J. Appl. Numer. Optim. 6 (2024), No. 2, pp. 153-175
Available online at http://jano.biemdas.com
https://doi.org/10.23952/jano.6.2024.2.01

PRIMAL AND DUAL SECOND-ORDER NECESSARY OPTIMALITY CONDITIONS IN BILEVEL PROGRAMMING

ZAKARYA DARDOUR[1], LAHOUSSINE LAFHIM[1], EL MOSTAFA KALMOUN[2,*]

[1]*Department of Mathematics, Sidi Mohammed Ben Abdellah University, Fez, Morocco*
[2]*School of Science and Engineering, Al Akhawayn University in Ifrane, Ifrane, Morocco*

Abstract. The purpose of this paper is to derive primal and dual second-order necessary optimality conditions for a standard bilevel optimization problem with both smooth and nonsmooth data. The approach involves utilizing two different reformulations of the hierarchical model as a single-level problem under a partial calmness assumption. The first reformulation consists on the use of the value function of the lower-level problem, which is then tackled by using second-order directional derivatives. However, for the dual conditions, this approach is not suitable except for cases that the value function is smooth. Therefore, we adopt a second approach that relies on the Ψ-reformulation. In both cases, the obtained necessary optimality conditions can be expressed according to the problem data. Finally, some examples are given to illustrate the proven results.

Keywords. Bilevel optimization; Clarke's generalized derivative; Nonsmooth optimization; Partial calmness; Second-order directional derivative.

2020 Mathematics Subject Classification. 90C46, 90C29.

1. Introduction

Bilevel programming has emerged as a significant area of research in modern optimization theory. It involves a hierarchical structure comprising two decision levels, upper and lower, with the constraint region of the upper-level problem determined implicitly by the solution set to the lower-level problem. These applications are used in many different fields, including finance, economics, chemistry, and logistics.

The standard bilevel optimization problem, which we consider throughout this paper, is defined as follows

$$\min_{x,y} \{F(x,y) \mid G(x,y) \leq 0, y \in S(x)\}, \tag{BOP}$$

also known as the upper-level problem, and the lower-level problem

$$\min_{y} \{f(x,y) \mid g(x,y) \leq 0\}. \tag{$P[x]$}$$

*Corresponding author.
E-mail address: E.Kalmoun@aui.ma (E.M. Kalmoun).
Received 26 May 2023; Accepted 24 October 2023; Published online 1 April 2024.

The optimal solution mapping of the lower-level problem ($P[x]$) is represented by the set-valued mapping $S:\mathbb{R}^n \rightrightarrows \mathbb{R}^m$, which is defined as follows:

$$\forall x \in \mathbb{R}^n: \quad S(x) := \underset{y}{\operatorname{argmin}}\{f(x,y) \mid g(x,y) \leq 0\}. \tag{1.1}$$

Problem (BOP) has a two-level structure, upper-level problem with the objective function $F:\mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, and constraint function $G:\mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^P$ of components $G_1,\ldots,G_p:\mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$. The lower-level problem ($P[x]$) of (BOP) has as objective function $f:\mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, and constraint function $g:\mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^q$ of components $g_1,\ldots,g_q:\mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$. Note that the minimization in (BOP) is done with respect to x and y.

Bilevel programs are mathematically difficult to study, as they are generally irregular, non-convex, and non-smooth. A reformulation of the hierarchical model as a single-level program results in surrogate problems that have an inherent lack of smoothness, convexity, and regularity. To address this issue, various methods were devised and developed to determine necessary optimality conditions for bilevel optimization problems.

The use of second-order optimality conditions can be very helpful in both understanding the properties of the solution and analyzing the convergence of solution algorithms. In particular, second-order conditions can help to determine whether a local optimum is also a global optimum, which is important to ensure the quality of a solution. They can also provide information about the sensitivity of the solution to changes in the problem data, which is efficient in understanding the robustness of the solution.

Overall, the use of second-order conditions in bilevel optimization holds great promise for improving the understanding and solution of these challenging optimization problems. There are numerous results on second-order conditions, including [1, 6, 15] for problems with C^2 and $C^{1.1}$ data, and [16, 23] for problems with only C^1 data.

In the literature, few authors investigated the second-order necessary optimality conditions for bilevel optimization problems. Dempe and Gadhi [9] utilized an approximation of the contingent cone to the feasible set of the bilevel problem and described it by using a support function to an auxiliary set-valued mapping. This approach enables them to establish second-order necessary and sufficient optimality conditions for the optimistic case of bilevel. Lafhim [22] investigated second-order necessary and sufficient optimality conditions for the optimistic case of a bilevel multiobjective programming problem under a generalized Abadie constraint qualification without the assumption that the lower-level problem satisfies the Mangasarian Fromovitz constraint qualification (*MFCQ*) by using the optimal value function of the lower-level problem. However, all these optimality conditions may not always capture the inherent complexities of bilevel optimization because they do not rely on the initial problem data.

To obtain second-order necessary optimality conditions in terms of initial data of problem (BOP), one may reformulate it as a single-level optimization problem and apply optimality conditions to the single-level problem under a partial calmness assumption considered by Ye and Zhu [31], using two different approaches. The first approach for reformulation involves utilizing the value function of the bilevel optimization problem and applying second-order directional derivatives. However, this method may not be appropriate for the dual conditions, except in cases that the value function is smooth. As a result, we chosen to use the second approach, which involves relying on the Ψ-reformulation, introduced by Ye in [29] as well as the concept of partial calmness, considered by Ye and Zhu in [31].

The remaining parts of this paper are organized as follows. In Section 2, we briefly summarize some reminders, definitions and regularity conditions used in rest of the paper. We use lower-level value function reformulation (*LLVF*) and the concept of partial calmness to derive second-order optimality conditions in Section 3. Section 4 is dedicated to the Ψ-reformulation of (BOP) to obtain second-order necessary optimality conditions (Primal and dual) for the bilevel optimization problem under consideration. Finally, in Section 5, we provide a conclusion and discuss future directions for research.

2. Preliminaries

2.1. **Notations and constraint qualifications.** Next, we give some constructions that we need in the sequel. We set $I=\{1,\dots,p\}$ and $J=\{1,\dots,q\}$.

Let us define the set of feasible points of lower-level problem ($P[x]$)

$$\Pi := \{(x,y)\in\mathbb{R}^n\times\mathbb{R}^m \mid g(x,y)\le 0\},$$

and the set-valued map

$$Y(x) := \{y\in\mathbb{R}^m \mid (x,y)\in\Pi\},\quad \forall x\in\mathbb{R}^n,$$

which provides all lower-level feasible points for a given value of x. Let

$$domY = \{x\in\mathbb{R}^n \mid Y(x)\neq\emptyset\} \quad \text{and} \quad \mathrm{gph}Y = \{(x,y)\mid x\in\mathbb{R}^n, y\in Y(x)\}.$$

For our lower-level problem ($P[x]$), we define the lower-level Lagrangian function ℓ by

$$\ell(x,y,w) := f(x,y)+w^\top g(x,y),\quad \forall (x,y,w)\in\mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}^q.$$

For any vector $(x,y)\in\mathrm{gph}Y$, the set of lower-level Lagrange multipliers is defined as

$$\Lambda(x,y) := \left\{w\in\mathbb{R}^q \mid \nabla_y\ell(x,y,w)=0, w\ge 0, g(x,y)\le 0, w^\top g(x,y)=0\right\}.$$

For $d_1\in\mathbb{R}^n$, let

$$\Lambda^2(x,y,d_1) := \left\{w\in\Lambda(x,y) \mid \langle\nabla_x\ell(x,y,w),d_1\rangle = \max_{\omega\in\Lambda(x,y)}\langle\nabla_x\ell(x,y,\omega),d_1\rangle\right\}.$$

We denote by

$$\Omega := \{(x,y)\in\mathbb{R}^n\times\mathbb{R}^m \mid G(x,y)\le 0, g(x,y)\le 0\},$$

the set of feasible points of (BOP). For every feasible point $(\bar{x},\bar{y})\in\Omega$, we define

$$\bar{I}^G = I^G(\bar{x},\bar{y}) := \{i\in I \mid G_i(\bar{x},\bar{y})=0\},\ \text{and}$$
$$\bar{I}^g = I^g(\bar{x},\bar{y}) := \left\{j\in J \mid g_j(\bar{x},\bar{y})=0\right\},$$

as the set of active constraints for the upper-level problem and the lower-level problem, respectively.

The following standard conditions, called, *LLICQ*, *LSOSC*, and *LMFCQ* are applied to the lower-level problem.

1. LLICQ. The lower-level linear independence constraint qualification (*LLICQ*) is verified at a point $(\bar{x},\bar{y})\in\mathrm{gph}Y$ if the family of gradient vectors

$$\{\nabla_y g_j(\bar{x},\bar{y})\ ,\ j\in\bar{I}^g\}$$

is linearly independent.

2. LMFCQ. The lower-level Mangasarian-Fromovitz-Constraint-Qualification (*LMFCQ*) is said to hold at a point $(\bar{x},\bar{y}) \in \text{gph}Y$ if there is a vector $d \in \mathbb{R}^m$ such that

$$\nabla_y g_j(\bar{x},\bar{y})^T d < 0 \ , \ \forall j \in \bar{I}^g.$$

Since (*LMFCQ*) confirms the positive linear independence of the gradients of active lower-level constraints, it is obviously weaker than (*LLICQ*). For further discussion on these conditions, we refer to Bazaraa et al. [4] and Janin [19].

3. LSOSC. We define the lower-level critical cone at $(\bar{x},\bar{y}) \in \text{gph}Y$ as

$$\mathscr{C}^l(\bar{x},\bar{y}) = \left\{ d_2 \in \mathbb{R}^m \ \middle| \ \begin{array}{ll} \nabla_y f(\bar{x},\bar{y})^T d_2 & = 0 \\ \nabla_y g_j(\bar{x},\bar{y})^T d_2 & \leq 0 \ , \ \ j \in \bar{I}^g \end{array} \right\}.$$

The lower-level second-order sufficient condition (*LSOSC*) is verified at $(\bar{x},\bar{y})$ if we have

$$\forall d_2 \in \mathscr{C}^l(\bar{x},\bar{y}) \setminus \{0\} \ , \ \exists w \in \Lambda(\bar{x},\bar{y}) \ : \ d_2^T \nabla^2_{yy} \ell(\bar{x},\bar{y},w) d_2 > 0.$$

2.2. **Generalized differentiation.** In this subsection, we review some basic first-order directional differentiability concepts; see, e.g., [7, 27]. To proceed, let $\psi : \mathbb{R}^n \to \overline{\mathbb{R}}$ and $\bar{x} \in \text{dom}\psi$. For a direction $d \in \mathbb{R}^n$, the limits

$$\psi_d^+(\bar{x},d) := \limsup_{t\downarrow 0} \frac{\psi(\bar{x}+td)-\psi(\bar{x})}{t} \quad \text{and} \quad \psi_d^-(\bar{x},d) := \liminf_{t\downarrow 0} \frac{\psi(\bar{x}+td)-\psi(\bar{x})}{t}$$

are, respectively, referred to as the upper and lower Dini directional derivatives of ψ at $\bar{x}$ in the direction d. We denote the directional derivative of ψ at $\bar{x}$ in the direction d by

$$\psi'(\bar{x},d) := \lim_{t\downarrow 0} \frac{\psi(\bar{x}+td)-\psi(\bar{x})}{t}. \tag{2.1}$$

This first-order derivative is in Gâteaux's sense. In a similar way, we define the upper and lower Hadamard directional derivatives of ψ at $\bar{x}$ in the direction d as

$$\psi_H^+(\bar{x},d) := \limsup_{t\downarrow 0, u\to d} \frac{\psi(\bar{x}+tu)-\psi(\bar{x})}{t} \quad \text{and} \quad \psi_H^-(\bar{x},d) := \liminf_{t\downarrow 0, u\to d} \frac{\psi(\bar{x}+tu)-\psi(\bar{x})}{t}.$$

When it exists, we call

$$\psi^H(\bar{x},d) := \lim_{t\downarrow 0, u\to d} \frac{\psi(\bar{x}+tu)-\psi(\bar{x})}{t}$$

the directional Hadamard direvative of ψ at $\bar{x}$. Also, we recall that the Clarke directional derivative of ψ at $\bar{x}$ in the direction d is given by

$$\psi^\circ(\bar{x},d) := \limsup_{t\downarrow 0, x\to \bar{x}} \frac{\psi(x+td)-\psi(x)}{t}.$$

We say that ψ is directionally differentiable at $\bar{x}$ if, for each $d \in \mathbb{R}^n$, $\psi'(\bar{x},d)$ exists. Similarly, ψ is Hadamard or Clarke directionally differentiable at $\bar{x}$ if the corresponding directional derivative $\psi^H(\bar{x},d)$ or $\psi^\circ(\bar{x},d)$ exists for each $d \in \mathbb{R}^n$.

In the following result, we list some important properties of the above directional derivatives.

Proposition 2.1. *Let $\psi : \mathbb{R}^n \to \overline{\mathbb{R}}$ and $\bar{x} \in$ domψ. Let ψ be locally Lipschitz continuous at $\bar{x}$. Then*

(1) $\psi_d^+(\bar{x},\cdot)$ and $\psi_d^-(\bar{x},\cdot)$ are bounded and continuous w.r.t the second argument.

(2) For any $d \in \mathbb{R}^n$, $\psi_H^+(\bar{x},d) = \psi_d^+(\bar{x},d)$, and $\psi_d^-(\bar{x},d) = \psi_H^-(\bar{x},d)$.

(3) If ψ is directionally differentiable at $\bar{x}$, then it is also Hadamard directionally differentiable at $\bar{x}$. Furthermore, $\psi'(\bar{x},\cdot)$ and $\psi^H(\bar{x},\cdot)$ coincide.

(4) If ψ is continuously differentiable at $\bar{x}$, then, for any $d \in \mathbb{R}^n$, all of the generalized derivatives mentioned above are equal to $\nabla\psi(\bar{x})^\top d$, i.e.,

$$\psi'(\bar{x},d) = \psi^H(\bar{x},d) = \psi^\circ(\bar{x},d) = \nabla\psi(\bar{x})^\top d.$$

Recall that a locally Lipschitz function $\psi : \mathbb{R}^n \to \overline{\mathbb{R}}$ is said to be regular at a point $\bar{x} \in dom\psi$ if the directional derivative $\psi'(\bar{x},d)$ exists in every direction $d \in \mathbb{R}^n$ and $\psi'(\bar{x},d) = \psi^\circ(\bar{x},d)$ in any direction $d \in \mathbb{R}^n$ (Clarke [7]).

Below, we look at a second-order directional derivative that was used to analyze second-order necessary optimality conditions of optimization problems; see [5]. ψ is said to be second-order directionally differentiable at a point $\bar{x}$ if both limit (2.1) and

$$\psi''(\bar{x},d,w) := \lim_{t\downarrow 0} \frac{\psi\left(\bar{x}+td+\frac{1}{2}t^2w\right) - \psi(\bar{x}) - t\psi'(\bar{x},d)}{\frac{1}{2}t^2} \tag{2.2}$$

exist for each choice of $d,w \in \mathbb{R}^n$. If this limit exists, it is called second-order directional derivative of ψ at the point $\bar{x}$ with respect to the directions d and w. Note that, for $w = 0$, definition (2.2) coincides with the proposal for a second directional derivative, due to Demyanov and Pevnyi [11]

$$\psi''(\bar{x},d) := \lim_{t\downarrow 0} \frac{\psi(\bar{x}+td) - \psi(\bar{x}) - t\psi'(\bar{x},d)}{\frac{1}{2}t^2}.$$

Finally, we introduce the Páles and Zeidan's [26] second-order upper generalized directional derivative of ψ at $\bar{x}$ in direction $d \in \mathbb{R}^n$ by

$$\psi^{\circ\circ}(\bar{x},d) = \limsup_{t\downarrow 0} 2\frac{\psi(\bar{x}+td) - \psi(\bar{x}) - t\psi^\circ(\bar{x},d)}{t^2}.$$

2.3. **Partial calmness and exact penalization.** In the following section, we present uniform parametric error bounds and demonstrate their usefulness in obtaining exact penalty formulations for a general optimization problem. Consider the following optimization problem

$$\min_x \{h(x) \mid \alpha(x) = 0,\ \beta(x) \le 0,\ x \in C\}, \tag{2.3}$$

where $h : \mathbb{R}^n \to \mathbb{R}$, $\alpha : \mathbb{R}^n \to \mathbb{R}$, $\beta : \mathbb{R}^n \to \mathbb{R}^m$, and C is a closed subset of $\mathbb{R}^n$. We assume that both α and β are lower semicontinuous. The associated partially perturbed problem can be formulated as:

$$\min_x \{h(x) \mid \alpha(x) = \varepsilon,\ \beta(x) \le 0,\ x \in C\}, \tag{2.4}$$

where $\varepsilon \in \mathbb{R}$. The following definition was initially presented in [30].

Definition 2.1. (Partial calmness) Let $\bar{x}$ solve (2.3). We say that (2.3) satisfies the partial calmness property at $\bar{x}$ if there exist positive constants κ, δ such that, for all $\varepsilon \in \delta\mathbb{B}$ and all $x \in \bar{x} + \delta\mathbb{B}$ that are feasible for (2.4), the following inequality holds:

$$h(x) - h(\bar{x}) + \kappa|\alpha(x)| \ge 0,$$

where $\mathbb{B}$ denotes the open unit ball in $\mathbb{R}^n$.

The condition of partial calmness differs from the calmness condition introduced by Clarke and Rockafellar (see, for example, [7]), in which only the equality constraint $\alpha(x)=0$ is perturbed. Calmness was shown to be closely related to the concept of ”exact penalization” in [7, Prop. 6.4.3]. Specifically, if $\bar{x}$ is a local solution to (2.3) and problem (2.3) is calm at $\bar{x}$, then $\bar{x}$ is a local solution to a penalized problem.

The following proposition demonstrates that the concept of partial calmness is equivalent to local exact penalization.

Proposition 2.2. [12, Lemma 5] *Under the assumptions that h is continuous at $\bar{x}$, and $\bar{x}$ is a local minimum to (2.3) and (2.3) is partially calm at $\bar{x}$, there exists $\kappa^*>0$ such that, for all $\kappa \geq \kappa^*$, $\bar{x}$ is a local minimum to the following penalized problem*

$$\begin{cases} \min\limits_{x} h(x)+\kappa|\alpha(x)| \\ \beta(x) \leq 0 \\ x \in C. \end{cases} \tag{2.5}$$

Moreover, any local minima of (2.5) with $\kappa > \kappa^*$ in a neighborhood of $\bar{x}$, where $\bar{x}$ is a local minimum, is also local minima to (2.3).

2.4. **R-regularity.** Let $v \in \mathbb{R}^m$ and $D \subset \mathbb{R}^m$. We can represent $\rho(v,D)$ as the smallest distance between v and any point y in the set D, i.e.,

$$\rho(v,D)=\inf_{y\in D}|v-y|,$$

where $|\cdot|$ denotes the Euclidean norm.

Let $\Gamma:\mathbb{R}^n \rightrightarrows \mathbb{R}^m$ be a multi-valued mapping, and let $z_0=(x_0,y_0)\in \text{gph}\Gamma$.

- Γ is said to be R-regular at z_0 if there exist numbers $\alpha>0, \delta_1>0$, and $\delta_2>0$ such that
$$\rho(y,\Gamma(x)) \leq \alpha \max\{0, g_i(x,y)\}$$
for all $x \in x_0+\delta_1\mathbb{B}$ and $y \in y_0+\delta_2\mathbb{B}$.
- The set $\Gamma(x_0)$ is said to be R-regular at y_0 if there exist numbers $\alpha>0$ and $\delta>0$ such that
$$\rho(y,\Gamma(x_0)) \leq \alpha \max\{0, g_i(x_0,y)\},$$
for all $y \in y_0+\delta\mathbb{B}$.

Remark 2.1.
- It should be noted that if Γ satisfies R-regularity for $z=(x,y)\in \text{gph}\Gamma$, it is necessary that x belongs to the interior of $dom\Gamma$.
- It is known that if the Mangasarian-Fromovitz regularity condition holds at a point $z=(x,y)\in\text{gph}\Gamma$, then the R-regularity of the multi-valued mapping Γ also holds at that point. However, the converse is not generally true; see [25].

Among the various Lipschitz-type properties of multi-valued mappings, two important ones are pseudo-Lipschitz continuity [2] and upper pseudo- Lipschitz continuity (also known as calmness) [17, 18].

The multi-function Γ can be characterized by several Lipschitz-type properties:

- Γ is pseudo-Lipschitz at z_0 if there exist neighborhoods V and W of x_0 and y_0, and a number $l>0$, such that
$$\Gamma(x_1)\cap W \subset \Gamma(x_2)+l\,|x_2-x_1|\,\mathbb{B},$$
for all $x_1,x_2 \in V$.

- Γ is upper pseudo-Lipschitz (or calm) at z_0 if there exist neighborhoods V and W of x_0 and y_0, and a number $l > 0$, such that

$$\Gamma(x) \cap W \subset \Gamma(x_0) + l\,|x - x_0|\,\mathbb{B},$$

for all $x \in V$.

- Γ is locally upper Lipschitz at $z_0 = (x_0, y_0)$ if there exist neighborhoods V and W of x_0 and y_0, and a number $l > 0$, such that

$$\Gamma(x) \cap W \subseteq y_0 + l\,|x - x_0|\,\mathbb{B},$$

for all $x \in V$.

It is clearly seen that any pseudo-Lipschitz multifunction is also upper pseudo-Lipschitz. However, pseudo-lipschitz does not implies, in general, local upper Lipschitzity; see [21] for a counter example. The pseudo-Lipschitz property is also known as the local Lipschitz-like property or the Aubin property of multifunctions [20].

3. Second-order Necessary Optimality Conditions via ($LLVF$) Reformulation

The aim of this section is to derive primal and dual second-order necessary optimality conditions for a point $(\bar{x}, \bar{y})$ to be a local solution for (BOP) with nonsmooth data. To proceed, we employ the optimal value function $\varphi : \mathbb{R}^n \to \mathbb{R}$ defined by

$$\forall x \in \mathbb{R}^n : \quad \varphi(x) := \inf_y \{f(x,y) \mid g(x,y) \leq 0\},$$

which is associated with the lower-level problem ($P[x]$) with the set of optimal solutions given in (1.1). Exploiting the lower-level value function φ, we obtain the classical equivalent single-level problem of (BOP):

$$\min_{x,y} \{F(x,y) \mid G(x,y) \leq 0, f(x,y) - \varphi(x) \leq 0, g(x,y) \leq 0\}. \tag{$LLVF$}$$

Typically, optimization problem ($LLVF$) does not satisfy the standard constraint qualifications in nonsmooth programming, and it is not regular at all feasible points.

In order to establish necessary optimality conditions for (BOP), we employ a partial calmness condition for nonsmooth optimization problem ($LLVF$). This condition penalizes the challenging constraint $f(x,y) - \varphi(x) \leq 0$ by adding it to the objective function. To this end, we study the penalized problem

$$\begin{cases} \min\limits_{x,y} \kappa F(x,y) + f(x,y) - \varphi(x) \\ G(x,y) \leq 0 \\ g(x,y) \leq 0. \end{cases} \tag{LLVF[κ]}$$

Functions F, $\left(G_i, i \in \bar{I}^G\right)$ and $\left(g_j, j \in \bar{I}^g\right)$ are supposed to be locally Lipschitz near $(\bar{x}, \bar{y})$. Letting $\bar{z} = (\bar{x}, \bar{y}) \in \Omega$ and $d = (d_1, d_2) \in \mathbb{R}^n \times \mathbb{R}^m$, we consider the sets

$$I_0^G(\bar{z}, d) = \{i \in \bar{I}^G : G_i^\circ(\bar{z}, d) = 0\}, \text{ and}$$

$$I_0^g(\bar{z}, d) = \{j \in \bar{I}^g : g_j^\circ(\bar{z}, d) = 0\}.$$

Finally, we define the critical cone (set of critical directions) at the point $\bar{z} \in \Omega$ as

$$\mathscr{C}(\bar{z}) = \left\{ d = (d_1, d_2) \in \mathbb{R}^n \times \mathbb{R}^m \;\middle|\; \begin{array}{rl} \kappa F^\circ(\bar{z}, d) + f^\circ(\bar{z}, d) - \varphi'(\bar{x}, d_1) & \leq 0 \\ G_i^\circ(\bar{z}, d) & \leq 0 \ , \ i \in \bar{I}^G \\ g_j^\circ(\bar{z}, d) & \leq 0 \ , \ j \in \bar{I}^g \end{array} \right\}.$$

We are now in a position to state the initial theorem, which establishes the primal second-order necessary optimality conditions for (BOP).

Theorem 3.1. *Let $\bar{z} = (\bar{x}, \bar{y})$ be a local solution to (BOP), assumed to be partially calm at $\bar{z}$. Suppose that the functions F, f, $(G_i,\ i \in \bar{I}^G)$, and $(g_j,\ j \in \bar{I}^g)$ are locally Lipschitz and the functions $(G_i,\ i \notin \bar{I}^G)$ and $(g_j,\ j \notin \bar{I}^g)$ are continuous at $\bar{z}$. Moreover, suppose that the first and second-order directional derivatives of φ exist at $\bar{x}$ and the LMFCQ condition is satisfied at every point $(\bar{x}, y) \in gph\ S$ and the set Π is non-empty and compact. Then, for every $d = (d_1, d_2) \in \mathscr{C}(\bar{z})$, the system*

$$\kappa\left(F^\circ(\bar{z}, r) + F^{\circ\circ}(\bar{z}, d)\right) + f^\circ(\bar{z}, r) + f^{\circ\circ}(\bar{z}, d) - \varphi'(\bar{x}, r_1) - \varphi''(\bar{x}, d_1) < 0 \tag{3.1}$$

$$G_i^\circ(\bar{z}, r) + G_i^{\circ\circ}(\bar{z}, d) < 0, \quad i \in I_0^G(\bar{z}, d) \tag{3.2}$$

$$g_j^\circ(\bar{z}, r) + g_j^{\circ\circ}(\bar{z}, d) < 0, \quad j \in I_0^g(\bar{z}, d) \tag{3.3}$$

has no solution $r = (r_1, r_2) \in \mathbb{R}^n \times \mathbb{R}^m$.

Proof. Let $\bar{z} = (\bar{x}, \bar{y})$ be a local minimum solution to (BOP) at which (BOP) is partially calm. Then, from Proposition 2.2, there exists $\kappa > 0$ such that $\bar{z}$ is a local minimum to (LLVF[κ]). Since Π is non-empty and compact and *LMFCQ* condition is satisfied at every point $(\bar{x}, y) \in$ gph S, then one sees from [8, Theorem 4.14] that φ is locally Lipschitz continuous at $\bar{x}$.

To establish the conclusion of the theorem, we assume by contradiction that there exists a critical direction $d = (d_1, d_2) \in \mathscr{C}(\bar{z})$ such that system (3.1)-(3.2)-(3.3) has a solution $r = (r_1, r_2) \in \mathbb{R}^n \times \mathbb{R}^m$. The proof consists of several steps.

Step 1: Let us prove that there exists $\eta > 0$ such that $\bar{z} + td + \frac{1}{2}t^2 r$ is a feasible point to (LLVF[κ]) for all $t \in [0, \eta)$.

To proceed, three cases will be considered for both functions G and g:

(a) For every $i \in I \backslash \bar{I}^G$, we have $G_i(\bar{z}) < 0$. Hence, by continuity, there exists $\eta_1^i > 0$ such that $G_i\left(\bar{z} + td + \frac{1}{2}t^2 r\right) < 0$, for all $t \in \left[0, \eta_1^i\right)$.

(b) For every $i \in \bar{I}^G \backslash I_0^G(\bar{z}, d)$, we have $G_i^\circ(\bar{z}, d) < 0$. We need to show that there exists $\eta_2^i > 0$ such that, for all $t \in \left[0, \eta_2^i\right)$,

$$G_i\left(\bar{z} + td + \frac{1}{2}t^2 r\right) < G_i(\bar{z}) = 0.$$

Suppose by contradiction that, for any $\eta > 0$, there exists $0 < t_\eta < \eta$ such that

$$G_i\left(\bar{z} + t_\eta d + \frac{1}{2}t_\eta^2 r\right) \geq G_i(\bar{z}).$$

Let $\eta_n > 0$ be a sequence convergent to 0 as $n \to \infty$ and $t_n \in [0, \eta_n)$ such that

$$G_i\left(\bar{z} + t_n d + \frac{1}{2}t_n^2 r\right) - G_i(\bar{z}) \geq 0.$$

To simplify the writing, we set in the following $u_n^\eta = \bar{z} + t_n d$ and $v_n^\eta = \bar{z} + t_n d + \frac{1}{2}t_n^2 r$. The last inequality yields

$$\begin{aligned} 0 &\leq \limsup_{t_n \to 0} \frac{G_i\left(v_n^\eta\right) - G_i(\bar{z})}{t_n} \\ &\leq \limsup_{t_n \to 0} \frac{1}{t_n}\left[G_i\left(u_n^\eta\right) - G_i(\bar{z})\right] + \limsup_{t_n \to 0} \frac{1}{t_n}\left[G_i\left(v_n^\eta\right) - G_i\left(u_n^\eta\right)\right] \\ &\leq G_i^\circ(\bar{z}, d) + \limsup_{t_n \to 0} L_i t_n \|r\| \end{aligned}$$

as G_i is locally Lipschitz of constant $L_i > 0$. Thus $G_i^\circ(\bar{z}, d) \geq 0$, which contradicts the fact that $i \notin I_0^G(\bar{z}, d)$.

(c) Let $i \in I_0^G(\bar{z}, d)$. Since $G_i^\circ(\bar{z}, d) = 0$, there exist $\eta_3^i > 0$ such that, for all $t \in \left[0, \eta_3^i\right)$,

$$G_i\left(\bar{z} + td + \frac{1}{2}t^2 r\right) < G_i(\bar{z}) = 0.$$

To demonstrate this, let us suppose for the sake of contradiction that, for every $\eta > 0$, there exists some $0 < t_\eta < \eta$ such that

$$G_i\left(\bar{z} + t_\eta d + \frac{1}{2}t_\eta^2 r\right) \geq G_i(\bar{z}) = 0.$$

Let $\eta_n > 0$ be a sequence convergent to 0 as $n \to \infty$ and $t_n \in [0, \eta_n)$. Then,

$$\begin{aligned} 0 &\leq G_i\left(v_n^\eta\right) - G_i(\bar{z}) \\ &= \tfrac{t_n^2}{2}\left[\tfrac{2}{t_n^2}\left(G_i\left(v_n^\eta\right) - G_i\left(u_n^\eta\right)\right)\right] + \tfrac{t_n^2}{2}\left[\tfrac{2}{t_n^2}\left(G_i\left(u_n^\eta\right) - G_i(\bar{z}) - t_n G_i^\circ(\bar{z}, d)\right)\right]. \end{aligned}$$

We divide the inequality above by $t_n^2/2$ and we take the upper limit as $t_n \to 0^+$. Then

$$0 \leq G_i^\circ(\bar{z}, r) + G_i^{\circ\circ}(\bar{z}, d).$$

This contradicts the assumption that r is a solution to system (3.2).

Consequently, setting $\eta^1 = \min_{i \in I} \eta_1^i$, for all $i \in I$, we have $G_i\left(\bar{z} + td + \frac{1}{2}t^2 r\right) \leq 0$ for all $t \in \left[0, \eta^1\right)$. As in the case (a)-(c), we obtain the existence of $\eta^2 = \min_{j \in J} \eta_2^j$ such that, for all $j \in J$ and for all $t \in \left[0, \eta^2\right)$, $g_j(\bar{z} + td + \frac{1}{2}t^2 r) < 0$.

Finally, there exists $\eta > 0$ with $\eta = \min\{\eta^1, \eta^2\}$ such that $\bar{z} + td + \frac{1}{2}t^2 r$ is a feasible point of (LLVF[κ]) for all $t \in [0, \eta)$.

Step 2: We show that there exists $\nu > 0$, for all $t \in [0, \nu)$,

$$\kappa F(\bar{z} + td + \frac{1}{2}t^2 r) + f(\bar{z} + td + \frac{1}{2}t^2 r) - \varphi(\bar{x} + td_1 + \frac{1}{2}t^2 r_1) < \kappa F(\bar{z}) + f(\bar{z}) - \varphi(\bar{x}).$$

We proceed in two cases.

(a) $\kappa F^\circ(\bar{z}, d) + f^\circ(\bar{z}, d) - \varphi'(\bar{x}, d_1) < 0$.

Suppose by contradiction that, for any $\nu > 0$, there exists $t_\nu \in [0, \nu)$ such that

$$\kappa F\left(\bar{z} + t_\nu d + \frac{1}{2}t_\nu^2 r\right) + f\left(\bar{z} + t_\nu d + \frac{1}{2}t_\nu^2 r\right) - \varphi\left(\bar{x} + t_\nu d_1 + \frac{1}{2}t_\nu^2 r_1\right) \geq \kappa F(\bar{z}) + f(\bar{z}) - \varphi(\bar{x}).$$

Let $\nu_n > 0$ be a sequence convergent to 0 as $n \to \infty$ and $t_n \in [0, \nu_n)$ such that

$$\kappa F\left(\bar{z}+t_nd+\frac{1}{2}t_n^2r\right)+f\left(\bar{z}+t_nd+\frac{1}{2}t_n^2r\right)-\varphi\left(\bar{x}+t_\nu d_1+\frac{1}{2}t_n^2r_1\right) \geq \kappa F(\bar{z})+f(\bar{z})-\varphi(\bar{x}).$$

For each n, we set $u_n^\nu = \bar{z}+t_nd$, $v_n^\nu = \bar{z}+t_nd+\frac{1}{2}t_n^2r$, $a_n^\nu = \bar{x}+t_nd_1$, and $b_n^\nu = \bar{x}+t_nd_1+\frac{1}{2}t_n^2r_1$. It follows that

$$\begin{aligned}0 &\leq \limsup_{t_n\to 0}\left(\kappa\frac{F(v_n^\nu)-F(\bar{z})}{t_n}+\frac{f(v_n^\nu)-f(\bar{z})}{t_n}-\frac{\varphi(b_n^\nu)-\varphi(\bar{x})}{t_n}\right)\\ &\leq \kappa\limsup_{t_n\to 0}\frac{F(v_n^\nu)-F(\bar{z})}{t_n}+\limsup_{t_n\to 0}\frac{f(v_n^\nu)-f(\bar{z})}{t_n}-\liminf_{t_n\to 0}\frac{\varphi(b_n^\nu)-\varphi(\bar{x})}{t_n}.\end{aligned}$$

Inserting $\pm t_n^{-1}F(u_n^\nu)$, $\pm t_n^{-1}f(u_n^\nu)$, and $\pm t_n^{-1}\varphi(a_n^\nu)$ in the last inequality, we obtain

$$\begin{aligned}0 \leq\ &\kappa\limsup_{t_n\to 0}\frac{F(u_n^\nu)-F(\bar{z})}{t_n}+\kappa\limsup_{t_n\to 0}\frac{F(v_n^\nu)-F(u_n^\nu)}{t_n}\\ &+\limsup_{t_n\to 0}\frac{f(u_n^\nu)-f(\bar{z})}{t_n}+\limsup_{t_n\to 0}\frac{f(v_n^\nu)-f(u_n^\nu)}{t_n}\\ &-\liminf_{t_n\to 0}\frac{\varphi(a_n^\nu)-\varphi(\bar{x})}{t_n}-\liminf_{t_n\to 0}\frac{\varphi(b_n^\nu)-\varphi(a_n^\nu)}{t_n}.\end{aligned}$$

Using the definition of various directional derivatives, we have

$$0 \leq \kappa F^\circ(\bar{z},d)+f^\circ(\bar{z},d)-\varphi'(\bar{x},d_1)+\limsup_{t_n\to 0}L_Ft_n\|r\|+\limsup_{t_n\to 0}L_ft_n\|r\|+\limsup_{t_n\to 0}L_\varphi t_n\|r_1\|,$$

as F, f, and φ are locally Lipschitz of constant $L_F > 0$, $L_f > 0$, and $L_\varphi > 0$ respectively.

Thus $\kappa F^\circ(\bar{z},d)+f^\circ(\bar{z},d)-\varphi'(\bar{x},d_1) \geq 0$, which contradicts the previously assumed fact that $\kappa F^\circ(\bar{z},d)+f^\circ(\bar{z},d)-\varphi'(\bar{x},d_1) < 0$.

(b) $\kappa F^\circ(\bar{z},d)+f^\circ(\bar{z},d)-\varphi'(\bar{x},d_1) = 0$. Using the same arguments as in **(a)**, one obtains

$$\begin{aligned}0 &\leq \kappa(F(v_n^\nu)-F(\bar{z}))+f(v_n^\nu)-f(\bar{z})-\varphi(b_n^\nu)+\varphi(\bar{x})\\ &= \kappa(F(v_n^\nu)-F(u_n^\nu))+\kappa(F(v_n^\nu)-\kappa F(\bar{z})-F^\circ(\bar{z},d))\\ &\quad+(f(v_n^\nu)-f(u_n^\nu))+(f(u_n^\nu)-f(\bar{z})-t_nf^\circ(\bar{z},d))\\ &\quad-(\varphi(b_n^\nu)-\varphi(a_n^\nu))-\left(\varphi(a_n^\nu)-\varphi(\bar{x})-t_n\varphi'(\bar{x},d_1)\right).\end{aligned}$$

We divide the inequality above by $t_n^2/2$. Take the limit as $t_n \to 0^+$, we see that

$$0 \leq \kappa(F^\circ(\bar{z},r)+F^{\circ\circ}(\bar{z},d))+f^\circ(\bar{z},r)+f^{\circ\circ}(\bar{z},d)-\varphi'(\bar{x},r_1)-\varphi''(\bar{x},d_1).$$

This is in contradiction with the assumption that r satisfies system (3.1).

Step 3: In conclusion, taking $\varepsilon = \min\{\eta, \nu\}$, we see that, for all $t \in [0, \varepsilon)$,

$$\begin{aligned}\kappa F\left(\bar{z}+td+\frac{1}{2}t^2r\right)+f\left(\bar{z}+td+\frac{1}{2}t^2r\right)-\varphi\left(\bar{x}+td_1+\frac{1}{2}t^2r_1\right) &< \kappa F(\bar{z})+f(\bar{z})-\varphi(\bar{x})\\ G_i\left(\bar{z}+td+\frac{1}{2}t^2r\right) &< 0,\ \text{for all } i \in I\\ g_j\left(\bar{z}+td+\frac{1}{2}t^2r\right) &< 0,\ \text{for all } j \in J.\end{aligned}$$

This implies that $\bar{z}$ cannot be a local minimizer to problem (LLVF[κ]). Contradiction. □

Remark 3.1. It is worth noting that Theorem 3.1 does not require the differentiability of value function φ. The only requirement on φ is that it is locally Lipschitz continuous at $\bar{x}$ and that its first and second-order directional derivatives exist at $\bar{x}$. While differentiability of φ may be important in some specific cases, it does not affect the validity of the theorem. In fact, under certain regularity conditions, the estimations of the first and second-order directional derivatives of φ can be obtained. In the following, we provide several corollaries derived from Theorem 3.1, which involve an approximation of the first directional derivative of φ using the initial problem data.

Assuming in addition to the conditions of Theorem 3.1 that *LLICQ* holds at all points $(\bar{x}, y) \in$ gph S, it is possible to obtain an upper bound for the directional derivative of φ as shown in [14, corollary 4.4] by Gauvin and Dubeau. Note that condition *LMFCQ* is implicitly satisfied by adding condition *LLICQ*. Based on this bound, the following corollary can be easily derived.

Corollary 3.1. *Let $\bar{z} = (\bar{x}, \bar{y})$ be a local solution to problem (BOP), assumed to be partially calm at $\bar{z}$. Suppose that the functions F, f, $(G_i,\ i \in \bar{I}^G)$, and $(g_j,\ j \in \bar{I}^g)$ are locally Lipschitz and the functions $(G_i,\ i \notin \bar{I}^G)$ and $(g_j,\ j \notin \bar{I}^g)$ are continuous at $\bar{z}$. Moreover, suppose that the second-order directional derivative of φ exists at $\bar{x}$ and the conditions LLICQ is satisfied at every point $(\bar{x}, y) \in gph\ S$ and the set Π is non-empty and compact. Then, for every $d = (d_1, d_2) \in \mathscr{C}(\bar{z})$, there is no $r = (r_1, r_2) \in \mathbb{R}^n \times \mathbb{R}^m$, which solves system (3.2), (3.3), and*

$$\kappa\left(F^{\circ}(\bar{z}, r) + F^{\circ\circ}(\bar{z}, d)\right) + \nabla f(\bar{z})^T r + f^{\circ\circ}(\bar{z}, d) - \min_{y \in S(\bar{x})} \langle \nabla_x \ell(\bar{x}, y, w), r_1 \rangle - \varphi''(\bar{x}, d_1) < 0.$$

If functions f and $(g_j,\ j \in J)$ are convex with respect to the variable y for each value of x, and the lower-level problem is convex, then the assumptions of Theorem 3.1 ensure that φ is directionally differentiable; see [8, Theorem 4.16]. The existing formula for the directional derivative produces slightly superior outcomes compared to the one presented in Corollary 3.1.

Corollary 3.2. *Let $\bar{z} = (\bar{x}, \bar{y})$ be a local solution to problem (BOP), assumed to be partially calm at $\bar{z}$. Suppose that the functions F, f, $(G_i,\ i \in \bar{I}^G)$, and $(g_j,\ j \in \bar{I}^g)$ are locally Lipschitz and the functions $(G_i,\ i \notin \bar{I}^G)$, and $(g_j,\ j \notin \bar{I}^g)$ are continuous at $(\bar{x}, \bar{y})$. Moreover, suppose that the second-order directional derivative of φ exist at $\bar{x}$ and the LMFCQ condition is satisfied at every point $(\bar{x}, y) \in gph\ S$ and the set Π is non-empty and compact and the lower-level problem ($P[x]$) is convex with respect to y. Then, for every $d = (d_1, d_2) \in \mathscr{C}(\bar{z})$, there is no $r = (r_1, r_2) \in \mathbb{R}^n \times \mathbb{R}^m$ which solves system (3.2), (3.3) and*

$$\kappa\left(F^{\circ}(\bar{z}, r) + F^{\circ\circ}(\bar{z}, d)\right) + \nabla f(\bar{z})^T r + f^{\circ\circ}(\bar{z}, d) - \min_{y \in S(\bar{x})} \max_{w \in \Lambda(\bar{x}, y)} \langle \nabla_x \ell(\bar{x}, y, w), r_1 \rangle - \varphi''(\bar{x}, d_1) < 0.$$

According to [8, Corollary 4.7], if the lower-level problem is fully convex, it is unnecessary to take into account the minimum over all lower-level solutions that are associated with the reference point, which makes the situation even more favorable.

Corollary 3.3. *Let $\bar{z} = (\bar{x}, \bar{y})$ be a local solution to problem (BOP), assumed to be partially calm at $\bar{z}$. Suppose that the functions F, f, $(G_i,\ i \in \bar{I}^G)$, and $(g_j,\ j \in \bar{I}^g)$ are locally Lipschitz and the functions $(G_i,\ i \notin \bar{I}^G)$ and $(g_j,\ j \notin \bar{I}^g)$ are continuous at $\bar{z}$. Moreover, suppose that the second-order directional derivatives of φ exist at $\bar{x}$ and the LMFCQ condition is satisfied at every point $(\bar{x}, y) \in gph\ S$ and the set Π is non-empty and compact and the problem ($P[x]$) is convex with respect* *to* (x,y). *Then, for every* $d=(d_1,d_2)\in\mathscr{C}(\bar{z})$, *there is no* $r=(r_1,r_2)\in\mathbb{R}^n\times\mathbb{R}^m$ *which solves system* (3.2), (3.3), *and*

$$\kappa\left(F^{\circ}(\bar{z},r)+F^{\circ\circ}(\bar{z},d)\right)+\nabla f(\bar{z})^T r+f^{\circ\circ}(\bar{z},d)-\max_{w\in\Lambda(\bar{z})} w^T\nabla_x g(\bar{x},\bar{y})r_1-\varphi^{''}(\bar{x},d_1)<0.$$

The first result about the existence of the second-order directional derivative of the value function φ was obtained in [10] for the case when problem ($P[x]$) is convex with respect to y, the matrix $\nabla^2_{yy}f(\bar{x},\bar{y})$ is positively definite, and $Y(x)$ satisfies Slater regularity condition. In [28], it was shown that under the *LMFCQ* regularity condition at all points $\bar{z}=(\bar{x},\bar{y})\in\{\bar{x}\}\times S(\bar{x})$ the first and second-order derivatives of φ exist at $\bar{x}$ if *LSOSC* holds. However, these works almost always utilized various second-order sufficient optimality conditions and, consequently, only studied the problems with $S(\bar{x})=\{\bar{y}\}$.

In the following, we use first-order and second-order Hadamard derivative estimates of the value function φ established by Minchenko and Tarakanov [25, Theorem 3.1 and Theorem 4.1] with non-single-valued solutions. For this purpose, let $\bar{x}\in\mathbb{R}^n$, $d_1\in\mathbb{R}^n$, and assume

$\mathscr{H}(1)$**:** Y is R-regular;
$\mathscr{H}(2)$**:** the solution map S is upper pseudo-Lipschitz at all points of $\{\bar{x}\}\times S(\bar{x})$;
$\mathscr{H}(3)$**:** the set $S(\bar{x},d_1)$ is non-empty;
$\mathscr{H}(4)$**:** the functions f, and $(g_j,\ j\in J)$ are $\mathscr{C}^2$-differentiable.

Here,

$$S(\bar{x},d_1)=\{(y,d_2)\mid y\in S(\bar{x}),d_2\in\Upsilon((\bar{x},y),d_1),\varphi'(\bar{x},d_1)=\langle\nabla f(\bar{x},y),d\rangle\},$$

and

$$\Upsilon(x,y,d_1)=\{d_2\in\mathbb{R}^m\mid\langle\nabla g_i(x,y),(d_1,d_2)\rangle\geq 0\}.$$

From [25, Lemma 2.3, Theorem 3.1 and Theorem 4.1], assuming that $\mathscr{H}(1)$-$\mathscr{H}(4)$ are satisfied, we obtain that the value function φ is Lipschitzian in some neighbourhood of $\bar{x}$ and we obtain the following estimates for the first and second-order directional derivatives of φ at $\bar{x}$ in a direction $d_1\in\mathbb{R}^n$

$$\varphi'(\bar{x},d_1)=\inf_{\bar{y}\in S(\bar{x})}\max_{w\in\Lambda(\bar{x},\bar{y})}\langle\nabla_x\ell(\bar{x},\bar{y},w),d_1\rangle,$$

and

$$\varphi^{''}(\bar{x},d_1)=\inf_{(\bar{y},d_2)\in S(\bar{x},d_1)}\max_{w\in\Lambda^2(\bar{x},\bar{y},d_1)}\langle d,\nabla^2\ell(\bar{x},\bar{y},w)d\rangle.$$

Here, $d=(d_1,d_2)$. Using this estimates, the following theorem follows easily.

Theorem 3.2. *Let* $\bar{z}=(\bar{x},\bar{y})$ *be a local solution to problem* (BOP), *assumed to be partially calm at* $\bar{z}$. *Suppose that the functions* F, f, $(G_i,\ i\in\bar{I}^G)$ *and* $(g_j,\ j\in\bar{I}^g)$ *are locally Lipschitz and the functions* $(G_i,\ i\notin\bar{I}^G)$ *and* $(g_i,\ i\notin\bar{I}^g)$ *are continuous at* $(\bar{x},\bar{y})$. *Moreover, suppose that assumptions* $\mathscr{H}(1)$-$\mathscr{H}(4)$ *hold. Then, for every* $d=(d_1,d_2)\in\mathscr{C}(\bar{z})$, *there is no* $r=(r_1,r_2)\in\mathbb{R}^n\times\mathbb{R}^m$ *which solves system* (3.2)-(3.3) *and*

$$\kappa\left(F^{\circ}(\bar{z},r)+F^{\circ\circ}(\bar{z},d)\right)+f^{\circ}(\bar{z},r)+f^{\circ\circ}(\bar{z},d)-\inf_{\bar{y}\in S(\bar{x})}\max_{w\in\Lambda(\bar{z})}\langle\nabla_x\ell(\bar{z},w),r_1\rangle$$

$$-\inf_{(\bar{y},d_2)\in S(\bar{x},d_1)}\max_{w\in\Lambda^2(\bar{z},d_1)}\langle d,\nabla^2\ell(\bar{z},w)d\rangle<0.$$

Example 3.1. Let us consider the following bilevel optimization problem:

$$\min_{x,y}\ \{F(x,y)=|x|+|y| \mid G_1(x,y)=x-1\le 0,\ G_2(x,y)=-x-1\le 0,\ y\in S(x)\}. \qquad \text{(BOP)}$$

For every $x\in\mathbb{R}$, let $S(x)$ denote the set of optimal solution mapping of the problem $(P[x])$

$$\min_{y}\ \{f(x,y)=x^2+y^2 \mid g_1(x,y)=y^2-y\le 0,\ g_2(x,y)=x-y\le 0,\ g_3(x,y)=-x-y\le 0\}.$$

On the one hand, it can be observed that $(0,0)$ is the only global optimum to (BOP) with $S(x)=\{|x|\}$ and

$$\varphi(x)=\inf_{y}\ \{f(x,y)|g(x,y)\le 0\}=2x^2.$$

The penalized problem reformulation of (BOP) is given by:

$$\begin{cases}\min\limits_{x,y}\ \kappa F(x,y)+f(x,y)-\varphi(x)\\ G_1(x,y)=x-1\le 0\\ G_2(x,y)=-x-1\le 0\\ g_1(x,y)=y^2-y\le 0\\ g_2(x,y)=x-y\le 0\\ g_3(x,y)=-x-y\le 0,\end{cases} \qquad \text{(LLVF}[\kappa])$$

where $\kappa>0$.

We now show that $\bar z=(\bar x,\bar y)=(0,0)$ verifies the necessary conditions established in Theorem 3.1. We have $\bar I^g=\{1,2,3\}$. Since $S(x)=\{|x|\}$, we have $(\bar x,y)\in$ gph S. This implies $(\bar x,y)=(0,0)$ and for $d\in\mathbb{R}$, one has

$$\nabla_y g_1(0,0)^T d=-d<0\ ,\ \ \nabla_y g_2(0,0)^T d=-d<0\ \text{ and }\ \nabla_y g_3(0,0)^T d=-d<0.$$

Hence, there is $d\in\mathbb{R}^+$ verifying $(\nabla_y g_j(\bar x,\bar y)^T d<0\ ,\ \forall j\in\bar I^g)$. Thus, *LMFCQ* holds at $\bar z=(0,0)$. We also have $\Pi=\{(x,y)\in\mathbb{R}^2 \mid g(x,y)\le 0\}=[-1,1]\times[0,1]$ is non-empty and compact.

On the other hand, we have $\bar I^G=\emptyset$ and $\bar I^g=\{1,2,3\}$. Let $v=(v_1,v_2)\in\mathbb{R}^2$. Then

$$\kappa F^\circ(\bar z,v)+f^\circ(\bar z,v)-\varphi'(\bar x,v_1)=\kappa|v_1|+\kappa|v_2|$$

and

$$g_1^\circ(\bar z,v)=-v_2,\ g_2^\circ(\bar z,v)=v_1-v_2\ \text{ and }\ g_3^\circ(\bar z,v)=-v_1-v_2.$$

Let $d=(d_1,d_2)\in\mathscr{C}(\bar z)$. It follows that

$$\begin{cases}\kappa|d_1|+\kappa|d_2|\le 0\\ -d_2\le 0\\ d_1-d_2\le 0\\ -d_1-d_2\le 0.\end{cases}$$

Thus $d=(0,0)$ is the unique critical direction. As a result

$$I_0^G(\bar z,d)=\emptyset\ \text{ and }\ I_0^g(\bar z,d)=\{1,2,3\},$$

and consequently

$$\kappa F^{\circ\circ}(\bar z,v)+f^{\circ\circ}(\bar z,v)-\varphi''(\bar x,v_1)=2v_2-2v_1,\ g_1^{\circ\circ}(\bar z,v)=2v_2,\ g_2^{\circ\circ}(\bar z,v)=0\ \text{ and }\ g_1^{\circ\circ}(\bar z,v)=0.$$

Now, we analyze the following system

$$\begin{cases} \kappa\left(F^{\circ}(\bar{z},r)+F^{\circ\circ}(\bar{z},d)\right)+f^{\circ}(\bar{z},r)+f^{\circ\circ}(\bar{z},d)-\varphi^{'}(\bar{x},r_1)-\varphi^{''}(\bar{x},d_1)=\kappa|r_1|+\kappa|r_2|<0 \\ g_1^{\circ}(\bar{z},r)+g_1^{\circ\circ}(\bar{z},d)=-r_2<0 \\ g_2^{\circ}(\bar{z},r)+g_2^{\circ\circ}(\bar{z},d)=r_1-r_2<0 \\ g_3^{\circ}(\bar{z},r)+g_3^{\circ\circ}(\bar{z},d)=-r_1-r_2<0 \end{cases}$$

Since $\kappa>0$, the above system has no solution. This proves that $\bar{z}=(0,0)$ is a candidate of being the optimum solution of problem (BOP) according to Theorem 3.1.

Remark 3.2. It is worth mentioning that, in general, the directional derivative $\varphi'(x,.)$ is neither convex nor lower semi-continuous. Deriving the dual necessary conditions via linear duality technique seems to be impossible, even if the directional derivative of φ is sub-additive (which is the case for regular functions). However, there are some special cases where φ is not only continuously differentiable but also second-order directionally differentiable.

The following proposition assumes that the lower-level solution being considered is unique and is determined through a standard second-order sufficient condition and convexity.

Proposition 3.1. [24, Proposition 4] *Let $(\bar{x},\bar{y})\in\Omega$, where LLICQ and LSOSC hold, and let $\overline{w}\in\mathbb{R}^q$ be unique lower-level Lagrange multiplier. Assume that, for each $x\in\mathbb{R}^n$, $f(x,\cdot):\mathbb{R}^n\to\mathbb{R}$ is convex and $g(x,\cdot):\mathbb{R}^n\to\mathbb{R}^q$ is componentwise convex. Then, the following assertions hold:*

(i) *The optimal value function φ is continuously differentiable at $\bar{x}$, and it holds*

$$\nabla\varphi(\bar{x})=\nabla_x\ell(\bar{x},\bar{y},\overline{w}).$$

(ii) *φ is second-order directionally differentiable at $\bar{x}$, and, for every $d_1\in\mathbb{R}^n$, it holds*

$$\varphi''(\bar{x},d_1)=\inf_{d_2\in\mathbb{R}^m}\left\{d^T\nabla^2\ell(\bar{x},\bar{y},\overline{w})\,d \;\middle|\; \begin{array}{l}\nabla g_j(\bar{x},\bar{y})^T d=0,\ j\in\bar{I}^g,\ \overline{w}_j>0 \\ \nabla g_j(\bar{x},\bar{y})^T d\le 0,\ j\in\bar{I}^g,\ \overline{w}_j=0\end{array}\right\}.$$

Taking into account the conditions and the notations of the previous proposition, the critical cone becomes

$$\mathscr{C}(\bar{z})=\left\{\; d\in\mathbb{R}^n\times\mathbb{R}^m \;\middle|\; \begin{array}{rll}\kappa F^{\circ}(\bar{z},d)+f^{\circ}(\bar{z},d)-\nabla\varphi(\bar{x})^T d_1 & \le 0 & \\ G_i^{\circ}(\bar{z},d) & \le 0, & i\in\bar{I}^G \\ g_j^{\circ}(\bar{z},d) & \le 0, & j\in\bar{I}^g\end{array}\right\}.$$

In the following theorem, we establish dual second-order necessary optimality conditions in the simple case that the optimal value function φ is continuously differentiable. To proceed, we need the following reduced critical cone

$$\mathscr{RC}(\bar{z})=\left\{\; d\in\mathscr{C}(\bar{z}) \;\middle|\; \begin{array}{rl}\kappa F^{\circ\circ}(\bar{z},d)+f^{\circ\circ}(\bar{z},d)-\varphi^{''}(\bar{x},d_1) & <+\infty \text{ if } \Theta(\bar{z},d)=0 \\ G_i^{\circ\circ}(\bar{z},d) & <+\infty,\ i\in I_0^G(\bar{z},d) \\ g_j^{\circ\circ}(\bar{z},d) & <+\infty,\ j\in I_0^g(\bar{z},d)\end{array}\right\}$$

with $\Theta(\bar{z},d)=\kappa F^{\circ}(\bar{z},d)+f^{\circ}(\bar{z},d)-\nabla_x\ell(\bar{x},\bar{y},\overline{w})^T d_1$.

Theorem 3.3. *Let $\bar{z}=(\bar{x},\bar{y})$ be a local solution to (BOP), assumed to be partially calm at $\bar{z}$ and where LLICQ and LSOSC hold. Assume that the functions F, f, $(G_i, i\in\bar{I}^G)$ and $(g_j, j\in\bar{I}^g)$ are locally Lipschitz, Gâteaux differentiable and regular at $\bar{z}$, and the functions $(G_i, i\notin\bar{I}^G)$ and $(g_j, j\notin\bar{I}^g)$ are continuous at $\bar{z}$. Moreover, suppose that for each $x\in\mathbb{R}^n$, $f(x,\cdot):\mathbb{R}^n\to\mathbb{R}$ is convex* *and $g(x,\cdot):\mathbb{R}^n\to\mathbb{R}^q$ is componentwise convex, and let $\overline{w}\in\mathbb{R}^q$ be the unique lower-level Lagrange multiplier. Then, for any nonzero critical direction $d=(d_1,d_2)\in\mathscr{RC}(\bar{z})$ there exist non-negative multipliers $(\lambda,\alpha,\beta)\in\mathbb{R}_+\times\mathbb{R}^p_+\times\mathbb{R}^q_+$ not all zeros with*

$$\lambda\kappa\nabla F(\bar{z})+\lambda\nabla f(\bar{z})-\lambda\nabla_x\ell(\bar{x},\bar{y},\overline{w})+\sum_{i\in\bar{I}^G}\alpha_i\nabla G_i(\bar{z})+\sum_{j\in\bar{I}^g}\beta_j\nabla g_j(\bar{z})=0, \tag{3.4}$$

$$\lambda\kappa F^{\circ\circ}(\bar{z},d)+\lambda f^{\circ\circ}(\bar{z},d)-\lambda\varphi''(\bar{x},d_1)+\sum_{i\in\bar{I}^G}\alpha_i G_i^{\circ\circ}(\bar{z},d)+\sum_{j\in\bar{I}^g}\beta_j g_j^{\circ\circ}(\bar{z},d)\geq 0, \tag{3.5}$$

$$\lambda\kappa\nabla F(\bar{z})^T d+\lambda\nabla f(\bar{z})^T d-\lambda\nabla_x\ell(\bar{x},\bar{y},\overline{w})^T d_1=0, \tag{3.6}$$

$$\alpha_i G_i(\bar{z})=0, i\in I \text{ and } \quad \beta_j g_j(\bar{z})=0,\quad j\in J, \tag{3.7}$$

$$\alpha_i\nabla G_i(\bar{z})^T d=0, i\in\bar{I}^G \text{ and } \beta_j\nabla g_j(\bar{z})^T d=0,\quad j\in\bar{I}^g. \tag{3.8}$$

Proof. By utilizing Theorem 3.1, we can deduce that there is no solution to the following system in r

$$\kappa\left(\nabla F(\bar{z})^T r+F^{\circ\circ}(\bar{z},d)\right)+\nabla f(\bar{z})^T r+f^{\circ\circ}(\bar{z},d)-\nabla_x\ell(\bar{x},\bar{y},\overline{w})^T r_1-\varphi''(\bar{x},d_1)<0$$

$$\nabla G_i(\bar{z})^T r+G_i^{\circ\circ}(\bar{z},d)<0,\quad i\in I_0^G(\bar{z},d)$$

$$\nabla g_j(\bar{z})^T r+g_j^{\circ\circ}(\bar{z},d)<0,\quad j\in I_0^g(\bar{z},d),$$

for any fixed critical direction d such that $d\in\mathscr{RC}(\bar{z})$.

By discarding inequalities from the system where the corresponding upper generalized directional derivative of the second-order is $-\infty$, we can set

$$A=\begin{pmatrix}\kappa\nabla F(\bar{z})^T+\nabla f(\bar{z})^T-\nabla_x\ell(\bar{x},\bar{y},\overline{w})^T\\ \nabla G_i(\bar{z})^T\ ,\ i\in I_0^G(\bar{z},d)\\ \nabla g_j(\bar{z})^T\ ,\ j\in I_0^g(\bar{z},d)\end{pmatrix}$$

and

$$b=\begin{pmatrix}-\kappa F^{\circ\circ}(\bar{z},d)-f^{\circ\circ}(\bar{z},d)+\varphi''(\bar{x},d_1)\\ -(G_i)^{\circ\circ}(\bar{z},d)\ ,\ i\in I_0^G(\bar{z},d)\\ -\left(g_j\right)^{\circ\circ}(\bar{z},d)\ ,\ j\in I_0^g(\bar{z},d)\end{pmatrix}.$$

It follows that $Ar<b$ does not have a solution r. Or equivalently, the linear constrained program $\max\ \{s\mid Ar+\hat{s}\leq b\}$ has a solution $\bar{s}\leq 0$. Here $\hat{s}$ is the vector with all components equal to s. Hence, from the duality theory, the program

$$\min\left\{b^T\zeta, A^T\zeta=0,\quad\sum\zeta_k=1,\zeta_k\geq 0\right\},$$

admits a non-positive optimal value $b^T\bar{\zeta}\leq 0$ with $A^T\bar{\zeta}=0$, $\sum\bar{\zeta}_k=1$, and $\bar{\zeta}_k\geq 0$. Thus, the system

$$A^T\zeta=0, b^T\zeta\leq 0,\zeta=(\lambda,\alpha,\beta)\geq 0,\zeta\neq 0$$

has a solution, where $(\lambda,\alpha,\beta)\in\mathbb{R}\times\mathbb{R}^p\times\mathbb{R}^q$. Thus, there exists $(\lambda,\alpha,\beta)\in\mathbb{R}_+\times\mathbb{R}^p_+\times\mathbb{R}^q_+$ not all zeros such that (3.4) and (3.5) hold true, while taking $\lambda=0$, $\alpha_i=0$ and $\beta_j=0$ for the multipliers that correspond to the functions that have the second-order directional derivative is $-\infty$.

On the other hand, we take $\alpha_i=0$ if $i\notin I_0^G(\bar{z},d)$ and $\beta_j=0$ if $j\notin I_0^g(\bar{z},d)$. In addition, since F, f, $(G_i, i\in\bar{I}^G)$ and $(g_j, j\in\bar{I}^g)$ are locally Lipschitz, Gâteaux differentiable, and regular at $\bar{z}$, then, for $i\in I_0^G(\bar{z},d)$ we have $\nabla G_i(\bar{z})^T d=G_i^0(\bar{z},d)=0$, for $j\in I_0^g(\bar{z},d)$ we have $\nabla g_j(\bar{z})^T d=g_j^0(\bar{z},d)=0$

and $\kappa\nabla F(\bar{z})^T d + \nabla f(\bar{z})^T d - \nabla_x \ell(\bar{x},\bar{y},\overline{w})^T d_1 = 0$. Consequently, the multipliers (λ,α,β) satisfy inequalities (3.6)-(3.8). This completes the proof. □

4. SECOND-ORDER NECESSARY OPTIMALITY CONDITIONS VIA Ψ-REFORMULATION

In the previous section, we derived dual second-order necessary optimality conditions by using reformulation (*LLVF*). However, this was only possible in the case that φ is differentiable under constraint qualifications *LLICQ* and *LSOSC*, and we assumed that the lower-level Lagrange multipliers were a singleton.

To generalize this special case, we consider a new single-level problem in this section that is equivalent to (*LLVF*) at the optimal solution. Our approach involves replacing the function $f-\varphi$ with a new function Ψ, which is Lipschitz near the optimal solution, without any requirements on the *MFCQ* of the lower-level problem.

Let $x \in \mathbb{R}^n$, and assume that the feasible region $Y(x)$ is a closed bounded set. According to Ye [29], (BOP) is locally equivalent to the following single-level problem

$$\min_{x,y} \{F(x,y) \mid \Psi(x,y) \le 0,\ G(x,y) \le 0,\ g(x,y) \le 0\}, \qquad \text{(LL}\Psi\text{R)}$$

where

$$\Psi(x,y) = \max_{z\in Y(x)} \sigma(x,y,z),$$

and

$$\sigma(x,y,z) = \min\{f(x,y) - f(x,z), -\max_{j\in J} g_j(x,z)\}.$$

The function $\Psi(.,.)$ is locally Lipschitz [29, Proposition 2.2].

The following lemma will be used to prove that if $(\bar{x},\bar{y})$ is a global solution to (BOP), then $(\bar{x},\bar{y})$ is also a global solution to single-level problem (LLΨR). The argument is similar to that used in [29] to prove local equivalence.

Lemma 4.1.

1) $$\left\{ \begin{array}{c} (x,y) \in \mathbb{R}^n \times \mathbb{R}^m \\ y \in Y(x) \ \ and \ \ f(x,y) - \varphi(x) < 0 \end{array} \right\} = \left\{ \begin{array}{c} (x,y) \in \mathbb{R}^n \times \mathbb{R}^m \\ y \in Y(x) \ \ and \ \ \Psi(x,y) < 0 \end{array} \right\} = \varnothing.$$

2) $$\left\{ \begin{array}{c} (x,y) \in \mathbb{R}^n \times \mathbb{R}^m \\ y \in Y(x) \ \ and \ \ f(x,y) - \varphi(x) = 0 \end{array} \right\} = \left\{ \begin{array}{c} (x,y) \in \mathbb{R}^n \times \mathbb{R}^m \\ y \in Y(x) \ \ and \ \ \Psi(x,y) = 0 \end{array} \right\}.$$

3) *If $(\bar{x},\bar{y})$ is a feasible point to* (BOP)*, then the solution set of problem $\max_{z\in Y(x)} \sigma(\bar{x},\bar{y},z)$ is given by $S(\bar{x})$.*

Proof. **1)** Since $Y(x)$ is compact, we have

$$\left\{ \begin{array}{c} (x,y) \in \mathbb{R}^n \times \mathbb{R}^m \\ y \in Y(x) \ \text{ and } \ \Psi(x,y) < 0 \end{array} \right\} = \left\{ \begin{array}{c} (x,y) \in \mathbb{R}^n \times \mathbb{R}^m : \\ y \in Y(x) \ \text{ and } \ [\sigma(x,y,z) < 0,\ \forall z \in Y(x)] \end{array} \right\}.$$

Since

$$z \in Y(x) \Longleftrightarrow g(x,z) \le 0 \Longleftrightarrow \max_{i=1}^{q} g_i(x,z) \le 0,$$

one sees that

$$\sigma(x,y,z) < 0, \quad \forall z \in Y(x) \Longleftrightarrow f(x,y) - f(x,z) < 0, \quad \forall z \in Y(x)$$
$$\Longleftrightarrow f(x,y) < \varphi(x),$$

which means

$$\left\{ \begin{array}{c} (x,y) \in \mathbb{R}^n \times \mathbb{R}^m \\ y \in Y(x) \text{ and } \Psi(x,y) < 0 \end{array} \right\} = \left\{ \begin{array}{c} (x,y) \in \mathbb{R}^n \times \mathbb{R}^m \\ y \in Y(x) \text{ and } f(x,y) - \varphi(x) < 0 \end{array} \right\}.$$

In veiw of $\left\{ \begin{array}{c} (x,y) \in \mathbb{R}^n \times \mathbb{R}^m \\ y \in Y(x) \text{ and } f(x,y) - \varphi(x) < 0 \end{array} \right\} = \emptyset$, one obtains the result.

2) We prove a double inclusion.

a) Let $x \in \mathbb{R}^n$ and $y \in Y(x)$ with $f(x,y) - \varphi(x) = 0$. We obtain that y is a global solution to problem ($P[x]$) w.r.t. x. By **1)**, we obtain $\Psi(x,y) \geq 0$. Precisely, we have $\Psi(x,y) = 0$. Indeed, if $\Psi(x,y) > 0$, we have $z \in Y(x)$ and $\sigma(x,y,z) > 0$. Thus

$$f(x,y) - f(x,z) > 0 \text{ and } -\max_{i=1}^{q} g_i(x,z) > 0.$$

It follows that $f(x,y) - f(x,z) > 0$ and $g(x,y) \leq 0$. Then, $z \in Y(x)$ such that $f(x,y) > f(x,z)$. This is a contradiction, because y is a global solution to problem ($P[x]$) w.r.t. x. One concludes that

$$\left\{ \begin{array}{c} (x,y) \in \mathbb{R}^n \times \mathbb{R}^m \\ y \in Y(x) \text{ and } f(x,y) - \varphi(x) = 0 \end{array} \right\} \subseteq \left\{ \begin{array}{c} (x,y) \in \mathbb{R}^n \times \mathbb{R}^m \\ y \in Y(x) \text{ and } \Psi(x,y) = 0 \end{array} \right\}.$$

b) We now prove the opposite inclusion. Let $x \in \mathbb{R}^n$ and $y \in Y(x)$ be such that $\Psi(x,y) = 0$. Then, for every $z \in Y(x)$, one has

$$\min\left\{ f(x,y) - f(x,z), -\max_{i=1}^{q} g_i(x,z) \right\} \leq 0.$$

Since $z \in Y(x) \Longleftrightarrow g(x,z) \leq 0 \Longleftrightarrow \max_{i=1}^{q} g_i(x,z) \leq 0$, one has $f(x,y) - \varphi(x) \leq 0$. By **1)**, one deduces that $f(x,y) - \varphi(x) = 0$. Finally,

$$\left\{ \begin{array}{c} (x,y) \in \mathbb{R}^n \times \mathbb{R}^m \\ y \in Y(x) \text{ and } f(x,y) - \varphi(x) = 0 \end{array} \right\} \supseteq \left\{ \begin{array}{c} (x,y) \in \mathbb{R}^n \times \mathbb{R}^m \\ y \in Y(x) \text{ and } \Psi(x,y) = 0 \end{array} \right\}.$$

3) Let $(\bar{x},\bar{y})$ be a feasible point to (BOP). Note that $\bar{y} \in S(\bar{x})$. Letting $\bar{z} \in S(\bar{x})$, we have

$$f(\bar{x},\bar{y}) - f(\bar{x},\bar{z}) = 0 \ \text{ and } \ \max_{i=1}^{q} g_i(\bar{x},\bar{z}) \leq 0.$$

Thus $\sigma(\bar{x},\bar{y},\bar{z}) = 0$. Using the fact that $S(\bar{x}) \subset Y(\bar{x})$, it is sufficient to show that

$$\sigma(\bar{x},\bar{y},z) \leq 0 \ \text{ for all } \ z \in Y(\bar{x}).$$

By contradiction, we suppose that there exists $z \in Y(x)$ such that $\sigma(\bar{x},\bar{y},z) > 0$. Then

$$f(\bar{x},\bar{y}) - f(\bar{x},z) > 0 \ \text{ and } \ -\max_{i=1}^{q} g_i(\bar{x},z) > 0.$$

Hence, $f(\bar{x},\bar{y}) > f(\bar{x},z)$ and $g(\bar{x},z) \leq 0$. A contradiction with $\bar{y} \in S(\bar{x})$. □

To obtain necessary optimality conditions for (BOP), we use a partial calmness condition for nonsmooth optimization problem (LLΨR). This condition adds the constraint $\Psi(x,y) \leq 0$ to the objective function. To this end, we study the penalized problem

$$\begin{cases} \min\limits_{x,y} \mu F(x,y) + \Psi(x,y) \\ G(x,y) \leq 0 \\ g(x,y) \leq 0. \end{cases} \qquad \text{(LL}\Psi\text{R}[\mu])$$

Using the above arguments together with the notation in Section 3, the critical cone at the point $\bar{z} = (\bar{x}, \bar{y}) \in \Omega$ is given by

$$\mathscr{C}(\bar{z}) = \left\{ d \in \mathbb{R}^n \times \mathbb{R}^m \;\middle|\; \begin{array}{rll} \mu F^\circ(\bar{z},d) + \Psi^\circ(\bar{z},d) & \leq 0 & \\ G_i^\circ(\bar{z},d) & \leq 0 \;, & i \in \bar{I}^G \\ g_j^\circ(\bar{z},d) & \leq 0 \;, & j \in \bar{I}^g \end{array} \right\}.$$

Theorem 4.1. *Let $\bar{z} = (\bar{x}, \bar{y})$ be a local solution to problem* (BOP)*, assumed to be partially calm at $\bar{z}$. Suppose that the functions $(G_i, i \notin \bar{I}^G)$ and $(g_j, j \notin \bar{I}^g)$ are continuous at $\bar{z}$, and the functions F, $(G_i, i \in \bar{I}^G)$ and $(g_j, j \in \bar{I}^g)$ are locally Lipschitz. Moreover, suppose that the Clarke second-order directional derivative of Ψ exists at $\bar{z}$. Then, for every $d \in \mathscr{C}(\bar{z})$, there is no $r \in \mathbb{R}^n \times \mathbb{R}^m$ which solves the system*

$$\mu \left(F^\circ(\bar{z},r) + F^{\circ\circ}(\bar{z},d)\right) + \Psi^\circ(\bar{z},r) + \Psi^{\circ\circ}(\bar{z},d) < 0 \tag{4.1}$$

$$G_i^\circ(\bar{z},r) + G_i^{\circ\circ}(\bar{z},d) < 0, \quad i \in I_0^G(\bar{z},d) \tag{4.2}$$

$$g_j^\circ(\bar{z},r) + g_j^{\circ\circ}(\bar{z},d) < 0, \quad j \in I_0^g(\bar{z},d). \tag{4.3}$$

Proof. The proof is similar to that of Theorem 3.1 (Steps 1 and 3 are the same). Thus we detail Step 2 only.

Let $\bar{z} = (\bar{x}, \bar{y})$ be a solution to (BOP) where partial calmness holds. It follows from Proposition 2.2 that there exists $\mu > 0$ such that $\bar{z}$ is a solution to (LLΨR[μ]).

Step 2: We prove that there exists $\nu > 0$ such that

$$\mu F(\bar{z} + td + \frac{1}{2}t^2 r) + \Psi(\bar{z} + td + \frac{1}{2}t^2 r) < \mu F(\bar{z}) + \Psi(\bar{z})$$

for all $t \in [0, \nu)$. We proceed in two cases.

(a) If $\mu F^\circ(\bar{z},d) + \Psi^\circ(\bar{z},d) < 0$, we aim to demonstrate the existence of $\nu > 0$ by verifying, for all $t \in [0, \nu)$,

$$\mu F\left(\bar{z} + td + \frac{1}{2}t^2 r\right) + \Psi\left(\bar{z} + td + \frac{1}{2}t^2 r\right) < \mu F(\bar{z}) + \Psi(\bar{z}).$$

Suppose by contradiction that, for any $\nu > 0$, there exists $t_\nu \in [0, \nu)$ such that

$$\mu F\left(\bar{z} + t_\nu d + \frac{1}{2}t_\nu^2 r\right) + \Psi\left(\bar{z} + t_\nu d + \frac{1}{2}t_\nu^2 r\right) \geq \mu F(\bar{z}) + \Psi(\bar{z}).$$

Let $\nu_n > 0$ be a sequence convergent to 0 as $n \to \infty$ and $t_n \in (0, \nu_n)$ such that

$$\mu F\left(\bar{z} + t_n d + \frac{1}{2}t_n^2 r\right) - \mu F(\bar{z}) + \Psi\left(\bar{z} + t_n d + \frac{1}{2}t_n^2 r\right) - \Psi(\bar{z}) \geq 0.$$

For each n, we set $u_n^{\nu} = \bar{z} + t_n d$ and $v_n^{\nu} = \bar{z} + t_n d + \frac{1}{2}t_n^2 r$. Then

$$\begin{aligned}
0 \leq & \limsup_{t_n \to 0} \left(\frac{\mu F(v_n^{\nu}) - \mu F(\bar{z})}{t_n} + \frac{\Psi(v_n^{\nu}) - \Psi(\bar{z})}{t_n} \right) \\
& \leq \limsup_{t_n \to 0} \frac{1}{t_n} [\mu F(u_n^{\nu}) - \mu F(\bar{z})] + \limsup_{t_n \to 0} \frac{1}{t_n} [\mu F(v_n^{\nu}) - \mu F(u_n^{\nu})] \\
& + \limsup_{t_n \to 0} \frac{1}{t_n} [\Psi(u_n^{\nu}) - \Psi(\bar{z})] + \limsup_{t_n \to 0} \frac{1}{t_n} [\Psi(v_n^{\nu}) - \Psi(u_n^{\nu})] \\
0 \leq & \mu F^{\circ}(\bar{z}, d) + \Psi^{\circ}(\bar{z}, d) + \limsup_{t_n \to 0} L_F t_n \|r\| + \limsup_{t_n \to 0} L_{\Psi} t_n \|r\|
\end{aligned}$$

as F and Ψ are locally Lipschitz of constant $L_F > 0$ and $L_{\Psi} > 0$, respectively. Thus $\mu F^{\circ}(\bar{z}, d) + \Psi^{\circ}(\bar{z}, d) \geq 0$, which contradicts the fact that $\mu F^{\circ}(\bar{z}, d) + \Psi^{\circ}(\bar{z}, d) < 0$. We see that there exists $\nu > 0$, for all $t \in (0, \nu)$,

$$\mu F\left(\bar{z} + td + \frac{1}{2}t^2 r\right) + \Psi\left(\bar{z} + td + \frac{1}{2}t^2 r\right) < \mu F(\bar{z}) + \Psi(\bar{z}).$$

(b) If $\mu F^{\circ}(\bar{z}, d) + \Psi^{\circ}(\bar{z}, d) = 0$, then there exists $\nu > 0$, for all $t \in [0, \nu)$,

$$\mu F\left(\bar{z} + td + \frac{1}{2}t^2 r\right) + \Psi\left(\bar{z} + td + \frac{1}{2}t^2 r\right) < \mu F(\bar{z}) + \Psi(\bar{z}).$$

To show this, we assume by contradiction that, for any $\nu > 0$, there exists $0 \leq t_{\nu} < \nu$ such that

$$\mu F\left(\bar{z} + t_{\nu} d + \frac{1}{2}t_{\nu}^2 r\right) + \Psi\left(\bar{z} + t_{\nu} d + \frac{1}{2}t_{\nu}^2 r\right) \geq \mu F(\bar{z}) + \Psi(\bar{z}).$$

Let $\nu_n > 0$ be a sequence convergent to 0 as $n \to \infty$ and $t_n \in [0, \nu_n)$ such that

$$\mu F(v_n^{\nu}) + \Psi(v_n^{\nu}) \geq \mu F(\bar{z}) + \Psi(\bar{z}).$$

Then

$$\begin{aligned}
0 \leq & \mu F(v_n^{\nu}) - \mu F(\bar{z}) + \Psi(v_n^{\nu}) - \Psi(\bar{z}) \\
= & \frac{t_n^2}{2} \left[\frac{2}{t_n^2} (\mu F(v_n^{\nu}) - \mu F(u_n^{\nu})) \right] + \frac{t_n^2}{2} \left[\frac{2}{t_n^2} (\mu F(u_n^{\nu}) - \mu F(\bar{z}) - t_n \mu F^{\circ}(\bar{z}, d)) \right] \\
+ & \frac{t_n^2}{2} \left[\frac{2}{t_n^2} (\Psi(v_n^{\nu}) - \Psi(u_n^{\nu})) \right] + \frac{t_n^2}{2} \left[\frac{2}{t_n^2} (\Psi(u_n^{\nu}) - \Psi(\bar{z}) - t_n \Psi^{\circ}(\bar{z}, d)) \right].
\end{aligned}$$

After dividing the above inequality by $t_n^2/2$ and taking the upper limit as $t_n \to 0+$, we obtain

$$0 \leq \mu F^{\circ}(\bar{z}, r) + \mu F^{\circ\circ}(\bar{z}, d) + \Psi^{\circ}(\bar{z}, r) + \Psi^{\circ\circ}(\bar{z}, d),$$

which contradicts the fact that r is an optimal solution to system (4.1). We obtain that there exists $\nu > 0$ such that, for all $t \in (0, \nu)$,

$$\mu F\left(\bar{z} + td + \frac{1}{2}t^2 r\right) + \Psi\left(\bar{z} + td + \frac{1}{2}t^2 r\right) < \mu F(\bar{z}) + \Psi(\bar{z}).$$

□

Remark 4.1. Since the differentiability of the value function Ψ is not a requirement for Theorem 4.1 because Ψ is locally Lipschitz at $\bar{z}$, then the Clarke first-order derivative exists. The sole condition on Ψ is the existence of its second-order directional derivative at an optimal solution. Ye [29] provided an upper approximation of the Clark generalized gradient at an optimal solution and Auslender and Cominetti [3] gave estimates to the first and second-order directional derivatives under *LMFCQ* and *LSOSC*, although the Clarke second-order derivative of Ψ may only be relevant in certain cases, and it does not impact the validity of Theorem 4.1.

Example 4.1. Consider the following bilevel optimization problem:

$$\min_{x,y} \{F(x,y) = |x| + y^2 \mid G(x,y) = -x \leq 0,\ y \in S(x)\}, \tag{BOP}$$

where, for each $x \in \mathbb{R}$, $S(x)$ is the set of optimal solutions to the following parametric optimization problem:

$$\min_{y} \{f(x,y) = x^2 - y \mid g_1(x,y) = y \leq 0,\ g_2(x,y) = -y - x \leq 0\}. \tag{P[x]}$$

On easily remarks that $(0,0)$ is the unique global optimum to (BOP). Before showing that $\bar{z} = (\bar{x},\bar{y}) = (0,0)$ verifies the necessary conditions of our theorem, we need to check that *LMFCQ* condition is not satisfied at the feasible point $\bar{z} = (0,0)$ and that the set of lower-level Lagrange multipliers is not a singleton, and therefore Theorem 3.1 cannot be applied.

We have $\bar{I}^g = \{1,2\}$, and, for each $d \in \mathbb{R}$,

$$\nabla_y g_1(0,0)^T d = d < 0 \text{ and } \nabla_y g_2(0,0)^T d = -d < 0,$$

which is impossible. Thus *LMFCQ* does not hold at $\bar{z} = (0,0)$. On the other, letting $w = (w_1,w_2) \in \Lambda(0,0)$, we have

$$\ell(x,y,w) = x^2 - y + w_1 y + w_2(-y-x) \text{ and } \nabla_y \ell(0,0,w) = -1 + w_1 - w_2 = 0 \text{ and } w^T g(0,0) = 0.$$

Consequently, $\Lambda(0,0) = \{(w_2+1,w_2) \mid w_2 \in \mathbb{R}\}$ is not a singleton.

We now analyze the feasible point $\bar{z} = (0,0)$ by checking the second-order necessary conditions of Theorem 4.1. We have $S(x) = \{0\}$ and $Y(x) = \{y \in \mathbb{R} \mid g(x,y) \leq 0\} = [-x,0]$. Then

$$\Psi(x,y) = \max_{z \in Y(x)} \sigma(x,y,z) = \sigma(x,y,0) = \min\{-y,0\}.$$

The penalized problem reformulation of (BOP) is given by:

$$\begin{cases} \min_{x,y} \mu F(x,y) + \Psi(x,y) \\ G(x,y) = -x \leqslant 0 \\ g_1(x,y) = y \leqslant 0 \\ g_2(x,y) = -y - x \leqslant 0 \end{cases} \tag{LLΨR[μ]}$$

where $\mu > 0$. We have $\bar{I}^G = \{1\}$, $\bar{I}^g = \{1,2\}$, and

$$\mu F^\circ(\bar{z},v) + \Psi^\circ(\bar{z},v) = \mu|v_1| - v_2$$

and

$$G^\circ(\bar{z},v) = -v_1,\ \ g_1^\circ(\bar{z},v) = v_2 \text{ and } g_2^\circ(\bar{z},v) = -v_1 - v_2.$$

Let $d=(d_1,d_2)\in\mathscr{C}(\bar z)$. We have

$$\begin{cases} \mu|d_1|-d_2\leqslant 0\\ -d_1\leqslant 0\\ d_2\leqslant 0\\ -d_1-d_2\leqslant 0.\end{cases}$$

Thus $d=(0,0)$ is the unique critical direction. As a result, $I_0^G(\bar z,d)=\{1\}$ and $I_0^g(\bar z,d)=\{1,2\}$, and hence

$$\mu F^{\circ\circ}(\bar z,v)+\Psi^{\circ\circ}(\bar z,v)=2\mu v_2^2,\ \ G^{\circ\circ}(\bar z,v)=0,\ \ g_1^{\circ\circ}(\bar z,v)=0 \ \text{ and }\ g_2^{\circ\circ}(\bar z,v)=0.$$

Finally, we find the following system which does not admit any solution

$$\begin{cases} \mu\left(F^{\circ}(\bar z,r)+F^{\circ\circ}(\bar z,d)\right)+\Psi^{\circ}(\bar z,r)+\Psi^{\circ\circ}(\bar z,d)=\mu|r_1|-r_2<0\\ G^{\circ}(\bar z,r)+G^{\circ\circ}(\bar z,d)=-r_1<0\\ g_1^{\circ}(\bar z,r)+g_1^{\circ\circ}(\bar z,d)=r_2<0\\ g_2^{\circ}(\bar z,r)+g_2^{\circ\circ}(\bar z,d)=-r_1-r_2<0.\end{cases}$$

This proves that $\bar z=(0,0)$ is a possible global optimum of (BOP).

In the following theorem, we establish dual second-order necessary optimality conditions. To proceed, let $\bar z=(\bar x,\bar y)\in\Omega$ and define the following reduced critical cone

$$\mathscr{RC}(\bar z)=\left\{ d\in\mathscr{C}(\bar z)\ \middle|\ \begin{array}{rl} \mu F^{\circ\circ}(\bar z,d)+\Psi^{\circ\circ}(\bar z,d) & <\infty \text{ if } \mu F^{\circ}(\bar z,d)+\Psi^{\circ}(\bar z,d)=0\\ G_i^{\circ\circ}(\bar z,d) & <\infty\ ,\ i\in I_0^G(\bar z,d)\\ g_j^{\circ\circ}(\bar z,d) & <\infty\ ,\ j\in I_0^g(\bar z,d)\end{array}\right\}.$$

Theorem 4.2. *Let $\bar z=(\bar x,\bar y)$ be a local minimizer to (BOP), assumed to be partially calm at $\bar z$. Suppose that the functions F, $(G_i, i\in\bar I^G)$ and $(g_j, j\in\bar I^g)$ are locally Lipschitz, regular and Gâteaux differentiable at $\bar z$, and the functions $(G_i, i\notin\bar I^G)$ and $(g_j, j\notin\bar I^g)$ are continuous at $\bar z$. Moreover, suppose that the second-order directional derivative of Ψ exists at $\bar z$. Then, for every critical direction $d\in\mathscr{RC}(\bar z)$ that is not equal to zero, there exist multipliers $(\lambda,\alpha,\beta)\in\mathbb{R}_+\times\mathbb{R}_+^p\times\mathbb{R}_+^q$ that are non-negative and not all zeros with*

$$\alpha_i G_i(\bar z)=0, i\in I \ \text{ and }\ \beta_j g_j(\bar z)=0,\quad j\in J \tag{4.4}$$

$$\lambda\mu\nabla F(\bar z)^T d+\lambda\Psi^{\circ}(\bar z,d)=0 \tag{4.5}$$

$$\alpha_i\nabla G_i(\bar z)^T d=0, i\in\bar I^G \ \text{ and }\ \beta_j\nabla g_j(\bar z)^T d=0,\quad j\in\bar I^g \tag{4.6}$$

$$\lambda\mu F^{\circ\circ}(\bar z,d)+\lambda\Psi^{\circ\circ}(\bar z,d)+\sum_{i\in\bar I^G}\alpha_i G_i^{\circ\circ}(\bar z,d)+\sum_{j\in\bar I^g}\beta_j g_j^{\circ\circ}(\bar z,d)\geq 0. \tag{4.7}$$

Proof. Let d be any fixed critical direction. Note that the following system in r has no solution

$$\begin{array}{rl} \mu\nabla F(\bar z)^T r+F^{\circ\circ}(\bar z,d)+\Psi^{\circ}(\bar z,r)+\Psi^{\circ\circ}(\bar z,d)<0, & \text{if } 0\in I_0(\bar z,d)\\ \nabla G_i(\bar z)^T r+G_i^{\circ\circ}(\bar z,d)<0, & i\in I_0(\bar z,d)\\ \nabla g_j(\bar z)^T r+g_j^{\circ\circ}(\bar z,d)<0, & j\in I_0(\bar z,d).\end{array}$$

According to [13], the following system in $(\lambda,\alpha,\beta)\in\mathbb{R}\times\mathbb{R}^p\times\mathbb{R}^q$

$$\lambda\mu(\nabla F(\bar z)^T r+F^{\circ\circ}(\bar z,d))+\lambda(\Psi^{\circ}(\bar z,r)+\Psi^{\circ\circ}(\bar z,d)+$$
$$\sum_{i\in I_0(\bar z,d)}\alpha_i(\nabla G_i(\bar z)^T r+G_i^{\circ\circ}(\bar z,d))+\sum_{j\in I_0(\bar z,d)}\beta_j(\nabla g_j(\bar z)^T r+g_j^{\circ\circ}(\bar z,d))\geq 0$$

has a solution for all $r \in \mathbb{R}^n$. For $r = d$, we find

$$\lambda\mu F^{\circ\circ}(\bar{z},d) + \lambda\Psi^{\circ\circ}(\bar{z},d) + \sum_{i\in I_0(\bar{z},d)} \alpha_i G_i^{\circ\circ}(\bar{z},d) + \sum_{j\in I_0(\bar{z},d)} \beta_j g_j^{\circ\circ}(\bar{z},d) \geq 0.$$

By taking $\alpha_i = 0$ if $i \notin I_0^G(\bar{z},d)$; and $\beta_j = 0$ if $j \notin I_0^g(\bar{z},d)$, the systems become

$$\lambda\mu F^{\circ\circ}(\bar{z},d) + \lambda\Psi^{\circ\circ}(\bar{z},d) + \sum_{i\in \bar{I}^G} \alpha_i G_i^{\circ\circ}(\bar{z},d) + \sum_{j\in \bar{I}^g} \beta_j g_j^{\circ\circ}(\bar{z},d) \geq 0.$$

This shows (4.4) and (4.7).

On the other hand, since $(G_i, i \in \bar{I}^G)$ and $(g_j, j \in \bar{I}^g)$ are locally Lipschitz, Gâteaux differentiable, and regular at $\bar{z}$, then, for $i \in I_0^G(\bar{z},d)$, we have $\nabla G_i(\bar{z})^T d = G_i^0(\bar{z},d) = 0$. For $j \in I_0^g(\bar{z},d)$, we have $\nabla g_j(\bar{z})^T d = g_j^0(\bar{z},d) = 0$, which proves system (4.6).

Finally, to show (4.5), we take $\lambda = 0$ if $\mu F^\circ(\bar{z},d) + \Psi^\circ(\bar{z},d) < 0$. □

Conclusion

In this paper, we investigated second-order necessary optimality conditions for standard bilevel optimization problems with smooth and nonsmooth data. We used two different single-level reformulations of the hierarchical model under a partial calmness assumption and derived primal and dual conditions in terms of the initial problem data. We also gave some examples to illustrate our results. Our work improves some corresponding existing results in [9, 22] and provides new insights into the second-order properties of bilevel optimization problems. As future research directions, we plan to explore necessary optimality conditions for an isolated local minimizer of order two and for KKT-reformulation of (BOP).

References

[1] B. Aghezzaf, M. Hachimi, Second order optimality conditions in multiobjective optimization problems, J. Optim. Theory Appl. 102 (1999), 37–50.

[2] J. Aubin, I. Ekeland, Applied Nonlinear Analysis, Courier Corporation, 2006.

[3] A. Auslender, R. Cominetti, First and second order sensitivity analysis of nonlinear programs under directional constraint qualification conditions, Optimization, 21 (1990), 351-363.

[4] M. S. Bazaraa, D. H. Sherali, M. C. Shetty, Nonlinear Programming: Theory and Algorithms, John Wiley and Sons, 2013.

[5] A. Ben-Tal, J. Zowe, Necessary and sufficient optimality conditions for a class of nonsmooth minimization problems, Math. Program. 24 (1982), 70-91.

[6] S. Bolintineau, M. El Maghri, Second order efficiency conditions and sensitivity of efficient points, J. Optim. Theory Appl. 98 (1998), 569–592.

[7] F. H. Clarke, Nonsmooth analysis and optimization, Proceedings of the International Congress of Mathematicians, pp. 847-853, 1983.

[8] S. Dempe, Foundations of Bilevel Programming, Springer Science and Business Media, 2002.

[9] S. Dempe, N. Gadhi, Second order optimality conditions for bilevel set optimization problems, J. Global Optim. 47 (2010), 233-245.

[10] V. F. Demyanov, Minimax: Differentiability in Directions, Leningrad University Publ. Leningrad, 1974.

[11] V. F. Demyanov, A. B. Pevnyi, First and second marginal values of mathematical programming problems, Doklady Akademii Nauk. Russian Acad. Sci. pp. 277-280, 1972.

[12] N. Dinh, B. Mordukhovich, T. T. Nghia, Subdifferentials of value functions and optimality conditions for DC and bilevel infinite and semi-infinite programs, Math. Program. 123 (2010), 101-138.

[13] F. Flores-Bazan, F. Flores-Bazan, C. Vera, Gordan-type alternative theorems and vector optimization revisited, Recent Developments in Vector Optimization, pp. 29-59, Springer, Berlin, Heidelberg, 2012.

[14] J. Gauvin, F. Dubeau, Differential properties of the marginal function in mathematical programming, Math. Program. Stud. 19 (1982), 101–119.
[15] A. Guerraggio, D. T. Luc, Optimality conditions for C1,1 vector optimization problems, J. Optim. Theory Appl. 109 (2001), 615–629.
[16] A. Guerraggio, D. T. Luc, N. B. Minh, Second order optimality conditions for C1 multiobjective programming problems, Acta Math. Vietnam 26 (2001), 257–268.
[17] R. Henrion, A. Jourani, J. Outrata, On the calmness of a class of multifunctions, SIAM J. Optim. 13 (2002), 603-618.
[18] A. D. Ioffe, J. Outrata, On metric and calmness qualification conditions in subdifferential calculus, Set-Valued Anal. 16 (2008), 199-227.
[19] R. Janin, Directional derivative of the marginal function in nonlinear programming. In: Fiacco, A.V. (eds) Sensitivity, Stability and Parametric Analysis. Mathematical Programming Studies, vol 21. Springer, Berlin, Heidelberg, 1984.
[20] A. A. Khan, C. Tammer, C. Zalinescu, Set-valued Optimization, Springer-Verlag, Berlin, 2016.
[21] D. Klatte, B. Kummer, Nonsmooth equations in optimization: regularity, calculus, methods and applications, Nonconvex Optimization and Its Applications, Springer, New York, 2002.
[22] L. Lafhim, First-and second-order optimality conditions in optimistic bilevel set-valued programming, Optimization, 71 (2022), 1955-1981.
[23] D. T. Luc, Second order optimality conditions for problems with continuously differentiable data, Optimization 51 (2002), 497–510.
[24] P. Mehlitz, A. Zemkoho, Sufficient optimality conditions in bilevel programming, Math. Oper. Res. 46 (2021), 1573-1598.
[25] L. Minchenko, A. Tarakanov, On second-order directional derivatives of value functions, Optimization, 64 (2015), 389-407.
[26] Z. Pales, V.M. Zeidan, Nonsmooth optimum problems with constraints, SIAM J. Control Optim. 32 (1994), 1476-1502.
[27] A. Shapiro, On concepts of directional differentiability, J. Optim. Theory Appl. 66 (1990), 477-487.
[28] A. Shapiro, Sensitivity analysis of nonlinear programs and differentiability properties of metric projections, SIAM J. Control Optim. 26 (1988), 628–645.
[29] J. J. Ye, Constraint qualifications and KKT conditions for bilevel programming problems, Math. Oper. Res. 31 (2006), 811-824.
[30] J. J. Ye, D.L. Zhu, Optimality conditions for bilevel programming problems, Optimization, 33 (1995), 9-27.
[31] J. J. Ye, D. L. Zhu, Q. J. Zhu, Exact penalization and necessary optimality conditions for generalized bilevel programming problems, SIAM J. Optim. 7 (1997), 481-507.