Math. Program., Ser. A (2013) 137:37–64
DOI 10.1007/s10107-011-0481-z

FULL LENGTH PAPER

Global error bounds for piecewise convex polynomials

Guoyin Li

Received: 29 April 2010 / Accepted: 19 July 2011 / Published online: 4 August 2011

Abstract In this paper, by examining the recession properties of convex polynomials, we provide a necessary and sufficient condition for a piecewise convex polynomial to have a Hölder-type global error bound with an explicit Hölder exponent. Our result extends the corresponding results of Li (SIAM J Control Optim 33(5):1510–1529, 1995) from piecewise convex quadratic functions to piecewise convex polynomials.

Keywords Error bound · Piecewise convex polynomial · Convex optimization · Sensitivity analysis

Mathematics Subject Classification (2000) 90C31 · 90C25 · 65K10

1 Introduction

Let $f : \mathbb{R}^n \to \mathbb{R} \bigcup \{+\infty\}$ be an extended value function. Define the solution set by $S := \{x : f(x) \le 0\}$. We are interested in finding tractable conditions for the existence of a constant τ such that

$$d(x, S) \le \tau \left([f(x)]_+ + [f(x)]_+^{\delta}\right) \quad \text{for all} \quad x \in \mathbb{R}^n, \tag{1.1}$$

where $d(x, S)$ denotes the Euclidean distance between x and the set S, $\delta > 0$ and $[\alpha]_+ = \max\{\alpha, 0\}$. If f satisfies the above inequality (1.1), then we say f has a global

This research was partially supported by a grant from the Australian Research Council.

G. Li (✉)
Department of Applied Mathematics, University of New South Wales,
Sydney 2052, Australia
e-mail: g.li@unsw.edu.au

error bound. This inequality bounds the distance from an arbitrary point $x \in \mathbb{R}^n$ (which explains the term "global") to the set S in terms of a constant multiple of a computable "residual function" which measures the violation of the constraint "$S := \{x : f(x) \leq 0\}$". The study of error bound has attracted a lot of researchers and has found many important applications (see [3,20,34,43] for excellent surveys). In particular, it has been used in sensitivity analysis of linear programming/linear complementary problem. It has also been used as termination criteria for decent algorithms.

The first error bound result is due to Hoffman. He showed that f has a global error bound with the exponent $\delta = 1$ (which we refer it as a Lipschitz-type global error bound) if f can be expressed as the maximum of finitely many affine functions. After the important work of Hoffman, a lot of researchers have devoted themselves to the study of global error bound. For example, when f is convex and S is bounded with nonempty interior, Robinson (cf. [35]) established that f has a Lipschitz-type global error bound. Under the Slater condition and an asymptotic constraint qualification, Mangasarian (cf. [29]) established the same result when f is the maximum of finitely many differentiable convex functions. Later on, Auslender and Crouzeix (cf. [3]) extended Mangsarian's result to cover possible non-differentiable convex functions. Li and Klatte [18,19] also achieved the Lipschitz-type global error bound by imposing some appropriate conditions in terms of the Hausdorff continuity of the set S. Besides, Deng [11,13] obtained the Lipschitz-type global error bound by assuming a Slater condition on the recession function. For some other work, see [8,12,14,32]. All the results we mentioned above are only for the case when (1.1) holds with $\delta = 1$, and they typically require the Slater condition. On the other hand, if the Slater condition is not satisfied, Luo and Luo [27] and later, Wang and Pang [40] showed that f has a global error bound with some exponent $\delta < 1$ (which we often call it a Hölder-type global error bound) if f can be expressed as the maximum of finitely many convex quadratic functions. Moreover, Li [26] established that f has a Hölder-type global error bound with exponent $\delta = 1/2$ if f is a piecewise convex quadratic function and f itself is also convex.

The purpose of this paper is to extend the above Li's error bound result [26] from piecewise convex quadratic functions to piecewise convex polynomials (which occurs naturally in approximation theory, for example, see [39]). In this paper, we obtain a simple *necessary and sufficient condition* for a piecewise convex polynomial f to have a Hölder-type global error bound with an explicit Hölder exponent δ. We achieve this by first establishing a global error bound result for a convex polynomial over polyhedral constraints, and then, gluing each pieces together. Much of our study on error bound is in the spirit of [8,11] and is motivated from the recent work on extension of Frank-Wolfe Theorem [7,33] (some other approaches and related references for studying error bound can be found in [9,10,14,16,17,20,21,24,25,30,31,34,41,42]). As we will see later, an advantage of our approach is that the corresponding Hölder exponent δ in the Hölder-type global error bound result can be determined explicitly.

The organization of this paper is as follows. In Sect. 2, we collect some definitions and basic results of convex functions. In Sect. 3, we provide a necessary and sufficient condition characterizing when a piecewise convex polynomial has a Hölder-type global error bound with an explicit Hölder exponent. Finally, in Sect. 4, we conclude our paper and point out some possible future research directions.

2 Preliminaries

Throughout this paper, $\mathbb{R}^n$ denotes Euclidean space with dimension n. The corresponding inner product (resp. norm) in $\mathbb{R}^n$ is defined by $\langle x, y\rangle = x^T y$ for any $x, y \in \mathbb{R}^n$ ($\|x\| = (x^T x)^{1/2}$, for any $x \in \mathbb{R}^n$). We use $\mathbb{B}(x, \epsilon)$ (resp. $\overline{\mathbb{B}}(x, \epsilon)$) to denote the open (resp. closed) ball with center x and radius ϵ. For a set A in $\mathbb{R}^n$, the interior (resp. relative interior, closure, convex hull, affine hull) of A is denoted by $\text{int}A$ (resp. $\text{ri}A$, $\overline{A}$, $\text{co}A$, $\text{aff}A$). If A is a subspace, the orthogonal complement of A is denoted by $A^{\perp}$ and is defined as $A^{\perp} := \{d : a^T d = 0,\ \forall a \in A\}$. Let A be a closed convex set in $\mathbb{R}^n$. The indicator function $\delta_A : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is defined by

$$\delta_A(x) := \begin{cases} 0, & \text{if } \ x \in A, \\ +\infty, & \text{otherwise.} \end{cases} \tag{2.1}$$

The (convex) normal cone of A at a point $x \in \mathbb{R}^n$ is defined as

$$N_A(x) = \begin{cases} \{y \in \mathbb{R}^n : y^T(a - x) \le 0 \text{ for all } a \in A\}, & \text{if } x \in A, \\ \emptyset, & \text{otherwise.} \end{cases}$$

For a function f on $\mathbb{R}^n$, the effective domain and the epigraph are respectively defined by $\text{dom} f := \{x \in \mathbb{R}^n : f(x) < +\infty\}$ and $\text{epi} f := \{(x, r) \in \mathbb{R}^n \times \mathbb{R} : f(x) \le r\}$. We say f is proper if $f(x) > -\infty$ for all $x \in \mathbb{R}^n$ and $\text{dom} f \neq \emptyset$. For each $\epsilon \in \mathbb{R}$, we use $[f \le \epsilon]$ (resp. $[f = \epsilon]$) to denote the level set $\{x \in \mathbb{R}^n : f(x) \le \epsilon\}$ (resp. $\{x : f(x) = \epsilon\}$). Moreover, if $\liminf_{x' \to x} f(x') \ge f(x)$ for all $x \in \mathbb{R}^n$, then we say f is a lower semicontinuous function. A function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is said to be convex if

$$f((1 - \mu)x + \mu y) \le (1 - \mu) f(x) + \mu f(y) \quad \text{for all} \quad \mu \in [0, 1] \quad \text{and} \quad x, y \in \mathbb{R}^n.$$

Let f be a proper lower semicontinuous convex functions on $\mathbb{R}^n$. The (convex) subdifferential of f at $x \in \mathbb{R}^n$ is defined by

$$\partial f(x) = \begin{cases} \left\{z \in \mathbb{R}^n : z^T(y - x) \le f(y) - f(x)\, \forall\, y \in \mathbb{R}^n\right\}, & \text{if } x \in \text{dom} f, \\ \emptyset, & \text{otherwise.} \end{cases} \tag{2.2}$$

The (right) directional derivative of f in the direction h is defined by $f'_+(x; h) = \lim_{t \to 0^+} \frac{f(x+th) - f(x)}{t}$. A useful property relates the subdifferential and the directional derivative is that

$$f'_+(x; h) \ge \sup\{z^T h : z \in \partial f(x)\}.$$

Moreover, if we further assume that either f is continuous at x or $f = g + \delta_P$ where g is a continuously differentiable convex function and P is an affine set, then the above inequality can be strengthen to an equality, i.e.,

$$f'_+(x; h) = \sup\left\{z^T h : z \in \partial f(x)\right\}. \tag{2.3}$$

As usual, for any proper lower semicontinuous convex function f on $\mathbb{R}^n$, its conjugate function $f^*:\mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is defined by $f^*(x^*) = \sup_{x\in\mathbb{R}^n}\{\langle x^*, x\rangle - f(x)\}$ for all $x^* \in \mathbb{R}^n$ (cf. [15,23,36]). Let f be a proper function on $\mathbb{R}^n$. Its associated recession function f^∞ is defined by

$$f^\infty(v) = \liminf_{t\to\infty,\, v'\to v} \frac{f(tv')}{t} \quad \text{for all } v \in \mathbb{R}^n. \tag{2.4}$$

If f is further assumed to be lower semicontinuous and convex, then one has (cf. [3, Proposition 2.5.2])

$$\begin{aligned} f^\infty(v) &= \lim_{t\to\infty} \frac{f(x+tv) - f(x)}{t} \\ &= \sup_{t>0} \frac{f(x+tv) - f(x)}{t} \quad \text{for all } x \in \operatorname{dom} f. \end{aligned} \tag{2.5}$$

As usual, we say $f : \mathbb{R}^n \to \mathbb{R}$ is a (real) polynomial if there exists $r \in \mathbb{N}$ such that

$$f(x) = \sum_{0\le|\alpha|\le r} \lambda_\alpha x^\alpha$$

where $\lambda_\alpha \in \mathbb{R}$, $x = (x_1, \ldots, x_n)$, $x^\alpha := x_1^{\alpha_1} \cdots x_n^{\alpha_n}$, $\alpha_j \in \mathbb{N} \cup \{0\}$ and $|\alpha| := \sum_{j=1}^n \alpha_j$. The corresponding constant r is called the degree of f and is denoted by $\deg(f)$.

Definition 2.1 Recall that a continuous function on $\mathbb{R}^n$ is said to be a piecewise convex polynomial if there exist finitely many polyhedra $P_1, \ldots, P_k$ with $\bigcup_{j=1}^k P_j = \mathbb{R}^n$ such that

the restriction of f on each P_j is a convex polynomial.

For a piecewise convex polynomial f, let f_j be the restriction of f on P_j. The degree of f is denoted by $\deg(f)$ and is defined as the maximum of the degree of each f_j.

For a piecewise convex polynomial f, if $\deg(f) = 1$ then it is usually referred as a piecewise affine function. Moreover, if $\deg(f) = 2$, then it is usually referred as a piecewise convex quadratic function.

Two simple and useful examples of piecewise convex polynomials on $\mathbb{R}^n$ are:

- $f(x) = p(x) + L\|[Ax+b]_+\|^\alpha$, $\alpha \in \{1\} \cup \{2n : n \in \mathbb{N}\}$,
  where p is a convex polynomial on $\mathbb{R}^n$, $A = (a_1, \ldots, a_m)^T \in \mathbb{R}^{m\times n}$, $b = (b_1, \ldots, b_m) \in \mathbb{R}^m$ and $[Ax+b]_+ = (\max\{a_1^T x - b_1, 0\} \ldots, \max\{a_m^T x - b_m, 0\})^T \in \mathbb{R}^m$;
- $f(x) = (q \circ g)(x)$, $\forall x \in \mathbb{R}^n$,
  where q is a convex polynomial on $\mathbb{R}^m$, $g = (g_1, \ldots, g_m)$ is a vector value function and each $g_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m$, is a piecewise affine function.

It should be noted that, in general, a piecewise convex polynomial need not to be smooth or convex. For a simple example, consider $f(x) := \min\{x, 1\}$ for all $x \in \mathbb{R}$. In this case, it is clear that f is a piecewise convex polynomial which is neither smooth nor convex.

Finally, we summarize some basic properties of convex polynomials in the following two lemmas. For this first lemma, part (1) is taken from [33, Corollary 4.1 and Lemma 2.4], part (2) is due to [3, Propsition 3.2.1] and part (3) is from [22, Theorem 3.4]. The second lemma is a Frank-Wolfe type result for a convex polynomial system which was established in [7].

Lemma 2.1 *Let f be a convex polynomial on $\mathbb{R}^n$. Then the following statements hold:*

(1) *Let $v \in \mathbb{R}^n$ be such that $f^{\infty}(v) = 0$. Then $f(x + tv) = f(x)$ for all $t \in \mathbb{R}$ and for all $x \in \mathbb{R}^n$.*
(2) *Let x_1, x_2 be two points in $\mathbb{R}^n$. If f takes a constant value on $C := [x_1, x_2]$, then f takes the same constant value on* aff C.
(3) *The function f is asymptotic well behaved in the sense that*

$$\nabla f(x_k) \to 0 \;\Rightarrow\; f(x_k) \to \inf f.$$

Lemma 2.2 *Let $f_0, f_1, \ldots, f_m$ be convex polynomials on $\mathbb{R}^n$. Let $C := \bigcap_{j=1}^{m} [f_j \leq 0]$. Suppose that* $\inf_{x \in C} f_0(x) > -\infty$. *Then,* $\operatorname{argmin}_{x \in C} f_0(x) \neq \emptyset$.

3 Global error bound results

In this section, we study the global error bound results. To begin with, we formally recall the following definitions of error bound.

Definition 3.1 Let f be a proper lower semicontinuous function on $\mathbb{R}^n$. We say f has a

(1) Lipschitz-type global error bound if there exists $\tau > 0$ such that

$$d(x, [f \leq 0]) \leq \tau [f(x)]_+ \quad \text{for all } x \in \mathbb{R}^n,$$

where $[\alpha]_+$ denotes the number $\max\{\alpha, 0\}$.

(2) Hölder-type global error bound if there exist $\tau, \delta > 0$ such that

$$d(x, [f \leq 0]) \leq \tau \big([f(x)]_+ + [f(x)]_+^{\delta}\big) \quad \text{for all } x \in \mathbb{R}^n$$

(The corresponding δ satisfying the preceding inequality is called the Hölder exponent).

(3) Hölder-type local error bound if there exist $r, \tau, \delta > 0$ such that

$$d(x, [f \leq 0]) \leq \tau \big([f(x)]_+ + [f(x)]_+^{\delta}\big) \quad \text{for all } x \in [f \leq r].$$

To avoid the triviality, throughout this section, we assume that $\emptyset \neq [f \leq 0] \neq \mathbb{R}^n$.

Before we proceed to the study of error bound results, we first introduce a definition and summarize some existing results.

Definition 3.2 Let $n, d \in \mathbb{N}$. Define $\kappa(n, d) := (d-1)^n + 1$.

The following necessary and sufficient condition for a Lipschitz-type global error bound plays an important role in our later analysis (cf [20, Theorem 1]). Here, we state a simplified version of it which is convenient for us.

Lemma 3.1 *Let f be a proper lower semicontinuous convex function on $\mathbb{R}^n$ and let $\tau > 0$. Then the following statements are equivalent:*

(1) $d(x, [f \le 0]) \le \tau [f(x)]_+$ *for all $x \in \mathbb{R}^n$;*
(2) $\inf\{f'_+(x; h) : h \in N_{[f\le 0]}(x), \|h\| = 1\} \ge \tau^{-1}$ *for all $x \in \mathrm{bd}([f \le 0])$.*

In general, verifying the above necessary and sufficient condition may not be an easy task. Some simple and tractable sufficient conditions are listed below where (1) and (2) are taken from [20, Corollary 2] and (3) is from [22, Theorem 4.1].

Lemma 3.2 *Let f be a proper lower semicontinuous convex function on $\mathbb{R}^n$. Then f has a Lipschitz-type global error bound if one of the following conditions holds:*

(1) *$[f \le 0]$ is compact, and the Slater condition holds, i.e., there exists $x_0 \in \mathbb{R}^n$ such that $f(x_0) < 0$.*
(2) *There exists $v \in \mathbb{R}^n$ such that $f^\infty(v) < 0$.*
(3) *f is a separable function (in the sense that $f(x) = \sum_{i=1}^n f_i(x_i)$ where $x = (x_1, \ldots, x_n)$ and each f_i is a proper function on $\mathbb{R}$) and the Slater condition holds.*

Finally, we also list three important results which will be used later on. The first one is a Hölder-type global error bound result for a single convex polynomial which was established in our recent paper [22, Theorem 4.10]. The second one is known as the linear regularity for finitely many polyhedral sets. The last one is the fact that the Slater condition implies the basic constraint qualification.

Lemma 3.3 *Let f be a convex polynomial on $\mathbb{R}^n$ with degree d. Then, f has a Hölder-type global error bound with exponent $\kappa(n, d)^{-1}$, i.e., there exists a constant $\tau > 0$ such that*

$$d(x, [f \le 0]) \le \tau \left([f(x)]_+ + [f(x)]_+^{\kappa(n,d)^{-1}}\right) \quad \text{for all } x \in \mathbb{R}^n. \tag{3.1}$$

Lemma 3.4 (cf. [4, Corollary 5.26]) *Let $m \in \mathbb{N}$ and let $A_1, \ldots, A_m$ be polyhedral sets in $\mathbb{R}^n$. Then $\{A_1, \ldots, A_m\}$ is linear regular in the sense that there exists a constant $\tau > 0$ such that*

$$d\left(x, \bigcap_{i=1}^m A_i\right) \le \tau \sum_{i=1}^m d(x, A_i) \quad \text{for all } x \in \mathbb{R}^n.$$

Lemma 3.5 (cf. [43, Corollary 2.9.5]) *Let f be a proper lower semicontinuous and convex function satisfying the Slater condition, i.e., there exists $x_0 \in \mathbb{R}^n$ such that $f(x_0) < 0$. Then the following basic constraint qualification (BCQ) holds:*

$$N_{[f\le 0]}(x) = \bigcup \{\lambda \partial f(x) : \lambda \ge 0, \lambda f(x) = 0\} \quad \text{for all} \quad x \in [f \le 0].$$

3.1 Global error bound for convex polynomials over polyhedral constraints

In this subsection, we establish a global error bound result for convex polynomials over polyhedral constraints. More explicitly, we study the global error bound for a convex function f which takes the following form:

$$f(x) = g(x) + \delta_P(x) = \begin{cases} g(x), & \text{if } x \in P, \\ +\infty, & \text{else.} \end{cases} \tag{3.2}$$

where g is a convex polynomial on $\mathbb{R}^n$ and P is a polyhedron in $\mathbb{R}^n$. The result provided in this subsection will serve as a basis for establishing the global error bound result for piecewise convex polynomials. Below, we first establish a Lipschitz-type global error bound result for a convex function f with the form (3.2), under the Slater condition.

Proposition 3.1 *Let g be a convex polynomial on $\mathbb{R}^n$ and let P be a polyhedron in $\mathbb{R}^n$. Let $f := g + \delta_P$. Suppose that the Slater condition holds, i.e., there exists $x_0 \in \mathbb{R}^n$ such that $f(x_0) < 0$.*[1] *Then, f has a Lipschitz-type global error bound, i.e., there exists a constant $\tau > 0$ such that*

$$d(x, [f \leq 0]) \leq \tau [f(x)]_+ \quad \textit{for all } x \in \mathbb{R}^n. \tag{3.3}$$

Proof Let $P = \{x : a_i^T x \leq b_i, i = 1, \ldots, m\}$. We prove (3.3) by induction on the dimension n (of the underlying space). If $n = 1$, then f is a one dimensional proper lower semicontinuous convex function satisfying the Slater condition. Then, Lemma 3.2(3) implies that the conclusion holds for $n = 1$. Suppose that (3.3) holds for any convex function f with the form $f = g + \delta_P$ where g is a convex polynomial on $\mathbb{R}^s$ and P is a polyhedron in $\mathbb{R}^s$ such that f satisfy the Slater condition. Now, let us consider the case when $n = s + 1$. If $[f \leq 0]$ is compact, then the conclusion holds immediately by Lemma 3.2(1). So, we assume without loss of generality that $[f \leq 0] = [g \leq 0] \cap P$ is unbounded. Then, there exists a sequence $\{x_k\} \subseteq [g \leq 0] \cap P$ such that $\|x_k\| \to +\infty$. By passing to a subsequence if necessary, we may further assume that

$$\frac{x_k}{\|x_k\|} \to v \tag{3.4}$$

for some $v \in \mathbb{R}^{s+1}$ with $\|v\| = 1$. It can be verified that $a_i^T v \leq 0$ for all $i = 1, \ldots, m$, and $g^\infty(v) \leq 0$ (to see this, from (2.4), we see that

$$g^\infty(v) = \liminf_{t \to \infty,\, v' \to v} \frac{g(tv')}{t} \leq \liminf_{n \to \infty} \frac{g\left(\|x_n\| \frac{x_n}{\|x_n\|}\right)}{\|x_n\|} = \liminf_{n \to \infty} \frac{g(x_n)}{\|x_n\|} \leq 0.)$$

[1] This is equivalent to the existence of $x_0 \in P$ with $g(x_0) < 0$.

Suppose that $g^\infty(v) < 0$. It follows from (2.5) that

$$\sup_{t>0} \frac{g(x+tv)-g(x)}{t} = g^\infty(v) < 0 \quad \text{for all } x \in \mathbb{R}^{s+1}.$$

Note that, for any $x \in \operatorname{dom} f = P,\ x + tv \in P\ \forall t > 0$, and so,

$$f^\infty(v) = \sup_{t>0} \frac{f(x+tv)-f(x)}{t} = \sup_{t>0} \frac{g(x+tv)-g(x)}{t} < 0.$$

Therefore, the conclusion follows from Lemma 3.2(2). Now, suppose that $g^\infty(v) = 0$. Then, Lemma 2.1(1) gives us that

$$g(x+tv) = g(x) \text{ for any } x \in \mathbb{R}^{s+1} \text{ and for any } t \in \mathbb{R}. \tag{3.5}$$

To show (3.3), we only need to show

$$d(x, [g \le 0] \cap P) \le \tau [g(x)]_+ \quad \text{for all } x \in P. \tag{3.6}$$

Now, we consider the following two cases: Case 1, $a_i^T v = 0$ for all $i = 1, \dots, m$; Case 2, there exists $i_0 \in \{1, \dots, m\}$ such that $a_{i_0}^T v < 0$.

Suppose that Case 1 holds. Define $A = v^\perp := \{d \in \mathbb{R}^{s+1} : d^T v = 0\}$. Since $v \ne 0$, we have $\dim A = s$. Thus, there exists a full rank matrix $Q \in \mathbb{R}^{(s+1)\times s}$ such that $\{Qz : z \in \mathbb{R}^s\} = A$. Note that $\mathbb{R}^{s+1} = A \oplus \operatorname{span}\{v\}$ where $\oplus$ denotes the direct sum and $\operatorname{span}\{v\} := \{tv : t \in \mathbb{R}\}$. For any $x \in \mathbb{R}^{s+1}$, one has

$$x = \operatorname{Pr}_A(x) + \operatorname{Pr}_{\operatorname{span}\{v\}}(x), \tag{3.7}$$

where Pr is the usual Euclidean projection. From (3.5), we see that $g(x) = g(\operatorname{Pr}_A(x))$ for all $x \in \mathbb{R}^{s+1}$. Letting $h : \mathbb{R}^s \to \mathbb{R}$ be defined by $h(z) := g(Qz)$, it follows that h is a convex polynomial on $\mathbb{R}^s$. Let $P' = \{z \in \mathbb{R}^s : Qz \in \operatorname{Pr}_A(P)\}$. Then we see that P' is a polyhedron in $\mathbb{R}^s$. Since $\{Qz : z \in \mathbb{R}^s\} = A$, there exists $z_0 \in \mathbb{R}^s$ such that $Qz_0 = \operatorname{Pr}_A(x_0)$. This gives us that $z_0 \in P'$ and

$$h(z_0) = g(Qz_0) = g(\operatorname{Pr}_A(x_0)) = g(x_0) < 0.$$

Thus, the induction hypothesis implies that there exists $\mu_1 > 0$ such that

$$d(z, [h \le 0] \cap P') \le \mu_1 [h(z)]_+ \quad \text{for all } z \in P'.$$

We now show that

$$Q([h \le 0]) + \operatorname{span}\{v\} = [g \le 0] \text{ and } Q(P') + \operatorname{span}\{v\} = P. \tag{3.8}$$

Indeed, to see the first relation of (3.8), let $z \in [h \le 0]$ and $t \in \mathbb{R}$. Then, from (3.5), we see that $g(Qz + tv) = g(Qz) = h(z) \le 0$. Thus, $Q([h \le 0]) + \operatorname{span}\{v\} \subseteq [g \le 0]$.

To see the converse inclusion, let $x \in [g \leq 0]$. From (3.7), we have $x = \mathrm{Pr}_A(x) + \mathrm{Pr}_{\mathrm{span}\{v\}}(x)$. Since $\mathrm{Pr}_A(x) \in A = \{Qz : z \in \mathbb{R}^s\}$, there exists $z_0 \in \mathbb{R}^s$ such that $\mathrm{Pr}_A(x) = Qz_0$. Thus, we have $x = Qz_0 + \mathrm{Pr}_{\mathrm{span}\{v\}}(x)$. Noting that $\mathrm{Pr}_{\mathrm{span}\{v\}}(x) \in \mathrm{span}\{v\}$, it suffices to show that $z_0 \in [h \leq 0]$. To see this, since $g(x) = g(\mathrm{Pr}_A(x))$, we have $h(z_0) = g(Qz_0) = g(\mathrm{Pr}_A(x)) = g(x) \leq 0$. To see the second relation, take $x \in P$. From (3.7), we have $x = \mathrm{Pr}_A(x) + \mathrm{Pr}_{\mathrm{span}\{v\}}(x)$. As $\mathrm{Pr}_A(P) = Q(P')$, we see that $P \subseteq Q(P') + \mathrm{span}\{v\}$. Conversely, let $y \in Q(P') = \mathrm{Pr}_A(P)$ and $z \in \mathrm{span}\{v\}$. Then, there exists $x \in P$ such that $y = \mathrm{Pr}_A(x)$. It follows from (3.7) that $y - x \in \mathrm{span}\{v\}$, and so, $y + z = x + \big((y - x) + z\big) \in P + \mathrm{span}\{v\}$. Now, recall that $a_i^T v = 0,\ i = 1, \ldots, m$ and so, $p + tv \in P$ for any $p \in P$ and $t \in \mathbb{R}$. Thus, $y + z \in P$.

It follows from $x \in P$ that there exist $\overline{z} \in P'$ and $q \in \mathbb{R}$ such that $x = Q\overline{z} + qv$. Moreover, (3.8) also implies that $Q([h \leq 0] \cap P') + \mathrm{span}\{v\} \subseteq [g \leq 0] \cap P$, and so,

$$\begin{aligned} d(x, [g \leq 0] \cap P) &\leq d\big(Q\overline{z} + qv, Q([h \leq 0] \cap P') + \mathrm{span}\{v\}\big) \\ &\leq d\big(Q\overline{z}, Q([h \leq 0] \cap P')\big) \\ &\leq \|Q\| d(\overline{z}, [h \leq 0] \cap P') \\ &\leq \mu_1 \|Q\|\, [h(\overline{z})]_+, \end{aligned}$$

where $\|Q\| = \sup\{\|Qz\|_{\mathbb{R}^{s+1}} : \|z\| = 1, z \in \mathbb{R}^s\}$. Since $g(x) = g(Q\overline{z} + qv) = g(Q\overline{z}) = h(\overline{z})$ and $\|Q\| > 0$ (as Q is of full rank), we have

$$d(x, [g \leq 0] \cap P) \leq \mu_1 \|Q\|\, [g(x)]_+ \quad \text{for all } x \in P.$$

Thus, the conclusion holds in this case.

Suppose that Case 2 holds. From Lemma 3.1, it suffices to show that there exists $\tau > 0$ such that

$$\inf\{f'_+(x; h) : h \in N_{[f \leq 0]}(x), \|h\| = 1\} \geq \tau^{-1} \quad \text{for all } x \in \mathrm{bd}([f \leq 0]) \tag{3.9}$$

To see (3.9), fix a point $x \in \mathrm{bd}([f \leq 0])$ with $g(x) < 0$. Then, $x \in \mathrm{bd} P$ and so, $N_{[f \leq 0]}(x) = N_P(x) \neq \{0\}$. It follows that

$$x + th \notin P \quad \text{for all } h \in N_{[f \leq 0]}(x), \|h\| = 1 \text{ and } t > 0.$$

Thus,

$$\inf\{f'_+(x; h) : h \in N_{[f \leq 0]}(x), \|h\| = 1\} = +\infty \quad \text{for all } x \in \mathrm{bd}([f \leq 0]) \text{ with } g(x) < 0.$$

Therefore, to see (3.9), we only need to show that there exists $\tau > 0$ such that

$$\begin{aligned} &\inf\{f'_+(x; h) : h \in N_{[f \leq 0]}(x), \|h\| = 1\} \\ &\quad \geq \tau^{-1} \quad \text{for all } x \in \mathrm{bd}([f \leq 0]) \text{ with } g(x) = 0. \end{aligned} \tag{3.10}$$

Note that $[f \leq 0] = [g \leq 0] \cap P$,

$$\begin{aligned} f'_+(x; u) &\geq \sup \left\{ w^T u : w \in \partial f(x) \right\} \\ &= \sup \left\{ w^T u : w \in \nabla g(x) + N_P(x) \right\} \\ &= \sup \left\{ w^T u : w = \nabla g(x) + \sum_{i \in I(x)} \lambda_i a_i \text{ where } \lambda_i \geq 0 \right\} \text{ for all } u \in \mathbb{R}^n, \end{aligned}$$

where $I(x) := \{1 \leq i \leq m : a_i^T x = b_i\}$ and (from Lemma 3.5 that)

$$N_{[f \leq 0]}(x) = \bigcup_{\alpha \geq 0} \alpha \partial f(x) = \bigcup_{\alpha \geq 0, \lambda_i \geq 0} \alpha \left(\nabla g(x) + \sum_{i \in I(x)} \lambda_i a_i \right).$$

It follows that for all $x \in \mathrm{bd}([f \leq 0])$ with $g(x) = 0$ and $h \in N_{[f \leq 0]}(x)$ with $\|h\| = 1$,

$$h = \frac{\nabla g(x) + \sum_{i \in I(x)} \overline{\lambda}_i a_i}{\left\| \nabla g(x) + \sum_{i \in I(x)} \overline{\lambda}_i a_i \right\|}$$

for some $\overline{\lambda}_i \geq 0$ and so,

$$f'_+(x; h) \geq \left\| \nabla g(x) + \sum_{i \in I(x)} \overline{\lambda}_i a_i \right\|.$$

Therefore, to see (3.10), it remains to show

$$\left\| \nabla g(x) + \sum_{i \in I(x)} \lambda_i a_i \right\| \geq \tau^{-1} \quad \text{for all } \lambda_i \geq 0, x \in \mathrm{bd}([g \leq 0] \cap P) \text{ with } g(x) = 0.$$

We proceed by the method of contradiction and suppose that there exist $\{x^k\} \subseteq \mathrm{bd}([g \leq 0] \cap P)$ with $g(x^k) = 0$ and $\lambda_i^k \geq 0$, $i \in I(x^k)$ such that $\nabla g(x^k) + \sum_{i \in I(x^k)} \lambda_i^k a_i \to 0$. By passing to a subsequence if necessary, we may assume that $I(x^k) \equiv I \neq \emptyset$. This implies that $a_i^T x^k = b_i$ for all $i \in I$, and

$$\nabla g(x^k) + \sum_{i \in I} \lambda_i^k a_i \to 0. \tag{3.11}$$

Recall that $\frac{x^k}{\|x^k\|} \to v$. Define $I_1 = \{i \in I : a_i^T v = 0\}$ and $I_2 = \{i \in I : a_i^T v < 0\}$. Note from (3.5) that $\nabla g(x)^T v = 0$ for all $x \in \mathbb{R}^n$. As $a_i^T v = 0$ for all $i \in I_1$, it follows that

$$\sum_{i\in I_2}\lambda_i^k a_i^T v = \nabla g(x^k)^T v + \sum_{i\in I}\lambda_i^k a_i^T v \to 0.$$

Since $a_i^T v < 0$ for all $i \in I_2$, we have $\lim_{n\to\infty}\lambda_i^k = 0$ for all $i \in I_2$. It then follows from (3.11) that

$$\nabla g(x^k) + \sum_{i\in I_1}\lambda_i^k a_i \to 0. \tag{3.12}$$

Next, we observe that $I_1 \neq \emptyset$ (otherwise, $I_1 = \emptyset$ and so $\nabla g(x^k) \to 0$. Since g is a convex polynomial, g is asymptotic well behaved (see Lemma 2.1(3)). Thus, $g(x^k) \to \inf g$. This together with $g(x^k) = 0$ implies that $\inf g = 0$. This makes contradiction as $g(x_0) < 0$). Consider $P_{I_1} = \{x : a_i^T x = b_i, i \in I_1\}$. We now claim that there exists

$$x^* \in P_{I_1} \quad \text{such that} \quad g(x^*) < 0. \tag{3.13}$$

To see this claim, we proceed by the method of contradiction and suppose that g is nonnegative on P_{I_1}. From the Lagrangian duality, there exist $\gamma_i \in \mathbb{R},\ i \in I_1$, such that $0 \le \inf_{x\in P_{I_1}} g(x) = \inf_{x\in\mathbb{R}^n}\{g(x)+\sum_{i\in I_1}\gamma_i(a_i^T x - b_i)\}$. So, $g(x)+\sum_{i\in I_1}\gamma_i(a_i^T x - b_i) \ge 0$ for all $x \in \mathbb{R}^n$. As $g(x^k) = 0$ and $x^k \in P_{I_1}$, $\inf_{x\in P_{I_1}} g(x) = 0$, we see that $g(x^k) + \sum_{i\in I_1}\gamma_i(a_i^T x^k - b_i) = 0$. It follows that the convex function $q(x) := g(x) + \sum_{i\in I_1}\gamma_i(a_i^T x - b_i)$ attains its minimum at x^k for all $k \in \mathbb{N}$. Hence,

$$\nabla q(x^k) = \nabla g(x^k) + \sum_{i\in I_1}\gamma_i a_i = 0 \quad \text{for all } k \in \mathbb{N}$$

and so, $\nabla g(x^k) \equiv -\sum_{i\in I_1}\gamma_i a_i$ is a constant vector. This together with (3.12) implies that $\|\sum_{i\in I_1}\lambda_i^k a_i\|$ is bounded. Then, by the Hoffman error bound, we may assume without loss of generality that λ_i^k is bounded (Indeed, consider the function $A : \mathbb{R}^{|I_1|} \to \mathbb{R}$, defined by $A(\lambda) = \|\sum_{i\in I_1}\lambda_i a_i\|_\infty$ where $\|\cdot\|_\infty$ is the usual l^∞ norm. Let $S = \{\lambda \in \mathbb{R}^{|I_1|} : A(\lambda) = 0\}$. It is clear that S is a subspace. Let $\lambda^k = (\lambda_i^k)_{i\in I_1} \in \mathbb{R}^{|I_1|}$. Then, we can decompose $\lambda^k = \mu^k + \gamma^k$ where $\mu^k \in S^\perp$ and $\gamma^k \in S$. From the definition of S, we have $\sum_{i\in I_1}\lambda_i^k a_i = \sum_{i\in I_1}\mu_i^k a_i + \sum_{i\in I_1}\gamma_i^k a_i = \sum_{i\in I_1}\mu_i^k a_i$. Moreover, by the Hoffman error bound, there exists $\tau > 0$ such that $\|\mu^k\| = d(\lambda^k, S) \le \tau A(\lambda^k) = \tau\|\sum_{i\in I_1}\lambda_i a_i\|_\infty$. Thus, μ^k is bounded, and so, by replacing λ^k with μ^k if necessary, we can always assume that λ^k is bounded.) By passing to a subsequence if necessary, we have $\lambda_i^k \to \overline{\lambda}_i \ge 0$, and so, $\nabla g(x^{k_0}) + \sum_{i\in I_1}\overline{\lambda}_i a_i = 0$, where k_0 is any natural number. This implies that x^{k_0} is a stationary point of the convex function $w(x) = g(x) + \sum_{i\in I_1}\overline{\lambda}_i(a_i^T x - b_i)$, and so, $g(x) + \sum_{i\in I_1}\overline{\lambda}_i(a_i^T x - b_i) \ge 0$ for all $x \in \mathbb{R}^n$. However, this is impossible as $x_0 \in P$ and $g(x_0) < 0$. So, noticing (3.13) and $a_i^T v = 0$ for all $i \in I_1$, and proceeding much as in case 1, we can find $\alpha > 0$ such that, for all $x \in P_{I_1}$,

$$d(x, [g \leq 0] \cap P_{I_1}) \leq \alpha [g(x)]_+.$$

Let $\overline{f} = g + \delta_{P_{I_1}}$. It follows that $\overline{f}$ is a proper lower semicontinuous convex function and

$$d(x, [\overline{f} \leq 0]) \leq \alpha [\overline{f}(x)]_+ \quad \text{for all} \quad x \in \mathbb{R}^n.$$

Thus, from Lemma 3.1 again, we have for all $x \in \text{bd}([\overline{f} \leq 0])$

$$\inf\{\overline{f}'_+(x; u) : u \in N_{[\overline{f} \leq 0]}(x), \|u\| = 1\} \geq \alpha^{-1}.$$

Since $a_i^T x^k = b_i$, $\forall\, i \in I_1$, we have $x^k \in \text{bd}([\overline{f} \leq 0])$ and (by (2.3))

$$\begin{aligned} \overline{f}'_+(x^k; u) &= \sup\left\{w^T u : w \in \partial \overline{f}(x^k)\right\} \\ &= \sup\left\{\nabla g(x)^T u + v^T u : v \in N_{P_{I_1}}(x^k)\right\} \\ &= \sup\left\{\left(\nabla g(x^k) + \sum_{i \in I_1} \mu_i a_i\right)^T u : \mu_i \in \mathbb{R}\right\}. \end{aligned}$$

This implies that

$$\begin{aligned} &\max\left\{\left(\nabla g(x^k) + \sum_{i \in I_1} \mu_i a_i\right)^T u : \mu_i \in \mathbb{R}\right\} \\ &\qquad \geq \alpha^{-1} \quad \text{for all } u \in N_{[\overline{f} \leq 0]}(x^k) \text{ with } \|u\| = 1. \end{aligned} \tag{3.14}$$

Let $L = \{x \in \mathbb{R}^{s+1} : a_i^T x = 0, i \in I_1\}$. Then $L^\perp = \left\{\sum_{i \in I_1} \lambda_i a_i : \lambda_i \in \mathbb{R}\right\}$. Since $\mathbb{R}^{s+1} = L \oplus L^\perp$, we can write $\nabla g(x^k) = w^k + v^k$ where $w^k = \text{Pr}_L\big(\nabla g(x^k)\big)$ and $v^k \in \text{Pr}_{L^\perp}\big(\nabla g(x^k)\big)$. Then we have

$$\|w^k\| = \min\left\{\left\|\nabla g(x^k) + \sum_{i \in I_1} \lambda_i a_i\right\| : \lambda_i \in \mathbb{R}\right\}$$

and

$$\nabla g(x^k)^T w^k - \|w^k\|^2 = \left(\nabla g(x^k) - w^k\right)^T w^k = (v_k)^T w_k = 0. \tag{3.15}$$

Moreover, for all $x \in [\overline{f} \leq 0] = [g \leq 0] \cap P_{I_1}$, we see that

$$(w^k)^T (x - x^k) = \left(\nabla g(x^k) - v^k\right)^T (x - x^k) \leq 0,$$

and so, $w^k \in N_{[\overline{f} \leq 0]}(x^k)$. Note that $w_k \neq 0$ (otherwise, $\nabla g(x^k) \in L^\perp$ and so, g attains its minimum on P_{I_1} at x^k. This implies that g is nonnegative on P_{I_1} which is impossible). Then, we see that $\frac{w^k}{\|w^k\|} \in N_{[\overline{f} \leq 0]}(x^k)$. It follows from $w^k \in L$ (and so, $a_i^T w^k = 0 \,\forall i \in I_1$), (3.15) and (3.14) that

$$\begin{aligned}\alpha^{-1} &\leq \sup\left\{\left(\nabla g(x^k) + \sum_{i \in I_1} \mu_i a_i\right)^T \left(\frac{w^k}{\|w^k\|}\right) : \mu_i \in \mathbb{R}\right\} \\ &= \|w^k\| = \min\left\{\left\|\nabla g(x^k) + \sum_{i \in I_1} \lambda_i a_i\right\| : \lambda_i \in \mathbb{R}\right\}\end{aligned}$$

This contradicts (3.12) and completes the proof. □

We have seen that the Slater condition guarantees a Lipschitz-type global error bound. In general, the Lipschitz-type global error bound might fail if the Slater condition is not satisfied. For a simple example, let $n = 1$, $g(x) = x^2$ and $P = \mathbb{R}$. Then, $f = g + \delta_P = g$. It is clear that $[f < 0] = \emptyset$ and so, the Slater condition is not satisfied. On the other hand, consider $x_k = 1/k$ $(k \in \mathbb{N})$. It is easy to verify that $\frac{d(x_k, [f \leq 0])}{[f(x_k)]_+} \to +\infty$, and so, the Lipschitz-type global error bound fails. Thus, it is natural to ask what happens if the Slater condition is not satisfied.

Below, we will provide an answer for this question. More explicitly, we will show that, a Hölder-type global error bound with Hölder exponent $\kappa(n, d)^{-1}$ holds regardless whether the Slater condition holds or not. This Hölder-type global error bound is achieved by applying the standard Lagrange multiplier technique and Lemma 3.3.

Theorem 3.1 *Let g be a convex polynomial on $\mathbb{R}^n$ with degree d and let P be a polyhedron in $\mathbb{R}^n$. Let $f = g + \delta_P$. Then, f has a Hölder-type global error bound with exponent $\kappa(n, d)^{-1}$, i.e., there exists a constant $\tau > 0$ such that*

$$d(x, [f \leq 0]) \leq \tau \left([f(x)]_+ + [f(x)]_+^{\kappa(n,d)^{-1}}\right) \quad \textit{for all } x \in \mathbb{R}^n. \tag{3.16}$$

Proof From the preceding proposition, we may assume without loss of generality that f is nonnegative. This gives us that g is nonnegative on P. Let $P = \{x : a_i^T x \leq b_i, i = 1, \ldots, m\}$. Consider the following minimization problem:

$$(P) \min g(x) \quad \text{s.t. } a_i^T x \leq b_i, \quad i = 1, \ldots, m.$$

Take $\overline{x} \in [g \leq 0] \cap P = [g = 0] \cap P$. Then, $\overline{x}$ is a global minimizer of (P). Thus, there exist $\lambda_i \geq 0,\ i = 1, \ldots, m$ such that

$$\nabla g(\overline{x}) + \sum_{i=1}^{m} \lambda_i a_i = 0 \text{ and } \lambda_i \left(a_i^T \overline{x} - b_i\right) = 0, \quad i = 1, \ldots, m.$$

Define $\phi(x) := g(x) + \sum_{i=1}^{m} \lambda_i (a_i^T x - b_i)$. Then, we see that ϕ is a convex polynomial with degree d satisfying $\nabla \phi(\overline{x}) = 0$. This implies that ϕ attains its minimum at $\overline{x}$ and $\min \phi = \phi(\overline{x}) = 0$. Thus, from Lemma 3.3, there exists $\mu > 0$ such that

$$d(x, [\phi = 0]) = d(x, [\phi \leq 0]) \leq \mu \left(\phi(x) + \phi(x)^{\kappa(n,d)^{-1}}\right) \ \forall\, x \in \mathbb{R}^n. \qquad (3.17)$$

Let $L = \{x : (\lambda_i a_i)^T x = \lambda_i b_i, i = 1, \ldots, m\}$. Next, we claim that

$$[\phi = 0] \cap P \cap L = [g \leq 0] \cap P. \qquad (3.18)$$

Granting this and noting that $[\phi = 0]$, P and L are all polyhedral sets,[2] Lemma 3.4 implies that there exists $\tau_0 > 0$ such that

$$\begin{aligned} d(x, [g \leq 0] \cap P) &= d(x, [\phi = 0] \cap P \cap L) \\ &\leq \tau_0 \big(d(x, [\phi = 0]) + d(x, P) + d(x, L)\big). \end{aligned}$$

From the Hoffman's lemma, we see that $d(x, L) \leq \alpha \sum_{i=1}^{m} |\lambda_i (a_i^T x - b_i)|$. This implies that for all $x \in P$

$$d(x, [g \leq 0] \cap P) \leq \tau_0 d(x, [\phi = 0]) + \alpha \tau_0 \sum_{i=1}^{m} \left|\lambda_i \left(a_i^T x - b_i\right)\right|.$$

Noting that $\phi(x) \leq g(x) = [g(x)]_+$ for all $x \in P$, this together with (3.17) implies that

$$\begin{aligned} d(x, [g \leq 0] \cap P) \leq \mu \tau_0 &\left([g(x)]_+ + [g(x)]_+^{\kappa(n,d)}\right) \\ &+ \alpha \tau_0 \sum_{i=1}^{m} \left|\lambda_i \left(a_i^T x - b_i\right)\right| \ \forall\, x \in P. \end{aligned}$$

Since ϕ is nonnegative, we see that for all $x \in P$

$$\sum_{i=1}^{m} \left|\lambda_i \left(a_i^T x - b_i\right)\right| = -\sum_{i=1}^{m} \lambda_i \left(a_i^T x - b_i\right) = g(x) - \phi(x) \leq g(x).$$

[2] To see $[\phi = 0]$ is a polyhedron, we assume without loss of generality that $[\phi = 0]$ is not a singleton. Noting that $[\phi = 0] = \{x : \phi(x) = \inf \phi\}$ is convex, it follows from Lemma 2.1 (2) that $[\phi = 0]$ is affine and so is a polyhedral set.

Thus, letting $\tau = \tau_0(\mu + \alpha)$, we have

$$d(x, [g \leq 0] \cap P) \leq \tau \left([g(x)]_+ + [g(x)]_+^{\kappa(n,d)^{-1}}\right) \ \forall\, x \in P,$$

and so, (3.16) holds. Finally, to see (3.18), we first note that $[g \leq 0] \cap P \subseteq [\phi \leq 0] \cap P \cap L = [\phi = 0] \cap P \cap L$ always holds. To see the reverse inclusion, take $z \in [\phi = 0] \cap P \cap L$. Then, $\lambda_i(a_i^T z - b_i) = 0$. On the other hand, since ϕ is nonnegative and $\phi(z) = 0$, we have

$$\nabla\phi(z) = \nabla g(z) + \sum_{i=1}^{m} \lambda_i a_i = 0.$$

This implies that z satisfies the KKT condition of the problem (P). Note that (P) is a convex problem with linear constraint. It follows that z is a global minimizer of (P) and so (3.18) follows. □

As a corollary, we provide the following Hölder-type error bound result on a compact set. In the special case when f is a convex quadratic function, i.e., $d = 2$ (and so, $\kappa(n, d) = 2$), this result has been presented in [28, Theorem 4.1]. It should be emphasized that [28, Theorem 4.1] established a more general result. More explicitly, they established that the conclusion is true when the Hessian of the quadratic function is copositive on a set which is a translation of the corresponding polyhedron P.

Corollary 3.1 *Let f be a convex polynomial on $\mathbb{R}^n$ with degree d and let P be a polyhedron in $\mathbb{R}^n$ with the form $P = \{x \in \mathbb{R}^n : a_i^T x \leq b_i, i = 1, \ldots, m\}$. Then, there exists a constant $\tau > 0$ such that*

$$d(x, [f \leq 0] \cap P) \leq \tau\Bigg([f(x)]_+ + [f(x)]_+^{\kappa(n,d)^{-1}} + (1 + \|\nabla f(x)\|) \sum_{i=1}^{m} [a_i^T x - b_i]_+ + \|\nabla f(x)\|^{\kappa(n,d)^{-1}} \left(\sum_{i=1}^{m} \left[a_i^T x - b_i\right]_+\right)^{\kappa(n,d)^{-1}} \Bigg) \quad \textit{for all } x \in \mathbb{R}^n. \tag{3.19}$$

In particular, for any compact set K, there exists a constant $\tau_K > 0$ such that

$$d(x, [f \leq 0] \cap P) \leq \tau_K \left([f(x)]_+^{\kappa(n,d)^{-1}} + \left(\sum_{i=1}^{m} \left[a_i^T x - b_i\right]_+\right)^{\kappa(n,d)^{-1}} \right), \quad \textit{for all } x \in K. \tag{3.20}$$

Proof From the preceding theorem and the Hoffman error bound, we can find τ_1, $\tau_2 > 0$ such that

$$d(z, [f \leq 0] \cap P) \leq \tau_1([f(z)]_+ + [f(z)]_+^{\kappa(n,d)^{-1}}) \ \forall\, z \in P \text{ and } d(x, P) \leq \tau_2 \sum_{i=1}^{m} \left[a_i^T x - b_i\right]_+ \ \forall\, x \in \mathbb{R}^n.$$

Fix an arbitrary $x \in \mathbb{R}^n$. Let $z \in P$ be such that $z = \Pr_P(x)$. Note from the triangle inequality that

$$d(x, [f \le 0] \cap P) \le d(z, [f \le 0] \cap P) + \|x - z\| = d(z, [f \le 0] \cap P) + d(x, P).$$

Thus, we have

$$d(x, [f \le 0] \cap P) \le \tau_1 \left([f(z)]_+ + [f(z)]_+^{\kappa(n,d)^{-1}} \right) + \tau_2 \sum_{i=1}^m \left[a_i^T x - b_i \right]_+ .$$

Now, by the convexity of f, we have

$$f(z) \le f(x) + \nabla f(x)^T (z - x).$$

From the following elementary inequalities:

$$\forall\, q \le 1 \text{ and } a \le b, \quad [a]_+^q \le [b]_+^q \text{ and } [a+b]_+^q \le [a]_+^q + [b]_+^q,$$

it follows that for any $q \le 1$

$$\begin{aligned}
[f(z)]_+^q \le [f(x) + \nabla f(x)^T (z - x)]_+^q &\le [f(x)]_+^q + [\nabla f(x)^T (z - x)]_+^q \\
&\le [f(x)]_+^q + \|\nabla f(x)\|^q \, \|z - x\|^q \\
&= [f(x)]_+^q + \|\nabla f(x)\|^q \, d(x, P)^q \\
&\le [f(x)]_+^q + \tau_2^q \|\nabla f(x)\|^q \left(\sum_{i=1}^m \left[a_i^T x - b_i \right]_+ \right)^q .
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
d(x, [f \le 0] \cap P) &\le \tau_1 \left([f(z)]_+ + [f(z)]_+^{\kappa(n,d)^{-1}} \right) + \tau_2 \sum_{i=1}^m \left[a_i^T x - b_i \right]_+ \\
&\le \tau_1 \Bigg([f(x)]_+ + \tau_2 \|\nabla f(x)\| \sum_{i=1}^m \left[a_i^T x - b_i \right]_+ + [f(x)]_+^{\kappa(n,d)^{-1}} \\
&\quad + \tau_2^{\kappa(n,d)^{-1}} \|\nabla f(x)\|^{\kappa(n,d)^{-1}} \left(\sum_{i=1}^m [a_i^T x - b_i]_+ \right)^{\kappa(n,d)^{-1}} \Bigg) + \tau_2 \sum_{i=1}^m \left[a_i^T x - b_i \right]_+ \\
&\le \tau \Bigg([f(x)]_+ + [f(x)]_+^{\kappa(n,d)^{-1}} + (1 + \|\nabla f(x)\|) \sum_{i=1}^m \left[a_i^T x - b_i \right]_+ \\
&\quad + \|\nabla f(x)\|^{\kappa(n,d)^{-1}} \left(\sum_{i=1}^m \left[a_i^T x - b_i \right]_+ \right)^{\kappa(n,d)^{-1}} \Bigg), \qquad (3.21)
\end{aligned}$$

where $\tau = \tau_1 \left(1 + \tau_2 + \tau_2^{\kappa(n,d)^{-1}} \right)$. Thus, (3.19) is shown. To see the last assertion, let K be a compact set. Then, there exist $M_1, M_2, M_3 > 0$ such that for all $x \in K$

$$\|\nabla f(x)\| \leq M_1,\ [f(x)]_+ \leq M_2 [f(x)]_+^{\kappa(n,d)^{-1}}$$

and

$$\sum_{i=1}^{m} \left[a_i^T x - b_i\right]_+ \leq M_3 \left(\sum_{i=1}^{m} [a_i^T x - b_i]_+\right)^{\kappa(n,d)^{-1}}.$$

Therefore, the assertion follows from (3.21). □

Let f be a convex polynomial and let P be a polyhedron defined as $P = \{x : a^T x \leq b_i,\ i = 1, \ldots, m\}$. As another corollary, we obtain a continuity result of a set-valued mapping $A : \mathbb{R} \to 2^{\mathbb{R}^n}$, defined by

$$\begin{aligned} A(\lambda) :&= [f \leq \lambda] \cap P \\ &= \{x : f(x) \leq \lambda \text{ and } a^T x \leq b_i,\ i = 1, \ldots, m\},\ \lambda \in \mathbb{R}. \end{aligned} \tag{3.22}$$

Corollary 3.2 *Consider the set-valued mapping $A : \mathbb{R} \to 2^{\mathbb{R}^n}$ defined as in (3.22). Then, the mapping A is Hausdorff upper semicontinuous at 0 in the sense that*

$$\lim_{\lambda \to 0^+} \sup_{x \in A(\lambda)} d(x, A(0)) = 0.$$

Proof Let $\overline{f} = f + \delta_P$. It follows that $A(0) = [\overline{f} \leq 0]$, and $\overline{f}$ agrees f on P. Applying Theorem 3.1 to $\overline{f}$, this implies that there exists $\tau > 0$ such that

$$\begin{aligned} d(x, A(0)) = d(x, [\overline{f} \leq 0]) &\leq \tau \left([\overline{f}(x)]_+ + [\overline{f}(x)]_+^{\kappa(n,d)^{-1}}\right) \\ &= \tau \left([f(x)]_+ + [f(x)]_+^{\kappa(n,d)^{-1}}\right) \quad \text{for any } x \in P. \end{aligned} \tag{3.23}$$

Let $\{\lambda^k\}$ be an arbitrary sequence such that $\lambda^k \geq 0$ and $\lambda^k \to 0$ and fix any $\{x^k\} \subseteq A(\lambda^k)$. Noting that $A(\lambda^k) \subseteq P$, (3.23) implies that

$$d(x^k, A(0)) \leq \tau \left([f(x^k)]_+ + [f(x^k)]_+^{\kappa(n,d)^{-1}}\right) \leq \tau \left(\lambda^k + (\lambda^k)^{\kappa(n,d)^{-1}}\right) \to 0.$$

Thus, the conclusion follows.

The study of the continuity of a set-valued mapping is an important topic of set-valued analysis and variational analysis [6,37,38]. It has attracted a lot of researchers and has found many important applications. In particular, it is well known that a group of researchers from the Moscow University has constructed subtle examples (see [1,5,6,38]) of set-valued mapping M with the form

$$\begin{aligned} \lambda = (\lambda_0, \lambda_1 \ldots, \lambda_m) \in \mathbb{R}^{m+1} &\mapsto M(\lambda) \\ &:= \left\{x \in \mathbb{R}^n : f(x) \leq \lambda_0 \text{ and } a_i^T x \leq b_i + \lambda_i,\ i = 1, \ldots, m\right\}, \end{aligned}$$

where $f : \mathbb{R}^n \mapsto \mathbb{R}$ is a convex polynomial, $a_i \in \mathbb{R}^n$ and $b_i \in \mathbb{R}$, such that the set-valued mapping M is not Hausdorff upper semicontinuous at 0. It is worthy noting that all examples given in [1,2,5,6,38] work with varying $\lambda_1, \ldots, \lambda_m$, while our result in Corollary 3.2 works with a fixed polyhedral set P, that is, $\lambda_i \equiv 0,\ i = 1, \ldots, m$ (see (3.22)). Below, we present an example illustrating this subtle difference. More explicitly, making use of an example presented in [1,2], we construct two set-valued mapping A_1 and A_2 with the following form

$$\lambda = (\lambda_0, \lambda_1, \lambda_2) \in \mathbb{R}^3 \mapsto A_1(\lambda) := \left\{x \in \mathbb{R}^4 : f(x) \leq \lambda_0,\ a_i^T x \leq \lambda_i,\ i = 1, 2\right\}$$

and

$$\lambda \in \mathbb{R} \mapsto A_2(\lambda) := \left\{x \in \mathbb{R}^4 : f(x) \leq \lambda,\ a_i^T x \leq 0,\ i = 1, 2\right\},$$

where f is a convex polynomial, and show that A_1 is not Hausdorff upper semicontinuous at 0 while A_2 is Hausdorff upper semicontinuous at 0.

Example 3.1 We begin by recalling an example which was presented in [1,2]. Let $F : \mathbb{R}^3 \to \mathbb{R}$ be defined by

$$\begin{aligned} F(x_0, x_1, x_2) = \mu_0 x_0^8 x_1^{16} x_2 + \mu_1 x_0^{10} x_1^{12} x_2^2 + \mu_2 x_1^{32} x_2^{12} + \mu_3 x_2^{204} \\ + \mu_4 x_1^{34} + \mu_5 x_0^{12} x_1^8 + \mu_6 x_0^6 x_1^{20} + \mu_7 x_0^{16} + \mu_8 x_1^{32}, \end{aligned}$$

where $\mu_0 < 0$ and $\mu_i \in \mathbb{R},\ i = 1, \ldots, 8$ are positive such that F is convex.[3] Let $f_0(x_0, x_1, x_2, x_3) = F(x_0, x_1, x_2) - x_3$ and $g(x_0, x_1, x_2, x_3) = x_2^2$. It is clear that f_0 and g are both convex polynomials. Consider the associated set-valued mapping $M : \mathbb{R}^2 \to 2^{\mathbb{R}^4}$, defined by

$$\begin{aligned} M(\lambda) := \{x = (x_0, x_1, x_2, x_3) \in \mathbb{R}^4 : f_0(x) \leq \lambda_1 \\ \text{and } g(x) \leq \lambda_2\}, \quad \lambda = (\lambda_1, \lambda_2) \in \mathbb{R}^2. \end{aligned}$$

In [1,2] (see also [5, Example 6]), they showed that the set-valued mapping M is not Hausdorff upper-semicontinuous at 0[4] by constructing $\gamma > 0,\ \mu^k = (\mu_1^k, \mu_2^k) \to 0$ and $a^k \in M(\mu^k)$ such that

$$d(a^k, M(0)) \to \gamma. \tag{3.24}$$

[Construction of the set-valued mapping]
Now, define two closed related set-valued mapping $A_1 : \mathbb{R}^3 \to 2^{\mathbb{R}^4}$ and $A_2 : \mathbb{R} \to 2^{\mathbb{R}^4}$

[3] Possible choice of $\mu_0, \ldots, \mu_8$ is examined and explicitly given in [6] (see also [1,2,5]).

[4] Indeed, they have shown a stronger result, namely,

$$\{b : \exists \lambda^k \neq 0, \lambda^k \to 0 \text{ such that } \sup_{x \in M(\lambda^k)} d(x, M(0)) \to b\} = \{0, \gamma\}.$$

as follows

$$A_1(\lambda) := \{x = (x_0, x_1, x_2, x_3) \in \mathbb{R}^4 : f_0(x) \le \lambda_1, x_2 \le \lambda_2 \text{ and } -x_2 \le \lambda_3\}, \quad \lambda = (\lambda_1, \lambda_2, \lambda_3) \in \mathbb{R}^3$$

and

$$A_2(\lambda) := \{x = (x_0, x_1, x_2, x_3) \in \mathbb{R}^4 : f_0(x) \le \lambda, x_2 \le 0 \text{ and } -x_2 \le 0\}, \quad \lambda \in \mathbb{R}.$$

[Verifying that A_1 is not Hausdorff upper-semicontinuous at 0]
Recall that $\mu^k = \left(\mu_1^k, \mu_2^k\right) \to 0$ and $a^k \in M(\mu^k)$. This implies that $a^k \in M(\mu^k) \subseteq A_1\left(\mu_1^k, \sqrt{\mu_2^k}, \sqrt{\mu_2^k}\right)$ and $\left(\mu_1^k, \sqrt{\mu_2^k}, \sqrt{\mu_2^k}\right) \to 0$. Noting that $A_1(0) = M(0)$, it follows from (3.24) that

$$d(a^k, A_1(0)) \to \gamma > 0.$$

This shows that A_1 is not Hausdorff upper-semicontinuous at 0.
[Verifying the Hausdorff upper semicontinuity of A_2 at 0]

Now, consider the polyhedron $P = \{(x_0, x_1, x_2, x_3) \in \mathbb{R}^4 : x_2 = 0\}$ and let $f = f_0 + \delta_P$. We first verify that f has a global error bound. To see this, we first observe that

$$[f \le 0] = [f_0 \le 0] \cap P = \{(x_0, x_1, x_2, x_3) : x_2 = 0 \text{ and } \mu_4 x_1^{34} + \mu_5 x_0^{12} x_1^8 + \mu_6 x_0^6 x_1^{20} + \mu_7 x_0^{16} + \mu_8 x_1^{32} \le x_3\}.$$

Define a new polynomial h as follows

$$h(a_0, a_1, a_2) = \mu_4 a_1^{34} + \mu_5 a_0^{12} a_1^8 + \mu_6 a_0^6 a_1^{20} + \mu_7 a_0^{16} + \mu_8 a_1^{32} - a_2.$$

Clearly, h is convex (as $h(a_0, a_1, a_2) = f_0(a_0, a_1, 0, a_2)$) and

$$h^\infty(0, 0, 1) = \sup_{t>0} \frac{h(a + t(0, 0, 1)) - h(a)}{t} = -1 < 0 \quad \text{for all } a \in \mathbb{R}^3.$$

Thus, Lemma 3.2 (2) guarantees that there exists $\tau > 0$ such that

$$d(a, [h \le 0]) \le \tau [h(a)]_+ \quad \text{for all } a \in \mathbb{R}^3.$$

Moreover, from the definition of h, we see that

$$[f \le 0] = \{(x_0, x_1, x_2, x_3) : x_2 = 0 \text{ and } h(x_0, x_1, x_3) \le 0\}.$$

For any $x = (x_0, x_1, x_2, x_3) \in P$, we have $x_2 = 0$ and so,

$$d(x, [f \le 0]) = d((x_0, x_1, x_3), [h \le 0]) \\ \le \tau[h(x_0, x_1, x_3)]_+ = \tau[f_0(x_0, x_1, 0, x_3)]_+.$$

In other words, for any $x = (x_0, x_1, x_2, x_3) \in P$, $d(x, [f \le 0]) \le \tau[f(x)]_+$. As f takes the value $+\infty$ outside the polyhedron P, it follows that

$$d(x, [f \le 0]) \le \tau[f(x)]_+ \quad \text{for all } x \in \mathbb{R}^n.$$

Thus, f has a global error bound. Now, let $\{\lambda^k\}$ be an arbitrary sequence such that $\lambda^k \ge 0$ and $\lambda^k \to 0$, and fix any $x^k \in A_2(\lambda^k) \subseteq P$. As f has a global error bound, we have

$$d(x^k, A_2(0)) = d(x^k, [f \le 0]) \le \tau[f(x)]_+ \le \tau\lambda^k \to 0.$$

This shows that $\lim_{\lambda \to 0^+} \sup_{x \in A_2(\lambda)} d(x, A_2(0)) = 0$, that is, A_2 is Hausdorff upper semicontinuous at 0.

3.2 Global error bound for piecewise convex polynomial

In this subsection, we provide a necessary and sufficient condition characterizing when a piecewise convex polynomial has a Hölder-type global error bound with an explicit Hölder exponent. To begin with, as a simple application of Lemma 2.2, we first provide a Frank-Wolfe type result for a piecewise convex polynomial.

Lemma 3.6 *Let f be a piecewise convex polynomial on $\mathbb{R}^n$. Suppose that* $\inf f > -\infty$. *Then, we have* $\operatorname{argmin} f \neq \emptyset$.

Proof Since f is a piecewise convex polynomial, there exist finitely many polyhedra $P_1, \ldots, P_k$ with $\bigcup_{j=1}^k P_j = \mathbb{R}^n$ such that the restriction of f on each P_j is a convex polynomial. As $\inf f > -\infty$, we see that $\inf_{x \in P_j} f(x) > -\infty$, $j = 1, \ldots, k$. Thus, by Lemma 2.2, we have $\operatorname{argmin}_{x \in P_j} f(x) \neq \emptyset$. Note that $\inf f = \min_{1 \le j \le k} \min_{x \in P_j} f(x)$. Therefore, we see that $\operatorname{argmin} f \neq \emptyset$. □

Now, we establish a Hölder-type local error bound result for piecewise convex polynomials. Later, we will use this result to examine a sufficient and necessary condition for the Hölder-type global error bound results of piecewise convex polynomials.

Proposition 3.2 *Let f be a piecewise convex polynomial on $\mathbb{R}^n$ with degree d. Then, f has a Hölder-type local error bound with exponent $\kappa(n, d)^{-1}$, i.e., there exist $r > 0$ and $\tau > 0$ such that*

$$d(x, [f \le 0]) \le \tau\left([f(x)]_+ + [f(x)]_+^{\kappa(n,d)^{-1}}\right) \quad \text{for all } x \in [f \le r]. \tag{3.25}$$

Proof Since f is a piecewise convex polynomial on $\mathbb{R}^n$ with degree d, there exist finitely many polyhedra $P_1, \ldots, P_k$ with $\bigcup_{j=1}^k P_j = \mathbb{R}^n$ such that the restriction of f on each P_j is a convex polynomial with degree at most d. Let $J = \{1, \ldots, k\}$,

$$J_1 := \{j \in J : P_j \cap [f \leq 0] \neq \emptyset\} \quad \text{and} \quad J_2 := \{j \in J : P_j \cap [f \leq 0] = \emptyset\}.$$

Note that for each $j \in J_2$, $f(x) > 0$ for all $x \in P_j$. Let $P_j = \{x : (a_j^l)^T x \leq b_j^l, l = 1, \ldots, m_j\}$ where $m_j \in \mathbb{N}$ and $j \in J_2$. It follows from Lemma 2.2 that $\inf_{P_j} f$ is attained, and so, $\inf_{x \in P_j} f(x) > 0$. Thus, we have

$$\lambda := \inf \left\{ f(x) : x \in \bigcup_{j \in J_2} P_j \right\} > 0.$$

Now, let $r = \lambda/2 > 0$. Then, by the construction of r, we have

$$[f \leq r] \subseteq \bigcup \{P_j : j \in J_1\}. \tag{3.26}$$

On the other hand, let f_j $(j \in J)$ be convex polynomials on $\mathbb{R}^n$ such that it agrees f on P_j, i.e., $f_j + \delta_{P_j} = f + \delta_{P_j}$. Applying Theorem 3.1 to $h_j := f_j + \delta_{P_j}$, we see that, for each $j \in J_1$, there exists $\tau_j > 0$ such that for all $x \in P_j$,

$$\begin{aligned} d(x, [f \leq 0] \cap P_j) &= d(x, [h_j \leq 0]) \\ &\leq \tau_j \left([h_j(x)]_+ + [h_j(x)]_+^{\kappa(n,d)^{-1}} \right) \\ &= \tau_j \left([f(x)]_+ + [f(x)]_+^{\kappa(n,d)^{-1}} \right). \end{aligned}$$

Let $\tau := \max\{\tau_j : j \in J_1\}$. From $\bigcup_{j \in J_1} \left([f \leq 0] \cap P_j\right) = [f \leq 0]$, we have

$$d(x, [f \leq 0]) \leq \tau([f(x)]_+ + [f(x)]_+^{\kappa(n,d)^{-1}}) \quad \text{for all } x \in \bigcup \{P_j : j \in J_1\}.$$

Therefore, the conclusion follows by (3.26). □

As a corollary, we obtain the following error bound type result near the set $[f \leq 0]$.

Corollary 3.3 *Let f be a piecewise convex polynomial on $\mathbb{R}^n$ with degree d. Then, for any $\delta > 0$, there exists $\tau > 0$ such that*

$$d(x, [f \leq 0]) \leq \tau \left([f(x)]_+ + [f(x)]_+^{\kappa(n,d)^{-1}} \right) \quad \textit{for all } x \textit{ with } d(x, [f \leq 0]) \leq \delta. \tag{3.27}$$

Proof From Proposition 3.2, we can find $r > 0$ and $\tau_1 > 0$ such that

$$d(x, [f \leq 0]) \leq \tau_1 \left([f(x)]_+ + [f(x)]_+^{\kappa(n,d)^{-1}} \right) \quad \text{for all } x \in [f \leq r].$$

If $x \in [f \leq r]$ and $d(x, [f \leq 0]) \leq \delta$, then we have

$$d(x, [f \leq 0]) \leq \tau \left([f(x)]_+ + [f(x)]_+^{\kappa(n,d)^{-1}} \right).$$

On the other hand, suppose that $x \in [f > r]$ and $d(x, [f \le 0]) \le \delta$. Then, we obtain that

$$d(x, [f \le 0]) \le \delta \le \frac{\delta}{r + r^{\kappa(n,d)^{-1}}} \left([f(x)]_+ + [f(x)]_+^{\kappa(n,d)^{-1}} \right).$$

Therefore, the conclusion follows with $\tau = \tau_1 + \frac{\delta}{r + r^{\kappa(n,d)^{-1}}}$. □

As another corollary, we obtain the following result estimating the distance between a point to the solution set $\operatorname{argmin} f$. In the special case when $d = 2$ (and so, $\kappa(n, d) = 2$), this result has been presented in [26].

Corollary 3.4 *Let f be a piecewise convex polynomial on $\mathbb{R}^n$ with degree d. Then, there exist $r > \inf f$ and $\tau > 0$ such that*

$$d(x, \operatorname{argmin} f) \le \tau \left((f(x) - \inf f) + (f(x) - \inf f)^{\kappa(n,d)^{-1}} \right) \quad \textit{for all } x \in [f \le r]. \tag{3.28}$$

Proof Without loss of generality, we may assume that $\inf f > -\infty$ (Otherwise, the right hand side of (3.28) equals $+\infty$ and the conclusion follows). Then Lemma 3.6 implies that $\operatorname{argmin} f \neq \emptyset$. Now, applying Proposition 3.2 to $g := f - \inf f$, we can find $\delta > 0$ and $\tau > 0$ such that

$$\begin{aligned} d(x, \operatorname{argmin} f) = d(x, [g \le 0]) &\le \tau \left([g(x)]_+ + [g(x)]_+^{\kappa(n,d)^{-1}} \right) \\ &= \tau \left((f(x) - \inf f) + (f(x) - \inf f)^{\kappa(n,d)^{-1}} \right) \quad \text{for all } x \in [g \le \delta]. \end{aligned}$$

Letting $r = \inf f + \delta$, then the conclusion follows as $[g \le \delta] = [f \le r]$. □

The preceding proposition (Proposition 3.2) illustrates that Hölder-type *local* error bound holds for piecewise convex polynomials. However, in general, the following one dimensional example shows that a *global* error bound might fail.

Example 3.2 Consider the following piecewise convex polynomial $f : \mathbb{R} \to \mathbb{R}$, defined by

$$f(x) = \begin{cases} 1, & \text{if } x \ge 1, \\ x^4, & \text{if } x < 1. \end{cases}$$

Clearly, $[f \le 0] = \{0\}$. Consider $x_k = k$. Then $d(x_k, [f \le 0]) = k$ but $f(x_k) = 1$. Therefore, there is no global error bound in this case. Notably, in this example, the following implication fails

$$d(x, [f \le 0]) \to \infty \ \Rightarrow \ f(x) \to +\infty.$$

The above example prompt us to examine a condition ensuring a global error bound for piecewise convex polynomial. Below, we state our main result in this subsection, i.e. a necessary and sufficient condition characterizing when a piecewise convex polynomial has a Hölder-type global error bound with an explicit Hölder exponent.

Theorem 3.2 *Let f be a piecewise convex polynomial on $\mathbb{R}^n$ with degree d. Then, the following statements are equivalent:*

(1) $d(x,[f \le 0]) \to +\infty \;\Rightarrow\; f(x) \to +\infty$.
(2) *There exists $\tau > 0$ such that*

$$d(x,[f \le 0]) \le \tau \left([f(x)]_+ + [f(x)]_+^{\kappa(n,d)^{-1}}\right) \quad \text{for all } x \in \mathbb{R}^n. \qquad (3.29)$$

Proof [(2) $\Rightarrow$ (1)] This part is direct.

[(1) $\Rightarrow$ (2)] Since f is a piecewise convex polynomial on $\mathbb{R}^n$ with degree d, there exist finitely many polyhedra $P_1, \ldots, P_k$ with $\bigcup_{j=1}^k P_j = \mathbb{R}^n$ such that the restriction of f on each P_j is a convex polynomial with degree at most d. Let $J = \{1, \ldots, k\}$ and

$$J_\infty = \left\{ j \in J : \sup_{x \in P_j} f(x) = +\infty \right\}.$$

Take $\lambda > 0$ large enough such that

$$\sup \left\{ f(x) : x \in \bigcup_{j \in J \setminus J_\infty} P_j \right\} \le \lambda \text{ and } \inf\{f(x) : x \in P_j\} < \lambda \quad \text{for all } j \in J_\infty.$$

By the construction of λ, $P_j \cap [f \le \lambda] \neq \emptyset$ for each $j \in J_\infty$. Now, let f_j $(j \in J)$ be convex polynomials on $\mathbb{R}^n$ such that they agree f on P_j, i.e., $f_j + \delta_{P_j} = f + \delta_{P_j}$. For each $j \in J_\infty$, applying Theorem 3.1 to $h_j := f_j - \lambda + \delta_{P_j}$, we can find $\gamma_j > 0$ such that, for all $x \in P_j$,

$$\begin{aligned} d(x, P_j \cap [f \le \lambda]) = d(x,[h_j \le 0]) &\le \gamma_j \left([h_j(x)]_+ + [h_j(x)]_+^{\kappa(n,d)^{-1}}\right) \\ &= \gamma_j \left([f(x)-\lambda]_+ + [f(x)-\lambda]_+^{\kappa(n,d)^{-1}}\right). \end{aligned}$$

Letting $\tau_1 = \max\{\gamma_j : j \in J_\infty\}$, it follows that

$$\begin{aligned} d(x,[f \le \lambda]) \le d(x, P_j \cap [f \le \lambda]) \le \tau_1([f(x)-\lambda]_+ \\ +[f(x)-\lambda]_+^{\kappa(n,d)^{-1}}) \quad \text{for all } x \in \bigcup_{j \in J_\infty} P_j. \end{aligned} \qquad (3.30)$$

On the other hand, from our assumption, we can find $r > 0$ such that

$$d(x,[f \le 0]) > r \;\Rightarrow\; f(x) > \lambda.$$

This implies that $[f \le \lambda] \subseteq \{x : d(x, [f \le 0]) \le r\}$. Now, by Corollary 3.4, we can find $\tau_2 > 0$ such that

$$d(x, [f \le 0]) \le \tau_2 \left([f(x)]_+ + [f(x)]_+^{\kappa(n,d)^{-1}}\right) \quad \text{for all } x \text{ with } d(x, [f \le 0]) \le r.$$

This implies that

$$d(x, [f \le 0]) \le \tau_2 \left([f(x)]_+ + [f(x)]_+^{\kappa(n,d)^{-1}}\right) \quad \text{for all } x \in [f \le \lambda]. \quad (3.31)$$

Now, we show that, for all $x \in \mathbb{R}^n$,

$$d(x, [f \le 0]) \le 2^{1-\kappa(n,d)^{-1}}(\tau_1 + \tau_2)\left(f(x) + f(x)^{\kappa(n,d)^{-1}}\right). \quad (3.32)$$

Clearly, if $x \in [f \le \lambda]$, then (3.32) holds. So, it remains to consider the case when $x \in [f > \lambda]$. Take an arbitrary $x_0 \in [f > \lambda]$. From our construction of λ, we have $x_0 \in \bigcup_{j \in J_\infty} P_j$. Let $x_1 \in [f \le \lambda]$ be such that $\|x_0 - x_1\| = d(x_0, [f \le \lambda])$. Then, we see that $f(x_1) = \lambda > 0$ and (by (3.30) and $x_0 \in \bigcup_{j \in J_\infty} P_j$)

$$\begin{aligned} \|x_0 - x_1\| = d(x_0, [f \le \lambda]) &\le \tau_1 \left([f(x_0) - \lambda]_+ + [f(x_0) - \lambda]_+^{\kappa(n,d)^{-1}}\right) \\ &= \tau_1 \left(\big(f(x_0) - f(x_1)\big) + \big(f(x_0) - f(x_1)\big)^{\kappa(n,d)^{-1}}\right). \end{aligned}$$

Now, by (3.31), we have

$$d(x_1, [f \le 0]) \le \tau_2 \left([f(x_1)]_+ + [f(x_1)]_+^{\kappa(n,d)^{-1}}\right) = \tau_2 \left(f(x_1) + f(x_1)^{\kappa(n,d)^{-1}}\right).$$

Thus, the triangle inequality implies that

$$\begin{aligned} d(x_0, [f \le 0]) &\le d(x_1, [f \le 0]) + \|x_0 - x_1\| \\ &\le (\tau_1 + \tau_2)\Bigg(\big(f(x_0) - f(x_1)\big) + \big(f(x_0) - f(x_1)\big)^{\kappa(n,d)^{-1}} \\ &\quad + f(x_1) + f(x_1)^{\kappa(n,d)^{-1}}\Bigg) \\ &= (\tau_1 + \tau_2)\Bigg(f(x_0) + \big(f(x_0) - f(x_1)\big)^{\kappa(n,d)^{-1}} + f(x_1)^{\kappa(n,d)^{-1}}\Bigg). \end{aligned}$$

From the concavity of x^q $(q \le 1)$, we have $(\frac{a+b}{2})^q \ge \frac{1}{2}(a^q + b^q)$ for all $a, b > 0$, and so, $(a+b)^q \ge 2^{q-1}(a^q + b^q)$ for all $a, b > 0$. This implies that

$$\big(f(x_0) - f(x_1)\big)^{\kappa(n,d)^{-1}} + f(x_1)^{\kappa(n,d)^{-1}} \le 2^{1-\kappa(n,d)^{-1}} f(x_0)^{\kappa(n,d)^{-1}}.$$

It follows that $d(x_0, [f \le 0]) \le 2^{1-\kappa(n,d)^{-1}}(\tau_1 + \tau_2)\big(f(x_0) + f(x_0)^{\kappa(n,d)^{-1}}\big)$ and so, (3.32) holds.

Next, we provide two verifiable sufficient conditions ensuring the implication "$d(x, [f \le 0]) \to \infty \;\Rightarrow\; f(x) \to +\infty$", and so, we obtain Hölder-type global error bound results under these two sufficient conditions. In the special case when $d = 2$ (and so, $\kappa(n, d) = 2$), part (2) of this result has been presented in [26].[5]

Corollary 3.5 *Let f be a piecewise convex polynomial on $\mathbb{R}^n$ with degree d. Suppose that one of the following two condition holds:*

(1) *f is coercive in the sense that $\|x\| \to +\infty \;\Rightarrow\; f(x) \to +\infty$;*
(2) *f is convex.*

Then, there exists $\tau > 0$ such that

$$d(x, [f \le 0]) \le \tau \left([f(x)]_+ + [f(x)]_+^{\kappa(n,d)^{-1}} \right) \quad \text{for all } x \in \mathbb{R}^n. \tag{3.33}$$

Proof To see the conclusion, from Theorem 3.2, we only need to show

$$d(x, [f \le 0]) \to \infty \;\Rightarrow\; f(x) \to +\infty. \tag{3.34}$$

[Proof of (1)] To see (3.34), suppose that f is coercive and $d(x, [f \le 0]) \to \infty$. Since $d(x, [f \le 0]) \to \infty$, we have $\|x\| \to +\infty$. Thus, the coercive assumption of f implies that $f(x) \to +\infty$.

[Proof of (2)] To see (3.34), we proceed by the method of contradiction and suppose that there exist $\{x_k\} \subseteq \mathbb{R}^n$ and $M > 0$ such that

$$d(x_k, [f \le 0]) \to \infty \text{ and } 0 < f(x_k) \le M.$$

Let $y_k \in [f \le 0]$ be such that $\|x_k - y_k\| = d(x_k, [f \le 0])$. Then $f(y_k) = 0$. Clearly, we have

$$\frac{f(x_k) - f(y_k)}{\|x_k - y_k\|} \to 0.$$

Since $\|x_k - y_k\| \to \infty$, we may assume that $\|x_k - y_k\| > \delta$ for all $k \ge 1$. Let $\lambda_k = \frac{\delta}{\|x_k - y_k\|} \in (0, 1)$ and $z_k = y_k + \lambda_k(x_k - y_k)$. Then, for all $x \in [f \le 0]$, we have

$$\|x_k - y_k\| \le \|x_k - x\| \le \|x_k - z_k\| + \|z_k - x\| = (1 - \lambda_k)\|x_k - y_k\| + \|z_k - x\|.$$

This implies that $\|z_k - x\| \ge \lambda_k \|x_k - y_k\| = \|z_k - y_k\|$ for all $x \in [f \le 0]$. This implies that $d(z_k, [f \le 0]) = \|z_k - y_k\| = \lambda_k \|x_k - y_k\| = \delta$. On the other hand, by the convexity of f, we see that

$$\begin{aligned} f(z_k) &\le \lambda_k f(x_k) + (1 - \lambda_k) f(y_k) \\ &= f(y_k) + \lambda_k (f(x_k) - f(y_k)) = \delta \frac{f(x_k) - f(y_k)}{\|x_k - y_k\|} \to 0. \end{aligned}$$

[5] In [26], Li denoted all piecewise convex quadratic function which is also itself a convex function by convex piecewise quadratic function and showed that Hölder-type global error bound results holds for convex piecewise quadratic function with Holder exponent $1/2$.

This implies that $d(z_k, [f \le 0]) = \delta > 0$ and $[f(z_k)]_+ \to 0$ which contradicts Corollary 3.2. Thus, the conclusion follows.

Corollary 3.6 *Let f be a piecewise convex polynomial on $\mathbb{R}^n$ with degree d. Suppose that f is convex. Then, for any $r \geq \inf f$, there exists $\tau > 0$ such that*

$$d(x, \operatorname{argmin} f) \le \tau (f(x) - \inf f)^{\kappa(n,d)^{-1}} \text{ for all } x \in [f \le r]. \tag{3.35}$$

Proof First of all, we may assume that $\inf f > -\infty$. Then, from Lemma 3.6, $\operatorname{argmin} f \neq \emptyset$. Applying part (2) of the preceding Corollary with $g = f - \inf$, we can find $\tau > 0$ such that for any $x \in \mathbb{R}^n$

$$d(x, \operatorname{argmin} f) \le \tau \left((f(x) - \inf f) + (f(x) - \inf f)^{\kappa(n,d)^{-1}} \right).$$

Fix an arbitrary $r \ge \inf f$. Note that, for any $x \in [f \le r]$, $f(x) - \inf f \le M(f(x) - \inf f)^{\kappa(n,d)^{-1}}$ where $M = (r - \inf f)^{1-\kappa(n,d)^{-1}}$. Thus, the conclusion follows.

4 Conclusions and remarks

By examining the recession properties of convex polynomials, we provide a necessary and sufficient condition for a piecewise convex polynomial to have a Hölder-type global error bound with an explicit Hölder exponent. Our result extends the corresponding results of [26] from piecewise convex quadratic functions to piecewise convex polynomials.

As the error bound theory often has important implication in the sensitivity analysis and error estimation for optimization methods, it would be interesting to investigate whether our new error bound results will give effective global error estimates for some particular methods in solving a convex optimization problem, such as the proximal point method and the bundle method. These will be our further research directions and will be examined in a forthcoming paper.

Acknowledgments The author would like to express his sincere thanks to the referees and the associate editor for their constructive comments and valuable suggestions which have contributed to the preparation of this paper. Moreover, the author also would like to thank Professor L.Q. Qi, Professor K.F. Ng, Professor V. Jeyakumar and Professor W.H. Yang for many helpful suggestions and stimulated discussions.

References

1. Andronov, V.G., Belousov, E.G.: Types of upper discontinuities of convex polynomial mappings. Comput. Math. Math. Phys. **41**, 943–959 (2001)
2. Andronov, V.G., Belousov, E.G.: On a type of Hausdorff discontinuity of convex polynomial mappings. Comput. Math. Math. Phys. **42**, 1086–1094 (2002)
3. Auslender, A., Teboulle, M.: Asymptotic Cones and Functions in Optimization and Variational Inequalities. Springer Monographs in Mathematics. Springer, New York (2003)
4. Bauschke, H.H., Borwein, J.M.: On projection algorithms for solving convex feasibility problems. SIAM Rev. **38**(3), 367–426 (1996)

5. Belousov, E.G.: On types of Hausdorff discontinuity from above for convex closed mappings. Optimization **49**, 303–325 (2001)
6. Belousov, E.G., Andronov, V.G.: Solvability and stability of problems of polynomial programming (in Russian). Moscow University Publisher, Moscow (1993)
7. Belousov, E.G., Klatte, D.: A Frank–Wolfe type theorem for convex polynomial programs. Comput. Optim. Appl. **22**, 37–48 (2002)
8. Burke, J.V., Deng, S.: Weak sharp minima revisited part I: basic theory. Control Cybern. **31**, 439–469 (2002)
9. Burke, J.V., Deng, S.: Weak sharp minima revisited, II: application to linear regularity and error bounds. Math. Program. Ser. B **104**(2–3), 235–261 (2005)
10. Burke, J.V., Deng, S.: Weak sharp minima revisited, III: error bounds for differentiable convex inclusions. Math. Program. Ser. B **116**(1–2), 37–56 (2009)
11. Deng, S.: Computable error bounds for convex inequality systems in reflexive Banach spaces. SIAM. J. Optim. **7**, 274–279 (1997)
12. Deng, S.: Pertubation analysis of a condition number for convex inequality systems and global error bounds for analytic systems. Math. Program. **83**, 263–276 (1998)
13. Deng, S.: Lausanne 1997. Well-posed problems and error bounds in optimization. Reformulation: nonsmooth, piecewise smooth, semismooth and smoothing methods. Appl. Optim. **22**, 117–126 (1999)
14. He, Y.R., Sun, J.: Second-order sufficient conditions for error bounds in Banach spaces. SIAM J. Optim. **17**(3), 795–805 (2006)
15. Hiriart-Urruty, J.B., Lémarachal, C.: Convex analysis and minimization algorithms. II. Advanced theory and bundle methods. Grundlehren der Mathematischen Wissenschaften Fundamental Principles of Mathematical Sciences, 306, pp. xviii+346. Springer, Berlin (1993)
16. Jeyakumar, V., Lee, G.M., Li, G.Y.: Alternative theorems for quadratic inequality systems and global quadratic optimization. SIAM J. Optim. **20**(2), 983–1001 (2009)
17. Jeyakumar, V., Li, G.Y.: Necessary global optimality conditions for nonlinear programming problems with polynomial constraints. Math. Program. Ser. A **126**(2), 393–399 (2011)
18. Klatte, D.: Hoffman's error bound for systems of convex inequalities. Mathematical programming with data perturbations. Lecture Notes in Pure and Appl. Math., vol. 195, pp. 185–199. Dekker, New York (1998)
19. Klatte, D., Li, W.: Asymptotic constraint qualifications and global error bounds for convex inequalities. Math. Program. **84**, 137–160 (1999)
20. Lewis, A.S., Pang, J.S.: Error bounds for convex inequality systems generalized convexity, generalized monotonicity. In: Crouzeix, J.P., Martinez-Legaz, J.E., Volle, M. (eds.) pp. 75–110 (1998)
21. Li, G.Y., Jeyakumar, V.: Qualification-free optimality conditions for convex programs with separable inequality constraints. J. Convex Anal. **16**(3–4), 845–856 (2009)
22. Li, G.: On the asymptotic well behaved functions and global error bound for convex polynomials. SIAM J. Optim. **20**(4), 1923–1943 (2010)
23. Li, G., Ng, K.F.: On extension of Fenchel duality and its application. SIAM J. Optim. **19**, 1489–1509 (2008)
24. Li, G., Ng, K.F.: Error bounds of generalized D-gap functions for nonsmooth and nonmonotone variational inequality problems. SIAM J. Optim. **20**(2), 667–690 (2009)
25. Li, G., Tang, C.M., Wei, Z.X.: Error bound results for generalized D-gap functions of nonsmooth variational inequality problems. J. Comput. Appl. Math. **233**, 2795–2806 (2010)
26. Li, W.: Error bounds for piecewise convex quadratic programs and applications. SIAM J. Control Optim. **33**(5), 1510–1529 (1995)
27. Luo, X.D., Luo, Z.Q.: Extension of Hoffman's error bound to polynomial systems. SIAM J. Optim. **4**(2), 383–392 (1994)
28. Luo, Z.Q., Pang, J.S.: Error bounds for analytic systems and their applications. Math. Programm. **67**, 1–28 (1994)
29. Mangasarian, O.L.: A condition number for differentiable convex inequalities. Math. Oper. Res. **10**(2), 175–179 (1985)
30. Ng, K.F., Zheng, X.Y.: Error bounds for lower semicontinuous functions in normed spaces. SIAM J. Optim. **12**(1), 117 (2001)
31. Ng, K.F., Zheng, X.Y.: Global error bounds with fractional exponents. Error bounds in mathematical programming. Math. Program. Ser. B **88**(2), 357–370 (2000)

32. Ngai, H.V., Théra, M.: Error bounds for convex differentiable inequality systems in Banach spaces. Math. Program. Ser. B **104**(2–3), 465–482 (2005)
33. Obuchowska, W.T.: On generalizations of the Frank–Wolfe theorem to convex and quasi-convex programmes. Comput. Optim. Appl. **33**, 349–364 (2006)
34. Pang, J.S.: Error bounds in mathematical programming. Math. Program. **79**(2), 299–332 (1997)
35. Robinson, S.M.: An application of error bounds for convex programming in a linear space. SIAM J. Control **13**, 271–273 (1975)
36. Rockafellar, R.T.: Convex Analysis. Princeton University Press, Princeton, N.J (1970)
37. Rockafellar, R.T., Wets, R.: Variational analysis, Grundlehren der Mathematischen Wissenschaften (Fundamental Principles of Mathematical Sciences), vol. 317. Springer, Berlin (1998)
38. Shironin, V.M.: On Hausdorff continuity of convex and convex polynomial mappings (in Russian). In: Belousov, E.G., Bank, B. (eds.) Mathematical Optimization: Questions of Solvability and Stability, Chapter V. Moscow Univeristy Publ., Moscow (1986)
39. Tikhomirov, V.M.: Analysis. II. Convex analysis and approximation theory. In: Gamkrelidze, R.V. (ed.) A translation of Sovremennye problemy matematiki. Fundamentalnye napravleniya, Tom 14, Akad. Nauk SSSR, Vsesoyuz. Inst. Nauchn. i Tekhn. Inform., Moscow, 1987, Translation by D. Newton. Encyclopaedia of Mathematical Sciences, 14. Springer, Berlin (1990)
40. Wang, T., Pang, J.S.: Global error bounds for convex quadratic inequality systems. Optimization **31**(1), 1–12 (1994)
41. Wu, Z.L., Ye, J.J.: On error bounds for lower semicontinuous functions. Math. Program Ser. A. **92**(2), 301–314 (2002)
42. Yang, W.H.: Error bounds for convex polynomials. SIAM J. Optim. **19**(4), 1633–1647 (2008)
43. Zalinescu, C.: Convex Analysis in General Vector Spaces. World Scientific, Singapore (2002)