JOHANNES KEPLER UNIVERSITY LINZ

Institute of Computational Mathematics

New Sharp Necessary Optimality Conditions for Mathematical Programs with Equilibrium Constraints

Helmut Gfrerer
Institute of Computational Mathematics, Johannes Kepler University
Altenberger Str. 69, 4040 Linz, Austria

Jane J. Ye
Department of Mathematics and Statistics, University of Victoria
Victoria, B.C., Canada V8W 2Y2

NuMa-Report No. 2018-05 December 2018

A–4040 LINZ, Altenbergerstraße 69, Austria

Technical Reports before 1998:

1995

95-1 Hedwig Brandstetter
Was ist neu in Fortran 90? March 1995

95-2 G. Haase, B. Heise, M. Kuhn, U. Langer
Adaptive Domain Decomposition Methods for Finite and Boundary Element Equations. August 1995

95-3 Joachim Schöberl
An Automatic Mesh Generator Using Geometric Rules for Two and Three Space Dimensions. August 1995

1996

96-1 Ferdinand Kickinger
Automatic Mesh Generation for 3D Objects. February 1996

96-2 Mario Goppold, Gundolf Haase, Bodo Heise und Michael Kuhn
Preprocessing in BE/FE Domain Decomposition Methods. February 1996

96-3 Bodo Heise
A Mixed Variational Formulation for 3D Magnetostatics and its Finite Element Discretisation. February 1996

96-4 Bodo Heise und Michael Jung
Robust Parallel Newton-Multilevel Methods. February 1996

96-5 Ferdinand Kickinger
Algebraic Multigrid for Discrete Elliptic Second Order Problems. February 1996

96-6 Bodo Heise
A Mixed Variational Formulation for 3D Magnetostatics and its Finite Element Discretisation. May 1996

96-7 Michael Kuhn
Benchmarking for Boundary Element Methods. June 1996

1997

97-1 Bodo Heise, Michael Kuhn and Ulrich Langer
A Mixed Variational Formulation for 3D Magnetostatics in the Space $H(rot) \cap H(div)$ February 1997

97-2 Joachim Schöberl
Robust Multigrid Preconditioning for Parameter Dependent Problems I: The Stokes-type Case. June 1997

97-3 Ferdinand Kickinger, Sergei V. Nepomnyaschikh, Ralf Pfau, Joachim Schöberl
Numerical Estimates of Inequalities in $H^{\frac{1}{2}}$. August 1997

97-4 Joachim Schöberl
Programmbeschreibung NAOMI 2D und Algebraic Multigrid. September 1997

From 1998 to 2008 technical reports were published by SFB013. Please see
`http://www.sfb013.uni-linz.ac.at/index.php?id=reports`
From 2004 on reports were also published by RICAM. Please see
`http://www.ricam.oeaw.ac.at/publications/list/`
For a complete list of NuMa reports see
`http://www.numa.uni-linz.ac.at/Publications/List/`

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 (MPEC) 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 (MPCC) reformulation.

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

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

1 Introduction

A mathematical program with variational constraints (MPVIC) is an optimization problem in which the essential constraints are defined by a parametric variational inequality. Since many equilibrium phenomena that arise from engineering and economics are characterized by either an optimization problem or a variational inequality, MPVIC is also referred to as mathematical program with equilibrium constraints (MPEC); see e.g. [21, 23] and the references within.

Consider MPVICs in the following form:

$$\begin{aligned} \text{(MPVIC)} \qquad \min_{x,y} \quad & F(x,y) \\ \text{s.t.} \quad & \langle \phi(x,y), y'-y\rangle \geq 0 \quad \forall y' \in \Gamma, & (1)\\ & G(x,y) \leq 0, & (2) \end{aligned}$$

where $\Gamma := \{y \,|\, g(y) \leq 0\}$ and $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, $\phi : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$, $G : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p$ are assumed to be continuously differentiable and $g : \mathbb{R}^m \to \mathbb{R}^q$ is twice continuously differentiable.

If Γ is convex and the function $g(y)$ is either affine or a certain constraint qualification such as the Slater condition holds for the constraint $g(y) \leq 0$, then by the Kurash-Kuhn-Tucker (KKT) condition,

$$(1) \text{ holds } \iff \exists \lambda : 0 = \phi(x,y) + \nabla g(y)^T \lambda, \quad 0 \leq -g(y) \perp \lambda \geq 0.$$

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

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

This observation leads to the following mathematical program with a complementarity constraint (MPCC):

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

During the last two decades, there has been much progress made in both theories and algorithms for solving MPCCs.

However there are much difficulties involved in the reformulation of MPCCs. First of all, if g is not a convex function then KKT condition is in general necessary but not sufficient, and so one can not justify using MPCC approach. Secondly, even if g is a convex function, if there are more than one multiplier for solving the variational inequality (1), then MPCC is not equivalent to the original (MPVIC) (see Dempe and Dutta [4] in the context of bilevel programs). Precisely, it may be possible that for a local solution $(\bar{x}, \bar{y}, \bar{\lambda})$ of (MPCC), the pair $(\bar{x}, \bar{y})$ is not a local solution of (MPVIC).

Based on the above discussion, in this paper we consider MPECs of the form

$$
\begin{aligned}
(MPEC) \qquad \min_{x,y} \quad & F(x,y) \\
\text{s.t.} \quad & 0 \in \phi(x,y) + \widehat{N}_\Gamma(y), \\
& G(x,y) \le 0,
\end{aligned}
\tag{3}
$$

where $\widehat{N}_\Gamma(y)$ denotes the so-called *regular normal cone* to the set Γ at y (see Definition 1). 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 hence problem (3) is equivalent to problem (MPVIC).

In the MPEC-literaure, there exist a lot of constraint qualifications ensuring the Mordukhovich (M-)stationarity of locally optimal solutions. The weakest one of these constraint qualifications is MPEC-GCQ (Guignard constraint qualification), see [8] for a proof of M-stationarity of locally optimal solutions under MPEC-GCQ. For the problem (MPEC), it was shown by Ye and Ye [27] that calmness of the perturbation mapping associated with the constraints of (MPEC) (called pseudo upper-Lipschitz continuity in [27]) guarantees M-stationarity of solutions. Recently [2] has compared the two formulations (MPEC) and (MPCC) in terms of calmness. The authors pointed out there that, very often, the calmness condition related to (MPEC) is satisfied at some $(\bar{x}, \bar{y})$ while the one for (MPCC) is not fulfilled at $(\bar{x}, \bar{y}, \lambda)$ for certain multiplier λ. In particular, [2, Example 6] shows that it may be possible that the constraint for (MPEC) satisfies the calmness condition at $(\bar{x}, \bar{y}, 0)$ while the one for the corresponding (MPCC) does not satisfy the calmness condition at $(\bar{x}, \bar{y}, \lambda, 0)$ for any multiplier λ.

In our recent paper [18], we have shown that if multipliers are not unique, then the MPEC Mangasarian-Fromovitz constraint qualification (MFCQ) never holds for problem (MPCC). Moreover we have presented an example for which MPEC-GCQ is violated at $(\bar{x}, \bar{y}, \lambda, 0)$ for any multiplier λ while the calmness holds for the corresponding (MPEC) at $(\bar{x}, \bar{y}, 0)$. In [18, Theorem 5], we have provided a checkable sufficient condition for the calmness of the perturbed feasible region of (MPEC).

In a summary, the upshot is that (MPCC) may not be an equivalent reformulation of the original (MPVIC) even when g is a convex function while (MPEC) is, and it is easier for (MPEC) to satisfy constraint qualifications than (MPCC). Continuing the work in [18], in this paper we aim at developing a new necessary optimality condition for problem (MPEC), which can be applied even when none of constraint qualifications for (MPCC) hold. For reference we review the M-necessary optimality condition for (MPCC) here.

Theorem 1 (M-stationary condition for MPCC)**.** *Assume that $(\bar{x}, \bar{y}, \bar{\lambda})$ is a local minimizer for the problem (MPCC). Assume that a MPCC-GCQ holds at $(\bar{x}, \bar{y}, \bar{\lambda})$. Then there exist $w \in \mathbb{R}^m, \xi \in \mathbb{R}^q, \sigma \in \mathbb{R}^p_+$ such that*

$$0 = \nabla_x F(\bar{x}, \bar{y}) - \nabla_x \phi(\bar{x}, \bar{y})^T w + \nabla_x G(\bar{x}, \bar{y})^T \sigma, \tag{4a}$$

$$0 = \nabla_y F(\bar{x}, \bar{y}) - \nabla_y \phi(\bar{x}, \bar{y})^T w + \nabla_y G(\bar{x}, \bar{y})^T \sigma - \nabla^2 (\bar{\lambda}^T g)(\bar{y}) w + \nabla g(\bar{y})^T \xi, \tag{4b}$$

$$\xi_i = 0 \ \textit{if} \ g_i(\bar{y}) < 0, \bar{\lambda}_i = 0, \tag{4c}$$

$$\nabla g_i(\bar{y})^T w = 0 \ \textit{if} \ g_i(\bar{y}) = 0, \bar{\lambda}_i > 0, \tag{4d}$$

$$\textit{either} \ \xi_i > 0, \nabla g_i(\bar{y})^T w < 0 \ \textit{or} \ \xi_i \nabla g_i(\bar{y})^T w = 0 \ \textit{if} \ g_i(\bar{y}) = \bar{\lambda}_i = 0, \tag{4e}$$

$$0 = \sigma_i G_i(\bar{x}, \bar{y}), \ i = 1, \ldots, p. \tag{4f}$$

We organize our paper as follows. Section 2 contains the preliminaries from variational geometry and variational analysis. In secion 3, we present some optimality conditions 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. Finally in Section 4, 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. $\mathcal{S}_{\mathbb{R}^q}$ is the unit sphere in $\mathbb{R}^q$. 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 set 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 $\operatorname{conv} \Omega$ and $\operatorname{cl} \Omega$ 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 $P : \mathbb{R}^d \to R^s$ with $s > 1$, we denote by $\nabla P(z)$ the Jacobian matrix of P at z and for any give $w, v \in \mathbb{R}^d$, $w^T \nabla P(\bar{z}) v$ be the vector in 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, 22, 25] 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*

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

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

$$\begin{aligned} &\widehat{N}_\Omega(\bar{z}) := (T_\Omega(\bar{z}))^\circ \\ \textit{and} \qquad &N_\Omega(\bar{z}) := \Big\{ z^* \,|\, \exists z_k \overset{\Omega}{\to} \bar{z} \ \textit{and} \ z_k^* \to z^* \ \textit{such that} \ z_k^* \in \widehat{N}_\Omega(z_k) \ \forall k \Big\}, \end{aligned}$$

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 $(\bar{z}, \bar{w}) \in \operatorname{gph} M$.*

(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*

$$\mathrm{d}(z, M^{-1}(\bar{w})) \le \kappa \mathrm{d}(\bar{w}, M(z)) \ \ \forall z \in Z. \tag{5}$$

(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*

$$\mathrm{d}(z, M^{-1}(w)) \le \kappa \mathrm{d}(w, M(z)) \ \ \forall (w, z) \in W \times Z. \tag{6}$$

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. [18, 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 2 (Mordukhovich criterion)**.** *(cf. [25, 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*

$$\nabla P(\bar{z})^T w^* = 0, \ w^* \in N_D\big(P(\bar{z})\big) \implies w^* = 0. \tag{7}$$

For verifying the property of metric subregularity there are some sufficient conditions known, see e.g. [9, 10, 11, 12, 14].

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 $C = \{z \,|\, \langle a_i, z\rangle \le \alpha_i,\ i = 1, \ldots, p\}$.*

(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*

$$N_\Omega(\bar{z}) = \bigcup_{w \in T_\Omega(\bar{z})} \widehat{N}_{T_\Omega(\bar{z})}(w). \tag{8}$$

Proof. Follows from [12, Lemma 2.2]. □

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 $y \in D$,

$$\widehat{N}_{T_D(y)}(0) = \big(T_{T_D(y)}(0)\big)^\circ = \big(T_D(y)\big)^\circ = \widehat{N}_D(y). \tag{9}$$

The following rules for calculating polar cones will be useful.

Proposition 2. *([24, 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 $(A \cap B)^\circ = \operatorname{cl}(A^\circ + B^\circ)$.*

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. [24, 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 $(\bar{z}, \bar{z}^*) \in \operatorname{gph} N_C$. Then for all $z \in C$ sufficiently close to $\bar{z}$ we have*

$$T_C(z) \supseteq T_C(\bar{z}), \quad N_C(z) \subseteq N_C(\bar{z}). \tag{10}$$

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

$$\operatorname{gph} N_C \cap W = ((\bar{z}, \bar{z}^*) + \operatorname{gph} N_{\mathcal{K}_C(\bar{z}, \bar{z}^*)}) \cap W. \tag{11}$$

In particular we have

$$T_{\operatorname{gph} N_C}(\bar{z}, \bar{z}^*) = \operatorname{gph} N_{\mathcal{K}_C(\bar{z}, \bar{z}^*)}. \tag{12}$$

Further,

$$\widehat{N}_{\operatorname{gph} N_C}(\bar{z}, \bar{z}^*) = (\mathcal{K}_C(\bar{z}, \bar{z}^*))^\circ \times \mathcal{K}_C(\bar{z}, \bar{z}^*) \tag{13}$$

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

$$(F_1 - F_2)^\circ \times (F_1 - F_2) \tag{14}$$

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

Definition 5 (Recession Cone)**.** *([24, 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 [24, 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

$$\mathcal{L}(C) \subseteq \mathcal{L}\big(T_C(\bar{z})\big). \tag{15}$$

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

$$\begin{aligned} \mathcal{L}\big(\mathcal{K}_C(\bar{z}, \bar{z}^*)\big) &= \big(T_C(\bar{z}) \cap [\bar{z}^*]^\perp\big) \cap \big(- T_C(\bar{z}) \cap [\bar{z}^*]^\perp\big) = T_C(\bar{z}) \cap (-T_C(\bar{z})) \cap [\bar{z}^*]^\perp \\ &= \mathcal{L}\big(T_C(\bar{z})\big) \end{aligned} \tag{16}$$

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

$$\mathcal{L}\big(T_C(\bar{z})\big)^\perp = \big(T_C(\bar{z}) \cap (-T_C(\bar{z}))\big)^\circ = \operatorname{cl}\big(N_C(\bar{z}) - N_C(\bar{z})\big) = N_C(\bar{z})^+, \tag{17}$$

and hence

$$\mathcal{L}\big(T_C(\bar{z})\big) = \big(N_C(\bar{z})^+\big)^\perp. \tag{18}$$

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

$$\mathcal{L}(\operatorname{gph} N_C) = \mathcal{L}(C) \times (C^+)^\perp. \tag{19}$$

If C is convex polyhedral, then for every $(\bar z, \bar z^*) \in \operatorname{gph} N_C$ we obtain by virtue of (12), (19), (16) and (18)

$$\begin{aligned}\mathcal{L}\big(T_{\operatorname{gph} N_C}(\bar z, \bar z^*)\big) &= \mathcal{L}\big(\operatorname{gph} N_{\mathcal{K}_C(\bar z, \bar z^*)}\big) = \mathcal{L}\big(\mathcal{K}_C(\bar z, \bar z^*)\big) \times \big(\mathcal{K}_C(\bar z, \bar z^*)^+\big)^\perp \\ &= \mathcal{L}(T_C(\bar z)) \times \big(\mathcal{K}_C(\bar z, \bar z^*)^+\big)^\perp = \big(N_C(\bar z)^+\big)^\perp \times \big(\mathcal{K}_C(\bar z, \bar z^*)^+\big)^\perp.\end{aligned} \tag{20}$$

Further, for every $(\delta\bar z, \delta\bar z^*) \in \operatorname{gph} N_{\mathcal{K}_C(\bar z, \bar z^*)}$, by (15) we have

$$\begin{aligned}\mathcal{L}\big(T_{\operatorname{gph} N_{\mathcal{K}_C(\bar z, \bar z^*)}}(\delta\bar z, \delta\bar z^*)\big) &\supseteq \mathcal{L}\big(\operatorname{gph} N_{\mathcal{K}_C(\bar z, \bar z^*)}\big) = \mathcal{L}\big(\mathcal{K}_C(\bar z, \bar z^*)\big) \times \big(\mathcal{K}_C(\bar z, \bar z^*)^+\big)^\perp \\ &= \mathcal{L}\big(T_C(\bar z)\big) \times \big(\mathcal{K}_C(\bar z, \bar z^*)^+\big)^\perp = \mathcal{L}\big(T_{\operatorname{gph} N_C}(\bar z, \bar z^*)\big),\end{aligned} \tag{21}$$

where the equalities follow from (20).

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*

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

Then

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

and for every $v \in K$ we have

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

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

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

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

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

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

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

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

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

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

3 Optimality conditions for a set-constrained optimization problem

In this section we consider an optimization problem of the form

$$\begin{array}{ll} \min & f(z) \\ \text{s.t.} & P(z) \in D, \end{array} \tag{22}$$

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

Let $\bar{z}$ be a local minimizer and $\Omega := \{z \,|\, P(z) \in D\}$ the feasible region for the problem (22). Then $\nabla f(\bar{z})^T u \geq 0 \;\; \forall u \in T_\Omega(\bar{z})$ and so the following basic optimality condition holds:

$$0 \in \nabla f(\bar{z}) + \widehat{N}_\Omega(\bar{z}).$$

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

$$T^{\text{lin}}_{P,D}(\bar{z}) := \{u \in \mathbb{R}^d \,|\, \nabla P(\bar{z})u \in T_D(P(\bar{z}))\}.$$

It is well known that the inclusions

$$T_\Omega(\bar{z}) \subseteq T^{\text{lin}}_{P,D}(\bar{z}), \;\; \widehat{N}_\Omega(\bar{z}) \supseteq \left(T^{\text{lin}}_{P,D}(\bar{z})\right)^\circ, \tag{23}$$

$$\left(T^{\text{lin}}_{P,D}(\bar{z})\right)^\circ \supseteq \nabla P(\bar{z})^T \widehat{N}_D\big(P(\bar{z})\big) \tag{24}$$

are always valid, cf. [25, Theorems 6.31, 6.14]. In order to ensure equality in (23) one has to impose some constraint qualification.

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

(i) (cf. [8]) We say that the generalized Abadie constraint qualification *(GACQ) holds at $\bar{z}$ if*

$$T_\Omega(\bar{z}) = T^{\text{lin}}_{P,D}(\bar{z}). \tag{25}$$

(ii) (cf. [8]) We say that the generalized Guignard constraint qualification *(GGCQ) holds at $\bar{z}$ if*

$$\widehat{N}_\Omega(\bar{z}) = \left(T^{\text{lin}}_{P,D}(\bar{z})\right)^\circ. \tag{26}$$

(iii) (cf. [15]) 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:

$$\text{MSCQ} \Longrightarrow \text{GACQ} \Longrightarrow \text{GGCQ}.$$

Indeed, the first implication follows from [20, 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 $\widehat{N}_\Omega(\bar{z}) = \big(T^{\text{lin}}_{P,D}(\bar{z})\big)^\circ$, but it is very difficult to verify in general. On the other hand, MSCQ is a little bit stronger than GGCQ but there are effective tools for verifying.

Now let us examine conditions ensuring equality in (24). By [24, Corollary 16.3.2] we have

$$\text{cl}\left(\nabla P(\bar{z})^T \widehat{N}_D\big(P(\bar{z})\big)\right) = \big\{w \,|\, \nabla P(\bar{z})w \in \text{cl}\,\text{conv}\, T_D\big(P(\bar{z})\big)\big\}^\circ$$

showing that we can expect equality in (24) only under some restrictive assumption whenever $T_D\big(P(\bar{z})\big)$ is not convex. In the following theorem we provide such assumptions.

Theorem 3. *Assume that $\bar{z}$ is feasible for the problem (22) and assume that*

$$\nabla P(\bar{z})\mathbb{R}^d + \mathcal{L}\left(T_D\big(P(\bar{z})\big)\right) = \mathbb{R}^s. \tag{27}$$

Then the mapping $u \rightrightarrows M(u) := \nabla P(\bar{z})u - T_D\big(P(\bar{z})\big)$ is metrically regular around $(0,0)$ and (24) holds with equality.

Proof. By using the Mordukhovich criterion (7) to show the metric regularity of $M(\cdot)$, we shall show that

$$\lambda \in \ker \nabla P(\bar{z})^T \cap N_{T_D(P(\bar{z}))}(0) \Longrightarrow \lambda = 0, \tag{28}$$

where $\ker \nabla P(\bar{z})^T := \{\lambda \,|\, \nabla P(\bar{z})^T\lambda = 0\}$ is the kernel of the linear operator $\nabla P(\bar{z})^T$. Consider $\lambda \in \ker \nabla P(\bar{z})^T \cap N_{T_D(P(\bar{z}))}(0)$. By the definition of the limiting normal cone we can find sequences $y_k \overset{T_D(P(\bar{z}))}{\longrightarrow} 0$ and $\lambda_k \to \lambda$ such that $\lambda_k \in \widehat{N}_{T_D\left(P(\bar{z})\right)}(y_k)$ for all k. By the definition of $\mathcal{L}\left(T_D\big(P(\bar{z})\big)\right)$ we have $y_k + \mathcal{L}\left(T_D\big(P(\bar{z})\big)\right) \subseteq T_D\big(P(\bar{z})\big)$. We now show that

$$\mathcal{L}\left(T_D\big(P(\bar{z})\big)\right) \subseteq T_{T_D\left(P(\bar{z})\right)}(y_k). \tag{29}$$

Let $v \in \mathcal{L}\left(T_D\big(P(\bar{z})\big)\right)$. Then for every real t we have $tv \in \mathcal{L}\left(T_D\big(P(\bar{z})\big)\right)$ as well and hence $y_k + tv \in T_D\big(P(\bar{z})\big)$ which implies that $v \in T_{T_D\left(P(\bar{z})\right)}(y_k)$. Hence (29) holds. Taking polar on both sides of (29), we have

$$\lambda_k \in \widehat{N}_{T_D\left(P(\bar{z})\right)}(y_k) = (T_{T_D\left(P(\bar{z})\right)}(y_k))^\circ \subseteq \left(\mathcal{L}\left(T_D\big(P(\bar{z})\big)\right)\right)^\circ = \left(\mathcal{L}\left(T_D\big(P(\bar{z})\big)\right)\right)^\perp.$$

It follows that

$$\lambda \in \ker \nabla P(\bar{z})^T \cap \left(\mathcal{L}\left(T_D\big(P(\bar{z})\big)\right)\right)^\perp = \left(\nabla P(\bar{z})\mathbb{R}^d + \mathcal{L}\left(T_D\big(P(\bar{z})\big)\right)\right)^\perp = \{0\},$$

where the first equality follows from the calculus rules for polars in Proposition 2 and the second equality follows from taking the polar on both sides of (27). Consequently, (28) holds and the metric regularity of M around $(0,0)$ is established.

Now we can apply [17, Theorem 4] to the set $\Omega' := T^{\rm lin}_{P,D}(\bar z) = \{w \,|\, \nabla P(\bar z)w \in T_D\big(P(\bar z)\big)\}$ to obtain

$$\widehat N_{\Omega'}(0) = \nabla P(\bar z)^T \widehat N_{T_D\left(P(\bar z)\right)}(0) = \nabla P(\bar z)^T \widehat N_D\big(P(\bar z)\big),$$

where the second equality follows from (9). Taking into account

$$\widehat N_{\Omega'}(0) = \big(T_{\Omega'}(0)\big)^\circ = (\Omega')^\circ = \big(T^{\rm lin}_{P,D}(\bar z)\big)^\circ,$$

the claimed equation $\big(T^{\rm lin}_{P,D}(\bar z)\big)^\circ = \nabla P(\bar z)^T \widehat N_D\big(P(\bar z)\big)$ follows. $\square$

Based on Definition 8(ii), Theorem 3 has the following easy corollary. The result improves [13, Theorem 1] in that GGCQ is assumed instead of the stronger condition MSCQ.

Corollary 1. *Assume that both GGCQ and* (27) *hold at* $\bar z \in \Omega$. *Then*

$$\widehat N_\Omega(\bar z) = \nabla P(\bar z)^T \widehat N_D\big(P(\bar z)\big).$$

Consequently, at any local minimal solution $\bar z$ *of problem (22), the S-stationary condition holds:*

$$0 \in \nabla f(\bar z) + \nabla P(\bar z)^T \widehat N_D\big(P(\bar z)\big).$$

Remark 1. *In the special case* $D = \mathbb{R}^s_-$ *we have* $T_D\big(P(\bar z)\big) = \{w \,|\, \nabla P_i(\bar z)w \le 0, i \in \mathcal{I}\big(P(\bar z)\big)\}$ *and* $\mathcal{L}\left(T_D\big(P(\bar z)\big)\right) = \{w \,|\, \nabla P_i(\bar z)w = 0,\ i \in \mathcal{I}\big(P(\bar z)\big)\}$, *where* $\mathcal{I}\big(P(\bar z)\big) := \{i \in \{1,\ldots,s\} \,|\, P_i(\bar z) = 0\}$ *denotes the index set of constraints active at* $P(\bar z)$, *and* $P(z) = \big(P_1(z),\ldots,P_s(z)\big)$. *Thus* (27) *is equivalent to the requirement that the gradients of the active constraints are linearly independent. For this reason we will refer to* (27) *as* Generalized Linear Independence Condition *(GLIC). Note that GLIC is not a constraint qualification because it does not imply* $\widehat N_\Omega(\bar z) = \big(T^{\rm lin}_{P,D}(\bar z)\big)^\circ$ *in general.*

Note that condition (27) may be restrictive. If it does not hold, but MSCQ holds at $\bar z$. Then it is well-known that $\widehat N_\Omega(\bar z) \subseteq N_\Omega(\bar z) \subseteq \nabla P(\bar z)^T N_D\big(P(\bar z)$ and hence the M-stationary condition

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

holds at any local optimal solution $\bar z$. For the case where set D is simple, e.g., $D := \{(a,b) \,|\, 0 \le a \perp b \ge 0\}$, the complementarity cone, the limiting normal cone can be calculated using the variational analysis (cf. [22]) and one obtains the classical M-stationary condition for MPCC. However, for more complicated set D, usually very strong assumptions are required for using these calculus rules limiting considerably their applicability.

Recently an alternative approach is taken by Gfrerer in [13]. Under GGCQ, by (26) for every regular normal $z^* \in \widehat N_\Omega(\bar z)$ the point $u = 0$ is a global minimizer for the problem

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

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

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

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

$$0 \in \nabla f(\bar z) + \nabla P(\bar z)^T N_{T_D\left(P(\bar z)\right)}(0), \tag{30}$$

which is shaper than the M-stationarity condition since $N_{T_D\left(P(\bar z)\right)}(0) \subseteq N_D\big(P(\bar z)\big)$, cf. [25, Proposition 6.27(a)]. Although (30) is a sharper condition than the M-stationary condition, it still involves the limiting normal cone and so may be hard to calculate. In [13, Propositions 1,2], Gfrerer derived the following new necessary condition which can be considered as a refinement of the necessary optimality condition (30). 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 $T_D\big(P(\bar z)\big)$ is locally polyhedral at 0, condition (31) coincides with the condition (30).

Theorem 4. *Let $\bar z$ be a local optimal solution for problem* (22). *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

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

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

$$\nabla P(\bar z)\bar u \notin \mathcal{L}(T_D(P(\bar z))), \tag{32}$$

$$\nabla f(\bar z)^T \bar u = 0, \tag{33}$$

$$0 \in \nabla f(\bar z) + \widehat{N}_{T^{\rm lin}_{P,D}(\bar z)}(\bar u) \tag{34}$$

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

$$\|t\| \le C\|v\| \ \forall (v,t) \in \operatorname{gph} M_c \tag{35}$$

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

$$\nabla P(\bar z)\bar u \in \{\bar v\} \times M(\bar v). \tag{36}$$

Remark 2. *(i) Note that the assumptions of Theorem 4 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 [13, Lemmas 4].*

(ii) Note that when $T_D(P(\bar z))$ is locally polyhedral near $\nabla P(\bar z)\bar u$, then the condition (i) holds. Otherwise if $T_D(P(\bar z))$ is not locally polyhedral near $\nabla P(\bar z)\bar u$, then the condition (ii) can hold. In this case, (33) *implies together with GGCQ and the basic optimality condition (3) that $\bar u$ is a global solution of the problem*

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

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 [13, 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 4 once more to the above problem. If $T_{T_D(P(\bar{z}))}(\nabla P(\bar{z})\bar{u})$ is polyhedral, then Theorem 4(i) applies and we obtain the existence of $\omega \in T_{T_D(P(\bar{z}))}(\nabla P(\bar{z})\bar{u})$ such that*

$$0 \in \nabla f(\bar{z}) + \nabla P(\bar{z})^T \widehat{N}_{T_{T_D(P(\bar{z}))}(\nabla P(\bar{z})\bar{u})}(\omega). \tag{37}$$

In this case (37) would be the necessary optimality condition which is sharper than condition (31). In this paper we aim at finding some sufficient conditions under which the tangent cone $T_{T_D(P(\bar{z}))}(\nabla P(\bar{z})\bar{u})$ would be polyhedral and hence the above optimality condition holds for (MPEC). However if $T_{T_D(P(\bar{z}))}(\nabla P(\bar{z})\bar{u})$ is not polyhedral, then the process can continue. The interested reader is referred to [13] for the discussion for what might have happened after applying Theorem 4 repeatedly.

4 New optimality conditions for MPEC

In this section, we apply Theorem 4 to obtain a necessary optimality condition for the MPEC (3). Note that (3) can be written in the form (22) via

$$\begin{aligned} \min_{x,y} \quad & F(x,y) \\ \text{subject to} \quad & P(x,y) := \begin{pmatrix} (y, -\phi(x,y)) \\ G(x,y) \end{pmatrix} \in D := \operatorname{gph} \widehat{N}_\Gamma \times \mathbb{R}^p_-, \end{aligned} \tag{38}$$

where $\Gamma := \{y \,|\, g(y) \le 0\}$. To apply Theorem 4, we will need the following assumptions at a local solution $(\bar{x}, \bar{y})$ to problem (38).

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

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

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

Note that by Remark 2 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 [18, Theorem 5]. To apply Theorem 4, we will need to compute the tangent cone $T_D((\bar{y},\bar{y}^*), G(\bar{x},\bar{y}))$. By [18, Proposition 1] we have

$$T_{\operatorname{gph} \widehat{N}_\Gamma \times \mathbb{R}^p_-}((\bar{y},\bar{y}^*), G(\bar{x},\bar{y})) = T_{\operatorname{gph} \widehat{N}_\Gamma}(\bar{y},\bar{y}^*) \times T_{\mathbb{R}^p_-}(G(\bar{x},\bar{y})) \tag{39}$$

and hence we only need to describe how to compute $T_{\operatorname{gph} \widehat{N}_\Gamma}(\bar{y},\bar{y}^*)$. For this purpose, we need some notation. We denote by the critical cone for Γ at $(\bar{y},\bar{y}^*)$ as

$$\bar{K}_\Gamma := \{v \,|\, \nabla g(\bar{y})v \in T_{\mathbb{R}^q_-}(g(\bar{y}))\} \cap [\bar{y}^*]^\perp, \tag{40}$$

which is convex polyhedral. Indeed, 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 hence $\bar{K}_\Gamma$ is indeed equal to $\mathcal{K}_\Gamma(\bar{y},\bar{y}^*)$, the critical cone for Γ at $(\bar{y}, \bar{y}^*)$. Further we define the *multiplier set* for $\bar{y}$ as the polyhedral convex set defined by

$$\bar{\Lambda} := \{\lambda \in N_{\mathbb{R}^q_-}(g(\bar{y})) \,|\, \nabla g(\bar{y})^T \lambda = \bar{y}^*\}. \tag{41}$$

Indeed, the multiplier set is the set of all Lagrange multipliers for the lower level system (1) when $(x, y) = (\bar{x}, \bar{y})$. Moreover for every $v \in \bar{K}_\Gamma$, we define the *directional multiplier set* as

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

which is also a polyhedral convex set. By [15, 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,

$$\lambda \in \bar{\Lambda}(v) \Longleftrightarrow v^T \nabla^2 g(\bar{y}) v \in N_{\bar{\Lambda}}(\lambda). \tag{42}$$

Theorem 5 (cf. [18, Theorem 4])**.** *Let $\bar{y} \in \Gamma := \{y \,|\, g(y) \le 0\}$ 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*

$$\begin{aligned} & T_{\operatorname{gph} \widehat{N}_\Gamma}(\bar{y}, \bar{y}^*) \\ = \;& \left\{ (v, v^*) \in \mathbb{R}^{2m} \,|\, \exists\, \lambda \in \bar{\Lambda}(v) \;\; \textit{with} \;\; v^* \in \nabla^2 (\lambda^T g)(\bar{y}) v + N_{\bar{K}_\Gamma}(v) \right\} \\ = \;& \left\{ (v, v^*) \in \mathbb{R}^{2m} \,|\, \exists\, \lambda \in \bar{\Lambda}(v) \cap \kappa \|\bar{y}^*\| \mathcal{B}_{\mathbb{R}^q} \;\; \textit{with} \;\; v^* \in \nabla^2 (\lambda^T g)(\bar{y}) v + N_{\bar{K}_\Gamma}(v) \right\}, \end{aligned} \tag{43}$$

where $\kappa > 0$ is certain constant.

By taking into account (42) we obtain the alternative representation

$$\begin{aligned} & T_{\operatorname{gph} \widehat{N}_\Gamma}(\bar{y}, \bar{y}^*) \\ = \;& \left\{ (v, \nabla^2 (\lambda^T g)(y) v + z^*) \in \mathbb{R}^{2m} \,|\, (z^*, v^T \nabla^2 g(\bar{y}) v) \in N_{\bar{K}_\Gamma \times \bar{\Lambda}}(v, \lambda) \right\} \\ = \;& \left\{ (v, \nabla^2 (\lambda^T g)(y) v + z^*) \in \mathbb{R}^{2m} \,|\, (z^*, v^T \nabla^2 g(\bar{y}) v) \in N_{\bar{K}_\Gamma \times \bar{\Lambda}}(v, \lambda), \; \|\lambda\| \le \kappa \|\bar{y}^*\| \right\}. \end{aligned} \tag{44}$$

In order to apply Theorem 4, for any $(\bar{v}, \bar{v}^*, a) := w \in T_D(P(\bar{x}, \bar{y}))$, we need to compute $\widehat{N}_{T_D(P(\bar{x}, \bar{y}))}(w)$. But by (39), we have

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

Hence we will compute $\widehat{N}_{T_{\operatorname{gph} \widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)}(\bar{v}, \bar{v}^*)$. Similarly by virtue of (37) we also need to compute $\widehat{N}_{T_{T_{\operatorname{gph} \widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)}(\bar{v}, \bar{v}^*)}(\delta\bar{v}, \delta\bar{v}^*)$.

4.1 Regular normals to tangent directions

Throughout this subsection let $(\bar{v}, \bar{v}^*) \in T_{\operatorname{gph} \widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)$ be given. We first recall the following result.

Proposition 5 ([13, Proposition 3])**.** *If there is an open neighborhood V of $\bar{v}$ and a set $\tilde{\Lambda} \subset \bar{\Lambda}$ such that*

$$\bar{\Lambda}(v) = \tilde{\Lambda} \quad \forall v \in (\bar{K}_\Gamma \setminus \{\bar{v}\}) \cap V, \tag{45}$$

then

$$T_{\operatorname{gph}\widehat{N}_\Gamma}(\bar y,\bar y^*)\cap(V\times\mathbb{R}^m)=\big\{\big(v,\nabla^2(\tilde\lambda^Tg)(\bar y)v+z^*\big)\,|\,z^*\in N_{\bar K_\Gamma}(v)\big\}\cap(V\times\mathbb{R}^m),\tag{46}$$

where $\tilde\lambda\in\tilde\Lambda$ is an arbitrarily fixed multiplier. In particular, $T_{\operatorname{gph}\widehat{N}_\Gamma}(\bar y,\bar y^*)$ is locally polyhedral near $(\bar v,\bar v^*)$ for every $\bar v^*$ satisfying $(\bar v,\bar v^*)\in T_{\operatorname{gph}\widehat{N}_\Gamma}(\bar y,\bar y^*)$ and

$$\widehat{N}_{T_{\operatorname{gph}\widehat{N}_\Gamma}(\bar y,\bar y^*)}(\bar v,\bar v^*)=\big\{(w^*,w)\,|\,(w^*+\nabla^2(\tilde\lambda^Tg)(\bar y)w,w)\in\big(\mathcal{K}_{\bar K_\Gamma}(\bar v,\bar z^*)\big)^\circ\times\mathcal{K}_{\bar K_\Gamma}(\bar v,\bar z^*)\big\},\tag{47}$$

where $\bar z^*:=\bar v^*-\nabla^2(\tilde\lambda^Tg)(\bar y)\bar v$.

Consider the following definition:

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*

$$\nabla^2(\mu^Tg)(\bar y)\bar v\in(N_{\bar K_\Gamma}(\bar v))^+,\ \mu\in\big(\bar\Lambda(\bar v)\big)^+\implies\mu=0.$$

In the case where the directional multiplier set $\bar\Lambda(\bar v)$ is a singleton, $\big(\bar\Lambda(\bar v)\big)^+=\{0\}$ and hence g is 2-nondegenerate in this direction $\bar v$. In particular, if $\bar\Lambda$ is a singleton then g is non-degenerate in any direction $\bar v$. The concept of 2-nongeneracy is discussed more in detail at the end of section 4.3.

Lemma 2. *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*

$$\mathcal{H}(\bar v):=\{\nabla^2(\mu^Tg)(\bar y)\bar v\,|\,\mu\in\big(\bar\Lambda(\bar v)\big)^+\}+(N_{\bar K_\Gamma}(\bar v))^+.\tag{48}$$

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

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

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

$$\bar v^*=\nabla^2(\bar\lambda^Tg)(\bar y)\bar v+\bar z^*.\tag{49}$$

Proof. By the definition, the mapping $\mathcal{A}_{\bar v}$ is surjective and therefore we only have to show injectivity. Consider elements $(\mu,\bar z^*)\in\big(\bar\Lambda(\bar v)\big)^+\times(N_{\bar K_\Gamma}(\bar v))^+$ satisfying $\mathcal{A}_{\bar v}(\mu,\bar z^*)=0$. Then $\nabla^2(\mu^Tg)(\bar y)\bar v=-\bar z^*\in(N_{\bar K_\Gamma}(\bar v))^+$ 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 $(\bar v,\bar v^*)\in T_{\operatorname{gph}\widehat{N}_\Gamma}(\bar y,\bar y^*)$. The existence of $(\bar\lambda,\bar z^*)\in\bar\Lambda(\bar v)\times N_{\bar K_\Gamma}(\bar v)$ fulfilling (49) follows from Theorem 5. In order to prove the uniqueness of the representation (49), consider $(\lambda_1,z_1^*),(\lambda_2,z_2^*)\in\bar\Lambda(\bar v)\times N_{\bar K_\Gamma}(\bar v)$ such that

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

implying

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

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 $z_2^*=z_1^*$. $\square$

The following lemma will we useful for our analysis:

Lemma 3. *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 $\tilde{P}(\bar{z}) \in \tilde{D} := \operatorname{gph} N_C$ be given. Further assume that*

$$\ker \nabla \tilde{P}(\bar{z})^T \cap (L_1 \times L_2) = \{(0,0)\}, \tag{50}$$

where $L_1 \supseteq \big(N_C(\tilde{P}_1(\bar{z}))\big)^+$ and $L_2 \supseteq \big(\mathcal{K}_C(\tilde{P}_1(\bar{z}), \tilde{P}_2(\bar{z}))\big)^+$. Then the mapping $z \rightrightarrows \tilde{P}(z) - \tilde{D}$ is metrically regular around $(\bar{z}, 0)$,

$$T_{\{z \,|\, \tilde{P}(z) \in \tilde{D}\}}(\bar{z}) = \{w \,|\, \nabla \tilde{P}(\bar{z}) w \in T_{\tilde{D}}(\tilde{P}(\bar{z}))\} \tag{51}$$

and

$$\begin{aligned} \widehat{N}_{\{z \,|\, \tilde{P}(z) \in \tilde{D}\}}(\bar{z}) &= \nabla \tilde{P}(\bar{z})^T \widehat{N}_{\tilde{D}}(\tilde{P}(\bar{z})) \\ &= \nabla \tilde{P}_1(\bar{z})^T \big(\mathcal{K}_C(\tilde{P}_1(\bar{z}), \tilde{P}_2(\bar{z}))\big)^\circ + \nabla \tilde{P}_2(\bar{z})^T \mathcal{K}_C(\tilde{P}_1(\bar{z}), \tilde{P}_2(\bar{z})). \end{aligned} \tag{52}$$

Proof. In order to prove metric regularity of the mapping $\tilde{P}(\cdot) - \tilde{D}$ we invoke the Mordukhovich criterion (7), which reads in our case as

$$\nabla \tilde{P}_1(\bar{z})^T w^* + \nabla \tilde{P}_2(\bar{z})^T w = 0, \; (w^*, w) \in N_{\operatorname{gph} N_C}(\tilde{P}(\bar{z})) \;\Rightarrow\; (w^*, w) = 0. \tag{53}$$

Consider $(w^*, w) \in N_{\operatorname{gph} N_C}(\tilde{P}(\bar{z}))$. 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 (16) and (17). Thus (53) follows from (50) 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 (51) follows from [20, Proposition 1]. In order to show (52) we will invoke Theorem 3. From (20) 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 (50) we obtain

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

where the second equality follows from Proposition 2. Hence the assumption of Theorem 3 is fulfilled and the first equation in (52) follows, whereas the second equation is a consequence of (13). □

Proposition 6. *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 $(\bar{v}, \bar{v}^*) \in T_{\operatorname{gph} \widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)$ we have*

$$T_{T_{\operatorname{gph} \widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)}(\bar{v}, \bar{v}^*) = \left\{ (u, u^*) \,|\, \exists \mu, \zeta^* \; s.t. \; \begin{array}{l} u^* = \nabla^2(\bar{\lambda}^T g)(\bar{y}) u + \nabla^2(\mu^T g)(\bar{y}) \bar{v} + \zeta^*, \\ (u, \mu, \zeta^*, 2\bar{v}^T \nabla^2 g(\bar{y}) u) \in \operatorname{gph} N_{\tilde{K}(\bar{v}, \bar{v}^*)} \end{array} \right\}, \tag{54}$$

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*

$$\tilde K(\bar v, \bar v^*) := \mathcal{K}_{\bar K_\Gamma \times \bar\Lambda}(\bar v, \bar\lambda, \bar z^*, \bar v^T \nabla^2 g(\bar y)\bar v). \tag{55}$$

Further,

$$\begin{aligned}&\widehat N_{T_{\operatorname{gph}\widehat N_\Gamma}(\bar y, \bar y^*)}(\bar v, \bar v^*)\\ &= \left\{(w^*, w) \mid \begin{array}{c}\exists \eta \text{ s.t. } \big(w^* + \nabla^2(\bar\lambda^T g)(\bar y)w - 2\nabla^2(\eta^T g)(\bar y)\bar v, \bar v^T\nabla^2 g(\bar y)w, w, \eta\big)\\ \in \big(\tilde K(\bar v, \bar v^*)\big)^\circ \times \tilde K(\bar v, \bar v^*)\end{array}\right\}. \end{aligned}\tag{56}$$

Proof. Let $\bar v^*$ with $(\bar v, \bar v^*) \in T_{\operatorname{gph}\widehat N_\Gamma}(\bar y, \bar y^*)$ be fixed and let $\mathcal{R}$ denote the set on the right hand side of equation (54). In a first step we will show that $T_{T_{\operatorname{gph}\widehat N_\Gamma}(\bar y, \bar y^*)}(\bar v, \bar v^*) \subseteq \mathcal{R}$. Let $(u, u^*) \in T_{T_{\operatorname{gph}\widehat N_\Gamma}(\bar y, \bar y^*)}(\bar v, \bar v^*)$ and consider sequences $t_k \searrow 0$, $(u_k, u_k^*) \to (u, u^*)$ with $(\bar v + t_k u_k, \bar v^* + t_k u_k^*) \in T_{\operatorname{gph}\widehat N_\Gamma}(\bar y, \bar y^*)$. By (43) 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 $\bar v^* = \nabla^2(\bar\lambda^T g)(\bar y)\bar v + \bar z^*$, it follows that

$$\nabla^2((\lambda_k - \bar\lambda)^T g)(\bar y)\bar v + z_k^* - \bar z^* = t_k\big(u_k^* - \nabla^2(\lambda_k^T g)(\bar y)u_k\big). \tag{57}$$

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 (10) and $\bar\Lambda(\bar v + t_k u_k) \subseteq \bar\Lambda(\bar v)$ by [17, Lemma 3] implying $\lambda_k - \bar\lambda \in \big(\bar\Lambda(\bar v)\big)^+$ and $z_k^* - \bar z^*$ in $(N_{\bar K_\Gamma}(\bar v))^+$. Thus from (57), we have

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

By the boundedness of λ_k we conclude $t_k\big(u_k^* - \nabla^2(\lambda_k^T g)(\bar y)u_k\big) \to 0$. Hence, by Lemma 2 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

$$u^* = \nabla^2(\bar\lambda^T g)(\bar y)u + \mathcal{A}_{\bar v}(\mu, \zeta^*) = \nabla^2(\bar\lambda^T g)(\bar y)u + \nabla^2(\mu^T g)(\bar y)\bar v + \zeta^*. \tag{58}$$

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 (42), 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 (12). Thus combining the above inclusion and (58), we have that $(u, u^*) \in \mathcal{R}$ and the inclusion $T_{T_{\operatorname{gph}\widehat N_\Gamma}(\bar y, \bar y^*)}(\bar v, \bar v^*) \subseteq \mathcal{R}$ is shown.

Now we show the reverse inclusion in (54). First by applying Lemma 3 , we wish to show that

$$\begin{aligned}\Theta &:= T_{\{(v,\lambda,z^*)\,|\,\tilde P(v,\lambda,z^*)\in\tilde D\}}(\bar v,\bar\lambda,\bar z^*) = \{(u,\mu,\zeta^*)\,|\,\nabla\tilde P(\bar v,\bar\lambda,\bar z^*)(u,\mu,\zeta^*)\in T_{\tilde D}(\tilde P(\bar v,\bar\lambda,\bar z^*))\}\\ &= \{(u,\mu,\zeta^*)\,|\,(u,\mu,\zeta^*,2\bar v^T\nabla^2 g(\bar y)u)\in T_{\mathrm{gph}\,N_{\bar K_\Gamma\times\bar\Lambda}}(\bar v,\bar\lambda,\bar z^*,\bar v^T\nabla^2 g(\bar y)\bar v)\},\end{aligned} \tag{59}$$

where $\tilde P_1(v,\lambda,z^*) := (v,\lambda)$, $\tilde P_2(v,\lambda,z^*) := (z^*, v^T\nabla^2 g(\bar y)v)$, $\tilde D := \mathrm{gph}\,N_{\bar K_\Gamma\times\bar\Lambda}$. Let $\bar z := (\bar v,\bar\lambda,\bar z^*)$. Define $L_1 := \big(N_{\bar K_\Gamma}(\bar v)\big)^+\times\mathbb{R}^q \supseteq \big(N_{\bar K_\Gamma\times\bar\Lambda}(\tilde P_1(\bar z))\big)^+$ and $L_2 := \mathbb{R}^m\times\big(\bar\Lambda(\bar v)\big)^+$. Since

$$\mathcal{K}_{\bar\Lambda}(\bar\lambda,\bar v^T\nabla^2 g(\bar y)\bar v) = T_{\bar\Lambda}(\bar\lambda)\cap[\bar v^T\nabla^2 g(\bar y)\bar v]^\perp = T_{\bar\Lambda(\bar v)}(\bar\lambda), \tag{60}$$

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 $\alpha\ge 0$ sufficiently small, we have that

$$L_2 := \mathbb{R}^m\times\big(\bar\Lambda(\bar v)\big)^+ \supseteq \big(\mathcal{K}_{\bar K_\Gamma\times\bar\Lambda}(\tilde P_1(\bar z),\tilde P_2(\bar z))\big)^+.$$

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 $\nabla^2(\mu^T g)(\bar y)\bar v = -w^*/2\in\big(N_{\bar K_\Gamma}(\bar v)\big)^+$ and by the assumed 2-nondegeneracy we obtain $\mu=0$ and consequently $w^*=0$. Because we also have $\lambda^*=0$ and $z=0$, (50) is verified and by (51) we obtain (59).

Now let $(u,u^*)\in\mathcal{R}$. Then there exist μ,ζ^* such that

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

where the second equation follows from (12). It follows from (59) and the definition of the tangent cone that there exist sequences $t_k\searrow 0$ and $(u_k,\mu_k,\zeta_k^*)\to(u,\mu,\zeta^*)$ such that

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

By (44) it follows that $(\bar v+t_ku_k,\nabla^2\big((\bar\lambda+t_k\mu_k)^T g\big)(\bar y)(\bar v+t_ku_k)+\bar z^*+t_k\zeta_k^*)\in T_{\mathrm{gph}\,\widehat N_\Gamma}(\bar y,\bar y^*)$ implying $(u,u^*)\in T_{T_{\mathrm{gph}\,\widehat N_\Gamma}(\bar y,\bar y^*)}(\bar v,\bar v^*)$. Hence (54) is shown.

To show (56), note that

$$\begin{aligned}&\widehat N_{T_{\mathrm{gph}\,\widehat N_\Gamma}(\bar y,\bar y^*)}(\bar v,\bar v^*)\\ &= (T_{T_{\mathrm{gph}\,\widehat N_\Gamma}(\bar y,\bar y^*)}(\bar v,\bar v^*))^\circ\\ &= (\{(u,u^*) : u^* = \nabla^2(\bar\lambda^T g)(\bar y)u+\nabla^2(\mu^T g)(\bar y)\bar v+\zeta^*,(u,\mu,\zeta^*)\in\Theta\})^\circ\\ &= \{(w^*,w)\,|\,\langle w^*,u\rangle+\langle w,\nabla^2(\bar\lambda^T g)(\bar y)u+\nabla^2(\mu^T g)(\bar y)\bar v+\zeta^*\rangle\le 0\quad\forall(u,\mu,\zeta^*)\in\Theta\}\\ &= \{(w^*,w)\,|\,(w^*+\nabla^2(\bar\lambda^T g)(\bar y)w,\bar v^T\nabla^2 g(\bar y)w,w)\in\Theta^\circ\},\end{aligned} \tag{61}$$

where the second equality follows from (54). By (52) together with (13) we have

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

and (56) follows from (61). □

Unless $\bar{\Lambda}$ is a singleton, g can not be 2-nondegenerate in direction $\bar{v} = 0$. Hence, Proposition 6 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 [17, Lemma 3] 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 7. *Let $\bar{v}^* \in \bar{K}_\Gamma^\circ$. Then*

$$\begin{aligned} &\widehat{N}_{T_{\mathrm{gph}\, \widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)}(0, \bar{v}^*) && (62) \\ &\subseteq \bigcap_{\bar{v} \in \mathcal{L}(\bar{K}_\Gamma)} \Big\{ (w^*, w) \,|\, \exists \bar{\lambda} \in \operatorname{conv} \Sigma(\bar{v}, \bar{v}^*) : \ (w^* + \nabla^2(\bar{\lambda}^T g)(\bar{y}) w, w) \in \big(\mathcal{K}_{\bar{K}_\Gamma}(0, \bar{v}^*)\big)^\circ \times \mathcal{K}_{\bar{K}_\Gamma}(0, \bar{v}^*) \Big\}. \end{aligned}$$

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

$$\widehat{N}_{T_{\mathrm{gph}\, \widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)}(0, \bar{v}^*) \subseteq \bigcap_{\bar{\lambda} \in \bar{\Lambda}(\bar{v})} \widehat{N}_{T_{\mathrm{gph}\, \widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)}(\bar{v}, \nabla^2(\bar{\lambda}^T g)(\bar{y}) \bar{v} + \bar{v}^*). \tag{63}$$

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 (43) with $v = 0$, we have $(0, \bar{v}^* + \alpha z^*) \in T_{\mathrm{gph}\, \widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)$ and thus $\langle w^*, 0 \rangle + \langle w, \bar{v}^* + \alpha z^* - \bar{v}^* \rangle \le 0$ 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 $\bar{\lambda} \in \operatorname{conv} \Sigma(\bar{v}, \bar{v}^*)$ such that

$$w^* + \nabla^2(\bar{\lambda}^T g)(\bar{y}) w \in \big(\mathcal{K}_{\bar{K}_\Gamma}(0, \bar{v}^*)\big)^\circ. \tag{64}$$

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 (42),

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

Therefore by (44), $(\pm\alpha\bar{v}, \pm\alpha\nabla^2(\lambda^T g)(\bar{y})\bar{v} + \bar{v}^*) \in T_{\mathrm{gph}\, \widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)$, $\forall \alpha > 0$ sufficiently small, $\forall \lambda \in \bar{\Lambda}(\bar{v})$. By the definition of the regular normal cone we conclude

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

and therefore

$$\langle w^*, \bar{v} \rangle + \langle w, \nabla^2(\lambda^T g)(\bar{y})\bar{v} \rangle = 0 \quad \forall \lambda \in \bar{\Lambda}(\bar{v}). \tag{65}$$

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 (43) we have

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

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

$$\begin{aligned}\limsup_{\alpha\searrow 0} &\frac{\langle w^*, \alpha(\bar{v} + \beta u)\rangle + \langle w, \alpha\nabla^2(\lambda^T g)(\bar{y})(\bar{v} + \beta u) + \bar{v}^* - \bar{v}^*\rangle}{\alpha}\\ &= \beta\big(\langle w^*, u\rangle + \langle w, \nabla^2(\lambda^T g)(\bar{y})u\rangle\big) \le 0,\end{aligned}$$

where the equality follows from (65). 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 (64) holds. Putting all together, (62) follows.

Let $\bar{v} \in \mathcal{L}(\bar{K}_\Gamma)$. We now show (63) 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 6 there are elements μ, ζ^* such that

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

By taking into account (59), there are sequences $t_k \searrow 0$, $(u_k, \mu_k, \zeta_k^*) \to (u, \mu, \zeta^*)$ 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 (42), $(\bar{v} + t_k u_k)^T\nabla^2 g(\bar{y})(\bar{v} + t_k u_k) \in N_{\bar{\Lambda}}(\bar{\lambda} + t_k\mu_k)$ 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 $\big(\alpha(\bar{v} + t_k u_k), \alpha\nabla^2\big((\bar{\lambda} +$

$t_k\mu_k)^Tg\big)g(\bar y)(\bar v+t_ku_k)+\bar v^*+\alpha t_k^*\zeta_k^*\Big)\in T_{\mathrm{gph}\,\widehat N_\Gamma}(\bar y,\bar y^*)$ by (43), implying

$$\begin{aligned}0 &\geq \lim_{\alpha\searrow 0}\frac{\langle w^*,\alpha(\bar v+t_ku_k)\rangle+\langle w,\alpha\nabla^2\big((\bar\lambda+t_k\mu_k)^Tg\big)(\bar y)(\bar v+t_ku_k)+\bar v^*+\alpha t_k\zeta_k^*-\bar v^*\rangle}{\alpha}\\ &= \langle w^*,(\bar v+t_ku_k)\rangle+\langle w,\nabla^2\big((\bar\lambda+t_k\mu_k)^Tg\big)(\bar y)(\bar v+t_ku_k)+t_k\zeta_k^*\rangle\\ &= t_k\big(\langle w^*,u_k\rangle+\langle w,\nabla^2\big((\bar\lambda+t_k\mu_k)^Tg\big)(\bar y)u_k+\nabla^2(\mu_k^Tg)(\bar y)\bar v+\zeta_k^*\rangle\big)\end{aligned}$$

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

$$0\geq\langle w^*,u\rangle+\langle w,\nabla^2(\bar\lambda^Tg)(\bar y)u+\nabla^2(\mu^Tg)(\bar y)\bar v+\zeta^*\rangle=\langle w^*,u\rangle+\langle w,u^*\rangle.$$

Thus $(w^*,w)\in(T_{T_{\mathrm{gph}\,\widehat N_\Gamma}(\bar y,\bar y^*)}(\bar v,\nabla(\bar\lambda^Tg)(\bar y)\bar v+\bar v^*))^\circ$ and the inclusion (63) follows. □

4.2 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 (49), i.e., $\bar v^*=\nabla^2(\bar\lambda^Tg)(\bar y)\bar v+\bar z^*$, and let according to (54) $(\bar\mu,\bar\zeta^*)$ denote some element with

$$\delta\bar v^*=\nabla^2(\bar\lambda^Tg)(\bar y)\delta\bar v+\nabla^2(\bar\mu^Tg)(\bar y)\bar v+\bar\zeta^*,\tag{66}$$

$$(\delta\bar v,\bar\mu,\bar\zeta^*,2\bar v^T\nabla^2g(\bar y)\delta\bar v)\in\mathrm{gph}\,N_{\tilde K(\bar v,\bar v^*)}=T_{\mathrm{gph}\,N_{\bar K_\Gamma\times\bar\Lambda}}(\bar v,\bar\lambda,\bar z^*,\bar v^T\nabla^2g(\bar y)\bar v),\tag{67}$$

where the equality in (67) follows from (12). Note that by definition,

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

and hence it follows that $\bar\mu\in\mathcal K_{\bar\Lambda}(\bar\lambda,\bar v^T\nabla^2g(\bar y)\bar v)=T_{\bar\Lambda(v)}(\bar\lambda)$ where the equality follows from (60), $\bar\zeta^*\in N_{\mathcal K_{\bar K_\Gamma}(\bar v,\bar z^*)}(\delta\bar v)\subseteq(\mathcal K_{\bar K_\Gamma}(\bar v,\bar z^*))^\circ\subseteq\big(N_{\bar K_\Gamma}(\bar v)\big)^+$. By (66), $\mathcal A_{\bar v}(\bar\mu,\bar\zeta^*)=\delta\bar v^*-\nabla^2(\bar\lambda^Tg)(\bar y)\delta\bar v$ and from Lemma 2 we conclude that $(\bar\mu,\bar\zeta^*)$ are unique.

Proposition 8. *Under the assumption stated in the beginning of this subsection, we have*

$$\begin{aligned}&T_{T_{T_{\mathrm{gph}\,\widehat N_\Gamma}(\bar y,\bar y^*)}(\bar v,\bar v^*)}(\delta\bar v,\delta\bar v^*)\\ &\quad=\left\{(u,u^*)\,|\,\exists\delta\mu,\delta\zeta^*:\ \begin{array}{l}u^*=\nabla^2(\bar\lambda^Tg)(\bar y)u+\nabla^2(\delta\mu^Tg)(\bar y)\bar v+\delta\zeta^*)\\(u,\delta\mu,\delta\zeta^*,2\bar v^T\nabla^2g(\bar y)u)\in\mathrm{gph}\,N_{\tilde K(\bar v,\bar v^*,\delta\bar v,\delta\bar v^*)}\end{array}\right\}\end{aligned}\tag{68}$$

and

$$\begin{aligned}&\widehat N_{T_{T_{\mathrm{gph}\,\widehat N_\Gamma}(\bar y,\bar y^*)}(\bar v,\bar v^*)}(\delta\bar v,\delta\bar v^*)\\ &\quad=\left\{(w^*,w)\,|\,\exists\eta:\ \begin{array}{l}\big(w^*+\nabla^2(\bar\lambda^Tg)(\bar y)w-2\nabla^2(\eta^Tg)(\bar y)\bar v,\bar v^T\nabla^2g(\bar y)w,w,\eta\big)\\ \in\big(\tilde K(\bar v,\bar v^*,\delta\bar v,\delta\bar v^*)\big)^\circ\times\tilde K(\bar v,\bar v^*,\delta\bar v,\delta\bar v^*)\end{array}\right\},\end{aligned}\tag{69}$$

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

Proof. We use similar arguments as in the proof of Proposition 6. Let $\mathcal{R}$ denote the set on the right hand side of (68) and consider $(u,u^*) \in T_{T_{\mathrm{gph}\,\widehat{N}_\Gamma(\bar y,\bar y^*)}(\bar v,\bar v^*)}(\delta\bar v,\delta\bar v^*)$ 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 6 there are elements μ_k, ζ_k^* such that

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

where the equality in the second inclusion follows from (12). By taking into account (66) we obtain after rearranging

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

Similarly as shown in the paragraph before Proposition 8, 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 $\mu_k - \bar\mu \in \big(\bar\Lambda(\bar v)\big)^+$. Further, $\bar\zeta^* \in N_{\mathcal{K}_{\bar K_\Gamma}(\bar v,\bar z^*)}(\delta\bar v) \subseteq \big(N_{\bar K_\Gamma}(\bar v)\big)^+$ and $\zeta_k^* \in N_{\mathcal{K}_{\bar K_\Gamma}(\bar v,\bar z^*)}(\delta\bar v + t_k u_k) \subseteq \big(N_{\bar K_\Gamma}(\bar v)\big)^+$ implying $\zeta_k^* - \bar\zeta^* \in \big(N_{\bar K_\Gamma}(\bar v)\big)^+$. Thus, by Lemma 2 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 (68). Let $(u,u^*) \in \mathcal{R}$. Then there exist $\delta\mu$ and $\delta\zeta^*$ such that

$$\begin{aligned}u^* &= \nabla^2(\bar\lambda^T g)(\bar y)u + \nabla^2(\delta\mu^T g)(\bar y)\bar v + \delta\zeta^*, &(70)\\ (u,&\delta\mu,\delta\zeta^*,2\bar v^T\nabla^2 g(\bar y)u) \in \mathrm{gph}\,N_{\tilde K(\bar v,\bar v^*,\delta\bar v,\delta\bar v^*)}\\ &= T_{\mathrm{gph}\,N_{\tilde K(\bar v,\bar v^*)}}(\delta\bar v,\bar\mu,\bar\zeta^*,2\bar v^T\nabla^2 g(\bar y)\delta\bar v), &(71)\end{aligned}$$

where the equality in the second inclusion follows from (12) 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 (71) 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 (54) and taking into account (66) and (70), 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 (68) is proven.

In order to show (69) note that by (68), $\widehat{N}_{T_{T_{\mathrm{gph}\,\widehat{N}_\Gamma}(\bar y,\bar y^*)}(\bar v,\bar v^*)}(\delta\bar v,\delta\bar v^*)$ is the collection of all (w^*,w) fulfilling

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

for all

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

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 3 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 6 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 6 at $(\bar{v}, \bar{\lambda}, \bar{z}^*)$. Further, from (20) together with (21) and the definition of $\tilde{K}(\bar{v}, \bar{v}^*)$ we obtain

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

Applying (20) 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 3 to obtain

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

where the second equality follows from (12), and hence (69) follows. □

4.3 The new optimality condition

To establish the main optimality condition in Theorem 6, we first apply Theorem 4 to the problem (3) to obtain the following lemma.

Lemma 4. *Assume that $(\bar{x}, \bar{y})$ is a local minimizer for the problem (3) 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*

$$(\delta\bar{v}, \delta\bar{v}^*) \in T_{T_{\operatorname{gph} \widehat{N}_\Gamma}(\bar{y},\bar{y}^*)}\big(\delta y, -\nabla\phi(\bar{x}, \bar{y})(\delta x, \delta y)\big), \quad \delta a \in T_{T_{\mathbb{R}^m_-}(G(\bar{x},\bar{y}))}\big(\nabla G(\bar{x}, \bar{y})(\delta x, \delta y)\big), \tag{72}$$

together with multipliers

$$(w^*, w) \in \widehat{N}_{T_{T_{\operatorname{gph} \widehat{N}_\Gamma}(\bar{y},\bar{y}^*)}(\delta y, -\nabla\phi(\bar{x},\bar{y})(\delta x,\delta y))}(\delta\bar{v}, \delta\bar{v}^*), \;\; \sigma \in N_{T_{T_{\mathbb{R}^m_-}(G(\bar{x},\bar{y}))}(\nabla G(\bar{x},\bar{y})(\delta x,\delta y))}(\delta a) \tag{73}$$

such that

$$\nabla F(\bar{x}, \bar{y})^T(\delta x, \delta y) = 0, \tag{74a}$$

$$\nabla_x F(\bar{x}, \bar{y}) - \nabla_x\phi(\bar{x}, \bar{y})^T w + \nabla_x G(\bar{x}, \bar{y})^T\sigma = 0, \tag{74b}$$

$$\nabla_y F(\bar{x}, \bar{y}) + w^* - \nabla_y\phi(\bar{x}, \bar{y})^T w + \nabla_y G(\bar{x}, \bar{y})^T\sigma = 0, \tag{74c}$$

$$\delta y = 0 \;\Rightarrow\; \delta x = 0, \tag{74d}$$

$$\delta y \neq 0 \;\Rightarrow\; T_{\operatorname{gph} \widehat{N}_\Gamma}(\bar{y}, \bar{y}^*) \text{ is not locally polyhedral near } \big(\delta y, -\nabla\phi(\bar{x}, \bar{y})(\delta x, \delta y)\big). \tag{74e}$$

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

$$(w^* + \nabla^2(\bar{\lambda}^T g)(\bar{y})w, w) \in \big(\mathcal{K}_{\bar{K}_\Gamma}(0, \delta\bar{v}^*)\big)^\circ \times \mathcal{K}_{\bar{K}_\Gamma}(0, \delta\bar{v}^*). \tag{75}$$

Proof. Let $z := (x,y), \bar{z} = (\bar{x}, \bar{y}), P(x,y) := \begin{pmatrix} (y, -\phi(x,y)) \\ G(x,y) \end{pmatrix}, D := \operatorname{gph} \widehat{N}_\Gamma \times \mathbb{R}^p_-$. Assumption 1 ensures that Theorem 4 is applicable and so one of Theorem 4(i) and Theorem 4(ii) holds.

If Theorem 4(i) is fulfilled, then there exists a direction $\omega = (\delta\bar{v}, \delta\bar{v}^*, \delta a) \in T_D(P(\bar{z})) = T_{\operatorname{gph} \widehat{N}_\Gamma}(\bar{y}, \bar{y}^*) \times T_{\mathbb{R}^m_-}(G(\bar{x}, \bar{y}))$ and a multiplier

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

such that $0 = \nabla F(\bar{x}, \bar{y}) + \nabla P(\bar{x}, \bar{y})^T \omega^*$. By virtue of (9), we see that the conditions (74a)-(74c) are fulfilled with $\delta x = 0$, $\delta y = 0$. Otherwise Theorem 4(ii) is fulfilled, i.e., there is a direction $\bar{u} = (\delta x, \delta y)$ with

$$\begin{aligned} \nabla P(\bar{z})\bar{u} &= \big(\delta y, -\nabla\phi(\bar{x}, \bar{y})(\delta x, \delta y), \nabla G(\bar{x}, \bar{y})(\delta x, \delta y)\big) \\ &\in T_D\big(P(\bar{z})\big) = T_{\operatorname{gph} \widehat{N}_\Gamma}(\bar{y}, \bar{y}^*) \times T_{\mathbb{R}^p_-}(G(\bar{x}, \bar{y})) \end{aligned} \tag{76}$$

fulfilling (32), $\nabla F(\bar{x}, \bar{y})(\delta x, \delta y) = 0$ which is (74a) and (34) 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 $T_{\operatorname{gph} \widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)$ 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 (44) we see that $T_{\operatorname{gph} \widehat{N}_\Gamma \times \mathbb{R}^p_-}((\bar{y}, \bar{y}^*), G(\bar{x}, \bar{y}))$ 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

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

fulfills (35). Further, the graphs of M_p and M_c are closed cones and from (36) we conclude that $\delta y \neq 0$ by taking account of (76). Next we utilize (31), 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 the 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 2(ii), we can apply Theorem 4 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 [13, 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 $\Omega = \{z \,|\, P(z) \in D\}$ is replaced by the set $\{u \,|\, \nabla P(\bar{z})u \in T_D\big(P(\bar{z})\big)\}$, whose linearized tangent cone at $\bar{u}$ is $\{u \,|\, \nabla P(\bar{z})u \in T_{T_D(P(\bar{z}))}(\nabla P(\bar{z})\bar{u})\}$. Since

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

and g is 2-nondegenerate in direction $\delta y \neq 0$, by (54) the set $T_{T_D(P(\bar z))}(\nabla P(\bar z)\bar u)$ is polyhedral and therefore only the first alternative of Theorem 4 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

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

with $0 \in \nabla F(\bar z) + \nabla P(\bar z)^T\omega^*$ which results in (74b) and (74c).

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

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

If $\delta\bar v = 0$ and $\mathcal L(\bar K_\Gamma) \neq \{0\}$, then by (63) 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 (75) follows from (62) by taking $\bar v = 0$. □

Now we are ready to state and prove our main optimality condition for the problem (3). The main task is to interpret the formulas for the tangent cones and the regular normal cones in Propositions 6-8 appeared in Lemma 4 in terms of problem data.

Theorem 6. *Assume that $(\bar x, \bar y)$ is a local minimizer for the problem (3) 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*

$$\nabla_x F(\bar x,\bar y) - \nabla_x\phi(\bar x,\bar y)^T w + \nabla_x G(\bar x,\bar y)^T\sigma = 0, \tag{77a}$$

$$\begin{aligned}&\nabla_y F(\bar x,\bar y) - \nabla_y\phi(\bar x,\bar y)^T w + \nabla_y G(\bar x,\bar y)^T\sigma\\ &\qquad - \nabla^2(\bar\lambda^T g)(\bar y)w + 2\nabla^2(\eta^T g)(\bar y)\bar v \in -(\mathcal F_1^v - \mathcal F_2^v)^\circ,\end{aligned} \tag{77b}$$

$$\bar v^T\nabla^2 g(\bar y)w \in (\mathcal F_1^\lambda - \mathcal F_2^\lambda)^\circ, \tag{77c}$$

$$\bar v^T\nabla^2 g(\bar y)\delta v \in T_{\bar\Lambda(\bar v)}(\bar\lambda)^\circ, \mathcal F_1^\lambda = T_{\bar\Lambda(\bar v)}(\bar\lambda) \cap [\bar v^T\nabla^2 g(\bar y)\delta v]^\perp, \tag{77d}$$

$$\sigma_i G_i(\bar x,\bar y) = 0,\ i = 1,\ldots,p. \tag{77e}$$

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

$$\nabla F(\bar x,\bar y)^T(\delta x, \bar v) = 0, \tag{78a}$$

$$\nabla\phi(\bar x,\bar y)(\delta x, \bar v) + \nabla^2(\bar\lambda^T g)(\bar y)\bar v + \bar z^* = 0, \tag{78b}$$

$$\nabla G_i(\bar x,\bar y)^T(\delta x, \bar v) \le 0,\ \sigma_i\nabla G_i(\bar x,\bar y)^T(\delta x, \bar v) = 0,\ \forall i: G_i(\bar x,\bar y) = 0, \tag{78c}$$

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 (77a) holds and (77e) follows from the observation that

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

Case I: $\delta y = 0$**.** Then we also have $\delta x = 0$ by (74d) and thus

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

Subcase Ia: $\delta\bar{v} \neq 0$**.** Set $\bar{v} = \delta\bar{v}$ and by Lemma 2 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 (56) there is some $\eta \in T_{\bar{\Lambda}(\bar{v})}(\bar{\lambda})$ such that

$$w^* + \nabla^2(\bar{\lambda}^T g)(\bar{y})w - 2\nabla^2(\eta^T g)(\bar{y})\bar{v} \in \mathcal{K}_{\bar{K}_\Gamma}(\bar{v},\bar{z}^*)^\circ, \tag{79a}$$
$$w \in \mathcal{K}_{\bar{K}_\Gamma}(\bar{v},\bar{z}^*), \tag{79b}$$
$$\bar{v}^T\nabla^2 g(\bar{y})w \in T_{\bar{\Lambda}(\bar{v})}(\bar{\lambda})^\circ. \tag{79c}$$

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 (77c) follows from (79c), (77d) is fulfilled and (77b) follows from (74c) and (79a).

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 (75) 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 $\big(\bar{v}, -\nabla\phi(\bar{x},\bar{y})(\delta x,\bar{v})\big) \in T_{\operatorname{gph}\widehat{N}_\Gamma}(\bar{y},\bar{y}^*)$ and by Lemma 2 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 (66) and (67), 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

$$\begin{aligned}&\delta\bar{v}^* = \nabla^2(\bar{\lambda}^T g)(\bar{y})\delta v + \nabla^2(\bar{\mu}^T g)(\bar{y})\bar{v} + \bar{\zeta}^*,\\ &(\delta v, \bar{\mu}, \bar{\zeta}^*, 2\bar{v}^T\nabla^2 g(\bar{y})\delta v) \in \operatorname{gph} N_{\tilde{K}(\bar{v},-\nabla\phi(\bar{x},\bar{y})(\delta x,\bar{v}))}.\end{aligned}$$

Further, by (69) there is some η such that

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

By taking into account

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

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 $\delta v \in \operatorname{ri} \mathcal{F}_2^v$, $\bar{\mu} \in \operatorname{ri} \mathcal{F}_2^\lambda$ 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 (77) 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. □

In the following remark we summarize some comments on the optimality conditions of Theorem 6.

Remark 3.

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 [17, 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 [15, Proposition 4.3(iii)]. As a consequence it follows from Proposition 5 that $T_{\operatorname{gph} \widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)$ is locally polyhedral near $(\bar{v}, \bar{v}^*)$ for every $\bar{v}^*$ satisfying $(\bar{v}, \bar{v}^*) \in T_{\operatorname{gph} \widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)$. Thus by Theorem 6 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 $\bar{\lambda} \in \operatorname{ri} \bar{\Lambda}(\bar{v})$ 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. [18, Examples 1,2])**.** *Consider the MPEC*

$$\begin{array}{lll} \min\limits_{x,y} & F(x,y) := x_1 - \frac{3}{2}y_1 + x_2 - \frac{3}{2}y_2 - y_3 & \\ s.t. & 0 \in \phi(x,y) + N_\Gamma(y), & (80) \\ & G_1(x,y) = G_1(x) := -x_1 - 2x_2 \leq 0, & \\ & G_2(x,y) = G_2(x) := -2x_1 - x_2 \leq 0, & \end{array}$$

where

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

As it was demonstrated in [18], $\bar{x} = (0,0)$ and $\bar{y} = (0,0,0)$ is the unique global solution and Assumption 1 is fulfilled. Straightforward calculations yield

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

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*

$$\begin{pmatrix} \mu_1 v_1 \\ \mu_2 v_2 \\ 0 \end{pmatrix} \in \big(N_{\bar{K}_\Gamma}(v)\big)^+ = \big(\{0\} \times \{0\} \times \mathbb{R}\big)^+ = \{0\} \times \{0\} \times \mathbb{R},\ \mu \in \big(\bar{\Lambda}(v)\big)^+ \ \Rightarrow\ \mu = (0,0) \tag{81}$$

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* $\big(\bar{\Lambda}(v)\big)^+ = \{0\}$. *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* (81) *holds in this case as well. We claim that the optimality conditions of Theorem 6 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 3,* $\mathcal{K}_{\bar{K}_\Gamma}(\bar{v}, \bar{z}^*) = \bar{K}_\Gamma$, $\mathcal{F}_1^v, \mathcal{F}_2^v$ *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* (77a), (77b), (77c) *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, (77d) *holds because of* $\delta v = 0$ *and* $\mathcal{F}_2^v$ *is a subspace, and* (77e) *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 6 are fulfilled. Note that for this example Theorem 1 cannot be applied because MPCC-GCQ does not hold at* $(\bar{x}, \bar{y}, \lambda)$ *for any* $\lambda \in \bar{\Lambda}$, *cf. [18].*

At the end of this section we want to formulate the necessary optimality conditions in Theorem 6 in terms of index sets instead of faces. To this end let us define the index sets

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

By the definition of the critical cone in (40), we have

$$\bar{K}_\Gamma = \Big\{ v \,|\, \nabla g_i(\bar{y})^T v \leq 0 \text{ if } g_i(\bar{y}) = 0 \Big\} \cap [\bar{y}^*]^\perp.$$

Since by the definition of the multiplier set (41),

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

we have

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

Hence it is obvious that

$$\bar{K}_\Gamma = \Big\{ v \,|\, \nabla g_i(\bar{y})^T v \begin{cases} = 0 & i \in \bar{J}^+(\bar{\Lambda}) \\ \le 0 & i \in \bar{I} \setminus \bar{J}^+(\bar{\Lambda}) \end{cases} \Big\}. \tag{82}$$

Theorem 7. *Assume that $(\bar{x}, \bar{y})$ is a local minimizer for the problem (3) 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*

$$0 = \nabla_x F(\bar{x}, \bar{y}) - \nabla_x \phi(\bar{x}, \bar{y})^T w + \nabla_x G(\bar{x}, \bar{y})^T \sigma, \tag{83a}$$

$$0 = \nabla_y F(\bar{x}, \bar{y}) - \nabla_y \phi(\bar{x}, \bar{y})^T w + \nabla_y G(\bar{x}, \bar{y})^T \sigma - \nabla^2(\bar{\lambda}^T g)(\bar{y}) w + \nabla g(\bar{y})^T \xi + 2\nabla^2(\eta^T g)(\bar{y})\bar{v}, \tag{83b}$$

$$\xi_i = 0 \text{ if } i \notin \mathcal{I}, \tag{83c}$$

$$\xi_i \ge 0, \nabla g_i(\bar{y})^T w \le 0 \text{ if } i \in \mathcal{I} \setminus \mathcal{I}^+, \tag{83d}$$

$$\nabla g_i(\bar{y})^T w = 0 \text{ if } i \in \mathcal{I}^+, \tag{83e}$$

$$\nabla g(\bar{y})^T \eta = 0, \ \eta_i = 0, i \notin \mathcal{J}, \ \eta_i \ge 0, i \in \mathcal{J} \setminus \mathcal{J}^+, \tag{83f}$$

$$0 = \sigma_i G_i(\bar{x}, \bar{y}), \ i = 1, \ldots, p. \tag{83g}$$

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

$$\nabla g_i(\bar{y})^T \delta v = 0, \ i \in \bar{J}^+(\bar{\Lambda}), \ \nabla g_i(\bar{y})^T \delta v \le 0, \ i \in \bar{I}(\bar{v}) \setminus \bar{J}^+(\bar{\Lambda}), \tag{83h}$$

$$\mathcal{I} = \{ i \in \bar{I}(\bar{v}) \,|\, \nabla g_i(\bar{y})^T \delta v = 0 \}, \tag{83i}$$

$$\nabla g_i(\bar{y})^T s_{\delta v} + \bar{v}^T \nabla^2 g_i(\bar{y}) \delta v = 0, i \in \bar{J}^+(\bar{\lambda}), \ \nabla g_i(\bar{y}) s_{\delta v} + \bar{v}^T \nabla^2 g_i(\bar{y})^T \delta v \le 0, i \in \bar{J}^+(\bar{\Lambda}(\bar{v})) \setminus \bar{J}^+(\bar{\lambda}) \tag{83j}$$

$$\mathcal{J} = \{ i \in \bar{J}^+(\bar{\Lambda}(\bar{v}))) \,|\, \nabla g_i(\bar{y})^T s_{\delta v} + \bar{v}^T \nabla^2 g_i(\bar{y}) \delta v = 0 \} \tag{83k}$$

$$\nabla g(\bar{y})^T \bar{\mu} = 0, \ \bar{\mu}_i = 0, i \notin \mathcal{J}, \bar{\mu}_i \ge 0, i \in \mathcal{J} \setminus \bar{J}^+(\bar{\lambda}) \tag{83l}$$

$$\mathcal{J}^+ = \bar{J}^+(\bar{\lambda}) \cup \{ i \in \mathcal{J} \setminus \bar{J}^+(\bar{\lambda}) \,|\, \bar{\mu}_i > 0 \} \tag{83m}$$

$$\nabla g_i(\bar{y})^T s_w + \bar{v}^T \nabla^2 g_i(\bar{y}) w = 0, i \in \mathcal{J}^+, \ \nabla g_i(\bar{y})^T s_w + \bar{v}^T \nabla^2 g_i(\bar{y}) w \le 0, i \in \mathcal{J} \setminus \mathcal{J}^+. \tag{83n}$$

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) $\bar{v} \neq 0$; 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 $\mathcal{I} \neq \bar{I}(\bar{v})$ or $\mathcal{J}^+ \neq \bar{J}^+(\bar{\lambda})$ or $\mathcal{J} \neq \bar{J}^+(\bar{\Lambda}(\bar{v}))$ then $\bar{v} \neq 0$ 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 (78a)-(78c) *hold 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. 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 6. The index sets $\mathcal{I}, \mathcal{I}^+, \mathcal{J}, \mathcal{J}^+$ were chosen such that

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

Then $(\mathcal{F}_1^v - \mathcal{F}_2^v)^\circ = \{\nabla g(\bar{y})^T \xi \mid \xi_i = 0, i \notin \mathcal{I},\ \xi_i \geq 0, i \in \mathcal{I} \setminus \mathcal{I}^+\}$ and conditions (83a)-(83g) follow immediately. (83h) states that $\delta v \in T_{\bar{K}_\Gamma}(\bar{v})$ whereas (83i) results from the requirement $\delta v \in \operatorname{ri} \mathcal{F}_2^v$ 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 $\mathcal{I}^+$. (83j) states that $\bar{v}^T \nabla^2 g(\bar{y}) \delta v \in T_{\bar{\Lambda}(\bar{v})}(\bar{\lambda})^\circ$ and condition (83k) results from the second part of (77d). The point $\bar{\mu}$ denotes any point in $\operatorname{ri} F_2^\lambda$ yielding the condition (83m) by Proposition 4. Finally, condition (83n) is equivalent to (77c). □

As discussed in the introduction, the problem (MPCC) may not be equivalent to the original (MPVIC) in the case where the lower level problem does not have unique multiplier. It is easy to see that since the multiplier $\bar{\lambda}$ in our theorem is not arbitrary as in the one for (MPCC), one can not compare our new necessary condition with the M-stationary condition for (MPCC) unless the lower level multiplier is unique. We now compare our necessary optimality conditions with M-optimality condition for (MPCC) in the case where the multiplier set $\bar{\Lambda} = \{\bar{\lambda}\}$ is a singleton. In this case the condition that g is 2-nondegenerate in every nonzero critical direction $0 \neq v \in \bar{K}_\Gamma$ at $(\bar{y}, \bar{y}^*)$ holds automatically. By [2, Proposition 2], the assumption of MSCQ for (MPEC) is weaker than the corresponding one for (MPCC). Further, by Proposition 5 the set $T_{\operatorname{gph} \widehat{N}_\Gamma}(\bar{y}, \bar{y}^*)$ is polyhedral and therefore by Theorem 7, 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 7 and take, as already pointed out in Remark 3, $\eta = 0$. Moreover, we can take $\delta v = 0$, $s_{\delta v} = 0$ and $\bar{\mu} = 0$ in order to fulfill conditions (83h)-(83m). 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 (83n). Therefore we have the following corollary.

Corollary 2. *Assume that $(\bar{x}, \bar{y})$ is a local minimizer for the problem (3) 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*

$$0 = \nabla_x F(\bar{x}, \bar{y}) - \nabla_x \phi(\bar{x}, \bar{y})^T w + \nabla_x G(\bar{x}, \bar{y})^T \sigma, \tag{84a}$$
$$0 = \nabla_y F(\bar{x}, \bar{y}) - \nabla_y \phi(\bar{x}, \bar{y})^T w + \nabla_y G(\bar{x}, \bar{y})^T \sigma - \nabla^2(\bar{\lambda}^T g)(\bar{y}) w + \nabla g(\bar{y})^T \xi, \tag{84b}$$
$$\xi_i = 0 \text{ if } i \notin \bar{I}(\bar{v}), \tag{84c}$$
$$\xi_i \geq 0, \nabla g_i(\bar{y})^T w \leq 0 \text{ if } i \in \bar{I}(\bar{v}) \setminus \mathcal{I}^+, \tag{84d}$$
$$\nabla g_i(\bar{y})^T w = 0 \text{ if } i \in \mathcal{I}^+, \tag{84e}$$
$$0 = \sigma_i G_i(\bar{x}, \bar{y}),\ i = 1, \ldots, p. \tag{84f}$$

Suppose that $(\bar{x}, \bar{y}, \bar{\lambda})$ satisfies the optimality condition in Corollary 2 and let $w \in \mathbb{R}^m$, $\xi \in \mathbb{R}^q$ and $\sigma \in \mathbb{R}^p_+$ be those found in Corollary 2. Then (4a)-(4b) and (4f) 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 $\bar{J}^+(\bar{\lambda}) = \{i | g_i(\bar{y}) = 0, \bar{\lambda}_i > 0\}$, (84c) and (84e) 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 (4c)-(4e) 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 Theorem 1 imply the necessary optimality conditions of Corollary 2 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 (4), 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 (84) are fulfilled. However, the following example demonstrates that the optimality conditions of Corollary 2 are sharper, when $\bar{\Lambda}$ is a singleton but LICQ fails.

Example 2. *Consider the problem*

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

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

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 Theorem 1 amount to

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

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 2 with $w = (0, 1)$, and $\xi = (-1, 0, 0)$does not hold.

By applying [18, Theorem 5] 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 $\nabla g_1(\bar{y})w = w_1$, (84e) 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 (84c) 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 [16], 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 [16] 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 [16, Proposition 3] is stronger than the 2-nondegeneracy assumption we use as stated in Definition 9. To see this recall the concept of *2-regularity*: 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). 2-regularity was initiated (and named) by Tret'yakov [26] in the case of zero Jacobian and then was strongly developed by Avakov [1]. The sufficient condition for 2-LICQ in direction $\bar{v} \in \bar{K}_\Gamma$ used in [16, Proposition 3] now states that for every index set $\hat{J}$ with $\bar{J}^+(\bar{\Lambda}(\bar{v})) \subseteq \hat{J} \subseteq \bar{I}(\bar{v})$ satisfying

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

the mapping $(g_i)_{i\in\hat{J}}$ is 2-regular in direction $\bar{v}$. Such an index set $\hat{J}$ always exists, e.g. by duality theory of linear programming $\hat{J} = \bar{J}^+(\bar{\Lambda}(\bar{v}))$ is a possible choice. Now choose $\hat{J}$ 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 $\nabla g_i(\bar{y})$, $i \in \hat{J}$. 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 $(g_i)_{i\in\hat{J}}$ is equivalent to the statement

$$\sum_{i\in\hat{J}}(\eta_i\nabla g_i(\bar{y}) + \mu_i\nabla^2 g_i(\bar{y})) = 0,\ \sum_{i\in\hat{J}}\mu_i\nabla g_i(\bar{y}) = 0\ \Rightarrow \mu_i = 0,\ i \in \hat{J}. \tag{85}$$

By taking into account

$$\begin{aligned}\big(N_{\bar{K}_\Gamma}(\bar{v})\big)^+ &= \{\sum_{i\in\bar{I}(\bar{v})}\eta_i\nabla g_i(\bar{y}) \mid \eta_i \in \mathbb{R}, i \in \bar{I}(\bar{v})\} = \{\sum_{i\in\hat{J}}\eta_i\nabla g_i(\bar{y}) \mid \eta_i \in \mathbb{R}, i \in \hat{J}\},\\ \big(\bar{\Lambda}(\bar{v})\big)^+ &= \{\mu \in \mathbb{R}^q \mid \nabla g(\bar{y})^T\mu = 0,\ \bar{v}^T\nabla^2(\mu^T g)(\bar{y})\bar{v} = 0, \mu_i = 0,\ i \notin \bar{J}^+(\bar{\Lambda}(\bar{v}))\}\end{aligned}$$

it follows that (85) implies 2-nondegeneracy. This proves our claim that the sufficient condition for 2-LICQ is stronger than 2-nondegeneracy.

Further one can show that the necessary conditions of Theorem 7 are stronger than the M-stationary conditions which one could obtain with the M-stationary conditions of [16, Theorem 4] insofar as an additional condition on δx is included in Theorem 7.

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

[8] 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.

[9] 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.

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

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

[12] 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.

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

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

[15] 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.

[16] 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.

[17] 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.

[18] 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.

[19] R. Henrion, A. Jourani, J. V. Outrata, *On the calmness of a class of multifunctions*, SIAM J. Optim., 13 (2002), pp. 603–618.

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

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

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

[23] 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.

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

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

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

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

[28] J.J. Ye, J.C. Zhou, *Verifiable sufficient conditions for the error bound property of second-order cone complementarity problems*, Math. Program., 171 (2017), pp. 361-395.

Latest Reports in this series

2009 - 2016

[..]

2017

2017-01	Katharina Rafetseder and Walter Zulehner *A Decomposition Result for Kirchhoff Plate Bending Problems and a New Discretization Approach*	March 2017
2017-02	Clemens Hofreither *A Black-Box Algorithm for Fast Matrix Assembly in Isogeometric Analysis*	April 2017
2017-03	Ulrich Langer, Martin Neumüller and Ioannis Toulopoulos *Multipatch Space-Time Isogeometric Analysis of Parabolic Diffusion Problems*	May 2017
2017-04	Jarle Sogn and Walter Zulehner *Schur Complement Preconditioners for Multiple Saddle Point Problems of Block Tridiagonal Form with Application to Optimization Problems*	July 2017
2017-05	Helmut Gfrerer and Boris S. Mordukhovich *Second-Order Variational Analysis of Parametric Constraint and Variational Systems*	November 2017
2017-06	Katharina Rafetseder and Walter Zulehner *On a New Mixed Formulation of Kirchhoff Plates on Curvilinear Polygonal Domains*	November 2017
2017-07	Matúš Benko, Helmut Gfrerer and Jiří V. Outrata *Calculus for Directional Limiting Normal Cones and Subdifferentials*	December 2017
2017-08	Helmut Gfrerer *Linearized M-Stationarity Conditions for General Optimization Problems*	December 2017

2018

2018-01	Jarle Sogn and Stefan Takacs *Robust Multigrid Solvers for the Biharmonic Problem in Isogeometric Analysis*	January 2018
2018-02	Irina Georgieva and Clemens Hofreither *Greedy Low-Rank Approximation in Tucker Format of Tensors and Solutions of Tensor Linear Systems*	March 2018
2018-03	Katharina Rafetseder and Walter Zulehner *A New Mixed Isogeometric Approach to Kirchhoff-Love Shells*	March 2018
2018-04	Matúš Benko, Helmut Gfrerer and Boris S. Mordukhovich *Characterizations of Tilt-Stable Minimizers in Second-Order Cone Programming*	September 2018
2018-05	Helmut Gfrerer and Jane J. Ye *New Sharp Necessary Optimality Conditions for Mathematical Programs with Equilibrium Constraints*	December 2018

From 1998 to 2008 reports were published by SFB013. Please see
`http://www.sfb013.uni-linz.ac.at/index.php?id=reports`
From 2004 on reports were also published by RICAM. Please see
`http://www.ricam.oeaw.ac.at/publications/list/`
For a complete list of NuMa reports see
`http://www.numa.uni-linz.ac.at/Publications/List/`