New constraint qualifications for mathematical programs with equilibrium constraints via variational analysis

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. Compared with the usual way of formulating MPEC through a KKT condition, this formulation has the advantage that it does not involve extra multipliers as new variables, and it usually requires weaker assumptions on the problem data. Using the so-called first order sufficient condition for metric subregularity, we derive verifiable sufficient conditions for the metric subregularity of the involved set-valued mapping, or equivalently the calmness of the perturbed generalized equation mapping.

Key words: mathematical programs with equilibrium constraints, constraint qualification, metric subregularity, calmness.

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

1 Introduction

A mathematical program with equilibrium constraints (MPEC) usually refers to an optimization problem in which the essential constraints are defined by a parametric variational inequality or complementarity system. Since many equilibrium phenomena that arise from engineering and economics are characterized by either an optimization problem or a variational inequality, this justifies the name mathematical program with equilibrium constraints ([27, 30]). During the last two decades, more and more applications for MPECs have been found and there has been much progress made in both theories and algorithms for solving MPECs.

For easy discussion, consider the following mathematical program with a variational inequality constraint

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

where $C \subseteq \mathbb{R}^n \times \mathbb{R}^m$, $\Gamma := \{y \in \mathbb{R}^m | g(y) \leq 0\}$, $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, $\phi : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$, and $g : \mathbb{R}^m \to \mathbb{R}^q$ are sufficiently smooth. If the set Γ is convex, then MPVIC can be equivalently

*Institute of Computational Mathematics, Johannes Kepler University Linz, A-4040 Linz, Austria, e-mail: helmut.gfrerer@jku.at.

†Department of Mathematics and Statistics, University of Victoria, Victoria, B.C., Canada V8W 2Y2, e-mail: janeye@uvic.ca.

written as a mathematical program with a generalized equation constraint

$$\text{(MPGE)} \qquad \min_{(x,y)\in C} \quad F(x,y) \qquad \text{s.t.} \quad 0 \in \phi(x,y) + N_\Gamma(y),$$

where $N_\Gamma(y)$ is the normal cone to set Γ at y in the sense of convex analysis. If g is affine or certain constraint qualification such as the Slater condition holds for the constraint $g(y) \le 0$, then it is known that $N_\Gamma(y) = \nabla g(y)^T N_{\mathbb{R}^q_-}(g(y))$. Consequently

$$0 \in \phi(x,y) + N_\Gamma(y) \Longleftrightarrow \exists \lambda : 0 \in \big(\phi(x,y) + \nabla g(y)^T\lambda, g(y)\big) + N_{\mathbb{R}^m \times \mathbb{R}^q_+}(y,\lambda), \tag{2}$$

where λ is referred to as a multiplier. This suggests to consider the mathematical program with a complementarity constraint

$$\text{(MPCC)} \qquad \min_{(x,y)\in C, \lambda \in \mathbb{R}^q} \quad F(x,y) \qquad \text{s.t.} \quad 0 \in \big(\phi(x,y) + \nabla g(y)^T\lambda, g(y)\big) + N_{\mathbb{R}^m \times \mathbb{R}^q_+}(y,\lambda).$$

In the case where the equivalence (2) holds for a unique multiplier λ for each y, (MPGE) and (MPCC) are obviously equivalent, while in the case where the multipliers are not unique, the two problems are not necessarily equivalent if the local optimal solutions are considered (see Dempe and Dutta [8] 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 (MPGE). This is a serious drawback for using the MPCC reformulation, since a numerical method computing a stationary point for (MPCC) may not have anything to do with the solution to the original MPEC. This shows that whenever possible, one should consider solving problem (MPGE) instead of problem (MPCC). Another fact we want to mention is that in many equilibrium problems, the constraint set Γ or the function g may not be convex. In this case, if y solves the variational inequality (1), then $y' = y$ is a global minimizer of the optimization problem: $\min\limits_{y'} \ \langle \phi(x,y), y' \rangle$ s.t. $y' \in \Gamma$, and hence by Fermat's rule (see, e.g., [34, Theorem 10.1]) it is a solution of the generalized equation

$$0 \in \phi(x,y) + \widehat{N}_\Gamma(y), \tag{3}$$

where $\widehat{N}_\Gamma(y)$ is the regular normal cone to Γ at y (see Definition 1). In the nonconvex case, by replacing the original variational inequality constraint (1) by the generalized equation (3), the feasible region is enlarged and the resulting MPGE may not be equivalent to the original MPVIC. However, if the solution $(\bar{x}, \bar{y})$ of MPGE is feasible for the original MPVIC, then it must be a solution of the original MPVIC; see [2] for this approach in the context of bilevel programs.

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

$$\begin{array}{lll} \text{(MPEC)} & \min & F(x,y) \\ & \text{s.t.} & 0 \in \phi(x,y) + \widehat{N}_\Gamma(y), \\ & & G(x,y) \le 0, \end{array}$$

where Γ is possibly non-convex and $G : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p$ is smooth.

Besides of the issue of equivalent problem formulations, one has to consider constraint qualifications as well. This task is of particular importance for deriving optimality conditions. For the problem (MPCC), there exist a lot of constraint qualifications from the MPEC-literature ensuring the Mordukhovich (M-)stationarity of locally optimal solutions. The weakest one of these constraint qualifications appears to be MPEC-GCQ (Guignard constraint qualification) as introduced by Flegel and Kanzow [11], see [12] for a proof of M-stationarity of locally optimal solutions under MPEC-GCQ. For the problem (MPEC), it was shown by Ye and Ye [37] that calmness of the perturbation mapping associated with the constraints of (MPEC) (called pseudo upper-Lipschitz continuity in [37]) guarantees M-stationarity of solutions. [1] 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, [1, 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 this paper, we first show that if multipliers are not unique, then the MPEC Mangasarian-Fromovitz constraint qualification (MFCQ) never holds for problem (MPCC). Then we present 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 $(\bar{x}, \bar{y}, 0)$. Note that in contrast to [1, Example 6], Γ in our example is even convex. However, very little is known about how to verify the calmness for (MPEC) when the multiplier λ is not unique. When ϕ, g and G are affine, calmness follows simply by Robinson's result on polyhedral multifunctions [33]. Another approach is to verify calmness by showing the stronger Aubin property (also called pseudo Lipschitz continuity or Lipschitz-like property) via the so-called Mordukhovich criterion, cf. [29]. However, the Mordukhovich criterion involves the limiting coderivate of $\widehat{N}_\Gamma(\cdot)$, which is very difficult to compute in the case of nonunique λ; see [20].

The main goal of this paper is to derive a simply verifiable criterion for the so-called metric subregularity constraint qualification (MSCQ); see Definition 5, which is equivalent to calmness. Our sufficient condition for MSCQ involves only first-order derivatives of ϕ and G and derivatives up to the second-order of g at $(\bar{x}, \bar{y})$, and is therefore efficiently checkable. Our approach is mainly based on the sufficient conditions for metric subregularity as recently developed in [13, 14, 15, 16], and some implications of metric subregularity which can be found in [18, 21]. A special feature is that the imposed constraint qualification on both the lower level system $g(y) \leq 0$ and the upper level system $G(x, y) \leq 0$ is only MSCQ, which is much weaker than MFCQ usually required.

We organize our paper as follows. Section 2 contains the preliminaries and preliminary results. In section 3, we discuss the difficulties involved in formulating MPECs as (MPCC). Section 4 gives new verifiable sufficient conditions for MSCQ.

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 differentiable mapping $P : \mathbb{R}^d \to \mathbb{R}^s$, we denote by $\nabla P(z)$ the Jacobian matrix of P at z if $s > 1$ and the gradient vector if $s = 1$. For a function $f : \mathbb{R}^d \to \mathbb{R}$, we denote by $\nabla^2 f(\bar{z})$ the Hessian matrix of f at $\bar{z}$. 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 Basic definitions and preliminary results

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 [7, 29, 34] and in the papers we refer to.

Definition 1. *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*

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

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

$$\widehat{N}_\Omega(\bar{z}) := \left\{ v^* \in \mathbb{R}^d \middle| \, \limsup_{z \xrightarrow{\Omega} \bar{z}} \frac{\langle v^*, z - \bar{z}\rangle}{\|z - \bar{z}\|} \leq 0 \right\}$$

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

respectively.

Note that $\widehat{N}_\Omega(\bar{z}) = (T_\Omega(\bar{z}))^\circ$, and 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.

It is easy to see that $u \in T_\Omega(\bar{z})$ if and only if $\liminf_{t\downarrow 0} t^{-1}\mathrm{d}(\bar{z} + tu, \Omega) = 0$. Recall that a set Ω is said to be geometrically derivable at a point $\bar{z} \in \Omega$ if the tangent cone coincides with the derivable cone at $\bar{x}$, i.e., $u \in T_\Omega(\bar{z})$ if and only if $\lim_{t\downarrow 0} t^{-1}\mathrm{d}(\bar{z} + tu, \Omega) = 0$; see e.g., [34]. From the definitions of various tangent cones, it is easy to see that if a set Ω is Clarke regular in the sense of [7, Definition 2.4.6], then it must be geometrically derivable and the converse relation is in general false. The following proposition therefore improves the rule of tangents to product sets given in [34, Proposition 6.41]. The proof is omitted since it follows from the definitions of the tangent cone and the geometrical derivability of the set.

Proposition 1 (Rule of Tangents to Product Sets)**.** *Let $\Omega = \Omega_1 \times \Omega_2$ with $\Omega_1 \subseteq \mathbb{R}^{d_1}, \Omega_2 \subseteq \mathbb{R}^{d_2}$ closed. Then at any $\bar{z} = (\bar{z}_1, \bar{z}_2)$ with $\bar{z}_1 \in \Omega_1, \bar{z}_2 \in \Omega_2$, one has*

$$T_\Omega(\bar{z}) \subseteq T_{\Omega_1}(\bar{z}_1) \times T_{\Omega_2}(\bar{z}_2).$$

Furthermore equality holds if at least one of sets Ω_1, Ω_2 is geometrically derivable.

The following directional version of the limiting normal cone was introduced in [14].

Definition 2. *Given a set $\Omega \subseteq \mathbb{R}^d$, a point $\bar{z} \in \Omega$ and a direction $w \in \mathbb{R}^d$, the* limiting normal cone *to Ω in direction w at $\bar{z}$ is defined by*

$$N_\Omega(\bar{z}; w) := \left\{ z^* | \exists t_k \downarrow 0, w_k \to w, z_k^* \to z^* : z_k^* \in \widehat{N}_\Omega(\bar{z} + t_k w_k) \ \forall k \right\}.$$

By definition, it is easy to see that $N_\Omega(\bar z;0) = N_\Omega(\bar z)$ and $N_\Omega(\bar z;u) = \emptyset$ if $u \notin T_\Omega(\bar z)$. Further by [15, Lemma 2.1], if Ω is the union of finitely many closed convex sets, then one has the following relationship between the limiting normal cone and its directional version.

Proposition 2. *[15, Lemma 2.1] Let $\Omega \subseteq \mathbb{R}^d$ be the union of finitely many closed convex sets, $\bar z \in \Omega, u \in \mathbb{R}^d$. Then $N_\Omega(\bar z;u) \subseteq N_\Omega(\bar z) \cap \{u\}^\perp$ and equality holds if the set Ω is convex and $u \in T_\Omega(\bar z)$.*

Next we consider constraint qualifications for a constraint system of the form

$$z \in \Omega := \{z \mid P(z) \in D\}, \tag{4}$$

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

Definition 3 (cf. [12])**.** *Let $\bar z \in \Omega$ where Ω is defined as in (4) with P smooth, and $T^{\rm lin}_\Omega(\bar z)$ be the linearized cone of Ω at $\bar z$ defined by*

$$T^{\rm lin}_\Omega(\bar z) := \{w | \nabla P(\bar z)w \in T_D(P(\bar z))\}. \tag{5}$$

We say that the generalized Abadie constraint qualification *(GACQ) and the* generalized Guignard constraint qualification *(GGCQ) hold at $\bar z$, if*

$$T_\Omega(\bar z) = T^{\rm lin}_\Omega(\bar z) \text{ and } (T_\Omega(\bar z))^\circ = (T^{\rm lin}_\Omega(\bar z))^\circ,$$

respectively.

It is obvious that GACQ implies GGCQ which is considered as the weakest constraint qualification. In the case of a standard nonlinear program, GACQ and GGCQ reduce to the standard definitions of Abadie and Guignard constraint qualification, respectively. Under GGCQ, any locally optimal solution to a disjunctive problem, i.e., an optimization problem where the constraint has the form (4) with the set D equal to the union of finitely many polyhedral convex sets, must be M-stationary (see e.g., [12, Theorem 7]).

GACQ and GGCQ are weak constraint qualifications, but they are usually difficult to verify. Hence, we are interested in constraint qualifications that are effectively verifiable, and yet not too strong. The following notion of metric subregularity is the base of the constraint qualification, which plays a central role in this paper.

Definition 4. *Let $M : \mathbb{R}^d \rightrightarrows \mathbb{R}^s$ be a set-valued mapping and let $(\bar z, \bar w) \in \operatorname{gph} M$. We say that M is* metrically subregular *at $(\bar z, \bar w)$ if there exist a neighborhood W of $\bar z$ and a positive number $\kappa > 0$ such that*

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

The metric subregularity property was introduced in [26] for single-valued maps under the terminology "regularity at a point", and the name of "metric subregularity" was suggested in [9]. Note that the metrical subregularity at $(\bar z, 0) \in \operatorname{gph} M$ is also referred to as the existence of a local error bound at $\bar z$. It is easy to see that M is metrically subregular at $(\bar z, \bar w)$ if and only if its inverse set-valued map M^{-1} is calm at $(\bar w, \bar z) \in \operatorname{gph} M^{-1}$, i.e., there exist a neighborhood W of $\bar z$, a neighborhood V of $\bar w$ and a positive number $\kappa > 0$ such that

$$M^{-1}(w) \cap V \subseteq M^{-1}(\bar w) + \kappa \|w - \bar w\| \mathcal{B} \quad \forall z \in W.$$

While the term for the calmness of a set-valued map was first coined in [34], it was introduced as the pseudo-upper Lipschitz continuity in [37], taking into account that it is weaker than both the pseudo Lipschitz continuity of Aubin [5] and the upper Lipschitz continuity of Robinson [31, 32] .

More general constraints can be easily written in the form $P(z) \in D$. For instance, a set $\Omega = \{z \,|\, P_1(z) \in D_1, 0 \in P_2(z) + Q(z)\}$, where $P_i : \mathbb{R}^d \to \mathbb{R}^{s_i}$, $i = 1, 2$ and $Q : \mathbb{R}^d \rightrightarrows \mathbb{R}^{s_2}$ is a set-valued map, can also be written as

$$\Omega = \{z \,|\, P(z) \in D\} \text{ with } P(z) := \begin{pmatrix} P_1(z) \\ (z, -P_2(z)) \end{pmatrix}, \; D := D_1 \times \operatorname{gph} Q.$$

We now show that for both representations of Ω, the properties of metric subregularity for the maps describing the constraints are equivalent.

Proposition 3. *Let $P_i : \mathbb{R}^d \to \mathbb{R}^{s_i}$, $i = 1, 2$, $D_1 \subseteq \mathbb{R}^{s_1}$ be closed and $Q : \mathbb{R}^d \rightrightarrows \mathbb{R}^{s_2}$ be a set-valued map with a closed graph. Further assume that P_1 and P_2 are Lipschitz near $\bar{z}$. Then the set-valued map*

$$M_1(z) := \begin{pmatrix} P_1(z) - D_1 \\ P_2(z) + Q(z) \end{pmatrix}$$

is metrically subregular at $(\bar{z}, (0, 0))$ if and only if the set-valued map

$$M_2(z) := \begin{pmatrix} P_1(z) \\ (z, -P_2(z)) \end{pmatrix} - D_1 \times \operatorname{gph} Q$$

is metrically subregular at $(\bar{z}, (0, 0, 0))$.

Proof. Assume without loss of generality that the image space $\mathbb{R}^{s_1} \times \mathbb{R}^{s_2}$ of M_1 is equipped with the norm $\|(y_1, y_2)\| = \|y_1\| + \|y_2\|$, whereas we use the norm $\|(y_1, z, y_2)\| = \|y_1\| + \|z\| + \|y_2\|$ for the image space $\mathbb{R}^{s_1} \times \mathbb{R}^d \times \mathbb{R}^{s_2}$ of M_2. If M_2 is metrically subregular at $(\bar{z}, (0, 0, 0))$, then there are a neighborhood W of $\bar{z}$ and a constant κ such that for all $z \in W$, we have

$$\begin{aligned} \mathrm{d}(z, \Omega) &\leq \kappa \mathrm{d}\big((0, 0, 0), M_2(z)\big) \\ &= \kappa\big(\mathrm{d}(P_1(z), D_1) + \inf\{\|z - \tilde{z}\| + \| - P_2(z) - \tilde{y}\| \,|\, (\tilde{z}, \tilde{y}) \in \operatorname{gph} Q\}\big) \\ &\leq \kappa\big(\mathrm{d}(P_1(z), D_1) + \inf\{\| - P_2(z) - \tilde{y}\| \,|\, \tilde{y} \in Q(z)\}\big) = \kappa \mathrm{d}\big((0, 0), M_1(z)\big), \end{aligned}$$

which shows the metric subregularity of M_1. Now assume that M_1 is metrically subregular at $(\bar{z}, (0, 0))$, and hence we can find a radius $r > 0$ and a real κ such that

$$\mathrm{d}(z, \Omega) \leq \kappa \mathrm{d}\big((0, 0), M_1(z)\big) \; \forall z \in \mathcal{B}(\bar{z}; r).$$

Further, assume that P_1, P_2 are Lipschitz with modulus L on $\mathcal{B}(\bar{z}; r)$, and consider $z \in \mathcal{B}(\bar{z}; r/(2 + L))$. Since $\operatorname{gph} Q$ is closed, there are $(\tilde{z}, \tilde{y}) \in \operatorname{gph} Q$ with

$$\|z - \tilde{z}\| + \| - P_2(z) - \tilde{y}\| = \mathrm{d}\big((z, -P_2(z)), \operatorname{gph} Q\big).$$

Then

$$\|z - \tilde{z}\| \leq \mathrm{d}\big((z, -P_2(z)), \operatorname{gph} Q\big) \leq \|z - \bar{z}\| + \| - P_2(z) + P_2(\bar{z})\| \leq (1 + L)\|z - \bar{z}\|,$$

implying $\|\bar z-\tilde z\|\le \|\bar z-z\|+\|z-\tilde z\|\le (2+L)\|z-\bar z\|\le r$ and

$$\begin{aligned}
\mathrm{d}(\tilde z,\Omega) &\le \kappa \mathrm{d}\big((0,0),M_1(\tilde z)\big)=\kappa\Big(\mathrm{d}(P_1(\tilde z),D_1)+\mathrm{d}\big(-P_2(\tilde z),Q(\tilde z)\big)\Big)\\
&\le \kappa\big(\mathrm{d}(P_1(\tilde z),D_1)+\|-P_2(\tilde z)-\tilde y\|\big)\\
&\le \kappa\big(2L\|z-\tilde z\|+\mathrm{d}(P_1(z),D_1)+\|-P_2(z)-\tilde y\|\big).
\end{aligned}$$

Taking into account $\mathrm{d}(z,\Omega)\le \mathrm{d}(\tilde z,\Omega)+\|z-\tilde z\|$, we arrive at

$$\begin{aligned}
\mathrm{d}(z,\Omega) &\le \kappa\max\{2L+\frac{1}{\kappa},1\}\big(\mathrm{d}(P_1(z),D_1)+\|z-\tilde z\|+\|-P_2(z)-\tilde y\|\big)\\
&= \kappa\max\{2L+\frac{1}{\kappa},1\}\mathrm{d}\big((0,0,0),M_2(z)\big),
\end{aligned}$$

establishing metric subregularity of M_2 at $(\bar z,(0,0,0))$. □

Since the metric subregularity of the set-valued map $M(z):=P(z)-D$ at $(\bar z,0)$ implies GACQ holding at $\bar z$ (see e.g., [23, Proposition 1]), it can serve as a constraint qualification. Following [17, Definition 3.2], we define it as a constraint qualification below.

Definition 5 (metric subregularity constraint qualification). *Let $P(\bar z)\in D$. 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 $(\bar z,0)$, or equivalently the perturbed set-valued map $M^{-1}(w):=\{z|w\in P(z)-D\}$ is calm at $(0,\bar z)$.*

There exist several sufficient conditions for MSCQ in the literature. Here are the two most frequently used ones. The first case is when the linear CQ holds, i.e., when P is affine and D is the union of finitely many polyhedral convex sets. The second case is when the no nonzero abnormal multiplier constraint qualification (NNAMCQ) holds at $\bar z$ (see e.g., [36]):

$$\nabla P(\bar z)^T\lambda=0,\ \lambda\in N_D(P(\bar z)) \quad\Longrightarrow\quad \lambda=0. \tag{7}$$

It is known that NNAMCQ is equivalent to MFCQ in the case of standard nonlinear programming. Condition (7) appears under different terminologies in the literature; e.g., while it is called NNAMCQ in [36], it is referred to as generalized MFCQ (GMFCQ) in [12].

The linear CQ and NNAMCQ may be still too strong for some problems to hold. Recently, some new constraint qualifications for standard nonlinear programs that are stronger than MSCQ and weaker than the linear CQ and/or NNAMCQ have been introduced in the literature; see e.g., [3, 4]. These CQs include the relaxed constant positive linear dependence condition (RCPLD) (see [25, Theorem 4.2]), the constant rank of the subspace component condition (CRSC) (see [25, Corollary 4.1]), and the quasinormality [24, Theorem 5.2].

In this paper, we will use the following sufficient conditions.

Theorem 1. *Let $\bar z\in\Omega$ where Ω is defined as in (4). MSCQ holds at $\bar z$ if one of the following conditions is fulfilled:*

- *First-order sufficient condition for metric subregularity (FOSCMS) for the system $P(z)\in D$ with P smooth, cf. [16, Corollary 1] : for every $0\neq w\in T^{\mathrm{lin}}_\Omega(\bar z)$, one has*

$$\nabla P(\bar z)^T\lambda=0,\ \lambda\in N_D(P(\bar z);\nabla P(\bar z)w) \quad\Longrightarrow\quad \lambda=0.$$

- *Second-order sufficient condition for metric subregularity (SOSCMS) for the inequality system $P(z) \in \mathbb{R}^s_-$ with P twice Fréchet differentiable at $\bar z$, cf. [13, Theorem 6.1]: For every $0 \neq w \in T^{\rm lin}_\Omega(\bar z)$, one has*

$$\nabla P(\bar z)^T\lambda = 0,\ \lambda \in N_{\mathbb{R}^s_-}(P(\bar z)),\ w^T\nabla^2(\lambda^T P)(\bar z)w \geq 0 \quad\Longrightarrow\quad \lambda = 0.$$

In the case $T^{\rm lin}_\Omega(\bar z) = \{0\}$, FOSCMS is satisfied automatically. By definition of the linearized cone (5), $T^{\rm lin}_\Omega(\bar z) = \{0\}$ means that

$$\nabla P(\bar z)w = \xi, \quad \xi \in T_D(P(\bar z)) \implies w = 0.$$

By the graphical derivative criterion for strong metric subregularity [10, Theorem 4E.1], this is equivalent to saying that the set-valued map $M(z) = P(z) - D$ is strongly metrically subregular (or equivalently, its inverse is isolated calm) at $(\bar z, 0)$. When the set D is convex, by the relationship between the limiting normal cone and its directional version in Proposition 2,

$$N_D(P(\bar z); \nabla P(\bar z)w) = N_D(P(\bar z)) \cap \{\nabla P(\bar z)w\}^\perp.$$

Consequently, in the case where $T^{\rm lin}_\Omega(\bar z) \neq \{0\}$ and D is convex, FOSCMS reduces to NNAMCQ. Indeed, suppose that $\nabla P(\bar z)^T\lambda = 0$ and $\lambda \in N_D(P(\bar z))$. Then $\lambda^T(\nabla P(\bar z)w) = 0$. Hence $\lambda \in N_D(P(\bar z); \nabla P(\bar z)w)$, which implies from FOSCMS that $\lambda = 0$. Hence for convex D, FOSCMS is equivalent to saying that either the strong metric subregularity or the NNAMCQ (7) holds at $(\bar z, 0)$. In the case of an inequality system $P(z) \leq 0$ and $T^{\rm lin}_\Omega(\bar z) \neq \{0\}$, SOSCMS is obviously weaker than NNAMCQ.

In many situations, the constraint system $P(z) \in D$ can be splitted into two parts such that one part can be easily verified to satisfy MSCQ. For example,

$$P(z) = (P_1(z), P_2(z)) \in D = D_1 \times D_2, \tag{8}$$

where $P_i : \mathbb{R}^d \to \mathbb{R}^{s_i}$ are smooth and $D_i \subseteq \mathbb{R}^{s_i}$, $i = 1, 2$ are closed, and for one part, let us say $P_2(z) \in D_2$, it is known in advance that the map $P_2(\cdot) - D_2$ is metrically subregular at $(\bar z, 0)$. In this case, the following theorem is useful.

Theorem 2. *Let $P(\bar z) \in D$ with P smooth and D closed, and assume that P and D can be written in the form* (8) *such that the set-valued map $P_2(z) - D_2$ is metrically subregular at $(\bar z, 0)$. Further assume for every $0 \neq w \in T^{\rm lin}_\Omega(\bar z)$, one has*

$$\nabla P_1(\bar z)^T\lambda^1 + \nabla P_2(\bar z)^T\lambda^2 = 0,\ \lambda^i \in N_{D_i}(P_i(\bar z); \nabla P_i(\bar z)w)\ i = 1, 2 \implies \lambda^1 = 0.$$

Then MSCQ holds at $\bar z$ for the system $P(z) \in D$.

Proof. Let the set-valued maps M, $M_i (i = 1, 2)$ be given by $M(z) := P(z) - D$ and $M_i(z) = P_i(z) - D_i (i = 1, 2)$, respectively. Since P_1 is assumed to be smooth, it is also Lipschitz near $\bar z$ and thus M_1 has the Aubin property around $(\bar z, 0)$. Consider any direction $0 \neq w \in T^{\rm lin}_\Omega(\bar z)$. By [14, Definition 2(3.)], the limit set critical for directional metric regularity $\mathrm{Cr}_{\mathbb{R}^{s_1}} M((\bar z, 0); w)$ with respect to w and $\mathbb{R}^{s_1}$ at $(\bar z, 0)$ is defined as the collection of all elements $(v, z^*) \in \mathbb{R}^s \times \mathbb{R}^d$ such that there are sequences $t_k \downarrow 0$, $(w_k, v_k, z^*_k) \to (w, v, z^*)$, $\lambda_k \in \mathcal{S}_{\mathbb{R}^s}$ and a real $\beta > 0$ such that $(-z^*_k, \lambda_k) \in \widehat N_{\mathrm{gph}\, M}(\bar z + t_k w_k, t_k v_k)$ and $\|\lambda^1_k\| \geq \beta$ for all k, where $\lambda_k = (\lambda^1_k, \lambda^2_k) \in \mathbb{R}^{s_1} \times \mathbb{R}^{s_2}$. We claim that $(0, 0) \notin \mathrm{Cr}_{\mathbb{R}^{s_1}} M((\bar z, 0); w)$. Assume on the contrary that $(0,0) \in \mathrm{Cr}_{\mathbb{R}^{s_1}} M((\bar{z},0);w)$ and consider the corresponding sequences $(t_k, w_k, v_k, z_k^*, \lambda_k)$. The sequence λ_k is bounded and by passing to a subsequence we can assume that λ_k converges to some $\lambda = (\lambda^1, \lambda^2)$ satisfying $\|\lambda^1\| \geq \beta > 0$. Since $(-z_k^*, \lambda_k) \in \widehat{N}_{\mathrm{gph}\, M}(\bar{z} + t_k w_k, t_k v_k)$, it follows from [34, Exercise 6.7] that $-\lambda_k \in \widehat{N}_D(P(\bar{z} + t_k w_k) - t_k v_k)$ and $-z_k^* = \nabla P(\bar{z} + t_k w_k)^T(-\lambda_k)$ which implies that $-\lambda \in N_D(P(\bar{z}); \nabla P(\bar{z})w)$ and $0 = \nabla P(\bar{z})^T(-\lambda) = \nabla P_1(\bar{z})^T(-\lambda^1) + \nabla P_2(\bar{z})^T(-\lambda^2)$. From [16, Lemma 1], we also conclude $-\lambda^i \in N_{D_i}(P_i(\bar{z}); \nabla P_i(\bar{z})w)$, resulting in a contradiction to the assumption of the theorem. Hence our claim $(0,0) \notin \mathrm{Cr}_{\mathbb{R}^{s_1}} M((\bar{z},0);w)$ holds true, and by [14, Lemmas 2, 3, Theorem 6], it follows that M is metrically subregular in direction w at $(\bar{z},0)$, where directional metric subregularity is defined in [14, Definition 1]. Since by definition, M is metrically subregular in every direction $w \notin T_\Omega^{\mathrm{lin}}(\bar{z})$, we conclude from [15, Lemma 2.7] that M is metrically subregular at $(\bar{z},0)$. □

We now discuss some consequences of MSCQ. First, we have the following change of coordinate formula for normal cones.

Proposition 4. *Let $\bar{z} \in \Omega := \{z | P(z) \in D\}$ with P smooth and D closed. Then*

$$\widehat{N}_\Omega(\bar{z}) \supseteq \nabla P(\bar{z})^T \widehat{N}_D(P(\bar{z})). \tag{9}$$

Further, if MSCQ holds at $\bar{z}$ for the system $P(z) \in D$, then

$$\widehat{N}_\Omega(\bar{z}) \subseteq N_\Omega(\bar{z}) \subseteq \nabla P(\bar{z})^T N_D(P(\bar{z})). \tag{10}$$

In particular, if MSCQ holds at $\bar{z}$ for the system $P(z) \in D$ with convex D, then

$$\widehat{N}_\Omega(\bar{z}) = N_\Omega(\bar{z}) = \nabla P(\bar{z})^T N_D(P(\bar{z})). \tag{11}$$

Proof. The inclusion (9) follows from [34, Theorem 6.14]. The first inclusion in (10) follows immediately from the definitions of the regular/limiting normal cone, whereas the second one follows from [22, Theorem 4.1]. When D is convex, the regular normal cone coincides with the limiting normal cone, and hence (11) follows by combining (9) and (10). □

In the case where $D = \mathbb{R}^{s_1}_- \times \{0\}^{s_2}$, it is well-known in nonlinear programming theory that MFCQ or equivalently NNAMCQ is a necessary and sufficient condition for compactness of the set of Lagrange multipliers. In the case where $D \neq \mathbb{R}^{s_1}_- \times \{0\}^{s_2}$, NNAMCQ also implies the boundedness of the multipliers. However MSCQ is weaker than NNAMCQ, and hence the set of Lagrange multipliers may be unbounded if MSCQ holds but NNAMCQ fails. However, Theorem 3 shows that under MSCQ one can extract some uniformly compact subset of the multipliers.

Definition 6 (cf. [18])**.** *Let $\bar{z} \in \Omega := \{z | P(z) \in D\}$ with P smooth and D closed. We say that the* bounded multiplier property *(BMP) holds at $\bar{z}$ for the system $P(z) \in D$, if there are some modulus $\kappa \geq 0$ and some neighborhood W of $\bar{z}$ such that for every $z \in W \cap \Omega$ and every $z^* \in N_\Omega(z)$, there is some $\lambda \in \kappa \|z^*\| \mathcal{B}_{\mathbb{R}^s} \cap N_D(P(z))$ satisfying*

$$z^* = \nabla P(z)^T \lambda.$$

The following theorem gives a sharper upper estimate for the normal cone than (10).

Theorem 3. *Let $\bar{z} \in \Omega := \{z \,|\, P(z) \in D\}$, and assume that MSCQ holds at the point $\bar{z}$ for the system $P(z) \in D$. Let W denote an open neighborhood of $\bar{z}$, and let $\kappa \geq 0$ be a real such that*

$$\mathrm{d}(z, \Omega) \leq \kappa \mathrm{d}(P(z), D) \quad \forall z \in W.$$

Then

$$N_\Omega(z) \subseteq \left\{ z^* \in \mathbb{R}^d \,|\, \exists \lambda \in \kappa \|z^*\| \mathcal{B}_{\mathbb{R}^s} \cap N_D(P(z)) \text{ with } z^* = \nabla P(z)^T \lambda \right\} \quad \forall z \in W.$$

In particular BMP holds at $\bar{z}$ for the system $P(z) \in D$.

Proof. Under the assumption, the set-valued map $M(z) := P(z) - D$ is metrically subregular at $(\bar{z}, 0)$. The definition of the metric subregularity justifies the existence of the open neighborhood W and the number κ in the assumption. Hence for each $z \in M^{-1}(0) \cap W = \Omega \cap W$, the map M is also metrically subregular at $(z, 0)$, and by applying [21, Proposition 4.1] we obtain

$$N_\Omega(z) = N_{M^{-1}(0)}(z; 0) \subseteq \{z^* \,|\, \exists \lambda \in \kappa \|z^*\| \mathcal{B}_{\mathbb{R}^s} : (z^*, \lambda) \in N_{\mathrm{gph}\, M}((z, 0); (0, 0))\}.$$

It follows from [34, Exercise 6.7] that

$$N_{\mathrm{gph}\, M}((z, 0); (0, 0)) = N_{\mathrm{gph}\, M}((z, 0)) = \{(z^*, \lambda) \,|\, -\lambda \in N_D(P(z)), z^* = \nabla P(z)^T(-\lambda)\}.$$

Hence the assertion follows. $\square$

3 Failure of MPCC-tailored constraint qualifications for problem (MPCC)

In this section, we discuss difficulties involved in MPCC-tailored constraint qualifications for the problem (MPCC) by considering the constraint system for problem (MPCC) in the following form

$$\widetilde{\Omega} := \left\{ (x, y, \lambda) : \begin{array}{l} 0 = h(x, y, \lambda) := \phi(x, y) + \nabla g(y)^T \lambda, \\ 0 \geq g(y) \perp -\lambda \leq 0 \\ G(x, y) \leq 0 \end{array} \right\},$$

where $\phi : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ and $G : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p$ are continuously differentiable and $g : \mathbb{R}^m \to \mathbb{R}^q$ is twice continuously differentiable.

Given a triple $(\bar{x}, \bar{y}, \bar{\lambda}) \in \widetilde{\Omega}$ we define the following index sets of active constraints:

$$\begin{aligned} \mathcal{I}_g &:= \mathcal{I}_g(\bar{y}, \bar{\lambda}) := \{i \in \{1, \dots, q\} \,|\, g_i(\bar{y}) = 0, \bar{\lambda}_i > 0\}, \\ \mathcal{I}_\lambda &:= \mathcal{I}_\lambda(\bar{y}, \bar{\lambda}) := \{i \in \{1, \dots, q\} \,|\, g_i(\bar{y}) < 0, \bar{\lambda}_i = 0\}, \\ \mathcal{I}_0 &:= \mathcal{I}_0(\bar{y}, \bar{\lambda}) := \{i \in \{1, \dots, q\} \,|\, g_i(\bar{y}) = 0, \bar{\lambda}_i = 0\}, \\ \mathcal{I}_G &:= \mathcal{I}_G(\bar{x}, \bar{y}) := \{i \in \{1, \dots, p\} \,|\, G_i(\bar{x}, \bar{y}) = 0\}. \end{aligned}$$

Definition 7 ([35])**.** *We say that MPCC-MFCQ holds at $(\bar{x}, \bar{y}, \bar{\lambda})$ if the gradient vectors*

$$\nabla h_i(\bar{x}, \bar{y}, \bar{\lambda}), i = 1, \dots, m, \;\; (0, \nabla g_i(\bar{y}), 0), i \in \mathcal{I}_g \cup \mathcal{I}_0, \;\; (0, 0, e_i), i \in \mathcal{I}_\lambda \cup \mathcal{I}_0, \tag{12}$$

where e_i denotes the unit vector with the ith component equal to 1, are linearly independent, and there exists a vector $(d_x, d_y, d_\lambda) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^q$ orthogonal to the vectors in (12) and such that

$$\nabla G_i(\bar{x}, \bar{y})(d_x, d_y) < 0, i \in \mathcal{I}_G.$$

We say that MPCC-LICQ holds at $(\bar{x}, \bar{y}, \bar{\lambda})$ if the gradient vectors

$$\nabla h_i(\bar{x}, \bar{y}, \bar{\lambda}), i = 1, \dots, m,\ (0, \nabla g_i(\bar{y}), 0), i \in \mathcal{I}_g \cup \mathcal{I}_0,\ (0, 0, e_i), i \in \mathcal{I}_\lambda \cup I_0,\ (\nabla G_i(\bar{x}, \bar{y}), 0), i \in \mathcal{I}_G$$

are linearly independent.

MPCC-MFCQ implies that for every partition (β_1, β_2) of $\mathcal{I}_0$, the branch

$$\begin{cases} \phi(x, y) + \nabla g(y)^T \lambda = 0, \\ g_i(y) = 0, \lambda_i \geq 0, i \in \mathcal{I}_g,\ \lambda_i = 0, g_i(y) \leq 0, i \in \mathcal{I}_\lambda, \\ g_i(y) = 0, \lambda_i \geq 0, i \in \beta_1,\ g_i(y) \leq 0, \lambda_i = 0, i \in \beta_2, \\ G(x, y) \leq 0 \end{cases} \tag{13}$$

satisfies MFCQ at $(\bar{x}, \bar{y}, \bar{\lambda})$.

We now show that MPCC-MFCQ never holds for (MPCC) if the lower level program has more than one multiplier.

Proposition 5. *Let $(\bar{x}, \bar{y}, \bar{\lambda}) \in \widetilde{\Omega}$, and assume that there exists a second multiplier $\hat{\lambda} \neq \bar{\lambda}$ such that $(\bar{x}, \bar{y}, \hat{\lambda}) \in \widetilde{\Omega}$. Then for every partition (β_1, β_2) of $\mathcal{I}_0$, the branch (13) does not fulfill MFCQ at $(\bar{x}, \bar{y}, \bar{\lambda})$.*

Proof. Since $\nabla g(\bar{y})^T(\hat{\lambda} - \bar{\lambda}) = 0$, $(\hat{\lambda} - \bar{\lambda})_i \geq 0$, $i \in \mathcal{I}_\lambda \cup \beta_2$ and $\hat{\lambda} - \bar{\lambda} \neq 0$, the assertion follows immediately. □

Since MPCC-MFCQ is stronger than the MPCC-LICQ, we have the following corollary immediately.

Corollary 1. *Let $(\bar{x}, \bar{y}, \bar{\lambda}) \in \widetilde{\Omega}$, and assume that there exists a second multiplier $\hat{\lambda} \neq \bar{\lambda}$ such that $(\bar{x}, \bar{y}, \hat{\lambda}) \in \widetilde{\Omega}$. Then MPCC-LICQ fails at $(\bar{x}, \bar{y}, \bar{\lambda})$.*

It is worth noting that our result in Proposition 5 is only valid under the assumption that $g(y)$ is independent of x. In the case of bilevel programming where the lower level problem has a constraint dependent of the upper level variable, an example given in [28, Example 4.10] shows that if the multiplier is not unique, then the corresponding MPCC-MFCQ may hold at some of the multipliers and fail to hold at others.

Definition 8 (see e.g., [12])**.** *Let $(\bar{x}, \bar{y}, \bar{\lambda})$ be feasible for (MPCC). We say MPCC-ACQ and MPCC-GCQ hold if*

$$T_{\widetilde{\Omega}}(\bar{x}, \bar{y}, \bar{\lambda}) = T^{\text{lin}}_{\text{MPCC}}(\bar{x}, \bar{y}, \bar{\lambda}) \text{ and } \widehat{N}_{\widetilde{\Omega}}(\bar{x}, \bar{y}, \bar{\lambda}) = (T^{\text{lin}}_{\text{MPCC}}(\bar{x}, \bar{y}, \bar{\lambda}))^\circ,$$

respectively, where

$$T^{\text{lin}}_{\text{MPCC}}(\bar{x}, \bar{y}, \bar{\lambda}) := \left\{ (u, v, \mu) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^q \,\middle|\, \begin{array}{l} \nabla_x \phi(\bar{x}, \bar{y})u + \nabla_y(\phi + \nabla_y(\lambda^T g))(\bar{x}, \bar{y})v + \nabla g(\bar{y})^T \mu = 0, \\ \nabla g_i(\bar{y})v = 0, i \in \mathcal{I}_g,\ \mu_i = 0, i \in \mathcal{I}_\lambda, \\ \nabla g_i(\bar{y})v \leq 0, \mu_i \geq 0, \mu_i \nabla g_i(\bar{y})v = 0, i \in \mathcal{I}_0, \\ \nabla G_i(\bar{x}, \bar{y})(u, v) \leq 0, i \in \mathcal{I}_G \end{array} \right\}$$

is the MPEC linearized cone at $(\bar{x}, \bar{y}, \bar{\lambda})$.

Note that MPCC-ACQ and MPCC-GCQ are the GACQ and GGCQ for the equivalent formulation of the set $\widetilde{\Omega}$ in the form of $P(z) \in D$, with D involving the complementarity set

$$D_{cc} := \{(a, b) \in \mathbb{R}^q_- \times \mathbb{R}^q_- | a^T b = 0\},$$

respectively. MPCC-MFCQ implies MPCC-ACQ (cf. [11]) and from definition, it is easy to see that MPCC-ACQ is stronger than MPCC-GCQ. Under MPCC-GCQ, it is known that a local optimal solution of (MPCC) must be a M-stationary point ([12, Theorem 14]). Although MPCC-GCQ is weaker than most of other MPCC-tailored constraint qualifications, the following example shows that the constraint qualification MPCC-GCQ still can be violated when the multiplier for the lower level is not unique. In contrast to [1, Example 6], all the constraints are convex .

Example 1. *Consider the MPEC*

$$\begin{aligned} \min_{x,y} \quad & F(x, y) := x_1 - \frac{3}{2}y_1 + x_2 - \frac{3}{2}y_2 - y_3 \\ s.t. \quad & 0 \in \phi(x, y) + N_\Gamma(y), \\ & G_1(x, y) = G_1(x) := -x_1 - 2x_2 \le 0, \\ & G_2(x, y) = G_2(x) := -2x_1 - x_2 \le 0, \end{aligned} \tag{14}$$

where

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

Let $\bar{x} = (0, 0)$ and $\bar{y} = (0, 0, 0)$. The lower level inequality system $g(y) \le 0$ is convex satisfying the Slater condition, and therefore y is a solution to the parametric generalized equation (14) if and only if $y' = y$ is a global minimizer of the optimization problem: $\min_{y'} \ \langle \phi(x, y), y' \rangle$ s.t. $y' \in \Gamma$, and if and only if there is a multiplier λ fulfilling KKT-conditions

$$\begin{pmatrix} y_1 - x_1 + \lambda_1 y_1 \\ y_2 - x_2 + \lambda_2 y_2 \\ -1 + \lambda_1 + \lambda_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \tag{15}$$

$$0 \ge y_3 + \frac{1}{2}y_1^2 \perp -\lambda_1 \le 0,$$

$$0 \ge y_3 + \frac{1}{2}y_2^2 \perp -\lambda_2 \le 0.$$

Let $\mathcal{F} := \{x \,|\, G_1(x) \le 0, G_2(x) \le 0\}$. Then $\mathcal{F} = \mathcal{F}_1 \cup \mathcal{F}_2 \cup \mathcal{F}_3$ where

$$\begin{aligned} \mathcal{F}_1 &:= \left\{(x_1, x_2) \in \mathbb{R}^2 \,|\, 2|x_2| \le x_1\right\}, \\ \mathcal{F}_2 &:= \left\{(x_1, x_2) \in \mathbb{R}^2 \,|\, \frac{x_1}{2} \le x_2 \le 2x_1\right\}, \\ \mathcal{F}_3 &:= \left\{(x_1, x_2) \in \mathbb{R}^2 \,|\, 2|x_1| \le x_2\right\}. \end{aligned}$$

Straightforward calculations yield that for each $x \in \mathcal{F}$, there exists a unique solution $y(x)$, which is given by

$$y(x) = \begin{cases} (\frac{x_1}{2}, x_2, -\frac{1}{8}x_1^2) & \text{if } x \in \mathcal{F}_1, \\ (\frac{x_1+x_2}{3}, \frac{x_1+x_2}{3}, -\frac{1}{18}(x_1 + x_2)^2) & \text{if } x \in \mathcal{F}_2, \\ (x_1, \frac{x_2}{2}, -\frac{1}{8}x_2^2) & \text{if } x \in \mathcal{F}_3. \end{cases}$$

Further, at $\bar{x} = (0,0)$ we have $y(\bar{x}) = (0,0,0)$ and the set of the multipliers is

$$\Lambda := \{\lambda \in \mathbb{R}^2_+ | \lambda_1 + \lambda_2 = 1\},$$

while for all $x \neq (0,0)$ the gradients of the lower level constraints active at $y(x)$ are linearly independent and the unique multiplier is given by

$$\lambda(x) = \begin{cases} (1,0) & \text{if } x \in \mathcal{F}_1, \\ (\frac{2x_1 - x_2}{x_1 + x_2}, \frac{2x_2 - x_1}{x_1 + x_2}) & \text{if } x \in \mathcal{F}_2, \\ (0,1) & \text{if } x \in \mathcal{F}_3. \end{cases} \tag{16}$$

Since

$$F(x, y(x)) = \begin{cases} \frac{1}{4}x_1 - \frac{1}{2}x_2 + \frac{1}{8}x_1^2 & \text{if } x \in \mathcal{F}_1, \\ \frac{1}{18}(x_1 + x_2)^2 & \text{if } x \in \mathcal{F}_2, \\ \frac{1}{4}x_2 - \frac{1}{2}x_1 + \frac{1}{8}x_2^2 & \text{if } x \in \mathcal{F}_3, \end{cases}$$

and $\mathcal{F} = \mathcal{F}_1 \cup \mathcal{F}_2 \cup \mathcal{F}_3$, we see that $(\bar{x}, \bar{y})$ is a globally optimal solution of the MPEC.

The original problem is equivalent to the following MPCC:

$$\begin{aligned} \min_{x,y,\lambda} \quad & x_1 - \frac{3}{2}y_1 + x_2 - \frac{3}{2}y_2 - y_3 \\ s.t. \quad & x, y, \lambda \text{ fulfill } (15), \\ & -2x_1 - x_2 \leq 0, \\ & -x_1 - 2x_2 \leq 0. \end{aligned}$$

The feasible region of this problem is

$$\widetilde{\Omega} = \bigcup_{\bar{x} \neq x \in \mathcal{F}} \{(x, y(x), \lambda(x))\} \cup (\{(\bar{x}, \bar{y})\} \times \Lambda).$$

Any $(\bar{x}, \bar{y}, \lambda)$ where $\lambda \in \Lambda$ is a globally optimal solution. However it is easy to verify that unless $\lambda_1 = \lambda_2 = 0.5$, any $(\bar{x}, \bar{y}, \lambda)$ is not even a weak stationary point, implying by [12, Theorem 7] that MPCC-GCQ and consequently MPCC-ACQ fails to hold. Now consider $\lambda = (0.5, 0.5)$. The MPEC linearized cone $T^{\text{lin}}_{\text{MPCC}}(\bar{x}, \bar{y}, \lambda)$ is the collection of all (u, v, μ) such that

$$\begin{pmatrix} 1.5v_1 - u_1 \\ 1.5v_2 - u_2 \\ \mu_1 + \mu_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{aligned} & v_3 = 0, \\ & -2u_1 - u_2 \leq 0, \ -u_1 - 2u_2 \leq 0. \end{aligned} \tag{17}$$

Next, we compute the actual tangent cone $T_{\widetilde{\Omega}}(\bar{x}, \bar{y}, \lambda)$. Consider sequences $t_k \downarrow 0$, $(u^k, v^k, \mu^k) \to (u, v, \mu)$ such that $(\bar{x}, \bar{y}, \lambda) + t_k(u^k, v^k, \mu^k) \in \widetilde{\Omega}$. If $u^k \neq 0$ for infinitely many k, then $\bar{x} + t_k u^k \neq 0$, and hence $(\bar{y} + t_k v^k, \lambda + t_k \mu^k) = (y(\bar{x} + t_k u^k), \lambda(\bar{x} + t_k u^k))$ for those k. Since $\lambda = (0.5, 0.5)$, it follows from (16) that $\bar{x} + t_k u^k \in \mathcal{F}_2$ for infinitely many k, implying, by passing to a subsequence if necessary,

$$v = \lim_{k \to \infty} \frac{y(\bar{x} + t_k u^k) - \bar{y}}{t_k} = \frac{1}{3}(u_1 + u_2, u_1 + u_2, 0)$$

and

$$\begin{aligned}\mu &= \lim_{k\to\infty} \frac{\lambda(\bar{x}+t_k u^k)-\lambda}{t_k} = \lim_{k\to\infty} \frac{(\frac{2u_1^k-u_2^k}{u_1^k+u_2^k}, \frac{2u_2^k-u_1^k}{u_1^k+u_2^k}) - (0.5, 0.5)}{t_k} \\ &= \lim_{k\to\infty} 1.5\frac{(\frac{u_1^k-u_2^k}{u_1^k+u_2^k}, \frac{u_2^k-u_1^k}{u_1^k+u_2^k})}{t_k}.\end{aligned}$$

Hence, $v_1 = v_2 = \frac{1}{3}(u_1+u_2)$, $v_3 = 0$ and $\mu_1+\mu_2 = 0$. Also from (17), we have $u_1 = u_2$ since $v_1 = v_2$ and the tangent cone $T_{\widetilde{\Omega}}(\bar{x},\bar{y},\lambda)$ is always a subset of the MPEC linearized cone $T^{\text{lin}}_{\text{MPCC}}(\bar{x},\bar{y},\lambda)$ (see e.g., [11, Lemma 3.2]). Further, since $\bar{x}+t_k u^k \in \mathcal{F}_2$, we must have $u_1 \geq 0$. If $u^k = 0$ for all but finitely many k, then we have $v^k = 0$ and $\lambda + t_k\mu^k \in \Lambda$, implying $\mu_1+\mu_2 = 0$. Putting all together, we obtain that the actual tangent cone $T_{\widetilde{\Omega}}(\bar{x},\bar{y},\lambda)$ to the feasible set is the collection of all (u,v,μ) satisfying

$$\begin{aligned}&u_1 = u_2 \geq 0, v_1 = v_2 = \frac{2}{3}u_1,\\ &v_3 = 0, \mu_1+\mu_2 = 0.\end{aligned}$$

Now it is easy to see that $T_{\widetilde{\Omega}}(\bar{x},\bar{y},\lambda) \neq T^{\text{lin}}_{\text{MPCC}}(\bar{x},\bar{y},\lambda)$. Moreover, since both $T_{\widetilde{\Omega}}(\bar{x},\bar{y},\lambda)$ and $T^{\text{lin}}_{\text{MPCC}}(\bar{x},\bar{y},\lambda)$ are convex polyhedral sets, one also has $(T_{\widetilde{\Omega}}(\bar{x},\bar{y},\lambda))^\circ \neq (T^{\text{lin}}_{\text{MPCC}}(\bar{x},\bar{y},\lambda))^\circ$, and thus MPEC-GCQ does not hold for $\lambda = (0.5, 0.5)$ as well.

4 Sufficient condition for MSCQ

As we discussed in the introduction and section 3, there are much difficulties involved in formulating an MPEC as (MPCC). In this section, we turn our attention to problem (MPEC) with the constraint system defined in the following form

$$\Omega := \left\{ (x,y) : \begin{array}{l} 0 \in \phi(x,y) + \widehat{N}_\Gamma(y) \\ G(x,y) \leq 0 \end{array} \right\}, \tag{18}$$

where $\Gamma := \{y \in \mathbb{R}^m | g(y) \leq 0\}$, $\phi : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ and $G : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p$ are continuously differentiable, and $g : \mathbb{R}^m \to \mathbb{R}^q$ is twice continuously differentiable. Let $(\bar{x},\bar{y})$ be a feasible solution of problem (MPEC). We assume that MSCQ is fulfilled for the constraint $g(y) \leq 0$ at $\bar{y}$. Then by definition, MSCQ also holds for all points $y \in \Gamma$ near $\bar{y}$, and by Proposition 4 the following equations hold for such y:

$$N_\Gamma(y) = \widehat{N}_\Gamma(y) = \nabla g(y)^T N_{\mathbb{R}^q_-}(g(y)),$$

where $N_{\mathbb{R}^q_-}(g(y)) = \{\lambda \in \mathbb{R}^q_+ \mid \lambda_i = 0, i \notin \mathcal{I}(y)\}$ and $\mathcal{I}(y) := \{i \in \{1,\ldots,q\} \mid g_i(y) = 0\}$ is the index set of active inequality constraints.

For the sake of simplicity, we do not include equality constraints in either the upper or the lower level constraints. We are using MSCQ as the basic constraint qualification for both the upper and the lower level constraints, and this allows us to write an equality constraint $h(x) = 0$ equivalently as two inequality constraints $h(x) \leq 0$, $-h(x) \leq 0$ without affecting MSCQ.

In the case where Γ is convex, MSCQ is proposed in [37] as a constraint qualification for the M-stationary condition. Two types of sufficient conditions were given for MSCQ. One is the case when all involved functions are affine, and the other is when metric regularity holds. In this section, by making use of FOSCMS for the split system in Theorem 2, we derive some new sufficient condition for MSCQ for the constraint system (18). Applying the new constraint qualification to the problem in Example 1, we show that in contrast to the MPCC reformulation, under which even the weakest constraint qualification MPEC-GCQ fails at $(\bar{x}, \bar{y}, \lambda)$ for all multipliers λ, the MSCQ holds at $(\bar{x}, \bar{y})$ for the original formulation.

In order to apply FOSCMS in Theorem 2, we need to calculate the linearized cone $T^{\rm lin}_{\Omega}(\bar{z})$, and consequently we need to calculate the tangent cone $T_{{\rm gph}\widehat{N}_\Gamma}(\bar{y}, -\phi(\bar{x}, \bar{y}))$. We now perform this task. First we introduce some notations. Given vectors $y \in \Gamma$, $y^* \in \mathbb{R}^m$, consider the *set of multipliers*

$$\Lambda(y, y^*) := \big\{\lambda \in \mathbb{R}^q_+ \big| \; \nabla g(y)^T\lambda = y^*, \lambda_i = 0, i \notin \mathcal{I}(y)\big\}. \tag{19}$$

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

$$I^+(\lambda) := \big\{i \in \{1, \ldots, q\} \big| \; \lambda_i > 0\big\} \;\text{ for }\; \lambda = (\lambda_1, \ldots, \lambda_q) \in \mathbb{R}^q_+. \tag{20}$$

Denote by $\mathcal{E}(y, y^*)$ the collection of all the *extreme points* of the closed and convex set of multipliers $\Lambda(y, y^*)$, and recall that $\lambda \in \Lambda(y, y^*)$ belongs to $\mathcal{E}(y, y^*)$ if and only if the family of gradients $\{\nabla g_i(y) \,|\, i \in I^+(\lambda)\}$ is linearly independent. Further $\mathcal{E}(y, y^*) \neq \emptyset$ if and only if $\Lambda(y, y^*) \neq \emptyset$. To proceed further, recall the notion of the *critical cone* to Γ at $(y, y^*) \in {\rm gph}\,\widehat{N}_\Gamma$ given by $K(y, y^*) := T_\Gamma(y) \cap \{y^*\}^\perp$, and define the *multiplier set in a direction* $v \in K(y, y^*)$ by

$$\Lambda(y, y^*; v) := \arg\max\big\{v^T\nabla^2(\lambda^T g)(y)v \,|\, \lambda \in \Lambda(y, y^*)\big\}. \tag{21}$$

Note that $\Lambda(y, y^*; v)$ is the solution set of a linear optimization problem, and therefore $\Lambda(y, y^*; v) \cap \mathcal{E}(y, y^*) \neq \emptyset$ whenever $\Lambda(y, y^*; v) \neq \emptyset$. Further, we denote the corresponding optimal function value by

$$\theta(y, y^*; v) := \max\big\{v^T\nabla^2(\lambda^T g)(y)v \,|\, \lambda \in \Lambda(y, y^*)\big\}. \tag{22}$$

The critical cone to Γ has the following two expressions.

Proposition 6. *(see e.g., [17, Proposition 4.3]) Suppose that MSCQ holds for the system $g(y) \in \mathbb{R}^q_-$ at y. Then the critical cone to Γ at $(y, y^*) \in {\rm gph}\,\widehat{N}_\Gamma$ is a convex polyhedron that can be explicitly expressed as*

$$K(y, y^*) = \{v | \nabla g(y)v \in T_{\mathbb{R}^q_-}(g(y)), v^T y^* = 0\}.$$

Moreover, for any $\lambda \in \Lambda(y, y^*)$,

$$K(y, y^*) = \left\{v \,|\nabla g(y)v \left\{\begin{array}{l} = 0 \;\textit{if}\; \lambda_i > 0 \\ \leq 0 \;\textit{if}\; \lambda_i = 0 \end{array}\right\}\right\}.$$

Based on the expression for the critical cone, it is easy to see that the normal cone to the critical cone has the following expression.

Lemma 1. *[19, Lemma 1] Assume MSCQ holds at y for the system $g(y) \in \mathbb{R}^q_-$. Let $v \in K(y, y^*), \lambda \in \Lambda(y, y^*)$. Then*

$$N_{K(y,y^*)}(v) = \{\nabla g(y)^T\mu | \mu^T\nabla g(y)v = 0, \mu \in T_{N_{\mathbb{R}^q_-}(g(y))}(\lambda)\}.$$

We are now ready to calculate the tangent cone to the graph of $\widehat{N}_\Gamma$. This result will be needed in the sufficient condition for MSCQ, and it is also of an independent interest. The first equation in the formula (23) was first shown in [19, Theorem 1] under the extra assumption that the metric regularity holds locally uniformly except for $\bar{y}$, whereas in [6] this extra assumption was removed.

Theorem 4. *Given $\bar{y} \in \Gamma$, assume that MSCQ holds at $\bar{y}$ for the system $g(y) \in \mathbb{R}^q_-$. Then there are a real $\kappa > 0$ and a neighorhood V of $\bar{y}$ such that for any $y \in \Gamma \cap V$ and any $y^* \in \widehat{N}_\Gamma(y)$, the tangent cone to the graph of $\widehat{N}_\Gamma$ at (y, y^*) can be calculated by*

$$\begin{aligned} &T_{\mathrm{gph}\,\widehat{N}_\Gamma}(y, y^*) \qquad (23)\\ &= \big\{(v, v^*) \in \mathbb{R}^{2m} \big|\ \exists\, \lambda \in \Lambda(y, y^*; v) \ \ \textit{with} \ \ v^* \in \nabla^2(\lambda^T g)(y)v + N_{K(y,y^*)}(v)\big\}\\ &= \big\{(v, v^*) \in \mathbb{R}^{2m} \big|\ \exists\, \lambda \in \Lambda(y, y^*; v) \cap \kappa\|y^*\|\mathcal{B}_{\mathbb{R}^q} \ \ \textit{with} \ \ v^* \in \nabla^2(\lambda^T g)(y)v + N_{K(y,y^*)}(v)\big\}, \end{aligned}$$

where the critical cone $K(y, y^*)$ and the normal cone $N_{K(y,y^*)}(v)$ can be calculated as in Proposition 6 and Lemma 1, respectively, and the set $\mathrm{gph}\,\widehat{N}_\Gamma$ is geometrically derivable at (y, y^*).

Proof. Since MSCQ holds at $\bar{y}$ for the system $g(y) \in \mathbb{R}^q_-$, we can find an open neighborhood V of $\bar{y}$ and and a real $\kappa > 0$ such that

$$\mathrm{d}(y, \Gamma) \le \kappa \mathrm{d}(g(y), \mathbb{R}^q_-)\ \forall y \in V, \qquad (24)$$

which means that MSCQ holds at every $y \in \Gamma \cap V$. Therefore $K(y, y^*)$ and $N_{K(y,y^*)}(v)$ can be calculated as in Proposition 6 and Lemma 1, respectively. By the proof of the first part of [19, Theorem 1], we obtain that for every $y^* \in \widehat{N}_\Gamma(y)$,

$$\begin{aligned} &\big\{(v, v^*) \in \mathbb{R}^{2m} \big|\ \exists\, \lambda \in \Lambda(y, y^*; v) \cap \kappa\|y^*\|\mathcal{B}_{\mathbb{R}^q} \ \ \text{with} \ \ v^* \in \nabla^2(\lambda^T g)(y)v + N_{K(y,y^*)}(v)\big\}\\ &\subseteq \big\{(v, v^*) \in \mathbb{R}^{2m} \big|\ \exists\, \lambda \in \Lambda(y, y^*; v) \ \ \text{with} \ \ v^* \in \nabla^2(\lambda^T g)(y)v + N_{K(y,y^*)}(v)\big\}\\ &\subseteq \big\{(v, v^*) \in \mathbb{R}^{2m} \big|\ \lim_{t \downarrow 0} t^{-1}\mathrm{d}\big((y + tv, y^* + tv^*), \mathrm{gph}\,\widehat{N}_\Gamma\big) = 0\big\}\\ &\subseteq T_{\mathrm{gph}\,\widehat{N}_\Gamma}(y, y^*). \end{aligned}$$

We now show the reversed inclusion

$$\begin{aligned} &T_{\mathrm{gph}\,\widehat{N}_\Gamma}(y, y^*) \qquad (25)\\ &\subseteq \big\{(v, v^*) \in \mathbb{R}^{2m} \big|\ \exists\, \lambda \in \Lambda(y, y^*; v) \cap \kappa\|y^*\|\mathcal{B}_{\mathbb{R}^q} \ \ \text{with} \ \ v^* \in \nabla^2(\lambda^T g)(y)v + N_{K(y,y^*)}(v)\big\}. \end{aligned}$$

Although the proof technique is essentially the same as [19, Theorem 1], for completeness we provide the detailed proof. Consider $y \in \Gamma \cap V$, $y^* \in \widehat{N}_\Gamma(y)$ and let $(v, v^*) \in T_{\mathrm{gph}\,\widehat{N}_\Gamma}(y, y^*)$. Then by definition of the tangent cone, there exist sequences $t_k \downarrow 0, v_k \to v, v_k^* \to v^*$ such that $y_k^* := y^* + t_k v_k^* \in \widehat{N}_\Gamma(y_k)$, where $y_k := y + t_k v_k$. By passing to a subsequence if necessary we can assume that $y_k \in V\ \forall k$ and that there is some index set $\widetilde{\mathcal{I}} \subseteq \mathcal{I}(y)$ such that $\mathcal{I}(y_k) = \widetilde{\mathcal{I}}$ hold for all k. For every $i \in \mathcal{I}(y)$, we have

$$g_i(y_k) = g_i(y) + t_k \nabla g_i(y)v_k + o(t_k) = t_k \nabla g_i(y)v_k + o(t_k) \begin{cases} = 0 & \text{if } i \in \widetilde{\mathcal{I}},\\ \le 0 & \text{if } i \in \mathcal{I}(y) \setminus \widetilde{\mathcal{I}}. \end{cases} \qquad (26)$$

Dividing by t_k and passing to the limit we obtain

$$\nabla g_i(y)v \begin{cases} = 0 & \text{if } i \in \widetilde{\mathcal{I}}, \\ \leq 0 & \text{if } i \in \mathcal{I}(y) \setminus \widetilde{\mathcal{I}}, \end{cases} \tag{27}$$

which means $v \in T_\Gamma^{\text{lin}}(y)$. Since MSCQ holds at every $y \in \Gamma \cap V$, we have that the GACQ holds at y as well, and hence $v \in T_\Gamma(y)$.

Since (24) holds and $y_k \in V$, $y_k^* \in \widehat{N}_\Gamma(y_k) = N_\Gamma(y_k)$, by Theorem 3, there exists a sequence of multipliers $\lambda^k \in \Lambda(y_k, y_k^*) \cap \kappa \|y_k^*\| \mathcal{B}_{\mathbb{R}^q}$ as $k \in \mathbb{N}$. Consequently, we assume that there exists $c_1 \geq 0$ such that $\|\lambda^k\| \leq c_1$ for all k. Let

$$\Psi_{\widetilde{\mathcal{I}}}(y^*) := \{\lambda \in \mathbb{R}^q | \nabla g(y)^T \lambda = y^*, \lambda_i \geq 0, i \in \widetilde{\mathcal{I}}, \lambda_i = 0, i \notin \widetilde{\mathcal{I}}\}. \tag{28}$$

By Hoffman's Lemma, there is some constant β such that for every $y^* \in \mathbb{R}^m$ with $\Psi_{\widetilde{\mathcal{I}}}(y^*) \neq \emptyset$, one has

$$\mathrm{d}(\lambda, \Psi_{\widetilde{\mathcal{I}}}(y^*)) \leq \beta(\|\nabla g(y)^T \lambda - y^*\| + \sum_{i \in \widetilde{\mathcal{I}}} \max\{-\lambda_i, 0\} + \sum_{i \notin \widetilde{\mathcal{I}}} |\lambda_i|) \quad \forall \lambda \in \mathbb{R}^q. \tag{29}$$

Since

$$\nabla g(y)^T \lambda^k - y^* = t_k v_k^* + (\nabla g(y) - \nabla g(y_k))^T \lambda^k$$

and $\|\nabla g(y) - \nabla g(y_k)\| \leq c_2 \|y_k - y\| = c_2 t_k \|v_k\|$ for some $c_2 \geq 0$, by (29) we can find for each k some $\widetilde{\lambda}^k \in \Psi_{\widetilde{\mathcal{I}}}(y^*) \subseteq \Lambda(y, y^*)$ with $\|\widetilde{\lambda}^k - \lambda^k\| \leq \beta t_k(\|v_k^*\| + c_1 c_2 \|v_k\|)$. Taking $\mu^k := (\lambda^k - \widetilde{\lambda}^k)/t_k$ we have that (μ^k) is uniformly bounded. By passing to subsequence if necessary, we assume that (λ^k) and (μ^k) are convergent to some $\lambda \in \Lambda(y, y^*) \cap \kappa \|y^*\| \mathcal{B}_{\mathbb{R}^q}$, and some μ, respectively. Obviously the sequence $(\widetilde{\lambda}^k)$ converges to λ as well. Since $\lambda_i^k = \widetilde{\lambda}_i^k = 0, i \notin \widetilde{\mathcal{I}}$, by virtue of (27) we have ${\mu^k}^T \nabla g(y) v = 0 \ \forall k$, implying

$$\mu \in (\nabla g(y) v)^\perp. \tag{30}$$

Taking into account ${\lambda^k}^T g(y_k) = 0$ and (26), we obtain

$$0 = \lim_{k \to \infty} \frac{{\lambda^k}^T g(y_k)}{t_k} = \lim_{k \to \infty} {\lambda^k}^T \nabla g(y) v_k = {y^*}^T v.$$

Therefore combining the above with $v \in T_\Gamma(y)$, we have

$$v \in K(y, y^*). \tag{31}$$

Further, we have for all $\lambda' \in \Lambda(y, y^*)$, since $\widetilde{\lambda}^k \in \Lambda(y, y^*)$,

$$\begin{aligned} 0 &\leq (\widetilde{\lambda}^k - \lambda')^T g(y_k) = (\widetilde{\lambda}^k - \lambda')^T (g(y) + t_k \nabla g(y) v_k + \frac{1}{2} t_k^2 v_k^T \nabla^2 g(y) v_k + o(t_k^2)) \\ &= (\widetilde{\lambda}^k - \lambda')^T (\frac{1}{2} t_k^2 v_k^T \nabla^2 g(y) v_k + o(t_k^2)). \end{aligned}$$

Dividing by t_k^2 and passing to the limit, we obtain $(\lambda - \lambda')^T v^T \nabla^2 g(y) v \geq 0 \quad \forall \lambda' \in \Lambda(y, y^*)$, and hence $\lambda \in \Lambda(y, y^*; v)$.

Since

$$y_k^* = \nabla g(y)^T \widetilde{\lambda}^k + t_k v_k^* = \nabla g(y_k)^T \lambda^k,$$

we obtain

$$\begin{aligned} v^* &= \lim_{k\to\infty} v_k^* = \lim_{k\to\infty} \frac{\nabla g(y_k)^T\lambda^k - \nabla g(y)^T\widetilde{\lambda}^k}{t_k} \\ &= \lim_{k\to\infty} \frac{(\nabla g(y_k) - \nabla g(y))^T\lambda^k + \nabla g(y)^T(\lambda^k - \widetilde{\lambda}^k)}{t_k} \\ &= \nabla^2(\lambda^T g)(y)v + \nabla g(y)^T\mu. \end{aligned}$$

If $\mu \in T_{N_{\mathbb{R}^q_-}(g(y))}(\lambda)$, since (30) holds, by using Lemma 1 we have $\nabla g(y)^T\mu \in N_{K(y,y^*)}(v)$, and hence the inclusion (25) is proved. Otherwise, by taking into account

$$T_{N_{\mathbb{R}^q_-}(g(y))}(\lambda) = \left\{ \mu \in \mathbb{R}^q \mid \begin{array}{ll} \mu_i \geq 0 & \text{if } \lambda_i = 0 \text{ and } i \in \mathcal{I}(y) \\ \mu_i = 0 & \text{if } i \notin \mathcal{I}(y) \end{array} \right\},$$

and $\mu_i = 0$, $i \notin \widetilde{I}$, the set $J := \{i \in \widetilde{\mathcal{I}} \mid \lambda_i = 0, \mu_i < 0\}$ is not empty. Since μ^k converges to μ, we can choose some index $\bar{k}$ such that $\mu_i^{\bar{k}} = (\lambda_i^{\bar{k}} - \widetilde{\lambda}_i^{\bar{k}})/t_{\bar{k}} \leq \mu_i/2\ \forall i \in J$. Set $\widetilde{\mu} := \mu + 2(\widetilde{\lambda}^{\bar{k}} - \lambda)/t_{\bar{k}}$. Then for all i with $\lambda_i = 0$, we have $\widetilde{\mu}_i \geq \mu_i$ and for all $i \in J$, we have

$$\widetilde{\mu}_i = \mu_i + 2(\widetilde{\lambda}_i^{\bar{k}} - \lambda_i)/t_{\bar{k}} \geq \mu_i + 2(\widetilde{\lambda}_i^{\bar{k}} - \lambda_i^{\bar{k}})/t_{\bar{k}} \geq 0,$$

and therefore $\widetilde{\mu} \in T_{N_{\mathbb{R}^q_-}(g(y))}(\lambda)$. Observing that $\nabla g(y)^T\widetilde{\mu} = \nabla g(y)^T\mu$ because of $\lambda, \widetilde{\lambda}^{\bar{k}} \in \Lambda(y, y^*)$, and taking into account Lemma 1 we have $\nabla g(y)^T\widetilde{\mu} \in N_{K(y,y^*)}(v)$, and hence the inclusion (25) is proved. This finishes the proof of the theorem.

□

Since the regular normal cone is the polar of the tangent cone, the following characterization of the regular normal cone of $\operatorname{gph}\widehat{N}_\Gamma$ follows from the formula for the tangent cone in Theorem 4.

Corollary 2. *Assume that MSCQ is satisfied for the system $g(y) \leq 0$ at $\bar{y} \in \Gamma$. Then, there is a neighborhood V of $\bar{y}$ such that for every $(y, y^*) \in \operatorname{gph}\widehat{N}_\Gamma$ with $y \in V$ the following assertion holds: given any pair $(w^*, w) \in \widehat{N}_{\operatorname{gph}\widehat{N}_\Gamma}(y, y^*)$ we have $w \in K(y, y^*)$ and*

$$\langle w^*, w\rangle + w^T\nabla^2(\lambda^T g)(y)w \leq 0 \quad \textit{whenever} \quad \lambda \in \Lambda(y, y^*; w). \tag{32}$$

Proof. Choose V such that (23) holds true for every $y \in \Gamma \cap V$, and consider any $(y, y^*) \in \operatorname{gph}\widehat{N}_\Gamma$ with $y \in V$ and $(w^*, w) \in \widehat{N}_{\operatorname{gph}\widehat{N}_\Gamma}(y, y^*)$. By the definition of the regular normal cone, we have $\widehat{N}_{\operatorname{gph}\widehat{N}_\Gamma}(y, y^*) = \big(T_{\operatorname{gph}\widehat{N}_\Gamma}(y, y^*)\big)^\circ$ and, since $\{0\} \times N_{K(y,y^*)}(0) \subseteq T_{\operatorname{gph}\widehat{N}_\Gamma}(y, y^*)$, we obtain

$$\langle w^*, 0\rangle + \langle w, v^*\rangle \leq 0\ \forall v^* \in N_{K(y,y^*)}(0) = K(y, y^*)^\circ,$$

implying $w \in \operatorname{cl}\operatorname{conv} K(y, y^*) = K(y, y^*)$. By (23), we have $(w, \nabla^2(\lambda^T g)(y)w) \in T_{\operatorname{gph}\widehat{N}_\Gamma}(y, y^*)$ for every $\lambda \in \Lambda(y, y^*; w)$, and therefore the claimed inequality

$$\langle w^*, w\rangle + \langle w, \nabla^2(\lambda^T g)(y)w\rangle = \langle w^*, w\rangle + w^T\nabla^2(\lambda^T g)(y)w \leq 0$$

follows. □

The following result will be needed in the proof of Theorem 5.

Lemma 2. *Given $\bar{y} \in \Gamma$, assume that MSCQ holds at $\bar{y}$. Then there is a real $\kappa' > 0$ such that for any $y \in \Gamma$ sufficiently close to $\bar{y}$, any normal vector $y^* \in \widehat{N}_\Gamma(y)$ and any critical direction $v \in K(y, y^*)$ one has*

$$\Lambda(y, y^*; v) \cap \mathcal{E}(y, y^*) \cap \kappa' \|y^*\| \mathcal{B}_{\mathbb{R}^q} \neq \emptyset. \tag{33}$$

Proof. Let $\kappa > 0$ be chosen according to Theorem 4, and consider $y \in \Gamma$ as close to $\bar{y}$ such that MSCQ holds at y and (23) is valid for every $y^* \in \widehat{N}_\Gamma(y)$. Consider $y^* \in \widehat{N}_\Gamma(y)$ and a critical direction $v \in K(y, y^*)$. By [17, Proposition 4.3] we have $\Lambda(y, y^*; v) \neq \emptyset$ and, by taking any $\lambda \in \Lambda(y, y^*; v)$, we obtain from Theorem 4 that $(v, v^*) \in T_{\mathrm{gph}\, \widehat{N}_\Gamma}(y, y^*)$ with $v^* = \nabla^2(\lambda^T g)(y)v$. Applying Theorem 4 once more, we see that $v^* \in \nabla^2(\tilde{\lambda}^T g)(y)v + N_{K(y,y^*)}(v)$ with $\tilde{\lambda} \in \Lambda(y, y^*; v) \cap \kappa \|y^*\| \mathcal{B}_{\mathbb{R}^q}$, showing that $\Lambda(y, y^*; v) \cap \kappa \|y^*\| \mathcal{B}_{\mathbb{R}^q} \neq \emptyset$. Next consider a solution $\bar{\lambda}$ of the linear optimization problem

$$\min \sum_{i=1}^{q} \lambda_i \quad \text{subject to } \lambda \in \Lambda(y, y^*; v).$$

We can choose $\bar{\lambda}$ as an extreme point of the polyhedron $\Lambda(y, y^*; v)$, implying $\bar{\lambda} \in \mathcal{E}(y, y^*)$. Since $\Lambda(y, y^*; v) \subseteq \mathbb{R}^q_+$, we obtain

$$\|\bar{\lambda}\| \leq \sum_{i=1}^{q} |\bar{\lambda}_i| = \sum_{i=1}^{q} \bar{\lambda}_i \leq \sum_{i=1}^{q} \tilde{\lambda}_i \leq \sqrt{q} \|\tilde{\lambda}\| \leq \sqrt{q} \kappa \|y^*\|,$$

and hence (33) follows with $\kappa' = \kappa \sqrt{q}$. □

We are now in position to state a verifiable sufficient condition for MSCQ to hold for problem (MPEC).

Theorem 5. *Given $(\bar{x}, \bar{y}) \in \Omega$ defined as in (18), assume that MSCQ holds both for the lower level problem constraints $g(y) \leq 0$ at $\bar{y}$ and for the upper level constraints $G(x, y) \leq 0$ at $(\bar{x}, \bar{y})$. Further assume that*

$$\nabla_x G(\bar{x}, \bar{y})^T \eta = 0,\ \eta \in N_{\mathbb{R}^p_-}(G(\bar{x}, \bar{y})) \quad \Longrightarrow \quad \nabla_y G(\bar{x}, \bar{y})^T \eta = 0, \tag{34}$$

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

$$\nabla G(\bar{x}, \bar{y})(u, v) \in T_{\mathbb{R}^p_-}(G(\bar{x}, \bar{y})), \tag{35}$$

$$(v, -\nabla_x \phi(\bar{x}, \bar{y})u - \nabla_y \phi(\bar{x}, \bar{y})v) \in T_{\mathrm{gph}\, \widehat{N}_\Gamma}(\bar{y}, -\phi(\bar{x}, \bar{y})), \tag{36}$$

$$-\nabla_x \phi(\bar{x}, \bar{y})^T w + \nabla_x G(\bar{x}, \bar{y})^T \eta = 0,\ \eta \in N_{\mathbb{R}^p_-}(G(\bar{x}, \bar{y})),\ \eta^T \nabla G(\bar{x}, \bar{y})(u, v) = 0, \tag{37}$$

$$\nabla g_i(\bar{y})w = 0, i \in I^+(\lambda),\ w^T \left(\nabla_y \phi(\bar{x}, \bar{y}) + \nabla^2(\lambda^T g(\bar{y})\right) w - \eta^T \nabla_y G(\bar{x}, \bar{y})w \leq 0, \tag{38}$$

where the tangent cone $T_{\mathrm{gph}\, \widehat{N}_\Gamma}(\bar{y}, -\phi(\bar{x}, \bar{y}))$ can be calculated as in Theorem 4. Then the multifunction M_{MPEC} defined by

$$M_{\mathrm{MPEC}}(x, y) := \begin{pmatrix} \phi(x, y) + \widehat{N}_\Gamma(y) \\ G(x, y) - \mathbb{R}^p_- \end{pmatrix} \tag{39}$$

is metrically subregular at $\big((\bar{x}, \bar{y}), 0\big)$.

Proof. By Proposition 3, it suffices to show that the multifunction $P(x,y) - D$ with P and D given by

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

is metrically subregular at $\big((\bar{x},\bar{y}),0\big)$. We now invoke Theorem 2 with

$$P_1(x,y) := (y, -\phi(x,y)),\ P_2(x,y) := G(x,y),\ D_1 := \operatorname{gph} \widehat{N}_\Gamma, D_2 := \mathbb{R}^p_-.$$

By the assumption, $P_2(x,y) - D_2$ is metrically subregular at $\big((\bar{x},\bar{y}),0\big)$. Assume to the contrary that $P(\cdot,\cdot) - D$ is not metrically subregular at $\big((\bar{x},\bar{y}),0\big)$. Then by Theorem 2, there exist $0 \neq z = (u,v) \in T^{\text{lin}}_\Omega(\bar{x},\bar{y})$ and a directional limiting normal $z^* = (w^*, w, \eta) \in \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^p$ such that $\nabla P(\bar{x},\bar{y})^T z^* = 0$, $(w^*,w) \in N_{\operatorname{gph} \widehat{N}_\Gamma}(P_1(\bar{x},\bar{y}); \nabla P_1(\bar{x},\bar{y})z)$, $\eta \in N_{\mathbb{R}^p_-}\big(G(\bar{x},\bar{y}); \nabla G(\bar{x},\bar{y})(u,v)\big)$ and $(w^*,w) \neq 0$.

Hence

$$0 = \nabla P(\bar{x},\bar{y})^T z^* = \begin{pmatrix} -\nabla_x \phi(\bar{x},\bar{y})^T w + \nabla_x G(\bar{x},\bar{y})^T \eta \\ w^* - \nabla_y \phi(\bar{x},\bar{y})^T w + \nabla_y G(\bar{x},\bar{y})^T \eta \end{pmatrix}. \tag{40}$$

Since $z = (u,v) \in T^{\text{lin}}_\Omega(\bar{x},\bar{y})$, by the rule of tangents to product sets from Proposition 1, we obtain

$$\nabla P(\bar{x},\bar{y})z = \begin{pmatrix} (v, -\nabla_x \phi(\bar{x},\bar{y})u - \nabla_y \phi(\bar{x},\bar{y})v) \\ \nabla G(\bar{x},\bar{y})(u,v) \end{pmatrix} \in T_{\operatorname{gph} \widehat{N}_\Gamma}(\bar{y},\bar{y}^*) \times T_{\mathbb{R}^p_-}\big(G(\bar{x},\bar{y})\big),$$

where $\bar{y}^* := -\phi(\bar{x},\bar{y})$. It follows that $(v, -\nabla_x \phi(\bar{x},\bar{y})u - \nabla_y \phi(\bar{x},\bar{y})v) \in T_{\operatorname{gph} \widehat{N}_\Gamma}(\bar{y},\bar{y}^*)$, and consequently by Theorem 4 we have $v \in K(\bar{y},\bar{y}^*)$. Further we deduce from Proposition 2 that

$$\eta \in N_{\mathbb{R}^p_-}(G(\bar{x},\bar{y})),\ \eta^T \nabla G(\bar{x},\bar{y})(u,v) = 0.$$

So far we have shown that u, v, η, w fulfill (35)-(37). Further we have $w \neq 0$, because if $w = 0$, then by virtue of (34) and (40) we would obtain $\nabla_x G(\bar{x},\bar{y})^T \eta = 0$, $\nabla_y G(\bar{x},\bar{y})^T \eta = 0$ and consequently $w^* = 0$, contradicting $(w^*,w) \neq 0$. If we can show the existence of $\lambda \in \Lambda(\bar{y},\bar{y}^*;v) \cap \mathcal{E}(\bar{y},\bar{y}^*)$ such that (38) holds, then we have obtained the desired contradiction to our assumptions, and this would complete the proof.

Since $(w^*,w) \in N_{\operatorname{gph} \widehat{N}_\Gamma}(P_1(\bar{x},\bar{y}); \nabla P_1(\bar{x},\bar{y})z)$, by the definition of the directional limiting normal cone, there are sequences $t_k \downarrow 0$, $d_k = (v_k, v_k^*) \in \mathbb{R}^m \times \mathbb{R}^m$ and $(w_k^*, w_k) \in \mathbb{R}^m \times \mathbb{R}^m$ satisfying $(w_k^*, w_k) \in \widehat{N}_{\operatorname{gph} \widehat{N}_\Gamma}(P_1(\bar{x},\bar{y}) + t_k d_k)\ \forall k$ and $(d_k, w_k^*, w_k) \to (\nabla P_1(\bar{x},\bar{y})z, w^*, w)$. That is, $(y_k, y_k^*) := (\bar{y},\bar{y}^*) + t_k(v_k, v_k^*) \in \operatorname{gph} \widehat{N}_\Gamma$, $(w_k^*, w_k) \in \widehat{N}_{\operatorname{gph} \widehat{N}_\Gamma}(y_k, y_k^*)$ and $(v_k, v_k^*) \to (v, -\nabla_x \phi(\bar{x},\bar{y})u - \nabla_y \phi(\bar{x},\bar{y})v)$. By passing to a subsequence if necessary, we can assume that MSCQ holds for $g(y) \leq 0$ at y_k for all k, and by invoking Corollary 2 we obtain $w_k \in K(y_k, y_k^*)$, and

$$w_k^{*T} w_k + w_k^T \nabla^2(\lambda^T g)(y_k) w_k \leq 0 \text{ whenever } \lambda \in \Lambda(y_k, y_k^*; w_k). \tag{41}$$

By Lemma 2, we can find a uniformly bounded sequence $\lambda^k \in \Lambda(y_k, y_k^*; w_k) \cap \mathcal{E}(y_k, y_k^*)$. In particular, following from the proof of Lemma 2, we can choose λ^k as an optimal solution of the linear optimization problem

$$\min \sum_{i=1}^q \lambda_i \ \text{ subject to } \ \lambda \in \Lambda(y_k, y_k^*; w_k). \tag{42}$$

By passing once more to a subsequence if necessary, we can assume that λ^k converges to $\bar{\lambda}$, and we easily conclude $\bar{\lambda} \in \Lambda(\bar{y}, \bar{y}^*)$ and ${w^*}^T w + w^T \nabla^2(\bar{\lambda}^T g)(\bar{y}) w \leq 0$, which together with $w^* - \nabla_y \phi(\bar{x}, \bar{y})^T w + \nabla_y G(\bar{x}, \bar{y})^T \eta = 0$ (see (40)) results in

$$w^T \left(\nabla_y \phi(\bar{x}, \bar{y}) + \nabla^2(\bar{\lambda}^T g)(\bar{y}) \right) w - \eta^T \nabla_y G(\bar{x}, \bar{y}) w \leq 0. \tag{43}$$

Further, we can assume that $I^+(\bar{\lambda}) \subseteq I^+(\lambda^k)$ and therefore, because of $\lambda^k \in N_{\mathbb{R}^q_-}(g(y_k))$, $\bar{\lambda}^T g(y_k) = {\lambda^k}^T g(y_k) = 0$. Hence for every $\lambda \in \Lambda(\bar{y}, \bar{y}^*)$ we obtain

$$\begin{aligned} 0 &\geq (\lambda - \bar{\lambda})^T g(y_k) \\ &= (\lambda - \bar{\lambda})^T g(\bar{y}) + \nabla((\lambda - \bar{\lambda})^T g)(\bar{y})(y_k - \bar{y}) \\ &\quad + \frac{1}{2}(y_k - \bar{y})^T \nabla^2((\lambda - \bar{\lambda})^T g)(\bar{y})(y_k - \bar{y}) + o(\|y_k - \bar{y}\|^2) \\ &= \frac{t_k^2}{2} v_k^T \nabla^2((\lambda - \bar{\lambda})^T g)(\bar{y}) v_k + o(t_k^2 \|v_k\|^2). \end{aligned}$$

Dividing by $t_k^2/2$ and passing to the limit yields $0 \geq v^T \nabla^2((\lambda - \bar{\lambda})^T g)(\bar{y}) v$, and thus $\bar{\lambda} \in \Lambda(\bar{y}, \bar{y}^*; v)$. Since $w_k \in K(y_k, y_k^*)$, by Proposition 6, we have $\nabla g_i(y_k) w_k = 0$, $i \in I^+(\lambda^k)$, from which $\nabla g_i(\bar{y}) w = 0$, $i \in I^+(\bar{\lambda})$ follows.

It is known that the polyhedron $\Lambda(\bar{y}, \bar{y}^*)$ can be represented as the sum of the convex hull of its extreme points $\mathcal{E}(\bar{y}, \bar{y}^*)$ and its recession cone $\mathcal{R} := \{\lambda \in N_{\mathbb{R}^q_-}(g(\bar{y})) | \nabla g(\bar{y})^T \lambda = 0\}$. We show by contradiction that $\bar{\lambda} \in \operatorname{conv} \mathcal{E}(\bar{y}, \bar{y}^*)$. Assuming on the contrary that $\bar{\lambda} \notin \operatorname{conv} \mathcal{E}(\bar{y}, \bar{y}^*)$, then $\bar{\lambda}$ has the representation $\bar{\lambda} = \lambda^c + \lambda^r$ with $\lambda^c \in \operatorname{conv} \mathcal{E}(\bar{y}, \bar{y}^*)$ and $\lambda^r \neq 0$ belongs to the recession cone $\mathcal{R}$, i.e.

$$\lambda^r \in N_{\mathbb{R}^q_-}(g(\bar{y})), \ \nabla g(\bar{y})^T \lambda^r = 0. \tag{44}$$

Since $\lambda^k \in \Lambda(y_k, y_k^*; w_k)$, it is a solution to the linear program:

$$\begin{aligned} \max_{\lambda \geq 0} \quad & w_k^T \nabla^2(\lambda^T g)(y_k)(w_k) \\ s.t. \quad & \nabla g(y_k)^T \lambda = y_k^*, \\ & \lambda^T g(y_k) = 0. \end{aligned}$$

By the duality theory of linear programming, for each k there is some $r_k \in \mathbb{R}^m$ verifying

$$\nabla g_i(y_k) r_k + w_k^T \nabla^2 g_i(y_k) w_k \leq 0, \ \lambda_i^k(\nabla g_i(y_k) r_k + w_k^T \nabla^2 g_i(y_k) w_k) = 0, \ i \in \mathcal{I}(y_k).$$

Since $\Lambda(y_k, y_k^*; w_k) = \{\lambda \in \Lambda(y_k, y_k^*) \,|\, w_k^T \nabla^2(\lambda^T g)(y_k) w_k \geq \theta(y_k, y_k^*; w_k)\}$ and λ^k solves (42), again by the duality theory of linear programming we can find for each k some $s_k \in \mathbb{R}^m$ and $\beta_k \in \mathbb{R}_+$ such that

$$\nabla g_i(y_k) s_k + \beta_k w_k^T \nabla^2 g_i(y_k) w_k \leq 1, \ \lambda_i^k(\nabla g_i(y_k) s_k + \beta_k w_k^T \nabla^2 g_i(y_k) w_k - 1) = 0, \ i \in \mathcal{I}(y_k).$$

Next we define for every k the elements $\tilde{\lambda}^k \in \mathbb{R}^q_+$, $\xi_k^* \in \mathbb{R}^m$ by

$$\begin{aligned} \tilde{\lambda}_i^k &:= \begin{cases} \lambda_i^r & \text{if } i \in I^+(\lambda^r), \\ \frac{1}{k} & \text{if } i \in I^+(\lambda^k) \setminus I^+(\lambda^r), \\ 0 & \text{else}, \end{cases} \\ \xi_k^* &:= \nabla g(y_k)^T \tilde{\lambda}^k. \end{aligned} \tag{45}$$

Since $I^+(\lambda^r) \subseteq I^+(\bar\lambda) \subseteq I^+(\lambda^k)$, we obtain $I^+(\tilde\lambda^k) = I^+(\lambda^k)$, $\tilde\lambda^k \in N_{\mathbb{R}^q_-}(g(y_k))$ and $\xi_k^* \in N_\Gamma(y_k)$. Thus $w_k \in K(y_k, \xi_k^*)$ by Proposition 6 and

$$\nabla g_i(y_k) r_k + w_k^T \nabla^2 g_i(y_k) w_k \le 0,\ \tilde\lambda_i^k(\nabla g_i(y_k) r_k + w_k^T \nabla^2 g_i(y_k) w_k) = 0,\ i \in \mathcal{I}(y_k),$$

implying $\tilde\lambda^k \in \Lambda(y_k, \xi_k^*; w_k)$ by duality theory of linear programming. Moreover, because of $I^+(\tilde\lambda^k) = I^+(\lambda^k)$ we also have

$$\nabla g_i(y_k) s_k + \beta_k w_k^T \nabla^2 g_i(y_k) w_k \le 1,\ \tilde\lambda_i^k(\nabla g_i(y_k) s_k + \beta_k w_k^T \nabla^2 g_i(y_k) w_k - 1) = 0,\ i \in \mathcal{I}(y_k),$$

implying that $\tilde\lambda^k$ is solution of the linear program

$$\min \sum_{i=1}^q \lambda_i \ \text{ subject to } \ \lambda \in \Lambda(y_k, \xi_k^*; w_k),$$

and, together with $\Lambda(y_k, \xi_k^*; w_k) \subseteq \mathbb{R}^q_+$,

$$\min\{\|\lambda\| \mid \lambda \in \Lambda(y_k, \xi_k^*; w_k)\} \ge \frac{1}{\sqrt{q}} \min\{\sum_{i=1}^q \lambda_i \mid \lambda \in \Lambda(y_k, \xi_k^*; w_k)\} \ge \frac{\sum_{i=1}^q \lambda_i^r}{\sqrt{q}} := \beta > 0.$$

Taking into account that $\lim_{k\to\infty} \tilde\lambda^k = \lambda^r$ and (44), (45), we conclude $\lim_{k\to\infty} \|\xi_k^*\| = 0$, showing that for every real κ' we have

$$\Lambda(y_k, \xi_k^*; w_k) \cap \mathcal{E}(y_k, \xi_k^*) \cap \kappa' \|\xi_k^*\| \mathcal{B}_{\mathbb{R}^q} \subseteq \Lambda(y_k, \xi_k^*; w_k) \cap \kappa' \|\xi_k^*\| \mathcal{B}_{\mathbb{R}^q} = \emptyset$$

for all k sufficiently large contradicting the statement of Lemma 2. Hence $\bar\lambda \in \operatorname{conv} \mathcal{E}(\bar y, \bar y^*)$ and thus $\bar\lambda$ admits a representation as convex combination

$$\bar\lambda = \sum_{j=1}^N \alpha_j \hat\lambda^j \ \text{ with } \ \sum_{j=1}^N \alpha_j = 1,\ 0 < \alpha_j \le 1,\ \hat\lambda^j \in \mathcal{E}(\bar y, \bar y^*),\ j = 1, \ldots, N.$$

Since $\bar\lambda \in \Lambda(\bar y, \bar y^*; v)$, we have $\theta(\bar y, \bar y^*; v) = v^T \nabla^2(\bar\lambda^T g)(\bar y) v = \sum_{j=1}^N \alpha_j v^T \nabla^2({\hat\lambda^j}^T g)(\bar y) v$ implying, together with $v^T \nabla^2({\hat\lambda^j}^T g)(\bar y) v \le \theta(\bar y, \bar y^*; v)$, that $v^T \nabla^2({\hat\lambda^j}^T g)(\bar y) v = \theta(\bar y, \bar y^*; v)$. Consequently $\hat\lambda^j \in \Lambda(\bar y, \bar y^*; v)$. It follows from (43) that

$$\begin{aligned}
&\sum_{j=1}^N \alpha_j \left(w^T \big(\nabla_y \phi(\bar x, \bar y) + \nabla^2({\hat\lambda^j}^T g)(\bar y)\big) w - \eta^T \nabla_y G(\bar x, \bar y) w \right) \\
&\qquad = w^T \left(\nabla_y \phi(\bar x, \bar y) + \nabla^2(\bar\lambda^T g)(\bar y)\right) w - \eta^T \nabla_y G(\bar x, \bar y) w \le 0,
\end{aligned}$$

and hence there exists some index $\bar j$ with

$$w^T \left(\nabla_y \phi(\bar x, \bar y) + \nabla^2({\hat\lambda^{\bar j}}^T g)(\bar y)\right) w - \eta^T \nabla_y G(\bar x, \bar y) w \le 0.$$

Further, by Proposition 6, we have $\nabla g_i(\bar y) w = 0\ \forall i \in I^+(\bar\lambda) \supseteq I^+(\hat\lambda^{\bar j})$, and we see that (38) is fulfilled with $\lambda = \hat\lambda^{\bar j}$. $\square$

Example 2 (Example 1 revisited). *Instead of reformulating the MPEC as a (MPCC), we consider the MPEC in the original form (MPEC). Since for the constraints $g(y) \leq 0$ of the lower level problem, MFCQ is fulfilled at $\bar{y}$ and the gradients of the upper level constraints $G(x, y) \leq 0$ are linearly independent, MSCQ holds for both constraint systems. Condition* (34) *is obviously fulfilled due to $\nabla_y G(x, y) = 0$. Setting $\bar{y}^* := -\phi(\bar{x}, \bar{y}) = (0, 0, 1)$, as in Example 1 we obtain*

$$\Lambda(\bar{y}, \bar{y}^*) = \{(\lambda_1, \lambda_2) \in \mathbb{R}^2_+ \mid \lambda_1 + \lambda_2 = 1\}.$$

Since $\nabla g_1(\bar{y}) = \nabla g_2(\bar{y}) = (0, 0, 1)$ and for every $\lambda \in \Lambda(\bar{y}, \bar{y}^*)$ either $\lambda_1 > 0$ or $\lambda_2 > 0$, we deduce

$$W(\lambda) := \{w \in \mathbb{R}^3 \mid \nabla g_i(\bar{y})w = 0,\ i \in I^+(\lambda)\} = \mathbb{R}^2 \times \{0\} \quad \forall \lambda \in \Lambda(\bar{y}, \bar{y}^*).$$

Since

$$w^T \left(\nabla_y \phi(\bar{x}, \bar{y}) + \nabla^2(\lambda^T g)(\bar{y})\right) w - \eta^T \nabla_y G(\bar{x}, \bar{y}) w = (1 + \lambda_1) w_1^2 + (1 + \lambda_2) w_2^2 \geq 0,$$

there cannot exist $0 \neq w \in W(\lambda)$ and $\lambda \in \Lambda(\bar{y}, \bar{y}^*)$ fulfilling (38)*. Hence by virtue of Theorem 5, MSCQ holds at $(\bar{x}, \bar{y})$.*

Acknowledgements

The research of the first author was supported by the Austrian Science Fund (FWF) under grants P26132-N25 and P29190-N32. The research of the second author was partially supported by NSERC. The authors would like to thank the two anonymous reviewers for their extremely careful review and valuable comments that have helped us to improve the presentation of the manuscript.

References

[1] L. Adam, R. Henrion, J. Outrata, *On M-stationarity conditions in MPECs and the associated qualification conditions*, preprint (2016).

[2] G.B. Allende, G. Still, *Solving bilevel programs with the KKT-approach*, Math. Program., 138 (2013), pp. 309-332.

[3] R. Andreani, G. Haeser, M.L. Schuverdt, P.J.S. Silva, *A relaxed constant positive linear dependence constraint qualification and applications*, Math. Program. 135(2012), pp. 255–273.

[4] R. Andreani, G. Haeser, M.L. Schuverdt, P.J.S. Silva, *Two new weak constraint qualification and applications*, SIAM J. Optim., 22(2012), pp. 1109–1135.

[5] J.-P. Aubin, *Lipschitz behavior of solutions to convex minimization problems*, Math. Oper. Res., 9 (1994), pp. 87–111.

[6] N.H. Chieu, L.V. Hien, *Computation of graphical derivative for a class of normal cone mappings under a very weak condition*, preprint (2016), to appear in SIOPT.

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

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

[9] A.L. Dontchev, R.T. Rockafellar, *Regularity and conditioning of solution mappings in variational analysis*, Set-Valued Var. Anal., 12 (2004), pp. 79-109.

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

[11] M.L. Flegel, C. Kanzow, *Abadie-type constraint qualification for mathematical programs with equilibrium constraints*, J. Optim. Theory Appl., 124 (2005), pp. 595-614.

[12] 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.

[13] 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.

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

[15] 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.

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

[17] 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.

[18] H. Gfrerer, B.S. Mordukhovich, *Robinson stability of parametric constraint systems via variational analysis*, to appear in SIOPT, preprint available at arXiv:1609.02238.

[19] 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.

[20] H. Gfrerer and J.V. Outrata, *On computation of limiting coderivatives of the normal-cone mapping to inequality systems and their applications*, Optimization, DOI 10.1080/02331934.2015.1066372.

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

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

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

[24] L. Guo, J.J. Ye, J. Zhang, *Mathematical programs with geometric constraints in Banach spaces: enhanced optimality, exact penalty, and sensitivity*, SIAM J. Optim., 23 (2013), pp. 2295–2319.

[25] L. Guo, J. Zhang, G-H. Lin, *New results on constraint qualifications for nonlinear extremum problems and extensions*, J. Optim. Theory Appl., 163 (2014), pp. 737-754.

[26] A.D. Ioffe, *Regular points of Lipschitz functions*, Trans. Amer. Math. Soc., 251 (1979), pp. 61-69.

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

[28] P. Mehlitz and G. Wachsmuth, *Weak and strong stationarity in generalized bilevel programming and bilevel optimal control*, Optimization, 65 (2016), pp. 907-935.

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

[30] 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.

[31] S.M. Robinson, *Stability theory for systems of inequality constraints, part I: linear systems*, SIAM J. Numer. Anal., 12 (1975), pp. 754-769.

[32] S.M. Robinson, *Stability theory for systems of inequality constraints, part II: differentiable nonlinear systems*, SIAM J. Numer. Anal., 13(1976), pp. 497-513.

[33] S.M. Robinson, *Some continuity properties of polyhedral multifunctions*, SIAM J. Control Optim., 28 (1981), pp. 206-214.

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

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

[36] J.J. Ye, *Constraint qualification and necessary optimality conditions for optimization problems with variational inequality constraints*, SIAM J. Optim., 10 (2000), pp. 943-962.

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