J Glob Optim (2012) 52:509–535
DOI 10.1007/s10898-011-9765-3

Exact penalty and error bounds in DC programming

Hoai An Le Thi · Tao Pham Dinh · Huynh Van Ngai

Received: 1 August 2011 / Accepted: 4 August 2011 / Published online: 24 September 2011

Abstract In the present paper, we are concerned with conditions ensuring the exact penalty for nonconvex programming. Firstly, we consider problems with concave objective and constraints. Secondly, we establish various results on error bounds for systems of DC inequalities and exact penalty, with/without error bounds, in DC programming. They permit to recast several class of difficult nonconvex programs into suitable DC programs to be tackled by the efficient DCA.

Keywords DC programming · Concave programming · DCA · Subdifferential · Exact penalty · Local and global error bounds · Reformulation

Mathematics Subject Classification (2000) 49J52 · 90C20 · 90C26 · 90C30 · 90C46

The paper is dedicated to the memory of our friend Reiner Horst.

H. A. Le Thi (✉)
Laboratory of Theoretical and Applied Computer Science—LITA EA 3097, UFR MIM,
University of Paul Verlaine—Metz, Ile du Saulcy, 57045 Metz, France
e-mail: lethi@univ-metz.fr

T. Pham Dinh
LMI-National Institute for Applied Sciences-Rouen, BP08-Avenue de l'Université,
76801 Saint-Etienne-du-Rouvray, France
e-mail: pham@insa-rouen.fr

H. V. Ngai
Department of Mathematics, University of Quynhon, 170 An Duong Vuong, Qui Nhon, Vietnam
e-mail: ngaivn@yahoo.com

1 Introduction

We consider the nonconvex program:

$$\min f(x) \quad s.t. \quad \begin{cases} g_i(x) \leq 0 \ \ (i = 1, \ldots, M); \\ g_i(x) = 0 \ \ (i = M+1, \ldots, N); \quad (NCP) \\ x \in C. \end{cases} \tag{1}$$

here C is a closed convex subset of $\mathbb{R}^n$ and f, g_i are DC functions (Difference of Convex functions) on C.

A way to make the study of (NCP) easier is to penalize difficult constraints, e.g., nonconvex constraints. Usually, the penalty problem of (NCP) is

$$\min \big(f(x) + \tau p(x)\big) : \ \ x \in C\big\} \quad (NCP_\tau) \tag{2}$$

where $\tau > 0$ and $p(x) := \sum_{i=1}^{M} g_i^+(x) + \sum_{i=M+1}^{N} |g_i(x)|$ with $g^+ := \max(g, 0)$.

We say that the exact penalty holds if there exists a nonnegative number $\tau_0 \geq 0$ such that for all $\tau > \tau_0$ the problems (NCP), (NCP_τ) have the same optimal value and the same optimal solution set. It is well known that such a property generally holds in convex programming ([29–34,36–38,41–43,47,48,52,54,55,67,69]). For nonconvex programming—to the best of our knowledge—there are few existing exact penalty results in the literature. In [51], Luo et al. have established a general exact penalty result for problems with the subanalytic data. In which, $p(x)^\gamma$, $\gamma \in (0, 1)$, is instead of $p(x)$ in the objective function of the problem (NCP_τ). However, from the computational viewpoint, that exact penalty result might be disadvantageous since the exponential parameter γ is unknown and furthermore, the penalty functions have not special structure to be enhanced.

Let $\Gamma_0(\mathbb{R}^n)$ denote the convex cone of all lower semicontinuous proper convex functions on $\mathbb{R}^n$. The vector space of DC functions, $DC(\mathbb{R}^n) = \Gamma_0(\mathbb{R}^n) - \Gamma_0(\mathbb{R}^n)$, is quite large to contain almost real life objective functions and is closed under all the operations usually considered in Optimization. Throughout the paper, $\|.\|$ denotes the Euclidean norm on $\mathbb{R}^n$, unless otherwise specified.

Consider the standard DC program

$$\alpha := \inf\{f(x) := g(x) - h(x) \ : \ x \in \mathbb{R}^n\} \quad (P_{dc})$$

with $g, h \in \Gamma_0(\mathbb{R}^n)$.

Remark that the closed convex constraint set C is incorporated in the first convex DC component g with the help of its indicator function χ_C ($\chi_C(x) := 0$ if $x \in C$, $+\infty$ otherwise). Recall that in DC programming we have $+\infty - (+\infty) = +\infty$ and that the finite optimal value α implies $dom g := \{x \in \mathbb{R}^n : g(x) < +\infty\} \subset dom h$. The function f is called a DC function (resp. polyhedral DC function) on $\mathbb{R}^n$ (resp. if either g or h are polyhedral convex). Polyhedral DC programs, being DC programs with polyhedral DC objective functions, are at the heart of DC programming.

DC programming and DCA (DC Algorithms) have been introduced by Pham Dinh Tao in 1985 and extensively developed by Le Thi Hoai An and Pham Dinh Tao since 1994. DC programming plays a key role in nonconvex programming because almost nonconvex programs encountered in practice are DC programs. Based on local optimality conditions and DC duality, DCA is one of a few efficient algorithms for nonconvex programs, especially for smooth/nonsmooth large-scale ones. Due to its local character it cannot guarantee the globality of computed solutions for general DC programs. However, we observe that, with a suitable starting point, it converges quite often to global solutions. In practice, DCA was successfully applied to a lot of different and various nonconvex programs proved to be more robust and more efficient than related standard methods, especially in the large-scale setting. On the other hand, it is worth noting that, with appropriate DC decompositions, DCA permits to find again standard optimization algorithms for convex and nonconvex programming ([8], [1–22,63,64] and references therein). In global approaches to nonconvex programming, DC programming have been investigated by Reiner Horst and Hoang Tuy [26], Reiner Horst, Panos Pardalos and Nguyen Van Thoai [23], Reiner Horst [24], Reiner Horst et al. [25], Reiner Horst and Nguyen Van Thoai [27]. In practice, to check globality of solutions computed by DCA, we proposed to combine DCA with global optimization techniques like Branch-and-Bound scheme (B&B) using respectively DC/SDP relaxation techniques [3,4,6,13,21]. The combined algorithm allows checking globality of solutions computed by DCA and restarting it if necessary, and, consequently, accelerates the B&B approach. Exact penalty then is crucial insofar as solutions computed by DCA applied to equivalent penalized problems provide upper bounds for the optimal value of the original problem.

The goal of the present paper is to establish exact penalty results with/without error bounds of nonconvex constraint sets. These results are based on the concavity of objective/constraint functions and local error bounds. They permit us to recast several classes of problems (NCP) into the more suitable DC programming framework. More precisely, under appropriate conditions, (NCP) is equivalent to a DC program.

The paper is organized as follows. The next section presents new general exact penalty results (without available error bounds for nonconvex feasible sets) for nonconvex programs with concave objective and constraints functions over bounded polyhedral convex sets. These results encompass that of An-Tao-Muu in [2] as a special case and allow the statement of equivalent penalized polyhedral DC programs. In Sect. 3, we investigate new error bound results for DC inequality systems of the form:

$$S = \left\{x \in C : \quad g(x) - h(x) \leq 0\right\},$$

where C is a closed convex set in $\mathbb{R}^n$ and g, h are convex functions on C, by using, local error bounds, compactness, polyhedrality of C on the one hand, and differentiability of h, polyhedrality of g on the other hand. Finally, consequent exact penalty results in DC programming are reported in the last section.

2 Exact penalty in concave programming

We first investigate exact penalty for nonconvex programs having concave objective functions and bounded polyhedral convex constraint sets with additional concave constraint functions. These results are not derived from error bounds of feasible solution sets because the question concerning their error bounds is still open. We get back to it in Sect. 3.1.

Let K be a nonempty bounded polyhedral convex subset of $\mathbb{R}^n$ and let f, g be finite functions on K. Recall that the function f is said to be continuous relative to K, if its restriction to K is a continuous function. The notions of relative upper semicontinuity and relative lower semicontinuity are defined similarly. In this section, we are concerned with the following nonconvex programs:

$$\alpha := \min\left\{f(x) : x \in K, \; g(x) = 0\right\} \quad (P) \tag{3}$$

$$\min\left\{f(x) : x \in K, \; g(x) \leq 0\right\} \quad (P') \tag{4}$$

In [2], it has been established an exact penalty result for the problem (P) under the hypothesis that g is nonnegative on K. We consider (P) here without this nonnegativity assumption. Problem (P) could be rewritten equivalently

$$\alpha := \min\left\{f(x) \;:\; x \in C,\; g(x) \le 0\right\} \quad (P) \tag{5}$$

where C is the convex set defined by

$$C := \{x \in K, g(x) \ge 0\} \tag{6}$$

on which g is nonnegative. It is clear that C is a polyhedral convex set iff $-g$ is a polyhedral convex function on K. In such a case, exact penalty relative to the nonconvex constraint $g(x) \le 0$ holds, due to [2].

According to DC programming and DCA approaches, to handle nonconvex constraints, we make use of a penalty function to bring these nonconvex constraints into the objective function. For this goal, the penalty problem (P_τ) of (P) is defined by, for $\tau > 0$,

$$\alpha(\tau) := \min\left\{f(x) + \tau g(x) \;:\; x \in K,\; g(x) \ge 0\right\} \quad (P_\tau). \tag{7}$$

The following result constitutes a substantial extension of the key theorem in ([2]-Theorem 2). Let us denote by $V(K)$ the vertex set of K and by $\mathcal{P}$ and $\mathcal{P}_\tau$ the optimal solution sets of $(P)-$ (5) and $(P_\tau)-$ (7), respectively.

Theorem 1 *Let K be a nonempty bounded polyhedral convex set in $\mathbb{R}^n$ and let f, g be finite concave functions continuous relative to K. Suppose that the feasible set of (P) is nonempty. Then there exists $\tau_0 \ge 0$ such that for all $\tau > \tau_0$, the problems $(P)-$ (5) and $(P_\tau)-$ (7) have the same optimal value and the same optimal solution set. Furthermore, we can take $\tau_0 = \frac{f(x_0)-\alpha(0)}{m}$ with $m = \min\{g(x) \;:\; x \in V(K), g(x) > 0\}$ and any $x_0 \in K$, $g(x_0) = 0$.*

Proof First, note that the convention $\min_\emptyset g(x) = +\infty$ will be used in the sequel. For $\tau \ge 0$, since the functions f, $f + \tau g$, being continuous concave functions on K, then $\mathcal{P}$ and $\mathcal{P}_\tau$ are nonempty. Set $C = \{x \in K : g(x) \ge 0\}$. Denote by $E(C)$ the set of all extremal points of C. Then $C = coE(C)$. It is obvious that

$$E(C) \subset \{x \in V(K) :\; g(x) > 0\} \cup \{x \in K \;:\; g(x) = 0\}.$$

We first prove that there exists $\tau_0 \ge 0$ such that $\mathcal{P} \subset \mathcal{P}_\tau$ and $\alpha(\tau) = \alpha$ for all $\tau > \tau_0$. Indeed, if $\{x \in V(K) :\; g(x) > 0\} = \emptyset$ then $E(C) \subset \{x \in K \;:\; g(x) = 0\}$. Since $f + \tau g$ is concave, then there is $z \in E(C) \subset \{x \in K,\; g(x) = 0\}$ such that $z \in \mathcal{P}_\tau$. Hence, for any $x \in \mathcal{P}$, we have

$$f(x) \le f(z) \le f(y) + \tau g(y) \;\text{ for all }\; y \in C.$$

That is, $x \in \mathcal{P}_\tau$. It follows that $\mathcal{P} \subset \mathcal{P}_\tau$ and $\alpha(\tau) = \alpha$.

We now consider the case $\{x \in V(K),\; g(x) > 0\} \ne \emptyset$. Set $\tau_0 = \frac{f(x_0)-\alpha(0)}{m}$ and $x_0 \in K$ with $g(x_0) = 0$. Let $\tau > \tau_0$. By again the concavity of $f + \tau g$, there exists $z \in \mathcal{P}_\tau \cap E(C)$. It follows that $g(z) = 0$. In fact, if $g(z) > 0$ then $z \in V(K)$. Therefore, $g(z) \ge m > 0$ and

$$f(z) + \tau g(z) > f(z) + \tau_0 g(z) \ge f(z) + f(x_0) - \alpha(0) \ge f(x_0),$$

a contradiction. Hence $z \in \mathcal{P} \cap \mathcal{P}_\tau$ and $\alpha(\tau) = \alpha$. By using the same argument as for the above case, we also obtain $\mathcal{P} \subset \mathcal{P}_\tau$ for all $\tau > \tau_0$.

Next, let $\tau > \tau_0$. For any $z \in \mathcal{P}_\tau$, to see that $z \in \mathcal{P}$, we assume for contradiction that $g(z) > 0$. Then by taking some τ' with $\tau_0 < \tau' < \tau$, one has

$$\alpha(\tau) = f(z) + \tau g(z) > f(z) + \tau' g(z) \geq \alpha(\tau') = \alpha.$$

This contradicts $\alpha(\tau) = \alpha$. Thus, $\mathcal{P} = \mathcal{P}_\tau$ and the proof is complete. □

Let us now consider the problem (P'):

$$\min\left\{f(x) \,:\, x \in K,\ g(x) \leq 0\right\} \qquad (P'). \tag{8}$$

Obviously, (P') is equivalent to the following one:

$$\min\left\{f(x) \,:\, (x,t) \in K \times [0,\beta],\ g(x) + t = 0\right\} \quad (\overline{P'}) \tag{9}$$

where $\beta \geq \max\{-g(x) :\ x \in K\}$ and such the equivalence is in the following usual sense: *If x is a solution of (P')then $(x, -g(x))$ is a solution of $(\overline{P'})$.Conversely, if (x,t) is a solution of $(\overline{P'})$ then x is a solution of (P').*

In virtue of Theorem 1, there is $\tau_0 \geq 0$ such that for all $\tau > \tau_0$, the problem $(\overline{P'})$ is equivalent to the following one:

$$\min\left\{f(x) + \tau(t + g(x)) \,:\, (x,t) \in K \times [0,\beta],\ g(x) + t \geq 0\right\} (\overline{P'_\tau}). \tag{10}$$

Thus we obtain the following corollary.

Corollary 1 *Under the assumptions of Theorem 1, the problems (P')−(8) and $(\overline{P'_\tau})$−(10) are equivalent in the sense given in Theorem 1.*

It is obvious that $(\overline{P'_\tau})$−(10) is a DC program. As shown by the corollary, it is worth noting that if f, g are finite concave functions on a nonempty bounded polyhedral convex set K, then Problem (P') − (8) is equivalent to $(\overline{P'_\tau})$−(10).

By applying Corollary 1, we list below the nonconvex programs (with DC objective functions and DC constraint functions) frequently encountered in practice that we can advantageously recast into more suitable equivalent DC programs (in the sense given Theorem 1):

(a) *Polyhedral convex objective function and concave constraint function*

Consider the problem with a nonempty feasible set:

$$\min\left\{f(x) \,:\, x \in K,\ g(x) \leq 0\right\}, \tag{11}$$

where K is a polyhedral convex set and $f \in \Gamma_0(\mathbb{R}^n)$ is a polyhedral convex function, i.e., it is a sum of a pointwise supremum of a finite collection of affine functions and an indicator funtion of a nonempty polyhedral convex set D

$$\begin{aligned} f(x) :&= \max\{\langle x, y_i\rangle - \alpha_i : i = 1, \ldots, r\} + \chi_D(x) \\ &= \overline{f}(x) + \chi_D(x), \end{aligned} \tag{12}$$

such that $D \cap K$ is bounded and g is a finite concave function continuous relative to $D \cap K$. Obviously, this problem is equivalent to

$$\min\left\{t \,:\, \overline{f}(x) \leq t,\ (x,t) \in [D \cap K] \times [a,b],\ g(x) \leq 0\right\}, \tag{13}$$

where $a \leq \min\{f(x) : x \in [D \cap K]\}$ and $b \geq \max\{f(x) : x \in [D \cap K]\}$. According to Corollary 1, if the feasible set of (11) is nonempty there is τ_0 such that for all $\tau > \tau_0$, Problem (13) is equivalent to following DC program (with $\beta \geq \max_{x \in [D \cap K]}(-g(x))$):

$$\min\left\{t + \tau(s + g(x)) \,:\, \overline{f}(x) \leq t,\ (x,t,s) \in [D \cap K] \times [a,b] \times [0,\beta],\ g(x) + s \geq 0\right\}, \tag{14}$$

that can be reduced to

$$\min\left\{\overline{f}(x)+\tau(s+g(x))\ :\ (x,s)\in[D\cap K]\times[0,\beta],\ g(x)+s\geq 0\right\} \tag{15}$$

(b) *Concave objective function and DC (whose first DC component is polyhedral convex function) constraint function*

Let f be a finite concave function on a polyhedral convex set K. Let g be a polyhedral convex function defined as in (12) with $domg := \{x \in \mathbb{R}^n : g(x) < +\infty\} = D$ such that $D \cap K$ is bounded nonempty, and h be a finite convex function on $D \cap K$. Let us now consider the following problem with a polyhedral DC constraint:

$$\min\left\{f(x)\ :\ x\in K,\ g(x)-h(x)\leq 0\right\}. \tag{16}$$

This problem is equivalently transformed into

$$\min\left\{f(x)\ :\ (x,t)\in[D\cap K]\times[a,b],\ \overline{g}(x)-t\leq 0,\ t-h(x)\leq 0\right\}, \tag{17}$$

where $a \leq \min_{x\in[D\cap K]} g(x)$, $b \geq \max_{x\in[D\cap K]} g(x)$. In virtue of Corollary 1, if the feasible set of (16) is nonempty then there is $\tau_0 \geq 0$ such that for all $\tau > \tau_0$ Problem (17) is equivalent to the DC program ($\beta \geq \max_{x\in[D\cap K]} h(x) - a$):

$$\begin{aligned}\min\Big\{f(x)+\tau(s+t-h(x))\ \ :\ \ &(x,t,s)\in[D\cap K]\times[a,b]\times[0,\beta],\\ &\overline{g}(x)-t\leq 0,\ s+t-h(x)\geq 0\Big\}.\end{aligned} \tag{18}$$

(c) *DC objective function and DC constraint function (whose first DC component is polyhedral convex function)*

Let K be a nonempty polyhedral convex set in $\mathbb{R}^n$. Consider the nonconvex program with a polyhedral DC objective function $f = g_0 - h_0$ and DC constraint function:

$$\min\{g_0(x)-h_0(x) : x\in K,\ g_1(x)-h_1(x)\leq 0\} \tag{19}$$

where g_0, g_1 are polyhedral convex functions on $\mathbb{R}^n$, i.e.,

$$\begin{aligned}g_0(x) :&= \max\{\langle x, y_i\rangle-\alpha_i : i=1,\ldots,q\}+\chi_D(x)\\ &=\overline{g_0}(x)+\chi_D(x),\end{aligned} \tag{20}$$

$$\begin{aligned}g_1(x) :&= \max\{\langle x, z_i\rangle-\alpha_i : i=1,\ldots,r\}+\chi_R(x)\\ &=\overline{g_1}(x)+\chi_R(x),\end{aligned}$$

with D, R being nonempty polyhedral convex sets and h_0, h_1 are a finite convex function continuous relative to $D \cap K \cap R$. Assume that the polyhedral convex set $D \cap K \cap R$ is bounded nonempty.

Problem (19) is equivalent to

$$\begin{aligned}\min\{s-h_0(x)\ :\ &(x,s,t)\in D\cap K\cap R\times[a_0,b_0]\times[a_1,b_1],\ \overline{g_0}(x)\leq s,\ \overline{g_1}(x)\leq t,\\ &t-h_1(x)\leq 0\}\end{aligned} \tag{21}$$

where a_0, b_0, a_1, b_1 are constant numbers verifying

$$a_0 \leq \min\{\overline{g_0}(x) : x\in D\cap K\cap R\}\leq \max\{\overline{g_0}(x) : x\in D\cap K\cap R\}\leq b_0$$
$$a_1 \leq \min\{\overline{g_1}(x) : x\in D\cap K\cap R\}\leq \max\{\overline{g_1}(x) : x\in D\cap K\cap R\}\leq b_1$$

Hence, it follows from Corollary 1 that if the feasible set of Problem (21) is nonempty then there is $\tau_0 \geq 0$ such that for all $\tau > \tau_0$ this problem is equivalent to

$$\min\{s-h_0(x)+\tau[u+t-h_1(x)] : (x,s,t,u) \in D\cap K\cap R\times[a_0,b_0]\times[a_1,b_1]\times[0,c], \\ \overline{g_0}(x)\le s,\ \overline{g_1}(x)\le t,\ u+t-h_1(x)\ge 0\} \tag{22}$$

with c being a constant number satisfying $c \ge \max\{h_1(x) : x \in D\cap K\cap R\} - a_1$.

Finally, Problem (19) is equivalent to the DC program

$$\min\{f(x,u,t) = \overline{g_0}(x) - h_0(x) + \tau[u+t-h_1(x)] : \\ (x,t,u) \in D\cap K\cap R\times[a_1,b_1]\times[0,c],\ \overline{g_1}(x)\le t,\ u+t-h_1(x)\ge 0\} \tag{23}$$

3 Error bounds for DC inequality systems

Throughout the paper $\|.\|$ denotes the Euclidean norm on $\mathbb{R}^n$ unless otherwise specified, and $B(c,r) := \{x \in \mathbb{R}^n : \|x-c\| \le r\}$. Let C be a nonempty closed subset in $\mathbb{R}^n$ and let h be a finite function on C. We consider the solution set of the following inequality system

$$S = \{x \in C : h(x) \le 0\} \tag{24}$$

Recall that an error bound of S is an inequality of the form

$$d(x,S) \le \tau h^+(x) \quad \text{for all } x \in C, \tag{25}$$

where $d(x,S) = \inf_{z\in S}\|x-z\|$, and τ is a positive number. For $x_0 \in S$, if the inequality in (25) holds for all x in a neighborhood of x_0, then we say that S has an error bound around x_0. Instead of (25), an inequality of the form

$$d(x,S) \le \tau h^+(x)^\gamma \text{ for all } x \in C, \tag{26}$$

with $\tau > 0$ and $0 < \gamma < 1$ will be called an error bound for (24) with the exponent γ. Note that if an error bound of S holds for a norm it holds for every norm because all norms on $\mathbb{R}^n$ are equivalent. Error bounds have important applications in many areas of mathematical programming, e.g., in sensitivity analysis, in convergence analysis of some algorithms, and in exact penalty, etc. The reader is referred to the suvey papers [28,45,66] and the references therein for the summary of the theory and applications of error bounds. For convex inequality systems, many results on error bounds have been established in the literature (see, e.g., [46,56–62] and references therein) and several conditions ensuring the error bounds have been investigated in different contexts. However, from the computational point of view these existing conditions are hard to check in practical problems.

Recently, in [50,51], Luo et al. by adapting the very deep result of Lojasiewicz [46], have established the error bound result with exponent for subanalytic systems and the application to establish the exact penaltization results for optimizations problems with equilibrium constraints. In [49,50], the authors also have sharply estimated the exponential parameter for nonconvex quadratic systems under appropriate conditions.

In this section, we derive several new results on error bounds for the inequality system (24) when C is a suitable closed convex set and h is a concave function on C. These systems play an important role in global optimization.

In general, a system which has local error bounds at all points in its solution set does not necessarily attain the global error bound. However, under the compactness hypothesis, such an implication holds as will show the following proposition that will be needed thereafter.

Theorem 2 *Let C be a nonempty compact set in $\mathbb{R}^n$. Let h be a finite function lower semi-continuous relative to C and $S := \{x \in C : h(x) \le 0\} \ne \emptyset$. Suppose that for each $z \in S$, there exist $\tau(z)$, $\gamma(z) > 0$ and $\epsilon(z) > 0$ such that*

$$d(x,S) \le \tau(z) h^+(x)^{\gamma(z)} \quad \textit{for all } x \in B(z,\epsilon(z)) \cap C. \tag{27}$$

Then there exist $\tau,\ \gamma > 0$ *such that*

$$d(x,S) \le \tau h^+(x)^{\gamma} \quad \textit{for all } x \in C. \tag{28}$$

Proof Since h is lower semicontinuous relative to C and C is compact, then S is a compact set too. For each $z \in S$, let $0 < \epsilon(z) < 1$, $\tau(z)$, $\gamma(z)$ such that (27) is satisfied. By the compactness, there exist $z_1, \ldots, z_m$ in S such that

$$S \subset \bigcup_{i=1}^{m} B(z_i, \epsilon(z_i)/2).$$

Set $\epsilon = \min\{\epsilon(z_i) : i = 1, \ldots, m\}$; $\gamma = \min\{\gamma(z_i) : i = 1, \ldots, m\}$ and $\tau = \max\{\tau(z_i) : i = 1, \ldots, m\}$. Let $x \in C$ such that $d(x,S) \le \epsilon/2$. Then we can find $z \in S$ such that $\|x - z\| \le \epsilon/2$. Therefore, there is an index $i \in \{1, \ldots, m\}$ such that $z \in B(z_i, \epsilon(z_i)/2)$. It follows that

$$\|x - z_i\| \le \|x - z\| + \|z - z_i\| \le \epsilon/2 + \epsilon(z_i)/2 \le \epsilon(z_i).$$

Hence, we obtain

$$d(x,S) \le \tau(z_i) h^+(x)^{\gamma(z_i)} \le \tau h^+(x)^{\gamma}$$

if $h^+(x) < 1$; otherwise, one has $d(x,S) \le h^+(x)^{\gamma}$. Now let $x \in C$ with $d(x,S) > \epsilon/2$. We say that there is a $\eta > 0$ such that

$$h(x) \ge \eta \ \text{ for all } x \in C,\ d(x,S) > \epsilon/2.$$

Indeed, if this is not the case, one selects a sequence $\{x_k\} \subset C$ such that $h(x_k) \le \eta_k$ and $d(x_k, S) > \epsilon/2$ for all k, where $\{\eta_k\}$ is a sequence of positive numbers with $\lim_{k\to+\infty} \eta_k = 0$. By the compactness, without loss of generality, assume that $\{x_k\}$ converges to some $x^* \in C$. Then $h(x^*) \le 0$ by lower semicontinuity of h. That is, $x^* \in S$ and a contradiction that $d(x_k, S) \le \|x_k - x^*\| \to 0$. Since C is bounded, then $d(\cdot, S)$ is bounded above on C, say, by some $r > 0$. Thus for all $x \in C$ with $d(x,S) > \epsilon/2$, one has

$$d(x,S) \le r \le \frac{r}{\eta^{\gamma}} h(x)^{\gamma}.$$

By taking $\tau^* = \max\{\tau, 1, r/\eta^{\gamma}\}$, we obtain

$$d(x,S) \le \tau^* h^+(x)^{\gamma} \quad \text{for all } x \in C.$$

The proof is complete. □

Remark 1 We see from the proof above that if (27) is satisfied for a common exponent γ which does not depend on z, then (28) holds also with this exponential parameter.

3.1 Error bounds for DC inequality systems

It is well known (see Sect. 4) that error bounds of feasible solution sets of nonconvex programs provide, in an elegant and deep way, exact penalty for those problems with Lipschitz objective functions. This subsection is concerned with error bounds for DC inequality systems.

Let $f : \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$ be a lower semicontinuous function on $\mathbb{R}^n$. The lower directional derivative (or contingent derivative or Hadamard derivative) of f at $x \in domf := \{x \in \mathbb{R}^n : f(x) < +\infty\}$ is given by

$$f'(x;u) = \liminf_{(t,v)\to(0_+,u)} \frac{f(x+tv)-f(x)}{t}.$$

If f is differentiable at x with the derivative $\nabla f(x)$ then $f'(x;u) = \langle \nabla f(x), u\rangle$. When f is convex or concave then the lower directional derivative coincides with the usual directional derivative. The *Clarke* derivative of a locally Lipschitz f at x is given by the following formula.

$$f^0(x;u) := \limsup_{(t,v)\to(0_+,x)} \frac{f(v+tu)-f(v)}{t}.$$

The *Clarke subdifferential* of f at x is defined by

$$\partial^0 f(x) := \{y \in \mathbb{R}^n : \langle y, u\rangle \le f^0(x;u) \ \forall u \in \mathbb{R}^n\}.$$

It is well known that when f is a convex function and continuous at x then *Clarke subdifferential* of f at x coincides with the subdifferential of f at x in the sense of convex analysis—which is denoted by $\partial f(x)$—and $\partial^0(-f)(x) = -\partial f(x)$. More generally, by observing from the sum rule for the Clarke subdifferential, if g, h are convex and continuous at x then one has

$$\partial^0(g-h)(x) \subset \partial g(x) - \partial h(x).$$

We present first a general sufficient condition for the error bound in terms of the lower directional derivative. This condition can be derived from the existing sufficient conditions using generalized derivatives in the literature (see [28,44,57,65]). However, for the reader's convenience, we give here a simple proof in the context of finite dimension.

Proposition 1 *Let C be a nonempty closed convex set in $\mathbb{R}^n$, let h be a finite function lower semicontinuous relative to C and $S := \{x \in C : h(x) \le 0\} \neq \emptyset$.*

(i) *Given $x_0 \in S$. If there exists $\mu > 0$, $\epsilon > 0$ such that for each $y \in [B(x_0,\epsilon)\cap C] \setminus S$ there exists $z \in C, z \neq y$ such that $h'(y; z-y) < -\mu\|z-y\|$ then one has*

$$d(x,S) \le \frac{1}{\mu} h^+(x) \ \textit{for all} \ x \in B(x_0,\epsilon/2)\cap C.$$

(ii) *In addition, if C is bounded, h is continuous relative to C and for all $x \in S$ with $h(x) = 0$, the set*

$$\{(\mu,\epsilon) > (0,0) : \forall y \in [B(x,\epsilon)\cap C]\setminus S \quad \exists z_y \in C \textit{ such that } h'(y; z_y - y) < -\mu\|z_y - y\|\}$$

is nonempty, then there is $\tau > 0$ such that

$$d(x,S) \le \tau h^+(x) \ \textit{for all} \ x \in C.$$

Proof (i) Let $\mu > 0, \epsilon > 0$ as in (i). Let $\overline{x} \in B(x_0,\epsilon/2)\cap C$ with $h(\overline{x}) > 0$. For every $\xi > 1$, consider the following function defined by $\varphi(x) = h^+(x) + \xi \frac{h(\overline{x})}{d(\overline{x},S)}\|x-\overline{x}\|,\ x \in C$. Then φ

is lower semicontinuous relative to C, bounded below and coercive (i.e., $\lim_{\|x\|\to\infty} \varphi(x) = +\infty$) on C. Hence, the minimum of φ relative to C is attained at $y \in C$

$$h^+(y) + \xi \frac{h(\overline{x})}{d(\overline{x}, S)} \|y - \overline{x}\| \le h^+(x) + \xi \frac{h(\overline{x})}{d(\overline{x}, S)} \|x - \overline{x}\|, \ \forall x \in C.$$

It implies that

$$h^+(y) \le h^+(x) + \xi \frac{h(\overline{x})}{d(\overline{x}, S)} \|x - y\|, \ \forall x \in C. \tag{29}$$

By setting $x = \overline{x}$ in the first inequality, we get

$$\xi \frac{h(\overline{x})}{d(\overline{x}, S)} \|y - \overline{x}\| \le h^+(\overline{x}).$$

Thus $\|y - \overline{x}\| < d(\overline{x}, S) \le \epsilon/2$. It follows that $y \in [B(x_0, \epsilon) \cap C] \setminus S$. On the other hand, from the relation (29), we derive that

$$\frac{-h^+(y + t(z - y))) + h^+(y)}{t\|z - y\|} \le \xi \frac{h(\overline{x})}{d(\overline{x}, S)}, \ \forall t \in (0, 1), \ z \in C.$$

According to the assumption, there exists $z_y \in C$ such that $h'(y; z_y - y) < -\mu\|z_y - y\|$. By setting $z = z_y$ and letting $t \to 0_+$, and since $\xi > 1$ is arbitrary, we obtain the desired inequality

$$\frac{h(\overline{x})}{d(\overline{x}, S)} \ge \frac{-h'(y; z_y - y)}{\|z_y - y\|} > \mu > 0$$

and complete the proof of (i).

For the assertion (ii), by virtue of Theorem 2, it suffices to show that for each $x_0 \in S$, there exist $\tau, \epsilon > 0$ such that

$$d(x, S) \le \tau h^+(x) \ \text{ for all } x \in B(x_0, \epsilon) \cap C.$$

Indeed, let $x_0 \in S$. If $h(x_0) < 0$ then by the continuity of h, there is $\epsilon > 0$ such that $h(x) < 0$ for all $x \in B(x_0, \epsilon) \cap C$. Therefore, $d(x, S) = h^+(x) = 0$ for all $x \in B(x_0, \epsilon) \cap C$.

The case $h(x_0) = 0$ follows directly from part (i). □

Remark 2 Proposition 1 remains valid if we replace the conditions (i) (resp. (ii) by the following conditions (iii) (resp. (iv)):

(iii) Let $x_0 \in S$. If there exits $z \in C$ such that the function $x \to h'(x; z - x)$ is upper semicontinuous at x_0 and $h'(x_0; z - x_0) < 0$.

(iv) In addition, if C is bounded, h is continuous relative to C and for all $x_0 \in S$ with $h(x_0) = 0$,there exits $z \in C$ such that the function $x \to h'(x; z - x)$ is upper semicontinuous at x_0 and $h'(x_0; z - x_0) < 0$.

Error bound for the relative complementary of a Euclidean ball with respect to another one

Let us present now a nice application of Proposition 1 in pointing out an error bound for this nonconvex set, which is of great interest in practice.

Let $a_1, a_2 \in \mathbb{R}^n$ and $r_1, r_2 > 0$ be given. Denote by S the solution set of the system:

$$S := \{x \in \mathbb{R}^n : \ \|x - a_1\|^2 - r_1^2 \le 0, \|x - a_2\|^2 - r_2^2 \ge 0\}. \tag{30}$$

Set $C = B(a_1, r_1)$ and $h(x) = -\|x - a_2\|^2 + r_2^2$.

Proposition 2 *If S is nonempty, then there is $\tau > 0$ such that*

$$d(x, S) \leq \tau(r_2^2 - \|x - a_2\|^2)^+ \text{ for all } x \in C. \tag{31}$$

Proof We shall show Proposition 2 by using the condition (iv) in Remark 2. Let $x_0 \in S$ with $h(x_0) = 0$. We shall distinguish different cases:

(1) If $\|x_0 - a_1\| < r_1$ then, for $t > 0$ is sufficiently small, one has $z := x_0 + t(x_0 - b) \in C$ and $h'(x_0; z - x_0) = -2t\|x_0 - b\|^2 < 0$.
(2) If $\|x_0 - a_1\| = r_1$ and $a_1 \neq a_2$ then there exists $u \in \mathbb{R}^n$ with $\|u\| = 1$ such that $\langle u, x_0 - a_1\rangle < 0$ and $\langle u, x_0 - a_2\rangle > 0$. Hence for $t > 0$ sufficiently small, we also have

$$z := x_0 + tu \in C \text{ and } h'(x_0; z - x_0) = -2t\langle x_0 - b, u\rangle < 0.$$

(3) If $\|x_0 - a_1\| = r_1$ and $a_1 = a_2 = a$ then we have $r_1 \geq r_2$.Two subcases should be considered:

(3.1) If $r_1 > r_2$ then $\|x_0 - a\| < r_1$ for all $x_0 \in S$ with $h(x_0) = 0$ and we return to the case 1).
(3.2) If $r_1 = r_2 = r$ then there no $z \in C$ such that $h'(x_0; z - x_0) < 0$, i.e., the condition (iv) in Remark 2 doesn't hold. However we can state two error bounds for S in this subcase as follows:

The first error bound is given by

$$d(x, S) \leq (r - \|x - a\|) \text{ for all } x \in C, \tag{32}$$

which is a nonsmooth error bound, while the second smooth error bound

$$d(x, S) \leq \frac{1}{r}(r^2 - \|x - a\|^2) \text{ for all } x \in C \tag{33}$$

comes from the fact that

$$h(x) = -\|x - a\|^2 + r^2 = (r - \|x - a\|)(r + \|x - a\|) \geq rd(x, S) \text{ for all } x \in C \setminus S$$

□

Remark 3 Proposition 1 establishes existence of error bounds without providing the value of the parameter τ, which is important for using exact penalty techniques in a computational point of view. The error bounds (32) and (33) are then very interesting in practice. Note that if $\{x \in \mathbb{R}^n : \|x - a_2\| - r_2 \leq 0\}$ is contained in C,then S has the following error bounds

$$d(x, S) \leq (r - \|x - a_2\|)^+ \text{ for all } x \in C,$$

and

$$d(x, S) \leq \frac{1}{r}(r^2 - \|x - a_2\|^2)^+ \text{ for all } x \in C$$

Next, we give a condition ensuring the error bound for DC inequality systems which is a slight extension of the well known Slater one for convex systems (when $h = 0$). The systems of that type play an important role in DC programming. Let C be a nonempty closed convex set in $\mathbb{R}^n$ and let g, h be finite convex functions on C. Consider the following set:

$$S = \{x \in C : g(x) - h(x) \leq 0\}.$$

Theorem 3 *Let $C \subset \mathbb{R}^n$ be a nonempty compact convex set, g be a finite convex function continuous relative to C and h be a differentiable convex function on C. If for each $x_0 \in S$ with $g(x_0) - h(x_0) = 0$ the set*

$$\{z \in C : g(z) - \langle \nabla h(x_0), z - x_0 \rangle - h(x_0) < 0\} \tag{34}$$

is nonempty, then there exists $\tau > 0$ such that

$$d(x, S) \leq \tau[g(x) - h(x)]^+ \text{ for all } x \in C.$$

Proof Let $x_0 \in S$ and let $z \in C$ satisfy (34). By the assumptions, the function $z \to g(z) - \langle \nabla h(y), z - y \rangle - h(y)$ is continuous and negative at x_0. Hence there exist $\mu > 0$, $\epsilon > 0$ such that

$$g(z) - \langle \nabla h(y), z - y \rangle - h(y) < -\mu < 0,$$

for all $y \in B(x_0, \epsilon) \cap C$. For each $y \in B(x_0, \epsilon) \cap S$, defining the set $S(y)$ by

$$S(y) = \{x \in C : \; g(x) - \langle \nabla h(y), \; x - y \rangle - h(y) \leq 0\}.$$

By the convexity of h, $S(y) \subset S$. Consequently,

$$d(x, S) \leq d(x, S(y)) \text{ for all } x \in C, \; y \in B(x_0, \epsilon) \cap S.$$

Let now $x \in C \setminus S$. Then $g(x) - \langle \nabla h(x), x - y \rangle - h(y) > 0$. Therefore, there exist $x_y := tx + (1 - t)z \in [x, z]$, $t \in (0, 1)$ such that $g(x_y) - \langle \nabla h(y), x_y - y \rangle - h(y) = 0$. Thus $x_y \in S(y)$ and

$$d(x, S) \leq \|x - x_y\| = (1 - t)\|x - z\|. \tag{35}$$

On the other hand,

$$\begin{aligned} 0 = g(x_y) - \langle \nabla h(y), x_y - y \rangle - h(y) &\leq t[g(x) - \langle \nabla h(y), x - y \rangle - h(y)] \\ &+ (1 - t)[g(z) - \langle \nabla h(y), z - y \rangle - h(y)]. \end{aligned}$$

It follows that

$$(1 - t) \leq \frac{g(x) - \langle \nabla h(y), x - y \rangle - h(y)}{\mu}. \tag{36}$$

From (35) and (36), we obtain

$$d(x, S) \leq d(x, S(y)) \leq \frac{g(x) - \langle \nabla h(y), x - y \rangle - h(y)}{\mu} \|x - z\|, \tag{37}$$

for all $x \in C, x \notin S(y)$. Next, take $\delta > 0$ such that

$$\frac{\|x_0 - z\| + \epsilon}{\mu} \delta < 1/2. \tag{38}$$

Since h is continuously differentiable on C, then there exists $\epsilon' > 0$ with $\epsilon' < \epsilon$ such that

$$\|\nabla h(u) - \nabla h(v)\| < \delta \text{ for all } u, v \in C, \; \|u - v\| < \epsilon'.$$

Let $x \in B(x_0, \epsilon'/2) \cap C$ be given. Pick $y \in S$ such that $\|x - y\| = d(x, S)$. Obviously, $y \in B(x_0, \epsilon)$. Thanks to the mean value theorem, we can find $y' \in [x, y]$ such that

$$h(x) - h(y) = \langle \nabla h(y'), x - y \rangle.$$

Hence

$$h(x) - h(y) \leq \langle \nabla h(y), x - y \rangle + \delta \|x - y\|. \tag{39}$$

Combining the relations (37), (38), (39), we derive that

$$\begin{aligned} d(x, S) &\leq \frac{\|x - z\|}{\mu}[g(x) - h(x)] + \frac{\|x - z\|}{\mu}\delta\|x - y\| \\ &\leq \frac{\|x_0 - z\| + \epsilon}{\mu}[g(x) - h(x)] + \frac{d(x, S)}{2}. \end{aligned} \tag{40}$$

Hence

$$d(x, S) \leq \tau[g(x) - h(x)] \ \text{ for all } \ x \in B(x_0, \epsilon'/2) \cap C,$$

where $\tau = 2(\|x_0 - z\| + \epsilon)/\mu$. In view of Theorem 2, the proof is complete. □

In the case of systems consisting of quadratic inequalities, the conditions in the preceding proposition can be rewritten more simply as follows.

Corollary 2 *Let C be a nonempty compact convex subset of $\mathbb{R}^n$ and*

$$f_i(x) := \frac{1}{2}\langle x, Q_i x \rangle + \langle q_i, x \rangle + r_i, \quad i = 1, \ldots, m$$

where Q_i are symmetric $n \times n$ matrices, $q_i \in \mathbb{R}^n$, $r_i \in \mathbb{R}$. Let

$$f(x) := max\{f_i(x) : i = 1, \ldots, m\}.$$

If for each $x \in S := \{x \in C : f(x) \leq 0\}$ with $f(x) = 0$, there holds

$$\{z \in C : f(z) + \frac{\lambda}{2}\|z - x\|^2 < 0\} \neq \emptyset,$$

where $\lambda := \max\{\rho(Q_i) : i = 1, \ldots, m\}$ and $\rho(Q_i)$ is the spectral radius of Q_i, then there exists $\tau > 0$ such that

$$d(x, S) \leq \tau f^+(x) \ \ \text{for all} \ \ x \in C.$$

Proof The function f is a DC function **on** $\mathbb{R}^n$ whose DC decomposition given below, according to [70], is particularly interesting:

$$f(x) = \max\left\{\frac{1}{2}\langle x, (Q_i + \lambda I)x \rangle + \langle q_i, x \rangle + r_i : i = 1, \ldots, m\right\} - \frac{\lambda}{2}\|x\|^2 := g(x) - h(x),$$

where I denotes the identity matrix. Obviously, the functions g, h are convex. The conclusion follows immediately from Proposition 3 and the following equality.

$$g(z) - \langle \nabla h(x), z - x \rangle - h(x) = f(z) + \frac{\lambda}{2}\|z - x\|^2.$$

□

Consider now the following system of the concave inequalities:

$$S = \{x \in \mathbb{R}^n : \ h_i(x) \leq 0, \ i = 1, \ldots, m\},$$

where $h_i : \mathbb{R}^n \to \mathbb{R}$, $(i = 1, \ldots, m)$ are m concave functions. Setting ($\Im := \{1, \ldots, m\}$)

$$h(x) := \max\{h_1(x), \ldots, h_m(x)\}, \ \Im(x) := \{i \in \{1, \ldots, m\} \ : h_i(x) = h(x)\}$$

Proposition 3 *Let h_i, $(i = 1, \ldots, m)$, h and S be defined as above. If S and the set*

$$\{z \in \mathbb{R}^n : h_i(z) > 0 \text{for } i = 1, \ldots, m\} \tag{41}$$

are nonempty, then for every $\rho > 0$, there exists $\tau > 0$ such that

$$d(x, S) \leq \tau h^+(x) \ \textit{ for all } x \in \mathbb{R}^n, \|x\| \leq \rho.$$

Moreover, if $\mathbb{R}^n \setminus S$ is bounded then the inequality holds for all $x \in \mathbb{R}^n$.

Proof By the finiteness and the concavity of h_i on $\mathbb{R}^n$, the functions h_i, $(i = 1, \ldots, m)$, are continuous on $\mathbb{R}^n$, so is the function h. By compactness, we only need to show that S has a local error bound at each $x \in S$. Indeed, for any $x \in S$, if $h(x) < 0$ then the conclusion trivially holds. Suppose $h(x) = 0$. Thanks to [44,57,72] and several other contributions, it suffices to prove that $0 \notin \partial^0 h(x)$. By [35] (Proposition 2.3.12), one has

$$\partial^0 h(x) \subset co\{-\partial(-h_i)(x) \quad : \quad i \in \Im(x)\}.$$

Let $y \in \partial^0 h(x)$. Then there exist $\lambda_i \in [0, 1]$, $\sum_{i \in \Im(x)} \lambda_i = 1$; $y_i \in -\partial(-h_i)(x)$, $i \in \Im(x)$ such that $y = \sum_{i \in \Im(x)} \lambda_i y_i$. Hence

$$\langle y, z - x\rangle = \sum_{i \in \Im(x)} \lambda_i \langle y_i, z - x\rangle \geq \sum_{i \in \Im(x)} \lambda_i (h_i(z) - h_i(x)) = h(z) > 0.$$

Therefore, $y \neq 0$ and which completes the proof. □

In the case of $m = 1$ concave inequality, if the assumption (41) doesn't hold then $S = \mathbb{R}^n$ and it is trivial that $d(x, S) \leq \tau h^+(x)$ for all $x \in \mathbb{R}^n$. Hence one has

Corollary 3 *Let h be a finite concave function on $\mathbb{R}^n$ and let $S := \{x \in \mathbb{R}^n : h(x) \leq 0\}$ be nonempty. Then for every $\rho > 0$, there exists $\tau > 0$ such that*

$$d(x, S) \leq \tau h^+(x) \ \textit{ for all } x \in \mathbb{R}^n, \|x\| \leq \rho.$$

Moreover if $\mathbb{R}^n \setminus S$ is bounded, then

$$d(x, S) \leq \tau h^+(x) \ \textit{ for all } x \in \mathbb{R}^n.$$

Remark 4 Let $E := \{x \in \mathbb{R}^n : \|x\| = r\}$ that can be written as $E := \{x \in C : h_1(x) = r - \|x\| \leq 0\}$ where $C := \{x \in \mathbb{R}^n : \|x\| \leq r\}$. According to Corollary 4 there is $\tau_1 > 0$ such that

$$d(x, E) \leq \tau_1 h_1(x) \ \text{ for all } x \in C. \tag{42}$$

Similarly using the concave function $h_2(x) := r^2 - \|x\|^2 \geq 0$ to define $E := \{x \in C : h_2(x) \leq 0\}$ yields the error bound with $\tau_2 > 0$

$$d(x, E) \leq \tau_2 h_2(x) \text{ for all } x \in C \tag{43}$$

Since

$$2rh_1(x) \geq h_2(x) := r^2 - \|x\|^2 = [r + \|x\|][r - \|x\|] = [r + \|x\|]h_1(x) \geq rh_1(x) \ \ \forall x \in C, \tag{44}$$

there hold

(i) If $\tau_1 > 0$ verifies (42) then $\tau_2 = \frac{\tau_1}{r}$ verifies (43)
(ii) Conversely if $\tau_2 > 0$ verfies (43) then $\tau_1 = 2r\tau_2$ verifies (42)

Consider now the feasible set $F := \{x \in B(c,r) : h_3(x) := r - \|x - c\| \geq 0\}$ in Proposition 2. Since $B(c,r) = c + C$, $F = c + S$ and the distance is invariant with respect to translations, it follows from (42) that there $\tau_3 > 0$ such that

$$d(x, F) \leq \tau_3 h_3(x) \ \text{ for all } x \in B(c,r) \tag{45}$$

Likewise, the concave function $h_4(x) := r^2 - \|x - c\|^2$ leads to the error bound with $\tau_4 > 0$

$$d(x, F) \leq \tau_4 h_4(x) \ \text{ for all } x \in B(c,r) \tag{46}$$

and the parameters τ_3, τ_4 satisfy the properties (i) and (ii)as τ_1, τ_2. Note at last that the error bound of F given by (32) and (33) are practically advantageous insofar as the parameters τ_3, τ_4 are known.

3.2 Error bounds for concave inequality systems over polyhedral convex sets

Let now K be a nonempty bounded polyhedral convex set and let h be a finite concave function on K. Consider the following set:

$$S = \{x \in K : \ h(x) \leq 0\}.$$

The constraint "$h(x) \leq 0$" is often called reverse convex constraint. Of course, the preceding results on exact penalty in Sect. 2 can be applicable for this system. However, under additional assumptions, we can establish some results on error bounds which are easy to verify in many applications. It is worth noting that the sets of that type figure in many practical problems (see, e.g., [1,2,8,11]).

Theorem 4 *Let K be a nonempty bounded polyhedral convex set in $\mathbb{R}^n$ and let h be a differentiable concave function on K. If S is nonempty, then there exists $\tau > 0$ such that*

$$d(x, S) \leq \tau h^+(x) \ \textit{ for all } \ x \in K. \tag{47}$$

Proof We prove this by induction on the dimension $\dim K$ of K. When $\dim K = 0$, then K has a single element. Then, by the nonemptiness of S, $S = K$ and (47) holds trivially with $t_0 = 0$. Suppose that (47) holds for every nonempty bounded polyhedral convex set K with $\dim K \leq k$. Now, let K be a bounded polyhedral convex set with $\dim K = k + 1$. Denote by $\mathcal{F}(K)$ the set of faces F of K such that $F \neq K$, and by $\Im(K)$ the set of faces F of K such that $F \neq K$ and $F \cap S \neq \emptyset$. Obviuously, $\Im(K)$ is a nonempty finite set. By the induction hypothesis, for each $F \in \Im(K)$, there exists $t(F) \geq 0$ such that

$$d(x, S \cap F) \leq t(F) h^+(x) \quad \text{for all } x \in F \tag{48}$$

Let us set

$$t_1 := \max_{F \in \Im(K)} t(F) \in (0, +\infty).$$

For each $F \in \mathcal{F}(K) \setminus \Im(K)$, then $h(x)$ is positive on F. Since h is continuous relative on the compact set K, then h attains a minimizer on that each F with $m(F) := \min_{x \in F} h(x) > 0$. Denote by

$$m := \min_{F \in \mathcal{F}(K)\setminus\Im(K)} m(F) > 0 \text{ if } \mathcal{F}(K) \setminus \Im(K) \neq \emptyset, \text{ otherwise } m = +\infty;$$
$$M := \max_{x \in K} g(x) \text{ and } \Delta := \operatorname{diam} K = \max\{\|x - y\| : x, y \in K\}.$$

If $M = 0$ then $S = K$ and then the proof is done. Suppose that $M > 0$ and pick $\epsilon \in (0, M/\Delta)$. Taking

$$t_0 = \max\left\{t_1, \frac{1}{\epsilon}, \frac{\Delta}{M - \epsilon\Delta}, \frac{\Delta}{m}\right\}$$

and let $t > t_0$. Let $x \in K \setminus S$ be given and let y^* be a solution of the problem

$$\min\{\|x - y\| + th^+(x) : \ x \in K\}.$$

We need to prove that $y^* \in S$. Suppose this isn't the case. We show that $y^* \notin \operatorname{ri}(K)$, where, $\operatorname{ri}(K)$ denotes the relative interior of K. Indeed, by working on the Affine hull $\operatorname{Aff}(K)$ of K instead of $\mathbb{R}^n$, without loss of generality, we can suppose that $\operatorname{Aff}(K) = \mathbb{R}^n$ and therefore $\operatorname{ri}(K) = \operatorname{int}(K)$. Assume to the contrary that $y^* \in \operatorname{int}(K)$. Then y^* is a local minimizer of the function $y \mapsto \|x - y\| + th(y)$. Hence, one has

$$t\|\nabla h(y^*)\| \leq 1.$$

Thus, by the concavity of h, one obtains

$$h(z) - h(y^*) \leq \langle \nabla h(y^*), z - y^* \rangle \leq \|\nabla h(y^*)\| \|z - y^*\| \leq \epsilon\Delta, \text{ for all } z \in K.$$

Consequently, $h(y^*) \geq M - \epsilon\Delta > 0$. Pick a $z \in S$, since y^* is a minimizer of the function $\|x - \cdot\| + th^+$ on K, one derives that

$$t \leq \frac{\|x - z\| - \|x - y^*\|}{h(y^*)} \leq \frac{\|z - y^*\|}{M - \epsilon\Delta} \leq \frac{\Delta}{M - \epsilon\Delta} \leq t_0,$$

a contradiction. Hence $y^* \notin \operatorname{ri}(K)$. It follows that there exists $F \in \mathcal{F}(K)$ such that $y^* \in F$. Then F must belong to $\Im(K)$ because of the definition of m, t_0 and that

$$h(y^*) \leq \frac{\|x - z\| - \|x - y^*\|}{t} \leq \frac{\Delta}{t} < \frac{\Delta}{t_0} \leq m.$$

Thus, one has $y^* \in \operatorname{argmin}\{\|x - y\| + th^+(y) : \ y \in F\}$ and

$$th^+(y^*) \leq \|x - y\| - \|x - y^*\| \leq \|y - y^*\| \quad \text{for all} \quad y \in S \cap F.$$

Hence, $th^+(y^*) \leq d(y^*, S \cap F)$. This together with the relation (48) imply $y^* \in S \cap F$. As a result, $th(x) \geq \|x - y^*\| \geq d(x, S)$ and the proof is completed. □

As before, let K be a given nonempty bounded polyhedral convex set. We now consider the function h defined by

$$h(x) = \sum_{j=1}^{m} \min\{h_{ij}(x) : \ i \in \Im_j\} \tag{49}$$

where $\Im_1, \ldots, \Im_m$ are finite index sets and h_{ij} are differentiable concave functions on K. The functions of this type are not differentiable. By applying Theorem 4, we next show that the conclusion of Theorem 4 remains valid for the function h given by (49).

Theorem 5 *Let K be a nonempty bounded polyhedral convex set and let h be defined as above. Suppose that the set $S := \{x \in K : \; h(x) \le 0\}$ is nonempty. Then there exists $\tau > 0$ such that*

$$d(x, S) \le \tau h^+(x) \;\; \textit{for all} \;\; x \in K.$$

Proof For any $(x, j) \in K \times \{1, \ldots, m\}$, let

$$\Im(j, x) := \left\{ i \in \Im_j : \; h_{ij}(x) = \min_{i \in \Im_j} h_{ij}(x) \right\}.$$

For any $(x, j) \in K \times \{1, \ldots, m\}$, pick $i(j, x) \in \Im(j, x)$ and define the set $S(x)$ by

$$S(x) := \left\{ y \in K : \; \sum_{j=1}^{m} h_{i(j,x),j}(y) \le 0 \right\}.$$

Obviously, $S(x) \subset S$, for any $x \in K$. First, we show that there exists $\epsilon > 0$ such that $S(x)$ is nonempty for any $x \in K$ with $h(x) \le \epsilon$. Indeed, suppose for a contradiction that this does not hold. We can select a sequence $\{x^p\}$ such that $S(x^p) = \emptyset$, $\lim_{p \to \infty} h(x^p) \le 0$ and that (relabeling if necessary) for each $j = 1, \ldots, m$, $i(j, x^p)$ coincide for all p. By the compactness of K, without loss of generality, assume that x^p converges to $x^* \in K$. One has $\sum_{j=1}^{m} h_{i(j,x^p),j}(x^*) \le 0$. Therefore, $x^* \in S(x^p)$ for all p, contrary to the emptiness of $S(x^p)$. According to Theorem 4 and since the index sets $\Im_j$ are finite, there is $\tau > 0$ such that

$$d(y, S(x)) \le \tau \sum_{j=1}^{m} [h_{i(j,x),j}]^+(y), \quad \text{for all} \;\; x, y \in K, \; h^+(x) \le \epsilon.$$

Hence, for all $x \in K$ satisfying $h(x) \le \epsilon$, one has

$$d(x, S) \le d(x, S(x)) \le \tau \sum_{j=1}^{m} [h_{i(j,x),j}]^+(x) = \tau h^+(x).$$

Since K is bounded, there is $\rho > 0$ such that $d(x, S) \le \rho$ for all $x \in K$. Therefore, $d(x, S) \le (\rho/\epsilon) h(x)$, for all $x \in K$ satisfying $h(x) > \epsilon$. The proof is complete. □

Combining Theorem 5 and Hoffman's result on the error bound for linear inequality system, we state the following result.

Theorem 6 *Let A be a $n \times m$ matrix and let $a \in \mathbb{R}^m$. Suppose that h is a concave function defined by (49) in which, h_{ij} are differentiable concave functions on $\mathbb{R}^n$. Assume that the set*

$$S := \{x \in \mathbb{R}^n : \;\; Ax + a \le 0, \; h(x) \le 0\} \qquad \textit{is nonempty}.$$

Then, for any $\rho > 0$, there exists $\tau > 0$ such that

$$d(x, S) \le \tau \left(\|[Ax + a]^+\| + h^+(x) \right) \;\; \textit{for all} \;\; x \in \mathbb{R}^n, \; \|x\| \le \rho.$$

Proof Because all norms on a finite-dimensional vector space are equivalent, it suffices to prove the theorem with the norm $\| \cdot \| := \| \cdot \|_1$. Let $\rho > 0$ be given and $x^* \in S$. Consider the following truncated set of S :

$$S_1 := S \cap \{x \in \mathbb{R}^n : \;\; \|x\| \le \|x^*\| + 2\rho\}.$$

For any $x, y \in \mathbb{R}^n$ with $\|x\| < \rho$ and $\|y\| > \|x^*\| + 2\rho$, one has

$$\|x - y\| \geq \|y\| - \|x\| > \|x^*\| + 2\rho - \rho \geq \|x - x^*\|.$$

Therefore, $d(x, S_1) = d(x, S)$, for all $x \in \mathbb{R}^n$ with $\|x\| \leq \rho$. Denote by

$$K := \{x \in \mathbb{R}^n : \quad Ax + a \leq 0, \ \|x\| \leq \|x^*\| + 2\rho\}.$$

Then, K is a bounded polyhedral convex set. In virtue of Theorem 5, there exists $\tau > 0$ such that

$$d(x, S) = d(x, S_1) \leq \tau h^+(x) \quad \text{for all} \ \ x \in K.$$

For $x \in \mathbb{R}^n$ with $\|x\| \leq \rho$, let us take $y \in K$ such that $d(x, K) = \|x - y\|$. Denote by L the Lipschitz constant of h on $(\|x^*\| + 2\rho)B_{\mathbb{R}^n}$ ($B_{\mathbb{R}^n}$ stands for the unit ball in $\mathbb{R}^n$). One has

$$d(x, S) \leq \|x - y\| + d(y, S) \leq \|x - y\| + \tau h^+(y) \leq (\tau L + 1)d(x, K) + \tau h^+(x). \tag{50}$$

On the other hand, according to Hoffman's result on the error bound for linear inequality systems [40], there exists $\beta > 0$ such that

$$d(x, K) \leq \beta \|[Ax + a]^+\| \quad \text{for all} \ \ x \in \mathbb{R}^n, \ \|x\| \leq \rho. \tag{51}$$

From the inequalities (50), (51), one obtains

$$d(x, S) \leq (\tau L + 1)\beta \|[Ax + a]^+\| + \tau h^+(x) \quad \text{for all} \ \ x \in \mathbb{R}^n, \ \|x\| \leq \rho.$$

□

A direct application of Theorem 6 is to derive the error bound for the following *linear complementarity* problem:

(LCP) *Find* $x \in \mathbb{R}^n$ *such that* $Ax + a \geq 0$, $Bx + b \geq 0$ *and* $\langle Ax + a, Bx + b\rangle = 0$,

where A, B are real $n \times n$ matrices and $a, b \in \mathbb{R}^n$. Let $A_i, B_i (j = 1, \ldots, n)$ are the row vectors of A, B, respectively and $a := (a_1, \ldots, a_n)^T$, $b := (b_1, \ldots, b_n)^T$. Denote by S the solution set of (LCP). It is obvious that

$$S = \left\{ x \in \mathbb{R}^n : \quad Ax + a \geq 0, \ Bx + b \geq 0, \ h(x) := \sum_{i=1}^{n} \min\{A_i x + a_i, B_i x + b_i\} \leq 0 \right\}.$$

Corollary 4 *For any* $\rho > 0$, *there exists* $\tau > 0$ *such that the following inequality holds*

$$d(x, S) \leq \tau \left(\|[-Ax - a]^+\| + \|[-Bx - b]^+\| + \left[\sum_{i=1}^{n} \min\{A_i x + a_j, B_i x + b_i\} \right]^+ \right)$$

for all $x \in \mathbb{R}^n$ *with* $\|x\| \leq \rho$.

In [53], it was established a error bound result for linear monotone complementary problems. It is worth to note that the error bound result in Corollary 4 does not require the added assumption of monotonicity.

Let us end this section with two other error bounds resulting from Theorem 6. The first one involves the convex quadratic constraints

$$g_i(x) := \frac{1}{2}\langle Q_i x, x\rangle + \langle q_i, x\rangle + r_i, \quad i = 1, \ldots, m \tag{52}$$

where $q_i \in \mathbb{R}^n$, $r_i \in \mathbb{R}$ and Q_i are symmetric positive semidefinite $n \times n$ matrices. Let K be a nonempty bounded polyhedral convex set in $\mathbb{R}^n$. Consider the following inequality system:

$$S = \big\{x \in K : \; g_i(x) \le 0, \; h_p(x) \le 0, \; (i, p) \in \{1, \dots, m\} \times \{1, \dots, \ell\}\big\},$$

where h_p are functions of the same type as in Theorem 5, i.e.,

$$h_p(x) := \sum_{j \in \Im_p} \min\{h_{ij}(x) : \; i \in \Im_{j,p}\},$$

where $\Im_p, \Im_{j,p}$ are finite index sets and h_{ij} are differentiable concave functions on K. Set

$$S_1 = \{x \in K : \;\; g_i(x) \le 0, \; i = 1, \dots, m\}. \tag{53}$$

Theorem 5 yields the following result:

Corollary 5 *Let g_i, h_p be defined as above. In addition, suppose that S is nonempty, g_i, $(i = 1, \dots, m)$, are nonnegative on K and h_p, $(p = 1, \dots, \ell)$, are nonnegative on S_1. Then there exists $\tau > 0$ such that*

$$d(x, S) \le \tau \sum_{p=1}^{l} h_p(x) \;\; \textit{for all} \;\; x \in S_1.$$

Proof As shown by Wang-Pang in [71]: if the convex quadratic function g_i are nonnegative on K, then the set S_1 can be written as follows:

$$S_1 = \big\{x \in K : \;\; Q_i(x - z) = 0, \;\; \langle Q_i z + q_i, x - z\rangle \le 0, i = 1, \dots, m\big\}. \tag{54}$$

where z is an arbitrary element of S_1. Thus, S_1 is a nonempty bounded polyhedral convex set. Since all h_p are nonnegative on S_1, there holds

$$S = \left\{x \in S_1 : \;\; h(x) := \sum_{p=1}^{l} h_p(x) \le 0\right\}.$$

Hence, the conclusion follows directly from Theorem 5. □

Finally the second error bound derived from Theorem 5 is related to finite feasible sets

Corollary 6 *Let K be a nonempty bounded polyhedral convex set and let h be as in Theorem 5. Assume that the set $S := \{x \in K \; : \; h(x) \le 0\}$ has finitely many elements. Then for all closed set $C \subset K$ such that $S^{'} = \{x \in C \; : \; h(x) \le 0\}$ is nonempty, there exists $\tau > 0$ such that*

$$d(x, S') \le \tau h^+(x) \;\;\; \textit{for all} \;\;\; x \in C.$$

Proof In virtue of Theorem 5, there exists $\tau_1 > 0$ such that

$$d(x, S) \le \tau_1 h^+(x) \quad \text{for all} \quad x \in K.$$

If $S \subset C$ then $S' = S$ and the conclusion holds trivially. Otherwise, set $E := S \setminus C, \alpha := \min_{x \in E} d(x, C)$ and $\beta := \max_{x,y \in K} \|x - y\|$. Then $\alpha > 0$ because E is a finite set. Let $x \in C$ and $z \in S$ such that $\|x - z\| = d(x, S)$. If $z \in S'$ then $d(x, S') = \|x - z\| \le \tau_1 h^+(x)$. Otherwise, $z \in E$ and $\tau_1 h^+(x) \ge \|x - z\| \ge \alpha$. Therefore, $d(x, S') \le \beta \le \frac{\tau_1 \beta}{\alpha} h^+(x)$, and the proof is complete. □

Remark 5 It is easy to see that the assumption on the finiteness of the set S is satisfied if h is strictly concave and nonnegative on K. For example, consider the feasible set in $0-1$ programming:

$$S := \{x \in C : x \in \{0,1\}^n\}$$

where C is a closed subset of $\mathbb{R}^n$. One can rewrite $S = \{x \in C \cap [0,1]^n : h(x) \leq 0\}$ where $h(x) := \langle e, x\rangle - \|x\|^2$, and $e \in \mathbb{R}^n$ is the vector of ones. Obviously, the assumptions of Corollary 6 are satisfied.

Note that the concave (but not strictly) penalty function

$$\overline{h}(x) = \sum_{i=1}^{n} \min\{x_i, 1 - x_i\}$$

also verifies the assumptions of Corollary 6.

4 Exact penalty in DC programming via error bounds

To complement the results concerning exact penalty in concave programming of Sect. 2, we use the error bounds results in Sect. 3 to establish below some exact penalty properties in DC programming. Recall that, throughout this section, two problems are said to be equivalent if they have the same optimal value and the same optimal solution set.

First, we give a general exact penalty result which is a refinement of that mentioned by Clarke in [35]. Recall that a real-valued function f defined on a set C in $\mathbb{R}^n$ is said to be Lipschitz on C, if there exists a nonnegative scalar L such that

$$|f(x) - f(y)| \leq L\|x - y\|$$

for all $x, y \in C$. Also f is said to be locally Lipschitz relative to C at some $x \in C$ if for some $\epsilon > 0$, f is Lipschitz on $B(x, \epsilon) \cap C$.

Let θ be a finite DC function on a closed convex set C, i.e.

$$\theta(x) = \varphi(x) - \psi(x) \ \ \forall x \in C, \tag{55}$$

where φ and ψ belong to $\Gamma_0(\mathbb{R}^n)$ such that

$$C \subset dom\ \varphi := \{x \in \mathbb{R}^n : \varphi(x) < +\infty\} \subset dom\ \psi \tag{56}$$

According to [39,68], if

(i) $dom\ \varphi$ and $dom\ \psi$ have the same dimension and
(ii) C is bounded and contained in the relative interior of $dom\ \varphi$,then the DC function θ is Lipschitz on C.

Let f, h be real-valued functions defined on C. Consider the minimization problem whose optimal solution set is denote by $\mathcal{P}$:

$$\alpha = \inf\{f(x) : \ x \in C, \ h(x) \leq 0\} \qquad (P), \tag{57}$$

that we can write as

$$\alpha = \inf\{f(x) : \ x \in S\}, \tag{58}$$

where

$$S := \{x \in C : h(x) \leq 0\} \tag{59}$$

Let $g : C \to \mathbb{R}$ be a nonnegative function such that S can be expressed by

$$S := \{x \in C : g(x) \leq 0\} \tag{60}$$

Such a function g must verify

$$g(x) = 0 \text{ if and only if } x \in S \tag{61}$$

Exact penalty in mathematical programming usually deals with (Sect. 3)

$$g(x) := \left[h^{+}(x)\right]^{\gamma} \tag{62}$$

For $\tau \geq 0$, we define the problems (P_τ) by

$$\alpha(\tau) = \inf\{f(x) + \tau g(x) : \ x \in C\} \quad (P_\tau), \tag{63}$$

whose optimal solution set is denoted by $\mathcal{P}_\tau$.

Proposition 4 *Let f be a Lipschitz function on C with constant L and let g be a nonnegative finite function on C such that $S := \{x \in C : h(x) \leq 0\} = \{x \in C : g(x) \leq 0\}$. If S is nonempty and there exists some $\ell > 0$ such that*

$$d(x, S) \leq \ell g(x) \ \ \textit{for all} \ \ x \in C,$$

then one has:

(i) $\alpha(\tau) = \alpha$ *and* $\mathcal{P} \subset \mathcal{P}_\tau$ *for all* $\tau \geq L\ell$
(ii) $\mathcal{P}_\tau = \mathcal{P}$ *for all* $\tau > L\ell$

Proof (i) We have

$$\begin{aligned} \alpha(\tau) &= \inf\{f(x) + \tau g(x) : \ x \in C\} \\ &\min\{\inf\{f(x) \ : \ x \in S\}, \inf\{f(x) + \tau g(x) : \ x \in C \setminus S\}\} \\ &\leq \alpha \end{aligned} \tag{64}$$

Hence $\alpha(\tau) = \alpha$ if $\alpha = -\infty$. Suppose now that $\alpha \in \mathbb{R}$. We shall prove the opposite inequality of (64) for $\tau \geq L\ell$. By definition of α, for every $\epsilon > 0$ there exists $z \in S$ such that

$$\alpha \leq f(z) \leq \alpha + \epsilon \tag{65}$$

For any $x \in C$, pick a sequence $\{x_k\}$ in S such that $d(x, S) = \lim_{k \longrightarrow \infty} \|x - x_k\|$. By assumption, for all k, one has

$$f(z) - \epsilon \leq f(x_k) \leq f(x) + L\|x - x_k\|.$$

Letting $k \to +\infty$ yields for

$$f(z) - \epsilon \leq f(x) + Ld(x, S) \leq f(x) + \tau h(x),$$

and so

$$f(z) \leq \alpha(\tau) + \epsilon \tag{66}$$

It implies that $\alpha \leq \alpha(\tau)$. Hence $\alpha = \alpha(\tau)$ and the inclusion $\mathcal{P} \subset \mathcal{P}_\tau$ then is immediate.

(ii) Assume now $\tau > L\ell$ and let $z \in \mathcal{P}_\tau$. If $z \notin S$, then $g(z) > 0$ and there exists a sequence $\{z_k\}$ in S such that $\|z - z_k\| \to d(z, S)$. Since $\tau g(z) > Ld(z, S)$, then for k sufficiently large, one has

$$f(z) + \tau h(z) > f(z) + L\|z - z_k\| \geq f(z_k).$$

This contradicts the fact $z \in \mathcal{P}_\tau$. Hence, $z \in S$ and $f(z) \leq f(x)$ for all $x \in S$, i.e., $z \in \mathcal{P}$.□

The above proof is quite standard and follows the line in [35]. However, it is worth noting that our proof, unlike that of [35], needs neither the nonemptiness of the optimal solution set of (P) nor the closedness of its feasible set S.

4.1 Exact penalty for concave inequalities constraints

The error bound results in Sect. 3 and Theorem 4 allow to state nice exact penalty properties in mathematical programming with concave constraints.

The first exact penalty property concerns the couple of nonconvex programs:

$$\inf\{f(x) : x \in K,\ h(x) \leq 0\} \quad (P) \tag{67}$$

and its penalized program

$$\inf\{f(x) + \tau h^+(x) : x \in K\} \quad (P_\tau), \tag{68}$$

where K is a nonempty bounded polyhedral convex set in $\mathbb{R}^n$ and h is a concave function on K defined by (49)

$$h(x) = \sum_{j=1}^{m} \min\{h_{ij}(x) :\ \ i \in \Im_j\}.$$

For each $j = 1, \ldots, m$, $\Im_j$ is a finite index sets and for $i \in \Im_j$, h_{ij} is differentiable concave functions on K.

Theorem 7 *Assume that the feasible set* $S := \{x \in K : h(x) \leq 0\}$ *of* (67) *is nonempty. If* f *is a Lipschitz function on* K *then there exist* $\tau_0 > 0$ *such that for all* $\tau > \tau_0$ *the problems* (67) *and* (68) *are equivalent.*

Proof It follows directly from Proposition 4 and Theorem 5. □

Next we consider the following problem (P) with additional convex quadratic constraints:

$$\min\{f(x)\ :\ x \in K,\ h(x) \leq 0,\ g_i(x) \leq 0,\ i = 1, \ldots, k\}, \tag{P}$$

where K is a nonempty bounded polyhedral convex set, $g_i(x)$, $(i = 1, \ldots, k)$, are convex quadratic functions

$$g_i(x) := \frac{1}{2}\langle Q_i x, x\rangle + \langle q_i, x\rangle + r_i, \quad i = 1, \ldots, k$$

already encountered in (52) and the function h defined by

$$h(x) := \sum_{j=1}^{m} \min\{\langle a_i, x\rangle + b_i\ :\ i \in \Im_j\}, \tag{69}$$

with $\Im_j$, $(j = 1, \ldots, m)$,being finite index sets and $a_i \in \mathbb{R}^n$, $b_i \in \mathbb{R}$. Note that $-h$ is polyhedral convex.

Theorem 8 *Let K, g_i, $(i = 1, \dots, k)$, h be defined as above and let f be a Lipschitz function on K. Assume further that the feasible set of (P) is nonempty. Then there exists $\tau_0 > 0$ and an integer $s \le k$ such that (P) is equivalent to*

$$\min\{f(x) + \tau[h^+(x)]^{\frac{1}{2^s}} \ : \ x \in K, \ \ g_i(x) \le 0, \ i = 1, \dots, k\}, \qquad (\mathrm{P}_\tau(\tfrac{1}{2^s}))$$

for all $\tau > \tau_0$. In particular we can take $s = k$.

Proof Denote by S and C the feasible sets of (P) and $(\mathrm{P}_\tau(\frac{1}{2^s}))$ respectively

$$C := \{x \in K : \ g_i(x) \le 0, \ i = 1, \dots, k\}$$
$$S := \{x \in C : h(x) \le 0\}$$

It is clear that C is a nonempty bounded closed convex set.

Thanks to Proposition 4, it is enough to show that there exists $\tau > 0$ and an integer $s \le k$ such that

$$d(x, S) \le \tau[h^+(x)]^{\frac{1}{2^s}} \text{ for all } x \in C.$$

For each $j = 1, \dots, m$ and for $x \in S$ pick $i(j, x) \in \Im_j$ such that

$$\langle a_{i(j,x)}, x\rangle + b_{i(j,x)} = \min\{\langle a_i, x\rangle + b_i \ : \ i \in \Im_j\}$$

and set

$$S(x) = \left\{ y \in C : \sum_{j=1}^{m} \langle a_{i(j,x)}, y\rangle + b_{i(j,x)} \le 0 \right\}$$

Obviously, $S(x) \subset S$ for all $x \in C$. Similarly to the proof of Theorem 5, we prove the existence of $\epsilon > 0$ such that $S(x) \ne \emptyset$ for all $x \in C$ with $h(x) \le \epsilon$. Notice that $S(x)$ is the solution set of an affine/convex quadratic inequality system. According to Wang-Pang (Theorem 5 in [71]), and by the finiteness of the index sets $\Im_j$, there exists an integer $s \le k$ and $\tau > 0$ such that

$$d(y, S(x)) \le \tau \left[\left(\sum_{j=1}^{m} \langle a_{i(j,x)}, x\rangle + b_{i(j,x)} \right)^{+} \right]^{\frac{1}{2^s}}$$

whenever $x \in C$ with $S(x) \ne \emptyset$ and $y \in C$. Hence, $d(x, S) \le d(x, S(x)) \le \tau[h^+(x)]^{\frac{1}{2^s}}$ whenever $h(x) \le \epsilon$. If $h(x) > \epsilon$, then we have

$$d(x, S) \le (r/\epsilon^{\frac{1}{2^s}})[h^+(x)]^{\frac{1}{2^s}}$$

for any $r \ge \max\{d(x, S) : x \in C\}$. The proof is complete. □

Application to 0–1 Programming. Consider the following problems in mixed $0-1$ programming:

$$\min\{f(x, y) \ : \ (x, y) \in K, \ \ x \in \{0, 1\}^n\}, \qquad (\mathrm{P1})$$

and more generally,

$$\min\{f(x, y) \ : \ (x, y) \in K, \ \ g_i(x, y) \le 0, \ i = 1, \dots, k, \ \ x \in \{0, 1\}^n\}, \qquad (\mathrm{P2})$$

where K is a nonempty bounded polyhedral convex set in $\mathbb{R}^n \times \mathbb{R}^m$, g_i, $i = 1, \dots, k$ are convex quadratic functions on $\mathbb{R}^n \times \mathbb{R}^m$. We define the functions h and $\overline{h}$ by

$$h(x) = \sum_{i=1}^{n} x_i(1 - x_i), \quad \overline{h}(x) = \sum_{i=1}^{n} \min\{x_i, 1 - x_i\},$$

where $x = (x_1, \dots, x_n) \in \mathbb{R}^n$. Obviously, ($P_1$), ($P_2$) are equivalent to the following, respectively,

$$\min\{f(x, y) \ : \ (x, y) \in K \cap ([0, 1]^n \times \mathbb{R}^m), \ \ h(x) \le 0\}, \tag{P'_1}$$

$$\min\{f(x, y) \ : \ (x, y) \in K \cap ([0, 1]^n \times \mathbb{R}^m), \ \ g_i(x, y) \le 0, \ i = 1, \dots, k, \ \ h(x) \le 0\}. \tag{P'_2}$$

Observe that the set $C := K \cap ([0, 1]^n \times \mathbb{R}^m)$ is also a nonempty bounded polyhedral convex set in $\mathbb{R}^n \times \mathbb{R}^m$, and $h, \overline{h}$ are nonnegative concave function on C. In view of Theorems 5 and 8, if f is Lipschitz on C then there exist $\tau_0 > 0$ such that (P_1), (P_2) are equivalent to the following problems($P_{1,\tau}$), ($P_{2,\tau}(1/2^k)$) whenever $\tau > \tau_0$, respectively

$$\min\{f(x, y) + \tau h(x) \ : \ (x, y) \in K \cap ([0, 1]^n \times \mathbb{R}^m)\}, \tag{$P'_{1,\tau}$}$$

$$\min\{f(x, y) + \tau[\overline{h}(x)]^{1/2^k} \ : \ (x, y) \in K \cap ([0, 1]^n \times \mathbb{R}^m), \ g_i(x, y) \le 0, \ i = 1, \dots, k.\} \tag{$P_{2,\tau}(1/2^k)$}$$

By observing that $\overline{h}(x) \le 2h(x)$ for all $x \in [0, 1]^n$, (P_2), ($P_{2,\tau}(1/2^k)$) are equivalent to

$$\min\{f(x, y) + 2\tau h(x)^{1/2^k} \ : \ (x, y) \in K \cap ([0, 1]^n \times \mathbb{R}^m), \ g_i(x, y) \le 0, \ i = 1, \dots, k.\} \tag{$P_{2,\tau}(1/2^k)$}$$

5 Conclusion

We present in this paper new exact penalty results, with/without error bounds, in DC programming, that strongly rely on the concavity of the objective functions and constraints within a bounded polyhedral convex set. The resulting penalty equivalent DC programs seem to be quite suitable to the DCA for their solution.

Acknowledgments The authors would like to thank the referees for their useful comments and suggestions which have improved the presentation of the revised paper.

References

1. DC Programming and DCA: http://lita.sciences.univ-metz.fr/~lethi/
2. Le Thi, H. A., Pham Dinh, T., Muu, L.D.: Exact penalty in DC programming. Vietnam J. Math. **27**, 169–178 (1999)
3. Le Thi, H. A., Pham Dinh, T.: Solving a class of linearly constrained indefinite quadratic problems by DC algorithms. J. Glob. Optim. **11**, 253–285 (1997)
4. Le Thi, H. A., Pham Dinh, T.: A branch-and-bound method via DC optimization and ellipsoidal technique for box constrained nonconvex quadratic programming problems. J. Glob. Optim. **13**, 171–206 (1998)
5. Le Thi, H. A., Pham Dinh, T.: A continuous approach for large scale linearly constrained quadratic zero-one programming. Optimization **50**, 93–120 (2001)

6. Le Thi, H. A.: An efficient algorithm for globally minimizing a quadratic function under convex quadratic constraints. Math. Program. Ser. A **87**(3), 401–426 (2000)
7. Le Thi, H. A., Pham Dinh, T.: A continuous approach for large-scale constrained quadratic zero-one programming. Optimization **45**(3), 1–28 (2001)
8. Le Thi, H.A., Pham Dinh, T.: DC programming: theory, algorithms and applications. The State of the Art (28 pages). In: Proceedings of the First International Workshop on Global Constrained Optimization and Constraint Satisfaction (Cocos' 02). Valbonne-Sophia Antipolis, France, October 2–4 (2002)
9. Le Thi, H. A., Pham Dinh, T.: Large scale global molecular optimization from distance matrices by a DC optimization appoach. SIAM J. Optim. **14**(1), 77–116 (2003)
10. Le Thi, H. A., Pham Dinh, T., Muu, L.D.: Simplicially constrained DC optimization over the efficient and weakly efficient sets. J. Optim. Theory Applications **117**(3), 503–531 (2003)
11. Le Thi, H. A., Pham Dinh, T.: The DC (difference of convex functions) programming and DCA revisited with DC models of real world nonconvex optimization problems. Ann. Oper. Res. **133**, 23–48 (2005)
12. Le Thi, H. A., Minh, L.H., Pham Dinh, T.: Optimization based DC programming and DCA for hierarchical clustering. Eur. J. Oper. Res. **183**, 1067–1085 (2007)
13. Le Thi, H. A., Pham Dinh, T.: A continuous approach for the concave cost supply problem via DC programming and DCA. Discret. Appl. Math. **156**, 325–338 (2008)
14. Pham Dinh, T., Le Thi, H. A., Akoa, F.: Combining DCA and interior point techniques for large-scale nonconvex quadratic programming optimization. Methods Softw. **23**(4), 609–629 (2008)
15. Le Thi, H. A., Minh, L.H., Vinh, N.V., Pham Dinh, T.: A DC programming approach for feature selection in support vector machines learning. Adv. Data Analysis Classif. **2**(3), 259–278 (2008)
16. Le Thi, H. A., Pham Dinh, T., Thoai, N.V., Nam, N.C.: DC optimization techniques for solving a class of nonlinear bilevel programs. J. Glob. Optim. **44**(3), 313–337 (2009)
17. Minh, L.H., Le Thi, H. A., Pham Dinh, T., Bouvry, P.: A combined DCA-GA for constructing highly nonlinear balanced boolean functions in cryptography. J. Glob. Optim. **47**(4), 597–614 (2010)
18. Thiao, M., Phanm Dinh, T., Le Thi, H. A.: A DC programming approach for sparse eigenvalue problem, ICML 2010, pp. 1063–1070
19. Tao, P.D., Niu, Y.-S.: An efficient DC programming approach for portfolio decision with higher moments. Comput. Optim. Appl. 1–30 (2010). doi:10.1007/s10589-010-9383-x
20. Le Thi, H. A., Pham Dinh, T., Yen, N.D.: Properties of two DC algorithms for quadratic programming. J. Glob. Optim. **49**(3), 481–495 (2010)
21. Pham Dinh, T., Nam, N.C., Le Thi, H. A.: An efficient combination of DCA and B&B using DC/SDP relaxation for globally solving binary quadratic programs. J. Glob. Optim. **48**(4), 595–632 (2010)
22. Le Thi, H. A., Mahdi, M., Pham Dinh, T., Joaquim, J.: A DC programming approach for solving the symmetric eigenvalue complementarity problem. Comput. Optim. Appl. 1–21 (2010). doi:10.1007/s10589-010-9388-5
23. Horst, R., Pardalos, P.M., Thoai, N.V.: Introduction to Global Optimization. Kluwer, Dordrecht, Holland (1995)
24. Horst, R.: Deterministic global optimization with partition sets whose feasibility is not known: application to concave minimization, reverse convex constraints, DC programming, and Lipschitz optimization. J. Optim. Theory Appl. **58**, 11–37 (1988)
25. Horst, R., Phong, T.Q., Thoai, N.V., De Vries, J.: On solving a DC programming problem by a sequence of linear programs. J. Glob. Optim. **1**, 183–203 (1991)
26. Horst, R., Tuy, H.: Global optimization (deterministic approaches). 3rd edn. Springer, Berlin, Germany (1996)
27. Horst, R., Thoai, N.V.: DC programming: overview. J. Optim. Theory Appl. **103**(1), 1–43 (1999)
28. Aze, D.: A survey on error bounds for lower semicontinuous functions. In: Proceedings of the ESAIM (2003)
29. Aze, D., Corvellec, J.-N.: Characterization of error bounds for lower semicontinuous functions on metric spaces. ESAIM Control Optim. Calc. Var **10**, 409–425 (2004)
30. Auslender, A., Crouzeix, J.P.: Global regularily theorems. Math. Oper. Res **13**, 243–253 (1998)
31. Bartelt, M., Li, W.: Exact order of Hoffman's error bounds for elliptic quadratic inequalities derived from vector-valued Chebyshev approximation. Math. Program. Ser. B, **88**(2) 223–253 (2000)
32. Bosch, P., Jourani, A., Henrion, R.: Sufficient conditions for error bounds and applications. Appl. Math. Optim. **50**, 161–181 (2004)
33. Burke, J.V.: An exact penalization viewpoint of contrained optimization. SIAM J. Control Optim. **29**, 968–998 (1991)
34. Burke, J.V., Tseng, P.: A unified analysis of Hoffman's bound via Fenchel duality. SIAM J. Optim. **6**, 265–282 (1996)
35. Clarke, F.H.: Optimization and Nonsmooth Analysis. Wiley, New York (1983)

36. Cornejo, O., Jourani, A., Zalinescu, C.: Conditioning and upper-Lipschitz inverse subdifferentials in nonsmooth optimization. J. Optim. Theory Appl. **95**, 127–148 (1997)
37. Dedieu, J.P.: Penalty functions in subanalytic optimization. Optimization **26**, 27–32 (1992)
38. Fabian, M., Henrion, R., Kruger, A.Y., Outrata, J.V.: Error bounds: necessary and sufficient conditions. Set-Valued Var. Anal. **18**, 121–149 (2010)
39. Hiriart-Urruty, J.B., Lemaréchal, C.: Convex Analysis and Minimization Algorithms, Parts I & II. Springer, Verlag, Berlin, Heidelberg (1991)
40. Hoffman, A.J.: On approximate solutions of systems of linear inequalities. J. Res. Nat. Bur. Stand. **49**, 263–265 (1952)
41. Ioffe, A.: Regular points of Lipschitz functions. Trans. Am. Soc. **251**, 61–69 (1979)
42. Tuy, H.: Convex Analysis and Global Optimization. Kluwer, Acadelic Publisher, Boston, Dordrecht, London (2000)
43. Klatte, D., Li, W.: Asymptotic constraint qualifications and error bounds for convex inequalities. Math. Program. **84**, 137–160 (1999)
44. Jourani, A.: Hoffman's error bound, local controlability and sensivity analysis. SIAM J. Control Optim. **38**, 947–970 (2000)
45. Lewis, A.S., Pang, J.S.: Error bound for convex inequality systems. In: Crouzeix, J.P. et al. (eds.) Generalized convexity and generalized monotonicity. Proceedings of the Fifth Symposium on Generalized Convexity, pp. 75–100. Kluwer, Dordrecht (1997)
46. Lojasiewicz, M.S.: Sur le problème de la division. Stud. Math. **18**, 87–136 (1959)
47. Li, W.: Abadie's constraint qualification, metric regularity, and error bound for differentiable convex inequalities. SIAM J. Optim. **7**(4), 966–978 (1997)
48. Luo, X.D., Luo, Z.Q.: Extension of Hoffman's error bound to polynominal systems. SIAM J. Optim. **4**, 383–392 (1994)
49. Luo, Z.Q., Sturm, J.F.: Error bounds for quadratic systems. In: Frenk, H. et al (eds.) High performance optimization, pp. 383–404. Kluwer, Dordrecht, Netherlands (2000)
50. Luo, Z.Q., Pang, J.S.: Error bounds for the analytic systems and their applications. Math. Program. **67**, 1–28 (1994)
51. Luo, Z.Q., Pang, J.S., Ralph, D., Wu, S-Q.: Exact penalization and stationarity conditions of mathematical programs with equilibrium constraints. Math. Program. **75**, 19–76 (1996)
52. Magasarian, O.L.: A condition number for differentiable convex inequalities. Math. Oper. Res. **10**, 175–179 (1985)
53. Magasarian, O.L., Shiau, T.H.: Error bounds for monotone linear complementarity problems. Math. Program. **36**, 81–89 (1986)
54. Mordukhovich, B.S: Variational analysis and generalized differentiation. I. Basic theory. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 330. Springer, Berlin (2006)
55. Mordukhovich, B.S: Variational analysis and generalized differentiation. II. Applications. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 331. Springer, Berlin, (2006)
56. Ngai, H.V., Théra, M.: On necessary conditions for non-Lipschitz optimization problems. SIAM J. Optim. **12**(3), 656–668 (2002)
57. Ngai, H.V., Théra, M.: Errors bounds and implicit multifunction theorem in smooth Banach spaces and applications to optimization. Set-Valued Anal. **12**, 195–223 (2004)
58. Ngai, H.V., Théra, M.: Error bounds for convex differentiable inequality systems in Banach spaces. Math. Program. **104**, 465–482 (2005)
59. Ngai, H.V., Théra, M.: Error bounds in metric spaces and application to the perturbation stability of metric regularity. SIAM J. Optim. **19**(1), 1–20 (2008)
60. Ngai, H.V., Théra, M.: Error bounds for systems of lower semicontinuous functions in Asplund spaces. Math. Program. **116**, 397–427 (2009)
61. Ngai, H.V., Kruger, A., Théra, M.: Stability of error bounds for semi-infinite convex constraint systems. SIAM J. Optim. **20**(4), 2080–2096 (2010)
62. Ngai, H.V., Kruger, A., Théra, M.: Stability of error bounds for convex constraint systems in Banach spaces. SIAM J. Optim. **20**, 3280–3296
63. Pham Dinh, T., Le Thi, H. A.: Convex analysis approach to D.C. programming: theory, algorithms and applications. Acta Mathe. Vietnamica **22**, 289–355 (1997)
64. Pham Dinh, T., Le Thi, H. A.: A DC optimization algorithm for solving the trust region subproblem. SIAM J. Optim. **8**(2), 476–505 (1998)
65. Ng, K.F., Zheng, X.Y.: Error bound for lower semicontinuous functionsin normed spaces. SIAM J.Optim. **12**, 1–17 (2001)

66. Pang, J.S.: Error bounds in mathematical programming. Math. Program. Ser. B **79**, 232–299 (1997)
67. Penot, J.P.,: Well-behavior, well-posedness and nonsmooth analysis. In: Proceedings of the ć6th International Conference on Math. Meth. in Ope. Research, Pliska Stud. Math. Bulgar. **12**, pp. 141–190 (1998)
68. Rockafellar, R.T.: Convex Analysis. Princeton University Press, (1970)
69. Robinson, S.M.: An application of error bound for convex programming in a linear space. SIAM J. Control Optim. **13**, 271–273 (1975)
70. Phong, T.Q., Pham Dinh, T., Le Thi, H. A.: A new method for solving DC programming problem. Applications to fuel mixture nonconvex optimization problem. J. Glob. Optim. **6**, 87–107 (1995)
71. Wang, T., Pang, J.S.: Global error bound for convex quadratic inequality systems. Optimization **31**, 1–12 (1994)
72. Wu, Z., Ye, J.: On errors bounds for lower semicontinuous functions. Math. Program. **92**, 301–314 (2002)