Math. Program., Ser. A (2011) 126:259–279
DOI 10.1007/s10107-009-0277-6

FULL LENGTH PAPER

Strong KKT conditions and weak sharp solutions in convex-composite optimization

Xi Yin Zheng · Kung Fu Ng

Received: 11 December 2007 / Accepted: 17 March 2009 / Published online: 17 April 2009

Abstract Using variational analysis techniques, we study convex-composite optimization problems. In connection with such a problem, we introduce several new notions as variances of the classical KKT conditions. These notions are shown to be closely related to the notions of sharp or weak sharp solutions. As applications, we extend some results on metric regularity of inequalities from the convex case to the convex-composite case.

Keywords Convex-composite optimization · Strong KKT condition · Sharp solution · Weak sharp solution

Mathematics Subject Classification (2000) 90C31 · 90C25 · 49J52 · 46B20

1 Introduction

Let X be a Banach space and $\phi : X \to \mathbb{R} \cup \{+\infty\}$ be a proper lower semicontinuous function. Recall that ϕ is said to have a sharp minimum at $\bar{x} \in X$ if there exist two

---

This research was supported by the National Natural Science Foundation of P. R. China (10761012) and an Earmarked Grant from the Research Grant Council of Hong Kong.

X. Y. Zheng (✉)
Department of Mathematics, Yunnan University,
650091 Kunming, People's Republic of China
e-mail: xyzheng@ynu.edu.cn

K. F. Ng
Department of Mathematics, The Chinese University of Hong Kong,
Shatin, New Territory, Hong Kong
e-mail: kfng@math.cuhk.edu.hk

positive numbers η and δ such that

$$\eta\|x-\bar{x}\| \leq \phi(x)-\phi(\bar{x}) \quad \forall x \in B(\bar{x}, \delta),$$

where $B(\bar{x}, \delta)$ denotes the open ball with center $\bar{x}$ and radius δ. Let Ω be a subset of X, $\lambda := \inf\{\phi(x) : x \in \Omega\}$ and $S := \{x \in \Omega : \phi(x) = \lambda\}$. Suppose that λ is finite. Following Burke and Ferris [5], we say

(a) S is a set of weak sharp minima for ϕ over Ω if there exists $\eta > 0$ such that

$$\eta d(x, S) \leq \phi(x)-\lambda \quad \forall x \in \Omega;$$

(b) $\bar{x} \in \Omega$ is a local weak sharp minimum for ϕ over Ω if there exist $\eta, \delta \in (0, +\infty)$ such that

$$\eta d(x, S(\bar{x})) \leq \phi(x)-\phi(\bar{x}) \quad \forall x \in \Omega \cap B(\bar{x}, \delta),$$

where $S(\bar{x}) := \{x \in \Omega : \phi(x) = \phi(\bar{x})\}$. The notions of sharp minima and weak sharp minima have many important consequences for convergence analysis and stability analysis of many algorithms. The readers can look at [3–5,9,25,28,29] and references therein for the history and motivation for the study of sharp minima and weak sharp minima. In terms of normal cones and subdifferentials, Burke and Ferris [5] established some valuable duality characterizations for weak sharp minima in finite dimensional spaces; Burke and Deng [3], with the help of the Fenchel dual technique, extended these results to a infinite dimensional space setting and established results on local weak sharp minima in a Hilbert space; the authors [28], using the Banach space geometrical technique, provided some characterizations for a local weak sharp minimum in a general Banach space. All the works mentioned above are under the convexity assumption. In this paper, we will relax the convexity assumption by considering the "convex-composite" situation, that is, the functions ϕ involved are given in the form $\phi = \psi \circ f$, where f is a smooth function from a Banach space X to another Banach space Y and ψ is a convex real-valued function on Y. Such a function (which is usually referred to as a convex-composite function) is not necessarily convex but shares many interesting and useful properties with convex functions. The class of such functions is huge (in particular it contains amenable functions due to Poliquin and Rockafellar (see [23, P. 442])) and these functions arise naturally in mathematical programming (see Rockafellar [21] where he gave many interesting examples showing that a wide spectrum of problems can be cast in terms of convex-composite functions).

Let $\phi_0, \ldots, \phi_m$ be proper lower semicontinuous functions on X and let C be a closed set in X. Consider the following constrained optimization problem

$$\begin{aligned} &\min \phi_0(x) \\ &\text{s.t. } \phi_i(x) \leq 0, \quad i = 1, \ldots, m, \\ &\qquad x \in C. \end{aligned} \tag{1.1}$$

Taking account of not only the objective function ϕ_0 *but also* the constraint functions $\phi_1, \ldots, \phi_m$ and the geometric constraint set C, we adopt the following penalty functions (of l^1 type):

$$p_\tau(x) := \phi_0(x) + \tau \sum_{i=1}^{m} [\phi_i(x)]_+ + \tau d(x, C) \quad \forall x \in X, \tag{1.2}$$

where $\tau > 0$ is constant and $[\phi_i(x)]_+ = \max\{\phi_i(x), 0\}$. It is natural and useful to consider relationship between solutions of (1.1) and unconstrained minimizers of the above penalty functions. Let Ω denote the feasible set of the above optimization problem, that is, $\Omega = \{x \in C : \ \phi_i(x) \le 0, \ i = 1, \ldots, m\}$. Suppose that ϕ_0 is Lipschitz and that the constraint system of (1.1) has an error bound in the following sense: there exists $\eta > 0$ such that

$$\eta d(x, \Omega) \le \sum_{i=1}^{m} [\phi_i(x)]_+ + d(x, C) \quad \forall x \in X;$$

it is known and easy to verify (cf. [6, Proposition 2.4.3]) that $\bar{x} \in \Omega$ is a solution of (1.1) if and only if it is an unconstrained minimizer of some penalty function p_τ. In this paper, we consider weak sharp solutions of (1.1) in terms of weak sharp minima of the penalty functions p_τ.

When each ϕ_i is smooth, under some constraint qualification (e.g. the Slater condition or the Mangasarian–Fromowitz condition), it is well known that if a feasible point $\bar{x}$ is a local solution of (1.1) then the Karush–Kuhn–Tucker (KKT in short) condition is satisfied at $\bar{x}$, namely there exist $\lambda_i \in \mathbb{R}_+$ such that

$$0 \in \phi_0'(\bar{x}) + \sum_{i \in I(\bar{x})} \lambda_i \phi_i'(\bar{x}) + N(C, \bar{x}),$$

that is,

$$0 \in \phi_0'(\bar{x}) + \sum_{i \in I(\bar{x})} \mathbb{R}_+ \phi_i'(\bar{x}) + N(C, \bar{x}), \tag{1.3}$$

where $I(\bar{x}) := \{1 \le i \le m : \ \phi_i(\bar{x}) = 0\}$. In the nonsmooth setting, one can take the Clarke–Rockafellar subdifferentials $\partial \phi_i(\bar{x})$ (for the definition and some basic properties, see Sect. 2) to replace the derivatives $\phi_i'(\bar{x})$, and the KKT condition can then be rewritten as

$$0 \in \partial \phi_0(\bar{x}) + \sum_{i \in I(\bar{x})} \mathbb{R}_+ \partial \phi_i(\bar{x}) + N(C, \bar{x}), \tag{1.4}$$

where $\mathbb{R}_+ \partial \phi_i(\bar{x}) = \bigcup_{r \ge 0} r \partial \phi_i(\bar{x})$ and $0 \partial \phi_i(\bar{x}) := \partial^\infty \phi(\bar{x})$ denotes the singular subdifferential. It is known that if the Slater condition is satisfied and each ϕ_i is locally

Lipschitz then (1.1) satisfies the KKT condition at every local solution of (1.1) (for the details see [6, Sects. 6.1, 6.3]). As a strengthened condition of the Lagrange multiplier, the KKT condition plays an important role in mathematical programming and has been extensively studied by many authors (see [2,12,22,26] and references therein). As a condition stronger than (1.4), we say that (1.1) satisfies the strong KKT condition at a feasible point $\bar{x}$ if

$$0 \in \text{int}\left(\partial\phi_0(\bar{x}) + \sum_{i\in I(\bar{x})} \mathbb{R}_+\partial\phi_i(\bar{x}) + N(C,\bar{x})\right). \tag{1.5}$$

We show that if $\bar{x} \in \Omega$ is a sharp minimum of some penalty function p_τ and each ϕ_i is locally Lipschtz (without any constraint qualification) then $\bar{x}$ is an isolated solution of (1.1) and the strong KKT is satisfied at $\bar{x}$.

In this paper, our main aim is to study (1.1) under the assumption (which is always assumed in the remainder of this paper) that

$$\phi_i(x) := \psi_i(f_i(x)) \quad \forall x \in X, \tag{1.6}$$

where each $f_i : X \to Y$ is a smooth function and $\psi_i : Y \to \mathbb{R} \cup \{+\infty\}$ is a proper lower semicontinuous convex function ($i = 0, 1, \ldots, m$), and that C is a closed convex set in X. Many authors have studied the convex-composite problems (see [10,11,14,15,19,27]). In contrast to (1.5), now we define another stronger condition: (1.1) is said to satisfy the strong KKT$^+$ condition at a feasible point $\bar{x}$ if

$$0 \in \text{int}\left((f_0'(\bar{x}))^*(\partial\psi_0(f_0(\bar{x}))+\sum_{i\in I(\bar{x})}(f_i'(\bar{x}))^*(\mathbb{R}_+\partial\psi_i(f_i(\bar{x}))+N(C,\bar{x})\right), \tag{1.7}$$

We prove, in Sect. 3, that if (1.1) satisfies the strong KKT$^+$ condition at a feasible point $\bar{x}$ then it is a sharp minimum of some penalty function p_τ. In particular, in the case when each ϕ_i is locally Lipschtz and each $f_i'(\bar{x})$ is surjective, $\bar{x} \in \Omega$ is a sharp minimum of some penalty function p_τ if and only if (1.1) satisfies the strong KKT condition at $\bar{x}$. Moreover, we also prove that (1.1) satisfies the strong KKT condition at a feasible point $\bar{x}$ if and only if there exist $\eta, r, \tau \in (0, +\infty)$ such that

$$\eta B_{X^*} \subset rB_{X^*} \cap \partial\phi_0(\bar{x})+\sum_{i\in I(u)} rB_{X^*} \cap [0, \tau]\partial\phi_i(\bar{x})+\tau B_{X^*} \cap N(C,\bar{x}), \tag{1.8}$$

where B_{X^*} denotes the closed unit ball of the dual space X^*.

For a feasible point of (1.1), let

$$S_{\bar{x}} := \{x \in \Omega : \ \phi_0(x) \le \phi_0(\bar{x})\}.$$

It is easy to verify that if $\bar{x}$ is a local solution of (1.1) then each point of $S_{\bar{x}}$ close to $\bar{x}$ is also a local solution of (1.1). Clearly, $S_{\bar{x}} \cap B(\bar{x}, \delta) = \{\bar{x}\}$ for some $\delta > 0$ if and only if $\bar{x}$ is a local isolated solution of (1.1). Note that $N_c(S_{\bar{x}}, \bar{x}) = \hat{N}(S_{\bar{x}}, \bar{x}) = X^*$ if $\bar{x}$ is an isolated solution of (1.1), where $N_c(\cdot, \cdot)$ and $\hat{N}(\cdot, \cdot)$ denote respectively the Clarke normal cone and Fréchet normal cone (see Sect. 2 for their definitions). In order to take care the non-isolated solution case, we make the following definition as a natural extension of (1.8): for any $\bar{x} \in \Omega$, (1.1) is said to satisfy the quasi-strong KKT condition around $\bar{x}$ if there exist $\eta, r, \tau, \delta \in (0, +\infty)$ such that for all $u \in S_{\bar{x}} \cap B(\bar{x}, \delta)$,

$$\eta B_{X^*} \cap N_c(S_{\bar{x}}, u) \subset rB_{X^*} \cap \partial\phi_0(u) + \sum_{i \in I(u)} rB_{X^*} \cap [0, \tau]\partial\phi_i(u) \\ + \tau B_{X^*} \cap N(C, u). \tag{1.9}$$

Similarly, we say that (1.1) satisfies the quasi-strong KKT$^+$ condition around $\bar{x} \in \Omega$ if there exist $\eta, r, \tau, \delta \in (0, +\infty)$ such that for all $u \in S_{\bar{x}} \cap B(\bar{x}, \delta)$,

$$\eta B_{X^*} \cap N_c(S_{\bar{x}}, u) \subset A(r, \tau; u) + \tau B_{X^*} \cap N(C, u), \tag{1.10}$$

where

$$A(r, \tau; u) := (f_0'(u))^*(rB_{Y^*} \cap \partial\psi_0(f_0(u))) \\ + \sum_{i \in I(u)} (f_i'(u))^*(rB_{Y^*} \cap [0, \tau]\partial\psi_i(f_i(u))).$$

With the Fréchet normal cone replacing the Clarke normal cone in (1.9), one has a weaker notion: we say that (1.1) satisfies the sub-quasi-strong KKT condition around $\bar{x} \in \Omega$ if there exist $\eta, r, \tau, \delta \in (0, +\infty)$ such that for all $u \in S_{\bar{x}} \cap B(\bar{x}, \delta)$,

$$\eta B_{X^*} \cap \hat{N}(S_{\bar{x}}, u) \subset rB_{X^*} \cap \partial\phi_0(u) \\ + \sum_{i \in I(u)} rB_{X^*} \cap [0, \tau]\partial\phi_i(u) + \tau B_{X^*} \cap N(C, u). \tag{1.11}$$

In Sect. 4, we prove that if (1.1) satisfies the quasi-strong KKT$^+$ condition around $\bar{x} \in \Omega$ then there exists $\tau > 0$ such that the following results hold: (1) $\bar{x}$ is a local weak sharp minimum of p_τ over X, and (2) each x close to $\bar{x}$ is a local minimizer of p_τ over X if and only if x is a local solution of (1.1). Under the assumption that $\phi_1, \ldots, \phi_m$ are locally Lipschitz around $\bar{x}$, we prove (a) if $\bar{x}$ is a local weak sharp minimum of a certain penalty function p_τ over X and every local minimizer close to $\bar{x}$ of p_τ over X is feasible for (1.1) then (1.1) satisfies the sub-quasi-strong KKT condition around $\bar{x}$, and (b) if $f_0 = \cdots = f_m$ and $f_0'(\bar{x})$ is surjective then the conditions being quasi-strong KKT$^+$, quasi-strong KKT and sub-quasi-strong KKT are mutually equivalent. As applications, in Sect. 5, we extend some existing results on metric regularity for inequalities from the convex case to the nonconvex case.

2 Preliminaries

To facilitate our discussion, we review some standard notions in variational analysis. Let X be a Banach space with the closed unit ball denoted by B_X, and let X^* denote the dual space of X. For a closed subset A of X, let $\mathrm{int}(A)$, $\mathrm{cl}(A)$ and $\mathrm{bd}(A)$ respectively denote the interior, closure and boundary of A. For $a \in A$, let $T_c(A, a)$ denote the Clarke tangent cone of A at a, which is defined by

$$T_c(A, a) = \liminf_{x \xrightarrow{A} a, t \to 0^+} \frac{A - x}{t},$$

where $x \xrightarrow{A} a$ means that $x \to a$ with $x \in A$. Thus, $v \in T_c(A, a)$ if and only if, for each sequence $\{a_n\}$ in A converging to a and each sequence $\{t_n\}$ in $(0, \infty)$ decreasing to 0, there exists a sequence $\{v_n\}$ in X converging to v such that $a_n + t_n v_n \in A$ for all n.

We denote by $N_c(A, a)$ the Clarke normal cone of A at a, that is,

$$N_c(A, a) := \{x^* \in X^* : \langle x^*, h\rangle \leq 0 \text{ for all } h \in T_c(A, a)\}.$$

Let $\hat{N}(A, a)$ denote the Fréchet normal cone of A at a, that is,

$$\hat{N}(A, a) := \left\{ x^* \in X^* : \limsup_{x \xrightarrow{A} a} \frac{\langle x^*, x - a\rangle}{\|x - a\|} \leq 0 \right\}.$$

If A is convex, then $T_c(A, a) = \limsup\limits_{t \to 0^+} \frac{A-a}{t}$ and

$$N_c(A, a) = \hat{N}(A, a) = \{x^* \in X^* : \langle x^*, x\rangle \leq \langle x^*, a\rangle \text{ for all } x \in A\}.$$

The following approximate projection result for a general closed set (recently established in [31]) will be useful in the proofs of our main results.

Lemma 2.1 *Let X be a Banach (resp. Asplund) space and A a closed nonempty subset of X. Let $\gamma \in (0, 1)$. Then for any $x \notin A$ there exist $a \in \mathrm{bd}(A)$ and $a^* \in N_c(A, a)$ (resp $a^* \in \hat{N}(A, a)$) with $\|a^*\| = 1$ such that*

$$\gamma\|x - a\| < \min\{d(x, A), \langle a^*, x - a\rangle\}.$$

Let $\phi : X \to \mathbb{R} \cup \{+\infty\}$ be a proper lower semicontinuous function,

$$\mathrm{dom}(\phi) := \{x \in X : \phi(x) < +\infty\} \text{ and } \mathrm{epi}(\phi) := \{(x, t) \in X \times \mathbb{R} : \phi(x) \leq t\}.$$

For $x \in \text{dom}(\phi)$ and $h \in X$, let $\phi^{\uparrow}(x, h)$ denote the generalized directional derivative introduced by Rockafellar [20], that is,

$$\phi^{\uparrow}(x,h) := \lim_{\varepsilon \downarrow 0} \limsup_{z \xrightarrow{\phi} x, t\downarrow 0} \inf_{w \in h+\varepsilon B_X} \frac{\phi(z+tw)-\phi(z)}{t},$$

where the expression $z \xrightarrow{\phi} x$ means that $z \to x$ and $\phi(z) \to \psi(x)$. Let $\partial\phi(x)$ denote the Clarke–Rockafellar subdifferential of ϕ at x, that is,

$$\partial\phi(x) := \{x^* \in X^* : \langle x^*, h\rangle \leq \phi^{\uparrow}(x,h) \quad \forall h \in X\}.$$

Recall that the Fréchet subdifferential of ϕ at $x \in \text{dom}(\phi)$ is defined as

$$\hat{\partial}\phi(x) := \left\{x^* \in X^* : \liminf_{z \to x} \frac{\phi(z)-\phi(x)-\langle x^*, z-x\rangle}{\|z-x\|} \geq 0\right\}.$$

It is well known (cf. [17]) that

$$\hat{\partial}\phi(x) \subset \partial\phi(x). \tag{2.1}$$

When ϕ is convex, the Clarke–Rockafellar and Fréchet subdifferentials reduce to the one in the sense of convex analysis, that is,

$$\partial\phi(x) = \hat{\partial}\phi(x) = \{x^* \in X^* : \langle x^*, y-x\rangle \leq \phi(y)-\phi(x) \quad \forall y \in X\} \quad \forall x \in \text{dom}(\phi).$$

For a closed set A in X, let δ_A denote the indicator function of A. It is known (see [6]) that

$$N_c(A,a) = \partial\delta_A(a) \quad \forall a \in A$$

and

$$\partial\phi(x) = \{x^* \in X^* : (x^*, -1) \in N_c(\text{epi}(\phi), (x, \phi(x)))\} \quad \forall x \in \text{dom}(\phi).$$

Let $\partial^{\infty}\phi(x)$ denote the singular subdifferential of ϕ at $x \in \text{dom}(\phi)$ and be defined by

$$\partial^{\infty}\phi(x) := \{x^* \in X^* : (x^*, 0) \in N_c(\text{epi}(\phi), (x, \phi(x)))\}.$$

If ϕ is locally Lipschitz near x then $\partial^{\infty}\phi(x) = \{0\}$.

The following lemma (which is a consequence of [31, Proposition 2.2] and [17, Theorem 1.17]) will be useful for us.

Lemma 2.2 *Let X, Y be Banach spaces, $f : X \to Y$ a smooth function and let A be a closed convex subset of Y. Let $u \in f^{-1}(A)$ and suppose that $f'(u)$ is surjective. Then*

$$N_c(f^{-1}(A), u) = (f'(u))^*(N(A, f(u))).$$

In the case when ψ is directionally Lipschitz, the following lemma is known (see [6, Theorem 2.99]).

Lemma 2.3 *Let X, Y be Banach spaces, $f : X \to Y$ a smooth function and $\psi : Y \to \mathbb{R} \cup \{+\infty\}$ a proper lower semicontinuous convex function. Let $\phi(x) := \psi(f(x))$ for all $x \in X$. Let $\bar{x} \in \mathrm{dom}(\phi)$ and suppose that $f'(\bar{x})$ is surjective. Then, there exists $\delta > 0$ such that for any $u \in \mathrm{dom}(\phi) \cap B(\bar{x}, \delta)$,*

$$\partial\phi(u) = (f'(u))^*(\partial\psi(f(u))) \text{ and } \partial^\infty\phi(u) = (f'(u))^*(\partial^\infty\psi(f(u))). \quad (2.2)$$

Proof Since f is smooth and $f'(\bar{x})$ is surjective, there exists $\delta > 0$ such that $f'(u)$ is surjective for all $u \in B(\bar{x}, \delta)$. Define $F : X \times \mathbb{R} \to Y \times \mathbb{R}$ by

$$F(x, t) := (f(x), t) \quad \forall (x, t) \in X \times \mathbb{R}.$$

Then, $F'(x, t)(h, r) = (f'(x)(h), r)$ for all $(x, t), (h, r) \in X \times \mathbb{R}$ and so $F'(u, t)$ is surjective for all $(u, t) \in B(\bar{x}, \delta) \times \mathbb{R}$. Noting that $\mathrm{epi}(\phi) = F^{-1}(\mathrm{epi}(\psi))$, the convexity of ψ and Lemma 2.2 imply that for all $u \in \mathrm{dom}(\phi) \cap B(\bar{x}, \delta)$,

$$\begin{aligned} N_c(\mathrm{epi}(\phi), (u, \phi(u))) &= (F'(u, \phi(u)))^*(N(\mathrm{epi}(\psi), (f(u), \psi(f(u)))) \\ &= \{((f'(u))^*(y^*), r) : (y^*, r) \in N(\mathrm{epi}(\psi), (f(u), \psi(f(u))))\}. \end{aligned}$$

It follows that (2.2) holds.

Remark In the case when ψ is convex and f is smooth, one always has that

$$(f'(x))^*(\partial\psi(f(x))) \subset \partial\phi(x) \quad \forall x \in \mathrm{dom}(\phi)$$

[cf. [23, Theorem 10.6] and (2.1)]. This and (2.1) imply the following implications hold:

[strong KKT$^+$ $\Rightarrow$ strong KKT] and [quasi-strong KKT $\Rightarrow$sub-quasi-strong KKT].

3 Strong KKT condition and sharp solution

Recall our standing assumption (1.6). The definition of sharp minimum for the penalty functions and that of the strong KKT$^+$ and strong KKT conditions for constraint optimization problem (1.1) were defined in Sect. 1. Their interrelationships will be explored in this section.

Proposition 3.1 *Let $\bar{x}$ be a point in the feasible set Ω of (1.1). Then (1.1) satisfies the strong KKT$^+$ condition at $\bar{x}$ if and only if there exist $\eta, \gamma, \tau \in (0, +\infty)$ such that*

$$\begin{aligned} \eta B_{X^*} \subset (f_0'(\bar{x}))^*(\gamma B_{Y^*} \cap \partial\psi_0(f_0(\bar{x}))) + \sum_{i \in I(\bar{x})} (f_i'(\bar{x}))^*(\gamma B_{X^*} \cap [0,\ \tau]\partial\psi_i(f_i(\bar{x}))) \\ +\tau B_{X^*} \cap N(C, \bar{x}). \end{aligned} \quad (3.1)$$

Proof The sufficiency part is trivial. To prove the necessity part, suppose that (1.1) satisfies the strong KKT$^+$ condition. Then there exists $r > 0$ such that

$$rB_{X^*} \subset (f_0'(\bar{x}))^*(\partial\psi_0(f_0(\bar{x}))) + \sum_{i\in I(\bar{x})} (f_i'(\bar{x}))^*(\mathbb{R}_+\partial\psi_i(f_i(\bar{x}))) + N(C,\bar{x}).$$

Hence,

$$rB_{X^*} \subset \bigcup_{k,n=1}^{\infty} (A_{k,n} + nB_{X^*} \cap N(C,\bar{x})), \tag{3.2}$$

where

$$A_{k,n} := (f_0'(\bar{x}))^*(kB_{Y^*} \cap \partial\psi_0(f_0(\bar{x}))) + \sum_{i\in I(\bar{x})} (f_i'(\bar{x}))^*(kB_{Y^*} \cap [0,\ n]\partial\psi_i(f_i(\bar{x}))).$$

Since $(f_i'(\bar{x}))^*$ is weak*–weak* continuous and since both $kB_{Y^*} \cap \partial\psi_0(f_0(\bar{x}))$ and $kB_{Y^*} \cap [0,\ n]\partial\psi_i(f_i(\bar{x}))$ are weak*-compact, $A_{k,n}$ is weak*-compact. Therefore, $A_{k,n} + nB_{X^*} \cap N(C,\bar{x})$ is weak*-compact and so is (norm-) closed. Since every open subspace of a complete metric space is a Baire space (cf. [18, Theorem 48.2 and Lemma 48.4]), $\text{int}(rB_{X^*})$ is a Baire space. It follows from (3.2) that there exist $x_0^* \in \text{int}(rB_{X^*})$, $r_0 > 0$, k_0 and n_0 such that

$$x_0^* + r_0B_{X^*} \subset A_{k_0,n_0} + n_0B_{X^*} \cap N(C,\bar{x}).$$

On the other hand, (3.2) implies that there exist natural numbers k_1 and n_1 such that $-x_0^* \in A_{k_1,n_1} + n_1B_{X^*} \cap N(C,\bar{x})$. Hence, by convexity, one has

$$\frac{r_0}{2}B_{X^*} \subset A_{\tilde{k},\tilde{n}} + \tilde{n}B_{X^*} \cap N(C,\bar{x})$$

with $\tilde{k} = \max\{k_0, k_1\}$ and $\tilde{n} = \max\{n_0, n_1\}$. This completes the proof.

Similarly, we can prove the following proposition.

Proposition 3.2 *Let $\bar{x} \in \Omega$. Then (1.1) satisfies the strong KKT condition at $\bar{x}$ if and only if there exist $\eta, \gamma, \tau \in (0,\ +\infty)$ such that*

$$\eta B_{X^*} \subset \gamma B_{X^*} \cap \partial\phi_0(\bar{x}) + \sum_{i\in I(\bar{x})} \gamma B_{X^*} \cap [0,\ \tau]\partial\phi_i(\bar{x}) + \tau B_{X^*} \cap N(C,\bar{x}).$$

The following theorem provides a relationship between the strong KKT$^+$/KKT condition and the sharp minimum property for the penalty function p_τ given by (1.2).

Theorem 3.1 *Let $\bar{x} \in \Omega$. The following statements are valid.*

(i) *If* (1.1) *satisfies the strong KKT*$^+$ *condition at* $\bar{x}$ *then* $\bar{x}$ *is an isolate solution of* (1.1) *and it is a sharp minimum of some penalty function* p_τ.
(ii) *Suppose that* $\phi_1, \ldots, \phi_m$ *are locally Lipschitz around* $\bar{x}$. *Then* (1.1) *satisfies the strong KKT condition at* $\bar{x}$ *whenever there exists* $\tau > 0$ *such that* $\bar{x}$ *is a sharp minimum of the penalty function* p_τ.
(iii) *Suppose that* $\phi_1, \ldots, \phi_m$ *are locally Lipschitz around* $\bar{x}$ *and that each* $f_i'(\bar{x})$ *is surjective. Then* (1.1) *satisfies the strong KKT condition at* $\bar{x}$ *if and only if* $\bar{x}$ *is a sharp minimum of some penalty function* p_τ.

Proof Suppose that (1.1) satisfies the strong KKT$^+$ condition at $\bar{x}$. By Proposition 3.1, there exist $\eta, \gamma, \tau \in (0, +\infty)$ such that (3.1) holds. Fix an ε in $(0, \frac{\eta}{(m+1)\gamma})$ and take $\delta > 0$ such that

$$\|f_i(x) - f_i(\bar{x}) - f_i'(\bar{x})(x - \bar{x})\| \le \varepsilon\|x - \bar{x}\| \quad \forall x \in B(\bar{x}, \delta).$$

Hence

$$\langle y^*, f_i'(\bar{x})(x - \bar{x})\rangle \le \langle y^*, f_i(x) - f_i(\bar{x})\rangle + \|y^*\|\varepsilon\|x - \bar{x}\| \quad \forall (y^*, x) \in Y^* \times B(\bar{x}, \delta),$$

that is, for all $(y^*, x) \in Y^* \times B(\bar{x}, \delta)$,

$$\langle (f_i'(\bar{x}))^*(y^*), x - \bar{x}\rangle \le \langle y^*, f_i(x) - f_i(\bar{x})\rangle + \|y^*\|\varepsilon\|x - \bar{x}\|. \tag{3.3}$$

To prove (i), let $x \in B(\bar{x}, \delta)$. By the Hahn–Banach theorem (cf. [24, P. 59, Corollary]), there exists $x^* \in \eta B_{X^*}$ such that

$$\eta\|x - \bar{x}\| = \langle x^*, x - \bar{x}\rangle. \tag{3.4}$$

By (3.1), there exist $y_0^* \in \gamma B_{Y^*} \cap \partial\psi_0(f_0(\bar{x}))$, $y_i^* \in \gamma B_{Y^*} \cap [0, \ \tau]\partial\psi_i(f_i(\bar{x}))$ and $y^* \in \tau B_{X^*} \cap N(C, \bar{x})$ such that

$$x^* = (f_0'(\bar{x}))^*(y_0^*) + \sum_{i \in I(\bar{x})} (f_i'(\bar{x}))^*(y_i^*) + y^*. \tag{3.5}$$

Noting that $\partial d(\cdot, C)(\bar{x}) = B_{X^*} \cap N(C, \bar{x})$ (by the convexity of C), one has

$$\langle y^*, x - \bar{x}\rangle \le \tau d(x, C). \tag{3.6}$$

By (3.3) and the convexity of ψ_0, we have

$$\begin{aligned}\langle (f_0'(\bar{x}))^*(y_0^*), x - \bar{x}\rangle &\le \langle y_0^*, f_0(x) - f_0(\bar{x})\rangle + \|y_0^*\|\varepsilon\|x - \bar{x}\| \\ &\le \phi_0(x) - \phi_0(\bar{x}) + \gamma\varepsilon\|x - \bar{x}\|.\end{aligned} \tag{3.7}$$

Let

$$I_0(\bar{x}) := \{i \in I(\bar{x}) : \ y_i^* \in \partial^\infty\psi_i(f_i(\bar{x}))\}.$$

Noting that $\partial^{\infty}\psi_i(f_i(\bar{x})) = N(\mathrm{dom}(\psi_i), f_i(\bar{x}))$ (because ψ_i is convex),

$$\langle y_i^*, f_i(u) - f_i(\bar{x})\rangle \leq 0 \quad \forall i \in I_0(\bar{x}) \text{ and } \forall u \in \mathrm{dom}(\phi_i)$$

Hence,

$$\langle y_i^*, f_i(u) - f_i(\bar{x})\rangle \leq \tau[\phi_i(u)]_+ \quad \forall i \in I_0(\bar{x}) \text{ and } \forall u \in X.$$

This and (3.3) imply that

$$\langle (f_i'(\bar{x}))^*(y_i^*), x - \bar{x}\rangle \leq \tau[\phi_i(x)]_+ + \gamma\varepsilon\|x - \bar{x}\| \quad \forall i \in I_0(\bar{x}). \tag{3.8}$$

Now let $i \in I(\bar{x})\backslash I_0(\bar{x})$. Then, there exist $t_i \in (0, \ \tau]$ and $v_i^* \in \partial\psi_i(f_i(\bar{x}))$ such that $y_i^* = t_i v_i^*$. By (3.3) and the convexity of ψ_i, it is easy to verify that

$$\begin{aligned} \langle (f_i'(\bar{x}))^*(y_i^*), x - \bar{x}\rangle &\leq \langle y_i^*, f_i(x) - f_i(\bar{x})\rangle + \gamma\varepsilon\|x - \bar{x}\| \\ &= t_i\langle v_i^*, f_i(x) - f_i(\bar{x})\rangle + \gamma\varepsilon\|x - \bar{x}\| \\ &\leq t_i\phi_i(x) + \gamma\varepsilon\|x - \bar{x}\| \\ &\leq \tau[\phi_i(x)]_+ + \gamma\varepsilon\|x - \bar{x}\|. \end{aligned}$$

It follows from (3.5), (3.6), (3.7) and (3.8) that

$$\langle x^*, x - \bar{x}\rangle \leq \phi_0(x) - \phi_0(\bar{x}) + \tau \sum_{i \in I(\bar{x})} [\phi_i(x)]_+ + (m+1)\gamma\varepsilon\|x - \bar{x}\| + \tau d(x, C)$$

This and (3.4) imply that

$$\begin{aligned} (\eta - (m+1)\gamma\varepsilon)\|x - \bar{x}\| &\leq \phi_0(x) - \phi_0(\bar{x}) + \tau\left(\sum_{i \in I(\bar{x})} [\phi_i(x)]_+ + d(x, C)\right) \\ &= p_\tau(x) - p_\tau(\bar{x}). \end{aligned}$$

Hence, $\bar{x}$ is a sharp minimum of p_τ. Noting that $p_\tau(x) = \phi_0(x)$ for all $x \in \Omega$, it follows that (i) holds.

To prove (ii), suppose that $\eta, \tau, \delta \in (0, \ +\infty)$ such that

$$\eta\|x - \bar{x}\| \leq \phi_0(x) - \phi_0(\bar{x}) + \tau\sum_{i=1}^{m}[\phi_i(x)]_+ + \tau d(x, C) \quad \forall x \in B(\bar{x}, \delta).$$

Let x^* be an arbitrary point in B_{X^*}. Then,

$$\begin{aligned} \langle \eta x^*, x - \bar{x}\rangle \leq \eta\|x - \bar{x}\| &\leq \phi_0(x) - \phi_0(\bar{x}) \\ &+\tau\sum_{i=1}^{m}[\phi_i(x)]_+ + \tau d(x, C) \quad \forall x \in B(\bar{x}, \delta). \end{aligned}$$

Noting $\bar{x} \in \Omega$, this implies that $\bar{x}$ is a local minimizer of the function

$$x \mapsto -\langle \eta x^*, x - \bar{x} \rangle + \phi_0(x) - \phi_0(\bar{x}) + \tau \sum_{1 \le i \le n} [\phi_i(x)]_+ + \tau d(x.C).$$

Since ϕ_i $(1 \le i \le m)$ is locally Lipschtz at $\bar{x}$, it follows from [6, P. 105, Corollary 1] and [6, Proposition 2.3.12] that

$$\begin{aligned} \eta x^* &\in \partial \phi_0(\bar{x}) + \tau \sum_{i=1}^{n} \partial [\phi_i]_+(\bar{x}) + \tau \partial d(\cdot, C)(\bar{x}) \\ &\subset \partial \phi_0(\bar{x}) + \tau \sum_{i \in I(\bar{x})} [0,\ 1] \partial \phi_i(\bar{x}) + \tau B_{X^*} \cap N(C, \bar{x}). \end{aligned}$$

Therefore,

$$\eta B_{X^*} \subset \partial \phi_0(\bar{x}) + \sum_{i \in I(\bar{x})} \mathbb{R}_+ \partial \phi_i(\bar{x}) + N(C, \bar{x}).$$

This shows that (ii) holds.
(iii) is immediate from (i), (ii) and Lemma 2.3. The proof is completed.

Remark The full smoothness assumption of each f_i was used only for the part (iii). Under the assumption of Theorem 3.1 (iii), $\phi_i = \psi_i \circ f_i$ is an amenable function (cf. [23, P. 442]).

4 Relationship between quasi-strong KKT conditions and weak sharp solutions

In this section, we consider the non-isolated solution case. We explore interrelationships between local weak sharp minima of a certain penalty function p_τ and quasi-strong KKT conditions. Let $\bar{x}$ be in the feasible set Ω of optimization problem (1.1), and recall that $S_{\bar{x}} := \{x \in \Omega : \ \phi_0(x) \le \phi_0(\bar{x})\}$. For $\tau > 0$, let

$$S_{\bar{x}}(\tau) := \{x \in X : \ p_\tau(x) \le p_\tau(\bar{x})\}.$$

Clearly, $S_{\bar{x}} \subset S_{\bar{x}}(\tau)$. Also it is easy to verify that if $\bar{x}$ is an isolated local minimum of a certain penalty function p_τ over X then $\bar{x}$ is an isolated local solution of (1.1), that is, there exists $\delta > 0$ such that

$$S_{\bar{x}}(\tau) \cap B(\bar{x}, \delta) = S_{\bar{x}} \cap B(\bar{x}, \delta) = \{\bar{x}\}.$$

Now suppose that $\bar{x}$ is a nonisolated local minimum of a certain penalty function p_τ over X. Thus there exists $\delta > 0$ such that $S_{\bar{x}}(\tau) \cap B(\bar{x}, \delta)$ is not a singleton and each point x in $S_{\bar{x}}(\tau) \cap B(\bar{x}, \delta)$ is a minimum of p_τ on $B(\bar{x}, \delta)$. In this case, a natural question arises: whether or not there exists $\delta > 0$ such that

$$S_{\bar{x}}(\tau) \cap B(\bar{x}, \delta) = S_{\bar{x}} \cap B(\bar{x}, \delta); \tag{4.1}$$

clearly, (4.1) means that each point x in $S_{\bar{x}}(\tau) \cap B(\bar{x}, \delta)$ is a local solution of (1.1). Unfortunately, the answer to the above question is not positive. Indeed, (4.1) does not necessarily hold for all $\delta > 0$ even when $\bar{x}$ is a local weak sharp minimum of p_τ over X (see the example at the end of this section). This leads us to define the following notion: $\bar{x} \in \Omega$ is said to be a local weak sharp solution of (1.1) if there exist $\tau, \delta \in (0, +\infty)$ such that (4.1) holds and $\bar{x}$ is a local weak sharp minimum of p_τ over X. Since $S_{\bar{x}}$ is a closed subset of $S_{\bar{x}}(\tau)$, it is clear that $\bar{x}$ is a local weak sharp solution of (1.1) if and only if there exist $\eta, \tau, \sigma \in (0, +\infty)$ such that

$$\eta d(x, S_{\bar{x}}) \leq \phi_0(x) - \phi_0(\bar{x}) + \tau \sum_{i=1}^{m} [\phi_i(x)]_+ + \tau d(x, C) \quad \forall x \in B(\bar{x}, \sigma). \tag{4.2}$$

Theorem 4.1 *Let $\bar{x} \in \Omega$ and suppose that (1.1) satisfies the quasi-strong KKT$^+$ condition around $\bar{x}$. Then $\bar{x}$ is a local weak sharp solution of (1.1).*

Proof By the assumption, there exist $\eta, r, \tau, \delta \in (0, +\infty)$ such that (1.10) holds for all $u \in S_{\bar{x}} \cap B(\bar{x}, \delta)$. Since each f_i is smooth, we can assume further that

$$\|f_i(x) - f_i(u) - f_i'(u)(x - u)\| \leq \frac{\eta}{2r(m+1)} \|x - u\| \quad \forall x, u \in B(\bar{x}, \delta) \tag{4.3}$$

(consider smaller δ if necessary). Let $x \in B(\bar{x}, \frac{\delta}{2})$. We need only show that

$$\frac{1}{2} \eta d(x, S_{\bar{x}}) \leq \phi_0(x) - \phi_0(\bar{x}) + \tau \sum_{i=1}^{m} [\phi_i(x)]_+ + \tau d(x, C). \tag{4.4}$$

For this, we can assume that $x \notin S_{\bar{x}}$. Then $0 < d(x, S_{\bar{x}}) \leq \|x - \bar{x}\| < \frac{\delta}{2}$. Let $\gamma \in (\max\{\frac{2d(x, S_{\bar{x}})}{\delta}, \frac{1}{2}\}, 1)$. By Lemma 2.1, there exist $u \in S_{\bar{x}}$ and $u^* \in N_c(S_{\bar{x}}, u) \cap B_{X^*}$ such that

$$\gamma \|x - u\| \leq \min\{d(x, S_{\bar{x}}), \langle u^*, x - u \rangle\}. \tag{4.5}$$

Thus, $u \in S_{\bar{x}} \cap B(\bar{x}, \delta)$. It follows from (1.10) that there exist $y_0^* \in r B_{Y^*} \cap \partial \psi_0(f_0(u))$,

$$y_i^* \in r B_{Y^*} \cap [0, \tau] \partial \psi_i(f_i(u)) \quad \forall i \in I(u) \tag{4.6}$$

and $y^* \in \tau B_{X^*} \cap N(C, u)$ such that

$$\eta u^* = (f_0'(u))^*(y_0^*) + \sum_{i \in I(u)} (f_i'(u))^*(y_i^*) + y^*.$$

By (4.3) and the convexity of ψ_0, one has

$$\begin{aligned}\langle \eta u^*, x-u\rangle &= \langle y_0^*, f_0'(u)(x-u)\rangle + \sum_{i\in I(u)}\langle y_i^*, f_i'(u)(x-u)\rangle + \langle y^*, x-u\rangle\\ &\le \langle y_0^*, f_0(x)-f_0(u)\rangle + \sum_{i\in I(u)}\langle y_i^*, f_i(x)-f_i(u)\rangle + \langle y^*, x-u\rangle + \frac{\eta}{2}\|x-u\|\\ &\le \phi_0(x)-\phi_0(u) + \sum_{i\in I(u)}\langle y_i^*, f_i(x)-f_i(u)\rangle + \langle y^*, x-u\rangle + \frac{\eta}{2}\|x-u\|\\ &= \phi_0(x)-\phi_0(\bar{x}) + \sum_{i\in I(u)}\langle y_i^*, f_i(x)-f_i(u)\rangle + \langle y^*, x-u\rangle + \frac{\eta}{2}\|x-u\|.\end{aligned}$$

By (4.5) and $d(x, S_{\bar{x}}) \le \|x-u\|$, it follows that

$$\begin{aligned}\left(\gamma-\frac{1}{2}\right)\eta d(x, S_{\bar{x}}) &\le \phi_0(x)-\phi_0(\bar{x}) + \sum_{i\in I(u)}\langle y_i^*, f_i(x)-f_i(u)\rangle + \langle y^*, x-u\rangle\\ &\le \phi_0(x)-\phi_0(\bar{x}) + \sum_{i\in I(u)}\langle y_i^*, f_i(x)-f_i(u)\rangle + \tau d(x, C)\end{aligned}\tag{4.7}$$

(the last inequality holds because $y^* \in \tau B_{X^*}\cap N(C, u) = \tau\partial d(\cdot, C)(u)$). Next we show that

$$\langle y_i^*, f_i(x)-f_i(u)\rangle \le \tau[\phi_i(x)]_+ \quad \forall i \in I(u) \tag{4.8}$$

To do this, let $i \in I(u)$. By (4.6), we have two cases to consider: (a) $y_i^* \in \partial^\infty\psi_i(f_i(u))$ or (b) $y_i^* = t_i v_i^*$ for some $t_i \in (0, \ \tau]$ and $v_i^* \in \partial\psi_i(f_i(u))$. In the later case, we have

$$\begin{aligned}\langle y_i^*, f_i(x)-f_i(u)\rangle &= t_i\langle v_i^*, f_i(x)-f_i(u)\rangle\\ &\le t_i(\phi_i(x)-\phi_i(u))\\ &= t_i\phi_i(x) \le \tau[\phi_i(x)]_+\end{aligned}$$

and so (4.8) holds. In the case (a), by the definition of the singular subdifferential, the convexity of ψ_i implies that

$$\langle y_i^*, f_i(z)-f_i(u)\rangle \le 0 \quad \forall z \in f_i^{-1}(\mathrm{dom}(\psi_i))$$

and so $\langle y_i^*, f_i(x)-f_i(u)\rangle \le \tau[\phi_i(x)]_+$. Therefore (4.8) is true. By (4.7) and (4.8), one has

$$\begin{aligned}\left(\gamma-\frac{1}{2}\right)\eta d(x, S_{\bar{x}}) &\le \phi_0(x)-\phi_0(\bar{x}) + \sum_{i\in I(u)}[\phi_i(x)]_+ + \tau d(x, C)\\ &= \phi_0(x)-\phi_0(\bar{x}) + \sum_{i=1}^{m}[\phi_i(x)]_+ + \tau d(x, C).\end{aligned}$$

This gives (4.4) by letting $\gamma \to 1$. The proof is completed.

Theorem 4.2 *Let $\bar{x} \in \Omega$ and $\tau > 0$. Suppose that $\bar{x}$ is a local weak sharp solution of (1.1) and that $\phi_1, \ldots, \phi_m$ are locally Lipschtz around $\bar{x}$. Then (1.1) satisfies the sub-quasi-strong KKT condition around $\bar{x}$.*

Proof Take $\eta, \sigma, L \in (0, +\infty)$ such that (4.2) holds and

$$|\phi_i(x_1) - \phi_i(x_2)| \le L\|x_1 - x_2\| \text{ for all } x_1, x_2 \in B(\bar{x}, \sigma) \text{ and } 1 \le i \le m. \quad (4.9)$$

Let $u \in S_{\bar{x}} \cap B(\bar{x}, \frac{\sigma}{2})$ and $u^* \in \hat{N}(S_{\bar{x}}, u) \cap B_{X^*}$. Noting that $\hat{N}(S_{\bar{x}}, u) \cap B_{X^*} = \hat{\partial} d(\cdot, S_{\bar{x}})(u)$ (cf. [17, Corollary 1.96]), it follows that for any $\varepsilon > 0$ there exists $\sigma_1 \in (0, \frac{\sigma}{2})$ such that

$$\langle u^*, x - u\rangle \le d(x, S_{\bar{x}}) + \varepsilon\|x - u\| \quad \forall x \in B(u, \sigma_1).$$

This and (4.2) imply that

$$\begin{aligned}\langle \eta u^*, x - u\rangle \le \phi_0(x) - \phi_0(\bar{x}) + \tau \sum_{i=1}^{m} [\phi_i(x)]_+ + \tau d(x, C) \\ + \eta\varepsilon\|x - u\| \quad \forall x \in B(u, \sigma_1).\end{aligned}$$

Since $\phi_0(\bar{x}) = \phi_0(u)$ [by (4.2) and the fact that $u \in S_{\bar{x}}$], this means that u is a minimum of the function g over $B(u, \sigma_1)$ defined by

$$\begin{aligned}g(x) := -\langle \eta u^*, x - u\rangle + \phi_0(x) - \phi_0(\bar{x}) + \tau \sum_{i=1}^{m} [\phi_i(x)]_+ \\ + \tau d(x, C) + \eta\varepsilon\|x - u\| \quad \forall x \in X.\end{aligned}$$

Then $0 \in \partial g(u)$. Thus, by [6, P. 105, Corollary 1] and [6, Proposition 2.3.12], one has

$$\begin{aligned}0 &\in -\eta u^* + \partial\phi_0(u) + \tau \sum_{i=1}^{m} \partial[\phi_i]_+(u) + \tau\partial d(\cdot, C)(u) + \eta\varepsilon\partial\| \cdot - u\|(u) \\ &= -\eta u^* + \partial\phi_0(u) + \tau \sum_{i \in I(u)} [0, 1]\partial\phi_i(u) + \tau B_{X^*} \cap N(C, u) + \eta\varepsilon B_{X^*}.\end{aligned}$$

Letting $\varepsilon \to 0$, one has $\eta u^* \in \partial\phi_0(u) + \sum_{i \in I(u)} [0, \tau]\partial\phi_i(u) + \tau B_{X^*} \cap N(C, u)$. This and (4.9) imply that

$$\eta u^* \in r B_{X^*} \cap \partial\phi_0(u) + \sum_{i \in I(u)} r B_{X^*} \cap [0, \tau]\partial\phi_i(u) + \tau B_{X^*} \cap N(C, u)$$

with $r = (1 + mL)\tau + \eta$. This shows that (1.11) holds for all $u \in B(\bar{x}, \frac{\sigma}{2})$ and hence the proof is completed.

Remark We cannot use $\bar{x}$ is a local weak sharp minimum of a certain penalty p_τ over X" to replace the corresponding assumption in Theorem 4.2 (see the example at the end of this section).

Proposition 4.1 *Let $\bar{x} \in \Omega$ and suppose that $C = X$, $Y_i = Y$, $f_i = f$ $(0 \le i \le m)$ with $f'(\bar{x})(X) = Y$. Then the following statement are equivalent.*

(i) (1.1) *satisfies the quasi-strong KKT$^+$ condition around $\bar{x}$.*
(ii) (1.1) *satisfies the quasi-strong KKT condition around $\bar{x}$.*
(iii) (1.1) *satisfies the sub-quasi-strong KKT condition around $\bar{x}$.*

Proof By the smoothness of f and $f'(\bar{x})(X) = Y$, there exists $\delta > 0$ such that $f'(u)$ is surjective for all $u \in B(\bar{x}, \delta)$. We claim that

$$N_c(S_{\bar{x}}, u) = \hat{N}(S_{\bar{x}}, u) \quad \forall u \in S_{\bar{x}} \cap B(\bar{x}, \delta). \tag{4.10}$$

Let $A := \{y \in Y : \ \psi_0(y) \le \psi_0(f(\bar{x}))$ and $\psi_i(y) \le 0$ for $1 \le i \le m\}$. Then A is a closed convex subset of Y and $S_{\bar{x}} = f^{-1}(A)$. By Lemma 2.2, one has

$$N_c(S_{\bar{x}}, u) = (f'(u))^*(\hat{N}(A, f(u)) \quad \forall u \in S_{\bar{x}} \cap B(\bar{x}, \delta).$$

This and [17, Theorem 1.17] imply that (4.10) holds. From (4.10) and Lemma 2.3, it is easy to verify that (i)⇔(ii)⇔(iii). The proof is completed.

Let $\psi : Y \to \mathbb{R} \cup \{+\infty\}$ be a proper lower semicontinuous convex function; recall that for any $y \in \mathrm{dom}(\psi)$ and $v \in Y$, the directional derivative $d^+\psi(y, v)$ always exists (as an extended real number) and is defined by $d^+\psi(y, v) = \lim\limits_{t \to 0^+} \frac{\psi(y+tv)-\psi(y)}{t}$.

Proposition 4.2 *Let $\bar{x}$, C, $Y_i = Y$, $f_i = f$ be as in Proposition (4.1) and suppose that $\phi_1, \ldots, \phi_m$ are locally Lipschitz around $\bar{x}$. Let $\eta, \tau \in (0, \ +\infty)$. Then the following statements are equivalent.*

(i) $N_c(S_{\bar{x}}, \bar{x}) \cap \eta B_{X^*} \subset \partial\phi_0(\bar{x}) + \sum\limits_{i \in I(\bar{x})} [0, \ \tau]\partial\phi_i(\bar{x})$.

(ii) $\eta d(h, T_c(S_{\bar{x}}, \bar{x})) \le d^+\psi_0(f(\bar{x}), f'(\bar{x})(h)) + \tau \sum\limits_{i \in I(\bar{x})} [d^+\psi_i(f(\bar{x}), f'(\bar{x})(h))]_+$ *for all $h \in X$.*

Proof Since $f'(\bar{x})$ is surjective, the local Lipschitz assumption implies that for $1 \le i \le m$, $f(\bar{x}) \in \mathrm{int}(\mathrm{dom}(\psi_i))$ and so $d^+\psi_i(f(\bar{x}), \cdot)$ is Lipschitz. Let $\tilde{\phi}_i(h) := d^+\psi_i(f(\bar{x}), f'(\bar{x})(h))$ for all $h \in X$ $(i = 0, 1, \ldots, m)$. By Lemma 2.3, one has

$$\partial\tilde{\phi}_i(0) = (f'(\bar{x}))^*(\partial d^+\psi_i(f(\bar{x}), \cdot)(0)) = (f'(\bar{x}))^*(\partial\psi_i(f(\bar{x}))) = \partial\phi_i(\bar{x}). \tag{4.11}$$

First, suppose that (ii) holds. Similar to the proof of Theorem 4.2, one has

$$\hat{N}(T_c(S_{\bar{x}}, \bar{x}), 0) \cap \eta B_{X^*} \subset \partial\tilde{\phi}_0(0) + \sum_{i \in I(\bar{x})} [0, \ \tau]\partial\tilde{\phi}_i(0). \tag{4.12}$$

Since $T_c(S_{\bar{x}}, 0)$ is convex,

$$\hat{N}(T_c(S_{\bar{x}}, \bar{x}), 0) = N_c(T_c(S_{\bar{x}}, \bar{x}), 0) = N_c(S_{\bar{x}}, \bar{x}).$$

It follows from (4.11) and (4.12) that (i) holds.

Now suppose that (i) holds. Let $h \in X \backslash T_c(S_{\bar{x}}, \bar{x})$ and $\gamma \in (0, \ 1)$. Then, by Lemma 2.1, there exist $u \in T_c(S_{\bar{x}}, \bar{x})$ and $u^* \in N_c(T_c(S_{\bar{x}}, \bar{x}), u)) \cap B_{X^*}$ such that

$$\gamma \|h - u\| \le \langle u^*, h - u\rangle$$

Since $T_c(S_{\bar{x}}, \bar{x})$ is a closed convex cone,

$$N_c(T_c(S_{\bar{x}}, \bar{x}), u)) \subset N_c(T_c(S_{\bar{x}}, \bar{x}), 0)) = N_c(S_{\bar{x}}, \bar{x})$$

and $\langle u^*, u\rangle = 0$. Hence, $\gamma d(h, T_x(S_{\bar{x}}, \bar{x})) \le \gamma \|x - u\| \le \langle u^*, h\rangle$; moreover (i) and (4.11) imply that there exist $x_i^* \in \partial \psi_i(f(\bar{x}))$ and $t_i \in [0, \ \tau]$ such that

$$\eta u^* = (f'(\bar{x}))^*(x_0^*) + \sum_{i \in I(\bar{x})} t_i (f'(\bar{x}))^*(x_i^*).$$

It follows that

$$\begin{aligned}
\gamma \eta d(h, T_c(S_{\bar{x}}(\tau), \bar{x})) &\le \langle (f'(\bar{x}))^*(x_0^*) + \sum_{i \in I(\bar{x})} t_i (f'(\bar{x}))^*(x_i^*), h\rangle \\
&= \langle x_0^*, f'(\bar{x})(h)\rangle + \sum_{i \in I(\bar{x})} t_i \langle x_i^*, f'(\bar{x})(h)\rangle \\
&\le d^+\psi_0(f(\bar{x}, f'(\bar{x})(h)) + \sum_{i \in I(\bar{x})} t_i d^+\psi_i(f(\bar{x}), f'(\bar{x})(h)) \\
&\le d^+\psi_0(f(\bar{x}, f'(\bar{x})(h)) + \sum_{i \in I(\bar{x})} \tau [d^+\psi_i(f(\bar{x}), f'(\bar{x})(h))]_+.
\end{aligned}$$

Letting $\gamma \to 1$, one sees that (i) holds. The proof is complete.

Remark It is easy to verify that $\bar{x} \in \Omega$ is a local weak sharp solution of (1.1) if and only if there exist $\tau, \delta \in (0, \ +\infty)$ such that $S_{\bar{x}}(\tau) \cap B(\bar{x}, \delta)$ is contained in Ω and $\bar{x}$ is a local weak sharp minimum of p_τ over the entire space X. On the other hand, when $\bar{x} \in \Omega$ is a local weak sharp minimum of some penalty function p_τ over X, $\bar{x}$ is not necessarily a local weak sharp solution of (1.1), and it may even happen that all other local minimizers $x \ne \bar{x}$ of p_τ over X may be infeasible for (1.1). For example, let $X = \mathbb{R}$, $m = 1$, $\phi_0(x) = -x^2$ and $\phi_1(x) = x^2$ for for all $x \in \mathbb{R}$, and let $C = \mathbb{R}$. Then 0 is the unique feasible point. Let $\tau = 1$; then $p_\tau(x) = 0$ for all $x \in \mathbb{R}$. In particular, $\bar{x} = 0$ is a weak sharp minimum of p_τ over $\mathbb{R}$. On the other hand, noting that $S_{\bar{x}} = \{0\}$ and

$$\lim_{x \to 0} \frac{p'_\tau(x) - p'_\tau(0)}{d(x, S_{\bar{x}})} = \lim_{x \to 0} (-1 + \tau')|x| = 0 \quad \forall \tau' \in [0, \ +\infty),$$

the feasible point 0 is not a local weak sharp solution of (1.1).

Even in the convex case, we do not know whether a point $\bar{x}$ in Ω is a weak sharp solution of (1.1) if there exists $\tilde{\tau} > 0$ such that $\bar{x}$ is a local weak sharp minimum of p_τ over the entire space X for all $\tau \in [\tilde{\tau}, +\infty)$.

5 Applications to metric regularity for convex-composite inequalities

Let $\phi : X \to \mathbb{R} \cup \{+\infty\}$ and consider the following inequality

$$\phi(x) \leq 0. \tag{5.1}$$

Let S denote the solution set of (5.1), that is, $S := \{x \in X : \phi(x) \leq 0\}$. Following Lewis and Pang [13], we say that (5.1) is metrically regular at $\bar{x} \in S$ if there exist $\eta, \delta \in (0, +\infty)$ such that

$$\eta d(x, S) \leq [\phi(x)]_+ \quad \forall x \in B(\bar{x}, \delta). \tag{5.2}$$

In some references (cf. [4,28]), the metric regularity of (5.1) is named as the local error bound property of (5.1). It is known (cf. [28]) and easy to verify that (5.1) is metrically regular at $\bar{x}$ if and only if $\bar{x}$ is a weak sharp minimum of $[\phi]_+$ over X.

In terms of the normal cones of the solution set and the subdifferential of the concerned function in the solution set, many authors (cf. [3,5–8,13,16,28,29] and references therein) studied the metric regularity (with various names) of inequality (5.1). For example, under the assumption when $X = \mathbb{R}^n$ and ϕ is convex, Lewis and Pang [13] proved that the metric regularity of (5.1) at $\bar{x} \in S$ implies $N(S, \bar{x}) = \mathrm{cl}(\mathbb{R}_+\partial\phi(\bar{x}))$, a condition which is of course weaker than the well known basic constraint qualification (BCQ for short) of (5.1) at $\bar{x}$: $N(S, \bar{x}) = \mathbb{R}_+\partial\phi(\bar{x})$. In the case when ϕ is a smooth convex function, Li [16] proved that (5.1) is metrically regular at $\bar{x}$ if and only if (5.1) satisfies BCQ at each point of S close to $\bar{x}$. Following their line of investigation, using a Banach space geometrical technique, the authors [29] proved that under the convexity assumption on ϕ, (5.1) is metrically regular at $\bar{x}$ if and only if there exist $\tau, \delta \in (0, +\infty)$ such that for all $u \in S \cap B(\bar{x}, \delta)$,

$$N(S, u) \cap B_{X^*} \subset [0, \tau]\partial\phi(u). \tag{*}$$

With the help of the Fenchel dual technique, Burke and Deng [3] proved the same result for the special case when X is a Hilbert space; the readers can find other interesting characterization as well as results on $(*)$ and BCQ in their paper [3]. The above inclusion $(*)$, first appeared in Burke and Ferris [5], is stronger than the BCQ of (5.1) at u and so is named as strong BCQ in [29]. By giving a counterexample, the authors [29] also showed that (5.1) is not necessarily metrically regular at $\bar{x}$ if (5.1) satisfies the Strong BCQ only at the point $\bar{x}$ alone. Later, Hu [7,8] further considered relationships among the BCQ, the strong BCQ and the metric regularity for a convex inequality. Recently, Burke and Deng [4] considered the differentiable convex inclusion, and the authors [30] extended the above-mentioned results to more general setting, namely for the following so-called generalized convex equation

$$b \in F(x) \text{ subject to } x \in C$$

(where F is a convex multifunction between two Banach spaces X and Y, $b \in Y$ and C is a closed convex subset of X). All these previous works are under the convexity assumption. As applications of Theorems 4.1, 4.2 and 4.3, we can now establish the corresponding results (on the metric regularity) for inequality (5.1) in a nonconvex situation: ϕ is not necessarily convex but convex-composite.

In the remainder of this section, let $\phi = \psi \circ f$, where f is a smooth function between Banach spaces X and Y and $\psi : Y \to \mathbb{R}$ is a continuous convex function. The corresponding inequality (5.1) is said to satisfy
(α) strong BCQ around $\bar{x}$ if there exist $\tau, \delta \in (0, +\infty)$ such that

$$N_c(S, u) \cap B_{X^*} \subset [0, \tau]\partial\phi(u) \quad \forall u \in S \cap B(\bar{x}, \delta); \tag{5.3}$$

(β) sub-strong BCQ around $\bar{x}$ if the same as (α) but the Clarke normal cone $N_c(S, u)$ is replaced with the Fréchet normal cone $\hat{N}(S, u)$, that is,

$$\hat{N}(S, u) \cap B_{X^*} \subset [0, \tau]\partial\phi(u) \quad \forall u \in S \cap B(\bar{x}, \delta); \tag{5.4}$$

(γ) strong BCQ$^+$ around $\bar{x}$ if there exist $\tau, \delta \in (0, +\infty)$ such that

$$N_c(S, u) \cap B_{X^*} \subset (f'(u))^*([0, \tau]\partial\psi(f(u))) \quad \forall u \in S \cap B(\bar{x}, \delta). \tag{5.5}$$

Theorem 5.1 *Let $\phi = \psi \circ f$ and S be explained above. Let $\bar{x} \in S$ and consider the following statements*:

(i) *Inequality* (5.1) *is metrically regular at $\bar{x}$.*
(ii) (5.1) *satisfies the strong BCQ$^+$ around $\bar{x}$.*
(iii) (5.1) *satisfies the strong BCQ around $\bar{x}$.*
(iv) (5.1) *satisfies the sub-strong BCQ around $\bar{x}$.*
(v) *There exist $\eta, \delta \in (0, +\infty)$ such that*

$$\eta d(h, T_c(S, u)) \le [d^+\psi(f(\bar{x}), f'(x)(h))]_+$$

for all $u \in S \cap B(\bar{x}, \delta)$ and all $h \in X$.
Then the following assertions are valid.

(a) *(ii)⇒(iii)⇒(iv) and (ii)⇒(i) always hold.*
(b) *Suppose that $f'(\bar{x})$ is surjective. Then (i)–(v) are equivalent.*

Proof (a) Note that $(f'(x))^*(\partial\psi(f(x)) \subset \partial\phi(x)$ for all $x \in X$ and $\hat{N}(S, u) \subset N_c(S, u)$ for all $u \in S$. The implications (ii)⇒(iii)⇒(iv) are trivial. To prove (ii)⇒(i), suppose that there exist $\tau, \delta \in (0, +\infty)$ such that (5.5) holds. Let $C = X$, $\phi_0 = 0$ and $\phi_i = \phi$ for $i = 1, \ldots, m$. Then (1.1) reduces to inequality (5.1), $S_{\bar{x}} = S$, and $\bar{x}$ is a weak sharp solution of (1.1) if and only if (5.1) is metrically regular at $\bar{x}$. Since f is smooth and ψ is continuous and convex, we can assume that there exists $L > 0$ such that

$$\sup\{\|y^*\| : \ y^* \in \partial\psi(f(u)) \text{ for all } u \in B(\bar{x}, \delta)\} \leq L$$

(considering smaller δ if necessarily). It follows from (5.5) that

$$N_c(S, u) \cap B_{X^*} \subset (f'(u))^*(L\tau B_{Y^*} \cap [0, \ \tau]\partial\psi(f(u))) \quad \forall u \in S \cap B(\bar{x}, \delta).$$

Hence (1.1) satisfies the quasi-strong KKT$^+$ around $\bar{x}$. This and Theorem 4.1 imply that $\bar{x}$ is a weak sharp solution of (1.1), and hence (5.1) is metrically regular at $\bar{x}$. This shows that (a) holds. Moreover, by Propositions 4.1 and 4.2, we have that (i–v) are equivalent. The proof is completed.

Theorem 5.1b extends the corresponding results in [3,5,29] (cf. [3, Theorem 2.2], [5, Theorem 5.2] and [29, Theorems 2.2 and 2.3]), where they considered the special case when $X = Y$ and f is the identity mapping of X.

The study of the present paper proceeds in terms of the normal cones of the solution set and the subdifferentials of ϕ at points in the solution set. Different from our study, some authors considered the metric regularity of inequality (5.1) in terms of various properties of ϕ at points *outside the solution set*; the readers can see [1] and references therein for details. Let $F : X \rightrightarrows Y$ be a multifunction and recall another important metric regularity: F is called to be metrically regular at $(\bar{x}, \bar{y})$ if there exist $\eta, \delta \in (0, \ +\infty)$ such that

$$\eta d(x, F^{-1}(y)) \leq d(y, F(x)) \quad \forall (x, y) \in B(\bar{x}, \delta) \times B(\bar{y}, \delta).$$

Such a metric regularity is stronger and has been well studied. Ioffe [9] gave an excellent summary for the metric regularity of multifunctions.

Acknowledgments The authors wish to thank the referees for many valuable comments and for reference [1,4,9] as well as drawing authors' attention to amenable functions due to Poliquin and Rockafellar and the use of penalty functions.

References

1. Aze, D., Corvellec, J.-N.: Characterizations of error bounds for lower semicontinuous fucntions on metric spaces. ESAIM Control Optim. Calc. Var. **10**, 409–425 (2004)
2. Bazaraa, M., Sherali, H., Shetty, C.: Nonlinear Programming, Theory and Algorithms, 2nd edn. Wiley, New York (1993)
3. Burke, J.V., Deng, S.: Weak sharp minima revisited Part I: basic theory. Control Cybern. **31**, 439–469 (2002)
4. Burke, J.V., Deng, S.: Weak sharp minima revisited Part III: error bounds for differentiable convex inclusions, Math. Program. Ser. B (2009, in Press)
5. Burke, J.V., Ferris, M.C.: Weak sharp minima in mathematical programming. SIAM J. Optim. Control **36**, 1340–1359 (1993)
6. Clarke, F.H.: Optimization and Nonsmooth Analysis. Wiley, New York (1983)

7. Hu, H.: Characterizations of the strong basic constraint qualifications. Math. Oper. Res. **30**, 956–965 (2005)
8. Hu, H.: Characterizations of local and global error bounds for convex inequalities in Banach spaces. SIAM J. Optim. **18**, 309–321 (2007)
9. Ioffe, A.D.: Metric regularity and subdifferential calculus. Russ. Math. Survers **55**, 501–558 (2000)
10. Jeyakumar, V., Luc, D.T., Tinh, P.N.: Convex composite non-Lipschitz programming. Math. Program. **92**, 177–195 (2002)
11. Jeyarkumar, V., Yang, X.Q.: Convex composite multiobjective nonsmooth programming. Math. Program. 59 (1993)
12. Kuhn, H.W., Tucker, A.W.: Nonlinear programming. In: Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, Berkeley, pp. 81–492 (1950)
13. Lewis, A.S., Pang, J.S.: Error bounds for convex inequality systems. In: Crouzeix, J.P. (ed.), Generalized Convexity, Proceedings of the Fifth Sysposium on Generalized Convexity, Luminy Marseille, pp. 75–110 (1997)
14. Li, C., Ng, K.F.: Majoring functions and convergence of the Gauss–Newton method for convex composite optimization, **18**, pp. 613–642 (2007)
15. Li, C., Wang, X.: On convergence of the Gauss–Newton method for convex composite optimization. Math. Program. **91**, 349–356 (2002)
16. Li, W.: Abadie's constraint qualification, metric regularity, and error bounds for differentiable convex inequalities. SIAM J. Optim. **7**, 966–978 (1997)
17. Mordukhovich, B.S.: Variational Analysis and Generalized Differentiation I. Springer, Berlin (2006)
18. Munkres, J.R.: Topology. Prentice-Hall, Engle-wood Cliffs (1975)
19. Penot, J.P.: Optimality conditions in mathematical programming and composite optimization. Math. Program. **67**, 225–245 (1994)
20. Rockafellar, R.T.: Generalized directional derivatives and subgradient of nonconvex functions. Can. J. Math. **32**, 157–180 (1980)
21. Rockafellar, R.T.: First- and second-order epidifferentiability in nonlinear programming. Trans. Am. Math. Soc. **307**, 75–108 (1988)
22. Rockafellar, R.T.: Lagrange multipliers and optimality. SIAM Rev. **35**, 183–238 (1993)
23. Rockafellar, R.T., Wets, R.J.B.: Variational Analysis. Springer, Heidelberg (1998)
24. Rudin, W.: Functional Analysis. McGraw-Hill, New York (1973)
25. Studniarski, M., Ward, D.E.: Weak sharp minima: characterizations and sufficient conditions. SIAM J. Optim. Control **38**, 219–236 (1999)
26. Tapia, R.A., Trosset, M.W.: An extension of the Karush–Kuhn–Tucker necessity conditions to infinite programming. SIAM Rev. **36**, 1–17 (1994)
27. Womersley, R.S.: Local properties of algorithms for minimizing nonsmooth composite functions. Math. Program. **32**, 69–89 (1985)
28. Zalinescu, C.: Weak sharp minima, well-behaving functions and global error bounds for convex inequalities in Banach spaces. In: Proceedings of the 12th Baikal International Conference on Optimization Methods and Their Applications. Irkutsk, Russia, pp. 272–284 (2001)
29. Zheng, X.Y., Ng, K.F.: Metric regularity and constraint qualifications for convex inequalities on Banach spaces. SIAM J. Optim. **14**, 757–772 (2003)
30. Zheng, X.Y., Ng, K.F.: Metric subregularity and constraint qualifications for convex generalized equations in Banach spaces. SIAM J. Optim. **18**, 437–460 (2007)
31. Zheng, X.Y., Ng, K.F.: Linear regularity for a collection of subsmooth sets in Banach spaces. SIAM J. Optim. **19**, 62–76 (2008)