JOHANNES KEPLER UNIVERSITY LINZ

Institute of Computational Mathematics

Optimality Conditions for Disjunctive Programs Based on Generalized Differentiation with Application to Mathematical Programs with Equilibrium Constraints

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

NuMa-Report No. 2013-02 March 2013

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/`

Optimality conditions for disjunctive programs based on generalized differentiation with application to mathematical programs with equilibrium constraints

Helmut Gfrerer*

Abstract

We consider optimization problems with a disjunctive structure of the constraints. Prominent examples of such problems are mathematical programs with equilibrium constraints or vanishing constraints. Based on the concepts of directional subregularity and their characterization by means of objects from generalized differentiation, we obtain the new stationarity concept of extended M-stationarity, which turns out to be an equivalent dual characterization of B-stationarity. These results are valid under a very weak constraint qualification of Guignard type which is usually very difficult to verify. We also state a new constraint qualification which is a little bit stronger but verifiable. Further we present second-order optimality conditions, both necessary and sufficient. Finally we apply these results to the special case of mathematical programs with equilibrium constraints and compute explicitly all the objects from generalized differentiation. For this type of problems we also introduce the concept of strongly M-stationarity which builds a bridge between S-stationarity and M-stationarity.

Key words. Optimality conditions, M-stationarity, metric subregularity
AMS subject classification. 49J53 49K27 90C48

1 Introduction

In this paper we consider mathematical programs with disjunctive constraints of the form

$$\min_{x\in\mathbb{R}^n} f(x) \quad \text{subject to} \quad F(x)\in\Omega, \tag{1}$$

where the functions $f:\mathbb{R}^n\to\mathbb{R}$, $F:\mathbb{R}^n\to\mathbb{R}^m$ are continuously differentiable and $\Omega=\cup_{i=1}^{\bar p}P_i$ is the union of finitely many convex polyhedra P_i.

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

A prominent example for such programs are mathematical programs with equilibrium constraints (MPEC for short)

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & f(x) \\ \text{subject to} \quad & g(x) \le 0, \\ & h(x) = 0, \\ & G_i(x) \ge 0, H_i(x) \ge 0, G_i(x)H_i(x) = 0,\ i = 1,\dots,q \end{aligned} \tag{2}$$

with functions $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to R^l$, $h : \mathbb{R}^n \to \mathbb{R}^p$, and $G,H : R^n \to \mathbb{R}^q$. Note that the complementarity conditions $G_i(x) \ge 0, H_i(x) \ge 0, G_i(x)H_i(x) = 0$ can be equivalently rewritten in the form

$$-(G_i(x), H_i(x)) \in Q_{\mathrm{EC}},\ i = 1,\dots,q, \tag{3}$$

where

$$Q_{\mathrm{EC}} := \{(a,b) \in \mathbb{R}^2_- \,|\, ab = 0\} \tag{4}$$

is the union of the 2 polyhedral sets $\mathbb{R}_- \times \{0\}$ and $\{0\} \times \mathbb{R}_-$. Hence the MPEC is of the form (1) with

$$F(x) = \begin{pmatrix} g(x) \\ h(x) \\ -G_1(x) \\ -H_1(x) \\ \vdots \\ -G_q(x) \\ -H_q(x) \end{pmatrix},\ \Omega = \mathbb{R}^l_- \times \{0\}^p \times Q^q_{\mathrm{EC}} \tag{5}$$

is the union of 2^q polyhedral sets. Here, the minus signs are used only for convenience of the subsequent analysis.

MPECs have their origin in bilevel programming and arise in many applications in economic, engineering and natural sciences. We refer to the monographs [32, 37] fur further details.

MPECs are known to be difficult optimization problems because due to the complementarity conditions $G_i(x) \ge 0$, $H_i(x) \ge 0$, $G_i(x)H_i(x) = 0$ many of the standard constraint qualifications of nonlinear programming are violated at any feasible point. Hence the usual Karush-Kuhn-Tucker conditions fail to hold at a local minimzer and various first-order optimality conditions such as Abadie (A-), Bouligand (B-), Clarke (C-), Mordukhovich (M-) and Strong (S-) stationarity conditions have been studied in the literature [6, 8, 29, 36, 35, 43, 44, 46, 45]. B-stationarity expresses the first-order necessary condition that there does not exist a feasible descent direction at a local optimum. However this is very difficult to verify and hence the other stationarity concepts have been introduced. S-stationarity is sufficient for B-stationarity but it only holds only under some strong constraint qualification such as MPEC-LICQ (Linear Independence Constraint Qualification). A slightly weaker stationary concept is M-stationarity which holds under fairly mild assumptions. However, it is to be noted that M-stationarity (and therefore also the weaker concepts of A- and C-stationarity) does not preclude the occurrence of feasible descent directions.

Compared with the first-order optimality conditions, very little has been done done with the second-order optimality conditions for MPECs. In [43] necessary and sufficient conditions based on the concept of S-stationarity have been stated. In [15] second-order necessary conditions in terms of S- and C-multipliers were stated and some consequences of a strong second-order sufficient conditions based on M-multipliers were given.

Another example for programs with disjunctive constraints arise from programs with vanishing constraints (MPVC)

$$H_i(x) \geq 0,\ G_i(x)H_i(x) \leq 0, i = 1,\dots,q \tag{6}$$

which can be equivalently formulated as

$$(G_i(x), H_i(x)) \in Q_{\mathrm{VC}}$$

with Q_{VC} being the union of the 2 polyhedral sets $\mathbb{R}_- \times \mathbb{R}_+$ and $\mathbb{R} \times \{0\}$. For more details on MPVCs and optimality conditions we refer the reader to [1, 2, 22, 23, 24, 28].

For S- and M-stationarity conditions for mathematical programs with disjunctive constraints we refer to [7].

The aim of this paper is to present a unified theory of optimality conditions based on the concepts of generalized differentiation by Mordukhovich [33, 34]. In fact, by the Mordukhovich criterion [33, Theorem 4.18], the M-stationarity conditions state that a certain multifunction built by the objective and the constraints is not metrically regular. Our optimality conditions rely on the observation that at a local minimizer for every critical direction such a multifunction cannot have a certain regular behaviour. They are obtained by applying the characterizations of directional metric regularity, subregularity, and mixed regularity/subregularity as can be found in the recent papers [10, 11, 12].

In section 2 we recall the basic definitions of the different versions of regularity and their characterizations by means of generalized differentiation. In section 3 we state various optimality conditions for the problem (1). We obtain first-order optimality conditions called *extended M-stationarity conditions* that state that for every critical direction a certain M-stationarity condition (with possibly different multiplier) has to be fulfilled. We will show that this condition is an equivalent dual characterization of B-stationarity. Further we introduce a new constraint qualification based on directional metric subregularity which appears to be rather weak. Second-order optimality conditions, both necessary and sufficient are presented.

In section 4 we apply these results to MPECs by explicitly calculating the objects from generalized differentiation. Since extended M-stationarity is still difficult to verify, we present also the weaker necessary condition of *strong M-stationarity* which builds a bridge between S- and M-stationarity and seems to be well suited for numerical purposes.

In what follows we denote by $\mathscr{B}_{\mathbb{R}^n} := \{x \in \mathbb{R}^n \mid \|x\| \leq 1\}$ the closed unit ball. For a mapping $F : \mathbb{R}^n \to \mathbb{R}^m$ we denote by $\nabla F(\bar{x})$ the Jacobian.

2 Preliminaries

We start by recalling several definitions and results from variational analysis: Let $\Omega \subset \mathbb{R}^n$ be an arbitrary closed set and $x \in \Omega$. The *contingent* (also called *Bouligand* or *tangent*) *cone* to Ω at x, denoted by $T(x;\Omega)$, is given by

$$T(x;\Omega) := \{u \in \mathbb{R}^n \,|\, \exists (u_k) \to u, (t_k) \downarrow 0 : x + t_k u_k \in \Omega\}.$$

We denote by

$$\hat{N}(x;\Omega) = \{\xi \in \mathbb{R}^n \,|\, \limsup_{x' \xrightarrow{\Omega} x} \frac{\xi^T(x'-x)}{\|x'-x\|} \le 0\} \tag{7}$$

the *Fréchet* (*regular*) *normal cone* to Ω. Finally, the *Mordukhovich* (*basic/limiting*) *normal cone* to Ω at x is defined by

$$N(x;\Omega) := \{\xi \,|\, \exists (x_k) \xrightarrow{\Omega} x,\ (\xi_k) \to \xi : \xi_k \in \hat{N}(x_k;\Omega) \forall k\}.$$

If $x \notin \Omega$ we put $T(x;\Omega) = \emptyset$, $\hat{N}(x;\Omega) = \emptyset$ and $N(x;\Omega) = \emptyset$.

The Mordukhovich normal cone is generally nonconvex whereas the Fréchet normal cone is always convex. In the case of a convex set Ω, both the Fréchet normal cone and the Mordukhovich normal cone coincide with the standard normal cone from convex analysis and moreover, the contingent cone is equal to the tangent cone in the sense of convex analysis.

Note that $\xi \in \hat{N}(x;\Omega) \Leftrightarrow \xi^T u \le 0\ \forall u \in T(x;\Omega)$, i.e. $\hat{N}(x;\Omega)$ is the polar cone of $T(x;\Omega)$.

Given a multifunction $M : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ and a point $(\bar{x},\bar{y}) \in \operatorname{gph} M := \{(x,y) \in X \times Y \,|\, y \in M(x)\}$ from its graph, the *coderivative* of M at $(\bar{x},\bar{y})$ is a multifunction $D^*M(\bar{x},\bar{y}) : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ with the values $D^*M(\bar{x},\bar{y})(\eta) := \{\xi \in \mathbb{R}^n \,|\, (\xi,-\eta) \in N((\bar{x},\bar{y});\operatorname{gph} M)\}$, i.e. $D^*M(\bar{x},\bar{y})(\eta)$ is the collection of all $\xi \in \mathbb{R}^n$ for which there are sequences $(x_k,y_k) \to (\bar{x},\bar{y})$ and $(\xi_k,\eta_k) \to (u,v)$ with $(\xi_k,-\eta_k) \in \hat{N}((x_k,y_k);\operatorname{gph} M)$.

For more details we refer to the monographs [33, 42]

The following directional versions of these limiting constructions were introduced in [11]. Given a direction $u \in \mathbb{R}^n$, the Mordukhovich normal cone to a subset $\Omega \subset \mathbb{R}^n$ in direction u at $x \in \Omega$ is defined by

$$N(x;\Omega;u) := \{\xi \in \mathbb{R}^n \,|\, \exists (t_k) \downarrow 0,\ (u_k) \to u,\ (\xi_k) \to \xi : \xi_k \in \hat{N}(x + t_k u_k;\Omega) \forall k\}.$$

For a multifunction $M : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ and a direction $(u,v) \in \mathbb{R}^n \times \mathbb{R}^m$, the coderivative of M in direction (u,v) at $(\bar{x},\bar{y}) \in \operatorname{gph} M$ is defined as the multifunction $D^*M((\bar{x},\bar{y});(u,v)) : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ given by $D^*M((\bar{x},\bar{y});(u,v))(\eta) := \{\xi \in \mathbb{R}^n \,|\, (\xi,-\eta) \in N((\bar{x},\bar{y});\operatorname{gph} M;(u,v))\}$.

Note that the directional versions of the Mordukhovich normal cone and the coderivative as defined in [11] were introduced for general Banach spaces and therefore look somewhat different. In particular, in [11] it was distinguished between normal, mixed and reversed mixed coderivatives. However, in finite dimensional spaces weak-* and strong convergence coincide and hence this distinction is superfluous in our setting. In fact the definitions above are equivalent with the definitions from [11].

Note that by the definition we have $N(x;\Omega;0) = N(x;\Omega)$ and $D^*_N M((\bar{x},\bar{y});(0,0)) = D^*_N M(\bar{x},\bar{y})$. Further $N(x;\Omega;u) \subset N(x;\Omega)$ for all u and $N(x;\Omega;u) = \emptyset$ if $u \notin T(x;\Omega)$.

The following two lemmas give characterizations of the directional Mordukhovich normal cone.

Lemma 1. *Let $\Omega \subset \mathbb{R}^n$ be the union of finitely many closed convex sets Ω_i, $i = 1, \ldots, \bar{p}$, $\bar{x} \in \Omega$, $u \in \mathbb{R}^n$. Then*

$$N(\bar{x};\Omega;u) \subset \{\xi \in N(\bar{x};\Omega) \,|\, \xi^T u = 0\}. \tag{8}$$

If Ω is convex and $u \in T(\bar{x};\Omega)$ then the inclusion holds with equality and therefore $N(\bar{x};\Omega;u) = N(u;T(\bar{x};\Omega))$.

Proof. Let $\xi \in N(\bar{x};\Omega;u)$ be arbitrarily fixed and consider sequences $(t_k) \downarrow 0$, $(u_k) \to u$, $(\xi_k) \to \xi$ with $\xi_k \in \hat{N}(\bar{x} + t_k u_k;\Omega)$ for all k. Let $I_k := \{i \in \{1, \ldots, \bar{p}\} \,|\, \bar{x} + t_k u_k \in \Omega_i\}$. Since there are only finitely many subsets of $\{1, \ldots, \bar{p}\}$, by passing to a subsequence we can assume that there is some index set I such that $I_k = I$ for all k. Let $\bar{i} \in I$ arbitrarily fixed. Since $\Omega_{\bar{i}}$ is closed we have $\bar{x} \in \Omega_{\bar{i}}$ and therefore $\xi_k^T(\bar{x} - (\bar{x} + t_k u_k)) = -t_k \xi_k^T u_k \le 0$, implying $\xi^T u \ge 0$. Since $(t_k) \downarrow 0$, we can find for each k some index $j(k) \ge k$ with $t_{j(k)}/t_k \le \frac{1}{k}$. Then for all k we have $\xi_{j(k)}^T(\bar{x} + t_k u_k - (\bar{x} + t_{j(k)} u_{j(k)})) \le 0$ and therefore

$$0 \ge \lim_{k\to\infty} \xi_{j(k)}^T (u_k - \frac{t_{j(k)}}{t_k} u_{j(k)}) = \xi^T u$$

also holds. To show the assertion about equality in the case of convex Ω, let $u \in T(\bar{x};\Omega)$ and $\xi \in N(\bar{x};\Omega)$ with $\xi^T u = 0$ be arbitrarily fixed. Then we can find sequences $(t_k) \downarrow 0$ and $(u_k) \to u$ with $x_k := \bar{x} + t_k u_k \in \Omega$ for all k and, by passing to a subsequence if necessary, we can assume $k\|u - u_k\| \to 0$. Because of

$$-\xi^T(x - x_k) = -\xi^T(x - \bar{x}) + t_k \xi^T u_k = -\xi^T(x - \bar{x}) + t_k \xi^T (u_k - u) \ge t_k \xi^T (u_k - u) \;\forall x \in \Omega,$$

by invoking Ekeland's variational principle, there is for every k some $\tilde{x}_k \in \Omega$ such that $\|\tilde{x}_k - x_k\| \le k t_k \xi^T (u - u_k)$ and $\tilde{x}_k$ is a global minimizer of the problem

$$\min_{x\in\Omega} -\xi^T(x - x_k) + \frac{1}{k}\|x - \tilde{x}_k\|.$$

By the well known first order optimality conditions from convex analysis (see, e.g., [41, Theorem 27.4]) there is some element $\eta_k \in \mathscr{B}_{\mathbb{R}^n}$ such that $\xi - \frac{1}{k}\eta_k =: \xi_k \in N(\tilde{x}_k;\Omega)$, showing $\lim_{k\to\infty} \xi_k = \xi$. Since we also have

$$\limsup_{k\to\infty} \|\frac{\tilde{x}_k - \bar{x}}{t_k} - u\| \le \limsup_{k\to\infty} (\|\frac{\tilde{x}_k - x_k}{t_k}\| + \|\frac{x_k - \bar{x}}{t_k} - u\|) \le \limsup_{k\to\infty} (k\|\xi\|\|u_k - u\| + \|u_k - u\|) = 0,$$

$\xi \in N(x;\Omega;u)$ follows. □

Lemma 2. *Let $\Omega \subset \mathbb{R}^n$ be the union of finitely many polyhedra P_i, $i = 1, \ldots, \bar{p}$ and let $\bar{x} \in \Omega$ and $u \in T(\bar{x};\Omega)$. Then*

$$N(\bar{x};\Omega;u) = \bigcup_{v \in T(u;T(\bar{x};\Omega))} \hat{N}(v;T(u;T(\bar{x};\Omega))). \tag{9}$$

Proof. Let the polyhedra P_i $i=1,\ldots,\bar{p}$ be represented by

$$P_i=\{x\in\mathbb{R}^n\,|\,a_{ij}^Tx\le b_i,\ i=1,\ldots,m_i\}$$

and denote $\bar{\mathscr{P}}:=\{i\in\{1,\ldots,\bar{p}\}\,|\,\bar{x}\in P_i\}$ and $\bar{\mathscr{A}_i}:=\{j\in\{1,\ldots,m_i\}\,|\,a_{ij}^T\bar{x}=b_i\}$ for $i\in\bar{\mathscr{P}}$. Then

$$T(\bar{x};\Omega)=\bigcup_{i\in\bar{\mathscr{P}}}T(\bar{x};P_i)=\bigcup_{i\in\bar{\mathscr{P}}}\{v\in\mathbb{R}^n\,|\,a_{ij}^Tv\le 0,\ j\in\bar{\mathscr{A}_i}\}$$

and denoting $\mathscr{P}(u):=\{i\in\bar{\mathscr{P}}\,|\,u\in T(\bar{x};P_i)\}$ and $\mathscr{A}_i(u):=\{j\in\bar{\mathscr{A}_i}\,|\,a_{ij}^Tu=0\}$ for $i\in\mathscr{P}(u)$ we have

$$T(u;T(\bar{x};\Omega))=\bigcup_{i\in\mathscr{P}(u)}\{v\in\mathbb{R}^n\,|\,a_{ij}^Tv\le 0,\ j\in\mathscr{A}_i(u)\}. \tag{10}$$

Now let $v\in T(u;T(\bar{x};\Omega))$ and $\xi\in\hat{N}(v;T(u;T(\bar{x};\Omega)))$ be arbitrarily fixed and let $\mathscr{P}^v:=\{i\in\mathscr{P}(u)\,|\,v\in T(u;T(\bar{x};P_i))\}$ and $\mathscr{A}_i^v:=\{j\in\mathscr{A}(u)\,|\,a_{ij}^Tv=0\}$, $i\in\mathscr{P}^v$. Then $\xi\in\bigcap_{i\in\mathscr{P}^v}\hat{N}(v;T(u;T(\bar{x};P_i)))$ and thus for each $i\in\mathscr{P}^v$ there are nonnegative numbers $\mu_{ij}\ge 0$, $j\in\mathscr{A}_i^v$ such that $\xi=\sum_{j\in\mathscr{A}_i^v}\mu_{ij}a_{ij}$. We claim that for all $t>0$ sufficiently small $\xi\in\hat{N}(\bar{x}+tu+t^2v;\Omega)$. To prove this claim it is sufficient to show $\mathscr{P}^v=\{i\in\{1,\ldots,\bar{p}\}\,|\,\bar{x}+tu+t^2v\in P_i\}$ and $\mathscr{A}_i^v=\{j\in\{1,\ldots,m_i\}\,|\,a_{ij}^T(\bar{x}+tu+t^2v)=b_j\}$, $i\in\mathscr{P}^v$ for all $t>0$ sufficiently small, since then

$$\hat{N}(\bar{x}+tu+t^2v;\Omega)=\bigcap_{i\in\mathscr{P}^v}\{\sum_{j\in\mathscr{A}_i^v}\mu_{ij}a_{ij}\,|\,\mu_{ij}\ge 0, j\in\mathscr{A}_i^v\}.$$

Let $i\in\mathscr{P}^v$ and $j\in\{1,\ldots,m_i\}$. The index set $\{1,\ldots,m_i\}$ can be partitioned into the four sets $J_1=\{1,\ldots,m_i\}\setminus\bar{\mathscr{A}_i}$, $J_2:=\bar{\mathscr{A}_i}\setminus\mathscr{A}_i(u)$, $J_3:=\mathscr{A}_i(u)\setminus\mathscr{A}_i^v$ and $\mathscr{A}_i^v$. If $j\in\mathscr{A}_i^v$ we have $a_{ij}^Tv=a_{ij}^Tu=a_{ij}\bar{x}-b_j=0$, if $j\in J_3$ we have $a_{ij}^Tv<a_{ij}^Tu=a_{ij}\bar{x}-b_j=0$, if $j\in J_2$ we have $a_{ij}^Tu<a_{ij}\bar{x}-b_j=0$ and finally $a_{ij}\bar{x}-b_j<0$ for $j\in J_1$. This shows $a_{ij}^T(\bar{x}+tu+t^2v)-b_j=0$, $j\in\mathscr{A}_i^v$ and $a_{ij}^T(\bar{x}+tu+t^2v)-b_j<0$, $j\in J_1\cup J_2\cup J_3$ for all $t>0$ sufficiently small and we can conclude $\mathscr{P}^v\subset\{i\in\{1,\ldots,\bar{p}\}\,|\,\bar{x}+tu+t^2v\in P_i\}$ and $\mathscr{A}_i^v=\{j\in\{1,\ldots,m_i\}\,|\,a_{ij}^T(\bar{x}+tu+t^2v)=b_j\}$, $i\in\mathscr{P}^v$. For $i\notin\mathscr{P}^v$ we either have $i\in I_1:=\{1,\ldots,\bar{p}\}\setminus\bar{\mathscr{P}}$ or $i\in I_2:=\bar{\mathscr{P}}\setminus\mathscr{P}(u)$ or $i\in I_3:=\mathscr{P}(u)\setminus\mathscr{P}^v$. If $i\in I_1$ there is some $j\in\{1,\ldots,m_i\}$ with $a_{ij}^T\bar{x}-b_j>0$, if $i\in I_2$ there is some $j\in\bar{\mathscr{A}_i}$ with $0=a_{ij}^T\bar{x}-b_j<a_{ij}^Tu$ and finally for $i\in I_3$ there is some $j\in\mathscr{A}_i(u)$ with $0=a_{ij}^T\bar{x}-b_j=a_{ij}^Tu<a_{ij}^Tv$. Hence there is some j with $a_{ij}^T(\bar{x}+tu+t^2v)>0$ for all $t>0$ sufficiently small and $\mathscr{P}^v\supset\{i\in\{1,\ldots,\bar{p}\}\,|\,\bar{x}+tu+t^2v\in P_i\}$ follows and our claim is proved. Since $\lim_{t\downarrow 0}t^{-1}(\bar{x}+tu+t^2v-\bar{x})=u$ we obtain $\xi\in N(\bar{x};\Omega;u)$ and $\bigcup_{v\in T(u;T(\bar{x};\Omega))}\hat{N}(v;T(u;T(\bar{x};\Omega)))\subset N(\bar{x};\Omega;u)$ follows.

To show the reverse inclusion, let $\xi\in N(\bar{x};\Omega;u)$ and consider sequences $(t_k)\downarrow 0$, $(u_k)\to u$ and $(\xi_k)\to\xi$ such that $\xi_k\in\hat{N}(\bar{x}+t_ku_k;\Omega)$. Then for all k sufficiently large we have $\mathscr{P}^k:=\{i\in\{1,\ldots,\bar{p}\}\,|\,\bar{x}+t_ku_k\in P_i\}\subset\mathscr{P}(u)$ and $\mathscr{A}_i^k:=\{j\in\{1,\ldots,m_i\}\,|\,a_{ij}^T(\bar{x}+t_ku_k)=b_i\}=\{j\in\bar{\mathscr{A}_i}\,|\,a_{ij}^Tu_k=0\}\subset\mathscr{A}_i(u)$, $i\in\mathscr{P}^k$ and

$$\xi_k\in\bigcap_{i\in\mathscr{P}^k}\hat{N}(\bar{x}+t_ku_k;P_i)=\bigcap_{i\in\mathscr{P}^k}\{\sum_{j\in\mathscr{A}_i^k}\mu_{ij}a_{ij}\,|\,\mu_{ij}\ge 0\}.$$

It follows that $a_{ij}^T u_k \le 0,\ j \in \mathscr{A}_i(u),\ i \in \mathscr{P}^k$ and hence $u_k \in T(u;T(\bar{x};P_i)),\ i \in \mathscr{P}^k$. Since there are only finitely many subsets of $\{1,\ldots,\bar{p}\}$ and $\{1,\ldots,m_i\},\ i \in \{1,\ldots,\bar{p}\}$ we can assume, by eventually passing to a subsequence, that there are index sets $\mathscr{P}^\xi \subset \mathscr{P}(u),\ \mathscr{A}_i^\xi \subset \mathscr{A}_i(u),\ i \in \mathscr{P}^\xi$ such that $\mathscr{P}^k = \mathscr{P}^\xi,\ \mathscr{A}_i^k = \mathscr{A}_i^\xi,\ i \in \mathscr{P}^\xi$ for all k. Because the normal cones $\hat{N}(\bar{x}+t_k u_k;P_i)$, $i \in \mathscr{P}^\xi$ are closed, we obtain $\xi \in \hat{N}(\bar{x}+t_k u_k;P_i),\ i \in \mathscr{P}^\xi$. Now let k be arbitrarily fixed. For every $i \in \mathscr{P}(u) \setminus \mathscr{P}^\xi$ there is some index $j_i \in \mathscr{A}_i(u)$ with $a_{ij_i}^T u_k > 0$ and therefore we can find $\delta > 0$ such that for every $v \in T(u;T(\bar{x};\Omega)) \cap (u_k + \delta \mathscr{B}_{\mathbb{R}^n})$ we have $v \notin T(\bar{x};P_i),\ i \in \mathscr{P}(u) \setminus \mathscr{P}^\xi$ showing $\mathscr{P}^v := \{i \in \mathscr{P}(u) \,|\, v \in T(u;T(\bar{x};P_i))\} \subset \mathscr{P}^\xi$. Thus, for every $i \in \mathscr{P}^v$ there are nonnegative numbers $\mu_{ij} \ge 0,\ j \in \mathscr{A}_i^\xi$ with $\xi = \sum_{j \in \mathscr{A}_i^\xi} \mu_{ij} a_{ij}$ implying

$$\xi^T(v-u_k) = \sum_{j \in \mathscr{A}_i^\xi} \mu_{ij} a_{ij}^T (v-u_k) = \sum_{j \in \mathscr{A}_i^\xi} \mu_{ij} a_{ij}^T v \le 0$$

because of $a_{ij}^T v \le 0,\ j \in \mathscr{A}_i(u) \supset \mathscr{A}_i^\xi$ and we conclude $\xi \in \hat{N}(u_k, T(u;T(\bar{x};\Omega)))$. This finishes the proof. □

In what follows we consider the notions of metric regularity and subregularity, respectively, and its characterization by coderivatives and Mordukhovich normal cones.

Recall that a multifunction $M : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is called *metrically regular* with modulus $\kappa > 0$ near the point $(\bar{x},\bar{y}) \in \operatorname{gph} M$ from its graph, provided there exist neighborhoods U of $\bar{x}$ and V of $\bar{y}$ such that

$$\mathrm{d}(x, M^{-1}(y)) \le \kappa \mathrm{d}(y, M(x)) \quad \forall (x,y) \in U \times V. \tag{11}$$

Here $\mathrm{d}(x,\Omega)$ denotes the usual distance between a point x and a set Ω.

It is well known that metric regularity of the multifunction M near $(\bar{x},\bar{y})$ is equivalent to the Aubin property of the inverse multifunction M^{-1}. A multifunction $S : R^m \rightrightarrows R^n$ has the *Aubin property* with modulus $L \ge 0$ near some point $(\bar{y},\bar{x}) \in \operatorname{gph} S$, if there are neighborhoods U of $\bar{x}$ and V of $\bar{y}$ such that

$$S(y_1) \cap U \subset S(y_2) + L\|y_1 - y_2\| \mathscr{B}_{\mathbb{R}^n} \quad \forall y_1, y_2 \in V.$$

We refer to the monographs [33, 34, 30, 42] and the survey [26] for an extensive treatment of these subjects and the related notions of *pseudo-Lipschitz continuity*, *Lipschitz-like property* and *openness with a linear rate*.

Metric regularity can be equivalently characterized by the so-called *Mordukhovich criterion* (cf. [33, Theorem 4.18], [42, Theorem 9.43]):

Theorem 1. *For a multifunction $M : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ with closed graph and any $(\bar{x},\bar{y}) \in \operatorname{gph} M$ the following statements are equivalent:*

1. *M is metrically regular near $(\bar{x},\bar{y})$.*

2. $\ker D^*M(\bar{x},\bar{y}) = \{0\}$, *i.e.* $0 \in D^*M(\bar{x},\bar{y})(\eta) \Rightarrow \eta = 0$.

Applying this criterion to multifunctions of the form $M(x) = F(x) - \Omega$ we obtain the following collorary, see e.g. [42]:

Corollary 1. *Let $M : R^n \to \mathbb{R}^m$, $M(x) = F(x) - \Omega$ be a multifunction, where $F : \mathbb{R}^n \to \mathbb{R}^m$ is continuously differentiable, $\Omega \subset \mathbb{R}^m$ is closed and let $\bar{x} \in \mathbb{R}^n$ be given with $F(\bar{x}) \in \Omega$. Then M is metrically regular near $(\bar{x}, 0)$ if and only if*

$$\nabla F(\bar{x})^T \lambda = 0,\ \lambda \in N(F(\bar{x}); \Omega)\ \Rightarrow \lambda = 0 \tag{12}$$

Among other things metric regularity is important in the context of constraint qualifications:

Example 1. *Consider a system of inequalities and equalities*

$$g(x) \le 0,\ h(x) = 0$$

with continuously differentiable functions $g : \mathbb{R}^n \to \mathbb{R}^l$, $h : \mathbb{R}^n \to \mathbb{R}^p$. Recall that at a solution $\bar{x}$ the Mangasarian Fromovitz constraint qualification *(MFCQ) is said to hold, if the gradients $\nabla h_i(\bar{x})$ are linearly independent and there exists a vector $z \in \mathbb{R}^n$ with*

$$\nabla h(\bar{x})z = 0, \nabla g_i(\bar{x})z < 0,\ i \in I(\bar{x}),$$

where $I(\bar{x}) = \{i \,|\, g_i(\bar{x}) = 0\}$ denotes the index set of active inequalities.

It is well known [39] that MFCQ is equivalent with metric regularity of the multifunction $M : \mathbb{R}^n \rightrightarrows \mathbb{R}^l \times \mathbb{R}^p$

$$M(x) := (g(x), h(x)) - \mathbb{R}^l_- \times \{0\}^p$$

near $(\bar{x}, 0)$. Straightforward calculations yield that condition (12) reads as

$$\sum_{i \in I(\bar{x})} \nabla g_i(\bar{x}) \eta^g_i + \sum_{i=1}^{p} \nabla h_i(\bar{x}) \eta^h_i = 0, \eta^g_i \ge 0, i \in I(\bar{x})\ \Rightarrow \eta^g_i = 0, i \in I(\bar{x}), \eta^h_i = 0,\ i = 1, \ldots, p$$

which is also called positive linear independence constraint qualification.

Condition (12) appears under different names in the literature. E.g. in [14], [44], [46] it is called *no nonzero abnormal multiplier constraint qualification* (NNAMCQ), whereas in [7] it is called *generalized Mangasarian-Fromovitz constraint qualification* (GMFCQ).

When fixing $y = \bar{y}$ in (11) we obtain the weaker property of *metric subregularity* of M at $(\bar{x}, \bar{y})$, i.e. we require the estimate

$$\mathrm{d}(x, M^{-1}(\bar{y})) \le \kappa \mathrm{d}(\bar{y}, M(x)) \quad \forall x \in U \tag{13}$$

with some neighborhood U of $\bar{x}$ and a positive real $\kappa > 0$.

The metric subregularity property was introduced by Ioffe [25],[26] using the terminology "regularity at a point". The notation "metric subregularity" was suggested in [4]. It is well known [4] that metric subregularity of M at $(\bar{x}, \bar{y})$ is equivalent to *calmness* of the inverse multifunction M^{-1} at $(\bar{y}, \bar{x})$. Criteria for subregularity and calmness, respectively, can be found e.g.

in the papers [5], [10], [13], [16],[17],[19],[20],[27],[31], [47], [48]. An important subclass of multifunctions which are known to be metrically subregular at every point of its graph, is given by polyhedral multifunctions, i.e. multifunctions whose graph is the union of finitely many polyhedral sets. This result is due to Robinson [40]. An important special case of polyhedral multifunctions is given by linear systems, where subregularity is a consequence of Hoffman's error bound [21], whereas, as pointed out in the example above, metric regularity is equivalent to MFCQ.

We consider also the following concept of mixed metric regularity/subregularity for multifunctions M composed by two multifunctions $M_i:\mathbb{R}^n\rightrightarrows\mathbb{R}^{m_i}$, $i=1,2$, i.e. M has the form

$$M=(M_1,M_2):\mathbb{R}^n\to\mathbb{R}^{m_1}\times\mathbb{R}^{m_2},\ M(x)=M_1(x)\times M_2(x).$$

We say that $M=(M_1,M_2)$ is *mixed metrically regular subregular* at a point $(\bar{x},(\bar{y}_1,\bar{y}_2))\in\operatorname{gph}M$, if there are neighborhoods U of $\bar{x}$ and V_1 of $\bar{y}_1$ such that

$$\mathrm{d}(x,M^{-1}(y_1,\bar{y}_2))\leq\kappa\mathrm{d}((y_1,\bar{y}_2),M(x))\quad\forall(x,y_1)\in U\times V_1.$$

Clearly, mixed metric regularity/subregularity of (M_1,M_2) implies metric subregularity of M.

Theorem 2. *Let $M_i:R^n\to\mathbb{R}^{m_i}$, $M_i(x)=F_i(x)-\Omega_i$, $i=1,2$ be two multifunctions, where $F_i:\mathbb{R}^n\to\mathbb{R}^{m_i}$ is continuously differentiable and $\Omega_i\subset\mathbb{R}^{m_i}$ is the union of finitely many convex polyhedra and let $F_i(\bar{x})\in\Omega_i$. Assume that M_2 is metrically subregular at $(\bar{x},0)$ and that*

$$\nabla F_1(\bar{x})^T\eta_1+\nabla F_2(\bar{x})^T\eta_2=0,\ \eta_i\in N(F_i(\bar{x});\Omega_i),i=1,2\ \Rightarrow\ \eta_1=0.$$

Then the multifunction $M=(M_1,M_2)$ is mixed regular/subregular at $(\bar{x},(0,0))$

Proof. Consider the multifunction $S:\mathbb{R}^n\rightrightarrows\mathbb{R}^{m_1}\times\mathbb{R}^{m_2}$, $S(x)=(-\Omega_1)\times(-\Omega_2)$. Since F_1 is continuously differentiable, it is also Lipschitz near $\bar{x}$ and therefore M_1 has the Aubin property near $(\bar{x},0)$. In finite dimensions every linear operator has closed range. Hence we can invoke [12, Lemma 2.6] together with [12, Theorem 4.3] to obtain that the condition

$$0\in\nabla F_1(\bar{x})^T\eta_1+\nabla F_2(\bar{x})^T\eta_2+D^*S(\bar{x},(-F_1(\bar{x}),-F_2(\bar{x})))(\eta_1,\eta_2)\ \Rightarrow\ \eta_1=0$$

is sufficient for mixed regularity/subregularity of M. Since

$$D^*S(\bar{x},(-F_1(\bar{x}),-F_2(\bar{x})))(\eta_1,\eta_2)=\begin{cases}\{0\} & \text{if } \eta_i\in N(F_i(\bar{x});\Omega_i),\ i=1,2,\\ \emptyset & \text{else,}\end{cases}$$

the assertion follows. □

For our analysis we also need the notion of directional metric subregularity. To define this property it is convenient to introduce the following neighborhoods of directions: Given a direction $u\in\mathbb{R}^n$ and positive numbers $\varepsilon,\delta>0$, the set $V_{\varepsilon,\delta}(u)$, is given by

$$V_{\varepsilon,\delta}(u):=\{z\in\varepsilon\mathscr{B}_{\mathbb{R}^n}\mid\big\|\|u\|z-\|z\|u\big\|\leq\delta\|z\|\,\|u\|\}. \tag{14}$$

This can also be written in the form

$$V_{\varepsilon,\delta}(u)=\begin{cases}\{0\}\cup\{z\in\varepsilon\mathscr{B}_{\mathbb{R}^n}\setminus\{0\}\mid \left\|\frac{z}{\|z\|}-\frac{u}{\|u\|}\right\|\le\delta\} & \text{if } u\neq 0,\\ \varepsilon\mathscr{B}_{\mathbb{R}^n} & \text{if } u=0.\end{cases}$$

Given $u\in\mathbb{R}^n$, the multifunction $M:\mathbb{R}^n\rightrightarrows\mathbb{R}^m$ is said to be *metrically subregular in direction* u at $(\bar{x},\bar{y})\in\operatorname{gph}M$, if there are positive reals $\rho>0,\delta>0$ and $\kappa>0$ such that

$$\mathrm{d}(x,M^{-1}(\bar{y}))\le\kappa\mathrm{d}(\bar{y},M(x)) \tag{15}$$

holds for all $x\in\bar{x}+V_{\rho,\delta}(u)$.

Note that metric subregularity in direction 0 is equivalent to the property of metric subregularity.

Theorem 3. *Let $M_i:R^n\to\mathbb{R}^{m_i}$, $M_i(x)=F_i(x)-\Omega_i$, $i=1,2$ be two multifunctions, where F_i is continuously differentiable and Ω_i is the union of finitely many convex polyhedra and let $F_i(\bar{x})\in\Omega_i$. Further, given $u\in\mathbb{R}^n$ assume that M_2 is metrically subregular in direction u at $(\bar{x},0)$.*

1. *If*

$$\nabla F_1(\bar{x})^T\eta_1+\nabla F_2(\bar{x})^T\eta_2=0,\ \eta_i\in N(F_i(\bar{x});\Omega_i;\nabla F_i(\bar{x})u),i=1,2\ \Rightarrow\ \eta_1=0,$$

then the multifunction $M=(M_1,M_2)$ is metrically subregular in direction u at $(\bar{x},(0,0))$.

2. *Assume that F_i, $i=1,2$ are twice Fréchetdifferentiable at $\bar{x}$ and $u\neq 0$. If*

$$u^T\nabla^2(\eta_1^TF_1+\eta_2^TF_2)(\bar{x})u<0$$

holds for all $(\eta_1,\eta_2)\in N(F_1(\bar{x});\Omega_1;\nabla F_1(\bar{x})u)\times N(F_2(\bar{x});\Omega_2;\nabla F_2(\bar{x})u)$ with $\nabla F_1(\bar{x})^T\eta_1+\nabla F_2(\bar{x})^T\eta_2=0$ and $\eta_1\neq 0$, then the multifunction $M=(M_1,M_2)$ is metrically subregular in direction u at $(\bar{x},(0,0))$.

Proof. Using similar arguments as in the proof of Theorem 2, the assertion follows from [12, Theorem 4.3] together with [12, Lemma 2.6] by taking into account

$$\begin{aligned}&D^*S((\bar{x},(-F_1(\bar{x}),-F_2(\bar{x})));(u,(-\nabla F_1(\bar{x})u,-\nabla F_2(\bar{x})u)))(\eta_1,\eta_2)\\&=\begin{cases}\{0\} & \text{if } \eta_i\in N(F_i(\bar{x});\Omega_i;\nabla F_i(\bar{x})u),\ i=1,2,\\ \emptyset & \text{else,}\end{cases}\end{aligned}$$

□

Characterization of directional metric subregularity also yields a characterization of metric subregularity:

Lemma 3. *Let $M:\mathbb{R}^n\rightrightarrows\mathbb{R}^m$ be a multifunction and $(\bar{x},\bar{y})\in\operatorname{gph}M$. Then M is metrically subregular at $(\bar{x},\bar{y})$ if and only if it is metrically subregular in every direction $u\neq 0$ at $(\bar{x},\bar{y})$.*

Proof. The "only if"-part is obviously true by the definition. We prove the if-part by contraposition. Assume that M is not metrically subregular at $(\bar{x},\bar{y})$. Then we can find a sequence $(x_k)\to\bar{x}$ satisfying $\mathrm{d}(x_k,M^{-1}(\bar{y}))>k\mathrm{d}(\bar{y},M(x_k))$ for all k. Since $\bar{x}\in M^{-1}(\bar{y})$ we conclude $x_k\neq\bar{x}$ and therefore $u_k=\frac{x_k-\bar{x}}{\|x_k-\bar{x}\|}$ is well defined. By eventually passing to a subsequence, we can assume that the sequence (u_k) converges to some element $u\in\mathbb{R}^n$ with $\|u\|=1$. Now let $\rho>0$ and $\delta>0$ be arbitrarily fixed. Then for all k sufficiently large we have $x_k\in\bar{x}+\rho B_{\mathbb{R}_n}$ and

$$\left\|\frac{x_k-\bar{x}}{\|x_k-\bar{x}\|}-\frac{u}{\|u\|}\right\|=\|u_k-u\|\leq\delta,$$

showing $x_k\in\bar{x}+V_{\rho,\delta}(u)$. Hence M is not metrically subregular in direction u. □

3 Optimality conditions for the disjunctive program

Now we apply the results of the preceding section to the problem (1). We denote the feasible region of (1) by $\mathscr{F}$ and for a feasible point $\bar{x}\in\mathscr{F}$ we define the *linearized cone* by

$$T_{\mathrm{lin}}(\bar{x}):=\{u\in\mathbb{R}^n\,|\,\nabla F(\bar{x})u\in T(F(\bar{x});\Omega)\}$$

and the *cone of critical directions* by

$$\mathscr{C}(\bar{x}):=\{u\in T_{\mathrm{lin}}(\bar{x})\,|\,\nabla f(\bar{x})u\leq 0\}.$$

Note that always $0\in\mathscr{C}(\bar{x})$ and that $T(\bar{x};\mathscr{F})\subset T_{\mathrm{lin}}(\bar{x})$.

To state our optimality conditions in a general framework we consider for arbitrary $\eta\in\mathbb{R}^n$ and $\bar{x}\in\mathscr{F}$ the multifunction $\mathscr{M}^{\eta,\bar{x}}:=(\mathscr{M}_1^{\eta,\bar{x}},\mathscr{M}_2):\mathbb{R}^n\to\mathbb{R}\times\mathbb{R}^m$ given by

$$\mathscr{M}_1^{\eta,\bar{x}}(x)=f(x)-f(\bar{x})+(\eta^T(x-\bar{x}))^3-\mathbb{R}_-,\ \mathscr{M}_2(x):=M(x).$$

Proposition 1. *Let $\bar{x}$ be a local minimizer for* (1)*. Then $\mathscr{M}^{0,\bar{x}}$ is not mixed regular/subregular at $(\bar{x},0)$ and for every nonzero critical direction $0\neq u\in\mathscr{C}(\bar{x})$ there exists some η such that $\mathscr{M}^{\eta,\bar{x}}$ is not metrically subregular in direction u.*

Proof. Follows from [12, Proposition 5.1]. □

Throughout this section we assume that for every $u\in T_{\mathrm{lin}}(\bar{x})$ the constraint mapping $M(x)=F(x)-\Omega$ can be split into two parts $M=(M_1,M_2):\mathbb{R}^n\rightrightarrows\mathbb{R}^{m_1}\times\mathbb{R}^{m_2}$ with $M_i(x)=F_i(x)-\Omega_i$ and $m=m_1+m_2$, where for each $i\in\{1,2\}$ the mapping $F_i:\mathbb{R}^n\to\mathbb{R}^{m_i}$ is continuously differentiable, $\Omega_i\subset\mathbb{R}^{m_i}$ is the union of finitely many convex polyhedra, $F=(F_1,F_2)$, $\Omega=\Omega_1\times\Omega_2$ and M_2 is metrically subregular in direction u at the point $(\bar{x},0)$ with modulus $\kappa_2(u)$.

This assumption is e.g. automatically fulfilled if F_2 is affine linear, because then M_2 is a polyhedral multifunction and therefore metrically subregular at every point of its graph [40]. If we cannot identify some part of the multifunction which is metrically subregular in the considered direction u then we can simply take $m_2=0$. Note that this splitting is not unique and to ease the notation we also suppress the dependence on u.

We define the *generalized Lagrangian* $\mathscr{L} : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}$ by

$$\mathscr{L}(x, \lambda_0, \lambda) := \lambda_0 f(x) + \lambda^T F(x)$$

Given $\bar{x} \in \mathscr{F}$, $u \in T_{\text{lin}}(\bar{x})$ and $\alpha \geq 0$, we define the sets of multipliers

$$\Lambda^{\lambda_0}(\bar{x};u) := \left\{ \lambda = (\lambda_1, \lambda_2) \in \mathbb{R}^{m_1} \times \mathbb{R}^{m_2} \mid \begin{array}{l} \lambda \in N(F(\bar{x}); \Omega; \nabla F(\bar{x})u), \\ \nabla_x \mathscr{L}(\bar{x}, \lambda_0, \lambda) = 0, \\ \lambda_0 + \|\lambda_1\| \neq 0 \end{array} \right\}$$

and

$$\hat{\Lambda}^{\lambda_0}(\bar{x};u) := \left\{ \lambda = (\lambda_1, \lambda_2) \in \mathbb{R}^{m_1} \times \mathbb{R}^{m_2} \mid \begin{array}{l} \lambda \in \hat{N}(\nabla F(\bar{x})u; T(F(\bar{x}); \Omega)), \\ \nabla_x \mathscr{L}(\bar{x}, \lambda_0, \lambda) = 0, \\ \lambda_0 + \|\lambda_1\| \neq 0 \end{array} \right\}.$$

By Lemma 2 we have

$$\hat{N}(\nabla F(\bar{x})u; T(F(\bar{x}); \Omega)) = \hat{N}(0; T(\nabla F(\bar{x})u; T(F(\bar{x}); \Omega))) \subset N(F(\bar{x}); \Omega; \nabla F(\bar{x})u)$$

and thus $\hat{\Lambda}^{\lambda_0}(\bar{x};u) \subset \Lambda^{\lambda_0}(\bar{x};u)$.

If $u = 0$ we set $\Lambda^{\lambda_0}(\bar{x}) := \Lambda^{\lambda_0}(\bar{x};0)$, $\hat{\Lambda}^{\lambda_0}(\bar{x}) := \hat{\Lambda}^{\lambda_0}(\bar{x};0)$.

We see from the definition that the splitting $M = (M_1, M_2)$ only influences the sets $\Lambda^0(\bar{x};u)$ and $\hat{\Lambda}^0(\bar{x};u)$ by the requirement that certain components of the multipliers are not all zero.

Note that $N(F(\bar{x}); \Omega; \nabla F(\bar{x})u) = N(F_1(\bar{x}); \Omega_1; \nabla F_1(\bar{x})u) \times N(F_2(\bar{x}); \Omega_2; \nabla F_2(\bar{x})u)$ and that for every critical direction u and every $\lambda_0 \geq 0$ such that $\Lambda^{\lambda_0}(\bar{x};u) \neq \emptyset$ we have

$$\lambda_0 \nabla f(\bar{x})u = -\lambda^T \nabla F(\bar{x})u = 0 \tag{16}$$

because of Lemma 1.

We are now in the position to state our main result on first-order and second-order necessary optimality conditions:

Theorem 4. *Let $\bar{x}$ be a local minimizer for the problem* (1) *and let $u \in \mathscr{C}(\bar{x})$. Then there exists $\lambda_0 \geq 0$ such that $\Lambda^{\lambda_0}(\bar{x};u) \neq \emptyset$. If f and F are twice Fréchetdifferntiable at $\bar{x}$ then there exist some $\lambda \in \Lambda^{\lambda_0}(\bar{x};u)$ with*

$$u^T \nabla_x^2 \mathscr{L}(\bar{x}, \lambda_0, \lambda)u \geq 0. \tag{17}$$

If M is metrically subregular in direction u at $(\bar{x}, 0)$ then these conditions hold with $\lambda_0 = 1$.

Proof. First consider the case $u = 0$: From Proposition 1 we know that $(\mathscr{M}_1^{0,\bar{x}}, M)$ is not mixed regular/subregular at $(\bar{x}, 0)$. If M is subregular at $(\bar{x}, 0)$, then it follows from Theorem 2 that there exist $0 \neq \lambda_0 \in N(0; \mathbb{R}_-)$ and $\lambda = (\lambda_1, \lambda_2) \in N(F(\bar{x}); \Omega) = N(F_1(\bar{x}); \Omega_1) \times N(F_2(\bar{x}); \Omega_2)$ with $\lambda_0 \nabla f(\bar{x})^T + \nabla F(\bar{x})^T \lambda = 0$. Hence $\lambda_0 > 0$ and since $N(F(\bar{x}), \Omega)$ is a cone, it follows that $\frac{\lambda}{\lambda_0} = (\frac{\lambda_1}{\lambda_0}, \frac{\lambda_2}{\lambda_0}) \in \Lambda^1(\bar{x})$. If $M = (M_1, M_2)$ is not metrically subregular at $(\bar{x}, 0)$, then it is also not mixed regular/subregular at $(\bar{x}, 0)$ and by Theorem 2 we obtain $\Lambda^0(\bar{x}) \neq \emptyset$. Of course, the second-order condition (17) is trivially fulfilled for $u = 0$.

Now let $u \neq 0$. If M is metrically subregular in direction u at $(\bar{x}, 0)$, we choose η such that $(\mathscr{M}_1^{\eta,\bar{x}}, M)$ is not metrically subregular in direction u according to Proposition 1 and apply Theorem 3. Therefore there exists $0 \neq \lambda_0 \in N(0; \mathbb{R}_-; \nabla f(\bar{x})u)$ and $\lambda = (\lambda_1, \lambda_2) \in N(F(\bar{x}); \Omega; \nabla F(\bar{x})u) = N(F_1(\bar{x}); \Omega_1; \nabla F_1(\bar{x})u) \times N(F_2(\bar{x}); \Omega_2; \nabla F_2(\bar{x})u)$ with $\lambda_0 \nabla f(\bar{x})^T + \nabla F(\bar{x})^T \lambda = 0$. In addition, if F and f are twice Fréchetdifferentiable, we can assume that

$$u^T \nabla_x x^2 (\lambda_0 f + \lambda^T F)(\bar{x})u = u^T \nabla_x^2 \mathscr{L}(\bar{x}, \lambda_0, \lambda)u \geq 0.$$

Then $\lambda_0 > 0$, $\frac{\lambda}{\lambda_0} = (\frac{\lambda_1}{\lambda_0}, \frac{\lambda_2}{\lambda_0}) \in \Lambda^1(\bar{x}; u)$ and $u^T \nabla_x^2 \mathscr{L}(\bar{x}, 1, \frac{\lambda}{\lambda_0})u \geq 0$.

If M is not metrically subregular in direction u, we can apply Theorem 3 to (M_1, M_2) to conclude that the assertions hold with $(\lambda_1, \lambda_2) \in \Lambda^0(\bar{x}; u)$. □

Definition 1. *Let $\bar{x}$ be feasible for the problem* (1). *We say that*

1. *$\bar{x}$ is* B-stationary *(Bouligand-stationary) if*

$$\nabla f(\bar{x})u \geq 0 \ \forall u \in T(\bar{x}; \mathscr{F}),$$

2. *$\bar{x}$ is* M-stationary *(Mordukhovich-stationary) if*

$$\Lambda^1(\bar{x}) \neq \emptyset,$$

3. *$\bar{x}$ is* extended M-stationary *if*

$$\Lambda^1(\bar{x}; u) \neq \emptyset \ \forall u \in \mathscr{C}(\bar{x}),$$

4. *$\bar{x}$ is S-stationary (strongly stationary) if*

$$\hat{\Lambda}^1(\bar{x}) \neq \emptyset.$$

It is well known that a local minimizer is B-stationary. The B-stationarity condition expresses that at a local minimizer there does not exist any feasible descent direction.

Our definition of B-stationarity corresponds to the definition of B-stationarity for MPECs as can be found in the monograph [32]. The definitions of M-stationarity and S-stationarity were introduced in [7] and are in accordance with the definitions for MPECs [43].

Lemma 4. *If $\bar{x}$ is S-stationary then $\nabla f(\bar{x})u \geq 0 \ \forall u \in T_{\text{lin}}(\bar{x})$ and consequently $\bar{x}$ is B-stationary.*

Proof. Consider an arbitrarily fixed direction $u \in T_{\text{lin}}(\bar{x})$. Since $\nabla F(\bar{x})u \in T(F(\bar{x}); \Omega)$, for every $\lambda \in \hat{\Lambda}^1(\bar{x}) \subset \hat{N}(0; T(F(\bar{x}); \Omega))$ we have $-\nabla f(\bar{x})u = \lambda^T \nabla F(\bar{x})u =\leq 0$. □

From Theorem 4 it follows that a local minimizer is M-stationary if there exist one critical direction u such that the multifunction M associated with the constraints is metrically subregular in direction u. Further, if M is metrically subregular in every critical direction u, then a local minimizer is also extended M-stationary. Note that the requirement that M is metrically subregular in one respectively any critical direction is not a constraint qualification in general, since the objective function is also involved in the definition of critical directions. The only exception is the trivial critical direction $u = 0$, because metric subregularity of M in direction 0 means metric subregularity of M.

Obviously extended M-stationarity implies M-stationarity. We will now show that under a suitable weak constraint qualification extended M- stationarity is equivalent to B-stationarity.

Definition 2. *(cf. [7]) We say that the* generalized *(or dual)* Guignard constraint qualification *(GGCQ) holds at the feasible point $\bar{x} \in \mathscr{F}$ if*

$$\hat{N}(\bar{x};\mathscr{F}) = \hat{N}(0;T_{\text{lin}}(\bar{x})).$$

Recall that a polyhedral cone is finitely generated [41, §19]. For each $i = 1,\ldots,\bar{p}$ the set P_i is polyhedral and therefore both the tangent cone $T(\bar{x};\Omega_i)$ and the cone $L_i := \{u \in \mathbb{R}^n \,|\, \nabla F(\bar{x})u \in T(\bar{x},P_i)\}$ are polyhedral cones and consequently finitely generated. Hence, $\operatorname{conv}(\bigcup_{i=1}^{\bar{p}} L_i) = \operatorname{conv}\{u \in \mathbb{R}^n \,|\, \nabla F(\bar{x})u \in T(\bar{x};\Omega)\}$ is also finitely generated, at least by the union of the generators for L_i, but maybe by a smaller set. That is, there exist a set $\mathscr{U} = \{u_1,\ldots,u_N\} \subset T_{\text{lin}}(\bar{x})$ such that

$$\operatorname{conv} T_{\text{lin}}(\bar{x}) = \{\sum_{i=1}^{N} \alpha_i u_i \,|\, \alpha_i \geq 0, i = 1,\ldots,N\} \tag{18}$$

Theorem 5. *Assume that GGCQ is satisfied at the point $\bar{x} \in \mathscr{F}$ feasible for the problem* (1) *and let* $\operatorname{conv} T_{\text{lin}}(\bar{x})$ *be finitely generated by the set $\mathscr{U} = \{u_1,\ldots,u_N\} \subset T_{\text{lin}}(\bar{x})$. Then the following statements are equivalent:*

(a) $\bar{x}$ is B-stationary.

(b) $\nabla f(\bar{x})u \geq 0 \ \forall u \in T_{\text{lin}}(\bar{x})$.

(c) $\bar{x}$ is extended M-stationary.

(d) For every direction $u \in \mathscr{U} \cap \mathscr{C}(\bar{x})$ there holds $\Lambda^1(\bar{x};u) \neq \emptyset$.

Proof. Using the equivalences

$$\bar{x} \text{ is B-stationary} \Leftrightarrow -\nabla f(\bar{x}) \in \hat{N}(\bar{x};\mathscr{F}), \quad -\nabla f(\bar{x}) \in \hat{N}(0;T_{\text{lin}}(\bar{x})) \Leftrightarrow \nabla f(\bar{x})u \geq 0 \ \forall u \in T_{\text{lin}}(\bar{x})$$

we obtain (a)⇔(b) from GGCQ.

Next we show (b)⇒(c). Statement (b) means that $u = 0$ is a solution of the problem

$$\min \nabla f(\bar{x})u \quad \text{subject to} \quad \nabla F(\bar{x})u \in T(\bar{x};\Omega).$$

Since Ω is the union of finitely many polyhedra, there is a neighborhood U of $F(\bar{x})$ such that $\Omega \cap U = (F(\bar{x}) + T(F(\bar{x});\Omega)) \cap U$ and thus $u = 0$ is a local minimizer of the problem

$$\min \nabla f(\bar{x})u \quad \text{subject to} \quad F(\bar{x}) + \nabla F(\bar{x})u \in \Omega \tag{19}$$

The constraint mapping $u \rightrightarrows F(\bar{x}) + \nabla F(\bar{x})u - \Omega$ is a polyhedral multifunction and therefore metrically subregular at 0 by Robinson's result [42]. Hence we can apply Theorem 4 to obtain that 0 is extended M-stationary for the problem (19). But it is easy to see that extended M-stationarity of $u = 0$ for the problem (19) is equivalent to extended M-stationarity of $\bar{x}$ for the problem (1) and the assertion follows.

The implication (c)⇒(d) is obviously true. Finally we show (d)⇒(b). Since $\Lambda^1(\bar{x};u) \neq \emptyset$ implies $\nabla f(\bar{x})u = 0$ we see that (d) implies $\nabla f(\bar{x})u \geq 0\ \forall u \in \mathscr{U}$. Since $\operatorname{conv} T_{\text{lin}}(\bar{x})$ is generated by $\mathscr{U}$ we obtain $\nabla f(\bar{x})u \geq 0\ \forall u \in \operatorname{conv} T_{\text{lin}}(\bar{x})$ and (b) follows. □

GGCQ is very difficult to verify in general. Hence we present another constraint qualification stronger than GGCQ but verifiable:

Definition 3. *We say that the* weak directional metric subregularity constraint qualification *(WDMSCQ) is satisfied at the point $\bar{x}$ feasible for* (1)*, if there is a finite set $\mathscr{U} \subset T_{\text{lin}}(\bar{x})$ generating* $\operatorname{conv} T_{\text{lin}}(\bar{x})$ *such that $M(x) = F(x) - \Omega$ is metrically subregular in every direction $u \in \mathscr{U}$ at* $(\bar{x}, 0)$.

Proposition 2. *WDMSCQ ⇒ GGCQ.*

Proof. Since $T(\bar{x};\mathscr{F}) \subset T_{\text{lin}}(\bar{x})$ the inclusion $\hat{N}(\bar{x};\mathscr{F}) \supset \hat{N}(0;T_{\text{lin}}(\bar{x}))$ always hold. We prove $\hat{N}(\bar{x};\mathscr{F}) \subset \hat{N}(0;T_{\text{lin}}(\bar{x}))$ by contraposition. Assume that there exist some $\xi \in \hat{N}(\bar{x};\mathscr{F}) \setminus \hat{N}(0;T_{\text{lin}}(\bar{x}))$. Then $\xi^T u > 0$ for some $u \in T_{\text{lin}}(\bar{x})$. Since u can be represented as a nonnegative linear combination of $u_1, \ldots, u_N$, there exist $\tilde{u} \in \mathscr{U}$ with $\xi^T \tilde{u} > 0$. Because Ω is the union of finitely many polyhedral sets, there is some neighborhood U of $F(\bar{x})$ such that $\Omega \cap U = (F(\bar{x}) + T(F(\bar{x});\Omega)) \cap U$ and therefore $F(\bar{x}) + t\nabla F(\bar{x})\tilde{u} \in \Omega$ for all $t \geq 0$ sufficiently small. Since M is assumed to be metrically subregular in direction $\tilde{u}$ there is some $\kappa > 0$ such that

$$\mathrm{d}(\bar{x}+t\tilde{u},\mathscr{F}) = \mathrm{d}(\bar{x}+t\tilde{u},M^{-1}(0)) \leq \kappa \mathrm{d}(0,M(\bar{x}+t\tilde{u})) \leq \kappa\|F(\bar{x})+t\nabla F(\bar{x})\tilde{u} - F(\bar{x}+t\tilde{u})\|$$

holds for all $t \geq 0$ sufficiently small. This implies that for every $t > 0$ we can find some $x_t \in \mathscr{F}$ satisfying

$$0 \leq \limsup_{t\downarrow 0} \|\frac{x_t - \bar{x}}{t} - \tilde{u}\| = \limsup_{t\downarrow 0} \|\frac{x_t - (\bar{x}+t\tilde{u})}{t}\| \leq \limsup_{t\downarrow 0} \kappa \frac{\|F(\bar{x})+t\nabla F(\bar{x})\tilde{u} - F(\bar{x}+t\tilde{u})\|}{t} = 0.$$

Hence $\tilde{u} \in T(\bar{x},\mathscr{F})$ and because of $\xi \in \hat{N}(\bar{x};\mathscr{F})$ we have $\xi^T \tilde{u} \leq 0$ contradicting $\xi^T \tilde{u} > 0$. □

Note that Theorem 3 provides pointbased conditions to verify WDMSCQ. We reformulate these conditions in the following lemma:

Lemma 5. *Let $\bar{x}$ be feasible for the problem* (1) *and let $u \in T_{\text{lin}}(\bar{x})$. If either*

1. *$\Lambda^0(\bar{x};u) = \emptyset$, or*
2. *F is twice Fréchetdifferentiable at $\bar{x}$ and $u^T \nabla_x^2 \mathscr{L}(\bar{x},0,\lambda)u < 0\ \forall \lambda \in \Lambda^0(\bar{x};u)$,*

then M is metrically subregular in direction u.

This lemma states that, if for a critical direction u either the first-order necessary optimality condition or the second-order necessary optimality conditions cannot be fulfilled with multiplier $\lambda_0 = 0$, then the constraint mapping M is metrically subregular in direction u.

Example 2. *Consider the nonlinear programming problem*

$$\begin{array}{rcl} \min & & -x_1 \\ -x_1 + |x_1|^{\frac{3}{2}} & \leq & 0, \\ -x_2 & \leq & 0, \\ (x_1 - 2x_2)(x_2 - 2x_1) & \leq & 0, \end{array}$$

at $\bar{x} = (0,0)$. Then $\bar{x}$ is not a local minimizer and we will demonstrate how this can be verified by our necessary conditions.

Consider the multifunction $\tilde{M}(x) := \tilde{F}(x) - \tilde{\Omega} := (x_1 - 2x_2)(x_2 - 2x_1) - \mathbb{R}_-$ and let $u = (u_1, u_2) \in T_{\text{lin}}(\bar{x}) = \mathbb{R}^2_+$ with $(u_1 - 2u_2)(u_2 - 2u_1) < 0$. We shall now show by using Lemma 5 that $\tilde{M}$ is metrically subregular in direction u at $(\bar{x}, 0)$. Straightforward calculations yield that the corresponding set of multipliers is $\tilde{\Lambda}^0(\bar{x}; u) = \mathbb{R}_+ \setminus \{0\}$ and for every $\lambda > 0$ we have $u^T \nabla^2_x \mathscr{L}(\bar{x}, 0, \lambda) u = 2\lambda(u_1 - 2u_2)(u_2 - 2u_1) < 0$ establishing directional metric subregularity.

Hence, for $u = (u_1, u_2) \in T_{\text{lin}}(\bar{x})$ we can use the splitting of the constraint mapping

$$\begin{aligned} M_1(x) &= \begin{pmatrix} -x_1 + |x_1|^{\frac{3}{2}} \\ -x_2 \end{pmatrix} - \mathbb{R}^2_-, \ M_2(x) = (x_1 - 2x_2)(x_2 - 2x_1) - \mathbb{R}_- \ \text{if } (u_1 - 2u_2)(u_2 - 2u_1) < 0, \\ M_1(x) &= \begin{pmatrix} -x_1 + |x_1|^{\frac{3}{2}} \\ -x_2 \\ (x_1 - 2x_2)(x_2 - 2x_1) \end{pmatrix} - \mathbb{R}^3_-, \ M_2(x) = \{0\} \ \text{if } (u_1 - 2u_2)(u_2 - 2u_1) \geq 0. \end{aligned}$$

Now we consider the critical direction $u = (1,0) \in \mathscr{C}(\bar{x}) = T_{\text{lin}}(\bar{x}) = \mathbb{R}^2_+$. The Lagrange function is given by

$$\mathscr{L}(x, \lambda_0, \lambda_1, \lambda_2, \lambda_3) = -\lambda_0 x_1 + \lambda_1(-x_1 + |x_1|^{\frac{3}{2}}) - \lambda_2 x_2 + \lambda_3(x_1 - 2x_2)(x_2 - 2x_1)$$

and for $\lambda_0 \geq 0$ the set $\Lambda^{\lambda_0}(\bar{x}; u)$ is given by

$$\Lambda^{\lambda_0}(\bar{x}; (1,0)) = \left\{ (\lambda_1, \lambda_2, \lambda_3) \mid \begin{array}{c} \nabla_x \mathscr{L}(\bar{x}, \lambda_0, \lambda_1, \lambda_2, \lambda_3) = (-\lambda_0 - \lambda_1, -\lambda_2) = 0 \\ (\lambda_1, \lambda_2, \lambda_3) \in N((-1,0,0); \mathbb{R}^3_-) = \{0\} \times R^2_+ \\ \lambda_0 + |\lambda_1| + |\lambda_2| > 0 \end{array} \right\} = \emptyset.$$

Hence the first-order optimality conditions of Theorem 4 are violated and $\bar{x}$ is not a local minimizer.

Obviously $T_{\text{lin}}(\bar{x}) = \mathbb{R}^2_+$ is generated by the two directions $u = (1,0)^T$ and $v = (0,1)^T$. We have just established $\Lambda^0(\bar{x}; u) = \emptyset$ showing metric subregularity of the constraint mapping in direction u by Lemma 5. In the same way one can also show metric subregularity in direction v and thus WDMSCQ is fulfilled.

Note that the mapping $M(x) = F(x) - \Omega$ is not metrically subregular. E.g., consider the points $x_t := (t,t)$ for arbitrary $t > 0$ satisfying $\lim_{t\downarrow 0}(x_t - \bar{x})/t = (1,1)$, $F(x_t) = (-t + t^{\frac{3}{2}}, -t, t^2) \notin \Omega$, $\mathrm{d}(0, M(x_t)) = t^2$ and $\mathrm{d}(x_t, M^{-1}(0)) = t/\sqrt{5}$ for $0 < t < 1$, showing that M is not metrically subregular in direction $(1,1)$. Similar arguments show that metric subregularity also fail to hold in every direction u with $(u_1 - 2u_2)(u_2 - 2u_1) \geq 0$.

Further, we have $T(\bar{x}; \mathscr{F}) \neq T_{\text{lin}}(\bar{x})$, i.e. the so-called Abadie constriant qualification *fails to hold.*

We consider now second-order sufficient conditions. Consider the following definition owing to Penot [38]:

Definition 4. *We say that $\bar{x} \in \mathbb{R}^n$ is an* essential local minimizer of second order *for problem* (1), *if $\bar{x}$ is feasible and there exists some neighborhood U of $\bar{x}$ and some real $\beta > 0$ such that*

$$\max\{f(x) - f(\bar{x}), \mathrm{d}(F(x), \Omega)\} \geq \beta \|x - \bar{x}\|^2 \ \forall x \in U.$$

Obviously at an essential local minimizer of second order the following *quadratic growth condition* is fulfilled:

$$f(x) \geq f(\bar{x}) + \beta \|x - \bar{x}\|^2 \ \forall x \in \mathscr{F} \cap U.$$

This quadratic growth condition is also sufficient for $\bar{x}$ to be an essential local minimizer of second order, if the constraint mapping M is metrically subregular at $(\bar{x}, 0)$.

Theorem 6. *Assume that $\bar{x}$ is a local minimizer but not an essential local minimizer of second order for the problem* (1). *Then there exists a twice continuously differentiable function $h = (\delta f, \delta F) : \mathbb{R}^n \to \mathbb{R} \times \mathbb{R}^m$ with $h(\bar{x}) = 0$, $\nabla h(\bar{x}) = 0$, $\nabla^2 h(\bar{x}) = 0$ such that $\bar{x}$ is not a local minimizer for the problem*

$$\min (f + \delta f)(\bar{x}) \text{ subject to } (F + \delta F)(x) \in \Omega.$$

Proof. Follows from the proof of [9, Theorem 3.5] by recognizing that convexity of Ω is not needed in that proof. □

From this statement it follows that the property of $\bar{x}$ being an essential local minimizer of second order is the weakest possible sufficient second-order optimality condition which uses solely function values and derivatives up to order 2 at the point $\bar{x}$.

For each $u \in T_{\text{lin}}(\bar{x})$ we now denote by $\mathscr{P}(u)$ the index set

$$\mathscr{P}(u) := \{i \in 1, \ldots, \bar{p} \,|\, \nabla F(\bar{x})u \in T(\bar{x}; P_i)\}$$

Since $T(\bar{x}; \Omega) = \bigcup_{i=1}^{\bar{p}} T(\bar{x}; P_i)$ we have $\mathscr{P}(u) \neq \emptyset$ for every $u \in T_{\text{lin}}(\bar{x})$.

Lemma 6. *Let $\bar{x}$ be feasible for the problem* (1) *and let f, F be twice Fréchet differentiable at $\bar{x}$. Then the following statements are equivalent:*

(a) $\bar{x}$ is an essential local minimizer of second order.

(b) *For every nonzero critical direction* $0 \neq u \in \mathscr{C}(\bar{x})$ *and every* $i \in \mathscr{P}(u)$ *there exists some multiplier* $(\lambda_0, \lambda) \in \hat{N}(\nabla f(\bar{x})u; \mathbb{R}_-) \times \hat{N}(\nabla F(\bar{x})u; T(F(\bar{x}); P_i))$ *with* $\nabla_x \mathscr{L}(\bar{x}, \lambda_0, \lambda) = 0$ *and*

$$u^T \nabla_x^2 \mathscr{L}(\bar{x}, \lambda_0, \lambda) u > 0. \tag{20}$$

(c) *For every nonzero critical direction* $0 \neq u \in \mathscr{C}(\bar{x})$ *there does not exist* $v \in \mathbb{R}^n$ *with*

$$\nabla f(\bar{x})v + \frac{1}{2}u^T \nabla^2 f(\bar{x})u \quad \in \quad T(\nabla f(\bar{x})u; \mathbb{R}_-) \tag{21}$$

$$\nabla F(\bar{x})v + \frac{1}{2}u^T \nabla^2 F(\bar{x})u \quad \in \quad T(\nabla F(\bar{x})u; T(F(\bar{x}); \Omega)) \tag{22}$$

Proof. Let $\bar{\mathscr{P}} := \{i \in \{1, \ldots, \bar{p}\} \,|\, F(\bar{x}) \in P_i\}$. Then there is a neighborhood U of $\bar{x}$ such that $\mathrm{d}(F(x), \Omega) = \min_{i \in \bar{\mathscr{P}}} \mathrm{d}(F(x), P_i)\ \forall x \in U$ and therefore $\bar{x}$ is an essential local minimizer of second order if and only if for each $i \in \bar{\mathscr{P}}$ the point $\bar{x}$ is an essential local minimizer of second order for the problem

$$\min f(x) \text{ subject to } F(x) \in P_i$$

Hence, by using [9, Theorems 5.4,5.11] we obtain that $\bar{x}$ is an essential local minimizer of second order if and only if for each $i \in \bar{\mathscr{P}}$ and every u with $\nabla f(\bar{x})u \leq 0$ and $\nabla F(\bar{x})u \in T(F(\bar{x}); P_i)$ there is some multiplier $(\lambda_0, \lambda) \in \mathbb{R}_+ \times \hat{N}(F(\bar{x}); P_i)$ with $\nabla_x \mathscr{L}(\bar{x}, \lambda_0, \lambda) = 0$ and $u^T \nabla_x^2 \mathscr{L}(\bar{x}, \lambda_0, \lambda) u > 0$. Hence $\lambda_0 \nabla f(\bar{x})u \leq 0$ and $\lambda^T \nabla F(\bar{x})u \leq 0$ and from $\nabla_x \mathscr{L}(\bar{x}, \lambda_0, \lambda)u = 0$ we conclude $\lambda_0 \nabla f(\bar{x})u = -\lambda^T \nabla F(\bar{x})u = 0$. Since for a closed convex cone K and $k \in K$ we have $\hat{N}(k; K) = \{\xi \in \hat{N}(0; K) \,|\, \xi^T k = 0\}$, we conclude $(\lambda_0, \lambda) \in \hat{N}(\nabla f(\bar{x})u; \mathbb{R}_-) \times \hat{N}(\nabla F(\bar{x})u; T(F(\bar{x}); P_i))$ and this establishes the equivalence (a)⇔(b).

To show the equivalence (b)⇔ (c) we use the well-known Farkas Lemma: statement (b) is equivalent that for every $0 \neq u \in \mathscr{C}(\bar{x})$ and each $i \in \mathscr{P}(u)$ the system

$$\nabla f(\bar{x})v + \frac{1}{2}u^T \nabla^2 f(\bar{x})u \quad \in \quad T(\nabla f(\bar{x})u; \mathbb{R}_-)$$

$$\nabla F(\bar{x})v + \frac{1}{2}u^T \nabla^2 F(\bar{x})u \quad \in \quad T(\nabla F(\bar{x})u; T(F(\bar{x}); P_i))$$

does not have a solution v. Noting that $T(\nabla F(\bar{x})u; T(F(\bar{x}); \Omega)) = \bigcup_{i \in \mathscr{P}(u)} T(\nabla F(\bar{x})u; T(F(\bar{x}); P_i))$ we conclude (b)⇔(c). □

Theorem 7. *Let* $\bar{x}$ *be feasible for the problem* (1) *and let* f, F *be twice Fréchet differentiable at* $\bar{x}$. *If* $\bar{x}$ *is an essential local minimizer of second order then for every critical direction* $0 \neq u \in \mathscr{C}(\bar{x})$ *there is some pair* $(\lambda_0, \lambda) \in \mathbb{R}_+ \times \mathbb{R}^m$ *with* $\lambda \in \Lambda^{\lambda_0}(\bar{x}; u)$ *such that*

$$u^T \nabla_x^2 \mathscr{L}(\bar{x}, \lambda_0, \lambda) u > 0. \tag{23}$$

Conversely, if for every critical direction $0 \neq u \in \mathscr{C}(\bar{x})$ *there is some pair* $(\lambda_0, \lambda) \in \mathbb{R}_+ \times \mathbb{R}^m$ *fulfilling* $\lambda \in \hat{\Lambda}^{\lambda_0}(\bar{x}; u)$ *and* (23)*, then* $\bar{x}$ *is an essential local minimizer of second order.*

Proof. Firstly assume that $\bar{x}$ is an essential local minimizer of second order and let $0 \neq u \in \mathscr{C}(\bar{x})$ be arbitrarily fixed. For each $i \in \mathscr{P}(u)$ the set $(\nabla f(\bar{x}), \nabla F(\bar{x}))\mathbb{R}^n + \mathbb{R}_- \times T(\nabla F(\bar{x})u; T(F(\bar{x}); P_i))$ is closed because $T(\nabla F(\bar{x})u; T(F(\bar{x}); P_i))$ is a polyhedral cone and therefore also

$$\Xi := (\nabla f(\bar{x}), \nabla F(\bar{x}))\mathbb{R}^n + \mathbb{R}_- \times T(\nabla F(\bar{x})u; T(F(\bar{x}); \Omega))$$

is closed as the union of these finitely many sets. Owing to Lemma 6 we have

$$\zeta := \frac{1}{2}(u^T \nabla^2 f(\bar{x})u, u^T \nabla^2 F(\bar{x})u) \notin \Xi$$

and hence we can find $\bar{v} \in \mathbb{R}^n$ and $(\bar{\phi}, \bar{\psi}) \in \mathbb{R}_- \times T(\nabla F(\bar{x})u; T(F(\bar{x}); \Omega))$ such that

$$\bar{\xi} := (-\nabla f(\bar{x})\bar{v} + \bar{\phi}, -\nabla F(\bar{x})\bar{v} + \bar{\psi}) \in \Xi$$

is the orthogonal projection of ζ on Ξ and $0 \neq (\lambda_0, \lambda) := \zeta - \bar{\xi} \in \hat{N}(\bar{\xi}; \Xi)$. Since Ξ is the union of finitely many polyhedra there is a neighborhood U of $\bar{\xi}$ such that $(\lambda_0, \lambda)^T(\xi - \bar{\xi}) \leq 0\, \forall \xi \in \Xi \cap U$. Hence $\lambda \in \hat{N}(\bar{\psi}; T(\nabla F(\bar{x})u; T(F(\bar{x}); \Omega))) \subset N(F(\bar{x}); \Omega; \nabla F(\bar{x})u)$ by Lemma 2, $\lambda_0 \in \hat{N}(\bar{\phi}; \mathbb{R}_-)$, and $\lambda_0 \nabla f(\bar{x})(v - \bar{v}) + \lambda^T \nabla F(\bar{x})(v - \bar{v}) \leq 0$ for all v belonging to some neighborhood of $\bar{v}$ yielding $\lambda_0 \geq 0$, $\lambda_0 \bar{\phi} = 0$ and $\nabla_x \mathscr{L}(\bar{x}, \lambda_0, \lambda) = 0$. Since $T(\nabla F(\bar{x})u; T(F(\bar{x}); \Omega))$ is a cone we also obtain $\lambda^T \bar{\psi} = 0$ and consequently

$$(\lambda_0, \lambda)^T \bar{\xi} = \nabla_x \mathscr{L}(\bar{x}, \lambda_0, \lambda)\bar{v} - \lambda_0 \bar{\phi} - \lambda^T \bar{\psi} = 0$$

and

$$0 < \|\zeta - \bar{\xi}\|^2 = (\lambda_0, \lambda)^T(\zeta - \bar{\xi}) = (\lambda_0, \lambda)^T \zeta = \frac{1}{2} u^T \nabla_x^2 \mathscr{L}(\bar{x}, \lambda_0, \lambda)u.$$

There remains to show $\lambda_0 + \|\lambda_1\| > 0$. Observe that $\bar{x}$ is an essential local minimizer of second order if and only if it is an essential local minimizer of second order for the problem

$$\min f^\alpha := \alpha f(x) \text{ subject to } F^\alpha(x) := (\alpha F_1(x), F_2(x)) \in (\alpha \Omega_1) \times \Omega_2 =: \Omega^\alpha \tag{24}$$

for every $\alpha > 0$. Let α be chosen such that $\kappa_2(u)\alpha(\|\nabla f(\bar{x})\| + \|\nabla F(\bar{x})\|) < 1$ and consider the construction of (λ_0, λ) as before. We claim that $\lambda_0 + \|\lambda_1\| > 0$. Indeed, if $\lambda_0 + \|\lambda_1\| = 0$ then $\alpha(\nabla f(\bar{x})\bar{v} + \frac{1}{2}u^T \nabla^2 f(\bar{x})u) \in \mathbb{R}_-$, $\alpha(\nabla F_1(\bar{x})\bar{v} + \frac{1}{2}u^T \nabla^2 F_1(\bar{x})u) \in T(\nabla F_1(\bar{x})u; T(F_1(\bar{x}); \Omega_1))$ and

$$\|\zeta - \bar{\xi}\| = \|(\lambda_0, \lambda)\| = \|\lambda_2\| = \|\nabla F_2(\bar{x})\bar{v} + \frac{1}{2}u^T \nabla^2 F_2(\bar{x})u - \bar{\psi}_2\|$$

Since $\bar{\psi}_2 \in T(\nabla F_2(\bar{x})u; T(F_2(\bar{x}); \Omega_2))$ and $T(F_2(\bar{x}); \Omega_2)$ respectively Ω_2 are the union of finitely many polyhedra we have $\nabla F_2(\bar{x})u + t\bar{\psi}_2 \in T(F_2(\bar{x}); \Omega_2)$ and consequently $F_2(\bar{x}) + t(\nabla F_2(\bar{x})u + t\bar{\psi}_2) \in \Omega_2$ for all $t \geq 0$ sufficiently small. It follows that

$$\begin{aligned}
&\mathrm{d}(F_2(\bar{x} + tu + t^2\bar{v}); \Omega_2) \\
&\quad \leq \|F_2(\bar{x}) + t\nabla F_2(\bar{x})u + t^2(\nabla F_2(\bar{x})\bar{v} + \frac{1}{2}u^T \nabla F_2(\bar{x})u + r(t)) - (F_2(\bar{x}) + t\nabla F_2(\bar{x})u + t^2\bar{\psi}_2)\| \\
&\quad \leq t^2(\|\lambda_2\| + \|r(t)\|)
\end{aligned}$$

with $\lim_{t\to 0} r(t) = 0$. Now choose an arbitrary sequence $(t_k) \downarrow 0$. Since M_2 is assumed to be metrically subregular in direction u we can find for every k some w_k with $\|w_k\| \leq \kappa_2(u)(\|\lambda_2\| + \|r(t_k)\|)$ such that $F_2(\bar{x} + t_k u + t_k^2(\bar{v} + w_k)) \in \Omega_2$. Passing to a subsequence if necessary we can assume that (w_k) converges to some $\bar{w}$ with $\|\bar{w}\| \leq \kappa_2(u)\|\lambda_2\|$ and it follows that $F_2(\bar{x} + t_k u + t_k^2(\bar{v} + w_k)) - F_2(\bar{x}) \in T(F_2(\bar{x}); \Omega_2)$, $t_k^{-1}(F_2(\bar{x} + t_k u + t_k^2(\bar{v} + w_k)) - F_2(\bar{x})) - \nabla F_2(\bar{x})u \in T(\nabla F_2(\bar{x})u; T(F_2(\bar{x}); \Omega_2))$ and therefore also

$$\begin{aligned}\nabla F_2(\bar{x})(\bar{v}+\bar{w}) + \frac{1}{2}u^T \nabla F_2(\bar{x})u &= \lim_{k\to\infty} t_k^{-1}(t_k^{-1}(F_2(\bar{x}+t_k u + t_k^2(\bar{v}+w_k)) - F_2(\bar{x})) - \nabla F_2(\bar{x})u) \\ &\in T(\nabla F_2(\bar{x})u; T(F_2(\bar{x}); \Omega_2))\end{aligned}$$

But then we obtain

$$\mathrm{d}(\zeta, \Xi) \leq \|(\alpha \nabla f(\bar{x})\bar{w}, \alpha \nabla F_1(\bar{x})\bar{w}, 0)\| \leq \alpha(\|\nabla f(\bar{x})\| + \|\nabla F_1(\bar{x})\|)\|\bar{w}\| < \|\lambda_2\| = \|\zeta - \bar{\xi}\|$$

contradicting the choice of $\bar{\xi}$ as the projection of ζ on Ξ. Hence $\lambda_0 + \|\lambda_1\| \neq 0$ and it follows that $(\alpha\lambda_1, \lambda_2) \in \Lambda^{\alpha\lambda_0}(\bar{x}; u)$ and $u^T \nabla_x^2 \mathscr{L}(\bar{x}; \alpha\lambda_0, (\alpha\lambda_1, \lambda_2))u > 0$.

To show the second assertion we use the equivalence (a)$\Leftrightarrow$(b) of Lemma 6. Let $0 \neq u \in \mathscr{C}(\bar{x})$ be arbitrarily fixed and choose $\lambda_0 \geq 0$ and $\lambda \in \hat{\Lambda}^{\lambda_0}(\bar{x}; u)$ with $u^T \nabla_x^2 \mathscr{L}(\bar{x}, \lambda_0, \lambda)u > 0$. By the definition of $\hat{\Lambda}^{\lambda_0}(\bar{x}; u)$ we have $\nabla_x \mathscr{L}(\bar{x}, \lambda_0, \lambda) = 0$ and we will now show that $(\lambda_0, \lambda) \in \hat{N}(\nabla f(\bar{x})u; \mathbb{R}_-) \times \hat{N}(\nabla F(\bar{x})u; T(F(\bar{x}); P_i))$ for each $i \in \mathscr{P}(u)$. Because of $\hat{\Lambda}^{\lambda_0}(\bar{x}; u) \subset \Lambda^{\lambda_0}(\bar{x}; u)$ and (16) we have $\lambda_0 \in \hat{N}(\nabla f(\bar{x})u; \mathbb{R}_-)$. Further

$$\begin{aligned}\lambda &\in \hat{N}(\nabla F(\bar{x})u; T(F(\bar{x}); \Omega)) = \hat{N}(0; T(\nabla F(\bar{x})u; T(F(\bar{x}); \Omega))) \\ &= \hat{N}(0; \bigcup_{i\in\mathscr{P}(u)} T(\nabla F(\bar{x})u; T(F(\bar{x}); P_i))) = \bigcap_{i\in\mathscr{P}(u)} \hat{N}(0; T(\nabla F(\bar{x})u; T(F(\bar{x}); P_i))) \\ &= \bigcap_{i\in\mathscr{P}(u)} \hat{N}(\nabla F(\bar{x})u; T(F(\bar{x}); P_i))\end{aligned}$$

and thus our assertion is proved. □

We will now show that under a certain directional linear independence constraint qualification condition the sets $\hat{\Lambda}^{\lambda_0}(\bar{x}; u)$ and $\Lambda^{\lambda_0}(\bar{x}; u)$ coincide.

Definition 5. *Let $u \in T_{\mathrm{lin}}(\bar{x})$. We say that the* linear independence constraint qualification condition in direction *u (LICQ(u)) holds at $\bar{x}$ if there is some subspace $L \subset R^m$ such that*

$$T(\nabla F(\bar{x})u; T(F(\bar{x}); \Omega)) + L \subset T(\nabla F(\bar{x})u; T(F(\bar{x}); \Omega))$$

and

$$\nabla F(\bar{x})\mathbb{R}^n + L = \mathbb{R}^m.$$

Note that LICQ(0) is related to the nondegeneracy condition [3, (4.172)]

Lemma 7. *Assume that LICQ(u) holds for $u \in T_{\mathrm{lin}}(\bar{x})$. Then $\Lambda^0(\bar{x}; u) = \emptyset$*

Proof. Assume that there is some $\lambda \in \Lambda^0(\bar{x};u)$. Then $\lambda \neq 0$ and there is some $v \in \mathbb{R}^n$ and $w \in L$ such that $\nabla F(\bar{x})v + w = \lambda$, implying

$$0 < \lambda^T \lambda = \lambda^T \nabla F(\bar{x})v + \lambda^T w = \lambda^T w$$

because of $\nabla_x \mathscr{L}(\bar{x};0,\lambda) = \lambda^T \nabla F(\bar{x}) = 0$. By Lemma 2 there is some $z \in T(\nabla F(\bar{x})u; T(F(\bar{x};\Omega)))$ with $\lambda \in \hat{N}(z; T(\nabla F(\bar{x})u; T(F(\bar{x};\Omega))))$ and therefore, since $z + L \subset T(\nabla F(\bar{x})u; T(F(\bar{x});\Omega))$, we obtain the contradiction $\lambda^T w = 0$. □

Proposition 3. *Assume that $\bar{x}$ is B-stationary for the problem* (1). *Then for every critical direction $u \in \mathscr{C}(\bar{x})$ fulfilling LICQ(u) there is a unique element $\lambda_u \in \mathbb{R}^m$ such that*

$$\Lambda^1(\bar{x};u) = \hat{\Lambda}^1(\bar{x};u) = \{\lambda_u\}.$$

Proof. Consider an arbitrarily fixed direction $u \in \mathscr{C}(\bar{x})$ satisfying LICQ(u). We claim that

$$\min\{\nabla f(\bar{x})^T v \,|\, \nabla F(\bar{x})v \in T(\nabla F(\bar{x})u; T(F(\bar{x});\Omega))\} = 0. \tag{25}$$

Indeed, if there were $\bar{v}$ with $\nabla F(\bar{x})\bar{v} \in T(\nabla F(\bar{x})u; T(F(\bar{x});\Omega))$ and $\nabla f(\bar{x})\bar{v} < 0$, then for every $\alpha > 0$ sufficiently small we have $\nabla F(\bar{x})(u + \alpha\bar{v}) \in T(F(\bar{x});\Omega)$ and consequently $F(\bar{x}) + t\nabla F(\bar{x})(u + \alpha\bar{v}) \in \Omega$ for all $t > 0$ sufficiently small, since both Ω and $T(F(\bar{x});\Omega)$ are the union of finitely many polyhedra. By Lemma 7 and Lemma 5 we have that $M(x) = F(x) - \Omega$ is metrically subregular in direction u at $(\bar{x},0)$ and therefore there are $\rho > 0$, $\delta > 0$, $\kappa > 0$ such that for all $x \in \bar{x} + V_{\rho,\delta}(u)$ the inequality (15) holds. We can choose $\alpha > 0$ small enough such that for all $t > 0$ sufficiently small we have $\bar{x} + t(u + \alpha\bar{v}) \in \bar{x} + V_{\rho,\delta}(u)$ and $F(\bar{x}) + t\nabla F(\bar{x})(u + \alpha\bar{v}) \in \Omega$ implying the existence of $w(t)$ with $F(\bar{x} + t(u + \alpha\bar{v} + w(t))) \in \Omega$ and

$$t\|w(t)\| \leq \kappa \mathrm{d}(F(\bar{x} + t(u + \alpha\bar{v})),\Omega) \leq \kappa\|F(\bar{x} + t(u + \alpha\bar{v})) - F(\bar{x}) - t\nabla F(\bar{x})(u + \alpha\bar{v})\|.$$

Thus $\lim_{t\downarrow 0} w(t) = 0$ and $u + \alpha\bar{v} \in T(\bar{x};\mathscr{F})$ follows. But $\nabla f(\bar{x})(u + \alpha\bar{v}) \leq \alpha\nabla f(\bar{x})\bar{v} < 0$ contradicting B-stationarity of $\bar{x}$ and hence our claim is proved. The constraint mapping of (25) is a polyhedral multifunction and hence metrically subregular. Applying the M-stationarity condition at $v = 0$ yields the existence of some multiplier $\tilde{\lambda} \in N(0; T(\nabla F(\bar{x})u; T(F(\bar{x});\Omega)))$ with $\nabla f(\bar{x})^T + \nabla F(\bar{x})^T \tilde{\lambda} = 0$. By Lemma 2 we conclude $N(0; T(\nabla F(\bar{x})u; T(F(\bar{x});\Omega))) \subset N(F(\bar{x});\Omega;\nabla F(\bar{x})u)$ and $\tilde{\lambda} \in \Lambda^1(\bar{x};u) \neq \emptyset$ follows. Next we show that $\Lambda^1(\bar{x};u)$ is a singleton. Assume on the contrary that there are two different elements $\lambda^i \in \Lambda^1(\bar{x};u)$, $i = 1,2$. By Lemma 2 we have $\lambda^i \in \hat{N}(z^i; T(\nabla F(\bar{x})u; T(F(\bar{x});\Omega)))$ with $z^i \in T(\nabla F(\bar{x})u; T(F(\bar{x});\Omega))$ and because of $z^i + L \subset T(\nabla F(\bar{x})u; T(F(\bar{x});\Omega))$ we conclude that λ^i belongs to $L^\perp$, $i = 1,2$. Further $\nabla F(\bar{x})^T \lambda^i = -\nabla f(\bar{x})^T$, $i = 1,2$ and we obtain the contradiction $0 \neq \lambda^1 - \lambda^2 \in \ker \nabla F(\bar{x})^T \cap L^\perp = \mathrm{Im}\,\nabla F(\bar{x})^\perp \cap L^\perp = (\mathrm{Im}\,F(\bar{x}) + L)^\perp = \mathbb{R}^{m\perp} = \{0\}$. Since $\hat{\Lambda}^1(\bar{x};u) \subset \Lambda^1(\bar{x};u)$, it suffices now to show $\hat{\Lambda}^1(\bar{x};u) \neq \emptyset$ in order to prove $\Lambda^1(\bar{x};u) = \hat{\Lambda}^1(\bar{x};u) = \{\tilde{\lambda}\}$. We claim that

$$\min\{\nabla f(\bar{x})^T v \,|\, \nabla F(\bar{x})v \in \mathrm{conv}\, T(\nabla F(\bar{x})u; T(F(\bar{x});\Omega))\} = 0. \tag{26}$$

Assume on the contrary that there is some $\bar{v}$ with $\nabla F(\bar{x})\bar{v} \in \mathrm{conv}\, T(\nabla F(\bar{x})u; T(F(\bar{x});\Omega))$ and $\nabla f(\bar{x})^T \bar{v} < 0$. Then $\nabla F(\bar{x})\bar{v}$ can be represented as a convex combination $\sum_{i=1}^k \mu_i z_i$ of elements

$z_1,\dots,z_k \in T(\nabla F(\bar{x})u; T(F(\bar{x});\Omega))$. Each element z_i can be written in the form $\nabla F(\bar{x})v_i + w_i$ with $w_i \in L$ and we obtain $\nabla F(\bar{x})v_i = z_i - w_i \in T(\nabla F(\bar{x})u; T(F(\bar{x});\Omega))$ and consequently $\nabla f(\bar{x})v_i \geq 0$ because of (25). Then, using $\tilde{\lambda} \in L^\perp$ we obtain the contradiction

$$0 > \nabla f(\bar{x})\bar{v} = -\tilde{\lambda}^T \nabla F(\bar{x})\bar{v} = -\sum_{i=1}^k \mu_i \tilde{\lambda}^T (\nabla F(\bar{x})v_i + w_i) = \sum_{i=1}^k \mu_i \nabla f(\bar{x})v_i \geq 0.$$

Therefore (26) holds true and since $\operatorname{conv} T(\nabla F(\bar{x})u; T(F(\bar{x});\Omega))$ is a polyhedral cone as the convex hull of the union of finitely many polyhedral cones, we obtain that the constraint mapping $v \rightrightarrows \nabla F(\bar{x})v - \operatorname{conv} T(\nabla F(\bar{x})u; T(F(\bar{x});\Omega))$ is metrically subregular at $(0,0)$ Applying now the M-stationarity condition at $v = 0$ yields the existence of some multiplier

$$\begin{aligned} \hat{\lambda} &\in N(0; \operatorname{conv} T(\nabla F(\bar{x})u; T(F(\bar{x});\Omega))) = \hat{N}(0; \operatorname{conv} T(\nabla F(\bar{x})u; T(F(\bar{x});\Omega))) \\ &= \hat{N}(0; T(\nabla F(\bar{x})u; T(F(\bar{x});\Omega))) = \hat{N}(\nabla F(\bar{x})u; T(F(\bar{x});\Omega)) \end{aligned}$$

with $\nabla f(\bar{x})^T + \nabla F(\bar{x})^T \hat{\lambda} = 0$ and therefore $\hat{\lambda} \in \hat{\Lambda}^1(\bar{x};u) \neq \emptyset$ and this completes the proof. □

As a byproduct of Proposition 3 we obtain that B-stationarity implies S-stationarity if LICQ(0) is fulfilled.

We now state a second-order sufficient condition in terms of multipliers belonging to $\Lambda^1(\bar{x};u)$.

Theorem 8. *Assume that $\bar{x}$ is an extended M-stationary solution for* (1)*, f and F are twice Fréchetdifferentiable at $\bar{x}$ and that for every nonzero critical direction $0 \neq u \in \mathscr{C}(\bar{x})$ one has*

$$u^T \nabla_x^2 \mathscr{L}(\bar{x}, 1, \lambda) u > 0 \; \forall \lambda \in \Lambda^1(\bar{x};u). \tag{27}$$

Then $\bar{x}$ is an essential local minimizer of second order.

Proof. By contraposition. Assuming on the contrary that $\bar{x}$ is not an essential local minimizer, by Lemma 6 we can find $0 \neq u \in \mathscr{C}(\bar{x})$ and $v \in \mathbb{R}^n$ fulfilling (21), (22). We now claim that the problem

$$\min_v \nabla f(\bar{x})v \text{ subject to } \nabla F(\bar{x})v + \frac{1}{2}u^T \nabla^2 F(\bar{x})u \in T(\nabla F(\bar{x})u; T(F(\bar{x});\Omega)) \tag{28}$$

has an optimal solution. If there would not exist an optimal solution, because the feasible region is not empty because of (22), we could find a sequence (v^k) feasible for (28) such that $\nabla f(\bar{x})v^k \to -\infty$. Consider the sequence $\tilde{v}^k := v^k/|\nabla f(\bar{x})v^k|$. Then $\mathrm{d}(\nabla F(\bar{x})\tilde{v}^k, T(\nabla F(\bar{x})u; T(F(\bar{x});\Omega))) \to 0$ and since $v \rightrightarrows \nabla F(\bar{x})v - T(\nabla F(\bar{x})u; T(F(\bar{x});\Omega))$ is a polyhedral multifunction and therefore metrically subregular at $(0,0)$, there is a sequence $(\hat{v}^k)$ with $\nabla F(\bar{x})\hat{v}^k \in T(\nabla F(\bar{x})u; T(F(\bar{x};\Omega)$ and $\lim_{k\to\infty}(\tilde{v}^k - \hat{v}^k) = 0$. Fixing $\bar{v} := \hat{v}^k$ for k sufficiently large we have $\nabla f(\bar{x})\bar{v} < -\frac{1}{2}$. Since $\nabla F(\bar{x})(u + \alpha\bar{v}) \in T(F(\bar{x};\Omega)$ for $\alpha > 0$ sufficiently small we have $u + \alpha\bar{v} \in T_{\text{lin}}(\bar{x})$. Together with $\nabla f(\bar{x})(u + \alpha\bar{v}) < -\frac{\alpha}{2} < 0$ we have $u + \alpha\bar{v} \in \mathscr{C}(\bar{x})$ and thus $\Lambda^1(\bar{x}; u + \alpha\bar{v}) \neq \emptyset$ by extended M-stationarity of $\bar{x}$. But from (16) we obtain the contradiction $\nabla f(\bar{x})(u + \alpha\bar{v}) = 0$. Hence the problem (28) has an optimal solution $\tilde{v}$. Since the constraint mapping is a polyhedral multifunction and therefore metrically subregular at $(\tilde{v},0)$, we can apply the M-stationarity conditions at $\tilde{v}$ to find a multiplier

$$\lambda \in N(\nabla F(\bar{x})\tilde{v}+\frac{1}{2}u^T\nabla^2 F(\bar{x})u; T(\nabla F(\bar{x})u; T(F(\bar{x});\Omega))) \subset N(F(\bar{x});\Omega;\nabla F(\bar{x})u)$$

with $\nabla f(\bar{x})^T+\nabla F(\bar{x})^T\lambda = \nabla_x\mathscr{L}(\bar{x};1,\lambda)^T=0$ showing $\lambda\in\Lambda^1(\bar{x};u)$. Using extended M-stationarity of $\bar{x}$ and (16) we obtain $\nabla f(\bar{x})u=0$ and therefore $\nabla f(\bar{x})\tilde{v}+\frac{1}{2}u^T\nabla^2 f(\bar{x})u\leq 0$ because of (21). Because $T(\nabla F(\bar{x})u;T(F(\bar{x});\Omega))$ is a cone we have $\lambda^T(\nabla F(\bar{x})\tilde{v}+\frac{1}{2}u^T\nabla^2F(\bar{x})u)=0$ and thus

$$\begin{aligned} 0 &\geq \nabla f(\bar{x})\tilde{v}+\frac{1}{2}u^T\nabla f(\bar{x})u+\lambda^T(\nabla F(\bar{x})\tilde{v}+\frac{1}{2}u^T\nabla^2F(\bar{x})u)\\ &= \nabla_x\mathscr{L}(\bar{x},1,\lambda)\tilde{v}+\frac{1}{2}u^T\nabla_x^2\mathscr{L}(\bar{x},1,\lambda)u=\frac{1}{2}u^T\nabla_x^2\mathscr{L}(\bar{x},1,\lambda)u \end{aligned}$$

contradicting (27). □

Remark 1. *Following [14, Definition 3.2] the point $\bar{x}$ is said to fulfill the* strong second-order sufficient condition *(SSOSC) for* (1) *if $\Lambda^1(\bar{x})\neq\emptyset$ and for every nonzero critical direction $0\neq u\in\mathscr{C}(\bar{x})$ one has*

$$u^T\nabla_x^2\mathscr{L}(\bar{x},1,\lambda)u>0\;\forall\lambda\in\Lambda^1(\bar{x}).$$

However note that this condition is not sufficient for $\bar{x}$ to be a local minimizer as can be easily seen from the example

$$\min -x_1+x_1^2+x_2^2 \text{ subject to } (-x_1,-x_2)\in Q_{\text{EC}}.$$

In order to make (SSOSC) sufficient for $\bar{x}$ being a local minimizer, in view of Theorem 8 we have to replace the M-stationarity condition $\Lambda^1(\bar{x})\neq\emptyset$ by the extended M-stationarity condition $\Lambda^1(\bar{x};u)\neq\emptyset\;\forall 0\neq u\in\mathscr{C}(\bar{x})$.

4 Applications to MPECs

We now want to apply the results of the preceding section to the MPEC (2), or more exactly, to the problem (1) with F and Ω given by (5). By straightforward calculation we can obtain the formulas for the Fréchet normal cone, the Mordukhovich normal cone and the contingent cone of the set Q_{EC} defined in (4) as follows:

Lemma 8. *For all $a=(a_1,a_2)\in Q_{\text{EC}}$ we have*

$$\hat{N}(a;\Omega_{\text{EC}})=\left\{(\xi_1,\xi_2)\;\middle|\;\begin{array}{ll}\xi_2=0 & \text{if } 0=a_1>a_2\\ \xi_1\geq 0,\xi_2\geq 0 & \text{if } a_1=a_2=0\\ \xi_1=0 & \text{if } a_1<a_2=0\end{array}\right\},$$

$$N(a;\Omega_{\text{EC}})=\begin{cases}\hat{N}(a;\Omega_{\text{EC}}) & \text{if } a\neq(0,0)\\ \{(\xi_1,\xi_2)\,|\,\text{either } \xi_1>0,\ \xi_2>0 \text{ or } \xi_1\xi_2=0\} & \text{if } a=(0,0),\end{cases}$$

$$T(a;\Omega_{\text{EC}}) = \left\{ (u_1,u_2) \mid \begin{array}{ll} u_1 = 0 & \textit{if } 0 = a_1 > a_2 \\ u_1 \le 0, u_2 \le 0, u_1u_2 = 0 & \textit{if } a_1 = a_2 = 0 \\ u_2 = 0 & \textit{if } a_1 < a_2 = 0 \end{array} \right\}$$

and for all $u = (u_1,u_2) \in T(a;\Omega_{\text{EC}})$ *we have*

$$T(u;T(a;\Omega_{\text{EC}})) = \begin{cases} T(a;\Omega_{\text{EC}}) & \textit{if } a \neq (0,0) \\ T(u;\Omega_{\text{EC}})) & \textit{if } a = (0,0), \end{cases}$$

$$\hat{N}(u;T(a;\Omega_{\text{EC}})) = \begin{cases} \hat{N}(a;\Omega_{\text{EC}}) & \textit{if } a \neq (0,0) \\ \hat{N}(u;\Omega_{\text{EC}}) & \textit{if } a = (0,0), \end{cases}$$

$$N(a;\Omega_{\text{EC}};u) = \begin{cases} N(a;\Omega_{\text{EC}}) & \textit{if } a \neq (0,0) \\ N(u;\Omega_{\text{EC}}) & \textit{if } a = (0,0). \end{cases}$$

In what follows, we denote by $\bar{x}$ a point feasible for the MPEC (2). Further we assume throughout this section that the mappings f, g, h, G, H are continuously Fréchet differentiable, twice Fréchet differentiable at $\bar{x}$ and that there are numbers $1 \le l_1 \le l$, $1 \le p_1 \le p$, $1 \le q_1 \le q$ such that the components

$$g_i(x), i = l_1+1,\ldots,l, \quad h_i(x), i = p_1+1,\ldots,p, \quad G_i(x), H_i(x), i = q_1+1\ldots,q$$

are affine linear. In what follows, for every direction $u \in T_{\text{lin}}(\bar{x})$ the multifunction M_2 which is assumed to be metrically subregular in direction u is build by the linear parts of the constraints.

Denoting

$$\begin{aligned} \bar{I}_g &:= \{i \in \{1,\ldots,l\} \mid g_i(\bar{x}) = 0\}, \\ \bar{I}^{+0} &:= \{i \in \{1,\ldots,q\} \mid G_i(\bar{x}) > 0 = H_i(\bar{x})\}, \\ \bar{I}^{0+} &:= \{i \in \{1,\ldots,q\} \mid G_i(\bar{x}) = 0 < H_i(\bar{x})\}, \\ \bar{I}^{00} &:= \{i \in \{1,\ldots,q\} \mid G_i(\bar{x})0 = H_i(\bar{x})\}, \end{aligned}$$

the cone $T_{\text{lin}}(\bar{x})$ is given by

$$T_{\text{lin}}(\bar{x}) = \left\{ u \in \mathbb{R}^n \mid \begin{array}{l} \nabla g_i(\bar{x})u \le 0,\ i \in \bar{I}_g, \\ \nabla h_i(\bar{x})u = 0,\ i = 1,\ldots,p, \\ \nabla G_i(\bar{x})u = 0,\ i \in \bar{I}^{0+}, \\ \nabla H_i(\bar{x})u = 0,\ i \in \bar{I}^{+0}, \\ -(\nabla G_i(\bar{x})u, \nabla H_i(\bar{x})u) \in Q_{\text{EC}},\ i \in \bar{I}^{00} \end{array} \right\}.$$

The generalized Lagrangian reads as

$$\mathscr{L}(x,\lambda_0,\lambda) = \lambda_0 f(x) + \lambda^{g^T} g(x) + \lambda^{h^T} h(x) - \lambda^{G^T} G(x) - \lambda^{H^T} H(x),$$

where $\lambda_0 \in \mathbb{R}$, $\lambda := (\lambda^g, \lambda^h, \lambda^G, \lambda^H) \in \mathbb{R}^l \times \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R}^q$.

Given $u \in T_{\text{lin}}(\bar{x})$ we define

$$\begin{aligned}
I_g(u) &:= \{i \in \bar{I}_g \,|\, \nabla g_i(\bar{x})u = 0\} \\
I^{+0}(u) &:= \{i \in \bar{I}^{00} \,|\, \nabla G_i(\bar{x})u > 0 = \nabla H_i(\bar{x})u\}, \\
I^{0+}(u) &:= \{i \in \bar{I}^{00} \,|\, \nabla G_i(\bar{x})u = 0 < \nabla H_i(\bar{x})u\}, \\
I^{00}(u) &:= \{i \in \bar{I}^{00} \,|\, \nabla G_i(\bar{x})u = 0 = \nabla H_i(\bar{x})u\}.
\end{aligned}$$

Then for $\lambda_0 \geq 0$ we have

$$\Lambda^{\lambda_0}(\bar{x};u) = \left\{ \lambda = (\lambda^g, \lambda^h, \lambda^G, \lambda^H) \,\middle|\, \begin{array}{l} \nabla_x \mathscr{L}(\bar{x}, \lambda_0, \lambda) = 0 \\ \lambda_i^g \geq 0, \lambda_i^g g_i(x) = 0,\ i \in \{1, \dots, l\} \\ \lambda_i^g = 0,\ i \in \bar{I}_g \setminus I_g(u) \\ \lambda_i^H = 0,\ i \in \bar{I}^{0+} \cup I^{0+}(u) \\ \lambda_i^G = 0,\ i \in \bar{I}^{+0} \cup I^{+0}(u) \\ \text{either } \lambda_i^G > 0, \lambda_i^H > 0 \text{ or } \lambda_i^G \lambda_i^H = 0,\ i \in I^{00}(u) \\ \lambda_0 + \sum_{i=l}^{l_1} \lambda_i^g + \sum_{i=1}^{p_1} |\lambda_i^h| + \sum_{i=1}^{q_1} (|\lambda_i^G| + |\lambda_i^H|) > 0 \end{array} \right\}$$

and

$$\hat{\Lambda}^{\lambda_0}(\bar{x};u) = \left\{ \lambda = (\lambda^g, \lambda^h, \lambda^G, \lambda^H) \,\middle|\, \begin{array}{l} \nabla_x \mathscr{L}(\bar{x}, \lambda_0, \lambda) = 0 \\ \lambda_i^g \geq 0, \lambda_i^g g_i(x) = 0,\ i \in \{1, \dots, l\} \\ \lambda_i^g = 0,\ i \in \bar{I}_g \setminus I_g(u) \\ \lambda_i^H = 0,\ i \in \bar{I}^{0+} \cup I^{0+}(u) \\ \lambda_i^G = 0,\ i \in \bar{I}^{+0} \cup I^{+0}(u) \\ \lambda_i^G \geq 0, \lambda_i^H \geq 0,\ i \in I^{00}(u) \\ \lambda_0 + \sum_{i=l}^{l_1} \lambda_i^g + \sum_{i=1}^{p_1} |\lambda_i^h| + \sum_{i=1}^{q_1} (|\lambda_i^G| + |\lambda_i^H|) > 0 \end{array} \right\}.$$

Further it follows that LICQ(u) is fulfilled if and only if the family of gradients

$$\begin{aligned}
&\{\nabla g_i(\bar{x}) \,|\, i \in I_g(u)\} \cup \{\nabla h_i(\bar{x}) \,|\, i \in \{1, \dots, p\}\} \cup \{\nabla G_i(\bar{x}) \,|\, i \in \bar{I}^{0+} \cup I^{0+}(u) \cup I^{00}(u)\} \\
&\cup \{\nabla H_i(\bar{x}) \,|\, i \in \bar{I}^{+0} \cup I^{+0}(u) \cup I^{00}(u)\}
\end{aligned}$$

is linearly independent. It is easy to see that LICQ(0) is exactly the well-known MPEC LICQ condition.

Example 3. *Consider the problem*

$$\begin{aligned}
\min_{x=(x_1,x_2,x_3)} f(x) &:= x_1 + x_2 - 2x_3 \\
g_1(x) &:= -x_1 - x_3 \leq 0 \\
g_2(x) &:= -x_2 + x_3 \leq 0 \\
-(G_1(x), H_1(x)) &:= -(x_1, x_2) \in Q_{\text{EC}}
\end{aligned}$$

Then $\bar{x} = (0,0,0)$ is not a local minimizer, because for every $\alpha > 0$ the point $x^\alpha := (0, \alpha, \alpha)$ is feasible and $f(x^\alpha) = -\alpha < 0$. Indeed, for the critical direction $u = (0,1,1)$ we have $\Lambda^1(\bar{x};u) = \emptyset$ and therefore $\bar{x}$ is not an extended M-stationary solution and consequently not a local minimizer, since all problem functions are linear and the constraint mapping is thus metrically subregular. However, $\bar{x}$ is M-stationary since $\Lambda^1(\bar{x}) = \{(1,3,0,-2)\}$.

To demonstrate the results on second-order optimality conditions of the preceding section we consider the following example:

Example 4. *Consider the parameter dependent problem*

$$\begin{aligned} P(a) \qquad \min_{x_1,x_2} f(x_1,x_2) &:= -x_1+\frac{1}{2}x_2^2\\ g_1(x_1,x_2) &:= ax_1^2-x_2\le 0,\\ -(G_1(x_1,x_2),H_1(x_1,x_2)) &:= -(x_1,x_2)\in Q_{\rm EC}\end{aligned}$$

where $a\in\mathbb{R}$. Then it is easy to see that $\bar x=(0,0)$ is a local minimizer, if and only if $a>0$. Let us verify this by using our theory.

We have

$$\mathscr{L}(x,\lambda_0,(\lambda^g,\lambda^G,\lambda^H))=\lambda_0(-x_1+\frac{1}{2}x_2^2)+\lambda^g(ax_1^2-x_2)-\lambda^G x_1-\lambda^H x_2$$

and for every a it follows that $\Lambda^0(\bar x)=\{(\alpha,0,-\alpha\,|\,\alpha>0\}$, implying that metric regularity of the constraint mapping and therefore also LICQ(0) are violated. Further we have

$$T_{\rm lin}(\bar x)=\{(u_1,u_2)\,|\,-u_2\le 0,\ (-u_1,-u_2)\in Q_{\rm EC}\}=-Q_{\rm EC}$$

and $\mathscr{C}(\bar x)=T_{\rm lin}(\bar x)$, i.e. we have to analyze the problem with respect to the two critical directions $(1,0)$ *and* $(0,1)$.

1. $u=(1,0)$*: Then $\Lambda^1(\bar x;u)=\emptyset$ and $\Lambda^0(\bar x;u)=\hat\Lambda^0(\bar x;u)=\Lambda^0(\bar x)$ and taking $\lambda=(\alpha,0,-\alpha)$ with $\alpha>0$ we have*

$$u^T\nabla_x^2\mathscr{L}(\bar x,0,\lambda)u=2\alpha a\begin{cases}<0 & \text{if } a<0,\\ >0 & \text{if } a>0\end{cases}. \tag{29}$$

   By the second-order conditions (17) *we conclude that $\bar x$ is not a local minimizer for $a<0$. In case $a=0$ the constraint mapping is polyhedral and hence metrically subregular. Since $\Lambda^1(\bar x,u)=\emptyset$ we can also conclude from Theorem 4 that $\bar x$ is not a local minimizer in case $a=0$.*

2. $u=(0,1)$*: In this case LICQ(u) is fulfilled and we have $\hat\Lambda^1(\bar x;u)=\Lambda^1(\bar x;u)=\{(0,-1,0)\}$. Since $u^T\nabla_x^2\mathscr{L}(\bar x,1,(0,-1,0))u=1>0$, together with* (29)*, we conclude from Theorem 7 that $\bar x$ is a essential local minimizer of second order in case $a>0$.*

Since extended M-stationarity is usually difficult to verify in practice, we now introduce the following concept of *strong M-stationarity*, which builds a bridge between M-stationarity and S-stationarity. In what follows we note by $r(\bar x)$ the rank of the family of gradients

$$\{\nabla g_i(\bar x)\,|\,i\in\bar I_g\}\cup\{\nabla h_i(\bar x)\,|\,i\in\{1,\ldots,p\}\}\cup\{\nabla G_i(\bar x)\,|\,i\in\bar I^{0+}\cup\bar I^{00}\}\cup\{\nabla H_i(\bar x)\,|\,i\in\bar I^{+0}\cup\bar I^{00}\}. \tag{30}$$

Definition 6. *1. A triple of index sets (J_g, J_G, J_H), $J_g \subset \bar{I}_g$, $J_G \subset \bar{I}^{0+} \cup \bar{I}^{00}$, $J_H \subset \bar{I}^{+0} \cup \bar{I}^{00}$ is called a* MPEC working set *for the MPEC* (2)*, if $J_G \cup J_H = \{1,\dots,q\}$,*

$$|J_g| + p + |J_G| + |J_H| = r(\bar{x})$$

and the family of gradients

$$\{\nabla g_i(\bar{x}) \,|\, i \in J_g\} \cup \{\nabla h_i(\bar{x}) \,|\, i \in \{1,\dots,p\}\} \cup \{\nabla G_i(\bar{x}) \,|\, i \in J_G\} \cup \{\nabla H_i(\bar{x}) \,|\, i \in J_H\}$$

is linearly independent.

2. The point $\bar{x}$ is called strongly M-stationary *for the MPEC* (2)*, if there is a MPEC working set (J_g, J_G, J_H) together with a multplier $\lambda = (\lambda^g, \lambda^h, \lambda^G, \lambda^H) \in \Lambda^1(\bar{x})$ satisfying*

$$\lambda_i^g = 0,\ i \in \{1,\dots,l\} \setminus J_g, \tag{31}$$
$$\lambda_i^G = 0,\ i \in \{1,\dots,q\} \setminus J_G, \tag{32}$$
$$\lambda_i^H = 0,\ i \in \{1,\dots,q\} \setminus J_H, \tag{33}$$
$$\lambda_i^G \geq 0, \lambda_i^H \geq 0,\ i \in J_G \cap J_H. \tag{34}$$

Note that the condition $J_G \cup J_H = \{1,\dots,q\}$ implies $\bar{I}^{0+} \subset J_G$ and $\bar{I}^{+0} \subset J_H$.

Theorem 9. *Assume that $\bar{x}$ is extended M-stationary for the problem* (1) *with F and Ω given by* (5) *and assume that there exists some MPEC working set. Then $\bar{x}$ is strongly M-stationary.*

Proof. Since $\bar{x}$ is extended M-stationary, we have $\Lambda^1(\bar{x}) \neq \emptyset$ and therefore $\nabla f(\bar{x})$ can be represented as a linear combination of the gradients (30). It follows that for every MPEC working set $J = (J_g, J_G, J_H)$ there is a unique multiplier $\lambda(J) = (\lambda^g, \lambda^h, \lambda^G, \lambda^H)$ satisfying (31)-(33) and $\nabla_x \mathscr{L}(\bar{x}, 1, \lambda(J)) = 0$. Now let $J^0 = (J_g^0, J_G^0, J_H^0)$ be an arbitrarily fixed working set and choose $b = (b^g, b^G, b^H) \in \mathbb{R}^l_+ \times \mathbb{R}^q_- \times \mathbb{R}^q_-$ with $b_i^g = 0$, $i \in J_g^0$, $b_i^G = 0$, $i \in J_G^0$ and $b_i^H = 0$, $i \in J_H^0$ such that for all $u \in \mathbb{R}^n$ the family of gradients

$$\begin{aligned} &\{\nabla g_i(\bar{x}) \,|\, i \in \bar{I}_g, \nabla g_i(\bar{x})u = b_i^g\} \cup \{\nabla h_i(\bar{x}) \,|\, i \in \{1,\dots,p\}, \nabla h_i(\bar{x})u = 0\} \\ &\cup \{\nabla G_i(\bar{x}) \,|\, i \in \bar{I}^{0+} \cup \bar{I}^{00}, \nabla G_i(\bar{x})u = b_i^G\} \cup \{\nabla H_i(\bar{x}) \,|\, i \in \bar{I}^{+0} \cup \bar{I}^{00}, \nabla H_i(\bar{x})u = b_i^H\} \end{aligned} \tag{35}$$

is linearly independent. Such a vector b exists by the following arguments. For every triple of index sets $K = (K_g, K_G, K_H)$, $K_g \subset \bar{I}^g$, $K_G \subset \bar{I}^{0+} \cup \bar{I}^{00}$, $K_H \subset \bar{I}^{+0} \cup \bar{I}^{00}$ let $\mathscr{B}(K)$ denote a basis for the subspace

$$\left\{\mu = \{(\mu^g, \mu^h, \mu^G, \mu^H) \,\middle|\, \begin{array}{l} \nabla_x \mathscr{L}(\bar{x}, 0, \mu) = 0, \\ \mu_i^g = 0,\ i \in \{1,\dots,l\} \setminus K_g, \\ \mu_i^G = 0,\ i \in \{1,\dots,q\} \setminus K_G, \\ \mu_i^H = 0,\ i \in \{1,\dots,q\} \setminus K_H \end{array} \right\},$$

where $\mathscr{B}(K)$ is eventually empty. By the definition of a MPEC working set, for every basis element $\mu = (\mu^g, \mu^h, \mu^G, \mu^H)$ there must be either an index $i \in \bar{I}_g \setminus J_g^0$ with $\mu_i^g \neq 0$ or an index

$i \in (\bar{I}^{0+} \cup \bar{I}^{00}) \setminus J_G^0$ with $\mu_i^G \neq 0$ or an index $i \in (\bar{I}^{+0} \cup \bar{I}^{00}) \setminus J_H^0$ with $\mu_i^H \neq 0$. The union of the bases $\bigcup_K \mathscr{B}(K)$ consists of finitely many elements and therefore we can find $b = (b^g, b^G, b^H) \in \mathbb{R}_+^l \times \mathbb{R}_-^q \times \mathbb{R}_-^q$ with $b_i^g = 0$, $i \in J_g^0$, $b_i^G = 0$, $i \in J_G^0$ and $b_i^H = 0$, $i \in J_H^0$ with

$$b^{g^T} \mu^g + b^{G^T} \mu^G + b^{H^T} \mu^H \neq 0 \; \forall (\mu^g, \mu^h, \mu^G, \mu^H) \in \bigcup_K \mathscr{B}(K)$$

We claim that this vector b has the required property. If there would exist $u \in \mathbb{R}^n$ such that (35) does not hold, by taking $K_g := \{i \in \bar{I}_g \,|\, \nabla g_i(\bar{x})u = b_i^g\}$, $K_G := \{i \in \bar{I}^{0+} \cup \bar{I}^{00} \,|\, \nabla G_i(\bar{x})u = b_i^G\}$, $K_H := \{i \in \bar{I}^{+0} \cup \bar{I}^{00} \,|\, \nabla H_i(\bar{x})u = b_i^H\}$, there is some element $\mu = (\mu^g, \mu^h, \mu^G, \mu^H) \in \mathscr{B}(K_g, K_G, K_H)$ with

$$0 = \nabla_x \mathscr{L}(\bar{x}, 0, \mu)u = b^{g^T} \mu^g + b^{G^T} \mu^G + b^{H^T} \mu^H \neq 0,$$

a contradiction, and therefore our claim is proved. Now consider the following algorithm:

$u := 0$, $J := J^0$, $(\lambda^g, \lambda^h, \lambda^G, \lambda^H) := \lambda(J)$;
`while` $((\exists i \in J_g : \lambda_i^g < 0) \vee (\exists i \in J_G \cap J_H : \lambda_i^G < 0 \vee \lambda_i^H < 0))$
$\{$ `if` $(\exists i_0 \in J_g : \lambda_{i_0}^g < 0)$
$J_g := J_g \setminus \{i_0\}$;
`else`
$\{$ select $i_0 \in J_G \cap J_H$ with $\lambda_{i_0}^G < 0$ or $\lambda_{i_0}^H < 0$;
`if` $(\lambda_{i_0}^G < 0)$
$J_G := J_G \setminus \{i_0\}$;
`else`
$J_H := J_H \setminus \{i_0\}$;
$\}$
Compute search direction d with $\nabla f(\bar{x})d = -1$, $\nabla g_i(\bar{x})d = 0, i \in J_g$, $\nabla h_i(\bar{x})d = 0, i = 1, \ldots, p$, $\nabla G_i(\bar{x})d = 0, i \in J_G$, $\nabla H_i(\bar{x})d = 0, i \in J_H$;
Compute step length

$$\hat{\alpha}_j = \min\{ \min_{\substack{i \in \bar{I}_g \setminus J_g \\ \nabla g_i(\bar{x})d > 0}} \{\frac{b_i^g - \nabla g_i(\bar{x})u}{\nabla g_i(\bar{x})d}\}, \min_{\substack{i \in \bar{I}^{00} \setminus J_G \\ \nabla G_i(\bar{x})d < 0}} \{\frac{b_i^G - \nabla G_i(\bar{x})u}{\nabla G_i(\bar{x})d}\}, \min_{\substack{i \in \bar{I}^{00} \setminus J_H \\ \nabla H_i(\bar{x})d < 0}} \{\frac{b_i^H - \nabla H_i(\bar{x})u}{\nabla H_i(\bar{x})d}\}\};$$

//The index j indicates the constraint to enter the MPEC working set
Either set $J := J_g \cup \{j\}$ or $J_G := J_G \cup \{j\}$ or $J_H := J_H \cup \{j\}$, depending in which part the minimum is attained when computing $\hat{\alpha}_j$;
$u := u + \hat{\alpha}_j d$, compute $(\lambda^g, \lambda^h, \lambda^G, \lambda^H) := \lambda(J)$;
$\}$

This algorithm is very close to the well-known pivoting algorithms from linear programming.

At the beginning of each cycle (J_g, J_G, J_H) constitutes a MPEC working set and we have

$$\begin{aligned}
&\nabla g_i(\bar{x})u \le b_i^g,\ i \in \bar{I},\\
&\nabla h_i(\bar{x})u = 0,\ i = 1,\dots,p,\\
&\nabla G_i(\bar{x})u = b_i^G = 0,\ i \in \bar{I}^{0+},\\
&\nabla G_i(\bar{x})u \ge b_i^G,\ i \in \bar{I}^{00}\\
&\nabla H_i(\bar{x})u = b_i^H = 0,\ i \in \bar{I}^{+0},\\
&\nabla H_i(\bar{x})u \ge b_i^H,\ i \in \bar{I}^{00}\\
&(\nabla G_i(\bar{x})u - b_i^G)(\nabla H_i(\bar{x})u - b_i^H) = 0,\ i = 1,\dots,q
\end{aligned}$$

and

$$J_g = \{i \in \bar{I}_g \,|\, \nabla g(\bar{x})u = b_i^g\},\ J_G = \{i \in \{1,\dots q\} \,|\, \nabla G_i(\bar{x})u = b_i^H\},\ J_H = \{i \in \{1,\dots q\} \,|\, \nabla H_i(\bar{x})u = b_i^H\}.$$

The computation of the search direction d in line 12 is possible, because after removing index i_0 from the MPEC working set the family of gradients

$$\{\nabla f(\bar{x})\} \cup \{\nabla g_i(\bar{x}) \,|\, i \in J_g\} \cup \{\nabla h_i(\bar{x}) \,|\, i = 1,\dots,p\} \cup \{\nabla G_i(\bar{x}) \,|\, i \in J_G\} \cup \{\nabla H_i(\bar{x}) \,|\, i \in J_H\}$$

is linearly independent. The minimum when computing $\hat{\alpha}_j$ must be attained, because otherwise the direction d would fulfill $d \in T_{\text{lin}}(\bar{x})$ and $\nabla f(\bar{x})d < 0$ contradicting extended M-stationarity of $\bar{x}$. Further, our construction of b guarantees that the index j is unique and $\hat{\alpha}_j$ is strictly positive.

Since the value $\nabla f(\bar{x})u$ strictly decreases in each cycle and only a finite number of MPEC working sets exist, the algorithm always terminates in a finite number of steps and the outcome $J = (J_g, J_G, J_H)$ together with $\lambda(J)$ proves strong M-stationarity of $\bar{x}$. □

The proof of this theorem is constructive because the algorithm can be implemented in practice. A random choice of b with

$$b_i^g > 0, i \in \bar{I} \setminus J_g^0,\ b_i^G < 0,\ i \in \bar{I}^{00} \setminus J_G^0,\ b_i^H < 0,\ i \in \bar{I}^{00} \setminus J_H^0$$

and fixing the other components to 0 will yield a suitable vector b with probability 1, as can be easily seen from the arguments used in the proof. Moreover, the unlikely case of a wrong choice of b can be easily detected during the course of the algorithm and then we can modify b to meet the requirements. Of course, one has to implement an exit in case that $\hat{\alpha}_j = \infty$, i.e. $\{i \in \bar{I}_g \setminus J_g \,|\, \nabla g_i(\bar{x})d > 0\} = \{i \in \bar{I}^{00} \setminus J_G \,|\, \nabla G_i(\bar{x})d < 0\} = \{i \in \bar{I}^{00} \setminus J_H \,|\, \nabla H_i(\bar{x})d < 0\} = \emptyset$, since then the computed direction d is a descent direction.

The assumption, that one MPEC working set exists, is fulfilled, if there are index sets $\tilde{J}_G, \tilde{J}_H \subset \bar{I}^{00}$ with $\tilde{J}_G \cup \tilde{J}_H = \bar{I}^{00}$ such that the family of gradients

$$\{\nabla h_i(\bar{x}) \,|\, i = 1,\dots,p\} \cup \{\nabla G_i(\bar{x}) \,|\, i \in I^{0+} \cup \tilde{J}_G\} \cup \{\nabla H_i(\bar{x}) \,|\, i \in I^{+0} \cup \tilde{J}_H\}$$

is linearly independent. This is a rather weak assumption, it is e.g. fulfilled if for one direction $u \in T_{\text{lin}}(\bar{x})$ the first-order condition for directional metric subregularity $\Lambda^0(\bar{x}; u) = \emptyset$ is fulfilled.

The following theorem justifies the definition of strongly M-stationary solutions.

Theorem 10. *Let $\bar{x}$ be feasible for* (2) *and assume that LICQ(0) is fulfilled at $\bar{x}$. Then $\bar{x}$ is strongly M-stationary if and only if it is S-stationary.*

Proof. The statement follows immediately from the fact that under LICQ(0) there exist exactly one MPEC working set and this set fulfills $J_g = \bar{I}_g$, $J_G = \bar{I}^{0+} \cup \bar{I}^{00}$, $J_H = \bar{I}^{+0} \cup \bar{I}^{00}$. □

Using similar arguments it can also be shown that under the weaker condition *partial MPEC LICQ* [46] the concepts of strongly M-stationarity and S-stationarity are equivalent. However, for other conditions ensuring S-stationarity like the *intersection property* [7] or the condition found in [8], the relation between strong M-stationarity and S-stationarity is still unknown.

References

[1] W. ACHTZIGER, T. HOHEISEL, C. KANZOW, *A smoothing-regularization approach to mathematical programs with vanishing constraints*, Comput. Optim. Appl., to appear.

[2] W. ACHTZIGER, C. KANZOW, *Mathematical programs with vanishing constraints: Optimality conditions and constraint qualifications*, Math. Program., 114 (2008), pp. 69-99.

[3] J. F. BONNANS, A. SHAPIRO, *Perturbation analysis of optimization problems*, Springer, New York, 2000.

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

[5] M. J. FABIAN, R. HENRION, A. Y. KRUGER, J. V. OUTRATA, *Error bounds: necessary and sufficient conditions*, Set-Valued Anal.18 (2010), pp. 121-149.

[6] M. L. FLEGEL, C. KANZOW, *A Fritz John approach to first order optimality conditions for mathematical programs with equilibrium constraints*, Optimization, 52 (2003), pp. 277-286.

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

[8] M. FUKUSHIMA, J. S. PANG, *Complementarity constraint qualifications and simplified B-stationary conditions for mathematical programs with equilibrium constraints*, Comput. Optim. Appl., 13 (1999), pp. 111-136.

[9] H. GFRERER, *Second-order optimality conditions for scalar and vector optimization problems in Banach spaces*, SIAM J. Control Optim., 45 (2006), pp. 972–997.

[10] 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.

[11] H. GFRERER, *On directional metric regularity, subregularity and optimality conditions for nonsmooth mathematical programs*, (2012), Set-Valued Var. Anal., to appear.

[12] H. GFRERER, *On directional metric subregularity and second-order optimality conditions for a class of nonsmooth mathematical programs*, SIAM J. Optim., to appear.

[13] H. GFRERER, *On metric pseudo-(sub)regularity of multifunctions and optimality conditions for degenerated mathematical programs* , preprint, Univ.Linz (2012), submitted.

[14] L. GUO, G. LIN, J. J. YE, *Stability analysis for parametric mathematical programs with geometric constraints and its applications*, SIAM J. Optim., 22 (2012), pp. 11511176.

[15] L. GUO, G. LIN, J. J. YE, *Second-Order Optimality Conditions for Mathematical Programs with Equilibrium Constraints*, J. Optim. Theory Appl., to appear.

[16] R. HENRION, A. JOURANI, *Subdifferential conditions for calmness of convex constraints*, SIAM J. Optim., 13(2002), pp. 520–534.

[17] R. HENRION, A. JOURANI, J. V. OUTRATA, *On the calmness of a class of multifunctions*, SIAM J. Optim., 13 (2002), pp. 603–618.

[18] R. HENRION, B. S. MORDUKHOVICH, N. M. NAM, *Second-order analysis of polyhedral systems in finite and infinite dimensions with applications to robust stability of variational inequalities*, SIAM J. Optim., 20 (2010), pp. 2199–2237.

[19] R. HENRION, J. V. OUTRATA, *A subdifferential condition for calmness of multifunctions*, J. Math. Anal. Appl., 258 (2001), pp. 110–130.

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

[21] A. HOFFMAN, *On approximate solutions of systems of linear inequalities*, Journal of Research of the National Bureau of Standards, 49 (1952), pp. 263–265.

[22] T. HOHEISEL, C. KANZOW *Stationary conditions for mathematical programs with vanishing constraints using weak constraint qualifications*, J. Math. Anal. Appl. 337 (2008), pp. 292-310.

[23] T. HOHEISEL, C. KANZOW, *On the Abadie and Guignard constraint qualification for mathematical progams with vanishing constraints*, Optimization 58 (2009), pp. 431-448.

[24] T. HOHEISEL, C. KANZOW, J. V. OUTRATA, *Exact penalty results for mathematical programs with vanishing constraints*, Nonlinear Anal., 72 (2010), pp. 2514-2526.

[25] A. D. IOFFE, *Necessary and sufficient conditions for a local minimum 1: A reduction theorem and first order conditions*, SIAM J. Control Optim., 17 (1979), pp. 245–250.

[26] A. D. IOFFE, *Metric regularity and subdifferential calculus*, Russian Math. Surveys, 55 (2000), pp. 501–558.

[27] A. D. IOFFE, V. OUTRATA, *On metric and calmness qualification conditions in subdifferential calculus*, Set-Valued Anal.16 (2008), pp. 199–227.

[28] A. F. IZMAILOV, M. V. SOLODOV, *Mathematical programs with vanishing constraints: optimality conditions, sensitivity and a relaxation method*, J. Optim. Theory Appl., 142 (2009), pp. 501-532.

[29] C. KANZOW, A. SCHWARTZ, *Mathematical programs with equilibrium constraints: enhanced Fritz John conditions, new constraint qualifications and improved exact penalty results*, SIAM J. Optim. 20 (2010), pp. 2730-2753.

[30] D. KLATTE, B. KUMMER, *Nonsmooth equations in optimization. Regularity, calculus, methods and applications*, Nonconvex Optimization and its Applications 60, Kluwer Academic Publishers, Dordrecht, Boston, London, 2002.

[31] B. KUMMER, *Inclusions in general spaces: Hoelder stability, solution schemes and Ekeland's principle*, J. Math. Anal. Appl. 358 (2009), pp. 327–344.

[32] Z. Q. LUO, J. S. PANG, D. RALPH, *Mathematical Programs with Equilibrium Constraints*, Cambridge University Press, Cambridge, New York, Melbourne, 1996.

[33] B. S. MORDUKHOVICH, *Variational analysis and generalized differentiation I: Basic theory*, Springer, Berlin, Heidelberg, 2006.

[34] B. S. MORDUKHOVICH, *Variational analysis and generalized differentiation II: Applications*, Springer, Berlin, Heidelberg, 2006.

[35] J. V. OUTRATA, *Optimality conditions for a class of mathematical programs with equilibrium constraints*, Math. Oper. Res., 24 (1999), pp. 627-644.

[36] J. V. OUTRATA, *A generalized mathematical program with equilibrium constraints*, SIAM J. Control Optim., 38 (2000), pp. 1623-1638

[37] J. V. OUTRATA, M. KOČVARA, J. ZOWE, *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints*, Nonconvex Optimization and its Applications, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998.

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

[39] S. M. ROBINSON, *Stability theorems for systems of inequalities, Part II: differentiable nonlinear systems*, SIAM J. Numer. Anal., 13 (1976), pp. 497–513.

[40] S. M. ROBINSON, *Some continuity properties of polyhedral multifunctions*, Math. Program. Studies, 14 (1981), pp. 206–214.

[41] R. T. ROCKAFELLAR, *Convex analysis*, Princeton, New Jersey, 1970.

[42] R. T. ROCKAFELLAR, R. J-B. WETS, *Variational analysis*, Springer, Berlin, 1998.

[43] H. SCHEEL, S. SCHOLTES, *Mathematical programs with complementarity constraints: Stationarity, optimality, and sensitivity*, Math. Oper. Res., 25 (2000), pp. 1-22.

[44] J. J. YE, *Constraint qualifications and necessary optimality conditions for optimization problems with variational inequality constraints*, SIAM J. Optim., 10 (2000), pp. 943–962.

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

[46] J. J. YE,*Necessary and sufficient optimality conditions for mathematical programs with equilibrium constraints*, J. Math. Anal. Appl., 307 (2005), pp. 350-369.

[47] X. Y. ZHENG, K. F. NG, *Metric subregularity and constraint qualifications for convex generalized equations in Banach spaces*, SIAM J. Optim., 18 (2007), pp. 437–460.

[48] X. Y. ZHENG, K. F. NG, *Metric subregularity and calmness for nonconvex generalized equations in Banach spaces*, SIAM J. Optim. 20 (2010), pp. 2119-2136.

Latest Reports in this series

2009 - 2011

[..]

2012

[..]

2012-02	Michael Kolmbauer and Ulrich Langer *A Robust Preconditioned MinRes-Solver for Time-Periodic Eddy Current Problems*	January 2012
2012-03	Wolfgang Krendl, Valeria Simoncini and Walter Zulehner *Stability Estimates and Structural Spectral Properties of Saddle Point Problems*	February 2012
2012-04	Helmut Gfrerer *On Directional Metric Regularity, Subregularity and Optimality Conditions for Nonsmooth Mathematical Programs*	February 2012
2012-05	Veronika Pillwein and Stefan Takacs *A Multigrid Fourier Analysis of a Multigrid Method using Symbolic Computation*	April 2012
2012-06	Stefan Takacs and Walter Zulehner *Convergence Analysis of All-at-once Multigrid Methods for Elliptic Control Problems Under Partial Elliptic Regularity*	June 2012
2012-07	Helmut Gfrerer *On Directional Metric Subregularity and Second-Order Optimality Conditions for a Class of Nonsmooth Mathematical Programs*	August 2012
2012-08	Michael Kolmbauer and Ulrich Langer *Efficient Solvers for Some Classes of Time-Periodic Eddy Current Optimal Control Problems*	November 2012
2012-09	Clemens Hofreither, Ulrich Langer and Clemens Pechstein *FETI Solvers for Non-Standard Finite Element Equations Based on Boundary Integral Operators*	November 2012
2012-10	Helmut Gfrerer *On Metric Pseudo-(sub)Regularity of Multifunctions and Optimality Conditions for Degenerated Mathematical Programs*	December 2012
2012-11	Clemens Pechstein and Clemens Hofreither *A Rigorous Error Analysis of Coupled FEM-BEM Problems with Arbitrary Many Subdomains*	December 2012
2012-12	Markus Eslitzbichler, Clemens Pechstein and Ronny Ramlau *An H^1-Kaczmarz Reconstructor for Atmospheric Tomography*	December 2012
2012-13	Clemens Pechstein *On Iterative Substructuring Methods for Multiscale Problems*	December 2012

2013

2013-01	Ulrich Langer and Monika Wolfmayr *Multiharmonic Finite Element Analysis of a Time-Periodic Parabolic Optimal Control Problem*	January 2013
2013-02	Helmut Gfrerer *Optimality Conditions for Disjunctive Programs Based on Generalized Differentiation with Application to Mathematical Programs with Equilibrium Constraints*	March 2013

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/`