Error Bounds for Convex Inequality Systems

Adrian S. Lewis*
Department of Combinatorics and Optimization
University of Waterloo
Waterloo, Ontario, Canada N2L 3G1
Email: aslewis@orion.uwaterloo.ca

Jong-Shi Pang†
Department of Mathematical Sciences
The Johns Hopkins University
Baltimore, Maryland 21218-2689, U.S.A
Email: jsp@vicp.mts.jhu.edu

July 1996

Abstract

Using convex analysis, this paper gives a systematic and unified treatment for the existence of a global error bound for a convex inequality system. We establish a necessary and sufficient condition for a closed convex set defined by a closed proper convex function to possess a global error bound in terms of a natural residual. We derive many special cases of the main characterization, including the case where a Slater assumption is in place. Our results show clearly the essential conditions needed for convex inequality systems to satisfy global error bounds; they unify and extend a large number of existing results on global error bounds for such systems.

*The research of this author was based on work supported by the Natural Sciences and Engineering Research Council of Canada.

†The research of this author was based on work supported by the National Science Foundation under grant CCR-9213739 and the Office of Naval Research under grant N00014-93-1-0228.

1 Introduction

Let $f : \Re^n \to \Re \cup \{\infty\}$ be an extended-valued closed proper convex function. Define the closed convex set

$$\mathcal{S} \equiv \{x \in \Re^n : f(x) \leq 0\}, \tag{1}$$

which we shall assume to be nonempty throughout the paper. The main goal of this paper is to derive a necessary and sufficient condition for the existence of a constant $\gamma > 0$ such that

$$\operatorname{dist}(x, \mathcal{S}) \leq \gamma f(x)_+, \quad \text{for all } x \in \Re^n, \tag{2}$$

where "dist" is the distance function from a variable point to a fixed set (with the usual Euclidean norm) and $a_+ \equiv \max(0, a)$. An inequality of the kind (2) is called a *global error bound* for the set $\mathcal{S}$. This inequality bounds the distance from an arbitrary point $x \in \Re^n$ (thus 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 $f(x) \leq 0$. By appropriately specializing the function f, we unify and extend a wide variety of error bound results known to date for convex inequality systems.

The study of error bounds has in recent years received a lot of attention in the mathematical programming literature. We begin with a brief survey of previous papers related to our work. Robinson [34] established a global error bound for an arbitrary closed convex set S in a normed space under the assumption that S is bounded and has a nonempty interior. Mangasarian [30] considered a closed convex set $S \subseteq \Re^n$ defined by a system of finitely many differentiable convex inequalities and established a global error bound under a Slater assumption and an asymptotic constraint qualification (ACQ). Auslender and Crouzeix [2] extended Mangasarian's result to nondifferentiable functions. Luo and Luo [26] considered convex quadratic inequality systems and established global error bounds for them, assuming only the Slater condition (and no other constraint qualification). Klatte [23] studied the connection between Hausdorff continuity of solutions to perturbed inequality systems and global error bound of the unperturbed system. Li [25] obtained some interesting characterizations of error bounds on compact sets for convex differentiable inequalities in terms of a constraint qualification. Recently, Deng [13, 14] established error bounds for closed convex sets defined by closed proper convex functions in a Banach space, assuming a Slater condition on the associated recession functions. Deng and Hu [16] obtained error bound results for semidefinite programs. Additional work related to ours will be cited throughout the paper.

All the error bounds mentioned above are of the "unit exponent (i.e. Lipschitzian) type"; that is, the distance is bounded by a constant multiple of the residual. There are error bounds of the fractional exponent (i.e. Hölderian) type for other kinds of inequality systems; these results require an additional residual term which is equal to the original residual raised to a certain fractional power. The latter systems include polynomial systems [26], analytic systems [27] and their generalizations to "subanalytic sets" [29], convex quadratic inequalities without the Slater assumption [36], convex piecewise quadratic systems [24], and the solution system of a monotone linear complementarity problem [32]. In addition, there are many other studies on error bounds and their fundamental roles in various aspects of mathematical programming. A partial survey that contains more details of all the above-mentioned error bounds and their applications is documented in Chapter 2 of the monograph [28].

2 Preliminary Discussion

We begin our technical discussion with an example to illustrate that a global error bound (2) does not, in general, exist for the set $\mathcal{S}$. This example is constructed using a well-known function in convex analysis, denoted $\tilde{f}$ below, which appears in [35] and was subsequently used by Auslender and Crouzeix [3].

Example 1. Consider the closed proper convex function:

$$\tilde{f}(x_1, x_2) \equiv \begin{cases} \frac{x_1^2}{x_2} & \text{if } x_2 > 0 \\ 0 & \text{if } x_1 = x_2 = 0 \\ \infty & \text{otherwise.} \end{cases}$$

Let $f(x) \equiv \tilde{f}(x) - 1$. Define

$$x^k \equiv \begin{pmatrix} k \\ k^2 \end{pmatrix} + \sqrt{k} \begin{pmatrix} 1 \\ -\frac{1}{2k} \end{pmatrix}, \quad k = 1, 2, \ldots.$$

It is easy to verify that

$$\mathrm{dist}(x^k, \mathcal{S}) = \sqrt{k}\sqrt{1 + \frac{1}{4k^2}}$$

which approaches ∞ as k tends to infinity. Nevertheless, we have

$$f(x^k) \;=\; \frac{(k+\sqrt{k})^2}{k^2 - 1/2\sqrt{k}} - 1 \;>\; 0,$$

which approaches 0 as $k \to \infty$. Consequently, there can not exist any constants γ, $\rho > 0$ such that for all k,

$$\mathrm{dist}(x^k, \mathcal{S}) \;\leq\; \gamma\, f(x^k)^\rho.$$

In the terminology of [3], the above function f is not well-posed. Specifically, we say that a convex function $f : \Re^n \to \Re \cup \{\infty\}$ is *well-posed* if every stationary sequence of f is a minimizing sequence; that is with ∂f denoting the subdifferential operator, we have

$$[\lim_{k\to\infty} a^k = 0,\; a^k \in \partial f(x^k),\; \forall k\,] \;\Rightarrow\; \lim_{k\to\infty} f(x^k) = \inf_{x\in\Re^n} f(x).$$

This condition clearly fails for the above function $\tilde{f}$ and sequence $\{x^k\}$. Subsequently, we will show that if f is a well-posed closed proper convex function and satisfies a Slater assumption, then the global error bound (2) holds; see Corollary 1.

Another example of a convex function that fails the global error bound is given by

$$f(x_1, x_2) \;\equiv\; x_1 + \sqrt{x_1^2 + x_2^2}, \quad (x_1, x_2) \in \Re^2.$$

Unlike the function in Example 1, this function f is finite valued everywhere. It is easy to see that

$$\mathcal{S} \;=\; \{(x_1, 0) \,\in\, \Re^2 \,:\, x_1 \,\leq\, 0\}.$$

Thus $\mathcal{S}$ has an empty interior. The sequence of points $\{x^k \equiv (-k, 1)\}$ has the property that $\mathrm{dist}(x^k, \mathcal{S}) = 1$ for all k but

$$f(x^k)_+ \;=\; f(x^k) \;=\; -k + \sqrt{k^2+1} \to 0, \quad \text{as } k \to \infty.$$

Thus no error bound can hold for this sequence.

A major portion of our results will be concerned with a convex set of the following extended type:

$$\mathcal{S} \;\equiv\; \{x \,\in\, \mathcal{C} \,:\, f(x) \,\leq\, 0\}, \tag{3}$$

where $\mathcal{C}$ is a closed convex set in $\Re^n$. For a set of this type, we seek a global error bound of the form:

$$\operatorname{dist}(x, \mathcal{S}) \leq \gamma \max(f(x)_+, \operatorname{dist}(x, \mathcal{C})), \quad \text{for all } x \in \Re^n. \tag{4}$$

In the important case where $\mathcal{C}$ is a polyhedron, the well-known Hoffman bound for a polyhedral set [20] yields that $\operatorname{dist}(x, \mathcal{C})$ is bounded, above and below, by constant multiples of the amount of violation of the linear inequality constraints which represent $\mathcal{C}$.

The Slater assumption plays an important role in global error bounds for convex inequality systems; it postulates the existence of a vector $\hat{x} \in \mathcal{C}$ such that $f(\hat{x}) < 0$. Notice $\mathcal{S}$ may still have empty interior. Below, we give a set $\mathcal{S}$ of the type (3) which satisfies the Slater condition but for which no global error bound exists.

Example 2. We consider the space $\mathcal{M}^2$ of 2×2 symmetric matrices (which we may identify as $\Re^3$), with the Frobenious norm

$$\|M\| \equiv \sqrt{\text{trace } M^2}, \quad \text{for } M \in \mathcal{M}^2.$$

For any $M \in \mathcal{M}^2$, define

$$f(M) \equiv -\lambda_{\min}(M),$$

where $\lambda_{\min}$ denotes the smallest eigenvalue. Then f is a finite-valued, convex function on $\Re^3$. Let

$$\mathcal{C} \equiv \{M \in \mathcal{M}^2 : m_{11} = 1\}.$$

The set $\mathcal{S}$ then consists of all 2×2 symmetric positive semidefinite matrices with the first diagonal entry equal to 1. Clearly the Slater assumption is satisfied for $\mathcal{S}$. Let

$$X^k \equiv \begin{pmatrix} 1 & k+1 \\ k+1 & k^2 \end{pmatrix} \in \mathcal{C}, \quad k = 0, 1, \ldots.$$

By a direct computation, it can be verified that

$$f(X^k) = \frac{\sqrt{(k^2+1)^2 + 4(2k+1)} - (k+1)}{2} > 0,$$

which approaches 0 as k tends to infinity. Nevertheless $\operatorname{dist}(X^k, \mathcal{S})$ tends to 1 as $k \to \infty$. Consequently, a global error bound (even of fractional exponent type) can not hold for $\mathcal{S}$.

There are many occasions where we are interested in an error bound of a convex inequality system with the test vector x restricted to lie in a given proper subset T of $\Re^n$. The next example shows that such a bound is not guaranteed to exist even for a compact set T. In this example, the Slater condition is violated; as we shall see later, this is the main cause for the nonexistence of the error bound; cf. Proposition 3.

Example 3. Consider the same function f and space $\mathcal{M}^2$ defined in Example 2. Let

$$\mathcal{C} \equiv \{M \in \mathcal{M}^2 : m_{11} = 0\}.$$

Then we have

$$\mathcal{S} = \left\{ \begin{pmatrix} 0 & 0 \\ 0 & \alpha \end{pmatrix} : \alpha \geq 0 \right\}.$$

Let for all scalars $\varepsilon > 0$, the matrix

$$X^\varepsilon \equiv \begin{pmatrix} \varepsilon & \sqrt{\varepsilon} \\ \sqrt{\varepsilon} & 1 \end{pmatrix} \notin \mathcal{C}$$

is symmetric positive semidefinite. For $\varepsilon > 0$ sufficiently small,

$$\operatorname{dist}(X^\varepsilon, \mathcal{S}) = \sqrt{2\varepsilon + \varepsilon^2};$$

but

$$\operatorname{dist}(X^\varepsilon, \mathcal{C}) = \varepsilon.$$

Consequently, an error bound (of unit exponent type) can not hold for the set $\mathcal{S}$ with reference to the compact set $\{X^\varepsilon : \varepsilon \in [0,1]\}$. Nevertheless an error bound of a square-root type holds. Indeed the latter conclusion is not incidental; see Proposition 6.

Error bounds of constraint systems on compact sets are closely related to the concept of metric regularity of these systems. Specifically, the system (3) (or the set $\mathcal{S}$) is said to be *metrically regular* at a feasible vector $x \in \mathcal{S}$ if a local error bound holds in some neighborhood of x, i.e., there exist positive scalars γ and δ such that

$$\operatorname{dist}(y, \mathcal{S}) \leq \gamma \max(\operatorname{dist}(y, \mathcal{C}), f(y)_+), \quad \text{for all } y \text{ satisfying } \|y - x\| \leq \delta.$$

The article [7] contains extensive references on this regularity concept and the fundamental role it plays in the theory of exact penalty functions for constrained optimization problems. In the recent work, Li [25] obtained an interesting characterization of metric regularity of a convex differentiable inequality system in terms of the Abadie constraint qualification [1]. We will generalize one part of Li's result later.

A set $\mathcal{S}$ of the form (3) is the intersection of two closed convex sets. In [6], two subsets C and D of $\Re^n$ are defined to be "linearly regular" if there exists a constant $\sigma > 0$ such that

$$\operatorname{dist}(x, C \cap D) \le \sigma \max(\operatorname{dist}(x, C), \operatorname{dist}(x, D)), \quad \text{for all } x \in \Re^n;$$

the sets C and D are said to be "boundedly linearly regular" if for every bounded subset T of $\Re^n$, a constant $\sigma > 0$ exists such that

$$\operatorname{dist}(x, C \cap D) \le \sigma \max(\operatorname{dist}(x, C), \operatorname{dist}(x, D)), \quad \text{for all } x \in T.$$

These two regularity concepts are related to the error bounds established in the present paper; nevertheless, the regularity concepts are not concerned with bounding $\operatorname{dist}(x, f^{-1}(-\infty, 0])$ in terms of $f(x)_+$; the latter task is a major focus of our work. (A word of notation: $f^{-1}(-\infty, 0]$ is the inverse image of the closed interval $(-\infty, 0]$ under f; similarly, we write $f^{-1}(0)$ for the zero set of f.)

Convex analysis is the main tool used in this paper. Consequently, our approach is closely related to the work of Burke and Tseng [9] and Deng [13, 14] who also relied on convex analysis for the derivation of error bounds. In what follows, we review some standard notation and basic properties of a closed proper convex function $f : \Re^n \to \Re \cup \{\infty\}$; see [21, 35]. We have seen the subdifferential ∂f already; the conjugate function of f is denoted f^*; the recession function of f is denoted f_∞. The domain of finiteness of f is denoted $\operatorname{dom}(f)$. This set, $\operatorname{dom}(f)$ is closed and convex. Inequality (2) is clearly equivalent to

$$\operatorname{dist}(x, \mathcal{S}) \le \gamma f(x)_+, \quad \text{for all } x \in \operatorname{dom}(f).$$

A similar remark applies to (4). The directional derivative of f at a vector x along a direction d is denoted $f'(x; d)$. For a set-valued operator F from $\Re^n$ into subsets of $\Re^n$, $\operatorname{dom}(F)$ consists of all vectors $x \in \Re^n$ for which $F(x) \neq \emptyset$. The relative interior and boundary of a convex set $S \subseteq \Re^n$ are denoted $\operatorname{ri}(S)$ and $\operatorname{rb}(S)$ respectively; the closure and boundary of S is denoted $\operatorname{cl}(S)$ and $\operatorname{bd}(S)$ respectively; the tangent cone and normal cone of S at a vector $x \in S$ are denoted by $\mathcal{T}(x; S)$ and $\mathcal{N}(x; S)$ respectively. The recession cone and polar cone of S are denoted S_∞ and S^o, respectively.

For a closed proper convex function f, we have the following inclusions:

$$\mathrm{ri}(\mathrm{dom}(f)) \subseteq \mathrm{dom}(\partial f) \subseteq \mathrm{dom}(f) \subseteq \mathrm{cl}(\mathrm{dom}(f)).$$

For any vector $x \in \mathrm{dom}(\partial f)$, we have

$$f'(x; d) = \sup\{u^T d : u \in \partial f(x)\}, \quad \text{for all } d \in \Re^n.$$

In particular, it follows that

$$f'(x; d) \geq d^T d, \quad \text{for all } d \in \partial f(x). \tag{5}$$

Throughout this paper, these and other known results from convex analysis will be used freely; we rely on the texts [21, 35] where details of the results can be found. For any two subsets C and D of $\Re^n$, we define

$$\mathrm{dist}(C, D) \equiv \inf_{x \in C} \mathrm{dist}(x, D) = \inf_{(x,y) \in C \times D} \|x - y\|.$$

Finally, we let $\mathbb{B}_n$ denote the Euclidean unit ball in $\Re^n$; we will omit the subscript n in $\mathbb{B}_n$ if the dimension of the space is obvious from the context.

3 Main Results

We divide our analysis into two cases, depending on whether the set $\mathcal{C}$ is the entire space $\Re^n$ or a proper subset of $\Re^n$. The reason for this division is because the former case is somewhat simpler and illustrates some ideas more clearly.

3.1 The unconstrained case

We first deal with the set $\mathcal{S}$ given by (1). The following result gives necessary and sufficient conditions for the global error bound (2) to hold.

Theorem 1 *Let $f : \Re^n \to \Re \cup \{\infty\}$ be a closed proper convex function and let $\mathcal{S} \equiv f^{-1}(-\infty, 0]$. For any scalar $\gamma > 0$, the following statements are equivalent.*

(a) The global error bound (2) holds.

(b) For all $\bar{x} \in f^{-1}(0)$ and $d \in \mathcal{N}(\bar{x}; \mathcal{S})$,

$$f'(\bar{x}; d) \geq \gamma^{-1} \|d\|. \tag{6}$$

(c) For all $\bar{x} \in f^{-1}(0)$ and $d \in \mathcal{T}(\bar{x}; \mathrm{dom}(f)) \cap \mathcal{N}(\bar{x}; \mathcal{S})$, (6) holds.

Proof. (a) $\Rightarrow$ (b). Assume that (2) holds. Let $\bar{x} \in f^{-1}(0)$ and $d \in \mathcal{N}(\bar{x}; \mathcal{S})$ be given. Thus for all $\tau > 0$,

$$\mathrm{dist}(\bar{x} + \tau d, \mathcal{S}) \;=\; \tau \, \|d\|.$$

By (2), we deduce

$$f(\bar{x} + \tau d) \;\geq\; \gamma^{-1} \, \tau \, \|d\|.$$

Dividing by $\tau > 0$, recalling $f(\bar{x}) = 0$, and passing to the limit $\tau \to 0$, we deduce

$$f'(\bar{x}; d) \;\geq\; \gamma^{-1} \, \|d\|,$$

which is the desired inequality (6). Thus statement (b) holds.

(b) $\Rightarrow$ (c). This is obvious.

(c) $\Rightarrow$ (a). Let $x \notin \mathcal{S}$ be given. Without loss of generality we may assume that $f(x) < \infty$. Consider the projection problem:

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}(x - u)^T (x - u) \\ \text{subject to} & u \in \mathcal{S}. \end{array} \tag{7}$$

The projected point, denoted $\Pi_{\mathcal{S}}(x)$, is the unique vector $\bar{x} \in \mathcal{S}$ satisfying $x - \bar{x} \in \mathcal{N}(\bar{x}; \mathcal{S})$. Since $\mathrm{dom}(f)$ is convex, $x - \bar{x} \in \mathcal{T}(\bar{x}; \mathrm{dom}(f))$. We claim that $f(\Pi_{\mathcal{S}}(x)) = 0$. Indeed if $f(\Pi_{\mathcal{S}}(x)) < 0$, then for $\tau > 0$ sufficiently small, the vector $\Pi_{\mathcal{S}}(x) + \tau \, (x - \Pi_{\mathcal{S}}(x))$ belongs to $\mathcal{S}$ and its distance to x is less than that from $\Pi_{\mathcal{S}}(x)$ to x; this contradicts the nearest point property of $\Pi_{\mathcal{S}}(x)$. Hence $f(\Pi_{\mathcal{S}}(x)) = 0$. By assumption, we deduce

$$f'(\Pi_{\mathcal{S}}(x); x - \Pi_{\mathcal{S}}(x)) \;\geq\; \gamma^{-1} \, \|x - \Pi_{\mathcal{S}}(x)\|.$$

We have

$$f(x)_+ \;=\; f(x) \;=\; f(x) - f(\Pi_{\mathcal{S}}(x)) \;\geq\; f'(\Pi_{\mathcal{S}}(x); x - \Pi_{\mathcal{S}}(x)),$$

where the last inequality follows from the convexity of f. Consequently, it follows that

$$\mathrm{dist}(x; \mathcal{S}) \;\leq\; \gamma \, f(x)_+,$$

as desired. Q.E.D.

Remarks. Working in Banach spaces and assuming $\mathcal{N}(\bar{x}; \mathcal{S}) = \Re_+ \partial f(\bar{x})$ for all $\bar{x} \in f^{-1}(0)$, Deng [14, 15] established a sufficient condition in terms of subgradients (and not directional derivatives) for the global error bound (2). See Corollary 1 below. To the best of our knowledge, necessary conditions for such global error bounds have not appeared in the literature before.

Michael Ferris pointed out that a condition analogous to statement (c) in Theorem 1 also appeared in the context of characterization of weak sharp minima for convex programs; see [8, Theorem 2.6]. In general, close connections between error bounds and weak sharp minima are known [17, 18]. In what follows, we explain one such connection in the context of the unconstrained system $f(x) \leq 0$. For this purpose, assume that the Slater condition fails to hold; thus $f(x) \geq 0$ for all $x \in \Re^n$. In this case $f^{-1}(0)$ is the set of (unconstrained) global minimizers of f and (see the proof on Theorem 1)

$$\operatorname{dist}(x, \mathcal{S}) = \operatorname{dist}(x, f^{-1}(0)), \quad \text{for all } x \in \operatorname{dom}(f).$$

Hence an error bound holds for the set $\mathcal{S}$ if and only if $f^{-1}(0)$ is a set of weak sharp (unconstrained) minima of f. More generally, for an arbitrary convex function f with a finite (unconstrained) infimum value $f_{\inf} > -\infty$, the set $\mathcal{S} \equiv \{x \in \Re^n : f(x) \leq f_{\inf}\}$ is the set of weak sharp minima of f if and only if an error bound holds for $\mathcal{S}$. This discussion extends to the constrained system $f(x) \leq 0$, $x \in \mathcal{C}$ which we will not repeat. Due to the close connection between error bounds and weak sharp minima, the results obtained in this paper have direct implications for the theory of weak sharp minima of optimization problems. We should also point out that by Theorem 2.6 in [8], for f and $\mathcal{S}$ as given in Theorem 1, a vector $\bar{x} \in \mathcal{S}$ and a scalar $\alpha > 0$ satisfy

$$f'(\bar{x}; d) \geq \alpha \|d\|, \quad \forall\, d \in \mathcal{T}(\bar{x}; \operatorname{dom}(f)) \cap \mathcal{N}(\bar{x}; \mathcal{S}),$$

if and only if $\bar{x}$ and d satisfy

$$\alpha\, \mathbb{B} \subseteq \partial f(\bar{x}) + [\, \mathcal{T}(\bar{x}; \operatorname{dom}(f)) \cap \mathcal{N}(\bar{x}; \mathcal{S})\,]^o.$$

In what follows, we present several sufficient conditions for statement (c) of Theorem 1 to hold, and thus for the set $\mathcal{S} \equiv f^{-1}(-\infty, 0]$ to possess a global error bound. These conditions highlight the important role played by several fundamental objects associated with a convex function in the existence of error bounds for the level sets of the given function. For this purpose, we first prove a lemma.

Lemma 1 *Let $f : \Re^n \to \Re \cup \{\infty\}$ be a closed proper convex function and let $\mathcal{S} \equiv f^{-1}(-\infty, 0]$. Let L be the linear subspace of $\Re^n$ parallel to the affine hull of* $\mathrm{dom}(f)$. *If there exists a vector $\hat{x} \in \Re^n$ such that $f(\hat{x}) < 0$, then for every $\bar{x} \in f^{-1}(0) \cap \mathrm{ri}(\mathrm{dom}(f))$,*

$$\mathcal{N}(\bar{x}; \mathcal{S}) \cap L = \Re_+ \partial f(\bar{x}) \cap L;$$

consequently,

$$\mathcal{N}(\bar{x}; \mathcal{S}) \cap \mathcal{T}(\bar{x}; \mathrm{dom}(f)) \subseteq \Re_+ \partial f(\bar{x}).$$

Proof. Let $\bar{x} \in f^{-1}(0) \cap \mathrm{ri}(\mathrm{dom}(f))$ be given. To simplify the notation, we assume without loss of generality that $\bar{x}$ is the origin. Under this assumption, L coincides with the affine hull of $\mathrm{dom}(f)$. We have

$$\Re^n = L \oplus L^\perp,$$

where $\oplus$ denotes the direct sum operation and $L^\perp$ denotes the orthogonal complement of L. Define a new function g on L simply by letting g be the restriction of f to L. Then g is a closed proper convex function defined on L satisfying the Slater assumption; moreover the origin is an interior point of $\mathrm{dom}(g) \subseteq L$. By Corollary 23.7.1 in [35], it follows that in the space L,

$$\mathcal{N}(0, g^{-1}(-\infty, 0]) = \Re_+ \partial g(0).$$

Notice two things: (1) $\mathcal{N}(0; \mathcal{S}) = \mathcal{N}(0; g^{-1}(-\infty, 0]) \oplus L^\perp$ and (2) considered a subset of $\Re^n$, $\partial g(0)$ is contained in $\partial f(0)$. Putting these observations together yields

$$\mathcal{N}(0; \mathcal{S}) \cap L \subseteq \Re_+ \partial f(0) \cap L.$$

Since the reverse inclusion clearly holds, equality follows. This establishes the first assertion of the lemma.

The proof of the last inclusion of the lemma follows easily from the observation that $\mathcal{T}(\bar{x}; \mathrm{dom}(f))$ is contained in L. Q.E.D.

The following corollary contains the promised sufficient conditions for the set $\mathcal{S}$ to possess a global error bound.

Corollary 1 *Let $f : \Re^n \to \Re \cup \{\infty\}$ be a closed proper convex function and let $\mathcal{S} \equiv f^{-1}(-\infty, 0]$. Assume $f^{-1}(0) \subseteq \mathrm{ri}(\mathrm{dom}(f))$. Consider the following statements:*

(a) (Global error bound) the global error bound (2) holds for some scalar $\gamma > 0$;

(b) (Strong Slater) $0 \notin \mathrm{cl}(\partial f(f^{-1}(0)))$;

(c) (Well-posedness and Slater) f is well-posed and satisfies the Slater condition (i.e., $f(\hat{x}) < 0$ for some $\hat{x} \in \Re^n$);

(d) (Slater and a subcase of well-posedness) $0 \in \mathrm{ri}(\mathrm{dom}(f^*)) \setminus \partial f(f^{-1}(0))$;

(e) (Recession) $0 \notin \mathrm{cl}(\mathrm{dom}(f^*))$, or equivalently, there exists a vector $d \in \Re^n$ satisfying $f_\infty(d) < 0$;

(f) (Slater and positive homogeneity) f is positively homogeneous (i.e., $f(\tau d) = \tau f(d)$ for all $\tau \geq 0$) and satisfies the Slater condition.

The implications below hold:

$$\begin{array}{ccccccc} (f) & \Rightarrow & (e) & \Rightarrow & (b) & \Rightarrow & (a) \\ & & & & \Uparrow & & \\ & & (d) & \Rightarrow & (c). & & \end{array}$$

Proof. Since f is convex, it follows that the Slater condition holds if and only if $0 \notin \partial f(f^{-1}(0))$.

If (b) holds, then so does the Slater condition. Moreover, we have

$$\delta \equiv \mathrm{dist}(0, \mathrm{cl}(\partial f(f^{-1}(0))) > 0.$$

Thus for all vectors $\bar{x} \in f^{-1}(0)$ and $d \in \partial f(\bar{x})$, we have

$$\|d\| \geq \delta;$$

which implies, by (5),

$$f'(\bar{x}; d) \geq \delta \, \|d\|.$$

By Lemma 1, it follows that statement (c) of Theorem 1 holds with $\gamma \equiv \delta^{-1}$. Hence the error bound (2) holds under condition (b) here.

By Theorem 2.3 in [4], (d) implies (c). We show that (c) implies (b). Suppose (c) holds but there exist sequences $\{\bar{x}^k\} \subset f^{-1}(0)$ and $\{d^k\}$ such that for each k, $d^k \in \partial f(\bar{x}^k)$ and $\{d^k\}$ converges to the zero vector. Thus $\{d^k\}$ is a stationary sequence. Since f is well posed, it follows that

$$\inf_{x \in \Re^n} f(x) = \lim_{k \to \infty} f(\bar{x}^k) = 0.$$

But this contradicts the Slater condition.

The equivalence of the two statements in (e) follows from Corollary 13.3.4 in [35]. Since

$$\partial f(f^{-1}(0)) \subseteq \text{range } \partial f = \text{dom}(\partial f^*) \subseteq \text{dom}(f^*),$$

(e) implies (b). Finally, we show that (f) implies (e). Since f is positively homogeneous, we have $f(0) = 0$. Let d satisfy $f(d) < 0$; then

$$f_\infty(d) = \lim_{\tau \to \infty} \frac{f(\tau d)}{\tau} = f(d) < 0.$$

Thus (e) holds. Q.E.D.

We make several remarks about the assumptions in the above corollary. The assumption $f^{-1}(0) \subseteq \text{ri}(\text{dom}(f))$ does not necessarily imply $f^{-1}(-\infty, 0] \subseteq \text{ri}(\text{dom}(f))$. This can be seen from the function f in Example 1. For this function, the former assumption is clearly satisfied because

$$f^{-1}(0) = \{(x_1, x_2) \in \Re^2 : x_1^2 = x_2 > 0\} \subseteq \Re \times \Re_{++} = \text{int}(\text{dom}(f));$$

but the origin is contained in $f^{-1}(-\infty, 0]$ and not in $\text{ri}(\text{dom}(f))$ which is equal to $\text{int}(\text{dom}(f))$. The recession condition in (e), i.e., there exists $d \in \Re^n$ such that $f_\infty(d) < 0$, was introduced in [13]; in the subsequent paper [14], it was noted that this condition is equivalent to $0 \notin \text{cl}(\text{dom}(f^*))$. Condition (c) was inspired by the work [3, 4] concerning well-posed convex functions. For more discussion on the role of error bounds in the theory of well-posed optimization problems, the reader is referred to the paper [10]. Conditions (c) and (e) are not compatible. For example, under condition (e), f can not attain its infimum value; whereas this is possible under (c). The strong Slater condition (b) has not appeared in the literature before. Clearly condition (a) does not imply (b); for instance, take $f(x) \equiv \max(x, 0)$, $x \in \Re$.

Under the assumption that $f^{-1}(0) \subseteq \text{int}(\text{dom}(f))$ and the Slater assumption, the result above shows that a global error bound holds if either

$$0 \in \text{ri}(\text{dom}(f^*)), \quad \text{or} \quad 0 \notin \text{cl}(\text{dom}(f^*)).$$

The critical case is thus

$$0 \in \text{rb}(\text{dom}(f^*)).$$

Notice that the function f in Example 1 satisfies the two stated assumptions and

$$f^*(y_1, y_2) = \begin{cases} 1 & \text{if } y_1^2/4 + y_2 \le 0, \\ \infty & \text{otherwise,} \end{cases}$$

so $0 \in \text{rb}(\text{dom}(f^*))$.

3.2 The constrained case

We now extend Theorem 1 to a convex set of the form (3) where $\mathcal{C}$ is a closed convex proper subset of $\Re^n$ (not necessarily a polyhedron).

Theorem 2 *Let $f : \Re^n \to \Re \cup \{\infty\}$ be a closed proper convex function and let $\mathcal{C}$ be a closed convex subset of $\Re^n$ satisfying either the projection condition,*

$$\Pi_{\mathcal{C}}(\mathrm{dom}(f)) \subseteq \mathrm{dom}(f), \tag{8}$$

or the interiority condition:

$$\{x \in \mathcal{C} : f(x) < 0\} \subseteq \mathrm{int}(\mathrm{dom}(f)). \tag{9}$$

Let $\mathcal{S} \equiv \mathcal{C} \cap f^{-1}(-\infty, 0]$. For a scalar $\gamma \geq 1$, the global error bound (4) holds if and only if for all $\bar{x} \in f^{-1}(0) \cap \mathcal{C}$ and $d \in \mathcal{N}(\bar{x}; \mathcal{S})$,

$$\max \{ \mathrm{dist}(d, \mathcal{T}(\bar{x}; \mathcal{C})),\, f'(\bar{x}; d) \} \geq \gamma^{-1} \|d\|. \tag{10}$$

Proof. Write $g(x) \equiv \mathrm{dist}(x, \mathcal{C}) = \|x - \Pi_{\mathcal{C}}(x)\|$. Then g is a convex function on $\Re^n$; moreover, the directional derivative of g at a vector $\bar{x} \in \mathcal{C}$ along an arbitrary direction $d \in \Re^n$ is given by

$$g'(\bar{x}; d) = \|d - \Pi'_{\mathcal{C}}(\bar{x}; d)\| = \mathrm{dist}(d, \mathcal{T}(\bar{x}; \mathcal{C})), \tag{11}$$

where the last equality follows from the fact that $\Pi'_{\mathcal{C}}(\bar{x}; d) = \Pi_{\mathcal{T}(\bar{x};\mathcal{C})}(d)$ by [21, Proposition 5.3.5]. Define the convex function

$$\tilde{f}(x) \equiv \max(f(x), g(x)), \quad x \in \Re^n.$$

We have $\mathcal{S} = \{x \in \Re^n : \tilde{f}(x) \leq 0\}$; moreover,

$$\tilde{f}(x)_+ = \max(f(x)_+, \mathrm{dist}(x, \mathcal{C})).$$

Noting (11), we deduce that for arbitrary vectors $\bar{x} \in f^{-1}(0) \cap \mathcal{C}$ and $d \in \Re^n$,

$$\tilde{f}'(\bar{x}; d) = \max\{ f'(\bar{x}; d), \mathrm{dist}(d, \mathcal{T}(\bar{x}; \mathcal{C})) \}.$$

Suppose that the error bound (4) holds; that is for all $x \in \Re^n$,

$$\mathrm{dist}(x, \mathcal{S}) \leq \gamma \tilde{f}(x)_+, \tag{12}$$

By Theorem 1, we have

$$\tilde{f}'(\bar{x}; d) \geq \gamma^{-1} \|d\|$$

for all $\bar{x} \in \tilde{f}^{-1}(0)$ and all $d \in \mathcal{N}(\bar{x}; \mathcal{S})$. Since

$$f^{-1}(0) \cap \mathcal{C} \subseteq \tilde{f}^{-1}(0),$$

(10) holds for all $(\bar{x}, d) \in (f^{-1}(0) \cap \mathcal{C}) \times \mathcal{N}(\bar{x}; \mathcal{S})$.

Conversely, suppose (10) holds for some $\gamma \geq 1$ and all $(\bar{x}, d) \in (\mathcal{C} \cap f^{-1}(0)) \times \mathcal{N}(\bar{x}; \mathcal{S})$. Thus for any such pair $(\bar{x}, d)$, we have

$$\tilde{f}'(\bar{x}; d) \geq \gamma^{-1} \|d\|. \tag{13}$$

Since $\mathcal{C} \cap f^{-1}(0)$ is only a subset of $\tilde{f}^{-1}(0)$, Theorem 1 is not directly applicable to yield the desired error bound (12). Proceeding as in the proof of Theorem 1, let $x \in \mathrm{dom}(f)$ be arbitrary and consider the projection problem:

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}(x-u)^T(x-u) \\ \text{subject to} & u \in \mathcal{S} = \{u : \tilde{f}(u) \leq 0\}. \end{array}$$

Now if the projected vector $\bar{x} \equiv \Pi_{\mathcal{S}}(x)$, which must belong to $\mathcal{C}$, satisfies $f(\bar{x}) = 0$, then by the proof of Theorem 1, the inequality (13) yields the desired error bound (12). If $\bar{x}$ satisfies $f(\bar{x}) < 0$, then $\bar{x}$ must be the projection of x onto $\mathcal{C}$. Indeed, let $x' \equiv \Pi_{\mathcal{C}}(x)$ and assume that $x' \neq \bar{x}$. Then we have

$$\|x - x'\| < \|x - \bar{x}\|.$$

Moreover, on the one hand, if the projection condition (8) holds, then $x' \in \mathrm{dom}(f)$; on the other hand, if the interiority condition (9) holds, then $\bar{x} \in \mathrm{int}(\mathrm{dom}(f))$. In either case, it follows that for $\tau > 0$ sufficiently small, the vector $\bar{x} + \tau(x' - \bar{x})$ belongs to $\mathcal{S}$ and, because of the above strict inequality, its distance to x is less than that from $\bar{x}$ to x; a contradiction. Hence $x' = \bar{x}$. Consequently,

$$\mathrm{dist}(x, \mathcal{S}) = \mathrm{dist}(x, \mathcal{C}) \leq \gamma \max(f(x)_+, \mathrm{dist}(x, \mathcal{C})),$$

because $\gamma \geq 1$ by assumption. Q.E.D.

Remarks. The condition (8) is trivially satisfied if either $\mathcal{C} \equiv \Re^n$ (the unconstrained case) or f is finite valued. The condition (9) holds if the set $\{x \in \Re^n : f(x) < 0\}$ lies in the interior of $\mathrm{dom}(f)$; the latter condition depends on f only. For example, it holds if f is essentially smooth [35, Section 26]. The assumption that $\gamma \geq 1$ is not particularly restrictive: the existence of an error bound of the form (4) for *some* $\gamma > 0$ is clearly equivalent to its existence for some $\gamma \geq 1$.

The above result makes it clear that the normal cone $\mathcal{N}(\bar{x};\mathcal{S})$ for vectors $\bar{x} \in \mathcal{C} \cap f^{-1}(0)$ plays an important role in the existence of a global error bound for the set $\mathcal{S}$. (See also Theorem 1.) In what follows we examine this cone and the left-hand quantity in (10) more closely. Since $\mathcal{S}$ is the intersection of $\mathcal{C}$ and $Q \equiv f^{-1}(-\infty, 0]$, it follows that

$$\mathcal{N}(\bar{x};\mathcal{S}) \supseteq \mathcal{N}(\bar{x};\mathcal{C}) + \mathcal{N}(\bar{x};Q) \supseteq \mathcal{N}(\bar{x};\mathcal{C}) + \Re_+ \partial f(\bar{x}).$$

Under an appropriate constraint qualification, equality holds throughout the above expression. In what follows we give a consequence of such an equality.

Proposition 1 *Let $f : \Re^n \to \Re \cup \{\infty\}$ be a closed proper convex function and let $\mathcal{C}$ be a closed convex subset of $\Re^n$ satisfying the projection condition (8). Let $\mathcal{S} \equiv \mathcal{C} \cap f^{-1}(-\infty, 0]$. If $\bar{x} \in \mathcal{C} \cap f^{-1}(0)$ is such that*

$$\mathcal{N}(\bar{x};\mathcal{S}) = \mathcal{N}(\bar{x};\mathcal{C}) + \Re_+ \partial f(\bar{x}), \tag{14}$$

then for all nonzero $d \in \mathcal{N}(\bar{x};\mathcal{S})$,

$$\max\{\, \mathrm{dist}(d, \mathcal{T}(\bar{x};\mathcal{C})),\ f'(\bar{x};d)\,\} > 0.$$

Proof. Suppose $\bar{x} \in f^{-1}(0) \cap \mathcal{C}$ and $0 \neq d \in \mathcal{N}(\bar{x};\mathcal{S})$ are such that

$$\max\{\, \mathrm{dist}(d, \mathcal{T}(\bar{x};\mathcal{C})),\ f'(\bar{x};d)\,\} \leq 0.$$

Then $d \in \mathcal{T}(\bar{x};\mathcal{C})$. Assume that (14) holds. Write $d \equiv u + \lambda v$, where $(u, v) \in \mathcal{N}(\bar{x};\mathcal{C}) \times \partial f(\bar{x})$ and $\lambda \geq 0$. We have

$$0 < d^T d = d^T u + \lambda\, d^T v \leq \lambda\, d^T v \leq \lambda\, f'(\bar{x};d) \leq 0,$$

which is a contradiction. Q.E.D.

By refining the proof of the above proposition, we can establish an important corollary of Theorem 2 which provides several sufficient conditions for the set (3) to possess a global error bound. Before stating this corollary, we should point out that we could have formulated Theorem 2 in a form analogous to Theorem 1, i.e., with three equivalent statements rather than just two in the present form. With this refinement of Theorem 2, the assumption $\mathcal{C} \cap f^{-1}(0) \subseteq \mathrm{int}(\mathrm{dom}(f))$ in the following corollary can be replaced by the weaker assumption $\mathcal{C} \cap f^{-1}(0) \subseteq \mathrm{ri}(\mathrm{dom}(f))$; the resulting corollary will then be a complete generalization of Corollary 1 which pertains to the special case $\mathcal{C} = \Re^n$. Nevertheless, we will not be concerned with this minor technical point because in the next section, most of the applications of the corollary below are all defined by a finite-valued f.

Corollary 2 *Let $f : \Re^n \to \Re \cup \{\infty\}$ be a closed proper convex function and let $\mathcal{C}$ be a closed convex subset of $\Re^n$ satisfying either the projection condition (8) or the interiority condition (9). Let $\mathcal{S} \equiv \mathcal{C} \cap f^{-1}(-\infty, 0]$. Assume $\mathcal{S} \neq \emptyset$ and $\mathcal{C} \cap f^{-1}(0) \subseteq \operatorname{int}(\operatorname{dom}(f))$. Consider the following statements:*

(a) (Global error bound) there exists a scalar $\gamma > 0$ such that

$$\operatorname{dist}(x, \mathcal{S}) \;\leq\; \gamma \max(\operatorname{dist}(x, \mathcal{C}),\, f(x)_+), \quad \textit{for all } x \in \Re^n.$$

(b) (Strong Slater) there exists a constant $\kappa > 0$ such that for vectors $\bar{x} \in f^{-1}(0) \cap \mathcal{C}$, scalars $\lambda > 0$, and pairs $(u, v) \in \mathcal{N}(\bar{x}; \mathcal{C}) \times \partial f(\bar{x})$ satisfy

$$\|u\| + \lambda \;\leq\; \kappa \,\|\, u + \lambda v \,\|;$$

(c) (Bounded subgradients and a constraint qualification) the set $\partial f(\mathcal{C} \cap f^{-1}(0))$ is bounded and

$$\inf\{\operatorname{dist}(-\mathcal{N}(\bar{x}; \mathcal{C}), \partial f(\bar{x})) \;:\; \bar{x} \in \mathcal{C} \cap f^{-1}(0)\} > 0; \tag{15}$$

(d) (Generalized Robinson condition) the Slater condition holds and $\operatorname{cl}(\mathcal{C} \cap f^{-1}(0))$ is bounded and contained in $\operatorname{int}(\operatorname{dom}(f))$;

(e) (Recession) there exists a vector $d \in \mathcal{S}_\infty$ satisfying $f_\infty(d) < 0$.

The following implications hold:

$$\begin{array}{ccccccc} (d) & \Rightarrow & (c) & \Rightarrow & (b) & \Rightarrow & (a) \\ & & \Uparrow & & & & \\ & & (e). & & & & \end{array}$$

Proof. We shall prove the implications in the reverse order. So assume (b). As in the proof of Corollary 1, we verify that the Slater condition holds for the set $\mathcal{S}$. Indeed if this condition fails to hold, then every vector in $\mathcal{S}$ must be a zero of f and is a global minimum of f on the set $\mathcal{C}$. Consequently, for every $x \in \mathcal{S}$, we have $x \in \mathcal{C} \cap f^{-1}(0)$ and there exists $(u, v) \in \mathcal{N}(x; \mathcal{C}) \times \partial f(x)$ such that

$$0 \;=\; u + v.$$

But this violates assumption (b) because by this assumption, $\|u + v\| \geq \kappa^{-1}$. Consequently, the Slater condition holds for $\mathcal{S}$. It follows that (14) holds for every $\bar{x} \in \mathcal{C} \cap f^{-1}(0)$.

To prove the validity of the global error bound, we assume by way of contradiction (by Theorem 2) that there exist sequences $\{\bar{x}^k\} \subset f^{-1}(0) \cap \mathcal{C}$ and $\{d^k\}$ such that for each k, $0 \neq d^k \in \mathcal{N}(\bar{x}^k; \mathcal{S})$ and

$$\max\{\, \mathrm{dist}(d^k, \mathcal{T}(\bar{x}^k; \mathcal{C})),\, f'(\bar{x}^k; d^k)\,\} \leq \frac{\|d^k\|}{k}.$$

Let $\tilde{d}^k \equiv d^k / \|d^k\|$ and let $w^k \in \mathcal{T}(\bar{x}^k; \mathcal{C})$ be such that $\mathrm{dist}(\tilde{d}^k, \mathcal{T}(\bar{x}^k; \mathcal{C})) = \|\tilde{d}^k - w^k\|$. Then

$$\lim_{k \to \infty} \|\tilde{d}^k - w^k\| = 0.$$

For each k, we may write

$$\tilde{d}^k = u^k + \lambda_k v^k,$$

for some $(u^k, v^k) \in \mathcal{N}(\bar{x}^k; \mathcal{C}) \times \partial f(\bar{x}^k)$ and $\lambda_k \geq 0$. By assumption (b), we have

$$\|u^k\| + \lambda_k \leq \kappa.$$

Thus the sequences $\{\lambda_k\}$ and $\{u^k\}$ are bounded. As in the proof of Proposition 1, we have

$$(\tilde{d}^k)^T v^k \leq f'(\bar{x}^k, \tilde{d}^k) \leq 1/k,$$

which implies

$$\limsup_{k \to \infty} \lambda_k (\tilde{d}^k)^T v^k \leq 0.$$

We have

$$\begin{aligned} 1 &= (\tilde{d}^k)^T \tilde{d}^k = (\tilde{d}^k)^T u^k + \lambda_k (\tilde{d}^k)^T v^k \\ &= (\tilde{d}^k - w^k)^T u^k + (w^k)^T u^k + \lambda_k (\tilde{d}^k)^T v^k \\ &\leq (\tilde{d}^k - w^k)^T u^k + \lambda_k (\tilde{d}^k)^T v^k, \end{aligned}$$

which yields a contradiction because the limsup of the right-hand side is nonpositive as k tends to infinity. Applying Theorem 2, we conclude that the set $\mathcal{S}$ has a global error bound under assumption (b).

We next show that condition (c) implies (b). Assume (b) fails. Then there exist sequences of vectors $\{\bar{x}^k\} \subset \mathcal{C} \cap f^{-1}(0)$ and $\{(u^k, v^k)\}$ and scalars $\{\lambda^k\}$ with $(u^k, v^k) \in \mathcal{N}(\bar{x}^k; \mathcal{C}) \times \partial f(\bar{x}^k)$ and $\lambda_k > 0$ for each k such that

$$\|u^k + \lambda_k v^k\| \leq k^{-1}(\|u^k\| + \lambda_k).$$

The sequence $\{v^k\}$ is bounded by assumption (c); hence so is the sequence $\{\lambda_k^{-1} u^k\}$. Let $\tilde{u}^k \equiv \lambda_k^{-1} u^k$. Clearly each $\tilde{u}^k$ remains in $\mathcal{N}(\bar{x}^k; \mathcal{C})$; since

$$\|\tilde{u}^k + v^k\| \;\le\; k^{-1}(\|\tilde{u}^k\| + 1),$$

it follows that $\{\tilde{u}^k + v^k\}$ converges to zero but this contradicts the condition (15).

We next show that condition (d) implies (c). Under assumption (d), the set $\partial f(\mathcal{C} \cap f^{-1}(0))$ is bounded by Theorem 24.7 in [35]. To verify (15), assume on the contrary that there exist sequences of vectors $\{\bar{x}^k\} \subset \mathcal{C} \cap f^{-1}(0)$ and $\{(u^k, v^k)\}$ with $(u^k, v^k) \in \mathcal{N}(\bar{x}^k; \mathcal{C}) \times \partial f(\bar{x}^k)$ such that $\|u^k + v^k\| \to 0$. We have

$$(\hat{x} - \bar{x}^k)^T (u^k + v^k) \;\le\; f(\hat{x}) - f(\bar{x}^k) \;=\; f(\hat{x}).$$

The sequence $\{\bar{x}^k\}$ is bounded; hence the left-hand inner product converges to zero; but since $f(\hat{x}) < 0$, we obtain a contradiction. Hence (c) holds.

Finally, we show that (e) also implies (c). Let $\bar{x} \in \mathcal{C} \cap f^{-1}(0)$ and $(u, v) \in \mathcal{N}(\bar{x}; \mathcal{C}) \times \partial f(\bar{x})$ be arbitrary. We have $\bar{x} + \tau d \in \mathcal{C}$ for all $\tau > 0$; thus $d^T u \le 0$. Moreover, for any $\tau > 0$,

$$\tau\, d^T v \;\le\; f(\bar{x} + \tau d) - f(\bar{x}) \;\le\; \tau f_\infty(d),$$

whence $d^T v \le f_\infty(d)$. This implies that

$$\|v\| \;\le\; \frac{-f_\infty(d)}{\|d\|}.$$

Hence the set $\partial f(\mathcal{C} \cap f^{-1}(0))$ is bounded. If (15) fails to hold, then there exist sequences of vectors $\{\bar{x}^k\} \subset \mathcal{C} \cap f^{-1}(0)$ and $\{(u^k, v^k)\}$ with $(u^k, v^k) \in \mathcal{N}(\bar{x}^k; \mathcal{C}) \times \partial f(\bar{x}^k)$ such that $\|u^k + v^k\| \to 0$. We have

$$d^T (u^k + v^k) \;\le\; f_\infty(d), \quad \text{for all } d.$$

The left-hand inner product approaches zero whereas the right-hand constant is negative. Thus (c) holds. Q.E.D.

Assumption (b) in Corollary 2 is closely related to the ACQ for a system of convex inequalities studied in [30, 26, 2]. When $\mathcal{C} = \Re^n$, this assumption (b) reduces to the strong Slater assumption in Corollary 1. We will subsequently give some interesting cases where assumption (c) in Corollary 2 holds easily. Z.Q. Luo communicated to us that he was able to establish the conclusion of Corollary 2 under assumption (d) with $\mathcal{C}$ being polyhedral and

f finite valued, based on Robinson's 1975 result [34]. In general, for a closed proper convex function f, the set $f^{-1}(0)$ may not be closed (cf. e.g. Example 1); so in assumption (d) we need the closure of the set $\mathcal{C} \cap f^{-1}(0)$ (and not just the set itself) to be contained in $\operatorname{int}(\operatorname{dom}(f))$ in order to deduce the boundedness of $\partial f(\mathcal{C} \cap f^{-1}(0))$. Assumption (e) in Corollary 2 generalizes the corresponding assumption in Corollary 1.

Corollary 1 gives sufficient conditions for global error bounds to hold. In what follows, we present a result that gives a necessary condition for a (local) error bound to hold in a neighborhood of a point, i.e., for the set $\mathcal{S}$ to be metrically regular at that point. This result is inspired by the work of Li [25]. In essence, the result below says that a necessary condition for metric regularity to hold is the validity of the normal cone identity (14), provided that the cone on the right-hand side is closed.

Proposition 2 *Let $f : \Re^n \to \Re \cup \{\infty\}$ be a closed proper convex function and let $\mathcal{C}$ be a closed convex subset of $\Re^n$. Let $\mathcal{S} \equiv \mathcal{C} \cap f^{-1}(-\infty, 0]$. Let $\bar{x} \in \mathcal{C} \cap f^{-1}(0)$ and assume $\bar{x} \in \operatorname{dom}(\partial f)$. If (3) is metrically regular at $\bar{x}$, then*

$$\mathcal{N}(x; \mathcal{S}) = \operatorname{cl}(\mathcal{N}(\bar{x}; \mathcal{C}) + \Re_+ \partial f(x)). \tag{16}$$

If in addition $\bar{x} \in \operatorname{int}(\operatorname{dom}(f))$, then the closure can be dropped from the right-hand set.

Proof. Consider first the case where $\mathcal{C} = \Re^n$. In this case, it suffices to show $\mathcal{N}(\bar{x}; \mathcal{S}) \subseteq \operatorname{cl}(\Re_+ \partial f(\bar{x}))$. This clearly holds if a Slater point exists, by Theorem 23.7 in [35]. (In fact, this theorem yields $\mathcal{N}(\bar{x}; \mathcal{S}) = \Re_+ \partial f(\bar{x})$ if $\bar{x} \in \operatorname{int}(\operatorname{dom}(f))$.) So assume that no such point exists. Then $f(x) \geq 0$ for all $x \in \operatorname{dom}(f)$; thus $f(x) = f(x)_+$ for all $x \in \Re^n$. Let γ be the multiplicative constant associated with the local error bound at $\bar{x}$. Thus $\operatorname{dist}(x, \mathcal{S}) \leq \gamma \, f(x)$ for all x sufficiently close to $\bar{x}$.

Choose $y \in \mathcal{N}(\bar{x}; \mathcal{S})$. So $\bar{x}$ is a constrained global minimizer of the linear function $-y^T x$ over $x \in \mathcal{S}$. Hence by a fundamental exact penalty function result proved in [11], it follows that there exists a scalar $c > 0$ such that $\bar{x}$ is an unconstrained global minimizer of the function $x \mapsto -y^T x + c \operatorname{dist}(x, \mathcal{S})$. In particular, for all x sufficiently close to $\bar{x}$, we have

$$-y^T \bar{x} \;\leq\; -y^T x + c \operatorname{dist}(x, \mathcal{S}) \;\leq\; -y^T x + c \; \gamma f(x);$$

this implies, by convexity of f, that for all $x \in \Re^n$,

$$f(x) \;\geq\; (c \, \gamma)^{-1} y^T (x - \bar{x}).$$

Recalling $f(\bar{x}) = 0$, we conclude that $y \in (c\,\gamma)^{-1}\partial f(\bar{x})$. This establishes the inclusions $\mathcal{N}(\bar{x};\mathcal{S}) \subseteq \Re_+\partial f(\bar{x}) \subseteq \mathrm{cl}(\Re_+\partial f(\bar{x}))$.

Suppose $\mathcal{C}$ is a proper subset of $\Re^n$. Consider the function $\tilde{f}$ defined by

$$\tilde{f}(x) \equiv \max(\,\mathrm{dist}(x,\mathcal{C}),\, f(x)\,), \quad x \in \Re^n.$$

It follows from the previous case that

$$\mathcal{N}(\bar{x};\mathcal{S}) = \mathrm{cl}(\Re_+\,\partial\tilde{f}(\bar{x})).$$

Since $\partial\tilde{f}(\bar{x})$ is equal to the convex hull of $\partial f(\bar{x})$ and $\mathcal{N}(\bar{x};\mathcal{C}) \cap \mathbb{B}$, where $\mathbb{B}$ is the unit Euclidean ball in $\Re^n$ (see [21]), we have

$$\Re_+\partial\tilde{f}(\bar{x}) = \mathcal{N}(\bar{x};\mathcal{C}) + \Re_+\partial f(\bar{x}).$$

Combining the last two equations, we readily obtain the desired equality (16).

Finally, the above argument also shows that the closure can be dropped if $\bar{x} \in \mathrm{int}(\mathrm{dom}(f))$. Q.E.D.

Remark. Proposition 2 (with $\mathcal{C} = \Re^n$) is proved independently by Sien Deng. Alternatively, this proposition can also be proved by applying a standard separation theorem of convex sets, without invoking the exact penalty function result of Clarke.

4 Special cases

We discuss several special cases of the results obtained in the last section. For the most part, the results presented below are new.

4.1 Compact test sets

Our first application of part (c) of Corollary 2 is to the situation of a compact test set. We state this situation more precisely in the following result.

Proposition 3 *Let $f : \Re^n \to \Re \cup \{\infty\}$ be a closed proper convex function and let $\mathcal{C}$ be a closed convex subset of $\Re^n$. Let $\mathcal{S} \equiv \mathcal{C} \cap f^{-1}(-\infty, 0]$. Assume $\mathcal{C} \cap f^{-1}(0) \subseteq \mathrm{int}(\mathrm{dom}(f))$ and there exists a vector $\hat{x} \in \mathcal{C}$ satisfying $f(\hat{x}) < 0$. Then for every compact set T in $\Re^n$ for which $\partial f(\hat{T} \cap f^{-1}(0))$ is bounded and $\Pi_{\hat{T}}(\mathrm{dom}(f)) \subseteq \mathrm{dom}(f)$, where $\hat{T}$ is the convex hull of $\{\hat{x}\} \cup \Pi_{\mathcal{C}}(T)$, there exists a scalar $\gamma > 0$ such that*

$$\mathrm{dist}(x,\mathcal{S}) \le \gamma \max(\,\mathrm{dist}(x,\mathcal{C}),\, f(x)_+\,), \quad \textit{for all } x \in T. \qquad (17)$$

Proof. The set $\hat{T}$ is compact and convex. By following the proof of (d) implying (c) in Corollary 2, we can verify that (15) holds with $\mathcal{C}$ replaced by $\hat{T}$. Hence by this corollary, it follows that there exists a scalar $\gamma > 0$ such that for all $x \in \Re^n$,

$$\operatorname{dist}(x, \hat{\mathcal{S}}) \leq \gamma \max(f(x)_+, \operatorname{dist}(x, \hat{T})),$$

where $\hat{\mathcal{S}} \equiv \hat{T} \cap f^{-1}(\infty, 0]$. Since $\hat{\mathcal{S}}$ is a subset of $\mathcal{S}$, we have

$$\operatorname{dist}(x, \mathcal{S}) \leq \operatorname{dist}(x, \hat{\mathcal{S}}).$$

Moreover if x is in T, then $\Pi_{\mathcal{C}}(x) \in \Pi_{\mathcal{C}}(T)$; thus $\operatorname{dist}(x, \hat{T}) = \operatorname{dist}(x, \mathcal{C})$. Consequently, we have for all $x \in T$,

$$\operatorname{dist}(x, \mathcal{S}) \leq \gamma \max(\operatorname{dist}(x, \mathcal{C}), f(x)_+),$$

as desired. Q.E.D.

Ideally, we would like the above result to remain valid with the Slater assumption replaced by the normal cone identity (14). If this strengthening of Proposition 3 is indeed correct, then combining with Proposition 2, we would have obtained a characterization of error bounds holding on compact sets in terms of the identity (14); in turn, such a characterization would completely generalize the result of Li [25] for the special case where the function f is the pointwise maximum of a finite number of convex differentiable functions. Regrettably, we are presently not able to establish Proposition 3 under the weaker assumption (14) for an arbitrary closed proper convex function f.

4.2 The ice-cream cone

The second application concerns the case where the function f is very specific; namely,

$$f(x) \equiv \sqrt{\sum_{i=1}^{n-1} x_i^2} - x_n, \quad x \in \Re^n. \tag{18}$$

For this function, the set $f^{-1}(-\infty, 0]$ is the (homogeneous) convex cone commonly known as the "ice-cream" cone, which we denote S_{ice}. It is easy to see that $f^{-1}(0) = \operatorname{bd}(S_{\text{ice}})$ and

$$\partial f(f^{-1}(0)) \subseteq \mathbb{B}_{n-1} \times \{-1\},$$

where $\mathbb{B}_{n-1}$ denotes the unit Euclidean ball in $\Re^{n-1}$. Thus $\partial f(f^{-1}(0))$ is bounded. We note that

$$\operatorname{dist}(x, S_{\text{ice}}) = \begin{cases} \frac{f(x)_+}{\sqrt{2}}, & \text{if } x \notin -S_{\text{ice}}, \\ \|x\| \left(\in \left[\frac{f(x)_+}{\sqrt{2}},\, f(x)_+ \right] \right), & \text{if } x \in -S_{\text{ice}}, \end{cases}$$

which implies

$$\frac{f(x)_+}{\sqrt{2}} \leq \operatorname{dist}(x, S_{\text{ice}}) \leq f(x)_+, \quad \text{for all } x \in \Re^n;$$

so $f(x)_+$ is "equivalent" to the distance function to S_{ice}.

Proposition 4 *Let S_{ice} be the ice-cream cone in $\Re^n$ and $\mathcal{C}$ a closed convex set in $\Re^n$. Assume that $\mathcal{C} \cap \operatorname{int}(S_{\text{ice}}) \neq \emptyset$. Then for every compact set $T \subset \Re^n$, there exists a scalar $\gamma > 0$ such that for all $x \in T$,*

$$\operatorname{dist}(x, \mathcal{C} \cap S_{\text{ice}}) \leq \gamma \max \left(\operatorname{dist}(x, \mathcal{C}), \left(\sqrt{\sum_{i=1}^{n-1} x_i^2} - x_n \right)_+ \right).$$

If in addition,

$$\operatorname{dist} \left(\bigcup_{\bar{x} \in \mathcal{C} \cap \operatorname{bd}(S_{\text{ice}})} \mathcal{N}(\bar{x}; \mathcal{C}),\, (\mathbb{B}_{n-1} \times \{1\}) \right) > 0, \tag{19}$$

then the above error bound holds for $T = \Re^n$.

Proof. With f given by (18), the first conclusion follows from Proposition 3. We next show that the additional assumption (19) implies the assumption of Corollary 2(c). Once this is established, the desired global error bound (for $T = \Re^n$) follows readily from this corollary. We have noted that $\partial f(f^{-1}(0))$, and hence $\partial f(\mathcal{C} \cap f^{-1}(0))$, is bounded. Thus it remains to show that (15) holds. For any $(u, v) \in \mathcal{N}(\bar{x}; \mathcal{C}) \times \partial f(\bar{x})$, where $\bar{x} \in \mathcal{C} \cap f^{-1}(0)$ is arbitrary, we have $-v \in \mathbb{B}_{n-1} \times \{1\}$, hence

$$\|u + v\| \geq \operatorname{dist} \left(\bigcup_{\bar{x} \in \mathcal{C} \cap \operatorname{bd}(S_{\text{ice}})} \mathcal{N}(\bar{x}; \mathcal{C}),\, (\mathbb{B}_{n-1} \times \{1\}) \right).$$

Consequently (15) follows. Q.E.D.

It was pointed out by Z.Q. Luo that if the ice-cream cone $S_{\rm ice}$ is represented by the (nonhomogeneous) convex inequality,

$$\sum_{i=1}^{n-1} x_i^2 - \tau \,\leq\, 0, \quad \text{with } (x_1, \ldots, x_{n-1}, \tau) \in \Re^n,$$

and if $\mathcal{C}$ is a polyhedron in the $(x_1, \ldots, x_{n-1}, \tau)$ space, then a global error bound for the set $\mathcal{C} \cap S_{\rm ice}$ follows from the result of Luo and Luo [26], under the sole assumption that $\mathcal{C} \cap \operatorname{int}(S_{\rm ice}) \neq \emptyset$. This result, apart from the polyhedral restriction on $\mathcal{C}$, is different from Proposition 4 because an inequality system in the variables $(x_1, \ldots, x_{n-1}, \tau)$ is not of the same type as the system in the variables $(x_1, \ldots, x_{n-1}, x_n)$, under the change of variables $\tau = x_n^2$. The following example shows that the distance assumption (19) is, in general, needed for a global error bound to hold for $\mathcal{C} \cap S_{\rm ice}$ if $\mathcal{C}$ is a proper subset of $\Re^n$ (in this example, $\mathcal{C}$ is an affine subspace).

Example 4. Take $n = 3$ and $\mathcal{C} \equiv \{x \in \Re^3 : x_3 = x_1 + 1\}$. Consider the sequence of vectors $\{x^k\}$ where

$$x^k \,\equiv\, \begin{pmatrix} k \\ 1 + \sqrt{1+2k} \\ 1+k \end{pmatrix} \in \mathcal{C}, \quad k = 1, 2, \ldots.$$

As in Example 2, it can be shown that

$$f(x^k) \,=\, \sqrt{k^2 + 2 + 2k + 2\sqrt{1+2k}} - (1+k) \,>\, 0,$$

where f is the function (18), and $f(x^k) \to 0$ as $k \to \infty$; nevertheless, $\operatorname{dist}(x^k, \mathcal{C} \cap S_{\rm ice}) \to 1$ as $k \to \infty$. For this example, (19) fails to hold because $\mathcal{N}(\bar{x}; \mathcal{C}) = \{x \in \Re^3 : x_3 + x_1 = 0, x_2 = 0\}$ (for any $\bar{x} \in \mathcal{C}$) clearly intersects $\mathbb{B}_2 \times \{1\}$.

4.3 The positive semidefinite cone

The third special case concerns a positive semidefinite system. Specifically, let $\mathcal{M}^n$ denote the vector space of all $n \times n$ symmetric matrix with real entries. This space with the trace inner product is isomorphic to the finite-dimensional Euclidean space $\Re^{n(n+1)/2}$. Let $\mathcal{M}^n_+$ and $\mathcal{M}^n_{++}$ denote, respectively, the set of all $n \times n$ symmetric positive semidefinite and definite matrices. Let $\mathcal{C}$ be a given closed convex subset of $n \times n$ symmetric matrices; we are interested in obtaining a global error bound for the convex set $\mathcal{C} \cap \mathcal{M}^n_+$. Let $\lambda_{\min}(X)$ be the smallest eigenvalue of a matrix $X \in \mathcal{M}^n$. By defining the convex function

$$f(X) \equiv -\lambda_{\min}(X), \quad X \in \mathcal{M}^n,$$

which is everywhere finite valued on $\mathcal{M}^n$, we can apply the results in the last subsection. It is easy to see (for example by diagonalization) that since

$$f(X)_+ \le \operatorname{dist}(X, \mathcal{M}^n_+) \le \sqrt{n}\, f(X)_+, \quad \text{for all } X \in \mathcal{M}_+,$$

the function f_+ is thus equivalent to the distance function to $\mathcal{M}^n_+$. The set $f^{-1}(0)$ consists of the singular, symmetric, positive semidefinite matrix; we denote this set by $\mathcal{M}^n_0$. We have (see [33]):

$$\partial f(\mathcal{M}^n) = \text{ convex hull of } \{uu^T \,:\, u^T u = 1\} = \mathcal{M}^n_+(1),$$

where $\mathcal{M}^n_+(1)$ is the subset of $\mathcal{M}^n_+$ consisting of all symmetric positive semidefinite matrices with unit trace. Thus $\partial f(\mathcal{S})$ is bounded. Similar to Proposition 4, we have the following result whose proof is omitted.

Proposition 5 *Let $\mathcal{C}$ be a closed convex subset of $n \times n$ symmetric matrices. Assume that $\mathcal{C} \cap \mathcal{M}^n_{++} \neq \emptyset$. Then for every compact set $T \subset \mathcal{M}^n$, there exists a scalar $\gamma > 0$ such that for all $X \in T$,*

$$\operatorname{dist}(X, \mathcal{C} \cap \mathcal{M}^n_+) \le \gamma \max\left(\operatorname{dist}(X, \mathcal{C}), (\lambda_{\min}(X))_-\right).$$

If in addition,

$$\operatorname{dist}\left(\bigcup_{X \in \mathcal{C} \cap \mathcal{M}^n_0} \mathcal{N}(X; \mathcal{C}), \mathcal{M}^n_+(1) \right) > 0,$$

then the above error bound holds for $T = \mathcal{M}^n$.

Since the cone $\mathcal{M}^2_+$ is isomorphic to the (homogeneous) ice-cream cone in $\Re^3$, by appropriately modifying Example 4, we see that the additional distance assumption is again needed for the second conclusion of Proposition 5 to hold.

Due to independent interest, we present an error bound for a positive semidefinite system involving analytic functions, without assuming the Slater condition or convexity of these functions. Specifically, let

$$\mathcal{C} \equiv \{X \in \mathcal{M}^n \,:\, g_i(X) \le 0,\ i = 1, \ldots, m\}$$

where each $g_i : \mathcal{M}^n \to \Re$ is analytic. Note that this set $\mathcal{C}$ is not convex in general; also, although we have not explicitly included (analytic) equality constraints in the above representation of $\mathcal{C}$, the result below clearly extends to the case where such equations are present.

Proposition 6 *Let $\mathcal{C}$ be as given above. If $\mathcal{C} \cap \mathcal{M}^n_+ \neq \emptyset$, then for every compact set $T \subset \mathcal{M}^n$, there exist positive scalars γ and ρ such that for all $x \in T$,*

$$\operatorname{dist}(X, \mathcal{C} \cap \mathcal{M}^n_+) \leq \gamma \left[\max(\operatorname{dist}(X, \mathcal{C}), (\lambda_{\min}(X))_-) \right]^\rho . \tag{20}$$

Proof. The proof is divided into two main steps. In the first step, we express $\mathcal{C} \cap \mathcal{M}^n_+$ as the solution set of a finite system of inequalities defined by analytic functions; this expression allows us to use the error bound for an analytic system [27] to get a preliminary error bound for $\mathcal{C} \cap \mathcal{M}^n_+$. In the second step, we use the interlacing property of eigenvalues [22, Theorem 4.3.8] to bound from above the residual of this equivalent analytic system in terms of the desired residual in (20).

Since a symmetric matrix is positive semidefinite if and only if all its principal minors are nonnegative, it follows that a symmetric matrix X belongs to $\mathcal{C} \cap \mathcal{M}^n_+$ if and only if

$$\begin{aligned} g_i(X) &\leq 0, \quad \forall\, i = 1, \ldots, m \\ \det(X_\alpha) &\geq 0, \quad \forall\, \alpha \subseteq N \equiv \{1, \ldots, n\}, \end{aligned}$$

where det denotes the determinant of a matrix and X_α is the principal submatrix of X indexed by α. The above is a finite system of inequalities defined by analytic functions of X. Hence by Theorem 2.2 in [27], for every compact set $T \subset \mathcal{M}^n$, there exist positive scalars γ' and ρ' such that for all $X \in T$,

$$\operatorname{dist}(X, \mathcal{C} \cap \mathcal{M}^n_+) \leq$$

$$\gamma' \left[\max\left(\max\{(g_i(X))_+ \,:\, i = 1, \ldots, m\}, \max\{(\det(X_\alpha))_- \,:\, \alpha \subseteq N\} \right) \right]^{\rho'} .$$

Since T is compact and g_i is analytic, it follows that there exists a positive constant $L > 0$ such that for all $X \in T$,

$$\max\{(g_i(X))_+ \,:\, i = 1, \ldots, m\} \leq L \operatorname{dist}(X, \mathcal{C}).$$

Since T is bounded, it follows that there exists a constant $r > 0$ such that for all $X \in T$, all $\alpha \subseteq N$, and all eigenvalues λ of X_α,

$$|\lambda| \leq r.$$

Fix an arbitrary subset α of N; we derive an upper bound for $(\det(X_\alpha))_-$ in terms of $\lambda_{\min}(X)_-$. Without loss of generality, assume that $\det(X_\alpha)$ is negative. Hence $\lambda_{\min}(X_\alpha) < 0$. Let

$$\lambda_{\min}(X_\alpha) = \lambda_1 \leq \ldots \leq \lambda_k,$$

be the eigenvalues of X_α, where k is the cardinality of α. By the interlacing property of eigenvalues, we have

$$0 > \lambda_{\min}(X_\alpha) \geq \lambda_{\min}(X).$$

Consequently,

$$\det(X_\alpha) = \prod_{i=1}^{k} \lambda_i \geq \lambda_1 \prod_{i=2}^{k} |\lambda_i| \geq r^{k-1} \lambda_{\min}(X).$$

Hence

$$(\det(X_\alpha))_- \leq r^{k-1} (\lambda_{\min}(X))_-.$$

This is enough for us to obtain the desired error bound (20). Q.E.D.

As an application of Proposition 6, we derive an error bound for the solution set of a complementarity problem over the cone of symmetric positive semidefinite matrices. Specifically, we consider the problem of finding a matrix $X \in \mathcal{M}^n$ such that

$$\mathcal{M}^n_+ \ni X \perp F(X) \in \mathcal{M}^n_+, \tag{21}$$

where $F : \mathcal{M}^n \to \mathcal{M}^n$ is a given mapping from $\mathcal{M}^n$ into itself, and $X \perp Y$ means that the two matrices X and Y are perpendicular in the trace inner product. We say that the matrix-valued map F is analytic if each of its entry functions $F_{ij}(X)$ is an analytic, real-valued function. Let $\mathcal{S}$ denote the solution set of the complementarity problem (21) and let

$$r(X) \equiv \max(\,(\lambda_{\min}(X))_-,\ (\lambda_{\min}(F(X)))_-,\ |\operatorname{tr}(XF(X))|\,), \quad X \in \mathcal{M}^n,$$

be the residual function of this problem, where "tr" is a shorthand for trace.

Corollary 3 *Let $F : \mathcal{M}^n \to \mathcal{M}^n$ be an analytic map. Assume $\mathcal{S} \neq \emptyset$. For every compact subset $T \subset \mathcal{M}^n$, there exist positive scalars γ and ρ such that for every $X \in T$,*

$$\operatorname{dist}(X, \mathcal{S}) \leq \gamma\, r(X)^\rho.$$

Proof. Clearly, a matrix X belongs to $\mathcal{S}$ if and only if the pair $(X, Y) \in \mathcal{M}^n \times \mathcal{M}^n$ satisfies the following analytic system in positive semidefinite matrices:

$$Y - F(X) = 0, \quad \operatorname{tr}(XY) = 0, \quad (X, Y) \in \mathcal{M}^n_+ \times \mathcal{M}^n_+.$$

The desired conclusion now follows easily from this observation and Proposition 6. Q.E.D.

4.4 Convex differentiable inequalities

The last application of the results in Section 3 concerns an inequality system defined by finitely many convex differentiable functions. This case has been treated extensively in the references cited at the end of this paper. We give here a necessary and sufficient condition for a global error bound to hold for such a system.

Specifically, for a given positive integer m and for each $i = 1, \ldots, m$, let $g_i : \Re^n \to \Re$ be a convex, continuously differentiable function. Define

$$\mathcal{S} \equiv \bigcap_{i=1}^{m} g_i^{-1}(-\infty, 0],$$

which we assume is nonempty. Associated with this set, let

$$f(x) \equiv \max_{1 \le i \le m} g_i(x), \quad x \in \Re^n. \tag{22}$$

Then f is a finite-valued, convex function on $\Re^n$; and $f^{-1}(-\infty, 0] = \mathcal{S}$. Thus Theorem 1 is applicable. For each $x \in \Re^n$, let

$$\mathcal{I}(x) \equiv \{ i : g_i(x) = f(x) \}.$$

It is well-known that for all $x \in \Re^n$ $\partial f(x)$ is equal to the convex hull of the gradients

$$\{ \nabla g_i(x) : i \in \mathcal{I}(x) \}.$$

Let $A(x)$ denote the $n \times |\mathcal{I}(x)|$ matrix with these gradients as its columns. For each $x \in \mathcal{S}$, we have

$$\mathcal{N}(\bar{x}; \mathcal{S}) \supseteq \left\{ \sum_{i \in \mathcal{I}(\bar{x})} \lambda_i \nabla g_i(\bar{x}) : \lambda_i \ge 0, \, i \in \mathcal{I}(\bar{x}) \right\}. \tag{23}$$

Moreover, for all vectors $d \in \Re^n$,

$$f'(x;d) = \max_{i \in \mathcal{I}(x)} \nabla g_i(x)^T d.$$

By postulating that equality holds in (23), i.e., assuming the Abadie constraint qualification, we obtain the following result which follows immediately from Theorem 1. In view of the above discussion, no proof is required for this result. Notice that the Slater condition is not assumed in this result.

Proposition 7 *For $i = 1, \ldots, m$, let $g_i : \Re^n \to \Re$ be a convex, continuously differentiable function. Let*

$$\mathcal{S} \equiv \{x \in \Re^n : g_i(x) \leq 0, \; i = 1, \ldots, m\}. \tag{24}$$

Assume that $\mathcal{S} \neq \emptyset$ and equality holds in (23) for all $\bar{x} \in \mathcal{S}$ with $g_i(\bar{x}) = 0$ for at least one i. For any constant $\gamma > 0$,

$$\operatorname{dist}(x, \mathcal{S}) \leq \gamma \max_{1 \leq i \leq m} (g_i(x))_+, \quad \textit{for all } x \in \Re^n \tag{25}$$

if and only if

$$\inf_{\bar{x} \in f^{-1}(0)} \max_{i \in \mathcal{I}(\bar{x})} \left\{ \nabla g_i(\bar{x})^T A(\bar{x}) \lambda \; : \; \lambda \in \Re_+^{|\mathcal{I}(\bar{x})|}, \; \| A(\bar{x}) \lambda \|_2 = 1 \right\} \geq \gamma^{-1}.$$

As a final result of this paper, we give a sufficient condition for the above set $\mathcal{S}$ to satisfy a global error bound under a "strong Slater" constraint qualification introduced recently by Mangasarian [31]. This result does not require the functions g_i to be differentiable; it is first proved in the reference by a multiplier approach. Our purpose of including the result is to illustrate how it can readily be proved by invoking Corollary 1.

Proposition 8 *Let $\mathcal{S}$ be given by (24) where each g_i is a finite convex function. Assume that*

$$\operatorname{int}(\mathcal{S}) \equiv \{x \in \Re^n : g_i(x) < 0, \; i = 1, \ldots, m\}$$

is nonempty and there exists a scalar $\delta > 0$ such that

$$\forall \, \bar{x} \in \partial \mathcal{S}, \; \exists \, \bar{z} \in \operatorname{int}(\mathcal{S}) : \frac{\| \bar{x} - \bar{z} \|}{\min_{1 \leq i \leq m} -g_i(\bar{z})} \leq \delta. \tag{26}$$

Then there exists a constant $\gamma > 0$ such that (25) holds.

Proof. Define f by (22). We claim that (26) implies $0 \notin \mathrm{cl}(\partial f(f^{-1}(0)))$; thus the desired error bound follows from Corollary 1. Assume that the claim is false. Thus there exist sequences $\{x^k\}$ and $\{a^k\}$ such that

$$\lim_{k \to \infty} a^k \,=\, 0,$$

and for each k, $f(x^k) = 0$ and $a^k \in \partial f(x^k)$. By assumption, let $z^k \in \mathrm{int}(\mathcal{S})$ be such that

$$\|\, x^k - z^k \,\| \,\leq\, \delta \min_{1 \leq i \leq m} -g_i(z^k),$$

or equivalently,

$$-f(z^k) \,\geq\, \delta^{-1} \,\|\, x^k - z^k \,\|.$$

Since $a^k \in \partial f(x^k)$, we have

$$f(z^k) \,=\, f(z^k) - f(x^k) \,\geq\, (a^k)^T(z^k - x^k) \,\geq\, -\|a^k\| \,\|\, z^k - x^k \,\|,$$

which implies that $\|a^k\| \geq \delta^{-1}$ for all k; a contradiction. Q.E.D.

5 Conclusions and Open Questions

In this paper, we have developed a systematic treatment of error bounds for convex inequality systems and related problems. We conclude the paper with several open issues that we have not been able to resolve at this time. One issue is finding a useful converse of Proposition 2 (characterizing metric regularity via the normal cone identity), applicable to our main models. Another issue concerns the determination (or estimation) of the exponent ρ in the error bounds of positive semidefinite systems failing the Slater assumption ; cf. Proposition 6 and Corollary 3. Presently, we have no idea of what this exponent ought to be except in trivial cases. Finally, it would perhaps be useful to be able to compute (or obtain upper bounds for) the multiplicative constant γ throughout the error bounds. Our analysis suggests that for the general system (3), this constant can be obtained by considering the optimization problem:

$$\inf_{\bar{x} \in \mathcal{C} \cap f^{-1}(0)} \; \sup_{d \in \mathcal{N}(\bar{x};\mathcal{S}),\, \|d\|=1} f'(\bar{x}; d).$$

Since the directional derivative $f'(\bar{x}; d)$ is typically not even continuous in its first argument, the above optimization problem is expected to be very difficult. This difficulty is compounded by the presence of the normal cone

$\mathcal{N}(\bar{x};\mathcal{S})$ which also depends on $\bar{x}$. Nevertheless, good estimates of γ could perhaps be derived in interesting applications by understanding more about this directional derivative and the normal cone $\mathcal{N}(\bar{x};\mathcal{S})$ as the vector $\bar{x}$ varies.

At the completion of this paper, the authors receive the preprint [19] in which a novel application of error bounds to the accurate identification of active constraints in nonlinear programming is described. Further applications of error bounds to mathematical programs in general deserve to be investigated and can be expected to be fruitful.

Acknowledgements. The authors gratefully acknowledge several stimulating discussions with Professors Zhi-Quan Luo and Sien Deng on the subject of this paper. In particular, Professor Luo's valuable insights and constructive comments have let us gain better understanding of the results obtained herein. We thank Professor Michael Ferris who reminded us of his work on weak sharp minima and the connection between our Theorem 1 and his joint results with Professor James Burke. We gratefully acknowledge Professor Chi-Kwong Li who helped us with the second step in the proof of Proposition 6. We also thank Professor Jon Borwein for his suggestions that led to the elegant proof of Proposition 2. Finally, we thank Professor Olvi Mangasarian for making available to us his paper [31] which is completed at about the same time as our work.

References

[1] J. Abadie, "On the Kuhn-Tucker theorem", in J. Abadie, ed., *Nonlinear Programming*, North Holland Publishing Company, Amsterdam (1967), pp. 21–36.

[2] A. Auslender and J.P. Crouzeix, "Global regularity theorems", *Mathematics of Operations Research* 13 (1988) 243-253.

[3] A. Auslender and J.P. Crouzeix, "Well-behaved asymptotical convex functions", *Analyse Non-Linéaire*, Gauthiers-Villars, Paris (1989) pp. 101-122.

[4] A. Auslender, R. Cominetti and J.P. Crouzeix, "Convex functions with unbounded level sets and applications to duality theory", *SIAM Journal on Optimization* 3 (1993) 669-687.

[5] H.H. Bauschke and J.M. Borwein, "On projection algorithms for solving convex feasibility problems", *SIAM Review* 38 (1996).

[6] H.H. Bauschke, J.M. Borwein and A.S. Lewis, "On the method of cyclic projections for convex sets in Hilbert Space", Technical report, Center for Experimental & Constructive Mathematics, Simon Fraser University, Victoria (February, 1994).

[7] J.V. Burke, "An exact penalization viewpoint of constrained optimization", *SIAM Journal on Control and Optimization* 29 (1991) 968–998.

[8] J.V. Burke and M.C. Ferris, "Weak sharp minima in mathematical programming", *SIAM Journal on Control and Optimization* 31 (1993) 1340–1359.

[9] J.V. Burke and P. Tseng, "A unified analysis of Hoffman's bound via Fenchel duality", *SIAM Journal on Optimization* 6 (1996) 265–282.

[10] C.C. Chou, K.F. Ng, and J.S. Pang, "Minimizing and stationary sequences of optimization problems", manuscript, Department of Mathematical Sciences, The Johns Hopkins University, Baltimore (September 1995).

[11] F.H. Clarke, *Optimization and Nonsmooth Analysis*, John Wiley, New York (1983).

[12] R.W. Cottle, J.S. Pang, and R.E. Stone, *The Linear Complementarity Problem*, Academic Press, Boston (1992).

[13] S. Deng, "Computable error bounds for convex inequality systems in reflexive Banach Spaces", *SIAM Journal on Optimization*, forthcoming.

[14] S. Deng, "Perturbation analysis of a condition number for convex systems", manuscript, Department of Mathematics, Northern Illinois University, DeKalb (revised September 1995).

[15] S. Deng, "Global error bounds for convex inequality systems in Banach Spaces", manuscript, Department of Mathematics, Northern Illinois University, DeKalb (October 1995).

[16] S. Deng and H. Hu, "Computable error bounds for semidefinite programming", manuscript, Department of Mathematics, Northern Illinois University, DeKalb.

[17] M.C. Ferris and O.L. Mangasarian, "Minimum principle sufficiency", *Mathematical Programming* 57 (1992) 1–14.

[18] M.C. Ferris and J.S. Pang, "Nondegenerate solutions and related Concepts in affine variational inequalities", *SIAM Journal on Control and Optimization* 34 (1996) 244–263.

[19] F. Facchinei, A. Fischer, and C. Kanzow, "On the accurate identification of active constraints", manuscript, Dipartimento di Informatica e Sistemistica, Università di Roma "La Sapienza", Rome (July 1996).

[20] A.J. Hoffman, "On approximate solutions of systems of linear inequalities", *Journal of Research of the National Bureau of Standards* 49 (1952) 263-265.

[21] J.P. Hiriart-Urruty and C. Lemaréchal, *Convex Analysis and Minimization Algorithms, I & II*, Springer-Verlag, Berlin (1993).

[22] R.A. Horn and C.A. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge (1985).

[23] D. Klatte, "Lipschitz stability and Hoffman's error bounds for convex inequality systems", manuscript, Institut für Operations Research, Universität Zürich (1995).

[24] W. Li, "Error bounds for piecewise convex quadratic programs and applications", *SIAM Journal on Control and Optimization* 33 (1995) 1510-1529.

[25] W. Li, "Abadie's constraint qualification, metric regularity, and error bounds for differentiable convex inequalities", manuscript, Department of Mathematics and Statistics, Old Dominon University, Norfolk (1995).

[26] X.D. Luo and Z.Q. Luo, "Extensions of Hoffman's error bound to polynomial systems", *SIAM Journal on Optimization* 4 (1994) 383-392.

[27] Z.Q. Luo and J.S. Pang, "Error bounds for analytic systems and their applications", *Mathematical Programming* 67 (1995) 1-28.

[28] Z.Q. Luo, J.S. Pang, and D. Ralph, *Mathematical Programs With Equilibrium Constraints*, Cambridge University Press, Cambridge (1996), forthcoming.

[29] Z.Q. Luo, J.S. Pang, D. Ralph, and S.Q. Wu, "Exact penalization and stationarity conditions of mathematical programs with equilibrium constraints", *Mathematical Programming*, forthcoming.

[30] O.L. Mangasarian, "A condition number for differentiable convex inequalities", *Mathematics of Operations Research* 10 (1985) 175-179.

[31] O.L. Mangasarian, "Error bounds for nondifferentiable convex inequalities under a strong Slater constraint qualification", Mathematical Programming Technical Report 96-04, Computer Sciences Department, University of Wisconsin, Madison, Wisconsin (July 1996).

[32] O.L. Mangasarian and T.H. Shiau, "Error bounds for monotone linear complementarity problems", *Mathematical Programming* 36 (1986) 81-89.

[33] M.L. Overton, "On minimizing the maximum eigenvalue of a symmetric matrix", *SIAM Journal on Matrix Analysis and Applications* 9 (1988) 256-268.

[34] S.M. Robinson, "An application of error bounds for convex programming in a linear space", *SIAM Journal on Control* 13 (1975) 271-273.

[35] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton (1970).

[36] T. Wang and J.S. Pang, "Global error bounds for convex quadratic inequality systems", *Optimization* 31 (1994) 1-12.