1-1-2016

Variational Analysis And Stability In Optimization

M. Ebrahim Sarabi
Wayne State University,

VARIATIONAL ANALYSIS AND STABILITY IN OPTIMIZATION

by

M. EBRAHIM SARABI

DISSERTATION

Submitted to the Graduate School

of Wayne State University,

Detroit, Michigan

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

2016

MAJOR: (Applied) MATHEMATICS

Approved By:

Advisor	Date

DEDICATION

To my parents

ACKNOWLEDGEMENTS

It would be impossible to adequately express my gratitude and appreciation for my advisor, Professor Boris Mordukhovich. His ceaseless devotion, support, and encouragement has been the driving force behind all my accomplishments during my doctoral studies. He is the inspiration for the kind of mathematician I aspire to be. It was my great honor to have him as my advisor during the years at Wayne State University.

I am also indebted to my mentor Professor Dinh The Luc from whom I learned a lot about optimization and variational analysis.

I would like to thank the rest of my thesis committee: Professors Alper Murat, Peiyong Wang, George Yin, and Sheng Zhang for taking their time to serve on my dissertation committee.

My sincere thanks also go to the entire Department of Mathematics at Wayne State University for their kind support in a number of ways. Specially, I wish to thank Christopher Leirstein and Christopher Nazelli for their invaluable feedbacks and help with my teaching.

Most importantly, none of this would have been possible without the love and endless care of my family. I would like to thank my parents for their love, encouragement, and support.

TABLE OF CONTENTS

Dedication . . . . **ii**

Acknowledgements . . . . **iii**

CHAPTER 1: INTRODUCTION **1**

CHAPTER 2: FULL STABILITY FOR UNCONSTRAINED PROBLEMS **5**

2.1 Tools of Variational Analysis 5

2.2 Full Stability 8

2.3 Characterizations of Full Stability 13

CHAPTER 3: STABILITY ANALYSIS OF UNCONSTRAINED PROBLEMS **27**

3.1 Full Stability of Constrained Optimization Problems 29

3.2 Relationships of Full Stability with Other Stability Notions 34

CHAPTER 4: SECOND-ORDER ANALYSIS OF PIECEWISE LINEAR FUNCTIONS **42**

4.1 Reducibility, Nondegeneracy and Second-Order Qualification 45

4.2 Full Stability in Composite Optimization 55

4.3 Full Stability in Constrained Minimax Problems 62

4.4 Strong Regularity and Strong Stability in Composite Models 66

CHAPTER 5: CRITICAL MULTIPLIERS FOR COMPOSITE PROBLEMS **77**

5.1 Critical Multipliers for Composite Optimization Problems 84

5.2 Full Stability and Critical Multipliers 95

References **98**

Abstract **103**

Autobiographical Statement **104**

CHAPTER 1 INTRODUCTION

Lipschitzian stability of locally optimal solutions with respect to small parameter perturbations is undoubtedly important in optimization theory allowing us to recognize robust solutions and support computational work from the viewpoints of justifying numerical algorithms, their convergence properties, stopping criteria, etc. There are several versions of Lipschitzian stability in optimization; see, e.g., the books [1, 5, 10, 27, 53] and the references therein. The focus of this dissertation is on what is known as *full stability* of locally optimal solutions introduced by Levy, Poliquin and Rockafellar [20]. This notion emerged as a far-going extension of *tilt stability* of local minimizers in the sense of Poliquin and Rockafellar [45]; see Chapter 2 for the precise definitions and more discussions. It seems that full stability is probably the most fundamental stability notion for locally optimal solutions, from both theoretical and practical points of view, particularly in connection with numerical methodology and applications [20, 45].

In [20], the authors derived necessary and sufficient conditions for fully stable minimizers of parameterized optimization problems written in the unconstrained format with extended-real-valued and prox-regular cost functions. They expressed these conditions in terms of a partial modification of the *second-order subdifferential* (or generalized Hessian) in the sense of Mordukhovich [26], which was previously used in [45] for characterizations of tilt stability. As mentioned in [20], implementing this approach in particular classes of constrained optimization problems important for the theory and applications requires the developments of *second-order subdifferential calculus* for the constructions involved, which was challenging and not available at that time. Partly such a calculus has been developed in the recent paper by Mordukhovich and Rockafellar [38] with applications to tilt stability therein.

Quite recently, Mordukhovich and Nghia [31] have developed a new approach to both Lipschitzian and Hölderian (introduced therein) full stability in finite and infinite dimensions and applied it to deriving constructive characterization of full stability in NLPs, infinite-dimensional problems with polyhedric constraints, and optimal control problems governed by semilinear elliptic equations without any nondegeneracy assumption. Moreover, they extended their results in [33] for constrained optimization problems under the nondegeneracy assumption using the uniform second-order growth condition.

The main goal of this dissertation is to obtain *complete characterizations* of full stability for remarkable classes of constrained optimization problems expressing these characterizations entirely in terms of the problem data. The classes under consideration include general models given in *composite formats* of optimization (particularly with fully amenable compositions), and consequently for classical problems of *nonlinear programming* (NLP) with $\mathcal{C}^2$ equality and inequality constraints. The key machinery is based on *exact* (equality type) second-order calculus rules for the aforementioned constructions taken partly from [38] and also the new ones derived in this dissertation.

The rest of the dissertation is organized as follows. In Chapter 2 we review the basic generalized differential tools of variational analysis used in formulations and proofs of the main results. We then presents definitions of full stability and related notions for optimization problems written in the *unconstrained* extended-real-valued format. We discuss the second-order necessary and sufficient conditions for full stability of local minimizers in this setting [20] and establish relationships between full stability of local minimizers and the new notion of *partial strong metric regularity* (PSMR) of the corresponding subdifferential mappings. Then these conditions are characterized via a certain *uniform second-order growth condition* (USOGC).

Chapter 3 addresses a general class of constrained optimization problems covering those in conic programming, establishes new properties of fully stable minimizers, and provides new proofs of major second-order characterizations of fully stable minimizers under reducibility and partial nondegeneracy conditions. In particular, the developed approach allows us to describe the framework of canonical perturbations, where full stability is equivalent to tilt stability under an appropriate parametric reduction. Then we investigate relationships between full stability of local minimizers for general constrained optimization problems with $\mathcal{C}^2$-smooth data and Lipschitzian single-valued localizations of solution maps to the corresponding KKT (Karush-Kuhn-Tucker) systems. The obtained results with self-contained proofs ensure, in particular, the equivalence between full stability of local minimizers and Robinson's strong regularity [47] of the associated generalized equations as well as strong Lipschitz stability of stationary points with respect to $\mathcal{C}^2$-smooth perturbations.

The major result of Chapter 4 is to establish a certain equivalence between qualification conditions used in [28, 38] for deriving by different approaches the exact second-order chain rule for fully amenable compositions involving convex piecewise linear (CPWL) functions. Being important for its own sake, the key ingredient of this result (together with the explicit calculation of the second-order subdifferential of CPWL functions) is the proof of the so-called $\mathcal{C}^\infty$*-reducibility* of CPWL functions via *linear* transformations that are used then in the formulation of partial nondegeneracy. As a by-product of the obtained equivalence, we completely clarify the essence of the powerful second-order chain rule that is largely employed in the subsequent material. Next we present the explicit composite SSOSC characterization of *fully stable* local minimizers in the partially nondegenerate *composite framework* of optimization involving CPWL functions. Then we effectively apply the general composite result to characterizing full stability of local optimal solutions of *minimax* problems with *polyhedral constraints*. Finally, we finish our study of composite optimization models with CPWL functions with characterizing *strong regularity* of the associated KKT systems and *strong stability* of the related stationary points under perturbations. We prove that these differently defined notions are *equivalent* to full stability of the corresponding local minimizers under partial nondegeneracy being therefore completely characterized by the aforementioned SSOSC.

The last chapter discusses a very recent concept of critical multipliers for composite optimization problems. We first formulate the later concept in this framework and then characterize it via second-order subdifferentials. Then we show that full stability of local minimizers can rule out critical multipliers for problems of composite optimization.

CHAPTER 2 FULL STABILITY FOR UNCONSTRAINED PROBLEMS

2.1 Tools of Variational Analysis

In this chapter we briefly overview some basic constructions of generalized differentiation in variational analysis, which are widely used in what follows. Then we discuss the concept of full stability in the general framework of unconstrained optimization problems. Finally, we provide some characterizations of this concept, which open the door to proceed in the subsequent chapters. We start with recalling the corresponding first-order subdifferentials as well as associated objects of variational geometry. Given $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ finite at $\bar{x}$, its *regular subdifferential* (known also as the presubdifferential and as the Fréchet or viscosity subdifferential) at $\bar{x}$ is

$$\widehat{\partial}\varphi(\bar{x}) := \Big\{ v \in \mathbb{R}^n \Big| \ \liminf_{x \to \bar{x}} \frac{\varphi(x) - \varphi(\bar{x}) - \langle v, x - \bar{x} \rangle}{\|x - \bar{x}\|} \ge 0 \Big\}. \tag{2.1}$$

While $\widehat{\partial}\varphi(\bar{x})$ reduces to a singleton $\{\nabla\varphi(\bar{x})\}$ if φ is Fréchet differentiable at $\bar{x}$ and to the classical subdifferential of convex analysis if φ is convex, the set (2.1) may often be empty for nonconvex and nonsmooth functions as, e.g., for $\varphi(x) = -|x|$ at $\bar{x} = 0 \in \mathbb{R}$. Another serious disadvantage of (2.1) is the failure of standard calculus rules inevitably required in the theory and applications of variational analysis including those to optimization and equilibria.

The picture dramatically changes when we perform a limiting procedure over the mapping $x \mapsto \widehat{\partial}\varphi(x)$ as $x \xrightarrow{\varphi} \bar{x}$ that leads us to the (basic first-order) *subdifferential* of φ at $\bar{x}$ defined by

$$\partial\varphi(\bar{x}) := \mathop{\mathrm{Lim\,sup}}_{x \xrightarrow{\varphi} \bar{x}} \widehat{\partial}\varphi(x) \tag{2.2}$$

and known also as the general, or limiting, or Mordukhovich subdifferential; it was first introduced in [24] in an equivalent way. In contrast to (2.1), the subgradient set (2.2) is often nonconvex (e.g., $\partial\varphi(0) = \{-1, 1\}$ for $\varphi(x) = -|x|$) while enjoying a *full calculus* based on *variational/extremal principles*, which replace separation arguments in the absence of convexity.

We need also another first-order subdifferential construction for $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ finite at $\bar{x}$, which is complemented to (2.2) in the case of non-Lipschitzian functions. The *singular/horizon subdifferential* of φ at $\bar{x}$ is defined by

$$\partial^\infty\varphi(\bar{x}) := \operatorname*{Lim\,sup}_{\substack{x \xrightarrow{\varphi} \bar{x} \\ \lambda \downarrow 0}} \lambda\widehat{\partial}\varphi(x). \tag{2.3}$$

We know that $\partial^\infty\varphi(\bar{x}) = \{0\}$ if and only if φ is locally Lipschitzian around $\bar{x}$, provided that it is lower semicontinuous (l.s.c.) around this point.

Recall further some constructions of variational geometry needed in what follows and associated with the subdifferential ones defined above. Given a set $\emptyset \neq \Omega \subset \mathbb{R}^n$, consider its indicator function $\delta(x;\Omega)$ equal to 0 for $x \in \Omega$ and to ∞ otherwise. For any fixed $\bar{x} \in \Omega$, the *regular normal cone* to Ω at $\bar{x}$ is

$$\widehat{N}(\bar{x};\Omega) := \widehat{\partial}\delta(\bar{x};\Omega) = \Big\{ v \in \mathbb{R}^n \Big| \limsup_{x \xrightarrow{\Omega} \bar{x}} \frac{\langle v, x - \bar{x}\rangle}{\|x - \bar{x}\|} \le 0 \Big\} \tag{2.4}$$

and the (basic, limiting) *normal cone* to Ω at $\bar{x}$ is $N(\bar{x};\Omega) := \partial\delta(\bar{x};\Omega)$. It follows from (2.2) and (2.4) that the normal cone $N(\bar{x};\Omega)$ admits the limiting representation

$$N(\bar{x};\Omega) = \operatorname*{Lim\,sup}_{x \xrightarrow{\Omega} \bar{x}} \widehat{N}(x;\Omega) \tag{2.5}$$

via the Painlevé-Kuratowski outer limit. Observe also the *duality/polarity* correspondence

$$\widehat{N}(\bar{x};\Omega) = T(\bar{x};\Omega)^* := \Big\{ v \in \mathbb{R}^n \Big| \langle v, w\rangle \le 0 \ \text{ for all } \ w \in T(\bar{x};\Omega) \Big\} \tag{2.6}$$

between the regular normal cone (2.4) and the *tangent cone* to Ω at $\bar{x} \in \Omega$ defined by

$$T(\bar{x};\Omega) := \Big\{ w \in \mathbb{R}^n \Big| \ \exists\, x_k \xrightarrow{\Omega} \bar{x},\ \alpha_k \ge 0 \ \text{ with } \ \alpha_k(x_k - \bar{x}) \to w \ \text{ as } \ k \to \infty \Big\} \tag{2.7}$$

and known also as the Bouligand-Severi contingent cone to Ω at this point. Note that the basic normal cone (2.5) *cannot* be tangentially generated in a polar form (2.6), since it is intrinsically nonconvex while the polar T^* to any set T is always convex. In what follows we may also use the subindex set notation like $N_\Omega(\bar{x})$, $T_\Omega(\bar{x})$, etc. for the constructions involved.

Given further a mapping $F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, define its *coderivative* [25] at $(\bar{x},\bar{y}) \in F$ by

$$D^*F(\bar{x},\bar{y})(v) := \Big\{u \in \mathbb{R}^n \Big|\ (u,-v) \in N\big((\bar{x},\bar{y});F\big)\Big\}, \quad v \in \mathbb{R}^m, \tag{2.8}$$

via the normal cone (2.5) to the graph F. The set-valued mapping $D^*F(\bar{x},\bar{y})\colon \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ is clearly positive-homogeneous. Moreover, if the mapping $F\colon \mathbb{R}^n \to \mathbb{R}^m$ is single-valued (then we omit $\bar{y} = F(\bar{x})$ in the coderivative notation) and *strictly differentiable* at $\bar{x}$ (which is automatic when it is $\mathcal{C}^1$ around this point), then the coderivative (2.8) is also single-valued and reduces to the *adjoint* derivative operator

$$D^*F(\bar{x})(v) = \big\{\nabla F(\bar{x})^* v\big\}, \quad v \in \mathbb{R}^m, \tag{2.9}$$

with the operator symbol '$*$' on the right-hand side of (2.9) standing for the matrix transposition in finite dimensions. It is worth noting that the coderivative values in (2.8) are often nonconvex sets due to the intrinsic nonconvexity of the normal cone on the right-hand side therein. Observe furthermore that this nonconvex normal cone is taken to a *graphical* set. Thus its convexification in (2.8), which reduces to the convexified/Clarke normal cone to the set in question, creates serious troubles; see Rockafellar [50] and Mordukhovich [27, Subsection 3.2.4] for more details.

Coming back to extended-real-valued functions, let us present their *second-order* subdifferential constructions, which are at the heart of the variational techniques developed in this dissertation. Given $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ finite at $\bar{x}$, pick a subgradient $\bar{y} \in \partial\varphi(\bar{x})$ and follow Mordukhovich [26] to introduce the *second-order subdifferential* (or generalized Hessian) of φ at $\bar{x}$ relative to $\bar{y}$ by

$$\partial^2\varphi(\bar{x},\bar{y})(u) := (D^*\partial\varphi)(\bar{x},\bar{y})(u), \quad u \in \mathbb{R}^n \tag{2.10}$$

via the coderivative (2.8) of the first-order subdifferential mapping (2.2). Observe that for

$\varphi \in \mathcal{C}^2$ with the (symmetric) Hessian matrix $\nabla^2\varphi(\bar{x})$ we have

$$\partial^2\varphi(\bar{x})(u) = \Big\{\nabla^2\varphi(\bar{x})u\Big\} \text{ for all } u \in \mathbb{R}^n.$$

From now on we focus on an appropriate partial counterpart of (2.10) for functions $\varphi\colon \mathbb{R}^n \times \mathbb{R}^d \to \overline{\mathbb{R}}$ of two variables $(x, w) \in \mathbb{R}^n \times \mathbb{R}^d$. Consider the partial first-order subgradient mapping

$$\partial_x\varphi(x, w) := \Big\{\text{set of subgradients } v \text{ of } \varphi_w := \varphi(\cdot, w) \text{ at } x\Big\} = \partial\varphi_w(x), \tag{2.11}$$

take $(\bar{x}, \bar{w})$ with $\varphi(\bar{x}, \bar{w}) < \infty$, and define the *extended partial second-order subdifferential* of φ with respect to x at $(\bar{x}, \bar{w})$ relative to some $\bar{y} \in \partial_x\varphi(\bar{x}, \bar{w})$ by

$$\widetilde{\partial}^2_x\varphi(\bar{x}, \bar{w}, \bar{y})(u) := (D^*\partial_x\varphi)(\bar{x}, \bar{w}, \bar{y})(u), \quad u \in \mathbb{R}^n. \tag{2.12}$$

This second-order construction was first employed by Levy, Poliquin and Rockafellar [20] for characterizing full stability of extended-real-valued functions in the unconstrained format of optimization. Note that the second-order construction (2.12) is different from the standard partial second-order subdifferential

$$\partial^2_x\varphi(\bar{x}, \bar{w}, \bar{y})(u) := (D^*\partial\varphi_{\bar{w}})(\bar{x}, \bar{y})(u) = \partial^2\varphi_{\bar{w}}(\bar{x}, \bar{y})(u), \quad u \in \mathbb{R}^n,$$

of $\varphi = \varphi(x, w)$ with respect to x at $(\bar{x}, \bar{w})$ relative to $\bar{y} \in \partial_x\varphi(\bar{x}, \bar{w})$, even in the classical $\mathcal{C}^2$ setting. Indeed, for such functions φ with $\bar{y} = \nabla_x\varphi(\bar{x}, \bar{w})$ we have

$$\partial^2_x\varphi(\bar{x}, \bar{w})(u) = \Big\{\nabla^2_{xx}\varphi(\bar{x}, \bar{w})u\Big\} \text{ while}$$

$$\widetilde{\partial}^2_x\varphi(\bar{x}, \bar{w})(u) = \Big\{(\nabla^2_{xx}\varphi(\bar{x}, \bar{w})u, \nabla^2_{xw}\varphi(\bar{x}, \bar{w})u)\Big\} \text{ for all } u \in \mathbb{R}^n. \tag{2.13}$$

2.2 Full Stability

Let $\varphi\colon \mathbb{R}^n \times \mathbb{R}^d \to \overline{\mathbb{R}} = (-\infty, \infty]$ be a proper extended-real-valued function of two variables $(x, w) \in \mathbb{R}^n \times \mathbb{R}^d$. Throughout this chapter we assume, unless otherwise stated, that φ is *lower semicontinuous* around the reference points of its effective domain

$$\operatorname{dom}\varphi := \Big\{(x,w)\in\mathbb{R}^n\times\mathbb{R}^d \big|\ \varphi(x,w)<\infty\Big\}.$$

Following Levy, Poliquin and Rockafellar [20], consider the two-parametric unconstrained problem of minimizing the perturbed function φ defined by

$$\text{minimize}\ \ \varphi(x,w)-\langle v,x\rangle\ \ \text{over}\ \ x\in\mathbb{R}^n \tag{2.14}$$

and label it as $\mathcal{P}(w,v)$. In this parameterized optimization problem, the vector $w\in\mathbb{R}^d$ signifies general parameter perturbations (called *basic* perturbations in [20]) while the linear parametric shift of the objective with $v\in\mathbb{R}^n$ in (2.14) represents the so-called *tilt* perturbations. We consider the following fairly general type of *quantitative/Lipschitzian stability* of local minimizers for the parameterized family $\mathcal{P}(w,v)$ of the optimization problems (2.14) with respect to parameter perturbations (w,v) varying around the given nominal parameter value $(\bar w,\bar v)$ corresponding to the unperturbed problem $\mathcal{P}(\bar w,\bar v)$. Feasible solutions to $\mathcal{P}(w,v)$ are the points $x\in\mathbb{R}^n$ such that the function value $\varphi(x,w)$ is finite.

Let $\bar x$ be a feasible solution to the unperturbed problem $\mathcal{P}(\bar w,\bar v)$. For any number $\gamma>0$ we consider the (local) optimal value function

$$m_\gamma(w,v):=\inf_{\|x-\bar x\|\le\gamma}\Big\{\varphi(x,w)-\langle v,x\rangle\Big\},\quad (w,v)\in\mathbb{R}^d\times\mathbb{R}^n, \tag{2.15}$$

for the perturbed optimization problem (2.14) and then the corresponding parametric family of optimal solution sets to (2.14) given by

$$M_\gamma(w,v):=\operatorname{argmin}_{\|x-\bar x\|\le\gamma}\Big\{\varphi(x,w)-\langle v,x\rangle\Big\},\quad (w,v)\in\mathbb{R}^d\times\mathbb{R}^n. \tag{2.16}$$

A point $\bar x$ is said to be a *locally optimal solution* to $\mathcal{P}(\bar w,\bar v)$ if $\bar x\in M_\gamma(\bar w,\bar v)$ for some small $\gamma>0$. We recall the following notion of Lipschitzian stability for locally optimal solutions to the unperturbed problem $\mathcal{P}(\bar w,\bar v)$ introduced in [20].

Definition 2.1 (full stability). *A point $\bar x$ is a* FULLY STABLE *locally optimal solution* *to problem* $\mathcal{P}(\bar{w},\bar{v})$ *if there exist a number* $\gamma>0$ *and neighborhoods* W *of* $\bar{w}$ *and* V *of* $\bar{v}$ *such that the mapping* $(w,v)\mapsto M_\gamma(w,v)$ *is single-valued and Lipschitz continuous with* $M_\gamma(\bar{w},\bar{v})=\{\bar{x}\}$ *and the function* $(w,v)\mapsto m_\gamma(w,v)$ *is likewise Lipschitz continuous on* $W\times V$.

Tilt stability of local minimizers $\bar{x}$ introduced earlier by Poliquin and Rockafellar [45] corresponds to Definition 2.1 under the fixed basic parameter $w=\bar{w}$, i.e., it imposes single-valued Lipschitzian behavior of $v\to M_\gamma(\bar{w},v)$ with respect to tilt perturbations v in (2.14). Observe that in this case the Lipschitz continuity of the optimal value functions $m_\gamma(\bar{w},v)$ is automatic in the finite-dimensional setting under consideration, since it follows from (2.15) that $m_\gamma(\bar{w},v)$ is finite and concave in v. Note also that the idea of considering stability from the viewpoint of single-valued Lipschitzian behavior goes back to Robinson [47] being mainly motivated by applications to numerical algorithms in optimization.

Below, we give an equivalent description of full stability and show that indeed the single-valuedness property in Definition 2.1 can be dropped provided that we replace the Lipschitz continuity with an appropriate Lipschitzian concept for set-valued mappings. To this end, recall that $F\colon\mathbb{R}^n\rightrightarrows\mathbb{R}^m$ is *Lipschitz-like* (or has the Aubin property) around $(\bar{x},\bar{y})\in F$ with modulus $l\ge 0$ if there are neighborhoods U of $\bar{x}$ and V of $\bar{y}$ such that

$$F(x)\cap V\subset F(u)+l\|x-u\|I\!B\ \text{ whenever }\ x,u\in U. \tag{2.17}$$

Recall also that $F\colon\mathbb{R}^n\rightrightarrows\mathbb{R}^n$ is *monotone* provided that

$$\langle y_2-y_1,x_2-x_1\rangle\ge 0\ \text{ for all }\ (x_1,y_1),\ (x_2,y_2)\in F.$$

Theorem 2.2 (equivalent description of full stability). *Definition* 2.1 *can be equivalently reformulated by replacing the single-valuedness and Lipschitz continuity of the solution map* M_γ *around* $(\bar{w},\bar{v})$ *by the Lipschitz-like property of this mapping around* $(\bar{w},\bar{v},\bar{x})$.

Proof. By the Lipschitz-like property of M_γ around $(\bar{w},\bar{v},\bar{x})$, find a neighborhood triple

(W,V,U) for $(\bar{w},\bar{v},\bar{x})$ such that (2.17) holds for M_γ in this notation. Fix $w\in W$ and define

$$\varphi_w(\cdot)=\varphi(\cdot,w),\quad \widetilde{\varphi}_w:=\varphi_w+\delta_{I\!B_\gamma(\bar{x})},\quad g_w:=\widetilde{\varphi}^*_w,$$

where the latter stands for the conjugate of $\widetilde{\varphi}_w$. Thus g_w is convex and being expressed as

$$g_w(v)=\max_{x\in I\!B_\gamma(\bar{x})}\Big\{\langle v,x\rangle-\varphi_w(x)\Big\},\quad v\in\mathbb{R}^n. \tag{2.18}$$

Considering further the set-valued mapping $T_w\colon\mathbb{R}^n\rightrightarrows\mathbb{R}^n$ defined by $T_w(v):=M_\gamma(w,v)$ on V and $T_w(v):=\emptyset$ otherwise, we claim that it is monotone. Indeed, pick $x_i\in T_w(v_i)$ with $v_i\in V$ as $i=1,2$ and get from (2.18) the relationships

$$\begin{aligned}\langle x_1-x_2,v_1-v_2\rangle &= \langle x_1,v_1\rangle-\langle x_2,v_1\rangle-\langle x_1,v_2\rangle+\langle x_2,v_2\rangle\\ &= \big[g_w(v_1)-\langle x_2,v_1\rangle+\varphi_w(x_2)\big]+\big[g_w(v_2)-\langle x_1,v_2\rangle+\varphi_w(x_1)\big]\ge 0,\end{aligned}$$

which verify the claimed monotonicity of $T_{\bar{w}}$. Since the mapping $v\mapsto T_{\bar{w}}(v)=M_\gamma(\bar{w},v)$ is Lipschitz-like (i.e., surely lower semicontinuous) around $(\bar{v},\bar{x})$, the classical Kenderov theorem [16] tells us that $T_{\bar{w}}(\bar{v})=\{\bar{x}\}$, and hence $M_\gamma(\bar{w},\bar{v})=\{\bar{x}\}$. The same arguments work for any $(w,v)\in W\times V$ justifying therefore the single-valuedness of M_γ on $W\times V$. $\triangle$

To formulate the main result of [20] on characterizing full stability of local minimizers in problem $\mathcal{P}(\bar{w},\bar{v})$ with an extended-real-valued φ in finite dimensions, we need to recall the following important notions of variational analysis; cf. [20, 44, 53] for more details. A lower semicontinuous function $\varphi(x,w)$ is *prox-regular* in x at $\bar{x}$ for $\bar{v}$ with *compatible parameterization* by w at $\bar{w}$ if $\bar{v}\in\partial_x\varphi(\bar{x},\bar{w})$ and there exist neighborhoods U of $\bar{x}$, W of $\bar{w}$, and V of $\bar{v}$ together with numbers $\varepsilon>0$ and $\gamma\ge 0$ such that

$$\begin{aligned}&\varphi(u,w)\ge\varphi(x,w)+\langle v,u-x\rangle-\frac{\gamma}{2}\|u-x\|^2\ \text{ for all }\ u\in U\\ &\text{when }\ v\in\partial_x\varphi(x,w)\cap V,\ x\in U,\ w\in W,\ \varphi(x,w)\le\varphi(\bar{x},\bar{w})+\varepsilon.\end{aligned} \tag{2.19}$$

Furthermore, $\varphi(x,w)$ is called to be *subdifferentially continuous* at $(\bar{x},\bar{w},\bar{v})$ if it is continuous as a function of (x,w,v) on the partial subdifferential graph $\partial_x\varphi$ at this point. If both of these properties hold simultaneously, we say that φ is *continuously prox-regular* in x at $\bar{x}$ for $\bar{v}$ with *compatible parameterization* by w at $\bar{w}$, or simply that this function is *parametrically continuously prox-regular* at $(\bar{x}, \bar{w}, \bar{v})$.

It is known from [20] that the class of parametrically continuously prox-regular functions $\varphi\colon \mathbb{R}^n \times \mathbb{R}^d \to \overline{\mathbb{R}}$ at $(\bar{x}, \bar{w}, \bar{v})$ with $\bar{v} \in \partial_x \varphi(\bar{x}, \bar{w})$ is fairly large including, in particular, all extended-real-valued functions $\varphi(x, w)$ that are *strongly amenable* in x at $\bar{x}$ with *compatible parametrization* by w at $\bar{w}$ in the following sense: There are $\Phi\colon \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}^m$ and $\theta\colon \mathbb{R}^m \to \overline{\mathbb{R}}$ such that $\varphi(x, w) = \theta(\Phi(x, w))$ and Φ is $\mathcal{C}^2$ around $(\bar{x}, \bar{w})$ while θ is convex, proper, l.s.c., and finite at $\Phi(\bar{x}, \bar{w})$ under the first-order qualification condition

$$\partial^\infty \theta\big(\Phi(\bar{x}, \bar{w})\big) \cap \ker \nabla_x \Phi(\bar{x}, \bar{w})^* = \{0\}. \tag{2.20}$$

The parametric continuous prox-regularity of such functions is proved in [20, Proposition 2.2], where it is shown in addition that the parametric strong amenability of φ formulated above ensures the validity of the *basic constraint qualification*:

$$(0, q) \in \partial^\infty \varphi(\bar{x}, \bar{w}) \Longrightarrow q = 0. \tag{2.21}$$

The strong amenability property and its parametric expansion hold not only in the obvious cases of $\mathcal{C}^2$ and convex functions but in dramatically larger frameworks typically encountered in finite-dimensional variational analysis and optimization; see [20,21,45,53]. The main result of [20, Theorem 2.3] is as follows.

Theorem 2.3 (characterization of full stability in unconstrained extended-real-valued format). *Let $\bar{x}$ be a feasible solution to the unperturbed problem $\mathcal{P}(\bar{w}, \bar{v})$ in* (2.14) *at which the first-order necessary optimality condition $\bar{v} \in \partial_x \varphi(\bar{x}, \bar{w})$ and the basic constraint qualification* (2.21) *are satisfied. Assume in addition that φ is parametrically continuously prox-regular at $(\bar{x}, \bar{w}, \bar{v})$. Then $\bar{x}$ is a fully stable locally optimal solution to $\mathcal{P}(\bar{x}, \bar{w})$ if and* *only if the following second-order conditions hold:*

$$(0,q) \in D^*(\partial_x\varphi)(\bar{x},\bar{w},\bar{v})(0) \Longrightarrow q = 0, \tag{2.22}$$

$$[(p,q) \in D^*(\partial_x\varphi)(\bar{x},\bar{w},\bar{v})(u),\ u \neq 0] \Longrightarrow \langle p,u\rangle > 0 \tag{2.23}$$

via the extended second-order subdifferential mapping (2.12).

In the subsequent chapters we employ Theorem 2.3 to obtain verifiable necessary and sufficient conditions for full stability of local minimizers in favorable classes of constrained optimization problems in terms of the problem data. Achieving it requires the implementation and development of *second-order subdifferential calculus* as well as *precise calculating* the partial second-order subdifferential constructions for the corresponding functions involved.

2.3 Characterizations of Full Stability

We proceed in this section with establishing useful relationships between *full stability* of local minimizers in the unconstrained format of (2.14) with an extended-real-valued function $\varphi(x,w)$ and an appropriate version of the so-called "strong metric regularity" of the partial subdifferential mapping $\partial_x\varphi$. Recall [5] that a set-valued mapping $F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is *strongly metrically regular* at $(\bar{x},\bar{y}) \in F$ if the inverse mapping F^{-1} admits a Lipschitzian single-valued *localization* around $(\bar{x},\bar{y})$, i.e., there are neighborhood U of $\bar{x}$ and V of $\bar{y}$ and a single-valued Lipschitz continuous mapping $f\colon V \to U$ such that $f(\bar{y}) = \bar{x}$ and $F^{-1}(y)\cap U = \{f(y)\}$ for all $y \in V$. This notion is an abstract version of *Robinson's strong regularity* for variational inequalities and nonlinear programming problems [47].

Close relationships (equivalences under appropriate constraint qualifications) between *tilt* stability and strong regularity have been established by Mordukhovich and Rockafellar [38] and Mordukhovich and Outrata [34] in the framework of nonlinear programming and by Lewis and Zhang [22] and Drusvyatskiy and Lewis [6] via strong metric regularity of subdifferential mappings for extended-real-valued objective functions in the general unconstrained format of nonparametric optimization. Based on [20], we now extend the latter results to the parametric framework of (2.14) while establishing the equivalence between *full* stability of locally optimal solutions to (2.14) and an appropriate notion of *partial* strong metric regularity for the corresponding partial subdifferential mapping of the function $\varphi(x,w)$ therein. We also establish characterizations of these notions via a certain partial *second-order growth* condition.

Given a function $\varphi\colon \mathbb{R}^n\times\mathbb{R}^d\to\overline{\mathbb{R}}$, consider its partial first-order subdifferential mapping $\partial_x\varphi\colon \mathbb{R}^n\times\mathbb{R}^d\to\mathbb{R}^n$ and define the *partial inverse* of $\partial_x\varphi$ by

$$S(w,v):=\Big\{x\in\mathbb{R}^n\big|\;v\in\partial_x\varphi(x,w)\Big\},\tag{2.24}$$

where the subdifferential is understood in the basic sense (2.2).

Definition 2.4 (partial strong metric regularity). *Given $(\bar x,\bar w)\in\operatorname{dom}\varphi$ and $\bar v\in\partial_x\varphi(\bar x,\bar w)$, we say that the partial subdifferential mapping $\partial_x\varphi\colon\mathbb{R}^n\times\mathbb{R}^d\rightrightarrows\mathbb{R}^n$ is* PARTIALLY STRONGLY METRICALLY REGULAR (*abbr. PSMR*) *at $(\bar x,\bar w,\bar v)$ if its partial inverse* (2.24) *admits a Lipschitzian single-valued localization around this point.*

Note that the notion introduced in Definition 2.4 is different from the (total) strong metric regularity of $\partial_x\varphi$ at $(\bar x,\bar w,\bar v)$ discussed above, since its concerns Lipschitzian localizations of the *partial* inverse S instead of the inverse mapping $(\partial_x\varphi)^{-1}$.

Theorem 2.5 (full stability versus partial strong metric regularity). *Given a function $\varphi\colon\mathbb{R}^n\times\mathbb{R}^d\to\overline{\mathbb{R}}$ with $(\bar x,\bar w)\in\operatorname{dom}\varphi$, consider the unperturbed problem $\mathcal{P}(\bar w,\bar v)$ in* (2.14) *with some $\bar v\in\partial_x\varphi(\bar x,\bar w)$ and let $\bar x$ be a locally optimal solution to $\mathcal{P}(\bar w,\bar v)$, i.e., $\bar x\in M_\gamma(\bar w,\bar v)$ for some number $\gamma>0$ in* (2.16). *Assume that the basic constraint qualification* (2.21) *is satisfied at $(\bar x,\bar w)$. The following assertions hold:*

(i) *If $\partial_x\varphi$ is PSMR at $(\bar x,\bar w,\bar v)$, then $\bar x$ is a fully stable local minimizer for $\mathcal{P}(\bar w,\bar v)$ and the function φ is prox-regular in x at $\bar x$ with compatible parameterization by w at $\bar w$.*

(ii) *Conversely, if φ is parametrically continuously prox-regular at $(\bar{x},\bar{w},\bar{v})$ and if $\bar{x}$ is a fully stable local minimizer for $\mathcal{P}(\bar{w},\bar{v})$, then $\partial_x\varphi$ is PSMR at $(\bar{x},\bar{w},\bar{v})$.*

Proof. To justify assertion (i), assume that the partial subdifferential mapping $\partial_x\varphi$ is PSMR at $(\bar{x},\bar{w},\bar{v})$ and fix the number $\gamma>0$ from the formulation of the theorem. Then it follows from Definition 2.4 that there exist neighborhoods U of $\bar{x}$, V of $\bar{v}$, and W of $\bar{w}$ such that for all $(w,v)\in W\times V$ the localization $S(w,v)\cap U$ is single-valued. Without loss of generality suppose $I\!B_\gamma(\bar{x})\subset U$. We claim that

$$M_\gamma(\bar{w},\bar{v})=\{\bar{x}\}.$$

Indeed, by the stationary condition in (2.14) and the assumed PSMR property we have

$$\varphi_0(\bar{x},\bar{w})-\langle\bar{v},\bar{x}\rangle<\varphi_0(x,\bar{w})-\langle\bar{v},x\rangle\ \text{ for all }\ x\in\operatorname{int}I\!B_\gamma(\bar{x}). \tag{2.25}$$

If there is $\widetilde{x}\in M_\gamma(\bar{w},\bar{v})$ with $\widetilde{x}\neq\bar{x}$ and $\|\widetilde{x}-\bar{x}\|=\gamma$, then replacing $M_\gamma(\bar{w},\bar{v})$ by $M_{\gamma/2}(\bar{w},\bar{v})$ gives us $M_{\gamma/2}(\bar{w},\bar{v})=\{\bar{x}\}$. Thus we can suppose that $M_\gamma(\bar{w},\bar{v})=\{\bar{x}\}$. Invoking now the basic constraint qualification (2.21) and employing [20, Proposition 3.5] ensure the Lipschitz continuity around $(\bar{w},\bar{v})$ of the optimal value function m_γ from (2.15) and allow us to find $\eta>0$ with

$$M_\gamma(w,v)\subset\operatorname{int}I\!B_\gamma(\bar{x})\ \text{ whenever }\ (w,v)\in\operatorname{int}I\!B_\eta(\bar{w})\times\operatorname{int}I\!B_\eta(\bar{v}).$$

Thus we have under the assumptions made that

$$M_\gamma(w,v)\subset S(w,v)\cap\operatorname{int}I\!B_\gamma(\bar{x})\ \text{ for all }\ (w,v)\in\operatorname{int}I\!B_\eta(\bar{w})\times\operatorname{int}I\!B_\eta(\bar{v}), \tag{2.26}$$

which in fact holds as *equality* by the single-valuedness of the right-hand side and the nonemptiness of the left-hand one, implying hence that M_γ is single-valued and Lipschitz continuous around $(\bar{w},\bar{v})$. This means that $\bar{x}$ is a fully stable local minimizer of $\mathcal{P}(\bar{w},\bar{v})$ by Definition 2.1.

To complete the proof of assertion (i), it remains to justify the claimed parametric prox-regularity of φ at $(\bar{x},\bar{w})$. Take any $x \in \operatorname{int} I\!B_\gamma(\bar{x})$, $w \in \operatorname{int} I\!B_\eta(\bar{w})$, and $v \in \partial_x\varphi(x,w) \cap \operatorname{int} I\!B_\eta(\bar{v})$ with the positive numbers γ,η found above. Then $x \in M_\gamma(w,v)$ by the equality in (2.26), and thus we get from the construction of M_γ in (2.16) that

$$\varphi(u,w) \ge \varphi(x,w) + \langle v, u-x\rangle \quad \text{whenever} \quad u \in \operatorname{int} I\!B_\gamma(\bar{x}),$$

which obviously implies by (2.19) the desired parametric prox-regularity of φ.

To justify assertion (ii), observe that it follows from the second part of [20, Theorem 2.3] that (2.26) holds as equality with some numbers $\gamma,\eta > 0$ provided that φ is parametrically continuously prox-regular at $(\bar{x},\bar{w},\bar{v})$. Since $\bar{x}$ is now assumed to be a fully stable local minimizer in $\mathcal{P}(\bar{w},\bar{v})$, this ensures the single-valued Lipschitzian localization of S_φ around $(\bar{w},\bar{v},\bar{x})$ and thus justifies the PSMR property of the partial subdifferential mapping $\partial_x\varphi$ at $(\bar{x},\bar{w},\bar{v})$. $\triangle$

Next we derive *necessary and sufficient* conditions for PSMR from Definition 2.4 and full stability properties in the case of general extended-real-valued functions via a partial version of the so-called *uniform second-order (quadratic) growth condition.*

Definition 2.6 (uniform second-order growth condition). *Given $\varphi\colon \mathbb{R}^n \times \mathbb{R}^d \to \overline{\mathbb{R}}$ finite at $(\bar{x},\bar{w})$ and given a partial subgradient $\bar{v} \in \partial_x\varphi(\bar{x},\bar{w})$, we say that the* UNIFORM SECOND-ORDER GROWTH CONDITION *(abbr. USOGC) holds for φ at $(\bar{x},\bar{w},\bar{v})$ if there exist a constant $\eta > 0$ and neighborhoods U of $\bar{x}$, W of $\bar{w}$, and V of $\bar{v}$ such that for any $(w,v) \in W \times V$ there is a point $x_{wv} \in U$ (necessarily unique) satisfying $v \in \partial_x\varphi(x_{wv},w)$ and*

$$\varphi(u,w) \ge \varphi(x_{wv},w) + \langle v, u-x_{wv}\rangle + \eta\|u-x_{wv}\|^2 \quad \textit{whenever} \quad u \in U. \tag{2.27}$$

Note that for problems of conic programming with $\mathcal{C}^2$ data this notion appeared in a different while equivalent form in [1, Definition 5.16] as the “uniform second-order (quadratic) growth condition with respect to the $\mathcal{C}^2$-smooth parameterization." Its version “with respect to the tilt parameterization" was employed in [1, Theorem 5.36] for characterizing tilt-stable minimizers of conic programs and then in [22, Theorem 6.3] and [6, Theorem 3.3] in more general settings of extended-real-valued functions.

Let us employ USOGC from Definition 2.6 to characterize fully stable local minimizer of $\mathcal{P}(\bar{w},\bar{v})$. To achieve this goal, we use the following lemma obtained in [20, Lemma 5.2].

Lemma 2.7 (uniform second-order growth for convex functions). *Let $f\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ be a proper, l.s.c., and convex function whose conjugate f^* is differentiable on* $\mathrm{int}\mathbb{B}_\gamma(\bar{v})$ *for some $\bar{v} \in \mathbb{R}^n$ and $\gamma > 0$, and let the gradient of f^* be Lipschitz continuous on* $\mathrm{int}\mathbb{B}_\gamma(\bar{v})$ *with constant $\sigma > 0$. Then for any $(x,v) \in (\partial f) \cap [\mathrm{int}\mathbb{B}_{\frac{\sigma\gamma}{4}}(\bar{x}) \times \mathrm{int}\,\mathbb{B}_{\frac{\gamma}{2}}(\bar{v})]$ with $\bar{x} \in \mathbb{R}^n$ we have*

$$f(u) \geq f(x) + \langle v, u-x\rangle + \frac{1}{2\sigma}\|u-x\|^2 \ \ \textit{whenever} \ \ u \in \mathbb{B}_{\frac{\sigma\gamma}{4}}(\bar{x}). \tag{2.28}$$

Proof. Consider the open set $O := \{v \in \mathbb{R}^n |\ \mathbb{B}_{\frac{\gamma}{2}}(v) \subset \mathrm{int}\mathbb{B}_\gamma(\bar{v})\}$. Then by [20, Lemma 5.2] for all $v \in \partial f(x) \cap O$ we get the estimate

$$f(u) \geq f(x) + \langle v, u-x\rangle + \frac{1}{2\sigma}\|u-x\|^2 \ \text{ whenever } \ \|u-x\| \leq \frac{\gamma\sigma}{2},$$

which implies (2.28) for the corresponding pairs (x,v). $\triangle$

Recall that a mapping $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is *local maximal monotone* around $(\bar{x},\bar{v})$, if there exist neighborhoods $U \times V$ of $(\bar{x},\bar{v})$ such that every monotone mapping $S : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ with $(T) \cap (U \times V) \subset S$ satisfies $(T) \cap (U \times V) = (S) \cap (U \times V)$.

Theorem 2.8 (relationships between full stability and uniform second-order growth). *Let $\varphi\colon \mathbb{R}^n \times \mathbb{R}^d \to \overline{\mathbb{R}}$ be l.s.c. with $\bar{v} \in \partial_x\varphi(\bar{x},\bar{w})$ for some $(\bar{x},\bar{w}) \in \mathrm{dom}\,\varphi$. The following assertions hold:*

(i) *If $\bar{x}$ is a fully stable local minimizer of the unperturbed problem $\mathcal{P}(\bar{w},\bar{v})$ in* (2.14) *and the basic constraint qualification* (2.21) *is satisfied at $(\bar{x},\bar{w})$, then USOGC of Definition* 2.6 *holds at $(\bar{x},\bar{w},\bar{v})$.*

(ii) *Conversely, assume that φ is parametrically continuously prox-regular at $(\bar{x},\bar{w},\bar{v})$ and that USOGC holds at this point with the mapping $(w,v) \mapsto x_{wv}$ in Definition* 2.6 *being* *locally Lipschitzian around* $(\bar w,\bar v)$. *Then* $\partial_x\varphi$ *is PSMR at* $(\bar x,\bar w,\bar v)$.

Proof. To justify (i), let $\bar x$ be a fully stable locally optimal solution to problem $\mathcal{P}(\bar w,\bar v)$. Then there is a number $\gamma>0$ such that the mapping $(w,v)\mapsto M_\gamma(w,v)$ form (2.16) is single-valued and Lipschitz continuous on $\mathrm{int}\,I\!B_\gamma(\bar w)\times\mathrm{int}\,I\!B_\gamma(\bar v)$ with some constant $\sigma>0$. For any fixed $w\in\mathrm{int}\,I\!B_\gamma(\bar w)$ consider the function $\varphi_w(\cdot)=\varphi(\cdot,w)$ and define

$$\bar\varphi_w:=\varphi_w+\delta_{I\!B_\gamma(\bar x)},\quad g_w:=\bar\varphi_w^*,\ \text{ and }\ h_w:=g_w^*. \tag{2.29}$$

We easily get from (2.16) and the definition of g_w that

$$M_\gamma(w,v)=\mathrm{argmin}_{x\in I\!B_\gamma(\bar x)}\big\{\varphi(x,w)-\langle v,x\rangle\big\}\in\partial g_w(v)\ \text{ for }\ v\in\mathrm{int}\,I\!B_\gamma(\bar v). \tag{2.30}$$

Indeed, it follows from the constructions above the function g_w is convex and is expressed as

$$g_w(v)=\max_{x\in I\!B_\gamma(\bar x)}\big\{\langle v,x\rangle-\varphi_w(x)\big\}.$$

This readily implies the relationships

$$\begin{aligned} g_w(v')-g_w(v) &\ge \langle v',M_\gamma(w,v)\rangle-\varphi_w(M_\gamma(w,v))-\langle v,M_\gamma(w,v)\rangle+\varphi_w(M_\gamma(w,v))\\ &= \langle v'-v,M_\gamma(w,v)\rangle\ \text{ for all }\ v'\in\mathbb{R}^n, \end{aligned}$$

which yields (2.30) holds. Consider further the mapping $T_w(\cdot):=M_\gamma(w,\cdot)$ and show that it is monotone on $\mathrm{int}\,I\!B_\gamma(\bar v)$. To check it, pick $x_i\in T_w(v_i)$ with $v_i\in\mathrm{int}\,I\!B_\gamma(\bar v)$ as $i=1,2$ and get from (2.30) that

$$\begin{aligned} \langle x_1-x_2,v_1-v_2\rangle &= \langle x_1,v_1\rangle-\langle x_2,v_1\rangle-\langle x_1,v_2\rangle+\langle x_2,v_2\rangle\\ &= \big[g_w(v_1)-\langle x_2,v_1\rangle+\varphi_w(x_2)\big]+\big[g_w(v_2)-\langle x_1,v_2\rangle+\varphi_w(x_1)\big]\ge 0. \end{aligned}$$

Since T_w is (Lipschitz) continuous, it is locally maximal monotone around $(\bar v,\bar x)$ relative to $\mathrm{int}\,I\!B_\gamma(\bar v)\times\mathrm{int}\,I\!B_\gamma(\bar x)$; see [53, Example 12.7]. Remembering next that the subdifferential mappings for convex functions are also maximal monotone, we conclude from (2.30) that

$$(\partial h_w)^{-1}(v)=\partial g_w(v)=T_w(v)\ \text{ for all }\ v\in\mathrm{int}\,I\!B_\gamma(\bar v), \tag{2.31}$$

where the first equality is duo to the relation $h_w := g_w^*$. Thus g_w is Fréchet differentiable on $\text{int}\, I\!B_\gamma(\bar v)$ and its gradient mapping ∇g_w is Lipschitz continuous with constant σ on this set. Now we are in a position of applying Lemma 2.7 to the function $f := h_w$ with $h_w^* = g_w^{**} = g_w$. This gives us the estimate

$$h_w(u) \geq h_w(x) + \langle v, u - x\rangle + \frac{1}{2\sigma}\|u - x\|^2 \quad \text{whenever} \quad u \in \text{int}\, I\!B_{\frac{\sigma\gamma}{4}}(\bar x) \tag{2.32}$$

for all $(x, v) \in (\partial h_w) \cap [\text{int}\, I\!B_{\frac{\sigma\gamma}{4}}(\bar x) \times \text{int}\, I\!B_{\frac{\gamma}{2}}(\bar v)]$. Observe that, since the Lipschitz constant σ does not depend on the w, the estimate in (2.32) is uniform with respect to w in the selected neighborhood of $\bar w$. Also we can assume without loss of generality that $\text{int}\, I\!B_{\frac{\sigma\gamma}{4}}(\bar x) \subset \text{int}\, I\!B_{\frac{\gamma}{2}}(\bar x)$.

Take now $x \in (\partial h_w)^{-1}(v) = \partial g_w(v) = T_w(v)$ and get from the single-valuedness of the set $T_w(v)$ by its construction above that

$$h_w(T_w(v)) = h_w(x) = \varphi_w(x) = \varphi(x, w).$$

This allows us to deduce from (2.32) that

$$\varphi(u, w) \geq \varphi(x, w) + \langle v, u - x\rangle + \frac{1}{2\sigma}\|u - x\|^2 \tag{2.33}$$

whenever $(x, v) \in (\partial h_w) \cap [\text{int}\, I\!B_{\frac{\sigma\gamma}{4}}(\bar x) \times \text{int}\, I\!B_{\frac{\gamma}{2}}(\bar v)]$ and $u \in \text{int}\, I\!B_{\frac{\sigma\gamma}{4}}(\bar x)$. Since $\bar x$ is a full stable minimizer of (2.14), we get $\bar x \in M_\gamma(\bar w, \bar v)$ for some $\gamma < \frac{\sigma\gamma}{4}$. Taking into account the basic constraint qualification (2.21) together with [20, Proposition 3.5], we obtain $M_\gamma(w, v) \in \text{int}\, I\!B_\gamma(\bar x)$ for any $(w, v) \in \text{int}\, I\!B_\gamma(\bar w) \times \text{int}\, I\!B_\gamma(\bar v)$. Thus for any $(w, v) \in \text{int}\, I\!B_\gamma(\bar w) \times \text{int}\, I\!B_\gamma(\bar v)$, we can find $x_{wv} \colon = M_\gamma(w, v) \in \text{int}\, I\!B_\gamma(\bar x)$. Remember that $(\partial g_w) \cap [\text{int}\, I\!B_\gamma(\bar x) \times \text{int}\, I\!B_\gamma(\bar v)] = (\partial T_w) \cap [\text{int}\, I\!B_\gamma(\bar x) \times \text{int}\, I\!B_\gamma(\bar v)]$. Since $(v, x_{wv}) \in (\partial g_w) \cap [\text{int}\, I\!B_\gamma(\bar x) \times \text{int}\, I\!B_\gamma(\bar v)]$, we therefore get

$$x_{wv} = T_w(v) = \partial g_w(v) = (\partial h_w)^{-1}(v),$$

and hence $(x_{wv}, v) \in (\partial h_w) \cap [\text{int}\, I\!B_{\frac{\sigma\gamma}{4}}(\bar x) \times \text{int}\, I\!B_{\frac{\gamma}{2}}(\bar v)]$. Letting $x \colon = x_{wv}$ in (2.32) justifies validity of USOGC for φ at $(\bar{x}, \bar{w}, \bar{v})$ and thus ends the proof of (i).

Next we justify assertion (ii) observing by Theorem 2.5 that it suffices to show that the mapping $\partial_x \varphi$ is PSMR at $(\bar{x}, \bar{w}, \bar{v})$ under the assumptions made. To proceed, fix the neighborhoods U of $\bar{x}$, W of $\bar{w}$, and V of $\bar{v}$ for which the second-order growth condition (2.27) holds and thus gives us the single-valued and Lipschitz continuous mapping $s : W \times V \to U$ defined by $s(w, v) := x_{wv}$. Denote $T_w(\cdot) := s(w, \cdot)$ and pick any vectors $v_i \in T_w^{-1}(x_i)$ with $v_i \in V$ and $x_i \in U$ for $i = 1, 2$. By (2.27) with $\eta = (2\sigma)^{-1}$ for some $\sigma > 0$ we get the estimates

$$\begin{aligned} \varphi(x_2, w) &\geq \varphi(x_1, w) + \langle v_1, x_2 - x_1 \rangle + \frac{1}{2\sigma}\|x_2 - x_1\|^2, \\ \varphi(x_1, w) &\geq \varphi(x_2, w) + \langle v_2, x_1 - x_2 \rangle + \frac{1}{2\sigma}\|x_2 - x_1\|^2, \end{aligned}$$

which tell us that the mapping T_w^{-1} is locally *strongly monotone* with constant σ^{-1}; see [53, Definition 12.53]. Hence T_w is locally monotone relative to V and U and in fact is locally maximal monotone relative to $V \times U$ due to its continuity. Note that if $(v, x) \in T_w$, then $v \in \partial \varphi_w(x)$.

Let $F_w : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be the mapping for which F_w^{-1} is the intersection of $\partial \varphi_w$ and $U \times V$. We have $T_w \subset F_w$ and thus get the inclusions

$$T_w^{-1}(x) \subset F_w^{-1}(x) \subset \partial \varphi_w(x) \quad \text{whenever} \quad x \in U. \tag{2.34}$$

It follows from the parametric continuous prox-regularity of φ that the mapping $\partial \varphi_w$ are locally *hypomonotone* whenever $w \in W$ with the same constant r from (2.19), and so the mapping $F_w^{-1} + tI$ is locally strongly monotone with constant $t - r$ for any fixed $t > r$; see [53, Example 12.28]. Since T_w^{-1} is locally strongly monotone with constant σ^{-1}, we keep this property for the mapping $T_w^{-1} + tI$ with constant $\sigma^{-1} + t$. Hence the mappings $(F_w^{-1} + tI)^{-1}$ and $(T_w^{-1} + tI)^{-1}$ are *single-valued* on their domains. Furthermore, it follows from (2.34) that $(T_w^{-1}+tI)^{-1}\subset(F_w^{-1}+tI)^{-1}$. Assume $\gamma>0$ is small enough such that

$$\mathbb{B}_\gamma(\bar{x})\subset U,\quad \mathbb{B}_{\frac{t\gamma}{4}}(\bar{v})\subset V,\quad s(W\times V)\subset \operatorname{int}\mathbb{B}_{\frac{\gamma}{4}}(\bar{x}), \tag{2.35}$$

otherwise we can shrink the neighborhoods W and V. Let $U' := \operatorname{int}\mathbb{B}_\gamma(\bar{x})$ and $O := J_t(V\times U')$ where J is the bilinear mapping defined by $J_t(v,u)=(v+tu,u)$. Therefore O is a neighborhood of $(\bar{v}+t\bar{x},\bar{x})$ due to open mapping theorem. Without loss of generality, we can assume further that $\mathbb{B}_{t\gamma}(\bar{x}+t\bar{v})\times\mathbb{B}_\gamma(\bar{x})\subset O$. Employing now (4.63) yields

$$(\partial h_w+tI)^{-1}\cap O=(T_w^{-1}+tI)^{-1}\cap O. \tag{2.36}$$

Let $w\in W$ be arbitrary and fixed and then define the function $p_w(x) := h_w(x)+\frac{t}{2}\|x\|^2$ when $x\in\mathbb{B}_\gamma(\bar{x})$ and ∞ otherwise, where h_w is defined by (2.29). Therefore p_w is proper, l.s.c, and strongly convex with modulus t, which assures us that p_w^* is a proper, l.s.c, and convex function. Appealing now to [53, Proposition 12.60] confirms that the conjugate function p_w^* is differentiable on $\mathbb{R}^n$ and its gradient mapping is also Lipschitz continuous with constant $\frac{1}{t}$. We claim now that for any $\widehat{v}\in\operatorname{int}\mathbb{B}_{\frac{t\gamma}{8}}(\bar{v}+t\bar{x})$, we can find a *unique* $\widehat{x}\in\operatorname{int}\mathbb{B}_\gamma(\bar{x})$ with $\widehat{x}=\nabla p_w^*(\widehat{v})$. To this aim, assume that $\widehat{v}\in\operatorname{int}\mathbb{B}_{\frac{t\gamma}{8}}(\bar{v}+t\bar{x})$. Since p_w^* is differentiable at $\widehat{v}$, we can choose $\widehat{x} := \nabla p_w^*(\widehat{v})$. We need to show that $\widehat{x}\in\operatorname{int}\mathbb{B}_\gamma(\bar{x})$. Using convexity of p_w^* allows us to obtain

$$\widehat{x}=\nabla p_w^*(\widehat{v})=\partial p_w^*(\widehat{v})=\operatorname{argmin}_{x\in\mathbb{B}_\gamma(\bar{x})}\Big\{p_w(x)-\langle\widehat{v},x\rangle\Big\}, \tag{2.37}$$

which in turn implies that $\widehat{x}=\nabla p_w^*(\widehat{v})$ if and only if $0\in\partial p_w(\widehat{x})-\widehat{v}+N_{\mathbb{B}_\gamma(\bar{x})}(\widehat{x})$. The latter inclusion amounts to $\widehat{x}\in\mathbb{B}_\gamma(\bar{x})$ and $\widehat{v}-\alpha(\widehat{x}-\bar{x})\in\partial p_w(\widehat{x})$ for some $\alpha\geq 0$ with $\alpha(\widehat{x}-\bar{x})\in N_{\mathbb{B}_\gamma(\bar{x})}(\widehat{x})$. Remember that $\widehat{x}\in\mathbb{B}_\gamma(\bar{x})$. We claim now that $\|\widehat{x}-\bar{x}\|<\gamma$. By contradiction, we assume that $\|\widehat{x}-\bar{x}\|=\gamma$. Pick $x'\in\mathbb{B}_{\frac{\gamma}{8}}(\bar{x})$, which in turn verifies that $(\widehat{v},x')\in O$. It is easy to see $J_t(v,x')=(\widehat{v},x')$ with $v := \widehat{v}-tx'$. Using the above assumptions, we obtain

$$\|v-\bar v\| \le \|\widehat v-(\bar v+t\bar x)\| + t\|x'-\bar x\| \le \frac{t\gamma}{8}+\frac{t\gamma}{8}=\frac{t\gamma}{4}, \tag{2.38}$$

and hence $v \in I\!B_{\frac{t\gamma}{4}}(\bar v) \subset V$ due (2.35). Since we have $(w,v)\in W\times V$, we find x_{wv} so that $s(w,v)=x_{wv}$, which demonstrates that $x_{wv}\in I\!B_{\frac{\gamma}{4}}(\bar x)$. Applying now (4.63) confirms that $v\in\partial h_w(x_{wv})$, which brings us to $v+tx_{wv}\in\partial p_w(x_{wv})$ due to the first order calculus rule for convex functions. Recalling that ∂p_w is t-strongly monotone and $\widehat v-\alpha(\widehat x-\bar x)\in\partial p_w(\widehat x)$, we deduce that

$$\Big\langle \widehat v-\alpha(\widehat x-\bar x)-(v+tx_{wv}),\widehat x-x_{wv}\Big\rangle \ge t\|\widehat x-x_{wv}\|^2,$$

which in turn leads us to

$$\Big\langle \widehat v-(v+tx_{wv}),\widehat x-x_{wv}\Big\rangle + \Big\langle -\alpha(x_{wv}-\bar x),\widehat x-x_{wv}\Big\rangle \ge (t+\alpha)\|\widehat x-x_{wv}\|^2. \tag{2.39}$$

Taking now into account (2.35) together with (2.39), we come up to

$$\begin{aligned}\|\widehat v-(v+tx_{wv})\| &\ge (t+\alpha)\|\widehat x-x_{wv}\|-\alpha\|x_{wv}-\bar x\|\\ &\ge (t+\alpha)\Big[\|\widehat x-\bar x\|-\|\bar x-x_{wv}\|\Big]-\tfrac{\alpha\gamma}{4}\\ &\ge (t+\alpha)\Big[\gamma-\tfrac{\gamma}{4}\Big]-\tfrac{\alpha\gamma}{4}=\tfrac{3t\gamma}{4}+\tfrac{\alpha\gamma}{2}.\end{aligned} \tag{2.40}$$

On the other hand, appealing to (2.35) and (2.38) tells us that

$$\begin{aligned}\|\widehat v-(v+tx_{wv})\| &\le \|\widehat v-(\bar v+t\bar x)\|+\|v-\bar v\|+t\|\bar x-x_{wv}\|\\ &< \tfrac{t\gamma}{8}+\tfrac{2t\gamma}{8}+\tfrac{2t\gamma}{8}=\tfrac{5t\gamma}{8},\end{aligned} \tag{2.41}$$

which clearly contradicts (2.40) and thus proves the claim. Remember that $(F_w^{-1}+tI)^{-1}$ is single-valued on its domain. Using the justified claim together with (2.36), we get

$$(\partial h_w+tI)^{-1}\cap O'=(T_w^{-1}+tI)^{-1}\cap O'=(F_w^{-1}+tI)^{-1}\cap O' \tag{2.42}$$

with $O' :\ = \mathrm{int}\, I\!B_{\frac{t\gamma}{8}}(\bar x)\times \mathrm{int}\, I\!B_{\gamma}(\bar x)\subset O$. Therefore we arrive at

$$(\partial h_w)^{-1}\cap J_t^{-1}(O')=T_w\cap J_t^{-1}(O')=(\partial\varphi_w)^{-1}\cap J_t^{-1}(O'),\quad w\in W. \tag{2.43}$$

Since $J_t^{-1}(O')$ is a neighborhood of $(\bar{v}, \bar{x})$, there exists $\beta > 0$ such that $\text{int}\, I\!B_\beta(\bar{v}) \times \text{int}\, I\!B_\beta(\bar{x}) \subset J_t^{-1}(O')$. Recalling finally definition (2.24) of the partial inverse S, we easily deduce from (2.43) that

$$S(w,v) \cap \text{int}\, I\!B_\beta(\bar{x}) = \{s(w,v)\} \text{ whenever } (w,v) \in W \times \text{int}\, I\!B_\beta(\bar{v})$$

for the mapping s defined at the beginning of the proof of (ii). This means that s is a Lipschitzian single-valued localization of S, and thus $\partial_x \varphi$ is PSMR at $(\bar{x}, \bar{w}, \bar{v})$ by Definition 2.4. △

Finally, we apply Theorem 2.3 to derive necessary and sufficient conditions for full stability in *composite models* of optimization written in the form

$$\text{minimize } \varphi(x) := \varphi_0(x) + \theta(\varphi_1(x), \ldots, \varphi_m(x)) = \varphi_0(x) + \theta(\Phi(x)) \text{ over } x \in \mathbb{R}^n, \tag{2.44}$$

where $\varphi_0\colon \mathbb{R}^n \to \overline{\mathbb{R}}$, $\theta\colon \mathbb{R}^m \to \overline{\mathbb{R}}$, and $\Phi(x) := (\varphi_1(x), \ldots, \varphi_m(x))$ is a mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$. Written in the unconstrained form, problem (2.44) is actually a problem of constrained optimization with the cost function φ_0 and the set of feasible solutions given by

$$X := \{x \in \mathbb{R}^n |\; (\varphi_1(x), \ldots, \varphi_m(x)) \in Z\} \text{ with } Z := \{z \in \mathbb{R}^m |\; \theta(z) < \infty\}.$$

Observe that the results presented in this section for problem (2.44) can be easily transferred to problem of this type with additional geometric constraints given by $x \in \Omega$ via a polyhedral set $\Omega \subset \mathbb{R}^n$. Indeed the only change needed to be done is replacing the mapping Φ in (2.44) by $x \mapsto (x, \varphi_1(x), \ldots, \varphi_m(x))$ and the set Z above by the convex polyhedron $\Omega \times Z$. As discussed in [52, 53], the composite format (2.44) is a general and convenient framework, from both theoretical and computational viewpoints, to accommodate a variety of particular models in constrained optimization. Note that conventional nonlinear programs with s inequality constraints and $m - s$ equality constraints can be written in form

$$\text{minimize } \varphi_0(x) + \delta_Z(\Phi(x)) \text{ over } x \in \mathbb{R}^n \tag{2.45}$$

via the indicator functions of the set $Z = \mathbb{R}^s_- \times \{0\}^{m-s}$.

Consider now the *fully perturbed* version $\mathcal{P}(w,v)$ of (2.44) with two parameters $(w,v) \in \mathbb{R}^d \times \mathbb{R}^n$ standing, respectively, for *basic* and *tilt* perturbations:

$$\text{minimize } \ \varphi(x,w) - \langle v, x\rangle \ \text{ over } \ x \in \mathbb{R}^n \ \text{ with } \ \varphi(x,w) := \varphi_0(x,w) + (\theta \circ \Phi)(x,w) \tag{2.46}$$

and $\Phi(x,w) = (\varphi_1(x,w),\ldots,\varphi_m(x,w))$ defined on $\mathbb{R}^n \times \mathbb{R}^d$. Our characterization of full stability for problem (2.45) utilizes the exact chain rule for the extended second-order subdifferential obtained in [38, Theorem 3.1] under the *full rank* condition (2.47) on Φ. For simplicity we suppose that the all the functions φ_i for $i = 0,\ldots,m$ are $\mathcal{C}^2$ around the reference points, although it is sufficient to assume that φ_i are merely smooth with strictly differentiable derivatives. Observe also that such properties are sometimes needed only *partially* with respect to the decision variable x; see the formulations and proofs below. It is worth noting that in the next theorem we use the second-order subdifferential of $\theta = \theta(z)$ and the special form (2.13) of the *extended* partial second-order subdifferential for the $\mathcal{C}^2$ functions $\varphi_i = \varphi_i(x,w)$.

Theorem 2.9 (characterizing fully stable local minimizers for composite problems under full rank condition). *Let $\bar{x}$ be a feasible solution to the unperturbed problem $\mathcal{P}(\bar{w},\bar{v})$ in (2.46) with some $\bar{w} \in \mathbb{R}^d$ and $\bar{v} \in \partial_x\varphi(\bar{x},\bar{w})$, where $\varphi_0, \Phi \in \mathcal{C}^2$ around $(\bar{x},\bar{w})$ under the validity of the full rank condition*

$$\operatorname{rank} \nabla_x\Phi(\bar{x},\bar{w}) = m. \tag{2.47}$$

Assume further that the outer function θ is continuously prox-regular at $\bar{z} := \Phi(\bar{x},\bar{w})$ for the unique vector $\bar{y}$ satisfying the relationships

$$\nabla_x\Phi(\bar{x},\bar{w})^*\bar{y} = \bar{v} - \nabla_x\varphi_0(\bar{x},\bar{w}) \ \text{ and } \ \bar{y} \in \partial\theta(\bar{z}). \tag{2.48}$$

Then $\bar{x}$ is a fully stable local minimizer for $\mathcal{P}(\bar{w},\bar{v})$ if and only if we have the implication

$$[(p,q)\in\mathcal{T}(\bar{x},\bar{w},\bar{v})(u),\ u\neq 0]\Longrightarrow\langle p,u\rangle>0 \tag{2.49}$$

for the set-valued mapping $\mathcal{T}(\bar{x},\bar{w},\bar{v})\colon\mathbb{R}^n\rightrightarrows\mathbb{R}^n\times\mathbb{R}^d$ defined by

$$\begin{aligned}\mathcal{T}(\bar{x},\bar{w},\bar{v})(u):\ &=\Big(\nabla^2_{xx}\varphi_0(\bar{x},\bar{w})u,\nabla^2_{xw}\varphi_0(\bar{x},\bar{w})u\Big)\\ &+\Big(\nabla^2_{xx}\langle\bar{y},\Phi\rangle(\bar{x},\bar{w})u,\nabla^2_{xw}\langle\bar{y},\Phi\rangle(\bar{x},\bar{w})u\Big)\\ &+\Big(\nabla_x\Phi(\bar{x},\bar{w}),\nabla_w\Phi(\bar{x},\bar{w})\Big)^*\partial^2\theta(\bar{z},\bar{y})(\nabla_x\Phi(\bar{x},\bar{w})u),\quad u\in\mathbb{R}^n.\end{aligned}$$

Proof. We apply the characterization of full stability from Theorem 2.3 to the function $\varphi(x,w)$ in (2.46). Observe first that the condition $\bar{v}\in\partial_x\varphi(\bar{x},\bar{w})$ on the tilt perturbation can be equivalently written as

$$\bar{v}\in\partial_x\varphi_0(\bar{x},\bar{w})+\nabla_x\Phi(\bar{x},\bar{w})^*\partial\theta(\bar{z}). \tag{2.50}$$

Indeed, this follows from first-order rules for φ in (2.46) under the full rank assumption on $\nabla_x\Phi(\bar{x},\bar{w})$; see, e.g., [27, Propositions 1.107(ii) and 1.112(i)]. Employing further the calculus of prox-regularity from [46, Theorem 2.1 and 2.2], which can be easily extended to the parametric case under consideration, allows us to conclude that the composite function φ is parametrically continuously prox-regular at $(\bar{x},\bar{w},\bar{v})$.

Let us show next that the basic constraint qualification (2.21) is automatically satisfied, under the assumptions made, for the function φ given in (2.46). Indeed, by the smoothness of φ_0 the constraint qualification (2.21) is clearly equivalent to

$$(0,q)\in\partial^\infty(\theta\circ\Phi)(\bar{x},\bar{w})\Longrightarrow q=0. \tag{2.51}$$

Employing in (2.51) the chain rule for (2.3) from [27, Proposition 1.107(ii)] reduces it to the implication

$$\Big[\nabla_x\Phi(\bar{x},\bar{w})^*p=0,\ \nabla_w\Phi(\bar{x},\bar{w})^*p=q,\ p\in\partial^\infty\theta(\bar{z})\Big]\Longrightarrow q=0,$$

which obviously holds due to the full rank condition (2.47).

Now we are ready to apply the characterization of full stability from Theorem 2.3 to the function φ in (2.46). Let us first check that condition (2.22) is automatically satisfied in the setting under consideration. To proceed, apply to this composite function φ the second-order sum rule from [27, Proposition 1.121] and then the second-order chain rule from [38, Theorem 3.1], which tell us that (2.22) is equivalent to

$$\Big[(0,q)\in\Big(\nabla_x\Phi(\bar{x},\bar{w}),\nabla_w\Phi(\bar{x},\bar{w})\Big)^*\partial^2\theta(\bar{z},\bar{y})(0)\Big]\Longrightarrow q=0, \tag{2.52}$$

where the uniqueness of the vector $\bar{y}$ satisfying (2.48) follows from the full rank condition (2.47). The last implication can be rewritten as

$$\Big[\nabla_x\Phi(\bar{x},\bar{w})^*p=0,\ \nabla_w\Phi(\bar{x},\bar{w})^*p=q,\ p\in\partial^2\theta(\bar{z},\bar{y})(0)\Big]\Longrightarrow q=0,$$

which surely holds by the full rank of $\nabla_x\Phi(\bar{x},\bar{w})$ in (2.47). To complete the proof of the theorem, it remains finally to observe that condition (2.23) in Theorem 2.3 reduces to that of (2.49) imposed in this theorem due to the aforementioned second-order sum and chain rules from [27, Proposition 1.121] and [38, Theorem 3.1] applied to the function φ in (2.46). $\triangle$

CHAPTER 3 STABILITY ANALYSIS OF UNCONSTRAINED PROBLEMS

This chapter addresses the following constrained optimization problem:

$$\text{minimize} \;\; \varphi_0(x) \;\; \text{subject to} \;\; \Phi(x) \colon = \big(\varphi_1(x), \ldots, \varphi_m(x)\big) \in \Theta, \tag{3.1}$$

where all the functions $\varphi_i \colon \mathbb{R}^n \to \mathbb{R}$ for $i = 0, \ldots, m$ are $\mathcal{C}^2$-smooth around the reference points, and where $\Theta \subset \mathbb{R}^m$ is a closed convex set in $\mathbb{R}^m$. Problems of this type belong to *conic programming* provided that Θ is a subcone of $\mathbb{R}^m$. Note that classical nonlinear programs (NLPs) with s inequality and $m - s$ equality constraints correspond to (3.1) for $\Theta = \mathbb{R}^s_- \times \{0\}^{m-s}$.

To study full stability of local minimizers in (3.1) by reducing it to the extended unconstrained format (2.14), consider the two-parametric version of (3.1) written as

$$\mathcal{P}(w, v): \quad \text{minimize} \;\; \varphi_0(x, w) + \delta_\Theta(\Phi(x, w)) - \langle v, x \rangle \;\; \text{over} \;\; x \in \mathbb{R}^n \tag{3.2}$$

with the *basic* parameter $w \in \mathbb{R}^d$ and the *tilt* parameter $v \in \mathbb{R}^n$ under the the same $\mathcal{C}^2$-smooth assumptions on φ_0 and Φ with respect to both variables. Let

$$\varphi(x, w) := \varphi_0(x, w) + \delta_\Theta(\Phi(x, w)) \;\; \text{with} \;\; x \in \mathbb{R}^n, \; w \in \mathbb{R}^d \tag{3.3}$$

and fix in what follows a triple $(\bar{x}, \bar{w}, \bar{v})$ such that $\Phi(\bar{x}, \bar{w}) \in \Theta$ and $\bar{v} \in \partial_x \varphi(\bar{x}, \bar{w})$. Recall that the (partial) *Robinson constraint qualification* (abbr. RCQ) holds for (3.1) at $(\bar{x}, \bar{w})$ with $\Phi(\bar{x}, \bar{w}) \in \Theta$ if we have

$$N_\Theta(\Phi(\bar{x}, \bar{w})) \cap \ker \nabla_x \Phi(\bar{x}, \bar{w})^* = \{0\}. \tag{3.4}$$

Note that for NLPs the full rank condition (2.47) reduces to the classical *linear independence constraint qualification* (LICQ). For the general constrained problem (3.1) this condition does not depend on the underlying set Θ and thus readily calls for a possible improvement.

We now recall two conditions from [1, Definition 3.135], widely recognized in the framework of (3.1), and extend Theorem 2.9 to this general setting by reducing it to the full rank case (2.47).

(RC) The closed and convex set $\Theta \subset \mathbb{R}^m$ is said to be $\mathcal{C}^2$-*reducible* at $\bar{z} = \Phi(\bar{x}, \bar{w}) \in \Theta$ to the closed and convex set $\Xi \subset \mathbb{R}^p$ if there is a neighborhood U of $\bar{z}$ and a $\mathcal{C}^2$-smooth mapping $h\colon U \to \mathbb{R}^p$ such that $\delta_\Theta(z) = \delta_\Xi(h(z))$ for all $z \in U$ and the derivative operator $\nabla h(\bar{z})\colon \mathbb{R}^m \to \mathbb{R}^p$ is surjective. If this holds for all $z \in \Theta$, then we say that Θ is $\mathcal{C}^2$-reducible to Ξ. The reduction is *pointed* if the tangent cone $T_\Xi(h(\bar{z}))$ is a pointed cone. Without loss of generality we assume that $h(\bar{z}) = 0$.

(ND) We say that $(\bar{x}, \bar{w})$ in **(RC)** is a *partially nondegenerate point* for Φ with respect to Θ if

$$\nabla_x \Phi(\bar{x}, \bar{w})\mathbb{R}^n + \text{lin}\{T_\Theta(\bar{z})\} = \mathbb{R}^m, \tag{3.5}$$

where $\text{lin}\{T_\Theta(\bar{z})\}$ signifies the largest linear subspace contained in $T_\Theta(\bar{z})$.

It is well known that the reducibility condition **(RC)** holds for many important classes of problems in constrained optimization. This includes the cases when Θ is a *polyhedral* set, a *Lorentz* (second-order, ice-cream) cone, and the cone of *positive semidefinite* matrices; see, e.g., [1]. The nondegeneracy condition **(ND)** is more restrictive. It follows from [1, Proposition 4.73] that (3.5) can be equivalently reformulated in the dual form

$$\text{span}\{N_\Theta(\bar{z})\} \cap \ker \nabla_x \Phi(\bar{x}, \bar{w})^* = \{0\}, \tag{3.6}$$

which shows that it reduces to LICQ for the case of NLPs, being however essentially less restrictive than the latter even for polyhedral sets Θ as in [39, Example 6.9].

To proceed further, impose **(RC)** and deduce from it that the original constraint $\Phi(x, w) \in \Theta$ in (3.1) is locally equivalent to $h(\Phi(x, w)) \in \Xi$. This allows us to conclude that problem $\mathcal{P}(w, v)$ in (3.2) locally around $(\bar{x}, \bar{w})$ amounts to the *reduced* problem as follows:

$$\mathcal{P}_r(w,v) \quad \begin{cases} \text{minimize } \ \varphi_0(x,w) - \langle v,x\rangle \ \text{ subject to } \ x\in\mathbb{R}^n, \\ \Psi(x,w) := h(\Phi(x,w)) \in \Xi, \end{cases} \tag{3.7}$$

which can be equivalently rewritten as

$$\text{minimize } \ \varphi_0(x,w) + \delta_\Xi(\Psi(x,w)) - \langle v,x\rangle \ \text{ over } \ x\in\mathbb{R}^n. \tag{3.8}$$

3.1 Full Stability of Constrained Optimization Problems

We begin with the following result, which shows that full stability issues for (3.2) and (3.7) are equivalent. Moreover, we show that the *full rank* condition for the reduced problem is ensured by the validity of **(ND)** for the original one.

Proposition 3.1 (full stability and nondegeneracy in the original and reduced problems). *Let $\bar{x}$ be a feasible solution to $\mathcal{P}(\bar{w},\bar{v})$ in* (3.2) *along the fixed parameter pair $(\bar{w},\bar{v})$, and let condition* **(RC)** *hold. Then $\bar{x}$ is a fully stable locally optimal solution to $\mathcal{P}(\bar{w},\bar{v})$ if and only if it is a fully stable locally optimal solution to the reduced problem $\mathcal{P}_r(\bar{w},\bar{v})$. Furthermore, the validity in addition of* **(ND)** *for $(\bar{x},\bar{w})$ implies the surjectivity of $\nabla_x\Psi(\bar{x},\bar{w})$ for Ψ in* (3.7).

Proof. The claimed equivalence follows directly from representation (3.8) of the reduced problem with Ψ from (3.7) and the definition of full stability. To prove the second part of the proposition, assume **(ND)** for $(\bar{x},\bar{w})$ in (3.2) and get by **(RC)** and [1, Proposition 4.73] that

$$\operatorname{lin}\big\{T_\Theta(\bar{z})\big\} = T_\Omega(\bar{z}) \ \text{ with } \ \Omega := \big\{z\in U\big|\ h(z)=0\big\},$$

where U is given in **(RC)**. Taking into account the representation of the tangent cone to Ω from [53, Example 6.8], the nondegeneracy condition (3.5) reduces now to

$$\nabla_x\Phi(\bar{x},\bar{w})\mathbb{R}^n + \ker\nabla h(\bar{z}) = \mathbb{R}^m.$$

Using this together with the surjectivity of $\nabla h(\bar z)$ we get by the classical chain rule that

$$\nabla_x\Psi(\bar x,\bar w)\mathbb{R}^n=\nabla h(\bar z)\nabla_x\Phi(\bar x,\bar w)\mathbb{R}^n=\nabla h(\bar z)\big(\nabla_x\Phi(\bar x,\bar w)\mathbb{R}^n+\ker\nabla h(\bar z)\big)=\nabla h(\bar z)\mathbb{R}^m=\mathbb{R}^p,$$

which justifies the surjectivity of $\nabla_x\Psi(\bar x,\bar w)$ and completes the proof of the proposition. $\triangle$

Observe further from the standard subdifferential sum and chain rules [27, 53] applied to (3.3) under (3.4) that the stationary condition $\bar v\in\partial_x\varphi(\bar x,\bar w)$ for (3.2) yields

$$\bar v\in\nabla_x\varphi_0(\bar x,\bar w)+\nabla_x\Phi(\bar x,\bar w)^*N_\Theta(\Phi(\bar x,\bar w)). \tag{3.9}$$

This leads us to the KKT system (2.48), which can be equivalently rewritten as

$$\bar v=\nabla_x L(\bar x,\bar w,\bar\lambda),\quad \bar\lambda\in N_\Theta(\Phi(\bar x,\bar w)) \tag{3.10}$$

via the *Lagrangian* $L(x,w,\lambda):=\varphi_0(x,w)+\langle\lambda,\Phi(x,w)\rangle$ for (3.2). It is well known (see, e.g., [1, Proposition 4.75]) that (3.10) admits a unique Lagrange multiplier under the validity of **(ND)**.

Similarly we define the KKT system associated with *reduced* problem (3.7) by

$$\bar v=\nabla_x L_r(\bar x,\bar w,\bar\mu),\quad \bar\mu\in N_\Xi(\Psi(\bar x,\bar w)), \tag{3.11}$$

where L_r is the *Lagrangian* for (3.7) given by $L_r(x,w,\mu):=\varphi_0(x,w)+\langle\mu,\Psi(x,w)\rangle$. This system surely has a unique solution due to the full rank result of Proposition 3.1.

The next important result provides a second-order subdifferential characterization of full stability for $\mathcal{P}(\bar w,\bar v)$ at nondegenerate solutions by reducing it to the full rank setting of Theorem 2.9. Our proof is essentially different from the original one given recently in [33, Theorem 5.6], which is based on the unform quadratic growth characterization of Robinson's strong regularity of the associated KKT system/generalized equation obtained in [1, Theorem 5.24].

Theorem 3.2 (second-order subdifferential characterization of full stability of

nondegenerate solutions in constrained optimization). *Let $\bar{x}$ be a feasible solution to the unperturbed problem $\mathcal{P}(\bar{w},\bar{v})$ in* (3.2) *with some $\bar{w}\in\mathbb{R}^d$ and $\bar{v}$ from* (3.9)*. Assume further* **(RC)** *and* **(ND)** *hold, and let $\bar{\lambda}$ be a unique vector satisfying* (3.10)*. Then $\bar{x}$ is a fully stable local minimizer of $\mathcal{P}(\bar{w},\bar{v})$ if and only if we have*

$$\langle u,\nabla^2_{xx}L(\bar{x},\bar{w},\bar{\lambda})u\rangle+\langle q,\nabla_x\Phi(\bar{x},\bar{w})u\rangle>0 \tag{3.12}$$

for all $q\in\partial^2\delta_\Theta(\bar{z},\bar{\lambda})(\nabla_x\Phi(\bar{x},\bar{w})u)$ with $u\neq 0$.

Proof. Starting with verifying the "only if" part, let $\bar{x}$ be a fully stable local minimizer for $\mathcal{P}(\bar{w},\bar{v})$ and hence for the reduced problem $\mathcal{P}_r(\bar{w},\bar{v})$ in (3.7) by the first part of Proposition 3.1. The second part of this proposition ensures that $\nabla_x\Psi(\bar{x},\bar{w})$ is surjective under the assumptions made. Then Theorem 2.9 tells us that implication (2.49) holds with replacing $\mathcal{T}(\bar{x},\bar{w},\bar{v})$ by the set-valued mapping $\widehat{\mathcal{T}}_r(\bar{x},\bar{w},\bar{v})\colon\mathbb{R}^n\rightrightarrows\mathbb{R}^n\times\mathbb{R}^d$ defined by

$$\begin{aligned}\widehat{\mathcal{T}}_r(\bar{x},\bar{w},\bar{v})(u):\;&=\Big(\nabla^2_{xx}\varphi_0(\bar{x},\bar{w})u,\nabla^2_{xw}\varphi_0(\bar{x},\bar{w})u\Big)\\&+\Big(\nabla^2_{xx}\langle\bar{\mu},\Psi\rangle(\bar{x},\bar{w})u,\nabla^2_{xw}\langle\bar{\mu},\Phi\rangle(\bar{x},\bar{w})u\Big)\\&+\Big(\nabla_x\Psi(\bar{x},\bar{w}),\nabla_w\Psi(\bar{x},\bar{w})\Big)^*\partial^2\delta_\Xi(\bar{z},\bar{\mu})(\nabla_x\Psi(\bar{x},\bar{w})u),\quad u\in\mathbb{R}^n,\end{aligned}$$

where $\bar{\mu}$ is a unique solution of the reduced KKT system (3.11). Using now the second-order chain rule from [38, Theorem 3.1] under the full rank assumption leads us to

$$\widehat{\mathcal{T}}_r(\bar{x},\bar{w},\bar{v})(u)=\Big(\nabla^2_{xx}\varphi_0(\bar{x},\bar{w})u,\nabla^2_{xw}\varphi_0(\bar{x},\bar{w})u\Big)+D^*\partial_x(\delta_\Xi\circ\Psi)(\bar{x},\bar{w},\bar{v})(u). \tag{3.13}$$

On the other hand, it follows from **(RC)** that $(\delta_\Xi\circ\Psi)(x,w)=(\delta_\Theta\circ\Phi)(x,w)$ for all (x,w) around $(\bar{x},\bar{w})$. Using this together with (3.13), we get

$$\widehat{\mathcal{T}}_r(\bar{x},\bar{w},\bar{v})(u)=\Big(\nabla^2_{xx}\varphi_0(\bar{x},\bar{w})u,\nabla^2_{xw}\varphi_0(\bar{x},\bar{w})u\Big)+D^*\partial_x(\delta_\Theta\circ\Phi)(\bar{x},\bar{w},\bar{v})(u).$$

Finally, the result of [43, Theorem 7] held under **(ND)** ensures that

$$\widehat{\mathcal{T}}_r(\bar{x},\bar{w},\bar{v})(u)=\mathcal{T}(\bar{x},\bar{w},\bar{v})(u),\quad u\in\mathbb{R}^n, \tag{3.14}$$

which justifies together with (2.49) that condition (3.12) is satisfied.

To verify now the "if" part, assume the validity of (3.12) and deduce from (3.14) that it also holds for $\widehat{\mathcal{T}}_r(\bar{x},\bar{w},\bar{v})$; hence we get implication (2.49) for the latter mapping. By the surjectivity of $\nabla_x\Psi(\bar{x},\bar{w})$ it follows from Theorem 2.9 that $\bar{x}$ is a fully stable local minimizer of the reduced problem $\widehat{\mathcal{P}}(\bar{w},\bar{v})$ and thus for the original problem $\mathcal{P}(\bar{w},\bar{v})$ by Proposition 3.1.$\triangle$

Remark 3.3 (enhanced second-order condition). An important point hidden in the proof of Theorem 3.2 and used below is that assumptions **(RC)** and **(ND)** ensure the validity of implication (2.22), which was established previously under the full rank condition; see (2.52) in the proof of Theorem 2.9. To elaborate it more, take $(0,q)\in\mathcal{T}(\bar{x},\bar{w},\bar{v})(0)$ and observe from the discussion above that it yields

$$\nabla_x\Phi(\bar{x},\bar{w})^*p=0 \ \text{ and } \ \nabla_w\Phi(\bar{x},\bar{w})^*p=q \ \text{ with some } \ p\in\partial^2\delta_\Theta(\bar{z},\bar{\lambda})(0). \tag{3.15}$$

Employing **(RC)**, we get $\Theta\cap U=h^{-1}(\Xi)\cap U$ in the notation therein. It follows from [27, Theorem 1.17] by the surjectivity of $\nabla_x\Psi(\bar{x},\bar{w})$ for $\Psi=h\circ\Phi$ that $N_\Theta(\bar{z})=\nabla h(\bar{z})^*N_\Xi(\Psi(\bar{x},\bar{w}))$. Appealing now to [38, Theorem 3.1] gives us $d\in\partial^2\delta_\Xi(\Psi(\bar{x},\bar{w}),\bar{\mu})(0)$ such that $p=\nabla h(\bar{z})^*d$, where $\bar{\mu}\in N_\Xi(\Psi(\bar{x},\bar{w}))$ is the unique solution to the reduced KKT system (3.11) satisfying $\bar{\lambda}=\nabla h(\bar{z})^*\bar{\mu}$. Thus $\nabla_x\Psi(\bar{x},\bar{w})^*d=0$ due to (3.15), which shows that $d=0$ and hence $p=0$. Substitution $p=0$ into (3.15) justifies (2.22).

Recall that the validity of implication (2.22) under assumptions **(RC)** and **(ND)** was first proved in [39, Theorem 6.6] for *mathematical programs with polyhedral constraint* (i.e., when Θ in (3.1) is as polyhedral set) and then in [37, Lemma 4.5] for *second-order cone programs* when Θ stands for the Lorentz second-order/ice-cream cone.

As a consequence of the discussions in Remark 3.3, we show next that the validity of **(ND)** under **(RC)** implies the following *second-order qualification condition* (SOQC)

$$\partial^2\delta_\Theta(\bar{z},\bar{\lambda})(0)\cap\ker\nabla_x\Phi(\bar{x},\bar{w})^*=\{0\} \tag{3.16}$$

from [38], where $\bar{\lambda}$ is the unique solution of KKT system (3.10) at the given triple $(\bar{x}, \bar{w}, \bar{v})$. Note that the converse implication holds when Θ is either a polyhedral convex set [35, Proposition 6.1], or the Lorentz second-order cone [37, Theorem 3.6], or the SDP cone $\mathcal{S}^m_+$ (this can be derived from [3]), while in general it still remains an open question.

Corollary 3.4 (second-order qualification condition under nondegeneracy). *Let $\bar{\lambda}$ be a unique vector satisfying* (3.10) *for the triple* $(\bar{x}, \bar{w}, \bar{v})$ *from Theorem* 2.9*, and let conditions* **(RC)** *and* **(ND)** *be fullfiled. Then SOCQ* (3.16) *holds.*

Proof. Take $p \in \partial^2\delta_\Theta(\bar{z}, \bar{\lambda})(0)$ with $\nabla_x\Phi(\bar{x}, \bar{w})^*p = 0$ and find from the discussions in Theorem 3.2 and Remark 3.3 a vector $d \in \partial^2\delta_\Xi(\Psi(\bar{x}, \bar{w}), \bar{\mu})(0)$ such that $p = \nabla h(\bar{z})^*d$ in the notation above. This gives us $\nabla_x\Psi(\bar{x}, \bar{w})^*d = 0$ and hence $d = 0$ by the surjectivity of $\nabla_x\Psi(\bar{x}, \bar{w})$. It shows that $p = 0$ and completes the proof. $\triangle$

The next consequence of Theorem 3.2 opens a technical gate for obtaining the main result of Section 3.2 given in Theorem 3.8. To proceed, consider the following *canonically perturbed* version of problem (3.1) with parametric pairs $(v_1, v_2) \in \mathbb{R}^n \times \mathbb{R}^m$:

$$\widetilde{\mathcal{P}}_{\bar{w}}(v_1, v_2) \quad \begin{cases} \text{minimize} \;\; \varphi_0(x, \bar{w}) - \langle v_1, x\rangle \;\; \text{subject to} \;\; x \in \mathbb{R}^n \\ \Phi(x, \bar{w}) + v_2 \in \Theta. \end{cases} \tag{3.17}$$

Corollary 3.5 (full stability with respect to canonical perturbations). *Let $\bar{x}$ be a feasible solution to the unperturbed problem $\mathcal{P}(\bar{w}, \bar{v})$ in* (3.2) *with some $\bar{w} \in \mathbb{R}^d$ and $\bar{v}$ from* (3.9), and assumptions **(RC)** and **(ND)** be satisfied. Then $\bar{x}$ is a fully stable local minimizer of $\mathcal{P}(\bar{w}, \bar{v})$ if and only if it is a fully stable local minimizer of $\widetilde{\mathcal{P}}_{\bar{w}}(\bar{v}, 0)$.

Proof. We can easily to see that the nondegeneracy condition **(ND)** for $\mathcal{P}(\bar{w}, \bar{v})$ at $\bar{x}$ is equivalent to the validity of this condition for $\widetilde{\mathcal{P}}_{\bar{w}}(\bar{v}, 0)$. It follows from Theorem 3.2 that the full stability of the local minimizer $\bar{x}$ in both problems $\mathcal{P}(\bar{w}, \bar{v})$ and $\widetilde{\mathcal{P}}_{\bar{w}}(\bar{v}, 0)$ amounts to the validity of the same second-order condition (3.12). This justifies the claimed equivalence. △

Looking at the problem $\widetilde{\mathcal{P}}_{\bar{w}}(\bar{v},0)$ in Corollary 3.5, observe that it corresponds to just the *tilt* perturbation of the original problem (3.2) with the *fixed* basic parameter $w=\bar{w}$. The latter problem can be written as $\mathcal{P}_{\bar{w}}(v)$. Thus we have the following consequence of Corollary 3.5 about the relationship between full and tilt stability under the assumptions made.

Corollary 3.6 (reduction of full stability to tilt stability at nondegenerate solutions.) *Consider the setting of Corollary* 3.5*. Then the full stability of the local minimizer* $\bar{x}$ *for the original problem* $\mathcal{P}(\bar{w},\bar{v})$ *is equivalent to its tilt stability in problem* $\mathcal{P}_{\bar{w}}(\bar{v})$*.*

Proof. It follows from the discussion above that both stability notions are characterized by the same second-order condition (3.12) under the **(RC)** and **(ND)** assumptions made. △

3.2 Relationships of Full Stability with Other Stability Notions

This section addresses relationships between full stability of our basic problem $\mathcal{P}(\bar{w},\bar{v})$ in (3.2) and other well-recognized stability notions in constrained optimization and associated variational systems. We develop a largely self-contained approach to such relationships based on the reduction procedure of Section 3.1, which allows us to establish new equivalences and also to provide new proofs of some recently discovered results in this direction.

We first present a rather simple description of full stability in $\mathcal{P}(\bar{w},\bar{v})$ via a Lipschitzian single-valued localization of the parameterized collection of stationary points therein. Recall that a set-valued mapping $F\colon\mathbb{R}^n\rightrightarrows\mathbb{R}^m$ admits a *single-valued graphical localization* around $(\bar{x},\bar{y})\in F$ provided that there exist neighborhoods U of $\bar{x}$ and V of $\bar{y}$ together with a single-valued mapping $f\colon U\to V$ such that $F\cap(U\times V)=f$.

Proposition 3.7 (equivalence between full stability of $\mathcal{P}(\bar{w},\bar{v})$ and Lipschitzian localization of parameterized stationary points). *Let* $\bar{x}$ *be a feasible solution to the* *unperturbed problem* $\mathcal{P}(\bar{w},\bar{v})$ *in* (3.2) *with some* $\bar{w}\in\mathbb{R}^d$ *and* $\bar{v}$ *from* (3.9), and let RCQ (3.4) hold. Then $\bar{x}$ is a fully stable locally optimal solution to problem $\mathcal{P}(\bar{w},\bar{v})$ if and only if $\bar{x}\in M_\gamma(\bar{w},\bar{v})$ for some $\gamma>0$ and the set-valued mapping

$$S(w,v):=\Big\{x\in\mathbb{R}^n\Big|\ v\in\nabla_x\varphi_0(x,w)+\nabla_x\Phi(x,w)^*N_\Theta(\Phi(x,w))\Big\} \tag{3.18}$$

admits a Lipschitzian single-valued graphical localization around $(\bar{w},\bar{v},\bar{x})$.

Proof. Applying the corresponding characterization of full stability in the general unconstrained format of Theorem 2.5 with the extended-real-valued function φ from (3.3), we conclude that the basic constraint qualification imposed on Theorem 2.5 holds due to the assumed RCQ by [20, Proposition 2.2]. This ends the proof. △

Unless otherwise stated, in the rest of this section we take $\bar{v}=0$ in the KKT system (3.10) without less of generality. Consider the *generalized equation* (GE)

$$\begin{bmatrix} v \\ 0 \end{bmatrix} \in \begin{bmatrix} \nabla_x L(x,w,\lambda) \\ -\Phi(x,w) \end{bmatrix} + \begin{bmatrix} 0 \\ N_\Theta^{-1}(\lambda) \end{bmatrix}, \tag{3.19}$$

which is indeed the KKT system for problem $\mathcal{P}(w,v)$ in (3.2). Let $(\bar{x},\bar{\lambda})$ be a solution to (3.19) with $(w,v)=(\bar{w},0)$ and define the *partial linearization* of (3.19) at $(\bar{x},\bar{\lambda})$ by

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \in \begin{bmatrix} \nabla^2_{xx} L(\bar{x},\bar{w},\bar{\lambda})(x-\bar{x})+\nabla_x\Phi(\bar{x},\bar{w})^*(\lambda-\bar{\lambda}) \\ -\Phi(\bar{x},\bar{w})-\nabla_x\Phi(\bar{x},\bar{w})(x-\bar{x}) \end{bmatrix} + \begin{bmatrix} 0 \\ N_\Theta^{-1}(\lambda) \end{bmatrix}. \tag{3.20}$$

Recall [47] that $(\bar{x},\bar{\lambda})$ is a *strongly regular* solution to the KKT system (3.19) if the solution map to (3.20) has a Lipschitz continuous single-valued localization around $(0,0)\in\mathbb{R}^n\times\mathbb{R}^m$.

Theorem 3.8 (full stability of $\mathcal{P}(\bar{w},\bar{v})$ and local single-valuedness and Lipschitz continuity of solution maps to basic and reduced KKT systems). *Let $\bar{x}$ be a feasible solution to the unperturbed problem $\mathcal{P}(\bar{w},\bar{v})$ in* (3.2) *with some $\bar{w}\in\mathbb{R}^d$ and $\bar{v}=0$ from* (3.9) under the validity of the reducibility **(RC)** and RCQ conditions. The following are equivalent:

(i) $\bar{x}$ is a fully stable locally optimal solution to $\mathcal{P}(\bar{w}, \bar{v})$ satisfying **(ND)**.

(ii) $\bar{x} \in M_\gamma(\bar{w}, \bar{v})$ for some $\gamma > 0$ and the solution map $S^r_{KKT} \colon (w, v) \mapsto (x, \mu)$ for the reduced KKT system (3.11) is single-valued and Lipschitz continuous around $(\bar{w}, \bar{v}, \bar{x}, \bar{\mu})$.

(iii) $\bar{x} \in M_\gamma(\bar{w}, \bar{v})$ for some $\gamma > 0$ and the solution map $S_{KKT} \colon (w, v) \mapsto (x, \lambda)$ for the KKT system (3.10) is single-valued and Lipschitz continuous around $(\bar{w}, \bar{v}, \bar{x}, \bar{\lambda})$.

(iv) $\bar{x} \in M_\gamma(\bar{w}, \bar{v})$ for some $\gamma > 0$ and $(\bar{x}, \bar{\lambda})$ is a strongly regular solution to (3.19).

Proof. To verify implication (**i**) $\Longrightarrow$ (**ii**), we get from (**i**) and the first part of Proposition 3.1 that $\bar{x}$ is a fully stable locally optimal solution to the reduced problem (3.7). The local single-valuedness of the solution map $S^r_{KKT} \colon (w, v) \mapsto (x, \mu)$ for the reduced KKT system (3.11) in **(ii)** was established above as a consequence of the imposed **(RC)** and **(ND)** assumptions ensuring the full rank condition for $\nabla_x \Psi(\bar{x}, \bar{w})$ by the second part of Proposition 3.1.

Next we verify that the mapping $S^r_{KKT} \colon (w, v) \mapsto (x, \mu)$ is Lipschitz continuous around $(\bar{w}, \bar{v}, \bar{x}, \bar{\mu})$. Note that the Lipschitz continuity of $(w, v) \mapsto x_{wv}$ comes directly from the full stability of $\bar{x}$ in (3.7). To justify this property for the mapping $(w, v) \mapsto \mu_{wv}$, pick $w_1, w_2 \in W$ and $v_1, v_2 \in V$ and then find $\mu_{w_i v_i} \in N_\Xi(c_i)$ with $c_i := \Psi(x_{w_i v_i}, w_i)$ for $i = 1, 2$ satisfying

$$\begin{cases} v_2 = \nabla_x \varphi_0(x_{w_2 v_2}, w_2) + \nabla_x \Psi(x_{w_2 v_2}, w_2)^* \mu_{w_2 v_2}, \\ v_1 = \nabla_x \varphi_0(x_{w_1 v_1}, w_1) + \nabla_x \Psi(x_{w_1 v_1}, w_1)^* \mu_{w_1 v_1}. \end{cases}$$

It shows therefore that the validity of the following equality

$$\begin{aligned} \nabla_x \Psi(x_{w_2 v_2}, w_2)^* (\mu_{w_2 v_2} - \mu_{w_1 v_1}) \quad &= \quad \Big(\nabla_x \Psi(x_{w_1 v_1}, w_1) - \nabla_x \Psi(x_{w_2 v_2}, w_2) \Big)^* \mu_{w_1 v_1} \\ &+ \quad \nabla_x \varphi_0(x_{w_1 v_1}, w_1) - \nabla_x \varphi_0(x_{w_2 v_2}, w_2) + v_2 - v_1. \end{aligned} \tag{3.21}$$

By shrinking the neighborhoods W and V if necessary, we can always assume the surjectivity of $\nabla_x \Psi(x_{w_i v_i}, w_i)$ due to this property of $\nabla_x \Psi(\bar{x}, \bar{w})$. Thus it follows from the standard surjectivity result of [27, Lemma 1.18] that for any $(w, v) \in W \times V$ there is $\kappa_{wv} > 0$ such that

$$\|\nabla_x\Psi(x_{w_2v_2},w_2)^*(\mu_{w_2v_2}-\mu_{w_1v_1})\|\ge\kappa_{w_2v_2}\|\mu_{w_2v_2}-\mu_{w_1v_1}\|\ge\kappa\|\mu_{w_2v_2}-\mu_{w_1v_1}\|, \tag{3.22}$$

where $\kappa:=\inf\{\kappa_{wv}|\ (w,v)\in W\times V\}$. Furthermore, it is easy to conclude from the surjectivity of $\nabla_x\Psi(\bar{x},\bar{w})$ that $\kappa>0$ and that there is $\rho<\infty$ such that $\|\mu_{wv}\|\le\rho$ for all $(w,v)\in W\times V$. Denoting by $\ell>0$ is a common Lipschitz constant for the mappings $\nabla_x\varphi_0$, $\nabla_x\Phi$, ∇h, Φ, and $(w,v)\mapsto x_{wv}$ on $W\times V$, we derive from (3.21) and (3.22) the estimates

$$\begin{aligned}\|\mu_{w_2v_2}-\mu_{w_1v_1}\| &\le \tfrac{1}{\kappa}\Big(\|\nabla_x\Psi(x_{w_1v_1},w_1)-\nabla_x\Psi(x_{w_2v_2},w_2)\|\cdot\|\mu_{w_1v_1}\|\\ &+ \|\nabla_x\varphi_0(x_{w_1v_1},w_1)-\nabla_x\varphi_0(x_{w_2v_2},w_2)\|+\|v_2-v_1\|\Big)\\ &\le \tfrac{\gamma}{\kappa}\Big[\rho\ell^2\Big(\|x_{w_2v_2}-x_{w_1v_1}\|+\|w_2-w_1\|\Big)\\ &+ \ell\Big(\|x_{w_2v_2}-x_{w_1v_1}\|+\|w_2-w_1\|\Big)+\|v_2-v_1\|\Big].\end{aligned}$$

which imply the local Lipschitz continuity of $(w,v)\mapsto\mu_{wv}$ and thus justify **(ii)**.

Next we show that assuming the local Lipschitz continuity of $S^r_{KKT}\colon(w,v)\mapsto(x,\mu)$ around $(\bar{w},\bar{v},\bar{x},\bar{\mu})$ in **(ii)** implies this property for $S_{KKT}\colon(w,v)\mapsto(x,\lambda)$ around $(\bar{w},\bar{v},\bar{x},\bar{\lambda})$ with $\bar{\lambda}:=\nabla h(\bar{x})^*\bar{\mu}$ and $\bar{z}:=\Phi(\bar{x},\bar{w})$ in **(iii)**. Similarly to the above, it remains to verify that the mapping $(w,v)\mapsto\lambda_{wv}$ is Lipschitz continuous around $(\bar{w},\bar{v})$. To proceed, take any $w_i\in W$, $v_i\in V$ and form $\lambda_{w_iv_i}:=\nabla h(z_i)^*\mu_{w_iv_i}$ with $z_i:=\Phi(x_{w_iv_i},w_i)$ for $i=1,2$. Then we have

$$\begin{aligned}\|\lambda_{w_2v_2}-\lambda_{w_1v_1}\| &= \|\nabla h(z_2)^*\mu_{w_2v_2}-\nabla h(z_1)^*\mu_{w_1v_1}\|\\ &\le \|\nabla h(z_2)^*\|\cdot\|\mu_{w_2v_2}-\mu_{w_1v_1}\|+\|\nabla h(z_2)-\nabla h(z_1)\|\cdot\|\mu_{w_1v_1}\|\\ &\le \tau\|\mu_{w_2v_2}-\mu_{w_1v_1}\|+\rho\ell^2\Big(\|x_{w_2v_2}-x_{w_1v_1}\|+\|w_2-w_1\|\Big)\end{aligned}$$

in the notation above, where $\tau>0$ is an upper bound of $\|\nabla h(z)^*\|$ for all z sufficiently close to $\bar{z}$. This justifies the claimed local Lipschitz continuity of S_{KKT} and thus verifies **(iii)**.

Our next implication to prove is **(iii)**$\Longrightarrow$**(i)**. Taking into account Proposition 3.7 and the form of the KKT system (3.10), it remains to check that (**iii**) ensures the validity of **(ND)**.

This kind of relationships has been well understood in optimization theory (see, e.g., [1]); we present a complete proof in our setting for the reader's convenience.

Arguing by contradiction, suppose that **(ND)** in the equivalent form (3.5) does not hold and thus find $0 \neq \vartheta \in \mathbb{R}^m$ so that $\nabla_x \Phi(\bar{x}, \bar{w})^* \vartheta = 0$ and that $\vartheta \in \text{span}\{N_\Theta(\bar{z})\}$ with $\bar{z} = \Phi(\bar{x}, \bar{w})$. By **(iii)** we have $S_{KKT}(\bar{w}, \bar{v}) = \{(\bar{x}, \bar{\lambda})\}$ for some $\bar{\lambda} \in \mathbb{R}^m$. If $\bar{\lambda} \in \text{ri}\, N_\Theta(\bar{z})$, with "ri" standing for the relative interior of a convex set, then $\bar{\lambda} + t\vartheta \in N_\Theta(\bar{z})$ for any small $t > 0$. Indeed, it is easy to see that $\text{span}\{N_\Theta(\bar{z})\} = \text{aff}\,\{N_\Theta(\bar{z})\}$ and hence $\bar{\lambda} + t\vartheta \in \text{aff}\,\{N_\Theta(\bar{z})\}$ when $t > 0$ is sufficiently small. Employing this, we get $(\bar{x}, \bar{\lambda} + t\vartheta) \in S_{KKT}(\bar{w}, \bar{v})$, which contradicts the the aforementioned uniqueness of Lagrange multipliers in (3.10) and so justifies **(ND)** in this case. In the remaining case of $\bar{\lambda} \notin \text{ri}\, N_\Theta(\bar{z})$, pick $\xi \in \text{ri}\, N_\Theta(\bar{z}) \neq \emptyset$ and get from the well-known result of convex analysis (see, e.g., [53, Proposition 2.40]) that $\bar{\lambda} + t(\xi - \bar{\lambda}) \in \text{ri}\, N_\Theta(\bar{z})$ for any $t \in (0, 1)$. Putting $v_t = t\nabla_x \Phi(\bar{x}, \bar{w})^*(\xi - \bar{\lambda})$ for $t > 0$ sufficiently small, we obtain that $(\bar{x}, \bar{\lambda} + t(\xi - \bar{\lambda})) \in S_{KKT}(\bar{w}, v_t)$. Since $\bar{\lambda} + t(\xi - \bar{\lambda}) \in \text{ri}\, N_\Theta(\bar{z})$, it again justifies **(ND)** by the arguments above and thus confirms the validity of assertion **(i)**.

To verify now implication (**i**) $\Longrightarrow$ (**iv**), take $\bar{x}$ from **(i)** and deduce from Corollary 3.5 that $\bar{x}$ is a fully stable locally optimal solution to problem $\widetilde{\mathcal{P}}_{\bar{w}}(\bar{v}, 0)$ defined by (3.17) with $\bar{v} = 0$. Note that the KKT system for the parametric problem $\widetilde{\mathcal{P}}_{\bar{w}}(v_1, v_2)$ is given by

$$\begin{bmatrix} v_1 \\ \\ v_2 \end{bmatrix} \in \begin{bmatrix} \nabla_x L(x, \bar{w}, \lambda) \\ \\ -\Phi(x, \bar{w}) \end{bmatrix} + \begin{bmatrix} 0 \\ \\ N_\Theta^{-1}(\lambda) \end{bmatrix}, \tag{3.23}$$

where (v_1, v_2) varies around $(\bar{v}, 0) \in \mathbb{R}^n \times \mathbb{R}^m$. It follows from the implication (**i**) $\Longrightarrow$ (**iii**) established above that the solution map $\widetilde{S}_{KKT} \colon (v_1, v_2) \mapsto (x, \lambda)$ for (3.23) is single-valued and Lipschitz continuous around $(\bar{v}, 0)$. Observe that the generalized equation (3.20) can be treated as a (partial) linearization of the KKT system (3.23). Taking into account that (3.23) is a *canonically* perturbed system, we conclude that the local single-valuedness and Lipschitz continuity of its solution map is *equivalent* to these properties of solutions to its linearization (3.20); see, e.g., [5, Theorem 2B.10]). The latter justifies the strong regularity of the KKT system (3.10) around $(\bar{x}, \bar{\lambda})$ according to the definition above taken from [47].

To complete the proof of the theorem, it remains to show that (**iv**) $\Longrightarrow$ (**i**). Take $\bar{x}$ satisfying **(iv)** with some $\bar{\lambda}$. Then the arguments of the preceding paragraph tell us that the solution map $\widetilde{S}_{KKT}$ for the KKT system (3.23) is single-valued and Lipschitz continuous around $(\bar{v}, 0)$. Employing now in this setting the implication (**iii**) $\Longrightarrow$ (**i**) established above ensures that $\bar{x}$ is a fully stable locally optimal solution to problem $\widetilde{\mathcal{P}}_{\bar{w}}(\bar{v}, 0)$ satisfying **(ND)**. Thus it is a fully stable locally optimal solution to the original problem $\mathcal{P}(\bar{w}, \bar{v})$ by Corollary 3.5. $\triangle$

Note that the equivalence **(i)**$\Longleftrightarrow$**(iv)** of Theorem 3.8 has been recently proved in [33, Theorem 5.6] by using a more sophisticated device based on characterizing strong regularity in [1] via the uniform quadratic growth condition with respect to the so-called $\mathcal{C}^2$-smooth parametrization defined below. Furthermore, the latter growth condition has been employed in [33] to characterize yet another stability notion known as strong Lipschitzian stability. In theorem 3.9 we relate this notion to full stability by using a new approach via Theorem 3.2 and Proposition 3.7. Note that the first part of Theorem 3.9 does not impose **(RC)** in contrast to [33, Theorem 5.6].

To proceed, fix $\bar{w} \in \mathbb{R}^d$ and consider the constrained optimization problem

$$\mathcal{P}_{\bar{w}}: \quad \text{minimize } \varphi_0(x, \bar{w}) \text{ subject to } \Phi(x, \bar{w}) \in \Theta \tag{3.24}$$

with the data from (3.1). We say that the pair $(\vartheta(x,u), \Upsilon(x,u))$ with $u \in \mathbb{R}^s$ and $\vartheta\colon \mathbb{R}^n \times \mathbb{R}^s \to \mathbb{R}$, $\Upsilon\colon \mathbb{R}^n \times \mathbb{R}^s \to \mathbb{R}^m$ is a *$\mathcal{C}^2$-smooth parametrization* of $(\varphi_0(x, \bar{w}), \Phi(x, \bar{w}))$ in (3.24) at $\bar{u} \in \mathbb{R}^s$ if $\varphi_0(x, \bar{w}) = \vartheta(x, \bar{u})$ and $\Phi(x, \bar{w}) = \Upsilon(x, \bar{u})$ for all $x \in \mathbb{R}^n$, where both ϑ and Υ are twice continuously differentiable. Define the family of parametric optimization problems:

$$\widehat{\mathcal{P}}(u) \quad \begin{cases} \text{minimize } \ \vartheta(x,u) \ \text{ subject to } \ x \in \mathbb{R}^n, \\ \Upsilon(x,u) \in \Theta. \end{cases}$$

We say [1, Definition 5.33] that a stationary point $\bar{x}$ of $\mathcal{P}_{\bar{w}}$ is *strongly Lipschitz stable* with respect to the $\mathcal{C}^2$-smooth parametrization $(\vartheta(x,u), \Upsilon(x,u))$ of $(\varphi_0(x,\bar{w}), \Phi(x,\bar{w}))$ in (3.24) at $\bar{u} \in \mathbb{R}^s$ if there are neighborhoods U of $\bar{u}$ and X of $\bar{x}$ such that for any $u \in U$ each problem $\widehat{\mathcal{P}}(u)$ has a unique stationary point $x(u) \in X$ and the mapping $u \longmapsto x(u)$ is Lipschitz continuous around $\bar{u}$. If it holds for any $\mathcal{C}^2$-smooth parameterizations of $(\varphi_0(x,\bar{w}), \Phi(x,\bar{w}))$ in (3.24) at $\bar{u} \in \mathbb{R}^s$, then $\bar{x}$ is called strongly Lipschitz stable. This notion is a Lipschitzian counterpart of the Kojima's *strong stability* [18], where the mapping $u \longmapsto x(u)$ is merely continuous.

Theorem 3.9 (full stability vs. strong Lipschitzian stability in constrained optimization. *Let $\bar{x}$ be a Lipschitz stable locally optimal solution to problem $\mathcal{P}_{\bar{w}}$ in the framework of Proposition* 3.7. *Then it is a fully stable locally optimal solution to problem $\mathcal{P}(\bar{w},\bar{v})$ with $\bar{v} = 0$. The converse implication holds provided that both* **(RC)** *and* **(ND)** *conditions are satisfied.*

Proof. To justify the first part of the theorem, take a Lipschitz stable locally optimal solution to (3.24). It is easy to see that $(\varphi_0(x,w) - \langle x, v\rangle, \Phi(x,w))$ is a $\mathcal{C}^2$-smooth parametrization of $(\varphi_0(x,\bar{w}), \Phi(x,\bar{w}))$ in (3.24) at $\bar{u} := (\bar{w}, 0) \in \mathbb{R}^d \times \mathbb{R}^n$. Let $x(u)$ be a unique stationary point $x(u)$ for any $u = (w,v)$ close enough to $\bar{u}$ and so that the mapping $u \longmapsto x(u)$ is Lipschitz continuous around $\bar{u}$. This tells us that the set-valued mapping

$$S(u) \colon = \Big\{ x \in \mathbb{R}^n \Big| \ v \in \nabla_x \varphi_0(x,w) + \nabla_x \Phi(x,w)^* N_\Theta(\Phi(x,w)) \Big\}$$

has a Lipschitzian single-valued graphical localization around $(\bar{u}, \bar{x})$. Employing now Proposition 3.7, we deduce that $\bar{x}$ is a fully stable locally optimal solution to problem $\mathcal{P}(\bar{w}, 0)$.

To prove the converse implication of the theorem, suppose that $\bar{x}$ is a fully stable locally optimal solution to problem $\mathcal{P}(\bar{w},0)$ under the validity of **(RC)** and **(ND)**. By Theorem 3.2 we have the second-order characterization (3.12). Take now an arbitrary $\mathcal{C}^2$-smooth parametrization of $(\varphi_0(x,\bar{w}),\Phi(x,\bar{w}))$ in (3.24) at $\bar{u}\in\mathbb{R}^s$. This yields the equalities $\nabla_x\varphi_0(\bar{x},\bar{w}) = \nabla_x\vartheta(\bar{x},\bar{u})$, $\nabla_x\Phi(\bar{x},\bar{w}) = \nabla_x\Upsilon(\bar{x},\bar{u})$ as well as those for the corresponding second-order derivatives. Thus we have (3.12) for problem $\widehat{\mathcal{P}}(\bar{u})$, which ensures that $\bar{x}$ is a fully stable locally optimal solution to this problem. Then it follows from Proposition 3.7 that the set-valued mapping

$$S(u,v):=\Big\{x\in\mathbb{R}^n\Big|\ v\in\nabla_x\vartheta(x,u)+\nabla_x\Upsilon(x,u)^*N_\Theta(\Upsilon(x,u))\Big\}$$

admits a Lipschitzian single-valued graphical localization around $(\bar{u},0)$. Letting now $x(u):=S(u,0)$, we get that $x(u)$ is a stationary point for problem $\widehat{\mathcal{P}}(u)$ and that the mapping $u\longmapsto x(u)$ is locally Lipschitz continuous around $\bar{u}$. This verifies the strong Lipschitzian stability of $\bar{x}$ in (3.24) and thus completes the proof of the theorem. $\triangle$

CHAPTER 4 SECOND-ORDER ANALYSIS OF PIECEWISE LINEAR FUNCTIONS

In this chapter we mainly address variational theory and applications of the class of *convex piecewise linear* (CPWL) extended-real-valued functions [53] playing an important role in many aspects of variational analysis and optimization. Having in hands recently obtained [41] explicit calculations of the *second-order subdifferentials* (or generalized Hessians) of such function in the sense of [26], we present here some of their applications to second-order variational analysis and parametric optimization. Proceeding in this direction requires us to deal not only with CPWL functions per se but mainly with *fully amenable compositions* involving such functions, which play an underlying role in many aspects of variational analysis, optimization, and stability.

Recall [51, 53] that $\theta\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ is *piecewise linear* if its domain $\operatorname{dom}\theta$ is nonempty and can be represented as the union of finitely many convex polyhedral sets so that on each of these pieces θ is given by $\langle a, x\rangle - \alpha$ with some $\alpha \in \mathbb{R}$ and $a \in \mathbb{R}^n$. Observing that such functions are not necessarily convex, we focus on the study of *convex* piecewise linear (CPWL) functions, which admit the following equivalent descriptions [53, Theorem 2.49]. For simplicity we write in what follows $\theta \in CPWL$ whenever θ belongs to this class.

Proposition 4.1 (convex piecewise linear functions). *The following are equivalent:*

(i) $\theta\colon \mathbb{R}^m \to \overline{\mathbb{R}}$ *is a convex and piecewise linear function labeled as* $\theta \in CPWL$.

(ii) *The epigraph* $\operatorname{epi}\theta$ *is a convex polyhedron in* $\mathbb{R}^{m+1}$.

(iii) *There are* $\alpha_i \in \mathbb{R}$ *and* $a_i \in \mathbb{R}^n$ *for* $i \in T_1\colon = \{1, \ldots, l\}$ *such that* φ *is represented by*

$$\theta(x) = \begin{cases} \max\Big\{\langle a_1, z\rangle - \alpha_1, \ldots, \langle a_l, z\rangle - \alpha_l\Big\} & \text{if } z \in \operatorname{dom}\theta, \\ \infty & \text{otherwise} \end{cases} \tag{4.1}$$

with some $l \in I\!N$, where the set $\operatorname{dom}\theta$ *is a convex polyhedron given by*

$$\operatorname{dom}\theta = \Big\{ z \in \mathbb{R}^m \Big| \ \langle d_i, z\rangle \le \beta_i \ \ \textit{for all} \ \ i \in T_2 := \{1, \ldots, p\} \Big\} \tag{4.2}$$

with some $d_i \in \mathbb{R}^n$, $\beta_i \in \mathbb{R}$, and $p \in I\!N$.

It follows from (4.1) that any $\theta \in CPWL$ can be represented in the *sum form*

$$\theta(z) = \max\Big\{ \langle a_1, z\rangle - \alpha_1, \ldots, \langle a_l, z\rangle - \alpha_l \Big\} + \delta(z; \operatorname{dom}\theta), \quad z \in \mathbb{R}^m, \tag{4.3}$$

where both summands are nonsmooth. Note that CPWL functions may be given in other forms different from (4.3), e.g., as the support function of a convex polyhedron

$$\theta(x) = \sigma_P(z) := \sup\{\langle h, z\rangle |\ h \in P\},$$

which is conjugate to the indicator function of P. Thus σ_P is CPWL by [53, Theorem 11.14(a)].

It is observed in [41, Proposition 3.2] that, besides (4.2), the domain of θ admits the representation $\operatorname{dom}\theta = \bigcup_{i=1}^{l} C_i$ with l taken from (4.1) and the sets C_i, $i \in T_1$, defined by

$$C_i := \Big\{ z \in \operatorname{dom}\theta \Big| \ \langle a_j, z\rangle - \alpha_j \le \langle a_i, z\rangle - \alpha_i, \ \ \text{for all} \ \ j \in T_1 \Big\}. \tag{4.4}$$

Consider next the corresponding active index subsets in (4.4) and (4.2) given by

$$K(\bar z) := \Big\{ i \in T_1 \Big| \ \bar z \in C_i \Big\} \ \ \text{and} \ \ I(\bar z) := \Big\{ i \in T_2 \Big| \ \langle d_i, \bar z\rangle = \beta_i \Big\} \tag{4.5}$$

and recall the formula for $\partial\theta(\bar z)$ at $\bar z \in \operatorname{dom}\theta$ obtained in [41, Proposition 3.3]:

$$\partial\theta(\bar z) = \operatorname{co}\Big\{ a_i \Big| \ i \in K(\bar z) \Big\} + N(\bar z; \operatorname{dom}\theta) = \operatorname{co}\Big\{ a_i \Big| \ i \in K(\bar z) \Big\} + \Big\{ d_i \Big| \ i \in I(\bar z) \Big\}. \tag{4.6}$$

Then for any $(\bar z, \bar v) \in \partial\theta$ we get from (4.6) that $\bar v = \bar v_1 + \bar v_1$, where

$$\bar v_1 = \sum_{i \in K(\bar z)} \bar\lambda_i a_i \ \ \text{with} \ \ \sum_{i \in K(\bar z)} \bar\lambda_i = 1, \ \bar\lambda_i \ge 0 \ \ \text{and} \ \ \bar v_2 = \sum_{i \in I(\bar z)} \bar\mu_i d_i \ \ \text{with} \ \ \bar\mu_i \ge 0. \tag{4.7}$$

Recall the well-known tangent cone representation

$$T(\bar{z};\operatorname{dom}\theta)=\Big\{x\in\mathbb{R}^n\Big|\ \langle d_i,\bar{x}\rangle\le 0\ \text{ for all }\ i\in I(\bar{z})\Big\} \tag{4.8}$$

for $\bar{z}\in\operatorname{dom}\theta$. Corresponding to (4.7), define the index subsets of positive multipliers by

$$J_+(\bar{z},\bar{v}_1):=\Big\{i\in K(\bar{z})\Big|\ \bar{\lambda}_i>0\Big\},\quad J_+(\bar{z},\bar{v}_2):=\Big\{i\in I(\bar{z})\Big|\ \bar{\mu}_i>0\Big\} \tag{4.9}$$

and then consider the following sets defined entirely via the parameters in (4.1) and (4.2) along arbitrary index subsets $P_1\subset Q_1\subset T_1$ and $P_2\subset Q_2\subset T_2$:

$$\begin{aligned}\mathcal{F}_{\{P_1,Q_1\},\{P_2,Q_2\}} :&= \operatorname{span}\Big\{a_i-a_j\ \Big|\ i,j\in P_1\Big\}\\ &+\Big\{a_i-a_j\ \Big|\ (i,j)\in(Q_1\setminus P_1)\times P_1\Big\}\\ &+\operatorname{span}\Big\{d_i\ \Big|\ i\in P_2\Big\}+\Big\{d_i\ \Big|\ i\in Q_2\setminus P_2\Big\},\end{aligned} \tag{4.10}$$

$$\begin{aligned}\mathcal{G}_{\{P_1,Q_1\},\{P_2,Q_2\}}:=\Big\{u\in\mathbb{R}^n\Big|\ &\langle a_i-a_j,u\rangle=0\ \text{ if }\ i,j\in P_1,\\ &\langle a_i-a_j,u\rangle\le 0\ \text{ if }\ (i,j)\in(Q_1\setminus P_1)\times P_1,\\ &\langle d_i,u\rangle=0\ \text{ if }\ i\in P_2,\ \text{ and }\ \langle d_i,u\rangle\le 0\ \text{ if }\ i\in Q_2\setminus P_2\ \Big\}.\end{aligned} \tag{4.11}$$

Now we are ready to formulate the precise calculation formulas for the second-order subdifferential of CPWL functions. In the notation above we have from [41, Theorem 5.1] that

$$\partial^2\theta(\bar{z},\bar{v})(u)=\Big\{w\Big|(w,-u)\in\mathcal{F}_{\{P_1,Q_1\},\{P_2,Q_2\}}\times\mathcal{G}_{\{P_1,Q_1\},\{P_2,Q_2\}},(P_1,Q_1,P_2,Q_2)\in\mathcal{A}\Big\} \tag{4.12}$$

for any $u\in\mathbb{R}^m$, where the set $\mathcal{A}$ of index quadruples is defined by

$$\begin{aligned}\mathcal{A}:=\Big\{(P_1,Q_1,P_2,Q_2)\Big|\ \ &P_1\subset Q_1\subset K,\ P_2\subset Q_2\subset I,\\ &(P_1,P_2)\in D(\bar{z},\bar{v}),\ H_{\{Q_1,Q_2\}}\ne\emptyset\Big\}\end{aligned} \tag{4.13}$$

with $K:=K(\bar{z})$, $I:=I(\bar{z})$, $H_{\{Q_1,Q_2\}}:=\{z\in\operatorname{dom}\theta\mid K(z)=Q_1,\ I(z)=Q_2\}$, and

$$D(\bar{z},\bar{v}):=\Big\{(P_1,P_2)\subset K\times I\ \Big|\ \bar{v}\in\operatorname{co}\{a_i|\ i\in P_1\}+\{d_i\mid i\in P_2\}\Big\}.$$

Furthermore, [41, Theorem 5.2] gives us the domain formula

$$\operatorname{dom}\partial^2\theta(\bar{z},\bar{v})=\Big\{u\,\Big|\,\langle a_i-a_j,u\rangle=0\ \text{ for }\ i,j\in\Gamma(J_1)\ \text{ and }\ \langle d_t,u\rangle=0\ \text{ for }\ t\in\Gamma(J_2)\Big\},\tag{4.14}$$

where the index sets $\Gamma(J_1)$ and $\Gamma(J_2)$ are defined by

$$\begin{aligned}\Gamma(J_1)&:=\Big\{i\in K\,\Big|\,\langle a_i-a_j,u\rangle=0\ \text{ for all }\ j\in J_1\ \text{ and }\ u\in\mathcal{G}_{\{J_1,K\},\{J_2,I\}}\Big\},\\ \Gamma(J_2)&:=\Big\{t\in I\,\Big|\,\langle d_t,u\rangle=0\ \text{ for all }\ u\in\mathcal{G}_{\{J_1,K\},\{J_2,I\}}\Big\}\end{aligned}\tag{4.15}$$

with the notation $J_1:=J_+(\bar{z},\bar{v}_1)$ and $J_2:=J_+(\bar{z},\bar{v}_2)$ built upon (4.7) and (4.9).

In the subsequent sections of this chapter, we will often consider compositions $\theta\circ\Phi$ of CPWL outer functions $\theta\colon\mathbb{R}^m\to\overline{\mathbb{R}}$ and inner mappings $\Phi\colon\mathbb{R}^n\times\mathbb{R}^d\to\mathbb{R}^m$ that are $\mathcal{C}^2$-smooth around some $(\bar{x},\bar{w})$ with $\bar{z}:=\Phi(\bar{x},\bar{w})\in\operatorname{dom}\theta$ under the first-order qualification condition

$$\partial^\infty\theta(\bar{z})\cap\ker\nabla_x\Phi(\bar{x},\bar{w})^*=\{0\}.\tag{4.16}$$

Such compositions form an important subclass of functions known as *fully amenable* in x at $\bar{x}$ with compatible parametrization by w at $\bar{w}$ (we will drop in what follows the latter parametrization expression for brevity), which are defined in this way with using more general convex *piecewise linear-quadratic* outer functions θ; see [53] for more details.

4.1 Reducibility, Nondegeneracy and Second-Order Qualification

The main goal of this section is to establish relationships between the second-order qualification condition introduced in [53] in order to derive the exact second-order chain rule for fully amenable compositions with CPWL outer functions and the partial nondegeneracy condition of a completely different nature that was employed in [28] to get the same second-order chain rule. In this way we obtain below some auxiliary results of their independent interest.

Considering first a fully amenable composition $\psi = \theta \circ \Phi$ as defined at the end of the last section, recall that the *second-order qualification condition* (SOCQ) holds for ψ in x at $(\bar{x}, \bar{w})$ if

$$\partial^2\theta(\bar{z}, v)(0) \cap \ker \nabla_x \Phi(\bar{x}, \bar{w})^* = \{0\} \;\text{ for all } v \in M(\bar{x}, \bar{w}, \bar{q}), \tag{4.17}$$

where $\bar{q} \in \partial_x \psi(\bar{x}, \bar{w})$ is a fixed partial subgradient of ψ in x at $(\bar{x}, \bar{w})$, and where

$$M(\bar{x}, \bar{w}, \bar{q}) := \Big\{ v \in \mathbb{R}^m \;\Big|\; v \in \partial\theta(\bar{z}) \;\text{ with }\; \nabla_x \Phi(\bar{x}, \bar{w})^* v = \bar{q} \Big\}. \tag{4.18}$$

Note that the imposed qualification condition (4.16) ensures, by using the well-known first-order subdifferential chain rule [27, 53], that $M(\bar{x}, \bar{w}, \bar{q}) \neq \emptyset$.

For a given $\theta \colon \mathbb{R}^n \to \overline{\mathbb{R}}$, denote by $S(z)$ a subspace of $\mathbb{R}^m$ parallel to the *affine hull* $\operatorname{aff} \partial\theta(z)$ of the subdifferential $\partial\theta(z)$, $z \in \mathbb{R}^m$. The next theorem provides a precise calculation of the second-order subdifferential for CPWL functions at the *origin* $0 \in \mathbb{R}^m$ entirely via the initial data in (4.1) and (4.2), relates it to the subspace $S(\bar{z})$ defined above, and gives an effective representation of SOQC in (4.17) convenient for our further analysis and applications.

Theorem 4.2 (second-order subdifferential of CPWL functions at the origin and SOQC representation). *Let $\psi = \theta \circ \Phi$ be a fully amenable composition of $\theta \colon \mathbb{R}^m \to \overline{\mathbb{R}}$ and $\Phi \colon \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}^m$ with $\theta \in CPWL$ and $\bar{z} = \Phi(\bar{x}, \bar{w})$, and let $S(\bar{z})$ be a subspace of $\mathbb{R}^m$ parallel to the affine hull $\operatorname{aff} \partial\theta(\bar{z})$. Then we following assertions hold:*

(i) *For all $\bar{v} \in \partial\theta(\bar{z})$ we have the representation*

$$\partial^2\theta(\bar{z}, \bar{v})(0) = \operatorname{span}\Big\{ a_i - a_j \;\Big|\; i, j \in K(\bar{z}) \Big\} + \operatorname{span}\Big\{ d_i \;\Big|\; i \in I(\bar{z}) \Big\} \tag{4.19}$$

via the data in (4.1) and (4.2) with the active index sets $K(\bar{z})$ and $I(\bar{z})$ defined in (4.5).

(ii) *Furthermore, we have $\partial^2\theta(\bar{z}, \bar{v})(0) = S(\bar{z})$ independently of $\bar{v} \in \partial\theta(\bar{z})$.*

(iii) *The SOQC property* (4.17) *can be equivalently written as*

$$S(\bar{z}) \cap \ker \nabla_x \Phi(\bar{x}, \bar{w})^* = \{0\} \tag{4.20}$$

independently of $\bar{v} \in \partial\theta(\bar{z})$ *and* $\bar{q} \in \partial_x \psi(\bar{x}, \bar{w})$ *in* (4.18).

Proof. To verify first the inclusion "$\subset$" in (4.19), pick $y \in \partial^2\theta(\bar{z}, \bar{v})(0)$ and find by (4.12) an index quadruple $(P_1, Q_1, P_2, Q_2) \in \mathcal{A}$ from (4.13) such that

$$(y, 0) \in \mathcal{F}_{\{P_1,Q_1\},\{P_2,Q_2\}} \times \mathcal{G}_{\{P_1,Q_1\},\{P_2,Q_2\}}.$$

We immediately deduce from representation (4.10) that

$$\mathcal{F}_{\{P_1,Q_1\},\{P_2,Q_2\}} \subset \text{span}\Big\{a_i - a_j \;\Big|\; i, j \in K(\bar{z})\Big\} + \text{span}\Big\{d_i \;\Big|\; i \in I(\bar{z})\Big\},$$

which justifies the inclusion "$\subset$" in (4.19). To derive further the opposite inclusion "$\supset$" therein, take any vector $y \in \text{span}\{a_i - a_j \mid i, j \in K(\bar{z})\} + \text{span}\{d_i \mid i \in I(\bar{z})\}$ and then put $P_1 = Q_1 := K(\bar{z})$ and $P_2 = Q_2 := I(\bar{z})$. Since $\bar{z} \in H_{\{Q_1,Q_2\}}$ in (4.19), it follows that $(P_1, Q_1, P_2, Q_2) \in \mathcal{A}$. Employing again the second-order formula (4.12) tells us that $(y, 0) \in N((\bar{z}, \bar{v}); \partial\theta)$ and hence yields $y \in \partial^2\theta(\bar{z}, \bar{v})(0)$, which thus verifies assertion (i).

To prove assertion (ii), observe that $S(\bar{z}) = \text{aff}\,\partial\theta(\bar{z}) - a_t$ for some $t \in K(\bar{z})$. Picking $y \in S(\bar{z})$ gives us $y + a_t = \sum_{i=1}^s \alpha_i c_i$ for some vectors $c_i \in \partial\theta(\bar{z})$ and some number $s > 0$ with $\sum_{i=1}^s \alpha_i = 1$. It follows from (4.6) that $c_i = c_{1i} + c_{2i}$ with $c_{1i} \in \text{co}\,\{a_r \mid r \in K(\bar{z})\}$ and $c_{2i} \in N(\bar{z}; \text{dom}\,\theta)$ for $i = 1, \ldots, s$. Therefore we arrive at the representation

$$y = \sum_{i=1}^{s} \alpha_i c_i - a_t = \sum_{i=1}^{s} \alpha_i (c_{1i} - a_t) + \sum_{i=1}^{s} \alpha_i c_{2i}. \tag{4.21}$$

It is clear that $c_{1i} - a_t \in \text{span}\{a_i - a_j \mid i, j \in K(\bar{z})\}$ by $c_{1i}, a_t \in \text{co}\,\{a_r \mid r \in K(\bar{z})\}$. Thus we get

$$c_i - a_t \in \text{span}\Big\{a_i - a_j \;\Big|\; i, j \in K(\bar{z})\Big\} + \text{span}\Big\{d_i \;\Big|\; i \in I(\bar{z})\Big\}.$$

Using this together with (4.21) and (4.19) justifies the inclusion $S(\bar{z}) \subset \partial^2\theta(\bar{z}, \bar{v})(0)$ in (ii).

To verify the opposite inclusion therein, take $y \in \partial^2\theta(\bar{z},\bar{v})(0)$ and express it by (4.19) as

$$y = \sum_{(i,j)\in A_1\times A_2} \alpha_{i,j}(a_i - a_j) + \sum_{t\in A_3} \beta_t d_t$$

with some index subsets $A_1, A_2 \subset K(\bar{z})$ and $A_3 \subset I(\bar{z})$. Select now $B_1, B_2 \subset A_1 \times A_2$ and $B_3, B_4 \subset A_3$ so that $A_1 \times A_2 = B_1 \cup B_2$, and $A_3 = B_3 \cup B_4$, and

$$\begin{aligned} &\alpha_{i,j} \ge 0 \ \text{ whenever } \ (i,j) \in B_1 \quad \text{and} \quad \alpha_{i,j} < 0 \ \text{ whenever } \ (i,j) \in B_2,\\ &\beta_t \ge 0 \quad \text{whenever} \quad t \in B_3 \qquad \text{and} \quad \beta_t < 0 \quad \text{whenever} \quad t \in B_4. \end{aligned} \tag{4.22}$$

In this way we represent the given vector y as $y = y' - b$ with

$$y' := \sum_{(i,j)\in B_1} \alpha_{i,j} a_i - \sum_{(i,j)\in B_2} \alpha_{i,j} a_j + \sum_{t\in B_3} \beta_t d_t \ \text{ and } \ b := \sum_{(i,j)\in B_1} \alpha_{i,j} a_j - \sum_{(i,j)\in B_2} \alpha_{i,j} a_i + \sum_{t\in B_4} (-\beta_t) d_t.$$

Denoting $\alpha := \sum_{(i,j)\in B_1} \alpha_{i,j} - \sum_{(i,j)\in B_2} \alpha_{i,j}$, deduce from (4.22) that $\alpha \ge 0$. For $\alpha > 0$ we get

$$\frac{1}{\alpha} y' \in \operatorname{aff} \partial\theta(\bar{z}) \ \text{ and } \ \frac{1}{\alpha} b \in \operatorname{aff} \partial\theta(\bar{z}). \tag{4.23}$$

It follows from the construction of $S(\bar{z})$ and the second inclusion in (4.23) that we have the equality $S(\bar{z}) = \operatorname{aff} \partial\theta(\bar{z}) - \frac{1}{\alpha} b$, and so the first one in (4.23) yields $\frac{1}{\alpha} y \in S(\bar{z})$. This shows that $y \in S(\bar{z})$ since $S(\bar{z})$ is a subspace, and thus we get $\partial^2\theta(\bar{z},\bar{v})(0) \subset S(\bar{z})$ in the case of $\alpha > 0$. Considering now the remaining case of $\alpha = 0$ gives us the expression $y = \sum_{t\in A_3} \beta_t d_t$, which implies by (4.22) that

$$y = a_t + \sum_{t\in B_3} \beta_t d_t - \Big(a_t + \sum_{t\in B_4} (-\beta_t) d_t\Big) \in S(\bar{z}) \ \text{ with some } \ t \in K(\bar{z})$$

due to $a_t + \sum_{t\in B_4} (-\beta_t) d_t \in \operatorname{aff} \partial\theta(\bar{z})$ and therefore verifies assertion (ii). This immediately implies condition (4.20) in (iii) by comparing it with the SOCQ definition in (4.17). $\triangle$

Note that while the precise calculation of $\partial^2\theta(\bar{z},\bar{v})(0)$ in Theorem 4.2(i) is new, assertion (ii) therein follows from the proof of Theorem 4.3 in [38] by using the representation of $\partial^2\theta(\bar{z},\bar{v})(0)$ for piecewise linear-quadratic functions $\theta\colon \mathbb{R}^m \to \overline{\mathbb{R}}$ established in [38, Theorem 4.1]. The proof of the latter result in [38] is based on the tangential approach from [53] being significantly more involved in comparison with the one given above.

It is also worth mentioning as a by product of the above calculations that the validity of SOQC for fully amenable compositions with CPWL outer functions yields the fulfillment of the first-order qualification condition (4.16) in the definition of such compositions. To see this, recall that $\partial^\infty\theta(\bar{z}) = N(\bar{z}; \operatorname{dom}\theta)$ for convex functions and thus get the inclusion $\partial^\infty\theta(\bar{z}) \subset \partial^2\theta(\bar{z}, \bar{v})(0)$ whenever $\bar{v} \in \partial\theta(\bar{z})$ by comparing (4.19) with that of $N(\bar{z}; \operatorname{dom}\theta) = \{d_i : \; i \in I(\bar{z})\}$.

Next we consider the concept of *nondegeneracy*. It was first initiated for *sets* in [48] as a polyhedral counterpart of the classical linear independence constraint qualification (LICQ) in nonlinear programming. Note that even for mathematical programs with equilibrium constraints (MPPCs) this nondegeneracy condition may be strictly weaker than LICQ; see [39] for equivalent descriptions for MPPCs and particularly Example 6.7 therein. Nondegeneracy and associated *reducibility* notions for general sets were comprehensively studied in [1] based on the previous paper of these authors. For the case of extended-real-valued *functions* the notion of $\mathcal{C}^2$-reducibility and the corresponding notion of partial nondegeneracy was formulated in [28] in order to derive the aforementioned second-order subdifferential chain rule; see below.

Following this pattern, we say that a function $\theta\colon \mathbb{R}^m \to \overline{\mathbb{R}}$ is *$\mathcal{C}^2$-reducible* (resp. *$\mathcal{C}^\infty$-reducible*) to a function $\vartheta\colon \mathbb{R}^s \to \overline{\mathbb{R}}$ at $\bar{z}$ with $s \le m$ if there exists a $\mathcal{C}^2$-smooth (resp. $\mathcal{C}^\infty$-smooth) mapping $h\colon \mathbb{R}^m \to \mathbb{R}^s$ with the surjective derivative $\nabla h(\bar{z})$ such that $\theta(z) = (\vartheta \circ h)(z)$ for all z around $\bar{z}$.

Our next result shows that any function $\theta \in CPWL$ on $\mathbb{R}^m$ is $\mathcal{C}^\infty$-reducible to some function $\vartheta \in CPWL$ on $\mathbb{R}^s$ by using actually a *linear* surjective operator $h\colon \mathbb{R}^m \to \mathbb{R}^s$. From now on we assume that $0 \in \operatorname{aff}\partial\theta(\bar{z})$ at $\bar{z} \in \operatorname{dom}\theta$, which tells us that $S(\bar{z}) = \operatorname{aff}\partial\theta(\bar{z})$.

In fact this assumption does *not restrict the generality* in dealing with the second-order subdifferential. Indeed, we always have $S(\bar{z}) = \text{aff}\,\partial\theta(\bar{z}) - b_{\bar{z}}$ for some $b_{\bar{z}} \in \text{aff}\,\partial\theta(\bar{z})$. Defining then $\bar{\theta}(z) := \theta(z) - \langle b_{\bar{z}}, z\rangle$ shows that $0 \in \text{aff}\,\partial\bar{\theta}(z)$ and $\partial^2\theta(\bar{z},\bar{y}) = \partial^2\bar{\theta}(\bar{z},\bar{y} - b_{\bar{z}})$ for any $\bar{v} \in \partial\theta(\bar{z})$.

Lemma 4.3 ($\mathcal{C}^\infty$-reducibility of piecewise linear functions). *Let $\theta\colon \mathbb{R}^m \to \overline{\mathbb{R}}$ be CPWL, let $\bar{z} \in \text{dom}\,\theta$, and let $s := \dim S(\bar{z}) \le m$. Then θ is $\mathcal{C}^\infty$-reducible at $\bar{z}$ to a CPWL function $\vartheta\colon \mathbb{R}^s \to \overline{\mathbb{R}}$ via a linear operator $h(z) := Bz$ generated by some $s \times m$ matrix B.*

Proof. It follows from [41, Proposition 3.3(i)] that $\partial\theta(z) \subset \partial\theta(\bar{z})$ for all $z \in O$ in some neighborhood of $\bar{z}$. Denote by A the matrix of a linear isometry from $\mathbb{R}^m$ into $\mathbb{R}^s \times \mathbb{R}^{m-s}$ under which $A^*(S(\bar{z})) = \mathbb{R}^s \times \{0\}$. Define the function $\xi\colon \mathbb{R}^m \to \overline{\mathbb{R}}$ by

$$\xi(y) := \theta(Ay) \ \text{ for all } \ y \in \mathbb{R}^m \tag{4.24}$$

and get by [53, Proposition 3.55(b)] that ξ is proper, convex, and piecewise linear on $\mathbb{R}^m$. Applying the first-order chain rule of convex analysis to (4.24) gives us

$$\partial\xi(y) = A^*\partial\theta(z) \ \text{ with } \ Ay = z. \tag{4.25}$$

Denote $U := A^{-1}(O)$ and deduce from the classical open mapping theorem that U is a neighborhood of $\bar{y} := A^{-1}\bar{z}$. Suppose that $\alpha > 0$ is so small that $I\!B_\alpha(\bar{y}) \subset U$ for the ball centered at $\bar{y}$ with radius α and put $O' := A(\text{int}\,I\!B_\alpha(\bar{y}))$, which is a neighborhood of $\bar{z}$ by the open mapping theorem. Then $S(z) = \text{aff}\,\partial\theta(z) + b_z$ with some $b_z \in \mathbb{R}^m$ for each $z \in O$, and furthermore $b_{\bar{z}} = 0$ as discussed before the formulation of the lemma. This tells us by the above relationships that

$$v = (v_1, \ldots, v_m) \in \partial\xi(y) = A^*\partial\theta(z) \subset A^*\partial\theta(\bar{z}) \subset A^*(S(\bar{z})) - A^*b_{\bar{z}} \subset \mathbb{R}^s \times \{0\} \tag{4.26}$$

for all $y \in U$, which implies that the last $m - s$ elements of any $v \in \partial\xi(y)$ are zeros whenever $y \in U$. Construct now the desired $s \times m$ matrix B claimed in the lemma from the $m \times m$

matrix A^{-1} by deleting the last $m-s$ rows of the latter. We define the corresponding function $\vartheta\colon \mathbb{R}^s \to \overline{\mathbb{R}}$ by using ξ in (4.24) as follows: take $y = (y_s, y_{m-s}) = (x, y_{m-s}) \in \mathbb{R}^s \times \mathbb{R}^{m-s}$ and put

$$\vartheta(x) := \xi(x, \bar{y}_{m-s}) = \xi(y_s, \bar{y}_{m-s}) \text{ for all } x \in \mathbb{R}^s, \tag{4.27}$$

where $\bar{y}_{m-s}$ is the last $m-s$ elements of the vector $\bar{y} = A^{-1}\bar{z}$. Since ξ is proper, so is the function ϑ in (4.27). It is easy to see that ϑ is piecewise linear and the convexity of ξ implies the convexity of ϑ. To justify the statement of the lemma, it remains to verify the representation

$$\theta(z) = (\vartheta \circ B)(z) \text{ for all } z \in O'. \tag{4.28}$$

Let us do it by observing first that $y \in \operatorname{int} I\!B_\alpha(\bar{y})$ whenever $y = A^{-1}z$ generated by $z \in O'$. It follows from (4.24), (4.27), and the definition of B that $(\vartheta \circ B)(z) = \xi(y_s, \bar{y}_{m-s})$ in the notation above, where $(y_s, \bar{y}_{n-s}) \in \operatorname{int} I\!B_\alpha(\bar{y})$. Thus (4.28) would be implied by the relationship

$$\xi(y_s, \bar{y}_{m-s}) = \xi(y_s, y_{m-s}) \text{ for any } y = (y_s, y_{m-s}) = A^{-1}z, \quad z \in O'. \tag{4.29}$$

Since (4.29) is trivial when both values $\xi(y_s, y_{m-s})$ and $\xi(y_s, \bar{y}_{m-s})$ are infinite, suppose without loss of generality that $\xi(y_s, y_{n-s})$ is a real number. The polyhedrality of $\operatorname{epi}\xi$ ensures that the function ξ is l.s.c., and hence we can apply to it the approximate mean value inequality from [27, Corollary 3.50]. This allows us to find a point $c \in \mathbb{R}^m$ on the segment connecting (y_s, y_{m-s}) and $(y_s, \bar{y}_{m-s})$ as well as a sequence $v_k \in \partial\xi(u_k)$ with $u_k \to c$ and $\xi(u_k) \to \xi(c)$ so that

$$\xi(y_s, \bar{y}_{m-s}) - \xi(y_s, y_{m-s}) \le \liminf_{k\to\infty} \langle v_k, (0_s, \bar{y}_{m-s} - y_{m-s})\rangle. \tag{4.30}$$

It follows from (4.26) that $u_k \in \operatorname{int} I\!B_\alpha(\bar{y}) \subset U$ and so $\langle v_k, (0_s, \bar{y}_{m-s} - y_{m-s})\rangle = 0$ for all

$k \in I\!N$ sufficiently large. In the same way we get the opposite inequality

$$\xi(y_s, y_{m-s}) - \xi(y_s, \bar{y}_{m-s}) \le 0$$

and combining the latter with (4.30) arrive at (4.29), which completes the proof. △

Now we are ready to formulate, following [28], the notion of *nondegeneracy* of one mapping relative to another one used for deriving the second-order chain rule. Observe that, although this notion is formulated for two arbitrary mappings, its application to second-order analysis mainly concerns amenable compositions $\theta \circ \Phi$ of $\theta\colon \mathbb{R}^m \to \overline{\mathbb{R}}$ and $\Phi\colon \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}^m$ while defining nondegenerate points of $\Phi\colon \mathbb{R}^n \times \mathbb{R}^d$ relative to the mapping $h\colon \mathbb{R}^m \to \mathbb{R}^s$ that furnishes the appropriate *reducibility* of the outer function θ. Thus in our case of $\theta \in CPWL$ we deal with *linear* mapping $h(z) = Bz$ that appears in the $\mathcal{C}^\infty$-reducibility assertion of Lemma 4.3.

Having this in mind, it is said that $(\bar{x}, \bar{w}) \in \mathbb{R}^n \times \mathbb{R}^d$ is a *partial nondegenerate point* of $\Phi\colon \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}^m$ in x relative to $h\colon \mathbb{R}^m \to \mathbb{R}^s$ if

$$\nabla_x \Phi(\bar{x}, \bar{w})\mathbb{R}^n + \ker \nabla h(\bar{z}) = \mathbb{R}^m \quad \text{with} \quad \bar{z} = \Phi(\bar{x}, \bar{w}) \tag{4.31}$$

under the corresponding differentiability assumptions on Φ and h. The next theorem based on the previous results of this section reveals that, in the case of fully amenable compositions with CPWL outer functions, the SOQC property (4.20) of $\theta \circ \Phi$ is *equivalent* to the nondegeneracy condition (4.31) *provided* that the mapping $h\colon \mathbb{R}^m \to \mathbb{R}^s$ with $s = \dim S(\bar{z})$ therein is the linear transformation $h(z) = Bz$ constructed in Lemma 4.3 to realize the $\mathcal{C}^\infty$-reducibility of θ.

It is worth mentioning that this line of equivalency between the corresponding SOQC and nondegeneracy properties is a continuation of the results previously established in [39] in connection with mathematical programs with polyhedral constraints and in [37] in connection with second-order cone programs (SOCPs), where (in both cases) the nondegeneracy condition of a mapping relative to the underlying set (polyhedron and second-order cone, respectively) was understood in the sense of [1] via the tangent cone to this set. The crucial difference of our case is that we implement the general nondegeneracy/reducibility notion [28] relative to a mapping and emphasize the *linearity* of this mapping in the CPWL setting under consideration.

Theorem 4.4 (relationship between SOQC and nondegeneracy for fully amenable compositions with CPWL outer functions). *Let $\psi = \theta \circ \Phi$ be a fully amenable composition finite at $(\bar{x}, \bar{w}) \in \mathbb{R}^n \times \mathbb{R}^d$, let $\theta\colon \mathbb{R}^m \to \overline{\mathbb{R}}$ be CPWL, and let B be an $s \times m$ matrix constructed in Lemma* 4.3*. Then the SOQC property* (4.20) *holds at $(\bar{x}, \bar{w})$ if and only if this point is partially nondegenerate* (4.31) *for Φ relative to $h(z) = Bz$ with $s = \dim S(\bar{z})$.*

Proof. Since $0 \in \operatorname{aff} \partial\theta(\bar{z})$ as discussed before the formulation of Lemma 4.3, we have $S(\bar{z}) = \operatorname{aff} \partial\theta(\bar{z})$. This lemma gives us a CPWL function $\vartheta\colon \mathbb{R}^s \to \overline{\mathbb{R}}$ and a mapping $h(z) = Bz$ from $\mathbb{R}^m$ to $\mathbb{R}^s$ such that $\theta(z) = (\vartheta \circ h)(z)$ for all $z \in \mathbb{R}^m$ sufficiently close to $\bar{z}$. Assuming that SOQC holds at $(\bar{x}, \bar{w})$ and taking the orthogonal complements of both sides in (4.20), we arrive at

$$\nabla_x \Phi(\bar{x}, \bar{w})\mathbb{R}^n + S(\bar{z})^\perp = \mathbb{R}^m. \tag{4.32}$$

To deduce from (4.32) the partial nondegeneracy condition (4.31) with $h(z) = Bz$, it suffices to show that $\ker \nabla h(\bar{z}) = S(\bar{z})^\perp$, which reads as $\ker B = S(\bar{z})^\perp$. Indeed, picking $u \in \ker B$ and taking into account that $A^*(S(\bar{z})) = \mathbb{R}^s \times \{0\}$ in the proof of the lemma yield

$$0 = \langle A^{-1}u, A^*p\rangle = \langle u, (A^{-1})^* A^* p\rangle = \langle u, p\rangle \quad \text{for any} \quad p \in S(\bar{z}),$$

which tells us that $u \in S(\bar{z})^\perp$, and so $\ker B \subset S(\bar{z})^\perp$. The opposite inclusion $S(\bar{z})^\perp \subset \ker B$ can be checked similarly, which shows therefore that SOQC$\Longrightarrow$partial nondegeneracy. The same arguments allow us to verify via (4.32) the reverse implication partial nondegeneracy$\Longrightarrow$SOQC and thus complete the proof of the theorem. $\triangle$

The final result of this section presents the second-order chain rule for the *partial* second-order subdifferential (denoted below as $D^*\partial_x\psi$) of fully amenable compositions $\psi = \theta\circ\Phi$ with CPWL outer functions. This result was first obtained in [38, Theorem 4.3] for nonparametric compositions and then in [39, Theorem 4.1] in the general parametric case. Both proofs in [38, 39] are involved, being based on the difficult Theorem 4.1 from [38]. The new proof given below is much simpler based on the equivalency result of Theorem 4.4 and the second-order chain rule obtained in [28, Theorem 3.6] under nondegeneracy condition in the Banach setting.

Corollary 4.5 (second-order chain rule for parametric compositions with CPWL outer functions). *Let $\psi = \theta \circ \Phi$ be a fully amenable composition with a CPWL outer function $\theta\colon \mathbb{R}^m \to \overline{\mathbb{R}}$ and an inner mapping $\Phi\colon \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}^m$ that is $\mathcal{C}^2$-smooth around $(\bar{x}, \bar{w})$. Then the validity of SOQC in* (4.20) *ensures that for any $\bar{q} \in \partial_x\psi(\bar{x}, \bar{y})$ the set $M(\bar{x}, \bar{w}, \bar{q})$ from* (4.18) *is a singleton $\{\bar{v}\}$ and we have the following second-order chain rule whenever $u \in \mathbb{R}^n$:*

$$\begin{aligned}(D^*\partial_x\psi)(\bar{x}, \bar{w}, \bar{q})(u) &= \Big(\nabla^2_{xx}\langle\bar{v}, \Phi\rangle(\bar{x}, \bar{w})u, \nabla^2_{xw}\langle\bar{v}, \Phi\rangle(\bar{x}, \bar{w})u\Big)\\ &\quad + \Big(\nabla_x\Phi(\bar{x}, \bar{w}), \nabla_w\Phi(\bar{x}, \bar{w})\Big)^*\partial^2\theta(\bar{z}, \bar{q})(\nabla_x\Phi(\bar{x}, \bar{w})u).\end{aligned} \tag{4.33}$$

Proof. For any $\bar{v}_1, \bar{v}_2 \in M(\bar{x}, \bar{w}, \bar{q})$ we have $\bar{v}_1 - \bar{v}_2 \in \ker\nabla_x\Phi(\bar{x}, \bar{w})^*$. Since furthermore $\bar{v}_1, \bar{v}_2 \in \partial\theta(\bar{z})$, it follows from (4.6) and Theorem 4.2(i) that $\bar{v}_1, \bar{v}_2 \in S(\bar{z})$. Applying now SOCQ (4.20) gives us $\bar{v}_1 = \bar{v}_2$, and so $M(\bar{x}, \bar{w}, \bar{q}) = \{\bar{v}\}$. Then we get from Lemma 4.3 that θ is $\mathcal{C}^\infty$-reducible by the linear mapping $h(z) = Bz$, and hence $(\bar{x}, \bar{w})$ is a partial nondegenerate point (4.31) of Φ relative to this mapping $h\colon \mathbb{R}^m \to \mathbb{R}^s$ with $s = \dim S(\bar{z})$. To arrive finally at the chain rule (4.33), it remains to apply [28, Theorem 3.6] and thus complete the proof. $\triangle$

Note that Corollary 4.5 clarifying the meaning of [38, Theorem 4.3] and [39, Theorem 4.3] can be viewed as a realization of the second-order chain rule from [28, Theorem 3.6] in the case of CPWL outer functions under the fulfillment of SOQC, which corresponds to a *linear reduction mapping* $h\colon \mathbb{R}^n \to \mathbb{R}^s$ in the nondegeneracy condition (4.31). The result of the latter theorem justifies the validity of (4.33) under (4.31) when h is merely a $\mathcal{C}^2$*-smooth* mapping that furnishes the required reducibility of θ.

4.2 Full Stability in Composite Optimization

In this section we proceed with applications of second-order generalized differentiation to problems of *composite optimization* given in the form:

$$\text{minimize } \ \varphi_0(x) + \theta(\Phi(x)) \ \text{ subject to } \ x \in \mathbb{R}^n \ \text{ with } \ \Phi(x) := \big(\varphi_1(x), \ldots, \varphi_m(x)\big), \tag{4.34}$$

where $\theta\colon \mathbb{R}^m \to \overline{\mathbb{R}}$ is a CPWL extended-real-valued function, and where all $\varphi_i\colon \mathbb{R}^n \to \mathbb{R}$, $i = 0, \ldots, m$, are $\mathcal{C}^2$-smooth around the reference optimal solution. This class of problems encompasses conventional problems of nonlinear programming (NLPs) as well as constrained and unconstrained minimax problems. It also includes the following major subclass of *extended nonlinear programs* (ENLPs) introduced in [52]:

$$\text{minimize } \ \varphi_0(x) + (\theta \circ \Phi)(x) \ \text{ with } \ \theta(z) := \sup_{p \in P} \langle p, z \rangle, \quad x \in \mathbb{R}^n, \tag{4.35}$$

where P is a convex polyhedron and thus θ in (4.35) is piecewise linear; see [41] for more details. Consider now the two-parametric version of (4.34) constructed by

$$\mathcal{P}^{\text{com}}(w, v): \quad \text{minimize } \ \varphi_0(x, w) + \theta(\Phi(x, w)) - \langle v, x \rangle \ \text{ subject to } \ x \in \mathbb{R}^n, \tag{4.36}$$

where the perturbed functions $\varphi_0(x, w)$ and $\Phi(x, w) = (\varphi_1(x, w), \ldots, \varphi_m(x, w))$ are $\mathcal{C}^2$-smooth with respect to both variables. Denote

$$\varphi(x, w) := \varphi_0(x, w) + \theta(\Phi(x, w)) \ \text{ for } \ (x, w) \in \mathbb{R}^n \times \mathbb{R}^d \tag{4.37}$$

and then fix a number $\gamma > 0$ and a triple $(\bar{x}, \bar{w}, \bar{v})$ with $\Phi(\bar{x}, \bar{w}) \in \operatorname{dom}\theta$ and $\bar{v} \in \partial_x \varphi(\bar{x}, \bar{w})$.

In this section we establish new second-order characterizations of full stability for local optimal solutions to problems of composite optimization (4.36) with CPWL outer functions therein. In particular, the results established below cover those in [39, 40] while being independent from characterizations obtained in [31–33, 37] for optimization and variational problems that cannot be represented in the composite form (4.34) with a CPWL outer function θ.

To proceed, denote $\bar{z} := \Phi(\bar{x}, \bar{w}) \in \operatorname{dom}\theta$ and recall from Lemma 4.3 that θ is reducible at $\bar{z}$ to some CPWL function $\vartheta\colon \mathbb{R}^s \to \overline{\mathbb{R}}$ with $s = \dim S(\bar{z})$ by using a linear mapping $h(z) = Bz$ with the $s \times m$ matrix B constructed in that lemma. Thus we have the representation $\theta(z) = (\vartheta \circ B)(z)$ for all z near $\bar{z}$ generating the mapping $\Psi(x, w) := (B \circ \Phi)(x, w)$. This tells us that the problem $\mathcal{P}^{\text{com}}(w, v)$ from (4.36) is locally equivalent around $(\bar{x}, \bar{w})$ to the following *reduced* problem:

$$\mathcal{P}_r^{\text{com}}(w, v): \quad \text{minimize} \;\; \varphi_0(x, w) + \vartheta(\Psi(x, w)) - \langle v, x\rangle \;\; \text{subject to} \;\; x \in \mathbb{R}^n. \tag{4.38}$$

We will see below that the reduced problem (4.38) is very instrumental in deriving the explicit second-order characterization of full stability of local minimizers in composite optimization obtained in this section as well as other important results established in the subsequent sections of the paper. The main assumption we need in what follows is the following *nondegeneracy condition* discussed in Section 4.1:

ND: A pair $(\bar{x}, \bar{w})$ is a partial nondegenerate point (4.31) of Φ from (4.36) in x relative to the linear mapping $h(z) = Bz$, where B is the $s \times m$ matrix constructed in the proof of Lemma 4.3 with $s = \dim S(\bar{z})$, $\bar{z} = \Phi(\bar{x}, \bar{w})$.

We know from Theorem 4.4 that condition ND is equivalent to the SOCQ property (4.20) in the framework of the composite optimization problem (4.36).

The next proposition is a composite optimization counterpart of Proposition 3.1 obtained for constrained optimization problems with $\theta = \delta_\Theta$, the indicator function of a $\mathcal{C}^2$-reducible closed and convex set Θ.

Proposition 4.6 (full stability and nondegeneracy in the original and reduced problems). *Let $\bar{x}$ be a feasible solution to $\mathcal{P}^{com}(\bar{w}, \bar{v})$ from* (4.36) *for the parameter pair $(\bar{w}, \bar{v}) \in \mathbb{R}^d \times \mathbb{R}^n$. Then $\bar{x}$ is a fully stable locally optimal solution to $\mathcal{P}^{com}(\bar{w}, \bar{v})$ if and only if it is a fully stable locally optimal solution to the reduced problem ${\mathcal{P}_r}^{com}(\bar{w}, \bar{v})$. Furthermore, the validity of* ND *for $(\bar{x}, \bar{w})$ implies the surjectivity* (*full rank*) *of the partial Jacobian matrix $\nabla_x \Psi(\bar{x}, \bar{w})$, where $\Psi = B \circ \Phi$.*

Proof. The claimed equivalence for full stability follows directly from the above observation that problems $\mathcal{P}^{\text{com}}(w, v)$ and $\mathcal{P}_r^{\text{com}}(w, v)$ are locally the same. Let us verify the part of the proposition concerning nondegeneracy. Supposing that ND holds gives us

$$\nabla_x \Phi(\bar{x}, \bar{w})\mathbb{R}^n + \ker \nabla h(\bar{z}) = \nabla_x \Phi(\bar{x}, \bar{w})\mathbb{R}^n + \ker B = \mathbb{R}^m.$$

It yields by applying the classical chain rule that

$$\nabla_x \Psi(\bar{x}, \bar{w})\mathbb{R}^n = B\nabla_x \Phi(\bar{x}, \bar{w})\mathbb{R}^n = B\big(\nabla_x \Phi(\bar{x}, \bar{w})\mathbb{R}^n + \ker B\big) = B\mathbb{R}^m = \mathbb{R}^s,$$

which justifies the surjectivity of $\nabla_x \Psi(\bar{x}, \bar{w})$ and thus completes the proof. $\triangle$

Recall that the equivalence between ND and SOCQ implies that the first-order qualification condition (4.16) automatically holds under ND; see the discussion after the proof of Theorem 4.2. This ensures, by the well-known first-order subdifferential chain rule, that the stationary condition $\bar{v} \in \partial_x \varphi(\bar{x}, \bar{w})$ via φ from (4.37) can be equivalently written as

$$\bar{v} \in \nabla_x \varphi_0(\bar{x}, \bar{w}) + \nabla_x \Phi(\bar{x}, \bar{w})^* \partial\theta(\Phi(\bar{x}, \bar{w})). \tag{4.39}$$

This allows us to consider the corresponding KKT system for problem $\mathcal{P}^{\text{com}}(w, v)$ given in the form

$$\begin{cases} v = \nabla_x L(x, w, \lambda), \quad \lambda \in \partial\theta(\Phi(x, w)) \\ \text{with } \ L(x, w, \lambda) := \varphi_0(x, w) + \langle \lambda, \Phi(x, w) \rangle. \end{cases} \tag{4.40}$$

Similarly, the KKT system for the reduced problem $\mathcal{P}_r^{\text{com}}(w, v)$ from (4.38) is given by

$$\begin{cases} v = \nabla_x L_r(x, w, \mu), \quad \mu \in \partial\vartheta(\Psi(x, w)) \\ \text{with } \ L_r(x, w, \mu) := \varphi_0(x, w) + \langle \mu, \Psi(x, w) \rangle. \end{cases} \tag{4.41}$$

It is not hard to observe from the reducibility $\theta(z) = (\vartheta \circ B)(z)$ around $\bar{z}$ together with the full rank property of B that Lagrange multipliers λ of (4.40) and μ of (4.41) are related by $\lambda = B^*\mu$. The next proposition establishes the uniqueness of solutions to (4.40) under the validity of ND. It is a composite optimization counterpart of [1, Proposition 4.75] in optimization problems with constraints $\Phi(x, z) \in \Theta$ under the corresponding reducibility and nondegeneracy conditions.

Proposition 4.7 (uniqueness of Lagrange multipliers for composite problems under ND). *Let $\bar{x}$ be a feasible solution to $\mathcal{P}^{com}(\bar{w}, \bar{v})$ for the parameter pair $(\bar{w}, \bar{v})$ with $\bar{v}$ from* (4.39) *and $(\bar{z}, \bar{v}) \in \partial\theta$, let* ND *hold, and let $\theta \in CPWL$. Then the set of Lagrange multipliers*

$$\left\{ \bar{\lambda} \in \partial\theta(\Phi(\bar{x}, \bar{w})) \;\middle|\; \bar{v} = \nabla_x L(\bar{x}, \bar{w}, \bar{\lambda}) \right\} \tag{4.42}$$

for the KKT system (4.40) *is singleton.*

Proof. Pick two vectors λ_1, λ_2 from set (4.42). It follows from the structure of (4.42) and the subdifferential description (4.6) for CPWL functions that $\lambda_1 - \lambda_2 \in \ker \nabla_x \Phi(\bar{x}, \bar{w})^*$ and

$$\lambda_s = \sum_{i \in K(\bar{z})} \eta_{si} a_i + \sum_{i \in I(\bar{z})} \tau_{si} d_i \ \text{ with } \ \sum_{i \in K(\bar{z})} \eta_{si} = 1, \ \ \eta_{si}, \tau_{si} \geq 0 \ \text{ for } \ s = 1, 2.$$

Then employing assertions (i) and (ii) of Theorem 4.2, we get

$$\begin{aligned}\lambda_1-\lambda_2 &= \sum_{i\in K(\bar z)}\eta_{1i}a_i+\sum_{i\in I(\bar z)}\tau_{1i}d_i-\sum_{i\in K(\bar z)}\eta_{2i}a_i-\sum_{i\in I(\bar z)}\tau_{2i}d_i\\ &= \sum_{j\in K(\bar z)}\eta_{2j}\sum_{i\in K(\bar z)}\eta_{1i}(a_i-a_j)+\sum_{i\in I(\bar z)}\tau_{1i}d_i-\sum_{i\in I(\bar z)}\tau_{2i}d_i\in S(\bar z)\end{aligned}$$

thus showing that $\lambda_1=\lambda_2$ by SOCQ (4.20), which is equivalent to ND. △

Now we in a position to introduce a new second-order condition formulated entirely via the initial data of the composite optimization problem (4.36) and then to show that this condition provides a complete characterization of full stability of local minimizers therein under the validity of ND. This condition is crucial in stability issues for composite optimization playing here the role similar to the *strong second-order sufficient condition* (SSOSC) in the sense of Robinson [47] for classical NLPs; so we keep this name in what follows while just adding “composite."

Definition 4.8 (composite SSOSC). *Given* $(\bar x,\bar w,\bar v,\bar\lambda)\in\mathbb{R}^n\times\mathbb{R}^d\times\mathbb{R}^n\times\mathbb{R}^m$ *with* $\bar v$ *satisfying* (4.39) *and* $\bar\lambda$ *satisfying* (4.40)*, we say that the* COMPOSITE SSOSC *holds at this point if*

$$\langle u,\nabla^2_{xx}L(\bar x,\bar w,\bar\lambda)u\rangle>0\ \ \textit{for all}\ \ 0\neq u\in\mathcal{S},\tag{4.43}$$

where L *is the Lagrangian for* (4.36) *taken from* (4.40)*, and where the subspace* $\mathcal{S}$ *is defined by*

$$\mathcal{S}:=\Big\{u\in\mathbb{R}^n\ \Big|\ \begin{array}{l}\langle a_i-a_j,\nabla_x\Phi(\bar x,\bar w)u\rangle=0\ \ \textit{for}\ \ i,j\in\Gamma(J_1),\\ \langle d_t,\nabla_x\Phi(\bar x,\bar w)u\rangle=0\ \ \textit{for}\ \ t\in\Gamma(J_2)\end{array}\Big\}\tag{4.44}$$

via the index sets $\Gamma(J_1)$ *and* $\Gamma(J_2)$ *taken from* (4.15).

Observe the following description of the subspace (4.44) of the positive definiteness of the Lagrangian Hessian in the composite SSOSC:

$$u\in\mathcal{S}\Longleftrightarrow\nabla_x\Phi(\bar x,\bar w)u\in\operatorname{dom}\partial^2\theta(\bar x,\bar v),\tag{4.45}$$

which is implied by (4.14) and reveals the second-order nature of this subspace. The composite SSOSC reduces to Robinson's SSOSC for NLPs by putting $\Gamma(J_1) = \emptyset$ and $\Gamma(J_2) = J_2$ in (4.43) and (4.44). Accordingly, it reduces to [39, Definition 6.4] and [39, Definition 7.2] in the corresponding settings of MPPCs and ENLPs, respectively.

The next lemma is important, together with the second-order subdifferential chain rule, for deriving the aforementioned characterization of full stability of local minimizers in (4.36).

Lemma 4.9 (second-order subdifferential property of CPWL functions). *Take a pair* $(\bar{z}, \bar{v}) \in \theta$ *for a CPWL function* θ. *Then we have* $0 \in \partial^2\theta(\bar{z}, \bar{v})(u)$ *whenever* $u \in \operatorname{dom} \partial^2\theta(\bar{z}, \bar{v})$.

Proof. Pick $u \in \operatorname{dom} \partial^2\theta(\bar{z}, \bar{v})$ and find $w \in \partial^2\theta(\bar{z}, \bar{v})(u)$, which by the coderivative definition (2.8) means $(w, -u) \in N((\bar{z}, \bar{v}), \partial\theta)$. Applying now formula (4.12) gives us a quadruple $(P_1, Q_1, P_2, Q_2) \in \mathcal{A}$ such that

$$w \in \mathcal{F}_{\{P_1,Q_1\},\{P_2,Q_2\}} \quad \text{and} \quad -u \in \mathcal{G}_{\{P_1,Q_1\},\{P_2,Q_2\}}.$$

Since we always have $0 \in \mathcal{F}_{\{P_1,Q_1\},\{P_2,Q_2\}}$, it follows that

$$(0, -u) \in \mathcal{F}_{\{P_1,Q_1\},\{P_2,Q_2\}} \times \mathcal{G}_{\{P_1,Q_1\},\{P_2,Q_2\}},$$

which implies by (4.12) the claimed inclusion $0 \in \partial^2\theta(\bar{z}, \bar{v})(u)$. △

We now proceed with establishing the main result of this section, which provides a complete characterization of fully stable local minimizers of $\mathcal{P}^{\text{com}}(\bar{w}, \bar{v})$ entirely via the initial data.

Theorem 4.10 (second-order characterization of full stability in composite optimization). *Let* $\bar{x}$ *be a feasible solution to* $\mathcal{P}^{\text{com}}(\bar{w}, \bar{v})$ *from* (4.36) *for the parameter pair* $(\bar{w}, \bar{v})$ *with* $\bar{v}$ *from* (4.39), *let* $\theta \in CPWL$, *and let* $(\bar{z}, \bar{v}) \in \partial\theta$ *with* $\bar{z} = \Phi(\bar{x}, \bar{v})$. *Under the validity of ND, let* $\bar{\lambda}$ *be a unique solution of the KKT system* (4.40). *Then* $\bar{x}$ *is a fully stable local minimizer of* $\mathcal{P}^{\text{com}}(\bar{w}, \bar{v})$ *if and only if the composite SSOSC from Definition* 4.8

is satisfied.

Proof. If $\bar{x}$ is a fully stable local minimizer of $\mathcal{P}^{\rm com}(\bar{w},\bar{v})$, then it is also a fully stable local minimizer of the reduced problem $\mathcal{P}_r^{\rm com}(\bar{w},\bar{v})$ by Proposition 4.6. It follows from this proposition that the partial Jacobian matrix $\nabla_x\Psi(\bar{x},\bar{w})$ of $\Psi=B\circ\Phi$ has full rank. Employing [39, Theorem 5.1] tells us that full stability of $\bar{x}$ for the reduced problem $\mathcal{P}_r^{\rm com}(\bar{w},\bar{v})$ is equivalent to

$$[(p,q)\in\mathcal{T}_r(\bar{x},\bar{w},\bar{v})(u),\ u\neq 0]\Longrightarrow\langle p,u\rangle>0 \tag{4.46}$$

via the set-valued mapping $\mathcal{T}_r(\bar{x},\bar{w},\bar{v})\colon\mathbb{R}^m\rightrightarrows\mathbb{R}^m\times\mathbb{R}^d$ defined by

$$\begin{aligned}\mathcal{T}_r(\bar{x},\bar{w},\bar{v})(u)\ &:=\Big(\nabla^2_{xx}\varphi_0(\bar{x},\bar{w})u,\nabla^2_{xw}\varphi_0(\bar{x},\bar{w})u\Big)\\ &+\Big(\nabla^2_{xx}\langle\bar{\mu},\Psi\rangle(\bar{x},\bar{w})u,\nabla^2_{xw}\langle\bar{\mu},\Psi\rangle(\bar{x},\bar{w})u\Big)\\ &+\Big(\nabla_x\Psi(\bar{x},\bar{w}),\nabla_w\Psi(\bar{x},\bar{w})\Big)^*\partial^2\vartheta(\bar{z},\bar{\mu})(\nabla_x\Psi(\bar{x},\bar{w})u),\quad u\in\mathbb{R}^m,\end{aligned}$$

where $\bar{\mu}$ is a unique solution to the reduced KKT system (4.41) for $(x,w,v):=(\bar{x},\bar{w},\bar{v})$. The full rank of $\nabla_x\Psi(\bar{x},\bar{w})$ allows us to use the second-order chain rule from [38, Theorem 3.1] and get

$$\mathcal{T}_r(\bar{x},\bar{w},\bar{v})(u)=\Big(\nabla^2_{xx}\varphi_0(\bar{x},\bar{w})u,\nabla^2_{xw}\varphi_0(\bar{x},\bar{w})u\Big)+D^*\partial_x(\vartheta\circ\Psi)(\bar{x},\bar{w},\bar{v})(u).$$

By the representation $(\vartheta\circ\Psi)(x,w)=(\theta\circ\Phi)(x,w)$ around $(\bar{x},\bar{w})$ we have

$$\mathcal{T}_r(\bar{x},\bar{w},\bar{v})(u)=\Big(\nabla^2_{xx}\varphi_0(\bar{x},\bar{w})u,\nabla^2_{xw}\varphi_0(\bar{x},\bar{w})u\Big)+D^*\partial_x(\theta\circ\Phi)(\bar{x},\bar{w},\bar{v})(u). \tag{4.47}$$

Applying now the second-order chain rule from Corollary 4.5 to the composition $\theta\circ\Phi$ in (4.47) together with (4.46) tells us that $\bar{x}$ being a fully stable local minimizer of the reduced problem $\mathcal{P}_r^{\rm com}(\bar{w},\bar{v})$ is equivalent to the validity of the inequality

$$\langle u,\nabla^2_{xx}L(\bar{x},\bar{w},\bar{\lambda})u\rangle+\langle q,\nabla_x\Phi(\bar{x},\bar{w})u\rangle>0\ \text{ for all }\ q\in\partial^2\theta(\bar{z},\bar{\lambda})(\nabla_x\Phi(\bar{x},\bar{w})u),\ u\neq 0. \tag{4.48}$$

Pick $0\neq u\in\mathcal{S}$ and get by (4.45) that $\nabla_x\Phi(\bar{x},\bar{w})u\in\operatorname{dom}\partial^2\theta(\bar{z},\bar{v})$. Thus it follows from

Lemma 4.9 that $0 \in \partial^2\theta(\bar{z},\bar{\lambda})(\nabla_x\Phi(\bar{x},\bar{w})u)$ implying by (4.48) that

$$\langle u, \nabla^2_{xx}L(\bar{x},\bar{w},\bar{\lambda})u\rangle = \langle u, \nabla^2_{xx}L(\bar{x},\bar{w},\bar{\lambda})u\rangle + \langle 0, \nabla_x\Phi(\bar{x},\bar{w})u\rangle > 0.$$

This shows that the composite SSOSC is satisfied and thus verifies the "only if" statement.

To justify next the "if" part of the theorem, take $u \neq 0$ and $q \in \partial^2\theta(\bar{z},\bar{\lambda})(\nabla_x\Phi(\bar{x},\bar{w})u)$, which yields $u \in \mathcal{S}$. Then [45, Theorem 2.1] together with the convexity of θ ensures that $\langle q, \nabla_x\Phi(\bar{x},\bar{w})u\rangle \geq 0$, and hence we have

$$\langle u, \nabla^2_{xx}L(\bar{x},\bar{w},\bar{\lambda})u\rangle + \langle q, \nabla_x\Phi(\bar{x},\bar{w})u\rangle \geq \langle u, \nabla^2_{xx}L(\bar{x},\bar{w},\bar{\lambda})u\rangle > 0$$

by the assumed composite SSOSC. This implies by (4.48) that $\bar{x}$ is a fully stable local minimizer of the reduced problem $\mathcal{P}^{\rm com}_r(\bar{w},\bar{v})$. Appealing finally to Proposition 4.6 shows that $\bar{x}$ is a fully stable local minimizer of problem $\mathcal{P}^{\rm com}(\bar{w},\bar{v})$ and thus completes the proof of the theorem. △

The obtained characterization extends the results of [39, Theorem 6.6] for MPPCs, of [39, Theorem 7.3] for ENLPs, and of [40, Theorem 6.3] for unconstrained minimax problems. An important advantage of Theorem 4.10 is that it allows us to characterize full stability of local minimizers in (nonsmooth) minimax problems with polyhedral constraints, which is done in the next section while cannot be obtained by using the developments in [39, 40].

4.3 Full Stability in Constrained Minimax Problems

This section deals with applications of Theorem 4.10 and second-order subdifferential calculations from [41] to characterizing fully stable local minimizers for the following class of *minimax* problems with *polyhedral constraints*:

$$\text{minimize} \quad \max\{\varphi_1(x),\ldots,\varphi_l(x)\} \quad \text{over} \quad \Upsilon(x) := (\zeta_1(x),\ldots,\zeta_r(x)) \in Z \quad \text{with} \quad r+l=m, \tag{4.49}$$

where the functions $\varphi_i\colon \mathbb{R}^n \to \mathbb{R}$ for $i = 1,\ldots,l$ and $\zeta_s\colon \mathbb{R}^n \to \mathbb{R}$ for $s = 1,\ldots,r$ are $\mathcal{C}^2$-smooth around the reference points, and where the convex polyhedron $Z \subset \mathbb{R}^r$ is given by

$$Z := \left\{ y \in \mathbb{R}^r \;\middle|\; \langle c_t, y\rangle \le \tau_t \;\text{ for all }\; t = 1,\ldots,p \right\} \tag{4.50}$$

with $(c_t, \tau_t) \in \mathbb{R}^r \times \mathbb{R}$ for $t = 1,\ldots,p$. The minimax counterpart of the perturbed problem $\mathcal{P}^{\text{com}}(w,v)$ from the previous section is written now as

$$\text{minimize}\;\; \max\{\varphi_1(x,w),\ldots,\varphi_l(x,w)\} + \delta(\Upsilon(x,w); Z) - \langle v, x\rangle \;\text{ subject to }\; x \in \mathbb{R}^n \tag{4.51}$$

with $(w,v) \in \mathbb{R}^d \times \mathbb{R}^n$. We say that $x \in \mathbb{R}^n$ is a feasible point to it (4.51) if $\Upsilon(x,w) \in Z$. Note that problem (4.51) differs from $\mathcal{P}^{\text{com}}(w,v)$ in (4.36) due to *nonsmoothness* of all the summands in (4.51) but $\langle v, x\rangle$. Let us show that nevertheless (4.51) can be reduced to the composite form (4.36) as follows. Consider the mapping $\Phi\colon \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}^{l+r} = \mathbb{R}^m$ given by

$$\Phi(x,w) := (\Xi(x,w), \Upsilon(x,w)) \;\text{ for all }\; (x,w) \in \mathbb{R}^n \times \mathbb{R}^d \tag{4.52}$$

with the mapping Υ taken from (4.49) and $\Xi(x,w) := (\varphi_1(x,w),\ldots,\varphi_l(x,w))$. Remembering that $r + l = m$, define the extended-real-valued function $\theta\colon \mathbb{R}^{l+r} \to \overline{\mathbb{R}}$ by

$$\begin{cases} \theta(x) := \max\left\{\langle a_1, x\rangle,\ldots,\langle a_l, x\rangle\right\} + \delta(x; \mathcal{Z}) \;\text{ for }\; x \in \mathbb{R}^{l+r} = \mathbb{R}^m \\ \text{with }\;\; \mathcal{Z} := \left\{ x \in \mathbb{R}^{l+r} \;\middle|\; \langle d_t, x\rangle \le \tau_t \;\text{ for }\; t = 1,\ldots,p \right\}, \end{cases} \tag{4.53}$$

where the generating vectors a_i and d_t are constructed from the unit orths $e_i \in \mathbb{R}^l$ and the vectors $c_t \in \mathbb{R}^r$ from (4.50) by, respectively,

$$a_i := (e_i, 0) \;\text{ for }\; i = 1,\ldots,l \;\text{ and }\; d_t := (0, c_t) \;\text{ for }\; t = 1,\ldots,p. \tag{4.54}$$

Observe the θ from (4.53) is a CPWL function in the summation form (4.3). Thus we can represent the constrained minimax problem (4.51) in the composite optimization form (4.36) written as

$$\text{minimize} \;\; (\theta \circ \Phi)(x,w) - \langle v, x\rangle \;\; \text{subject to} \;\; x \in \mathbb{R}^n \tag{4.55}$$

with θ taken from (4.53) with parameters (4.54) and the $\mathcal{C}^2$-smooth mapping Φ defined by (4.52).

Now we can apply Theorem 4.10 to (4.55) and derive in this way a second-order characterization of full stability of local solutions to the minimax problem (4.51) via its initial data. Prior to that, let us specify the nondegeneracy condition ND for problem (4.55) and presents it in terms of the original minimax problem (4.51) without appealing to the matrix B from the proof of Lemma 4.3.

Denote $\bar{z}_1 := \Xi(\bar{x},\bar{w})$ and $\bar{z}_2 := \Upsilon(\bar{x},\bar{w}) \in Z$ and construct the index sets

$$\begin{aligned} \mathcal{K}(\bar{z}_1): \;\; &= \;\; \Big\{ i \in \{1,\ldots,l\} \;\Big|\; \max\{\varphi_1(\bar{x},\bar{w}),\ldots,\varphi_l(\bar{x},\bar{w})\} = \varphi_i(\bar{x},\bar{w}) \Big\}, \\ \mathcal{I}(\bar{z}_2): \;\; &= \;\; \Big\{ t \in \{1,\ldots,p\} \;\Big|\; \langle c_t, \bar{z}_2\rangle = \tau_t \Big\} \end{aligned} \tag{4.56}$$

via the data of (4.49) and (4.50). It is easy to observe that $\mathcal{K}(\bar{z}_1) = K(\bar{z})$ and $\mathcal{I}(\bar{z}_2) = I(\bar{z})$ for the index sets defined in (4.5) for the function θ from (4.53) with $\bar{z} := (\bar{z}_1,\bar{z}_2) \in \operatorname{dom}\theta$.

Proposition 4.11 (equivalent form of condition ND for constrained minimax problems). *Let $\bar{x}$ be a feasible solution to* (4.51) *corresponding to $(\bar{w},\bar{v})$, and let $\bar{z} = (\bar{z}_1,\bar{z}_2)$ with $\bar{z}_1 = \Xi(\bar{x},\bar{w})$, $\bar{z}_2 = \Upsilon(\bar{x},\bar{w})$, and $(\bar{z},\bar{v}) \in \partial\theta$, where the mappings Ξ and Υ and the CPWL function θ are defined by* (4.52) *and* (4.53)*, respectively. Then the nondegeneracy condition ND in the framework of the minimax problem* (4.51) *can be equivalently written as*

$$\mathcal{D} \cap \ker\Big(\nabla_x \Xi(\bar{x},\bar{w})^*, \nabla_x \Upsilon(\bar{x},\bar{w})^*\Big) = \{0\} \tag{4.57}$$

with $\mathcal{D} := \{(y_1,y_2) \in \mathbb{R}^l \times \mathbb{R}^r \mid y_1 \in \operatorname{span}\{e_i - e_j \mid i,j \in \mathcal{K}(\bar{z}_1)\}$ and $y_2 \in \operatorname{span}\{c_t \mid t \in \mathcal{I}(\bar{z}_2)\}\}$, where $e_i \in \mathbb{R}^l$ are the unit vectors, and where the index sets $\mathcal{K}(\bar{z}_1)$ and $\mathcal{I}(\bar{z}_2)$ are defined in (4.56).

Proof. Applying the nondegeneracy condition ND to the composite optimization form (4.55) of the minimax problem (4.55) and using Theorem 4.4 on the equivalence of ND to SOCQ give us

$$\partial^2\theta(\bar z,\bar v)(0)\cap\ker\Phi(\bar x,\bar w)^*=\{0\}$$

with Φ and θ taken from (4.52) and (4.53), respectively. Then the second-order calculations of Theorem 4.2 together with the equalities $\mathcal{K}(\bar z_1)=K(\bar z)$ and $\mathcal{I}(\bar z_2)=I(\bar z)$ reveal that

$$\begin{aligned}&\partial^2\theta(\bar z,\bar v)(0)=\operatorname{span}\Big\{a_i-a_j\;\Big|\;i,j\in K(\bar z)\Big\}+\operatorname{span}\Big\{d_t\;\Big|\;t\in I(\bar z)\Big\}\\&=\Big\{(y_1,y_2)\in\mathbb{R}^l\times\mathbb{R}^n\;\Big|\;y_1\in\operatorname{span}\{e_i-e_j\mid i,j\in\mathcal{K}(\bar z_1)\},\;y_2\in\operatorname{span}\{c_t\mid t\in\mathcal{I}(\bar z_2)\}\Big\},\end{aligned}\tag{4.58}$$

where the vectors a_i and d_t are taken from (4.54). Observing that

$$\nabla_x\Phi(\bar x,\bar w)=\begin{pmatrix}\nabla_x\Xi(\bar x,\bar w)\\ \nabla_x\Upsilon(\bar x,\bar w)\end{pmatrix}$$

and combining this with representation (4.58) justify the equivalent form (4.57) of the ND condition in the minimax problem under consideration. △

After these adjustments, we are now ready to derive a characterization of fully stable local minimizers of (4.51). The KKT system associated with (4.51) can be expressed as

$$\begin{cases}\bar v=\displaystyle\sum_{i=1}^{l}\bar\lambda_i\nabla_x\varphi_i(\bar x,\bar w)+\sum_{s=1}^{r}\bar\mu_s\nabla_x\zeta_s(\bar x,\bar w)\\ \text{with}\quad\bar\lambda_i\ge 0,\ \displaystyle\sum_{i=1}^{l}\bar\lambda_i=1,\ (\bar\mu_1,\ldots,\bar\mu_r)\in N(\bar z_2;Z),\end{cases}\tag{4.59}$$

where $\bar z_2=\Upsilon(\bar x,\bar w)$, and where Z is taken from (4.50). The following definition is an adaptation of the composite SSOSC for the minimax problem (4.51).

Definition 4.12 (minimax SSOSC). *Given $\varpi:=(\bar\lambda,\bar\mu)\in\mathbb{R}^l\times\mathbb{R}^r$ from (4.59), we say* *that the* MINIMAX SSOSC *holds at* $(\bar{x}, \bar{w}, \bar{v}, \varpi)$ *with* $\bar{v}$ *satisfying* (4.59) *if*

$$\sum_{i=1}^{l} \bar{\lambda}_i \langle u, \nabla^2_{xx}\varphi_i(\bar{x}, \bar{w})u\rangle + \sum_{s=1}^{r} \bar{\mu}_s \langle u, \nabla^2_{xx}\zeta_s(\bar{x}, \bar{w})u\rangle > 0 \ \textit{ for all } \ 0 \neq u \in \mathcal{S}, \tag{4.60}$$

where the subspace $\mathcal{S}$ *is defined by*

$$\mathcal{S} := \Big\{ u \in \mathbb{R}^n \ \Big| \ \langle \nabla_x \varphi_i(\bar{x}, \bar{w}), u\rangle = \gamma \ \textit{ for } \ i \in \Gamma(J_1) \ \textit{ and } \ \langle d_t, \nabla_x \Upsilon(\bar{x}, \bar{w})u\rangle = 0 \ \textit{ for } \ t \in \Gamma(J_2) \Big\}$$

via the index sets $\Gamma(J_1)$ *and* $\Gamma(J_2)$ *taken from* (4.15) *and some constant* $\gamma \in \mathbb{R}$.

The next result extends [38, Theorem 6.3] to the case of *constrained* minimax problems.

Theorem 4.13 (characterization of fully stable solutions to constrained minimax problems). *Let* $\bar{x}$ *be a feasible solution to the minimax problem* (4.51) *corresponding to* $(\bar{w}, \bar{v})$ *with* $\bar{v} \in \partial_x(\theta \circ \Phi)(\bar{x}, \bar{w})$, *where* θ *and* Φ *are taken from* (4.53) *and* (4.52), *respectively. Assume that the ND condition* (4.57) *holds, and let* $\varpi = (\bar{\lambda}, \bar{\mu}) \in \mathbb{R}^l \times \mathbb{R}^r$ *be a unique solution to* (4.59). *Then* $\bar{x}$ *is a fully stable local minimizer of* (4.51) *if and only if the minimax SSOSC from* (4.60) *holds.*

Proof. Taking into account the previous considerations in this section, we now implement Theorem 4.10 in the constrained minimax setting by observing that the Lagrangian for problem (4.51) is $L(\bar{x}, \bar{w}, \varpi) = \langle \varpi, \Phi(\bar{x}, \bar{w})\rangle$. Therefore we get

$$\nabla^2_{xx} L(\bar{x}, \bar{w}, \varpi) = \sum_{i=1}^{l} \bar{\lambda}_i \nabla^2_{xx}\varphi_i(\bar{x}, \bar{w}) + \sum_{s=1}^{r} \bar{\mu}_s \nabla^2_{xx}\zeta_s(\bar{x}, \bar{w}).$$

Furthermore, it is easy to see that the set $\mathcal{S}$ from Definition 4.12 is an adaptation of the set $\mathcal{S}$ from (4.44) to the minimax problem (4.51). Thus the claimed second-order characterization of full stability in (4.51) readily follows from the equivalence in Theorem 4.10. △

4.4 Strong Regularity and Strong Stability in Composite Models

In this section, we continue our study of *composite optimization* problems of type (4.34) with CPWL outer functions θ therein. Our main goal is to establish relationships between full stability of local minimizers in (4.36) and some other stability/regularity notions for perturbed versions of (4.34) and associated (linearized and nonlinearized) KKT systems. The notions under consideration in what follows revolve around Robinson's *strong regularity* [47] and Kojima's *strong stability* [18]. Involving the nondegeneracy condition ND in composite optimization and employing the reduction approach as above, we show that these notions are actually equivalent in our setting while being also equivalent to full stability of local minimizers under appropriate choices of perturbations. In this way we derive explicit second-order characterizations of strong regularity and strong stability via the composite SSOSC introduced in Definition 4.8. All the results below can be specified in the case of constrained minimax problems (4.49) with replacing the composite SSOSC by its minimax counterpart from Definition 4.12.

To begin with, we rewrite the KKT system (4.40) for (4.36) as the *generalized equation*

$$\begin{bmatrix} v \\ 0 \end{bmatrix} \in \begin{bmatrix} \nabla_x L(x,w,\lambda) \\ -\Phi(x,w) \end{bmatrix} + \begin{bmatrix} 0 \\ (\partial\theta)^{-1}(\lambda) \end{bmatrix} \tag{4.61}$$

and denote by $S_{KKT}\colon (w,v) \mapsto (x,\lambda)$ the solution map to (4.61).

Robinson's idea [47] to define the property of strong regularity for generalized equations involved considering Lipschitzian single-valued localizations of solution maps to appropriate *linearizations*. This idea was further developed and applied in many publications; see, e.g., the books [1,5,10,17] and the references therein. We keep such a definition of strong regularity in the case of (4.61) and study it later on in this section. However, it is more convenient for us to start with a similar property for the solution map S_{KKK} of the KKT system (4.61) *itself*, without any linearization, and characterize it via the composite SSOSC.

Definition 4.14 (SVLL property of KKT systems). *We say that the KKT system* (4.61) *associated with the composite optimization problem* (4.36) *has the* SINGLE-VALUED LIPSCHITZIAN LOCALIZATION (SVLL) *property at* $(\bar{x},\bar{\lambda},\bar{w},\bar{v}) \in S_{KKT}$ *if its solution map* $S_{KKT}\colon (w,v) \mapsto (x,\lambda)$ *admits a Lipschitzian single-valued graphical localization around*

$(\bar{w}, \bar{v}, \bar{x}, \bar{\lambda})$.

The next theorem shows the SVLL property of (4.61) is characterized by the simultaneous fulfillment of the composite SSOSC and the nondegeneracy condition ND in composite optimization. It extends the corresponding result of [2, Theorem 4.10] and [4, Theorem 6] for NLPs; see also commentaries in [2, 5, 17] on related developments in this direction.

Theorem 4.15 (characterization of SVLL property via ND and composite SSOSC). *Let $\bar{x}$ be a feasible solution to problem $\mathcal{P}^{\text{com}}(\bar{w}, \bar{v})$ in* (4.36) *with some $\bar{w} \in \mathbb{R}^d$ and $\bar{v}$ from* (4.39), where $\theta \in CPWL$ and Φ is $\mathcal{C}^2$-smooth around $(\bar{x}, \bar{w})$. Consider the following statements:

(i) The SVLL property from Definition 4.14 holds and we have $\bar{x} \in M_\gamma(\bar{w}, \bar{v})$ for the argminimum set (2.16) with some $\gamma > 0$.

(ii) The composite SSOSC (4.43) and the nondegeneracy condition ND are satisfied.

Then we have (ii)$\Longrightarrow$(i), while the converse application holds if in addition the first-order qualification condition (4.16) is fulfilled.

Proof. Suppose first that (ii) holds and deduced from Theorem 4.10 that $\bar{x}$ is a fully stable locally optimal solution to $\mathcal{P}^{\text{com}}(\bar{w}, \bar{v})$. It follows from Proposition 4.6 that $\bar{x}$ is also a fully stable locally optimal solution to the reduced problem $\mathcal{P}_r^{\text{com}}(\bar{w}, \bar{v})$. Similarly to (4.61), we can write the KKT system for the reduced problem (4.41) in the generalized equation form

$$\begin{bmatrix} v \\ 0 \end{bmatrix} \in \begin{bmatrix} \nabla_x L_r(x, w, \mu) \\ -\Psi(x, w) \end{bmatrix} + \begin{bmatrix} 0 \\ (\partial\vartheta)^{-1}(\mu) \end{bmatrix} \tag{4.62}$$

and denote by $S_{KKT}^r\colon (w, v) \mapsto (x, \mu)$ its solution map. Lemma 4.3 gives us the representation $\theta = \vartheta \circ B$ with some mapping $\vartheta \in CPWL$ and the $s \times n$ matrix B constructed therein. Remembering that $\lambda = B^*\mu$, we split the proof of (ii)$\Longrightarrow$(i) into several steps.

Step 1: *The conditions in* (ii) *imply that the solution map $S_{KKT}^r\colon (w, v) \mapsto (x, \mu)$ has the*

SVLL property around $(\bar{w},\bar{v},\bar{x},\bar{\mu})$.

We start the proof of this fact by recalling that the full stability of $\bar{x}$ in the reduced problem $\mathcal{P}_r^{\rm com}(\bar{w},\bar{v})$ ensures by Theorem 2.5 that the set-valued mapping

$$S_r(w,v) := \Big\{x \in \mathbb{R}^n \;\Big|\; v \in \nabla_x\varphi_0(x,w) + \nabla_x\Psi(x,w)^*\partial\vartheta(\Psi(x,w))\Big\}$$

admits a Lipschitzian single-valued graphical localization around $(\bar{w},\bar{v},\bar{x})$. Employing this together with the surjectivity of $\nabla_x\Psi(\bar{x},\bar{w})$, which comes from the second part of Proposition 4.6, tells us that the mapping $S^r_{KKT}\colon (w,v)\mapsto(x,\mu)$ is single-valued around $(\bar{w},\bar{v},\bar{x},\bar{\mu})$. The Lipschitz continuity of $(w,v)\mapsto x_{wv} =: x$ around $(\bar{w},\bar{v})$ is a direct consequence of the full stability of $\bar{x}$ in the reduced problem $\mathcal{P}_r^{\rm com}(\bar{w},\bar{v})$. Let the latter property hold in some neighborhoods W of $\bar{w}$ and V of $\bar{v}$. To verify the same property for the mapping $(w,v)\mapsto\mu_{wv} =: \mu$, pick $w_1,w_2\in W$ and $v_1,v_2\in V$ and thus find $\mu_{w_iv_i}\in\partial\vartheta(c_i)$ with $c_i := \Psi(x_{w_iv_i},w_i)$ for $i=1,2$ satisfying

$$\begin{cases} v_2 = \nabla_x\varphi_0(x_{w_2v_2},w_2) + \nabla_x\Psi(x_{w_2v_2},w_2)^*\mu_{w_2v_2},\\[2mm] v_1 = \nabla_x\varphi_0(x_{w_1v_1},w_1) + \nabla_x\Psi(x_{w_1v_1},w_1)^*\mu_{w_1v_1}.\end{cases}$$

This allows us to obtain the equality

$$\begin{aligned}\nabla_x\Psi(x_{w_2v_2},w_2)^*(\mu_{w_2v_2}-\mu_{w_1v_1}) \;&=\; \Big(\nabla_x\Psi(x_{w_1v_1},w_1)-\nabla_x\Psi(x_{w_2v_2},w_2)\Big)^*\mu_{w_1v_1}\\ &+\;\; \nabla_x\varphi_0(x_{w_1v_1},w_1)-\nabla_x\varphi_0(x_{w_2v_2},w_2)+v_2-v_1.\end{aligned}$$

We can assume that $\nabla_x\Psi(x_{w_iv_i},w_i)$ is surjective because of this property for $\nabla_x\Psi(\bar{x},\bar{w})$. It follows from [27, Lemma 1.18] that for any $(w,v)\in W\times V$ there is $\kappa_{wv}>0$ such that

$$\|\nabla_x\Psi(x_{w_2v_2},w_2)^*(\mu_{w_2v_2}-\mu_{w_1v_1})\| \ge \kappa_{w_2v_2}\|\mu_{w_2v_2}-\mu_{w_1v_1}\| \ge \kappa\|\mu_{w_2v_2}-\mu_{w_1v_1}\|,$$

where $\kappa := \inf\{\kappa_{wv}|\;(w,v)\in W\times V\}$. Now we claim that $\kappa>0$. Indeed, assuming $\kappa=0$ gives us $(w_k,v_k)\to(\bar{w},\bar{v})$ such that $\kappa_{w_kv_k}\to 0$ as $k\to\infty$. Appealing to [27, Lemma 1.18], observe that $\kappa_{w_kv_k} = \inf\{\|\nabla_x\Psi(x_{w_kv_k},w_k)^*y\|\;|\;\|y\|=1\}$. This allows us to find y_k with

$\|y_k\| = 1$ and

$$\|\nabla_x \Psi(x_{w_k v_k}, w_k)^* y_k\| < \kappa_{w_k v_k} + \frac{1}{k}. \tag{4.63}$$

Suppose next without loss of generality that $y_k \to \bar{y}$ as $k \to \infty$ with $\|\bar{y}\| = 1$. Passing to limit in (4.63), we deduce that $\nabla_x \Psi(\bar{x}, \bar{w})^* \bar{y} = 0$. Taking then into account the surjectivity of $\nabla_x \Psi(\bar{x}, \bar{w})$, we arrive at $\bar{y} = 0$, which is a contradiction telling us that $\kappa > 0$. By the surjectivity of $\nabla_x \Psi(\bar{x}, \bar{w})$ there is $\rho < \infty$ so that $\|\mu_{wv}\| \le \rho$ for all $(w, v) \in W \times V$. Denoting by $\ell > 0$ a common Lipschitz constant for the mappings $\nabla_x \varphi_0$, $\nabla_x \Psi$, and $(w, v) \mapsto x_{wv}$ on $W \times V$ yields

$$\begin{aligned}
\|\mu_{w_2 v_2} - \mu_{w_1 v_1}\| &\le \kappa^{-1}\Big(\|\nabla_x \Psi(x_{w_1 v_1}, w_1) - \nabla_x \Psi(x_{w_2 v_2}, w_2)\| \cdot \|\mu_{w_1 v_1}\| \\
&+ \|\nabla_x \varphi_0(x_{w_1 v_1}, w_1) - \nabla_x \varphi_0(x_{w_2 v_2}, w_2)\| + \|v_2 - v_1\|\Big) \\
&\le \kappa^{-1}\Big[\rho\ell\Big(\|x_{w_2 v_2} - x_{w_1 v_1}\| + \|w_2 - w_1\|\Big) \\
&+ \ell\Big(\|x_{w_2 v_2} - x_{w_1 v_1}\| + \|w_2 - w_1\|\Big) + \|v_2 - v_1\|\Big],
\end{aligned}$$

which justifies the claimed local Lipschitz continuity of the mapping $(w, v) \mapsto \mu_{wv}$.

Step 2: *The conditions in* (ii) *imply that the SVLL property of* (4.61) *is satisfied* $(\bar{x}, \bar{\lambda}, \bar{w}, \bar{v})$.

To verify it, remember that the conditions in (ii) ensure by Theorem 4.10 that $\bar{x}$ is a fully stable local minimizer of $\mathcal{P}^{\mathrm{com}}(\bar{w}, \bar{v})$. Thus it follows from Theorem 2.5 that the set-valued mapping

$$S(w, v) := \Big\{ x \in \mathbb{R}^n \,\Big|\, v \in \nabla_x \varphi_0(x, w) + \nabla_x \Phi(x, w)^* \partial\theta(\Phi(x, w)) \Big\} \tag{4.64}$$

is single-valued and locally Lipschitzian around $(\bar{w}, \bar{v}, \bar{x})$. Since ND holds, the Lagrange multiplier in (4.61) is unique, and therefore the mapping $S_{KKT} \colon (w, v) \mapsto (x, \lambda)$ is single-valued around $(\bar{w}, \bar{v}, \bar{x}, \bar{\lambda})$. Furthermore, the Lipschitz continuity of $(w, v) \mapsto x_{wv} =: x$ around $(\bar{w}, \bar{v})$ follows from full stability of $\bar{x}$. Taking the neighborhoods W and V from Step 1, pick $w_i \in W$ and $v_i \in V$, $i = 1, 2$. Using then the relationship $\lambda = B^* \mu$, for each

i find a unique multiplier $\mu_{w_iv_i} \in \partial\vartheta(c_i)$ with $c_i := \Psi(x_{w_iv_i}, w_i)$ such that $\lambda_{w_iv_i} := B^*\mu_{w_iv_i}$. This leads us to

$$\begin{aligned}\|\lambda_{w_2v_2} - \lambda_{w_1v_1}\| &= \|B^*\mu_{w_2v_2} - B^*\mu_{w_1v_1}\| \\ &\leq \|B^*\| \cdot \|\mu_{w_2v_2} - \mu_{w_1v_1}\|,\end{aligned}$$

which thus justifies the local Lipschitz continuity of the mapping $(w, v) \mapsto \lambda_{wv}$ due to Step 1. This completes the proof of implication (ii)$\Longrightarrow$(i).

To verify the converse implication (i)$\Longrightarrow$(ii), suppose that the SVLL condition holds and pick $\eta \in S(\bar{z}) \cap \ker \nabla_x\Phi(\bar{x}, \bar{w})^*$ with $\bar{z} = \Phi(\bar{x}, \bar{w})$. Since $S(\bar{z}) = \operatorname{aff} \partial\theta(\bar{z})$, we get $\eta \in \operatorname{aff} \partial\theta(\bar{z})$ and deduce that $S_{KKT}(\bar{w}, \bar{v}) = \{(\bar{x}, \bar{\lambda})\}$ for some $\bar{\lambda} \in \mathbb{R}^m$ by taking into account the imposed qualification condition (4.16). If $\bar{\lambda} \in \operatorname{ri} \partial\theta(\bar{z})$ with "ri" standing for the relative interior of a convex set, then $\bar{\lambda} + t\eta \in \partial\theta(\bar{z})$ for any small $t > 0$, which tells us that $(\bar{x}, \bar{\lambda} + t\eta) \in S_{KKT}(\bar{w}, \bar{v})$. Employing now the single-valuedness of the mapping S_{KKT}, we get $\eta = 0$, and hence condition ND holds in this case. Suppose now that $\bar{\lambda} \notin \operatorname{ri} \partial\theta(\bar{z})$ and, taking into account that $\operatorname{ri} \partial\theta(\bar{z}) \neq \emptyset$, pick $\eta \in \operatorname{ri} \partial\theta(\bar{z})$. It follows from [53, Proposition 2.40] that $\bar{\lambda} + t(\eta - \bar{\lambda}) \in \operatorname{ri} \partial\theta(\bar{z})$ for any $t \in (0, 1)$. Letting $v_t := t\nabla_x\Phi(\bar{x}, \bar{w})^*(\eta - \bar{\lambda})$ for $t > 0$ small enough gives us $(\bar{x}, \bar{\lambda} + t(\eta - \bar{\lambda})) \in S_{KKT}(\bar{w}, \bar{v} + v_t)$. Remember that $\bar{\lambda} + t(\eta - \bar{\lambda}) \in \operatorname{ri} \partial\theta(\bar{z})$, which allows us to repeat the above arguments and to justify the validity of ND.

To finish the proof, it is not hard to see that by SVLL the set-valued mapping $S(w, v)$ in (4.64) is single-valued and locally Lipschitzian around $(\bar{w}, \bar{v}, \bar{x})$. Remembering that $\bar{x} \in M_\gamma(\bar{w}, \bar{v})$ in (i) and appealing to Theorem 2.5, with taking into account that the qualification condition imposed therein follows from the justified ND, tell us that $\bar{x}$ is a fully stable local minimizer of $\mathcal{P}^{\text{com}}(\bar{w}, \bar{v})$. Thus SSOSC holds by Theorem 4.10, and we complete the proof of the theorem. △

Now we proceed with the definition and second-order characterization of Robinson's strong regularity for the KKT system (4.61) associated with problem $\mathcal{P}^{\text{com}}(\bar{w}, \bar{v})$ of composite optimization.

Definition 4.16 (strong regularity of KKT in composite optimization). *Let* $(\bar{x}, \bar{\lambda})$ *be a solution to* (4.61) *for* $(w, v) = (\bar{w}, \bar{v})$ *with* $\bar{v} = 0$. *We say that* $(\bar{x}, \bar{\lambda})$ *is* STRONGLY REGULAR *for KKT* (4.61) *if the solution map to the linearized system at* $(\bar{x}, \bar{\lambda})$ *defined by*

$$\begin{bmatrix} v_1 \\ \\ v_2 \end{bmatrix} \in \begin{bmatrix} \nabla^2_{xx} L(\bar{x}, \bar{w}, \bar{\lambda})(x - \bar{x}) + \nabla_x \Phi(\bar{x}, \bar{w})^*(\lambda - \bar{\lambda}) \\ \\ -\Phi(\bar{x}, \bar{w}) - \nabla_x \Phi(\bar{x}, \bar{w})(x - \bar{x}) \end{bmatrix} + \begin{bmatrix} 0 \\ \\ (\partial\theta)^{-1}(\lambda) \end{bmatrix}$$

admits a Lipschitzian single-valued graphical localization around $(0, 0, \bar{x}, \bar{\lambda}) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^m$.

Our subsequent goal is to establish relationships between the KKT strong regularity from Definition 4.16 and full stability of local minimizers in composite optimization. We show that these notions are actually equivalent under nondegeneracy; see the precise formulation in Theorem 4.18. The result obtained below continues the line of equivalencies developed recently for various problems of constrained optimization in [33,39,40] while being new for the composite optimization problems studied in the paper. To proceed, we consider the following *canonically perturbed* version of problem (4.34) with parametric pairs $(v_1, v_2) \in \mathbb{R}^n \times \mathbb{R}^m$:

$$\widetilde{\mathcal{P}}^{\text{com}}_{\bar{w}}(v_1, v_2): \quad \text{minimize } \ \varphi_0(x, \bar{w}) + \theta(\Phi(x, \bar{w}) + v_2) - \langle v_1, x \rangle \ \text{ subject to } \ x \in \mathbb{R}^n. \tag{4.65}$$

The next lemma important in what follows reduces the study of full stability in the original optimization problem (4.36) to that in the canonically perturbed one (4.65) under nondegeneracy. Its proof is based on the criterion of full stability obtained in Theorem 4.10 and allows us to deal with generalized equations of type (4.61) whose set-valued parts depend on parameters.

Lemma 4.17 (full stability with respect to canonical perturbations). *Let* $\bar{x}$ *be a feasible solution to the composite optimization problem* $\mathcal{P}^{com}(\bar{w}, \bar{v})$ *in* (4.36) *with some* $\bar{w} \in \mathbb{R}^d$ *and* $\bar{v}$ *from* (4.39) under the nondegeneracy condition ND. Then $\bar{x}$ is a fully stable local minimizer of $\mathcal{P}^{\text{com}}(\bar{w},\bar{v})$ if and only if it is a fully stable local minimizer of $\widetilde{\mathcal{P}}^{\text{com}}_{\bar{w}}(\bar{v},0)$ in (4.65).

Proof. It is easy to see that the nondegeneracy condition ND for the canonically perturbed problem (4.65) agrees with the one for the fully perturbed problem (4.36). Suppose now that $\bar{x}$ is a fully stable local minimizer of $\widetilde{\mathcal{P}}^{\text{com}}_{\bar{w}}(\bar{v},0)$ and then apply Theorem 4.10 to conclude that it is equivalent to the validity of the following inequality:

$$\langle u, \nabla^2_{xx} L_{\bar{w}}(\bar{x},0,\bar{\lambda})u\rangle > 0 \;\text{ for all }\; 0 \neq u \in \mathcal{S}, \tag{4.66}$$

where the subspace $\mathcal{S}$ is defined in (4.44) and $L_{\bar{w}}$ is the Lagrangian associated with problem (4.65) given by $L_{\bar{w}}(x,v_2,\lambda) := \varphi_0(x,\bar{w}) + \langle \lambda, \Phi(x,\bar{w}) + v_2\rangle$. Therefore we have $\nabla^2_{xx} L_{\bar{w}}(\bar{x},0,\bar{\lambda}) = \nabla^2_{xx} L(\bar{x},\bar{w},\bar{\lambda})$ with L coming from (4.40), which indeed tells us that $\bar{x}$ is a fully stable local minimizer of $\mathcal{P}^{\text{com}}(\bar{w},\bar{v})$. The converse implication of the lemma is verified similarly. △

We are now ready to establish the aforementioned relationships between full stability of local minimizers in composite optimization and strong regularity of the associated KKT systems.

Theorem 4.18 (relationships between full stability and strong regularity in composite optimization). *Let $\bar{x}$ be a feasible solution to $\mathcal{P}^{\text{com}}(\bar{w},\bar{v})$ in* (4.36) *with some $\bar{w} \in \mathbb{R}^d$ and $\bar{v} = 0$ from* (4.39). Assume that the qualification condition (4.16) holds. Then the following are equivalent:

(i) $\bar{x}$ is a fully stable locally optimal solution to $\mathcal{P}^{\text{com}}(\bar{w},\bar{v})$ satisfying ND.

(ii) $\bar{x} \in M_\gamma(\bar{w},\bar{v})$ for some $\gamma > 0$ and $(\bar{x},\bar{\lambda})$ is a strongly regular solution to (4.61).

Proof. We first verify implication (ii)$\Longrightarrow$(i). It has been well recognized (see, e.g., [5, Theorem 2B.10] that strong regularity of the KKT system (4.61) at $(\bar{x},\bar{\lambda})$ is equivalent to the fact that the KKT system associated with the canonically perturbed problem (4.65) and given by

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \in \begin{bmatrix} \nabla_x L(x, \bar{w}, \lambda) \\ -\Phi(x, \bar{w}) \end{bmatrix} + \begin{bmatrix} 0 \\ (\partial\theta)^{-1}(\lambda) \end{bmatrix}$$

admits a Lipschitz continuous single-valued graphical localization around $(0, 0, \bar{x}, \bar{\lambda}) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^m$. Thus it results from Theorem 4.15 that the composite SSOSC from (4.66) and the nondegeneracy condition ND in the setting of $\widetilde{\mathcal{P}}^{\rm com}_{\bar{w}}(\bar{v}, 0)$ are satisfied. As mentioned in the proof of Lemma 4.17, the nondegeneracy conditions ND for both problems $\widetilde{\mathcal{P}}^{\rm com}_{\bar{w}}(\bar{v}, 0)$ and $\mathcal{P}^{\rm com}(\bar{w}, \bar{v})$ are the same, and therefore Theorem 4.10 says that $\bar{x}$ is a fully stable local minimizer for $\widetilde{\mathcal{P}}^{\rm com}_{\bar{w}}(\bar{v}, 0)$. Employing now Lemma 4.17 tells us that $\bar{x}$ is a fully stable local minimizer for the original problem $\mathcal{P}^{\rm com}(\bar{w}, \bar{v})$ as well, which justifies that (ii)$\Longrightarrow$(i). By similar arguments we verify the converse implication and thus complete the proof of the theorem. $\triangle$

As a by-product of the obtained equivalence and the characterization of full stability of local minimizers in Theorem 4.10, we get the composite SSOSC characterization of strong regularity for the associated KKT system (4.61). The results of this type for various problems of constrained optimization with $\mathcal{C}^2$-smooth data can be found in [1, 4, 17, 33, 39] via appropriate SSOSC and nondegeneracy conditions. Note that, in contrast to full stability, the corresponding nondegeneracy condition is *necessary* for strong regularity. Some second-order characterizations of full stability *without nondegeneracy* have been recently established in [31] for NLPs.

The last part of this section is devoted to studying relationships between strong regularity in the sense of Definition 4.16 and strong Lipschitzian stability in the sense of Kojima [18]. The concept of strong Lipschitzian stability was considered before only for problems of constrained optimization with $\mathcal{C}^2$-smooth data. Here we extend its to the general framework of composite optimization problems and then show that it is indeed equivalent to strong regularity of the corresponding KKT system. Note that relationships between strong regularity and strong stability were first studied in [15] for classical NLPs and then further developed for more general constrained problems in [1, 17, 33, 40]; see also the references therein.

To proceed in our composite optimization setting, suppose without loss of generality that $\bar{v} = 0$ and say that the pair $(\xi(x,u), \Upsilon(x,u))$ with $u \in \mathbb{R}^q$, $\xi\colon \mathbb{R}^n \times \mathbb{R}^q \to \mathbb{R}$, and $\Upsilon\colon \mathbb{R}^n \times \mathbb{R}^q \to \mathbb{R}^m$ is a $\mathcal{C}^2$-*smooth parametrization* of $(\varphi_0(x,\bar{w}), \Phi(x,\bar{w}))$ in $\mathcal{P}(\bar{w},0)$ at $\bar{u} \in \mathbb{R}^q$ if $\varphi_0(x,\bar{w}) = \xi(x,\bar{u})$ and $\Phi(x,\bar{w}) = \Upsilon(x,\bar{u})$ for all $x \in \mathbb{R}^n$, where both functions ξ and Υ are twice continuously differentiable. Consider now the family of the parametric optimization problems given by

$$\widehat{\mathcal{P}}(u): \quad \text{minimize } \ \xi(x,u) + \theta(\Upsilon(x,u)) \ \text{ subject to } \ x \in \mathbb{R}^n.$$

Definition 4.19 (strong Lipschitzian stability for composite optimization problems). *A stationary point $\bar{x}$ of problem $\mathcal{P}^{\text{com}}(\bar{w},0)$ from* (4.36) *is called* STRONGLY LIPSCHITZ STABLE *with respect to the given $\mathcal{C}^2$-smooth parametrization $(\xi(x,u), \Upsilon(x,u))$ of $(\varphi_0(x,\bar{w}), \Phi(x,\bar{w}))$ in $\mathcal{P}(\bar{w},0)$ at $\bar{u} \in \mathbb{R}^q$ if there are neighborhoods U of $\bar{u}$ and O of $\bar{x}$ such that for any $u \in U$ each problem $\widehat{\mathcal{P}}(u)$ has a unique stationary point $x(u) \in O$ and the mapping $u \longmapsto x(u)$ is locally Lipschitz continuous around $\bar{u}$. If it holds for any $\mathcal{C}^2$-smooth parameterization of $(\varphi_0(x,\bar{w}), \Phi(x,\bar{w}))$ in $\mathcal{P}(\bar{w},0)$ at $\bar{u} \in \mathbb{R}^q$, then the stationary point $\bar{x}$ is called strongly Lipschitz stable.*

The next theorem provides exact relationships between strong regularity and strong stability.

Theorem 4.20 (equivalence between strong regularity and strong Lipschitzian stability for composite optimization problems). *Let $\bar{x}$ be a feasible solution to the unperturbed problem $\mathcal{P}^{\text{com}}(\bar{w},\bar{v})$ in* (4.36) *with some $\bar{w} \in \mathbb{R}^d$ and $\bar{v} = 0$ from* (4.39). Assume further that the qualification condition (4.16) holds. Then the following are equivalent:

(i) $\bar{x}$ is a Lipschitz stable local optimal minimizer of $\mathcal{P}^{\text{com}}(\bar{w},0)$ satisfying ND.

(ii) $\bar{x} \in M_\gamma(\bar{w}, \bar{v})$ for some $\gamma > 0$ and $(\bar{x}, \bar{\lambda})$ is a strongly regular solution to (4.61).

Proof. Suppose that (i) holds. Since $(\varphi_0(x,w) - \langle x, v\rangle, \Phi(x,w))$ is a $\mathcal{C}^2$-smooth parametrization of $(\varphi_0(x,\bar{w}), \Phi(x,\bar{w}))$ in problem $\mathcal{P}^{\text{com}}(\bar{w}, 0)$ at the point $\bar{u} := (\bar{w}, 0) \in \mathbb{R}^d \times \mathbb{R}^n$, we find some neighborhoods U of $\bar{u}$ and O of $\bar{x}$ such that for any $u = (w, v) \in U$ there exists a unique stationary point $x(u)$ of $\widehat{\mathcal{P}}(u)$ for which the mapping $u \longmapsto x(u)$ is Lipschitz continuous around $(\bar{u}, \bar{x})$. This shows that the set-valued mapping

$$S(u) := \left\{ x \in \mathbb{R}^n \;\middle|\; v \in \nabla_x \varphi_0(x,w) + \nabla_x \Phi(x,w)^* \partial\theta(\Phi(x,w)) \right\}$$

admits a Lipschitzian single-valued graphical localization around $(\bar{u}, \bar{x})$. Employing Theorem 2.5, we see that $\bar{x}$ is a fully stable locally optimal solution to problem $\mathcal{P}^{\text{com}}(\bar{w}, 0)$, which in turn yields the validity of (ii) due to Theorem 4.18.

To prove the converse implication (ii) $\Longrightarrow$ (i), let $(\bar{x}, \bar{\lambda})$ be a strongly regular solution to the KKT system (4.61). This tells us that $\bar{x}$ is a fully stable local minimizer of $\mathcal{P}^{\text{com}}(\bar{w}, 0)$ due to Theorem 4.18 and that the nondegeneracy condition ND is satisfied. Pick now an arbitrary $\mathcal{C}^2$-smooth parametrization $(\xi(x,u), \Upsilon(x,u))$ of $(\varphi_0(x,\bar{w}), \Phi(x,\bar{w}))$ in $\mathcal{P}^{\text{com}}(\bar{w}, 0)$ at $\bar{u} \in \mathbb{R}^q$, which gives us the equalities $\nabla_x \varphi_0(\bar{x}, \bar{w}) = \nabla_x \xi(\bar{x}, \bar{u})$ and $\nabla_x \Phi(\bar{x}, \bar{w}) = \nabla_x \Upsilon(\bar{x}, \bar{u})$ together with those for the corresponding second-order derivatives. Therefore the composite SSOSC from (4.43) is satisfied for problem $\widehat{\mathcal{P}}(\bar{u})$, which in turn implies that $\bar{x}$ is a fully stable local minimizer of problem $\widehat{\mathcal{P}}(\bar{u})$. Employing now Theorem 2.5, we deduce that the set-valued mapping

$$S(u,v) := \left\{ x \in \mathbb{R}^n \;\middle|\; v \in \nabla_x \xi(x,u) + \nabla_x \Upsilon(x,u)^* \partial\theta(\Upsilon(x,u)) \right\}$$

admits a Lipschitzian single-valued graphical localization around $(\bar{u}, 0, \bar{x})$. Defining $x(u) := S(u, 0)$, conclude that it is a stationary point for problem $\widehat{\mathcal{P}}(u)$ and that the mapping $u \longmapsto x(u)$ is locally Lipschitzian around $(\bar{u}, \bar{x})$. This verifies (i) and completes the proof of theorem. $\triangle$

CHAPTER 5 CRITICAL MULTIPLIERS FOR COMPOSITE PROBLEMS

This chapter concerns some applications of full stability in numerical algorithms. To this end, we first introduce the concept of critical multipliers for composite optimization problems. This concept has been recently investigated in [14] for classical problems of nonlinear programming. It is shown there that critical multipliers are largely responsible for slow convergence of major primal-dual numerical algorithms including the Newton method and the sequential quadratic programming method. Therefore it is crucial from the numerical viewpoint to rule out the existence of critical multipliers. In this chapter, we show how full stability can be implement to rule out such multipliers.

Assume that the mapping $f\colon \mathbb{R}^n \to \mathbb{R}^n$ is $\mathcal{C}^1$-smooth, and that the mapping $\Phi\colon \mathbb{R}^n \to \mathbb{R}^m$ is $\mathcal{C}^2$-smooth around $\bar{x}$. Letting $(p_1, p_2) \in \mathbb{R}^n \times \mathbb{R}^m$ and $\theta \in CPWL$, consider the parameterized generalized equations

$$\begin{bmatrix} p_1 \\ \\ p_2 \end{bmatrix} \in \begin{bmatrix} \Psi(x,v) \\ \\ -\Phi(x) \end{bmatrix} + \begin{bmatrix} 0 \\ \\ (\partial\theta)^{-1}(v) \end{bmatrix} \tag{5.1}$$

with the mapping Ψ defined by

$$\Psi(x,v)\colon = f(x) + \nabla\Phi(x)^* v, \qquad (x,v) \in \mathbb{R}^n \times \mathbb{R}^m.$$

The generalized equation (5.1) is a broad framework for many KKT systems of optimization problems including composite optimization problems, which were recently studied in [42], and nonlinear programming problems. Following the tradition for optimization problems, if (x, v) is a solution of the generalized equation (5.1) associated with the parameters (p_1, p_2), then v is referred as a Lagrange multiplier associated with the primal solution x. Given $\bar{x} \in \mathbb{R}^n$, define the set of Lagrange multipliers associated with $\bar{x}$ by

$$\Lambda(\bar{x})\colon = \Big\{ v \in \mathbb{R}^m \;\Big|\; \Psi(\bar{x}, v) = 0, \;\; v \in \partial\theta(\bar{z}) \Big\} \quad \text{with } \; \bar{z} = \Phi(\bar{x}). \tag{5.2}$$

Pick $\bar{v} \in \partial\theta(\bar{z})$ with $\bar{z} = \Phi(\bar{x})$ and introduce the *critical cone* for the CPWL function θ by

$$\mathcal{K}(\bar{z},\bar{v}) := \Big\{ w \in T(\bar{z}; \operatorname{dom}\theta) \;\Big|\; \langle \bar{v}, w\rangle = d\theta(\bar{z})(w) \Big\}, \tag{5.3}$$

where the subderivative function $d\theta(\bar{z})\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ is defined by

$$d\theta(\bar{z})(\bar{w})\colon = \liminf_{\substack{w\to\bar{w}\\ t\downarrow 0}} \frac{\theta(\bar{z}+tw)-\theta(\bar{z})}{t}.$$

It was proved in [53, Theorem 10.21] that the subderivative function for the CPWL function θ can be simplified as

$$d\theta(\bar{z})(\bar{w}) = \lim_{t\downarrow 0} \frac{\theta(\bar{z}+t\bar{w})-\theta(\bar{z})}{t}. \tag{5.4}$$

The term "critical cone" exploited above for the set $\mathcal{K}(\bar{z},\bar{v})$ from (5.3) was inspired by the same concept for convex polyhedra. Indeed, when we have $l = 0$ in representation (4.3), the CPWL function θ reduces to the indicator function $\delta(\cdot\,;\Omega)$ of the polyhedral set $\Omega\colon = \operatorname{dom}\theta$. This implies that $d\theta(\bar{z})(\bar{w}) = 0$ for any $\bar{w} \in T(\bar{z};\Omega)$ and hence the set $\mathcal{K}(\bar{z},\bar{v})$ from (5.3) has a representation of the form

$$\mathcal{K}(\bar{z},\bar{v}) = T(\bar{z};\Omega) \cap \bar{v}^{\perp},$$

which is the the well-known definition for the critical cone of the convex polyhedral set Ω. Below, we prove that the critical cone $\mathcal{K}(\bar{z},\bar{v})$ for CPWL functions can be entirely expressed via their parameters from (4.3).

Proposition 5.1 (equivalent description of the critical cone for CPWL functions). *Let $\theta \in CPWL$ with $(\bar{z},\bar{v}) \in \partial\theta$, and let $\bar{v}_1$, $\bar{v}_2$ be from* (4.7) *such that $\bar{v} = \bar{v}_1 + \bar{v}_2$. Denote by $K\colon = K(\bar{z})$, $I\colon = I(\bar{z})$, $J_1 := J_+(\bar{z},\bar{v}_1)$, and $J_2 := J_+(\bar{z},\bar{v}_2)$ the index sets from* (4.5), *and* (4.9), *respectively. Then the critical cone* $\mathcal{K}(\bar{z},\bar{v})$ *can be equivalently expressed by*

$$\mathcal{K}(\bar{z},\bar{v}) = \Big\{u\in\mathbb{R}^m \Big|\ \begin{array}{l}\langle a_i-a_j,u\rangle = 0 \ \textit{ if } \ i,j\in J_1,\\[1ex] \langle a_i-a_j,u\rangle \le 0 \ \textit{ if } \ (i,j)\in(K\setminus J_1)\times J_1,\\[1ex] \langle d_i,u\rangle = 0 \ \textit{ if } \ i\in J_2, \ \ \textit{and} \ \langle d_i,u\rangle\le 0 \ \textit{ if } \ i\in I\setminus J_2 \end{array}\Big\}. \tag{5.5}$$

Proof. Take $u\in\mathcal{K}(\bar{z},\bar{v})$. We claim that $\langle a_s,u\rangle = \langle a_i,u\rangle$ whenever $i\in J_1$. In fact, it follows from [53, Theorem 10.21] that $\operatorname{dom} d\theta(\bar{z}) = T(\bar{z};\operatorname{dom}\theta)$. Since we have $u\in T(\bar{z};\operatorname{dom}\theta)$, we find sequences $t_k\downarrow 0$ and $u_k\to u$ such that $\bar{z}+t_ku_k\in\operatorname{dom}\theta$. Thus we obtain a constant index subset $P\subset K$ so that $K(\bar{z}+t_ku_k) = P$ for all k. Picking $s\in P$ and using the equivalent description (5.4) of the subderivative function for CPWL functions, we arrive at

$$d\theta(\bar{z})(u) = \langle a_s,u\rangle. \tag{5.6}$$

Pick $i\in K$ and $s\in P$ and observe that $\langle a_i,\bar{z}+t_ku_k\rangle-\alpha_i\le\langle a_s,\bar{z}+t_ku_k\rangle-\alpha_s$, which leads us to

$$\langle a_i,u\rangle\le\langle a_s,u\rangle \quad \text{whenever} \quad i\in K,\ s\in P. \tag{5.7}$$

Moreover, we deduce from (4.8) and $u\in T(\bar{z};\operatorname{dom}\theta)$ that $\langle\bar{v}_2,u\rangle\le 0$. Employing this together with (5.6) and (5.7), we get

$$\langle a_s,u\rangle = d\theta(\bar{z})(u) = \langle\bar{v},u\rangle\le\langle\bar{v}_1,u\rangle = \sum_{i\in J_1}\bar{\lambda}_i\langle a_i,u\rangle\le\sum_{i\in J_1}\bar{\lambda}_i\langle a_s,u\rangle = \langle a_s,u\rangle, \tag{5.8}$$

which justifies the claim. Thus we get $\langle a_i-a_j,u\rangle = 0$ whenever $i,j\in J_1$. Assume now $(i,j)\in(K\setminus J_1)\times J_1$. Pick $s\in P$ and observe from (5.7) that $\langle a_i,u\rangle\le\langle a_s,u\rangle$. Since $\langle a_s,u\rangle = \langle a_j,u\rangle$, we accomplish that $\langle a_i-a_j,u\rangle\le 0$. Finally, we infer from (5.8) that $\langle\bar{v}_2,u\rangle = 0$. Combining this with the inequality $\langle\bar{v}_2,u\rangle\le 0$, we arrive at $\langle d_i,u\rangle = 0$ for $i\in J_2$ and $\langle d_i,u\rangle\le 0$ when $i\in I\setminus J_2$. These justify the inclusion "$\subset$" in (5.5).

To prove the opposite inclusion, let u be an element from right side of (5.5). Thus we deduce from (4.8) that $u\in T(\bar{z};\operatorname{dom}\theta)$. Appealing to [53, Exercize 8.4] that $\langle\bar{v},u\rangle\le$

$d\theta(\bar{z})(u)$. On the other hand, it is not hard to see that $d\theta(\bar{z})(u) = \langle a_r, u\rangle$ for some $r \in K$. Take $i \in J_1$ and get

$$\langle a_r, u\rangle \le \langle a_i, u\rangle = \langle \bar{v}_1, u\rangle = \langle \bar{v}, u\rangle.$$

This shows that $\langle \bar{v}, u\rangle = d\theta(\bar{z})(u)$ and hence we arrive at $u \in \mathcal{K}(\bar{z}, \bar{v})$. $\triangle$

Definition 5.2 (critical multipliers). *Assume that* $(\bar{x}, \bar{v})$ *is a solution of the generalized equation* (5.1) *for the parameters* $(\bar{p}_1, \bar{p}_2) = (0, 0)$, *and that* $\theta \in CPWL$ *with* $(\bar{z}, \bar{v}) \in \partial\theta$ *and* $\bar{z} = \Phi(\bar{x})$. *We say that the Lagrange multiplier* $\bar{v} \in \Lambda(\bar{x})$ *is* CRITICAL *for the generalized equation* (5.1) *if there exists a pair* $(\xi, \eta) \in \mathbb{R}^n \times \mathbb{R}^m$ *with* $\xi \neq 0$ *for which the following conditions are satisfied:*

- **(i)** $\eta \in \widehat{D}^*(\partial\theta)(\bar{x}, \bar{v})(-\nabla\Phi(\bar{x})\xi)$.
- **(ii)** $\nabla_x\Psi(\bar{x}, \bar{v})\xi + \nabla\Phi(\bar{x})^*\eta = 0$.
- **(iii)** $\langle \eta, \nabla\Phi(\bar{x})\xi\rangle = 0$.

Moreover, the Lagrange multiplier $\bar{v} \in \Lambda(\bar{x})$ *is called* NONCRITICAL *provided that it is not critical.*

It is worth noticing that the values of the precoderivative of CPWL functions were recently calculated in [41, Theorem 4.3] via their parameters from (4.3). Indeed, it was justified there that

$$\widehat{D}^*(\partial\theta)(\bar{z}, \bar{v})(u) = \Big(\mathcal{K}(\bar{z}, \bar{v})\Big)^* \quad \text{for any} \ \ u \in \text{dom}\,(\widehat{D}^*\partial\theta)(\bar{z}, \bar{v}) = -\mathcal{K}(\bar{z}, \bar{v}), \tag{5.9}$$

where '$*$' signifies the polar cone of $\mathcal{K}(\bar{z}, \bar{v})$. Moreover, we showed that

$$\begin{aligned}\Big(\mathcal{K}(\bar{z}, \bar{v})\Big)^* \ &= \text{span}\Big\{a_i - a_j \Big| \ i, j \in J_1\Big\} + \Big\{a_i - a_j \Big| \ (i, j) \in (K \setminus J_1) \times J_1\Big\} \\ &+ \text{span}\Big\{d_i \Big| \ i \in J_2\Big\} + \Big\{d_i \Big| \ i \in I \setminus J_2\Big\},\end{aligned}$$

where the index sets $K\!: = K(\bar{z})$, $I\!: = I(\bar{z})$, $J_1 := J_+(\bar{x}, \bar{v}_1)$, and $J_2 := J_+(\bar{z}, \bar{v}_2)$ are from (4.5), and (4.9), respectively. The original definition of critical multipliers was appeared in [14, Definition 1.41] when the CPWL function θ is the indicator function $\delta(\cdot\,;\Omega)$ of the polyhedral set $\Omega\colon=\mathbb{R}^s\times\mathbb{R}^{n-s}_-$. The latter cone is used in problems of nonlinear programming with s equality and $n-s$ inequality constraints. It is not hard to see that Definition 5.2 reduces to the one appeared in [14, Definition 1.41] when the CPWL function θ is as described above. It is worth noticing that the validity of the *second order sufficient condition*

$$\langle\nabla_x\Psi(\bar{x},\bar{v})u,u\rangle>0\quad\text{for all}\quad 0\neq u\in\mathbb{R}^n\quad\text{with}\quad\nabla\Phi(\bar{x})u\in\mathcal{K}(\bar{z},\bar{v})\tag{5.10}$$

for $\bar{v}\in\Lambda(\bar{x})$ results in that the Lagrange multiplier $\bar{v}$ is *noncritical.* The interest on keeping the term "second-order sufficient condition" for condition (5.10) resides in the following observation. Consider the following composite optimization problem:

$$\text{minimize}\ \ \varphi_0(x)+\theta(\Phi(x))\ \ \text{subject to}\ \ x\in\mathbb{R}^n,\tag{5.11}$$

where $\theta\in CPWL$, and where $\varphi_0\colon\mathbb{R}^n\to\mathbb{R}$ and $\Phi\colon\mathbb{R}^n\to\mathbb{R}^m$ are $\mathcal{C}^2$-smooth around the reference optimal solution. Pick $v\in\partial\theta(x)$ for $x\in\mathbb{R}^n$ and define the Lagrangian of problem (5.11) by

$$L(x,v)\colon=\varphi_0(x)+\langle\Phi(x),v\rangle.$$

Moreover, introduce the set of Lagrange multipliers associated with the feasible solution $\bar{x}$ for problem (5.11) by

$$\Lambda_{\text{com}}(\bar{x})\colon=\Big\{v\in\mathbb{R}^m\ \Big|\ L(\bar{x},v)=0,\ \ v\in\partial\theta(\bar{z})\Big\}\quad\text{with}\ \ \bar{z}=\Phi(\bar{x}).\tag{5.12}$$

The next proposition reveals that the validity of the *second order sufficient condition*

$$\langle\nabla^2_{xx}L(\bar{x},\bar{v})u,u\rangle>0\quad\text{for all}\quad 0\neq u\in\mathbb{R}^n\quad\text{with}\quad\nabla\Phi(\bar{x})u\in\mathcal{K}(\bar{z},\bar{v})\tag{5.13}$$

ensures that $\bar{x}$ is a unique local optimal solution for problem (5.11). It is important to point out that condition (5.13) boils down to the well-known second order sufficient condition for problems of nonlinear programming.

Proposition 5.3 (sufficient condition for local optimal solutions of composite optimization problems). *Let $\theta \in CPWL$ and $\Phi(\bar{x}) \in \operatorname{dom}\theta$, and let $\bar{v} \in \Lambda_{com}(\bar{x})$. Assume further that the second order sufficient condition* (5.13) *is satisfied. Then $\bar{x}$ is a unique local optimal solution to problem* (5.11).

Proof. Suppose that $\bar{x}$ is not a unique local minimizer for problem (5.11). This allows us to find a sequence x_k with $x_k \to \bar{x}$ as $k \to \infty$ for which we have

$$\varphi_0(x_k) + \theta(\Phi(x_k)) \le \varphi_0(\bar{x}) + \theta(\Phi(\bar{x})) \quad \text{and} \quad \Phi(x_k) \in \operatorname{dom}\theta. \tag{5.14}$$

Letting $\bar{z} := \Phi(\bar{x})$ and $z_k := \Phi(x_k)$, we have $K(z_k) \subset K(\bar{z})$ for all k sufficiently large. Extracting a subsequence of z_k if necessary, pick without loss of generality a constant index subset $P \subset K(\bar{z})$ so that $K(z_k) = P$ for all k. Define $u_k := \frac{x_k - \bar{x}}{\|x_k - \bar{x}\|}$ and assume without loss of generality that $u_k \to \bar{u}$ as $k \to \infty$ for some $\bar{u} \in \mathbb{R}^n$. Pick $r \in P$ and observe by the inclusion $P \subset K(\bar{z})$ that

$$\theta(\Phi(\bar{x})) = \langle a_r, \Phi(\bar{x})\rangle - \alpha_r \qquad \text{and} \qquad \theta(\Phi(x_k)) = \langle a_r, \Phi(x_k)\rangle - \alpha_r.$$

Hence we get

$$\Big(\varphi_0(x_k) - \varphi_0(\bar{x})\Big) + \langle a_r, z_k - \bar{z}\rangle \le 0, \tag{5.15}$$

which leads us to the relationship

$$\nabla\varphi_0(\bar{x})\bar{u} + \langle a_r, \nabla\Phi(\bar{x})\bar{u}\rangle \le 0 \tag{5.16}$$

for any $r \in P$. Since $\bar{v}_1 \in \partial\theta(\bar{z})$ and $a_r \in \partial\theta(z_k)$ for any $r \in P$, we deduce from the convexity of θ that

$$\langle \bar{v}_1, z_k - \bar{z}\rangle \le \langle a_r, z_k - \bar{z}\rangle.$$

Combining this with (5.16), we arrive at

$$\nabla\varphi_0(\bar{x})\bar{u} + \langle \bar{v}_1, \nabla\Phi(\bar{x})\bar{u}\rangle \le 0. \tag{5.17}$$

Moreover, by the inclusion $I(z_k) \subset I(\bar{z})$, we obtain

$$\langle \bar{v}_2, z_k - \bar{z} \rangle \le 0 \qquad \text{and} \qquad \langle \bar{v}_2, \nabla\Phi(\bar{x})\bar{u} \rangle \le 0. \tag{5.18}$$

Remember that $\bar{v} \in \Lambda_{\text{com}}(\bar{x})$, which says that $\nabla\varphi_0(\bar{x}) + \nabla\Phi(\bar{x})^*\bar{v} = 0$. Taking it into account together with (5.17) and (5.18), we come up to the relationships

$$\langle a_j, \nabla\Phi(\bar{x})\bar{u} \rangle = -\nabla\varphi_0(\bar{x})\bar{u} \quad \text{for} \quad j \in J_1 \quad \text{and} \quad \langle d_t, \nabla\Phi(\bar{x})\bar{u} \rangle = 0 \quad \text{for} \quad t \in J_2. \tag{5.19}$$

Take $i \in K$ and conclude from the convexity of θ that

$$\langle a_i, z_k - \bar{z} \rangle \le \theta(z_k) - \theta(\bar{z}) \le -\Big(\varphi_0(x_k) - \varphi_0(\bar{x})\Big)$$

by which we arrive at

$$\langle a_i, \nabla\Phi(\bar{x})\bar{u} \rangle \le -\nabla\varphi_0(\bar{x})\bar{u} = \langle a_j, \nabla\Phi(\bar{x})\bar{u} \rangle \quad \text{for} \quad j \in J_1.$$

Employing this along with (5.19) tells us that $\nabla\Phi(\bar{x})\bar{u} \in \mathcal{K}(\bar{z}, \bar{v})$ with $\bar{u} \neq 0$. Since we have $\bar{v} = \bar{v}_1 + \bar{v}_2$, it follows from (5.15) and (5.18) that

$$\Big(\varphi_0(x_k) - \varphi_0(\bar{x})\Big) + \langle \bar{v}, z_k - \bar{z} \rangle \le 0.$$

Implementing now the Taylor expansion formula together with $\bar{v} \in \Lambda_{\text{com}}(\bar{x})$ leads us to

$$\langle \nabla^2_{xx} L(\bar{x}, \bar{v})\bar{u}, \bar{u} \rangle \le 0,$$

being a contradiction with the second-order sufficient condition (5.13). This justifies the result. $\triangle$

It is worth noticing that the assumption $\bar{v} \in \Lambda_{\text{com}}(\bar{x})$ in Proposition 5.3 can be satisfied under the validity of the well-known qualification condition

$$\partial^\infty\theta(\Phi(\bar{x})) \cap \ker \nabla\Phi(\bar{x})^* = \{0\}, \tag{5.20}$$

which reduces to the Robinson constraint qualification for problems of constrained optimization. Moreover, we can guarantee the validity of the latter assumption provided that the set-valued mapping $F\colon \mathbb{R}^n \times \mathbb{R} \rightrightarrows \mathbb{R}^m \times \mathbb{R}$ defined by

$$F(x,\alpha)\colon = \operatorname{epi}\theta - (\Phi(x),\alpha)$$

is *metrically subregular* at $(\bar{x},\theta(\Phi(\bar{x})),0,0)$; see [13, pp. 210] for more details. It is interesting to observe that the qualification condition (5.20) is equivalent to the set-valued mapping F being *metrically regular* around $(\bar{x},\theta(\Phi(\bar{x})),0,0)$.

5.1 Critical Multipliers for Composite Optimization Problems

We continue with the following extension of [41, Theorem 3.4]. The message of the following important theorem for problems of nonlinear programming is trivial. Indeed, let the CPWL function θ be the indicator function $\delta(\cdot\,;\Omega)$ of the polyhedral set $\Omega\colon= \mathbb{R}^s \times \mathbb{R}^{p-s}_{-}$ with $s \geq 0$. This tells us that $l = 0$ and $d_i = e_i$ in representation (4.3), where $e_i \in \mathbb{R}^m$ is the unit vector such that the ith component of it is 1 while the others are 0. The following result in this particular case can be easily proved because the vectors d_i, $i \in T_2$, are linear independent. However, in the general framework of CPWL functions it needs to be taken care rigourously.

Theorem 5.4 (description of points in the subdifferential graph of CPWL functions). *Let $\theta \in CPWL$ with $(\bar{z},\bar{v}) \in \partial\theta$. Then there exists a neighborhood O of $(\bar{z},\bar{v})$ such that for any $(z,v) \in (\partial\theta) \cap O$ we have $J_+(\bar{z},\bar{v}_1) \subset K(x)$ and $J_+(\bar{z},\bar{v}_2) \subset I(z)$, where $\bar{v}_1$ and $\bar{v}_2$ are taken from* (4.7)*, and where and $J_+(\bar{z},\bar{v}_1)$ and $J_+(\bar{z},\bar{v}_2)$ are given by* (4.9).

Proof. We split the proof into the following major steps with keeping all the notation above.

Claim 1: *Let $\bar{v} = \sum_{i\in P}\eta_i a_i + \sum_{i\in Q}\tau_i d_i$ with some $\tau_i,\eta_i \geq 0$ satisfying $\sum_{i\in P}\eta_i = 1$, $P \subset K(\bar{z})$, and $Q \subset I(\bar{z})$. Then we have the equality*

$$\sum_{i\in P}\eta_i\alpha_i + \sum_{i\in Q}\tau_i\beta_i = \sum_{i\in K(\bar{z})}\bar{\lambda}_i\alpha_i + \sum_{i\in I(\bar{z})}\bar{\mu}_i\beta_i, \tag{5.21}$$

where the multipliers $\bar{\lambda}_i$ and $\bar{\mu}_i$ are taken from (4.7).

To verify this claim, suppose that $\bar{v} = \widehat{v}_1 + \widehat{v}_2$ for $\widehat{v}_1 = \sum_{i\in P} \eta_i a_i$ and $\widehat{v}_2 = \sum_{i\in Q} \tau_i d_i$ with $\sum_{i\in P} \eta_i = 1$ and $\eta_i, \tau_i \geq 0$. Fix $j \in P$ and observe that $\langle a_j, \bar{z}\rangle - \alpha_j = \langle a_i, \bar{z}\rangle - \alpha_i$ for any $i \in K(\bar{z})$ and $P \subset K(\bar{z})$. This tells us that

$$\langle a_j, \bar{z}\rangle - \alpha_j = \langle \bar{v}_1, \bar{z}\rangle - \sum_{i\in K(\bar{z})} \bar{\lambda}_i \alpha_i$$

with $\bar{v}_1$ taken from (4.7), which implies in turn that

$$\langle \widehat{v}_1, \bar{z}\rangle - \sum_{j\in P} \eta_j \alpha_j = \langle \bar{v}_1, \bar{z}\rangle - \sum_{i\in K(\bar{z})} \bar{\lambda}_i \alpha_i. \tag{5.22}$$

Since $Q \subset I(\bar{z})$, this allows us to deduce that

$$\langle \widehat{v}_2, \bar{z}\rangle = \sum_{i\in Q} \tau_i \beta_i \text{ and } \langle \bar{v}_2, \bar{z}\rangle = \sum_{i\in I(\bar{z})} \bar{\mu}_i \beta_i, \tag{5.23}$$

where $\bar{v}_2$ is from (4.7). Combining (5.22) and (5.23) with $\widehat{v}_1 + \widehat{v}_2 = \bar{v}_1 + \bar{v}_2$ justifies the claim.

Suppose now that the conclusion of the theorem does not hold and thus find a sequence $(z_k, v_k) \in \partial\theta$ such that $(z_k, v_k) \to (\bar{z}, \bar{v})$ as $k \to \infty$ while either $J_+(\bar{z}, \bar{v}_1) \not\subset K(z_k)$ or $J_+(\bar{z}, \bar{v}_2) \not\subset I(z_k)$ for all $k \in I\!N$. We suppose that $J_+(\bar{z}, \bar{v}_1) \not\subset K(z_k)$ and $J_+(\bar{z}, \bar{v}_2) \not\subset I(z_k)$ for all $k \in I\!N$. The other cases can be handled similarly.

Taking into account that the sets $J_+(\bar{z}, \bar{v}_1)$ and $J_+(\bar{z}, \bar{v}_2)$ are finite and considering a subsequence of z_k if necessary, we find $s \in J_+(\bar{z}, \bar{v}_1)$ and $s' \in J_+(\bar{z}, \bar{v}_2)$ so that $s \notin K(z_k)$ and $s' \notin I(z_k)$ for all $k \in I\!N$. Furthermore, it is not hard to see that $K(z_k) \subset K(\bar{z})$ and $I(z_k) \subset I(\bar{z})$ for k sufficiently large. Extracting similarly another subsequence, pick without loss of generality constant index subsets $P \subset K(\bar{z})$ and $Q \subset I(\bar{z})$ so that $K(z_k) = P$ and $I(z_k) = Q$ for all k. Select $j \in P$ and observe that $z_k \in C_j$, which implies by (4.4) that

$$\langle a_j, z_k\rangle - \alpha_j \geq \langle a_i, z_k\rangle - \alpha_i \text{ for all } i \in K(\bar{z}). \tag{5.24}$$

On the other hand, the construction in (4.4) and the conditions $z_k \notin C_s$, $z_k \in \operatorname{dom}\theta$ allow us to select $t \in T_1$ independently of k and so that $\langle a_s, z_k\rangle - \alpha_s < \langle a_t, z_k\rangle - \alpha_t$ for all $k \in I\!N$.

Claim 2: *We have $t \in K(\bar{z})$ for the index $t \in T_1$ selected above.*

Indeed, suppose by contradiction that $t \notin K(\bar{z})$. Combining this with $s \in K(\bar{z})$ tells us that $\langle a_t, \bar{z}\rangle - \alpha_t < \langle a_s, \bar{z}\rangle - \alpha_s$, and thus $\langle a_t, z_k\rangle - \alpha_t < \langle a_s, z_k\rangle - \alpha_s$ for all k sufficiently large. This clearly contradicts the choice of the index t and hence justifies the claim.

For the selected $s \in J_+(\bar{z}, \bar{v}_1)$ define now the index set

$$D_s := \Big\{ t \in T_1 \Big| \; \langle a_s, z_k\rangle - \alpha_s < \langle a_t, z_k\rangle - \alpha_t \;\text{ for all }\; k \in I\!N \Big\}.$$

It follows from Claim 2 that $\emptyset \neq D_s \subset K(\bar{z})$. We continue with the next assertion.

Claim 3: *$P \subset D_s$, where P was selected so that $K(z_k) = P$ for all $k \in I\!N$.*

Assuming the contrary, find $j \in P$ such that $j \notin D_s$ and pick $t \in D_s$. Employing this gives us

$$\langle a_s, z_k\rangle - \alpha_s < \langle a_t, z_k\rangle - \alpha_t \;\text{ for all }\; k \in I\!N. \tag{5.25}$$

Since $j \notin D_s$, there exists a number $k_0 \in I\!N$ for which we have

$$\langle a_j, x_{k_0}\rangle - \alpha_j \leq \langle a_s, x_{k_0}\rangle - \alpha_s. \tag{5.26}$$

Combining (5.25) for $k = k_0$ together with (5.26) leads us to the strict inequality

$$\langle a_j, x_{k_0}\rangle - \alpha_j < \langle a_t, x_{k_0}\rangle - \alpha_t,$$

which contradicts (5.24) due to $t \in K(\bar{z})$ and thus verifies the claim.

We proceed with proof of the theorem with the following claim.

Claim 4: *We have $\bar{v} \notin \mathrm{co}\Big\{ a_i \Big| \; i \in P \Big\} + \Big\{ d_i \Big| \; i \in Q \Big\}$.*

To verify the claim, suppose on the contrary that there exist vectors $\widehat{v}_1 \in \mathrm{co}\{a_i |\; i \in P\}$ and $\widehat{v}_2 \in \mathrm{cone}\{d_i |\; i \in Q\}$ such that $\bar{v} = \widehat{v}_1 + \widehat{v}_2$. This allows us to find numbers $\tau_i, \eta_i \geq 0$ with $\sum_{i\in P} \eta_i = 1$ such that $\widehat{v}_1 : = \sum_{i\in P} \eta_i a_i$ and $\widehat{v}_2 : = \sum_{i\in Q} \tau_i d_i$. Pick $t \in D_s \cap P$, which can be done by Claim 3. It follows from $t \in D_s$ that

$$\langle a_s, z_k\rangle - \alpha_s < \langle a_t, z_k\rangle - \alpha_t \ \text{ for all } \ k$$

while $t \in P$ results in the inequality

$$\langle a_i, z_k\rangle - \alpha_i \le \langle a_t, z_k\rangle - \alpha_t \ \text{ whenever } \ i \in K(\bar{z}).$$

Using these two facts together with $s \in J_+(\bar{z}, \bar{v}_1)$ yields

$$\langle \bar{v}_1, z_k\rangle - \sum_{i\in K(\bar{z})} \bar{\lambda}_i \alpha_i < \langle a_t, z_k\rangle - \alpha_t, \tag{5.27}$$

where $\bar{v}_1$ is from (4.7) and the multipliers $\bar{\lambda}_i$ are taken from (4.7). Remembering that $\langle a_i, z_k\rangle - \alpha_i = \langle a_j, z_k\rangle - \alpha_j$ for all $i, j \in P$ and taking (5.27) into account ensure that

$$\langle \bar{v}_1, z_k\rangle - \sum_{i\in K(\bar{z})} \bar{\lambda}_i \alpha_i < \langle \widehat{v}_1, z_k\rangle - \sum_{i\in P} \eta_i \alpha_i. \tag{5.28}$$

On the other hand, we know that $z_k \in \operatorname{dom}\theta$, $s' \in J_+(\bar{z}, \bar{v}_2)$, and $s' \notin I(z_k)$ which leads us to

$$\langle \bar{v}_2, z_k\rangle < \sum_{i\in I(\bar{z})} \bar{\mu}_i \beta_i \ \text{ and } \ \langle \widehat{v}_2, z_k\rangle = \sum_{i\in Q} \tau_i \beta_i. \tag{5.29}$$

Using (5.28) and (5.29) together with $\bar{v}_1 + \bar{v}_2 = \widehat{v}_1 + \widehat{v}_2$ gives us

$$\sum_{i\in K(\bar{z})} \bar{\lambda}_i \alpha_i + \sum_{i\in I(\bar{z})} \bar{\mu}_i \beta_i > \sum_{i\in P} \eta_i \alpha_i + \sum_{i\in Q} \tau_i \beta_i. \tag{5.30}$$

Appealing finally to Claim 1 along with the inclusions $Q \subset I(\bar{z})$ and $P \subset K(\bar{z})$, we arrive at a contradiction with (5.30) and hence verify this claim.

Now we are ready to finish the proof of the theorem. Remember that $(z_k, v_k) \xrightarrow{\partial\theta} (\bar{z}, \bar{v})$, which yields $v_k \in \partial\theta(z_k)$ for all $k \in I\!N$. It follows from (4.6) that $\partial\theta(z_k) = \operatorname{co}\{a_i |\ i \in P\} + \{d_i |\ i \in Q\}$ due to $K(z_k) = P$ and $I(z_k) = Q$. Hence we have $\bar{v} \in \operatorname{co}\{a_i |\ i \in P\} + \{d_i |\ i \in Q\}$ thus contradicting Claim 4 and showing that the assumption made after Claim 1 cannot be correct, while the opposite is the conclusion of the theorem. $\triangle$

Next, we provide a characterization for noncritical Lagrange multipliers via *upper Lipschitzian property* for the generalized equation (5.1). This is an extension of [14, Proposition 1.43] that provides such a characterization when the CPWL function θ in the framework of the generalized equation (5.1) is the indicator function $\delta(\cdot\,;\Omega)$ of the convex polyhedral set $\Omega\colon=\mathbb{R}^s\times\mathbb{R}^{n-s}_-$.

Theorem 5.5 (characterization of critical multipliers via upper Lipschitizan property of the KKT system). *Let $(\bar{x},\bar{v})$ be a solution of the generalized equation (5.1) for the parameters $(\bar{p}_1,\bar{p}_2)=(0,0)\in\mathbb{R}^n\times\mathbb{R}^m$. Let $\bar{z}\colon=\Phi(\bar{x})$ and denote by $K\colon=K(\bar{z})$, $I\colon=I(\bar{z})$, $J_1:=J_+(\bar{x},\bar{v}_1)$, and $J_2:=J_+(\bar{z},\bar{v}_2)$ the index sets from (4.5), and (4.9), respectively. Then the following properties are equivalent:*

(i) *The Lagrange multiplier $\bar{v}$ is noncritical.*

(ii) *There exist some number $\ell\geq 0$, and neighborhoods V of $\bar{p}_1\colon=0\in\mathbb{R}^n$ and W of $\bar{p}_2\colon=0\in\mathbb{R}^m$ such that for any solution $(x_{p_1p_2},v_{p_1p_2})$ of the generalized equation (5.1), associated with the pair $(p_1,p_2)\in V\times W$, close to $(\bar{x},\bar{v})$ we have*

$$\|x_{p_1p_2}-\bar{x}\|+\mathrm{dist}(v_{p_1p_2};\Lambda(\bar{x}))\leq\ell(\|p_1\|+\|p_2\|). \tag{5.31}$$

Proof. To justify implication (ii) $\Longrightarrow$ (i), assume by contradiction that there exists a pair $(\xi,\eta)\in\mathbb{R}^n\times\mathbb{R}^m$ with $\xi\neq 0$ for which properties (i) $-$ (iii) in Definition 5.2 are satisfied. Let $t>0$ and define $(x_t,v_t):=(\bar{x}+t\xi,\bar{v}+t\eta)$. Thus for sufficiently small t we get

$$\begin{aligned}\Psi(x_t,v_t)-\Psi(\bar{x},\bar{v}) &= \Big(f(x_t)-\nabla f(\bar{x})\Big)\\ &+ \Big(\nabla\Phi(x_t)-\nabla\Phi(\bar{x})\Big)^*\bar{v}+t\nabla\Phi(x_t)^*\eta\\ &= t\nabla f(\bar{x})\xi+o(t)+t(\nabla^2\Phi(\bar{x})\xi)^*\bar{v}+t\nabla\Phi(\bar{x})^*\eta+o(t)\\ &= t\Big(\nabla_x\Psi(\bar{x},\bar{v})\xi+\nabla\Phi(\bar{x})^*\eta\Big)+o(t)=o(t).\end{aligned}$$

Since $\Psi(\bar{x}, \bar{v}) = 0$, we conclude that

$$\Psi(x_t, v_t) = p_{1t} \quad \text{with} \quad p_{1t} := o(t). \tag{5.32}$$

Remember that $\Phi(x_t) = \Phi(\bar{x}) + t\nabla\Phi(\bar{x})\xi + o(t)$. Letting $z_t := \Phi(\bar{x}) + t\nabla\Phi(\bar{x})\xi$, we obtain

$$z_t = \Phi(x_t) + p_{2t} \quad \text{with} \quad p_{2t} := o(t). \tag{5.33}$$

It is not hard to see that $z_t \in \operatorname{dom}\theta$ for sufficiently small t, where the $\operatorname{dom}\theta$ was given by (4.2). To proceed we need to prove the following claim.

Claim 1. Given z_t as defined above and t sufficiently small, we have $J_1 \subset K(z_t)$ and $J_2 \subset I(z_t)$.

To prove the second inclusion $J_2 \subset I(z_t)$, take $i \in J_2$ and see that

$$\langle d_i, z_t \rangle = \langle d_i, \Phi(\bar{x}) \rangle + t\langle d_i, \nabla\Phi(\bar{x})\xi \rangle = 0$$

due to $\nabla\Phi(\bar{x})\xi \in \mathcal{K}(\bar{z}, \bar{v})$ and $J_2 \subset I(\bar{z})$. To prove the inclusion $J_1 \subset K(z_t)$, pick $i \in J_1$. To finish the proof of the claim, we have to show that $z_t \in C_i$, where the polyhedral set C_i is taken from (4.4). To see this, take $t \in K$ and then get $\langle a_i - a_t, \bar{z} \rangle = \alpha_i - \alpha_t$. It follows from $\nabla\Phi(\bar{x})\xi \in \mathcal{K}(\bar{z}, \bar{v})$ that $\langle a_i - a_t, \nabla\Phi(\bar{x})\xi \rangle \le 0$. These lead us to $\langle a_i - a_t, z_t \rangle \le \alpha_i - \alpha_t$ for $t \in K$. Similarly, we can show that $\langle a_i - a_t, z_t \rangle \le \alpha_i - \alpha_t$ for $t \in T_1 \setminus K$. Therefore we arrive at $\langle a_i - a_t, z_t \rangle \le \alpha_i - \alpha_t$ for $t \in T_1$ and hence $z_t \in C_i$.

We next prove that $v_t \in \partial\theta(z_t)$ when t is sufficiently small. To this end, it follows from Definition 5.2(i), and from (5.9) and Proposition 5.1 that

$$\eta \in \Big(\mathcal{K}(\bar{z}, \bar{v})\Big)^* = \Big(\mathcal{G}_{\{K,J_1\},\{I,J_2\}}\Big)^* = \mathcal{F}_{\{K,J_1\},\{I,J_2\}}.$$

Thus by (4.10) we get $\eta = \eta_1 + \eta_2$ so that

$$\begin{aligned}
&\eta_1 := \sum_{i,j\in J_1} \beta_{ij}(a_i - a_j) + \sum_{(i,j)\in(K\setminus J_1)\times J_1} \rho_{ij}(a_i - a_j) \quad \text{and} \quad \eta_2 := \sum_{s\in J_2} \tau_{1s} d_s + \sum_{s\in I\setminus J_2} \tau_{2s} d_s,\\
&\beta_{ij} \in \mathbb{R} \quad \text{for} \quad i,j \in J_1, \qquad\qquad \text{and} \quad \tau_{1s} \in \mathbb{R} \quad \text{for} \quad s \in J_2,\\
&\rho_{ij} \geq 0 \quad \text{for} \quad (i,j) \in (K \setminus J_1) \times J_1, \qquad \text{and} \quad \tau_{2s} \geq 0 \quad \text{for} \quad s \in I \setminus J_2.
\end{aligned} \tag{5.34}$$

We know that $K(z_t) \subset K(\bar{z})$ and $I(z_t) \subset I(\bar{z})$ for sufficiently small t. Pick $i_0 \in K(\bar{z}) \setminus K(z_t)$ and $j \in J_1$. Thus we deduce from Claim 1 that $j \in K(z_t)$, which together with $\nabla\Phi(\bar{x})\xi \in \mathcal{K}(\bar{z}, \bar{v})$ brings us to $\langle a_{i_0} - a_j, \nabla\Phi(\bar{x})\xi\rangle < 0$. This implies by Definition 5.2(iii) that $\rho_{i_0 j} = 0$ in (5.34); therefore we accomplish by (4.7) that

$$\begin{aligned}
v_{1t} := \bar{v}_1 + t\eta_1 &= \sum_{i\in K} \bar{\lambda}_i a_i + t\sum_{i,j\in J_1} \beta_{ij}(a_i - a_j) + t\sum_{(i,j)\in(K\setminus J_1)\times J_1} \rho_{ij}(a_i - a_j)\\
&= \sum_{i\in J_1} \bar{\lambda}_i a_i + t\sum_{i,j\in J_1} \beta_{ij}(a_i - a_j) + t\sum_{(i,j)\in(K(z_k)\setminus J_1)\times J_1} \rho_{ij}(a_i - a_j).
\end{aligned} \tag{5.35}$$

Taking $t > 0$ sufficiently small, we can find $\lambda'_{ti} \geq 0$, $i \in K(z_t)$, so that $\sum_{i\in K(z_t)} \lambda'_{ti} = \sum_{i\in K} \bar{\lambda}_i = 1$ and

$$v_{1t} = \sum_{i\in J_1} \lambda'_{ti} a_i + \sum_{i\in K(z_t)\setminus J_1} \lambda'_{ti} a_i. \tag{5.36}$$

Similarly, pick $s_0 \in I(\bar{z}) \setminus I(z_t)$ and observe by $\nabla\Phi(\bar{x})\xi \in \mathcal{K}(\bar{z}, \bar{v})$ that $\langle d_{s_0}, \nabla\Phi(\bar{x})\xi\rangle < 0$. So we obtain from Definition 5.2(iii) that $\tau_{2s_0} = 0$ in (5.34), which says that

$$\begin{aligned}
v_{2t} := \bar{v}_2 + t\eta_2 &= \sum_{s\in I} \bar{\mu}_s d_s + t\sum_{s\in J_2} \tau_{1s} d_s + t\sum_{s\in I\setminus J_2} \tau_{2s} d_s\\
&= \sum_{s\in I\setminus J_2} (\bar{\mu}_s + t\tau_{2s}) d_s + \sum_{s\in J_2} (\bar{\mu}_s + t\tau_{1s}) d_s\\
&= \sum_{s\in I(z_t)\setminus J_2} (t\tau_{2s}) d_s + \sum_{s\in J_2} (\bar{\mu}_s + t\tau_{1s}) d_s.
\end{aligned} \tag{5.37}$$

Taking into account (5.36) and (5.37) together with the inclusions $J_1 \subset K(z_t)$ and $J_2 \subset I(z_t)$

due to Claim 1 tells us that

$$v_t = v_{1t} + v_{2t} \in \text{co}\Big\{a_i \,\Big|\, i \in K(z_t)\Big\} + \Big\{\sum_{s\in I(z_t)} \mu_s d_s \,\Big|\, \mu_s \geq 0\Big\} = \partial\theta(z_t).$$

Using this along with (5.32) and (5.33) tells us that (x_t, v_t) is a solution for the generalized equation (5.1) associated with the parameters (p_{1t}, p_{2t}); therefore by (ii) we come up to

$$t\|\xi\| = \|x_t - \bar{x}\| \leq \ell(\|p_{1t}\| + \|p_{2t}\|) = \ell\|o(t)\|.$$

This confirms that $\xi = 0$, which is a contradiction, and hence completes the proof of implication (ii) $\Longrightarrow$ (i).

To prove implication (i) $\Longrightarrow$ (ii), it suffices to show that there exist some number $\ell \geq 0$ and neighborhoods V of $\bar{p}_1 = 0 \in \mathbb{R}^n$ and W of $\bar{p}_2 = 0 \in \mathbb{R}^m$ such that for any solution $(x_{p_1p_2}, v_{p_1p_2})$ of the generalized equation (5.1), associated with the parameters $(p_1, p_2) \in V \times W$, close to $(\bar{x}, \bar{v})$ we have

$$\|x_{p_1p_2} - \bar{x}\| \leq \ell(\|p_1\| + \|p_2\|). \tag{5.38}$$

Assume that estimate (5.38) holds. We next prove that there exists some number $\ell' \geq 0$ such that

$$\text{dist}(v_{p_1p_2}; \Lambda(\bar{x})) \leq \ell'(\|x_{p_1p_2} - \bar{x}\| + \|p_1\| + \|p_2\|), \tag{5.39}$$

which together with (5.38) justifies (5.31). To furnish it, observe that $\partial\theta(\bar{z})$ is a convex polyhedral set; therefore by [5, Theorem 2E.2] we find a positive number r, a matrix $A \in \mathbb{R}^{m\times r}$, and a vector $q \in \mathbb{R}^r$ such that the convex polyhedral set $\partial\theta(\bar{z})$ have an equivalent representation of the form

$$\partial\theta(\bar{z}) = \Big\{y \in \mathbb{R}^m \,\Big|\, Ay \leq q\Big\}.$$

Define now the set

$$\mathcal{D}_{\bar{x}}(\epsilon, \tau) = \Big\{v \in \mathbb{R}^m \,\Big|\, \Psi(\bar{x}, v) = \epsilon, \;\; Av \leq \tau\Big\}, \tag{5.40}$$

where $\epsilon \in \mathbb{R}^n$ and $\tau \in \mathbb{R}^r$. It is easy to observe that $\mathcal{D}_{\bar{x}}(0,q) = \Lambda(\bar{x})$, where the set $\Lambda(\bar{x})$ comes from (5.2). Let $\rho \geq 0$ be a Lipschitz constant for the mappings f and $\nabla\Phi$. Appealing now to the Hoffman lemma (see [5, Lemma 3C.4]) together with the fact that $v_{p_1p_2} \in \partial\theta(\Phi(x_{p_1p_2}) + p_2) \subset \partial\theta(\bar{z})$ for any $(p_1, p_2) \in V \times W$, we arrive at

$$\begin{aligned} \operatorname{dist}(v_{p_1p_2};\Lambda(\bar{x})) = \operatorname{dist}(v_{p_1p_2};\mathcal{D}_{\bar{x}}(0,q)) &\leq \kappa\|\Psi(\bar{x}, v_{p_1p_2})\| \\ &\leq \kappa\Big(\|\Psi(\bar{x}, v_{p_1p_2}) - \Psi(x_{p_1p_2}, v_{p_1p_2})\| \\ &+ \|\Psi(x_{p_1p_2}, v_{p_1p_2})\|\Big) \\ &\leq \kappa\Big(\rho\|x_{p_1p_2} - \bar{x}\| + \|p_1\|\Big) \\ &\leq \kappa\Big(\rho\|x_{p_1p_2} - \bar{x}\| + \|p_1\| + \|p_2\|\Big), \end{aligned} \tag{5.41}$$

which justifies (5.39). Now we turn to proving (5.38). Suppose on the contrary that for any $k \in I\!N$ there exist a pair $(p_{1k}, p_{2k}) \in I\!B_{\frac{1}{k}}(\bar{v}) \times I\!B_{\frac{1}{k}}(\bar{w})$ with $(\bar{p}_1, \bar{p}_2) = (0,0) \in \mathbb{R}^n \times \mathbb{R}^m$ and a solution (x_k, v_k) of the generalized equation (5.1) associated with the pair (p_{1k}, p_{2k}), converging to $(\bar{x}, \bar{v})$, such that

$$\frac{\|x_k - \bar{x}\|}{\|p_{1k}\| + \|p_{2k}\|} \to \infty \quad \text{as} \quad k \to \infty,$$

which amounts to

$$\frac{\|p_{1k}\| + \|p_{2k}\|}{\|x_k - \bar{x}\|} \to 0 \quad \text{as} \quad k \to \infty.$$

This tells us that $p_{1k} = o(\|x_k - \bar{x}\|)$ and $p_{2k} = o(\|x_k - \bar{x}\|)$. Let $z_k :\ = \Phi(x_k) + p_{2k}$ and observe by (5.1) that $(z_k, v_k) \in \partial\theta$. Applying Theorem 5.4, we conclude that $J_1 \subset K(z_k) \subset K(\bar{z})$ and $J_2 \subset I(z_k) \subset I(\bar{z})$. Passing to a subsequence of (z_k, v_k) if necessary, we can assume without loss of generality that there exist subsets $P \subset K(\bar{z})$ and $Q \subset I(\bar{z})$ such that

$$P = K(z_k) \quad \text{and} \quad Q = I(z_k) \ \text{ whenever } \ k \in I\!N. \tag{5.42}$$

Remember that (x_k, v_k) is a solution of the generalized equation (5.1) associated with the pair (p_{1k}, p_{2k}), so we deduce that

$$\begin{aligned} o(\|x_k-\bar{x}\|)=p_{1k} &= \Psi(x_k,v_k) \\ &= \Psi(x_k,\bar{v})-\Psi(\bar{x},\bar{v})+\nabla\Phi(x_k)^*(v_k-\bar{v}) \\ &= \nabla_x\Psi(\bar{x},\bar{v})(x_k-\bar{x})+\nabla\Phi(\bar{x})^*(v_k-\bar{v})+o(\|x_k-\bar{x}\|). \end{aligned} \tag{5.43}$$

Employing [41, Proposition 3.2], we find $\lambda_{ik} \geq 0$ with $i \in P$, and $\mu_{ik} \geq 0$ with $i \in Q$ for which v_k has a representation of the form $v_k = v_{1k} + v_{2k}$, where

$$v_{1k}=\sum_{i\in P}\lambda_{ik}a_i \quad \text{and} \quad v_{2k}=\sum_{i\in Q}\mu_{ik}d_i \quad \text{with} \quad \sum_{i\in P}\lambda_{ik}=1.$$

This together with (4.7) and (5.43) implies that

$$\begin{aligned} -\nabla_x\Psi(\bar{x},\bar{v})\frac{(x_k-\bar{x})}{\|x_k-\bar{x}\|}+\frac{o(\|x_k-\bar{x}\|)}{\|x_k-\bar{x}\|} &= \frac{1}{\|x_k-\bar{x}\|}\nabla\Phi(\bar{x})^*\Big[(v_{1k}-\bar{v}_1)+(v_{2k}-\bar{v}_2)\Big] \\ = \frac{1}{\|x_k-\bar{x}\|}\nabla\Phi(\bar{x})^*\Big[\Big(\sum_{i\in P}\lambda_{ik}a_i-\sum_{j\in J_1}\bar{\lambda}_j a_j\Big) &+ \Big(\sum_{i\in Q}\mu_{ik}d_i-\sum_{j\in J_2}\bar{\mu}_j d_j\Big)\Big] \\ = \frac{1}{\|x_k-\bar{x}\|}\nabla\Phi(\bar{x})^*\Big[\sum_{i\in P}\lambda_{ik}\sum_{j\in J_1}\bar{\lambda}_j(a_i-a_j) &+ \Big(\sum_{i\in Q}\mu_{ik}d_i-\sum_{j\in J_2}\bar{\mu}_j d_j\Big)\Big] \\ \in \nabla\Phi(\bar{x})^*\Big(\operatorname{span}\Big\{a_i-a_j\Big|\, i,j\in J_1\Big\} &+ \Big\{a_i-a_j\Big|\,(i,j)\in(P\setminus J_1)\times J_1\Big\} \\ &+ \Big\{d_j\,\Big|\, j\in Q\setminus J_2\Big\}+\operatorname{span}\Big\{d_j\,\Big|\, j\in J_2\Big\}\Big). \end{aligned} \tag{5.44}$$

Assume without loss of generality that $\frac{x_k-\bar{x}}{\|x_k-\bar{x}\|} \to \xi$ as $k \to \infty$. Because the set on the right-hand side of (5.44) is closed, by passing to the limit we get

$$\begin{aligned} -\nabla_x\Psi(\bar{x},\bar{v})\xi \in \nabla\Phi(\bar{x})^*\Big(\operatorname{span}\Big\{a_i-a_j\Big|\, i,j\in J_1\Big\} &+ \Big\{a_i-a_j\Big|\,(i,j)\in(P\setminus J_1)\times J_1\Big\} \\ &+ \Big\{d_j\,\Big|\, j\in Q\setminus J_2\Big\}+\operatorname{span}\Big\{d_j\,\Big|\, j\in J_2\Big\}\Big). \end{aligned}$$

Therefore we find some vector $\eta = \eta_1 + \eta_2$ with

$$\begin{aligned} &\eta_1 \in \operatorname{span}\Big\{a_i-a_j\Big|\, i,j\in J_1\Big\}+\Big\{a_i-a_j\Big|\,(i,j)\in(P\setminus J_1)\times J_1\Big\}, \\ &\eta_2 \in \Big\{d_j\,\Big|\, j\in Q\setminus J_2\Big\}+\operatorname{span}\Big\{d_j\,\Big|\, j\in J_2\Big\} \end{aligned}$$

for which we have $\nabla_x\Psi(\bar{x},\bar{v})\xi + \nabla\Phi(\bar{x})^*\eta = 0$. This tells us that

$$\eta_1 = \sum_{i,j\in J_1}\gamma_{ij}(a_i - a_j) + \sum_{(i,j)\in(P\setminus J_1)\times J_1}\gamma'_{ij}(a_i - a_j) \quad \text{and} \quad \eta_2 = \sum_{t\in J_2}\tau_t d_t + \sum_{t\in Q\setminus J_2}\tau'_t d_t \tag{5.45}$$

for some numbers $\gamma_{ij} \in \mathbb{R}$, $\gamma'_{ij} \geq 0$, $\tau_t \geq 0$, and $\tau'_t \in R$. We now claim that $\nabla\Phi(\bar{x})\xi \in \mathcal{K}(\bar{z},\bar{v})$, which together with (5.45) confirms that properties (i) and (ii) in Definition 5.2 are satisfied. To prove the claim, let $i,j \in J_1$ and conclude by the inclusion $J_1 \subset P \subset K(\bar{z})$ together with (4.7) that

$$\langle a_i - a_j, z_k - \bar{z}\rangle = \langle a_i - a_j, \Phi(x_k) + p_{2k} - \Phi(\bar{x})\rangle = 0,$$

so we deduce from $p_{2k} = o(\|x_k - \bar{x}\|)$ that

$$\langle a_i - a_j, \nabla\Phi(\bar{x})\frac{x_k - \bar{x}}{\|x_k - \bar{x}\|} + \frac{o(x_k - \bar{x})}{\|x_k - \bar{x}\|}\rangle = 0.$$

This results in

$$\langle a_i - a_j, \nabla\Phi(\bar{x})\xi\rangle = 0 \qquad \text{whenever} \quad i,j \in J_1. \tag{5.46}$$

Assume now $i \in K \setminus J_1$ and $j \in J_1$; by the similar arguments as above, we can justify that

$$\langle a_i - a_j, \nabla\Phi(\bar{x})\xi\rangle \leq 0 \qquad \text{whenever} \quad (i,j) \in (K \setminus J_1) \times J_1. \tag{5.47}$$

Pick now $t \in J_2$ and observe that $\langle d_t, \Phi(x_k) + p_{2k} - \Phi(\bar{x})\rangle = 0$ because of the inclusion $J_2 \subset Q \subset I$. Combining this with $p_{2k} = o(\|x_k - \bar{x}\|)$ allows us to get

$$\langle d_t, \nabla\Phi(\bar{x})\frac{x_k - \bar{x}}{\|x_k - \bar{x}\|} + \frac{o(x_k - \bar{x})}{\|x_k - \bar{x}\|}\rangle = 0,$$

so we come up to

$$\langle d_t, \nabla\Phi(\bar{x})\xi\rangle = 0 \qquad \text{whenever} \quad t \in J_2. \tag{5.48}$$

For any $t \in I \setminus J_2$ we have $\langle d_t, \Phi(x_k) + p_{2k} - \Phi(\bar{x})\rangle \leq 0$, and by the similar arguments it yields $\langle d_t, \nabla\Phi(\bar{x})\xi\rangle \leq 0$. Using this together with (5.46)-(5.48), we accomplish that $\nabla\Phi(\bar{x})\xi \in \mathcal{K}(\bar{z},\bar{v})$ via representation (5.5), hence justifies the claim. It is not hard to see that (5.46) holds if J_1 is replaced by P. Similarly, inequality (5.48) is still true provided that J_2 is replaced by Q. Employing these observations along with (5.45), we arrive at $\langle \eta, \nabla\Phi(\bar{x})\xi\rangle = 0$, which says that the Lagrange multiplier $\bar{v}$ is critical, a contradiction. This finishes the proof. $\triangle$

5.2 Full Stability and Critical Multipliers

In this section we consider the perturbed version of problem (5.11) defined by

$$\text{minimize} \;\; \varphi_0(x, p_2) + \theta(\Phi(x, p_2)) - \langle p_1, x\rangle \;\; \text{subject to} \;\; x \in \mathbb{R}^n \tag{5.49}$$

with $(p_1, p_2) \in \mathbb{R}^n \times \mathbb{R}^d$. Fix $\gamma > 0$ and $(\bar{x}, \bar{p}_1, \bar{p}_1)$ with $\Phi(\bar{x}, \bar{p}_2) \in \operatorname{dom}\theta$. To proceed instead of working with the fully perturbed problem (5.49), we need to restrict our attention to the canonical perturbed version of problem (5.11) given by

$$\text{minimize} \;\; \varphi_0(x) + \theta(\Phi(x) + p_2) - \langle p_1, x\rangle \;\; \text{subject to} \;\; x \in \mathbb{R}^n \tag{5.50}$$

with $(p_1, p_2) \in \mathbb{R}^n \times \mathbb{R}^m$. We next demonstrate that the full stability for the canonical perturbed problem (5.11) excludes the existence of critical multipliers, which are the main source for *slow primal convergence* in the Newton-type algorithms.

Theorem 5.6 (excluding critical multipliers via full stability). *Let $\bar{x}$ be a feasible solution to* (5.50) *for the parameter pair $(\bar{p}_1, \bar{p}_2) = (0, 0) \in \mathbb{R}^n \times \mathbb{R}^m$, and let $\theta \in CPWL$. Let $\bar{x}$ be a fully stable local minimizer of the canonical perturbed problem* (5.50)*. Then the Lagrange multiplier set $\Lambda_{com}(\bar{x})$ from* (5.12) *does not include any critical multipliers.*

Proof. We first show that the assumed full stability implies the validity of the constraint qualification (5.20). To this end, let $\eta \in \partial^\infty\theta(\Phi(\bar{x})) \cap \ker\nabla\Phi(\bar{x})^*$. Since the CPWL function θ is convex, we have $\partial^\infty\theta(\Phi(\bar{x})) = N(\Phi(\bar{x}); \operatorname{dom}\theta)$. Select $p_1 = \bar{p}_1 = 0$ and $p_2 = t\eta$ with $t \downarrow 0$. By assumption, there exist $\ell \geq 0$ and a solution of problem (5.50), denoted by $x_{p_1p_2}$, for which we have $\|x_{p_1p_2} - \bar{x}\| \leq \ell\|p_2\| = \ell t\|\eta\|$. It follows from $\Phi(x_{p_1p_2}) + p_2 \in \operatorname{dom}\theta$ and

$\Phi(x_{p_1p_2}) = \Phi(\bar{x}) + \nabla\Phi(\bar{x})^*(x_{p_1p_2} - \bar{x}) + o(\|x_{p_1p_2} - \bar{x}\|)$ that

$$\begin{aligned} 0 &\geq \langle \eta, \Phi(x_{p_1p_2}) + p_2 - \Phi(\bar{x}) \rangle \\ &= \langle \eta, \nabla\Phi(\bar{x})^*(x_{p_1p_2} - \bar{x}) + o(\|x_{p_1p_2} - \bar{x}\|) + p_2 \rangle \\ &= \langle \eta, o(\|x_{p_1p_2} - \bar{x}\|) \rangle + t\|\eta\|^2, \end{aligned}$$

which tells us that $\eta = 0$. This justifies the claim.

Pick $\bar{v} \in \Lambda_{\text{com}}(\bar{x})$, where the Lagrange multiplier set $\Lambda_{\text{com}}(\bar{x})$ is defined by (5.12). We need to prove that $\bar{v}$ is noncritical. Consider the KKT system of problem (5.50), which can be written as a generalized equation

$$\begin{bmatrix} p_1 \\ p_2 \end{bmatrix} \in \begin{bmatrix} \nabla_x L(x,v) \\ -\Phi(x) \end{bmatrix} + \begin{bmatrix} 0 \\ (\partial\theta)^{-1}(v) \end{bmatrix} \tag{5.51}$$

with $\nabla_x L(x,v) \colon = \nabla\varphi_0(x) + \nabla\Phi(x)^* v$ being the Lagrange function for problem (5.50). By Theorem 5.5 it suffices to show that there exist some number $\ell \geq 0$ and neighborhoods V of $\bar{p}_1 = 0 \in \mathbb{R}^n$ and W of $\bar{p}_2 = 0 \in \mathbb{R}^m$ such that for any solution $(x_{p_1p_2}, v_{p_1p_2})$ of the generalized equation (5.51), associated with the pair $(p_1, p_2) \in V \times W$ and close to $(x_{\bar{p}_1\bar{p}_2}, v_{\bar{p}_1\bar{p}_2}) = (\bar{x}, \bar{v})$, the upper Lipschitzian estimate (5.31) holds. To this end, since $\bar{x}$ is a fully stable local minimizer of problem (5.50), it follows from [39, Proposition 6.1] that there exist some neighborhoods $\widetilde{V} \times \widetilde{W}$ of $(\bar{p}_1, \bar{p}_2)$ and $\widetilde{U}$ of $\bar{x}$ for which the set-valued mapping

$$S(p_1, p_1) \colon = \left\{ x \in \mathbb{R}^n \;\middle|\; p_1 \in \nabla\varphi_0(x) + \nabla\Phi(x)^*\partial\theta(\Phi(x) + p_2) \right\}$$

admits a Lipschitzian single-valued graphical localization on $\widetilde{V} \times \widetilde{W} \times \widetilde{U}$. This amounts to saying that there exists a Lipschitzian single-valued mapping $s \colon \widetilde{V} \times \widetilde{W} \to \widetilde{U}$ such that $S \cap (\widetilde{V} \times \widetilde{W} \times \widetilde{U}) = s$. Letting now $V = \widetilde{V}$ and $W = \widetilde{W}$, pick any solution $(x_{p_1p_2}, v_{p_1p_2})$ of the generalized equation (5.51), associated with the pair $(p_1, p_2) \in V \times W$, close to $(\bar{x}, \bar{v})$. This implies that $x_{p_1p_2} \in S(p_1, p_2)$; therefore we can find some number $\ell \geq 0$ such that

$$\|x_{p_1p_2} - \bar{x}\| = \|x_{p_1p_2} - x_{\bar{p}_1\bar{p}_2}\| \leq \ell(\|p_1 - \bar{p}_1\| + \|p_2 - \bar{p}_2\|) = \ell(\|p_1\| + \|p_2\|).$$

As we showed in the proof of Theorem 5.5, the above estimate justifies the validity of estimate (5.39), which therefore leads us to the upper Lipschitzian estimate (5.31). This completes the proof. △

REFERENCES

[1] J. F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer, New York, 2000.

[2] J. F. Bonnans and A. Sulem, Pseudopower expansion of solutions of generalized equations and constrained optimization problems, *Math. Program.* **70** (1995), 123–148.

[3] C. Ding, D. Sun and J. J. Ye, First order optimality conditions for mathematical programs with semidefinite cone complementarity constraints, *Math. Program.* **147** (2014), 539–579.

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

[5] A. L. Dontchev and R. T. Rockafellar, *Implicit Functions and Solution Mappings: A View from Variational Analysis*, Springer, Dordrecht, 2009.

[6] D. Drusvyatskiy and A. S. Lewis, Tilt stability, uniform quadratic growth, and strong metric regularity of the subdifferential, *SIAM J. Optim.* **23** (2013), 256–267.

[7] D. Drusvyatskiy, B. S. Mordukhovich and T. T. A. Nghia, Second-order growth, tilt stability, and metric regularity of the subdifferential, *J. Convex Anal.* **21** (2014), 1165–1192.

[8] A. C. Eberhard and R. Wenczel, A study of tilt-stable optimality and sufficient conditions, *Nonlinear Anal.* **75** (2012), 1260–1281.

[9] K. Emich and R. Henrion, A simple formula for the second-order subdifferential of maximum functions, *Vietnam J. Math.* **42** (2014), 467–478.

[10] F. Facchinei and J. S. Pang, *Finite-Dimesional Variational Inequalities and Complementarity Problems*, Springer, New York, 2003.

[11] R. Henrion, A. Y. Kruger and J. V. Outrata, Some remarks on stability of generalized equations, *J. Optim. Theory Appl.* **159** (2013), 681–697.

[12] R. Henrion, J. V. Outrata and T. Surowiec, On regular coderivatives in parametric equilibria with non-unique multipliers, *Math. Program.* **136** (2012), 111–131.

[13] A. D. Ioffe and J. V. Outrata, On Metric and Calmness Qualification Conditions in Subdifferential Calculus, *Set-Valued Anal.* **16** (2008), 199–227.

[14] A. F. Izmailov and M. V. Solodov, *Newton-Type Methods for Optimization and Variational Problems*, Springer, New York, 2014.

[15] H. Th. Jongen, W. Möbert, J. J. Rückmann and K. Tammer, Implicit functions and sensitivity of stationary points, *Linear Algebra Appl.* **95** (1987), 97–109.

[16] P. Kenderov, Semi-continuity of set-valued monotone mappings, *Fund. Math.* **88** (1975), 61–69.

[17] D. Klatte and B. Kummer, *Nonsmooth Equations in Optimization*, Kluwer, Dordrecht, 2002.

[18] M. Kojima, Strongly stable stationary solutions in nonlinear programming, in *Analysis and Computation of Fixed Points* (S. M. Robinson, ed.), pp. 93–138, Academic Press, New York, 1980.

[19] B. Kummer, Newton's method based on generalized derivatives for nonsmooth functions: Convergence analysis, in: *Advances in Optimization* (W. Oettli and D. Pallaschke, eds.), pp. 171–194, *Lecture Notes Econ. Math. Sci.* **382**, Springer, Berlin, 1992.

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

[21] A. B. Levy and R. T. Rockafellar, Variational conditions and the proto-differentiation of partial subgradient mappings, *Nonlin. Anal.* **26** (1996), 1951–1964.

[22] A. S. Lewis and S. Zhang, Partial smoothness, tilt stability, and generalized Hessians, *SIAM J. Optim.* **23** (2013), 74–94.

[23] S. Lu, Implications of the constant rank constraint qualification, *Math. Program.* **126** (2011), 365–392.

[24] B. S. Mordukhovich, Maximum principle in problems of time optimal control with nonsmooth constraints, *J. Appl. Math. Mech.* **40** (1976), 960–969.

[25] B. S. Mordukhovich, Metric approximations and necessary optimality conditions for general classes of extremal problems, *Soviet Math. Dokl.* **22** (1980), 526–530.

[26] B. S. Mordukhovich, Sensitivity analysis in nonsmooth optimization, in *Theoretical Aspects of Industrial Design* (D. A. Field and V. Komkov, eds.), pp. 32–46, *Proceed. Applied Math.* **58**, SIAM, Philadelphia, 1992.

[27] B. S. Mordukhovich, *Variational Analysis and Generalized Differentiation, I: Basic Theory; II: Applications*, Springer, Berlin, 2006.

[28] B. S. Mordukhovich, N. M. Nam and N. T. Y. Nhi, Partial second-order subdifferentials in variational analysis and optimization, *Numer. Func. Anal. Optim.* **35** (2014), 1113–1151.

[29] B. S. Mordukhovich and T. T. A. Nghia, Second-order variational analysis and characterizations of tilt-stable optimal solutions in infinite-dimensional spaces, *Nonlinear Anal.* **86** (2013), 159–180.

[30] B. S. Mordukhovich and T. T. A. Nghia, Second-order characterizations of tilt stability with applications to nonlinear programming, *Math. Program.* **149** (2015), 83–104.

[31] B. S. Mordukhovich and T. T. A. Nghia, Full Lipschitzian and Hölderian stability in optimization with applications to mathematical programming and optimal control, *SIAM J. Optim.* **24** (2014), 1344–1381.

[32] B. S. Mordukhovich and T. T. A. Nghia, Local strong maximal monotonicity and full stability for parametric variational systems, *SIAM J. Optim.* **26** (2016), 1032–1059.

[33] B. Mordukhovich, T. T. A. Nghia and R. T. Rockafellar, Full stability in finite-dimensional optimization, *Math. Oper. Res.* **40** (2015), 226–252.

[34] B. S. Mordukhovich and J. V. Outrata, Tilt stability in nonlinear programming under Mangasarian-Fromovitz constraint qualification, *Kybernetika* **49** (2013), 446–464.

[35] B. S. Mordukhovich, J. V. Outrata and H. Ramírez C., Second-order variational analysis in conic programming with applications to optimality and stability, *Set-Valued Var. Anal.* **23** (2015), 687–704.

[36] B. S. Mordukhovich, J. V. Outrata and H. Ramírez C., Graphical derivatives of solution maps to parameterized equilibria with conic constraints, *SIAM J. Optim.* **25** (2015), 76–101.

[37] B. S. Mordukhovich, J. V. Outrata and M. E. Sarabi, Full stability of locally optimal solution in second-order cone programming, *SIAM J. Optim.* **24** (2014), 1581–1613.

[38] B. S. Mordukhovich and R. T. Rockafellar, Second-order subdifferential calculus with application to tilt stability in optimization, *SIAM J. Optim.* **22** (2012), 953–986.

[39] B. S. Mordukhovich, R. T. Rockafellar and M. E. Sarabi, Characterizations of full stability in constrained optimization, *SIAM J. Optim.* **23** (2013), 1810–1849.

[40] B. S. Mordukhovich and M. E. Sarabi, Variational analysis and full stability of optimal solutions to constrained and minimax problems, *Nonlinear Anal.* **121** (2015), 36–53.

[41] B. S. Mordukhovich and M. E. Sarabi, Variational analysis of piecewise linear functions with applications to optimization and stability, I: Second-order theory, *Nonlinear Anal.* **132** (2016), 240–273.

[42] B. Mordukhovich and M. E. Sarabi, Second-order analysis of piecewise linear functions with applications to optimization and stability, *J. Optim. Theory Appl.*, (2016), 1-23.

[43] J. V. Outrata and H. Ramírez C., On the Aubin property of critical points to perturbed second-order cone programs, *SIAM J. Optim.* **21** (2011), 798–823.

[44] R. A. Poliquin and R. T. Rockafellar, Prox-regular functions in variational analysis, *Trans. Amer. Math. Soc.* **348** (1996), 1805–1838.

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

[46] R. A. Poliquin and R. T. Rockafellar, A calculus of prox-regularity, *J. Convex Anal.* **17** (2010), 203–210.

[47] S. M. Robinson, Strongly regular generalized equations, *Math. Oper. Res.* **5** (1980), 43–62.

[48] S. M. Robinson, Local structure of feasible sets in nonlinear programming, II: Nondegeneracy, *Math. Program. Stud.* **22** (1984), 217–230

[49] S. M. Robinson, Aspects of the projector on prox-regular sets, in: *Variational Analysis and Applications* (F. Giannessi and A. Maugeri, eds.), *Nonconvex Optim. Appl.* **79** (2005), 963–973.

[50] R. T. Rockafellar, Maximal monotone relations and the second derivatives of nonsmooth functions, *Ann. Inst. H. Poincaré: Analyse Non Linéaire* **2** (1985), 167–184.

[51] R. T. Rockafellar, First- and second-order epi-differentiability in nonlinear programming, *Trans. Amer. Math. Soc.* **307** (1988), 75–108.

[52] R. T. Rockafellar, Extended nonlinear programming, in: *Nonlinear Optimization and Related Topics* (G. Di Pillo and F. Giannessi. eds.), pp. 381–399, *Applied Optimization* **36**, Kluwer Academic Publishers, Dordrecht, 2000.

[53] R. T. Rockafellar and R. J-B. Wets, *Variational Analysis*, Springer, Berlin, 2006.

ABSTRACT

VARIATIONAL ANALYSIS AND STABILITY IN OPTIMIZATION

by

M. Ebrahim Sarabi

August 2016

Advisor: Dr. Boris. S. Mordukhovich

Major: Mathematics (Applied)

Degree: Doctor of Philosophy

The dissertation is devoted to the study of the so-called *full Lipschitzian stability* of local solutions to finite-dimensional parameterized problems of constrained optimization, which has been well recognized as a very important property from both viewpoints of optimization theory and its applications. Employing second-order subdifferentials of variational analysis, we obtain necessary and sufficient conditions for fully stable local minimizers in general classes of constrained optimization problems including problems of composite optimization as well as problems of nonlinear programming with twice continuously differentiable data. Based on our recent explicit calculations of the second-order subdifferential for convex piecewise linear functions, we establish relationships between nondegeneracy and second-order qualification for fully amenable compositions involving piecewise linear functions and obtain new applications of the developed second-order theory to full stability in composite optimization and constrained minimax problems, strong regularity of associate generalized equations and strong stability of stationary points for composite optimization. Finally, we discuss the important concept of critical multipliers for composite optimization problems and characterize it via second-order subdifferentials. Then we demonstrate that full stability can rule out the existence of critical multipliers in the mentioned framework.

AUTOBIOGRAPHICAL STATEMENT

M. Ebrahim Sarabi

Education

- *Ph.D. in Applied Mathematics*, Wayne State University, August (2016)
- *M.S. in Applied Mathematics*, Wayne State University, August (2014)
- *M.S. in Applied Mathematics*, Kharazmi University, Tehran, Iran (2006)
- *B.S. in Applied Mathematics*, Kharazmi University, Tehran, Iran (2004)

Awards

- *Karl W. and Helen L. Folley Endowed Scholarship*, Department of Mathematics, Wayne State University, 2015.
- *Karl W. and Helen L. Folley Endowed Scholarship*, Department of Mathematics, Wayne State University, 2014.
- *AMS Travel Grant Support*, 2015-2016.
- *Thomas C. Rumble University Graduate Fellowship*, Wayne State University, September 2014-May 2015.
- *University Graduate Research Fellowship*, Wayne State University, September 2011-September 2012.

Publications and Preprints

1. B. Mordukhovich and M. E. Sarabi, *Critical multipliers for composite optimization problems and its applications*, In preparation.
2. B. Mordukhovich and M. E. Sarabi, *Stability analysis for composite optimization problems and parametric variational systems*, submitted.
3. B. Mordukhovich and M. E. Sarabi, *Generalized differentiation of piecewise linear functions in second-order variational analysis*, Nonlinear Anal., 132 (2016), 240–273.
4. B. Mordukhovich and M. E. Sarabi, *Second-order analysis of piecewise linear functions with applications to optimization and stability*, J. Optim. Theory Appl., (2016), 1-23.
5. B. Mordukhovich, J. Outrata and M. E. Sarabi, *Full stability in second-order cone programming*, SIAM J. Optim., 24 (2014), 1581–1613
6. B. Mordukhovich and M. E. Sarabi, *Variational analysis and full stability of optimal solutions to constrained and minimax problems*, Nonlinear Anal., 121 (2015), 36–53
7. B. S. Mordukhovich, R. T. Rockafellar and M. E. Sarabi, *Characterizations of full stability in constrained optimization*, SIAM J. Optim., 23 (2013), 1810–1849