SIAM J. OPTIM.
Vol. 14, No. 3, pp. 621–645

FIRST-ORDER AND SECOND-ORDER CONDITIONS FOR ERROR BOUNDS*

ZILI WU† AND JANE J. YE†

Abstract. For a lower semicontinuous function f on a Banach space X, we study the existence of a positive scalar μ such that the distance function d_S associated with the solution set S of $f(x) \leq 0$ satisfies

$$d_S(x) \leq \mu \max\{f(x), 0\}$$

for each point x in a neighborhood of some point x_0 in X with $f(x) < \epsilon$ for some $0 < \epsilon \leq +\infty$. We give several sufficient conditions for this in terms of an abstract subdifferential and the Dini derivatives of f. In a Hilbert space we further present some second-order conditions. We also establish the corresponding results for a system of inequalities, equalities, and an abstract constraint set.

Key words. error bounds, existence of solutions, inequality systems, lower Dini derivatives, abstract subdifferentials, first-order conditions, second-order conditions

AMS subject classifications. 49J52, 90C26, 90C31

DOI. 10.1137/S1052623402412982

1. Introduction.

Let (X, d) be a metric space. For a proper and lower semicontinuous (l.s.c.) function $f : X \to (-\infty, \infty]$, denote the solution set of the inequality system $f(x) \leq 0$ by

$$S := \{x \in X : f(x) \leq 0\}$$

and the distance from a point $x \in X$ to the set S by

$$d_S(x) := \inf\{d(x, s) : s \in S\}$$

if S is nonempty. By convention, $d_S(x) = +\infty$ if S is empty.

Let T be a nonempty subset of X and let γ be a positive scalar. We say that the inequality system $f(x) \leq 0$ has an *error bound* of the pair (S, T) with exponent γ if the set S is nonempty and there exists a scalar $\mu > 0$ such that

$$d_S(x) \leq \mu [f(x)_+]^\gamma \quad \text{for all } x \in T,$$

where $f(x)_+ := \max\{f(x), 0\}$. For the case $\gamma = 1$, if

$$T = f^{-1}(0, \epsilon) := \{x \in X : 0 < f(x) < \epsilon\}$$

for some $0 < \epsilon < +\infty$ ($\epsilon = +\infty$), we simply say that the system $f(x) \leq 0$ (or the set S) has a *local* (*global*) *error bound*; if

$$T = B(x_0, \delta) := \{x \in X : d(x, x_0) < \delta\}$$

for some $x_0 \in S$ and $0 < \delta$, the set S is said to be *metrically regular* at x_0.

*Received by the editors August 12, 2002; accepted for publication (in revised form) May 29, 2003; published electronically December 19, 2003.
http://www.siam.org/journals/siopt/14-3/41298.html

†Department of Mathematics and Statistics, University of Victoria, Victoria, BC, Canada V8W 3P4 (ziliwu@email.com, janeye@math.uvic.ca). The research of the second author was supported by an NSERC grant.

Error bounds have important applications in sensitivity analysis of mathematical programming and in convergence analysis of some algorithms. In his seminal paper [8], Hoffman showed that a linear inequality system has a global error bound. For nonlinear inequality systems, the existence of error bounds usually requires some conditions. Most earlier results about error bounds are related to a continuous or convex system on R^n. The reader is referred to the recent survey papers [11, 14] and the references therein for a summary of the theory and applications of error bounds.

Recently Ng and Zheng [15, 16] and Wu and Ye [21, 22] studied l.s.c. inequality systems and presented several sufficient conditions for error bounds in terms of the lower Dini derivative and an abstract subdifferential. These results are mainly established for the case $T = f^{-1}(0, \epsilon)$ $(0 < \epsilon \leq +\infty)$. The first purpose of this paper is to extend and develop the above first-order conditions to the case $T = B(x_0, \delta) \cap f^{-1}(0, \epsilon)$, where $x_0 \in X$, $0 < \epsilon \leq +\infty$ and $0 < \delta \leq +\infty$. We do not assume that x_0 lies in the solution set S nor that δ is $+\infty$. However, our results are applicable to the cases $x_0 \in S$ and $\delta = +\infty$; that is, they serve as sufficient conditions not only for regularity (when $x_0 \in S$) but also for error bounds (when $\delta = +\infty$). The second purpose is to present a second-order sufficient condition for the existence of error bounds with exponents $1/2$ in a Hilbert space from which we can further obtain sufficient conditions for nonconvex quadratic systems. Our third purpose is to specify the first-order and second-order conditions for the following system of inequalities, equalities, and an abstract set:

$$\begin{aligned} &g_i(x) \leq 0 \text{ for all } i \in I := \{1, \ldots, m\}, \\ &h_j(x) = 0 \text{ for all } j \in J := \{1, \ldots, n\}, \\ &\quad x \in C, \end{aligned}$$

where g_i and $|h_j|$ are l.s.c. and C is a nonempty closed subset of X.

It is worth pointing out that, unlike other error bound results, the nonemptiness of the solution set of an inequality system in ours comes as a conclusion instead of an assumption. Therefore, we can also use them as sufficient conditions for the existence of its solutions.

Apart from the above notation, the following concepts on nonsmooth analysis also are needed in this paper (see, e.g., [3, 4, 17]):

Let X be a normed linear space, let x and v be in X, and let $f : X \to (-\infty, +\infty]$ be finite at x.

- The lower Dini derivative of f at x in the direction v is

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

- The upper Dini derivative of f at x in the direction v is

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

- The *Clarke derivative* of f at x in the direction v is

$$f^\circ(x; v) := \limsup_{\substack{y \to x \\ t \to 0^+}} \frac{f(y + tv) - f(y)}{t}.$$

- The *Clarke subdifferential* of f at x is

$$\partial^\circ f(x) := \{\xi \in X^* : \langle \xi, v\rangle \le f^\circ(x; v) \text{ for all } v \in X\}.$$

When X is a Hilbert space, we say that a vector $\xi \in X$ is a *proximal subgradient* of f at x provided that there exist positive scalars M and δ such that

$$f(y) \ge f(x) + \langle \xi, y - x\rangle - M\|y - x\|^2 \text{ for all } y \in B(x, \delta).$$

The set of all such ξ, denoted by $\partial^\pi f(x)$, is referred to as the *proximal subdifferential* of f at x.

For each $\xi \in \partial^\pi f(x)$, we define the following second-order subderivatives:

$$\begin{aligned}
d_L^2 f(x|\xi)(u) &:= \liminf_{t\to 0^+} \frac{f(x+tu) - f(x) - t\langle \xi, u\rangle}{t^2},\\
d_-^2 f(x|\xi)(u) &:= \liminf_{\substack{v\to u\\ t\to 0^+}} \frac{f(x+tv) - f(x) - t\langle \xi, v\rangle}{t^2},\\
d_+^2 f(x|\xi)(u) &:= \limsup_{\substack{v\to u\\ t\to 0^+}} \frac{f(x+tv) - f(x) - t\langle \xi, v\rangle}{t^2}.
\end{aligned}$$

Usually, for $u \in X$ and $\xi \in \partial^\pi f(x)$, we have

$$d_-^2 f(x|\xi)(u) \le d_L^2 f(x|\xi)(u) \le d_+^2 f(x|\xi)(u).$$

If f is a C^2 function with its first-order and second-order derivatives at x denoted by $\nabla f(x)$ and $\nabla^2 f(x)$, respectively, then, since $\partial^\pi f(x) = \{\nabla f(x)\}$, these second-order subderivatives coincide with each other and satisfy

$$d_L^2 f(x|\nabla f(x))(u) = d_-^2 f(x|\nabla f(x))(u) = d_+^2 f(x|\nabla f(x))(u) = \frac{1}{2}\langle \nabla^2 f(x)u, u\rangle.$$

For other second-order subderivatives, the reader is referred to [7, 17] and the references therein.

For a nonempty set C in a normed linear space X, ψ_C denotes the indicator function associated with the set C defined as below:

$$\psi_C(x) = \begin{cases} 0 & \text{if } x \in C,\\ +\infty & \text{otherwise.}\end{cases}$$

2. Sufficient conditions in terms of subdifferentials. We recall the concept of an abstract subdifferential introduced in [21].

DEFINITION 2.1. *Let X be a Banach space, and let $f : X \to (-\infty, +\infty]$ be l.s.c. at $x \in X$ with $f(x) < +\infty$. A subset of X^*, denoted by $\partial_\omega f(x)$, is called a* ∂_ω-subdifferential of f *at x if it has the following properties:*

(ω_1) $\partial_\omega g(x) = \partial_\omega f(x)$ *if $g = f$ near x.*

(ω_2) $0 \in \partial_\omega f(x)$ *when f attains a local minimum at x.*

(ω_3) $\partial_\omega f(x) \subseteq L\overline{B}^*$ *if f is convex and Lipschitz of L near x.*

(ω_4) *If $g : X \to (-\infty, +\infty]$ is Lipschitz near x, then for each $\xi \in \partial_\omega(f + g)(x)$ and each $\delta > 0$ there exist $x_1, x_2 \in B(x, \delta)$ such that*

$$-\delta < f(x_1)-f(x) < \delta, \quad -\delta < g(x_2)-g(x) < \delta, \quad \textit{and } \xi \in \partial_\omega f(x_1)+\partial_\omega g(x_2)+\delta B^*,$$

where B^* is the open unit ball in X^* and $\overline{B}^*$ is its closure.

As indicated in [21], ∂_ω-subdifferentials include the Clarke subdifferential and the Michel–Penot subdifferential in a Banach space, the Fréchet subdifferential in an Asplund space, the proximal subdifferential in a Hilbert space, and the lower Dini subdifferential in R^n. Thus these subdifferentials can be taken as ∂_ω-subdifferentials in our main result of this section below whose proof is based on Ioffe's technique [9].

THEOREM 2.2. *Let X be a Banach space and let $f : X \to (-\infty, +\infty]$ be l.s.c. Suppose that, for some $x_0 \in X$, $0 < \delta \leq +\infty$, $0 < \mu < +\infty$, and $0 < \epsilon \leq \delta/(2\mu)$, the set $B(x_0, \delta/2) \cap f^{-1}(-\infty, \epsilon)$ is nonempty and*

$$\|\xi\|_* \geq \mu^{-1} \text{ for all } \xi \in \partial_\omega f(x) \text{ and each } x \in B(x_0, \delta) \cap f^{-1}(0, \epsilon).$$

Then $S := \{x \in X : f(x) \leq 0\}$ is nonempty and

$$d_S(x) \leq \mu f(x)_+ \text{ for all } x \in B\left(x_0, \frac{\delta}{2}\right) \cap f^{-1}(-\infty, \epsilon).$$

Moreover, if $x_0 \in S$, then the condition $0 < \epsilon \leq \delta/(2\mu)$ can be replaced with $0 < \epsilon \leq +\infty$.

Proof. Obviously it suffices to prove that

$$d_S(x) \leq \mu f(x)_+ \text{ for all } x \in B\left(x_0, \frac{\delta}{2}\right) \cap f^{-1}(-\infty, \epsilon)$$

since this together with the nonemptiness of the set $B(x_0, \delta/2) \cap f^{-1}(-\infty, \epsilon)$ implies the nonemptiness of S.

Suppose that there were $u \in B(x_0, \delta/2) \cap f^{-1}(-\infty, \epsilon)$ such that $d_S(u) > \mu f(u)_+$. Then $u \notin S$ and hence $0 < f(u) < \epsilon$. In addition, we can choose $t > 1$ and $\alpha > 0$ such that

$$d_S(u) > t\mu f(u) := \gamma \ \text{ and } \left\{ \begin{array}{ll} \max\{\gamma, \|u - x_0\|\} \leq \frac{\delta}{2+\alpha} & \text{for } 0 < \epsilon \leq \delta/(2\mu); \\ \|u - x_0\| \leq \frac{\delta}{2+\alpha} & \text{for } x_0 \in S \text{ and } 0 < \epsilon \leq +\infty. \end{array} \right. \tag{1}$$

Thus $f(u)_+ = f(u) = \gamma(t\mu)^{-1}$ and hence

$$f(u)_+ \leq \inf_{v \in X} f(v)_+ + \gamma(t\mu)^{-1}.$$

Note that the function $f(\cdot)_+$ is l.s.c. and bounded below. Applying Ekeland's variational principle [5] to $f(\cdot)_+$ with $\sigma = \gamma(t\mu)^{-1}$ and $\lambda = \gamma$, we find $x \in X$ satisfying

$$f(x)_+ \leq f(u)_+, \tag{2}$$

$$\|x - u\| \leq \gamma, \tag{3}$$

$$f(v)_+ + (t\mu)^{-1} h(v) \geq f(x)_+ \text{ for all } v \in X, \tag{4}$$

where $h(v) := \|v - x\|$. It follows from (1), (2), and (3) that $0 < f(x) < \epsilon$.

On the other hand, (4) implies that the function $f(v)_+ + (t\mu)^{-1}h(v)$ attains its minimum on X at x. Hence, by property (ω_2) in Definition 2.1,

$$0 \in \partial_\omega [f(x)_+ + (t\mu)^{-1} h(x)]. \tag{5}$$

Since f is l.s.c. and $0 < f(x)$, there exists $\delta_1 > 0$ such that

$$0 < f(y) \text{ for all } y \in B(x, \delta_1).$$

Thus, by property (ω_1) in Definition 2.1 and (5),

$$0 \in \partial_\omega (f + (t\mu)^{-1} h)(x). \tag{6}$$

Let $\epsilon_1 := \min\{f(x), (1-t^{-1})\mu^{-1}, \delta_1, \epsilon - f(u), \alpha\delta(2+\alpha)^{-1}\} > 0$. Then by property (ω_4) in Definition 2.1 and (6) there exist x_1 and x_2 both in $B(x, \epsilon_1)$ such that

$$f(x) - \epsilon_1 < f(x_1) < f(x) + \epsilon_1$$

and

$$0 \in \partial_\omega f(x_1) + \partial_\omega ((t\mu)^{-1} h)(x_2) + \epsilon_1 B^*.$$

These inequalities with (2) mean that $x_1 \in B(x, \epsilon_1) \cap f^{-1}(0, \epsilon)$. The inclusion, by property (ω_3) in Definition 2.1, implies that there exists $\xi \in \partial_\omega f(x_1)$ such that

$$\|\xi\|_* < (t\mu)^{-1} + \epsilon_1 \le (t\mu)^{-1} + (1 - t^{-1})\mu^{-1} = \mu^{-1},$$

which contradicts the assumption since $x_1 \in f^{-1}(0, \epsilon)$ and, by the triangle inequality and (1),

$$\|x_1 - x_0\| \le \|x_1 - x\| + \|x - u\| + \|u - x_0\| < \epsilon_1 + \gamma + \frac{\delta}{2+\alpha}$$
$$\le \begin{cases} \frac{\alpha\delta}{2+\alpha} + \frac{2\delta}{2+\alpha} = \delta & \text{for } 0 < \epsilon \le \delta/(2\mu); \\ \frac{(1+\alpha)\delta}{2+\alpha} + d_S(u) \le \frac{(1+\alpha)\delta}{2+\alpha} + \|u - x_0\| \le \delta & \text{for } x_0 \in S \text{ and } 0 < \epsilon \le +\infty. \end{cases} \qquad \square$$

Remark 2.1. Note that the nonemptiness of S in Theorem 2.2 is a natural result of the inequality for error bounds and the nonemptiness of the set $B(x_0, \delta/2) \cap f^{-1}(-\infty, \epsilon)$. It is worth comparing Theorem 2.2 with [22, Theorem 4], in which the nonemptiness of S can follow from an existence theorem of minimum in [18]. When f is regular, $f^-(x; v) = f^\circ(x; v)$ holds for each $x \in X$ and $v \in X$. The condition that $f^-(x; h_x) \le -\mu^{-1}$ for some $\mu > 0$, each $x \in f^{-1}(0, \epsilon)$, and corresponding h_x in [22, Theorem 4] turns into $f^\circ(x; h_x) \le -\mu^{-1}$, which implies that $\|\xi\|_* \ge \mu^{-1}$ for each $\xi \in \partial^\circ f(x)$. So the corresponding result of [22, Theorem 4] can be obtained from Theorem 2.2 by taking $\delta = +\infty$ and $\partial_\omega = \partial^\circ$. Hence Theorem 2.2 provides a weaker condition for the existence of solutions for an inequality system than that in [22, Theorem 4].

Theorem 2.2 is an extension of [21, Theorem 3.1] in that $B(x_0, \delta) \cap f^{-1}(0, \epsilon)$, not just $B(x_0, \delta)$ or $f^{-1}(0, \epsilon)$, can be taken as a test set T. In particular, for the case where the test set $T = f^{-1}(0, \epsilon)$, Theorem 2.2 is a refinement of [21, Theorem 3.1], in which the nonemptiness of S is a part of the assumption, not of the conclusion. In addition, the inequality $d_S(x) \le \mu f(x)_+$ in [21, Theorem 3.1] holds only for all $x \in X$ with $f(x) < \epsilon/2$ instead of for all $x \in X$ with $f(x) < \epsilon$, as in Theorem 2.2. We thank Dr. Qiji Jim Zhu for his help in the proof of this improvement.

For an l.s.c. function f on a Hilbert space X, the *limiting subdifferential* $\partial_L f(x)$ of f at $x \in dom\, f$ is a set defined by

$$\partial_L f(x) := \{w\text{-}\lim \xi_i : \xi_i \in \partial^\pi f(x_i), x_i \to x, f(x_i) \to f(x)\}.$$

That is, $\partial_L f(x)$ consists of all vectors, each of which is the weak limit (that is what w-$\lim \xi_i$ signifies) of a weak convergent sequence $\{\xi_i\}$, where $\xi_i \in \partial^\pi f(x_i)$ with $x_i \to x$

and $f(x_i) \to f(x)$. It is easy to check that the limiting subdifferential satisfies (ω_1)–(ω_3) in Definition 2.1. In addition, if at least one of functions f and g is Lipschitz near x, then

$$\partial_L(f+g)(x) \subseteq \partial_L f(x) + \partial_L g(x)$$

[4, Proposition 10.1, p. 62]; that is, the sum rule holds. So the limiting subdifferential is a ∂_ω-subdifferential and Theorem 2.2 is applicable to it. The following is a version of Theorem 2.2 with $\partial_\omega = \partial_L$ and f replaced with $f + \psi_C$.

COROLLARY 2.3. *Let X be a Hilbert space, let C be a closed subset of X, and let $f_i : X \to R$ be locally Lipschitz continuous for each $i \in I$. Denote*

$$f(x) = \max\{f_i(x) : i \in I\} \quad \text{and} \quad I(x) := \{i \in I : f_i(x) = f(x)\} \text{ for } x \in X.$$

Suppose that, for some $x_0 \in X$, $0 < \delta \leq +\infty$, $0 < \mu < +\infty$, and $0 < \epsilon \leq \delta/(2\mu)$, the set $C \cap B(x_0, \delta/2) \cap f^{-1}(-\infty, \epsilon)$ is nonempty and

$$\|\xi\|_* \geq \mu^{-1} \text{ for all } \xi \in co\{\partial_L f_i(x) : i \in I(x)\} + N_C^L(x) \text{ for all } x \in C \cap B(x_0, \delta) \cap f^{-1}(0, \epsilon),$$

where $co\, A$ denotes the convex hull of a set A and $N_C^L(x) := \partial_L \psi_C(x)$. Then $S := \{x \in C : f(x) \leq 0\}$ is nonempty and

$$d_S(x) \leq \mu f(x)_+ \text{ for all } x \in C \cap B\left(x_0, \frac{\delta}{2}\right) \cap f^{-1}(-\infty, \epsilon).$$

Moreover, if $x_0 \in S$, then the condition $0 < \epsilon \leq \delta/(2\mu)$ can be replaced with $0 < \epsilon \leq +\infty$.

Proof. For each $\xi \in \partial_L f(x)$, by the conclusion in [4, Problem 11.17, p. 65] and the sum rule, there exist $\gamma_i \geq 0\,(i \in I(x))$ with $\sum_{i \in I(x)} \gamma_i = 1$ such that

$$\xi \in \partial_L \left(\sum_{i \in I(x)} \gamma_i f_i \right)(x) \subseteq co\{\partial_L f_i(x) : i \in I(x)\}.$$

Hence applying Theorem 2.2 to $\partial_\omega = \partial_L$ with f replaced with $f + \psi_C$ completes the proof. $\square$

Next we use Theorem 2.2 to prove a result about the regularity of a set at a point.

THEOREM 2.4. *Let X be a separable Hilbert space, C a closed subset of X, and $x_0 \in C$. Suppose that $g : X \to R^m$ and $h : X \to R^n$ are Lipschitz near x_0 and*

$$f(x) = \max_{i,\,j}\{g_i(x), |h_j(x)|\}.$$

If the constraint qualification

$$\left.\begin{array}{r} 0 \leq \gamma \in R^m,\ \gamma_i[g_i(x_0) - f(x_0)] = 0,\ i \in I \\ \lambda \in R^n,\ \lambda_j[|h_j(x_0)| - f(x_0)] = 0,\ j \in J \\ 0 \in \partial_L[\langle \gamma, g\rangle + \langle \lambda, h\rangle](x_0) + N_C^L(x_0) \end{array}\right\} \Rightarrow \gamma = 0 \text{ and } \lambda = 0$$

(where $N_C^L(x_0) := \partial_L \psi_C(x_0)$) is satisfied at x_0, then there exist $0 < \delta < +\infty$ and $0 < \mu < +\infty$ such that

$$\|\xi\| \geq \mu^{-1} \text{ for all } \xi \in \partial^\pi(f + \psi_C)(x) \text{ and each } x \in C \cap B(x_0, \delta) \cap f^{-1}(0, +\infty).$$

Moreover, if the set $C \cap B(x_0, \delta/2) \cap f^{-1}[0, \epsilon)$ *is nonempty for some* $0 < \epsilon \le \delta/(2\mu)$, *then the set* $S := \{x \in C : g(x) \le 0,\ h(x) = 0\}$ *is nonempty and*

$$d_S(x) \le \mu f(x) \text{ for all } x \in C \cap B\left(x_0, \frac{\delta}{2}\right) \cap f^{-1}(0, \epsilon).$$

In particular, if $x_0 \in S$, *then* S *is metrically regular at* x_0. *If* x_0 *lies in the interior of* C, *then the above conclusions hold in every Hilbert space* X.

Proof. Suppose that there did not exist $0 < \delta < +\infty$ and $0 < \mu < +\infty$ such that

$$\|\xi\| \ge \mu^{-1} \text{ for all } \xi \in \partial^\pi(f + \psi_C)(x) \text{ and each } x \in C \cap B(x_0, \delta) \cap f^{-1}(0, +\infty).$$

Then there would exist sequences

$$C \ni x_k \to x_0, \quad f(x_k) > 0, \quad \xi_k \in \partial^\pi(f + \psi_C)(x_k), \quad \|\xi_k\| \to 0.$$

If x_0 is in the interior of C, then $\xi_k \in \partial^\pi f(x_k)$ and $\|\xi_k\| \to 0$ imply that $0 \in \partial_L f(x_0)$. Thus there exist $0 \le \gamma \in R^m$ and $\lambda \in R^n$ such that

$$\gamma_i[g_i(x_0) - f(x_0)] = 0 \text{ for } i \in I, \tag{7}$$

$$\lambda_j[|h_j(x_0)| - f(x_0)] = 0 \text{ for } j \in J, \tag{8}$$

$$\sum_{i=1}^m \gamma_i + \sum_{j=1}^n |\lambda_j| = 1, \tag{9}$$

$$0 \in \partial_L[\langle \gamma, g\rangle + \langle \lambda, h\rangle](x_0)$$

(see [4, Problem 1.11.17, p. 65]), which contradicts the assumption.

If x_0 is not in the interior of C, then since f is Lipschitz near x_k when k is large enough, by [4, Theorem 1.8.3, p. 56], there exist $y_k \to x_0$, $C \ni z_k \to x_0$, $\eta_k \in \partial^\pi f(y_k)$, and $\zeta_k \in \partial^\pi \psi_C(z_k)$ such that $f(y_k) > 0$ and

$$\xi_k \in \eta_k + \zeta_k + B\left(x_0, \frac{1}{k}\right). \tag{10}$$

Since $\partial^\pi f(y_k) \subseteq \partial_L f(y_k)$, for k large enough so that y_k enters some prescribed neighborhood of x_0 on which f is Lipschitz, there exist $0 \le \gamma^k \in R^m$ and $\lambda^k \in R^n$ such that

$$\gamma_i^k[g_i(y_k) - f(y_k)] = 0 \text{ for } i \in I,$$

$$\lambda_j^k[|h_j(y_k)| - f(y_k)] = 0 \text{ for } j \in J,$$

$$\sum_{i=1}^m \gamma_i^k + \sum_{j=1}^n |\lambda_j^k| = 1,$$

$$\eta_k \in \partial_L[\langle \gamma^k, g\rangle + \langle \lambda^k, h\rangle](y_k).$$

By extracting convergent subsequences of $\{\gamma^k\}$ and $\{\lambda^k\}$ (we do not relabel them) and taking the limit of (γ^k, λ^k), we obtain a nonzero $(\gamma, \lambda) \in R^m \times R^n$ satisfying (7)–(9).

Note that

$$\partial_L[\langle \gamma^k, g\rangle + \langle \lambda^k, h\rangle](y_k) \subseteq \partial^\circ[\langle \gamma^k, g\rangle + \langle \lambda^k, h\rangle](y_k)$$

and the set on the right-hand side is contained in a ball of the form $L\overline{B}_*$ (for some positive L) which is weak* compact when k is large enough. There is a weakly convergent subsequence of $\{\eta_k\}$ (without relabeling) corresponding to (γ^k, λ^k) whose weak limit lies in

$$\partial_L[\langle \gamma, g\rangle + \langle \lambda, h\rangle](x_0)$$

since X is a separable Hilbert space (see [4, Problem 1.11.21, p. 66]).

In addition, corresponding to η_k, by (10), the sequence $\{\zeta_k\}$ contains a weakly convergent subsequence with its limit belonging to $N_C^L(x_0)$. Therefore we have

$$0 \in \partial_L[\langle \gamma, g\rangle + \langle \lambda, h\rangle](x_0) + N_C^L(x_0),$$

but (γ, λ) is nonzero. This is again a contradiction.

The rest follows immediately from the conclusion shown above and from Theorem 2.2. $\square$

Remark 2.2. Theorem 2.4 is a refinement of [4, Theorem 3.8, p. 131] in that x_0 may not be in S and an abstract constraint set is allowed. In a general Banach space, one relevant result about metrical regularity in terms of Clarke subdifferentials can be found in [3, Theorem 6.6.1]. However, in Hilbert space where limiting subdifferential is applicable, our constraint qualification is weaker than that in [3, Theorem 6.6.1].

If $x_0 \in S$, $g_1, \ldots, g_m, h_1, h_2, \ldots, h_n$ are all C^1 functions and $C = X$, the constraint qualification in Theorem 2.4 is equivalent to the Mangasarian–Fromovitz constraint qualification in mathematical programming. In particular, if

$$\nabla g_1(x_0), \ldots, \nabla g_m(x_0), \nabla h_1(x_0), \ldots, \nabla h_n(x_0)$$

are linearly independent, then the Mangasarian–Fromovitz constraint qualification is satisfied at x_0.

Example 2.1. For $x \in R^3$, let

$$f_1(x) := ax_1 + g_1(x_2, x_3), \quad f_2(x) = bx_2 + g_2(x_3), \quad f_3(x) = cx_3,$$

where a, b, and c are nonzero constants while g_1 and g_2 are locally Lipschitz continuous. Since, for any point $x_0 \in R^3$, $\nabla f_1(x_0), \nabla f_2(x_0), \nabla f_3(x_0)$ are linearly independent, by Theorem 2.4, the system $S = \{x \in R^3 : f(x) \leq 0\}$ with $f(x) := \max\{f_1(x), f_2(x), f_3(x)\}$ is metrically regular at any $x_0 \in S$.

Note that for an l.s.c. convex function f on a Banach space X the Clarke subdifferential of f at $x \in X$ reduces to the subdifferential of f at x in the sense of convex analysis given by

$$\partial f(x) := \{\xi \in X^* : \langle \xi, y - x\rangle \leq f(y) - f(x) \text{ for all } y \in X\}.$$

It has been shown in [22] that for a convex inequality system a global error bound exists iff a local error bound does, and many first-order sufficient conditions for the existence of error bounds become necessary as well. In the following result, we use $\partial f(x)$ to develop the sufficient condition stated in Theorem 2.2 into a necessary one for a convex system.

THEOREM 2.5. *Let X be a Banach space, let $f : X \to (-\infty, +\infty]$ be l.s.c. and convex, and let $S := \{x \in X : f(x) \leq 0\}$. Then for some $x_0 \in X$ and $0 < \mu < +\infty$ the following are equivalent:*

(i) *For some* $0 < \delta \le +\infty$, *each* $0 < \epsilon \le \delta/(2\mu)$, *and each* $\delta' \in (0, \delta)$ *the set* $B(x_0, \delta') \cap f^{-1}(-\infty, \epsilon)$ *is nonempty and*

$$\|\xi\|_* \ge \mu^{-1} \textit{ for all } \xi \in \partial f(x) \textit{ and each } x \in B(x_0, \delta) \cap f^{-1}(0, \epsilon).$$

(ii) *For some* $0 < \delta \le +\infty$, *each* $0 < \epsilon \le \delta/(2\mu)$, *and each* $\delta' \in (0, \delta)$ *the set* $B(x_0, \delta') \cap f^{-1}(-\infty, \epsilon)$ *is nonempty and*

$$d_S(x) \le \mu f(x)_+ \textit{ for all } x \in B(x_0, \delta) \cap f^{-1}(-\infty, \epsilon).$$

In particular, if $x_0 \in S$, *then* (i) *and* (ii) *are equivalent to each other with "each* $0 < \epsilon \le \delta/(2\mu)$*" in both replaced by "some* $0 < \epsilon \le +\infty$*."*

Proof. (i) $\Rightarrow$ (ii) This is immediate from Theorem 2.2 by taking $\partial_\omega f(x) = \partial f(x)$.

(ii) $\Rightarrow$ (i) Let $x \in B(x_0, \delta) \cap f^{-1}(0, \epsilon)$. Then $d_S(x) > 0$ and for any $\xi \in \partial f(x)$ we have

$$\|\xi\|_* \cdot \|y - x\| \ge -\langle \xi, y - x \rangle \ge -[f(y) - f(x)] \ge f(x) \text{ for all } y \in S.$$

This implies $\|\xi\|_* \cdot d_S(x) \ge f(x)$, from which we have

$$\|\xi\|_* \ge \frac{f(x)}{d_S(x)} \ge \mu^{-1}.$$

Therefore the desired inequality follows. $\square$

3. Second-order conditions. In mathematical programming, it is known that a second-order sufficient condition implies strict local minimum of order 2. This idea can be applied to error bounds. For a nonnegative function $f : R^n \to R$, consider the inequality system $S = \{x \in R^n : f(x) \le 0\}$. If $x_0 \in S$, f is twice continuously differentiable near x_0, and there exist $\mu > 0$ and $\delta > 0$ such that

$$\langle \nabla^2 f(x') u, u \rangle \ge \mu^{-1} \quad \text{for each unit vector } u \in R^n \text{ and } x' \in B(x_0, \delta), \tag{11}$$

then for each $x \in B(x_0, \delta)$, by the Taylor expansion, there exists $x' \in [x_0, x]$ such that

$$f(x) = \frac{1}{2} \langle \nabla^2 f(x')(x - x_0), x - x_0 \rangle,$$

which along with (11) implies that

$$f(x) \ge \frac{1}{2\mu} \|x - x_0\|^2.$$

Thus

$$d_S^2(x) \le 2\mu f_+(x) \text{ for all } x \in B(x_0, \delta).$$

Note that under the above assumption, S must be a singleton. In studying weak sharp minima, several authors, including Bonnans and Ioffe [1, 2] and Ward [20] have extended the above result to include the case where f is not twice continuously differentiable and the solution set S is not a singleton by using certain second-order subderivatives. In the following main result in this section, we present a second-order sufficient condition for the existence of error bound with exponent 1/2. Note that if f

is nonnegative and twice continuously differentiable, then our second-order condition in Theorem 3.1 amounts to

$$\langle \nabla^2 f(x)u_x, u_x\rangle \le -4\mu^{-1} \qquad \text{for some unit vector } u_x \in X \text{ and each } x \notin S.$$

Hence, unlike the second-order condition of type (11), which requires certain convexity, our second-order condition is suitable for nonconvex systems.

THEOREM 3.1. *Let X be a Hilbert space and let $f : X \to (-\infty, +\infty]$ be l.s.c. Suppose that, for some $x_0 \in X$, $0 < \delta \le +\infty$, $0 < \mu$, and $0 < \epsilon \le (2\mu)^{-1}(\delta/2)^2$, the set $B(x_0, \delta/2) \cap f^{-1}(-\infty, \epsilon)$ is nonempty and that, for each $x \in B(x_0, \delta) \cap f^{-1}(0, \epsilon)$, one of the following is satisfied for each $\xi \in \partial^\pi f(x)$ with $\|\xi\| \le \min\{2\sqrt{2\epsilon}\,\mu^{-1/2}, \delta\mu^{-1}\}$:*

(i) *There exists a unit vector u_x such that $d^2_- f(x|\xi)(u_x) \le -2\mu^{-1}$.*

(ii) *There exist sequences $t_n \to 0^+$ in R and $\{u_n\}$ in X such that $\lim_{n\to+\infty} \|u_n\| = 1$ and*

$$\lim_{n\to+\infty} \frac{f(x + t_n u_n)_+ - f(x)_+ - t_n\langle \xi, u_n\rangle}{t_n^2} \le -2\mu^{-1}.$$

Then $S := \{x \in X : f(x) \le 0\}$ is nonempty and

$$d_S^2(x) \le 2\mu f(x)_+ \text{ for all } x \in B\left(x_0, \frac{\delta}{2}\right) \cap f^{-1}(0, \epsilon).$$

Moreover, if $x_0 \in S$, then the condition $0 < \epsilon \le (2\mu)^{-1}(\delta/2)^2$ can be replaced with $0 < \epsilon \le +\infty$.

Proof. By the definition of the second-order subderivative, condition (i) implies condition (ii). Hence it suffices to prove the theorem under condition (ii).

We now prove the theorem by contradiction. Suppose that there were $u \in B(x_0, \delta/2) \cap f^{-1}(0, \epsilon)$ such that $d_S^2(u) > 2\mu f(u)_+$. We choose $t > 1$ such that

$$\text{(12)} \qquad 4\gamma := 2t\mu f(u) < \begin{cases} \min\{d_S^2(u), (\frac{\delta}{2})^2\} & \text{for } 0 < \epsilon \le (2\mu)^{-1}(\frac{\delta}{2})^2; \\ d_S^2(u) & \text{for } x_0 \in S \text{ and } 0 < \epsilon \le +\infty. \end{cases}$$

Thus $f(u) = 2\gamma(t\mu)^{-1}$ and hence

$$f(u)_+ \le \inf_{v\in X} f(v)_+ + 2\gamma(t\mu)^{-1}.$$

Note that the function $f(\cdot)_+$ is l.s.c. and bounded below. Applying smooth variational principle [4, Theorem 4.2, p. 43] to $f(\cdot)_+$ with $\sigma = 2\gamma(t\mu)^{-1}$ and $\lambda = \sqrt{\gamma}$, we find $x, y \in X$ satisfying

$$\|y - u\| < \lambda, \quad \|x - y\| < \lambda, \quad f(x)_+ \le f(u)_+$$

and

$$\text{(13)} \qquad f(v)_+ + 2(t\mu)^{-1}h(v) \ge f(x)_+ + 2(t\mu)^{-1}h(x) \text{ for all } v \in X,$$

where $h(v) := \|v - y\|^2$. Thus

$$\|x - u\| \le \|x - y\| + \|y - u\| < 2\lambda = 2\sqrt{\gamma} < d_S(u)$$

and, by the triangle inequality and (12),

$$\begin{aligned} \|x - x_0\| &\le \|x - u\| + \|u - x_0\| \\ &< \begin{cases} 2\sqrt{\gamma} + \frac{\delta}{2} < \min\{d_S(u), \frac{\delta}{2}\} + \frac{\delta}{2} \le \delta & \text{for } 0 < \epsilon \le (2\mu)^{-1}(\frac{\delta}{2})^2; \\ d_S(u) + \|u - x_0\| < \delta & \text{for } x_0 \in S \text{ and } 0 < \epsilon \le +\infty \end{cases} \end{aligned}$$

and hence $x \in B(x_0, \delta) \cap f^{-1}(0, \epsilon)$.

On the other hand, from (13) and [4, Proposition 1.2.11, p. 38], we have

$$0 \in \partial^\pi (f(x) + 2(t\mu)^{-1} h(x)) = \partial^\pi f(x) + 2(t\mu)^{-1}\{2(x-y)\}. \tag{14}$$

This implies that $\xi := 4(t\mu)^{-1}(y-x) \in \partial^\pi f(x)$ and hence, by (12),

$$\begin{aligned}\|\xi\| &\le 4\|y-x\|(t\mu)^{-1} < 4\lambda(t\mu)^{-1} = 4\sqrt{\gamma}(t\mu)^{-1} = 2\sqrt{2t\mu f(u)}(t\mu)^{-1}\\ &\le 2\min\left\{\sqrt{2t\mu\epsilon}, d_S(u), \frac{\delta}{2}\right\}(t\mu)^{-1} < \min\{2\sqrt{2\epsilon}\,\mu^{-1/2}, \delta\mu^{-1}\}.\end{aligned}$$

So for the sequences $\{t_n\}$ and $\{u_n\}$ given in condition (ii) corresponding to ξ, by (13), we have

$$\begin{aligned}&\lim_{n\to+\infty} \frac{f(x+t_n u_n)_+ - f(x)_+ - t_n\langle \xi, u_n\rangle}{t_n^2}\\ &= \lim_{n\to+\infty} \frac{f(x+t_n u_n)_+ + 2(t\mu)^{-1}h(x+t_n u_n) - f(x)_+ - 2(t\mu)^{-1}h(x)}{t_n^2} - 2(t\mu)^{-1}\\ &\ge -2(t\mu)^{-1} > -2\mu^{-1},\end{aligned}$$

which contradicts condition (ii). ◻

To put first-order and second-order conditions together, we will use the following relation between a global error bound and a local error bound.

PROPOSITION 3.2. *Let (X, d) be a metric space, let $f : X \to (-\infty, +\infty]$ be proper, and let $S := f^{-1}(-\infty, 0]$. Then the following are equivalent:*

(i) *There exist $0 < \epsilon_1 < \epsilon_2 \le +\infty$ and $0 < \mu_1, \mu_2 < +\infty$ such that*

$$\begin{aligned}d_S(x) &\le \mu_1 f(x)_+ \quad \text{for all } x \in f^{-1}(0, \epsilon_1) \text{ and}\\ d_{S_1}(x) &\le \mu_2 f(x)_+ \quad \text{for all } x \in f^{-1}[\epsilon_1, \epsilon_2),\end{aligned}$$

where $S_1 := f^{-1}(-\infty, \epsilon_1)$.

(ii) *There exist $0 < \epsilon \le +\infty$ and $0 < \mu < +\infty$ such that*

$$d_S(x) \le \mu f(x)_+ \quad \text{for all } x \in f^{-1}(0, \epsilon).$$

Proof. The implication (ii) ⇒ (i) is immediate. We only need to show (i) ⇒ (ii). Let $0 < \epsilon_1 < \epsilon_2 \le +\infty$ and $0 < \mu_1, \mu_2 < +\infty$ satisfy

$$\begin{aligned}d_S(x) &\le \mu_1 f(x)_+ \quad \text{for all } x \in f^{-1}(0, \epsilon_1) \text{ and}\\ d_{S_1}(x) &\le \mu_2 f(x)_+ \quad \text{for all } x \in f^{-1}[\epsilon_1, \epsilon_2),\end{aligned}$$

where $S_1 := f^{-1}(-\infty, \epsilon_1)$. Note that for any fixed $x \in f^{-1}[\epsilon_1, \epsilon_2)$ and each $y \in f^{-1}(-\infty, \epsilon_1)$ we have

$$d_S(x) \le d_S(y) + d(x,y) \le \mu_1 f(y)_+ + d(x,y) \le \mu_1\epsilon_1 + d(x,y).$$

Taking the inferior of the right-hand side expression in the above inequalities for y over $f^{-1}(-\infty, \epsilon_1)$ yields $d_S(x) \le \mu_1\epsilon_1 + d_{S_1}(x)$. And hence

$$d_S(x) \le \mu_1\epsilon_1 + \mu_2 f(x)_+ \le (\mu_1 + \mu_2) f(x)_+ = \mu f(x)_+$$

for $\mu := \mu_1 + \mu_2$. Therefore, (ii) holds for $\epsilon = \epsilon_2$. ◻

Remark 3.1. When X is a normed linear space and f is convex, we can prove that the nonemptiness of S and the first inequality in (i) of Proposition 3.2 imply the second inequality in it. So Proposition 3.2 is an extension of [22, Proposition 2], which states that for a convex system a local error bound implies a global error bound.

Next, we use Proposition 3.2 and Theorems 2.2 and 3.1 to give a mixed condition.

THEOREM 3.3. *Let X be a Hilbert space, and let $f : X \to (-\infty, +\infty]$ be continuous. Denote*

$$D(\mu) := \{x \in X : 0 < f(x) \text{ and } \|\xi\| \le \mu^{-1} \text{ for some } \xi \in \partial^{\pi} f(x)\} \text{ for } \mu > 0.$$

Suppose that there exist $0 < \epsilon_1 < \epsilon_2 \le +\infty$ and $0 < \mu_1, \mu_2$ such that the set $f^{-1}(-\infty, \epsilon_1)$ is nonempty and the following conditions hold:

(i) *$D(\mu_1) \subseteq f^{-1}(\epsilon_1, \epsilon_2)$.*

(ii) *For each $x \in f^{-1}(\epsilon_1, \epsilon_2)$ there exists a unit vector u_x such that*

$$d^2_- f(x|\xi)(u_x) \le -2\mu_2^{-1} \text{ for all } \xi \in \partial^{\pi} f(x) \text{ with } \|\xi\| \le 2\sqrt{2(\epsilon_2 - \epsilon_1)/\mu_2}.$$

Then $S := \{x \in X : f(x) \le 0\}$ is nonempty and

$$d_S(x) \le \mu f(x) \text{ for all } x \in f^{-1}(0, \epsilon_2),$$

where $\mu = \mu_1 + (2\mu_2/\epsilon_1)^{1/2}$.

Proof. Since condition (i) implies that

$$\|\xi\| > \mu_1^{-1} \text{ for all } \xi \in \partial^{\pi} f(x) \text{ and each } x \in f^{-1}(0, \epsilon_1),$$

applying Theorem 2.2 to the function f with $\partial_\omega = \partial^\pi$, we obtain that S is nonempty and

$$d_S(x) \le \mu_1 f(x)_+ \text{ for all } x \in f^{-1}(-\infty, \epsilon_1).$$

This also holds for all $x \in X$ satisfying $f(x) = \epsilon_1$ by the continuity of f and d_S.

Next, by applying Theorem 3.1 to the function $f(\cdot) - \epsilon_1$, we have

$$d_{f^{-1}(-\infty,\epsilon_1]}(x) \le \sqrt{2\mu_2[f(x) - \epsilon_1]} < \sqrt{\frac{2\mu_2}{\epsilon_1}}\, f(x) \text{ for all } x \in f^{-1}(\epsilon_1, \epsilon_2).$$

Thus, by Proposition 3.2, for $\mu = \mu_1 + (2\mu_2/\epsilon_1)^{1/2}$ we have

$$d_S(x) \le \mu f(x)_+ \text{ for all } x \in f^{-1}(0, \epsilon_2). \qquad \square$$

Remark 3.2. Only in a Hilbert space is Theorem 3.3 established, unlike [7, Theorem 3.2], which is given in a Banach space for the case $\epsilon_2 = +\infty$. However, the function f in [7, Theorem 3.2] needs to be not only continuous but also Gâteaux differentiable, while the inequality $d^2_- f(x|\xi)(u_x) \le -2\mu_2^{-1}$ in (ii) is required to hold for each x in $D(\mu_1)\backslash f^{-1}(-\infty, \epsilon_1]$ and for all points in the corresponding interval $(x, x + Tu_x)$ for some $T > 0$. In Theorem 3.3, we do not restrict ϵ_2 to equal $+\infty$ nor require the condition $d^2_- f(x|\xi)(u_x) \le -2\mu_2^{-1}$ to be satisfied in the interval $(x, x + Tu_x)$ for each $x \in f^{-1}(\epsilon_1, \epsilon_2]$.

In what follows, we use Theorem 3.1 to develop sufficient conditions for a system of inequalities, equalities, and an abstract constraint to have error bounds in terms of the second-order subderivatives of the functions involved and certain tangent cones to the abstract constraint set.

We first review some concepts about tangent cone and contingent cone briefly. For a closed subset C in a Banach space X and $x \in C$, the *tangent cone* to C at x, denoted $T_C(x)$, is defined as

$$T_C(x) := \{v \in X : d_C^\circ(x; v) = 0\},$$

and the *contingent* (or the *Bouligand tangent*) *cone* to C at x, denoted $K_C(x)$, is given by

$$K_C(x) := \{v \in X : d_C^-(x; v) = 0\}.$$

It is well known that $v \in T_C(x)$ iff, for every sequence x_n in C converging to x and sequence t_n in $(0, +\infty)$ decreasing to 0, there is a sequence v_n in X converging to v such that $x_n + t_n v_n \in C$ for all n and that $v \in K_C(x)$ iff there exist $v_n \to v$ and $t_n \to 0^+$ such that $x + t_n v_n \in C$. Therefore we have the inclusive relation $T_C(x) \subseteq K_C(x)$.

We also recall that a vector v is *hypertangent* to the set C at the point x in C if there exists $0 < \epsilon$ such that

$$y + tw \in C \quad \text{for all } y \in B(x, \epsilon) \cap C,\ w \in B(v, \epsilon),\ t \in (0, \epsilon).$$

[3, Theorem 2.4.8] states that if the set of hypertangents to the set C at x is nonempty, then it coincides with $\operatorname{int} T_C(x)$, the interior of $T_C(x)$.

The above concepts turn out to be important for us to use Theorem 3.1 to give sufficient conditions for an inequality system with an abstract constraint set to have error bounds.

THEOREM 3.4. *Let X be a Hilbert space, let C be a nonempty closed set in X, and let $f : X \to (-\infty, +\infty]$ be an l.s.c. function. Suppose that, for some $x_0 \in X$, $0 < \delta \leq +\infty$, $0 < \mu$, and $0 < \epsilon \leq (2\mu)^{-1}(\delta/2)^2$, the set $C \cap B(x_0, \delta/2) \cap f^{-1}(-\infty, \epsilon)$ is nonempty and that, for each $x \in C \cap B(x_0, \delta) \cap f^{-1}(0, \epsilon)$, there exists a unit vector $u_x \in X$ such that*

(i) *u_x is hypertangent to C at x and satisfies*

$$d_-^2 f(x|\xi)(u_x) \leq -2\mu^{-1}$$

for each $\xi \in \partial^\pi(f + \psi_C)(x)$ with $\|\xi\| \leq \min\{2\sqrt{2\epsilon}\,\mu^{-1/2}, \delta\mu^{-1}\}$; or

(ii) *$u_x \in K_C(x)$ and satisfies*

$$d_+^2 f(x|\xi)(u_x) \leq -2\mu^{-1}$$

for all $\xi \in X$ with $\|\xi\| \leq \min\{2\sqrt{2\epsilon}\,\mu^{-1/2}, \delta\mu^{-1}\}$ and $\langle \xi, u_x\rangle \leq f^+(x; u_x)$.

Then $S := \{x \in C : f(x) \leq 0\}$ is nonempty and

$$d_S^2(x) \leq 2\mu f(x)_+ \text{ for all } x \in C \cap B\left(x_0, \frac{\delta}{2}\right) \cap f^{-1}(0, \epsilon).$$

Moreover, if $x_0 \in S$, then the condition $0 < \epsilon \leq (2\mu)^{-1}(\delta/2)^2$ can be replaced with $0 < \epsilon \leq +\infty$.

Proof. Let $x \in C \cap B(x_0, \delta) \cap f^{-1}(0, \epsilon)$. Based on Theorem 3.1, it suffices to show that there exists a unit vector $u_x \in X$ such that

$$d_-^2(f + \psi_C)(x|\xi)(u_x) \leq -2\mu^{-1}$$

for each $\xi \in \partial^\pi(f + \psi_C)(x)$ with $\|\xi\| \leq \min\{2\sqrt{2\epsilon}\,\mu^{-1/2}, \delta\mu^{-1}\}$.

Now if u_x is a unit hypertangent vector in (i), then, for each $\xi \in \partial^\pi(f + \psi_C)(x)$ with $\|\xi\| \le \min\{2\sqrt{2\epsilon}\,\mu^{-1/2}, \delta\mu^{-1}\}$, we have sequences $u_n \to u_x$ and $t_n \to 0^+$ such that $x + t_n u_n \in C$ and

$$\begin{aligned} d_-^2(f+\psi_C)(x|\xi)(u_x) &\le \lim_{n\to+\infty} \frac{f(x+t_nu_n) - f(x) - t_n\langle \xi, u_n\rangle}{t_n^2} \\ &= d_-^2 f(x|\xi)(u_x) \le -2\mu^{-1}. \end{aligned}$$

If $u_x \in K_C(x)$ is a unit vector in (ii), then there exist sequences $u_n \to u_x$ and $t_n \to 0^+$ such that $x + t_n u_n \in C$. It follows that for each $\xi \in \partial^\pi(f+\psi_C)(x)$ there exists some $M > 0$ such that

$$f(x+t_nu_n) - f(x) \ge t_n\langle \xi, u_n\rangle - Mt_n^2\|u_n\|^2$$

for sufficiently large n. This implies that $\langle \xi, u_x\rangle \le f^+(x; u_x)$ for each $\xi \in \partial^\pi(f + \psi_C)(x)$, that is,

$$\partial^\pi(f+\psi_C)(x) \subseteq \{\xi \in X : \langle \xi, u_x\rangle \le f^+(x;u_x)\}.$$

Thus for each $\xi \in \partial^\pi(f+\psi_C)(x)$ with $\|\xi\| \le \min\{2\sqrt{2\epsilon}\,\mu^{-1/2}, \delta\mu^{-1}\}$ we have

$$\begin{aligned} d_-^2(f+\psi_C)(x|\xi)(u_x) &\le \limsup_{n\to+\infty} \frac{f(x+t_nu_n) - f(x) - t_n\langle \xi, u_n\rangle}{t_n^2} \\ &\le d_+^2 f(x|\xi)(u_x) \le -2\mu^{-1}. \end{aligned}$$

The proof is therefore complete. □

Remark 3.3. From the above proof we see that Theorem 3.4 is a direct result of Theorem 3.1. Note that if x is an interior point of a closed subset C of X, then the set of hypertangents to the set C at x is just X. In particular, when $C = X$, each unit vector u_x is hypertangent to C at $x \in X$. In this case Theorem 3.4 reduces to Theorem 3.1. So they are in fact equivalent.

To apply Theorem 3.1 to a system of inequalities, we first give a result about the proximal subdifferential of the pointwise maxima function of several functions.

PROPOSITION 3.5. *Let $f_i : X \to R$ be Lipschitz near x for each $i \in I$. Denote*

$$f(x) = \max\{f_i(x) : i \in I\} \quad \text{and} \quad I(x) := \{i \in I : f_i(x) = f(x)\} \text{ for } x \in X.$$

Suppose that $\partial^\pi f_i(x) = \partial^\circ f_i(x)$ for each $i \in I(x)$. Then

$$\partial^\pi f(x) = co\{\partial^\pi f_i(x) : i \in I(x)\} = \partial^\circ f(x),$$

where $co\, A$ is the convex hull of a set A.

Proof. Since $\partial^\pi f_i(x) = \partial^\circ f_i(x)$ for each $i \in I(x)$, by [3, Proposition 2.3.12], we have

$$\partial^\pi f(x) \subseteq \partial^\circ f(x) \subseteq co\{\partial^\circ f_i(x) : i \in I(x)\} = co\{\partial^\pi f_i(x) : i \in I(x)\}.$$

So it suffices to show that $co\{\partial^\pi f_i(x) : i \in I(x)\} \subseteq \partial^\pi f(x)$.

For any fixed $i \in I(x)$ and $\xi_i \in \partial^\pi f_i(x)$, there exist $M > 0$ and $\delta > 0$ such that

$$f_i(y) - f_i(x) + M\|y - x\|^2 \ge \langle \xi_i, y - x\rangle \text{ for all } y \in B(x, \delta).$$

It follows that

$$f(y) - f(x) + M\|y - x\|^2 \geq \langle \xi_i, y - x\rangle \text{ for all } y \in B(x, \delta),$$

which implies that $\xi_i \in \partial^\pi f(x)$. Since i and ξ_i are arbitrary, $\partial^\pi f_i(x) \subseteq \partial^\pi f(x)$ for each $i \in I(x)$. In addition, $\partial^\pi f(x)$ is convex, so for any $\lambda_i \geq 0$ with $\sum_{i\in I(x)} \lambda_i = 1$,

$$\sum_{i\in I(x)} \lambda_i \xi_i \in \partial^\pi f(x).$$

This is what we need to prove. □

THEOREM 3.6. *Let X be a Hilbert space, and let $f_i : X \to R$ be an l.s.c. function for each $i \in I$. Denote*

$$f(x) = \max\{f_i(x) : i \in I\} \quad \textit{and} \quad I(x) := \{i \in I : f_i(x) = f(x)\} \textit{ for } x \in X.$$

Suppose that, for some $x_0 \in X$, $0 < \delta \leq +\infty$, $0 < \mu$, and $0 < \epsilon \leq (2\mu)^{-1}(\delta/2)^2$, the set $B(x_0, \delta/2) \cap f^{-1}(-\infty, \epsilon)$ is nonempty and that, for each $x \in B(x_0, \delta) \cap f^{-1}(0, \epsilon)$ and each $i \in I(x)$,

(i) *f_i is Lipschitz near x and $\partial^\pi f_i(x) = \partial^\circ f_i(x)$; and*

(ii) *there exists a unit vector u_x such that $d_L^2 f_j(x|\xi_j)(u_x) \leq -2\mu^{-1}$ and*

$$\lim_{t\to 0^+} \frac{f_i(x + tu_x) - [f_j(x + tu_x) - t\langle \xi_j, u_x\rangle] - t\langle \xi_k, u_x\rangle}{t^2} = 0$$

for some $j \in I(x)$ and $\xi_j \in \partial^\pi f_j(x)$, each $i \in I(x)$ and $\xi_i \in \partial^\pi f_i(x)$, and each $k \in I(x)$ and $\xi_k \in \partial^\pi f_k(x)$.

Then $S := \{x \in X : f(x) \leq 0\}$ is nonempty and

$$d_S^2(x) \leq 2\mu f(x)_+ \textit{ for all } x \in B\left(x_0, \frac{\delta}{2}\right) \cap f^{-1}(0, \epsilon).$$

Moreover, if $x_0 \in S$, then the condition $0 < \epsilon \leq (2\mu)^{-1}(\delta/2)^2$ can be replaced with $0 < \epsilon \leq +\infty$.

Proof. Let $x_0 \in X$, $0 < \delta \leq +\infty$, $0 < \mu$ and let the set $B(x_0, \delta/2) \cap f^{-1}(-\infty, \epsilon)$ be nonempty for some $0 < \epsilon < (2\mu)^{-1}(\delta/2)^2$. If, for $x \in B(x_0, \delta) \cap f^{-1}(0, \epsilon)$, $\partial^\pi f_i(x) = \partial^\circ f_i(x)$ for each $i \in I(x)$, then, for $\xi \in \partial^\pi f(x)$, by Proposition 3.5, $\xi = \sum_{i\in I(x)} \lambda_i \xi_i$ for some $\lambda_i \geq 0$ and $\xi_i \in \partial^\pi f_i(x)$ with $i \in I(x)$ and $\sum_{i\in I(x)} \lambda_i = 1$.

If u_x is the unit vector stated in the assumption, then

$$\begin{aligned}
d_L^2 f(x|\xi)(u_x) &= \liminf_{t\to 0^+} \frac{\max\{f_i(x + tu_x) : i \in I(x)\} - f(x) - t\langle \xi, u_x\rangle}{t^2} \\
&= \liminf_{t\to 0^+} \frac{\max\{f_i(x + tu_x) : i \in I(x)\} - f(x) - t\sum_{i\in I(x)} \lambda_i\langle \xi_i, u_x\rangle}{t^2} \\
&= \liminf_{t\to 0^+} \sum_{i\in I(x)} \lambda_i \frac{\max\{f_i(x + tu_x) : i \in I(x)\} - f(x) - t\langle \xi_i, u_x\rangle}{t^2} \\
&\leq \liminf_{t\to 0^+} \frac{f_j(x + tu_x) - f_j(x) - t\langle \xi_j, u_x\rangle}{t^2} \\
&+ \lim_{t\to 0^+} \sum_{i\in I(x)} \frac{|f_i(x + tu_x) - [f_j(x + tu_x) - t\langle \xi_j, u_x\rangle] - t\langle \xi_k, u_x\rangle|}{t^2} \\
&= d_L^2 f_j(x|\xi_j)(u_x) \leq -2\mu^{-1}.
\end{aligned}$$

Thus the conclusion follows from Theorem 3.1. ∎

Remark 3.4. From the proof of Theorem 3.6 we see that condition (i) can be replaced with the condition that f_i be continuous at x and $\partial^\pi f(x) = co\{\partial^\pi f_i(x) : i \in I(x)\}$.

THEOREM 3.7. *Let X be a Hilbert space and, for each $i \in I$, let $f_i : X \to R$ be C^1 and satisfy $\partial^\pi f_i(x) = \partial^\circ f_i(x)$ for $x \in X$. Denote*

$$f(x) = \max\{f_i(x) : i \in I\} \quad \textit{and} \quad I(x) := \{i \in I : f_i(x) = f(x)\} \textit{ for } x \in X.$$

Suppose that, for some $x_0 \in X$, $0 < \delta \le +\infty$, $0 < \mu$, and $0 < \epsilon \le (2\mu)^{-1}(\delta/2)^2$, the set $B(x_0, \delta/2) \cap f^{-1}(-\infty, \epsilon)$ is nonempty and that, for each $x \in B(x_0, \delta) \cap f^{-1}(0, \epsilon)$, there exists a unit vector u_x such that

$$\langle \nabla f_i(x), u_x\rangle = \langle \nabla f_j(x), u_x\rangle \quad \textit{and} \quad d^2_+ f_i(x|\nabla f_i(x))(u_x) \le -2\mu^{-1} \textit{ for all } i, j \in I(x).$$

Then $S := \{x \in X : f(x) \le 0\}$ is nonempty and

$$d^2_S(x) \le 2\mu f(x)_+ \textit{ for all } x \in B\left(x_0, \frac{\delta}{2}\right) \cap f^{-1}(0, \epsilon).$$

Moreover, if $x_0 \in S$, then the condition $0 < \epsilon \le (2\mu)^{-1}(\delta/2)^2$ can be replaced with $0 < \epsilon \le +\infty$.

Proof. Let x_0, δ, μ, and ϵ be given as in the assumption. For each $x \in B(x_0, \delta) \cap f^{-1}(0, \epsilon)$ and $i \in I(x)$, f_i is C^1 and $\partial^\pi f_i(x) = \{\nabla f_i(x)\} = \partial^\circ f_i(x)$, so for $\xi \in \partial^\pi f(x)$, by Proposition 3.5, $\xi = \sum_{i\in I(x)} \lambda_i \nabla f_i(x)$ for some $\lambda_i \ge 0$ with $i \in I(x)$ and $\sum_{i\in I(x)} \lambda_i = 1$.

If u_x is the vector in the assumption, then there exists $t_n \to 0$ such that

$$\begin{aligned}
d^2_L f(x|\xi)(u_x) &= \lim_{n\to+\infty} \frac{\max\{f_i(x + t_n u_x) : i \in I(x)\} - f(x) - t_n\langle \xi, u_x\rangle}{t_n^2} \\
&= \lim_{n\to+\infty} \frac{\max\{f_i(x + t_n u_x) : i \in I(x)\} - f(x) - t_n \sum_{i\in I(x)} \lambda_i\langle \nabla f_i(x), u_x\rangle}{t_n^2} \\
&= \lim_{n\to+\infty} \sum_{i\in I(x)} \lambda_i \frac{\max\{f_i(x + t_n u_x) : i \in I(x)\} - f(x) - t_n\langle \nabla f_i(x), u_x\rangle}{t_n^2} \\
&\le \limsup_{n\to+\infty} \max\left\{\frac{f_i(x + t_n u_x) - f_i(x) - t_n\langle \nabla f_i(x), u_x\rangle}{t_n^2} : i \in I(x)\right\} \\
&\le \max\{d^2_+ f_i(x|\nabla f_i(x))(u_x) : i \in I(x)\} \le -2\mu^{-1};
\end{aligned}$$

that is, we have $d^2_L f(x|\xi)(u_x) \le -2\mu^{-1}$. Therefore, upon using Theorem 3.1 to f, the conclusion follows. ∎

We now consider a system of quadratic inequalities

$$S = \{x \in R^n : f_1(x) \le 0, \dots, f_m(x) \le 0\},$$

where $f_i(x) = x^t Q_i x + b_i^t x + c_i$, Q_i is a real $n \times n$ symmetric matrix, $b_i \in R^n$, and $c_i \in R$ for each $i \in I$ with x^t denoting the transpose of x. For the convex quadratic system, i.e., when each Q_i is positive semidefinite, Luo and Luo [12] and Wang and Pang [19] show that the nonemptiness of S implies the existence of a positive integer $d \le n + 1$ and a positive scalar μ such that

$$d_S(x) \le \mu\left[f(x)_+ + f(x)_+^{\frac{1}{2^d}}\right] \text{ for all } x \in R^n, \tag{15}$$

where $f(x) = \max\{f_i(x) : i \in I\}$. Furthermore, if S contains an interior point, then $d = 0$.

For a nonconvex quadratic system, there are very few existing error bound results. For the special case of a single quadratic function, Luo and Sturm [13] show that (15) holds with d equal to 1; Ng and Zheng [15] further prove that for a single quadratic function, global error bounds with either exponents 1 or 1/2 hold, and they also classify the cases for exponents being 1 or 1/2. In the following theorem we apply Theorem 3.7 to derive a sufficient condition for a nonconvex quadratic system. It is worth pointing out that even for the case of a single quadratic system our theorem offers something new since an error bound is explicitly given in terms of the eigenvalues of matrices.

COROLLARY 3.8. *For each $i \in I$, let*

$$f_i(x) = x^t Q_i x + b_i^t x + c_i \quad \text{for } x \in R^n,$$

where Q_i is a real $n \times n$ symmetric matrix, $b_i \in R^n$, and $c_i \in R$. Denote

$$f(x) = \max\{f_i(x) : i \in I\} \quad \text{and} \quad I(x) := \{i \in I : f_i(x) = f(x)\} \text{ for } x \in R^n.$$

Suppose that for each $x \in f^{-1}(0, +\infty)$ and for each Q_i there exists a negative eigenvalue λ_i with a common eigenvector u and $\langle 2Q_i x + b_i, u\rangle = \langle 2Q_j x + b_j, u\rangle$ for all $i, j \in I(x)$. Then $S := \{x \in R^n : f(x) \le 0\}$ is nonempty and

$$d_S^2(x) \le -\frac{4}{\lambda} f(x)_+ \text{ for all } x \in R^n,$$

where $\lambda = \max\{\lambda_i : i \in I(x)\}$. In particular, if $I = \{1\}$ and λ_1 and λ_2 are the smallest eigenvalue and the largest eigenvalue of Q_1 with $\lambda_1 < 0 < \lambda_2$, then $S := \{x \in R^n : f_1(x) = 0\}$ is nonempty and

$$d_S^2(x) \le -\frac{4}{\lambda} |f_1(x)| \text{ for all } x \in R^n,$$

where $\lambda = \max\{\lambda_1, -\lambda_2\}$.

Proof. Let u be a common eigenvector of Q_i corresponding to an eigenvalue $\lambda_i < 0$ for all $i \in I(x)$. Then we have

$$f_i(\alpha u) = \lambda_i \alpha^2 u^t Q_i u + \alpha b_i^t u + c_i < 0$$

for sufficiently large positive scalar α. This implies that $S := \{x \in R^n : f(x) \le 0\}$ is nonempty.

Denote $u_x := \frac{u}{\|u\|}$. Then

$$\langle \nabla f_i(x), u_x\rangle = \langle 2Q_i x + b_i, u_x\rangle = \langle 2Q_j x + b_j, u_x\rangle = \langle \nabla f_j(x), u_x\rangle \text{ and}$$

$$d_+^2 f_i(x|\nabla f_i(x))(u_x) = u_x^t Q_i u_x = \lambda_i \le \lambda.$$

Thus, by Theorem 3.7,

$$d_S^2(x) \le -\frac{4}{\lambda} f(x)_+ \text{ for all } x \in R^n.$$

Now if $I = \{1\}$, we consider

$$f(x) := \max\{f_1(x), -f_1(x)\} = \begin{cases} f_1(x) & \text{if } f_1(x) \ge 0, \\ -f_1(x) & \text{if } f_1(x) < 0. \end{cases}$$

It is easy to see that $S := \{x \in R^n : f_1(x) = 0\}$ is nonempty and that $\nabla f(x) = \nabla f_1(x)$ for $x \in f^{-1}(0, +\infty)$ and $\nabla f(x) = -\nabla f_1(x)$ for $x \in f^{-1}(-\infty, 0)$. If u_1 and u_2 are unit eigenvectors corresponding to λ_1 and λ_2, respectively, then, for each $x \in f_1^{-1}(0, +\infty)$,

$$d_+^2 f(x|\nabla f(x))(u_1) = u_1^t Q_1 u_1 = \lambda_1 \leq \lambda$$

and, for each $x \in f_1^{-1}(-\infty, 0)$,

$$d_+^2 f(x|\nabla f(x))(u_2) = -u_2^t Q_1 u_2 = -\lambda_2 \leq \lambda.$$

Therefore it follows from Theorem 3.7 that

$$d_S^2(x) \leq -\frac{4}{\lambda}|f_1(x)| \text{ for all } x \in R^n. \qquad \square$$

Example 3.1. For $x \in R^2$, define

$$f_1(x) = x^t Q_1 x + b_1^t x + 1 \quad \text{and} \quad f_2(x) = x^t Q_2 x + b_2^t x,$$

where

$$Q_1 = \begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix},\ b_1 = \begin{pmatrix} 3 \\ 0 \end{pmatrix},\ Q_2 = \begin{pmatrix} -1 & 0 \\ 0 & -4 \end{pmatrix},\ \text{and } b_2 = \begin{pmatrix} 3 \\ -1 \end{pmatrix}.$$

It is easy to see that $\lambda = -1$ is a common eigenvalue of Q_1 and Q_2 with a common eigenvector $u = (1, 0)^t$ and that $\langle 2Q_1x + b_1, u\rangle = -2x_1 + 3 = \langle 2Q_2x + b_2, u\rangle$. Therefore, by Corollary 3.8, $S := \{x \in R^2 : f_1(x) \leq 0,\ f_2(x) \leq 0\}$ is nonempty and

$$d_S^2(x) \leq 4\max\{f_1(x), f_2(x)\}_+ \text{ for all } x \in R^2.$$

4. Sufficient conditions in lower Dini derivatives. We note that in a general Banach space the lower Dini subdifferential is not always a ∂_ω-subdifferential (see [10]). Thus Theorem 2.2 is not applicable to the lower Dini subdifferential in a general Banach space. However, in this case the lower Dini derivative $f^-(x;\cdot)$ of function f at x turns out to be more convenient for us to present a sufficient condition for error bounds to exist. For this we first prove one of the main results in this section.

THEOREM 4.1. *Let (X, d) be a metric space and let $f : X \to (-\infty, +\infty]$ be an l.s.c. function. For some $0 < \epsilon \leq +\infty$ and $0 < \mu < +\infty$ we consider the following statements:*

(i) *If the set $f^{-1}(-\infty, \epsilon)$ is nonempty and for each $x \in f^{-1}(0, \epsilon)$ there exists a point $y \in f^{-1}[0, \epsilon)$ such that*

$$0 < d(x, y) \leq \mu[f(x) - f(y)],$$

then $S := \{x \in X : f(x) \leq 0\}$ is nonempty and

$$d_S(x) \leq \mu f(x)_+ \text{ for all } x \in f^{-1}(-\infty, \epsilon).$$

(ii) *If for some $x_0 \in X$ and $0 < \delta < +\infty$ the set $B(x_0, \delta) \cap f^{-1}(-\infty, \epsilon)$ is nonempty and for some $0 < \rho < 1$ and each $x \in B(x_0, \delta) \cap f^{-1}(0, \epsilon)$ there exists a point $y \in f^{-1}[0, \epsilon)$ such that*

$$d(y, x_0) \leq \max\{\rho\delta, d(x, x_0)\} \quad \text{and} \quad 0 < d(x, y) \leq \mu[f(x) - f(y)],$$

then $S := \{x \in X : f(x) \leq 0\}$ is nonempty and

$$d_S(x) \leq \mu f(x)_+ \text{ for all } x \in B(x_0, \delta) \cap f^{-1}(-\infty, \epsilon).$$

(iii) *If for some nonempty closed subset C of X the set $C \cap f^{-1}(-\infty, \epsilon)$ is nonempty and for some $0 < \mu < +\infty$ and each $x \in C \cap f^{-1}(0, \epsilon)$ there exists a sequence $\{x_n\} \subseteq C \backslash \{x\}$ such that*

$$\lim_{n \to +\infty} \frac{f(x_n)_+ - f(x)_+}{\|x_n - x\|} \leq -\mu^{-1}, \tag{16}$$

then $S := \{x \in C : f(x) \leq 0\}$ is nonempty and

$$d_S(x) \leq \mu f(x)_+ \text{ for all } x \in C \cap f^{-1}(-\infty, \epsilon).$$

In a metric space X, (i) $\Rightarrow$ (ii)*; if X is also complete, then both* (i) *and* (ii) *hold. In a normed space X,* (i) $\Leftrightarrow$ (iii)*; hence* (iii) *holds in a Banach space X.*

Proof. We first prove that (i) implies (ii) in a metric space X.

Let $x_0 \in X$, $0 < \delta < +\infty$, and $0 < \rho < 1$. For each $m \in N$ (the set of natural numbers) such that $B(x_0, \rho\delta) \subseteq \overline{B}_m(x_0, \delta) := \overline{B}(x_0, (1 - 1/m)\delta)$ and each $x \in \overline{B}_m(x_0, \delta) \cap f^{-1}(0, \epsilon)$ there exists y with the properties stated in (ii) such that $y \in \overline{B}_m(x_0, \delta)$ since

$$d(y, x_0) \leq \max\{\rho\delta, d(x, x_0)\} \leq \left(1 - \frac{1}{m}\right)\delta.$$

Upon applying (i) to the function $f + \psi_{\overline{B}_m(x_0,\delta)}$, we obtain that $S_m := \overline{B}_m(x_0, \delta) \cap f^{-1}(-\infty, 0]$ is nonempty and

$$d_{S_m}(x) \leq \mu f(x)_+ \text{ for all } x \in \overline{B}_m(x_0, \delta) \cap f^{-1}(-\infty, \epsilon).$$

This implies that (ii) holds since for each $x \in B(x_0, \delta) \cap f^{-1}(-\infty, \epsilon)$ there exists an m stated above such that $x \in \overline{B}_m(x_0, \delta) \cap f^{-1}(-\infty, \epsilon)$ and $d_S(x) \leq d_{S_m}(x)$.

Now it is known from [22, Theorem 3] that (i) holds in a complete metric space, so (ii) also holds in a complete metric space.

Next, we prove that (i) and (iii) are equivalent in a normed space X.

Suppose that (i) is true. To prove (iii) to be also true, it suffices to show that for any $\lambda > 1$ and $x \in C \cap f^{-1}(0, \epsilon)$ there exists a point $y \in C \cap (f_+)^{-1}[0, \epsilon)$ such that

$$0 < \|x - y\| \leq \lambda\mu[f(x)_+ - f(y)_+].$$

Let $\lambda > 1$ be fixed. For each $x \in C \cap f^{-1}(0, \epsilon)$, by assumption, there exists a sequence $\{x_n\} \subseteq C \backslash \{x\}$ satisfying (16). Hence for sufficiently large n we have

$$\frac{f(x_n)_+ - f(x)_+}{\|x_n - x\|} \leq -(\lambda\mu)^{-1},$$

that is,

$$0 < \|x_n - x\| \leq \lambda\mu[f(x)_+ - f(x_n)_+].$$

So we can take $y = x_n$ for any such an n.

Now, to prove (iii) $\Rightarrow$ (i), we suppose that for each $x \in f^{-1}(0, \epsilon)$ there exists a point $y \in f^{-1}[0, \epsilon)$ such that

$$0 < \|x - y\| \leq \mu[f(x) - f(y)].$$

By taking $x_n = y$ we have

$$\lim_{n\to+\infty} \frac{f(x_n)_+ - f(x)_+}{\|x_n - x\|} = \lim_{n\to+\infty} \frac{f(y) - f(x)}{\|y - x\|} \le -\mu^{-1}.$$

It follows from statement (iii) with $C = X$ that S is nonempty and satisfies

$$d_S(x) \le \mu f(x)_+ \text{ for all } x \in f^{-1}(-\infty, \epsilon).$$

Therefore (i) is valid.

As we indicated above, (i) holds in a complete metric space, so (iii) holds in a Banach space. $\square$

Based on Theorem 4.1, we present some sufficient conditions in terms of Dini derivatives of involved functions and tangent cones to a set as below.

THEOREM 4.2. *Let X be a Banach space and let C be a nonempty closed subset in X. Suppose that $f : X \to (-\infty, +\infty]$ is an l.s.c. function and that for some $0 < \epsilon \le +\infty$ the set $C \cap f^{-1}(-\infty, \epsilon)$ is nonempty. If for some $0 < \mu$ and each $x \in C \cap f^{-1}(0, \epsilon)$ there exists a unit hypertangent vector u_x to C at x such that $f^-(x; u_x) \le -\mu^{-1}$ or $u_x \in K_C(x)$ such that $f^+(x; u_x) \le -\mu^{-1}$, then $S := \{x \in C : f(x) \le 0\}$ is nonempty and*

$$d_S(x) \le \mu f(x)_+ \text{ for all } x \in C \cap f^{-1}(-\infty, \epsilon).$$

Proof. For some $0 < \epsilon \le +\infty$, let $x \in C \cap f^{-1}(0, \epsilon)$. If u_x is a unit hypertangent vector to C at x satisfying $f^-(x; u_x) \le -\mu^{-1}$, then there exist sequences $u_n \to u_x$ and $t_n \to 0^+$ such that

$$\lim_{n\to+\infty} \frac{f(x + t_n u_n) - f(x)}{t_n} = f^-(x; u_x) \le -\mu^{-1}$$

and $x + t_n u_n \in C$. If $u_x \in K_C(x)$ and $f^+(x; u_x) \le -\mu^{-1}$, then there exist sequences $u_n \to u_x$ and $t_n \to 0^+$ such that $x + t_n u_n \in C$, for which we have

$$\liminf_{n\to+\infty} \frac{f(x + t_n u_n) - f(x)}{t_n} \le f^+(x; u_x) \le -\mu^{-1}.$$

Now for the above sequences $u_n \to u_x$ and $t_n \to 0^+$ we have $x_n := x + t_n u_n \in C\backslash\{x\}$ and

$$\liminf_{n\to+\infty} \frac{f(x_n)_+ - f(x)_+}{\|x_n - x\|} = \liminf_{n\to+\infty} \frac{f(x_n) - f(x)}{\|x_n - x\|} \le -\mu^{-1}.$$

Hence there exists a subsequence $\{x_{n_k}\}$ satisfying the condition (iii) in Theorem 4.1. Therefore the conclusion holds. $\square$

Similar to Theorem 3.4, Theorem 4.2 has the following equivalent result.

THEOREM 4.3. *Let X be a Banach space, and let $f : X \to (-\infty, +\infty]$ be an l.s.c. function. Suppose that for some $0 < \epsilon \le +\infty$ the set $f^{-1}(-\infty, \epsilon)$ is nonempty and that for some $0 < \mu$ and each $x \in f^{-1}(0, \epsilon)$ there exists a unit vector u_x in X such that $f^-(x; u_x) \le -\mu^{-1}$. Then $S := \{x \in X : f(x) \le 0\}$ is nonempty and*

$$d_S(x) \le \mu f(x)_+ \text{ for all } x \in f^{-1}(-\infty, \epsilon).$$

In what follows we use Theorem 4.2 to establish error bounds for a system containing functions f and $g_i : X \to (-\infty, +\infty]$ $(i \in I)$ for which we denote

$$g(x) := \max\{g_i(x) : i \in I\} \quad \text{and} \quad I(x) := \{i \in I : g_i(x) = g(x)\} \text{ for } x \in X.$$

THEOREM 4.4. *Let C be a nonempty closed subset in a Banach space X, let $f : X \to (-\infty, +\infty]$ be l.s.c., and let $g_i : X \to (-\infty, +\infty)$ be locally Lipschitz for each $i \in I$. Denote*

$$C_0 := \{x \in C : g_i(x) \leq 0 \text{ for each } i \in I\}.$$

Suppose that for some $0 < \epsilon \leq +\infty$ the set $C_0 \cap f^{-1}(-\infty, \epsilon)$ is nonempty. If, for some $0 < \mu$ and each $x \in C_0 \cap f^{-1}(0, \epsilon)$, there exists a unit vector $u_x \in K_C(x)$ such that $f^+(x; u_x) \leq -\mu^{-1}$ and, for each $x \in C_0 \cap f^{-1}(0, \epsilon)$ with $g(x) = 0$ and each $i \in I(x)$, $g_i^+(x; u_x) < 0$, then $S := \{x \in C_0 : f(x) \leq 0\}$ is nonempty and

$$d_S(x) \leq \mu f(x)_+ \text{ for all } x \in C_0 \cap f^{-1}(-\infty, \epsilon).$$

Proof. Let $x \in C_0 \cap f^{-1}(0, \epsilon)$ and let $u_x \in K_C(x)$ be the unit vector in the assumption. Then there exist sequences $u_n \to u_x$ and $t_n \to 0^+$ such that $x + t_n u_n \in C$. According to Theorem 4.2 , it suffices to show that $u_x \in K_{C_0}(x)$.

If $g(x) < 0$, then, by the continuity of g, $g(x + t_n u_n) \leq 0$ when n is large enough. This implies that $x + t_n u_n \in C_0$ when n is large enough and hence $u_x \in K_{C_0}(x)$.

If $g(x) = 0$, then, by the definition of $g_i^+(x; u_x)$, there are $\delta > 0$ and $\epsilon > 0$ such that for each $i \in I(x)$ and all $t \in (0, \delta)$ we have

$$g_i(x + tu_x) - g_i(x) \leq -\epsilon t.$$

Since g_i is Lipschitz near x, there exists a constant L_i such that

$$\begin{aligned} g_i(x + t_n u_n) &\leq g_i(x + t_n u_x) + L_i t_n \|u_n - u_x\| \\ &\leq g_i(x) + t_n(-\epsilon + L_i \|u_n - u_x\|) \leq 0 \end{aligned}$$

for sufficiently large n. It follows that $x + t_n u_n \in C_0$ when n is large enough. Thus u_x belongs to $K_{C_0}(x)$. $\square$

PROPOSITION 4.5. *Let x be a point in a closed subset C of a Banach space X, let $f_i : X \to (-\infty, +\infty)$ be Lipschitz near x, let $g_i : X \to (-\infty, +\infty]$ be Fréchet differentiable at x for each $i \in I$, and let $h_j : X \to (-\infty, +\infty)$ be continuous in a neighborhood of x and Fréchet differentiable at x for each $j \in J$ with the Fréchet derivative $\nabla h(x) = (\nabla h_1(x), \ldots, \nabla h_n(x))^t$ being surjective. Denote*

$$C_1 := \{x \in C : (f_i + g_i)(x) \leq 0 \text{ for } i \in I \text{ and } h_j(x) = 0 \text{ for } j \in J\} \text{ and}$$

$$I(x) := \{i \in I : (f_i + g_i)(x) = 0\}.$$

Suppose that $x \in C_1$ and there exists $v^* \in X$ such that $f_i^\circ(x; v^*) + g_i'(x; v^*) < 0$ for each $i \in I(x)$ and $h_j'(x; v^*) = 0$ for each $j \in J$. If the set of hypertangents to the set C at x is nonempty, then

$$\{v \in \operatorname{int} T_C(x) : f_i^+(x; v) + g_i'(x; v) \leq 0,\ i \in I(x);\ h_j'(x; v) = 0,\ j \in J\} \subseteq K_{C_1}(x).$$

Proof. First, for $v \in \operatorname{int} T_C(x)$ satisfying $f_i^+(x; v) + g_i'(x; v) < 0$ for each $i \in I(x)$ and $h_j'(x; v) = 0$ for each $j \in J$, we prove that $v \in K_{C_1}(x)$.

Since $\nabla h(x)$ is surjective, by the correction theorem of Halkin [6, Theorem F] and its proof, there exist a neighborhood U of x and a continuous mapping ξ from U into X such that $\xi(x) = 0$, $\nabla \xi(x) = 0$, and

$$h_j(y + \xi(y)) = \langle \nabla h_j(x), y - x \rangle \text{ for all } y \in U \text{ and each } j \in J.$$

By taking $y = x + sv$ we have, for $t > 0$ small enough and all $s \in (0, t)$,

$$h_j(x + sv + \xi(x + sv)) = \langle \nabla h_j(x), sv \rangle = 0 \text{ for each } j \in J.$$

Note that $\xi(x) = 0$ and $\nabla\xi(x) = 0$, so $\xi(x + tv)/t \to 0$ as $t \to 0$. By the inequality $f_i^+(x; v) + g_i'(x; v) < 0$, we can take $\epsilon > 0$ and $t > 0$ small enough such that

$$(f_i + g_i)(x + sv + \xi(x + sv)) \le (f_i + g_i)(x) - \epsilon s = -\epsilon s$$

for all $s \in (0, t)$ and each $i \in I(x)$. Also, if $v \in \operatorname{int} T_C(x)$, then v is hypertangent to C at x. Hence

$$x + sv + \xi(x + sv) = x + s\left[v + \frac{\xi(x + sv)}{s}\right] \in C \text{ for all } s \in (0, t)$$

when $t > 0$ is small enough. This implies that $v \in K_{C_1}(x)$.

Now, if $v^* \in X$ satisfies $f_i^\circ(x; v^*) + g_i'(x; v^*) < 0$ for each $i \in I(x)$ and $h_j'(x; v^*) = 0$ for each $j \in J$, then, for $v \in \operatorname{int} T_C(x)$ with $f_i^+(x; v) + g_i'(x; v) \le 0$ for each $i \in I(x)$ and $h_j'(x; v) = 0$ for each $j \in J$, we can take $t > 0$ small enough such that, for all $s \in (0, t)$, $v + sv^* \in \operatorname{int} T_C(x)$ and

$$\begin{aligned} & f_i^+(x; v + sv^*) + g_i'(x; v + sv^*) \\ & \qquad \le f_i^+(x; v) + g_i'(x; v + sv^*) + \sup_{u \in X}[f_i^+(x; u + sv^*) - f_i^+(x; u)] \\ & \qquad \le f_i^+(x; v) + g_i'(x; v) + s[f_i^\circ(x; v^*) + g_i'(x; v^*)] < 0 \end{aligned}$$

for each $i \in I(x)$ and

$$h_j'(x; v + sv^*) = \langle \nabla h_j(x), v + sv^* \rangle = 0$$

for each $j \in J$. By the conclusion of the above paragraph, we have $v + sv^* \in K_{C_1}(x)$ for all $s > 0$ small enough. This implies that $v \in K_{C_1}(x)$ since $K_{C_1}(x)$ is closed. □

Combining Theorem 4.2 with Proposition 4.5, we obtain the following result.

THEOREM 4.6. *Let C be a nonempty closed subset in a Banach space X, let $f : X \to (-\infty, +\infty]$ be l.s.c., let $f_i : X \to (-\infty, +\infty)$ be locally Lipschitz and $g_i : X \to (-\infty, +\infty]$ Fréchet differentiable on C for each $i \in I$, and let $h_j : X \to (-\infty, +\infty]$ be continuous on C for each $j \in J$. Denote*

$$C_1 := \{x \in C : (f_i + g_i)(x) \le 0 \textit{ for } i \in I \textit{ and } h_j(x) = 0 \textit{ for } j \in J\} \textit{ and}$$

$$I(x) := \{i \in I : (f_i + g_i)(x) = 0\} \textit{ for } x \in C_1.$$

Suppose that for some $0 < \epsilon \le +\infty$ the set $C_1 \cap f^{-1}(-\infty, \epsilon)$ is nonempty, that, for each $x \in C_1 \cap f^{-1}(0, \epsilon)$, h_j is Fréchet differentiable at x for each $j \in J$ with the Fréchet derivative $\nabla h(x) = (\nabla h_1(x), \ldots, \nabla h_n(x))^t$ being surjective and there exists $v_x^* \in X$ such that $f_i^\circ(x; v_x^*) + g_i(x; v_x^*) < 0$ for each $i \in I(x)$ and $h_j'(x; v_x^*) = 0$ for each $j \in J$, and that there exists a unit hypertangent vector u_x to the set C at x such that $f_i^+(x; u_x) + g_i'(x; u_x) \le 0$ for each $i \in I(x)$, $h_j'(x; u_x) = 0$ for each $j \in J$ and $f^+(x; u_x) \le -\mu^{-1}$ for some $0 < \mu$ independent of x. Then $S := \{x \in C_1 : f(x) \le 0\}$ is nonempty and

$$d_S(x) \le \mu f(x)_+ \textit{ for all } x \in C_1 \cap f^{-1}(-\infty, \epsilon).$$

In what follows, we consider an inequality system determined by several inequalities.

THEOREM 4.7. *Let C be a nonempty closed subset in a Banach space X and let $f_i : X \to R$ be continuous for each $i \in I$. Denote*

$$f(x) = \max\{f_i(x) : i \in I\} \quad \textit{and} \quad I(x) := \{i \in I : f_i(x) = f(x)\} \textit{ for } x \in X.$$

Suppose that, for some $0 < \epsilon \le +\infty$, the set $C \cap f^{-1}(-\infty, \epsilon)$ is nonempty and that, for some $0 < \mu$, each $x \in f^{-1}(0, \epsilon)$, and $i \in I(x)$, there exists a unit vector u_x such that

(i) *u_x is hypertangent to C at x, $f_j^-(x; u_x) \le -\mu^{-1}$ for some $j \in I(x)$ and*

$$\lim_{\substack{v \to u_x \\ t \to 0^+}} \frac{f_i(x + tv) - f_j(x + tv)}{t} = 0 \textit{ for each } i \in I(x); \textit{ or}$$

(ii) *$u_x \in K_C(x)$ and $f_i^+(x; u_x) \le -\mu^{-1}$ for each $i \in I(x)$.*

Then $S := \{x \in C : f(x) \le 0\}$ is nonempty and

$$d_S(x) \le \mu f(x)_+ \textit{ for all } x \in C \cap f^{-1}(0, \epsilon).$$

Proof. Let $0 < \mu$, $0 < \epsilon \le +\infty$, and $x \in C \cap f^{-1}(0, \epsilon)$. If u_x is a unit vector satisfying (i), then

$$\begin{aligned} f^-(x; u_x) &= \liminf_{\substack{v \to u_x \\ t \to 0^+}} \frac{\max\{f_i(x + tv) : i \in I(x)\} - f(x)}{t} \\ &\le \liminf_{\substack{v \to u_x \\ t \to 0^+}} \frac{f_j(x + tv) - f_j(x)}{t} + \lim_{\substack{v \to u_x \\ t \to 0^+}} \sum_{i \in I(x)} \frac{|f_i(x + tv) - f_j(x + tv)|}{t} \\ &= f_j^-(x; u_x) \le -\mu^{-1}, \end{aligned}$$

where the first equality is obtained by the continuity of f_i at x for each $i \in I$.

If u_x satisfies (ii), then there exist $u_n \to u_x$ and $t_n \to 0^+$ such that

$$\begin{aligned} f^+(x; u_x) &= \lim_{n \to +\infty} \frac{f(x + t_n u_n) - f(x)}{t_n} \\ &= \lim_{n \to +\infty} \max\left\{\frac{f_i(x + t_n u_n) - f_i(x)}{t_n} : i \in I(x)\right\} \\ &\le \max\{f_i^+(x; u_x) : i \in I(x)\} \le -\mu^{-1}. \end{aligned}$$

Therefore, from Theorem 4.2, the required result follows. ∎

COROLLARY 4.8. *For each $i \in I$, let $g_i : R^n \to R$ be differentiable and let $f_i(x) := g_i(x) + b_i^t x + c_i$, where $b_i = (b_{i1}, \ldots, b_{in})^t \in R^n$ and $c_i \in R$. Suppose that for all $i \in I$ and some $j \in J$ the coordinates b_{ij} have the same sign and all g_i's are independent of the jth coordinate x_j of $x \in R^n$. Then*

$$S := \{x \in R^n : f_i(x) \le 0 \textit{ for all each } i \in I\}$$

is nonempty and for some $0 < \mu$ there holds $d_S(x) \le \mu f(x)_+$ for all $x \in R^n$.

Proof. In fact, for $x \in f^{-1}(0, +\infty)$ and $i \in I(x)$ and $u_x := (0, \ldots, 0, -sgn\, b_{ij}, 0, \ldots, 0)^t$ we have

$$f_i'(x; u_x) = \langle \nabla f_i(x), u_x \rangle = \langle \nabla g_i(x) + b_i, u_x \rangle = -|b_{ij}| \text{ for each } i \in I(x).$$

Taking $\mu^{-1} = \min\{|b_{ij}| : i \in I\}$ and applying Theorem 4.7, we arrive at the conclusion. $\square$

Example 4.1. We consider the functions

$$\begin{aligned} f_1(x) &= g_1(x_1, x_2) + 2x_1 - x_2 + 3x_3, \\ f_2(x) &= g_2(x_1, x_2) + 2x_3, \\ f_3(x) &= g_3(x_1, x_2) + 2x_1 + 6x_3 - 4, \end{aligned}$$

where g_1, g_2, and g_3 are differentiable and independent of x_3. Since the coefficients of x_3 in f_i's are all positive and their minimum is 2,

$$S := \{x \in R^3 : f_i(x) \le 0 \text{ for } i = 1, 2, 3\}$$

is nonempty and $d_S(x) \le \frac{1}{2} f(x)_+$ holds for all $x \in R^3$.

Acknowledgment. We thank two anonymous referees for their valuable comments and suggestions.

REFERENCES

[1] J. F. Bonnans and A. D. Ioffe, *Second-order sufficiency and quadratic growth for non isolated minima*, Math. Oper. Res., 20 (1995), pp. 801–817.

[2] J. F. Bonnans and A. D. Ioffe, *Quadratic growth and stability in convex programming problems with multiple solutions*, J. Convex Anal., 2 (1995), pp. 41–57.

[3] F. H. Clarke, *Optimization and Nonsmooth Analysis*, Wiley-Interscience, New York, 1983; reprinted as Classics Appl. Math. 5, SIAM, Philadelphia, PA, 1990.

[4] F. H. Clarke, Yu. S. Ledyaev, R. J. Stern, and P. R. Wolenski, *Nonsmooth Analysis and Control Theory*, Grad. Texts in Math. 178, Springer-Verlag, New York, 1998.

[5] I. Ekeland, *On the variational principle,* J. Math. Anal. Appl., 47 (1974), pp. 324–353.

[6] H. Halkin, *Implicit functions and optimization problems without continuous differentiability of the data,* SIAM J. Control, 12 (1974), pp. 229–236.

[7] L. R. Huang and K. F. Ng, *On first and second-order conditions for error bounds,* preprint.

[8] A. J. Hoffman, *On approximate solutions of systems of linear inequalities,* J. Research Nat. Bur. Standards, 49 (1952), pp. 263–265.

[9] A. D. Ioffe, *Regular points of Lipschitz functions,* Trans. Amer. Math. Soc., 251 (1979), pp. 61–69.

[10] A. D. Ioffe, *Calculus of Dini subdifferentials of functions and contingent coderivatives of set-valued maps,* Nonlinear Anal., 8 (1984), pp. 517–539.

[11] A. S. Lewis and J. S. Pang, *Error bounds for convex inequality systems,* in Generalized Convexity, Generalized Monotonicity, J.-P. Crouzeix, J.-E. Martinez-Legaz, and M. Volle, eds., Kluwer Academic, Dordrecht, The Netherlands, 1998, pp. 75–110.

[12] X.-D. Luo and Z.-Q. Luo, *Extension of Hoffman's error bound to polynomial systems*, SIAM J. Optim., 4 (1994), pp. 383–392.

[13] Z.-Q. Luo and J. F. Sturm, *Error bounds for quadratic systems,* in High Performance Optimization, Appl. Optim. 33, Kluwer Academic, Dordrecht, The Netherlands, 2000, pp. 383–404.

[14] J. S. Pang, *Error bounds in mathematical programming,* Math. Programming, 79 (1997), pp. 299–332.

[15] K. F. Ng and X. Y. Zheng, *Global error bounds with fractional exponents,* Math. Program. Ser. B, 88 (2000), pp. 357–370.

[16] K. F. Ng and X. Y. Zheng, *Error bounds for lower semicontinuous functions in normed spaces*, SIAM J. Optim., 12 (2001), pp. 1–17.

[17] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis,* Grundlehren Math. Wiss. 317, Springer-Verlag, Berlin, Heidelberg, 1998.

[18] W. Takahashi, *Existence theorems generalizing fixed point theorems for multivalued mappings,* in Fixed Point Theory and Applications, M. A. Théra and J. B. Baillon, Pitman Res. Notes Math. Ser. 252, Longman Scientific and Technical, Harlow, UK, 1991, pp. 397–406.

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

[20] D. Ward, *Sufficient conditions for weak sharp minima of order two and directional derivatives of the value function*, in Mathematical Programming with Data Perturbations, Lecture Notes in Pure and Appl. Math. 195, Dekker, New York, 1998, pp. 419–436.

[21] Z. Wu and J. J. Ye, *Sufficient conditions for error bounds,* SIAM J. Optim., 12 (2001), pp. 421–435.

[22] Z. Wu and J. J. Ye, *On error bounds for lower semicontinuous functions,* Math. Program. Ser. A, 92 (2002), pp. 301–314.