Math. Program., Ser. B (2009) 116:37–56
DOI 10.1007/s10107-007-0130-8

FULL LENGTH PAPER

Weak sharp minima revisited, Part III: error bounds for differentiable convex inclusions

James V. Burke · Sien Deng

Received: 15 August 2005 / Accepted: 4 April 2006 / Published online: 28 April 2007

Abstract The notion of weak sharp minima unifies a number of important ideas in optimization. Part I of this work provides the foundation for the theory of weak sharp minima in the infinite-dimensional setting. Part II discusses applications of these results to linear regularity and error bounds for nondifferentiable convex inequalities. This work applies the results of Part I to error bounds for differentiable convex inclusions. A number of standard constraint qualifications for such inclusions are also examined.

Keywords Weak sharp minima · Convex inclusion · Affine convex inclusion · Constraint qualification · Error bounds · Calmness

Mathematics Subject Classification (2000) 90C25 · 90C31 · 49J52

We dedicate this paper to Professor A. Auslender on the occasion of his 65th birthday. We, and the optimization community at large, have greatly profited from the deep insight and intuition Professor Auslender has brought to the subject over his many years of his service.

J. V. Burke's research was supported in part by the National Science Foundation Grant No. DMS-0505712.

J. V. Burke (✉)
Department of Mathematics, University of Washington, Box # 354350,
Seattle, WA 98195–4350, USA
e-mail: burke@math.washington.edu

S. Deng
Department of Mathematical Sciences, Northern Illinois University,
DeKalb, IL 60115, USA
e-mail: deng@math.niu.edu

1 Introduction

We continue our study of weak sharp minima by focusing on applications to error bounds for differentiable convex inclusions. Given a Gateaux differentiable mapping $h\colon X \mapsto Y$ between normed linear spaces X and Y consider the problem of finding points $x \in X$ satisfying the inclusion

$$h(x) \in C, \quad \text{where } C \subset Y \text{ is non-empty, closed, and convex} \tag{1}$$

and h is concave with respect to the recession cone of C, i.e.,

$$h((1-\lambda)x + \lambda z) - [(1-\lambda)h(x) + \lambda h(z)] \in C^{\infty} \text{ for all } x, z \in X \text{ and } \lambda \in [0, 1], \tag{2}$$

where

$$C^{\infty} = \{ d \mid x + d \in C\ \forall x \in C \}.$$

We call such problems *convex inclusions*. When C is a cone, the study of convex inclusions and associated error bounds goes back to the early work of Robinson [37–39] and Ursescu [43]. Most results for general convex inclusions can be obtained from this case by a standard lifting technique [8, Sect. 5], but not always [15, Sect. 6]. The prototype convex inclusion is obtained by taking $Y = \mathbb{R}^m$ and $C = \mathbb{R}^s_- \times \{0\}^{m-s}$. The inclusion (1) is then equivalent to a finite system of inequalities and equations $h_i(x) \leq 0$, for $i = 1, 2, \ldots, s$ and $h_i(x) = 0$, for $i = s+1, \ldots, m$, where h_i denotes the ith component function of h. Here the concavity condition on h reduces to the requirement that each h_i is convex for $i = 1, 2, \ldots, s$ and affine for $i = s+1, \ldots, m$.

Denote the solution set for (1) by

$$\Sigma = \{ x \mid h(x) \in C \},$$

and assume that $\Sigma \neq \emptyset$ throughout the paper. Let $\|\cdot\|$ denote the norm on the underlying space. We say that the inclusion (1) satisfies a *local error bound* at a point $\bar{x} \in \Sigma$ if there exists a neighborhood U of $\bar{x}$ and an $\alpha > 0$ such that

$$\alpha \operatorname{dist}(x \mid \Sigma) \leq \operatorname{dist}(h(x) \mid C) \quad \forall\, x \in U, \tag{3}$$

where $\operatorname{dist}(z \mid \tilde{C}) = \inf \left\{ \|z - s\| \,\middle|\, s \in \tilde{C} \right\}$ for any subset $\tilde{C}$ of a normed linear space. If $U = X$, then this inequality is said to be a *global error bound* for (1). The parameter α in (3) is called the *modulus* of the error bound.

If h is continuous, the existence of a local error bound is equivalent to the notion of calmness [42] of the constraint set-valued mapping $\tilde{\Sigma}$ from Y into X at $(0, \bar{x})$ where

$$\tilde{\Sigma}(y) = \{ x \mid h(x) + y \in C \} \tag{4}$$

and $\tilde{\Sigma}(0) = \Sigma$, as observed by Henrion and Outrata [23, pp. 438]. For more on the calmness property and its role in variational analysis, see [9,16,21,22,42] and references therein. The existence of a local error bound is also related to the notion of metric regularity for constraint systems [42, Chap. 9]. Indeed, in some references the existence of a local error bound is called metric regularity [30]. In [7] this notion is referred to as *weak metric regularity* and in [18] it is called *metric subregularity*. We use the term "local error bound" in order to more closely align our study with the large body of work on error bounds for systems of convex equations and inequalities (e.g., see [3,6,11,15,17,25–31,33–35,38,46,47] and Sect. 4.3 of [4]). Nondifferentiable convex inequalities are studied in [3,6,28,29,35] and in Part II [11] of this work.

Infinite systems of convex inequalities can be studied within the framework of (1) using convex multi-function methods [37–39,43] as well as exact penalization [8] and weak sharp minima [10] techniques. More recently such systems have also been studied in [46] and [47] which contain several old and some new results.

In Sect. 2 we begin by quickly reviewing the necessary tools from Part I on weak sharp minima. The body of the results are contained in Sect. 3. We first recall the basic results from [8] on pairs (h, C) satisfying (2). The fundamental subdifferential characterization theorem for error bounds is then established as an immediate consequence of results in Part I [10]. As a result, we obtain a characterization of the *calmness* of the set-valued mapping $\tilde{\Sigma}$ at a point. We then focus on the hypotheses that imply the conditions established in the subdifferential characterization hold. These hypotheses are called *constraint qualifications*. We start with the case where C is assumed to be a polyhedral set in finite dimensional space. In this case we easily recover Hoffman's bound [25] and Li's result [30] on the equivalence of the Abadie constraint qualification and the existence of a local error bound as special cases. We then examine a spectrum of constraint qualifications appearing in the literature and establish the relationships between them. We conclude with a few observations on affine convex inclusions establishing a number of verifiable sufficient conditions under which an affine convex inclusion has a local or global error bound.

The notation that we employ is consistent with that used in Parts I and II, and is for the most part the same as that in [2,19,40,41]. A partial list is provided below for the reader's convenience.

Denote the dual space of X by X^*. When X is endowed with the weak topology and X^* with the weak* topology then the spaces X and X^* are said to be *paired in duality* by the continuous bi-linear form $\langle x^*, x\rangle = x^*(x)$ defined on $X^* \times X$ [41]. Denote the norm on X^* by $\|\cdot\|_\circ$: $\|z\|_\circ = \sup_{x\in\mathbb{B}}\langle z,\, x\rangle$, where $\mathbb{B} = \{x \in X \mid \|x\| \le 1\}$ is the unit ball in X. We will use the notation $\mathbb{B}$ for the unit ball of whatever space we are discussing. If there is a possibility of confusion, we will write $\mathbb{B}_Z$ for the unit ball in the normed linear space Z. Given a set $\tilde{C}$ in either X or X^*, the set $cl\left(\tilde{C}\right)$ is the closure of this set in the norm topology, and given a set E in X^*, the set $cl^*(E)$ is the closure in the weak* topology.

For a non-empty subset $\tilde{C}$ of any normed linear space Y, denote the indicator function of $\tilde{C}$ and the support function of $\tilde{C}$ by $\psi_{\tilde{C}}(\cdot)$ and $\psi^*_{\tilde{C}}(\cdot)$, respectively. Thus, in particular, $\|z\|_o = \psi^*_{\mathbb{B}}(z)$. The barrier cone of a convex set $\tilde{C}$ is the set bar $\left(\tilde{C}\right) =$

dom $\left(\psi^*_{\tilde{C}}(\cdot)\right)$. The lineality of $\tilde{C}$, denoted by lin $\left(\tilde{C}\right)$ is the smallest subspace L such that $x + L \subset \tilde{C}$ for all $x \in \tilde{C}$. Clearly, lin $\left(\tilde{C}\right) = $ lin $\left(\tilde{C}^{\infty}\right)$. The norm-topology interior of $\tilde{C}$ is int $\left(\tilde{C}\right)$, and the boundary of $\tilde{C}$ is bdry $(\tilde{C}) = cl\ \left(\tilde{C}\right) \backslash \mathrm{int}\ \left(\tilde{C}\right)$. When Y is finite-dimensional, ri $\left(\tilde{C}\right)$ is the interior of $\tilde{C}$ relative to the smallest affine set containing $\tilde{C}$. The cone generated by $\tilde{C}$ is cone $(\tilde{C}) = \cup_{\lambda \geq 0}\{\lambda \tilde{C}\}$. The linear span of $\tilde{C}$ is denoted by span $\left(\tilde{C}\right)$, and the affine hull of $\tilde{C}$ is denoted by aff $(\tilde{C})$.

Define the *projection* of a point $x \in X$ onto the set $\tilde{C}$, denoted $P(x \mid \tilde{C})$, as the set of all points in $\tilde{C}$ that are closest to x as measured by the norm $\|\cdot\|$: $P(x \mid \tilde{C}) = \left\{ y \in \tilde{C} \;\middle|\; \|x - y\| = \mathrm{dist}\,(x \mid \tilde{C}) \right\}$. For non-empty sets $\tilde{C} \subset X$ and $E \subset X^*$, the polar of $\tilde{C}$ and E are given, respectively, by the sets $\tilde{C}^\circ = \{x^* \in X^* \mid \langle x^*, x\rangle \leq 1\ \forall x \in \tilde{C}\}$, $E^\circ = \{x \in X \mid \langle x^*, x\rangle \leq 1\ \forall x^* \in E\}$, respectively. Thus, in particular, $\mathbb{B}^\circ \subset X^*$ is the unit ball associated with the dual norm $\|\cdot\|_o$. If either $\tilde{C}$ or E is a subspace, we also write $\tilde{C}^\circ = \tilde{C}^\perp$ and $E^\circ = E^\perp$. For a non-empty closed convex set $\tilde{C}$ in X, and $x \in \tilde{C}$, define the tangent cone to $\tilde{C}$ at x by $T_{\tilde{C}}\,(x) = cl\ \left(\bigcup_{t>0} \frac{\tilde{C}-x}{t}\right)$. The normal cone to $\tilde{C}$ at x is given by $N_{\tilde{C}}\,(x) = T_{\tilde{C}}\,(x)^\circ$. It is easy to see that $N_{\tilde{C}}\,(x) = \left\{x^* \in X^* | \langle x^*, y - x\rangle \leq 0,\ \ \text{for any } y \in \tilde{C}\right\}$.

Given a linear operator $T\colon X \mapsto Y$ between topological vector spaces X and Y, we denote the kernel (or null-space) of T and the range of T by Nul (T) and Ran (T), respectively.

Let $f\colon X \mapsto \overline{\mathbb{R}}$ be a lower semi-continuous convex function. The function $f^*\colon X^* \mapsto \overline{\mathbb{R}}$ defined by $f^*(x^*) = \sup_{x \in X}(\langle x^*, x\rangle - f(x))$ is called the convex conjugate of f. The subdifferential of f at x and the directional derivative of f at x in the direction d are denoted by $\partial f(x)$ and $f'(x; d)$ respectively.

2 Weak sharp minima

Let X be a normed linear space, $S \subset X$ a non-empty closed convex set, and $f\colon X \mapsto \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ a lower semi-continuous convex function. We assume that $S \cap$ *dom* $(f) \neq \emptyset$, where *dom* $(f) = \{x \in X \mid f(x) < \infty\}$, and

$$\bar{S} = \left\{ x \in S \;\middle|\; f(x) = \min_{y \in S} f(y) \right\} \neq \emptyset.$$

The set $\bar{S} \subset X$ is said to be a set of *weak sharp minima* for the function f over the set S with modulus $\alpha > 0$ if

$$f(\bar{x}) + \alpha \mathrm{dist}\,(x \mid \bar{S}) \leq f(x) \quad \text{for all } \bar{x} \in \bar{S} \text{ and } x \in S, \tag{5}$$

where $\mathrm{dist}\,(x \mid \bar{S}) = \inf_{\bar{x} \in \bar{S}} \|x - \bar{x}\|$, and $\|\cdot\|$ is the norm on X. The set $\bar{S}$ is said to be a set of *weak sharp minima at* $\bar{x} \in \bar{S}$ for f over the set S if there exists an $\epsilon > 0$ and $\alpha > 0$ such that

$$f(\bar{x}) + \alpha \mathrm{dist}\,(x \mid \bar{S}) \leq f(x) \quad \text{for all } x \in S \cap (\bar{x} + \epsilon \mathbb{B}).$$

Theorem 1 [10, Theorem 2.3 and Theorem 5.2] *Let f, S, and $\bar{S}$ be as in (5), and assume that the addition formula*

$$\partial(f + \psi_S)(x) = cl^* \left(\partial f(x) + N_S(x)\right), \tag{6}$$

holds for all $x \in \bar{S}$. Let $\alpha > 0$. Then the following are true.
(A) *The set $\bar{S}$ is a set of weak sharp minima for the function f over the set $S \subset X$ with modulus α if and only if the normal cone inclusion*

$$\alpha \mathbb{B}^\circ \bigcap N_{\bar{S}}(x) \subset cl^* \left(\partial f(x) + N_S(x)\right) \tag{7}$$

holds for all $x \in \bar{S}$.
(B) *Assume X is either Hilbert or finite-dimensional. Then $\bar{S}$ is a set of weak sharp minima at $\bar{x} \in \bar{S}$ for f over the set S if and only if there exist some $\alpha > 0$ and $\epsilon > 0$ such that (7) holds for all $x \in \bar{S} \cap int(\bar{x} + \epsilon \mathbb{B})$.*

It is shown in Part I that the normal cone inclusion (7) can be decomposed into two independent conditions. These conditions play a pivotal role in connecting the notion of weak sharp minima to a number of related ideas in the literature.

Lemma 1 [10, Lemma 3.1] *Let the basic assumptions of Theorem 1 hold. Given $x \in \bar{S}$, we have*

$$\alpha \mathbb{B}^\circ \cap N_{\bar{S}}(x) \subset cl^* \left(\partial f(x) + N_S(x)\right) \tag{8}$$

if and only if

$$\begin{aligned} &\mathrm{cone}\left(cl^*\left(\partial f(x) + N_S(x)\right)\right) = N_{\bar{S}}(x) \qquad \textit{and} \\ &\alpha \mathbb{B}^\circ \cap \left[\mathrm{cone}\left(cl^*\left(\partial f(x) + N_S(x)\right)\right)\right] \subset cl^*\left(\partial f(x) + N_S(x)\right). \end{aligned} \tag{9}$$

In addition, if the set $\partial f(x) + N_S(x)$ is weak* closed, then

$$\mathrm{cone}\left(cl^*\left(\partial f(x) + N_S(x)\right)\right) = \mathrm{cone}\left(\partial f(x)\right) + N_S(x).$$

The history and motivations for the study of weak sharp minima are well documented in [10,12,13,20,36,45], where one can find extensive references on weak sharp minima and related issues.

3 Error bounds for differentiable convex inclusions

Let X, Y, h, C, and Σ be as in (1) and (2). Observe that the inclusion (1) has an error bound at a point $\bar{x}$ if and only if the set Σ is a set of weak sharp minima for the function $f(x) = \mathrm{dist}\,(h(x) \mid C)$. Our goal is to apply the characterization results in Theorem 1 for weak sharp minima to study the convex inclusion (1) and in this way obtain error bound results.

Lemma 2 [8, Lemma 5.1] *The following statements are equivalent*:

1. $h\colon X \mapsto Y$ *is concave with respect the recession cone of* C.
2. *For each* $y^* \in \text{bar }(C)$, *the mapping* $h_{y^*}\colon X \mapsto \mathbb{R}$ *defined by* $h_{y^*}(x) = \langle y^*, h(x)\rangle$ *is a convex function of* x.

Moreover, either one of these statements implies that the function $f(x) = \text{dist}\,(h(x) \mid C)$ *is convex.*

We now focus on Gateaux differentiable convex inclusions.

Definition 1 Let $L(X, Y)$ denote the space of bounded linear transformations from X to Y with norm given by $\|A\| = \sup\{\,\|Ax\| \mid \|x\| \le 1\,\}$. We say that the mapping $h\colon X \mapsto Y$ is Gateaux differentiable at $x \in X$ if there exists $A \in L(X, Y)$ such that

$$\lim_{\lambda\searrow 0} \frac{1}{\lambda}\,\|h(x + \lambda d) - h(x) - \lambda A d\| = 0,$$

for every $d \in X$. Denote the transformation A by $h'(x)$. We say that h is continuously Gateaux differentiable on an open subset U of X if the mapping from X to $L(X, Y)$ given by $x \mapsto h'(x)$ is well-defined and continuous.

Gateaux differentiability at a point $x \in X$ can be expressed equivalently by the equation

$$h(x + \lambda d) = h(x) + \lambda h'(x)d + o(\lambda),$$

where the function $o(\lambda)$ depends on both x and d, and satisfies $\lim_{\lambda\searrow 0} \frac{o(\lambda)}{\lambda} = 0$. We list a few important consequences of differentiability for convex inclusions.

Lemma 3 *Consider the convex inclusion* (1) *and* (2) *where it is further assumed that the mapping* $h\colon X \mapsto Y$ *is Gateaux differentiable at the point* $\bar{x} \in X$. *Let* $\rho\colon Y \mapsto \mathbb{R}$ *denote the distance function* $\rho(y) = \text{dist}\,(y \mid C)$, *and let* $f\colon X \mapsto \mathbb{R}$ *be the convex function* $f(x) = \rho \circ h$.

1. *For every* $x \in X$, $h(\bar{x}) + h'(\bar{x})(x - \bar{x}) - h(x) \in C^{\infty}$.
2. $\Sigma \subset \left\{x \in X \;\middle|\; h(\bar{x}) + h'(\bar{x})(x - \bar{x}) \in C\right\}$.
3. *If* $\bar{x} \in \Sigma$, *then* $T_{\Sigma}(\bar{x}) \subset h'(\bar{x})^{-1} T_C(h(\bar{x}))$ *and* $h'(\bar{x})^* N_C(h(\bar{x})) \subset N_{\Sigma}(\bar{x})$.
4. *For all* $d \in X$ *we have* $f'(\bar{x}; d) = \rho'(h(\bar{x}); h'(\bar{x})d)$.
5. $\partial f(\bar{x}) = h'(\bar{x})^*(\mathbb{B}^{\circ} \cap N_C(h(x)))$.

Proof 1. Since $h(\bar{x} + \lambda(x - \bar{x})) - h(\bar{x}) - \lambda[h(x) - h(\bar{x})] \in C^{\infty}\ \forall\, \lambda \in [0, 1]$, we have $h(\bar{x}) + \frac{h(\bar{x}+\lambda(x-\bar{x})) - h(\bar{x})}{\lambda} - h(x) \in C^{\infty}\ \forall\, \lambda \in (0, 1]$. The result follows by taking the limit as $\lambda \searrow 0$.
2. This follows immediately from Part 1 and the definition of the recession cone.
3. This is an immediate consequence of Part 2 and [2, Theorem 16, pp. 174].

4. Since ρ is globally Lipschitz with Lipschitz constant 1, we have

$$\mid f(\bar{x}+\lambda d)-\rho(h(\bar{x})+\lambda h'(\bar{x})d)\mid \leq \mid h(\bar{x}+\lambda d)-(h(\bar{x})+\lambda h'(\bar{x})d)\mid = o(\lambda),$$

from which the result follows.
5. This follows immediately from [10, Theorem A.1, Parts 4 and 5]. □

The normal and tangent cone inclusions in Part 3 of this lemma can be strict. A condition under which equality holds is called a *constraint qualification* (CQ). The straightforward assumption that equality does hold is called the *Abadie constraint qualification* [1]. The Abadie CQ is closely related to the condition (9) in Lemma 1. Indeed, if we take $\bar{S}=\Sigma$ and $f(x)=\operatorname{dist}(h(x)\mid C)$ in this lemma, then these conditions are equivalent yielding the following restatement of Lemma 1 in this context. The proof is omitted since it follows immediately from Lemma 1 and the representation for the subdifferential of the distance function $\operatorname{dist}(h(x)\mid C)$ provided in Part 5 of Lemma 3.

Lemma 4 *Let the hypotheses of Lemma 3 hold and let $\alpha>0$. If $x\in\Sigma$ is a point at which h is Gateaux differentiable, then*

$$\alpha(\mathbb{B}_X)^\circ\cap N_\Sigma(x)\subset h'(x)^*((\mathbb{B}_Y)^\circ\cap N_C(h(x))) \tag{10}$$

if and only if

$$N_\Sigma(x)=h'(x)^*N_C(h(x)) \qquad \text{and} \tag{11}$$
$$\alpha(\mathbb{B}_X)^\circ\cap h'(x)^*N_C(h(x))\subset h'(x)^*((\mathbb{B}_Y)^\circ\cap N_C(h(x)))\,. \tag{12}$$

We use this lemma to obtain a characterization for the existence of an error bound for the inclusion (1).

Theorem 2 *Let the hypotheses of Lemma 3 hold and assume that h is Gateaux differentiable at every point of Σ. Then the following are true.*

1. *A global error bound holds for (1) with $\alpha>0$ if and only if (11) and (12) hold for all $x\in\Sigma$.*
2. *If it is further assumed that X is a Hilbert space, then (11) and (12) hold for all $x\in\Sigma\cap\operatorname{int}(\bar{x}+\epsilon\mathbb{B})$ for some $\epsilon>0$ if and only if the inclusion (1) has a local error bound at $\bar{x}$.*
3. *If both X and Y are assumed to be finite dimensional, then there exists $\epsilon>0$ and $\alpha_1>0$ such that both (11) and (12) hold for $\alpha=\alpha_1$ for all $x\in\Sigma\cap\operatorname{int}(\bar{x}+\epsilon\mathbb{B})$ if and only if the inclusion (1) has a local error bound at $\bar{x}$.*
4. *If both X and Y are assumed to be finite dimensional, then the convex inclusion $h(x)\in C$ has a local error bound at every point of Σ if and only if for each $r>0$ there exists $\alpha>0$ such that both (11) and (12) hold with for all $x\in r\mathbb{B}\cap\Sigma$.*

Remark 1 By Lemma 4, the condition that (11) and (12) hold can be replaced by the single condition that (10) holds in the statement of each of the results given in Theorem 2.

Remark 2 As noted in the Introduction, if h is assumed to be continuous, then the existence of local error bound for the inclusion (1) is equivalent to the calmness of the constraint set-valued mapping $\tilde{\Sigma}$ (see (4)) at $(0, \bar{x})$ [23], which is induced by the inclusion (1). Thus, Parts 2 and 3 provides a characterization for the *calmness* of the set-valued mapping $\tilde{\Sigma}$ at $(0, \bar{x})$. In addition, if the order structure induced by the recession cone C^∞ makes the space Y a B-lattice [44], then it can be shown that the concavity of h combined with hypothesis that h is locally order bounded implies that h is locally Lipschitz continuous. Further results on calmness, including cases where the mapping h is not assumed to be differentiable, can be found in [23].

Proof 1. It has already been observed that there is a global error bound for (1) with $\alpha > 0$ if and only if Σ is a set of weak sharp minima for $\operatorname{dist}(h(x) \mid C)$ with modulus $\alpha > 0$. Hence, by Part (A) of Theorem 1 and Part 5 of Lemma 3, there is a global error bound for (1) with $\alpha > 0$ if and only if (10) holds for all $x \in \Sigma$ which by Lemma 4 is equivalent to (11) and (12).
2. This follows immediately from Part (B) of Theorem 1 and Lemma 4.
3. This is an immediate consequence of Part (B) of Theorem 1 and Lemma 4.
4. This is an immediate consequence of [10, Theorem 6.3], [10, Corollary 6.4], and Lemma 4. □

These results show that in order to establish an error bound for the convex inclusion (1) one needs only to establish (11) (the Abadie CQ) and (12) on an appropriate set. We now examine these conditions in the finite–dimensional setting under the assumption that the set C is polyhedral.

3.1 The Abadie CQ in the polyhedral case:

Assume that $C \subset \mathbb{R}^m$ is polyhedral. We show that the condition (12) is satisfied in a neighborhood of every point in Σ. If in addition h is assumed to be affine, (12) is satisfied uniformly on Σ. We begin with a simple corollary to [40, Corollary 17.1.2].

Lemma 5 *Let $C \subset \mathbb{R}^m$ be the polyhedral set*

$$C = \left\{ y \;\middle|\; \left\langle a^j, y \right\rangle \leq \alpha_j, \;\; j = 1, 2, \ldots, r \right\},$$

where $a^j \in \mathbb{R}^m$ and $\alpha_j \in \mathbb{R}$ for $j = 1, 2, \ldots, r$. Given $y \in C$, $H \in \mathbb{R}^{n\times m}$, and $z \in H N_C(y)$ with $z \neq 0$, there exist linearly independent vectors $v_1, v_2, \ldots, v_s$ from the set of vectors

$$\left\{ H a^j \;\middle|\; j \in \{1, 2, \ldots, r\}, \;\; \text{and} \;\; \left\langle a^j, y \right\rangle = \alpha_j \right\}$$

such that $z = \sum_{t=1}^{s} \lambda_t v_t$.

Proof Let $I = \left\{ j \;\middle|\; \langle a^j, y\rangle = \alpha_j \right\}$. Then the normal cone to C at y is the smallest convex cone containing the vectors $\left\{ a^j \mid j \in I \right\}$. Apply [40, Corollary 17.1.2] to the set of vectors $S = \left\{ H a^j \mid j \in I \right\}$ to obtain the result. □

Proposition 1 *Let $C \subset \mathbb{R}^m$ be a polyhedral convex set and let $h: \mathbb{R}^n \mapsto \mathbb{R}^m$ be a continuously differentiable mapping that is concave with respect to C^∞. Then, given $\bar{x} \in \Sigma$, there exists $\kappa > 0$ and $\delta > 0$, depending on $\bar{x}$, such that (12) holds whenever $x \in \Sigma \cap (\bar{x} + \delta\mathbb{B})$. If it is further assumed that h is affine, then (12) holds for all $x \in \Sigma$.*

Proof We may assume $C = \left\{ y \,\middle|\, \left\langle a^j, y \right\rangle \le \alpha_j,\ j = 1, 2, \dots, r \right\}$, where $a^j \in \mathbb{R}^m$ and $\alpha_j \in \mathbb{R}$ for $j = 1, 2, \dots, r$. Suppose to the contrary that the result is false. Then there exist sequences $x^k \to \bar{x}$ and $\kappa_k \uparrow \infty$ such that $\{x^k\} \subset \Sigma$ and

$$\mathbb{B}^\circ \cap h'(x^k)^T N_C\left(h(x^k)\right) \subset h'(x^k)^T \left(\kappa_k \mathbb{B}^\circ \cap N_C\left(h(x^k)\right)\right)$$

for all $k = 1, 2, \dots$. That is, there exist $z^k \in h'(x^k)^T N_C\left(h(x^k)\right)$ with $\left\|z^k\right\|_o = 1$ such that $z^k \notin h'(x^k)^T (\kappa_k \mathbb{B}^\circ \cap N_C\left(h(x^k)\right))$ for all $k = 1, 2, \dots$. Set $v_j^k = h'(x^k)^T a^j$ for $j = 1, 2, \dots, r$ and all $k = 1, 2, \dots$. By Lemma 5, there exist index sets $J_k \subset \left\{ j \,\middle|\, \left\langle a^j, h(x^k) \right\rangle = \alpha_j \right\}$ and scalars $\lambda_j^k \ge 0$ for $j \in J_k$ such that $z^k = \sum_{j \in J_k} \lambda_j^k v_j^k$ where the vectors $\left\{ v_j^k \mid j \in J_k \right\}$ are linearly independent for all $k = 1, 2, \dots$. Due to the finiteness of the index sets, we assume with no loss of generality that there is an index set $J \subset \{1, 2, \dots, r\}$ such that $J_k = J$ for all $k = 1, 2, \dots$. Since z^k is not an element of $h'(x^k)^T (\kappa_k \mathbb{B}^\circ \cap N_C\left(h(x^k)\right))$ for all $k = 1, 2, \dots$, the vectors $w^k = \sum_{j \in J} \lambda_j^k a^j$ diverge. Thus, we may further assume, with no loss of generality, that for each $j \in J$ there is a $\bar{\lambda}_j \ge 0$ with $\lambda_j^k (\sum_{p \in J} \lambda_p^k)^{-1} \to \bar{\lambda}_j$, $z^k (\sum_{p \in J} \lambda_p^k)^{-1} \to 0$, and $\sum_{p \in J} \bar{\lambda}_p > 0$. Set $\tau_j = \frac{\bar{\lambda}_j}{\left\|\sum_{p \in J} \bar{\lambda}_p a^p\right\|_o}$ for $j \in J$, and $\bar{w} = \sum_{p \in J} \tau_p a^p$. We have $\|\bar{w}\|_o = 1$, $\bar{w} \in N_C\left(h(\bar{x})\right)$, and $0 = h'(\bar{x})^T \bar{w} = \sum_{j \in J} \tau_j v_j$ where $v_j = h'(\bar{x})^T a^j$ for $j = 1, 2, \dots, r$.

Consider the function $\phi(x) = \langle \bar{w}, h(x) \rangle$. By Lemma 2, we know that ϕ is a convex function of x. In addition, $\nabla\phi(\bar{x}) = h'(\bar{x})^T \bar{w} = 0$. Hence, $\bar{x}$ is a global solution to the problem $\min \phi(x)$. Suppose there exists $\hat{x} \in \Sigma$ such that $\left\langle \bar{w}, h(\hat{x}) \right\rangle = \phi(\hat{x}) > \phi(\bar{x}) = \langle \bar{w}, h(\bar{x}) \rangle$. Then, since $\bar{w} \in \mathbb{B}^\circ \cap N_C\left(h(\bar{x})\right)$, we have

$$\begin{aligned}
\operatorname{dist}\left(h(\hat{x}) \mid C\right) &= \sup_{\|w\|_o \le 1} \left\langle w, h(\hat{x}) \right\rangle - \psi_C^*(w) && \text{[10, Theorem A.1, Part 2]}\\
&\ge \left\langle \bar{w}, h(\hat{x}) \right\rangle - \psi_C^*(\bar{w})\\
&> \langle \bar{w}, h(\bar{x}) \rangle - \psi_C^*(\bar{w})\\
&= \operatorname{dist}\left(h(\bar{x}) \mid C\right) && \text{[10, Theorem A.1, Part 5]}\\
&= 0.
\end{aligned}$$

This contradicts the fact that $\hat{x} \in \Sigma$. Therefore, $\Sigma \subset \arg\min \phi(x)$, or equivalently, $0 = h'(x)^T \bar{w}$ for all $x \in \Sigma$. In particular, $0 = h'(x^k)^T \bar{w} = \sum_{j \in J} \tau_j v_j^k$. Since $\tau_j \neq 0$ for at least one $j \in J$, this contradicts the linear independence of the vectors $\left\{ v_j^k \mid j \in J \right\}$, which establishes the result.

Let us now assume that h is affine, then $h'(x) \equiv H \in \mathbb{R}^{m \times n}$ for all $x \in \mathbb{R}^n$. Since C is polyhedral, it has only finitely many faces, where $F \subset C$ is a face of C if and only if every convex subset of C whose relative interior meets F is contained in F. The relative interiors of the faces of C form a partition of C [40, Theorem 18.2]. In addition, if $x, y \in C$ lie in the relative interior of the same face of C, then their normal cones coincide [14, Theorem 2.3]. Therefore, to each face of C one can associate a unique normal cone, i.e., the unique normal cone associated with every point in the relative interior of that face. Thus, in particular, C has only finitely many distinct normal cones, one for each one of it faces. The first part of this result shows that for each such normal cone N there is a $\kappa_N > 0$ such that $\mathbb{B}^\circ \cap H^T N \subset H^T(\kappa_N \mathbb{B}^\circ \cap N)$. Setting κ equal to the maximum of this finite collection of κ_N's yields the result. $\square$

We give two elementary applications of this result. The first is Hoffman's bound [25] and the second is a recent result of Li [30, Theorem 3.5] concerning the equivalence of the existence of a local error bound and Abadie's CQ in the polyhedral case. We observe that as stated [30, Theorem 3.5] provides a characterization for the existence of local error bound at every point of Σ in terms of the Abadie CQ. However, on review of [30, Definition 2.1] one sees that the characterization is really a pointwise characterization. We give the pointwise characterization below.

Theorem 3 *Let $C \subset \mathbb{R}^m$ be a polyhedral convex set and let $h\colon \mathbb{R}^n \mapsto \mathbb{R}^m$ be a continuously differentiable mapping that is concave with respect to C^∞.*

1. *If h is affine, then a global error bound holds for the inclusion (1).*
2. *The inclusion (1) has a local error bound at $\bar{x} \in \Sigma$ if and only if the condition ACQ($U \cap \Sigma$) holds for some neighborhood U of $\bar{x}$.*

Proof 1. Suppose $h(x) = Ax + a$. Since $\psi_\Sigma(x) = \psi_C(Ax + a)$, we find from [40, Theorem 23.9] that (11) holds on all of Σ. In addition, Proposition 1 implies that (12) holds uniformly on Σ. Therefore, the global error bound follows from Part 1 of Theorem 2.
2. By Part 3 of Theorem 2, the existence of a local error bound at $\bar{x}$ implies that both (11) and (12) hold in a neighborhood of $\bar{x}$ where (11) is equivalent to ACQ.

Conversely, Proposition 1 states that there is a neighborhood $\hat{U}$ of $\bar{x}$ such that (12) holds on $\hat{U}$. Therefore, both (11) and (12) hold on the neighborhood $W = U \cap \hat{U}$. Consequently, Part 3 of Theorem 2 applies again to yield the existence of a local error bound for (1) at $\bar{x}$ $\square$

3.2 Other constraint qualifications:

A number of other constraint qualifications are possible for the inclusion (1). Some do not require the concavity of h with respect to C^∞. Let $\hat{\Sigma}$ be a subset of X. Typically $\hat{\Sigma}$ will be associated with some subset of the solution set $\Sigma = \{x \mid h(x) \in C\}$. Then, in finite dimensions, the most well know constraint qualifications are

LICQ($\hat{\Sigma}$): The linear independence constraint qualification:

$$\mathrm{Nul}\,(h'(x)) \cap \mathrm{span}\,(N_C\,(h(x))) = \{0\} \quad \forall\, x \in \hat{\Sigma},$$

where $\mathrm{Nul}\,(h'(x))$ denotes the null space of $h'(x)$.

MFCQ($\hat{\Sigma}$): The Mangasarian–Fromovitz constraint qualification:

$$\text{Nul}\,(h'(x)) \cap N_C\,(h(x)) = \{0\} \quad \forall\, x \in \hat{\Sigma}.$$

ACQ($\hat{\Sigma}$): The Abadie constraint qualification:

$$N_\Sigma\,(x) = h'(x)^T N_C\,(h(x)) \quad \forall\, x \in \hat{\Sigma}.$$

SCQ: The Slater constraint qualification:

There is a $\hat{x} \in \mathbb{R}^n$ such that $h(\hat{x}) \in \text{ri}\ (C)$.

Here we have stated the *dual* formulation of the constraint qualifications LICQ, MFCQ, and ACQ. The *primal* formulations are obtained by taking the polar of each of these expressions:

LICQ($\hat{\Sigma}$)*: $\text{Ran}\,(h'(x)) + \text{lin}\ (T_C\,(h(x))) = \mathbb{R}^m \quad \forall\, x \in \hat{\Sigma}.$
MFCQ($\hat{\Sigma}$)*: $\text{Ran}\,(h'(x)) + T_C\,(h(x)) = \mathbb{R}^m \quad \forall\, x \in \hat{\Sigma}.$
ACQ($\hat{\Sigma}$)*: $T_\Sigma\,(x) = h'(x)^{-1} T_C\,(h(x)) \quad \forall\, x \in \hat{\Sigma}.$

The equivalence of the primal and dual forms of both the LICQ and MFCQ follow from [40, Corollary 16.4.2], however, the primal and dual formulations of the ACQ are not necessarily equivalent. By [40, Corollary 16.3.2], one can show that ACQ($\hat{\Sigma}$) implies ACQ($\hat{\Sigma}$)*, but the converse implication holds if and only if the set $h'(x)^T N_C\,(h(x))$ is a closed set for all $x \in \hat{\Sigma}$. In the case when C is the cone $\mathbb{R}^s_- \times \{0\}^{m-s}$, it is straightforward to show that these definitions for the constraint qualifications LICQ, MFCQ, ACQ, and SCQ are equivalent to those used in the literature (for example, see [24]).

Using a constraint qualification suitable for the infinite–dimensional case, Maguregui [32, Theorem 2] establishes the metric regularity, and hence the existence of a local error bound, of the inclusion (1) without the assumption that h is concave with respect to C^∞. It is easy to show that Maguregui's CQ is equivalent to the primal version of the MFCQ in the finite–dimensional case (e.g., see [13, Lemma 3.2]). Hence, Maguregui's result shows that the MFCQ implies the existence of a local error bound (and much more). Since the LICQ implies the MFCQ, the same result holds for the LICQ.

For convex inclusions, it has long been known that the Slater CQ implies the existence of a local error bound when the set C is the cone $\mathbb{R}^s_- \times \{0\}^{m-s}$. We now establish this result in the general finite–dimensional case. For this purpose, it is useful to separate out that part of h which is guaranteed to be affine. A simple technique for doing this is illustrated in the following lemma.

Lemma 6 *Let $h\colon \mathbb{R}^n \mapsto \mathbb{R}^m$ and $C \subset \mathbb{R}^m$ be as given in (1) and (2). Let L be the subspace parallel to aff(C) and P be the orthogonal projection matrix onto L. Define $h_1\colon \mathbb{R}^n \mapsto \mathbb{R}^m$ and $h_2\colon \mathbb{R}^n \mapsto \mathbb{R}^m$ by $h_1(x) = Ph(x)$ and $h_2(x) = (I-P)h(x)$. Set $y^0 = (I-P)y$ for any $y \in$ aff(C), $\hat{C} = PC$, $\Sigma_1 = \left\{x \,\middle|\, h_1(x) \in \hat{C}\right\}$, and $\Sigma_2 = \left\{x \,\middle|\, h_2(x) = y^0\right\}$. Then h_2 is affine and $\Sigma = \Sigma_1 \cap \Sigma_2$.*

Proof We have $C^\infty \subset L$ and $h = h_1 + h_2$, hence, $h_2((1-\lambda)x + \lambda y) = (1-\lambda)h_2(x) + \lambda h_2(y)$ for all $x, y \in \mathbb{R}^n$ and $\lambda \in [0, 1]$. Therefore, h_2 is affine. Since the orthogonal projection onto aff (C) is given by $Pw + y^0$ for any $w \in \mathbb{R}^m$, we have $h(x) \in C$ if and only if $Ph(x) \in PC$ and $(I - P)h(x) = y^0$. This establishes the result. □

Theorem 4 *Let $h\colon \mathbb{R}^n \mapsto \mathbb{R}^m$ and $C \subset \mathbb{R}^m$ be as given in (1) and (2). Assume further that h is a continuously differentiable mapping. If the constraint qualification SCQ holds, then there exists a local error bound for (1) at every point of Σ.*

Proof We first assume that int $(C) \neq \emptyset$ and show that MFCQ(Σ) holds. Indeed, if MFCQ(Σ) does not hold, then there is an $\tilde{x} \in \Sigma$ and a non-zero vector $w \in N_C(h(\tilde{x}))$ with $h'(\tilde{x})^T w = 0$. Hence the convex function $\langle w, h(\cdot)\rangle$ attains its global minimum at $\tilde{x}$. But, by [40, Theorem 13.1], $\left\langle w, h(\hat{x})\right\rangle < \psi_C^*(w) = \langle w, h(\tilde{x})\rangle$, since $\hat{x} \in$ int (C). This contradiction implies that MFCQ(Σ) holds if int $(C) \neq \emptyset$.

Now assume that int $(C) = \emptyset$ and let $L,\ P,\ h_1,\ h_2,\ y^0,\ \hat{C},\ \Sigma_1$, and Σ_2 be as in Lemma 6, and choose $\bar{x} \in \Sigma$. Observe that $\Sigma_1 = \left\{x \,\middle|\, h_1(x) \in \hat{C}\right\} = \left\{x \,\middle|\, h_1(x) \in \hat{C} + L^\perp\right\}$. In addition, $h_1(\hat{x}) \in$ int $\left(\hat{C} + L^\perp\right)$, since $h(\hat{x}) \in$ ri (C). Therefore, the continuity of h implies that $\hat{x} \in$ int $(\Sigma_1) \cap \Sigma_2$. By Lemma 6, $\Sigma = \Sigma_1 \cap \Sigma_2$. Therefore, Bauschke's Theorem [5, Theorem 5.6.2] applies to yield the existence of a $\tau_1 > 0$ and $\epsilon_2 > 0$ such that

$$\mathrm{dist}_2\,(x \mid \Sigma) \le \tau_1 \max\{\mathrm{dist}_2\,(x \mid \Sigma_1), \mathrm{dist}_2\,(x \mid \Sigma_2)\} \quad \forall\, x \in (\bar{x} + \epsilon_1 \mathbb{B})\,. \tag{13}$$

Since $h_1(\hat{x}) \in$ int $\left(\hat{C} + L^\perp\right)$, we know from above that the convex inclusion $h_1(x) \in \hat{C} + L^\perp$ satisfies the constraint qualification MFCQ(Σ_1). Therefore, by Maguregui [32, Theorem 2], there exist $\tau_2 > 0$ and $\epsilon > 0$ such that

$$\mathrm{dist}_2\,(x \mid \Sigma_1) \le \tau_2 \mathrm{dist}_2\,(h_1(x) \mid \hat{C} + L^\perp) = \tau_2 \mathrm{dist}_2\,(h_1(x) \mid \hat{C}) \quad \forall\, x \in (\bar{x} + \epsilon_2 \mathbb{B})\,, \tag{14}$$

where $\mathrm{dist}_2\,(h_1(x) \mid \hat{C} + L^\perp) = \mathrm{dist}_2\,(h_1(x) \mid \hat{C})$ for all $x \in \mathbb{R}^n$ since $h_1(x) \in L$ for all $x \in \mathbb{R}^n$ and $\hat{C} \in L$. Also, since h_2 is affine, there exists $\tau_3 > 0$ such that

$$\mathrm{dist}_2\,(x \mid \Sigma_2) \le \tau_3 \left\| h_2(x) - y^0 \right\|_2 \quad \forall\, x \in \mathbb{R}^n\,. \tag{15}$$

Setting $\epsilon_0 = \min\{\epsilon_1, \epsilon_2\}$ and $\tau_0 = \tau_1 \max\{\tau_2, \tau_3\}$, the relations (13), (14), and (15) imply that

$$\mathrm{dist}_2\,(x \mid \Sigma) \le \tau_0 \max\{\mathrm{dist}_2\,(h_1(x) \mid \hat{C}), \left\| h_2(x) - y^0 \right\|_2\} \quad \forall\, x \in (\bar{x} + \epsilon \mathbb{B}). \tag{16}$$

Next observe that for every $x \in \mathbb{R}^n$ and $y \in C$,

$$\begin{aligned} \|h(x) - y\|_2^2 &= \|(h_1(x) - Py) + (h_2(x) - (I - P)y)\|_2^2 \\ &= \|h_1(x) - Py\|_2^2 + \left\| h_2(x) - y^0 \right\|_2^2 \\ &\ge \max\left\{ \|h_1(x) - Py\|_2^2, \left\| h_2(x) - y^0 \right\|_2^2 \right\} \end{aligned}$$

Taking the square-root on both sides of this inequality and then taking the infimum over all $y \in C$ yields the inequality

$$\operatorname{dist}_2(h(x) \mid C) \geq \max\{\operatorname{dist}_2(h_1(x) \mid \hat{C}), \left\|h_2(x) - y^0\right\|_2\} \quad \forall\, x \in \mathbb{R}^n. \tag{17}$$

Combining (17) and (16) yields the result for the 2-norm. This implies the result for an arbitrary norm by the equivalence of norms. □

We conclude this section by giving a result that establishes the relationships between the various constraint qualifications that we have discussed.

Proposition 2 *Let $h\colon \mathbb{R}^n \mapsto \mathbb{R}^m$ and $C \subset \mathbb{R}^m$ be as given in (1) and (2). Assume further that h is a continuously differentiable mapping. Then, for any $x \in \Sigma$, we have*

$$\textit{LICQ}(\{x\}) \;\Rightarrow\; \textit{MFCQ}(\{x\}) \;\Rightarrow\; \textit{SCQ} \;\Rightarrow \textit{ACQ}(\Sigma)\,.$$

If it is further assumed that $\operatorname{int}(C) \neq \emptyset$, *then SCQ ⇒ MFCQ($\Sigma$).*

Proof LICQ($\{x\}$) ⇒ MFCQ($\{x\}$): This follows immediately from the definitions.
MFCQ($\{x\}$) ⇒ SCQ: Let $x \in \Sigma$ be such that MFCQ($\{x\}$) is satisfied. If $h(x) \in \operatorname{ri}(C)$ we are done, so assume that $h(x) \notin \operatorname{ri}(C)$. By [13, Lemma 2.3], the condition MFCQ($\{x\}$) is equivalent to the condition that there exits $\epsilon, \mu > 0$ such that $0 \in \operatorname{int}(\mu h'(y)\mathbb{B} + (\operatorname{ri}(C) - h(y)))$ whenever $\|x - y\| \leq \epsilon$. Consequently,

$$\exists\, d \in \mu\mathbb{B} \text{ such that } h(x) + h'(x)d \in \operatorname{ri}(C)\,. \tag{18}$$

We will show that (18) is sufficient to establish SCQ.

Let us first assume that $\operatorname{int}(C) \neq \emptyset$, so that $h(x) + h'(x)d \in \operatorname{int}(C)$. Then there is an $\epsilon > 0$ such that $S = h(x) + h'(x)d + \epsilon\mathbb{B} \subset \operatorname{int}(C)$. By convexity, the angle θ between $h(x) + h'(x)d$ and any vector in the boundary cone (S) satisfies $0 < \sin\theta < 1$. Set $\delta = \sin\theta$. Then for $\lambda > 0$ sufficiently small, we have $\operatorname{dist}(h(x) + \lambda h'(x)d \mid \operatorname{bdry}(S)) = \lambda \left\|h'(x)d\right\|_2 \delta$. Now if $h(x + \lambda d) \notin \operatorname{int}(C)$ for all small $\lambda > 0$, then $\left\|h(x) + \lambda h'(x)d - h(x + \lambda d)\right\| \geq \operatorname{dist}(h(x) + \lambda h'(x)d \mid \operatorname{bdry}(S)) = \lambda \left\|h'(x)d\right\|_2 \delta$ for all small $\lambda > 0$. But this contradicts the fact that $\left\|h(x) + \lambda h'(x)d - h(x + \lambda d)\right\| = o(\lambda)$. Therefore, $h(x + \lambda d) \in \operatorname{int}(C)$ for all small $\lambda > 0$ which implies that SCQ is satisfied.

Now consider the general case and let $L,\ P,\ h_1,\ h_2,\ y^0,\ \hat{C},\ \Sigma_1$, and Σ_2 be as in Lemma 6. By Lemma 6, there exist $A \in \mathbb{R}^{m\times n}$ and $a \in \mathbb{R}^m$ such that $h_2(x) = Ax + a$. By (18), there is a $d \in \mathbb{R}^n$ such that $h(x) + h'(x)d \in \operatorname{ri}(C)$. In particular, this implies that $h_2(x + \lambda d) = A(x + \lambda d) + a = y^0$ for all $\lambda \in \mathbb{R}$, or equivalently, $d \in \operatorname{Nul}(A)$. In addition, we have also shown that $h_1(x + \lambda d) \in \operatorname{int}\left(C_1 + L^\perp\right)$ for all $\lambda > 0$ sufficiently small, since $h_1(x) + h_1'(x)d \in \operatorname{int}\left(C_1 + L^\perp\right)$. Therefore, $h(x + \lambda d) = h_1(x + \lambda d) + h_2(x + \lambda d) \in \operatorname{ri}(C_1) + \{y^0\} = \operatorname{ri}(C)$ for all $\lambda > 0$ sufficiently small. This establishes the result.
SCQ ⇒ ACQ(Σ): By Theorem 4, SCQ implies the existence of a local error bound at every point of Σ. Therefore, the result follows from Part 2 of Theorem 2.

$[\text{int }(C) \neq \emptyset] \Rightarrow [\text{SCQ} \Rightarrow \text{MFCQ}(\Sigma)]$: Suppose to the contrary that there is an $\tilde{x} \in C$ and a non-zero vector $w \in N_C(\tilde{x})$ with $h'(\tilde{x})^T w = 0$. Then the convex function $\langle w, h(\cdot)\rangle$ attains its global minimum at $\tilde{x}$. But, by [40, Theorem 13.1], $\langle w, h(\hat{x})\rangle < \psi_C^*(w) = \langle w, h(\tilde{x})\rangle$, since $\hat{x} \in \text{int }(C)$. This contradiction yields the result. □

The following examples show that the implications given in the Proposition 2 are complete.

Examples 1. (ACQ(Σ) $\not\Rightarrow$ SCQ) Let $h\colon \mathbb{R} \mapsto \mathbb{R}^2$ be given by $h(x) = (x, -x)^T$. Let $C = \mathbb{R}^2_-$. The $\Sigma = \{0\}$, $N_\Sigma(0) = \mathbb{R}$, $N_C(0) = \mathbb{R}^2_+$, and $h'(0)^T\mathbb{R}^2_+ = \mathbb{R}$. Therefore, the ACQ($\Sigma$) is satisfied, but SCQ is not.

2. (SCQ $\not\Rightarrow$ MFCQ($\{x\}$)) Let $h\colon \mathbb{R} \mapsto \mathbb{R}^2$ be given by $h(x) = (x, 0)^T$. Let $C = \mathbb{R} \times \{0\}$. Then SCQ is satisfied, but $\text{Nul}\,(h'(x)) \cap N_C(h(x)) = \{0\} \times \mathbb{R}$ for every $x \in \Sigma = \mathbb{R}$.
3. (MFCQ($\{x\}$) $\not\Rightarrow$ LICQ($\{x\}$)) Let $h\colon \mathbb{R} \mapsto \mathbb{R}^2$ be given by $h(x) = (x, x)^T$. Let $C = \mathbb{R}^2_-$. Then $\Sigma = \mathbb{R}_-$, $\text{Nul}\,(h'(0)) \cap N_C(h(0)) = \{0\}$, but LICQ($\{0\}$) does not hold.

3.3 Affine convex inclusions and Hoffman bounds:

It is instructive to consider the special case of affine convex inclusions since this is by far the most studied as well as being arguably the most important case with regard to applications. Global error bounds for affine convex inclusions are referred to as *Hoffman bounds* in honor of the seminal work of A. J. Hoffman [25]. Hoffman's original work focuses on the case where X and Y are finite dimensional and C is the negative orthant. Historically, this is the most intensively studied case. We do not attempt a review of the enormous literature on this case or even on the slightly more general polyhedral case in finite dimensions. Rather, our focus is on *semi-infinite* affine convex inclusions where the set C is only assumed to be convex. Specifically, given $A \in L(X, Y)$ and $b \in Y$ with X being a normed linear space and Y finite dimensional, we consider convex inclusions of the form

$$Ax - b \in C, \quad \text{where } C \subset Y \text{ non-empty, closed and convex.} \tag{19}$$

To our knowledge, the only study of global error bounds, or Hoffman bounds, for this general case is [15].

The goal in [15] is to obtain verifiable sufficient conditions under which a global error bound exists and to obtain sharp computable estimates for the modulus of the error bound. The main results in [15] are similar to the type stated previously in this section for general convex inclusions. For example, in the affine setting the MFCQ($\{x\}$) takes the form

$$\text{Nul}\,(A^*) \bigcap N_C(x) = \{0\}. \tag{20}$$

Thus, one might conjecture that by specifying that (20) hold at every point of C the existence of a global error bound would follow. This result indeed follows directly from [15, Theorem 9]. The precise statement given in [15] makes use of the relationship between the normal cone and the barrier cone as described in [11, Lemma 5].

Theorem 5 [15, Theorem 9] *Let $A \in L(X, Y)$, $b \in Y$, and $C \subset Y$ be as in* (19) (*Y finite dimensional*) *and satisfy*

$$Nul\,(A^*) \bigcap (C^\infty)^\circ = \{0\}. \tag{21}$$

Then the affine convex inclusion (19) *has a global error bound.*

Remark 3 Sharp estimates for the modulus of the error bound are also presented in [15, Theorem 9].

Remark 4 The condition (21) is equivalent to the dual statement

$$\mathrm{Ran}\,(A) + C^\infty = Y\ . \tag{22}$$

Remark 5 The actual statement of [15, Theorem 9] requires that $b \in \mathrm{Ran}\,(A) + (C_1 \cap \mathrm{ri}\ (C_2))$ where $C = C_1 \bigcap C_2 \neq \emptyset$ with both C_1 and C_2 closed and convex, and C_1 polyhedral. However, due to the equivalence of (21) and (22) a straightforward separation argument shows that (21) implies $\mathrm{Ran}\,(A) + (C_1 \cap \mathrm{ri}\ (C_2)) = Y$ so the hypotheses of [15, Theorem 9] are trivially satisfied. The condition that $b \in \mathrm{Ran}\,(A) + (C_1 \cap \mathrm{ri}\ (C_2))$ is the counterpart of the Slater CQ in this setting.

The sufficient condition (21) given in Theorem 5 does not require knowledge of the solution set $\Sigma = \{x \mid Ax - b \in C\}$. In contrast, the necessary and sufficient condition derived from Part 1 of Theorem 2 makes explicit use of Σ.

Theorem 6 *A global error bound holds for* (19) (*Y finite dimensional*) *with modulus $\alpha > 0$ if and only if for every $x \in \Sigma = \{x \mid Ax - b \in C\}$ one has*

$$N_\Sigma\,(x) = A^* N_C\,(Ax + b)$$

and

$$\alpha \mathbb{B}^\circ \cap A^* N_C\,(Ax + b) \subset A^*(\mathbb{B}^\circ \cap N_C\,(Ax + b)). \tag{23}$$

If C is a cone and $b = 0$, a result stronger than Theorem 6 is possible.

Theorem 7 *Consider the affine convex inclusion* (19) *with Y possibly infinite-dimensional, $b = 0$, and C assumed to be a non-empty, closed, and convex cone. Then the following statements are equivalent.*

1. *A global error bound holds for the inclusion* (19).
2. *The convex inclusion* (19) *has a local error bound at every point in Σ.*
3. *There is an $\alpha > 0$ such that*

$$\alpha \mathbb{B}^\circ \cap \Sigma^\circ \subset A^*(\mathbb{B}^\circ \cap C^\circ). \tag{24}$$

Finally, if Y is finite dimensional, then the constraint qualification

$$Nul\,(A^*) \bigcap C^\circ = \{0\} \tag{25}$$

implies that a global error bound holds for the convex inclusion (19).

Proof Let $\rho(x) = \operatorname{dist}(Ax \mid C)$. Since C is a cone, Part 3 of [10, Theorem A.1] states that $\rho(x) = \psi^*_{\mathbb{B}^\circ \cap C^\circ}(Ax)$. Thus, in particular,

$$\partial\rho(0) = \mathbb{B}^\circ \cap C^\circ. \tag{26}$$

Since $A \in L(X, Y)$, ρ is positively homogeneous which implies that

$$\rho = \rho^\infty \quad \text{and} \quad \Sigma^\infty = \Sigma \tag{27}$$

is a convex cone. In particular, Σ being a cone implies that

$$N_\Sigma(0) = \Sigma^\circ. \tag{28}$$

Clearly, Part 1 implies Part 2. Part 3 follows from Part 2 by applying Part 1 of [10, Theorem 5.2] at $x = 0$ and using (26) and (28). The fact that Part 3 implies Part 1 follows from [10, Theorem 4.1] along with the relations (27).

Finally, when Y is finite dimensional, the constraint qualification (25) is equivalent to (21) since $C^\infty = C$. The result now follows from Theorem 5 and (22). $\square$

This result has implications for the notion of linear regularity introduced in [11, Sect. 3].

Corollary 1 *Let $\{C_i \mid i = 1, \ldots, N\}$ be a collection of non-empty closed convex cones in the normed linear space X and suppose that the convex set $C = \bigcap_{i=1}^N C_i$ is non-empty. Then the following statements are equivalent*:

1. *The collection $\{C_i \mid i = 1, \ldots, N\}$ is linearly regular.*
2. *The collection $\{C_i \mid i = 1, \ldots, N\}$ is boundedly linearly regular.*
3. *There is an $\alpha > 0$ such that*

$$\alpha\mathbb{B}^\circ \cap C^\circ \subset \sum_{i=1}^N (\mathbb{B}^\circ \cap C_i^\circ). \tag{29}$$

In addition, if any one of these conditions holds, then the set $\sum_{i=1}^N C_i^\circ$ is w^*–closed. Finally, if X is finite dimensional, then the constraint qualification

$$[z^j \in C_j^\circ \; j = 1, \ldots, N, \; \sum_{i=1}^N z^j = 0] \iff [z^j = 0, \; j = 1, \ldots, N], \tag{30}$$

implies that the collection $\{C_i \mid i = 1, \ldots, N\}$ is linearly regular.

Proof Define $\hat{C} = C_1 \times \cdots \times C_N \subset X \times \cdots \times X = X^N$ and let $A \in L(X, X^N)$ be given by $Ax = (x, x, \ldots, x)$. Then $x \in C$ if and only if $Ax \in \hat{C}$. We define the norm on X^N to be $\left\|(x^1, x^2, \ldots, x^N)\right\| = \max\left\{\left\|x^j\right\| \mid j = 1, \ldots, N\right\}$. Now apply Theorem 7 to obtain the equivalence of statements 1, 2, and 3.

Next observe that statement 3 implies that

$$C^\circ = \text{cone}\,(\alpha\mathbb{B}^\circ \cap C^\circ) \subset \text{cone}\,(\sum_{i=1}^{N}(\mathbb{B}^\circ \cap C_i^\circ)) = \sum_{i=1}^{N} C_i^\circ \subset C^\circ.$$

Therefore, $\sum_{i=1}^{N} C_i^\circ$ is w*–closed since C° is.

Finally, when X is finite dimensional, the pointedness condition (30) is equivalent to the condition $\text{Nul}\,(A^*) \cap \hat{C}^\circ = \{0\}$. The last part of Theorem 7 implies that the collection $\{\, C_i \mid i = 1, \ldots, N\,\}$ is linearly regular. □

A key ingredient in Theorem 7 is the positive homogeneity of the distance function $\text{dist}\,(Ax \mid C)$. The following example shows that, in general, Theorem 7 fails for inclusion $Ax + b \in C$ even when C is a convex cone. The example shows that even if (19) has a local error bound at every point, a global error bound may not exist.

Example 1 Consider the inclusion $Ax + b \in C = C_1 \times C_2$, where

$$C_1 = \{x \in R^3 \mid \sqrt{x_1^2 + x_2^2} \le x_3\}, \qquad C_2 = \{0\},$$

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & -1 \end{pmatrix}, \quad \text{and} \quad b = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Let $\Sigma = \{x \in R^3 \mid Ax + b \in C\}$. A simple computation shows that $\Sigma = \{x \in C_1 \mid x_3 = x_1 + 1\}$, $\Sigma^\infty = \{x \in R^3 \mid x_1 = x_3 \ge 0, x_2 = 0\} = \text{cone}\,([1, 0, 1]^T)$, and $ri(C) = int(C_1) \times C_2$. Since $A\bar{x} + b \in ri(C)$, where $\bar{x} = [0, 0, 1]^T$, by Theorem 4, the inclusion $Ax + b \in C$ has a local error bound at every point of Σ. From $\Sigma = \arg\min_{x \in R^3} \text{dist}\,(Ax + b \mid C)$, and $(\text{dist}\,(Ax + b \mid C))^\infty = (\text{dist}\,(Ax \mid C))^\infty = \text{dist}\,(Ax \mid C)$ [24, Proposition 3.2.9], we know that $\Sigma^\infty = \arg\min_{x \in R^3} \text{dist}\,(Ax \mid C)$. We claim that Σ^∞ is not a set of weak sharp minima for $\text{dist}\,(Ax \mid C)$. To see this, let $x^n = [n, 1, n]^T$ for $n = 1, 2, \ldots$ Then $\text{dist}\,(x^n \mid \Sigma^\infty) = 1$, and

$$\begin{aligned}
\text{dist}\,(Ax^n \mid C^\infty) &= \text{dist}\,([n, 1, n, 0]^T \mid C) = \psi^*_{\mathbb{B}^\circ \cap C_1^\circ}(x^n) \\
&= \max\left\{ nx_1 + x_2 + nx_3 \;\middle|\; \sqrt{x_1^2 + x_2^2} \le -x_3, -1 \le x_3 \le 0 \right\} \\
&\le \max\left\{ \sqrt{n^2 + 1}\sqrt{x_1^2 + x_2^2} + nx_3 \;\middle|\; \begin{array}{c} \sqrt{x_1^2 + x_2^2} \le -x_3, \\ -1 \le x_3 \le 0 \end{array} \right\} \\
&= \sqrt{n^2 + 1} - n = \frac{1}{\sqrt{n^2 + 1} + n}.
\end{aligned}$$

It follows that $\text{dist}\,(Ax^n \mid C) \to 0$ as $n \to \infty$. Hence there is no $\alpha > 0$ such that

$$\alpha \text{dist}\,(x \mid \Sigma^\infty) \leq \text{dist}\,(Ax \mid C) \quad \forall x \in R^3.$$

By [10, Theorem 4.2], Σ cannot be a set of weak sharp minima for $\text{dist}\,(Ax + b \mid C)$, which is equivalent to saying that there is no global error bound for the inclusion $Ax+b \in C$. Finally, we note that in this example, $(1, 0, -1, -1)^T \in \text{Nul}\,(A^*) \cap C^\circ$, i.e. (21) fails.

The stark difference in the behavior of the solution set to affine convex cone inclusions for $b = 0$ and $b \neq 0$ is intriguing. In the case when $b \neq 0$, one suspects the existence of a simple geometric condition under which a local error bound at every point is equivalent to the existence of a global error bound. We present a partial result in this direction that applies to general affine convex inclusions in finite dimensions. This result follows from [10, Theorem 6.5] and, as in Theorem 5, is based on a hypothesis concerning the relationship between the sets $\text{Ran}\,(A)$ and C^∞.

Proposition 3 *Suppose that X and Y are finite dimensional and consider the affine convex inclusion* (19). *If*

$$Ran\,(A) \cap C^\infty = \{0\}, \tag{31}$$

then the inclusion (19) *has a global error bound if and only if it has a local error bound at every point of* Σ.

Proof We have that $\Sigma^\infty = \{\, y \mid Ay \in C^\infty \}$. This, combined with the hypothesis (31), implies that $\Sigma^\infty = \text{Nul}\,(A)$. Let P be the orthogonal projector onto $\text{Nul}\,(A)^\perp$ and set $\hat{\Sigma} = P\Sigma$. It is straightforward to show that

$$\Sigma = \hat{\Sigma} + \text{Nul}\,(A).$$

Therefore, $\hat{\Sigma}^\infty \subset \Sigma^\infty = \text{Nul}\,(A)$, and so $\hat{\Sigma}^\infty = \{0\}$. Hence, by [40, Theorem 8.4], the set $\hat{\Sigma}$ is bounded. [10, Theorem 6.5] now applies to yield the result. □

Acknowledgments The authors thank an anonymous referee for very carefully reading the original submission and for supplying helpful comments which significantly improved the presentation.

References

1. Abadie, J.: On the Kuhn–Tucker Theorem. In: Abadie, J. (ed.) Nonlinear Programming, pp. 21–36. North Holland (1967)
2. Aubin, J.-P., Ekeland, I.: Applied Nonlinear Analysis. Wiley–Interscience, New York (1984)
3. Auslender, A., Crouzeix, J.-P.: Global regularity theorem. Math. Oper. Res. **13**, 243–253 (1988)
4. Auslender, A., Teboulle, M.: Asymptotic Cones and Functions in Optimization and Variational Inequalities. Springer, Heidelberg (2003)
5. Bauschke, H.: Projection algorithms and monotone operators. Ph.D. Thesis, Simon Fraser University, Department of Mathematics, Burnaby, British Columbia, V5A 1S6, Canada (1996)
6. Bauschke, H., Borwein, J., Li, W.: Strong conical hull intersection property, bounded linear regularity, Jameson's property (G), and error bounds in convex optimization. Math. Prog. **86**, 135–160 (1999)
7. Borwein, J., Lewis, A.: Convex Analysis and Nonlinear Optimization, Theory and Examples CMS Books in Mathematics. Springer, New York (2000)

8. Burke, J.V.: An exact penalization viewpoint of constrained optimization. SIAM J. Control Optim. **29**, 968–998 (1991)
9. Burke, J.V.: Calmness and exact penalization. SIAM J. Control Optim. **29**, 493–497 (1991)
10. Burke, J.V., Deng, S.: Weak sharp minima revisited, part I: Basic theory. Control Cybernetics **31**, 439–469 (2002)
11. Burke, J.V., Deng, S.: Weak sharp minima revisited, part II: Application to linear regularity and error bounds. Math. Program. Ser. B **104**, 235–261 (2005)
12. Burke, J.V., Ferris, M.C.: Weak sharp minima in mathematical programming. SIAM J. Control Optim. **31**, 1340–1359 (1993)
13. Burke, J.V., Ferris, M.C.: A Gauss–Newton method for convex composite optimization. Math. Prog. **71**, 179–194 (1995)
14. Burke, J.V., Moré, J.J.: On the identification of active constraints. SIAM J. Numer. Anal. **25**, 1197–1211 (1988)
15. Burke, J.V., Tseng, P.: A unified analysis of Hoffman's bound via Fenchel duality. SIAM J. Optim. **6**, 265–282 (1996)
16. Clarke, F.H.: A new approach to Lagrange multipliers. Math. Oper. Res. **2**, 165–174 (1976)
17. Deng, S.: Global error bounds for convex inequality systems in Banach spaces. SIAM J. Control Optim. **36**, 1240–1249 (1998)
18. Dontchev, A.L., Rockafellar, R.T.: Regularity and conditioning of solution mappings in variational analysis. Set-Valued Anal. **12**, 79–109 (2004)
19. Ekeland, I., Temam, R.: Convex analysis and variational problems. North Holland (1976)
20. Ferris, M.C.: Weak sharp minima and penalty functions in mathematical programming. Tech. Report 779, Computer Sciences Department, University of Wisconsin, Madison, Wisconsin (1988)
21. Henrion, R, Jourani, A.: Subdifferential conditions for calmness of convex constraints. SIAM J. Optim. **13**, 520–534 (2002)
22. Henrion, R., Outrata, J.: A subdifferential criterion for calmness of multifunctions. J. Math. Anal. Appl. **258**, 110–130 (2001)
23. Henrion, R., Outrata, J.: Calmness of constraint systems with applications. Math. Prog. Ser. B. **104**, 437–464 (2005)
24. Hiriart-Urruty, J.-B., Lemaréchal, C.: Convex Analysis and Minimization Algorithms I, volume 306 of Grundlehren der Mathematischen Wissenschaften. Springer, Heidelberg (1993)
25. Hoffman, A.J.: On approximate solutions to systems of linear inequalities. J. Res. Nat. Bur. Stand. **49**, 263–265 (1952)
26. Jourani, A.: Hoffman's error bounds, local controllability, and sensitivity analysis. SIAM J. Control Optim. **38**, 947–970 (2000)
27. Klatte, D.: Hoffman's error bound for systems of convex inequalities. In: Fiacco, A. V. (ed.) Mathematical Programming with Data Perturbations, pp. 185–199. Marcel Dekker Publ., Moscow (1997)
28. Klatte, D., Li, W.: Asymptotic constraint qualifications and global error bounds for convex inequalities. Math. Prog. **84**, 137–160 (1999)
29. Lewis, A. S., Pang, J.-S.: Error bounds for convex inequality systems. In: Crouzeix, J. P., Martinez-Legaz, J.-E., Volle, M. (ed.) Proceedings of the Fifth International Symposium on Generalized Convexity held in Luminy June 17-21, 1996, pp. 75–101, Kluwer Academic Publishers, Dordrecht (1998)
30. Li, W.: Abadie's constraint qualification, metric regularity, and error bounds for differentiable convex inequalities. SIAM J. Optim. **7**, 966–978 (1997)
31. Luo, X.D., Luo, Z.Q.: Extension of hoffman's error bound to polynomial systems. SIAM J. Optim. **4**, 383–392 (1994)
32. Maguregui, J.: Regular multivalued functions and algorithmic applications. Ph.D. Thesis, University of Wisconsin at Madison, Madison, WI (1977)
33. Ng, K.F., Yang, W.H.: Regularities and their relations to error bounds. Math. Prog. Ser. A **99**, 521–538 (2004)
34. Ngai, H.V., Thera, M.: Error bounds for convex differentiable inequality systems in Banach spaces. Math. Program. Ser. B **104**, 465–482 (2005)
35. Pang, J.-S.: Error bounds in mathematical programming. Math. Prog. **79**, 299–333 (1997)
36. Polyak, B.T.: Sharp minima. In: Proceedings of the IIASA Workshop on Generalized Lagrangians and Their Applications: Laxenburg, Austria. Institute of Control Sciences Lecture Notes, Moscow (1979)
37. Robinson, S.M.: Normed convex processes. Trans. Amer. Math. Soc. **174**, 127–140 (1972)
38. Robinson, S.M.: An application of error bounds for convex programming in a linear space. SIAM J. Control **13**, 271–273 (1975)

39. Robinson, S.M.: Regularity and stability for convex multivalued functions. Math. Oper. Res. **1**, 130–143 (1976)
40. Rockafellar, R.T.: Convex Analysis. Princeton University Press, Princeton (1970)
41. Rockafellar, R.T.: Conjugate Duality and Optimization. SIAM, Philadelphia (1974)
42. Rockafellar, R.T., Wets, R.J.-B.: Variational Analysis. Springer, Heidelberg (1998)
43. Ursescu, C.: Multifunctions with closed convex graph. Czech. Math. J. **25**, 438–441 (1975)
44. Yosida, K.: Functional Analysis. Springer, Heidelberg (1980)
45. Zălinescu, C.: Weak sharp minima, well-behaving functions and global error bounds for convex inequalities in Banach spaces. In: Bulatov, V., Baturin, V. (ed.) Proceedings of the 12th Baikal International Conference on Optimization Methods and their Applications, pp. 272–284, Institute of System Dynamics and Control Theory of SB RAS, Irkutsk (2001)
46. Zheng, X.Y., Ng, K.F.: Metric regularity and constraint qualifications for convex inequalities on Banach spaces. SIAM J. Optim. **14**, 757–772 (2004)
47. Zheng, X.Y., Ng, K.F.: Error moduli for conic convex systems on Banach spaces. Math. Oper. Res. **29**, 213–228 (2004)