第49卷第5期 中国科学技术大学学报 Vol. 49, No. 5
2019年5月 JOURNAL OF UNIVERSITY OF SCIENCE AND TECHNOLOGY OF CHINA May 2019

文章编号:0253-2778(2019)05-0351-17

Variational analysis for pessimistic semivectorial bilevel programming with nonsmooth data

LIU Bingbing[1,2], CHEN Sugen[1]

(1. *School of Mathematics and Computational Science*, *Anqing Normal University*, *Anqing* 246133, *China*;
2. *School of Management*, *University of Science and Technology of China*, *Hefei* 230026, *China*)

Abstract: Using variational analysis theory developed recently by Mordukhovich, the pessimistic semivectorial bilevel programming problem (PSBPP) was investigated. PSBPP was first transformed into a scalar bilevel optimization problem with the help of a scalarization method. Furthermore, using single-level and two-level optimal value functions reformulations and generalized differentiation calculus of Mordukhovich, the first-order necessary optimality conditions were established for the resulting scalar bilevel optimization problem and thus for the PSBPP with nonsmooth data.

Key words: pessimistic semivectorial bilevel programming problem; necessary optimality condition; Lipschitz continuous; optimal value function reformulation; sensitivity analysis

CLC number: O221; O224 **Document code**: A doi:10.3969/j.issn.0253-2778.2019.05.002

2010 Mathematics Subject Classification: 90C26; 90C30; 90C31; 90C46

Citation: LIU Bingbing, CHEN Sugen. Variational analysis for pessimistic semivectorial bilevel programming with nonsmooth data[J]. Journal of University of Science and Technology of China, 2019,49(5):351-367.
刘兵兵,陈素根. 非光滑悲观半向量双层规划的变分分析[J]. 中国科学技术大学学报,2019,49(5):351-367.

非光滑悲观半向量双层规划的变分分析

刘兵兵[1,2],陈素根[1]

(1.安庆师范大学数学与计算科学学院,安徽安庆 246133;2.中国科学技术大学管理学院,安徽合肥 230026)

摘要:利用最近由Mordukhovich发展的变分分析理论,研究了悲观半向量双层规划问题,得到了在非光滑情形下的悲观半向量双层规划问题的必要最优性条件.为了得到该最优性条件,首先借助于标量化方法将悲观半向量双层规划问题转化为一个标量的双层优化问题.进而利用单层和两层值函数构造和Mordukhkvich广义微分计算规则,研究得到了所得的标量双层优化问题的一阶必要最优性条件,进而根据原悲观半向量双层规划问题与所得的标量双层优化问题的等价命题得到了原问题在非光滑情形下的一阶必要最优性条件.

关键词:悲观半向量双层规划问题;必要最优性条件;李普希兹连续;最优值函数构造;灵敏度分析

Received: 2017-05-31; **Revised**: 2018-01-04

Foundation item: Supported by the National Natural Science Foundation of China (61702012), the Scientific Research Foundation of the Higher Education Institutions of Anhui Province of China (KJ2017A361), the University Outstanding Young Talent Support Project of Anhui Province of China (gxyq2017026).

Biography: LIU Bingbing (corresponding author), male, born in 1980, PhD/associate Prof. Research field: Operational management, optimization theory and method. E-mail:lbb1224@mail.ustc.edu.cn

0 Introduction

In this study, we consider the following a class of bilevel programming problem in which the upper-level objective function is scalar and the lower-level objective function is vectorial[1], i. e., the semivectorial bilevel programming problem (SBPP):

$$\text{“}\min_{x}\text{”} F(x,z)$$
$$\text{s. t. } G(x)\leqslant 0,\ z\in\Psi_{\text{wef}}(x) \tag{1}$$

where function $F:R^n\times R^m\to R$ is the upper-level objective function and $G:R^n\to R^q$ denotes the upper-level constraint function. Let the set $\{x \mid G(x)\leqslant 0\}$ be nonempty and closed. The set-valued mapping Ψ_{wef} denotes the weakly efficient optimal solution mapping of the following lower-level multiobjective optimization problem:

$$R^{\nu}_{+}-\min_{z} f(x,z)$$
$$\text{s. t. } g(x,z)\leqslant 0 \tag{2}$$

where the function $f:R^n\times R^m\to R^{\nu}$ is the lower-level multiobjective function and $g:R^n\times R^m\to R^p$ is the lower-level constraint function. The term "$R^{\nu}_{+}-\min$" in (2) is used to symbolize that vector values in the lower-level problem are in the sense of weak Pareto minima with respect to an order induced by the positive orthant of R^{ν}. In order to ensure that the results in this article are correct, we make two hypotheses as follows.

Assumption 0.1 The set $\{x\in R^n \mid G(x)\leqslant 0\}$ is nonempty and compact.

Assumption 0.2 For any x verifying $G(x)\leqslant 0$, the set $\{z\in R^m \mid g(x,z)\leqslant 0\}$ is nonempty and compact.

As we know, the weakly efficient solution set $\Psi_{\text{wef}}(x)$ of the lower-level problem (2) in general has more than one solution. Thus, the notion of an optimal solution for the bilevel programming problem may be ambiguous, which is why the word "min" is written in quotes in (1). In order to overcome this ambiguity, we can consider the optimistic formulation and the pessimistic formulation for the problem (1). The definition of the optimistic and pessimistic semivectorial bilevel programming problems can be found in Ref. [2]. For studies on the optimistic semivectorial bilevel programming problem (OSBPP) the reader is referred to Refs. [1,3-10]. The research review on these literatures can be found in Ref. [2] and is thus omitted.

In contrast to OSBPP, the study on the pessimistic semivectorial bilevel programming problem (PSBPP) can be found in Refs. [2,11-12]. Firstly, we give the reformulation of PSBPP as

$$\min_{x}\max_{z} F(x,z)$$
$$\text{s. t. } G(x)\leqslant 0,\ z\in\Psi_{\text{wef}}(x) \tag{3}$$

In Ref. [11], Bonnel and Morgan developed optimality conditions for a class of bilevel optimal control problem in the optimistic case and pessimistic case, respectively. Nie[12] defined the conservative optimal decision for the PSBPP by using the weighting method. However, the detailed optimality conditions have not been established. Hence, in Ref. [2], using the optimal value function formulation and generalized differentiation calculus of Mordukhovich, Liu et al. developed the detailed first-order necessary optimality conditions for PSBPP under the assumption of the upper-level and lower-level objective functions and constraint functions are continuously differentiable and the lower-level problem is strictly convex. As an application, the necessary optimality conditions for the PSBPP with linear lower-level multiobjective function with respect to the lower-level variables were established. On the other hand, penalty function method[13], K-th best algorithm[14] and maximum entropy approach[15] are currently adopted as solving approaches to the pessimistic problem of general bilevel programming.

It is just the smooth setting under which the necessary optimality conditions were developed in Ref. [2]. In this paper, we intend to extend those results in Ref. [2] to the nonsmooth setting case. For this purpose, by using the generalized differentiation calculus of Mordukhovich, we develop a sensitivity analysis of the lower-level negative value function and the lower-level optimal solution set mapping, respectively. Furthermore, we also develop a sensitivity analysis of the maximization bilevel optimal value functions $\varphi_p(x,y)$ and $\varphi_{pp}(x)$ in the nonsmooth setting. Based on the above results, the first-order necessary optimality conditions are established under the assumption of the basic CQ (see Section 2) holds. The results proposed in this paper extend the results in Ref. [2]. Thus, all results in Ref. [2] are the special case of the results obtained in this paper.

The rest of the paper is organized as follows. In Section 1, we recall the definitions of weakly efficient solutions and Pareto minima, the relevant notions and properties from variational analysis will also be introduced. In Section 2, we transform the PSBPP into a single-level generalized minimax optimization problem with constraints by means of the optimal value function reformulation. In Section 3, in the nonsmooth setting, we first develop the sensitivity analysis estimation of the lower-level negative value function and the lower-level optimal solution set mapping, respectively. Based on these results, we also develop the sensitivity analysis estimation for the maximization bilevel value function. As the most important results, the first-order necessary optimality conditions for the PSBPP are established when all the functions involved are locally Lipschitz continuous. We finish with some conclusions in Section 4.

1 Preliminaries

In this section, we mainly recall some basic definitions of weakly efficient solutions and Pareto minima. Some relevant notions and properties from variational analysis are also introduced.

1.1 Weakly efficient solution and Pareto minima

Definition 1.1 Let set $C \subset R^n$ be a closed convex cone with nonempty interior, if $C \cap -C = \{0\}$, we call C pointed convex cone. We denote a partial order by $\precsim_C$ in R^n induced by C.

Definition 1.2 Let set $A \subseteq R^n$ be nonempty. A point $z^* \in A$ is said to be Pareto (resp. weak Pareto) minima of A with respect to C if

$$A \subset z^* + [(R^n \backslash (-C)) \cup \{0\}]$$
$$(\text{resp. } A \subset z^* + (R^n \backslash -\text{int } C))$$

where 'int' denotes the topological interior of the set in question.

Let us consider the following multiobjective optimization problem with respect to $\precsim_C$:

$$C - \min f(x)$$
$$\text{s.t. } x \in X \tag{4}$$

where f represents a vector-valued function and X the nonempty feasible set. For a nonempty set $A \subset X$, the image of A by f is defined by $f(A) := \{f(x) \mid x \in A\}$.

Definition 1.3 A point $x^* \in X$ is said to be an efficient (resp. weakly efficient) optimal solution of problem (4) if $f(x^*)$ is a Pareto (resp. weak Pareto) minima of $f(X)$.

Definition 1.4 A point $x^* \in X$ is said to be a local efficient (resp. weakly local efficient) optimal solution of problem (4) if there exists a neighborhood U of x^* such that $f(x^*)$ is a Pareto (resp. weak Pareto) minima of $f(X \cap U)$.

Definition 1.5 A vector-valued function $f: R^n \to R^m$ is said to convex with respect to a partial $\precsim_C$ induced by a pointed, closed and convex cone C, if we have

$$f(\lambda x_1 + (1-\lambda)x_2) \precsim_C \lambda f(x_1) + (1-\lambda) f(x_2),$$
$$\forall \lambda \in (0,1), \ \forall x_1, x_2 \in R^n.$$

1.2 Tools from variational analysis

Definition 1.6 Given a point $\bar{x}$, $\limsup_{x\to\bar{x}} \Gamma(x)$ is said to be the Kuratowski-Painlevé outer/upper limit of a set-valued mapping $\Gamma: R^n \rightrightarrows R^m$ at $\bar{x}$, if

$$\limsup_{x\to\bar{x}} \Gamma(x) = \{v \in R^m \mid \exists x_k \to \bar{x},$$
$$v_k \to v \text{ with } v_k \in \Gamma(x_k) \text{ as } k \to \infty\}.$$

Definition 1.7 For an extended real-valued function $\Psi: R^n \to \bar{R}$, $\hat{\partial}\Psi(\bar{x})$ is said to be the Fréchet subdifferential of Ψ at a point $\bar{x}$ of its domain, if

$\hat{\partial}\Psi(\bar{x}) =$

$$\{v \in R^n \mid \liminf_{x \to \bar{x}} \frac{\Psi(x) - \Psi(\bar{x}) - \langle v, x - \bar{x} \rangle}{\| x - \bar{x} \|} \geqslant 0\}.$$

Definition 1.8 Given a point $\bar{x}$, $\partial\Psi(\bar{x})$ is said to be the basic/Mordukhovich subdifferential of Ψ at $\bar{x}$, if

$$\partial\Psi(\bar{x}) = \limsup_{x \to \bar{x}} \hat{\partial}\Psi(\bar{x}) \tag{5}$$

If Ψ is convex, $\partial\Psi(\bar{x})$ is reduced to the subdifferential in the sense of convex analysis:

$$\partial\Psi(\bar{x}) = \{v \in R^n \mid \Psi(x) - \Psi(\bar{x}) \geqslant \langle v, x - \bar{x} \rangle, \forall x \in R^n\},$$

where $\partial\Psi(\bar{x})$ is nonempty and compact when Ψ is locally Lipschitz continuous, its convex hull is the Clarke subdifferential $\bar{\partial}\Psi(\bar{x})$, i. e.

$$\bar{\partial}\Psi(\bar{x}) = \mathrm{co}\, \partial\Psi(\bar{x}) \tag{6}$$

where "co" denotes the convex hull of the set in question. Via this link between the basic and Clarke subdifferential, we have the following convex hull property:

$$\mathrm{co}\, \partial(-\Psi)(\bar{x}) = -\mathrm{co}\, \partial\Psi(\bar{x}) \tag{7}$$

where Ψ is Lipschitz continuous near $\bar{x}$.

Definition 1.9 $\partial_x\Psi(\bar{x}, \bar{y})$ is said to be the partial basic (resp. Clarke) subdifferential of Ψ with respect to x, if we have

$$\partial_x\Psi(\bar{x}, \bar{y}) = \partial\Psi(\cdot, \bar{y})(\bar{x})$$

$$(\text{resp. } \bar{\partial}_x\Psi(\bar{x}, \bar{y}) = \bar{\partial}\Psi(\cdot, \bar{y})(\bar{x})).$$

The partial basic (resp. Clarke) subdifferential with respect to y can be defined analogously as follows

$$\partial_y\Psi(\bar{x}, \bar{y}) = \partial\Psi(\bar{y}, \cdot)(\bar{y})$$

$$(\text{resp. } \bar{\partial}_y\Psi(\bar{x}, \bar{y}) = \bar{\partial}\Psi(\bar{y}, \cdot)(\bar{y})).$$

Definition 1.10 Given a point $\bar{x} \in \Omega$, $N_\Omega(\bar{x})$ is said to be the basic/Mordukhovich normal cone to a set $\Omega \subset R^n$ at $\bar{x}$, if

$$N_\Omega(\bar{x}) = \limsup_{x \to \bar{x}(x \in \Omega)} \hat{N}_\Omega(\bar{x}) \tag{8}$$

where $\hat{N}_\Omega(\bar{x})$ represents the prenormal/Fréchet normal cone to a set Ω at $\bar{x}$ defined by

$$\hat{N}_\Omega(\bar{x}) = \{v \in R^n \mid \limsup_{x \to \bar{x}(x \in \Omega)} \frac{\langle v, x - \bar{x} \rangle}{\| x - \bar{x} \|} \leqslant 0\}.$$

The set Ω will be said to be regular at $\bar{x} \in \Omega$ if $N_\Omega(\bar{x}) = \hat{N}_\Omega(\bar{x})$ holds.

For the lower semicontinuous function Ψ with the epigraph $\mathrm{epi}\Psi$, we can equivalently define the basic/Mordukhovich subdifferential (5) using the normal cone (8) by

$$\partial\Psi(\bar{x}) = \{v \in R^n \mid (v, -1) \in N_{\mathrm{epi}\Psi}(\bar{x}, \Psi(\bar{x})).$$

The singular subdifferential of Ψ at point $\bar{x}$ ($\in \mathrm{dom}\Psi$) is denoted by

$$\partial^\infty\Psi(\bar{x}) = \{v \in R^n \mid (v, 0) \in N_{\mathrm{epi}\Psi}(\bar{x}, \Psi(\bar{x}))\}.$$

If Ψ is lower semicontinuous near $\bar{x}$, then $\partial^\infty\Psi(\bar{x}) = \{0\}$ if and only if Ψ is locally Lipschitz continuous near $\bar{x}$. Given a set-valued mapping Ξ: $R^n \to 2^{R^m}$ with its graph

$$\mathrm{gph}\Xi := \{(x, y) \in R^n \times R^m \mid y \in \Xi(x)\},$$

recall the notion of coderivative for Ξ at $(\bar{x}, \bar{y}) \in \mathrm{gph}\Xi$ is defined by

$$D^*\Xi(\bar{x}, \bar{y})(v) := \{u \in R^n \mid (u, -v) \in N_{\mathrm{gph}\Xi}(\bar{x}, \bar{y})\} \text{ for } v \in R^m \tag{9}$$

via the normal cone (8) to the graph of Ξ. If Ξ is single-valued and locally Lipschitz continuous near $\bar{x}$, its coderivative can be denoted analytically as $D^*\Xi(\bar{x})(v) = \partial\langle v, \Xi\rangle(\bar{x})$ for $v \in R^m$, via the basic subdifferential (5) of the Lagrangian scalarization $\langle v, \Xi\rangle(x) := \langle v, \Xi(x)\rangle$, where the component $\bar{y}(= \Xi(\bar{x}))$ is omitted in the coderivative notation for single-valued mappings. This implies that the coderivative can be represented as $D^*\Xi(\bar{x})(v) = \{\nabla\Xi(\bar{x})^{\mathrm{T}}v\}$ for $v \in R^m$, where Ξ is strictly differentiable at point $\bar{x}$, and $\nabla\Xi(\bar{x})$ denotes its Jacobian matrix at $\bar{x}$.

Definition 1.11 A set-valued mapping Ξ is said to be inner semicompact at $\bar{x}$ with $\Xi(\bar{x}) \neq \emptyset$, if for every sequence $x_k \to \bar{x}$ with $\Xi(x_k) \neq \emptyset$, there exists a sequence of $y_k \in \Xi(x_k)$ which contains a convergent subsequence as $k \to \infty$.

Definition 1.12 A set-valued mapping Ξ is said to be inner semicontinuous at $(\bar{x}, \bar{y}) \in \mathrm{gph}\Xi$, if for every sequence $x_k \to \bar{x}$ there exists a sequence of $y_k \in \Xi(x_k)$ that converges to $\bar{y}$ as $k \to \infty$.

From Definitions 1.11 and 1.12, it is clear that Ξ is inner semicontinuous at $(\bar{x}, \bar{y})$, if Ξ is inner semicompact at $\bar{x}$ with $\Xi(\bar{x}) = \{\bar{y}\}$. Generally speaking, the inner semicontinuity which is much stronger than the inner semicompactness and is a necessary condition for the Lipschitz-like/Aubin property, which means that there exist two neighborhoods U of $\bar{x}$ and V of $\bar{y}$, and a constant $\kappa > 0$ such that $\forall x, u \in U$ and $y \in \Xi(u) \cap V$,

$$d(y, \Xi(x)) \leqslant \kappa \| u - x \| \tag{10}$$

where d means the distance from a point to a set in R^m. When $V = R^m$ in (10), this property is reduced to the classical local Lipschitz continuity of Ξ near $\bar{x}$. A complete characterization of the Lipschitz-like/Aubin property (10), and hence a sufficient condition for the inner semicontinuity of Ξ at $(\bar{x}, \bar{y})$, is given for closed graph mappings by the following coderivative/Mordukhovich criterion (see Refs. [16, Theorem 5.7] and [17, Theorem 9.40]):

$$D^* \Xi(\bar{x}, \bar{y})(0) = \{0\} \tag{11}$$

In addition, the infimum of all $\kappa > 0$ for which (10) holds is equal to the coderivative norm $\| D^* \Xi(\bar{x}, \bar{y}) \|$ as a positively homogeneous mapping $D^* \Xi(\bar{x}, \bar{y})$. Set $x = \bar{x}$ in (10), the resulting weaker property is known as calmness of Ξ at $(\bar{x}, \bar{y})$ [18], which is used to derive the sensitivity analysis of the lower-level optimal solution mapping of the problem in the sequel.

2 Optimal value function reformulation for PSBPP

In this section, we shall briefly present the one-level formulation of the PSBPP (3), the detailed reformulation process can be found in Ref. [2]. For this purpose, we first use the scalarization technique to transform the problem (2) into an usual one-level scalar optimization problem, which consists of solving the following parametric problem:

$$\min_z f(x, y, z) = \langle y, f(x, z) \rangle$$
$$\text{s.t.} \quad g(x, z) \leqslant 0 \tag{12}$$

where the parameter y is a nonnegative point of the unit sphere, i.e.,

$$y \in Y = \{ y \in R^\nu \mid y \geqslant 0, \| y \| = 1 \} \tag{13}$$

For a given upper-level variable x, the weakly efficient solution set $\Psi_{\text{wef}}(x)$ of the lower-level problem (2) is not in general a singleton, hence it is difficult to choose the best point $z(x)$ on $\Psi_{\text{wef}}(x)$. Furthermore, we consider the set Y (13) as a new constraint set for the upper-level problem[9]. For all $(x, y) \in X \times Y$ (where $X := \{ x \in R^n \mid G(x) \leqslant 0 \}$), we denote by $\Psi(x, y)$ the solution set of the problem (12). When the weakly efficient solutions are considered for the lower-level problem (2), the relationship (see e.g. Ref. [19]) relates the solution set of this problem and that of (12) as follows.

Lemma 2.1 Assume the functions $g(x, \cdot)$ and $f(x, \cdot)$ are R^p_+-convex and R^ν_+-convex for all $x \in X$, respectively. Then $\Psi_{\text{wef}}(x) = \Psi(x, Y) := \bigcup \{ \Psi(x, y) \mid y \in Y \}$.

Hence, the PSBPP (3) can be replaced by the following classical pessimistic bilevel programming problem:

$$\min_x \max_y \max_z F(x, z)$$
$$\text{s.t.} \quad (x, y) \in X \times Y, z \in \Psi(x, y) \tag{14}$$

where the set Y (13) on the new parameter of the lower-level problem acts like additional upper-level constraints. The problem (14) can be reformulated as the following generalized minimax problem:

$$\min_x \max_y \{ \varphi_p(x, y) \mid y \in Y, x \in X \},$$

by defining the following maximization bilevel optimal value function (see e.g. Ref. [20])

$$\varphi_p(x, y) = \max_z \{ F(x, z) \mid z \in \Psi(x, y) \}.$$

Furthermore, the problem (14) can be further expressed as single-level optimization problem:

$$\min_x \{ \varphi_{pp}(x) \mid x \in X \} \tag{15}$$

if we define also the maximization another bilevel optimal value function by

$$\varphi_{pp}(x) = \max_y \{ \varphi_p(x, y) \mid y \in Y \}.$$

In the following, we give the theorem of the existence of the solution to the problem (14).

Lemma 2.2 If the set $\{(x, y, z) \mid (x, y) \in X \times Y, g(x, z) \leqslant 0\}$ is nonempty and compact, and for each $x \in X$, the Mangasarian-Fromowitz constraint qualification (MFCQ) holds. Suppose that the lower-level solution set mapping $\Psi(x, y)$ is lower semicontinuous at all points $(x, y) \in X \times Y$. Then, the problem (14) has an optimal solution.

The link between the PSBPP (3) and (14) will be given in the next result (also see Ref. [2, Proposition 3.1]).

Lemma 2.3 Consider the problems (3) and (2), where the lower-level constraint function $g(x, \cdot)$ is R^p_+-convex and $f(x, \cdot)$ is R^v_+-convex for all $x \in X$. Assume that Ψ is lower semicontinuous on $X \times Y$, then the following assertions hold.

(i) Let $(\bar{x}, \bar{z})$ be a local (resp. global) optimal solution of the problem (3). Then, for all $\bar{y} \in Y$ with $\bar{z} \in \Psi(\bar{x}, \bar{y})$, the point $(\bar{x}, \bar{y}, \bar{z})$ is a local (resp. global) optimal solution of the problem (14).

(ii) Let $(\bar{x}, \bar{y}, \bar{z})$ be a local (resp. global) optimal solution of the problem (14). Assume the set-valued mapping Ψ is closed-valued. Then $(\bar{x}, \bar{z})$ is a local (resp. global) optimal solution of the problem (3).

Now we give the optimal value function reformulation for the pessimistic bilevel programming problem (14) as follows.

$$
\begin{gathered}
\min \varphi_{pp}(x) \\
\text{s.t. } x \in X, \\
\begin{cases}
\varphi_{pp}(x) = \max\limits_{y}\{\varphi_p(x, y) \mid y \in Y\}, \\
\varphi_p(x, y) = \max\limits_{z}\{F(x, z) \mid z \in \Psi(x, y)\}, \\
\Psi(x, y) = \{z \in R^m \mid f(x, y, z) - \varphi(x, y) \leqslant 0, g(x, z) \leqslant 0\}, \\
\varphi(x, y) = \min\limits_{z}\{f(x, y, z) \mid g(x, z) \leqslant 0\}
\end{cases}
\end{gathered}
\tag{16}
$$

Based on this result, we will attempt to derive the necessary optimality conditions of PSBPP (3) via deriving those of the auxiliary problem (14). Obviously, if we set the minimization optimal value function as $\varphi^o_p(x, y) = \min\limits_{z}\{-F(x, z) \mid z \in \Psi(x, y)\}$, then, for all $(x, y) \in X \times Y$, we have

$$\varphi_p(x, y) = -\varphi^o_p(x, y) \tag{17}$$

The minimization of another optimal value function can be set as $\varphi^o_{pp}(x) = \min\limits_{y}\{-\varphi_p(x, y) \mid y \in Y\}$, then, for all $x \in X$, we also have

$$\varphi_{pp}(x) = -\varphi^o_{pp}(x) \tag{18}$$

By (17) and (18), we analyze $\varphi_p(x, y)$ and $\varphi_{pp}(x)$ via analyzing $\varphi^o_p(x, y)$ and $\varphi^o_{pp}(x)$, respectively. For these purposes, we consider the following a general 'abstract' framework of the marginal function:

$$\mu(x) = \min_{y}\{\Psi(x, y) \mid y \in \Xi(x)\} \tag{19}$$

where $\Psi: R^n \times R^m \to \overline{R}$ and $\Xi: R^n \to 2^{R^m}$. Denote the argminimum mapping in (19) by $\Xi_o(x) = \operatorname{argmin}\{\Psi(x, y) \mid y \in \Xi(x)\} = \{y \in \Xi(x) \mid \Psi(x, y) \leqslant \mu(x)\}$. We summarize in the next theorem some known results on the general value functions needed in the paper (see Refs. [16, Theorem 1.108] and [21, Theorem 5.2]).

Lemma 2.4 Let the value function μ given in (19) be finite at $\bar{x}$ with $\Xi_o \neq \emptyset$. The following assertions hold:

(i) Let Ξ_o be inner semicompact at $\bar{x}$, assume that Ξ is a closed-graph at $\bar{x}$, and that Ψ is lower semicontinuous on gph$G$ when $x = \bar{x}$. Then μ is lower semicontinuous at $\bar{x}$ and the upper bound for its basic subdifferential is given as follows:

$$\partial\mu(\bar{x}) \subset \{x^* \mid (x^*, 0) \in \bigcup_{\bar{y} \in \Xi_o(\bar{x})} \partial(\varphi(\bar{x}, \bar{y}) + \delta((\bar{x}, \bar{y}); \text{gph}\Xi))\}.$$

If in addition Ξ is Lipschitz-like around $(\bar{x}, \bar{y})$ for all vectors $\bar{y} \in \Xi_o(\bar{x})$, then we also have the Lipschitz continuity of μ around $\bar{x}$.

(ii) Let Ξ_o be inner semicontinuous at $(\bar{x}, \bar{y})$. Then μ is lower semicontinuous at $\bar{x}$ and the upper bound for its basic subdifferential is given as follows:

$$\partial\mu(\bar{x}) \subset \{x^* \mid (x^*, 0) \in \partial(\varphi(\bar{x}, \bar{y}) + \delta((\bar{x}, \bar{y}); \text{gph}\Xi))\}.$$

If in addition Ξ is Lipschitz-like around $(\bar{x},\bar{y})$, then μ is Lipschitz continuous around $\bar{x}$.

By Eqs. (17), (18) and (7), we easily have

$$\partial\varphi_p(x,y)\subset -\operatorname{co}\partial\varphi_p^o(x,y) \tag{20}$$

$$\partial\varphi_{pp}(x)\subset -\operatorname{co}\partial\varphi_{pp}^o(x) \tag{21}$$

By Lemma 2.4, we can estimate the upper bound of the subdifferential of the bilevel optimal value function $\varphi_p(x,y)$ (resp. $\varphi_{pp}(x)$) via estimating the subdifferential of $\partial\{\varphi_p^o\}(x,y)$ (resp. $\partial\{\varphi_{pp}^o\}(x)$). In the next section, based on specific structures of the set-valued mapping Ξ, our aim is to give detailed upper bounds for $D^*\Xi(\bar{x},\bar{y})$ in terms of the problem data. Verifiable rules for Ξ to be Lipschitz-like will also be provided. Further, we present the sensitivity analysis for the maximization bilevel optimal value function $\partial\varphi_p(x,y)$ and $\partial\varphi_{pp}(x)$. Based on these results, we develop the necessary optimality conditions for the problem (14) and thus for PSBPP (3).

3 Necessary optimality conditions for PSBPP

In this section, we shall establish the first-order necessary optimality conditions for PSBPP. For this goal, we need to develop the necessary optimality conditions for the optimal value function reformulation (16) of the problem (14). In the next, we first recall that the solution set mapping of the lower-level problem (12) as

$$\Psi(x,y)=\{z\in R^m \mid f(x,y,z)-\varphi(x,y)\leqslant 0,\ g(x,z)\leqslant 0\} \tag{22}$$

with φ denoting the optimal value function associated with the lower-level problem (12), i.e.,

$$\varphi(x,y)=\min_z\{f(x,y,z)\mid g(x,z)\leqslant 0\} \tag{23}$$

Here, we employ the lower-level value function approach[21] to the sensitivity analysis of the bilevel value function $\varphi_p^o(x,y)$. So we have the lower-level optimal value function reformulation of $\varphi_p^o(x,y)$ as $\varphi_p^o(x,y)=\min_z\{-F(x,z)\mid g(x,z)\leqslant 0, f(x,y,z)-\varphi(x,y)\leqslant 0\}$. Since the basic subdifferential $\partial\varphi$ does not satisfy the plus symmetry, an appropriate estimate of $\partial(-\varphi)$ is needed to proceed with this approach. By the well-known convex hull property (7), the estimate of $\partial(-\varphi)$ can be done.

In order to develop the sensitivity analysis of the negative value function in the lower-level problem (12), we first recall the nonsmooth counterparts of the lower- and upper-level regularity conditions, which are defined, respectively, as

$$\left.\begin{array}{l}\sum_{i=1}^{p}\beta_i x_i^*=0;\\ \beta_i\geqslant 0,\beta_i g_i(\bar{x},\bar{z})=0, i=1,\cdots,p;\\ (x_i^*,z_i^*)\in\bar{\partial}g_i(\bar{x},\bar{y}),\ i=1,\cdots,p\end{array}\right\}\Rightarrow \beta_i=0,\ i=1,\cdots,p \tag{24}$$

$$\left.\begin{array}{l}0\in\sum_{j=1}^{q}\alpha_j\partial G_j(\bar{x})=0;\\ \alpha_j\geqslant 0,\alpha_j G_j(\bar{x})=0, j=1,\cdots,q\end{array}\right\}\Rightarrow \alpha_j=0,\ j=1,\cdots,q \tag{25}$$

A particularity of the new constraint set Y (13), that the related Lagrange multipliers can be completely eliminated from the optimality conditions, is given in the next lemma.

Lemma 3.1[6, Lemma 4.2] The set of vectors $(\bar{x},\bar{y},\bar{z})\in R^n\times R^\nu\times R^m$, $\gamma,z_s\in R^\nu$ and $\mu,\lambda,\eta_s\in R$ with $s=1,\cdots,n+\nu+1$, satisfies the system

$$\begin{cases}\lambda f(\bar{x},\bar{z})-\lambda\sum_{s=1}^{n+\nu+1}\eta_s f(\bar{x},z_s)-\gamma+\mu\bar{y}=0,\\ \gamma\geqslant 0,\gamma^{\mathrm T}\bar{y}=0,\ \|\bar{y}\|=1,\end{cases}$$

if and only if the following inequality holds:

$$\lambda\Big\{\Big[\sum_{k=1}^{\nu}y_k\Big(f_k(\bar{x},\bar{z})-\sum_{s=1}^{n+\nu+1}\eta_s f_k(\bar{x},z_s)\Big)\Big]\bar{y}-\Big[f(\bar{x},\bar{z})-\sum_{s=1}^{n+\nu+1}\eta_s f(\bar{x},z_s)\Big]\Big\}\leqslant 0 \tag{26}$$

3.1 Sensitivity analysis of lower-level negative value function $\partial(-\varphi)$

In this subsection, we intend to develop the sensitivity analysis of the negative value function in the lower-level problem (12).

Theorem 3.1 The following assertions hold for the negation of the value function φ in (23).

(i) If f and g are Lipschitz continuous near the point $(\bar{x},z)$ and the solution map $\Psi(x,y)$ in (22) is inner semicompact at $(\bar{x},\bar{y})$ for all $(\bar{x},\bar{y},z)\in\operatorname{gph}\Psi$ satisfying (24), then φ is Lipschitz continuous around $(\bar{x},\bar{y})$, and the following inclusion holds:

$$\partial(-\varphi)(\bar{x},\bar{y})\subset\left\{\begin{bmatrix}-\sum_{s=1}^{n+\nu+1}\eta_s\left(\sum_{k=1}^{\nu}\bar{y}_k x_{ks}^*+\sum_{i=1}^{p}\beta_i^s w_{is}^*\right)\\ -\sum_{s=1}^{n+\nu+1}\eta_s f(\bar{x},z_s)\end{bmatrix}\;\middle|\;\begin{array}{l}\sum_{s=1}^{n+\nu+1}\eta_s=1, z_s\in\Psi(\bar{x},\bar{y}),\eta_s\geqslant 0,\ s=1,\cdots,n+\nu+1,\\ (x_{ks}^*,z_{ks}^*)\in\partial f_k(\bar{x},z_s),k=1,\cdots,\nu;\ s=1,\cdots,n+\nu+1,\\ (w_{is}^*,v_{is}^*)\in\partial g_i(\bar{x},z_s),i=1,\cdots,p;\ s=1,\cdots,n+\nu+1,\\ \sum_{k=1}^{\nu}\bar{y}_k z_{ks}^*+\sum_{i=1}^{p}\beta_i^s v_{is}^*=0,\ s=1,\cdots,n+\nu+1,\\ \beta_i^s\geqslant 0,\beta_i^s g_i(\bar{x},z_s)=0,i=1,\cdots,p;\ s=1,\cdots,n+\nu+1.\end{array}\right\}$$

(ii) If f and g are Lipschitz continuous near the point $(\bar{x},\bar{z})$ and assume that $(\bar{x},\bar{y},\bar{z})\in\mathrm{gph}\Psi$ with $\bar{x}\in\mathrm{dom}\varphi$ satisfying (24) and that either Ψ is inner semicontinuous at this point or f and g are convex. Then φ is Lipschitz continuous around $(\bar{x},\bar{y})$, and the following inclusion holds:

$$\partial(-\varphi)(\bar{x},\bar{y})\subset\left\{\begin{bmatrix}-\sum_{k=1}^{\nu}\bar{y}_k\hat{x}_k^*-\sum_{i=1}^{p}\beta_i\hat{w}_i^*\\ -f(\bar{x},\bar{z})\end{bmatrix}\;\middle|\;\begin{array}{l}\sum_{k=1}^{\nu}\bar{y}_k\hat{z}_k^*+\sum_{i=1}^{p}\beta_i\hat{v}_i^*=0,\\ (\hat{x}_k^*,\hat{z}_k^*)\in\bar{\partial}f_k(\bar{x},\bar{z}),k=1,\cdots,\nu,\\ (\hat{w}_i^*,\hat{v}_i^*)\in\bar{\partial}g_i(\bar{x},\bar{z}),i=1,\cdots,p,\\ \beta_i\geqslant 0,\beta_i g_i(\bar{x},\bar{z})=0,i=1,\cdots,p.\end{array}\right\}$$

Proof The local Lipschitz continuity of φ is justified from Ref. [18, Theorem 5.2] under the fulfillment of (24) in both the inner semicontinuity and inner semicompactness cases.

To prove the subdifferential inclusion of (i), firstly note that since for all $1,\cdots,\nu$, f_k is Lipschitz continuous near $(\bar{x},z)$, $z\in\Psi(\bar{x},\bar{y})$, by Ref. [6, Eq. (4.39)], while taking into account that $\bar{y}_k\geqslant 0$ for $k=1,\cdots,\nu$, the basic subdifferential of $f(\bar{x},\bar{y},\bar{z})$ can be expressed as the following inclusion relation:

$$\partial f(\bar{x},\bar{y},\bar{z})\subset f^\circ(\bar{x},\bar{z})+\left\{\sum_{k=1}^{\nu}\bar{y}_k(x_k^*,0,z_k^*)\mid(x_k^*,z_k^*)\in\partial f_k(\bar{x},\bar{z}),k=1,\cdots,\nu\right\}\tag{27}$$

where $f^\circ(\bar{x},\bar{z})=(\overbrace{0,\cdots,0}^{n\text{-times}},f(\bar{x},\bar{z}),\overbrace{0,\cdots,0}^{m\text{-times}})$.

In addition, we recall the estimate of $\partial\varphi(\bar{x},\bar{y})$ from Ref. [9, Theorem 7 (ii)] under the assumptions of (i) as follows:

$$\partial\varphi(\bar{x},\bar{y})\subset\bigcup_{\bar{z}\in\Psi(\bar{x},\bar{y})}\left\{(x^*,y^*)\;\middle|\;\begin{array}{l}(x^*,y^*,0)\in\partial f(\bar{x},\bar{y},\bar{z})+\sum_{i=1}^{p}\beta_i\partial_{x,y,z}g_i(\bar{x},\bar{z}),\\ \beta_i\geqslant 0,\beta_i g_i(\bar{x},\bar{z})=0,i=1,\cdots,p\end{array}\right\}\tag{28}$$

Combining inclusions (27) and (28), we have

$$\partial\varphi(\bar{x},\bar{y})\subset\bigcup_{\bar{z}\in\Psi(\bar{x},\bar{y})}\left\{\begin{bmatrix}\sum_{k=1}^{\nu}\bar{y}_k x_k^*+\sum_{i=1}^{p}\beta_i w_i^*\\ f(\bar{x},\bar{z})\end{bmatrix}\;\middle|\;\begin{array}{l}\sum_{k=1}^{\nu}\bar{y}_k z_k^*+\sum_{i=1}^{p}\beta_i v_i^*=0,\\ \beta_i\geqslant 0,\beta_i g_i(\bar{x},z)=0,\\ (x_k^*,z_k^*)\in\partial f_k(\bar{x},\bar{z}),\\ (w_i^*,v_i^*)\in\partial g_i(\bar{x},\bar{z}),\\ k=1,\cdots,\nu;\ i=1,\cdots,p\end{array}\right\}\tag{29}$$

The claimed estimate of $\partial(-\varphi)$ follows from here by combining (7) and Carathéodory's Theorem.

If Ψ is inner semicontinuous at $(\bar{x},\bar{y},\bar{z})$, we have the Clarke subdifferential of φ by Ref. [22, Theorem 5.4] that

$$\bar{\partial}\varphi(\bar{x},\bar{y}) \subset \bigcup \left\{ \begin{bmatrix} \sum_{k=1}^{l} \bar{y}_k \hat{x}_k^* + \sum_{i=1}^{p} \beta_i \hat{w}_i^* \\ f(\bar{x},\bar{z}) \end{bmatrix} \left| \begin{array}{l} \sum_{k=1}^{l} \bar{y}_k \hat{z}_k^* + \sum_{i=1}^{p} \beta_i \hat{v}_i^* = 0, \\ (\hat{x}_k^*, \hat{z}_k^*) \in \bar{\partial} f_k(\bar{x},\bar{z}), \\ (\hat{w}_i^*, \hat{v}_i^*) \in \bar{\partial} g_i(\bar{x},\bar{z}), \\ \beta_i \geqslant 0, \beta_i g_i(\bar{x},\bar{z}) = 0, \\ k = 1,\cdots,l; i = 1,\cdots,p. \end{array} \right. \right\}$$

This implies the subdifferential inclusion of (ii) by (6) and (7). This completes the proof.

3.2 Sensitivity analysis of the lower-level optimal solution set mapping Ψ

In this subsection, we shall present an upper estimate for the coderivative of the solution set mapping Ψ given in (22) and establish its Lipschitz-like property. For this purpose, we first present the calmness property. By Eq. (9), calculating the coderivative of Ψ, we must compute the limiting normal cone to the graph of Ψ:

$$\text{gph}\Psi = \{(x,y,z) \in K \mid f(x,y,z) - \varphi(x,y) \leqslant 0\} \quad \text{with } K = \{(x,y,z) \mid g(x,z) \leqslant 0\} \tag{30}$$

in terms of the initial data. To proceed this way by using the conventional results of the generalized differential calculus requires the fulfillment of the following basic qualification condition,

$$\partial(f-g)(\bar{x},\bar{y},\bar{z}) \cap (-N_k(\bar{x},\bar{y},\bar{z})) = \varnothing \tag{31}$$

However, it is shown in Ref. [23, Theorem 3.1] that condition (31) fails in common situations; in particular, when φ is locally Lipschitz around the point in question. The weaker assumption which can help circumventing this difficulty is given as follows:

$$\Phi(\nu) = \{(x,y,z) \in K \mid f(x,y,z) - \varphi(x,y) \leqslant \nu\} \text{ is calm at } (0,\bar{x},\bar{y},\bar{z}) \tag{32}$$

The condition (32) is automatically satisfied if f and g are linear. Furthermore, (32) holds at $(\bar{x},\bar{y},\bar{z})$ for the locally Lipschitzian function φ if we pass to the boundary of the normal cone in (31), that is, if the following qualification condition holds

$$\partial(f-g)(\bar{x},\bar{y},\bar{z}) \cap (-bdN_k(\bar{x},\bar{y},\bar{z})) = \varnothing \tag{33}$$

with K being semismooth, in particular, convex. The condition (33) seems to be especially effective for the so-called simple convex bilevel programming problems. For more details, the readers can be refer to Refs. [23-24]. It is deserved that for the latter case, the condition (33) can be further weakened by passing to the boundary of the subdifferential of f. For estimating the coderivative of Ψ, we present an additional qualification condition:

$$[(\lambda,\beta_i) \in \Lambda_z(\bar{x},\bar{y},\bar{z},0), x^* \in (-\varphi)(\bar{x},\bar{y})] \Rightarrow$$
$$\lambda x^* = \left\{ -(\lambda \sum_{k=1}^{l} \bar{y}_k x_k^* + \sum_{i=1}^{p} \beta_i w_i^*) \left| \begin{array}{l} (x_k^*, z_k^*) \in \partial f_k(\bar{x},\bar{z}), \\ (w_i^*, v_i^*) \in \partial g_i(\bar{x},\bar{z}), \\ k = 1,\cdots,\nu; i = 1,\cdots,p \end{array} \right. \right\} \tag{34}$$

where $\Lambda_z(\bar{x},\bar{y},\bar{z},z^*)$ is a particular multipliers set, that is,

$$\Lambda_z(\bar{x},\bar{y},\bar{z},z^*)=\left\{(\lambda,\beta_i)\left|\begin{array}{l}\lambda\geqslant 0,\ \beta_i\geqslant 0,\ \beta_i g_i(\bar{x},\bar{z})=0,\\ z^*+\lambda\sum_{k=1}^{\nu}\bar{y}_k z_k^*+\sum_{i=1}^{p}\beta_i v_i^*=0,\\ k=1,2,\cdots,\nu;\ i=1,2,\cdots,p\end{array}\right.\right\} \tag{35}$$

As it will be very clear in the proof of Theorem 3.2, they are sufficient conditions for the coderivative criterion (11) to hold for the set-valued mappings $K(x)=\{z\mid g(x,z)\leqslant 0\}$ and $\Psi(x,y)$ in (22), respectively, provided that the calmness condition (32) is satisfied. Under some mild conditions, the lower-level regularity condition (24) can ensure that the condition (35) holds at point $(\bar{x},\bar{y},\bar{z})$. More details on these types of conditions and more generally on the development of coderivatives can be found in Refs. [16,18]. In the following, we give a lower-level Lagrange multipliers set which will play an important role in the sequel.

$$\Lambda(\bar{x},\bar{y},\bar{z})=\left\{\gamma_i\geqslant 0\left|\begin{array}{l}\gamma_i g_i(\bar{x},\bar{z})=0,i=1,2,\cdots,p,\\ \sum_{k=1}^{\nu}\bar{y}_k z_k^*+\sum_{i=1}^{p}\gamma_i v_i^*=0\end{array}\right.\right\}.$$

In the following, we shall present the coderivative estimate and Lipschitz-like property of lower-level solution set mapping Ψ.

Theorem 3.2 (i) For all $(\bar{x},\bar{y},z)\in \mathrm{gph}\Psi$, let the conditions (24) and (32) hold at this point, and let Ψ (22) be inner semicompact at $(\bar{x},\bar{y})$. Then, for all $z\in R^m$, (33) and the following inclusion hold:

$$D^*\Psi(\bar{x},\bar{y},\bar{z})(z)\subset\bigcup_{z_s\in\Psi(\bar{x},\bar{y})}\ \bigcup_{(\lambda,\beta_i)\in\Lambda_z(\bar{x},\bar{y},\bar{z},z)}\ \bigcup_{\beta_i^s\in\Lambda(\bar{x},\bar{y},\bar{z})}\left\{\begin{array}{l}\lambda\Big(\sum_{k=1}^{\nu}\bar{y}_k x_k^*-\sum_{s=1}^{n+\nu+1}\eta_s\big(\sum_{k=1}^{\nu}\bar{y}_k x_{ks}^*+\sum_{i=1}^{p}\beta_i^s w_{is}^*\big)\Big)+\sum_{i=1}^{p}\beta_i w_i^*,\\ \lambda f(\bar{x},z)-\lambda\sum_{s=1}^{n+\nu+1}\eta_s f(\bar{x},z_s),\\ \sum_{s=1}^{n+\nu+1}\eta_s=1,\eta_s\geqslant 0,s=1,\cdots,n+\nu+1\end{array}\right\}$$

with $\sum_{s=1}^{n+\nu+1}\eta_s=1$ and $\eta_s\geqslant 0$ for $s=1,\cdots,n+\nu+1$. If in addition the condition (34) holds at $(\bar{x},\bar{y},z_s)$, then Ψ is Lipschitz-like around this point.

(ii) Let the solution map Ψ (22) be inner semicontinuous at $(\bar{x},\bar{y},\bar{z})\in \mathrm{gph}\Psi$, and let the qualification conditions (24) and (32) hold at this point. Then, for all $z^*\in R^m$ we have

$$D^*\Psi(\bar{x},\bar{y},\bar{z})(z^*)\subset\bigcup_{(\lambda,\beta_i)\in\Lambda_z(\bar{x},\bar{y},\bar{z},z^*)}\ \bigcup_{\gamma_i\in\Lambda(\bar{x},\bar{y},\bar{z})}\left\{\lambda\sum_{k=1}^{\nu}\bar{y}_k x_k^*-\lambda\Big(\sum_{k=1}^{\nu}\bar{y}_k\hat{x}_k^*+\sum_{i=1}^{p}\gamma_i\hat{w}_i^*\Big)+\sum_{i=1}^{p}\beta_i w_i^*\right\} \tag{36}$$

If in addition the condition (34) holds at $(\bar{x},\bar{y},\bar{z})$, then Ψ is Lipschitz-like around this point.

Proof We first show the proof for (i). It follows from Theorem 3.1(i) that the lower-level function φ is Lipschitz continuous around $(\bar{x},\bar{y})$ under assumption condition (24) and the inner semicompactness assumptions. If we add the calmness property (32), then we have

$$N_{\mathrm{gph}\Psi}(\bar{x},\bar{y},\bar{z})\subset\bigcup_{z_s\in\Psi(\bar{x},\bar{y})}\bigcup_{\lambda\geqslant 0}\{\lambda(\partial f(\bar{x},\bar{y},\bar{z})+\partial(-\varphi)(\bar{x},\bar{y})\times\{0\})+N_k(\bar{x},\bar{y},\bar{z})\},$$

by Ref. [25, Theorem 4.1] taking into account that the constraint $f(x,y,z)-\varphi(x,y)\leqslant 0$ is working at point $(\bar{x},\bar{y},\bar{z})$. By (30), we have

$$N_k(\bar{x},\bar{y},\bar{z})=\Big\{\sum_{i=1}^{p}\beta_i v_i^* \mid \beta_i \geqslant 0,\ \beta_i g_i(\bar{x},\bar{z})=0, i=1,\cdots,p\Big\} \tag{37}$$

which holds under the validity of the condition (24) at point $(\bar{x},\bar{y},z_s)$ for $s=1,\cdots,n+\nu+1$. Combining the definition of the coderivative (9), we derive the coderivative estimate (35). Further, by Eq. (35) and the coderivative criterion (11) for the Lipschitz-like property, the coderivative criterion holds provided that

$$\left.\begin{array}{l} x^* \in \lambda(\sum_{k=1}^{\nu}\bar{y}_k x_k^* + \partial(-\varphi)(\bar{x},\bar{y})) + \sum_{i=1}^{p}\beta_i w_i^*, \\ (\lambda,\beta_i)\in \Lambda_z(\bar{x},\bar{y},\bar{z},0) \end{array}\right\} \Rightarrow x^*=0 \tag{38}$$

Next we begin to prove (ii). According to Theorem 3.1(ii), the lower-level function φ is Lipschitz continuous around $(\bar{x},\bar{y})$ under the condition (24) and the inner semicontinuous assumptions. If we add the calmness property (32), then we have

$$N_{\mathrm{gph}\Psi}(\bar{x},\bar{y},\bar{z})\subset\bigcup_{\lambda\geqslant 0}\{\lambda(\partial f(\bar{x},\bar{y},\bar{z})+\partial(-\varphi)(\bar{x},\bar{y})\times\{0\})+N_k(\bar{x},\bar{y},\bar{z})\},$$

by Ref. [25, Theorem 4.1] taking into account that the constraint $f(x,y,z)-\varphi(x,y)\leqslant 0$ is working at point $(\bar{x},\bar{y},\bar{z})$. By (30), the equality (37) holds. Combining the definition of the coderivative (9), we derive the coderivative estimate (36). Further, by (36) and the coderivative criterion (11) for the Lipschitz-like property, the coderivative criterion holds provided that Eq. (38) holds. This completes the proof.

Remark 3.1 If the functions f and g are convex with respect to z, the inner semicontinuity of Ψ can be dropped in Theorem 3.2(ii).

3.3 Sensitivity analysis of maximization bilevel optimal value functions φ_p and φ_{pp}

For the sake of simplicity, we first define the upper-level optimal solution set mapping as follows:

$$\Psi_o(x,y)=\{z\in\Psi(x,y)\mid -F(x,z)-\varphi_p^o(x,y)\leqslant 0\} \tag{39}$$

In the rest of this paper, we always assume that the set $\Psi_o(x,y)$ is nonempty. The following results illustrate the local sensitivity analysis of the bilevel value function φ_p.

Theorem 3.3 Considering φ_p defined in Section 2 and (39), the following assertions hold:

(i) Assume that Ψ_o is inner semicompact at $(\bar{x},\bar{y})$, the condition (24) holds at $(\bar{x},\bar{y},z)\in \mathrm{gph}\Psi$, while the condition (32) holds at $(\bar{x},\bar{y},z)$ for all $z\in\Psi_o(\bar{x},\bar{y})$. Then the following inclusion holds:

$$\partial\varphi_p(\bar{x},\bar{y})\subset\bigcup_{z_t\in\Psi(\bar{x},\bar{y})}\bigcup_{z_s\in\Psi(\bar{x},\bar{y})}\bigcup_{(\lambda_t,\beta_i^t)\in\Lambda_z(\bar{x},\bar{y},\bar{z},z_t)}\bigcup_{\beta_i^s\in\Lambda(\bar{x},\bar{y},\bar{z})}\left\{(x_t^*,y_t^*)\ \middle|\ \begin{array}{l} x_t^*\in\sum_{t=1}^{n+\nu+1}\rho_t\{x_{Ft}^*-\lambda_t(\sum_{k=1}^{\nu}\bar{y}_k x_{kt}^*-\sum_{s=1}^{n+\nu+1}\eta_s(\sum_{k=1}^{\nu}\bar{y}_k x_{ks}^*+\sum_{i=1}^{p}\beta_i^s w_{is}^*))+\sum_{i=1}^{p}\beta_i^t w_{it}^*\}; \\ y_t^*\in\lambda_t f(\bar{x},z_t)-\lambda_t\sum_{s=1}^{n+\nu+1}\eta_s f(\bar{x},z_s); \\ (x_{Ft}^*,0)\in\partial F(\bar{x},\bar{z}), \text{for } t=1,\cdots,n+\nu+1; \\ \sum_{t=1}^{n+\nu+1}\rho_t=1, \rho_t\geqslant 0, \text{ for } t=1,\cdots,n+\nu+1; \\ \sum_{s=1}^{n+\nu+1}\eta_s=1, \eta_s\geqslant 0, \text{ for } s=1,\cdots,n+\nu+1 \end{array}\right\}$$

If in addition (34) is satisfied at $(\bar{x},\bar{y},z)$ for all $z \in \Psi_o(\bar{x},\bar{y})$, then φ_p is Lipschitz continuous around $(\bar{x},\bar{y})$.

(ii) Assume that Ψ_o is inner semicontinuous at $(\bar{x},\bar{y},\bar{z})$, the conditions (24) and (32) hold at this point. Furthermore, assume that the set $\mathrm{co}N_{\mathrm{gph}\Psi}(\bar{x},\bar{y},\bar{z})$ is closed. Then the following inclusion holds:

$$\partial\varphi_p(\bar{x},\bar{y}) \subset \bigcup_{(\lambda_t,\beta_i^t)\in\Lambda_z(\bar{x},\bar{y},\bar{z},z^*)} \bigcup_{\gamma_i^t\in\Lambda(\bar{x},\bar{y},\bar{z})} \left\{ \begin{array}{l} x_F^* - \sum_{t=1}^{n+\nu+1}\rho_t\{\lambda_t\sum_{k=1}^{\nu}\bar{y}_k x_{kt}^* - \lambda_t(\sum_{k=1}^{\nu}\bar{y}_k\hat{x}_{kt}^* + \sum_{i=1}^{p}\gamma_i^t\hat{w}_{it}^*) + \sum_{i=1}^{p}\beta_i^t w_{it}^*\} \mid \\ (x_F^*,0)\in\partial F(\bar{x},\bar{z}), \sum_{t=1}^{n+\nu+1}\rho_t = 1, \rho_t \geqslant 0, \\ \text{for } t=1,\cdots,n+\nu+1 \end{array} \right\} \tag{40}$$

If in addition (34) is satisfied at point $(\bar{x},\bar{y},\bar{z})$, then φ_p is Lipschitz continuous around $(\bar{x},\bar{y})$.

Proof We first provide the proof of (i). To justify (i), by Lemma 2.4(i) and subdifferential chain rules and related calculus in Ref. [16], we have

$$\partial\varphi_p^o(\bar{x},\bar{y}) \subset \bigcup_{z_t\in\Psi_o(\bar{x},\bar{y})} \left\{ \begin{pmatrix} \partial_x(-F(\bar{x},z_t)) \\ \partial_y(-F(\bar{x},z_t)) \end{pmatrix} + D^*\Psi(\bar{x},\bar{y},z_t)(\partial_z(-F(\bar{x},z_t))) \right\}$$

under the inner semicompactness assumption on Ψ_o. Since $\Psi_o(x,y)\subset\Psi(x,y)$ for all $(x,y)\in X\times Y$, the lower-level optimal solution map Ψ in (22) is also inner semicompact at $(\bar{x},\bar{y},z_t)\in\{\mathrm{gph}\Psi_o\}$. Hence, by the subdifferential of the lower-level negation value function $-\varphi$ in Theorem 3.1(i) and the coderivative of Ψ in Theorem 3.2 (i), combining with (20) and Carathéodory's Theorem, we can derive the upper estimate of $\varphi_p(\bar{x},\bar{y})$.

To prove the local Lipschitz continuity of $\varphi_p(\bar{x},\bar{y})$ in (i) under the condition (34), the latter condition implies the Lipschitz-like property of Ψ around $(\bar{x},\bar{y},z_t)$. Thus the desired result is obtained from Lemma 2.4(i).

For justifying (ii), since the function F is Lipschitz continuous and Ψ_o is inner semicontinuous at $(\bar{x},\bar{y},\bar{z})$, we get[22, Theorem 5.1]

$$\bar{\partial}\varphi_p^o(\bar{x},\bar{y}) \subset \left\{ \begin{pmatrix} \bar{\partial}_x(-F(\bar{x},\bar{z})) \\ \bar{\partial}_y(-F(\bar{x},\bar{z})) \end{pmatrix} + \bar{D}^*\Psi(\bar{x},\bar{y},\bar{z})(\bar{\partial}_z(-F(\bar{x},\bar{z}))) \right\}.$$

Combining (36) and Carathéodory's Theorem, we can derive the estimation for the coderivative $\bar{D}^*\Psi(\bar{x},\bar{y},z_t)(\bar{\partial}_z(-F(\bar{x},\bar{z})))$:

$$\bar{D}^*\Psi(\bar{x},\bar{y},\bar{z})(\bar{\partial}_z(-F(\bar{x},\bar{z}))) \subset \bigcup_{(\lambda_t,\beta_i^t)\in\Lambda_z(\bar{x},\bar{y},\bar{z},z^*)} \bigcup_{\gamma_i^t\in\Lambda(\bar{x},\bar{y},\bar{z})} \left\{ \begin{array}{l} \sum_{t=1}^{n+\nu+1}\rho_t\{\lambda_t\sum_{k=1}^{\nu}\bar{y}_k x_{kt}^* - \\ \lambda_t(\sum_{k=1}^{\nu}\bar{y}_k\hat{x}_{kt}^* + \sum_{i=1}^{p}\gamma_i^t\hat{w}_{it}^*) + \\ \sum_{i=1}^{p}\beta_i^t w_{it}^*\} \mid \sum_{t=1}^{n+\nu+1}\rho_t = 1, \\ \rho_t \geqslant 0, \\ \text{for } t=1,\cdots,n+\nu+1 \end{array} \right\}.$$

The latter inclusion implies that $\overline{N}_{\mathrm{gph}\Psi}(\bar{x},\bar{y},\bar{z}) = \mathrm{co}N_{\mathrm{gph}\Psi}(\bar{x},\bar{y},\bar{z})$ provided that the set $\mathrm{co}N_{\mathrm{gph}\Psi}(\bar{x},\bar{y},\bar{z})$ is closed. Combining the above two results, by (6) and (20), we can justify (40). To justify the local Lipschitz continuity of $\varphi_p(\bar{x},\bar{y})$ in (ii) under the condition (34), the latter condition implies the Lipschitz-like property of Ψ around $(\bar{x},\bar{y},z_t)$, for $t=1,\cdots,n+\nu+1$. This completes the proof.

In the following, we shall develop a local sensitivity analysis of the maximization bilevel optimal value function $\varphi_{pp}(x)$ in (15). For this purpose, we need to estimate $\partial\varphi_{pp}^o(x)$. Combining (18) and the conclusion on the sensitivity analysis for the value function of the nonparametric minimax problem (see Ref. [26, Lemma 3.3]), we have $\partial\varphi_{pp}^o(x) \subseteq \partial(-\varphi_p)(x,y) = \partial\varphi_p^o(x,y)$. By (21), we derive that $\partial\varphi_{pp}(x) \subset -\mathrm{co}\partial\varphi_p^o(x,y)$ and $\varphi_{pp}(x)$ is Lipschitz continuous around $\bar{x}$ under the corresponding conditions of Theorem 3.3.

Based on the above analysis, we present the estimation of the subdifferential of the maximization bilevel optimal value function $\varphi_{pp}(x)$ in (15).

Corollary 3.1 Considering the definition of $\varphi_{pp}(x)$ and (39), the following assertions hold:

(i) Assume that Ψ_o is inner semicompact at $(\bar{x},\bar{y})$, the condition (24) holds at $(\bar{x},\bar{y},z) \in \mathrm{gph}\Psi$, while the condition (32) holds at $(\bar{x},\bar{y},z)$ for all $z \in \Psi_o(\bar{x},\bar{y})$. Then the following inclusion holds:

$$\partial\varphi_{pp}(\bar{x}) \subset \bigcup_{z_t\in\Psi(\bar{x},\bar{y})}\ \bigcup_{z_s\in\Psi(\bar{x},\bar{y})}\ \bigcup_{(\lambda_t,\beta_i^t)\in\Lambda_z(\bar{x},\bar{y},\bar{z},z_t)}\ \bigcup_{\beta_i^s\in\Lambda(\bar{x},\bar{y},\bar{z})}$$

$$\left\{(x_t^*,y_t^*)\ \middle|\ \begin{array}{l} x_t^* \in \sum\limits_{t=1}^{n+\nu+1}\rho_t\{x_{Ft}^* - \lambda_t(\sum\limits_{k=1}^{\nu}\bar{y}_k x_{kt}^* - \sum\limits_{s=1}^{n+\nu+1}\eta_s(\sum\limits_{k=1}^{\nu}\bar{y}_k x_{ks}^* + \sum\limits_{i=1}^{p}\beta_i^s w_{is}^*)) + \sum\limits_{i=1}^{p}\beta_i^t w_{it}^*\};\\ y_t^* \in \lambda_t f(\bar{x},z_t) - \lambda_t\sum\limits_{s=1}^{n+\nu+1}\eta_s f(\bar{x},z_s);\\ (x_{Ft}^*,0) \in \partial F(\bar{x},\bar{z}),\text{for } t=1,\cdots,n+\nu+1;\\ \sum\limits_{t=1}^{n+\nu+1}\rho_t = 1,\rho_t \geqslant 0,\ \text{for } t=1,\cdots,n+\nu+1;\\ \sum\limits_{s=1}^{n+\nu+1}\eta_s = 1,\eta_s \geqslant 0,\ \text{for } s=1,\cdots,n+\nu+1. \end{array}\right\}$$

If in addition (34) is satisfied at $(\bar{x},\bar{y},z)$ for all $z \in \Psi_o(\bar{x},\bar{y})$, then φ_{pp} is Lipschitz continuous around $\bar{x}$.

(ii) Assume that Ψ_o is inner semicontinuous at $(\bar{x},\bar{y},\bar{z})$, the conditions (24) and (32) hold at this point. Furthermore, assume that the set $\mathrm{co}N_{\mathrm{gph}\Psi}(\bar{x},\bar{y},\bar{z})$ is closed. Then the following inclusion holds:

$$\partial\varphi_{pp}(\bar{x}) \subset \bigcup_{(\lambda_t,\beta_i^t)\in\Lambda_z(\bar{x},\bar{y},\bar{z},z^*)}\ \bigcup_{\gamma_i^t\in\Lambda(\bar{x},\bar{y},\bar{z})} \left\{\begin{array}{l} x_F^* - \sum\limits_{t=1}^{n+\nu+1}\rho_t\{\lambda_t\sum\limits_{k=1}^{\nu}\bar{y}_k x_{kt}^* - \lambda_t(\sum\limits_{k=1}^{\nu}\bar{y}_k\hat{x}_{kt}^* + \sum\limits_{i=1}^{p}\gamma_i^t\hat{w}_{it}^*) + \sum\limits_{i=1}^{p}\beta_i^t w_{it}^*\}\ \mid\\ (x_F^*,0) \in \partial F(\bar{x},\bar{z}),\ \sum\limits_{t=1}^{n+\nu+1}\rho_t = 1,\rho_t \geqslant 0,\\ \text{for } t=1,\cdots,n+\nu+1 \end{array}\right\}$$

If in addition (34) is satisfied at $(\bar{x},\bar{y},\bar{z})$, then φ_{pp} is Lipschitz continuous around $\bar{x}$.

3.4 Necessary optimality conditions for the problem (14) and PSBPP

In this subsection, we shall establish the necessary optimality conditions for the optimal value reformulation (16) of the minimax problem (14) in the nonsmooth setting using the above sensitivity analysis results. Furthermore, with the help of Lemma 2.3, we present the necessary optimality conditions for the PSBPP with Lipschitz continuous data.

Theorem 3.4 Let $(\bar{x},\bar{y})$ be an upper-level regular local optimal solution to the problem (14), whereas F and G are Lipschitz continuous at $(\bar{x},\bar{z})$ and $\bar{x}$, respectively. Let $X\times Y$ be closed. Then, the following assertions hold:

(i) Let Ψ_o be inner semicompact at $(\bar{x},\bar{y})$ while for all $z\in\Psi(\bar{x},\bar{y})$ and the point $(\bar{x},z)$ is lower-level regular, let f and g be Lipschitz continuous at $(\bar{x},z)$, $z\in\Psi(\bar{x},\bar{y})$, and let the conditions (32) and (34) be satisfied at all points $(\bar{x},\bar{y},z)$ with $z\in\Psi_o(\bar{x},\bar{y})$. Then there exist $\lambda_t\geqslant 0$, α, β^t, β^s, η_s, ρ_t, $(x^*_{Ft},z^*_{Ft})\in\partial F(\bar{x},\bar{z})$, with $t=1,\cdots,n+\nu+1$ and $z_t,z_s\in\Psi(\bar{x},\bar{y})$ with $t,s=1,\cdots,n+\nu+1$ such that (26) and the following conditions hold:

$$\left\{\begin{array}{l}\sum_{t=1}^{n+\nu+1}\rho_t\{x^*_{Ft}-\lambda_t(\sum_{k=1}^{\nu}\bar{y}_k x^*_{kt}-\sum_{s=1}^{n+\nu+1}\eta_s(\sum_{k=1}^{\nu}\bar{y}_k x^*_{ks}+\sum_{i=1}^{p}\beta^s_i w^*_{is}))+\sum_{i=1}^{p}\beta^t_i w^*_{it}\}+\sum_{j=1}^{q}\alpha_j x^*_{Gj}=0,\\ \text{for } t=1,\cdots,n+\nu+1:\ z^*_{Ft}-\lambda_t\sum_{k=1}^{\nu}\bar{y}_k z^*_{kt}-\sum_{i=1}^{p}\beta^t_i v^*_{it}=0,\\ \text{for } s=1,\cdots,n+\nu+1:\ \sum_{k=1}^{\nu}\bar{y}_k z^*_{ks}+\sum_{i=1}^{p}\beta^s_i v^*_{is}=0,\\ \text{for } j=1,\cdots,q:\ \alpha_j\geqslant 0,\ \alpha_j G_j(\bar{x})=0,\\ \text{for } t=1,\cdots,n+\nu+1;\ i=1,\cdots,p:\ \beta^t_i\geqslant 0,\beta^t_i g_i(\bar{x},z_t)=0,\\ \text{for } s=1,\cdots,n+\nu+1;\ i=1,\cdots,p:\ \beta^s_i\geqslant 0,\beta^s_i g_i(\bar{x},z_s)=0,\\ \text{for } t=1,\cdots,n+\nu+1:\ \rho_t\geqslant 0,\ \sum_{t=1}^{n+\nu+1}\rho_t=1,\\ \text{for } s=1,\cdots,n+\nu+1:\ \eta_s\geqslant 0,\ \sum_{s=1}^{n+\nu+1}\eta_s=1,\\ \text{where we have the following inclusions:}\\ \text{for } t=1,\cdots,n+\nu+1:\ (x^*_{Ft},z^*_{Ft})\in\partial F(\bar{x},\bar{z}),\\ \text{for } j=1,\cdots,q:\ x^*_{Gj}\in\partial G_j(\bar{x}),\\ \text{for } k=1,\cdots,\nu;\ t=1,\cdots,n+\nu+1:\ (x^*_{kt},z^*_{kt})\in\partial F(\bar{x},z_t),\\ \text{for } k=1,\cdots,\nu;\ s=1,\cdots,n+\nu+1:\ (x^*_{ks},z^*_{ks})\in\partial F(\bar{x},z_s),\\ \text{for } i=1,\cdots,p;\ t=1,\cdots,n+\nu+1:\ (w^*_{it},v^*_{it})\in\partial g_i(\bar{x},z_t),\\ \text{for } i=1,\cdots,p;\ s=1,\cdots,n+\nu+1:\ (w^*_{is},v^*_{is})\in\partial g_i(\bar{x},z_s)\end{array}\right. \tag{41}$$

The relationships (26) and (41) considered together are called the KM-stationarity conditions.

(ii) Let Ψ_o be inner semicontinuous at $(\bar{x},\bar{y},\bar{z})$, $(\bar{x},\bar{z})$ be lower-level regular, f and g be Lipschitz continuous at $(\bar{x},\bar{z})$, and let the conditions (32) and (34) be satisfied at $(\bar{x},\bar{y},\bar{z})$ and the set $\mathrm{co}N_{\mathrm{gph}\Psi}(\bar{x},\bar{y},\bar{z})$ be closed. Then there exist $(x^*_F,z^*_F)\in\partial F(\bar{x},\bar{z})$, $\lambda_t\geqslant 0$, α, β^t and γ^t such that the following conditions hold:

$$\left\{\begin{array}{l} x_F^* - \sum_{t=1}^{n+\nu+1}\rho_t\{\lambda_t\sum_{k=1}^{\nu}\bar{y}_k x_{kt}^* - \lambda_t(\sum_{k=1}^{\nu}\bar{y}_k\hat{x}_{kt}^* + \sum_{i=1}^{p}\gamma_i^t\hat{w}_{is}^*) + \sum_{i=1}^{p}\beta_i^t w_{it}^*\} + \sum_{j=1}^{q}\alpha_j x_{Gj}^* = 0, \\ \text{for } t=1,\cdots,n+\nu+1:\ z_F^* - \lambda_t\sum_{k=1}^{\nu}\bar{y}_k z_{kt}^* - \sum_{i=1}^{p}\beta_i^t v_{it}^* = 0, \\ \text{for } t=1,\cdots,n+\nu+1:\ \sum_{k=1}^{\nu}\bar{y}_k\hat{z}_{kt}^* + \sum_{i=1}^{p}\gamma_i^t\hat{v}_{it}^* = 0, \\ \text{for } j=1,\cdots,q:\ \alpha_j \geqslant 0,\ \alpha_j G_j(\bar{x}) = 0, \\ \text{for } t=1,\cdots,n+\nu+1;\ i=1,\cdots,p:\ \gamma_i^t \geqslant 0, \gamma_i^t g_i(\bar{x},\bar{z}) = 0, \\ \text{for } t=1,\cdots,n+\nu+1;\ i=1,\cdots,p:\ \beta_i^t \geqslant 0, \beta_i^t g_i(\bar{x},\bar{z}) = 0, \\ \text{for } t=1,\cdots,n+\nu+1:\ \rho_t \geqslant 0,\ \sum_{t=1}^{n+\nu+1}\rho_t = 1, \\ \text{where we have the following inclusions:} \\ (x_F^*, z_F^*) \in \partial F(\bar{x},\bar{z}), \\ \text{for } j=1,\cdots,q:\ x_{Gj}^* \in \partial G_j(\bar{x}), \\ \text{for } k=1,\cdots,\nu;\ t=1,\cdots,n+\nu+1:\ (x_{kt}^*, z_{kt}^*) \in \partial F(\bar{x},\bar{z}), \\ \text{for } k=1,\cdots,\nu;\ t=1,\cdots,n+\nu+1:\ (\hat{x}_{kt}^*, \hat{z}_{kt}^*) \in \bar{\partial} F(\bar{x},\bar{z}), \\ \text{for } i=1,\cdots,p;\ t=1,\cdots,n+\nu+1:\ (w_{it}^*, v_{it}^*) \in \partial g_i(\bar{x},\bar{z}), \\ \text{for } i=1,\cdots,p;\ t=1,\cdots,n+\nu+1:\ (\hat{w}_{it}^*, \hat{v}_{it}^*) \in \bar{\partial} g_i(\bar{x},\bar{z}) \end{array}\right. \tag{42}$$

The relationships in (42) are called the KN-stationarity conditions.

Proof Under the assumptions of (ii), the bilevel value function φ_{pp} is Lipschitz continuous near $\bar{x}$. Since X is closed, one has from Ref. [16, Proposition 5.3] that $0 \in \partial\varphi_{pp}(\bar{x}) + N_x(\bar{x})$. By the inner semicontinuity of Ψ_o at $(\bar{x},\bar{y},\bar{z})$ and the upper regularity (25), we have

$$N_x(\bar{x}) = \Big\{\sum_{j=1}^{q}\alpha_j\partial G_j(\bar{x}) \mid \alpha_j \geqslant 0, \alpha_j G_j(\bar{x}) = 0, j=1,\cdots,q\Big\} \tag{43}$$

Combining Corollary 3.1(ii) and (43), Theorem 3.4(ii) is easily derived. If Ψ_o is inner semicompact around $(\bar{x},\bar{y})$, the condition (26) holds at all point $(\bar{x},\bar{y},z)$ with $z \in \Psi(\bar{x},\bar{y})$, and that (32) and (34) are satisfied at all point $(\bar{x},\bar{y},z)$ with $z \in \Psi_o(\bar{x},\bar{y})$. Thus, by Corollary 3.1(i), we obtain the conclusion (i). This completes the proof.

Remark 3.2 The prefixes 'KN' and 'KM' in Theorem 3.4 reflect the difference between the KKT-type optimality conditions via the inner semicompactness and inner semicontinuity of the upper-level optimal solution set mapping Ψ_o, respectively. For the notions 'KM-stationarity' and 'KN-stationarity', the readers can refer to Ref. [27]. Under the inner semicontinuity of the lower-level optimal set valued mapping Ψ, the necessary optimality conditions (ii) in Theorem 3.4 are in fact those of the problem:

$$\begin{array}{ll} \min_x \max_y \max_z F(x,z) \\ \text{s.t.} \quad x \in X, z \in \Psi(x,y). \end{array}$$

This means that the above framework, the constraints described by Y (13) can be dropped and the condition that set $X \times Y$ is closed is reduced to that the set X is closed, the latter is immediately reached by Assumption 0.1, while deriving the necessary optimality conditions of PSBPP.

By Lemma 2.3(i) and Theorem 3.4, the necessary optimality conditions for PSBPP are derived when the involved functions are locally Lipschitz continuous.

Corollary 3.2 Let $(\bar{x},\bar{z})$ be a local optimal solution of problem (3), where F and G_j, $j=1$,

$\cdots, q$, are Lipschitz continuous at $(\bar{x}, \bar{z})$ and $\bar{z}$, respectively. For all $x \in X$, $f(x, \cdot)$ and $g(x, \cdot)$ are R_+^ν- and R_+^p-convex, respectively. Let $\bar{x}$ be upper-level regular. Then, the following assertions hold:

(i) (KM-stationarity conditions) Let Ψ_o be inner semicompact at $(\bar{x}, \bar{y})$ while for all $z \in \Psi(\bar{x}, \bar{y})$ and the point $(\bar{x}, z)$ are lower-level regular. Let f and g be Lipschitz continuous at $(\bar{x}, z)$, $z \in \Psi(\bar{x}, \bar{y})$, and let the conditions (32) and (34) be satisfied at all points $(\bar{x}, \bar{y}, z)$ with $z \in \Psi_o(\bar{x}, \bar{y})$ and the set $X \times Y$ be closed. Then there exist $\lambda_t \geqslant 0$, α, β^t, β^s, η_s, ρ_t, $(x_{Ft}^*, z_{Ft}^*) \in \partial F(\bar{x}, \bar{z})$, with $t = 1, \cdots, n + \nu + 1$ and $z_s \in \Psi(\bar{x}, \bar{y})$ with $t, s = 1, \cdots, n + \nu + 1$ such that (26) and (41) hold.

(ii) (KN-stationarity conditions) Let Ψ_o be inner semicontinuous at $(\bar{x}, \bar{y}, \bar{z})$ and the point $(\bar{x}, \bar{z})$ be lower-level regular. Let f and g be Lipschitz continuous at $(\bar{x}, \bar{z})$, $X \times Y$ be closed, and let the conditions (32) and (34) be satisfied at $(\bar{x}, \bar{y}, \bar{z})$ and the set $\mathrm{co}N_{\mathrm{gph}\Psi}(\bar{x}, \bar{y}, \bar{z})$ be closed. Then there exist $(x_F^*, z_F^*) \in \partial F(\bar{x}, \bar{z})$, $\lambda_t \geqslant 0$, α, β^t and γ^t such that (42) holds.

4 Conclusion

In this paper, we present the first-order necessary optimality conditions for the PSBPP with nonsmooth data. With the help of the scalarization method and optimal value function reformulation, we first transform the original problem into a generalized minimax optimization problem with constraints. Then, by using the variational analysis and generalized differential calculus of Mordukhovich, the first-order necessary optimality conditions are established for PSBPP in the nonsmooth setting. Our results enrich greatly the optimization theory ofbilevel programming problems and especially for PSBPP. Furthermore, our results may inspire the design of new solving methods for PSBPP and extensively expand its application fields. In fact, all the results in this paper can easily be extended to the more general operator constraints in the sense of Mordukhovich. In addition, it is also worth mentioning that the first-order necessary optimality conditions for PSBPP, in which all the functions involved are fully convex with respect to their variables, can be established based on our results.

Acknowledgments The authors are grateful for the thorough remarks of the anonymous referees, whose comments led to an improved version of this paper.

References

[1] CALVETE H I, GALÉ C. On linear bilevel problems with multiple objectives at the lower level[J]. Omega, 2011, 39(1): 33-40.

[2] LIU B, WAN Z, CHEN J, et al. Optimality conditions for pessimistic semivectorial bilevel programming problems[J]. Journal of Inequalities and Applications, 2014, 2014: 41.

[3] BONNEL H. Optimality conditions for the semivectorial bilevel optimization problem[J]. Pacific Journal of Optimization, 2006, 2(3): 447-467.

[4] BONNEL H, MORGAN J. Semivectorial bilevel optimization problem: Penalty approach[J]. Journal of Optimization Theory and Applications, 2006, 131(3): 365-382.

[5] ANKHILI Z, MANSOURI A. An exact penalty on bilevel programs with linear vector optimization lower level[J]. European Journal of Operational Research, 2009, 197: 36-41.

[6] DEMPE S, GADHI N, ZEMKOHO A B. New optimality conditions for the semivectorial bilevel optimization problem[J]. Journal of Optimization Theory and Applications, 2013, 157(1): 54-74.

[7] EICHFELDER G. Multiobjective bilevel optimization [J]. Mathematical Programming, 2010, 123: 419-449.

[8] LYU Y, WAN Z. A solution method for the optimistic linear semivectorial bilevel optimization problem[J]. Journal of Inequalities and Applications, 2014, 2014: 164.

[9] MORDUKHOVICH B S, NAM M N, YEN N D. Subgradients of marginal functions in parametric mathematical programming [J]. Mathematical Programming, 2009, 116: 369-396.

[10] ZHENG Y, WAN Z. A solution method for semivectorial bilevel programming problem via penalty

method[J]. Journal of Applied Mathematics and Computing, 2011, 37: 207-219.

[11] BONNEL H, MORGAN J. Optimality conditions for semivectorial bilevel convex optimal control problem [C]// Computational and Analytical Mathematics in Honor of Jonathan Borwein's 60th Birthday. Springer Proceedings in Mathematics & Statistics, Vol 50. New York: Springer, 2013: 45-78.

[12] NIE P. A note on bilevel optimization problems[J]. International Journal of Applied Mathematical Sciences, 2005, 2(1): 31-38.

[13] ABOUSSOROR A, MANSOURI A. Weak linear bilevel programming problems: Existence of solutions via a penalty method[J]. Journal of Mathematical Analysis and Applications, 2005, 304(1): 399-408.

[14] ZHENG Y, FANG D, WAN Z. A solution approach to the weak linear bilevel programming problems[J]. Optimization, 2016, 65(7): 1437-1449.

[15] ZHENG Y, ZHUO X, CHEN J. Maximum entropy approach for solving pessimistic bilevel programming problems[J]. Wuhan University Journal of Natural Sciences, 2017, 22(1): 63-67.

[16] MORDUKHOVICH B S. Variational Analysis and Generalized Differentiation. Ⅰ: Basic Theory. Ⅱ: Applications[M]. Berlin: Springer, 2006.

[17] ROBINSON S M. Some continuity properties of polyhedral multifunctions [C]// Mathematical Programming at Oberwolfach. Mathematical Programming Studies, Vol 14. Berlin/ Heidelberg: Springer, 1981: 206-214.

[18] ROCKAFELLAR R T, WETS R J-B. Variational Analysis[M]. Berlin: Springer, 1998.

[19] EHRGOTT M. Multicriteria Optimization[M]. 2nd ed. Berlin: Springer, 2005.

[20] DEMPE S, MORDUKHOVICH B S, ZEMKOHO A B. Sensitivity analysis for two-level value functions with applications to bilevel programming[J]. SIAM Journal on Optimization, 2012, 22(4): 1309-1343.

[21] MORDUKHOVICH B S, NAM N M. Variational stability and marginal functions via generalized differentiation [J]. Mathematics of Operations Research, 2005, 30(4): 800-816.

[22] MORDUKHOVICH B S, NAM M N, PHAN H M. Variational analysis of marginal functions with applications to bilevel programming[J]. Journal of Optimization Theory and Applications, 2012, 152(3): 557-586.

[23] DEMPE S, ZEMKOHO A B. The generalized Mangasarian-Fromowitz constraint qualification and optimality conditions for bilevel optimization problem [J]. Journal of Optimization Theory and Applications, 2011, 148: 433-441.

[24] DEMPE S, DINH N, DUTTA J. Optimality conditions for a simple convex bilevel porgramming problem[C]// Variational Analysis and Generalized Differentiation in Optimization and Control. Springer Optimization and Its Applications, Vol 47. New York: Springer, 2010: 149-161.

[25] HENRION R, JOURANI A, OUTRATA J V. On the calmness of a class of multifunctions[J]. SIAM Journal on Optimization, 2002, 13(2): 603-618.

[26] ZEMKOHO A B. A simple approach to optimality conditions in minmax programming[J]. Optimization, 2014, 63: 385-401.

[27] ZEMKOHO A B. Bilevel programming: Reformulations, regularity, and stationarity [D]. Freiberg, Germany: Faculty of Mathematics and Computer Science, TU Bergakademie Freiberg, 2012.