




Extension of the value function reformulation to multiobjective bilevel optimization

Lahoussine Lafhim¹ · Alain Zemkoho² 

Received: 9 December 2021 / Accepted: 19 October 2022
© The Author(s) 2022

Abstract

We consider a multiobjective bilevel optimization problem with vector-valued upper- and lower-level objective functions. Such problems have attracted a lot of interest in recent years. However, so far, scalarization has appeared to be the main approach used to deal with the lower-level problem. Here, we utilize the concept of frontier map that extends the notion of optimal value function to our parametric multiobjective lower-level problem. Based on this, we build a tractable constraint qualification that we use to derive necessary optimality conditions for the problem. Subsequently, we show that our resulting necessary optimality conditions represent a natural extension from standard optimistic bilevel programs with scalar objective functions.

Keywords Multiobjective bilevel optimization · Frontier map · Strong domination property · Coderivative · Optimality conditions

Mathematics Subject Classification 90C26 · 90C31 · 90C46 · 49K99

1 Introduction

A standard bilevel optimization problem involves the minimization of a real-valued function under a constraint set partly constrained by the optimal solution set of a parametric optimization problem with a scalar objective function; see, e.g., [10] for the most recent surveys on the topic. However, in the last two to three decades, significant attention has been paid to the generalization of this model to the case where

✉ Alain Zemkoho
a.b.zemkoho@soton.ac.uk

Lahoussine Lafhim
lahoussine.lafhim@usmba.ac.ma

¹ Laboratoire LASMA, Department of Mathematics, Sidi Mohammed Ben Abdellah University, Fes, Morocco

² School of Mathematical Sciences, University of Southampton, Southampton, UK

the upper- and/or lower-level objective functions are vector-valued. This is precisely the main focus of the work in this paper, as we consider the multiobjective optimization problem

$$\min_{x,y} F(x,y) := (F_1(x,y), \dots, F_p(x,y))^T \text{ s.t. } x \in X, \ y \in S(x), \quad (\text{MUL})$$

where the vector valued function $F : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^p$ (with $p, m, n \in \mathbb{N}$, $p \geq 2$) represents the upper-level objective function, while $X \subseteq \mathbb{R}^n$ corresponds to the upper-level feasible set. As for the set-valued mapping $S : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, it collects the optimal solutions of the vector lower-level problem

$$\min_y f(x,y) := (f_1(x,y), \dots, f_q(x,y))^T \text{ s.t. } y \in Y(x) \quad (\text{L}[x])$$

for a given $x \in \mathbb{R}^n$. Here, $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^q$ (with $q, m, n \in \mathbb{N}$, $q \geq 2$) represents the vector lower-level objective function, while the set-valued map $Y : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ describes the lower-level feasible set.

We associate with the (multiobjective) lower-level problem (L[x]) the corresponding frontier map $\Phi : \mathbb{R}^n \rightrightarrows \mathbb{R}^q$ defined by

$$\Phi(x) := \text{Eff/WEff} (f(x, Y(x)); \mathbb{R}_+^q), \quad (1.1)$$

where the notation Eff/WEff is used to reflect the fact that optimality in (1.1) is in the sense of efficient Pareto (Eff) or weakly efficient Pareto (WEff) optimality. In the sequel, we will simply write $\Phi^E(x) = \text{Eff}(f(x, Y(x)); \mathbb{R}_+^q)$ (resp. $\Phi^W(x) = \text{WEff}(f(x, Y(x)); \mathbb{R}_+^q)$) when referring to Pareto (resp. weakly Pareto) efficiency in situations where it is necessary to distinguish between these two concepts, which are defined in the next section. Obviously, based on (1.1), the set-valued mapping $S : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ can be rewritten as

$$S(x) := \{y \in Y(x) : f(x,y) \in \Phi(x)\} \text{ for } x \in \mathbb{R}^n. \quad (1.2)$$

Hence, our problem (MUL)–(L[x]) can be equivalently written as

$$\min_{x,y} F(x,y) := (F_1(x,y), \dots, F_p(x,y))^T \text{ s.t. } x \in X, \ y \in Y(x), \ f(x,y) \in \Phi(x). \quad (1.3)$$

Note that in the case where $q = 1$, meaning that our lower-level problem (L[x]) is simply a standard scalar objective parametric optimization problem, then the frontier map Φ reduces to the corresponding optimal value function. And therefore, problem (1.3) will become the standard lower-level value function (LLVF) reformulation well-known in bilevel optimization with scalar objective functions; see, e.g., [5, 8, 9, 29, 30], for more details on this class of the problem. Hence, clearly, (1.3) is a natural extension of the LLVF reformulation to the multiobjective bilevel optimization problem (MUL)–(L[x]); thus we labelled it as such throughout this paper.

The number of publications around problem (MUL)–(L[x]) or the semivectorial version of the problem, where only the lower-level is multiobjective has been growing significantly over the last decade. Recent surveys on the subject include [11,

25], where overviews of different types of solution algorithms are given. However, our main interest here is on constructing necessary optimality conditions for problem (MUL)–(L[x]); a common point of most works on optimality conditions of this problem is that they rely on some form of scalarization to deal with the multiobjective nature of the lower-level problem (L[x]); for recent surveys on the subject, see, e.g., [6, 7] and references therein.

Additionally, in the latter references, the LLVF reformulation is common after the scalarization step, although [32] provides a different perspective. Subsequently, as in the case where $p = 1$ and $q = 1$, the standard approach to develop necessary optimality conditions for the corresponding model, after scalarization, has been the concept of partial calmness introduced in [29]. However, given that partial calmness is, in some sense, equivalent to partial exact penalization of the corresponding value function constraint, it is unclear how such an approach can be directly applied to (1.3) when $p > 1$ or $q > 1$. Hence, our first main focus in this paper (see Sect. 3) is to study the possibility to apply the concept of calmness of set-valued mapping, which is closely related to partial calmness [9, 15]. In Sect. 3, we construct a tractable framework for a set-valued mapping tailored to (MUL)–(L[x]) to be used as constraint qualification (CQ) for the problem. In Sect. 4, we show how this CQ can be used to develop necessary optimality conditions for problem (MUL)–(L[x]). As a by-product of the regularity condition studied in Sect. 4, we provide a new sufficient condition to derive stability for the optimal solution set-valued mapping S (1.2); i.e., for the estimation of its coderivative and Lipschitz-likeness.

Before we move to the development of the main results in Sects. 3 and 4. Finally, in Sect. 5, we apply our results to problems with smooth constraint functionals.

2 Preliminaries

2.1 Tools from variational analysis

In this subsection, we present basic tools from variational analysis that will be used throughout the paper; more on the material covered here can be found in [22, 23], for example. For some point $x \in \mathbb{R}^n$ and a scalar $\epsilon > 0$,

$$\mathbb{U}_\epsilon(x) := \{y \in \mathbb{R}^n \mid \|y - x\| < \epsilon\} \quad \text{and} \quad \mathbb{B}_\epsilon(x) := \{y \in \mathbb{R}^n \mid \|y - x\| \leq \epsilon\}$$

denote the open and closed ϵ -ball around x , respectively. For brevity, we make use of $\mathbb{U}_n = \mathbb{U}_1(0)$ and $\mathbb{B}_n = \mathbb{B}_1(0)$. For a set-valued mapping $Y : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, its Painlevé-Kuratowski outer/upper limit at a point \bar{x} is defined by

$$\limsup_{x \rightarrow \bar{x}} Y(x) := \{x^* \in \mathbb{R}^m : \exists x_k \rightarrow \bar{x}, x_k^* \rightarrow x^* \text{ with } x_k^* \in Y(x_k) \text{ for all } k \in \mathbb{N}\}.$$

Next, consider a set $\Omega \subset \mathbb{R}^n$, which is assumed to be closed around a point $\bar{x} \in \Omega$. The Fréchet normal cone to Ω at $\bar{x} \in \Omega$ is defined by

$$\hat{N}(\bar{x}; \Omega) := \left\{ x^* \in \mathbb{R}^n : \limsup_{\substack{\Omega \\ x \rightarrow \bar{x}}} \frac{\langle x^*, x - \bar{x} \rangle}{\|x - \bar{x}\|} \leq 0 \right\}, \quad (2.1)$$

where $x \xrightarrow{\Omega} \bar{x}$ means that $x \rightarrow \bar{x}$ and $x \in \Omega$. Based on this concept, we can introduce the limiting/Mordukhovich normal cone $N(\bar{x}; \Omega)$ to Ω at \bar{x} , which can be obtained by taking the sequential Painlevé-Kuratowski upper limits of the Fréchet normal cone in (2.1):

$$N(\bar{x}; \Omega) := \limsup_{\substack{\Omega \\ x \rightarrow \bar{x}}} \hat{N}(x; \Omega).$$

If $\bar{x} \notin \Omega$, it is standard to set $N(\bar{x}; \Omega) := \emptyset$. We obviously have $\hat{N}(\bar{x}; \Omega) \subset N(\bar{x}; \Omega)$ and if the inclusion holds as equality, then we say that Ω is normally regular at \bar{x} . The class of normally regular sets includes convex ones and many other important sets in the field of variational analysis and optimization; see, e.g., [22], for more details.

Let $Y : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ be a set-valued mapping with its graph

$$\text{gph } Y := \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m : y \in Y(x)\},$$

the normal coderivative of Y at $(\bar{x}, \bar{y}) \in \text{gph } Y$ is defined by

$$D^*Y(\bar{x}, \bar{y})(y^*) := \{x^* \in \mathbb{R}^n : (x^*, -y^*) \in N((\bar{x}, \bar{y}); \text{gph } Y)\} \text{ for all } y^* \in \mathbb{R}^m. \quad (2.2)$$

When Y is a single-valued mapping, to simplify the notation, one writes $D^*Y(\bar{x})(y^*)$ instead of $D^*Y(\bar{x}, Y(\bar{x}))(y^*)$. Furthermore, for a function $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ that is strictly differentiable at (\bar{x}, \bar{y}) , we have the representation

$$D^*f(\bar{x})(y^*) = \{\nabla f(\bar{x})^\top y^*\} \text{ for all } y^* \in \mathbb{R}^m.$$

We conclude this subsection with some further properties of set-valued mappings. Consider a set-valued mapping $Y : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$. It will be said to be Lipschitz-like around (\bar{x}, \bar{y}) if there exist neighbourhoods U of \bar{x} , V of \bar{y} , and a constant $l > 0$ such that

$$Y(x) \cap V \subseteq Y(u) + l \|u - x\| \mathbb{B}_m \text{ for all } x, u \in U.$$

The weaker concept of calmness is said to hold for a set-valued map Y at some point (\bar{x}, \bar{y}) if there exist neighbourhoods U of \bar{x} , V of \bar{y} , and a constant $l > 0$ such that

$$d(y, Y(\bar{x})) \leq l \|x - \bar{x}\| \text{ for all } x \in U \text{ and } y \in V \cap Y(x).$$

Considering the continuous functions $h_i : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ for $i = 1, \dots, q$, we associate the set-valued mapping Y defined by

$$Y(x) := \{y \in \mathbb{R}^m : h_i(x, y) \leq 0, i = 1, \dots, q\} \text{ for } x \in \mathbb{R}^n. \quad (2.3)$$

Υ (2.3) will be said to be R -regular at (\bar{x}, \bar{y}) w.r.t. $\Omega \subseteq \mathbb{R}^n$ if there are some positive numbers σ , and δ such that for all $(x, y) \in \cup_{\delta}(\bar{x}, \bar{y}) \cap \Omega \times \mathbb{R}^m$,

$$d(y, \Upsilon(x)) \leq \sigma \max \{0, \max \{h_i(x, y) \mid i = 1, \dots, q\}\}. \quad (2.4)$$

For more details on R -regularity, see [21] and references therein.

A set-valued mapping $\Upsilon : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ will be said to be order semicontinuous at a point $(\bar{x}, \bar{y}) \in \text{gph } \Upsilon$, if for any sequence $(x_k, y_k) \in \text{epi } \Upsilon$ converging to (\bar{x}, \bar{y}) , there is a sequence $(x_k, z_k) \in \text{gph } \Upsilon$ with $y_k - z_k \in \mathbb{R}_+^m$ such that (z_k) contains a subsequence converging to \bar{y} . Here, $\text{epi } \Upsilon$ corresponds to the epigraph of Υ with respect to the ordering cone \mathbb{R}_+^m :

$$\text{epi } \Upsilon := \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m : y \in \Upsilon(x) + \mathbb{R}_+^m\},$$

Obviously, Υ will be order semicontinuous around $(\bar{x}, \bar{y}) \in \text{gph } \Upsilon$ if there exists a neighbourhood U of (\bar{x}, \bar{y}) such that Υ is order semicontinuous at any $(x, y) \in U \cap \text{gph } \Upsilon$.

2.2 Multiobjective optimization concepts

Let $C \subset \mathbb{R}^n$ be a pointed closed convex cone, with nonempty interior, introducing a partial order denoted by \leq_C in \mathbb{R}^n .

Definition 2.1 Let Ω be a nonempty set of \mathbb{R}^n . $x \in \Omega$ is said to be a Pareto (resp. weak Pareto) efficient/minimal vector of Ω w.r.t. C if

$$\Omega \subset x + [(\mathbb{R}^n \setminus (-C)) \cup \{0\}] \quad (\text{resp. } \Omega \subset x + (\mathbb{R}^n \setminus -\text{int } C)),$$

where “int” denotes the topological interior of the set in question.

In the sequel, the set of all the Pareto (resp. weak Pareto) efficient/minimal vectors of a set Ω w.r.t. C is denoted by $\text{Eff}(\Omega; C)$ (resp. $\text{WEff}(\Omega; C)$). Let us now consider the following multiobjective optimization problem with respect to the partial order introduced by the pointed, closed, and convex cone C :

$$\min f(x) \text{ s.t. } x \in \Omega, \quad (2.5)$$

where f represents a vector-valued function defined on \mathbb{R}^n and Ω the nonempty feasible set. For a nonempty set $N \subset \Omega$, the image of N by f is defined by

$$f(N) := \{f(x) : x \in N\}.$$

Definition 2.2 A point $\bar{x} \in \Omega$ is said to be a Pareto (resp. weakly Pareto) optimal solution of problem (2.5) if $f(\bar{x})$ is a Pareto (resp. weak Pareto) minimal vector of $f(\Omega)$, i.e., $f(\bar{x}) \in \text{Eff}(f(\Omega); C)$ (resp. $f(\bar{x}) \in \text{WEff}(f(\Omega); C)$).

Similarly, a point $\bar{x} \in \Omega$ is said to be a local Pareto (resp. weakly local Pareto) optimal solution of problem (2.5) if there exists a neighborhood U of \bar{x} such that $f(\bar{x})$ is a Pareto (resp. weak Pareto) minimal vector of $f(U \cap \Omega)$. For our analysis of the multiobjective bilevel program (MUL), we will use either the concept of weakly efficient solution for the upper-level problem or the concept of efficient solution for the upper-level problem, and similarly, for the lower-level problem, both notions of efficient optimal solution and weakly efficient solution will be applied.

3 Generalized value function constraint qualification

We start this section by introducing the main constraint qualification that will be used to derive necessary optimality conditions for problem (1.3).

Definition 3.1 The generalized value function constraint qualification (GVFCQ) holds at (\bar{x}, \bar{y}) if the set-valued mapping $\Psi : \mathbb{R}^n \times \mathbb{R}^q \rightrightarrows \mathbb{R}^n \times \mathbb{R}^m$ defined by

$$\Psi(u, v) := \left\{ (x, y) \in \text{gph } Y : \begin{pmatrix} x \\ f(x, y) \end{pmatrix} + \begin{pmatrix} u \\ v \end{pmatrix} \in \text{gph } \Phi \right\}, \quad (3.1)$$

is calm at the point $(0, 0, \bar{x}, \bar{y})$.

Note that if the lower-level problem $(L[x])$ has a scalar objective function, then replacing the frontier map Φ by the corresponding optimal value function φ , then the value function constraint qualification (VFCQ) in this case is obtained by replacing (3.1) with

$$\Psi_{\varphi}(v) := \{(x, y) \in \text{gph } Y : f(x, y) - \varphi(x) \leq v\}. \quad (3.2)$$

Clearly, we have $\Psi_{\varphi}(v) := \Psi(0, -v)$ if $\text{gph } \Phi$ is replaced in (3.1) by the hypograph of φ .

It is well-known that in bilevel programs with scalar objective functions, the VFCQ implies that the partial calmness condition holds in the case where the lower-level feasible set is unperturbed [15]. Moreover, to the best of our knowledge, the VFCQ is the weakest CQ that ensures that partial calmness holds. Hence, since partial cannot be defined for (1.3), due to the multiobjective nature of the objective functions in (MUL)– $(L[x])$, it makes sense to consider the GVFCQ as the natural candidate for tractable CQ for the problem under consideration. For the remainder of this section, we focus our attention on constructing sufficient conditions ensuring that GVFCQ can hold.

We start with an extension of the uniform weak sharp minimum condition, which enables an extension of a relationship already well-known to be valid in standard bilevel optimization problems with scalar objective functions [15, 28–30].

Definition 3.2 The local uniform weak sharp minimum (LUWSM) condition holds at (\bar{x}, \bar{y}) , for the family problems $(L[x])_{x \in X}$, if there exist $\epsilon > 0$ and $\lambda > 0$ such that

$$\forall (x, y) \in V_\epsilon(\bar{x}, \bar{y}) : y \in Y(x) \implies \lambda d(y, S(x)) \leq d(f(x, y); \Phi(x)).$$

If $V_\epsilon(\bar{x}, \bar{y})$ is replaced by the whole space $\mathbb{R}^n \times \mathbb{R}^m$ in this definition, we simply say that the uniformly weak sharp minimum (UWSM) condition holds at (\bar{x}, \bar{y}) .

Theorem 3.3 *Let $(\bar{x}, \bar{y}) \in \text{gph}S$ and f be locally Lipschitzian around (\bar{x}, \bar{y}) with constant L and assume that Φ is Lipschitz-like around (\bar{x}, \bar{z}) , with $\bar{z} = f(\bar{x}, \bar{y})$. If the LUWSM condition holds at (\bar{x}, \bar{y}) , then the GVFCQ is satisfied at (\bar{x}, \bar{y}) .*

Proof Based on the assumptions, there exist $l > 0$ and $\delta > 0$ such that

$$\Phi(x_0) \cap (\bar{z} + \delta \mathbb{B}) \subset \Phi(x_1) + l \|x_0 - x_1\| \mathbb{B} \text{ for all } x_0, x_1 \in \bar{x} + \delta \mathbb{B}_q. \quad (3.3)$$

Let $0 < \epsilon < \frac{\delta}{2}$ and $\lambda > 0$ be the constants from Definition 3.2 and $u \in \epsilon \mathbb{B}_n$, $v \in \epsilon \mathbb{B}_q$, and $(x, y) \in (\bar{x}, \bar{y}) + \epsilon \mathbb{B}_{n \times m}$ such that $(x, y) \in \Psi(u, v)$. Since $\Psi(0, 0) = \text{gph}S$, then

$$d((x, y), \Psi(0, 0)) = d((x, y), \text{gph}S) \leq d(y, S(x)). \quad (3.4)$$

By the local uniform weak sharp minimum condition, we have

$$d(y, S(x)) \leq \lambda^{-1} d(f(x, y); \Phi(x)). \quad (3.5)$$

Since, f is locally Lipschitzian around (\bar{x}, \bar{y}) with constant L and radius α , then setting $\beta = \min \left\{ \alpha, \frac{\delta}{4L} \right\}$ leads to

$$\begin{aligned} \|v + f(x, y) - f(\bar{x}, \bar{y})\| &\leq \|v\| + L(\|x - \bar{x}\| + \|y - \bar{y}\|), \\ &\leq \epsilon + L\left(\frac{\delta}{4L} + \frac{\delta}{4L}\right), \\ &\leq \frac{\delta}{2} + \frac{\delta}{2}, \\ &= \delta \end{aligned}$$

for all $(x, y) \in (\bar{x}, \bar{y}) + \beta \mathbb{B}_{n \times m}$. Thus, $v + f(x, y) \in \bar{z} + \delta \mathbb{B}_q$. Taking $x_0 = x + u$ and $x_1 = x$ while considering that $x + u \in \bar{x} + \delta \mathbb{B}_n$, it follows from (3.3) that there exist $z \in \Phi(x)$ such that

$$\|v + f(x, y) - z\| \leq l \|u\|.$$

Consequently,

$$d(f(x, y); \Phi(x)) \leq \|f(x, y) - z\| \leq l \|u\| + \|v\|. \quad (3.6)$$

Setting $\tau = \max(l, 1)$ and combining (3.4), (3.5), and (3.6), it follows that

$$d((x, y), \Psi(0, 0)) \leq \lambda^{-1} \tau \|(u, v)\|$$

for all $(u, v) \in \epsilon \mathbb{B}_{n \times q}$ and $(x, y) \in \Psi(u, v) \cap ((\bar{x}, \bar{y}) + \epsilon \mathbb{B}_{n \times m})$. Hence, the result follows. \square

To provide a concrete case where the LUWSM condition holds, we consider the parametric linear multiobjective optimization problem

$$\min_y Cy \quad \text{s.t.} \quad Ax + By \leq d, \quad (3.7)$$

where $d \in \mathbb{R}^k$, $C \in \mathbb{R}^q \times \mathbb{R}^m$, $A \in \mathbb{R}^k \times \mathbb{R}^n$ and $B \in \mathbb{R}^k \times \mathbb{R}^m$. To state the corresponding result, let Γ denotes the simplex defined by

$$\Gamma := \left\{ \alpha \in \mathbb{R}^q : \alpha \geq 0, \sum_{i=1}^q \alpha_i = 1 \right\}. \quad (3.8)$$

Proposition 3.4 *Consider a family of problems $(L[x])_{x \in X}$ defined in (3.7) with $X \subseteq \mathbb{R}^n$, and let the corresponding version of the set-valued mapping S (1.2) for problem (3.7) be uniformly bounded on X ; i.e., there exists some $k > 0$ such that for all $x \in X$ and $y \in S(x)$, $\|y\| \leq k$. Furthermore, suppose that there exists a constant $\delta > 0$ such that for all $\alpha \in \Gamma$, $x \in X$, and $y \in S(x)$, we have $\alpha^T Cy \geq \delta$. Then the UWSM condition holds.*

Proof Let $x \in X$ and consider the family of sets

$$S_\alpha(x) := \arg \min_y \{ \alpha^T Cy : Ax + By \leq d \} \quad \text{for } \alpha \in \Gamma. \quad (3.9)$$

Given that the set-valued mapping $G(x) = \{y \in \mathbb{R}^m : Ax + By \leq d\}$ is a polyhedral and convex-valued, it follows from [1] (see also [20, Theorem 3.3, pp 96]) that there are finitely many vectors $\alpha_1(x), \dots, \alpha_s(x)$ of the set Γ (3.8) such that we have

$$S(x) = \bigcup_{j=1}^s S_{\alpha_j(x)}(x). \quad (3.10)$$

Let $y \in G(x)$. If $y \in S(x)$, then (3.2) holds true. Otherwise, considering any $y \in G(x) \setminus S(x)$, we have $0 \notin Cy - \Phi(x)$. Now, let $z \in \Phi(x)$, then there is some $\tilde{y} \in S(x)$ such that $z = C\tilde{y}$ and $Cy - C\tilde{y} \neq 0$. On the other side, setting $a = \alpha_j^T(x)C$ and $b = \alpha_j^T(x)C\tilde{y}$ and using Hoffman's lemma (see [16, Theorem 1]) it follows from (3.9) and (3.10)

$$d(y, S(x)) \leq d(y, S_{\alpha_j(x)}(x)) \leq k\delta^{-1}\alpha_j^T(x)(Cy - C\tilde{y}),$$

where k is the constant appearing in uniform boundedness of S . Hence,

$$\begin{aligned} d(y, S(x)) &\leq k\delta^{-1} \|\alpha_j(x)\|_1 \|Cy - C\tilde{y}\|, \\ &\leq \lambda^{-1} \|Cy - C\tilde{y}\|, \end{aligned}$$

where $\lambda^{-1} = k\delta^{-1}$ and $\|\alpha_j(x)\|_1 = 1$. It follows from the last inequality that

$$Cy - C\tilde{y} \notin \lambda d(y, S(x)) \mathbb{U}_q.$$

Finally, since $z = C\tilde{y}$ is arbitrary, we get $Cy - \Phi(x) \cap \lambda d(y, S(x)) \mathbb{U}_q = \emptyset$. This means that for all $z \in \Phi(x)$, $Cy - z \notin \lambda d(y, S(x)) \mathbb{U}_q$. Consequently, $\lambda d(y, S(x)) \leq \|Cy - z\|$

for all $z \in \Phi(x)$. This implies that $\lambda d(y, S(x)) \leq d(Cy, \Phi(x))$. Hence, the result follows. \square

Next, we provide an example where all the assumptions of proposition 3.4 are satisfied.

Example 3.1 Setting $X := [4, \infty) \times [3, \infty)$ and considering problem (3.7) with

$$C := \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}, \quad A := \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ -1 & 0 \\ 0 & -1 \end{pmatrix}, \quad B := \begin{pmatrix} 1 & 0 \\ -1 & 0 \\ 0 & 2 \\ 0 & -1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{and} \quad d := \begin{pmatrix} 4 \\ -1 \\ 6 \\ -2 \\ 0 \\ 0 \end{pmatrix}, \quad (3.11)$$

we can easily check that for any $x \in X$ and y such that $Ax + By \leq d$, taking any $(\mu, \nu) \in \mathbb{R}_+^2$ such that $\mu + \nu = 1$, we have the inequality

$$(\mu, \nu)Cy = 2\mu y_1 + \nu y_2 \geq 2.$$

In case the uniform boundedness of the set-valued mapping S required in Proposition 3.4 is not satisfied, we can use the following alternative result.

Proposition 3.5 Consider a family of problems $(L[x])_{x \in X}$ defined in (3.7) with $X \subseteq \mathbb{R}^n$ such that for all $x \in X$ and $j \in \{1, \dots, s\}$, the sets $S_{\alpha_j}(x)$ from (3.10) are unbounded. Furthermore, suppose that there exist $\delta > 0$ and a unit vector $z \in \mathbb{R}^q$ such that for all a constant $\alpha \in \Gamma$ and $x \in X$, we have $\alpha^T Cz \geq \delta > 0$. Then the UWSM condition holds.

Proof Its follows on the path of Proposition 3.4. We shall argue in the same way as above and use [16, Theorem 2] instead of [16, Theorem 1]. \square

The next result provides a sufficient condition for the existence of uniform weak sharp minimum tailored to a more general multiobjective bilevel optimization problem.

Theorem 3.6 The UWSM condition holds for any general family of problems $(L[x])_{x \in X}$, where f is Lipschitz continuous in y uniformly in $x \in X$, the set $Y(x)$ is closed for any fixed $x \in X$, and there exists a strictly positive number λ such that we have

$$\|\varsigma\| \geq \lambda^{-1}, \quad \forall \varsigma \in \partial_y \langle y^*, f \rangle(x, y) + N(y, Y(x)), \\ \forall y^* \in N(f(x, y), z - \mathbb{R}_+^q), \quad z \in \Phi(x), \quad (x, y) \in \text{gph } Y, \quad y \notin S(x).$$

Proof Consider any closed subset Λ of \mathbb{R}^m , a locally Lipschitz function $\phi : \mathbb{R}^m \rightarrow \mathbb{R}^q$ with constant L , a vector $z \in \mathbb{R}^q$, and the set

$$\Xi(\phi, z) = \{y \in \Lambda : \phi(y) \leq z\}$$

and the function

$$\phi_z^+(y) = d(\phi(y), z - \mathbb{R}_+^q) = \max_{i=1}^q (\phi_i(y) - z_i)_+,$$

where the distance function is defined by the max norm on \mathbb{R}^q and $a_+ = \max\{a, 0\}$. Now, let us show that if there exist $\lambda > 0$ and $0 < \epsilon \leq +\infty$ such that

$$\|\varsigma\| \geq \lambda^{-1} \quad (3.12)$$

for all $\varsigma \in \partial\langle y^*, \phi \rangle(y) + N(y, \Lambda)$, $y^* \in N(\phi(y), z - \mathbb{R}_+^q)$, $y \in \Lambda$, and $0 < \phi_i(y) - z_i < \epsilon$ for some i , then we have

$$d(y, \Xi(\phi, z)) \leq \lambda \phi_z^+(y), \quad \forall y \in \Lambda \text{ such that } \phi_z^+(y) < \epsilon(1 + L\lambda)^{-1}. \quad (3.13)$$

First, by contraposition, suppose that there exist $\bar{y} \in \Lambda$ such that

$$\lambda \phi_z^+(\bar{y}) < d(\bar{y}, \Xi(\phi, z)) \quad \text{and} \quad \phi_z^+(\bar{y}) < \epsilon(1 + L\lambda)^{-1}. \quad (3.14)$$

It is obvious, by choosing suitable $r > 1$, that the following inequalities hold

$$\delta < d(\bar{y}, \Xi(\phi, z)) \quad \text{and} \quad \phi_z^+(\bar{y}) < \epsilon(1 + rL\lambda)^{-1}$$

with $\delta = r\lambda\phi_z^+(\bar{y})$. Now, observing that

$$\phi_z^+(\bar{y}) \leq \inf_{y \in \Lambda} \phi_z^+(y) + \delta(r\lambda)^{-1},$$

one can deduce that

$$\psi(\bar{y}) \leq \inf_{y \in \Lambda} \psi(y) + \epsilon$$

with $\psi(y) = \phi_z^+(y) + \delta_\Lambda(y)$, δ_Λ is the indicator function of the set Λ and $\epsilon = \delta(r\lambda)^{-1}$. Hence, applying the variational principle of Ekeland we find $v \in \Lambda$ such that

$$\begin{cases} \|v - \bar{y}\| \leq \delta, \\ \psi(v) \leq \psi(y) + (r\lambda)^{-1} \|y - v\| \quad \text{for all } y \in \Lambda. \end{cases} \quad (3.15)$$

Hence, v is a minimum of the function $y \mapsto \psi(y) + (r\lambda)^{-1} \|y - v\|$ and we get, by exploiting the chain rule, that

$$0 \in \partial\phi_z^+(v) + N(v, \Lambda) + (r\lambda)^{-1} \mathbb{B}_m. \quad (3.16)$$

In view of [22, Theorem 1.97 and Corrolary 3.43] it follows that

$$\partial\phi_z^+(v) \subset \bigcup_{y^* \in N(\phi(v), z - \mathbb{R}_+^q)} \partial\langle y^*, \phi \rangle(v).$$

Consequently, there exist $y^* \in N(\phi(v), z - \mathbb{R}_+^q)$ and $\varsigma \in \partial \langle y^*, \phi \rangle(v) + N(v, \Lambda)$ such that (3.16) yields

$$\|\varsigma\| \leq (r\lambda)^{-1} < \lambda^{-1}.$$

According to (3.14), (3.15) and $v \in \Lambda$, we have $v \notin \Xi(\phi, z)$. Consequently, $0 < \phi_i(v) - z_i < \epsilon$ for some i . On the other hand, since $\|v - \bar{y}\| \leq \delta$, the condition (3.14) guarantees that

$$\begin{aligned} \phi_z^+(v) &\leq \phi_z^+(\bar{y}) + L \|v - \bar{y}\|, \\ &\leq \phi_z^+(\bar{y}) + L\delta, \\ &\leq \phi_z^+(\bar{y})(1 + Lr\lambda), \\ &\leq \epsilon(1 + Lr\lambda)^{-1}(1 + Lr\lambda), \\ &\leq \epsilon. \end{aligned}$$

Since $\phi_i(v) - z_i \leq \phi_z^+(v)$, we deduce that $\|\varsigma\| \leq (r\lambda)^{-1} < \lambda^{-1}$ and $\phi_i(v) - z_i \leq \epsilon$, which contradict (3.12) and justifies the required inclusion (3.13).

Secondly, taking $\phi(v) = f(x, y)$, $\Lambda = Y(x)$, $z \in \Phi(x)$, and observing that

$$\Xi(\phi, z) = \{y \in Y(x) \mid f(x, y) \leq z\} \subset S(x),$$

it holds that

$$\begin{aligned} d(y, S(x)) &\leq d(y, \Xi(z, f)), \\ &\leq \lambda d(f(x, y), z - \mathbb{R}_+^q), \\ &\leq \lambda d(f(x, y), z). \end{aligned}$$

Since, z is arbitrary in $\Phi(x)$, then $d(y, S(x)) \leq \lambda d(f(x, y), \Phi(x))$. \square

To conclude this section, we provide another sufficient condition for the LUWSM condition based on the R-regularity concept introduced in Subsection 2.1. To proceed, observe that the lower-level efficient solution mapping S (1.2) can be rewritten as

$$S(x) = \{y \in \mathbb{R}^m : d(f(x, y), \Phi(x)) \leq 0, d((x, y), \text{gph } Y) \leq 0\}.$$

Hence, we will say that the R-regularity constraint qualification (RRCQ) holds at the point $(\bar{x}, \bar{y}) \in \text{gph } S$ if S is R-regular (2.4) at (\bar{x}, \bar{y}) w.r.t. $\text{dom } S$.

Proposition 3.7 *If RRCQ holds at (\bar{x}, \bar{y}) and there is some neighborhood $U \subset \mathbb{R}^n$ of \bar{x} such that $\text{dom } Y \cap U = \text{dom } S \cap U$, then LUWSM is satisfied at (\bar{x}, \bar{y}) .*

Proof Fix $(\bar{x}, \bar{y}) \in \text{gph } S$. Since, the mapping S is R-regular at (\bar{x}, \bar{y}) w.r.t. $\text{dom } S$, there exist $\sigma > 0$ and $\epsilon > 0$ such that for all $(x, y) \in \mathbb{U}_\epsilon(\bar{x}, \bar{y}) \cap (\text{dom } S \times \mathbb{R}^m)$ we have the inequality

$$d(y, S(x)) \leq \sigma \max\{0, d(f(x, y), \Phi(x)), d((x, y), \text{gph } Y)\}.$$

From the definition of the frontier map, for any $(x, y) \in \bigcup_{\epsilon} (\bar{x}, \bar{y})$ with $d((x, y), \text{gph } Y) = 0$, we have the inequality $d(f(x, y), \Phi(x)) \geq 0$. Hence, for all $(x, y) \in \bigcup_{\epsilon} (\bar{x}, \bar{y}) \cap (\text{dom } S \times \mathbb{R}^m)$,

$$y \in Y(x) \implies d(y, S(x)) \leq \sigma d(f(x, y), \Phi(x)).$$

On the other hand, we can choose an open ball around (\bar{x}, \bar{y}) which is contained in $\bigcup_{\epsilon} (\bar{x}, \bar{y}) \cap (\text{dom } S \times \mathbb{R}^m)$. Dividing (3) by σ , we get the result. \square

Finally, note that the relationships between all the constraint qualifications discussed above are summarized in Fig. 1.

4 Necessary optimality conditions

Our aim in this section is to use the GVFCQ, introduced and studied in the previous section, to derive necessary optimality conditions for problem (MUL)–(L[x]). To proceed, we consider the set

$$\Pi := (X \times \mathbb{R}^m) \cap \text{gph } S \subset \mathbb{R}^n \times \mathbb{R}^m \quad (4.1)$$

and the set-valued mapping

$$\Sigma(x) := f(x, Y(x)) := \{f(x, y) : y \in Y(x)\} \quad \text{for } x \in \mathbb{R}^n. \quad (4.2)$$

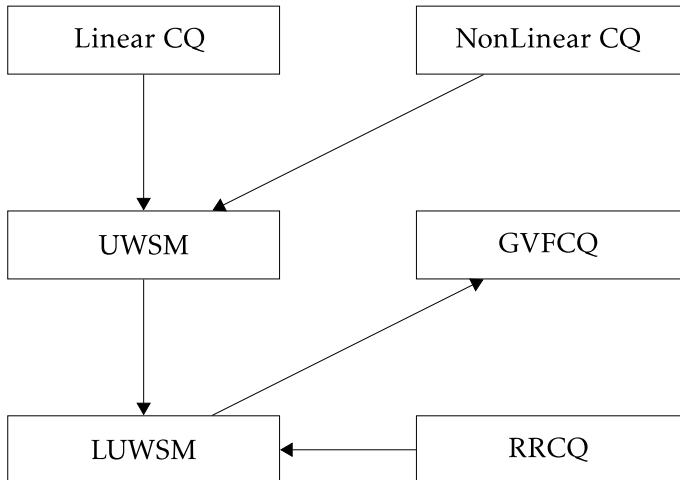


Fig. 1 Here, *Linear CQ* refers to the assumptions in Proposition 3.4 or Proposition 3.5, while *NonLinear CQ* represents the collection of assumptions in Theorem 3.6. As for *UWSM* (resp. *LUWSM*), it is the abbreviation used for the *uniform weak sharp minimum* (resp. *local uniform weak sharp minimum*) condition. Finally, *GVFCQ* denotes the generalized value function constraint qualification and *RRCQ* corresponds to the R-regularity of the set-valued mapping S , see (MUL)–(L[x]), according to (2.4)

In the process, the upper estimates for coderivatives of the optimal solution set-valued mapping S and the frontier map Φ will also be useful. To specifically compute an estimate of the coderivative of Φ^E (see (1.1) and related discussion), we additionally need the following *strong domination property* for the lower-level problem $(L[x])$:

$$f(x, Y(x)) \subset \Phi^E(x) + \mathbb{R}_+^q \quad \forall x \in X. \quad (4.3)$$

This property has been used in the literature with different names; for example, it is used in [26], where is called \mathbb{R}_+^q -minicomplete property, and utilized to estimate the contingent derivative of the set-valued mapping Σ . However, we borrow our vocabulary from the following weaker *domination property* used in [17]:

$$f(x, Y(x)) + \mathbb{R}_+^q = \Phi^E(x) + \mathbb{R}_+^q.$$

To construct an estimate of the coderivative of Φ^E in the next result, we also need the limiting qualification condition at a reference point (\bar{x}, \bar{y}) :

$$D^*S(\bar{x}, \bar{y}) \cap (-N(\bar{x}, X)) = \{0\}. \quad (4.4)$$

Theorem 4.1 *Let $(\bar{x}, \bar{y}) \in \text{gph } S$. Suppose that f is locally Lipschitzian around (\bar{x}, \bar{y}) , that the graph of the image map Σ is locally compact around \bar{x} , that Y is locally closed around (\bar{x}, \bar{y}) with $\bar{z} = f(\bar{x}, \bar{y})$ and the strong domination property (4.3) is satisfied. Suppose in addition that Y is Lipschitz-like around (\bar{x}, \bar{y}) . Then, it holds that*

$$D^*\Phi^E(\bar{x}, \bar{z})(z^*) \subset \bigcup_{(x^*, y^*) \in D^*f(\bar{x}, \bar{y})(y^*)} [x^* + D^*Y(\bar{x}, \bar{y})(y^*)], \quad \text{for all } z^* \in \mathbb{R}^q \quad (4.5)$$

and Φ^E is Lipschitz-like around (\bar{x}, \bar{z}) . Furthermore, if the function f is strictly differentiable at (\bar{x}, \bar{y}) , then for any $z^* \in \mathbb{R}^q$, we have

$$D^*\Phi^E(\bar{x}, \bar{z})(z^*) \subset \nabla_x f(\bar{x}, \bar{y})^* z^* + D^*Y(\bar{x}, \bar{y})(\nabla_y f(\bar{x}, \bar{y})^* z^*).$$

Proof First, observe that the image map Σ in (4.2) has a composite form. Hence, applying to this composition the coderivative chain rule from [22, Theorem 3.18(i)] for the locally Lipschitzian cost mapping $f(x, y)$, we get

$$D^*\Sigma(\bar{x}, \bar{z})(z^*) \subset \bigcup_{(x^*, y^*) \in D^*f(\bar{x}, \bar{y})(z^*)} [x^* + D^*Y(\bar{x}, \bar{y})(y^*)], \quad z^* \in \mathbb{R}^q. \quad (4.6)$$

Fix $z^* \in \mathbb{R}^q$ and let us prove that $D^*\Phi^E(\bar{x}, \bar{z})(z^*) \subset D^*\Sigma(\bar{x}, \bar{z})(z^*)$. Let $x^* \in D^*\Phi^E(\bar{x}, \bar{z})(z^*)$. Based on (2.2), there are sequences $(x_k, z_k) \xrightarrow{\text{gph } \Phi^E} (\bar{x}, \bar{z})$ and $(x_k^*, z_k^*) \rightarrow (x^*, z^*)$ such that

$$\limsup_{(x_{k_s}, z_{k_s}) \xrightarrow{\text{gph } \Phi^E} (x_k, z_k)} \frac{\langle x^*, x_{k_s} - x_k \rangle - \langle z^*, z_{k_s} - z_k \rangle}{\|x_{k_s} - x_k\| + \|z_{k_s} - z_k\|} \leq 0.$$

We claim that in some neighborhood U of (\bar{x}, \bar{z}) for any $(x_{k_s}, z_{k_s}) \in U$ such that

$$(x_{k_s}, z_{k_s}) \xrightarrow{\text{gph } \Phi^E} (x_k, z_k), \quad \text{one has} \quad (x_{k_s}, z_{k_s}) \xrightarrow{\text{gph } \Sigma} (x_k, z_k).$$

Indeed, suppose the contrary to our claim, that there exists

$$(x_{k_s}, z_{k_s}) \in \text{gph } \Sigma \setminus \text{gph } \Phi^E \quad \text{such that} \quad (x_{k_s}, z_{k_s}) \rightarrow (x_k, z_k).$$

It follows immediately from the strong domination property (4.3) that $(x_{k_s}, z_{k_s}) \in \text{epi } \Phi^E$. Since Σ is locally compact around \bar{x} it follows from [19, Proposition 4.3 (iv)] that Φ^E is locally order semicontinuous around (\bar{x}, \bar{z}) . Hence, for $(x_{k_s}, z_{k_s}) \in \text{epi } \Phi^E$, there exists a sequence $(x_{k_s}, t_{k_s}) \in \text{gph } \Phi^E$ such that $z_{k_s} \in t_{k_s} + \mathbb{R}_+^q$. Then applying [18, Theorem 1.3], we get a contradiction while considering the fact that $D^*\Sigma(\bar{x}, \bar{z})(0) = \{0\}$, which results from the Lipschitz-likeness of Y and the inclusion in (4.6). The above arguments ensures that $x^* \in D^*\Sigma(\bar{x}, \bar{z})(z^*)$. Combining this with (4.6), the desired result follows. \square

Note that our formula in (4.5) is the same as the one obtained in [17]. However, in the later reference, it is required that z^* be in the interior of the corresponding cone; such a requirement is very restrictive and will not make it possible to construct the necessary optimality conditions, which represent our main goal in this section. Furthermore, under the strong domination property (4.3), the paper [27] provides an estimate of the coderivative of Φ^E for all z^* in the uniformly positive polar to cone \mathbb{R}_+^q defined by

$$K_{up}^* := \{\alpha \in \mathbb{R}^q : \exists \beta > 0, \langle \alpha, z \rangle \geq \beta \|z\|, \forall z \in \mathbb{R}_+^q\}.$$

As it can be seen in Theorem 4.1, our estimate of the coderivative of Φ^E is calculated at any point $z^* \in \mathbb{R}^q$, thus enabling an easy derivation of optimality conditions for problem (MUL)–(L[x]), as it will be clear by the end of this section. It is also important to note that a version of the strong domination property can well be defined for Φ^W . However, it is unclear how it would help in obtaining an estimate of the coderivative of Φ^W analogous to the one derived in Theorem 4.1 for efficient Pareto points.

The next proposition gives a sufficient condition for the family of parametric linear programming problems (3.7) to satisfy the strong domination property (4.3).

Proposition 4.2 *Assume that for all $x \in X$, the set $Y(x) = \{y \in \mathbb{R}^m : Ax + By \leq d\}$ is bounded. Then, problem (3.7) satisfies the strong domination property (4.3).*

Proof Fix $x \in X$ and let $\bar{y} \in Y(x)$. If $\bar{y} \in S(x)$, then since $C\bar{y} \in C\bar{y} + \mathbb{R}_+^q$, we have the inclusion $f(x, Y(x)) \subset \Phi^E(x) + \mathbb{R}_+^q$. Suppose that $\bar{y} \notin S(x)$ and consider the set

$$\mathcal{DP}(x, \bar{y}) = \{y \in G(x) \mid C\bar{y} - Cy \in \mathbb{R}_+^q\}.$$

Since $Y(x)$ is bounded, $\mathcal{DP}(x, \bar{y})$ is also bounded. Hence, its support function $\sigma(\cdot, \mathcal{DP}(x, \bar{y}))$ is defined everywhere. Now, choose $u \in \text{int } \mathbb{R}_+^q$. Then, from [24,

Corollary 23.5.3] there exists $z \in \mathcal{DP}(x, \bar{y})$ such that $\langle C^T u, z \rangle = \sigma(C^T u, \mathcal{DP}(x, \bar{y}))$. We claim that $z \in S(x)$. Indeed, suppose, in contrary to our claim, that there exist $v \in Y(x)$ such that

$$Cv - Cz \in -\mathbb{R}_+^q \text{ and } Cv \neq Cz.$$

Or equivalently, that

$$Cv - Cz \in -\mathbb{R}_+^q \setminus \{0\} \text{ for some } v \in Y(x).$$

On the one side, $C\bar{y} - Cv = C\bar{y} - Cz + Cz - Cv \in \mathbb{R}_+^q + \mathbb{R}_+^q \setminus \{0\} \subset \mathbb{R}_+^q$. Consequently, $v \in \mathcal{DP}(x, \bar{y})$. On the other side, since, $u \in \text{int } \mathbb{R}_+^q$, one has $0 < \langle u, Cv - Cz \rangle$. Thus, $\sigma(C^T u, \mathcal{DP}(x, \bar{y})) < \langle C^T u, v \rangle$. Which is a contradiction. Finally, we have $z \in \mathcal{DP}(x, \bar{y})$ and $Cz \in \Phi^E(x)$, it follows that $C\bar{y} - Cz \in \mathbb{R}_+^q$, which concludes the proof. \square

Now, we come to the final step before the statement of the main result of this section; i.e., we provide an estimate for the coderivative of the lower-level optimal solution set-valued mapping S (1.2) under the GVFCQ (3.1).

Proposition 4.3 *Consider the lower-level optimal solution set-valued mapping S (1.2) and suppose that f is locally Lipschitz continuous and the set $\text{gph } Y$ and $\text{gph } \Phi$ are closed. Furthermore, assume that the GVFCQ holds at (\bar{x}, \bar{y}) . Then it holds that*

$$D^*S(\bar{x}, \bar{y})(y^*) \subset \bigcup_{(u^*, v^*) : u^* \in D^*\Phi(\bar{x}, f(\bar{x}, \bar{y}))(-v^*)} \bigcup_{(a^*, b^*) \in D^*f(\bar{x}, \bar{y})(v^*)} \{u^* + a^* + D^*Y(\bar{x}, \bar{y})(y^* + b^*)\}.$$

If additionally, f is strictly differentiable, then we have

$$D^*S(\bar{x}, \bar{y})(y^*) \subset \bigcup_{(u^*, v^*) : u^* \in D^*\Phi(\bar{x}, f(\bar{x}, \bar{y}))(-v^*)} \{u^* + \nabla_x f(\bar{x}, \bar{y})^\top v^* + D^*Y(\bar{x}, \bar{y})(y^* + \nabla_y f(\bar{x}, \bar{y})^\top v^*)\}.$$

Proof Note that the graph of S (1.2) can be written as

$$\text{gph } S = \Omega \cap \psi^{-1}(\Lambda) \text{ with } \Omega := \text{gph } Y, \psi(x, y) := \begin{pmatrix} x \\ f(x, y) \end{pmatrix}, \text{ and } \Lambda := \text{gph } \Phi.$$

Then, based on the assumptions made, it follows from [14, Theorem 4.1] that

$$N((\bar{x}, \bar{y}); \text{gph } S) \subset \bigcup_{(u^*, v^*) \in N(\psi(\bar{x}, \bar{y}); \Lambda)} D^*\psi(\bar{x}, \bar{y})(u^*, v^*) + N((\bar{x}, \bar{y}); \Omega).$$

Hence, considering the definition of the concept of coderivative in (2.2), we have

$$\begin{aligned} D^*S(\bar{x}, \bar{y})(y^*) &\subset \bigcup_{(u^*, v^*) \in N(\psi(\bar{x}, \bar{y}); \Lambda)} \{x^* \in \mathbb{R}^n \mid (x^*, -y^*) \in N((\bar{x}, \bar{y}); \Omega) + D^*\psi(\bar{x}, \bar{y})(u^*, v^*)\} \\ &\subset \bigcup_{(u^*, v^*) \in N(\psi(\bar{x}, \bar{y}); \Lambda)} \{x^* \in \mathbb{R}^n \mid (x^* - u^*, -y^*) \in N((\bar{x}, \bar{y}); \Omega) + D^*f(\bar{x}, \bar{y})(v^*)\} \\ &\subset \bigcup_{(u^*, v^*) \in N(\psi(\bar{x}, \bar{y}); \Lambda)} \{x^* \in \mathbb{R}^n \mid \exists (a^*, b^*) \in D^*f(\bar{x}, \bar{y})(v^*) : \\ &\quad x^* - u^* - a^* \in D^*Y(\bar{x}, \bar{y})(y^* + b^*)\} \end{aligned}$$

with the second inclusion resulting from

$$D^*\psi(\bar{x}, \bar{y})(u^*, v^*) \subset \begin{pmatrix} u^* \\ 0 \end{pmatrix} + D^*f(\bar{x}, \bar{y})(v^*).$$

Clearly, the last inclusion in the above sequence of inclusions gives the desired result for the upper bound of $D^*S(\bar{x}, \bar{y})(y^*)$ when f is locally Lipschitz continuous. The case where f is strictly differentiable obviously follows from $D^*f(\bar{x}, \bar{y})(v^*) = \nabla f(\bar{x}, \bar{y})^\top v^*$. \square

What is nice with this result is not the construct process of the proof, which is not necessarily new, but its reliance on the GVFCQ and the corresponding rich set of sufficient conditions provided in the previous section. Such an approach does not seem to have been used in the literature to construct an estimate of the coderivative of the optimal solution set-valued mapping of a parametric multiobjective optimization problem.

We are now ready to state one of the main results of this paper, providing new necessary optimality conditions for the multiobjective bilevel optimization problem (MUL)–(L[x]).

Theorem 4.4 *Let (\bar{x}, \bar{y}) be a local efficient/weakly efficient Pareto point for problem (MUL)–(L[x]). We assume that the function F and f are Lipschitz continuous around (\bar{x}, \bar{y}) and suppose that X , $\text{gph } S$, $\text{gph } Y$, and $\text{gph } \Phi$ are closed sets. Furthermore, assume that the GVFCQ holds at (\bar{x}, \bar{y}) . Then, there exist vectors $v^* \in \mathbb{R}^q$ and $w^* \in \mathbb{R}_+^p$ with $\|w^*\| = 1$ such that*

$$0 \in \partial\langle w^*, F \rangle(\bar{x}, \bar{y}) + \partial\langle v^*, f \rangle(\bar{x}, \bar{y}) + D^*\Phi(\bar{x}, f(\bar{x}, \bar{y}))(-v^*) \times \{0\} + N((\bar{x}, \bar{y}); \text{gph } Y) + N(\bar{x}; X) \times \{0\}. \quad (4.7)$$

If $\Phi = \Phi^E$ in (4.7) and additionally, $\text{gph } \Sigma$ is locally compact around \bar{x} , Y is locally closed and Lipschitz-like around (\bar{x}, \bar{y}) , and the strong domination property (4.3) is satisfied. Then, there exist vectors $v^ \in \mathbb{R}^q$, $(\alpha^*, \beta^*) \in \partial\langle -v^*, f \rangle(\bar{x}, \bar{y})$, and $w^* \in \mathbb{R}_+^p$ with $\|w^*\| = 1$ such that*

$$(-\alpha^*, 0) \in \partial\langle w^*, F \rangle(\bar{x}, \bar{y}) + \partial\langle v^*, f \rangle(\bar{x}, \bar{y}) + D^*Y(\bar{x}, \bar{y})(\beta^*) \times \{0\} + N((\bar{x}, \bar{y}); \text{gph } Y) + N(\bar{x}; X) \times \{0\}. \quad (4.8)$$

If additionally, F and f are strictly differentiable at the point (\bar{x}, \bar{y}) , then there exist vectors $v^ \in \mathbb{R}^q$ and $w^* \in \mathbb{R}_+^p$ with $\|w^*\| = 1$ such that we have*

$$0 \in \nabla_x F(\bar{x}, \bar{y})^\top w^* + D^*Y(\bar{x}, \bar{y})(-\nabla_y f(\bar{x}, \bar{y})^\top v^*) + D^*Y(\bar{x}, \bar{y})(\nabla_y F(\bar{x}, \bar{y})^\top w^* + \nabla_y f(\bar{x}, \bar{y})^\top v^*) + N(\bar{x}; X). \quad (4.9)$$

Proof First start by noting that based on (4.1), problem (MUL)–(L[x]) can be rewriting as

$$\min_{x,y} F(x,y) := (F_1(x,y), \dots, F_p(x,y))^T \quad \text{s.t.} \quad (x,y) \in \Pi.$$

Since the set Π is closed, as intersection of two closed sets, and the function F is Lipschitz continuous around the point (\bar{x}, \bar{y}) , which is a local efficient/weakly efficient Pareto point for problem (MUL)–(L[x]), there exists $w^* \in \mathbb{R}_+^p$ with $\|w^*\| = 1$ such that we have

$$\partial \langle w^*, F \rangle (\bar{x}, \bar{y}) + N((\bar{x}, \bar{y}); \Pi), \quad (4.10)$$

according to [2, Theorem 5.3]. Hence, it suffices now to calculate an upper estimate of the normal cone to Π . Using the intersection rule from [23, Corollary 3.5], one has

$$N((\bar{x}, \bar{y}); \Pi) \subset N((\bar{x}, \bar{y}); X \times \mathbb{R}^m) + N((\bar{x}, \bar{y}); \text{gph } S) \quad (4.11)$$

as X and $\text{gph } S$ are locally closed around \bar{x} and (\bar{x}, \bar{y}) , respectively, and provided that

$$N((\bar{x}, \bar{y}); X \times \mathbb{R}^m) \cap (-N((\bar{x}, \bar{y}); \text{gph } S)) = \{0\}. \quad (4.12)$$

is satisfied. Considering the coderivative calculus rules, we can easily show that the fulfilment of (4.4) implies that (4.12) holds. Then combining (4.10) and (4.11), one gets

$$0 \in \partial \langle w^*, F \rangle (\bar{x}, \bar{y}) + N((\bar{x}, \bar{y}); X \times \mathbb{R}^m) + N((\bar{x}, \bar{y}); \text{gph } S). \quad (4.13)$$

Hence, there exist $(x^*, y^*) \in \partial \langle w^*, F \rangle (\bar{x}, \bar{y})$ and $c^* \in N(\bar{x}; X)$ such that

$$(-x^* - c^*, -y^*) \in N((\bar{x}, \bar{y}); \text{gph } S) \quad \text{or equivalently,} \quad -x^* - c^* \in D^*S(\bar{x}, \bar{y})(y^*). \quad (4.14)$$

Thanks to the upper estimate of coderivative of the optimal solution set-valued mapping S in Proposition 4.3, we can find vectors u^*, v^*, a^* and b^* such that

$$\left. \begin{aligned} u^* &\in D^*\Phi(\bar{x}, f(\bar{x}, \bar{y}))(-v^*) \\ (a^*, b^*) &\in D^*f(\bar{x}, \bar{y})(v^*) \\ -x^* - c^* &\in u^* + a^* + D^*Y(\bar{x}, \bar{y})(y^* + b^*) \end{aligned} \right\} \quad (4.15)$$

given that $\text{gph } Y$ and $\text{gph } \Phi$ are closed, and f is Lipschitz continuous around (\bar{x}, \bar{y}) . Then combining (4.13), (4.14), and (4.15), we immediately arrive at (4.7).

As for the inclusion in (4.8), it follows from Theorem 4.1 that one has an upper estimate for coderivative of frontier map Φ^E with respect to local Pareto optimality concept

$$D^*\Phi^E(\bar{x}, \bar{z})(-v^*) \subseteq \bigcup_{(\alpha^*, \beta^*) \in D^*f(\bar{x}, \bar{y})(-v^*)} \left[\alpha^* + D^*Y(\bar{x}, \bar{y})(\beta^*) \right] \quad (4.16)$$

considering the assumptions made. Substituting (4.16) into (4.7), it follows that we can find $(\alpha^*, \beta^*) \in D^*f(\bar{x}, \bar{y})(-v^*)$ such that we have (4.8), which obviously leads to (4.9) under the additional differentiability assumptions. \square

Remark 4.5 Recall that the CQ (4.4) is automatically satisfied at (\bar{x}, \bar{y}) provided that problem (MUL)–(L[x]) has no upper-level constraints (i.e., $X = \mathbb{R}^n$) or the lower-level optimal solution set-valued mapping S is Lipschitz-like around (\bar{x}, \bar{y}) , which is automatically the case if an upper estimate of $D^*S(\bar{x}, \bar{y})(0)$ from Proposition 4.3 equals to zero, thanks to the Mordukhovich criterion [22, 23]; also see [31, 32] for further details and references.

To have a clear view of the fact that the necessary optimality conditions obtained in Theorem 4.4 represent a natural extension of the those from a standard optimistic bilevel optimization problem, consider problem (MUL)–(L[x]) with $p = 1$ and $q = 1$. Let (\bar{x}, \bar{y}) be a local optimal solution of the problem in this case. If the point satisfies the corresponding version of CQ (4.12) and F and f are strictly differentiable, then we have

$$0 \in \nabla_x F(\bar{x}, \bar{y}) + D^*S(\bar{x}, \bar{y})(\nabla_y F(\bar{x}, \bar{y})) + N(\bar{x}; X). \quad (4.17)$$

This inclusion obviously coincides with (4.13) in this context where w^* reduces to 1. Secondly, if φ denotes the optimal value function of the corresponding parametric optimization problem (L[x]) and we additionally suppose that the function φ is lower semicontinuous around \bar{x} and the set-valued mapping

$$\Psi_\varphi(v) := \{(x, y) \in \text{gph } Y : f(x, y) - \varphi(x) + v = 0\}. \quad (4.18)$$

is calm at $(0, \bar{x}, \bar{y})$, then condition (4.17) can be detailed further to obtain

$$0 \in \nabla F(\bar{x}, \bar{y}) + \nabla f(\bar{x}, \bar{y})v^* + \partial\langle -v^*, \varphi \rangle(\bar{x}) \times \{0\} + N((\bar{x}, \bar{y}); \text{gph } Y) + N(\bar{x}; X) \times \{0\} \quad (4.19)$$

for some $v^* \in \mathbb{R}$. Similarly, this coincides with (4.7) for $w^* = 1$. Note that the set-valued mapping (4.18) is slightly different from (3.2), as in the latter case, we instead have an inequality on the perturbed value function constraint. Of course, using the version of the set-valued mapping in (3.2) would have led to $v^* \geq 0$.

Finally, still in the case $p = 1$ and $q = 1$, if S is inner semicontinuous and has a closed graph around (\bar{x}, \bar{y}) , and φ is Lipschitz continuous around \bar{x} , then from (4.19), we have inclusion (4.9) with the corresponding $w^* = 1$. For more background details on the constructions and relevant concepts above, in the context of standard optimistic optimization, interested readers are referred to [32], where, unlike in (4.9), a scalarization approach is used to deal with the lower-level multiobjective problem, as, to the best of our knowledge, it is the case for all previous references on necessary optimality conditions for multiobjective bilevel optimization.

5 Application to smooth constraint functionals

Let us consider the multiobjective bilevel optimization problem (MUL)–(L[x]) in the case where the upper- and lower-level feasible sets are defined by

$$X := \{x \in \mathbb{R}^n : G(x) \leq 0\} \text{ and } Y(x) := \{y \in \mathbb{R}^m : g(x, y) \leq 0\},$$

respectively, with $G : \mathbb{R}^n \rightarrow \mathbb{R}^r$ and $g : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^s$ being continuously differentiable functions. The upper-level regularity condition will be said to hold at \bar{x} if there exists a vector $d \in \mathbb{R}^n$ such that we have

$$\nabla G_i(\bar{x})^\top d < 0 \quad \text{for all } i \in I_G(\bar{x}) := \{i \in \{1, \dots, r\} : G_i(\bar{x}) = 0\}.$$

Similarly, the lower-level regularity condition will be satisfied at (\bar{x}, \bar{y}) if there exists a vector $d \in \mathbb{R}^{n+m}$ that verifies

$$\nabla_y g_j(\bar{x}, \bar{y})^\top d < 0 \quad \text{for all } j \in I_g(\bar{x}, \bar{y}) := \{j \in \{1, \dots, s\} : g_j(\bar{x}, \bar{y}) = 0\}.$$

Obviously, these upper- and lower-level regularity conditions correspond to the Mangasarian-Fromovitz constraint qualification for the feasible set of the corresponding (upper- or lower-) level of our problem (MUL)–(L[x]). It is well-known that under the upper- and lower-level regularity conditions, we respectively have

$$N(\bar{x}; X) = \{\nabla G(\bar{x})^\top u : u \geq 0, u^\top G(\bar{x}) = 0\}$$

and

$$D^*Y(\bar{x}, \bar{y})(y^*) = \{\nabla_x g(\bar{x}, \bar{y})^\top v : -y^* = \nabla_y g(\bar{x}, \bar{y})^\top v, v \geq 0, v^\top g(\bar{x}, \bar{y}) = 0\}.$$

Now, in addition to all the assumptions of Theorem 4.4, we assume that the upper- and lower-level regularity conditions at \bar{x} and (\bar{x}, \bar{y}) , respectively, it follows from (4.9) that there exist $w^* \in \mathbb{R}_+^p$ with $\|w^*\| = 1$, $v^* \in \mathbb{R}^q$, $u \in \mathbb{R}^r$, and $v \in \mathbb{R}^s$ satisfying the relationships

$$\nabla_y f(\bar{x}, \bar{y})^\top (-v^*) + \nabla_y g(\bar{x}, \bar{y})^\top v = 0, \quad (5.1)$$

$$u \geq 0, \quad G(\bar{x}) \leq 0, \quad u^\top G(\bar{x}) = 0, \quad (5.2)$$

$$v \geq 0, \quad g(\bar{x}, \bar{y}) \leq 0, \quad v^\top g(\bar{x}, \bar{y}) = 0, \quad (5.3)$$

and such that

$$-\nabla_x F(\bar{x}, \bar{y})^* w^* - \nabla G(\bar{x})^\top u - \nabla_x g(\bar{x}, \bar{y})^\top v \in D^*Y(\bar{x}, \bar{y})(\nabla_y F(\bar{x}, \bar{y})^* w^* + \nabla_y f(\bar{x}, \bar{y})^* v^*).$$

Then from a second application of the above coderivative formula for $D^*Y(\bar{x}, \bar{y})(y^*)$ to the latter inclusion, it follows that we can find $w \in \mathbb{R}^s$, $w^* \in \mathbb{R}_+^p$ with $\|w^*\| = 1$, $u \in \mathbb{R}^r$, and $v \in \mathbb{R}^s$ such that the relationships (5.1)–(5.3) hold together with

$$\nabla F(\bar{x}, \bar{y})^\top w^* + \begin{bmatrix} \nabla G(\bar{x})^\top u \\ 0 \end{bmatrix} + \nabla g(\bar{x}, \bar{y})^\top (v + w) = 0, \quad (5.4)$$

$$w \geq 0, \quad g(\bar{x}, \bar{y}) \leq 0, \quad w^\top g(\bar{x}, \bar{y}) = 0. \quad (5.5)$$

The optimality conditions (5.1)–(5.5) are very similar to their standard optimistic bilevel optimization problem counterpart with scalar objective functions, as it can

be seen in [5, 9], for example. The corresponding conditions in the latter paper have been shown in the recent papers [12, 13, 33] to be suitable to efficiently solve standard optimistic bilevel optimization problem. Hence, the extension of the methods in these papers to multiobjective bilevel programs will be explored in future works.

Finally, to end this section, we provide the following illustrative example, where the lower-level problem is the linear parametric multiobjective problem from Example 3.1.

Example 5.1 Consider problem (MUL)–(L[x]), where $F : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}^p$ is any differentiable function and X and the lower-level problem are defined as in Example 3.1. We can easily check that for all $x \in X$, $Y(x) := \{y \in \mathbb{R}^2 \mid Ax + By \leq d\}$ is bounded. Hence, the strong domination property (4.3) is satisfied according to Proposition 4.2. Furthermore, all the other assumptions of Theorem 4.4 hold. Hence, for any local efficient Pareto point (\bar{x}, \bar{y}) of the problem,

$$\begin{aligned} \nabla_x F(\bar{x}, \bar{y})^\top w^* - \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} - \begin{pmatrix} v_5 + w_5 \\ v_6 + w_6 \end{pmatrix} &= 0, \\ \nabla_y F(\bar{x}, \bar{y})^\top w^* + \begin{pmatrix} v_1 + w_1 - v_2 - w_2 + v_5 + w_5 \\ 2(v_3 + w_3) - v_4 - w_4 + v_6 + w_6 \end{pmatrix} &= 0, \\ \begin{pmatrix} 2v_1^* \\ v_2^* \end{pmatrix} - \begin{pmatrix} v_1 - v_2 + v_5 \\ 2v_3 - v_4 + v_6 \end{pmatrix} &= 0, \\ u_1 \geq 0, \quad y_1 \geq 4, \quad u_1(y_1 - 4) &= 0, \\ u_2 \geq 0, \quad y_2 \geq 3, \quad u_2(y_2 - 3) &= 0, \\ v \geq 0, \quad A\bar{x} + B\bar{y} \leq d, \quad v(A\bar{x} + B\bar{y} - d) &= 0, \\ w \geq 0, \quad A\bar{x} + B\bar{y} \leq d, \quad w(A\bar{x} + B\bar{y} - d) &= 0, \end{aligned}$$

for some $u \in \mathbb{R}^2$, $v \in \mathbb{R}^6$, $w \in \mathbb{R}^6$, $v^* \in \mathbb{R}^2$, and $w^* \in \mathbb{R}_+^p$ with $\|w^*\| = 1$. Note that here, the Pareto efficient solution concept is also considered for the lower-level problem. The matrices A , B , and d in the last two lines of this system are given in (3.11).

Acknowledgements The work of AZ is supported by the EPSRC grant EP/V049038/1 and the Alan Turing Institute for Data Science and Artificial Intelligence under the EPSRC grant EP/N510129/1

Data availability No data is needed in this paper.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Arrow, K.J., Barankin, E.W., Blackwell, D.: Admissible points of convex sets, in Contributions to the Theory of Games, H. W. Kuhn and A. W. Tucker, eds., Princeton University Press, Princeton, New Jersey, vol 2, 87–91 (1953)
2. Bao, T.Q., Mordukhovich, B.S.: Relative Pareto minimizers for multiobjective problems: existence and optimality conditions. *Math. Program.* **122**(2), 301–47 (2010)
3. Bao, T.Q., Mordukhovich, B.S.: Necessary conditions for super minimizers in constrained multiobjective optimization. *J. Global Optim.* **43**, 533–552 (2009)
4. Bonnel, H.: Optimality conditions for the semivectorial bilevel optimization problem. *Pacific J. Optim.* **2**(3), 447–467 (2006)
5. Dempe, S., Dutta, J., Mordukhovich, B.S.: New necessary optimality conditions in optimistic bilevel programming. *Optimization* **56**(5–6), 577–604 (2007)
6. Dempe, S., Gadhi, N., Zemkoho, A.B.: New optimality conditions for the semivectorial bilevel optimization problem. *J. Optim. Theory Appl.* **157**, 54–74 (2013)
7. Dempe, S., Mehrlitz, P.: Semivectorial bilevel programming versus scalar bilevel programming. *Optimization* **69**(4), 657–679 (2020)
8. Dempe, S., Mordukhovich, B.S., Zemkoho, A.B.: Sensitivity analysis for two-level value functions with applications to bilevel programming. *SIAM J. Optim.* **22**(4), 1309–1343 (2012)
9. Dempe, S., Zemkoho, A.B.: The generalized mangasarian-fromowitz constraint qualification and optimality conditions for bilevel programs. *J. Optim. Theory Appl.* **148**(1), 46–68 (2011)
10. Dempe, S., Zemkoho, A.B.: (eds.), *Bilevel Optimization: Advances and Next Challenges*, Springer (2020)
11. Eichfelder, G.: *Methods for Multiobjective Bilevel Optimization*. In S. Dempe and A. Zemkoho (eds.), *Bilevel Optimization* Springer, Cham. 423–449 (2020)
12. Fischer, A., Zemkoho, A.B., Zhou, S.: Semismooth Newton-type method for bilevel optimization: Global convergence and extensive numerical experiments. *Optimization methods and software*, in press, (2021) <https://doi.org/10.1080/10556788.2021.1977810>
13. Fliege, J., Tin, A., Zemkoho, A.B.: Gauss Newton-type methods for bilevel optimization. *Comput. Optim. Appl.* **78**, 793–824 (2021)
14. Henrion, R., Jourani, A., Outrata, J.: On the calmness of a class of multifunctions. *SIAM J. Optim.* **13**(2), 603–618 (2002)
15. Henrion, R., Surowiec, T.: On calmness conditions in convex bilevel programming. *Appl. Anal.* **90**(6), 951–970 (2011)
16. Hui, H., Qing, W.: On approximate solutions of infinite systems of linear inequalities. *Linear Algebra Appl.* **114**(115), 429–438 (1989)
17. Huy, N.Q., Mordukhovich, B.S., Yao, J.C.: Coderivatives of frontier and solution maps in parametric multiobjective optimization. *Taiwan. J. Math.* **12**, 2083–2111 (2008)
18. Levy, A.B.: Nonsingularity conditions for multifunctions. *Set-Valued Variational Anal.* **7**, 89–99 (1999)
19. Li, S.J., Xue, X.W.: Sensitivity analysis of gap functions for vector variational inequality via coderivatives. *Optimization* **63**(7), 1075–1098 (2014)
20. Luc, D.T.: *Theory of Vector Optimization*. Springer-Verlag, Berlin (1989)
21. Mehrlitz, P., Minchenko, L.I.: R-regularity of set-valued mappings under the relaxed constant positive linear dependence constraint qualification with applications to parametric and bilevel optimization. *Set-Valued Variational Anal.* (2020). <https://doi.org/10.1007/s11228-021-00578-0>
22. Mordukhovich, B.S.: *Variational Analysis and Generalized Differentiation. I: Basic Theory*. Springer, Berlin (2006)
23. Mordukhovich, B.S.: *Variational Analysis and Generalized Differentiation. II. Applications*. Springer, Berlin (2006)
24. Rockafellar, R.T.: *Convex Analysis*. Princeton University Press, Princeton, New Jersey (1970)
25. Sinha, A., Malo, P., Deb, K.: A review on bilevel optimization: from classical to evolutionary approaches and applications. *IEEE Trans. Evol. Comput.* **22**(2), 276–295 (2017)
26. Tanino, T.: Sensitivity analysis in multiobjective optimization. *J. Optim. Appl.* **56**(3), 479–499 (1988)
27. Xue, X.W., Li, S.J., Lia, C.M., Yao, J.C.: Sensitivity analysis of parametric vector set-valued optimization problems via coderivatives. *Taiwan. J. Math.* **15**(6), 2533–2554 (2011)

28. Ye, J.J.: New uniform parametric error bounds. *J. Optim. Theory Appl.* **98**(1), 197–219 (1998)
29. Ye, J.J., Zhu, D.L.: Optimality conditions for bilevel programming problems. *Optimization* **33**, 9–27 (1995)
30. Ye, J.J., Zhu, D.L.: A note on optimality conditions for bilevel programming problems. *Optimization* **39**, 361–366 (1997)
31. Zemkoho, A.B.: Estimates of generalized Hessians for optimal value functions in mathematical programming. *Set-valued and variational analysis*, in press, (2021) <https://doi.org/10.1007/s11228-021-00623-yarXiv:1710.05887>
32. Zemkoho, A.B.: Solving ill-posed bilevel programs. *Set-Valued Var. Anal.* **24**, 423–448 (2016)
33. Zemkoho, A.B., Zhou, S.: Theoretical and numerical comparison of the Karush-Kuhn-Tucker and value function reformulations in bilevel optimization. *Comput. Optim. Appl.* **78**(2), 625–674 (2021)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.