATLAB R2016b and the arising optimization problem is solved with Gurobi 7 [8], which we also use for a comparison. All tests are run on a personal computer with two cores à 2.3 GHz and 8 GB RAM.

In Table 1 we consider the relative optimality gap $(\widehat{v} - \check{v}_{\sigma}^s)/\widehat{v}$ of FRA-SOR applied to different instances of the BKP. The results seem to indicate that the optimality gap is independent of the problem size m . However, we see a strong dependency of the optimality gap on the upper bound U . This is caused by the fact that U controls the expected grid fineness, which plays a crucial role in the error bound obtained for FRA-SOR.

Note that the error bound given in Example 2 actually bounds the absolute optimality gap, and that this bound decreases linearly with finer grids. Thus, for the current setting this result predicts a hyperbolic decrease of the relative optimality gap with increasing values of U . This is confirmed by Fig. 1.

As mentioned above, solving the BKP to optimality is an NP-hard optimization problem. Instead, for nonempty enlarged inner parallel sets the main effort of our feasible rounding approach consists of solving a continuous linear optimization problem which can be done in polynomial time. This fact is demonstrated in Table 2 and Fig. 2 where we see that especially for the larger test instances FRA-SOR is able to find very good feasible points in reasonable time. Their relative optimality gaps (cf. Table 1) are of order 10^{-3} , that is, the additional time that Gurobi needs to identify a global optimal point only yields a marginal benefit.

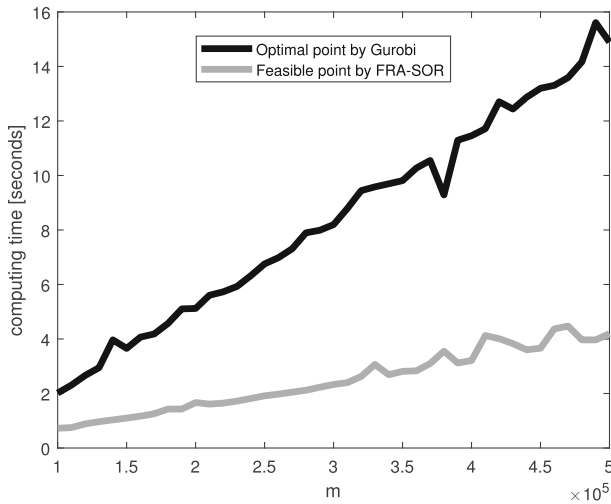


Fig. 2 Computing time in seconds for an optimal point by Gurobi and a feasible point by FRA-SOR for $U = 1000$ and different choices of m

6 Conclusions

This article assesses the quality of a point generated by a feasible rounding approach for mixed-integer nonlinear optimization problems. To this end, its optimality gap is estimated by a-posteriori as well as a-priori bounds, and the latter are shown to decrease at least linearly with increasing integer grid fineness.

The bounded knapsack problem illustrates our findings computationally. Detailed numerical results for the application of the feasible rounding approach to problems from the MIPLIB and MINLPLib libraries, which motivated the current research, are reported in [19, 20].

Acknowledgements Open Access funding provided by Projekt DEAL. The authors are grateful to the anonymous referees and the editor for their precise and substantial remarks on earlier versions of this manuscript.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Auslender, A., Crouzeix, J.-P.: Global regularity theorems. *Math. Oper. Res.* **13**, 243–253 (1988)
2. Auslender, A., Teboulle, M.: *Asymptotic Cones and Functions in Optimization and Variational Inequalities*. Springer, New York (2003)

3. Azé, D.: A survey on error bounds for lower semicontinuous functions. In: Penot, J.P. (ed.) Proceedings of 2003 MODE-SMAI Conference. ESAIM: Proceedings, vol. 13, pp. 1–17. EDP Sciences, Les Ulis (2003)
4. Danna, E., Rothberg, E., Le Pape, C.: Exploring relaxation induced neighborhoods to improve MIP solutions. *Math. Program.* **102**, 71–90 (2005)
5. Dantzig, G.B.: Discrete-variable extremum problems. *Oper. Res.* **5**, 266–277 (1957)
6. Deng, S.: Computable error bounds for convex inequality systems in reflexive Banach spaces. *SIAM J. Optim.* **7**, 274–279 (1997)
7. Güler, O., Hoffman, A.J., Rothblum, U.G.: Approximations to solutions to systems of linear inequalities. *SIAM J. Matrix Anal. Appl.* **16**, 688–696 (1995)
8. Gurobi Optimization, L.L.C.: Gurobi Optimizer Reference Manual. <http://www.gurobi.com> (2019)
9. Hoffman, A.J.: On approximate solutions of systems of linear inequalities. *J. Res. Natl. Bur. Stand.* **49**, 263–265 (1952)
10. Kellerer, H., Pferschy, U., Pisinger, D.: *Knapsack Problems*. Springer, Berlin (2004)
11. Klatté, D.: Eine Bemerkung zur parametrischen quadratischen Optimierung. *Sektion Mathematik der Humboldt-Universität zu Berlin. Seminarbericht Nr. 50*, 174–185 (1983)
12. Lewis, A.S., Pang, J.-S.: Error bounds for convex inequality systems. In: Crouzeix, J.-P., Martinez-Legaz, J.-E., Volle, M. (eds.) *Generalized Convexity, Generalized Monotonicity: Recent Results. Nonconvex Optimization and Its Applications*, vol. 27, pp. 75–110. Kluwer Academic Publishers (1996)
13. Li, G.: Global error bounds for piecewise convex polynomials. *Math. Program.* **137**, 37–64 (2013)
14. Li, G., Mordukhovich, B.S., Pham, T.S.: New fractional error bounds for polynomial systems with applications to Hölderian stability in optimization and spectral theory of tensors. *Math. Program.* **153**, 333–362 (2015)
15. Luo, X.-D., Luo, Z.-Q.: Extension of Hoffman’s error bound to polynomial systems. *SIAM J. Optim.* **4**, 383–392 (1994)
16. Luo, Z.-Q., Pang, J.-S.: Error bounds for analytic systems and their applications. *Math. Program.* **67**, 1–28 (1994)
17. Mangasarian, O.L.: A condition number for differentiable convex inequalities. *Math. Oper. Res.* **10**, 175–179 (1985)
18. Mangasarian, O.L., Shiau, T.-H.: Lipschitz continuity of solutions of linear inequalities, programs and complementarity problems. *SIAM J. Control Optim.* **25**, 583–595 (1987)
19. Neumann, C., Stein, O., Sudermann-Merx, N.: A feasible rounding approach for mixed-integer optimization problems. *Comput. Optim. Appl.* **72**, 309–337 (2019)
20. Neumann, C., Stein, O., Sudermann-Merx, N.: Granularity in nonlinear mixed-integer optimization. *J. Optim. Theory Appl.* **184**, 433–465 (2020)
21. Ngai, H.V., Kruger, A., Théra, M.: Stability of error bounds for semi-infinite convex constraint systems. *SIAM J. Optim.* **20**, 2080–2096 (2010)
22. Pang, J.-S.: Error bounds in mathematical programming. *Math. Program.* **79**, 299–332 (1997)
23. Pisinger, D.: Where are the hard knapsack problems? *Comput. Oper. Res.* **32**, 2271–2284 (2005)
24. Robinson, S.M.: An application of error bounds for convex programming in a linear space. *SIAM J. Control Optim.* **13**, 271–273 (1975)
25. Stein, O.: Error bounds for mixed integer linear optimization problems. *Math. Program.* **156**, 101–123 (2016)
26. Stein, O.: Error bounds for mixed integer nonlinear optimization problems. *Optim. Lett.* **10**, 1153–1168 (2016)