SIAM J. OPTIM.
Vol. 29, No. 4, pp. 2625–2649

DIRECTIONAL QUASI-/PSEUDO-NORMALITY AS SUFFICIENT CONDITIONS FOR METRIC SUBREGULARITY*

KUANG BAI†, JANE J. YE‡, AND JIN ZHANG§

Abstract. In this paper we study sufficient conditions for metric subregularity of a set-valued map which is the sum of a single-valued continuous map and a locally closed subset. First we derive a sufficient condition for metric subregularity which is weaker than the so-called first-order sufficient condition for metric subregularity (FOSCMS) by adding an extra sequential condition. Then we introduce directional versions of quasi-normality and pseudo-normality which are stronger than the new weak sufficient condition for metric subregularity but weaker than classical quasi-normality and pseudo-normality. Moreover we introduce a nonsmooth version of the second-order sufficient condition for metric subregularity and show that it is a sufficient condition for the new sufficient condition for metric subregularity to hold. An example is used to illustrate that directional pseudo-normality can be weaker than FOSCMS. For the class of set-valued maps where the single-valued mapping is affine and the abstract set is the union of finitely many convex polyhedral sets, we show that pseudo-normality and hence directional pseudo-normality holds automatically at each point of the graph. Finally we apply our results to complementarity and Karush–Kuhn–Tucker systems.

Key words. directional limiting normal cones, metric subregularity, calmness, error bounds, directional pseudo-normality, directional quasi-normality, complementarity systems

AMS subject classifications. 49J52, 49J53, 90C30, 90C31, 90C33

DOI. 10.1137/18M1232498

1. Introduction.

In this paper, we study stability analysis of the system of the form

$$P(x) \in \Lambda, \tag{1}$$

where $\mathscr{X}$, $\mathscr{Y}$ are finite-dimensional Hilbert spaces, $P : \mathscr{X} \to \mathscr{Y}$ is continuous near the point of interest and Λ is a subset of $\mathscr{Y}$ which is closed near the point of interest. Throughout the paper, unless otherwise specified, we assume that $\mathscr{Y}$ is an m-dimensional Hilbert space with inner product $\langle \cdot, \cdot \rangle$ equipped with the orthogonal basis $\mathscr{E} = \{e_1, \ldots, e_m\}$. Without loss of generality, throughout this paper for any $y \in \mathscr{Y}$ we denote $\langle y, e_i \rangle$ by y_i, $i = 1, \ldots, m$.

Since the set Λ is not required to be convex, the system represented by $P(x) \in \Lambda$ is very general and many systems can be formulated in this form. In particular, various variational inequalities/complementarity systems can be reformulated in this form. For example, consider the cone complementarity system defined as

$$\mathcal{K} \ni \Phi(x) \perp \Psi(x) \in \mathcal{K},$$

*Received by the editors December 12, 2018; accepted for publication (in revised form) June 19, 2019; published electronically October 22, 2019.
https://doi.org/10.1137/18M1232498
Funding: The research of the first author was partially supported by the China Scholarship Council. The second author's research was partially supported by the NSERC. The third author's research is supported by NSFC grants 11601458, 11971220, and 11871269.
†Department of Mathematics and Statistics, University of Victoria, BC, V8W 2Y2 Canada (kuangbai@uvic.ca).
‡Corresponding author. Department of Mathematics and Statistics, University of Victoria, BC, V8W 2Y2 Canada (janeye@uvic.ca).
§Department of Mathematics, Southern University of Science and Technology, Shenzhen, People's Republic of China (zhangj9@sustech.edu.cn).

where $\mathcal{K}$ is a convex cone in $\mathscr{Y}$, $\Phi, \Psi : \mathscr{X} \to \mathscr{Y}$, and $y \perp z$ means that $\langle y, z\rangle = 0$. Then the cone complementarity system can be reformulated in the form (1) by defining $P(x) := (\Phi(x), \Psi(x))$ and the complementarity set

$$\Lambda := \{(y, z) \in \mathscr{Y} \times \mathscr{Y} | \mathcal{K} \ni y \perp z \in \mathcal{K}\}.$$

Note that although $\mathcal{K}$ is convex, the complementarity set is not convex.

Denote by $G(x) := P(x) - \Lambda$ a set-valued map induced by the system $P(x) \in \Lambda$. An important stability issue to study is the metric subregularity. We say that the set-valued map G is metrically subregular at $(\bar{x}, 0) \in gphG$, where

$$gphG := \{(x, y)|y \in G(x)\}$$

is the graph of G, if there exist $\kappa \geq 0$ and a neighborhood U of $\bar{x}$ such that

$$d(x, G^{-1}(0)) \leq \kappa d(P(x), \Lambda) \quad \forall x \in U,$$

where $d(x, C)$ denotes the distance between a point x and a set C and $G^{-1}(y) := \{x|y \in G(x)\}$ denotes the inverse of G at y.

The concept of metric subregularity was introduced by Ioffe [24] using the terminology "regularity at a point." The terminology "metric subregularity" was suggested by Dontchev and Rockafellar in [4, Definition 3.1]. This property is also referred to as an error bound property since it enables us to estimate the distance from a point x near $\bar{x}$ to the set of solutions to the system (1) by its residue $d(P(x), \Lambda)$, which is much easier to deal with; see, e.g., [8, 51, 50, 52, 6, 39] and the references therein for related results and applications. Metric subregularity is a weaker condition than the more familiar property of metric regularity, which requires the existence of $\kappa \geq 0$ and U, V, neighborhoods of $\bar{x}, 0$, respectively, such that

$$d(x, G^{-1}(y)) \leq \kappa d(P(x), \Lambda) \quad \forall x \in U,\ y \in V,$$

and strong metric subregularity (see, e.g., [5]), which requires the existence of $\kappa \geq 0$ and U, a neighborhood of $\bar{x}$ such that

$$\|x - \bar{x}\| \leq \kappa d(P(x), \Lambda) \quad \forall x \in U.$$

It is well known (see, e.g., [4, Theorem 3.2]) that the metric subregularity of a set-valued map is equivalent to the calmness of its inverse map, which means that there exist $\kappa \geq 0$ and neighborhoods U of $\bar{x}$ and V of 0 such that

$$G^{-1}(y) \cap U \subseteq G^{-1}(0) + \kappa\|y\|\mathbb{B} \quad \forall y \in V,$$

where $\|\cdot\|$ and $\mathbb{B}$ denote the norm and the closed unit ball in $\mathscr{Y}$, respectively. The concept of calmness was first introduced by J. J. Ye and X. Y. Ye in [55, Definition 2.8] under a different name, "pseudo-upper-Lipschitz continuity," and the terminology of "calmness" was coined by Rockafellar and Wets in [46]. Note that the calmness property is part of the property required in the notion of pseudo-Lipschitz continuity introduced by Klatte [30]. As suggested by the name "pseudo-upper-Lipschitz continuity," the concept of calmness is weaker than both the pseudo-Lipschitz continuity (or Aubin continuity) introduced by Aubin [1] and the upper-Lipschitz continuity introduced by Robinson [42, 43, 44]. Analogous to the fact that a set-valued map is metrically subregular if and only if its inverse map is calm, it is well known that the metric regularity of a set-valued map is equivalent to the pseudo-Lipschitz continuity of its inverse map (see [35, Theorem 1.49]).

Metric subregularity/calmness plays an important role in optimization. It serves as a constraint qualification and a sufficient condition for exact penalty; see, e.g., [3, 24, 23, 25, 31, 48, 53, 55]. As pointed out in [26], metric subregularity/calmness is also an important tool in the subdifferential calculus of nonsmooth analysis. More recently, it has been discovered that it serves as a sufficient condition for linear convergence of certain numerical algorithms [32, 49] and quadratic convergence of the Newton-type method [7].

Although the metric subregularity/calmness/error bound condition is very important, it is by no means easy to verify. For a long time, there have been only two major *checkable* sufficient conditions: one is derived by Robinson's multifunction theory and the other is by Mordukhovich's criteria. By Robinson's multifunction theory [44], if the linear constraint qualification (linear CQ) holds, i.e., $P(x)$ is affine and Λ is the union of finitely many polyhedral convex sets, then the set-valued map $G(x) = P(x) - \Lambda$ must be a polyhedral multivalued function and so is its inverse map G^{-1}. Hence the set-valued map G^{-1} must be upper Lipschitz and hence calm. Recall that in optimization we call a multiplier abnormal if it is a multiplier corresponding to an optimality system where the objective function vanishes. Assuming P is continuously differentiable (C^1), if the no nonzero abnormal multiplier constraint qualification (NNAMCQ) holds, i.e., there is no nonzero abnormal multiplier ζ such that

$$0 = \nabla P(\bar{x})^* \zeta, \quad \zeta \in N_\Lambda(P(\bar{x})), \tag{2}$$

where $N_\Lambda(\cdot)$ is the limiting normal cone, ∇P denotes the Fréchet derivative of P, and * denotes the adjoint, then Mordukhovich's criteria for metric regularity (see, e.g., [46, Theorem 9.40]) holds and so does metric subregularity. These two criteria are relatively strong since they are actually sufficient conditions for stronger stability concepts. And therefore there are many situations where these sufficient conditions do not hold but the systems are still metrically subregular. In general metric subregularity is weaker than NNAMCQ but for the case of a differentiable convex inequality system, which is (1) with P convex and differentiable and Λ a nonnegative orthant, Li [33] has shown that all the following conditions are equivalent: metric subregularity, Abadie's constraint qualification, the Slater condition, and the Mangasarian–Fromovitz constraint qualification (MFCQ) (which is equivalent to NNAMCQ in this case).

Over the last fifteen years or so, some results for characterizing metric subregularity/calmness for general set-valued maps have been obtained; see, e.g., [19, 20, 21, 22, 59]. Recently the concept of a directional limiting normal cone which is in general a smaller set than the limiting normal cone was introduced [16, 10]. Based on the result for general set-valued maps in [10], Gfrerer and Klatte [14, Corollary 1] showed that metric subregularity holds for system (1) at $\bar{x}$ under the first-order sufficient condition for metric subregularity (FOSCMS): assuming $P(x)$ is C^1, if for each nonzero direction u satisfying $\nabla P(\bar{x})u \in T_\Lambda(P(\bar{x}))$ there is no nonzero ζ such that

$$0 = \nabla P(\bar{x})^* \zeta, \quad \zeta \in N_\Lambda(P(\bar{x}); \nabla P(\bar{x})u),$$

where $T_\Lambda(\cdot)$ and $N_\Lambda(y; d)$ are the tangent cone and the limiting normal cone at y in direction d (see Definition 2.2). Moreover, if $P(x)$ is strictly differentiable and twice directionally differentiable and Λ is the union of finitely many polyhedral convex sets, it was shown in [11, Theorem 4.3] that metric subregularity holds at $(\bar{x}, 0)$ under the following second-order sufficient condition for metric subregularity (SOSCMS): for each nonzero direction u satisfying $\nabla P(\bar{x})u \in T_\Lambda(P(\bar{x}))$, there exists no $\zeta \neq 0$ such that

$$\begin{cases} 0 = \nabla P(\bar{x})^*\zeta, \quad \zeta \in N_\Lambda(P(\bar{x}); \nabla P(\bar{x})u), \\ \langle \zeta, P''(\bar{x}; u)\rangle \geq 0, \end{cases}$$

where $P''(\bar{x}; u)$ denotes the second-order derivative of $P(x)$ at $\bar{x}$ in the direction u. Some sufficient conditions for the metric subregularity/calmness/error bound condition for special complementarity systems based on the FOSCMS have been obtained in [15, 58].

Another direction in the effort of weakening the NNAMCQ is to add some extra conditions to (2). In the case where P is continuously differentiable at $\bar{x}$, we say that quasi-normality and pseudo-normality hold at $\bar{x}$ if there exists no $\zeta \neq 0$ such that (2) holds and

$$\exists (x^k, s^k, \zeta^k) \to (\bar{x}, P(\bar{x}), \zeta) \text{ s.t. } \zeta^k \in N_\Lambda(s^k) \text{ and } \zeta_i(P_i(x^k) - s_i^k) > 0 \text{ if } \zeta_i \neq 0,$$
$$\exists (x^k, s^k, \zeta^k) \to (\bar{x}, P(\bar{x}), \zeta) \text{ s.t. } \zeta^k \in N_\Lambda(s^k) \text{ and } \langle \zeta, P(x^k) - s^k\rangle > 0,$$

respectively. It is obvious that pseudo-normality implies quasi-normality. For a system with equality and inequality constraints where all constraint functions are C^{1+}, which means that the gradients are locally Lipschitz, Minchenko and Tarakanov [34, Theorem 2.1] showed that quasi-normality implies the existence of a local error bound or equivalently metric subregularity/calmness at $\bar{x}$. In [56, Theorem 5], this result is extended to systems with continuously differentiable equality constraint functions and subdifferentially regular inequality constraint functions and a regular constraint set. Quasi-normality/pseudo-normality for the general system in the form (1) was introduced by Guo, Ye, and Zhang [17, Definition 4.2] and proved to be a sufficient condition for error bound/metric subregularity/calmness in [17, Theorem 5.2] under the Lipschitz continuity of P and the closeness of the set Λ only.

The main purpose of this paper is to combine the two approaches of weakening the NNAMCQ, i.e., to replace the limiting normal cone by the directional normal cone as in FOSCMS and SOSCMS and to add extra conditions as in quasi-/pseudo-normality and prove that our weaker sufficient conditions are still sufficient for verifying metric subregularity/calmness.

Our assumptions are very general. We only assume the continuity of the mapping $P(x)$. Indeed, it is natural to study the case where $P(x)$ is only continuous, since it will widen the range of applications of formation (1). For example, consider the recovery of an unknown vector $x \in \mathbb{R}^n$ (such as a signal or an image) from noisy data $b \in \mathbb{R}^m$ by minimizing with respect to x a regularized cost function

$$F(x, b) = f(x, b) + \mu g(x), \tag{3}$$

where typically $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is a data-fidelity term and $g : \mathbb{R}^n \to \mathbb{R}$ is a nonsmooth regularization term, with $\mu > 0$ a parameter. One usual choice for the data-fidelity term is

$$f(x, b) = \sum_{i=1}^m |a_i^T x - b_i|^\rho$$

with $a_i \in \mathbb{R}^n$ and ρ in the range $(0, \infty]$; see, e.g., [38, 40, 41]. Apparently when ρ takes a value in the interval $(1, 2)$, the optimality condition of minimizing function (3) with respect to x can be described by $0 \in \nabla_x f(x,b) + \partial g(x)$, where $\partial g(x)$ denotes a certain subdifferential of g at x, which can be reformulated as $P(x) \in \Lambda$, where $P(x)$ is continuous. Therefore, thanks to the equivalence between calmness of different reformulations established in [15, Proposition 3], our results can be used to study the calmness of the optimality condition system of minimizing (3) without imposing an unnecessarily stronger condition.

We organize our paper as follows. Section 2 gives the preliminaries and preliminary results. In section 3, we propose the weak sufficient condition for metric subregularity and show that it is sufficient for metric subregularity. In section 4, we propose the concepts of directional quasi-/pseudo-normality and show that they are stronger than the new sufficient condition for metric subregularity. Moreover, in this section it is shown that the SOSCMS implies pseudo-normality. In section 5 we apply our results to complementarity systems and Karush–Kuhn–Tucker (KKT) systems.

2. Preliminaries and preliminary results.

In this section, we gather some preliminaries on variational analysis and nonsmooth analysis that will be used in the following sections. We only give concise definitions and results that will be needed in this paper. For more detailed information on the subject, the reader is referred to Mordukhovich [35], and Rockafellar and Wets [46].

First, we give the definition of tangent cones and normal cones.

DEFINITION 2.1 (tangent cones and normal cones; see, e.g., [46, Definition 6.1]). *Given a set $\Omega \subseteq \mathscr{Y}$ and a point $\bar{y} \in \Omega$, the tangent cone to Ω at $\bar{y}$ is defined as*

$$T_\Omega(\bar{y}) := \{d \in \mathscr{Y} | \exists t_k \downarrow 0,\ d_k \to d \ s.t.\ \bar{y} + t_k d_k \in \Omega\ \forall k\}.$$

The derivable cone to Ω at $\bar{y}$ is defined as

$$T^i_\Omega(\bar{y}) := \{d \in \mathscr{Y} | \forall t_k \downarrow 0, \exists d_k \to d \ s.t.\ \bar{y} + t_k d_k \in \Omega\ \forall k\}.$$

A set Ω is said to be geometrically derivable if the tangent cone coincides with the derivable cone at each point of Ω, or equivalently if $\lim_{t\downarrow 0} t^{-1} d(\bar{y} + tu, \Omega) = 0$.

The regular normal cone and the limiting normal cone to Ω at $\bar{y}$ are defined as

$$\widehat{N}_\Omega(\bar{y}) := \left\{ \zeta \in \mathscr{Y} \middle| \limsup_{y \xrightarrow{\Omega} \bar{y}} \frac{\langle \zeta, y - \bar{y}\rangle}{\|y - \bar{y}\|} \le 0 \right\}$$

and

$$N_\Omega(\bar{y}) := \left\{ \zeta \in \mathscr{Y} \middle| \exists\ y_k \xrightarrow{\Omega} \bar{y},\ \zeta_k \to \zeta \ such\ that\ \zeta_k \in \widehat{N}_\Omega(y_k)\ \forall k \right\},$$

respectively, where $y_k \xrightarrow{\Omega} \bar{y}$ means $y_k \to \bar{y}$ and for each k, $y_k \in \Omega$.

Recently a directional version of limiting normal cones was introduced in [16, Definition 2.3] and extended to general Banach spaces in [10].

DEFINITION 2.2 (directional normal cones; see [10, Definition 2]). *Given a point $\bar{y} \in \mathscr{Y}$ and a direction $d \in \mathscr{Y}$, the limiting normal cone to Ω at $\bar{y}$ in direction d is defined by*

$$N_\Omega(\bar{y}; d) := \left\{ \zeta \in \mathscr{Y} \middle| \exists\ t_k \downarrow 0,\ d_k \to d,\ \zeta_k \to \zeta \ s.t.\ \zeta_k \in \widehat{N}_\Omega(\bar{y} + t_k d_k)\ \forall k \right\}.$$

From the definition, it is obvious that $N_\Omega(\bar{y}; d) = \emptyset$ if $d \notin T_\Omega(\bar{y})$ and $N_\Omega(\bar{y}; d) \subseteq N_\Omega(\bar{y})$.

Proposition 2.1 (see [58, Proposition 3.3]). *Let $\Omega := \Omega_1 \times \cdots \times \Omega_l$, where $\Omega_i \subseteq \mathbb{R}^{n_i}$ are closed for $i = 1, \ldots, l$ and $n = n_1 + \cdots + n_l$. Consider a point $\bar{y} = (\bar{y}_1, \ldots, \bar{y}_l) \in \Omega$ and a direction $d = (d_1, \ldots, d_l) \in \mathbb{R}^n$. Then*

$$\begin{aligned} T_\Omega(\bar{y}) &\subseteq T_{\Omega_1}(\bar{y}_1) \times \cdots \times T_{\Omega_l}(\bar{y}_l), \\ N_\Omega(\bar{y}; d) &\subseteq N_{\Omega_1}(\bar{y}_1; d_1) \times \cdots \times N_{\Omega_l}(\bar{y}_l; d_l). \end{aligned}$$

The equality holds if all except at most one of Ω_i for $i = 1, \ldots, l$ are directionally regular at y_i in the sense of [58, Definition 3.3].

We give the definition of some subdifferentials below.

Definition 2.3 (subdifferentials; see, e.g., [35]). *Let $f : \mathscr{X} \to [-\infty, +\infty]$ and $\bar{x}$ is a point where f is finite. Then*

- *the Fréchet (regular) subdifferential of f at $\bar{x}$ is the set*

$$\widehat{\partial} f(\bar{x}) := \left\{ \xi \in \mathscr{X} \,\middle|\, \liminf_{h \to 0} \frac{f(\bar{x} + h) - f(\bar{x}) - \langle \xi, h \rangle}{\|h\|} \geq 0 \right\},$$

- *the limiting (Mordukhovich or basic) subdifferential of f at $\bar{x}$ is the set*

$$\partial f(\bar{x}) := \left\{ \xi \in \mathscr{X} \,\middle|\, \exists x_k \to \bar{x},\ f(x_k) \to f(\bar{x}),\ \text{and } \xi_k \to \xi \text{ with } \xi_k \in \widehat{\partial} f(x_k) \right\}.$$

Recently, based on the concept of the directional limiting normal cone, the following directional version of the limiting subdifferential was introduced in [2].

Definition 2.4 (directional subdifferentials; see [2]). *Let $f : \mathcal{X} \to [-\infty, +\infty]$ and $\bar{x}$ be a point where f is finite. Then the limiting subdifferential of f at $\bar{x}$ in direction $(u, \zeta) \in \mathscr{X} \times \mathbb{R}$ is defined as*

$$\begin{aligned} \partial f(\bar{x}; (u, \zeta)) := \Big\{ \xi \in \mathscr{X} \,\Big|\, & \exists t_k \downarrow 0,\ u^k \to u,\ \zeta^k \to \zeta,\ \xi^k \to \xi, \\ & f(\bar{x}) + t_k \zeta^k = f(\bar{x} + t_k u^k),\ \xi^k \in \widehat{\partial} f(\bar{x} + t_k u^k) \Big\}. \end{aligned}$$

Remark 2.1. Let f be continuously differentiable at $\bar{x}$. Then $\partial f(\bar{x}; (u, \zeta)) \neq \emptyset$ if and only if $\zeta = \nabla f(\bar{x}) u$, in which case

$$\partial f(\bar{x}; (u, \zeta)) = \partial f(\bar{x}) = \{\nabla f(\bar{x})\}.$$

Definition 2.5 (graphical derivatives; see, e.g., [5]). *For a set-valued map $G : \mathscr{X} \rightrightarrows \mathscr{Y}$ and a pair (x, y) with $y \in G(x)$, the graphical derivative of G at x for y is the set-valued map $DG(x|y) : \mathscr{X} \rightrightarrows \mathscr{Y}$ whose graph is the tangent cone to gphG at (x, y):*

$$v \in DG(x|y)(u) \Leftrightarrow (u, v) \in T_{gphG}(x, y).$$

Thus, $v \in DG(x|y)(u)$ if and only if there exist sequences $u_k \to u$, $v_k \to v$, and $\tau_k \downarrow 0$ such that $y + \tau_k v_k \in G(x + \tau_k u_k)$ for all k.

For a single-valued mapping $P : \mathscr{X} \to \mathscr{Y}$, its graphical derivative at x for $y = P(x)$ is

$$DP(x)(u) := \left\{ \xi \,\middle|\, \exists t_k \downarrow 0, u_k \to u \text{ s.t. } \lim_{k \to +\infty} \frac{P(x + t_k u_k) - P(x)}{t_k} = \xi \right\}. \tag{4}$$

Moreover if $P(x)$ is Hadamard directionally differentiable at x, then its graphical derivative is equal to the directional derivative: for any $u \in \mathscr{X}$,

$$DP(x)(u) = P'(x; u) := \lim_{t\downarrow 0,\, u'\to u} \frac{P(x+tu') - P(x)}{t}.$$

The following sum rule extends the sum rule in [5, Proposition 4A.2] by allowing $P(x)$ to be only continuous.

PROPOSITION 2.2. *Let $G(x) := -P(x) + \Lambda$ and $P(\bar{x}) \in \Lambda$, where $P(x) : \mathscr{X} \to \mathscr{Y}$ is a continuous singled-valued map.*

Then either

$$DG(\bar{x}|0)(u) \subseteq -DP(\bar{x})(u) + T_\Lambda(P(\bar{x})) \tag{5}$$

or there exists $\zeta \neq 0$ such that

$$\zeta \in DP(\bar{x})(0) \cap T_\Lambda(P(\bar{x})).$$

If either $P(x)$ is Hadamard directionally differentiable at $\bar{x}$ or Λ is geometrically derivable, then (5) *holds as an equality.*

Proof. By definition, $v \in DG(\bar{x}|0)(u)$ if and only if $(u, v) \in T_{gphG}(\bar{x}, 0)$. It follows from the definition of the tangent cone that there exist sequences $(u_k, v_k) \to (u, v)$ and $\tau_k \downarrow 0$ such that $(\bar{x}, 0) + \tau_k(u_k, v_k) \in gphG$, which means that there exists $s_k \in \Lambda$ such that $\tau_k v_k = -P(\bar{x} + \tau_k u_k) + s_k$.

Case (i) ($\{\frac{P(\bar{x}+\tau_k u_k)-P(\bar{x})}{\tau_k}\}$ is bounded). Then without loss of generality we may assume that $\lim_{k\to+\infty} \frac{P(\bar{x}+\tau_k u_k)-P(\bar{x})}{\tau_k} = \xi$. Therefore we have

$$v = \lim_{k\to+\infty} v_k = -\lim_{k\to+\infty} \frac{P(\bar{x}+\tau_k u_k) - P(\bar{x})}{\tau_k} + \lim_{k\to+\infty} \frac{s_k - P(\bar{x})}{\tau_k}.$$

Since $s_k \in \Lambda$, we have

$$\lim_{k\to+\infty} \frac{s_k - P(\bar{x})}{\tau_k} \in T_\Lambda(P(\bar{x})).$$

Hence $v \in -DP(\bar{x})(u) + T_\Lambda(P(\bar{x}))$.

Case (ii) ($\{\frac{P(\bar{x}+\tau_k u_k)-P(\bar{x})}{\tau_k}\}$ is unbounded). Without loss of generality, assume that

$$\lim_{k\to+\infty} \frac{\|P(\bar{x}+\tau_k u_k) - P(\bar{x})\|}{\tau_k} = \infty.$$

Define $t_k := \|P(\bar{x} + \tau_k u_k) - P(\bar{x})\|$.

Since

$$\left\{\frac{P(\bar{x}+\tau_k u_k) - P(\bar{x})}{t_k}\right\} = \left\{\frac{P(\bar{x}+\tau_k u_k) - P(\bar{x})}{\|P(\bar{x}+\tau_k u_k) - P(\bar{x})\|}\right\}$$

is bounded, we may without loss of generality assume $\lim_{k\to+\infty}\{\frac{P(\bar{x}+\tau_k u_k)-P(\bar{x})}{t_k}\} = \zeta$. By definition of $DP(\bar{x})(0)$ and the fact that $\lim_{k\to\infty} \frac{\tau_k}{t_k} = 0$, we have

$$0 \neq \zeta = \lim_{k\to+\infty} \frac{P(\bar{x}+\tau_k u_k) - P(\bar{x})}{t_k} = \lim_{k\to+\infty} \frac{P(\bar{x}+t_k(\frac{\tau_k}{t_k}u_k)) - P(\bar{x})}{t_k} \in DP(\bar{x})(0).$$

Since $v_k \to v$ and $\lim_{k\to\infty} \frac{\tau_k}{t_k} = 0$, we have

$$0 = \lim_{k\to\infty} \frac{\tau_k}{t_k} v_k = -\lim_{k\to+\infty} \frac{P(\bar{x}+\tau_k u_k) - P(\bar{x})}{t_k} + \lim_{k\to+\infty} \frac{s_k - P(\bar{x})}{t_k}$$
$$= -\zeta + \lim_{k\to+\infty} \frac{s_k - P(\bar{x})}{t_k}.$$

Therefore $\zeta = \lim_{k\to+\infty} \frac{s_k - P(\bar{x})}{t_k}$, which implies that $\zeta \in T_\Lambda(P(\bar{x}))$.

Conversely, let $v \in -DP(\bar{x})(u) + T_\Lambda(P(\bar{x}))$. Then there exist $\xi \in DP(\bar{x})(u)$ and $\zeta \in T_\Lambda(P(\bar{x}))$ such that $v = -\xi + \zeta$.

If $P(x)$ is Hadamard directionally differentiable at $\bar{x}$, then the limit

$$\xi = \lim_{t\downarrow 0,\, u'\to u} \frac{P(\bar{x}+tu') - P(\bar{x})}{t}$$

exists and there exist sequences $\tau_k \downarrow 0$, $s_k \xrightarrow{\Lambda} P(\bar{x})$ such that

$$\zeta = \lim_{k\to+\infty} \frac{s_k - P(\bar{x})}{\tau_k}.$$

Define

$$v_k = -\frac{P(\bar{x}+\tau_k u_k) - P(\bar{x})}{\tau_k} + \frac{s_k - P(\bar{x})}{\tau_k} = \frac{-P(\bar{x}+\tau_k u_k) + s_k}{\tau_k}.$$

Then $\lim_{k\to\infty} v_k = v$ and $\tau_k v_k \in -P(\bar{x}+\tau_k u_k) + \Lambda$ for all k. Hence $v \in DG(\bar{x}|0)(u)$.

Now suppose that Λ is geometrically derivable. let $\tau_k \downarrow 0$, $u_k \to u$ be sequences such that

$$\xi = \lim_{k\downarrow\infty} \frac{P(\bar{x}+\tau_k u_k) - P(\bar{x})}{\tau_k}.$$

Since Λ is geometrically derivable, there exists $s_k \in \Lambda$ such that

$$\zeta = \lim_{k\to+\infty} \frac{s_k - P(\bar{x})}{\tau_k}.$$

Define

$$v_k = -\frac{P(\bar{x}+\tau_k u_k) - P(\bar{x})}{\tau_k} + \frac{s_k - P(\bar{x})}{\tau_k} = \frac{-P(\bar{x}+\tau_k u_k) + s_k}{\tau_k}.$$

Then $\lim_{k\to\infty} v_k = v$ and $\tau_k v_k \in G(\bar{x}+\tau_k u_k)$ for all k. Hence $v \in DG(\bar{x}|0)(u)$. ∎

DEFINITION 2.6 (coderivatives and directional coderivatives; see [2] and [35, Definition 1.32]). *For a set-valued map* $G : \mathscr{X} \rightrightarrows \mathscr{Y}$ *and a point* $(\bar{x},\bar{y}) \in gphG := \{(x,y) \in \mathscr{X}\times\mathscr{Y} | y \in G(x)\}$, *the Fréchet coderivative (precoderivative) of* G *at* $(\bar{x},\bar{y})$ *is a multifunction* $\widehat{D}^*G(\bar{x},\bar{y}) : \mathscr{Y} \rightrightarrows \mathscr{X}$ *defined as*

$$\widehat{D}^*G(\bar{x},\bar{y})(\zeta) := \left\{\eta \in \mathscr{X} \middle| (\eta,-\zeta) \in \widehat{N}_{gphG}(\bar{x},\bar{y})\right\};$$

the limiting (Mordukhovich) coderivative of G *at* $(\bar{x},\bar{y})$ *is a multifunction* $D^*G(\bar{x},\bar{y}) : \mathscr{Y} \rightrightarrows \mathscr{X}$ *defined as*

$$D^*G(\bar{x},\bar{y})(\zeta) := \{\eta \in \mathscr{X} | (\eta,-\zeta) \in N_{gphG}(\bar{x},\bar{y})\}.$$

The symbol $D^*G(\bar{x})$ is used when G is single valued. The limiting coderivative of G at $(\bar{x}, \bar{y})$ in direction $(u, \xi) \in \mathscr{X} \times \mathscr{Y}$ is defined as

$$D^*G(\bar{x}, \bar{y}; (u, \xi))(\zeta) := \{\eta \in \mathscr{X} \,|\, (\eta, -\zeta) \in N_{gphG}(\bar{x}, \bar{y}; (u, \xi))\}.$$

Similarly the symbol $D^*G(\bar{x}; (u, \xi))$ is used when G is single valued.

Remark 2.2. In the special case when $P : \mathscr{X} \to \mathscr{Y}$ is a single-valued map which is Lipschitz continuous at $\bar{x}$, by [35, Theorem 3.28], the coderivative is related to the limiting subdifferential in the following way:

$$D^*P(\bar{x})(\zeta) = \partial\langle P, \zeta\rangle(\bar{x}) \quad \text{for all } \zeta \in \mathscr{Y}.$$

By [2, Proposition 5.1], if P is Lipschitz near $\bar{x}$ in direction u, then $D^*P(\bar{x}; (u, \xi))(\zeta) \neq \emptyset$ if and only if $\xi \in DP(\bar{x})(u)$, in which case

$$D^*P(\bar{x}; (u, \xi))(\zeta) = \partial\langle P, \zeta\rangle(\bar{x}; (u, \langle \xi, \zeta\rangle)).$$

Let $P : \mathscr{X} \to \mathscr{Y}$ be C^1. By [2, Remark 2.1], one has $DP(\bar{x})(u) = \nabla P(\bar{x})u$ and thus $D^*P(\bar{x}; (u, \xi))(\zeta) \neq \emptyset$ if and only if $\xi = \nabla P(\bar{x})u$, in which case

$$D^*P(\bar{x}; (u, \xi))(\zeta) = D^*P(\bar{x})(\zeta) = \nabla P(\bar{x})^*\zeta.$$

To state our main results, given $P : \mathscr{X} \to \mathscr{Y}$ and $\Lambda \subseteq \mathscr{Y}$, we define the extended linearization cone as

$$\widetilde{\mathcal{L}}(x) := \{(u, \xi) \in \mathscr{X} \times \mathscr{Y} \,|\, \xi \in DP(x)(u) \cap T_\Lambda(P(x))\}. \tag{6}$$

It is easy to see that the projection of $\widetilde{\mathcal{L}}(x)$ onto the space $\mathscr{X}$ is the linearization cone defined by $\mathcal{L}(x) := \{u \in \mathscr{X} \,|\, \exists \xi \text{ such that } \xi \in DP(x)(u) \cap T_\Lambda(P(x))\}$. When P is differentiable at x, $DP(x)(u) = \nabla P(x)u$ and hence in this case

$$\widetilde{\mathcal{L}}(x) = \{(u, \nabla P(x)u) : 0 \in -\nabla P(x)u + T_\Lambda(P(x))\}.$$

PROPOSITION 2.3. *Let $P : \mathscr{X} \to \mathscr{Y}$ be continuous and $\Lambda \subseteq \mathscr{Y}$. Then*

$$\widetilde{\mathcal{L}}(\bar{x}) = \{(0, 0)\} \implies DG(\bar{x}|0)^{-1}(0) = \{0\}. \tag{7}$$

Proof. By virtue of Proposition 2.2, when $\widetilde{\mathcal{L}}(\bar{x}) = \{(0, 0)\}$, one must have

$$DG(\bar{x}|0)(u) \subseteq -DP(\bar{x})(u) + T_\Lambda(P(\bar{x})).$$

Suppose that $u \in DG(\bar{x}|0)^{-1}(0)$. Then equivalently, $0 \in DG(\bar{x}|0)(u)$. Hence $0 \in -DP(\bar{x})(u) + T_\Lambda(P(\bar{x}))$ or equivalently $DP(x)(u) \cap T_\Lambda(P(x)) \neq \emptyset$. Since $\widetilde{\mathcal{L}}(\bar{x}) = \{(0, 0)\}$, it means that $\forall u \neq 0$, $DP(x)(u) \cap T_\Lambda(P(x)) = \emptyset$. Hence we must have $u = 0$. ◻

PROPOSITION 2.4. *Let $P : \mathscr{X} \to \mathscr{Y}$ be continuous and $\Lambda \subseteq \mathscr{Y}$ be closed near $\bar{x} \in \mathscr{X}$. If $\widetilde{\mathcal{L}}(\bar{x}) = \{(0, 0)\}$, then $G(x) = P(x) - \Lambda$ is strongly metrically subregular at $(\bar{x}, 0)$.*

Proof. By [5, Theorem 4C.1], G is strongly metrically subregular at $(\bar{x}, 0)$ if and only if $DG(\bar{x}|0)^{-1}(0) = \{0\}$. The result then follows from applying Proposition 2.3. ◻

3. Weak sufficient condition for metric subregularity. In this section we will derive a sufficient condition for metric subregularity of the system $P(x) \in \Lambda$, where $P(x) : \mathscr{X} \to \mathscr{Y}$ is a continuous single-valued map and $\Lambda \subseteq \mathscr{Y}$ is locally closed. Recall that no ζ satisfying condition (10) alone is the so-called first-order sufficient condition for metric subregularity (FOSCMS), as established by Gfrerer and Klatte in [14, Corollary 1] for the case where P is smooth and extended to the nonsmooth but calmness case in [2, Proposition 2.2]. Our sufficient condition in Theorem 3.1 improves the FOSCMS in [2, Proposition 2.2] in two aspects. First, we allow $P(x)$ to be only continuous instead of being calm. Secondly, even in the case where $P(x)$ is calm, our condition is weaker in that the extra condition of the existence of sequences $(u_k, v_k, \zeta_k) \to (u, 0, \zeta)$ and $t_k \downarrow 0$ satisfying (11) and (12) is required.

We will derive our result based on the following sufficient conditions for metric subregularity for general set-valued maps by Gfrerer in [12].

LEMMA 3.1 (see [12, Corollary 1 and Remarks 1 and 2]). *Let $G : \mathscr{X} \rightrightarrows \mathscr{Y}$ be a closed set-valued map, and take a point $(\bar{x}, \bar{y}) \in gphG$. Assume that for any direction $u \in \mathscr{X}$, there do not exist sequences $t_k \downarrow 0$, $\|(u_k, v_k)\| = 1$, $\|y_k^*\| = 1$ with $\|u_k\| \to 1$, $\|u\| u_k \to u$, $v_k \to 0$, $x_k^* \to 0$ satisfying*

$$(x_k^*, -y_k^*) \in \widehat{N}_{gphG}(x_k', y_k'), \quad x_k' \notin G^{-1}(\bar{y}),$$

and

$$\lim_{k\to\infty} \frac{\langle y_k^*, y_k' - \bar{y}\rangle}{\|y_k' - \bar{y}\|} = 1,$$

where $x_k' := \bar{x} + t_k u_k \neq \bar{x}$, $y_k' := \bar{y} + t_k v_k \neq \bar{y}$. Then G is metrically subregular at $(\bar{x}, \bar{y})$.

Note that as commented in [12, Remark 2], if the condition $x_k' \notin G^{-1}(\bar{y})$ is omitted then the resulting sufficient condition is stronger but may be easier to verify. However, in [37, Example 1], it was shown that sometimes these kinds of conditions can not be omitted in order to show the metric subregularity.

LEMMA 3.2. *Let P be a single-valued map from $\mathscr{X}$ to $\mathscr{Y}$ and Λ be a subset of $\mathscr{Y}$. Define $G(x) := P(x) - \Lambda$, $y = P(x) - s$ for some $s \in \Lambda$. Then $(x^*, -y^*) \in \widehat{N}_{gphG}(x, y)$ implies that*

$$x^* \in \widehat{D}^* P(x)(y^*), \quad y^* \in \widehat{N}_\Lambda(P(x) - y).$$

Proof. Since $(x^*, -y^*) \in \widehat{N}_{gphG}(x, y)$, by definition for any $\epsilon > 0$,

$$\langle x^*, x' - x\rangle + \langle -y^*, y' - y\rangle \leq \epsilon \|(x' - x, y' - y)\| \tag{8}$$

for any $(x', y') \in gphG$ which is sufficiently close to (x, y). Let $y' := P(x) - s'$, $s' \in \Lambda$. Then when s' is close to s, $y' = P(x) - s'$ is close to $y = P(x) - s$. Hence, fixing $x' = x$ in (8) we obtain that for any $\epsilon > 0$ and any $s' \in \Lambda$ sufficiently close to s,

$$\langle -y^*, s - s'\rangle \leq \epsilon \|s - s'\| \Leftrightarrow \langle y^*, s' - s\rangle \leq \epsilon \|s - s'\|.$$

This means that $y^* \in \widehat{N}_\Lambda(s) = \widehat{N}_\Lambda(P(x) - y)$.

On the other hand, let $x' \in \mathscr{X}$ and $y' := P(x') - s$. Then $y' \in G(x')$ and when $(x', P(x'))$ is close to $(x, P(x))$, (x', y') is close to (x, y). Hence, by (8) we have

$$\langle x^*, x' - x\rangle + \langle -y^*, P(x') - P(x)\rangle \leq \epsilon \|(x' - x, P(x') - P(x))\|$$

for any $(x', P(x'))$ which is close to $(x, P(x))$. This means that

$$(x^*, -y^*) \in \widehat{N}_{gphP}(x, P(x))$$

or equivalently $x^* \in \widehat{D}^* P(x)(y^*)$. The proof of the lemma is therefore complete. ◻

Applying Lemmas 3.1 and 3.2, we obtain the following sufficient condition for metric subregularity.

PROPOSITION 3.1. *Let $P : \mathscr{X} \to \mathscr{Y}$ be a single-valued map and $\Lambda \subseteq \mathscr{Y}$ be closed. Let $G(x) := P(x) - \Lambda$ and $P(\bar{x}) \in \Lambda$. Assume that $G(x)$ is a set-valued map which is closed around $\bar{x}$ and suppose that for any direction $u \in \mathscr{X}$, there do not exist sequences $t_k \downarrow 0$, $\|(u_k, v_k)\| = 1$, $\|y_k^*\| = 1$ with $\|u_k\| \to 1$, $\|u\|u_k \to u$, $v_k \to 0$, $x_k^* \to 0$ satisfying*

$$x_k^* \in \widehat{D}^* P(\bar{x} + t_k u_k)(y_k^*), \quad y_k^* \in \widehat{N}_\Lambda(P(\bar{x} + t_k u_k) - t_k v_k), \quad P(\bar{x} + t_k u_k) \notin \Lambda,$$

and

$$\lim_{k\to\infty} \frac{\langle y_k^*, v_k\rangle}{\|v_k\|} = 1. \tag{9}$$

Then G is metrically subregular at $(\bar{x}, 0)$.

Note that by [35, Theorem 1.38], when P is Fréchet differentiable but not necessarily Lipschitz continuous, we have $\widehat{D}^* P(x)(y^*) = \{\nabla P(x)^* y^*\}$.

THEOREM 3.1. *Let $P : \mathscr{X} \to \mathscr{Y}$ be continuous and $\Lambda \subseteq \mathscr{Y}$ be closed at $\bar{x} \in \mathscr{X}$. Suppose that the weak sufficient condition for metric subregularity (WSCMS) holds at $\bar{x}$, i.e., for all $(0, 0) \neq (u, \xi) \in \widetilde{\mathcal{L}}(\bar{x})$, there exists no unit vector ζ, sequences $(u_k, v_k, \zeta_k) \to (u, 0, \zeta)$ and $t_k \downarrow 0$ satisfying*

$$0 \in D^* P(\bar{x}; (u, \xi))(\zeta), \quad \zeta \in N_\Lambda(P(\bar{x}); \xi), \tag{10}$$

$$\zeta_k \in \widehat{N}_\Lambda(s_k),\ s_k = P(\bar{x} + t_k u_k) - t_k v_k,\ P(\bar{x} + t_k u_k) \notin \Lambda, \tag{11}$$

$$\lim_{k\to\infty} \left\langle \zeta_k, \frac{v_k}{\|v_k\|} \right\rangle = 1. \tag{12}$$

Then $G(x) = P(x) - \Lambda$ is metrically subregular at $(\bar{x}, 0)$.

Proof. If $\widetilde{\mathcal{L}}(\bar{x}) = \{(0, 0)\}$, then by Proposition 2.4, G is strongly metrically subregular and hence metrically subregular at $(\bar{x}, 0)$. We now prove the result for the $\widetilde{\mathcal{L}}(\bar{x}) \neq \{(0, 0)\}$ case by contradiction. To the contrary, suppose that $P(x) - \Lambda$ is not metrically subregular at $(\bar{x}, 0)$. By Proposition 3.1, there exist $u \in \mathscr{X}$ and sequences $t_k \downarrow 0$, $\|(u_k, v_k)\| = 1$, $\|y_k^*\| = 1$ with $\|u_k\| \to 1$, $\|u\|u_k \to u$, $v_k \to 0$, $x_k^* \to 0$ such that

$$(x_k^*, -y_k^*) \in \widehat{N}_{gphP}(\bar{x} + t_k u_k, P(\bar{x} + t_k u_k)), \quad y_k^* \in \widehat{N}_\Lambda(P(\bar{x} + t_k u_k) - t_k v_k), \tag{13}$$

and (9) holds.

Since we have $\|y_k^*\| = 1$, $\|(u_k, v_k)\| = 1$, and $v_k \to 0$, passing to a subsequence if necessary, we assume that $\lim_{k\to\infty} y_k^* = \zeta$, $\lim_{k\to\infty} u_k = u$ for certain $\|u\| = 1$. It follows that $\|\zeta\| = 1$.

Case (1) ($\{\frac{P(\bar{x}+t_k u_k)-P(\bar{x})}{t_k}\}$ is bounded). Then without loss of generality we may assume that $\lim_{k\to+\infty} \frac{P(\bar{x}+t_k u_k)-P(\bar{x})}{t_k} = \xi$. Thus, letting $\xi_k := \frac{P(\bar{x}+t_k u_k)-P(\bar{x})}{t_k}$, we have $P(\bar{x} + t_k u_k) = P(\bar{x}) + t_k \xi_k$. Combining with (13) we get

$$(x_k^*, -y_k^*) \in \widehat{N}_{gphP}((\bar{x}, P(\bar{x})) + t_k(u_k, \xi_k)), \quad y_k^* \in \widehat{N}_\Lambda(P(\bar{x} + t_k u_k) - t_k v_k).$$

Since $(u_k,\xi_k)\to(u,\xi)$ as $k\to\infty$, we have

$$(0,-\zeta)\in N_{gphP}((\bar{x},P(\bar{x}));(u,\xi)),\quad \zeta\in N_\Lambda(P(\bar{x});\xi).$$

Also, from the proof of Proposition 2.2, we see that $\xi\in DP(\bar{x})(u)\cap T_\Lambda(P(\bar{x}))$ and hence $(u,\xi)\in\tilde{\mathcal{L}}(\bar{x})$.

In summary for Case (1), we have obtained a nonzero vector ζ, a nonzero vector $(u,\xi)\in\widetilde{L}(\bar{x})$, and sequences $(u_k,v_k,y_k^*)\to(u,0,\zeta)$ and $t_k\downarrow 0$ such that

$$\begin{aligned}&0\in D^*P(\bar{x};(u,\xi))(\zeta),\quad \zeta\in N_\Lambda(P(\bar{x});\xi),\\&y_k^*\in\widehat{N}_\Lambda(s_k),\quad s_k=P(\bar{x}+t_ku_k)-t_kv_k,\\&\lim_{k\to\infty}\left\langle y_k^*,\frac{v_k}{\|v_k\|}\right\rangle=1,\end{aligned}$$

which contradicts the assumption in (WSCMS). Thus $P(x)-\Lambda$ is metrically subregular at $(\bar{x},0)$.

Case (2) ($\{\frac{P(\bar{x}+t_ku_k)-P(\bar{x})}{t_k}\}$ is unbounded). Without loss of generality, assume that $\lim_{k\to+\infty}\frac{\|P(\bar{x}+t_ku_k)-P(\bar{x})\|}{t_k}=\infty$. Define

$$\begin{aligned}\tau_k&:=\|t_ku_k\|+\|P(\bar{x}+t_ku_k)-P(\bar{x})\|,\quad &u_k'&:=\frac{t_ku_k}{\tau_k},\\ \xi_k&:=\frac{P(\bar{x}+\tau_ku_k')-P(\bar{x})}{\tau_k},&v_k'&:=\frac{t_kv_k}{\tau_k}.\end{aligned}$$

Since $t_k/\tau_k\le t_k/\|P(\bar{x}+t_ku_k)-P(\bar{x})\|$, we have $t_k/\tau_k\to 0$ and hence $v_k'\to 0$ and $u_k'\to 0$. Since $\{\xi_k\}$ is bounded, taking a subsequence if necessary, we have

$$\xi:=\lim_{k\to\infty}\xi_k.$$

Then with $t_ku_k=\tau_ku_k'$ and $P(\bar{x}+t_ku_k)=P(\bar{x})+\tau_k\xi_k$, by (13) we get

$$(x_k^*,-y_k^*)\in\widehat{N}_{gphP}((\bar{x},P(\bar{x}))+\tau_k(u_k',\xi_k)),\quad y_k^*\in\widehat{N}_\Lambda(P(\bar{x}+\tau_ku_k')-\tau_kv_k').$$

Since $s_k=P(\bar{x}+\tau_ku_k')-\tau_kv_k'$, we know that

$$\begin{aligned}\lim_{k\to\infty}\frac{s_k-P(\bar{x})}{\tau_k}&=\lim_{k\to\infty}\frac{P(\bar{x}+\tau_ku_k')-\tau_kv_k'-P(\bar{x})}{\tau_k}\\&=\lim_{k\to\infty}\frac{P(\bar{x}+\tau_ku_k')-P(\bar{x})}{\tau_k}=\xi.\end{aligned}$$

Thus, $\xi\in DP(\bar{x})(0)\cap T_\Lambda(P(\bar{x}))$, which means $(0,\xi)\in\tilde{\mathcal{L}}(\bar{x})$. With $x_k^*\to 0$, we have

$$0\in D^*P(\bar{x};(0,\xi))(\zeta),\quad \zeta\in N_\Lambda(P(\bar{x});\xi).$$

By (9), we can easily obtain that

$$\lim_{k\to\infty}\left\langle y_k^*,\frac{v_k'}{\|v_k'\|}\right\rangle=\lim_{k\to\infty}\left\langle y_k^*,\frac{t_kv_k}{\|t_kv_k\|}\right\rangle=\lim_{k\to\infty}\left\langle y_k^*,\frac{v_k}{\|v_k\|}\right\rangle=1.$$

In summary for Case (2), we obtain a nonzero vector ζ, a nonzero vector $(0,\xi)\in\tilde{\mathcal{L}}(\bar{x})$, and sequences $(u_k',v_k',y_k^*)\to(0,0,\zeta)$ and $\tau_k\downarrow 0$ such that

$$\begin{aligned}&0\in D^*P(\bar{x};(0,\xi))(\zeta),\quad \zeta\in N_\Lambda(P(\bar{x});\xi),\\&y_k^*\in\hat{N}_\Lambda(s_k),\quad s_k=P(\bar{x}+\tau_ku_k')-\tau_kv_k',\\&\lim_{k\to\infty}\left\langle y_k^*,\frac{v_k'}{\|v_k'\|}\right\rangle=1,\end{aligned}$$

which contradicts the assumption in (WSCMS). Thus $P(x)-\Lambda$ is metrically subregular at $(\bar{x}, 0)$. ∎

As an immediate consequence, if we discard the sequential conditions (11) and (12) in WSCMS, we derive from Theorem 3.1 the following sufficient condition for metric subregularity in the form of FOSCMS. The result improves [2, Proposition 2.2] in that P is only assumed to be continuous instead of being calm.

COROLLARY 3.1. *Let $P : \mathscr{X} \to \mathscr{Y}$ be continuous and $\Lambda \subseteq \mathscr{Y}$ be closed at $\bar{x} \in \mathscr{X}$. Suppose that FOSCMS holds at $\bar{x}$, i.e., for all (u, ξ) such that $\xi \in DP(\bar{x})(u) \cap T_\Lambda(P(\bar{x}))$,*

$$0 \in D^*P(\bar{x};(u,\xi))(\zeta), \quad \zeta \in N_\Lambda(P(\bar{x});\xi) \quad \Longrightarrow \quad \zeta = 0.$$

Then $G(x) = P(x) - \Lambda$ is metrically subregular at $(\bar{x}, 0)$.

4. Directional quasi-/pseudo-normality. As we mentioned in the introduction, quasi-/pseudo-normality are also sufficient for metric subregularity. In this section we propose directional versions of quasi-/pseudo-normality and show that they are slightly stronger than the WSCMS. Moreover we show that the SOSCMS implies pseudo-normality. Our results are based on the following observations.

PROPOSITION 4.1. *Let $P : \mathscr{X} \to \mathscr{Y}$, $(u^k, v^k, \zeta^k) \to (u, 0, \zeta)$, $t_k \downarrow 0$ with $\|\zeta\| = 1$, and $s^k = P(\bar{x} + t_k u^k) - t_k v^k$. Then the condition*

$$\lim_{k\to\infty} \left\langle \zeta^k, \frac{v^k}{\|v^k\|} \right\rangle = 1 \tag{14}$$

implies

$$\zeta_i(P_i(\bar{x} + t_k u^k) - s_i^k) > 0 \quad \forall i \in I := \{i : \zeta_i \neq 0\}, \tag{15}$$

which implies

$$\langle \zeta, P(\bar{x} + t_k u^k) - s^k \rangle > 0. \tag{16}$$

Proof. Suppose that (14) holds. Since

$$\begin{aligned}
&\left\| \frac{\zeta^k}{\|\zeta^k\|} - \frac{v^k}{\|v^k\|} \right\|^2 \\
&= \left\langle \frac{\zeta^k}{\|\zeta^k\|} - \frac{v^k}{\|v^k\|}, \frac{\zeta^k}{\|\zeta^k\|} - \frac{v^k}{\|v^k\|} \right\rangle \\
&= \frac{\|\zeta^k\|^2}{\|\zeta^k\|^2} - 2\left\langle \frac{\zeta^k}{\|\zeta^k\|}, \frac{v^k}{\|v^k\|} \right\rangle + \frac{\|v^k\|^2}{\|v^k\|^2} \\
&= 2 - \frac{2}{\|\zeta^k\|} \left\langle \zeta^k, \frac{v^k}{\|v^k\|} \right\rangle,
\end{aligned}$$

$\lim_{k\to\infty} \langle \zeta^k, \frac{v^k}{\|v^k\|} \rangle = 1$ and $\lim_{k\to\infty} \|\zeta^k\| = \|\zeta\| = 1$, we have

$$\lim_{k\to\infty} \left\| \frac{\zeta^k}{\|\zeta^k\|} - \frac{v^k}{\|v^k\|} \right\| = 0.$$

Consequently, $\lim_{k\to\infty} \frac{v^k}{\|v^k\|} = \frac{\zeta}{\|\zeta\|}$. Thus when k is large enough, for each $i = 1, \ldots, m$, with $\zeta_i \neq 0$, v_i^k have the same sign as ζ_i. This means

$$\zeta_i v_i^k > 0 \quad \forall i \in I := \{i : \zeta_i \neq 0\},$$

which implies (15). Since $\zeta \neq 0$, (15) obviously implies (16). ◻

We are now in a position to define the concept of directional quasi-/pseudo-normality.

DEFINITION 4.1 (directional quasi-/pseudo-normality). *Let $P : \mathscr{X} \to \mathscr{Y}$ with $P(\bar{x}) \in \Lambda$.*

(a) *We say that directional quasi-normality holds at $\bar{x}$ if for all*

$$(0,0) \neq (u,\xi) \in \widetilde{\mathcal{L}}(\bar{x}) := \{(u,\xi) \in \mathscr{X} \times \mathscr{Y} | \xi \in DP(x)(u) \cap T_\Lambda(P(x))\},$$

there exists no $\zeta \neq 0$ such that

$$0 \in D^*P(\bar{x};(u,\xi))(\zeta), \quad \zeta \in N_\Lambda(P(\bar{x});\xi), \tag{17}$$

and

$$\begin{cases} \exists (u^k, s^k, \zeta^k) \to (u, P(\bar{x}), \zeta) \text{ and } t_k \downarrow 0 \\ \text{s.t. } \zeta^k \in \widehat{N}_\Lambda(s^k) \text{ and } \zeta_i(P_i(\bar{x} + t_k u^k) - s_i^k) > 0 \text{ if } \zeta_i \neq 0. \end{cases}$$

(b) *We say that directional pseudo-normality holds at $\bar{x}$ if for all $(0,0) \neq (u,\xi) \in \widetilde{\mathcal{L}}(\bar{x})$, there exists no $\zeta \neq 0$ such that (17) holds and*

$$\begin{cases} \exists (u^k, s^k, \zeta^k) \to (u, P(\bar{x}), \zeta) \text{ and } t^k \downarrow 0 \\ \text{s.t. } \zeta^k \in \widehat{N}_\Lambda(s^k) \text{ and } \langle \zeta, P(\bar{x} + t_k u^k) - s^k \rangle > 0. \end{cases}$$

By virtue of Proposition 4.1, directional pseudo-normality is stronger than directional quasi-normality. And consequently from Theorem 3.1, they can provide sufficient conditions for metric subregularity.

COROLLARY 4.1. *Let $P : \mathscr{X} \to \mathscr{Y}$, $P(\bar{x}) \in \Lambda$, where $P(x)$ is continuous at $\bar{x}$ and Λ is closed near $\bar{x}$. If either directional pseudo-normality or directional quasi-normality holds at $\bar{x}$, then the set-valued map $G(x) = P(x) - \Lambda$ is metrically subregular at $(\bar{x}, 0)$.*

By definition, directional quasi-/pseudo-normality is weaker than quasi-/pseudo-normality, and the following example shows that it is weaker than both quasi-normality and FOSCMS.

Example 4.1 (FOSCMS fails but directional pseudo-normality holds). Consider the constraint system defined by $P(x) = (x, -x^2) \in \Lambda$, where

$$\Lambda := \{(x,y) | y \leq 0 \text{ or } y \leq x\}.$$

The point $\bar{x} = 0$ is feasible since $(0,0) \in \Lambda$. We have

$$P(\bar{x}) = (0,0), \quad \nabla P(\bar{x}) = \begin{pmatrix} 1 \\ -2\bar{x} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad T_\Lambda(P(\bar{x})) = \Lambda,$$

and the linearized cone $\mathcal{L}(\bar{x}) = \{u \in \mathbb{R} | 0 \in -\nabla P(\bar{x})u + T_\Lambda(P(\bar{x}))\} = \mathbb{R}$. Let $\bar{u} := -1 \in \mathcal{L}(\bar{x})$, $\zeta := (0,1)$, and $(x^k, y^k) = P(\bar{x}) + \frac{1}{k}\nabla P(\bar{x})\bar{u} = (-\frac{1}{k}, 0)$. Then $\nabla P(\bar{x})^T\zeta = 0$ and for each k, $\zeta \in N_\Lambda(x^k, y^k)$. Thus $\nabla P(\bar{x})^T\zeta = 0$ and $\zeta \in N_\Lambda(P(\bar{x}); \nabla P(\bar{x})\bar{u})$. Hence the FOSCMS fails at $\bar{x}$.

However, we can prove that directional pseudo-normality holds at $\bar{x}$. We prove it by contradiction. Assume that directional pseudo-normality fails at $\bar{x}$. Then there exist $0 \neq u \in \mathcal{L}(\bar{x})$, $0 \neq \zeta \in N_\Lambda(P(\bar{x}); \nabla P(\bar{x})u)$, and a sequence $\{u^k, s^k, \zeta^k\}$ converging to $(u, P(\bar{x}), \zeta)$ and $t_k \downarrow 0$ such that

$$\nabla P(\bar{x})^T\zeta = 0, \quad \sum_{i=1}^{2} \zeta_i(P_i(\bar{x} + t_k u^k) - s_i^k) > 0, \quad \zeta^k \in \widehat{N}_\Lambda(s^k). \tag{18}$$

Solving $\nabla P(\bar{x})^T\zeta = 0$, we obtain $\zeta_1 = 0$. Moreover since $N_\Lambda(P(\bar{x})) = \{0\} \times \mathbb{R}_+ \cup \{(-r, r) | r \geq 0\}$, we have $\zeta \in \{0\} \times \mathbb{R}_{++}$. Since $\zeta^k \to \zeta$ and $\zeta^k \in \widehat{N}_\Lambda(s^k)$, we must have $\zeta^k \in \{0\} \times \mathbb{R}_{++}$ and $s^k \in \{0\} \times \mathbb{R}_+$. Thus we obtain

$$\sum_{i=1}^{2} \zeta_i(P_i(z^k) - s_i^k) = \zeta_2(P_2(z^k) - s_2^k) = \lambda(-(z^k)^2 - s_2^k) \leq 0,$$

where $z^k := \bar{x} + t_k u^k$. But this contradicts (18). Hence directional pseudo-normality holds at $\bar{x}$.

We now consider the case where Λ is the union of finitely many convex polyhedral sets in $\mathscr{Y}$, i.e., $\Lambda := \bigcup_{i=1}^{p} \Lambda_i$, where

$$\Lambda_i := \{y \in \mathscr{Y} | \langle \lambda_{ij}, y \rangle \leq b_{ij},\ j = 1, \ldots, m_i\}, \quad i = 1, \ldots, p,$$

with $\lambda_{ij} \in \mathscr{Y},\ b_{ij} \in \mathbb{R}$ for $j = 1, \ldots, m_i$, are convex polyhedral sets. As noted in the introduction, by Robinson's multifunction theory [44], we know that when P is affine and Λ is the union of finitely many convex polyhedral sets, the set-valued map G^{-1} is upper Lipschitz continuous and hence calm at each point of the graph. What is more, we now show that the pseudo-normality always holds. To our knowledge, this result has never been shown in the literature before.

The following results will be needed in the proof. For every $s \in \Lambda$, we denote by $p(s) := \{i = 1, \ldots, p | s \in \Lambda_i\}$ the index set of the convex polyhedral sets containing s. Then we have from [11] that

$$T_\Lambda(s) = \bigcup_{i \in p(s)} T_{\Lambda_i}(s), \qquad \widehat{N}_\Lambda(s) = \bigcap_{i \in p(s)} \widehat{N}_{\Lambda_i}(s). \tag{19}$$

Proposition 4.2. *Let $P : \mathscr{X} \to \mathscr{Y}$. Suppose that $P(x)$ is affine and Λ is the union of finitely many convex polyhedral sets defined as above. Then for any feasible point $\bar{x}$ satisfying $P(\bar{x}) \in \Lambda$, pseudo-normality holds.*

Proof. We prove the proposition by contradiction. Assume that pseudo-normality does not hold at $\bar{x}$. Then there exists $\zeta \neq 0$ such that

$$\begin{cases} 0 = \nabla P(\bar{x})^*\zeta, \quad \zeta \in N_\Lambda(P(\bar{x})), \\ \exists (x^k, s^k, \zeta^k) \to (\bar{x}, P(\bar{x}), \zeta) \\ \ \text{s.t. } \zeta^k \in \widehat{N}_\Lambda(s^k),\ \langle \zeta, P(x^k) - s^k \rangle > 0. \end{cases}$$

As $s^k \to P(\bar{x})$ when $k \to \infty$ and $s^k \in \Lambda = \bigcup_{i=1}^{p} \Lambda_i$, by virtue of (19), taking a subsequence if necessary, there exists $i \in \{1, \ldots, p\}$ such that for k sufficiently large,

$s^k \in \Lambda_i,\ P(\bar{x}) \in \Lambda_i,\ \zeta^k \in N_{\Lambda_i}(s^k)$. Define $J(s^k) := \{j = 1, \ldots, m_i | \langle \lambda_{ij}, s^k \rangle = b_{ij}\}$ and $J(P(\bar{x})) := \{j = 1, \ldots, m_i | \langle \lambda_{ij}, P(\bar{x}) \rangle = b_{ij}\}$. Since $\zeta^k \neq 0$, s^k is not an interior point of Λ_i and hence the index set $J(s^k)$ is not empty. Since $s^k \to P(\bar{x})$, we have $J(s^k) \subseteq J(P(\bar{x}))$ when k is sufficiently large. Hence without loss of generality, we can find a nonempty set $J \subseteq J(P(\bar{x}))$ such that $J(s^k) \equiv J$ for all k large enough. Define $C := \{\lambda_{ij} | j \in J\}$. Then we have $\zeta^k \in cone(C)$, where

$$cone(C) := \left\{ \sum_{j \in J} c_j \lambda_{ij} | c_j \geq 0 \ \forall j \in J \right\}$$

denotes the conic hull of C. It follows that $\zeta \in cone(C)$. Since when k large enough, for each $j \in J$, $\langle \lambda_{ij}, P(\bar{x}) - s^k \rangle = b_{ij} - b_{ij} = 0$, we obtain $\langle \zeta, P(\bar{x}) - s^k \rangle = 0$. Thus for sufficiently large k, we have

$$\begin{aligned} 0 &> \langle \zeta, s^k - P(x^k) \rangle + \langle \zeta, P(\bar{x}) - s^k \rangle \\ &= \langle \zeta, P(\bar{x}) - P(x^k) \rangle \\ &= \langle \zeta, \nabla P(\bar{x})(\bar{x} - x^k) \rangle, \end{aligned}$$

which contradicts the condition that $0 = \nabla P(\bar{x})^* \zeta$. Thus pseudo-normality holds at $\bar{x}$. ☐

For a single-valued mapping $P : \mathscr{X} \to \mathscr{Y}$ which is C^1 at $\bar{x}$ and $u \in \mathscr{Y}$, we define its second-order graphical derivative of $P(x)$ at $\bar{x}$ in direction u as

$$D^2 P(\bar{x})(u) := \left\{ l \in \mathscr{Y} \middle| \exists t_k \downarrow 0,\ u^k \to u \text{ s.t. } l = \lim_{k \to \infty} \frac{P(\bar{x} + t_k u^k) - P(\bar{x}) - t_k \nabla P(\bar{x}) u^k}{\frac{1}{2} t_k^2} \right\}.$$

In [11, Theorem 4.3], a second-order sufficient condition for metric subregularity (SOSCMS) is presented for a split system in Banach spaces where one of the system is metrically subregular. Specializing the result in [11, Theorem 4.3] to our system (1), we may conclude that if $P(x)$ is C^1 and directionally second-order differentiable, Λ is the union of finitely many convex polyhedral sets and SOSCMS as stated in Theorem 4.1 holds, then the system is directionally pseudo-normal. In Theorem 4.1, we extend this result to the case where $P(x)$ is C^1 and $\nabla P(x)$ is directionally calm at $\bar{x}$ in each nonzero direction u lying in the linearization cone, which means that there exist positive numbers $\epsilon,\ \delta,\ L_u$ such that

$$\|\nabla P(\bar{x} + tu') - \nabla P(\bar{x})\| \leq L_u \|tu'\| \quad \text{for all } 0 < t < \epsilon,\ \|u' - u\| < \delta.$$

Moreover, we show that SOSCMS implies directional pseudo-normality.

Theorem 4.1. *Let $P(\bar{x}) \in \Lambda$, where $P(x)$ is C^1, Λ is the union of finitely many convex polyhedral sets in $\mathscr{Y}$, and $\nabla P(x)$ is directionally calm at $\bar{x}$ in each direction $0 \neq u$ such that $\nabla P(\bar{x}) u \in T_\Lambda(P(\bar{x}))$. Suppose the SOSCMS holds at $\bar{x}$, i.e., for all $0 \neq u$ such that $\nabla P(\bar{x}) u \in T_\Lambda(P(\bar{x}))$, there exists no $\zeta \neq 0$ such that*

$$\begin{cases} \nabla P(\bar{x})^* \zeta = 0,\ \zeta \in N_\Lambda(P(\bar{x}); \nabla P(\bar{x}) u), \\ \exists l \in D^2 P(\bar{x})(u) \ s.t. \ \langle \zeta, l \rangle \geq 0. \end{cases}$$

Then $\bar{x}$ is directionally pseudo-normal at $\bar{x}$.

Proof. We prove that SOSCMS is stronger than directional pseudo-normality by contradiction. Assume there exist $0 \neq u$ such that $\nabla P(\bar{x})u \in T_\Lambda(P(\bar{x}))$ and $\zeta \neq 0$ such that

$$\begin{cases} \nabla P(\bar{x})^*\zeta = 0, \quad \zeta \in N_\Lambda(P(\bar{x}); \nabla P(\bar{x})u), \\ \exists (u^k, s^k, \zeta^k) \to (u, P(\bar{x}), \zeta) \text{ and } t_k \downarrow 0 \\ \text{ s.t. } \zeta^k \in \widehat{N}_\Lambda(s^k), \ \sum_{i=1}^m \zeta_i(P_i(\bar{x} + t_k u^k) - s_i^k) > 0. \end{cases}$$

Notice that $\langle P(x), e_j \rangle$, where e_j is in the orthogonal basis $\mathscr{E}$, is a function on $\mathscr{X}$. By the mean value theorem, for each j and k, there exist $0 < c_j^k < t_k$ such that

$$\langle P(\bar{x}+t_k u^k) - P(\bar{x}), e_j \rangle = \langle \nabla P(\bar{x}+c_j^k u^k)(\bar{x}+t_k u^k - \bar{x}), e_j \rangle = \langle \nabla P(\bar{x}+c_j^k u^k) t_k u^k, e_j \rangle.$$

Thus

$$\begin{aligned} &\left\langle \frac{P(\bar{x} + t_k u^k) - P(\bar{x}) - t_k \nabla P(\bar{x}) u^k}{\frac{1}{2} t_k^2}, e_j \right\rangle \\ &= \frac{1}{2t_k} \left(\frac{\langle P(\bar{x} + t_k u^k) - P(\bar{x}), e_j \rangle}{t_k} - \langle \nabla P(\bar{x}) u^k, e_j \rangle \right) \\ &= \frac{2}{t_k} \left(\frac{\langle \nabla P(\bar{x} + c_j^k u^k) u^k, e_j \rangle t_k}{t_k} - \langle \nabla P(\bar{x}) u^k, e_j \rangle \right) \\ &= \frac{2}{t_k} (\langle \nabla P(\bar{x} + c_j^k u^k) u^k, e_j \rangle - \langle \nabla P(\bar{x}) u^k, e_j \rangle). \end{aligned}$$

Since $\nabla P(x)$ is directionally calm at $\bar{x}$ in direction u, there exists $L_u > 0$ such that for each j and sufficiently large k,

$$\begin{aligned} &\left\| \frac{2}{t_k} (\langle \nabla P(\bar{x} + c_j^k u^k) u^k, e_j \rangle - \langle \nabla P(\bar{x}) u^k, e_j \rangle) \right\| \\ &\le \frac{2L_u \|\bar{x} + c_j^k u^k - \bar{x}\| \|u^k\|}{t_k} \\ &\le \frac{2L_u t_k \|u^k\|^2}{t_k} = 2L_u \|u^k\|^2. \end{aligned}$$

This implies that the sequence $\{\langle \frac{P(\bar{x}+t_k u^k) - P(\bar{x}) - t_k \nabla P(\bar{x}) u^k}{\frac{1}{2} t_k^2}, e_j \rangle\}$ is bounded. Consequently, the sequence $\{\frac{P(\bar{x}+t_k u^k) - P(\bar{x}) - t_k \nabla P(\bar{x}) u^k}{\frac{1}{2} t_k^2}\}$ is bounded. Taking a subsequence if necessary, there exists l such that

$$l := \lim_{k \to \infty} \frac{P(\bar{x} + t_k u^k) - P(\bar{x}) - t_k \nabla P(\bar{x}) u^k}{\frac{1}{2} t_k^2} \in D^2 P(\bar{x})(u).$$

It follows that

$$\begin{aligned} 0 &< \langle \zeta, P(\bar{x} + t_k u^k) - s^k \rangle \\ &= \langle \zeta, P(\bar{x} + t_k u^k) - P(\bar{x}) + P(\bar{x}) - s^k \rangle \\ &= \left\langle \zeta, t_k \nabla P(\bar{x}) u^k + \frac{t_k^2}{2} l + o(t_k^2) \right\rangle + \langle \zeta, P(\bar{x}) - s^k \rangle. \end{aligned} \tag{20}$$

By assumption, $\nabla P(\bar{x})^*\zeta = 0$, which means $\langle \zeta, t_k \nabla P(\bar{x}) u^k \rangle = 0$. And since $s^k \to P(\bar{x})$ as $k \to \infty$, taking a subsequence if necessary, there exists $j \in \{1, \ldots, p\}$ such that for k sufficiently large, $s^k \in \Lambda_j,\ P(\bar{x}) \in \Lambda_j,\ \zeta^k \in N_{\Lambda_j}(s^k)$. Since Λ_j is convex polyhedral, similar to the discussion in the proof of Proposition 4.2, we have $\langle \zeta, P(\bar{x}) - s^k \rangle = 0$. Thus for k large enough, by (20) we have

$$\begin{aligned} 0 &< \left\langle \zeta, t_k \nabla P(\bar{x}) u^k + \frac{t_k^2}{2} l + o(t_k^2) \right\rangle + \langle \zeta, P(\bar{x}) - s^k \rangle \\ &\leq \frac{t_k^2}{2} \langle \zeta, l + o(1) \rangle. \end{aligned}$$

Then we obtain that $\exists l \in D^2 P(\bar{x})(u)$ such that $\langle \zeta, l \rangle \geq 0$. But this contradicts the assumption of the SOSCMS. The contradiction proves that the SOSCMS implies directional pseudo-normality. ∎

Since directional calmness is obviously weaker than calmness, the following corollary follows immediately from Theorem 4.1. We say that $P(x)$ is $C^{1,c}$ at $\bar{x}$ if $P(x)$ is C^1 at $\bar{x}$ and $\nabla P(x)$ is calm at $\bar{x}$, i.e., there exist $\kappa > 0$ and a neighborhood U of $\bar{x}$ such that $\|\nabla P(x) - \nabla P(\bar{x})\| \leq \kappa \|x - \bar{x}\|$ for all $x \in U$.

Corollary 4.2. *Let $P(\bar{x}) \in \Lambda$, where P is $C^{1,c}$ and Λ is the union of finitely many convex polyhedral sets in $\mathscr{Y}$. Suppose SOSCMS holds at $\bar{x}$. Then $\bar{x}$ is directionally pseudo-normal.*

In summary, we have shown the following implications:

SOSCMS $\Longrightarrow$ directional pseudo-normality $\Longrightarrow$ directional quasi-normality $\Longrightarrow$ WSCMS $\Longrightarrow$ metric subregularity.

5. Applications to complementarity and KKT systems.

In this section we apply our results to complementarity and KKT systems. When directional quasi-/pseudo-normality are applied to these systems we derive expressions that are much simpler and moreover can be directly compared with classical quasi-/pseudo-normality.

First we consider the complementarity system (CS) formulated as follows:

$$\text{(CS)} \qquad H(x) = 0, \quad 0 \leq \Phi(x) \perp \Psi(x) \geq 0,$$

where $H(x) : \mathbb{R}^n \to \mathbb{R}^d,\ \Phi, \Psi : \mathbb{R}^n \to \mathbb{R}^m$. For simplicity of explanation, we omit possible inequality and abstract constraints and moreover we assume that all functions are continuously differentiable. The results can be extended to the general case in a straightforward manner.

Define $\Omega_{EC} := \{(a, b) \in \mathbb{R}_+ \times \mathbb{R}_+ | ab = 0\}$. For any set C and any positive integer m we denote by C^m the m-Cartesian product of C. (CS) can be rewritten as

$$(H(x), (\Phi_1(x), \Psi_1(x)), \ldots, (\Phi_m(x), \Psi_m(x))) \in \{0\}^d \times \Omega_{EC}^m.$$

To derive the precise form of the directional quasi-/pseudo-normality, we review the formulas for the regular normal cone, the limiting normal cone, the tangent cone, and the directional limiting normal cone of the set Ω_{EC}.

Lemma 5.1 (see [13, Lemma 4.1]). *The Fréchet normal cone to Ω_{EC} is*

$$\widehat{N}_{\Omega_{EC}}(a, b) = \left\{ -(\gamma, \nu) \left| \begin{array}{ll} \nu = 0 & \text{if } 0 = a < b, \\ \gamma \geq 0,\ \nu \geq 0 & \text{if } a = b = 0, \\ \gamma = 0 & \text{if } a > b = 0 \end{array} \right. \right\},$$

the limiting normal cone is

$$N_{\Omega_{EC}}(a,b) = \left\{ \begin{array}{ll} \widehat{N}_{\Omega_{EC}}(a,b) & \text{if } (a,b) \neq (0,0), \\ \{-(\gamma,\nu) | \text{either } \gamma > 0,\ \nu > 0 \text{ or } \ \gamma\nu = 0\} & \text{if } (a,b) = (0,0) \end{array} \right\},$$

and the tangent cone is

$$T_{\Omega_{EC}}(a,b) = \left\{ (d_1,d_2) \left| \begin{array}{ll} d_1 = 0 & \text{if } 0 = a < b, \\ (d_1,d_2) \in \Omega_{EC} & \text{if } a = b = 0, \\ d_2 = 0 & \text{if } a > b = 0 \end{array} \right. \right\}.$$

For all $d = (d_1,d_2) \in T_{\Omega_{EC}}(a,b)$, the directional limiting normal cone to Ω_{EC} in direction d is

$$N_{\Omega_{EC}}((a,b);d) = \left\{ \begin{array}{ll} N_{\Omega_{EC}}(a,b) & \text{if } (a,b) \neq (0,0), \\ N_{\Omega_{EC}}(d_1,d_2) & \text{if } (a,b) = (0,0) \end{array} \right\}.$$

Let $\bar{x}$ be a feasible point of the system (CS). We define the index sets

$$\begin{aligned} I_{00} &:= I_{00}(\bar{x}) := \{i | \Phi_i(\bar{x}) = 0,\ \Psi_i(\bar{x}) = 0\}, \\ I_{0+} &:= I_{0+}(\bar{x}) := \{i | \Phi_i(\bar{x}) = 0,\ \Psi_i(\bar{x}) > 0\}, \\ I_{+0} &:= I_{+0}(\bar{x}) := \{i | \Phi_i(\bar{x}) > 0,\ \Psi_i(\bar{x}) = 0\}, \end{aligned}$$

and define the linearization cone as

$$\mathcal{L}(\bar{x}) := \left\{ u \in \mathbb{R}^n | \left| \begin{array}{ll} 0 = \nabla H_i(\bar{x})u, & i = 1,\ldots,d, \\ 0 = \nabla \Phi_i(\bar{x})u, & i \in I_{0+}, \\ 0 = \nabla \Psi_i(\bar{x})u, & i \in I_{+0}, \\ (\nabla \Phi_i(\bar{x})u, \nabla \Psi_i(\bar{x})u) \in \Omega_{EC}, & i \in I_{00} \end{array} \right. \right\}.$$

Given $u \in \mathcal{L}(\bar{x})$ we define

$$\begin{aligned} I_{+0}(u) &:= \{i \in I_{00} | \nabla \Phi_i(\bar{x})u > 0 = \nabla \Psi_i(\bar{x})u\}, \\ I_{0+}(u) &:= \{i \in I_{00} | \nabla \Phi_i(\bar{x})u = 0 < \nabla \Psi_i(\bar{x})u\}, \\ I_{00}(u) &:= \{i \in I_{00} | \nabla \Phi_i(\bar{x})u = 0 = \nabla \Psi_i(\bar{x})u\}. \end{aligned}$$

Let $\bar{x}$ be a feasible point of (CS). By Definition 4.1 and Proposition 2.1, since the complementarity set Ω_{EC} is directionally regular, (CS) is directionally quasi- or pseudo-normal if and only if for all directions $0 \neq u \in \mathcal{L}(\bar{x})$ there exists no $(\eta,\gamma,\nu) \neq 0$ such that

$$0 = \nabla H(\bar{x})^T \eta - \nabla \Phi(\bar{x})^T \gamma - \nabla \Psi(\bar{x})^T \nu, \tag{21}$$

$$-(\gamma_i,\nu_i) \in N_{\Omega_{EC}}(\Phi_i(\bar{x}),\Psi_i(\bar{x}); \nabla \Phi_i(\bar{x})u, \nabla \Psi_i(\bar{x})u), \quad i = 1,\ldots,m, \tag{22}$$

$$\begin{aligned} &\exists (u^k,h^k,\phi^k,\psi^k,\eta^k,\gamma^k,\nu^k) \to (u, H(\bar{x}), \Phi(\bar{x}), \Psi(\bar{x}), \eta, \gamma, \nu),\ t_k \downarrow 0 \\ &\text{such that} \left\{ \begin{array}{l} \eta^k \in N_{\{0\}^d}(h^k), -(\gamma_i^k,\nu_i^k) \in \widehat{N}_{\Omega_{EC}}(\phi_i^k,\psi_i^k), \quad i = 1,\ldots,m, \\ \text{if } \eta_i \neq 0,\ \eta_i(H_i(\bar{x}+t_k u^k) - h_i^k) > 0, \\ \text{if } \gamma_i \neq 0,\ \gamma_i(\Phi_i(\bar{x}+t_k u^k) - \phi_i^k) < 0, \\ \text{if } \nu_i \neq 0,\ \nu_i(\Psi_i(\bar{x}+t_k u^k) - \psi_i^k) < 0, \end{array} \right. \end{aligned} \tag{23}$$

or

$$\eta^T(H(\bar{x}+t_k u^k) - h^k) - \gamma^T(\Phi(\bar{x}+t_k u^k) - \phi^k) - \nu^T(\Psi(\bar{x}+t_k u^k) - \psi^k) > 0,$$

respectively.

By the formula for the directional limiting normal cone in Lemma 5.1, (22) is equivalent to (ii) in the following definition. Since $\eta^k \in N_{\{0\}^d}(h^k)$, we have $h^k = 0$. Suppose $\gamma_i \neq 0$. Then for sufficiently large k, $\gamma_i^k \neq 0$. Since $-(\gamma_i^k, \nu_i^k) \in \widehat{N}_{\Omega_{EC}}(\phi_i^k, \psi_i^k)$ we must have $\phi_i^k = 0$. Similarly if $\nu_i \neq 0$, we must have $\psi_i^k = 0$. Based on these discussions, directional quasi-normality for (CS) can be written in the following form which is much more concise.

DEFINITION 5.1. *Let $\bar{x}$ be a feasible solution of* (CS). *$\bar{x}$ is said to be directionally quasi- or pseudo-normal if for all directions $0 \neq u \in \mathcal{L}(\bar{x})$ there exists no $(\eta, \gamma, \nu) \neq 0$ such that*

(i) $0 = \nabla H(\bar{x})^T \eta - \nabla \Phi(\bar{x})^T \gamma - \nabla \Psi(\bar{x})^T \nu$;

(ii) $\gamma_i = 0,\ i \in I_{+0} \cup I_{+0}(u)$; $\nu_i = 0,\ i \in I_{0+} \cup I_{0+}(u)$; *either* $\gamma_i >,\ \nu_i > 0$ *or* $\gamma_i \nu_i = 0,\ i \in I_{00}(u)$;

(iii) $\exists u^k \to u$ *and* $t_k \downarrow 0$ *such that*

$$\begin{cases} \text{if } \eta_i \neq 0,\ \eta_i H_i(\bar{x} + t_k u^k) > 0, \\ \text{if } \gamma_i \neq 0,\ \gamma_i \Phi_i(\bar{x} + t_k u^k) < 0, \\ \text{if } \nu_i \neq 0,\ \nu_i \Psi_i(\bar{x} + t_k u^k) < 0, \end{cases}$$

or

$$\eta^T H(\bar{x} + t_k u^k) - \gamma^T \Phi(\bar{x} + t_k u^k) - \nu^T \Psi(\bar{x} + t_k u^k) > 0,$$

respectively.

Remark 5.1. In Definition 5.1, if we only require that there exists no $(\eta, \gamma, \nu) \neq 0$ satisfying condition (i), then it reduces to the linearly independent constraint qualification (MPEC-LICQ) (see [47]). If we only require that there exists no $(\eta, \gamma, \nu) \neq 0$ satisfying condition (i) and change (ii) to

$$\gamma_i = 0,\ i \in I_{+0};\ \nu_i = 0,\ i \in I_{0+},\ \text{either } \gamma_i \geq 0,\ \nu_i \geq 0 \text{ or } \gamma_i \nu_i = 0,\ i \in I_{00},$$

then it reduces to MPEC-NNAMCQ [54, Definition 2.10]. If we omit (iii), then it reduces to FOSCMS. If we take u to be any direction, then it reduces to the MPEC quasi-/pseudo-normality first given in [29, Definition 3.2] and extended to the Lipschitz continuous case in [57, Definition 5]. Since for the set Ω_{EC} and any $0 \neq d \in T_{\Omega_{EC}}(0,0)$, the directional normal cone $N_{\Omega_{EC}}((0,0); d)$ is strictly smaller than the limiting normal cone $N_{\Omega_{EC}}(0,0)$, if there exists some $u \in \mathcal{L}(\bar{x})$ such that $(\nabla G(\bar{x})u, \nabla H(\bar{x})u) \neq (0,0)$, then directional quasi-/pseudo-normality will be strictly weaker than standard quasi-/pseudo-normality.

We now consider the following KKT system of an optimization problem with equality and inequality constraints:

$$\begin{aligned} &\nabla_x L(x, \mu, \lambda) = 0, \\ &\mu \geq 0, \quad g(x) \leq 0, \quad \langle g(x), \mu \rangle = 0, \\ &h(x) = 0, \end{aligned}$$

where $f : \mathbb{R}^p \to \mathbb{R},\ g : \mathbb{R}^p \to \mathbb{R}^m,\ h : \mathbb{R}^p \to \mathbb{R}^n$ are twice continuously differentiable, $\mu \in \mathbb{R}^m,\ \lambda \in \mathbb{R}^n$, and $L(x, \mu, \lambda) := f(x) + \mu^T g(x) + \lambda^T h(x)$ is the Lagrange function. Denote the feasible set of the KKT system by $\mathcal{F}_{KKT}$. We say that the error bound property holds at $(x^*, \mu^*, \lambda^*) \in \mathcal{F}_{KKT}$ if there exist $\alpha > 0$ and U, a neighborhood of (x^*, μ^*, λ^*), such that

$$d_{\mathcal{F}_{KKT}}(x, \mu, \lambda) \leq \alpha \max\{\|\nabla_x L(x, \mu, \lambda)\|, \|h(x)\|, \|\min\{\mu, -g(x)\}\|\} \quad \forall (x, \mu, \lambda) \in U. \tag{24}$$

It is easy to see that this error bound property can be derived from the metric subregularity/calmness of the KKT system and hence directional quasi-/pseudo-normality is a sufficient condition. Such an error bound property is crucial to the quadratic convergence of the Newton-type method (see [7]). The classical sufficient conditions for the error bound property are either MFCQ combined with the second-order sufficient condition (SOSC) or requiring g, h to be affine and f to be quadratic (see, e.g., [45]). These sufficient conditions were weakened in [9, 18] but still require SOSC. Recently, weaker sufficient conditions have been proposed including the existence of noncritical multipliers, a concept introduced by Izmailov for pure equality constraint in [27], extended by Izmailov and Solodov [28, Definition 2] to problems with inequalities, and further extended to a general variational system by Mordukhovich and Sarabi [36, Definition 3.1]. Note that as shown in [28, Proposition 3], the existence of noncritical multipliers is equivalent to a stronger type of error bound property: existence of $\alpha > 0$ and U, a neighborhood of (x^*, μ^*, λ^*), such that

$$\|x - \bar{x}\| + d_{\mathcal{M}(\bar{x})}(\mu, \lambda) \le \alpha \max\{\|\nabla_x L(x, \mu, \lambda)\|, \|h(x)\|, \|\min\{\mu, -g(x)\}\|\} \qquad \forall (x, \mu, \lambda) \in U,$$

where $\mathcal{M}(\bar{x}) := \{(\mu, \lambda) : 0 = \nabla_x L(\bar{x}, \mu, \lambda),\ \mu \ge 0,\ \langle g(\bar{x}), \mu\rangle = 0\}$ denotes the set of multipliers. Obviously this is a stronger error bound property than the error bound property (24).

The KKT system is a special case of (CS) with

$$H(x, \mu, \lambda) := (\nabla_x L(x, \mu, \lambda), h(x)), \quad \Phi(x, \mu, \lambda) := -g(x), \quad \Psi(x, \mu, \lambda) := \mu.$$

Let $(\bar{x}, \bar{\mu}, \bar{\lambda})$ be a feasible point of the KKT system. We define the following index sets:

$$\begin{aligned}
I_{00} &:= I_{00}(\bar{x}, \bar{\mu}, \bar{\lambda}) := \{i | g_i(\bar{x}) = 0,\ \bar{\mu}_i = 0\},\\
I_{0+} &:= I_{0+}(\bar{x}, \bar{\mu}, \bar{\lambda}) := \{i | g_i(\bar{x}) = 0,\ \bar{\mu}_i > 0\},\\
I_{+0} &:= I_{+0}(\bar{x}, \bar{\mu}, \bar{\lambda}) := \{i | -g_i(\bar{x}) > 0,\ \bar{\mu}_i = 0\}.
\end{aligned}$$

The linearized cone for the KKT system is

$$\mathcal{L}(\bar{x}, \bar{\mu}, \bar{\lambda}) := \left\{ u = (u^x, u^\mu, u^\lambda) \;\middle|\; \begin{array}{ll}
0 = \nabla^2_{xx} L(\bar{x}, \bar{u}, \bar{v}) u^x + \nabla g(\bar{x})^T u^\mu + \nabla h(\bar{x})^T u^\lambda, & \\
0 = \nabla h(\bar{x}) u^x, & \\
0 = \nabla g_i(\bar{x}) u^x, & i \in I_{0+},\\
0 = u_i^\mu, & i \in I_{+0},\\
u_i^\mu \ge 0, \nabla g_i(\bar{x}) u^x \le 0, \text{ and } u_i^\mu \nabla g_i(\bar{x}) u^x = 0, & i \in I_{00}
\end{array} \right\}.$$

Given $u \in \mathcal{L}(\bar{x}, \bar{\mu}, \bar{\lambda})$ we define the index sets

$$\begin{aligned}
I_{+0}(u) &:= \{i \in I_{00} | -\nabla g_i(\bar{x}) u^x > 0 = u_i^\mu\},\\
I_{0+}(u) &:= \{i \in I_{00} | \nabla g_i(\bar{x}) u^x = 0 < u_i^\mu\},\\
I_{00}(u) &:= \{i \in I_{00} | \nabla g_i(\bar{x}) u^x = 0 = u_i^\mu\}.
\end{aligned}$$

Then by Definition 5.1, we propose the following definition of directional quasi-normality for the KKT system.

DEFINITION 5.2. *Let $(\bar{x}, \bar{\mu}, \bar{\lambda})$ be a feasible point of the KKT system. $(\bar{x}, \bar{\mu}, \bar{\lambda})$ is said to be directionally quasi- or pseudo-normal if for all directions*

$$0 \neq \bar{u} := (\bar{u}^x, \bar{u}^\mu, \bar{u}^\lambda) \in \mathcal{L}(\bar{x}, \bar{\mu}, \bar{\lambda})$$

there exists no $(\xi, \zeta, \eta) \neq 0$ such that

(i) $0 = \nabla^2_{xx} L(\bar{x}, \bar{\mu}, \bar{\lambda})\xi + \nabla h(\bar{x})^T \eta + \nabla g(\bar{x})^T \zeta$;

(ii) $\nabla h(\bar{x})\xi = 0$;

(iii) $\zeta_i = 0,\ i \in I_{+0} \cup I_{+0}(\bar{u})$; $\nabla g_i(\bar{x})\xi = 0,\ i \in I_{0+} \cup I_{0+}(\bar{u})$; *either* $\zeta_i > 0,\ \nabla g_i(\bar{x})\xi > 0$ *or* $\zeta_i \nabla g_i(\bar{x})\xi = 0,\ i \in I_{00}(\bar{u})$;

(iv) $\exists u_k := (u^x_k, u^\mu_k, u^\lambda_k) \to \bar{u}$ *and* $t_k \downarrow 0$ *such that*

$$\begin{cases} \text{if } \xi_i \neq 0,\ \xi_i \nabla_x L_i((\bar{x}, \bar{\mu}, \bar{\lambda}) + t_k u_k) > 0, \\ \text{if } \eta_i \neq 0,\ \eta_i h_i(\bar{x} + t_k u^x_k) > 0, \\ \text{if } \zeta_i \neq 0,\ \zeta_i g_i(\bar{x} + t_k u^x_k) > 0, \\ \text{if } (\nabla g(\bar{x})\xi)_i \neq 0,\ (\nabla g(\bar{x})\xi)_i(\bar{u}^\mu_i + t_k (u^\mu_k)_i) < 0; \end{cases}$$

or

$$\xi^T \nabla_x L((\bar{x}, \bar{\mu}, \bar{\lambda}) + t_k u_k) + \eta^T h(\bar{x} + t_k u^x_k) - (\nabla g(\bar{x})\xi)^T (\bar{u} + t_k u^\mu_k) > 0$$

respectively.

Remark 5.2. Let $(\bar{x}, \bar{\mu}, \bar{\lambda})$ be a feasible point to the KKT system. By [28, Definition 2], $(\bar{\mu}, \bar{\lambda}) \in \mathcal{M}(\bar{x})$ is a critical multiplier associated with $\bar{x}$ if there exists (ξ, ζ, η) with $\xi \neq 0$ satisfying that

$$\begin{cases} 0 = \nabla^2_{xx} L(\bar{x}, \bar{\mu}, \bar{\lambda})\xi + \nabla h(\bar{x})^T \eta + \nabla g(\bar{x})^T \zeta, & \\ 0 = \nabla h(\bar{x})\xi, & \\ 0 = \nabla g_i(\bar{x})\xi, & i \in I_{0+}, \\ 0 = \zeta_i, & i \in I_{+0}, \\ \zeta_i \geq 0,\ \nabla g_i(\bar{x})\xi \leq 0 \text{ and } \zeta_i \nabla g_i(\bar{x})\xi = 0, & i \in I_{00}. \end{cases}$$

Note that from Definition 5.2, we can see that even if $(\bar{\mu}, \bar{\lambda})$ is a critical multiplier with $\bar{x}$, it is still possible for directional quasi-normality to hold. In particular, let (ξ, ζ, η) satisfy Definition 5.2 with $\xi \neq 0$. Suppose that for $i \in I_{00}(\bar{u})$, it is not possible to have $\zeta_i > 0,\ \nabla g_i(\bar{x})\xi > 0$. Then $(\bar{\mu}, \bar{\lambda}) \in \mathcal{M}(\bar{x})$ is a critical multiplier associated with $\bar{x}$.

Acknowledgments. The authors would like to thank Helmut Gfrerer for an inspiring discussion on an earlier version of this paper, and the anonymous referees for their helpful suggestions and comments. The alphabetical order of the authors indicates their equal contributions to the paper.

REFERENCES

[1] J.-P. AUBIN, *Lipschitz behavior of solutions to convex minimization problems*, Math. Oper. Res., 9 (1984), pp. 87–111, https://doi.org/10.1287/moor.9.1.87.

[2] M. BENKO, H. GFRERER, AND J. V. OUTRATA, *Calculus for directional limiting normal cones and subdifferentials*, Set-Valued Var. Anal., (2018), pp. 1–33, https://doi.org/10.1007/s11228-018-0492-5.

[3] J. BURKE, *Calmness and exact penalization*, SIAM J. Control. Optim., 29 (1991), pp. 493–497, https://doi.org/10.1137/0329027.

[4] A. DONTCHEV AND R. ROCKAFELLAR, *Regularity and conditioning of solution mappings in variational analysis*, Set-Valued Anal., 12 (2004), pp. 79–109, https://doi.org/10.1023/B:SVAN.0000023394.19482.30.

[5] A. L. Dontchev and R. T. Rockafellar, *Implicit Functions and Solution Mappings*, Springer, New York, 2009, https://doi.org/10.1007/978-0-387-87821-8.
[6] M. J. Fabian, R. Henrion, A. Y. Kruger, and J. V. Outrata, *Error bounds: Necessary and sufficient conditions*, Set-Valued Var. Anal., 18 (2010), pp. 121–149, https://doi.org/10.1007/s11228-010-0133-0.
[7] F. Facchinei, A. Fischer, and M. Herrich, *An LP-Newton method: Nonsmooth equations, KKT systems, and nonisolated solutions*, Math. Program., 146 (2014), pp. 1–36, https://doi.org/10.1007/s10107-013-0676-6.
[8] F. Facchinei and J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Vol. I, Springer Ser. Oper. Res. Financ. Eng., Springer, New York, 2003.
[9] A. Fischer, *Local behavior of an iterative framework for generalized equations with nonisolated solutions*, Math. Program., 94 (2002), pp. 91–124, https://doi.org/10.1007/s10107-002-0364-4.
[10] H. Gfrerer, *On directional metric regularity, subregularity and optimality conditions for nonsmooth mathematical programs*, Set-Valued Var. Anal., 21 (2013), pp. 151–176, https://doi.org/10.1007/s11228-012-0220-5.
[11] H. Gfrerer, *On directional metric subregularity and second-order optimality conditions for a class of nonsmooth mathematical programs*, SIAM J. Optim., 23 (2013), pp. 632–665, https://doi.org/10.1137/120891216.
[12] H. Gfrerer, *On metric pseudo-(sub) regularity of multifunctions and optimality conditions for degenerated mathematical programs*, Set-Valued Var. Anal., 22 (2014), pp. 79–115, https://doi.org/10.1007/s11228-013-0266-z.
[13] H. Gfrerer, *Optimality conditions for disjunctive programs based on generalized differentiation with application to mathematical programs with equilibrium constraints*, SIAM J. Optim., 24 (2014), pp. 898–931, https://doi.org/10.1137/130914449.
[14] H. Gfrerer and D. Klatte, *Lipschitz and Hölder stability of optimization problems and generalized equations*, Math. Program., 158 (2016), pp. 35–75, https://doi.org/10.1007/s10107-015-0914-1.
[15] H. Gfrerer and J. J. Ye, *New constraint qualifications for mathematical programs with equilibrium constraints via variational analysis*, SIAM J. Optim., 27 (2017), pp. 842–865, https://doi.org/10.1137/16M1088752.
[16] I. Ginchev and B. S. Mordukhovich, *On directionally dependent subdifferentials*, C. R. Bulg. Acad. Sci., 64 (2011), pp. 497–508.
[17] L. Guo, J. J. Ye, and J. Zhang, *Mathematical programs with geometric constraints in Banach spaces: Enhanced optimality, exact penalty, and sensitivity*, SIAM J. Optim., 23 (2013), pp. 2295–2319, https://doi.org/10.1137/130910956.
[18] W. W. Hager and S. M. Gowda, *Stability in the presence of degeneracy and error estimation*, Math. Program., 85 (1999), pp. 181–192, https://doi.org/10.1007/s101070050051.
[19] R. Henrion and A. Jourani, *Subdifferential conditions for calmness of convex constraints*, SIAM J. Optim., 13 (2002), pp. 520–534, https://doi.org/10.1137/S1052623401386071.
[20] R. Henrion, A. Jourani, and J. V. Outrata, *On the calmness of a class of multifunctions*, SIAM J. Optim., 13 (2002), pp. 603–618, https://doi.org/10.1137/S1052623401395553.
[21] R. Henrion and J. Outrata, *A subdifferential condition for calmness of multifunctions*, J. Math. Anal. Appl., 258 (2001), pp. 110–130, https://doi.org/10.1006/jmaa.2000.7363.
[22] R. Henrion and J. V. Outrata, *Calmness of constraint systems with applications*, Math. Program., 104 (2005), pp. 437–464, https://doi.org/10.1007/s10107-005-0623-2.
[23] A. D. Ioffe, *Necessary and sufficient conditions for a local minimum.* 1: *A reduction theorem and first order conditions*, SIAM J. Control Optim., 17 (1979), pp. 245–250, https://doi.org/10.1137/0317019.
[24] A. D. Ioffe, *Regular points of Lipschitz functions*, Trans. Amer. Math. Soc., 251 (1979), pp. 61–69, https://doi.org/10.2307/1998683.
[25] A. D. Ioffe, *Metric regularity and subdifferential calclulus*, Russian Math. Surveys, 55 (2000), pp. 501–558, https://doi.org/10.1070/RM2000v055n03ABEH000292.
[26] A. D. Ioffe and J. V. Outrata, *On metric and calmness qualification conditions in subdifferential calculus*, Set-valued Anal., 16 (2008), pp. 199–227, https://doi.org/10.1007/s11228-008-0076-x.
[27] A. F. Izmailov, *On the analytical and numerical stability of critical Lagrange multipliers*, Comput. Math. Math. Phys., 45 (2005), pp. 966–982.
[28] A. F. Izmailov and M. V. Solodov, *Stabilized SQP revisited*, Math. Program., 133 (2012), pp. 93–120, https://doi.org/10.1007/s10107-010-0413-3.
[29] C. Kanzow and A. Schwartz, *Mathematical programs with equilibrium constraints: Enhanced Fritz John-conditions, new constraint qualifications, and improved exact penalty results*,

SIAM J. Optim., 20 (2010), pp. 2730–2753, https://doi.org/10.1137/090774975.
[30] D. Klatte, *A note on quantitative stability results in nonlinear optimization*, in Proceedings of the 19. Jahrestagung Mathematische Optimierung, Seminarbericht 90, Humboldt-Universität Berlin, Sektion Mathematik, 1987, pp. 77–86.
[31] D. Klatte and B. Kummer, *Constrained minima and Lipschitzian penalties in metric spaces*, SIAM J. Optim., 13 (2002), pp. 619–633, https://doi.org/10.1137/S105262340139625X.
[32] D. Klatte and B. Kummer, *Optimization methods and stability of inclusions in Banach spaces*, Math. program., 117 (2009), pp. 305–330, https://doi.org/10.1007/s10107-007-0174-9.
[33] W. Li, *Abadie's constraint qualification, metric regularity, and error bounds for differentiable convex inequalities*, SIAM J. Optim., 7 (1997), pp. 966–978, https://doi.org/10.1137/S1052623495287927.
[34] L. Minchenko and A. Tarakanov, *On error bounds for quasinormal programs*, J. Optim. Theory Appl., 148 (2011), pp. 571–579.
[35] B. S. Mordukhovich, *Variational Analysis and Generalized Differentiation* I: *Basic Theory*, Grundlehren Math. Wiss. 330, Springer, Berlin, 2006, https://doi.org/10.1007/3-540-31247-1.
[36] B. S. Mordukhovich and M. E. Sarabi, *Critical multipliers in variational systems via second-order generalized differentiation*, Math. Program., 169 (2018), pp. 605–648, https://doi.org/10.1007/s10107-017-1155-2.
[37] H. V. Ngai and P. N. Tinh, *Metric subregularity of multifunctions: First and second order infinitesimal characterizations*, Math. Oper. Res., 40 (2015), pp. 703–724, https://doi.org/10.1287/moor.2014.0691.
[38] M. Nikolova, *Minimizers of cost-functions involving nonsmooth data-fidelity terms. Application to the processing of outliers*, SIAM J. Numer. Anal., 40 (2002), pp. 965–994, https://doi.org/10.1137/S0036142901389165.
[39] J.-P. Penôt, *Error bounds, calmness and their applications in nonsmooth analysis*, in Nonlinear Analysis and Optimization II: Optimization, Contemp. Math. 514, A. Leizarowitz, B. S. Mordukhovich, I. Shafrir, and A. J. Zaslavski, eds., American Mathematical Society, Providence, RI, 2010, pp. 225–247, https://doi.org/10.1090/conm/514/10110.
[40] T. T. Pham and R. J. deFigueiredo, *Maximum likelihood estimation of a class of non-Gaussian densities with application to I_p deconvolution*, IEEE Trans. Acoust. Speech Signal Process., 37 (1989), pp. 73–82, https://doi.org/10.1109/29.17502.
[41] J. R. Rice and J. S. White, *Norms for smoothing and estimation*, SIAM Rev., 6 (1964), pp. 243–256, https://doi.org/10.1137/1006061.
[42] S. M. Robinson, *Stability theory for systems of inequalities. Part* I: *Linear systems*, SIAM J. Numer. Anal., 12 (1975), pp. 754–769, https://doi.org/10.1137/0712056.
[43] S. M. Robinson, *Stability theory for systems of inequalities. Part* II: *Differentiable nonlinear systems*, SIAM J. Numer. Anal., 13 (1976), pp. 497–513, https://doi.org/10.1137/0713043.
[44] S. M. Robinson, *Some continuity properties of polyhedral multifunctions*, Mathematical Programming at Oberwolfach, Math. Program. Stud. 14, Springer, Berlin, 1981, pp. 206–214, https://doi.org/10.1007/BFb0120929.
[45] S. M. Robinson, *Generalized equations and their solutions, part* II: *Applications to nonlinear programming*, in Optimality and Stability in Mathematical Programming, Math. Program. Stud. 19, Springer, Berlin, 1982, pp. 200–221, https://doi.org/10.1007/BFb0120989.
[46] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Grundlehren Math. Wiss. 317, Springer, Berlin, 1998, https://doi.org/10.1007/978-3-642-02431-3.
[47] S. Scholtes and M. Stöhr, *How stringent is the linear independence assumption for mathematical programs with complementarity constraints?*, Math. Oper. Res., 26 (2001), pp. 851–863, https://doi.org/10.1287/moor.26.4.851.10007.
[48] M. Studniarski and D. E. Ward, *Weak sharp minima: Characterizations and sufficient conditions*, SIAM J. Control Optim., 38 (1999), pp. 219–236, https://doi.org/10.1137/S0363012996301269.
[49] X. Wang, J. Ye, X. Yuan, S. Zeng, and J. Zhang, *Perturbation Techniques for Convergence Analysis of Proximal Gradient Method and Other First-Order Algorithms Via Variational Analysis*, preprint, https://arxiv.org/abs/1810.10051, 2018.
[50] Z. Wu and J. Y. Jane, *On error bounds for lower semicontinuous functions*, Math. Program., 92 (2002), pp. 301–314, https://doi.org/10.1007/s101070100278.
[51] Z. Wu and J. J. Ye, *Sufficient conditions for error bounds*, SIAM J. Optim., 12 (2001), pp. 421–435, https://doi.org/10.1137/S1052623400371557.
[52] Z. Wu and J. J. Ye, *First-order and second-order conditions for error bounds*, SIAM J. Optim., 14 (2003), pp. 621–645, https://doi.org/10.1137/S1052623402412982.
[53] J. J. Ye, *Constraint qualifications and necessary optimality conditions for optimization prob-*

lems with variational inequality constraints, SIAM J. Optim., 10 (2000), pp. 943–962, https://doi.org/10.1137/S105262349834847X.

[54] J. J. Ye, *Necessary and sufficient optimality conditions for mathematical programs with equilibrium constraints*, J. Math. Anal. Appl., 307 (2005), pp. 350–369, https://doi.org/10.1016/j.jmaa.2004.10.032.

[55] J. J. Ye and X. Y. Ye, *Necessary optimality conditions for optimization problems with variational inequality constraints*, Math. Oper. Res., 22 (1997), pp. 977–997, https://doi.org/10.1287/moor.22.4.977.

[56] J. J. Ye and J. Zhang, *Enhanced Karush–Kuhn–Tucker condition and weaker constraint qualifications*, Math. Program., 139 (2013), pp. 353–381, https://doi.org/10.1007/s10107-013-0667-7.

[57] J. J. Ye and J. Zhang, *Enhanced Karush–Kuhn–Tucker conditions for mathematical programs with equilibrium constraints*, J. Optim. Theory Appl., 163 (2014), pp. 777–794, https://doi.org/10.1007/s10957-013-0493-3.

[58] J. J. Ye and J. Zhou, *Verifiable sufficient conditions for the error bound property of second-order cone complementarity problems*, Math. Program., 171 (2018), pp. 361–395, https://doi.org/10.1007/s10107-017-1193-9.

[59] X. Y. Zheng and K. F. Ng, *Metric subregularity and constraint qualifications for nonconvex generalized equations in Banach spaces*, SIAM J. Optim., 20 (2010), pp. 2119–2136, https://doi.org/10.1137/090772174.