Second-order optimality conditions for non-convex set-constrained optimization problems

Helmut Gfrerer* Jane J. Ye† Jinchuan Zhou‡

January 13, 2020

Abstract. In this paper we study second-order optimality conditions for non-convex set-constrained optimization problems. For a convex set-constrained optimization problem, it is well-known that second-order optimality conditions involve the support function of the second-order tangent set. In this paper we propose two approaches for establishing second-order optimality conditions for the non-convex case. In the first approach we extend the concept of the support function so that it is applicable to general non-convex set-constrained problems, whereas in the second approach we introduce the notion of the directional regular tangent cone and apply classical results of convex duality theory. Besides the second-order optimality conditions, the novelty of our approach lies in the systematic introduction and use, respectively, of directional versions of well-known concepts from variational analysis.

Keywords: second-order tangent sets, second-order optimality conditions, lower generalized support function, directional metric subregularity, directional normal cones, directional regular tangent cones, directional Robinson's constraint qualification, directional non-degeneracy.

AMS subject classifications. 90C26, 90C46, 49J53.

1 Introduction

Second-order optimality conditions have long been recognized as an important tool in optimization theory and algorithms. In this paper we aim at developing second-order optimality conditions for a set-constrained optimization problem in the form of

$$(P) \qquad \min \quad f(x) \quad \text{s.t.} \quad g(x) \in \Lambda, \tag{1}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}^m$ are twice continuously differentiable, and Λ is a closed subset in $\mathbb{R}^m$. For the case where Λ is convex, a complete theory of second-order necessary and sufficient optimality conditions has been developed by Bonnans, Cominetti and Shapiro in [3] and the results have been reviewed in the monograph of Bonnans and Shapiro [5].

*Institute of Computational Mathematics, Johannes Kepler University Linz, A-4040 Linz, Austria, e-mail: helmut.gfrerer@jku.at. This author's research was supported by the Austrian Science Fund (FWF) under grant P29190-N32.

†Department of Mathematics and Statistics, University of Victoria, Victoria, B.C., Canada V8W 2Y2, e-mail: janeye@uvic.ca. The research of this author was partially supported by NSERC.

‡Department of Statistics, School of Mathematics and Statistics, Shandong University of Technology, Zibo 255049, P.R. China, e-mail: jinchuanzhou@163.com. This author's work is supported by National Natural Science Foundation of China (11771255, 11801325) and Young Innovation Teams of Shandong Province (2019KJI013).

In recent years, some important problem classes which can be reformulated in the form of problem (P) with non-convex Λ have attracted much attention from the optimization community. These problems include the mathematical program with complementarity constraints (MPCC) (see e.g. [23]), the mathematical program with second-order cone complementarity constraints (SOC-MPCC) (see e.g. [34]) and the mathematical program with semi-definite cone complementarity constraints (SDC-MPCC) (see e.g. [8]). Unlike the first-order optimality conditions for which much research works have been appeared, there is very little research done with the second-order optimality conditions for MPCC, SOC-MPCC and SDC-MPCC, let alone the general non-convex set-constrained problem (1). The classical second-order necessary optimality condition for MPCCs was given in [32, Theorem 7(1)] under the MPCC strict Mangasarian-Fromovitz constraint qualification (SMFCQ). Some weaker second-order necessary optimality conditions for MPCCs were derived in [21]. For the case when Λ is the union of finitely many convex polyhedral sets, which comprises MPCCs, strong second-order necessary optimality conditions were given in [14] under a directional metric subregularity constraint qualification which is much weaker than SMFCQ. Recently a second-order necessary optimality condition is derived in [7, Theorem 5.1] for SOC-MPCCs under the nondegeneracy condition (equivalently the generalized linear independence constraint qualification, generalized LICQ).

To our knowledge, there is no work dealing with the second-order optimality condition for the general non-convex set-constrained problem in the form (1). The main purpose of this paper is to fill this gap.

In case of non-convex set Λ, first-order necessary optimality conditions can be derived by means of variational analysis. From the work of Bonnans, Cominetti and Shapiro [3], it is well known that the second-order optimality condition must involve in some way the second-order tangent set to the set Λ, and the non-convexity of this second-order tangent set is also an issue. We consider two different approaches for handling non-convex second-order tangent sets. In the first approach, solely based on non-convex variational analysis, we first show that directional metric subregularity of the feasible set mapping carries over to the second-order linearized subproblem. Then we introduce the concept of *lower generalized support function* (which coincides with the support function in the convex case) in order to state the second-order necessary optimality conditions in Theorem 2.

In the second approach, convex duality plays an essential role. We first introduce the *directional regular (Clarke) tangent cone* to Λ and state its relations with the directional limiting normal cone and the second-order tangent set. Using these relations we introduce a new constraint qualification called *directional Robinson's constraint qualification* and we can carry over the ideas already employed in [3] to obtain the second-order necessary optimality conditions of Corollary 4. We show that these second-order conditions are equivalent with primal second-order conditions and are in general stronger than the one obtained with our first approach. However, they also require a stronger constraint qualification and their practical use is limited by the fact that we have not only one condition for every critical direction, but for every convex set contained in the second-order tangent set. If we further strengthen the constraint qualification to some *directional non-degeneracy condition*, this drawback vanishes and we can state a single condition involving the support function of the second-order tangent set.

Second-order optimality condition have in general the form that for every critical direction some condition is fulfilled. It seems to be that Penot [28] was the first who recognized that only some directional form of a constraint qualification (directional metric subregularity) is required for stating the necessary conditions. We pursue this approach and, as a byproduct of the second-order optimality conditions, we introduce and analyze a lot of new directional objects like the directional regular tangent cone, the directional Robinson's constraint qualification and directional non-degeneracy. These results are of its own interest.

We organize our paper as follows. Section 2 contains the preliminaries and preliminary results. In Sections 3 and 4, we derive the primal and dual form of second-order necessary optimality conditions, respectively. Section 5 discusses second-order sufficient conditions for optimality. In Section 6 we present four examples which illustrate our second-order necessary and sufficient conditions.

2 Preliminaries and preliminary results

In this section we clarify the notations, recall some background material we need from variational analysis and develop some preliminary results.

The unit sphere in $\mathbb{R}^n$ is denoted by $\mathbb{S}$ and the open and closed unit balls are denoted by $\mathbb{B}$ and $\mathcal{B}$ respectively. For a set C, denote by $\operatorname{int} C$, $\operatorname{ri} C$, $\operatorname{cl} C$, $\operatorname{bd} C$, $\operatorname{co} C$, $C^{\rm comp}$ its interior, relative interior, closure, boundary, convex hull, and its complement, respectively. For a closed set $C \subseteq \mathbb{R}^n$, let C° and $\sigma_C(x)$ or $\sigma(x|C)$ stand for the polar cone and the support function of C, respectively, i.e., $C^\circ := \{v|\ \langle v, w\rangle \le 0,\ \forall w \in C\}$ and $\sigma_C(x) = \sigma(x|C) := \sup\{\langle x, x'\rangle | x' \in C\}$ for $x \in \mathbb{R}^n$. Let $o(\lambda) : \mathbb{R}_+ \to \mathbb{R}^m$ stand for a mapping with the property that $o(\lambda)/\lambda \to 0$ when $\lambda \downarrow 0$. $z \stackrel{S}{\to} x$ means $z \in S$ and $z \to x$. For a mapping $\Phi : \mathbb{R}^n \to \mathbb{R}^m$, we denote by $\nabla\Phi(x) \in \mathbb{R}^{m\times n}$ the Jacobian and by $\nabla^2\Phi(x)$ the second order derivative as defined by

$$u^T\nabla^2\Phi(x) := \lim_{t\to 0} \frac{\nabla\Phi(x+tu) - \nabla\Phi(x)}{t} \quad \forall u \in \mathbb{R}^n.$$

Hence, for a scalar mapping $f : \mathbb{R}^n \to \mathbb{R}$, $\nabla^2 f(x)$ can be identified with the Hessian and for a mapping $\Phi : \mathbb{R}^n \to \mathbb{R}^m$, we have

$$\nabla^2\Phi(x)(d,d) := d^T\nabla^2\Phi(x)d = (d^T\nabla^2\Phi_1(x)d, \ldots, d^T\nabla^2\Phi_m(x)d)^T \quad \forall d \in \mathbb{R}^n.$$

Let $\Phi : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ be a set-valued mapping. We denote by $\limsup_{x'\to x} \Phi(x')$ and $\liminf_{x'\to x} \Phi(x')$ the Painlevé-Kuratowski upper and lower limit, i.e.,

$$\limsup_{x'\to x} \Phi(x) := \left\{ v \in \mathbb{R}^m \middle| \exists x_k \to x, v_k \to v \text{ with } v_k \in \Phi(x_k) \right\}$$
$$\liminf_{x'\to x} \Phi(x) := \left\{ v \in \mathbb{R}^m \middle| \forall x_k \to x, \exists v_k \to v \text{ with } v_k \in \Phi(x_k) \right\},$$

respectively.

Definition 1 (Tangent cones)**.** *[4, Definitions 2.54 and 3.28] Given $S \subseteq \mathbb{R}^n$ and $x \in S$, the regular/Clarke and (Bouligand-Severi)* tangent/contingent cone *to S at x are defined respectively by*

$$\widehat{T}_S(x) \ := \ \liminf_{\substack{x' \stackrel{S}{\to} x \\ t\downarrow 0}} \frac{S - x'}{t} = \left\{ d \in \mathbb{R}^n \ \middle|\ \forall\, t_k \downarrow 0,\ x_k \stackrel{S}{\to} x,\ \exists d_k \to d \ \text{ with } \ x_k + t_k d_k \in S \right\},$$

$$T_S(x) \ := \ \limsup_{t\downarrow 0} \frac{S - x}{t} = \left\{ d \in \mathbb{R}^n \ \middle|\ \exists\ t_k \downarrow 0,\ d_k \to d \ \text{ with } \ x + t_k d_k \in S \right\}.$$

For $x \in S$ and $d \in T_S(x)$, the outer second-order tangent set to S in the direction d is defined by

$$\begin{aligned} T_S^2(x;d) &:= \limsup_{t\downarrow 0} \frac{S-x-td}{\frac{1}{2}t^2} \\ &= \left\{ w \in \mathbb{R}^n \mid \exists\ t_k \downarrow 0, v_k \to w \text{ such that } x + t_k d + \frac{1}{2}t_k^2 v_k \in S \right\}. \end{aligned}$$

Alternatively, the contingent cone and the second-order tangent set can be written in the form

$$T_S(x) = \left\{ d \in \mathbb{R}^n \,\middle|\, \exists\ t_k \downarrow 0, \operatorname{dist}(x + t_k d, S) = o(t_k) \right\}, \tag{2}$$

$$T_S^2(x;d) = \left\{ w \in \mathbb{R}^n \,\middle|\, \exists\ t_k \downarrow 0, \operatorname{dist}(x + t_k d + \frac{1}{2}t_k^2 w, S) = o(t_k^2) \right\}, \tag{3}$$

respectively; see [5, (2.87) and (3.50)]. The regular tangent cone is always a closed convex cone. The tangent cone is always a closed cone and it is a closed convex cone provided that the set S is convex. However the outer second-order tangent set may be a non-convex set even when the set S is convex (see [5, Example 3.35]). While the tangent cone contains zero always, the second-order tangent set may not be a cone and it may be empty (see e.g. [5, Example 3.29]).

We now introduce a concept of directional regular/Clarke tangent cone which we will need later. The following definition is motivated by the formula $\widehat{T}_S(x) = \liminf_{x' \xrightarrow{S} x} T_S(x')$, whenever S is locally closed at x, cf. [31, Theorem 6.26].

Definition 2 (Directional regular/Clarke tangent cone)**.** *Given $S \subseteq \mathbb{R}^n$, $x \in S$ and $d \in \mathbb{R}^n$, the regular/Clarke tangent cone to S at x in direction d is defined by*

$$\begin{aligned} \widehat{T}_S(x;d) :=& \liminf_{\substack{t\downarrow 0, d' \to d \\ x+td' \in S}} T_S(x+td') \\ =& \left\{ v \in \mathbb{R}^n \,\middle|\, \forall t_k \downarrow 0, d_k \to d, x + t_k d_k \in S, \exists v_k \to v \textit{ with } v_k \in T_S(x + t_k d_k) \right\}. \end{aligned}$$

It is easy to see from definition that for a closed set S the directional version of the regular tangent cone contains the non-directional one and it coincides with the non-directional one when the direction is equal to zero, i.e., $\widehat{T}_S(x;d) \supseteq \widehat{T}_S(x)$ and $\widehat{T}_S(x;0) = \widehat{T}_S(x)$.

We now derive some properties of first and second-order tangent sets. The formula for the second-order tangent set extends the one in [31, Proposition 13.12].

Proposition 1. *Given a closed set $S \subseteq \mathbb{R}^n$, for every $x \in S$ and every $d \in T_S(x)$ one has*

$$T_{T_S(x)}(d) + \widehat{T}_S(x;d) = T_{T_S(x)}(d), \quad T_S^2(x;d) + \widehat{T}_S(x;d) = T_S^2(x;d).$$

Proof. The inclusion $\supseteq$ in both equations is clear, since $0 \in \widehat{T}_S(x;d)$. In order to show the inclusion $\subseteq$ in the first equation, consider $w \in T_{T_S(x)}(d)$ and $v \in \widehat{T}_S(x;d)$ and we prove that $w + v \in T_{T_S(x)}(d)$ by contradiction. To the contrary, suppose that $w + v \notin T_{T_S(x)}(d)$. Then by virtue of (2) there is some $\epsilon > 0$ and some $\bar{t} > 0$ such that

$$\operatorname{dist}\big(d + t(w+v), T_S(x)\big) \geq 4\epsilon t \quad \forall t \in (0, \bar{t}).$$

Consequently, $d + t(w + v) \notin T_S(x)$ and hence for every $t \in (0, \bar{t})$ there is some $\bar{\alpha}_t > 0$ with

$$\text{dist}\big(x + \alpha(d + t(w + v)), S\big) \geq 3\epsilon\alpha t \quad \forall \alpha \in (0, \bar{\alpha}_t).$$

Since $w \in T_{T_S(x)}(d)$, by definition there are sequences $t_k \downarrow 0$ and $w_k \to w$ such that $d + t_k w_k \in T_S(x)$ for all k. Thus, for every k there is a sequence $\alpha_i^k \downarrow 0$ and $d_i^k \to d$, as $i \to \infty$ satisfying $x + \alpha_i^k(d_i^k + t_k w_k) \in S$ for all i. For every k sufficiently large we have $t_k < \bar{t}$, $\|w_k - w\| < \epsilon$ and we can find some index $i(k)$ such that $\alpha_{i(k)}^k < \min\{\frac{1}{k}, \bar{\alpha}_{t_k}\}$ and $\|d_{i(k)}^k - d\| < \epsilon t_k$. It follows together with Lipschitz property of the distance function that

$$\begin{aligned}
&\text{dist}\big(x + \alpha_{i(k)}^k(d_{i(k)}^k + t_k(w_k + v)), S\big) \\
&\geq \text{dist}\big(x + \alpha_{i(k)}^k(d + t_k(w + v)), S\big) - \alpha_{i(k)}^k(\|d_{i(k)}^k - d\| + t_k\|w_k - w\|) \\
&> 3\epsilon\alpha_{i(k)}^k t_k - 2\epsilon\alpha_{i(k)}^k t_k = \epsilon\alpha_{i(k)}^k t_k
\end{aligned}$$

implying

$$(x_k + \tau_k(v + \epsilon\mathbb{B})) \cap S = \emptyset$$

with $x_k := x + \alpha_{i(k)}^k(d_{i(k)}^k + t_k w_k) \in S$ and $\tau_k := \alpha_{i(k)}^k t_k$. By the proof of [31, Theorem 6.26], there exists $\tilde{x}_k \in S \cap (x_k + \tau_k(\|v\| + \epsilon)\mathbb{B})$ such that

$$\text{dist}(v, T_S(\tilde{x}_k)) \geq \epsilon. \tag{4}$$

Since

$$\begin{aligned}
\|\tilde{x}_k - (x + \alpha_{i(k)}^k d)\| &\leq \|\tilde{x}_k - x_k\| + \|x_k - (x + \alpha_{i(k)}^k d)\| \\
&\leq \alpha_{i(k)}^k \Big(t_k(\|v\| + \epsilon) + \|d_{i(k)}^k - d\| + t_k\|w_k\|\Big) \\
&= o(\alpha_{i(k)}^k),
\end{aligned}$$

we have $\tilde{x}_k = x + \alpha_{i(k)}^k d_k$ with some sequence $d_k \to d$. Together with (4) this implies that $v \notin \widehat{T}_S(x; d)$ which contradicts the assumption that $v \in \widehat{T}_S(x; d)$ and hence we have proved that $w + v \in T_{T_S(x)}(d)$. Indeed, to the contrary if $v \in \widehat{T}_S(x; d)$, then by definition of the directional regular tangent cone, for the sequence $\alpha_{i(k)}^k \downarrow 0, d_k \to d$, $\tilde{x}_k \in S$, there must exists a sequence $v_k \to v$ with $v_k \in T_S(\tilde{x}_k)$, contradicting (4).

We show the inclusion $\subseteq$ in the second equation in a similar way. Let $w \in T_S^2(x; d), v \in \widehat{T}_S(x; d)$ and we prove that $w + v \in T_S^2(x; d)$ by contradiction. To the contrary, suppose that $w + v \notin T_S^2(x; d)$. Then by virtue of (3) there exists $\epsilon > 0, t_k \downarrow 0, w_k \to w$ such that $x + t_k d + \frac{1}{2}t_k^2 w_k \in S$ and

$$\text{dist}(x + t_k d + \frac{1}{2}t_k^2(w_k + v), S) \geq \frac{1}{2}t_k^2\epsilon.$$

Denote by $x_k := x + t_k d + \frac{1}{2}t_k^2 w_k, \tau_k := \frac{1}{2}t_k^2$, the above inequality is equivalent to saying that

$$(x_k + \tau_k(v + \epsilon\mathbb{B})) \cap S = \emptyset.$$

Now we can proceed similar as before to obtain the contradiction $v \notin \widehat{T}_S(x; d)$. $\square$

Definition 3 (Normal Cones). *(See e.g. [25]) Given $S \subseteq \mathbb{R}^n$ and $x \in S$, the regular/Fréchet normal cone to S at x is given by*

$$\widehat{N}_S(x) := \left\{ v \in \mathbb{R}^n \,\middle|\, \langle v, x' - x \rangle \leq o\big(\|x' - x\|\big),\ \forall x' \in S \right\};$$

the limiting/Mordukhovich normal cone to S at x is defined as

$$N_S(x) := \limsup_{x' \xrightarrow{S} x} \widehat{N}_S(x'),$$

and the Clarke normal cone to S at x is $N_S^c(x) := \operatorname{cl}\operatorname{co} N_S(x)$.

The limiting normal cone is in general non-convex whereas the Fréchet normal cone is always convex. In the case of a convex set S, both the Fréchet normal cone and the limiting normal cone coincide with the normal cone in the sense of convex analysis, i.e.,

$$N_S(x) := \left\{ v \in \mathbb{R}^n \,\middle|\, \langle v, x' - x \rangle \leq 0,\ \forall x' \in S \right\}.$$

Recently a directional version of limiting normal cones were introduced in [11] and extended to general Banach spaces in [13].

Definition 4 (Directional Limiting Normal Cones). *Given a set $S \subseteq \mathbb{R}^n$, a point $x \in S$ and a direction $d \in \mathbb{R}^n$, the limiting normal cone to S in direction d at x is defined by*

$$N_S(x; d) := \limsup_{t \downarrow 0, d' \to d} \widehat{N}_S(x + td') = \left\{ v | \exists t_k \downarrow 0, d_k \to d, v_k \to v \text{ with } v_k \in \widehat{N}_S(x + t_k d_k) \right\}.$$

From definition, it is obvious that $N_S(x; d) = \emptyset$ if $d \notin T_S(x)$, $N_S(x; d) \subseteq N_S(x)$, and $N_S(x; 0) = N_S(x)$. When S is convex and $d \in T_S(x)$ there holds

$$N_S(x; d) = N_S(x) \cap \{d\}^{\perp} = N_{T_S(x)}(d), \tag{5}$$

cf. [14, Lemma 2.1]. The following result is the directional counterpart of the fact that the limiting normal cone mapping is outer semicontinuous (see e.g.[31, Proposition 6.6]).

Proposition 2. *Given a set $S \subseteq \mathbb{R}^n$, a point $x \in S$ and a direction $d \in \mathbb{R}^n$, one has*

$$N_S(x; d) = \limsup_{t \downarrow 0, d' \to d} N_S(x + td') = \left\{ v | \exists t_k \downarrow 0, d_k \to d, v_k \to v \text{ with } v_k \in N_S(x + t_k d_k) \right\}.$$

Proof. The inclusion $N_S(x; d) \subseteq \limsup_{t \downarrow 0, d' \to d} N_S(x + td')$ follows easily from the fact that for every x' we have $\widehat{N}_S(x') \subseteq N_S(x')$. In order to show the reverse inclusion, consider sequences $t_k \downarrow 0$, $d_k \to d$ and $v_k \to v$ with $v_k \in N_S(x + t_k d_k)$. By the definition of limiting normals, for every k there exist sequences $x_k^i \to x + t_k d_k$ and $v_k^i \to v_k$ as $i \to \infty$ with $v_k^i \in \widehat{N}_S(x_k^i)$. Using a standard diagonal process, for every k we can find some index $i(k)$ satisfying

$$\|x_k^{i(k)} - (x + t_k d_k)\| \leq \frac{t_k}{k},\ \|v_k^{i(k)} - v_k\| \leq \frac{1}{k}.$$

Setting $d_k' := (x_k^{i(k)} - x)/t_k$, it follows that $\|d_k' - d_k\| \leq \frac{1}{k}$ and consequently $d_k' \to d$. Since $\lim_{k\to\infty} v_k^{i(k)} = \lim_{k\to\infty} v_k = v$ and $v_k^{i(k)} \in \widehat{N}_S(x + t_k d_k')$, $v \in N_S(x; d)$ follows. Hence, the inclusion $N_S(x; d) \supseteq \limsup_{t \downarrow 0, d' \to d} N_S(x + td')$ is also established and the proof is complete. □

From the definition of the Clarke normal cone in Definition 3, it is natural to define the directional Clarke normal cone as follows.

Definition 5 (Directional Clarke Normal Cones)**.** *Given a set $S \subseteq \mathbb{R}^n$, a point $x \in S$ and a direction $d \in \mathbb{R}^n$, the Clarke normal cone to S in direction d at x is defined by*

$$N_S^c(x; d) := \operatorname{cl} \operatorname{co} N_S(x; d).$$

Similarly to the directional limiting normal cone, we also have $N_S^c(x; d) = \emptyset$ if $d \notin T_S(x)$, $N_S^c(x; d) \subseteq N_S^c(x)$ and $N_S^c(x; 0) = N_S^c(x)$.

Similar to the standard tangent-normal polarity (see [31, Theorem 6.28], [6]), we have the following directional tangent-normal polarity.

Proposition 3 (Directional Tangent-Normal Polarity)**.** *For a closed set S, $x \in S$, and $d \in \mathbb{R}^n$, one has*

$$\widehat{T}_S(x; d) = N_S(x; d)^\circ = N_S^c(x; d)^\circ, \quad \widehat{T}_S(x; d)^\circ = N_S^c(x; d).$$

In particular, the directional regular tangent cone $\widehat{T}_S(x; d)$ is closed and convex.

Proof. Firstly we show $\widehat{T}_S(x; d) \subseteq N_S(x; d)^\circ$. For any given $w \in \widehat{T}_S(x; d)$, take $v \in N_S(x; d)$. By the definition of directional normal cone, there exist $v_n \to v$ with $v_n \in \widehat{N}_S(x + t_n d_n)$ for some $t_n \downarrow 0$ and $d_n \to d$ and $x + t_n d_n \in S$. Since $w \in \widehat{T}_S(x; d)$, for this sequence, by the definition of directional regular tangent cone, there exists $w_n \to w$ with $w_n \in T_S(x + t_n d_n)$. It follows that $\langle w_n, v_n \rangle \leq 0$ since $v_n \in \widehat{N}_S(x + t_n d_n) = (T_S(x + t_n d_n))^\circ$. Taking the limit yields $\langle w, v \rangle \leq 0$, which implies that $w \in N_S(x; d)^\circ$. Hence $\widehat{T}_S(x; d) \subseteq N_S(x; d)^\circ$.

Secondly we show $\widehat{T}_S(x; d) \supseteq N_S(x; d)^\circ$. Suppose that $w \notin \widehat{T}_S(x; d)$. Then there exist $t_n \downarrow 0$, $d_n \to d$, $x + t_n d_n \in S$ such that $w \notin \liminf_{n \to \infty} T_S(x + t_n d_n)$. Hence by [31, Exercise 4.2(a)], $\limsup_{n \to \infty} d(w, T_S(x + t_n d_n)) > 0$, i.e., there exists $\epsilon > 0$ and some subsequence $\mathcal{N} \subseteq \{1, 2, \ldots, \}$ such that $\operatorname{dist}(w, T_S(x + t_n d_n)) > \epsilon$ for all $n \in \mathcal{N}$. According to [31, Proposition 6.27(b)], there exists $v_n \in N_S(\bar{x} + t_n d_n) \cap \mathbb{S}$ such that

$$\langle w, v_n \rangle = \operatorname{dist}(w, T_S(\bar{x} + t_n d_n)) > \epsilon \tag{6}$$

for all $n \in \mathcal{N}$. Since v_n is bounded, we can assume by further taking subsequence if necessary that $v_n \to v$ (as $n \to \infty$ and $n \in \mathcal{N}$). Clearly by Proposition 2, $v \in N_S(\bar{x}; d)$. Hence we obtain $\langle w, v \rangle \geq \epsilon$ by (6). So $w \notin N_S(x; d)^\circ$.

Thus we have shown $\widehat{T}_S(x; d)^{\text{comp}} \subseteq (N_S(x; d)^\circ)^{\text{comp}}$ and the inclusion $\widehat{T}_S(x; d) \supseteq N_S(x; d)^\circ$ follows. We conclude $\widehat{T}_S(x; d) = N_S(x; d)^\circ$ and therefore the directional regular tangent cone is closed and convex as a polar cone.

The rest of proofs follow from the fact that the directional Clarke normal cone is closed and convex as well. $\square$

In this paper, we rely on the following stability property of a set-valued map in developing our results.

Definition 6. *[13, Definition 1] Let $\varphi : \mathbb{R}^n \to \mathbb{R}^m$, $C \subseteq \mathbb{R}^m$ and $\varphi(\bar{x}) \in C$. We say that the set-valued map $M(x) := \varphi(x) - C$ is metrically subregular (MS) at $(\bar{x}, 0)$ in direction $d \in \mathbb{R}^n$, if there exist $\kappa, \rho, \delta > 0$ such that*

$$\operatorname{dist}(x, M^{-1}(0)) \leq \kappa \operatorname{dist}(\varphi(x), C), \quad \forall x \in \bar{x} + V_{\rho,\delta}(d), \tag{7}$$

where

$$V_{\rho,\delta}(d) \quad := \quad \left\{ w \in \rho\mathbb{B} \,\middle|\, \big\| \|d\| w - \|w\| d \big\| \leq \delta \|w\| \|d\| \right\}$$

$$= \begin{cases} \rho\mathbb{B} & \text{if } d = 0 \\ \{0\} \cup \{w \in \rho\mathbb{B} \backslash \{0\} | \| \frac{w}{\|w\|} - \frac{d}{\|d\|} \| \le \delta\} & \text{if } d \neq 0 \end{cases}$$

is a directional neighborhood of the direction d. In the case where $d = 0$, we simple say that M is metrically subregular at $(\bar{x}, 0)$ or metric subregularity constraint qualification (MSCQ) holds at $\bar{x}$.

The infimum of κ over all such combinations of κ, ρ and δ fulfilling (7) is called the modulus of the respective property.

It is well-known that the metric subregularity of a set-valued map M at $(\bar{x}, 0)$ is equivalent to the property of calmness/pseudo upper-Lipschitz continuity of the inverse map M^{-1} at $(0, \bar{x})$; see [31, 33] for definition and [9] for discussions about the equivalence.

Proposition 4. *[16, Theorem 1] Let $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ be continuously differentiable, $C \subseteq \mathbb{R}^m$ be closed and $\varphi(\bar{x}) \in C$. The set-valued map $M(x) := \varphi(x) - C$ is metrically subregular at $(\bar{x}, 0)$ in direction d satisfying $\nabla\varphi(\bar{x})d \in T_C(\varphi(\bar{x}))$ if the first order sufficient condition for metric subregularity (FOSCMS) for direction d holds:*

$$\nabla\varphi(\bar{x})^T\lambda = 0,\ \lambda \in N_C(\varphi(\bar{x}); \nabla\varphi(\bar{x})d) \implies \lambda = 0.$$

Recently, weaker sufficient conditions than FOSCMS such as the directional quasi/pseudo-normality was introduced in [1]. More sufficient conditions based on directional normal cones or/and for specific systems can be found e.g. in [16, 19, 35].

Classical sufficient conditions for metric subregularity include the case where φ is affine and C is the union of finitely many polyhedral sets by Robinson's polyhedral multifunction theory [29] and the no nonzero abnormal multiplier constraint qualification (NNAMCQ) holds:

$$\nabla\varphi(\bar{x})^T\lambda = 0,\ \lambda \in N_C(\varphi(\bar{x})) \Longrightarrow \lambda = 0,$$

by Mordukhovich criteria for metric regularity (see e.g., [31, Theorem 9.40]). Note that when C is convex, by [5, Corollary 2.98], NNAMCQ is equivalent to Robinson's constraint qualification [5, (2.178)]

$$0 \in \text{int}\{\varphi(\bar{x}) + \nabla\varphi(\bar{x})\mathbb{R}^n - C\},$$

which in turn is equivalent to

$$\nabla\varphi(\bar{x})\mathbb{R}^n + T_C(\varphi(\bar{x})) = \mathbb{R}^m$$

in finite dimensions.

In the following result, we show that the directional metric subregularity of the mapping $M(x) := \varphi(x) - C$ is carried over to its linearized mapping, the so-called graphical derivative. Recall that for a set-valued mapping $M : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ the *graphical derivative* to M at a point $(\bar{x}, \bar{y}) \in \text{gph}\, M$ is the mapping $DM(\bar{x}, \bar{y}) : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ satisfying

$$\text{gph}\, DM(\bar{x}, \bar{y}) = T_{\text{gph}\, M}(\bar{x}, \bar{y}),$$

resulting in $D(\varphi(\cdot) - C)(\bar{x}, \bar{y})(w) = \nabla\varphi(\bar{x})w - T_C(\varphi(\bar{x}) - \bar{y})$.

Lemma 1. *Let $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ be continuously differentiable, $C \subseteq \mathbb{R}^m$ be closed and $\varphi(\bar{x}) \in C$. If $M(x) := \varphi(x) - C$ is metrically subregular at $(\bar{x}, 0)$ in direction*

$$d \in \nabla\varphi(\bar{x})^{-1}(T_C(\varphi(\bar{x})) := \{d | \nabla\varphi(\bar{x})d \in T_C(\varphi(\bar{x}))\}$$

with modulus $\bar{\kappa}$, then the graphical derivative $DM(\bar{x}, 0)(w) = \nabla\varphi(\bar{x})w - T_C(\varphi(\bar{x}))$ is metrically subregular at $(d, 0)$ with modulus no larger than $\bar{\kappa}$.

Proof. Choose $\kappa, \rho, \delta > 0$ such that (7) holds and consider $w \in V_{\rho,\delta}(d)$ together with $v \in T_C(\varphi(\bar{x}))$ satisfying $\text{dist}\big(\nabla\varphi(\bar{x})w, T_C(\varphi(\bar{x}))\big) = \|\nabla\varphi(\bar{x})w - v\|$. Then $tw \in V_{\rho,\delta}(d)$ for all $t \in [0,1]$ and, by picking a sequence $t_k \downarrow 0$ with $\text{dist}(\varphi(\bar{x}) + t_k v, C) = o(t_k)$, we have

$$\begin{array}{rcl} \text{dist}(\bar{x} + t_k w, M^{-1}(0)) & \leq & \kappa \text{dist}(\varphi(\bar{x} + t_k w), C) \\ & = & \kappa\big(\text{dist}(\varphi(\bar{x}) + t_k \nabla\varphi(\bar{x})w + o(t_k), C)\big) \\ & = & \kappa\big(\text{dist}(\varphi(\bar{x}) + t_k \nabla\varphi(\bar{x})w, C) + o(t_k)\big) \\ & \leq & \kappa\big(t_k \|\nabla\varphi(\bar{x})w - v\| + o(t_k)\big), \end{array}$$

where the second equality follows from the Lipschitz property of the distance function.

Thus we can find a sequence x_k satisfying $\varphi(x_k) \in C$ and

$$\|x_k - (\bar{x} + t_k w)\| \leq \kappa\big(t_k \|\nabla\varphi(\bar{x})w - v\| + o(t_k)\big). \tag{8}$$

It follows that $(x_k - \bar{x})/t_k$ is bounded and, by possibly passing to a subsequence, we may assume that $(x_k - \bar{x})/t_k$ converges to some w'. Dividing (8) by t_k and passing to the limit we obtain $\|w - w'\| \leq \kappa\|\nabla\varphi(\bar{x})w - v\|$. Further,

$$\text{dist}(\varphi(\bar{x}) + t_k \nabla\varphi(\bar{x})w', C) \leq \|\varphi(\bar{x}) + t_k \nabla\varphi(\bar{x})w' - \varphi(x_k)\| = o(t_k)$$

showing $\nabla\varphi(\bar{x})w' \in T_C(\varphi(\bar{x}))$ and thus $w' \in DM(\bar{x}, 0)^{-1}(0)$. Since the directional neighborhood $V_{\rho,\delta}(d)$ is also a neighborhood of d in the classical sense, we can find some $\rho' > 0$ such that $d + V_{\rho',\delta}(0) = d + \rho'\mathbb{B} \subseteq V_{\rho,\delta}(d)$. Thus we have shown that for all $w \in d + V_{\rho',\delta}(0)$ there holds

$$\text{dist}(w, DM(\bar{x}, 0)^{-1}(0)) \leq \|w - w'\| \leq \kappa\|\nabla\varphi(\bar{x})w - v\| = \kappa \text{dist}\big(\nabla\varphi(\bar{x})w, T_C(\varphi(\bar{x}))\big)$$

showing metric subregularity of $DM(\bar{x}, 0)$ at $(d, 0)$ □

2.1 Uniform MSCQ for the second-order linearized mapping

From now on we denote by $\mathcal{F}$ the feasible region of problem (P), i.e., $\mathcal{F} := \{x \,|\, g(x) \in \Lambda\} = M^{-1}(0)$, where $M(x) := g(x) - \Lambda$ is the feasible mapping. If the feasible mapping $M(x)$ is metrically subregular at $(\bar{x}, 0)$, then

$$T_{\mathcal{F}}(\bar{x}) = DM(\bar{x}, 0)^{-1}(0) = \{d \,|\, \nabla g(\bar{x})d \in T_\Lambda(g(\bar{x}))\},$$

see, e.g., [22, Proposition 1] or [18, Corollary 4.2], where $DM(\bar{x}, 0)(d) = \nabla g(\bar{x})d - T_\Lambda(g(\bar{x}))$ denotes the graphical derivative of M at $(\bar{x}, 0)$. Moreover, by [15, Lemmas 3 and 4] (a variant of [12, Proposition 2.1]), there is some $\kappa > 0$ such that

$$\text{dist}(d, T_{\mathcal{F}}(\bar{x})) \leq \kappa \text{dist}(0, DM(\bar{x}, 0)(d)) = \kappa \text{dist}(\nabla g(\bar{x})d, T_\Lambda(g(\bar{x}))) \quad \forall d \in \mathbb{R}^n,$$

which is some kind of uniform metric subregularity of the graphical derivative. We will now show that an analogous relation holds for the second-order tangent set $T^2_{\mathcal{F}}(\bar{x}; d)$ and the second-order linearized mapping

$$D^2M(\bar{x}, 0; d)(w) := \nabla g(\bar{x})w + \nabla^2 g(\bar{x})(d, d) - T^2_\Lambda\big(g(\bar{x}); \nabla g(\bar{x})d\big). \tag{9}$$

To prove the result, we need the following lemma.

Lemma 2. *For any $x \in \mathbb{R}^n$ and a set-valued mapping $C : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$, one has*

$$\liminf_{u' \to u} \operatorname{dist}(x, C(u')) = \operatorname{dist}(x, \limsup_{u' \to u} C(u')).$$

Proof. Let $\{u_n\}$ be a sequence satisfying

$$\liminf_{n \to \infty} \operatorname{dist}(x, C(u_n)) = \liminf_{u' \to u} \operatorname{dist}(x, C(u')).$$

Then according to [31, Exercise 4.8] we have

$$\liminf_{n \to \infty} \operatorname{dist}(x, C(u_n)) = \operatorname{dist}(x, \limsup_{n \to \infty} C(u_n)) \geq \operatorname{dist}(x, \limsup_{u' \to u} C(u')).$$

Hence

$$\liminf_{u' \to u} \operatorname{dist}(x, C(u')) \geq \operatorname{dist}(x, \limsup_{u' \to u} C(u')).$$

Conversely, take r satisfying $r > \operatorname{dist}(x, \limsup_{u' \to u} C(u'))$. Then there exists $x' \in \limsup_{u' \to u} C(u')$ such that $\|x - x'\| < r$. Since $x' \in \limsup_{u' \to u} C(u')$, then there exists $u_n \to u$ and $x'_n \in C(u_n)$ and $x'_n \to x'$. So $\|x - x'_n\| < r$ as n large enough. Hence $\operatorname{dist}(x, C(u_n)) \leq \|x - x'_n\| < r$. So

$$r \geq \liminf_{n \to \infty} \operatorname{dist}(x, C(u_n)) \geq \liminf_{u' \to u} \operatorname{dist}(x, C(u')).$$

Due to the arbitrariness of $r > \operatorname{dist}(x, \limsup_{u' \to u} C(u'))$, we obtain

$$\operatorname{dist}(x, \limsup_{u' \to u} C(u')) \geq \liminf_{u' \to u} \operatorname{dist}(x, C(u')).$$

□

Proposition 5. *Let $\bar{x} \in \mathcal{F}$ and suppose that the set-valued map $M(x) := g(x) - \Lambda$ is metrically subregular at $(\bar{x}, 0)$ in direction d with modulus κ. Then*

$$d \in T_{\mathcal{F}}(\bar{x}) \;\Leftrightarrow\; \nabla g(\bar{x})d \in T_\Lambda(g(\bar{x})) \tag{10}$$

and for any d satisfying $g(\bar{x})d \in T_\Lambda(g(\bar{x}))$, one has

$$T^2_{\mathcal{F}}(\bar{x}; d) = \{w \mid \nabla g(\bar{x})w + \nabla^2 g(\bar{x})(d, d) \in T^2_\Lambda\big(g(\bar{x}); \nabla g(\bar{x})d\big)\} = D^2 M(\bar{x}, 0; d)^{-1}(0) \tag{11}$$

with $D^2 M(\bar{x}, 0; d)$ given by (9). Moreover,

$$\begin{aligned} \operatorname{dist}(w, T^2_{\mathcal{F}}(x; d)) &\leq \kappa \operatorname{dist}\Big(\nabla g(\bar{x})w + \nabla g^2(\bar{x})(d, d), T^2_\Lambda\big(g(\bar{x}); \nabla g(\bar{x})d\big)\Big) \\ &= \kappa \operatorname{dist}\big(0, D^2 M(\bar{x}, 0; d)(w)\big) \;\forall w \in \mathbb{R}^n. \end{aligned} \tag{12}$$

Proof. The equivalence (10) follows from [18, Proposition 4.1]. Next we will prove the inequality (12). Let $w \in \mathbb{R}^n$ be fixed and consider $\kappa' > \kappa$. It follows that for $t > 0$ sufficiently small

$$\begin{aligned} &\operatorname{dist}(\bar{x} + td + \frac{1}{2}t^2 w, \mathcal{F}) \\ &\leq \;\; \kappa' \operatorname{dist}(g(\bar{x} + td + \frac{1}{2}t^2 w), \Lambda) \end{aligned}$$

$$
\begin{aligned}
&= \kappa' \mathrm{dist}\Big(g(\bar{x}) + t\nabla g(\bar{x})d + \frac{1}{2}t^2\big(\nabla g(\bar{x})w + \nabla^2 g(\bar{x})(d,d)\big) + o(t^2), \Lambda\Big) \\
&= \kappa' \mathrm{dist}\Big(g(\bar{x}) + t\nabla g(\bar{x})d + \frac{1}{2}t^2\big(\nabla g(\bar{x})w + \nabla^2 g(\bar{x})(d,d)\big), \Lambda\Big) + o(t^2),
\end{aligned}
$$

where the first and second equalities follow from Taylor expansion and the Lipschitz continuity of the distance function. Dividing both sides of the above inequality by $\frac{1}{2}t^2$ we obtain

$$
\mathrm{dist}\left(w, \frac{\mathcal{F} - \bar{x} - td}{\frac{1}{2}t^2}\right) \le \kappa' \mathrm{dist}\left(\nabla g(\bar{x})w + \nabla^2 g(\bar{x})(d,d), \frac{\Lambda - g(\bar{x}) - t\nabla g(\bar{x})d}{\frac{1}{2}t^2}\right) + \frac{o(t^2)}{\frac{1}{2}t^2}.
$$

Taking the inf-limits on the both sides of the above inequality and using Lemma 2 yields

$$
\begin{aligned}
&\mathrm{dist}(w, T^2_{\mathcal{F}}(\bar{x}; d)) \\
&= \mathrm{dist}\left(w, \limsup_{t\downarrow 0} \frac{\mathcal{F} - \bar{x} - td}{\frac{1}{2}t^2}\right) = \liminf_{t\downarrow 0} \mathrm{dist}\left(w, \frac{\mathcal{F} - \bar{x} - td}{\frac{1}{2}t^2}\right) \\
&\le \kappa' \liminf_{t\downarrow 0} \mathrm{dist}\left(\nabla g(\bar{x})w + \nabla^2 g(\bar{x})(d,d), \frac{\Lambda - g(\bar{x}) - t\nabla g(\bar{x})d}{\frac{1}{2}t^2}\right) \\
&= \kappa' \mathrm{dist}\left(\nabla g(\bar{x})w + \nabla^2 g(\bar{x})(d,d), \limsup_{t\downarrow 0} \frac{\Lambda - g(\bar{x}) - t\nabla g(\bar{x})d}{\frac{1}{2}t^2}\right) \\
&= \kappa' \mathrm{dist}\Big(\nabla g(\bar{x})w + \nabla^2 g(\bar{x})(d,d), T^2_{\Lambda}(g(\bar{x}); \nabla g(\bar{x})d)\Big).
\end{aligned}
$$

Since $\kappa' > \kappa$ can be chosen arbitrarily close to κ, the bound (12) follows. From (12) we may conclude

$$
T^2_{\mathcal{F}}(\bar{x}; d) \supseteq \{w \mid \nabla g(\bar{x})w + \nabla^2 g(\bar{x})(d,d) \in T^2_{\Lambda}\big(g(\bar{x}); \nabla g(\bar{x})d\big)\}.
$$

There remains to show the reverse inclusion. Consider $w \in T^2_{\mathcal{F}}(\bar{x}; d)$ together with sequences $t_k \downarrow 0$ and $w_k \to w$ with $\bar{x} + t_k d + \frac{1}{2}t_k^2 w_k \in \mathcal{F}$. Then

$$
g(\bar{x} + t_k d + \frac{1}{2}t_k^2 w_k) = g(\bar{x}) + t_k \nabla g(\bar{x})d + \frac{1}{2}t_k^2\big(\nabla g(\bar{x})w + \nabla^2 g(\bar{x})(d,d)\big) + o(t_k^2) \in \Lambda
$$

implying $\nabla g(\bar{x})w + \nabla^2 g(\bar{x})(d,d) \in T^2_{\Lambda}\big(g(\bar{x}); \nabla g(\bar{x})d\big)$. This shows the inclusion $\subseteq$ in (11) and the proof of the proposition is complete. □

The chain rule (11) was derived in [31, Proposition 13.13] under the assumption of metric regularity of M and in [24] under (non-directional) metric subregularity, see also [26]. For a related result under directional metric subregularity we refer to [28, Proposition 4.1].

It is known that the second-order tangent set may be empty even if the set considered is a convex set [5]. As a consequence of Proposition 5, we can show that $T^2_{\mathcal{F}}$ and T^2_{Λ} are either empty or nonempty at the same time under the metric subregularity.

Corollary 1. *Suppose that the set-valued map $M(x) := g(x) - \Lambda$ is metrically subregular at $(\bar{x}, 0)$ in direction d with $\bar{x} \in \mathcal{F}$ and $\nabla g(\bar{x})d \in T_{\Lambda}(g(\bar{x}))$. Then $T^2_{\mathcal{F}}(x; d) \neq \emptyset$ if and only if $T^2_{\Lambda}\big(g(\bar{x}); \nabla g(\bar{x})d\big) \neq \emptyset$.*

Proof. Suppose first that $T^2_{\Lambda}\big(g(\bar{x}); \nabla g(\bar{x})d\big) \neq \emptyset$. Then $T^2_{\mathcal{F}}(x; d) \neq \emptyset$ by virtue of (12), because otherwise the left-hand side of (12) must be equal to infinity while the right-hand side is finite which is impossible. The reverse statement follows immediately from (11). □

2.2 On estimates of the normal cone of the tangent sets

In this subsection we give some estimates on the limiting normal cone to the first and the second-order tangent set which will be used in the necessary optimality condition we are developing.

Lemma 3. *Let S be a closed subset in $\mathbb{R}^n$, $x \in S$, $d \in T_S(x)$ and $w \in T_S^2(x,d)$. Then*

$$N_{T_S(x)}(d) \subseteq N_S(x;d), \tag{13}$$

$$N_{T_S^2(x;d)}(w) \subseteq N_S(x;d). \tag{14}$$

Proof. We only prove (14) since (13) can be proved similarly. Take $v \in N_{T_S^2(x,d)}(w)$. Note that $T_S^2(x;d) = \limsup_{t\downarrow 0} \frac{S-x-td}{\frac{1}{2}t^2}$. It follows from [31, Exercise 6.18] that there exists $t_k \downarrow 0$, $w_k \in T_k := \frac{S-x-t_k d}{\frac{1}{2}t_k^2}$ and $v_k \in \widehat{N}_{T_k}(w_k)$ with $v_k \to v$ and $w_k \to w$. Since $T_k = \{w | x + t_k d + \frac{1}{2}t_k^2 w \in S\}$, by the change of coordinates formula in [31, Exercise 6.7] we have $v_k \in \widehat{N}_{T_k}(w_k) = \widehat{N}_S\big(x + t_k d + \frac{1}{2}t_k^2 w_k\big) = \widehat{N}_S\big(x + t_k(d + \frac{1}{2}t_k w_k)\big)$. Hence $v \in N_S(x;d)$. □

The inclusion (13) can be strict. For example, take $S := \{0, 1, \frac{1}{2}, \frac{1}{n}, \dots\}$. It is easy to see that $T_S(0) = \mathbb{R}_+$ and $N_{T_S(0)}(0) = \mathbb{R}_-$. Take $d = 1 \in T_S(0)$. Then $N_S(0;d) \supseteq \limsup_{n\to\infty} \widehat{N}_S(\frac{1}{n}) = \mathbb{R}$, since $\frac{1}{n}$ is an isolated point in S. So $N_S(0;d) = \mathbb{R}$. Hence $\{0\} = N_{T_S(x)}(d) \subsetneq N_S(x;d) = \mathbb{R}$ at $x = 0$.

By (5), the inclusion (13) holds as an equality whenever S is convex. However (14) can fail to be an equality even if S is polyhedral; see e.g. Example 1.

Example 1. *Let $S = \mathbb{R}^2_+$ and $x = (0,0)$. Then $T_S(x) = S$. Take $d = (1,0) \in T_S(x)$. Then $T_S^2(x;d) = T_{T_S(x)}(d) = T_S(d) = \mathbb{R} \times \mathbb{R}_+$. Take $w = (0,1) \in T_S^2(x;d)$. Then $N_{T_S^2(x;d)}(w) = \{(0,0)\}$ and $N_{T_S(x)}(d) = N_S(d) = \{0\} \times \mathbb{R}_-$. So $N_{T_S^2(x;d)}(w) \subsetneq N_{T_S(x)}(d)$.*

3 Primal form of second-order necessary optimality conditions

In this section we derive the primal form of second-order necessary optimality conditions for the general problem (P) under directional metric subregularity. Recall the following basic second-order necessary condition.

Theorem 1 (cf. [27, Corollary 1.3])**.** *Let $\bar{x}$ be a locally optimal solution of (P). Then for all $d \in T_{\mathcal{F}}(\bar{x})$ with $\nabla f(\bar{x})d = 0$ one has*

$$\nabla f(\bar{x})w + \nabla^2 f(\bar{x})(d,d) \geq 0, \quad \forall w \in T_{\mathcal{F}}^2(\bar{x};d).$$

We shall now apply this theorem under the assumption of directional metric subregularity. Define the critical cone at $\bar{x}$ as:

$$C(\bar{x}) := \Big\{ d \,|\, \nabla g(\bar{x})d \in T_\Lambda(g(\bar{x})), \nabla f(\bar{x})d \leq 0 \Big\}.$$

Lemma 4. *Let $\bar{x}$ be a locally optimal solution of (P). Suppose that the feasible set mapping $g(x) - \Lambda$ is metrically subregular at $(\bar{x}, 0)$ in direction $d \in C(\bar{x})$. Then $\nabla f(\bar{x})d = 0$.*

Proof. Since $\bar{x}$ is a locally optimal solution of (P), we have

$$\nabla f(\bar{x})d \geq 0 \qquad \forall d \in T_{\mathcal{F}}(\bar{x}).$$

By Proposition 5 we have $d \in T_{\mathcal{F}}(\bar{x}) \iff \nabla g(\bar{x})d \in T_\Lambda(g(\bar{x}))$ and hence $\nabla f(\bar{x})d \geq 0$. By definition of the critical cone, it follows that $\nabla f(\bar{x})d = 0$. $\square$

The following second-order necessary optimality condition in primal form follows now immediately from Lemma 4, Theorem 1 and Proposition 5. It improves [5, Lemma 3.44] in that Λ does not need to be convex and the result holds under the directional metric subregularity instead of Robinson's constraint qualification.

Corollary 2. *Let $\bar{x}$ be a locally optimal solution of (P). Suppose that the feasible set mapping $g(x) - \Lambda$ is metrically subregular at $(\bar{x}, 0)$ in direction $d \in C(\bar{x})$. Then for all w satisfying $\nabla g(\bar{x})w + \nabla^2 g(\bar{x})(d, d) \in T^2_\Lambda(g(\bar{x}); \nabla g(\bar{x})d)$ one has*

$$\nabla f(\bar{x})w + \nabla^2 f(\bar{x})(d, d) \geq 0.$$

4 Dual form of second-order optimality conditions

In this section we will derive the dual form of second-order necessary optimality conditions for the general problem (P). By Corollary 2 we know, that at a local solution $\bar{x}$ of (P) for every critical direction d satisfying a directional metric subregularity constraint qualification, the optimal value of the program

$$\begin{aligned} \min_w \quad & \nabla f(\bar{x})w + \nabla^2 f(\bar{x})(d, d) \\ \text{s.t.} \quad & \nabla g(\bar{x})w + \nabla^2 g(\bar{x})(d, d) \in T^2_\Lambda(g(\bar{x}); \nabla g(\bar{x})d) \end{aligned}$$

is nonnegative. The second-order necessary conditions for (P) presented below are necessary conditions and charcterizations, respectively, of this fact.

Recall that for a set S, its support function is defined as $\sigma_S(\lambda) := \sup_{u \in S} \lambda^T u$. Suppose that the supremum $\sigma_S(\lambda)$ is achieved at $\bar{u} \in S$. Since $\bar{u} \in S$ is an optimal solution for $\sup_{u \in S} \lambda^T u$ if and only if $\lambda \in N_S(\bar{u})$ as S is convex, the support function in convex case can be represented as

$$\sigma_S(\lambda) = \lambda^T u \;\text{ if } \lambda \in N_S(u) \text{ for some } u \in S. \tag{15}$$

Inspired by the above expression for the support function when the set is convex and the supremum is achieved, we define the following function which will play an important role for our analysis. It turns out that this function is in general smaller and coincides with the support function when the set is convex.

Definition 7. *Given a nonempty closed set $S \subseteq \mathbb{R}^n$ we define the* lower generalized support function to S *as the mapping $\hat{\sigma}_S : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ by*

$$\hat{\sigma}_S(\lambda) := \liminf_{\tilde{\lambda} \to \lambda} \inf_u \{\tilde{\lambda}^T u \mid \tilde{\lambda} \in N_S(u)\}.$$

Further, for every subset $A \subseteq \mathbb{R}^n$ we define the lower generalized support function to S with respect to A *as the mapping $\hat{\sigma}_{S,A} : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ by*

$$\hat{\sigma}_{S,A}(\lambda) := \liminf_{\tilde{\lambda} \to \lambda} \inf_u \{\tilde{\lambda}^T u \mid u \in N_S^{-1}(\tilde{\lambda}) \cap A\}.$$

By convention, $\hat{\sigma}_S(\lambda) := +\infty$ if $S = \emptyset$. By the definition, we have $\hat{\sigma}_S \le \hat{\sigma}_{S,A}$ for every subset $A \subseteq \mathbb{R}^n$ and $\hat{\sigma}_{S,B} \le \hat{\sigma}_{S,A}$ whenever $A \subseteq B \subseteq \mathbb{R}^n$.

We can show that the limiting normal cone in the above definition of $\hat{\sigma}_S$ can be replaced by the regular normal cone.

Lemma 5. *Let $S \subseteq \mathbb{R}^n$ be closed. Then*

$$\hat{\sigma}_S(\lambda) = \liminf_{\tilde{\lambda}\to\lambda} \inf_u \{\tilde{\lambda}^T u \mid \tilde{\lambda} \in \widehat{N}_S(u)\}, \quad \forall \lambda \in \mathbb{R}^n.$$

Proof. It follows easily that by the definition of the limiting normal cone that we have $\widehat{N}_S(u) \subseteq N_S(u)$ and for every $u \in S$, every $\tilde{\lambda} \in N_S(u)$ and every $\epsilon > 0$ we can find u_ϵ and $\tilde{\lambda}_\epsilon \in \widehat{N}_S(u_\epsilon)$ such that

$$\|u - u_\epsilon\| \le \epsilon, \ \|\tilde{\lambda} - \tilde{\lambda}_\epsilon\| \le \epsilon, \ |\tilde{\lambda}^T u - \tilde{\lambda}_\epsilon^T u_\epsilon| \le \epsilon.$$

□

In the following proposition we show that the lower generalized support function is always less than or equal to the support function and that both functions coincide when the underlying set is convex.

Proposition 6. *1. For every nonempty closed set $S \subseteq \mathbb{R}^n$ one has*

$$\hat{\sigma}_S(\lambda) \le \sigma_S(\lambda), \ \forall \lambda.$$

2. For every nonempty closed convex set $S \subseteq \mathbb{R}^n$ one has

$$\hat{\sigma}_S(\lambda) = \sigma_S(\lambda), \ \forall \lambda.$$

Proof. 1. If $\sigma_S(\lambda) = \infty$, then there is nothing to prove. Now assume that $\sigma_S(\lambda) < \infty$.

Then for any $\epsilon > 0$, there exists an ϵ-optimal solution, say u_ϵ satisfying

$$-\lambda^T u_\epsilon + \delta_S(u_\epsilon) < -\sigma_S(\lambda) + \epsilon,$$

where $\delta_S(\cdot)$ denotes the indicator function of set S. By Ekeland's variational principle, for any $\mu > 0$ satisfying $\mu(\|u_\epsilon\| + 1) \le \epsilon$, there exists $\tilde{u}_\epsilon \in u_\epsilon + \epsilon/\mu\mathbb{B}$ with $-\lambda^T \tilde{u}_\epsilon + \delta_S(\tilde{u}_\epsilon) \le -\lambda^T u_\epsilon + \delta_S(u_\epsilon)$ and $\arg\min_u\{-\lambda^T u + \delta_S(u) + \mu\|u - \tilde{u}_\epsilon\|\} = \{\tilde{u}_\epsilon\}$. According to the first-order optimality conditions, we have

$$0 \in -\lambda + \partial\big(\delta_S(u) + \mu\|u - \tilde{u}_\epsilon\|\big)|_{u=\tilde{u}_\epsilon} \subseteq -\lambda + \mu\mathcal{B} + N_S(\tilde{u}_\epsilon),$$

where the second inclusion follows from the subdifferential sum rule [31, Corollary 10.9]. Hence there exists λ'_ϵ such that $\|\lambda'_\epsilon - \lambda\| \le \mu \le \epsilon$ and $\lambda'_\epsilon \in N_S(\tilde{u}_\epsilon)$. Note that $\lambda'_\epsilon \to \lambda$ as $\epsilon \to 0$ and

$$\|\lambda'_\epsilon - \lambda\|\|\tilde{u}_\epsilon\| \le \mu(\|u_\epsilon\| + \epsilon/\mu) = \mu\|u_\epsilon\| + \epsilon \le 2\epsilon. \tag{16}$$

Thus

$$\hat{\sigma}_S(\lambda) \;=\; \liminf_{\tilde{\lambda}\to\lambda} \inf_u \{\tilde{\lambda}^T u \mid \tilde{\lambda} \in N_S(u)\} \le \liminf_{\epsilon\to 0^+} \inf_u \{(\lambda'_\epsilon)^T u \mid \lambda'_\epsilon \in N_S(u)\}$$

$$
\begin{aligned}
&\le \liminf_{\epsilon\to 0^+} \langle \lambda'_\epsilon, \tilde{u}_\epsilon\rangle = \liminf_{\epsilon\to 0^+} \langle \lambda, \tilde{u}_\epsilon\rangle + \langle \lambda'_\epsilon - \lambda, \tilde{u}_\epsilon\rangle \\
&\le \liminf_{\epsilon\to 0^+} \sigma_S(\lambda) + \|\lambda'_\epsilon - \lambda\|\|\tilde{u}_\epsilon\| \\
&\le \liminf_{\epsilon\to 0^+} \sigma_S(\lambda) + 2\epsilon \\
&= \sigma_S(\lambda),
\end{aligned}
$$

where we have used the fact (16).

2. By virtue of 1, we only need to prove the inequality $\hat{\sigma}_S(\lambda) \ge \sigma_S(\lambda)$. Consider an arbitrary λ. If $\hat{\sigma}_S(\lambda) = \infty$ there is nothing to show. Hence we can assume $\hat{\sigma}_S(\lambda) < \infty$. Then we can find a sequences λ_k converging to λ such that

$$
\hat{\sigma}_S(\lambda) = \liminf_{\tilde{\lambda}\to\lambda} \inf_u \{\tilde{\lambda}^T u \mid \tilde{\lambda} \in N_S(u)\} = \lim_{k\to\infty} \inf_u \{\lambda_k^T u \mid \lambda_k \in N_S(u)\}.
$$

By convexity of S we have by (15) that $\lambda_k^T u = \sigma_S(\lambda_k)$ whenever $\lambda_k \in N_S(u)$. It follows by the above and the lower semi-continuity of the support function that

$$
\hat{\sigma}_S(\lambda) = \liminf_{\tilde{\lambda}\to\lambda} \inf_u \{\tilde{\lambda}^T u \mid \tilde{\lambda} \in N_S(u)\} = \liminf_{\lambda_k\to\lambda} \sigma_S(\lambda_k) \ge \sigma_S(\lambda).
$$

□

Let the Lagrange function of problem (P) be

$$
L(x, \lambda) := f(x) + g(x)^T \lambda.
$$

Consider the following directional Mordukhovich (M-) multiplier set:

$$
\Lambda(\bar{x}; d) := \{\lambda \,|\, \nabla_x L(\bar{x}, \lambda) = 0,\ \lambda \in N_\Lambda(g(\bar{x}); \nabla g(\bar{x})d)\}.
$$

The following directional first-order necessary optimality condition holds at a local minimizer under the directional metric subregularity.

Proposition 7. *Let $\bar{x}$ be a local optimal solution of problem (P). Suppose that the set-valued map $M(x) := g(x) - \Lambda$ is metrically subregular at $(\bar{x}, 0)$ in direction d with $d \in C(\bar{x})$. Then the directional M-multiplier set $\Lambda(\bar{x}; d)$ is nonempty.*

Proof. By [13, Theorem 7] there is some λ satisfying

$$
0 \in \nabla f(\bar{x}) + D^* M((\bar{x}, 0); (d, 0))(\lambda)
$$

where the directional limiting coderivative $D^* M((\bar{x}, 0); (d, 0))$ is defined by

$$
u^* \in D^* M((\bar{x}, 0); (d, 0))(\lambda) \Longleftrightarrow (u^*, -\lambda) \in N_{\operatorname{gph} M}((\bar{x}, 0); (d, 0)).
$$

Since $\operatorname{gph} M = \{(x, y) |\, g(x) - y \in \Lambda\}$, we obtain from [2, Corollary 3.2]

$$
N_{\operatorname{gph} M}((\bar{x}, 0); (d, 0)) \subseteq \{(\nabla g(\bar{x})^T \mu, -\mu) \ \mid\ \mu \in N_\Lambda\big(g(\bar{x}); \nabla g(\bar{x})d\big)\}
$$

yielding the assertion of the proposition. □

In the following theorem we give a second-order necessary optimality condition for problem (P) in terms of directional M-multipliers under the directional metric subregularity condition.

Theorem 2. *Let $\bar{x}$ be a local optimal solution of problem (P). Suppose that the set-valued map $M(x) := g(x) - \Lambda$ is metrically subregular at $(\bar{x}, 0)$ in direction d with $d \in C(\bar{x})$ and $T^2_\Lambda\big(g(\bar{x}); \nabla g(\bar{x})d\big) \neq \emptyset$. Then there exists a directional M-multiplier $\lambda \in \Lambda(\bar{x}; d)$ such that for every set A satisfying $A \supseteq \nabla g(\bar{x}) T^2_{\mathcal{F}}(\bar{x}; d) + \nabla^2 g(\bar{x})(d, d)$ one has*

$$\nabla^2_{xx} L(\bar{x}, \lambda)(d, d) - \hat{\sigma}_{T^2_\Lambda(g(\bar{x}); \nabla g(\bar{x})d), A}(\lambda) \geq 0. \tag{17}$$

In particular, there exists a multiplier $\lambda \in \Lambda(\bar{x}; d)$ such that

$$\nabla^2_{xx} L(\bar{x}, \lambda)(d, d) - \hat{\sigma}_{T^2_\Lambda(g(\bar{x}); \nabla g(\bar{x})d)}(\lambda) \geq 0. \tag{18}$$

Proof. Note that by Corollary 1, under the assumptions of the theorem we have $T^2_{\mathcal{F}}(\bar{x}; d) \neq \emptyset$. Consider for every $\epsilon > 0$ the optimization problem

$$\min_w \nabla f(\bar{x}) w + \frac{\epsilon}{2} \|w\|^2 \text{ subject to } w \in T^2_{\mathcal{F}}(\bar{x}; d). \tag{19}$$

Since $\epsilon > 0$, problem (19) has a globally optimal solution w_ϵ because the objective is coercive and the set $T^2_{\mathcal{F}}(\bar{x}; d)$ is closed and nonempty. We claim that $\lim_{\epsilon \downarrow 0} \epsilon \|w_\epsilon\| = 0$. Indeed, by Theorem 1 together with the optimality of w_ϵ we obtain

$$-\nabla^2 f(\bar{x})(d, d) + \frac{\epsilon}{2}\|w_\epsilon\|^2 \leq \nabla f(\bar{x}) w_\epsilon + \frac{\epsilon}{2}\|w_\epsilon\|^2 \leq \nabla f(\bar{x}) \bar{w} + \frac{\epsilon}{2}\|\bar{w}\|^2$$

for arbitrarily chosen $\bar{w} \in T^2_{\mathcal{F}}(\bar{x}; d)$, yielding

$$\epsilon \|w_\epsilon\| \leq \sqrt{2\epsilon}\Big(\nabla f(\bar{x}) \bar{w} + \nabla^2 f(\bar{x})(d, d) + \frac{\epsilon}{2}\|\bar{w}\|^2\Big)^{\frac{1}{2}}.$$

Taking the limit on the both sides of the above inequality yields $\epsilon \|w_\epsilon\| \to 0$ as $\epsilon \downarrow 0$. By [31, Theorem 6.12], the basic first-order optimality condition for problem (19) at w_ϵ

$$-\nabla f(\bar{x}) - \epsilon w_\epsilon \in N_{T^2_{\mathcal{F}}(\bar{x}; d)}(w_\epsilon) \tag{20}$$

is fulfilled. By Proposition 5, $T^2_{\mathcal{F}}(\bar{x}; d) = \{w \,|\, P(w) \in D\}$, where $P(w) := \nabla g(\bar{x}) w + \nabla^2 g(\bar{x})(d, d)$, $D := T^2_\Lambda\big(g(\bar{x}); \nabla g(\bar{x})d\big)$, and MSCQ holds at $w_\epsilon \in T^2_{\mathcal{F}}(\bar{x}; d)$ for the system $P(w) \in D$ with modulus κ which is the modulus of metric subregularity of M at $(\bar{x}, 0)$ in direction d. It follows by [19, Theorem 3] that

$$N_{T^2_{\mathcal{F}}(\bar{x}; d)}(w_\epsilon) \subseteq \{z | \exists \lambda_\epsilon \in \kappa \|z\| \mathcal{B} \cap N_D(P(w_\epsilon)) \text{ with } z = \nabla P(w_\epsilon)^T \lambda_\epsilon\}.$$

By virtue of (20) and the above inclusion, there is some multiplier

$$\lambda_\epsilon \in \kappa \|\nabla f(\bar{x}) + \epsilon w_\epsilon\| \mathcal{B} \cap N_{T^2_\Lambda(g(\bar{x}); \nabla g(\bar{x})d)}\big(\nabla g(\bar{x}) w_\epsilon + \nabla^2 g(\bar{x})(d, d)\big) \tag{21}$$

such that

$$\nabla f(\bar{x}) + \epsilon w_\epsilon + \nabla g(\bar{x})^T \lambda_\epsilon = 0. \tag{22}$$

Since $\epsilon w_\epsilon \to 0$ as shown above, λ_ϵ is bounded as ϵ sufficiently small. Hence we can take a sequence of positive numbers ϵ_k converging to 0 such that the corresponding sequence of multipliers λ_{ϵ_k} converges to some λ. Taking limits as $\epsilon_k \to 0$ in (22) we obtain

$$\nabla f(\bar{x}) + \nabla g(\bar{x})^T \lambda = 0.$$

By (21) and Lemma 3, we have

$$\lambda_\epsilon \in N_{T^2_\Lambda(g(\bar{x});\nabla g(\bar{x})d)}\big(\nabla g(\bar{x})w_\epsilon + \nabla^2 g(\bar{x})(d,d)\big) \subseteq N_\Lambda(g(\bar{x});\nabla g(\bar{x})d).$$

Taking limits as $\epsilon_k \to 0$, we obtain $\lambda \in N_\Lambda(g(\bar{x});\nabla g(\bar{x})d)$ and consequently $\lambda \in \Lambda(\bar{x};d)$.

Now consider any set $A \supseteq \nabla g(\bar{x})T^2_{\mathcal{F}}(\bar{x};d) + \nabla^2 g(\bar{x})(d,d)$. Setting $u_k := \nabla g(\bar{x})w_{\epsilon_k} + \nabla^2 g(\bar{x})(d,d)$ we have $\lambda_{\epsilon_k} \in N_{T^2_\Lambda(g(\bar{x});\nabla g(\bar{x})d)}(u_k)$ and $u_k \in A$. Taking into account (22) and $\nabla f(\bar{x})w_{\epsilon_k} + \nabla^2 f(\bar{x})(d,d) \geq 0$ by virtue of Theorem 1 we obtain

$$\begin{aligned}
\nabla^2_{xx}L(\bar{x},\lambda)(d,d) &= \lim_{k\to\infty}\left\{\nabla^2_{xx}L(\bar{x},\lambda_{\epsilon_k})(d,d) + \big(\nabla f(\bar{x}) + \epsilon_k w_{\epsilon_k} + \nabla g(\bar{x})^T\lambda_{\epsilon_k}\big)^T w_{\epsilon_k}\right\}\\
&= \lim_{k\to\infty}\Big(\nabla f(\bar{x})w_{\epsilon_k} + \nabla^2 f(\bar{x})(d,d) + \lambda_{\epsilon_k}^T u_k + \epsilon_k\|w_{\epsilon_k}\|^2\Big)\\
&\geq \limsup_{k\to\infty}\lambda_{\epsilon_k}^T u_k \geq \limsup_{k\to\infty}\inf\{\lambda_{\epsilon_k}^T u \mid u \in N^{-1}_{T^2_\Lambda(g(\bar{x});\nabla g(\bar{x})d)}(\lambda_{\epsilon_k}) \cap A\}\\
&\geq \liminf_{\tilde{\lambda}\to\lambda}\inf\{\tilde{\lambda}^T u \mid u \in N^{-1}_{T^2_\Lambda(g(\bar{x});\nabla g(\bar{x})d)}(\tilde{\lambda}) \cap A\}\\
&= \hat{\sigma}_{T^2_\Lambda(g(\bar{x});\nabla g(\bar{x})d),A}(\lambda).
\end{aligned}$$

In particular if we take $A = \mathbb{R}^m$, then $\hat{\sigma}_{T^2_\Lambda(g(\bar{x});\nabla g(\bar{x})d),A} = \hat{\sigma}_{T^2_\Lambda(g(\bar{x});\nabla g(\bar{x})d)}$ and hence (18) holds. □

Among all possible choices for A, the second-order necessary condition (17) is strongest for $A = A_{\text{opt}} := \nabla g(\bar{x})T^2_{\mathcal{F}}(\bar{x};d) + \nabla^2 g(\bar{x})(d,d)$ and weakest for $A = \mathbb{R}^m$. There is also an intermediate choice of $A_{\text{mid}} := \nabla g(\bar{x})\mathbb{R}^n + \nabla^2 g(\bar{x})(d,d)$. The optimal choice A_{opt} involves the second-order tangent cone $T^2_{\mathcal{F}}(\bar{x};d)$ which is in general hard to compute and, moreover, if the second-order cone $T^2_{\mathcal{F}}(\bar{x};d)$ is known we may use the primal optimality condition Theorem 1 instead. On the other hand, choosing $A = \mathbb{R}^m$ results in the weakest optimality condition but more trackable lower generalized support function $\hat{\sigma}_{T^2_\Lambda(g(\bar{x});\nabla g(\bar{x})d)}$. The intermediate choice $A_{\text{mid}} := \nabla g(\bar{x})\mathbb{R}^n + \nabla^2 g(\bar{x})(d,d)$ may result in stronger optimality conditions and slightly harder to calculate $\hat{\sigma}_{T^2_\Lambda(g(\bar{x});\nabla g(\bar{x})d),A_{\text{mid}}}$ than $\hat{\sigma}_{T^2_\Lambda(g(\bar{x});\nabla g(\bar{x})d)}$.

Note that in Theorem 2, even the first order optimality condition is stronger than the classical M-stationary condition since the directional limiting normal cone is in general smaller than the nondirectional limiting normal cone. However in the case where Λ is convex, the directional M-stationary condition in a critical direction coincides with the classical stationary condition and the directional M-multiplier set $\Lambda(\bar{x};d)$ coincides with the classical multiplier set

$$\Lambda(\bar{x}) := \{\lambda \in N_\Lambda(g(\bar{x})) \,|\, \nabla_x L(\bar{x},\lambda) = 0\}$$

for every critical direction $d \in C(\bar{x})$. Indeed, the inclusion $\Lambda(\bar{x};d) \subseteq \Lambda(\bar{x})$ obviously holds. Now pick any $\lambda \in \Lambda(\bar{x})$. Since $\nabla g(\bar{x})d \in T_\Lambda(g(\bar{x}))$ and $\lambda \in N_\Lambda(g(\bar{x}))$, we have

$$0 \geq \lambda^T\nabla g(\bar{x})d = -\nabla f(\bar{x})d \geq 0$$

implying $\lambda \in \{\nabla g(\bar{x})d\}^\perp$. Owing to (5) we conclude $\lambda \in N_\Lambda(g(\bar{x});\nabla g(\bar{x})d)$ and $\lambda \in \Lambda(\bar{x};d)$ follows. We now specialize Theorem 2 under the additional assumption that Λ is convex and $T^2_\Lambda(g(\bar{x});\nabla g(\bar{x})d)$ is convex. In this case, we obtain the same second order necessary optimality as the classical result of [5, Theorem 3.45] under the directional metric subregularity which is weaker than the Robinson's constraint qualification.

Corollary 3. *Let $\bar{x}$ be a local optimal solution of problem (P) where Λ is convex. Suppose that the set-valued map $M(x) := g(x) - \Lambda$ is metrically subregular at $(\bar{x}, 0)$ in direction d with $d \in C(\bar{x})$. If the second-order tangent cone $T^2_\Lambda(g(\bar{x}); \nabla g(\bar{x})d)$ is convex, then there exists a multiplier $\lambda \in \Lambda(\bar{x})$ such that*

$$\nabla^2_{xx} L(\bar{x}, \lambda)(d, d) - \sigma_{T^2_\Lambda(g(\bar{x});\nabla g(\bar{x})d)}(\lambda) \geq 0.$$

Proof. If $T^2_\Lambda\big(g(\bar{x}); \nabla g(\bar{x})d\big) = \emptyset$, then $\sigma_{T^2_\Lambda(g(\bar{x});\nabla g(\bar{x})d)}(\lambda) = -\infty$ by convention and so there is nothing to prove. The case of $T^2_\Lambda\big(g(\bar{x}); \nabla g(\bar{x})d\big) \neq \emptyset$ follows from (18) and Proposition 6(2). □

Consider the following directional Clarke (C-) multiplier set:

$$\Lambda^c(\bar{x}; d) := \{\lambda \,|\, \nabla_x L(\bar{x}, \lambda) = 0,\ \lambda \in N^c_\Lambda(g(\bar{x}); \nabla g(\bar{x})d)\}.$$

It is clear that the set of directional C-multipliers is closed convex and in general larger than the set of directional M-multipliers.

In what follows, we derive a second-order necessary optimality condition for problem (P) in terms of directional C-multipliers under the constraint qualification condition

$$\nabla g(\bar{x})^T \lambda = 0, \quad \lambda \in N^c_\Lambda(g(\bar{x}); \nabla g(\bar{x})d) \implies \lambda = 0 \tag{23}$$

which we will call *directional Robinson's constraint qualification* (DirRCQ) in direction d. Condition (23) is stronger than FOSCMS in direction d, under which by virtue of Proposition 4 the constraint mapping $M(x) = g(x) - \Lambda$ is metrically subregular in direction d.

Lemma 6. *The following three statements are equivalent:*

(i) DirRCQ in direction d holds.

(ii)
$$\nabla g(\bar{x})\mathbb{R}^n + \widehat{T}_\Lambda(g(\bar{x}); \nabla g(\bar{x})d) = \mathbb{R}^m. \tag{24}$$

(iii) The set $\Lambda^c(\bar{x}; d)$ is compact, whenever it is nonempty.

Proof. Condition (23) can be equivalently written as

$$\ker \nabla g(\bar{x})^T \cap N^c_\Lambda(g(\bar{x}); \nabla g(\bar{x})d) = \{0\}. \tag{25}$$

By taking polars on both sides of the equation and using the rule for polar cones [31, Corollary 11.25] and the fact that $\big(N^c_\Lambda(g(\bar{x}); \nabla g(\bar{x})d)\big)^\circ = \widehat{T}_\Lambda(g(\bar{x}); \nabla g(\bar{x})d)$, we can see that condition (23) is equivalent to saying $\mathrm{cl}\Big(\nabla g(\bar{x})\mathbb{R}^n + \widehat{T}_\Lambda(g(\bar{x}); \nabla g(\bar{x})d)\Big) = \mathbb{R}^m$. Obviously the set $\nabla g(\bar{x})\mathbb{R}^n + \widehat{T}_\Lambda(g(\bar{x}); \nabla g(\bar{x})d)$ is convex and thus $\mathrm{ri}\,(\nabla g(\bar{x})\mathbb{R}^n + \widehat{T}_\Lambda(g(\bar{x}); \nabla g(\bar{x})d)) = \mathbb{R}^m$ by [30, Theorem 6.3]. It follows that condition (24) holds and thus the implication "(i)⇒(ii)" is established. In order to show the reverse implication, just note that by taking polars on both sides of (24) we obtain (25). Finally, the equivalence between (i) and (iii) follows from [30, Theorem 8.4] together with the fact that the recession cone to $\Lambda^c(\bar{x}; d)$ is exactly the set on the left hand side of (25). □

Proposition 8. *Let $\bar{x}$ be feasible for the problem (P). Suppose that $d \in C(\bar{x})$ satisfies $T^2_\Lambda(g(\bar{x}); \nabla g(\bar{x})d) \neq \emptyset$ and DirRCQ. Then the following three statements are equivalent:*

(i) The primal second-order necessary condition

$$\nabla f(\bar{x})w + \nabla^2 f(\bar{x})(d, d) \geq 0, \quad \forall w \in T^2_{\mathcal{F}}(\bar{x}; d)$$

of Theorem 1 holds.

(ii) For every $u \in T^2_\Lambda(g(\bar{x}); \nabla g(\bar{x})d)$, there exists $\lambda_u \in \Lambda^c(\bar{x}; d)$ such that

$$\nabla^2_{xx} L(\bar{x}, \lambda_u)(d, d) - \lambda_u^T u \geq 0.$$

(iii) For every nonempty convex subset $C \subseteq T^2_\Lambda(g(\bar{x}); \nabla g(\bar{x})d)$, there exists $\lambda \in \Lambda^c(\bar{x}; d)$ such that

$$\nabla^2_{xx} L(\bar{x}, \lambda)(d, d) - \sigma_C(\lambda) \geq 0.$$

Proof. Since condition (23) implies the metric subregularity in direction d, by Proposition 5 we have

$$w \in T^2_{\mathcal{F}}(\bar{x}; d) \Longleftrightarrow \nabla g(\bar{x})w + \nabla^2 g(\bar{x})(d, d) \in T^2_\Lambda(g(\bar{x}); \nabla g(\bar{x})d). \tag{26}$$

We first show the implication "(i)⇒(ii)". Take $u \in T^2_\Lambda(g(\bar{x}); \nabla g(\bar{x})d)$. Then

$$u + \widehat{T}_\Lambda(g(\bar{x}); \nabla g(\bar{x})d) \subseteq T^2_\Lambda(g(\bar{x}); \nabla g(\bar{x})d)$$

by virtue of Proposition 1. Since $\nabla f(\bar{x})w + \nabla^2 f(\bar{x})(d, d) \geq 0$ for all $w \in T^2_{\mathcal{F}}\big(\bar{x}; d\big)$, the following conic linear program

$$\begin{aligned} \min_{w} \quad & \nabla f(\bar{x})w + \nabla^2 f(\bar{x})(d, d) \\ \text{s.t.} \quad & \nabla g(\bar{x})w + \nabla^2 g(\bar{x})(d, d) \in u + \widehat{T}_\Lambda(g(\bar{x}); \nabla g(\bar{x})d) \end{aligned} \tag{27}$$

has nonnegative optimal value. The dual program of the conic linear program (27) is

$$\begin{aligned} \max_{\lambda \in (\widehat{T}_\Lambda(g(\bar{x}); \nabla g(\bar{x})d))^\circ} \quad & \nabla^2_{xx} L(\bar{x}, \lambda)(d, d) - \lambda^T u \\ s.t. \quad & \nabla f(\bar{x}) + \nabla g(\bar{x})^T \lambda = 0. \end{aligned}$$

Since by Proposition 3, $(\widehat{T}_\Lambda(g(\bar{x}); \nabla g(\bar{x})d))^\circ = N^c_\Lambda(g(\bar{x}); \nabla g(\bar{x})d)$, the above dual problem can be equivalently written as

$$\max_{\lambda \in \Lambda^c(\bar{x}; d)} \{\nabla^2_{xx} L(\bar{x}, \lambda)(d, d) - \lambda^T u\}.$$

By Lemma 6, condition (23) is equivalent to (24) and it is easy to see that the latter implies

$$0 \in \operatorname{int}\{\nabla g(\bar{x})\mathbb{R}^n + \nabla^2 g(\bar{x})(d, d) - \widehat{T}_\Lambda(g(\bar{x}); \nabla g(\bar{x})d)\},$$

cf. [5, (2.354)]. Consequently by [5, Theorem 2.187], there is no dual gap and the dual program has an optimal solution λ_u such that

$$\max_{\lambda \in \Lambda^c(\bar{x}; d)} \{\nabla^2_{xx} L(\bar{x}, \lambda)(d, d) - \lambda^T u\} = \nabla^2_{xx} L(\bar{x}, \lambda_u)(d, d) - \lambda_u^T u \geq 0.$$

This proves "(i)⇒(ii)".

In order to show the reverse implication, take $w \in T^2_{\mathcal{F}}(\bar{x}; d)$ together with $u := \nabla g(\bar{x})w + \nabla^2 g(\bar{x})(d, d) \in T^2_\Lambda(g(\bar{x}); \nabla g(\bar{x})w)$, where the existence of such w is guaranteed by virtue of (26), and $\lambda_u \in \Lambda^c(\bar{x}; d)$ fulfilling $\nabla^2_{xx} L(\bar{x}, \lambda_u)(d, d) - \lambda_u^T u \geq 0$. Then

$$\nabla f(\bar{x})w + \nabla^2 f(\bar{x})(d, d) = -\lambda_u^T \nabla g(\bar{x})w + \nabla^2 f(\bar{x})(d, d) = \nabla^2_{xx} L(\bar{x}, \lambda_u)(d, d) - \lambda_u^T u \geq 0$$

showing "(ii)⇒(i)".
Since "(iii)⇒(ii)" always hold, there remains to show "(ii)⇒(iii)". Consider a nonempty convex subset $C \subseteq T^2_\Lambda(g(\bar{x}); \nabla g(\bar{x})d)$. Since $\sigma_C = \sigma_{\operatorname{cl} C}$, we may assume that C is closed. For every $u \in C$ the corresponding λ_u according to (ii) fulfills $\lambda_u \in \Lambda^c(\bar{x}; d)$ and hence $\Lambda^c(\bar{x}; d)$ is nonempty. We conclude from Lemma 6 that $\Lambda^c(\bar{x}; d)$ is compact and therefore by the minimax theorem in [30, Corollary 37.3.2] we have

$$\begin{aligned} \inf_{u \in C} \sup_{\lambda \in \Lambda^c(\bar{x}; d)} \nabla^2_{xx} L(\bar{x}, \lambda)(d, d) - \lambda^T u &= \sup_{\lambda \in \Lambda^c(\bar{x}; d)} \inf_{u \in C} \nabla^2_{xx} L(\bar{x}, \lambda)(d, d) - \lambda^T u & (28) \\ &= \sup_{\lambda \in \Lambda^c(\bar{x}; d)} \nabla^2_{xx} L(\bar{x}, \lambda)(d, d) - \sigma_C(\lambda). & (29) \end{aligned}$$

Due to (ii) the quantity on left hand side of (28) is nonnegative. On the other hand, the supremum in (29) is attained at some λ, since $\Lambda^c(\bar{x}; d)$ is compact and $-\sigma_C(\cdot)$ is upper semi-continuous. This completes the proof. □

The following second-order necessary optimality condition follows immediately by Theorem 1 and Proposition 8.

Corollary 4. *Let $\bar{x}$ be a local optimal solution of problem (P). Suppose that $d \in C(\bar{x})$ satisfies $T^2_\Lambda(g(\bar{x}); \nabla g(\bar{x})d) \neq \emptyset$ and DirRCQ. Then, for every nonempty convex subset $C \subseteq T^2_\Lambda(g(\bar{x}); \nabla g(\bar{x})d)$ there is some $\lambda \in \Lambda^c(\bar{x}; d)$ such that*

$$\nabla^2_{xx} L(\bar{x}, \lambda)(d, d) - \sigma_C(\lambda) \geq 0.$$

A close look at the proof of Theorem 2 shows that the second-order necessary conditions stated therein are implied by the primal second-order necessary condition of Theorem 1. Hence, in view of Proposition 8, the second-order necessary condition of Corollary 4 are stronger than the one of Theorem 2. However, the constraint qualification DirRCQ used in Corollary 4 is also stronger than the one of Theorem 2.

We now want to compare Corollary 4 with the classical result of [5, Theorem 3.45] under the additional assumption that Λ is convex. By (5) and using the convexity of $N_\Lambda(g(\bar{x}); \nabla g(\bar{x})d)$, for every d with $\nabla g(\bar{x})d \in T_\Lambda(g(\bar{x}))$ there holds

$$\begin{aligned} \nabla g(\bar{x})^T \lambda = 0,\ \lambda \in N^c_\Lambda(g(\bar{x}); \nabla g(\bar{x})d) \quad &\Leftrightarrow \quad \nabla g(\bar{x})^T \lambda = 0,\ \lambda \in N_\Lambda(g(\bar{x})),\ \lambda \in \{\nabla g(\bar{x})d\}^\perp \\ &\Leftrightarrow \quad \nabla g(\bar{x})^T \lambda = 0,\ \lambda \in N_\Lambda(g(\bar{x})) \\ &\Leftrightarrow \quad \lambda \in (\nabla g(\bar{x})\mathbb{R}^n)^\perp \cap N_\Lambda(g(\bar{x})). \end{aligned}$$

Thus, by [5, Proposition 2.97] the directional Robinson's constraint qualification is equivalent to the non-directional one

$$\nabla g(\bar{x})\mathbb{R}^n + T_\Lambda(g(\bar{x})) = \mathbb{R}^m.$$

Since we also have $\Lambda^c(\bar{x}; d) = \Lambda(\bar{x})$ as pointed out above, in case of convex Λ the Corollary 4 is equivalent with [5, Theorem 3.45].

When the directional C-multiplier set $\Lambda^c(\bar{x}; d) = \{\lambda_0\}$ is a singleton, it is easy to see from Proposition 8(ii) and Theorem 1 that

$$\nabla^2_{xx} L(\bar{x}, \lambda_0)(d, d) - \sigma_{T^2_\Lambda(g(\bar{x}); \nabla g(\bar{x})d)}(\lambda_0) \geq 0 \tag{30}$$

is a necessary second-order condition at a local minimizer $\bar{x}$. Note that by definition $\sigma_{T^2_\Lambda(g(\bar{x}); \nabla g(\bar{x})d)}(\lambda_0) = -\infty$ whenever $T^2_\Lambda(g(\bar{x}); \nabla g(\bar{x})d) = \emptyset$. In this case the above second-order optimality condition holds automatically. We now try to enhance the condition (23) so that the directional C-multiplier set $\Lambda^c(\bar{x}; d)$ is a singleton. As we will show below this is achieved by the directional non-degeneracy condition

$$\nabla g(\bar{x})^T \lambda = 0, \ \lambda \in \text{span } N_\Lambda(g(\bar{x}); \nabla g(\bar{x})d) \implies \lambda = 0 \tag{31}$$

defined for $\bar{x}$ feasible for the problem (P) and direction d satisfying $\nabla g(\bar{x})d \in T_\Lambda(g(\bar{x}))$, where span S denotes the affine hull of the set S.

Recall that by [19, Definition 6] we call a subspace L the generalized linearity space of set C and denote it by $\mathcal{L}(C)$ provided that it is the largest subspace L such that $C + L \subseteq C$. Note that when C is a convex set, the generalized linearity space is reduced to the linearity space ([30, page 65]) and in the case when C is a closed convex cone, we have $\mathcal{L}(C) = (-C) \cap C$.

Given $d \in C(\bar{x})$, define the set of *strong multipliers in direction d* as

$$\Lambda^s(\bar{x}; d) := \{\lambda \in \widehat{N}_{T_\Lambda(g(\bar{x}))}(\nabla g(\bar{x})d) \,|\, \nabla_x L(\bar{x}, \lambda) = 0\}.$$

By (13) we have $\Lambda^s(\bar{x}; d) \subseteq \Lambda(\bar{x}; d)$.

Lemma 7. *Let $\bar{x}$ be a local optimal solution of problem (P) and suppose that the directional non-degeneracy condition* (31) *holds for a critical direction $d \in C(\bar{x})$. Then*

$$\Lambda^s(\bar{x}; d) = \Lambda(\bar{x}; d) = \Lambda^c(\bar{x}; d) = \{\lambda_0\}$$

is a singleton.

Proof. Since $d \in C(\bar{x})$ implies that $\nabla g(\bar{x})d \in T_\Lambda(g(\bar{x}))$ and (31) implies that

$$\nabla g(\bar{x})^T \lambda = 0, \ \lambda \in N_\Lambda(g(\bar{x}); \nabla g(\bar{x})d) \Longrightarrow \lambda = 0,$$

FOSCMS for direction d holds and by Proposition 4, the set-valued map $g(x) - \Lambda$ is metrically subregular at $(\bar{x}, 0)$ in direction d. By definition there are reals $\kappa, \rho, \delta > 0$ such that

$$\text{dist}(x, \mathcal{F}) \leq \kappa \text{dist}(g(x), \Lambda) \ \forall x \in \bar{x} + V_{\rho,\delta}(d). \tag{32}$$

We now show by contradiction that d is a locally optimal solution of the program

$$\min_{d'} \nabla f(\bar{x})d' \quad \text{subject to} \quad \nabla g(\bar{x})d' \in T_\Lambda(g(\bar{x})). \tag{33}$$

Assume on the contrary that there is a sequence $d_k \to d$ satisfying $\nabla g(\bar{x})d_k \in T_\Lambda(g(\bar{x}))$ and $\nabla f(\bar{x})d_k < 0 = \nabla f(\bar{x})d$. Then we can find some index $\bar{k}$ and some $\bar{t} > 0$ such that $\|\bar{t}d_{\bar{k}}\| < \rho$ and

$$\big\|\|d\|d_{\bar{k}} - \|d_{\bar{k}}\|d\big\| \leq \big\|\|d\|d_{\bar{k}} - \|d\|d\big\| + \big|\|d_{\bar{k}}\| - \|d\|\big|\|d\| \leq 2\|d_{\bar{k}} - d\|\|d\| \leq \delta\|d_{\bar{k}}\|\|d\|$$

implying $td_{\bar{k}} \in \mathcal{V}_{\rho,\delta}(d)\ \forall t \in [0, \bar{t}]$. By (32), there is some sequence $t_n \downarrow 0$ such for all $t_n < \bar{t}$ we can find some $x_n \in \mathcal{F}$ satisfying

$$\|x_n - (\bar{x} + t_n d_{\bar{k}})\| \le \kappa \mathrm{dist}(g(\bar{x} + t_n d_{\bar{k}}), \Lambda) = \kappa\big(\mathrm{dist}(g(\bar{x}) + t_n \nabla g(\bar{x}) d_{\bar{k}}, \Lambda) + o(t_n)\big) = o(t_n),$$

where the last equality follows from the fact that $\nabla g(\bar{x}) d_{\bar{k}} \in T_\Lambda(g(\bar{x}))$. It follows that $f(x_n) = f(\bar{x}) + t_n \nabla f(\bar{x}) d_{\bar{k}} + o(t_n) < f(\bar{x})$ for all t_n sufficiently small contradicting the optimality of $\bar{x}$ for the problem (P). Hence d is a local minimizer for the problem (33) and the basic optimality condition [31, Theorem 6.12]

$$-\nabla f(\bar{x}) \in \widehat{N}_{\nabla g(\bar{x})^{-1}\left(T_\Lambda(g(\bar{x}))\right)}(d) \tag{34}$$

is fulfilled.

By taking polars in both sides of the directional non-degeneracy condition (31), we have

$$\nabla g(\bar{x})\mathbb{R}^n + \big(\mathrm{span} N_\Lambda(g(\bar{x}); \nabla g(\bar{x}) d)\big)^\circ = \mathbb{R}^m.$$

Since

$$\mathrm{span}\ N_\Lambda(g(\bar{x}); \nabla g(\bar{x}) d) = \mathrm{span}\ \mathrm{cl}\,\mathrm{co}\, N_\Lambda(g(\bar{x}); \nabla g(\bar{x}) d) = \mathrm{span} N^c_\Lambda(g(\bar{x}); \nabla g(\bar{x}) d)$$

and the Clarke directional normal cone is a closed convex cone, we have

$$\begin{aligned}
&\big(\mathrm{span}\ N_\Lambda(g(\bar{x}); \nabla g(\bar{x}) d)\big)^\circ \\
&= \big(\mathrm{span}\ N^c_\Lambda(g(\bar{x}); \nabla g(\bar{x}) d)\big)^\circ = \big(N^c_\Lambda(g(\bar{x}); \nabla g(\bar{x}) d) - N^c_\Lambda(g(\bar{x}); \nabla g(\bar{x}) d)\big)^\circ \\
&= N^c_\Lambda(g(\bar{x}); \nabla g(\bar{x}) d)^\circ \cap \big(-N^c_\Lambda(g(\bar{x}); \nabla g(\bar{x}) d)\big)^\circ \\
&= \widehat{T}_\Lambda(g(\bar{x}); \nabla g(\bar{x}) d) \cap \big(-\widehat{T}_\Lambda(g(\bar{x}); \nabla g(\bar{x}) d)\big) = \mathcal{L}(\widehat{T}_\Lambda(g(\bar{x}); \nabla g(\bar{x}) d)),
\end{aligned}$$

where the last equality follows from the fact that the directional regular tangent cone is a closed convex cone. Hence we have shown that the directional non-degeneracy condition (31) is equivalent to

$$\nabla g(\bar{x})\mathbb{R}^n + \mathcal{L}(\widehat{T}_\Lambda(g(\bar{x}); \nabla g(\bar{x}) d)) = \mathbb{R}^m.$$

Note that $\mathcal{L}(\widehat{T}_\Lambda(g(\bar{x}); \nabla g(\bar{x}) d)) \subseteq \mathcal{L}(T_{T_\Lambda(g(\bar{x}))}(\nabla g(\bar{x}) d))$ because

$$T_{T_\Lambda(g(\bar{x}))}(\nabla g(\bar{x}) d) + \widehat{T}_\Lambda(g(\bar{x}); \nabla g(\bar{x}) d) = T_{T_\Lambda(g(\bar{x}))}(\nabla g(\bar{x}) d)$$

by Proposition 1. It follows that

$$\nabla g(\bar{x})\mathbb{R}^n + \mathcal{L}(T_{T_\Lambda(g(\bar{x}))}(\nabla g(\bar{x}) d)) = \mathbb{R}^m$$

and, since the mapping $w \rightrightarrows \nabla g(\bar{x}) w - T_\Lambda(g(\bar{x}))$ is metrically subregular at $(d, 0)$ by Lemma 1, we may invoke [17, Theorem 4] to obtain $\widehat{N}_{\nabla g(\bar{x})^{-1}\left(T_\Lambda(g(\bar{x}))\right)}(d) = \nabla g(\bar{x})^T \widehat{N}_{T_\Lambda(g(\bar{x}))}(\nabla g(\bar{x}) d)$. By (34), it follows that $\emptyset \ne \Lambda^s(\bar{x}; d) \subseteq \Lambda(\bar{x}; d) \subseteq \Lambda^c(\bar{x}; d)$. Since

$$N^c_\Lambda(g(\bar{x}); \nabla g(\bar{x}) d) - N^c_\Lambda(g(\bar{x}); \nabla g(\bar{x}) d) = \mathrm{span}\ N^c_\Lambda(g(\bar{x}); \nabla g(\bar{x}) d) = \mathrm{span}\ N_\Lambda(g(\bar{x}); \nabla g(\bar{x}) d),$$

condition (31) ensures that $\Lambda^c(\bar{x}; d)$ is a singleton $\{\lambda_0\}$. Hence the claimed result $\Lambda^s(\bar{x}; d) = \Lambda(\bar{x}; d) = \Lambda^c(\bar{x}; d) = \{\lambda_0\}$ follows. □

Corollary 5. *Let $\bar{x}$ be a local optimal solution of problem (P). Suppose that for $d \in C(\bar{x})$ the directional non-degeneracy condition* (31) *holds. Then $\Lambda^s(\bar{x}; d) = \Lambda(\bar{x}; d) = \Lambda^c(\bar{x}; d) = \{\lambda_0\}$ and the second-order condition* (30) *is fulfilled.*

The following result extends the second-order necessary optimality condition for convex set-constrained problems as in [5, Proposition 3.46] to allow the set Λ to be non-convex. Note that in [5, Proposition 3.46], both Robinson's constraint qualification and the uniqueness of the multipliers are required while we derive our result under a nondegeneracy condition which is stronger than Robinson's CQ but guarantees the uniqueness of the multipliers.

Corollary 6. *Let $\bar{x}$ be a local optimal solution of problem (P) and assume that the non-degeneracy condition*

$$\nabla g(\bar{x})^T \lambda = 0, \ \lambda \in \text{span } N_\Lambda(g(\bar{x})) \implies \lambda = 0 \tag{35}$$

holds. Then there is a unique multiplier λ_0 satisfying the first-order optimality conditions

$$\nabla_x L(\bar{x}, \lambda_0) = 0, \ \lambda_0 \in \widehat{N}_\Lambda(g(\bar{x})). \tag{36}$$

Further, for all $d \in C(\bar{x})$ we have

$$\nabla^2_{xx} L(\bar{x}, \lambda_0) - \sigma_{T^2_\Lambda(g(\bar{x}); \nabla g(\bar{x})d)}(\lambda_0) \geq 0.$$

Proof. The existence and uniqueness of the multiplier λ_0 fulfilling (36) follows from Lemma 7 applied with $d = 0$. Further,

$$\{\lambda_0\} = \{\lambda \in \widehat{N}_\Lambda(g(\bar{x})) \mid \nabla_x L(\bar{x}, \lambda) = 0\} \subseteq \{\lambda \in N_\Lambda(g(\bar{x})) \mid \nabla_x L(\bar{x}, \lambda) = 0\} = \Lambda(\bar{x})$$

and the non-degeneracy condition (35) ensures that $\Lambda(\bar{x}) = \{\lambda_0\}$. Further note that for every $d \in C(\bar{x})$ the condition (35) implies (31) and $\emptyset \neq \Lambda(\bar{x}; d) \subseteq \Lambda(\bar{x}) = \{\lambda_0\}$. This shows $\Lambda(\bar{x}; d) = \{\lambda_0\}$ and the second statement follows from Corollary 5. □

Remark 1. *According to [10], the first-order optimality conditions* (36) *are called* S-stationarity conditions*.*

5 Second-order sufficient conditions

We now consider sufficient conditions for optimality. We need the following definition of an upper second order approximation set of Λ which is a special case of the definition given in [5, Definition 3.82].

Definition 8. *Let $\bar{x}$ be a feasible solution of problem (P) and $d \in C(\bar{x})$. We say that a closed set $\mathcal{A}(d)$ is an* upper second-order approximation set *for Λ at $g(\bar{x})$ in direction $\nabla g(\bar{x})d \in T_\Lambda(g(\bar{x}))$, if for any sequence $y_n \in \Lambda$ of the form $y_n := g(\bar{x}) + t_n \nabla g(\bar{x})d + \frac{1}{2}t_n^2(\nabla g(\bar{x})w_n + a_n)$, where $t_n \downarrow 0$ and $\{a_n\}$ being a convergent sequence and $\{w_n\}$ satisfying $t_n w_n \to 0$, the following condition holds*

$$\lim_{n\to\infty} \text{dist}(\nabla g(\bar{x})w_n + a_n, \mathcal{A}(d)) = 0.$$

Consider the so-called *generalized Lagrangian* $L^g : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}$ defined by

$$L^g(x, \alpha, \lambda) = \alpha f(x) + g(x)^T \lambda.$$

It is easy to check that the following variant of [5, Theorem 3.83] holds for a non-convex set Λ.

Theorem 3. *Let $\bar{x}$ be a feasible point of (P). Assume that every $d \in C(\bar{x})$ corresponds to $\mathcal{A}(d)$, an upper second-order approximation set for Λ at $g(\bar{x})$ in direction d. Further assume that for every $d \in C(\bar{x})\backslash\{0\}$ there is some $(\alpha, \lambda) \in \mathbb{R} \times \mathbb{R}^m$ satisfying*

$$\alpha \geq 0, \ \alpha \nabla f(\bar{x})d = 0, \ \nabla_x L^g(\bar{x}, \alpha, \lambda) = 0 \tag{37}$$

and

$$\nabla^2_{xx} L^g(\bar{x}, \alpha, \lambda)(d, d) - \sigma_{\mathcal{A}(d)}(\lambda) > 0. \tag{38}$$

Then the second order growth condition holds at $\bar{x}$, i.e., there exists a neighborhood U of $\bar{x}$ and $\delta > 0$ such that

$$f(x) \geq f(\bar{x}) + \delta \|x - \bar{x}\|^2 \quad \forall x \in U \ s.t. \ g(x) \in \Lambda.$$

The second-order condition (38) has the following two implications. Firstly, if $\mathcal{A}(d) \neq \emptyset$ it is easy to see that $(\alpha, \lambda) \neq (0, 0)$. Secondly, we have $\lambda \in N^c_\Lambda(g(\bar{x}); \nabla g(\bar{x})d)$ whenever $T^2_\Lambda(g(\bar{x}); \nabla g(\bar{x})d) \neq \emptyset$. Indeed, by the definition we have $\emptyset \neq T^2_\Lambda(g(\bar{x}); \nabla g(\bar{x})d) \subseteq \mathcal{A}(d)$ for every upper second-order approximation set and hence $\lambda \in \big(\widehat{T}_\Lambda(g(\bar{x}); \nabla g(\bar{x})d)\big)^\circ = N^c_\Lambda(g(\bar{x}); \nabla g(\bar{x})d)$ by virtue of Propositions 1 and 3, since otherwise

$$\sigma_{\mathcal{A}(d)}(\lambda) \geq \sigma_{T^2_\Lambda(g(\bar{x}); \nabla g(\bar{x})d)}(\lambda) = \infty.$$

In general $T^2_\Lambda(g(\bar{x}); \nabla g(\bar{x})d)$ may not be an upper second-order approximation set for Λ at $g(\bar{x})$ in direction $\nabla g(\bar{x})d$. But if $T^2_\Lambda(g(\bar{x}); \nabla g(\bar{x})d)$ is an upper second-order approximation set for Λ at $g(\bar{x})$ in direction $\nabla g(\bar{x})d$, then we say that Λ is outer second-order regular at $g(\bar{x})$ in direction $\nabla g(\bar{x})d$; see [5, Definition 3.85].

Combining theorems 2 and 3, we obtain immediately the following "no-gap" necessary and sufficient optimality conditions under the outer second order regularity of Λ. Thus [5, Theorem 3.86] is extended to the non-convex set Λ and Robinson' constraint qualification is weakened to directional metric subregularity.

Theorem 4. *Let $\bar{x}$ be a feasible solution of (P) and assume that Λ is outer second-order regular at $g(\bar{x})$ in direction $\nabla g(\bar{x})d$ for every $d \in C(\bar{x})\backslash\{0\}$. If for every $d \in C(\bar{x})\backslash\{0\}$ there is some $(\alpha, \lambda) \in \mathbb{R} \times \mathbb{R}^m$ satisfying (37) and*

$$\nabla^2_{xx} L^g(\bar{x}, \alpha, \lambda)(d, d) - \sigma_{T^2_\Lambda(g(\bar{x}); \nabla g(\bar{x})d)}(\lambda) > 0, \tag{39}$$

then the second order growth condition holds at $\bar{x}$. If, in addition, the feasible mapping $M(x) = g(x) - \Lambda$ is metrically subregular at $(\bar{x}, 0)$ in direction d, $\nabla f(\bar{x})d = 0$ and $T^2_\Lambda(g(\bar{x}); \nabla g(\bar{x})d)$ is convex for every $d \in C(\bar{x})\backslash\{0\}$, then the second order conditions

$$\sup_{\lambda \in \Lambda(\bar{x}; d)} \Big(\nabla^2_{xx} L(\bar{x}, \lambda)(d, d) - \sigma_{T^2_\Lambda(g(\bar{x}); \nabla g(\bar{x})d)}(\lambda) \Big) > 0, \forall d \in C(\bar{x})\backslash\{0\} \tag{40}$$

are necessary and sufficient for the second order growth condition at the point $\bar{x}$.

Proof. The sufficiency of (40) for the quadratic growth condition follows from (39) by taking $\alpha = 1$. There remains to show the necessity of (40) for the quadratic growth condition. Assume that $f(x) \geq f(\bar{x}) + \delta\|x - \bar{x}\|^2$ holds for all feasible x sufficiently close to $\bar{x}$ for some $\delta > 0$ and consider $d \in C(\bar{x})\backslash\{0\}$. Then $\Lambda(\bar{x}; d) \neq \emptyset$ by Proposition 7 and $\bar{x}$ is a local minimizer of the problem

$$\min_x f(x) - \delta\|x - \bar{x}\|^2 \quad \text{subject to} \quad g(x) \in \Lambda.$$

By Theorem 2, there is some $\lambda \in \Lambda(\bar{x}; d)$ such that

$$\nabla^2_{xx} L(\bar{x}, \lambda)(d, d) - 2\delta\|d\|^2 - \hat{\sigma}_{T^2_\Lambda(g(\bar{x});\nabla g(\bar{x})d)}(\lambda) \geq 0.$$

But by assumption $T^2_\Lambda(g(\bar{x}); \nabla g(\bar{x})d)$ is convex and hence

$$\hat{\sigma}_{T^2_\Lambda(g(\bar{x});\nabla g(\bar{x})d)}(\lambda) = \sigma_{T^2_\Lambda(g(\bar{x});\nabla g(\bar{x})d)}(\lambda)$$

by Proposition 6(2), and (40) follows, provided $T^2_\Lambda(g(\bar{x}); \nabla g(\bar{x})d) \neq \emptyset$. On the other hand, if $T^2_\Lambda(g(\bar{x}); \nabla g(\bar{x})d) = \emptyset$ then (40) automatically holds because the support function of the empty set is identical $-\infty$ by definition. □

6 Examples

In this section we use some examples to illustrate our theory. We will apply our theory to the class of SOC-MPCCs and MPCCs in a forthcoming paper [20].

Example 2. *Consider the one-dimensional problem*

$$\min_{x\in\mathbb{R}} -\frac{1}{2}x^2 \quad \text{s.t.} \quad g(x) := (x^2, x) \in \Lambda := C_1 \cup C_2$$

at the reference point $\bar{x} = 0$, *where* $C_1 := \{(x_1, x_2) \,|\, (x_1 - 1)^2 + x_2^2 \leq 1\}$ *and* $C_2 := \{(x_1, x_2) \,|\, (x_1 + 1)^2 + x_2^2 \leq 1\}$ *are unit circles with center* $(1, 0)$ *and* $(-1, 0)$*, respectively. Clearly, the feasible region is* $\mathcal{F} = [-1, 1]$ *and thus* $\bar{x}$ *is not a local minimizer. We want to check whether we can reject* $\bar{x}$ *as a local minimizer by our theory. First at all note that the feasible set mapping* $g(x) - \Lambda$ *is metrically subregular at* $(\bar{x}, 0)$ *because all* $x \in \mathbb{R}$ *sufficiently close to* $\bar{x}$ *are feasible. Straightforward calculations yield*

$$T_\Lambda(g(\bar{x})) = T_{C_1}(g(\bar{x})) \cup T_{C_2}(g(\bar{x})) = (\mathbb{R}_+ \times \mathbb{R}) \cup (\mathbb{R}_- \times \mathbb{R}) = \mathbb{R}^2,$$

$\widehat{N}_\Lambda(g(\bar{x})) = \{(0, 0)\}$, *and* $C(\bar{x}) = \mathbb{R}$. *Consider the critical direction* $d = 1$*, the discussion for the opposite direction* $d = -1$ *could be performed similarly. Utilizing [5, Propositions 3.30 and 3.37] it is not difficult to show that*

$$\begin{aligned} T^2_\Lambda(g(\bar{x}); \nabla g(\bar{x})d) &= T^2_{C_1}(g(\bar{x}); \nabla g(\bar{x})d) \cup T^2_{C_2}(g(\bar{x}); \nabla g(\bar{x})d) \\ &= (\{t \mid t \geq 1\} \times \mathbb{R}) \cup (\{t \mid t \leq -1\} \times \mathbb{R}). \end{aligned}$$

Since $\nabla g(\bar{x})0 + \nabla^2 g(\bar{x}, \bar{x})(d, d) = (2, 0) \in T^2_\Lambda(g(\bar{x}); \nabla g(\bar{x})d)$*, we obtain* $0 \in T^2_{\mathcal{F}}(\bar{x}; d)$ *from* (11)*. By observing* $\nabla f(\bar{x})0 + \nabla^2 f(\bar{x})(d, d) = -1 < 0$*, we may conclude from Theorem 1 that* $\bar{x}$ *is not a local minimizer.*

Next let us apply Theorem 2. Since for all sufficiently small $x > 0$ *the point* $y^1(x) = (1 - \sqrt{1 - x^2}, x)$ *belongs to* C_1 *but not to* C_2*, we obtain* $\widehat{N}_\Lambda(y^1(x)) = \widehat{N}_{C_1}(y^1(x)) =$

$\mathbb{R}_+(-\sqrt{1-x^2}, x)$. *Together with* $\lim_{x\downarrow 0} \frac{y^1(x)-(0,0)}{x} = (0,1) = \nabla g(\bar{x})d$ *we obtain* $\mathbb{R}_+(-1,0) = \mathbb{R}_- \times \{0\} \subseteq N_\Lambda(g(\bar{x}); \nabla g(\bar{x})d)$. *Similar arguments using the points* $y^2(x) = (\sqrt{1-x^2}-1, x)$ *show* $\mathbb{R}_+ \times \{0\} \subseteq N_\Lambda(g(\bar{x}); \nabla g(\bar{x})d)$ *and in fact, it follows that* $N_\Lambda(g(\bar{x}); \nabla g(\bar{x})d) = \mathbb{R} \times \{0\}$. *It follows that for every* $\lambda \in N_\Lambda(g(\bar{x}); \nabla g(\bar{x})d)$ *we have* $\nabla g(\bar{x})^T \lambda = 0$ *and consequently* $\Lambda(\bar{x}; d) = N_\Lambda(g(\bar{x}); \nabla g(\bar{x})d) = \mathbb{R} \times \{0\}$. *By taking*

$$A = \nabla g(\bar{x})\mathbb{R} + \nabla^2 g(\bar{x})(d,d) = \{2\} \times \mathbb{R} \subset \operatorname{int} T^2_\Lambda(g(\bar{x}); \nabla g(\bar{x})d)$$

we have

$$\hat{\sigma}_{T^2_\Lambda(g(\bar{x});\nabla g(\bar{x})d),A}(\lambda) = \begin{cases} 0 & \text{if } \lambda = (0,0), \\ \infty & \text{else,} \end{cases}$$

because $u \in N^{-1}_{T^2_\Lambda(g(\bar{x});\nabla g(\bar{x})d)}(\lambda) \cap A = \emptyset$ *whenever* $\lambda \neq (0,0)$. *Thus we obtain*

$$\nabla^2_{xx} L(\bar{x}, \lambda) - \hat{\sigma}_{T^2_\Lambda(g(\bar{x});\nabla g(\bar{x})d),A}(\lambda) = \begin{cases} -1 & \text{if } \lambda = (0,0), \\ -\infty & \text{else} \end{cases}$$

violating condition (17). *Thus we can also conclude from Theorem 2 that* $\bar{x}$ *is not a local minimizer.*

On the other hand, for $\lambda := (1,0) \in \Lambda(\bar{x}; d)$ *we have* $N^{-1}_{T^2_\Lambda(g(\bar{x});\nabla g(\bar{x})d)}(\lambda) = \{-1\} \times \mathbb{R}$ *and* $\hat{\sigma}_{T^2_\Lambda(g(\bar{x});\nabla g(\bar{x})d)}(\lambda) = -1$ *follows. Hence*

$$\nabla^2_{xx} L(\bar{x}, \lambda) - \hat{\sigma}_{T^2_\Lambda(g(\bar{x});\nabla g(\bar{x})d)}(\lambda) = -1 - (-1) = 0$$

and the second-order necessary condition (18) *is fulfilled.*

Note that we can not apply Theorem 4 because condition (23) *fails to hold because of* $\nabla g(\bar{x})^T \lambda = 0$ *with* $\lambda = (1,0) \in N_\Lambda(g(\bar{x}); \nabla g(\bar{x})d)$.

Example 3. *Consider the MPCC*

$$\begin{aligned} \min \quad & f(x_1,x_2,x_3) := x_1 - x_2 + x_3 + \frac{1}{2}x_2^2 - x_3^2 \\ \text{s.t.} \quad & g(x_1,x_2,x_3) := \begin{pmatrix} -x_1 \\ -x_2 \\ -4x_1 + x_2 - x_3 \\ -3x_2 - x_3 \end{pmatrix} \in \Lambda := D_{CC} \times \mathbb{R}_- \times \mathbb{R}_-, \end{aligned}$$

at $\bar{x} = (0,0,0)$, *where* $D_{CC} := \{(a,b) \in \mathbb{R}^2_- \,|\, ab = 0\}$ *denotes the complementarity cone in* $\mathbb{R}^2$. *By considering points of the form* $(0,t,t)$ *with* $t > 0$ *it is easy to see that* $\bar{x}$ *is not a local minimizer and we want to verify this with our theory. We claim that* $C(\bar{x}) = \{(0,t,t) \,|\, t \geq 0\}$. *Indeed, the inclusion "*$\supseteq$*" obviously holds and we only have to verify the opposite inclusion. Since* g *is linear and* Λ *is a cone, every critical direction* d *must fulfill* $g(d) \in \Lambda$, *from which, by considering the third and fourth component of the system, the inequality*

$$d_2 - d_3 \leq \min\{4d_1, 4d_2\} = 0$$

follows. Together with $\nabla f(\bar{x})d = d_1 - (d_2 - d_3) \leq 0$ *and* $d_1 \geq 0$ *we conclude* $d_1 = 0$ *and* $d_3 = d_2 \geq 0$ *proving our claim. Now consider a critical direction* $d = (0,t,t)$ *with* $t > 0$. *By using [14, Lemma 4.1] we obtain*

$$N_\Lambda(g(\bar{x}); \nabla g(\bar{x})d) = \mathbb{R} \times \{0\} \times \mathbb{R}_+ \times \{0\} = N^c_\Lambda(g(\bar{x}); \nabla g(\bar{x})d)$$

and thus both DirRCQ (23) *and the directional non-degeneracy condition* (31) *are fulfilled. Straightforward calculation yield that* $\Lambda(\bar{x}; d) = \Lambda^c(\bar{x}; d) = \Lambda^s(\bar{x}; d) = \{\lambda_0\}$ *with* $\lambda_0 = (-3, 0, 1, 0)$. *Further,* $T^2_\Lambda(g(\bar{x}); \nabla g(\bar{x})d) = \{0\} \times \mathbb{R} \times \mathbb{R}_- \times \mathbb{R}$ *and* $\sigma_{T^2_\Lambda(g(\bar{x});\nabla g(\bar{x})d)}(\lambda_0) = 0$. *Hence*

$$\nabla^2_{xx} L(\bar{x}, \lambda_0)(d, d) - \sigma_{T^2_\Lambda(g(\bar{x});\nabla g(\bar{x})d)}(\lambda_0) = t^2 - 2t^2 = -t^2 < 0$$

and we can reject $\bar{x}$ *as a local minimizer by means of Corollary 5. Further we could reject* $\bar{x}$ *also by Theorem 2 together with Proposition 6(2) due to convexity of the second-order tangent set* $T^2_\Lambda(g(\bar{x}); \nabla g(\bar{x})d)$. *However, we cannot apply Corollary 6 because condition* (35) *is not fulfilled. This example demonstrates that directional non-degeneracy condition* (31)*is strictly weaker than the non-directional condition* (35)*, which, in terms of the literature on MPCC, is equivalent to MPCC-LICQ.*

Example 4. *Consider the problem*

$$\begin{aligned}
\min \quad & f(x_1, x_2, x_3) := x_1 - x_2 + x_3 + \frac{1}{2}x_2^2 - x_3^2 \\
\text{s.t.} \quad & g(x_1, x_2, x_3) := \begin{pmatrix} -x_1 \\ -x_2 \\ -4x_1 + x_2 - x_3 + x_2^2 \\ -3x_2 - x_3 \end{pmatrix} \in \Lambda := D_{CC} \times \mathbb{R}_- \times \mathbb{R}_-,
\end{aligned}$$

at $\bar{x} = (0, 0, 0)$, *which differs from the problem in the preceding example only by the presence of the term* x_2^2 *in the third component of* g. *Thus* $C(\bar{x})$ *remains unchanged and it is easy to see that* Λ *is outer second-order regular at* $g(\bar{x})$ *in direction* $\nabla g(\bar{x})d$ *for every critical direction* $d = (0, t, t)$ *with* $t > 0$. *Further,*

$$\nabla^2_{xx} L(\bar{x}, \lambda_0)(d, d) - \sigma_{T^2_\Lambda(g(\bar{x});\nabla g(\bar{x})d)}(\lambda_0) = t^2 > 0$$

and therefore the sufficient second-order condition (39) *is fulfilled implying that* $\bar{x}$ *is a strictly local minimizer satisfying the second-order growth condition.*

Example 5. *Consider the program*

$$\begin{aligned}
\min \quad & f(x_1, x_2) = -x_1 + \frac{1}{2}x_2^2 \\
\text{s.t.} \quad & g(x_1, x_2) = \begin{pmatrix} -x_1 \\ -x_2 \\ x_1^2 - x_2 \end{pmatrix} \in \Lambda := D_{cc} \times \mathbb{R}_-
\end{aligned}$$

at $\bar{x} = (0, 0)$. *Then* $C(\bar{x}) = (\{0\} \times \mathbb{R}_+) \cup (\mathbb{R}_+ \times \{0\})$ *and* Λ *is outer second-order regular at* $g(\bar{x})$ *in direction* $\nabla g(\bar{x})d$ *for every critical direction* $d \in C(\bar{x})$. *If* $d = (t, 0)$, $t > 0$, *then* $N_\Lambda(g(\bar{x}); \nabla g(\bar{x})d) = \{0\} \times \mathbb{R} \times \mathbb{R}_+$ *and* $T^2_\Lambda(g(\bar{x}); \nabla g(\bar{x})d) = \mathbb{R} \times \{0\} \times \mathbb{R}_-$ *and condition* (39) *is fulfilled with* $\alpha = 0$ *and* $\lambda = (0, -1, 1)$. *On the other hand, for* $d = (0, t)$, $t > 0$, *we have* $N_\Lambda(g(\bar{x}); \nabla g(\bar{x})d) = \mathbb{R} \times \{0\} \times \{0\}$ *and* $T^2_\Lambda(g(\bar{x}); \nabla g(\bar{x})d) = \{0\} \times \mathbb{R} \times \mathbb{R}$. *Now* (39) *is fulfilled with* $\alpha = 1$, $\lambda = (-1, 0, 0)$ *and we have verified that* $\bar{x}$ *is a strictly local minimizer fulfilling the second-order growth condition.*

References

[1] K. Bai, J.J. Ye and J. Zhang, *Directional quasi-/pseudo-normality as sufficient conditions for metric subregularity*, SIAM J. Optim., 29 (2019), pp. 2625-2649.

[2] M. Benko, H. Gfrerer and J.V. Outrata, *Calculus for directional limiting normal cones and subdifferentials*, Set-Valued Var. Anal., 27 (2019), pp. 713–745.

[3] J.F. Bonnans, R. Cominetti and A. Shapiro, *Second order optimality conditions based on parabolic second order tangent sets*, SIAM J. Optim., 9 (1999), pp. 466-492.

[4] J.F. Bonnans and H.C. Ramírez, *Perturbation analysis of second-order cone programming problems*, Math. Program., 104 (2005), pp. 205-227.

[5] J.F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer, 2000.

[6] F.H. Clarke, *Optimization and Nonsmooth Analysis*, Wiley-Interscience, New York, 1983.

[7] J.S. Chen, J.J. Ye, J. Zhang and J.C. Zhou, *Exact formula for the second-order tangent set of the second-order cone complementarity set*, SIAM J. Optim., to appear.

[8] C. Ding, D.F. Sun and J.J. Ye, *First order optimality conditions for mathematical programs with semidefinite cone complementarity constraints*, Math. Program., 147 (2014), pp. 539-579.

[9] A.L. Dontchev and R.T. Rockafellar, *Implicit Functions and Solution Mappings*, Springer, Heidelberg, 2014.

[10] M. L. Flegel, C. Kanzow and J. V. Outrata, *Optimality conditions for disjunctive programs with application to mathematical programs with equilibrium constraints*, Set-Valued Anal., 15 (2007), pp. 139–162.

[11] I. Ginchev and B.S. Mordukhovich, *On directionally dependent subdifferentials*, C.R. Bulg. Acad. Sci., 64 (2011), pp. 497-508.

[12] H. Gfrerer, *First order and second order characterizations of metric subregularity and calmness of constraint set mappings*, SIAM J. Optim., 21 (2011), pp. 1439-1474.

[13] H. Gfrerer, *On directional metric regularity, subregularity and optimality conditions for nonsmooth mathematical programs*, Set-Valued Var. Anal., 21 (2013), pp. 151-176.

[14] 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.

[15] H. Gfrerer *Linearized M-stationary conditions for general optimization problems*, Set-Valued Var. Anal., 27 (2019), pp. 819-840.

[16] H. Gfrerer and D. Klatte, *Lipschitz and Hölder stability of optimization problems and generalized equations*, Math. Program., 158 (2016), pp. 35-75.

[17] H. Gfrerer and J.V. Outrata, *On computation of generalized derivatives of the normal-cone mapping and their applications*, Math. Oper. Res., 41 (2016), pp. 1535-1556.

[18] H. Gfrerer and J.V. Outrata, *On Lipschitzian properties of implicit multifunctions*, SIAM J. Optim., 26 (2016), pp. 2160-2189.

[19] 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.

[20] H. Gfrerer, J.J. Ye, J. Zhang and J.C. Zhou, *Second-order optimality conditions for mathematical programs with second-order cone complementarity constraints*, forthcoming.

[21] L. Guo, G-H Lin and J.J. Ye, *Second-order optimality conditions for mathematical programs with equilibrium constraints*, J. Optim. Theory Appl., 158 (2013), pp. 33-64.

[22] R. Henrion and J.V. Outrata, *Calmness of constraint systems with applications*, Math. Program., 104 (2005), pp. 437–464.

[23] Z.Q. Luo, J-S. Pang and D. Ralph, *Mathematical Programs with Equilibrium Constraints*, Cambridge University Press, Cambridge, 1996.

[24] P. Mehlitz, *On the linear independence constraint qualification in disjunctive programming*, Optimization, doi:10.1080/02331934.2019.1679811.

[25] B.S. Mordukhovich, *Variational Analysis and Generalized Differentiation, Vol. 1: Basic Theory, Vol. 2: Applications*, Springer, Berlin, 2006.

[26] A. Mohammadi, B. S. Mordukhovich, E. Sarabi, *Parabolic regularity in geometric variational analysis*, preprint, 2019, arXiv:1909.00241.

[27] J.-P. Penot, *Optimality conditions in mathematical programmig and composite optimization*, Math. Program., 67 (1994), pp. 225–245.

[28] J.-P. Penot, *Second-order conditions for optimization problems with constraints*, SIAM J. Contr. Optim., 37 (1998), pp. 303–318.

[29] S.M. Robinson, *Some continuity properties of polyhedral multifunctions*, Math. Program. Stud., 14 (1981), pp. 206-214.

[30] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, 1970.

[31] R.T. Rockafellar and R.J. Wets, *Variational Analysis*, Springer, New York, 1998.

[32] H.S. Scheel and S. Scholtes, *Mathematical programs with complementarity constraints: stationarity, optimality, and sensitivity*, Math. Oper. Res., 25 (2000), pp. 1-22.

[33] 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.

[34] J.J. Ye and J.C. Zhou, *First order optimality conditions for mathematical programs with second-order cone complementarity constraints*, SIAM J. Optim., 26 (2016), pp. 2820-2846.

[35] J.J. Ye and J.C. Zhou, *Verifiable sufficient conditions for the error bound property of second-order cone complementarity problems*, Math. Program., 171 (2018), pp. 361-395.