



A Steepest Descent Method for Set Optimization Problems with Set-Valued Mappings of Finite Cardinality

Gemayqzel Bouza¹ · Ernest Quintana² · Christiane Tammer³

Received: 18 January 2021 / Accepted: 9 June 2021 / Published online: 25 June 2021
© The Author(s) 2021

Abstract

In this paper, we study a first-order solution method for a particular class of set optimization problems where the solution concept is given by the set approach. We consider the case in which the set-valued objective mapping is identified by a finite number of continuously differentiable selections. The corresponding set optimization problem is then equivalent to find optimistic solutions to vector optimization problems under uncertainty with a finite uncertainty set. We develop optimality conditions for these types of problems and introduce two concepts of critical points. Furthermore, we propose a descent method and provide a convergence result to points satisfying the optimality conditions previously derived. Some numerical examples illustrating the performance of the method are also discussed. This paper is a modified and polished version of Chapter 5 in the dissertation by Quintana (On set optimization with set relations: a scalarization approach to optimality conditions and algorithms, Martin-Luther-Universität Halle-Wittenberg, 2020).

Keywords Set optimization · Robust vector optimization · Descent method · Stationary point

Mathematics Subject Classification 49J53 · 90C29 · 90C46 · 90C47

Communicated by Nguyen Dong Yen.

✉ Ernest Quintana
ernest.quintana-aporicio@tu-ilmenau.de

Gemayqzel Bouza
gema@matcom.uh.cu

Christiane Tammer
christiane.tammer@mathematik.uni-halle.de

- ¹ University of Havana, Havana, Cuba
- ² Technical University of Ilmenau, Ilmenau, Germany
- ³ Martin-Luther University of Halle-Wittenberg, Halle, Germany

1 Introduction

Set optimization is the class of mathematical problems that consists in minimizing set-valued mappings acting between two vector spaces, in which the image space is partially ordered by a given closed, convex and pointed cone. There are two main approaches for defining solution concepts for this type of problems, namely the vector approach and the set approach. In this paper, we deal with the last of these concepts. The main idea of this approach lies on defining a preorder on the power set of the image space and to consider minimal solutions of the set-valued problem accordingly. Research in this area started with the works of Young [46], Nishnianidze [41] and Kuroiwa [35,36], in which the first set relations for defining a preorder were considered. Furthermore, Kuroiwa [34] was the first who considered set optimization problems where the solution concept is given by the set approach. Since then, research in this direction has expanded immensely due to its applications in finance, optimization under uncertainty, game theory and socioeconomics. We refer the reader to [29] for a comprehensive overview of the field.

The research topic that concerns us in this paper is the development of efficient algorithms for the solution of set optimization problems. In this setting, the current approaches in the literature can be roughly clustered into four different groups:

- Derivative-free methods [23,24,30].

In this context, the derived algorithms are descent methods and use a derivative-free strategy [7]. These algorithms are designed to deal with unconstrained problems, and they assume no particular structure of the set-valued objective mapping. The first method of this type was described in [23]. There, the case in which both the epigraphical and hypographical multifunctions of the set-valued objective mapping have convex values was analyzed. This convexity assumption was then relaxed in [30] for the so-called upper set less relation. Finally, in [24], a new method with this strategy was studied. An interesting feature of the algorithm in this reference is that, instead of choosing only one descent direction at every iteration, it considers several of them at the same time. Thus, the method generates a tree with the initial point as the root and the possible solutions as leaves.

- Algorithms of a sorting type [17,18,31,32].

The methods in this class are specifically designed to treat set optimization problems with a finite feasible set. Because of this, they are based on simple comparisons between the images of the set-valued objective mapping. In [31,32], the algorithms are extensions of those by Jahn [21,26] for vector optimization problems. They use a so-called forward and backward reduction procedures that, in practice, avoid making many of these previously mentioned comparisons. Therefore, these methods perform more efficiently than a naive implementation in which every pair of sets must be compared. More recently, in [17,18], an extension of the algorithm by Günther and Popovici [16] for vector problems was studied. The idea now is to, first, find an enumeration of the images of the set-valued mapping whose values by a scalarization using a strongly monotone functional are increasing. In a second step, a forward iteration procedure is performed. Due to the presorting step, these methods enjoy an almost optimal computational complexity, compare [33].

- Algorithms based on scalarization [11,12,19,20,27,44].
The methods in this group follow a scalarization approach and are derived for problems where the set-valued objective mapping has a particular structure that comes from the so-called robust counterpart of a vector optimization problem under uncertainty, see [20]. In [11,19,20], a linear scalarization was employed for solving the set optimization problem. Furthermore, the ϵ -constraint method was extended too in [11,19], for the particular case in which the ordering cone is the nonnegative orthant. Weighted Chebyshev scalarization and some of its variants (augmented, min-ordering) were also studied in [19,27,44].
- Branch and bound [12].
The algorithm in [12] is also designed for uncertain vector optimization problems, but in particular it is assumed that only the decision variable is the source of uncertainty. There, the authors propose a branch and bound method for finding a box covering of the solution set.

The strategy that we consider in this paper is different to the ones previously described and is designed for dealing with unconstrained set optimization problems in which the set-valued objective mapping is given by a finite number of continuously differentiable selections. Our motivation for studying problems with this particular structure is twofold:

- Problems of this type have important applications in optimization under uncertainty.
Indeed, set optimization problems with this structure arise when computing robust solutions to vector optimization problems under uncertainty, if the so-called uncertainty set is finite, see [20]. Furthermore, the solvability of problems with a finite uncertainty set is an important component in the treatment of the general case with an infinite uncertainty set, see the cutting plane strategy in [40] and the reduction results in [3, Proposition 2.1] and [11, Theorem 5.9].
- Current algorithms in the literature pose different theoretical and practical difficulties when solving these types of problems.
Indeed, although derivative-free methods can be directly applied in this setting, they suffer from the same drawbacks as their counterparts in the scalar case. Specifically, because they make no use of first-order information (which we assume is available in our context), we expect them to perform slower in practice than a method that uses these additional properties. Even worse, in the set-valued setting, there is now an increased cost on performing comparisons between sets, which was almost negligible for scalar problems. On the other hand, the algorithms of a sorting type described earlier cannot be used in our setting since they require a finite feasible set. Similarly, the branch and bound strategy is designed for problems that do not fit the particular structure that we consider in this paper, and so it cannot be taken into account. Finally, we can also consider the algorithms based on scalarization in our context. However, the main drawback of these methods is that, in general, they are not able to recover all the solutions of the set optimization problem. In fact, the ϵ -constraint method, which is known to overcome this difficulty in standard multiobjective optimization, will fail in this setting.

Thus, we address in this paper the need of a first-order method that exploits the particular structure of the set-valued objective mapping previously mentioned and does not have the same drawbacks of the other approaches in the literature.

The rest of the paper is structured as follows. We start in Sect. 2 by introducing the main notations, basic concepts and results that will be used throughout the paper. In Sect. 3, we derive optimality conditions for set optimization problems with the aforementioned structure. These optimality conditions constitute the basis of the descent method described in Sect. 4, where the full convergence of the algorithm is also obtained. In Sect. 5, we illustrate the performance of the method on different test instances. We conclude in Sect. 6 by summarizing our results and proposing ideas for further research.

2 Preliminaries

We start this section by introducing the main notations used in the paper. First, the class of all nonempty subsets of \mathbb{R}^m will be denoted by $\mathcal{P}(\mathbb{R}^m)$. Furthermore, for $A \in \mathcal{P}(\mathbb{R}^m)$, we denote by $\text{int } A$, $\text{cl } A$, $\text{bd } A$ and $\text{conv } A$ the interior, closure, boundary and convex hull of the set A , respectively. All the considered vectors are column vectors, and we denote the transpose operator with the symbol \top . On the other hand, $\|\cdot\|$ will stand for either the Euclidean norm of a vector or for the standard spectral norm of a matrix, depending on the context. We also denote the cardinality of a finite set A by $|A|$. Finally, for $k \in \mathbb{N}$, we put $[k] := \{1, \dots, k\}$.

We next consider the most important definitions and properties involved in the results of the paper. Recall that a set $K \in \mathcal{P}(\mathbb{R}^m)$ is said to be a cone if $ty \in K$ for every $y \in K$ and every $t \geq 0$. Moreover, a cone K is called convex if $K + K = K$, pointed if $K \cap (-K) = \{0\}$, and solid if $\text{int } K \neq \emptyset$. An important related concept is that of the dual cone. For a cone K , this is the set

$$K^* := \{v \in \mathbb{R}^m \mid \forall y \in K : v^\top y \geq 0\}.$$

Throughout, we suppose that $K \in \mathcal{P}(\mathbb{R}^m)$ is a cone.

It is well known (see [14]) that when K is convex and pointed, it generates a partial order \preceq on \mathbb{R}^m as follows:

$$y \preceq z : \iff z - y \in K. \quad (1)$$

Furthermore, if K is solid, one can also consider the so-called strict order \prec which is defined by

$$y \prec z : \iff z - y \in \text{int } K. \quad (2)$$

In the following definition, we collect the concepts of minimal and weakly minimal elements of a set with respect to \preceq .

Definition 2.1 Let $A \in \mathcal{P}(\mathbb{R}^m)$ and suppose that K is closed, convex, pointed and solid.

(i) The set of minimal elements of A with respect to K is defined as

$$\text{Min}(A, K) := \{y \in A \mid (y - K) \cap A = \{y\}\}.$$

(ii) The set of weakly minimal elements of A with respect to K is defined as

$$\text{WMin}(A, K) := \{y \in A \mid (y - \text{int } K) \cap A = \emptyset\}.$$

The following proposition will be often used.

Proposition 2.1 ([22, Theorem 6.3 c]) *Let $A \in \mathcal{P}(\mathbb{R}^m)$ be compact, and K be closed, convex and pointed. Then, A satisfies the so-called domination property with respect to K , that is,*

$$A + K = \text{Min}(A, K) + K.$$

The Gerstewitz scalarizing functional will play also an important role in the main results.

Definition 2.2 Let K be closed, convex, pointed and solid. For a given element $e \in \text{int } K$, the Gerstewitz functional associated with e and K is $\psi_e : \mathbb{R}^m \rightarrow \mathbb{R}$ defined as

$$\psi_e(y) := \min\{t \in \mathbb{R} \mid te \in y + K\}. \tag{3}$$

Useful properties of this functional are summarized in the next proposition.

Proposition 2.2 ([29, Section 5.2]) *Let K be closed, convex, pointed and solid, and consider an element $e \in \text{int } K$. Then, the functional ψ_e satisfies the following properties:*

- (i) ψ_e is sublinear and Lipschitz on \mathbb{R}^m .
- (ii) ψ_e is both monotone and strictly monotone with respect to the partial order \preceq , that is,

$$\forall y, z \in \mathbb{R}^m : y \preceq z \implies \psi_e(y) \leq \psi_e(z)$$

and

$$\forall y, z \in \mathbb{R}^m : y \prec z \implies \psi_e(y) < \psi_e(z),$$

respectively.

- (iii) ψ_e satisfies the so-called representability property, that is,

$$-K = \{y \in \mathbb{R}^m \mid \psi_e(y) \leq 0\}, \quad -\text{int } K = \{y \in \mathbb{R}^m \mid \psi_e(y) < 0\}.$$

We next introduce the set relations between the nonempty subsets of \mathbb{R}^m that will be used in the definition of the set optimization problem we consider. We refer the reader to [25,28] and the references therein for other set relations.

Definition 2.3 [37] For the given cone K , the lower set less relation \preceq^ℓ is the binary relation defined on $\mathcal{P}(\mathbb{R}^m)$ as follows:

$$\forall A, B \in \mathcal{P}(\mathbb{R}^m) : A \preceq^\ell B : \iff B \subseteq A + K.$$

Similarly, if K is solid, the strict lower set less relation \prec^ℓ is the binary relation defined on $\mathcal{P}(\mathbb{R}^m)$ by:

$$\forall A, B \in \mathcal{P}(\mathbb{R}^m) : A \prec^\ell B : \iff B \subseteq A + \text{int } K.$$

Remark 2.1 Note that for any two vectors $y, z \in \mathbb{R}^m$ the following equivalences hold:

$$\{y\} \preceq^\ell \{z\} \iff y \preceq z, \quad \{y\} \prec^\ell \{z\} \iff y \prec z.$$

Thus, the restrictions of \preceq^ℓ and \prec^ℓ to the singletons in $\mathcal{P}(\mathbb{R}^m)$ are equivalent to \preceq and \prec , respectively.

We are now ready to present the set optimization problem together with a solution concept based on set relations.

Definition 2.4 Let $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ be a given set-valued mapping taking only nonempty values, and suppose that K is closed, convex, pointed and solid. The set optimization problem with these data is formally represented as

$$\preceq^\ell - \min_{x \in \mathbb{R}^n} F(x), \quad (\mathcal{SP}_\ell)$$

and a solution is understood in the following sense: We say that a point $\bar{x} \in \mathbb{R}^n$ is a local weakly minimal solution of (\mathcal{SP}_ℓ) if there exists a neighborhood U of \bar{x} such that the following holds:

$$\nexists x \in U : F(x) \prec^\ell F(\bar{x}).$$

Moreover, if we can choose $U = \mathbb{R}^n$ above, we simply say that \bar{x} is a weakly minimal solution of (\mathcal{SP}_ℓ) .

Remark 2.2 A related problem to (\mathcal{SP}_ℓ) that is relevant in our paper is the so-called vector optimization problem [22,38]. There, for a vector-valued mapping $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, one considers

$$\preceq - \min_{x \in \mathbb{R}^n} f(x),$$

where a point \bar{x} is said to be a weakly minimal solution if

$$f(\bar{x}) \in \text{WMin}(f[\mathbb{R}^n], K)$$

(corresponding to Definition 2.1). Taking into account Remark 2.1, it is easy to verify that this solution concept coincides with ours for (\mathcal{SP}_ℓ) when the set-valued mapping F is given by $F(x) := \{f(x)\}$ for every $x \in \mathbb{R}^n$.

We conclude the section by establishing the main assumption employed in the rest of the paper for the treatment of (\mathcal{SP}_ℓ) :

Assumption 1 Suppose that $K \in \mathcal{P}(\mathbb{R}^m)$ is a closed, convex, pointed and solid cone and that $e \in \text{int } K$ is fixed. Furthermore, consider a reference point $\bar{x} \in \mathbb{R}^n$, given vector-valued functions $f^1, f^2, \dots, f^p : \mathbb{R}^n \rightarrow \mathbb{R}^m$ that are continuously differentiable, and assume that the set-valued mapping F in (\mathcal{SP}_ℓ) is defined by

$$F(x) := \left\{ f^1(x), f^2(x), \dots, f^p(x) \right\}.$$

3 Optimality Conditions

In this section, we study optimality conditions for weakly minimal solutions of (\mathcal{SP}_ℓ) under Assumption 1. These conditions are the foundation on which the proposed algorithm is built. In particular, because of the resemblance of our method with standard gradient descent in the scalar case, we are interested in Fermat rules for set optimization problems. Recently, results of this type were derived in [5], see also [2]. There, the optimality conditions involve the computation of the limiting normal cone [39] of the set-valued mapping F at different points in its graph. However, this is a difficult task in our case because the graph of F is the union of the graphs of the vector-valued functions f^i , and to the best of our knowledge, there is no exact formula for finding the normal cone to the union of sets (at a given point) in terms of the initial data. Thus, instead of considering the results from [5], we exploit the particular structure of F and the differentiability of the functionals f^i to deduce new necessary conditions.

We start by defining some index-related set-valued mappings that will be of importance. They make use of the concepts introduced in Definition 2.1.

Definition 3.1 The following set-valued mappings are defined:

- (i) The active index of minimal elements associated with F is $I : \mathbb{R}^n \rightrightarrows [p]$ given by

$$I(x) := \{i \in [p] \mid f^i(x) \in \text{Min}(F(x), K)\}.$$

- (ii) The active index of weakly minimal elements associated with F is $I_W : \mathbb{R}^n \rightrightarrows [p]$ defined as

$$I_W(x) := \{i \in [p] \mid f^i(x) \in \text{WMin}(F(x), K)\}.$$

- (iii) For a vector $v \in \mathbb{R}^m$, we define $I_v : \mathbb{R}^n \rightrightarrows [p]$ as

$$I_v(x) := \{i \in I(x) \mid f^i(x) = v\}.$$

It follows from the definition that $I_v(x) = \emptyset$ whenever $v \notin \text{Min}(F(x), K)$ and that

$$\forall x \in \mathbb{R}^n : I(x) = \bigcup_{v \in \text{Min}(F(x), K)} I_v(x). \quad (4)$$

Definition 3.2 The map $\omega : \mathbb{R}^n \rightarrow \mathbb{R}$ is defined as the cardinality of the set of minimal elements of F , that is,

$$\omega(x) := |\text{Min}(F(x), K)|.$$

Furthermore, we set $\bar{\omega} := \omega(\bar{x})$.

From now on, we consider that, for any point $x \in \mathbb{R}^n$, an enumeration $\{v_1^x, \dots, v_{\omega(x)}^x\}$ of the set $\text{Min}(F(x), K)$ has been chosen in advance.

Definition 3.3 Let $x \in \mathbb{R}^n$, and consider the enumeration $\{v_1^x, \dots, v_{\omega(x)}^x\}$ of the set $\text{Min}(F(x), K)$. The partition set of x is defined as

$$P_x := \prod_{j=1}^{\omega(x)} I_{v_j^x}(x),$$

where $I_{v_j^x}(x)$ is given in Definition 3.1 (iii) for $j \in [\omega(x)]$.

The optimality conditions for (SP_ℓ) we will present are based on the following idea: from the particular structure of F , we will construct a family of vector optimization problems that, together, locally represent (SP_ℓ) (in a sense to be specified) around the point which must be checked for optimality. Then, (standard) optimality conditions are applied to the family of vector optimization problems. The following lemma is the key step in that direction.

Lemma 3.1 Let $\tilde{K} \in \mathcal{P}(\mathbb{R}^{m\bar{\omega}})$ be the cone defined as

$$\tilde{K} := \prod_{j=1}^{\bar{\omega}} K, \quad (5)$$

and let us denote by $\leq_{\tilde{K}}$ and $<_{\tilde{K}}$ the partial order and the strict order in $\mathbb{R}^{m\bar{\omega}}$ induced by \tilde{K} , respectively (see (2)). Furthermore, consider the partition set $P_{\bar{x}}$ associated with \bar{x} and define, for every $a = (a_1, \dots, a_{\bar{\omega}}) \in P_{\bar{x}}$, the function $\tilde{f}^a : \mathbb{R}^n \rightarrow \prod_{j=1}^{\bar{\omega}} \mathbb{R}^m$ as

$$\tilde{f}^a(x) := \begin{pmatrix} f^{a_1}(x) \\ \vdots \\ f^{a_{\bar{\omega}}}(x) \end{pmatrix}. \quad (6)$$

Then, \bar{x} is a local weakly minimal solution of (\mathcal{SP}_ℓ) if and only if, for every $a \in P_{\bar{x}}$, \bar{x} is a local weakly minimal solution of the vector optimization problem

$$\leq_{\tilde{K}} - \min_{x \in \mathbb{R}^n} \tilde{f}^a(x). \tag{VP}_a$$

Proof We argue by contradiction in both cases. First, assume that \bar{x} is a local weakly minimal solution of (\mathcal{SP}_ℓ) and that, for some $a \in P_{\bar{x}}$, \bar{x} is not a local weakly minimal solution of (\mathcal{VP}_a) . Then, we could find a sequence $\{x_k\}_{k \geq 1} \subseteq \mathbb{R}^n$ such that $x_k \rightarrow \bar{x}$ and

$$\forall k \in \mathbb{N} : \tilde{f}^a(x_k) <_{\tilde{K}} \tilde{f}^a(\bar{x}). \tag{7}$$

Hence, we deduce that

$$\begin{aligned} \forall k \in \mathbb{N} : F(\bar{x}) &\stackrel{\text{(Proposition 2.1)}}{\subseteq} \{f^{a_1}(\bar{x}), \dots, f^{a_{\tilde{\omega}}}(\bar{x})\} + K \\ &\stackrel{(7)}{\subseteq} \{f^{a_1}(x_k), \dots, f^{a_{\tilde{\omega}}}(x_k)\} + \text{int } K + K \\ &\subseteq F(x_k) + \text{int } K. \end{aligned}$$

Since this is equivalent to $F(x_k) <^\ell F(\bar{x})$ for every $k \in \mathbb{N}$ and $x_k \rightarrow \bar{x}$, it contradicts the weak minimality of \bar{x} for (\mathcal{SP}_ℓ) .

Next, suppose that \bar{x} is a local weakly minimal solution of (\mathcal{VP}_a) for every $a \in P_{\bar{x}}$, but not a local weakly minimal solution of (\mathcal{SP}_ℓ) . Then, we could find a sequence $\{x_k\}_{k \geq 1} \subseteq \mathbb{R}^n$ such that $x_k \rightarrow \bar{x}$ and $F(x_k) <^\ell F(\bar{x})$ for every $k \in \mathbb{N}$. Consider the enumeration $\{v_{\tilde{\omega}}^{\bar{x}}, \dots, v_1^{\bar{x}}\}$ of the set $\text{Min}(F(\bar{x}), K)$. Then,

$$\forall j \in [\tilde{\omega}], k \in \mathbb{N}, \exists i_{(j,k)} \in [p] : f^{i_{(j,k)}}(x_k) < v_j^{\bar{x}}. \tag{8}$$

Since the indexes $i_{(j,k)}$ are being chosen on the finite set $[p]$, we can assume without loss of generality that $i_{(j,k)}$ is independent of k , that is, $i_{(j,k)} = \bar{i}_j$ for every $k \in \mathbb{N}$ and some $\bar{i}_j \in [p]$. Hence, taking the limit in (8) when $k \rightarrow +\infty$, we get

$$\forall j \in [\tilde{\omega}] : f^{\bar{i}_j}(\bar{x}) \leq v_j^{\bar{x}}. \tag{9}$$

Because $v_j^{\bar{x}} \in \text{Min}(F(\bar{x}), K)$, it follows from (9) that $f^{\bar{i}_j}(\bar{x}) = v_j^{\bar{x}}$ and that $\bar{i}_j \in I(\bar{x})$ for every $j \in [\tilde{\omega}]$. Consider now the tuple $\bar{a} := (\bar{i}_1, \dots, \bar{i}_{\tilde{\omega}})$. Then, it can be verified that $\bar{a} \in P_{\bar{x}}$. Moreover, from (8) we deduce that $\tilde{f}^{\bar{a}}(x_k) <_{\tilde{K}} \tilde{f}^{\bar{a}}(\bar{x})$ for every $k \in \mathbb{N}$. Since $x_k \rightarrow \bar{x}$, this contradicts the weak minimality of \bar{x} for (\mathcal{VP}_a) when $a = \bar{a}$. □

We now establish the necessary optimality condition for (\mathcal{SP}_ℓ) that will be used in our descent method.

Theorem 3.1 Suppose that \bar{x} is a local weakly minimal solution of (\mathcal{SP}_ℓ) . Then,

$$\forall a \in P_{\bar{x}}, \exists \mu_1, \mu_2, \dots, \mu_{\bar{w}} \in K^* : \sum_{j=1}^{\bar{w}} \nabla f^{a_j}(\bar{x}) \mu_j = 0, (\mu_1, \dots, \mu_{\bar{w}}) \neq 0 \quad (10)$$

Conversely, assume that f^i is K -convex for each $i \in I(\bar{x})$, that is,

$$\forall i \in I(\bar{x}), x_1, x_2 \in \mathbb{R}^n, t \in [0, 1] : f^i(tx_1 + (1-t)x_2) \leq tf^i(x_1) + (1-t)f^i(x_2).$$

Then, condition (10) is also sufficient for the local weak minimality of \bar{x} .

Proof By Lemma 3.1, we get that \bar{x} is a local weakly minimal solution of (\mathcal{VP}_a) for every $a \in P_{\bar{x}}$. Applying now [38, Theorem 4.1] for every $a \in P_{\bar{x}}$, we get

$$\forall a \in P_{\bar{x}}, \exists \mu \in \tilde{K}^* \setminus \{0\} : \nabla \tilde{f}^a(\bar{x}) \mu = 0. \quad (11)$$

Since $\tilde{K}^* = \prod_{j=1}^{\bar{w}} K^*$, it is easy to verify that (11) is equivalent to the first part of the statement.

In order to see the sufficiency under convexity, assume that \bar{x} satisfies (10). Note that for any $a \in P_{\bar{x}}$, the function \tilde{f}^a is \tilde{K} -convex, provided that each f^i is K -convex for every $i \in I(\bar{x})$. Then, in this case, it is well known that (11) is equivalent to \bar{x} being a local weakly minimal solution of (\mathcal{VP}_a) for every $a \in P_{\bar{x}}$, see [15]. Applying now Lemma 3.1, we obtain that \bar{x} is a local weakly minimal solution of (\mathcal{SP}_ℓ) . \square

Based on Theorem 3.1, we define the following concepts of stationarity for (\mathcal{SP}_ℓ) .

Definition 3.4 We say that \bar{x} is a stationary point of (\mathcal{SP}_ℓ) if there exists a nonempty set $Q \subseteq P_{\bar{x}}$ such that the following assertion holds:

$$\forall a \in Q, \exists \mu_1, \mu_2, \dots, \mu_{\bar{w}} \in K^* : \sum_{j=1}^{\bar{w}} \nabla f^{a_j}(\bar{x}) \mu_j = 0, (\mu_1, \dots, \mu_{\bar{w}}) \neq 0. \quad (12)$$

In that case, we also say that \bar{x} is stationary with respect to Q . If, in addition, we can choose $Q = P_{\bar{x}}$ in (12), we simply call \bar{x} a strongly stationary point.

Remark 3.1 It follows from Definition 3.4 that a point \bar{x} is stationary for (\mathcal{SP}_ℓ) if and only if

$$\exists a \in P_{\bar{x}}, \mu_1, \mu_2, \dots, \mu_{\bar{w}} \in K^* : \sum_{j=1}^{\bar{w}} \nabla f^{a_j}(\bar{x}) \mu_j = 0, (\mu_1, \dots, \mu_{\bar{w}}) \neq 0.$$

Furthermore, a strongly stationary point of (SP_ℓ) is also stationary with respect to Q for every nonempty set $Q \subseteq P_{\bar{x}}$. Moreover, from Theorem 3.1, it is clear that stationarity is also a necessary optimality condition for (SP_ℓ) .

In the following example, we illustrate a comparison of our optimality conditions with previous ones in the literature for standard optimization problems.

Example 3.1 Suppose that in Assumption 1 we have $m = 1, K = \mathbb{R}_+$. Furthermore, consider the functional $f : \mathbb{R}^n \rightarrow \mathbb{R}$ defined as

$$f(x) := \min_{i \in [p]} f^i(x)$$

and problem (SP_ℓ) associated with these data. Hence, in this case,

$$P_{\bar{x}} = I(\bar{x}) = \{i \in [p] \mid f^i(\bar{x}) = f(\bar{x})\}.$$

It is then easy to verify that the following statements hold:

(i) \bar{x} is strongly stationary for (SP_ℓ) if and only if

$$\forall i \in I(\bar{x}) : \nabla f^i(\bar{x}) = 0.$$

(ii) \bar{x} is stationary for (SP_ℓ) if and only if

$$\exists i \in I(\bar{x}) : \nabla f^i(\bar{x}) = 0.$$

On the other hand, it is straightforward to verify that \bar{x} is a weakly minimal solution of (SP_ℓ) if and only if \bar{x} is a solution of the problem

$$\min_{x \in \mathbb{R}^n} f(x). \tag{P}$$

Moreover, if we denote by $\widehat{\partial}f(\bar{x})$ and $\partial f(\bar{x})$ the Fréchet and Mordukhovich subdifferential of f at point \bar{x} , respectively (see [39]), it follows from [39, Proposition 1.114] that the inclusions

$$0 \in \widehat{\partial}f(\bar{x}) \tag{13}$$

and

$$0 \in \partial f(\bar{x}) \tag{14}$$

are necessary for \bar{x} being a solution of (P). A point \bar{x} satisfying (13) and (14) is said to be Fréchet and Mordukhovich stationary for (P), respectively. Furthermore, from [10, Proposition 5] and [39, Proposition 1.113], we have

$$\widehat{\partial}f(\bar{x}) = \bigcap_{i \in I(\bar{x})} \{\nabla f^i(\bar{x})\} \tag{15}$$

and

$$\partial f(\bar{x}) \subseteq \bigcup_{i \in I(\bar{x})} \{\nabla f^i(\bar{x})\}, \tag{16}$$

respectively. Thus, from (13), (15) and (i), we deduce that

(iii) \bar{x} is strongly stationary for (\mathcal{SP}_ℓ) if and only if \bar{x} is Fréchet stationary for (\mathcal{P}) .

Similarly, from (14), (16) and (ii), we find that:

(iii) If \bar{x} is Mordukhovich stationary for (\mathcal{P}) , then \bar{x} is stationary for (\mathcal{SP}_ℓ) .

We close the section with the following proposition that presents an alternative characterization of stationary points.

Proposition 3.1 *Let $Q \subseteq P_{\bar{x}}$ be given. Then, \bar{x} is stationary for (\mathcal{SP}_ℓ) with respect to Q if and only if*

$$\forall a \in Q, u \in \mathbb{R}^n, \exists j \in [\bar{\omega}] : \nabla f^{aj}(\bar{x})^\top u \notin -\text{int } K. \tag{17}$$

Proof Suppose first that \bar{x} is stationary with respect to Q . Fix now $a \in Q, u \in \mathbb{R}^n$, and consider the vectors $\mu_1, \mu_2, \dots, \mu_{\bar{\omega}} \in K^*$ that satisfy (12). We argue by contradiction. Assume that

$$\forall j \in [\bar{\omega}] : \nabla f^{aj}(\bar{x})^\top u \in -\text{int } K. \tag{18}$$

From (18) and the fact that $(\mu_1, \dots, \mu_{\bar{\omega}}) \in \left(\prod_{j=1}^{\bar{\omega}} K^*\right) \setminus \{0\}$, we deduce that

$$\left(\mu_1^\top \left(\nabla f^{a1}(\bar{x})^\top u\right), \dots, \mu_{\bar{\omega}}^\top \left(\nabla f^{a\bar{\omega}}(\bar{x})^\top u\right)\right) \in -\mathbb{R}_+^{\bar{\omega}} \setminus \{0\}. \tag{19}$$

Hence, we get

$$0 \stackrel{(12)}{=} \left(\sum_{j=1}^{\bar{\omega}} \nabla f^{aj}(\bar{x}) \mu_j\right)^\top u = \sum_{j=1}^{\bar{\omega}} \mu_j^\top \left(\nabla f^{aj}(\bar{x})^\top u\right) \stackrel{(19)}{<} 0,$$

a contradiction.

Suppose now that (17) holds, and fix $a \in Q$. Consider the functional \tilde{f}^a and the cone \tilde{K} from Lemma 3.1, together with the set

$$A := \left\{ \nabla \tilde{f}^a(\bar{x})^\top u \mid u \in \mathbb{R}^n \right\}.$$

Then, we deduce from (17) that

$$A \cap \text{int } \tilde{K} = \emptyset.$$

Applying now Eidelheit’s separation theorem for convex sets [22, Theorem 3.16], we obtain some $(\mu_1, \dots, \mu_{\bar{\omega}}) \in \left(\prod_{j=1}^{\bar{\omega}} \mathbb{R}^m\right) \setminus \{0\}$ such that

$$\forall u \in \mathbb{R}^n, v_1, \dots, v_{\bar{\omega}} \in K : \left(\sum_{j=1}^{\bar{\omega}} \nabla f^{a_j}(\bar{x})\mu_j\right)^\top u \leq \sum_{j=1}^{\bar{\omega}} \mu_j^\top v_j. \tag{20}$$

By fixing $\bar{j} \in [\bar{\omega}]$ and substituting $u = 0, v_j = 0$ for $j \neq \bar{j}$ in (20), we obtain

$$\forall v_{\bar{j}} \in K : \mu_{\bar{j}}^\top v_{\bar{j}} \geq 0.$$

Hence, $\mu_{\bar{j}} \in K^*$. Since \bar{j} was chosen arbitrarily, we get that $(\mu_1, \dots, \mu_{\bar{\omega}}) \in \left(\prod_{j=1}^{\bar{\omega}} K^*\right) \setminus \{0\}$. Define now

$$\bar{u} := \sum_{j=1}^{\bar{\omega}} \nabla f^{a_j}(\bar{x})\mu_j.$$

To finish the proof, we need to show that $\bar{u} = 0$. In order to see this, substitute $u = \bar{u}$ and $v_j = 0$ for each $j \in [\bar{\omega}]$ in (20). Then, we obtain

$$\left\| \sum_{j=1}^{\bar{\omega}} \nabla f^{a_j}(\bar{x})\mu_j \right\|^2 \leq 0.$$

Hence, it can only be $\bar{u} = 0$, and statement (12) is true. □

4 Descent Method and Its Convergence Analysis

Now, we present the solution approach. It is clearly based on the result shown in Lemma 3.1. At every iteration, an element a in the partition set of the current iterate point is selected, and then, a descent direction for (\mathcal{VP}_a) will be found using ideas from [6,15]. However, one must be careful with the selection process of the element a in order to guarantee convergence. Thus, we propose a specific way to achieve this. After the descent direction is determined, we follow a classical backtracking procedure of Armijo type to determine a suitable step size, and we update the iterate in the desired direction. Algorithm 1 formally describes our method.

Remark 4.1 Algorithm 1 extends the approaches proposed in [6,15] for vector optimization optimization problems to case (\mathcal{SP}_ℓ) . The main difference is that, in Step 2, the authors use the well known Hiriart-Urruty functional and the support of a so-called generator of the dual cone instead of ψ_e , respectively. However, in our framework,

Algorithm 1 Descent Method in Set Optimization

Step 0. Choose $x_0 \in \mathbb{R}^n$, $\beta, \nu \in (0, 1)$, and set $k := 0$.

Step 1. Compute

$$M_k := \text{Min}(F(x_k), K), \quad P_k := P_{x_k}, \quad \omega_k := |\text{Min}(F(x_k), K)|.$$

Step 2. Find

$$(a_k, u_k) \in \underset{(a,u) \in P_k \times \mathbb{R}^n}{\text{argmin}} \max_{j \in [\omega_k]} \{ \psi_e(\nabla f^{a,j}(x_k)^\top d) \} + \frac{1}{2} \|u\|^2.$$

Step 3. If $u_k = 0$, Stop. Otherwise, go to Step 4.

Step 4. Compute

$$t_k := \max_{q \in \mathbb{N} \cup \{0\}} \left\{ \nu^q \mid \forall j \in [\omega_k] : f^{a_k,j}(x_k + \nu^q u_k) \leq f^{a_k,j}(x_k) + \beta \nu^q \nabla f^{a_k,j}(x_k)^\top u_k \right\}.$$

Step 5. Set $x_{k+1} := x_k + t_k u_k$, $k := k + 1$ and go to Step 2.

the functional ψ_e is a particular case of those employed in the other methods, see [4, Corollary 2]. Thus, the equivalence in the case of vector optimization problems of the three algorithms is obtained.

Now, we start the convergence analysis of Algorithm 1. Our first lemma describes local properties of the active indexes.

Lemma 4.1 *Under our assumptions, there exists a neighborhood U of \bar{x} such that the following properties are satisfied (some of them under additional conditions to be established below) for every $x \in U$:*

- (i) $I_W(x) \subseteq I_W(\bar{x})$,
- (ii) $I(x) \subseteq I(\bar{x})$, provided that $\text{Min}(F(\bar{x}), K) = \text{WMin}(F(\bar{x}), K)$,
- (iii) $\forall v \in \text{Min}(F(\bar{x}), K) : \text{Min}(\{f^i(x)\}_{i \in I_v(\bar{x})}, K) \subseteq \text{Min}(F(x), K)$,
- (iv) For every $v_1, v_2 \in \text{Min}(F(\bar{x}), K)$ with $v_1 \neq v_2$:

$$\text{Min}(\{f^i(x)\}_{i \in I_{v_1}(\bar{x})}, K) \cap \text{Min}(\{f^i(x)\}_{i \in I_{v_2}(\bar{x})}, K) = \emptyset,$$

- (v) $\omega(x) \geq \omega(\bar{x})$.

Proof It suffices to show the existence of the neighborhood U for each item independently, as we could later take the intersection of them to satisfy all the properties.

(i) Assume that this is not satisfied in any neighborhood U of \bar{x} . Then, we could find a sequence $\{x_k\}_{k \geq 1} \subseteq \mathbb{R}^n$ such that $x_k \rightarrow \bar{x}$ and

$$\forall k \in \mathbb{N} : I_W(x_k) \setminus I_W(\bar{x}) \neq \emptyset. \tag{21}$$

Because of the finite cardinality of all possible differences in (21), we can assume without loss of generality that there exists a common $\bar{i} \in [p]$ such that

$$\forall k \in \mathbb{N} : \bar{i} \in I_W(x_k) \setminus I_W(\bar{x}). \tag{22}$$

In particular, (22) implies that $\bar{i} \in I_W(x_k)$. Hence, we get

$$\forall k \in \mathbb{N}, i \in [p] : f^i(x_k) - f^{\bar{i}}(x_k) \in -(\mathbb{R}^m \setminus \text{int } K).$$

Since $\mathbb{R}^m \setminus \text{int } K$ is closed, taking the limit when $k \rightarrow +\infty$ we obtain

$$\forall i \in [p] : f^i(\bar{x}) - f^{\bar{i}}(\bar{x}) \in -(\mathbb{R}^m \setminus \text{int } K).$$

Hence, we deduce that $f^{\bar{i}}(\bar{x}) \in \text{WMin}(F(\bar{x}), K)$ and $\bar{i} \in I_W(\bar{x})$, a contradiction to (21).

(ii) Consider the same neighborhood U on which statement (i) holds. Note that, under the given assumption, we have $I_W(\bar{x}) = I(\bar{x})$. This, together with statement (i), implies:

$$\forall x \in U : I(x) \subseteq I_W(x) \subseteq I_W(\bar{x}) = I(\bar{x}).$$

(iii) For this statement, it is also sufficient to show that the neighborhood U can be chosen for any point in the set $\text{Min}(F(\bar{x}), K)$. Hence, fix $v \in \text{Min}(F(\bar{x}), K)$ and assume that there is no neighborhood U of \bar{x} on which the statement is satisfied. Then, we could find sequences $\{x_k\}_{k \geq 1} \subseteq \mathbb{R}^n$ and $\{i_k\}_{k \geq 1} \subseteq I_v(\bar{x})$ such that $x_k \rightarrow \bar{x}$ and

$$\forall k \in \mathbb{N} : f^{i_k}(x_k) \in \text{Min}(\{f^i(x_k)\}_{i \in I_v(\bar{x})}, K) \setminus \text{Min}(F(x_k), K). \tag{23}$$

Since $I_v(\bar{x})$ is finite, we deduce that there are only a finite number of different elements in the sequence $\{i_k\}$. Hence, we can assume without loss of generality that there exists $\bar{i} \in I_v(\bar{x})$ such that $i_k = \bar{i}$ for every $k \in \mathbb{N}$. Then, (23) is equivalent to

$$\forall k \in \mathbb{N} : f^{\bar{i}}(x_k) \in \text{Min}(\{f^i(x_k)\}_{i \in I_v(\bar{x})}, K) \setminus \text{Min}(F(x_k), K). \tag{24}$$

From (24), we get in particular that $f^{\bar{i}}(x_k) \notin \text{Min}(F(x_k), K)$ for every $k \in \mathbb{N}$. This, together with the domination property in Proposition 2.1 and the fact that the sets $I(x_k)$ are contained in the finite set $[p]$, allows us to obtain without loss of generality the existence of $\tilde{i} \in I(\bar{x})$ such that

$$\forall k \in \mathbb{N} : f^{\tilde{i}}(x_k) \leq f^{\bar{i}}(x_k), f^{\tilde{i}}(x_k) \neq f^{\bar{i}}(x_k). \tag{25}$$

Now, taking the limit in (25) when $k \rightarrow +\infty$, we obtain $f^{\tilde{i}}(\bar{x}) \leq f^{\bar{i}}(\bar{x}) = v$. Since v is a minimal element of $F(\bar{x})$, it can only be $f^{\tilde{i}}(\bar{x}) = v$ and, hence, $\tilde{i} \in I_v(\bar{x})$. From this, the first inequality in (25), and the fact that $f^{\tilde{i}}(x_k) \in \text{Min}(\{f^i(x_k)\}_{i \in I_v(\bar{x})}, K)$ for every $k \in \mathbb{N}$, we get that $f^{\tilde{i}}(x_k) = f^{\bar{i}}(x_k)$ for all $k \in \mathbb{N}$. This contradicts the second part of (25), and hence, our statement is true.

(iv) It follows directly from the continuity of the functionals $f^i, i \in [p]$.

(v) The statement is an immediate consequence of (iii) and (iv). □

For the main convergence theorem of our method, we will need the notion of regularity of a point for a set-valued mapping.

Definition 4.1 We say that \bar{x} is a regular point of F if the following conditions are satisfied:

- (i) $\text{Min}(F(\bar{x}), K) = \text{WMin}(F(\bar{x}), K)$,
- (ii) the cardinality functional ω introduced in Definition 3.2 is constant in a neighborhood of \bar{x} .

Remark 4.2 Since we will analyze the stationarity of the regular limit points of the sequence generated by Algorithm 1, the following points must be addressed:

- Notice that, by definition, the regularity property of a point is independent of our optimality concept. Thus, by only knowing that a point is regular, we cannot infer anything about whether it is optimal or not.
- The concept of regularity seems to be linked to the complexity of comparing sets in a high-dimensional space. For example, in case $m = 1$ or $p = 1$, every point in \mathbb{R}^n is regular for the set-valued mapping F . Indeed, in these cases, we have $\omega(x) = 1$ and

$$\text{Min}(F(x), K) = \text{WMin}(F(x), K) = \begin{cases} \left\{ \min_{i \in [p]} f^i(x) \right\} & \text{if } m = 1, \\ \{f^1(x)\} & \text{if } p = 1 \end{cases}$$

for all $x \in \mathbb{R}^n$.

A natural question is whether regularity is a strong assumption to impose on a point. In that sense, given the finite structure of the sets $F(x)$, condition (i) in Definition 4.1 seems to be very reasonable. In fact, we would expect that, for most practical cases, this condition is fulfilled at almost every point. For condition (ii), a formalized statement is derived in Proposition 4.1.

Proposition 4.1 *The set*

$$S := \{x \in \mathbb{R}^n \mid \omega \text{ is locally constant at } x\}$$

is open and dense in \mathbb{R}^n .

Proof (i) The openness is trivial. Suppose now that S is not dense in \mathbb{R}^n . Then, $\mathbb{R}^n \setminus (\text{cl } S)$ is nonempty and open. Furthermore, since ω is bounded above, the real number

$$p_0 := \max_{x \in \mathbb{R}^n \setminus (\text{cl } S)} \omega(x)$$

is well defined. Consider the set

$$A := \left\{ x \in \mathbb{R}^n \mid \omega(x) \leq p_0 - \frac{1}{2} \right\}.$$

From Lemma 4.1 (v), it follows that ω is lower semicontinuous. Hence, A is closed as it is the sublevel set of a lower semicontinuous functional, see [43, Lemma 1.7.2]. Consider now the set

$$U := (\mathbb{R}^n \setminus \text{cl } S) \cap (\mathbb{R}^n \setminus A).$$

Then, U is a nonempty open subset of $\mathbb{R}^n \setminus \text{cl } S$. This, together with the definition of A , gives us $\omega(x) = p_0$ for every $x \in U$. However, this contradicts the fact that ω is not locally constant at any point of $\mathbb{R}^n \setminus \text{cl } S$. Hence, S is dense in \mathbb{R}^n . \square

An essential property of regular points of a set-valued mapping is described in the next lemma.

Lemma 4.2 *Suppose that \bar{x} is a regular point of F . Then, there exists a neighborhood U of \bar{x} such that the following properties hold for every $x \in U$:*

- (i) $\omega(x) = \bar{\omega}$,
- (ii) *there is an enumeration $\{w_1^x, \dots, w_{\bar{\omega}}^x\}$ of $\text{Min}(F(x), K)$ such that*

$$\forall j \in [\bar{\omega}] : I_{w_j^x}(x) \subseteq I_{v_j^{\bar{x}}}(\bar{x}).$$

In particular, without loss of generality, we have $P_x \subseteq P_{\bar{x}}$ for every $x \in U$.

Proof Let U be the neighborhood of \bar{x} from Lemma 4.1. Since \bar{x} is a regular point of F , we assume without loss of generality that ω is constant on U . Hence, property (i) is fulfilled. Fix now $x \in U$ and consider the enumeration $\{v_1^{\bar{x}}, \dots, v_{\bar{\omega}}^{\bar{x}}\}$ of $\text{Min}(F(\bar{x}), K)$. Then, from properties (iii) and (iv) in Lemma 4.1 and the fact that $\omega(x) = \bar{\omega}$, we deduce that

$$\forall j \in [\bar{\omega}] : \left| \text{Min} \left(\{f^i(x)\}_{i \in I_{v_j^{\bar{x}}}(\bar{x})}, K \right) \right| = 1. \tag{26}$$

Next, for $j \in [\bar{\omega}]$, we define w_j^x as the unique element of the set

$$\text{Min} \left(\{f^i(x)\}_{i \in I_{v_j^{\bar{x}}}(\bar{x})}, K \right).$$

Then, from (26), property (iii) in Lemma 4.1 and the fact that ω is constant on U , we obtain that $\{w_1^x, \dots, w_{\bar{\omega}}^x\}$ is an enumeration of the set $\text{Min}(F(x), K)$.

It remains to show now that this enumeration satisfies (ii). In order to see this, fix $j \in [\bar{\omega}]$ and $\bar{i} \in I_{w_j^x}(x)$. Then, from the regularity of \bar{x} and property (ii) in Lemma 4.1, we get that $I(x) \subseteq I(\bar{x})$. In particular, this implies $\bar{i} \in I(\bar{x})$. From this and (4), we have the existence of $j' \in [\bar{\omega}]$ such that $\bar{i} \in I_{v_{j'}^{\bar{x}}}(\bar{x})$. Hence, we deduce that

$$w_j^x = f^{\bar{i}}(x) \in \left\{ f^i(x) \right\}_{i \in I_{v_{j'}^{\bar{x}}}(\bar{x})}. \tag{27}$$

Then, from (26), (27) and the definition of $w_{j'}^x$, we find that $w_{j'}^x \preceq w_j^x$. Moreover, because $w_{j'}^x, w_j^x \in \text{Min}(F(x), K)$, it can only be $w_{j'}^x = w_j^x$. Thus, it follows that $j = j'$, since $\{w_1^x, \dots, w_\omega^x\}$ is an enumeration of the set $\text{Min}(F(x), K)$. This shows that $\bar{i} \in I_{w_j^x}(\bar{x})$, as desired. \square

For the rest of the analysis, we need to introduce the parametric family of functionals $\{\varphi_x\}_{x \in \mathbb{R}^n}$, whose elements $\varphi_x : P_x \times \mathbb{R}^n \rightarrow \mathbb{R}$ are defined as follows:

$$\forall a \in P_x, u \in \mathbb{R}^n : \varphi_x(a, u) := \max_{j \in [\omega(x)]} \{ \psi_e(\nabla f^{a_j}(x)^\top u) \} + \frac{1}{2} \|u\|^2, \tag{28}$$

where the functional ψ_e is given by (3). It is easy to see that, for every $x \in \mathbb{R}^n$ and $a \in P_x$, the functional $\varphi_x(a, \cdot)$ is strongly convex in \mathbb{R}^n , that is, there exists a constant $\alpha > 0$ such that the inequality

$$\varphi_x(a, tu + (1 - t)u') + \alpha t(1 - t)\|u - u'\|^2 \leq t\varphi_x(a, u) + (1 - t)\varphi_x(a, u')$$

is satisfied for every $u, u' \in \mathbb{R}^n$ and $t \in [0, 1]$. According to [13, Lemma 3.9], the functional $\varphi_x(a, \cdot)$ attains its minimum over \mathbb{R}^n , and this minimum is unique. In particular, we can check that

$$\forall x \in \mathbb{R}^n, a \in P_x : \min_{u \in \mathbb{R}^n} \varphi_x(a, u) \leq 0 \tag{29}$$

and that, if $u_a \in \mathbb{R}^n$ is such that $\varphi_x(a, u_a) = \min_{u \in \mathbb{R}^n} \varphi_x(a, u)$, then

$$\varphi_x(a, u_a) = 0 \iff u_a = 0. \tag{30}$$

Taking into account that P_x is finite, we also obtain that φ_x attains its minimum over the set $P_x \times \mathbb{R}^n$. Hence, we can consider the functional $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$ given by

$$\phi(x) := \min_{(a,u) \in P_x \times \mathbb{R}^n} \varphi_x(a, u). \tag{31}$$

Then, because of (29), it can be verified that

$$\forall x \in \mathbb{R}^n : \phi(x) \leq 0. \tag{32}$$

Furthermore, if $(a, u) \in P_x \times \mathbb{R}^n$ is such that $\phi(x) = \varphi_x(a, u)$, it follows from (30) (see also [15]) that

$$\phi(x) = 0 \iff u = 0. \tag{33}$$

In the following two propositions, we show that Algorithm 1 is well defined. We start by proving that, if Algorithm 1 stops in Step 3, a stationary point was found.

Proposition 4.2 Consider the functionals $\varphi_{\bar{x}}$ and ϕ given in (28) and (31), respectively. Furthermore, let $(\bar{a}, \bar{u}) \in P_{\bar{x}} \times \mathbb{R}^n$ be such that $\phi(\bar{x}) = \varphi_{\bar{x}}(\bar{a}, \bar{u})$. Then, the following statements are equivalent:

- (i) \bar{x} is a strongly stationary point of (SP_{ℓ}) ,
- (ii) $\phi(\bar{x}) = 0$,
- (iii) $\bar{u} = 0$.

Proof The result will be a consequence of [6, Proposition 2.2] where, using the Hiriart-Urruty functional, a similar statement is proved for vector optimization problems.

Consider the cone \tilde{K} given by (5), the vector $\tilde{e} := \begin{pmatrix} e \\ \vdots \\ e \end{pmatrix} \in \text{int } \tilde{K}$, and the scalarizing functional $\psi_{\tilde{e}}$ associated with \tilde{e} and \tilde{K} , see Definition 2.2. Then, for any $v_1, \dots, v_{\bar{\omega}} \in \mathbb{R}^m$ and $v := \begin{pmatrix} v_1 \\ \vdots \\ v_{\bar{\omega}} \end{pmatrix}$, we get

$$\begin{aligned} \psi_{\tilde{e}}(v) &= \min\{t \in \mathbb{R} \mid t\tilde{e} \in v + \tilde{K}\} \\ &= \min\{t \in \mathbb{R} \mid \forall j \in [\bar{\omega}] : te \in v_j + K\} \\ &= \max_{j \in [\bar{\omega}]} \psi_e(v_j). \end{aligned} \tag{34}$$

From [4, Theorem 4], we know that $\psi_{\tilde{e}}$ is an Hiriart-Urruty functional. Hence, for a fixed $a \in P_{\bar{x}}$ we can apply [6, Proposition 2.2] to (\mathcal{VP}_a) to obtain that

$$\bar{x} \text{ is a stationary point of } (\mathcal{VP}_a) \iff \min_{u \in \mathbb{R}^n} \psi_{\tilde{e}}\left(\nabla \tilde{f}^a(\bar{x})^\top u\right) + \frac{1}{2}\|u\|^2 = 0. \tag{35}$$

Thus, we deduce that

$$\begin{aligned} \bar{x} \text{ is strongly stationary} &\stackrel{\text{(Remark 3.1)}}{\iff} \forall a \in P_{\bar{x}} : \bar{x} \text{ is stationary for } (\mathcal{VP}_a) \\ &\stackrel{\text{(35)+(28)+(34)}}{\iff} \forall a \in P_{\bar{x}} : \min_{u \in \mathbb{R}^n} \varphi_{\bar{x}}(a, u) = 0 \\ &\iff \min_{(a,u) \in P_{\bar{x}} \times \mathbb{R}^n} \varphi_{\bar{x}}(a, u) = 0 \\ &\stackrel{\text{(31)}}{\iff} \phi(\bar{x}) = 0 \\ &\stackrel{\text{(33)}}{\iff} \bar{u} = 0, \end{aligned}$$

as desired. □

Remark 4.3 A similar statement to the one in Proposition 4.2 can be made for stationary points of (SP_{ℓ}) . Indeed, for a set $Q \subseteq P_{\bar{x}}$, consider a point $(\bar{a}_Q, \bar{u}_Q) \in Q \times \mathbb{R}^n$ such that $\varphi_{\bar{x}}(\bar{a}_Q, \bar{u}_Q) = \min_{(a,u) \in Q \times \mathbb{R}^n} \varphi_{\bar{x}}(a, u)$. Then, by replacing $P_{\bar{x}}$ by Q in the proof of Proposition 4.2, we can show that the following statements are equivalent:

- (i) \bar{x} is stationary for (\mathcal{SP}_ℓ) with respect to Q ,
- (ii) $\min_{(a,u) \in Q \times \mathbb{R}^n} \varphi_{\bar{x}}(a, u) = 0$,
- (iii) $\bar{u}_Q = 0$.

Next, we show that the line search in Step 4 of Algorithm 1 terminates in finitely many steps.

Proposition 4.3 Fix $\beta \in (0, 1)$ and consider the functionals $\varphi_{\bar{x}}$ and ϕ given in (28) and (31), respectively. Furthermore, let $(\bar{a}, \bar{u}) \in P_{\bar{x}} \times \mathbb{R}^n$ be such that $\phi(\bar{x}) = \varphi_{\bar{x}}(\bar{a}, \bar{u})$ and suppose that \bar{x} is not a strongly stationary point of (\mathcal{SP}_ℓ) . The following assertions hold:

- (i) There exists $\tilde{t} > 0$ such that

$$\forall t \in (0, \tilde{t}], j \in [\bar{\omega}] : f^{\bar{a}_j}(\bar{x} + t\bar{u}) \preceq f^{\bar{a}_j}(\bar{x}) + \beta t \nabla f^{\bar{a}_j}(\bar{x})^\top \bar{u}.$$

- (ii) Let \tilde{t} be the parameter in statement (i). Then,

$$\forall t \in (0, \tilde{t}] : F(\bar{x} + t\bar{u}) \preceq^\ell \left\{ f^{\bar{a}_j}(\bar{x}) + \beta t \nabla f^{\bar{a}_j}(\bar{x})^\top \bar{u} \right\}_{j \in [\bar{\omega}]} \prec^\ell F(\bar{x}).$$

In particular, \bar{u} is a descent direction of F at \bar{x} with respect to the preorder \preceq^ℓ .

Proof (i) Assume that (i) does not hold. Then, we could find a sequence $\{t_k\}_{k \geq 1}$ and $\bar{j} \in [\bar{\omega}]$ such that $t_k \rightarrow 0$ and

$$\forall k \in \mathbb{N} : f^{\bar{a}_{\bar{j}}}(\bar{x} + t_k \bar{u}) - f^{\bar{a}_{\bar{j}}}(\bar{x}) - \beta t_k \nabla f^{\bar{a}_{\bar{j}}}(\bar{x})^\top \bar{u} \notin -K. \tag{36}$$

As $(\mathbb{R}^m \setminus -K) \cup \{0\}$ is a cone, we can multiply (36) by $\frac{1}{t_k}$ for each $k \in \mathbb{N}$ to obtain

$$\forall k \in \mathbb{N} : \frac{f^{\bar{a}_{\bar{j}}}(\bar{x} + t_k \bar{u}) - f^{\bar{a}_{\bar{j}}}(\bar{x})}{t_k} - \beta \nabla f^{\bar{a}_{\bar{j}}}(\bar{x})^\top \bar{u} \notin -K. \tag{37}$$

Taking now the limit in (37), when $k \rightarrow +\infty$ we get

$$(1 - \beta) \nabla f^{\bar{a}_{\bar{j}}}(\bar{x})^\top \bar{u} \notin -\text{int } K.$$

Since $\beta \in (0, 1)$, this is equivalent to

$$\nabla f^{\bar{a}_{\bar{j}}}(\bar{x})^\top \bar{u} \notin -\text{int } K. \tag{38}$$

On the other hand, since \bar{x} is not strongly stationary, we can apply Proposition 4.2 to obtain that $\bar{u} \neq 0$ and that $\phi(\bar{x}) < 0$. This implies that $\varphi_{\bar{x}}(\bar{a}, \bar{u}) < 0$, and hence,

$$\max_{j \in [\bar{\omega}]} \left\{ \psi_e \left(\nabla f^{\bar{a}_j}(\bar{x})^\top \bar{u} \right) \right\} < -\frac{1}{2} \|\bar{u}\|^2 < 0.$$

From this, we deduce that

$$\forall j \in [\bar{\omega}] : \psi_e \left(\nabla f^{\bar{a}_j}(\bar{x})^\top \bar{u} \right) < 0$$

and, by Proposition 2.2 (iii),

$$\forall j \in [\bar{\omega}] : \nabla f^{\bar{a}_j}(\bar{x})^\top \bar{u} \in -\text{int } K. \tag{39}$$

However, this is a contradiction to (38), and hence, the statement is proven.

(ii) From (39), we know that

$$\forall j \in [\bar{\omega}], t \in (0, \tilde{t}] : f^{\bar{a}_j}(\bar{x}) + \beta t \nabla f^{\bar{a}_j}(\bar{x})^\top \bar{u} < f^{\bar{a}_j}(\bar{x}). \tag{40}$$

Then, it follows that

$$\begin{aligned} \forall t \in (0, \bar{t}] : F(\bar{x}) &\stackrel{\text{(Proposition 2.1)}}{\subseteq} \left\{ f^{\bar{a}_1}(\bar{x}), \dots, f^{\bar{a}_{\bar{\omega}}}(\bar{x}) \right\} + K \\ &\stackrel{(40)}{\subseteq} \left\{ \nabla f^{\bar{a}_j}(\bar{x}) + \beta t \nabla f^{\bar{a}_j}(\bar{x})^\top \bar{u} \right\}_{j \in [\bar{\omega}]} + \text{int } K \\ &\stackrel{\text{(Statement (i))}}{\subseteq} \left\{ f^{\bar{a}_j}(\bar{x} + t\bar{a}_1), \dots, f^{\bar{a}_j}(\bar{x} + t\bar{a}_{\bar{\omega}}) \right\}_{j \in [\bar{\omega}]} + \text{int } K \\ &\subseteq F(\bar{x} + t\bar{u}) + \text{int } K, \end{aligned}$$

as desired. □

We are now ready to establish the convergence of Algorithm 1.

Theorem 4.1 *Suppose that Algorithm 1 generates an infinite sequence for which \bar{x} is an accumulation point. Furthermore, assume that \bar{x} is regular for F . Then, \bar{x} is a stationary point of (\mathcal{SP}_ℓ) . If in addition $|P_{\bar{x}}| = 1$, then \bar{x} is a strongly stationary point of (\mathcal{SP}_ℓ) .*

Proof Consider the functional $\zeta : \mathcal{P}(\mathbb{R}^m) \rightarrow \mathbb{R} \cup \{-\infty\}$ defined as

$$\forall A \in \mathcal{P}(\mathbb{R}^m) : \zeta(A) := \inf_{y \in A} \psi_e(y).$$

The proof will be divided in several steps:

Step 1: We show the following result:

$$\forall k \in \mathbb{N} \cup \{0\} : (\zeta \circ F)(x_{k+1}) \leq (\zeta \circ F)(x_k) + \beta t_k \left[\phi(x_k) - \frac{1}{2} \|u_k\|^2 \right]. \tag{41}$$

Indeed, because of the monotonicity property of ψ_e in Proposition 2.2 (ii), the functional ζ is monotone with respect to the preorder \preceq^ℓ , that is,

$$\forall A, B \in \mathcal{P}(\mathbb{R}^m) : A \preceq^\ell B \implies \zeta(A) \leq \zeta(B).$$

On the other hand, from Proposition 4.3 (ii), we deduce that

$$\forall k \in \mathbb{N} \cup \{0\} : F(x_k + t_k u_k) \preceq^\ell \left\{ f^{a_{k,j}}(x_k) + \beta t_k \nabla f^{a_{k,j}}(x_k)^\top u_k \right\}_{i \in [\omega_k]}.$$

Hence, using the monotonicity of ζ and the sublinearity of ψ_e from Proposition 2.2 (i), we obtain for any $k \in \mathbb{N} \cup \{0\}$:

$$\begin{aligned} (\zeta \circ F)(x_{k+1}) &\leq \min_{j \in [\omega_k]} \left\{ \psi_e \left(f^{a_{k,j}}(x_k) + \beta t_k \nabla f^{a_{k,j}}(x_k)^\top u_k \right) \right\} \\ &\leq \min_{j \in [\omega_k]} \left\{ \psi_e \left(f^{a_{k,j}}(x_k) \right) + \beta t_k \psi_e \left(\nabla f^{a_{k,j}}(x_k)^\top u_k \right) \right\} \\ &\leq \min_{j \in [\omega_k]} \left\{ \psi_e \left(f^{a_{k,j}}(x_k) \right) + \beta t_k \max_{j' \in [\omega_k]} \left\{ \psi_e \left(\nabla f^{a_{k,j'}}(x_k)^\top u_k \right) \right\} \right\} \\ &= (\zeta \circ F)(x_k) + \beta t_k \max_{j \in [\omega_k]} \left\{ \psi_e \left(\nabla f^{a_{k,j}}(x_k)^\top u_k \right) \right\}. \end{aligned}$$

The above inequality, together with the definition of ϕ in (31), implies (41).

On the other hand, since \bar{x} is an accumulation point of the sequence $\{x_k\}_{k \geq 0}$, we can find a subsequence \mathcal{K} in \mathbb{N} such that $x_k \xrightarrow{\mathcal{K}} \bar{x}$.

Step 2: The following inequality holds

$$\forall k \in \mathbb{N} \cup \{0\} : F(\bar{x}) \preceq^\ell F(x_k). \tag{42}$$

Indeed, from Proposition 4.3 (ii), we can guarantee that the sequence $\{F(x_k)\}_{k \geq 0}$ is decreasing with respect to the preorder \preceq^ℓ , that is,

$$\forall k \in \mathbb{N} \cup \{0\} : F(x_{k+1}) \preceq^\ell F(x_k). \tag{43}$$

Fix now $k \in \mathbb{N}$, and $i \in [p]$. Then, according to (43), we have

$$\forall k' \in \mathcal{K}, k' \geq k, \exists i_{k'} \in [p] : f^{i_{k'}}(x_{k'}) \preceq f^i(x_k). \tag{44}$$

Since there are only a finite number of possible values for $i_{k'}$, we assume without loss of generality that there is $\bar{i} \in [p]$ such that $i_{k'} = \bar{i}$ for every $k' \in \mathcal{K}, k' \geq k$. Hence, (44) is equivalent to

$$\forall k' \in \mathcal{K}, k' \geq k : f^{\bar{i}}(x_{k'}) - f^i(x_k) \in -K. \tag{45}$$

Taking the limit now in (45) when $k' \xrightarrow{\mathcal{K}} +\infty$, we find that

$$f^i(x_k) \in f^{\bar{i}}(\bar{x}) + K.$$

Since i was chosen arbitrarily in $[p]$, this implies the statement.

Step 3: We prove that the sequence $\{u_k\}_{k \in \mathcal{K}}$ is bounded.

In order to see this, note that, since x_k is not a stationary point, we have by Proposition 4.2 that $\phi(x_k) < 0$ for every $k \in \mathbb{N} \cup \{0\}$. By the definition of a_k and u_k , we then have

$$\forall k \in \mathbb{N} \cup \{0\} : \varphi_{x_k}(a_k, u_k) < 0. \tag{46}$$

Let ρ be the Lipschitz constant of ψ_e from Proposition 2.2 (i). Then, we deduce that

$$\begin{aligned} \forall k \in \mathbb{N} \cup \{0\} : \|u_k\|^2 &\stackrel{((46)+(28))}{<} -2 \max_{j \in [\omega_k]} \left\{ \psi_e \left(\nabla f^{a_k,j}(\bar{x})^\top u_k \right) \right\} \\ &= 2 \max_{j \in [\omega_k]} \left\{ \left| \psi_e \left(\nabla f^{a_k,j}(\bar{x})^\top u_k \right) \right| \right\} \\ &\stackrel{(\text{Proposition } 2.2(i))}{\leq} 2\rho \max_{j \in [\omega_k]} \left\{ \left\| \nabla f^{a_k,j}(\bar{x})^\top u_k \right\| \right\} \\ &\leq 2\rho \|u_k\| \max_{j \in [\omega_k]} \left\{ \left\| \nabla f^{a_k,j}(x_k) \right\| \right\}. \end{aligned}$$

Hence,

$$\forall k \in \mathbb{N} \cup \{0\} : \|u_k\| \leq 2\rho \max_{j \in [\omega_k]} \left\{ \left\| \nabla f^{a_k,j}(x_k) \right\| \right\}. \tag{47}$$

Since $\{x_k\}_{k \in \mathbb{K}}$ is bounded, the statement follows from (47).

Step 4: We show that \bar{x} is stationary.

Fix $\kappa \in \mathbb{N}$. Then, it follows from (41) that

$$\forall k \in \mathbb{N} : -\beta t_k \max_{j \in [\omega_k]} \left\{ \psi_e \left(\nabla f^{a_k,j}(x_k)^\top u_k \right) \right\} \leq (\zeta \circ F)(x_k) - (\zeta \circ F)(x_{k+1}). \tag{48}$$

Adding this inequality for $k = 0, \dots, \kappa$, we obtain

$$-\beta \sum_{k=0}^{\kappa} t_k \max_{j \in [\omega_k]} \left\{ \psi_e \left(\nabla f^{a_k,j}(x_k)^\top u_k \right) \right\} \leq (\zeta \circ F)(x_0) - (\zeta \circ F)(x_{\kappa+1}). \tag{49}$$

On the other hand, similarly to (39) in the proof of Proposition 4.3 (i), we obtain that

$$\forall k \in \mathbb{N} \cup \{0\}, j \in [\omega_k] : \nabla f^{a_k,j}(x_k)^\top u_k \in -\text{int } K. \tag{50}$$

In particular, applying Proposition 2.2 (iii) in (50), we find that

$$\forall k \in \mathbb{N} \cup \{0\}, j \in [\omega_k] : \psi_e \left(\nabla f^{a_k,j}(x_k)^\top u_k \right) < 0. \tag{51}$$

We then have

$$\begin{aligned}
 0 &\stackrel{(51)}{<} - \sum_{k=0}^{\kappa} t_k \max_{j \in [\omega_k]} \left\{ \psi_e \left(\nabla f^{a_k, j}(x_k)^\top u_k \right) \right\} \stackrel{(49)}{\leq} \frac{(\zeta \circ F)(x_0) - (\zeta \circ F)(x_{\kappa+1})}{\beta} \\
 &\stackrel{(42)}{\leq} \frac{(\zeta \circ F)(x_0) - (\zeta \circ F)(\bar{x})}{\beta}.
 \end{aligned}$$

Taking now the limit in the previous inequality when $\kappa \rightarrow +\infty$, we deduce that

$$0 \leq - \sum_{k=0}^{\infty} t_k \max_{j \in [\omega_k]} \left\{ \psi_e \left(\nabla f^{a_k, j}(x_k)^\top u_k \right) \right\} < +\infty.$$

In particular, this implies

$$\lim_{k \rightarrow \infty} t_k \max_{j \in [\omega_k]} \left\{ \psi_e \left(\nabla f^{a_k, j}(x_k)^\top u_k \right) \right\} = 0. \tag{52}$$

Since there are only a finite number of subsets of $[p]$ and \bar{x} is regular for F , we can apply Lemma 4.2 to obtain, without loss of generality, the existence of $Q \subseteq P_{\bar{x}}$ and $\bar{a} \in Q$ such that

$$\forall k \in \mathcal{K} : \omega_k = \bar{\omega}, P_{x_k} = Q, a_k = \bar{a}. \tag{53}$$

Furthermore, since the sequences $\{t_k\}_{k \geq 1}, \{u_k\}_{k \in \mathcal{K}}$ are bounded, we can also assume without loss of generality the existence of $\bar{t} \in \mathbb{R}, \bar{u} \in \mathbb{R}^n$ such that

$$t_k \xrightarrow{\mathcal{K}} \bar{t}, u_k \xrightarrow{\mathcal{K}} \bar{u}. \tag{54}$$

The rest of the proof is devoted to show that \bar{x} is a stationary point with respect to Q . First, observe that by (53) and the definition of a_k , we have

$$\forall a \in Q, k \in \mathcal{K}, u \in \mathbb{R}^n : \phi(x_k) = \phi_{x_k}(\bar{a}, u_k) \leq \phi_{x_k}(a, u).$$

Then, taking into account that $\omega_k = \bar{\omega}$ in (53), we can take the limit when $k \xrightarrow{\mathcal{K}} +\infty$ in the above expression to obtain

$$\forall a \in Q, u \in \mathbb{R}^n : \phi_{\bar{x}}(\bar{a}, \bar{u}) \leq \phi_{\bar{x}}(a, u).$$

Equivalently, we have

$$(\bar{a}, \bar{u}) \in \underset{(a, u) \in Q \times \mathbb{R}^n}{\operatorname{argmin}} \phi_{\bar{x}}(a, u). \tag{55}$$

Next, we analyze two cases:

Case 1: $\bar{t} > 0$.

According to (52) and (53), we have in this case

$$\lim_{k \xrightarrow{\mathcal{K}} +\infty} \max_{j \in [\bar{\omega}]} \left\{ \psi_e \left(\nabla f^{\bar{a}_j}(x_k)^\top u_k \right) \right\} = 0. \tag{56}$$

Then, it follows that

$$\begin{aligned} 0 &\leq \frac{1}{2} \|\bar{u}\|^2 \\ &\stackrel{\text{(53)} + \text{(54)} + \text{(56)}}{=} \lim_{k \xrightarrow{\mathcal{K}} +\infty} \max_{j \in [\bar{\omega}]} \left\{ \psi_e \left(\nabla f^{\bar{a}_j}(x_k)^\top u_k \right) \right\} + \frac{1}{2} \|u_k\|^2 \\ &= \lim_{k \xrightarrow{\mathcal{K}} +\infty} \phi(x_k) \\ &\stackrel{\text{(32)}}{\leq} 0, \end{aligned}$$

from which we deduce $\bar{u} = 0$. This, together with (55) and Remark 4.3, implies that \bar{x} is a stationary point with respect to Q .

Case 2: $\bar{t} = 0$.

Fix an arbitrary $\kappa \in \mathbb{N}$. Since $t_k \xrightarrow{\mathcal{K}} 0$, for $k \in \mathcal{K}$ large enough v^κ does not satisfy Armijo’s line search criteria in Step 4 of Algorithm 1. By (53) and the finiteness of $\bar{\omega}$, we can assume without loss of generality the existence of $\bar{j} \in [\bar{\omega}]$ such that

$$\forall k \in \mathcal{K} : f^{\bar{a}_{\bar{j}}}(x_k + v^\kappa u_k) \not\leq f^{\bar{a}_{\bar{j}}}(x_k) + \beta v^\kappa \nabla f^{\bar{a}_{\bar{j}}}(x_k)^\top u_k.$$

From this, it follows that

$$\forall k \in \mathcal{K} : \frac{f^{\bar{a}_{\bar{j}}}(x_k + v^\kappa u_k) - f^{\bar{a}_{\bar{j}}}(x_k)}{v^\kappa} - \beta \nabla f^{\bar{a}_{\bar{j}}}(x_k)^\top u_k \notin -K.$$

Now, taking the limit when $k \xrightarrow{\mathcal{K}} +\infty$, we obtain

$$\frac{f^{\bar{a}_{\bar{j}}}(\bar{x} + v^\kappa \bar{u}) - f^{\bar{a}_{\bar{j}}}(\bar{x})}{v^\kappa} - \beta \nabla f^{\bar{a}_{\bar{j}}}(\bar{x})^\top \bar{u} \notin -\text{int } K.$$

Next, taking the limit when $\kappa \rightarrow +\infty$, we get

$$(1 - \beta) \nabla f^{\bar{a}_{\bar{j}}}(\bar{x})^\top \bar{u} \notin -\text{int } K.$$

Since $\beta \in (0, 1)$, we deduce that $\nabla f^{\bar{a}_{\bar{j}}}(\bar{x})^\top \bar{u} \notin -\text{int } K$ and, according to Proposition 2.2 (iii), this is equivalent to

$$\psi_e(\nabla f^{\bar{a}_{\bar{j}}}(\bar{x})^\top \bar{u}) \geq 0. \tag{57}$$

Finally, we find that

$$0 \stackrel{(57)}{\leq} \psi_e(\nabla f^{\bar{a}j}(\bar{x})^\top \bar{u}) \leq \varphi_{\bar{x}}(\bar{a}, \bar{u}) \stackrel{(55)}{=} \min_{(a,u) \in Q \times \mathbb{R}^n} \varphi_{\bar{x}}(a, u) \stackrel{(29)}{\leq} 0,$$

which implies

$$\min_{(a,u) \in Q \times \mathbb{R}^n} \varphi_{\bar{x}}(a, u) = 0. \quad (58)$$

The stationarity of \bar{x} follows then from (58) and Remark 4.3. The proof is complete. \square

5 Implementation and Numerical Illustrations

In this section, we report some preliminary numerical experience with the proposed method. Algorithm 1 was implemented in Python 3 and the experiments were done in a PC with an Intel(R) Core(TM) i5-4200U CPU processor and 4.0 GB of RAM. We describe in the following some details of the implementation and the experiments:

- We considered instances of problem (\mathcal{SP}_ℓ) only for the case in which K is the standard ordering cone, that is, $K = \mathbb{R}_+^m$. In addition, we choose the parameter $e \in \text{int } K$ for the scalarizing functional ψ_e as $e = (1, \dots, 1)^\top$.
- The parameters β and ν for the line search in Step 4 of the method were chosen as $\beta = 0.0001$, $\nu = 0.500$.
- The stopping criteria employed were that $\|u_k\| < 0.0001$, or a maximum number of 200 iterations were reached.
- For finding the set $\text{Min}(F(x_k), K)$ at the k^{th} -iteration in Step 1 of the algorithm, we implemented the method developed by Günther and Popovici in [16]. This procedure requires a strongly monotone functional $\psi : \mathbb{R}^m \rightarrow \mathbb{R}$ with respect to the partial order \leq for a so-called presorting phase. In our implementation, we used ψ defined as follows:

$$\forall v \in \mathbb{R}^m : \psi(v) := \sum_{i=1}^m v_i.$$

The other possibility for finding the set $\text{Min}(F(x_k), K)$ would be to use the method introduced by Jahn in [21,22,26] with ideas from [45]. However, as mentioned in “Introduction,” the first approach has better computational complexity. Thus, the algorithm proposed in [16] was a clear choice.

- At the k^{th} iteration in Step 2 of the algorithm, we worked with the modeling language CVXPY 1.0 [1,8] for the solution of the problem

$$\min_{(a,u) \in P_k \times \mathbb{R}^n} \varphi_{x_k}(a, u).$$

Since the variable a is constrained to be in the discrete set P_k , we proceeded as follows: Using the solver ECOS [9] within CVXPY, we compute for every $a \in P_k$ the unique solution u_a of the strongly convex problem

$$\min_{u \in \mathbb{R}^n} \varphi_{x_k}(a, u).$$

Then, we set

$$(a_k, u_k) := \operatorname{argmin}_{a \in P_k} \varphi_{x_k}(a, u_a).$$

- For each test instance considered in the experimental part, we generated initial points randomly on a specific set and run the algorithm. We define as solved those experiments in which the algorithm stopped because $\|u_k\| < 0.0001$, and declared that a strongly stationary point was found. For a given experiment, its final error is the value of $\|u_k\|$ at the last iteration. The following variables are collected for each test instance:
 - **Solved:** this value indicates the number of initial points for which the problem was solved.
 - **Iterations:** this is a 3-tuple (min, mean, max) that indicates the minimum, the mean and the maximum of the number of iterations in those instances reported as solved.
 - **Mean CPU Time:** Mean of the CPU time(in seconds) among the solved cases.

Furthermore, for clarity, all the numerical values will be displayed for up to four decimal places.

Now, we proceed to the different instances on which our algorithm was tested. Our first test instance can be seen as a continuous version of an example in [17].

Test Instance 5.1 We consider $F : \mathbb{R} \rightrightarrows \mathbb{R}^2$ defined as

$$F(x) := \left\{ f^1(x), \dots, f^5(x) \right\},$$

where, for $i \in [5]$, $f^i : \mathbb{R} \rightarrow \mathbb{R}^2$ is given as

$$f^i(x) := \left(\frac{x}{2} \sin(x) \right) + \sin^2(x) \left[\frac{(i-1)}{4} \begin{pmatrix} 1 \\ -1 \end{pmatrix} + \left(1 - \frac{(i-1)}{4} \right) \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right].$$

The objective values in this case are discretized segments moving around a curve and being contracted (dilated) by a factor dependent on the argument. We generated 100 initial points x_0 randomly on the interval $[-5\pi, 5\pi]$ and run our algorithm. Some of the metrics are listed in Table 1. As we can see, in this case all the runs terminated finding a strongly stationary point. Moreover, we observed that for this problem not too many iterations were needed.

Table 1 Performance of Algorithm 1 in Test Instance 5.1

Test instance 5.1		
Solved	Iterations	Mean CPU time
100	(0, 13.97, 71)	0.1872

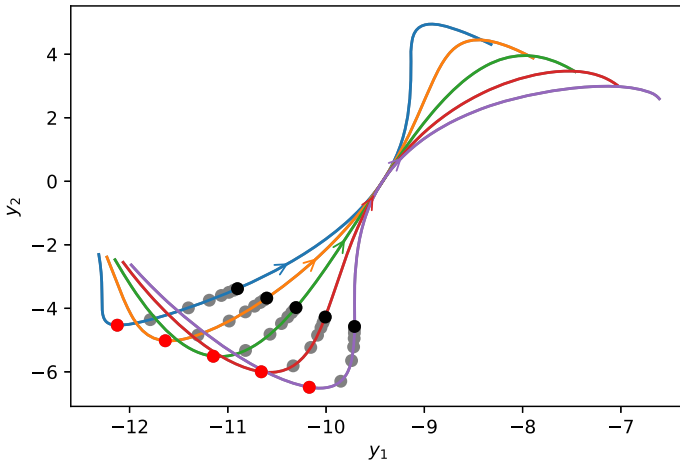


Fig. 1 Sequence generated in the image space by Algorithm 1 for a selected starting point in Test Instance 5.1

In Fig. 1, the sequence $\{F(x_k)\}_{k \in \{0\} \cup [7]}$ generated by Algorithm 1 for a selected starting point is shown. In this case, strong stationarity was declared after seven iterations. The traces of the curves f^i for $i \in [5]$ are displayed, with arrows indicating their direction of movement. Moreover, the sets $F(x_0)$ and $F(x_7)$ are represented by black and red points, respectively, and the elements of the sets $F(x_k)$ with $k \in [6]$ are in gray color. The improvements of the objective values after every iteration are clearly observed.

Test Instance 5.2 In this example, we start by taking a uniform partition \mathcal{U}_1 of 10 points of the interval $[-1, 1]$, that is,

$$\mathcal{U}_1 = \{-1, -0.7778, -0.5556, -0.3333, -0.1111, 0.1111, 0.3333, 0.5556, 0.7778, 1\}.$$

Then, the set $\mathcal{U} := \mathcal{U}_1 \times \mathcal{U}_1$ is a mesh of 100 points of the square $[-1, 1] \times [-1, 1]$. Let $\{u_1, \dots, u_{100}\}$ be an enumeration of \mathcal{U} and consider the points

$$l_1 := \begin{pmatrix} 0 \\ 0 \end{pmatrix}, l_2 := \begin{pmatrix} 8 \\ 0 \end{pmatrix}, l_3 := \begin{pmatrix} 0 \\ 8 \end{pmatrix}.$$

Table 2 Performance of Algorithm 1 in Test Instance 5.2

Test instance 5.2		
Solved	Iterations	Mean CPU time
100	(0, 1.32, 2)	0.0637

We define, for $i \in [100]$, the functional $f^i : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ as

$$f^i(x) := \frac{1}{2} \begin{pmatrix} \|x - l_1 - u_i\|^2 \\ \|x - l_2 - u_i\|^2 \\ \|x - l_3 - u_i\|^2 \end{pmatrix}.$$

Finally, the set-valued mapping $F : \mathbb{R}^2 \rightrightarrows \mathbb{R}^3$ is defined by

$$F(x) := \{f^1(x), \dots, f^{100}(x)\}.$$

Note that problem (SP_ℓ) corresponds in this case to the robust counterpart of a vector-valued facility location problem under uncertainty [20], where \mathcal{U} represents the uncertainty set with respect to the facility sites l_1, l_2, l_3 . Furthermore, with the aid of Theorem 3.1, it is possible to show that a point \bar{x} is a local weakly minimal solution of (SP_ℓ) if and only if

$$\bar{x} \in \text{conv} \{l_j + u_i \mid (i, j) \in I(\bar{x}) \times \{1, 2, 3\}\}.$$

Thus, in particular, the local weakly minimal solutions lie on the set

$$C := \text{conv} ((l_1 + \mathcal{U}) \cup (l_2 + \mathcal{U}) \cup (l_3 + \mathcal{U})). \tag{59}$$

In this test instance, 100 initial points x_0 were generated in the square $[-50, 50] \times [-50, 50]$, and Algorithm 1 was ran in each case. A summary of the results is presented in Table 2. Again, for any initial point, the sequence generated by the algorithm stopped with a local solution to our problem. Perhaps, the most noticeable parameter recorded in this case is the number of iterations required to declare the solution. Indeed, in most cases, only 1 iteration was enough, even when the starting point was far away from the locations l_1, l_2, l_3 .

In Fig. 2, the set of solutions found in this experiment are shown in red. The locations l_1, l_2, l_3 are represented by black points and the elements of the set $(l_1 + \mathcal{U}) \cup (l_2 + \mathcal{U}) \cup (l_3 + \mathcal{U})$ are colored in gray. We can observe, as expected, that all the local solutions found are contained in the set C given in (59).

Our last test example comes from [24].

Test Instance 5.3 For $i \in [100]$, we consider the functional $f^i : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ defined as

$$f^i(x) := \begin{pmatrix} e^{\frac{x_1}{2}} \cos(x_2) + x_1 \cos(x_2) \cos^3\left(\frac{2\pi(i-1)}{100}\right) - x_2 \sin(x_2) \sin^3\left(\frac{2\pi(i-1)}{100}\right) \\ e^{\frac{x_2}{20}} \sin(x_1) + x_1 \sin(x_2) \cos^3\left(\frac{2\pi(i-1)}{100}\right) + x_2 \cos(x_2) \sin^3\left(\frac{2\pi(i-1)}{100}\right) \end{pmatrix}.$$

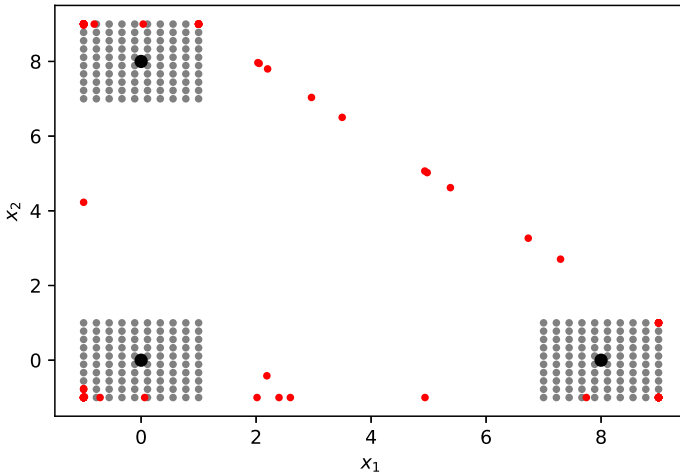


Fig. 2 Solutions found (in red) in the argument space for Test Instance 5.2

Table 3 Performance of Algorithm 1 in Test Instance 5.3

Test instance 5.3		
Solved	Iterations	Mean CPU time
88	(0, 11.9091 , 110)	0.8492

Hence, $F : \mathbb{R}^2 \rightrightarrows \mathbb{R}^2$ is given by

$$F(x) := \{f^1(x), \dots, f^{100}(x)\}.$$

The images of the set-valued mapping in this example are discretized, shifted, rotated and deformed rhombuses, as shown in Fig. 3. We generated randomly 100 initial points in the square $[-10\pi, 10\pi] \times [-10\pi, 10\pi]$ and ran our algorithm. A summary of the results is given in Table 3. In this case, only for 88 initial points a solution was found. In the rest of the occasions, the algorithm stopped because the maximum number of iterations was reached. Further examination in these unsolved cases revealed that, except for two of the initial points, the final error was of the order of 10^{-1} (even 10^{-3} and 10^{-4} in half of the cases). Thus, perhaps only a few more iterations were needed in order to declare strong stationarity.

Figure 3 illustrates the sequence $\{F(x_k)\}_{k \in \{0\} \cup [18]}$ generated by Algorithm 1 for a selected starting point. Strong stationarity was declared after 18 iterations in this experiment. The sets $F(x_0)$ and $F(x_{18})$ are represented by black and red points, respectively, and the elements of the sets $F(x_k)$ with $k \in [17]$ are in gray color. Similarly to the other test instances, we can observe that at every iteration the images decrease with respect to the preorder \leq^ℓ .

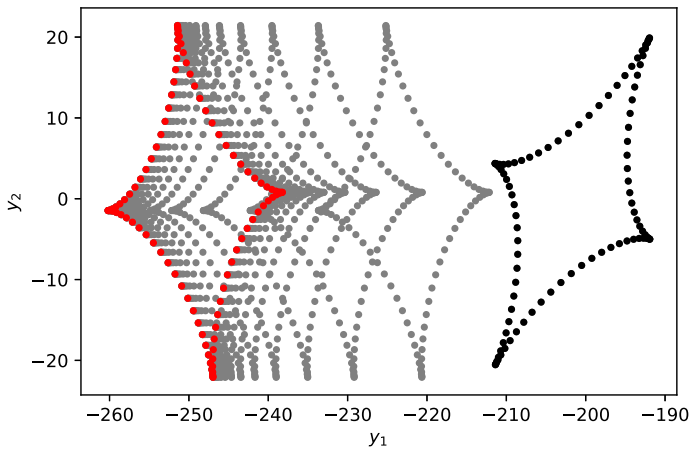


Fig. 3 Sequence generated in the image space by Algorithm 1 for a selected starting point in Test Instance 5.3

6 Conclusions

In this paper, we considered set optimization problems with respect to the lower less set relation, where the set-valued objective mapping can be decomposed into a finite number of continuously differentiable selections. The main contributions are the tailored optimality conditions derived using the first-order information of the selections in the decomposition, together with an algorithm for the solution of the problems with this structure. An attractive feature of our method is that we are able to guarantee convergence toward points satisfying the previously mentioned optimality conditions. To the best of our knowledge, this would be the first procedure having such property in the context of set optimization. Finally, because of the applications of problems with this structure in the context of optimization under uncertainty, ideas for further research include the development of cutting plane strategies for general set optimization problems, as well as the extension of our results to other set relations.

Funding Open Access funding enabled and organized by Projekt DEAL.

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. Agrawal, A., Verschueren, R., Diamond, S., Boyd, S.: A rewriting system for convex optimization problems. *J. Control Decis.* **5**(1), 42–60 (2018)

2. Amahroq, T., Oussarhan, A.: Lagrange multiplier rules for weakly minimal solutions of compact-valued set optimization problems. *Asia-Pac. J. Oper. Res.* **36**(4), 22 (2019)
3. Ben-Tal, A., Nemirovski, A.: Robust convex optimization. *Math. Oper. Res.* **23**(4), 769–805 (1998)
4. Bouza, G., Quintana, E., Tammer, C.: A unified characterization of nonlinear scalarizing functionals in optimization. *Vietnam J. Math.* **47**(3), 683–713 (2019)
5. Bouza, G., Quintana, E., Tuan, V.A., Tammer, C.: The Fermat rule for set optimization problems with Lipschitzian set-valued mappings. *J. Nonlinear Conv. Anal.* **21**(5), 1137–1174 (2020)
6. Chuong, T.D., Yao, J.-C.: Steepest descent methods for critical points in vector optimization problems. *Appl. Anal.* **91**(10), 1811–1829 (2012)
7. Conn, A.R., Scheinberg, K., Vicente, L.N.: Introduction to derivative-free optimization. In: MPS/SIAM (2009)
8. Diamond, S., Boyd, S.: CVXPY: a python-embedded modeling language for convex optimization. *J. Mach. Learn. Res.* **17**(83), 1–5 (2016)
9. Domahidi, A., Chu, E., Boyd, S.: ECOS: an SOCP solver for embedded systems. In: European Control Conference (ECC), pp. 3071–3076 (2013)
10. Eberhard, A., Roshchina, V., Sang, T.: Outer limits of subdifferentials for min-max type functions. *Optimization* **68**(7), 1391–1409 (2019)
11. Ehrgott, M., Ide, J., Schöbel, A.: Minmax robustness for multi-objective optimization problems. *Eur. J. Oper. Res.* **239**(1), 17–31 (2014)
12. Eichfelder, G., Niebling, J., Rocktäschel, S.: An algorithmic approach to multiobjective optimization with decision uncertainty. *J. Glob. Optim.* **77**, 3–25 (2019)
13. Geiger, C., Kanzow, C.: Numerische Verfahren zur Lösung Unrestringierter Optimierungsaufgaben. Springer, Berlin (1999)
14. Göpfert, A., Riahi, H., Tammer, C., Zălinescu, C.: Variational Methods in Partially Ordered Spaces. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, vol. 17. Springer, New York (2003)
15. Graña Drummond, L.M., Svaiter, B.F.: A steepest descent method for vector optimization. *J. Comput. Appl. Math.* **175**(2), 395–414 (2005)
16. Günther, C., Popovici, N.: New algorithms for discrete vector optimization based on the Graef–Younes method and cone-monotone sorting functions. *Optimization* **67**(7), 975–1003 (2018)
17. Günther, C., Köbis, E., Popovici, N.: Computing minimal elements of finite families of sets w.r.t. preorder relations in set optimization. *J. Appl. Numer. Optim.* **1**(2), 131–144 (2019)
18. Günther, C., Köbis, E., Popovici, N.: On strictly minimal elements w.r.t. preorder relations in set-valued optimization. *Appl. Set-Valued Anal. Optim.* **1**(3), 205–219 (2019)
19. Ide, J., Köbis, E.: Concepts of efficiency for uncertain multi-objective optimization problems based on set order relations. *Math. Methods Oper. Res.* **80**(1), 99–127 (2014)
20. Ide, J., Köbis, E., Kuroiwa, D., Schöbel, A., Tammer, C.: The relationship between multi-objective robustness concepts and set-valued optimization. *Fixed Point Theory Appl.* **2014**, 83 (2014). <https://doi.org/10.1186/1687-1812-2014-83>
21. Jahn, J.: Multiobjective search algorithm with subdivision technique. *Comput. Optim. Appl.* **35**(2), 161–175 (2006)
22. Jahn, J.: *Vector Optimization: Theory, Applications, and Extensions*, 2nd edn. Springer, Berlin (2011)
23. Jahn, J.: A derivative-free descent method in set optimization. *Comput. Optim. Appl.* **60**(2), 393–411 (2015)
24. Jahn, J.: A derivative-free rooted tree method in nonconvex set optimization. *Pure Appl. Funct. Anal.* **3**(4), 603–623 (2018)
25. Jahn, J., Ha, T.X.D.: New order relations in set optimization. *J. Optim. Theory Appl.* **148**(2), 209–236 (2011)
26. Jahn, J., Rathje, U.: Graef–Younes method with backward iteration. In: Küfer, K.H., Rommelfanger, H.C., Tammer, C., Winkler, K. (eds.) *Multicriteria Decision Making and Fuzzy Systems—Theory, Methods and Applications*, pp. 75–81. Shaker Verlag, Aachen (2006)
27. Jiang, L., Cao, J., Xiong, L.: Generalized multiobjective robustness and relations to set-valued optimization. *Appl. Math. Comput.* **361**, 599–608 (2019)
28. Karaman, E., Soyterem, M., Atasever, G.I., Tozkan, D., Küçük, M., Küçük, Y.: Partial order relations on family of sets and scalarizations for set optimization. *Positivity* **22**(3), 783–802 (2018)
29. Khan, A.A., Tammer, C., Zălinescu, C.: *Set-Valued Optimization: An Introduction with Applications*. Springer, Heidelberg (2015)

30. Köbis, E., Köbis, M.A.: Treatment of set order relations by means of a nonlinear scalarization functional: a full characterization. *Optimization* **65**(10), 1805–1827 (2016)
31. Köbis, E., Kuroiwa, D., Tammer, C.: Generalized set order relations and their numerical treatment. *Appl. Anal. Optim.* **1**(1), 45–65 (2017)
32. Köbis, E., Le Thanh, T.: Numerical procedures for obtaining strong, strict and ideal minimal solutions of set optimization problems. *Appl. Anal. Optim.* **2**(3), 423–440 (2018)
33. Kung, H.T., Luccio, F., Preparata, F.P.: On finding the maxima of a set of vectors. *J. Assoc. Comput. Mach.* **22**(4), 469–476 (1975)
34. Kuroiwa, D.: Some criteria in set-valued optimization. Number 985, pp. 171–176 (1997). *Investigations on nonlinear analysis and convex analysis (Japanese)* (Kyoto, 1996)
35. Kuroiwa, D.: The natural criteria in set-valued optimization. *Sūrikaiseikikenkyūsho Kōkyūroku*, (1031):85–90 (1998). *Research on nonlinear analysis and convex analysis (Japanese)* (Kyoto, 1997)
36. Kuroiwa, D.: On set-valued optimization. In: *Proceedings of the Third World Congress of Nonlinear Analysts, Part 2* (Catania, 2000), vol. 47, pp. 1395–1400 (2001)
37. Kuroiwa, D., Tanaka, T., Ha, T.X.D.: On cone convexity of set-valued maps. *Nonlinear Anal.* **30**(3), 1487–1496 (1997)
38. Luc, D.T.: *Theory of vector optimization*. In: *Lecture Notes in Economics and Mathematical Systems*, vol. 319. Springer, Berlin (1989)
39. Mordukhovich, B.S.: Variational analysis and generalized differentiation I: basic theory. In: Mordukhovich, B.S. (ed.) *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*, vol. 330. Springer, Berlin (2006)
40. Mutapcic, A., Boyd, S.: Cutting-set methods for robust convex optimization with pessimizing oracles. *Optim. Methods Softw.* **24**(3), 381–406 (2009)
41. Nishnianidze, Z.G.: Fixed points of monotone multivalued operators. *Soobshch. Akad. Nauk Gruzin. SSR* **114**(3), 489–491 (1984)
42. Quintana, E.: *On set optimization with set relations: a scalarization approach to optimality conditions and algorithms*. Dissertation, Martin-Luther-Universität Halle-Wittenberg (2020)
43. Schirotzek, W.: *Nonsmooth Analysis*. Universitext. Springer, Berlin (2007)
44. Schmidt, M., Schöbel, A., Thom, L.: Min-ordering and max-ordering scalarization methods for multi-objective robust optimization. *Eur. J. Oper. Res.* **275**(2), 446–459 (2019)
45. Younes, Y.M.: *Studies on discrete vector optimization*. Dissertation, University of Demiatta (1993)
46. Young, R.C.: The algebra of many-valued quantities. *Math. Ann.* **104**(1), 260–290 (1931)

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