7-1-2006

Charactarizations of Linear Suboptimality for Mathematical Programs with Equilibrium Constraints

Boris S. Mordukhovich
Wayne State University, boris@math.wayne.edu

CHARACTERIZATIONS OF LINEAR SUBOPTIMALITY FOR MATHEMATICAL PROGRAMS WITH EQUILIBRIUM CONSTRAINTS

B. S. MORDUKHOVICH

WAYNE STATE UNIVERSITY

Detroit, MI 48202

Department of Mathematics
Research Report

2006 Series
#7

This research was partly supported by the National Science Foundation and the Australian Research Council

CHARACTERIZATIONS OF LINEAR SUBOPTIMALITY FOR MATHEMATICAL PROGRAMS WITH EQUILIBRIUM CONSTRAINTS[1]

B. S. MORDUKHOVICH [2]

Dedicated to Boris Polyak in honor of his 70th birthday

Abstract. The paper is devoted to the study of a new notion of linear suboptimality in constrained mathematical programming. This concept is different from conventional notions of solutions to optimization-related problems, while seems to be natural and significant from the viewpoint of modern variational analysis and applications. In contrast to standard notions, it admits complete characterizations via appropriate constructions of generalized differentiation in nonconvex settings. In this paper we mainly focus on various classes of mathematical programs with equilibrium constraints (MPECs), whose principal role has been well recognized in optimization theory and its applications. Based on robust generalized differential calculus, we derive new results giving pointwise necessary and sufficient conditions for linear suboptimality in general MPECs and its important specifications involving variational and quasivariational inequalities, implicit complementarity problems, etc.

Key words. nonsmooth optimization–variational analysis–generalized differentiation–mathematical programs with equilibrium constraints–linear suboptimality–necessary and sufficient conditions

Mathematics Subject Classification (2000): 90C30, 49J52, 49J53

1 Introduction

It is well known that, excepts convex programming and related problems with a convex structure, *necessary* conditions are usually *not sufficient* for conventional notions of optimality. Observe also that major necessary optimality conditions in all the branches of the classical and modern optimization theory (e.g., Lagrange multipliers and Karush-Kuhn-Tucker conditions in nonlinear programming, the Euler-Lagrange equation in the calculus of variations, the Pontryagin maximum principle in optimal control, etc.) are expressed in *dual* forms involving *adjoint* variables. At the same time, the very *notions of optimality*, in both scalar and vector frameworks, are formulated of course in *primal* terms.

Besides the standard definition of local minimizers in mathematical programming, there are other notions of minima important for optimization theory and applications. Let us first mention the concept of *sharp minima* introduced by Polyak [21] from certain numerical viewpoints and then applied (together with its "weak sharp" counterpart) by many researchers and practitioners to various aspects of optimization; see, e.g., the recent paper by Burke and Deng [4] with its references and discussions.

Another interesting modification of local minima was introduced, under the name of *tilt-stable local minima*, by Poliquin and Rockafellar [20] motivated by applications to sensitivity

[1]Research was partially supported by the National Science Foundation under grant DMS-0304989 and by the Australian Research Council under grant DP-0451168.

[2]Department of Mathematics, Wayne State University, Detroit, Michigan 48202, USA; boris@math.wayne.edu

analysis. The main result of [20] establishes a *second-order characterization* of tilt-stable minimizers for an extended-real-function φ on $I\!R^n$ via the positive definiteness of its second-order subdifferential (Hessian) matrix in the sense of [13]. For $\varphi \in C^2$, the latter condition reduces to the positive definiteness of the classical Hessian matrix $\nabla^2\varphi(\bar{x})$; see also the subsequent paper [12] for further developments and applications to nonsmooth optimization.

A challenging question is to find certain modified–weaker–notions of local optimality so that *first-order* necessary conditions known for the previously recognized notions become *necessary and sufficient* in the new framework. Such a study has been started by Kruger for *unconstrained* optimization problems (see [8, 9]), where the corresponding notions are called "extended optimality" and "weak stationarity." It seems that the main difference between the conventional notions and those studied in [8, 9] and then in the author's book [15, Chapter 5] is that the latter relate to a certain (sub)optimality not at the point in question but in a *neighborhood* of it, and that they involve a *linear rate* in the sense precisely defined below. To some extent, this is similar to the linear rate in *Lipschitz continuity* (in contrast merely to continuity) as well as in modern concepts of *metric regularity* and *linear openness*, which distinguishes them from the classical regularity and openness notions of nonlinear analysis. On this basis we suggested in [15] to use the name of "linear suboptimality" for general multiobjective optimization problems and of "linear subminimality" for constrained problems of minimization.

As has been fully recognized only in the framework of modern variational analysis (even regarding the classical settings), the *linear rate nature* of the fundamental properties involving Lipschitz continuity, metric regularity, and openness for single-valued and set-valued mappings is the *key issue* allowing us to derive *complete characterizations* of these properties via appropriate tools of generalized differentiation; see the books [14, 23] and their references. Precisely the same linear rate essence of (sub)optimality studied in this paper is the *driving force* ensuring the possibility to justify the validity of both well-known and recently discovered necessary optimality conditions for the conventional notions as *necessary and sufficient* conditions for the new notion of linear suboptimality and its modifications.

In contrast to [8, 9], where dual criteria for "extended optimality" and "weak stationarity" are obtained in "fuzzy" forms involving Fréchet-like constructions at points *nearby* the reference ones, in [15, Chapter 5] and in this paper we pay the main attention to *pointwise/pointbased* conditions for linear suboptimality expressed via the basic *robust* generalized differential constructions of [14, 15] *exactly at* the points in question. Besides the latter being more convenient for applications, the pointbased conditions allow us to employ the well-developed (full) *calculus* enjoyed by the robust constructions, which particularly gives us the possibility to study problems with various *constrained* structures important for both the optimization theory and its applications. A number of results in this direction were derived in [15, Section 5.4] for mathematical programs with conventional functional and geometric constraints as well as for certain problems of multiobjective optimization.

In this paper we study a new notion of (geometric) linear suboptimality/subminimality for problems of constrained minimization, which is induced by the concept of linear subextremality for set systems (see Section 3) and is *different* in general from the notion of linear subminimality studied in [15, Section 5.4] in the case of *constrained* minimization problems.

For convenience we use the name of *linear suboptimality* (vs. "linear subminimality" in [15]) for the new notion under consideration observing that both of these notions agree under certain *regularity* assumptions, which however are rather restrictive; cf. [14, 15].

Another important issue that distinguishes this paper from the corresponding developments of [15] is that now we pay the main attention to *mathematical programs with equilibrium constraints* (MPECs), which were not considered in [15, Section 5.4] even from the point of linear subminimality. The latter broad class of mathematical programs has been well recognized in both optimization theory and a variety applications. We refer the reader to the fundamental books [12, 19] and to subsequent numerous publications for many theoretical and computational developments in the area of MPECs and their significant applications to economics, mechanics, engineering, finance, etc. It is important to emphasize that the majority of MPECs are *intrinsically nonsmooth*, which makes them substantially different from conventional classes of problems in mathematical programming.

The focus and the principal contributions of this paper are to obtain *complete characterizations* of linear suboptimality for general MPECs and their several specifications. This is done in the paper by employing advanced tools of *variational analysis* and *generalized differentiation.* Observe that, from the viewpoint of deriving *necessary and sufficient* conditions for linear suboptimality, we need calculus rules of not merely the (right) *inclusion type* as required by the majority of applications (in particular, to necessary optimality conditions), but largely of the *equality type* that are available as well [14] for our basic generalized differential constructions way *beyond convexity.* Furthermore, in *infinite-dimensional* spaces one also needs calculus of the so-called *sequential normal compactness* (SNC) properties (automatic in finite dimensions), which is strongly developed in the book [14]. Based on these calculi (including that for the *second-order* subdifferentials), we obtain verifiable characterizations of linear suboptimality for the broad classes of MPECs under consideration.

The rest of the paper is organized as follows. In Section 2 we present basic preliminaries from variational analysis and generalized differentiation widely used in the sequel. In Section 3 we discuss the concept of *linear subextremality* for systems of sets and present its full characterization via the (exact) *extremal principle*, which are at the heart of the geometric approach to variational analysis and generalized differentiation; see [14, 15] and the discussions below.

Section 4 is devoted to the study of linear suboptimality in general minimization problems with equilibrium constraints. We first define this notion from both geometric and analytic viewpoints and apply it to the so-called "abstract MPECs" with equilibrium constraints described in the form

$$y \in S(x) \quad \text{with} \quad S\colon X \rightrightarrows Y, \tag{1.1}$$

where $S(\cdot)$ is a set-valued mapping from the space of parameters $x \in X$ to the space of decision variables $y \in Y$. In particular, $S(x)$ may model sets of optimal solutions (or stationary points, or KKT points) to *lower-level* problems of parametric mathematical programming, while may also arise from different sources: parametric complementarity problems, variational inequalities and their extensions as well as from other kinds of equilibrium conditions; see [12, 19] and the references therein. The main attention is paid in Section 4 to the case when the mapping $S(\cdot)$ in (1.1) is given as solution maps to parametric *generalized equations/variational conditions* of the type

$$0 \in f(x,y) + Q(x,y) \tag{1.2}$$

with a single-valued mapping f and a set-valued mapping Q. Model (1.2), which has been well recognized as a convenient framework for describing equilibrium constraints, was introduced by Robinson [22] for the case of normal cone mappings $Q(y) = N(y;\Omega)$ to convex sets, when (1.2) reduces to parametric variational inequalities.

Finally, in Section 5 we derive necessary and sufficient conditions for linear suboptimality in the most important specifications of MPECs when the set-valued part Q is given in one of the following *composite subdifferential forms*

$$Q(x,y) = \partial(\psi \circ g)(x,y), \quad Q(x,y) = (\partial\psi \circ g)(x,y), \tag{1.3}$$

where the first form in (1.3) involves the subdifferential (in the sense discussed in Section 2) of the composition of an extended-real-valued function $\psi\colon W \to \overline{I\!R} := (-\infty,\infty]$ and a single-valued mapping $g\colon X \times Y \to W$, while the second form therein is the composition of the subgradient set-valued mapping $\partial\psi\colon W \rightrightarrows W^*$ and a single-valued mapping $g\colon X \times Y \to W$. The composite forms in (1.3), which are clearly different from each other, cover the majority of equilibrium constraints in (1.2) that are the most interesting from the viewpoints of both optimization theory and applications.

The notation of this paper is basically standard; cf. [14, 23]. In particular, $I\!B$ stands for the unit closed ball of the space in question, while $B_r(x)$ signifies the ball centered at x with radius $r > 0$. As usual, $I\!N := \{1,2,\ldots\}$. Given a set-valued mapping $F\colon X \rightrightarrows X^*$ between a Banach space X and its topological dual X^*, we denote by

$$\operatorname*{Lim\,sup}_{x\to\bar{x}} F(x) := \Big\{x^* \in X^* \Big| \ \exists \text{ sequences } x_k \to \bar{x} \text{ and } x_k^* \xrightarrow{w^*} x^* \text{ with } x_k^* \in F(x_k) \text{ for all } k \in I\!N\Big\}$$

the *sequential* Painlevé-Kuratowski upper/outer limit of F as $x \to \bar{x}$ with respect to the norm topology of X and the weak* topology w^* of X^*.

2 Tools of Variational Analysis

We first recall the generalized differential constructions of variational analysis used in what follows; see the book [14] with the references and discussions therein and also [2, 15, 23] for some related and additional material.

Given a nonempty subset Ω of a Banach space X and a point $\bar{x} \in \Omega$, the (basic, limiting) *normal cone* to Ω at $\bar{x}$ is

$$N(\bar{x};\Omega) := \operatorname*{Lim\,sup}_{\substack{x \xrightarrow{\Omega} \bar{x} \\ \varepsilon \downarrow 0}} \widehat{N}_\varepsilon(x;\Omega), \tag{2.1}$$

where $x \xrightarrow{\Omega} \bar{x}$ means that $x \to \bar{x}$ with $x \in \Omega$, and where

$$\widehat{N}_\varepsilon(x;\Omega) := \Big\{x^* \in X^* \Big| \limsup_{u \xrightarrow{\Omega} x} \frac{\langle x^*, u - x\rangle}{\|u - x\|} \le \varepsilon\Big\} \tag{2.2}$$

is the set of ε-*normals* to Ω at $x \in \Omega$. When $\varepsilon = 0$ in (2.2), $\widehat{N}(x;\Omega) := \widehat{N}_0(x;\Omega)$ is a convex cone called the *prenormal cone* or the *Fréchet normal cone* to Ω at x. We can equivalently put $\varepsilon = 0$ in (2.1) if Ω is locally closed around $\bar{x}$ and the space X is *Asplund*, i.e., each separable subspace of X has a separable dual. The latter class includes all spaces with a Fréchet differentiable renorm, particularly every reflexive space. On the other hand, there are Asplund spaces that fail to have even a Gâteaux differentiable renorm; see [5, 14] for more details, discussions, and references.

In contrast to (2.2), the normal cone (2.1) is often *nonconvex* enjoying nevertheless *full calculus* in the framework of Asplund spaces, while a number of useful calculus results are also available in arbitrary Banach spaces; see [14, 15]. This calculus is mainly based on *extremal/variational principles* that replace convex separation theorems in nonconvex settings. Accordingly, similar well-developed calculi hold true for the associated subgradients of extended-real-valued functions and coderivatives of set-valued mappings defined below.

A set $\Omega \subset X$ is *normally regular* at $\bar{x} \in \Omega$ if

$$N(\bar{x};\Omega) = \widehat{N}(\bar{x};\Omega). \tag{2.3}$$

Besides convex sets, this property is satisfied in other important settings, particularly for sets described by smooth equalities and inequalities under the Mangasarian-Fromovitz constraint qualification. The reader can find more information about (2.3) and other notions of set regularity in [3, 14, 23] and the references therein. Note however that the normal regularity (2.3) *fails* for sets homeomorphic to *graphs* of single-valued nonsmooth Lipschitzian mappings, which is particularly the case of maximal monotone operators; see [14, Subsections 1.2.2 and 3.2.4].

Considering next a set-valued mapping $F\colon X \rightrightarrows Y$ between Banach spaces and a point $(\bar{x},\bar{y}) \in \operatorname{gph} F$ from its graph

$$\operatorname{gph} F := \big\{(x,y) \in X \times Y \big| \, y \in F(x)\big\},$$

we define the (normal) *coderivative* $D^*F(\bar{x},\bar{y})\colon Y^* \rightrightarrows X^*$ of F at $(\bar{x},\bar{y})$ by

$$D^*F(\bar{x},\bar{y})(y^*) := \big\{x^* \in X^* \big| \, (x^*,-y^*) \in N\big((\bar{x},\bar{y});\operatorname{gph} F\big)\big\} \tag{2.4}$$

and drop $\bar{y} = f(\bar{x})$ in (2.4) for single-valued mappings $F = f\colon X \to Y$. Observe that

$$D^*f(\bar{x})(y^*) = \big\{\nabla f(\bar{x})^* y^*\big\} \quad \text{for all} \;\; y^* \in Y^*$$

in any Banach spaces provided that $f\colon X \to Y$ is single-valued and *strictly differentiable* at $\bar{x}$ (in particular, when it is continuously differentiable around this point). This means that the coderivative (2.4) is a proper extension of the *adjoint* derivative operator to nonsmooth and set-valued mappings.

Furthermore, it comes directly from definitions (2.4) and (2.1) that the coderivative $D^*F(\bar{x},\bar{y})$ admits the limiting representation

$$D^*F(\bar{x},\bar{y})(y^*) := \Big\{x^* \in X^* \Big|\ \exists\, \varepsilon_k \downarrow 0,\ (x_k,y_k) \overset{\mathrm{gph}\,F}{\longrightarrow} (\bar{x},\bar{y}),\ (x_k^*,y_k^*) \overset{w^*}{\longrightarrow} (x^*,y^*) \text{ with } (x_k^*,-y_k^*) \in \widehat{N}_{\varepsilon_k}\big((x_k,y_k);\mathrm{gph}\,F\big) \text{ as } k \in I\!N\Big\},$$

where we can equivalently let $\varepsilon_k = 0$ for all $k \in I\!N$ if both spaces X and Y are Asplund and if the graph of F is locally closed around $(\bar{x},\bar{y})$. For some results of this paper we need the following property

$$D^*F(\bar{x},\bar{y})(y^*) := \Big\{x^* \in X^* \Big|\ \exists\, \varepsilon_k \downarrow 0,\ (x_k,y_k) \overset{\mathrm{gph}\,F}{\longrightarrow} (\bar{x},\bar{y}),\ x_k^* \overset{w^*}{\longrightarrow} x^*,\ \|y_k^* - y^*\| \to 0 \text{ with } (x_k^*,-y_k^*) \in \widehat{N}_{\varepsilon_k}\big((x_k,y_k);\mathrm{gph}\,F\big) \text{ as } k \in I\!N\Big\} \tag{2.5}$$

called [14] the *strong coderivative normality* of F at $(\bar{x},\bar{y})$; as above, one can put $\varepsilon_k \equiv 0$ in (2.5) in the Asplund space setting. Comparing (2.5) with the above limiting description of the coderivative (2.4), it is easy to see that the strong coderivative normality means that the coderivative $D^*F(\bar{x},\bar{y})(y^*)$ does not change if we replace the weak* convergence $y_k^* \overset{w^*}{\to} y^*$ by the norm convergence $y_k^* \to y^*$ in Y^*, while keeping the weak* convergence $x_k^* \overset{w^*}{\to} x^*$ in X^*. Of course, property (2.5) is automatic when Y is finite-dimensional. It also holds for broad classes of single-valued and set-valued mappings with values in infinite-dimensional spaces. We refer the reader to Proposition 4.9 from the book [14] summarizing important classes of mappings that enjoy the coderivative normality property (2.5). They particularly include mappings that are *N-regular* at $(\bar{x},\bar{y})$ (i.e., those whose graphs are normally regular (2.3) at $(\bar{x},\bar{y})$; hence both convex-graph and strictly differentiable ones), also the so-called "strictly Lipschitzian mappings," etc.

Consider further an extended-real-valued function $\varphi\colon X \to \overline{I\!R}$ finite at $\bar{x}$ and the associated *epigraphical multifunction* $E_\varphi\colon X \rightrightarrows I\!R$ given by

$$E_\varphi(x) := \big\{\mu \in I\!R \big|\ \mu \ge \varphi(x)\big\} \ \text{ with } \ \mathrm{gph}\,E_\varphi = \mathrm{epi}\,\varphi.$$

Then the *basic subdifferential* $\partial\varphi(\bar{x})$ and the *singular subdifferential* $\partial^\infty\varphi(\bar{x})$ of φ at $\bar{x}$ can be defined via the coderivative (2.4) of E_φ by, respectively,

$$\partial\varphi(\bar{x}) := D^*E_\varphi\big(\bar{x},\varphi(\bar{x})\big)(1) \ \text{ and } \ \partial^\infty\varphi(\bar{x}) := D^*E_\varphi\big(\bar{x},\varphi(\bar{x})\big)(0). \tag{2.6}$$

If the space X is Asplund and if φ is lower semicontinuous (l.s.c.) around $\bar{x}$, then one has the analytic representation of both constructions in (2.6) by

$$\partial\varphi(\bar{x}) = \mathop{\mathrm{Lim\,sup}}_{x \overset{\varphi}{\to} \bar{x}} \widehat{\partial}\varphi(x), \quad \partial^\infty\varphi(\bar{x}) = \mathop{\mathrm{Lim\,sup}}_{\substack{x \overset{\varphi}{\to} \bar{x} \\ \lambda \downarrow 0}} \lambda\widehat{\partial}\varphi(x) \tag{2.7}$$

via the so-called *Fréchet subdifferential*

$$\widehat{\partial}\varphi(\bar{x}) := \Big\{x^* \in X^* \Big|\ \liminf_{x \to \bar{x}} \frac{\varphi(x) - \varphi(\bar{x}) - \langle x^*, x - \bar{x}\rangle}{\|x - \bar{x}\|} \ge 0\Big\},$$

which is also known as the subdifferential in the sense of *viscosity solutions* of φ at $\bar{x}$. The symbol $x \xrightarrow{\varphi} \bar{x}$ in (2.7) signifies that $x \to \bar{x}$ with $\varphi(x) \to \varphi(\bar{x})$. Observe that $\partial^\infty\varphi(\bar{x}) = \{0\}$ if φ is locally Lipschitzian around $\bar{x}$.

Finally in this section, we define and discuss certain "sequential normal compactness" properties of sets, set-valued mappings, and extended-real-valued functions needed for establishing the results of the paper in *infinite-dimensional* settings. Given a set $\Omega \subset X$, we say that it is *sequentially normally compact* (SNC) at $\bar{x} \in \Omega$ if for any sequences $\varepsilon_k \downarrow 0$, $x_k \xrightarrow{\Omega} \bar{x}$, and $x_k^* \xrightarrow{w^*} 0$ one has

$$\|x_k^*\| \to 0 \text{ provided that } x_k^* \in \widehat{N}_{\varepsilon_k}(x_k;\Omega) \text{ as } k \to \infty,$$

where ε_k can be equivalently omitted when X is Asplund and Ω is locally closed around $\bar{x}$. It is automatic when Ω is *compactly epi-Lipschitzian* (CEL) around $\bar{x}$ in the sense of Borwein and Strójwas [1], while in general the SNC requirement may be essentially weaker than the CEL one; see [7] for various examples in Banach and Asplund spaces.

Accordingly, a set-valued mapping $F\colon X \rightrightarrows Y$ is *SNC* at $(\bar{x},\bar{y}) \in \operatorname{gph} F$ if its graph is SNC at this point. Given an extended-real-valued function $\varphi\colon X \to \overline{I\!R}$ finite at $\bar{x}$, we sat that it is *sequentially epi-compact* (SNEC) at this point if its epigraph is SNC at $(\bar{x},\varphi(\bar{x}))$. Note that latter property is automatic if φ is locally Lipschitzian around $\bar{x}$.

The SNC property and its modifications automatically hold in finite dimensions being among the *most essential ingredients* of infinite-dimensional variational analysis and generalized differentiation. They are unavoidably present in calculus rules for robust generalized differential constructions discussed above and in the corresponding optimality conditions. It is important to emphasize that a well-developed *full calculus* is available for such properties (mostly in Asplund spaces while also in general Banach space settings), in the sense that they are known to be preserved while various operations are performed on sets, set-valued mappings, and extended-real-valued functions under natural qualification conditions; see [14, 15] for more details and applications.

3 Linear Suboptimality from the Extremal Principle

Following the *geometric approach* to variational analysis and generalized differentiation [14, 15], we start with extremal properties of sets and then proceed with solutions to constrained optimization problems. Given two subsets Ω_1 and Ω_2 of a normed space X, recall [10] that $\bar{x} \in \Omega_1 \cap \Omega_2$ is a *local extremal point* of the set system $\{\Omega_1,\Omega_2\}$ if there exists a neighborhood U of $\bar{x}$ such that for any $\varepsilon > 0$ there is $a \in \varepsilon I\!B$ with

$$(\Omega_1 + a) \cap \Omega_2 \cap U = \emptyset.$$

Loosely speaking, the local extremality of sets at a common point means that they can be locally "pushed apart" by a small perturbation (translation) of one of them.

It is clear that every *boundary point* $\bar{x}$ of a closed set Ω is a local extremal point of the pair $\{\Omega, \{\bar{x}\}\}$. In general, this geometric concept of extremality covers conventional notions of optimal solutions to various problems of scalar and vector/multiobjective optimization, equilibria, etc. To illustrate it, let us consider a local optimal solution $\bar{x}$ to the following problem of *constrained optimization*:

$$\text{minimize } \varphi(x) \text{ subject to } x \in \Omega \subset X.$$

Then one can easily check that $(\bar{x}, \varphi(\bar{x}))$ is a local extremal point of the set system $\{\Omega_1, \Omega_2\}$ in $X \times I\!R$ with $\Omega_1 = \operatorname{epi}\varphi$ and $\Omega_2 = \Omega \times \{\varphi(\bar{x})\}$. More examples of extremal systems of sets related and also not related to optimization can be found in [14, 15],

It is not hard to observe that $\bar{x} \in \Omega_1 \cap \Omega_2$ is a *local extremal point* of the set system $\{\Omega_1, \Omega_2\}$ *if and only if*

$$\vartheta\big(\Omega_1 \cap B_r(\bar{x}), \Omega_2 \cap B_r(\bar{x})\big) = 0 \quad \text{for some} \quad r > 0, \tag{3.1}$$

where the *measure of overlapping* $\vartheta(\Omega_1, \Omega_2)$ for the sets Ω_1, Ω_2 is defined by

$$\vartheta(\Omega_1, \Omega_2) := \sup\big\{\nu \geq 0 \big|\ \nu I\!B \subset \Omega_1 - \Omega_2\big\}$$

Modifying the constant $\vartheta(\cdot, \cdot)$ in (3.1), Kruger introduced (under the name of "extended extremality" in [8] and "weak stationarity" in [9]) the new notion of local extremality for set systems in normed spaces, which in fact reflects a certain amount of *linear subextremality*; cf. Section 1 and the discussion below.

Definition 3.1 (linear subextremality of sets). *Given $\Omega_1, \Omega_2 \subset X$ and $\bar{x} \in \Omega_1 \cap \Omega_2$, we say that the set system $\{\Omega_1, \Omega_2\}$ is* LINEARLY SUBEXTREMAL *around the point $\bar{x}$ if*

$$\vartheta_{\rm lin}(\Omega_1, \Omega_2, \bar{x}) := \liminf_{\substack{x_i \xrightarrow{\Omega_i} \bar{x} \\ r \downarrow 0}} \frac{\vartheta\big([\Omega_1 - x_1] \cap r I\!B, [\Omega_2 - x_2] \cap r I\!B\big)}{r} = 0 \tag{3.2}$$

with $i = 1, 2$ under the lim inf *sign in* (3.2).

It is clear that the set extremality in the sense of (3.1) implies the linear subextremality in the sense of (3.2), but not vice versa. Let us discuss some specific features of linear subextremality for set systems that distinguish this notion from the concept of (3.1).

(a) The constant $\vartheta_{\rm lin}(\Omega_1, \Omega_2, \bar{x})$ defined in (3.2), in contrast to the one from (3.1), involves a *linear rate* of set perturbations as $r \downarrow 0$. Therefore, condition (3.2) describes a *local nonoverlapping at linear rate* for the sets Ω_1 and Ω_2, while condition (3.1) corresponds to a local nonoverlapping of these sets with an *arbitrary rate* as $r \downarrow 0$,

(b) Condition (3.2) requires not the *precise* local nonoverlapping of the given sets but up to their *infinitesimally small deformations.*

(c) Condition (3.2) does not require that the sets Ω_1 and Ω_2 nonoverlap *exactly at* the point $\bar{x}$. Moreover, it is easy to observe from the relations in (b) that (3.2) holds if, given any neighborhood U of $\bar{x}$, there are points $x_1 \in \Omega_1 \cap U$ and $x_2 \in \Omega_2 \cap U$ ensuring an *approximate nonoverlapping* of the translated sets $\Omega_1 - x_1$ and $\Omega_2 - x_2$ with a linear rate.

One of the most important results in the geometric theory of variational analysis and its applications is the so-called *extremal principle* providing *necessary* conditions for local extremal points of systems of closed sets. Its first versions were formulated in [10], while the most advanced result on the *exact* (pointbased) extremal principle is given in [14, Theorem 2.22] in the following form:

The Extremal Principle. *Let $\bar{x} \in \Omega_1 \cap \Omega_2$ be a local extremal point of the set system $\{\Omega_1, \Omega_2\}$, where Ω_1 and Ω_2 are locally closed subsets of an Asplund space X. Assume that either Ω_1 or Ω_2 is SNC at $\bar{x}$. Then there is $x^* \in X^*$ satisfying*

$$x^* \in N(\bar{x}; \Omega_1) \cap \big(-N(\bar{x}; \Omega_2)\big), \quad \|x^*\| = 1. \tag{3.3}$$

The next theorem, which follows from [15, Theorem 5.89] and [14, Theorem 2.22(ii], shows that the above conditions of the *extremal principle* are *necessary* not only for local extremal points of $\{\Omega_1, \Omega_2\}$ but also for a less restrictive notion of *linear subextremality*, providing actually a *characterization* of Asplund spaces. Moreover, these conditions happen to be *necessary and sufficient* for linear subextremality in finite-dimensional spaces.

Theorem 3.2 (linear subextremality via the extremal principle). *Let Ω_1 and Ω_2 be nonempty subsets of a Banach space X, and let $\bar{x} \in \Omega_1 \cap \Omega_2$. Assume that both Ω_1 and Ω_2 are locally closed around $\bar{x}$ and that one of them is sequentially normally compact at this point. The following assertions hold:*

(i) *If X is Asplund and if the system $\{\Omega_1, \Omega_2\}$ is linearly subextremal around $\bar{x}$, then there is $x^* \in X^*$ satisfying the relationships of the extremal principle* (3.3).

(ii) *Furthermore, if relationships* (3.3) *are satisfied for every set system $\{\Omega_1, \Omega_2\}$ linearly subextremal around $\bar{x}$, then the space X is Asplund.*

(iii) *Let $\dim X < \infty$. Then the system $\{\Omega_1, \Omega_2\}$ is linearly subextremal around $\bar{x}$ if and only if the relationships of the extremal principle* (3.3) *are satisfied.*

In [15, Section 5.4], the reader can find applications of Theorem 3.2 to various problems of constrained multiobjective optimization in finite and infinite dimensions. In what follows we provide new applications of this theorem and *calculus results* available for our basic generalized differential constructions to mathematical programs with equilibrium constraints.

4 Linear Suboptimality for General MPECs

Consider first the following problem of *constrained optimization* with geometric constraints:

$$\text{minimize } \varphi(x) \text{ subject to } x \in \Omega \subset X, \tag{4.1}$$

where $\varphi\colon X \to \overline{I\!R}$ is an extended-real-valued function on the normed space X. We have seen in Section 3 that a local optimal solution $\bar{x}$ (in the standard sense) to problem (4.1) geometrically corresponds to the fact that $(\bar{x}, \varphi(\bar{x}))$ is a *local extremal point* to the set system $\{\operatorname{epi}\varphi, \Omega \times \{\varphi(\bar{x})\}\}$ in $X \times I\!R$. Following this geometric idea and having in mind the subsequent application of the extremal principle, we define the notion of local suboptimality for (4.1) based on local subextremality of the above set system formulated in Definition 3.1.

Definition 4.1 (linear suboptimality for constrained minimization problems). *Let $\bar{x}$ be a feasible solution to problem* (4.1). *We say that it is* LINEARLY SUBOPTIMAL *with respect to* (φ, Ω) *if the point* $(\bar{x}, \varphi(\bar{x}))$ *is linearly subextremal for the set system* $\{\Omega_1, \Omega_2\}$ *with* $\Omega_1 := \operatorname{epi} \varphi$ *and* $\Omega_2 := \Omega \times \{\varphi(\bar{x})\}$.

It is not hard to check that $\bar{x}$ is linearly suboptimal to (4.1) in the sense of Definition 4.1 if it is linearly subminimal to (4.1) in the sense of [15, Definition 5.101], i.e.,

$$\limsup_{\substack{x \xrightarrow{\Omega} \bar{x} \\ \varphi(x) \to \varphi(\bar{x}) \\ r \downarrow 0}} \inf_{u \in B_r(x) \cap \Omega} \frac{\varphi(u) - \varphi(x)}{r} = 0, \tag{4.2}$$

which is in the line of Kruger [9] who formulated the latter property (as "weak inf-stationarity") with $\Omega = X$ and obtained its fuzzy/neighborhood characterization in terms of Fréchet subgradients applied actually to the extended-real-valued "condensed" function

$$\varphi_\Omega(x) := \varphi(x) + \delta(x; \Omega),$$

where $\delta(\cdot; \Omega)$ stands for the indicator function of a set. In [15, Section 5.4], we derived certain *exact/pointbased* characterizations of linear subminimality in the constrained (4.1) involving our basic constructions from Section 2. The most valuable "separated" characterizations of linear subminimality in (4.1), i.e., those expressed via the initial data φ and Ω of the given problem, are required certain *regularity* assumptions on φ and Ω; namely, the normal regularity of Ω at $\bar{x}$ and the "lower regularity" $\partial\varphi(\bar{x}) = \widehat{\partial}\varphi(\bar{x})$ of φ at this point. The regularity assumptions ensure in fact that the notions of linear suboptimality and linear subminimality agree; compare, e.g., [15, Corollary 5.107] and the proof of Theorem 4.2 below. Let us mention that the notion of linear suboptimality from Definition 4.1 is *always* different from/weaker than linear subminimality (or weak inf-stationarity) unless the formulated regularity assumptions are satisfied.

Let us emphasize a serious advantage of the above suboptimality notions in comparison with the convention concept of optimality: namely, the intrinsic *stability* of the new notions the with respect to perturbations of the initial data, which is not the case of standard optimality; cf. [9] and [15, Subsection 5.4.3] with the examples and discussions therein. Furthermore, for *smooth* problems of unconstrained minimization we known that linear suboptimality/subminimality reduces to the classical *stationarity* $\nabla\varphi(\bar{x}) = 0$ that is the base for constructing *descent directions* and other numerical devices. Thus the weakened suboptimality/stationarity notions under consideration can be potentially used in this framework, which we intend to investigate in more details in our future research.

The crucial advantage of linear suboptimality from Definition 4.1 in comparison with linear subminimality (4.2) in constrained optimization is that now we are able to justify that these "separated" condition happen to be *necessary and sufficient* for linear suboptimality *without imposing any regularity assumptions.* Since the main goal of this paper is to study mathematical programs with equilibrium constraints, we consider in what follows only problems that fell into this category, while the reader can develop this line for other constrained optimization problems studied, e.g., in [15].

Let us start with the *abstract MPEC* given in the form:

$$\text{minimize } \varphi(x,y) \text{ subject to } y \in S(x), \tag{4.3}$$

where $\varphi\colon X \times Y \to \overline{I\!R}$ is an extended-real-valued function while $S\colon X \rightrightarrows Y$ is a set-valued mapping. A point $(\bar{x},\bar{y})$ feasible to (4.3) is *linearly suboptimal* for this problem if it is linearly suboptimal with respect to (φ,Ω) in the sense of Definition 4.1 with $\Omega = \operatorname{gph} S \subset X \times Y$.

Theorem 4.2 (linear suboptimality in abstract MPECs). *Let $(\bar{x},\bar{y})$ be a feasible solution to the abstract MPEC* (4.3). *Assume that both spaces X and Y are Asplund, that the function φ is finite at $(\bar{x},\bar{y})$ and l.s.c. around this point, and that the mapping S is locally closed-graph around $(\bar{x},\bar{y})$. The following assertions hold:*

(i) *Impose the qualification condition*

$$\partial^\infty\varphi(\bar{x},\bar{y}) \cap \big[-N\big((\bar{x},\bar{y});\operatorname{gph} S\big)\big] = \{0\} \tag{4.4}$$

and assume that either S is SNC at $(\bar{x},\bar{y})$, or φ is SNEC at this point; observe that both the SNEC property of φ and the qualification condition (4.4) *are automatic when φ is locally Lipschitzian around $(\bar{x},\bar{y})$. Then the inclusion*

$$0 \in \partial\varphi(\bar{x},\bar{y}) + N\big((\bar{x},\bar{y});\operatorname{gph} S\big) \tag{4.5}$$

provides a necessary condition for linear suboptimality of $(\bar{x},\bar{y})$ to (4.3).

(ii) *Assume in addition to* (4.4) *that both spaces X and Y be finite-dimensional. Then condition* (4.5) *is necessary and sufficient for linear suboptimality of $(\bar{x},\bar{y})$ to* (4.3).

Proof. According to our definition of linear suboptimality in the case of MPECs, this notion for $(\bar{x},\bar{y})$ in (4.3) means that the point $(\bar{x},\bar{y},\varphi(\bar{x},\bar{y}))$ is *linearly subextremal* for the system of *locally closed* sets defined by

$$\Omega_1 := \operatorname{epi}\varphi, \quad \Omega_2 := \operatorname{gph} S \times \{\varphi(\bar{x},\bar{y})\} \text{ in } X \times Y \times I\!R. \tag{4.6}$$

Note that the product space $X \times Y \times I\!R$ is *Asplund* as a product of Asplund spaces; see, e.g., [5]. Observe also that the SNC property of Ω_1 at $(\bar{x},\bar{y},\varphi(\bar{x},\bar{y}))$ in (4.6) reduces to the SNEC property of φ at $(\bar{x},\bar{y})$, while the SNC property of Ω_2 at this point is equivalent to the SNC property of the mapping S at $(\bar{x},\bar{y})$. Thus we can apply to (4.6), under the assumptions made in the first part of this theorem, the results of Theorem 3.2 based on the extremal principle. Employing assertion (i) of the latter theorem, we find a dual-space triple $(x^*,y^*,\lambda) \in X^* \times Y^* \times I\!R$ such that

$$0 \neq (x^*,y^*,\lambda) \in N\big((\bar{x},\bar{y},\varphi(\bar{x},\bar{y}));\operatorname{epi}\varphi\big) \cap \big[-N\big((\bar{x},\bar{y},\varphi(\bar{x},\bar{y}));\operatorname{gph} S \times \{\varphi(\bar{x},\bar{y})\}\big)\big].$$

Taking into account the constructions of basic and singular subgradients in (2.6) and the simple product formula for basic normals (which holds as equality), we conclude that $\lambda \neq 0$ under the qualification condition (4.4) and then derive the necessary condition (4.5) for linear suboptimality of $(\bar{x},\bar{y})$ to the MPEC problem (4.3). This gives (i).

To justify assertion (ii) of the theorem, we can repeat the above arguments applying assertion (ii) of Theorem 3.2 instead of assertion (i). $\triangle$

Our main interest in what follows is to establish *characterizations* of linear suboptimality for MPECs (4.3) with equilibrium constraints $y \in S(x)$ given as solution maps to parametric generalized equations of type (1.2). On the other hand, the set-valued mapping $S(\cdot)$ in Theorem 4.2 may be described by different relations, e.g., via *functional constraints* of equality and inequality types defined by

$$S(x) := \big\{ y \in Y \big|\ \varphi_i(x,y) \le 0,\ i = 1,\dots,m;\ \ \varphi_i(x,y) = 0,\ i = m+1,\dots,m+r \big\} \tag{4.7}$$

with the real-valued functions φ_i. Let us present a consequence of Theorem 4.2(ii) for constraints given by (4.7) confining ourself to the case when the cost function φ is locally Lipschitzian around the reference point. Other results of this type can be derived from Theorem 4.2(ii) and *equality* type representations of the normal cone to graphs of set-valued mappings $S\colon X \rightrightarrows Y$ given my more general relations; see [14, Subsection 4.3.1].

Corollary 4.3 (characterization of linear suboptimality for problems with parametric functional constraints). *Given a feasible solution to problem* (4.3) *with the functional constraints, assume that φ is locally Lipschitzian around $(\bar{x},\bar{y})$, that φ_i are strictly differentiable at this point for all $i = 1,\dots,m+r$, and that the spaces X and Y are finite-dimensional. Impose the Mangasarian-Fromovitz constraint qualification:*

(a) $\nabla\varphi_{m+1}(\bar{x},\bar{y}),\dots,\nabla\varphi_{m+r}(\bar{x},\bar{y})$ *are linearly independent, and*

(b) *there is $u \in X \times Y$ satisfying*

$$\langle \nabla\varphi_i(\bar{x},\bar{y}), u\rangle < 0, \quad i \in \{1,\dots,m\} \cap I(\bar{x},\bar{y}),$$

$$\langle \nabla\varphi_i(\bar{x},\bar{y}), u\rangle = 0, \quad i = m+1,\dots,m+r,$$

where $I(\bar{x},\bar{y}) := \big\{ i = 1,\dots,m+r \big|\ \varphi_i(\bar{x},\bar{y}) = 0 \big\}$. Then $(\bar{x},\bar{y})$ is linearly suboptimal to (4.3), (4.7) *if and only if one has the inclusion*

$$-\sum_{i=1}^{m+r} \lambda_i \nabla\varphi_i(\bar{x},\bar{y}) \in \partial\varphi(\bar{x},\bar{y})$$

with some multipliers $\lambda_i \in I\!R$ satisfying the sign and complementary slackness conditions

$$\lambda_i \ge 0 \ \ \textit{and} \ \ \lambda_i\varphi_i(\bar{x},\bar{y}) = 0 \ \ \textit{for all} \ \ i = 1,\dots,m.$$

Proof. It follows from Theorem 4.2(ii) due to the normal cone representation

$$N\big((\bar{x},\bar{y}); \operatorname{gph} S\big) = \Big\{ \sum_{i \in I(\bar{x},\bar{y})} \lambda_i \nabla\varphi_i(\bar{x},\bar{y}) \Big|\ \ \lambda_i \in I\!R \ \text{ for all } \ i \in I(\bar{x},\bar{y}) \ \text{ and } \ \lambda_i \ge 0 \ \text{ if } \ \lambda_i \in \{1,\dots,m\} \cap I(\bar{x},\bar{y}) \Big\}$$

for the set-valued mapping S defined by (4.7) under the assumed Mangasarian-Fromovitz constraint qualification; see [14, Corollary 4.45]. Note that the qualification condition (4.4) is fulfilled due to $\partial^\infty \varphi(\bar{x}, \bar{y}) = \{0\}$ for locally Lipschitzian functions. $\triangle$

Now we consider a general class of MPECs with equilibrium constraints described by solution maps to *generalized equations*:

$$\text{minimize } \varphi(x,y) \text{ subject to } 0 \in f(x,y) + Q(x,y), \tag{4.8}$$

where $f\colon X \times Y \to Z$ is a single-valued mapping while $Q\colon X \times Y \rightrightarrows Z$ is a set-valued mapping between Banach spaces. Since the main purpose of this paper is to provide *characterizations* of linear suboptimality for MPEC optimization problems from the viewpoint of the *extremal principle* for set systems, we restrict ourself—in view of Theorem 3.2(ii) characterizing set extremality—to the case of finite-dimensional spaces X and Y. Note that the space Z in (4.8) may be infinite-dimensional. To distinguish terminologically between f and Q in (4.8), it is convenient to use the terms *base* and *field* for the single-valued part f and the set-valued part Q of (4.8), respectively; see [14, 15].

As mentioned in Section 1, the generalized equation formalism was introduced by Robinson [22] and has been well recognized as a very useful model for many issues in optimization and its applications. Note that, till the recent time, the main attention has been paid to the case of generalized equations with *parameter-independent fields* when $Q = Q(y)$ in (4.8). This particularly covers equilibrium constraints given by parametric *variational inequalities* and *complementarity problems.* On the other hand, there are broad classes of optimization-related and equilibrium problems that reduce to generalized equations with *parameter-dependent fields.* Among such problems, we mention the so-called *quasivariational inequalities*, where the generated convex set is *moving*, i.e., depends on parameters.

The next theorem provides *necessary and sufficient* conditions for linear suboptimality for MPECs (4.8) with equilibrium constraints governed by generalized equations in both cases of parameter-independent and parameter-dependent fields. The key issue to derive all the statements of this theorem involves, besides the principal characterization of Theorem 3.2(ii), the usage of *calculus rules with equalities* allowing us to *compute* the normal cone (2.1) to the graph of the *solution map*

$$S(x) := \big\{ y \in Y \big|\ 0 \in f(x,y) + Q(x,y) \big\} \tag{4.9}$$

to the parametric generalized equation in (4.8). The theorem below applies to both cases of parameter-independent and parameter-dependent fields illuminating some specific features of the parameter-independent case from the viewpoint of linear suboptimality. For simplicity we assume that the cost function φ is (4.8) is locally Lipschitzian around the reference point.

Theorem 4.4 (characterizing linear suboptimality for MPECs governed by generalized equations). *Let $(\bar{x}, \bar{y})$ be a feasible solution to (4.8), where $f\colon X \times Y \to Z$ is strictly differentiable at $(\bar{x}, \bar{y})$, where $Q\colon X \rightrightarrows Z$ is closed-graph around $(\bar{x}, \bar{y}, \bar{z})$ with $\bar{z} := -f(\bar{x}, \bar{y}) \in Q(\bar{x}, \bar{y})$, and where $\dim(X \times Y) < \infty$ while Z is arbitrarily Banach. Suppose also that the cost function $\varphi\colon X \times Y \to \overline{I\!R}$ is locally Lipschitzian around $(\bar{x}, \bar{y})$. Assume in addition that:*

(a) *either the partial derivative operator* $\nabla_x f(\bar{x},\bar{y})\colon X\times Y\to Z$ *is surjective/onto and the field* $Q=Q(y)$ *is parameter-independent;*

(b) *or* Z *is Asplund, the mapping* $Q=Q(x,y)$ *is* N*-regular and SNC at* $(\bar{x},\bar{y},\bar{z})$*, and the following Fredholm qualification condition holds: the adjoint generalized equation*

$$0\in\nabla f(\bar{x},\bar{y})^* z^*+D^*Q(\bar{x},\bar{y},\bar{z})(z^*) \tag{4.10}$$

has only the trivial solution $z^*=0$.

Then $(\bar{x},\bar{y})$ *is linearly suboptimal to* (4.8) *if and only if there is* $z^*\in Z^*$ *satisfying*

$$-\nabla f(\bar{x},\bar{y})^* z^*\in\partial\varphi(\bar{x},\bar{y})+D^*Q(\bar{x},\bar{y},\bar{z})(z^*). \tag{4.11}$$

Proof. The MPEC problem (4.8) under consideration is obviously equivalent to the abstract MPEC (4.3) with the equilibrium constraint mapping $S\colon X\rightrightarrows Y$ given by the solution map (4.9) to the underlying generalized equation. To characterize linear suboptimality of $(\bar{x},\bar{y})$ to MPEC (4.8), we apply assertion (ii) of Theorem 4.2 for the case of S defined by (4.9). It is to see that the graph of S is locally closed around $(\bar{x},\bar{y})$ due to the general assumptions on f and Q and that the qualification condition (4.4) is fulfilled by the local Lipschitz continuity of φ around $(\bar{x},\bar{y})$. It remains to compute the the normal cone $N((\bar{x},\bar{y});\operatorname{gph}S)$ for mapping (4.9) via the corresponding (generalized) differential constructions for f and Q.

To proceed, we consider first case (a) of the theorem and apply the normal cone representation for the graph of S from [14, Theorem 4.44(i)], which is as follows:

$$N\big((\bar{x},\bar{y});\operatorname{gph}S\big)=\Big\{(x^*,y^*)\in X^*\times Y^*\Big|\ \exists z^*\in Z^*\ \text{with}\ x^*=\nabla_x f(\bar{x},\bar{y})^* z^*,\ y^*\in\nabla_y f(\bar{x},\bar{y})z^*+D^*Q(\bar{y},\bar{z})(z^*)\Big\}$$

under the assumptions made in (a). Thus (4.5) reduces to (4.11), where $Q=Q(y)$.

To establish the characterization of linear suboptimality in case (b) with $Q=Q(x,y)$, we employ assertion (ii) of Theorem 4.44 from [14], which gives

$$N\big((\bar{x},\bar{y});\operatorname{gph}S\big)=\Big\{(x^*,y^*)\in X^*\times Y^*\Big|\ \exists z^*\in Z^*\ \text{such that}\ \Big(x^*-\nabla_x f(\bar{x},\bar{y})^* z^*,\ y^*-\nabla_y f(\bar{x},\bar{y})^* z^*\Big)\in D^*Q(\bar{x},\bar{y},\bar{z})(z^*)\Big\}$$

under the assumptions made in (b). Thus we arrive at (4.11) from (4.5). $\triangle$

Let us present efficient specifications of the results of Theorem 4.4 for the case of fields Q in (4.8) having *convex graphs.* This assumption automatically implies the N-regularity of Q at any point of the graph and ensures explicit form of the corresponding conditions characterizing linear suboptimality in the MPEC (4.8). To proceed, we need to introduce some notation.

Given $Q\colon X\times Y\rightrightarrows Z$ and $f\colon X\times Y\to Z$ strictly differentiable at $(\bar{x},\bar{y})$, consider the *linearized* set-valued operator $\Lambda\colon X\times Y\rightrightarrows Z$ with

$$\Lambda(x,y):=f(\bar{x},\bar{y})+\nabla_x f(\bar{x},\bar{y})(x-\bar{x})+\nabla_y f(\bar{x},\bar{y})(y-\bar{y})+Q(x,y).$$

Corollary 4.5 (linear suboptimality for MPECs with convex-graph fields). *Let $(\bar{x},\bar{y})$ be a feasible solution to the MPEC* (4.8), *where $f\colon X\times Y\to Z$ is strictly differentiable at $(\bar{x},\bar{y})$, where $Q\colon X \rightrightarrows Z$ is convex-graph and also closed-graph around $(\bar{x},\bar{y},\bar{z})$ with $\bar{z}=-f(\bar{x},\bar{y})\in Q(\bar{x},\bar{y})$, and where* $\dim(X\times Y)<\infty$ *while Z is arbitrarily Banach. Suppose also that the cost function $\varphi\colon X\times Y\to \overline{I\!R}$ is locally Lipschitzian around $(\bar{x},\bar{y})$. Assume in addition that either the requirements in* (a) *from Theorem* 4.4 *hold, or*

—Z is Asplund, $Q=Q(x,y)$ is SNC at $(\bar{x},\bar{y},\bar{z})$, and

$$N(0;\mathrm{rge}\,\Lambda)=\{0\}\ \ \textit{with}\ \ \mathrm{rge}\,\Lambda:=\Lambda(X,Y). \tag{4.12}$$

Then $(\bar{x},\bar{y})$ is linearly suboptimal to (4.8) *if and only if*

$$0\in\partial\varphi(\bar{x},\bar{y})+\mathrm{proj}_{X^*\times Y^*}N(0;\mathrm{rge}\,\Xi)\},\tag{4.13}$$

where $\Xi(x,y):=\big(x-\bar{x},y-\bar{y},\Lambda(x,y)\big)$ and proj *stands for the corresponding projection.*

Proof. Based on [15, Theorem 4.45] and the normal cone structure for convex sets, we conclude that the *Fredholm qualification condition* in Theorem 4.4(b) is equivalent to (4.12) and that the normal cone to the graph of the solution map S admits the representation:

$$N\big((\bar{x},\bar{y});\mathrm{gph}\,S\big)=\mathrm{proj}_{X^*\times Y^*}N(0;\mathrm{rge}\,\Xi).$$

This shows that characterization (4.5) reduces to (4.13) in the setting under consideration and thus completes the proof of the corollary. $\triangle$

The qualification condition (4.12) obviously holds if $0\in\mathrm{int}(\mathrm{rge}\,\Lambda)$, which is actually equivalent to (4.12) if the range of Λ is locally closed around $\bar{w}=0$ and SNC at this point. Note that, due to convexity, the SNC property of the sets $\mathrm{rge}\,\Lambda$ and $\mathrm{gph}\,Q$ can be characterized via their *finite codimensionality* by [14, Theorem 1.21].

Let us mention a special case of (4.8) when Q is given by

$$Q(x,y):=\begin{cases} E & \text{if } (x,y)\in\Omega,\\ \emptyset & \text{otherwise,}\end{cases}\tag{4.14}$$

where $E\subset Z$ and $\Omega\subset X\times Y$ are closed convex sets. In this case the interiority condition $0\in\mathrm{int}(\mathrm{rge}\,\Lambda)$ reduces to

$$0\in\mathrm{int}\Big\{f(\bar{x},\bar{y})+\nabla f(\bar{x},\bar{y})\big(\Omega-(\bar{x},\bar{y})\big)+E\Big\}$$

When $Q=Q(y)$ in (4.14), the corresponding constraint qualification (4.12) automatically holds under the classical *Robinson qualification condition*

$$0\in\mathrm{int}\Big\{f(\bar{x},\bar{y})+\nabla_y f(\bar{x},\bar{y})\big(\Omega-\bar{y}\big)+E\Big\}.$$

5 Classes of MPECs with Subdifferential Structures

In the final section of the paper we establish characterizations of linear suboptimality for MPECs (4.8) given in the *subdifferential form*, where the set-valued field Q of the equilibrium constraint (4.9) is described by a subdifferential operator

$$Q(x,y) = \partial\vartheta(x,y) \tag{5.1}$$

applied to an extended-real-valued l.s.c. function $\vartheta\colon X \times Y \to \overline{I\!R}$. In what follows we study the case when $\partial\vartheta$ in (5.1) is the *basic subdifferential* equivalently defined in (2.6) and (2.7), while other subdifferentials can be considered as well in this framework.

It has been well recognized that subdifferential forms of generalized equations are the most convenient for modeling *variational systems* arising, e.g., from describing sets of *optimal solutions* to parametric optimization problems as well as Lagrange multipliers, KKT vectors, complementarity conditions, variational inequalities of different types, their hemivariational and quasivariational extensions, etc. In particular, generalized equations (1.2) defining equilibrium constraints in (4.8) reduce to the classical *variational inequalities* (known also as variational inequalities of the first kind) given by

$$\text{find } \ y \in \Omega \ \text{ with } \ \langle f(y), v - y\rangle \geq 0 \ \text{ for all } \ v \in \Omega \tag{5.2}$$

when $Q(y) = N(y;\Omega)$ is the normal cone mapping generated by a *convex* set Ω. The normal cone operator in (5.2) can be rewritten in the subdifferential form $Q(y) = \partial\delta(y;\Omega)$ via the extended-real-valued *indicator function* $\delta(\cdot;\Omega)$ of the set Ω equal to 0 for $x \in \Omega$ and ∞ otherwise. The classical *complementarity* problem corresponds to (5.2) when Ω is the nonnegative orthant in $I\!R^n$. We refer the reader to the books [12, 15, 19] and to the bibliographies therein for detailed discussions of other kinds of variational systems that can be presented as generalized equations of the subdifferential type and then treated as equilibrium constraints for various MPECs.

The main attention in the section is paid to characterizations of linear suboptimality for general classes of MPECs with equilibrium constraints of the subdifferential type, where the field Q is given in two *composite forms* involving our basic first-order subdifferential given in (2.6) and (2.7).

The first class of such MPECs under consideration concerns equilibrium constraints with subdifferential fields (5.2) in the corresponding generalized equations given via the *composite potential* (we borrow the mechanical terminology; see [14] for more discussions) $\vartheta = \psi \circ g$ in (5.2), where $g\colon X \times Y \to W$ and $\psi\colon W \to \overline{I\!R}$. On the other words, we study MPECs of the following type:

$$\text{minimize } \ \varphi(x,y) \ \text{ subject to } \ 0 \in f(x,y) + \partial(\psi \circ g)(x,y), \tag{5.3}$$

where the range space for f and $Q = \partial(\psi \circ g)$ in (5.3) is either $X^* \times Y^*$ when $g = g(x,y)$, or Y^* when $g = g(y)$. Note that the composite potential structure in (5.3) is typically encountered in problems of constrained parametric optimization and related topics. In particular, this is the case for the broad class of *amenable* functions $\vartheta\colon I\!R^n \to \overline{I\!R}$ admitting the representation $\vartheta(u) = (\psi \circ g)(u)$, where $\psi\colon I\!R^m \to \overline{I\!R}$ is convex and l.s.c. while the mapping $g\colon I\!R^n \to I\!R^m$ is C^1 around the point in question under the qualification condition

$$\partial^\infty \psi\big(g(\bar u)\big) \cap \ker \nabla g(\bar u)^* = \{0\};$$

see [23] and also [14] for more discussions and references.

The next theorem contains *necessary and sufficient* conditions (i.e., characterization) for linear suboptimality in MPECs (5.3). To proceed in this direction, we need to use the concept of the second-order subdifferential for extended-real-valued functions together with appropriate results (chain rules) of the second-order subdifferential calculus.

Given an extended-real-valued function $\varphi\colon X \to \overline{I\!R}$ finite at $\bar x$ and a basic first-order subgradient $\bar y \in \partial\varphi(\bar x)$ from (2.6), recall [13, 14] that the *second-order subdifferential* of φ at $\bar x$ relative to $\bar v$ is the mapping $\partial^2\varphi(\bar x, \bar v)\colon X^{**} \rightrightarrows X^*$ with the values

$$\partial^2\varphi(\bar x, \bar v)(u) := (D^*\partial\varphi)(\bar x, \bar v)(u), \quad u \in X^{**}, \tag{5.4}$$

i.e., it is defined as the coderivative (2.4) of the first-order subdifferential mapping (an extension of the classical derivative-of-derivative approach in second-order differentiation). If $\varphi \in C^2$ near $\bar x$, we have

$$\partial^2\varphi(\bar x)(u) = \big\{\nabla^2\varphi(\bar x)^* u\big\} = \big\{\nabla^2\varphi(\bar x)u\big\} \text{ for all } u \in X^{**},$$

where the Hessian operator $\nabla^2\varphi(\bar x)$ is known to be *symmetric* in the case of reflexive spaces $X = X^{**}$. We refer the reader to the books [14, 15] for a developed theory and various applications of the second-order subdifferential construction (5.4) and its modifications.

To derive characterizations of linear suboptimality for MPECs of type (5.3), we involve *second-order* subdifferential *chain rules* giving a representation of $D^*Q = \partial^2(\psi \circ g)$ via the initial data (ψ, g). Again, we may apply only those calculus results that ensure chain rules as *equalities*. For brevity and by taking into account that the N-regularity property does not hold for broad classes of *subdifferential* mappings with nonsmooth potentials (see [14] for more discussions and references), we restrict ourselves to case (a) of Theorem 4.4 combined with second-order subdifferential calculus from [14].

Theorem 5.1 (characterizing linear suboptimality for MPECs with composite potentials). *Let $Q(y) = \partial(\psi \circ g)(y)$ under the assumptions imposed in case* (a) *of Theorem 4.4, where $g\colon Y \to W$ and $\psi\colon W \to \overline{I\!R}$ with a Banach space W. Suppose in addition that $g \in C^1$ near $\bar y$ with the surjective derivative $\nabla g(\bar y)\colon Y \to W$, that $\nabla g(\cdot)$ is strictly differentiable at $\bar y$, and that the graph of $\partial\psi\colon W \rightrightarrows W^*$ is locally norm-closed in $W \times W^*$ around $(\bar w, \bar v)$, where $\bar w := g(\bar y)$ and where $\bar v \in W^*$ is the unique linear functional satisfying*

$$-f(\bar x, \bar y) = \nabla g(\bar y)^* \bar v;$$

note that the closed-graph property of $\partial\psi$ is automatic if $\dim W < \infty$ *and if ψ is either continuous or amenable around the reference point.*

Then $(\bar x, \bar y)$ is linearly suboptimal to the MPEC formulated in (5.3) *if and only if there is a vector $u \in Y^{**} = Y$ satisfying the relationship*

$$-\nabla f(\bar x, \bar y)^* u \in \partial\varphi(\bar x, \bar y) + \Big(0, \nabla^2\langle \bar v, g\rangle(\bar y)u + g(\bar y)^*\partial^2\psi(\bar w, \bar v)\big(\nabla g(\bar y)u\big)\Big). \tag{5.5}$$

Proof. First of all, observe that the graph of $Q(y) = \partial(\psi \circ g)(y)$ is locally closed around $(\bar{y}, f(\bar{x}, \bar{y}))$ under the general assumptions of the theorem. Furthermore, the required graph-closedness for the subdifferential mapping $\partial\psi(\cdot)$ is automatic if W is finite-dimensional and if φ is either continuous or amenable around $\bar{x}$. In the first case it easily follows from representation (2.7) of the basic subdifferential, while the in the second case it is justified in [23]. Note that more general conditions ensuring the closed-graph property of $\partial\psi(\cdot)$ in infinite dimensions are established in Theorem 3.60 of [14].

To prove the theorem, it remains to compute the coderivative $D^*Q(\bar{y}, \bar{z})(u)$ with $u \in Y$ and $\bar{z} = -f(\bar{x}, \bar{y})$ for the composite subdifferential mapping $Q(y) = \partial(\psi \circ g)(y)$ in terms of the initial data ψ and g of the MPEC (5.3) under consideration. By the second-order subdifferential chain rule from [14, Theorem 1.127], we have the *equality*

$$\partial^2(\psi \circ g)(\bar{y}, \bar{z}) = \nabla^2\langle \bar{v}, g\rangle(\bar{y})u + \nabla g(\bar{y})^*\partial\psi(\bar{w}, \bar{v})\big(\nabla g(\bar{y})u\big) \quad \text{for all} \quad u \in Y \tag{5.6}$$

under the assumptions made in theorem. Substituting (5.6) into relationship (4.11) of Theorem 4.4 and taking into account that $Q = Q(y)$ in our setting, we arrive at characterization (5.5) of linear suboptimality for MPEC (5.3). △

When the cost function φ in (5.3) is *strictly differentiable* at $(\bar{x}, \bar{y})$ (as well as in some other cases; see [14]), we can *decompose* the subdifferential of φ at (5.5) and thus derive the following specification of Theorem 5.1.

Corollary 5.2 (linear suboptimality for MPECs with composite potentials and smooth costs). *Let the cost function φ be strictly differentiable at $(\bar{x}, \bar{y})$, in addition to the other assumptions of Theorem* 5.1. *Then $(\bar{x}, \bar{y})$ is linearly suboptimal to MPEC* (5.3) *if and only if*

$$0 \in \nabla_y\varphi(\bar{x}, \bar{y}) + \nabla_y f(\bar{x}, \bar{y})^*u + \nabla^2\langle \bar{v}, g\rangle(\bar{y})u + g(\bar{y})^*\partial^2\psi(\bar{w}, \bar{v})\big(\nabla g(\bar{y})u\big), \tag{5.7}$$

where the vector $u \in Y$ is uniquely defined by

$$-\nabla_x\varphi(\bar{x}, \bar{y}) = \nabla_x f(\bar{x}, \bar{y})^*u. \tag{5.8}$$

Proof. Since we now have

$$\partial\varphi(\bar{x}, \bar{y}) = \big\{\big(\nabla_x\varphi(\bar{x}, \bar{y}), \nabla_y\varphi(\bar{x}, \bar{y})\big)\big\},$$

inclusion (5.5) reduces to the simultaneous fulfillment of (5.7) and (5.8). The uniqueness of $u \in Y$ in (5.8) follows from the surjectivity of the partial derivative operator $\nabla_x f(\bar{x}, \bar{y})$, which is equivalent to the injectivity of the adjoint operator $\nabla_x f(\bar{x}, \bar{y})^*$. △

The second special class of MPECs we consider in this section involves equilibrium constraints given in the form

$$0 \in f(x, y) + (\partial\psi \circ g)(x, y), \tag{5.9}$$

where $f\colon X \times Y \to W^*$, $\psi\colon W \to \overline{I\!R}$, and $g\colon X \times Y \to W$. Following [14], we label (5.9) as generalized equations with *composite subdifferential fields.*

Besides the classical variational inequalities and related problems mentioned above, model (5.9) covers a broad range of parametric variational systems important in optimization/equilibrium theory and applications. In particular, framework (5.9) includes perturbed *implicit complementarity* problems of the type: find $y \in Y$ satisfying

$$f(x,y) \geq 0, \quad y - g(x,y) \geq 0, \quad \big\langle f(x,y), y - g(x,y)\big\rangle = 0,$$

where the inequalities are understood in the sense of some order on Y (e.g., componentwisely in finite-dimensions). Problems of this kind frequently arise in a large spectrum of mathematical models involving various types of economic and mechanical equilibria; see the books [12, 15, 19] and the references therein.

Theorem 5.3 (characterization of linear suboptimality for MPECs with composite subdifferential fields). *Let $Q(y) = (\partial\psi \circ g)(y)$ under the assumptions in case* (a) *of Theorem* 4.4, *where $\psi\colon W \to \overline{I\!R}$, and $g\colon X \times Y \to W$ for a Banach space W. Assume in addition that g is strictly differentiable at $\bar y$ with the surjective derivative $\nabla g(\bar y)$ and that the graph of $\partial\psi\colon W \rightrightarrows W^*$ is locally norm-closed in $W \times W^*$ around $(\bar w, \bar z)$, where $\bar w := g(\bar y)$ and $\bar z := -f(\bar x, \bar y)$; the latter is automatic when $\dim W < \infty$ and when ψ is either continuous or amenable around $\bar w$. Then $(\bar x, \bar y)$ is linearly suboptimal to the MPEC*

$$\text{minimize } \varphi(\bar x, \bar y) \text{ subject to } 0 \in f(x,y) + (\partial\psi \circ g)(x,y) \tag{5.10}$$

if and only if there is a linear functional $u \in W^{**}$ such that

$$-\nabla f(\bar x, \bar y)^* u \in \partial\varphi(\bar x, \bar y) + \Big(0, \nabla g(\bar y)^* \partial^2\psi(\bar w, \bar z)(u)\Big). \tag{5.11}$$

Proof. Apply Theorem 4.4 in case (a) to the field mapping $Q(y) = (\partial\psi \circ g)(y)$. Similarly to Theorem 5.1, check that the graph of Q is locally closed under the assumptions made. Using the coderivative chain rule from [14, Theorem 1.66] and definition (5.4) of the second-order subdifferential, we get the equalities

$$D^*(\partial\psi \circ g)(\bar y, \bar z)(u) = \nabla g(\bar y)\big(D^*\partial\psi\big)(\bar w, \bar z)(u) = \nabla g(\bar y)\partial^2\psi(\bar w, \bar z)(u) \tag{5.12}$$

whenever $u \in W^{**}$. Substituting (5.12) into (4.11), and taking into account that $Q = Q(y)$, we arrive at the characterization (5.11) of linear suboptimality for MPEC (5.10). $\triangle$

When the cost function φ in (5.10) is strictly differentiable, characterization (5.11) admits the following decomposition and simplification.

Corollary 5.4 (linear suboptimality for MPECs with composite subdifferential fields and smooth costs). *Assume in addition to Theorem* 5.3 *that φ is strictly differentiable at $(\bar x, \bar y)$. Then $(\bar x, \bar y)$ is linearly suboptimal to MPEC* (5.9) *if and only if*

$$0 \in \nabla_y\varphi(\bar x, \bar y) + \nabla_y f(\bar x, \bar y)^* u + \nabla g(\bar y)^* \partial^2\psi(\bar w, \bar z)(u),$$

where $u \in W^{**}$ is uniquely defined by

$$-\nabla_x\varphi(\bar x, \bar y) = \nabla_x f(\bar x, \bar y)^* u.$$

Proof. Similar to the case of Corollary 5.2. $\triangle$

We refer the reader to [6, 14, 16, 17, 18, 24, 25] and the discussions therein for efficient *calculations* of the second-order subdifferential $\partial^2\psi$ for favorable classes of extended-real-valued functions typically encountered in optimization-related problems and their various applications, particularly to mechanical and economical models. This makes it possible to utilize the characterizations of linear suboptimality for the MPECs studied above to specific classes of MPEC problems arising in optimization/equilibrium theory and applications. Note also that the characterizations established in this paper can be extended to MPECs involving, along with equilibrium constraints, constraints of other types, e.g., $(x, y) \in \Omega$. This can be done, following the above procedures, by using additional *equality* calculus rules of generalized differentiation developed in [14].

References

[1] Borwein, J.M., Strójwas, H.M.: Tangential approximations. Nonlinear Anal. **9**, 1347–1366 (1985)

[2] Borwein, J.M., Zhu, Q.J.: Techniques of Variational Analysis. Canadian Mathematical Society Series, Springer, New York, 2005

[3] Bounkhel, M., Thibault, L.: On various notions of regularity of sets in nonsmooth analysis. Nonlinear Anal. **48**, 223–246 (2002)

[4] Burke, J.V., Deng, S.: Weak sharp minima revisted, II: Application to linear regularity and error bounds. Math. Prog. **104**, 235–261 (2005)

[5] Deville, R., Godefroy, G., Zizler, V.: Smoothness and Renorming in Banach spaces. Wiley, New York, 1997

[6] Dontchev, A.L., Rockafellar, R.T.: Characterizations of strong regularity for variational inequalities over polyhedral convex sets. SIAM J. Optim. **7**, 1087–1105 (1996)

[7] Fabian, M., Mordukhovich, B.S.: Sequential normal compactness versus topological normal compactness in variational analysis. Nonlinear Anal. **54**, 1057–1067 (2003)

[8] Kruger, A.Y.: Strict (ε, δ)-semidifferentials and extremality conditions. Optimization **51**, 539–554 (2002)

[9] Kruger, A.Y.: Weak stationarity: eliminating the gap between necessary and sufficient conditions. Optimization **53**, 147–164 (2004)

[10] Kruger, A.Y., Mordukhovich, B.S.: Extremal points and the Euler equation in nonsmooth optimization. Dokl. Akad. Nauk BSSR **24**, 684–687 (1980)

[11] Levy, A.B., Poliquin, R.A., Rockafellar, R.T.: Stability of locally optimal solutions. SIAM J. Optim. **10**, 580–604 (2000)

[12] Luo, Z.Q., Pang, J.-S., Ralph, D: Mathematical Programs with Equilibrium Constraints. Cambridge University Press, Cambridge, UK, 1996

[13] Mordukhovich, B.S.: Sensitivity analysis in nonsmooth optimization. In: 'Theoretical Aspects of Industrial Design,' Field/Komkov (eds.), SIAM Proc. Appl. Math. **58**, 32–46 (1992)

[14] Mordukhovich, B.S.: Variational Analysis and Generalized Differentiation, I: Basic Theory. Grundlehren Series (Fundamental Principles of Mathematical Sciences) **330**, Springer, Berlin, 2006

[15] Mordukhovich, B.S.: Variational Analysis and Generalized Differentiation, II: Applications. Grundlehren Series (Fundamental Principles of Mathematical Sciences) **331**, Springer, Berlin, 2006

[16] Mordukhovich, B.S., Outrata, J.V.: Second-order subdifferentials and their applications. SIAM J. Optim. **12**, 139–169 (2001)

[17] Mordukhovich, B.S., Outrata, J.V., Červinka, M.: Equilibrium problems with complementarity constraints: Case study with applications to oligopolistic markets. Optimization (to appear)

[18] Outrata, J.V.: Optimality conditions for a class of mathematical programs with equilibrium constraints. Math. Oper. Res. **24**, 627–644 (1999)

[19] Outrata, J.V., Kočvara, M., Zowe, J.: Nonsmooth Approach to Optimization Problems with Equilibrium Constraints. Kluwer, Dordrecht, The Netherlands, 1998

[20] Poliquin, R.A., Rockafellar, R.T.: Tilt stability of a local minimum. SIAM J. Optim. **8**, 287–299 (1998)

[21] Polyak, B.T.: Introduction to Optimization. Optimization Software, New York, 1987

[22] Robinson, S.M.: Generalized equations and their solutions, I: Basic theory. Math. Progr. Study **10**, 128–141 (1979)

[23] Rockafellar, R.T., Wets, R.J.-B.: Variational Analysis. Grundlehren Series (Fundamental Principles of Mathematical Sciences) **317**, Springer, Berlin, 1998

[24] Ye, J.J.: Constraint qualifications and necessary optimality conditions for optimization problems with variational inequality constraints. SIAM J. Optim. **10**, 943–962 (1999)

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