New sharp necessary optimality conditions for mathematical programs with equilibrium constraints

Helmut Gfrerer[*]   Jane J. Ye[†]

Abstract

In this paper, we study the mathematical program with equilibrium constraints formulated as a mathematical program with a parametric generalized equation involving the regular normal cone. We derive a new necessary optimality condition which is sharper than the usual M-stationary condition and is applicable even when no constraint qualifications hold for the corresponding mathematical program with complementarity constraints reformulation.

Key words: mathematical programs with equilibrium constraints, constraint qualifications, necessary optimality conditions

AMS subject classification: 49J53, 90C30, 90C33, 90C46.

1   Introduction

In this paper we consider the *mathematical program with equilibrium constraints* (MPEC) of the form

$$\text{(MPEC)} \qquad \min_{x,y} \quad F(x,y) \tag{1}$$
$$\text{s.t.} \quad 0 \in \phi(x,y) + \widehat{N}_\Gamma(y),$$
$$G(x,y) \le 0,$$

where $\Gamma := \{y \,|\, g(y) \le 0\}$ and $\widehat{N}_\Gamma(y)$ denotes the so-called *regular normal cone* to the set Γ at y (see Definition 1). Here we assume that $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, $\phi : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$, $G : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p$ are continuously differentiable and $g : \mathbb{R}^m \to \mathbb{R}^q$ is twice continuously differentiable.

In the case where Γ is convex, $\widehat{N}_\Gamma(y) = N_\Gamma(y)$ is the normal cone in the sense of convex analysis and (MPEC) is equivalent to the *mathematical program with variational inequality constraints* (MPVIC) which arised in many applications from engineering and economics; see e.g. [19, 21] and the references within.

Up to now a common approach for handling (MPEC) is its reformulation as a *mathematical program with complementarity constraints* (MPCC). If at a point $y \in \Gamma$ a certain constraint qualification is fulfilled, then by the Karush-Kuhn-Tucker (KKT) condition,

$$0 \in \phi(x,y) + \widehat{N}_\Gamma(y) \iff \exists \lambda : 0 = \phi(x,y) + \nabla g(y)^T \lambda, \ \ 0 \leq -g(y) \perp \lambda \geq 0.$$

This observation yields the program

$$\begin{aligned}
\text{(MPCC)} \quad \min_{x,y,\lambda} \quad & F(x,y) \\
\text{s.t.} \quad & 0 = \phi(x,y) + \nabla g(y)^T \lambda, \\
& 0 \leq -g(y) \perp \lambda \geq 0, \\
& G(x,y) \leq 0,
\end{aligned}$$

which has been considered extensively in the literature during the last three decades. However, since in (MPCC) the minimization is over the original variable x, y as well as the multiplier λ, it is not equivalent to the original problem (MPEC) in general, cf.[4]. Moreover as discussed in [2, 17], there are many difficulties involved in using reformulation (MPCC) and therefore it is favorable to consider (MPEC) instead of (MPCC).

In Ye and Ye [26], the calmness/pseudo upper-Lipschitz continuity of the perturbed feasible mapping of (MPEC) has been proposed and proven to be a constraint qualification for the Mordukhovich (M-) stationarity to hold at a minimizer. The calmness of the perturbed feasible mapping of (MPEC) is known to be equivalent to the subregularity of the set-valued map

$$M_{\text{MPEC}}(x,y) := \left(\begin{array}{c} \phi(x,y) + \widehat{N}_\Gamma(y) \\ G(x,y) - \mathbb{R}^p_- \end{array} \right),$$

which we refer to be the *metric subregularity constraint qualification* (MSCQ). In [17, Theorem 5] (Theorem 4 in this paper), a concrete sufficient condition in terms of the problem data is provided for MSCQ. Continuing the work in [17], in this paper we aim at developing a new sharp necessary optimality condition for problem (MPEC).

Recently [12] (see Theorem 2 in this paper) derived a new necessary optimality condition for an optimization problem with a set-constraint in the form of $P(z) \in D$ where P is continuously differentiable and D is a closed set. The new optimality condition is derived in terms of the so-called *linearized Mordukhovich (M-) stationary condition* which is stronger than the usual M-stationarity condition.

It is easy to see that the constraint of (MPEC) can be rewritten in the form

$$P(x,y) := \left(\begin{array}{c} (y, -\phi(x,y)) \\ G(x,y) \end{array} \right) \in D := \text{gph}\,\widehat{N}_\Gamma \times \mathbb{R}^p_-$$

and hence (MPEC) can be treated as an optimization problem with the above set-constraint. In [12, Theorem 5], under some constraint qualifications on the lower level constraint $g(y) \leq 0$, which can be guaranteed to hold under the constant rank constraint qualification (CRCQ), the linearized M-stationary condition for (MPEC) is derived. In this paper, we drop this constraint qualification and we derive the linearized M-necessary optimality condition under the so-called 2-nondegeneracy condition on $g(y) \leq 0$.

We organize our paper as follows. Section 2 contains the preliminaries from variational geometry and variational analysis. In Section 3, we recall the linearized M- optimality conditions for the optimization problem with a set-constraint. In Section 4, we discuss constraint

qualifications for (MPEC). In Section 5, under the 2-nondegeneracy condition, we derive formula for regular normal cones to tangent directions that will be used in applying the necessary optimality condition from Section 3. Finally in Section 6, we reformulate (MPEC) in the form of an optimization problem with a set-constraint and apply the necessary optimality condition from Section 3.

The following notation will be used throughout the paper. We denote by $\mathcal{B}_{\mathbb{R}^q}$ the closed unit ball in $\mathbb{R}^q$ while when no confusion arises we denote it by $\mathcal{B}$. By $\mathcal{B}(\bar{z}; r)$ we denote the closed ball centered at $\bar{z}$ with radius r. For a matrix A, we denote by A^T its transpose. The inner product of two vectors x, y is denoted by $x^T y$ or $\langle x, y \rangle$ and by $x \perp y$ we mean $\langle x, y \rangle = 0$. For $\Omega \subseteq \mathbb{R}^d$ and $z \in \mathbb{R}^d$, we denote by $\mathrm{d}(z, \Omega)$ the distance from z to Ω. The polar cone of a set Ω is $\Omega^\circ := \{x | x^T v \leq 0 \ \forall v \in \Omega\}$ and $\Omega^\perp$ denotes the orthogonal complement to Ω. For a set Ω, we denote by $\mathrm{conv}\,\Omega$ and $\mathrm{cl}\,\Omega$ the convex hull and the closure of Ω, respectively. For a function $f : \mathbb{R}^d \to \mathbb{R}$, we denote by $\nabla f(\bar{z})$ the gradient vector of f at $\bar{z}$ and $\nabla^2 f(\bar{z})$ the Hessian matrix of f at $\bar{z}$. For a mapping $P : \mathbb{R}^d \to \mathbb{R}^s$ with $s > 1$, we denote by $\nabla P(z)$ the Jacobian matrix of P at z and for any given $w, v \in \mathbb{R}^d$, $w^T \nabla P(\bar{z})v$ is the vector in $\mathbb{R}^s$ with the ith component equal to $w^T \nabla^2 P_i(\bar{z})v, i = 1, \ldots, s$. Let $M : \mathbb{R}^d \rightrightarrows \mathbb{R}^s$ be an arbitrary set-valued mapping. We denote its graph by $\mathrm{gph} M := \{(z, w) | w \in M(z)\}$. $o : \mathbb{R}_+ \to \mathbb{R}$ denotes a function with the property that $o(\lambda)/\lambda \to 0$ when $\lambda \downarrow 0$.

2   Preliminaries from variational geometry and variational analysis

In this section, we gather some preliminaries and preliminary results in variational analysis that will be needed in the paper. The reader may find more details in the monographs [3, 20, 24] and in the papers we refer to.

Definition 1 (Tangent cone and normal cone). *Given a set $\Omega \subseteq \mathbb{R}^d$ and a point $\bar{z} \in \Omega$, the (Bouligand-Severi) tangent/contingent cone to Ω at $\bar{z}$ is a closed cone defined by*

$$T_\Omega(\bar{z}) := \limsup_{t \downarrow 0} \frac{\Omega - \bar{z}}{t} = \left\{ u \in \mathbb{R}^d \,\Big|\, \exists t_k \downarrow 0, \ u_k \to u \ \ with \ \ \bar{z} + t_k u_k \in \Omega \ \forall \ k \right\}.$$

The (Fréchet) regular normal cone and the (Mordukhovich) limiting/basic normal cone to Ω at $\bar{z} \in \Omega$ are closed cones defined by

$$\widehat{N}_\Omega(\bar{z}) := (T_\Omega(\bar{z}))^\circ$$

$$and \qquad N_\Omega(\bar{z}) := \left\{ z^* \,|\, \exists z_k \overset{\Omega}{\to} \bar{z} \ and \ z_k^* \to z^* \ such \ that \ z_k^* \in \widehat{N}_\Omega(z_k) \ \forall k \right\},$$

respectively.

When the set Ω is convex, the tangent/contingent cone and the regular/limiting normal cone reduce to the classical tangent cone and normal cone of convex analysis, respectively.

Definition 2 (Metric regularity and subregularity). *Let $M : \mathbb{R}^d \rightrightarrows \mathbb{R}^s$ be a set-valued mapping and let $(\bar{z}, \bar{w}) \in \mathrm{gph}\,M$.*

(i) We say that M is metrically subregular *at $(\bar{z}, \bar{w})$ if there exist a neighborhood Z of $\bar{z}$ and a positive number $\kappa > 0$ such that*

$$\mathrm{d}(z, M^{-1}(\bar{w})) \leq \kappa \mathrm{d}(\bar{w}, M(z)) \ \ \forall z \in Z.$$

(ii) We say that M is metrically regular around $(\bar{z}, \bar{w})$ if there exist neighborhoods Z of $\bar{z}$, W of $\bar{w}$ and a positive number $\kappa > 0$ such that

$$\mathrm{d}(z, M^{-1}(w)) \leq \kappa \mathrm{d}(w, M(z)) \;\; \forall (w, z) \in W \times Z.$$

It is well-known that metric subregularity of M at $(\bar{z}, \bar{w})$ is equivalent with the property of *calmness* of the inverse mapping M^{-1} at $(\bar{w}, \bar{z})$, cf. [6], whereas metric regularity of M around $(\bar{z}, \bar{w})$ is equivalent with the *Aubin property* of the inverse mapping M^{-1} around $(\bar{w}, \bar{z})$. It follows immediately from the definition that metric regularity of M around $(\bar{z}, \bar{w})$ implies metric subregularity. Further, metric subregularity of M at $(\bar{z}, \bar{w})$ is equivalent with metric subregularity of the mapping $z \to (z, \bar{w}) - \operatorname{gph} M$ at $(\bar{z}, (0, 0))$, cf. [17, Proposition 3].

Metric regularity can be verified via the so-called *Mordukhovich-criterion*. We give here only reference to a special case which is used in the sequel.

Theorem 1 (Mordukhovich criterion). *(cf. [24, Example 9.44]) Let $P : \mathbb{R}^d \to \mathbb{R}^s$ be continuously differentiable, let $D \subseteq \mathbb{R}^s$ be closed and let $P(\bar{z}) \in D$. Then the mapping $z \rightrightarrows P(z) - D$ is metrically regular around $(\bar{z}, 0)$ if and only if*

$$\nabla P(\bar{z})^T w^* = 0, \;\; w^* \in N_D\big(P(\bar{z})\big) \;\; \Longrightarrow \;\; w^* = 0. \tag{2}$$

For verifying the property of metric subregularity there are some sufficient conditions known, see e.g. [8, 9, 10, 11, 13].

Definition 3 (Critical cone). *For a closed set $\Omega \subseteq \mathbb{R}^d$, a point $z \in \Omega$ and a regular normal $z^* \in \widehat{N}_\Omega(z)$ we denote by*

$$\mathcal{K}_\Omega(z, z^*) := T_\Omega(z) \cap [z^*]^\perp$$

the critical cone *to Ω at (z, z^*)*

In this paper polyhedrality will play an important role.

Definition 4 (Polyhedrality). *1. Let $C \subseteq \mathbb{R}^d$.*

(a) We say that C is convex polyhedral, *if it can be written as the intersection of finitely many halfspaces, i.e. there are elements $(a_i, \alpha_i) \in \mathbb{R}^d \times \mathbb{R}$, $i = 1, \ldots, p$ such that $C = \{z \mid \langle a_i, z \rangle \leq \alpha_i, \; i = 1, \ldots, p\}$.*

(b) C is said to be polyhedral, *if it is the union of finitely many convex polyhedral sets.*

(c) Given a point $c \in C$, we say that C is locally polyhedral near c *if there is a neighborhood W of c and a polyhedral set $\tilde{C}$ such that $C \cap W = \tilde{C} \cap W$.*

2. A mapping $M : \mathbb{R}^d \rightrightarrows \mathbb{R}^s$ is called polyhedral, *if its graph $\operatorname{gph} M$ is a polyhedral set.*

Lemma 1. *Let $\Omega \subseteq \mathbb{R}^d$ be locally polyhedral near some point $\bar{z} \in \Omega$. Then*

$$N_\Omega(\bar{z}) = \bigcup_{w \in T_\Omega(\bar{z})} \widehat{N}_{T_\Omega(\bar{z})}(w).$$

Proof. Follows from [11, Lemma 2.2]. $\qquad\square$

We recall some properties of closed cones.

Proposition 1. *Let K be a closed cone in $\mathbb{R}^d$. Then $T_K(0) = K$.*

A consequence of the above property is that for any $y \in D$,

$$\widehat{N}_{T_D(y)}(0) = \left(T_{T_D(y)}(0)\right)^\circ = \left(T_D(y)\right)^\circ = \widehat{N}_D(y). \tag{3}$$

The following rules for calculating polar cones will be useful.

Proposition 2. *([23, Corollary 16.4.2]) Let A, B be nonempty convex cones in $\mathbb{R}^d$. Then $(A + B)^\circ = A^\circ \cap B^\circ$. If both A and B are closed then $(A \cap B)^\circ = \mathrm{cl}\,(A^\circ + B^\circ)$.*

In this paper we use Proposition 2 solely in situations, when the closure operation in the last formula can be omitted, namely, when either $A^\circ + B^\circ$ is a subspace or when both A and B are convex polyhedral cones, cf. [23, Corollaries 19.2.2, 19.3.2].

In the following proposition we collect some important facts about the normal cone mapping to a convex polyhedral set C, which can be extracted from [5].

Proposition 3 (Normal cone to a convex polyhedral set)**.** *Let $C \subseteq \mathbb{R}^d$ be a polyhedral convex set and $(\bar{z}, \bar{z}^*) \in \mathrm{gph}\, N_C$. Then for all $z \in C$ sufficiently close to $\bar{z}$ we have*

$$T_C(z) \supseteq T_C(\bar{z}), \quad N_C(z) \subseteq N_C(\bar{z}). \tag{4}$$

Further, there exists a neighborhood W of $(\bar{z}, \bar{z}^*)$ such that

$$\mathrm{gph}\, N_C \cap W = ((\bar{z}, \bar{z}^*) + \mathrm{gph}\, N_{\mathcal{K}_C(\bar{z}, \bar{z}^*)}) \cap W.$$

In particular we have

$$T_{\mathrm{gph}\, N_C}(\bar{z}, \bar{z}^*) = \mathrm{gph}\, N_{\mathcal{K}_C(\bar{z}, \bar{z}^*)}. \tag{5}$$

Further,

$$\widehat{N}_{\mathrm{gph}\, N_C}(\bar{z}, \bar{z}^*) = (\mathcal{K}_C(\bar{z}, \bar{z}^*))^\circ \times \mathcal{K}_C(\bar{z}, \bar{z}^*) \tag{6}$$

and the limiting normal cone $N_{\mathrm{gph}\, N_C}(\bar{z}, \bar{z}^*)$ is the union of all sets of the form

$$(F_1 - F_2)^\circ \times (F_1 - F_2)$$

where $F_2 \subseteq F_1$ are faces of $\mathcal{K}_C(\bar{z}, \bar{z}^*)$.

Definition 5 (Recession Cone)**.** *([23, page 61]) Let $C \subseteq \mathbb{R}^d$ be a closed convex set. The recession cone of C is a closed convex cone defined as $0^+C := \{y \in \mathbb{R}^d | x + \lambda y \in C \quad \forall \lambda \geq 0, x \in C\}$.*

Definition 6 (Generalized lineality space)**.** *Given an arbitrary set $C \subseteq \mathbb{R}^d$, we call a subspace L the generalized lineality space of C and denote it by $\mathcal{L}(C)$ provided that it is the largest subspace $L \subseteq \mathbb{R}^d$ such that $C + L \subseteq C$.*

Note that $\mathcal{L}(C)$ is well defined because for two subspaces L_1, L_2 fulfilling $C + L_i \subseteq C$, $i = 1, 2$ we have $C + L_1 + L_2 = (C + L_1) + L_2 \subseteq C + L_2 \subseteq C$ and hence we can always find a largest subspace satisfying $C + L \subseteq C$ since the dimension $\mathbb{R}^d$ is finite. Note that since 0 is in every subspace we have $C + L \supseteq C$ and thus $C + \mathcal{L}(C) = C$. In the case where C is a convex set, the generalized lineality space reduces to the lineality space as defined in [23, page 65] and can be calculated as $\mathcal{L}(C) = (-0^+C) \cap 0^+C$. In the case where C is a convex cone, the lineality space of C is the largest subspace contained in C and can be calculated as $\mathcal{L}(C) = (-C) \cap C$.

By definition of the generalized lineality space and the tangent cone, it is easy to verify that for every $\bar{z} \in C$ we have

$$\mathcal{L}(C) \subseteq \mathcal{L}\big(T_C(\bar{z})\big). \tag{7}$$

For a closed convex set C and $(\bar{z}, \bar{z}^*) \in \operatorname{gph} N_C$ we have $\mathcal{L}\big(T_C(\bar{z})\big) \subseteq [\bar{z}^*]^\perp$ and thus

$$\mathcal{L}\big(\mathcal{K}_C(\bar{z}, \bar{z}^*)\big) = \big(T_C(\bar{z}) \cap [\bar{z}^*]^\perp\big) \cap \big(-T_C(\bar{z}) \cap [\bar{z}^*]^\perp\big) = T_C(\bar{z}) \cap (-T_C(\bar{z})) \cap [\bar{z}^*]^\perp \tag{8}$$
$$= \mathcal{L}\big(T_C(\bar{z})\big)$$

Definition 7 (Affine Hull). *For a closed convex set C we denote by*

$$C^+ := \operatorname{span}(C - C)$$

the unique subspace parallel to the affine hull of C.

If K is a closed convex cone then we always have $K^+ = K - K$.

Let C be a closed convex set. Then the tangent cone $T_C(\bar{z})$ is a closed convex cone for any $\bar{z} \in C$. Since $C \subset \bar{z} + T_C(\bar{z})$ and $T_C(\bar{z}) = \limsup_{t \downarrow 0}(C - \bar{z})/t \subseteq \operatorname{span}(C - C) = C^+$ for any $\bar{z} \in C$, we have

$$C^+ = \operatorname{span}(C - C) \subseteq \operatorname{span}(T_C(\bar{z}) - T_C(\bar{z})) = T_C(\bar{z}) - T_C(\bar{z}) \subseteq C^+ - C^+ = C^+.$$

It follows that for a closed convex set C and every $\bar{z} \in C$ we have

$$C^+ = T_C(\bar{z}) - T_C(\bar{z}) = T_C(\bar{z})^+ \tag{9}$$

and for every $(\bar{z}, \bar{z}^*) \in \operatorname{gph} N_C$

$$\mathcal{K}_C(\bar{z}, \bar{z}^*)^+ = T_C(\bar{z}) \cap [\bar{z}^*]^\perp - T_C(\bar{z}) \cap [\bar{z}^*]^\perp \subseteq T_C(\bar{z})^+ \cap [\bar{z}^*]^\perp.$$

For every closed convex set C and every $\bar{z} \in C$ we have by virtue of Proposition 2 that

$$\mathcal{L}\big(T_C(\bar{z})\big)^\perp = \big(T_C(\bar{z}) \cap (-T_C(\bar{z}))\big)^\circ = \operatorname{cl}\big(N_C(\bar{z}) - N_C(\bar{z})\big) = N_C(\bar{z})^+, \tag{10}$$

and hence

$$\mathcal{L}\big(T_C(\bar{z})\big) = \big(N_C(\bar{z})^+\big)^\perp. \tag{11}$$

Further, by virtue of Proposition 2 we have

$$\mathcal{L}(N_C(\bar{z})) = \big(T_C(\bar{z})\big)^\circ \cap \big(-T_C(\bar{z})\big)^\circ = (T_C(\bar{z}) - T_C(\bar{z}))^\circ = (T_C(\bar{z})^+)^\perp = (C^+)^\perp$$

implying

$$\mathcal{L}(\operatorname{gph} N_C) = \mathcal{L}(C) \times (C^+)^\perp. \tag{12}$$

If C is convex polyhedral, then for every $(\bar{z}, \bar{z}^*) \in \operatorname{gph} N_C$ we obtain by virtue of (5), (12), (8) and (11)

$$\mathcal{L}\big(T_{\operatorname{gph} N_C}(\bar{z}, \bar{z}^*)\big) = \mathcal{L}\big(\operatorname{gph} N_{\mathcal{K}_C(\bar{z}, \bar{z}^*)}\big) = \mathcal{L}\big(\mathcal{K}_C(\bar{z}, \bar{z}^*)\big) \times \big(\mathcal{K}_C(\bar{z}, \bar{z}^*)^+\big)^\perp \tag{13}$$
$$= \mathcal{L}(T_C(\bar{z})) \times \big(\mathcal{K}_C(\bar{z}, \bar{z}^*)^+\big)^\perp = \big(N_C(\bar{z})^+\big)^\perp \times \big(\mathcal{K}_C(\bar{z}, \bar{z}^*)^+\big)^\perp.$$

Further, for every $(\delta \bar{z}, \delta \bar{z}^*) \in \mathrm{gph}\, N_{\mathcal{K}_C(\bar{z}, \bar{z}^*)}$, by (7) we have

$$\mathcal{L}\big(T_{\mathrm{gph}\, N_{\mathcal{K}_C(\bar{z}, \bar{z}^*)}}(\delta\bar{z}, \delta\bar{z}^*)\big) \supseteq \mathcal{L}\big(\mathrm{gph}\, N_{\mathcal{K}_C(\bar{z}, \bar{z}^*)}\big) = \mathcal{L}\big(\mathcal{K}_C(\bar{z}, \bar{z}^*)\big) \times \big(\mathcal{K}_C(\bar{z}, \bar{z}^*)^+\big)^\perp \tag{14}$$

$$= \mathcal{L}\big(T_C(\bar{z})\big) \times \big(\mathcal{K}_C(\bar{z}, \bar{z}^*)^+\big)^\perp = \mathcal{L}\big(T_{\mathrm{gph}\, N_C}(\bar{z}, \bar{z}^*)\big),$$

where the equalities follow from (13).

In the following proposition we recall some basic properties of convex polyhedral cones.

Proposition 4. *Consider two finite index sets I_1, I_2, vectors $a_i \in \mathbb{R}^n$, $i \in I_1 \cup I_2$ and let*

$$K := \left\{ v \,\Big|\, a_i^T v \begin{cases} = 0 & i \in I_1, \\ \leq 0 & i \in I_2 \end{cases} \right\}.$$

Then

$$\mathcal{L}(K) = \left\{ v \,\Big|\, a_i^T v = 0, \ i \in I_1 \cup I_2 \right\}, \quad K^+ \subseteq \left\{ v \,\Big|\, a_i^T v = 0, \ i \in I_1 \right\}$$

and for every $v \in K$ we have

$$T_K(v) = \left\{ u \,\Big|\, a_i^T u \begin{cases} = 0 & i \in I_1 \\ \leq 0 & i \in I(v) \setminus I_1 \end{cases} \right\}, \quad N_K(v) = \left\{ \sum_{i \in I(v)} \mu_i a_i \,\Big|\, \mu_i \geq 0, i \in I(v) \setminus I_1 \right\}$$

where $I(v) := \{ i \in I_1 \cup I_2 \,|\, a_i^T v = 0 \}$. Further, for every $z^* \in N_K(v)$ there is an index set $\mathcal{I}$ with $I_1 \subseteq \mathcal{I} \subseteq I(v)$ such that

$$\mathcal{K}_K(v, z^*) = \left\{ u \,\Big|\, a_i^T u \begin{cases} = 0 & i \in \mathcal{I} \\ \leq 0 & i \in I(v) \setminus \mathcal{I} \end{cases} \right\}$$

and vice versa. The faces of K are given by the sets

$$\mathcal{F} = \left\{ u \,\Big|\, a_i^T u \begin{cases} = 0 & i \in \mathcal{I} \\ \leq 0 & i \in I_2 \setminus \mathcal{I} \end{cases} \right\}, \quad \text{where } \mathcal{I} \text{ satisfies } \ I_1 \subseteq \mathcal{I} \subseteq I_1 \cup I_2.$$

For all $v \in K$, the following face of K defined by

$$\mathcal{F}_v := \left\{ u \,\Big|\, a_i^T u \begin{cases} = 0 & i \in I(v) \\ \leq 0 & i \in I_2 \setminus I(v) \end{cases} \right\}$$

is the unique face satisfying $v \in \mathrm{ri}\, \mathcal{F}_v$. Consequently, for all $v \in K$ and all faces $\mathcal{F}_1, \mathcal{F}_2$ of K such that $v \in \mathrm{ri}\, \mathcal{F}_2 \subseteq \mathcal{F}_1$ there is some index set $\mathcal{I}$, $I_1 \subset \mathcal{I} \subseteq I(v)$ such that

$$\mathcal{F}_1 - \mathcal{F}_2 = \left\{ u \,\Big|\, a_i^T u \begin{cases} = 0 & i \in \mathcal{I} \\ \leq 0 & i \in I(v) \setminus \mathcal{I} \end{cases} \right\},$$

which is the same as saying that there is some $z^* \in N_K(v)$ with $\mathcal{F}_1 - \mathcal{F}_2 = \mathcal{K}_K(v, z^*)$.

The following lemma will be useful for our analysis:

Lemma 2. *Let $\tilde{P} = (\tilde{P}_1, \tilde{P}_2) : \mathbb{R}^d \to \mathbb{R}^s \times \mathbb{R}^s$ be continuously differentiable, let $C \subseteq \mathbb{R}^s$ be a polyhedral convex set and let $\bar{z} \in \mathbb{R}^d$ with $\tilde{P}(\bar{z}) \in \tilde{D} := \mathrm{gph}\, N_C$ be given. Further assume that we are given two subspaces $L_1 \supseteq \big(N_C(\tilde{P}_1(\bar{z}))\big)^+$ and $L_2 \supseteq \big(\mathcal{K}_C(\tilde{P}_1(\bar{z}), \tilde{P}_2(\bar{z}))\big)^+$ such that*

$$\ker \nabla \tilde{P}(\bar{z})^T \cap (L_1 \times L_2) = \{(0,0)\}. \tag{15}$$

Then the mapping $z \rightrightarrows \tilde{P}(z) - \tilde{D}$ is metrically regular around $(\bar{z}, 0)$,

$$T_{\{z \,|\, \tilde{P}(z) \in \tilde{D}\}}(\bar{z}) = \{w \,|\, \nabla \tilde{P}(\bar{z}) w \in T_{\tilde{D}}(\tilde{P}(\bar{z}))\} \tag{16}$$

and

$$\begin{aligned}
\widehat{N}_{\{z \,|\, \tilde{P}(z) \in \tilde{D}\}}(\bar{z}) &= \nabla \tilde{P}(\bar{z})^T \widehat{N}_{\tilde{D}}(\tilde{P}(\bar{z})) \\
&= \nabla \tilde{P}_1(\bar{z})^T \big(\mathcal{K}_C(\tilde{P}_1(\bar{z}), \tilde{P}_2(\bar{z}))\big)^\circ + \nabla \tilde{P}_2(\bar{z})^T \mathcal{K}_C(\tilde{P}_1(\bar{z}), \tilde{P}_2(\bar{z})).
\end{aligned} \tag{17}$$

Proof. In order to prove metric regularity of the mapping $\tilde{P}(\cdot) - \tilde{D}$ we invoke the Mordukhovich criterion (2), which reads in our case as

$$\nabla \tilde{P}_1(\bar{z})^T w^* + \nabla \tilde{P}_2(\bar{z})^T w = 0, \ (w^*, w) \in N_{\mathrm{gph}\, N_C}(\tilde{P}(\bar{z})) \ \Rightarrow \ (w^*, w) = 0. \tag{18}$$

Consider $(w^*, w) \in N_{\mathrm{gph}\, N_C}(\tilde{P}(\bar{z}))$. By Proposition 3 there are faces F_1, F_2 of the convex polyhedral cone $\tilde{K} := \mathcal{K}_C(\tilde{P}_1(\bar{z}), \tilde{P}_2(\bar{z}))$ such that $(w^*, w) \in (F_1 - F_2)^\circ \times (F_1 - F_2)$. Since the lineality space of a convex polyhedral cone is always contained in any of its faces, we have $\mathcal{L}(\tilde{K}) \subseteq F_1 - F_2 \subset \tilde{K}^+$, from which we obtain

$$(w^*, w) \in \mathcal{L}(\tilde{K})^\circ \times \tilde{K}^+ = \mathcal{L}(\tilde{K})^\perp \times \tilde{K}^+ = \big(N_C(\tilde{P}_1(\bar{z}))\big)^+ \times \big(\mathcal{K}_C(\tilde{P}_1(\bar{z}), \tilde{P}_2(\bar{z}))\big)^+ \subseteq L_1 \times L_2,$$

where the second equality follows by using (8) and (10). Thus (18) follows from (15) and the claimed property of metric regularity is established. Metric regularity in turn implies MSCQ for the system $\tilde{P}(z) \in \tilde{D}$ at $\bar{z}$ and (16) follows from [18, Proposition 1]. In order to show (17) we will invoke [16, Theorem 4]. From (13) we deduce

$$\mathcal{L}\big(T_{\tilde{D}}(\tilde{P}(\bar{z}))\big) = \Big(N_C\big(\tilde{P}_1(\bar{z})\big)^+\Big)^\perp \times \Big(\mathcal{K}_C\big(\tilde{P}_1(\bar{z}), \tilde{P}_2(\bar{z})\big)^+\Big)^\perp \supseteq L_1^\perp \times L_2^\perp$$

and together with (15) we obtain

$$\begin{aligned}
\mathbb{R}^s \times \mathbb{R}^s &= \Big(\ker \nabla \tilde{P}(\bar{z})^T \cap (L_1 \times L_2)\Big)^\perp = \mathrm{Range}\, \nabla \tilde{P}(\bar{z}) + (L_1^\perp \times L_2^\perp) \\
&\subseteq \mathrm{Range}\, \nabla \tilde{P}(\bar{z}) + \mathcal{L}\big(T_{\tilde{D}}(\tilde{P}(\bar{z}))\big) \subseteq \mathbb{R}^s \times \mathbb{R}^s,
\end{aligned}$$

where the second equality follows from Proposition 2. Hence the assumption of [16, Theorem 4] is fulfilled and the first equation in (17) follows, whereas the second equation is a consequence of (6). $\qquad\square$

3 Optimality conditions for a set-constrained optimization problem

In this section we consider an optimization problem of the form

$$\begin{aligned}
\min \quad & f(z) \\
\mathrm{s.t.} \quad & P(z) \in D,
\end{aligned} \tag{19}$$

where $f : \mathbb{R}^d \to \mathbb{R}$ and $P : \mathbb{R}^d \to \mathbb{R}^s$ are continuously differentiable and $D \subseteq \mathbb{R}^s$ is closed.

Let $\bar{z}$ be a local minimizer and $\Omega := \{z \,|\, P(z) \in D\}$ the feasible region for the problem (19). Then $\nabla f(\bar{z})^T u \geq 0 \quad \forall u \in T_\Omega(\bar{z})$ and so the following basic/geometric optimality condition holds:
$$0 \in \nabla f(\bar{z}) + \widehat{N}_\Omega(\bar{z}). \tag{20}$$

To express the basic optimality condition in terms of the problem data $P(\cdot)$ and D, one needs to estimate the regular normal cone $\widehat{N}_\Omega(\bar{z})$. Given $\bar{z} \in \Omega$ we denote the *linearized tangent cone to Ω at $\bar{z}$* by
$$T_{P,D}^{\mathrm{lin}}(\bar{z}) := \{u \in \mathbb{R}^d \,|\, \nabla P(\bar{z})u \in T_D(P(\bar{z}))\}.$$

It is well known that the inclusions
$$T_\Omega(\bar{z}) \subseteq T_{P,D}^{\mathrm{lin}}(\bar{z}), \;\; \widehat{N}_\Omega(\bar{z}) \supseteq \big(T_{P,D}^{\mathrm{lin}}(\bar{z})\big)^\circ, \tag{21}$$
$$\big(T_{P,D}^{\mathrm{lin}}(\bar{z})\big)^\circ \supseteq \nabla P(\bar{z})^T \widehat{N}_D\big(P(\bar{z})\big) \tag{22}$$

are always valid, cf. [24, Theorems 6.31, 6.14]. Hence, if both inclusions hold with equality, the formula $\widehat{N}_\Omega(\bar{z}) = \nabla P(\bar{z})^T \widehat{N}_D\big(P(\bar{z})\big)$ is at our disposal and the optimality condition (20) reads as
$$0 \in \nabla f(\bar{x}) + \nabla P(\bar{z})^T \widehat{N}_D\big(P(\bar{z})\big),$$

which is also known as *strong (S-) stationarity condition*, cf. [7]. In order to ensure equality in (21) one has to impose some constraint qualification.

Definition 8. *Let $P(\bar{z}) \in D$.*

(i) *(cf. [7]) We say that the* generalized Abadie constraint qualification *(GACQ) holds at $\bar{z}$ if*
$$T_\Omega(\bar{z}) = T_{P,D}^{\mathrm{lin}}(\bar{z}).$$

(ii) *(cf. [7]) We say that the* generalized Guignard constraint qualification *(GGCQ) holds at $\bar{z}$ if*
$$\widehat{N}_\Omega(\bar{z}) = \big(T_{P,D}^{\mathrm{lin}}(\bar{z})\big)^\circ. \tag{23}$$

(iii) *(cf. [14]) We say that the* metric subregularity constraint qualification (MSCQ) *holds at $\bar{z}$ for the system $P(z) \in D$ if the set-valued map $M(z) := P(z) - D$ is metrically subregular at $(\bar{z}, 0)$.*

There hold the following implications:
$$\mathrm{MSCQ} \implies \mathrm{GACQ} \implies \mathrm{GGCQ}.$$

Indeed, the first implication follows from [18, Proposition 1] whereas the second one is an immediate consequence of the definition of the regular normal cone. GGCQ is the weakest of the three constraint qualifications ensuring $\widehat{N}_\Omega(\bar{z}) = \big(T_{P,D}^{\mathrm{lin}}(\bar{z})\big)^\circ$, but it is very difficult to verify it in general. On the other hand, MSCQ is stronger than GGCQ but there are effective tools for verifying it.

Now let us consider inclusion (22). By [23, Corollary 16.3.2] we have
$$\mathrm{cl}\left(\nabla P(\bar{z})^T \widehat{N}_D\big(P(\bar{z})\big)\right) = \big\{w \,|\, \nabla P(\bar{z})w \in \mathrm{cl\,conv}\, T_D\big(P(\bar{z})\big)\big\}^\circ$$

showing that we can expect equality in (22) only under some restrictive assumption whenever $T_D(P(\bar{z}))$ is not convex. Such an assumption is e.g. provided by [16, Theorem 4]. If it does not hold, but MSCQ holds at $\bar{z}$, then it is well-known that $\widehat{N}_\Omega(\bar{z}) \subseteq N_\Omega(\bar{z}) \subseteq \nabla P(\bar{z})^T N_D(P(\bar{z}))$ and hence the M-stationary condition

$$0 \in \nabla f(\bar{z}) + \nabla P(\bar{z})^T N_D(P(\bar{z}))$$

holds at any local optimal solution $\bar{z}$. For the case where the set D is simple, e.g., $D := \{(a,b) \mid 0 \le a \perp b \ge 0\}$, the complementarity cone, the limiting normal cone can be calculated using the variational analysis (cf. [20]) and one obtains the classical M-stationary condition for MPCC. However, for more complicated set D, e.g., $D := \mathrm{gph}\,\widehat{N}_\Gamma \times \mathbb{R}^p_-$, usually very strong assumptions are required for using these calculus rules limiting considerably their applicability; see e.g. Gfrerer and Outrata [15, Theorem 4] .

Recently an alternative approach is taken by Gfrerer in [12]. Under GGCQ, by (23) for every regular normal $z^* \in \widehat{N}_\Omega(\bar{z})$ the point $u = 0$ is a global minimizer for the problem

$$\min_u -z^{*T}u \quad \text{subject to} \quad \nabla P(\bar{z})u \in T_D(P(\bar{z})).$$

Provided that the mapping $u \rightrightarrows \nabla P(\bar{z})u - T_D(P(\bar{z}))$ is metrically subregular at $(0,0)$ we can apply the M-stationarity conditions to this linearized problem which read as

$$z^* \in \nabla P(\bar{z})^T N_{T_D(P(\bar{z}))}(0).$$

Thus we obtain the inclusion $\widehat{N}_\Omega(\bar{z}) \subseteq \nabla P(\bar{z})^T N_{T_D(P(\bar{z}))}(0)$. This results in a necessary optimality condition

$$0 \in \nabla f(\bar{z}) + \nabla P(\bar{z})^T N_{T_D(P(\bar{z}))}(0), \tag{24}$$

which is sharper than the M-stationarity condition since $N_{T_D(P(\bar{z}))}(0) \subseteq N_D(P(\bar{z}))$, cf. [24, Proposition 6.27(a)]. Although (24) is a sharper condition than the M-stationary condition, it still involves the limiting normal cone and so may be hard to calculate. In [12, Propositions 1,2], Gfrerer derived the following linearized M-necessary optimality condition which can be considered as a refinement of the necessary optimality condition (24). The condition is easier to calculate since it involves only the regular normal cone. In fact by virtue of Lemma 1, in the case where $T_D(P(\bar{z}))$ is locally polyhedral at 0, condition (25) coincides with condition (24).

Theorem 2. *Let $\bar{z}$ be a local optimal solution for problem* (19). *Assume that GGCQ holds at $\bar{z}$ and the mapping $u \rightrightarrows \nabla P(\bar{z})u - T_D(P(\bar{z}))$ is metrically subregular at $(0,0)$. Then one of the following two conditions is fulfilled:*

(i) *There is $\omega \in T_D(P(\bar{z}))$ such that*

$$0 \in \nabla f(\bar{z}) + \nabla P(\bar{z})^T \widehat{N}_{T_D(P(\bar{z}))}(\omega). \tag{25}$$

(ii) *There is $\bar{u} \in T^{\mathrm{lin}}_{P,D}(\bar{z})$ such that*

$$\nabla P(\bar{z})\bar{u} \notin \mathcal{L}(T_D(P(\bar{z}))), \tag{26}$$
$$\nabla f(\bar{z})^T \bar{u} = 0, \tag{27}$$
$$0 \in \nabla f(\bar{z}) + \widehat{N}_{T^{\mathrm{lin}}_{P,D}(\bar{z})}(\bar{u}) \tag{28}$$

and $T_D(P(\bar{z}))$ is not locally polyhedral near $\nabla P(\bar{z})\bar{u}$.

If in addition $T_D(P(\bar{z}))$ is the graph of a set-valued mapping $M = M_c + M_p$, where $M_c, M_p :$ $\mathbb{R}^r \rightrightarrows \mathbb{R}^{s-r}$ are set-valued mappings whose graphs are closed cones, M_p is polyhedral and there is some real C such that

$$\|t\| \leq C\|v\| \ \forall(v,t) \in \operatorname{gph} M_c \tag{29}$$

then there is some $\bar{v} \neq 0$ such that

$$\nabla P(\bar{z})\bar{u} \in \{\bar{v}\} \times M(\bar{v}). \tag{30}$$

Remark 1. *(i) Note that the assumptions of Theorem 2 are fulfilled if MSCQ holds at $\bar{z}$. Indeed, MSCQ implies GGCQ and metric subregularity of $u \rightrightarrows \nabla P(\bar{z})u - T_D(P(\bar{z}))$ at $(0,0)$ follows from [12, Lemma 4].*

(ii) Note that when $T_D(P(\bar{z}))$ is locally polyhedral near $\nabla P(\bar{z})\bar{u}$, then condition (i) holds. Otherwise, if $T_D(P(\bar{z}))$ is not locally polyhedral near $\nabla P(\bar{z})\bar{u}$, then condition (ii) can hold. In this case, (27) implies together with GGCQ and the basic optimality condition (20) that $\bar{u}$ is a global solution of the problem

$$\min \nabla f(\bar{z})^T u \quad subject \ to \ \nabla P(\bar{z})u \in T_D(P(\bar{z})).$$

Now since the graph of the mapping $u \rightrightarrows \nabla P(\bar{z})u - T_D(P(\bar{z}))$ is a closed cone, by virtue of [12, Lemma 3], the metric subregularity of $u \rightrightarrows \nabla P(\bar{z})u - T_D(P(\bar{z}))$ at $(0,0)$ implies the metric subregularity of the same mapping at $(\bar{u}, 0)$. Hence we can apply Theorem 2 once more to the above problem. If $T_{T_D(P(\bar{z}))}(\nabla P(\bar{z})\bar{u})$ is polyhedral, then Theorem 2(i) applies and we obtain the existence of $\omega \in T_{T_D(P(\bar{z}))}(\nabla P(\bar{z})\bar{u})$ such that

$$0 \in \nabla f(\bar{z}) + \nabla P(\bar{z})^T \widehat{N}_{T_{T_D(P(\bar{z}))}(\nabla P(\bar{z})\bar{u})}(\omega). \tag{31}$$

In this case (31) would be the necessary optimality condition which is sharper than condition (25). In this paper we aim at finding some sufficient conditions, i.e., the 2-nondegeneracy condition introduced in Subsection 5.1 below, under which the tangent cone $T_{T_D(P(\bar{z}))}(\nabla P(\bar{z})\bar{u})$ is polyhedral and hence the above optimality condition holds for (MPEC). However, in general $T_{T_D(P(\bar{z}))}(\nabla P(\bar{z})\bar{u})$ is not polyhedral, and the process could continue. The interested reader is referred to [12] for the discussion for what might have happened after applying Theorem 2 repeatedly.

4 Constraint qualifications for the new optimality conditions

Note that (MPEC) can be written in the form (19) via

$$\min_{x,y} \quad F(x,y) \tag{32}$$

$$\text{subject to} \quad P(x,y) := \begin{pmatrix} (y, -\phi(x,y)) \\ G(x,y) \end{pmatrix} \in D := \operatorname{gph} \widehat{N}_\Gamma \times \mathbb{R}^p_-,$$

where $\Gamma := \{y \,|\, g(y) \leq 0\}$. To apply Theorem 2, we will need the following assumptions at a local solution $(\bar{x}, \bar{y})$ to problem (32).

Assumption 1. *(i) MSCQ holds for the lower level constraint $g(y) \in \mathbb{R}^q_-$ at $\bar{y}$.*

(ii) GGCQ holds at $(\bar{x}, \bar{y})$ and the mapping

$$(u, v) \quad \rightrightarrows \quad \nabla P(\bar{x}, \bar{y})(u, v) - T_D((\bar{y}, \bar{y}^*), G(\bar{x}, \bar{y}))$$
$$= \quad \begin{pmatrix} (v, -\nabla\phi(\bar{x}, \bar{y}))(u, v) \\ \nabla G(\bar{x}, \bar{y})(u, v) \end{pmatrix} - T_{\mathrm{gph}\,\widehat{N}_\Gamma \times \mathbb{R}^p_-}((\bar{y}, \bar{y}^*), G(\bar{x}, \bar{y}))$$

is metrically subregular at $((0, 0), 0)$, where $\bar{y}^* := -\phi(\bar{x}, \bar{y})$.

It is well-known that if either g is affine or the Mangasarian-Fromovitz constraint qualification (MFCQ) holds at $\bar{z}$ then Assumption 1(i) holds. By Remark 1, Assumption 1(ii) is fulfilled if MSCQ holds for the system $P(x, y) \in D$ at $(\bar{x}, \bar{y})$. A point-based sufficient condition for the validity of MSCQ for this system is given by [17, Theorem 5]. We now describe this condition. When MSCQ holds at $\bar{y}$ for the system $g(y) \in \mathbb{R}^q_-$, we have $T_\Gamma(\bar{y}) = \{v \,|\, \nabla g(\bar{y})v \in T_{\mathbb{R}^q_-}(g(\bar{y}))\}$ and thus the critical cone $\mathcal{K}_\Gamma(\bar{y}, \bar{y}^*)$ amounts to

$$\bar{K}_\Gamma := \mathcal{K}_\Gamma(\bar{y}, \bar{y}^*) = \{v \,|\, \nabla g(\bar{y})v \in T_{\mathbb{R}^q_-}(g(\bar{y}))\} \cap [\bar{y}^*]^\perp, \tag{33}$$

which is convex polyhedral. Further we define the *multiplier set* for $\bar{y}$ as the polyhedral convex set defined by

$$\bar{\Lambda} := \Lambda(\bar{y}, \bar{y}^*) := \{\lambda \in N_{\mathbb{R}^q_-}(g(\bar{y})) \,|\, \nabla g(\bar{y})^T \lambda = \bar{y}^*\}. \tag{34}$$

For a multiplier λ, the corresponding collection of *strict complementarity indexes* is denoted by

$$I^+(\lambda) := \{i \in \{1, \dots, q\} \,|\, \lambda_i > 0\} \quad \text{for} \quad \lambda = (\lambda_1, \dots, \lambda_q) \in \mathbb{R}^q_+.$$

Denote by $\mathcal{E}(\bar{y}, \bar{y}^*)$ the collection of all the extreme points of the closed and convex set of multipliers $\Lambda(\bar{y}, \bar{y}^*)$ and recall that $\lambda \in \Lambda(\bar{y}, \bar{y}^*)$ belongs to $\mathcal{E}(\bar{y}, \bar{y}^*)$ if and only the family of gradients $\{\nabla g_i(\bar{y}) \,|\, i \in I^+(\lambda)\}$ is linearly independent. Moreover for every $v \in \mathcal{K}_\Gamma(\bar{y}, \bar{y}^*)$, we define the *directional multiplier set* as

$$\bar{\Lambda}(v) := \Lambda(\bar{y}, \bar{y}^*; v) := \arg\max\left\{v^T \nabla^2(\lambda^T g)(\bar{y})v \,|\, \lambda \in \bar{\Lambda}\right\},$$

which is also a polyhedral convex set. By [14, Proposition 4.3(iii)] we have $\bar{\Lambda}(v) \neq \emptyset \; \forall v \in \bar{K}_\Gamma$ under Assumption 1. Note that since $\bar{\Lambda}$ is a closed convex set and the objective of the above problem is linear, by the optimality condition,

$$\lambda \in \bar{\Lambda}(v) \Longleftrightarrow v^T \nabla^2 g(\bar{y})v \in N_{\bar{\Lambda}}(\lambda). \tag{35}$$

Theorem 3 (cf. [17, Theorem 4]). *Let $\bar{y} \in \Gamma := \{y \,|\, g(y) \leq 0\}$ and $\bar{y}^* = -\phi(\bar{x}, \bar{y})$. Assume that MSCQ holds at $\bar{y}$ for the system $g(y) \in \mathbb{R}^q_-$. Then the tangent cone to the graph of $\widehat{N}_\Gamma$ at $(\bar{y}, \bar{y}^*)$ can be calculated by*

$$T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar{y}, \bar{y}^*) \tag{36}$$
$$= \left\{(v, v^*) \in \mathbb{R}^{2m} \,|\, \exists \lambda \in \bar{\Lambda}(v) \; \text{with} \; v^* \in \nabla^2(\lambda^T g)(\bar{y})v + N_{\bar{K}_\Gamma}(v)\right\}$$
$$= \left\{(v, v^*) \in \mathbb{R}^{2m} \,|\, \exists \lambda \in \bar{\Lambda}(v) \cap \kappa\|\bar{y}^*\|\mathcal{B}_{\mathbb{R}^q} \; \text{with} \; v^* \in \nabla^2(\lambda^T g)(\bar{y})v + N_{\bar{K}_\Gamma}(v)\right\},$$

where $\kappa > 0$ is certain constant. Or equivalently

$$T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar{y}, \bar{y}^*) \tag{37}$$
$$= \left\{(v, \nabla^2(\lambda^T g)(y)v + z^*) \in \mathbb{R}^{2m} \,|\, (z^*, v^T \nabla^2 g(\bar{y})v) \in N_{\bar{K}_\Gamma \times \bar{\Lambda}}(v, \lambda)\right\}$$
$$= \left\{(v, \nabla^2(\lambda^T g)(y)v + z^*) \in \mathbb{R}^{2m} \,|\, (z^*, v^T \nabla^2 g(\bar{y})v) \in N_{\bar{K}_\Gamma \times \bar{\Lambda}}(v, \lambda), \; \|\lambda\| \leq \kappa\|\bar{y}^*\|\right\}.$$

Note that the equivalence of (36) and (37) is s due to (35).

We now in a position to review a sufficent condition for Assumption 1 to hold.

Theorem 4 ([17, Theorem 5]). *Let $(\bar{x}, \bar{y})$ be a feasible solution of the system $P(x,y) \in D$. Assume that MSCQ holds both for the lower level problem constraints $g(y) \leq 0$ at $\bar{y}$ and for the upper level constraints $G(x,y) \leq 0$ at $(\bar{x}, \bar{y})$. Further assume that*

$$\nabla_x G(\bar{x}, \bar{y})^T \eta = 0, \ \eta \in N_{\mathbb{R}^p_-}(G(\bar{x}, \bar{y})) \quad \Longrightarrow \quad \nabla_y G(\bar{x}, \bar{y})^T \eta = 0,$$

and assume that there do not exist $(u,v) \neq 0$, $\lambda \in \Lambda(\bar{y}, -\phi(\bar{x}, \bar{y}); v) \cap \mathcal{E}(\bar{y}, -\phi(\bar{x}, \bar{y}))$, $\eta \in \mathbb{R}^p_+$ and $w \neq 0$ satisfying

$$\nabla G(\bar{x}, \bar{y})(u,v) \in T_{\mathbb{R}^p_-}(G(\bar{x}, \bar{y})),$$
$$(v, -\nabla_x \phi(\bar{x}, \bar{y})u - \nabla_y \phi(\bar{x}, \bar{y})v) \in T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar{y}, -\phi(\bar{x}, \bar{y})),$$
$$-\nabla_x \phi(\bar{x}, \bar{y})^T w + \nabla_x G(\bar{x}, \bar{y})^T \eta = 0, \ \eta \in N_{\mathbb{R}^p_-}(G(\bar{x}, \bar{y})), \ \eta^T \nabla G(\bar{x}, \bar{y})(u,v) = 0,$$
$$\nabla g_i(\bar{y})w = 0, i \in I^+(\lambda), \ w^T \left(\nabla_y \phi(\bar{x}, \bar{y}) + \nabla^2(\lambda^T g(\bar{y})) \right) w - \eta^T \nabla_y G(\bar{x}, \bar{y})w \leq 0.$$

Then MSCQ for the system $P(x,y) \in D$ holds at $(\bar{x}, \bar{y})$.

5 Computing regular normals to tangent directions and tangents of tangents

In this section, we apply Theorem 2 to obtain a necessary optimality condition for program (32) which is equivalent to the MPEC (1). In order to apply Theorem 2, for any $(\bar{v}, \bar{v}^*, a) := w \in T_D(P(\bar{x}, \bar{y}))$, we need to compute $\widehat{N}_{T_D(P(\bar{x}, \bar{y}))}(w)$. Using [17, Proposition 1] together with [24, Proposition 6.41], we obtain

$$\widehat{N}_{T_D(P(\bar{x}, \bar{y}))}(w) = \widehat{N}_{T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)}(\bar{v}, \bar{v}^*) \times \widehat{N}_{T_{\mathbb{R}^p_-}(G(\bar{x}, \bar{y}))}(a).$$

Hence the aim of this section is to compute the regular tangent cone to the tangent directions $\widehat{N}_{T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)}(\bar{v}, \bar{v}^*)$. Similarly by virtue of (31) we also need to compute the regular tangent cone to the tangents of tangents $\widehat{N}_{T_{T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)}(\bar{v}, \bar{v}^*)}(\delta\bar{v}, \delta\bar{v}^*)$.

As discussed in the introduction, under a certain constraint qualification such as CRCQ, formulas for $\widehat{N}_{T_D(P(\bar{x}, \bar{y}))}(w)$ and the resulting optimality condition for (1) are derived in [12, Proposition 3, Theorem 5]. In this paper we use a different approach. Given $(\bar{v}, \bar{v}^*) \in T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)$, by the formula for the tangent cone (36), there is some $\lambda \in \bar{\Lambda}(\bar{v})$ and $z^* \in N_{\bar{K}_\Gamma}(\bar{v})$ such that $\bar{v}^* \in \nabla^2(\lambda^T g)(\bar{y})\bar{v} + z^*$. Suppose that such representation is unique, i.e., there is a unique $\bar{\lambda} \in \bar{\Lambda}(\bar{v})$ and $\bar{z}^* \in N_{\bar{K}_\Gamma}(\bar{v})$ such that $\bar{v}^* = \nabla^2(\bar{\lambda}^T g)(\bar{y})\bar{v} + \bar{z}^*$. Then

$$(\bar{v}, \bar{v}^*) \in T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar{y}, \bar{y}^*) \Longleftrightarrow \bar{v}^* = \nabla^2(\bar{\lambda}^T g)(\bar{y})\bar{v} + \bar{z}^*.$$

The uniqueness allows the efficient calculation of the regular normal cone to tangent directions. To guarantee this uniqueness we perform our analysis under the assumption of 2-nondegeneracy on g as introduced in the next subsection.

5.1 2-nondegeneracy

Definition 9. *Let $\bar{v} \in \bar{K}_\Gamma$. We say that g is 2-nondegenerate in direction $\bar{v}$ at $(\bar{y}, \bar{y}^*)$ if*

$$\nabla^2(\mu^T g)(\bar{y})\bar{v} \in (N_{\bar{K}_\Gamma}(\bar{v}))^+, \ \mu \in (\bar{\Lambda}(\bar{v}))^+ \implies \mu = 0.$$

In the case where the directional multiplier set $\bar{\Lambda}(\bar{v})$ is a singleton, $(\bar{\Lambda}(\bar{v}))^+ = \{0\}$ and hence g is 2-nondegenerate in this direction $\bar{v}$. In particular, if $\bar{\Lambda}$ is a singleton then g is 2-nondegenerate in any direction $\bar{v}$. We now provide a formulation of 2-nondegeneracy in terms of index sets. To this end let us define

$$\bar{I} := \{i \in \{1, \ldots, q\} \mid g_i(\bar{y}) = 0\}, \quad \bar{I}(v) := \{i \in \bar{I} \mid \nabla g_i(\bar{y})^T v = 0\}, v \in \bar{K}_\Gamma,$$
$$\bar{J}^+(\lambda) := \{i \in \bar{I} \mid \lambda_i > 0\}, \lambda \in \bar{\Lambda}, \quad \bar{J}^+(\Xi) := \bigcup_{\lambda \in \Xi} \bar{J}^+(\lambda) \text{ for any } \Xi \subseteq \bar{\Lambda}.$$

By the definition of the critical cone in (33), we have

$$\bar{K}_\Gamma = \left\{ v \mid \nabla g_i(\bar{y})^T v \leq 0, \ i \in \bar{I} \right\} \cap [\bar{y}^*]^\perp.$$

Since by the definition of the multiplier set (34),

$$\lambda \in \bar{\Lambda} \iff \lambda \in N_{\mathbb{R}^q_-}(g(\bar{y})), \quad \bar{y}^* = \nabla g(\bar{y})^T \lambda$$

we have

$$v \in [\bar{y}^*]^\perp \iff 0 = \bar{y}^{*T} v = \langle \nabla g(\bar{y})^T \lambda, v \rangle = \lambda^T \nabla g(\bar{y}) v.$$

Hence it is obvious that for every $\lambda \in \bar{\Lambda}$ we have

$$\bar{K}_\Gamma = \left\{ v \mid \nabla g_i(\bar{y})^T v \begin{cases} = 0 & i \in \bar{J}^+(\lambda) \\ \leq 0 & i \in \bar{I} \setminus \bar{J}^+(\lambda) \end{cases} \right\}$$

yielding

$$\bar{K}_\Gamma = \left\{ v \mid \nabla g_i(\bar{y})^T v \begin{cases} = 0 & i \in \bar{J}^+(\bar{\Lambda}) \\ \leq 0 & i \in \bar{I} \setminus \bar{J}^+(\bar{\Lambda}) \end{cases} \right\} = \left\{ v \mid \nabla g_i(\bar{y})^T v \begin{cases} = 0 & i \in \bar{J}^+(\bar{\Lambda}(\bar{v})) \\ \leq 0 & i \in \bar{I} \setminus \bar{J}^+(\bar{\Lambda}(\bar{v})) \end{cases} \right\}.$$

Thus

$$N_{\bar{K}_\Gamma}(\bar{v}) = \{ \sum_{i \in \bar{I}(\bar{v})} \eta_i \nabla g_i(\bar{y}) \mid \eta_i \geq 0, i \in \bar{I}(\bar{v}) \setminus \bar{J}^+(\bar{\Lambda}(\bar{v})) \}$$

Now choose $\hat{J}$ with $\bar{J}^+(\bar{\Lambda}(\bar{v})) \subseteq \hat{J} \subseteq \bar{I}(\bar{v})$ large enough such that for every $j \in \bar{I}(\bar{v}) \setminus \hat{J}$ the gradient $\nabla g_j(\bar{y})$ linearly depend on $\nabla g_i(\bar{y})$, $i \in \hat{J}$. It follows that

$$(N_{\bar{K}_\Gamma}(\bar{v}))^+ = \{ \sum_{i \in \bar{I}(\bar{v})} \eta_i \nabla g_i(\bar{y}) \mid \eta_i \in \mathbb{R}, i \in \bar{I}(\bar{v}) \} = \{ \sum_{i \in \hat{J}} \eta_i \nabla g_i(\bar{y}) \mid \eta_i \in \mathbb{R}, i \in \hat{J} \}.$$

Next we claim that

$$(\bar{\Lambda}(\bar{v}))^+ = L := \{ \mu \in \mathbb{R}^q \mid \nabla g(\bar{y})^T \mu = 0, \ \bar{v}^T \nabla^2(\mu^T g)(\bar{y})\bar{v} = 0, \mu_i = 0, \ i \notin \bar{J}^+(\bar{\Lambda}(\bar{v})) \}.$$

Indeed, for every pair $\lambda^1, \lambda^2 \in \bar{\Lambda}(\bar{v})$, we have $\bar{y}^* = \nabla g(\bar{y})^T \lambda^1 = \nabla g(\bar{y})^T \lambda^2$ and

$$\bar{v}^T \nabla^2 (\lambda_1^T g)(\bar{y}) \bar{v} = \bar{v}^T \nabla^2 (\lambda_2^T g)(\bar{y}) \bar{v},$$

which implies that

$$\nabla g(\bar{y})^T (\lambda^1 - \lambda^2) = 0, \ \bar{v}^T \nabla^2 ((\lambda^1 - \lambda^2)^T g)(\bar{y}) \bar{v} = 0, \lambda_i^1 - \lambda_i^2 = 0, \ i \notin \bar{J}^+ (\bar{\Lambda}(\bar{v}))$$

showing $\left(\bar{\Lambda}(\bar{v}) \right)^+ \subseteq L$. To show the reverse inclusion, take any $\mu \in L$ and any $\bar{\lambda} \in \bar{\Lambda}(\bar{v})$. Then $\bar{\lambda} + \alpha\mu \geq 0$ for all $\alpha > 0$ sufficiently small. It is easy to see that $\bar{\lambda} + \alpha\mu \in \bar{\Lambda}(\bar{v})$ implying $L \subseteq \left(\bar{\Lambda}(\bar{v}) \right)^+$. Thus our claim holds true and we obtain that g is *2-nondegenerate in direction* $\bar{v}$ *at* $(\bar{y}, \bar{y}^*)$ if and only if

$$\sum_{i \in \hat{J}} \eta_i \nabla g_i(\bar{y}) + \sum_{i \in \bar{J}^+ (\bar{\Lambda}(\bar{v}))} \mu_i \nabla^2 g_i(\bar{y}) \bar{v} = 0, \quad \sum_{i \in \bar{J}^+ (\bar{\Lambda}(\bar{v}))} \mu_i \nabla g_i(\bar{y}) = 0 \ \Rightarrow \mu_i = 0, \ i \in \bar{J}^+ (\bar{\Lambda}(\bar{v})). \tag{38}$$

We now want to compare 2-nondegeneracy with the notion of *2-regularity* which was initiated (and named) by Tret'yakov [25] in the case of zero Jacobian and then was strongly developed by Avakov [1]. A twice continuously differentiable mapping $h : \mathbb{R}^m \to \mathbb{R}^l$ is called 2-regular at a point $\bar{y} \in \mathbb{R}^m$ in direction $v \in \mathbb{R}^m$ if for all $\alpha \in \mathbb{R}^l$ the system

$$\nabla h(\bar{y}) u + v^T \nabla^2 h(\bar{y}) w = \alpha, \ \nabla h(\bar{y}) w = 0$$

has a solution (u, w). We claim that 2-regularity of $(g_i)_{i \in \hat{J}}$ implies 2-nondegeneracy of g in direction $\bar{v}$. Indeed, by the Farkas lemma 2-regularity of $(g_i)_{i \in \hat{J}}$ in direction $\bar{v}$ is equivalent to the statement

$$\sum_{i \in \hat{J}} (\eta_i \nabla g_i(\bar{y}) + \mu_i \nabla^2 g_i(\bar{y}) \bar{v}) = 0, \ \sum_{i \in \hat{J}} \mu_i \nabla g_i(\bar{y}) = 0 \ \Rightarrow \mu_i = 0, \ i \in \hat{J}.$$

and it is easy to see that this condition implies (38).

The following lemma states some important consequences of 2-nondgeneracy.

Lemma 3. *Assume that g is 2-nondegenerate in the critical direction $\bar{v} \in \bar{K}_\Gamma$ at $(\bar{y}, \bar{y}^*)$ and define the subspace*

$$\mathcal{H}(\bar{v}) := \{ \nabla^2 (\mu^T g)(\bar{y}) \bar{v} \mid \mu \in \left(\bar{\Lambda}(\bar{v}) \right)^+ \} + (N_{\bar{K}_\Gamma}(\bar{v}))^+.$$

Then the linear mapping $\mathcal{A}_{\bar{v}} : \left(\bar{\Lambda}(\bar{v}) \right)^+ \times (N_{\bar{K}_\Gamma}(\bar{v}))^+ \to \mathcal{H}(\bar{v})$ given by

$$\mathcal{A}_{\bar{v}}(\mu, z^*) := \nabla^2 (\mu^T g)(\bar{y}) \bar{v} + z^*$$

is a bijection. In particular, for every $\bar{v}^*$ with $(\bar{v}, \bar{v}^*) \in T_{\text{gph} \, \widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)$ there are unique elements $\bar{\lambda} \in \bar{\Lambda}(\bar{v})$ and $\bar{z}^* \in N_{\bar{K}_\Gamma}(\bar{v})$

$$\bar{v}^* = \nabla^2 (\bar{\lambda}^T g)(\bar{y}) \bar{v} + \bar{z}^*. \tag{39}$$

Proof. By the definition, the mapping $\mathcal{A}_{\bar{v}}$ is surjective and therefore we only have to show injectivity. Consider elements $(\mu, \bar{z}^*) \in \left(\bar{\Lambda}(\bar{v}) \right)^+ \times (N_{\bar{K}_\Gamma}(\bar{v}))^+$ satisfying $\mathcal{A}_{\bar{v}}(\mu, \bar{z}^*) = 0$. Then $\nabla^2 (\mu^T g)(\bar{y}) \bar{v} = -\bar{z}^* \in (N_{\bar{K}_\Gamma}(\bar{v}))^+$ and by the assumed 2-nondegeneracy of g in direction $\bar{v}$ we obtain $\mu = 0$ and consequently $\bar{z}^* = 0$. Thus $\mathcal{A}_{\bar{v}}$ is injective.

In order to show the second statement consider $\bar{v}^*$ with $(\bar{v}, \bar{v}^*) \in T_{\text{gph}\,\widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)$. The existence of $(\bar{\lambda}, \bar{z}^*) \in \bar{\Lambda}(\bar{v}) \times N_{\bar{K}_\Gamma}(\bar{v})$ fulfilling (39) follows from Theorem 3. In order to prove uniqueness of the representation (39), consider $(\lambda_1, z_1^*), (\lambda_2, z_2^*) \in \bar{\Lambda}(\bar{v}) \times N_{\bar{K}_\Gamma}(\bar{v})$ such that

$$\bar{v}^* = \nabla^2(\lambda_j^T g)(\bar{y})\bar{v} + z_j^*, \ j = 1, 2$$

implying

$$\mathcal{A}_{\bar{v}}(\lambda_2 - \lambda_1, z_2^* - z_1^*) = \nabla^2((\lambda_2 - \lambda_1)^T g)(\bar{y})\bar{v} + z_2^* - z_1^* = 0.$$

Then $\lambda_2 - \lambda_1 \in \big(\bar{\Lambda}(\bar{v})\big)^+$ and $z_2^* - z_1^* \in (N_{\bar{K}_\Gamma}(\bar{v}))^+$ and by the injectivity of $\mathcal{A}_{\bar{v}}$ we obtain $\lambda_2 = \lambda_1$ and $z_2^* = z_1^*$. $\hfill\square$

5.2 Regular normals to tangent directions

Throughout this subsection let $(\bar{v}, \bar{v}^*) \in T_{\text{gph}\,\widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)$ be given. The main purpose of this section is to compute the regular normal cone of the tangent directions $\widehat{N}_{T_{\text{gph}\,\widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)}(\bar{v}, \bar{v}^*)$.

Proposition 5. *Assume that g is 2-nondegenerate in the critical direction $\bar{v} \in \bar{K}_\Gamma$ at $(\bar{y}, \bar{y}^*)$. Then for every $\bar{v}^*$ with $(\bar{v}, \bar{v}^*) \in T_{\text{gph}\,\widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)$ we have*

$$T_{T_{\text{gph}\,\widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)}(\bar{v}, \bar{v}^*) = \left\{ (u, u^*) \,\Big|\, \exists \mu, \zeta^* \ s.t. \ \begin{matrix} u^* = \nabla^2(\bar{\lambda}^T g)(\bar{y})u + \nabla^2(\mu^T g)(\bar{y})\bar{v} + \zeta^*, \\ (u, \mu, \zeta^*, 2\bar{v}^T \nabla^2 g(\bar{y})u) \in \text{gph}\, N_{\tilde{K}(\bar{v}, \bar{v}^*)} \end{matrix} \right\}, \ (40)$$

where $(\bar{\lambda}, \bar{z}^*) \in \bar{\Lambda}(\bar{v}) \times N_{\bar{K}_\Gamma}(\bar{v})$ is the unique element fulfilling $\bar{v}^* = \nabla^2(\bar{\lambda}^T g)(\bar{y})\bar{v} + \bar{z}^*$ and

$$\tilde{K}(\bar{v}, \bar{v}^*) := \mathcal{K}_{\bar{K}_\Gamma \times \bar{\Lambda}}(\bar{v}, \bar{\lambda}, \bar{z}^*, \bar{v}^T \nabla^2 g(\bar{y})\bar{v}).$$

Further,

$$\widehat{N}_{T_{\text{gph}\,\widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)}(\bar{v}, \bar{v}^*)$$
$$= \left\{ (w^*, w) \,\Big|\, \begin{matrix} \exists \eta \ s.t. \ \big(w^* + \nabla^2(\bar{\lambda}^T g)(\bar{y})w - 2\nabla^2(\eta^T g)(\bar{y})\bar{v}, \bar{v}^T \nabla^2 g(\bar{y})w, w, \eta\big) \\ \in \big(\tilde{K}(\bar{v}, \bar{v}^*)\big)^\circ \times \tilde{K}(\bar{v}, \bar{v}^*) \end{matrix} \right\}. \ (41)$$

Proof. Let $\bar{v}^*$ with $(\bar{v}, \bar{v}^*) \in T_{\text{gph}\,\widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)$ be fixed and let $\mathcal{R}$ denote the set on the right hand side of equation (40).

Step 1. In this step we will show that $T_{T_{\text{gph}\,\widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)}(\bar{v}, \bar{v}^*) \subseteq \mathcal{R}$. Let $(u, u^*) \in T_{T_{\text{gph}\,\widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)}(\bar{v}, \bar{v}^*)$. Then by definition of the tangent cone, there exists sequences $t_k \downarrow 0$, $(u_k, u_k^*) \to (u, u^*)$ with $(\bar{v} + t_k u_k, \bar{v}^* + t_k u_k^*) \in T_{\text{gph}\,\widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)$. By (36) there are sequences $\lambda_k \in \bar{\Lambda}(\bar{v} + t_k u_k) \cap \kappa\|\bar{y}^*\|\mathcal{B}_{\mathbb{R}^q}$ and $z_k^* \in N_{\bar{K}_\Gamma}(\bar{v} + t_k u_k)$ such that

$$\bar{v}^* + t_k u_k^* = \nabla^2(\lambda_k^T g)(\bar{y})(\bar{v} + t_k u_k) + z_k^*.$$

Moreover, since by Lemma 3 there are unique elements $\bar{\lambda} \in \bar{\Lambda}(\bar{v})$ and $\bar{z}^* \in N_{\bar{K}_\Gamma}(\bar{v})$ satisfying $\bar{v}^* = \nabla^2(\bar{\lambda}^T g)(\bar{y})\bar{v} + \bar{z}^*$, it follows that

$$\nabla^2((\lambda_k - \bar{\lambda})^T g)(\bar{y})\bar{v} + z_k^* - \bar{z}^* = t_k\big(u_k^* - \nabla^2(\lambda_k^T g)(\bar{y})u_k\big). \tag{42}$$

16

For all k sufficiently large we have $N_{\bar{K}_\Gamma}(\bar{v} + t_k u_k) \subseteq N_{\bar{K}_\Gamma}(\bar{v})$ by (4) and $\bar{\Lambda}(\bar{v} + t_k u_k) \subseteq \bar{\Lambda}(\bar{v})$ by [16, Lemma 3]. Hence we have $\lambda_k - \bar{\lambda} \in (\bar{\Lambda}(\bar{v}))^+$ and $z_k^* - \bar{z}^*$ in $(N_{\bar{K}_\Gamma}(\bar{v}))^+$. Thus from (42), we have

$$\mathcal{A}_{\bar{v}}(\lambda_k - \bar{\lambda}, z_k^* - \bar{z}^*) := \nabla^2((\lambda_k - \bar{\lambda})^T g)(\bar{y})\bar{v} + z_k^* - \bar{z}^* = t_k(u_k^* - \nabla^2(\lambda_k^T g)(\bar{y})u_k).$$

By the boundedness of λ_k we conclude $t_k(u_k^* - \nabla^2(\lambda_k^T g)(\bar{y})u_k) \to 0$. Hence, by Lemma 3 we have $(\lambda_k - \bar{\lambda}, z_k^* - \bar{z}^*) \to (0,0)$ and

$$(\mu, \zeta^*) := \lim_{k \to \infty}(\mu_k, \zeta_k^*) = \lim_{k \to \infty} \mathcal{A}_{\bar{v}}^{-1}(u_k^* - \nabla^2(\lambda_k^T g)(\bar{y})u_k) = \mathcal{A}_{\bar{v}}^{-1}(u^* - \nabla^2(\bar{\lambda}^T g)(\bar{y})u),$$

where $\mu_k := \frac{\lambda_k - \bar{\lambda}}{t_k}$ and $\zeta_k^* := \frac{z_k^* - \bar{z}^*}{t_k}$. Thus

$$u^* = \nabla^2(\bar{\lambda}^T g)(\bar{y})u + \mathcal{A}_{\bar{v}}(\mu, \zeta^*) = \nabla^2(\bar{\lambda}^T g)(\bar{y})u + \nabla^2(\mu^T g)(\bar{y})\bar{v} + \zeta^*. \tag{43}$$

Since $z_k^* \in N_{\bar{K}_\Gamma}(\bar{v} + t_k u_k)$ and $\lambda_k \in \bar{\Lambda}(\bar{v} + t_k u_k)$ which is equivalent to saying that $(\bar{v} + t_k u_k)^T \nabla^2 g(\bar{y})(\bar{v} + t_k u_k)) \in N_{\bar{\Lambda}}(\lambda_k)$ by virtue of (35), we have

$$(\bar{v} + t_k u_k, \lambda_k, z_k^*, (\bar{v} + t_k u_k)^T \nabla^2 g(\bar{y})(\bar{v} + t_k u_k)) \in \operatorname{gph} N_{\bar{K}_\Gamma \times \bar{\Lambda}}.$$

It follows that from definition of tangent cone and the above that

$$(u, \mu, \zeta^*, 2\bar{v}^T \nabla^2 g(\bar{y})u) \in T_{\operatorname{gph} N_{\bar{K}_\Gamma \times \bar{\Lambda}}}(\bar{v}, \bar{\lambda}, \bar{z}^*, \bar{v}^T \nabla^2 g(\bar{y})\bar{v}) = \operatorname{gph} N_{\tilde{K}(\bar{v}, \bar{v}^*)},$$

where the equation follows from (5). Thus combining the above inclusion and (43), we have that $(u, u^*) \in \mathcal{R}$ and the inclusion $T_{T_{\operatorname{gph} \hat{N}_\Gamma}(\bar{y}, \bar{y}^*)}(\bar{v}, \bar{v}^*) \subseteq \mathcal{R}$ is shown.

Step 2. Now we show the reverse inclusion $T_{T_{\operatorname{gph} \hat{N}_\Gamma}(\bar{y}, \bar{y}^*)}(\bar{v}, \bar{v}^*) \supseteq \mathcal{R}$ in (40). Let $(u, u^*) \in \mathcal{R}$. Then there exist μ, ζ^* such that

$$u^* = \nabla^2(\bar{\lambda}^T g)(\bar{y})u + \nabla^2(\mu^T g)(\bar{y})\bar{v} + \zeta^*,$$
$$(u, \mu, \zeta^*, 2\bar{v}^T \nabla^2 g(\bar{y})u) \in \operatorname{gph} N_{\tilde{K}(\bar{v}, \bar{v}^*)} = T_{\operatorname{gph} N_{\bar{K}_\Gamma \times \bar{\Lambda}}}(\bar{v}, \bar{\lambda}, \bar{z}^*, \bar{v}^T \nabla^2 g(\bar{y})\bar{v}),$$

where the second equation follows from (5).

First by applying Lemma 2 , we wish to show that

$$\Theta := T_{\{(v, \lambda, z^*) \,|\, \tilde{P}(v, \lambda, z^*) \in \tilde{D}\}}(\bar{v}, \bar{\lambda}, \bar{z}^*) = \{(u, \mu, \zeta^*) \,|\, \nabla\tilde{P}(\bar{v}, \bar{\lambda}, \bar{z}^*)(u, \mu, \zeta^*) \in T_{\tilde{D}}(\tilde{P}(\bar{v}, \bar{\lambda}, \bar{z}^*))\}$$
$$= \{(u, \mu, \zeta^*) \,|\, (u, \mu, \zeta^*, 2\bar{v}^T \nabla^2 g(\bar{y})u) \in T_{\operatorname{gph} N_{\bar{K}_\Gamma \times \bar{\Lambda}}}(\bar{v}, \bar{\lambda}, \bar{z}^*, \bar{v}^T \nabla^2 g(\bar{y})\bar{v})\}, \tag{44}$$

where $\tilde{P}_1(v, \lambda, z^*) := (v, \lambda)$, $\tilde{P}_2(v, \lambda, z^*) := (z^*, v^T \nabla^2 g(\bar{y})v)$, $\tilde{D} := \operatorname{gph} N_{\bar{K}_\Gamma \times \bar{\Lambda}}$. Let $\bar{z} := (\bar{v}, \bar{\lambda}, \bar{z}^*)$. Define $L_1 := (N_{\bar{K}_\Gamma}(\bar{v}))^+ \times \mathbb{R}^q \supseteq (N_{\bar{K}_\Gamma \times \bar{\Lambda}}(\tilde{P}_1(\bar{z})))^+$ and $L_2 := \mathbb{R}^m \times (\bar{\Lambda}(\bar{v}))^+$. Note that

$$\mathcal{K}_{\bar{\Lambda}}(\bar{\lambda}, \bar{v}^T \nabla^2 g(\bar{y})\bar{v}) = T_{\bar{\Lambda}}(\bar{\lambda}) \cap [\bar{v}^T \nabla^2 g(\bar{y})\bar{v}]^\perp = T_{\bar{\Lambda}(\bar{v})}(\bar{\lambda}), \tag{45}$$

where the second equality follows from the fact that $\mu \in T_{\bar{\Lambda}}(\bar{\lambda}) \cap [\bar{v}^T \nabla^2 g(\bar{y})\bar{v}]^\perp$ if and only if $\bar{\lambda} + \alpha\mu \in \bar{\Lambda}(\bar{v})$ for all $\alpha \geq 0$ sufficiently small. It follows together with (9) that

$$L_2 := \mathbb{R}^m \times (\bar{\Lambda}(\bar{v}))^+ \supseteq (\mathcal{K}_{\bar{K}_\Gamma \times \bar{\Lambda}}(\tilde{P}_1(\bar{z}), \tilde{P}_2(\bar{z})))^+.$$

Next consider $(w^*, \lambda^*, z, \mu) \in L_1 \times L_2$ satisfying

$$\nabla \tilde{P}(\bar{z})^T(w^*, \lambda^*, z, \mu) = (w^* + 2\nabla^2(\mu^T g)(\bar{y})\bar{v}, \lambda^*, z) = (0, 0, 0).$$

Then $\nabla^2(\mu^T g)(\bar{y})\bar{v} = -w^*/2 \in \left(N_{\bar{K}_\Gamma}(\bar{v})\right)^+$ and by the assumed 2-nondegeneracy we obtain $\mu = 0$ and consequently $w^* = 0$. Because we also have $\lambda^* = 0$ and $z = 0$, (15) is verified and by (16) we obtain (44).

It follows from (44) that $(u, \mu, \zeta^*) \in \Theta$. Consequently by the definition of the tangent cone $T_{\{(v, \lambda, z^*) \mid \tilde{P}(v, \lambda, z^*) \in \tilde{D}\}}(\bar{v}, \bar{\lambda}, \bar{z}^*)$, there exist sequences $t_k \downarrow 0$ and $(u_k, \mu_k, \zeta_k^*) \to (u, \mu, \zeta^*)$ such that

$$\begin{aligned}
&\tilde{P}(\bar{v} + t_k u_k, \bar{\lambda} + t_k \mu_k, \bar{z}^* + t_k \zeta_k^*) \\
&= (\bar{v} + t_k u_k, \bar{\lambda} + t_k \mu_k, \bar{z}^* + t_k \zeta_k^*, (\bar{v} + t_k u_k)^T \nabla^2 g(\bar{y})(\bar{v} + t_k u_k)) \in \operatorname{gph} N_{\bar{K}_\Gamma \times \bar{\Lambda}}.
\end{aligned}$$

By (37) it follows that $(\bar{v} + t_k u_k, \nabla^2((\bar{\lambda} + t_k \mu_k)^T g)(\bar{y})(\bar{v} + t_k u_k) + \bar{z}^* + t_k \zeta_k^*) \in T_{\operatorname{gph} \widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)$ implying $(u, u^*) \in T_{T_{\operatorname{gph} \widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)}(\bar{v}, \bar{v}^*)$. Hence (40) is shown.

Step 3. To show (41), note that

$$\begin{aligned}
&\widehat{N}_{T_{\operatorname{gph} \widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)}(\bar{v}, \bar{v}^*) \\
&= (T_{T_{\operatorname{gph} \widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)}(\bar{v}, \bar{v}^*))^\circ \\
&= (\{(u, u^*) : u^* = \nabla^2(\bar{\lambda}^T g)(\bar{y})u + \nabla^2(\mu^T g)(\bar{y})\bar{v} + \zeta^*, (u, \mu, \zeta^*) \in \Theta\})^\circ \\
&= \{(w^*, w) \mid \langle w^*, u \rangle + \langle w, \nabla^2(\bar{\lambda}^T g)(\bar{y})u + \nabla^2(\mu^T g)(\bar{y})\bar{v} + \zeta^* \rangle \leq 0 \quad \forall (u, \mu, \zeta^*) \in \Theta\} \\
&= \{(w^*, w) \mid (w^* + \nabla^2(\bar{\lambda}^T g)(\bar{y})w, \bar{v}^T \nabla^2 g(\bar{y})w, w) \in \Theta^\circ\},
\end{aligned} \tag{46}$$

where the second equality follows from (40). By (17) together with (6) we have

$$\begin{aligned}
\Theta^\circ &= (T_{\{(v, \lambda, z^*) \mid \tilde{P}(v, \lambda, z^*) \in \tilde{D}\}}(\bar{v}, \bar{\lambda}, \bar{z}^*))^\circ \\
&= \widehat{N}_{\{(v, \lambda, z^*) \mid \tilde{P}(v, \lambda, z^*) \in \tilde{D}\}}(\bar{v}, \bar{\lambda}, \bar{z}^*)) \\
&= \nabla \tilde{P}(\bar{v}, \bar{\lambda}, \bar{z}^*)^T (\tilde{K}(\bar{v}, \bar{v}^*))^\circ \times \tilde{K}(\bar{v}, \bar{v}^*) \\
&= \{(v^* + 2\nabla^2(\eta^T g)(\bar{y})\bar{v}, \xi, v) \mid (v^*, \xi, v, \eta) \in (\tilde{K}(\bar{v}, \bar{v}^*))^\circ \times \tilde{K}(\bar{v}, \bar{v}^*)\}
\end{aligned}$$

and (41) follows from (46). $\qquad \square$

Unless $\bar{\Lambda}$ is a singleton, g can not be 2-nondegenerate in direction $\bar{v} = 0$. Hence, Proposition 5 might not be useful in case when $\bar{v} = 0$ and $\bar{\Lambda}$ contains more than one element. We now want to cover this situation. We denote for every $\bar{v} \in \bar{K}_\Gamma$, $\bar{v}^* \in N_{\bar{K}_\Gamma}(\bar{v})$ by $\Sigma(\bar{v}, \bar{v}^*)$ a nonempty subset of the extreme points of $\bar{\Lambda}(\bar{v})$ such that for every direction $u \in \mathcal{K}_{\bar{K}_\Gamma}(\bar{v}, \bar{v}^*)$ we have

$$\Sigma(\bar{v}, \bar{v}^*) \cap \bar{\Lambda}(\bar{v} + \beta u) \neq \emptyset \text{ for all } \beta > 0 \text{ sufficiently small.}$$

We can always choose $\Sigma(\bar{v}, \bar{v}^*)$ as the collection of all extreme points of $\bar{\Lambda}(\bar{v})$, because by [22, Lemma 3.5] we have $\bar{\Lambda}(v) \subseteq \bar{\Lambda}(\bar{v})$ for every v sufficiently close to $\bar{v}$ and the set $\bar{\Lambda}(v)$ is a face of $\bar{\Lambda}(\bar{v})$ whose extreme points are also extreme points of $\bar{\Lambda}(\bar{v})$. However, it might be advantageous to choose $\Sigma(\bar{v}, \bar{v}^*)$ smaller to get a sharper inclusion in the following proposition.

Proposition 6. *Let* $\bar{v}^* \in \bar{K}_\Gamma^\circ$. *Then*

$$\widehat{N}_{T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar{y},\bar{y}^*)}(0,\bar{v}^*) \tag{47}$$

$$\subseteq \bigcap_{\bar{v}\in\mathcal{L}(\bar{K}_\Gamma)} \Big\{(w^*,w)\,|\,\exists\bar{\lambda}\in\mathrm{conv}\,\Sigma(\bar{v},\bar{v}^*):\ \big(w^*+\nabla^2(\bar{\lambda}^T g)(\bar{y})w,w\big)\in\big(\mathcal{K}_{\bar{K}_\Gamma}(0,\bar{v}^*)\big)^\circ\times\mathcal{K}_{\bar{K}_\Gamma}(0,\bar{v}^*)\Big\}.$$

Moreover, for every $\bar{v}\in\mathcal{L}(\bar{K}_\Gamma)$ *such that* g *is 2-nondegenerate at* $(\bar{y},\bar{y}^*)$ *in direction* $\bar{v}$ *we have*

$$\widehat{N}_{T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar{y},\bar{y}^*)}(0,\bar{v}^*)\subseteq\bigcap_{\bar{\lambda}\in\bar{\Lambda}(\bar{v})}\widehat{N}_{T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar{y},\bar{y}^*)}(\bar{v},\nabla^2(\bar{\lambda}^T g)(\bar{y})\bar{v}+\bar{v}^*). \tag{48}$$

Proof. Let $(w^*,w)\in\widehat{N}_{T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar{y},\bar{y}^*)}(0,\bar{v}^*)$ and let $\bar{v}\in\mathcal{L}(\bar{K}_\Gamma)=\bar{K}_\Gamma\cap(-\bar{K}_\Gamma)$ be arbitrarily fixed. We first show that $w\in\mathcal{K}_{\bar{K}_\Gamma}(0,\bar{v}^*)$. For every $z^*\in T_{\bar{K}_\Gamma^\circ}(\bar{v}^*)$ we have $\bar{v}^*+\alpha z^*\in\bar{K}_\Gamma^\circ=N_{\bar{K}_\Gamma}(0)$ for $\alpha>0$ small enough. But by (36) with $v=0$, we have $(0,\bar{v}^*+\alpha z^*)\in T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar{y},\bar{y}^*)$ and thus $\langle w^*,0\rangle+\langle w,\bar{v}^*+\alpha z^*-\bar{v}^*\rangle\leq 0$ implying $w\in\big(T_{\bar{K}_\Gamma^\circ}(\bar{v}^*)\big)^\circ=N_{\bar{K}_\Gamma^\circ}(\bar{v}^*)=\mathcal{K}_{\bar{K}_\Gamma}(0,\bar{v}^*)$.

Next we show that there exists $\bar{\lambda}\in\mathrm{conv}\,\Sigma(\bar{v},\bar{v}^*)$ such that

$$w^*+\nabla^2(\bar{\lambda}^T g)(\bar{y})w\in\big(\mathcal{K}_{\bar{K}_\Gamma}(0,\bar{v}^*)\big)^\circ. \tag{49}$$

Note that $-\bar{v}\in\bar{K}_\Gamma$, $\bar{\Lambda}(\bar{v})=\bar{\Lambda}(-\bar{v})$, and since $\bar{K}_\Gamma$ is a convex polyhedral cone,

$$\bar{v}^*\in\bar{K}_\Gamma^\circ=T_{\bar{K}_\Gamma}(\bar{v})^\circ=N_{\bar{K}_\Gamma}(\bar{v})=N_{\bar{K}_\Gamma}(-\bar{v}).$$

Moreover by (35),

$$\lambda\in\bar{\Lambda}(\bar{v})\Leftrightarrow\bar{v}^T\nabla^2 g(\bar{y})\bar{v}\in N_{\bar{\Lambda}}(\lambda).$$

Therefore by (37), $(\pm\alpha\bar{v},\pm\alpha\nabla^2(\lambda^T g)(\bar{y})\bar{v}+\bar{v}^*)\in T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar{y},\bar{y}^*)$, $\forall\alpha>0$ sufficiently small, $\forall\lambda\in\bar{\Lambda}(\bar{v})$. By the definition of the regular normal cone we conclude

$$\limsup_{\alpha\searrow 0}\frac{\langle w^*,\pm\alpha\bar{v}\rangle+\langle w,\pm\alpha\nabla^2(\lambda^T g)(\bar{y})\bar{v}+\bar{v}^*-\bar{v}^*\rangle}{\alpha}=\pm\big(\langle w^*,\bar{v}\rangle+\langle w,\nabla^2(\lambda^T g)(\bar{y})\bar{v}\rangle\big)\leq 0$$

and therefore

$$\langle w^*,\bar{v}\rangle+\langle w,\nabla^2(\lambda^T g)(\bar{y})\bar{v}\rangle=0\quad\forall\lambda\in\bar{\Lambda}(\bar{v}). \tag{50}$$

Consider $u\in\mathcal{K}_{\bar{K}_\Gamma}(\bar{v},\bar{v}^*)$ and choose $\beta>0$ sufficiently small such $\Sigma(\bar{v},\bar{v}^*)\cap\bar{\Lambda}(\bar{v}+\beta u)\neq\emptyset$. Then $u\in T_{\bar{K}_\Gamma}(\bar{v})$ and $u^T\bar{v}^*=0$. It follows that $\bar{v}+\beta u\in\bar{K}_\Gamma$ for $\beta>0$ small and hence $\langle\bar{v}^*,\bar{v}+\beta u\rangle=0$ due to the fact that $\bar{v}^*\in N_{\bar{K}_\Gamma}(\bar{v})$. Hence $\bar{v}^*\in N_{\bar{K}_\Gamma}(\bar{v}+\beta u)$. Let $\lambda\in\Sigma(\bar{v},\bar{v}^*)\cap\bar{\Lambda}(\bar{v}+\beta u)$ and $\alpha>0$. Since $\lambda\in\bar{\Lambda}(\alpha(\bar{v}+\beta u))$ and $\bar{v}^*\in N_{\bar{K}_\Gamma}(\bar{v}+\beta u)$, by (36) we have

$$\big(\alpha(\bar{v}+\beta u),\alpha\nabla^2(\lambda^T g)(\bar{y})(\bar{v}+\beta u)+\bar{v}^*\big)\in T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar{y},\bar{y}^*).$$

It follows by definition for the regular normal cone $\widehat{N}_{T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar{y},\bar{y}^*)}(0,\bar{v}^*)$ that

$$\limsup_{\alpha\downarrow 0}\frac{\langle w^*,\alpha(\bar{v}+\beta u)\rangle+\langle w,\alpha\nabla^2(\lambda^T g)(\bar{y})(\bar{v}+\beta u)+\bar{v}^*-\bar{v}^*\rangle}{\alpha}$$
$$=\beta\big(\langle w^*,u\rangle+\langle w,\nabla^2(\lambda^T g)(\bar{y})u\rangle\big)\leq 0,$$

where the equality follows from (50). Hence

$$\langle w^*, u\rangle + \langle w, \nabla^2(\lambda^T g)(\bar{y})u\rangle \le 0 \quad \forall u \in \mathcal{K}_{\bar{K}_\Gamma}(\bar{v}, \bar{v}^*), \lambda \in \Sigma(\bar{v}, \bar{v}^*)$$

and by taking into account that $\operatorname{conv}\Sigma(\bar{v}, \bar{v}^*)$ is compact as the convex hull of a finite set, we obtain

$$\begin{aligned}
0 \;&\ge\; \max_{u \in \mathcal{K}_{\bar{K}_\Gamma}(\bar{v}, \bar{v}^*) \cap \mathcal{B}_{\mathbb{R}^m}} \min_{\lambda \in \operatorname{conv}\Sigma(\bar{v}, \bar{v}^*)} \langle w^*, u\rangle + \langle w, \nabla^2(\bar{\lambda}^T g)(\bar{y})u\rangle \\
&=\; \min_{\lambda \in \operatorname{conv}\Sigma(\bar{v}, \bar{v}^*)} \max_{u \in \mathcal{K}_{\bar{K}_\Gamma}(\bar{v}, \bar{v}^*) \cap \mathcal{B}_{\mathbb{R}^m}} \langle w^*, u\rangle + \langle w, \nabla^2(\bar{\lambda}^T g)(\bar{y})u\rangle.
\end{aligned}$$

Hence there is $\bar{\lambda} \in \operatorname{conv}\Sigma(\bar{v}, \bar{v}^*)$ such that $\max_{u \in \mathcal{K}_{\bar{K}_\Gamma}(\bar{v}, \bar{v}^*)} \langle w^*, u\rangle + \langle w, \nabla^2(\bar{\lambda}^T g)(\bar{y})u\rangle \le 0$. Since $\bar{v} \in \mathcal{L}(\bar{K}_\Gamma)$, we have $T_{\bar{K}_\Gamma}(\bar{v}) = \bar{K}_\Gamma$ and $\mathcal{K}_{\bar{K}_\Gamma}(\bar{v}, \bar{v}^*) = \mathcal{K}_{\bar{K}_\Gamma}(0, \bar{v}^*)$. Therefore (49) holds. Putting all together, (47) follows.

Let $\bar{v} \in \mathcal{L}(\bar{K}_\Gamma)$. We now show (48) under the assumption that g is 2-nondegenerate in direction $\bar{v}$ at $(\bar{y}, \bar{y}^*)$. Let $(w^*, w) \in \widehat{N}_{T_{\operatorname{gph}\widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)}(0, \bar{v}^*)$. Fixing $\bar{\lambda} \in \bar{\Lambda}(\bar{v})$, we wish to prove that $(w^*, w) \in \widehat{N}_{T_{\operatorname{gph}\widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)}(\bar{v}, \nabla^2(\bar{\lambda}^T g)(\bar{y})\bar{v} + \bar{v}^*) = (T_{T_{\operatorname{gph}\widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)}(\bar{v}, \nabla^2(\bar{\lambda}^T g)(\bar{y})\bar{v} + \bar{v}^*))^\circ$. So consider $(u, u^*) \in T_{T_{\operatorname{gph}\widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)}(\bar{v}, \nabla^2(\bar{\lambda}^T g)(\bar{y})\bar{v} + \bar{v}^*)$. By Proposition 5 there are elements μ, ζ^* such that

$$\begin{aligned}
u^* &= \nabla^2(\bar{\lambda}^T g)(\bar{y})u + \nabla^2(\mu^T g)(\bar{y})\bar{v} + \zeta^*, \\
(u, \mu, \zeta^*, 2\bar{v}^T \nabla^2 g(\bar{y})u) &\in \operatorname{gph} N_{\bar{K}(\bar{v}, \nabla^2(\bar{\lambda}^T g)(\bar{y})\bar{v} + \bar{v}^*)} = T_{\operatorname{gph} N_{\bar{K}_\Gamma \times \bar{\Lambda}}}(\bar{v}, \bar{\lambda}, \bar{v}^*, \bar{v}^T \nabla^2 g(\bar{y})\bar{v}).
\end{aligned}$$

By taking into account (44), there are sequences $t_k \downarrow 0$, $(u_k, \mu_k, \zeta_k^*) \to (u, \mu, \zeta^*)$ such that for each k, $\bar{v}^* + t_k \zeta_k^* \in N_{\bar{K}_\Gamma}(\bar{v} + t_k u_k)$, $(\bar{v} + t_k u_k)^T \nabla^2 g(\bar{y})(\bar{v} + t_k u_k) \in N_{\bar{\Lambda}}(\bar{\lambda} + t_k \mu_k)$. Note that by (35), $(\bar{v} + t_k u_k)^T \nabla^2 g(\bar{y})(\bar{v} + t_k u_k) \in N_{\bar{\Lambda}}(\bar{\lambda} + t_k \mu_k)$ if and only if $\bar{\lambda} + t_k \mu_k \in \bar{\Lambda}(\bar{v} + t_k u_k)$, and so

$$\bar{\lambda} + t_k \mu_k \in \bar{\Lambda}(\bar{v} + t_k u_k), \quad \bar{v}^* + t_k \zeta_k^* \in N_{\bar{K}_\Gamma}(\bar{v} + t_k u_k).$$

The set $N_{\bar{K}_\Gamma}(\bar{v} + t_k u_k)$ is a face of $\bar{K}_\Gamma^\circ$ and since the polyhedral convex cone $\bar{K}_\Gamma$ only has finitely many faces, after passing to a subsequence we can assume that $N_{\bar{K}_\Gamma}(\bar{v} + t_k u_k) = F \; \forall k$ for some face F of $\bar{K}_\Gamma^\circ$. Since F is closed, we obtain $\bar{v}^* \in F$ and thus $\bar{v}^* + \alpha t_k \zeta_k^* \in F = N_{\bar{K}_\Gamma}(\bar{v} + t_k u_k)$ $\forall k$, $\forall \alpha \in [0, 1]$. Hence, for every k and every $\alpha \in [0, 1]$ we have $\left(\alpha(\bar{v} + t_k u_k), \alpha\nabla^2((\bar{\lambda} + t_k \mu_k)^T g)(\bar{y})(\bar{v} + t_k u_k) + \bar{v}^* + \alpha t_k^* \zeta_k^*\right) \in T_{\operatorname{gph}\widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)$ by (36), implying

$$\begin{aligned}
0 \;&\ge\; \lim_{\alpha \downarrow 0} \frac{\langle w^*, \alpha(\bar{v} + t_k u_k)\rangle + \langle w, \alpha\nabla^2((\bar{\lambda} + t_k \mu_k)^T g)(\bar{y})(\bar{v} + t_k u_k) + \bar{v}^* + \alpha t_k \zeta_k^* - \bar{v}^*\rangle}{\alpha} \\
&=\; \langle w^*, (\bar{v} + t_k u_k)\rangle + \langle w, \nabla^2((\bar{\lambda} + t_k \mu_k)^T g)(\bar{y})(\bar{v} + t_k u_k) + t_k \zeta_k^*\rangle \\
&=\; t_k\left(\langle w^*, u_k\rangle + \langle w, \nabla^2((\bar{\lambda} + t_k \mu_k)^T g)(\bar{y})u_k + \nabla^2(\mu_k^T g)(\bar{y})\bar{v} + \zeta_k^*\rangle\right)
\end{aligned}$$

for all k. Dividing by t_k and passing to the limit we obtain

$$0 \ge \langle w^*, u\rangle + \langle w, \nabla^2(\bar{\lambda}^T g)(\bar{y})u + \nabla^2(\mu^T g)(\bar{y})\bar{v} + \zeta^*\rangle = \langle w^*, u\rangle + \langle w, u^*\rangle.$$

Thus $(w^*, w) \in (T_{T_{\operatorname{gph}\widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)}(\bar{v}, \nabla(\bar{\lambda}^T g)(\bar{y})\bar{v} + \bar{v}^*))^\circ$ and the inclusion (48) follows. $\qquad\square$

5.3 Regular normals to tangents of tangent cones

Throughout this subsection let $(\bar{v}, \bar{v}^*) \in T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)$ and $(\delta\bar{v}, \delta\bar{v}^*) \in T_{T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar{y},\bar{y}^*)}(\bar{v}, \bar{v}^*)$ be given and we assume that g is 2-nondegenerate in direction $\bar{v}$ at $(\bar{y}, \bar{y}^*)$. Further let $(\bar{\lambda}, \bar{z}^*) \in \bar{\Lambda}(\bar{v}) \times N_{\bar{K}_\Gamma}(\bar{v})$ denote the unique element fulfilling (39), i.e., $\bar{v}^* = \nabla^2(\bar{\lambda}^T g)(\bar{y})\bar{v} + \bar{z}^*$, and let according to (40) $(\bar{\mu}, \bar{\zeta}^*)$ denote some element with

$$\delta\bar{v}^* = \nabla^2(\bar{\lambda}^T g)(\bar{y})\delta\bar{v} + \nabla^2(\bar{\mu}^T g)(\bar{y})\bar{v} + \bar{\zeta}^*, \tag{51}$$

$$(\delta\bar{v}, \bar{\mu}, \bar{\zeta}^*, 2\bar{v}^T \nabla^2 g(\bar{y})\delta\bar{v}) \in \mathrm{gph}\, N_{\tilde{K}(\bar{v},\bar{v}^*)} = T_{\mathrm{gph}\, N_{\bar{K}_\Gamma \times \bar{\Lambda}}}(\bar{v}, \bar{\lambda}, \bar{z}^*, \bar{v}^T \nabla^2 g(\bar{y})\bar{v}), \tag{52}$$

where the equality in (52) follows from (5). Note that by definition,

$$\tilde{K}(\bar{v}, \bar{v}^*) := \mathcal{K}_{\bar{K}_\Gamma \times \bar{\Lambda}}(\bar{v}, \bar{\lambda}, \bar{z}^*, \bar{v}^T \nabla^2 g(\bar{y})\bar{v})$$

and hence it follows that $\bar{\mu} \in \mathcal{K}_{\bar{\Lambda}}(\bar{\lambda}, \bar{v}^T \nabla^2 g(\bar{y})\bar{v}) = T_{\bar{\Lambda}(v)}(\bar{\lambda})$ where the equality follows from (45), $\bar{\zeta}^* \in N_{\mathcal{K}_{\bar{K}_\Gamma}(\bar{v}, \bar{z}^*)}(\delta\bar{v}) \subseteq (\mathcal{K}_{\bar{K}_\Gamma}(\bar{v}, \bar{z}^*))^\circ \subseteq (N_{\bar{K}_\Gamma}(\bar{v}))^+$. By (51), $\mathcal{A}_{\bar{v}}(\bar{\mu}, \bar{\zeta}^*) = \delta\bar{v}^* - \nabla^2(\bar{\lambda}^T g)(\bar{y})\delta\bar{v}$ and from Lemma 3 we conclude that $(\bar{\mu}, \bar{\zeta}^*)$ are unique.

Proposition 7. *Under the assumption stated in the beginning of this subsection, we have*

$$T_{T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar{y},\bar{y}^*)}(\bar{v},\bar{v}^*)(\delta\bar{v}, \delta\bar{v}^*)$$
$$= \left\{ (u, u^*) \,\Big|\, \exists \delta\mu, \delta\zeta^* : \begin{array}{l} u^* = \nabla^2(\bar{\lambda}^T g)(\bar{y})u + \nabla^2(\delta\mu^T g)(\bar{y})\bar{v} + \delta\zeta^*) \\ (u, \delta\mu, \delta\zeta^*, 2\bar{v}^T \nabla^2 g(\bar{y})u) \in \mathrm{gph}\, N_{\tilde{K}(\bar{v},\bar{v}^*,\delta\bar{v},\delta\bar{v}^*)} \end{array} \right\} \tag{53}$$

and

$$\widehat{N}_{T_{T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar{y},\bar{y}^*)}(\bar{v},\bar{v}^*)}(\delta\bar{v}, \delta\bar{v}^*)$$
$$= \left\{ (w^*, w) \,\Big|\, \exists \eta : \begin{array}{l} (w^* + \nabla^2(\bar{\lambda}^T g)(\bar{y})w - 2\nabla^2(\eta^T g)(\bar{y})\bar{v}, \bar{v}^T \nabla^2 g(\bar{y})w, w, \eta) \\ \in (\tilde{K}(\bar{v}, \bar{v}^*, \delta\bar{v}, \delta\bar{v}^*))^\circ \times \tilde{K}(\bar{v}, \bar{v}^*, \delta\bar{v}, \delta\bar{v}^*) \end{array} \right\}, \tag{54}$$

where $\tilde{K}(\bar{v}, \bar{v}^*, \delta\bar{v}, \delta\bar{v}^*) := \mathcal{K}_{\tilde{K}(\bar{v},\bar{v}^*)}(\delta\bar{v}, \bar{\mu}, \bar{\zeta}^*, 2\bar{v}^T \nabla^2 g(\bar{y})\delta\bar{v})$.

Proof. We use similar arguments as in the proof of Proposition 5. Let $\mathcal{R}$ denote the set on the right hand side of (53) and consider $(u, u^*) \in T_{T_{T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar{y},\bar{y}^*)}(\bar{v},\bar{v}^*)}(\delta\bar{v}, \delta\bar{v}^*)$ together with sequences $t_k \downarrow 0$ and $(u_k, u_k^*) \to (u, u^*)$ with $(\delta\bar{v} + t_k u_k, \delta\bar{v}^* + t_k u_k^*) \in T_{T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar{y},\bar{y}^*)}(\bar{v}, \bar{v}^*)$. By Proposition 5 there are elements μ_k, ζ_k^* such that

$$\delta\bar{v}^* + t_k u_k^* = \nabla^2(\bar{\lambda}^T g)(\bar{y})(\delta\bar{v} + t_k u_k) + \nabla^2(\mu_k^T g)(\bar{y})\bar{v} + \zeta_k^*,$$
$$(\delta\bar{v} + t_k u_k, \mu_k, \zeta_k^*, 2\bar{v}^T \nabla^2 g(\bar{y})(\delta\bar{v} + t_k u_k)) \in \mathrm{gph}\, N_{\tilde{K}(\bar{v},\bar{v}^*)} = T_{\mathrm{gph}\, N_{\bar{K}_\Gamma \times \bar{\Lambda}}}(\bar{v}, \bar{\lambda}, \bar{z}^*, \bar{v}^T \nabla^2 g(\bar{y})\bar{v}),$$

where the equality in the second inclusion follows from (5). By taking into account (51) we obtain after rearranging

$$u_k^* - \nabla^2(\bar{\lambda}^T g)(\bar{y})u_k = \nabla^2\Big(\frac{(\mu_k - \bar{\mu})^T}{t_k} g\Big)(\bar{y})\bar{v} + \frac{\zeta_k^* - \bar{\zeta}^*}{t_k}.$$

Similarly as shown in the paragraph before Proposition 7, we can show that both $\bar{\mu}$ and μ_k belong to $\mathcal{K}_{\bar{\Lambda}}(\bar{\lambda}, \bar{v}^T \nabla^2 g(\bar{y})\bar{v}) = T_{\bar{\Lambda}(v)}(\bar{\lambda})$. Hence we obtain $\mu_k - \bar{\mu} \in (\bar{\Lambda}(\bar{v}))^+$. Further,

$\bar{\zeta}^* \in N_{\mathcal{K}_{\bar{K}_\Gamma}(\bar{v},\bar{z}^*)}(\delta\bar{v}) \subseteq \left(N_{\bar{K}_\Gamma}(\bar{v})\right)^+$ and $\zeta_k^* \in N_{\mathcal{K}_{\bar{K}_\Gamma}(\bar{v},\bar{z}^*)}(\delta\bar{v}+t_k u_k) \subseteq \left(N_{\bar{K}_\Gamma}(\bar{v})\right)^+$ implying $\zeta_k^* - \bar{\zeta}^* \in \left(N_{\bar{K}_\Gamma}(\bar{v})\right)^+$. Thus, by Lemma 3 the sequences $\frac{\mu_k-\bar{\mu}}{t_k}$ and $\frac{\zeta_k^*-\bar{\zeta}^*}{t_k}$ converge to some elements $\delta\mu$ and $\delta\zeta^*$, respectively, with $u^* - \nabla^2(\bar{\lambda}^T g)(\bar{y})u = \nabla^2(\delta\mu^T g)(\bar{y})\bar{v} + \delta\zeta^*$ and

$$(u, \delta\mu, \delta\zeta^*, 2\bar{v}^T\nabla^2 g(\bar{y})u) \in T_{\mathrm{gph}\,N_{\tilde{K}(\bar{v},\bar{v}^*)}}(\delta\bar{v}, \bar{\mu}, \bar{\zeta}^*, 2\bar{v}^T\nabla^2 g(\bar{y})\delta\bar{v}) = \mathrm{gph}\,N_{\tilde{K}(\bar{v},\bar{v}^*,\delta\bar{v},\delta\bar{v}^*)}$$

verifying $(u, u^*) \in \mathcal{R}$.

Now we prove the reverse inclusion of (53). Let $(u, u^*) \in \mathcal{R}$. Then there exist $\delta\mu$ and $\delta\zeta^*$ such that

$$u^* = \nabla^2(\bar{\lambda}^T g)(\bar{y})u + \nabla^2(\delta\mu^T g)(\bar{y})\bar{v} + \delta\zeta^*, \tag{55}$$

$$(u, \delta\mu, \delta\zeta^*, 2\bar{v}^T\nabla^2 g(\bar{y})u) \in \mathrm{gph}\,N_{\tilde{K}(\bar{v},\bar{v}^*,\delta\bar{v},\delta\bar{v}^*)}$$
$$= T_{\mathrm{gph}\,N_{\tilde{K}(\bar{v},\bar{v}^*)}}(\delta\bar{v}, \bar{\mu}, \bar{\zeta}^*, 2\bar{v}^T\nabla^2 g(\bar{y})\delta\bar{v}), \tag{56}$$

where the equality in the second inclusion follows from (5) and the notation $\tilde{K}(\bar{v},\bar{v}^*,\delta\bar{v},\delta\bar{v}^*) := \mathcal{K}_{\tilde{K}(\bar{v},\bar{v}^*)}(\delta\bar{v}, \bar{\mu}, \bar{\zeta}^*, 2\bar{v}^T\nabla^2 g(\bar{y})\delta\bar{v})$. Since $\tilde{K}(\bar{v},\bar{v}^*) := \mathcal{K}_{\bar{K}_\Gamma \times \bar{\Lambda}}(\bar{v}, \bar{\lambda}, \bar{z}^*, \bar{v}^T\nabla^2 g(\bar{y})\bar{v})$ is a convex polyhedral set, $\mathrm{gph}\,N_{\tilde{K}(\bar{v},\bar{v}^*)}$ is polyhedral, it follows by (56) that

$$(\delta\bar{v}+tu, \bar{\mu}+t\delta\mu, \bar{\zeta}^*+t\delta\zeta^*, 2\bar{v}^T\nabla^2 g(\bar{y})(\delta\bar{v}+tu)) \in \mathrm{gph}\,N_{\tilde{K}(\bar{v},\bar{v}^*)}$$

for all $t > 0$ sufficiently small. By (40) and taking into account (51) and (55), it follows that

$$(\delta\bar{v}+tu, \delta\bar{v}^*+tu^*) \in T_{T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar{y},\bar{y}^*)}(\bar{v},\bar{v}^*)$$

from which we can conclude $(u, u^*) \in T_{T_{T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar{y},\bar{y}^*)}(\bar{v},\bar{v}^*)}(\delta\bar{v}, \delta\bar{v}^*)$. Thus (53) is proven.

In order to show (54) note that by (53), $\widehat{N}_{T_{T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar{y},\bar{y}^*)}(\bar{v},\bar{v}^*)}(\delta\bar{v}, \delta\bar{v}^*)$ is the collection of all (w^*, w) fulfilling

$$0 \geq \langle w^*, u \rangle + \langle w, \nabla^2(\bar{\lambda}^T g)(\bar{y})u + \nabla^2(\delta\mu^T g)(\bar{y})\bar{v} + \delta\zeta^* \rangle$$
$$= \langle w^* + \nabla^2(\bar{\lambda}^T g)(\bar{y})w, u \rangle + w^T\nabla^2(\delta\mu^T g)(\bar{y})\bar{v} + w^T\delta\zeta^*$$

for all

$$(u, \delta\mu, \delta\zeta^*) \in \Theta := \{(u, \delta\mu, \delta\zeta^*) \,|\, (u, \delta\mu, \delta\zeta^*, 2\bar{v}^T\nabla^2 g(\bar{y})u) \in \mathrm{gph}\,N_{\tilde{K}(\bar{v},\bar{v}^*,\delta\bar{v},\delta\bar{v}^*)}\},$$

which is the same as $(w^* + \nabla^2(\bar{\lambda}^T g)(\bar{y})w, w^T\nabla^2 g(\bar{y})\bar{v}, w) \in \Theta^\circ$. In order to compute Θ° we use Lemma 2 with the linear mappings $\tilde{P}_1(u, \delta\mu, \delta\zeta^*) := (u, \delta\mu)$, $\tilde{P}_2(u, \delta\mu, \delta\zeta^*) := (\delta\zeta^*, 2\bar{v}^T\nabla^2 g(\bar{y})u)$ and $C = \tilde{K}(\bar{v},\bar{v}^*)$ and $\bar{z} = (\delta\bar{v}, \bar{\mu}, \bar{\zeta}^*)$. Indeed, for the subspaces L_1, L_2 defined in the proof of Proposition 5 we have shown $\ker\nabla\tilde{P}(\bar{z}) \cap (L_1 \times L_2) = \{0\}$, where we have to take into account that $\nabla\tilde{P}$ coincides with the derivative of the mapping $\tilde{P}$ used in the proof of Proposition 5 at $(\bar{v}, \bar{\lambda}, \bar{z}^*)$. Further, from (13) together with (14) and the definition of $\tilde{K}(\bar{v},\bar{v}^*)$ we obtain

$$L_1^\perp \times L_2^\perp \subseteq \mathcal{L}\big(T_{\mathrm{gph}\,N_{\bar{K}_\Gamma \times \bar{\Lambda}}}(\bar{v}, \bar{\lambda}, \bar{z}^*, \bar{v}^T\nabla^2 g(\bar{y})\bar{v})\big)$$
$$\subseteq \mathcal{L}\big(T_{\mathrm{gph}\,N_{\mathcal{K}_{\bar{K}_\Gamma \times \bar{\Lambda}}(\bar{v},\bar{\lambda},\bar{z}^*,\bar{v}^T\nabla^2 g(\bar{y})\bar{v})}}(\delta\bar{v}, \bar{\mu}, \bar{\zeta}^*, 2\bar{v}^T\nabla^2 g(\bar{y})\delta\bar{v})\big)$$
$$= \mathcal{L}\big(T_{\mathrm{gph}\,N_{\tilde{K}(\bar{v},\bar{v}^*)}}(\delta\bar{v}, \bar{\mu}, \bar{\zeta}^*, 2\bar{v}^T\nabla^2 g(\bar{y})\delta\bar{v})\big).$$

22

Applying (13) once more we obtain

$$L_1 \supseteq N_{\tilde{K}(\bar{v},\bar{v}^*)}(\delta\bar{v},\bar{\mu}), \quad L_2 \supseteq \mathcal{K}_{\tilde{K}(\bar{v},\bar{v}^*)}(\delta\bar{v},\bar{\mu},\bar{\zeta}^*,2\bar{v}^T\nabla^2 g(\bar{y})\delta\bar{v}).$$

Hence we can apply Lemma 2 to obtain

$$\widehat{N}_{\{z\,|\,\tilde{P}(z)\in\tilde{D}\}}(\bar{z})$$
$$= \{w\,|\,\nabla\tilde{P}(\bar{z})w \in T_{\tilde{D}}(\tilde{P}(\bar{z}))\}^\circ = \{w\,|\,\nabla\tilde{P}(\bar{z})w \in \mathrm{gph}\,N_{\tilde{K}(\bar{v},\bar{v}^*,\delta\bar{v},\delta\bar{v}^*)}\}^\circ = \Theta^\circ$$
$$= \{(v^*+2\nabla^2(\eta^T g)(\bar{y})\bar{v},\xi,v)\,|\,(v^*,\xi,v,\eta)\in\tilde{K}(\bar{v},\bar{v}^*,\delta\bar{v},\delta\bar{v}^*)^\circ\times\tilde{K}(\bar{v},\bar{v}^*,\delta\bar{v},\delta\bar{v}^*)\},$$

where the second equality follows from (5), and hence (54) follows. $\qquad\square$

6 New optimality condition for (MPEC)

To establish the main optimality condition in Theorem 5, we first apply Theorem 2 to problem (1) to obtain the following lemma.

Lemma 4. *Assume that $(\bar{x},\bar{y})$ is a local minimizer for problem* (1) *fulfilling Assumption 1. Further assume that g is 2-nondegenerate in every nonzero critical direction $0 \neq v \in \bar{K}_\Gamma$ at $(\bar{y},\bar{y}^*)$ with $\bar{y}^* := -\phi(\bar{x},\bar{y})$. Then there are a direction $(\delta x,\delta y)$ and elements*

$$(\delta\bar{v},\delta\bar{v}^*) \in T_{T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar{y},\bar{y}^*)}\big(\delta y,-\nabla\phi(\bar{x},\bar{y})(\delta x,\delta y)\big), \quad \delta a \in T_{T_{\mathbb{R}_-^m}(G(\bar{x},\bar{y}))}\big(\nabla G(\bar{x},\bar{y})(\delta x,\delta y)\big),$$

together with multipliers

$$(w^*,w) \in \widehat{N}_{T_{T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar{y},\bar{y}^*)}(\delta y,-\nabla\phi(\bar{x},\bar{y})(\delta x,\delta y))}(\delta\bar{v},\delta\bar{v}^*), \quad \sigma \in N_{T_{T_{\mathbb{R}_-^m}(G(\bar{x},\bar{y}))}(\nabla G(\bar{x},\bar{y})(\delta x,\delta y))}(\delta a)$$

such that

$$\nabla F(\bar{x},\bar{y})^T(\delta x,\delta y) = 0, \tag{57a}$$
$$\nabla_x F(\bar{x},\bar{y}) - \nabla_x\phi(\bar{x},\bar{y})^T w + \nabla_x G(\bar{x},\bar{y})^T\sigma = 0, \tag{57b}$$
$$\nabla_y F(\bar{x},\bar{y}) + w^* - \nabla_y\phi(\bar{x},\bar{y})^T w + \nabla_y G(\bar{x},\bar{y})^T\sigma = 0, \tag{57c}$$
$$\delta y = 0 \ \Rightarrow\ \delta x = 0, \tag{57d}$$
$$\delta y \neq 0 \ \Rightarrow\ T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar{y},\bar{y}^*) \text{ is not locally polyhedral near } \big(\delta y,-\nabla\phi(\bar{x},\bar{y})(\delta x,\delta y)\big). \tag{57e}$$

Further, if $\mathcal{L}(\bar{K}_\Gamma) \neq \{0\}$ then $(\delta y,\delta\bar{v}) \neq (0,0)$. Otherwise, if $\mathcal{L}(\bar{K}_\Gamma) = \{0\}$ and $(\delta y,\delta\bar{v}) = (0,0)$ then there is some $\bar{\lambda} \in \Sigma(0,\delta\bar{v}^*)$ such that

$$(w^* + \nabla^2(\bar{\lambda}^T g)(\bar{y})w,w) \in \big(\mathcal{K}_{\bar{K}_\Gamma}(0,\delta\bar{v}^*)\big)^\circ \times \mathcal{K}_{\bar{K}_\Gamma}(0,\delta\bar{v}^*). \tag{58}$$

Proof. Let $z := (x,y), \bar{z} = (\bar{x},\bar{y}), P(x,y) := \begin{pmatrix} (y,-\phi(x,y)) \\ G(x,y) \end{pmatrix}, D := \mathrm{gph}\,\widehat{N}_\Gamma \times \mathbb{R}_-^p$. Assumption 1 ensures that Theorem 2 is applicable and so one of Theorem 2(i) and Theorem 2(ii) holds.

If Theorem 2(i) is fulfilled, then there exists a direction $\omega = (\delta\bar{v},\delta\bar{v}^*,\delta a) \in T_D(P(\bar{z})) = T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar{y},\bar{y}^*) \times T_{\mathbb{R}_-^m}(G(\bar{x},\bar{y}))$ and a multiplier

$$\omega^* = (w^*,w,\sigma) \in \widehat{N}_{T_D(P(\bar{z}))}(w) = \widehat{N}_{T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar{y},\bar{y}^*)}(\delta\bar{v},\delta\bar{v}^*) \times N_{T_{\mathbb{R}_-^m}(G(\bar{x},\bar{y}))}(\delta a)$$

such that $0 = \nabla F(\bar{x}, \bar{y}) + \nabla P(\bar{x}, \bar{y})^T \omega^*$. By virtue of (3), we see that the conditions (57a)-(57c) are fulfilled with $\delta x = 0$, $\delta y = 0$. Otherwise Theorem 2(ii) is fulfilled, i.e., there is a direction $\bar{u} = (\delta x, \delta y)$ with

$$\nabla P(\bar{z})\bar{u} = \big(\delta y, -\nabla\phi(\bar{x}, \bar{y})(\delta x, \delta y), \nabla G(\bar{x}, \bar{y})(\delta x, \delta y)\big) \tag{59}$$
$$\in T_D\big(P(\bar{z})\big) = T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar{y}, \bar{y}^*) \times T_{\mathbb{R}^p_-}\big(G(\bar{x}, \bar{y})\big)$$

fulfilling (26), $\nabla F(\bar{x}, \bar{y})(\delta x, \delta y) = 0$ which is (57a) and (28) such that $T_D(P(\bar{z}))$ is not locally polyhedral near $\nabla P(\bar{z})\bar{u}$, which is equivalent to the requirement that $T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)$ is not locally polyhedral near $(\delta y, -\nabla\phi(\bar{x}, \bar{y})(\delta x, \delta y))$ due to the polyhedrality of $T_{\mathbb{R}^p_-}\big(G(\bar{x}, \bar{y})\big)$. From (37) we see that $T_{\mathrm{gph}\,\widehat{N}_\Gamma \times \mathbb{R}^p_-}((\bar{y}, \bar{y}^*), G(\bar{x}, \bar{y}))$ is the graph of a set-valued mapping $M = M_c + M_p$, where $M_p(v) := N_{\bar{K}_\Gamma}(v) \times T_{\mathbb{R}^p_-}(G(\bar{x}, \bar{y}))$ is polyhedral and

$$M_c(v) := \{\nabla^2(\lambda^T g)(y)v \mid \lambda \in \bar{\Lambda}(v) \cap \kappa\|\bar{y}^*\|\mathcal{B}_{\mathbb{R}^q}\} \times \{0\}$$

fulfills (29). Further, the graphs of M_p and M_c are closed cones and from (30) we conclude that $\delta y \neq 0$ by taking account of (59). Next we utilize (25), which says $\nabla F(\bar{z})^T \bar{u} = 0$. Since by the assumed GGCQ we have

$$\nabla F(\bar{z})^T u \geq 0 \ \forall u \text{ s.t. } \nabla P(\bar{z})u \in T_D(P(\bar{z})),$$

$\bar{u}$ is a global minimizer of problem

$$\min \nabla F(\bar{z})^T u \quad \text{subject to } \nabla P(\bar{z})u \in T_D(P(\bar{z})).$$

Similarly as in Remark 1(ii), we can apply Theorem 2 once more to the above problem, because metric subregularity of $u \rightrightarrows \nabla P(\bar{z})u - T_D(P(\bar{z}))$ at $(0,0)$ implies metric subregularity at $(\bar{u}, 0)$ by [12, Lemma 3] and therefore also GGCQ for the system $\nabla P(\bar{z})u \in T_D(P(\bar{z}))$ at $\bar{u}$. This means that the set $\Omega = \{z \mid P(z) \in D\}$ is replaced by the set $\{u \mid \nabla P(\bar{z})u \in T_D\big(P(\bar{z})\big)\}$, whose linearized tangent cone at $\bar{u}$ is $\{u \mid \nabla P(\bar{z})u \in T_{T_D(P(\bar{z}))}(\nabla P(\bar{z})\bar{u})\}$. Since

$$T_{T_D(P(\bar{z}))}(\nabla P(\bar{z})\bar{u}) = T_{T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)}\big(\delta y, -\nabla\phi(\bar{x}, \bar{y})(\delta x, \delta y)\big) \times T_{T_{\mathbb{R}^m_-}(G(\bar{x}, \bar{y}))}\big(\nabla G(\bar{x}, \bar{y})(\delta x, \delta y)\big)$$

and g is 2-nondegenerate in direction $\delta y \neq 0$, by (40) the set $T_{T_D(P(\bar{z}))}(\nabla P(\bar{z})\bar{u})$ is polyhedral and therefore only the first alternative of Theorem 2 is possible. Hence there is a direction $\omega = (\delta\bar{v}, \delta\bar{v}^*, \delta a) \in T_{T_D(P(\bar{z}))}(\nabla P(\bar{z})\bar{u})$ and a multiplier

$$\omega^* = (w^*, w, \sigma) \in \widehat{N}_{T_{T_D(P(\bar{z}))}(\nabla P(\bar{z})\bar{u})}(\omega)$$
$$= \widehat{N}_{T_{T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)}(\delta y, -\nabla\phi(\bar{x}, \bar{y})(\delta x, \delta y))}(\delta\bar{v}, \delta\bar{v}^*) \times N_{T_{T_{\mathbb{R}^m_-}(G(\bar{x}, \bar{y}))}(\nabla G(\bar{x}, \bar{y})(\delta x, \delta y))}(\delta a)$$

with $0 \in \nabla F(\bar{z}) + \nabla P(\bar{z})^T \omega^*$ which results in (57b) and (57c).

Now consider the case when $\delta y = 0$. In this case we must have $\delta x = 0$. Then by Proposition 1,

$$(w^*, w) \in \widehat{N}_{T_{T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)}(0,0)}(\delta\bar{v}, \delta\bar{v}^*) = \widehat{N}_{T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)}(\delta\bar{v}, \delta\bar{v}^*).$$

24

If $\delta\bar{v} = 0$ and $\mathcal{L}(\bar{K}_\Gamma) \neq \{0\}$, then by (48) we also have

$$(w^*, w) \in \widehat{N}_{T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar{y},\bar{y}^*)}\big(\bar{v}, \nabla^2(\bar{\lambda}^T g)(\bar{y})\bar{v} + \delta\bar{v}^*\big)$$

for every $0 \neq \bar{v} \in \mathcal{L}(\bar{K}_\Gamma) \neq \{0\}$ and every $\bar{\lambda} \in \bar{\Lambda}(\bar{v})$ and therefore we can assume $\delta\bar{v} \neq 0$. Otherwise, if $\delta\bar{v} = 0$ and $\mathcal{L}(\bar{K}_\Gamma) = \{0\}$ then (58) follows from (47) by taking $\bar{v} = 0$. $\qquad\square$

Now we are ready to state and prove our main optimality condition for problem (1). The main task is to interpret the formulas for the tangent cones and the regular normal cones in Propositions 5-7 appearing in Lemma 4 in terms of problem data.

Theorem 5. *Assume that $(\bar{x}, \bar{y})$ is a local minimizer for problem (1) fulfilling Assumption 1. Further assume that g is 2-nondegenerate in every nonzero critical direction $0 \neq v \in \bar{K}_\Gamma$ at $(\bar{y}, \bar{y}^*)$, where $\bar{y}^* := -\phi(\bar{x}, \bar{y})$. Then there are $\bar{v} \in \bar{K}_\Gamma$, $\bar{z}^* \in N_{\bar{K}_\Gamma}(\bar{v})$, $\bar{\lambda} \in \bar{\Lambda}(\bar{v})$, two faces $\mathcal{F}_1^v, \mathcal{F}_2^v$ of $\mathcal{K}_{\bar{K}_\Gamma}(\bar{v}, \bar{z}^*)$ with $\mathcal{F}_2^v \subseteq \mathcal{F}_1^v$, $\delta v \in \mathrm{ri}\,\mathcal{F}_2^v$, two faces $\mathcal{F}_1^\lambda, \mathcal{F}_2^\lambda$ of $T_{\bar{\Lambda}(\bar{v})}(\bar{\lambda})$ with $\mathcal{F}_2^\lambda \subseteq \mathcal{F}_1^\lambda$, $w \in \mathcal{F}_1^v - \mathcal{F}_2^v$, $\eta \in \mathcal{F}_1^\lambda - \mathcal{F}_2^\lambda$ and $\sigma \in \mathbb{R}_+^p$ such that*

$$\nabla_x F(\bar{x}, \bar{y}) - \nabla_x \phi(\bar{x}, \bar{y})^T w + \nabla_x G(\bar{x}, \bar{y})^T \sigma = 0, \tag{60a}$$

$$\nabla_y F(\bar{x}, \bar{y}) - \nabla_y \phi(\bar{x}, \bar{y})^T w + \nabla_y G(\bar{x}, \bar{y})^T \sigma \tag{60b}$$
$$- \nabla^2(\bar{\lambda}^T g)(\bar{y})w + 2\nabla^2(\eta^T g)(\bar{y})\bar{v} \in -(\mathcal{F}_1^v - \mathcal{F}_2^v)^\circ,$$

$$\bar{v}^T \nabla^2 g(\bar{y})w \in (\mathcal{F}_1^\lambda - \mathcal{F}_2^\lambda)^\circ, \tag{60c}$$

$$\bar{v}^T \nabla^2 g(\bar{y})\delta v \in T_{\bar{\Lambda}(\bar{v})}(\bar{\lambda})^\circ, \mathcal{F}_1^\lambda = T_{\bar{\Lambda}(\bar{v})}(\bar{\lambda}) \cap [\bar{v}^T \nabla^2 g(\bar{y})\delta v]^\perp, \tag{60d}$$

$$\sigma_i G_i(\bar{x}, \bar{y}) = 0, \ i = 1, \ldots, p. \tag{60e}$$

Furthermore, if $\mathcal{F}_1^v - \mathcal{F}_2^v = \mathcal{K}_{\bar{K}_\Gamma}(\bar{v}, \bar{z}^*)$ and $\mathcal{F}_1^\lambda - \mathcal{F}_2^\lambda = T_{\bar{\Lambda}(\bar{v})}(\bar{\lambda})$ then one of the following two cases must occur: case (a) $\bar{v} \neq 0$; case (b) $\bar{v} = 0$ and $\mathcal{L}(\bar{K}_\Gamma) = \{0\}$ and $\bar{\lambda} \in \Sigma(0, \bar{z}^*)$. Otherwise, if $\mathcal{F}_1^v - \mathcal{F}_2^v \neq \mathcal{K}_{\bar{K}_\Gamma}(\bar{v}, \bar{z}^*)$ or $\mathcal{F}_1^\lambda - \mathcal{F}_2^\lambda \neq T_{\bar{\Lambda}(\bar{v})}(\bar{\lambda})$ then $\bar{v} \neq 0$ and there is some $\delta x \in \mathbb{R}^n$ such that

$$\nabla F(\bar{x}, \bar{y})^T(\delta x, \bar{v}) = 0, \tag{61a}$$

$$\nabla \phi(\bar{x}, \bar{y})(\delta x, \bar{v}) + \nabla^2(\bar{\lambda}^T g)(\bar{y})\bar{v} + \bar{z}^* = 0, \tag{61b}$$

$$\nabla G_i(\bar{x}, \bar{y})^T(\delta x, \bar{v}) \leq 0, \ \sigma_i \nabla G_i(\bar{x}, \bar{y})^T(\delta x, \bar{v}) = 0, \ \forall i : G_i(\bar{x}, \bar{y}) = 0, \tag{61c}$$

and $T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)$ is not locally polyhedral near $(\bar{v}, -\nabla\phi(\bar{x}, \bar{y})(\delta x, \bar{v}))$.

Proof. Consider $\delta x, \delta y, \delta\bar{v}, \delta\bar{v}^*, w^*, w, \delta a$ and σ as in Lemma 4. Then (60a) holds and (60e) follows from the observation that

$$\sigma \in N_{T_{\mathbb{R}_-^m}(G(\bar{x},\bar{y}))}(\nabla G(\bar{x},\bar{y})(\delta x, \delta y))(\delta a) \subseteq N_{T_{\mathbb{R}_-^m}(G(\bar{x},\bar{y}))}\left(\nabla G(\bar{x},\bar{y})(\delta x, \delta y)\right) \subseteq N_{\mathbb{R}_-^m}(G(\bar{x},\bar{y})).$$

Case I: $\delta y = 0$. Then we also have $\delta x = 0$ by (57d) and thus

$$(\delta\bar{v}, \delta\bar{v}^*) \in T_{T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar{y},\bar{y}^*)}(0,0) = T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar{y}, \bar{y}^*), \ (w^*, w) \in \widehat{N}_{T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar{y},\bar{y}^*)}(\delta\bar{v}, \delta\bar{v}^*).$$

Subcase Ia: $\delta\bar{v} \neq 0$. Set $\bar{v} = \delta\bar{v}$ and by Lemma 3 there are unique elements $\bar{\lambda} \in \bar{\Lambda}(\bar{v})$ and $\bar{z}^* \in N_{\bar{K}_\Gamma}(\bar{v})$ such that $\delta\bar{v}^* = \nabla^2(\bar{\lambda}^T g)(\bar{y})\bar{v} + \bar{z}^*$. Since

$$\tilde{K}(\bar{v}, \delta\bar{v}^*) = \mathcal{K}_{\bar{K}_\Gamma \times \bar{\Lambda}}(\bar{v}, \bar{\lambda}, \bar{z}^*, \bar{v}^T \nabla^2 g(\bar{y})\bar{v}) = \mathcal{K}_{\bar{K}_\Gamma}(\bar{v}, \bar{z}^*) \times \mathcal{K}_{\bar{\Lambda}}(\bar{\lambda}, \bar{v}^T \nabla^2 g(\bar{y})\bar{v})$$

and $\mathcal{K}_{\bar{\Lambda}}(\bar{\lambda}, \bar{v}^T \nabla^2 g(\bar{y})\bar{v}) = T_{\bar{\Lambda}(\bar{v})}(\bar{\lambda})$, by (41) there is some $\eta \in T_{\bar{\Lambda}(\bar{v})}(\bar{\lambda})$ such that

$$w^* + \nabla^2(\bar{\lambda}^T g)(\bar{y})w - 2\nabla^2(\eta^T g)(\bar{y})\bar{v} \in \mathcal{K}_{\bar{K}_\Gamma}(\bar{v}, \bar{z}^*)^\circ, \tag{62a}$$

$$w \in \mathcal{K}_{\bar{K}_\Gamma}(\bar{v}, \bar{z}^*), \tag{62b}$$

$$\bar{v}^T \nabla^2 g(\bar{y})w \in T_{\bar{\Lambda}(\bar{v})}(\bar{\lambda})^\circ. \tag{62c}$$

Set $\delta v = 0$, $\mathcal{F}_2^v = \{0\}$, $\mathcal{F}_1^v = \mathcal{K}_{\bar{K}_\Gamma}(\bar{v}, \bar{z}^*)$, $\mathcal{F}_2^\lambda = \{0\}$, $\mathcal{F}_1^\lambda = T_{\bar{\Lambda}(\bar{v})}(\bar{\lambda})$ implying $w \in \mathcal{F}_1^v - \mathcal{F}_2^v = \mathcal{K}_{\bar{K}_\Gamma}(\bar{v}, \bar{z}^*)$ and $\eta \in \mathcal{F}_1^\lambda - \mathcal{F}_2^\lambda = T_{\bar{\Lambda}(\bar{v})}(\bar{\lambda})$. Then (60c) follows from (62c), (60d) is fulfilled and (60b) follows from (57c) and (62a).

Subcase Ib: $\delta\bar{v} = 0$. By Lemma 4 the case $\delta\bar{v} = 0$ is only possible when $\mathcal{L}(\bar{K}_\Gamma) = \{0\}$ and in this case there is some $\bar{\lambda} \in \Sigma(0, \delta\bar{v}^*)$ such that (58) holds. It follows that the conditions of the theorem are fulfilled with $\bar{v} = 0$, $\bar{z}^* = \delta\bar{v}^*$, $\eta = 0$, $\delta v = 0$, $\mathcal{F}_2^v = \{0\}$, $\mathcal{F}_1^v = \mathcal{K}_{\bar{K}_\Gamma}(\bar{v}, \bar{z}^*)$, $\mathcal{F}_2^\lambda = \{0\}$, $\mathcal{F}_1^\lambda = T_{\bar{\Lambda}(\bar{v})}(\bar{\lambda})$.

Case II: $\delta y \neq 0$. In this case set $\bar{v} := \delta y$, $\delta v := \delta\bar{v}$. Then $(\bar{v}, -\nabla\phi(\bar{x}, \bar{y})(\delta x, \bar{v})) \in T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)$ and by Lemma 3 there are unique $\bar{\lambda} \in \bar{\Lambda}(\bar{v})$ and $\bar{z}^* \in N_{\bar{K}_\Gamma}(\bar{v})$ such that

$$-\nabla\phi(\bar{x}, \bar{y})(\delta x, \bar{v}) = \nabla^2(\bar{\lambda}^T g)(\bar{y})\bar{v} + \bar{z}^*.$$

In view of (51) and (52), there are unique $\bar{\mu} \in T_{\bar{\Lambda}(\bar{v})}(\bar{\lambda})$ and $\bar{\zeta}^* \in N_{\mathcal{K}_{\bar{K}_\Gamma}(\bar{v}, \bar{z}^*)}(\delta v)$ such that

$$\delta\bar{v}^* = \nabla^2(\bar{\lambda}^T g)(\bar{y})\delta v + \nabla^2(\bar{\mu}^T g)(\bar{y})\bar{v} + \bar{\zeta}^*,$$

$$(\delta v, \bar{\mu}, \bar{\zeta}^*, 2\bar{v}^T \nabla^2 g(\bar{y})\delta v) \in \mathrm{gph}\, N_{\tilde{K}(\bar{v}, -\nabla\phi(\bar{x}, \bar{y})(\delta x, \bar{v}))}.$$

Further, by (54) there is some η such that

$$\left(w^* + \nabla^2(\bar{\lambda}^T g)(\bar{y})w - 2\nabla^2(\eta^T g)(\bar{y})\bar{v}, \bar{v}^T \nabla^2 g(\bar{y})w, w, \eta\right)$$
$$\in \tilde{K}(\delta y, -\nabla\phi(\bar{x}, \bar{y})(\delta x, \bar{v}), \delta v, \delta\bar{v}^*)^\circ \times \tilde{K}(\bar{v}, -\nabla\phi(\bar{x}, \bar{y})(\delta x, \bar{v}), \delta v, \delta\bar{v}^*).$$

By taking into account

$$\tilde{K}(\bar{v}, -\nabla\phi(\bar{x}, \bar{y})(\delta x, \bar{v}), \delta v, \delta\bar{v}^*) = \mathcal{K}_{\mathcal{K}_{\bar{K}_\Gamma}(\bar{v}, \bar{z}^*) \times T_{\bar{\Lambda}(\bar{v})}(\bar{\lambda})}(\delta v, \bar{\mu}, \bar{\zeta}^*, 2\bar{v}^T \nabla^2 g(\bar{y})\delta v)$$
$$= \mathcal{K}_{\mathcal{K}_{\bar{K}_\Gamma}(\bar{v}, \bar{z}^*)}(\delta v, \bar{\zeta}^*) \times \mathcal{K}_{T_{\bar{\Lambda}(\bar{v})}(\bar{\lambda})}(\bar{\mu}, 2\bar{v}^T \nabla^2 g(\bar{y})\delta v),$$

we obtain $\delta v \in \mathcal{K}_{\bar{K}_\Gamma}(\bar{v}, \bar{z}^*)$, $\bar{\zeta}^* \in N_{\mathcal{K}_{\bar{K}_\Gamma}(\bar{v}, \bar{z}^*)}(\delta v)$,

$$(w^* + \nabla^2(\bar{\lambda}^T g)(\bar{y})w - 2\nabla^2(\eta^T g)(\bar{y})\bar{v}, w) \in \mathcal{K}_{\mathcal{K}_{\bar{K}_\Gamma}(\bar{v}, \bar{z}^*)}(\delta v, \bar{\zeta}^*)^\circ \times \mathcal{K}_{\mathcal{K}_{\bar{K}_\Gamma}(\bar{v}, \bar{z}^*)}(\delta v, \bar{\zeta}^*),$$
$$(\bar{v}^T \nabla^2 g(\bar{y})w, \eta) \in \mathcal{K}_{T_{\bar{\Lambda}(\bar{v})}(\bar{\lambda})}(\bar{\mu}, 2\bar{v}\nabla^2 g(\bar{y})\delta v)^\circ \times \mathcal{K}_{T_{\bar{\Lambda}(\bar{v})}(\bar{\lambda})}(\bar{\mu}, 2\bar{v}^T \nabla^2 g(\bar{y})\delta v).$$

By defining $\mathcal{F}_1^v := \mathcal{K}_{\bar{K}_\Gamma}(\bar{v}, \bar{z}^*) \cap [\bar{\zeta}^*]^\perp$, $\mathcal{F}_1^\lambda := T_{\bar{\Lambda}(\bar{v})}(\bar{\lambda}) \cap [\bar{v}^T \nabla^2 g(\bar{y})\delta v]^\perp$ and choosing $\mathcal{F}_2^v \subset \mathcal{F}_1^v$ and $\mathcal{F}_2^\lambda \subset \mathcal{F}_1^\lambda$ as those faces fulfilling $\delta v \in \mathrm{ri}\,\mathcal{F}_2^v$, $\bar{\mu} \in \mathrm{ri}\,\mathcal{F}_2^\lambda$ we obtain $\mathcal{K}_{\mathcal{K}_{\bar{K}_\Gamma}(\bar{v}, \bar{z}^*)}(\delta v, \bar{\zeta}^*) = \mathcal{F}_1^v - \mathcal{F}_2^v$ and $\mathcal{K}_{T_{\bar{\Lambda}(\bar{v})}(\bar{\lambda})}(\bar{\mu}, 2\bar{v}^T \nabla^2 g(\bar{y})\delta v) = \mathcal{F}_1^\lambda - \mathcal{F}_2^\lambda$. Hence (60) follows. Since we have $\bar{v} \neq 0$, the claimed properties follow when $\mathcal{F}_1^v - \mathcal{F}_2^v = \mathcal{K}_{\bar{K}_\Gamma}(\bar{v}, \bar{z}^*)$ and $\mathcal{F}_1^\lambda - \mathcal{F}_2^\lambda = T_{\bar{\Lambda}(\bar{v})}(\bar{\lambda})$. Otherwise, if $\mathcal{F}_1^v - \mathcal{F}_2^v \neq \mathcal{K}_{\bar{K}_\Gamma}(\bar{v}, \bar{z}^*)$ and $\mathcal{F}_1^\lambda - \mathcal{F}_2^\lambda \neq T_{\bar{\Lambda}(\bar{v})}(\bar{\lambda})$ the claimed properties follow as well. $\quad\square$

In the following remark we summarize some comments on the optimality conditions of Theorem 5.

Remark 2. *1. If $\bar{\Lambda}(\bar{v}) = \{\bar{\lambda}\}$ is a singleton then we have $\bar{\Lambda}(v) = \bar{\Lambda}(\bar{v})$ for all $v \in \bar{K}_\Gamma$ sufficiently close to $\bar{v}$. Indeed, by [16, Lemma 3] we have $\bar{\Lambda}(v) \subseteq \bar{\Lambda}(\bar{v})$ for every $v \in \bar{K}_\Gamma$ sufficiently close to $\bar{v}$ and $\bar{\Lambda}(v) \neq \emptyset$ for any $v \in \bar{K}_\Gamma$ by [14, Proposition 4.3(iii)]. As a consequence it follows from [12, Proposition 3] that $T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)$ is locally polyhedral near $(\bar{v}, \bar{v}^*)$ for every $\bar{v}^*$ satisfying $(\bar{v}, \bar{v}^*) \in T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)$. Thus by Theorem 5 we must have $\mathcal{F}_1^v - \mathcal{F}_2^v = \mathcal{K}_{\bar{K}_\Gamma}(\bar{v}, \bar{z}^*)$ and $\mathcal{F}_1^\lambda - \mathcal{F}_2^\lambda = T_{\bar{\Lambda}(\bar{v})}(\bar{\lambda}) = \{0\}$. Hence we have $\mathcal{F}_1^v = \mathcal{K}_{\bar{K}_\Gamma}(\bar{v}, \bar{z}^*)$, $\mathcal{F}_2^v = \mathcal{L}(\mathcal{K}_{\bar{K}_\Gamma}(\bar{v}, \bar{z}^*))$, $\mathcal{F}_1^\lambda = \mathcal{F}_2^\lambda = \{0\}$, $\eta = 0$ and hence $\delta v \in \mathcal{L}(\mathcal{K}_{\bar{K}_\Gamma}(\bar{v}, \bar{z}^*))$.*

2. If $\bar{K}_\Gamma$ is a subspace then for every $\bar{v} \in \bar{K}_\Gamma$, $\bar{z}^* \in N_{\bar{K}_\Gamma}(\bar{v})$ there holds $\mathcal{K}_{\bar{K}_\Gamma}(\bar{v}, \bar{z}^*) = \bar{K}_\Gamma$.

3. If $\mathcal{K}_{\bar{K}_\Gamma}(\bar{v}, \bar{z}^*)$ is a subspace then the only face of $\mathcal{K}_{\bar{K}_\Gamma}(\bar{v}, \bar{z}^*)$ is $\mathcal{K}_{\bar{K}_\Gamma}(\bar{v}, \bar{z}^*)$ itself and therefore $\mathcal{F}_1^v = \mathcal{F}_2^v = \mathcal{F}_1^v - \mathcal{F}_2^v = \mathcal{K}_{\bar{K}_\Gamma}(\bar{v}, \bar{z}^*)$. Similarly, if $\bar{\lambda} \in \mathrm{ri}\,\bar{\Lambda}(\bar{v})$ then $T_{\bar{\Lambda}(\bar{v})}(\bar{\lambda})$ is a subspace and $\mathcal{F}_1^\lambda = \mathcal{F}_2^\lambda = \mathcal{F}_1^\lambda - \mathcal{F}_2^\lambda = T_{\bar{\Lambda}(\bar{v})}(\bar{\lambda})$.

Example 1 (cf. [17, Examples 1,2]). *Consider the MPEC*

$$\min_{x,y} \quad F(x,y) := x_1 - \frac{3}{2}y_1 + x_2 - \frac{3}{2}y_2 - y_3$$

$$\mathrm{s.t.} \quad 0 \in \phi(x,y) + N_\Gamma(y),$$

$$G_1(x,y) = G_1(x) := -x_1 - 2x_2 \leq 0,$$

$$G_2(x,y) = G_2(x) := -2x_1 - x_2 \leq 0,$$

where

$$\phi(x,y) := \begin{pmatrix} y_1 - x_1 \\ y_2 - x_2 \\ -1 \end{pmatrix}, \quad \Gamma := \left\{ y \in \mathbb{R}^3 \,\middle|\, g_1(y) := y_3 + \frac{1}{2}y_1^2 \leq 0, \; g_2(y) := y_3 + \frac{1}{2}y_2^2 \leq 0 \right\}.$$

As it was demonstrated in [17], $\bar{x} = (0,0)$ and $\bar{y} = (0,0,0)$ is the unique global solution and Assumption 1 is fulfilled. Straightforward calculations yield

$$\bar{\Lambda} = \{\lambda \in \mathbb{R}_+^2 \mid \lambda_1 + \lambda_2 = 1\}, \quad \bar{K}_\Gamma = \mathbb{R}^2 \times \{0\}.$$

For every $v \in \mathbb{R}^3$, $\lambda \in \mathbb{R}^2$ we have $v^T \nabla^2(\lambda^T g)(\bar{y})v = \lambda_1 v_1^2 + \lambda_2 v_2^2$ yielding

$$\bar{\Lambda}(v) = \begin{cases} \{(1,0)\} & \text{if } |v_1| > |v_2|, \\ \bar{\Lambda} & \text{if } |v_1| = |v_2|, \\ \{(0,1)\} & \text{if } |v_1| < |v_2|. \end{cases}$$

We now show that the mapping g is 2-nondegenerate in every direction $0 \neq v \in \bar{K}_\Gamma$, i.e. we have to verify

$$\begin{pmatrix} \mu_1 v_1 \\ \mu_2 v_2 \\ 0 \end{pmatrix} \in \left(N_{\bar{K}_\Gamma}(v)\right)^+ = \left(\{0\} \times \{0\} \times \mathbb{R}\right)^+ = \{0\} \times \{0\} \times \mathbb{R}, \; \mu \in \left(\bar{\Lambda}(v)\right)^+ \; \Rightarrow \; \mu = (0,0) \tag{63}$$

for every $0 \neq v \in \bar{K}_\Gamma = \mathbb{R}^2 \times \{0\}$. If $\bar{\Lambda}(v)$ is a singleton this holds obviously true because then $\left(\bar{\Lambda}(v)\right)^+ = \{0\}$. But the only case when $\bar{\Lambda}(v)$ is not a singleton is when $|v_1| = |v_2|$, which together with $v \neq 0$ implies $|v_1| = |v_2| > 0$ and we see that (63) holds in this case as well.

We claim that the optimality conditions of Theorem 5 hold with $\bar{v} = (1, 1, 0)$, $\bar{z}^* = (0, 0, 0)$, $\bar{\lambda} = (\frac{1}{2}, \frac{1}{2})$, $\eta = (0, 0)$, $\delta v = (0, 0, 0)$, $\mathcal{F}_1^v = \mathcal{F}_2^v = \bar{K}_\Gamma$, $w = -(1, 1, 0)$, $\mathcal{F}_1^\lambda = \mathcal{F}_2^\lambda = T_{\bar{\Lambda}(\bar{v})}(\bar{\lambda}) = \{(\eta_1, \eta_2) \,|\, \eta_1 + \eta_2 = 0\}$ and $\sigma = (0, 0)$. Indeed, we obviously have $\bar{v} \in \bar{K}_\Gamma$, $\bar{z}^* \in N_{\bar{K}_\Gamma}(\bar{v})$ and, since $\bar{K}_\Gamma$ is a subspace and $\bar{\lambda} \in \operatorname{ri} \bar{\Lambda}(\bar{v})$, by Remark 2, $\mathcal{K}_{\bar{K}_\Gamma}(\bar{v}, \bar{z}^*) = \bar{K}_\Gamma$, $\mathcal{F}_1^v, \mathcal{F}_2^v$ are faces of $\mathcal{K}_{\bar{K}_\Gamma}(\bar{v}, \bar{z}^*)$, $\mathcal{F}_1^\lambda, \mathcal{F}_2^\lambda$ are faces of $T_{\bar{\Lambda}(\bar{v})}(\bar{\lambda})$ and $\delta v \in \operatorname{ri} \mathcal{F}_2^v = \bar{K}_\Gamma$, $w \in \mathcal{F}_1^v - \mathcal{F}_2^v = \bar{K}_\Gamma$, $\eta \in \mathcal{F}_1^\lambda - \mathcal{F}_2^\lambda = T_{\bar{\Lambda}(\bar{v})}(\bar{\lambda})$ and $\sigma \geq 0$. Conditions (60a), (60b), (60c) amount to

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} - \begin{pmatrix} \sigma_1 + 2\sigma_2 \\ 2\sigma_1 + \sigma_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} \frac{3}{2} \\ \frac{3}{2} \\ 1 \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \\ 0 \end{pmatrix} + \begin{pmatrix} \bar{\lambda}_1 w_1 \\ \bar{\lambda}_2 w_2 \\ 0 \end{pmatrix} - 2 \begin{pmatrix} \eta_1 \bar{v}_1 \\ \eta_2 \bar{v}_2 \\ 0 \end{pmatrix} \in (\mathcal{F}_1^v - \mathcal{F}_2^v)^\circ = \bar{K}_\Gamma^\perp = \{0\} \times \{0\} \times \mathbb{R}$$

$$\begin{pmatrix} \bar{v}_1 w_1 \\ \bar{v}_2 w_2 \end{pmatrix} \in (F_1^\lambda - \mathcal{F}_2^\lambda)^\circ = T_{\bar{\Lambda}(\bar{v})}(\bar{\lambda})^\perp = \mathbb{R} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

and it is easy to see that they are fulfilled. Further, (60d) holds because of $\delta v = 0$ and $\mathcal{F}_2^v$ is a subspace, and (60e) is fulfilled as well. Finally, we have $\mathcal{F}_1^v - \mathcal{F}_2^v = \mathcal{K}_{\bar{K}_\Gamma}(\bar{v}, \bar{z}^*)$, $\mathcal{F}_1^\lambda - \mathcal{F}_2^\lambda = T_{\bar{\Lambda}(\bar{v})}(\bar{\lambda})$ and $\bar{v} \neq 0$. Thus the optimality conditions of Theorem 5 are fulfilled.

At the end of this section we want to formulate the necessary optimality conditions in Theorem 5 in terms of index sets instead of faces. Recall the definitions of $\bar{I}$, $\bar{I}(v)$, $\bar{J}^+(\lambda)$, $\bar{J}^+(\Xi)$ given in Subsection 5.1.

Theorem 6. *Assume that $(\bar{x}, \bar{y})$ is a local minimizer for problem (1) fulfilling Assumption 1. Further assume that g is 2-nondegenerate in every nonzero critical direction $0 \neq v \in \bar{K}_\Gamma$ at $(\bar{y}, \bar{y}^*)$, where $\bar{y}^* := -\phi(\bar{x}, \bar{y})$. Then there are a critical direction $\bar{v} \in \bar{K}_\Gamma$, a multiplier $\bar{\lambda} \in \bar{\Lambda}(\bar{v})$, index sets $\mathcal{J}^+$, $\mathcal{J}$, $\mathcal{I}^+$, and $\mathcal{I}$ with $\bar{J}^+(\bar{\lambda}) \subseteq \mathcal{J}^+ \subseteq \mathcal{J} \subseteq \bar{J}^+(\bar{\Lambda}(\bar{v})) \subseteq \bar{J}^+(\bar{\Lambda}) \subseteq \mathcal{I}^+ \subseteq \mathcal{I} \subseteq \bar{I}(\bar{v})$ and elements $w \in \mathbb{R}^m$, $\eta, \xi \in \mathbb{R}^q$ and $\sigma \in \mathbb{R}_+^p$ such that*

$$0 = \nabla_x F(\bar{x}, \bar{y}) - \nabla_x \phi(\bar{x}, \bar{y})^T w + \nabla_x G(\bar{x}, \bar{y})^T \sigma, \tag{64a}$$

$$0 = \nabla_y F(\bar{x}, \bar{y}) - \nabla_y \phi(\bar{x}, \bar{y})^T w + \nabla_y G(\bar{x}, \bar{y})^T \sigma - \nabla^2(\bar{\lambda}^T g)(\bar{y}) w + \nabla g(\bar{y})^T \xi + 2\nabla^2(\eta^T g)(\bar{y})\bar{v}, \tag{64b}$$

$$\xi_i = 0 \ \text{if} \ i \notin \mathcal{I}, \tag{64c}$$

$$\xi_i \geq 0, \nabla g_i(\bar{y})^T w \leq 0 \ \text{if} \ i \in \mathcal{I} \setminus \mathcal{I}^+, \tag{64d}$$

$$\nabla g_i(\bar{y})^T w = 0 \ \text{if} \ i \in \mathcal{I}^+, \tag{64e}$$

$$\nabla g(\bar{y})^T \eta = 0, \ \eta_i = 0, i \notin \mathcal{J}, \ \eta_i \geq 0, i \in \mathcal{J} \setminus \mathcal{J}^+, \tag{64f}$$

$$0 = \sigma_i G_i(\bar{x}, \bar{y}), \ i = 1, \ldots, p. \tag{64g}$$

Moreover, there are $\delta v \in \mathbb{R}^m$, $s_{\delta v}, s_w \in \mathbb{R}^m$ and $\bar\mu \in \mathbb{R}^q$ such that

$$\nabla g_i(\bar y)^T \delta v = 0, \ i \in \bar J^+(\bar \Lambda), \ \nabla g_i(\bar y)^T \delta v \le 0, \ i \in \bar I(\bar v) \setminus \bar J^+(\bar \Lambda), \tag{64h}$$

$$\mathcal{I} = \{i \in \bar I(\bar v) \,|\, \nabla g_i(\bar y)^T \delta v = 0\}, \tag{64i}$$

$$\nabla g_i(\bar y)^T s_{\delta v} + \bar v^T \nabla^2 g_i(\bar y) \delta v = 0, i \in \bar J^+(\bar \lambda), \ \nabla g_i(\bar y) s_{\delta v} + \bar v^T \nabla^2 g_i(\bar y)^T \delta v \le 0, i \in \bar J^+(\bar \Lambda(\bar v)) \setminus \bar J^+(\bar \lambda) \tag{64j}$$

$$\mathcal{J} = \{i \in \bar J^+(\bar \Lambda(\bar v))) \,|\, \nabla g_i(\bar y)^T s_{\delta v} + \bar v^T \nabla^2 g_i(\bar y) \delta v = 0\} \tag{64k}$$

$$\nabla g(\bar y)^T \bar \mu = 0, \ \bar \mu_i = 0, i \notin \mathcal{J}, \bar \mu_i \ge 0, i \in \mathcal{J} \setminus \bar J^+(\bar \lambda) \tag{64l}$$

$$\mathcal{J}^+ = \bar J^+(\bar \lambda) \cup \{i \in \mathcal{J} \setminus \bar J^+(\bar \lambda) \,|\, \bar \mu_i > 0\} \tag{64m}$$

$$\nabla g_i(\bar y)^T s_w + \bar v^T \nabla^2 g_i(\bar y) w = 0, i \in \mathcal{J}^+, \ \nabla g_i(\bar y)^T s_w + \bar v^T \nabla^2 g_i(\bar y) w \le 0, i \in \mathcal{J} \setminus \mathcal{J}^+. \tag{64n}$$

Furthermore, if $\mathcal{I} = \bar I(\bar v)$, $\mathcal{J}^+ = \bar J^+(\bar \lambda)$ and $\mathcal{J} = \bar J^+(\bar \Lambda(\bar v))$ then one of the following two cases must occur: case (a) $\bar v \ne 0$; case (b) $\bar v = 0$, $\mathcal{L}(\bar K_\Gamma) = \{0\}$ and $\bar \lambda \in \Sigma(0, \bar z^*)$ for some $\bar z^* = \sum_{i \in \mathcal{I}} \nabla g_i(\bar y) \alpha_i$ with $\alpha_i > 0$, $i \in \mathcal{I}^+ \setminus \bar J^+(\bar \Lambda)$.

Otherwise, if either $\mathcal{I} \ne \bar I(\bar v)$ or $\mathcal{J}^+ \ne \bar J^+(\bar \lambda)$ or $\mathcal{J} \ne \bar J^+(\bar \Lambda(\bar v))$ then $\bar v \ne 0$ and there are some $\delta x \in \mathbb{R}^n$ and some $\bar z^* = \sum_{i \in \mathcal{I}} \nabla g_i(\bar y) \alpha_i$ with $\alpha_i > 0$, $i \in \mathcal{I}^+ \setminus \bar J^+(\bar \Lambda)$ such that conditions (61a)-(61c) hold and $T_{\text{gph}\,\widehat N_\Gamma}(\bar y, \bar y^*)$ is not locally polyhedral near $(\bar v, -\nabla\phi(\bar x, \bar y)(\delta x, \bar v))$.

Proof. Let $\bar v, \bar \lambda, \delta v, \eta, w, \sigma, \mathcal{F}_1^v, \mathcal{F}_2^v, \mathcal{F}_1^\lambda, \mathcal{F}_2^\lambda$ as in Theorem 5. The index sets $\mathcal{I}, \mathcal{I}^+, \mathcal{J}, \mathcal{J}^+$ were chosen such that

$$\mathcal{F}_1^v - \mathcal{F}_2^v = \{s \,|\, \nabla g_i(\bar y)^T s = 0, i \in \mathcal{I}^+, \ \nabla g_i(\bar y)^T s \le 0, i \in \mathcal{I} \setminus \mathcal{I}^+\},$$
$$F_1^\lambda - \mathcal{F}_2^\lambda = \{\mu \,|\, \nabla g(\bar y)^T \mu = 0, \ \mu_i = 0, i \notin \mathcal{J}, \ \mu_i \ge 0, i \in \mathcal{J} \setminus \mathcal{J}^+\}.$$

Then $(\mathcal{F}_1^v - \mathcal{F}_2^v)^\circ = \{\nabla g(\bar y)^T \xi \,|\, \xi_i = 0, i \notin \mathcal{I}, \ \xi_i \ge 0, i \in \mathcal{I} \setminus \mathcal{I}^+\}$ and conditions (64a)-(64g) follow immediately. (64h) states that $\delta v \in T_{\bar K_\Gamma}(\bar v)$ whereas (64i) results from the requirement $\delta v \in \text{ri}\,\mathcal{F}_2^v$ together with Proposition 4. The index set $\mathcal{I}^+$ is related with $\bar z^* \in N_{\bar K_\Gamma}(\bar v)$. Since we do not have any further condition on $\bar z^*$, the same applies to $\mathcal{I}^+$. (64j) states that $\bar v^T \nabla^2 g(\bar y) \delta v \in T_{\bar \Lambda(\bar v)}(\bar \lambda)^\circ$ and condition (64k) results from the second part of (60d). The point $\bar \mu$ denotes any point in ri F_2^λ yielding the condition (64m) by Proposition 4. Finally, condition (64n) is equivalent to (60c). $\qquad\square$

The optimality conditions of Theorem 6 dramatically simplify under the assumption that $\bar \Lambda = \{\bar \lambda\}$ is a singleton. In this case the condition that g is 2-nondegenerate in every nonzero critical direction $0 \ne v \in \bar K_\Gamma$ at $(\bar y, \bar y^*)$ holds automatically. Further, by the formula for the tangent cone in Theorem 3, it is easy to see that in this case the tangent cone $T_{\text{gph}\,\widehat N_\Gamma}(\bar y, \bar y^*)$ is polyhedral and therefore by Theorem 6, we must have

$$\bar J^+(\bar \lambda) = \mathcal{J}^+ = \mathcal{J} = \bar J^+(\bar \Lambda) \subseteq \mathcal{I}^+ \subseteq \mathcal{I} = \bar I(\bar v).$$

Let $w \in \mathbb{R}^m$, $\xi \in \mathbb{R}^q$ and $\sigma \in \mathbb{R}^p_+$ be those found in Theorem 6 and take, as already pointed out in Remark 2, $\eta = 0$. Moreover, we can take $\delta v = 0$, $s_{\delta v} = 0$ and $\bar \mu = 0$ in order to fulfill conditions (64h)-(64m). Finally, the assumption that $\bar \Lambda = \{\bar \lambda\}$ is a singleton implies that the gradients $\nabla g_i(\bar y)$, $i \in \bar J^+(\bar \lambda)$ are linearly independent and therefore there always exists an element s_w fulfilling (64n). Therefore we have the following corollary.

Corollary 1. *Assume that $(\bar{x}, \bar{y})$ is a local minimizer for problem (1) fulfilling Assumption 1. Further assume that the multiplier set $\bar{\Lambda} = \{\bar{\lambda}\}$ is a singleton. Then there are a critical direction $\bar{v} \in \bar{K}_\Gamma$, index set $\mathcal{I}^+$ with $\bar{J}^+(\bar{\lambda}) \subseteq \mathcal{I}^+ \subseteq \bar{I}(\bar{v})$ and elements $w \in \mathbb{R}^m$, $\xi \in \mathbb{R}^q$ and $\sigma \in \mathbb{R}^p_+$ such that*

$$0 = \nabla_x F(\bar{x}, \bar{y}) - \nabla_x \phi(\bar{x}, \bar{y})^T w + \nabla_x G(\bar{x}, \bar{y})^T \sigma, \tag{65a}$$

$$0 = \nabla_y F(\bar{x}, \bar{y}) - \nabla_y \phi(\bar{x}, \bar{y})^T w + \nabla_y G(\bar{x}, \bar{y})^T \sigma - \nabla^2(\bar{\lambda}^T g)(\bar{y})w + \nabla g(\bar{y})^T \xi, \tag{65b}$$

$$\xi_i = 0 \ \textit{if} \ i \notin \bar{I}(\bar{v}), \tag{65c}$$

$$\xi_i \geq 0, \nabla g_i(\bar{y})^T w \leq 0 \ \textit{if} \ i \in \bar{I}(\bar{v}) \setminus \mathcal{I}^+, \tag{65d}$$

$$\nabla g_i(\bar{y})^T w = 0 \ \textit{if} \ i \in \mathcal{I}^+, \tag{65e}$$

$$0 = \sigma_i G_i(\bar{x}, \bar{y}), \ i = 1, \dots, p. \tag{65f}$$

We now want to compare our optimality conditions for (MPEC) with the known M-stationarity conditions for (MPCC) defined as follows.

Definition 10 (M-stationary condition for (MPCC))**.** *Let $(\bar{x}, \bar{y}, \bar{\lambda})$ be a feasible solution for problem (MPCC). We say that $(\bar{x}, \bar{y}, \bar{\lambda})$ satisfies the M-stationary condition for (MPCC) if there exist $w \in \mathbb{R}^m, \xi \in \mathbb{R}^q, \sigma \in \mathbb{R}^p_+$ such that*

$$0 = \nabla_x F(\bar{x}, \bar{y}) - \nabla_x \phi(\bar{x}, \bar{y})^T w + \nabla_x G(\bar{x}, \bar{y})^T \sigma, \tag{66a}$$

$$0 = \nabla_y F(\bar{x}, \bar{y}) - \nabla_y \phi(\bar{x}, \bar{y})^T w + \nabla_y G(\bar{x}, \bar{y})^T \sigma - \nabla^2(\bar{\lambda}^T g)(\bar{y})w + \nabla g(\bar{y})^T \xi, \tag{66b}$$

$$\xi_i = 0 \ \textit{if} \ g_i(\bar{y}) < 0, \bar{\lambda}_i = 0, \tag{66c}$$

$$\nabla g_i(\bar{y})^T w = 0 \ \textit{if} \ g_i(\bar{y}) = 0, \bar{\lambda}_i > 0, \tag{66d}$$

either $\xi_i > 0, \nabla g_i(\bar{y})^T w < 0$ *or* $\xi_i \nabla g_i(\bar{y})^T w = 0$ *if* $g_i(\bar{y}) = \bar{\lambda}_i = 0,$ \tag{66e}

$$0 = \sigma_i G_i(\bar{x}, \bar{y}), \ i = 1, \dots, p. \tag{66f}$$

It is well know that problem (MPCC) may not be equivalent to problem (MPEC) in the case when the lower level problem does not have unique multiplier. Moreover, in this case it is also possible that at a locally optimal solution $(\bar{x}, \bar{y}, \bar{\lambda})$ even the weakest known constraint qualification, the MPCC-GCQ (Guignard constraint qualification), ensuring M-stationarity is not fulfilled. E.g., it was shown in [17] that for Example 1 MPCC-GCQ does not hold at $(\bar{x}, \bar{y}, \lambda)$ for any $\lambda \in \bar{\Lambda}$.

For the case when the multiplier set $\bar{\Lambda} = \{\bar{\lambda}\}$ is a singleton, we now compare our necessary optimality conditions of Corollary 1 with M-stationarity condition for (MPCC). By [2, Proposition 2], the assumption of MSCQ for (MPEC) is weaker than the corresponding one for (MPCC). Suppose that $(\bar{x}, \bar{y}, \bar{\lambda})$ satisfies the optimality condition in Corollary 1 and let $w \in \mathbb{R}^m$, $\xi \in \mathbb{R}^q$ and $\sigma \in \mathbb{R}^p_+$ be those found in Corollary 1. Then (66a)-(66b) and (66f) hold. Since

$$i \notin \bar{I}(\bar{v}) \iff \begin{array}{l} \text{either } g_i(\bar{y}) = 0, \nabla g_i(\bar{y})^T \bar{v} < 0, \bar{\lambda}_i = 0 \\ \text{or } g_i(\bar{y}) < 0, \bar{\lambda}_i = 0 \end{array}$$

and $\bar{J}^+(\bar{\lambda}) = \{i | g_i(\bar{y}) = 0, \bar{\lambda}_i > 0\}$, (65c) and (65e) implies that $\xi_i = 0$ if $\bar{\lambda}_i = 0$ and $\nabla g_i(\bar{y})^T w = 0$ if $\bar{\lambda}_i > 0$. It follows that (66c)-(66e) hold. Therefore $(\bar{x}, \bar{y}, \bar{\lambda})$ must satisfy the M-stationary condition for (MPCC) as well.

It is not difficult to show that the M-stationarity conditions of Definition 10 imply the necessary optimality conditions of Corollary 1 provided the *linear independence constraint*

qualification (LICQ) holds for the lower level problem at $\bar{y}$. Indeed, under LICQ the multiplier set $\bar{\Lambda} = \{\bar{\lambda}\}$ is a singleton. Given w, ξ and σ fulfilling (66), define

$$\mathcal{I}^+ := \bar{J}^+(\bar{\lambda}) \cup \{i \in \bar{I} \,|\, \xi_i < 0\}, \ \mathcal{I} := \mathcal{I}^+ \cup \{i \in \bar{I} \,|\, \nabla g_i(\bar{y})^T w \le 0, \xi_i > 0\}$$

and then find $\bar{v}$ fulfilling

$$\nabla g_i(\bar{y})^T \bar{v} = 0, i \in \mathcal{I}, \ \nabla g_i(\bar{y})^T \bar{v} = -1, i \in \bar{I} \setminus \mathcal{I}$$

which exists due to the imposed LICQ. It follows that $\bar{v} \in \bar{K}_\Gamma$, $\mathcal{I} = \bar{I}(\bar{v})$ and that the conditions (65) are fulfilled. Hence, under LICQ the optimality conditions of Corollary 1 are equivalent to the M-stationarity conditions of Definition 10. However, the following example demonstrates that the optimality conditions of Corollary 1 are sharper, when $\bar{\Lambda}$ is a singleton but LICQ fails.

Example 2. *Consider the problem*

$$\min_{x \in \mathbb{R}, y \in \mathbb{R}^2} x + y_1 + y_2$$

$$\textit{subject to } 0 \in \begin{pmatrix} x + 2y_1 \\ x + y_2 \end{pmatrix} + \widehat{N}_\Gamma(y) \textit{ where } \Gamma := \{y \in \mathbb{R}^2 \,|\, y_1 \le 0, \ -y_2 \le 0, y_1 + y_2 \le 0\}$$

at $\bar{x} = 0$, $\bar{y} = (0,0)$. Straightforward calculations yield that $\bar{\Lambda} = \{(0,0,0)\}$ and that $(\bar{x}, \bar{y})$ is not a local minimizer. However, the M-stationary conditions of Definition 10 amount to

$$0 = 1 - (w_1 + w_2),$$
$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 2w_1 \\ w_2 \end{pmatrix} + \begin{pmatrix} \xi_1 + \xi_3 \\ -\xi_2 + \xi_3 \end{pmatrix},$$
$$(w_1 < 0, \ \xi_1 > 0) \vee (w_1 \xi_1 = 0),$$
$$(-w_2 < 0, \ \xi_2 > 0) \vee (-w_2 \xi_2 = 0),$$
$$(w_1 + w_2 < 0, \ \xi_3 > 0) \vee ((w_1 + w_2)\xi_3 = 0),$$

where $(X) \vee (Y)$ denotes either X or Y holds, and are uniquely fulfilled with $w = (0,1)$ and $\xi = (-1,0,0)$. Now let us show that the optimality condition in Corollary 1 with $w = (0,1)$, and $\xi = (-1,0,0)$ does not hold.

By applying Theorem 4 we deduce that MSCQ and consequently Assumption 1 are fulfilled. Since the only multiplier is $\bar{\lambda} = (0,0,0)$, we have $\bar{J}^+(\bar{\lambda}) = \emptyset$ and $\bar{K}_\Gamma = \Gamma$. Since $w_1 = 0$ and $\nabla g_1(\bar{y})w = w_1$, (65e) holds if and only if $\{1\} \subseteq \mathcal{I}^+$. Since $\{1\} \subseteq \mathcal{I}^+ \subseteq \bar{I}(\bar{v})$, one must have $\nabla g_1(\bar{y})^T \bar{v} = \bar{v}_1 = 0$. But then (65c) means $\xi_3 = 0$ if $3 \notin \bar{I}(\bar{v})$, which in turn means that $\nabla g_3(\bar{y})^T \bar{v} = \bar{v}_1 + \bar{v}_2 < 0$. This is impossible since we cannot find $\bar{v} \in \bar{K}_\Gamma = \Gamma$ satisfying $\bar{v}_1 = 0$ and $\bar{v}_1 + \bar{v}_2 < 0$. This shows that the M-stationarity conditions do not correctly describe the faces of the critical cone.

Finally we want to compare our results with the ones of Gfrerer and Outrata [15], where the limiting normal cone of the normal cone mapping was computed and thus could be used to compute the conventional M-stationarity conditions for problem (MPEC). The assumption 2-LICQ used in [15] cannot be characterized by first-order and second-order derivatives of the constraint mapping g, however the sufficient condition for 2-LICQ as stated in [15, Proposition

3] is stronger than the 2-nondegeneracy assumption we use. The sufficient condition for 2-LICQ in direction $\bar{v} \in \bar{K}_\Gamma$ used in [15, Proposition 3] now states that for every index set J with $\bar{J}^+(\bar{\Lambda}(\bar{v})) \subseteq J \subseteq \bar{I}(\bar{v})$ satisfying

$$\nabla g_i(\bar{y})^T s + \bar{v}^T \nabla^2 g_i(\bar{y}) \bar{v} \begin{cases} = 0 & i \in J \\ \leq 0 & i \in \bar{I}(\bar{v}) \setminus J \end{cases} \quad \text{for some } s \in \mathbb{R}^m$$

the mapping $(g_i)_{i \in J}$ is 2-regular in direction $\bar{v}$. Such an index set J always exists, e.g. by duality theory of linear programming $J = \bar{J}^+(\bar{\Lambda}(\bar{v}))$ is a possible choice. Now choose J large enough such that for every $j \in \bar{I}(\bar{v}) \setminus J$ the gradient $\nabla g_j(\bar{y})$ linearly depend on $\nabla g_i(\bar{y})$, $i \in J$. so that J meets the requirements on the index set $\hat{J}$ used in Subsection 5.1 and we see that the assumption of 2-regularity of $(g_i)_{i \in J}$ in direction $\bar{v}$ implies 2-nondegeneracy of g in direction $\bar{v}$.

Further one can show that the necessary conditions of Theorem 6 are stronger than the M-stationary conditions which one could obtain with the M-stationary conditions of [15, Theorem 4] insofar as an additional condition on δx is included in Theorem 6.

References

[1] E. R. Avakov, *Extremum conditions for smooth problems with equality-type constraints*, USSR Comput. Math. Math. Phys., 25 (1985), pp. 24–32.

[2] L. Adam, R. Henrion, J. Outrata, *On M-stationarity conditions in MPECs and the associated qualification conditions*, Math. Program., 168 (2018), pp. 229-259.

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

[4] S. Dempe, J. Dutta, *Is bilevel programming a special case of a mathematical program with complementarity constraints?*, Math. Program., 131 (2012), pp. 37-48.

[5] A.L. Dontchev, R. T. Rockafellar, *Characterizations of strong regularity for variational inequalities over polyhedral convex sets*, SIAM J. Optim., 6 (1996), pp. 1087–1105.

[6] A.L. Dontchev, R. T. Rockafellar, *Regulartity and conditioning of solution mappings in variatonal anlysis*, Set-Valued Anal., 12 (2004), pp. 79–109.

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

[8] 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.

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

[10] H. Gfrerer, *On directional metric subregularity and second-order optimality conditions for a class of nonsmooth mathematical programs*, SIAM J. Optim., 23 (2013), pp. 632–665.

[11] 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.

[12] H. Gfrerer, *Linearized M-stationarity conditions for general optimization problems*, Set-Valued Var. Anal., https://doi.org/10.1007/s11228-018-0491-6.

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

[14] H. Gfrerer, B.S. Mordukhovich, *Complete characterizations of tilt stability in nonlinear programming under weakest qualification conditions*, SIAM J. Optim., 25 (2015), pp. 2081–2119.

[15] H. Gfrerer, J.V. Outrata, *On computation of limiting coderivatives of the normal-cone mapping to inequality systems and their applications*, Optimization, 65 (2016), pp. 671–700.

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

[17] H. Gfrerer, J.J. Ye, *New constraint qualifications for mathematical programs with equilibrium constraints via variational analysis*, SIAM J. Optim., 27 (2017), pp. 842-865.

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

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

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

[21] J.V. Outrata, M. Kočvara, J. Zowe, *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints: Theory, Applications and Numerical Results*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998.

[22] S. M. Robinson, *Local structure of feasible sets in nonlinear programming. Part II: Nondegeneracy*, Math. Program. Studies 22 (1984), pp. 217230.

[23] R.T. Rockafellar, *Convex analysis*, Princeton, New Jersey, 1970.

[24] R.T. Rockafellar, R. J-B. Wets, *Variational analysis*, Springer, Berlin, 1998.

[25] A.A. Tret'yakov, *Necessary and sufficient conditions for optimality of p-th order*, USSR Comput. Math. Math. Phys., 24 (1984), pp. 123–127.

[26] J.J. Ye, X.Y. Ye, *Necessary optimality conditions for optimization problems with variational inequality constraints*, Math. Oper. Res., 22 (1997), pp. 977-997.