Enhanced Optimality Conditions and New Constraint Qualifications for Nonsmooth Optimization Problems

by

Jin Zhang
B.A., Dalian University of Technology, 2007
M.Sc., Dalian University of Technology, 2010

A Dissertation Submitted in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Mathematics and Statistics

Enhanced Optimality Conditions and New Constraint Qualifications for Nonsmooth Optimization Problems

by

Jin Zhang
B.A., Dalian University of Technology, 2007
M.Sc., Dalian University of Technology, 2010

Supervisory Committee

Dr. Jane Ye, Supervisor
(Department of Mathematics and Statistics, University of Victoria)

Dr. Florin Diacu, Departmental Member
(Department of Mathematics and Statistics, University of Victoria)

Dr. G. Cornelis van Kooten, Outside Member
(Department of Economics, University of Victoria)

Supervisory Committee

Dr. Jane Ye, Supervisor
(Department of Mathematics and Statistics, University of Victoria)

Dr. Florin Diacu, Departmental Member
(Department of Mathematics and Statistics, University of Victoria)

Dr. G. Cornelis van Kooten, Outside Member
(Department of Economics, University of Victoria)

ABSTRACT

The main purpose of this dissertation is to investigate necessary optimality conditions for a class of very general nonsmooth optimization problems called the mathematical program with geometric constraints (MPGC). The geometric constraint means that the image of certain mapping is included in a nonempty and closed set.

We first study the conventional nonlinear program with equality, inequality and abstract set constraints as a special case of MPGC. We derive the enhanced Fritz John condition and from which, we obtain the enhanced Karush-Kuhn-Tucker (KKT) condition and introduce the associated pseudonormality and quasinormality condition. We prove that either pseudonormality or quasinormality with regularity implies the existence of a local error bound. We also give a tighter upper estimate for the Fréchet subdifferential and the limiting subdifferential of the value function in terms of quasinormal multipliers which is usually a smaller set than the set of classical normal multipliers.

We then consider a more general MPGC where the image of the mapping from a Banach space is included in a nonempty and closed subset of a finite dimensional space. We obtain the enhanced Fritz John necessary optimality conditions in terms of the approximate subdifferential. One of the technical difficulties in obtaining such a result in an infinite dimensional space is that no compactness result can be used to show the existence of local minimizers of a perturbed problem. We employ the celebrated Ekeland's variational principle to obtain the results instead. We then apply our results to the study of exact penalty and sensitivity analysis.

We also study a special class of MPCG named mathematical programs with equilibrium constraints (MPECs). We argue that the MPEC-linear independence constraint qualification is not a constraint qualification for the strong (S-) stationary condition when the objective function is nonsmooth. We derive the enhanced Fritz John Mordukhovich (M-) stationary condition for MPECs. From this enhanced Fritz John M-stationary condition we introduce the associated MPEC generalized pseudonormality and quasinormality condition and build the relations between them and some other widely used MPEC constraint qualifications. We give upper estimates for the subdifferential of the value function in terms of the enhanced M- and C-multipliers respectively.

Besides, we focus on some new constraint qualifications introduced for nonlinear extremum problems in the recent literature. We show that, if the constraint functions are continuously differentiable, the relaxed Mangasarian-Fromovitz constraint qualification (or, equivalently, the constant rank of the subspace component condition) implies the existence of local error bounds. We further extend the new result to the MPECs.

Contents

List of Figures

List of Abbreviations

MPGC	mathematical program with geometric constraints
NLP	nonlinear programming problem
MPEC	mathematical program with equilibrium constraints
NLSDP	nonlinear semidefinite programs
KKT	Karush-Kuhn-Tucker condition
CS	complementary slackness condition
CV	complementarity violation condition
WCG	weakly compactly generated
LICQ	linear independent constraint qualification
NNAMCQ	no nonzero abnormal multiplier constraint qualification
MFCQ	Mangasarian-Fromovitz constraint qualification
GCQ	Guignard constraint qualification
EGCQ	enhanced Guignard constraint qualification
CPLD	constant positive linear dependance constraint qualification
RCPLD	relaxed constant positive linear dependance constraint qualification
RCRCQ	relaxed constant rank constraint qualification
CRSC	rank of the subspace component condition
RMFCQ	relaxed Mangasarian-Fromovitz constraint qualification
CRMFCQ	constant rank Mangasarian-Fromovitz constraint qualification

List of Notations

Spaces and Orthants

$\mathbb{R}$	the real numbers
$\overline{\mathbb{R}}$	the extended-real numbers
$\mathbb{R}^n$	the n-dimensional real vector space
$\mathbb{R}^n_+$	the nonnegative orthant in $\mathbb{R}^n$
$\mathbb{R}^n_-$	the nonpositive orthant in $\mathbb{R}^n$
$\mathbb{X}^*$	the dual space of a Banach space $\mathbb{X}$
$\mathcal{S}^l$	the linear space of all $l \times l$ real symmetric matrices
$\mathcal{S}^l_-$	the cone of all $l \times l$ negative semidefinite matrices in $\mathcal{S}^l$

Sets

$\{x\}$	the set consisting of the vector x
int $\mathcal{C}$	interior of the set $\mathcal{C}$
cl $\mathcal{C}$	closure of the set $\mathcal{C}$
conv $\mathcal{C}$	convex hull of the set $\mathcal{C}$
cl*conv $\mathcal{C}$	weak* closure of the convex hull of the set $\mathcal{C}$
$\mathcal{C}^o$	polar of set $\mathcal{C}$
$\mathbb{B}(x, \epsilon)$	the closed ball centered at x with radius ϵ in $\mathbb{R}^n$
$\mathbb{B}$	the closed unit ball centered at 0 in $\mathbb{R}^n$
$\mathscr{E}$	orthogonal basis for an Euclidean space $\mathbb{Y}$
$\mathcal{B}_\delta(x)$	the open ball centered at x with radius δ
$\mathbb{B}_\mathbb{X}$	closed unit balls of the space $\mathbb{X}$
$\mathbb{B}_{\mathbb{X}^*}$	closed unit balls of the dual space $\mathbb{X}^*$ of $\mathbb{X}$
$A(x)$	set of active inequality constraints at x

Cones

$\mathcal{N}^{\pi}_{\Omega}(x)$	proximal normal cone to $\Omega \subseteq \mathbb{R}^n$ at x
$\widehat{\mathcal{N}}_{\Omega}(x)$	Fréchet normal cone to $\Omega \subseteq \mathbb{R}^n$ at x
$\mathcal{N}^{c}_{\Omega}(x)$	Clarke normal cone to $\Omega \subseteq \mathbb{R}^n$ at x
$T_{\Omega}(x)$	contingent cone to $\Omega \subseteq \mathbb{R}^n$ at x
$\widehat{\mathcal{N}}_{\epsilon}(x, \Omega)$	ϵ-normal cone to $\Omega \subseteq \mathbb{X}$ at x
$\mathcal{N}_{\Omega}(x)$	limiting normal cone to $\Omega \subseteq \mathbb{X}$ at x
$\mathcal{N}^{g}_{\Omega}(x)$	G-normal cone to $\Omega \subseteq \mathbb{X}$ at x
$\tilde{\mathcal{N}}^{g}_{\Omega}(x)$	nucleus cone to $\Omega \subseteq \mathbb{X}$ at x
$\mathcal{N}^{a}_{\Omega}(x)$	A-normal cone to $\Omega \subseteq \mathbb{X}$ at x
$\mathcal{T}_{\Omega}(x)$	contingent cone to $\Omega \subseteq \mathbb{X}$ at x
$\mathcal{T}^{w}_{\Omega}(x)$	weak contingent cone to $\Omega \subseteq \mathbb{X}$ at x
$\mathcal{T}^{c}_{\Omega}(x)$	Clarke tangent cone to $\Omega \subseteq \mathbb{X}$ at x

Sequences

$\limsup_{x \to x_0} \Phi(x)$	Painlevé-Kuratowski upper limit for a set-valued map Φ
$\liminf_{x \to x_0} \Phi(x)$	Painlevé-Kuratowski lower limit for a set-valued map Φ
$\overline{\mathrm{Lim\,sup}}_{x \to x_0}\ \Phi(x)$	topological counterpart of the Painlevé-Kuratowski upper limit for a set-valued map Φ

Functions

$\mathrm{dist}_{\mathcal{C}}(x)$	the distance between x and a closed set $\mathcal{C}$
$g^{+}(x)$	$\max\{0, g(x)\}$
$\partial^{\pi}\varphi(x)$	proximal subdifferential of a function $\varphi(x)$
$\partial^{c}\varphi(x)$	Clarke subdifferential of a function $\varphi(x)$
$\hat{\partial}_{\epsilon}\varphi(x)$	(Fréchet-like) ϵ-subdifferential of a function $\varphi(x)$

$\hat{\partial}\varphi(x)$	Fréchet subdifferential of a function $\varphi(x)$
$\partial\varphi(x)$	limiting subdifferential of a function $\varphi(x)$
$\partial^{\infty}\varphi(x)$	singular subdifferential of a function $\varphi(x)$
$D^{-}\varphi(x,d)$	lower Dini directional derivative of a function $\varphi(x)$
$\partial^{-}_{\epsilon}\varphi(x)$	Dini ϵ-subdifferential of a function $\varphi(x)$
$\partial^{a}\varphi(x)$	approximate subdifferential of a function $\varphi(x)$
$\varphi^{o}(x,d)$	Clarke's generalized directional derivative of a function $\varphi(x)$

ACKNOWLEDGEMENTS

I would like to thank:

my supervisor, Dr. Jane Ye, for her encouragement, patience and mentorship. Without her guidance and persistent help this dissertation would not have been possible.

Dr. Gui-Hua Lin, Dr. Xiao-Dai Dong, for the guidance and financial support.

UVIC, NSERC and China Scholarship Council for the fellowships and scholarships.

Last but not least, I would like to express my sincere gratitude to my family, for the support, kindness and generosity.

DEDICATION

To my late grandfather. His words of inspiration and encouragement in pursuit of excellence, still linger on.

Chapter 1

Introduction

This thesis is dedicated to a thorough investigation of various enhanced stationarity concepts and constraint qualifications for nonsmooth optimization problems and their applications. Only first-order necessary conditions are investigated. Sufficient conditions, are for the most part not considered, which remains a subject for future research. Nonetheless, it is our opinion that, at the time of print, this thesis contains an exhaustive discussion of the state of the art of the enhanced first-order theory for nonsmooth mathematical programming problems.

1.1 Background on enhanced optimality conditions

Consider the mathematical program with geometric constraints in $\mathbb{R}^n$:

$$(\mathrm{MPGC}_{\mathbb{R}^n}) \qquad \min_{x \in \mathcal{X}} \quad f(x) \qquad \text{s.t.} \quad F(x) \in \Lambda, \tag{1.1}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ and $F : \mathbb{R}^n \to \mathbb{R}^m$ are mappings, $\mathcal{X}$ and Λ are nonempty and closed subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively. The problem $\mathrm{MPGC}_{\mathbb{R}^n}$ is very general since it includes as special cases the conventional nonlinear program, the cone constrained program, the mathematical program with equilibrium constraints [49, 68], the problems considered in [27, 72], the semidefinite program, and the mathematical program with semidefinite cone complementarity constraints [21].

In the case when $F(x) := (h(x), g(x), x)$ and $\Lambda := \{0\}^p \times \mathbb{R}^q_- \times \mathcal{X}$, problem $\mathrm{MPGC}_{\mathbb{R}^n}$ is the nonlinear programming problem (NLP) in the form:

$$\text{(NLP)} \qquad \min \quad f(x) \qquad \text{s.t.} \quad x \in \mathcal{F},$$

where the feasible region $\mathcal{F}$ consists of equality and inequality constraints as well as an additional abstract set constraint $\mathcal{X} \subseteq \mathbb{R}^n$,

$$\mathcal{F} = \mathcal{X} \cap \{x : h_1(x) = 0, \ldots, h_p(x) = 0\} \cap \{x : g_1(x) \leq 0, \ldots, g_q(x) \leq 0\} \tag{1.2}$$

and all functions are assumed to be continuously differentiable.

In 1948, Fritz John [38] proposed the now well-known Fritz John necessary optimality condition for smooth optimization problems with inequality constraints only. In 1967, Mangasarian and Fromovitz [50] extended the Fritz John condition to smooth optimization problems with equality and inequality constraints (i.e. $\mathcal{X} = \mathbb{R}^n$). For the smooth case, Fritz John condition asserts that if x^* is a local optimal solution of problem (NLP) with $\mathcal{X} = \mathbb{R}^n$, then there exist scalars $\lambda_1^*, \ldots, \lambda_p^*$ and $\mu_0^*, \cdots, \mu_q^*$ not all zero, satisfying $\mu_j^* \geq 0$ for all $j = 0, 1, \ldots, q$ and

$$0 = \mu_0^* \nabla f(x^*) + \sum_{i=1}^{p} \lambda_i^* \nabla h_i(x^*) + \sum_{j=1}^{q} \mu_j^* \nabla g_j(x^*), \tag{1.3}$$

$$0 = \mu_j^* g_j(x^*), \tag{1.4}$$

where $\nabla\varphi(x)$ denotes the gradient of the function φ at x. Condition (1.4) is referred to as the complementary slackness condition (CS for short). We call a multiplier $(\lambda_1^*, \cdots, \lambda_p^*, \mu_1^*, \cdots, \mu_q^*)$ satisfying the Fritz John condition (1.3)-(1.4) with $\mu_0^* = 1$ and $\mu_0^* = 0$ a normal multiplier and an abnormal multiplier respectively. It follows from the Fritz John condition that if there is no nonzero abnormal multiplier then there must exist a normal multiplier. This simple corollary from the Fritz John condition leads to the so-called No Nonzero Abnormal Multiplier Constraint Qualification (NNAMCQ for short) or the so-called Basic Constraint Qualification for the Karush-Kuhn-Tucker (KKT for short) condition to hold at a local minimum. It was Mangasarian and Fromovitz who first pointed out that the Fritz John condition can be used to derive the KKT condition under the condition that the gradient vectors

$$\nabla h_i(x^*), i = 1, \ldots, p$$

are linearly independent and there exists a vector $d \in \mathbb{R}^m$ such that

$$\nabla h_i(x^*)^T d = 0 \quad i = 1, \ldots, p,$$

$$\nabla g_j(x^*)^T d < 0 \quad j \in A(x^*),$$

where $A(x^*) := \{j : g_j(x^*) = 0\}$ is the set of active inequality constraints at x^*, using the fact that the above condition is equivalent to the NNAMCQ by the Motzkin's transposition theorem. The above condition is now well-known as the Mangasarian-

Fromovitz Constraint Qualification (MFCQ).

The first but weaker versions of the enhanced Fritz John conditions were considered in a largely overlooked analysis by Hestenes [30] for the case of smooth optimization problem without an abstract set constraint. A version of the enhanced Fritz John condition first given by Bertsekas in [5] for a smooth problem with $\mathcal{X} = \mathbb{R}^n$ states that if x^* is a local optimal solution of problem (NLP) with $\mathcal{X} = \mathbb{R}^n$, then there exist scalars $\lambda_1^*, \ldots, \lambda_p^*$ and $\mu_0^* \geq 0, \cdots, \mu_q^* \geq 0$ not all zero satisfying (1.3) and the following sequential property: If the index set $I \cup J$ is nonempty, where

$$I = \{i | \lambda_i^* \neq 0\}, \qquad J = \{j \neq 0 | \mu_j^* > 0\},$$

then there exists a sequence $\{x^k\} \subseteq \mathcal{X}$ converging to x^* such that for all k,

$$f(x^k) < f(x^*), \qquad \lambda_i^* h_i(x^k) > 0,\ \forall i \in I, \qquad \mu_j^* g_j(x^k) > 0,\ \forall j \in J. \tag{1.5}$$

Condition (1.5) is stronger than the complementary slackness condition (1.4) since if $\mu_j^* > 0$, then according to condition (1.5), the corresponding jth inequality constraint must be violated arbitrarily close to x^*, implying that $g_j(x^*) = 0$. For this reason, the condition (1.5) is called the complementarity violation condition (CV for short) by Bertsekas and Ozdaglar [7].

Since the enhanced Fritz John condition is stronger than the classical Fritz John condition, it results in a stronger KKT condition under a weaker constraint qualification than the MFCQ. The enhanced Fritz John condition has been further extended to the case of smooth problem data with a convex abstract set constraint in Bertsekas [5] and with nonconvex set in Bertsekas and Ozdaglar [7] and Bertsekas, Nedić and Ozdaglar [6].

The first result on the enhanced Fritz John condition for nonsmooth problems with no abstract set constraint can be found in Bector, Chandra and Dutta [4] where the classical gradient is replaced by the Clarke subdifferential. Duality results for convex problems in terms of the enhanced Fritz John condition have also been studied by Bertsekas, Ozdaglar and Tseng in [9].

Moreover, if we denote

$$F(x) := \begin{pmatrix} g(x) \\ h(x) \\ \Psi(x) \end{pmatrix}, \quad \Lambda := \mathbb{R}^p_- \times \{0\}^q \times C^m, \tag{1.6}$$

where $\mathbb{R}_-$ denotes the nonpositive orthant $\{v \in \mathbb{R} \mid v \leq 0\}$ and

$$\Psi(x) := \begin{pmatrix} G_1(x) \\ H_1(x) \\ \vdots \\ G_m(x) \\ H_m(x) \end{pmatrix}, \quad C := \{\,(a,b) \in \mathbb{R}^2 \mid 0 \leq a \perp b \geq 0\}, \tag{1.7}$$

problem $\text{MPGC}_{\mathbb{R}^n}$ results in the mathematical program with equilibrium constraints formulated as follows:

$$\begin{array}{lll} \text{(MPEC)} & \min\limits_{x\in\mathcal{X}} & f(x) \\ & \text{s.t.} & h_i(x) = 0 \;\; i = 1,\dots,p, \quad g_j(x) \leq 0 \;\; j = 1,\dots,q, \\ & & G_l(x) \geq 0, H_l(x) \geq 0, G_l(x)H_l(x) = 0 \quad \forall l = 1,\dots,m. \end{array}$$

MPECs form a class of very important problems, since they arise frequently in applications; see [18, 49, 68]. MPECs are known to be a difficult class of optimization problems due to the fact that usual constraint qualifications, such as the LICQ and the MFCQ, are violated at any feasible point (see [87, Proposition 1.1]). Thus, the classical KKT condition is not always a necessary optimality condition for a MPEC. Alternatively, one can therefore use the Fritz John approach to derive necessary optimality conditions since these conditions do not require any constraint qualifications. However, it should also be noted that the standard Fritz John conditions applied to MPECs do not give much information regarding the signs of the Lagrange multipliers. Recently, Kanzow and Schwartz [42] studied the enhanced version of the Fritz John conditions.

1.2 Main contributions

The purpose of the thesis is mainly to develop enhanced stationarity conditions and introduce new constraint qualifications for nonsmooth optimization problems, including NLP, MPEC and MPGC. We may divide the thesis into two parts: The first part includes chapters 2-4 in which we study the enhanced optimality conditions and associated constraint qualifications, and the second part consists of chapter 5 in which we investigate some new constraint qualifications introduced in the recent literature. The chapter-to-chapter description of the thesis follows:

Chapter 2 For nonsmooth NLP we first derive the enhanced Fritz John condition. We then derive the enhanced KKT condition and introduce the associated pseudonormality and quasinormality condition. We prove that either pseudonormality or quasinormality with regularity on the constraint functions and the set constraint implies the existence of a local error bound. Finally we give a tighter upper estimate for the Fréchet subdifferential and the limiting subdifferential of the value function in terms of quasinormal multipliers which is usually a smaller set than the set of classical normal multipliers.

Chapter 3 For the first time, we obtain the enhanced Fritz John necessary optimality conditions for a nonsmooth mathematical program with geometric constraints where $F(x)$ is a mapping from a Banach space to a finite dimensional space. The enhanced Fritz John condition allows us to obtain the enhanced KKT condition under the pseudonormality and the quasinormality conditions. We then prove that the quasinomality is a sufficient condition for the existence of local error bounds. Finally we obtain a tighter upper estimate for the subdifferentials of the value function of the perturbed problem in terms of the enhanced multipliers.

Chapter 4 We first show that, unlike the smooth case, the mathematical program with equilibrium constraints linear independent constraint qualification is not a constraint qualification for the strong stationary condition when the objective function is nonsmooth. We argue that the strong stationary condition is unlikely for a mathematical program with equilibrium constraints with a nonsmooth objective function to hold at a local minimizer. We then focus on the study of the enhanced version of the Mordukhovich stationary condition, which is a weaker optimality condition than the strong stationary condition. We introduce the MPEC Pseudonormality, the MPEC Quasinormality, and the MPEC Constant Positive Linear Dependence, and show that the enhanced Mordukhovich stationary condition holds under them. Moreover we study the relations between the constraint qualifications and some other widely used constraints constraint qualifications for the MPEC. We also prove that quasinormality with regularity implies the existence of a local error bound. Finally, we give upper estimates for the subdifferential of the value function in terms of the enhanced M- and C-multipliers respectively.

Chapter 5 We show that, the relaxed Mangasarian-Fromovitz constraint qualification (or, equivalently, the constant rank of the subspace component condition) implies the existence of local error bounds. We further extend the new result to the MPEC. In particular, we show that the MPEC relaxed (or enhanced relaxed) constant positive linear dependence condition implies the existence of local M-PEC error bounds.

1.3 Backgrounds on nonsmooth analysis

This section contains some background material on nonsmooth analysis and preliminary results which will be used later. We give only concise definitions and results that will be needed in this thesis. For more detailed information on the subject our references are Borwein and Lewis [11], Borwein and Zhu [12], Clarke [16], Clarke, Ledyaev, Stern and Wolenski [17], Loewen [46], Mordukhovich [61,62] and Rockafellar and Wets [74].

We first give the following notations that will be used throughout the thesis. We denote by $\mathbb{B}(x^*, \epsilon)$ the closed ball centered at x^* with radius ϵ and $\mathbb{B}$ the closed unit ball centered at 0. For a set $\mathcal{C}$, we denote by int $\mathcal{C}$, cl $\mathcal{C}$, conv $\mathcal{C}$ its interior, closure and convex hull respectively. We let $\mathrm{dist}_{\mathcal{C}}(x^*)$ denote the distance of x^* to set $\mathcal{C}$. For a function $g : \mathbb{R}^n \to \mathbb{R}$, we denote by $g^+(x) := \max\{0, g(x)\}$ and if it is vector-valued then the maximum is taken componentwise. For a cone $\mathcal{N}$, we denote by $\mathcal{N}^o$ its polar.

For a set-valued map $\Phi : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$, we denote by

$$\limsup_{x \to x_0} \Phi(x) := \left\{ \xi \in \mathbb{R}^n : \begin{array}{l} \exists \text{ sequences } x_k \to x_0, \xi_k \to \xi, \\ \text{with } \xi_k \in \Phi(x_k) \ \ \forall k = 1, 2, \ldots \end{array} \right\}$$

$$\liminf_{x \to x_0} \Phi(x) := \left\{ \xi \in \mathbb{R}^n : \begin{array}{l} \forall \text{ sequences } x_k \to x_0, \exists \xi_k \in \Phi(x_k) \ \forall k = 1, 2, \ldots \\ \text{such that } \xi_k \to \xi \end{array} \right\},$$

the Painlevé-Kuratowski upper (outer) and lower (inner) limit respectively.

Definition 1 (Subdifferentials). *Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous (l.s.c.) function and $x_0 \in dom f := \{x \in \mathbb{R}^n : f(x) < +\infty\}$. The proximal subdifferential of f at x_0 is the set*

$$\partial^\pi f(x_0) := \left\{ \xi \in \mathbb{R}^n : \begin{array}{l} \exists \sigma > 0, \eta > 0 \ s.t. \\ f(x) \geq f(x_0) + \langle \xi, x - x_0 \rangle - \sigma \|x - x_0\|^2 \quad \forall x \in \mathcal{B}_\delta(x_0) \end{array} \right\}.$$

The Fréchet (regular) subdifferential of f at x_0 is the set

$$\hat{\partial} f(x_0) := \left\{ \xi \in \mathbb{R}^n : \liminf_{h \to 0} \frac{f(x_0 + h) - f(x_0) - \langle \xi, h \rangle}{\| h \|} \geq 0 \right\}.$$

The limiting (Mordukhovich or basic) subdifferential of f at x_0 is the set

$$\begin{aligned} \partial f(x_0) &:= \left\{ \xi \in \mathbb{R}^n : \exists x_k \to x_0, \ \textit{and} \ \xi_k \to \xi \ \textit{with} \ \xi_k \in \hat{\partial} f(x_k) \right\} \\ &= \{ \xi \in \mathbb{R}^n : \exists x_k \to x_0, \ \textit{and} \ \xi_k \to \xi \ \textit{with} \ \xi_k \in \partial^\pi f(x_k) \}. \end{aligned}$$

The singular limiting (Mordukhovich) subdifferential of f at x_0 is the set

$$\begin{aligned} \partial^\infty f(x_0) &:= \left\{ \xi \in \mathbb{R}^n : \exists x_k \to x_0, \ \textit{and} \ t_k \xi_k \to \xi \ \textit{with} \ \xi_k \in \hat{\partial} f(x_k), t_k \downarrow 0 \right\} \\ &= \{ \xi \in \mathbb{R}^n : \exists x_k \to x_0, \ \textit{and} \ t_k \xi_k \to \xi \ \textit{with} \ \xi_k \in \partial^\pi f(x_k), t_k \downarrow 0 \}. \end{aligned}$$

Let $f : \mathbb{R}^n \to \mathbb{R}$ be Lipschitz near x_0. The Clarke subdifferential (generalized gradient) of f at x_0 is the set

$$\partial^c f(x_0) = clconv \partial f(x_0).$$

In general we have the following inclusions, which may be strict:

$$\partial^\pi f(x_0) \subseteq \hat{\partial} f(x_0) \subseteq \partial f(x_0) \subseteq \partial^c f(x_0).$$

In the case where f is a convex function, all subdifferentials coincide with the subdifferential in the sense of convex analysis, i.e.,

$$\partial^{\pi} f(x_0) = \hat{\partial} f(x_0) = \partial f(x_0) = \partial^c f(x_0) = \{\xi : f(x) - f(x_0) \geq \langle \xi, x - x_0 \rangle \quad \forall x\}.$$

When f is strictly differentiable (see the definition, e.g. in Clarke [16]), $\partial f(x_0) = \partial^c f(x_0) = \{\nabla f(x_0)\}$. A l.s.c. function f is said to be subdifferentially regular ([61, Definition 1.91]) at x_0 if $\partial f(x_0) = \hat{\partial} f(x_0)$. It is known that for a locally Lipschitz continuous function, the subdifferential regularity is the same as the Clarke regularity (see [16, Definition 2.3.4] for the definition).

The following facts about the subdifferentials are well-known.

Proposition 1.3.1. *(i) A function $f : \mathbb{R}^n \to \mathbb{R}$ is Lipschitz near x_0 and $\partial f(x_0) = \{\zeta\}$ if and only if f is strictly differentiable at x_0 and the gradient of f at x_0 is equal to ζ.*

(ii) If a function $f : \mathbb{R}^n \to \mathbb{R}$ is Lipschitz near x_0 with positive constant L_f, then $\partial f(x_0) \subseteq L_f \mathbb{B}$.

(iii) A l.s.c. function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is Lipschitz near x_0 if and only if $\partial^{\infty} f(x_0) = \{0\}$.

(iv) Let $a \in \mathbb{R}$. Then

$$\partial \max\{0, a\} = \begin{cases} \{0\} & a < 0 \\ [0,1] & a = 0 \\ \{1\} & a > 0 \end{cases},$$

$$\partial |a| = \begin{cases} \{-1\} & a < 0 \\ [-1,1] & a = 0 \\ \{1\} & a > 0 \end{cases}.$$

Definition 2 (Proximal subdifferentiability)**.** *Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a l.s.c. function and $x_0 \in dom f$. We say that f is proximal subdifferentiable at x_0 if $\partial^\pi f(x_0) \neq \emptyset$.*

Proposition 1.3.2 (The Density Theorem)**.** *([17, Theorem 3.1]) Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a l.s.c. function. Then the set of points $x_0 \in dom f$ such that $\partial^\pi f(x_0) \neq \emptyset$ is dense in $dom f$.*

Definition 3 (Normal cones)**.** *Let Ω be a nonempty subset of $\mathbb{R}^n$ and $x_0 \in cl\Omega$. The convex cone*

$$\mathcal{N}^\pi_\Omega(x_0) := \left\{\xi \in \mathbb{R}^n : \exists \sigma > 0 \ s.t.\ \langle \xi, x - x_0\rangle \leq \sigma \|x - x_0\|^2 \quad \forall x \in \Omega\right\}$$

is called the proximal normal cone to Ω at x_0. The convex cone

$$\widehat{\mathcal{N}}_\Omega(x_0) := \left\{\xi \in \mathbb{R}^n : \limsup_{x \to x_0, x \in \Omega} \frac{\langle \xi, x - x_0\rangle}{\|x - x_0\|} \leq 0\right\}$$

is called the Fréchet (regular) normal cone to Ω at x_0. The nonempty cone

$$\mathcal{N}_\Omega(x_0) := \limsup_{x \to x_0} \widehat{\mathcal{N}}_\Omega(x_0) = \limsup_{x \to x_0} \mathcal{N}^\pi_\Omega(x_0)$$

is called the limiting (Mordukhovich or basic) normal cone to Ω at x_0. The Clarke normal cone is the closure of the convex hull of the limiting normal cone, i.e.,

$$\mathcal{N}^c_\Omega(x_0) = clconv\mathcal{N}_\Omega(x_0).$$

In general we have the following inclusions, which may be strict:

$$\mathcal{N}^\pi_\Omega(x_0) \subseteq \widehat{\mathcal{N}}_\Omega(x_0) \subseteq \mathcal{N}_\Omega(x_0) \subseteq \mathcal{N}^c_\Omega(x_0).$$

We say a set Ω is regular if $\widehat{\mathcal{N}}_\Omega(x) = \mathcal{N}_\Omega(x)$ for all $x \in \Omega$.

Lemma 1.1. [74, Theorem 6.11] *Let Ω be a nonempty subset of $\mathbb{R}^n$ and $x_0 \in cl\Omega$. A vector $\xi \in \widehat{\mathcal{N}}_\Omega(x_0)$ if and only if there is a function φ which is smooth on $\mathbb{R}^n$ with $-\nabla\varphi(x_0) = \xi$ and its global minimum on Ω is achieved uniquely at x_0.*

Proposition 1.3.3 (Tangent-normal polarity)**.** *([74, Theorem 6.26, Theorem 6.28]) Let Ω be a nonempty subset of $\mathbb{R}^n$ and $x_0 \in cl\Omega$.*

$$\mathcal{N}_\Omega(x_0)^o = \liminf_{x \xrightarrow{\Omega} x_0} T_\Omega(x),$$

where $T_\Omega(x) := \limsup_{\tau\downarrow 0} \frac{\Omega - x}{\tau}$ denotes the contingent cone to Ω at x.

Proposition 1.3.4 (Calculus rules)**.** *(i) Let $f : \mathbb{R}^n \to \mathbb{R}$ be Lipschitz near x_0 and $g : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be l.s.c. and finite at x_0. Let α, β be nonnegative scalars. Then*

$$\partial(\alpha f + \beta g)(x_0) \subseteq \alpha\partial f(x_0) + \beta\partial g(x_0).$$

(ii) [65, Corollary 3.4] Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be l.s.c. near x_0 and $g : \mathbb{R}^n \to \mathbb{R}$ be Lipschitz near x_0. Assume that $\hat{\partial} g(x_0) \neq \emptyset$ for all x near x_0. Then

$$\partial(f - g)(x_0) \subseteq \partial f(x_0) - \partial g(x_0).$$

(iii) Let $\varphi : \mathbb{R}^m \to \mathbb{R}^n$ be Lipschitz near x_0 and $f : \mathbb{R}^n \to \mathbb{R}$ be Lipschitz near $\varphi(x_0)$. Then

$$\partial(f \circ \varphi)(x_0) \subseteq \cup_{\xi \in \partial f(\varphi(x_0))} \partial\langle \xi, \varphi\rangle(x_0).$$

(iv) Let $f : \mathbb{R}^n \to \mathbb{R}$ be Lipschitz near x^* and $\mathcal{C}$ be a closed subset of $\mathbb{R}^n$. If x^* is a local minimizer of f on $\mathcal{C}$, then $0 \in \partial f(x^*) + \mathcal{N}_\mathcal{C}(x^*)$.

(v) Let $f : \mathbb{R}^n \to \mathbb{R}$ be Fréchet differentiable at x^* and $\mathcal{C}$ be a closed subset of $\mathbb{R}^n$. If x^* is a local minimizer of f on $\mathcal{C}$, then $0 \in \nabla f(x^*) + \widehat{\mathcal{N}}_\mathcal{C}(x^*)$.

Chapter 2

Enhanced Karush-Kuhn-Tucker condition and weaker constraint qualifications

2.1 Introduction

In this chapter we focus on the NLP problem (1.2). Unless otherwise indicated we assume throughout this chapter that $f, h_i(i = 1, \dots, p), g_j(j = 1, \dots, q) : \mathbb{R}^n \to \mathbb{R}$ are Lipschitz continuous around the point of interest and $\mathcal{X}$ is a nonempty closed subset of $\mathbb{R}^n$.

2.1.1 Motivation and contribution

One of the main results of this chapter is an improved version of the enhanced Fritz John condition for problem (NLP) with Lipschitz problem data based on the limiting subdifferential and limiting normal cone. Even in the case of a smooth problem,

---

This chapter is the content of Ye, J.J. and Zhang, J., "Enhanced Karush-Kuhn-Tucker condition and weaker constraint qualifications". Math. Program., series B., (2013). **139**, 353-381.

our improved enhanced Fritz John condition provides some new information. In our improved CV, we have an extra condition that the sequence $\{x^k\}$ can be found such that the functions $f, h_i(i \in I), g_j(j \in J)$ are proximal subdifferentiable at x^k (see Definition 2). Note that our improved CV is stronger than the original CV for the smooth problem since a continuously differentiable function may not be proximal subdifferentiable (a sufficient condition for a function to be proximal subdifferentiable is C^{1+}, i.e. the gradient of the function is locally Lipschitz).

Based on the enhanced Fritz John condition, Bertsekas and Ozdaglar [7] introduced the so-called pseudonormality and quasinormality as constraint qualifications that are weaker than the MFCQ. Since our improved enhanced Fritz John condition is stronger than the original enhanced Fritz John condition even in the smooth case, our pseudonormality and quasinormality conditions are even weaker than the original pseudonormality and quasinormality respectively and are much weaker than the NNAMCQ (which is in general weaker than the MFCQ in the nonsmooth case).

In recent years, it has been shown that constraint qualifications have strong connections with certain Lipschitz-like property of the set-valued map $\mathcal{F} : \mathbb{R}^{p+q} \rightrightarrows \mathbb{R}^m$ defined by the perturbed feasible region

$$\mathcal{F}(\alpha, \beta) := \{x \in \mathcal{X} : h(x) = \alpha, g(x) \leq \beta\},$$

where $h := (h_1, \ldots, h_p), g := (g_1, \ldots, g_q)$. For the case of a smooth optimization problem with $\mathcal{X} = \mathbb{R}^n$, by Mordukhovich's criteria for pseudo-Lipschitz continuity ([60,61]), MFCQ (or equivalently NNAMCQ) at a feasible point x^* is equivalent to the pseudo-Lipschitz continuity (or so-called Aubin continuity) of the set-valued map $\mathcal{F}(\alpha, \beta)$ around $(0, 0, x^*)$. Calmness of a set-valued map (introduced as the pseudo upper-Lipschitz continuity by Ye and Ye [85] and coined as calmness by Rockafellar and Wets [74]) is a much weaker condition than the pseudo-Lipschitz continuity. It is known that the calmness of the set-valued map $\mathcal{F}(\alpha, \beta)$ around $(0, 0, x^*)$ is equivalent to the existence of local error bound for the constraint region, i.e., the existence of positive constants c, δ such that

$$\text{dist}_{\mathcal{F}}(x) \leq c(\|h(x)\|_1 + \|g^+(x)\|_1) \quad \forall x \in \mathcal{B}_\delta(x^*) \cap \mathcal{X}. \tag{2.1}$$

In this chapter we show that either pseudonormality or quasinormality with regularity on the constraint functions and the set constraint implies that the set-valued map $\mathcal{F}(\alpha, \beta)$ is calm around the point $(0, 0, x^*)$. Hence pseudonormality and quasinormality are much weaker than the NNAMCQ.

NNAMCQ plays an important role in the sensitivity analysis. In particular it is a sufficient condition for the value function of a perturbed problem to be Lispchitz continuous (see e.g. [47, 48]). In this chapter we apply our improved enhanced KKT condition to derive an estimate for the Fréchet subdifferential and the limiting subdifferential of the value function. We provide a tighter upper estimate for the Fréchet subdifferential and the limiting subdifferentials of the value function in terms of the quasinormal multipliers. As a consequence we show that the value function is Lipschitz continuous under the perturbed quasinormality condition which is a much weaker condition than the NNAMCQ.

2.1.2 Scopes of the chapter

The rest of this chapter is organized as follows. In the next section, we derive the improved enhanced Fritz John condition. New constraint qualifications, the enhanced KKT and the relationship between pseudonormality and quasinormality are given in Section 2.3. Section 2.4 is devoted to the sufficient condition for the existence of local error bounds. In Section 2.5, the results is applied to the sensitivity analysis to provide a tighter upper estimate for the subdifferential of the value function.

2.2 Enhanced Fritz John necessary optimality condition

For nonsmooth problem (NLP), the classical Fritz John necessary optimality condition is generalized to one where the classical gradient is replaced by the generalized gradient by Clarke ([15], see also [16, Theorem 6.1.1]). The limiting subdifferential version of the Fritz John condition was first obtained by Mordukhovich in [59] (see also [78, Corollary 4.2] for more explicit expressions).

The following theorem strengthens the limiting subdifferential version of the Fritz John conditions by replacing the complementary slackness condition with a stronger condition [Theorem 2.1(iv)], and hence their effectiveness has been significantly enhanced. Although [Theorem 2.1(iv)] is slightly stronger than the complementarity violation condition of Bertsekas and Ozdaglar [7], for convenience we still refer to it as the complementarity violation condition (CV).

Theorem 2.1. *Let x^* be a local minimum of problem (NLP). Then there exist scalars $\mu_0^*, \lambda_1^*, \ldots, \lambda_p^*, \mu_1^*, \ldots, \mu_q^*$, satisfying the following conditions:*

(i) $0 \in \mu_0^* \partial f(x^*) + \sum_{i=1}^{p} \partial(\lambda_i^* h_i)(x^*) + \sum_{j=1}^{q} \mu_j^* \partial g_j(x^*) + \mathcal{N}_{\mathcal{X}}(x^*)$.

(ii) $\mu_j^* \geq 0, \qquad$ *for all* $j = 0, 1, \ldots, q$.

(iii) $\mu_0^*, \lambda_1^*, \ldots, \lambda_p^*, \mu_1^*, \ldots, \mu_q^*$ *are not all equal to 0.*

(iv) *The complementarity violation condition holds: If the index set $I \cup J$ is nonempty, where*

$$I = \{i | \lambda_i^* \neq 0\}, \qquad J = \{j \neq 0 | \mu_j^* > 0\},$$

then there exists a sequence $\{x^k\} \subseteq \mathcal{X}$ *converging to* x^* *such that for all* k*,*

$$f(x^k) < f(x^*), \qquad \lambda_i^* h_i(x^k) > 0,\ \forall i \in I, \qquad \mu_j^* g_j(x^k) > 0,\ \forall j \in J,$$

and $f, h_i (i \in I)$*,* $g_j (j \in J)$ *are all proximal subdifferentiable at* x^k*.*

Proof. Similar to the differentiable case in Bertsekas and Ozdaglar [7], we use a quadratic penalty function approach originated with McShane [52] to prove the result. For each $k = 1, 2, \ldots,$ we consider the penalized problem

$$\begin{array}{lll} (P_k) \qquad & \min & F^k(x) = f(x) + \frac{k}{2}\sum_{i=1}^{p}(h_i(x))^2 + \frac{k}{2}\sum_{j=1}^{q}(g_j^+(x))^2 + \frac{1}{2}\|x - x^*\|^2 \\ & \text{s.t.} & x \in \mathcal{X} \cap \mathbb{B}(x^*, \epsilon), \end{array}$$

where $\epsilon > 0$ is such that $f(x^*) \leq f(x)$ for all feasible x with $x \in \mathbb{B}(x^*, \epsilon)$. Since $\mathcal{X} \cap \mathbb{B}(x^*, \epsilon)$ is compact, by the Weierstrass theorem, an optimal solution $\mathbb{x}^k$ of the problem (P_k) exists. Consequently

$$\begin{aligned} f(\mathbb{x}^k) + \frac{k}{2}\sum_{i=1}^{p}(h_i(\mathbb{x}^k))^2 + \frac{k}{2}\sum_{j=1}^{q}(g_j^+(\mathbb{x}^k))^2 + \frac{1}{2}\|\mathbb{x}^k - x^*\|^2 = F^k(\mathbb{x}^k) \\ \leq F^k(x^*) = f(x^*). \end{aligned} \tag{2.2}$$

Since $f(\mathbb{x}^k)$ is bounded over $x \in \mathcal{X} \cap \mathbb{B}(x^*, \epsilon)$, we obtain from (2.2) that

$$\lim_{k\to\infty} |h_i(\mathbb{x}^k)| = 0, \qquad i = 1, \ldots, p,$$

$$\lim_{k\to\infty} |g_j^+(\mathbb{x}^k)| = 0, \qquad j = 1, \ldots, q$$

and hence every limit point $\bar{x}$ of $\{\mathbb{x}^k\}$ is feasible; i.e., $\bar{x} \in \mathcal{F}$. Furthermore, (2.2) yields

$$f(\mathbb{x}^k) + \frac{1}{2}\|\mathbb{x}^k - x^*\|^2 \le f(x^*), \qquad \forall k. \tag{2.3}$$

So by taking limit as $k \to \infty$, we obtain

$$f(\bar{x}) + \frac{1}{2}\|\bar{x} - x^*\|^2 \le f(x^*).$$

Since $\bar{x} \in \mathbb{B}(x^*, \epsilon)$ and $\bar{x}$ is feasible, we have $f(x^*) \le f(\bar{x})$, which combined with the preceding inequality yields $\|\bar{x} - x^*\| = 0$ so that $\bar{x} = x^*$. Thus, the sequence $\{\mathbb{x}_k\}$ converges to x^*, and it follows that $\mathbb{x}^k$ is an interior point of the closed ball $\mathbb{B}(x^*, \epsilon)$ for all k greater than some $\bar{k}$.

For $k > \bar{k}$, since $\mathbb{x}^k$ is an optimal solution of (P_k) and $\mathbb{x}^k$ is an interior point of the closed ball $\mathbb{B}(x^*, \epsilon)$, we have by the necessary optimality condition in terms of limiting subdifferential in Proposition 1.3.4 (iv) that

$$0 \in \partial F^k(\mathbb{x}^k) + \mathcal{N}_{\mathcal{X}}(\mathbb{x}^k).$$

Applying the calculus rules in Proposition 1.3.4 (i),(iii) to $\partial F^k(\mathbb{x}^k)$ we have the existence of multipliers

$$\xi_i^k := k h_i(\mathbb{x}^k), \qquad \zeta_j^k := k g_j^+(\mathbb{x}^k) \tag{2.4}$$

such that

$$0 \in \partial f(\mathbb{x}^k) + \sum_{i=1}^{p} \partial(\xi_i^k h_i)(\mathbb{x}^k) + \sum_{j=1}^{q} \zeta_j^k \partial g_j(\mathbb{x}^k) + (\mathbb{x}^k - x^*) + \mathcal{N}_{\mathcal{X}}(\mathbb{x}^k). \tag{2.5}$$

Denote by

$$\delta^k := \sqrt{1 + \sum_{i=1}^{p} (\xi_i^k)^2 + \sum_{j=1}^{q} (\zeta_j^k)^2},$$

$$\mu_0^k := \frac{1}{\delta^k}, \qquad \lambda_i^k := \frac{\xi_i^k}{\delta^k}, \quad i = 1, \ldots, p, \qquad \mu_j^k := \frac{\zeta_j^k}{\delta^k}, \quad j = 1, \ldots, q. \tag{2.6}$$

Then since $\delta^k > 0$, dividing (3.13) by δ^k, we obtain for all $k > \bar{k}$,

$$\begin{aligned} 0 \in \mu_0^k \partial f(\mathbb{x}^k) \;&+\; \sum_{i=1}^{p} \partial(\lambda_i^k h_i)(\mathbb{x}^k) + \sum_{j=1}^{q} \mu_j^k \partial g_j(\mathbb{x}^k) + \frac{1}{\delta^k}(\mathbb{x}^k - x^*) \\ &+\; \mathcal{N}_{\mathcal{X}}(\mathbb{x}^k). \end{aligned} \tag{2.7}$$

Since by construction we have

$$(\mu_0^k)^2 + \sum_{i=1}^{p} (\lambda_i^k)^2 + \sum_{j=1}^{q} (\mu_j^k)^2 = 1 \tag{2.8}$$

the sequence $\{\mu_0^k, \lambda_1^k, \ldots, \lambda_p^k, \mu_1^k, \ldots, \mu_q^k\}$ is bounded and must contain a subsequence that converges to some limit $\{\mu_0^*, \lambda_1^*, \ldots, \lambda_p^*, \mu_1^*, \ldots, \mu_q^*\}$.

Since h_i is Lipschitz near x^*, we have

$$\begin{aligned} \partial(\lambda_i^k h_i)(\mathbb{x}^k) \;&\subseteq\; \partial[(\lambda_i^k - \lambda_i^*) h_i](\mathbb{x}^k) + \partial(\lambda_i^* h_i)(\mathbb{x}^k) \quad \text{by Proposition 1.3.4 (i)} \\ &\subseteq\; L_{h_i} |\lambda_i^k - \lambda_i^*| \mathbb{B} + \partial(\lambda_i^* h_i)(\mathbb{x}^k) \quad \text{by Proposition 1.3.1 (ii)}, \end{aligned}$$

where L_{h_i} is the Lipschitz constant of h_i. Similarly,

$$\begin{aligned} \mu_0^k \partial f(\mathbb{x}^k) \;&\subseteq\; L_f |\mu_0^k - \mu_0^*| \mathbb{B} + \mu_0^* \partial f(\mathbb{x}^k), \\ \mu_j^k \partial g_j(\mathbb{x}^k) \;&\subseteq\; L_{g_j} |\mu_j^k - \mu_j^*| \mathbb{B} + \mu_j^* \partial g_j(\mathbb{x}^k), \end{aligned}$$

where L_f, L_{g_j} are the Lipschitz constants of f, g_j. Hence we have from (3.14) that

$$0 \in \mu_0^* \partial f(\mathbb{x}^k) + \sum_{i=1}^{p} \partial(\lambda_i^* h_i)(\mathbb{x}^k) + \sum_{j=1}^{q} \mu_j^* \partial g_j(\mathbb{x}^k) + \frac{1}{\delta^k}(\mathbb{x}^k - x^*)$$
$$+(L_f|\mu_0^k - \mu_0^*| + \sum_{i=1}^{p} L_{h_i}|\lambda_i^k - \lambda_i^*| + \sum_{j=1}^{q} L_{g_j}|\mu_j^k - \mu_j^*|)\mathbb{B} + \mathcal{N}_{\mathcal{X}}(\mathbb{x}^k).$$

Taking limit as $k \to \infty$, by the definition of the limiting subdifferential and the limiting normal cone (or the fact ∂f is outer semicontinuous [74, Proposition 8.7]), we see that μ_0^*, λ_i^* and μ_j^* must satisfy condition (i). From (2.4) and (2.6), μ_0^* and μ_j^* must satisfy condition (ii) and from (2.8), μ_0^*, λ_i^* and μ_j^* must satisfy condition (iii).

Finally, to show that condition (iv) is satisfied, assume that $I \cup J$ is nonempty (otherwise there is nothing to prove). Since $\lambda_i^k \to \lambda_i^*$ as $k \to \infty$ and $\lambda_i^* \neq 0$ for $i \in I$, for sufficiently large k, λ_i^k have the same sign as λ_i^*. Hence we must have $\lambda_i^* \lambda_i^k > 0$ for all $i \in I$ and sufficiently large k. Similarly $\mu_j^* \mu_j^k > 0$ for all $j \in J$ and sufficiently large k. Therefore from (2.4) and (2.6) we must have $\lambda_i^* h_i(\mathbb{x}^k) > 0$ for all $i \in I$ and $\mu_j^* g_j(\mathbb{x}^k) > 0$ for all $j \in J$ and $k \geq K_0$ for some positive integer K_0. Consequently since $I \cup J$ is nonempty, it follows that there exists either $i \in I$ such that $h_i(\mathbb{x}^k) \neq 0$ or $j \in J$ such that $g_j(\mathbb{x}^k) \neq 0$ for all $k \geq K_0$ and hence from (2.2) we have $f(\mathbb{x}^k) < f(x^*)$ for all $k \geq K_0$. It remains to show the proximal subdifferentiability of the functions $f, h_i(i \in I), g_j(j \in J)$ at $\mathbb{x}^k$. By the density theorem in Proposition 1.3.2, for each $\mathbb{x}^k$ with $k \geq K_0$, there exists a sequence $\{\mathbb{x}^{k,l}\} \subseteq \mathcal{X}$ with $\lim_{l \to \infty} \mathbb{x}^{k,l} = \mathbb{x}^k$ such that f, h_i, g_j are proximal subdifferentiable at $\mathbb{x}^{k,l}$. Since

$$f(\mathbb{x}^k) < f(x^*), \ \ \lambda_i^* h_i(\mathbb{x}^k) > 0, \forall i \in I, \quad \mu_j^* g_j(\mathbb{x}^k) > 0, \forall j \in J,$$

we have that and for all sufficiently large l,

$$f(\mathbb{x}^{k,l}) < f(x^*), \quad \lambda_i^* h_i(\mathbb{x}^{k,l}) > 0, \forall i \in I, \quad \mu_j^* g_j(\mathbb{x}^{k,l}) > 0, \forall j \in J.$$

For each $k \geq K_0$, choose an index l_k such that $l_1 < \ldots < l_{k-1} < l_k$ and

$$\lim_{k \to \infty} \mathbb{x}^{k,l_k} = x^*.$$

Consider the sequence $\{x^k\}$ defined by

$$x^k = \mathbb{x}^{(K_0+k),(l_{K_0+k})}, \qquad k = 1, 2, \ldots.$$

It follows from the preceding relations that $\{x^k\} \subseteq \mathcal{X}$,

$$\lim_{k \to \infty} x^k = x^*, \quad f(x^k) < f(x^*), \quad \lambda_i^* h_i(x^k) > 0, \forall i \in I, \quad \mu_j^* g_j(x^k) > 0, \forall j \in J,$$

and $f, h_i (i \in I)$, $g_j (j \in J)$ are all proximal subdifferentiable at x^k. $\square$

The condition (iv) is illustrated in Figure 2.1.

2.3 Enhanced KKT condition and weakened CQs

Based on the enhanced Fritz John condition, we define the following enhanced KKT condition.

Definition 4 (Enhanced KKT condition)**.** *Let x^* be a feasible point of the problem (NLP). We say the enhanced KKT condition holds at x^* if the enhanced Fritz John condition holds with $\mu_0^* = 1$.*

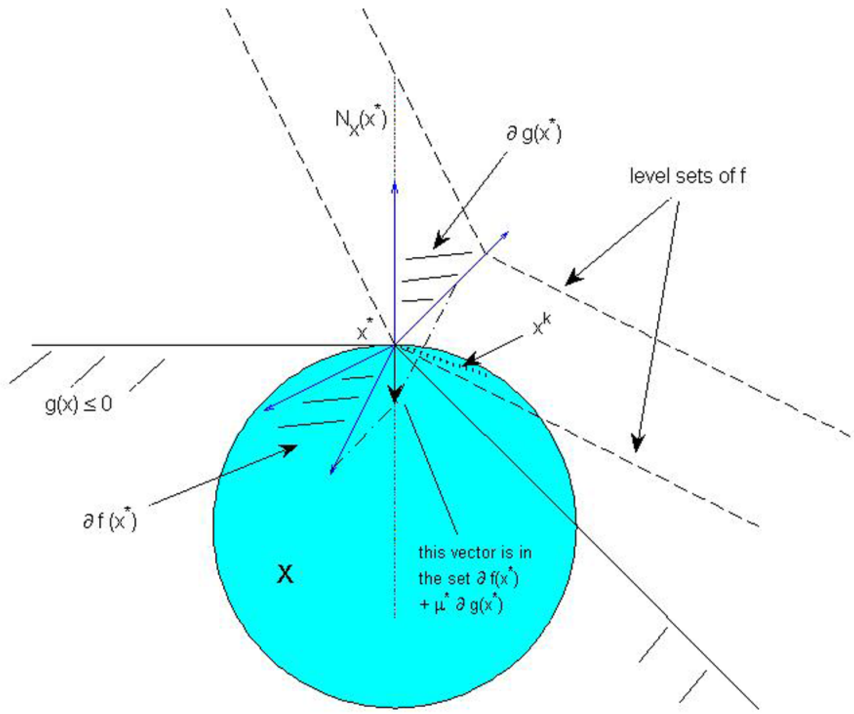

Figure 2.1: Existence of μ^* and $\{x^k\}$

Theorem 2.2. *Let x^* be a local minimum of problem (NLP). Suppose that there is no nonzero vector $(\lambda, \mu) \in \mathbb{R}^p \times \mathbb{R}^q_+$ such that*

$$0 \in \sum_{i=1}^{p} \partial(\lambda_i h_i)(x^*) + \sum_{j=1}^{q} \mu_j \partial g_j(x^*) + \mathcal{N}_{\mathcal{X}}(x^*), \tag{2.9}$$

and the CV condition defined in [Theorem 2.1(iv)] hold. Then the enhanced KKT condition holds at x^*.

Proof. Under the assumptions of the theorem, (i)-(iv) of Theorem 2.1 never hold if $\mu_0^* = 0$. Hence μ_0^* must be nonzero. The enhanced KKT condition then holds after a scaling. □

Note that the condition in Theorem 2.1 is not a constraint qualification since it involves the objective function f. However Theorem 2.2 leads to the introduction of some constraint qualifications for a weaker version of the enhanced KKT condition to hold. In the smooth case, the pseudonormality and the quasinormality are slightly weaker than the original definitions introduced by Bertsekas and Ozdaglar [7].

Definition 5. *Let x^* be in the feasible region $\mathcal{F}$.*

(a) *x^* is said to satisfy NNAMCQ if there is no nonzero vector $(\lambda, \mu) \in \mathbb{R}^p \times \mathbb{R}^q_+$ such that (2.9) and CS holds: $\mu_j g_j(x^*) = 0$ for all $j = 1, \ldots, q$.*

(b) *x^* is said to be pseudonormal (for the feasible region $\mathcal{F}$) if there is no vector $(\lambda, \mu) \in \mathbb{R}^p \times \mathbb{R}^q_+$ and no infeasible sequence $\{x^k\} \subseteq \mathcal{X}$ converging to x^* such that (2.9) and the pseudo-complementary slackness condition (pseudo-CS for short) hold: if the index set $I \cup J$ is nonempty, where $I = \{i|\lambda_i \neq 0\}, J = \{j|\mu_j > 0\}$, then for each k*
$$\sum_{i=1}^{p} \lambda_i h_i(x^k) + \sum_{j=1}^{q} \mu_j g_j(x^k) > 0,$$
and $h_i(i \in I)$, $g_j(j \in J)$ are all proximal subdifferentiable at x^k for each k.

(c) *x^* is said to be quasinormal (for the feasible region $\mathcal{F}$) if there is no nonzero vector $(\lambda, \mu) \in \mathbb{R}^p \times \mathbb{R}^q_+$ and no infeasible sequence $\{x^k\} \subseteq \mathcal{X}$ converging to x^* such that (2.9) and the quasi-complementary slackness condition (quasi-CS for short) hold: if the index set $I \cup J$ is nonempty, where $I = \{i|\lambda_i \neq 0\}, J = \{j|\mu_j > 0\}$, then for all $i \in I$, $j \in J$, $\lambda_i h_i(x^k) > 0$ and $\mu_j g_j(x^k) > 0$, and $h_i(i \in I)$, $g_j(j \in J)$ are all proximal subdifferentiable at x^k for each k.*

Since Quasi-CS $\Longrightarrow$ Pseudo-CS $\Longrightarrow$ CS, the following implications hold:

$$NNAMCQ \Longrightarrow Pseudonormality \Longrightarrow Quasinormality.$$

The first reverse implication is obviously not true. [7, Example 3] shows that the second reverse implication is not true either. We will show later that under the assumption that $\mathcal{N}_{\mathcal{X}}(x^*)$ is convex, quasinormality is in fact equivalent to a slightly weaker version of pseudonormality. In [7, Proposition 3.1] Bertsekas and Ozadaglar showed that any feasible point of a constraint region where the equality functions are linear and inequality functions are concave and smooth and there is no abstract constraint must be pseudonormal. In what follows we extend it to the nonsmooth case.

Proposition 2.3.1. *Suppose that h_i are linear and g_j are concave and $\mathcal{X} = \mathbb{R}^n$. Then any feasible point of problem (NLP) is pseudonormal.*

Proof. We prove it by contradiction. To the contrary, suppose that there is a feasible point x^* which is not pseudonormal. Then there exists nonzero vector $(\lambda, \mu) \in \mathbb{R}^p \times \mathbb{R}^q_+$ and a sequence $\{x^k\} \subseteq \mathcal{X}$ converging to x^* such that (2.9) and the following condition hold: for each k

$$\sum_{i=1}^{p} \lambda_i h_i(x^k) + \sum_{j=1}^{q} \mu_j g_j(x^k) > 0. \tag{2.10}$$

By the linearity of h_i and concavity of g_j, we have that for all $x \in \mathbb{R}^n$,

$$\begin{aligned} h_i(x) &= h_i(x^*) + \nabla h_i(x^*)^T (x - x^*) \quad i = 1, \ldots, p, \\ g_j(x) &\leq g_j(x^*) + \xi_j^T (x - x^*) \quad \forall \xi_j \in \partial g_j(x^*), j = 1, \ldots, q. \end{aligned}$$

By multiplying these two relations with λ_i and μ_j and by adding over i and j, respectively, we obtain that for all $x \in \mathbb{R}^n$ and all $\xi_j \in \partial g_j(x^*), j = 1, \ldots, q$,

$$\begin{aligned} &\sum_{i=1}^{p} \lambda_i h_i(x) + \sum_{j=1}^{q} \mu_j g_j(x) \\ &\leq \sum_{i=1}^{p} \lambda_i h_i(x^*) + \sum_{j=1}^{q} \mu_j g_j(x^*) + [\sum_{i=1}^{p} \lambda_i \nabla h_i(x^*) + \sum_{j=1}^{q} \mu_j \xi_j]^T (x - x^*) \\ &= [\sum_{i=1}^{p} \lambda_i \nabla h_i(x^*) + \sum_{j=1}^{q} \mu_j \xi_j]^T (x - x^*) \end{aligned}$$

where the last equality holds because we have

$$\lambda_i h_i(x^*) = 0 \text{ for all } i \text{ and } \sum_{j=1}^{q} \mu_j g_j(x^*) = 0.$$

By (2.9), since $\mathcal{N}_{\mathbb{R}^n}(x^*) = \{0\}$ there exists $\xi_j^* \in \partial g_j(x^*), j = 1, \ldots, q$ such that

$$\sum_{i=1}^{p} \lambda_i \nabla h_i(x^*) + \sum_{j=1}^{q} \mu_j \xi_j^* = 0.$$

Hence it follows that for all $x \in \mathbb{R}^n$,

$$\sum_{i=1}^{p} \lambda_i h_i(x) + \sum_{j=1}^{q} \mu_j g_j(x) \leq 0$$

which contradicts (4.4). Hence the proof is complete. □

Definition 6. *Let x^* be a feasible point of problem (NLP). We call a vector $(\lambda, \mu) \in \mathbb{R}^p \times \mathbb{R}^q_+$ satisfying the following weaker version of the enhanced KKT conditions a quasinormal multiplier:*

(i) $0 \in \partial f(x^*) + \sum_{i=1}^{p} \partial(\lambda_i^* h_i)(x^*) + \sum_{j=1}^{q} \mu_j^* \partial g_j(x^*) + \mathcal{N}_{\mathcal{X}}(x^*)$.

(ii) *There exists a sequence $\{x^k\} \subseteq \mathcal{X}$ converging to x^* such that the quasi-CS as defined in Definition 5 holds.*

Since the only difference of the quasinormality with the sufficient condition given in Theorem 2.2 is the condition $f(x^k) < f(x^*)$, it is obvious that the quasinormality is a constraint qualification for the weaker version of the enhanced KKT condition to hold and hence the following result follows immediately from Theorem 2.2 and the definitions of the three constraint qualifications.

Corollary 2.3. *Let x^* be a local minimizer of problem (NLP). Then if x^* satisfies*

NNAMCQ, or is pseudonormal, or is quasinormal, then the weaker version of the enhanced KKT condition holds at x^*.

It is known that NNAMCQ implies the boundedness of the set of all normal multipliers (see e.g. [40]). In what follows, we show that the set of all quasinormal multipliers are bounded under the quasinormality condition.

Theorem 2.4. *Let x^* be a feasible point for problem (NLP). If quasinormality holds at x^*, then the set of all quasinormal multipliers $M_Q(x^*)$ is bounded.*

Proof. To the contrary, suppose that $M_Q(x^*)$ is unbounded. Then there exists $(\lambda^n, \mu^n) \in M_Q(x^*)$ such that $\|(\lambda^n, \mu^n)\| \to \infty$ as n tends to infinity. By definition of a quasinormal multiplier, for each n, there exists a sequence $\{x_n^k\}_k \subseteq \mathcal{X}$ converging to x^* such that

$$0 \in \partial f(x^*) + \sum_{i=1}^{p} \partial(\lambda_i^n h_i)(x^*) + \sum_{j=1}^{q} \mu_j^n \partial g_j(x^*) + \mathcal{N}_{\mathcal{X}}(x^*), \tag{2.11}$$

$$\mu_j^n \geq 0, \quad \forall j = 1, \ldots, q, \tag{2.12}$$

$$\lambda_i^n h_i(x_n^k) > 0 \ \forall i \in I^n, \ \mu_j^n g_j(x_n^k) > 0 \ \forall j \in J^n, \tag{2.13}$$

$$h_i(i \in I^n), g_j(j \in J^n) \text{ are proximal subdifferential at } x_n^k, \tag{2.14}$$

where $I^n := \{i : \lambda_i^n \neq 0\}$ and $J^n := \{j : \mu_j^n > 0\}$.

Denote by $\xi^n := \frac{\lambda^n}{\|(\lambda^n, \mu^n)\|}$ and $\zeta^n := \frac{\mu^n}{\|(\lambda^n, \mu^n)\|}$. Assume without loss of generality that $(\xi^n, \mu^n) \to (\xi^*, \mu^*)$. Divide both sides of (2.11) by $\|(\lambda^n, \mu^n)\|$ and take the limit, we have

$$0 \in \sum_{i=1}^{p} \partial(\xi_i^* h_i)(x^*) + \sum_{j=1}^{q} \zeta_j^* \partial g_j(x^*) + \mathcal{N}_{\mathcal{X}}(x^*).$$

It follow from (2.12) that $\zeta_j^* \geq 0$, for all $j = 1, \ldots, q$. Finally, let

$$I = \{i : \xi_i^* \neq 0\}; \qquad J = \{j : \zeta_j^* > 0\}.$$

Then $I \cup J$ is nonempty. By virtue of (2.13), there are some N_0 such that for $n > N_0$, we must have $\xi_i^* h_i(x_n^k) > 0$ for all $i \in I$ and $\zeta_j^* g_j(x_n^k) > 0$ for all $j \in J$. Moreover by (2.14), $h_i(i \in I^n), g_j(j \in J^n)$ are proximal subdifferential at x_n^k. Thus there exist scalars $\{\xi_1^*, \ldots \xi_p^*, \zeta_1^*, \ldots, \zeta_q^*\}$ not all zero and a sequence $\{x_n^k\} \subseteq \mathcal{X}$ that satisfy the preceding relation an so violate the quasinormality of x^*. Hence the proof is complete.

□

Combining the proof techniques of Theorem 2.1 and [8, Lemma 2] in the following proposition we can extend [8, Lemma 2] to our nonsmooth problem.

Lemma 2.5. *If a vector $x^* \in \mathcal{F}$ is quasinormal, then all feasible vectors in a neighborhood of x^* are quasinormal.*

Proof. Assume that the claim is not true. Then we can find a sequence $\{x^k\} \subseteq \mathcal{F}$ such that $x^k \neq x^*$ for all k, $x^k \to x^*$ and x^* is not quasinormal for all k. This implies, for each k, the existence of scalars $\xi_1^k, \ldots, \xi_p^k, \zeta_1^k, \ldots, \zeta_q^k$ and a sequence $\{x^{k,l}\} \subseteq \mathcal{X}$ such that

(1) $0 \in \sum_{i=1}^p \partial(\xi_i^k h_i)(x^k) + \sum_{j=1}^q \zeta_j^k \partial g_j(x^k) + \mathcal{N}_{\mathcal{X}}(x^k)$,

(2) $\zeta_j^k \geq 0$, for all $j = 1, \ldots, q$,

(3) $\xi_1^k, \ldots \xi_p^k, \zeta_1^k, \ldots, \zeta_q^k$ are not all equal to 0,

(4) $\{x^{k,l}\}$ converges to x^k as $l \to \infty$, and for each l, $\xi_i^k h_i(x^{k,l}) > 0$ for all i with $\xi_i^k \neq 0$ and $\zeta_j g_j(x^{k,l}) > 0$ for all j with $\zeta_j^k > 0$, and for these i, j, h_i and g_j are Fréchet subdifferentiable at $x^{k,l}$.

For each k, denote

$$\delta^k = \sqrt{\sum_{i=1}^{p}(\xi_i^k)^2 + \sum_{j=1}^{q}(\zeta_j^k)^2}, \quad \lambda_i^k = \frac{\xi_i^k}{\delta^k}; 1 = 1, \dots, p; \quad \mu_j^k = \frac{\zeta_j^k}{\delta^k}, j = 1, \dots, q.$$

Since $\delta^k \neq 0$ and $\mathcal{N}_{\mathcal{X}}(x^k)$ is a cone, conditions (1) - (4) yields the following set of conditions that gold for each k for the scalars $\lambda_1^k, \dots, \lambda_p^k, \mu_1^k, \dots, \mu_q^k$:

(i)

$$0 \in \sum_{i=1}^{p} \partial(\lambda_i^k h_i)(x^k) + \sum_{j=1}^{q} \mu_j^k \partial g_j(x^k) + \mathcal{N}_{\mathcal{X}}(x^k), \tag{2.15}$$

(ii) $\mu_j^k \geq 0$, for all $j = 1, \dots, q$,

(iii) $\lambda_1^k, \dots \lambda_p^k, \mu_1^k, \dots, \mu_q^k$ are not all equal to 0,

(iv) $\{x^{k,l}\}$ converges to x^k as $l \to \infty$, and for each l, $\lambda_i^k h_i(x^{k,l}) > 0$ for all i with $\lambda_i^k \neq 0$ and $\mu_j g_j(x^{k,l}) > 0$ for all j with $\mu_j^k > 0$, and for these i, j, h_i and g_j are proximal subdifferentiable at $x^{k,l}$.

Since by construction we have

$$\sum_{i=1}^{p}(\lambda_i^k)^2 + \sum_{j=1}^{q}(\mu_j^k)^2 = 1, \tag{2.16}$$

the sequence $\{\lambda_1^k, \dots, \lambda_p^k, \mu_1^k, \dots, \mu_q^k\}$ is bounded and must contain a subsequence that converges to some nonzero limit $\{\lambda_1^*, \dots, \lambda_p^*, \mu_1^*, \dots, \mu_q^*\}$. Assume without loss of generality that $\{\lambda_1^k, \dots, \lambda_p^k, \mu_1^k, \dots, \mu_q^k\}$ converges to $\{\lambda_1^*, \dots, \lambda_p^*, \mu_1^*, \dots, \mu_q^*\}$. Taking the limit in (2.15), in analogy to Theorem 2.1, by [17, Theorem 3.8] and the closedness of normal cone, we see the limit must satisfy

$$0 \in \sum_{i=1}^{p} \partial(\lambda_i^* h_i)(x^*) + \sum_{j=1}^{q} \mu_j^* \partial g_j(x^*) + \mathcal{N}_{\mathcal{X}}(x^*).$$

Moreover, from conditions (ii)-(iii) and (2.16), it follows that $\mu_j^* \geq 0$, for all $j = 1, \ldots, q$, and $\lambda_1^*, \ldots, \lambda_p^*, \mu_1^*, \ldots, \mu_q^*$ are not all equal to 0. Finally, let

$$I = \{i | \lambda_i^* \neq 0\}; \qquad J = \{j | \mu_j^* > 0\}.$$

Then $I \cup J$ is nonempty and, it is easy to see there are some K_0 such that for $k > K_0$, we must have $\lambda_i^* \lambda_i^k > 0$ for all $i \in I$ and $\mu_j^* \mu_j^k > 0$ for all $j \in J$. From condition (iv) , it follows that for each $k > K_0$, there exists a sequence $\{x^{k,l}\} \subseteq \mathcal{X}$ with

$$\lim_{l \to \infty} x^{k,l} = x^k,$$

and for all l, $x^{k,l} \neq x^k$,

$$\lambda_i^* h_i(x^{k,l}) > 0, \forall i \in I, \qquad \mu_j^* g_j(x^{k,l}) > 0, \forall j \in J,$$

and for those index $i \in I$, $j \in J$, h_i, g_j are all proximal subdifferentiable at $x^{k,l}$. For each $k > K_0$, choose an index l_k such that $l_1 < \ldots < l_{k-1} < l_k$ and

$$\lim_{k \to \infty} x^{k,l_k} = x^*.$$

Consider the sequence $\{\varsigma^k\}$ defined by

$$\varsigma^k = x^{(K_0+k),(l_{K_0+k})}, \qquad k = 1, 2, \ldots.$$

It follows from the preceding relations that $\varsigma^k \subseteq \mathcal{X}$ and

$$\lim_{k\to\infty} \varsigma^k = x^*; \quad \lambda_i^* h_i(\varsigma^k) > 0, \forall i \in I; \qquad \mu_j^* g_j(\varsigma^k) > 0, \forall j \in J,$$

and for those index $i \in I$, $j \in J$, h_i, g_j are all Fréchet subdifferentiable at ς^k. The existence of scalars $\{\lambda_1^*, \ldots, \lambda_p^*, \mu_1^*, \ldots, \mu_q^*\}$ and sequence $\{\varsigma^k\} \subseteq \mathcal{X}$ satisfies the preceding relation violates the quasinormality of x^*, thus completing the proof. □

In the following result we obtain a specific representation of the limiting normal cone to the constraint region in terms of the set of quasinormal multipliers. Note that our result is sharper than the result of Bertsekas and Ozdaglar [8, Proposition 1] which gives a representation of the Fréchet normal cone in terms of the set of quasinormal multipliers for the case of smooth problems with a closed abstract set constraint. The result is also sharper than the one given by Henrion, Jourani and Outrata [29, Theorem 4.1] in which the representation is given in terms of the usual normal multipliers.

Proposition 2.3.2. *If $\bar{x}$ is quasinormal for $\mathcal{F}$, then*

$$\mathcal{N}_{\mathcal{F}}(\bar{x}) \subseteq \left\{ \sum_{i=1}^{p} \partial(\lambda_i h_i)(\bar{x}) + \sum_{j=1}^{q} \mu_j \partial g_j(\bar{x}) + \mathcal{N}_{\mathcal{X}}(\bar{x}) : (\lambda, \mu) \in M_Q(\bar{x}) \right\}.$$

Proof. Let v be a vector that belongs to $\mathcal{N}_{\mathcal{F}}(\bar{x})$. Then by definition, there are sequences $x^l \to \bar{x}$ and $v^l \to v$ with $v^l \in \widehat{\mathcal{N}}_{\mathcal{F}}(x^l)$ and $x^l \in \mathcal{F}$.

Step 1. By Lemma 2.5, for l sufficiently large, x^l is quasinormal for $\mathcal{F}$. By Lemma 1.1, for each l there exists a smooth function φ^l that achieves a strict global minimum over $\mathcal{F}$ at x^l with $-\nabla\varphi^l(x^l) = v^l$. Since x^l is a quasinormal vector of $\mathcal{F}$, by Theorem

2.2, the weaker version of the enhanced KKT condition holds for problem

$$\min \varphi^l(x) \quad \text{s.t. } x \in \mathcal{F}.$$

That is, there exists a vector $(\lambda^l, \mu^l) \in \mathbb{R}^p \times \mathbb{R}^q_+$ such that

$$v^l \in \sum_{i=1}^p \partial(\lambda_i^l h_i)(x^l) + \sum_{j=1}^q \mu_j^l \partial g_j(x^l) + \mathcal{N}_{\mathcal{X}}(x^l) \tag{2.17}$$

and a sequence $\{x^{l,k}\} \subseteq \mathcal{X}$ converging to x^l as $k \to \infty$ such that for all k, $\lambda_i^l h_i(x^{l,k}) > 0, \forall i \in I^l$, $\mu_j^l g_j(x^{l,k}) > 0, \forall j \in J^l$, and $h_i (i \in I^l)$, $g_j (j \in J^l)$ are proximal subdifferentiable at $x^{l,k}$, where $I^l = \{i : \lambda_i^l \neq 0\}$ and $J^l = \{j : \mu_j^l > 0\}$.

Step 2. We show that the sequence $\{\lambda_1^l, \ldots, \lambda_p^l, \mu_1^l, \ldots, \mu_q^l\}$ is bounded. To the contrary suppose that the sequence $\{\lambda_1^l, \ldots, \lambda_p^l, \mu_1^l, \ldots, \mu_q^l\}$ is unbounded. For every l, denote

$$\delta^l = \sqrt{1 + \sum_{i=1}^p (\lambda_i^l)^2 + \sum_{j=1}^q (\mu_j^l)^2}, \ \xi_i^l = \frac{\lambda_i^l}{\delta^l}, \ i = 1, \ldots, p, \ \zeta_j^l = \frac{\mu_j^l}{\delta^l}, \quad j = 1, \ldots, q.$$

Then from (4.5) it follows that

$$\frac{v^l}{\delta^l} \in \sum_{i=1}^p \partial(\xi_i^l h_i)(x^l) + \sum_{j=1}^q \zeta_j^l \partial g_j(x^l) + \mathcal{N}_{\mathcal{X}}(x^l).$$

Since the sequence $\{\xi_1^l, \ldots, \xi_p^l, \zeta_1^l, \ldots, \zeta_q^l\}$ is bounded, for the sake of simplicity, we may assume that $\{\xi_1^l, \ldots, \xi_p^l, \zeta_1^l, \ldots, \zeta_q^l\} \to \{\xi_1^*, \ldots, \xi_p^*, \zeta_1^*, \ldots, \zeta_q^*\} \neq 0$ as $l \to \infty$. Taking limits in the above inclusion, similar to the proof of Theorem 2.1 we obtain

$$0 \in \sum_{i=1}^p \partial(\xi_i^* h_i)(\bar{x}) + \sum_{j=1}^q \zeta_j^* \partial g_j(\bar{x}) + \mathcal{N}_{\mathcal{X}}(\bar{x}),$$

where $\zeta_j^* \geq 0$ for all $j = 1, \ldots, q$ and $\xi_1^*, \ldots, \xi_p^*, \zeta_1^*, \ldots, \zeta_q^*$ are not all zero. Let $i \in I^* := \{i : \xi_i^* \neq 0\}$. Since $\xi_i^l \to \xi_i^* \neq 0$ as $l \to \infty$, $\xi_i^l \neq 0$ and has the same sign as ξ_i^* for sufficiently large l. Consequently since $\xi_i^l h_i(x^{l,k}) > 0$ we have also $\xi_i^* h_i(x^{l,k}) > 0$ for all sufficiently large l and all k. Similarly let $j \in J^* := \{j : \zeta_j^* > 0\}$, we have $\zeta_j^* g_j(x^{l,k}) > 0$. Also similar to the proof of Theorem 2.1, by using the density theorem we can find a subsequence $\{x^{l,k_l}\} \subseteq \{x^{l,k}\} \subseteq \mathcal{X}$ converging to $\bar{x}$ as $l \to \infty$ such that for all sufficiently large l,

$$\xi_i^* h_i(x^{l,k_l}) > 0 \ \ \forall i \in I^*, \qquad \zeta_j^* g_i(x^{l,k_l}) > 0 \ \forall j \in J^*$$

and $h_i(x^{l,k_l})(i \in I^*), g_j(x^{l,k_l})(j \in J^*)$ are proximal subdifferentiable at x^{l,k_l}. But this is impossible since $\bar{x}$ is assumed to be quasinormal and hence the sequence $\{\lambda_1^l, \ldots, \lambda_p^l, \mu_1^l, \ldots, \mu_q^l\}$ must be bounded.

Step 3. By virtue of Step 2, without loss of generality, we assume that

$$\{\lambda_1^l, \ldots, \lambda_p^l, \mu_1^l, \ldots, \mu_q^l\} \text{ converges to } \{\lambda_1, \ldots, \lambda_p, \mu_1, \ldots, \mu_q\} \text{ as } l \to \infty.$$

Taking the limit in (4.5) as $l \to \infty$, we have

$$v \in \sum_{i=1}^{p} \partial(\lambda_i h_i)(\bar{x}) + \sum_{j=1}^{q} \mu_j \partial g_j(\bar{x}) + \mathcal{N}_{\mathcal{X}}(\bar{x}).$$

Similar to Step 2, we can find a subsequence $\{x^{l,k_l}\} \subseteq \{x^{l,k}\} \subseteq \mathcal{X}$ converging to $\bar{x}$ as $l \to \infty$ such that for all sufficiently large l, $\lambda_i h_i(x^{l,k_l}) > 0, \forall i \in I$, $\mu_j g_j(x^{l,k_l}) > 0, \forall j \in J$, and $h_i(i \in I)$, $g_j(j \in J)$ are proximal subdifferentiable at x^{l,k_l}, where $I = \{i : \lambda_i \neq 0\}$ and $J = \{j : \mu_j > 0\}$. $\square$

From Propositions 2.3.2 and calculus rule 1.3.4 (v), the following enhanced KKT necessary optimality condition for the case where the objective function is Fréchet differentiable (but may not be Lipschitz) follows immediately. Note that for a Fréchet differentiable function which is not Lipschitz continuous, the limiting subdifferential may not coincide with the usual gradient and hence the following result provides a sharper result for this case.

Corollary 2.6. *Let x^* be a local minimizer of problem (NLP) where the objective function f is Fréchet differentiable at x^*. If x^* either satisfies NNAMCQ, is pseudonormal, or is quasinormal, then the weaker version of the enhanced KKT condition holds.*

We close this section with a result showing that quasinormality and a weaker version of pseudonormality coincide under the condition that the normal cone is convex and the constraint functions are strictly differentiable at the point x^*. This result is an extension of a similar result of Bertsekas and Ozdaglar [7, Proposition 3.2] in that we do not require the function to be continuously differentiable at x^*.

Proposition 2.3.3. *Let $x^* \in \mathcal{F}$. Assume that for each $i = 1, \ldots, p$, $j = 1, \ldots, q$, $h_i(x)$, $g_j(x)$ are strictly differentiable at x^*, and the limiting normal cone $\mathcal{N}_{\mathcal{X}}(x^*)$ is convex. Then x^* is quasinormal if and only if the following weaker version of pseudonormality holds: there are no vector $(\lambda, \mu) \in \mathbb{R}^p \times \mathbb{R}^q_+$ and no sequence $\{x^k\} \subseteq \mathcal{X}$ converging to x^* such that*

(i) $0 \in \sum_{i=1}^{p} \lambda_i \nabla h_i(x^*) + \sum_{j=1}^{q} \mu_j \nabla g_j(x^*) + \mathcal{N}_{\mathcal{X}}(x^*)$.

(ii) $\lambda_i h_i(x^k) \geq 0$ *for all* i *and* $\mu_j g_j(x^k) \geq 0$ *for all* j*, and if the index sets* $I \cup J \neq \emptyset$ *where* $I = \{i|\lambda_i \neq 0\}$ $J = \{j|\mu_j > 0\}$ *then*

$$\sum_{i=1}^{p} \lambda_i h_i(x^k) + \sum_{j=1}^{q} \mu_j g_j(x^k) > 0, \qquad \forall k$$

and $h_i (i \in I)$, $g_j (j \in J)$ *are proximal subdifferentiable at* x^k.

Proof. It is easy to see that the weaker version of pseudonormality implies the quasinormality. So what we have to do is to show the converse. To the contrary, suppose that the quasinormality holds but the weaker version of pseudonormality does not hold. Then there exist scalars $\lambda_1, \dots, \lambda_p, \mu_1, \dots, \mu_q$ and a sequence $\{x^k\} \subseteq \mathcal{X}$ converging to x^* such that (i)-(ii) hold. Condition (ii) implies that $\lambda_i h_i(x^k) > 0$ for some $\bar{i}$ such that $\lambda_{\bar{i}} \neq 0$ or $\mu_j g_j(x^k) > 0$ for some $\bar{j}$ such that $\mu_{\bar{j}} > 0$. We now suppose that such $\bar{j}$ exists (the case where $\bar{j}$ does not exist but $\bar{i}$ exists can be similarly proved and we omit it here). Without loss of generality, we can assume $\bar{j} = 1$ and $\mu_1 = 1$ (otherwise we can normalize it) such that (i) holds:

$$-\Big(\nabla g_1(x^*) + \sum_{i=1}^{p} \lambda_i \nabla h_i(x^*) + \sum_{j=2}^{q} \mu_j \nabla g_j(x^*)\Big) \in \mathcal{N}_{\mathcal{X}}(x^*). \tag{2.18}$$

Since $g_1(x^k) > 0$ for all k, $\mu_2, \dots, \mu_q, \lambda_1, \dots, \lambda_p$ are not all equal to 0, otherwise it would contradicts the quasinormality of x^*. Besides, because $\mathcal{N}_{\mathcal{X}}(x^*)$ is closed and convex, by [7, Lemma 2.2] there exists a vector $\bar{d} \in \mathcal{N}_{\mathcal{X}}(x^*)^*$ with $\langle \bar{d}, \nabla g_1(x^*)\rangle < 0$, $\langle \bar{d}, \nabla g_j(x^*)\rangle > 0$ for all $j = 2, \dots, q$, such that $\mu_j > 0$ and $\langle \bar{d}, \lambda_i \nabla h_i(x^*)\rangle > 0$ for all $i = 1, \dots, p$, such that $\lambda_i \neq 0$.

In the remaining part of the proof, we show that the scalars

$$\mu_1 = 1, \mu_2, \dots, \mu_q, \lambda_1, \dots, \lambda_p$$

achieved above satisfy condition: $\lambda_i h_i(x^k) > 0 \;\; \forall i \in I := \{i : \lambda_i \neq 0\}$, $\mu_j g_j(x^k) > 0 \;\; \forall j \in J := \{j = 2, \dots, q : \mu_j > 0\}$ which would contradicts the fact that x^* is quasinormal. Since g_j and h_i are strictly differentiable at x^*, the gradients coincide with the limiting subdifferentials, i.e.,

$$\begin{aligned}\nabla(\mu_j g_j)(x^*) &= \lim_{k\to\infty} \rho_j^k \text{ for some } \rho_j^k \in \partial^\pi(\mu_j g_j)(x^k)\\ \nabla(\lambda_i h_i)(x^*) &= \lim_{k\to\infty} \varrho_j^k \text{ for some } \varrho_i^k \in \partial^\pi(\lambda_i h_i)(x^k).\end{aligned}$$

By Proposition 1.3.3, for vector $\bar{d} \in \mathcal{N}_{\mathcal{X}}(x^*)^*$ and the sequence x^k converging to x^* constructed above, there is a sequence $d^k \in \mathcal{T}_{\mathcal{X}}(x^k)$ such that $d^k \to \bar{d}$. By virtue of $x^k \to x^*$, $d^k \to \bar{d}$ and $\langle \bar{d}, \nabla(\mu_j g_j)(x^*)\rangle > 0$ for all $j = 2, \ldots, q$, with $\mu_j > 0$, $\langle \bar{d}, \nabla(\lambda_i h_i)(x^*)\rangle > 0$ for all $i = 1, \ldots, p$, with $\lambda_i \neq 0$, we have that, for all sufficiently large k, $\langle d^k, \rho_j^k\rangle > 0$ for all $j = 2, \ldots, q$, with $\mu_j > 0$, $\langle d^k, \varrho_i^k\rangle > 0$ for all $i = 1, \ldots, p$, with $\lambda_i \neq 0$. Since $d^k \in \mathcal{T}_{\mathcal{X}}(x^k)$, there exists a sequence $\{x^{k,l}\} \in \mathcal{X}$ such that, for each k, we have $x^{k,l} \neq x^k$ for all l and

$$x^{k,l} \to x^k, \qquad \frac{x^{k,l} - x^k}{\|x^{k,l} - x^k\|} \to \frac{d^k}{\|d^k\|}, \quad as\ l \to \infty,$$

h_i, g_j are proximal subdifferentiable at $x^{k,l}$. Since $\rho_j^k \in \partial^\pi(\mu_j g_j)(x^k) \subseteq \hat{\partial}(\mu_j g_j)(x^k)$, by definition of the Fréchet subdifferential, for some vector sequence υ converging to 0, and for each $j = 2, \ldots, q$, with $\mu_j > 0$,

$$\begin{aligned}\mu_j g_j(x^{k,l}) &\geq \mu_j g_j(x^k) + \langle x^{k,l} - x^k, \rho_j^k\rangle + o(\|x^{k,l} - x^k\|)\\ &\geq \mu_j \langle \frac{d^k}{\|d^k\|} + \upsilon, \rho_j^k\rangle \|x^{k,l} - x^k\| + o(\|x^{k,l} - x^k\|)\end{aligned}$$

where the second inequality above follows from the assumption that $\mu_j g_j(x^k) \geq 0$, for all j and x^k. It follows that, for l and k sufficiently large, there exists $x^{k,l} \in \mathcal{X}$ arbitrary close to x^k such that $\mu_j g_j(x^{k,l}) > 0$ and, g_j are proximal subdifferentiable at $x^{k,l}$ for all $j = 2, \ldots, q$, with $\mu_j > 0$. Similarly, for l and k sufficiently large, there exists $x^{k,l} \in \mathcal{X}$ arbitrary close to x^k such that $\lambda_i h_i(x^{k,l}) > 0$ and, h_i are proximal subdifferentiable at $x^{k,l}$ for all $i = 1, \ldots, p$ with $\lambda_i \neq 0$. □

2.4 Sufficient conditions for error bounds

In this section we show that either pseudonormality or quasinormality plus the subdifferential regularity condition on constraints implies the existence of local error bounds. Our results are new even for the smooth case.

In order to derive the desired error bound formula (2.1), let us first rewrite the constraint region (1.2) equivalently as follows:

$$\mathcal{F} = \{x \in \mathcal{X} : \|h(x)\|_1 + \|g^+(x)\|_1 = 0\}. \tag{2.19}$$

By [79, Theorem 3.3], to prove the desired error bound result we only need to derive the following estimation.

Lemma 2.7. *Let x^* be feasible for problem (NLP) such that pseudonormality holds. Then there are $\delta, c > 0$ such that*

$$\frac{1}{c} \leq \|\xi\|_1 \qquad \forall \xi \in \partial^\pi(\|h(x)\|_1 + \|g^+(x)\|_1 + \delta_{\mathcal{X}}(x)), x \in \mathcal{B}_{\frac{\delta}{2}}(x^*) \cap \mathcal{X}, x \notin \mathcal{F},$$

where $\delta_{\mathcal{X}}(x)$ denotes the indicator function of the set $\mathcal{X}$ at x.

Proof. To the contrary, assume that there exists a sequence $\{x^k\} \to x^*$ with $x^k \in \mathcal{X} \setminus \mathcal{F}$ and $\xi^k \in \partial^\pi(\|h\|_1 + \|g^+\|_1 + \delta_{\mathcal{X}})(x^k)$ for all $k \in \mathbb{N}$ such that $\|\xi^k\|_1 \to 0$. By the calculus rule in Proposition 1.3.4 (i), (iii) and Proposition 1.3.1 (iv), we can find bounded multipliers (μ^k, λ^k) with $\mu^k \geq 0$ such that

$$\xi^k \in \sum_{i=1}^{q} \partial(\lambda_i^k h_i)(x^k) + \sum_{j=1}^{q} \mu_j^k \partial g_j(x^k) + \mathcal{N}_{\mathcal{X}}(x^k) \tag{2.20}$$

for all $k \in \mathbb{N}$. Hence, we may assume without loss of generality that it converges to a limit (λ, μ). Taking the limit as $k \to \infty$ in (3.12) yields

$$0 \in \sum_{i=1}^{p} \partial(\lambda_i h_i)(x^*) + \sum_{j=1}^{q} \mu_j \partial g_j(x^*) + \mathcal{N}_{\mathcal{X}}(x^*).$$

In addition, by the existence of λ_i^k, μ_j^k and Proposition 1.3.1 (iv), for k large enough, it is easy to see that

$$\lambda_i h_i(x^k) \geq 0 \qquad \forall i = 1, \ldots, p,$$
$$\mu_j g_j(x^k) \geq 0 \qquad \forall j = 1, \ldots, q.$$

Since $x^k \notin \mathcal{F}$ for all k, at least one functional constraint has to be violated infinitely many times. Using again Proposition 1.3.1 (iv), it is easy to see that there exists at least one multiplier λ_i or μ_j not equal to zero, and the corresponding product is strictly positive for all k such that the constraint is violated, i.e. if constraint $h_i(x^k) = 0$ is violated for infinitely many k, we may have $\lambda_i \neq 0$ and $\lambda_i h_i(x^k) > 0$ for all those k, if the constraint $g_j(x^k) \leq 0$ is violated for infinitely many k, we may have $\mu_j > 0$ and $\mu_j g_j(x^k) > 0$ for all those k. Therefore

$$\sum_{i=1}^{p} \lambda_i h_i(x^k) + \sum_{j=1}^{q} \mu_j g_j(x^k) > 0$$

at least on a subsequence. Moreover by the density theorem in Proposition 1.3.2 $h_i (i \in I), g_j (j \in J)$ can be selected to be proximal subdifferentiable at x^k. This, however, implies that pseudonormality is violated in x^* since x^k is chosen from $\mathcal{X}$, a contradiction. $\square$

Using the local error bound result of [79, Theorem 3.3], we obtain the following error bound result.

Theorem 2.8. *Let x^* be feasible for problem (NLP) such that pseudonormality holds. Then the local error bound holds: there exist positive constants c and δ such that*

$$\mathrm{dist}_{\mathcal{F}}(x) \leq c(\|h(x)\|_1 + \|g^+(x)\|_1) \quad \forall x \in \mathcal{B}_\delta(x^*) \cap \mathcal{X}.$$

By Clarke's exact penalty principle [16, Proposition 2.4.3] we obtain the following exact penalty result immediately.

Corollary 2.9. *Let x^* be a local minimizer of problem (NLP). If pseudonormality holds at x^*, then x^* is a local minimizer of the penalized problem:*

$$\begin{array}{ll} \min & f(x) + \alpha(\|h(x)\|_1 + \|g^+(x)\|_1) \\ s.t. & x \in \mathcal{X}, \end{array}$$

where $\alpha \geq L_f c$, L_f is the Lipschitz constant of f and c is the error bound constant.

Notice that Corollary 2.9 even works for nonstrict local minima x^* in the nonsmooth case. However, we find that the exact penalty result in [7, Proposition 4.2], established in the smooth case, requires x^* to be a strict local minimum, and it is stated in [7, Example 7.7] that this assumption might be crucial. The example is the following:

$$\min \; f(x_1, x_2) := x_2 \;\; \text{s.t.} \;\; h(x_1, x_2) := x_2/(x_1^2 + 1) = 0.$$

The feasible points are of the form

$$(x_1, 0) \quad \text{with } x_1 \in \mathbb{R}.$$

And each feasible point is a local minimum. Since the gradient $\nabla h(x_1, x_2)$ is nonzero, every feasible point is quasinormal. The authors claim that pseudonormality at a nonstrict local minimum may not imply the exact penalty since for any $c > 0$,

$$\inf_{(x_1,x_2)\in\mathbb{R}^2}\{x_2 + c|x_2|/|x_1^2+1|\} = -\infty,$$

which shows that a local optimal solution of the original problem is not a global optimal solution of the penalized problem. However this example is not a counter example since $(x_1, 0)$ is a local minimum of the function $x_2 + c|x_2|/|x_1^2+1|$ for large enough $c > 0$. Since the limiting subdifferential agrees with the classical gradient when a function is strictly differentiable, we stress, that not only did we extend the exact penalty result in [7] to a more general case, but we also improved their result in another way. We have now answered positively the open question raised in [42] in which the authors ask whether or not the proof technique based on error bound and the exact penalty principle of Clarke (which is completely different from the one used in [7]) can be used to prove the exact penalty result in [7] with a nonstrict local optimum.

The following example shows that the converse of Theorem 2.9 does not hold. Since when the objective function is Lipschitz continuous, the existence of an exact penalty function implies the exact penalty. It also shows that the existence of an exact penalty function does not imply pseudonormality.

Example 1. *Consider the locally Lipschitz optimization problem*

$$\begin{aligned} \min \quad & f(x) = |x_1| + |x_2| \\ s.t. \quad & g(x) = |x_1| - x_2 \le 0 \\ & x \in \mathcal{X} \end{aligned}$$

where $\mathcal{X}$ *is denoted as* $\{(x_1,x_2)|x_1^2 + (x_2+1)^2 \le 1\}$. *At the only feasible point*

$x^* = (0,0)$, $\partial g(x^*) = \{(\zeta, -1) | -1 \leq \zeta \leq 1\}$ and $\mathcal{N}_{\mathcal{X}}(x^*) = \{t(0,1) | t \geq 0\}$. However, if we choose $\mu = 1$ and a sequence $\{x^k\}$ located in $\mathcal{X}$ where for each $k = 1, 2, \ldots$, $x^k = (\cos(\frac{\pi}{2} - \frac{\pi}{2k}), -1 + \sin(\frac{\pi}{2} - \frac{\pi}{2k}))$, we have $0 \in \mu\partial g(x^*) + \mathcal{N}_{\mathcal{X}}(x^*)$ and $\mu g(x^k) > 0$ for all $k > 1$. This implies x^* is not pseudonormal. However it is easy to see that the error bound holds:

$$\text{dist}_{\mathcal{C}}(x) \leq |x_1| - x_2 = |x_1| + |x_2| \qquad \forall x \in \mathcal{X}$$

with $\mathcal{C} = \{0, 0)\}$ and $\mathcal{X} = \{(x_1, x_2) : x_1^2 + (x_2 + 1)^2 \leq 1\}$.

Naturally, after showing that pseudonormality implies the existence of local error bound, we would like to explore the relation between quasinormality and the error bound property. In [57, Theorem 2.1], under the assumption that the constraint functions are C^{1+}, Minchenko and Tarakanov show that quasinormality implies the error bound for a smooth optimization problem with $\mathcal{X} = \mathbb{R}^n$. In what follows, we will show that quasinormality implies the error bound property for our nonsmooth optimization problem (NLP) under the condition that the constraint functions are subdifferential regular and the abstract constraint set is regular. Since a smooth function must be subdifferentially regular, our results show that the condition of C^{1+} for the constraint functions in Minchenko and Tarakanov [57, Theorem 2.1] can be removed.

Theorem 2.10. *Assume in the constraint system (1.2) that $\mathcal{X}$ is a nonempty closed regular set. Further let $x^* \in \mathcal{F}$, assume $h_i(x)$ are continuously differentiable, $g_j(x)$ are subdifferentially regular around x^* (automatically holds when g_j are convex or C^1 around x^*). If x^* is a quasinormal point of $\mathcal{F}$, then there exist positive numbers c*

and δ*, such that*

$$\mathrm{dist}_{\mathcal{F}}(x) \leq c(\|h(x)\|_1 + \|g^+(x)\|_1) \quad \forall x \in \mathcal{B}_\delta(x^*) \cap \mathcal{X}. \tag{2.21}$$

Proof. By assumption we can find $\delta_0 > 0$ such that $\hat{\partial} h_i(x)$ is nonempty and $g_j(x)$ are subdifferentially regular for all $x \in \mathcal{B}_{\delta_0}(x^*)$. Since the required assertion is always true if $x^* \in int\mathcal{F}$, we only need to consider the case when $x^* \in \partial\mathcal{F}$. In this case, (2.21) can be violated only for $x \notin \mathcal{F}$. Let us take some sequences $\{\mathbb{x}^k\}$ and $\{x^k\}$, such that $\mathbb{x}^k \to x^*$,$\mathbb{x}^k \in \mathcal{X} \setminus \mathcal{F}$, and $x^k = \prod_{\mathcal{F}}(\mathbb{x}^k)$, the projection of $\mathbb{x}^k$ onto the set $\mathcal{C}$. Note that $x^k \to x^*$, since $\|x^k - \mathbb{x}^k\| \leq \|\mathbb{x}^k - x^*\|$. For simplicity we may assume both $\{\mathbb{x}^k\}$ and $\{x^k\}$ belong to $\mathcal{B}_{\delta_0}(x^*) \cap \mathcal{X}$.

Since $\mathbb{x}^k - x^k \in \mathcal{N}^{\pi}_{\mathcal{F}}(x^k) \subseteq \widehat{\mathcal{N}}_{\mathcal{F}}(x^k)$, we have

$$\eta^k = \frac{\mathbb{x}^k - x^k}{\|\mathbb{x}^k - x^k\|} \in \widehat{\mathcal{N}}_{\mathcal{F}}(x^k).$$

Since x^* is quasinormal, from Lemma 2.5 it follows that the point x^k is also quasinormal for all sufficiently large k and, without loss of generality, we may assume that all x^k are quasinormal. Then, by Proposition 2.3.2 and Proposition 1.3.4 (iv), there exists a sequence $\{\xi_1^k, \ldots, \xi_p^k, \zeta_1^k, \ldots, \zeta_q^k\}$ with $\zeta_j^k \geq 0$, such that

$$\eta^k \in \sum_{i=1}^{p} \xi_i^k \nabla h_i(x^k) + \sum_{j=1}^{q} \zeta_j^k \partial g_j(x^k) + \mathcal{N}_{\mathcal{X}}(x^k), \tag{2.22}$$

and there exists a sequence $\{x^{k,l}\} \subseteq \mathcal{X}$, such that $x^{k,l} \to x^k$ as $l \to \infty$ and for all $l = 1, 2, \ldots$, $\xi_i^k h_i(x^{k,l}) > 0$ for $i \in I^k$; $\zeta_j^k g_j(x^{k,l}) > 0$ for $j \in J^k$, where $I^k = \{i : \xi_i^k \neq 0\}$ and $J^k = \{j : \zeta_j^k > 0\}$. As in the proof of Step 2 in Proposition 2.3.2, we can show that the quasinormality of x^* implies that the sequence $\{\xi_1^k, \ldots, \xi_p^k, \zeta_1^k, \ldots, \zeta_q^k\}$ is bounded.

Therefore, without loss of generality, we may assume $\{\xi_1^k, \ldots, \xi_p^k, \zeta_1^k, \ldots, \zeta_q^k\}$ converges to some vector $\{\xi_1^*, \ldots, \xi_p^*, \zeta_1^*, \ldots, \zeta_q^*\}$. Then there exists a number $M_0 > 0$, such that for all k, $\|(\xi^k, \zeta^k)\| \le M_0$.

Without loss of any generality, we may assume that $\mathbb{x}^k \in \mathcal{B}_{\frac{\delta_0}{2}}(x^*) \cap \mathcal{X} \setminus \mathcal{F}$ and $x^k \in \mathcal{B}_{\delta_0}(x^*) \cap \mathcal{X}$ for all k. Setting $(\bar{\xi}^k, \bar{\zeta}^k) = 2(\xi^k, \zeta^k)$, then from (2.22) for each k there exist $\rho_j^k \in \partial g_j(x^k), \forall j = 1, \ldots, q$, and $\omega^k \in \mathcal{N}_{\mathcal{X}}(x^k)$ such that

$$\frac{\mathbb{x}^k - x^k}{\|\mathbb{x}^k - x^k\|} = \frac{x^k - \mathbb{x}^k}{\|\mathbb{x}^k - x^k\|} + \sum_{i=1}^{p} \bar{\xi}_i^k \nabla h_i(x^k) + \sum_{j=1}^{q} \bar{\zeta}_j^k \rho_j^k + 2\omega^k.$$

We obtain from the discussion above that

$$\begin{aligned}
\|\mathbb{x}^k - x^k\| &= \frac{\langle \mathbb{x}^k - x^k, \mathbb{x}^k - x^k \rangle}{\|\mathbb{x}^k - x^k\|} \\
&= \Big\langle \frac{x^k - \mathbb{x}^k}{\|\mathbb{x}^k - x^k\|} + \sum_{i=1}^{p} \bar{\xi}_i^k \nabla h_i(x^k) + \sum_{j=1}^{q} \bar{\zeta}_j^k \rho_j^k + 2\omega^k, \mathbb{x}^k - x^k \Big\rangle \\
&\le \Big\langle \frac{x^k - \mathbb{x}^k}{\|\mathbb{x}^k - x^k\|} + \sum_{i=1}^{p} \bar{\xi}_i^k \nabla h_i(x^k) + \sum_{j=1}^{q} \bar{\zeta}_j^k \rho_j^k, \mathbb{x}^k - x^k \Big\rangle + o(\|\mathbb{x}^k - x^k\|) \\
&\le \sum_{i=1}^{p} \Big\langle \bar{\xi}_i^k \nabla h_i(x^k), \mathbb{x}^k - x^k \Big\rangle + \sum_{j=1}^{q} \Big\langle \bar{\zeta}_j^k \rho_j^k, \mathbb{x}^k - x^k \Big\rangle + o(\|\mathbb{x}^k - x^k\|) \\
&\le \sum_{i=1}^{p} \bar{\xi}_i^k \Big(h_i(\mathbb{x}^k) - o(\|\mathbb{x}^k - x^k\|) \Big) + \sum_{j=1}^{q} \bar{\zeta}_j^k \Big(g_j(\mathbb{x}^k) - o(\|\mathbb{x}^k - x^k\|) \Big) \\
&\qquad + o(\|\mathbb{x}^k - x^k\|) \\
&\le 2\Big| \sum_{i=1}^{p} \xi_i^k h_i(\mathbb{x}^k) + \sum_{j=1}^{q} \zeta_j^k g_j(\mathbb{x}^k) \Big| + 2\Big| \sum_{i=1}^{p} \xi_i^k + \sum_{j=1}^{q} \zeta_j^k + 1 \Big| o(\|\mathbb{x}^k - x^k\|) \\
&\le 2\Big| \sum_{i=1}^{p} \xi_i^k h_i(\mathbb{x}^k) + \sum_{j=1}^{q} \zeta_j^k g_j(\mathbb{x}^k) \Big| + \frac{1}{2}\|\mathbb{x}^k - x^k\|
\end{aligned}$$

where the first inequality comes from the fact that $\mathcal{X}$ is regular, the third one arises from the subdifferential regularity assumption of $h_i(x)$ and $g_j(x)$ in $\mathcal{B}_{\delta_0}(x^*) \cap \mathcal{X}$, and the last one is valid because without loss of generality, we may assume for k sufficiently large,

$$o(\|\mathbb{x}^k - x^k\|) \leq \frac{1}{4(M_0+1)}\|\mathbb{x}^k - x^k\|$$

since $\mathbb{x}^k - x^k \to 0$ as k tends to infinity. This means

$$\text{dist}_{\mathcal{F}}(\mathbb{x}^k) = \|\mathbb{x}^k - x^k\| \leq 4M_0\Big(\sum_{i=1}^{p}|h_i(\mathbb{x}^k)| + \sum_{i=1}^{q} g_j^+(\mathbb{x}^k)\Big).$$

Thus, for any sequence $\{\mathbb{x}^k\} \subseteq \mathcal{X}$ converging to x^* there exists a number $c > 0$ such that

$$\text{dist}_{\mathcal{F}}(\mathbb{x}^k) \leq c(\|h(\mathbb{x}^k)\|_1 + \|g^+(\mathbb{x}^k)\|_1) \quad \forall k = 1, 2, \ldots.$$

This further implies the error bound property at x^*. Indeed, suppose the contrary. Then there exists a sequence $\tilde{\mathbb{x}}^k \to x^*$, such that $\tilde{\mathbb{x}}^k \in \mathcal{X} \setminus \mathcal{F}$ and $\text{dist}_{\mathcal{F}}(\tilde{\mathbb{x}}^k) > c(\|h(\tilde{\mathbb{x}}^k)\|_1 + \|g^+(\tilde{\mathbb{x}}^k)\|_1)$ for all $k = 1, 2, \ldots$, which is a contradiction. □

A natural question to ask is: Is the quasinormality strictly stronger than the error bound property. This question has been answered positively in [57, Example 2.1], with a smooth optimization problem without an abstract set constraint.

2.5 Sensitivity analysis of value functions

In this section we consider the following perturbed optimization problem:

$$\begin{aligned} (\text{NLP}_\text{a}) \qquad \min \quad & \hat{f}(x,a) \\ \text{s.t.} \quad & x \in \mathcal{F}(a), \end{aligned} \tag{2.23}$$

with

$$\mathcal{F}(a) = \{x \in \mathcal{X} : \hat{h}(x,a) = 0, \hat{g}(x,a) \leq 0\}, \tag{2.24}$$

where $\mathcal{X}$ is closed subset of $\mathbb{R}^n$, $\hat{f} : \mathbb{R}^n \times \mathbb{R}^{\mathrm{n}} \to \mathbb{R}, \hat{h} : \mathbb{R}^n \times \mathbb{R}^{\mathrm{n}} \to \mathbb{R}^p, \hat{g} : \mathbb{R}^n \times \mathbb{R}^{\mathrm{n}} \to \mathbb{R}^q$ are Lipschitz continuous around $(\bar{x}, \bar{a})$.

In practice it is often important to know how well the model responds to the perturbation a. For this we need to consider, for instance, the value function $\mathcal{V}(a)$ related to the parametric optimization problem:

$$\mathcal{V}(a) := \inf_{x \in \mathcal{F}(a)} \hat{f}(x, a), \tag{2.25}$$

with the solution map $\mathcal{S}(\cdot)$ defined by

$$\mathcal{S}(a) := \{x \in \mathcal{F}(a) : \mathcal{V}(a) = \hat{f}(x, a)\}. \tag{2.26}$$

In the recent paper [66], Mordukhovich, Nam and Yen obtain some new results for computing and estimating the Fréchet subgradient of the value function in parametric optimization (2.23)-(2.24) with smooth and nonsmooth data using normal multipliers. In the following result we estimate the Fréchet subdifferential of the value function by using the quasinormal multipliers instead. Since the set of quasinormal multipliers are smaller than the set of normal multipliers, our estimate provides a tighter bound for the Fréchet subdifferential of the value function.

Let $M_Q^r(\bar{x}, \bar{a})$ denotes the set of vectors $(\lambda, \mu, \gamma) \in \mathbb{R}^p \times \mathbb{R}^q_+ \times \mathbb{R}$ such that

$$0 \in r\partial \hat{f}(\bar{x}, \bar{a}) + \sum_{i=1}^{p} \partial(\lambda_i \hat{h}_i)(\bar{x}, \bar{a}) + \sum_{j=1}^{q} \mu_i \partial \hat{g}_j(\bar{x}, \bar{a}) + (0, \gamma) + \mathcal{N}_{\mathcal{X}}(\bar{x}) \times \{0\}$$

and there exists a corresponding sequence $\{(x^k, a^k)\} \subseteq \mathcal{X} \times \mathbb{R}^{\mathrm{n}}$ converging to $(\bar{x}, \bar{a})$ such that $\lambda_i \hat{h}_i(x^k, a^k) > 0$ for all $i \in I := \{i : \lambda_i \neq 0\}$, $\mu_j \hat{g}_j(x^k, a^k) > 0$ for all $j \in J := \{\mu_j > 0\}$, and $\hat{h}_i (i \in I), \hat{g}_j (j \in J)$ are proximal subdifferentiable at (x^k, a^k) for each k.

Theorem 2.11. *Let $\mathcal{V}(a)$ be the value function as defined in (2.25) and $\bar{x} \in \mathcal{S}(\bar{a})$. Assume also that $(\bar{x}, \bar{a})$ is quasinormal for the constraint region*

$$\{(x, a) \in \mathcal{X} \times \mathbb{R}^{\mathrm{n}} : \hat{h}(x, a) = 0, \hat{g}(x, a) \leq 0\}.$$

Then one has the upper estimation:

$$\hat{\partial}\mathcal{V}(\bar{a}) \subseteq \{-\gamma : (\lambda, \mu, \gamma) \in M_Q^1(\bar{x}, \bar{a})\}. \tag{2.27}$$

Proof. There is nothing to prove if $\hat{\partial}\mathcal{V}(\bar{a}) = \emptyset$. Let $\gamma \in \hat{\partial}\mathcal{V}(\bar{a}) \neq \emptyset$. Then by definition of the Fréchet subdifferential, for arbitrary $\kappa > 0$, there exists $\delta_\kappa > 0$ such that

$$\mathcal{V}(a) - \mathcal{V}(\bar{a}) \geq \langle \gamma, a - \bar{a} \rangle - \kappa \|a - \bar{a}\| \qquad \forall a \in \mathcal{B}_{\delta_\kappa}(\bar{a}).$$

By definition of the value function, for every $x \in \mathcal{F}(a)$, we have $\hat{f}(x, a) \geq \mathcal{V}(a)$ and hence

$$\hat{f}(x, a) - \langle \gamma, a - \bar{a} \rangle + \kappa \|a - \bar{a}\| \geq \hat{f}(\bar{x}, \bar{a}) \qquad \forall x \in \mathcal{F}(a).$$

Thus, $(\bar{x}, \bar{a})$ is a local optimal solution to the optimization problem

$$\begin{aligned} \min \quad & \hat{f}(x, a) - \langle \gamma, a - \bar{a} \rangle + \kappa \|a - \bar{a}\| \\ \text{s.t.} \quad & \hat{h}_i(x, a) = 0, \quad i = 1, \ldots, p, \\ & \hat{g}_j(x, a) \leq 0, \quad j = 1, \ldots, q, \\ & (x, a) \in \mathcal{X} \times \mathbb{R}^{\mathrm{n}}. \end{aligned}$$

Since $(\bar{x}, \bar{a})$ is quasinormal by assumption, by the enhanced KKT condition (Theorem 2.3), there exist a vector $(\lambda, \mu) \in \mathbb{R}^p \times \mathbb{R}^q_+$ and a sequence $\{(x^k, a^k)\} \subseteq \mathcal{X} \times \mathbb{R}^n$

converging to $(\bar{x}, \bar{a})$ such that the following conditions hold:

$$0 \in \partial \hat{f}(\bar{x}, \bar{a}) + \sum_{i=1}^{p} \partial(\lambda_i \hat{h}_i)(\bar{x}, \bar{a}) + \sum_{j=1}^{q} \mu_j \partial \hat{g}_j(\bar{x}, \bar{a}) +$$

$$\mathcal{N}_{\mathcal{X} \times \mathbb{R}^{\mathrm{n}}}(\bar{x}, \bar{a}) - \begin{pmatrix} 0 \\ \gamma \end{pmatrix} + \kappa \begin{pmatrix} 0 \\ \mathbb{B} \end{pmatrix}, \tag{2.28}$$

$$\lambda_i \hat{h}_i(x^k, a^k) > 0 \ \forall i \in I, \quad \mu_j \hat{g}_j(x^k, a^k) > 0 \ \forall j \in J,$$

$$\hat{h}_i(i \in I), \hat{g}_j(j \in J) \text{ are proximal subdifferentiable at } (x^k, a^k).$$

The desired upper estimation follows since κ is arbitrary. □

We now give a tighter estimate for the limiting subdifferential of the value function in terms of the quasinormality.

Theorem 2.12. *Let $\mathcal{V}(a)$ be the value function as defined in (2.25). Suppose that the growth hypothesis holds, i.e., there exists $\delta > 0$ such that the set*

$$\{x \in \mathcal{X} : \hat{h}(x, \bar{a}) = \alpha, \ \hat{g}(x, \bar{a}) \leq \beta, \ \hat{f}(x, \bar{a}) \leq M, (\alpha, \beta) \in \delta \mathbb{B}\}$$

is bounded for each $M \in \mathbb{R}$. Assume that for each $\bar{x} \in \mathcal{S}(\bar{a})$, $(\bar{x}, \bar{a})$ is quasinormal for the constraint region

$$\{(x, a) \in \mathcal{X} \times \mathbb{R}^{\mathrm{n}} : \hat{h}(x, a) = 0, \ \hat{g}(x, a) \leq 0\}. \tag{2.29}$$

Then the value function $\mathcal{V}(a)$ is l.s.c. near $\bar{a}$ and

$$\partial \mathcal{V}(\bar{a}) \subseteq \bigcup_{\bar{x} \in \mathcal{S}(\bar{a})} \{-\gamma : (\lambda, \mu, \gamma) \in M_Q^1(\bar{x}, \bar{a})\}$$

$$\partial^{\infty} \mathcal{V}(\bar{a}) \subseteq \bigcup_{\bar{x} \in \mathcal{S}(\bar{a})} \{-\gamma : (\lambda, \mu, \gamma) \in M_Q^0(\bar{x}, \bar{a})\}$$

Proof. By [47, Theorem 3.6], the value function $\mathcal{V}(a)$ is lower semicontinuous near $\bar{a}$ under our assumption.

Step 1. Let v be a vector that belongs to $\partial\mathcal{V}(\bar{a})$, by definition there are sequences $a^l \to \bar{a}$ and $v^l \to v$ with $v^l \in \hat{\partial}\mathcal{V}(a^l)$. By the growth hypothesis, for l sufficiently large, we may find a solution $x^l \in \mathcal{S}(a^l)$. Following [16, Theorem 6.5.2], without loss of generality we may assume x^l converges to an element $\bar{x} \in \mathcal{S}(\bar{a})$. Since $(\bar{x}, \bar{a})$ is quasinormal and it is a limit point of the sequence $\{(x^l, a^l)\}$, by Lemma 2.5 we find that for sufficient large l, (x^l, a^l) is also quasinormal for the constraint region (2.29) and hence from Theorem 2.11 it follows that for each l there exist a vector $(\lambda^l, \mu^l) \in \mathbb{R}^p \times \mathbb{R}^q_+$ and a sequence $\{(x^{l,k}, a^{l,k})\}_k \subseteq \mathcal{X} \times \mathbb{R}^n$ converging to (x^l, a^l) as $k \to \infty$ such that

$$(0, v^l) \in \partial \hat{f}(x^l, a^l) + \sum_{i=1}^{p} \partial(\lambda_i^l \hat{h}_i)(x^l, a^l) + \sum_{j=1}^{q} \mu_j^l \partial \hat{g}_j(x^l, a^l) + \mathcal{N}_{\mathcal{X}}(x^l) \times \{0\}, \tag{2.30}$$

$$\lambda_i^l \hat{h}_i(x^{l,k}, a^{l,k}) > 0 \ \forall i \in I^l, \quad \mu_j^l \hat{g}_j(x^{l,k}, a^{l,k}) > 0 \ \forall j \in J^l, \tag{2.31}$$

$$\hat{h}_i(i \in I^l), \hat{g}_j(j \in J^l) \text{ are proximal subdifferentiable at } (x^{l,k}, a^{l,k}), \tag{2.32}$$

where $I^l := \{i : \lambda_i^l \neq 0\}, J^l := \{j : \mu_j^l > 0\}$. Similar as in Step 2 of the proof of Proposition 2.3.2, we may obtain the boundedness of the multipliers sequence $\{\lambda_1^l, \ldots, \lambda_p^l, \mu_1^l, \ldots, \mu_q^l\}$. Therefore, without loss of generality, we may assume $\{\lambda_1^l, \ldots, \lambda_p^l, \mu_1^l, \ldots, \mu_q^l\}$ converges to $\{\lambda_1, \ldots, \lambda_p, \mu_1, \ldots, \mu_q\}$. Taking the limit on both sides of (2.30), similar to Theorem 2.1, we obtain

$$(0, v) \ \in \ \partial \hat{f}(\bar{x}, \bar{a}) + \sum_{i=1}^{p} \partial(\lambda_i \hat{h}_i)(\bar{x}, \bar{a}) + \sum_{j=1}^{q} \mu_j \partial \hat{g}_j(\bar{x}, \bar{a}) + \mathcal{N}_{\mathcal{X}}(\bar{x}) \times \{0\}.$$

Also we find a sequence $\{(x^{l,k_l}, a^{l,k_l})\} \subseteq \mathcal{X} \times \mathbb{R}^n$ converging to $\bar{x}$ as $l \to \infty$ and is such that for all l, $\lambda_i \hat{h}_i(x^{l,k_l}, a^{l,k_l}) > 0, \forall i \in I$, $\mu_j g_j(x^{l,k_l}, a^{l,k_l}) > 0, \forall j \in J$, and h_i, g_j are proximal subdifferentiable at x^{l,k_l}, where $I = \{i|\lambda_i \neq 0\}$ and $J = \{j|\mu_j > 0\}$.

Step 2. Let $v \in \partial^\infty \mathcal{V}(\bar{a})$. By definition there are sequence $a^l \to \bar{a}$, $v^l \in \hat{\partial}\mathcal{V}(a^l)$ and $t^l \downarrow 0$ such that $t^l v^l \to v$. Similar as in Step 1, for each l there exist a vector $(\lambda^l, \mu^l) \in \mathbb{R}^p \times \mathbb{R}^q_+$ and a sequence $\{(x^{l,k}, a^{l,k})\} \subseteq \mathcal{X} \times \mathbb{R}^n$ converging to (x^l, a^l) such that (2.30)-(2.32) hold. Multiplying both sides of (2.30) by t^l we have

$$\begin{aligned}(0, t^l v^l) \in t^l \partial \hat{f}(x^l, a^l) \quad &+ \quad \sum_{i=1}^{p} \partial(t^l \lambda_i^l \hat{h}_i)(x^l, a^l) + \sum_{j=1}^{q} t^l \mu_j^l \partial \hat{g}_j(x^l, a^l) \\ &+\mathcal{N}_{\mathcal{X}}(\bar{x}) \times \{0\}. \end{aligned} \tag{2.33}$$

Since $(\bar{x}, \bar{a})$ is quasinormal for the constraint region (2.29), similarly as in Step 2 of the proof of Proposition 2.3.2, the sequence

$$\{t^l \lambda_1^l, \ldots, t^l \lambda_p^l, t^l \mu_1^l, \ldots, t^l \mu_q^l\}$$

must be bounded as $l \to \infty$. Without loss of generality assume that the limit is $\{\lambda_1, \ldots, \lambda_p, \mu_1, \ldots, \mu_q\}$. Talking limits in (2.33), we have

$$(0, v) \quad \in \quad \sum_{i=1}^{p} \lambda_i \partial \hat{h}_i(\bar{x}, \bar{a}) + \sum_{j=1}^{q} \mu_j \partial \hat{g}_j(\bar{x}, \bar{a}) + \mathcal{N}_{\mathcal{X}}(\bar{x}) \times \{0\}.$$

The rest of the proof is similar to Step 1. □

From Theorem 2.12 we derive the following very interesting result which significantly improves the classical result in that our sufficient condition is the perturbed quasinormality which is much weaker than the classical condition of NNAMCQ (see e.g. [47, Corollary 3.7]).

Corollary 2.13. *Let $\mathcal{V}(a)$ be the value function as defined in (2.25). Suppose that* *the growth hypothesis holds at each $\bar{x} \in \mathcal{S}(\bar{a})$.*

(i) Assume that $(\bar{x}, \bar{a})$ is quasinormal for the constraint region (2.29). If

$$\bigcup_{\bar{x} \in \mathcal{S}(\bar{a})} \{-\gamma : (\lambda, \mu, \gamma) \in M_Q^0(\bar{x}, \bar{a})\} = \{0\}, \tag{2.34}$$

then the value function $\mathcal{V}(a)$ is Lipschitz continuous around $\bar{a}$ with

$$\emptyset \neq \partial \mathcal{V}(\bar{a}) \subseteq \bigcup_{\bar{x} \in \mathcal{S}(\bar{a})} \{-\gamma : (\lambda, \mu, \gamma) \in M_Q^1(\bar{x}, \bar{a})\}.$$

In addtion to the above assumptions, if

$$\bigcup_{\bar{x} \in \mathcal{S}(\bar{a})} \{-\gamma : (\lambda, \mu, \gamma) \in M_Q^1(\bar{x}, \bar{a})\} = \{-\bar{\gamma}\}$$

for some $(\bar{\lambda}, \bar{\mu}, \bar{\gamma}) \in M_Q^1(\bar{x}, \bar{a})$, then $\mathcal{V}$ is strictly differentiable at $\bar{a}$ and $\nabla \mathcal{V}(\bar{a}) = -\bar{\gamma}$.

(ii) For the functions $\phi = \hat{f}, \pm \hat{h}_i, \hat{g}_j$, suppose that the partial limiting subdifferential property holds at $(\bar{x}, \bar{a})$:

$$\partial \phi(\bar{x}, \bar{a}) = \partial_x \phi(\bar{x}, \bar{a}) \times \partial_a \phi(\bar{x}, \bar{a}).$$

Also assume that $(\bar{x}, \bar{a})$ is quasinormal for the constraint region (2.29). If

$$\bigcup_{\bar{x} \in \mathcal{S}(\bar{a})} \left\{ \sum_{i=1}^{p} \partial_a (\lambda_i \hat{h}_i)(\bar{x}, \bar{a}) + \sum_{j=1}^{q} \mu_j \partial_a \hat{g}_j(\bar{x}, \bar{a}) : (\mu, \lambda) \in \widetilde{M}_Q^0(\bar{x}, \bar{a}) \right\} = \{0\} \tag{2.35}$$

then the value function $\mathcal{V}(a)$ is Lipschitz continuous around $\bar{a}$ and

$$\emptyset \neq \partial\mathcal{V}(\bar{a}) \subseteq \bigcup_{\substack{\bar{x}\in\mathcal{S}(\bar{a})\\(\mu,\lambda)\in\widetilde{M}_Q^1(\bar{x},\bar{a})}} \left\{\partial_a\hat{f}(\bar{x},\bar{a}) + \sum_{i=1}^{p}\partial_a(\lambda_i\hat{h}_i)(\bar{x},\bar{a}) + \sum_{j=1}^{q}\mu_i\partial_a\hat{g}_j(\bar{x},\bar{a})\right\} \tag{2.36}$$

where $\widetilde{M}_Q^r(\bar{x},\bar{a})$ denotes the set of perturbed quasinormal multipliers which are the set of vectors $(\lambda,\mu)\in\mathbb{R}^p\times\mathbb{R}^q_+$ such that

$$0 \in r\partial_x\hat{f}(\bar{x},\bar{a}) + \sum_{i=1}^{p}\partial_x(\lambda_i\hat{h}_i)(\bar{x},\bar{a}) + \sum_{j=1}^{q}\mu_j\partial_x\hat{g}_j(\bar{x},\bar{a}) + \mathcal{N}_{\mathcal{X}}(\bar{x})$$

and there exists a corresponding sequence $\{(x^k,a^k)\}\subseteq\mathcal{X}\times\mathbb{R}^{\mathrm{n}}$ converging to $(\bar{x},\bar{a})$ such that $\lambda_i\hat{h}_i(x^k,a^k)>0$ for all $i\in I:=\{i:\lambda_i\neq 0\}$, $\mu_j\hat{g}_j(x^k,a^k)>0$ for all $j\in J:=\{\mu_j>0\}$, and $\hat{h}_i(i\in I),\hat{g}_j(j\in J)$ are proximal subdifferentiable at (x^k,a^k) for each k.

(iii) Suppose that the partial limiting subdifferential property at $(\bar{x},\bar{a})$ holds as in (ii) and $\widetilde{M}_Q^0(\bar{x},\bar{a})=\{0\}$ for each $\bar{x}\in\mathcal{S}(\bar{a})$. Then the value function $\mathcal{V}(a)$ is Lipschitz continuous around $\bar{a}$ and (2.36) holds.

Proof. (i) It follows from Theorem 2.12 that

$$\bigcup_{\bar{x}\in\mathcal{S}(\bar{a})}\{-\gamma:(\lambda,\mu,\gamma)\in M_Q^0(\bar{x},\bar{a})\}=\{0\}$$

implies that $\partial^\infty\mathcal{V}(\bar{a})=\{0\}$. We conclude that the value function is Lipschitz around $\bar{a}$ by virtue of Proposition 1.3.1 (iii). The assertion about the strict differentiability then follows from Proposition 1.3.1 (i).

(ii) It is clear that under the partial limiting subdifferential property, (2.34) is equivalent to (2.35). The conclusion then follows from applying Theorem 2.12 and

Proposition 1.3.1 (iii).

(iii) follows immediately from (ii) and the fact that $\widetilde{M}_Q^0(\bar{x}, \bar{a}) = \{0\}$ implies the quasinormality of $(\bar{x}, \bar{a})$.

$\square$

Chapter 3

Mathematical programs with geometric constraints in Banach spaces: Enhanced optimality, exact penalty, and sensitivity

3.1 Introduction

In this chapter we extend the enhanced optimality, exact penalty, and sensitivity discussed in chapter 2 from $\mathbb{R}^n$ to a more general infinite dimensional space. Unless otherwise stated, we denote by $\mathbb{X}$ a Banach space and by $\mathbb{X}^*$ its dual space equipped with the weak* topology and by $\mathbb{Y}$ an Euclidean space together with the inner product $\langle \cdot\,, \cdot \rangle$ equipped with the orthogonal basis $\mathscr{E} = \{e_1, \ldots, e_m\}$. We study the following general mathematical program with geometric constraints such that the image of a

This chapter is the content of Guo, L., Ye, J.J. and Zhang, J., "Mathematical programs with geometric constraints in Banach spaces: enhanced optimality, exact penalty, and sensitivity." (2013). SIAM J. Optim., **23**, 2295-2319.

mapping from a Banach space is included in a closed subset of a finite dimensional space:

$$\text{(MPGC)} \qquad \min_{x\in\Omega} \quad f(x) \qquad \text{s.t.} \quad F(x) \in \Lambda, \tag{3.1}$$

where $f : \mathbb{X} \to \mathbb{R}$ and $F : \mathbb{X} \to \mathbb{Y}$ are Lipschitzian near the point of interest, Ω and Λ are nonempty and closed subsets of $\mathbb{X}$ and $\mathbb{Y}$ respectively.

The classical Fritz John (FJ) necessary optimality condition for (MPGC) with continuously differentiable functions $\{f, F\}$, $\Omega = \mathbb{X}$, and convex geometric constraint Λ takes the following form: There exist $r \geq 0$ and $\mu \in \mathbb{Y}$ not all equal to zero such that

$$0 = r\nabla f(x^*) + \nabla F(x^*)^*\mu \quad \text{and} \quad \mu \in \mathcal{N}_\Lambda(F(x^*)), \tag{3.2}$$

where $\nabla\varphi(x)$ is the Fréchet derivative of mapping φ at x, $\mathcal{A}^*$ denotes the adjoint of a linear operator $\mathcal{A}$, and $\mathcal{N}_\Lambda(y)$ denotes the normal cone of Λ at y in the sense of convex analysis [73]:

$$\mathcal{N}_\Lambda(y) = \begin{cases} \{d \in \mathbb{Y} \mid \langle d, z - y\rangle \leq 0 \ \forall z \in \Lambda\} & \text{if } y \in \Lambda, \\ \emptyset & \text{if } y \notin \Lambda. \end{cases}$$

From the FJ condition, it follows immediately that if x^* is a locally optimal solution of (MPGC) and the *no nonzero abnormal multiplier constraint qualification* (NNAMCQ) or *basic constraint qualification* (Basic CQ) [74] holds at x^*, i.e., there is no nonzero μ such that

$$0 = \nabla F(x^*)^*\mu \quad \text{and} \quad \mu \in \mathcal{N}_\Lambda(F(x^*)),$$

then there exist $r > 0$ (which can be taken as 1) and $\mu \in \mathbb{Y}$ such that the KKT condition holds

$$0 = \nabla f(x^*) + \nabla F(x^*)^* \mu \quad \text{and} \quad \mu \in \mathcal{N}_\Lambda(F(x^*)).$$

Since $\mathbb{Y}$ is assumed to be a finite dimensional space and Λ is a closed convex set, by virtue of [10, Corollary 2.98], the NNAMCQ is equivalent to the Robinson's CQ

$$0 \in int\{F(x^*) + \nabla F(x^*)\mathbb{X} - \Lambda\},$$

where "int" denotes the topological interior of a given set.

When $\{f, F\}$ are nonsmooth but locally Lipschitzian and Λ is not a convex set, the FJ condition can be obtained by replacing the usual derivatives and the normal cone in the sense of convex analysis with the limiting subdifferential and the limiting normal cone respectively if the underlying space $\mathbb{X}$ is an Asplund space (an Asplund space $\mathbb{X}$ is a Banach space such that every separable closed subspace of $\mathbb{X}$ has a separable dual, see Mordukhovich [61, 62]), and by the Clarke subdifferential and the Clarke normal cone respectively if $\mathbb{X}$ is a general Banach space (see Clarke [16]).

Although the NNAMCQ or the Basic CQ provides an easy way to verify constraint qualification, it may be fairly strong for some applications and in particular for certain classes of optimization problems such as bilevel programs, mathematical programs with equilibrium constraints [21, 49, 68, 86, 87], they are never satisfied.

3.1.1 Purpose and contribution

The main purpose of this chapter is to study the enhanced optimality condition for (MPGC) when the space $\mathbb{X}$ is a Banach space and $\mathbb{Y}$ is an Euclidean space. Such a result is new even for the classical smooth nonlinear program. There are two technical difficulties involved when the space $\mathbb{X}$ is not finite dimensional. First, unlike the finite dimensional case, the quadratic penalization approach in chapter 2 cannot be employed any more because the compactness of the closed unit ball is possibly invalid in a Banach space $\mathbb{X}$, which plays a key role in guaranteeing the existence of enhanced sequential approximating solution by using the Weierstrass theorem in chapter 2. Nevertheless, by virtue of the optimization process, for any $\epsilon > 0$, a problem in the form as (MPGC) always possesses an ϵ-optimal solution (see [10] for definition), provided that the optimal value of the problem is finite. Inspired by this fact, we employ the Ekeland's variational principle instead to construct a cluster of ϵ-optimal solutions, and each of them becomes the minimizer of a certain slightly perturbed problem. We then employ generalized calculus to obtain necessary optimality conditions for the perturbed problem. The second difficulty lies in applying the basic calculus rules and passing to the limit as ϵ tends to zero. When the space $\mathbb{X}$ is finite dimensional, the limiting subdifferential and the limiting normal cone have nice calculus rules and are known to be closed as set-valued maps. The nice calculus rules and the robust property allow one to obtain the desired result. However, when $\mathbb{X}$ is an infinite dimensional Banach space, the limiting subdifferential for locally Lipschitzian functions may even be empty, and hence the basic calculus rules may fail and the robust property may not hold in general. To cope with the second difficulty, we use the approximate subdifferential developed by Ioffe [35, 36] instead. The approximate subdifferential seems to be the most natural analytic tool in our situation since it has fairly rich calculus rule for locally Lipschitzian functions and the approximate subdifferential and the approximate normal cone are known to be closed as set-valued maps. Moreover, the approximate subdifferential for locally Lipschitzian functions is minimal (as a set) among all subdifferentials that have desired properties, and is in general smaller than the Clarke subdifferential. When the underlying space $\mathbb{X}$ is a weakly compactly generated (WCG) Asplund space, the approximate subdifferential coincides with the limiting subdifferential [61, Theorem 3.59] and hence in this case, we obtain the desired result in terms of limiting subdifferential. Recall that $\mathbb{X}$ is WCG if there is a weakly compact set $\mathbb{K} \subseteq \mathbb{X}$ such that $\mathbb{X}$ is equal to the closure of the span of $\mathbb{K}$. Canonical examples of WCG Aspund spaces are reflexive Banach spaces, see, e.g., [61] for further discussions.

In recent years, it has been shown that constraint qualifications have strong connections with the stability of feasible region under certain perturbation p:

$$\mathscr{F}(p) := \{x \in \Omega \,|\, F(x,p) \in \Lambda\},$$

where p is in a topological space $\mathbb{P}$. For the case of a smooth optimization problem with convex geometric constraint Λ and $\mathbb{X} = \Omega$, it is known that the Robinson's CQ at $x^* \in \mathcal{F}(p^*)$ implies the stability for the constraint region (see [10, Theorem 2.87]), i.e., the existence of a neighborhood $\mathcal{U}$ of (x^*, p^*) such that for all $(x,p) \in \mathcal{U} \cap (\mathbb{X} \times \mathbb{P})$,

$$\mathrm{dist}_{\mathscr{F}(p)}(x) = O(\mathrm{dist}_\Lambda(F(x,p))),$$

and hence the existence of local error bounds, i.e., there exist positive constants $\{\kappa, \delta_0\}$ such that

$$\mathrm{dist}_{\mathscr{F}(p^*)}(x) \leq \kappa \mathrm{dist}_\Lambda(F(x,p^*)) \quad \forall x \in \mathcal{B}_{\delta_0}(x^*) \cap \mathbb{X}.$$

In fact, the above stability results still hold in an infinite dimensional space even if the set Λ is not convex and $\mathbb{X}$ is replaced with a closed subset Ω under the NNAMCQ which can be easily derived by using the error bound result as in [80, Theorem 2.4].

Error bounds have important applications in sensitivity analysis of mathematical programming and in convergence analysis of some algorithms. In his seminal paper [31], Hoffman showed that a linear inequality system in a finite dimensional space has a global error bound. Such a result was generalized to an infinite dimensional Banach space by Ioffe [34]. For a general constraint system, the existence of error bounds usually requires some conditions. As we discussed above, the Robinson's CQ and the NNAMCQ imply a local error bound for (MPGC). Therefore, the error bound estimates can be obtained straightforwardly for smooth nonlinear programs and nonlinear semidefinite programs (NLSDP) with the constraint systems taking the geometric forms respectively (see [10, Example 2.92, Example 2.93]). Very recently, for the case of nonsmooth (NLP), it was shown in chapter 2 that either the pseudonormality or the quasinormality with regularity on the constraints implies the existence of local error bounds, which extends the result in [57] where all constraints are assumed to be twice continuously differentiable. In this chapter, we show that a local error bound for nonsmooth (MPGC) exists under the quasinormality, which generalizes and improves all earlier results since except the constraint qualification, neither additional regularity condition nor continuous differentiability assumption is required.

3.2 Preliminaries

We denote by $\mathscr{F}$ the feasible region of (MPGC) and denote by $\mathcal{B}_\delta(x) := \{y \in \mathbb{X} \mid \|y - x\| < \delta\}$ the open ball centered at x with radius $\delta > 0$. As usual, $\mathbb{B}_{\mathbb{X}}$ and $\mathbb{B}_{\mathbb{X}^*}$ stand for the closed unit balls of the space $\mathbb{X}$ and its dual $\mathbb{X}^*$ respectively.

In addition to section 1.3, we next summarize some preliminary material in variational analysis in infinite dimensional spaces that will be needed in this chapter. We refer the reader to $[16, 35, 36, 61, 62, 77]$ for more details and discussions.

For a set-valued map $S : \mathbb{X} \rightrightarrows \mathbb{X}^*$, unless specified, we denote by

$$\operatorname*{Lim\,sup}_{x \to x^*} S(x) := \{v \in \mathbb{X}^* | \exists \text{ sequences } x_k \to x^* \text{ and } v_k \xrightarrow{w^*} v \\ \text{with } v_k \in S(x_k) \text{ for all k}\}.$$

the sequential Painlevé–Kuratowski upper limit with respect to the norm topology of $\mathbb{X}$ and the weak* topology of $\mathbb{X}^*$.

Given $\Omega \subseteq \mathbb{X}$ and $\epsilon \geq 0$, define the collection of ϵ-normals to Ω at $x^* \in \Omega$ by

$$\widehat{\mathcal{N}}_\epsilon(x^*, \Omega) := \{v \in \mathbb{X}^* \mid \limsup_{x \xrightarrow{\Omega} x^*} \frac{\langle v, x - x^* \rangle}{\|x - x^*\|} \leq \epsilon\}. \tag{3.3}$$

where $x \xrightarrow{\Omega} x^*$ means that $x \to x^*$ with $x \in \Omega$. When $\epsilon = 0$, elements of (3.3) are called Fréchet normals and their collection, denoted by $\widehat{\mathcal{N}}_\Omega(x^*)$, is the prenormal cone to Ω at x^*. The basic/limiting normal cone $\mathcal{N}_\Omega(x^*)$ to Ω at x^* is defined as

$$\mathcal{N}_\Omega(x^*) := \operatorname*{Lim\,sup}_{x \xrightarrow{\Omega} x^*, \epsilon \downarrow 0} \widehat{\mathcal{N}}_\epsilon(x, \Omega).$$

If $\mathbb{X}$ is an Asplund space, then the limiting normal cone has the following simpler expression (see [61, Theorem 2.35])

$$\mathcal{N}_\Omega(x^*) := \operatorname*{Lim\,sup}_{x \xrightarrow{\Omega} x^*} \widehat{\mathcal{N}}_\Omega(x).$$

For $\Omega \subseteq \mathbb{X}$ and $x^* \in \Omega$, the contingent cone $\mathcal{T}_\Omega(x^*)$ to Ω at x^* is the set defined by

$$\mathcal{T}_\Omega(x^*) := \operatorname*{Lim\,sup}_{t \to 0, t \geq 0} \frac{\Omega - x^*}{t}, \tag{3.4}$$

where the "Lim sup" is taken with respect to the norm topology of $\mathbb{X}$. If the "Lim sup" in (3.4) is taken with respect to the weak topology of $\mathbb{X}$, then the resulting construction, denoted by $\mathcal{T}_\Omega^w(x^*)$, is called the weak contingent cone to Ω at x^*.

The Clarke tangent cone to Ω at x^* is defined by

$$\mathcal{T}_\Omega^c(x^*) := \{v \mid \forall x^k \to x^*, \forall t_k \downarrow 0, \exists v^k \to v^* \text{ s.t. } x^k + t_k v^k \in \Omega \quad \forall k\},$$

and the Clarke normal cone to Ω at x^* is the dual to the Clarke tangent cone to Ω at x^*, i.e.,

$$\mathcal{N}_\Omega^c(x^*) := \mathcal{T}_\Omega^c(x^*)^o,$$

where $C^o := \{x \mid \langle x, v\rangle \leq 0 \quad \forall v \in C\}$ denotes the polar of set C. In the general Banach space setting, we have

$$cl^* conv \mathcal{N}_\Omega(x^*) \subseteq \mathcal{N}_\Omega^c(x^*),$$

where "cl*conv" denotes the weak* closure of the convex hull and the inclusion relationship above holds with equality when $\mathbb{X}$ is an Asplund space.

Let $\varphi : \mathbb{X} \to \overline{\mathbb{R}}$ be an extended-real-valued function with $\varphi(x^*)$ finite. The set

$$\hat{\partial}_\epsilon \varphi(x^*) := \{v \in \mathbb{X}^* \mid \liminf_{x \to x^*} \frac{\varphi(x) - \varphi(x^*) - \langle v, x - x^*\rangle}{\|x - x^*\|} \geq -\epsilon\}$$

is called the (Fréchet-like) ϵ-subdifferential of φ at x^*. When $\epsilon = 0$, the Fréchet-like ϵ-subdifferential reduces to the Fréchet subdifferential, denoted by $\hat{\partial}\varphi(x^*)$. The basic/limiting subdifferential of φ at x^* is defined by

$$\partial \varphi(x^*) := \operatorname*{Lim\,sup}_{x \xrightarrow{\varphi} x^*, \epsilon \downarrow 0} \hat{\partial}_\epsilon \varphi(x),$$

where $x \xrightarrow{\varphi} x^*$ means that $x \to x^*$ and $\varphi(x) \to \varphi(x^*)$. The singular subdifferential of $\varphi : \mathbb{X} \to \overline{\mathbb{R}}$ at x^* is defined by

$$\partial^\infty \varphi(x^*) := \underset{x \xrightarrow{\varphi} x^*, \epsilon \downarrow 0, t \downarrow 0}{\operatorname{Lim\,sup}} \ t\hat{\partial}_\epsilon \varphi(x).$$

If $\mathbb{X}$ is an Asplund space, then we have the following simpler form ([61, Theorems 2.34 and 2.38])

$$\partial \varphi(x^*) := \underset{x \xrightarrow{\varphi} x^*}{\operatorname{Lim\,sup}} \, \hat{\partial} \varphi(x) \quad \text{and} \quad \partial^\infty \varphi(x^*) := \underset{x \xrightarrow{\varphi} x^*, t \downarrow 0}{\operatorname{Lim\,sup}} \, t\hat{\partial} \varphi(x).$$

Next we introduce the approximate subdifferential developed by Ioffe [35, 36]. The lower Dini directional derivative of φ at x^* along the direction d is given by

$$D^- \varphi(x^*, d) := \liminf_{d' \to d, t \downarrow 0} \frac{\varphi(x^* + td') - \varphi(x^*)}{t},$$

and the Dini ϵ-subdifferential of φ at x^* is defined by

$$\partial_\epsilon^- \varphi(x^*) := \{v \in \mathbb{X}^* \mid \langle v, d \rangle \leq D^- \varphi(x^*, d) + \epsilon \|d\| \quad \forall d \in \mathbb{X}\}.$$

As usual, we set $\partial_\epsilon^- \varphi(x^*) := \emptyset$ if $|\varphi(x^*)| = \infty$. The approximate subdifferential of φ at x^* is given by

$$\partial^a \varphi(x^*) := \bigcap_{L \in \mathcal{L}} \overline{\underset{x \xrightarrow{\varphi} x^*}{\operatorname{Lim\,sup}}} \, \partial_0^- (\varphi + \delta(\cdot, L))(x) = \bigcap_{L \in \mathcal{L}, \epsilon > 0} \overline{\underset{x \xrightarrow{\varphi} x^*}{\operatorname{Lim\,sup}}} \, \partial_\epsilon^- (\varphi + \delta(\cdot, L))(x),$$

where $\mathcal{L}$ is the collection of all finite dimensional subspaces of $\mathbb{X}$, $\delta(\cdot, L)$ is the indictor function of L, and $\overline{\operatorname{Lim\,sup}}$ stands for the topological counterpart of the Painlevé-Kuratowski upper limit with sequences replaced by nets. The G-normal cone $\mathcal{N}^g$ and its nucleus $\tilde{\mathcal{N}}^g$ to Ω at x^* are defined by

$$\mathcal{N}^g_\Omega(x^*) = cl^*\tilde{\mathcal{N}}^g_\Omega(x^*) \quad \text{and} \quad \tilde{\mathcal{N}}^g_\Omega(x^*) := \bigcup_{\lambda>0} \lambda\partial^a \text{dist}_\Omega(x^*).$$

The A-normal cone to Ω at x^* is defined by

$$\mathcal{N}^a_\Omega(x^*) := \partial^a\delta(x^*, \Omega).$$

It follows from [36, Proposition 3.4], [61, Section 2.5.2, Page 238], and [36, Proposition 3.3] that

$$\mathcal{N}^c_\Omega(x^*) = cl^*conv\mathcal{N}^g_\Omega(x^*), \quad \mathcal{N}_\Omega(x^*) \subseteq \tilde{\mathcal{N}}^g_\Omega(x^*), \quad \text{and} \quad \mathcal{N}^g_\Omega(x^*) \subseteq \mathcal{N}^a_\Omega(x^*).$$

Clearly,

$$\mathcal{N}_\Omega(x^*) \subseteq \tilde{\mathcal{N}}^g_\Omega(x^*) \subseteq \mathcal{N}^g_\Omega(x^*) \subseteq \mathcal{N}^c_\Omega(x^*).$$

If Ω is convex, then $\mathcal{N}(x^*) = \tilde{\mathcal{N}}^g_\Omega(x^*) = \mathcal{N}^g(x^*) = \mathcal{N}^c(x^*)$ is the normal cone of Ω at x^* in the sense of convex analysis.

Now we introduce the Clarke subdifferential of locally Lipschitian functions. In this paragraph we assume that φ is Lipschitian near x^*. Recall that the Clarke's generalized derivative of φ at x^* along the direction d is defined by

$$\varphi^o(x^*, d) := \limsup_{x\to x^*, t\downarrow 0} \frac{\varphi(x+td) - \varphi(x)}{t}.$$

The Clarke subdifferential of φ at x^* is defined by

$$\partial^c\varphi(x^*) := \{v \in \mathbb{X}^* \mid \langle v, d\rangle \leq \varphi^o(x^*, d) \quad \forall d \in \mathbb{X}\}.$$

In the general Banach space setting, we have

$$cl^* conv \partial\varphi(x^*) \subseteq \partial^c \varphi(x^*),$$

where the inclusion relationship above holds with equality when $\mathbb{X}$ is an Asplund space. It follows from [61, Section 2.5.2, Page 238] and [35, Proposition 3.3] that

$$\partial\varphi(x^*) \subseteq \partial^a \varphi(x^*) \quad \text{and} \quad \partial^a \varphi(x^*) \subseteq \partial^c \varphi(x^*).$$

If, in addition, φ is convex, then $\partial^a \varphi(x) = \partial\varphi(x) = \partial^c \varphi(x)$ is the same as the subdifferential of φ at x^* in the sense of convex analysis.

The following propositions provide a summary of some of the important properties of the approximate subdifferential, see [24, 35, 36, 41, 61]. For a set-valued map $S : \mathbb{X} \rightrightarrows \mathbb{X}^*$, we say S is closed if its graph is closed in the appropriate topology.

Proposition 3.2.1. *Let $f : \mathbb{X} \to \mathbb{R}$ be Lipschitzian near x^* with positive modulus L_f. Then the following results hold:*

(i) *[35, Proposition 3.3] and [16, Proposition 2.1.2] $cl^* conv \partial^a f(x^*) = \partial^c f(x^*) \subseteq L_f \mathbb{B}_{\mathbb{X}^*}$.*

(ii) *[24, Theorem 1.1] If x^* is a local minimizer of f on $\mathbb{X}$, then $0 \in \partial^a f(x^*)$.*

Proposition 3.2.2. *[24, Theorem 1.4] Let $f : \mathbb{X} \to \mathbb{R}$ be Lipschitzian near x^*. Then the set-valued map*

$$(\lambda, x) \to \partial^a (\lambda f)(x)$$

is closed at (λ^*, x^*), i.e.,

$$\partial^a (\lambda^* f)(x^*) = \operatorname*{Lim\,sup}_{\lambda \to \lambda^*, x \xrightarrow{f} x^*} \partial^a (\lambda f)(x) \quad \forall \lambda^* \in \mathbb{R}.$$

For a locally Lipschitzian function in WCG Asplund spaces, at each point the limiting subdifferential set coincides with the approximate subdifferential set [61, Theorem 3.59]. Thus, the limiting subdifferential enjoys the robust property as in Proposition 3.2.2 in the WCG Asplund setting. Note that even for a locally Lipschitzian function, the limiting subdifferential might not enjoy the robustness property in a non-WCG Banach space (see [61, Example 3.61]).

Proposition 3.2.3. *[35, Proposition 2.3] The A-normal cone mapping* $\mathcal{N}^a_\Omega(\cdot) = \partial^a\delta(\cdot,\Omega)$ *is closed, i.e.,*

$$\mathcal{N}^a_\Omega(x^*) = \operatorname*{Lim\,sup}_{x\xrightarrow{\Omega}x^*}\mathcal{N}^a_\Omega(x) \qquad \forall x^* \in \Omega.$$

Proposition 3.2.4 (Calculus rules)**.** *(i) [35, Corollary 4.1.1] Let* $f, g : \mathbb{X} \to \mathbb{R} \cup \{+\infty\}$ *be lower semicontinuous near* x^**, finite at* x^**, and at least one of them is Lipschitzian near* $x^* \in \mathbb{X}$*. Let* α, β *be positive scalars. Then*

$$\partial^a(\alpha f + \beta g)(x^*) \subseteq \alpha\partial^a f(x^*) + \beta\partial^a g(x^*),$$

where $\gamma \cdot \emptyset := \emptyset$ *for any nonzero scalar* γ*.*

(ii) [41, Theorem 2.5 and Remark (2)] Let $\varphi : \mathbb{X} \to \mathbb{Y}$ *be Lipschitzian near* x^* *and* $f : \mathbb{Y} \to \mathbb{R}$ *be Lipschitzian near* $\varphi(x^*)$*. Then* $f \circ \varphi$ *is Lipschitzian near* x^* *and*

$$\partial^a(f \circ \varphi)(x^*) \subseteq \cup_{\xi\in\partial^a f(\varphi(x^*))}\partial^a\langle \xi, \varphi\rangle(x^*).$$

(iii) [24, Corollary 1.2] Let $f_i : \mathbb{X} \to \mathbb{R}$ $(i = 1, \dots, n)$ *be Lipschitzian near* x^* *and* $f(x) := \max\{f_i(x) \mid i = 1, \dots, n\}$*. Then* $f(x)$ *is Lipschitzian near* $x^* \in \mathbb{X}$ *and*

$$\partial^a f(x^*) \subseteq conv\{\partial^a f_i(x^*) \mid i \in \mathcal{I}(x^*)\},$$

where $\mathcal{I}(x^*) := \{i \mid f_i(x^*) = f(x^*)\}$ *is the set of active indices.*

3.3 Enhanced Fritz John condition

For nonsmooth problem (MPGC), the classical Fritz John necessary optimality condition is generalized to one where the classical gradient is replaced by the limiting subdifferential (see Mordukhovich [61]) and the Clarke subdifferential (see Clarke [16]), respectively. The following theorem strengthens the classical Fritz John condition (i.e., conditions (i)-(ii) of Theorem 3.1) through a stronger sequential condition (iii) of Theorem 3.1, and hence their effectiveness has been significantly enhanced. Taking limits in (3.5) it is easy to see that condition (ii) is included in condition (iii). In order to emphasize the enhanced properties, however, we keep the redundant condition (ii) in Theorem 3.1. Note that the following result depends on the chosen basis $\mathscr{E} = \{e_1, \ldots, e_m\}$ and, since $\mathbb{Y}$ is assumed to be finite dimensional, the limiting normal cone of Λ coincides with the nucleus of the G-normal cone of Λ at any point [61, Theorem 3.59(ii)].

Theorem 3.1. *Let* x^* *be a local minimizer of problem* (MPGC)*. Then for all choices of bases* $\mathscr{E}$*, there exist a scalar* $r \geq 0$ *and a vector* $\eta^* \in \mathbb{Y}$ *not all zero, such that the following conditions hold:*

(i) $0 \in r\partial^a f(x^*) + \sum_{i=1}^m \partial^a \langle \eta^*, e_i \rangle \langle F, e_i \rangle (x^*) + \tilde{\mathcal{N}}_\Omega^g(x^*)$*;*

(ii) $\eta^* \in \mathcal{N}_\Lambda(F(x^*))$;

(iii) *If the index set* $\mathcal{I} := \{i \mid \langle \eta^*, e_i \rangle \neq 0\}$ *is nonempty, then there exists a sequence*

$\{(x^k, y^k, \eta^k)\} \subseteq \Omega \times \Lambda \times \mathbb{Y}$ *converging to* $(x^*, F(x^*), \eta^*)$ *such that for all* k,

$$f(x^k) < f(x^*),$$

$$\eta^k \in \mathcal{N}_\Lambda(y^k), \tag{3.5}$$

$$\langle \eta^*, e_i \rangle \langle F(x^k) - y^k, e_i \rangle > 0 \quad \forall i \in \mathcal{I}. \tag{3.6}$$

Proof. Without loss of generality we may assume that x^* is a global minimizer of problem (MPGC). First of all, we observe that if x^* is a local minimizer of the problem

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & x \in \Omega, \end{aligned} \tag{3.7}$$

then by the Clarke exact penalty principle [16, Proposition 2.4.3], there exists $\kappa > 0$ such that x^* is a local minimizer for

$$\min \quad f(x) + \kappa \text{dist}_\Omega(x).$$

Then by Proposition 3.2.1(ii) and Proposition 3.2.4(i), we have

$$0 \in \partial^a f(x^*) + \kappa \partial \text{dist}_\Omega(x^*) \subseteq \partial^a f(x^*) + \tilde{\mathcal{N}}^g_\Omega(x^*).$$

Hence the proof is complete by letting $r = 1$ and $\eta^* = 0$. In the following, we assume that x^* is not a local minimizer of problem (3.7). By introducing a slack variable $y \in \mathbb{Y}$ for the geometric constraint $F(x) \in \Lambda$, we first reformulate problem (MPGC) as follows:

$$\begin{array}{lll} (\text{MPGC})' & \min & f(x) \\ & \text{s.t.} & F(x) - y = 0, \\ & & x \in \Omega, y \in \Lambda. \end{array}$$

Then (x^*, y^*) with $y^* = F(x^*)$ is a global minimizer for problem $(\text{MPGC})'$. For each $k = 1, 2, \ldots,$ we consider the function $F^k : \mathbb{X} \times \mathbb{Y} \to \mathbb{R}$ defined by

$$F^k(x, y) := \max\{f(x) - f(x^*) + \frac{1}{2k}, |\langle F(x) - y, e_1 \rangle|, \ldots, |\langle F(x) - y, e_m \rangle|\}.$$

Since (x^*, y^*) is a global minimizer of $(\text{MPGC})'$, we have

$$F^k(x, y) > 0 \qquad \forall (x, y) \in \Omega \times \Lambda,$$

which, together with $F^k(x^*, y^*) = \frac{1}{2k}$, implies

$$F^k(x^*, y^*) < \inf_{(x,y) \in \Omega \times \Lambda} F^k(x, y) + \frac{1}{k}. \tag{3.8}$$

Since x^* is not a local minimizer of problem (3.7), then there exists a sequence $\{\tilde{x}^k\}$ with $\tilde{x}^k \in \Omega$ such that $\tilde{x}^k \to x^*$ and

$$f(\tilde{x}^k) < f(x^*) \quad \forall k.$$

Then $\|F(\tilde{x}^k) - y^*\| = \|F(\tilde{x}^k) - F(x^*)\| \to 0$. Thus we can choose the sequence $\{\tilde{x}^k\}$ such that

$$F^k(\tilde{x}^k, y^*) < \frac{1}{2k} = F^k(x^*, y^*). \tag{3.9}$$

This and (3.8) imply

$$F^k(\tilde{x}^k, y^*) < \inf_{(x,y)\in\Omega\times\Lambda} F^k(x,y) + \frac{1}{k}.$$

Clearly, F^k is Lipschitzian near (x^*, y^*) and hence, by the Ekeland's variational principle (see, e.g., [61, Theorem 2.26], [77, Corollary 8.2.6]), there exists $(x^k, y^k) \in \Omega \times \Lambda$ such that

$$\begin{cases} \|(x^k, y^k) - (\tilde{x}^k, y^*)\| \le \frac{1}{\sqrt{k}}, \\ F^k(x^k, y^k) \le F^k(\tilde{x}^k, y^*), \\ F^k(x^k, y^k) \le F^k(x,y) + \frac{1}{\sqrt{k}}\|(x,y) - (x^k, y^k)\| \qquad \forall (x,y) \in \Omega \times \Lambda. \end{cases} \tag{3.10}$$

It follows that $(x^k, y^k) \to (x^*, y^*)$ as $k \to \infty$ and for each k, $(x,y) = (x^k, y^k)$ is a global minimizer of the problem

$$\begin{aligned} \min \qquad & \tilde{F}^k(x,y) := F^k(x,y) + \frac{1}{\sqrt{k}}\|(x,y) - (x^k, y^k)\| \\ \text{s.t.} \qquad & (x,y) \in \Omega \times \Lambda. \end{aligned}$$

Then by the Clarke exact penalty principle, there exists $\kappa \ge 0$ such that $(x,y) = (x^k, y^k)$ is a global minimizer of the problem

$$\begin{aligned} \min \qquad & \tilde{F}^k(x,y) + \kappa \text{dist}_{\Omega\times\Lambda}(x,y) \\ \text{s.t.} \qquad & (x,y) \in \mathbb{X} \times \mathbb{Y}. \end{aligned}$$

Thus, we have by the necessary optimality condition (Proposition 3.2.1(ii)) and calculus rule (Proposition 3.2.4(i)) that

$$0 \in \partial^a \tilde{F}^k(x^k, y^k) + \kappa \partial^a \text{dist}_{\Omega}(x^k) \times \kappa \partial^a \text{dist}_{\Lambda}(y^k). \tag{3.11}$$

Applying the calculus rules (Proposition 3.2.4(iii)), there exist nonnegative scalars $\{r_k, \hat{\eta}_1^k, \ldots, \hat{\eta}_m^k\}$ such that for each k,

$$r^k + \sum_{i=1}^{m} \hat{\eta}_i^k = 1 \tag{3.12}$$

and

$$\begin{aligned} 0 \;&\in\; r^k \begin{pmatrix} \partial^a f(x^k) \\ 0 \end{pmatrix} + \sum_{i=1}^{m} \hat{\eta}_i^k \partial^a |\psi_i|(x^k, y^k) + \frac{1}{\sqrt{k}} \begin{pmatrix} \mathbb{B}_{\mathbb{X}^*} \\ \mathbb{B}_{\mathbb{Y}} \end{pmatrix} \\ &+\; \kappa \begin{pmatrix} \partial^a \mathrm{dist}_\Omega(x^k) \\ \partial^a \mathrm{dist}_\Lambda(y^k) \end{pmatrix}, \end{aligned} \tag{3.13}$$

where $\psi_i(x, y) := \langle F(x) - y, e_i \rangle$ and $\hat{\eta}_i^k = 0$ if i is not an active index. Since the active indices only count in the maximum rule and $F^k(x^k, y^k) > 0$, we may assume that for each k, $\psi_i(x^k, y^k) = \langle F(x^k) - y^k, e_i \rangle = 0$ implies $\hat{\eta}_i^k = 0$, otherwise we can choose a subsequence. Define

$$\tilde{\eta}_i^k := (\mathrm{sign}\, \langle F(x^k) - y^k, e_i \rangle) \hat{\eta}_i^k,$$

where $\mathrm{sign}\, 0 = 0$. We then obtain by the chain rule (Proposition 3.2.4(ii)) that

$$\hat{\eta}_i^k \partial^a |\psi_i|(x^k, y^k) = \begin{pmatrix} \partial^a \langle F, \tilde{\eta}_i^k e_i \rangle (x^k) \\ -\tilde{\eta}_i^k e_i \end{pmatrix}.$$

This and (3.13) imply that

$$0 \in r^k \begin{pmatrix} \partial^a f(x^k) \\ 0 \end{pmatrix} + \sum_{i=1}^{m} \begin{pmatrix} \partial^a \langle F, \tilde{\eta}_i^k e_i \rangle (x^k) \\ -\tilde{\eta}_i^k e_i \end{pmatrix} + \begin{pmatrix} \frac{\mathbb{B}_{\mathbb{X}^*}}{\sqrt{k}} \\ \frac{\mathbb{B}_{\mathbb{Y}}}{\sqrt{k}} \end{pmatrix} + \kappa \begin{pmatrix} \partial^a \mathrm{dist}_\Omega(x^k) \\ \partial^a \mathrm{dist}_\Lambda(y^k) \end{pmatrix}$$

that is,

$$\begin{cases} 0 \in r^k \partial^a f(x^k) + \sum_{i=1}^m \partial^a \tilde{\eta}_i^k \langle F, e_i \rangle (x^k) + \frac{1}{\sqrt{k}} \mathbb{B}_{\mathbb{X}^*} + \kappa \partial^a \mathrm{dist}_\Omega(x^k), \\ \tilde{\eta}^k \in \frac{1}{\sqrt{k}} \mathbb{B}_{\mathbb{Y}} + \kappa \partial^a \mathrm{dist}_\Lambda(y^k), \end{cases} \tag{3.14}$$

where $\tilde{\eta}^k := \sum_{i=1}^m \tilde{\eta}_i^k e_i$.

Since by construction we have $r^k + \sum_{i=1}^m |\tilde{\eta}_i^k| = 1$, the sequence $\{(r^k, \tilde{\eta}_1^k, \ldots, \tilde{\eta}_m^k)\}$ is bounded and must contain a subsequence that converges to some limit $(r, \bar{\eta}_1, \ldots, \bar{\eta}_m)$, where $r \geq 0$ and $(r, \bar{\eta}_1, \ldots, \bar{\eta}_m) \neq 0$. By virtue of the closedness of the subdifferential (Proposition 3.2.2), it follows from (3.14) that

$$\begin{cases} 0 \in r \partial^a f(x^*) + \sum_{i=1}^m \partial^a \langle \eta^*, e_i \rangle \langle F, e_i \rangle (x^*) + \kappa \partial^a \mathrm{dist}_\Omega(x^*), \\ \eta^* \in \kappa \partial^a \mathrm{dist}_\Lambda(y^*), \end{cases} \tag{3.15}$$

where $\eta^* = \sum_{i=1}^m \bar{\eta}_i e_i$. Thus,

$$\begin{cases} 0 \in r \partial^a f(x^*) + \sum_{i=1}^m \partial^a \langle \eta^*, e_i \rangle \langle F, e_i \rangle (x^*) + \tilde{\mathcal{N}}_\Omega^g(x^*), \\ \eta^* \in \tilde{\mathcal{N}}_\Lambda^g(F(x^*)). \end{cases}$$

To show that condition (iii) is satisfied, assume that $\mathcal{I} \neq \emptyset$ (otherwise there is nothing to prove). Since

$$\tilde{\eta}^k \in \frac{1}{\sqrt{k}} \mathbb{B}_{\mathbb{Y}} + \kappa \partial^a \mathrm{dist}_\Lambda(y^k) \subseteq \frac{1}{\sqrt{k}} \mathbb{B}_{\mathbb{Y}} + \tilde{\mathcal{N}}_\Lambda^g(y^k),$$

there exists $\rho^k \in \mathbb{B}_{\mathbb{Y}}$ such that

$$\eta^k := \tilde{\eta}^k + \frac{1}{\sqrt{k}} \rho^k \in \tilde{\mathcal{N}}_\Lambda^g(y^k).$$

Since $\tilde{\eta}^k \to \eta^*$ as $k \to \infty$, it is easy to see that $\eta^k \to \eta^*$. Since $\tilde{\eta}_i^k \to \bar{\eta}_i = \langle \eta^*, e_i \rangle \neq 0$

for each $i \in \mathcal{I}$, $\tilde{\eta}_i^k$ has the same sign as $\langle \eta^*, e_i \rangle$ for sufficiently large k. Hence we must have $\langle \eta^*, e_i \rangle \tilde{\eta}_i^k > 0$ for all $i \in \mathcal{I}$ and sufficiently large k. By the definition, $\tilde{\eta}_i^k$ have the same sign as $\langle F(x^k) - y^k, e_i \rangle$, therefore we must have $\langle \eta^*, e_i \rangle \langle F(x^k) - y^k, e_i \rangle > 0$ for all $i \in \mathcal{I}$ and sufficiently large k. Moreover, it follows from the definition of F^k and (3.9)–(3.10) that

$$\begin{aligned} f(x^k) - f(x^*) + \frac{1}{2k} &\leq F^k(x^k, y^k) \\ &\leq F^k(\tilde{x}^k, y^*) \\ &< \frac{1}{2k} \end{aligned}$$

and hence $f(x^k) < f(x^*)$. The proof is complete by noting that the limiting normal cone of Λ coincides with the nucleus of the G-normal cone of Λ at any point in the finite dimensional setting [61, Theorem 3.59(ii)]. □

Since for any function φ and set S, it must hold that (see, e.g., [36, Proposition 3.4])

$$\partial^g \varphi(x) \subseteq \partial^c \varphi(x) \text{ and } \tilde{\mathcal{N}}_S^g(x) \subseteq \mathcal{N}_S^c(x),$$

the following holds immediately.

Corollary 3.2. *Let x^* be a local minimizer of problem* (MPGC)*. Then there exist a scalar $r \geq 0$ and a vector $\eta^* \in \mathbb{Y}$ not all zero, such that conditions (ii)–(iii) of Theorem 3.1 hold and*

$$0 \in r\partial^c f(x^*) + \sum_{i=1}^{m} \partial^c \langle \eta^*, e_i \rangle \langle F, e_i \rangle (x^*) + \mathcal{N}_\Omega^c(x^*).$$

Since in the WCG Asplund space setting, the limiting subdifferential and limiting normal cone coincide with the approximate subdifferential and the nucleus of the G-normal cone respectively [61, Theorem 3.59], we have the following result immediately.

Corollary 3.3. *Assume that $\mathbb{X}$ is a WCG Asplund space. Let x^* be a local minimizer of problem* (MPGC)*. Then there exist a scalar $r \geq 0$ and a vector $\eta^* \in \mathbb{Y}$ not all zero, such that conditions (ii)–(iii) of Theorem 3.1 hold and*

$$0 \in r\partial f(x^*) + \sum_{i=1}^{m} \partial\langle \eta^*, e_i\rangle\langle F, e_i\rangle(x^*) + \mathcal{N}_\Omega(x^*).$$

We now specialize Theorem 3.1 to problem $\mathrm{NLP}_{\mathrm{Banach}}$ where $f : \mathbb{X} \to \mathbb{R}, h : \mathbb{X} \to \mathbb{R}^p, g : \mathbb{X} \to \mathbb{R}^q$ are Lipschitzian near the optimal solution and Ω is a nonempty closed subset of $\mathbb{X}$. Let

$$F(x) := (h(x), g(x)) \quad \text{and} \quad \Lambda := \{0\}^p \times \mathbb{R}^q_-, \tag{3.16}$$

By virtue of Theorem 3.1, we are now able to establish the enhanced Fritz John condition for the nonsmooth nonlinear programs in a Banach space, which improves Theorem 2.1 in chapter 2. Note that the set Λ in (3.16) is a convex cone.

Corollary 3.4. *Let x^* be a local minimizer of problem* $\mathrm{NLP}_{\mathrm{Banach}}$*. Then there exist $r \geq 0, \lambda^* \in \mathbb{R}^p, \mu^* \in \mathbb{R}^q$ not all zero, such that*

(a) $0 \in r\partial^a f(x^*) + \sum_{i=1}^p \partial^a(\lambda_i^* h_i)(x^*) + \sum_{j=1}^q \mu_j^* \partial^a g_j(x^*) + \tilde{\mathcal{N}}_\Omega^g(x^*)$;

(b) $0 \leq -g(x^*) \perp \mu^* \geq 0$;

(c) *If $(\lambda^*, \mu^*) \neq 0$, then there exists a sequence $\{x^k\} \subseteq \Omega$ converging to x^* such that for all k, $f(x^k) < f(x^*)$ and*

$$\lambda_i^* \neq 0 \Longrightarrow \lambda_i^* h_i(x^k) > 0, \;\; \mu_j^* > 0 \Longrightarrow g_j(x^k) > 0.$$

Proof. Letting F and Λ be defined as in (3.16), it is not hard to see from Theorem 3.1 and the explicit expression for the normal cone $\mathcal{N}_\Lambda(F(x^*))$ that there exist $r \geq 0, \lambda^* \in$

$\mathbb{R}^p, \mu^* \in \mathbb{R}^q$ not all zero, such that conditions (a)–(b) hold and there exists a sequence $\{(x^k, \hat{y}^k, \tilde{y}^k, \lambda^k, \mu^k)\} \in \Omega \times \{0\}^p \times \mathbb{R}^q_- \times \mathbb{R}^p \times \mathbb{R}^q$ converging to $(x^*, h(x^*), g(x^*), \lambda^*, \eta^*)$ such that for all k, $f(x^k) < f(x^*)$,

$$(\lambda^k, \mu^k) \in \mathcal{N}_{\{0\}^p \times \mathbb{R}^q_-}(\hat{y}^k, \tilde{y}^k), \tag{3.17}$$

and

$$\langle \eta^*, e_i \rangle \neq 0 \Longrightarrow \langle \eta^*, e_i \rangle \langle F(x^k) - y^k, e_i \rangle > 0, \tag{3.18}$$

where $\eta^* := (\lambda^*, \mu^*)$ and $F(x^k) := (h(x^k), g(x^k))$. Since $\hat{y}^k = 0$, it is easy to see from (3.18) that

$$\lambda_i^* \neq 0 \Longrightarrow \lambda_i^* h_i(x^k) > 0.$$

If $\mu_j^* > 0$, then it follows from (3.18) that $g_j(x^k) > \tilde{y}_j^k$. We next show that there exists a subsequence $\{\tilde{y}_j^{k_\imath}\}_{\imath \in \mathbb{N}}$ such that $\tilde{y}_j^{k_\imath} = 0 \; \forall \imath \in \mathbb{N}$. Assume to the contrary that $\tilde{y}_j^k < 0$ for all sufficiently large k and then it follows from (3.17) that $\mu_j^k = 0$, which implies that $\mu_j^* = 0$ by taking a limit as $k \to \infty$. This contradicts assumption $\mu_j^* > 0$ and hence we have

$$\mu_j^* > 0 \Longrightarrow g_j(x^{k_\imath}) > 0 \quad \forall \imath \in \mathbb{N}.$$

Therefore, condition (c) also holds by choosing and resetting this subsequence. The proof is complete. □

Our next task is to specialize our result to the nonlinear semidefinite program:

$$\text{(NLSDP)} \qquad \min_{x \in \mathbb{X}} \quad f(x) \qquad \text{s.t.} \quad H(x) \in \mathcal{S}^l_-,$$

where $f : \mathbb{X} \to \mathbb{R}$ and $H : \mathbb{X} \to \mathcal{S}^l$, $\mathcal{S}^l$ is the linear space of all $l \times l$ real symmetric matrices equipped with the usual Frobenius inner product $\langle \cdot , \cdot \rangle$, and $\mathcal{S}^l_-$ is the cone of all $l \times l$ negative semidefinite matrices in $\mathcal{S}^l$. Note that for simplicity, we omit the usual equality and inequality constraints since they can be handled as in the usual nonlinear program. For $A \in \mathcal{S}^l$, we denote by $\lambda(A) \in \mathbb{R}^l$ the vector of its eigenvalues ordered in a decreasing order as follows:

$$\lambda_1(A) \geq \cdots \geq \lambda_l(A).$$

Clearly, (NLSDP) is equivalent to the problem

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & \lambda_1(H(x)) \leq 0. \end{array} \tag{3.19}$$

For $A \in \mathcal{S}^l$, the notation diag($\lambda(A)$)$\in \mathcal{S}^l$ is used for the diagonal matrix with the vector $\lambda(A)$ on the main diagonal. It is known that any $A \in \mathcal{S}^l$ admits an eigenvalue decomposition as follows:

$$A = U\text{diag}(\lambda(A))U^T$$

with a square orthogonal matrix $U = U(A)$ such that $U^TU = I$ whose columns are eigenvectors of A. Let $u_i(A)$ be the ith column of matrix $U(A)$. Note that since λ_1 is convex (see, e.g., [51, Proposition 1.1]), the approximate subdifferential coincides with the subdifferential in the sense of convex analysis.

Lemma 3.5. [45, 69] *The subdifferential of* $\lambda_1(A) : \mathcal{S}^l \to \mathbb{R}$ *in the sense of convex* *analysis is given by*

$$\begin{aligned}\partial^a \lambda_1(A) &= conv\{u_i(A)u_i(A)^T \mid i = 1, \ldots, d(A)\} \\ &= \{\sum_{i=1}^{d(A)} \tau_i u_i(A)u_i(A)^T \mid \sum_{i=1}^{d(A)} \tau_i = 1, \tau_i \geq 0 \ i = 1, \ldots, d(A)\},\end{aligned}$$

where $d(A)$ is the multiplicity of the largest eigenvalue of the matrix A.

We get the following results immediately by applying Corollary 3.4 to the problem (3.19). Note that we let $\mathcal{S}^l_+ = -\mathcal{S}^l_-$.

Corollary 3.6. *Assume that x^* is a local minimizer of problem* (NLSDP)*. Then there exist $r \geq 0$ and $\Gamma^* \in \mathcal{S}^l_+$, which are not both zero, such that*

(a) $0 \in r\partial^a f(x^*) + \partial^a \langle \Gamma^*, H\rangle(x^*)$;

(b) $\Gamma^* \in \mathcal{S}^l_+$, $\langle \Gamma^*, H(x^*)\rangle = 0$;

(c) *If $\Gamma^* \neq 0$, then there exists a sequence $\{x^k\}$ converging to x^* such that for all k, $f(x^k) < f(x^*)$ and $\lambda_1(H(x^k)) > 0$.*

Proof. Since x^* is a local minimizer of problem (3.19), it follows from Corollary 3.4 that there exist $\{r, \mu^*\}$ such that $(r, \mu^*) \neq 0$ and

(i) $0 \in r\partial^a f(x^*) + \mu^* \partial^a(\lambda_1 \circ H)(x^*)$;

(ii) $r \geq 0$, $0 \leq -\lambda_1(H(x^*)) \perp \mu^* \geq 0$;

(iii) If $\mu^* \neq 0$, then there exists a sequence $\{x^k\} \subseteq \mathbb{X}$ converging to x^* such that for all k, $f(x^k) < f(x^*)$ and $\lambda_1(H(x^k) > 0$.

It follows from Proposition 3.2.4(ii), Lemma 3.5, and (i) above that there exists

$$\begin{aligned}\Gamma^* &= \mu^* \sum_{i=1}^{d(H(x^*))} \tau_i^* u_i(H(x^*))u_i(H(x^*))^T \\ &\in \mu^* conv\{u_i(H(x^*))u_i(H(x^*))^T \mid i = 1, \ldots, d(H(x^*))\}\end{aligned}$$

such that

$$0 \in r\partial^a f(x^*) + \partial^a \langle \Gamma^*, H\rangle(x^*), \tag{3.20}$$

where $d(H(x^*))$ is the multiplicity of the largest eigenvalue of the matrix $H(x^*)$.

It is easy to see that $\Gamma^* \in \mathcal{S}^l_+$ and, from the definition of $\partial^a \lambda_1(H(x^*))$ and (ii)–(iii) above of this proof that

$$\begin{aligned}
\langle \Gamma^*, H(x^*)\rangle &= \Big\langle \mu^* \sum_{i=1}^{d(H(x^*))} \tau_i^* u_i(H(x^*))u_i(H(x^*))^T, H(x^*)\Big\rangle \\
&= \mu^* \sum_{i=1}^{d(H(x^*))} \tau_i^* \Big\langle u_i(H(x^*))u_i(H(x^*))^T, H(x^*)\Big\rangle \\
&= \mu^* \sum_{i=1}^{d(H(x^*))} \tau_i^* \Big\langle 1, u_i(H(x^*))^T H(x^*) u_i(H(x^*))\Big\rangle \\
&= \mu^* \lambda_1(H(x^*)) \sum_{i=1}^{d(H(x^*))} \Big(\tau_i^* u_i(H(x^*))^T u_i(H(x^*))\Big) \\
&= \mu^* \lambda_1(H(x^*)) \\
&= 0.
\end{aligned}$$

Then, conditions (a) and (b) in this corollary hold. We next show condition (c). From the definition of Γ^*, we have that

$$\Gamma^* \neq 0 \Longleftrightarrow \mu^* \neq 0.$$

Then from (iii) above we have the desired result. The proof is complete. $\square$

3.4 Enhanced KKT condition and weaker constraint qualification

Based on the enhanced Fritz John condition for problem (MPGC) in the previous section, we define the following enhanced KKT condition for problem (MPGC). We denote by $\mathcal{N}^e_\Lambda(F(x^*))$ the set of elements in the normal cone $\eta^* \in \mathcal{N}_\Lambda(F(x^*))$ such that there exists a sequence $\{(x^k, y^k, \eta^k)\} \subseteq \Omega \times \Lambda \times \mathbb{Y}$ converging to $(x^*, F(x^*), \eta^*)$ such that for all k,

$$\begin{aligned} &\eta^k \in \mathcal{N}_\Lambda(y^k) \\ &\langle \eta^*, e_i \rangle \neq 0 \Longrightarrow \langle \eta^*, e_i \rangle \langle F(x^k) - y^k, e_i \rangle > 0. \end{aligned}$$

Note that in this above case, if $\eta^* = 0$, then the existence of the approximate sequence is trivial.

Definition 7 (Enhanced KKT point)**.** Let x^* be a feasible point of the problem (MPGC).

(a) We say that x^* is an *enhanced KKT point* if there exists $\eta^* \in \mathcal{N}_\Lambda(F(x^*))$ such that

(i) $0 \in \partial^a f(x^*) + \sum_{i=1}^m \partial^a \langle \eta^*, e_i \rangle \langle F, e_i \rangle (x^*) + \tilde{\mathcal{N}}^g_\Omega(x^*)$,

(ii) If $\langle \eta^*, e_i \rangle \neq 0$, then there exists a sequence $\{x^k, y^k, \eta^k\} \subseteq \Omega \times \Lambda \times \mathbb{Y}$ converging to $(x^*, F(x^*), \eta^*)$ such that for all k,

$$\begin{aligned} &f(x^k) < f(x^*), \\ &\eta^k \in \mathcal{N}_\Lambda(y^k), \\ &\langle \eta^*, e_i \rangle \neq 0 \Longrightarrow \langle \eta^*, e_i \rangle \langle F(x^k) - y^k, e_i \rangle > 0. \end{aligned}$$

(b) We say that x^* is a *weaker enhanced KKT point* if there exists $\eta^* \in \mathcal{N}^e_\Lambda(F(x^*))$ such that (i) above holds.

It is clear that an enhanced KKT point is a weaker enhanced KKT point.

Definition 8. Let $x^* \in \mathscr{F}$.

(a) x^* is said to satisfy the *no nonzero abnormal multiplier constraint qualification* (NNAMCQ) if there is no nonzero vector $\eta^* \in \mathcal{N}_\Lambda(F(x^*))$ such that

$$0 \in \sum_{i=1}^{m} \partial^a \langle \eta^*, e_i \rangle \langle F, e_i \rangle (x^*) + \tilde{\mathcal{N}}^g_\Omega(x^*); \tag{3.21}$$

(b) x^* is said to be *pseudonormal* for $\mathscr{F}$ if there is no vector $\eta^* \in \mathcal{N}_\Lambda(F(x^*))$ such that (3.21) holds and there exists a sequence $\{(x^k, y^k, \eta^k)\} \subseteq \Omega \times \Lambda \times \mathbb{Y}$ converging to $(x^*, F(x^*), \eta^*)$ such that for each k,

$$\eta^k \in \mathcal{N}_\Lambda(y^k) \text{ and } \langle \eta^*, F(x^k) - y^k \rangle > 0;$$

(c) x^* is said to be *quasinormal* for $\mathscr{F}$ if there is no nonzero vector $\eta^* \in \mathcal{N}^e_\Lambda(F(x^*))$ such that (3.21) holds;

(d) x^* is said to satisfy the *enhanced Guignard constraint qualification* (EGCQ) if F is Fréchet differentiable at x^* and

$$\widehat{\mathcal{N}}_{\mathscr{F}}(x^*) \subseteq \nabla F(x^*)^* \mathcal{N}^e_\Lambda(F(x^*)) + \tilde{\mathcal{N}}^g_\Omega(x^*).$$

The relationships among the first three constraint qualifications are obvious:

$$\text{NNAMCQ} \Longrightarrow \text{pseudonormality} \Longrightarrow \text{quasinormality}.$$

The enhanced KKT condition under the quasinormality follows immediately from Theorem 3.1 and the definition of the quasinormality.

Theorem 3.7. *Let x^* be a local minimizer of problem* (MPGC)*. Suppose that x^* is quasinormal. Then x^* is an enhanced KKT point.*

We now make some comments on the EGCQ. It is well-known that $\mathcal{T}_{\mathscr{F}}(x^*)^o = \widehat{\mathcal{N}}_{\mathscr{F}}(x^*)$ in a finite dimensional space. We next consider the case of standard nonlinear constraints, i.e., $\mathcal{X} := \{x \in \Omega \mid F(x) \in \Lambda\}$ with $\Omega = \mathbb{R}^n$, $F(x)$ and Λ are defined as in (3.16). In this case,

$$\mathcal{L}_{\mathscr{F}}(x^*)^o = \nabla F(x^*)^* \mathcal{N}_\Lambda(F(x^*)),$$

where

$$\mathcal{L}_{\mathscr{F}}(x^*) := \{d \mid \nabla F(x^*)d \in \mathcal{T}_\Lambda(F(x^*))\}$$

is the linearized cone of $\mathscr{F}$ at x^*. Since the inclusion $\mathcal{N}^e_\Lambda(F(x^*)) \subseteq \mathcal{N}_\Lambda(F(x^*))$ may hold strictly, in the case of standard nonlinear constraints, the EGCQ is stronger than the condition $\mathcal{T}_{\mathscr{F}}(x^*)^o \subseteq \mathcal{L}_{\mathcal{X}}(x^*)^o$, which is the so-called Guignard constraint qualification (GCQ).

Next we show that the quasinormality implies the EGCQ in the case where $\mathbb{X}$ admits a Fréchet smooth renorm [61, Page 35]. To this end, we first show that the EGCQ is the weakest constraint qualification for weaker enhanced KKT points when the objective is Fréchet smooth in a Banach space.

Lemma 3.8. [61, Theorem 1.30] *Assume that $\mathbb{X}$ admits a Fréchet smooth renorm. Then for every $d \in \widehat{\mathcal{N}}_S(x^*)$, there is a concave Fréchet smooth function $\varphi : \mathbb{X} \to \mathbb{R}$ that achieves its global maximum relative to S uniquely at x^* and such that $\nabla\varphi(x^*) = d$.*

Theorem 3.9. *Suppose that $x^* \in \mathscr{F}$ is a local minimizer for the optimization problem $\min_{x\in\mathscr{F}} \theta(x)$, where θ is Fréchet differentiable at x^*, and*

$$\widehat{\mathcal{N}}_{\mathscr{F}}(x^*) \subseteq \nabla F(x^*)^* \mathcal{N}^e_\Lambda(F(x^*)) + \tilde{\mathcal{N}}^g_\Omega(x^*). \tag{3.22}$$

Then, x^* must be a weaker enhanced KKT point of $\min_{x\in\mathscr{F}} \theta(x)$. Conversely, assume that $\mathbb{X}$ admits a Fréchet smooth renorm and $x^* \in \mathscr{F}$ is a weaker enhanced KKT point of $\min\limits_{x\in\mathscr{F}} \theta(x)$ for any convex Fréchet smooth function θ at x^* with x^* being a local minimizer, then (3.22) holds.

Proof. Let x^* be locally optimal for problem $\min\limits_{x\in\mathscr{F}} \theta(x)$. Then it follows from [61, Proposition 5.1] that $-\nabla\theta(x^*) \in \widehat{\mathcal{N}}_{\mathscr{F}}(x^*)$. Thus if (3.22) holds, then

$$-\nabla\theta(x^*) \in \nabla F(x^*)^* \mathcal{N}^e_\Lambda(F(x^*)) + \tilde{\mathcal{N}}^g_\Omega(x^*)$$

and hence x^* is an enhanced KKT point of $\min\limits_{x\in\mathscr{F}} \theta(x)$.

Conversely suppose that if $x^* \in \mathscr{F}$ is a local minimizer for an optimization problem $\min\limits_{x\in\mathscr{F}} \theta(x)$ with convex Fréchet smooth objective functions, then x^* must be a weaker enhanced KKT point of the problem. Let $d \in \widehat{\mathcal{N}}_{\mathscr{F}}(x^*)$. By Lemma 3.8, there exists a convex Fréchet smooth function φ such that $-\nabla\varphi(x^*) = d$ and $\text{argmin}_{x\in\mathscr{F}}\varphi(x) = \{x^*\}$. It follows that x^* is a weaker enhanced KKT point of $\min\limits_{x\in\mathscr{F}} \varphi(x)$, i.e.,

$$-\nabla\varphi(x^*) \in \nabla F(x^*)^* \mathcal{N}^e_\Lambda(F(x^*)) + \mathcal{N}_\Omega(x^*).$$

Thus, $d = -\nabla\varphi(x^*) \in \nabla F(x^*)^* \mathcal{N}^e_\Lambda(F(x^*)) + \tilde{\mathcal{N}}^g_\Omega(x^*)$. Therefore, by the arbitrariness of $d \in \widehat{\mathcal{N}}_{\mathscr{F}}(x^*)$, (3.22) holds. □

The following result follows from Theorem 3.9.

Corollary 3.10. *Assume that $\{f, F\}$ are Fréchet differentiable at x^*. If x^* is a* *local minimizer of problem (MPGC) and the EGCQ holds at x^*, then x^* is a weaker enhanced KKT point.*

Corollary 3.11. *Assume that $\mathbb{X}$ admits a Fréchet smooth renorm and F is Fréchet differentiable at x^*. Then the quasinormality implies the EGCQ.*

Proof. It follows from Theorem 3.7 that for any locally Lipshitzian objective function f, if a local minimizer satisfies the quasinormality, then it is an enhanced KKT point. Since a convex Fréchet smooth function is locally Lipschitzian ([10, Proposition 2.107]), it follows from Theorem 3.9 that the EGCQ holds at this point. □

3.5 Error bound and exact penalty

In this section, we prove that a local error bound exists under the quasinormality in the general Banach space. Our results are new even for the finite dimensional space.

For nonsmooth finite dimensional (NLP) problem, the existence of a local error bound has been proved under the pseudonormality or under the quasinormality with extra regularity conditions on the constraint functions in chapter 2, where [79, Theorem 3.1] plays a significant role. In this section, we show that the quasinormality alone implies the existence of a local error bound without imposing any regularity conditions. We first establish the following estimate, which will lead to the possibility of applying [79, Theorem 3.1].

Lemma 3.12. *Let x^* be feasible for problem (MPGC) and*

$$\Phi(x,y) := \max_{1\le i\le m}\{|\psi_i(x,y)|\} \quad \text{with} \quad \psi_i(x,y) := \langle F(x)-y, e_i\rangle.$$

If x^* is quasinormal, then there exist $\delta > 0$ and $c > 0$ such that for all $(\xi, \upsilon) \in$

$\partial^a(\Phi + \mathrm{dist}_{\Omega\times\Lambda})(x, y)$ *with* $(x, y) \in \mathcal{B}_\delta(x^*, F(x^*)) \cap (\Omega \times \Lambda)$ *and* $x \notin \mathscr{F}$,

$$\|(\xi, \upsilon)\| \geq c.$$

Proof. Suppose to the contrary that there exists a sequence $\{(x^k, y^k)\} \subseteq \Omega \times \Lambda$ converging to $(x^*, F(x^*))$ with $x^k \notin \mathscr{F}$ and $(\xi^k, \upsilon^k) \in \partial^a(\Phi + \mathrm{dist}_{\Omega\times\Lambda})(x^k, y^k)$ such that $\|(\xi^k, \upsilon^k)\| \to 0$. Since $F(x^k) \notin \Lambda$ and $y^k \in \Lambda$ for all k, we have $\|F(x^k) - y^k\| > 0$ and hence $\Phi(x^k, y^k) > 0$. By the sum rule Proposition 3.2.4(i), we have

$$(\xi^k, \upsilon^k) \in \partial^a\Phi(x^k, y^k) + \partial^a\mathrm{dist}_\Omega(x^k) \times \partial^a\mathrm{dist}_\Lambda(y^k). \tag{3.23}$$

Since F is assumed to be locally Lipschitzian, applying the maximum rule (Proposition 3.2.4(iii)) in calculating the subdifferential of $\Phi(x, y) := \max_{1\leq i\leq m}\{|\psi_i(x, y)|\}$ at (x^k, y^k) yields the existence of nonnegative scalars $\{\hat{\mu}_1^k, \ldots, \hat{\mu}_m^k\}$ such that

$$\sum_{i=1}^m \hat{\mu}_i^k = 1 \quad \text{and} \quad \partial^a\Phi(x^k, y^k) \subseteq \sum_{i=1}^m \hat{\mu}_i^k \partial^a|\psi_i|(x^k, y^k), \tag{3.24}$$

where $\hat{\mu}_i^k = 0$ if i is not an active index. Since $\Phi(x^k, y^k) > 0$, any $i \in \{1, \ldots, m\}$ such that $\psi_i(x^k, y^k) = 0$ is not an active index. Hence, for all $i = 1, \ldots, m$, $\psi_i(x^k, y^k) = \langle F(x^k) - y^k, e_i\rangle = 0$ implies $\hat{\mu}_i^k = 0$. Define

$$\tilde{\mu}_i^k := (\mathrm{sign}\,\langle F(x^k) - y^k, e_i\rangle)\hat{\mu}_i^k.$$

We then obtain by the chain rule that

$$\hat{\mu}_i^k\partial^a|\psi_i|(x^k, y^k) = \begin{pmatrix} \partial^a\tilde{\mu}_i^k\langle F, e_i\rangle(x^k) \\ -\tilde{\mu}_i^k e_i \end{pmatrix}. \tag{3.25}$$

From (3.23)–(3.25), we obtain

$$\begin{pmatrix} \xi^k \\ \upsilon^k \end{pmatrix} \in \sum_{i=1}^m \begin{pmatrix} \partial^a \tilde{\mu}_i^k \langle F, e_i \rangle (x^k) \\ -\tilde{\mu}_i^k e_i \end{pmatrix} + \begin{pmatrix} \partial^a \mathrm{dist}_\Omega (x^k) \\ \partial^a \mathrm{dist}_\Lambda (y^k) \end{pmatrix},$$

that is,

$$\begin{cases} \xi^k \in \sum_{i=1}^m \partial^a \tilde{\mu}_i^k \langle F, e_i \rangle (x^k) + \partial^a \mathrm{dist}_\Omega (x^k), \\ \upsilon^k \in \sum_{i=1}^m \tilde{\mu}_i^k (-e_i) + \partial^a \mathrm{dist}_\Lambda (y^k). \end{cases} \tag{3.26}$$

Since by the construction $\sum_{i=1}^m |\tilde{\mu}_i^k| = 1$, the sequence $\{(\tilde{\mu}_1^k, \ldots, \tilde{\mu}_m^k)\}$ is bounded and must contain a subsequence that converges to some limit $(\bar{\mu}_1, \ldots, \bar{\mu}_m) \neq 0$. Taking limits as $k \to \infty$, by virtue of the closedness of the subdifferentials (Proposition 3.2.2), it follows from (3.26) that

$$\begin{cases} 0 \in \sum_{i=1}^m \partial^a \langle \mu^*, e_i \rangle \langle F, e_i \rangle (x^*) + \partial^a \mathrm{dist}_\Omega (x^*), \\ \mu^* \in \partial^a \mathrm{dist}_\Lambda (y^*), \end{cases}$$

where $\mu^* := \sum_{i=1}^m \bar{\mu}_i e_i$. Then we have

$$\begin{cases} 0 \in \sum_{i=1}^m \partial^a \langle \mu^*, e_i \rangle \langle F, e_i \rangle (x^*) + \tilde{\mathcal{N}}_\Omega^g (x^*), \\ \mu^* \in \tilde{\mathcal{N}}_\Lambda^g (F(x^*)). \end{cases}$$

Since $\mathbb{Y}$ is finite dimensional, by [61, Theorem 3.59], we have

$$\mu^* \in \tilde{\mathcal{N}}_\Lambda^g (F(x^*)) = \mathcal{N}_\Lambda (F(x^*)).$$

Since $\tilde{\mu}_i^k \to \bar{\mu}_i = \langle \mu^*, e_i \rangle \neq 0$ as $k \to \infty$ for $i \in \mathcal{J}$ where $\mathcal{J} := \{i \mid \langle \mu^*, e_i \rangle \neq 0\}$, $\tilde{\mu}_i^k$ has the same sign as $\langle \mu^*, e_i \rangle$ for sufficiently large k. Hence we must have $\langle \mu^*, e_i \rangle \tilde{\mu}_i^k > 0$

for all $i \in \mathcal{J}$ and sufficiently large k. By the definition, $\tilde{\mu}_i^k$ has the same sign as $\langle F(x^k) - y^k, e_i \rangle$, thus we must have $\langle \mu^*, e_i \rangle \langle F(x^k) - y^k, e_i \rangle > 0$ for all $i \in \mathcal{J}$ and sufficiently large k. Since $\upsilon^k \to 0$, $\mu^k := \sum_{i=1}^m \tilde{\mu}_i^k e_i + \upsilon^k \to \mu^*$ and then it follows from (3.26) and the fact that $\mathbb{Y}$ is finite dimensional that

$$\mu^k \in \tilde{\mathcal{N}}_\Lambda^g(y^k) = \mathcal{N}_\Lambda(y^k).$$

However, these facts above and $\mu^* \neq 0$ imply that the quasinormality is violated at x^* and hence a contradiction. □

Now we are ready to give the main result of this section about the existence of local error bounds.

Theorem 3.13. *Let x^* be feasible for problem (MPGC). Suppose that x^* is quasinormal. Then the local error bound holds, i.e., there exist $\delta_0 > 0$ and $\kappa > 0$ such that*

$$\mathrm{dist}_{\mathscr{F}}(x) \leq \kappa \mathrm{dist}_\Lambda(F(x)) \quad \forall x \in \mathcal{B}_{\delta_0}(x^*) \cap \Omega.$$

Proof. According to Lemma 3.12, there exist constants $\delta > 0$ and $\kappa > 0$ such that for all $(\xi, \upsilon) \in \partial^a(\Phi + \mathrm{dist}_{\Omega \times \Lambda})(x, y)$ with $(x, y) \in (\mathcal{B}_\delta(x^*) \times \mathcal{B}_\delta(F(x^*)) \cap (\Omega \times \Lambda)$ and $x \notin \mathscr{F}$,

$$\|(\xi, \upsilon)\| \geq \frac{1}{\kappa},$$

where $\Phi(x, y) = \max_{1 \leq i \leq m}\{|\langle F(x) - y, e_i \rangle|\}$. It follows from [79, Theorem 3.1] that for all $x \in \mathcal{B}_{\frac{\delta}{2}}(x^*) \cap \Omega$ and $y \in \mathcal{B}_{\frac{\delta}{2}}(F(x^*)) \cap \Lambda$,

$$\mathrm{dist}_{\mathcal{S}}(x, y) \leq \kappa \|F(x) - y\|, \tag{3.27}$$

where $\mathcal{S} := \{(x, y) \in \Omega \times \Lambda \mid F(x) - y = 0\}$. Let $\text{dist}_\Lambda(F(x)) = \|F(x) - y_x\|$ with $y_x \in \Lambda$. It follows from the continuity that there exists $\delta_0 \in (0, \frac{\delta}{2})$ such that if $x \in \mathcal{B}_{\delta_0}(x^*) \cap \Omega$, then $y_x \in \mathcal{B}_{\frac{\delta}{2}}(F(x^*))$. Thus, it follows from (3.27) that for each $x \in \mathcal{B}_{\delta_0}(x^*) \cap \Omega$,

$$\text{dist}_{\mathcal{S}}(x, y_x) \leq \kappa\|F(x) - y_x\| = \kappa\text{dist}_\Lambda(F(x)). \tag{3.28}$$

It is clear that for each x,

$$\text{dist}_{\mathscr{F}}(x) \leq \text{dist}_{\mathcal{S}}(x, y_x).$$

This and (3.28) imply that

$$\text{dist}_{\mathscr{F}} X(x) \leq \kappa\text{dist}_\Lambda(F(x)) \quad \forall x \in \mathcal{B}_{\delta_0}(x^*) \cap \Omega.$$

The proof is complete. □

As one of the main results, [57] has proven that the quasinormality implies the existence of a local error bound for smooth nonlinear programs in $\mathbb{R}^n$. Still in $\mathbb{R}^n$, Theorem 2.10 in chapter 2 extends [57, Theorem 2.1] to nonlinear programs with nonsmooth objective and constraints, and shows that the quasinormality implies local error bounds under some regularity conditions. Taking into account the previous results for problem (MPGC), we can now eliminate the extra regularity conditions and hence complete the investigation for local error bounds under the quasinormality for nonsmooth (NLP) problems in an infinite dimensional space. The improvement of our result owes much to the new approach constructing the enhanced sequential structure.

Corollary 3.14. *Let x^* be feasible for problem* (NLP$_{\text{Banach}}$). *Suppose that x^* is* *quasinormal. Then the local error bound holds, i.e., there are* $\delta > 0$ *and* $\kappa > 0$ *such that*

$$\mathrm{dist}_{\mathcal{F}_1}(x) \leq \kappa(\|h(x)\| + \|g^+(x)\|) \quad \forall x \in \mathcal{B}_\delta(x^*) \cap \Omega,$$

where $\mathcal{F}_1$ *is the feasible region of* (NLP$_{\text{Banach}}$).

We can also get the existence of local error bounds for the nonlinear semidefinite program (NLSDP) easily.

Corollary 3.15. *Let* x^* *be feasible for problem* (NLSDP)*. Suppose that* x^* *is quasinormal. Then the local error bound holds, i.e., there are* $\delta > 0$ *and* $\kappa > 0$ *such that*

$$\mathrm{dist}_{\mathcal{F}_2}(x) \leq \kappa\lambda_1(H(x))_+ \quad \forall x \in \mathcal{B}_\delta(x^*),$$

where $\mathcal{F}_2$ *is the feasible region of* (NLSDP)*.*

Taking Theorem 3.13 into account, we can now follow the Clarke's exact penalty principle [16, Proposition 2.4.3] and then get an exact penalty result for (MPGC) immediately.

Corollary 3.16. *Let* x^* *be a local minimizer of problem (MPGC). If the quasinormality holds at* x^**, then* x^* *is a local minimizer of the following penalized problem:*

$$\min_{x\in\Omega} \quad f(x) + \kappa L_f \mathrm{dist}_\Lambda(F(x))$$

where L_f *is the Lipschitzian constant of* f *near* x^* *and* κ *is the error bound constant.*

3.6 Sensitivity analysis

Mordukhovich and Nam [63], Mordukhovich et al [66], and Mordukhovich et al [64] studied the limiting subdifferential and singular subdifferential of value functions (or marginal functions) of a class of general optimization problems with abstract set-valued mapping constraints in Banach spaces and, Dempe et al [19] and [20] investigated the sensitivity of two-level value functions of pessimistic bilevel program and optimistic bilevel program in $\mathbb{R}^n$ respectively in terms of classical KKT multipliers by making use of the advanced tools of variational analysis [61]. In this section, we will study the sensitivity of value functions of (MPGC) and give much tighter upper estimate in terms of enhanced KKT multipliers. Consider the following parametric mathematical program with geometric constraints:

$$
\begin{aligned}
(\text{MPGC}_p) \qquad & \min_{x\in\Omega} \qquad && f(x,p) \\
& \quad \text{s.t.} && F(x,p) \in \Lambda,
\end{aligned}
$$

where $f : \mathbb{X} \times \mathbb{P} \to \mathbb{Y}$ and $F : \mathbb{X} \times \mathbb{P} \to \mathbb{Y}$ are locally Lipschitzian, and topological space $\mathbb{P}$ is assumed to be a Banach space in this section. Denote by $\mathscr{F}(p)$ the feasible region of problem (MPGC$_p$). We focus on the value function

$$
\mathcal{V}(p) := \inf\{f(x,p) \mid x \in \mathscr{F}(p)\}
$$

and the solution mapping

$$
\mathcal{O}(p) := \{x \in \mathscr{F}(p) \mid f(x,p) = \mathcal{V}(p)\}.
$$

To derive the sensitivity result in this section, we need to use the closedness of the approximal subdifferential and approximate normal cone. Since the nucleus of G-normal cone $\tilde{\mathcal{N}}^g_\Omega(x^*)$ as a set-valued map is not necessarily closed in Banach spaces, we consider the following slightly stronger quasinormality throughout this section by noting that the A-normal cone includes the nucleus of G-normal cone as a subset.

Definition 9. (x^*, p^*) is said to be *strongly quasinormal* for $\{(x,p) \in \Omega \times \mathbb{P} \,|\, F(x,p) \in \Lambda\}$ if there is no nonzero vector $\eta^* \in \mathbb{Y}$ such that

(1) $0 \in \sum_{i=1}^m \partial^a \langle \eta^*, e_i\rangle \langle F, e_i\rangle(x^*, p^*) + \mathcal{N}^a_\Omega(x^*) \times \{0\}, \quad \eta^* \in \mathcal{N}_\Lambda(F(x^*, p^*))$;

(2) There exists a sequence $\{x^k, p^k, y^k, \eta^k\}$ converging to $(x^*, p^*, F(x^*, p^*), \eta^*)$ such that for all k,

$$\eta^k \in \mathcal{N}_\Lambda(y^k), \quad \langle \eta^*, e_i\rangle \neq 0 \Longrightarrow \langle \eta^*, e_i\rangle \langle F(x^k, p^k) - y^k, e_i\rangle > 0.$$

The set of multipliers $\eta^* \in \mathcal{N}_\Lambda(F(x^*, p^*))$ satisfying (2) above is also denoted by $\mathcal{N}^e_\Lambda(F(x^*, p^*))$.

The following shows that the strong quasinormality is robust. Since the proof is similar to Lemma 1.1 in chapter 2 and [70, Lemma 2], we omit it here.

Proposition 3.6.1. *If the strong quasinormality holds at $(x^*, p^*) \in \{(x,p) \in \Omega \times \mathbb{P} \,|\, F(x,p) \in \Lambda\}$, then it holds at all feasible points near (x^*, p^*).*

It is well known that the MFCQ implies that the multiplier mapping is locally bounded (i.e., uniformly compact). The following shows that the strong quasinormality implies that the ϵ-quasinormality multiplier mapping

$$(\epsilon, x, p) \to \mathcal{M}^Q(\epsilon, x, p) :=$$
$$\{\eta \in \mathcal{N}^e_\Lambda(F(x,p)) \,|\, \epsilon \in \partial^a f(x,p) + \sum_{i=1}^m \partial^a \langle \eta, e_i\rangle \langle F, e_i\rangle(x,p) + \mathcal{N}^a_\Omega(x) \times \{0\}\}$$

is locally bounded. Since its proof is similar to Theorem 2.4 in chapter 2, we also omit it here.

Proposition 3.6.2. *If the strong quasinormality holds at*

$$(x^*, p^*) \in \{(x,p) \in \Omega \times \mathbb{P} \mid F(x,p) \in \Lambda\},$$

then the ϵ*-quasinormality multiplier mapping* $\mathcal{M}^Q$ *is locally bounded at* (ϵ^*, x^*, p^*)*, where* ϵ^* *is an arbitrary given element in* $\mathbb{X}^*$*.*

For sake of simplicity, given $\epsilon \geq 0$ and $r \geq 0$, we denote by $\mathcal{Q}^r_\epsilon(x^*, p^*)$ the set of vectors (η^*, ζ) satisfying

(i) $0 \in r\partial^a f(x^*, p^*) + \sum_{i=1}^m \partial^a \langle \eta^*, e_i\rangle\langle F, e_i\rangle(x^*, p^*) - (0, \zeta) + \mathcal{N}^a_\Omega(x^*) \times \epsilon\mathbb{B}_{\mathbb{P}^*}$ with $r \geq 0$ and $\eta^* \in \mathcal{N}_\Lambda(F(x^*, p^*))$,

(ii) If $\eta^* \neq 0$, then there exists a sequence $\{(x^k, p^k, y^k, \eta^k)\} \subseteq \Omega \times \mathbb{P} \times \Lambda \times \mathbb{Y}$ converging to $(x^*, p^*, F(x^*, p^*), \eta^*)$ such that for all k,

$$\begin{aligned}
&f(x^k) < f(x^*),\\
&\eta^k \in \mathcal{N}_\Lambda(y^k),\\
&\langle \eta^*, e_i\rangle \neq 0 \Longrightarrow \langle \eta^*, e_i\rangle\langle F(x^k, p^k) - y^k, e_i\rangle > 0.
\end{aligned}$$

Theorem 3.17. *Let* $x^* \in \mathcal{O}(p^*)$*. Assume that* (x^*, p^*) *is strongly quasinormal for the region* $\{(x,p) \in \Omega \times \mathbb{Y} \mid F(x,p) \in \Lambda\}$*. Then for any* $\epsilon > 0$*, we have*

$$\hat{\partial}_\epsilon \mathcal{V}(p^*) \subseteq \{\zeta \mid (\eta^*, \zeta) \in \mathcal{Q}^1_\epsilon(x^*, p^*)\}.$$

Proof. Let $\zeta \in \hat{\partial}_\epsilon \mathcal{V}(p^*)$. Then by the definition of ϵ-subdifferential, for each $\bar{\epsilon} > 0$, there exists $\delta_{\tilde{\epsilon}} > 0$ such that

$$\mathcal{V}(p) - \mathcal{V}(p^*) \geq \langle \zeta, p - p^* \rangle - (\epsilon + \tilde{\epsilon})\|p - p^*\| \quad \forall p \in \mathcal{B}_{\delta_{\tilde{\epsilon}}}(p^*).$$

By the definition of value functions, we have $f(x, p) \geq \mathcal{V}(p)$ for every $x \in \mathscr{F}(p)$ and hence

$$f(x, p) - \langle \zeta, p - p^* \rangle + (\epsilon + \tilde{\epsilon})\|p - p^*\| \geq f(x^*, p^*), \quad \forall x \in \mathscr{F}(p) \quad \forall p \in \mathcal{B}_{\delta_{\tilde{\epsilon}}}(p^*).$$

Thus, (x^*, p^*) is a locally optimal solution to the optimization problem

$$\begin{aligned} \min_{x \in \Omega, p \in \mathbb{P}} \quad & f(x, p) - \langle \zeta, p - p^* \rangle + (\epsilon + \tilde{\epsilon})\|p - p^*\| \\ \text{s.t.} \quad & F(x, p) \in \Lambda. \end{aligned}$$

Since (x^*, p^*) is strongly quasinormal for the above problem, it follows from Theorem 3.7 that there exist $\eta^* \in \mathcal{N}_\Lambda(F(x^*, p^*))$ and $\kappa \geq 0$ such that

(i) $0 \in \partial^a f(x^*, p^*) + \sum_{i=1}^m \partial^a \langle \eta^*, e_i \rangle \langle F, e_i \rangle (x^*, p^*) - (0, \zeta) + \mathcal{N}_\Omega^a(x^*) \times (\epsilon + \tilde{\epsilon})\mathbb{B}_{\mathbb{P}^*}$.

(ii) If $\eta^* \neq 0$, then there exists a sequence $\{(x^k, p^k, y^k, \eta^k)\} \subseteq \Omega \times \mathbb{P} \times \Lambda \times \mathbb{Y}$ converging to $(x^*, p^*, F(x^*, p^*), \eta^*)$ such that for all k,

$$\begin{aligned} & f(x^k) < f(x^*), \\ & \eta^k \in \mathcal{N}_\Lambda(y^k), \\ & \langle \eta^*, e_i \rangle \neq 0 \Longrightarrow \langle \eta^*, e_i \rangle \langle F(x^k, p^k) - y^k, e_i \rangle > 0. \end{aligned}$$

The desired result is obtained since $\tilde{\epsilon}$ is arbitrary. □

Definition 10. We say that the *inf-compactness* holds for (MPGC_p) with $p = p^*$ if there exist a number α and a compact set S such that for each p in some neighborhood of p^*, the level set

$$\{x \in \mathscr{F}(p) \mid f(x,p) \leq \alpha\}$$

is nonempty and contained in S.

Theorem 3.18. *Assume that the inf-compactness holds for problem* (MPGC)*. Suppose that for each $x^* \in \mathcal{O}(p^*)$, (x^*, p^*) is strongly quasinormal for the constraint region $\{(x,p) \in \Omega \times \mathbb{Y} \mid F(x,p) \in \Lambda)\}$. Then the value function $\mathcal{V}(p)$ is lower semicontinuous around p^* and*

$$\partial \mathcal{V}(p^*) \subseteq \bigcup_{x^* \in \mathcal{O}(p^*)} \{\zeta \mid (\eta^*, \zeta) \in \mathcal{Q}_0^1(x^*, p^*)\},$$
$$\partial^\infty \mathcal{V}(p^*) \subseteq \bigcup_{x^* \in \mathcal{O}(p^*)} \{\zeta \mid (\eta^*, \zeta) \in \mathcal{Q}_0^0(x^*, p^*)\}.$$

Proof. The lower semicontinuity follows from the proof of [10, Proposition 4.4] immediately. We complete the proof by considering the following two cases:

(a) Let $\zeta \in \partial \mathcal{V}(p^*)$. By the definition, there exist sequences $p^l \xrightarrow{\mathcal{V}} p^*$, $\epsilon_l \downarrow 0$, and $\zeta^l \xrightarrow{w^*} \zeta$ with $\zeta^l \in \hat{\partial}_{\epsilon_l} \mathcal{V}(p^l)$. Since the inf-compactness holds, the set $\{x \in \mathscr{F}(p^l) \mid f(x,p^l) \leq \alpha\}$ is nonempty when l is sufficiently large. By the inf-compactness again, there exists $x^l \in \mathcal{O}(p^l)$ and, without loss of generality, we may assume that $x^l \to x^*$. Since $\mathcal{V}(p^l) \to \mathcal{V}(p^*)$ and

$$\mathcal{V}(p^l) = f(x^l, p^l) \to f(x^*, p^*),$$

we have $f(x^*, p^*) = \mathcal{V}(p^*)$. Thus, $x^* \in \mathcal{O}(p^*)$. Since the strong quasinormality holds at (x^*, p^*) and $(p^l, x^l) \to (p^*, x^*)$, by Proposition 3.6.1, the strong quasinormality holds at (x^l, p^l) for each sufficiently large l. Thus, we have from

Theorem 3.17 that, for each sufficiently large l, there exist η^l and $\kappa \geq 0$ such that

(1) $(0, \zeta^l) \in \partial^a f(x^l, p^l) + \sum_{i=1}^m \partial^a \langle \eta^l, e_i\rangle\langle F, e_i\rangle(x^l, p^l) + \mathcal{N}_\Omega^a(x^l) \times \epsilon_l \mathbb{B}_{\mathbb{P}^*}$ with $\eta^l \in \mathcal{N}_\Lambda(F(x^l, p^l))$.

(2) If $\eta^l \neq 0$, then there exists a sequence $\{(x^{l,k}, p^{l,k}, y^{l,k}, \eta^{l,k})\} \subseteq \mathbb{X}\times\mathbb{P}\times\Lambda\times\mathbb{Y}$ converging to $(x^l, p^l, F(x^l, p^l), \eta^l)$ such that for all k,

$$f(x^{l,k}) < f(x^l),$$
$$\eta^{l,k} \in \mathcal{N}_\Lambda(y^{l,k}),$$
$$\langle \eta^l, e_i\rangle \neq 0 \Longrightarrow \langle \eta^l, e_i\rangle\langle F(x^{l,k}, p^{l,k}) - y^{l,k}, e_i\rangle > 0.$$

By the quasinormality assumption and Proposition 3.6.2, the sequence $\{\eta^l\}$ is bounded. Thus, without loss of generality, we may assume that $\{\eta^l\}$ converges to η^*. Taking a limit in (1) above, it is not hard to see from the weak* closedness of the approximate subdifferential and normal cone that

$$\begin{cases} (0, \zeta) \in \partial^a f(x^*, p^*) + \sum_{i=1}^m \partial^a \langle \eta^*, e_i\rangle\langle F, e_i\rangle(x^*, p^*) + \mathcal{N}_\Omega^a(x^*) \times \{0\}, \\ \eta^* \in \mathcal{N}_\Lambda(F(x^*, p^*)). \end{cases}$$

Also by the diagonal rule, we can find a sequence $\{(x^{l,k_l}, p^{l,k_l}, y^{l,k_l}, \eta^{l,k_l})\}$ converging to (x^*, p^*, y^*, η^*) as $l \to \infty$ such that for all l,

$$f(x^{l,k_l}) < f(x^*)$$
$$\eta^{l,k_l} \in \mathcal{N}_\Lambda(y^{l,k_l}),$$
$$\langle \eta^*, e_i\rangle \neq 0 \Longrightarrow \langle \eta^*, e_i\rangle\langle F(x^{l,k_l}, p^{l,k_l}) - y^{l,k_l}, e_i\rangle > 0.$$

Therefore, it follows that $(\eta^*, \zeta) \in \mathcal{Q}_0^1(x^*, p^*)$.

(b) Let $\zeta \in \partial^\infty \mathcal{V}(p^*)$. By the definition, there exist sequence $p^l \xrightarrow{\mathcal{V}} p^*$, $\epsilon_l \downarrow 0$, $\zeta^l \in \hat{\partial}_{\epsilon_l} \mathcal{V}(p^l)$, and $t_l \downarrow 0$ such that $t_l \zeta^l \to \zeta$. Similar as (a) in this proof, for each l sufficiently large l, there exist η^l and $\kappa \geq 0$ such that (1)–(2) in (a) of the proof hold. It is easy to get from (1) that

$$\begin{cases} (0, t_l \zeta^l) \in t_l \partial^a f(x^l, p^l) + \sum_{i=1}^m \partial^a \langle t_l \eta^l, e_i \rangle \langle F, e_i \rangle (x^l, p^l) + \mathcal{N}_\Omega^a(x^l) \times t_l \epsilon_l \mathbb{B}_{\mathbb{P}^*}, \\ t_l \eta^l \in \mathcal{N}_\Lambda(F(x^l, p^l)). \end{cases} \tag{3.29}$$

By the quasinormality assumption and Proposition 3.6.2, the sequence $\{t_l \eta^l\}$ is bounded. Without loss of generality, assume that $\{t_l \eta^l\}$ converges to η^*. Taking a limit as $k \to \infty$ in (4.16), we have from the weak* closedness of the approximate subdifferential and normal cone that

$$\begin{cases} (0, \zeta) \in \sum_{i=1}^m \partial^a \langle \eta^*, e_i \rangle \langle F, e_i \rangle (x^*, p^*) + \mathcal{N}_\Omega^a(x^*) \times \{0\}, \\ \eta^* \in \mathcal{N}_\Lambda(F(x^*, p^*)). \end{cases}$$

The rest of the proof is similar to (a).

The proof is complete. □

Chapter 4

Enhanced Karush-Kuhn-Tucker condition for mathematical programs with equilibrium constraints

4.1 Introduction

In this chapter, we study first-order necessary optimality conditions for the nonsmooth (MPEC). Since there are several different approaches to reformulate MPECs, various stationarity concepts such as Strong, Mordukhovich and Clarke (S, M and C)- stationarity arise (see [75, 82, 84, 85] for detailed discussions). The S-stationary condition, which is now well-known to be equivalent to the classical KKT conditions (see [23]), is the strongest among all stationary concepts for MPECs. For an MPEC with smooth problem data, it is shown that MPEC-LICQ is a constraint qualification

The content of this chapter is taken from Ye, J.J. and Zhang, J., "Enhanced Karush-Kuhn-Tucker condition for mathematical programs with equilibrium constraints." J. Optim. Theo. Appl., doi: 10.1007/s10957-013-0493-3, and part of Guo, L., Lin, G.H., Ye, J.J. and Zhang, J. (2014), "Sensitivity Analysis of the value function for parametric mathematical programs with equilibrium constraints." SIAM J. Optim., vol. 24, 1206-1237.

for S-stationarity (see [49, 81]). Moreover, MPEC-LICQ is a generic property [76], and hence it is not too stringent and can be satisfied for many smooth MPECs. It is tempting to assume that MPEC-LICQ is also a constraint qualification for MPECs, where the objective function is local Lipschitz but nonsmooth. In this chapter, we show through example that MPEC-LICQ is not a constraint qualification for MPECs, where the objective function is nonsmooth.

In this chapter, we not only extend Kanzow and Schwartz's results in [42] to the nonsmooth case, but also other than proving the enhanced M-stationary condition under the MPEC generalized quasi-normality and pseudo-normality for the nonsmooth MPEC, we derive some results which are new even for the smooth case. We show that if the equality functions and the complementarity functions are affine, the inequality function is concave and the abstract constraint set is polyhedral, then the MPEC generalized pseudo-normality holds at each feasible point. In [42], it was shown that the MPEC generalized pseudo-normality is a sufficient condition for the existence of a local error bound for a smooth MPEC. In this chapter, we improve this result by showing that the MPEC quasi-normality implies the existence of a local error bound under some reasonable conditions.

Recently, constraint qualifications such as quasi-normality (see Chapter 2), Constant Positive Linear Dependance (CPLD) (see [71]) and Relaxed Constant Positive Linear Dependence (RCPLD) (see [1]) have all been shown to provide weaker constraint qualifications than MFCQ. In this chapter we introduce a weaker version of the MPEC-CPLD and show that it is a stronger condition than the MPEC generalized quasi-normality. Consequently this weaker version of the MPEC-CPLD is also a constraint qualification for the enhanced M-stationary condition and a sufficient condition for the existence of a local error bound.

4.2 Enhanced stationary conditions for MPEC

In this chapter we study the MPEC problem (1.8), where $f, h_i(i = 1, \dots, p), g_j(j = 1, \dots, q) : \mathbb{R}^n \to \mathbb{R}$ are Lipschitz continuous around the point of interest, $G_l, H_l(l = 1, \dots, m) : \mathbb{R}^n \to \mathbb{R}$ are continuously differentiable, and $\mathcal{X}$ is a closed subset of $\mathbb{R}^n$. Let $\mathfrak{F}$ denote the feasible region. Let x^* be a feasible point of problem (MPEC). We define the following index sets:

$$\begin{aligned}
A(x^*) &:= \{j | g_j(x^*) = 0\}, \\
I_{00} := I_{00}(x^*) &:= \{l | G_l(x^*) = 0, H_l(x^*) = 0\}, \\
I_{0+} := I_{0+}(x^*) &:= \{l | G_l(x^*) = 0, H_l(x^*) > 0\}, \\
I_{+0} := I_{+0}(x^*) &:= \{l | G_l(x^*) > 0, H_l(x^*) = 0\}.
\end{aligned}$$

Recall that the MPEC-LICQ holds at a feasible point x^* if the gradient vectors

$$\{\nabla h_i(x^*) | i = 1, \dots, p\}, \{\nabla g_j(x^*) | j \in A(x^*)\},$$
$$\{\nabla G_l(x^*) | l \in I_{00} \cup I_{0+}\}, \{\nabla H_l(x^*) | l \in I_{00} \cup I_{+0}\}$$

are linearly independent (see [76]). The following example shows that MPEC-LICQ may not be a constraint qualification for S-stationary condition if the objective function is not differentiable.

Example 2. *Consider the MPEC:* $\min -y + |x - y|$ *subject to* $x \geq 0, y \geq 0, xy = 0$. *It is easy to see that* $(0, 0)$ *is a minimizer and MPEC-LICQ holds at every point of the feasible region. The S-stationary condition is the existence of* $\mu \geq 0, \nu \geq 0$ *such* *that*

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \beta \\ -1-\beta \end{pmatrix} + \mu \begin{pmatrix} -1 \\ 0 \end{pmatrix} + \nu \begin{pmatrix} 0 \\ -1 \end{pmatrix} \tag{4.1}$$

with $\beta \in [-1, 1]$ *being an element in the subdifferential of the convex function* $|\cdot|$ *at the origin. However,* (4.1) *never holds and hence* $(0, 0)$ *is not an S-stationary point.*

Remark 1. *We may construct a class of MPECs with nonsmooth objectives that have local minimizers satisfying MPEC-LICQ but not S-stationarity. Indeed, consider an MPEC with affine complementarity constraints:* $\min f(x)$ *s.t.* $0 \leq G(x) \perp H(x) \geq 0$*, where* $a \perp b$ *means that the vectors* a *and* b *are perpendicular. Since* $G(x)$ *and* $H(x)$ *are affine, a local optimal solution to the above MPEC is also a local optimal solution to the penalized problem*

$$\min \Big[f(x) + M(\|G(x) - y\| + \|H(x) - z\|) \Big] \textit{ s.t. } \quad 0 \leq y \perp z \geq 0$$

for some $M > 0$*, where* $\|\cdot\|$ *denotes the Euclidean norm. For the penalized problem above, MPEC-LICQ holds at each feasible point but the objective function is non-smooth. However, a local optimal solution is not always an S-stationary point for the penalized problem since otherwise it would also be an S-stationary point for the original problem as well, which may not be true.*

Nevertheless, the S-stationary condition is important for MPECs. It provides sharper optimality condition than the M- or the C-stationary condition. If MPEC LICQ is not a constraint qualification, are there any suitable constraint qualifications under which a local optimal solution of a MPEC with nonsmooth objective functions is an S-stationary point? Recall that in the case of MPEC, each feasible point violates MFCQ. However, what about other known weaker constraint qualifications? Izmailov et al. [37] conclude that RCPLD, and hence CPLD cannot be expected to hold with any frequency in cases of interest. As a matter of fact, Izmailov et al. [37] state that if RCPLD can hold, there remains no "degrees of freedom" for optimization, which in the context of MPEC is not the relevant case. Although it is possible to construct examples where RCPLD and even CPLD hold (as shown in [37] that CPLD holds for MPEC with all constraints being identically zero, which is of course extremely pathological), for practical MPECs with "degree of freedom" for optimization, (R)CPLD cannot be expected to hold.

It is natural to ask how likely the quasinormality can hold for an MPEC. To show it is possible for the quasinormality to hold we consider the following MPEC as a standard nonlinear programming problem with inequality constraints $G(x) \geq 0, H(x) \geq 0$ and equality constraint $G(x)H(x) = 0$.

Example 3. *Consider the MPEC:*

$$\begin{aligned} \min \quad & f(x) := |x| \\ s.t. \quad & G(x) \geq 0, H(x) \geq 0, G(x)H(x) = 0, \end{aligned}$$

where $H(x) = x$ and $G(x)$ is defined as follows:

$$G(x) := \begin{cases} -(x+1)^2 & x < -1, \\ 0 & -1 \leq x \leq 1, \\ -(x-1)^2 & x > 1. \end{cases}$$

It is easy to see the set of feasible solution is all points lying in the interval $[0, 1]$ and MPEC LICQ fails at $x^* = 0$. Suppose there are scalars $\lambda_G \geq 0, \lambda_H \geq 0, \lambda_{GH}$ such that

$$-\lambda_G G'(x^*) - \lambda_H H'(x^*) + \lambda_{GH} \left[G(x^*)H'(x^*) + H(x^*)G'(x^*) \right] = 0.$$

Then we can take $\lambda_H = 0$. *Since* $G'(x^*) = 0$ *and* $G(x^*)H'(x^*) + H(x^*)G'(x^*) = 0$, λ_G *and* λ_{GH} *can be chosen to be nonzero. However, since* $G(x) = 0$ *for all* x *in the neighborhood of* $x^* = 0$, *for any sequence* $\{x^k\} \to x^*$, *for* k *large enough, there always holds*

$$\lambda_{GH}G(x^k)H(x^k) = 0, \; for \; \lambda_{GH} \neq 0; \quad \lambda_G G(x^k) = 0, \; for \; \lambda_G > 0,$$

which means that x^* *is quasinormal.*

However, this example is also pathological. We next argue that for practical MPECs formulated as a standard nonlinear programming, quasinormality cannot be expected to hold. Consider MPEC formulated as a standard nonlinear programming and observe that the limiting subdifferentials of the constraints at any x has the form

$$\begin{pmatrix} \partial h(x) \\ \partial g(x) \\ -\nabla G(x) \\ -\nabla H(x) \\ (\nabla H(x))^T G(x) + (\nabla G(x))^T H(x) \end{pmatrix}. \tag{4.2}$$

Of interest is to consider a bi-active (we say that an index l is bi-active at x^* if $H_l(x^*) = G_l(x^*) = 0$) feasible point x^* of MPEC. Then the last row of the subdifferential (4.2) at x^* equals zero identically. Hence there always exist abnormal multipliers $(\lambda, \mu, \gamma, \nu, \tau) = (0, 0, 0, 0, e) \in \mathbb{R}^p \times \mathbb{R}^q_+ \times \mathbb{R}^m_+ \times \mathbb{R}^m_+ \times \mathbb{R}^m$ not all zero such that

$$\begin{aligned} 0 \in \quad & \sum_{i=1}^{p} \partial(\lambda_i h_i)(x^*) + \sum_{j=1}^{q} \mu_j \partial g_j(x^*) \\ & - \sum_{l=1}^{m} [(\gamma_l - \tau_l H_l(x^*) \nabla G_l(x^*) + (\nu_l - \tau_l G_l(x^*)) \nabla H_l(x^*)] + \mathcal{N}_{\mathcal{X}}(x^*), \end{aligned}$$

where $e = (1, \ldots, 1)$ is the vector of all 1s. Therefore, for quasinormality to hold at

x^*, $G_l(x^k)H_l(x^k)$ cannot be strictly positive for all $l \in \{1,\dots,m\}$ and any arbitrary sequence $\{x^k\}$ converging to x^*, which is extremely atypical.

The conclusion is that neither (R)CPLD nor quasinormality can be expected to hold for MPEC in cases of interest. In other words, the enhanced Fritz John condition in chapter 2 is not in general applicable for MPEC if treated as a NLP. Recently, Kanzow and Schwartz [42] have introduced the MPEC-tailored enhanced Fritz John condition for smooth MPEC with no abstract set constraint. In this chapter, we extend Kanzow and Schwartz's result into the more general nonsmooth case and hence leading up to several MPEC-tailored weaker CQs. However, the reason why we did not simply apply the result from chapter 2 but we exploit the particular structure of the complementarity constraints within our MPEC is to obtain suitable sign constraints on the multipliers (as it was done in [42]). As a matter of fact, direct application of chapter 2 would have led to conclusions with artificial slack variable y and z involved, which are less favorable than ours as they depend only on x. Note that even in the setting of standard MPEC, where all functions are smooth and there is no abstract constraint, our new enhanced Fritz John M-stationary condition is still stronger and provides more information than the original Fritz John M-stationary condition derived by Kanzow and Schwartz in [42] for the following reason: a continuously differentiable function may not be proximal subdifferentiable (a sufficient condition for a function to be proximal subdifferentiable is that the function is C^{1+}, i.e., the function is continuously differentiable and its gradient is Lipschitz continuous).

Theorem 4.1. *Let x^* be a local minimizer of problem* (MPEC)*. Then, there are multipliers $\alpha, \lambda, \mu, \gamma, \nu$ such that*

(i) $0 \in \alpha\partial f(x^*) + \sum_{i=1}^{p} \partial(\lambda_i h_i)(x^*) + \sum_{j=1}^{q} \mu_j \partial g_j(x^*) - \sum_{l=1}^{m} [\gamma_l \nabla G_l(x^*) + \nu_l \nabla H_l(x^*)] + \mathcal{N}_{\mathcal{X}}(x^*)$;

(ii) $\alpha \geq 0, \mu \geq 0$, $\gamma_l = 0$, $\forall l \in I_{+0}(x^*)$, $\nu_l = 0$, $\forall l \in I_{0+}(x^*)$, *and either* $\gamma_l > 0$,

$\nu_l > 0$ *or* $\gamma_l \nu_l = 0 \ \forall l \in I_{00}(x^*)$;

(iii) $\alpha, \lambda, \mu, \gamma, \nu$ *are not all equal to zero;*

(iv) *If* $\lambda, \mu, \gamma, \nu$ *are not all equal to zero, then there exists a sequence* $\{x^k\} \subseteq \mathcal{X}$ *converging to* x^* *such that for all* k,

$$\begin{aligned} & f(x^k) < f(x^*), \\ & \text{if } \lambda_i \neq 0, \ \text{then } \lambda_i h_i(x^k) > 0, \quad \text{if } \mu_j > 0, \ \text{then } \mu_j g_j(x^k) > 0, \\ & \text{if } \gamma_l \neq 0, \ \text{then } \gamma_l G_l(x^k) < 0, \quad \text{if } \nu_l \neq 0, \ \text{then } \nu_l H(x^k) < 0. \end{aligned}$$

Proof. The results can be proved by combining the techniques and the results in [42, Theorem 3.1] and Chapter 2. □

Based on the result above, we define the following enhanced M-stationary conditions.

Definition 11 (Enhanced M-stationary conditions)**.** *Let* x^* *be a feasible point of problem* (MPEC). *We say the enhanced M-stationary condition holds at* x^* *iff there are multipliers* $\lambda, \mu, \gamma, \nu$ *such that*

(i) $0 \in \partial f(x^*) + \sum_{i=1}^{p} \partial(\lambda_i h_i)(x^*) + \sum_{j=1}^{q} \mu_j \partial g_j(x^*) - \sum_{l=1}^{m} [\gamma_l \nabla G_l(x^*) + \nu_l \nabla H_l(x^*)] + \mathcal{N}_{\mathcal{X}}(x^*)$;

(ii) $\mu \geq 0$, $\gamma_l = 0$, $\forall l \in I_{+0}(x^*)$, $\nu_l = 0$, $\forall l \in I_{0+}(x^*)$, *and either* $\gamma_l > 0$, $\nu_l > 0$ *or* $\gamma_l \nu_l = 0 \ \forall l \in I_{00}(x^*)$;

(iv) *If* $\lambda, \mu, \gamma, \nu$ *are not all equal to zero, then there exists a sequence* $\{x^k\} \subseteq \mathcal{X}$

converging to x^* such that for all k,

$$\begin{aligned}
&\text{if } \lambda_i \neq 0, \text{ then } \lambda_i h_i(x^k) > 0,\\
&\text{if } \mu_j > 0, \text{ then } \mu_j g_j(x^k) > 0,\\
&\text{if } \gamma_l \neq 0, \text{ then } \gamma_l G_l(x^k) < 0,\\
&\text{if } \nu_l \neq 0, \text{ then } \nu_l H(x^k) < 0.
\end{aligned}$$

We call the multipliers $\lambda, \mu, \gamma, \nu$ the MPEC quasi-normal multipliers corresponding to x^*.

Motivated by Theorem 4.1 and the related discussion in Chapter 2, we now introduce some MPEC-variant CQs. Note that although Definition 12(d) is weaker than the MPEC-CPLD introduced in [32,43], where all functions involved are continuously differentiable and $\mathcal{X} = \mathbb{R}^n$, for convenience we still refer to it as MPEC-CPLD. The MPEC-RCPLD was first introduced in [28] and has been proven to be a sufficient condition for M-stationarity in [26].

Definition 12. *Let x^* be a feasible solution of problem* (MPEC)*.*

(a) *x^* is said to satisfy MPEC-NNAMCQ iff there is no nonzero vector $(\lambda, \mu, \gamma, \nu)$ such that*

  (i) $0 \in \sum_{i=1}^{p} \partial(\lambda_i h_i)(x^*) + \sum_{j=1}^{q} \mu_j \partial g_j(x^*) - \sum_{l=1}^{m} [\gamma_l \nabla G_l(x^*) + \nu_l \nabla H_l(x^*)] + \mathcal{N}_{\mathcal{X}}(x^*)$*;*

  (ii) $\mu \geq 0$, $\gamma_l = 0$, $\forall l \in I_{+0}(x^*)$*;* $\nu_l = 0$, $\forall l \in I_{0+}(x^*)$*, and either* $\gamma_l > 0$*,* $\nu_l > 0$ *or* $\gamma_l \nu_l = 0$ $\forall l \in I_{00}(x^*)$*.*

(b) *x^* is said to satisfy MPEC generalized pseudo-normality iff there is no nonzero vector $(\lambda, \mu, \gamma, \nu)$ such that*

(i) $0 \in \sum_{i=1}^{p} \partial(\lambda_i h_i)(x^*) + \sum_{j=1}^{q} \mu_j \partial g_j(x^*) - \sum_{l=1}^{m} [\gamma_l \nabla G_l(x^*) + \nu_l \nabla H_l(x^*)] + \mathcal{N}_{\mathcal{X}}(x^*)$;

(ii) $\mu \geq 0$, $\gamma_l = 0$, $\forall l \in I_{+0}(x^*)$; $\nu_l = 0$, $\forall l \in I_{0+}(x^*)$, *and either* $\gamma_l > 0$, $\nu_l > 0$ *or* $\gamma_l \nu_l = 0$ $\forall l \in I_{00}(x^*)$;

(iii) *There exists a sequence* $\{x^k\} \subseteq \mathcal{X}$ *converging to* x^* *such that for all* k,

$$\sum_{i=1}^{p} \lambda_i h_i(x^k) + \sum_{j=1}^{q} \mu_j g_j(x^k) - \sum_{l=1}^{m} [\gamma_l G_l(x^k) + \nu_l H_l(x^k)] > 0.$$

(c) x^* *is said to satisfy MPEC generalized quasi-normality iff there is no nonzero vector* $(\lambda, \mu, \gamma, \nu)$ *such that*

(i) $0 \in \sum_{i=1}^{p} \partial(\lambda_i h_i)(x^*) + \sum_{j=1}^{q} \mu_j \partial g_j(x^*) - \sum_{l=1}^{m} [\gamma_l \nabla G_l(x^*) + \nu_l \nabla H_l(x^*)] + \mathcal{N}_{\mathcal{X}}(x^*)$;

(ii) $\mu \geq 0$, $\gamma_l = 0$, $\forall l \in I_{+0}(x^*)$; $\nu_l = 0$, $\forall l \in I_{0+}(x^*)$, *and either* $\gamma_l > 0$, $\nu_l > 0$ *or* $\gamma_l \nu_l = 0$ $\forall l \in I_{00}(x^*)$.

(iii) *If* $\lambda, \mu, \gamma, \nu$ *are not all equal to zero, then there exists a sequence* $\{x^k\} \subseteq \mathcal{X}$ *converging to* x^* *such that, for all* k,

$$\begin{aligned} &\text{if } \lambda_i \neq 0, \text{ then } \lambda_i h_i(x^k) > 0, \\ &\text{if } \mu_j > 0, \text{ then } g_j(x^k) > 0, \\ &\text{if } \gamma_l \neq 0, \text{ then } \gamma_l G_l(x^k) < 0, \\ &\text{if } \nu_l \neq 0, \text{ then } \nu_l H(x^k) < 0. \end{aligned}$$

(d) *In addition to the basic assumptions for the problem* (MPEC), *suppose that* h, g *are continuously differentiable at* x^* *and* $\mathcal{X} = \mathbb{R}^n$. x^* *is said to satisfy MPEC-CPLD iff for any indices set* $I_0 \subseteq \mathfrak{P} := \{1, 2, \ldots, p\}$, $J_0 \subseteq A(x^*)$,

$L_0^G \subseteq I_{0+}(x^*) \cup I_{00}(x^*)$ *and* $L_0^H \subseteq I_{+0}(x^*) \cup I_{00}(x^*)$, *whenever there exist* λ_i, $\mu_j \geq 0$ *for all* $j \in J_0$, γ_l *and* ν_l *not all zero, such that*

$$0 = \sum_{i \in I_0} \lambda_i \nabla h_i(x^*) + \sum_{j \in J_0} \mu_j \nabla g_j(x^*) - \sum_{l \in L_0^G} \gamma_l \nabla G_l(x^*) - \sum_{l \in L_0^H} \nu_l \nabla H_l(x^*)$$

and either $\gamma_l \nu_l = 0$ *or* $\gamma_l > 0$, $\nu_l > 0$ $\forall l \in I_{00}(x^*)$, *there is a neighborhood* $U(x^*)$ *of* x^* *such that, for any* $x \in U(x^*)$,

$$(\{\nabla h_i(x) | i \in I_0\}, \{\nabla g_j(x) | j \in J_0\}, \{\nabla G_l(x) | l \in L_0^G\}, \{\nabla H_l(x) | l \in L_0^H\})$$

are linearly dependent.

(e) *In addition to the basic assumptions for the problem* (MPEC), *suppose that* h, g *are continuously differentiable at* x^* *and* $\mathcal{X} = \mathbb{R}^n$. *Let* $I_0 \subseteq \mathfrak{P}$ *be such that* $\{\nabla h_i(x^*)\}_{i \in I_0}$ *is a basis for span*$\{\nabla h_i(x^*)\}_{i \in \mathfrak{P}}$. x^* *is said to satisfy MPEC-RCPLD iff there is a neighborhood* $U(x^*)$ *of* x^* *such that*

(i) $\{\nabla h_i(x)\}_{i \in \mathfrak{P}}$ *has the same rank for every* $x \in U(x^*)$.

(ii) *For every* $J_0 \subseteq A(x^*)$, $L_0^G \subseteq I_{0+}(x^*) \cup I_{00}(x^*)$ *and* $L_0^H \subseteq I_{+0}(x^*) \cup I_{00}(x^*)$, *whenever there exist* λ_i, $\mu_j \geq 0$ *for all* $j \in J_0$, γ_l *and* ν_l *not all zero such that*

$$0 = \sum_{i \in I_0} \lambda_i \nabla h_i(x^*) + \sum_{j \in J_0} \mu_j \nabla g_j(x^*) - \sum_{l \in L_0^G} \gamma_l \nabla G_l(x^*) - \sum_{l \in L_0^H} \nu_l \nabla H_l(x^*),$$

and either $\gamma_l \nu_l = 0$ *or* $\gamma_l > 0$, $\nu_l > 0$ $\forall l \in I_{00}(x^*)$; *then the vectors*

$$(\{\nabla h_i(x) | i \in I_0\}, \{\nabla g_j(x) | j \in J_0\}, \{\nabla G_l(x) | l \in L_0^G\}, \{\nabla H_l(x) | l \in L_0^H\})$$

are linearly dependent for any $x \in U(x^*)$.

It is easy to see that MPEC-NNAMCQ $\Rightarrow$ MPEC generalized pseudo-normality $\Rightarrow$ MPEC generalized quasi-normality.

For the standard nonsmooth nonlinear program where the equality functions are linear, inequality functions are concave and there is no abstract constraint, Chapter 2 showed that the pseudo-normality holds automatically at any feasible point. In what follows, we extend this result to MPEC.

Theorem 4.2. *Suppose that h_i are linear, g_j are concave, G_l, H_l are all linear and $\mathcal{X}$ is polyhedral. Then any feasible point of problem* (MPEC) *is MPEC generalized pseudo-normal.*

Proof. We omit the abstract set $\mathcal{X}$ since it can be represented by several linear inequalities. We prove the theorem by contradiction. To the contrary, suppose that there is a feasible point x^* that is not MPEC generalized pseudo-normal. Then there exists nonzero vector $(\lambda, \mu, \gamma, \nu) \in \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R}^m \times \mathbb{R}^m$ and infeasible sequence $\{x^k\} \subseteq \mathcal{X}$ converging to x^* such that

$$0 \in \sum_{i=1}^{p} \lambda_i \nabla h_i(x^*) + \sum_{j=1}^{q} \mu_j \partial g_j(x^*) - \sum_{l=1}^{m} [\gamma_l \nabla G_l(x^*) + \nu_l \nabla H_l(x^*)], \tag{4.3}$$

where $\mu \geq 0, \mu_j = 0 \ \forall j \notin A(x^*)$, $\gamma_l = 0 \ \forall l \in I_{+0}(x^*)$, $\nu_l = 0 \ \forall l \in I_{0+}(x^*)$ and either $\gamma_l \nu_l = 0$ or $\gamma_l > 0$, $\nu_l > 0 \ \forall l \in I_{00}(x^*)$. Furthermore, for each k,

$$\sum_{i=1}^{p} \lambda_i h_i(x^k) + \sum_{j=1}^{q} \mu_j g_j(x^k) - \sum_{l=1}^{m} [\gamma_l G_l(x^k) + \nu_l H_l(x^k)] > 0. \tag{4.4}$$

By the linearity of h_i, G_l, H_l and concavity of g_j, we have that, for all $x \in \mathbb{R}^n$,

$$\begin{aligned} h_i(x) &= h_i(x^*) + \nabla h_i(x^*)^T(x - x^*) \quad i = 1, \ldots, p, \\ G_l(x) &= G_l(x^*) + \nabla G_l(x^*)^T(x - x^*) \quad l = 1, \ldots, m, \\ H_l(x) &= H_l(x^*) + \nabla H_l(x^*)^T(x - x^*) \quad l = 1, \ldots, m, \\ g_j(x) &\leq g_j(x^*) + \xi_j^T(x - x^*) \quad \forall \xi_j \in \partial g_j(x^*), j = 1, \ldots, q. \end{aligned}$$

By multiplying these four relations with λ_i, γ_l, ν_l and μ_j and by adding over i, l and j respectively, we obtain that, for all $x \in \mathbb{R}^m$ and all $\xi_j \in \partial g_j(x^*), j = 1, \ldots, q$,

$$\begin{aligned} &\sum_{i=1}^{p} \lambda_i h_i(x) + \sum_{j=1}^{q} \mu_j g_j(x) - \sum_{l=1}^{m} (\gamma_l G_l(x) + \nu_l H_l(x)) \\ &\leq \sum_{i=1}^{p} \lambda_i h_i(x^*) + \sum_{j=1}^{q} \mu_j g_j(x^*) - \sum_{l=1}^{m} (\gamma_l G_l(x^*) + \nu_l H_l(x^*)) \\ &\quad + \Big[\sum_{i=1}^{p} \lambda_i \nabla h_i(x^*) + \sum_{j=1}^{q} \mu_j \xi_j - \sum_{l=1}^{m} (\gamma_l \nabla G_l(x^*) + \nu_l \nabla H_l(x^*)) \Big]^T (x - x^*) \\ &= \Big[\sum_{i=1}^{p} \lambda_i \nabla h_i(x^*) + \sum_{j=1}^{q} \mu_j \xi_j - \sum_{l=1}^{m} (\gamma_l \nabla G_l(x^*) + \nu_l \nabla H_l(x^*)) \Big]^T (x - x^*), \end{aligned}$$

where the last equality holds because we have

$$\lambda_i h_i(x^*) = 0 \text{ for all } i \text{ and } \sum_{j=1}^{q} \mu_j g_j(x^*) = 0, \sum_{l=1}^{m} \gamma_l G_l(x^*) = 0, \sum_{l=1}^{m} \nu_l H_l(x^*) = 0.$$

By (4.3), there exists $\xi_j^* \in \partial g_j(x^*), j = 1, \ldots, q$ such that

$$\sum_{i=1}^{p} \lambda_i \nabla h_i(x^*) + \sum_{j=1}^{q} \mu_j \xi_j^* - \sum_{l=1}^{m} [\gamma_l \nabla G_l(x^*) + \nu_l \nabla H_l(x^*)] = 0.$$

Hence it follows that for all $x \in \mathbb{R}^n$, $\sum_{i=1}^{p} \lambda_i h_i(x) + \sum_{j=1}^{q} \mu_j g_j(x) - \sum_{l=1}^{m} [\gamma_l G_l(x) + \nu_l H_l(x)] \leq 0$, which contradicts (4.4). The proof is complete. $\square$

The CPLD was introduced by Qi and Wei in [71] and was used to analyze SQP algorithms. [3] showed that for smooth nonlinear programs, the CPLD condition implies the quasi-normality and hence is a constraint qualification as well. In what follows, we show that the MPEC-CPLD introduced in this chapter implies MPEC generalized quasi-normality. We first recall the following lemma, a proof of which may be found in [1, Lemma 1].

Lemma 4.3. *If $x = \sum_{i=1}^{m+p} \alpha_i \nu_i$ with $\nu_i \in \mathbb{R}^n$ for every i, $\{\nu_i\}_{i=1}^m$ is linearly independent and $\alpha_i \neq 0$ for every $i = m+1, \ldots, m+p$, then there exist $J \subseteq \{m+1, \ldots, m+p\}$ and scalars $\bar{\alpha}_i$ for every $i \in \{1, \ldots, m\} \cup J$ such that*

- $x = \sum_{i \in i, \ldots, m \cup J} \bar{\alpha}_i \nu_i$,
- $\alpha_i \bar{\alpha}_i > 0$ *for every* $i \in J$,
- $\{\nu_i\}_{i \in \{i, \ldots, m\} \cup J}$ *is linearly independent.*

Theorem 4.4. *Let x be a feasible solution of problem* (MPEC) *where h, g are continuously differentiable such that MPEC-CPLD holds. Then x is MPEC generalized quasi-normal.*

Proof. For brevity, we drop the equality and the inequality constraint in the proof, since the main difficulties are induced by the complementarity constraints. Assume that x is feasible and the MPEC-CPLD condition holds at x. If x satisfies MPEC-NNAMCQ, we are done. Suppose MPEC-NNAMCQ does not hold. Then, there exists a nonzero vector $(\gamma, \nu) \in \mathbb{R}^m \times \mathbb{R}^m$ such that $0 = -\sum_{l=1}^m [\gamma_l \nabla G_l(x) + \nu_l \nabla H_l(x)]$, $\gamma_l = 0\ \forall l \in I_{+0}(x)$, $\nu_l = 0\ \forall l \in I_{0+}(x)$ and either $\gamma_l \nu_l = 0$ or $\gamma_l > 0$, $\nu_l > 0\ \forall l \in I_{00}(x)$.

Define the index sets:

$$L_+^G(x) := \{l \in I_{0+}(x) | \gamma_l > 0\}, \ L_-^G(x) := \{l \in I_{0+}(x) | \gamma_l < 0\},$$

$$L_+^H(x) := \{l \in I_{+0}(x) | \nu_l > 0\}, L_-^H(x) := \{l \in I_{+0}(x) | \nu_l < 0\},$$

$$I_{00}^{++}(x) := \{l \in I_{00}(x) | \gamma_l > 0, \nu_l > 0\}, I_{00}^{+0}(x) := \{l \in I_{00}(x) | \gamma_l > 0, \nu_l = 0\},$$

$$I_{00}^{-0}(x) := \{l \in I_{00}(x) | \gamma_l < 0, \nu_l = 0\}, I_{00}^{0+}(x) := \{l \in I_{00}(x) | \gamma_l = 0, \nu_l > 0\},$$

$$I_{00}^{0-}(x) := \{l \in I_{00}(x) | \gamma_l = 0, \nu_l < 0\}.$$

Since (γ, ν) is a nonzero vector, the union of the above sets must be nonempty and we may write

$$\begin{aligned} 0 \quad = \quad & -[\sum_{l \in L_+^G(x)} \gamma_l \nabla G_l(x) + \sum_{l \in L_-^G(x)} \gamma_l \nabla G_l(x)] - [\sum_{l \in L_+^H(x)} \nu_l \nabla H_l(x) + \sum_{l \in L_-^H(x)} \nu_l \nabla H_l(x)] \\ & - \sum_{l \in I_{00}^{++}(x)} [\gamma_l \nabla G_l(x) + \nu_l \nabla H_l(x)] - [\sum_{l \in I_{00}^{+0}(x)} \gamma_l \nabla G_l(x) + \sum_{l \in I_{00}^{-0}(x)} \gamma_l \nabla G_l(x)] \\ & -[\sum_{l \in I_{00}^{0+}(x)} \nu_l \nabla H_l(x) + \sum_{l \in I_{00}^{0-}(x)} \nu_l \nabla H_l(x)]. \end{aligned}$$

Assume first that $L_+^G(x)$ is nonempty. Let $l_1 \in L_+^G(x)$. Then,

$$\begin{aligned} -\gamma_{l_1} \nabla G_{l_1}(x) \quad = \quad & \sum_{l \in L_+^G(x) \setminus \{l_1\}} \gamma_l \nabla G_l(x) + \sum_{l \in L_-^G(x)} \gamma_l \nabla G_l(x) \\ & +[\sum_{l \in L_+^H(x)} \nu_l \nabla H_l(x) + \sum_{l \in L_-^H(x)} \nu_l \nabla H_l(x)] \\ & + \sum_{l \in I_{00}^{++}(x)} [\gamma_l \nabla G_l(x) + \nu_l \nabla H_l(x)] \\ & +[\sum_{l \in I_{00}^{+0}(x)} \gamma_l \nabla G_l(x) + \sum_{l \in I_{00}^{-0}(x)} \gamma_l \nabla G_l(x)] \\ & +[\sum_{l \in I_{00}^{0+}(x)} \nu_l \nabla H_l(x) + \sum_{l \in I_{00}^{0-}(x)} \nu_l \nabla H_l(x)]. \end{aligned}$$

If $\nabla G_{i_1}(x) = 0$, the single-element set $\{\nabla G_{i_1}(x)\}$ is linearly dependent. By MPEC-CPLD, the set $\{\nabla G_{i_1}(y)\}$ must be linearly dependent for all y in some neighborhood of x. Therefore, $\nabla G_{i_1}(y) = 0$ for all y in an open neighborhood of x. Since $G_{i_1}(x) = 0$, this implies that $G_{i_1}(y) = 0$ for all y in that neighborhood. Hence for any sequence $x^k \to x$, $G_{i_1}(x^k) = 0$ always holds. That is, there is no sequence $x^k \to x$ such that $\lambda_{i_1} G_{i_1}(x^k) > 0$.

Assume now that $\nabla G_{i_1}(x) \neq 0$. Then, by Lemma 4.3, there exist index sets

$$
\begin{aligned}
&\bar{L}_+^G(x) \subseteq L_+^G(x) \setminus \{i_1\},\ \bar{L}_-^G(x) \subseteq L_-^G(x),\ \bar{L}_+^H(x) \subseteq L_+^H(x),\ \bar{L}_-^H(x) \subseteq L_-^H(x) \\
&\bar{I}_{00}^{++}(x) \subseteq I_{00}^{++}(x),\ \bar{I}_{00}^{+0}(x) \subseteq I_{00}^{+0}(x),\ \bar{I}_{00}^{-0}(x) \subseteq I_{00}^{-0}(x), \\
&\bar{I}_{00}^{0+}(x) \subseteq I_{00}^{0+}(x),\ \bar{I}_{00}^{0-}(x) \subseteq I_{00}^{0-}(x)
\end{aligned}
$$

such that the vectors

$$
\begin{aligned}
&\{\nabla G_l(x)\}_{l\in\bar{L}_+^G(x)}, \{\nabla G_l(x)\}_{l\in\bar{L}_-^G(x)}, \{\nabla H_l(x)\}_{l\in\bar{L}_+^H(x)}, \{\nabla H_l(x)\}_{l\in\bar{L}_-^H(x)}, \\
&\{\nabla G_l(x)\}_{l\in\bar{I}_{00}^{++}(x)}, \{\nabla H_l(x)\}_{l\in\bar{I}_{00}^{++}(x)}, \{\nabla G_l(x)\}_{l\in\bar{I}_{00}^{+0}(x)}, \{\nabla G_l(x)\}_{l\in\bar{I}_{00}^{-0}(x)}, \\
&\{\nabla H_l(x)\}_{l\in\bar{I}_{00}^{0+}(x)}, \{\nabla H_l(x)\}_{l\in\bar{I}_{00}^{0-}(x)}
\end{aligned}
$$

are linearly independent and

$$
\begin{aligned}
-\gamma_{l_1}\nabla G_{l_1}(x) = {} & [\sum_{l\in\bar{L}_+^G(x)} \bar{\gamma}_l \nabla G_l(x) + \sum_{l\in\bar{L}_-^G(x)} \bar{\gamma}_l \nabla G_l(x)] \\
& +[\sum_{l\in\bar{L}_+^H(x)} \bar{\nu}_l \nabla H_l(x) + \sum_{l\in\bar{L}_-^H(x)} \bar{\nu}_l \nabla H_l(x)] + \sum_{l\in\bar{I}_{00}^{++}(x)} [\bar{\gamma}_l \nabla G_l(x) + \bar{\nu}_l \nabla H_l(x)] \\
& +[\sum_{l\in\bar{I}_{00}^{+0}(x)} \bar{\gamma}_l \nabla G_l(x) + \sum_{l\in\bar{I}_{00}^{-0}(x)} \bar{\gamma}_l \nabla G_l(x)] + [\sum_{l\in\bar{I}_{00}^{0+}(x)} \bar{\nu}_l \nabla H_l(x) + \sum_{l\in\bar{I}_{00}^{0-}(x)} \bar{\nu}_l \nabla H_l(x)]
\end{aligned}
$$

with

$$\bar{\gamma}_l > 0, \forall l \in \bar{L}_+^G(x), \bar{\gamma}_l < 0, \forall l \in \bar{L}_-^G(x), \bar{\nu}_l > 0, \forall l \in \bar{L}_+^H(x), \nu_l < 0, \forall l \in \bar{L}_-^H(x),$$
$$\bar{\gamma}_l > 0, \nu_l > 0, \forall l \in \bar{I}_{00}^{++}(x), \bar{\gamma}_l > 0, \forall l \in \bar{I}_{00}^{+0}(x), \bar{\gamma}_l < 0, \forall l \in \bar{I}_{00}^{-0}(x),$$
$$\bar{\nu}_l > 0, \forall l \in \bar{I}_{00}^{0+}(x), \bar{\nu}_l < 0, \forall l \in \bar{I}_{00}^{0-}(x).$$

By the linear independence of the vectors and continuity arguments, the vectors

$$\{\nabla G_l(y)\}_{l\in \bar{L}_+^G(x)}, \{\nabla G_l(y)\}_{l\in \bar{L}_-^G(x)}, \{\nabla H_l(y)\}_{l\in \bar{L}_+^H(x)}, \{\nabla H_l(x)\}_{l\in \bar{L}_-^H(x)},$$
$$\{\nabla G_l(y)\}_{l\in \bar{I}_{00}^{++}(x)}, \{\nabla H_l(y)\}_{l\in \bar{I}_{00}^{++}(x)}, \{\nabla G_l(y)\}_{l\in \bar{I}_{00}^{+0}(x)}, \{\nabla G_l(y)\}_{l\in \bar{I}_{00}^{-0}(x)},$$
$$\{\nabla H_l(y)\}_{l\in \bar{I}_{00}^{0+}(x)}, \{\nabla H_l(y)\}_{l\in \bar{I}_{00}^{0-}(x)}$$

are linearly independent for all y in a neighborhood of x. However, by the MPEC-CPLD assumption, the vectors

$$\gamma_{i_1}\nabla G_{i_1}(y), \{\nabla G_l(y)\}_{l\in \bar{L}_+^G(x)}, \{\nabla G_l(y)\}_{l\in \bar{L}_-^G(x)}, \{\nabla H_l(y)\}_{l\in \bar{L}_+^H(x)}, \{\nabla H_l(x)\}_{l\in \bar{L}_-^H(x)},$$
$$\{\nabla G_l(y)\}_{l\in \bar{I}_{00}^{++}(x)}, \{\nabla H_l(y)\}_{l\in \bar{I}_{00}^{++}(x)}, \{\nabla G_l(y)\}_{l\in \bar{I}_{00}^{+0}(x)}, \{\nabla G_l(y)\}_{l\in \bar{I}_{00}^{-0}(x)},$$
$$\{\nabla H_l(y)\}_{l\in \bar{I}_{00}^{0+}(x)}, \{\nabla H_l(y)\}_{l\in \bar{I}_{00}^{0-}(x)}$$

are linearly dependent for all y in a neighborhood of x. Therefore, $\lambda_{i_1}\nabla G_{i_1}(y)$ must be a linear combination of the vectors for all y in a neighborhood of x. By [3, Lemma 3.2], there exists a smooth function φ defined in a neighborhood of $(0,\ldots,0)$ such that, for all y in a neighborhood of x,

$$\begin{aligned}\nabla\varphi(0,\ldots,0) = \quad & (\{\bar{\gamma}\}_{l\in \bar{L}_+^G(x)}, \{\bar{\gamma}\}_{l\in \bar{L}_-^G(x)}, \{\bar{\nu}\}_{l\in \bar{L}_+^H(x)}, \{\bar{\nu}\}_{l\in \bar{L}_-^H(x)}, \{\bar{\gamma}\}_{l\in \bar{I}_{00}^{++}(x)}, \\ & \{\bar{\nu}\}_{l\in \bar{I}_{00}^{++}(x)}, \{\bar{\gamma}\}_{l\in \bar{I}_{00}^{+0}(x)}, \{\bar{\gamma}\}_{l\in \bar{I}_{00}^{-0}(x)}, \{\bar{\nu}\}_{l\in \bar{I}_{00}^{0+}(x)}, \{\bar{\nu}\}_{l\in \bar{I}_{00}^{0-}(x)}),\end{aligned}$$

$$
\begin{aligned}
-\lambda_{i_1} G_{i_1}(y) = \quad & \varphi\Big(\{G_l(y)\}_{l\in \bar{L}_+^G(x)}, \{G_l(y)\}_{l\in \bar{L}_-^G(x)}, \{H_l(y)\}_{l\in \bar{L}_+^H(x)}, \{H_l(y)\}_{l\in \bar{L}_-^H(x)},\\
& \{G_l(y)\}_{l\in \bar{I}_{00}^{++}(x)}, \{H_l(y)\}_{l\in \bar{I}_{00}^{++}(x)}, \{G_l(y)\}_{l\in \bar{I}_{00}^{+0}(x)}, \{G_l(y)\}_{l\in \bar{I}_{00}^{-0}(x)},\\
& \{H_l(y)\}_{l\in \bar{I}_{00}^{0+}(x)}, \{H_l(y)\}_{l\in \bar{I}_{00}^{0-}(x)}\Big).
\end{aligned}
$$

Now suppose that $\{x^k\}$ is an infeasible sequence that converges to x and such that

$$
\begin{aligned}
& G_l(x^k) > 0, \forall l \in \bar{L}_+^G(x), \quad G_l(x^k) < 0, \forall l \in \bar{L}_-^G(x),\\
& H_l(x^k) > 0, \forall l \in \bar{L}_+^H(x), \quad H_l(x^k) < 0, \forall l \in \bar{L}_-^G(x),\\
& G_l(x^k) < 0, H_l(x^k) < 0, \forall l \in \bar{I}_{00}^{++}(x),\\
& G_l(x^k) < 0, \forall l \in \bar{I}_{00}^{+0}(x), \quad G_l(x^k) > 0, \forall l \in \bar{I}_{00}^{-0}(x),\\
& H_l(x^k) < 0, \forall l \in \bar{I}_{00}^{0+}(x), \quad H_l(x^k) > 0, \forall l \in \bar{I}_{00}^{0-}(x).
\end{aligned}
$$

By virtue of Tarloy's expansion of φ at $(0,\dots,0)$, for k large enough, we must have $-\lambda_{i_1} G_{i_1}(x^k) \geq 0$. Again, there is no sequence $x^k \to x$ such that $\lambda_{i_1} G_{i_1}(x^k) > 0$.

The proofs for the other cases are entirely analogous to the proof for this case. Therefore, MPEC-CPLD implies MPEC generalized quasi-normality. □

The following result follows immediately from Theorem 4.4 and the definitions of the three constraint qualifications.

Corollary 4.5. *Let x^* be a local minimizer of problem* (MPEC). *If x^* satisfies MPEC-CPLD, or is MPEC generalized pseudo-normal, or MPEC generalized quasi-normal, then x^* is an enhanced M-stationary point.*

4.3 MPEC error bound

As one of their main results, Kanzow and Schwartz proved in [42] that the MPEC generalized pseudo-normality implies the existence of a local error bound for smooth

MPECs. Combining the proof techniques of Theorem 2.1 in Chapter 2 and [42, Theorem 4.5], we can extend [42, Theorem 4.5] to the nonsmooth MPEC. The MPEC generalized quasi-normality is weaker than the MPEC generalized pseudo-normality. It is desirable to find conditions under which the existence of a local error bound holds under the MPEC generalized quasi-normality. We will answer this question in Theorem 2.10. Before we can do so, we need to prove some preliminary results, which will facilitate the proof of Theorem 2.10.

Lemma 4.6. *If a feasible point x^* is MPEC generalized quasi-normal, then all feasible points in a neighborhood of x^* are MPEC generalized quasi-normal.*

Proof. For simplicity, we drop the equality and the inequality constraints in the proof. Assume that the claim is not true. Then we can find a sequence $\{x^k\}$ such that $x^k \neq x^*$ for all k, $x^k \to x^*$ and x^k is not quasi-normal for all k. This implies, for each k, the existence of scalars $\{\gamma^k, \nu^k\}$ not zero and a sequence $\{x^{k,t}\} \subseteq \mathcal{X}$ such that

(1) $0 \in -\sum_{l=1}^m [\gamma_l^k \nabla G_l(x^k) + \nu_l^k \nabla H_l(x^k)] + \mathcal{N}_{\mathcal{X}}(x^k)$,

(2) $\gamma_l^k = 0 \ \forall l \in I_{+0}(x^k)$, $\nu_l^k = 0 \ \forall l \in I_{0+}(x^k)$ and either $\gamma_l^k \nu_l^k = 0$ *or* $\gamma_l^k > 0$, $\nu_l^k > 0 \ \forall l \in I_{00}(x^k)$,

(3) $\{x^{k,t}\}$ converges to x^k as $t \to \infty$, and for each t, $-\gamma_l^k G_l(x^{k,t}) > 0, \forall l \in \mathcal{G}^k$, $-\nu_l^k H_l(x^{k,t}) > 0, \forall l \in \mathcal{H}^k$, where $\mathcal{G}^k = \{l | \gamma_l^k \neq 0\}$ and $\mathcal{H}^k = \{l | \nu_l^k \neq 0\}$.

For each k, denote by $\tilde{\gamma}^k := \frac{\gamma^k}{\|(\gamma^k, \nu^k)\|}$, $\tilde{\nu}^k := \frac{\nu^k}{\|(\gamma^k, \nu^k)\|}$. Assume without any loss of generality that $(\tilde{\gamma}^k, \tilde{\nu}^k) \to (\gamma^*, \nu^*)$. Dividing both sides of (1) above by $\|(\gamma^k, \nu^k)\|$ and taking the limit, we have

(1) $0 \in -\sum_{l=1}^m [\gamma_l^* \nabla G_l(x^*) + \nu_l^* \nabla H_l(x^*)] + \mathcal{N}_{\mathcal{X}}(x^*)$,

(2) $\gamma_l^* = 0 \ \forall l \in I_{+0}(x^*)$, $\nu_l^* = 0 \ \forall l \in I_{0+}(x^*)$ and either $\gamma_l^* \nu_l^* = 0$ *or* $\gamma_l^* > 0$, $\nu_l^* > 0 \ \forall l \in I_{00}(x^*)$,

(3) $\{\varsigma^k\}$ converges to x^* as $k \to \infty$, and for each l, $-\gamma_l^* G_l(\varsigma^k) > 0$, $\forall l \in \mathcal{G}$, $-\nu_l^* H_l(\varsigma^k) > 0, \forall l \in \mathcal{H}$, where $\mathcal{G} = \{l | \gamma_l^* \neq 0\}, \mathcal{H} = \{l | \nu_l^* \neq 0\}$.

Indeed, for indices $l \in I_{00}(x^*)$, for each k,

$$\begin{aligned} \tilde{\gamma}_l^k = 0,\ \tilde{\nu}_l^k\ free, \qquad & if \quad l \in I_{+0}(x^k), \\ \tilde{\gamma}_l^k\ free,\ \tilde{\nu}_l^k = 0, \qquad & if \quad l \in I_{0+}(x^k), \\ either\ \tilde{\gamma}_l^k \tilde{\nu}_l^k = 0\ or\ \tilde{\gamma}_l^k > 0,\ \tilde{\nu}_l^k > 0, \qquad & if \quad l \in I_{00}(x^k), \end{aligned}$$

and hence that either $\gamma_l^* \nu_l^* = 0$ or $\gamma_l^* > 0$, $\nu_l^* > 0$ $\forall l \in I_{00}(x^*)$. The existence of scalars $\{\gamma^*, \nu^*\}$ and sequence $\{\varsigma^k\}$ violates the MPEC quasi-normality of x^*, thus completing the proof. □

In the following result, we obtain a specific representation of the limiting normal cone to the constraint region in terms of the set of MPEC quasi-normal multipliers.

Proposition 4.3.1. *If $\bar{x}$ is MPEC generalized quasi-normal for $\mathfrak{F}$, then $\mathcal{N}_{\mathfrak{F}}(\bar{x})$ belongs to the set*

$$\left\{ \sum_{i=1}^{p} \partial(\lambda_i h_i)(\bar{x}) + \sum_{j=1}^{q} \mu_j \partial g_j(\bar{x}) - \sum_{l=1}^{m} [\gamma_l \nabla G_t(\bar{x}) + \nu_l \nabla H_t(\bar{x})] + \mathcal{N}_{\mathcal{X}}(\bar{x}) | (\lambda, \mu, \gamma, \nu) \right\},$$

where $(\lambda, \mu, \gamma, \nu) \in M_Q(\bar{x})$ denotes the set of quasi-normal multipliers corresponding to $\bar{x}$.

Proof. For simplicity, we omit the equality and the inequality constraints in the proof. Let v be an element of set $\mathcal{N}_{\mathfrak{F}}(\bar{x})$. By definition, there are sequences $x^l \to \bar{x}$ and $v^l \to v$ with $v^l \in \widehat{\mathcal{N}}_{\mathfrak{F}}(x^l)$ and $x^l \in \mathfrak{F}$.

Step 1. By Lemma 4.6, for l sufficiently large, x^l is MPEC generalized quasi-normal. By [74, Theorem 6.11], for each l, there exists a smooth function φ^l that achieves a strict global minimum over $\mathfrak{F}$ at x^l with $-\nabla\varphi^l(x^l) = v^l$. Since x^l is a MPEC generalized quasi-normal vector of $\mathfrak{F}$, by Theorem 4.1, enhanced M stationary condition holds for problem $\min \varphi^l(x)$ s.t. $x \in \mathfrak{F}$. That is, there exists a vector (γ^l, ν^l) such that

$$v^l \in -\sum_{t=1}^{m}[\gamma_t^l \nabla G_t(x^l) + \nu_t^l \nabla H_t(x^l)] + \mathcal{N}_{\mathcal{X}}(\bar{x}), \tag{4.5}$$

with $\gamma_t^l = 0\ \forall t \in I_{+0}(x^l)$, $\nu_t^l = 0\ \forall l \in I_{0+}(x^l)$ and either $\gamma_t^l \nu_t^l = 0$ *or* $\gamma_t^l > 0$, $\nu_t^l > 0\ \forall l \in I_{00}(x^l)$. Moreover let $\mathcal{G}^l = \{l | \gamma_t^l \neq 0\}, \mathcal{H}^l = \{t | \nu_t^l \neq 0\}$, then there exists a sequence $\{x^{l,k}\}$ converging to x^l as $k \to \infty$ such that for all k, $-\gamma_t^l G_t(x^{l,k}) > 0, \forall t \in \mathcal{G}^l$, $-\nu_t^l H_t(x^{l,k}) > 0, \forall t \in \mathcal{H}^l$.

Step 2. We show that the sequence $\{\gamma^l, \nu^l\}$ is bounded. To the contrary, suppose that the sequence $\{\gamma^l, \nu^l\}$ is unbounded. For every l, denote by $\tilde{\gamma}^l := \frac{\gamma^l}{\|(\gamma^l,\nu^l)\|}$, $\tilde{\nu}^l := \frac{\nu^l}{\|(\gamma^l,\nu^l)\|}$. Assume without any loss of generality that $(\tilde{\gamma}^l, \tilde{\nu}^l) \to (\gamma^*, \nu^*)$. Dividing both sides of (4.5) by $\|(\gamma^l, \nu^l)\|$ and taking the limit; similarly to the proof of Lemma 4.6, we obtain

(1) $0 \in -\sum_{t=1}^m [\gamma_t^* \nabla G_t(\bar{x}) + \nu_t^* \nabla H_t(\bar{x})] + \mathcal{N}_{\mathcal{X}}(\bar{x})$,

(2) $\gamma_t^* = 0\ \forall t \in I_{+0}(\bar{x})$, $\nu_t^* = 0\ \forall t \in I_{0+}(\bar{x})$ and either $\gamma_t^* \nu_t^* = 0$ *or* $\gamma_t^* > 0$, $\nu_t^* > 0\ \forall t \in I_{00}(\bar{x})$,

(3) $\{\varsigma^l\}$ converges to $\bar{x}$ as $l \to \infty$, and for each l, $-\gamma_t^* G_t(\varsigma^l) > 0$, $\forall t \in \mathcal{G}$, $-\nu_t^* H_t(\varsigma^l) > 0, \forall t \in \mathcal{H}$, where $\mathcal{G} = \{t | \gamma_t^* \neq 0\}, \mathcal{H} = \{t | \nu_t^* \neq 0\}$.

However, this is impossible since $\bar{x}$ is assumed to be MPEC quasi-normal, and hence the sequence $\{\gamma^l, \nu^l\}$ must be bounded.

Step 3. By virtue of Step 2, without any loss of generality, we assume that $\{\gamma^l, \nu^l\}$

converges to $\{\gamma, \nu\}$ as $l \to \infty$. Taking the limit in (4.5) as $l \to \infty$, we have

$$v \in -\sum_{l=1}^{m}[\gamma_l \nabla G_l(\bar{x}) + \nu_l \nabla H_l(\bar{x})] + \mathcal{N}_{\mathcal{X}}(\bar{x}).$$

with $\gamma_t = 0\ \forall t \in I_{+0}(\bar{x})$, $\nu_t = 0\ \forall t \in I_{0+}(\bar{x})$ and either $\gamma_t \nu_t = 0$ *or* $\gamma_t > 0$, $\nu_t > 0\ \forall t \in I_{00}(\bar{x})$. Similarly to Step 2, we can find a subsequence $\{\varsigma^l\}$ that converges to $\bar{x}$ as $l \to \infty$, and for each l,

$$-\gamma_t G_t(\varsigma^l) > 0,\ \forall t \in \mathcal{G}, \qquad -\nu_t H_t(\varsigma^l) > 0, \forall t \in \mathcal{H},$$

where $\mathcal{G} = \{t | \gamma_t \neq 0\}, \mathcal{H} = \{t | \nu_t \neq 0\}$. □

Taking into account the previous two results, we are now able to obtain a local error bound result for MPECs under the MPEC quasi-normality.

Theorem 4.7. *Let $x^* \in \mathfrak{F}$, the feasible region of problem* (MPEC)*. Assume that h_i are C^1, $g_j(x)$ are subdifferentially regular around x^* in the sense of [61, Definition 1.91(i)] (automatically holds when g_j are convex or C^1 around x^*), $\mathcal{X}$ is a nonempty, closed and regular in the sense that $\mathcal{N}_{\mathcal{X}}(x) = \widehat{\mathcal{N}}_{\mathcal{X}}(x)$ for all $x \in \Omega$. If x^* is MPEC generalized quasi-normal and the strict complementarity condition holds at x^*, then there are $\delta, c > 0$ such that for each $x \in \mathcal{B}_{\frac{\delta}{2}}(x^*) \cap \mathcal{X}$,*

$$\text{dist}_{\mathfrak{F}}(x) \quad \leq \quad c(\|h(x)\|_1 + \|g^+(x)\|_1 + \sum_{l=1}^{m} dist_C(G_l(x), H_l(x))), \tag{4.6}$$

where $C := \{(a, b) \in \mathbb{R} | a \geq 0, b \geq 0, ab = 0\}$.

Proof. For simplicity, we omit the equality constraints in the proof. By assumption we can find $\delta_0 > 0$ such that $g_j(x)$ are subdifferentially regular for all $x \in \mathcal{B}_{\delta_0}(x^*)$. Since the required assertion is always true if x^* is in the interior of set $\mathfrak{F}$, we only need to consider the case when x^* is in the boundary of $\mathfrak{F}$. In this case, (4.6) can be violated only for $x \notin \mathfrak{F}$. Let us take some sequences $\{\mathbb{x}^k\}$ and $\{x^k\}$, such that $\mathbb{x}^k \to x^*$, $\mathbb{x}^k \in \mathcal{X} \setminus \mathfrak{F}$, and $x^k = \prod_{\mathfrak{F}}(\mathbb{x}^k)$, the projection of $\mathbb{x}^k$ onto the set $\mathfrak{F}$. Note that $x^k \to x^*$, since $\|x^k - \mathbb{x}^k\| \leq \|\mathbb{x}^k - x^*\|$. For simplicity, we may assume both $\{\mathbb{x}^k\}$ and $\{x^k\}$ belong to $\mathcal{B}_{\delta_0}(x^*) \cap \mathcal{X}$.

Since $\mathbb{x}^k - x^k \in \mathcal{N}^\pi_{\mathfrak{F}}(x^k) \subseteq \widehat{\mathcal{N}}_{\mathfrak{F}}(x^k)$, we have $\eta^k = \frac{\mathbb{x}^k - x^k}{\|\mathbb{x}^k - x^k\|} \in \widehat{\mathcal{N}}_{\mathfrak{F}}(x^k)$.

Since x^* is quasi-normal, it follows from Lemma 4.6 that the point x^k is also quasi-normal for all sufficiently large k and, without any loss of generality, we may assume that all x^k are quasi-normal. Then, employing Proposition 4.3.1, there exists a sequence $\{\mu^k, \gamma^k, \nu^k\}$ such that

$$\eta^k \in \sum_{j=1}^{q} \mu_j^k \partial g_j(x^k) - \sum_{l=1}^{m} [\gamma_l^k \nabla G_l(x^k) + \nu_l^k \nabla H_l(x^k)] + \mathcal{N}_{\mathcal{X}}(x^k), \tag{4.7}$$

$\mu^k \geq 0, \mu_j^k = 0\ \forall j \notin A(x^k)$, $\gamma_l^k = 0\ \forall l \in I_{+0}(x^k)$, $\nu_l^k = 0\ \forall l \in I_{0+}(x^k)$ and either $\gamma_l^k \nu_l^k = 0$ or $\gamma_l^k > 0$, $\nu_l^k > 0\ \forall l \in I_{00}(x^k)$, and there exists a sequence $\{x^{k,s}\} \subseteq \mathcal{X}$, such that $x^{k,s} \to x^k$ as $s \to \infty$ and for all s, $\mu_j^k g_j(x^{k,s}) > 0$ for $j \in J^k$ and $-\gamma_l^k G_l(x^{k,s}) > 0$, $\forall l \in \mathcal{G}^k$, $-\nu_l^k H_l(x^{k,s}) > 0, \forall l \in \mathcal{H}^k$, where $J^k = \{j | \mu_j^k > 0\}$ and $\mathcal{G}^k = \{l | \gamma_l^k \neq 0\}, \mathcal{H}^k = \{l | \nu_l^k \neq 0\}$. As in the proof of Step 2 in Proposition 4.3.1, we can show that the quasi-normality of x^* implies that the sequence $\{\mu^k, \gamma^k, \nu^k\}$ is bounded. Therefore, without any loss of generality, we may assume $\{\mu^k, \gamma^k, \nu^k\}$ converges to some vector $\{\mu^*, \gamma^*, \nu^*\}$. Then there exists a number $M_0 > 0$, such that for all k, $\|(\mu^k, \gamma^k, \nu^k)\| \leq M_0$. Without any loss of generality, we may assume that $\mathbb{x}^k \in \mathcal{B}_{\frac{\delta_0}{2}}(x^*) \setminus \mathfrak{F}$ and $x^k \in \mathcal{B}_{\delta_0}(x^*)$ for all k. Setting $(\bar{\mu}^k, \bar{\gamma}^k, \bar{\nu}^k) = 2(\mu^k, \gamma^k, \nu^k)$, then from (5.4), for each k, there exist $\rho_j^k \in \partial g_j(x^k), \forall j = 1, \ldots, q$ and $\omega^k \in \mathcal{N}_{\mathcal{X}}(x^k)$ such that

$$\frac{\mathbb{x}^k - x^k}{\|\mathbb{x}^k - x^k\|} = \frac{x^k - \mathbb{x}^k}{\|\mathbb{x}^k - x^k\|} + \sum_{j=1}^{q} \bar{\mu}_j^k \rho_j^k - \sum_{l=1}^{m} [\bar{\gamma}^k \nabla G_l(x^k) + \bar{\nu}_l^k \nabla H_l(x^k)] + \omega^k.$$

We obtain from the discussion above that

$$\begin{aligned}
&\|\mathbb{x}^k - x^k\| = \langle \frac{\mathbb{x}^k - x^k}{\|\mathbb{x}^k - x^k\|}, \mathbb{x}^k - x^k \rangle \\
&= \Big\langle \frac{x^k - \mathbb{x}^k}{\|\mathbb{x}^k - x^k\|}, \mathbb{x}^k - x^k \Big\rangle + \sum_{j=1}^{q} \Big\langle \bar{\mu}_j^k \rho_j^k, \mathbb{x}^k - x^k \Big\rangle \\
&\quad - \sum_{l=1}^{m} \Big\langle \bar{\gamma}_l^k \nabla G_l(x^k) + \bar{\nu}_l^k \nabla H_l(x^k), \mathbb{x}^k - x^k \Big\rangle + \Big\langle \omega^k, \mathbb{x}^k - x^k \Big\rangle \\
&\leq \sum_{j=1}^{q} \Big\langle \bar{\mu}_j^k \rho_j^k, \mathbb{x}^k - x^k \Big\rangle - \sum_{l=1}^{m} \Big\langle \bar{\gamma}_l^k \nabla G_l(x^k) + \bar{\nu}_l^k \nabla H_l(x^k), \mathbb{x}^k - x^k \Big\rangle + o(\|\mathbb{x}^k - x^k\|) \\
&\leq \sum_{j=1}^{q} \bar{\mu}_j^k \Big(g_j(\mathbb{x}^k) + o(\|\mathbb{x}^k - x^k\|) \Big) - \sum_{l=1}^{m} \bar{\gamma}_l^k \Big(G_l(\mathbb{x}^k) + o(\|\mathbb{x}^k - x^k\|) \Big) \\
&\quad - \sum_{l=1}^{m} \bar{\nu}_l^k \Big(H_l(\mathbb{x}^k) + o(\|\mathbb{x}^k - x^k\|) \Big) + o(\|\mathbb{x}^k - x^k\|) \\
&\leq 2\Big[\sum_{j=1}^{q} \mu_j^k g_j(\mathbb{x}^k) - \sum_{l=1}^{m} \Big(\gamma_l^k G_l(\mathbb{x}^k) + \nu_l^k H_l(\mathbb{x}^k) \Big) \Big] \\
&\quad + 2\Big| \sum_{j=1}^{q} \mu_j^k + \sum_{l=1}^{m} \gamma_l^k + \sum_{l=1}^{m} \nu_l^k + 1 \Big| o(\|\mathbb{x}^k - x^k\|) \\
&\leq 2\Big[\sum_{j=1}^{q} \mu_j^k g_j(\mathbb{x}^k) - \sum_{l=1}^{m} \Big(\gamma_l^k G_l(\mathbb{x}^k) + \nu_l^k H_l(\mathbb{x}^k) \Big) \Big] + \frac{1}{2}\|\mathbb{x}^k - x^k\|,
\end{aligned}$$

where the first inequality comes from the fact that $\mathcal{X}$ is regular, the second comes from the subdifferential regularity assumption of $g_j(x)$ in $\mathcal{B}_{\delta_0}(x^*)$, and the last one is valid because, without any loss of generality, we may assume for k sufficiently large,

$o(\|\mathbb{x}^k - x^k\|) \leq \frac{1}{4(M_0+1)}\|\mathbb{x}^k - x^k\|$ since $\mathbb{x}^k - x^k \to 0$ as k tends to infinity. This means

$$\mathrm{dist}_{\mathfrak{F}}(\mathbb{x}^k) = \|\mathbb{x}^k - x^k\| \leq 4M_0\Big(\sum_{i=1}^{q} g_j^+(\mathbb{x}^k) + \phi(G(\mathbb{x}^k), H(\mathbb{x}^k))\Big),$$

where

$$\phi(G(\mathbb{x}^k), H(\mathbb{x}^k)) = \sum_{l=1}^{m} \max\{-G_l(\mathbb{x}^k), -H_l(\mathbb{x}^k), G_l(\mathbb{x}^k) - H_l(\mathbb{x}^k), \min\{G_l(\mathbb{x}^k), H_l(\mathbb{x}^k)\}\}.$$

Thus, for any sequence $\{\mathbb{x}^k\} \subseteq \mathcal{X}$ converging to x^* there exists a number $c > 0$ such that

$$\mathrm{dist}_{\mathfrak{F}}(\mathbb{x}^k) \leq c(\|g^+(\mathbb{x}^k)\|_1 + \sum_{l=1}^{m} dist_C(G_l(\mathbb{x}^k), H_l(\mathbb{x}^k))) \quad \forall k = 1, 2, \ldots.$$

This further implies the error bound property at x^*. Indeed, suppose the contrary. Then there exists a sequence $\tilde{\mathbb{x}}^k \to x^*$, such that $\tilde{\mathbb{x}}^k \in \mathcal{X} \setminus \mathfrak{F}$ and

$$\mathrm{dist}_{\mathfrak{F}}(\tilde{\mathbb{x}}^k) > c(\|g^+(\tilde{\mathbb{x}}^k)\|_1 + \sum_{l=1}^{m} dist_C(G_l(\tilde{\mathbb{x}}^k), H_l(\tilde{\mathbb{x}}^k)))$$

for all $k = 1, 2, \ldots,$ which is a contradiction. □

4.4 Sensitivity analysis for nonsmooth MPEC

This section considers the following mathematical program with equilibrium constraints formulated as a mathematical program with complementarity constraints subject to perturbation p:

$$(\text{MPEC}_p) \qquad \min_{x\in\mathcal{X}} \quad f(x,p)$$
$$\text{s.t.} \quad g(x,p) \le 0,\ \ h(x,p) = 0,$$
$$0 \le G(x,p) \ \perp\ H(x,p) \ge 0,$$

where $f : \mathbb{R}^{n+\mathbb{n}} \to \mathbb{R},\ g : \mathbb{R}^{n+\mathbb{n}} \to \mathbb{R}^q,\ h : \mathbb{R}^{n+\mathbb{n}} \to \mathbb{R}^p$, and $G, H : \mathbb{R}^{n+\mathbb{n}} \to \mathbb{R}^m$, $\mathcal{X}$ is a nonempty and closed subset of $\mathbb{R}^n$. In this section, we assume that all the involved functions $\{f, g, h, G, H\}$ are only Lipschitzian around the point of interest. (MPEC$_p$) is called a parametric mathematical program with complementarity constrains. For given p, we denote by $\mathfrak{F}(p)$ the feasible region of (MPEC$_p$). The value function of (MPEC$_p$) is an extended-valued function defined by

$$\mathcal{V}(p) := \inf\{f(x,p) \mid x \in \mathfrak{F}(p)\}$$

and the optimal solution mapping is a set-valued mapping defined by

$$\mathcal{O}(p) := \{x \in \mathfrak{F}(p) \mid f(x,p) = \mathcal{V}(p)\}.$$

For sake of simplicity, we denote

$$F(x,p) := \begin{pmatrix} g(x,p) \\ h(x,p) \\ \Psi(x,p) \end{pmatrix}, \quad \Lambda := \mathbb{R}^q_- \times \{0\}^p \times C^m, \tag{4.8}$$

where $\mathbb{R}_-$ denotes the nonpositive orthant $\{v \in \mathbb{R} \mid v \leq 0\}$ and

$$\Psi(x,p) := \begin{pmatrix} G_1(x,p) \\ H_1(x,p) \\ \vdots \\ G_m(x,p) \\ H_m(x,p) \end{pmatrix}, \quad C := \{\,(a,b) \in \mathbb{R}^2 \mid 0 \leq a \perp b \geq 0\}. \tag{4.9}$$

Thus, the feasible region of (MPEC$_p$) can be rewritten as $\mathfrak{F}(p) := \{x \in \mathcal{X} \mid F(x,p) \in \Lambda\}$. For a given feasible point $x^* \in \mathfrak{F}(p^*)$, we define the following index sets:

$$\begin{cases} A^* := \{\, j \mid g_j(x^*,p^*) = 0\}, \\ I^*_{0+} := \{\, l \mid G_l(x^*,p^*) = 0 < H_l(x^*,p^*)\}, \\ I^*_{00} := \{\, l \mid G_l(x^*,p^*) = 0 = H_l(x^*,p^*)\}, \\ I^*_{+0} := \{\, l \mid G_l(x^*,p^*) > 0 = H_l(x^*,p^*)\}. \end{cases}$$

The *MPEC generalized Lagrangian function* of (MPEC$_p$) is given by

$$\mathscr{L}^r(x,p;\lambda,\mu,\gamma,\nu) := rf(x,p) + g(x,p)^T\mu + h(x,p)^T\lambda - G(x,p)^T\gamma - H(x,p)^T\nu. \quad r \geq 0,$$

Compared with the developments on optimality conditions, algorithms, and stability, there are only a few publications on the sensitivity of the value function for (MPEC$_p$). Lucet and Ye [47,48] addressed the sensitivity of the value function for optimization programs with variational inequality constraints (OPVIC), which includes MPEC as a special case. They established an upper estimate of the limiting subdifferential of value function in terms of the normal coderivative multipliers for OPVIC. For (MPEC$_p$), they gave upper estimates for the singular subdifferential and limiting subdifferential of the value function by C-, M-, and S-multipliers under the conditions that the growth condition and some normality conditions hold. In this section, we obtain some sharper upper estimates for the singular subdifferential and the limiting subdifferential of the value function for (MPEC$_p$) based on the enhanced Fritz John condition for MPECs under the weaker conditions that the restricted inf-compactness and some quasinormality conditions hold. For sake of simplicity, we denote

$$\begin{aligned}\partial_x \mathscr{L}^r(x,p;\lambda,\mu,\gamma,\nu) \;:=\; & r\partial_x f(x,p) + \sum_{j=1}^{q} \mu_j \partial_x g_j(x,p) + \sum_{i=1}^{p} \partial_x(\lambda_i h_i)(x,p) \\ & + \sum_{l=1}^{m} \partial_x(\gamma_l G_l)(x,p) + \sum_{l=1}^{m} \partial_x(\nu_l H_l)(x,p).\end{aligned}$$

Note that the limiting subdifferential of $\mathscr{L}^r$ at $(x,p,\lambda,\mu,\gamma,\nu)$ with respect to x is not equal to the right hand side of the above equation and, for simplicity, we use all plus signs in the formula above in contrast with the standard MPEC Lagrangian function.

4.4.1 Subdifferential via enhanced M-multipliers

Let us first give the enhanced Fritz John type M-stationary condition for (MPEC$_{p^*}$). In fact, Kanzow and Schwartz [42, Theorem 3.1] have presented the smooth enhanced Fritz John type M-stationary condition. In the following, we show that, for the nonsmooth case, any local minimizer for MPEC is also an enhanced Fritz John type M-stationary point.

An enhanced Fritz John optimality condition is given for a very general mathematical program with geometric constraints in Banach spaces in chapter 3. We next specialize this result to MPEC. By means of (4.8)–(4.9), (MPEC$_p$) can be rewritten as the more compact form

$$\begin{aligned}\min_{x\in\mathcal{X}} \quad & f(x,p) \\ \text{s.t.} \quad & F(x,p)\in\Lambda.\end{aligned} \tag{4.10}$$

Since problem (4.10) is of the form considered in chapter 3, we can specialize Corollary 3.3 in section 3.3 to problem (4.10).

Theorem 4.8. *Let* $x^* \in \mathfrak{F}(p^*)$ *be a local minimizer of* (MPEC_{p^*}). *Then there exists* $0 \neq (r, \lambda^*, \mu^*, \gamma^*, \nu^*)$ *with* $r \geq 0$ *such that*

(1) $0 \in \partial_x \mathscr{L}^r(x^*, p^*; \lambda^*, \mu^*, \gamma^*, \nu^*) + \mathcal{N}_{\mathcal{X}}(x^*)$, $\mu^* \geq 0$, $\mu_j^* = 0$ $(j \notin A^*)$, $\gamma_l^* = 0$ $(l \in I_{+0}^*)$, $\nu_i^* = 0$ $(l \in I_{0+}^*)$, $\gamma_l^* \nu_l^* = 0$ or $\gamma_l^* < 0$, $\nu_l^* < 0$ $(i \in I_{00}^*)$;

(2) *There exists a sequence* $\{x^k\} \subseteq \mathcal{X}$ *converging to* x^* *such that, for each* k,

$$\mu_j^* > 0 \Longrightarrow \mu_j^* g_j(x^k, p^*) > 0, \quad \lambda_i^* > 0 \Longrightarrow \lambda_i^* h_i(x^k, p^*) > 0$$
$$\gamma_l^* \neq 0 \Longrightarrow \gamma_l^* G_l(x^k, p^*) > 0, \quad \nu_l^* \neq 0, \Longrightarrow \nu_l^* H_l(x^k, p^*) > 0,$$

and $\{h, g, G, H\}$ *are all differentiable with respect to* x *at* (x^k, p^*).

Proof. Since x^* is a local minimizer of (4.10) for $p = p^*$, by Corollary 3.3 in chapter 3, there exist a scalar $r \geq 0$ and a vector η^*, not all zero, such that the following conditions hold, where $\{e_i \mid i = 1, \cdots, q+p+2m\}$ is the orthogonal basis of $\mathbb{R}^{q+p+2m}$:

(i) $0 \in r\partial_x f(x^*, p^*) + \sum_{i=1}^{q+p+2m} \partial_x \langle \eta^*, e_i \rangle \langle F, e_i \rangle (x^*, p^*) + \mathcal{N}_{\mathcal{X}}(x^*)$;

(ii) $\eta^* \in \mathcal{N}_\Lambda(F(x^*, p^*))$;

(iii) There exists a sequence $\{(x^k, y^k, \eta^k)\}$ converging to $(x^*, F(x^*), \eta^*)$ such that, for all k,

$$f(x^k, p^*) < f(x^*, p^*),$$
$$\eta^k \in \mathcal{N}_\Lambda(y^k), \tag{4.11}$$
$$\langle \eta^*, e_i \rangle \neq 0 \Longrightarrow \langle \eta^*, e_i \rangle \langle F(x^k, p^*) - y^k, e_i \rangle > 0. \tag{4.12}$$

Let $\eta := (\lambda, \mu, \gamma, \nu)$ and $y := (y^1, y^2, y^3, y^4)$ with appropriate dimensional components corresponding to (f, h, G, H). By the explicit expression of limiting normal cone $\mathcal{N}_\Lambda$ (see, e.g., [27, Proposition 5.1]), we have (1) immediately. We next show (2). It follows from (iii) that $\{x^k\} \subseteq \mathcal{X}$, $\{y^{1,k}\} \subseteq \mathbb{R}^q_-$, $\{y^{2,k}\} = \{0\}^p$, $\{(y^{3,k}, y^{4,k})\} \subseteq C^m$, and

$$\mu_j^* > 0 \Longrightarrow \mu_j^*(g_j(x^k, p^*) - y_j^{1,k}) > 0, \;\; \lambda_i^* > 0 \Longrightarrow \lambda_i^*(h_i(x^k, p^*) - y_i^{2,k}) > 0, \quad (4.13)$$

$$\gamma_l^* \neq 0 \Longrightarrow \gamma_l^*(G_l(x^k, p^*) - y_l^{3,k}) > 0, \;\; \nu_l^* \neq 0, \Longrightarrow \nu_l^*(H_l(x^k, p^*) - y_l^{4,k}) > 0. \quad (4.14)$$

Next, we show that, for each sufficiently large k, $\{y_j^{1,k}, y_i^{2,k}, y_l^{3,k}, y_l^{4,k}\}$ in (4.13)–(4.14) is equal to 0. Assume to the contrary that there exists a subsequence such that it does not hold. We first notice that $y^{2,k} = 0$. Thus, we consider the following three cases:

- If $y_j^{1,k} < 0$ for a subsequence $K_1 \subseteq \{1, 2, \dots\}$, then $\mathcal{N}_{\mathbb{R}_-}(y_j^{1,k}) = \{0\}$ $\forall k \in K_1$. Thus it follows from (4.11) that $\mu_j^k \to \mu_j^* = 0$ as $K_1 \ni k \to \infty$, which contradicts $\mu_j^* > 0$.

- If $y_l^{3,k} > 0$ for a subsequence $K_2 \subseteq \{1, 2, \dots\}$, then $y_l^{4,k} = 0, \forall k \in K_2$. Thus, it follows from the explicit expression of $\mathcal{N}_{\mathcal{X}}$ that $\gamma_l^k \to \gamma_l^* = 0$ as $K_2 \ni k \to \infty$, which contradicts $\gamma_l^* \neq 0$.

- Similarly as above we can show that it is impossible to have $y_l^{4,k} < 0, k \in K_3$ for some subsequence $K_3 \subseteq \{1, 2, \dots\}$.

So far we have shown (2) except the differentiability of $\{h, g, G, H\}$ with respect to x at (x^k, p^*). By the Rademacher's theorem, if a function ψ is Lipschitzian around x^*, then ψ is differentiable almost everywhere around x^*. Based on this fact, if $x^k \to x^*$ and $\psi(x^k) > 0$, then one can always find a sequence $\{\bar{x}^k\}$ with $\psi(\bar{x}^k) > 0$ such that for all k, ψ is differentiable at $\bar{x}^k$ and $\|\bar{x}^k - x^k\| \leq \frac{1}{k}$. Hence we have shown that the sequence $\{\bar{x}^k\}$ satisfies the condition (2). The proof of the theorem is complete by resetting x^k with $\bar{x}^k$ for each k. □

Definition 13. *Given $r \geq 0$ and $x^* \in \mathfrak{F}(p^*)$, we let $\mathscr{M}_M^r(x^*, p^*)$ denote the set of vectors $(\lambda, \mu, \gamma, \nu, \zeta)$ such that*

(i) $0 \in \partial_{(x,p)}\mathscr{L}^r(x^*, p^*; \lambda, \mu, \gamma, \nu) - (0, \zeta) + \mathcal{N}_{\mathcal{X}}(x^*) \times \{0\}$;

(ii) $\mu \geq 0$, $\mu_j = 0$ $(j \notin A^*)$, $\gamma_l = 0$ $(l \in I^*_{+0})$; $\nu_l = 0$ $(l \in I^*_{0+})$, $\gamma_l\nu_l = 0$ or $\gamma_l < 0$, $\nu_l < 0$ $(l \in I^*_{00})$;

(iii) *There exists a sequence $\{(x^k, p^k)\} \subseteq \mathcal{X} \times \mathbb{R}^n$ converging to (x^*, p^*) such that, for each k,*

$$\mu_j > 0 \Longrightarrow \mu_j g_j(x^k, p^k) > 0, \quad \lambda_i \neq 0 \Longrightarrow \lambda_i h_i(x^k, p^k) > 0,$$
$$\gamma_l \neq 0 \Longrightarrow \gamma_l G_l(x^k, p^k) > 0, \quad \nu_l \neq 0 \Longrightarrow \nu H_l(x^k, p^k) > 0,$$

and $\{h, g, G, H\}$ are all differentiable at (x^k, p^k).

In order to study the subdifferential of the value function of (MPEC$_p$), the following constraint qualification for ghp $\mathfrak{F}$ will be useful.

Definition 14. *Let $x^* \in \mathfrak{F}(p^*)$. We say that the MPEC M-quasinormality holds at $x^* \in \mathfrak{F}(p^*)$ if $(\lambda, \mu, \gamma, \nu, 0) \in \mathscr{M}_M^0(x^*, p^*) \Longrightarrow (\lambda, \mu, \gamma, \nu) = \{0\}$.*

Lemma 4.9. *Let $x^* \in \mathcal{O}(p^*)$. Assume that (p^*, x^*) is MPEC M-quasinormal for the constraint region* gph $\mathcal{X}$*. Then the following upper estimate holds:*

$$\hat{\partial}\mathcal{V}(p^*) \subseteq \{\zeta \mid (\lambda, \mu, \gamma, \nu, \zeta) \in \mathscr{M}_M^1(x^*, p^*)\} \tag{4.15}$$

Proof. Let $\zeta \in \hat{\partial}\mathcal{V}(p^*)$. Then, by the definition of Fréchet subdifferential, for an arbitrary $\epsilon > 0$, there exists $\delta_\epsilon > 0$ such that

$$\mathcal{V}(p) - \mathcal{V}(p^*) \geq \zeta^T(p - p^*) - \epsilon\|p - p^*\|, \quad \forall p \in \mathbb{B}(p^*, \delta_\epsilon).$$

By the definition of value function, we have $f(x,p) \geq \mathcal{V}(p)$ for every $x \in \mathfrak{F}(p)$ and hence

$$f(x,p) - \zeta^T(p - p^*) + \epsilon\|p - p^*\| \geq f(x^*, p^*), \quad \forall x \in \mathcal{X}(p), \forall p \in \mathbb{B}(p^*, \delta_\epsilon).$$

Thus, (x^*, p^*) is a locally optimal solution of the optimization problem

$$\begin{aligned} \min \quad & f(x,p) - \zeta^T(p - p^*) + \epsilon\|p - p^*\| \\ \text{s.t.} \quad & g(x,p) \leq 0, \; h(x,p) = 0, \\ & 0 \leq G(x,p) \;\perp\; H(x,p) \geq 0, \\ & (x,p) \in \mathcal{X} \times \mathbb{R}^{\mathrm{n}}. \end{aligned}$$

By Theorem 4.8 and the MPEC M-quasinormality assumption, there exists a vector $(\lambda, \mu, \gamma, \nu)$ such that the following conditions hold:

(i) $0 \in \partial_{(x,p)}\mathscr{L}^1(x^*, p^*, \lambda, \mu, \gamma, \nu) - (0, \zeta) + \mathcal{N}_{\mathcal{X}}(x^*) \times \{0\} + \epsilon \begin{pmatrix} 0 \\ \mathbb{B} \end{pmatrix}$, $\mu \geq 0$, $\mu_j = 0 \; (j \notin A^*)$, $\gamma_l = 0 \; (l \in I^*_{+0})$; $\nu_l = 0 \; (l \in I^*_{0+})$, $\gamma_l \nu_l = 0$ or $\gamma_l < 0, \; \nu_l < 0 \; (l \in I^*_{00})$;

(ii) There exists a sequence $\{(x^k, p^k)\} \subseteq \mathcal{X} \times \mathbb{R}^{\mathrm{n}}$ converging to (x^*, p^*) such that, for each k,

$$\mu_j > 0 \Longrightarrow \mu_j g_j(x^k, p^k) > 0, \quad \lambda_i \neq 0 \Longrightarrow \lambda_i h_i(x^k, p^k) > 0,$$

$$\gamma_l \neq 0 \Longrightarrow \gamma_l G_l(x^k, p^k) > 0, \quad \nu_l \neq 0 \Longrightarrow \nu_l H_l(x^k, p^k) > 0,$$

and $\{h, g, G, H\}$ are all differentiable at (x^k, p^k).

The desired upper estimate follows since ϵ is arbitrary. □

We now give a tighter estimate for the limiting subdifferential of the value function in terms of the enhanced M-multipliers than the one given in [47, 48]. To this end, we first give several lemmas.

The following lemma is similar to Lemma 2.5 in chapter 2 and Proposition 3.6.1 in chapter 3.

Lemma 4.10. *If a vector (p^*, x^*) is MPEC M-quasinormal for the constraint region* $\operatorname{gph}\mathfrak{F}$*, then there exists a neighborhood V of (p^*, x^*) such that all vectors $(p, x) \in$* $\operatorname{gph}\mathfrak{F} \cap V$ *are MPEC M-quasinormal.*

The following lemma can be obtained from the proof of [47, Lemma 3.4].

Lemma 4.11. *Assume that $\varphi : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is Lipschitzian around x^*. If $u^k \to u^*$, $v^k \to v^*$, and $x^k \to x^*$ with $v^k \to \partial(u^k\varphi)(x^k)$, then $v^* \in \partial(u^*\varphi)(x^*)$.*

To finally establish the estimate for subdifferential of value function, we also need the following restricted inf-compactness.

Definition 15. [16, Hypothesis 6.5.1] *We say that the restricted inf-compactness holds around p^* if $\mathcal{V}(p^*)$ is finite and there exist a compact Ω and a positive number ϵ_0 such that, for all $p \in \mathcal{B}_{\epsilon_0}(p^*)$ for which $\mathcal{V}(p) < \mathcal{V}(p^*) + \epsilon_0$, the problem* (MPEC_p) *has a solution in Ω.*

Theorem 4.12. *Assume that the restricted inf-compactness holds for ($MPEC_p$) around p^*. Then the value function $\mathcal{V}(p)$ is lower semicontinuous at p^*. Suppose further that, for each $x^* \in \mathcal{O}(p^*)$, (p^*, x^*) is MPEC M-quasinormal for the constraint region* $\operatorname{gph}\mathfrak{F}$.

Then

$$\partial \mathcal{V}(p^*) \subseteq \bigcup_{x^* \in \mathcal{O}(p^*)} \{\zeta \mid (\lambda, \mu, \gamma, \nu, \zeta) \in \mathscr{M}_M^1(x^*, p^*)\},$$
$$\partial^\infty \mathcal{V}(p^*) \subseteq \bigcup_{x^* \in \mathcal{O}(p^*)} \{\zeta \mid (\lambda, \mu, \gamma, \nu, \zeta) \in \mathscr{M}_M^0(x^*, p^*)\}.$$

Proof. The lower semicontinuity follows from the restricted inf-compactness immediately [15, Page 246]. We complete the proof by considering the following two cases:

(a) Let $\zeta \in \partial \mathcal{V}(p^*)$. By the definition, there exist sequences $p^{\mathbb{l}} \to_{\mathcal{V}} p^*$ and $\zeta^{\mathbb{l}} \to \zeta$ with $\zeta^{\mathbb{l}} \in \hat{\partial} \mathcal{V}(p^{\mathbb{l}})$. Since the restricted inf-compactness holds, $\mathcal{V}(p^*)$ is finite. Since $\mathcal{V}(p^{\mathbb{l}}) \to \mathcal{V}(p^*)$, we have $\mathcal{V}(p^{\mathbb{l}}) < \mathcal{V}(p^*) + \epsilon_0$ for each sufficiently large l. By the restricted inf-compactness again, there exists $x^{\mathbb{l}} \in \mathcal{O}(p^{\mathbb{l}})$ for each sufficiently large l and $\{x^{\mathbb{l}}\}$ is bounded. Without loss of generality, we assume that $x^{\mathbb{l}} \to x^*$. Since

$$\mathcal{V}(p^*) \leftarrow \mathcal{V}(p^{\mathbb{l}}) = f(x^{\mathbb{l}}, p^{\mathbb{l}}) \to f(x^*, p^*), \quad k \to \infty,$$

we have $f(x^*, p^*) = \mathcal{V}(p^*)$ and hence $x^* \in \mathcal{O}(p^*)$. Since the MPEC M-quasinormality holds at (x^*, p^*) and $(x^{\mathbb{l}}, p^{\mathbb{l}}) \to (x^*, p^*)$, by Lemma 4.10, the MPEC M-quasinormality holds at $(x^{\mathbb{l}}, p^{\mathbb{l}})$ for each sufficiently large l. Thus, it follows from Lemma 4.9 that, for each sufficiently large l, there exists a vector $(\lambda^{\mathbb{l}}, \mu^{\mathbb{l}}, u^{\mathbb{l}}, v^{\mathbb{l}})$ such that

(i) $(0, \zeta^{\mathbb{l}}) \in \partial_{(x,p)} \mathscr{L}^1(x^{\mathbb{l}}, p^{\mathbb{l}}; \lambda^{\mathbb{l}}, \mu^{\mathbb{l}}, u^{\mathbb{l}}, v^{\mathbb{l}}) + \mathcal{N}_{\mathcal{X}}(x^{\mathbb{l}}) \times \{0\}$;

(ii) $\mu^{\mathbb{l}} \geq 0$, $\mu_j^{\mathbb{l}} = 0$ $(j \notin A^{\mathbb{l}})$, $\gamma_l^{\mathbb{l}} = 0$ $(l \in I_{+0}^{\mathbb{l}})$; $\nu_l^{\mathbb{l}} = 0$ $(l \in I_{0+}^{\mathbb{l}})$, $\gamma_l^{\mathbb{l}} \nu_l^{\mathbb{l}} = 0$ or $\gamma_l^{\mathbb{l}} < 0$, $\nu_l^{\mathbb{l}} < 0$ $(l \in I_{00}^{\mathbb{l}})$;

(iii) there exists a sequence $\{(x^{\mathbb{l},k}, p^{\mathbb{l},k})\}_k$ converging to $(x^{\mathbb{l}}, p^{\mathbb{l}})$ as $k \to \infty$ such that

$$\mu_j^{\mathbb{l}} > 0 \Longrightarrow \mu_j^{\mathbb{l}} g_j(x^{\mathbb{l},k}, p^{\mathbb{l},k}) > 0, \quad \lambda_i^{\mathbb{l}} \neq 0 \Longrightarrow \lambda_i^{\mathbb{l}} h_i(x^{\mathbb{l},k}, p^{\mathbb{l},k}) > 0,$$
$$\gamma_\imath^{\mathbb{l}} \neq 0 \Longrightarrow u_\imath^l G_\imath(x^{\mathbb{l},k}, p^{\mathbb{l},k}) > 0, \quad \nu_\jmath^l \neq 0 \Longrightarrow v_\jmath^l H_\jmath(x^{\mathbb{l},k}, p^{\mathbb{l},k}) > 0,$$

and $\{h, g, G, H\}$ are all continuously differentiable at $(x^{\mathbb{l},k}, p^{\mathbb{l},k})$,

where $\{A^{\mathbb{l}}, I_{0+}^{\mathbb{l}}, I_{00}^{\mathbb{l}}, I_{+0}^{\mathbb{l}}\}$ is index sets corresponding to $(x^{\mathbb{l}}, p^{\mathbb{l}})$. Since (p^*, x^*) is MPEC M-quasinormal, by using the reduction to absurdity, we can show that the sequence $\{(\lambda^{\mathbb{l}}, \mu^{\mathbb{l}}, \gamma^{\mathbb{l}}, \nu^{\mathbb{l}})\}$ is bounded (see, e.g, Theorem 2.4 of chapter 2 or Proposition 3.6.2 of chapter 3. Thus, without loss of generality, we may assume that $\{(\lambda^{\mathbb{l}}, \mu^{\mathbb{l}}, \gamma^{\mathbb{l}}, \nu^{\mathbb{l}})\}$ converges to $(\lambda, \mu, \gamma, \nu)$. Taking a limit in (i)–(ii) above and noting that (ii) is the more compact form $(\lambda^{\mathbb{l}}, \gamma^{\mathbb{l}}, \nu^{\mathbb{l}}) \in \mathcal{N}_{\mathbb{R}^q_-}(g(x^{\mathbb{l}})) \times \mathcal{N}_{C^m}(G(x^{\mathbb{l}}), H(x^{\mathbb{l}}))$, it follows from Lemma 4.11 and the outer semiconituity of limiting subdifferential and limiting normal cone that

$$(0, \zeta) \in \partial_{(x,p)} \mathscr{L}^1(x^*, p^*; \lambda, \mu, \gamma, \nu) + \mathcal{N}_{\mathcal{X}}(x^*) \times \{0\},$$
$$\mu \geq 0,\ \mu_j = 0\ (j \notin A^*), \gamma_l = 0\ (l \in I_{+0}^*); \nu_l = 0\ (i \in I_{0+}^*),$$
$$\gamma_l \nu_l = 0 \text{ or } \gamma_l < 0,\ \nu_l < 0\ (l \in \mathcal{J}^*);$$

Moreover, by the diagonal rule, we can find a sequence $\{(x^{\mathbb{l},k_{\mathbb{l}}}, p^{\mathbb{l},k_{\mathbb{l}}})\}$ converging to (x^*, p^*) as $l \to \infty$ and, for all l,

$$\mu_j > 0 \Longrightarrow \mu_j g_j(x^{\mathbb{l},k_{\mathbb{l}}}, p^{\mathbb{l},k_{\mathbb{l}}}) > 0, \quad \lambda_i \neq 0 \Longrightarrow \lambda_i h_i(x^{\mathbb{l},k_{\mathbb{l}}}, p^{\mathbb{l},k_{\mathbb{l}}}) > 0,$$
$$\gamma_\imath \neq 0 \Longrightarrow \gamma_\imath G_\imath(x^{\mathbb{l},k_{\mathbb{l}}}, p^{\mathbb{l},k_{\mathbb{l}}}) > 0, \quad \nu_\jmath \neq 0 \Longrightarrow \nu_\jmath H_\jmath(x^{\mathbb{l},k_{\mathbb{l}}}, p^{\mathbb{l},k_{\mathbb{l}}}) > 0,$$

and $\{h, g, G, H\}$ are differentiable at $(x^{\mathbb{l},k_{\mathbb{l}}}, p^{\mathbb{l},k_{\mathbb{l}}})$. Therefore $(\lambda, \mu, \gamma, \nu, \zeta) \in \mathscr{M}_M^1(x^*, p^*)$.

(b) Let $\zeta \in \partial^\infty \mathcal{V}(p^*)$. By the definition, there exist sequences $p^{\mathbb{l}} \to_{\mathcal{V}} p^*$, $\zeta^{\mathbb{l}} \in$

$\hat{\partial}\mathcal{V}(p^l)$, and $t_l \downarrow 0$ such that $t_l\zeta^l \to \zeta$. Similarly as (a), for each l sufficiently large l, there exists a vector $(\lambda^l, \mu^l, \gamma^l, \nu^l)$ such that (i)–(ii) hold. Multiplying (i) by t_l implies

$$(0, t_l\zeta^l) \in \partial_{(x,p)}(t_l\mathscr{L}^1)(x^l, p^l; \lambda^l, \mu^l, \gamma^l, \nu^l) + \mathcal{N}_{\mathcal{X}}(x^l) \times \{0\}. \tag{4.16}$$

Since (p^*, x^*) is MPEC M-quasinormal, by using the reduction to absurdity, we can show that the sequence $\{t_l\lambda^l, t_l\mu^l, t_l\gamma^l, t_l\nu^l\}$ is bounded (see, e.g, Theorem 2.4 of chapter 2 or Proposition 3.6.2 of chapter 3. Without loss of generality, we may assume that $\{t_l\lambda^l, t_l\mu^l, t_l\gamma^l, t_l\nu^l\}$ converges to $\{\lambda, \mu, \gamma, \nu\}$. Taking a limit in (4.16), we have from Lemma 4.11 and the outer semiconituity of limiting subdifferential and limiting normal cone that

$$(0, \zeta) \in \partial_{(x,p)}\mathscr{L}^0(x^*, p^*; \lambda, \mu, \gamma, \nu) + \mathcal{N}_{\mathcal{X}}(x^*) \times \{0\}.$$

The rest of the proof is similar to (a). □

Corollary 4.13. *Assume that the restricted inf-compactness holds for* (MPEC$_p$) *around p^*. Suppose that, for each $x^* \in \mathcal{O}(p^*)$, (p^*, x^*) is MPEC M-quasinormal for the constraint region* gph $\mathfrak{F}$. *If*

$$\bigcup_{x^* \in \mathcal{O}(p^*)} \{\zeta \mid (\lambda, \mu, \gamma, \nu, \zeta) \in \mathscr{M}_M^0(x^*, p^*)\} = \{0\},$$

then the value function $\mathcal{V}$ is Lipschitzian around p^* with

$$\emptyset \neq \partial\mathcal{V}(p^*) \subseteq \bigcup_{x^* \in \mathcal{O}(p^*)} \{\zeta \mid (\lambda, \mu, \gamma, \nu, \zeta) \in \mathscr{M}_M^1(x^*, p^*)\}.$$

In addition to the above assumptions, if

$$\bigcup_{x^*\in\mathcal{O}(p^*)}\{\zeta \mid (\lambda,\mu,\gamma,\nu,\zeta)\in\mathscr{M}_M^1(x^*,p^*)\}=\{\zeta\},$$

then $\mathcal{V}$ is strictly differentiable at p^* and $\nabla\mathcal{V}(p^*)=\zeta$.

We now consider the special case where all the functions $\{f,g,h,G,H\}$ are differentiable. In this case, Definition 13(i) becomes

(i)^1 $0\in\nabla_x\mathscr{L}^r(x^*,p^*;\lambda,\mu,\gamma,\nu)+\mathcal{N}_{\mathcal{X}}(x^*)$; (i)^2 $\zeta=\nabla_p\mathscr{L}^r(x^*,p^*;\lambda,\mu,\gamma,\nu)$.

We define the set of the singular and nonsingular enhanced M-multipliers as the set of vectors (λ,μ,γ,ν)

(i) $0\in\nabla_x\mathscr{L}^r(x^*,p^*;\lambda,\mu,\gamma,\nu)+\mathcal{N}_{\mathcal{X}}(x^*)$;

(ii) $\mu\geq 0,\ \mu_j=0\ (j\notin A^*),\ \gamma_l=0\ (l\in I^*_{+0});\ \nu_l=0\ (l\in I^*_{0+}),\ \gamma_l\nu_l=0$ or $\gamma_l<0,\ \nu_l<0\ (l\in I^*_{00})$;

(iii) There exists a sequence $\{(x^k,p^k)\}\subseteq\mathcal{C}$ converging to x^* such that, for each k,

$$\mu_j>0\Longrightarrow\mu_jg_j(x^k,p^*)>0,\quad \lambda_i\neq 0\Longrightarrow\lambda_ih_i(x^k,p^*)>0,$$
$$\gamma_l\neq 0\Longrightarrow\gamma_lG_l(x^k,p^*)>0,\quad \nu_l\neq 0\Longrightarrow\nu_lH_l(x^k,p^*)>0,$$

and $\{h,g,G,H\}$ are all differentiable at (x^k,p^*).

$$\mathscr{M}_M^i(\zeta,x,p):=\{(\lambda,\mu,\gamma,\nu)\,|\,(\lambda,\mu,\gamma,\nu,\zeta)\in\mathcal{M}_M^r(x,p)\}.$$

When all the functions $\{f,g,h,G,H\}$ are differentiable, under this circumstance, the set $\mathscr{M}_M^i(\zeta,x,p)$ is actually the set of multipliers (λ,μ,γ,ν) which are independent of ζ. Thus, we denote the singular and nonsingular enhanced M-multipliers sets as $\mathscr{M}^i_M(x,p),\ i=0,1$, respectively.

The following example shows that our result Theorem 4.12 is much sharper than its M-counterpart [48, Theorem 4.4].

Example 4. Consider the problem

$$\begin{aligned} \min_x \quad & f(x) \equiv 1 \\ \text{s.t.} \quad & g(x) := x_1 + x_2 + p \leq 0, \\ & 0 \leq G(x) := x_1 + p \perp H(x) := x_2 \geq 0. \end{aligned}$$

It is clear that the value function $\mathcal{V}(p) \equiv 1$ and for $p^* = 0$, the unique feasible solution $(x_1^*, x_2^*) = (0,0)$ is the optimal solution. By solving the following singular and nonsingular M-stationarity systems for the parametric MPEC at $x^* \in \mathcal{X}(p^*)$

$$\mu \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \gamma \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \nu \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$
$$\mu \geq 0,\ \gamma < 0, \nu < 0 \text{ or } \gamma\nu = 0,$$

we find the sets of singular and nonsingular M-multipliers:

$$\mathbb{M}^r(x^*, p^*) := \{\mu(1,-1,-1) \mid \mu \geq 0\}, \quad r = 0, 1.$$

Since the set of singular M-multipliers contains nonzero vector, [48, Theorem 4.4] is not applicable and one cannot even get the Lipschitz continuity of the value function. However, for any sequence $(x^k, p^k) \to (x^*, p^*)$ and any multiplier $\mu(1,-1,-1)$ with

$\mu > 0$, the following system of inequalities does not hold

$$\mu(x_1^k + x_2^k + p^k) > 0,\ -\lambda(x_1^k + p^k) > 0,\ -\mu x_2^k > 0.$$

Thus, the sets of enhanced singular and nonsingular M-multipliers are $\mathcal{M}_M^r(x^*, p^*) = \{(0,0,0)\}$ for $r = 0, 1$, which are contained strictly in $\mathbb{M}^r(x^*, p^*)$ for $r = 0, 1$, respectively. Then the MPEC M-quasinormality holds at $x^* \in \mathfrak{F}(p^*)$. Since

$$\{\mu + \gamma \mid (\mu, \gamma, \nu) \in \mathcal{M}_M^0(x^*, p^*)\} = \{\mu + \gamma \mid (\mu, \gamma, \nu) \in \mathcal{M}_M^1(x^*, p^*)\} = \{0\},$$

by Corollary 4.13, we have that the value function $\mathcal{V}$ is strictly differentiable with $\nabla\mathcal{V}(p^*) = 0$.

4.4.2 Subdifferential via enhanced C-multipliers

In this subsection, we study the subdifferential of the value function in terms of the enhanced C-multipliers. To this end, we first give the nonsmooth enhanced Fritz John type C-stationarity condition for MPECs.

Lemma 4.14. [67, Theorems 7.5 and 7.6] *Let $g(x) := \max\{g_i(x) \mid i = 1, \cdots, m\}$, where $g_i : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$, and $I(\bar{x}) := \{i \mid g_i(\bar{x}) = g(\bar{x})\}$. Let $g_i, i = 1, \cdots, m$, be Lipschitzian around $\bar{x}$. Then g is Lipschitzian around $\bar{x}$ and*

$$\partial g(\bar{x}) \subseteq \bigcup\{\partial(\sum_{i \in I(\bar{x})} \lambda_i g_i)(\bar{x}) \mid \sum_{i \in I(\bar{x})} \lambda_i = 1, \lambda_i \geq 0, i \in I(\bar{x})\}.$$

Let $f(x) := \min\{f_i(x) \mid i = 1, \cdots, m\}$, where $f_i : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$, and $J(\bar{x}) := \{i \mid f_i(\bar{x}) = f(\bar{x})\}$. Assume that f_i is lower semicontinuous near $\bar{x}$ for $i \in J(\bar{x})$ and *lower semicontinuous at* $\bar{x}$ *for* $i \notin J(\bar{x})$*. Then* f *is lower semicontinuous near* $\bar{x}$ *and*

$$\partial f(\bar{x}) \subseteq \bigcup \{\partial f_i(\bar{x}) \mid i \in J(\bar{x})\}.$$

Theorem 4.15. *If* x^* *is a local minimizer of* (MPEC$_{p^*}$)*, then there exist nonzero vectors* $(r, \lambda, \mu, \gamma, \nu)$ *with* $r \geq 0$ *such that*

(i) $0 \in \partial_x \mathscr{L}^r(x^*, p^*; \lambda, \mu, \gamma, \nu) + \mathcal{N}_{\mathcal{X}}(x^*)$;

(ii) $\mu \geq 0$, $\mu_j = 0$ $(j \notin A^*)$, $\gamma_l = 0$ $(l \in I^*_{+0})$, $\nu_l = 0$ $(l \in I^*_{0+})$, $\gamma_l \nu_l \geq 0$ $(l \in I^*_{00})$;

(iii) *there exists a sequence* $\{x^k\} \subseteq \mathcal{X}$ *converging to* x^* *such that, for each* k*,*

$$\begin{aligned}
&\mu_j > 0 \Longrightarrow \mu_j g_j(x^k, p^*) > 0, \\
&\lambda_i \neq 0 \Longrightarrow \lambda_i h_i(x^k, p^*) > 0, \\
&\gamma_l \neq 0 \Longrightarrow \gamma_l \min(G_l(x^k, p^*), H_l(x^k, p^*)) > 0, \\
&\nu_l \neq 0 \Longrightarrow \nu_l \min(G_l(x^k, p^*), H_l(x^k, p^*)) > 0,
\end{aligned}$$

and $\{h(\cdot, p^*), g(\cdot, p^*), \min(G_l, H_l)(\cdot, p^*)\}$ *are all differentiable at* x^k*.*

Proof. Since the feasible region $\mathfrak{F}(p)$ of (MPEC$_p$) can be written as

$$\mathfrak{F}(p) = \{x \in \mathcal{X} \mid g(x,p) \leq 0, h(x,p) = 0, \min(G(x,p), H(x,p)) = 0\}.$$

Thus, by Theorem 2.1 of chapter 2 or Corollary 3.4 of chapter 3, there exist nonzero vectors (r, λ, μ, ξ) with $r \geq 0$ such that the following conditions hold:

(i) $0 \in r\partial_x f(x^*, p^*) + \sum_{j=1}^q \mu_j \partial_x g_j(x^*, p^*) + \sum_{i=1}^p \partial_x(\lambda_i h_i)(x^*, p^*)$
$+ \sum_{l=1}^m \partial_x(\xi_l \min(G_l, H_l))(x^*, p^*) + \mathcal{N}_{\mathcal{X}}(x^*)$, $\mu \geq 0$, $\mu_j = 0$ $(j \notin A^*)$;

(ii) There exists a sequence $\{x^k\} \subseteq \mathcal{X}$ converging to x^* such that, for each k,

$$\begin{aligned}
&\mu_j > 0 \Longrightarrow \mu_j g_j(x^k, p^*) > 0,\\
&\lambda_i \neq 0 \Longrightarrow \lambda_i h_i(x^k, p^*) > 0,\\
&\xi_l \neq 0 \Longrightarrow \xi_l \min(G_l(x^k, p^*), H_l(x^k, p^*)) > 0,
\end{aligned}$$

and $\{h(\cdot, p^*), g(\cdot, p^*), \min(G_l, H_l)(\cdot, p^*)\}$ are all differentiable at x^k.

We investigate $\partial_x(\xi_l \min(G_l, H_l))(x^*, p^*)$ in the following two cases:

(1) $\xi_l \geq 0$: It follow from Lemma 4.14 that

$$\begin{aligned}
&\partial_x(\xi_l \min(G_l, H_l))(x^*, p^*) = \xi_l \partial_x \min(G_l, H_l)(x^*, p^*)\\
&\subseteq \begin{cases}
\xi_l \partial_x G_l(x^*, p^*), & l \in I^*_{0+},\\
\xi_l \partial_x H_l(x^*, p^*), & l \in I^*_{+0},\\
\xi_l \partial_x G_l(x^*, p^*) \cup \xi_l \partial_x H_l(x^*, p^*), & l \in I^*_{00}.
\end{cases}
\end{aligned}$$

(2) $\xi_l < 0$: It follow from Lemma 4.14 that

$$\begin{aligned}
&\partial_x(\xi_l \min(G_l, H_l))(x^*, p^*) = \partial_x \max(\xi_l G_l, \xi_l H_l)(x^*, p^*)\\
&\subseteq \begin{cases}
\partial_x(\xi_l G_l)(x^*, p^*), & l \in I^*_{0+},\\
\partial_x(\xi_l H_l)(x^*, p^*), & l \in I^*_{+0},\\
\{\partial_x(\alpha \xi_l G_l)(x^*, p^*) + \partial_x((1-\alpha)\xi_l H_l)(x^*, p^*) \mid 0 \leq \alpha \leq 1\}, & l \in I^*_{00}.
\end{cases}
\end{aligned}$$

Therefore, by defining the multipliers $\{\gamma, \nu\}$ using $\{\xi, \alpha\}$, the desired result follows from (i)–(ii) and (1)–(2) immediately. □

Remark 2. In the above theorem, all the nonsmooth functions are required to be differentiable at a given sequence in our nonsmooth enhanced optimality conditions in contrast to the required proximal subdifferentiability at a given sequence in Theorem 2.1 of chapter 2. Because all the involved functions are required to be Lipschitzian around the point of interest, by Rademacher's theorem instead of the density theorem in [17, Theorem 3.1], the proximal subdifferentiability in Theorem 2.1 of chapter 2 can be replaced by the differentiability.

Definition 16. *Given* $r \geq 0$, *we let* $\mathscr{M}_C^r(x^*, p^*)$ *denote the set of vectors* $(\lambda, \mu, \gamma, \nu, \zeta)$ *at* $x^* \in \mathfrak{F}(p^*)$ *such that*

(i) $0 \in \partial_{(x,p)}\mathscr{L}^r(x^*, p^*, \lambda, \mu, \gamma, \nu) - (0, \zeta) + \mathcal{N}_{\mathcal{X}}(x^*) \times \{0\}$;

(ii) $\mu \geq 0$, $\mu_j = 0$ $(j \notin A^*)$, $\gamma_l = 0$ $(l \in I^*_{+0})$; $\nu_l = 0$ $(l \in I^*_{0+})$, $\gamma_l \nu_l \geq 0$ $(l \in I^*_{00})$;

(iii) *There exists a sequence* $\{(x^k, p^k)\} \subseteq \mathcal{X} \times \mathbb{R}^{\mathrm{n}}$ *converging to* (x^*, p^*) *such that, for each* k,

$$\mu_j > 0 \Longrightarrow \mu_j g_j(x^k, p^k) > 0,$$
$$\lambda_i \neq 0 \Longrightarrow \lambda_i h_i(x^k, p^k) > 0,$$
$$\gamma_l \neq 0 \Longrightarrow \gamma_l \min(G_l(x^k, p^k), H_l(x^k, p^k)) > 0,$$
$$\nu_l \neq 0 \Longrightarrow \nu_l \min(G_l(x^k, p^k), H_l(x^k, p^k)) > 0,$$

and $\{h, g, \min(G_l, H_l)\}$ *are all differentiable at* (x^k, p^k).

Definition 17. *We say that the MPEC C-quasinormality holds at* (p^*, x^*) *for the region* gph $\mathcal{X}$ *if* $(\lambda, \mu, \gamma, \nu, 0) \in \mathscr{M}_C^0(x^*, p^*) \Longrightarrow (\lambda, \mu, \gamma, \nu) = 0$.

It is not difficult to verify that the MPEC C-quasinormality persists in some feasible neighborhood.

Similarly to the previous subsection, we can easily get the following results.

Theorem 4.16. *Assume that the restricted inf-compactness holds for (MPEC$_p$) around p^*. Then the value function $\mathcal{V}(p)$ is lower semicontinuous at p^*. Suppose further that, for each $x^* \in \mathcal{O}(p^*)$, (p^*, x^*) is MPEC C-quasinormal for the constraint region* gph $\mathfrak{F}$. *Then*

$$\partial\mathcal{V}(p^*) \subseteq \bigcup_{x^* \in \mathcal{O}(p^*)} \{\zeta \mid (\lambda, \mu, \gamma, \nu, \zeta) \in \mathscr{M}_C^1(x^*, p^*)\},$$
$$\partial^\infty\mathcal{V}(p^*) \subseteq \bigcup_{x^* \in \mathcal{O}(p^*)} \{\zeta \mid (\lambda, \gamma, \nu, \zeta) \in \mathscr{M}_C^0(x^*, p^*)\}.$$

Corollary 4.17. *Assume that the restricted inf-compactness holds for* (MPEC$_p$) *around p^*. Suppose that, for each $x^* \in \mathcal{O}(p^*)$, (p^*, x^*) is MPEC C-quasinormal for the constraint region* gph $\mathcal{X}$. *If*

$$\bigcup_{x^* \in \mathcal{O}(p^*)} \{\zeta \mid (\lambda, \mu, \gamma, \nu, \zeta) \in \mathscr{M}_C^0(x^*, p^*)\} = \{0\},$$

then the value function $\mathcal{V}$ is Lipschitzian around p^* with

$$\emptyset \neq \partial\mathcal{V}(p^*) \subseteq \bigcup_{x^* \in \mathcal{O}(p^*)} \{\zeta \mid (\lambda, \mu, \gamma, \nu, \zeta) \in \mathscr{M}_C^1(x^*, p^*)\}.$$

In addition to the above assumptions, if

$$\bigcup_{x^* \in \mathcal{O}(p^*)} \{\zeta \mid (\lambda, \mu, \gamma, \nu, \zeta) \in \mathscr{M}_C^0(x^*, p^*)\} = \{\zeta\},$$

then $\mathcal{V}$ is strictly differentiable at p^* and $\nabla\mathcal{V}(p^*) = \zeta$.

We now consider the special case where all the functions $\{f, g, h, G, H\}$ are differentiable. In this case, Definition 16(i) becomes

(i)^1 $0 \in \nabla_x\mathscr{L}^r(x^*, p^*; \lambda, \mu, \gamma, \nu) + \mathcal{N}_\mathcal{X}(x^*)$; (i)^2 $\zeta = \nabla_p\mathscr{L}^r(x^*, p^*; \lambda, \mu, \gamma, \nu)$.

We define the set of the singular and nonsingular enhanced C-multipliers as the set of vectors $(\lambda, \mu, \gamma, \nu)$ satisfying (i)[1] and Definition 16(ii)–(iii), and denote them by $\mathcal{M}_C^r(x^*, p^*)$, $r = 0, 1$, respectively.

The following example shows that Theorem 4.16 is much sharper than its C-counterpart [48, Theorem 4.8].

Example 5. Consider the following example

$$\begin{aligned} \min_{x} \quad & f(x) := x_1^2 + x_2^2 \\ \text{s.t.} \quad & g(x) := x_1 + x_2 + p \geq 0, \\ & 0 \leq G(x) := x_1 + p \perp H(x) := x_2 \geq 0. \end{aligned}$$

The value function $\mathcal{V}(p) = \begin{cases} 0 & p \geq 0 \\ p^2 & p < 0 \end{cases}$ is a smooth function. For $p^* = 0$, the unique optimal solution $x^* = (0, 0)$. By solving the following singular and nonsingular C-stationarity systems for the parametric MPEC at $x^* \in \mathfrak{F}(p^*)$:

$$-\mu \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \gamma \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \nu \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \ \mu \geq 0, \ \gamma\nu \geq 0,$$

we find the sets of singular and nonsingular C-multipliers:

$$\mathbb{C}^i(x^*, p^*) := \{(1, 1, 1)\lambda \mid \lambda \geq 0\}, \quad i = 0, 1.$$

Since the set of singular C-multipliers contains nonzero vector, [48, Theorem 4.8] is not applicable and one cannot even get the Lipschitz continuity of the value function. However, for any sequence $(x^k, p^k) \to (x^*, p^*)$ and any multiplier $\mu(1, 1, 1)$ with $\mu > 0$, the following system of inequalities does not hold:

$$\mu(x_1^k + x_2^k + p^k) < 0, \ \mu \min(x_1^k + p^k, x_2^k) > 0. \tag{4.17}$$

Thus, the sets of singular and nonsingular enhanced C-multipliers are $\mathcal{M}_C^i(x^*, p^*) = \{0\}$ for $i = 0, 1$, which are contained strictly in $\mathbb{C}^i(x^*, p^*)$ for $i = 0, 1$, respectively. Therefore, the MPEC C-quasinormality holds at $x^* \in \mathfrak{F}(p^*)$. Since

$$\{\mu + \gamma \,|\, (\mu, \gamma, \nu) \in \mathcal{M}_C^0(x^*, p^*)\} = \{\mu + \gamma \,|\, (\mu, \gamma, \nu) \in \mathcal{M}_C^1(x^*, p^*)\} = \{0\},$$

by Corollary 4.17, we get that the value function is strictly differentiable at p^* with $\nabla \mathcal{V}(p^*) = 0$.

4.5 Conclusions

In this chapter, we have shown that the MPEC-LICQ is not a constraint qualification for S-stationary condition if the objective function is not differentiable. Moreover, we have derived the enhanced M-stationary condition and introduced the associated generalized pseudo-normality and quasi-normality conditions for nonsmooth MPECs. We have also introduced a weaker version of the MPEC-CPLD and shown that it implies the MPEC quasi-normality. We have shown the existence of a local error bound under either the MPEC generalized pseudo-normality or quasi-normality under the subdifferential regularity condition. Finally we give upper estimates for the subdifferential of the value function in terms of the enhanced M- and C-multipliers respectively.

Chapter 5

New results on constraint qualifications for nonlinear programming problems

5.1 Introduction

In this chapter we study the following nonlinear programming problem in $\mathbb{R}^n$:

$$\min \quad f(x) \quad \text{s.t.} \quad g(x) \leq 0,\ h(x) = 0, \tag{5.1}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}^q, h : \mathbb{R}^n \to \mathbb{R}^p$ are all continuously differentiable functions. Let $\mathcal{F}$ be the feasible region of problem (5.1). For $x^* \in \mathcal{F}$, we denote by $A(x^*)$ the index set of active inequality constraints at x^*, i.e., $A(x^*) := \{j : g_j(x^*) =$

This chapter is the content of Guo, L., Zhang, J. and Lin, G.H., "New Results on Constraint Qualifications for Nonlinear Extremum Problems and Extensions." J. Optim. Theo. Appl., doi: 10.1007/s10957-013-0510-6.

$0\}$, and by $\Lambda(x^*)$ the set of Lagrangian multipliers associated with x^*, i.e.,

$$\Lambda(x^*) := \{(\lambda, \mu) : \nabla_x L(x^*, \lambda, \mu) = 0, \mu \geq 0, \langle g(x^*), \mu \rangle = 0\},$$

where the Lagrange function $L(x, \lambda, \mu)$ is defined by

$$L(x, \lambda, \mu) := f(x) + \langle g(x), \mu \rangle + \langle h(x), \lambda \rangle.$$

It is well-known that one main role of constraint qualifications is to provide the validity of the Lagrange's principle, that is, to ensure the existence of Lagrangian multipliers associated with a local minimizer; see, e.g., [25, 40, 50]. Recently, some constraint qualifications such as the relaxed constant rank constraint qualification (RCRCQ) [56], the relaxed constant positive linear dependence condition (RCPLD) [1], the constant rank of the subspace component condition (CRSC) [2], the relaxed Mangasarian–Fromovitz constraint qualification (RMFCQ) and the constant rank Mangasarian-Fromovitz constraint qualification (CRMFCQ) [44] are introduced for nonlinear programming problems. These constraint qualifications are weaker than the standard LICQ and MFCQ and most of them have been extended to the MPECs; see chapter 4. The relations among these constraint qualifications are presented in [1,2,28] and the relations between these constraint qualifications and local error bounds are investigated in [2, 13, 55, 56].

In this chapter, we continue to study the above new constraint qualifications. We define the linearized cone of problem (5.1) at $x^* \in \mathcal{F}$ by

$$\mathcal{L}(x^*) := \{d : \langle \nabla g_j(x^*), d \rangle \leq 0 \ (j \in A(x^*)), \ \langle \nabla h_i(x^*), d \rangle = 0 \ (i = 1, \cdots, p)\}.$$

Definition 18. Given $A := \{a^1, \cdots, a^l\}$ and $B := \{b^1, \cdots, b^s\}$, (A, B) is said to be *positively linearly dependent* iff there exist α and β such that $\alpha \geq 0$, $(\alpha, \beta) \neq 0$, and

$$\sum_{i=1}^{l} \alpha_i a^i + \sum_{j=1}^{s} \beta_j b^j = 0.$$

Otherwise, (A, B) is said to be *positively linearly independent.*

Andreani et al. presented the RCPLD condition in [1], which is a relaxation of the CPLD introduced in [71], and they further proposed a weaker constraint qualification, called CRSC, in [2]. More recently, two relaxed constraint qualifications, called CRMFCQ and RMFCQ respectively, were introduced in [44].

Definition 19. Let $x^* \in \mathcal{F}$. (i) We say that the *constant rank Mangasarian-Fromovitz constraint qualification* (CRMFCQ) holds at x^* iff

(a) there exists $\delta > 0$ such that $\{\nabla h_i(x)\}_{i=1}^p$ has the same rank for each $x \in \mathcal{B}_\delta(x^*)$;

(b) there exists d such that

$$\langle \nabla g_j(x^*), d \rangle < 0 \ (j \in A(x^*)), \quad \langle \nabla h_i(x^*), d \rangle = 0 \ (i = 1, \cdots, p).$$

(ii) Let $\mathcal{I} \subseteq \{1, \cdots, p\}$ be such that $\{\nabla h_i(x^*)\}_{i \in \mathcal{I}}$ is a basis for $\operatorname{span}\{\nabla h_i(x^*)\}_{i=1}^p$. We say that the *relaxed constant positive linear dependence* (RCPLD) condition holds at x^* iff there exists $\delta > 0$ such that

(c) $\{\nabla h_i(x)\}_{i=1}^p$ has the same rank for each $x \in \mathcal{B}_\delta(x^*)$;

(d) for each $\mathcal{A} \subseteq A(x^*)$, if $(\{\nabla g_j(x^*) : j \in \mathcal{A}\}, \{\nabla h_i(x^*) : i \in \mathcal{I}\})$ is positively linearly dependent, then $\{\nabla g_j(x), \nabla h_i(x) : j \in \mathcal{A}, i \in \mathcal{I}\}$ is linearly dependent for each $x \in \mathcal{B}_\delta(x^*)$.

(iii) Let $\mathcal{J}_- := \{j \in A(x^*) : -\nabla g_j(x^*) \in \mathcal{L}(x^*)^o\}$. We say that the *constant rank of the subspace component* (CRSC) condition holds at x^* iff there exists $\delta > 0$ such that the family of gradients $\{\nabla g_j(x), \nabla h_i(x) : j \in \mathcal{J}_-, i \in \{1, \cdots, p\}\}$ has the same rank for every $x \in \mathcal{B}_\delta(x^*)$.

(iv) Let $\mathcal{A}_0 := \{j \in A(x^*) : \langle \nabla g_j(x^*), d\rangle = 0$ for all $d \in \mathcal{L}(x^*)\}$. We say that the *relaxed Mangasarian-Fromovitz constraint qualification* (RMFCQ) holds at x^* iff there exists $\delta > 0$ such that the family of gradients $\{\nabla g_j(x), \nabla h_i(x) : j \in \mathcal{A}_0, i \in \{1, \cdots, p\}\}$ has the same rank for every $x \in \mathcal{B}_\delta(x^*)$.

Note that the RMFCQ is actually the generalized Mangasarian-Fromovitz condition introduced in [53], which requires the following redundant condition: There exists d such that

$$\langle \nabla h_i(x^*), d\rangle = 0 \ (i \in \{1, \cdots, p\}), \quad \langle \nabla g_j(x^*), d\rangle = 0 \ (i \in \mathcal{A}_0),$$
$$\langle \nabla g_j(x^*), d\rangle < 0 \ (j \in A(x^*)\backslash\mathcal{A}_0);$$

see [44] for more details. It is not difficult to verify that $\mathcal{J}_- = \mathcal{A}_0$. Therefore, the CRSC coincides with the RMFCQ. It follows from [2, Theorem 4.3] that the RCPLD implies the CRSC (or, equivalently, the RMFCQ).

5.2 New results on constraint qualifications

In this section, we give some new results related to the constraint qualifications listed in the last section.

Theorem 5.1. *The CRMFCQ implies the RCPLD.*

Proof. Assume that the CRMFCQ holds at x^*. Let $\mathcal{I} \subseteq \{1, \cdots, p\}$ be such that $\{\nabla h_i(x^*)\}_{i\in\mathcal{I}}$ is a basis for span $\{\nabla h_i(x^*)\}_{i=1}^p$. Then, $\{\nabla h_i(x^*)\}_{i\in\mathcal{I}}$ is linearly independent. This, together with (b) in Definition 19, implies that the system $\{x :$

$g(x) \leq 0,\ h_i(x) = 0\ (i \in \mathcal{I})\}$ satisfies the MFCQ at x^*. By the Motzkin's transposition theorem, the family of gradients $(\{\nabla g_j(x^*) : j \in A(x^*)\}, \{\nabla h_i(x^*) : i \in \mathcal{I}\})$ is positively linearly independent. Thus, the condition (d) in Definition 19 holds. Since (a) coincides with (c) in Definition 19, the RCPLD holds at x^*. The proof is complete. □

Thus, we have the following relations:

$$\text{CRMFCQ} \Longrightarrow \text{RCPLD} \Longrightarrow \text{CRSC/RMFCQ}.$$

The following example from [1] shows that the converse of Theorem 5.1 is not true.

Example 6. Consider the constraint system

$$\{x \in \mathbb{R}^2 : h(x) = 0,\ g_1(x) \leq 0,\ g_2(x) \leq 0\},$$

where $h(x) := -(x_1+1)^2 - x_2^2 + 1$, $g_1(x) := x_1^2 + (x_2+1)^2 - 1$, and $g_2(x) := -x_2$. Pick a feasible point $x^* = (0,0)$. It is not hard to verify that the RCPLD holds at x^*. On the other hand, it is easy to verify that the following system in d has no solution:

$$\langle \nabla g_1(x^*), d\rangle < 0, \quad \langle \nabla g_2(x^*), d\rangle < 0.$$

This means that the CRMFCQ fails at x^*.

Guo and Lin [26] gave an upper estimate of the Fréchet normal cone in the setting of MPEC under the so-called MPEC-RCPLD condition. Since, for nonlinear problems, the MPEC-RCPLD reduces to the RCPLD, we have the following result immediately.

Theorem 5.2. *If the RCPLD holds at $x \in \mathcal{F}$, then*

$$\mathcal{N}_{\mathcal{F}}(x) \subseteq \nabla g(x)\mathcal{N}_{[-\infty,0]^q}(g(x)) + \nabla h(x)\mathcal{N}_{\{0\}^p}(h(x)).$$

Proof. If the RCPLD holds at x, by [26, Theorem 4.1], we have

$$\widehat{\mathcal{N}}_{\mathcal{F}}(x) \subseteq \nabla g(x)\mathcal{N}_{[-\infty,0]^q}(g(x)) + \nabla h(x)\mathcal{N}_{\{0\}^p}(h(x)).$$

Since the RCPLD persists in some feasible neighborhood of x (see [1, Theorem 4]), we have the desired conclusion. □

Kruger et al. [44, Theorem 2] showed that the CRMFCQ is robust, which means that, once it is satisfied at x^*, for any objective function f and any local minimizer x of problem (5.1) in a feasible neighborhood of x^*, x must be a stationary point of (5.1). In fact, we can show a stronger result, that is, it persists in some feasible neighborhood of x^* or, in other words, it is well-posed in the sense of [44, 54].

Theorem 5.3. *If the CRMFCQ holds at $x^* \in \mathcal{F}$, then there exists $\delta' > 0$ such that the CRMFCQ holds at each $x \in \mathcal{F} \cap \mathcal{B}_{\delta'}(x^*)$.*

Proof. Let δ and d be given as in Definition 19 (i). Let $\mathcal{I} \subseteq \{1, \cdots, p\}$ be such that $\{\nabla h_i(x^*)\}_{i\in\mathcal{I}}$ is a basis for span $\{\nabla h_i(x^*)\}_{i=1}^p$. Then $\{\nabla h_i(x^*)\}_{i\in\mathcal{I}}$ is linearly independent, which together with (b) in Definition 19, implies that the system $\mathcal{F}_1 := \{x : g(x) \le 0,\ h_i(x) = 0\ (i \in \mathcal{I})\}$ satisfies the MFCQ at x^*. Since the MFCQ persists in some feasible neighborhood of x^*, there exists $\delta_1 \in (0, \delta)$ such that, for each $x \in \mathcal{B}_{\delta_1}(x^*) \cap \mathcal{F}_1$,

- $\{\nabla h_i(x)\}_{i\in\mathcal{I}}$ is linearly independent;

– there exists d such that

$$\langle \nabla h_i(x), d\rangle = 0 \ (i \in \mathcal{I}), \quad \langle \nabla g_j(x), d\rangle < 0 \ (j \in A(x)). \tag{5.2}$$

On the other hand, since $\{\nabla h_i(x)\}_{i=1}^p$ has the same rank for each $x \in \mathcal{B}_\delta(x^*)$, we can choose $\delta_2 \in (0, \delta_1)$ such that $\{\nabla h_i(x)\}_{i\in\mathcal{I}}$ is a basis for span $\{\nabla h_i(x)\}_{i=1}^p$ for each $x \in \mathcal{B}_{\delta_2}(x^*)$. Thus, the vectors $\{\nabla h_i(x) : i \in \{1, \cdots, p\}\backslash\mathcal{I}\}$ can be represented by $\{\nabla h_i(x)\}_{i\in\mathcal{I}}$ and hence, by (5.2),

$$\langle \nabla h_i(x), d\rangle = 0 \ (i = 1, \cdots, p). \tag{5.3}$$

From (a) in Definition 19, $\mathcal{F} \subseteq \mathcal{F}_1$, and (5.2)–(5.3), we have the desired result by letting $\delta' = \frac{\delta_2}{2}$. This completes the proof. □

5.3 Local error bounds

In this section, we discuss the existence of local error bounds for the constraint system $\mathcal{F}$.

Definition 20. The system $\mathcal{F}$ is said to admit a *local error bound* at $x^* \in \mathcal{F}$ iff there exist $\delta > 0$ and $\kappa > 0$ such that

$$\mathrm{dist}_{\mathcal{F}}(x) \leq \kappa(\|h(x)\| + \|g^+(x))\|), \quad \forall x \in \mathcal{B}_\delta(x^*).$$

Since the above inequality automatically holds if $x \in \mathcal{F}$, we only need to focus on the infeasible points near the reference point. There are many approaches such as the primal space error bound criteria in terms of slopes and the dual criteria in terms of dual subdifferentials (see, e.g., [22, 79]) can be used to see whether a local error bound exists. Another way is to verify the boundedness of KKT multiplier sequence associated with the closest feasible solution sequence from any infeasible sequence to the feasible region [55, 56]. In this section, we employ the latter way due to our observation that the proof techniques used in the literature can be somehow improved.

Let $y \in \mathbb{R}^n \backslash \mathcal{F}$. Denote by $\prod_{\mathcal{F}}(y)$ the projection set of y onto $\mathcal{F}$ under the Euclidean norm, that is, $\prod_{\mathcal{F}}(y)$ is the solution set of the optimization problem

$$\min_{x} \quad f_y(x) := \|x - y\| \quad \text{s.t.} \quad x \in \mathcal{F}.$$

Note that, if $y \notin \mathcal{F}$, f_y is continuously differentiable with respect to x. For any given $y \notin \mathcal{F}$, we denote by $\Lambda_y(x)$ the set of multipliers of the above problem at x.

We next give a characterization of the existence of local error bounds for $\mathcal{F}$.

Theorem 5.4. *The following assertions are equivalent:*

(a) *The system $\mathcal{F}$ admits a local error bound at $x^* \in \mathcal{F}$.*

(b) *For each sequence $\{y^k\}$ converging to $x^* \in \mathcal{F}$ with $y^k \notin \mathcal{F}$, there exists a number $M > 0$ such that, for all k, the condition $\Lambda_M^k(x^k) = \{(\lambda, \mu) \in \Lambda_{y^k}(x^k) : \|(\lambda, \mu)\| \leq M\} \neq \emptyset$ with some point x^k in $\prod_{\mathcal{F}}(y^k)$.*

Proof. The implication (a)⇒(b) follows from [55, Theorem 2] immediately. We next show (b)⇒(a).

Let $y^k \to x^*$ with $y^k \notin \mathcal{F}$ and $x^k \in \prod_{\mathcal{F}}(y^k)$. It is obvious that $x^k \to x^*$ as $k \to \infty$. We then have $y^k - x^k \to 0$ as $k \to \infty$. If $x^* \in \text{int}\mathcal{F}$, the assertion is obviously true. Consider the case where $x^* \in \partial\mathcal{F}$. By the assumption, there exists a multiplier sequence $\{(\lambda^k, \mu^k)\}$ with $\lambda^k \geq 0$ and $\|(\lambda^k, \mu^k)\| \leq M$ for all k such that, for each k

and j,

$$\mu_j^k g_j(x^k) = 0, \qquad \frac{x^k - y^k}{\|x^k - y^k\|} + \sum_{j=1}^{q} \mu_j^k \nabla g_j(x^k) + \sum_{i=1}^{p} \lambda_i^k \nabla h_i(x^k) = 0. \tag{5.4}$$

We obtain from the above discussion that, for each k sufficiently large,

$$\begin{aligned}
\|x^k - y^k\| &= \Big\langle \sum_{j=1}^{q} \mu_j^k \nabla g_j(x^k) + \sum_{i=1}^{p} \lambda_i^k \nabla h_i(x^k), y^k - x^k \Big\rangle \\
&\le \sum_{j=1}^{q} \mu_j^k \Big(g_j(y^k) - g_j(x^k) + o(\|v^k - x^k\|) \Big) \\
&\qquad + \sum_{i=1}^{p} \lambda_i^k \Big(h_i(y^k) + o(\|y^k - x^k\|) \Big) \\
&= \sum_{j=1}^{q} \mu_j^k g_j(y^k) + \sum_{i=1}^{p} \lambda_i^k h_i(y^k) + \Big(\sum_{j=1}^{q} \mu_j^k + \sum_{i=1}^{p} \lambda_i^k \Big) o(\|y^k - x^k\|) \\
&\le \sum_{j=1}^{q} \mu_j^k g_j(y^k) + \sum_{i=1}^{p} \lambda_i^k h_i(y^k) + \frac{1}{2}\|y^k - x^k\|,
\end{aligned}$$

where the third equality follows from $\mu_j^k g_j(x^k) = 0$ for each k and i, and the last inequality follows from the fact that $\frac{o(\|y^k - x^k\|)}{\|y^k - x^k\|} \to 0$ as $k \to \infty$ and the boundedness of $\{(\lambda^k, \mu^k)\}$. This means

$$\mathrm{dist}_{\mathcal{F}}(y^k) = \|y^k - x^k\| \le 2M \Big(\sum_{j=1}^{q} g_j^+(y^k) + \sum_{i=1}^{p} |h_i(y^k)| \Big).$$

Thus, for any sequence $\{y^k\}$ converging to x^* with $y^k \notin \mathcal{F}$, there exists a number $M > 0$ such that, for each k sufficiently large,

$$\mathrm{dist}_{\mathcal{F}}(y^k) \le 2M(\|h(y^k)\|_1 + \|g^+(y^k)\|_1). \tag{5.5}$$

Suppose to the contrary that the local error bound condition does not hold, i.e., there exists $y^k \to x^*$ with $y^k \notin \mathcal{F}$ such that, for each k sufficiently large,

$$\text{dist}_{\mathcal{F}}(y^k) > k(\|h(y^k)\| + \|g^+(y^k)\|).$$

This contradicts (5.5) and hence the local error bound condition holds at $x^* \in \mathcal{F}$. $\square$

In the above theorem, to guarantee the existence of local error bounds, we only require the constraint functions to be once continuously differentiable. Therefore, Theorem 5.4 improves [56, Theorem 2] and [13, Lemma 4.2], where the constraint functions are assumed to be once continuously differentiable and have locally Lipschitzian derivatives. By virtue of Theorem 5.4, we can show that the RCPLD implies the local error bound condition if the constraint functions are continuously differentiable, which improves [1, Theorem 7] and [13, Theorem 4.1] because the constraint functions only need to be once continuously differentiable here.

Theorem 5.5. *If the RCPLD holds at $x^* \in \mathcal{F}$, then the local error bound condition holds at x^*, i.e., there exist $\delta > 0$ and $\kappa > 0$ such that*

$$\text{dist}_{\mathcal{F}}(x) \leq \kappa(\|h(x)\| + \|g^+(x)\|), \quad \forall x \in \mathcal{B}_\delta(x^*).$$

Proof. To make use of Theorem 5.4, it is sufficient to prove that condition (b) of Theorem 5.4 holds if the RCPLD holds. We only need to consider the case where $x^* \in \partial\mathcal{F}$. Let $x^k \to x^*$ with $x^k \notin \mathcal{F}$ and $\bar{x}^k \in \prod_{\mathcal{F}}(x^k)$. Obviously, we have $x^k - \bar{x}^k \in \widehat{\mathcal{N}}_{\mathcal{F}}(\bar{x}^k)$ and hence

$$\frac{x^k - \bar{x}^k}{\|x^k - \bar{x}^k\|} \in \widehat{\mathcal{N}}_{\mathcal{F}}(\bar{x}^k). \tag{5.6}$$

Since $\bar{x}^k \to x^*$ as $k \to \infty$ and the RCPLD persists in a feasible neighborhood of x^* (see [1, Theorem 4]), the RCPLD holds at $\bar{x}^k$ when k is sufficiently large and then, by (5.6) and Theorem 5.2, there exist $\lambda^k \geq 0$ and μ^k such that

$$\frac{x^k - \bar{x}^k}{\|x^k - \bar{x}^k\|} = \sum_{j \in A(\bar{x}^k)} \mu_j^k \nabla g_j(\bar{x}^k) + \sum_{i=1}^{p} \lambda_i^k \nabla h_i(\bar{x}^k). \tag{5.7}$$

Let $\mathcal{I} \subseteq \{1, \cdots, p\}$ be such that $\{\nabla h_i(x^*)\}_{i \in \mathcal{I}}$ is a basis for span $\{\nabla h_i(x^*)\}_{i=1}^p$. Then $\{\nabla h_i(x^*)\}_{i \in \mathcal{I}}$ is linearly independent. By the RCPLD assumption, $\{\nabla h_i(x)\}_{i=1}^p$ has the same rank for each $x \in \mathcal{B}_\delta(x^*)$. Thus, it is not hard to see that $\{\nabla h_i(\bar{x}^k)\}_{i \in \mathcal{I}}$ is a basis for span $\{\nabla h_i(\bar{x}^k)\}_{i=1}^p$ for each k sufficiently large. Thus, it follows from (5.7) and Lemma 4.3 that there exist $\mathcal{A}^k \subseteq A(\bar{x}^k)$ and $\bar{\mu}^k \geq 0, \bar{\lambda}^k$ such that

$$\frac{x^k - \bar{x}^k}{\|x^k - \bar{x}^k\|} = \sum_{j \in \mathcal{A}^k} \bar{\mu}_j^k \nabla g_j(\bar{x}^k) + \sum_{i \in \mathcal{I}} \bar{\lambda}_i^k \nabla h_i(\bar{x}^k),$$

and the vectors

$$\{\nabla h_i(\bar{x}^k),\ \nabla g_j(\bar{x}^k) : i \in \mathcal{I},\ j \in \mathcal{A}^k\} \tag{5.8}$$

are linearly independent. Without any loss of generality, we assume that $\mathcal{A}^k \equiv \mathcal{A}$ (otherwise, we can choose a subsequence). Thus, we have

$$\frac{x^k - \bar{x}^k}{\|x^k - \bar{x}^k\|} = \sum_{j \in \mathcal{A}} \bar{\mu}_j^k \nabla g_j(\bar{x}^k) + \sum_{i \in \mathcal{I}} \bar{\lambda}_i^k \nabla h_i(\bar{x}^k). \tag{5.9}$$

Let $\bar{\mu}_j^k = 0$ when $j \in \{1, \cdots, q\} \backslash \mathcal{A}$ and $\bar{\lambda}_i^k = 0$ when $i \in \{1, \cdots, p\} \backslash \mathcal{I}$. We next show that $\{\|(\bar{\lambda}^k, \bar{\mu}^k)\|\}$ is bounded. Without any loss of generality, we assume to the contrary that $\|(\bar{\lambda}^k, \bar{\mu}^k)\| \to \infty$ as $k \to \infty$ and $\frac{(\bar{\lambda}^k, \bar{\mu}^k)}{\|(\bar{\lambda}^k, \bar{\mu}^k)\|} \to (\lambda^*, \mu^*) \neq 0$. Dividing (5.9) by $\|(\bar{\lambda}^k, \bar{\mu}^k)\|$ and taking a limit, we have

$$\sum_{j \in \mathcal{A}} \mu_j^* \nabla g_j(x^*) + \sum_{i \in \mathcal{I}} \mu_i^k \nabla h_i(x^*) = 0.$$

This and (5.8) contradict the RCPLD assumption at x^*. Thus, $\{\|(\bar{\lambda}^k, \bar{\mu}^k)\|\}$ is bounded, i.e., there exists a number $M > 0$ such that the condition $\Lambda_M^k(x) = \{(\bar{\lambda}, \bar{\mu}) \in \Lambda_{v^k}(x) : \|(\bar{\lambda}, \bar{\mu})\| \leq M\} \neq \emptyset$ holds at $\bar{x}^k = \prod_{\mathcal{F}}(x^k)$. It follows from Theorem 5.4 that the system $\mathcal{F}$ admits a local error bound at x^*. □

By the fact that the RCPLD implies the existence of local error bounds, Andreani et al. [2] showed that, under the twice continuous differentiability of constraint functions, the CRSC (or, equivalently, the RMFCQ) implies the existence of local error bounds. Kruger et al. [44, Theorem 4] showed the same result under the locally Lipschitz continuity of derivatives of constraint functions.

Corollary 5.6. *If the CRSC (or, equivalently, the RMFCQ) holds at x^*, then there exist $\delta > 0$ and $\kappa > 0$ such that*

$$\text{dist}_{\mathcal{F}}(x) \leq \kappa(\|h(x)\| + \|g^+(x)\|), \quad \forall x \in \mathcal{B}_\delta(x^*).$$

Proof. First, as stated in [2, Lemma 5.3], the constraints in $\mathcal{J}_- := \{j \in A(x^*) : -\nabla g_j(x^*) \in \mathcal{L}(x^*)^o\}$ are actually equality constraints in a neighborhood of x^*. It is natural to consider the feasible set $\mathcal{F}^E$:

$$\mathcal{F}^E = \{x : h_i(x) = 0 \ \forall i \in \mathcal{I}, \quad g_j(x) = 0 \ \forall j \in \mathcal{J}_-, \quad g_j(x) < 0 \ \forall j \in A(x) \setminus \mathcal{J}_-\},$$

which is equivalent to the original feasible set $\mathcal{F}$ close to x^*. It is trivial that the CRSC point (with respect to $\mathcal{F}$) x^* verifies RCPLD as a feasible point of the set $\mathcal{F}^E$. Now, using Theorem 5.5, the system $\mathcal{F}^E$ admits a local error bound at x^*, which requires only once continuous differentiability of the involved constraint functions. Then following the proof idea of [2, Theorem 5.5], we obtain the desired result. □

Since the local error bound condition is equivalent to the calmness of the associated perturbed constraint mapping [39], we have the following result from [39, Proposition 1] and Corollary 5.6 immediately.

Corollary 5.7. *If the CRSC (or, equivalently, the RMFCQ) holds at* x^**, then the Abadie constraint qualification holds at* x^**, i.e.,* $\mathcal{T}_{\mathcal{F}}(x^*) = \mathcal{L}(x^*)$*.*

Moldovan and Pellegrini [58] introduced a regularity condition for nonlinear problems to ensure the validity of Lagrange's principle. From Corollary 5.7 and [58, Theorem 4.1], we can obtain the following result immediately.

Corollary 5.8. *If the CRSC (or, equivalently, the RMFCQ) holds at* x^**, then the regularity condition given in* [58] *holds at* x^**.*

5.4 Extensions to MPEC

In this section, we extend the results given in the previous sections to the MPEC problem defined in chapter 4.

As discussed in section 4.2 of chapter 4 that, since the feasible region of the MPEC problem contains some complementarity constraints, most constraint qualifications including the RCPLD are not satisfied with any degree of freedom. This means that Theorem 5.5 cannot be applied to MPEC directly. In order to derive similar results for MPEC, some new techniques based on the special structure of MPEC are needed.

In [13], Chieu and Lee gave lots of results under a new MPEC constraint qualification called MPEC-rCPLD and the condition that the constraint functions are continuously differentiable and their derivatives are locally Lipshitizian. In fact, we can improve Theorem 4.2 and Corollaries 4.1–4.2 given in [13] to the case where the constraint functions are only once continuously differentiable by making use of Theorem 5.4 instead of [13, Lemma 4.2].

Next we discuss the local error bound condition for the constraint region of the MPEC. Let $\mathfrak{F}$ denote the feasible region. In [26], the MPEC-RCPLD introduced in [28] was shown to be a constraint qualification for M-stationarity. In what follows, we next show that the MPEC-RCPLD implies the existence of local error bounds if the strict complementarity condition holds. We also give a sufficient condition to admit a local error bound in the setting of MPEC. Since the strict complementarity is a strong assumption, for a slightly stronger constraint qualification called MPEC-ERCPLD, we show that it can imply the existence of local error bounds.

Given a point x and three index sets $\mathcal{I}_1 \subseteq \{1, \cdots, p\}$, $\mathcal{I}_2, \mathcal{I}_3 \subseteq \{1, \cdots, m\}$, we denote

$$\mathcal{G}\,(x; \mathcal{I}_1, \mathcal{I}_2, \mathcal{I}_3) := \{\nabla h_i(x), \nabla G_\imath(x), \nabla H_\jmath(x) : j \in \mathcal{I}_1, \imath \in \mathcal{I}_2, \jmath \in \mathcal{I}_3\}.$$

Definition 21. Let $x^* \in \mathfrak{F}$ and $\mathcal{I}_1 \subseteq \{1, \cdots, p\}, \mathcal{I}_2 \subseteq I_{0+}, \mathcal{I}_3 \subseteq I_{+0}$ be the index sets such that $\mathcal{G}(x^*; \mathcal{I}_1, \mathcal{I}_2, \mathcal{I}_3)$ is a basis for span $\mathcal{G}(x^*; \{1, \cdots, p\}, I_{0+}, I_{+0})$. We say that the *MPEC relaxed constant positive linear dependence condition* (MPEC-RCPLD) holds at x^* iff there exists $\delta > 0$ such that

- $\mathcal{G}(x; \{1, \cdots, p\}, I_{0+}, I_{+0})$ has the same rank for each $x \in \mathcal{B}_\delta(x^*)$;
- for each $\mathcal{I}_4 \subseteq A(x^*)$ and $\mathcal{I}_5, \mathcal{I}_6 \subseteq I_{00}$, if there exist multipliers $\{\lambda, \mu, \gamma, \nu\}$ with $\mu_j \geq 0$ for each $j \in \mathcal{I}_4$ and either $\gamma_l \nu_l = 0$ or $\gamma_l > 0, \nu_l > 0$ for each $l \in I_{00}$, which are not all zero, such that

$$\sum_{j \in \mathcal{I}_4} \mu_j \nabla g_j(x^*) + \sum_{i \in \mathcal{I}_1} \lambda_i \nabla h_i(x^*) - \sum_{\imath \in \mathcal{I}_2 \cup \mathcal{I}_5} u_\imath \nabla G_\imath(x^*) - \sum_{\jmath \in \mathcal{I}_3 \cup \mathcal{I}_6} v_\jmath \nabla H_\jmath(x^*) = 0,$$

then, for any $x \in \mathcal{B}_\delta(x^*)$, the vectors

$$\{\nabla g_j(x) : j \in \mathcal{I}_4\}, \quad \{\nabla h_i(x) : \ i \in \mathcal{I}_1\},$$
$$\{\nabla G_\imath(x) : \ \imath \in \mathcal{I}_2 \cup \mathcal{I}_5\}, \quad \{\nabla H_\jmath(x) : \ \jmath \in \mathcal{I}_3 \cup \mathcal{I}_6\}$$

are linearly dependent.

In order to describe the local error bound condition for MPEC in a compact form, we rewrite the feasible region of the MPEC as $\mathfrak{F} = \{x : F(x) \in \Lambda\}$, where

$$F(x) := (g(x), h(x), \Psi(x))^T, \quad \Psi(x) := (G_1(x), H_1(x), \cdots, G_m(x), H_m(x)),$$

and

$$\Lambda :=]-\infty, 0]^p \times \{0\}^q \times C^m, \quad C := \{\, (a,b) \in \mathbb{R}^2 : 0 \le a \perp b \ge 0\}.$$

For any $(a,b) \in C$, it is easy to verify that (see, e.g., [42])

$$\mathcal{N}_C(a,b) \;=\; \left\{ (d_1, d_2) : \begin{array}{r} d_1 \in \mathbb{R}, d_2 = 0 \text{ if } a = 0 < b \\ d_1 = 0, d_2 \in \mathbb{R} \text{ if } a > 0 = b \\ \text{either } d_1 < 0, d_2 < 0 \text{ or } d_1 d_2 = 0 \text{ if } a = b = 0 \end{array} \right\}.$$

Theorem 5.9. *Suppose that the MPEC-RCPLD and the strict complementarity condition hold at $x^* \in \mathfrak{F}$. Then x^* satisfies a local error bound, i.e., there exist $\delta > 0$ and $c > 0$ such that*

$$\mathrm{dist}_{\mathfrak{F}}(x) \le c\, \mathrm{dist}_\Lambda(F(x)), \quad \forall x \in \mathcal{B}_\delta(x^*).$$

Proof. Suppose to the contrary that there exists $x^k \to x^*$ such that

$$\mathrm{dist}_{\mathfrak{F}}(x^k) > k\,\mathrm{dist}_{\Lambda}(F(x^k)). \tag{5.10}$$

It is obvious that $x^k \notin \mathfrak{F}$. Since $\mathfrak{F}$ is closed, we may choose $\bar{x}^k \in \prod_{\mathfrak{F}}(x^k)$, that is, $\bar{x}^k$ is an optimal solution of the problem

$$\min \quad f^k(x) := \|x^k - x\| \quad \text{s.t.} \quad g(x) \leq 0,\ h(x) = 0,\ 0 \leq G(x) \ \perp \ H(x) \geq 0. \tag{5.11}$$

Since $x^k \notin \mathfrak{F}$, f^k is continuously differentiable with respect to x. Since the MPEC-RCPLD persists in some feasible neighborhood of x^* (see [14, Theorem 4.3]) and $\bar{x}^k \to x^*$, the MPEC-RCPLD holds at $\bar{x}^k$ for each k sufficiently large. Due to the fact that the MPEC-RCPLD is a constraint qualification for M-stationarity [26, Corollary 4.1], $\bar{x}^k$ is an M-stationary point of problem (5.11) when k is sufficiently large. Since the MPEC-RCPLD holds at x^*, it follows from the proof of [28, Theorem 4.1] that the multiplier stability (see [27, Definition 3.3]) for problem (5.11) holds at x^*, that is, there exists a bounded multiplier sequence

$$\{(\mu^k, \lambda^k, -\gamma^k, -\nu^k) \in \mathcal{N}_{[-\infty,0]^q}(g(\bar{x}^k)) \times \mathcal{N}_{\{0\}^p}(h(\bar{x}^k)) \times \mathcal{N}_{C^m}(G(\bar{x}^k), H(\bar{x}^k))\}$$

such that, for each k,

$$\frac{x^k - \bar{x}^k}{\|\bar{x}^k - x^k\|} = \sum_{j=1}^{q} \mu_j^k \nabla g_j(\bar{x}^k) + \sum_{i=1}^{p} \lambda_i^k \nabla h_i(\bar{x}^k) - \sum_{\imath=1}^{m} \gamma_\imath^k \nabla G_\imath(\bar{x}^k) - \sum_{\jmath=1}^{m} \nu_\jmath^k \nabla H_\jmath(\bar{x}^k). \tag{5.12}$$

It is easy to see that, for each i, j, and $\imath$,

$$\mu_j^k g_j(\bar{x}^k) = 0, \quad \lambda_i^k h_i(\bar{x}^k) = 0, \quad \gamma_\imath^k G_\imath^k(\bar{x}^k) = 0, \quad \nu_\imath^k H_\imath^k(\bar{x}^k) = 0.$$

Let $\|(\lambda^k, \mu^k, \gamma^k, \nu^k)\| \leq M$ for each k. It follows from (5.12) that, for each k sufficiently large,

$$\begin{aligned}
\|x^k - \bar{x}^k\| &= \sum_{j=1}^{q} \mu_j^k \langle \nabla g_j(\bar{x}^k), x^k - \bar{x}^k \rangle + \sum_{i=1}^{p} \lambda_i^k \langle \nabla h_i(\bar{x}^k), x^k - \bar{x}^k \rangle \\
&\quad - \sum_{\imath=1}^{m} \gamma_\imath^k \langle \nabla G_\imath(\bar{x}^k), x^k - \bar{x}^k \rangle - \sum_{\imath=1}^{m} \nu_\imath^k \langle \nabla H_\imath(\bar{x}^k), x^k - \bar{x}^k \rangle \\
&= \sum_{j=1}^{q} \mu_j^k (g_j(x^k) - g_i(\bar{x}^k)) + \sum_{i=1}^{p} \lambda_i^k (h_i(x^k) - h_i(\bar{x}^k))) \\
&\quad - \sum_{\imath=1}^{m} \gamma_\imath^k (G_\imath(x^k) - G_\imath(\bar{x}^k)) - \sum_{\imath=1}^{m} \nu_\imath^k (H_\imath(x^k) - H_\imath(\bar{x}^k)) + o(\|x^k - \bar{x}^k\|) \\
&= \sum_{j=1}^{q} \mu_j^k g_j(x^k) + \sum_{i=1}^{p} \lambda_i^k h_i(x^k) - \sum_{\imath=1}^{m} \gamma_\imath^k G_\imath(x^k) - \sum_{\imath=1}^{m} \nu_\imath^k H_\jmath(x^k) \\
&\qquad + o(\|x^k - \bar{x}^k\|), \\
&\leq \sum_{j=1}^{q} \mu_j^k g_j(x^k) + \sum_{i=1}^{p} \lambda_i^k h_i(x^k) - \sum_{\imath=1}^{m} \gamma_\imath^k G_\imath(x^k) - \sum_{\imath=1}^{m} \nu_\imath^k H_\imath(x^k) \\
&\qquad + \frac{1}{2}\|x^k - \bar{x}^k\|. \qquad (5.13)
\end{aligned}$$

Clearly, for each i and j, we have

$$\mu_j^k g_j(x^k) \leq \mu_j^k g_j^+(x^k), \quad \lambda_i^k h_i(x^k) \leq |\lambda_i^k||h_i(x^k)|. \qquad (5.14)$$

We next show that

$$-\gamma_\imath^k G_\imath(x^k) - \nu_\imath^k H_\imath(x^k) \leq 2M\phi(G_\imath(x^k), H_\imath(x^k)), \qquad (5.15)$$

where

$$\phi(G_\imath(x^k), H_\imath(x^k)) := \max\left(-G_\imath(x^k), -H_\imath(x^k), -G_\imath(x^k) - H_\imath(x^k), \min(G_\imath(x^k), H_\imath(x^k))\right).$$

Note that $(-\gamma_\imath^k, -\nu_\imath^k) \in \mathcal{N}_C(G_\imath(\bar{x}^k), H_\imath(\bar{x}^k))$ implies

$$\gamma_\imath^k \nu_\imath^k = 0 \quad \text{or} \quad \gamma_\imath^k > 0, \nu_\imath^k > 0. \tag{5.16}$$

Consider the following four cases:

- If $\gamma_\imath^k G_\imath(x^k) \geq 0$ and $\nu_\imath^k H_\imath(x^k) \geq 0$, the above inequality (5.15) is trivial.
- If $\gamma_\imath^k G_\imath(x^k) < 0$ and $\nu_\imath^k H_\imath(x^k) > 0$, we can claim that $\gamma_\imath^k > 0$. In fact, otherwise, if $\gamma_\imath^k < 0$, we have $\nu_\imath^k = 0$ by (5.16), which gives a contradiction with $\nu_\imath^k H_\imath(x^k) > 0$. If $\gamma_\imath^k > 0$, we have

$$\begin{aligned} -\gamma_\imath^k G_\imath(x^k) - \nu_\imath^k H_\imath(x^k) &\leq -\gamma_\imath^k G_\imath(x^k) \\ &\leq \gamma_\imath^k \phi(G_\imath(x^k), H(x^k)) \\ &\leq M\phi(G_\imath(x^k), H_\imath(x^k)). \end{aligned}$$

- The case where $\gamma_\imath^k G_\imath(x^k) > 0$ and $\nu_\imath^k H_\imath(x^k) < 0$ is similar to the second case.
- If $\gamma_\imath^k G_\imath(x^k) \leq 0$ and $\nu_\imath^k H_\imath(x^k) \leq 0$, by (5.16), we only need to consider the following cases:
  - If $\gamma_\imath^k \geq 0$ and $\nu_\imath^k \geq 0$, then

$$\begin{aligned} -\gamma_\imath^k G_\imath(x^k) - \nu_\imath^k H_\imath(x^k) &\leq \gamma_\imath^k \phi(G_\imath(x^k), H_\imath(x^k)) + \nu_\imath^k \phi(G_\imath(x^k), H_\imath(x^k)) \\ &\leq 2M\phi(G_\imath(x^k), H_\imath(x^k)). \end{aligned}$$

  - If $\gamma_\imath^k < 0$ and $\nu_\imath^k = 0$, by the definition of $(\gamma_\imath^k, \nu_\imath^k)$, we have $H_\imath(\bar{x}^k) \geq G_\imath(\bar{x}^k) = 0$ and hence, due to $\bar{x}^k \to x^*$ and the strict complementarity, we have $H_\imath(x^*) > G_\imath(x^*)$. Since $x^k \to x^*$, we have $H_\imath(x^k) > G_\imath(x^k)$ for each k

sufficiently large. Then

$$\begin{aligned}
-\gamma_\imath^k G_\imath(x^k) - \nu_\imath^k H_\imath(x^k) &= -\gamma_\imath^k G_\imath(x^k) \\
&\leq MG_\imath(x^k) \\
&= M\min(G_\imath(x^k), H_\imath(x^k)) \\
&\leq M\phi(G_\imath(x^k), H_\imath(x^k)).
\end{aligned}$$

– The case where $\gamma_\imath^k = 0$ and $\nu_\imath^k < 0$ is similar to the above case.

In consequence, (5.15) holds. It follows from (5.13)–(5.15) that, for any k sufficiently large,

$$\begin{aligned}
\text{dist}_{\mathfrak{F}}(x^k) &= \|x^k - \bar{x}^k\| \\
&\leq 4M\Big(\sum_{j=1}^{q} g_j^+(x^k) + \sum_{i=1}^{p} |h_i(x^k)| + \sum_{\imath=1}^{m} \phi\big(G\imath(x^k), H\imath(x^k)\big)\Big).
\end{aligned} \tag{5.17}$$

From [42, Lemma 4.1], we have $\phi(G(x^k), H(x^k)) = \sum_{l=1}^{m} \text{dist}_C(G_l(x^k), H_l(x^k))$. It then follows from (5.17) that, for any k sufficiently large,

$$\text{dist}_{\mathfrak{F}}(x^k) \leq 4nM\text{dist}_\Lambda(F(x^k)),$$

which gives a contradiction with (5.10). Thus, the local error bound condition holds at x^* and so the proof is complete. □

In what follows, we give a sufficient condition to admit a local error bound in the setting of MPEC. For simplicity, if $y \notin \mathfrak{F}$, we denote by $\mathcal{M}_y(x)$ the set of M-multipliers of the following problem at x:

$$\begin{aligned} \min_{x} \quad & \|y - x\| \\ \text{s.t.} \quad & g(x) \leq 0, \;\; h(x) = 0, \\ & 0 \leq G(x) \;\perp\; H(x) \geq 0. \end{aligned}$$

That is, $\mathcal{M}_y(x)$ is the set of multipliers $(\mu, \lambda, \gamma, \nu) \in \mathcal{N}_{[-\infty,0]^q}(g(x)) \times \mathcal{N}_{\{0\}^p}(h(x)) \times \mathcal{N}_{C^m}(G(x), H(x))$ satisfying

$$\frac{y - x}{\|x - y\|} = \sum_{j=1}^{q} \mu_j^k \nabla g_j(x) + \sum_{i=1}^{p} \lambda_i^k \nabla h_i(x) - \sum_{\imath=1}^{m} \gamma_\imath^k \nabla G_\imath(x) - \sum_{\jmath=1}^{m} \nu_\jmath^k \nabla H_\jmath(x).$$

Theorem 5.10. *Suppose that, for each sequence $\{x^k\}$ converging to a strictly complementary solution x^* with $x^k \notin \mathcal{F}$, there exists a number $M > 0$ such that, for all k, the condition $\mathcal{M}_M^k(x) = \{(\lambda, \mu, u, v) \in \mathcal{M}_{x^k}(x) : \|(\lambda, \mu, \gamma, \nu)\| \leq M\} \neq \emptyset$ holds with some point x in $\prod_{\mathfrak{F}}(x^k)$. Then $\mathcal{X}$ admits a local error bound at x^*.*

Proof. It is not difficult to obtain the desired result from the proof process of Theorem 5.9. □

The strict complementarity condition in Theorems 5.9–5.10 is a strong assumption. However, for the following slightly stronger constraint qualification, called MPEC-r̃CPLD in [14], the strict complementarity assumption is redundant to ensure the same result to hold.

Definition 22. Let $x^* \in \mathfrak{F}$ and $\mathcal{I}_1 \subseteq \{1, \cdots, p\}$, $\mathcal{I}_2 \subseteq I_{0+}$, $\mathcal{I}_3 \subseteq I_{+0}$ be the index sets such that $\mathcal{G}(x^*; \mathcal{I}_1, \mathcal{I}_2, \mathcal{I}_3)$ is a basis for $\operatorname{span} \mathcal{G}(x^*; \{1, \cdots, p\}, I_{0+}^*, \mathcal{K}^*)$. We say that the *MPEC enhanced relaxed constant positive linear dependence condition* (MPEC-ERCPLD) holds at x^* iff there exists $\delta > 0$ such that

– $\mathcal{G}(x; \{1, \cdots, p\}, I_{0+}, I_{+0})$ has the same rank for each $x \in \mathcal{B}_\delta(x^*)$;

– for each $\mathcal{I}_4 \subseteq A(x^*)$ and $\mathcal{I}_5, \mathcal{I}_6 \subseteq I_{00}$, if there exist vectors $\{\lambda, \mu, \gamma, \nu\}$ with $\mu_j \geq 0$ for each $j \in \mathcal{I}_4$, which are not all zero, such that

$$\sum_{j \in \mathcal{I}_4} \mu_j \nabla g_j(x^*) + \sum_{i \in \mathcal{I}_1} \lambda_i \nabla h_i(x^*) - \sum_{\imath \in \mathcal{I}_2 \cup \mathcal{I}_5} u_\imath \nabla G_\imath(x^*) - \sum_{\jmath \in \mathcal{I}_3 \cup \mathcal{I}_6} v_\jmath \nabla H_\jmath(x^*) = 0,$$

then, for any $x \in \mathcal{B}_\delta(x^*)$, the vectors

$$\{\nabla g_j(x) : j \in \mathcal{I}_4\}, \quad \{\nabla h_i(x) : i \in \mathcal{I}_1\},$$
$$\{\nabla G_\imath(x) : \imath \in \mathcal{I}_2 \cup \mathcal{I}_5\}, \quad \{\nabla H_\jmath(x) : \jmath \in \mathcal{I}_3 \cup \mathcal{I}_6\}$$

are linearly dependent.

Theorem 5.11. *Suppose that the MPEC-ERCPLD holds at $x^* \in \mathfrak{F}$. Then x^* satisfies a local error bound.*

Proof. (Our proof technique is similar to [14, Theorem 3.2]. For the sake of completeness, we give a brief proof here.) Noting that the existence of local error bounds for a constraint system is equivalent to the calmness of the associated perturbed constraint system mapping (see, e.g., [39, page 438]), we investigate the calmness of the following perturbed constraint mapping:

$$\mathfrak{F}(p, q, r, s) := \left\{ x : \begin{array}{l} g(x) + p \leq 0, \;\; h(x) + q = 0 \\ 0 \leq G(x) + r \perp H(x) + s \geq 0 \end{array} \right\}.$$

Clearly, $\mathfrak{F}(0, 0, 0, 0) = \mathfrak{F}$. We next show the result by the mathematical induction.

First, if $m = 0$, which means that the constraint system $\mathfrak{F}$ has no complementarity constraints, then $\mathcal{X}$ reduces to an ordinary system of equalities and inequalities and it is easy to see that the RCPLD holds at x^*. By Theorem 5.5 and the equivalence of local error bounds and calmness, we get the desired result.

Suppose that the calmness condition holds at x^* for each $m \leq k$. In order to show that the calmness condition holds at x^* when $m = k+1$, we consider the constraint system mappings

$$\mathfrak{F}_1(p,q,r,s) := \left\{ x: \begin{array}{l} g(x)+p \leq 0, \ h(x)+q = 0 \\ G_{k+1}(x)+r_{k+1} = 0, \ H_{k+1}(x)+s_{k+1} \geq 0 \\ 0 \leq G_\imath(x)+r_\imath \perp H_\imath(x)+s_\imath \geq 0, \quad \imath \in \{1,\cdots,k\} \end{array} \right\}$$

and

$$\mathfrak{F}_2(p,q,r,s) := \left\{ x: \begin{array}{l} g(x)+p \leq 0, \ h(x)+q = 0 \\ G_{k+1}(x)+r_{k+1} \geq 0, \ H_{k+1}(x)+s_{k+1} = 0 \\ 0 \leq G_\imath(x)+r_\imath \perp H_\imath(x)+s_\imath \geq 0, \quad \imath \in \{1,\cdots,k\} \end{array} \right\}.$$

Denote by $\mathfrak{F}_i := \mathfrak{F}_i\,(0,0,0,0)$ for $i = 1,2$. It is easy to verify that $\mathfrak{F}_1(p,q,r,s) \cup \mathfrak{F}_2(p,q,r,s) = \mathfrak{F}(p,q,r,s)$. Since the existence of local error bounds is a kind of local property, without any loss of generality, we may assume that $x^* \in \mathfrak{F}_1 \cap \mathfrak{F}_2$. It is not hard to verify that the MPEC-ERCPLD holds at x^* for $\mathfrak{F}_1$ and $\mathfrak{F}_2$. Since both $\mathfrak{F}_1$ and $\mathfrak{F}_2$ have k complementarity constraints, by the induction hypothesis, there exist $\kappa > 0$ and $\delta > 0$ such that

$$\begin{array}{ll} \mathrm{dist}_{\mathfrak{F}_1}(x) \leq \kappa\|(p,q,r,s)\|, & \forall x \in \mathcal{B}_\delta(x^*) \cap \mathfrak{F}_1(p,q,r,s), \ \forall (p,q,r,s) \in \mathcal{B}_\delta(0), \\ \mathrm{dist}_{\mathfrak{F}_2}(x) \leq \kappa\|(p,q,r,s)\|, & \forall x \in \mathcal{B}_\delta(x^*) \cap \mathfrak{F}_2(p,q,r,s), \ \forall (p,q,r,s) \in \mathcal{B}_\delta(0). \end{array}$$

Thus, for any $x \in \mathcal{B}_\delta(x^*) \cap \mathfrak{F}(p,q,r,s)$ and $(p,q,r,s) \in \mathcal{B}_\delta(0)$, we have

$$\mathrm{dist}_{\mathfrak{F}}(x) = \min(\mathrm{dist}_{\mathfrak{F}_1}(x), \mathrm{dist}_{\mathfrak{F}_2}(x)) \leq \kappa\|(p,q,r,s)\|.$$

This completes the proof. □

Since the MPEC-RCPLD coincides with the MPEC-ERCPLD when the strict complementarity condition holds, Theorem 5.9 can also be considered as a corollary of Theorem 5.11.

From [39, Proposition 1], we have the following result immediately. Note that, by direct calculation, the linearized cone $\mathcal{L}_{\mathfrak{F}}(x^*) := \{d : \nabla F(x^*)^T d \in \mathcal{T}_\Lambda(F(x^*))\}$ is actually the MPEC linearized cone (see also [26, 27, 42]).

Corollary 5.12. *Suppose that the MPEC-ERCPLD holds at $x^* \in \mathfrak{F}$ or the MPEC-RCPLD holds at a strictly complementary point $x^* \in \mathfrak{F}$. Then the MPEC Abadie CQ holds at x^*, i.e., $\mathcal{T}_{\mathfrak{F}}(x^*) = \mathcal{L}_{\mathfrak{F}}(x^*)$.*

Since the objective function of the MPEC is continuously differentiable and hence locally Lipschitzian, it follows from the Clarke's exact penalty principle [16, Proposition 2.4.3] that the following exact penalty result holds.

Corollary 5.13. *Let x^* be a local minimizer of the MPEC. If the MPEC-ERCPLD holds at $x^* \in \mathfrak{F}$ or the MPEC-RCPLD holds at a strictly complementary point $x^* \in \mathfrak{F}$, then x^* is a local minimizer of the following penalized problem:*

$$\min f(x) + cL_f \mathrm{dist}_\Lambda(F(x))$$

where L_f is a Lipschitzian constant of f and c is the local error bound constant in Theorem 5.9.

From Corollary 5.13 and [74, Theorem 10.1], it is easy to get that, if x^* is a local minimizer of the MPEC, then

$$0 \in \partial f(x) + cL_f \partial \mathrm{dist}_\Lambda(F(x^*)),$$

and hence x^* is an M-stationary point by making use of the exact expression of $\partial \mathrm{dist}_\Lambda(F(x^*))$ [42, Lemma 4.2]. Moreover, it follows from [42, Lemma 4.1] that

$$\begin{aligned}\mathrm{dist}_\Lambda(F(x)) &= \sum_{j=1}^{q} \max(g_j(x), 0) + \sum_{i=1}^{p} |h_i(x)| + \sum_{\imath=1}^{m} \mathrm{dist}_C(G_\imath(x), H_\imath(x)) \\ &= \sum_{i=1}^{p} \max(g_i(x), 0) + \sum_{j=1}^{q} |h_j(x)| + \sum_{\imath=1}^{m} \Upsilon_\imath(x^k),\end{aligned}$$

where $\Upsilon_\imath(x^k) = \max\big(-G_\imath(x^k), -H_\imath(x^k), -G_\imath(x^k) - H_\imath(x^k), \min(G_\imath(x^k), H_\imath(x^k))\big)$. As a result, by smoothing the “max” and “min” functions, some approximation methods may be proposed to solve the MPEC.

Bibliography

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

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

[3] Andreani, R., Martinez, J.M. and Schuverdt, M.L. (2005). On the relation between constant positive linear dependence condition and quasinormality constraint qualification. J. Optim. Theo. Appl., **125**, 473-483.

[4] Bector, C.R, Chandra, S. and Dutta, J. (2004). *Principle of Optimization Theory.* Narosa Publishers, India and Alpha Science Publishers.

[5] Bertsekas, D.P. (1999). *Nonlinear Programming.* Athena Scientific Publishers.

[6] Bertsekas, D.P., Nedić, A. and Ozdaglar, A.E. (2003). *Convex Analysis and Optimization.* Belmont, MA, Athena Scientific.

[7] Bertsekas, D.P. and Ozdaglar, A.E. (2002). Pseudonormality and a Lagrange multiplier theory for constrained optimization. J. Optim. Theo. Appl., **114**, 287-343.

[8] Bertsekas, D.P. and Ozdaglar, A.E. (2004). The relation between pseudonormality and quasiregularity in constrained optimization. Optim. Meth. Softw., **19**, 493-506.

[9] Bertsekas, D.P., Ozdaglar, A.E. and Tseng, P. (2006). Enhanced Fritz John conditions for convex programming. SIAM J. Optim., **16**, 766-797.

[10] Bonnans, J.F. and Shapiro, A. (2000). *Perturbation Analysis of Optimization Problems*. Springer, New York.

[11] Borwein, J.M. and Lewis, A.S. (2000). *Convex Analysis and Nonlinear Optimization*. Canadian Mathematical Society Books in Mathematics, Springer, New York.

[12] Borwein, J.M. and Zhu, Q.J. (2005). *Techniques of Variational Analysis*. Canadian Mathematical Society Books in Mathematics, Springer, New York.

[13] Chieu, N.H. and Lee, G.M. (2013). A relaxed constant positive linear dependence constraint qualificaiton for mathematical programs with equilibrium constraints. J. Optim. Theo. Appl., **158**, 11-32.

[14] Chieu, N.H. and Lee, G.M. (2014). Constraint qualifications for mathematical programs with equilibrium constraints and their Local preservation property. J. Optim. Theo. Appl., doi : 10.1007/s10957-014-0546-2.

[15] Clarke, F.H. (1976). A new approach to Lagrange multipliers. Math. Oper. Res., **1**, 165-174.

[16] Clarke, F.H. (1990). *Optimization and Nonsmooth Analysis*. Wiley Interscience, New York, 1983; reprinted as vol. 5 of Classics Appl. Math. 5, SIAM, Philadelphia, PA.

[17] Clarke, F.H., Ledyaev, Y.S., Stern, R.J. and Wolenski, P.R. (1998). *Nonsmooth Analysis and Control Theory.* Springer, New York.

[18] Dempe, S. (2002). *Foundations of Bilevel Programming.* Kluwer Academic Publishers, Dordrecht.

[19] Dempe, S., Mordukhovich, B.S. and Zemkoho, A.B. (2014). Necessary optimality conditions in pessimistic bilevel programming. Optim., **63**, 505-533.

[20] Dempe, S., Mordukhovich, B.S. and Zemkoho, A.B. (2012). Sensitivity analysis for two-level value functions with applications to bilevel programming. SIAM J. Optim., **22**, 1309-1343.

[21] Ding, C., Sun, D. and Ye, J.J. (2013). First order optimality conditions for mathematical programs with semidefinite cone complementarity constraints. Math. Program, ser. A., doi : 10.1007/s10107-013-0735-z.

[22] Fabian, M.J., Henrion, R., Kruger, A.Y. and Outrata, J.V. (2010). Error bounds: Necessary and sufficient conditions. Set-Valued Var. Anal., **18**, 121-149.

[23] Flegel, M.L. and Kanzow, C. (2005). On the Guignard constraint qualification for mathematical programs with equilibrium constraints. Optim., **54**, 517-534.

[24] Glover, B.M. and Craven, B.D. (1994). A Fritz John optimality condition using the approximate subdifferential. J. Optim. Theo. Appl., **82**, 253-265.

[25] Gould, F.J. and Tolle, J.W. (1971). A necessary and sufficient qualification for constrained optimization. SIAM J. Appl. Math., **20**, 164-172.

[26] Guo, L. and Lin, G.H. (2013). Notes on some constraint qualifications for mathematical programs with equilibrium constraints. J. Optim. Theo. Appl., **156**, 600-616.

[27] Guo, L., Lin, G.H. and Ye, J.J. (2012). Stability analysis for parametric mathematical programs with geometric constraints and its applications. SIAM J. Optim., **22**, 1151-1176.

[28] Guo, L., Lin, G.H. and Ye, J.J. (2013). Second order optimality conditions for mathematical programs with equilibrium constraint. J. Optim. Theo. Appl., **158**, 33-64.

[29] Henrion, R., Jourani, A. and Outrata, J. (2002). On the calmness of a class of multifunctions. SIAM J. Optim., **13**, 603-618.

[30] Hestenes, M.R. (1975). *Optimization Theory: The Finite Dimensional Case.* Wiley, Now York.

[31] Hoffman, A.J. (1952). On approximate solutions of systems of linear inequalities. J. Res. Natl. Bur. Stand., **49**, 263-265.

[32] Hoheisel, T., Kanzow, C. and Schwartz, A. (2013). Theoretical and numerical comparison of relaxation methods for mathematical programs with complementarity constraints. Math. Program., Ser. A., **137**, 257-288.

[33] Hu, X.M. and Ralph, D. (2002). A note on sensitivity of value function of mathematical programs with complementarity constraints, Math. Program., **93**, 265-279.

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

[35] Ioffe, A.D. (1986). Approximate subdifferentials and applications, II: Functions on locally convex spaces. Mathematika., **33**, 111-128.

[36] Ioffe, A.D. (1989). Approximate subdifferentials and applications, III: The metric theory. Mathematatika., **36**, 1-38.

[37] Izmailov, A.F., Solodov, M.V. and Uskov, E.I. (2012). Global convergence of augmented Lagrangian methods applied to optimization problems with degenerate constraints, including problems with complementarity constraints. SIAM J. Optim., **22**, 1579-1606.

[38] John, F. (1948). Extremum problems with inequalities as side constraints. In "Studies and Essays, Courant Anniversary Volume" (K.O. Friedrichs, O.E. Neugebauer and J. J. Stoker, eds), 187-204. Wiley (interscience), New York.

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

[40] Jourani, A. (1994). Constraint qualification and Lagrange multipliers in nondifferentiable programming problems. J. Optim. Theo. Appl., **81**, 533-548.

[41] Jourani, A. and Thibault, L. (1993). The approximate subdifferential of composite functions. Bull. Aust. Math. Soc., **47**, 443-456.

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

[43] Kanzow, C. and Schwartz, A. (2013). A new regularization method for mathematical programs with complementarity constraints with strong convergence properties. SIAM J. Optim., **23**, 770-798.

[44] Kruger, A.Y., Minchenko, L. and Outrata, J.V. (2014). On relaxing the Mangasarian-Fromovitz constraint qualification. Positivity, **18**, 171-189.

[45] Lewis, A.S. (1999). Nonsmooth analysis of eigenvalues. Math. Program., **84**, 1-24.

[46] Loewen, P.D. (2003). *Optimal Control via Nonsmooth Analysis*. American Mathematical Society, Providence, Rhode Island.

[47] Lucet, Y. and Ye, J.J. (2001). Sensitivity analysis of the value function for optimization problems with variational inequality constraints. SIAM J. Contr. Optim., **40**, 699-723.

[48] Lucet, Y. and Ye, J.J. (2001). Erratum: sensitivity analysis of the value function for optimization problems with variational inequality constraints. SIAM J. Contr. Optim., **41**, 1315-1319.

[49] Luo, Z.Q., Pang, J.S. and Ralph, D. (1996). *Mathematical Programs with Equilibrium Constraints*. Cambridge University Press, Cambridge.

[50] Mangasarian, O.L. and Fromovitz, S. (1967). The Fritz John necessary optimality conditions in the presence of equality and inequality constraints. J. Math. Anal. Appl., **17**, 37-47.

[51] Maréchal, P. and Ye, J.J. (2009). Optimizing condition numbers. SIAM J. Optim., **20**, 935-947.

[52] McShane, E.J. (1973). The Lagrange multiplier rule. Amer. Math. Monthly., **80**, 922-925.

[53] Minchenko, L. and Stakhovski, S. (2010). About generalizing the Mangasarian-Fromovitz regularity condition. Doklady BGUIR., **8**, 104-109.

[54] Minchenko, L. and Stakhovski, S. (2011). On error bounds for quasinormal programs. J. Optim. Theo. Appl., **148**, 571-579.

[55] Minchenko, L. and Stakhovski, S. (2011). On relaxed constant rank regularity condition in mathematical programming. Optimization, **60**, 429-440.

[56] Minchenko, L. and Stakhovski, S. (2011). Parametric nonlinear programming problems under the relaxed constant rank condition. SIAM J. Optim., **21**, 314-332.

[57] Minchenko, L. and Turakanov, A. (2011). On error bounds for quasinormal programs. J. Optim. Theo. Appl., **148**, 571-579.

[58] Moldovan, A. and Pellegrini, L. (2009). On regularity for constrained extremum problems, part 2: Necessary optimality conditions. J. Optim. Theo. Appl., **142**, 165-183.

[59] Mordukhovich, B.S. (1980). Metric approximation and necessary optimality conditions for general classes of nonsmooth extremal problems. Soviet Math. Dokl., **22**, 526-530.

[60] Mordukhovich, B.S. (1992). Sensitivity analysis in nonsmooth analysis. in Theoretical Aspects of Industrial Design, edited by D.A. Field and V. Komkov, SIAM Proc. Appl. Math., **58**, 32-46, Philadelphia, Pennsylvania.

[61] Mordukhovich, B.S. (2006). *Variational Analysis and Generalized Differentiation I. Basic Theory*. Ser. Comprehensive Stud. Math., **330**, Springer, Berlin.

[62] Mordukhovich, B.S. (2006). *Variational Analysis and Generalized Differentiation II. Application*. Ser. Comprehensive Stud. Math., **331**, Springer, Berlin.

[63] Mordukhovich, B.S. and Nam, N.M. (2005). Variational stability and marginal functions via generalized differentiation. Math. Oper. Res., **30**, 800-816.

[64] Mordukhovich, B.S., Nam, N.M. and Phan H.M. (2012). Variational analysis of marginal functions with applications to bilevel programming. J. Optim. Theo. Appl., **152**, 557-586.

[65] Mordukhovich, B.S., Nam, N.M. and Yen, N.D. (2006). Fréchet subdifferential calculus and optimality conditions in nondifferentiable programming. Optim., **55**, 685-708.

[66] Mordukhovich, B.S., Nam, N.M. and Yen, N.D. (2009). Subgradients of marginal functions in parametric mathematical programming. Math. Program. Ser. B., **116**, 369-396.

[67] Mordukhovich, B.S. and Shao, Y.H. (1996). Nonsmooth sequential analysis in Asplund space, Trans. Amer. Math. Soc., **348**, 1235-1280.

[68] Outrata, J.V., Kocvara, M. and Zowe, J. (1998). *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints: Theory, Applications and Numerical Results*. Kluwer Academic Publishers, Boston.

[69] Overton, M.L. and Womersley, R.S. (1992). On the sum of the largest eigenvalues of a symmetric matrix. SIAM J. Matrix Anal. Appl., **13**, 41-45.

[70] Ozdaglar, A.E. and Bertsekas, D.P. (2004). The relation between pseudonormality and quasiregularity in constrained optimization. Optim. Methods Softw., **19**, 493-506.

[71] Qi, L. and Wei, Z. (2000). On the constant positive linear dependence condition and its application to SQP methods. SIAM J. Optim., **10**, 963-981 .

[72] Robinson, S.M. (1982). Generalized equations and their solution, part II: applications to nonlinear programming. Math. Program. Stud., **19**, 200-221.

[73] Rockafellar, R.T. (1970). *Convex Analysis*. Princeton University Press, Princeton, New Jersey.

[74] Rockafellar, R.T. and Wets, R. J.-B. (1998). *Variational Analysis*. Springer-Verlag, Berlin.

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

[76] Scheel, H.S. and Stöhr, M. (2001). How stringent is the linear independence assumption for mathematical programs with complementarity constraints. Math. Oper. Res., **26**, 851-863.

[77] Schirotzek, W. (2007). *Nonsmooth Analysis*. Springer, Berlin.

[78] Treiman, J.S. (1999). Lagrange multipliers for nonconvex generalized gradients with equality, inequality, and set constraints. SIAM J. Contr. Optim., **37**, 1313-1329.

[79] Wu, Z.L. and Ye, J.J. (2001). Sufficient conditions for error bounds. SIAM J. Optim., **12**, 421-435.

[80] Wu, Z.L. and Ye, J.J. (2003). First and second order conditions for error bounds. SIAM J. Optim., **14**, 621-645.

[81] Ye, J.J. (1999). Optimality conditions for optimization problems with complementarity constraints. SIAM J. Optim., **9**, 374-387.

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

[83] Ye, J.J. (2001). Multiplier rules under mixed assumptions of differentiability and Lipschitz continuity. SIAM J. Contr. Optim., **39**, 1441-1460.

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

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

[86] Ye, J.J. and Zhu, D.L. (1995). Optimality conditions for bilevel programming problems. Optim., **33**, 9-27.

[87] Ye, J.J., Zhu, D.L. and Zhu, Q.J. (1997). Exact penalization and necessary optimality conditions for generalized bilevel programming problems. SIAM J. Optim., **7**, 481-507.