Yokohama Publishers
ISSN 1349-8169 ONLINE JOURNAL

Pacific Journal of Optimization
Vol. 20, No. 3, pp. 475–488, 2024

NOTES ON LOWER-LEVEL DUALITY APPROACH FOR BILEVEL PROGRAMS*

Yu-Wei Li, Gui-Hua Lin† and Xide Zhu

Dedicated to Professor Masao Fukushima on the occasion of his 75th birthday

Abstract: This paper focuses on a new approach based on lower-level Wolfe and Mond-Weir duality for bilevel programs, which gives two new single-level reformulations called WDP and MDP, respectively. Different from the popular MPCC (i.e., mathematical program with complementarity constraints) approach, both WDP and MDP may satisfy the Mangasarian-Fromovitz constraint qualification at their feasible points. This paper aims at exploring whether these new reformulations satisfy other constraint qualifications such as Abadie CQ and Guignard CQ. In particular, some sufficient conditions to ensure Abadie CQ and Guignard CQ to hold for WDP and MDP are derived.

Key words: *bilevel program, Wolfe duality, Mond-Weir duality, abadie constraint qualification, guignard constraint qualification*

Mathematics Subject Classification: *90C30, 90C33, 90C46*

1 Introduction

Consider the bilevel program

$$\begin{array}{ll} \min & F(x,y) \\ \text{s.t.} & (x,y) \in \Omega,\ y \in S(x), \end{array} \tag{1.1}$$

where $F : \mathbb{R}^{n+m} \to \mathbb{R}$, $\Omega \subseteq \mathbb{R}^{n+m}$, and $S(x) \subseteq \mathbb{R}^m$ denotes the optimal solution set of the lower-level parameterized optimization problem

$$(\mathrm{P}_x) \qquad \begin{array}{ll} \min\limits_{y} & f(x,y) \\ \text{s.t.} & g(x,y) \le 0,\ h(x,y) = 0, \end{array}$$

where $f : \mathbb{R}^{n+m} \to \mathbb{R}$, g: $\mathbb{R}^{n+m} \to \mathbb{R}^p$, and h: $\mathbb{R}^{n+m} \to \mathbb{R}^q$. We denote the feasible region of P_x by $Y(x) := \{y \in \mathbb{R}^m \mid g(x,y) \le 0,\ h(x,y) = 0\}$ and assume that F is continuously differentiable, $\{f, g, h\}$ are twice continuously differentiable, and $S(x) \neq \emptyset$ for each x feasible to the upper-level program. Bilevel program was first introduced by Stackelberg in the

*This work was supported in part by the National Natural Science Foundation of China (Nos. 12071280, 11901380).

†Corresponding author.

DOI: https://doi.org/10.61208/pjo-2023-032

market economy [26] and many applications can be found in different fields such as supply chain management [2, 10], hyperparameter selection and meta learning in machine learning [11, 12, 18, 19, 33]. More details about bilevel programs and their recent developments can be found in [1, 3, 9, 21, 22, 25, 28] and the references therein.

As is well-known to us, solving bilevel programs numerically is very challenging. The most popular approach is to use the KKT conditions of the lower-level program to transform (1.1) into the following mathematical program with complementarity constraints:

$$\begin{array}{lll} \text{(MPCC)} & \min & F(x,y) \\ & \text{s.t.} & (x,y) \in \Omega,\ h(x,y) = 0, \\ & & \nabla_y f(x,y) + \nabla_y g(x,y)u + \nabla_y h(x,y)v = 0, \\ & & 0 \leq u \perp g(x,y) \leq 0, \end{array}$$

where $a \perp b$ means $a^T b = 0$. However, the MPCC does not satisfy the Mangasarian-Fromovitz constraint qualification (MFCQ) at any feasible point so that the well-developed optimization algorithms in nonlinear programming may be unstable in solving it. So far, many approximation algorithms have been proposed for solving the MPCC; see, e.g., [5–8, 15, 16, 20, 24, 27]. Another approach to solve bilevel programs is based on the lower-level value function. Since the value function does not have analytic expressions in general, it can not be solved directly by the popular optimization algorithms. For more details about this approach, we refer the readers to [17, 30–32] and the references therein.

Recently, two new methods based on lower-level duality were proposed for solving bilevel programs. The first one is based on the lower-level Wolfe duality in [13], where the bilevel program (1.1) is transformed equivalently to the single-level optimization problem

$$\begin{array}{lll} \text{(WDP)} & \min & F(x,y) \\ & \text{s.t.} & (x,y) \in \Omega,\ g(x,y) \leq 0,\ h(x,y) = 0, \\ & & f(x,y) - f(x,z) - u^T g(x,z) - v^T h(x,z) \leq 0, \\ & & \nabla_z f(x,z) + \nabla_z g(x,z)u + \nabla_z h(x,z)v = 0,\ u \geq 0. \end{array}$$

The second one is based on the lower-level Mond-Weir duality in [14], where (1.1) is transformed equivalently to the single-level optimization problem

$$\begin{array}{lll} \text{(MDP)} & \min & F(x,y) \\ & \text{s.t.} & (x,y) \in \Omega,\ g(x,y) \leq 0,\ h(x,y) = 0, \\ & & f(x,y) - f(x,z) \leq 0,\ u^T g(x,z) + v^T h(x,z) \geq 0, \\ & & \nabla_z f(x,z) + \nabla_z g(x,z)u + \nabla_z h(x,z)v = 0,\ u \geq 0. \end{array}$$

It was shown by two examples in [13, 14] that, unlike MPCCs, both WDP and MDP may satisfy the MFCQ at their feasible points. Numerical experiments on 150 linear bilevel programs generated randomly indicate that, although solving the WDP and MDP directly may not perform very well, relaxation methods based on them are more efficient than the MPCC-based methods. See [13, 14] for more details.

However, as shown in [13, 14], the MFCQ is still hard to hold for both WDP and MDP in many cases. Therefore, it is of great theoretical significance to explore whether these new reformulations satisfy other weaker constraint qualifications such as Abadie CQ and Guignard CQ. In this paper, we focus on this issue. Main contributions can be stated as follows:

- An interesting example is given to illustrate that the WDP and MDP reformulations have certain advantages over the MPCC reformulation.
- The WDP and MDP reformulations may satisfy Abadie and Guignard CQs under mild conditions. Furthermore, some sufficient conditions are given to ensure Abadie and Guignard CQs to hold for both WDP and MDP.

Before discussing Abadie and Guignard CQs for the WDP and MDP, we give an example, which comes from [29], to show that optimal solutions of bilevel program may be obtained by solving the KKT systems of the WDP and MDP, but can not be gotten by solving the corrsponding MPCC. This may be regarded as another superiority of the WDP and MDP reformulations compared with the MPCC reformulation.

Example 1.1. Consider the bilevel program

$$\begin{array}{ll} \min & (y+1)^2 \\ \text{s.t.} & -3 \le x \le 2, \quad y \in \arg\min\limits_{y}\{y^3 - 3y \mid y \ge x\}. \end{array} \tag{1.2}$$

This problem has a unique optimal solution $(\bar{x}, \bar{y}) = (-2, -2)$. Next, we discuss whether $(\bar{x}, \bar{y})$ corresponds to a stationary point of the MPCC, WDP, and MDP reformulations.

- The MPCC reformulation of (1.2) is

$$\begin{array}{ll} \min & (y+1)^2 \\ \text{s.t.} & -3 \le x \le 2, \quad 3y^2 - 3 - u = 0, \\ & 0 \le u \perp x - y \le 0. \end{array} \tag{1.3}$$

The corresponding optimal solution to (1.3) is $(\bar{x}, \bar{y}, \bar{u}) = (-2, -2, 9)$. This point being a weakly stationary point of (1.3) means that there exist $\lambda \in \mathbb{R}^2$ and $\mu, \alpha, \beta \in \mathbb{R}$ such that

$$\lambda_1 - \lambda_2 + \beta = 0, \tag{1.4}$$

$$2(\bar{y}+1) + 6\bar{y}\mu - \beta = 0, \tag{1.5}$$

$$-\mu - \alpha = 0, \tag{1.6}$$

$$0 \le \lambda \perp \begin{pmatrix} \bar{x} - 2 \\ -3 - \bar{x} \end{pmatrix} \le 0, \tag{1.7}$$

$$\alpha = 0 \ \text{ if } \bar{u} > 0, \qquad \beta = 0 \ \text{ if } \bar{x} - \bar{y} < 0. \tag{1.8}$$

Obviously, we have $\alpha = 0$ by (1.8) and then $\mu = 0$ by (1.6). Moreover, we have $\lambda = (0, 0)$ by (1.7) and hence $\beta = 0$ by (1.4). Substituting into (1.5) yields a contradiction, which means that $(\bar{x}, \bar{y}, \bar{u})$ is not a weakly stationary point of (1.3).

- The WDP reformulation of (1.2) is

$$\begin{array}{ll} \min & (y+1)^2 \\ \text{s.t.} & -3 \le x \le 2, \qquad 3z^2 - 3 - u = 0, \\ & y^3 - 3y - z^3 + 3z - u(x - z) \le 0, \quad u \ge 0, \quad x - y \le 0. \end{array} \tag{1.9}$$

One corresponding optimal solution to (1.9) is $(\bar{x}, \bar{y}, \bar{z}, \bar{u}) = (-2, -2, 1, 0)$. This point being a stationary point of (1.9) means that there exist $\lambda \in \mathbb{R}^2_+$, $\mu \in \mathbb{R}$, and $\alpha, \beta, \gamma \in \mathbb{R}_+$

such that

$$\lambda_1 - \lambda_2 - \alpha\bar{u} + \gamma = 0, \tag{1.10}$$
$$2(\bar{y}+1) + \alpha(3\bar{y}^2 - 3) - \gamma = 0, \tag{1.11}$$
$$6\bar{z}\mu = 0, \quad -\alpha(\bar{x}-\bar{z}) - \mu - \beta = 0, \tag{1.12}$$
$$0 \le \lambda \perp \begin{pmatrix} \bar{x}-2 \\ -3-\bar{x} \end{pmatrix} \le 0, \tag{1.13}$$
$$0 \le \alpha \perp \bar{y}^3 - 3\bar{y} - \bar{z}^3 + 3\bar{z} - \bar{u}(\bar{x}-\bar{z}) \le 0, \tag{1.14}$$
$$0 \le \beta \perp \bar{u} \ge 0, \quad 0 \le \gamma \perp \bar{x} - \bar{y} \le 0. \tag{1.15}$$

By (1.12) and (1.13), we have $\lambda = (0,0)$ and $\mu = 0$. Then, we have $\gamma = 0$ by (1.10). Substituting into (1.11) and (1.12) yields $\alpha = \frac{2}{9}$ and $\beta = \frac{2}{3}$. This means that $(\bar{x}, \bar{y}, \bar{z}, \bar{u})$ is a KKT point of (1.9).

- The MDP reformulation of (1.2) is

$$\begin{array}{ll} \min & (y+1)^2 \\ \text{s.t.} & -3 \le x \le 2, \quad 3z^2 - 3 - u = 0, \\ & y^3 - 3y - z^3 + 3z \le 0, \quad u(x-z) \ge 0, \\ & u \ge 0, \quad x - y \le 0. \end{array} \tag{1.16}$$

One corresponding optimal solution to (1.16) is $(\bar{x}, \bar{y}, \bar{z}, \bar{u}) = (-2, -2, 1, 0)$. This point being a stationary point of (1.16) means that there exist $\lambda \in \mathbb{R}^2_+$, $\mu \in \mathbb{R}$, and $\alpha, \beta, \gamma, \nu \in \mathbb{R}_+$ such that

$$\lambda_1 - \lambda_2 - \beta\bar{u} + \nu = 0, \quad 2(\bar{y}+1) + \alpha(3\bar{y}^2 - 3) - \nu = 0, \tag{1.17}$$
$$\alpha(-3\bar{z}^2 + 3) + \beta\bar{u} + 6\bar{z}\mu = 0, \quad -\beta(\bar{x}-\bar{z}) - \mu - \gamma = 0, \tag{1.18}$$
$$0 \le \lambda \perp \begin{pmatrix} \bar{x}-2 \\ -3-\bar{x} \end{pmatrix} \le 0, \quad 0 \le \nu \perp \bar{x} - \bar{y} \le 0, \tag{1.19}$$
$$0 \le \alpha \perp \bar{y}^3 - 3\bar{y} - \bar{z}^3 + 3\bar{z} \le 0, \quad 0 \le \beta \perp \bar{u}(\bar{x}-\bar{z}) \ge 0, \quad 0 \le \gamma \perp \bar{u} \ge 0. \tag{1.20}$$

It is easy to verify that (1.17)-(1.20) hold for $\lambda = (0,0)$, $\mu = 0$, $\alpha = \beta = \frac{2}{9}$, $\gamma = \frac{2}{3}$, and $\nu = 0$. Therefore, $(\bar{x}, \bar{y}, \bar{z}, \bar{u})$ is a KKT point of (1.16).

2 Abadie and Guignard Constraint Qualifications for WDP

As mentioned before, although the WDP may satisfy the MFCQ at its feasible point, this CQ is still hard to hold in many cases [13]. In this section, we focus on exploring whether the WDP satisfies other weak constraint qualifications such as Abadie CQ and Guignard CQ. For simplicity, since the upper-level constraints have no effect on subsequent analysis, we take them away from the WDP.

To proceed our discussion, we introduce the following associated MPCC related to the WDP by adding the complementary constraint:

$$\begin{array}{ll} \min & F(x,y) \\ \text{s.t.} & f(x,y) - f(x,z) - u^T g(x,z) - v^T h(x,z) \le 0, \\ & \nabla_z f(x,z) + \nabla_z g(x,z)u + \nabla_z h(x,z)v = 0, \quad h(x,y) = 0, \\ & 0 \le u \perp g(x,y) \le 0. \end{array} \tag{2.1}$$

Denote by $\mathcal{F}_{\text{WDP}}$ and $\mathcal{F}_{\text{MWDP}}$ the feasible regions of the WDP and (2.1), respectively. It is obvious that $\mathcal{F}_{\text{MWDP}} \subseteq \mathcal{F}_{\text{WDP}}$. Let (x, y, z, u, v) be a feasible point of (2.1). We define the following index sets:

$$\begin{aligned} &I_{0+} := \{i|\ g_i(x, y) = 0,\ u_i > 0\}, \quad I_{-0} := \{i|\ g_i(x, y) < 0,\ u_i = 0\}, \\ &I_{00} := \{i|\ g_i(x, y) = 0,\ u_i = 0\}, \quad I_g := \{i|\ g_i(x, y) = 0,\ i = 1, \cdots, p\}, \\ &I_u := \{i|\ u_i = 0,\ i = 1, \cdots, p\}. \end{aligned}$$

Then, we have the following relationship between stationary points of the WDP and (2.1).

Theorem 2.1. *Let $(x, y, z, u, v) \in \mathcal{F}_{\text{MWDP}}$. Then, (x, y, z, u, v) is a KKT point of the WDP if and only if (x, y, z, u, v) is an S-stationary point of* (2.1).

Proof. For simplicity, we remove the constraint $h(x, y) = 0$ away from the problems involved. Suppose that (x, y, z, u) is a KKT point of the WDP. Then, there exists $(\eta^g, \eta^u, \alpha, \beta) \in \mathbb{R}^{p+p+1+m}$ such that

$$\nabla_x F(x, y) + \alpha(\nabla_x f(x, y) - \nabla_x L(x, z, u)) + \nabla^2_{zx} L(x, z, u)\beta + \nabla_x g(x, y)\eta^g = 0, \tag{2.2}$$
$$\nabla_y F(x, y) + \alpha \nabla_y f(x, y) + \nabla_y g(x, y)\eta^g = 0, \tag{2.3}$$
$$\nabla^2_{zz} L(x, z, u)\beta = 0, \quad \nabla_z g(x, z)^T \beta - \alpha g(x, z) - \eta^u = 0, \tag{2.4}$$
$$0 \le \alpha \perp f(x, y) - L(x, z, u) \le 0, \tag{2.5}$$
$$0 \le \eta^g \perp g(x, y) \le 0, \tag{2.6}$$
$$0 \le \eta^u \perp u \ge 0, \tag{2.7}$$

where $L(x, z, u) := f(x, z) + u^T g(x, z)$. Set $\lambda^g := \eta^g$, $\lambda^u := \eta^u$, $\alpha' := \alpha$, and $\beta' := \beta$. When $i \in I_{-0}$ (i.e., $g_i(x, y) < 0$), we have from (2.6) that $\lambda_i^g = \eta_i^g = 0$. When $i \in I_{0+}$ (i.e., $u_i > 0$), we have from (2.7) that $\lambda_i^u = \eta_i^u = 0$. Thus, (x, y, z, u) is an S-stationary point of (2.1).

Conversely, suppose that (x, y, z, u) is an S-stationary point of (2.1), which means that there exists $(\lambda^g, \lambda^u, \alpha', \beta')$ satisfying (2.2)-(2.5) and

$$\lambda_i^g = 0 \quad (i \in I_{-0}), \tag{2.8}$$
$$\lambda_i^u = 0 \quad (i \in I_{0+}), \tag{2.9}$$
$$\lambda_i^g \ge 0,\ \lambda_i^u \ge 0 \quad (i \in I_{00}). \tag{2.10}$$

By setting $\eta^g := \lambda^g$, $\eta^u := \lambda^u$, $\alpha := \alpha'$, and $\beta := \beta'$, we obtain (2.2)-(2.7) immediately. This completes the proof. □

Note that (2.1) is equivalent to

$$\begin{aligned} \min \quad & F(x, y) \\ \text{s.t.} \quad & f(x, y) - f(x, z) - u^T g(x, z) - v^T h(x, z) \le 0, \\ & \nabla_z f(x, z) + \nabla_z g(x, z)u + \nabla_z h(x, z)v = 0, \quad h(x, y) = 0 \\ & u \ge 0, \quad g(x, y) \le 0, \quad u^T g(x, y) \ge 0. \end{aligned} \tag{2.11}$$

Denote by $\mathcal{F}_{\text{NWDP}}$ the feasible region of (2.11). We have $\mathcal{F}_{\text{MWDP}} = \mathcal{F}_{\text{NWDP}} \subseteq \mathcal{F}_{\text{WDP}}$.

Under the condition that the first constraint in (2.1) and (2.11) is active at (x, y, z, u, v), the linearization cones of $\mathcal{F}_{\mathrm{WDP}}$ and $\mathcal{F}_{\mathrm{NWDP}}$ at (x, y, z, u, v) can be written as

$$\mathcal{L}_{\mathcal{F}_{\mathrm{WDP}}}(x,y,z,u,v) = \left\{ \begin{array}{l|l} (dx,dy,dz,du,dv) & \big(\nabla_x f(x,y) - \nabla_x L(x,z,u,v)\big)^T dx + \nabla_y f(x,y)^T dy \\ \in \mathbb{R}^{n+m+m+p+q} & -g(x,z)^T du - h(x,z)^T dv \leq 0 \\ & \big(\nabla^2_{zx} L(x,z,u,v)\big)^T dx + \big(\nabla^2_{zz} L(x,z,u,v)\big)^T dz \\ & +\nabla_z g(x,z) du + \nabla_z h(x,z) dv = 0 \\ & du_i \geq 0, \quad (i \in I_u) \\ & \nabla_x g_i(x,y)^T dx + \nabla_y g_i(x,y)^T dy \leq 0, \quad (i \in I_g) \\ & \nabla_x h_i(x,y)^T dx + \nabla_y h_i(x,y)^T dy = 0, \quad (i = 1, \cdots, q) \end{array} \right\},$$

$$\mathcal{L}_{\mathcal{F}_{\mathrm{NWDP}}}(x,y,z,u,v) = \left\{ \begin{array}{l|l} (dx,dy,dz,du,dv) & \big(\nabla_x f(x,y) - \nabla_x L(x,z,u,v)\big)^T dx + \nabla_y f(x,y)^T dy \\ \in \mathbb{R}^{n+m+m+p+q} & -g(x,z)^T du - h(x,z)^T dv \leq 0 \\ & \big(\nabla^2_{zx} L(x,z,u,v)\big)^T dx + \big(\nabla^2_{zz} L(x,z,u,v)\big)^T dz \\ & +\nabla_z g(x,z) du + \nabla_z h(x,z) dv = 0 \\ & du_i \geq 0, \quad (i \in I_u) \\ & \nabla_x g_i(x,y)^T dx + \nabla_y g_i(x,y)^T dy \leq 0, \quad (i \in I_g) \\ & \nabla_x h_i(x,y)^T dx + \nabla_y h_i(x,y)^T dy = 0, \quad (i = 1, \cdots, q) \\ & -(\nabla_x g(x,y) u)^T dx - (\nabla_y g(x,y) u)^T dy - g(x,y)^T du \leq 0 \end{array} \right\}.$$

Remark 2.2. In the case that the first constraint in (2.1) and (2.11) is inactive at (x, y, z, u, v), the corresponding linearization conditions in the above cones need to be removed, which does not affect subsequent analysis. On the other hand, since $\nabla_x h(x,y)^T dx + \nabla_y h(x,y)^T dy = 0$ and $h(x,y) = 0$, the last item in $\mathcal{L}_{\mathcal{F}_{\mathrm{NWDP}}}(x,y,z,u,v)$ can be equivalently expressed as

$$\begin{aligned} &-(\nabla_x g(x,y)u + \nabla_x h(x,y)v)^T dx - (\nabla_y g(x,y)u + \nabla_y h(x,y)v)^T dy \\ &\qquad\qquad - g(x,y)^T du - h(x,y)^T dv \leq 0. \end{aligned}$$

It will be used in subsequent analysis.

We have the following results for the problems (2.1) and (2.11), which are preparatory to the establishment of Abadie CQ and Guignard CQ for the WDP.

Lemma 2.3. *We have* $\mathcal{L}_{\mathcal{F}_{\mathrm{NWDP}}}(x,y,z,u,v) = \mathcal{L}_{\mathcal{F}_{\mathrm{MWDP}}}(x,y,z,u,v)$ *for any feasible point* (x, y, z, u, v).

Proof. First of all, to simplify the proof, we skip all common constraints and so two cones can be rewritten as follows:

$$\mathcal{L}_{\mathcal{F}_{\mathrm{NWDP}}} = \left\{ (dx,dy,du) \left| \begin{array}{l} du_i \geq 0, \quad (i \in I_u = I_{00} \cup I_{-0}) \\ \nabla_x g_i(x,y)^T dx + \nabla_y g_i(x,y)^T dy \leq 0, \quad (i \in I_g = I_{00} \cup I_{0+}) \\ (\nabla_x g(x,y)u)^T dx + (\nabla_y g(x,y)u)^T dy + g(x,y)^T du \geq 0 \end{array} \right. \right\},$$

$$\mathcal{L}_{\mathcal{F}_{\mathrm{MWDP}}} = \left\{ (dx,dy,du) \left| \begin{array}{l} du_i = 0, \quad (i \in I_{-0}) \\ du_i \geq 0, \quad \nabla_x g_i(x,y)^T dx + \nabla_y g_i(x,y)^T dy \leq 0, \quad (i \in I_{00}) \\ \nabla_x g_i(x,y)^T dx + \nabla_y g_i(x,y)^T dy = 0, \quad (i \in I_{0+}) \end{array} \right. \right\}.$$

It is easy to see that $\mathcal{L}_{\mathcal{F}_{\rm MWDP}} \subseteq \mathcal{L}_{\mathcal{F}_{\rm NWDP}}$.

Let $(dx, dy, du) \in \mathcal{L}_{\mathcal{F}_{\rm NWDP}}$ be arbitrary. Obviously, we have

$$du_i \geq 0, \quad \nabla_x g_i(x,y)^T dx + \nabla_y g_i(x,y)^T dy \leq 0 \quad (i \in I_{00}). \tag{2.12}$$

Since

$$\begin{aligned} A &:= \sum_{i \in I_{00}} ((\nabla_x g_i(x,y)u_i)^T dx + (\nabla_y g_i(x,y)u_i)^T dy + g_i(x,y)du_i) = 0, \\ B &:= \sum_{i \in I_{0+}} ((\nabla_x g_i(x,y)u_i)^T dx + (\nabla_y g_i(x,y)u_i)^T dy + g_i(x,y)du_i) \\ &= \sum_{i \in I_{0+}} ((\nabla_x g_i(x,y)u_i)^T dx + (\nabla_y g_i(x,y)u_i)^T dy), \\ C &:= \sum_{i \in I_{-0}} ((\nabla_x g_i(x,y)u_i)^T dx + (\nabla_y g_i(x,y)u_i)^T dy + g_i(x,y)du_i) = \sum_{i \in I_{-0}} g_i(x,y)du_i, \end{aligned}$$

the inequality $(\nabla_x g(x,y)u)^T dx + (\nabla_y g(x,y)u)^T dy + g(x,y)^T du \geq 0$ is equivalent to

$$B + C \geq 0. \tag{2.13}$$

When $i \in I_{0+}$, we have $u_i > 0$ and $\nabla_x g_i(x,y)^T dx + \nabla_y g_i(x,y)^T dy \leq 0$, which implies $(\nabla_x g_i(x,y)u_i)^T dx + (\nabla_y g_i(x,y)u_i)^T dy \leq 0$ for $i \in I_{0+}$ and hence $B = \sum_{i \in I_{0+}} ((\nabla_x g_i(x,y)u_i)^T dx + (\nabla_y g_i(x,y)u_i)^T dy) \leq 0$. On the other hand, when $i \in I_{-0}$, we have $g_i(x,y) < 0$ and $du_i \geq 0$, which means $g_i(x,y)du_i \leq 0$ for $i \in I_{-0}$ and then $C = \sum_{i \in I_{-0}} g_i(x,y)du_i \leq 0$. As a result, we have $B + C \leq 0$, $B \leq 0$, and $C \leq 0$. By (2.13), we have $B = C = 0$, which implies

$$\nabla_x g_i(x,y)^T dx + \nabla_y g_i(x,y)^T dy = 0 \quad (i \in I_{0+}), \tag{2.14}$$

$$du_i = 0 \quad (i \in I_{-0}). \tag{2.15}$$

Combining (2.12) and (2.14)-(2.15), we have $\mathcal{L}_{\mathcal{F}_{\rm NWDP}} \subseteq \mathcal{L}_{\mathcal{F}_{\rm MWDP}}$, which means $\mathcal{L}_{\mathcal{F}_{\rm NWDP}} = \mathcal{L}_{\mathcal{F}_{\rm MWDP}}$. This completes the proof. □

Since $\mathcal{F}_{\rm MWDP} = \mathcal{F}_{\rm NWDP}$, we have the following corollary from Lemma 2.3 immediately.

Corollary 2.4. *The problem* (2.11) *satisfies Guignard (or Abadie) CQ at a feasible point* (x, y, z, u, v) *if and only if the problem* (2.1) *satisfies Guignard (or Abadie) CQ at* (x, y, z, u, v).

Based on the above analysis, we have the following result.

Theorem 2.5. *If Abadie CQ holds for* (2.1) *at* $(x, y, y, u, v) \in \mathcal{F}_{\rm MWDP}$, *then Abadie CQ also holds for the WDP at* (x, y, y, u, v).

Proof. By $\mathcal{F}_{\rm NWDP} \subseteq \mathcal{F}_{\rm WDP}$, we have

$$\mathcal{T}_{\mathcal{F}_{\rm NWDP}}(x,y,z,u,v) \subseteq \mathcal{T}_{\mathcal{F}_{\rm WDP}}(x,y,z,u,v) \subseteq \mathcal{L}_{\mathcal{F}_{\rm WDP}}(x,y,z,u,v).$$

Since (2.1) satisfies Abadie CQ at (x, y, y, u, v), by Lemma 2.3 and Corollary 2.4, we have

$$\mathcal{T}_{\mathcal{F}_{\rm MWDP}}(x,y,y,u,v) = \mathcal{L}_{\mathcal{F}_{\rm MWDP}}(x,y,y,u,v) = \mathcal{L}_{\mathcal{F}_{\rm NWDP}}(x,y,y,u,v) = \mathcal{T}_{\mathcal{F}_{\rm NWDP}}(x,y,y,u,v).$$

To show that the WDP satisfies Abadie CQ at (x,y,y,u,v), we only need to illustrate

$$\mathcal{L}_{\mathcal{F}_{\mathrm{NWDP}}}(x,y,y,u,v) \supseteq \mathcal{L}_{\mathcal{F}_{\mathrm{WDP}}}(x,y,y,u,v). \tag{2.16}$$

In fact, for any $d=(dx,dy,dz,du,dv) \in \mathcal{L}_{\mathcal{F}_{\mathrm{WDP}}}(x,y,y,u,v)$, it is easy to see that, compared to $\mathcal{L}_{\mathcal{F}_{\mathrm{WDP}}}(x,y,y,u,v)$, $\mathcal{L}_{\mathcal{F}_{\mathrm{NWDP}}}(x,y,y,u,v)$ has only one additional term

$$-(\nabla_x g(x,y)u)^T dx - (\nabla_y g(x,y)u)^T dy - g(x,y)^T du \le 0.$$

Considering the first item of $\mathcal{L}_{\mathcal{F}_{\mathrm{WDP}}}(x,y,y,u,v)$, we have

$$-(\nabla_x g(x,y)u + \nabla_x h(x,y)v)^T dx + \nabla_y f(x,y)^T dy - g(x,y)^T du - h(x,y)^T dv \le 0. \tag{2.17}$$

By $\nabla_y f(x,y)+\nabla_y g(x,y)u+\nabla_y h(x,y)v=0$ and Remark 2.2, we have $d\in\mathcal{L}_{\mathcal{F}_{\mathrm{NWDP}}}(x,y,y,u,v)$. This completes the proof. □

We further have the following result related to Guignard CQ.

Theorem 2.6. *If Guignard CQ holds for* (2.1) *at* $(x,y,y,u,v) \in \mathcal{F}_{\mathrm{MWDP}}$*, then Guignard CQ also holds for the WDP at* (x,y,y,u,v)*.*

Proof. By $\mathcal{F}_{\mathrm{NWDP}} \subseteq \mathcal{F}_{\mathrm{WDP}}$, we have

$$\mathcal{T}_{\mathcal{F}_{\mathrm{NWDP}}}(x,y,z,u,v) \subseteq \mathcal{T}_{\mathcal{F}_{\mathrm{WDP}}}(x,y,z,u,v) \subseteq \mathcal{L}_{\mathcal{F}_{\mathrm{WDP}}}(x,y,z,u,v),$$

which yields the inclusion

$$\mathcal{T}_{\mathcal{F}_{\mathrm{NWDP}}}(x,y,z,u,v)^\circ \supseteq \mathcal{T}_{\mathcal{F}_{\mathrm{WDP}}}(x,y,z,u,v)^\circ \supseteq \mathcal{L}_{\mathcal{F}_{\mathrm{WDP}}}(x,y,z,u,v)^\circ.$$

In order to verify Guignard CQ for the WDP at (x,y,y,u,v), we only need to show $\mathcal{T}_{\mathcal{F}_{\mathrm{NWDP}}}(x,y,y,u,v)^\circ \subseteq \mathcal{L}_{\mathcal{F}_{\mathrm{WDP}}}(x,y,y,u,v)^\circ$.

In fact, since Guignard CQ holds for (2.1) at (x,y,y,u,v), by Lemma 2.4, Guignard CQ also holds for (2.11) at (x,y,y,u,v), that is,

$$\mathcal{T}_{\mathcal{F}_{\mathrm{NWDP}}}(x,y,y,u,v)^\circ = \mathcal{L}_{\mathcal{F}_{\mathrm{NWDP}}}(x,y,y,u,v)^\circ.$$

Let $\xi=(\xi_x,\xi_y,\xi_z,\xi_u,\xi_v) \in \mathcal{T}_{\mathcal{F}_{\mathrm{NWDP}}}(x,y,y,u,v)^\circ = \mathcal{L}_{\mathcal{F}_{\mathrm{NWDP}}}(x,y,y,u,v)^\circ$ be arbitrary. By Remark 2.2 and Farkas' Lemma [23, Lemma 6.45], there exist multipliers $\xi_i^g \ge 0$ $(i \in I_g)$, ξ^h, $\xi_i^u \ge 0$ $(i \in I_u)$, β', $\alpha' \ge 0$, and $\gamma' \ge 0$ such that

$$\begin{aligned}
\xi_x &= -(\alpha'+\gamma')(\nabla_x g(x,y)u + \nabla_x h(x,y)v) + \nabla^2_{yx} L(x,y,u,v)\beta' \\
&\quad + \sum_{i\in I_g} \nabla_x g_i(x,y)\xi_i^g + \nabla_x h(x,y)\xi^h, \\
\xi_y &= (\alpha'+\gamma')\nabla_y f(x,y) + \sum_{i\in I_g} \nabla_y g_i(x,y)\xi_i^g + \nabla_y h(x,y)\xi^h, \\
\xi_z &= \nabla^2_{yy} L(x,y,u,v)\beta', \\
\xi_u &= -(\alpha'+\gamma')g(x,y) + \nabla_y g(x,y)^T\beta' - \sum_{i\in I_u} \xi_i^u I_{p\times i}, \\
\xi_v &= -(\alpha'+\gamma')h(x,y) + \nabla_y h(x,y)^T\beta',
\end{aligned}$$

where $I_{p\times i}$ denotes the *ith* column of the unit matrix $I_{p\times p}$. This implies that $\xi \in \mathcal{L}_{\mathcal{F}_{\mathrm{WDP}}}(x,y,y,u,v)^\circ$. Therefore, Guignard CQ holds for the WDP at (x,y,y,u,v). This completes the proof. □

Remark 2.7. See [4, Proposition 2.17 and Theorems 3.4-3.5] for some sufficient conditions to ensure (2.1) to satisfy Abadie CQ or Guignard CQ. For ease of understanding, we summarize some relations related to Guignard CQ as follows:

$$
\begin{aligned}
& \text{MPEC-LICQ of (2.1)} \\
\xrightarrow{[4]}\;\; & \text{Guignard CQ of (2.1)} \\
\xrightarrow{\text{Corollary 2.4}}\;\; & \text{Guignard CQ of (2.11)} \\
\xrightarrow[y=z]{\text{Theorem 2.6}}\;\; & \text{Guignard CQ of the WDP}
\end{aligned}
$$

Since the sufficient conditions given in [4] for (2.1) to satisfy Abadie CQ are somewhat complicated, we refer the reader to [4, Proposition 2.17] for details.

3 Abadie and Guignard Constraint Qualifications for MDP

In this section, we extend the results in the last section to the case of the MDP. To this end, we introduce the following two associated problems:

$$
\begin{array}{ll}
\min & F(x,y) \\
\text{s.t.} & f(x,y)-f(x,z)\le 0, \quad u^T g(x,z)+v^T h(x,z)\ge 0, \\
& \nabla_z f(x,z)+\nabla_z g(x,z)u+\nabla_z h(x,z)v=0, \quad h(x,y)=0, \\
& 0\le u\perp g(x,y)\le 0
\end{array}
\tag{3.1}
$$

and

$$
\begin{array}{ll}
\min & F(x,y) \\
\text{s.t.} & f(x,y)-f(x,z)\le 0, \quad u^T g(x,z)+v^T h(x,z)\ge 0, \\
& \nabla_z f(x,z)+\nabla_z g(x,z)u+\nabla_z h(x,z)v=0, \quad h(x,y)=0, \\
& u\ge 0, \quad g(x,y)\le 0, \quad u^T g(x,y)\ge 0.
\end{array}
\tag{3.2}
$$

Denote by $\mathcal{F}_{\text{MDP}}$, $\mathcal{F}_{\text{MMDP}}$, and $\mathcal{F}_{\text{NMDP}}$ the feasible sets of the MDP, (3.1), and (3.2) respectively. Obviously, we have $\mathcal{F}_{\text{NMDP}}=\mathcal{F}_{\text{MMDP}}\subseteq\mathcal{F}_{\text{MDP}}$. We also assume that the first two constraints are both active at (x,y,z,u,v) when defining the following linearization cones:

$$
\mathcal{L}_{\mathcal{F}_{\text{MDP}}}(x,y,z,u,v)=
$$

$$
\left\{
\begin{array}{l|l}
 & \big(\nabla_x f(x,y)-\nabla_x f(x,z)\big)^T dx+\nabla_y f(x,y)^T dy-\nabla_z f(x,z)^T dz\le 0 \\
 & \big(-\nabla_x g(x,z)u-\nabla_x h(x,z)v\big)^T dx-g(x,z)^T du-h(x,z)^T dv \\
 & \big(-\nabla_z g(x,z)u-\nabla_z h(x,z)v\big)^T dz\le 0 \\
(dx,dy,dz,du,dv) & \big(\nabla^2_{zx}L(x,z,u,v)\big)^T dx+\big(\nabla^2_{zz}L(x,z,u,v)\big)^T dz \\
\in\mathbb{R}^{n+m+m+p+q} & +\nabla_z g(x,z)du+\nabla_z h(x,z)dv=0 \\
 & du_i\ge 0, \quad (i\in I_u) \\
 & \nabla_x g_i(x,y)^T dx+\nabla_y g_i(x,y)^T dy\le 0, \quad (i\in I_g) \\
 & \nabla_x h_i(x,y)^T dx+\nabla_y h_i(x,y)^T dy=0, \quad (i=1,\cdots,q)
\end{array}
\right\},
$$

$$\mathcal{L}_{\mathcal{F}_{\rm NMDP}}(x,y,z,u,v)=$$
$$\left\{ \begin{array}{c|l} & \big(\nabla_x f(x,y)-\nabla_x f(x,z)\big)^T dx+\nabla_y f(x,y)^T dy-\nabla_z f(x,z)^T dz\le 0 \\ & \big(-\nabla_x g(x,z)u-\nabla_x h(x,z)v\big)^T dx-g(x,z)^T du-h(x,z)^T dv \\ & \big(-\nabla_z g(x,z)u-\nabla_z h(x,z)v\big)^T dz\le 0 \\ (dx,dy,dz,du,dv) & \big(\nabla^2_{zx}L(x,z,u,v)\big)^T dx+\big(\nabla^2_{zz}L(x,z,u,v)\big)^T dz \\ \in \mathbb{R}^{n+m+m+p+q} & +\nabla_z g(x,z)du+\nabla_z h(x,z)dv=0 \\ & du_i\ge 0,\quad (i\in I_u) \\ & \nabla_x g_i(x,y)^T dx+\nabla_y g_i(x,y)^T dy\le 0,\quad (i\in I_g) \\ & \nabla_x h_i(x,y)^T dx+\nabla_y h_i(x,y)^T dy=0,\quad (i=1,\cdots,q) \\ & -(\nabla_x g(x,y)u)^T dx-(\nabla_y g(x,y)u)^T dy-g(x,y)^T du\le 0 \end{array} \right\}.$$

In particular, the last item in $\mathcal{L}_{\mathcal{F}_{\rm NMDP}}(x,y,z,u,v)$ can be equivalently expressed as

$$\begin{aligned} &-(\nabla_x g(x,y)u+\nabla_x h(x,y)v)^T dx-(\nabla_y g(x,y)u \\ &\qquad +\nabla_y h(x,y)v)^T dy-g(x,y)^T du-h(x,y)^T dv\le 0. \end{aligned} \tag{3.3}$$

Similarly as in Section 2, we can show the following results for the MDP and the problems (2.1)-(2.11).

Theorem 3.1. *Let $(x,y,z,u,v)\in\mathcal{F}_{\rm MMDP}$. Then, (x,y,z,u,v) is a KKT point of the MDP if and only if (x,y,z,u,v) is an S-stationary point of* (3.1)*.*

Lemma 3.2. *For any $(x,y,z,u,v)\in\mathcal{F}_{\rm MMDP}$, we have $\mathcal{L}_{\mathcal{F}_{\rm NMDP}}(x,y,z,u,v)=\mathcal{L}_{\mathcal{F}_{\rm MMDP}}(x,y,z,u,v)$ and, in addition,* (3.2) *satisfies Abadie (or Guignard) CQ at (x,y,z,u,v) if and only if* (3.1) *satisfies Abadie (or Guignard) CQ at (x,y,z,u,v).*

Based on the above lemma, we have the following results related to Abadie and Guignard CQs.

Theorem 3.3. *If Abadie CQ holds for* (3.1) *at $(x,y,y,u,v)\in\mathcal{F}_{\rm MMDP}$, then Abadie CQ also holds for the MDP at (x,y,y,u,v).*

Proof. By $\mathcal{F}_{\rm NMDP}\subseteq\mathcal{F}_{\rm MDP}$, we have

$$\mathcal{T}_{\mathcal{F}_{\rm NMDP}}(x,y,z,u,v)\subseteq\mathcal{T}_{\mathcal{F}_{\rm MDP}}(x,y,z,u,v)\subseteq\mathcal{L}_{\mathcal{F}_{\rm MDP}}(x,y,z,u,v).$$

Since (3.1) satisfies Abadie CQ at (x,y,y,u,v), by Lemma 3.2, we have

$$\mathcal{L}_{\mathcal{F}_{\rm NMDP}}(x,y,y,u,v)=\mathcal{T}_{\mathcal{F}_{\rm NMDP}}(x,y,y,u,v).$$

Therefore, to prove that the MDP satisfies Abadie CQ at (x,y,y,u,v), we only need to show

$$\mathcal{L}_{\mathcal{F}_{\rm NMDP}}(x,y,y,u,v)\supseteq\mathcal{L}_{\mathcal{F}_{\rm MDP}}(x,y,y,u,v). \tag{3.4}$$

In fact, for any $d=(dx,dy,dz,du,dv)\in\mathcal{L}_{\mathcal{F}_{\rm MDP}}(x,y,y,u,v)$, it is easy to see that, compared to $\mathcal{L}_{\mathcal{F}_{\rm MDP}}(x,y,y,u,v)$, $\mathcal{L}_{\mathcal{F}_{\rm NMDP}}(x,y,y,u,v)$ has only one additional term $-(\nabla_x g(x,y)u)^T dx-(\nabla_y g(x,y)u)^T dy-g(x,y)^T du\le 0$. By (3.3), we have $d\in\mathcal{L}_{\mathcal{F}_{\rm NMDP}}(x,y,y,u,v)$. This completes the proof. □

Theorem 3.4. *If Guignard CQ holds for* (3.1) *at $(x,y,y,u,v)\in\mathcal{F}_{\rm MWDP}$, then Guignard CQ holds for the MDP at (x,y,y,u,v).*

Proof. By $\mathcal{F}_{\text{NMDP}} \subseteq \mathcal{F}_{\text{MDP}}$, we have

$$\mathcal{T}_{\mathcal{F}_{\text{NMDP}}}(x,y,z,u,v) \subseteq \mathcal{T}_{\mathcal{F}_{\text{MDP}}}(x,y,z,u,v) \subseteq \mathcal{L}_{\mathcal{F}_{\text{MDP}}}(x,y,z,u,v),$$

which yields the inclusion

$$\mathcal{T}_{\mathcal{F}_{\text{NMDP}}}(x,y,z,u,v)^\circ \supseteq \mathcal{T}_{\mathcal{F}_{\text{MDP}}}(x,y,z,u,v)^\circ \supseteq \mathcal{L}_{\mathcal{F}_{\text{MDP}}}(x,y,z,u,v)^\circ.$$

In order to verify that Guignard CQ holds for the MDP at (x,y,y,u,v), we only need to show $\mathcal{T}_{\mathcal{F}_{\text{NMDP}}}(x,y,y,u,v)^\circ \subseteq \mathcal{L}_{\mathcal{F}_{\text{MDP}}}(x,y,y,u,v)^\circ$. In fact, since Guignard CQ holds for (3.1) at (x,y,y,u,v), by Lemma 3.2, the problem (3.2) satisfies Guignard CQ at (x,y,y,u,v), that is,

$$\mathcal{T}_{\mathcal{F}_{\text{NMDP}}}(x,y,y,u,v)^\circ = \mathcal{L}_{\mathcal{F}_{\text{NMDP}}}(x,y,y,u,v)^\circ.$$

Pick $\xi = (\xi_x, \xi_y, \xi_z, \xi_u, \xi_v) \in \mathcal{T}_{\mathcal{F}_{\text{NMDP}}}(x,y,y,u,v)^\circ = \mathcal{L}_{\mathcal{F}_{\text{NMDP}}}(x,y,y,u,v)^\circ$ arbitrarily. By (3.3), there exist multipliers $\xi_i^g \geq 0$ $(i \in I_g)$, ξ^h, $\xi_i^u \geq 0$ $(i \in I_u)$, β', $\alpha_1' \geq 0$, $\alpha_2' \geq 0$, and $\gamma' \geq 0$ such that

$$\begin{aligned}
\xi_x &= -(\alpha_2' + \gamma')(\nabla_x g(x,y)u + \nabla_x h(x,y)v) + \nabla_{yx}^2 L(x,y,u,v)\beta' \\
&\quad + \sum_{i \in I_g} \nabla_x g_i(x,y)\xi_i^g + \nabla_x h(x,y)\xi^h, \\
\xi_y &= (\alpha_1' + \gamma')\nabla_y f(x,y) + \sum_{i \in I_g} \nabla_y g_i(x,y)\xi_i^g + \nabla_y h(x,y)\xi^h, \\
\xi_z &= -\alpha_1' \nabla_y f(x,y) - \alpha_2'(\nabla_y g(x,y)u + \nabla_y h(x,y)v) + \nabla_{yy}^2 L(x,y,u,v)\beta' \\
&= (\alpha_2' - \alpha_1')\nabla_y f(x,y) + \nabla_{yy}^2 L(x,y,u,v)\beta', \\
\xi_u &= -(\alpha_2' + \gamma')g(x,y) + \nabla_y g(x,y)^T \beta' - \sum_{i \in I_u} \xi_i^u I_{p \times i}, \\
\xi_v &= -(\alpha_2' + \gamma')h(x,y) + \nabla_y h(x,y)^T \beta'.
\end{aligned}$$

Note that $\alpha_2' + \gamma' \leftrightarrow \alpha_2$ and $\alpha_1' + \gamma' \leftrightarrow \alpha_1$ imply $\alpha_2' - \alpha_1' \leftrightarrow \alpha_2 - \alpha_1$, where the symbol $a \leftrightarrow b$ indicates that a and b correspond one-to-one. Specifically, set $\alpha_2' := \alpha_2 - \gamma'$ and $\alpha_1' := \alpha_1 - \gamma'$. Then, we have $\alpha_2' - \alpha_1' = \alpha_2 - \gamma' - (\alpha_1 - \gamma') = \alpha_2 - \alpha_1$. Therefore, we have $\xi \in \mathcal{L}_{\mathcal{F}_{\text{MDP}}}(x,y,y,u,v)^\circ$ and hence the MDP satisfies Guignard CQ at (x,y,y,u,v). This completes the proof. □

Remark 2.7 related to the WDP can also be applicable to the MDP by replacing (2.1) to (3.1). To end this section, we discuss the relationship between Abadie and Guignard CQs for the WDP and MDP. Note that, by the definitions of the WDP and MDP reformulations, we have $\mathcal{T}_{\mathcal{F}_{\text{MDP}}}(x,y,y,u,v) = \mathcal{T}_{\mathcal{F}_{\text{WDP}}}(x,y,y,u,v)$ and $\mathcal{L}_{\mathcal{F}_{\text{MDP}}}(x,y,y,u,v) = \mathcal{L}_{\mathcal{F}_{\text{WDP}}}(x,y,y,u,v)$. Thus, we can directly obtain the following result.

Corollary 3.5. *The WDP satisfies Guignard (or Abadie) CQ at a feasible point (x,y,y,u,v) if and only if the MDP satisfies Guignard (or Abadie) CQ at (x,y,y,u,v).*

4 Conclusions

This paper aims to explore whether the WDP and MDP reformulations for bilevel programs satisfy Abadie CQ or Guignard CQ. Some sufficient conditions to ensure Abadie CQ and Guignard CQ to hold for the WDP and MDP are given. These results can be regarded as a supplement to the previous works [13, 14]. In the future, we will continue to study this novel approach based on lower-level duality from theoretical and algorithmic perspectives.

Acknowledgments

The authors would like to thank two anonymous referees for their helpful and constructive comments that have helped us to improve the presentation of this paper.

References

[1] J.F. Bard, *Practical Bilevel Optimization: Algorithms and Applications*, Kluwer, Dordrecht, 1998.

[2] J.F. Camacho-Vallejo, L. López-Vera, A.E. Smith and J.L. González-Velarde, A tabu search algorithm to solve a green logistics bi-objective bi-level problem, *Ann. Oper. Res.* 316 (2022) 927–953.

[3] B. Colson, P. Marcotte and G. Savard, An overview of bilevel optimization, *Ann. Oper. Res.* 153 (2007) 235–256.

[4] M.L. Flegel, *Constraint Qualifications and Stationarity Concepts for Mathematical Programs with Equilibrium Constraints*, Ph.D. thesis, Institute of Applied Mathematics and Statistics, University of Würzburg, 2005.

[5] M. Fukushima, Z.Q. Luo and J.S. Pang, A globally convergent sequential quadratic programming algorithm for mathematical programs with linear complementarity constraints, *Comput. Optim. Appl.* 10 (1998) 5–34.

[6] L. Guo, G.H. Lin and J.J. Ye, Solving mathematical programs with equilibrium constraints, *J. Optim. Theory Appl.* 166 (2015) 234–256.

[7] T. Hoheisel, C. Kanzow and A. Schwartz, Theoretical and numerical comparison of relaxation methods for mathematical programs with complementarity constraints, *Math. Program.* 137 (2013) 257–288.

[8] A.F. Izmailov, A.L. Pogosyan and M.V. Solodov, Semismooth Newton method for the lifted reformulation of mathematical programs with complementarity constraints, *Comput. Optim. Appl.* 51 (2012) 199–221.

[9] V. Jeyakumar, J.B. Lasserre, G. Li and T.S. Pham, Convergent semidefinite programming relaxations for global bilevel polynomial optimization problems, *SIAM J. Optim.* 26 (2016) 753–780.

[10] R. Khorramfar, O.Y. Özaltın, K.G Kempf and R. Uzsoy, Managing product transitions: A bilevel programming approach, *INFORMS J. Comput.* 34 (2022) 2828–2844.

[11] G. Kunapuli, K.P. Bennett, J. Hu and J.S. Pang, Classification model selection via bilevel programming, *Optim. Methods Softw.* 23 (2008) 475–489.

[12] Q. Li, Z. Li and A. Zemkoho, Bilevel hyperparameter optimization for support vector classification: Theoretical analysis and a solution method, *Math. Methods Oper. Res.* 96 (2022) 315–350.

[13] Y.W. Li, G.H. Lin, J. Zhang and X. Zhu, A novel approach for bilevel programs based on Wolfe duality, *arXiv preprint arXiv:2302.06838* (2022).

[14] Y.W. Li, G.H. Lin and X. Zhu, Solving bilevel programs based on lower-level Mond-Weir duality, *INFORMS J. Comput.* (2024).

[15] G.H. Lin, X. Chen and M. Fukushima, Solving stochastic mathematical programs with equilibrium constraints via approximation and smoothing implicit programming with penalization, *Math. Program.* 116 (2009) 343–368.

[16] G.H. Lin and M. Fukushima, Hybrid approach with active set identification for mathematical programs with complementarity constraints, *J. Optim. Theory Appl.* 128 (2006) 1–28.

[17] G.H. Lin, M. Xu and J.J. Ye, On solving simple bilevel programs with a nonconvex lower level program, *Math. Program.* 144 (2014) 277–305.

[18] R. Liu, X. Liu, S. Zeng, J. Zhang and Y. Zhang, Hierarchical optimization-derived learning, *IEEE Trans. Pattern Anal. Mach. Intell.* 45 (2023) 14693–14708.

[19] R. Liu, X. Liu, S. Zeng, J. Zhang and Y. Zhang, Value-function-based sequential minimization for bi-level optimization, *IEEE Trans. Pattern Anal. Mach. Intell.* 45 (2023) 15930–15948.

[20] Z.Q. Luo, J.S. Pang and D. Ralph, *Mathematical Programs with Equilibrium Constraints*, Cambridge University Press, Cambridge, 1996.

[21] J. Nie, L. Wang and J.J. Ye, Bilevel polynomial programs and semidefinite relaxation methods, *SIAM J. Optim.* 27 (2017) 1728–1757.

[22] J. Nie, L. Wang, J.J. Ye and S. Zhong, A Lagrange multiplier expression method for bilevel polynomial optimization, *SIAM J. Optim.* 31 (2021) 2368–2395.

[23] R.T. Rockafellar and R.J.B. Wets, *Variational Analysis*, Springer, Berlin, 2004.

[24] S. Scholtes, Convergence properties of a regularization scheme for mathematical programs with complementarity constraints, *SIAM J. Optim.* 11 (2001) 918–936.

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

[26] H.V. Stackelberg, *The Theory of the Market Economy*, Oxford University Press, Oxford, 1952.

[27] S. Steffensen and M. Ulbrich, A new relaxation scheme for mathematical programs with equilibrium constraints, *SIAM J. Optim.* 20 (2010) 2504–2539.

[28] L.N. Vicente and P.H. Calamai, Bilevel and multilevel programming: A bibliography review, *J. Global Optim.* 5 (1994) 291–306.

[29] S. Vogel, *Zwei-Ebenen-Optimierungsaufgaben mit nichtkonvexer Zielfunktion in der unteren Ebene: Pfadverfolgung und Sprünge*, Ph.D. thesis, Department of Mathematics and Computer Science, Technische Universität Bergakademie Freiberg, 2002.

[30] M. Xu and J.J. Ye, A smoothing augmented Lagrangian method for solving simple bilevel programs, *Comput. Optim. Appl.* 59 (2014) 353–377.

[31] J.J. Ye and D.L. Zhu, Optimality conditions for bilevel programming problems, *Optimization* 33 (1995) 9–27.

[32] J.J. Ye and D.L. Zhu, New necessary optimality conditions for bilevel programs by combining the MPEC and value function approaches, *SIAM J. Optim.* 20 (2010) 1885–1905.

[33] J.J. Ye, X. Yuan, S. Zeng and J. Zhang, Difference of convex algorithms for bilevel programs with applications in hyperparameter selection, *Math. Program.* 198 (2023) 1583–1616.

Manuscript received 28 June 2023
revised 10 October 2023
accepted for publication 16 October 2023

Yu-wei Li
School of Management, Shanghai University
Shanghai 200444, China
E-mail address: yuwei_li@shu.edu.cn

Gui-Hua Lin
School of Management, Shanghai University
Shanghai 200444, China
E-mail address: guihualin@shu.edu.cn

Xide Zhu
School of Management, Shanghai University
Shanghai 200444, China
E-mail address: xidezhu@shu.edu.cn