ESTIMATES OF GENERALIZED HESSIANS FOR OPTIMAL VALUE FUNCTIONS IN MATHEMATICAL PROGRAMMING

ALAIN B. ZEMKOHO
School of Mathematics, University of Southampton
SO17 1BJ Southampton, UK
a.b.zemkoho@soton.ac.uk

Abstract. The *optimal value function* is one of the basic objects in the field of mathematical optimization, as it allows the evaluation of the variations in the *cost/revenue* generated while *minimizing/maximizing* a given function under some constraints. In the context of stability/sensitivity analysis, a large number of publications have been dedicated to the study of continuity and differentiability properties of the optimal value function. The differentiability aspect of works in the current literature has mostly been limited to first order analysis, with focus on estimates of its directional derivatives and subdifferentials, given that the function is typically nonsmooth. With the progress made in the last two to three decades in major subfields of optimization such as robust, minmax, semi-infinite and bilevel optimization, and their connection to the optimal value function, there is a crucial need for a *second order analysis of the generalized differentiability properties* of this function. This type of analysis will promote the development of robust solution methods, such as the Newton method, which is very popular in nonlinear optimization. The main goal of this paper is to provide results in this direction. In fact, we derive estimates of the *generalized Hessian* (also known as the second order subdifferential) for the optimal value function. Our results are based on two handy tools from parametric optimization, namely the optimal solution and Lagrange multiplier mappings, for which completely detailed estimates of their generalized derivatives are either well-known or can easily be obtained.

Contents

Date: October 16, 2017.
2010 *Mathematics Subject Classification*. 90C31, 90C30.
Key words and phrases. Parametric optimization, optimal value function, generalized Hessian, second order subdifferential, sensitivity analysis, coderivative.
This work is supported by an EPSRC grant with reference EP/P022553/1.

1. Introduction

1.1. **Aim of the work.** Considering the functions f and g defined from $\mathbb{R}^{n+m}$ to $\mathbb{R}$ and $\mathbb{R}^p$, respectively, we are interested in the parametric optimization problem

$$\min_y \{f(x,y)|\, g(x,y) \leq 0\}. \tag{1.1}$$

We only consider inequality constraints, in order to focus our attention on the main ideas. Note however that all the results in this paper remain valid, with the corresponding adjustments, if we include equality constraints to problem (1.1). Our focus will be on the optimal value function

$$\varphi(x) := \min_y \{f(x,y)|\, g(x,y) \leq 0\}, \tag{1.2}$$

related to the parametric optimization problem (1.1). Another object closely related to problem (1.1) is the optimal solution set-valued mapping $S : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ defined by

$$S(x) := \arg\min_y \{f(x,y)|\, g(x,y) \leq 0\}. \tag{1.3}$$

We assume throughout the paper that $S(x) \neq \emptyset$ for all $x \in \mathbb{R}^n$. As a consequence, φ will be finite-valued at all $x \in \mathbb{R}^n$. For most results obtained in this paper, we can easily accommodate the case where φ (1.2) is an extended real-valued function. But to concentrate on the main points, we let this specific case for future analysis.

The function φ (1.2) has payed a major role in the development and understanding of the structure of the underlying parametric optimization problem (1.1), and has been substantially analyzed in the literature. Initial work on stability/sensitivity analysis of optimization problems is almost as old as the field of optimization itself, given that early works on linear programming and the simplex method already provided interesting insights on the behavior of optimal values under perturbations; see, example, the 12th Chapter of the 1963 book by Dantzig [5].

The study of continuity and differential properties of φ, in the context of nonlinear optimization, which is our main focus here, grew dramatically following works by Fiacco [11], Gauvin and Dubeau [12], amongst many others. Recent publications on the topic include the papers [15, 16], where the tools by Mordukhovich are used to provide different types of subdifferentials estimates for φ.

It should be emphasized that most of the aforementioned works focus on the derivation of continuity properties and the estimation of directional derivatives and subdifferentials. As far as second order differentiation properties for φ are concerned, a few publications have been devoted to the estimation of second order-type directional derivatives; see, e.g. [3]. We are however not aware of any work pursuing generalized Hessian evaluations for φ. Fiacco [11, Chapter 3] provides Hessian formulas for φ in the case where S (1.3) is singled-valued and continuously differentiable. This assumption is very restrictive and cannot hold for most applications. We consider various scenarios in this paper, including the latter one, where our results coincide with those by Fiacco; see details in the next subsection and in Section 4.

It is also important to recall that the optimal value function naturally appears either in the constraints or objective functions of many mainstream optimization problems, including *robust* [1], *minmax* [4], *semi-infinite* [21], and *bilevel optimization* [6]. In order to be able to extend some standard optimization methods, including the Newton method, to such problems, it is fundamental to be able to understand the basic structure of the corresponding Hessian in the nonsmooth setting, given that φ is a typically nondifferentiable function.

Our main goal in this paper is to provide estimates of the Generalized Hessian of the optimal value function φ, in the sense of Mordukhovich. Note that for $\bar{x} \in \mathbb{R}^n$ and $\underline{x} \in \partial\varphi(\bar{x})$, the generalized Hessian of φ in the sense of Mordukhovich is defined by

$$\partial^2\varphi(\bar{x}|\underline{x})(\underline{x}^*) := D^*(\partial\varphi)(\bar{x}|\underline{x})(\underline{x}^*) \tag{1.4}$$

for all $\underline{x}^* \in \mathbb{R}^n$. Here, $\partial\varphi$ denotes the subdifferential of φ in any standard sense (convex analysis, Clarke or Mordukhovich) and $D^*(\partial\varphi)$ stands for the coderivative of Mordukhovich [14]. In the next subsection, we provide a general feeling of how the estimates developed in this paper look like. Further details on the mathematical tools and the proofs of the main results are provided in Sections 2, 3, and 4, respectively.

1.2. **Summary of the main results–outline of the paper.** We assume throughout the paper that the functions f and g, cf. (1.1), are twice continuously differentiable. If S (1.3) reduces to a single-valued function around a given point, the notation symbol will be adjusted to s. We use the set-valued map $\Lambda : \mathbb{R}^{n+m} \rightrightarrows \mathbb{R}^p$ to collect all the *Lagrange multipliers* of problem (1.1); i.e.,

$$\Lambda(x,y) := \left\{u \in \mathbb{R}^p | \nabla_y L(x,y,u) = 0,\, u \geq 0,\, g(x,y) \leq 0,\, u^\top g(x,y) = 0\right\}, \tag{1.5}$$

where $L(x,y,u) := f(x,y) + u^\top g(x,y)$ denotes the *Lagrange function* of problem (1.1). Similarly, λ will be used to represent Λ when it is single-valued.

If the constraint function g is independent from the parameter x and the corresponding optimal solution set-valued map S is single-valued and locally Lipschitz continuous around $\bar{x}$, then we show in Subsection 4.1 that under further appropriate conditions, the generalized Hessian of φ (1.2) can be obtained as

$$\partial^2\varphi(\bar{x})(x^*) \subseteq \nabla^2_{xx} f(\bar{x},\bar{y})x^* + \bar{\partial} s(\bar{x})^\top \nabla^2_{xy} f(\bar{x},\bar{y})x^*, \tag{1.6}$$

where $\bar{y} = s(\bar{x})$ and $\bar{\partial} s(\bar{x})^\top$ stands for the set of transposed Jacobians of s in the sense of Clarke, cf. (2.5). If we further impose the continuous differentiability on s at the point $\bar{x}$, then we can obtain the equality

$$\partial^2\varphi(\bar{x})(x^*) = \nabla^2_{xx} f(\bar{x},\bar{y})x^* + \nabla s(\bar{x})^\top \nabla^2_{xy} f(\bar{x},\bar{y})x^*, \tag{1.7}$$

which coincides with the result in [11, Corollary 3.4.2(c)]. Conditions ensuring that S is single-valued and locally Lipschitz continuous or continuously differentiable are recalled the Section 3. The results above can further be generalized to the case where the optimal solution mapping S (1.3) is multi-valued. Precisely, under appropriate conditions, we also obtain in Subsection 4.1 that

$$\partial^2\varphi(\bar{x}|\underline{x})(\underline{x}^*) \subseteq \bigcup_{y\in S(\bar{x}):\ \underline{x}=\nabla_x f(\bar{x},y)} \Big[\nabla^2_{xx} f(\bar{x},y)\underline{x}^* + D^*S(\bar{x}|y)\Big(\nabla^2_{xy} f(\bar{x},y)\underline{x}^*\Big)\Big]. \tag{1.8}$$

If we drop the assumption that the feasible set of the parametric optimization problem (1.1) is unperturbed, we can obtain the following estimate for the generalized Hessian of φ, provided that S (1.3) and Λ (1.5) are both single-valued and Lipschitz continuous around $\bar{x}$ and $(\bar{x},\bar{y})$, respectively, with $\bar{y}=s(\bar{x})$ and $\bar{u}=\lambda(\bar{x},\bar{y})$:

$$\begin{aligned}\partial^2\varphi(\bar{x}|\underline{x})(\underline{x}^*) \subseteq\ & \nabla^2_{xx}L(\bar{x},\bar{y},\bar{u})\underline{x}^* \\ &+ \bigcup_{(\zeta^*_x,\zeta_y)\in\partial\langle\nabla_x g(\bar{x},\bar{y})\underline{x}^*,\ \lambda\rangle(\bar{x},\bar{y})} \Big[\zeta^*_x + \partial\langle\zeta^*_y + \nabla^2_{xy}L(\bar{x},\bar{y},\bar{u})\underline{x}^*,\ s\rangle(\bar{x})\Big].\end{aligned} \tag{1.9}$$

Here, the symbol ∂ refers to the subdifferential, in the sense of Mordukhovich, cf. (2.4). Obviously, as in the case of (1.7), supposing that λ and s are both single-valued and differentiable functions, we will have

$$\begin{aligned}\partial^2\varphi(\bar{x}|\underline{x})(\underline{x}^*) = \nabla^2_{xx}L(\bar{x},\bar{y},\bar{u})\underline{x}^* + \nabla s(\bar{x})^\top\nabla^2_{xy}L(\bar{x},\bar{y},\bar{u})\underline{x}^* \\ +\Big[\nabla_x\lambda(\bar{x},\bar{y})^\top + \nabla s(\bar{x})^\top\nabla_y\lambda(\bar{x},\bar{y})^\top\Big]\nabla_x g(\bar{x},\bar{y})\underline{x}^*.\end{aligned} \tag{1.10}$$

Observing that $\nabla_x\lambda(\bar{x},\bar{y})^\top + \nabla s(\bar{x})^\top\nabla_y\lambda(\bar{x},\bar{y})^\top$ corresponds to the Jacobian of the function $x\mapsto\lambda(x,s(x))$, it is clear that (1.10) coincides with the formula obtained in [11, Corollary 3.4.1(c)]. More details on how the estimate in (1.9) is obtained are given in Subsection 4.2. For their generalizations to the case where S (1.3) is single-valued and Λ (1.5) set-valued, see Subsection 4.3. As for the case where S is set-valued and Λ is single-valued, see Subsection 4.4. Before moving to that, we first provide some background results, which are useful in their own right, in particular, in the understanding of the subdifferential of φ (1.2) and the *Aubin/Lipschitz-like* and *generalized differentiability* properties of the related mappings S (1.3) and Λ (1.5); cf. Section 3.

2. Notation and mathematical tools needed

We start this section with some notation and basic concepts used throughout the paper. We will use v_j, $j=1,\ldots,n$, to denote the jth component of a vector $v\in\mathbb{R}^n$, while $v^i\in\mathbb{R}^n$, $i=1,\ldots,m$, will represent the ith vector component of a vector of vectors $v\in\prod_{i=1}^m\mathbb{R}^n$. To avoid confusions at some points, we will use $\{0_n\}$ or 0_n for a n-dimensional zero vector. When a distinction is also necessary, we will use I_n for the $n\times n$ identity matrix. For a set-valued mapping $\Psi:\mathbb{R}^n\rightrightarrows\mathbb{R}^m$, we use the *lower case* form, $\psi:\mathbb{R}^n\to\mathbb{R}^m$, to symbolize the mapping when it is single-valued. Most notably, as already mentioned in the previous section, such transitions from *upper* to *lower case* will be used for the optimal solution and Lagrange multipliers set-valued mappings S (1.3) and Λ (1.5), in the forms s and λ, respectively. Additionally, for a set-valued mapping $\Psi:\mathbb{R}^n\rightrightarrows\mathbb{R}^m$, it will be said to be closed if its graph denoted by $\operatorname{gph}\Psi := \{(x,y)\in\mathbb{R}^{n+m}|\ y\in\Psi(x)\}$ is a closed subset of $\mathbb{R}^{n+m}$. Also recall that, as usual, for a set $C\subseteq\mathbb{R}^n$, $\operatorname{co} C$ will be used to denote the convex hull of C.

For a closed subset C of $\mathbb{R}^n$, the *Mordukhovich* (also known as basic or limiting) *normal cone* to C at one of its points $\bar{x}$ is the set

$$N_C(\bar{x}) := \left\{ v \in \mathbb{R}^n \,|\, \exists v_k \to v,\, x_k \to \bar{x}\, (x_k \in C) : v_k \in \widehat{N}_C(x_k) \right\}, \tag{2.1}$$

where $\widehat{N}_C$ denotes the dual of the contingent/Boulingand tangent cone to C. We have the following well-known result, which can be found, for example, in [13, 20].

Theorem 2.1. Let $C := \psi^{-1}(\Xi)$, where $\Xi \subseteq \mathbb{R}^m$ is a closed set and $\psi : \mathbb{R}^n \to \mathbb{R}^m$ a Lipschitz continuous function around $\bar{x}$, then we have

$$N_C(\bar{x}) \subseteq \bigcup \left\{ \partial\langle v, \psi\rangle(\bar{x}) \,\middle|\, v \in N_\Xi(\psi(\bar{x})) \right\}, \tag{2.2}$$

provided the following basic-type qualification condition is satisfied at $\bar{x}$:

$$[0 \in \partial\langle v, \psi\rangle(\bar{x}),\ v \in N_\Xi(\psi(\bar{x}))] \Longrightarrow v = 0. \tag{2.3}$$

Equality holds in (2.2), provided that the set Ξ is normally regular at $\psi(\bar{x})$, i.e., $N_\Xi(\psi(\bar{x})) = \widehat{N}_\Xi(\psi(\bar{x}))$. This is obviously the case if Ξ is a convex set.

In (2.2) and (2.3), the term $\partial\langle v, \psi\rangle(\bar{x})$ refers to the Mordukhovich subdifferential of the function $x \mapsto \sum_{i=1}^m v_i \psi_i(x)$ at $\bar{x}$. If $\psi : \mathbb{R}^n \to \mathbb{R}$, then the *Mordukhovich* (also known as basic or limiting) *subdifferential* of ψ at $\bar{x}$ can be defined by

$$\partial\psi(\bar{x}) := \left\{ \xi \in \mathbb{R}^n \,|\, (\xi, -1) \in N_{\mathrm{epi}\,\psi}(\bar{x}, \psi(\bar{x})) \right\}, \tag{2.4}$$

where epiψ stands for the epigraph of ψ. If ψ is Lipschitz continuous around $\bar{x}$, then we can also define the *Clarke* (or convexified) *subdifferential* of ψ at $\bar{x}$:

$$\bar{\partial}\psi(\bar{x}) := \mathrm{co}\, \partial\psi(\bar{x}). \tag{2.5}$$

In the case where ψ is convex, $\partial\psi(\bar{x})$ and $\bar{\partial}\psi(\bar{x})$ coincide with the subdifferential in the sense of convex analysis.

Using the above concept of basic normal cone, we now introduce the notion of *coderivative* for a given set-valued map $\Psi : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, at some point $(\bar{x}, \bar{y}) \in \mathrm{gph}\,\Psi$, which corresponds to a homogeneous mapping $D^*\Psi(\bar{x}|\bar{y}) : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$, defined by

$$D^*\Psi(\bar{x}|\bar{y})(y^*) := \left\{ x^* \in \mathbb{R}^n | (x^*, -y^*) \in N_{\mathrm{gph}\,\Psi}(\bar{x}, \bar{y}) \right\}, \tag{2.6}$$

for all $y^* \in \mathbb{R}^m$. Here, $N_{\mathrm{gph}\,\Psi}$ represents the basic normal cone (2.1) to gphΨ. The following chain rule from [13, Theorem 5.1] will be pivotal in this work.

Theorem 2.2. Let the set-valued mappings $\Phi : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ and $F : \mathbb{R}^m \rightrightarrows \mathbb{R}^q$ have closed graph. Furthermore, let $\bar{z} \in (F \circ \Phi)(\bar{x})$ and assume that the set-valued map

$$M(x, z) := \Phi(x) \cap F^{-1}(z) = \{ y \in \Phi(x) |\, z \in F(x) \} \tag{2.7}$$

is locally bounded around $(\bar{x}, \bar{z})$ and the qualification condition

$$D^*F(y|\bar{z})(0) \cap \mathrm{Ker}\, D^*F(\bar{x}|y) = \{0\} \ \ \forall y \in \Phi(\bar{x}) \cap F^{-1}(\bar{z}) \tag{2.8}$$

is fulfilled. Then for all $z^* \in \mathbb{R}^q$, we have

$$D^*(F \circ \Phi)(\bar{x}|\bar{z})(z^*) \subseteq \bigcup_{y \in \Phi(\bar{x}) \cap F^{-1}(\bar{z})} D^*\Phi(\bar{x}|y)(z^*) \circ D^*F(y|\bar{z})(z^*).$$

The following result from [9, Proposition 3.3], providing a coderivative estimate for a Cartesian product of finitely many set-valued mappings, will also be useful in the development of the main results of this paper.

Theorem 2.3. Consider the set-valued mappings $\Psi_i : \mathbb{R}^n \rightrightarrows \mathbb{R}^q$ for $i = 1, \ldots, p$, and define a Cartesian product mapping $\Psi : \mathbb{R}^n \rightrightarrows \mathbb{R}^{q \times p}$ by

$$\Psi(x) := \prod_{i=1}^{p} \Psi_i(x) = \Psi_1(x) \times \ldots \times \Psi_p(x).$$

Assume that $\operatorname{gph}\Psi_i$, $i = 1, \ldots, p$, is closed and the qualification condition

$$\left[\sum_{i=1}^{p} v^i = 0,\ v^i \in D^*\Psi_i(\bar{x}|\bar{y}^i)(0),\ i = 1, \ldots, p\right] \Longrightarrow v^1 = \ldots = v^p = 0 \tag{2.9}$$

is satisfied at $(\bar{x}, \bar{y})$ with $\bar{y} := (\bar{y}^i)_{i=1}^p \in \Psi(\bar{x})$. Then, for any $y^* := (y^{*i})_{i=1}^p \in \prod_{i=1}^p \mathbb{R}^q$,

$$D^*\Psi(\bar{x}|\bar{y})(y^*) \subseteq \sum_{i=1}^{p} D^*\Psi_i(\bar{x}|\bar{y}^i)(y^{*i}). \tag{2.10}$$

Equality holds in (2.10), if $\operatorname{gph}\Psi_i$ is normally regular at $(\bar{x}, \bar{y}^i)$, for $i = 1, \ldots, p$.

Next, we provide an estimate of the coderivative of a set-valued mapping defined by the convex hull of another set-valued mapping, which is not necessarily convex-valued. To proceed, consider $\Psi : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ and define $\Phi : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ by

$$\Phi(x) := \operatorname{co} \Psi(x). \tag{2.11}$$

Ψ is assumed to be nonconvex-valued at some points of $\mathbb{R}^n$, as "co" can obviously be dropped at points where the map is convex. An upper estimate of the coderivative of Φ in terms of the coderivative of Ψ can then be obtained as follows. To make the presentation of the result easier, we introduce the set

$$\begin{aligned}\Gamma(\bar{x}, \bar{y}) := \Big\{ (a, b) \in \mathbb{R}^{m+1} \times \textstyle\prod_{s=1}^{m+1} \mathbb{R}^m \Big|\ & a \geq 0,\ \textstyle\sum_{s=1}^{m+1} a_s = 1, \\ & \textstyle\sum_{s=1}^{m+1} a_s b^s = \bar{y},\ b := (b^s)_{s=1}^{n+1} \in \prod_{s=1}^{m+1} \Psi(\bar{x}) \Big\}.\end{aligned} \tag{2.12}$$

Proposition 2.4. Consider $(\bar{x}, \bar{y}) \in \operatorname{gph}\Phi$ and suppose that the set-valued mapping Ψ (2.11) is closed and locally bounded around $\bar{x}$. Furthermore, assume that (2.9), with $\Psi_s := \Psi$ for $s = 1, \ldots, m+1$, holds for all $(a, b) \in \Gamma(\bar{x}, \bar{y})$. Then for all $y^* \in \mathbb{R}^m$, we have

$$D^*\Phi(\bar{x}|\bar{y})(y^*) \subseteq \bigcup_{(a,b) \in \Gamma(\bar{x}, \bar{y})} \left[\sum_{s=1}^{m+1} D^*\Psi\left(\bar{x}|b^s\right)\left(a_s y^*\right)\right]. \tag{2.13}$$

Proof. Start by recalling that as $\Phi(x) \subseteq \mathbb{R}^m$ for all $x \in \mathbb{R}^n$, it follows from the well-known Theorem of Carathéodory that $\Phi(x)$ can be rewritten as

$$\Phi(x) = \left\{ \sum_{s=1}^{m+1} \eta_s v^s \,\middle|\, \eta_s \geq 0,\ v^s \in \Psi(x),\ s = 1, \ldots, m+1, \sum_{s=1}^{m+1} \eta_s = 1 \right\}.$$

Based on this expression, we can easily check that Φ can take the form

$$\Phi(x) = \ell \circ Q(x)$$
$$\text{with } \begin{cases} \ell(a,b) := \sum_{s=1}^{m+1} a_s b^s, \\ Q(x) := \Xi \times \prod_{s=1}^{m+1} \Psi(x), \\ \Xi := \left\{ a \in \mathbb{R}^{m+1} \middle|\, a \geq 0,\ \sum_{s=1}^{m+1} a_s = 1 \right\}. \end{cases}$$

Considering the continuous differentiability of ℓ, the closedness of the set Ξ and the set-valued mapping Ψ, it follows from the chain rule above, cf. Theorem 2.2, that for $\bar{x} \in \mathbb{R}^n$ and $\bar{y} \in \Phi(\bar{x})$, it holds that

$$D^*\Phi(\bar{x}|\bar{y})(y^*) \subseteq \bigcup_{(a,b)\in Q(\bar{x})\cap \ell^{-1}(\bar{y})} \left[D^*Q(\bar{x}|a,b)(\nabla \ell(a,b)^\top y^*) \right] \tag{2.14}$$

for $y^* \in \mathbb{R}^m$, provided that set-valued mapping $M(x,y) := \{(a,b) \in Q(x) | \, \ell(a,b) = y\}$, counterpart of (2.7) is locally bounded around $(\bar{x}, \bar{y})$. Obviously, from the definition of this mapping, $M(x,y) \subseteq \Xi \times \prod_{s=1}^{m+1} \Psi(x)$ for all (x,y). Hence, M is locally bounded around $(\bar{x}, \bar{y})$ given that Ξ is a bounded set and Ψ is assumed to be locally bounded around $\bar{x}$. Now observe that any $b := (b^s)_{s=1}^{n+1} \in \prod_{s=1}^{m+1} \mathbb{R}^m$ is a $m \times (m+1)$ matrix. We rearrange it as a $m^2 + m$-dimensional column vector and proceed with the following notation for the rest of the proof:

$$b := \left[b_1^1 \dots b_m^1 \, \dots \, b_1^{m+1} \dots b_m^{m+1} \right]^\top \quad \text{and} \quad \bar{b} := \begin{bmatrix} b_1^1 & \dots & b_1^{m+1} \\ \vdots & \dots & \vdots \\ b_m^1 & \dots & b_m^{m+1} \end{bmatrix}.$$

Then simple calculations show that $\nabla \ell(a,b) = \left[\bar{b} \ \ a_1 I_m \ \ \dots \ \ a_{m+1} I_m \right]$. Hence,

$$\nabla \ell(a,b)^\top y^* = \left[(\bar{b}^\top y^*)^\top,\ (a_1 y^*)^\top,\ \dots,\ (a_{m+1} y^*)^\top \right]^\top$$

for $y^* \in \mathbb{R}^m$. Considering this formula, the application of Theorem 2.3 to the set-valued mapping Q at the vector $\nabla \ell(a,b)^\top y^*$ leads to the inclusion

$$D^*Q(\bar{x}|a,b)(\nabla \ell(a,b)^\top y^*) \subseteq \sum_{s=1}^{m+1} D^*\Psi\left(\bar{x}|b^s\right)\left(a_s y^*\right), \tag{2.15}$$

given that the coderivative of the constant mapping defined by Ξ is $\{0_n\}$ and condition (2.9) is assumed to hold for all (a,b) such that $(a,b) \in Q(\bar{x})$ and $\ell(a,b) = \bar{y}$. Then combining (2.15) with (2.14), we have the result. □

Remark 2.5. Note that we also have from Theorem 2.3 that equality holds in (2.13), if $\operatorname{gph}\Psi$ is normally regular at $(\bar{x}, b^s)$ for $s = 1, \dots, m+1$. Furthermore, one can easily check that (2.13) is a natural extension of the coderivative of Φ (2.11) when Ψ is single-valued and differentiable. In fact, in the latter case, $D^*\Phi(\bar{x}|\bar{y})(y^*) = \nabla\Psi(\bar{x})^\top y^*$.

To close this section, we introduce the Aubin property (also known as *Lipschitz-like property*) that will used in the next section. A set-valued mapping $\Psi : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is *Lipschitz-like* at $(\bar{x}, \bar{y}) \in \operatorname{gph}\Psi$ if there are neighborhoods U of $\bar{x}$, V of $\bar{y}$, and a constant $\kappa > 0$ such that $d(y, \Psi(x)) \leq \kappa \|x - u\|$ for all $x, u \in U$ and $y \in \Psi(u) \cap V$, where d stands for the usual distance function. A closed set-valued mapping Ψ is Lipschitz-like around $(\bar{x},\bar{y})$ if and only if the condition

$$D^*\Psi(\bar{x}|\bar{y})(0) = \{0\}, \tag{2.16}$$

known as the *coderivative/Mordukhovich criterion*, is satisfied at $(\bar{x},\bar{y})$; cf. [14, Theorem 5.7] and [20, Theorem 9.40].

3. On the subdifferential of the optimal value function

To start this subsection, we recall that the fundamental goal of this paper is to develop *generalized Hessians* (also known as *second order subdifferentials*) of the optimal value function φ (1.2). Hence, it would be natural to first clarify the expressions of the *subdifferentials* or *first order subdifferentials*, to be precise, of this function. These quantities and further properties have been extensively studied in the literature; see, e.g., [3, 4, 11, 12, 15, 16, 23] and references therein. Below, we recall the relevant aspects of these properties while adding some crucial aspects based on the *concave-convexity*, that we define below. Before, note that from here on, the feasible set of the parametric problem (1.1) will be defined by the following set-valued mapping:

$$K(x) := \{y \in \mathbb{R}^m |\ g(x,y) \leq 0\}. \tag{3.1}$$

A function ψ defined from $\mathbb{R}^{n+m}$ to $\mathbb{R}$ by $(x,y) \mapsto \psi(x,y)$ will be said to be *concave-convex* if the function $\psi(.,y)$ is concave for all $y \in \mathbb{R}^m$, while the function $\psi(x,.)$ is convex for all $x \in \mathbb{R}^n$. Subsequently, problem (1.1) will be said to be convex-concave if the functions f and g_i, $i = 1,\ldots,p$, are concave-convex. Similarly, problem (1.1) will just be said to be convex if the latter functions are convex w.r.t. y. We will also use the *Mangasarian-Fromowitz constraint qualification* (*MFCQ*)

$$\left.\begin{array}{r} \nabla_y g(\bar{x},\bar{y})^\top u = 0 \\ u \geq 0,\ g(\bar{x},\bar{y}) \leq 0,\ u^\top g(\bar{x},\bar{y}) = 0 \end{array}\right\} \Longrightarrow u = 0 \tag{3.2}$$

and the *linear independence constraint qualification* (*LICQ*)

$$\sum_{i \in I(\bar{x},\bar{y})} u_i \nabla_y g_i(\bar{x},\bar{y}) = 0 \implies u_i = 0,\ i \in I(\bar{x},\bar{y}), \tag{3.3}$$

where $I(\bar{x},\bar{y}) := \{i = 1,\ldots,p |\ g_i(\bar{x},\bar{y}) = 0\}$. It is well-known that if the LICQ holds at $(\bar{x},\bar{y})$, then the MFCQ will automatically hold at the same point.

Theorem 3.1. Considering the optimal value function (1.2), it holds that:

(i) If Y defines a compact set such that $K(x) := Y$ for all $x \in \mathbb{R}^n$, then for all $x \in \mathbb{R}^n$,

$$\bar{\partial}\varphi(x) = \text{co}\{\nabla_x f(x,y) |\ y \in S(x)\}. \tag{3.4}$$

If additionally, Y is convex and f concave-convex, then for all $x \in \mathbb{R}^n$,

$$\bar{\partial}\varphi(x) = \{\nabla_x f(x,y) |\ y \in S(x)\}. \tag{3.5}$$

(ii) Let problem (1.1) be convex and S (1.3) single-valued (i.e., $S := s$) around $\bar{x}$. Furthermore, let gph K be compact and the MFCQ hold at $(\bar{x}, \bar{y})$ with $\bar{y} = s(\bar{x})$. Then for all x near $\bar{x}$, with $s(x) := y$, it holds that

$$\bar{\partial}\varphi(x) = \bigcup_{u \in \Lambda(x,y)} \left\{\nabla_x f(x,y) + \nabla_x g(x,y)^\top u\right\}. \tag{3.6}$$

(iii) If gph K is compact and the LICQ holds at $(\bar{x}, y)$, for all $y \in S(\bar{x})$, then near $\bar{x}$,

$$\bar{\partial}\varphi(x) = \text{ co} \bigcup_{y \in S(x)} \left\{\nabla_x f(x,y) + \nabla_x g(x,y)^\top u\right\} \tag{3.7}$$

with $u = \lambda(x,y)$. If additionally, problem (1.1) is concave-convex, then for all x near $\bar{x}$, with $u = \lambda(x,y)$ for $y \in S(x)$,

$$\bar{\partial}\varphi(x) = \bigcup_{y \in S(x)} \left\{\nabla_x f(x,y) + \nabla_x g(x,y)^\top u\right\}. \tag{3.8}$$

Proof. (i) Equality (3.4) is a well-known result by Danskin [4]. As for (3.5), the maximization case proven in [2] can easily be adapted to our minimization case in (1.2).

(ii) As the MFCQ holds at $(\bar{x}, y)$ with $\bar{y} = s(\bar{x})$, and remains persistent near some neighborhood of this point, then it is well-known that the formula (3.6) will hold in some neighborhood of $\bar{x}$, given that gph K is compact; see, e.g., [23].

(iii) Start by noting that as in the previous case, the LICQ being persistent near $(\bar{x}, y)$ for all $y \in S(\bar{x})$, where it holds, then we have (3.7) from Gauvin and Dubeau [12] given that gph K is compact. To prove (3.8), denote by $\Phi(x)$ the right-hand-side of the mapping in this formula. Then take $x^{*1}, x^{*2} \in \Phi(x)$, for a given value of $x \in \mathbb{R}^n$. As the LICQ holds at (x,y) for all $y \in S(x)$, it follows that for some $y^1, y^2 \in S(x)$, we can find some $u^1 = \lambda(x, y^1)$ and $u^2 = \lambda(x, y^2)$ such that we have $x^{*1} = \nabla_x L(x, y^1, u^1)$ and $x^{*2} = \nabla_x L(x, y^2, u^2)$. For all $\lambda \in (0,1)$, let $x^* := \lambda x^{*1} + (1-\lambda)x^{*2}$ and note that we have $y = \lambda y^1 + (1-\lambda)y^2 \in S(x)$, since problem (1.1) is convex. For the same reason, the Lagrangian function L is also convex w.r.t. y, for the fixed values u^1 and u^2, given that they are non-negative by definition. Hence, for a vector $h \in \mathbb{R}^n$ such that $(x+h, y^1, u^1) \in \text{gph}\,\Lambda$ and $(x+h, y^2, u^2) \in \text{gph}\,\Lambda$ [note that $\Lambda(x+h, y^1)$ and $\Lambda(x+h, y^2)$ might not necessarily be single-valued as it is that case for $\Lambda(x, y^1)$ and $\Lambda(x, y^2)$],

$$L(x+h, y, u) \le \lambda L(x+h, y^1, u) + (1-\lambda) L(x+h, y^2, u), \tag{3.9}$$

where $u = \lambda u^1 + (1-\lambda)u^2$. Since the functions f and g_i, $i = 1\ldots, q$, are concave w.r.t. x, then L is also concave w.r.t. x, for the fixed values u^1 and u^2, as they are non-negative by definition. Then since the functions f and g_i, $i = 1\ldots, q$, are all differentiable w.r.t. x, the same holds for L and it follows that we have

$$L(x+h, y^i, u) \le L(x, y^i, u) + \left\langle h, \nabla_x L(x, y^i, u)\right\rangle, \; i = 1, 2. \tag{3.10}$$

To proceed with the next step, recall that under the assumption that problem (1.1) is convex, it follows that for all $b \in S(b)$, the point (b, c) with $c \in \Lambda(a, b)$ is a saddle point of the Lagrangian function $L(a,.,.)$. Subsequently, it holds that

$$\begin{aligned}
L(x+h,y,u) &\leq \lambda L(x+h,y^1,u) + (1-\lambda)L(x+h,y^2,u), && (3.11)\\
&\leq \lambda L(x+h,y^1,u^1) + (1-\lambda)L(x+h,y^2,u^2), && (3.12)\\
&\leq \lambda\,[L(x,y^1,u^1) + \langle h, \nabla_x L(x,y^1,u^1)\rangle] \\
&\quad + (1-\lambda)\,[L(x,y^2,u^2) + \langle h, \nabla_x L(x,y^2,u^2)\rangle], && (3.13)\\
&= \lambda L(x,y^1,u^1) + (1-\lambda)L(x,y^2,u^2) + \langle h, x^*\rangle, \\
&\leq \lambda L(x,y,u^1) + (1-\lambda)L(x,y,u^2) + \langle h, x^*\rangle, && (3.14)\\
&= L(x,y,u) + \langle h, x^*\rangle, && (3.15)
\end{aligned}$$

where (3.11) corresponds to (3.9) and (3.12) is due to (y^1,u) and (y^2,u) being saddle points for $L(x+h,.,.)$, since $u^1 \in \Lambda(x+h,y^1)$ and $u^2 \in \Lambda(x+h,y^2)$. As for (3.13), it is follows from (3.10) while (3.14) is obtained from the fact that (y^1,u^1) and (y^2,u^2) are saddle points for $L(x,.,.)$, as $u^1 = \lambda(x,y^1)$ and $u^2 = \lambda(x,y^2)$; and equality (3.15) is due to the expression $u = \lambda u^1 + (1-\lambda)u^2$. Finally, as $L(.,y,u)$ is a concave function, x^* is one of its supergradients. Furthermore, as this function is also differentiable w.r.t. x, then it has only a single supergradient, which is $x^* = \nabla_x L(x,y,u)$. Considering the fact that $y \in S(x)$ and $u = \lambda(x,y)$, it follows that $x^* \in \Phi(x)$. □

It is important to recall that the compactness assumption on gph K can be relaxed by instead imposing some set-valued-type continuity properties on S (1.3); see, e.g., [3, 12, 15, 16, 23]. But for the purpose of simplifying the framework used in this paper, we do consider such relaxations here. However, most results in this paper will remain valid under such assumptions.

It is clear from Theorem 3.1 that the subdifferential of φ is a "function" of the Lagrange multipliers and optimal solution set-valued mappings. It is therefore natural to imagine that second order subdifferentials for φ will primarily depend on the generalized differentiation tools for this mappings. Hence, to get well prepared for our main results in the next section, we first provide some useful properties of these mappings here. We start with a coderivative estimate for Λ and deduce a condition ensuring that this mapping is *Lipschitz-like*; meaning that the *Aubin property* holds. From here on, we will also assume that the graph of the set-valued mapping Λ (1.5) is nonempty. Given that $S(x) \neq \emptyset$ for all $x \in \mathbb{R}^n$, the latter is automatically satisfied if there exist a point $(x,y) \in \text{gph}\, S$, where a constraint qualification, e.g., the MFCQ or LICQ, holds. Also recall that for a point $(\bar{x},\bar{y},\bar{u}) \in \text{gph}\,\Lambda$, we can define the partition

$$\begin{array}{lllll}
\eta & := & \eta(\bar{x},\bar{y},\bar{u}) & := & \{i=1,\ldots,p \,|\, \bar{u}_i = 0,\ g_i(\bar{x},\bar{y}) < 0\},\\
\theta & := & \theta(\bar{x},\bar{y},\bar{u}) & := & \{i=1,\ldots,p \,|\, \bar{u}_i = 0,\ g_i(\bar{x},\bar{y}) = 0\},\\
\nu & := & \nu(\bar{x},\bar{y},\bar{u}) & := & \{i=1,\ldots,p \,|\, \bar{u}_i > 0,\ g_i(\bar{x},\bar{y}) = 0\},
\end{array} \qquad (3.16)$$

of the indices of the constraints of the feasible set of problem (1.1). This allows us to introduce the following special class of multipliers, which permits an elegant presentation of the remaining results of this section:

$$\mho(\bar{x},\bar{y},\bar{u},u^*) := \left\{ (a,c) \,\middle|\, \begin{array}{r} u^*_\nu + \nabla_y g_\nu(x,y)a = 0,\ c_\eta = 0 \\ \forall i \in \theta : \left(u^*_i + \nabla_y g_i(x,y)a > 0 \wedge c_i > 0\right) \\ \vee c_i\left(u^*_i + \nabla_y g_i(x,y)a\right) = 0 \end{array} \right\} \tag{3.17}$$

with $(\bar{x},\bar{y},\bar{u}) \in \text{gph}\,\Lambda$ and $u^* \in \mathbb{R}^p$.

Proposition 3.2. Consider a point $(\bar{x},\bar{y},\bar{u}) \in \text{gph}\,\Lambda$ and suppose that we have

$$\left.\begin{array}{r} \nabla^2_{yx}L(\bar{x},\bar{y},\bar{u})a + \nabla_x g(\bar{x},\bar{y})^\top c = 0 \\ \nabla^2_{yy}L(\bar{x},\bar{y},\bar{u})a + \nabla_y g(\bar{x},\bar{y})^\top c = 0 \\ (a,c) \in \mho(\bar{x},\bar{y},\bar{u},0) \end{array}\right\} \Longrightarrow \left\{ \begin{array}{l} a = 0, \\ c = 0. \end{array}\right. \tag{3.18}$$

Then for all $u^* \in \mathbb{R}^p$, we have the following upper estimate:

$$D^*\Lambda(\bar{x},\bar{y}|\bar{u})(u^*) \subseteq \left\{ \left[\begin{array}{l} \nabla^2_{yx}L(\bar{x},\bar{y},\bar{u})a + \nabla_x g(\bar{x},\bar{y})^\top c \\ \nabla^2_{yy}L(\bar{x},\bar{y},\bar{u})a + \nabla_y g(\bar{x},\bar{y})^\top c \end{array}\right] \middle| (a,c) \in \mho(\bar{x},\bar{y},\bar{u},u^*) \right\}. \tag{3.19}$$

Furthermore, Λ is Lipschitz-like around $(\bar{x},\bar{y},\bar{u})$, provided we also have

$$(a,c) \in \mho(\bar{x},\bar{y},\bar{u},0) \Longrightarrow \left\{ \begin{array}{l} \nabla^2_{yx}L(\bar{x},\bar{y},\bar{u})a + \nabla_x g(\bar{x},\bar{y})^\top c = 0, \\ \nabla^2_{yy}L(\bar{x},\bar{y},\bar{u})a + \nabla_y g(\bar{x},\bar{y})^\top c = 0. \end{array}\right. \tag{3.20}$$

Proof. Observe that the set-valued mapping Λ can be written as

$$\begin{array}{l} \Lambda(x,y) = \{u \in \mathbb{R}^p |\ \psi(x,y,u) \in \Xi\} \\ \text{with} \left\{ \begin{array}{l} \psi(x,y,u) := \left[\nabla_y L(x,y,u)^\top, u^\top, -g(x,y)^\top\right], \\ \Xi := \{0_m\} \times \Theta, \\ \Theta := \{(a,b) \in \mathbb{R}^{2p} |\ a \geq 0,\ b \geq 0,\ a^\top b = 0\}. \end{array}\right. \end{array} \tag{3.21}$$

Let $(x^*,y^*) \in D^*\Lambda(\bar{x},\bar{y}|\bar{u})(u^*)$. Then, by the definition of the concept of coderivative (2.6), $(x^*,y^*,-u^*) \in N_{\text{gph}\,\Lambda}(\bar{x},\bar{y},\bar{u})$. Hence, it follows from Theorem 2.1 that there exists a vector $(a,b,c) \in N_\Xi(\psi(\bar{x},\bar{y},\bar{u}))$ such that

$$\left[\begin{array}{r} x^* \\ y^* \\ -u^* \end{array}\right] = \nabla\psi(\bar{x},\bar{y},\bar{u})^\top \left[\begin{array}{c} a \\ b \\ c \end{array}\right] = \left[\begin{array}{l} \nabla^2_{yx}L(\bar{x},\bar{y},\bar{u})a - \nabla_x g(\bar{x},\bar{y})^\top c \\ \nabla^2_{yy}L(\bar{x},\bar{y},\bar{u})a - \nabla_y g(\bar{x},\bar{y})^\top c \\ \nabla_y g(\bar{x},\bar{y})a + b \end{array}\right], \tag{3.22}$$

provided that the counterpart of (2.3) holds at $(\bar{x},\bar{y},\bar{u})$. For the latter requirement and the finalization of the estimate in (3.23), note that

$$N_\Xi(\psi(\bar{x},\bar{y},\bar{u})) = \mathbb{R}^m \times \left\{ (u^*,v^*) \in \mathbb{R}^{2p} : \begin{array}{ll} u^*_i = 0 & \forall i \in \nu \\ v^*_i = 0 & \forall i \in \eta \\ (u^*_i < 0 \wedge v^*_i < 0) \vee u^*_i v^*_i = 0 & \forall i \in \theta \end{array} \right\}.$$

Combining this equality with (3.22), one can easily check that the counterpart of (2.3) is satisfied under assumption (3.18). As for the Lipschitz-likeness of Λ around $(\bar{x},\bar{y})$, observe from the discussion above that

$$D^*\Lambda(\bar{x},\bar{y}|\bar{u})(0) \subseteq \left\{ \left[\begin{array}{c} \nabla^2_{yx}L(\bar{x},\bar{y},\bar{u})a + \nabla_x g(\bar{x},\bar{y})^\top c \\ \nabla^2_{yy}L(\bar{x},\bar{y},\bar{u})a + \nabla_y g(\bar{x},\bar{y})^\top c \end{array} \right] \middle| (a,c) \in \mho(\bar{x},\bar{y},\bar{u},0) \right\}. \tag{3.23}$$

Hence, under (3.20), $D^*\Lambda((\bar{x},\bar{y})|\bar{u})(0) = \{0\}$. This ensures that Λ is Lipschitz-like around $(\bar{x},\bar{y},\bar{u})$, based on the Mordukhovich coderivative criterion (2.16). □

Remark 3.3. The conclusions of Proposition 3.2 remain valid if the multipliers set $\mho(\bar{x},\bar{y},\bar{u},u^*)$ in (3.17) is replaced by the following one:

$$\overline{\mho}(\bar{x},\bar{y},\bar{u},u^*) := \left\{ (a,c) \middle| \begin{array}{c} u^*_\nu + \nabla_y g_\nu(x,y)a = 0,\ c_\eta = 0 \\ \forall i \in \theta:\ c_i\left(u^*_i + \nabla_y g_i(x,y)a\right) \geq 0 \end{array} \right\}. \tag{3.24}$$

The implications in (3.18) and (3.20) correspond to M(or *Mordukhovich*)-type conditions while (3.23) can be labeled as M-type estimate of the coderivative of Λ. Similarly, with (3.24), we will respectively have C(or *Clarke*)-type conditions and a C-type estimate for the coderivative of Λ. More details on constructions and vocabulary in this vein can be found, for example, in [7, 8].

Next, we provide a simple, yet powerful relationship between the coderivatives of S and Λ, that will allow the derivation of a complete estimate of the former based on Proposition 3.2.

Proposition 3.4. Suppose that the functions $f(x,.)$ and $g_i(x,.)$, $i = 1,\ldots,p$, are convex, for all $x \in \mathbb{R}^n$, and the MFCQ holds at $(\bar{x},\bar{y}) \in \text{gph}\, S$. Then for all $y^* \in \mathbb{R}^m$,

$$D^*S(\bar{x}|\bar{y})(y^*) \subseteq \bigcup_{u\in\Lambda(\bar{x},\bar{y})} \{x^* \in \mathbb{R}^n |\ (x^*,-y^*) \in D^*\Lambda(\bar{x},\bar{y}|u)(0)\}. \tag{3.25}$$

If in addition, the set-valued mapping S is closed and Λ is Lipschitz-like around $(\bar{x},\bar{y},u)$, for all $u \in \Lambda(\bar{x},\bar{y})$, then S is Lipschitz-like around $(\bar{x},\bar{y})$.

Proof. Under the assumption that the functions $f(x,.)$ and $g_i(x,.)$, $i = 1,\ldots,p$, are convex, for all $x \in \mathbb{R}^n$ and the MFCQ holds at $(\bar{x},\bar{y}) \in \text{gph}\, S$, it follows that near this point, the optimal solution set-valued mapping can take the form

$$S(x) = \Pi_1 \circ Q(x) \quad \text{with} \left\{ \begin{array}{l} \Pi_1(y,u) := y, \\ Q(x) := \{(y,u) |\ u \in \Lambda(x,y)\}. \end{array} \right.$$

Observe from the definition of the set-valued map Q and the function Π_1 that we have $\nabla\Pi_1(y,u)^\top y^* = \left[(y^*)^\top, 0_p^\top\right]^\top$ and $(y,u) \in Q(\bar{x}) \cap \Pi_1^{-1}(\bar{y})$ if and only if $u \in \Lambda(\bar{x},\bar{y})$. The set-valued mapping Λ is closed, given that the functions f and g are assumed to be continuously differentiable throughout the paper. Then applying the chain rule from Theorem 2.2 to the above expression of S,

$$D^*S(\bar{x}|\bar{y})(y^*) \subseteq \bigcup_{u\in\Lambda(\bar{x},\bar{y})} D^*Q(\bar{x}|\bar{y},u)(y^*,0), \tag{3.26}$$

provided that the set-valued mapping $M(x,y) := \{(z,u)|\,(z,u)\in Q(x),\ \Pi_1(z,u)=y\}$ is locally bounded around $(\bar{x},\bar{y})$. To show that this property actually holds, consider a sequence $(x^k,y^k,z^k,u^k)\in \text{gph}\ M$ such that $x^k\to\bar{x}$ and $y^k\to\bar{y}$. Then, $z^k=y^k$ and

$$\nabla_y L(x^k,y^k,u^k)=0,\ u^k\geq 0,\ g(x^k,y^k)\leq 0,\ (u^k)^\top g(x^k,y^k)=0. \tag{3.27}$$

Suppose that $u^k\geq k$ for all k. Then, without loss of generality, we can find a subsequence of $\{u^k\}$ with the same notation, provided there is no confusion such that $u^k/\|u^k\|\to\bar{u}$ and $\|\bar{u}\|=1$. Hence, dividing the corresponding terms in the system (3.27) containing u^k by $\|u^k\|$ and tending k to infinity, we arrive at

$$\nabla_y g(\bar{x},\bar{y})^\top\bar{u}=0,\ \bar{u}\geq 0,\ g(\bar{x},\bar{y})\leq 0,\ \bar{u}^\top g(\bar{x},\bar{y})=0. \tag{3.28}$$

Since the MFCQ is satisfied at $(\bar{x},\bar{y})$, we must have $\bar{u}=0$. Hence, contradicting the assumption that the sequence $\{u^k\}$ is divergent. This confirms that the set-valued mapping M above is locally bounded around $(\bar{x},\bar{y})$. On the other hand, we have

$$\begin{aligned} D^*Q(\bar{x}|\bar{y},u)(y^*,u^*) &= \left\{x^*|\,(x^*,-y^*,-u^*)\in N_{\text{gph}\ Q}(\bar{x},\bar{y},u)\right\}\\ &= \left\{x^*|\,(x^*,-y^*,-u^*)\in N_{\text{gph}\ \Lambda}(\bar{x},\bar{y},u)\right\}\\ &= \{x^*|\,(x^*,-y^*)\in D^*\Lambda(\bar{x},\bar{y}|u)(u^*)\}\end{aligned} \tag{3.29}$$

considering the fact that the graph of Q is the same as that of Λ. Combining the last equality in (3.29) with inclusion (3.26), we have (3.25). As for the Lipschitz-like property of S at $(\bar{x},\bar{y})$, this is based on inclusion (3.25), while applying the coderivative criterion (2.16), given that S is a closed set-valued mapping. □

Corollary 3.5. Suppose that the functions $f(x,.)$ and $g_i(x,.)$, $i=1,\ldots,p$, are convex, for all $x\in\mathbb{R}^n$, the MFCQ holds at $(\bar{x},\bar{y})\in\text{gph}\ S$, and the qualification condition (3.18) is satisfied at $(\bar{x},\bar{y},u)$ for all $u\in\Lambda(\bar{x},\bar{y})$. Then for all $y^*\in\mathbb{R}^m$,

$$\begin{aligned} D^*S(\bar{x}|\bar{y})(y^*) \subseteq \bigcup_{u\in\Lambda(\bar{x},\bar{y})}\ \bigcup_{(a,c)\in\mho(\bar{x},\bar{y},u,0)} &\Big\{\nabla^2_{xy}L(\bar{x},\bar{y},u)^\top a+\nabla_x g(\bar{x},\bar{y})^\top c\Big|\\ &y^*+\nabla^2_{yy}L(\bar{x},\bar{y},u)a+\nabla_y g(\bar{x},\bar{y})^\top c=0\Big\}.\end{aligned}$$

If in addition, the qualification condition (3.20) holds at $(\bar{x},\bar{y},u)$, for all $u\in\Lambda(\bar{x},\bar{y})$, then S is Lipschitz-like around $(\bar{x},\bar{y})$.

Proof. Obviously follows from an application of Propositions 3.4 and 3.2. □

This estimate of the coderivative of S was obtained in [17, Theorem 4.3] using a different approach. Also note that if the LICQ, the strict complementarity condition (SCC) and the strong second order sufficient condition (SSOSC) are satisfied, then the optimal solution set-valued mapping S (1.3) and the Lagrange multiplier mapping u

(as a function of just x) are locally unique and based on the implicit function theorem, their derivatives can be obtained from the system

$$\begin{bmatrix} \nabla^2_{yy}L(\bar{x},\bar{y},\bar{u}) & \nabla_y g_1(\bar{x},\bar{y})^\top & \dots & \nabla_y g_p(\bar{x},\bar{y})^\top \\ \bar{u}_1\nabla_y g_1(\bar{x},\bar{y}) & g_1(\bar{x},\bar{y}) & \dots & O \\ \vdots & \vdots & \ddots & \\ \bar{u}_p\nabla_y g_p(\bar{x},\bar{y}) & O & \dots & g_p(\bar{x},\bar{y}) \end{bmatrix} \begin{bmatrix} \nabla s(\bar{x}) \\ \nabla u(\bar{x}) \end{bmatrix} = \begin{bmatrix} \nabla^2_{xx}L(\bar{x},\bar{y},\bar{u}) \\ \bar{u}_1\nabla_x g_1(\bar{x},\bar{y}) \\ \vdots \\ \bar{u}_p\nabla_x g_p(\bar{x},\bar{y}) \end{bmatrix} \tag{3.30}$$

cf. [11, Chapter 3]. Under additional invertibility assumptions, complete expressions of $\nabla s(\bar{x})$ can be written in terms of the problem data; see, e.g., [23, Chapter 7]. If the SCC is dropped from the statement above, the best we can generally get is the locally Lipschtz continuity of the optimal solution function s, which however remains locally single-valued; see, e.g., [18], where the LICQ is further relaxed to assume that the MFCQ holds together with the constant rank constraint qualification (CRCQ). Under the latter class of assumptions, s is in fact a piecewise continuously differentiable (PC^1) function [6, Chapter 4]. Hence, its Clarke subdifferential can be obtained as

$$\bar{\partial}s(\bar{x}) = \text{co}\ \left\{s^i(\bar{x})|\ i\ \text{ s.t. }\ \bar{x}\in \text{cl int Supp}\ \left(s,s^i\right)\right\}, \tag{3.31}$$

where *cl int* denotes the closure of the interior of Supp $\left(s,s^i\right) := \left\{x|\ s(x) = s^i(x)\right\}$, cf. [6, Chapter 4]. A more detailed expression of (3.31) in terms of the corresponding problem data can be found in the latter reference. Further details on PC^1 functions can also be found in [22].

As for the mapping Λ, note that the formula in (3.30) leads to the Lagrange multiplier only as a function of x. We are not aware of any publication deriving differentiation properties for this mapping as a "function" of both x and y. It is well-known that Λ will be single-valued at any point $(\bar{x},\bar{y})$ where the LICQ is satisfied. However we limit our scope of analysis for this mapping to its coderivative and Lipschitz-likeness, as discussed in Proposition 3.2. Locally unique, possibly differentiable or locally Lipschitz continuity counterparts of Λ (i.e., λ) will be studied in a future work. For the results of Subsections 4.2 and 4.4, we assume that these properties can hold.

4. Generalized Hessian estimates for the optimal value function

4.1. Case where the feasible set is unperturbed.

In this subsection, we assume that the feasible set of problem (1.2) is independent of x. More precisely, for $g:\mathbb{R}^m\to\mathbb{R}^p$, our attention here will be on the following function assumed to be finite-valued:

$$\varphi(x) := \min_{y\in Y} f(x,y)\ \text{ with }\ Y := \{y\in\mathbb{R}^m|\ g(y)\leq 0\}. \tag{4.1}$$

Theorem 4.1. Suppose that Y is a convex and compact set and let the function f be concave-convex. Then for $\underline{x}\in\partial\varphi(\bar{x})$ and $\underline{x}^*\in\mathbb{R}^n$, we have (1.8).

Proof. Note that under the assumptions of the theorem, we have equality (3.5) from Theorem 3.1. This equality can equivalently be written as $\partial\varphi(x) = \nabla_x f\circ\Psi(x)$, where $\Psi(x) := \{x\}\times S(x)$. Further observe that the set-valued map Ψ is closed, given that the counterpart of S (1.3) for (4.1) can take the form $S(x) := \{y\in Y|\ f(x,y) = \varphi(x)\}$, and is thus closed as φ is locally Lipschitz continuous under the imposed continuous differentiability of the function f and compactness of the set Y. Also note that the set-valued mapping $M(x,z) := \{(a,b) | a = x,\ b \in S(x),\ \nabla_x f(a,b) = z\}$ is locally bounded around $(\bar{x},\underline{x})$, given that $M(X \times Z) \subseteq X \times Y$, for some neighborhoods X and Z of $\bar{x}$ and $\underline{x}$, respectively, with assumed X to be bounded. Hence, we have

$$\partial^2\varphi(\bar{x}|\underline{x})(\underline{x}^*) \subseteq \bigcup_{y\in S(\bar{x}):\ \underline{x}=\nabla_x f(\bar{x},y)} \left[D^*\Psi(\bar{x}|\bar{x},y)\left(\left[\begin{array}{c} \nabla^2_{xx} f(\bar{x},y)\underline{x}^* \\ \nabla^2_{yx} f(\bar{x},y)^\top \underline{x}^* \end{array}\right]\right)\right] \tag{4.2}$$

from the chain rule in Theorem 2.2. Finally, for the right-hand-side of (4.2), it follows that since S is closed as shown above, applying Theorem 2.3 to Ψ leads to

$$D^*\Psi(\bar{x}|\bar{x},y)\left(\left[\begin{array}{c} \nabla^2_{xx} f(\bar{x},y)\underline{x}^* \\ \nabla^2_{yx} f(\bar{x},y)^\top \underline{x}^* \end{array}\right]\right) \subseteq \nabla^2_{xx} f(\bar{x},y)\underline{x}^* + D^*S(\bar{x}|y)\left(\nabla^2_{yx} f(\bar{x},y)^\top \underline{x}^*\right) \tag{4.3}$$

given that the corresponding counterpart of qualification condition (2.9) is automatically satisfied, as $0 \in D^*S(\bar{x}|\bar{y})(0)$ by the positive homogeneity of the coderivative mapping. The proof is then completed by combining (4.2) and (4.3). □

Corollary 4.2. Let the assumptions of Theorem 4.1 hold. Furthermore, suppose that for all $y \in S(\bar{x})$ such that $\nabla_x f(\bar{x},y) = \underline{x}$, the MFCQ holds at y and the matrix $\nabla^2_{yx} f(\bar{x},y)$ is full rank. Then for $\underline{x} \in \partial\varphi(\bar{x})$ and $\underline{x}^* \in \mathbb{R}^n$, it holds that

$$\begin{aligned}\partial^2\varphi(\bar{x}|\underline{x})(\underline{x}^*) \subseteq \bigcup_{y\in S(\bar{x}):\ \underline{x}=\nabla_x f(\bar{x},y)} \ \bigcup_{u\in\Lambda(\bar{x},y)} \ \bigcup_{(a,c)\in\mho(\bar{x},y,u,0)} \Big\{ & \nabla^2_{xx} f(\bar{x},y)\underline{x}^* + \nabla^2_{yx} f(\bar{x},y)^\top a \Big| \\ & \nabla^2_{xy} f(\bar{x},y)\underline{x}^* + \nabla^2_{yy} L(\bar{x},y,u)a + \nabla g(y)^\top c = 0 \Big\}.\end{aligned}$$

Next, we drop the concave-convex assumption made in Theorem 4.1. Then we have the following result, where

$$\Gamma^\circ(\bar{x},\underline{x}) := \Big\{ (a,z) \in \mathbb{R}^{n+1} \times \textstyle\prod_{s=1}^{n+1} \mathbb{R}^m \Big| \quad a \geq 0,\ \sum_{s=1}^{m+1} a_s = 1, \\ \textstyle\sum_{s=1}^{m+1} a_s \nabla_x f(\bar{x},z^s) = \underline{x},\ z \in \prod_{s=1}^{n+1} S(\bar{x}) \Big\}$$

and

$$\Delta^\circ(\bar{x},z^s) := \{ y \in \mathbb{R}^m |\ y \in S(\bar{x}):\ \nabla_x f(\bar{x},y) = \nabla_x f(\bar{x},z^s) \}.$$

Theorem 4.3. Suppose that Y is a compact set and consider a point $(\bar{x},\underline{x})$ such that $\underline{x} \in \partial\varphi(\bar{x})$ and the implication

$$\left[\textstyle\sum_{s=1}^{m+1} v^s = 0,\ v^s \in \displaystyle\bigcup_{y\in\Delta^\circ(\bar{x},z^s)} D^*S(\bar{x}|y)(0),\ s = 1,\ldots,n+1 \right] \\ \Longrightarrow v^1 = \ldots = v^{n+1} = 0 \tag{4.4}$$

is satisfied at all $(a,z) \in \Gamma^\circ(\bar{x},\underline{x})$. Then, for all $\underline{x}^* \in \mathbb{R}^n$, it holds that

$$\partial^2\varphi(\bar{x}|\underline{x})(\underline{x}^*) \subseteq \\ \bigcup_{(a,z)\in\Gamma^\circ(\bar{x}|\underline{x})} \left\{ \sum_{s=1}^{n+1} \left[\bigcup_{y\in\Delta^\circ(\bar{x},z^s)} \left[a_s \nabla^2_{xx} f(\bar{x},y)\underline{x}^* + D^*S(\bar{x}|y)\left(a_s \nabla^2_{xy} f(\bar{x},y)\underline{x}^*\right)\right]\right]\right\}.$$

Proof. Let us first recall that under the compactness of Y and the continuously differentiability of f, we have from (3.4) that $\bar{\partial}\varphi(x) = \text{co } \nabla_x f \circ \Psi(x)$ with $\Psi(x) := \{x\} \times S(x)$. Next, recall that Ψ is closed, following the discussion in the proof of Theorem 4.1. Furthermore, one can easily check that Ψ is locally bounded around any point in $\mathbb{R}^n$. Now, consider a sequence $\{(a^k, c^k)\}$ with $c^k \in \nabla_x f \circ \Psi(a^k)$ such that $a^k \to \bar{a}$ and $c^k \to \bar{c}$. Obviously, we can find a sequence $\{b^k\}$ with $b^k \in \Psi(a^k)$ such that $\nabla_x f(b^k) = c^k$ for all k. By the local boundedness of Ψ, $\{b^k\}$ admits a convergent subsequence that we denote similarly, provided there is no confusion. By the closedness of Ψ, we have $b^k \to \bar{b} \in \Psi(\bar{a})$, for some $\bar{b} \in \mathbb{R}^{n+m}$. Additionally, note that $\nabla_x f(\bar{b}) = \bar{c}$, as f is assumed to be continuously differentiable throughout this paper. Thus $\bar{c} \in \nabla_x f \circ \Psi(\bar{a})$; confirming that $\nabla_x f \circ \Psi$ is a closed set-valued mapping. Next, we consider a sequence $x^k \to \bar{x}$ and any sequence $z^k \in \nabla_x f \circ \Psi(x^k)$. Then, we can find a sequence $\{y^k\}$ such that $y^k \in S(x^k)$ and $z^k = \nabla_x f(x^k, y^k)$. By definition of the counterpart of S (1.3) for (4.1), it follows that $y^k \in Y$ for all k. Hence, as Y is compact, it follows from the well-known Bolzano-Weierstrass Theorem that $\{y^k\}$ has a convergent subsequence, that we denote similarly, provided there is no confusion. Let $\bar{y}$ be the limit of this subsequence. Then, we have $\bar{y} \in S(\bar{x})$, given that S is closed as shown in Theorem 4.1. Subsequently, as f is continuously differentiable, the sequence $\{z^k\}$ converges to $\nabla_x f(\bar{x}, \bar{y})$. This shows that $\nabla_x f \circ \Psi$ is locally bounded around $\bar{x}$. It then follows from Proposition 2.4 that if the counterpart of (2.9) for the mapping in (3.4) holds at all $(a, b) \in \Gamma(\bar{x}, \underline{x})$, it holds that

$$\partial^2 \varphi(\bar{x}|\underline{x})(\underline{x}^*) \subseteq \bigcup_{(a,b)\in\Gamma(\bar{x}|\underline{x})} \left[\sum_{s=1}^{n+1} D^* \left(\nabla_x f \circ \Psi \right) (\bar{x}|b_s) \left(a_s \underline{x}^* \right) \right]. \tag{4.5}$$

Note that the mapping Γ in (4.5) corresponds to the counterpart of (2.12) for the convex hull set-valued mapping in (3.4). (4.5) leads to the estimate in the theorem, considering inclusions (4.2) and (4.3) and the fact that $(a, b) \in \Gamma(\bar{x}, \underline{x})$ if and only if

$$\exists z \text{ s.t. } (a, z) \in \Gamma^\circ(\bar{x}, \underline{x}) \text{ with } b = \left[\ \nabla_x f(\bar{x}, z^1)^\top, \ldots, \nabla_x f(\bar{x}, z^{n+1})^\top \ \right]^\top.$$

Finally, observe that (4.4) is a sufficient condition for the counterpart of qualification condition (2.9) for (3.4) to hold. □

Observe that based on the coderivative criterion (2.16), the qualification condition (4.4) is automatically satisfied if S is Lipschitz-like around $(\bar{x}, y)$ for all $y \in \Delta(\bar{x}, z^s)$, $(a, z) \in \Gamma^\circ(\bar{x}, \underline{x})$. Also, it is clear that the estimate of generalized Hessian of φ obtained in Theorem 4.1 is much tighter than the one derived in Theorem 4.3.

4.2. **Single-valued optimal solution and multipliers maps.** We assume throughout this subsection that the optimal solution mapping S (1.3) and the Lagrange multipliers mapping Λ (1.5) are all single-valued. Before we move to the general case, note that if s is single-valued in (4.1), we can get the following result, where the concave-convexity of f or the qualification condition (4.4) are not necessary.

Theorem 4.4. Consider the optimal value function φ (4.1) and let the corresponding optimal solution map S (1.3) be single-valued (i.e., $S := s$) and Lipschitz continuous around $\bar{x}$, where $s(\bar{x}) = \bar{y}$. Then,

$$\partial^2\varphi(\bar{x})(\underline{x}^*) \subseteq \nabla^2_{xx} f(\bar{x},\bar{y})\underline{x}^* + \partial\left\langle \nabla^2_{xy} f(\bar{x},\bar{y})\underline{x}^*, s\right\rangle(\bar{x}). \tag{4.6}$$

Proof. Consider the function $\psi : \mathbb{R}^n \to \mathbb{R}^{n+m}$ defined by $\psi(x) := (x^\top, s(x)^\top)^\top$. Then $\partial\varphi(x) := \nabla_x f \circ \psi(x)$ and we can check that

$$\begin{aligned}\partial^2\varphi(\bar{x})(x^*) &\subseteq \partial\left\langle \begin{bmatrix} \nabla^2_{xx} f(\bar{x},\bar{y})x^* \\ \nabla^2_{xy} f(\bar{x},\bar{y})x^* \end{bmatrix}, \psi\right\rangle(\bar{x}) \\ &\subseteq \nabla^2_{xx} f(\bar{x},\bar{y})x^* + \partial\left\langle \nabla^2_{xy} f(\bar{x},\bar{y})x^*, s\right\rangle(\bar{x}),\end{aligned}$$

while respectively using the chain and product rules in Theorems 2.2 and 2.3. □

Obviously, if s is single-valued and continuously differentiable at $\bar{x}$, we get the equality in (1.7). Also, it can be useful to observe that (1.6) provides an upper bound for the generalized Hessian of φ which is looser than the one in (4.6); cf. (2.5). Next, we show that under additional assumptions, inclusion (1.9) is valid in the case where the constraint function g effectively depends on both x and y.

Theorem 4.5. Consider φ (1.2) and suppose that gph K is compact and the MFCQ holds at $(\bar{x}, y)$, for all $y \in S(\bar{x})$. Further assume that the mappings S and Λ are single-valued (i.e., $S := s$ and $\Lambda := \lambda$) and Lipschitz continuous around $\bar{x}$ and $(\bar{x},\bar{y})$, respectively, with $\bar{y} = s(\bar{x})$ and $\bar{u} = \lambda(\bar{x},\bar{y})$. Then, we have inclusion (1.9).

Proof. It is clear that with gph K compact and the MFCQ satisfied at $(\bar{x}, y)$, for all $y \in S(\bar{x})$, we have from [12] that

$$\bar{\partial}\varphi(x) \subseteq \text{co} \bigcup_{y\in S(x)} \bigcup_{u\in\Lambda(x,y)} \left\{\nabla_x f(x,y) + \nabla_x g(x,y)^\top u\right\} \tag{4.7}$$

holds near $\bar{x}$. S and Λ being both single-valued around $\bar{x}$ and $(\bar{x},\bar{y})$, respectively, it follows from inclusion (4.7) that we have $\partial\varphi(x) = \nabla_x L \circ \phi \circ \psi(x)$ with ψ and ϕ respectively defined by $\phi(a,b) := [a^\top, b^\top, \lambda(a,b)^\top]^\top$ and $\psi(x) := [x^\top, s(x)^\top]^\top$. Hence,

$$\partial^2\varphi(\bar{x}|x)(x^*) = \partial\langle x^*, \nabla_x L \circ \phi \circ \psi\rangle(\bar{x}) \tag{4.8}$$

$$\subseteq \partial\langle \nabla(\nabla_x L)(\phi\circ\psi(\bar{x}))^\top x^*,\ \phi\circ\psi\rangle(\bar{x}) \tag{4.9}$$

$$\subseteq \bigcup_{y^*\in\partial\langle\nabla(\nabla_x L)(\phi\circ\psi(\bar{x}))^\top x^*,\ \phi\rangle(\psi(\bar{x}))} \partial\langle y^*, \psi\rangle(\bar{x}) \tag{4.10}$$

since the function $\nabla_x L$ is differentiable and ϕ and ψ are Lipschitz continuous around $\psi(\bar{x})$ and $\bar{x}$, respectively. One can easily observe that (4.9)–(4.10) result from the chain rule in Theorem 2.2. It now remains to evaluate the subdifferentials involved in (4.10). In fact, we can easily check that the following equalities hold:

$$\nabla(\nabla_x L)(\phi\circ\psi(\bar{x}))^\top x^* = \begin{bmatrix} \nabla^2_{xx} L(\bar{x}, s(\bar{x}), \lambda(\bar{x}, s(\bar{x})))x^* \\ \nabla^2_{xy} L(\bar{x}, s(\bar{x}), \lambda(\bar{x}, s(\bar{x})))x^* \\ \nabla_x g(\bar{x}, s(\bar{x}))x^* \end{bmatrix}, \tag{4.11}$$

$$\partial\langle\zeta^*, \phi\rangle(a,b) = \begin{bmatrix} \zeta^*_x \zeta^*_y \end{bmatrix} + \partial\langle\zeta^*_u, \lambda\rangle(a,b), \tag{4.12}$$

$$\partial\langle\zeta^*, \psi\rangle(\bar{x}) = \zeta^*_x + \partial\langle\zeta^*_y, s\rangle(\bar{x}). \tag{4.13}$$

It then follows from (4.11) and (4.12) that we have

$$y^* \in \partial\langle \nabla(\nabla_x L)(\phi\circ\psi(\bar{x}))^\top x^*,\ \phi\rangle(\psi(\bar{x})) \iff \left\{ \begin{array}{l} y^* = \left[\begin{array}{l} \nabla^2_{xx}L(\bar{x},s(\bar{x}),\lambda(\bar{x},s(\bar{x})))x^* \\ \nabla^2_{xy}L(\bar{x},s(\bar{x}),\lambda(\bar{x},s(\bar{x})))x^* \end{array}\right] + \zeta^* \\ \text{with } \zeta^* \in \partial\langle \nabla_x g(\bar{x},s(\bar{x}))x^*,\ \lambda\rangle(\bar{x},\bar{y}). \end{array}\right.$$

Substituting this, together with (4.13), in (4.10), we arrive at inclusion (1.9). □

If in addition to the assumptions made Theorem 4.5, we suppose that the functions s and λ are differentiable at $\bar{x}$ and $(\bar{x},\bar{y})$, respectively, then we have equality (1.10).

4.3. **Single-valued optimal solution map.** We assume here that the optimal solution mapping S (1.3) is single-valued; i.e., $S := s$. On the other hand, we let Λ (1.5) be set-valued. We can then estimate the generalized Hessian of φ (1.2) as follows.

Theorem 4.6. Suppose that gph $K := \{(x,y) \in \mathbb{R}^{n+m} |\ g(x,y) \leq 0\}$ is compact and S is singled-valued (i.e., $S := s$) and Lipschitz continuous around $\bar{x}$. Further assume that the MFCQ holds at $(\bar{x},\bar{y})$ with $\bar{y} = s(\bar{x})$ and the qualification condition

$$\left.\begin{array}{r} -a^* \in \partial\langle b^*, s\rangle(\bar{x}) \\ (a^*,b^*) \in D^*\Lambda(\bar{x},\bar{y}|u)(0) \end{array}\right\} \Longrightarrow \left\{\begin{array}{l} a^* = 0 \\ b^* = 0 \end{array}\right. \tag{4.14}$$

holds at $(\bar{x},\bar{y},u)$ with $\bar{y} = s(\bar{x})$ for any $u \in \Lambda(\bar{x},\bar{y})$. Then, for $\underline{x} \in \partial\varphi(\bar{x})$ and any $\underline{x}^* \in \mathbb{R}^n$, we have the following estimate for the second order subdifferential of φ:

$$\partial^2\varphi(\bar{x}|\underline{x})(\underline{x}^*) \subseteq \bigcup_{u\in\Lambda(\bar{x},\bar{y})} \bigcup_{(\zeta^*_x,\zeta^*_y)\in D^*\Lambda(\bar{x},\bar{y}|u)(\nabla_x g(\bar{x},\bar{y})\underline{x}^*)} \Big[\nabla^2_{xx}L(\bar{x},\bar{y},u)\underline{x}^* + \zeta^*_x + \partial\Big\langle \zeta^*_y + \nabla^2_{xy}L(\bar{x},\bar{y},u)\underline{x}^*,\ s\Big\rangle(\bar{x})\Big]. \tag{4.15}$$

Proof. Obviously, under the assumptions made, we equality (3.6), which can be rewritten as $\partial\varphi(x) = \nabla_x L\circ\Phi\circ\psi(x)$ with $\psi(x) := [x^\top, s(x)^\top]^\top$ and $\Phi(a,b) := \{(a,b)\}\times\Lambda(a,b)$. Consider the set-valued mapping $M(x,z) := \{(a,b,c) \in \Phi\circ\psi(x) |\ \nabla_x L(a,b,c) = z\}$ and a sequence $\{(x^k,z^k,a^k,b^k,c^k)\}$ with $(a^k,b^k,c^k) \in M(x^k,z^k)$. Then by definition,

$$\begin{array}{l} a^k = x^k,\ b^k = s(x^k),\ z^k = \nabla_x L(a^k,b^k,c^k), \\ \nabla_y L(a^k,b^k,c^k) = 0,\ c^k \geq 0,\ g(a^k,b^k) \leq 0,\ g(a^k,b^k)^\top c^k = 0. \end{array} \tag{4.16}$$

Suppose that $x^k \to \bar{x}$, $z^k \to \underline{x}$ and $\|c^k\| \geq k$ for all $k \in \mathbb{N}$. Then $b^k \to s(\bar{x})$, given that s is assumed to be locally Lipschitz continuous around $\bar{x}$. We can find a subsequence of $\{c^k\}$, with the same notation, provided there is no confusion, such that $c^k/\|c^k\|$ converges to some $\bar{c}$ with $\|\bar{c}\| = 1$. Inserting this subsequence in the second line of (4.16) and dividing the terms containing c^k by its norm, we arrive at (3.28) for the point $(\bar{x},\bar{y},\bar{c})$, as $k \to \infty$. Given that the MFCQ holds at $(\bar{x},\bar{y})$, it follows that we must have $\bar{c} = 0$. This contradicts the hypothesis that $\|c^k\| \geq k$ for all $k \in \mathbb{N}$. Hence, M is is locally bounded around $(\bar{x},\underline{x})$. Therefore, by the chain rule in Theorem 2.2,

$$\partial^2\varphi(\bar{x}|\underline{x})(\underline{x}^*) \subseteq \bigcup_{\substack{(a,b,u)\in\Phi\circ\psi(\bar{x}) \\ \underline{x}=\nabla_x L(a,b,u)}} \Big[D^*(\Phi\circ\psi)(\bar{x}|a,b,u)\Big(\nabla(\nabla_x L)(a,b,u)^\top \underline{x}^*\Big)\Big] \tag{4.17}$$

given that $\Phi \circ \psi$ is closed, as s is locally Lipschitz continuity around $\bar{x}$ and the functions f and g are continuously differentiable. Next, note that the set-valued mapping defined by $M_\circ(x,z) := \{(a,b) |\, (a,b) = \psi(x),\, z \in \Phi(a,b)\}$ is locally bounded around all the points $(\bar{x}, \bar{z})$ with $\bar{z} \in \Phi \circ \psi(\bar{x})$ and $\nabla_x L(\bar{z}) = \underline{x}$, given that s is Lipschitz continuous around $\bar{x}$ and for some neighborhoods X, A, B, C of $\bar{x}$, $\bar{a}$, $\bar{b}$, and $\bar{c}$, respectively, with X being a bounded set while A is a compact set, we have

$$M(D) = \bigcup_{(x,a,b,c)\in D} \{(\alpha,\beta) | \alpha = a = x, \beta = b = s(x), c \in \Lambda(x, s(x)\} \subseteq X \times s(X),$$

where $D := X \times A \times B \times C$. From the product rule in Theorem 2.3, it follows that for any $z^* = (x^*, y^*, u^*)$ and some $\bar{u} \in \Lambda(a,b)$ such that $\bar{z} = (a,b,\bar{u})$, we have

$$D^*\Phi(a,b|\bar{z})(z^*) \subseteq \begin{bmatrix} x^* \\ y^* \end{bmatrix} + D^*\Lambda(a,b|\bar{u})(u^*), \tag{4.18}$$

given that the counterpart of (2.9) is automatically satisfied. This obviously leads to $D^*\Phi(a,b|\bar{z})(0) \subseteq D^*\Lambda(a,b|\bar{u})(0)$. Then considering the fact that

$$\text{Ker } D^*\psi(\bar{x}|a,b) := \{(a^*,b^*) |\, 0 \in \partial\langle (a^*,b^*), \psi\rangle(\bar{x})\} = \{(a^*,b^*) |\, -a^* \in \partial\langle b^*, s\rangle(\bar{x})\},$$

it is clear that (4.14) is sufficient for $D^*\Phi(a,b|\bar{z})(0) \cap \text{Ker } D^*\psi(\bar{x}|a,b) = \{0\}$ to hold. Hence, from Theorem 2.2, take any $(a,b,u) \in \Phi \circ \psi(\bar{x})$ such that $\underline{x} = \nabla_x L(a,b,u)$,

$$\begin{aligned} &D^*(\Phi \circ \psi)(\bar{x}|a,b,c)\,(\nabla(\nabla_x L)(a,b,c)^\top \underline{x}^*) \\ &\qquad\subseteq \bigcup_{\zeta^* \in D^*\Phi(\bar{x},\bar{y}|\bar{x},\bar{y},u)(\nabla(\nabla_x L)(\bar{x},\bar{y},u)^\top \underline{x}^*)} \partial\langle \zeta^*,\, \psi\rangle(\bar{x}), \end{aligned} \tag{4.19}$$

given that $(a,b,u) \in \Phi \circ \psi(\bar{x})$ is equivalent to $a = \bar{x}$, $b = \bar{y}$ and $u \in \Lambda(\bar{x},\bar{y})$. Clearly, since $\partial\langle (a^*,b^*), \psi\rangle(\bar{x}) = a^* + \partial\langle b^*, s\rangle(\bar{x})$, we have (4.15) from a combination of inclusions (4.17), (4.18), and (4.19). □

The result of this theorem is clearly an extension of Theorem 4.5, as both estimates will coincide if Λ is locally single-valued and Lipschitz continuous. Note that the latter property ensures that qualification condition (4.14) automatically holds.

Corollary 4.7. Let all the assumptions of Theorem 4.6 hold and suppose that (3.18) is satisfied at $(\bar{x},\bar{y},u)$ for all $u \in \Lambda(\bar{x},\bar{y})$ with $\bar{y} = s(\bar{x})$. Then, for $\underline{x} \in \partial\varphi(\bar{x})$ and $\underline{x}^* \in \mathbb{R}^n$,

$$\begin{aligned} \partial^2\varphi(\bar{x}|\underline{x})(\underline{x}^*) \subseteq \bigcup_{u\in\Lambda(\bar{x},\bar{y})} \; \bigcup_{(a,c)\in\mho(\bar{x},\bar{y},u,\nabla_x g(\bar{x},\bar{y})\underline{x}^*)} &\Big[\nabla^2_{xx}L(\bar{x},\bar{y},u)\underline{x}^* + \nabla^2_{yx}L(\bar{x},\bar{y},u)a + \nabla_x g(\bar{x},\bar{y})^\top c \\ &+ \partial\Big\langle \nabla^2_{xy}L(\bar{x},\bar{y},u)\underline{x}^* + \nabla^2_{yy}L(\bar{x},\bar{y},u)a + \nabla_y g(\bar{x},\bar{y})^\top c,\ s\Big\rangle(\bar{x})\Big]. \end{aligned}$$

4.4. **Single-valued Lagrange multipliers map.** In this subsection, we suppose that the Lagrange multipliers set-valued mapping Λ is single-valued; i.e., $\Lambda := \lambda$.

Theorem 4.8. Suppose that gph K is compact and the LICQ holds at $(\bar{x}, y)$ for all $y \in S(\bar{x})$. Further assume that Λ is singled-valued and Lipschitz continuous around

$(\bar{x}, y)$ with $u = \lambda(\bar{x}, y)$ for all $y \in S(\bar{x})$ such that $\nabla_x L(\bar{x}, y, u) = \underline{x}$. If S is closed and locally bounded around $\bar{x}$ and problem (1.1) is concave-convex, then for $\underline{x}^* \in \mathbb{R}^n$,

$$\partial^2 \varphi(\bar{x}|\underline{x})(\underline{x}^*) \subseteq \bigcup_{y \in S(\bar{x}):\ \nabla_x L(\bar{x},y,u)=\underline{x}} \ \bigcup_{(\zeta_x^*, \zeta_y^*) \in \partial \langle \nabla_x g(\bar{x},y)\underline{x}^*,\ \lambda \rangle(\bar{x},y)} \Big[\nabla^2_{xx} L(\bar{x}, y, u)\underline{x}^* + \zeta_x^* + D^* S(\bar{x}|y)\left(\zeta_y^* + \nabla^2_{xy} L(\bar{x}, y, u)\underline{x}^*\right)\Big]. \tag{4.20}$$

Proof. Under the assumptions made, (3.8) holds. Hence, $\partial\varphi(x) = \nabla_x L \circ \phi \circ \Psi(x)$ with $\phi(a, b) := [a^\top, b^\top, \lambda(a, b)^\top]^\top$ and $\Psi(x) := \{x\} \times S(x)$. For some neighborhoods X, U, V, W of $\bar{x}$, $\bar{u}$, $\bar{v}$, and $\bar{w}$, respectively, with X chosen such that it is bounded and $S(X) \subseteq Y$, Y being a bounded set being (this is possible considering the local boundedness of S around $\bar{x}$). We have can easily check that $M(X \times U \times V \times W) \subseteq X \times S(X)$. Hence the mapping $M(x, z) := \{(a, b) \in \Psi(x)|\ \nabla_x L \circ \phi(a, b) = z\}$ is locally bounded around $(\bar{x}|\underline{x})$. Then since $\nabla_x L \circ \phi$ is single-valued and locally Lipschitz continuous around $\bar{x}$ and Ψ closed, as S is closed, it follows from Theorem 2.2 that

$$\begin{aligned} \partial^2 \varphi(\bar{x}|\underline{x})(\underline{x}^*) \ &\subseteq \bigcup_{y \in S(\bar{x}),\ \nabla_x L \circ \phi(\bar{x},y)=\underline{x}} D^*\Psi(\bar{x}|\bar{x}, y) \circ \partial \left\langle \underline{x}^*, \nabla_x L \circ \phi \right\rangle(\bar{x}, y) \\ &\subseteq \bigcup_{y \in S(\bar{x}):\ \nabla_x L \circ \phi(\bar{x},y)=\underline{x}} D^*\Psi(\bar{x}|\bar{x}, y) \circ \partial \left\langle \nabla(\nabla_x L)(\phi(\bar{x}, y))^\top \underline{x}^*, \phi \right\rangle(\bar{x}, y). \end{aligned}$$

Subsequently, by considering the estimate of the coderivative of ϕ and the one of Ψ from (4.3) in the current context, we have inclusion (4.20). $\square$

Corollary 4.9. Let the assumptions of Theorem 4.8 be satisfied and suppose that the qualification condition (3.18) is satisfied at $(\bar{x}, y, u)$ for all $y \in S(\bar{x})$ such that $\nabla_x L(\bar{x}, y, u) = \underline{x}$ and $u = \lambda(\bar{x}, \bar{y})$. Then for $\underline{x}^* \in \mathbb{R}^n$, it holds that

$$\begin{aligned} \partial^2 \varphi(\bar{x}|\underline{x})(\underline{x}^*) \subseteq \bigcup_{y \in S(\bar{x}):\ \nabla_x L(\bar{x},y,u)=\underline{x}} \ \bigcup_{(\zeta_x^*, \zeta_y^*) \in \partial \langle \nabla_x g(\bar{x},y)\underline{x}^*,\ \lambda \rangle(\bar{x},y)} \ \bigcup_{(a,c) \in \mho(\bar{x},y,u,0)} \Big\{ &\zeta_x^* + \nabla^2_{xx} L(\bar{x}, y, u)\underline{x}^* + \nabla^2_{xy} L(\bar{x}, y, u)^\top a + \nabla_x g(\bar{x}, y)^\top c \Big| \\ &\zeta_y^* + \nabla^2_{xy} L(\bar{x}, y, u)\underline{x}^* + \nabla^2_{yy} L(\bar{x}, y, u) a + \nabla_y g(\bar{x}, y)^\top c = 0 \Big\}. \end{aligned}$$

Similarly to Theorem 4.3, we now drop the concave-convex assumption made in Theorem 4.8. Then we have the following result, where

$$\begin{aligned} \Gamma^\lambda(\bar{x}, \underline{x}) := \Big\{ &(a, z, w) \in \mathbb{R}^{n+1} \times \textstyle\prod_{s=1}^{n+1} \mathbb{R}^m \times \prod_{s=1}^{n+1} \mathbb{R}^p \Big|\ a \geq 0,\ \sum_{s=1}^{n+1} a_s = 1, \\ &\textstyle\sum_{s=1}^{n+1} a_s \nabla_x L(\bar{x}, z^s, w^s) = \underline{x},\ z \in \prod_{s=1}^{n+1} S(\bar{x}),\ w \in \prod_{s=1}^{n+1} \{\lambda(\bar{x}, z^s)\} \Big\} \end{aligned}$$

and

$$\Delta^\lambda(\bar{x}, z^s, w^s) := \Big\{ (y, u)|\ y \in S(\bar{x}),\ \ u = \lambda(\bar{x}, y),\ \nabla_x L(\bar{x}, y, u) = \nabla_x L(\bar{x}, z^s, w^s) \Big\}.$$

Theorem 4.10. Suppose that gph K is compact and the LICQ holds at all (x, y) on the graph of S. Further assume that Λ is singled-valued and Lipschitz continuous around $(\bar{x}, y)$ with $u = \lambda(\bar{x}, y)$ for all $y \in S(\bar{x})$ such that $\nabla_x L(\bar{x}, y, u) = \underline{x}$. If S is closed and locally bounded around $\bar{x}$ and the qualification condition

$$\left[\sum_{s=1}^{m+1} v_s = 0,\ v_s \in \bigcup_{y\in\Delta(\bar{x},z^s,w^s)} D^*S\,(\bar{x}|y)\,(0),\ s=1,\dots,n+1\right] \Longrightarrow v_1 = \ldots = v_{n+1} = 0 \tag{4.21}$$

holds at all $(a,z,w)\in\Gamma^{\lambda}(\bar{x},\underline{x})$, where $\underline{x}\in\varphi(\bar{x})$. Then, for all $\underline{x}^*\in\mathbb{R}^n$, it holds that

$$\varphi^2(\bar{x}|\underline{x})(\underline{x}^*) \subseteq \bigcup_{(a,z,w)\in\Gamma^{\lambda}(\bar{x},\underline{x})} \left\{ \sum_{s=1}^{n+1}\left[\bigcup_{(y,u)\in\Delta^{\lambda}(\bar{x},z^s,w^s)} \ \bigcup_{(\zeta_x^*,\zeta_y^*)\in a_s\partial\langle\nabla_x g(\bar{x},y)x^*,\ \lambda\rangle(\bar{x},y)} \left[a_s\nabla^2_{xx}L(\bar{x},y,u)\underline{x}^* + \zeta_x^* + D^*S(\bar{x}|y)\left(\zeta_y^* + a_s\nabla^2_{xy}L(\bar{x},y,u)\underline{x}^*\right)\right]\right]\right\}.$$

Proof. Based on the compactness of gph K, the closedness of S and the Bolzano-Weierstrass Theorem we can easily show that $\nabla_x L\circ\phi\circ\Psi$ is closed and locally bounded around $\bar{x}$, thanks to the locally Lipschitz continuity of ϕ. The rest of the proof follows the steps of that of Theorem 4.3. $\square$

5. Applications to linear programming

5.1. **LHS perturbation.** Consider the following optimal value function with left-hand-side (LHS) perturbation:

$$\varphi(x) := \min_y \{x^\top y|\ Ay \le b\}.$$

Suppose that $Y := \{y|\ Ay\le b\}$ is compact and the matrix A is full rank. For a couple $(\bar{x},\underline{x})$, we have $\underline{x}\in\varphi(\underline{x})$ if and only if $\underline{x}\in S(\underline{x})$. We also have $\bar{u}=\lambda(\bar{x},\underline{x})$, as A is full rank. For $(\bar{x},\underline{x},\bar{u})$, consider the definitions of ν, η, and θ given in (3.16) for the constraint set Y, and introduce the corresponding set

$$\Xi(A) := \left\{(a,c)|\ A_\nu a = 0,\ c_\eta = 0,\ \forall i\in\theta:\ (A_i a > 0 \wedge c_i > 0) \vee c_i\,(A_i a) = 0\right\}.$$

Further note that the qualification condition (3.18) holds at the point $(\bar{x},\underline{x},\bar{u})$, as A has full rank. It therefore follows from Corollary 4.2 that

$$\partial^2\varphi(\bar{x}|\underline{x})(\underline{x}^*) \subseteq \left\{a|\ (a,c)\in\Xi(A),\ A^\top c + \underline{x}^* = 0\right\}.$$

5.2. **LHS and RHS perturbations.** Here, we consider the following optimal value function with both left-hand-side (LHS) and right-hand-side (RHS) perturbations:

$$\varphi(x) := \min_y \{x^\top y|\ Ay \le x\}.$$

Suppose that the set $\{(x,y)|\ Ay\le x\}$ is compact and the matrix A is full rank. For a couple $(\bar{x},\underline{x})$ such that $\underline{x}\in\varphi(\underline{x})$, take $y\in S(\underline{x})$. Then, as in the previous case, introduce ν, η, and θ at the point $(\bar{x},y,u)$. Then, at this point, we consider the set $\Xi(A)$ defined as in the previous subsection. It follows from Corollary 4.9 that

$$\partial^2\varphi(\bar{x}|\underline{x})(\underline{x}^*) \subseteq \bigcup_{(y,u)\in\Omega(\bar{x},\underline{x})}\ \bigcup_{(\zeta_x^*,\zeta_y^*)\in\partial\langle-\underline{x}^*,\ \lambda\rangle(\bar{x},y)}\ \bigcup_{(a,c)\in\Xi(A)} \left\{\zeta_x^* - c|\ \zeta_y^* + A^\top c = 0\right\}$$

given that as A has full rank, the qualification condition (3.18) holds at any point $(\bar{x}, y, u)$ with $(y,u) \in \Omega(\bar{x}, \underline{x}) := \{(y,u) |\; y \in S(\bar{x}),\ u = \lambda(\bar{x}, y) = -\underline{x}\}$.

6. Conclusions and future work

We have provided estimates for the generalized Hessian/second order subdifferential (1.4) of the optimal value (1.2) for the parametric optimization (1.1). The focus here was on four scenarios: (a) the case where the feasible set of problem (1.1) is independent of the parameter x (cf. Subsection 4.1); (b) the optimal solution S (1.3) and Lagrange multipliers mappings Λ (1.5) are single-valued (cf. Subsection 4.2); (c) S is single-valued and Λ is set-valued (cf. Subsection 4.3) and ; (d) S is set-valued and Λ is single-valued (cf. Subsection 4.4). From the exposition done in this paper, it is clear that there are still at least two gaps in the literature that need to be filled. The first one is a detailed study on conditions tailored to the Lagrange multipliers mapping Λ to ensure that it is locally single-valued and Lipschitz continuous, as a "function" of (x,y). Note that it is obvious that this mapping can be written as

$$\Lambda(x,y) := \{u \in \mathbb{R}^p |\; 0 \in \psi(x,y,u) + \Phi(u)\},$$

where the function ψ and the set-valued mapping Φ are respectively defined by $\psi(x,y,u) := -\left[\nabla_y L(x,y,u)^\top, g(x,y)^\top\right]^\top$ and $\Phi(u) := \{0_m\} \times N_{\mathbb{R}^p_+}(u)$. Many papers have been devoted to conditions ensuring that general mappings of the above form are locally single-valued and Lipschitz continuous; see, e.g., [10, 19]. As for the applications of such results to problems of the form (1.1), focus has mostly been on the deriving conditions where the Lagrange multipliers (and also the optimal solution) mapping is locally single-valued and Lipschitz continuous, as a "function" of x and not necessarily of (x,y), as needed in the context of this paper. We have provided an estimate of the coderivative of Λ in this paper, deduce conditions ensuring that it is Lipschitz-like, and also establish a nice relationship between the latter properties for Λ and those of the optimal solution mapping S (1.3); cf. Section 3.

Another open question from the work in this paper is how to derive estimates for the generalized Hessian of φ in the case where both S and Λ are set-valued mappings. In this case, it is well-known that the subdifferential of φ can be estimated as in (4.7). If additionally, problem (1.1) is concave-convex, then proceeding as in the case of Theorem 3.1(iii), it should be possible to prove that the convex hull operator can be dropped from (4.7). But the question is whether we can have equalities in the latter inclusions. If (4.7) holds as equality in some neighborhood of $\bar{x}$, where we are able to drop the convex hull operator, then under appropriate extensions of the results in this paper, it should be possible to prove that

$$\partial^2 \varphi(\bar{x}|\underline{x})(\underline{x}^*) \subseteq \bigcup_{(y,u)\in\Delta(\bar{x},\underline{x})} \ \bigcup_{(\zeta_x^*,\zeta_y^*)\in D^*\Lambda(\bar{x},y|u)(\nabla_x g(\bar{x},y)\underline{x}^*)} \left[\zeta_x^* + \nabla^2_{xx} L(\bar{x},y,u)\underline{x}^* + D^*S(\bar{x}|y)\left(\zeta_y^* + \nabla^2_{xy} L(\bar{x},y,u)\underline{x}^*\right)\right]$$

with $\Delta(\bar{x},\underline{x}) := \{(y,u) \in \mathbb{R}^m \times \mathbb{R}^p |\; y \in S(\bar{x}),\ u \in \Lambda(\bar{x},y),\ \underline{x} = \nabla_x L(\bar{x},y,u)\}$. The aforementioned points will be explored more carefully in future works.

References

[1] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski. *Robust optimization*. Princeton University Press, 2009.

[2] P. Bernhard and A. Rapaport. On a theorem of danskin with an application to a theorem of von neumann-sion. *Nonlinear Anal.*, 24(8):1163–1181, 1995.

[3] J.F. Bonnans and A. Shapiro. *Perturbation analysis of optimization problems*. New York: Springer, 2000.

[4] J.M. Danskin. *The theory of Max-Min and its application to weapons allocation problems.* Berlin-Heidelberg-New York: Springer-Verlag, 1967.

[5] G.B. Dantzig. *Linear Programming and Extensions*. Princeton University Press, 1963.

[6] S. Dempe. *Foundations of bilevel programming*. Dordrecht: Kluwer Academic Publishers, 2002.

[7] S. Dempe, B. Mordukhovich, and A.B. Zemkoho. Sensitivity analysis for two-level value functions with applications to bilevel programming. *SIAM J. Optim.*, 22(4):1309–1343, 2012.

[8] S. Dempe and A.B. Zemkoho. On the Karush-Kuhn-Tucker reformulation of the bilevel optimization problem. *Nonlinear Anal.*, 75(3):1202–1218, 2012.

[9] S. Dempe and A.B. Zemkoho. KKT reformulation and necessary conditions for optimality in nonsmooth bilevel optimization. *SIAM J. Optim.*, 24(4):1639–1669, 2014.

[10] A.L. Dontchev and W.W. Hager. Implicit functions, lipschitz maps, and stability in optimization. *Math. Oper. Res.*, 19(3):753–768, 1994.

[11] A.V. Fiacco. *Introduction to sensitivity and stability analysis in nonlinear programming*. New York - London: Academic Press, 1983.

[12] J. Gauvin and F. Dubeau. Differential properties of the marginal function in mathematical programming. *Math. Program. Study*, 18:101–119, 1982.

[13] B.S. Mordukhovich. Generalized differential calculus for nonsmooth and set-valued mappings. *J. Math. Anal. Appl.*, 183(1):250–288, 1994.

[14] B.S. Mordukhovich. *Variational analysis and generalized differentiation. I: Basic theory. II: Applications*. Berlin: Springer, 2006.

[15] B.S. Mordukhovich, M.N. Nam, and H.M. Phan. Variational analysis of marginal function with applications to bilevel programming problems. *J. Optim. Theory Appl.*, 152(3):557–586, 2012.

[16] B.S. Mordukhovich, M.N. Nam, and N.D. Yen. Subgradients of marginal functions in parametric mathematical programming. *Math. Program.*, 116(1-2):369–396, 2009.

[17] B.S. Mordukhovich and J.V. Outrata. Coderivative analysis of quasi-variational inequalities with applications to stability and optimization. *SIAM J. Optim.*, 18(2):389–412, 2007.

[18] D. Ralph and S. Dempe. Directional derivatives of the solution of a parametric nonlinear program. *Math. Program.*, 70(1):159–172, 1995.

[19] S.M. Robinson. Some continuity properties of polyhedral multifunctions. *Math. Program. Study*, 14:206–214, 1981.

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

[21] J.-J. Rückmann and A. Shapiro. Second-order optimality conditions in generalized semi-infinite programming. *Set-Valued Analysis*, 9(1):169–186, 2001.

[22] S. Scholtes. Introduction to piecewise differentiable equations. Technical Report No. 53/1994, Universität Karlsruhe, Institut für Statistik und Mathematische Wirtschaftstheorie, 1994.

[23] K. Shimizu, Y. Ishizuka, and J.F. Bard. *Nondifferentiable and two-level mathematical programming*. Dordrecht: Kluwer Academic Publishers, 1997.